"Construction and Testing of Large-Area GEM Detectors for the Forward Muon Endcap Upgrade of the CMS Experiment and Vector-Portal Search for Dark Matter Particles with Dimuon Pairs at $\sqrt{s} = 13$ TeV."

by

Mehdi Rahmani

Master of Science
Physics
Florida Institute of Technology
2018

Bachelor of Science
Physics
Shahid Behesti University
2015

A dissertation
submitted to the College of Engineering and Science
at Florida Institute of Technology
in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
in
Physics

Melbourne, Florida
May, 2022

We the undersigned committee
hereby approve the attached dissertation

"Construction and Testing of Large-Area GEM Detectors for the Forward Muon Endcap Upgrade of the CMS Experiment and Vector-Portal Search for Dark Matter Particles with Dimuon Pairs at $\sqrt{s} = 13$ TeV." by Mehdi Rahmani

---

Marcus Hohlmann, Ph.D.
Professor
Aerospace, Physics and Space Sciences
Major Advisor

Debasis Mitra, Ph.D.
Professor
Computer Engineering and Sciences

Hamid K. Rassoul, Ph.D.
Distinguished Professor
Dept. of Aerospace, Physics and Space Sciences

Ming Zhang, Ph.D.
Professor
Aerospace, Physics and Space Sciences

David C. Fleming, Ph.D.
Associate Professor and Department Head
Aerospace, Physics and Space Sciences

ABSTRACT

Title:

"Construction and Testing of Large-Area GEM Detectors for the Forward Muon Endcap Upgrade of the CMS Experiment and Vector-Portal Search for Dark Matter Particles with Dimuon Pairs at $\sqrt{s} = 13$ TeV."

Author:

Mehdi Rahmani

Major Advisor:

Marcus Hohlmann, Ph.D.

In this thesis, I present my contributions to the second major upgrade of the Compact Muon Solenoid at the Large Hadron Collider and the analysis that I performed on the data collected by the Compact Muon Solenoid during the 2018 and 2017 era. Alongside my team, I assembled and tested ten Gaseous Electron Multipliers to upgrade the forward region at the Compact Muon Solenoid, all of which are now installed at the dedicated stations. The campaign for producing these detectors succeeded in a two-year-long prototyping endeavor, where I examined the procedural steps for production to optimize the production and testing procedural efficiency. I dedicate the more extensive portion of this dissertation to the analysis of the dataset, corresponding to 59.7 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV recorded during 2018, wherein I interpret the model-independent results of the analysis in the context of a dark matter model, by setting 95% upper exclusion limits on the parameters of the model. The model involves a vector-portal i.e., a dark Z boson (Z_D), and its decay to dark scalars (s_D) which subsequently decay to a four muon final state ($pp \to Z_D \to s_D\overline{s_D} \to 4\mu$). While I conclude the 2018 analysis by declaring no significant deviation from the predicted background, to improve the background modeling and better statistics, I discuss the progress in analyzing the 2017 era dataset to be combined with the 2018 results.

Table of Contents

List of Figures

List of Tables

Acknowledgements

First, I must acknowledge the immense helpfulness and professionalism of Dr. Marcus Hohlmann, my mentor, teacher, and advisor. I remember the day I met him for the first time; as a *fresh-off-the-boat* international student, I was excited yet timid, unsure about what lay ahead of me. In that meeting, he instilled courage and confidence in me through expressing precisely what this scientific journey will look like and being honest about how challenging it will be and how confident he is in the successful fate of this journey. In my years of working with him, he provided vital guidance when faced with disappointment and frustration with a pressing scientific issue. And more importantly, when overwhelmed with the workload and mental illness, he offered support and understanding. There is a sentence from him that I will carry with me for the rest of my professional, scientific life: "All we have as scientists is our integrity." Hear, hear to that!

Next, my gratitude is to all my professors and colleagues during my academic education. My professors in Iran, such as dear Dr. Hamid Sepangi, who taught me how to inquire scientifically and showed me why pure sciences are *attractive*, above all. And all my professors at Florida Tech who provided excellent higher education. I have to thank Dr. Rassoul particularly, who mentored me through hard times. Special thanks to the post-doctorate scientists I have worked with, Stephano Colafranceschi and Aiwu Zhang. Thank you to Wei Shi of Texas A&M, who patiently taught me a

great deal about data analysis. Thanks to my colleague, Stephen Butalla, with whom I have collaborated on multiple fruitful projects and won awards.

I have much love and gratitude for my Family, my mother Mojgan, and my father Reza, who supported me with all their heart while enduring the great weight of *distance*. Infinite gratitude to my sister, Aylar, who was my best friend and confidant for all these years.

My thanks extend to all my friends, the ones in Iran and everyone here. Their support and friendship gave me the motivation and energy to complete this journey. Many thanks to Mohammad (Tahir), Chakavak (Chaka), her husband, and my dear friend Parhman (Pagham). Many thanks to Erfan, Kiarash (Kiar), Amir (Mohandes), and Aryan (Vezva), my friends and peers who came to the United States around the same time as me and shared this incredible experience with me. Much appreciation for all my friends here in the United States, whose companionship was a valuable asset, thanks to Levi, Angelo, Sarah, and Eric. Gratitude and appreciation to the small Iranian community here, my second family, in Melbourne, whom I turned to in my time of need; the beautiful people who cared for me when I was physically or mentally ill. Many thanks to Mohsen, Shayesteh, Mahsa, Samaneh, and Mahyar.

I have much gratitude to Dr. Bita Rahmati, my therapist, whose help was immense in maintaining my mental health through turbulent times.

Gratitude to Brendan Steffens, my peer and my friend with whom, shoulder by shoulder, I wrote this document. Brendan and I worked on our dissertations quite literally side by side. He kindly proofread my thesis, and I attempted to do the same for him. It is difficult to describe how helpful and assuring his presence was to the whole of this process.

And much appreciation and gratitude for all those who helped or provided support throughout my education years; without your help, this thesis would have been impossible.

Dedication

I dedicate this thesis to my parents. The man and the woman who selflessly sacrificed a great deal for my education and my success. The man and the woman who endured a great deal of anguish, inflicted by distance and the politics that restricted reunion for half a decade. The man and the woman who, in the face of uncertain times, courageously dared to hope for a better life for their son and offered all they had. I dedicate this thesis to my brilliant sister, my best friend, who completed this educational journey before me and lovingly provided moral support. And I dedicate this thesis to all the young scientists in Iran who, despite harsh geopolitical realities and financial difficulties surrounding them, devote their lives to the pursuit of knowledge and human progression.

Preface

Before you lies the document that presents my years of effort as a doctorate of philosophy student at the Florida Institute of Technology. The title of this dissertation is *"Construction and Testing of Large-Area GEM Detectors for the Forward Muon Endcap Upgrade of CMS Experiment and Vector-Portal Search for Dark Matter Particles with Dimuon Pairs at* $\sqrt{s} = 13$ *TeV"*. This is a long title; allow me to break it apart and give you an overview of each part. Effectively, my contribution to the vast landscape of high-energy physics is two-fold, separate but intertwined and codependent: first, the detector construction and testing *for* the upgrade of CMS, and second, analyzing the dataset collected during the year 2018 *by* CMS. In the below passages, I offer generalities about each portion of my contributions, provide the road map for this document, and demonstrate the crucial link between the two concepts.

The hardware sector is tied with the upgrade of the Compact Muon Detector (CMS) at the Large Hadron Collider (LHC). The LHC is the largest and the mightiest particle accelerator ever built, and its mission is to search for new physics. At the LHC, protons are accelerated and collide against one other. The resulting debris from these collisions is then studied in the search for new physics. CMS detector is one of the four major detectors built around the collision points, focusing on muon detection. The LHC is continuously enhanced to higher energies and higher collision rates. CMS detector goes through long shutdowns to meet the new challenges brought forth by

the enhancements of the LHC accelerator. Here is where my contribution comes in. I engaged in the prototyping process, construction, and testing of the Gas Electron Multiplier (GEM) detectors, designed to provide better muon detection rate and more accurate muon property measurements at the forward region of the LHC.

On the hardware side, in the first chapter, you will begin by getting introduced to the GEM detectors and operational principles, where I lay out the motivation for designing and installing GEMs at CMS and their designed function. Next, I describe the assembly procedure, the test steps, and the results I acquired for the detectors we produced at Florida Tech. Lastly, I conclude the chapter by presenting the fruit of my endeavor in the hardware sector to successfully assemble and test ten GEM detectors for CMS, all of which are currently installed and operational at CMS forward region.

The analysis venture in my research focuses on the notion of dark matter and the current highly active experimental search for a piece of concrete evidence for its non-gravitational interaction with the Standard Model (SM) matter. I carefully use the term *non-gravitational* here because there is ample observational evidence for the gravitational interaction of dark matter with celestial bodies. The hunt for dark matter has increasingly become the subject of research at the LHC. There are a few approaches to dark matter detection at the hadron colliders. My colleagues and I adopted the model-independent approach in which we consider the gross features of the data in search of physics beyond the standard model, independent of the theoretical frameworks. Once the model-independent results are solidified, I interpret the results within a simplified dark matter model, wherein the proton-proton collisions produce dark matter spin-0 particles (dark scalars) that subsequently decay to two muon pairs (dimuons). In chapter 3, I introduce this model as the *vector-portal* model.

You will notice that the analysis side of this dissertation is considerably more extensive and elaborate. Here is the road map for the content of this section: First, after a

comprehensive introduction to modern particle physics and the challenges that we face today in this realm —particularly, the lack of experimentally confirmed formulation for the existence of dark matter —I present the foundation for the benchmark model that I used for results interpretation. Next, I introduce CMS detector and its various parts and pieces and its inner working mechanisms. Once the introductory phase of the document comes to an end, I begin discussing the specific ingredients for the analysis, such as the samples that we used, including the simulated samples and the data collected by CMS in the year 2018. This is followed by establishing the set of event selection criteria designed to keep the events that interest us and disposing of the rest. The background modeling and estimation is a crucial subject that I address next. Here, I demonstrate what constitutes the expected SM background in the parameter space of interest; any excess above this estimated background could signal new physics. Lastly, uncertainty analysis, including theoretical and experimental, precedes the final results for the 2018 analysis in Chapter 8. In the last chapter, I argue for the improvement of the background modeling in a particular studied region, and I present our progress in adding the 2017 CMS data to the 2018 analysis.

I have been extremely fortunate to work alongside an incredible community around the fascinating machine that the LHC is. To build detectors that collect the data that a future young scientist uses to look for new physics has been a privilege to me; precisely how past scientists built the detectors that collected the data that I used for searching for new physics. That is the beauty of science, collaboration, and support for one thing, and one thing only: the progress of human knowledge about the fundamentals that shape their reality. As humble as it may be, I hope this dissertation contributes to this virtuous endeavor.

Chapter 1

Construction and Quality Control of GEM Detectors

At the moment of authoring this document, efforts are underway to transform the Large Hadron Collider (LHC) into the High Luminosity Large Hadron Collider (HL LHC), undergoing a series of upgrades in a multitude of subsystems. These upgrades will enable LHC to attain unprecedented luminosity of 2–3$\times 10^{34}$ cm^2 s^{-1} for Run 3 (starting in 2022), and it will be at least 5×10^{34} cm^2 s^{-1} when the High Luminosity Large Hadron Collider is operational during Run 4. The impressive higher luminosity at the LHC comes with certain caveats: current muon triggering, identification, and measurement will not be able to keep up, resulting in miss-identification and lower efficiencies in the muon system. The CMS collaboration has carefully designed and planned to install additional detector systems such as Gas Electron Multiplier (GEM) detector technology to meet this challenge. This chapter focuses on my work during the mass production of GEM detectors, including construction and quality control (QC). A total of 161 triple GEM detectors have been constructed by several construction

sites, including at Florida Institute of Technology, across the globe. Of 161 detectors produced, 156 detectors passed all tests, and 144 are now installed in CMS experiment. I engaged in the construction and QC of 10 GEM detectors installed at CMS GE1/1 station in the endcap. I participated in constructing and testing multiple prototypes that helped to define the final version.

1.1 Motivation

The muon systems at CMS take advantage of three muon subsystems: the Resistive Plate Chambers (RPCs) [12] in the barrel and endcaps, Cathode Strip Chambers (CSCs) [13] in the endcaps, and Drift Tubes (DTs) [14] in the barrel. The muon system is tasked with muon triggering, reconstruction, and identification. The upgrade of the LHC injector chain during the second long shut down (LS2) will warrant luminosities as high as 2×10^{34} cm^{-2}s^{-1}, which amounts to twice the design luminosity. The increase in luminosity implies increased pileup (PU) interactions –extra inelastic proton-proton collisions coinciding with the hard scattering collision (defined in Sec. 5.2.1) –to up to 100 PU and a higher radiation dose. Consequently, the overall performance of the current muon system in this condition is diminished. The muon measurements in the endcaps are especially affected as the multiple scattering in the return yoke, and the lower amplitude magnetic field (smaller bending angle) play a role. A higher multiple scattering rate alongside higher background rates in the forward region ($|\eta|$ >1.6) will result in higher Level 1 (L1) trigger rates that overwhelms the CSC system.

The CMS collaboration responded by introducing GEM technology to be installed at the designated GE1/1 endcap station along with other stations. This installation will help restore redundancy for tracking and triggering in the muon system. GEM detectors can operate in a high-rate environment to provide precise tracking, which

improves muon momentum resolution. The combination of GEM + CSC will accurately measure the muon track bending angle unaffected by multiple scattering as there is little material between the two detector systems, which can then be used in the L1 muon trigger to reduce the soft muon rate [1]. I give a comprehensive description of CMS detector, the muon system and its different stations (Sec. 4.6), and its inner-workings in Ch. 4. For the time being, you can find the GE1/1 location in the quadrant overview of CMS muon system in Fig. 4.4.

1.2 The GEM Technology at CMS

A triple-GEM chamber (detector) incorporates three *GEM foils*, i.e., Kapton foils coated with copper on both sides with an array of microscopic holes (typically 140µm pitch), separated by spacers and held between an anode readout board and a cathode drift board. Usage of GEM technology is perfectly appropriate for upgrading the muon system since it displays a rate capability well above the required ∼10 kHz/cm^2, a time resolution of ∼8 ns or better, and a spatial resolution of ∼200 μm. A GEM chamber utilizes electron amplification through microscopic holes in the GEM foils within a gas medium. The standard gas mixture for operating and testing this triple-GEM detector is Ar/CO_2 (70:30) [11].

A voltage of a few hundred volts is applied across the conductive copper layers of the GEM foil, creating a strong electric field (60-100 kV/cm) inside the holes. The passing charged particle, such as muon, ionizes the gas and produces electron-ion pairs. The produced primary electrons drift towards the GEM foil holes and gain more kinetic energy. Upon entering the hole, the electrons acquire enough kinetic energy to ignite secondary ionization in the gas and initiate what is known as an electron-ion *avalanche*. The produced avalanche amplifies through the remaining GEM layers and induces an

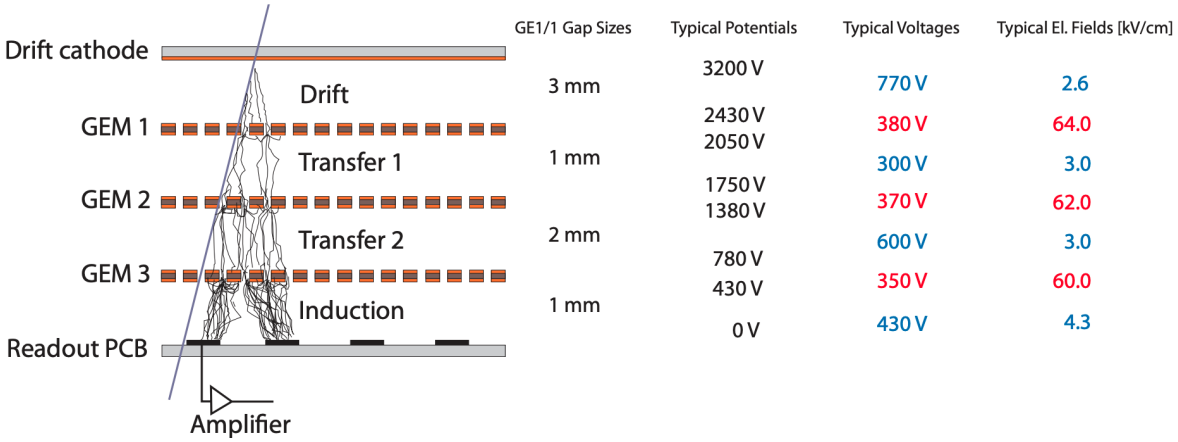


Figure 1.1: A schematic overview of a triple-GEM chamber (left) is illustrated. The typical voltage and electric field distribution across the three GEM foils are listed in the columns on the right [1].

electrical signal on the readout strips.

Instrumenting the detector with three GEM foils furthermore reduces the need for very high voltages across a single foil which could cause electrical breakdowns. The aggregate of three GEM foils delivers a total charge amplification of up to 10^5, also known as *gain* [15]. Structurally, the three GEM foils within a triple-GEM chamber are arranged as follows: 3 mm of drift region between drift cathode and first GEM, spaces of 1 mm and 2 mm in the electron transfer gaps between the GEM foils, and a 1 mm space in the signal induction region [1]. A schematic overview of a triple-GEM detector and the designated voltage distribution across the foils is shown in Fig 1.1.

The GE1/1 station instruments the pseudorapidity region $1.55 < |\eta| < 2.18$, as shown in Fig. 4.4. A pair of GEM detectors are joined to form a "super-chamber," which provides two position measurements for each muon track at GE1/1. To cover the entire ϕ range, 36 super-chambers, each covering 10°, are installed at each CMS endcap. Each endcap holds 18 long and 18 short super-chambers. As illustrated in Fig. 4.4, the super-chambers are installed in the gap between the hadron calorimeter and the CSC ME1/1 chambers.

The high-voltage (HV) is delivered to the GEM foils and the drift cathode from the drift board through a total of seven designated powering pins: two pins for each side of the three GEM foils, plus the drift cathode itself. The induced electrical signals on the readout strips are collected, amplified, digitized, and further processed by custom-designed 128-channel Application-specific integrated circuit (ASIC) circuits, called the *VFAT3 chips* [16]. The data are then communicated to off-GEM electronics via radiation tolerant optical fibers. The readout strips are segmented into three segments along the azimuthal ϕ-coordinate and eight segments along the pseudorapidity η-coordinate. The notation for referring to the resulting 24 readout segments is $(i\eta, i\phi)$, where $i\eta = 8$ points to the narrow end (closest to the beamline), and $i\eta = 1$ denotes the wide end of the detector.

The central building blocks of triple-GEM detectors are depicted in Fig. 1.2. A thorough discussion on the design of the GEM detectors is given in the GEM technical design report Ref. [1]. In the prototyping process, leading up to the final version, I assembled versions III, IV, and V of the GEM chambers and reported minor part deficiencies and proposed procedure optimizations. The reports can be found on CMS e-log service click this[1]. A photo of a version V (five) chamber assembled by Florida Tech is shown in Fig. 1.3.

1.3 Assembly

The preparation for assembly begins with the installation of gas input and output plugs onto the readout board (Fig. 1.5**(a,c)**) and insertion of the brass inserter onto the inner frames (Fig. 1.5**(d)**). I prepare the drift board (Fig. 1.5**(b)**) next by affixing the

[1]https://cmsonline.cern.ch/webcenter/portal/cmsonline?wc.contentSource=

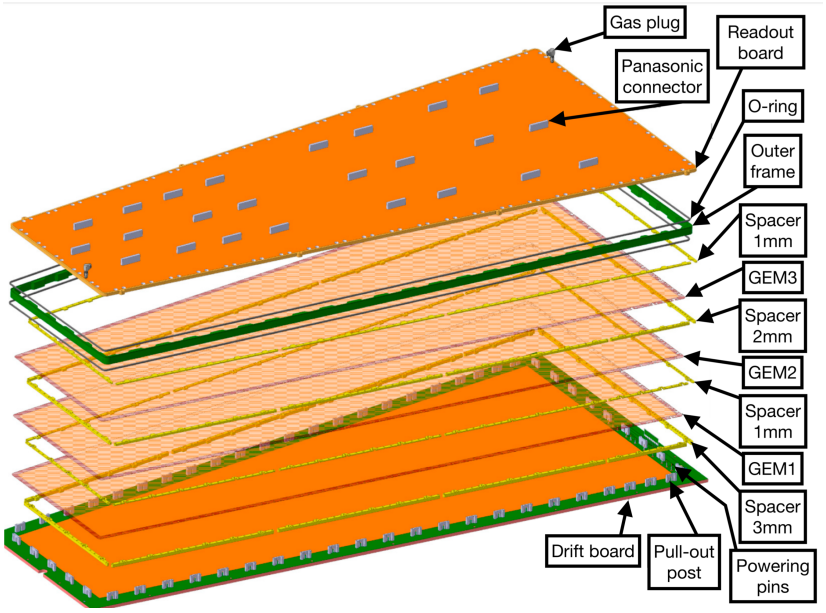


Figure 1.2: An exploded view of the triple-GEM detector is shown. The main components start from the bottom: drift board mounted all around with stainless steel pull-outs, used for stretching of the GEM foils, 3 mm frame (Spacer), first foil, 1 mm frame, second foil, 2 mm frame, third foil, 1 mm frame, first O-ring, external frame, second O-ring and the readout board. [2].

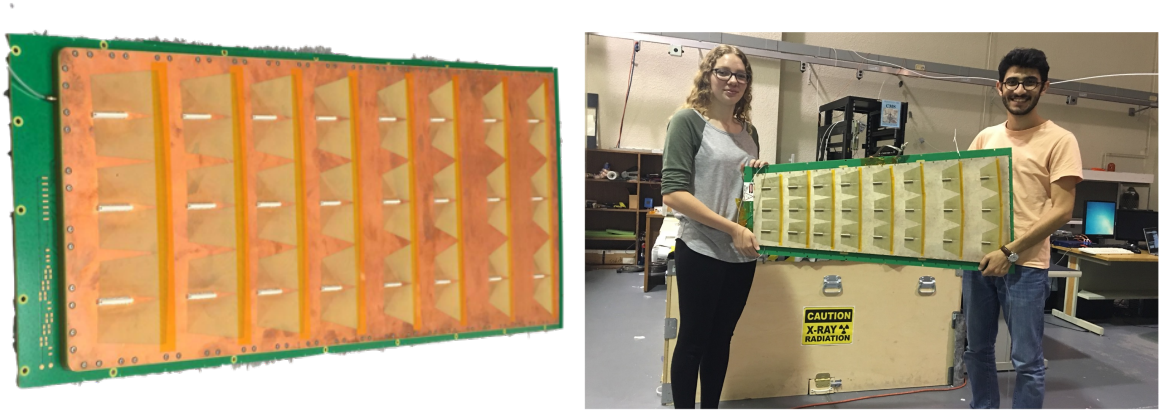


Figure 1.3: The figure on the left is the picture of version V of the GEM chamber prototypes, and the figure on the shows a fully assembled version III GEM chamber. In the image on the right, I am seen holding the right side of the detector, and undergraduate student, Sarah Arends, is holding the left side of the chamber.

stainless steel pull-outs (Fig. 1.5**(e)**) and soldering the HV connection pins (Fig. 1.5**(f)**), which power the drift board as well as the GEM foils.

The assembly and gas sealing process must be conducted in a clean room (rated at least a class 1000) and is done mechanically. A broad overview of Florida Tech's clean-room is shown in Fig. 1.4. This is a necessary practice, as airborne dust, particulate matter, or chemical contaminants, in a non-conditioned environment can land in the GEM foil holes and cause low resistance electrical shorts between the two electrodes of a GEM foil. For this reason, the foils, readout board, and drift board are regularly cleaned by an adhesive roller that picks up micro-sized particulates.

The complete assembly of the GEM chambers is conducted in the following order: the three foils are *stacked* (Fig. 1.6**(a)**) on top of each other, then placed between the unsegmented drift cathode and the segmented readout anode. I begin by assembling the *stack*, consisting of three GEM foils bound together by layers of the frame (inner frames) outside of the drift board. A plexiglass slate with designed holes for small metal pins supports the GEM foils and the frames as they are layered on top of one another. Once the stack is assembled, the excess Kapton is cut out from around the foil. The assembled stack is then carefully placed on the drift board. Every few centimeters on the internal frames, a square stainless steel nut is planted to provide a grip for stretching screws (Fig. 1.6**(b)**), mounted onto the pull-out posts. Next, I tension the foils using pull-out posts and stretching screws. The tension is provided by pull-out screws dragging the GEM foils in the stack against the posts. The screws are tightened manually, and the foil tension is standardized by the amount of torque applied with the "Torx screwdriver" to the pull-out screws (Fig. 1.6**(c)**). Once the foils are adequately tensioned, I ensure the tension is adequate and produces no electrical shorts caused by contact between the foils or GEM holes being distorted by faulty tension. For this purpose, I measure the foil resistance using a handheld Giga-Ohm insulation meter (Megger MIT485) (Fig. 1.6**(e)**). The meter applies an HV of 550 V across the foil to obtain the resistance and measures the corresponding leakage current. The outer

frame, equipped with a Viton o-ring, is placed around the fully stacked and tensioned foils to prevent gas leakage (Fig. 1.6**(d)**). Lastly, I close and seal the chamber by screwing the readout board onto the drift board as shown in Fig. 1.6**(f)**. We have documented the details of the assembly procedure in Ref. [2].

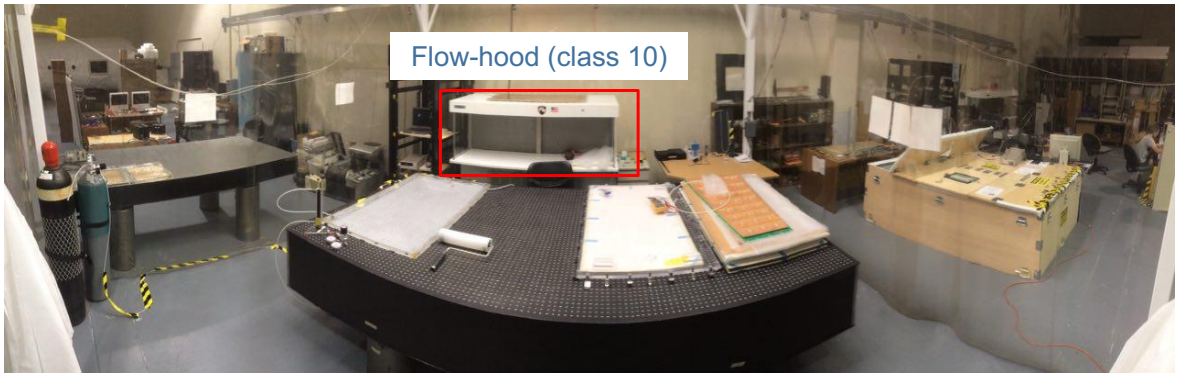
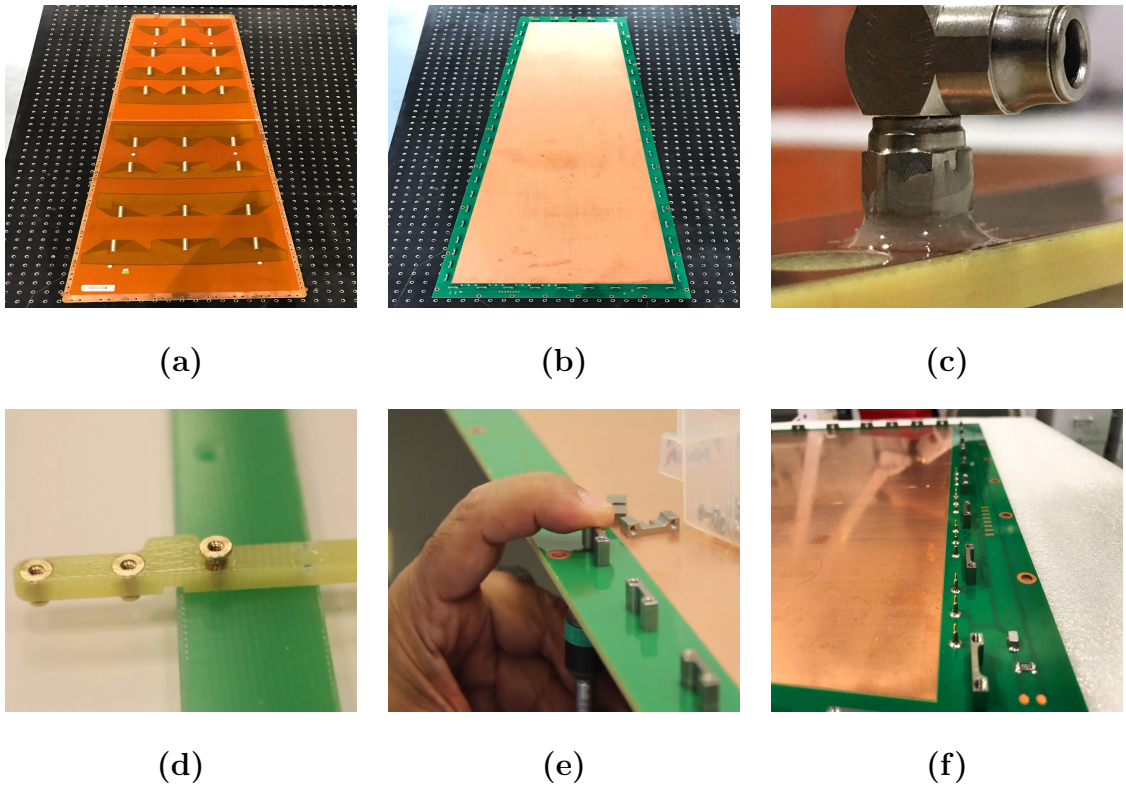


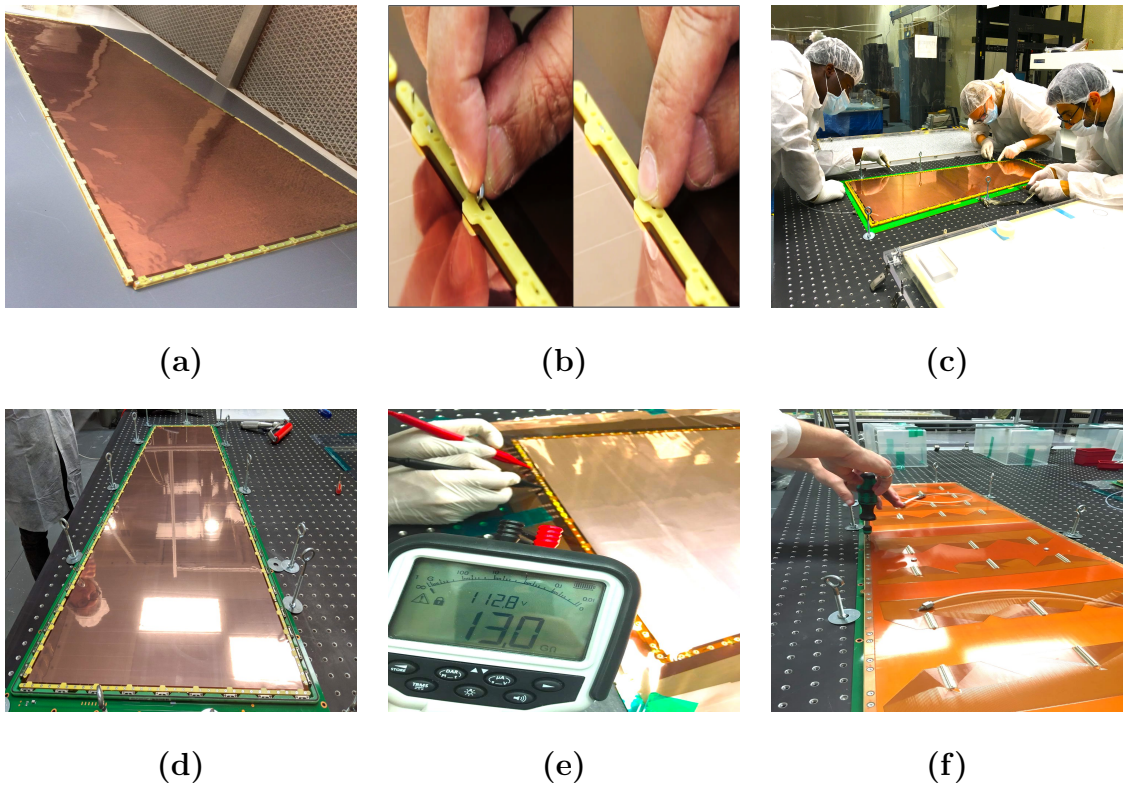
Figure 1.4: A broad panorama image of the clean room used for triple-GEM chamber assembly is shown. The stack assembly is performed on the optical table. The GEM stack is stored in the flow-hood while the drift board is cleaned.



(a) **(b)** **(c)**

(d) **(e)** **(f)**

Figure 1.5: **(a)** The readout board **(b)** The drift board **(c)** The gas plug is glued onto the readout board **(d)** The brass inserters are inserted into the inner-frames **(e)** The pull-outs are screwed onto the drift board **(f)** The HV pins



(a) **(b)** **(c)**

(d) **(e)** **(f)**

Figure 1.6: **(a)** An assembled GEM stack under the flow-hood **(b)** The square-nuts are inserted into the inner frame hinges **(c)** Students and I are stretching the GEM foils by using the stretching screws and Torx screwdrivers. **(d)** A stretched stack **(e)** Foils are regularly tested by applying high voltage to prevent electrical shorts **(f)** The readout board is mounted onto the drift board with the outer frame sandwiched between them.

1.4 Quality Control

Six countries from all over the globe participated in the assembly and QC campaign to build and test a total of 161 triple-GEM detectors. Of those detectors, 156 passed all tests, and 144 are now installed in CMS experiment. Florida Institute of Technology was able to participate in this campaign after following a two-year certification plan. The plan aimed to guarantee that the laboratories had adequate infrastructure and instrumentation for standardized production and concrete QCs. The GEM collaboration produced a series of documents outlining the assembly procedures and QC

instructions [2, 17].

The QC steps are designed to assess each aspect of the detector's performance and guarantee that the detectors will perform as expected in CMS environment. If a detector fails to satisfy the predefined standard expected results for a QC step, the chamber will not proceed to the next step unless the issue is resolved through further adjustments. The results for each QC step are regularly logged and systematically reviewed by the GEM community. Upon additional validation by the collaboration, the chambers that pass all the QC steps advance to shipment to CMS site. On-site, the final QC tests are performed on the chambers, super-chambers are constructed, final QCs are performed on the super-chamber, and installation at the endcaps ensues. I conducted the single-chamber QCs, as listed in Tab. 1.1. A broad picture of the Florida Tech site is shown in Fig. 1.7.

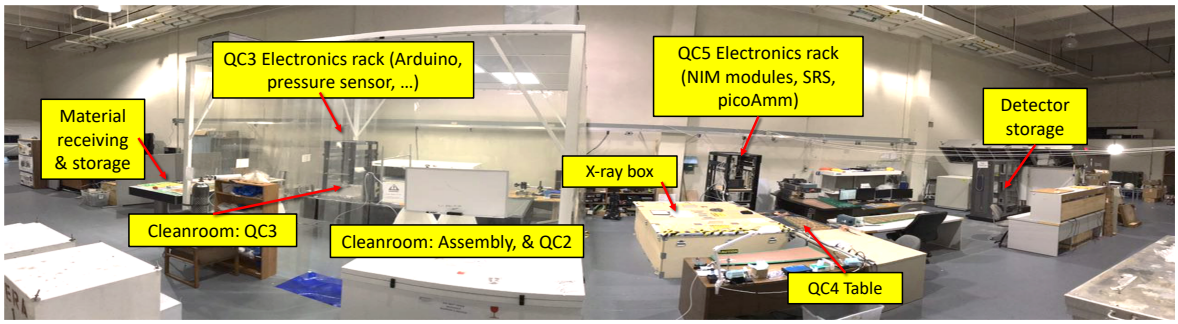


Figure 1.7: A panorama overview of the Florida Tech site for assembly and QC of the GEM chambers is shown. The workflow moves from left to right with product inspection, assembly, leakage current test, and gas leakage test in the cleanroom. Next, HV linearity and gain measurements are performed outside the cleanroom.

1.4.1 QC Step 2 – Resistance check

Motivation: The current leakage test is to validate that the holes on the GEM foils can tolerate sufficient electric field for electron multiplication without sparking any electrical shorts. As touched on in Sec. 1.3, there are three significant factors in creating an

Table 1.1: QC tests for a triple-GEM chamber to be installed at the GE1/1 station.

QC step	Description
1	Initial inspection of the chamber components
2	Electrical cleaning and current leakage test of the foils
3	Gas leakage test of the closed chamber
4	Linearity test of the HV divider + intrinsic noise test
5a	Effective gas gain measurement
5b	Gain uniformity test

electrical bridge (short) in a GEM foil: the presence of dust, chemical contaminants, and now I add the third cause, which is mechanical defects. To examine the foils for possible electrical shorts, I measure the total resistance between the two electrodes of the foil. Ideally, the resistance between the two electrodes should be very high as the Kapton foil is an insulator; nevertheless, there will be some surface current leakage along the walls of the hole and the natural moisture in the air. The environmental variables, such as humidity and temperature, can affect the measured current leakage. For this reason, I record the humidity and temperature together with the resistance of each foil.

Selection criteria: A voltage of 550 V is applied across the two electrodes (top and bottom) of each foil using a Megger for ten minutes. A foil is considered unhealthy if the resistance drops below 10 GΩ in a sub-50% humanity environment and the spark rate falls below 2 Hz during the last two to three minutes of the test. Depending on how many segments of the GEM foil suffers from electrical shorts, the measured resistance can be 10 MΩ —the total resistance of the protection resistors, soldered directly onto the segmented side of the GEM foil —or lower in the case of multiple shorts, as the protection resistors are summed in parallel. If the number of sparks exceeds this limit, or the resistance of the foils is under 10 GΩ, the foil needs to be cleaned again with the anti-static roller. Should the problem persist after several cleaning attempts,

the foil must be sent back to the production workshop and washed with deionized water.

Test setup: The foils are taped onto a plastic frame and leaned against a structure that I built for this purpose, using 80/20 aluminum bars. I subject the top and bottom of the foil to 550 V, provided by the Megger. We designed a PCB-based *HV clip* to deliver the HV from the Megger to the foils. I attach the HV clip to the designated electrode HV pads on the foil and keep it under voltage for as long as needed. The environmental indicators are recorded along with the resistance measurements. The photo of Fig. 1.8 shows my QC2 setup at Florida Tech.

Results: I present the QC2 test results performed over a GEM foil with the ID of *FOIL-ID-S-132/12* that we received at Florida Tech. The humidity in Florida is slightly higher (above 50%) than in Geneva. Here, I apply 550 V to each foil and measure the resistance. I use Ohm's law to calculate the current leakage. The total number of sparks is counted over ten minutes. If the resistance of all the three GEM foils is above 10 GΩ, the assembly can proceed. The results are presented in Tab. 1.2. The results summary for all Florida Tech-produced chambers is shown in Tab. 1.3.

1.4.2 QC Step 3 – Gas leakage test of the closed chamber

Motivation: The GEM chamber needs to be gas-tight as the intrusion of external pollutants could deteriorate the gain of the detector. For instance, the presence of the oxygen molecule in the chamber can absorb (steal) some of the primary or secondary electrons, resulting in a lower gain. Furthermore, the entrance of dust particles into the chamber could compromise the soundness of the GEM foils, leading to *sparking*, jeopardizing the HV stability across the system, and spoiling the overall performance

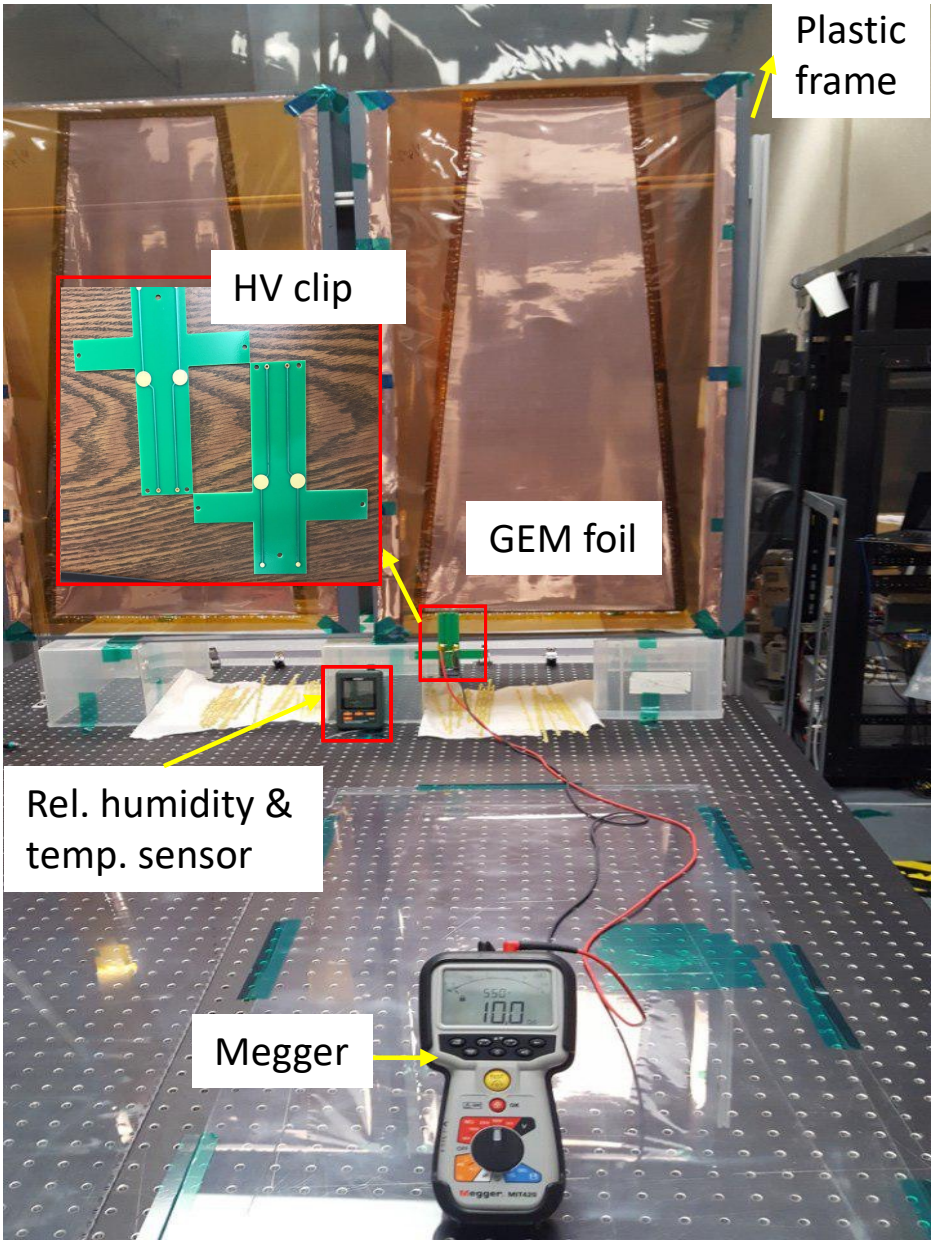


Figure 1.8: The setup at Florida Tech for the electrical resistance test of a GE1/1 GEM foil (QC2).

of the detector.

Selection criteria: The ultimate requirement to maintain an acceptable purity of the gas flow in the chamber is a leak rate of 1% of the total incoming gas flow rate.

Table 1.2: Typical QC2 results as measured by the author at Florida Tech. The environmental variables are: temperature (temp): 26.3 C°, atmospheric (atm) pressure: 1016.1 mb (millibar), Rel. Humidity: 53.6 %

FOIL-ID-S-132/12					
Time [min]	Voltage [V]	Resistance [G ohm]	Current [nA]	Sparks	Total sparks
0.5	550	4	550	1	1
1	550	8	250	1	2
2	550	9.8	152	1	3
3	550	3.8	134	0	3
4	550	8.9	112.2	1	4
5	550	12.5	95	0	4
6	550	13.4	90	0	4
7	550	14	71	0	4
8	550	15.5	65	0	4
9	550	16	61	0	4
10	550	17	61	0	4

Considering the volume of the GE1/1 chamber, this figure translates to a rate of 2.5 L/hr with a maximum internal over-pressure of 25 millibar. However, practically, directly measuring a leak rate of as little as 0.025 L/hr could present a challenge with conventional tooling available at our laboratory. Alternatively, I measure gas leakage of the closed chamber with both gas input and output blocked. Consequently, the time evolution of the internal pressure of the chamber can be expressed as an exponential,

$$P_{int} = P_0 e^{-t/\tau}, \tag{1.1}$$

where P_0 is the initial pressure, and τ is the time constant of the system. By using the ideal gas law, $PV = nRT$, with P being the absolute pressure of the system $(P_{int(t)} + P_{atm})$, we have,

Table 1.3: The Summary of QC step 2 results for all Florida Tech-produced chambers. For each chamber, the foil IDs, the resistance, and the environmental conditions of the test are listed.

Chmaber SN	Foil ID	Resistance (MΩ)	Temp (C°)	RH %	Pressure (mbar)
FIT0001	172/15 168/15 137/13	12 14 13	26.3	53.6	1016.1
FIT0002	167/15 170/15 171/15	17 11 14	27	51.2	1017.3
FIT0003	175/15 125/12 133/12	11 11 13	26.3	53	1016.1
FIT0004	172/17 186/16 132/12	14 13 10	27.2	52	1015.2
FIT0005	171/15 226/19 228/19	15 13 11	26.3	53.6	1016.7
FIT0006	207/18 215/18 218/18	10 13 17	26.3	53	1016.4
FIT0007	213/18 217/18 194/17	13 8 11	25.2	56.3	1013.4
FIT0008	214/18 191/17 199/17	11 13 11	26.9	55.2	1014.1
FIT0009	205/18 206/18 209/18	11 14 15	27.2	48.4	1015.1
FIT0010	166/15 208/18 210/18	11 11 13	26.1	43.8	1017.9

$$n = \frac{V}{RT}(P_{int(t)} + P_{atm}) \tag{1.2}$$

Substituting P_{int} from Eq. 1.1 in Eq. 1.2 and differentiating the resulting expression with respect to t, we arrive at a detector leak rate, the ultimate query for this QC step:

$$\frac{dn}{dt} = \frac{VP_0}{RT}\left(\frac{-1}{\tau}\right)e^{-t/\tau} \tag{1.3}$$

The expression above requires initial internal pressure and the time constant to

produce the leak rate. To extract the time constant, I plot the internal over-pressure against time. Fitting an exponential reveals the time constant. The condensed statement about the selection criteria is the following: Any chamber with a time constant greater than 3.04 hours passes the QC3 criteria to ensure that the leak rate remains below 1%. This criterion is derived experimentally by calculating the gas leak rate in GE1/1 chambers as a function of the internal over-pressure [11].

Test setup: The tooling for conducting this test is available at Florida Tech: a pressure regulator, atmospheric pressure indicator, ambient temperature indicator, Arduino, flow meter, two valves, and two flow meters. I connect the pressure sensor to the Arduino microcontroller, which enables the over-pressure monitoring of the chamber. Furthermore, I obtained the ambient temperature and pressure via sensors connected to the Arduino board. A diagram of the DAQ system is depicted in Fig. 1.9 (top). Before conducting the QC3 step on the chamber, I calibrate the system to make sure the system produces accurate measurements of the gas leak in the chamber. A diagram of the measurement setup is shown in Fig. 1.9 (bottom). At last, I pressurize the chambers to ∼25 mbar, close the gas input and output valves, and monitor the pressure loss for one hour.

Results: An exponential fit to the pressure drop against time will reveal the time constant. In the expression Eq. 1.1, the initial pressure P_0 is set to 25 mbar, and τ is the time constant that I extract from the fit. As an example of the typical results, the time constant in the exponential fit in Fig. 1.10 is 26.7 hr. This time constant is well above the selection criteria (3.04 hr); therefore, this chamber passes the QC3 test. In Fig. 1.11 I summarize the results of all ten detectors assembled at Florida Tech. for QC step 3.

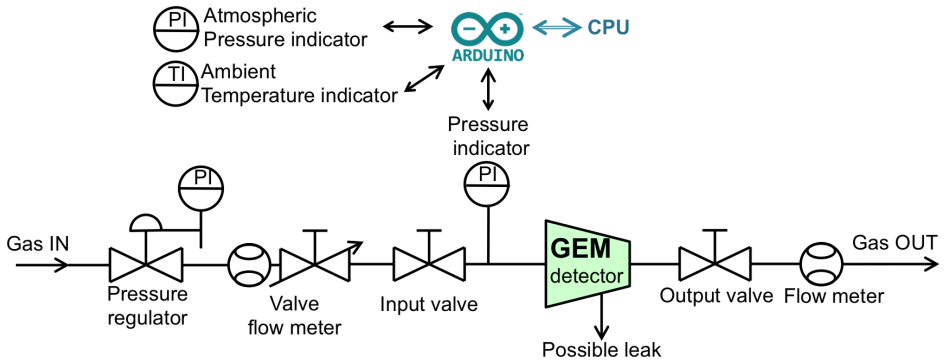


Figure 1.9: QC3 Gas leakage test setup taken from instructions given in CMS GEM internal note Ref. [2].

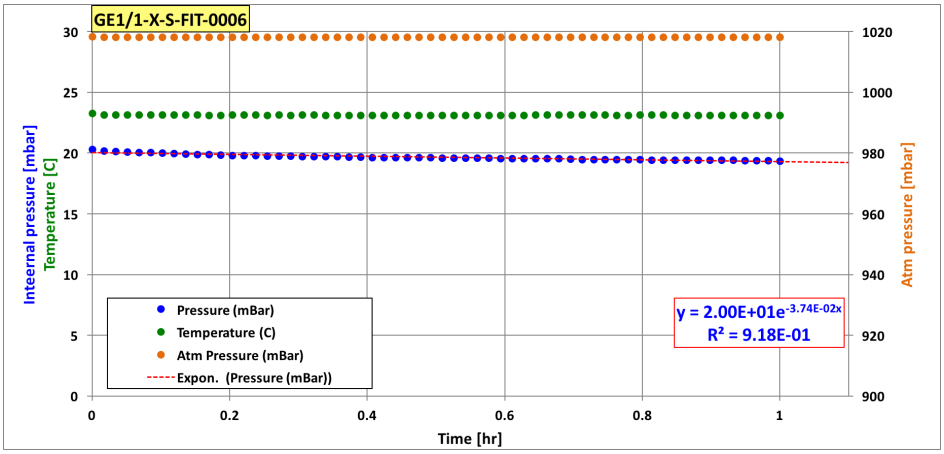


Figure 1.10: Typical gas leakage test results as measured by the author at Florida Tech.

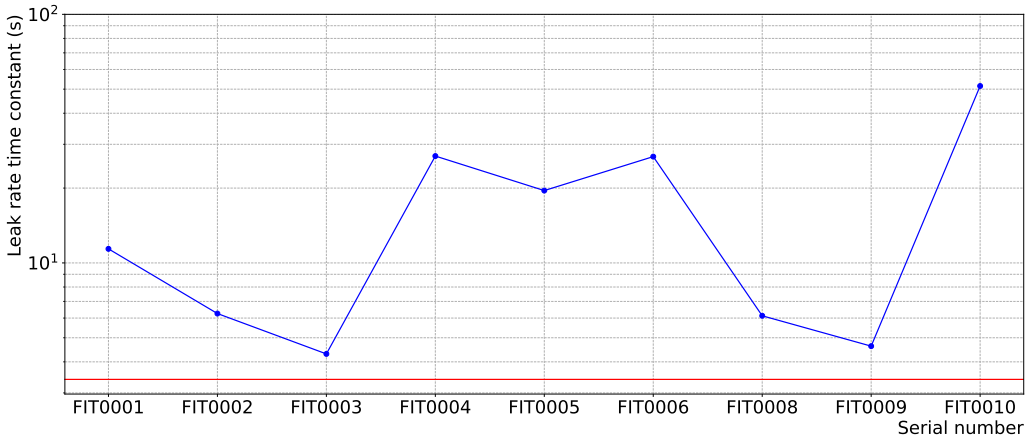


Figure 1.11: The summary of the results for QC step 3 for all Florida Tech-produced GEM chambers. The leak rate time constant for each Florida Tech chamber, identified by its serial number, is shown. The red line represents the selection criteria (3.04 hr). The FIT0007 chamber exhibited no pressure drop for the duration of the test (1 hr).

1.4.3 QC Step 4 - Linearity test of the HV divider

Motivation The HV has to be appropriately distributed among the GEM foils and the drift board to produce desired electric fields in each region of the GEM chamber. The QC step 4 is designed to test and validate the functionality of on-detector circuitry that distributes the HV. This stable and consistent response of the circuitry to HV is crucial for ensuring a steady, gain and avoiding discharges on the foils that could permanently damage the detector or the readout electronics.

As discussed in Sec. 1.3 and shown in Fig. 1.5(**f**), the GEM foils are powered through a set of HV pins soldered onto the drift board within the gas volume. The spring-loaded pins are pushed against their corresponding HV electrode pad on their designated GEM foil. PCB traces are routed from the HV pins to the outside of the gas volume and connected to the set of HV pads. The pads receive high voltage either through a multi-channel or single-channel power supply (CAEN N1470), delivered to the pins, and the pins power the GEM foils. The GE1/1 tripe-GEM foils are arranged in a 3/1/2/1 mm formation which dictates a specific HV distribution to reach desired

electric fields in the gaps and GEM foils. Whereas at CMS endcap, the voltage is supplied to the GEM detectors by a PC-controlled HV Multi-channel Power supply [1], a simple ceramic HV divider is used at the production sites. The HV divider is an 8-pin integrated resistor network with the first pin connected to the drift cathode, the second to seventh pins connected to the six electrodes of the 3 GEM foils, and the last pin connected to the ground [18]. I solder the HV divider onto the HV pads on the drift board and supply the HV through a single channel HV line, as shown in Fig. 1.12. The current I_d through each leg of the HV divider produces voltage drops $V_i = R_i \times I_d$ across the various resistors that are used to create the appropriate electric potentials needed to power the different electrodes of the detector. In Fig 1.13 the circuit diagram of the HV divider is depicted. To eliminate the high-frequency noise trafficked from the power supply to the chamber, the circuitry additionally includes a low-pass protection filter with equivalent resistance $R_{filter} = 0.3$ MΩ, bringing the total equivalent resistance of the circuit to 5.0 MΩ. The HV stability is quantified by monitoring two quantities: the deviation of the measured total resistance of the powering circuit with respect to the one predicted by Ohm's law and the noise-induced on the bottom of the third GEM foil in gas with high ionization energy (CO_2).

Selection criteria: I vary the supplied HV to the chamber in equal steps and record the corresponding current received through the powering circuit. By visualizing the result on an I-V plot, I should observe a linear behavior, attesting to the ohmic relation between the applied HV, and measure the current through the resistive HV divider. I derive the measured resistance of the powering circuit R_m by taking the average of the momentary resistance for each step. I am interested in the deviation of R_m from the nominal resistance R_n of the HV divider, measured by a simple multimeter.

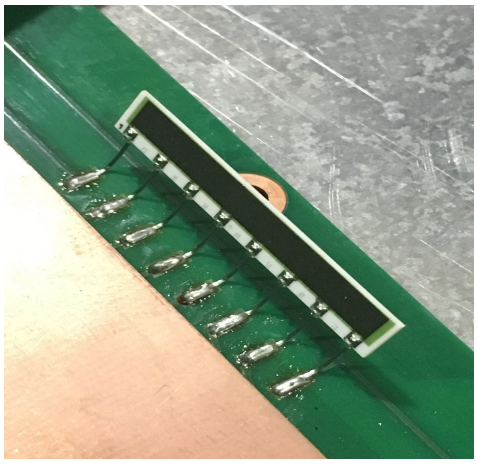


Figure 1.12: The ceramic HV divider is soldered onto the drift board.

