SEARCH FOR DARK MATTER DECAYING TO TWO DISPLACED MUONS PRODUCED IN PROTON-PROTON COLLISIONS AT 13 TEV WITH THE CMS DETECTOR, AND FOR DARK PHOTONS PRODUCED IN ELECTRON-POSITRON FIXED-TARGET COLLISIONS AT 500 MEV WITH THE PADME DETECTOR

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Andre Sterenberg Frankenthal

August 2020

SEARCH FOR DARK MATTER DECAYING TO TWO DISPLACED MUONS PRODUCED IN PROTON-PROTON COLLISIONS AT 13 TEV WITH THE CMS DETECTOR, AND FOR DARK PHOTONS PRODUCED IN ELECTRON-POSITRON FIXED-TARGET COLLISIONS AT 500 MEV WITH THE PADME DETECTOR

Andre Sterenberg Frankenthal, Ph.D.

Cornell University 2020

Two novel searches for dark matter using particle accelerators are presented. The first is a search for inelastically-coupled dark matter with the CMS detector at CERN, relying on 137 fb^{-1} of proton-proton collision data collected at 13 TeV center-of-mass energy between 2016 and 2018. The search strategy exploits the striking signature expected of inelastic dark matter: a pair of displaced, soft, and narrow muons collimated with missing transverse momentum and recoiled off an initial-state radiation jet. This is the first search for inelastic dark matter at a hadron collider. The second experiment is PADME, a small-scale detector to search for dark photons located in Frascati, Italy. PADME seeks to detect the production of dark photons in positron-electron collisions with a stationary diamond target and a 500 MeV positron beam. The missing-mass technique employed in the experiment relies on constraining all four-momenta in the system except for the dark photon and looking for a bump in the resulting invariant mass distribution corresponding to the dark photon's mass. The projected sensitivity for both experiments is compared in the context of highlighting the need for a comprehensive experimental search program for dark matter. Both analyses expect first public results by the end of 2020.

BIOGRAPHICAL SKETCH

Andre Sterenberg Frankenthal [redacted]. He first became interested in computer programming after being shown a cool website effect by his mom, later progressing to electronics in high school and finally landing on physics in college. His passion for physics was ignited after attending a summer school on plasma physics at the Weizmann Institute of Science in Rehovot, Israel, in July 2008. He graduated from the ORT Institute of Technology in Rio de Janeiro in December 2008.

Andre attended Reed College in Portland, Oregon, where he graduated with a Bachelor of Arts in physics in May 2013. In college, he became fascinated with dark matter after taking classes on particle physics and astrophysics and doing summer research at the Karlsruhe Institute of Technology, Germany, searching for dark matter with the EDELWEISS experiment. In his final year at Reed, he wrote a bachelor thesis on the quantum-mechanical dynamics of elementary particles under the Anti-de-Sitter central potential with Prof. Nelia Mann.

Andre subsequently began his graduate studies at Cornell University, specializing in experimental high-energy physics. After a not-so-brief stint working for the Muon g-2 experiment at Fermilab in Batavia, IL, he joined the CMS group under supervision of Prof. Jim Alexander. At CMS, he helped design and develop the next upgrade of CMS's pixel tracker, slated for installation in 2025. He ran computer simulations of several prototypes and studied the physics performance and trade-offs of several designs. He also served a leading role in test beam campaigns at Fermilab to study the performance of next-generation silicon pixel sensors. Finally, he developed from the ground up a search for inelastic dark matter, following up on his intense interest for dark matter physics.

In parallel with CMS, Andre and Prof. Alexander joined the PADME Collaboration to search for dark photons with a small-scale setup and a novel technique using a positron

beam. At PADME, Andre characterized the performance of the small-angle calorimeter with test beams at Laboratori Nazionali di Frascati in Italy, where PADME is also located. The small-angle calorimeter is a fast Cherenkov detector with strict timing requirements. Currently he is helping with the object reconstruction and data analysis efforts of the Collaboration.

Andre received his Master of Science in physics from Cornell in May 2017, and defended his dissertation on 31 July 2020, obtaining his doctoral degree in August 2020. He will next be joining the Physics department at Princeton University as a Dicke Fellow to continue his relentless search for dark matter.

To the many, many people hurt by the COVID-19 pandemic.

ACKNOWLEDGEMENTS

It is doubtful that I would have made it this far without the crucial support graciously extended to me by friends, family, and colleagues, across decades and across continents. The following acknowledgments unfortunately cover only a small subset of the people I would like to appreciate, lest this become a multiple-volume dissertation.

I would first like to thank my advisor, Jim Alexander, whom I credit almost exclusively for carrying me through the PhD. Thank you for believing in me and for taking me as a graduate student in my fourth year, when few others would have done the same. Your encouragement and support throughout these years have been of utmost vitality. Your energy and playful demeanor have made our casual chats so enjoyable and one of the absolute highlights of my time at Cornell—I will sorely miss them. Your intelligence and physics ingenuity have deepened my passion for this field and have played a large role in my decision to continue investigating dark matter into the future. Thank you for the great conversations, for the deep physics insights, and for the unwavering support. My hope as a future mentor is to emulate exactly what you have done at Cornell.

I would also like to thank Sarah How, who alongside Jim has provided an extremely solid foundation on which I could always rely for support. Your empathy and selflessness are a powerful beacon of light in these darker times and fill me with optimism that things will get better soon. Thank you for the interesting conversations, the spontaneous bits of knowledge, the entertaining Bethe House dinners, and the many fun socially-distanced pandemic gatherings.

On the CMS physics analysis side, I thank my collaborators Tres Reid, Bo Jayatilaka, Weinan Si and especially Allie Hall for helping out when I got stuck and for providing an environment conducive to fruitful and exciting work. I also immensely appreciate Kevin McDermott's extensive help in teaching me sometimes quite basic aspects of data analysis in CMS. I thank Juliette Alimena and Matthew Citron for their leadership work

and help in propelling the analysis forward as conveners of the long-lived exotica group at CMS. Finally, I am very grateful for Brian Shuve's invaluable help with theoretical aspects of the analysis and for his zen-like patience to explain quite basic things about MadGraph to me.

Related to CMS hardware projects, I sincerely thank Corrinne Mills for her help, leadership, and expertise in both test beam and simulation campaigns. Similarly, I enormously appreciate Julia Thom for making the US test beam effort get off the ground. Thanks to Lorenzo Uplegger for making his deep technical expertise available to us, and for the many beers at Fermilab's Users Center. I also want to thank Steve Wagner for helping organize the test beam campaigns at Fermilab and Stefan Spanier for consistently providing support and expert advice where needed.

I would also like to acknowledge and appreciate the CMS group at Cornell, which has made the latter part of my PhD extremely enjoyable. Cornell's CMS group has the rare combination of strong physics credentials and a friendly atmosphere. In particular, I thank my fellow current and former CMS graduate students Sam Bright-Thonney, Jorge Chaves, Jen Chu, Derek Cranshaw, Abhisek Datta, Divya Gadkari, Shaun Hogan, Nathan Mirman, Kevin McDermott, Dan Quach, Tres Reid, Shao Min Tan, Zhengcheng Tao, and Margaret Zientek; and also CMS postdocs Yangyang Cheng, Jose Monroy, Joey Reichert, Aurelijus Rinkevicius, Louise Skinnari, Livia Soffi, and Rui Zou. Finally, I thank CMS professors Jim Alexander, Ritchie Patterson, Anders Ryd, Julia Thom-Levy, and Peter Wittich for steering the group in the best possible direction. Additionally, I would like to give a nod to the theorists on the fourth floor for the very helpful discussions and thoughtful lectures throughout the years, in particular Csaba Csaki, Yuval Grossman, and Maxim Perelstein.

On the PADME side, I would like to thank the entire Collaboration for an awesome experience while doing some very exciting physics. I was always welcomed with open

arms in Frascati and met some truly nice people. Particular thanks go to graduate students Georgi Georgiev, Elizabeth Long, Isabella Oceano, Federica Oliva, Rado Simeonov, and Clara Taruggi; postdoc Gabriele Piperno; and professors and researchers Paola Gianotti, Venelin Kozhuharov, and especially Paolo Valente, who has been a supportive force throughout the PhD and whose beam accelerator expertise I greatly appreciate, and Mauro Raggi, who has played such a key role in spearheading the experiment and whose unparalled physics intuition I truly admire.

I also graciously thank my special committee who always made sure I was on track and provided extremely helpful advice about many different topics spanning a much broader range than just academic concerns. In addition to Jim Alexander, I thank Peter Wittich for his deep technical expertise and for always being willing to share it with me—I learned a great deal about FPGAs from you. I thank Csaba Csáki for the many fascinating QFT and particle physics classes that taught me to think deeply about the structure of fields and particles. I still owe you half a homework which I plan to complete later this year.