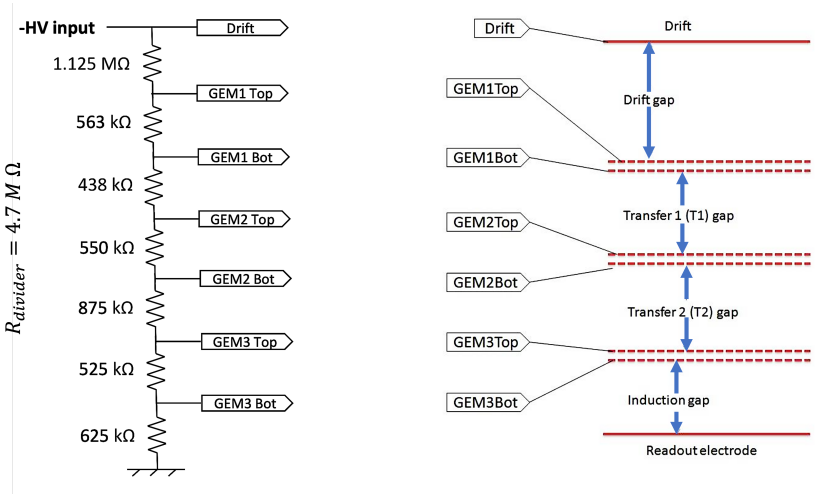


Figure 1.13: HV divider circuit diagram (left) for the 3/1/2/1 mm gap configuration and corresponding connections to GE1/1 chamber electrodes (right). For each of the chamber electrodes, there are two HV pins (one for redundancy). Note that additional 10 MΩ protection resistors are located on the segmented sides of all GEM foils [1].

$$D_R = \frac{R_m - R_n}{R_n} \tag{1.4}$$

The tolerance of the participating resistors in the total resistance of the powering circuit determines the allowable deviation (D_R). There is an aggregate of ten resistors in the HV divider and the HV filters, each rated with a 1% tolerance. Summing the

tolerances in quadrature reveals that a deviation of 3% can be tolerated. The selection criteria state that a chamber with a resistance deviation D_R% smaller than 3% can proceed to the next QC test.

Test setup: To conduct the QC4 test, I transfer the chamber from the cleanroom to the designated station for the test (Fig. 1.7). The chamber operated in safe mode, meaning that it is flushed with pure CO_2 with a flow rate of 5 l/hr to prevent sparking. The commercial CAEN power supply allows me to ramp up the HV in steps of 100 V up to 5 kV to monitor the current supplied to the HV divider.

Results: I fit the I-V curve with a linear function and compute the deviation of its slope with respect to the nominal value by using Eq. 1.4. In Fig. 1.14 the dotted curve is the I-V trend. I will address the blue curve in the upcoming section. The slope of the green curve is the measured (R_m). The deviation of this value from the nominal resistance (R_n) is below 3%. Consequently, the detector moves forward with the next QC step. In Fig. 1.15 I summarize the results of all ten detectors assembled at Florida Tech for QC step 4.

1.4.4 QC step 4 - Intrinsic noise rate measurement

Motivation: The triple-GEM chambers, operating at exceptionally high gains, are known to be prone to spurious discharges [19]. The *glow* discharges can occur without the ionizing source in the gas volume and therefore have to be kept at a minimum rate. This intrinsic noise should be well quantified and understood to reduce the misidentification probability. In this section, I show how the spurious signals are quantified. I conduct QC step 4 by flushing the chamber with CO_2, which has high enough high

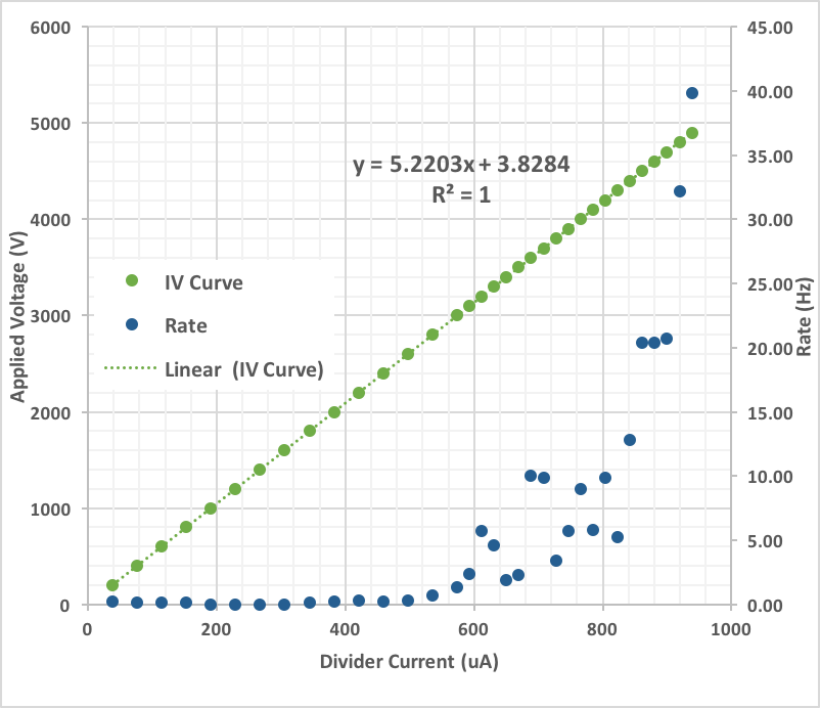


Figure 1.14: The HV as a function of the current is shown (green). The blue trend is the number of spurious signals as a function of HV. The results belong to FIT0004.

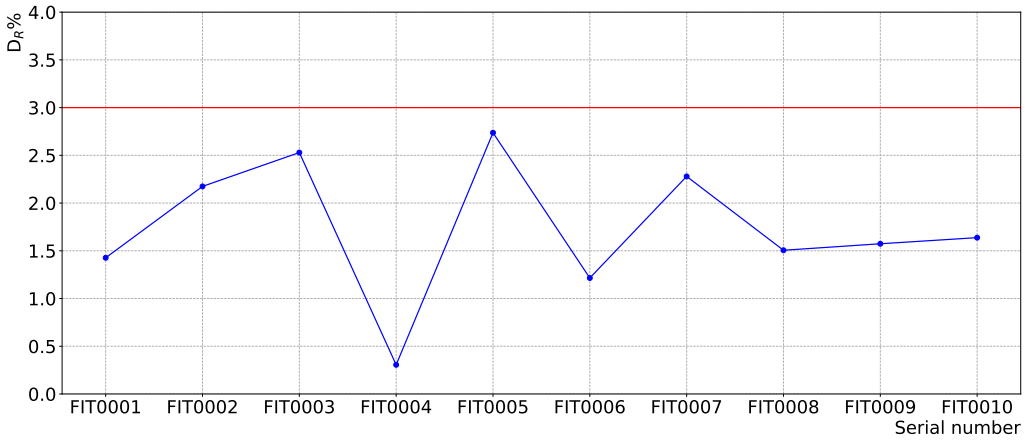


Figure 1.15: The summary of the results for QC step 4 for all Florida Tech-produced GEM chambers. The standard deviation between the measured and nominal HV divider resistance for each Florida Tech chamber, identified by its serial number, is shown. The red line represents the selection criteria (3%).

ionizing energy to be immune to ionization from the cosmic rays. By eliminating the probability of receiving a signal from the detector from the cosmics, I am left with

spurious signals. I monitor the spurious signal rate as a function of applied HV. The GEM collaboration has concluded that the origin of these spurious signals is a coronal discharge from the active area of the GEM along the internal frame (which holds the GEM foils inside the gas volume) to the ground through the anode strips where the signal is readout [11].

Selection criteria: The rate per surface area for intrinsic noise is studied by the GEM collaboration and published in the technical design Ref. [1]. The study establishes that across the entire 3500 – 4000 cm^2 surface area of a normally operating GE1/1 detector, the intrinsic rate should not exceed $\mathcal{O}$ (10^{-2}) Hz/cm. Hence, the selection criteria are the following: a detector passes the QC step 4 if the intrinsic noise rate for the entire detector does not exceed 100 Hz.

Test setup: First, I attach a copper-coated grounding plate to the backside of the readout board to mitigate the environmental RF noise pick-up. As shown in Fig. 1.6(**l**), each readout sector features a Panasonic connector for connecting an amplifier board. I ground each readout sector with Panasonic-to-LEMO adapters with 50 Ω termination and provide grounding by connecting the ground pad of each PCB to the grounding plate with copper tape, as shown in Fig. 1.16. I pick up the signal through the contiguous bottom of the third GEM foil through a decoupling CR differentiator circuit soldered onto the drift board of the GE1/1 detector. The high-pass filter (visible in Fig. 1.5(**f**)) consists of SMD (Surface Mount Devices) resistors and capacitors which are soldered to the drift board. The signals are then processed with a readout chain (Fig 1.17) composed of a charge-sensitive preamplifier, an amplifier, and a discriminator with the threshold set to suppress the environmental noise (typically around 140 mV). The resulting digital pulses go through a counting unit for the rate measurement.

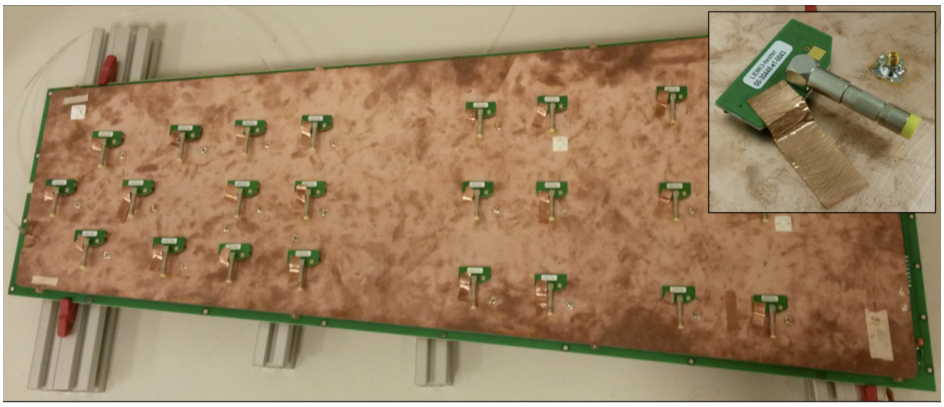


Figure 1.16: An overview of the GE1/1 readout board with the mounted grounding plate and the 50 Ω terminations. The spotty copper plane is an RF shielding plate that is not part of the detector. Spots are from manually handling the plate.

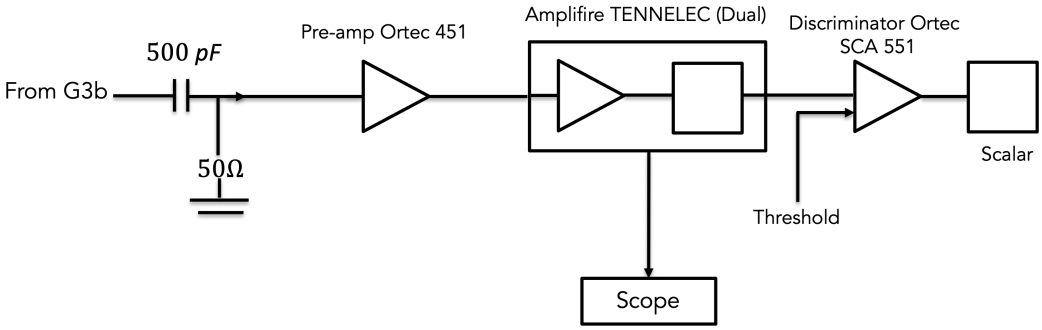


Figure 1.17: The QC step 4 electronic chain setup for data acquisition

Procedure and Results: The detector should be well-shielded, and the RF noise should be combated to fall below 100 mV. I achieve this by grounding every element of the signal processing chain and the grounding plate. Then, I set a −140 mV threshold on the discriminator and count the spurious signals. I ramp up the HV by steps of 200 V from 0 to 3000 V, then by steps of 100 V from 3000 V to 4900 V. Once I reach the maximum voltage, the number of spurious signals determines whether or not the detector passes the QC step. For the detector of Fig. 1.14 the maximum rate is well below the selection criteria (100 Hz). Consequently, the detector moves forward to QC step 5. In Fig. 1.18 I summarize the intrinsic noise measurements of all ten detectors

assembled at Florida Tech.

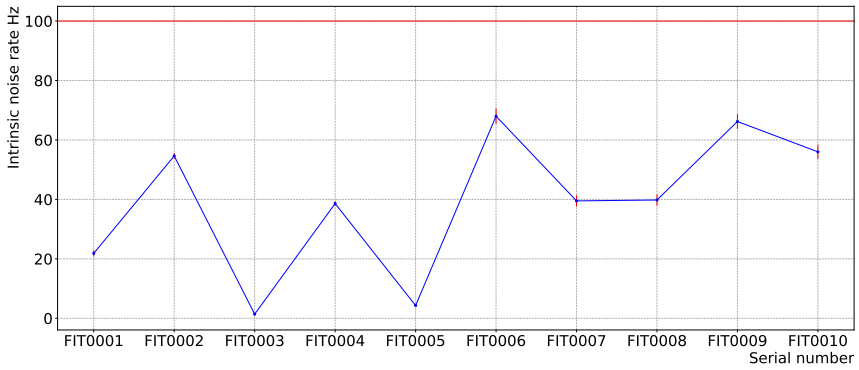


Figure 1.18: The intrinsic noise measurements for all Florida Tech-produced chambers. The red horizontal line represents the selection criteria: intrinsic noise rate = 100 Hz.

1.4.5 QC Step 5a - Effective gas gain measurement

Motivation: The gas gain is the critical parameter in determining the performance of the detector in the proportional mode and the high-rate capability of a gaseous detector. The time resolution of the detector and reconstruction efficiency depend on the effective gain of the detector. Two competing factors decide the gain of a GEM chamber: the amplification factor generated by each GEM foil and the fraction of the electrons lost to *recombination* or absorption by the GEM electrodes. The latter is often called the transparency factor. The *effective* gas gain is determined by measuring the induced current on readout strips, whereas the *absolute* gain takes the induced current on the cathode into account as well.

The effective gas gain of a GEM detector is determined experimentally in the QC process by comparing the primary ionization charge produced by the incoming charged particle in the drift gap to the final amplified charge collected from the readout electrode.

Selection criteria: The selection criteria for effective gas gain are dictated by the required time resolution of 8 ns and detection efficiency of 97% by CMS trigger system and the offline reconstruction algorithm to declare a successful muon hit detection. These figures have been experimentally translated to an effective gain requirement of 2×10^4 through multiple beam test campaigns [20]. To receive a sufficiently uniform gain from the detector system at CMS, we require every triple-GEM detector to have an effective gain within $\pm 1.1\sigma$ ($\pm 37\%$) of this nominal effective gas gain value (2×10^4). Further studies demonstrate that the operational region for GEM detectors is far from the breakdown voltage that would trigger discharges towards the readout electrode with a non-negligible probability [11, 21].

There is an intricacy to be addressed here. At the GE1/1 station, the effective gain of the chamber can be precisely adjusted to the desired value by varying HV settings applied by the PC-controlled HV multi-channel power supply. The gain adjustment with HV is indicated in Fig. 1.21, where the effective gain rises exponentially with the increase in applied voltage. While most Florida Tech GE1/1 chambers do not meet the selection criteria above, they are all accepted based on uniform gain performance across each chamber. Once in their designated position at the GE1/1 station, they are supplied with the appropriate voltage to exhibit the proper gain.

Test setup and procedure: My ultimate goal for this step is to measure effective gain as a function of the current through the HV divider. To measure the effective gain, I use a high-rate X-ray generator with an emission cone size of 120°, which makes it possible to irradiate the entire surface of the GEM chamber. For this purpose, we have a dedicated X-ray box at Florida Tech (Fig. 1.7), which is a custom-made radiation-sealed wooden box. The X-ray gun utilizes an electron gun, beaming onto a gold target,

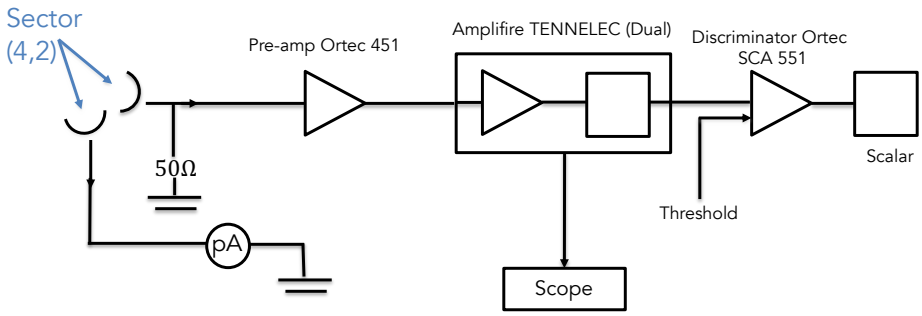


Figure 1.19: The signal processing chain for measuring the rate of incoming photoelectrons converted from X-ray photons. The chain is connected to the central readout sector ($i\eta$=4, $i\phi$=2).

leading to gold de-excitation mainly via L_α and L_β emission lines at 9.71 keV and 11.44 keV, respectively, on top of a Bremsstrahlung continuum. The incident X-ray photons are absorbed by the copper atoms on the drift board, which sequentially emit 8005 eV (8 keV) photons by electromagnetic fluorescence radiation. The fluorescence photons enter the gas volume –primarily the drift gap –and convert to electrons through the photoelectric effect. The primary current, amplification magnitude of which decide the effective gain, is $R \times N_{primary} \times e$; where R is the rate of photoelectron creation, $N_{primary}$ is the number of primary electrons produced by the primary ionization in the drift gap, and e is the electron charge. I measure the rate of produced photoelectrons (R) through a signal processing chain (Fig. 1.19), connected to the central readout sector ($i\eta$=4, $i\phi$=2).

The last ingredient for calculating the gain is the measured current induced on the readout strips. I measure the readout current I_{RO} separately by using a pico-ammeter connected to the same readout sector ($i\eta$ =4, $i\phi$ =2) as the rate measurement, as depicted in Fig. 1.19. The thermal noise for current and rate measurements is subtracted from the final results by pedestal measurements. In conclusion, the gain is calculated by the expression:

$$G = \frac{I_{RO}}{R \cdot N_{primary} \cdot e} \tag{1.5}$$

I propagate the error by assuming Poissonian uncertainty on the photoelectrons rate (R) and a standard deviation for 300 measurements of the readout current I_{RO}.

Measuring the number of primaries with the gold source X-Ray generator: We have good experience in measuring the gain in triple-GEM detectors, but we use an approximated value for the number of primaries, $N_{primary}$, which can cause an error in our gain calculations. I have calculated the number of primaries produced in the drift gap with a gold source X-ray generator. The number of primaries is given by [22],

$$N_{primary} = \frac{\Delta E}{W_I}, \tag{1.6}$$

where ΔE is the total energy deposit in the drift gap by an incoming photoelectron, and W_I is the ionization energy; the minimum energy required to unbound the loosest electron from its molecular/atomic orbit. The ionization energies for Ar and CO_2 are 25 eV and 34 eV, respectively [23]. To calculate the number of primary electrons produced by a photoelectron, I take X-ray fluorescence energy peak of Copper, 8005 eV, as ΔE and weight the ionization energies according to the gas composition in use for QC step 5: 70% of Ar and 30% of CO_2. By substituting these figures into Eq. 1.6, I have,

$$N_{primary} = \frac{8005\ eV}{(0.7 \times 25\ eV) + (0.3 \times 34\ eV)} = 289 \tag{1.7}$$

However, this naive perspective does not accurately portray what occurs in the drift gap, as copper fluorescence is not the only contributing process in energy deposition. In some cases, the incident photon, instead, extracts an electron from the inner K-shell

of the Ar, and the vacancy in the shell is filled by an outer electron, resulting in another photon (X-ray) emission (2.9 keV), leading to an escape peak in the energy spectrum. Furthermore, the fluorescence and escape peak lie on some background which is not negligible. To account for these effects, I use a Multichannel Analyzer (MCA), which receives voltages from the central readout sector of the chamber and sorts them into a spectrum. In Fig. 1.20, I show the output spectrum from the MCA.

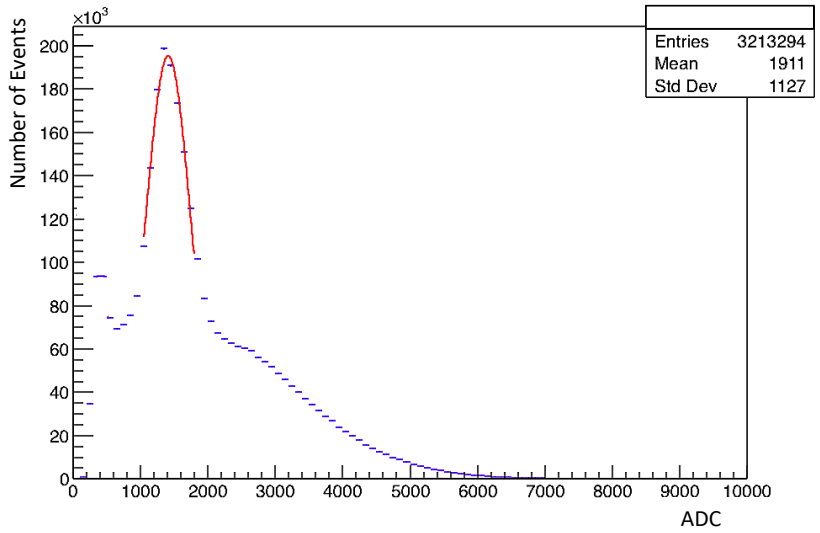


Figure 1.20: The Multichannel Analyzer output spectrum for a radiated chamber by a gold source X-ray generator. The peak is the copper fluorescence on top of a background continuum, including the Argon escape peak.

The MCA produces the energy spectrum as a function of analog-to-digital converter (ACD) units. By calibrating the mean of the main peak (1413 ADC) to the copper fluorescence peak (8005 eV), I find the spectrum mean (1911 ADC) to be at 10826 eV. I substitute 10826 eV for ΔE in Eq. 1.7:

$$N_{primary} = \frac{10826\ eV}{(0.7 \times 25\ eV) + (0.3 \times 34\ eV)} = 390 \tag{1.8}$$

Therefore, I use $N_{primary} = 390$ to calculate the effective gain for all the chambers assembled and tested at Florida Tech.

Results: The effective gain is calculated as a function of the current in the HV divider. The photoelectron rate and the effective gain are shown as functions of applied voltage in Fig. 1.21. The summary plot of Fig. 1.22 shows the measured effective gain for all Florida Tech-produced chambers.

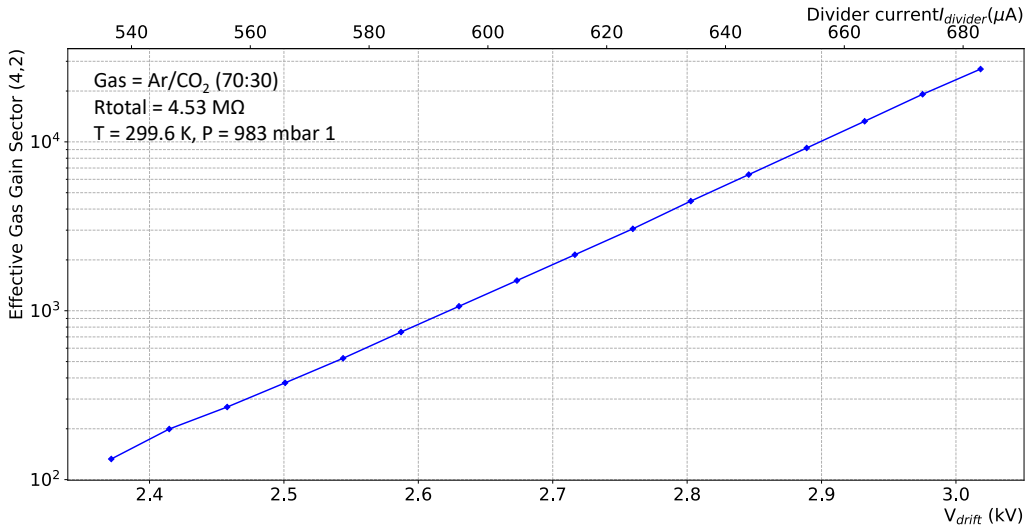


Figure 1.21: A typical effective gas gain, as measured by the author, as a function of the drift voltage and HV divider current. The drift voltage is corrected for pressure and temperature. The error bars on the effective gain measurements, invisible here due to their small amplitude, are included.

Environmental conditions effect on measured gain: The gas gain is the product of a series of Townsend avalanches in the gas volume. The Townsend coefficient, which represents the number of secondary electrons produced by the primary electron per unit path length, is dependent on the ratio between environmental temperature and pressure [24]. I normalize effective gain to a predefined reference pressure and temperature to account for varying environmental conditions across production sites. Rather than directly to the gain, I impose the normalization on the drift voltage (or divider current). The corrected drift voltage is then calculated according to the Eq. 1.9. The

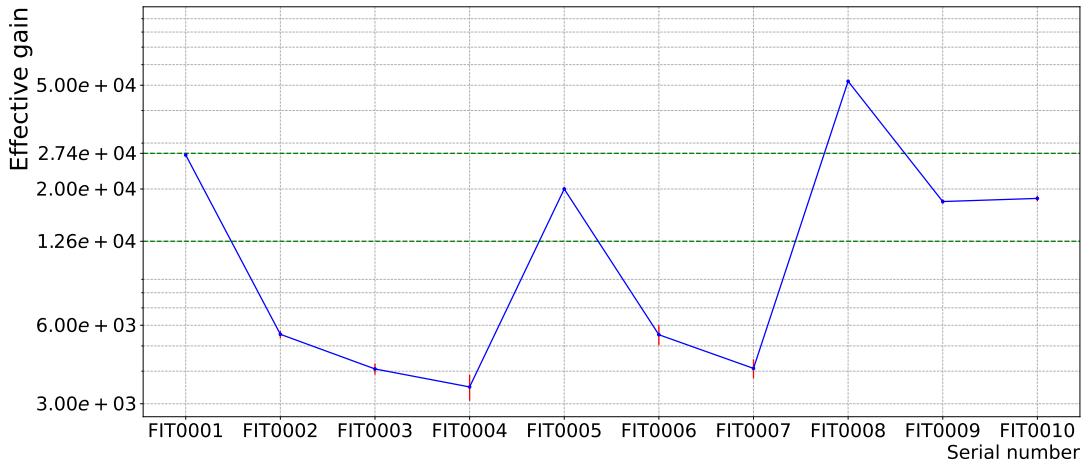


Figure 1.22: The effective measured gain for each Florida Tech-produced chamber is shown. To safely operate the detector while reaching full efficiency, the gain is measured at a maximum divider current of 710 μA. All results are presented after the environmental corrections. The region between the two green horizontal dashed lines represents the selection criteria. The chambers with measured effective gains outside the selection criteria region will be adjusted by varying the HV settings at the GE1/1 station.

corrected gain distribution across the site and the produced detectors is visualized in Fig. 1.23.

$$V_{drift}^{corrected} = V_{drift} \cdot \frac{T_X}{P_X} \cdot \frac{T_X}{P_X} \frac{T_0}{P_0} \tag{1.9}$$

1.4.6 QC Step 5b - Response Uniformity measurement

Motivations: The measured effective gain from the central readout sector ($i\eta$=4, $i\phi$=2) establishes a reference *absolute* gain of the chamber. Nevertheless, it is essential to study the gain variation across the active area of the chamber. A sufficiently uniform response is expected to avoid geometrical trigger or reconstruction biases [1]. The chamber is tested for uniform relative response across all readout strips in this QC step.

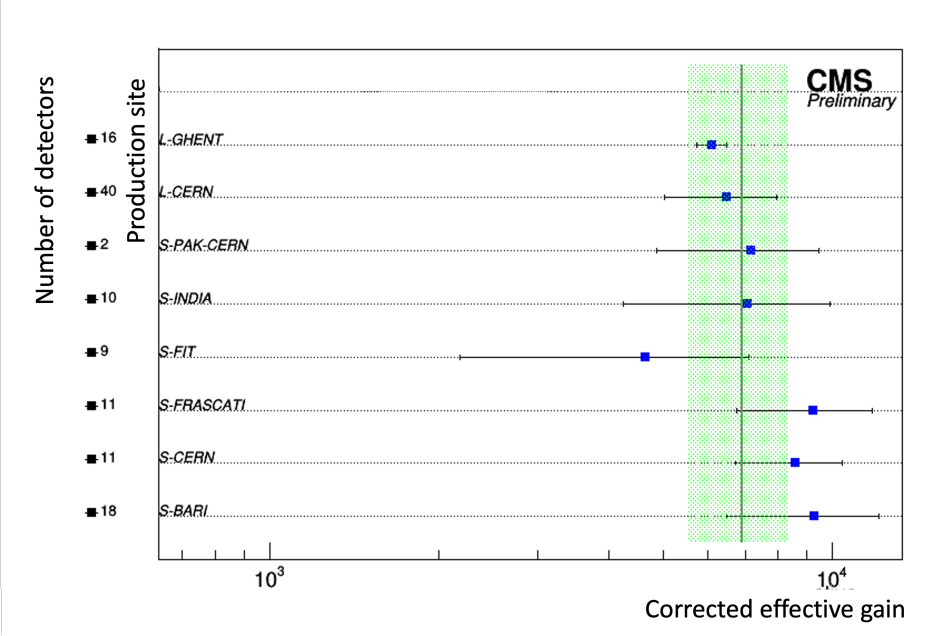


Figure 1.23: The average effective gain of numerous production detectors. The error on each mean is the standard deviation, and the green region is the one standard deviation around the global average. Lower measured gain at Florida Tech (FIT) are compensated for with HV settings adjustments.

Test setup: Due to the possible saturation of analog pipeline voltage 25 (APV25) at the full nominal gain of the detector, the GEM chambers under test are operated at a reduced gas gain (typically between 500 and 600), below the saturation level. I leave the chamber in the X-ray box for the uniformity test, as it is performed by irradiating the entire chamber with the X-ray generator, described in Sec. 1.4.5. I fix 24 APV25 analog readout chips [25] onto the readout Panasonic connectors according to the boss/subordinate sequence mapping provided by the collaboration. The APV25 on each readout sector amplifies the induced charge on the readout strips and sends them to analog-to-digital converters (ADCs) to digitize the signal. The digitized signals are recorded by front-end concentrator cards (FECs), which are components of the larger RD51 Scalable Readout System (SRS) [26]. A schematic overview of this setup is shown in Fig. 1.24.

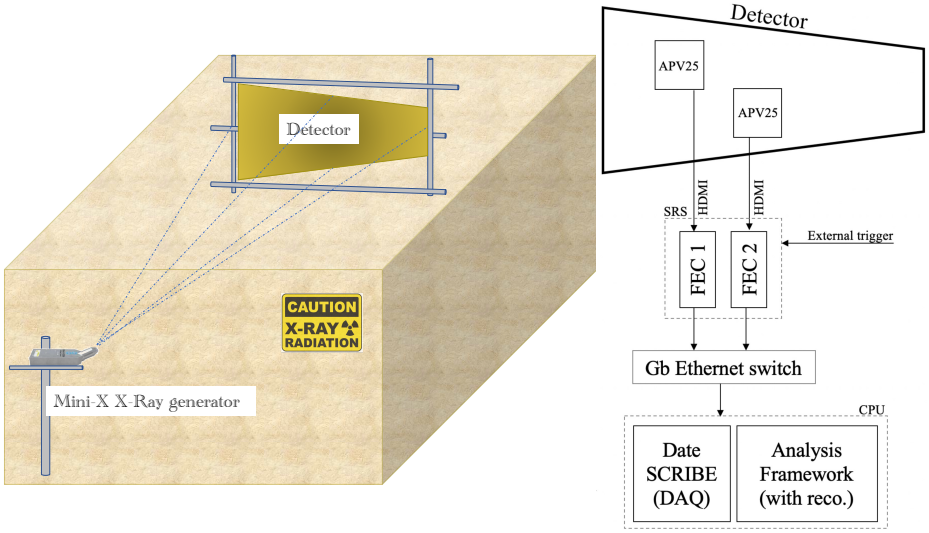


Figure 1.24: Schematic overview of the X-ray station for the QC5 response uniformity test at Florida Tech (left) and the typical data flow in the SRS DAQ from the front-end APV25 to the analysis framework (right).

To correctly capture signals, the triggering of the APV system would ideally come from the bottom of the third GEM foil in the chamber; however, due to the large capacitance of the GEM foil electrode, the electronic noise is typically too high to trigger the APV system reliably. Thus, random triggering is used instead. The usage of random triggering is valid as the X-ray generator rate is on the order of several MHz over the entire chamber, while the size of the APV25 acquisition time window is up to 50 ns shaping time, which means on the average at least one signal is recorded for each random trigger [25, 26].

Selection criteria: As I described in Sec. 1.4.5, the magnitude of the gas gain in a gaseous detector is determined by the combination of the amplification and the transparency factor. There are two factors to consider for gain variation across the chamber: varying hole geometry across the GEM foils (foil thickness and hole diameter) and varying applied electric field due to bending of the readout board or the drift board

after tensioning the foils and closing the chamber[11, 27]. The contributing factors and their contribution in terms of percent uncertainty are summarized in Tab. 1.4.

Table 1.4: The contributing parameters to the overall uncertainty in the gain uniformity response test [11]

Uncertainty	Description	Contribution%
$\sigma_{thickness}$	Gap size between the GEM foils	1
$\sigma_{diameter}$	Diameter of holes in the GEM foil	4.2
$\sigma_{drift\ bending}$	Bending of the drift PCB	25
$\sigma_{RO\ bending}$	Bending of the RO PCB	7.5

The total expected uncertainty on the effective gain across a GE1/1 chamber is derived by combining the values of Tab. 1.4 in quadrature:

$$\sigma_{total} = \sqrt{\sigma^2_{thickness} + \sigma^2_{diameter} + \sigma^2_{drift\ bending} + \sigma^2_{RO\ bending}} = 37.1\% \tag{1.10}$$

Therefore, the QC step 5b selection criteria for a GE1/1 chamber is: the standard deviation of the gain response across all readout strips must be below 37%.

Results: A GE1/1 readout board is divided into 768 regions called *slices* with each slice containing four readout strips. The charge collected from a cluster of four readout strips in a slice is called a strip cluster charge. The SRS system can produce an ADC spectrum for each slice or cluster. The prominent peak in an ADC spectrum for a cluster is the X-ray fluorescence photopeak of Copper. The fluorescence photopeak is located by fitting a Cauchy distribution to the ADC spectrum of the total measured strip-cluster charges. An example of a cluster ADC distribution and the photopeak fit is shown in Fig. 1.25.

The photopeaks of all cluster-strip charges are histogramed, and a Gaussian is fitted to it. The mean (μ) and the standard deviation (σ) of this distribution serve to measure

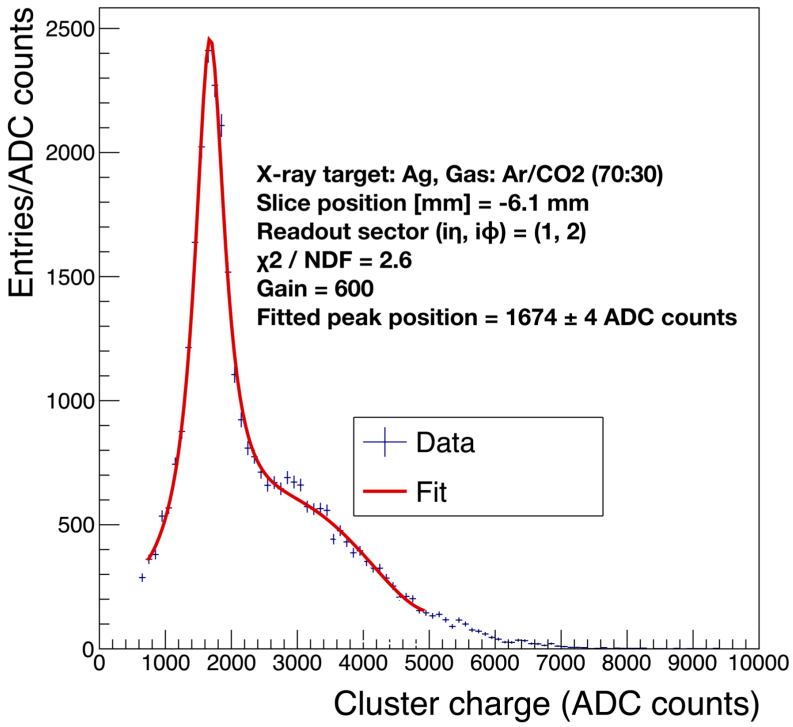


Figure 1.25: The ADC (charge) distribution of 4-strip cluster is shown. The peak is fitted with a Cauchy function and the underlying continuum is modeled with a fifth-order polynomial.

gain uniformity across the chamber. Specifically, the relative response uniformity is reported by the ratio $(\frac{\sigma}{\mu} \cdot 100\%)$. An example distribution for one chamber assembled and tested at Florida Tech is shown in Fig. 1.26. The target ration $\sigma/\mu \times 100 = 22\%$ is well below the passing threshold; therefore, the detector passes the test. The relative gain variation across a GE1/1 chamber can be further visualized in 2D as shown in Fig. 1.27. Fig. 1.28 summarizes the gain uniformity results for all Florida Tech-produced chambers.

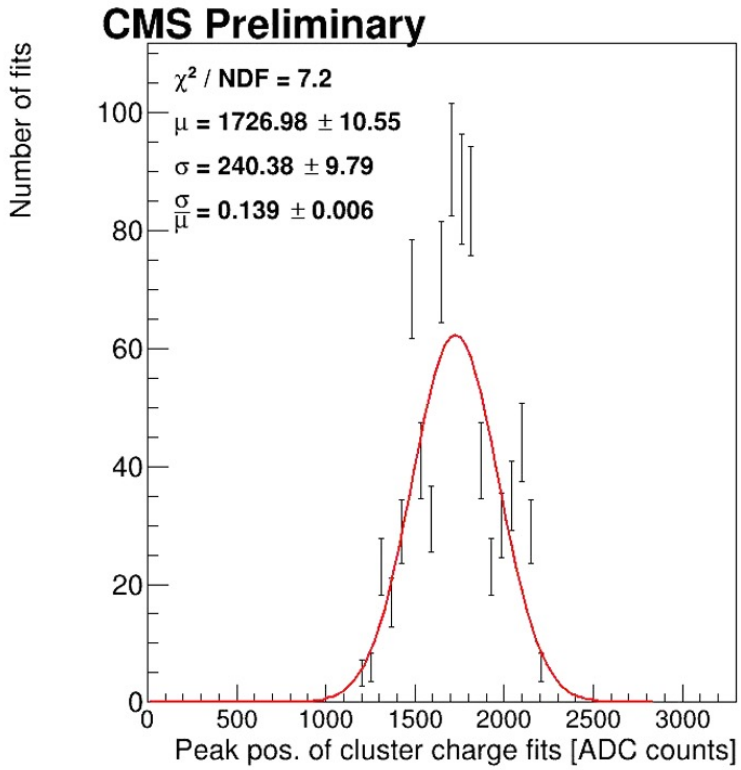


Figure 1.26: The means of the ADC spectra of the aggregate of strip-clusters is shown. The peak of the distribution is obtained through a Gaussian fit. The mean of the Gaussian represents the peak of the distribution. The ratio $\sigma/\mu \times 100 = 22\%$ represents the gain variation across the detector.

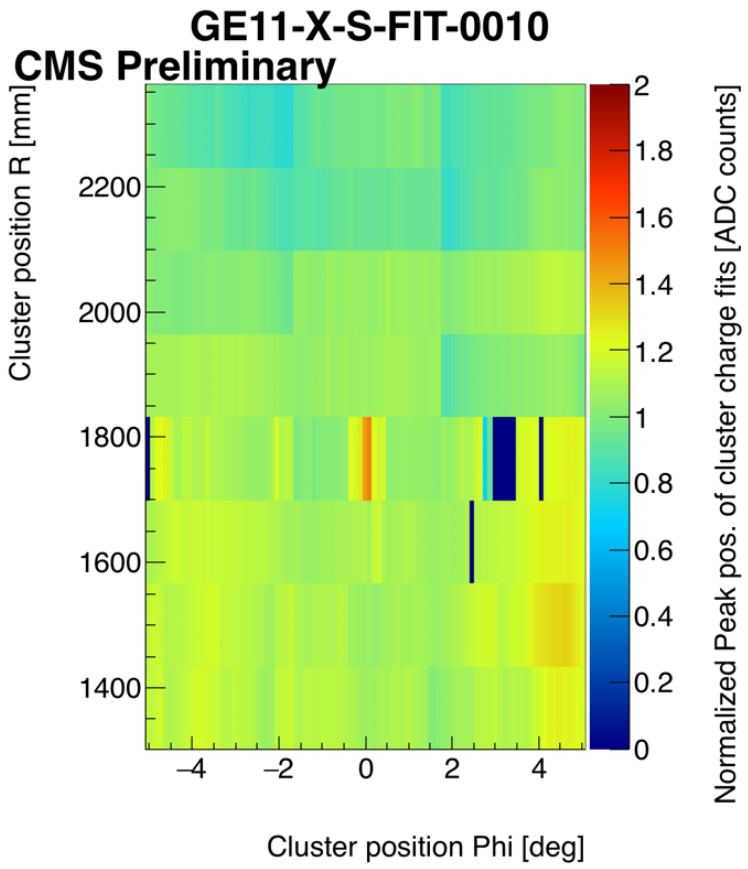


Figure 1.27: The relative gain variation across a FIT-assembled GE1/1 chamber (FIT-0010) is shown. The x-axis is plotted as the *angular* distance from the center of the chamber ($i\phi$) while the y-axis is the *radial* distance from the beamline. The binning in the horizontal axis corresponds to four-strip slices, while the vertical binning corresponds to the eight $i\eta$ sectors on the chamber. The color map is the normalized peak position of the cluster charges to the chamber average. The dark blue represents the strips with failed fits.

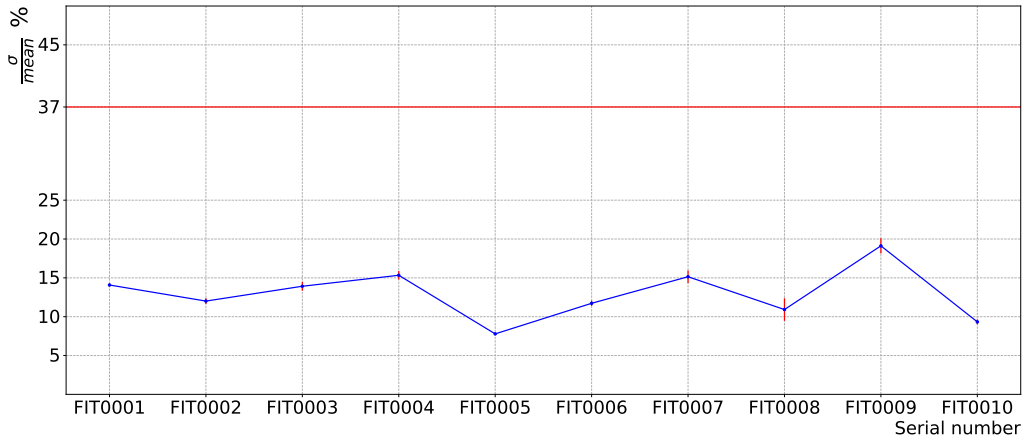


Figure 1.28: The summary of the results for QC step 5, gain uniformity, for all Florida Tech-produced GEM chambers. The red horizontal line represents the selection criteria: the standard deviation of the gain response across a chamber = 37%.

1.5 Concluding Remarks

Under the supervision of Dr. M. Hohlmann, my team and I assembled and thoroughly tested 10 GE1/1 chambers. The campaign for producing the final version chambers, which are all now installed at CMS endcap, proceeded with a prototyping campaign in which the various assembly processes, testing processes, and design issues were subjected to serious scrutiny. In this endeavor, several collaborating institutions from around the globe took on the challenge to mass-produce GE1/1 detectors. Our research group at the Florida Institute of Technology was among those able to demonstrate their excellent capability to the GEM collaboration at CMS for this collaboration. While for much of the process I was charged with the responsibility of the production, several undergraduate and graduate students participated and contributed to the successful execution of this practice. Additionally, the GEM collaboration actively engaged in a supporting role by conducting regular meetings and effectively responding to the needs and concerns of the satellite sites. The following image portrays the Florida Institute of Technology team after successfully assembling and testing 10 GE1/1 detectors. The gleeful individual with the victory fingers in the sky is Dr. M. Hohlmann, who provided vital direction and supervision for this journey.

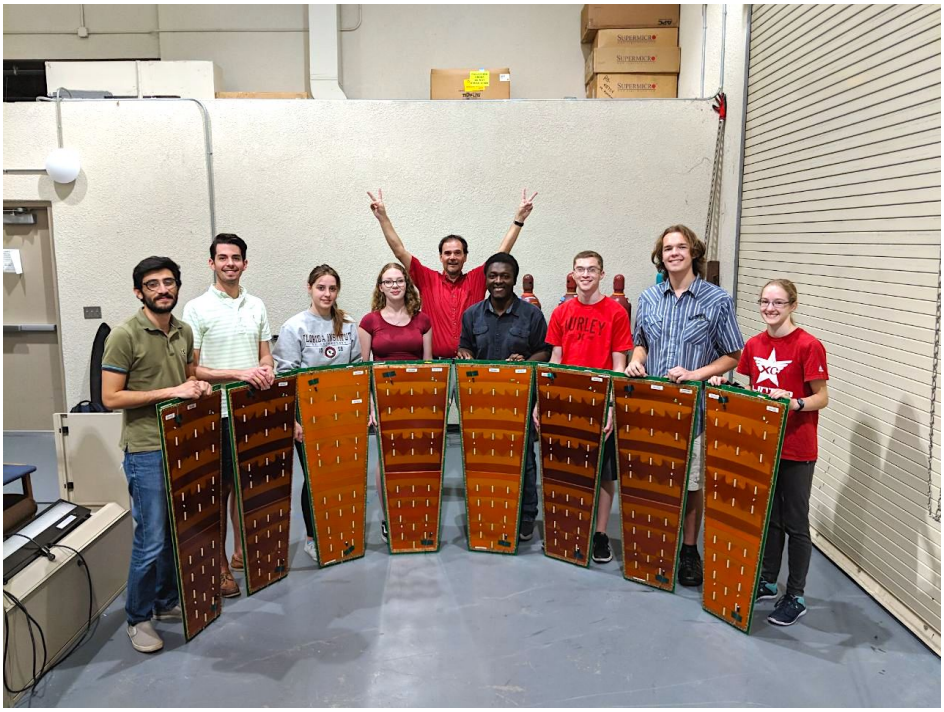


Figure 1.29: The victory photo of the Florida Institute of Technology GE1/1 detector production team. From left to right: The author, S.Butalla, J. Miksanek, S. Arends, M. Hohlmann, J. Collins, J. Hammond , M. Werbiskis, S. Wohlstadter

Chapter 2

The Standard Model

More than a century of particle physics has radically contributed to our contemporary understanding of the fundamental laws and systems that shape our reality. From the development of quantum mechanics to the experimental discovery of particles, from re-framing particles as fields and the introduction of Quantum Field Theory to the struggle to unify the fundamental forces of nature, we have come a long way. Each successful step forward has opened up a broader landscape of unknowns: dark matter, origins of the particles and forces, the validity of super-symmetry, to name a few of these exciting yet challenging questions that we still bear. In this chapter, I lay the foundation for the physics analysis portion of this thesis by introducing the stepping stone of modern particle physics, which is the standard model, developing the basics of quantum field theory, and conclude the chapter by presenting more recent notions in particle physics such as the electroweak theory and the Higgs mechanism.

2.1 The Standard Model

The Standard Model (SM) of particle physics is a robust mathematical framework that describes three fundamental forces: electromagnetic, weak, and strong. The gravitational force does not fit into the SM picture as the quantum field theory of gravity is not yet well understood. Decades of experimental endeavor are in impressive agreement with the SM with high prediction power [28]. As illustrated in Fig. 2.1, the SM classifies the fundamental particles into three generations (I, II, III) of *quarks* (in purple) and *leptons* (in green), ordered according to their masses: the lightest (most stable) particles belong to generation I, while the heavier (less-stable) particles belong to generations II and III. Quarks and leptons belong to a more general class called fermions as they all have a spin of $\frac{1}{2}$.

Quarks, fundamental particles that are building blocks of matter, combine to form composite particles called *hadrons*, the most stable of which are familiar particles such as protons and neutrons, the constituents of atomic nuclei. Quarks come in six flavours: up (u), charm (c), top (t), often categorized as *up-type* quarks, and the other class, *down-type* quarks, consist of down (d), strange (s), and bottom (b). Each up-type quark carries an electric charge of $\frac{2}{3}$ of elementary charge e, while each down-type carries $-\frac{1}{3}$. For each quark, there exists an anti-quark, denoted by a bar on top of the symbol referring to the quark ($\bar{q}$). As in the generic antimatter, anti-quarks have the same mass, mean lifetime, and spin as their respective quarks, but the electric charge and other charges have the opposite sign. The SM exhibits this *symmetry* by including six flavors of leptons in three generations. Charged leptons are: electron (e), muon (μ), and tau (τ); and for each there exist a charge-neutral *neutrino*: ν_e, ν_μ, and ν_τ.

There are three fundamental forces at work within the SM interpretation in the universe. The fundamental forces are characterized based on the types of particles that

experience the forces, relative strengths, the range over which the forces are effective, and the nature of the particles that mediate the forces. Not included in the SM, gravity, the most discernible of all in our day-to-day life, is, in fact, the weakest force among the fundamental forces of nature, even though, when it comes to range, gravity has an infinite range. The electromagnetic force, in principle, can be experienced over an infinite range despite being much stronger than gravity. With a subatomic range of effectiveness, the weak force is stronger than gravity but weaker than other forces. Lastly, the mightiest of all four, the strong force operates in exclusively subatomic ranges. Leaving gravity aside, as we do not have experimental evidence within the bounds of SM, there exists a force carrier for each of the fundamental forces. Force carries constitute the other general class of particles within the SM: bosons! Bosons, depicted in red in Fig. 2.1, have integer spins and mediate the forces that govern the interactions between the fermions. Photons, denoted by γ, mediate the electromagnetic force; it is a spin-1 boson with no mass. Z, W^+, W^- are massive spin-1 bosons and are the force carriers for the weak force. Mediators for the strong force are the spin-1 massless gluons (g).

I have now introduced all particles in Fig. 2.1, but one: the Higgs boson, displayed in yellow. A massive boson with the spin of 0. So far, I have associated every boson with its respective fundamental force.; where does this leave the Higgs boson? What is the justification for the existence of this spin-0 boson? All force carries emerge mass-less if the SM theoretical procedure is followed faithfully. While valid in the case of the photon (γ), experimental evidence exposes that weak force carriers Z, W^+, and W^-, bear masses. To explain the experimental evidence, Peter Higgs, Robert Brout, and François Englert proposed a solution [29, 30]. The *Brout-Englert-Higgs mechanism* asserts that the interaction with the omnipresent Higgs field is responsible for depositing mass in the weak bosons. In the coming sections, I attempt to briefly

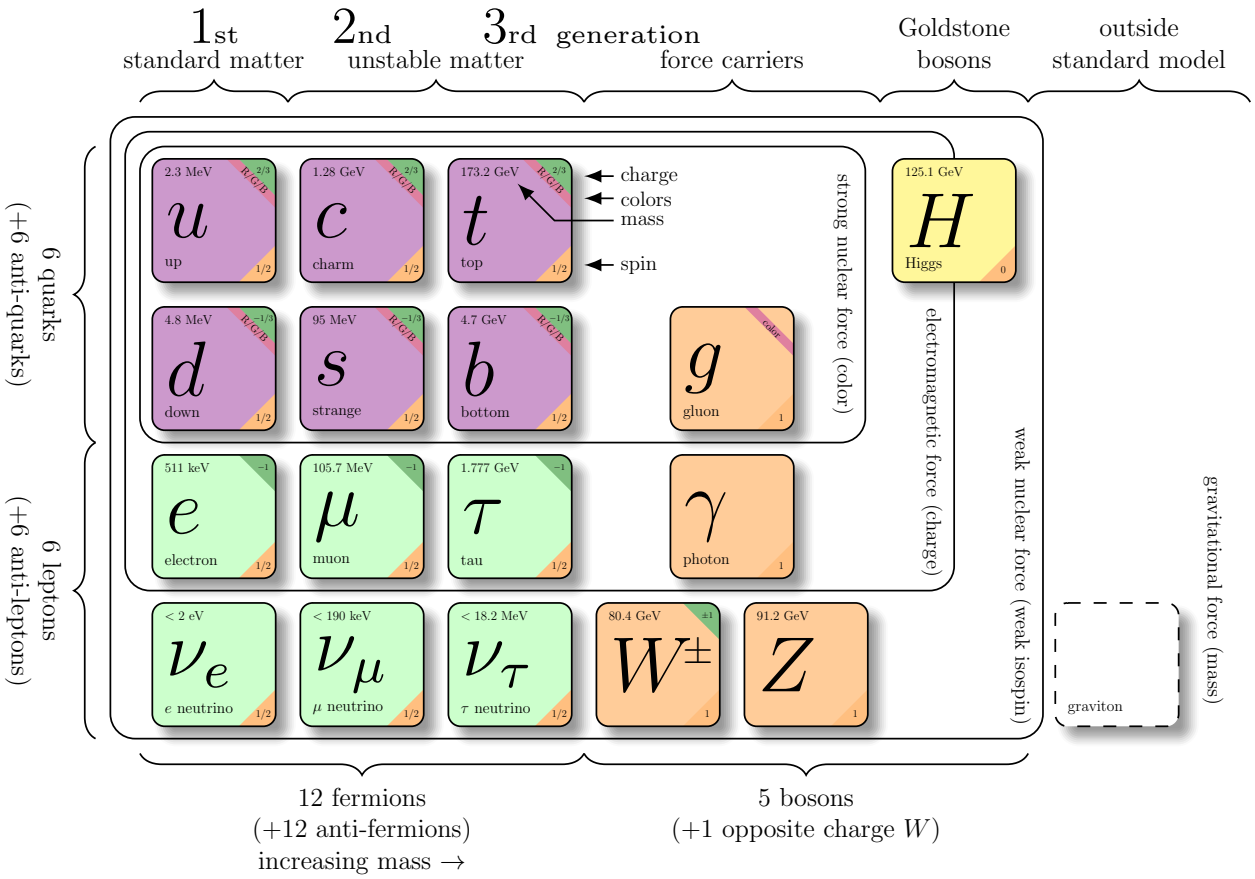


Figure 2.1: Elementary Particles in the Standard Model [3] are illustrated. The undiscovered gravity-associated boson lies outside of the standard model picture.

expand on some of the critical concepts of the present dissertation, introduced in the above general passage about the SM.

2.2 Quantum Field Theory

Quantum Field Theory (QFT) is a mathematical and conceptual framework for contemporary particle physics. Widely accepted, QFT can be described as an extension of Quantum Mechanics (QM) for systems of many particles (large degrees of freedom). The legendary Paul Dirac formulated the idea in his famous 1927 paper on *"The Quantum Theory of the Emission and Absorption of Radiation"* [31]. While it is conceivable

to tackle the problem of a single particle at the quantum level in QM, it is unattainable to describe a system of particles with a varying number of particles within the QM framework. This immediately becomes a desperate issue when describing interactions such as scattering, as particles are created and annihilated in the process. Moreover, classical QM is unqualified for describing electromagnetic phenomena since electromagnetic phenomena are inherently relativistic. More specifically, I have to confront relativistic QFT for this dissertation, where the scattering of the high-energy particles plays a central role. Historically the step toward QFT was achieved through an attempt to reconcile QM with Special Relativity Theory (SRT). Reconciliation is achieved by requiring the Schrödinger equation, i.e., the fundamental law for the temporal evolution of the quantum mechanical wave function, to remain invariant under Lorentz transformations. The Klein-Gordon or Dirac equations were derived in the 1920s to facilitate this process. Despite being a milestone, relativistic QM is not quite QFT; the defining factor that separates QM from QFT is degrees of freedom: QFT is a quantum description of systems with infinite degrees of freedom, more accurately what we refer as to *fields* [32, 33].