I would also like to raise a glass to all my friends who made grad school a fun and exciting endeavor and who provided much-needed support in difficult times. In particular, I thank Philip Burnham, Cari Cesarotti, Brendan Faeth, Ti-Yen Lan, Kevin McDermott, Katherine Quinn, Archishman Raju, Hao Shi, and Margaret Zientek. Thank you for the many fun times and important lessons learned, and for opening my eyes to the many struggles faced by graduate students daily.

The Hans Bethe House community at Cornell has been a lovely part of my graduate school experience. It is a rare and precious opportunity to meet such a group of empathetic and supportive people. In particular, I would like to thank the Bethe staff, Emily Pace, Erica Ostermann, Andrew Hicks, and especially Julia Thom-Levy, who has been such a wonderful source of support and caring throughout graduate school and whose capacity

for empathy is unmatched. I would also like to thank my Graduate Resident Fellow friends and in particular Amanda Recupero who managed to read through this entire dissertation without throwing it away and offered extremely helpful feedback (my usage of commas has decreased by 73% in the past two months). I thank in addition the many amazing House Fellows I had the honor of meeting and interacting with. You truly help make Bethe a vibrant community. I thank in particular Michelle Smith (and Jeremy and Alasdair), Yuval Grossman, and Sarah How.

Finally, I would like to appreciate the Cornell Physics administrative staff for all their amazing work supporting the needs of graduate students and for keeping the department a smoothly-running operation—particularly Kacey Acquilano, Deb Hatfield, Katerina Malisheva, John Miner, Sue Sullivan, Monica Wesley, and Craig Wiggers. I also give big thanks to Alan Giambattista and Nick Taylor for showing me what exceptional physics teaching looks like.

I would also like to acknowledge some important people from my college days, who contributed prominently to my career and reinforced to me the excitement of physics. Nelia Mann first introduced me to the fascinating world of particle physics with a deep-dive class and later taught me how to do serious research for my senior thesis. Johnny Powell was always so supportive of me throughout college and also first introduced me to the interesting open problems in astrophysics. Thank you for the awesome conversations and for the helpful advice, dough boy.

I also thank my college and high school friends who made the Brazil-US transition that much more enjoyable to me. Thanks in particular go to Margaret Black, Zach Brown, Guilherme Bruzzi, Alex Cohen, Svetoslav Ivanov, Mikhail Lepilov, Luis Lopes, Ivan Rivera, Heidi Rivera-Whitehouse, Laura Turcanu, and Roland Wu. I also immensely appreciate Alissa, Neal, and the Keny-Guyer family for providing me a cultural link to the US, exposing me to some virtuous American values, and always generously opening

up their home to me.

Going back further in time, I would like to thank Jill Otto for almost single-handedly setting me on my path to study in the US. None of this would have been possible without your generous and selfless support in preparing college applications, editing poorly-written essays, and teaching me about the ins-and-outs of college life in the US in only a matter of months. I am grateful beyond words for your investment in me. I hope someday to pass it forward and help some clueless Brazilian high-school student in achieving their own dreams to study in the US.

Similarly, I would like to thank Hugo Malajovich for his support throughout my high school years and for providing the resources I needed to succeed in my scholarly pursuits. ORT instilled in me a passion for all things science-y and tech-y, and I credit you for your tireless quest to keep students engaged and craving knowledge. Competing in programming olympiads is just one small example of the types of transformational opportunities that ORT under your guidance provided students of all backgrounds. Thank you for your dedication, which was not in vain. Also, thank you for turning a blind eye to my many, many late arrivals in the morning—this unfortunately has not changed.

Last but far from least, I thank my parents Claudio Frankenthal and Vivian Sterenberg and my sister Cintia Sterenberg Frankenthal for the vital support throughout the decades. Despite the harsh reality of having a child and sibling leave the country so abruptly with almost no forewarning and traveling thousands of miles away from home, you have never questioned the value of my endeavors and have always provided a safe, warm environment in which to pursue my goals. Thank you for your inexorable support and for instilling in me a passion for knowledge which persists, without signs of abating anytime soon.

Don't shoot for the stars; we already know what's there. Shoot for the space in between because that's where the real mystery lies.

Vera Rubin

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ΛCDM	Λ cold dark matter	14
ADC	analog-to-digital converter	246
ALP	axion-like particle	262
AMS-02	Alpha Magnetic Spectrometer	51
BGO	bismuth germanium oxide	225
BSM	beyond standard model	263
BTF	Beam Test Facility	220
CB	Crystal-Ball	244
CESR	Cornell Electron Storage Ring	265
CMB	Cosmic Microwave Background	5
CMS	Compact Muon Solenoid	66
CSC	cathode strip chamber	101
DAFNE	Double Annular Φ Factory for Nice Experiments	219
DAQ	data acquisition system	104
DM	dark matter	49
dof	degree of freedom	25
dSA	displaced standalone	120
DT	drift tube	101
EB	ECAL barrel	97
ECAL	electromagnetic calorimeter	97
EE	ECAL endcap	97
ES	ECAL pre-shower	98
EV	electron veto	228
EWSB	electroweak symmetry breaking	18
GCE	galactic center GeV excess	50
GM	global muons	120
HAHM	Hidden Abelian Higgs Model	109
HB	HCAL barrel	99
HCAL	hadronic calorimeter	99
HE	HCAL endcap	99
HEM	HCAL endcap minus	148
HEPV	high-energy positron veto	228
HF	HCAL forward	99
HLT	high-level trigger	106
HO	HCAL outer	99

iDM	inelastic dark matter	70
INFN	Instituto Nazionali di Fisica Nucleari	219
IPC	internal pair creation	257
ISR	initial-state radiation	77
ISS	International Space Station	51
JER	jet energy resolution	138
JES	jet energy scale	138
JLab	Thomas Jefferson National Accelerator Facility	265
L1	Level-1	104
LDMX	Light Dark Matter Experiment	274
LEP	Large Electron Positron collider	85
LHC	Large Hadron Collider	69
LLP	long-lived particle	66
LNF	Laboratori Nazionali di Frascati	219
LZ	LUX-ZEPLIN	53
MC	Monte Carlo	109
MET	missing transverse momentum	108
MG	MadGraph	109
MIP	minimum-ionizing particle	101
NFW	Navarro-Frenk-White	7
PADME	Positron Annihilation into Dark Matter Experiment	215
PAMELA	Payload for Antimatter Matter Exploration and Light-nuclei Astrophysics	51
PF	particle-flow	115
PMT	photomultiplier tube	53
POT	positron on target	221
pp	proton-proton	85
PS	proton synchrotron	87
PSB	proton synchrotron booster	87
PV	positron veto	227
QCD	quantum chromodynamics	18
QE	quantum efficiency	234
RF	radio-frequency	87
RPC	resistive plate chamber	101

SAC	small-angle calorimeter	223
SCM	standard cosmological model	14
SM	standard model	17
SPS	super proton synchrotron	87
SR	signal region	157
SSB	spontaneous symmetry breaking	18
SUSY	supersymmetry	66
TEC	tracker endcap	96
TIB	tracker inner barrel	96
TID	tracker inner disk	96
TOB	tracker outer barrel	96
VEV	vacuum expectation value	23
VR	validation region	163
WIMP	weakly interacting massive particle	31

LIST OF SYMBOLS

Δ^2_R	Curvature fluctuation amplitude	15
Λ	Cosmological constant	14
$\Omega_b h^2$	Baryon density	15
$\Omega_c h^2$	Cold dark matter density	15
η	Scalar spectral index	15
t_0	Age of the universe	15
z_{re}	Redshift of reionization	15
A'_μ	Dark photon field	37
L_{int}	Integrated luminosity: $L_{\mathrm{int}} = \int L\, dt$	86
L	Instantaneous luminosity	86
$SU(2)_L$	Part of electroweak symmetry of the SM	19
$SU(2)$	Lie group formed by special unitary matrices with unit determinant in 2D	18
$SU(3)_C$	Color symmetry of the standard model	18
$SU(3)$	Lie group formed by special unitary matrices with unit determinant in 3D	18
$U(1)_D$	Dark photon symmetry similar to hypercharge	37
$U(1)_Y$	Hypercharge symmetry of the standard model	19
$U(1)$	Lie group formed by unitary matrices in 1D	18
X_0	Radiation length	233
Δ	Mass splitting between dark matter states in iDM (in GeV)	72
α_D	Charge of $U(1)_D$ symmetry in the dark sector, equivalent to fine-structure constant in EM	73
α_{EM}	Electromagnetic fine structure constant, approximately 1/137	74
χ_1	Light dark matter state in iDM	70
χ_2	Heavy dark matter state in iDM	xviii
ϵ	Kinetic mixing coefficient	36
λ	Interaction length	233
τ	Peak separation in SAC waveform	252
$\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$	Missing transverse momentum vector	139
θ_w	Weak mixing angle	38
$c\tau$	Lifetime of χ_2, the heavy dark matter state	72
m_1	Mass of light dark matter χ_1 (in GeV)	72
m_2	Mass of heavy dark matter χ_2 (in GeV)	72
$m_{A'}$	Mass of dark photon mediator (in GeV)	43
r	Ratio of separation trough (h) and second largest peak (H)	252
v_{xy}	Dimuon vertex displacement	164
y	Coupling strength between dark sector and SM defined by $y \equiv \epsilon^2 \alpha_D \left(m_1/m_{A'}\right)^4$	75

$\lvert\Delta\phi_{\mathrm{MM}}\rvert$	Azimuthal separation between $p_{\mathrm{T}}^{\mathrm{miss}}$ and dimuon pair: $\lvert\Delta\phi(p_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{dimuon})\rvert$	176
$p_{\mathrm{T}}^{\mathrm{miss}}$	Magnitude of missing transverse momentum vector	63
PbF_2	Lead fluoride	227
SF57	Highly nonlinear lead silicate	231

Part I

Dark Matter

CHAPTER 1

INTRODUCTION

We do not understand 95% of the universe. It is a striking number for a civilization so thirsty for knowledge. Humans have advanced tremendously in their understanding of modern physics and cosmology during the past century, and yet it is fair to say that most of the energy content of the universe remains unaccounted for.

A large chunk of this mystery comes in the form of the *dark matter problem*. Dark matter is so little understood that its name contains essentially the only two properties we know about: it is dark, and it has mass. Both can be inferred from the fact that dark matter interacts gravitationally with other celestial bodies. It provides the remaining mass to explain the rotation velocity profiles of observed galaxies and it also causes some interesting optical distortions of light on its way to Earth. Therefore, it must be massive.

We have been searching for dark matter for a long time. Intensive experimental and theoretical efforts since the 1980s have failed to reveal conclusive evidence about its nature. We are not much closer to figuring it out than we were 50 years ago. We have searched for it in the sky, on Earth, and in the lab, but have come up empty-handed to our great frustration and befuddlement.

Dark matter has eluded us because it is so hard to see. Gravitational force constitutes its only known interaction, which is extremely feeble compared to the other fundamental forces. We have plenty of evidence that dark matter abounds in the galaxy and that it is constantly traveling through the Earth, and yet it fails to interact with our detectors.

If we hope to have a fighting chance to uncover the mysteries of dark matter with current technology, we must postulate the existence of additional forms of interaction with

ordinary particles apart from gravity. Barring some unlikely revolutionary experimental insight, we are just not sensitive enough to probe dark matter via its gravitational coupling to ordinary matter alone.

Fortunately, there have been exciting recent developments in the study of dark matter. We realized the possibility that dark matter may couple to ordinary matter via an additional, previously unknown force. This force, popularly referred to as the *fifth force*, could provide a bridge between dark and ordinary matter. It ought to be a faint force or it would have been detected already. But if real, it could provide a fitting seam to our patchy understanding of astrophysical and cosmological structures.

In its simplest formulation, the fifth force is similar in spirit to electromagnetism. As such, it has a carrier particle (a gauge boson) that transmits the force between different particles. This carrier is called the *dark photon*. The dark photon is to the fifth force what the photon is to electromagnetism.

The dark photon, if it exists, could help solve a range of open problems in physics (which I will later discuss). Furthermore, a dark photon term is not forbidden in the Lagrangian of the standard model of particle physics. Within this Lagrangian formulation of particle physics, if a term is not explicitly forbidden by some symmetry, then there must be a strong reason *not* to include it in the Lagrangian.

There are a variety of experimental probes at our disposal in the search for a dark photon. One promising strategy is to produce dark photons in the lab with the help of particle accelerators. If dark photons couple both to dark matter and to ordinary matter, then it might be possible to smash ordinary particles together and produce dark matter from the collision. This would not be the case if dark matter interacted only gravitationally. The basic assumption in the experimental searches for dark matter

presented in this dissertation is that dark matter can indeed be produced from ordinary particle collisions via the dark photon bridge.

This dissertation is focused solely on the dark matter problem. I will revisit the available cosmological and astrophysical evidence for dark matter, discuss and compare in detail modern and complementary experimental techniques to search for it, and present two new searches based on those techniques that also rely on the hypothesis of the dark photon.

The structure of the dissertation is as follows. The remainder of Part I is dedicated to the theory of dark photons and kinetic mixing (Chapter 3), as well as to an overview of the experimental techniques used to search for dark photons, with particular emphasis on accelerator-based approaches (Chapter 4). These are complementary to direct and indirect dark matter detection efforts and provide a controlled environment that is difficult to achieve with other strategies. Two concrete approaches are discussed: fixed-target experiments and collider experiments. In Parts II and III, an example of each of these strategies is presented in the form of a search for inelastic dark matter with the CMS detector and a direct search for dark photons with the PADME experiment in Frascati, Italy. Part IV concludes with a discussion on the prospects of future dark matter searches as well as a general prognosis of the field.

CHAPTER 2

EVIDENCE FOR DARK MATTER

This chapter provides an overview of the currently available evidence for dark matter. While there is overwhelming indication of dark matter's existence, the pieces of evidence obtained so far are of an indirect nature only. There is no "visual" confirmation of dark matter precisely because it is so difficult to observe. Most of the observations done and measurements taken rely on the gravitational interaction between dark matter and other astrophysical and cosmological structures. Therefore, the goal of current research on dark matter is to find direct evidence for it and determine its composition, whether it be one or more fundamental particles (a prominent hypothesis and the one explored in this dissertation), mini-black holes, or even other more exotic phenomena.

The chapter is divided into three sections corresponding to the three main pieces of evidence for dark matter: galactic rotation curves; gravitational lensing (and the related bullet cluster observation); and the Cosmic Microwave Background (CMB). They are also organized by cosmic scale, from shortest (galaxy), to intermediate (clusters of galaxies), to largest (cosmological).

2.1 Galactic rotation curves

The measurement of galactic rotation curves is the earliest and arguably most crucial piece of evidence in support of dark matter. A rotation profile curve measures the orbital velocity of stars and gas clouds in a galaxy as a function of distance from the galaxy's center. The observed rotation curves for a majority of galaxies deviate widely from theoretical predictions based on the amount of observed luminous matter inside them.

Fig. 2.1 shows data concerning the galaxy NGC3198, a typical example of a galaxy

exhibiting large discrepancy between measured and expected rotation curves [3]. The top plot presents the surface brightness measured as a function of distance from the galaxy. These data lead directly to a prediction of the distribution of orbital speeds in the galaxy shown by the solid line in the bottom plot (after adding an estimate of the gas contribution). However, the actual observed rotation curve is shown by the dotted line, which markedly differs from the prediction. In particular, the profile remains fairly flat even at very large distances from the center.

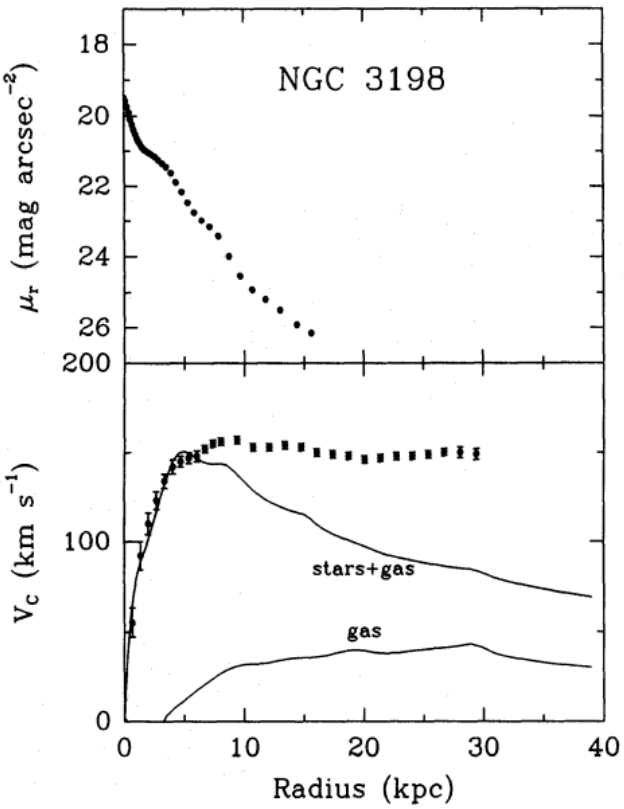


Figure 2.1: Surface brightness profile (top) and rotation velocity profile (bottom) as a function of radius from the center of the NGC3198 galaxy [3]. The dotted curves are measured data and the solid lines in the bottom plot indicate the expected rotation profile due to visible matter alone.

In fact, the flat tail distribution is a feature present in most galaxies across a diverse population of shapes, sizes, and comic locations. Fig. 2.2 highlights several observed rotation profiles, where the flat tails are evident across a large range of distances.