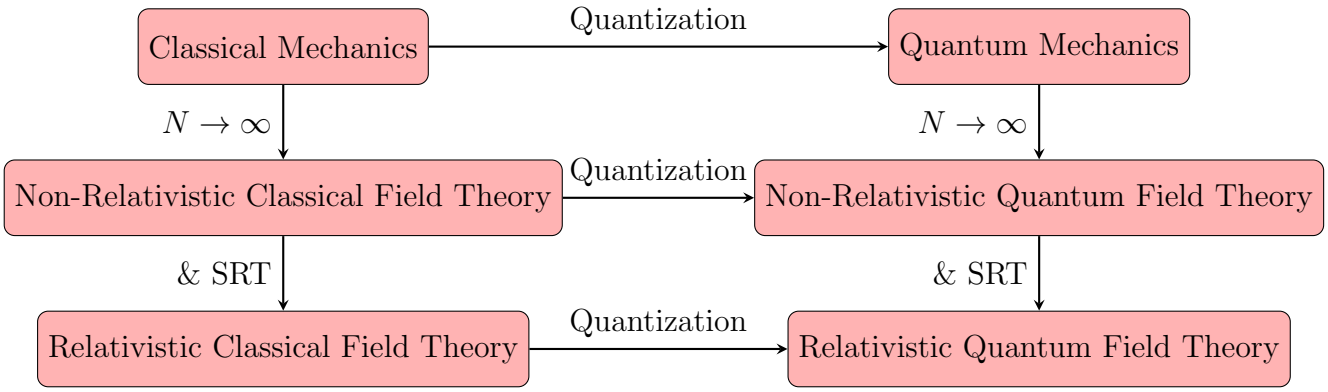


Figure 2.2: The info-graph shows how theories can be quantized to evolve to their quantum versions, how they can be transformed to their non-relativistic counterparts through considering an infinite number of degrees of freedom, and how adding SRT can transform field theories to their relativistic versions.

As shown in Fig. 2.2 the move from classical mechanics toward QM and subsequently to QFT starts with *quantization* which is achieved by establishing commutation relations. Analogous to QM, where commutation relations result in operator-valued physical states such as momentum (p) and spatial coordinates (q), commutation relations in QFT leads to operator-valued *fields*, namely ϕ for spatial and π for momentum coordinates. Having defined the Lagrangian, $\mathcal{L}$, the Hamiltonian equations transform as follows,

$$p = \partial\mathcal{L}/\partial\dot{q} \rightarrow \pi = \partial\mathcal{L}/\partial\dot{\phi}. \tag{2.1}$$

In Eq. 2.1, ϕ and π are conjugate fields as q and p are in classical Lagrangian. This leads to the commutation relations in Eq. 2.2:

$$\begin{aligned} [\phi(\mathbf{x},t),\pi(\mathbf{y},t)] &= i\delta(\mathbf{x}-\mathbf{y}) \quad (\hbar = 1) \\ [\phi(\mathbf{x},t),\phi(\mathbf{y},t)] &= [\pi(\mathbf{x},t),\pi(\mathbf{y},t)] = 0 \end{aligned} \tag{2.2}$$

Similar to its QM counterpart, the QFT commutation relations *quantize* the fields; nevertheless, it is essential to note a few subtle differences in the interpretations of commutations relations between QM and QFT. First, commutation relations in QM refer to quantum states with finite degrees of freedom, while in QFT, the fields are quantum objects of infinite degrees of freedom, each leading to their respective commutation relations. Second, there is a difference between the QM spatial state, $\Psi(\mathbf{x},\mathbf{t})$ and the QFT $\phi(\mathbf{x},\mathbf{t})$ as in QM observable operators act on $\Psi(\mathbf{x},\mathbf{t})$, whereas in QFT $\phi(\mathbf{x},t)$ is operator-valued itself and acts on spatial states. This clarification leads to a critical understating of the role of operators, states, and fields across QM and QFT. In QM, it is understood that quantum states carry a spatial and temporal substance,

translating into probabilities, and operators merely act on these states to transform them. In QFT, however, the quantum field operators can allow for spatial and temporal interpretation.

2.3 Interaction in Quantum Field Theory

As discussed in the above section, quantum field operators are the main playing ground in QFT. A free particle is generated when it comes to interaction with its field. What does interaction mean within this paradigm? To answer this question let us split the Hamiltonian into two parts: $H = H_0 + H_{int}$. Here, H_0 describes the free system, and H_{int} is the interaction Hamiltonian. In doing so, the free part Hamiltonian can be solved exactly. The new emerging terms can absorb this part by redefining the fields appropriately. We can derive commutation equations and equations of motions for interacting fields, the same way we can derive them for the free fields. Let us take the case of a simple scattering problem which is instrumental to this thesis. In the face of the scattering problem, QFT offers a matrix that holds all the prediction powers for the interaction, e.g., scattering cross-section. This matrix often referred to as the *S-Matrix* or the scattering matrix, is a unitary matrix connecting the incoming states ($|in\rangle$) to outgoing states ($\langle out|$) in the Hilbert space by specifying the *transition amplitude*. The square of this quantity can be interpreted as the probability of scattering interaction and consequently be tested against experimental results [34].

I cannot conclude this section without noting perturbation theory. What I have summarized in the previous section is only applicable to the case of free fields. Once we include interacting fields, we are immediately faced with the problem of infinite terms in the Hamiltonian. To mitigate the impossibly of the precise calculation of infinities in the problem and to reserve the possibility of validating the theory prediction against

experimental results, calculations in QFT are often considered only in the infinitesimal neighborhood of the *free* fields; this procedure is named *perturbative quantum field theory* (pQFT) [35].

2.4 Gauge Invariance

A gauge theory in QFT is defined as a type of field theory wherein the Lagrangian is kept intact under *gauge transformations*. A gauge transformation is a local symmetry transformation that is a smooth function of space $\mathbf{x}$ and time t. The family to which all these transformations belong —the local transformations that keep the Lagrangian invariant form a *Lie group* —is called a *symmetry group*, or *gauge group* under group theory. In other words, these transformations are a mapping of the Lagrangian to itself. A subset of a symmetry group such that every element of the group can be expressed as a combination of its members (via the *Lie algebra*) is called the *group generator* subset. Furthermore, for every group generator subset, there exists a corresponding vector field called the *gauge field*. Quantizing these gauge fields by deriving the commutation relations —as seen in the previous sections —for each field, there arises a *quantum* of the field. In a practical sense, the quantum of a vector field, such as the electromagnetic field, is the gauge boson of that field, which in the electromagnetic case is the photon (γ). Historically, the first widely recognized gauge theory is attributed to Pauli in his 1941 paper, *"Relativistic Field Theories of Elementary Particles"* [36].

Let us go over the simple case of electrodynamics, an intrinsically gauge invariant theory, and start with the classical formulation (Eq. 2.3), where $\mathbf{E}(\mathbf{x}, t)$ is the electric field and $\mathbf{B}(\mathbf{x}, t)$ is the magnetic field:

$$\begin{aligned} \mathbf{B} &= \nabla \times \mathbf{A} \\ \mathbf{E} &= -(\partial \mathbf{A}/\partial t) - \nabla \phi \end{aligned} \tag{2.3}$$

In Eq. 2.3, known as *Maxwell's* equations [37], $\mathbf{A}$ is the magnetic vector potential and ϕ is the scalar potential. The covariant formulation of Maxwell's equations is expressed by the *field tensor* $F^{\mu\nu}$ (Eq. 2.4) [38].

$$F^{\mu\nu} = \partial^{\mu} A^{\nu} - \partial^{\mu} A^{\nu} \tag{2.4}$$

Here $A^{\mu} = (\phi, \mathbf{A})$ is the electromagnetic 4-vector potential. From Eq. 2.4, it is immediately apparent that there is a freedom in choosing A^{μ}. Consider the following *gauge transformations*

$$\begin{aligned} \mathbf{A} &\rightarrow \mathbf{A} - \nabla \psi \\ \phi &\rightarrow +\phi + \partial \chi / \partial t; \end{aligned} \tag{2.5}$$

or more compactly, similar to Eq. 2.4,

$$A^{\mu} \rightarrow A^{\mu} + \partial \chi. \tag{2.6}$$

This suggests that no matter the choice of ψ or χ, the field tensor will have the same value. In other words, the field tensors stay invariant under this gauge transformation. The transformations of Eq.2.6 are local transformations as χ is a function of space and time.

To translate this to QFT notation, the 4-vector potential A^{μ} has to turn into a field operator [39]. The gauge transformation for a unitary group is expressed by a phase

rotation, operated by an exponential with an imaginary argument.

$$\phi(x) \rightarrow e^{-i\alpha(x)}\phi(x) \tag{2.7}$$

Ensuring that the Lagrangian stays invariant under the local gauge transformations, as prescribed by Yang-Mills gauge theory [40], new terms should be introduced to the Lagrangian. These new terms are the gauge fields that keep the Lagrangian invariant under the gauge transformations. The SM is indeed a gauge QFT, with symmetries of the unitary product group $U(1) \otimes SU(2) \otimes SU(3)$. From each of the three components of the unitary product group, a gauge field will emerge; quantizing the gauge field will introduce the gauge boson. For example, the gauge field for the electromagnetic $U(1)$ symmetry group is the electromagnetic four-potential, A_μ, with photon γ as its gauge boson. $SU(2)$ gives rise to three weak bosons (W^+, W^-, and Z) and $SU(3)$ generates eight gluons (g).

2.5 Electroweak Theory

In an ever-ongoing struggle to unify all fundamental force and symmetry of nature under a more general symmetry, the electromagnetic and weak gauge symmetries were successfully unified ($U(1) \otimes SU(2)$) in the 1960s by Sheldon Glashow, Abdus Salam, and Steven Weinberg (GSW) for which they won a Nobel prize in 1979 [41, 42, 43]. This unified theory is later appropriately called the electro-weak theory. The gauge transformation, following the structure prescribed by Eq. 2.7, for ensuring the $U(1)$ symmetry is,

$$\psi(x) \rightarrow e^{ig\zeta(x)\frac{Y}{2}}\psi(x), \tag{2.8}$$

where g is the coupling constant, a quantity defining the strength of an interaction, and $\zeta(x)$ is the signal function that specifies the local phase in each point in space-time. The gauge field that grants the invariance of the Lagrangian under this transformation is denoted by B_μ which couples to a brand-new charge called the *hypercharge* (Y).

Invariance under $SU(2)$ demands:

$$\phi(x) = e^{ig_w \alpha(x).\mathbf{T}} \phi(x) \tag{2.9}$$

The $\alpha(x)$'s are three functions that specify the local phases at each point in space-time, one for each generator and g_w is the coupling constant. Three gauge fields conserve the invariance of the Lagrangian for the weak interaction: W_1, W_2, W_3. Here, T is the $SU(2)$ generator, as Y was for $U(1)$. We fix the representation of $SU(2)$ to be $T^i = \frac{1}{2}\sigma^i$ (i.e., the fundamental representation). This will determine the Higgs field representation, when I discuss the Higgs mechanism. The σ^i are Pauli matrices, listed in Eq. 2.10.

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.10}$$

All generators for the unified $U(1) \otimes SU(2)$ need to satisfy the commutation relations. The results of the commutation relations fix Y, the generator of $U(1)$ to $\frac{1}{2}$, and the $SU(2)$ generators to Pauli matrices. Unlike $\psi(x)$ for $U(1)$, which is represented by a unitary matrix, $\phi(x)$ is represented by a doublet, called the iso-spin doublets. The weak isospin doublets must contain different flavors of fermions differing by one unit of electric charge to respect the electric charge conservation law. Two examples of weak iso-spin doubles are listed in Eq. 2.11.

$$\begin{bmatrix} \nu_e \\ e^- \end{bmatrix}, \begin{bmatrix} u \\ \bar{d} \end{bmatrix} \tag{2.11}$$

As a result of *symmetry breaking*, discussed in Sec. 2.6, the physical $W^{\pm}$ can be written as linear combination of the W_1 and W_2:

$$W^{\pm} = \frac{1}{\sqrt{2}}(W_1 \mp iW_2) \tag{2.12}$$

Moreover, as a consequence of spontaneous symmetry breaking, as described in Sec. 2.6, the electromagnetic photon γ and Z of weak interaction can also be written as linear combination of U(1) and SU(2) generators i.e., B and W_3:

$$\begin{bmatrix} \gamma \\ Z_0 \end{bmatrix} = \begin{bmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{bmatrix} \begin{bmatrix} B \\ W_3 \end{bmatrix} \tag{2.13}$$

$$e = g_W \sin\theta_W = g_Z \sin\theta_W \cos\theta_W \tag{2.14}$$

Here, e is the electric charge, g_W is the weak coupling constant for W boson, and g_Z is the weak coupling constant for Z boson. Furthermore, as a result, the electric charge can be formulated as a linear combination of the third component of the weak iso-spin, T_3, and the hypercharge:

$$q = T_3 + \frac{1}{2}Y \tag{2.15}$$

2.6 Higgs Mechanism

The equations of the electroweak theory correctly identify the gauge fields and their associated gauge bosons, namely the photon (γ), $W^{\pm}$ and the Z boson, with a significant caveat: All the bosons and fermions that interact with them emerge with no mass! Inclusion of mass terms in the electroweak Lagrangian kneecaps its invariance under local phase transformations such as Eq. 2.7. Finding theories with massive bosonic vector fields was brought to a halt before gauge theory and the Higgs mechanism was proposed [29]. The Higgs mechanism serves both to give the weak bosons, $W^{\pm}$ and Z, their masses and also the masses of the fermions that interact with them. To address the mass problem, we can extend the electroweak model by adding a single charged complex scalar field, the Higgs field, which couples to the electroweak force and demands that the Lagrangian stays invariant under gauge transformations. The Higgs field representation, as determined in Sec. 2.5, due to the necessity of commutation of all members of generator group of $U(1) \otimes SU(2)$, can be written down as a $SU(2)$ doublet:

$$\phi = \begin{bmatrix} \phi^{+} \\ \phi^{-} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{bmatrix} \tag{2.16}$$

The scalar potential generated by this field can be written as:

$$V(\phi) = \mu^2 \phi^{\dagger}\phi + \lambda(\phi^{\dagger}\phi)^2 \tag{2.17}$$

There could be two assumptions for μ^2 in Eq. 2.17: 1) $\mu^2 > 0$: this would mean that the $\phi = 0$ that minimizes $V(\phi)$, we call this state the *vacuum state* . In this case, the theory is quantum electrodynamics with massless bosons and a scalar field with the mass of μ with an unbroken symmetry. This assumption would not address the problem at hand, the massive bosons of the theory dictated by experimental results.

2) $\mu^2 < 0$: This is the familiar potential known as the *Sombrero* potential illustrated in Fig. 2.3. In this case, the minimum potential energy is not at $\phi = 0$ but rather an infinite set of degenerate minima [44]:

$$\langle\phi|\phi\rangle = \phi^\dagger\phi = -\frac{\mu^2}{2\lambda} = \frac{v^2}{2} \tag{2.18}$$

Here, the central concept is the average or the expectation value of the ϕ operator in the vacuum. Also denoted by v, referred to as vacuum expectation value *(VEV)*, this nonzero value underlies the Higgs mechanism of the SM. It is readily apparent that the phase of ϕ is irrelevant, and all choices yield the same energy. Hence the system has a massless degree of freedom corresponding to rotational fluctuations of the field around the brim of the sombrero, spontaneously breaking the global $SU(2) \otimes U(1)$ symmetry of the Lagrangian [4].

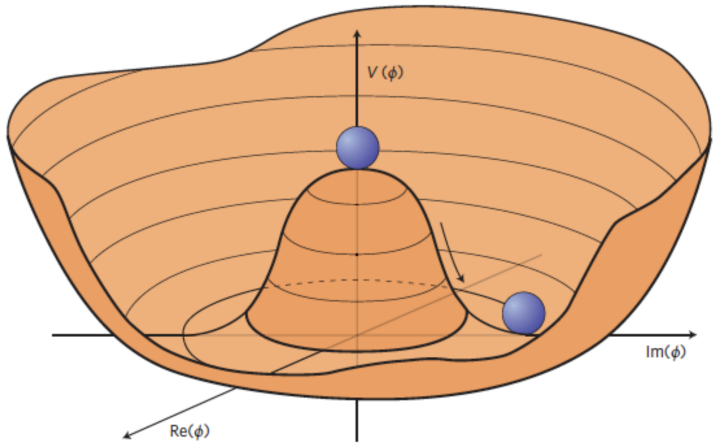


Figure 2.3: An effective sombrero potential that leads to spontaneous symmetry breaking in the electroweak theory [4].

The Lagrangian relevant to this field can be written as:

$$\mathcal{L}_{Higgs} = (\partial_\mu\phi)^\dagger(\partial^\mu\phi) - V(\phi) \tag{2.19}$$

If we parameterize our complex scalar field as the following:

$$\phi = (h + v)e^{i\alpha}, \tag{2.20}$$

where α is the massless real scalar field that generates a rotational degree of freedom, and h is a massive real scalar field. Eq. 2.20 suggests that the Lagrangian ($\mathcal{L}_{Higgs}$) is only dependent on the derivatives of α. If there were no other gauge fields coupled to ϕ, the scalar field α would be identified as the so called Goldstone boson corresponding to the spontaneous symmetry breaking [45]. However, there is a gauge field in the electroweak theory, demanding the local invariance:

$$\phi \rightarrow e^{i\theta(x)\phi} \tag{2.21}$$

And in the case of $U(1)$:

$$A_\mu \rightarrow A_\mu - \partial_\mu \theta(x) \tag{2.22}$$

If we set $\theta(x) = -\alpha(x)$, and plug it back into the Lagrangian, $\alpha(x)$ is eliminated from the Lagrangian, and we say that it has been "eaten" to give the vector fields mass. We can identify that the massive scalar field h is the SM Higgs field at this stage. The mass of the Higgs boson is given by:

$$m_H = \sqrt{2\lambda}v \tag{2.23}$$

The covariant derivative of the Lagrangian of Eq. 2.19 can be written as follows:

$$\partial_\mu \rightarrow D_\mu = \partial_{mu} + ig_W \mathbf{T} \cdot \mathbf{W}_\mu + ig' \frac{Y}{2} B_\mu, \tag{2.24}$$

where $\mathbf{T}, \mathbf{W}_\mu, Y$ and B_μ are the generators of $SU(2) \otimes U(1)$ that were already intro-

duced in this section.

Now we can explicitly derive the expressions shown in Eq. 2.12 and Eq. 2.13 by calculating the scalar kinetic energy expectation term; using the Higgs field, we can derive an expression for the physical gauge fields of $SU(2) \otimes U(1)$,

$$\begin{aligned} W^{\pm} &= \frac{1}{\sqrt{2}}(W_{\mu}^{1} + iW_{\mu}^{2}) \\ Z^{\mu} &= \frac{-g'B_{\mu} + gW^{3}{}_{\mu}}{\sqrt{g^2 + g'^2}} \\ A^{\mu} &= \frac{gB_{\mu} + g'W^{3}{}_{\mu}}{\sqrt{g^2 + g'^2}} \end{aligned} \tag{2.25}$$

Gauge bosons of the above fields acquire the following masses from the Higgs mechanism:

$$\begin{aligned} {M_W}^2 &= \frac{1}{4}g^2v^2 \\ {M_Z}^2 &= \frac{1}{4}(g^2 + g'^2) \\ M_A &= 0 \end{aligned} \tag{2.26}$$

The Higgs mechanism can promptly generate the fermion masses characterized by Yukawa couplings [46]. Dirac fermions can couple to the Higgs field and obtain mass. Their coupling, often referred to by g_f, is given below:

$$g_f = \sqrt{2}\frac{m_f}{v} \tag{2.27}$$

2.7 Concluding Remarks

This concludes my brief tour of the standard model as we currently understand it. In this chapter, I only touched on vital concepts for the rest of this thesis, while the landscape of the standard model is far more vast. For instance, I avoided discussing the intricate field of quantum chromodynamics; instead, I elaborate on the relevant concepts in quantum chromodynamics in the coming chapters when necessary. While the standard model has been a great success in various aspects of high-energy physics, efforts to understand new physics hinges upon migrating beyond standard model boundaries. In the next chapter, I take us to the edges of the standard model and discuss the main focus of this thesis: The mystery of Dark Matter!

Chapter 3

Beyond Standard Model

Scientific endeavor, according to my humble understanding, is Sisyphean struggle [47], if Sisyphus was rewarded with results and practicability. The SM of particle physics, although a successful theory in answering many of our questions, like any other successful theory in physics, opens up a vast field of questions, answers to which lies outside the said theory. With the SM before us, there are phenomena not explained (e.g., quantum gravity [48], dark matter [49, 50], dark energy [51, 52], neutrino masses [53], matter-antimatter asymmetry [54]), experimental results unexplained (e.g., the anomalous magnetic dipole moment of muon [33], B-meson decay anomalies [55]), theoretical predictions not observed (e.g., Koide formula [56]), theoretical problems unanswered (e.g., hierarchy problem [57] , number of parameters, Quantum triviality [58], Strong CP problem [59]). In this chapter, I discuss the theoretical basis of my analysis from the experimental perspective.

3.1 Dark Matter

In the present dissertation, the matter at hand is *Dark Matter* (DM). The terminology was coined by Jacobus Kapteyn in 1922 [60], in his study of the dynamics that rule our Galaxy. DM was defined as the type of matter that would interact gravitationally but would not reveal itself by any electromagnetic radiation. He tried to estimate the matter density near the Sun, including the density due to all stars near the galactic plane. Ultimately, however, Kapteyn concluded that the calculated density suffices to explain the vertical motions of stars near the plane of the Galaxy. The first claim of DM, however, came from Jan Oort in 1932, wherein he claims there is twice as much DM as visible near galactic plane [61]. His claim was also discredited later. The first observational evidence for the existence of DM is often attributed to Fritz Zwicky in 1933, a year later [62]. He correctly concluded that the measurements on the galaxy redshifts in the Coma cluster indicate the galaxy velocities are much larger than their expected escape velocity (from the cluster) if they are only being affected by visible matter.

Manifestly, in the present day, the evidence for the existence of DM is not limited to a single observation in 1932. Today we have ample observational evidence including but not limited to: discrepancy between expected and measured galactic rotation curve of Messier 33 [63], the mass-to-light ratios obtained by gravitational lensing (produces a map of DM around galaxy cluster) [64], acoustic peaks of small temperature anisotropies within the cosmic microwave background (CMB) [65].

3.2 Theoretical Candidates

By the 1970's DM as a concept became a force that had to be reckoned with by the scientific community. Most objections against the DM were discredited. Primarily

an excess over the expected visible matter both in the galaxies and clusters made it impossible to dismiss the idea of DM. With this concept came many questions that we are dealing with today. What is the nature of DM? What role has it played in the conception of the universe. What implications does its existence have on the rest of physics as we understand it today and many more questions yet to be answered [66]? Below is a theoretical review of some of the proposed ideas on the nature of DM.

3.2.1 Baryonic Dark Matter

Baryons are a subset of hadrons, with the other subset being the mesons. Baryons hold three or an odd number of quarks, whereas the mesons hold two. Protons and neutrons are baryons. Simply put, Baryonic DM is a type of DM that consists of baryons. Undetectable by its radiations, baryonic DM interacts with visible matter gravitationally. Supposed Massive Compact Halo Objects (MACHOs), dim brown dwarfs, white dwarfs, and neutron stars are often cited as candidates for Baryonic DM. If there is a baryonic-type DM, astronomical evidence suggests that it only accounts for a small portion of the total amount of DM in the universe. Models of Big Bang nucleosynthesis and observation of the CMB assert that if there are more baryons, including DM Baryons, then there should also be more helium, lithium, and heavier elements synthesized during the Big Bang. This model currently lacks popularity among the scientific community as observational data pour in and suggests that if baryonic DM was *the* candidate, it would have had to play a more dominant role in the dynamics of the Milky Way. That does not seem to be the case [67, 68, 69].

3.2.2 Nonbaryonic Dark Matter

The baryonic DM proposal is problematic on several fronts, and these problems were recognized by theorists and astronomers as early as the 1970s. What can be considered a breakthrough for non-baryonic DM popularity comes from the study of CMB. In the year 1964, CMB radiation was discovered [70], confirming the Big Bang theory. In the Big Bang theory, the *recombination* stage is referred to as the stage wherein the previously ionized gas from the primordial soup is cooled down and neutralized. From this stage onward, slight over-densities in the distribution of matter cause the attraction of even more matter via gravitation, causing a fluctuation in density. The Big Bang theory states that for the matter to have enough time to build up to all current observable structures such as Galaxies and clusters, the amplitude for density fluctuations at the recombination epoch must be of the order of 10^{-3} of total density. It also states that temperature fluctuates with the density and is of the same order. Naturally, astronomers and phenomenologists turn their attention to temperature fluctuations in the CMB. As the technology grew, it became exceedingly clear that the observation betrays the theoretical calculation by orders of magnitude. What could solve this mystery? In the late 1970S and early 1980s, astronomers argued that perturbations in a medium where heavy non-baryonic particles dominate could start much earlier than baryonic particles, providing enough time for the perturbation to grow in amplitude before the recombination state —large enough for structure formation that we can observe in the universe [66, 71, 72]. Hypothetical candidates for non-baryonic DM include: weakly interacting massive particles (WIMPs) [73] and hidden sector particles such as axions [74], sterile neutrino [75], and the dark photon [49, 76].

3.3 Experimental Search

The DM story is a journey that touches all sectors of physics. The observational evidence of dark matter kicked off a century of theoretical, experimental, and phenomenological endeavors that have occupied physics. The experimental struggle for detecting DM, which the present dissertation is a part of, can be classified into three different fronts: direct detection experiments, indirect detection, and collider searches.

3.3.1 Direct searches

Assuming that the DM particles are showering the Earth from the halo of the Milky Way [77], direct searches for DM aim to observe scattering between the DM and nuclei in a detector on Earth. Direct searches, typically in low-energy recoils (∼a few keVs), happen in deep underground laboratories. These experiments are usually situated underground to shield the experiment from the cosmic ray showers, hence reducing the background of the experiment. The signal for many of these experiments comes from scintillation light (photons), energy deposition (phonons) in bolometers, and ionization (electrons). Some examples of these experiment locations include the Stawell mine [78], the Soudan mine [79], the SNOLAB underground laboratory at Sudbury [80], the Gran Sasso National Laboratory [81], the Canfranc Underground Laboratory [82], and the China Jinping Underground Laboratory [83]. With no detection from the direct experiments so far, tight upper limits have been set on scattering cross-section between the target nuclei and the possible candidates for DM, such as WIMPs [84].

3.3.2 Indirect Searches

The search for DM can not only be pursued directly (scattering off of nuclei) but also indirectly. In the early evolution of the universe, there was a moment called the

freeze-out moment. This is when the temperature and density of the particles were low enough that DM particles can no longer find each other and annihilation ceased [85]. Indirect detection experiments search for the products of self-annihilation, i.e., the DM–anti-DM encounter, in outer space. The self-annihilation of DM can produce a flux of $\gamma-$rays, neutrinos, or anti-matter that can peak above the expected background in spectroscopy experiments. Alternatively, if DM is unstable, it can decay into SM products which can be looked for in our galaxy or other galaxies [86]. The Energetic Gamma Ray Experiment Telescope [87], the Fermi Gamma-ray Space Telescope [88], and the Alpha Magnetic Spectrometer on the International Space Station [89] are some examples of indirect searches for DM.

3.3.3 Collider Searches

An alternative approach to searching for DM is to produce it in the laboratory, which is the focus of this dissertation [90]. The Large Hadron Collider (LHC) is the prime machine that can be used for this purpose. There are several ways to approach detection within LHC, including:

Model-dependent searches: A more traditional analysis where a DM model is defined, a selection for the analysis is defined, and the SM background is fought. The background can be either estimated using MonteCarlo simulation or using actual data, determining a signal and control region, and extrapolating the background from the control region to the signal region. Given that the DM particle in question has survived the evolution of the universe, it is safe to assume that the DM particle is stable and relatively heavy.

Effective field theory approach (model-independent): Effective Field Theory

(EFTs) is a QFT, only relevant within a particular energy scale. The Lagrangian describing the theory contains only the terms that operate within that energy scale. The EFT approach has been used for Run I analyses at the LHC [91]. The EFT approach also has the cross-examination advantage where the results of direct and indirect experiments can be compared to each other.

Simplified models (model-independent): However useful, the EFT approach comes with an energy range limitation. The theory falls apart when the energies exceed the EFT cut-off energy scale. For direct and indirect experiments, where the probed energy does not exceed the TeV scale, and for Run I of the LHC, that does not pose a threat. For the higher energies, however, such as more contemporary LHC runs, probed energies are larger than the values of scale cut-off that can be excluded within the EFT framework. Here is where the simplified models can enter. In recent years the LHC community has engaged more and more with the simplified models [92, 93, 94, 95, 96]. These models are referred to as simplified because they are simply independent of Ultra-Violet (UV) extensions of the SM [97]. In simplified models, the mediator fields are generally classified into four groups: scalar fields, pseudo-scalar fields, vector fields, and axial-vector fields. The analysis parameters are relatively easier to track (in contrast with Ultraviolet-Complete SM): the messenger mass and width, the dark-matter particle mass, and the interaction couplings of one or two effective fields.

3.4 The Dark Sector

Candidates such as axions and dark photons discussed in Sec. 3.2.2 belong to a more recent and exciting theoretical and experimental exploration referred to as the Dark

Sector or the Hidden Sector, discussed in recent literature [98, 99, 100, 101, 102].

The conventional approach in the search for new particles, including DM particles, has been to consider them to be charged under at least some SM gauge symmetries. While this approach has been the basis of 50 years of theoretical and experimental development in particle physics, the experimental results so far do not corroborate this assumption. To overcome these underwhelming results, attention has increasingly turned toward models wherein new particles are not charged under SM gauge symmetries. Collectively, these models are referred to as the *dark sector* or hidden-sector models. Under this assumption, if the DM does not seemingly interact with the SM sector, the implication is that it resides in a *dark sector* of its own. In other words, if DM is not charged under SM symmetry groups, it might be charged under a dark (hidden) symmetry group. Depending on the model, this sector may contain dark bosons, dark fermions, and dark scalars. The dark states are assumed to interact through Yukawa couplings, mediator bosons, or both. To stay hopeful for the detection of DM at the LHC, the dark sector would have to go beyond the gravitational interaction with the SM and communicate with the SM through a weak *portal* — as the current terminology has it.

The portal may assume different forms based on the spin of the portal operator: spin-1 Vector, spin-1/2 Neutrinos, spin-0 Higgs (scalar), or Axions (pseudo-scalar). The focus of this dissertation is on the spin-1 Vector portal where a dark gauge boson interacts with an SM gauge boson through *kinetic mixing* between one dark and one visible Abelian gauge boson. This gauge boson is called the dark photon (or the dark Z).

A natural first attempt to glue the idea of dark matter to the standard theoretical framework of SM physics is to add generic candidate gauge symmetries to the known SM gauge group. Let us call this generic candidate G_D, and enlarge the SM gauge

group to $SU(3) \otimes SU(2)_L \otimes U(1)_Y \otimes G_D$. The first three terms are the familiar SM gauge groups for strong, weak, and electromagnetic interactions. If we want to stay hopeful and imagine dark matter can be observed through some interaction with SM particles, we should consider the possibility of the G_D *mixing* with at least one of the SM gauge bosons. The concept of mixing can result in the coupling of SM particles to dark matter particles through a *dark portal* — which essentially would be a *mixed* boson.

A minimal realization of this idea is the addition of a second *electromagnetism-like* Abelian $U(1)$ gauge symmetry to the SM, i.e. $G_D = U(1)_{Y_D}$. This introduces a *dark hypercharge*, Y_D, (analogous to Eq. 2.8) carried by dark matter particles and a corresponding dark vector gauge boson that mediates their interaction.

The dark sector reaps the benefits of simplified models: We can compare results across experiments and a more restricted number of parameters—just two in the case of the dark vector mediator (dark photon or dark Z), with the two parameters being the mass and the mixing parameter.

Adding the new gauge group to the SM generates a kinetic mixing term,

$$\mathcal{L}_{mix} = -\frac{1}{4}B^{\mu\nu}B_{\mu\nu} - \frac{1}{4}B'^{\,\mu\nu}B'_{\mu\nu} - \frac{\varepsilon}{2}B^{\mu\nu}B'_{\mu\nu}, \tag{3.1}$$

where $B^{\mu\nu}$ is the SM electromagnetic field tensor defined in Eq. 2.4, and $B'_{\mu\nu}$ is its counterpart in the dark sector. Here ε is called the kinetic mixing parameter, and it signifies the strength of mixing the electromagnetic field to the dark sector.

Before symmetry breaking, the field redefinition for the gauge sector is:

$$\begin{pmatrix} B'_\mu \\ B_\mu \end{pmatrix} = \begin{pmatrix} \sqrt{1-\chi^2} & 0 \\ -\chi & 1 \end{pmatrix} \begin{pmatrix} \hat{B}'_\mu \\ \hat{B}_\mu \end{pmatrix}, \tag{3.2}$$

where $\chi = -\frac{\varepsilon^2}{\cos^2\theta}$ and θ is the Weinberg mixing angle. The kinetic terms of Eq. 3.1 are diagonalized by the redefined fields. After electroweak symmetry breaking, the physical gauge fields $U(1) \otimes U(1)_D$ will be A^μ, Z^μ and ${Z_D}^\mu$ respectively, the photon, the Z boson, and the dark Z. The photon stays massless, while Z and Z_D mix. The mass eigenstates are given as follows [103]:

$$\begin{pmatrix} Z \\ Z_D \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} Z_0 \\ Z_{D,0} \end{pmatrix}, \tag{3.3}$$

where α is the mixing angle; the subscript 0 indicates that the gauge fields are not mass eigenstates yet. A more detailed discussion can be found in App. A. Lastly, a summary of 95% exclusion limits on the dark vector mediator, set by multiple experiments, is shown in Fig. 3.1.

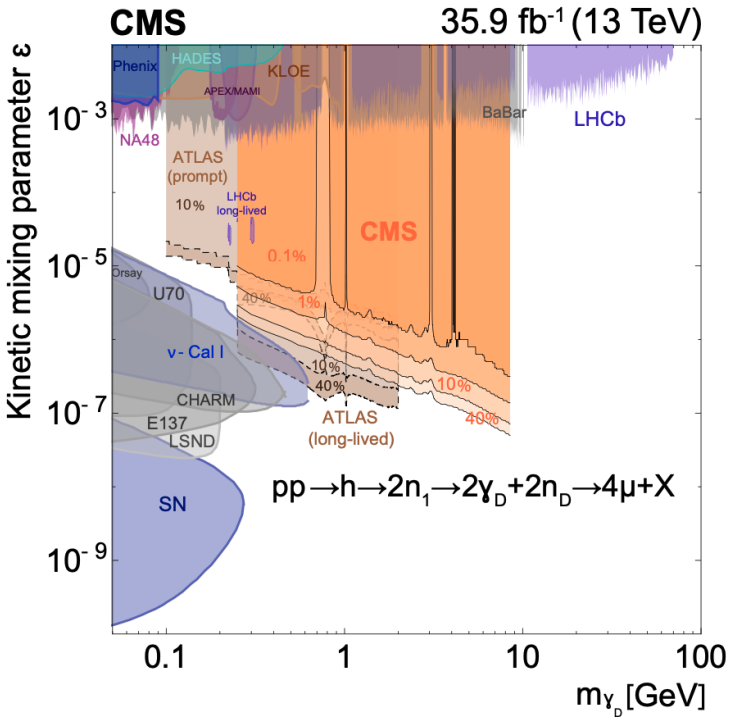


Figure 3.1: The 90% CL upper limits (black solid curves) on the dark vector mediator in the plane of parameters (m_{γ/Z_D} and ε) [5] are shown. The limits shown in light orange correspond to dataset recorded by CMS during the 2016 era.

3.5 Model Independent Search In this Thesis

In a fruitful collaboration with analysis communities from Texas A&M, Rice University, and the University of Sonora, I explored the pair production of new bosons at the LHC. As discussed in Sec. 3.3.3, model-independent searches benefit from cross-experiment or cross-modeling interpretation. In other words, the results of this search can be interpreted in the context of any model with the same final state. Our efforts substantially expanded the parameter space compared to the parent analysis for this search [5].

The precursor to our analysis presented a search for new light bosons decaying into muon pairs, corresponding to an integrated luminosity of 35.9 fb^{-1} at the center-of-mass energy $\sqrt{s} = 13$ TeV. The parameter space probed there was $0.25 < m_{Z_D} < 8.5$ GeV. Our analysis extents this parameter space to $0.25 < m_{Z_D} < 60$ GeV. The dataset corresponds to 59.7 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV recorded during 2018 at CMS.

The Feynman diagram of Fig. 3.2 illustrates the general idea behind this model-independent search, wherein proton-proton collisions at CMS produce two new bosons (a). These bosons then subsequently decay into a pair of oppositely charged muons, commonly referred to as di-muons. The grey circle in Fig. 3.2 can be interpreted as an indication of the model independence of the search: the results are decoupled from what happens in the "mystery box," as long as the final state involves two di-muons. The intermediate *portal*, illustrated by the encompassing grey circle, could be mediated by either a vector-portal or an SM-like Higgs scalar-portal that can further produce either new fermionic or scalar products, which finally can decay to the leptonic final state that can be detected at the LHC.

The discovery of such signatures can be considered a strong indication of dark

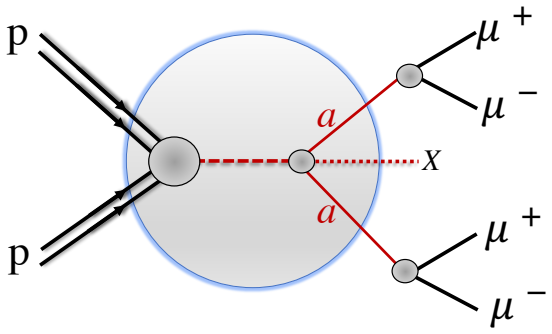


Figure 3.2: Schematic example of the proton-proton interaction that produces a pair of new bosons of which each decays into a muon pair is depicted. The dark sector interactions are indicated by the grey circle. The X particle is to signify any excess processes other than the four lepton final state.

matter particles. I use the term *indication* here, as what decays into muons is not stable. The DM particle proper, that accounts for the extra mass in the universe, has to be stable; otherwise, it has decayed to other products by now in the history of universe. Despite this fact, we continue to consider this search a DM search as the mediator of this interaction can be a portal to the dark sector and hence an indication of the existence of DM particle.

The model-independent approach to data analysis is not only beneficial, and as we discussed in Sec. 3.3.2, it proves to be necessary with growing data volume due to higher LHC luminosity. Our primary aim is to capture any discrepancy between the data and the SM prediction, regardless of the theoretical models. Previous model-independent searches stretch back to 2009 at the Tevatron [104], and more recent LHC searches can be found at [105, 106, 107].

We eventually do interpret the independent model results of this search in the context of the specific benchmark models. The advantage of using multiple benchmark models is verifying the independence of model parameters from the results. Four representative benchmark models were used in this overall analysis, performed by the working group. The Higgs (SM or non-SM) portal models include: the axion-like

particle (ALP) model and a supersymmetric model with additional dark sectors (i.e., dark SUSY), including cases with long-lived new particles and the next-to-minimal supersymmetric standard model (NMSSM). My contribution was to introduce a vector-portal model with dark scalars (s_{D}). I discuss this model in detail in the following section. For more information on the Higgs portal models see our the CMS analysis note AN-19-153 [108].

3.5.1 Benchmark Model: A Vector-Portal To the Dark Sector

As a part of data analysis effort, I explored the possibility of interpreting the model-independent results in the context of a non-Higgs portal model, where the dark matter particle could be produced through a spin-1 mediator, e.g., a Z-like dark boson, Z_{D} via kinetic mixing mechanism as described in Sec. 3.4. Further description of such models can be found in Refs. [109, 110, 111, 112, 113, 114]. The vector-portal model scans parameter space of mass, $5 < m < 60$ GeV/c^2, for prompt signatures, using events with four muons final state at CMS.

More specifically, my studies are on a vector-portal model, a simplified dark matter model, which includes dark scalar particles, s_{D}, that are produced through a vector portal and couple directly to SM particles, e.g. to muons: $pp \to Z_{\mathrm{D}} \to s_{\mathrm{D}}\overline{s_{\mathrm{D}}} \to \mu^+\mu^-\mu^+\mu^-$. In Sec. 2.6 I discussed the minimal Higgs model wherein, SU(2) $\otimes$ U(1) electroweak theory Higgs sector is merely comprised of one complex Higgs doublet. There is only one physical neutral Higgs scalar in the spectrum with this assumption. This minimal realization can be extended to include a two-Higgs-doublet and give rise to a charged Higgs sector and a neutral Higgs sector. In this case, a vertex with a Z boson can only exist with couplings with non-identical Higgses. This condition is mandated by bose symmetry, which dictates CP invariance for the vertex. The Bose symmetry states that under an exchange of identical particles the system must stay invariant. In other

words, the coupling of the Z boson to Higgs particles, or in general any two scalars, is allowed only when they have opposite CP quantum numbers [115]. To map the above principles to the dark sector, I require the dark scalar to be a complex scalar field that is not self-conjugate, i.e. the s_{D} and the $\overline{s_{\mathrm{D}}}$ are not the same particle (unlike the SM Higgs or photon), effectively differentiating the scalar fields and adhering to the opposite CP requirements. For the purpose of simplicity the branching fraction $\mathcal{B}$ of s_{D} to muons is considered to be 100%. The Feynman diagram for this process is shown in Fig. 3.3.

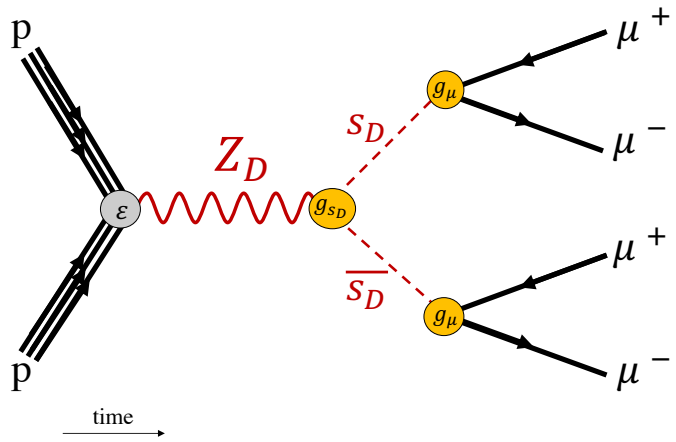


Figure 3.3: Z_{D} decays into a pair of scalar dark matter particles which then each subsequently decay into two oppositely charged muons.

The total Lagrangian is the sum $\mathcal{L}_{SM} + \mathcal{L}_{kinmix} + \mathcal{L}_{DM}$. The interaction Lagrangian for the above Feynman diagram is given by:

$$\mathcal{L}_{total} = \mathcal{L}_{SM} + \mathcal{L}_{mix} + i\frac{g_{s_{\mathrm{D}}}}{2}(\partial_\mu s_{\mathrm{D}}\partial^\mu \overline{s_{\mathrm{D}}})Z_{\mathrm{D}} + g_\mu(\overline{\mu}\mu)s_{\mathrm{D}} + g_\mu(\overline{\mu}\mu)\overline{s_{\mathrm{D}}} \tag{3.4}$$

As suggested in Eq. 3.4, the vector portal of the model, Z_{D}, couples to SM fermions through kinetic mixing with SM Z, and directly to the s_{D}. Typical branching fractions of the Z_{D} are studied using the MadWidth tool within the MadGraph infrastructure

and shown in Tab. 3.1.

Table 3.1: Branching fractions and widths of Z_{D} boson in the vector-portal model for $M_{Z_D} = 125$ GeV and $\varepsilon = 10^{-2}$

Z_D decays			
BR%	**Particle#1**	**Particle#2**	**Partial width (GeV)**
98.40	s_D	$\overline{s_D}$	7.25×10^{-2}
0.23	u	$\bar{u}$	1.70×10^{-4}
0.23	c	$\bar{c}$	1.70×10^{-4}
0.21	t	$\bar{t}$	1.52×10^{-4}
0.20	e^-	e^+	1.50×10^{-4}
0.20	μ^-	μ^+	1.50×10^{-4}
0.20	τ^-	τ^+	1.50×10^{-4}
0.07	d	$\bar{d}$	5.00×10^{-5}
0.07	s	$\bar{s}$	5.00×10^{-5}
0.07	b	$\bar{b}$	5.00×10^{-5}
0.04	ν_τ	$\overline{\nu_\tau}$	3.00×10^{-5}
0.04	ν_μ	$\overline{\nu_\mu}$	3.00×10^{-5}
0.04	ν_e	$\overline{\nu_e}$	3.00×10^{-5}
Total BR = 100%		Total width = 2.97×10^3(GeV)	

I perform scans over relevant parameters such as production cross-section, branching fraction, and kinematic and geometrical acceptance for different s_{D} masses for constant Z_{D} mass. Next, I repeat these scans for varying masses of Z_{D}. The product of the above parameters indicates the sensitivities of our search within this model. Fig. 3.4 illustrates the cross-sections (σ), that I obtained with MadGraph [116], for various masses of Z_{D}. It should be noted that the production cross-section remains constant as I scan over various masses of s_{D} as long as the Z_{D} mass remains unchanged. Furthermore, I derive scans over branching fractions from the MadWidth algorithm, as shown in Fig. 3.5. The product $\sigma(\mathrm{pp} \to \mathrm{Z_D}) \times \mathcal{B}(\mathrm{Z_D} \to \mathrm{s_D}\overline{\mathrm{s_D}})$ is shown in Fig. 3.6. I obtain the kinematic and geometrical acceptance of the muon selection, used in this

analysis for various masses of the $Z_{\rm D}$ and $s_{\rm D}$ particle, from the simulation of the final 4-muon state at the generator level using MadAnalysis, as shown in Fig. 3.7. I use cuts on the transverse momentum $p_{\rm T}$ and pseudo-rapidity η. Specifically, I require each muon to have $p_{\rm T} > 8$ GeV and $|\eta| < 2.4$. In additional, at least two muons need to be high-$p_{\rm T}$ and more central with $p_{\rm T} > 24$ GeV and $|\eta| < 2$. The full analysis selection is discussed at length in Ch. 6. Lastly, I perform a scan over the product of the three above parameters as an indicator of the sensitivity, as illustrated in Fig. 3.8. Here I set both $g_{s_{\rm D}}$ and g_μ to be 0.25. I focus on the region where the $s_{\rm D}$ signal mass is below 60 GeV, and $Z_{\rm D}$ masses are within 85−200 GeV.

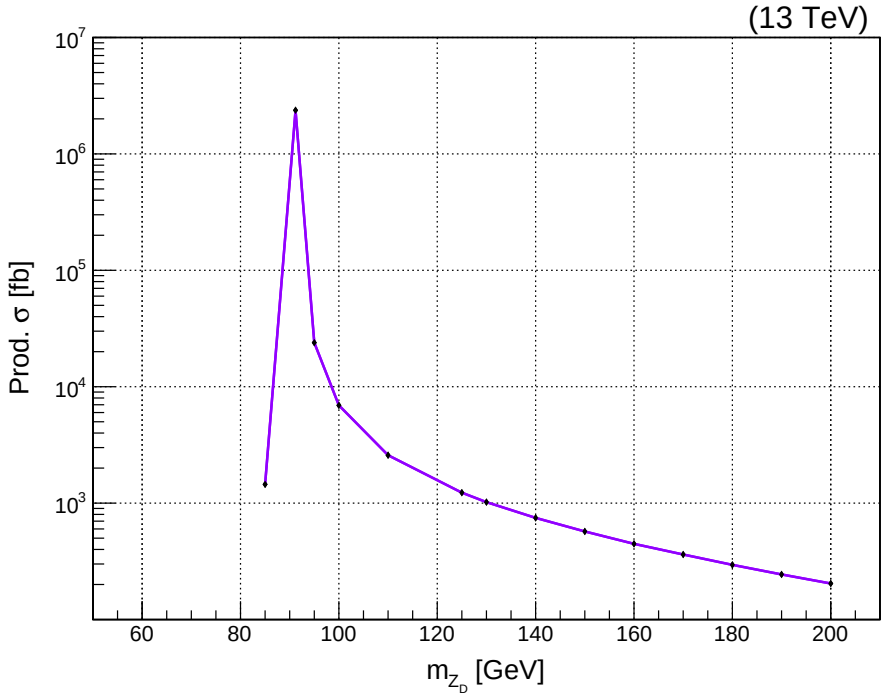


Figure 3.4: A scan of the production cross section for various masses of the $Z_{\rm D}$ particle. I choose the KM parameter to be $\varepsilon = 10^{-2}$ here. The values are calculated by MadGraph.

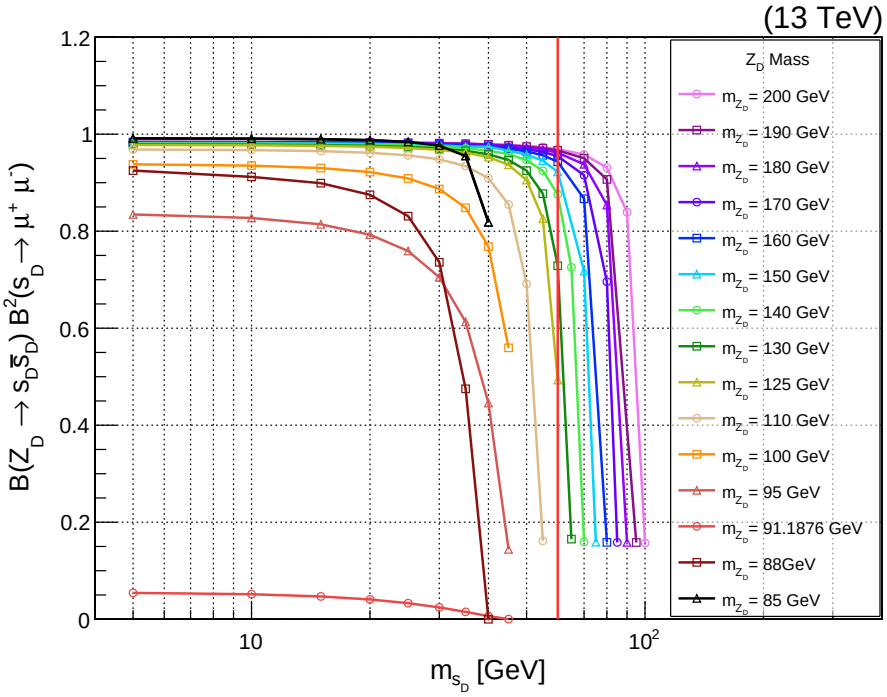


Figure 3.5: A scan of branching fraction over various masses of s_{D} and for various Z_{D} masses. I choose the KM parameter to be $\varepsilon = 10^{-2}$ here. The vertical red line shows the 60 GeV point for the s_{D} mass. The values are calculated by MadGraph.

3.6 Concluding Remarks

Once we step outside the standard model boundaries, we immediately find ourselves in an expansive realm, enriched with numerous possible theoretical frameworks, attempting to describe what has not been yet discovered. In this chapter, after introducing the dark matter problem and the high-energy physics angle at tackling the problem, I argued for conducting the search in a *model-independent* fashion. The model-independent method liberates us from the internal limits of individual theoretical paradigms and opens up a space for an unbiased search. I furthermore introduced a simplified dark matter model, the vector-portal model, with which I will interpret the model-independent results in Ch. 8.

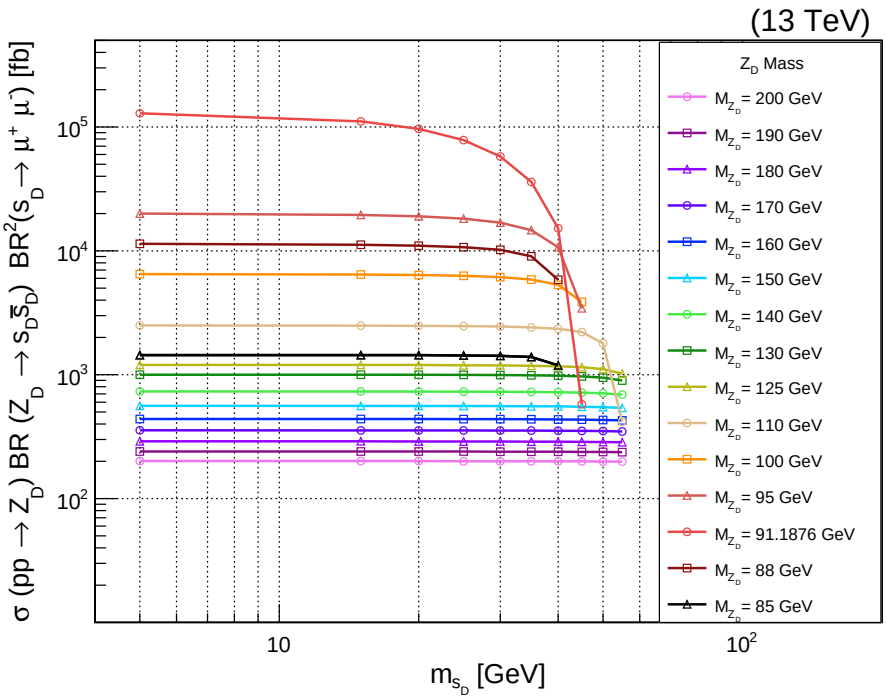


Figure 3.6: The $\sigma(\mathrm{pp} \to \mathrm{Z_D}) \times \mathcal{B}(\mathrm{Z_D} \to \mathrm{s_D}\overline{\mathrm{s_D}})$ for various masses of Z_D and s_D, calculated from MadGraph MC generator.

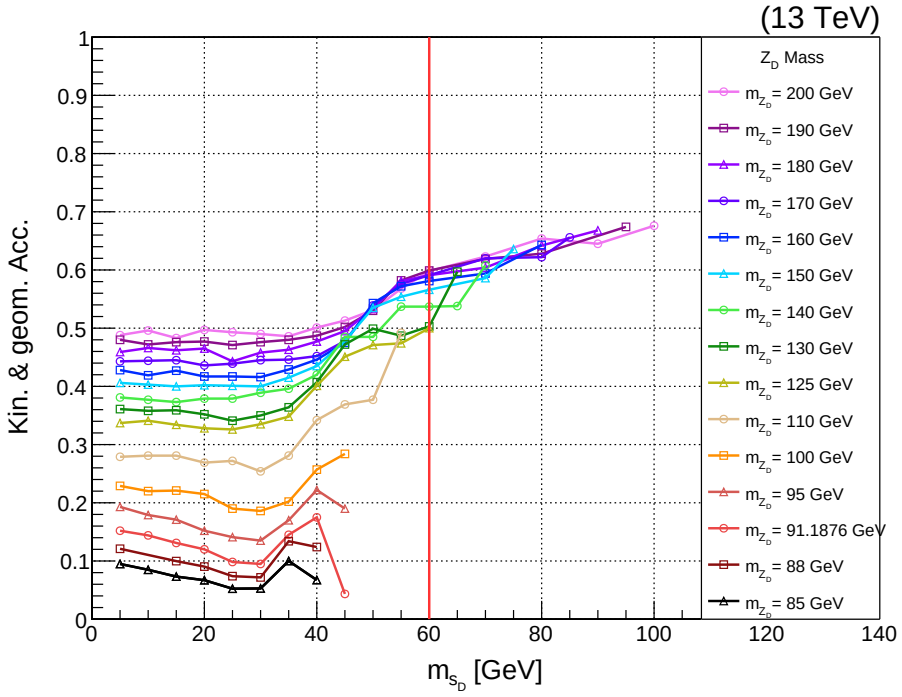


Figure 3.7: A scan of geometrical and kinematic acceptance of the muon selection used in this analysis over various masses of s_{D} and for various Z_{D} masses. I choose the KM parameter to be $\varepsilon = 10^{-2}$ here. The values are calculated by MadAnalysis.