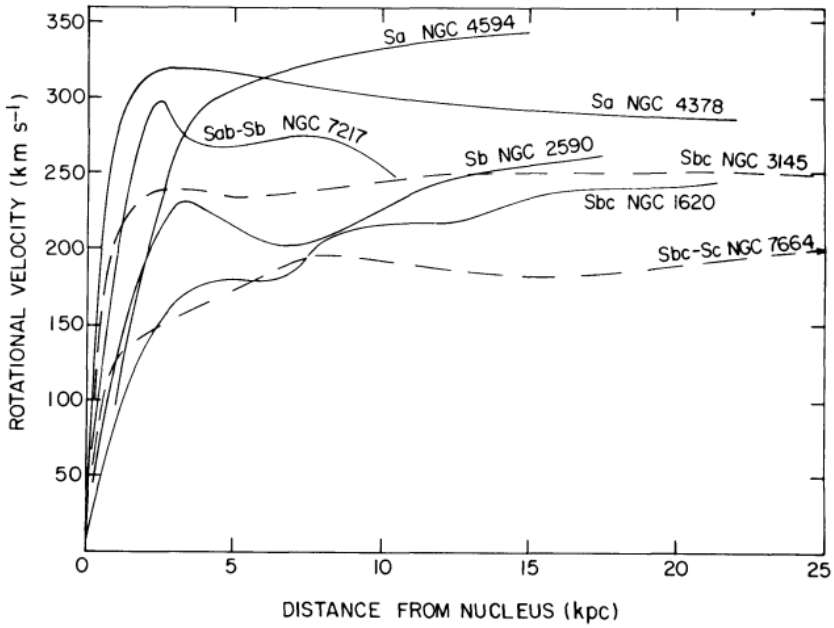


Figure 2.2: Measured rotation profiles of several astrophysical objects [4]. The unexpected flattening of the curves at large distances from the nucleus is a widespread phenomenon.

These measurements may actually be used to estimate the dark matter density distributions that would give rise to such rotation curves. With the help of N-body numerical simulations, it is possible to empirically determine the likely shape of dark matter distributions around galaxies. One prominent empirical result is the Navarro-Frenk-White (NFW) profile, characterized by the remarkably simple relation [5]:

$$\rho(r) = \frac{\rho_{crit}\,\delta_c}{(r/r_s)(1 + r/r_s)^2}, \tag{2.1}$$

where r is the distance from the center of the galaxy, r_s is a scale radius, δ_c is a dimensionless characteristic density, and $\rho_{crit} = 3H^2/8\pi G$ is the critical density necessary for cosmic closure.

NFW showed that it is possible to fit a majority of dark matter halos varying in mass and scale into this formula. The agreement is especially robust at larger distances

from a galaxy's center, but begins to deteriorate at smaller distances near the core. While profiles such as NFW predict steep increases in dark mater density very close to the center (forming a "cusp"), observations have indicated that instead the distribution seems to flatten out, leading to a flattened "core" profile [6]. This discrepancy, known as the core-cusp problem, is illustrated in Fig. 2.3. There is currently significant disagreement in the dark matter community about whether the core-cusp discrepancy is due to an incomplete understanding of astrophysical dynamics near the center of galaxies or to technical challenges in making observations and measurements in that narrow but bustling region.

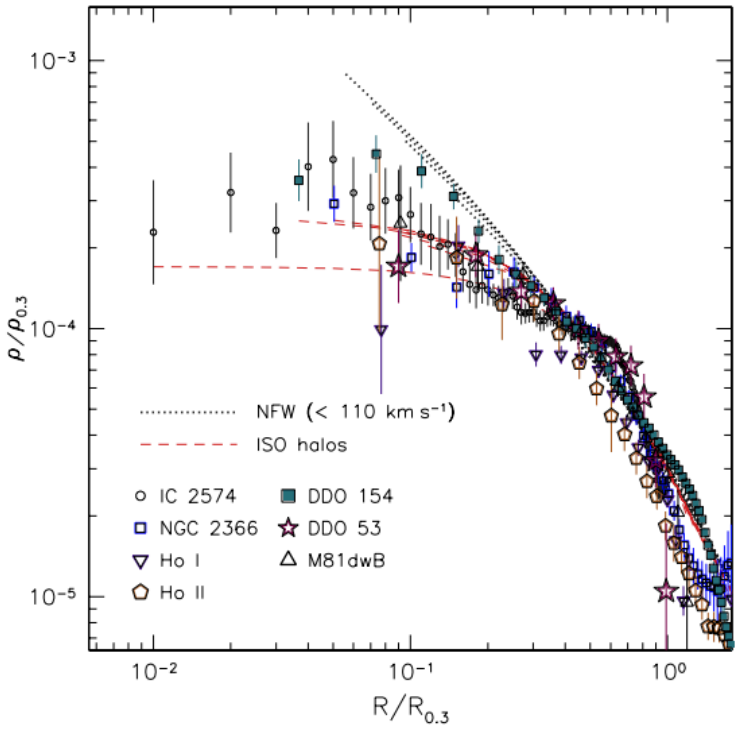


Figure 2.3: Inferred dark matter density profiles from rotation curves of several dwarf galaxies (data points) versus simulated profiles from N-body simulations (black dotted lines), near the nucleus [6]. Dwarf galaxies are known to be dominated by dark matter, making them good candidates for the study of rotation curves and dark matter distributions. The discrepancy between expected and inferred profiles illustrate the so-called core-cusp problem.

Nevertheless, there is robust evidence from galactic rotation curves that dark matter is necessary to account for the observed velocity profiles and furthermore that its distri-

bution must surround a galaxy and extend far away from the center, in fact much farther than the observed luminous matter. These are often referred to as dark matter halos.

2.2 Gravitational lensing and the bullet cluster

Since dark matter is massive, it exerts gravitational attraction. With his theory of general relativity published in 1915, Albert Einstein predicted that the gravitational force causes a corresponding curvature in four-dimensional spacetime. The curvature has several observable consequences, such as the bending of light rays traveling close to massive cosmic objects [7]. This phenomenon is known as gravitational lensing, because it is similar (though not identical) to the bending of light rays due to optical effects. Fig. 2.4 shows an observation of gravitational lensing due to a large galaxy cluster.

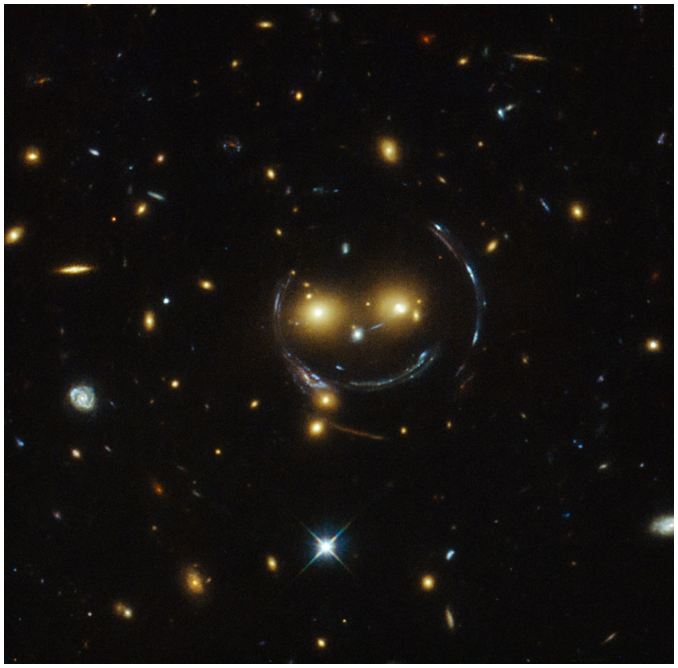


Figure 2.4: Gravitational lensing due to galaxy cluster SDSS J1038+4849 (image credit: NASA). The distorted shapes are due to light from galaxies behind the cluster getting bent by the gravitational field of the cluster, and are not an instrumentation effect.

If dark matter is massive, then it should also bend light rays just as galaxies and black holes do. This is inferred from examples such as the one in Fig. 2.5, an observation of the gravitational lensing of galaxy cluster Abell 2218. Unlike Fig. 2.4, here the distribution of visible matter is not sufficient to account for the observed gravitational lensing, so

there is an additional lensing component caused by dark matter [8].

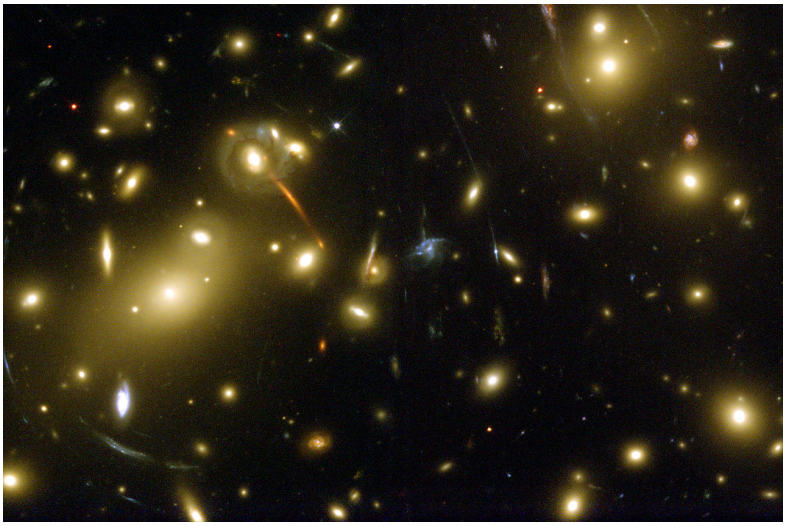


Figure 2.5: Inferred gravitational lensing due to dark matter in galaxy cluster Abell 2218 (image credit: NASA). The presence of dark matter can be determined by comparing the observed shape distortion with predictions due to visible matter alone (measured from the light emanating from the cluster) [8].

A stark example of gravitational lensing that contributed markedly to reaffirm the existence of dark matter is the bullet cluster [9]. The bullet cluster is a rare observation of two clusters of galaxies colliding. This collision can be probed both optically and via gravitational lensing, and they lead to largely discrepant conclusions. Fig. 2.6 shows the bullet cluster observed with light together with the inferred matter distribution computed from gravitational lensing measurements.

The bullet cluster is striking because it appears to have separated the dark matter halos surrounding both clusters of galaxies from the baryonic (or ordinary) matter inside. Typically, the dark matter halo is centered at the core of galaxies or cluster of galaxies, and spreads isotropically outward. But after the collision, these halos seem to have been displaced from the center, with the regular matter "lagging" behind the dark matter. Not only does this observation provide strong evidence for dark matter, it also supports the hypothesis that it is feebly-interacting, both with itself and with baryonic matter.

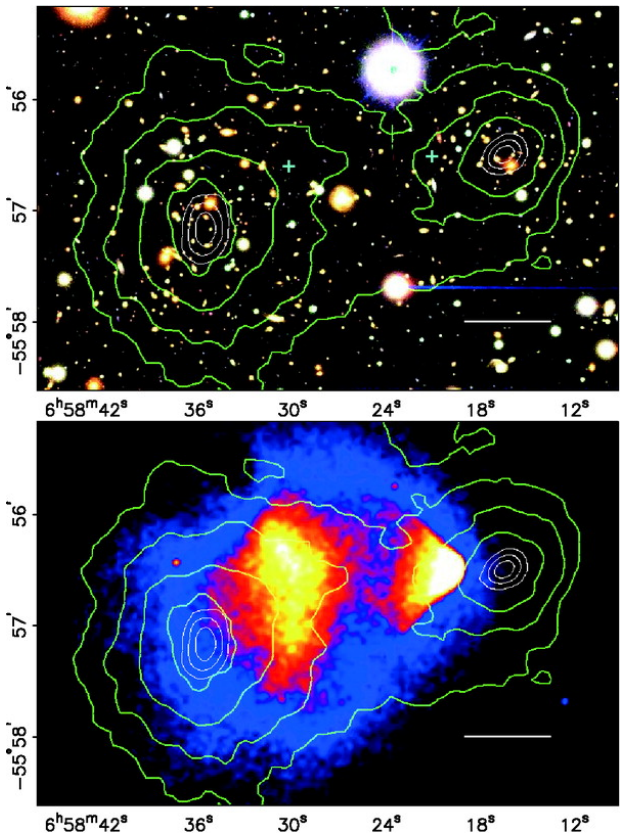


Figure 2.6: Inferred dark matter distribution from weak lensing measurements (top) and observed X-ray emission due to plasma (bottom) in galaxy cluster merger 1E0657-558, also known as the Bullet Cluster [9]. The striking discrepancy between the inferred dark matter and observed plasma distributions in this merger offers compelling evidence for dark matter.

This hypothesis would explain why the dark matter halo kept moving forward relatively unaffected by the collision while the visible matter (i.e. the gas clouds) slowed down.

2.3 Cosmic Microwave Background

In the early stage of the universe, when it was still hot and fast-expanding, it remained mostly opaque. There was so much energy that ionized hydrogen plasma permeated all of space, creating a wall that prevented photons from escaping their local environment.

When the universe finally expanded enough to cool down, protons and electrons in the plasma began to recombine, producing neutral hydrogen. This is the recombination epoch of the universe's early evolution. As neutral hydrogen began to form, the universe became transparent. At that moment, photons were allowed to escape and to roam the universe unimpeded.

Today, these first photons from the recombination epoch form a highly isotropic layer of radiation that can be measured from Earth. They have been traveling so long that their wavelengths have been extremely red-shifted, so their frequency distribution peaks in the microwave region of the electromagnetic spectrum. The photons form a faint glow of radiation uniformly spread between the starts and galaxies in the sky, known as the Cosmic Microwave Background (CMB). Fig. 2.7 shows a map of CMB temperature fluctuations in the distribution of photons first measured by the COBE Collaboration in the 1990s, and subsequently to higher precision by the WMAP (2000s) and Planck (2010s) Collaborations.

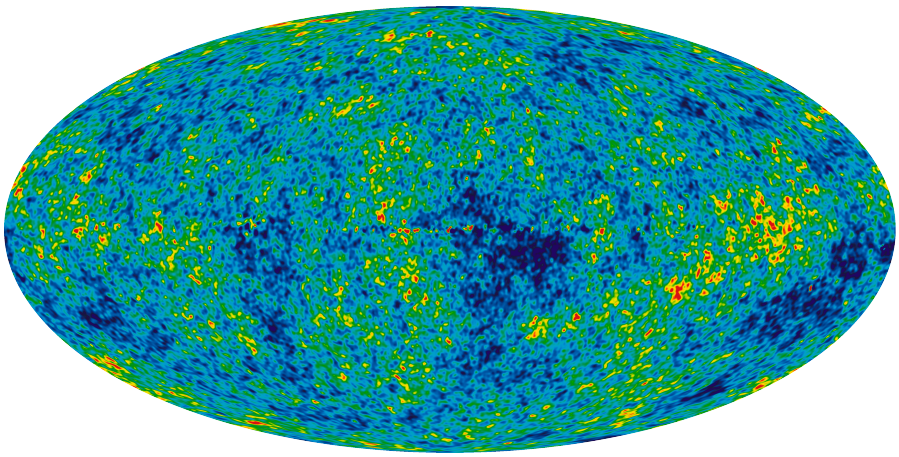


Figure 2.7: CMB temperature anisotropy map (image credit: NASA). Red and yellow spots (blue and green) indicate higher (lower) than average temperature (2.725 K). The anisotropy map corresponds to an imprint of cosmological-scale structures in the early evolution of the universe the moment before it became transparent.

Interpreted as thermal blackbody radiation, CMB photons correspond to a blackbody

emission with temperature 2.725 K [10]. This can be measured from the frequency distribution of CMB photons, shown in Fig. 2.8. Indeed the distribution agrees well with Planck's law of blackbody radiation, which predicts a varying frequency spectrum depending on the blackbody temperature.

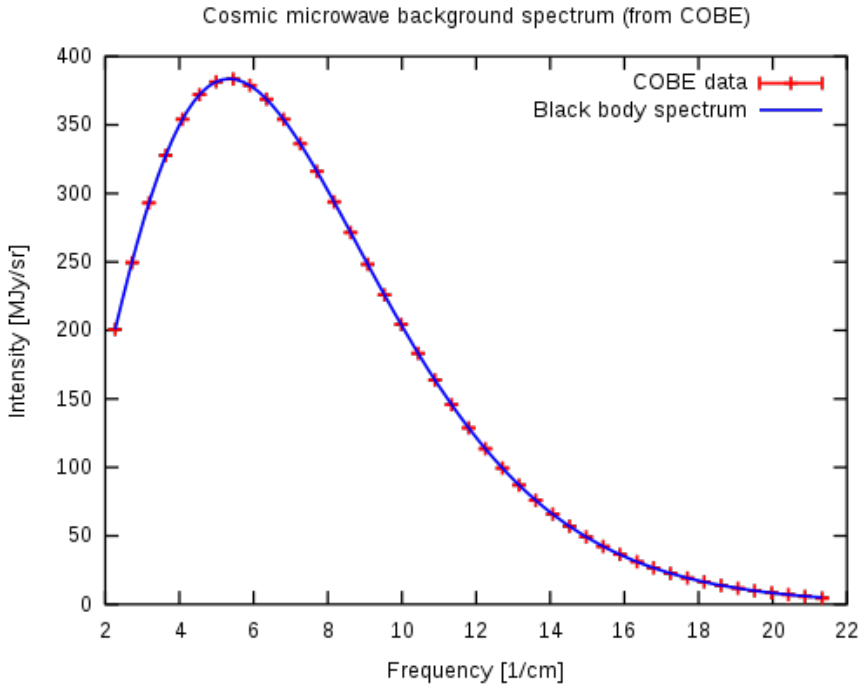


Figure 2.8: Frequency spectrum of the CMB radiation measured with COBE data (public domain). The spectrum matches precisely that expected from radiation of a blackbody with temperature 2.725 K.