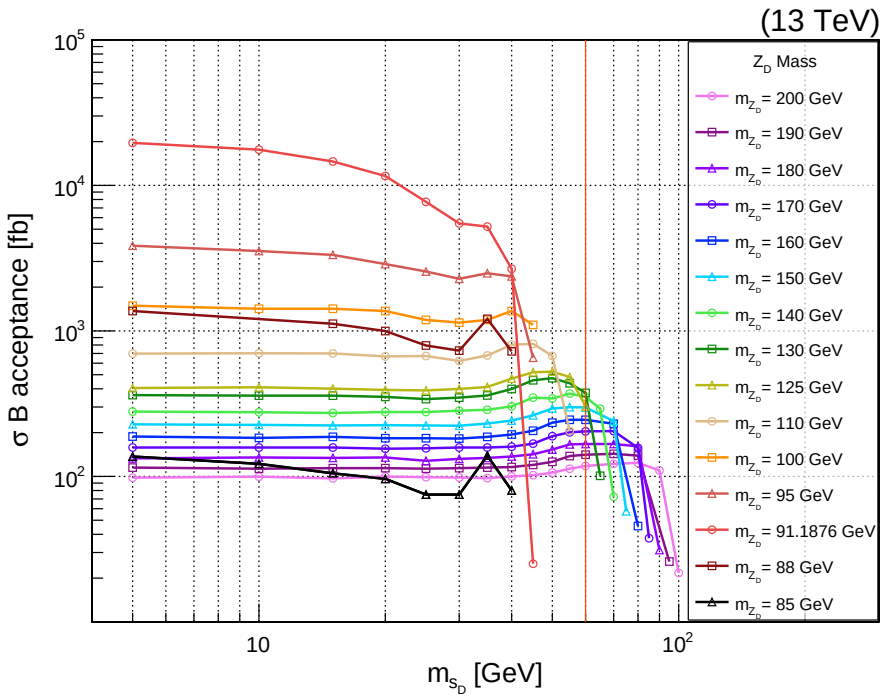


Figure 3.8: A scan of $\sigma \times \mathcal{B} \times$ acceptance, as an indication of sensitivity is shown over various masses of s_{D} and for various Z_{D} masses. I choose the KM parameter to be $\varepsilon = 10^{-2}$ here. Both $g_{s_{\mathrm{D}}}$ and g_μ are set to 0.25. The vertical red line shows the 60 GeV point for the s_{D} mass. The values are calculated by MadGraph.

Chapter 4

The Experimental Apparatus

Particle physicists must inquire about the most fundamental laws of physics and illuminate physics not yet explored. This requires probing of the smallest scales, which translates to high-energy physics. The known physics is the physics of already reached energies. New physics lies beyond the energies that have been achieved. This necessity justifies the century-old race toward higher energies by particle physicists. The accelerators are the primary tool that enables and aids us toward the higher energies. A particle accelerator utilizes the electromagnetic force to thrust charged particles to the near speed of light velocities and high energies at the scale of GeV and, more recently, TeV. The electromagnetic fields are used to guide and focus particles in two well-defined *beams*. In this chapter, I offer a concise introduction of the largest particle accelerator ever built and the detector used in data-collecting in this thesis.

4.1 The LHC Accelerator

The European Organization for Nuclear Research (Conseil Européen pour la Recherche Nucléaire in French), known as CERN, is the largest particle physics organization

globally. The organization was established in 1952 by 12 countries in western Europe and, in its early years, was operated by the University of Copenhagen under the direct supervision of Niels Bohr. Today, the organization is based in a northwest suburb of Geneva on the Franco-Swiss border and has 23 member states. The United States have been an observer state since 1997. CERN provides infrastructure and organizational planning for high-energy physics research, such as particle accelerators and laboratories. CERN is the host site for the Large Hadron Collider (LHC), the world's largest and most powerful particle accelerator ever built.

The LHC story began in 1976 when the high-energy physics community opened the discussion to build a Large Electron-Positron (LEP) collider at CERN. The LEP construction was completed in 1988 and installed in a 27 km tunnel on the Franco-Swiss border. While the LEP was operational, the CERN council approved the LHC proposal in 1994. The LEP was operational from 1989 to 2000 and its tunnel was re-purposed in 2001 to host the LHC [117]. The LHC is a crucial instrument in the high-energy physics endeavor. It provides energy high enough for completing the SM picture such as the Higgs mechanism and new physics such as supersymmetry and dark matter.

The principal idea behind an accelerator such as the LHC is utilizing the electromagnetic field to accelerate and steer particles. At the LHC, proton beams are injected into Radiofrequency (RF) cavities and boosted to the desired energies. Groups of protons are then sorted into packs of protons called "bunches." Once boosted, the beams are guided and focused by superconducting electromagnets at the near speed of light. The superconducting electromagnets are operational at low temperatures (-271.3°C) provided by a distribution system of superfluid liquid helium. Lastly, the beams are once again re-focused and *squeezed* right before the collisions to maximize the probability of head-on collisions. A proton in a bunch circumferences the LHC in about

90 microseconds (μs) (orbit frequency of ∼11 kHz and bunch crossing frequency of ∼40 MHz). There is a 25 ns time interval between the bunches called bunch-crossing (BX) [118].

The LHC is a synchrotron-type accelerator. In the synchrotron, the magnetic field strength is varied with time. Synchrotrons, unlike cyclotrons, are unable to accelerate particles from zero kinetic energies. Therefore the injected particles have to be pre-accelerated. Pre-acceleration is done in three steps: initially, the proton beam is boosted to 50 MeV in a linear accelerator (LINAC). The protons are then fed into the proton synchrotron booster (PSB) and accelerated to 1.4 GeV. The beam is further accelerated to 26 GeV by Proton Synchrotron (PS). The final boost is given by Super Proton Synchrotron (SPS), which increases the beam energy to 400 GeV and is sent to the LHC. A schematic representation of the CERN accelerator complex is shown in Fig. 4.1.

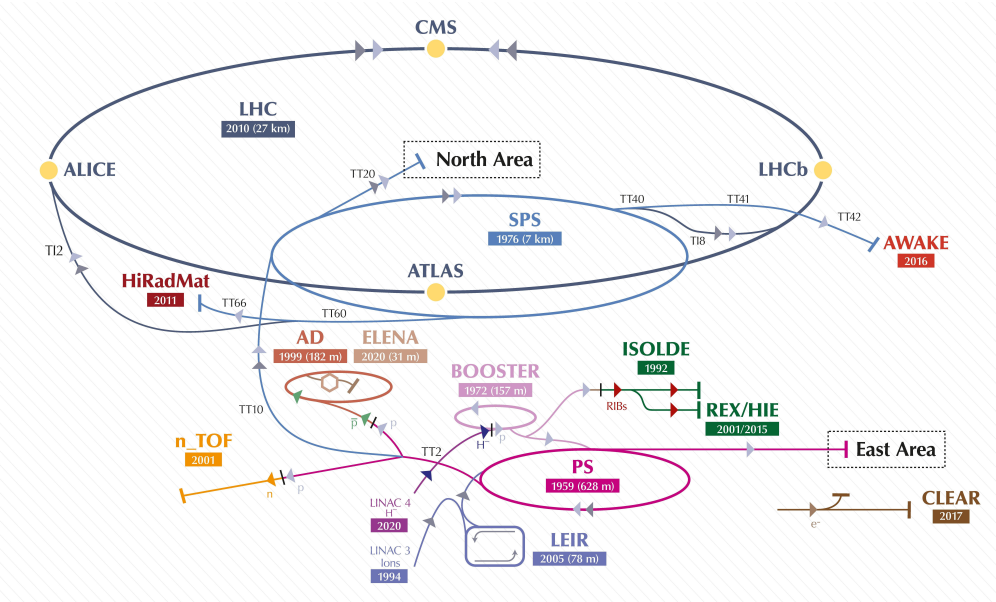


Figure 4.1: The CERN accelerator complex (image: CERN)

The designed center-of-mass of the LHC is 14 TeV and the designed luminosity is $10^{34}\text{cm}^{-2}\text{s}^{2}$. There are four major points where the two beams intersect. To each point,

there is a corresponding experiment: Compact Muon Solenoid (CMS), A Toroidal LHC Apparatus (ATLAS), A Large Ion Collider Experiment (ALICE), and Large Hadron Collider beauty (LHCb). Other experiments at the LHC include TOTEM, MoEDAL, LHCf, and FASER. CMS and ATLAS are the main general-purpose detectors, while the rest are designed for specific research [119, 120, 121, 122, 123, 124]. In this thesis, I have used the data collected by CMS in 2018 and 2017.

4.2 The CMS Detector

The CMS technical design was proposed in 1994, [125] but the concept of a compact detector for LHC based on a solenoid, the Compact Muon Solenoid (CMS) detector, was presented in October 1990 at the LHC workshop in Aachen [126]. This massive 14000-tonne machine started in 1998 above the ground, and the completed detector was lowered into the cavern in November 2006, piece by piece. Four years after the first proton-proton beam traveled the full 4-km radius circle of the LHC, the evidence for a particle at about 125 GeV was presented at a seminar and webcast was announced by the LHC scientists; later to be confirmed as the Higgs boson [127]. This global collaborative experiment is one of the largest scientific collaborations in the history of science, involving 5000 particle physicists, engineers, technicians, students, and support staff from 200 institutes in 50 countries. CMS consists of four major components: Magnet and Muon System, Tracking, and Calorimetry. Below Fig. 4.2 is a schematic overview of the CMS detector.

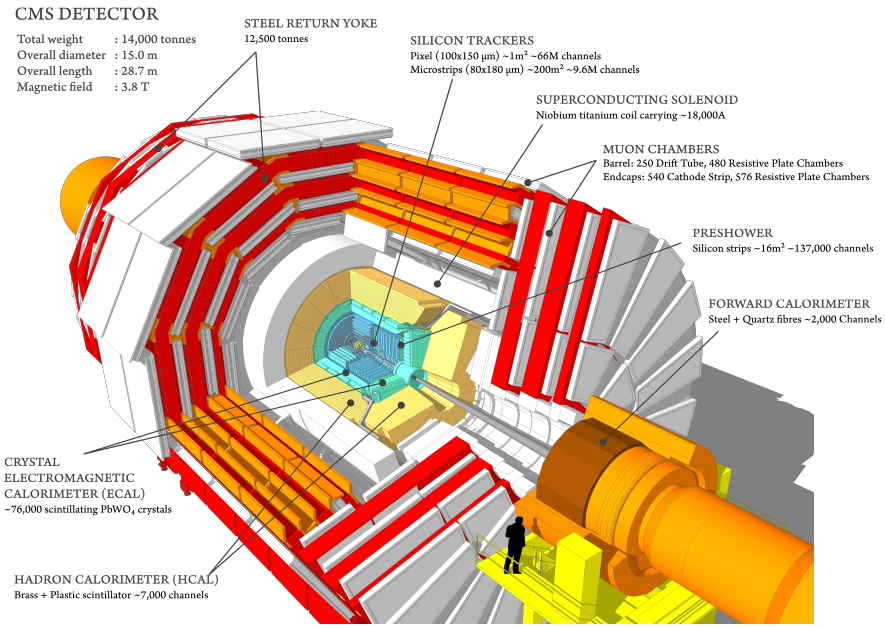


Figure 4.2: The components of the CMS detector [6]

4.3 Bending Particle Tracks

The central design idea upon which the CMS detector relies is choosing a solenoidal magnetic system for muon detection. A superconducting solenoid with 13 m of length and 5.9 m of radius produces a uniform magnetic field of up to 4 Tesla in the core and 2 Tesla outside the solenoid. The 1.8 m steel yokes, shown in red in Fig. 4.2, support the muon chambers and also guide the magnetic flux back into the coil. The magnetic field facilitates precise measurements of the momentum of the muon tracks. The magnetic field bends muon tracks in the transverse plane. The bending of the tracks is maximal inside the coil. In the return yoke, the magnetic field is in the opposite direction; as a result, the tracks are bent back. The bending and the direction of the tracks can be parametrized (sagitta) and used to solve the equation of motion for a relativistic charged particle in an electromagnetic field. The output will determine the momentum and the charge of the muons.

The bending angle at the point of interception with the chamber is the angle be-

tween the track tangent and a radius vector passing from the center through the point where the track passed the chamber. The *sin* of the bending angle ($\sin\theta_{bend}$) grows linearly with the radius inside of the coil; thus, the best momentum resolution is achieved by measuring the bending angle right after the coil.

$$\sin\theta_{bend} = \frac{a}{2R_{helix}} \tag{4.1}$$

Where a is the radius of the coil itself and R_{helix} is the radius of the charged particle trajectory (helix) in a magnetic field [128]. The helix radius is given by:

$$R_{helix} = \frac{p_t}{qB_z}, \tag{4.2}$$

where p_t is transverse momentum, q is electric charge, and B_z is the magnetic field in beam pipe direction. For high p_t, in Eq. 4.2, $R_{helix} >> a$ and as a result $\sin\theta_{bend} \approx \theta_{bend}$.

Now we can write p_t in terms of the bending angle:

$$p_t = \frac{aqB_z}{2\theta_{bend}} \tag{4.3}$$

4.4 Identifying Tracks

The tracking task in CMS is shared by two major systems: the inner trackers and the muon system. The inner tracker detectors, including the pixel and strip detectors, are the closest to the interaction region at CMS, which extends up to 24 cm away from the center of the beam pipe.

The pixel detector: The pixels are small silicon detectors, doped to create p-type and n-type semiconductors and patched together to form a p-n junction. An external

voltage is then applied to the junction to widen (100-300 μm) the depletion region. The silicon detector can detect a charged particle passing through the depletion zone by reading out the formed current from the electron-hole drift due to the electric field's depletion zone. The pixel detector, as the closest detector layer to the interaction point, is vital for p_t measurements, particle identification, and vertex reconstruction at CMS.

The pixel detector went through an upgrade at the end of 2016, adding a layer of pixels both in the barrel and the endcaps. This upgrade increased the total number of layers to 4 in the barrel pixel (BPIX) and increased the number of layers to 3 in the endcaps forward pixel (FPIX), with the innermost layer of the pixel tracker now sitting even closer to the collision point at a radius of 2.9 cm. This proximity to the interaction point required an upgrade to the beam pipe diameter. In the new configuration, the beam pip diameter could not exceed 45 mm. The original goal of the LHC aimed to reach instantaneous luminosity of 10^{34} cm^{-2}s^{-1} per bunch crossing. With the phase-1 upgrade of the accelerator, this luminosity is doubled. The upgrade of the pixel detector equips the tracker with the ability to buffer in high luminosity, minimizes the radiation degradation damage, and improves pattern recognition and track reconstruction. More on the pixel phase-1 upgrade can be found in Ref. [7]. A comparative layout for the upgrade is shown in Fig. 4.3.

The silicon strip detector: Behind the pixel detectors, the next tracker station that collision products pass through is ten layers of silicon strip detectors at the barrel, spanning a radius of 25-110 cm, and nine disks at each endcap. The silicon strip detectors are divided into four regions: the inner barrel part (TIB), the inner disks (TID), the outer barrel (TOB), and the outer endcaps (TEC). Much like the silicon pixels, the silicon strips take advantage of semiconducting properties of silicon. The *hits* from

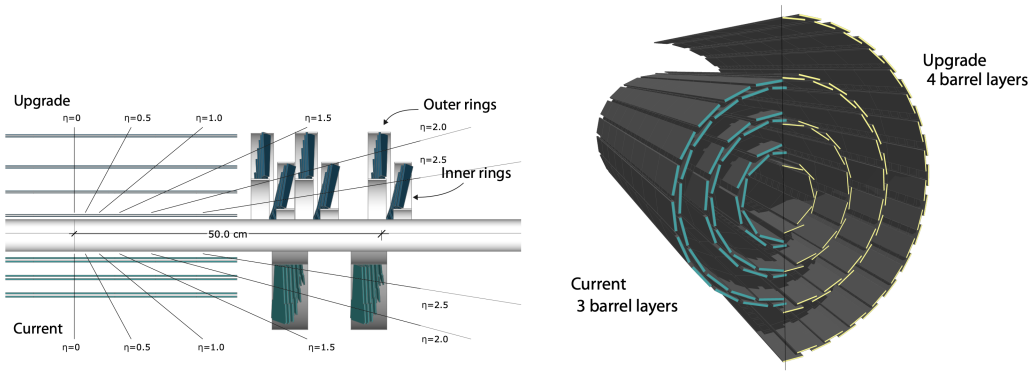


Figure 4.3: Left: comparative layout of the pixel detector between the layers and disks, before and after the upgrade of pixel detectors. Right: Transverse-oblique view comparing the pixel barrel layers in the upgraded detector versus pre-upgrade. [7]

the silicon strips are then read out and amplified by analog pipeline voltage (APV25) chips; a 128-channel analog pipeline ASIC. The silicon strips contribute to momentum and energy measurement for the charged tracks [129].

Muon detectors: The muon system covers the outermost section of the CMS detector. They also provide tracking and momentum measurements of the muons; however, pixel detectors were not a cost-effective choice because they cover a significantly larger volume. Instead, gaseous detectors were adopted for the muon system. A charged particle is tracked by fitting a curve to hits among the four muon stations, and combined with tracks from the tracker detector, so that the full path of the particle through the detector is revealed. A more comprehensive description of the muon system and its gas detectors is given in Sec. 4.6. An overview of track reconstruction methods for the muon system and the silicon tracker is given in Sec. 4.7.

4.5 Measuring Energy

In particle physics, calorimetry refers to measuring the energy deposition of a particle in a given medium. At CMS, calorimetry is done on electromagnetic particles such as electrons, positrons, photons, and hadronic particles such as protons, neutrons, pions, and kaons. The former types are detected and measured by the electromagnetic calorimeter (ECAL), and the latter case is done by the hadron calorimeter (HCAL).

The ECAL: The subdetector utilizes approximately 76000 lead tungstate ($PbWO_4$) crystals as scintillation materials for measuring the energy and position of electromagnetic particles. High-energy electromagnetic particles produce particle *showers* in the crystals upon smashing into them. Electromagnetic showering is a cyclic process wherein high energy photons (∼few MeV) interact with tungstate atoms through the pair production process. The emergent electron-positron (e^+e^-) lose most of their energy to Bremsstrahlung radiation. This cycle ignites a cascade in the crystal and continues until the Bremsstrahlung photon energies fall below the pair production energy threshold (1.02 MeV). The surviving e^+e^- with lower energies than pair production energies, lose further energy through processes other than Bremsstrahlung, such as multiple scattering or ionization. The electromagnetic showering, related to a certain search, can be described by two main parameters: the radiation length X_0 of the detector matter, by which we can determine the shower depth, and the Molière radius R_M, which characterizes the spread of the shower. Showering depth or, more simply, the cascade length scales linearly with X_0 and logarithmically with energy,

$$X = X_0 \frac{\ln E_0/E_c}{\ln 2}, \tag{4.4}$$

where E_c is the critical energy below which the Bremsstrahlung radiation is not the

dominant process for electrons to lose energy. On the other hand, the lateral spread of the shower is chiefly due to the multiple scattering of electrons after multiple scattering dominates. Up to the point where critical energy is achieved, the lateral spread of the showering cone is contained in a cylinder with a radius < 1 of the radiation length (X_0). When scattering becomes the dominant factor, the radius of the cone scale with the *Molière radius* R_M. A cylinder with a radius of R_M contains 90% of the shower. These parameters help us distinguish showers from one another and identify each electron, positron, and photon track [130].

To each $PbWO_4$ crystal, a photodetector is attached, which collects the light yield from the showers and converts and amplifies them to an electrical signal used to analyze the data. Specifically, Avalanche Photodiodes (APDs) are used for this light detection because they are designed to withstand the high rate and high electromagnetic fields present in the ECAL [131].

The HCAL: HCAL is tasked with measuring the energy and position of the charged and neutral hadrons, such as protons, pions, kaons, etc., at CMS. The HCAL consists of layers of dense material (brass or steel) interleaved with tiles of plastic fluorescent scintillators. The rapid light pulse signals produced by the scintillators are read out by hybrid photodiodes, delivered via wavelength-shifting fibers.

The incoming hadrons interact strongly with the nuclei of the calorimeter's material, provoking a hadronic shower. The development of the hadronic showers is relatively more complex than electromagnetic showers. Hadronic showers can be described in three main stages: hard collisions, spallation, and nuclear de-excitation. The incoming hadron initially loses energy by ionization of the material before colliding with the heavy nuclei of the detector material. The collision will initiate an intranuclear cascade where charged, and neutral pions (π^+, π^-, π) are produced alongside with excited nuclei.

Neutral pions can decay to photons ($\pi \rightarrow \gamma\gamma$), leading to an electromagnetic component in the hadronic shower. The de-excitation of the excited nuclei leads to the production of other particle debris, including neutrons, α particles, and nuclear fragments. The parameter that characterize nuclear interaction, counterpart to radiation length X_0 in electromagnetic showers, is *interaction length*, λ_{int},

$$\lambda_{int} = \frac{1}{\sigma n}, \tag{4.5}$$

where σ is the cross-section of the hadronic interaction with the material nucleus and n is the number of atoms per unit volume in the target material. The interaction length λ_{int} is generally significantly larger than radiation length X_0, which means the longitudinal dimension of a hadronic shower is larger than for an electromagnetic shower of corresponding energy. This is why in Fig. 4.2 the light blue ECAL is much thinner than the yellow HCAL. The lateral size of the hadronic showers is also larger than the electromagnetic showers. It is crucial to prevent any hadron *punch through* the HCAL and reach the muon system because that can obscure muon detection; therefore, the HCAL is designed to have at least $10\lambda_{int}$ of thickness.

The energy used to release protons and neutrons from binding energy to the calorimeter nuclei and nuclear recoil and delayed photon from nuclear de-excitation leaves the detector without being detected by the ECAL. Fluctuations in this *invisible energy* play an essential part in the degradation of the intrinsic energy resolution, resulting in a lower energy resolution compared to ECAL. To mitigate the low energy resolution, HCAL is a sampling detector, meaning alternating layers of the absorber and fluorescent scintillator are staggered to further ensure the hermeticity of the system. The energy of a given jet —narrow cone of energetic hadrons —is calculated by summing the energies over each layer, called *towers*. This *missing* energy in the transverse plane

(MET) can be inferred by ensuring the momentum conservation in each event. To measure the MET precisely, the hermetic property of HCAL is necessary as particles such as neutrinos, WIMPS, or any new physics particle can escape detection in the HCAL.

The strong interaction coupling, α_s, is a *running* coupling constant and is dependent on transferred momentum squared (Q^2). For small distances (high Q^2), strong coupling vanishes asymptotically, while for long-distance (low Q^2), the strong coupling grows, leading to asymptotically free quarks. The QCD is a model with color confinement, a phenomenon that dictates that color-charged particles such as quarks and gluons cannot be observed to exist freely. Consequently, they lump together to form hadrons. This process in high-energy physics is called hadronization. During parton showering in the HCAL, the interaction scale falls, and the strong interaction coupling rises, eventually triggering the process of *hadronization*. The tight cone that contains the hadrons produced by hadronization is called a *hadronic jet*. Because of the non-perturbative nature of QCD at the hadronization scale, the process is not fully understood. Nonetheless, Monte-Carlo simulations, such as PYTHIA and HERWIG [132] enable us to simulate the hadronization process and jet analysis, which is essential for studying new physics at the LHC [133]. The process of hadronization and its simulation is further discussed in Sec. 5.2.1.

4.6 Detecting Muons

As the name of CMS suggests, muon identification and momentum measurement is the key focus of the detector. The muon system is the outermost system of CMS, for muons' penetration range, is among the longest compared to other charged (hadronic) particles. The muon system is appropriately positioned behind the HCAL and has

the thick layers of HCAL acting as a hadron filter for its signals. Muons, similar to electrons, interact electromagnetically with muon system materials. Muons, and not electrons, are the main subject of the muon system since electrons, due to their significantly lower mass, lose most of their energy through Bremsstrahlung, resulting in a smaller ranger than muons. On the other hand, the much heavier muons lose their energy mainly through ionization and survive the ECAL and HCAL. The muon system is tasked with three major functionalities: muon identification, tracking (see Sec. 4.4), triggering (see Sec. 4.7), and momentum measurement.

The two major sections of the muon system are the *barrel* and the *endcap*. A muon is identified and measured by the fitted curve among hits in four muon station; Muon Station (MS) 1-4 [134]. Reconstructing the muon path is key to studying its features, such as momentum and direction. The magnetic field produced by the CMS magnet bends the tracks as muons are charged particles. This gives a measurement of its momentum because we know that particles traveling with more momentum bend less in a magnetic field.

The muon system is the outermost system at CMS; therefore, its volume coverage is much greater than the inner systems. Due to the greater volume, the usage of silicon-based detectors for tracking and momentum measurement is not cost-effective. Consequently, gaseous chambers are adopted for the muon system. The basics of the gas chamber operation are discussed in Ch. 1, where GEM detectors are introduced. In short, when a charged particle passes through a gas chamber, it ionizes the gas molecules and initiates an avalanche of charges via applied high voltage. Most gas chambers operate within the proportional region, where the charge collected by the readout system increases proportionally with a further increase in the detector voltage, while the number of primary ion pairs remains unchanged. This avalanche mechanism already provides sufficient signal amplification and improves the signal-to-noise

ratio within the detector. This feature reduces requirements for subsequent electronic amplification, further reducing the cost.

The barrel has 250 Drift Tubes (DTs) chambers, 480 Resistive Plate Chambers (RPCs), and the endcaps are armed with 540 Cathode Strip Chambers (CSCs), and 576 RPCs. The recent addition to CMS is the 72 modules, each containing two gas electron multiplier (GEM) detectors in order to detect muons that scatter at an angle of around 10° in relation to the beam axis. I have developed the concept and the overview of the construction and Quality Control (QC) of these gem detectors in Ch. 1. The GEM modules were not operational for the duration of my thesis. A slice of the quadrant of the CMS muon system in which each of the detectors is depicted in Fig. 4.4.

Protons are injected and boosted along the beam pipe, and their momentum is primarily in the direction of the pipeline before the collisions. Thus, any transverse momentum (p_T), momentum component in the perpendicular plane to beam pipe, is one of the paramount observable in the CMS experiment and expresses the momentum that arises from the *products* of the proton-proton collisions. In conjunction with the angle with which incoming products of the collisions encounter the detectors, we can accurately describe the features of the muons and the physics they represent. In context of LHC physics, the angle is delineated by the Lorentz invariant *pseudorapidity*, η, and is defined in Eq. 4.6,

$$\eta := -\ln\left[\tan\left(\frac{\theta_{cm}}{2}\right)\right], \tag{4.6}$$

where θ_{cm} is the center of mass polar angle with respect to the beamline. More on the CMS coordinate system can be found in App. E.

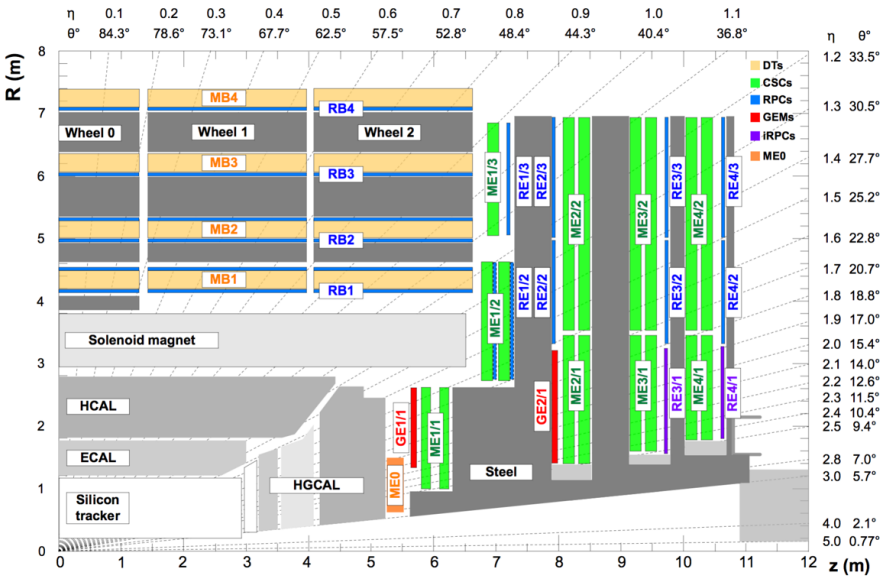


Figure 4.4: Quadrant of the CMS muon system showing present detectors, i.e. Drift Tubes (DTs), Resistive Plate Chambers (RPCs) and Cathode Strip Chambers (CSCs), and the locations of the proposed GEM detectors (not for duration of this thesis), i.e. the ME0, GE1/1 and GE2/1 stations with their respective η coordinates [8]

4.7 Triggering

Once the beams are operational, and it is time to record data, the struggle becomes storing the data and only storing the data that would be *interesting* for us. The interaction point at CMS can host as many as a billion proton-proton collisions per second when performing at its peak. That number of collisions per second, if stored, would translate to a corresponding data rate of ∼60 TB/s. It is not only practically impossible to record that volume of data; more importantly, it is unnecessary to do so. The interesting physics, the new phenomena, lies in energetic and hard proton-proton collisions; therefore, we target capturing those events and getting rid of the rest. We can record only potentially interesting events with a multi-level complex yet precise triggering system. The CMS experiment has adopted a two-level system. A Level-1 (L1) trigger utilizes a system of synchronized hardware which reduces the event rate from 40 MHz to 100 kHz. The rate is further reduced to several hundred Hz

through High-Level Triggering (HLT), which is purely software-based. The remaining events are read by the data acquisition system (DAQ) from all subsystems and are stored in a permanent database at CERN. A sound trigger system will offer high signal efficiency, significant reduction of high-rate unwanted processes, and high flexibility to react to changing experimental conditions. In short, finite storage capabilities and finite computer power determine what trigger rate can be afforded. In this section, we remain focused on the muon trigger system as this dissertation contains searches that look for muon signatures at CMS. Information about the rest of the trigger system can be found in Ref. [9].

4.7.1 The Level-1 Trigger

The level-1 (L1) CMS trigger system is essentially a synchronized system of custom-made electronics. The muon system, HCAL, and ECAL participate in the L1 trigger. The tracker system (pixel and strip detectors), however, is not used in the L1 trigger. Each muon detector system plays a role in the L1 trigger to ensure optimum coverage and redundancy. The hits (energy deposits) in DTs and CSCs —the $|\eta| < 1.2$ and $|\eta| > 0.9$ region —are processed via a track segment finder and transmitted to regional muon track finders (MTFs) by optical fibers. Track finders implement pattern recognition algorithms to identify muon candidates and approximately measure their p_t, using the bending angle. The RPC hits ($|\eta| < 1.6$) are directly sent to pattern logical gates of comparators that identify muon candidates. The designed GEM detector stations —not operational for the duration of this thesis —extend the forward coverage ($1.6 < |\eta| < 2.4$) and provide additional redundancy for CSCs and RPCs. The initial trigger output from the muon systems is called the trigger primitives (TPs). The TPs from DT and RPC hits from the same muon stations are integrated to form an output known as super-primitives which benefits from the spatial resolution of DTs and time

resolution of RPCs. Ultimately, the muon track segments from all muon detectors are collected and transmitted into the global muon trigger (GMT). The GMT then ranks muons according to their p_t and removes any duplicate tracks across regional segments. Later, the HLT algorithms use the menus offered by GMT with analysis-specific selection criteria. More detailed discussion on muon system L1 trigger system in can be found in Ref. [134]. An overview of the L1 trigger system is shown in Fig. 4.5.

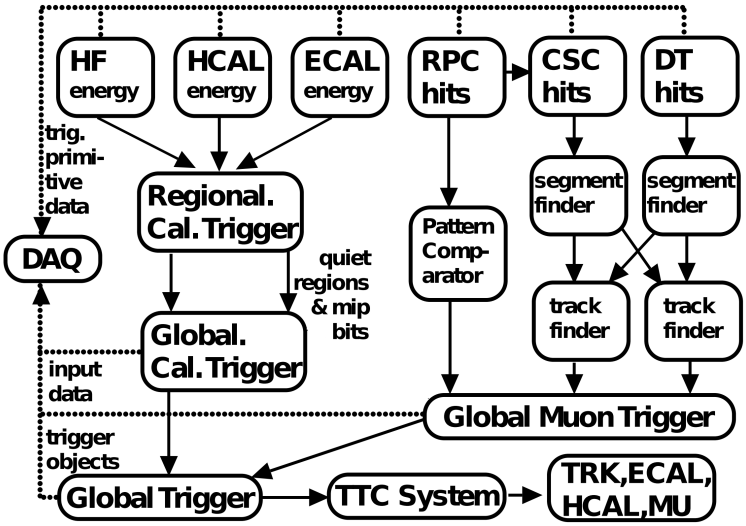


Figure 4.5: Overview of the CMS overall L1 trigger system [9]

At ECAL and HCAL, the energy deposits constitute the TPs. The two-tier calorimeter trigger algorithms receive and calibrate the TPs and then reconstruct the physics objects, e.g., leptons, jets, and missing energy. Similar to the muon trigger system, the regional triggers are then sent to global trigger (GT) processing. More in reference [135]. Decisions from muon and calorimeter trigger algorithms are combined by the GT, and the final decision is reached based on an OR logical gate. This is done by utilizing field-programmable gate arrays (FPGAs) on an advanced mezzanine card (AMC), more specifically AMC13, in a MicroTCA (μTCA) crate. Once the decision is made, the signal from all participating detectors is recorded by the DAQ system and stored in the buffer for later analysis.

4.7.2 Muon High Level Trigger

As discussed above, the information from the tracker is not used by the L1 trigger. At the HLT, however, the information from the muon system as well as tracker sub-detectors is combined to identify the muons and determine their p_t. The muon HLT is done in two main stages: Level-2 (L2) uses the muon system information only, while Level-3 (L3) combines the tracker information with muon sub-detectors. The two-level HLT improves the momentum measurement resolution of the accepted trigger objects.

A set of HLT paths for each analysis, wherein a collection of reconstruction and filtering modules appropriate for the said analysis selection, is chosen. The practical benefits of the HLT include reducing the CPU consumption in the final analysis, rejecting fake muon objects, and further reducing the trigger rate. Hence, the HLT menus continuously evolve to maintain relevance to a growing variety of offline analysis selections and criteria. One of the main techniques that combines the information from all the CMS sub-detectors for particle identification as well as some higher-level properties such as particle isolation is *particle-flow* (PF). A comprehensives introduction to the PF algorithm is given in the Refs. [136, 137].

L-2 Muon in HLT: The muon track reconstruction starts with the muon sub-detectors, CSC and DT *hits and segments* patterns, which are referred to as *seeds*. The L1 trigger decision determines what seeds are to be fed to the HLT algorithms. The azimuthal angle between hits from two segments is used to parametrize p_t for a given track. The Kalman filter technique is used to fully reconstruct muon tracks [138]. The Kalman filter technique is a recursive algorithm that performs pattern recognition and track fitting and uses the calorimeter isolation information to refine the muon p_t measurements.

L-3 Muon in HLT: The L3 muon reconstruction further improves the muon p_t resolution by making use of the incredible momentum and vertex reconstruction resolution of the inner tracker detector. There are three general steps to L3 muon HLT: seeding the tracker reconstruction starting with L2 information, reconstruction of tracker tracks, and lastly, the L2 muon seeds are propagated to the tracker, using a stepping helix propagation. This match-up (the tracker seeds and muon seeds) is done in both directions to maximize the reconstruction efficiency: outside-in (OI) and inside-out (IO). HLT algorithms have a time budget for their calculations set by available electronics so that all interesting objects can be triggered on and *parked* for further offline analysis. Once the L3 muons are fully reconstructed, their parameters, such as p_t and impact parameter, are used to filter the events for analysis.

During the data taking process at CMS, the data are labeled by the HLT they were selected by, e.g., `SingleMuon`, `DoubleMuon`, etc. These data contain physics objects such as tracks, muons, jets, and vertices. When saved, this *RAW* data events are sent for offline reconstruction. More details on the L3 algorithm can be found in Ref. [137].

4.7.3 Offline Muon Reconstruction

The data accepted by the HLT are recorded to *tapes* at CERN for the offline reconstruction. The label *Offline* reconstruction indicates that the reconstruction process starts after the data taking with CMS detectors has ended. The offline muon reconstruction benefits from similar algorithms such as PF and Kalman fitter at the HLT level, except there is more time available for further optimal calibration, alignment, and filtering on the event by event basis, appropriate for each specific search. This is achieved by using a modified configuration of the track reconstruction algorithms. For example, the CSC TP used in online muon reconstruction rely on a local electronic trigger where pattern recognition is performed, whereas in offline muon reconstruction, CMS reconstruction

software provides fits and charge weighting to form segments. These fits generally take more time to perform, and that is why they are done offline [139].

With the available HLT muons at the muon system, hits within various stations of the muon system are combined to form fully reconstructed local segments. Local segments are further combined until we have a complete picture of the muon sub-detectors. This picture will constitute the *seed* from the muon system. The seeds from the muon system are then fit together using the Kalman fitter, and the whole muon track through the muon system is reconstructed. The muons solely reconstructed within the muon system are *standalone* (SA) muons. At the tracker, tracks are extrapolated inside-out, and if a track matches with at least one triggered segment in the muon system, it will be labeled a *tracker muon*. The outside-in extrapolated track between the SA and the tracker muon is called a *global* muon. A description of the extrapolation and the parameters that are considered for the match-up is given in Ref. [139].

The CMS PF algorithm then accepts the reconstructed muons and applies a series of selections to them. Selections such as isolation and Single/DoubleMuon trigger types apply to SA, tracker, or global muons. Isolation information —energy deposits around a track —from the calorimeter is applied to the muon system tracks, and tracker isolation is applied to the extended tracks back to tracker to confirm the muon ID further. Muon ID criteria are developed to aid with analysis categorization. Muon IDs include loose, medium, tight, soft, and high-momentum muons. For instance, loose muons are used for prompt analysis at the primary vertex, which is the focus of this thesis. More details on the muon IDs can be found in Ref. [134].

There is further intricacy to p_t determination at CMS. The Tune-P algorithm chooses a final p_t measurement from several refits of the muon tracks based on statistical goodness-of-fit and relative p_t resolution. The fit that is used in my analysis is the Inner-Tracker-fit, where the p_t is determined solely based on tracker fit. This is

the choice for this analysis since our muons generally have $p_t < 50$ GeV. The muon system at CMS does not sufficiently contribute to soft muon p_t measurement due to multiple scattering and energy loss as they pass through the calorimeter and the flux-return yoke. Conversely, for muons with $p_t > 200$ GeV the muon system is used for p_t determination.

4.7.4 Muon Triggers in This Analysis

In the model-independent analysis, triggers are designed to accommodate high efficiency for each benchmark model. Since the possibility of displaced muons in the dark SUSY model was considered, an L2 muon trigger with p_t threshold of 23 GeV was introduced to make up for the trigger deficiency caused by inability of `DoubleMuon` triggers to capture displaced signatures. The other three triggers used in this analysis adopt a lower p_t threshold to maximize the overall signal efficiency. The trigger paths used for the 2018 model-independent analysis are shown in Tab. 4.1.

Table 4.1: The 2018 analysis HLT trigger paths. The p_t requirements of each trigger and the number of triggered muons in each HLT are expressed in the path labels.

Index	HLT path
1	HLT_DoubleL2Mu23NoVtx_2Cha_v*
2	HLT_Mu18_Mu9_SameSign_v*
3	HLT_TripleMu_12_10_5_v*
4	HLT_TrkMu12_DoubleTrkMu5NoFiltersNoVtx_v*

4.8 Concluding Remarks

The Large Hadron Collider has presented an unprecedented opportunity for unveiling unexplored high-energy scales. This territory is ripe with possibilities for discovering new physics, such as dark matter, the focus of this analysis. In this chapter, I detailed

the intricacies of this magnificent machine, mainly the Compact Muon Solenoid. As I outlined in this chapter, the operational mechanisms of the Compact Muon Solenoid, such as the trigger system and the muon identification, are integral concepts in the coming chapters. The trigger paths introduced in this chapter will be used for the analysis of 2018 data.

Chapter 5

Samples

There are two main inputs to data analysis in the search for new physics: the data collected by the experiment (CMS data) and the Monte Carlo (MC) simulation. We receive *reconstructed* data from CMS, as discussed in the previous chapter. I need the MC samples to simulate the benchmark models, which represent the anticipated *signals*. The auxiliary function of simulated signal samples is to confirm the model independence of the overall analysis. The MC simulation is correspondingly also used to simulate specific SM background processes. In addition to MC-generated background samples, for lower masses, *data-driven* samples are used. In this chapter, I go over the list of CMS data used in this analysis, the background samples, and the MC simulation of the vector-portal model.

5.1 Data Samples

In this analysis, our collaboration used the data recorded by CMS in 2018. More specifically, the proton-proton collisions at 13 TeV center-of-mass energy, corresponding to 59.9 fb^{-1} integrated luminosity. The analysis data sample includes only runs where

the LHC provided stable beams, the CMS silicon tracker, muon system, and trigger were performing well, and the luminosity was properly measured. We perform the signal search with fully reconstructed events stored in the `DoubleMuon` primary dataset (PD) shown in Tab. 5.1, based on triggers described in Sec. 4.7.4. More details about the dataset eras and number of events recorded are listed in Tab. 5.1.

Table 5.1: The `DoubleMuon` dataset used in 2018 analysis

Dataset Lables	Number of Events
/DoubleMuon/Run2018A-17Sep2018-v2/MINIAOD	75,499,908
/DoubleMuon/Run2018B-17Sep2018-v1/MINIAOD	35,057,758
/DoubleMuon/Run2018C-17Sep2018-v1/MINIAOD	34,565,869
/DoubleMuon/Run2018D-PromptReco-v2/MINIAOD	169,225,355
Total	314,348,890

5.2 Monte Carlo Samples

In its essence, a MC simulation is a statistical method for numerical integration computation. In other words, an MC simulation makes use of repeated random sampling to achieve numerical results. MC simulations are a factotum of experimental particle physics. The MC simulations are essential to various phases of physics analysis: event generation for theoretical models, cross-section calculation, optimization of data reconstruction software, detector responses and optimization, and evaluation of some of the systematic uncertainties, to name a few. Some of the well-known MC event generators include: Herwig, PYTHIA [132], Sherpa, POWHEG, and MADGRAPH5_aMC@NLO [140, 141]. The most well-known detector MC simulation tool is the GEANT4 simulation toolkit [142]. An essential input to the MC simulation are the Parton Distribution Functions (PDF).

5.2.1 Simulation Overview

The flow of the event generation in a typical analysis comprises four main steps: hard scattering process (fragmentation with the largest momentum transfer), parton-shower phase modeling (i.e., quarks and gluons), particles decay before hadronizing, hadronization, and unstable hadron decay (experimentally measured branching fraction ($\mathcal{B}$), phase-space, and distribution of the decay products).

Model implementation: I implemented the vector-portal benchmark model via the *FeynRules* [143] package. Here is the description of the package on their website:

> "*FeynRules* is a *Mathematica* package that allows the calculation of Feynman rules in momentum space for any QFT physics model. The user needs to provide FeynRules with the minimal information required to describe the new model contained in the so-called model file. This information is then used to calculate the Feynman rules associated with the Lagrangian. The Feynman rules calculated by the code can then be used to implement the new physics model into other existing tools, such as MC generators. This is done via a set of interfaces which are developed together and maintained by the corresponding MC authors."

I based my model on a simplified dark matter model publicly available on the *FeynRules* database [144]. The model features a dark sector s-channel vectorial mediator and DM candidate. In the initial model, the mediator directly couples to the quarks. After performing sensitivity studies, I concluded that for the LHC to be sensitive to this signature in Run II, the model has to be tweaked [145, 146]. To achieve reasonable reach, I tweaked the model to incorporate kinetic mixing of the dark Z (Z_{D}) with the SM Z and introduced a complex dark scalar s_{D}. The Feynman diagram of the resulting

model is shown in Fig. 3.3. After properly defining the Lagrangians and other relevant parameters in the model file (*.fr), I run the FeynRule in *Mathematica*, and produced Universal FeynRules Output (UFO) [147]. The UFO directory contains parameter files that are used by event generators to calculate the matrix element information related to the model. A more detailed discussion on model implementation in FeynRules is presented in App. C.

Parton Distribution Functions: The quarks, anti-quarks, and gluons within a hardon (protons at LHC) are collectively called *Partons*. The MC simulation demands parton distribution functions (PDF) as input to simulate the parton momentum distributions and perform factorization calculations. These partons each carry a fraction x of the total momentum of the proton at a squared energy scale Q^2. The distribution function of momenta of the partons inside the hadrons ($q = -Q^2$) is called a PDF. The perturbative QCD (pQCD) equations are effective at energies where the running coupling constant is well below one ($\alpha_s(Q^2) << 1$). The QCD equations at these energies are able to quantitatively predict the *rate* by which the PDFs change with Q^2. The QCD Parton distribution evolution equations [148, 149] are developed for different approximation levels such as Leading-Order (LO), Next-to-Leading-Order (NLO), and Next-to-Next-Leading-Order (NNLO). The QCD equations, although useful for determining the evolution of the PDFs with energy, cannot make definitive predictions about the absolute momentum fraction x at a given Q^2. This information is extracted from data by probing in a Deep Inelastic Scattering (DIS) experiment. The extracted PDFs are used in factorization theorems [150], where observable cross-sections are factorized by the PDFs. Experiments from all around the world have conducted these types of experiments and produced PDF sets. Some notable data sets include ABM [151], HERA [152], and NNPDF [153]. All vector-portal samples are generated using the

leading-order NNPDF3.1 set of PDF.

Hard Scattering Simulation: Event simulation starts with the hard scattering process where *highly energetic* hadrons within beams collide, and initial collision products are produced. A comparative description of various event generators is given in Ref.[154]. The momentum transfer fractions of partons within proton beams at the LHC are determined by random sampling of the PDFs. An integral is then performed over differential cross-sections, and hence production cross-section is calculated. This calculation is performed over the available phase space. The MC method is used to perform the integral in factorization equations and to obtain the momentum distribution of primary objects in the hard process.

The UFO output is fed to MadGraph5_aMC@NLO v2.6.5 for event generation. The resulting output is in the Les Houches Events (LHE) format [155]. An LHE file contains the kinematics of the events and the production cross-section information. The Figs. 3.4 3.8 are produced using the information extracted from the LHE files. The information from an LHE file can be cross-checked with theoretical calculations [156]. As a sanity check for the samples, the distribution of all four muon p_t at the LHE level is shown in Fig. 5.1.

Parton Showering: Hard scattering processes, as discussed above, carry high momentum transfer which translates to high acceleration. Accelerated partons emit QCD radiations; i.e., colored partons emit gluons that carry color charge, resulting in further radiation. The development of this QCD radiation is referred to as *parton showering*. The parton shower evolution consists of repeated *splitting* of a parton of type i into $j + k$, e.g., $q \rightarrow q + g$. The exact analytic calculation cross-section of the parton showering process is not computationally feasible because of inherent non-perturbative

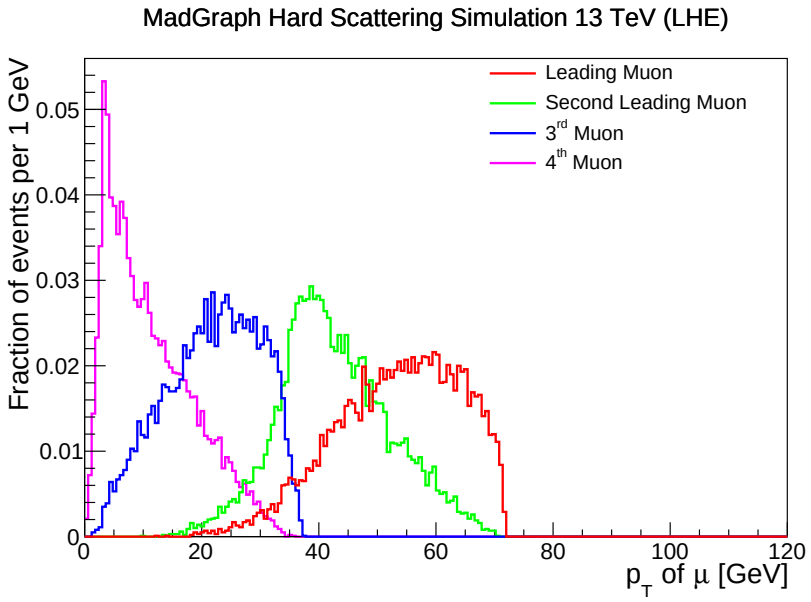


Figure 5.1: Results from hard scattering simulation, using MadGraph5_aMC@NLO, extracted from the LHE file. The p_t distributions of four generated muons for a MC signal sample of the vector-portal model are shown. This sample corresponds to m_{Z_D}=150 GeV and m_{s_D}=100 GeV.

nature of partons. Instead, approximations with varying orders are considered. In other words, the differential cross-section for the splits can be expanded to LO, NLO, NNLO, and ... . The resulting differential cross-sections are then averaged over the phase-space of interest. The role of the MC method in this process is to generate the phase-space values for each split based on the defined PDFs.

The virtual mass-squared q^2 (*virtuality*) of the partons evolve with the evolution of the parton shower. The virtuality of the hadrons is highest farthest away from being on-shell and declines until it reaches the momentum transfer squared Q^2. This is when the virtual particles start to hadronize at $q^2 = Q^2 \sim 1$ GeV^2.

The parton-showering step models the evolution of partons (i.e., quarks and gluons) radiated from all stages of the hard scattering, including the initial-state radiation (ISR) and the final-state radiation(FSR). The energy scale for showering is simulated down to $\mathcal{O}(1)$GeV [157].

The event generation process furthermore models the so-called *underlying event* activity (UE). In this context, the "underlying" activity refers to the activity that does not originate from the most energetic hard scattering process. Namely, the non-colliding partonic products emitted in the forward direction contribute to the final-state radiation, e.g., beam-beam remnants (BBR) and multi-parton interactions (MPI). The UE partons eventually fragment into hadrons through gluon radiation and quark-antiquark splittings, often in the form of jets. The UE is simulated approaching the QCD scale ($\Lambda_{QCD} \sim 0.2$GeV). The UE is modeled by generators such as PYTHIA (version 8.153), HERWIG++, and EPOS. In this analysis, PYTHIA 8 was used. The BBR and MPI are modeled using a set of parameters called a *tune*. We used CP5 tune [158] for our MC samples which uses the NNPDF3.1 [153] PDF set.

Hadronization & **Detector Response:** The legitimacy of the pQCD regiment is closely dependent on the strong running coupling constant α_s. The strong coupling constant value increases at lower hadron showering development scales, where the partons are at low momentum transfer, in a long-distance regime. Thus, the perturbation theory becomes invalid at low showering scales, and the system enters the non-perturbative phase. The final state hadrons are the products of this *hadronization* process. We currently cannot perform non-perturbative calculations to describe the hadronization process accurately. Instead, we rely on applying hadronization models such as the Lund String model[159] or the Cluster model in the event generators, such as PYTHIA and HERWIG[157]. Next, the generated events are hadronized with the event generators and then passed through the GEANT4-based CMS detector simulation and processed with the `CMSSW_10_2_X` release of the CMS software for 2018. The samples are then fully reconstructed using the same CMS software. The commands used within `CMSSW` for each reconstruction step is listed in App. B.

Luminosity & Reconstruction: The signals from the sub-detectors as simulated by GEANT4 are to be digitized for the L1 and HLT decision operations. This step is also done in the subsystems of CMSSW software. Characterizing trigger responses to the MC events and comparing them to the data is crucial as it constitutes a source of systematic uncertainty. An overall trigger scale factor is calculated to account for the differences in the trigger response in data versus MC. The MC events are weighted by the HLT scale factor. More details can be found in Sec. 8.1.

To simulate luminosity, we have to access the luminosity information of the data taking era, for which we wish to simulate the events. The number of interactions per BX determines the profile of luminosity. For each simulated event, a random number from the luminosity profile is chosen, representing the instantaneous luminosity of the event. Lastly, the events are reconstructed using the same software infrastructure that is used to reconstruct CMS recorded events, as discussed in Sec.4.7.3.

Pileup Simulation: Another challenge to data analysis is presented by the *PileUp* (PU) interactions. In each bunch crossing (BX) at CMS, a large number of extra inelastic proton-proton collisions coincide with the hard scattering collision. Thus, the PU events, by definition, involve proton-proton collisions that are typically softer than the primary head-on collision. PU is not be to be confused with UE objects such as MPI, which refers to softer interaction *within* the hard scattering collision. The PU events can be simulated as minimum bias (MB) events. MB refers to the events that are selected by the triggers with minimal requirements, i.e., introducing a minimal trigger bias [160].

Until 2018, the PU simulation was done by simulating the MB hits in the detectors and was mixed with the signal hits, using GEANT4 and sent for digitization simulation.

This is a time-expensive process since there are hundreds of MB events for each simulated signal event. Since 2018, this method was abandoned for a *premixing* scheme where the premixed MB events are digitized separately before they are used with the digitized signal event [161].

The designed average number of PU interactions at the LHC was 25 per crossing; however, during Run II in 2017, the PU surpassed 32 on average, reaching up to 50 PU interactions in 2018. A good MC simulation has to reasonably model the PU effects corresponding to actual data samples used in the analysis. The number of PU interactions per crossing can be histogrammed for MC and data. Nevertheless, the two distributions can be slightly different based on beam conditions, instantaneous luminosity, and other data-taking variables. This gap can be filled by *re-weighting* the MC events such that the MC pileup matches that in the data.

5.3 Signal MC Samples

I privately simulated all the MC samples for the vector-portal model (signal MC samples). I designed the simulated samples to conduct two-dimensional scans over the masses of s_{D} and Z_{D}. Two-dimensional scans, rather than only scanning over the mediator mass, are done because, unlike the other models in this search, the mediator Z_{D} in the vector-portal model is not the SM Higgs. The two distinct DM species in the model open up another dimension in the parameter space. The list of produced samples is tabulated in Tab. 5.2.

Table 5.2: Produced MC signal samples for the vector-portal model. A total of 130 samples are produced and analyzed.

m_{Z_D} (GeV)	m_{s_D} range (GeV)	m_{s_D} step (GeV)	Number of generated events
85	5-40	5	10000
91.18	5-45	5	10000
95	5-45	5	10000
100	5-45	5	10000
110	5-50	5	10000
125	5-55	5	10000
130	5-55	5	10000
140	5-55	5	10000
150	5-55	5	10000
160	5-55	5	10000
170	5-55	5	10000
180	5-55	5	10000
190	5-55	5	10000
200	5-55	5	10000

5.4 Background Samples

Even though I will dedicate an entire chapter to the crucial subject of background modeling, it is only appropriate to discuss here also the background samples that we used. The background MC samples for this analysis are centrally produced by the CMS working groups. Briefly, several MC samples were used for background estimation in the high-mass region (above the Υ resonance). This region is dominated by Drell-Yan (DY) production and by high-energy QED radiation that mimic our four-muon signal.

The production of two Z bosons (ZZ) and their subsequent four-lepton decays at the LHC account for one of the most dominant backgrounds for our analysis in the high-mass region. The first ZZ mode of production, via quark-antiquark interaction (qq→ZZ), is simulated at NLO in pQCD, using POWHEG 2.0 [162] and showered using PYTHIA 8, and the subsequent leptonic decays (ZZ→4L) are simulated under CMSSW_10_X. The second mode of ZZ production is simulated at LO with MCFM [163, 164], showered by PYTHIA 8 up to Parton level scale, and then fully reconstructed (GlueGlue→Contin→ZZ→4mu). The last considered mode of production for ZZ is

through Higgs boson production, associated with gluon-gluon fusion, with ZZ production ensuing (GluGluH→ZZ). The POWHEG2.0 generator also simulates the Higgs production at NLO in pQCD. The further decay to four-leptons is generated via the JHUgen generator [165, 166]. Alternatively, the Higgs boson can be produced in association with top-quark jets (TTJets_DiLept). The $t\bar{t}$ samples are simulated by MadGraph5_aMC@NLO and `Madspin` for spin correlations [167]. The QED DY processes with zero, one, or two jets (0J, 1J, 2J) are considered for this region. The DY samples are generated at NLO, and their production rate is normalized to the expected cross-section as calculated by MadGraph5_aMC@NLO. The list of MC background samples for the high-mass region is given in Tab. 5.3. The UE tune and the PDFs for these background samples are chosen to be identical to the signal samples.

Table 5.3: Background MC samples for the high-mass region

Dataset	$\sigma \cdot \mathcal{B}$ [fb]	Events
/ZZToAL-13TeV-powheg-pythia8_TuneCP5/RunIIAutumnl8MiniAOD-102X upgrade2018 realistic_v15-v1/MINIAODSIM	1.325E+03	19,089,600
/GluGluToContinToZZTo4mu_13TeV_MCFM701_pythia8/RunIIAutumn18 MiniAOD-102X_upgrade2018_realistic-vl5-v2/MINIAODSIM	1.59	911,500
/GluGluToZZTO4L_M125_13TeV-powheg2_JHUGenV7011_pythia8/RunII Autumn18MiniAOD-102X_upgrade2018_realistic-v15-v2/MINIAODSIM	1.218E+01	958000
/TTJets_DiLept_TuneCP5_13IeV-madgraphMLM-pythia8/RunIIAutumn18 MiniAOD-102X_upgrade2018_realistic-vl5-v1/MINIAODSIM	5.26E+04	28,701,360
/DY_JetsToLL_2J_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIIAutumn18 MiniAOD-102X_upgrade2018_realistic_v15-v1/MINIAODSIM	3.68E+05	6,1054,774
/DY_JetsToLL_1J_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIIAutumn18 MiniAOD-102X_upgrade2018_realistic_v15-v1/MINIAODSIM	1.016E+06	96,753,082
/DY_JetsToLL_0J_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIIAutumn18 MiniAOD-102X_upgrade2018_realistic_v15-v1/MINIAODSIM	5.351E+06	93,979,507

5.5 Concluding Remarks

The data and the Monte Carlo simulations are the principal ingredients of this analysis. In this chapter, I outlined the specific data samples collected by the Compact Muon detector used in my analysis as well as the Monte Carlo samples that I utilized for modeling the background. I also demonstrated particularities and the procedure by which these signal Mone Carlo samples are produced.