The CMB spectrum has been extensively studied in past decades and has become an essential tool in our understanding of modern cosmology and astrophysics. It helped establish and solidify the standard cosmological model (SCM), particularly after the discovery of the accelerating expansion of the universe [11]. In the SCM, the universe is made up of three distinct sources: baryonic matter, cold dark matter, and a cosmological constant Λ (also known as dark energy). This model is also referred to as Λ cold dark matter (ΛCDM). These three sources of energy have dominated at different times in the evolution of the universe, but currently it is estimated to consist of roughly 69% dark

energy, 26% dark matter, and only 5% baryonic matter. Fig. 2.9 depicts the evolution of these estimates in graphical form including the latest data collected and analyzed by the Planck Collaboration [12].

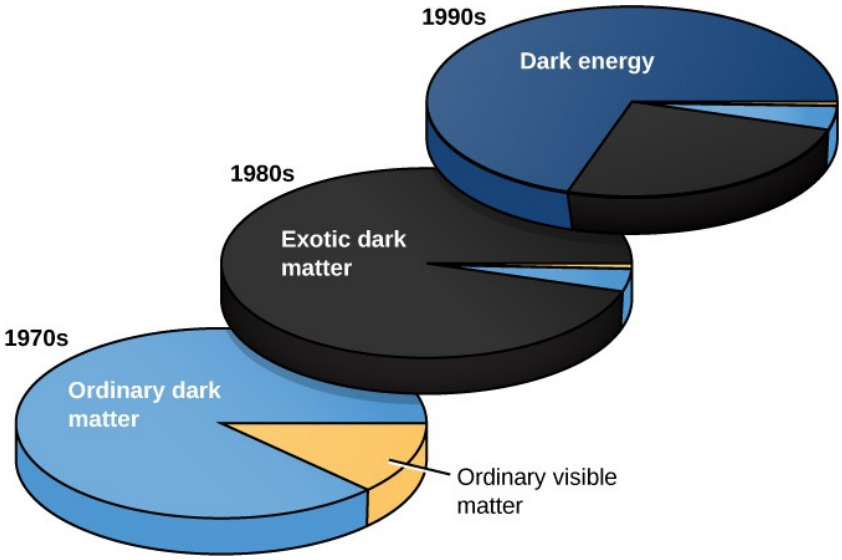


Figure 2.9: Evolution of our understanding of dark matter abundance in the universe [13]. 'Ordinary dark matter' mainly refers to black holes, while 'exotic dark matter' is the actual cosmological dark matter not well understood. (Access for free at `https://openstax.org/books/astronomy/pages/1-introduction`.)

The six base ΛCDM model parameters are [12]:[1]

- Baryon density ($\Omega_b h^2$),
- Dark matter density ($\Omega_c h^2$),
- Age of the universe (t_0),
- Redshift of reionization (z_{re}),
- Curvature fluctuation amplitude (Δ^2_R),
- Scalar spectral index (η).

[1]Many other parameters can be calculated from these, but there is also some variability in the choice of base parameters found in the literature.

The estimates of energy and matter abundance are produced by fitting the CMB data to the ΛCDM. The six parameters of the model are fitted by angular scale distributions obtained from CMB data. Fig. 2.10 shows the power spectrum of temperature fluctuations, overlaid on top of a fitted Λ model. The agreement is remarkable.

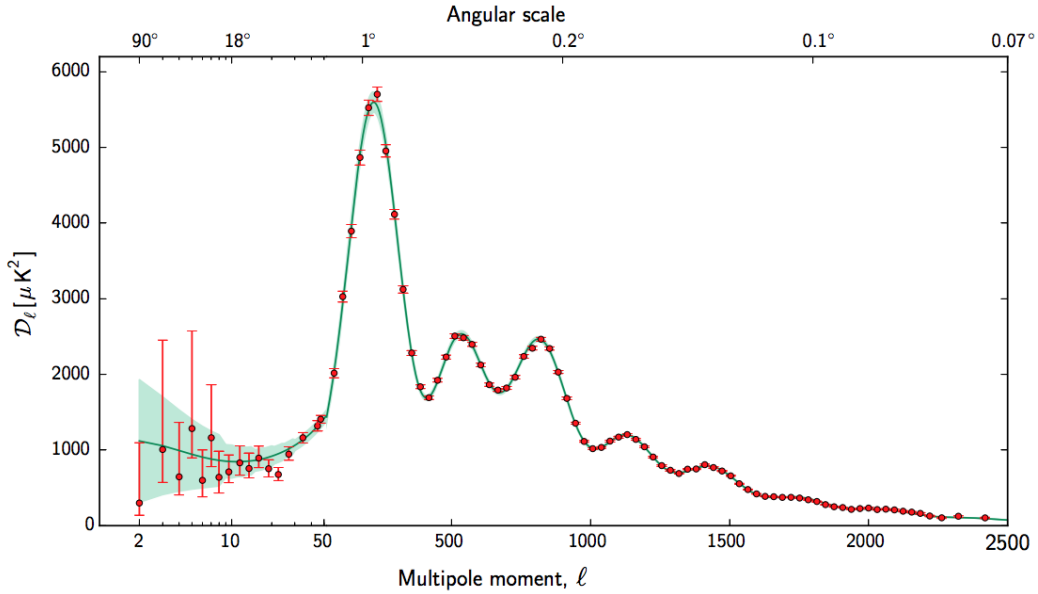


Figure 2.10: Multipole decomposition of CMB's power spectrum by the Planck Collaboration (2013) [14] (image credit: ESA and the Planck Collaboration). The data (dots) agree remarkably well with theoretical predictions based on the ΛCDM model (solid line).

From the fit, it is possible to extract the density of dark matter in the universe reported in Fig. 2.9. The best-fit value of $\Omega_c h^2$ is 0.2589 ± 0.0057 [12]. Thus the CMB provides compelling evidence for dark matter, going even further by allowing us to estimate the *amount* of dark matter in the universe.

Taking the evidence presented in this chapter into account, there is a strong case for the existence of dark matter. However, as mentioned before, none of this evidence provides direct insights into the *nature* of dark matter. The only facts we know about it are that it is massive, dark, and feebly-interacting. The elucidation of dark matter remains one of the most important open problems in twenty-first century physics and cosmology.

CHAPTER 3

DARK SECTOR AND KINETIC MIXING

This chapter introduces the theoretical framework underlying a hypothetical complex dark sector, which contains dark matter but can also include other exotic states such as dark photons, dark Higgs, and dark fermions. This framework brings our discussion of dark matter to a more rigorous footing. There are a variety of possible mediators connecting the standard model (SM) and the dark sector. These mediators are the carriers of new forces and effectively the "bridges" between dark matter and ordinary particles. I begin by summarizing the current understanding of particle physics, and then narrow the focus to specifically describe the dark photon mediator. I end by discussing the renormalizable extension of the SM that predicts the existence of dark photons.

3.1 The standard model of particle physics

The SM is the current best description of our understanding of elementary particles and fundamental forces. While still incomplete, it provides a unified framework for thinking about physics at the smallest scale. The SM is generally thought of as the lower-energy effective field theory of some currently unknown, higher-energy theory. As a phenomenological model, it both summarizes our understanding of the structure of matter and parameterizes our ignorance of more fundamental physics into a set of parameters that can be probed experimentally, allowing the extraction of information even if the higher-energy fundamental description is lacking.

The SM is built on the idea of fundamental symmetries of nature. These symmetries place tight rules on the possible shapes of interactions between particles. They also form the backbone of the spectrum of gauge bosons permitted in the SM. The local (or gauge)

symmetry of the SM is:[1]

$$SU(3)_C \times SU(2)_L \times U(1)_Y. \tag{3.1}$$

Here, the subscripts C, L and Y denote *color*, *left-handedness* (weak isospin), and *hypercharge* respectively. The numbers indicate the dimension of the fundamental representation of the corresponding Lie group: the color symmetry forms an $SU(3)$ Lie group, the weak isospin (or flavor) symmetry an $SU(2)$ Lie group, and hypercharge a single $U(1)$ Lie group.

In the SM, the strong interaction is described by quantum chromodynamics (QCD), which corresponds to the $SU(3)_C$ symmetry above. The gauge boson mediators of QCD are called gluons and there are 8 of them since they form an adjoint representation of $SU(3)$. The weak and electromagnetic interactions are described by the remaining symmetries, $SU(2)_L \times U(1)_Y$. There are 4 gauge bosons since the adjoint representations of $SU(3)$ and $U(1)$ are 3- and 1-dimensional, respectively. Notably, gravity is not included in the SM.

Eq. (3.1) is the full symmetry of the SM. However, this symmetry is broken down by the Higgs mechanism. The Higgs mechanism induces a breaking of the electroweak symmetry, which makes up the last two terms in Eq. (3.1):

$$SU(3)_C \times \left[SU(2)_L \times U(1)_Y\right] \rightarrow SU(3)_C \times U(1)_{EM}. \tag{3.2}$$

Electroweak symmetry breaking (EWSB) is an instance of a more general phenomenon known as spontaneous symmetry breaking (SSB) and in the case of the SM

[1]Unless otherwise noted, this section is adapted from Ref. [15] and Ref. [16].

it arises from the Higgs potential. This has a number of important effects in the phenomenology of the SM. Furthermore, it serves as inspiration for equivalent attempts to induce a symmetry breaking mechanism in the dark sector via a hypothetical dark Higgs [17].

The fermionic content of the SM consists of leptons and quarks. Leptons are singlets under $SU(3)_C$ rotations while quarks are not. The various charges or representations of fermions under the local symmetries of the SM are denoted in the format $(A, B)_C$, where A is the representation under $SU(3)_C$, B is the representation under $SU(2)_L$, and C is the charge under $U(1)_Y$. With this notation, the fermion fields in the SM are:

$$\begin{aligned} &Q_{Li}(3,2)_{+1/6}, \qquad U_{Ri}(3,1)_{+2/3}, \qquad D_{Ri}(3,1)_{-1/3}, \\ &L_{Li}(1,2)_{-1/2}, \qquad E_{Ri}(1,1)_{-1}, \qquad i = 1,2,3. \end{aligned} \tag{3.3}$$

There are three families of fermions and two left-handed doublets per family (Q and L). Using the first generation as example, $Q_{L1} = (\,u_L \;\; d_L\,)$ are the left-handed components of the up and down quarks, and $L_{L1} = (\,e_L \;\; \nu_{eL}\,)$ are the left-handed components of the electron and electron-neutrino respectively. Note that Q_{L1} transforms in the triplet representation of $SU(3)_C$, while L_{L1} transforms in the singlet representation (i.e. quarks participate in QCD while leptons do not). Finally, each quark (up-style and down-style) has a corresponding right-handed component which transforms in the singlet representation of $SU(2)_L$, namely U_{Ri} and D_{Ri}, whereas only charged leptons (i.e. electrons, muons, and taus) have a right-handed component, namely E_{Ri}. Neutrinos do not have a right-handed component in the SM, and are therefore predicted to be massless (an experimentally disproven prediction due to the observation of neutrino oscillations [18]).

The SM Lagrangian is organized into several *sectors*: Higgs, fermion, Yukawa, and gauge sectors. The SM Lagrangian density is written as a sum of these four sectors:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}} + \mathcal{L}_{\text{fermion}}. \tag{3.4}$$

Each sector is briefly described next.

3.1.1 Gauge sector

The gauge sector includes the kinetic terms for the gauge bosons in the SM. It also contains self-interaction terms from bosons of non-Abelian symmetries:

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}G^i_{\mu\nu}G^{\mu\nu i} - \frac{1}{4}W^i_{\mu\nu}W^{\mu\nu i} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}, \tag{3.5}$$

where G, W, and B are the field strength tensors for the three gauge symmetries: $SU(3)_C$, $SU(2)_L$, and $U(1)_Y$ respectively. The field strength tensors are defined as:

$$G^i_{\mu\nu} = \partial_\mu G^i_\nu - \partial_\nu G^i_\mu - g_s f_{ijk} G^j_\mu G^k_\nu, \quad i,j,k \in [1,8], \tag{3.6}$$

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g\epsilon_{ijk} W^j_\mu W^k_\nu, \quad i,j,k \in [1,3], \tag{3.7}$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \tag{3.8}$$

where G_μ is one of 8 gluons, W_μ is one of 3 interaction-basis weak bosons, and B_μ is the hypercharge gauge boson. Moreover, f_{ijk} are the structure constants of the $SU(3)_C$ algebra, and ϵ_{ijk} the structure constants of the $SU(2)_L$ algebra (in this case just the totally anti-symmetric tensor). Finally, g_s is the QCD strong coupling constant and g the weak coupling constant. The self-interaction terms arise from the last term in the first two equations, which couples several gauge bosons in accordance with the structure

constants. An equivalent description is that gauge bosons of non-Abelian symmetries are themselves charged under those symmetries. The structure constant of hypercharge is zero since $U(1)_Y$ is an Abelian group, so B_μ is not charged and no self-interactions arise.

3.1.2 Fermion sector

The fermion sector of the SM Lagrangian consists of kinetic terms for the fermionic content of the SM. These are just the covariant derivatives of the fermion fields:

$$\mathcal{L}_{\text{fermion}} = -i\overline{Q_{Li}}\not{D}Q_{Li} - i\overline{U_{Ri}}\not{D}U_{Ri} - i\overline{D_{Ri}}\not{D}D_{Ri} - i\overline{L_{Li}}\not{D}L_{Li} - i\overline{E_{Ri}}\not{D}E_{Ri}, \tag{3.9}$$

where the covariant derivatives are defined differently for each field based on its representations and charges under the various local symmetries:

$$\begin{aligned}
D^\mu Q_L &= \left(\partial^\mu + \frac{i}{2}g_s G_a^\mu \lambda_a + \frac{i}{2}g W_b^\mu \sigma_b + \frac{i}{6}g' B^\mu\right) Q_L, \\
D^\mu U_R &= \left(\partial^\mu + \frac{i}{2}g_s G_a^\mu \lambda_a + \frac{2i}{3}g' B^\mu\right) U_R, \\
D^\mu D_R &= \left(\partial^\mu + \frac{i}{2}g_s G_a^\mu \lambda_a - \frac{i}{3}g' B^\mu\right) D_R, \\
D^\mu L_L &= \left(\partial^\mu + \frac{i}{2}g W_b^\mu \sigma_b - \frac{i}{2}g' B^\mu\right) L_L, \\
D^\mu E_R &= \left(\partial^\mu - ig' B^\mu\right) E_R.
\end{aligned} \tag{3.10}$$

Here G^μ, W^μ, and B^μ are the gauge bosons; g_s, g, and g' are the strong, weak, and hypercharge coupling constants; and σ and λ are the Pauli and Gell-Mann matrices,

respectively. Pauli matrices are the generators of the $SU(2)$ algebra in the doublet representation, while Gell-Mann matrices are the generators of the $SU(3)$ algebra in the triplet representation.

Note that mass terms for fermions are not allowed in the SM as they would explicitly break gauge invariance. Instead, fermions acquire mass via spontaneous symmetry breaking of the Higgs field, as explained below.

3.1.3 Higgs sector

The Higgs sector contains the SM terms that relate to the Higgs. The Higgs field in the SM is a two-component complex scalar written as:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \tag{3.11}$$

which transforms in the doublet (or spinor) representation of $SU(2)_L$. In the notation of previous sections, this is $\phi(1,2)_{+1/2}$. The label on each component refers to the charge EM acquired after EWSB ($SU(2)_L \times U(1)_Y \rightarrow U(1)_{EM}$). The Higgs Lagrangian consists of the gauge-invariant kinetic term of the Higgs field plus a scalar potential:

$$\mathcal{L}_{\text{Higgs}} = \left(D^\mu \phi\right)^\dagger D_\mu \phi - V(\phi), \tag{3.12}$$

where $D_\mu \phi$ is the covariant derivative of the Higgs field:

$$D^\mu \phi = \left(\partial^\mu + \frac{i}{2} g W_a^\mu \sigma_a + \frac{i}{2} g' B^\mu\right) \phi, \tag{3.13}$$

and $V(\phi)$ is the Higgs scalar potential:

$$V(\phi) = +\mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2. \tag{3.14}$$

Here μ^2 must be a negative number in order to induce a non-zero minimum of the potential and hence a vacuum expectation value (VEV). The non-vanishing VEV of the Higgs potential is the source of SSB in the SM. The second term in Eq. (3.14) is the Higgs quartic self-interaction and λ is an empirical parameter regulating its strength.

Fig. 3.1 shows the shape of the Higgs potential when μ^2 is negative (the famous “Mexican Hat" potential).

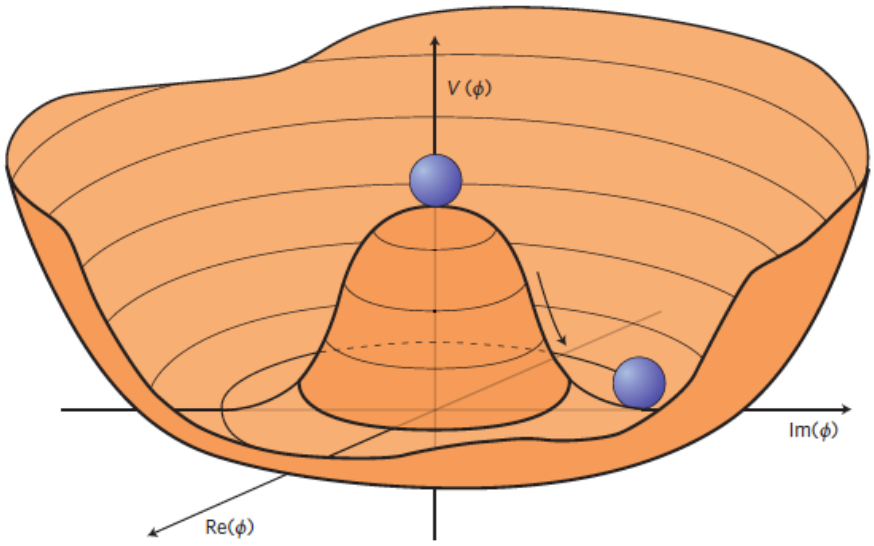


Figure 3.1: Shape of the Higgs potential $V(\phi)$ in Eq. (3.14) when $\mu^2 < 0$ is satisfied [19]. The minimum of the potential forms a ring of constant radius, the VEV. The actual location of the minimum of the field is arbitrary, giving rise to spontaneous symmetry breaking.