Chapter 6

Event Selection

Each analysis introduces a set criterion for rejecting the events that do not serve the objective of the analysis and for accepting events interesting enough to keep the hopes of discovery alive. For a model-independent analysis like ours, the primary aim is to ensure a model-independent performance of signal benchmark models once we pair the muons to form dimuons candidates. The following selection in this chapter is designed and optimized to ensure that each benchmark model stays independent of its parameter at each selection step. The surviving events later will serve to set model-independent limits for the parameter space of this search.

6.1 Basic Selection and Triggers

Since this is strictly a four-muon final state analysis, I impose the selections primarily on the muons. The `DoubleMuon` and `TripleMuon` triggers described in Sec. 4.7.4 accept events with baseline requirements on p_t and charge. The basic selections target the p_T and η of the muons in each triggered event. The muons that are subjected to the selection come from the `DoubleMuon` dataset listed in Tab. 5.1. To start with, at least

three muon candidate for selection must be a PF muon as described in Sec. 4.7.2 and pass the loose ID selection (be a global or tracker muon). Furthermore, the requirements on each muon p_t is dictated by the HLT signal trigger. The p_t cuts have to be at least above the trigger requirements to avoid bias in the analysis and ensure high efficiencies. For instance, in Tab. 4.7.4, the L2 `DoubleMuon` (index 1) imposes a 23 GeV requirement on the leading muon. Inspecting each signal trigger and requiring the efficiencies after each cut stays flat and high, we came up with the following cuts: 1) at least two muons with $p_t > 24$ GeV, and four muons with $p_t > 8$ GeV. The η cuts are designed to ensure that we capture the high-energy (forward) muons in the cms endcaps (two muons with $|\eta| < 2$ and four muons with $|\eta| < 2.4$). These cuts are applied to both data and MC events at generator (gen) and offline reconstructed (reco) level muons. Despite my particular analysis (the vector-portal model) being a prompt analysis, this analysis has benchmark models that allow for displaced muons (Long-Lived signatures). We also allow for at most one stand-alone muon, defined as the muons that are exclusively reconstructed as L2 muons in the muon system and not by the PF algorithm. This measure proves necessary as the reconstruction of the track efficiency is generally lower.

6.2 Muon Pairing

In this analysis, I search for two distinct muon pairs, commonly referred to as *dimuons*: each new boson —s_{D} in the vector-portal model —decays into two oppositely charged muons as shown in Fig. 3.2 and Fig. 3.3. One of the essential missions in this analysis is to pair the muons in the data correctly. In designing an algorithm, two basic facts of the model prove helpful: 1) the muon pairs come from neutral bosons, which means the two muons must have opposite charges to conserve charge conservation in the vertex,

and 2) the muon pairs come from the same parent-scalar s_{D}, which means they should have similar invariant mass. The flow chart in Fig. 6.1 shows the pairing algorithm.

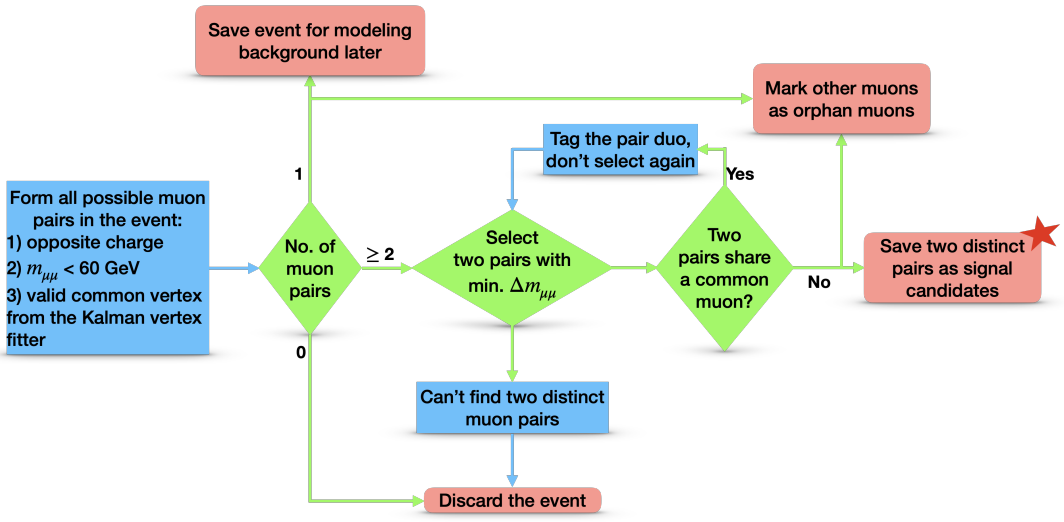


Figure 6.1: A flow-chart describing the muon-pairing algorithm in the model-independent analysis [10]

In the algorithm, for each event, all possible pairing arrangements are considered as long as the paired muons have opposite charges and the resulting invariant-mass is below 60 GeV ($m_{\mu\mu} < 60$ GeV $\sim m_h/2$). The invariant-mass upper restraint on a dimuon is chosen to be about five times the overall detector resolution at CMS to ensure that the detector systematic error is not affecting the final invariant-mass measurement results [134]. In this algorithm, the possibility of a muon being shared in multiple pairs is reserved. Then, each pair with a valid common vertex, fitted by a Kalman fitter, is saved. Next, the invariant mass for each of the formed muon pairs is computed, and the two dimuons candidates are sorted by invariant mass difference ($\Delta m_{\mu\mu}$), from lowest to the highest. We start with the lowest $\Delta m_{\mu\mu}$; if the formed dimuons do not share a common muon, then the dimuons will be saved as signal candidates. If they do share a common muon, then the next two dimuons in the $\Delta m_{\mu\mu}$ hierarchy will be considered. Such operations continue until two distinct dimuons are discovered. If no dimuon pairs

are found, or only one dimuon qualifies given the above conditions, we will mark all the remaining muons as *orphan* muons. The case of one-dimuon formation, conveniently, will come to use later in the background estimation studies. Worthy to notice, one could reject all the events with orphan muons; however, that would introduce unnecessary model dependence. Take the case of Higgs production associated with multiple highly-energetic jets. The energetic jet track could get reconstructed incorrectly as a muon object and pick up a signal muon stub. Here we would have an orphan muon in the event and rejecting the event would produce a bias for this particular event type.

6.3 Higher-level Selections

An additional selection set is imposed on the paired muons (dimuons). The cuts further trim the samples per the requirements of the search parameter space.

Fiducial volume cuts: Introduced in Sec. 4.4, the pixel system provides us with the possibility of tracking displaced signals (Long-Lived signatures). We impose fiducial cuts to accommodate the displaced signatures. The cuts on the volume of the detector are as follows: we require at least one pixel hit from at least one muon in each dimuon. That translates to transverse distance, $L_{xy} < 16$ cm with respect to the beamline and longitudinal distance $|L_z| < 51.6$ cm.

Dimuon fitted vertex probability cut: As touched on in Sec. 6.2 the dimuon vertices are fitted by the Kalman vertex fitter [168]. The Kalman vertex fitter algorithm returns a construction probability which indicates how likely it is that the two muons have a common vertex based on the chi-squared of the fit. We cut on this returned probability to reach an initial model-independent performance. There is a nuance

to be addressed here: the dimuon reconstruction probability is not a constant for the displaced muons. Empirical observation is that the reconstruction probability diminishes with increasing dimuon vertex displacement and opening angle. We define the following function and impose a cut to account for this loss of efficiency and recover more signal events. For each dimuon in an event, we require:

$$P_{\mu\mu} > P(L_{xy}, f(\Delta R), N_{SA}) \tag{6.1}$$

Here $P_{\mu\mu}$ is the dimuon vertex probability, L_{xy} is the vertex displacement in the transverse plane, $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ is the opening angle between the muon pair, and N_{SA} is the number of stand-alone muons for the LL signatures. The definition of the function is as follows:

$$P(L_{xy}, f(\Delta R), N_{SA}) = P_0 \times (1 - N_{SA}) \times e^{-\frac{L_{xy}}{R_0} \times f(\Delta R)} \tag{6.2}$$

Here P_0 represents the constant threshold probability of 20%, and R_0 indicates a constant length of 10 cm. The variable $f(\Delta R)$ is a polynomial function of 4^{th} degree. The $1 - N_{SA}$ factor guarantees acceptance for the events with one standalone muon because we expect particularly low reconstruction efficiency in those events.

Dimuon isolation cut: An isolation cut is applied to candidate dimuons to filter dimuons produced by heavy flavored quark decays such as $B\bar{B}$ and $t\bar{t}$. Decays from heavy flavor constitute one of our background sources and are discussed in the upcoming Ch. 7. The dimuon isolation is a function the dimuon p_t. To account for this dependency, *absolute* isolation is defined.

$$Iso_{\mu\mu} = \sum_{tracks} p_t(track) \tag{6.3}$$

The tracks that are summed over must meet the following requirements:

- $p_t > 0.5$ GeV
- $\Delta R(track, \mu\mu) < 0.4$
- $|z_{track} - z_{\mu\mu}| < 0.1$ cm
- tracks with all four muons, forming the dimuons, are excluded

All tracks in the vicinity of a dimuon ($\Delta R < 0.4$) are summed over as long as they have p_t lower than 0.5 GeV. The parameter z_{track} represents the z coordinate (along the beamline) of the primary vertex, while $z_{\mu\mu}$ is the propagation of the dimuon vertex back to the beamline. By requiring $|z_{track} - z_{\mu\mu}| < 0.1$ cm we are rejecting any non-prompt mediator (Z_{D}). Coincidentally, in the vector-portal model, the s_{D} scalars are prompt. Lastly, we exclude any muon track members of the reconstructed dimuon. By studying the absolute isolation distribution of formed dimuons in signal MC samples and comparing them with QCD enriched data, we discover that the cut, as described in the following equation, will discriminate the signal from the QCD background the best:

$$Iso_{\mu\mu,\Delta R<0.4} < 2.3\ GeV \tag{6.4}$$

We apply this cut to all dimuon candidates and reject those that fail it as background. Our approach toward the QED background is discussed in Sec. 7.3. One could apply isolation cuts directly to the reconstructed muons rather than the dimuons. However, that would introduce model dependence to the analysis [10].

Dimuon mass consistency cut: As discussed in Sec. 3.5.1, the dark scalars (s_{D}) are weakly coupled to the SM particles through kinetic mixing, resulting in narrow decay widths to the dimuons. Thus, the determining factor in determining the invariant mass

becomes the detector resolution. In Fig. 6.2, I show an example of the distribution of average invariant mass between the two dimuons.

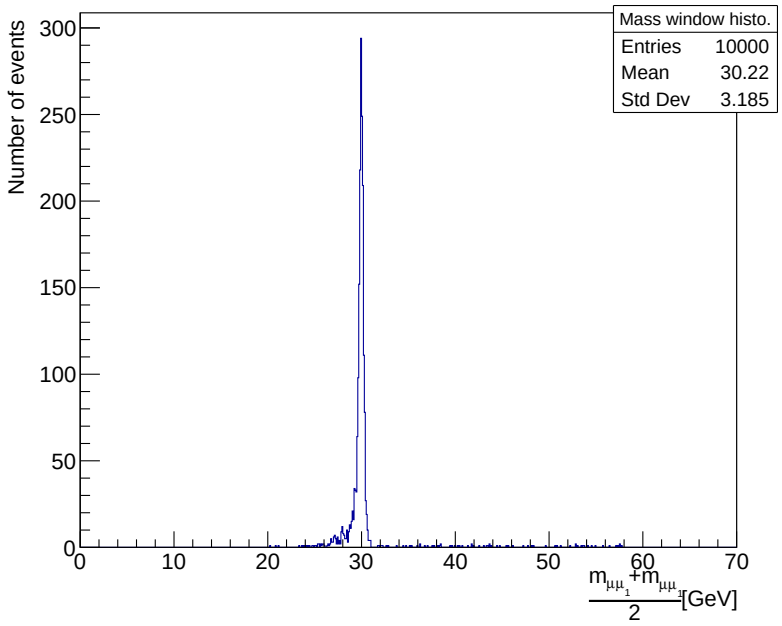


Figure 6.2: Invariant-mass average distribution for $m_{Z_{\rm D}} = 125$ GeV and $m_{s_{\rm D}} = 30$ GeV.

Since the dimuons are expected to originate from the same boson, I require that the selected two dimuons have consistent masses,

$$|m_{(\mu\mu)_1} - m_{(\mu\mu)_2}| < W(\tfrac{m_{(\mu\mu)_1}+m_{(\mu\mu)_2}}{2}), \tag{6.5}$$

where $W(\frac{m_{(\mu\mu)_1}+m_{(\mu\mu)_2}}{2})$ is a mass window as a function of the dimuon invariant masses $m_{(\mu\mu)_1}$ and $m_{(\mu\mu)_2}$.The mass window is set in such a way that 90% of all prompt signals and carves out a diagonal signal region (SR) in the two-dimensional plane of $m_{(\mu\mu)_1}$ and $m_{(\mu\mu)_2}$, as illustrated in Fig. 6.3. The remaining off-diagonal region is labeled the control region (CR).

More specifically, I generate a set of points and calculate the upper bound and lower

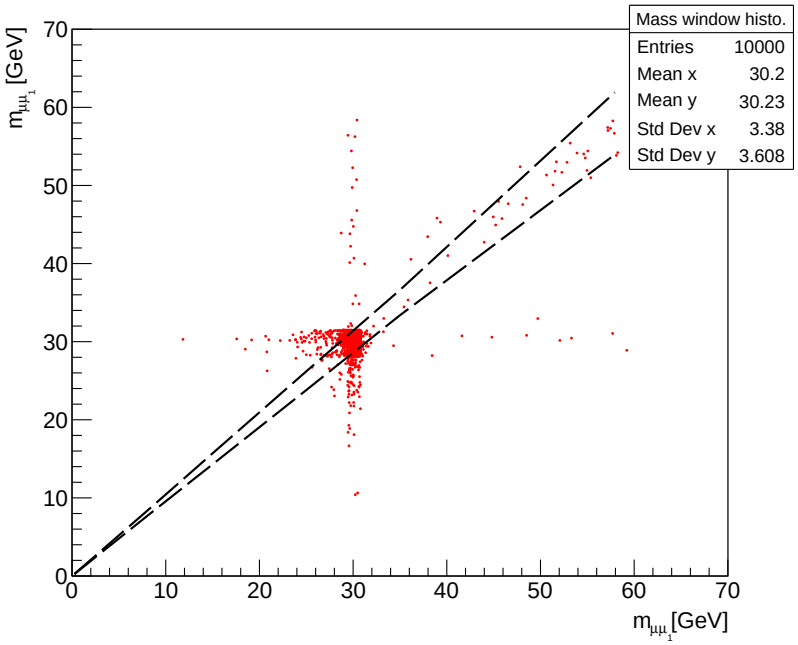


Figure 6.3: The invariant masses of the two dimuon pairs in each signal event plotted against each other for the signal sample $m_{Z_\mathrm{D}} = 125$ GeV and $m_{s_\mathrm{D}} = 30$ GeV. The dashed corridor signifies the mass window which later defines my signal region.

bound of the corridor and fit a linear function through it, as tabulated in Tab. 6.1.

Table 6.1: The mass window widths for the 2D dimuon invariant masses plane are listed for 11 points. The mass window width is required to guarantee 90% efficiency for the prompt vector-portal model.

m_{Z_D}	0.25	0.4	0.7	1	2	5	8.5	15	25	35	58
$W(\frac{m_{(\mu\mu)1}+m_{(\mu\mu)2}}{2})$	0.044	0.034	0.036	0.043	0.75	0.204	0.354	0.704	1.20	1.61	0.399

Other cuts: we also require at least one primary vertex of at least 4 tracks with a global z coordinate within 24 cm from CMS origin. The selection also includes vetoing Drell-Yan QED radiation which I further expand on in Sec. 7.3.

6.4 Model-Independent Verification

I present an example of a full cut-flow table in Tab. 6.2. I divide the table into generator level (gen) level and reconstruction (reco) level cuts.

Table 6.2: Exemplary cut-flow table for the signal sample $m_{Z_D} = 125$ GeV, $m_{s_D} = 30$ GeV. Each cut is quoted with their respective binomial statistical error.

Selection	Evts	Tot. Eff.	Rel. Eff.	Tot. Eff. Err.	Rel. Eff. Err.
Gen Level					
0 No cut	10000	1	1	0	0
1 one muon $p_t > 24$, $\|\eta\| < 2$	6301	0.63	0.63	0.005	0.005
2 two muon $p_t > 24$, $\|\eta\| < 2$	4160	0.416	0.66	0.004	0.006
3 three muon $p_t > 8$, $\|\eta\| < 2.4$	3714	0.371	0.893	0.005	0.005
4 four muon $p_t > 8$, $\|\eta\| < 2.4$	1912	0.191	0.515	0.004	0.008
5 Decay in Phase 1 pixdet	1912	0.191	1	0.00393	0
Reco Level					
6 one muon $p_t > 24$, $\|\eta\| < 2$	6352	0.635	0.635	0.005	0.005
7 two muon $p_t > 24$, $\|\eta\| < 2$	4180	0.418	0.658	0.005	0.006
8 three muon $p_t > 8$, $\|\eta\| < 2.4$	3765	0.377	0.901	0.005	0.005
9 four muon $p_t > 8$, $\|\eta\| < 2.4$	2012	0.201	0.534	0.004	0.008
10 Good primary vertex	2012	0.201	1	0.004	0
11 Two candidate dimuons	1075	0.108	0.534	0.003	0.011
12 Valid pixel hit	1075	0.108	1	0.003	0
13 Veto DY QED radiation	1073	0.107	0.998	0.003	0.00131
14 Dimuon isolation	896	0.0896	0.835	0.003	0.011
15 Signal HLT accepted	894	0.0894	0.998	0.003	0.001
16 Lxy cut SA-only Mu BKG	866	0.0866	0.969	0.003	0.006
17 Consistent dimuon mass	774	0.0774	0.894	0.003	0.01
$\varepsilon_{rec}/\alpha_{gen}$	0.405 ± 0.0161				

The gen level acceptance is defined as follows,

$$\alpha_{gen} = \frac{N_{gen\ selected}}{N_{events}} \tag{6.6}$$

where $N_{gen\ selected}$ is the number of events that survive the selection cuts, as described above, at gen level and N_{events} is the total number of events before the selection. Similarly, the reco level acceptance is formulated as:

$$\varepsilon_{reco} = \frac{N_{reco\ selected}}{N_{events}}, \tag{6.7}$$

Here $N_{reco\ selected}$ is the number of events surviving the selection cuts at reco level,

while N_{events} is again the total number of events. The parameters α_{gen} and ε_{reco} become useful devices for a standardized measure of model independence of the benchmark models. In other words, if the ratio $\varepsilon_{reco}/\alpha_{gen}$ is stable across the parameter space, it indicates that the benchmark model is independent of its parameters. Furthermore, the average value of $\varepsilon_{reco}/\alpha_{gen}$ across the parameter space can be used to compare the results amongst all benchmark models. I perform a scan in the m_{Z_D} plane as well as the m_{s_D}, calculating $\varepsilon_{reco}/\alpha_{gen}$ for each of my MC signal samples. An overall $\varepsilon_{reco}/\alpha_{gen}$ = 0.418±0.001 is found by calculating the average value from all signal MC samples. The results are illustrated in Fig. 6.4. I tabulate the efficiencies of each cut for every vector-portal model sample, discussed in App. D.

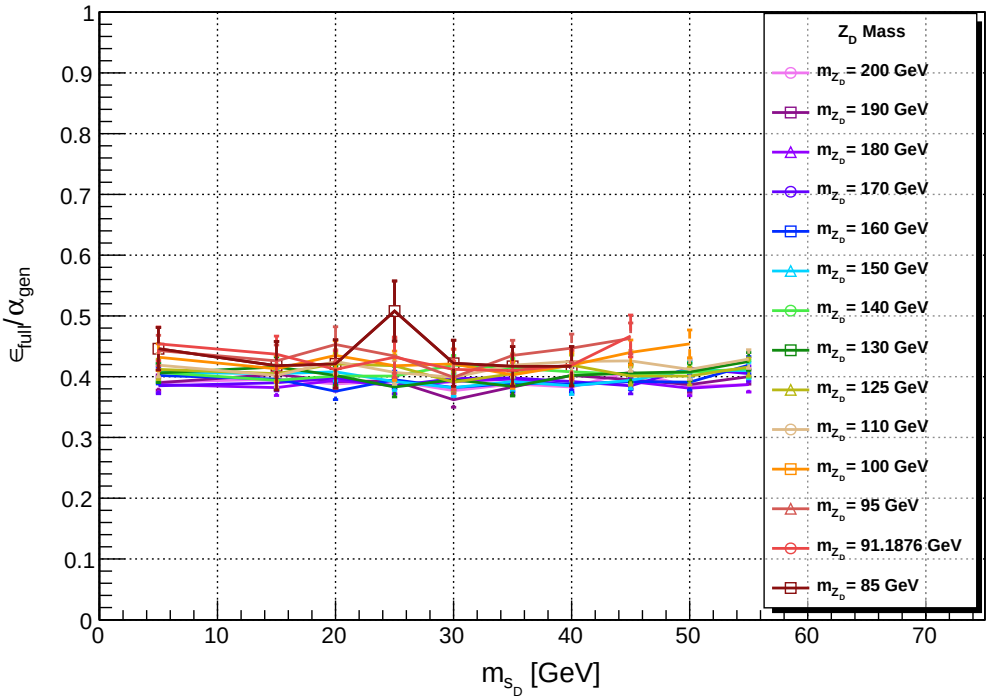


Figure 6.4: Total data selection efficiency ε_{reco} over generator level selection acceptance α_{gen}, $\varepsilon_{reco}/\alpha_{gen}$, as a function of the s_D mass for various Z_D masses in the vector portal model. The KM parameter, ε, is set to 10^{-2}.

6.5 Concluding Remarks

When presented with vast amounts of data, the selection criteria we can parse through the data becomes highly critical. For the case of my analysis, the selection also plays the central role in reserving the model-independence of the analysis. In this chapter, I discussed the details of various selection criteria and the algorithm that forms the muon pairs in the data and the Monte-Carlo samples.

Chapter 7

Background Estimation

To this point, I have described signal production and modeling. The background modeling and estimation in the signal region is the deciding element of the analysis. In this chapter, I aim to associate appropriate backgrounds to the different regions of the parameter space. I probe a parameter space that extends two orders of magnitude (0.12 - 60 GeV). I use MC and CMS data to model the QCD and QED background. The dimuon invariant mass parameter space is divided into two major regions: the *low-mass* region, defined as below Υ resonance masses (i.e., 0.21 - 9 GeV) and the *high-mass* region where the masses above the Υ resonance are considered (i.e., 11 - 60 GeV). The low-mass region is dominated by QCD background, such as multi-jet processes, especially contributions from $b\bar{b}$ and double semi-leptonic decay or decay via resonances ($\eta, \omega, \phi, J/\psi(1S), \psi(2S)$). The high-mass region is dominated by the electroweak background, such as ZZ production. The production leads to four-leptonic decays in both cases, mimicking our signal. The QCD processes often have high production cross-sections, whereas the QED cross-sections are usually smaller. The higher cross-section is expressed in data with higher statistics. Consequently, we prefer to use the data-driven method for QCD modeling rather than MC generated. In other words,

here it is harder to match the statistics in data with MC-generated samples. Additionally, the MC simulation can be less precise for the non-perturbative nature of the lower-energy QCD processes. On the contrary, using MC samples for the QED background modeling proves to be more efficient as the cross-section is relatively smaller and the processes are better understood and simulated.

7.1 The Low-Mass Region

The low-mass region extends 0.21 - 9 GeV. Here, the dominating $b\bar{b}$ decays to muon pairs are via double semi-leptonic decays or via resonances such as $\eta, \omega, \phi, J/\psi(1S), \psi(2S)$. An example Feynman diagram of this process is illustrated in Fig. 7.1. Another

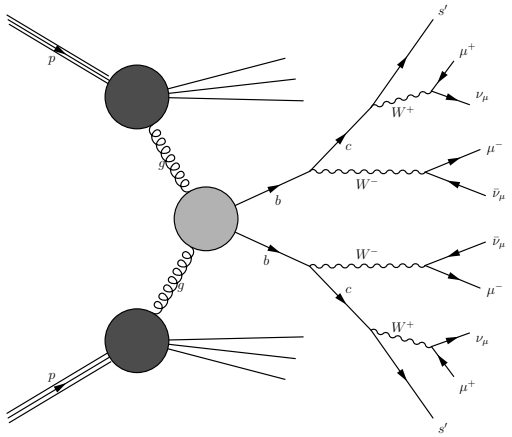


Figure 7.1: The Feynman diagram of two b quarks decaying into four muons via J/Ψ resonance is illustrated [10].

more minor contributing source of background in this region could be the pairing of a miss-identified charged track and a proper muon track from a $b\bar{b}$ semi-leptonic decay. The misidentification happens due to muon segments in the muon system occasionally wrongly attributing a muon ID to a nearby charged track. The probability of misidentification drops as the p_t rises [169]. We subject the background samples to the

same selection set as the signal samples. For instance, the dimuon isolation selection (Sec. 6.3) ensures that the muon pairs can come only from well-separated b jets.

A $b\bar{b}$ enriched control dataset is used for QCD modeling. As touched on in Sec. 6.3, 2018 `DoubleMuon` (Tab. 5.1), events with exactly one dimuon plus one or more orphan muons are saved during the selection process. The choice of an extra orphan muon is primarily set out by the triple muon triggers in Tab. 4.1. These events will now serve as the control data. If there is more than one orphan muon in the event, the orphan muon with the highest p_t will be selected. The basic selection for the selected events are: at least two *Loose* PF muons and at least one standalone muon, two muons with $p_t > 24$, $|\eta| < 2$, and three muon with $p_t > 8$, $\eta < 2.4$. The higher-level cuts are the same as Tab. 6.2.

The ultimate goal is to construct a dimuon versus dimuon invariant mass 2D background template. Assuming the b quarks decay independently into dimuons, the 2D background is constructed by the outer product of the fitted function of 1D dimuon invariant masses templates: $m_{\mu\mu_1}$ and $m_{\mu\mu_2}$. The $m_{\mu\mu_1}$ label is assigned to dimuons comprised of two high-p_t (> 24 GeV, $|\eta| < 2$) muons, while the rest of dimuons, with a high-p_t orphan muon, gain the $m_{\mu\mu_2}$ label. This procedure ensures that kinematic differences between signal events that have exactly two high-p_t dimuons or just one high-p_t dimuon are taken into account. The invariant mass distributions of the $\mu\mu_1$ and $\mu\mu_2$ are shown in Fig. 7.2.

Before introducing the fit parameters, we exclude the J/Ψ (from 2.72 GeV to 3.24 GeV) from the low-mass background modeling. The main reason is that if there is a signal in this region, it will be buried underneath the significant peak amplitude of J/Ψ. We utilize *the unbinned extended likelihood* method to fit these distributions, elaborated on in Ref. [170, 171]. By dividing the low-mass region into below and above

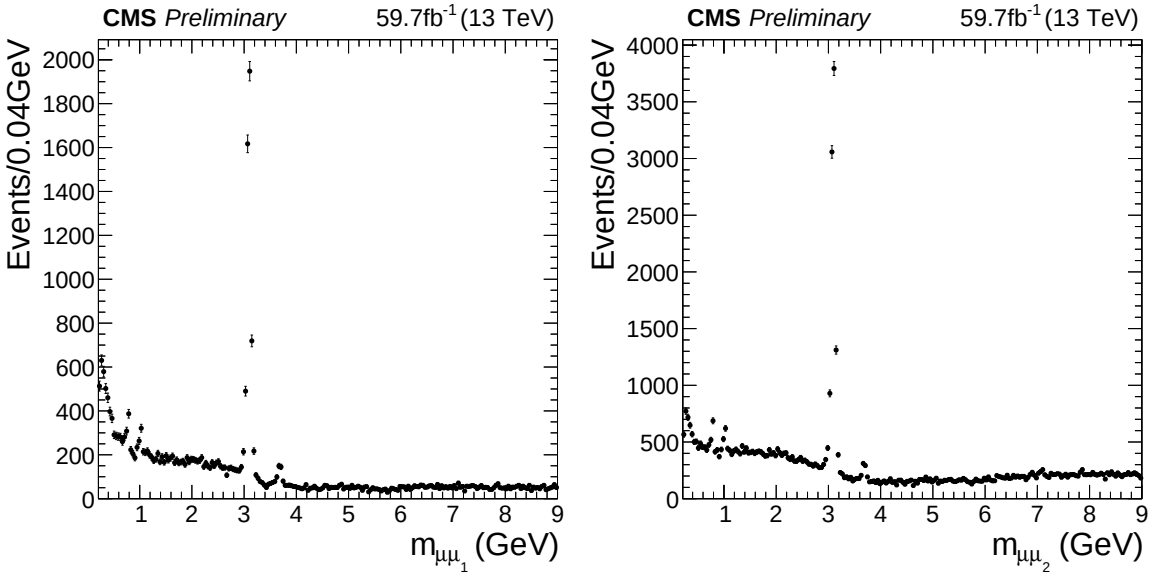


Figure 7.2: The invariant mass distribution in the low-mass control data is presented. Left shows the event distribution for $\mu\mu_1$ and right shows the event distribution for $\mu\mu_2$ [10].

the J/Ψ peak, the parametric fit functions are [10],

$$\begin{aligned}
&\text{Below J/}\psi\ f(m;p_\eta, p_\omega, p_\phi, p_B, p_{06}, ..., p_{66}) \\
&\quad = p_\eta G(m; m_{0,\eta}, \sigma_\eta) + p_\omega G(m; m_{0,\omega}, \sigma_\omega) + p_\phi G(m; m_{0,\phi}, \sigma_\phi) \\
&\quad + p_B B(m; p_{06}, ..., p_{66}) + p_{AH} G(m; m_{0,AH}, \sigma_{AH}) \\
&\quad + p_{E0}\left[\left(\frac{m}{m_{0,min}}\right)^2 - 1\right]^{p_{E1}} \times \exp\left[-p_{E2}\left[\left(\frac{m}{m_{0,min}}\right)^2 - 1\right]\right]
\end{aligned} \tag{7.1}$$

$$\begin{aligned}
&\text{Above J/}\psi\ f(m;p_\psi, p_B, p_{06}, ..., p_{66}) \\
&\quad = p_\psi G(m; m_{0,\psi}, \sigma_\psi) + p_B B(m; p_{06}, ..., p_{66})
\end{aligned} \tag{7.2}$$

where to each QCD resonances, η, ω, ϕ, and ψ, a Gaussian normalized to a unit integral is associated. The Gaussian function, $G(m; m_0, \sigma)$, is parametrized by m_0 which is the mass of the resonance, and σ is the resolution, fixed to the detector resolution at this mass. The mass (resolution) values are 0.548 (0.030) GeV for η, 0.782 (0.031) GeV for

ω, 1.019 (0.033) for ϕ, and finally 3.686 (σ_ψ) GeV for ψ, where σ_ψ is allowed to float in the fits. The $B(m; p_{06}, ..., p_{66})$ function, a series expansion in the Bernstein polynomial basis, describes the bulk shape in both regions. Only the normalization is allowed to vary for the above SM resonances, except for the ψ resonance, where the width is also free to vary. More can be found in our CMS analysis note AN-19-153 [108]. The fitted data, divided into below J/Ψ and above J/Ψ resonance, are shown in Fig. 7.3.

By extracting the fits for $f(m_{\mu\mu_1})$ and $f(m_{\mu\mu_2})$, we construct the 2D templates $f(m_{\mu\mu_1}) \otimes f(m_{\mu\mu_2})$, exhibited in Fig. 7.4.

7.2 Low-Mass Region Background Estimation

To estimate the number of background events in the signal region (SR), defined in Sec. 6.3 and illustrated in Fig. 6.3, events from 2018 data in the control region (CR) are overlaid on the 2D template constructed in this section. The SR remains *blinded* for now to avoid bias. The integral of the 2D function in the SR and CR is calculated. In Fig. 7.5, data events are represented by white circles. Using the number of events in the CR, we can estimate the number of expected events in the SR by using the integral ratio in SR and CR. The results are shown in Tab. 7.1.

Table 7.1: Estimated number of events in the SR in the low-mass region

Region	Integral ratio ($\frac{SR}{CR}$)	Events in CR	Expected events in SR	Stat. Uncrt.
Below J/Ψ	0.043/0.969	98	4.34	0.44
Above J/Ψ	0.035/0.965	66	6.16	0.76

7.3 The High-Mass Region

We designate the high-mass region to the masses above Υ resonance (11 - 60 GeV). Here, the electroweak processes dominate the background composition. The predomi-

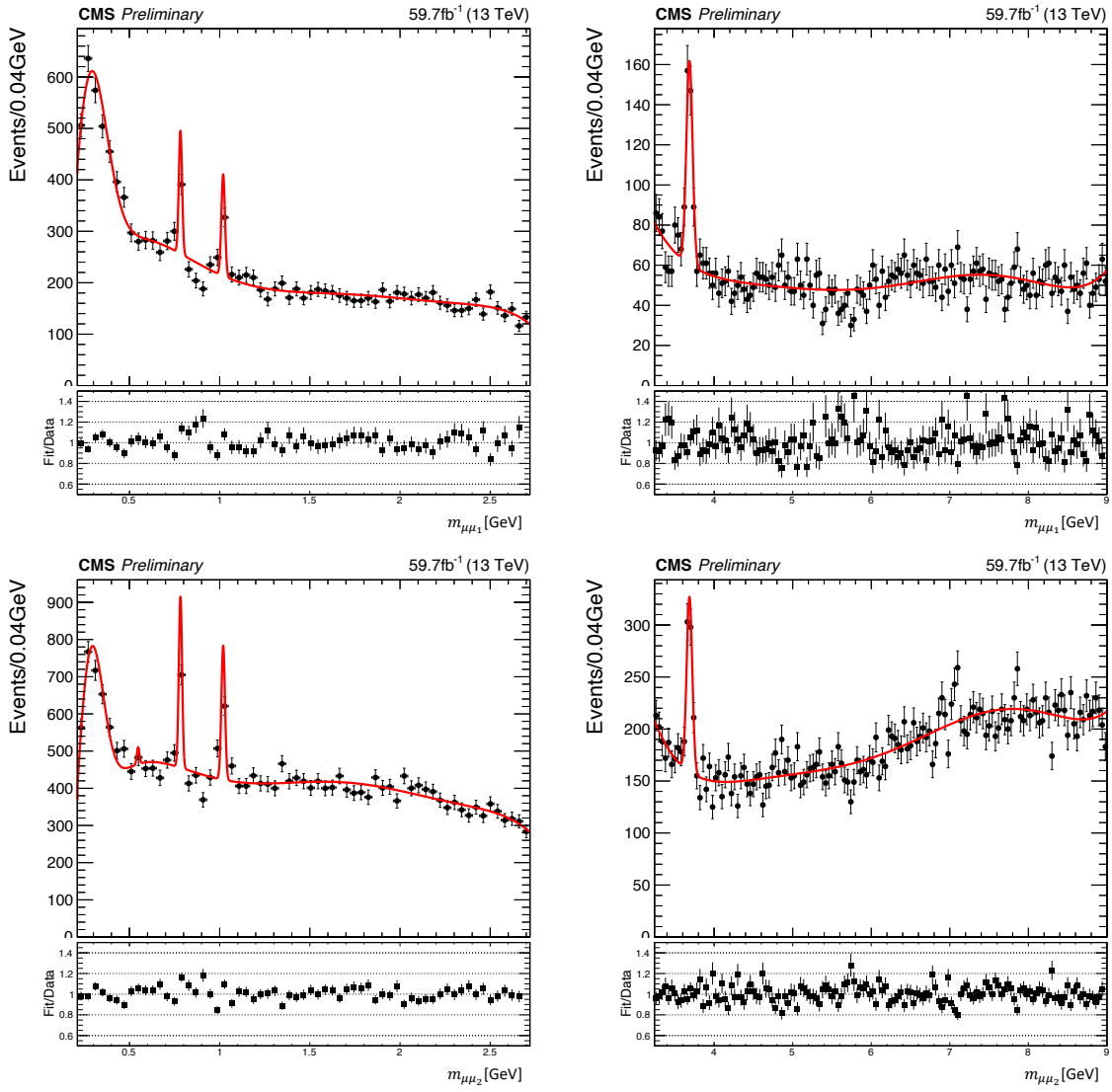


Figure 7.3: Fits to 2018 control data: $m_{\mu\mu_1}$ and fitted function below J/ψ, $\chi^2/\text{ndf} = 1.49$ (top left); $m_{\mu\mu 1}$ and fitted function above J/ψ, $\chi^2/\text{ndf} = 1.18$ (top right); $m_{\mu\mu 2}$ and fitted function below J/ψ, $\chi^2/\text{ndf} = 1.72$ (bottom left); $m_{\mu\mu 2}$ and fitted function above J/ψ, $\chi^2/\text{ndf} = 1.33$ (bottom right).

nant contribution comes from ZZ production and the subsequent decays to a four-lepton final state. A less significant contribution comes from heavy QCD $t\bar{t}$ decays. Normally, in a four-lepton analysis, the Drell-Yan (DY) process does not contribute to the background as the Z/γ mediator decays into two leptons. However, it is revealed that there is a scenario in which the muons can radiate a highly energetic photon that can convert

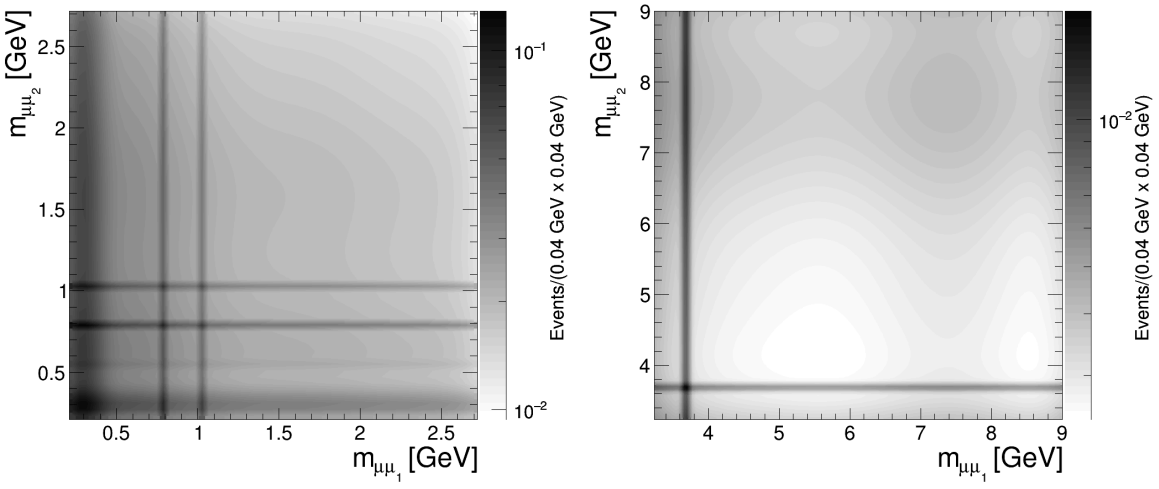


Figure 7.4: Invariant mass distribution in the control data for the QCD background determination is shown. Left shows the event distribution for below J/ψ and right shows event distribution for above J/ψ.

into another pair of muons. The muon products of the QED radiated photons can pair with a DY muons and mimic our signal. The Feynman diagram of Fig. 7.6 illustrates such process. Such events often fail the mass consistency requirement and reside in the CR. Since it is troublesome to determine the distribution shape of said events due to the limited statistics, it is crucial to veto them. To do so, we form alternative pairs, meaning once the selection algorithm forms dimuons, we pair the muons in each pair with the oppositely charged muons in the other pair. Once the alternative pairs are formed, the pair with the higher invariant mass is labeled as *leading* dimuon, while the dimuons with lower invariant mass are called the *trialing* dimuon. Invariant mass and ΔR cuts are then imposed on the alternative pairs, and the events which fail these cuts are vetoed. By comparing signal efficiencies of various cut thresholds on these variables, a final requirement of a trailing dimuon invariant mass larger than 3 GeV, or a trailing dimuon with ΔR larger than 0.2, is used to veto these background events. This cut has a negligible effect on signals.

The MC samples are run through the analysis selection, and the ones with surviving

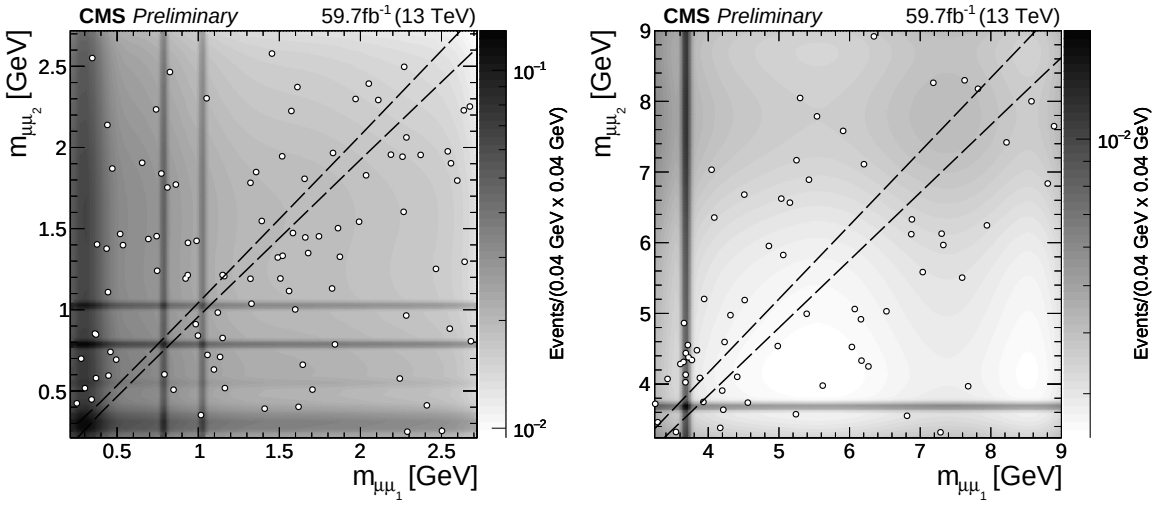


Figure 7.5: 2D QCD background templates and measured data points (white circles) for 2018 below (left) and above (right) the J/ψ. The darker vertical and horizontal lines are the small mass resonances shown in the 1D templates in Fig. 7.3. When multiplied into the 2D template, they manifest as darker lines. The mass window for the signal region is indicated. Events in this region are blinded at this stage of the analysis.

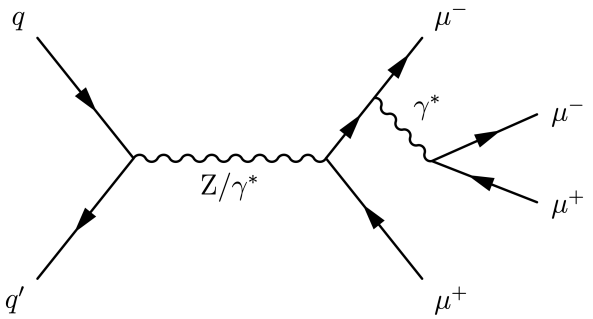


Figure 7.6: Feynman diagram of a DY process, in which a high-energy photon is radiated from the produced muon, which in turn can convert into a pair of muons [10].

events are used for background estimation. The MC events are scaled to the luminosity of the 2018 `DoubleMuon` data (59.9 fb^{-1}) based on their cross-section. The list of MC samples used for the high-mass background region is listed in Tab. 5.3. We practically compare the MC-produced stacked background distribution to the `DoubleMuon` data distribution and verify that MC adequately models the data in the CR. The verification in the CR tells us that the background modeling can be trusted in the SR as well. The distribution of MC and data in CR is shown in Fig. 7.7.

The agreement between data and scaled MC events is parameterized by calculating the data-to-MC ratio: 1.05 ± 0.12 (Tab. 7.2). We accept this figure as a good agreement between data and MC, and we proceed to estimate the background in SR.

7.4 High-Mass Region Background Estimation

With the assurance that the data and MC agree reasonably well in the CR, we can now use the same samples to estimate the background in the SR. In other words, when the MC background is subtracted from data in the SR, what remains is the potential signal for new physics and the basis for setting limits. The distribution of MC background is shown in Fig. 7.8.

Based on the MC-to-data ratio in the CR (1.05 ± 0.12) and having the number of MC entries in SR (11.67 ± 1.37), we can estimate the expected number of data entries due to background in the SR. Tab. 7.2 has the details of this estimate for the high-mass region.

Table 7.2: Estimated background events in the high-mass region

Region	MC entries	Data entries	Data/MC	Est. events in SR
CR	136 ± 11 (stat.)	143 ± 12 (stat.)	1.05 ± 0.12 (stat.)	N/A
SR	11.67 ± 1.37 (stat.)	SR blinded	N/A	12.28 ± 2.01 (stat.)

7.5 Concluding Remarks

In this chapter, I presented the background of this analysis. Understanding what processes contribute to the background is essential, as the shape of the background in the signal region contributes to the uncertainties. By reexamining the estimated background in both low-mass and high-mass regions, we notice that number of background events will not rise to statistically significant values to constitute the background for

the analysis. As the results of this observation, we declare this analysis as a close to "zero-background" analysis over the entirety of the probed parameter space. This practice, nevertheless, is necessary to develop an understanding of the possible background shapes.

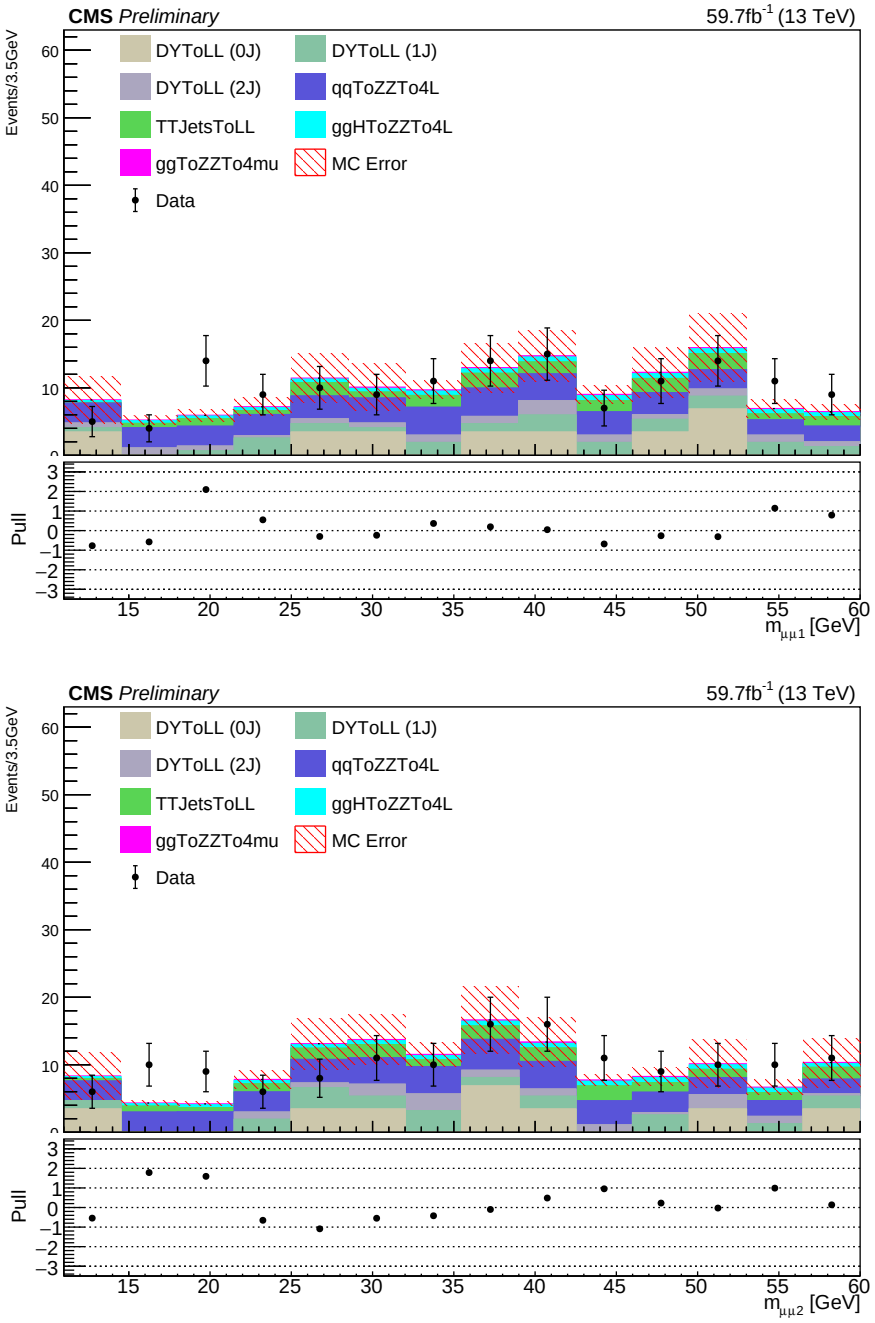


Figure 7.7: The distributions of simulated background versus the `DoubleMuon` dataset in the high-mass region in the control region [10].

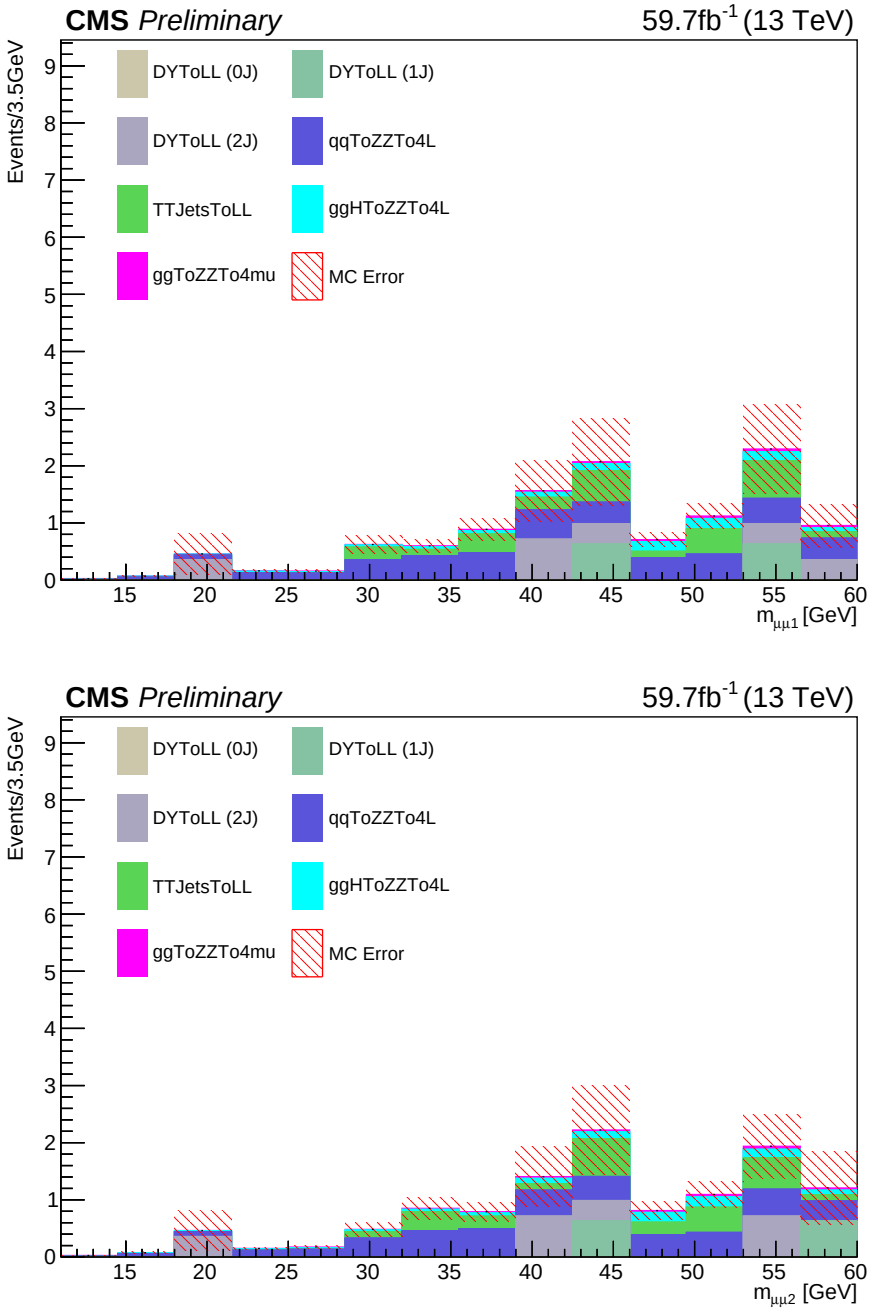


Figure 7.8: The distribution of simulated background in the high-mass region in the signal region is displayed. The signal region remains blinded at this stage of the analysis [10].

Chapter 8

Uncertainties and Results

In this chapter, having established the anticipated backgrounds, I move to unblind the SR and set limits on the vector-portal model parameters. The model-independent 95% confidence level (CL) upper limit on the product of cross-section times branching fraction squared times kinematic and geometric acceptance of the analysis at generator level is set based on the observed events across the parameter space. Thereafter, I interpret the results in the context of the vector-portal model. However, before limit setting, it is crucial to quantify the statistical and systematic uncertainties present in the experimental, theoretical, and background modeling sources in the analysis.

8.1 Uncertainty

The experimental uncertainty emerges from data taking, selection, and reconstruction processes. All uncertainties related to MC simulation are considered theoretical uncertainty. Lastly, I take the background normalization uncertainty into account.

8.1.1 Experimental Uncertainty

The experimental systematic uncertainties are primarily due to CMS measurement, such as the luminosity measurement for each run era, stored in the so-called "Lumi blocks," or offline reconstruction and selection procedures, e.g., trigger and muon ID efficiencies.

Luminosity: The luminosity at CMS is measured by the Hadronic Forward (HF) calorimeter. The absolute luminosity calibration is performed by Van der Meer (VdM) scans [172], involving scanning the LHC beams through one another to determine the size of the beams at their point of collision. I use a luminosity measurement uncertainty of 2.5%, provided centrally by CMS. A detailed description of luminosity uncertainty sources can be found in Ref. [172].

Pile-Up Distribution: As introduced in Sec. 5.2.1, the PU effect is due to more than one proton-proton collision taking place during each BX. The luminosity per-BX information with the total proton-proton inelastic nominal cross-section (σ_{nom}) is used to calculate the expected number of PU events per BX (PU = $\mathcal{L}_{BX} \times \sigma_{nom}$). The recommended σ_{nom} is 69.2 mb for 2018. The PU distribution per BX is derived for both MC and 2018 data, the ratio of which is used to *reweight* the MC event. To determine the systematic uncertainty, the cross-section is shifted 5% up and down from the nominal cross-section ($\sigma(pp)_{nom_{5\%}} = 65.8$ mb, $\sigma(pp)_{nom_{+5\%}} = 72.3$ mb) and the corresponding PU distributions are extracted [173]. The effect on the ε/α calculation is considered for three cases: 5% up, nominal, and 5% down. The greatest absolute deviation from the nominal value of ε/α is taken as the systematic uncertainty. We report systematic uncertainty due to the PU effect being 0.05%.

PU Effect: The PU affects some selection items, such as dimuon isolation, performances. To quantify the PU effect on the signal, we divide the PU range into three categories: low (PU $<$ 25), medium (25 $<$ PU $<$ 45), and high (PU $>$ 45). The variation of ε/α across these regions is studied, and a systematic uncertainty of 1.8% is adopted.

Selection: The uncertainty on Muon ID, provided by the Muon POG [174], is estimated at 0.6% per muon [175]. The HLT muon uncertainty is calculated by using the orthogonal method discussed in Sec. 9.5. The overall data-to-MC trigger scale factor is estimated to be 99.6% $\pm$ 0.6 (stat.) for the 2018 dataset. The uncertainty on the muon isolation is driven by underlying event (UE). The data-to-MC dimuon isolation uncertainty is calculated by the Tag-and-probe method described in [5] and estimated at 0.1% per dimuon. The mass-window cut assumes an uncertainty on average ε/α across all MC samples (0.418 $\pm$ 0.001) and is found to be 0.24%.

Reconstruction: There is systematic uncertainty associated with the reconstruction of muons near each other (close muons) in the muon system. We use the results reported in Ref. [107]. This uncertainty only applies to signal invariant mass points below 9 GeV. We use 2.6% per event. Similarly, close muons in the muon system uncertainty, there is an uncertainty associated with close muons to the tracker, also applied to mass points below 9 GeV. The corresponding systematic uncertainty was estimated as 1.2% per dimuon for these events.

8.1.2 Theoretical Uncertainty

Theoretical systematic uncertainty on the signal acceptance comes from uncertainties in PDFs, determining the strong coupling constant, α_s, and uncertainties evaluated by

varying QCD renormalization and factorization scales (μ_R and μ_F). The uncertainties related to the PDFs and the knowledge of the strong coupling constant α_s are calculated in Ref. [176]. We calculate the PDF uncertainty by following the PDF4LHC recommendations [177]. To this end, the QCD scale uncertainty is found by using the Monte Carlo for FeMtobarn processes tool (MCFM) [164]. Furthermore, the uncertainty due to the PDF choice is found by varying parametrization within the PDF, investigating alternative PDFs, and comparing the central values. When treated as uncorrelated and added in quadrature, these uncertainties are found to be 8% in total.

8.1.3 Background Uncertainty

Normalization uncertainties: For the region below the J/ψ, the nominal number of expected background events is 4.34 ± 0.44, which corresponds to a normalization uncertainty of 10.1%. For the region between the J/ψ and the Υ resonances, the nominal number of expected background events is 6.16 ± 0.76, which corresponds to a normalization uncertainty of 12.3%. For the region above the Υ, the nominal number of expected background events is 12.28 ± 2.01, which corresponds to a normalization uncertainty of 16.4%.

Background systematic uncertainties: The total number of expected backgrounds in the SR is calculated for a 5% up and 5% down deviation from the nominal 2.3 GeV isolation cut, used in Sec. 6.3. The background systematic uncertainties for below J/Ψ, above J/Ψ and high-mass region are 4.1%, 1.5%, and 2.3%, respectively.

High-mass background shape uncertainty: The shape of the background and its uncertainty in the high-mass region is extracted from a one-dimensional kernel density estimation PDF. The width of the Gaussian is adaptively calculated from the local

density of events, i.e., narrow for regions with high event density to preserve details and wide for regions with low event density to promote smoothness [178]. The resulting PDF models the distribution of the MC sample dataset as a superposition of Gaussian kernels, one for each data point, each contributing 1/N to the total integral of the PDF [179]. The kernel density estimation is adopted to obtain a smooth background shape based on a small number of estimated background events in the signal region. The kernel density estimation PDF for each of the MC processes is obtained and then summed up using the same weights as in Fig. 7.8. The weight is a multiplication of the MC-data scale factor and the number of MC events that pass all selection. The obtained smooth background PDF for $m_{\mu\mu_1}$ and $m_{\mu\mu_2}$ are shown as the blue curve in Fig. 8.1.

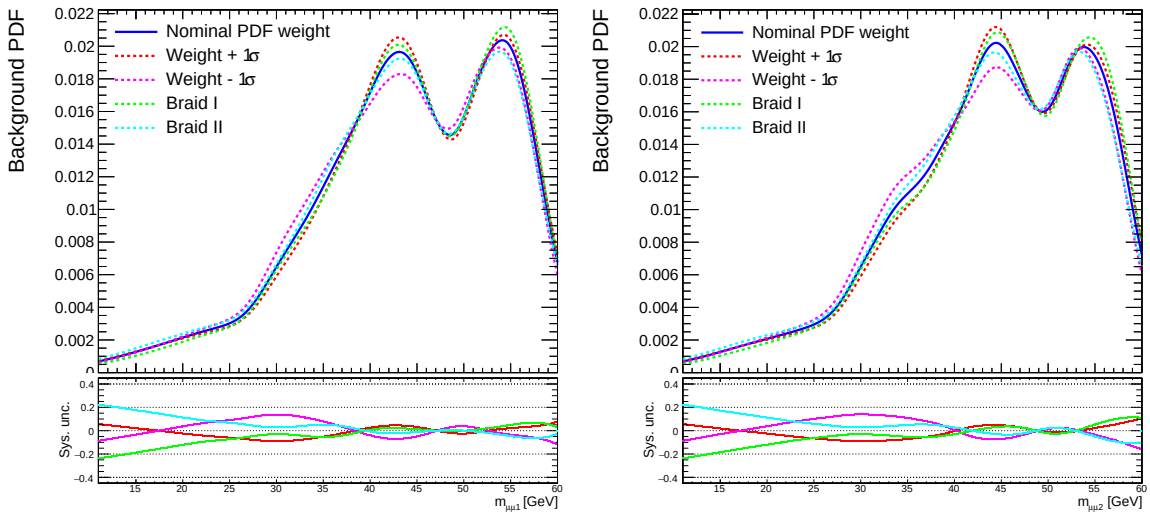


Figure 8.1: Kernel density estimation PDF, normalized to one, for $m_{\mu\mu_1}$ (left) and $m_{\mu\mu_2}$ (right) in the signal region. The PDF is plotted in the RooPlot framework and the variables $m_{\mu\mu_1}$ and $m_{\mu\mu_2}$ each have 98 bins (bin size is 0.5 GeV). The green dashed line (Braid I) use nominal MC weight + or − 1σ alternatively for DY+1J, DY+2J, ZZ→4L, TTJets→LL, ggH→ZZ→4L, and gg→ZZ→4mu processes, starting from nominal MC weight + 1σ for DY+1J. The dashed cyan line (Braid II) is similar to Braid I, however the alternation starts from nominal MC weight − 1σ for DY+1J. [10].

To evaluate the background shape uncertainty, the weight for each MC sample is

varied within the statistical uncertainty, and several other smooth shapes are obtained, as shown in Fig. 8.1.

8.1.4 Summary

Tab. 8.1 summarizes uncertainties that I use for setting upper limits on the signal search in the analysis. Systematic uncertainty and shape uncertainty is computed by Wei Shi, details of which are available in his dissertation Ref. [10], as well as our CMS analysis note AN-19-153 [108].

8.2 Results

We have arrived at the moment of truth in this section. By unblinding the signal region, there is a world in which I could claim evidence of the existence of DM, nay, the *discovery* of DM, on the condition that I observe a statistically significant signal. The less dramatic case, in case of lack of such a signal, would be setting a 95% CL upper limit on the signal. After I impose full selection on the 2018 data, thirty events in total survive the selection in the SR within the diagonal mass window. The observed events in the invariant mass space are listed in Tab. 8.2.

More particularly, for the low-mass region (below 9 GeV), I detect 4 and 6 events below and above the J/ψ, respectively. The observed number of events matches our expected number of background events in this region, i.e., 4.34 and 6.16 events, respectively. The unblinded data are shown in Fig. 8.2.

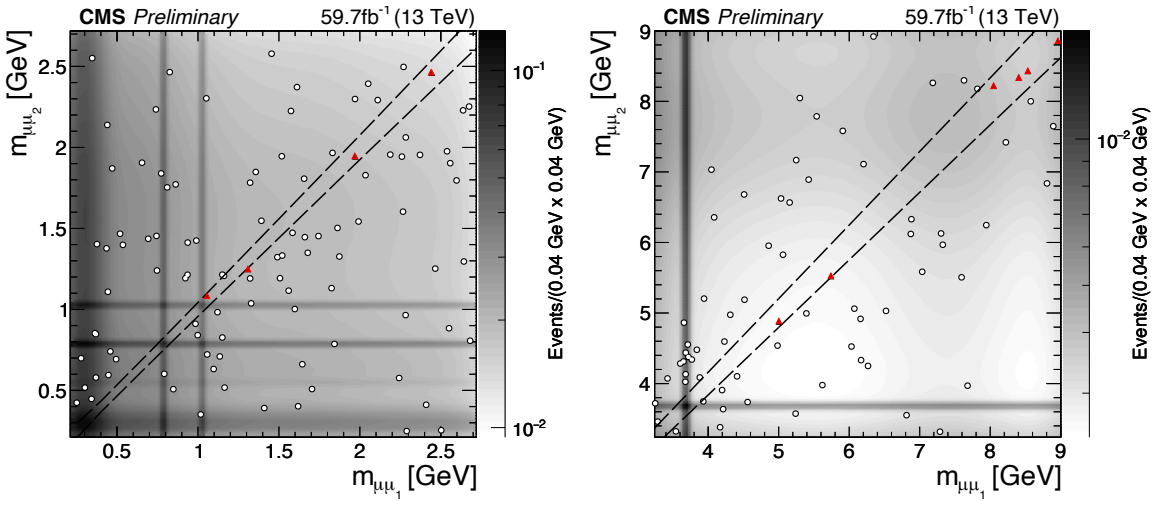


Figure 8.2: Two-dimensional distribution of the invariant masses below (left) and above (right) the J/ψ resonance. The grayscale heat maps show the normalized QCD background templates below the Υ resonances. White circles represent data events that pass all selection criteria but fall outside the SR (outlined by dashed lines). Red triangles represent data events passing all selection criteria, including the mass window cut.

We observe twenty events in the high-mass region, as shown in Fig. 8.3. The expected number of background events for this region is 12.28 $\pm$ 2.01 (stat.) $\pm$ 2.94 (sys.). The systematic error of 2.94 already includes the background shape systematic uncertainty of 23.8% obtained in Sec. 8.1.3. The statistical "pull" quantifies how much each observed variable is *pulled* away from its expected values and is defined as

$$pull(\theta) = \frac{\theta - \theta_e}{\sigma_\theta}, \tag{8.1}$$

where, θ is the observed parameter, θ_e is the estimated parameter, and σ_θ is the standard deviation. The invariant mass pulls in this region are within $2\sigma_\theta$ (only statistical errors considered) when they are plotted in the 1D histograms, as shown in Fig. 8.4.

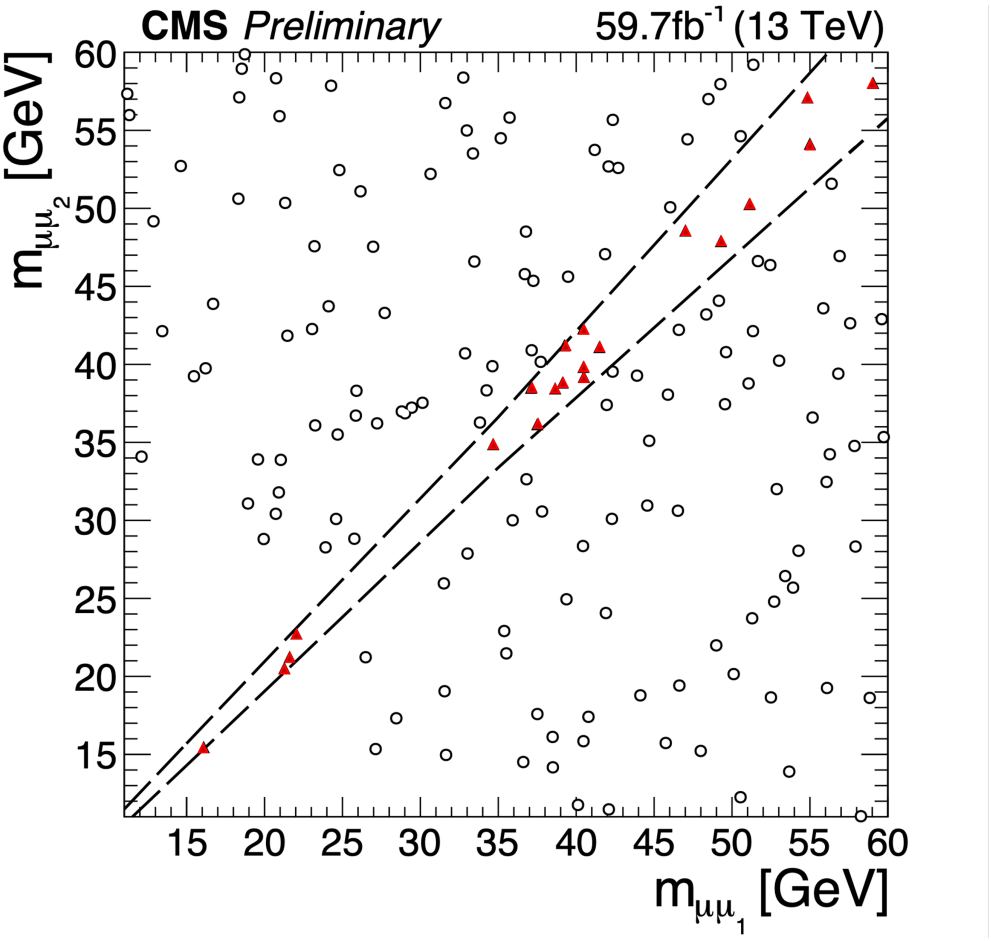


Figure 8.3: Two-dimensional distribution of the invariant masses above the Υ resonances. White circles represent data events that pass all selection criteria but fall outside the SR (outlined by dashed lines). Red triangles represent data events passing all selection criteria.

8.2.1 Model-Independent Limits

While the two-sigma excesses near 20 GeV and 40 GeV are intriguing, they are not statistically significant enough to claim presence of a signal. Consequently, I move to setting limits using the observed events. In Sec. 3.5, I lay out the argument for conducting the analysis within a model-independent framework. In this section, I present the methodology we use to set the model-independent limits. The results presented in this section can be used by any new physics model predicting the four-

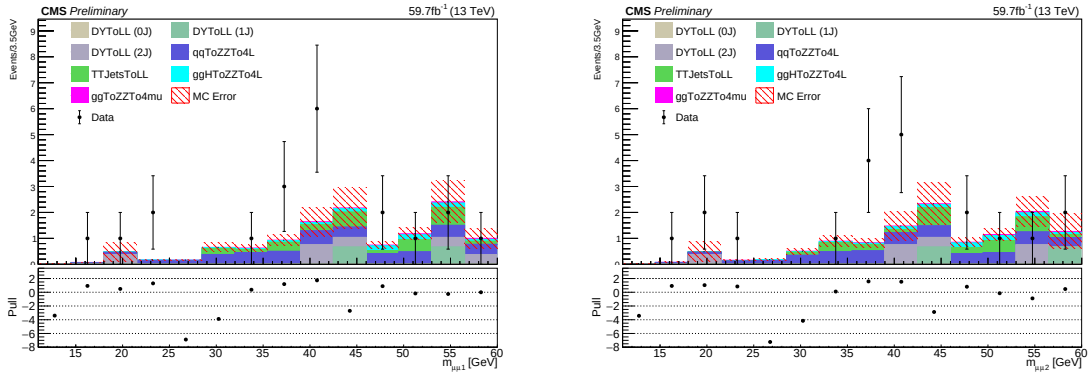


Figure 8.4: MC and data distributions for $m_{\mu\mu_1}$ (left) and $m_{\mu\mu_2}$ (right) in the high-mass signal region (above 11 GeV) after all analysis selection cuts are applied.

muon signature to set a 95% confidence level (CL) upper limit on the product of cross-section times branching fraction squared times kinematic and geometric acceptance of the analysis at generator level,

$$\sigma(pp \to 2a + X) \times \mathcal{B}^2(a \to 2\mu) \times \alpha_{\text{gen}} \leq \frac{N_{\text{evt}}}{\mathcal{L} \times r}, \tag{8.2}$$

where N_{evt} is the 95% CL upper limit on the number of observed events. The calculation uses the integrated luminosity $\mathcal{L} = 59.7$ fb^{-1}, as measured by CMS for the 2018 data-taking era, and takes the ratio, r, which is given by,

$$r = SF_{\varepsilon_{Full}} \times \varepsilon^{\text{MC}}_{\text{Full}}/\alpha_{\text{gen}}, \tag{8.3}$$

where the scale factor, $SF_{\epsilon_{Full}} = 0.996$, corrects for the experimental effects overlooked by the simulation. It is the multiplication of the muon ID, isolation, and trigger scale factors, summarized in Sec. 8.1.4. The other term, $\varepsilon^{\text{MC}}_{\text{Full}}/\alpha_{\text{gen}} = 0.418$, is the ratio of the full selection efficiency over the kinematic and geometric acceptance of the analysis at generator level averaged over all of the benchmark points as mentioned in the summary of Sec. 6.4.

The model-independent limits are set within the frequentist paradigm, based on the profile likelihood ratio as the test statistic. The profile likelihood ratio is used to determine how signal-like or background-like the data are[180, 181]. The systematic uncertainties are treated as nuisance parameters (parameter of the PDF). The ATLAS and CMS Collaborations provide the statistical tool to implement the test statistic via the LHC Higgs Combination Group, commonly referred to as "The Higgs Combine Tool." The details about the operational description of the Higgs combine tool are given in Refs. [182, 183, 184].

The determination of the MC signal and background shapes is necessary at this stage, as the limit computation is based on the *combination* of signal and background bins and channels. The signal shape is extracted from a 2D shape, constructed by multiplying two identical 1D *Crystal Ball* functions from ROOT `RooCBShape`. The Crystal Ball function is a PDF with common use in high-energy physics for modeling reconstructed object distributions such as reconstructed invariant mass, where some fraction of the energies and momenta may be lost to detection [185]. For each mass point in the model, we fit a Crystal Ball function to the average dimuon invariant mass, as done in the previous iteration of this analysis [5]. The background shape below 9 GeV is described in Sec. 7.1 and the background shape above 11 GeV is described in Sec. 8.1.3.

From here, the number of expected events is extracted from the signal and background shapes, and the test statistic (profile likelihood ratio) is calculated. Based on the test statistic result, toy experiments for various masses of the new boson are generated. The combine tool then outputs the 95% CL upper limits on N_{evt} for the dimuon mass points. Lastly, the obtained N_{evt} is used to derive the 95% CL upper limit on $\sigma(pp \to 2\mathrm{a} + \mathrm{X}) \times \mathcal{B}^2(\mathrm{a} \to 2\mu) \times \alpha_{\mathrm{gen}}$ as shown in Fig. 8.5.

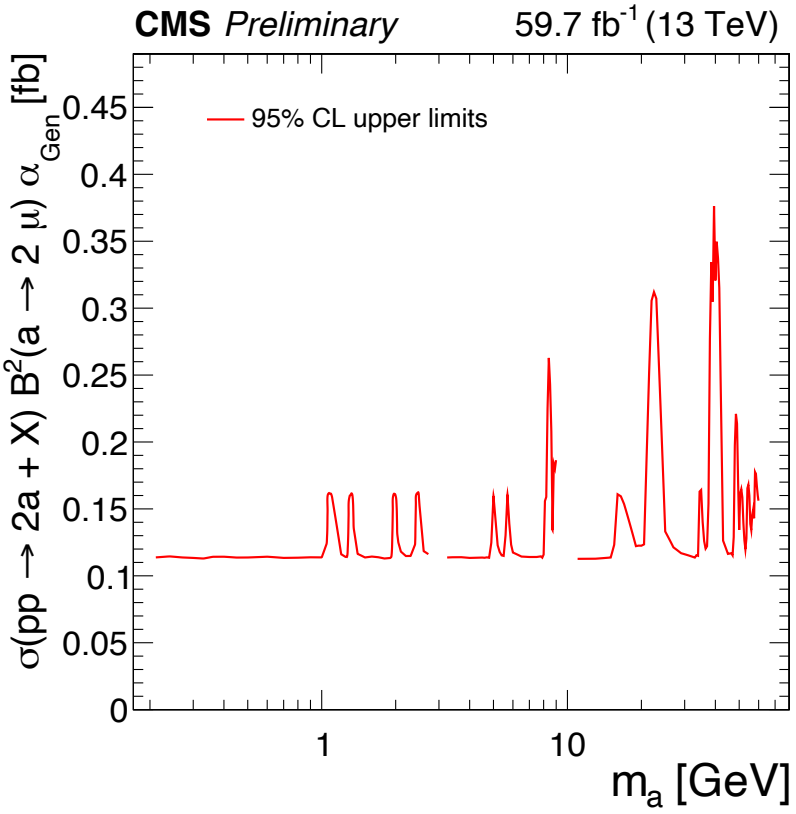


Figure 8.5: The model-independent 95% CL upper limit on the product of the cross section times branching fraction squared times acceptance at the generator level. The gaps in the line correspond to the regions near the J/ψ and Υ mass that are excluded from the analysis.

8.2.2 Limit Interpretation in the Vector-Portal Model

Given the model-independent results (N_{evts}), I use Eq. 8.2 to set a 95% CL limit on $\sigma(\mathrm{pp} \to \mathrm{Z_D})\mathcal{B}(\mathrm{Z_D} \to \mathrm{s_D}\overline{\mathrm{s_D}})\mathcal{B}^2(\mathrm{s_D} \to \mu^+\mu^-)$, the product of production cross section of the dark vector boson Z_D, branching fraction of Z_D decaying to a pair of dark scalar bosons s_D, and the squared branching fraction of s_D decaying to two muons as illustrated in Fig. 8.6. The limit curves exhibit a structure with an increase and a dip as the s_D mass approaches the kinematic limit of $m_{Z_D}/2$. These curves smoothen out as the Z_D mass increases. This behavior is in agreement with the behavior of the kinematic acceptance (see Fig. 3.7). An inspection of the p_T spectra of the four muons

in the simulated data shows that the dip in the kinematic acceptance is due to the specific shape of the p_T distributions of the fourth selected muon as a function of s_D mass as seen in Fig. 5.1.

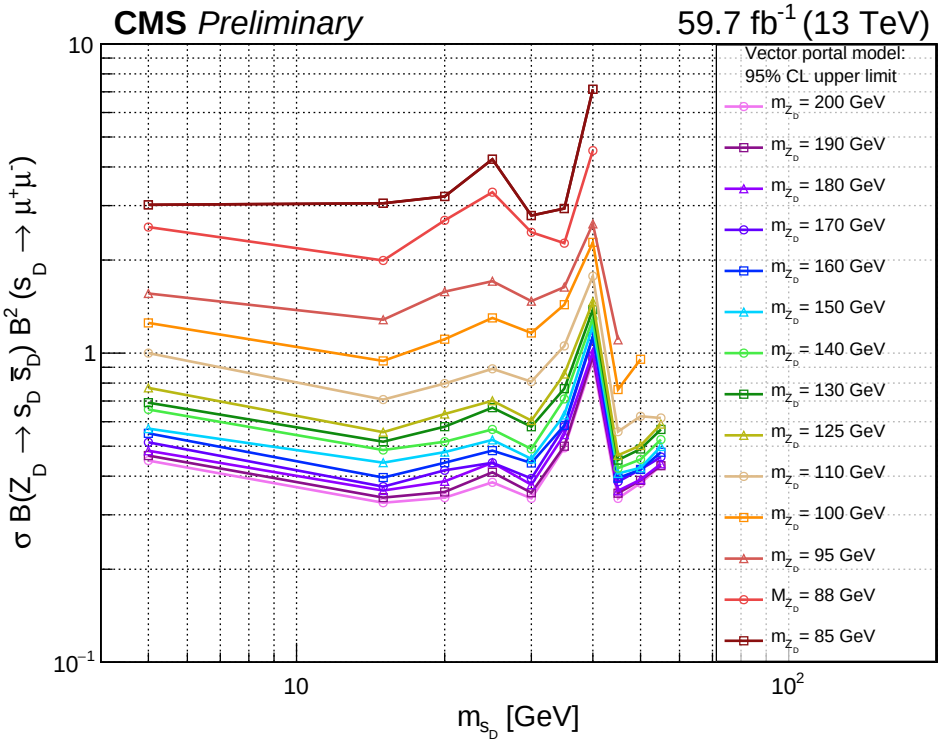


Figure 8.6: The observed 95% CL upper limits on $\sigma(\mathrm{pp} \to \mathrm{Z_D})\mathcal{B}(\mathrm{Z_D} \to \mathrm{s_D}\overline{\mathrm{s_D}})\mathcal{B}^2(\mathrm{s_D} \to \mu^+\mu^-)$ as a function of the dark scalar mass $m_{\mathrm{s_D}}$ and the dark vector boson mass $m_{\mathrm{Z_D}}$.

After setting limits on the cross-section, I translate these limits to a limit on $\varepsilon^2\mathcal{B}(\mathrm{Z_D} \to \mathrm{s_D}\overline{\mathrm{s_D}})\mathcal{B}^2(\mathrm{s_D} \to \mu^+\mu^-)$ as a function of the dark scalar mass $m_{\mathrm{s_D}}$ and the dark vector boson mass $m_{\mathrm{Z_D}}$. To do so, for each mass of Z_D I have scanned $\sigma(\mathrm{pp} \to \mathrm{Z_D})$ with four kinetic mixing parameters ($\varepsilon = 0.01, 0.001, 0.0001, 0.00001$). My analysis confirms a linear behavior of production cross-section with ε^2, for each mass of Z_D. By fitting each curve with a quadratic function ($\sigma = a \times \varepsilon^2$), I derive the values of the scale factor a. An example of such calculation is shown in Fig. 8.7. The scale factors are plotted against masses of Z_D in Fig. 8.8. The translated plot is shown in Fig. 8.9. Another way

of visualizing post fit observed limits is through the so-called *Brazilian band* plots. As an example, I show the Brazilian bands for a vector-portal sample in Fig. 8.10, where I compare the expected limits against the observed limits and visualize the one and two standard deviations around the expected limits.

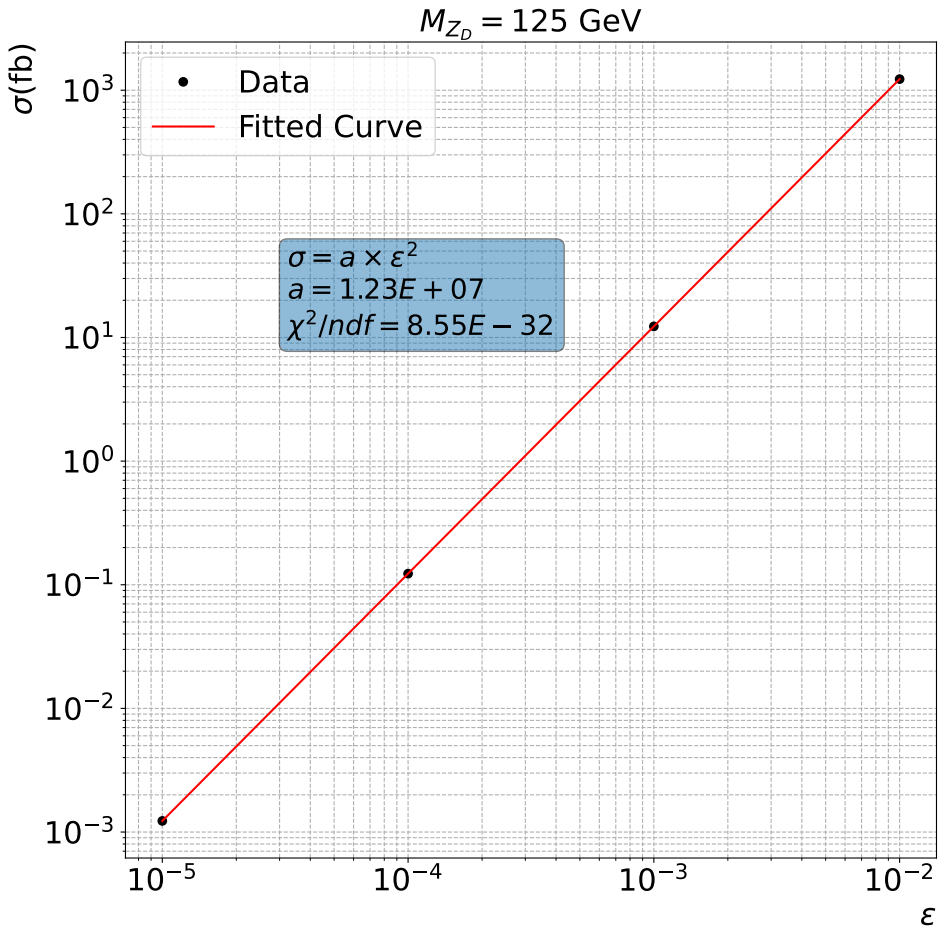


Figure 8.7: The cross-section is shown as a function of the kinetic mixing parameter ε, for $m_{Z_D} = 125$ GeV is shown. The resulting curve is fitted by a square function, and the constant of the function a is extracted. This constant serves as a scale factor to convert cross-section to kinetic mixing parameter ε.

Considering the relatively small variation of the limit values over the masses of s_D particles for each fixed mass of the Z_D boson, I can consider the average limits for s_D masses for each mass of Z_D and arrive at a singular point for each case. Fig. 8.11 shows

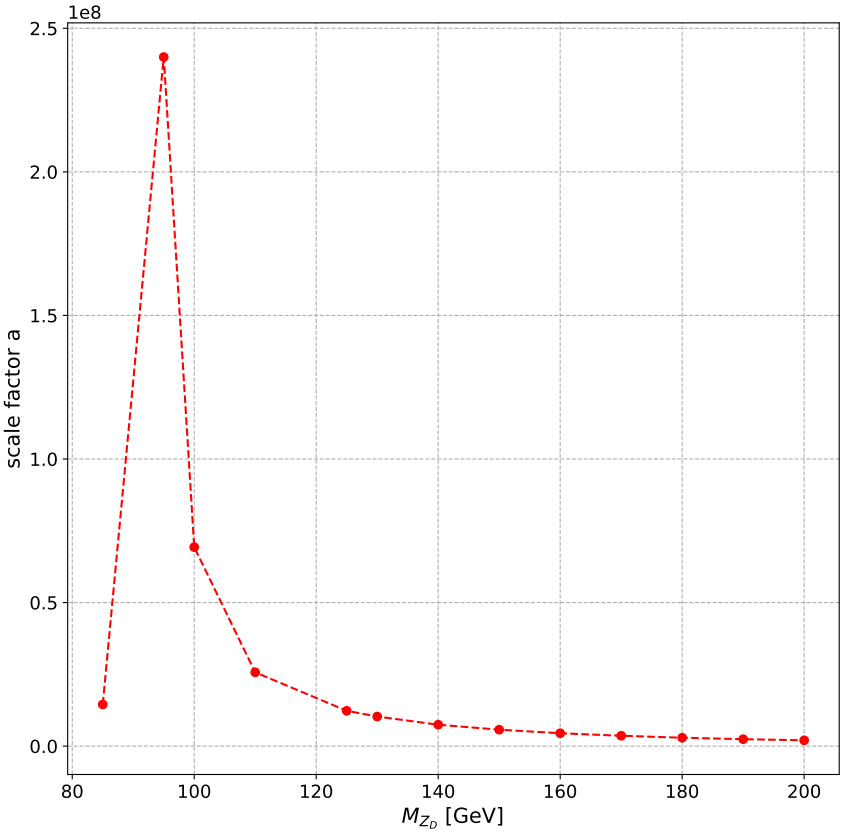


Figure 8.8: The behavior of the derived scale factors (a) in the $\sigma = a \times \varepsilon^2$ fits, with Z_D mass.

the excluded region under this simplification.

8.3 Concluding Remarks

After applying the full selection to data samples in this chapter, we found 30 events in the signal region. These events are shown within the diagonal mass window in Fig. 8.2 and Fig. 8.3. In particular, in the low-mass region, below J/Ψ resonance (9 GeV) masses and above J/Ψ resonance masses (9-11 GeV), we observed 4 and 6 events,

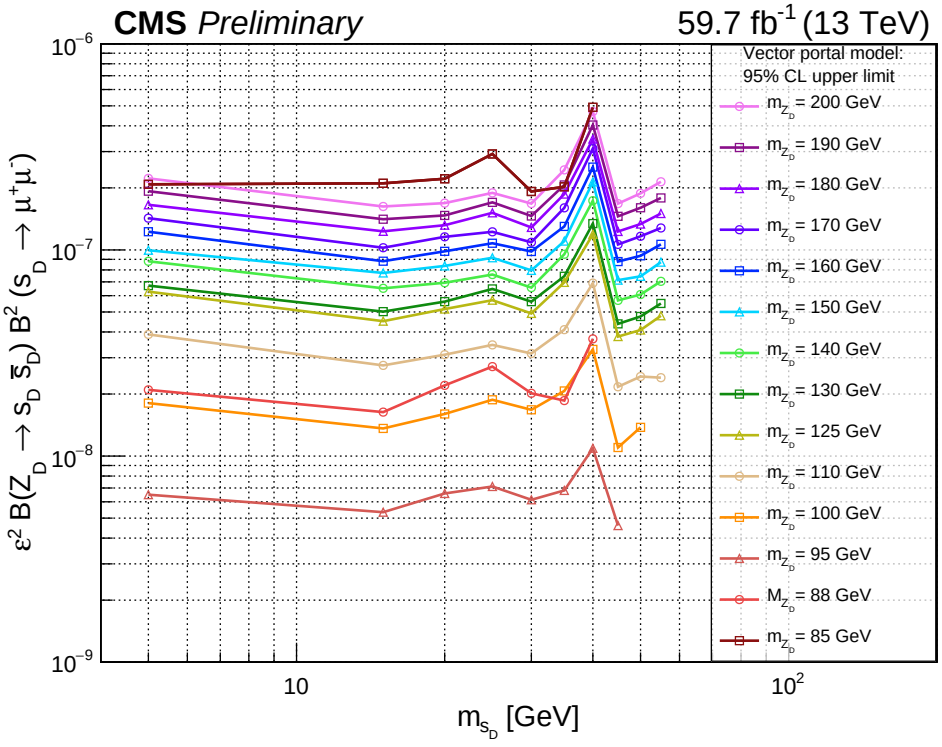


Figure 8.9: The 95% CL upper limits on $\varepsilon^2\mathcal{B}(\mathrm{Z_D} \to \mathrm{s_D}\overline{\mathrm{s_D}})\mathcal{B}^2(\mathrm{s_D} \to \mu^+\mu^-)$ as a function of the dark scalar mass $m_{\mathrm{s_D}}$ and the dark vector boson mass $m_{\mathrm{Z_D}}$.

respectively. The unblinded results in the low-mass region are statistically consistent with the predicted SM background. In the high mass region, above Υ resonance masses (11-60 GeV), we observed 20 events. While the observed number of events in the high-mass region can be considered consistent with the predicted background of 12.28 ± 2.01 (stat.) $\pm$ 2.94 (sys.) events, the pulls are within 2σ with only statistical error taken into consideration. Specifically, two bumps are observed around 22.5 GeV and 38.5 GeV in the Brazilian bands of Fig. 8.10. To better understand this apparent discrepancy, we launched the data analysis of the 2017 data recorded by CMS. Combining the results of 2017 with the 2018 results can provide better background modeling and better statistics in this region. In the upcoming chapter, I discuss the progress of the 2017 analysis.

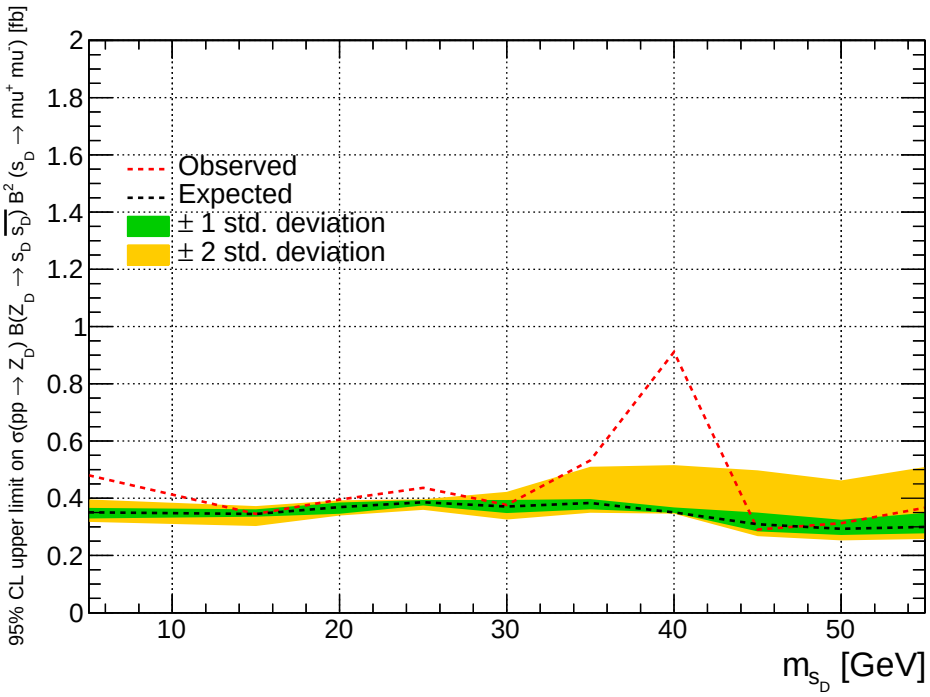


Figure 8.10: Brazilian bands for 95% CL limits with the vector-portal model samples with $m_{Z_D} = 125$ GeV are shown.

Table 8.1: Summary of systematic uncertainties on $\varepsilon_{full}/\alpha_{gen}$.

Source of uncertainty	Value
Experimental	
Integrated luminosity	2.5%
Muon ID	$4 \times 0.6\%$
Muon isolation	$2 \times 0.1\%$
Muon HLT	0.6%
Overlapping in tracker (signal below 9 GeV)	$2 \times 1.2\%$
Overlapping in muon system (signal below 9 GeV)	$2 \times 1.3\%$
Displaced track and vertex reconstruction	$2 \times 0.5\%$
PU distribution	0.05%
PU effect	1.8%
Dimuon mass consistency	0.24%
NNLO Higgs p_{T} re-weighting	2.0%
Total experimental uncertainty	8.5%
Theoretical	
PDF + α_s + QCD scales	8.0%
Higgs cross sec. and BR.	3.8%
Total theoretical uncertainty	8.8%
Background below J/ψ: 0.21-2.72 GeV	
Normalization	10.1%
Systematic	4.1%
Total background modeling uncertainty: below J/Ψ masses	10.9%
Background above J/ψ and below Υ: 3.24-9 GeV	
Normalization	12.3%
Systematic	1.5%
Total background modeling uncertainty: above J/Ψ masses	12.4%
Background above Υ: 11-60 GeV	
Normalization	16.4%
Systematic	2.3%
Shape	23.8%
Total background modeling uncertainty: above Υ masses	16.6%

Table 8.2: List of dimuon masses in observed events in the signal region.

Index	Run	Lumi	Event	$m_{\mu\mu_1}$ (GeV)	$m_{\mu\mu_2}$ (GeV)
Below J/ψ: 0.21-2.72 GeV					
1	325101	232	419558132	1.07823	1.08034
2	321817	246	436806495	1.3254	1.27405
3	321149	829	1258013039	1.99326	1.95203
4	321975	383	676490954	2.45786	2.45252
Between J/ψ and Υ: 3.24-9 GeV					
5	317320	1035	1528955010	5.03704	4.86024
6	317640	581	843743710	5.72808	5.54132
7	316060	587	615033944	8.05194	8.23958
8	316058	53	27396337	8.43057	8.34955
9	321732	976	1744000224	8.53801	8.41446
10	321007	790	1272669589	8.9618	8.87257
Above Υ: 11-60 GeV					
11	324237	156	236058213	16.0328	15.3048
12	315689	380	441164029	21.1458	20.7108
13	319854	107	158315307	21.7996	21.1412
14	317683	161	200017545	22.1098	22.6903
15	321818	647	1133062695	34.8751	34.646
16	321295	487	772995584	37.0113	38.517
17	316666	520	735147196	37.9536	36.4664
18	321732	361	670657863	38.8183	38.1312
19	321415	425	657554809	39.0521	41.1795
20	324021	198	297338909	39.1743	38.6899
21	319459	43	70546168	40.0932	42.109
22	319579	1688	2618444536	40.2104	39.1925
23	324765	376	722907726	40.2246	39.5777
24	323495	45	81623540	41.2983	41.1495
25	322492	155	268655418	47.1707	48.6395
26	319450	86	120342841	49.08	47.9448
27	319910	610	1081562887	51.4184	50.2041
28	322431	529	902730240	54.6371	54.2174
29	324970	629	1159743309	54.8353	57.102
30	322332	890	1564176321	59.0704	58.101

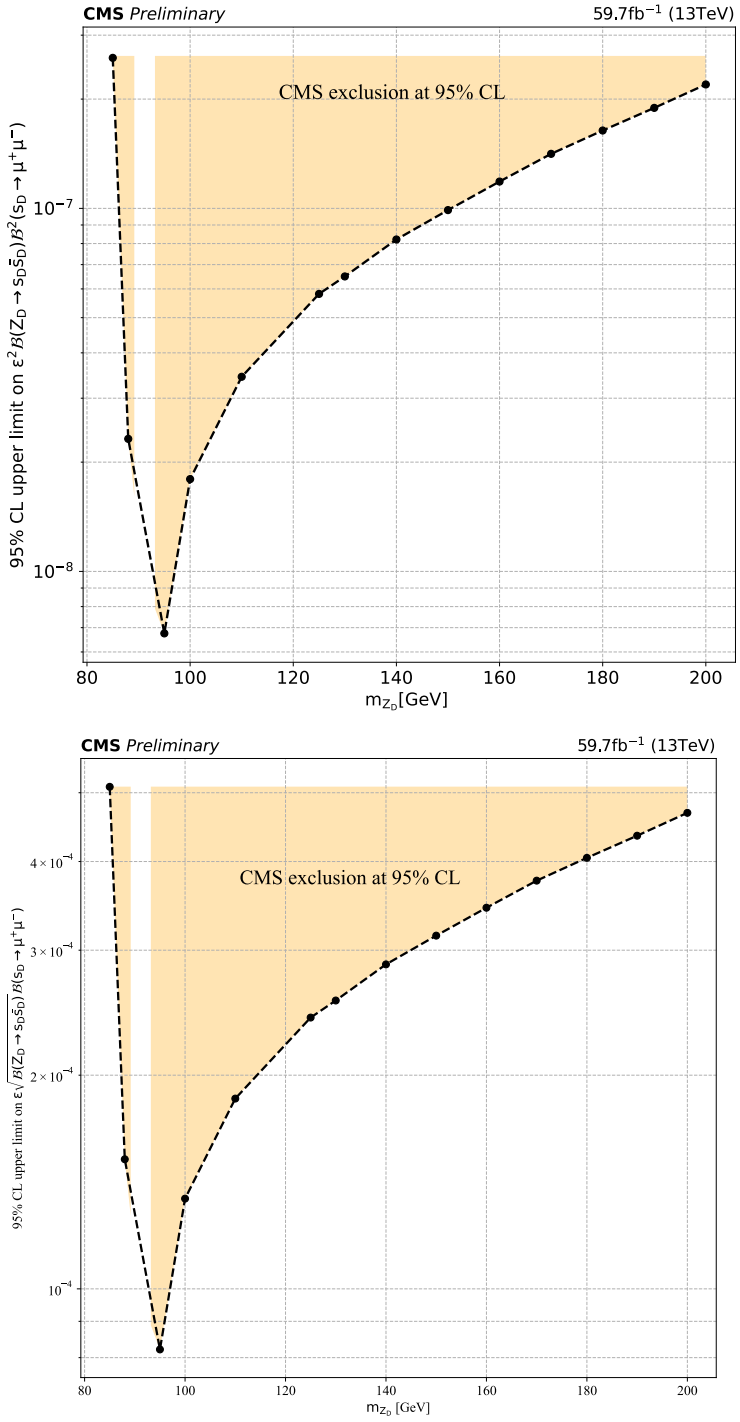


Figure 8.11: Each point represents limits averaged over the s_D masses vs. Z_D masses. The figure on the left illustrates the limits on $\varepsilon^2\mathcal{B}(\mathrm{Z_D} \to \mathrm{s_D}\overline{\mathrm{s_D}})\mathcal{B}^2(\mathrm{s_D} \to \mu^+\mu^-)$, and the figure on the right shows limits on $\varepsilon\sqrt{\mathcal{B}(\mathrm{Z_D} \to \mathrm{s_D}\overline{\mathrm{s_D}})}\mathcal{B}(\mathrm{s_D} \to \mu^+\mu^-)$. The shaded area under the curves signifies the excluded region for the averaged limits.

Chapter 9

The 2017 Analysis

A better understanding of the intriguing analysis results with localized 2-σ excesses, in which we used the dataset recorded by the CMS in the year 2018, demands extending the analysis to include the 2017 dataset and ultimately combining the results of the two years. In this chapter, I discuss my progress in analyzing the 2017 dataset corresponding to 41.5 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV.

9.1 Introduction

Even though we do not report any significant signal excess over the modeled background, we observe three and nine events in the 20-25 GeV and 35-40 GeV regions, respectively as show in Fig. 8.3. In contrast, the expected number of events in the 20-25 GeV and 35-40 GeV regions are estimated at ∼0.31 and ∼3.6, respectively. The Poissonian probability for 0.31 fluctuating to 3 is 0.00364 and for 3.6 fluctuating to 9 is 0.0076, which forced us to pause and further examine these region. One could argue that the background is not adequately modeled here, and further statistics are needed. To address this matter (with a bonus of an even stronger limit on the cross-section and

branching ratio) through the combination of 2018 and 2017 data, my colleagues and I launched a similar analysis for the 2017 data. The following chapter summarizes the 2017 analysis and my contribution to it. The 2017 analysis is done exclusively on the prompt signatures.

9.2 Samples

Data samples: The analysis data sample includes only the runs [1], in which the LHC was providing stable beams, the CMS silicon tracker, muon system, and the trigger system were performing well, and the luminosity was measured reliably. We conducted the signal search with fully reconstructed events, stored in the `DoubleMuon` primary dataset (PD) shown, as listed in Tab. 9.1.

Table 9.1: 2017 `DoubleMuon` dataset

Dataset	Events
/DoubleMuon/Run2017B-31Mar2018-v1/MINIAOD	14 501 767
/DoubleMuon/Run2017C-31Mar2018-v1/MINIAOD	49 636 525
/DoubleMuon/Run2017D-31Mar2018-v1/MINIAOD	23 075 733
/DoubleMuon/Run2017E-31Mar2018-v1/MINIAOD	51 589 091
/DoubleMuon/Run2017F-31Mar2018-v1/MINIAOD	79 756 560
Total	218 559 676

Simulated samples: The production of signal MC samples for the vector-portal model is procedurally identical to the 2018 production. However, the trigger set and the selection diverge from the 2018 analysis in the fashion I address in the following Sec. 9.3.

[1] `ReReco/Cert_294927-306462_13TeV_EOY2017ReReco_Collisions17_JSON_v1.txt`

9.3 Trigger and Event Selection

Trigger paths: The 2018 analysis high-level trigger set included the double L2 muon trigger (`HLT_DoubleL2Mu23NoVtx_2Cha_v*`), which was absent from the 2017 trigger menu. We insisted on using this trigger in 2018 because some of the benchmark models used for interpreting the model-independent results allow for displaced (long-lived) signatures. This particular trigger path recovers signal efficiency for the displaced models. This explains why we initially abstained from analyzing the 2017 dataset alongside the 2018 analysis. For the 2017 analysis, we only focus on prompt samples. Consequently, this trigger path in the 2017 is replaced by the available `DoubleMuon` `HLT_Mu23_Mu12_v*`. Tab. 9.2 shows the 2017 trigger paths. The replaced `DoubleMuon` trigger (index 1) is a double L2 muon trigger that requires a leading L2 muon with p_T threshold of 23 GeV and another L2 muon with p_{T} threshold of 12 GeV. The rest of the trigger paths are the same as used in the 2018 analysis. An additional complication is posed by the fact that not all of the triggers were available during the full run period. Tab. 9.2 shows the luminosity for each trigger.

Table 9.2: 2017 analysis signal triggers paths

Index	HLT Name	Luminosity [fb^{-1}]
1	`HLT_Mu23_Mu12_v*`	16.3
2	`HLT_Mu18_Mu9_SameSign_v*`	15.5
3	`HLT_TripleMu_12_10_5_v*`	41.5
4	`HLT_TrkMu12_DoubleTrkMu5NoFiltersNoVtx_v*`	36.7

Event selection: The offline reconstructed particle-flow (PF) muons [136] that are subjected to the event selection are taken from the "slimmedMuons" collection in the miniAOD dataset as listed in Tab. 9.1. PF muons used in this analysis are required to

satisfy the baseline *Loose* selection requirements [2]. The 2018 analysis allowed at most one of the four signal muons to be a standalone (SA) muon. Since the 2017 analysis considers only the prompt signal models, the event selection requires all four muons to be PF muons. The η selection is identical to the 2018 analysis, while the p_{T} selection is as follows: I require two muons with $p_{\mathrm{T}} > 13$ GeV and four muons with $p_{\mathrm{T}} > 8$ GeV. The lower p_{T} cuts are introduced to ensure better efficiencies given the lower statistics in the 2017 dataset. The cut on the vertex probability for 2018 is a function of ΔR, described in Eq. 6.1, whereas the vertex probability cut in 2017 is constant and equal to 0.15. The rest of the event selection criteria are the same as the 2018 sections discussed in Ch. 6. The pairing algorithm for forming the dimuons is identical to the 2018 algorithm.

Model-independence test: As I demonstrated in Sec. 6.4, the gen level acceptance (α_{gen}) to reconstructed muon acceptance (ϵ_{full}) ratio functions as a metric for model-independence of the analysis. All benchmark models must perform in such a way that the model stays independent of its parameters and exhibit a constant $\epsilon_{full}/\alpha_{gen}$ ratio across the parameter space. Furthermore, the achieved constant ratio must be comparable with the results obtained with other benchmark models. A similar figure as Fig. 6.4 is generated for the 2017 MC signal analysis and is shown in Fig. 9.1, which demonstrates that model independence is indeed maintained.

---

[2] `https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonId#Loose_Muon`

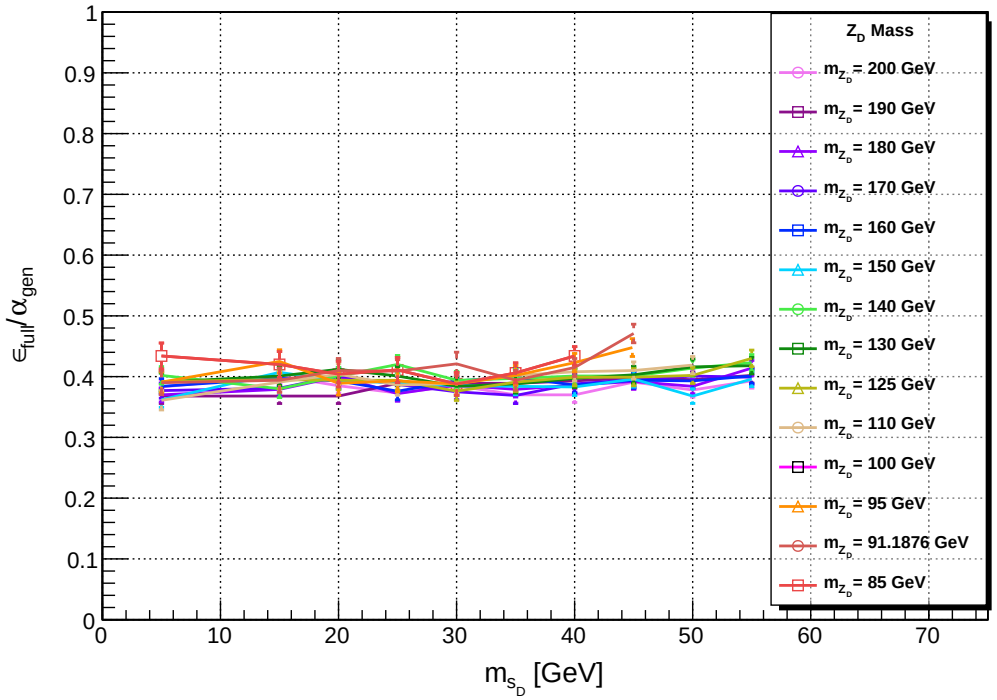


Figure 9.1: Total selection efficiency over generator level selection acceptance, $\varepsilon_{reco}/\alpha_{gen}$, as a function of the s_{D} mass for various Z_{D} masses in the vector portal model with the 2017 selection and trigger paths. The KM parameter, ε, is 10^{-2}

9.4 Background Modeling

The background modeling for 2017 follows the same categorization as the 2018 analysis. The masses below Υ resonances (below 9 GeV) are considered the low-mass region, and those above the Υ resonances (11-60 GeV) are tagged as the high-mass region.

The low-mass region: In the QCD background territory (below 9 GeV), similar to the 2018 analysis, the masses around the J/Ψ resonance are excluded (from 2.72 GeV to 3.24 GeV). Consequently, the low-mass region is further divided into the below J/Ψ region and the above J/Ψ region. The 1D distributions are fitted with Eq. 7.1 and are shown in Fig. 9.2. The 2D templates, constructed by multiplication of the fitted functions in 1D ($f(\mu\mu_1) \otimes f(\mu\mu_2)$), are shown Fig. 9.3. The estimated background in

the signal region is listed in Tab. 9.3.

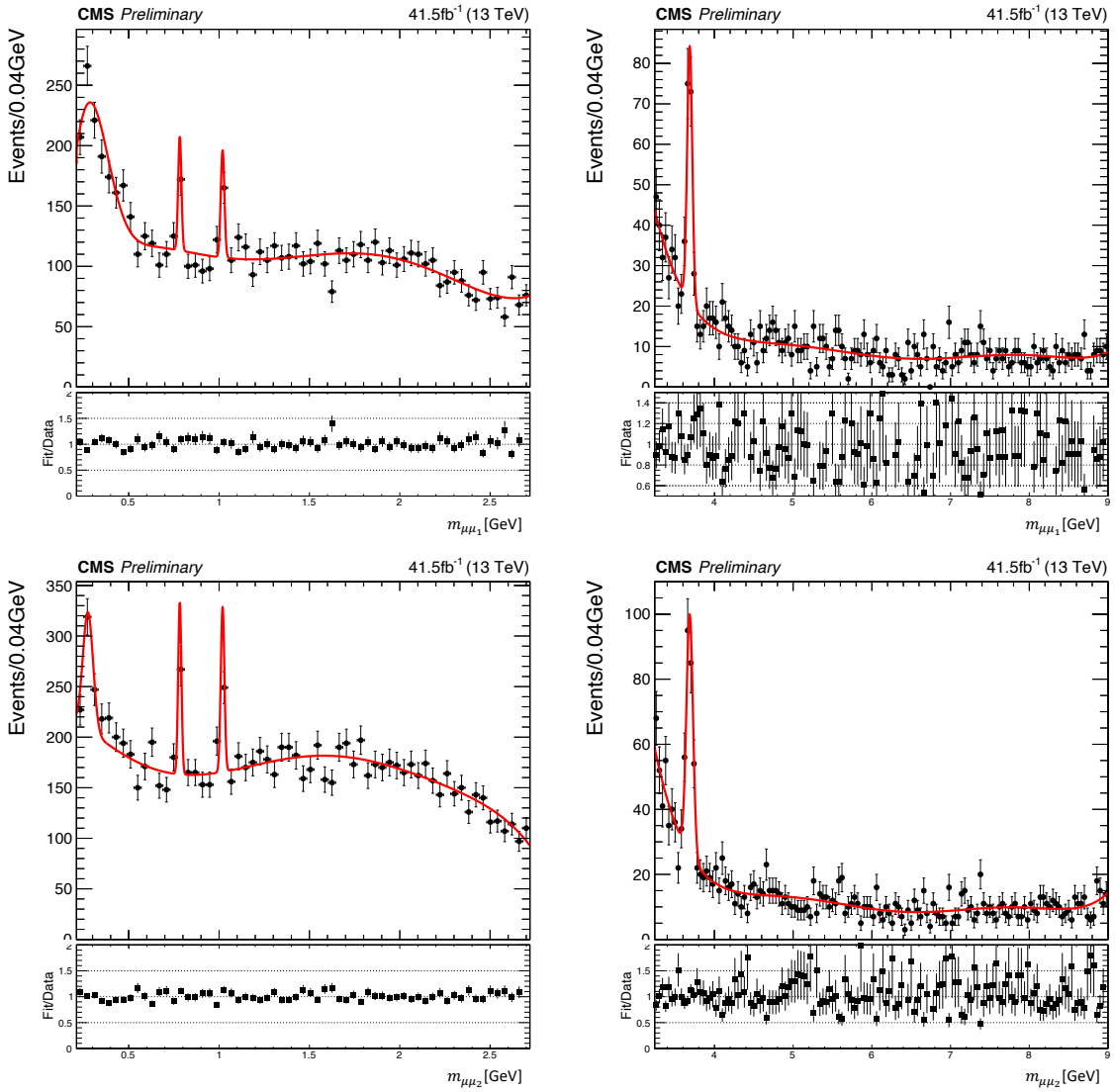


Figure 9.2: Fits to 2017 control data for the low-mass region backgrounds. Top left shows $m_{\mu\mu_1}$ and the fitted function below J/ψ masses. Top right shows $m_{\mu\mu_1}$ and the fitted function above J/ψ masses. Bottom left shows $m_{\mu\mu_2}$ and the fitted function below J/ψ masses. Bottom right shows $m_{\mu\mu_2}$ and the fitted function above J/ψ masses.

The high-mass region: The high-mass region background is dominated by electroweak background processes. In the 2018 analysis (Sec. 7.3), we utilized MC-generated events that accurately described the `DoubleMuon` data in the control region and estimated the background in the signal region. In the 2017 analysis, the MC fails to ade-

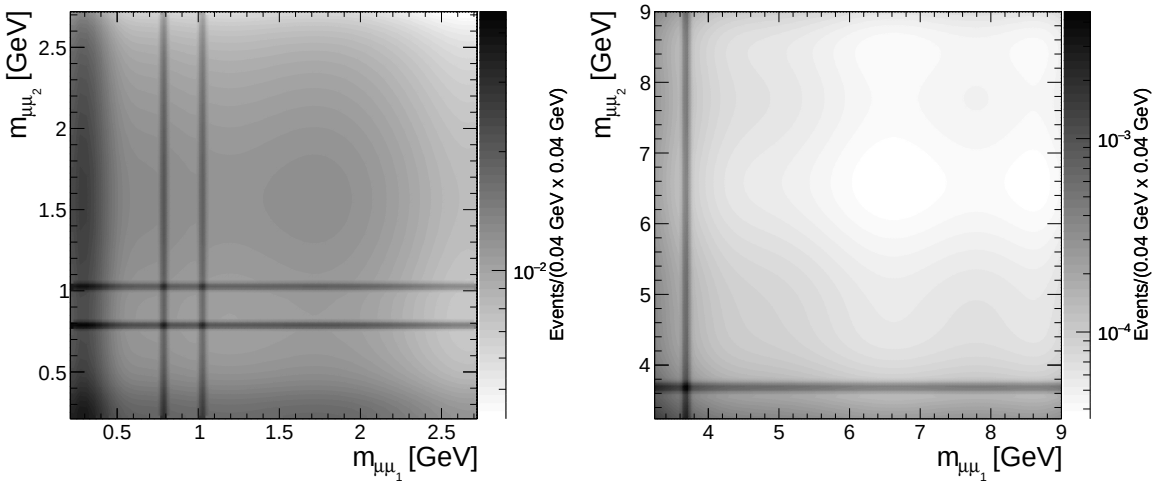


Figure 9.3: 2D QCD background templates below (left) and above (right) the J/ψ resonance in 2017. The darker vertical and horizontal lines represent the low-mass QCD resonances, as shown in the 1D templates in Fig. 9.2.

Table 9.3: The estimated background events in 2017 data in below Υ resonances in the control region and signal region.

Region	Integral ratio (R)	CR data events (N_{CR})	Est. events in SR	Stat. Unc.
Below J/ψ	0.045/0.965	49	2.26	0.32
Above J/ψ	0.088/0.918	2	0.19	0.14

quately describe the data in the control region due to lower statistics. Consequently, the MC cannot be trusted to describe the background in the signal region. Instead, for 2017, we employ a data-driven method as in the low-mass region. We use the 2017 `DoubleMuon` dataset in 1D, fit a Kernel Density Estimation probability density to $m_{\mu\mu_1}$ and $m_{\mu\mu_2}$ distributions as shown in Fig. 9.4, and lastly construct the 2D template as shown in Fig. 9.5. The estimated number of background events in the signal region is listed in Tab. 9.4.

Table 9.4: Estimated background events in 2017 data above Υ resonances in the control and signal regions

Region	Integral ratio (R)	CR data events (N_{CR})	Est. events in SR	Stat. Unc.
Above Υ	0.082/0.918	212	18.97	1.3

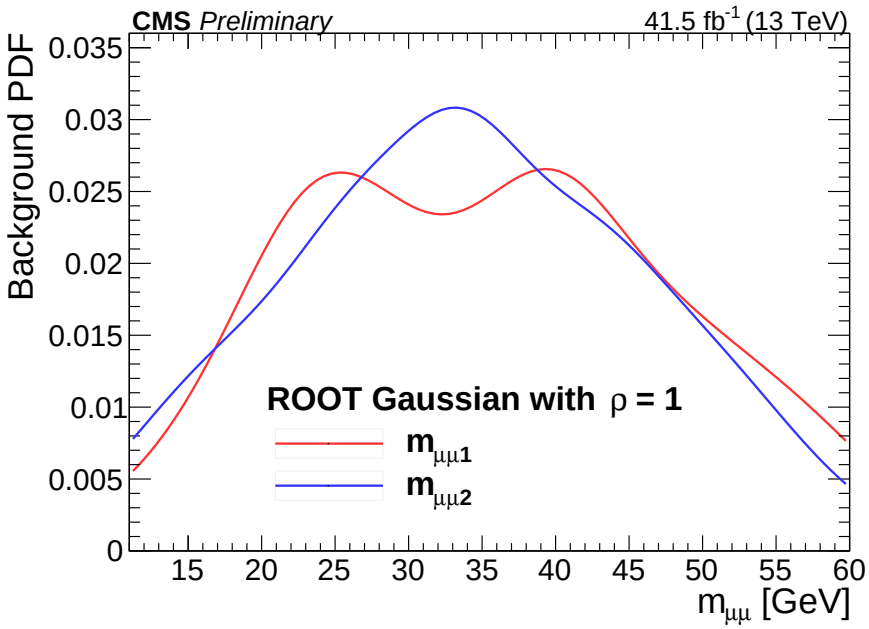


Figure 9.4: The kernel density PDFs used to describe the `DoubleMuon` dataset for each muon pair invariant mass in the high-mass region.

9.5 Signal Trigger Scale Factors

For 2017 data we use a combination of `DoubleMuon` and `TripleMuon` triggers as shown in Tab. 9.2. It is crucial to quantify the trigger efficiency of the signal triggers as it is used to scale the MC to data accurately. Conventional methods such as the tag-and-probe method cannot be used for these triggers. The reference trigger method cannot be used either since there is no shared `SingleMuon` trigger leg. Instead, I calculate the overall signal trigger efficiency at once on a control dataset via the *orthogonal trigger* method.

9.5.1 Methodology

The orthogonal method is the application of a set of `SingleMuon` triggers on `SingleMuon` datasets, instead of `DoubleMuon` datasets and multi-muon triggers, as used in the main analysis. Specifically, I use MC samples and data samples from CMS production for

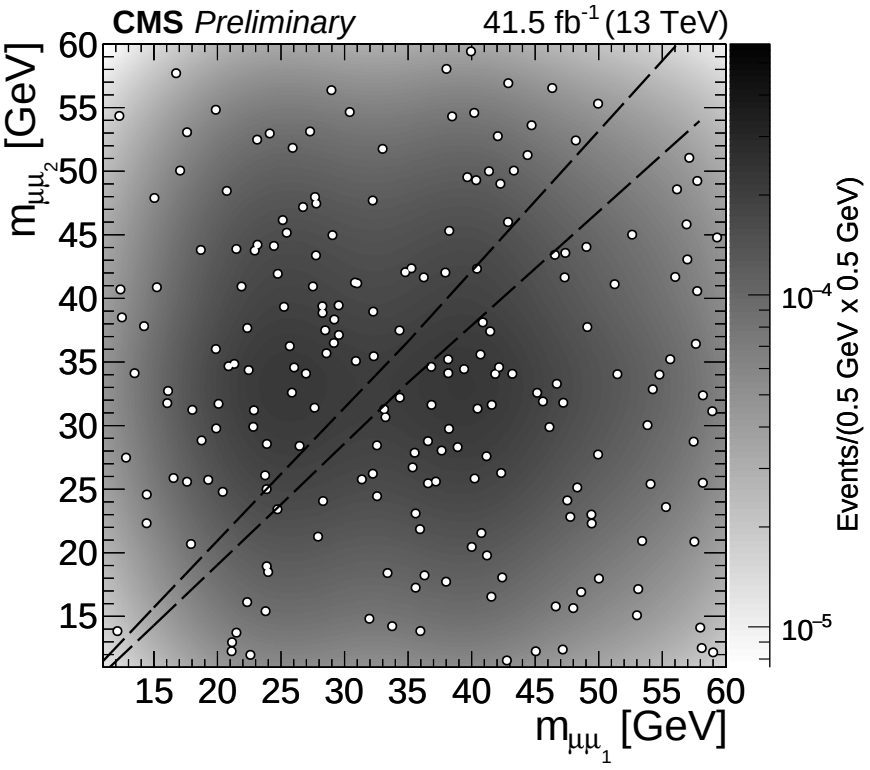


Figure 9.5: The `DoubleMuon` dataset is used to construct the 2D background template in the high-mass region. The white circles represent the `DoubleMuon` data points in the control region. The signal region remains blinded.

WZ $\rightarrow 3l\nu$ and ZZ $\rightarrow 4l$ processes. The 2017 data and simulated MC samples first go through a pre-selection process where events with at least three high-quality muons are selected. As these event types are expected to constitute a significant fraction of the events in the CR of the main analysis, they can be used to estimate the trigger scale factors as done in the previous iteration of this analysis [105] and the 2018 version of this analysis. I determine the efficiency of the signal triggers on events passing a set of selection criteria optimized to select WZ events. I perform this both on data and on MC simulated events. Next, I apply the signal HLT selection for the 2017 analysis to the surviving events and calculate its efficiency. Lastly, I derive the HLT scale factor by dividing the data signal HLT efficiency by the HLT efficiency for the simulated MC events.

9.5.2 Datasets

I use the MiniAOD single-muon samples shown below in Tab. 9.5. Events are pre-selected which have at least three muons with $p_t > 10$ GeV and at least one muon with $p_t > 20$ GeV, and are required to be in the run range specified in the JSON file [3]. Events with more than 3 muons with $p_t > 10$ GeV are rejected. The pre-selection accepts 4.7%. Additionally, I use WZ $\rightarrow 3l\nu$ and ZZ $\rightarrow 4l$ MC miniAOD files listed in Tab. 9.6. The MC miniAOD files are subjected to pre-selection and transformed into ntuples using our 2017 "ntuplizer" code with the correct reco-level variables, branches, counters and selectors, and HLT level variables.

To scale the simulated MC events to the luminosity of data (49.5 fb^{-1}), I need the cross-section for the produced MC samples: MC$_{SF} = \sigma \times$ lumi$_{data}$ / Number of events. To calculate the cross-sections for the ZZ and WZ processes I use the CMS developed tool GenXSecAnalyzer [4].

Table 9.5: `SingleMuon` data samples for the trigger scale factor studies.

Dataset name	Total Events	Pre-selected Events
/SingleMuon/Run2017B-31Mar2018-v1/MINIAOD	136 300 266	5 203 873
/SingleMuon/Run2017C-31Mar2018-v1/MINIAOD	165 652 756	6 775 048
/SingleMuon/Run2017D-31Mar2018-v1/MINIAOD	70 361 660	2 619 404
/SingleMuon/Run2017E-31Mar2018-v1/MINIAOD	154 630 534	7 588 986
/SingleMuon/Run2017F-09May2018-v1/MINIAOD	242 140 980	13 696 132
Total	769 086 196	35 883 443

Table 9.6: Monte Carlo samples for the trigger scale factor studies.

Abbreviation	Dataset name	Pre-selected events	Cross Section [pb]
WZTo3LNu	/WZTo3LNu_TuneCP5_13TeV-amcatnloFXFX pythia8/RunIIFall17MiniAODv2_ PU2017_12Apr2018_94X_mc2017_ realistic_v14-v1/MINIAODSIM	637 306	5.052 ± 0.004
ZZTo4L	ZZTo4L_13TeV_powheg_pythia8/RunIIFall17MiniAODv2 PU2017_12Apr2018_94X_mc2017 _realistic_v14-v2/MINIAODSIM	918 053	1.325 ± 0.001

[4] https://twiki.cern.ch/twiki/bin/view/CMS/HowToGenXSecAnalyzer

9.5.3 Event Selection

Initially, I require at least one of the orthogonal triggers listed in Tab. 9.7 to fire for further selection. Furthermore, events must have exactly three muons with $|\eta| < 2.4$ and with transverse momenta thresholds $p_{\mathrm{T},1} > 20$ GeV, $p_{\mathrm{T},2} > 20$ GeV, $p_{\mathrm{T},3} > 10$ GeV. Two muons with opposite charge and with an invariant mass compatible with the Z mass ($|m_{\mu\mu} - m_Z| < 15$ GeV) are paired. At least one pair is required in each event. The events are further cleaned by requiring that each muon pass the *tight* muon ID. To suppress muons from hadrons decaying in-flight (non-prompt muons), selections on the muon impact parameter and distance to the interaction point (IP) are applied. The impact parameter must be $|d_{xy,i}| < 0.005$ cm and the distance to the IP $|d_{z,i}| < 0.01$ cm for each muon. Non-isolated muons are rejected by applying a relative isolation cut of $Iso_{rel} < 0.1$ on each muon.

Tab. 9.8 shows the event selection in data and MC. The data are found to be consistent with the Monte Carlo prediction. Figs. 9.7-9.9 show the agreement between data and MC for several relevant event variables.

9.5.4 Results

I arrive at the final trigger efficiency on MC $\varepsilon_{WZTo3LNu} = 374.591/375.790 = 0.99$, $\varepsilon_{ZZTo4L} = 18.793/18.889 = 0.99$. Futhermore, I observe the efficiency in data to be $\varepsilon_{Data} = 399/426 = 0.94$. For the statistical uncertainty, I use *binomial* error estimation on these ratios. The resulting statistical uncertainty for the scale factor is 1.2%. I calculate the systematic uncertainty by varying the p_t selection configurations. Fig. 9.6 illustrates how the efficiencies and the SF change with respect to different p_t configurations. The standard deviation on the SF distribution is 0.25% and constitutes the systematic uncertainty.

Table 9.7: SingleMu Triggers used in the 2017 Analysis

Trigger Path
HLT_IsoMu20_eta2p1_LooseChargedIsoPFTau27_eta2p1_CrossL1
HLT_IsoMu20_eta2p1_LooseChargedIsoPFTau27_eta2p1_TightID_CrossL1
HLT_IsoMu20_eta2p1_MediumChargedIsoPFTau27_eta2p1_CrossL1
HLT_IsoMu20_eta2p1_MediumChargedIsoPFTau27_eta2p1_TightID_CrossL1
HLT_IsoMu20_eta2p1_TightChargedIsoPFTau27_eta2p1_CrossL1
HLT_IsoMu20_eta2p1_TightChargedIsoPFTau27_eta2p1_TightID_CrossL1
HLT_IsoMu20
HLT_IsoMu24_eta2p1_LooseChargedIsoPFTau20_SingleL1
HLT_IsoMu24_eta2p1_LooseChargedIsoPFTau20_TightID_SingleL1
HLT_IsoMu24_eta2p1_LooseChargedIsoPFTau35_Trk1_TightID_eta2p1_Reg_CrossL1
HLT_IsoMu24_eta2p1_LooseChargedIsoPFTau35_Trk1_eta2p1_Reg_CrossL1
HLT_IsoMu24_eta2p1_MediumChargedIsoPFTau20_SingleL1
HLT_IsoMu24_eta2p1_MediumChargedIsoPFTau20_TightID_SingleL1
HLT_IsoMu24_eta2p1_MediumChargedIsoPFTau35_Trk1_TightID_eta2p1_Reg_CrossL1
HLT_IsoMu24_eta2p1_MediumChargedIsoPFTau35_Trk1_eta2p1_Reg_CrossL1
HLT_IsoMu24_eta2p1_TightChargedIsoPFTau20_SingleL1
HLT_IsoMu24_eta2p1_TightChargedIsoPFTau20_TightID_SingleL1
HLT_IsoMu24_eta2p1_TightChargedIsoPFTau35_Trk1_TightID_eta2p1_Reg_CrossL1
HLT_IsoMu24_eta2p1_TightChargedIsoPFTau35_Trk1_eta2p1_Reg_CrossL1
HLT_IsoMu24_eta2p1
HLT_IsoMu24
HLT_IsoMu27
HLT_IsoMu30
HLT_L1SingleMu18
HLT_L1SingleMu25
HLT_L1_DoubleJet30_Mass_Min400_Mu10
HLT_L2Mu10
HLT_L2Mu50
HLT_Mu10_TrkIsoVVL_DiPFJet40_DEta3p5_MJJ750_HTT350_PFMETNoMu60
HLT_Mu15_IsoVVVL_PFHT450_CaloBTagCSV_4p5
HLT_Mu15_IsoVVVL_PFHT450_PFMET50
HLT_Mu15_IsoVVVL_PFHT450
HLT_Mu15_IsoVVVL_PFHT600
HLT_Mu20
HLT_Mu27
HLT_Mu3_PFJet40
HLT_Mu50_IsoVVVL_PFHT450
HLT_Mu50
HLT_Mu55
HLT_Mu8_TrkIsoVVL_DiPFJet40_DEta3p5_MJJ750_HTT300_PFMETNoMu60
HLT_OldMu100
HLT_TkMu100

Table 9.8: Table with event selection for the three-muon control region. MC samples are scaled to data sample size.

	Number of events surviving selection		
Selection	WZTo3LNu MC	ZZTo4Mu MC	Data
Pre-selection (if applicable)	12524	3140	35883443
Passes at least one orthogonal trigger	11766	2916	35883443
Exactly three muons & $\|\eta_i\| < 2.4$	2140	409	7835950
$\|\eta_i\| < 2.4$ $p_{\mathrm{T},1} > 20$ GeV, $p_{\mathrm{T},2} > 20$ GeV, $p_{\mathrm{T},3} > 10$ GeV	908	90	402582
Two muons with opposite charge	904	90	357734
$\|m_{\mu\mu} - m_Z| < 10$ GeV	722	51	214371
Tight muon ID	639	39	9097
$\|d_{xy,i}| < 0.005$ cm	610	34	4081
$\|d_{z,i}| < 0.01$ cm	531	27	2651
$\mathrm{RelIso}_i < 0.1$	376	19	426
Passes at least one signal trigger	375	19	399

This results in a trigger scale factor of $SF = \varepsilon_{data}/\varepsilon_{MC} = 0.937/0.99 = 0.941$ as given above. Finally, I estimate the overall trigger scale factor to be $94.1\% \pm 1.2\%$ (stat.) $\pm\, 0.26\%$ (syst.).

9.5.5 Accounting for Different Run Eras

One complication with the above results is that the dataset in 2017 is not triggered on with uniform efficiency. To account for this disparity, I divide the 2017 dataset into four different eras and observe the trigger efficiency for each. In Tab. 9.9, the four eras with their respective luminosity and available triggers are listed.

Table 9.9: The integrated luminosity and available signal triggers of each run era in the 2017 dataset.

Run era	Lumi (fb^{-1})	HLT_Mu18_Mu9_SS	HLT_Mu23_Mu12	HLT_TrkMu12_DoubleTrkMu5	HLT_TripleMu_12_10_5
Run B	4.79	No	No	No	Yes
Run C, D, E	23.19	No	No	Yes	Yes
Run F	13.53	Yes	Yes	Yes	Yes

Next, I apply the full selection, as discussed in Sec. 9.5.3, to each run era and separately measure the data and MC trigger efficiencies according to the available triggers in that run era. As tabulated in Tab. 9.10, for each run era, the ratio of data trigger efficiency to the total MC trigger efficiency determines the trigger scale

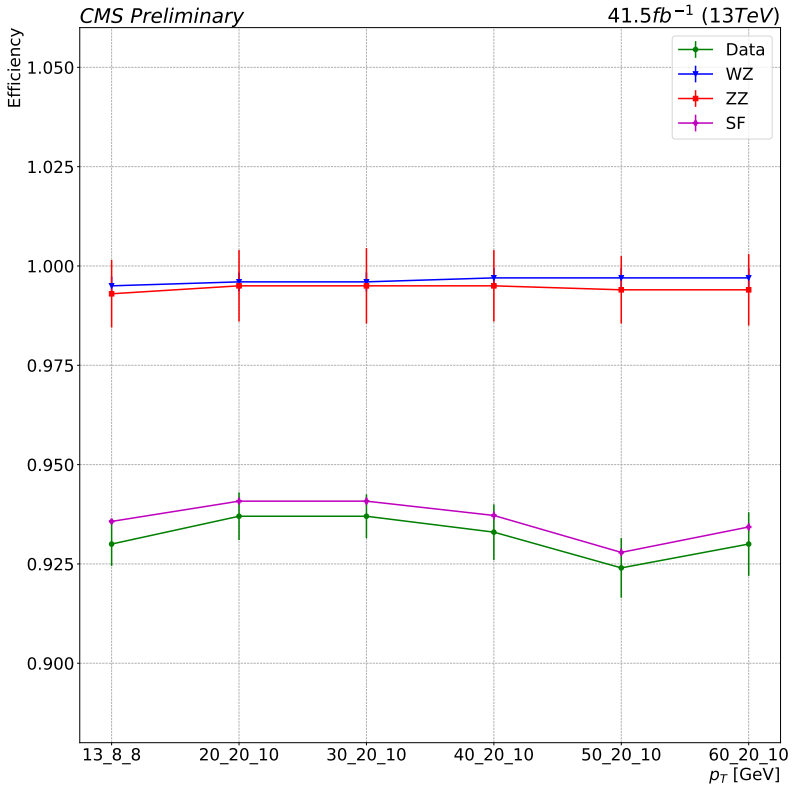


Figure 9.6: The behavior of data and MC relative efficiency with varying p_t configuration. The scale factor (SF) is included based on data and MC efficiencies.

factor for that era. Particularly, I compute the total MC trigger efficiency as the cross-section-weighted average of WZ $\rightarrow 3l\nu$ and ZZ $\rightarrow 4l$ efficiencies. For reference, the cross-sections for both processes are listed in Tab. 9.6. Finally, I assess the Luminosity-weighted average of the resulting scale factors, which emerges with a higher value than that of the *flat* treatment of the efficiencies as demonstrated in Eq. 9.1. This is an expected outcome since by dividing the data according to the available triggers in each era, the MC efficiency calculation also considers the inefficiencies from the missing triggers in data.

$$\overline{SF} = \frac{(4.79 \times 0.897) + (23.19 \times 0.988) + (13.53 \times 0.957)}{41.5} = 0.967 \qquad (9.1)$$

Table 9.10: The signal trigger efficiencies for the MC samples and 2017 dataset according to the available triggers in each run era.

Run	$WZ_{rel.\ eff.}$	$ZZ_{rel.\ eff.}$	$Data_{rel.\ eff.}$	SF
Run B	0.916	0.912	0.821	0.897
Run C, D, E	0.965	0.96	0.95	0.988
Run F	0.996	0.995	0.953	0.957

9.6 Concluding Remarks

The unblinding of the results in the signal region in the 2018 analysis called for further investigation into the number of observed events in the 20-25 GeV and 35-40 GeV invariant mass regions and its apparent discrepancy with the number of expected events in these regions. I have completed the model-independence test for the benchmark models, the background modeling, background estimation in the signal region, and the uncertainty analysis, including the calculation of the trigger scale factor. The next step is to unblind the signal region and set a model-independent 95% upper limit followed by limit setting on the vector-portal benchmark model. Lastly, the 2017 and 2018 results will have to be combined through the combine-tool to improve the statistics and the background modeling in the region in question.

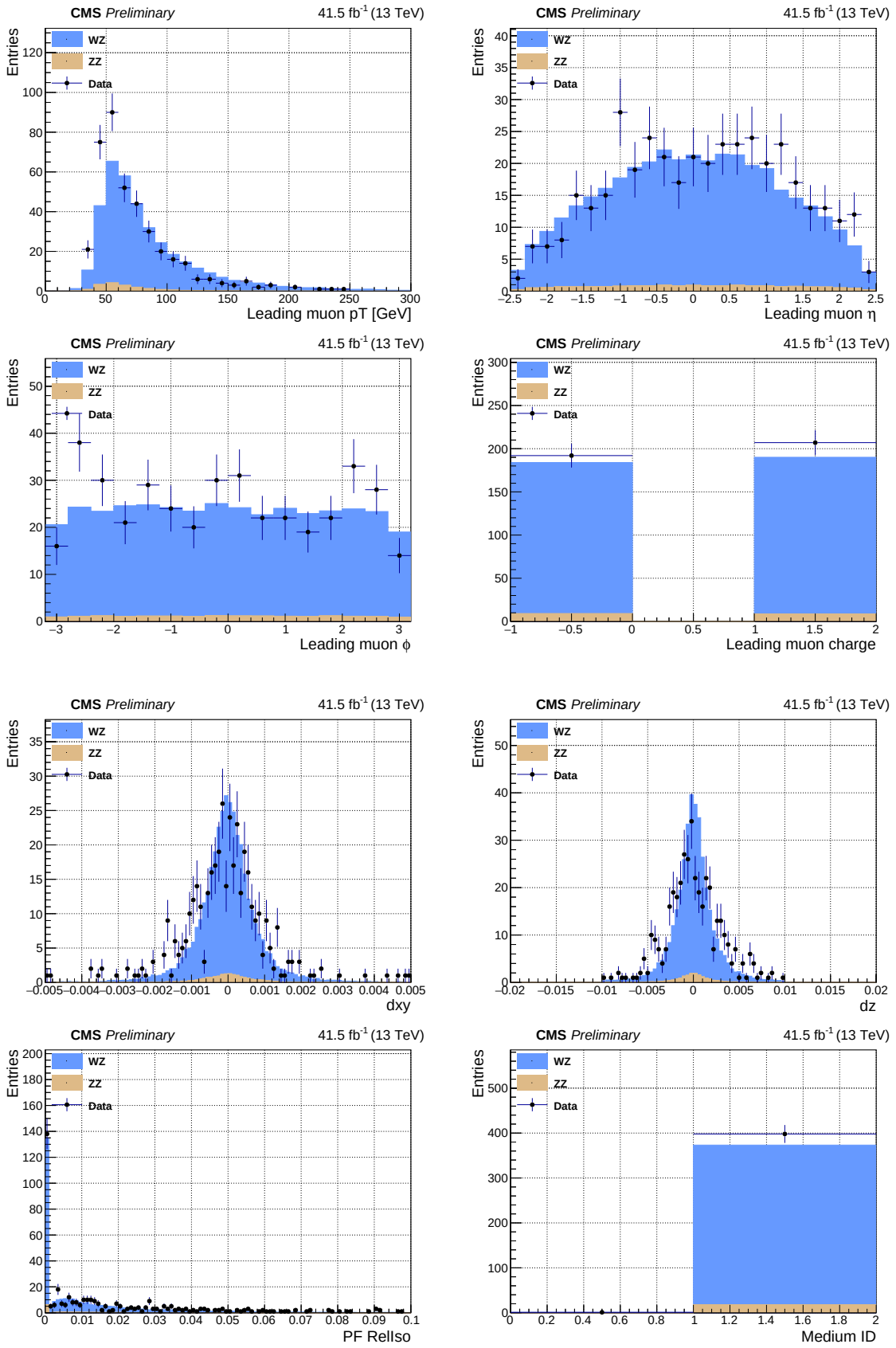


Figure 9.7: Data vs. Monte Carlo comparisons in the control region for 2017 data after all selections are applied. Properties of the leading muon.

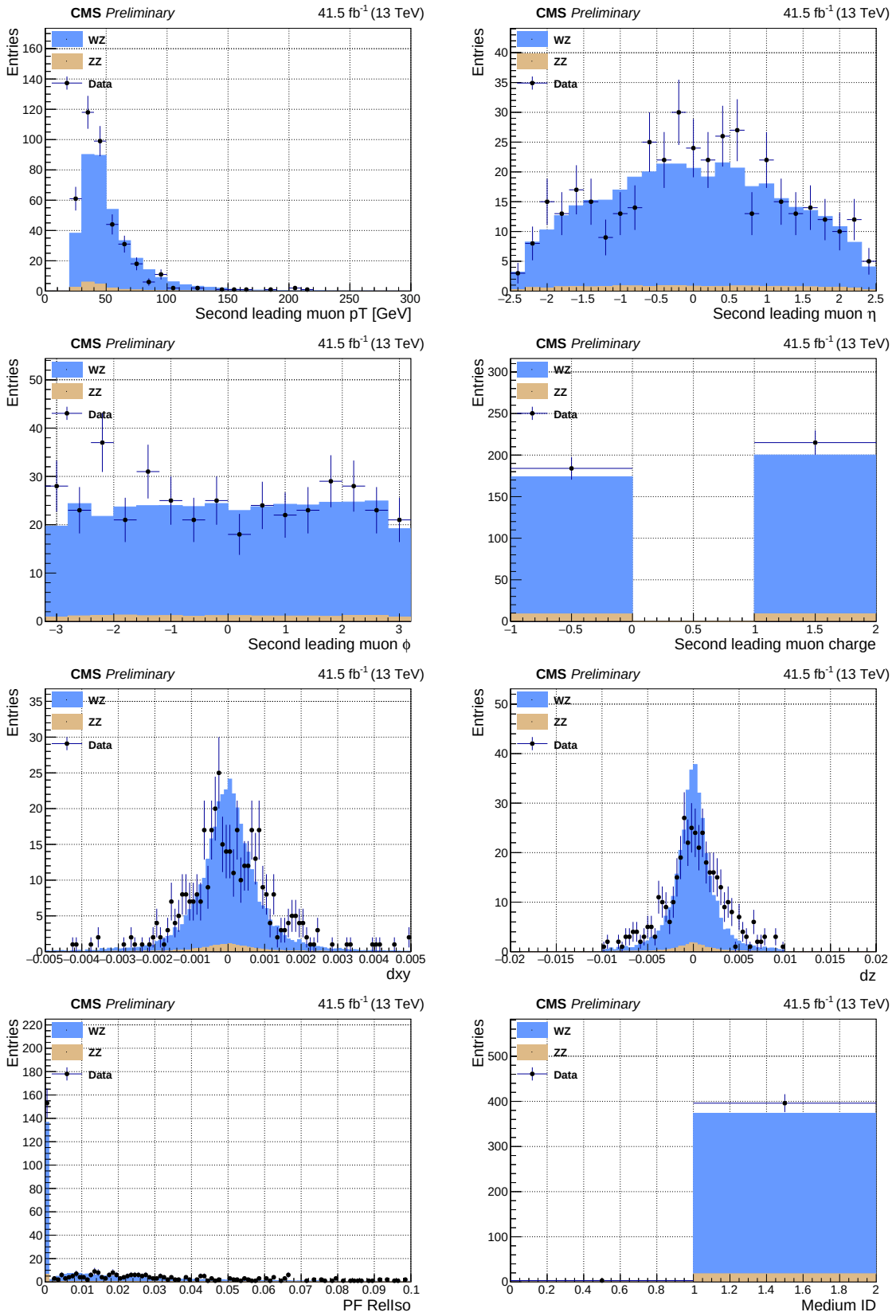


Figure 9.8: Data vs. Monte Carlo comparisons in the control region for 2017 data after all selections are applied. Properties of the sub-leading muon.

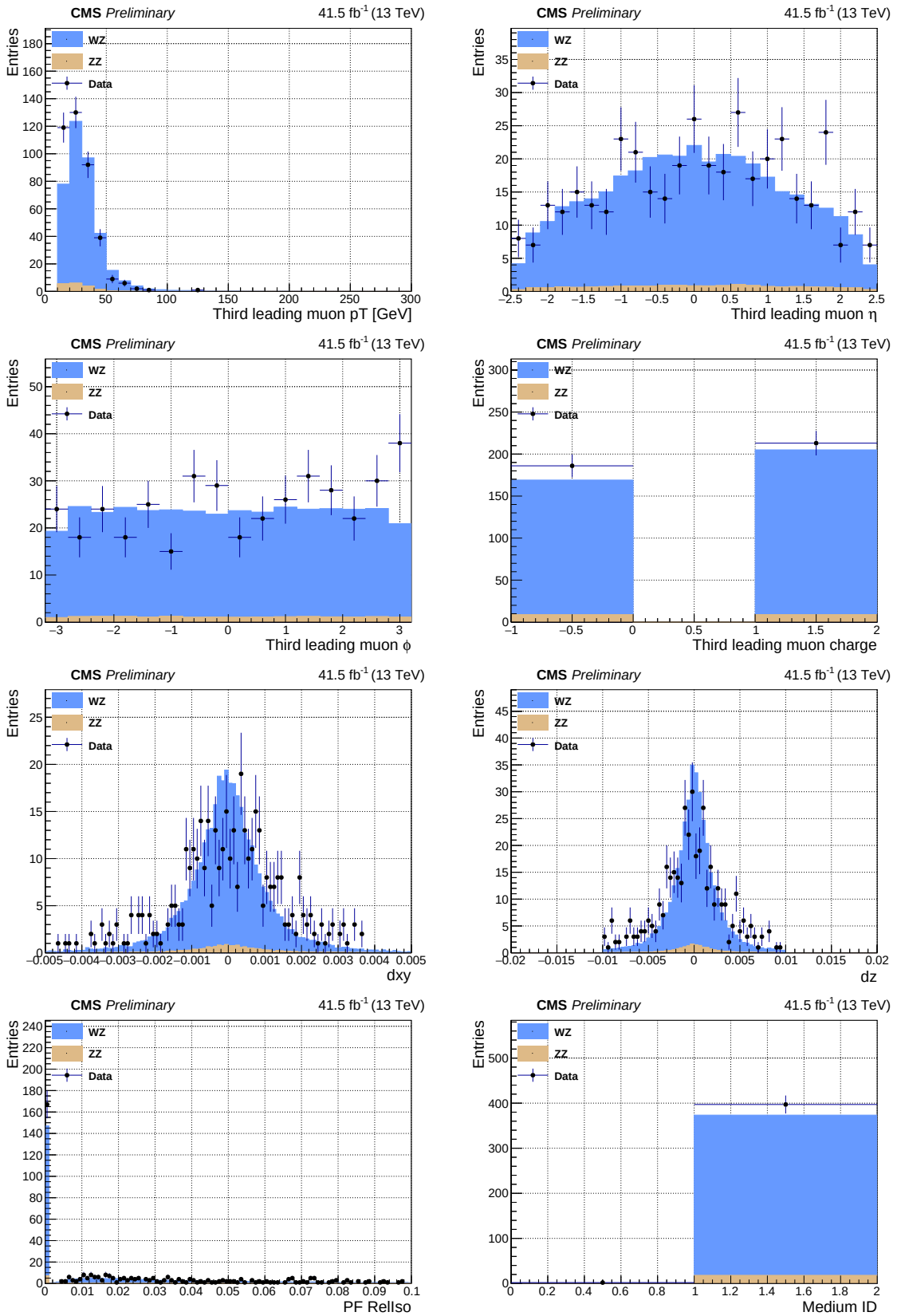


Figure 9.9: Data vs. Monte Carlo comparisons in the control region for 2017 data after all selections are applied. Properties of the third-leading muon.

A Final Word

Here, in this closing section, I offer a few suggestions on how to improve or extrapolate this work. I close by offering some insights afforded to me by my experience having worked both on hardware and analysis.

In this dissertation, I reported on my years of Ph.D. work on three major fronts:

- My participation in the upgrade of the CMS endcaps with GEM detectors
- My collaboration with Texas A&M and Rice University on a model-independent search for pair production of a new boson that decays to a pair of oppositely charged muons
- The interpretation of the model-independent search results within a dark matter model (vector-portal search)

The GEM technology has a great potential in the high-energy physics field. The GE1/1 was the entry project for the GEM technology to the CMS detector. At the time of authoring this document there are multiple GEM projects for the next grand upgrade of CMS, including the GE2/1 and ME0 projects. By working on the hardware side, I have learned the importance of community communication and patience when inevitable deadlocks appear. The research and development and eventually construction and testing of the GEM detectors taught me that, in the laboratory environment,

however inconvenient it might appear, it is necessary to abide by every recommended procedure as these procedures are the product of past experiences.

On the analysis side, the result of the 2018 dataset, as presented in Chap. 8, will be combined with the previously published results in Ref. [5] for new boson mass below 9 GeV. Additionally, my work, as presented in Chap 9, for the 2017 dataset, will be completed and combined with the results from the 2016 and 2018 data sets for a final result from Run 2 of the LHC. This combination will impose more stringent limit on $\sigma(pp \to 2a + X) \times \mathcal{B}^2(a \to 2\mu) \times \alpha_{gen}$. The resulting limits can then be translated into a 95% CL limit on the product of the production cross-section of the dark vector boson Z_D, branching fraction of Z_D decaying to a pair of dark scalar bosons s_D, and the squared branching fraction of s_D decaying to two muons, by following the methods that I laid out in Ch. 8.

The vector-portal model could be explored further to include dark fermions f_D particles and their eventual decays into muon pairs through an off-shell Z_D, $pp \to Z_D \to f_{D_1}\overline{f_{D_1}} \to f_{D_2}\overline{f_{D_2}}4\mu$, where f_{D_2} particles are stable and could be considered as a dark matter candidate. In the early stages of the model implementation in this analysis, I explored such a model; however, I temporarily adorned it because of complications in defining signal/control regions with this model. In my time as a post-doctoral researcher, I will attempt to bring this model to fruition.

The signal region of the search could be expanded by considering the production of two different new bosons in the entire 2D mass plane. In this case, the muon pairing algorithm has to be modified. One could pair muons based on the vertex probability or deploy machine-learning techniques based on the Monte-Carlo simulations and feature analysis.

Regarding the vector-portal model, the higher luminosity of the LHC and expanded fiducial volume at CMS, in the future, will provide the opportunity for allowing for

Long-Lived dark particles in the vector-portal model. The model can also be modified to allow for more decays channels for the dark scalar. Furthermore, one could bring *Higgs-mixing* into the picture along with the kinetic-mixing mechanism explored here.

My years of doctorate program were an exhilarating experience; to be on the edge of human knowledge is a priceless fortune. I suppose having worked both in hardware and data analysis endeavors, in collaboration with a community of 5000 scientists, from all around the globe, afforded me a rather unique perspective. I have often noticed that various sub-fields can benefit from better cross-group communications; not only different project at CMS, but also an effective communication line between theoreticians and experimentalists. There are already meetings and conference that could function as such communication lines, however my suggestion is to have a program specifically for graduate students to exchange notes.

Lastly, let me thank you, the reader, for making it this far into this document. My sincere hope is that you found it to be a pleasant read.

Bibliography

[1] A Colaleo, A Safonov, A Sharma, and M. Hohlmann Tytgat. CMS technical design report for the muon endcap GEM upgrade. Technical report, 2015.

[2] D Abbaneo, M Abbas, M Abbrescia, H Abdalla, A Ahmad, A Ahmed, W Ahmed, C Ali, I Asghar, P Aspell, et al. Layout and assembly technique of the GEM chambers for the upgrade of the CMS first muon endcap station. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 918:67–75, 2019.

[3] Wikimedia Commons. Standard model of elementary particles, 2004. [Online; accessed 19-October-2021].

[4] John Ellis, Mary K Gaillard, and Dimitri V Nanopoulos. An updated historical profile of the Higgs boson. *arXiv preprint arXiv:1504.07217*, 2015.

[5] Albert M Sirunyan, Armen Tumasyan, Wolfgang Adam, Federico Ambrogi, Ece Asilar, Thomas Bergauer, Johannes Brandstetter, Marko Dragicevic, Janos Erö, A Escalante Del Valle, et al. A search for pair production of new light bosons decaying into muons in proton-proton collisions at 13 TeV. *Physics Letters B*, 796:131–154, 2019.

[6] Tai Sakuma. Cutaway diagrams of CMS detector. In *J. Phys.: Conf. Ser. 513 022032*, number CMS-OUTREACH-2019-001, 2019.

[7] A Dominguez et al. CMS technical design report for the pixel detector upgrade. Technical report, Fermi National Accelerator Lab.(FNAL), Batavia, IL (United States), 2012.

[8] Duccio Abbaneo, M Abbas, M Abbrescia, M Abi Akl, O Aboamer, D Acosta, A Ahmad, W Ahmed, A Aleksandrov, P Altieri, et al. Quality control for the first large areas of triple-GEM chambers for the CMS endcaps. In *EPJ Web of Conferences*, volume 174, page 03003. EDP Sciences, 2018.

[9] CMS collaboration et al. The CMS trigger system. *preprint arXiv:1609.02366*, 2016.

[10] Wei Shi. *A Model Independent Search for New Bosons Decaying into Muons at LHC*. PhD thesis, Rice University, 2021.

[11] M Abbas, M Abbrescia, H Abdalla, A Abdelalim, S AbuZeid, A Agapitos, A Ahmad, A Ahmed, W Ahmed, C Aimè, et al. Quality control of mass-produced GEM detectors for the CMS GE1/1 muon upgrade. *preprint arXiv:2203.12037*, 2022.

[12] R Santonico and R Cardarelli. Development of resistive plate counters. *Nuclear Instruments and Methods in Physics Research*, 187(2-3):377–380, 1981.

[13] Jay Hauser, CMS Collaboration, et al. Cathode strip chambers for the CMS endcap muon system. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 384(1):207–210, 1996.

[14] G Abbiendi, N Amapane, C Battilana, C Fernandez Bedoya, Riccardo Bellan, M Bellato, AC Benvenuti, P Biallas, Sara Bolognesi, S Braibant, et al. The CMS muon barrel drift tubes system commissioning. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 598(1):192–195, 2009.

[15] Fabio Sauli. GEM: A new concept for electron amplification in gas detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 386(2-3):531–534, 1997.

[16] P Aspell, J Petäjäjärvi, P Vichoudis, J Kaspar, J Kopal, W Snoeys, J Rouet, V Avati, W Bialas, and E Radicioni. The VFAT production test platform for the TOTEM experiment. 2008.

[17] B. Dorney M. Bianco, J. A. Merlin. GE1/1 Quality Control Instructions. Technical report, 2016.

[18] Wenhao You, Yi Zhou, Cheng Li, Hongfang Chen, Ming Shao, Yongjie Sun, Zebo Tang, and Jianbei Liu. A large area GEM detector using an improved self-stretch technique. *preprint arXiv:1405.1872*, 2014.

[19] V Peskov and P Fonte. Research on discharges in micropattern and small gap gaseous detectors. *preprint arXiv:0911.0463*, 2009.

[20] M Abbas, M Abbrescia, H Abdalla, S Abu Zeid, A Agapitos, A Ahmad, A Ahmed, W Ahmed, S Amarjeet, I Asghar, et al. Performance of prototype GE1/ 1 chambers for the CMS muon spectrometer upgrade. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 972:164104, 2020.

[21] S Bachmann, Andrea Bressan, M Capeáns, M Deutel, S Kappler, B Ketzer, A Polouektov, L Ropelewski, F Sauli, E Schulte, et al. Discharge studies and prevention in the gas electron multiplier (GEM). *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 479(2-3):294–308, 2002.

[22] Fabio Sauli. *Gaseous radiation detectors: Fundamentals and applications*. Number 36. Cambridge University Press, 2014.

[23] YK Kim, KK Irikura, ME Rudd, MA Ali, PM Stone, JS Coursey, RA Dragoset, AR Kishore, KJ Olsen, AM Sansonetti, et al. Electron-impact cross sections for ionization and excitation. *NIST Standard Reference Database*, 107, 2010.

[24] Glenn F Knoll. *Radiation detection and measurement*. John Wiley and Sons, 2010.

[25] MJ French, LL Jones, Q Morrissey, A Neviani, R Turchetta, J Fulcher, G Hall, E Noah, M Raymond, G Cervelli, et al. Design and results from the APV25, a deep sub-micron cmos front-end chip for the CMS tracker. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 466(2):359–365, 2001.

[26] S Martoiu, H Muller, A Tarazona, and J Toledo. Development of the scalable readout system for micro-pattern gas detectors and other applications. *Journal of Instrumentation*, 8(03):C03015, 2013.

[27] S Bachmann, Andrea Bressan, Leszek Ropelewski, Fabio Sauli, A Sharma, and D Mörmann. Charge amplification and transfer processes in the gas electron multiplier. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 438(2-3):376–408, 1999.

[28] Particle Data Group, PAea Zyla, RM Barnett, J Beringer, O Dahl, DA Dwyer, DE Groom, C-J Lin, KS Lugovsky, E Pianori, et al. Review of particle physics. *Progress of Theoretical and Experimental Physics*, 2020(8):083C01, 2020.

[29] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508, 1964.

[30] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585, 1964.

[31] Paul Adrien Maurice Dirac. The quantum theory of the emission and absorption of radiation. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 114(767):243–265, 1927.

[32] Franz Mandl and Graham Shaw. *Quantum field theory*. John Wiley & Sons, 2010.

[33] Michael E Peskin. *An introduction to quantum field theory*. CRC press, 2018.

[34] Werner Heisenberg. Die beobachtbaren Größen in der Theorie der Elementarteilchen. In *Original Scientific Papers/Wissenschaftliche Originalarbeiten*, pages 611–636. Springer, 1989.

[35] Freeman J Dyson. The radiation theories of Tomonaga, Schwinger, and Feynman. *Physical Review*, 75(3):486, 1949.

[36] Wolfgang Pauli. Relativistic field theories of elementary particles. *Reviews of Modern Physics*, 13(3):203, 1941.

[37] James Clerk Maxwell. Viii. A dynamical theory of the electromagnetic field. *Philosophical transactions of the Royal Society of London*, (155):459–512, 1865.

[38] David J Griffiths. Introduction to Electrodynamics, 2005.

[39] Michael Redhead. The interpretation of gauge symmetry. In *Ontological aspects of quantum field theory*, pages 281–301. World Scientific, 2002.

[40] Chen-Ning Yang and Robert L Mills. Conservation of isotopic spin and isotopic gauge invariance. *Physical review*, 96(1):191, 1954.

[41] Sheldon L Glashow. The renormalizability of vector meson interactions. *Nuclear Physics*, 10:107–117, 1959.

[42] Abdus Salam. Weak and electromagnetic interactions. In *Selected Papers Of Abdus Salam: (With Commentary)*, pages 244–254. World Scientific, 1994.

[43] STeVen Weinberg. A model of leptons. *Physical review letters*, 19(21):1264, 1967.

[44] S Dawson. Introduction to electroweak symmetry breaking. *arXiv preprint hep-ph/9901280*, 1999.

[45] Jeffrey Goldstone. Field theories with superconductor solutions. *Il Nuovo Cimento (1955-1965)*, 19(1):154–164, 1961.

[46] Hideki Yukawa. On the interaction of elementary particles. I. *Proceedings of the Physico-Mathematical Society of Japan. 3rd Series*, 17:48–57, 1935.

[47] Albert Camus. *The myth of Sisyphus*. Penguin UK, 2013.

[48] John F Donoghue. The effective field theory treatment of quantum gravity. In *AIP Conference Proceedings*, volume 1483, pages 73–94. American Institute of Physics, 2012.

[49] Nima Arkani-Hamed, Douglas P Finkbeiner, Tracy R Slatyer, and Neal Weiner. A theory of dark matter. *Physical Review D*, 79(1):015014, 2009.

[50] Julio F Navarro. The structure of cold dark matter halos. In *Symposium-international astronomical union*, volume 171, pages 255–258. Cambridge University Press, 1996.

[51] Miao Li, Xiao-Dong Li, Shuang Wang, and Yi Wang. Dark energy. *arXiv preprint arXiv:1103.5870*, 2011.

[52] P James E Peebles and Bharat Ratra. The cosmological constant and dark energy. *Reviews of modern physics*, 75(2):559, 2003.

[53] J Schechter and José WF Valle. Neutrino masses in SU(2) U(1) theories. *Physical Review D*, 22(9):2227, 1980.

[54] Michael Dine and Alexander Kusenko. Origin of the matter-antimatter asymmetry. *Reviews of Modern Physics*, 76(1):1, 2003.

[55] Wolfgang Altmannshofer and David M Straub. Implications of b to s measurements. *arXiv preprint arXiv:1503.06199*, 2015.

[56] Piotr Zenczykowski. *Elementary particles and emergent phase space*. World Scientific, 2013.

[57] Francesco Vissani. Do experiments suggest a hierarchy problem? *Physical Review D*, 57(11):7027, 1998.

[58] David JE Callaway. Triviality pursuit: can elementary scalar particles exist? *Physics Reports*, 167(5):241–320, 1988.

[59] Thomas Mannel. Theory and phenomenology of CP violation. *Nuclear Physics-Section B-PS-Proceedings Supplements*, 167:115–119, 2007.

[60] JC Kapteyn. First attempt at a theory of the arrangement and motion of the sideral system. *Contributions from the Mount Wilson Observatory/Carnegie Institution of Washington*, 230:1–27, 1922.

[61] Jan H Oort. The force exerted by the stellar system in the direction perpendicular to the galactic plane and some related problems. *Bulletin of the Astronomical Institutes of the Netherlands*, 6:249, 1932.

[62] Fritz Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica physica acta*, 6:110–127, 1933.

[63] Edvige Corbelli and Paolo Salucci. The extended rotation curve and the dark matter halo of M33. *Monthly Notices of the Royal Astronomical Society*, 311(2):441–447, 2000.

[64] Sungwook E Hong, Donghui Jeong, Ho Seong Hwang, and Juhan Kim. Revealing the local cosmic web from galaxies by deep learning. *The Astrophysical Journal*, 913(1):76, 2021.

[65] Gary Hinshaw, JL Weiland, RS Hill, N Odegard, D Larson, CL Bennett, J Dunkley, B Gold, MR Greason, N Jarosik, et al. Five-year Wilkinson microwave anisotropy probe* observations: data processing, sky maps, and basic results. *The Astrophysical Journal Supplement Series*, 180(2):225, 2009.

[66] Jaan Einasto. Dark matter. *Brazilian Journal of Physics*, 43(5):369–374, 2013.

[67] Patrick Tisserand, L Le Guillou, C Afonso, JN Albert, J Andersen, R Ansari, É Aubourg, P Bareyre, JP Beaulieu, X Charlot, et al. Limits on the Macho content of the galactic halo from the EROS-2 survey of the magellanic clouds. *Astronomy & Astrophysics*, 469(2):387–404, 2007.

[68] David S Graff and Katherine Freese. Analysis of a Hubble space telescope search for red dwarfs: limits on baryonic matter in the galactic halo. *The Astrophysical Journal Letters*, 456(1):L49, 1996.

[69] Gerard Jungman, Marc Kamionkowski, and Kim Griest. Supersymmetric dark matter. *Physics Reports*, 267(5-6):195–373, 1996.

[70] Arno A Penzias and Robert Woodrow Wilson. A measurement of excess antenna temperature at 4080 Mc/s. *The Astrophysical Journal*, 142:419–421, 1965.

[71] R Cowsik and J McClelland. Gravity of neutrinos of non-zero mass in astrophysics. *The Astrophysical Journal*, 180:7–10, 1973.

[72] AD Chernin. The rest mass of primordial neutrinos, and gravitational instability in the hot universe. *Astronomicheskii Zhurnal*, 58:25–28, 1981.

[73] Katherine Garrett and Gintaras Duda. Dark matter: A primer. *Advances in Astronomy*, 2011, 2011.

[74] Roberto D Peccei and Helen R Quinn. Constraints imposed by CP conservation in the presence of pseudoparticles. *Physical Review D*, 16(6):1791, 1977.

[75] Alexey Boyarsky, M Drewes, T Lasserre, S Mertens, and O Ruchayskiy. Sterile neutrino dark matter. *Progress in Particle and Nuclear Physics*, 104:1–45, 2019.

[76] Marco Fabbrichesi, Emidio Gabrielli, and Gaia Lanfranchi. The dark photon. *arXiv preprint arXiv:2005.01515*, 2020.

[77] Risa H Wechsler and Jeremy L Tinker. The connection between galaxies and their dark matter halos. *Annual Review of Astronomy and Astrophysics*, 56:435–487, 2018.

[78] Phillip Urquijo. Searching for dark matter at the Stawell underground physics laboratory. In *EPJ Web of Conferences*, volume 123, page 04002. EDP Sciences, 2016.

[79] DS Akerib, MS Armel-Funkhouser, MJ Attisha, CN Bailey, L Baudis, Daniel A Bauer, PL Brink, R Bunker, B Cabrera, David O Caldwell, et al. Exclusion limits on the WIMP-nucleon cross section from the first run of the cryogenic dark matter search in the Soudan underground laboratory. *Physical Review D*, 72(5):052009, 2005.

[80] J Hall. The SNOLAB underground laboratory. In *Journal of Physics: Conference Series*, volume 1468, page 012252. IOP Publishing, 2020.

[81] L Paoluzi. The Gran Sasso national laboratory. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 279(1-2):133–136, 1989.

[82] A Bettini. The Canfranc underground laboratory (lsc). *The European Physical Journal Plus*, 127(9):1–7, 2012.

[83] Dennis Normile. Chinese scientists hope to make deepest, darkest dreams come true, 2009.

[84] Michael Klasen, Martin Pohl, and Günter Sigl. Indirect and direct search for dark matter. *Progress in Particle and Nuclear Physics*, 85:1–32, 2015.

[85] Russell Kirk. *Dark Matter Genesis*. PhD thesis, Royal Holloway, University of London, 2017.

[86] John Ellis, RA Flores, Katherine Freese, S Ritz, D Seckel, and Joseph Silk. Cosmic ray constraints on the annihilations of relic particles in the galactic halo. *Physics Letters B*, 214(3):403–412, 1988.

[87] FW Stecker, SD Hunter, and DA Kniffen. The likely cause of the EGRET GeV anomaly and its implications. *Astroparticle Physics*, 29(1):25–29, 2008.

[88] WB Atwood, Aous A Abdo, Markus Ackermann, W Althouse, B Anderson, M Axelsson, Luca Baldini, J Ballet, DL Band, Guido Barbiellini, et al. The large area telescope on the Fermi gamma-ray space telescope mission. *The Astrophysical Journal*, 697(2):1071, 2009.

[89] M Aguilar, Giovanni Alberti, B Alpat, A Alvino, G Ambrosi, K Andeen, H Anderhub, L Arruda, P Azzarello, A Bachlechner, et al. First result from the alpha magnetic spectrometer on the international space station: precision measurement of the positron fraction in primary cosmic rays of 0.5–350 GeV. *Physical Review Letters*, 110(14):141102, 2013.

[90] Gordon Kane and Scott Watson. Dark matter and LHC: What is the connection? *Modern Physics Letters A*, 23(26):2103–2123, 2008.

[91] Jessica Goodman, Masahiro Ibe, Arvind Rajaraman, William Shepherd, Tim MP Tait, and Hai-Bo Yu. Constraints on dark matter from colliders. *Physical Review D*, 82(11):116010, 2010.

[92] E Dudas, Y Mambrini, S Pokorski, and A Romagnoni. Invisible z' and dark matter. *Journal of High Energy Physics*, 2009(08):014, 2009.

[93] Jessica Goodman and William Shepherd. LHC bounds on UV-complete models of dark matter. *arXiv preprint arXiv:1111.2359*, 2011.

[94] Haipeng An, Xiangdong Ji, and Lian-Tao Wang. Light dark matter and z' dark force at colliders. *Journal of High Energy Physics*, 2012(7):1–25, 2012.

[95] Mads T Frandsen, Felix Kahlhoefer, Anthony Preston, Subir Sarkar, and Kai Schmidt-Hoberg. LHC and TeVatron bounds on the dark matter direct detection cross-section for vector mediators. *Journal of High Energy Physics*, 2012(7):1–26, 2012.

[96] Enrico Morgante. Simplified dark matter models. *Advances in High Energy Physics*, 2018, 2018.

[97] JW Moffat. Ultraviolet complete quantum field theory and particle model. *The European Physical Journal Plus*, 134(9):1–15, 2019.

[98] Jonathan L Feng, Manoj Kaplinghat, Huitzu Tu, and Hai-Bo Yu. Hidden charged dark matter. *Journal of Cosmology and Astroparticle Physics*, 2009(07):004, 2009.

[99] James D Wells. How to find a hidden world at the large hadron collider. *arXiv preprint:0803.1243*, 2008.

[100] James M Cline, Zuowei Liu, and Wei Xue. Millicharged atomic dark matter. *Physical Review D*, 85(10):101302, 2012.

[101] James M Cline, Zuowei Liu, Guy D Moore, and Wei Xue. Composite strongly interacting dark matter. *Physical Review D*, 90(1):015023, 2014.

[102] Clifford Cheung, Joshua T Ruderman, Lian-Tao Wang, and Itay Yavin. Kinetic mixing as the origin of a light dark-gauge-group scale. *Physical Review D*, 80(3):035008, 2009.

[103] David Curtin, Rouven Essig, Stefania Gori, and Jessie Shelton. Illuminating dark photons with high-energy colliders. *Journal of High Energy Physics*, 2015(2):1–45, 2015.

[104] D0 Collaboration et al. Search for NMSSM Higgs bosons in the h to aa to mumu mumutau tau channels using pp collisions at 1. 96 TeV. *Phys. Rev. Lett*, 103:061801, 2009.

[105] Vardan Khachatryan, Albert M Sirunyan, Armen Tumasyan, Wolfgang Adam, E Asilar, Thomas Bergauer, Johannes Brandstetter, Erica Brondolin, Marko Dragicevic, Janos Erö, et al. A search for pair production of new light bosons decaying into muons. *Physics Letters B*, 752:146–168, 2016.

[106] Serguei Chatrchyan, V Khachatryan, AM Sirunyan, A Tumasyan, W Adam, E Aguilo, T Bergauer, M Dragicevic, J Eroe, C Fabjan, et al. Search for a non-standard-model Higgs boson decaying to a pair of new light bosons in four-muon final states. *Physics Letters B*, 726(4-5):564–586, 2013.

[107] J Pivarski, A Safonov, and A Tatarinov. Search for collimated groups of muons. *CMS Note*, 2010:462, 2010.

[108] CMS Collaboration. A search for pair production of new light bosons decaying into muons in proton-proton collisions at 13 TeV. *Physics Letters B*, 796:131–154, September 2019. arXiv: 1812.00380.

[109] Lawrence Lee, Christian Ohm, Abner Soffer, and Tien-Tien Yu. Collider searches for long-lived particles beyond the standard model. 2018.

[110] Bob Holdom. Two U (1)'s and epsilon charge shifts. *Phys. Lett., B*, 166(2):196–198, 1986.

[111] James D. Wells. How to Find a Hidden World at the Large Hadron Collider. pages 283–298, 3 2008.

[112] J Aguilar-Saavedra, J Casas, J Quilis, and R Ruiz de Austri. Multilepton dark matter signals. *Journal of High Energy Physics*, 2020(1911.03486):1–24, 2020.

[113] Daniel Abercrombie et al. Dark Matter Benchmark Models for Early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum. *Phys. Dark Univ.*, 27:100371, 2020.

[114] Fady Bishara, Roberto Contino, and Juan Rojo. Higgs pair production in vector-boson fusion at the LHC and beyond. *Eur. Phys. J. C*, 77(7):481, 2017.

[115] John F. Gunion, Howard E. Haber, Gordon L. Kane, and Sally Dawson. *The Higgs Hunter's Guide*, volume 80. 2000.

[116] Johan Alwall, Claude Duhr, Benjamin Fuks, Olivier Mattelaer, Deniz Gizem Öztürk, and Chia-Hsien Shen. Computing decay rates for new physics theories with Feynrules and MadGraph 5_amc@ nlo. *Computer Physics Communications*, 197:312–323, 2015.

[117] S Myers and Emilio Picasso. The design, construction and commissioning of the CERN Large Electron–Positron collider. *Contemporary Physics*, 31(6):387–403, 1990.

[118] M Benedikt, P Collier, V Mertens, J Poole, and K Schindl. LHC design report, volume III, the LHC injector chain. *CERN, Geneva*, 2004.

[119] CMS collaboration et al. CMS physics technical design report, volume II: physics performance. *Journal of Physics G: Nuclear and Particle Physics*, 34(6):995, 2007.

[120] WW Armstrong, W Burris, DM Gingrich, P Green, LG Greeniaus, JC Hewlett, L Holm, JW Mcdonald, S Mullin, WC Olsen, et al. Atlas: Technical proposal for a general-purpose pp experiment at the large hadron collider at cern. 1994.

[121] H Dijkstra, Tatsuya Nakada, Hans Jürgen Hilke, and Thomas Ypsilantis. LHCb letter of intent, LHCb collaboration. Technical report, CERN-LHCb-95-001, 1995.

[122] W Kienzle, M Oriunno, Anne Laure Perrot, S Weisz, M Bozzo, A Buzzo, M Macri, A Santroni, G Settle, M Buenerd, et al. TOTEM, total cross section, elastic scattering and diffraction dissociation at the LHC: technical proposal. 1999.

[123] ALICE Collaboration et al. The ALICE experiment at the cern LHC. *Jinst*, 3(420):S08002, 2008.

[124] Lyndon Evans and Philip Bryant. LHC machine. *Journal of instrumentation*, 3(08):S08001, 2008.

[125] CMS collaboration et al. CERN/LHCC 94-38. *LHCC/P*, 1, 1994.

[126] Martti Pimiä. Compact muon solenoid. Technical report, CERN, 1990.

[127] Serguei Chatrchyan, Vardan Khachatryan, Albert M Sirunyan, Armen Tumasyan, Wolfgang Adam, Ernest Aguilo, Thomas Bergauer, M Dragicevic, J Erö, C Fabjan, et al. Observation of a new boson at a mass of 125 GeV with the cms experiment at the LHC. *Physics Letters B*, 716(1):30–61, 2012.

[128] Torsten Wildschek. *Design and simluation of the CMS first level muon trigger track finder*. CERN, 1998.

[129] Paolo Azzurri. The CMS silicon strip tracker. In *Journal of Physics: Conference Series*, volume 41, page 011. IOP Publishing, 2006.

[130] GL Bayatian, A Korablev, A Soha, O Sharif, M Chertok, W Mitaroff, F Pauss, V Genchev, M Wensveen, V Lemaitre, et al. CMS physics: Technical design report volume 1: Detector performance and software. Technical report, CMS-TDR-008-1, 2006.

[131] CMS collaboration et al. Performance and operation of the CMS electromagnetic calorimeter. *Journal of Instrumentation*, 5(03):T03010, 2010.

[132] Torbjörn Sjöstrand, Stefan Ask, Jesper R Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O Rasmussen, and Peter Z Skands. An introduction to pythia 8.2. *Computer physics communications*, 191:159–177, 2015.

[133] CMS collaboration et al. The hadron calorimeter technical design report. *CERN/LHCC*, 97:031, 1997.

[134] Albert M Sirunyan, CMS Collaboration, et al. Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at 13 TeV. 2018.

[135] CMS collaboration et al. Performance of the CMS level-1 trigger in proton-proton collisions at 13 TeV. *arXiv preprint arXiv:2006.10165*, 2020.

[136] Sirunyan et al. Particle-flow reconstruction and global event description with the CMS detector. *Journal of Instrumentation*, page 86, 2017.

[137] RICH LHCb. Technical design report. *CERN/LHCC*, 11:2001, 2001.

[138] Rudolf Fruhwirth. Application of Kalman filtering to track and vertex fitting. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 262(2-3):444–450, 1987.

[139] CMS collaboration et al. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *Journal of Instrumentation*, 9(10):P10009, 2014.

[140] Johan Alwall, R Frederix, S Frixione, V Hirschi, Fabio Maltoni, Olivier Mattelaer, H-S Shao, T Stelzer, P Torrielli, and M Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7):1–157, 2014.

[141] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. Madgraph 5: going beyond. *Journal of High Energy Physics*, 2011(6):1–40, 2011.

[142] Sea Agostinelli, John Allison, K al Amako, John Apostolakis, H Araujo, Pedro Arce, Makoto Asai, D Axen, Swagato Banerjee, GJNI Barrand, et al. Geant4—a simulation toolkit. *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.

[143] Adam Alloul, Neil D Christensen, Céline Degrande, Claude Duhr, and Benjamin Fuks. Feynrules 2.0—a complete toolbox for tree-level phenomenology. *Computer Physics Communications*, 185(8):2250–2300, 2014.

[144] Olivier Mattelaer and Eleni Vryonidou. Dark-matter production through loop-induced processes at the LHC: the s-channel mediator case. *The European Physical Journal C*, 75(9):1–12, 2015.

[145] Mehdi Rahmani, Marcus Hohlmann, W Shi, S Dildick, T Kamon, P Padley, A Safonov, T Ekafrawy, H Kim, and A Castaneda. Vector-portal search for long lived dark matter particles. *Bulletin of the American Physical Society*, 65, 2020.

[146] Mehdi Rahmani and Marcus Hohlmann. Vector-portal search for dark matter particles. *Bulletin of the American Physical Society*.

[147] Celine Degrande, Claude Duhr, Benjamin Fuks, David Grellscheid, Olivier Mattelaer, and Thomas Reiter. UFO–the universal Feynrules output. *Computer Physics Communications*, 183(6):1201–1214, 2012.

[148] Guido Altarelli and Giorgio Parisi. Asymptotic freedom in parton language. *Nuclear Physics B*, 126(2):298–318, 1977.

[149] Yuri L Dokshitzer. Calculation of the structure functions for deep inelastic scattering and ee annihilation by perturbation theory in quantum chromodynamics. *Zh. Eksp. Teor. Fiz*, 73:1216, 1977.

[150] Cyrille Marquet, Bo-Wen Xiao, and Feng Yuan. Semi-inclusive deep inelastic scattering at small-x. *Physics Letters B*, 682(2):207–211, 2009.

[151] S Alekhin, J Blümlein, and S Moch. The ABM parton distributions tuned to LHC data. *Physical Review D*, 89(5):054028, 2014.

[152] H Abramowicz, I Abt, L Adamczyk, M Adamus, R Aggarwal, S Antonelli, P Antonioli, A Antonov, M Arneodo, V Aushev, et al. Measurement of high-q2 charged current deep inelastic scattering cross sections with a longitudinally polarised positron beam at HERA. *The European Physical Journal C*, 70(4):945–963, 2010.

[153] Richard D Ball, Valerio Bertone, Stefano Carrazza, Luigi Del Debbio, Stefano Forte, Patrick Groth-Merrild, Alberto Guffanti, Nathan P Hartland, Zahari Kassabov, José I Latorre, et al. Parton distributions from high-precision collider data. *The European Physical Journal C*, 77(10):1–75, 2017.

[154] Johan Alwall, Stefan Höche, Frank Krauss, Nils Lavesson, Leif Lönnblad, F Maltoni, ML Mangano, M Moretti, CG Papadopoulos, F Piccinini, et al. Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions. *The European Physical Journal C*, 53(3):473–500, 2008.

[155] Johan Alwall, Alessandro Ballestrero, Paolo Bartalini, Sergey Belov, E Boos, Andy Buckley, Jonathan M Butterworth, L Dudko, Stefano Frixione, L Garren, et al. A standard format for Les Houches Event files. *Computer Physics Communications*, 176(4):300–304, 2007.

[156] Bryan Webber. Parton shower Monte Carlo event generators. *Scholarpedia*, 6(12):10662, 2011.

[157] Andy Buckley, Jonathan Butterworth, Stefan Gieseke, David Grellscheid, Stefan Höche, Hendrik Hoeth, Frank Krauss, Leif Lönnblad, Emily Nurse, Peter Richardson, et al. General-purpose event generators for LHC physics. *Physics Reports*, 504(5):145–233, 2011.

[158] Albert M Sirunyan, Armen Tumasyan, Wolfgang Adam, Federico Ambrogi, Ece Asilar, Thomas Bergauer, Johannes Brandstetter, Marko Dragicevic, Janos Erö, A Escalante Del Valle, et al. Extraction and validation of a new set of CMS PYTHIA 8 tunes from underlying-event measurements. *The European Physical Journal C*, 80(1):1–47, 2020.

[159] Bo Andersson and Part Andersson. *The lund model*. Number 7. Cambridge University Press, 1998.

[160] Rick Field. Min-bias and the underlying event at the LHC. *arXiv preprint arXiv:1110.5530*, 2011.

[161] M Hildreth. A new pileup mixing framework for CMS. In *Proceedings of this conference CHEP2015*, 2015.

[162] Tom Melia, Paolo Nason, Raoul Röntsch, and Giulia Zanderighi. W+ W-, WZ and ZZ production in the POWHEG BOX. *Journal of High Energy Physics*, 2011(11):1–22, 2011.

[163] John M Campbell, R Keith Ellis, and Ciaran Williams. Vector boson pair production at the LHC. *Journal of High Energy Physics*, 2011(7):1–36, 2011.

[164] John Campbell and Tobias Neumann. Precision phenomenology with MCFM. *Journal of High Energy Physics*, 2019(12):1–68, 2019.

[165] Andrei V Gritsan, Jeffrey Roskes, Ulascan Sarica, Markus Schulze, Meng Xiao, and Yaofu Zhou. New features in the JHU generator framework: constraining Higgs boson properties from on-shell and off-shell production. *Physical Review D*, 102(5):056022, 2020.

[166] CMS collaboration et al. Search for Higgs and Z boson decays to j/psi or upsilon pairs in the four-muon final state in proton-proton collisions at 13 TeV. *arXiv preprint arXiv:1905.10408*, 2019.

[167] Pierre Artoisenet, Rikkert Frederix, Olivier Mattelaer, and Robbert Rietkerk. Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations. *Journal of High Energy Physics*, 2013(3):1–19, 2013.

[168] T Speer, K Prokofiev, and R Frühwirth. Vertex fitting with the kalman filter formalism in the ORCA reconstruction program. In *CMS IN*, volume 8, page 2003, 2003.

[169] Albert M Sirunyan, Armen Tumasyan, Wolfgang Adam, Federico Ambrogi, Thomas Bergauer, Marko Dragicevic, J Ero, A Escalante Del Valle, Martin Flechl, R Fruhwirth, et al. Performance of the reconstruction and identification of high-momentum muons in proton-proton collisions at s= 13 TeV. 2020.

[170] Glen Cowan. *Statistical data analysis*. Oxford university press, 1998.

[171] Kyle Cranmer, Akira Shibata, Wouter Verkerke, Lorenzo Moneta, and George Lewis. Histfactory: A tool for creating statistical models for use with roofit and roostats. Technical report, 2012.

[172] CMS Collaboration et al. CMS luminosity measurement for the 2018 data-taking period at 13 TeV. CMS physics analysis summary CMS-pas-lum-18-002,(2019). *URL: https://cds. cern. ch/record/2676164*.

[173] CMS collaboration et al. Measurement of the inelastic proton-proton cross section at= 13 TeV. *arXiv preprint arXiv:1802.02613*, 2018.

[174] Muon ID uncertainty: CMS muon POG. https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffs2018. Accessed: 2022-01-24.

[175] CMS collaboration et al. Performance of cms muon reconstruction in pp collision events at 7 tev. *Journal of Instrumentation*, 7(10):P10002, 2012.

[176] Tongguang Cheng. Studies of Higgs boson production in the four-lepton final state at sqrt (s)= 13 TeV. Technical report, 2016.

[177] Jon Butterworth, Stefano Carrazza, Amanda Cooper-Sarkar, Albert De Roeck, Joël Feltesse, Stefano Forte, Jun Gao, Sasha Glazov, Joey Huston, Zahari Kassabov, et al. PDF4LHC recommendations for LHC run ii. *Journal of Physics G: Nuclear and Particle Physics*, 43(2):023001, 2016.

[178] Kernel density description. https://root.cern.ch/doc/master/classRooKeysPdf.html. Accessed: 2022-03-03.

[179] Kyle Cranmer. Kernel estimation in high-energy physics. *Computer Physics Communications*, 136(3):198 – 207, 2001.

[180] Jianqing Fan, Chunming Zhang, and Jian Zhang. Generalized likelihood ratio statistics and wilks phenomenon. *The Annals of statistics*, 29(1):153–193, 2001.

[181] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2):1–19, 2011.

[182] Serguei Chatrchyan et al. Observation of a new boson with mass near 125 GeV in pp collisions at 7 and 8 TeV. *JHEP*, 06:081, 2013.

[183] Procedure for the LHC Higgs boson search combination in summer 2011. 2011.

[184] Serguei Chatrchyan et al. Combined results of searches for the standard model Higgs boson in pp collisions at 7 TeV. *Phys. Lett. B*, 710:26, 2012.

[185] Souvik Das. A simple alternative to the crystal ball function. *arXiv preprint arXiv:1603.08591*, 2016.

[186] Chen-Ning Yang. Selection rules for the dematerialization of a particle into two photons. *Physical Review*, 77(2):242, 1950.

[187] Parton distributions from high-precision collider data. *Eur. Phys. J. C*, 77:663, 2017.

[188] Albert M Sirunyan et al. Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements. 2019.

[189] S. Agostinelli et al. GEANT4—a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003.

[190] A. M. Sirunyan et al. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12:P10003, 2017.

[191] Description and performance of track and primary-vertex reconstruction with the CMS tracker. *JINST*, 9:P10009, 2014.

[192] Shrihari Gopalakrishna, Sunghoon Jung, and James D Wells. Higgs boson decays to four fermions through an Abelian hidden sector. *Physical Review D*, 78(5):055002, 2008.

[193] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321, 1964.

[194] François Englert and Robert Brout. Broken symmetry and the mass of gauge vector mesons. *Physical Review Letters*, 13(9):321, 1964.

[195] Peter W Higgs. Broken symmetries and the masses of gauge bosons. *Physical Review Letters*, 13(16):508, 1964.

[196] Gerald S Guralnik, Carl R Hagen, and Thomas WB Kibble. Global conservation laws and massless particles. *Physical Review Letters*, 13(20):585, 1964.

[197] Sheldon L Glashow. Partial-symmetries of weak interactions. *Nuclear physics*, 22(4):579–588, 1961.

[198] Mark Thomson. *Modern particle physics*. Cambridge University Press, 2013.

[199] Jacobus Cornelius Kapteyn and PJ Van Rhijn. On the distribution of the stars in space especially in the high galactic latitudes. *The Astrophysical Journal*, 52:23, 1920.

[200] James H Jeans. The motions of stars in a Kapteyn universe. *Monthly Notices of the Royal Astronomical Society*, 82:122–132, 1922.

[201] Anaelle Halle, HongSheng Zhao, and Baojiu Li. A nonuniform dark energy fluid: perturbation equations. *The Astrophysical Journal Supplement Series*, 177(1):1, 2008.

[202] Yung-Fang Yu Yao and JT Kummer. Ion exchange properties of and rates of ionic diffusion in beta-alumina. *Journal of Inorganic and Nuclear Chemistry*, 29(9):2453–2475, 1967.

[203] Herbi Dreiner, Daniel Schmeier, and Jamie Tattersall. Contact interactions probe effective dark-matter models at the LHC. *EPL (Europhysics Letters)*, 102(5):51001, 2013.

[204] Serguei Chatrchyan, Vardan Khachatryan, AM Sirunyan, A Tumasyan, W Adam, T Bergauer, M Dragicevic, J Eroe, C Fabjan, M Friedl, et al. Search for light resonances decaying into pairs of muons as a signal of new physics. *Journal of High Energy Physics*, 2011(7):1–34, 2011.

[205] CERN_accelerator_complex_(cropped_2).jpeg (10197×6192).

[206] D Abbaneo, M Abbrescia, A Ahmad, W Ahmed, C Ali, PR Altieri, M Amr, I Asghar, P Aspell, Y Assran, et al. Operational experience with the GEM detector assembly lines for the CMS forward muon upgrade. *IEEE Transactions on Nuclear Science*, 65(11):2808–2816, 2018.

Appendix A

Background

Fields	Definition
B_d	$B_{d_\mu} \to s_w A_\mu - s_w c_\alpha Z_\mu + s_w s_\alpha Z_{d_\mu}$
X_d	$X_{d_\mu} \to s_\alpha Z_\mu + c_\alpha Z_{d_\mu}$
W_i	$W_{i_{\mu,1}} \to \frac{W_\mu + W_\mu^\dagger}{\sqrt{2}}$, $W_{i_{\mu,2}} \to \frac{-i(W_\mu + W_\mu^\dagger)}{\sqrt{2}}$, $W_{i_{\mu,3}} \to s_w A_\mu + c_\alpha c_w Z_\mu - c_w s_\alpha Z_{d_\mu}$
B	$B_\mu \to B_{d\mu} + \eta X_{d_\mu}$
X	$X_\mu \to \eta \epsilon X_{d_\mu}$

Table A.1: Non-physical fields and their definition for kinetic mixing model.

Lepton	**T³**	**Y**
ν_e	1/2	-1
e_L	-1/2	-1
e_R	0	-2
Quark	**T³**	**Y**
u_L	1/2	1/3
d_L	-1/2	1/3
u_R	0	4/3
d_R	0	-2/3

Table A.3: Weak iso-spin and hypercharge values.

Parameters	Definition	Description
η		$U(1)_X$ - $U(1)_Y$ mixing parameter (external)
c_w	$\frac{M_W}{M_Z}$	cos of the Weinberg angle
s_w	$\sqrt{1-{c_w}^2}$	sin of the Weinberg angle
c_α	$cos(\alpha)$	cos of mixing angle in weak sector
s_α	$sin(\alpha)$	sin of mixing angle in the weak sector
ϵ	$\frac{-1+\sqrt{1+4\eta^2}}{2\eta}$	kinetic mixing parameter

Table A.2: Parameters of the kinetic mixing model.

$$\alpha = -\frac{1}{2} ArcTan\left[\frac{2 s_w \eta}{1 - \Delta Z - {s_w}^2 \eta^2}\right] \tag{A.1}$$

$$\Delta Z = \frac{{M_X}^2}{{M_{Z_0}}^2} \tag{A.2}$$

Where M_X is X mass before mixing and M_{Z_0} is Z mass before mixing. ΔZ is referred to as the"*Ratio of scales.*"

The various parameters for a reference data point are tabulated in the following tables. The partital widths and branching ratios are calculated using $Madwidth$ tool within $Madgraph$. [116]

Appendix B

MC Simulation and Data Analysis Settings

B.1 2018

B.2 GEN-SIM

Release: CMSSW_10_2_3

```
git clone git@github.com:cms-tamu/MuJetAnalysis_Generator.git

cmsDriver.py MuJetAnalysis_Generator/Generator/python/
Pythia8GeneratorFilter_13TeV_NMSSM_Hto2Ato4mu_mH_100_mA_0p25_cfi.py \
--fileout file:output.root --mc --eventcontent RAWSIM --datatier GEN-SIM \
--conditions 102X_upgrade2018_realistic_v11 --beamspot Realistic25ns13TeVEarly2018Collision \
--step GEN,SIM --nThreads 8 --geometry DB:Extended --era Run2_2018 \
--python_filename HIG-RunIIFall18GS-00006_1_cfg.py --no_exec \
--customise Configuration/DataProcessing/Utils.addMonitoring -n 1685
```

B.3 DIGI-HLT

Release: CMSSW_10_2_5

```
cmsDriver.py step1 --filein file:input.root --fileout file:output.root \
--pileup_input "dbs:/Neutrino_E-10_gun/
RunIISummer17PrePremix-PUAutumn18_102X_upgrade2018_realistic_v15-v1/GEN-SIM-DIGI-RAW" \
--mc --eventcontent PREMIXRAW --datatier GEN-SIM-RAW
--conditions 102X_upgrade2018_realistic_v15 \
--step DIGI,DATAMIX,L1,DIGI2RAW,HLT:@relval2018 --procModifiers premix_stage2 --nThreads 8 \
--geometry DB:Extended --datamix PreMix --era Run2_2018 \
--python_filename HIG-RunIIAutumn18DRPremix-00013_1_cfg.py --no_exec \
--customise Configuration/DataProcessing/Utils.addMonitoring -n 2626
```

B.4 RECO

Release: CMSSW_10_2_5

```
cmsDriver.py step2 --filein file:input.root \
--fileout file:output.root --mc --eventcontent AODSIM \
--runUnscheduled --datatier AODSIM --conditions 102X_upgrade2018_realistic_v15 \
--step RAW2DIGI,L1Reco,RECO,RECOSIM,EI --procModifiers premix_stage2 --nThreads 8 \
--era Run2_2018 --python_filename HIG-RunIIAutumn18DRPremix-00013_2_cfg.py --no_exec \
--customise Configuration/DataProcessing/Utils.addMonitoring -n 2626
```

B.5 MINIAOD

Release: CMSSW_10_2_5

```
cmsDriver.py step1 --filein file:input.root --fileout file:output.root --mc \
--eventcontent MINIAODSIM --runUnscheduled --datatier MINIAODSIM \
--conditions 102X_upgrade2018_realistic_v15 --step PAT --nThreads 8 --geometry DB:Extended \
--era Run2_2018 --python_filename HIG-RunIIAutumn18MiniAOD-00013_1_cfg.py \
--no_exec --customise Configuration/DataProcessing/Utils.addMonitoring -n 8597
```

B.6 Analysis

Release: CMSSW_10_2_5

GlobalTag MC: 102X_upgrade2018_realistic_v19

GlobalTag Data: 102X_dataRun2_v11 (2018ABC), 102X_dataRun2_Prompt_v14 (2018D)

B.7 Conditions

beamspot: Realistic25ns13TeVEarly2018Collision

GEN-SIM: 102X_upgrade2018_realistic_v11 DR: 102X_upgrade2018_realistic_v15

B.8 Producing pileup

2018_25ns_JuneProjectionFull18_PoissonOOTPU

/Neutrino_E-10_gun/RunIISummer17PrePremix-PUFull18_102X_upgrade2018_realistic_v11-v1/GEN-SIM-DIGI-RAW

Appendix C

Implementation of Feynrules for Vector Portal Model Feynrules

The Feynrules software accepts the definition of fields, coupling constants and other parameters such as mass and particle widths, desired particle ID (PDGID), interaction Lagrangian for each vertex and the overall Lagrangian. The program uses the said information to compute the QFT Feynman rules and matrix elements.

C.1 New Gauge Group

The vector-portal model uses a new dark $U(1)_D$ Abelian gauge group:

```
U1D == { Abelian -> True, GaugeBoson -> Bp,
Charge -> Qd, CouplingConstant -> ee}
```

Where `Bp` is the B' in the main text 3.1, Qd is the dark hypercharge and the generator of the $U(1)_D$, and the coupling constant ee implies kinetic mixing. We are only considering the LO interaction matrix element calculation. This is expressed as

follows in the the Fynrule (*.fr) file:

```
M$InteractionOrderLimit = {{DMV, 4}};
```

Then, we have to provide the interaction hierarchy order with respect to the SM interaction. This information is used by the Feynrules to keep the interaction with the highest hierarchy at the LO, when there are multiple Feynman diagrams contributing to a process.

```
M$InteractionOrderHierarchy = {{QCD, 1}, {DMV, 2}, {QED, 2}};
```

Here the smaller number implies higher interaction hierarchy. The new dark sector is at the same level of interaction hierarchy as QED as it is similar to U(1) SM interaction and does not involve QCD products.

C.2 Physical Fields

Now, I define my the fields of my model. The dark scalar s_{D} is a complex scalar which means that it is a not a self-conjugated particle , i.e. not its own antiparticle.

```
S[8] == { ClassName -> SD,
SelfConjugate -> False,
Mass -> {MSD, 10.},
Width -> {WSD,1e-05},
PDG -> 5000512,
TeX -> Subscript[S,D],
FullName -> "Complex scalar DM" },
```

We set the mass to 10 GeV as a default mass of the s_{D}. The parameters of the model, however, are adjustable in the event generators.

Next, the gauge boson Z_{D} is defined as a self conjugated boson similar to the QED gauge bosons.

```
V[22] == { ClassName -> Zd,

SelfConjugate -> True,

Indices -> {},

Mass -> {MZd, 400},

Width -> {WZd, 1e-05},

PropagatorLabel -> "Zd",

PropagatorType -> Sine,

PropagatorArrow -> None,

PDG -> 1023,

FullName -> "Zd" },
```

C.3 Non-Physical Fields

The non-physical fields as defined in Tab A.1 are defined as follows:

```
V[210] == {

ClassName -> Bd,

SelfConjugate -> True,

Unphysical -> True,

Indices -> {},

Mass -> 0,

Width -> 0,

Definitions -> {Bd[mu_] :> cw A[mu] -sw ca Z[mu] + sw sa Zd[mu]}},
```

```
V[220] == {ClassName -> Xp,
SelfConjugate -> True,
Unphysical -> True,
Indices -> {},
Mass -> 0,
Width -> 0,
Definitions -> {Xp[mu_] :> sa Z[mu] + ca Zd[mu]}},

V[61] == {ClassName -> X,
SelfConjugate -> True,
Definitions -> {X[mu_] -> Eps Eta Xp[mu]},
Indices -> {},
Mass -> 0,
Unphysical -> True}
};
```

C.4 Parameters

The coupling of the s_{D} to the Z_{D} and the muons are *external* couplings:

```
gVSD == {
ParameterType -> External,
InteractionOrder -> {DMV, 1},
BlockName -> DMINPUTS,
TeX -> Subscript[g,VSD],
Description -> "SD-ZD vector coupling", Value -> 0.25 },
```

```
gSDMU == {
ParameterType -> External,
InteractionOrder -> {DMV, 1},
BlockName -> DMINPUTS,
TeX -> Subscript[g,SDMU],
Description -> "SD coupling mu",
Value -> 0.25 },
```

The parameters of the model as listed in the Tab. A.2 are defined in the following syntax. The tensorial parameters such as the below parameters are considered *internal*.

```
MZ0 == {
ParameterType -> Internal,
Value -> MZ,
Description -> "Z mass before mixing"},

MX =={
ParameterType -> Internal,
Value -> MZd,
Description -> "X mass before mixing"},

CapitalDeltaZ =={
ParameterType -> Internal,
Value -> MX^2/MZ0^2,
ParameterName -> DZ, Description -> "Ratio of scales"},
```

```
Thetaa == {
TeX -> Subscript[\[Theta], \[Alpha]],
ParameterType -> Internal,
Value -> ArcTan[-2 sw Eta/(1-sw^2 Eta^2 -CapitalDeltaZ)]/2,
ParameterName  -> alp,
Description -> "Mixing of the dark and weak sector"},
sa == {
TeX -> Subscript[s, \[Alpha]],
ParameterType -> Internal,
Value -> Sin[Thetaa],
Description -> "Sine of alp"},

ca == {
TeX -> Subscript[c, \[Alpha]],
ParameterType -> Internal,
Value -> Cos[Thetaa],
Description -> "Cosine of alp"},

Eta == {
ParameterType -> External,
BlockName -> HIDDEN,
ParameterName -> eta,
Value -> 0.01,
Description -> "U(1)X - U(1)Y mixing parameter"},
```

```
Eps == {
ParameterType -> Internal,
Value -> (Sqrt[1+4 Eta^2] - 1)/2/Eta,
ParameterName -> eps,
Description -> "kinetic mixing parameter"}
```

C.5 Lagrangians

The kinetic mixing Lagrangian of Eq. 3.1 in the Feynrules syntax manifest as the following:

```
Lmix := 1/4 (del[B[nu], mu] - del[B[mu], nu])^2
- 1/4 (del[Bd[nu], mu] - del[Bd[mu], nu])^2 +
Eps/2 (del[Bd[nu], mu] - del[Bd[mu], nu]) (del[B[nu], mu] -
del[B[mu], nu]);
```

Where the `del` signifies partial differentiation.

The last two terms of Eq. 3.4, the coupling of s_{D} to muons, are:

```
L1SDMU := mubar.(gSDMU).mu SD;
L1SDbarMU := mubar.(gSDMU).mu SDbar;
```

The third term in Eq.3.4, the s_{D} coupling to the Z_{D} is:

```
L1X := I gVSD/2 (SDbar del[SD,mu]-del[SDbar,mu] SD) Zd[mu];
```

Where I is i the imaginary unit number.

Feynrules has an issue with handling the difference between 3 and 3bar color representation. This leads to a lot of error in the QCD sector. To correct for this error

we *explicitly* define the quark-gluon interaction Lagrangian as[1]:

```
LqG := gs Ga[mu, s, r] T[a, i, j] uqbar[s, f, i].uq[r, f, j] G[mu, a] +
gs Ga[mu, s, r] T[a, i, j] dqbar[s, f, i].dq[r, f, j] G[mu, a];
```

The remaining interaction Lagrangian is the Z_{D} interaction with the SM sector through kinetic mixing. The SM particles, 6 quarks and 6 leptons, can be represented in left-handed doublets and right-handed singlets, each with 3 generations. For each of the hypercharge representation of the left-handed and right-handed quarks and leptons a vertex is defined. The resulting Lagrangian is then:

```
LBright := -2ee/cw B[mu]/2 lbar.Ga[mu].ProjP.l +  (*Y_lR=-2*)
4ee/3/cw B[mu]/2 uqbar.Ga[mu].ProjP.uq -          (*Y_uR=4/3*)
2ee/3/cw B[mu]/2 dqbar.Ga[mu].ProjP.dq;           (*Y_dR=-2/3*)
LBleft := -ee/cw B[mu]/2 vlbar.Ga[mu].ProjM.vl -  (*Y_LL=-1*)
ee/cw B[mu]/2 lbar.Ga[mu].ProjM.l  +              (*Y_LL=-1*)
ee/3/cw B[mu]/2 uqbar.Ga[mu].ProjM.uq +           (*Y_QL=1/3*)
ee/3/cw B[mu]/2 dqbar.Ga[mu].ProjM.dq ;           (*Y_QL=1/3*)
```

Lastly, the total Lagrangian is the summation of all above Lagrangins:

```
L1DM := L1X + LqG + LU1 + LBright + LBleft + L1SDMU + L1SDbarMU;
```

[1] https://bugs.launchpad.net/mg5amcnlo/+bug/1829266

Appendix D

Selection Efficiencies for Vector Portal MC Samples

Efficiencies for each cut for every s_{D} sample are shown in the tables below.

The table notes, describing each variable in each column, are as follows:

1 Total selection efficiency over generator level selection efficiency.

2 Eff. error on ϵ/α efficiency.

3 We search for new bosons with non-negligible lifetime that decays within the pixel detector volume.

4 The dimuon vertex fitting returns a probability based on the chi-square and degrees of freedom. We cut on the returned probability to achieve a preliminary model independent performance on signals.

5 Since the final state consists of oppositely charged muons as decay products of the mother boson, an algorithm is designed to pair muons and select 2 distinctive pairs as signal candidates. A diagram of the algorithm is shown in Fig. 15.

6 We use the absolute isolation of dimuons to suppress background events with muons from heavy-flavor $b\bar{b}$ and $t\bar{t}$ decays.

7 A muon from the DY process can radiate a photon which then immediately converts to a muon pair. Such events could mimic our signal. Thus, it is necessary to veto these events and meanwhile minimize the effect on the signal efficiency.

8 We use the absolute isolation of dimuons to suppress background events with muons from heavy-flavor $b\bar{b}$ and $t\bar{t}$ decays.

9 Refer to Tab. 4.1 for 2018 signal triggers.

10 We also allow at most one (i.e., ≤ 1) of the four signal muons to be a standalone (SA) muon only, meaning the muon is only reconstructed as a L2 muon and not reconstructed by the PF algorithm.

11 We define mass window cut based on signal dimuon mass spectrum resolution and cut the events that stand outside of this window.

Table D.1: Relative efficiencies for s_D model.

Ms_D	$\epsilon_{reco}/\alpha_{gen}$ [1]	Eff. error [2]	Phase 1 [3]	Good vertex [4]	2 cand. diμ [5]	Val. pix hit [6]	DY veto [7]	diμ iso. [8]	HLT acc. [9]	SA-only [10]	Mass window [11]
$MZ_D = 80$ GeV											
5	0.406	0.0449	1	1	0.425	1	0.931	0.82	0.97	0.977	0.92
15	0.45	0.0556	1	1	0.479	1	0.993	0.812	0.983	0.878	0.941
20	0.405	0.0527	1	1	0.418	1	0.983	0.835	1	0.885	0.976
25	0.407	0.0569	1	1	0.392	1	1	0.857	1	0.952	0.9
30	0.413	0.051	1	1	0.47	1	0.965	0.839	0.983	0.903	0.912
35	0.429	0.0417	1	1	0.539	1	0.986	0.839	0.994	0.926	0.926
$MZ_D = 85$ GeV											
5	0.446	0.035	1	1	0.462	1	0.955	0.883	0.985	0.981	0.925
15	0.418	0.04	1	1	0.473	1	0.991	0.842	0.989	0.892	0.939
20	0.421	0.0397	1	1	0.503	1	1	0.819	0.995	0.896	0.925
25	0.508	0.0495	1	1	0.541	1	1	0.868	0.995	0.94	0.924
30	0.422	0.038	1	1	0.514	1	0.988	0.834	0.991	0.904	0.926
35	0.417	0.0325	1	1	0.548	1	0.988	0.843	0.996	0.935	0.892
40	0.417	0.0325	1	1	0.548	1	0.988	0.843	0.996	0.935	0.892

Table D.2: Relative efficiencies for s_D model

Ms_D	$\epsilon_{reco}/\alpha_{gen}$ [1]	Eff. error [2]	Phase 1 [3]	Good vertex [4]	2 cand. diμ [5]	Val. pix hit [6]	DY veto [7]	diμ iso. [8]	HLT acc. [9]	SA-only [10]	Mass window [11]
$MZ_D = 91.1876$ GeV											
5	0.454	0.0279	1	1	0.509	0.998	0.978	0.852	0.995	0.988	0.935
15	0.437	0.0298	1	1	0.538	0.998	0.984	0.83	1	0.955	0.912
20	0.412	0.0308	1	1	0.492	1	0.997	0.843	0.997	0.939	0.91
25	0.432	0.0332	1	1	0.508	1	0.994	0.848	0.996	0.957	0.91
30	0.412	0.0332	1	1	0.493	1	0.98	0.836	0.996	0.952	0.919
35	0.41	0.026	1	1	0.522	1	0.988	0.844	0.998	0.958	0.902
40	0.419	0.0246	1	1	0.529	1	0.974	0.864	1	0.983	0.916
45	0.467	0.0345	1	1	0.533	1	0.994	0.896	1	0.983	0.931
$MZ_D = 95$ GeV											
5	0.443	0.0251	1	1	0.502	1	0.975	0.855	0.992	0.992	0.924
15	0.426	0.0262	1	1	0.522	0.998	0.988	0.876	0.991	0.965	0.902
20	0.453	0.0292	1	1	0.536	1	0.998	0.826	1	0.945	0.928
25	0.434	0.0283	1	1	0.536	1	0.998	0.846	0.995	0.952	0.901
30	0.399	0.0267	1	1	0.498	1	1	0.842	1	0.938	0.897
35	0.435	0.0249	1	1	0.528	0.998	0.988	0.872	1	0.96	0.912
40	0.447	0.0231	1	1	0.562	1	0.994	0.864	1	0.976	0.91
45	0.463	0.0254	1	1	0.577	1	0.981	0.868	1	0.98	0.916

Table D.3: Relative efficiencies for s_D model

Ms_D	$\epsilon_{reco}/\alpha_{gen}$ [1]	Eff. error [2]	Phase 1 [3]	Good vertex [4]	2 cand. diμ [5]	Val. pix hit [6]	DY veto [7]	diμ iso. [8]	HLT acc. [9]	SA-only [10]	Mass window [11]
$MZ_D = 100$ GeV											
5	0.432	0.0221	1	1	0.52	1	0.975	0.832	0.992	0.993	0.908
15	0.413	0.0221	1	1	0.531	1	0.949	0.857	0.993	0.984	0.91
20	0.435	0.0239	1	1	0.56	0.999	0.964	0.84	0.998	0.982	0.897
25	0.418	0.0241	1	1	0.522	1	0.974	0.842	0.996	0.973	0.92
30	0.421	0.0245	1	1	0.53	1	0.975	0.834	0.998	0.944	0.923
35	0.403	0.0223	1	1	0.536	1	0.963	0.85	0.998	0.972	0.879
40	0.419	0.0206	1	1	0.539	1	0.973	0.866	1	0.977	0.902
45	0.44	0.0204	1	1	0.554	1	0.989	0.862	0.999	0.98	0.916
$MZ_D = 110$ GeV											
5	0.419	0.0193	1	1	0.508	0.998	0.979	0.842	0.991	0.996	0.902
15	0.402	0.0188	1	1	0.535	1	0.992	0.804	0.994	0.974	0.916
20	0.425	0.0199	1	1	0.545	1	0.997	0.822	1	0.977	0.916
25	0.407	0.0196	1	1	0.539	1	0.998	0.841	0.999	0.968	0.867
30	0.399	0.0198	1	1	0.533	1	0.996	0.828	0.996	0.963	0.873
35	0.419	0.0196	1	1	0.548	1	0.997	0.838	0.995	0.959	0.901
40	0.425	0.0184	1	1	0.554	1	0.994	0.837	0.999	0.976	0.907
45	0.426	0.0171	1	1	0.537	1	0.996	0.858	0.999	0.971	0.924
50	0.412	0.0165	1	1	0.538	1	0.985	0.862	0.998	0.982	0.896
55	0.429	0.0154	1	1	0.559	1	0.981	0.885	1	0.987	0.884

Table D.4: Relative efficiencies for s_D model

Ms_D	$\epsilon_{reco}/\alpha_{gen}$ [1]	Eff. error [2]	Phase 1 [3]	Good vertex [4]	2 cand. diμ [5]	Val. pix hit [6]	DY veto [7]	diμ iso. [8]	HLT acc. [9]	SA-only [10]	Mass window [11]
$MZ_D = 125$ GeV											
5	0.412	0.0168	1	1	0.523	0.998	0.984	0.827	0.994	1	0.903
15	0.405	0.0123	1	1	0.534	1	0.989	0.833	0.999	0.979	0.894
20	0.421	0.0176	1	1	0.545	1	0.997	0.855	0.999	0.964	0.887
25	0.418	0.0177	1	1	0.558	1	0.999	0.814	1	0.966	0.894
30	0.391	0.0169	1	1	0.535	1	1	0.824	0.998	0.96	0.871
35	0.405	0.0172	1	1	0.534	1	0.998	0.835	0.998	0.969	0.894
40	0.419	0.0165	1	1	0.544	1	0.997	0.857	0.997	0.972	0.878
45	0.401	0.015	1	1	0.539	0.999	0.995	0.832	0.997	0.974	0.896
50	0.401	0.0147	1	1	0.544	1	0.993	0.849	0.998	0.981	0.888
55	0.416	0.0148	1	1	0.558	0.999	0.992	0.835	0.998	0.981	0.894
$MZ_D = 130$ GeV											
5	0.406	0.0157	1	1	0.528	0.998	0.985	0.815	0.996	0.996	0.91
15	0.416	0.0164	1	1	0.546	1	0.988	0.834	0.998	0.972	0.9
20	0.402	0.0164	1	1	0.542	1	0.996	0.82	0.998	0.976	0.885
25	0.383	0.0163	1	1	0.522	1	1	0.821	0.999	0.98	0.871
30	0.395	0.0166	1	1	0.528	0.999	0.998	0.838	0.999	0.973	0.867
35	0.384	0.0158	1	1	0.522	1	0.997	0.83	0.999	0.977	0.858
40	0.402	0.0156	1	1	0.526	1	0.998	0.848	0.998	0.98	0.877
45	0.406	0.0149	1	1	0.548	1	0.995	0.828	0.998	0.971	0.891
50	0.408	0.0145	1	1	0.539	1	0.993	0.845	0.998	0.978	0.893
55	0.425	0.0147	1	1	0.555	1	0.996	0.863	0.998	0.977	0.894

Table D.5: Relative efficiencies for s_D model

Ms_D	$\epsilon_{reco}/\alpha_{gen}$ [1]	Eff. error [2]	Phase 1 [3]	Good vertex [4]	2 cand. diμ [5]	Val. pix hit [6]	DY veto [7]	diμ iso. [8]	HLT acc. [9]	SA-only [10]	Mass window [11]
$MZ_D = 140$ GeV											
5	0.407	0.0154	1	1	0.53	0.999	0.988	0.827	0.993	1	0.894
15	0.394	0.0154	1	1	0.529	0.999	0.988	0.828	0.999	0.972	0.885
20	0.401	0.0154	1	1	0.538	1	0.997	0.828	0.998	0.979	0.882
25	0.401	0.0155	1	1	0.524	1	0.998	0.839	0.999	0.981	0.887
30	0.419	0.0159	1	1	0.538	1	0.999	0.859	1	0.974	0.886
35	0.414	0.0159	1	1	0.547	1	0.999	0.836	0.997	0.971	0.887
40	0.408	0.0153	1	1	0.538	1	0.999	0.829	0.999	0.982	0.899
45	0.404	0.0144	1	1	0.536	1	0.998	0.833	0.999	0.979	0.885
50	0.407	0.0139	1	1	0.55	1	0.995	0.838	0.999	0.979	0.877
55	0.413	0.0139	1	1	0.542	1	0.993	0.849	0.999	0.98	0.901
$MZ_D = 150$ GeV											
5	0.407	0.0143	1	1	0.532	0.999	0.983	0.831	0.993	0.999	0.892
15	0.405	0.0149	1	1	0.543	1	0.98	0.819	0.999	0.98	0.907
20	0.408	0.015	1	1	0.547	1	0.997	0.821	0.997	0.982	0.889
25	0.392	0.0147	1	1	0.523	1	0.999	0.836	0.999	0.981	0.871
30	0.382	0.0144	1	1	0.524	1	0.999	0.823	0.998	0.977	0.858
35	0.392	0.0145	1	1	0.539	1	0.999	0.823	0.999	0.98	0.866
40	0.385	0.0143	1	1	0.53	0.999	0.999	0.826	0.999	0.972	0.863
45	0.394	0.0139	1	1	0.545	1	0.996	0.822	0.999	0.968	0.878
50	0.411	0.0136	1	1	0.553	1	0.995	0.839	0.997	0.978	0.887
55	0.41	0.0135	1	1	0.55	0.999	0.994	0.847	1	0.974	0.887

Table D.6: Relative efficiencies for s_D model

Ms_D	$\epsilon_{reco}/\alpha_{gen}$ [1]	Eff. error [2]	Phase 1 [3]	Good vertex [4]	2 cand. diμ [5]	Val. pix hit [6]	DY veto [7]	diμ iso. [8]	HLT acc. [9]	SA-only [10]	Mass window [11]
$MZ_D = 160$ GeV											
5	0.402	0.014	1	1	0.542	0.996	0.991	0.82	0.994	1	0.878
15	0.396	0.0139	1	1	0.545	0.999	0.978	0.819	0.999	0.986	0.879
20	0.376	0.0137	1	1	0.524	1	0.996	0.816	1	0.98	0.862
25	0.395	0.0142	1	1	0.551	1	0.998	0.826	1	0.978	0.857
30	0.382	0.0142	1	1	0.529	1	0.999	0.834	0.999	0.979	0.846
35	0.389	0.0139	1	1	0.546	1	0.999	0.812	0.998	0.978	0.855
40	0.388	0.0138	1	1	0.529	1	0.996	0.827	0.998	0.988	0.863
45	0.392	0.0136	1	1	0.551	1	0.998	0.825	1	0.98	0.855
50	0.391	0.0131	1	1	0.527	1	0.998	0.836	0.997	0.98	0.881
55	0.419	0.0141	1	1	0.554	1	0.997	0.853	0.999	0.983	0.878
$MZ_D = 180$ GeV											
5	0.387	0.0128	1	1	0.529	0.998	0.988	0.818	0.993	1	0.878
15	0.382	0.013	1	1	0.516	1	0.977	0.826	0.999	0.985	0.889
20	0.394	0.0132	1	1	0.546	1	0.994	0.826	1	0.982	0.864
25	0.385	0.0133	1	1	0.53	1	0.997	0.821	1	0.982	0.869
30	0.392	0.0133	1	1	0.55	0.999	0.999	0.826	1	0.979	0.848
35	0.396	0.0136	1	1	0.557	1	0.999	0.828	0.999	0.969	0.852
40	0.386	0.013	1	1	0.527	1	0.999	0.837	0.999	0.976	0.865
45	0.392	0.013	1	1	0.548	1	1	0.825	0.999	0.972	0.865
50	0.381	0.0124	1	1	0.539	1	0.997	0.814	0.999	0.973	0.871
55	0.387	0.0122	1	1	0.53	1	0.996	0.83	0.998	0.986	0.878

Table D.7: Relative efficiencies for s_D model

Ms_D	$\epsilon_{reco}/\alpha_{gen}$ [1]	Eff. error [2]	Phase 1 [3]	Good vertex [4]	2 cand. diμ [5]	Val. pix hit [6]	DY veto [7]	diμ iso. [8]	HLT acc. [9]	SA-only [10]	Mass window [11]
$MZ_D = 190$ GeV											
5	0.39	0.0126	1	1	0.537	0.999	0.992	0.828	0.99	1	0.871
15	0.404	0.0131	1	1	0.552	0.999	0.978	0.83	0.999	0.984	0.883
20	0.392	0.0126	1	1	0.539	1	0.992	0.83	0.999	0.99	0.865
25	0.393	0.013	1	1	0.551	0.999	0.996	0.803	1	0.982	0.873
30	0.362	0.0123	1	1	0.521	1	0.997	0.816	1	0.982	0.837
35	0.383	0.0127	1	1	0.543	1	1	0.819	0.999	0.982	0.841
40	0.402	0.0131	1	1	0.553	0.999	1	0.819	1	0.983	0.873
45	0.395	0.0129	1	1	0.53	1	0.999	0.836	0.999	0.976	0.887
50	0.387	0.0125	1	1	0.536	1	0.998	0.828	0.999	0.981	0.861
55	0.4	0.0123	1	1	0.554	1	0.998	0.824	0.999	0.982	0.874
60	0.126	0.0061	1	1	0.212	1	0.983	0.835	0.999	0.921	0.772
$MZ_D = 200$ GeV											
5	0.391	0.0124	1	1	0.534	1	0.989	0.822	0.994	1	0.878
15	0.396	0.0127	1	1	0.542	1	0.979	0.82	0.997	0.985	0.886
20	0.388	0.0123	1	1	0.543	0.999	0.993	0.827	0.999	0.99	0.858
25	0.393	0.0126	1	1	0.545	1	0.994	0.824	0.999	0.985	0.863
30	0.377	0.0123	1	1	0.531	1	0.998	0.814	1	0.984	0.857
35	0.388	0.0127	1	1	0.545	1	0.999	0.838	1	0.978	0.84
40	0.383	0.0125	1	1	0.536	1	0.998	0.815	0.998	0.98	0.863
45	0.408	0.013	1	1	0.555	1	0.999	0.832	0.999	0.978	0.873
50	0.384	0.0123	1	1	0.53	1	0.998	0.822	0.999	0.981	0.867
55	0.387	0.0121	1	1	0.536	1	0.997	0.825	0.999	0.979	0.867

Appendix E

The CMS Coordinate System

Conventional coordinates: At the LHC there are four major designed interaction points (IP) where the proton-proton interactions occur. The center of the CMS detector is one of the interaction points. The origin of the coordinate system is placed on the nominal interaction point. In conventional terms, the beamline direction is marked as the z-axis and the x-y plane is the transverse to the beamline. The x-axis is the direction from the beamline pointing to the center of the LHC, the y-axis the direction from the beamline upwards, and ϕ and θ are respectively the azimuthal and the polar directions, as depicted in Fig. E.1.

Rapidity: In an accelerator such as the LHC the colliding particles (quarks, gluons) may not always have opposite charge and equal momentum, while their directions are always toward the common z-axis. This translates to a moving center-of-mass in the lab frame. In other words, the lab frame, with respect to the center-of-mass frame, is boosted with some velocity in the z direction: $v_z = \beta c$; where β is the ratio of the particles velocity to the speed of light and c is the speed of light. Therefore, the lab frame is not a preferable frame for analysis. To resolve this issue, we choose to operate within a coordinate defined by the Lorentz-invariant *rapidity* variable (Eq. E.1).

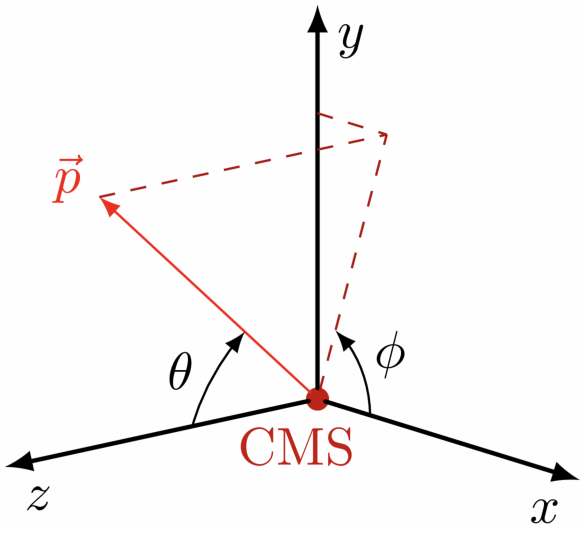


Figure E.1: The CMS coordinate system in conventional notations. The z axis represents the beamline direction.

$$y = 1/2 \ln \frac{E + p_z c}{E - p_z c} \tag{E.1}$$

This is a useful quantity as it generates values close to 0 for a particle moving in the $x - y$ plane (highly relativistic), and outputs $\to \infty$ for a particle moving down the beamline ($E \simeq p_z c$). Rapidity is often paired with the polar angle θ (the angle between the particle emission direction and the $x - y$-plane).

Pseudo-rapidity: While useful in theory, in practice, rapidity is hard to measure for highly relativistic particles. The rapidity definition demands the measurement for energy and total momentum of the particle. Regrettably, measuring the total momentum of a highly relativistic particle can be challenging. *Pseudo-rapidity*, a similar Lorentz invariant quantity, is defined to address this difficulty. The pseudo-rapidity is defined as follows,

$$\eta = -\ln tan\frac{\theta}{2}, \tag{E.2}$$

where θ is the polar angle (see Fig. E.1). It is far easier to measure pseudo-rapidity for a particle as it depends only on the polar angle. Based on Eq. E.2, for highly relativistic particles $y \simeq \eta$. Fig. E.2 demonstrates various values of η within the CMS coordinates. The conventional coordinate system is then expressed as (η, ϕ). A 3D overview of the CMS coordiate system is shown in Fig. E.3. Lastly, the standard way of expressing distance between particle tracks within CMS coordinates is by using η and ϕ and is denoted by ΔR.

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \tag{E.3}$$

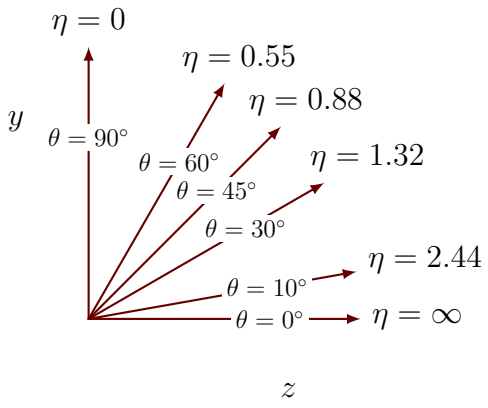


Figure E.2: Pseudorapidity representation on a 2D plane

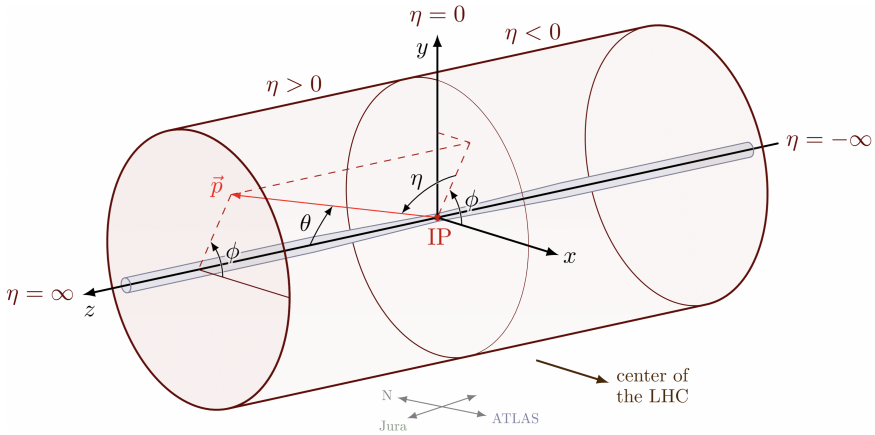


Figure E.3: The 3D overview of the CMS coordinate system