3.1.4 Yukawa sector

The Yukawa sector contains couplings between fermions and the Higgs field. After SSB, this gives rise to interaction terms between fermions and the Higgs boson, and also to fermion mass terms. The Lagrangian density for the Yukawa sector is:

$$\mathcal{L}_{\text{Yukawa}} = Y^u_{ij}\overline{Q_{Li}}U_{Rj}\widetilde{\phi} + Y^d_{ij}\overline{Q_{Li}}D_{Rj}\phi + Y^e_{ij}\overline{L_{Li}}E_{Rj}\phi + h.c., \tag{3.15}$$

where the indices $i, j = 1, 2, 3$ denote the fermion flavor (or family), and Y^u, Y^d, and Y^e are 3×3 complex Yukawa matrices. Finally, $\widetilde{\phi}$ is a special conjugate of the Higgs field that transforms according to $(1,2)_{-1/2}$ instead of $(1,2)_{+1/2}$ of the original Higgs field: $\widetilde{\phi}_a = \epsilon_{ab}\phi^*_b$.

The Yukawa matrices are completely general in the SM. It makes no prediction for what values they should take. These are left as parameters of the model and the actual values are determined experimentally, thereby *defining* the SM.

3.1.5 The Higgs mechanism and particle masses

When $\mu^2 < 0$ in Eq. (3.14), the Higgs potential takes the shape shown in Fig. 3.1. The minimum of the potential forms a ring of constant radius. By finding the value of the minimum of the potential, the VEV can be calculated as a function of μ and λ:

$$v = \sqrt{\frac{-\mu^2}{\lambda}}. \tag{3.16}$$

The degeneracy in values of the Higgs field that attain the minimum of the potential

means that the field, when finding its lowest-energy configuration, will arbitrarily pick a particular value along the ring that minimizes the potential. The decision of a particular choice of field value over all the others is said to spontaneously break the symmetry of the SM to the symmetry presented in Eq. (3.2). The non-zero VEV of the Higgs field is directly responsible for EWSB in the SM.

The VEV of the Higgs field and its conjugate, without loss of generality, may be written as:

$$\langle \phi \rangle = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix}, \qquad \langle \widetilde{\phi} \rangle = \begin{pmatrix} v/\sqrt{2} \\ 0 \end{pmatrix}, \tag{3.17}$$

which is equivalent to picking an arbitrary location around the ring of minimum VEV lying along the y-component of the 2D field space. It is customary to then write down the Higgs field based on this choice of VEV:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \tag{3.18}$$

Here $h(x)$ is a real-valued field and one of 4 degrees of freedom (dofs) from the original Higgs complex scalar doublet. The other 3 dofs are written as phases and then rotated away; these are incorporated as the longitudinal dofs of the three weak gauge bosons $W^{\pm}$ and Z, effectively making them massive. This can be seen by rewriting the derivative term of the Higgs field after SSB:

$$D^{\mu}\langle \phi \rangle = \frac{i}{\sqrt{8}} \left(g W_a^{\mu} \sigma_a + g' B^{\mu} \right) \begin{pmatrix} 0 \\ v \end{pmatrix} = \frac{i}{\sqrt{8}} \begin{pmatrix} g W_3^{\mu} + g' B^{\mu} & g \left(W_1^{\mu} - i W_2^{\mu} \right) \\ g \left(W_1^{\mu} + i W_2^{\mu} \right) & -g W_3^{\mu} + g' B^{\mu} \end{pmatrix} \begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{3.19}$$

where we have written out the Pauli matrices in the first term. The kinetic term, given by $(D_\mu\langle\phi\rangle)^\dagger (D^\mu\langle\phi\rangle)$, only has 4 non-vanishing terms, which are due to the VEV v. If we make the following definitions:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_1 \mp iW_2\right)_\mu, \quad Z_\mu^0 = \cos\theta_w W_{3\mu} - \sin\theta_w B_\mu, \quad A_\mu^0 = \sin\theta_w W_{3\mu} + \cos\theta_w B_\mu, \tag{3.20}$$

with $\tan\theta_w \equiv g'/g$ the *weak-mixing angle*, then this product can be written in a diagonal form:

$$\frac{1}{4}g^2v^2\, W^{+\mu}W_\mu^- + \frac{1}{8}\left(g^2 + g'^2\right)v^2\, Z^\mu Z_\mu. \tag{3.21}$$

The masses of the diagonal states can be read off right away:

$$m_W^2 = \frac{1}{4}g^2v^2, \qquad m_Z^2 = \frac{1}{4}\left(g^2 + g'^2\right)v^2, \qquad m_A^2 = 0. \tag{3.22}$$

Thus the Higgs Mechanism gives mass to three bosons, $W^\pm$ and Z, leaving only the physical photon, A, massless. We typically say that the three dofs from the Higgs field were "eaten" by the gauge bosons in order to become massive (since massless bosons only have two polarization dofs while massive ones have three).

Fermions also acquire mass from the non-zero VEV of the Higgs. In this case, the Yukawa sector (Eq. (3.15)) is rewritten as:

$$\mathcal{L}_{\text{Yukawa}} = Y_{ij}^u\overline{Q_{Li}}U_{Rj}\langle\widetilde{\phi}\rangle + Y_{ij}^d\overline{Q_{Li}}D_{Rj}\langle\phi\rangle + Y_{ij}^e\overline{L_{Li}}E_{Rj}\langle\phi\rangle + \text{h.c.} \tag{3.23}$$

Rotating to a basis where all of the Yukawa matrices are diagonal gives:[2]

$$\hat{Y}^u = \begin{pmatrix} y_u & 0 & 0 \\ 0 & y_c & 0 \\ 0 & 0 & y_t \end{pmatrix}, \qquad \hat{Y}^d = \begin{pmatrix} y_d & 0 & 0 \\ 0 & y_s & 0 \\ 0 & 0 & y_b \end{pmatrix}, \qquad \hat{Y}^e = \begin{pmatrix} y_e & 0 & 0 \\ 0 & y_\mu & 0 \\ 0 & 0 & y_\tau \end{pmatrix}. \tag{3.24}$$

The first (second) component of the $\langle\phi\rangle$ ($\langle\widetilde{\phi}\rangle$) field gives rise to mass terms for the fermions, while the second (first) component gives rise to interaction terms between the Higgs boson and fermions. The mass part of the Yukawa Lagrangian becomes:

$$\begin{aligned} \mathcal{L}^{\text{masses}}_{\text{Yukawa}} = & \frac{y_u v}{\sqrt{2}} \overline{u_L}\, u_R + \frac{y_c v}{\sqrt{2}} \overline{c_L}\, c_R + \frac{y_t v}{\sqrt{2}} \overline{t_L}\, t_R + \\ & \frac{y_d v}{\sqrt{2}} \overline{d_L}\, d_R + \frac{y_s v}{\sqrt{2}} \overline{s_L}\, s_R + \frac{y_b v}{\sqrt{2}} \overline{b_L}\, b_R + \\ & \frac{y_e v}{\sqrt{2}} \overline{e_L}\, e_R + \frac{y_\mu v}{\sqrt{2}} \overline{\mu_L}\, \mu_R + \frac{y_\tau v}{\sqrt{2}} \overline{\tau_L}\, \tau_R \,. \end{aligned} \tag{3.25}$$

Here I labeled each component of the $SU(2)_L$ doublets and singlets from Eq. (3.3) according to their eigenstate in the rotated basis; this gives rise to states with definite mass: u, d, c, s, t, and b quarks, as well as e, μ, and τ leptons. The masses of each state can be directly read off from Eq. (3.25):

$$\begin{aligned} m_u &= \frac{y_u v}{\sqrt{2}}, & m_c &= \frac{y_c v}{\sqrt{2}}, & m_t &= \frac{y_t v}{\sqrt{2}}, \\ m_d &= \frac{y_d v}{\sqrt{2}}, & m_s &= \frac{y_s v}{\sqrt{2}}, & m_b &= \frac{y_b v}{\sqrt{2}}, \\ m_e &= \frac{y_e v}{\sqrt{2}}, & m_\mu &= \frac{y_\mu v}{\sqrt{2}}, & m_\tau &= \frac{y_\tau v}{\sqrt{2}}, \end{aligned} \tag{3.26}$$

[2]This is always possible with two bi-unitary matrices, the product of which is the CKM matrix [15].

In this remarkable way, the Higgs mechanism, i.e. the non-zero VEV of the Higgs field, gives rise to gauge boson and fermion masses in the SM without explicitly breaking either gauge invariance or chiral symmetry. The masses are all proportional to the VEV and to the Yukawa coupling constants but are not predicted by the SM. They are left as parameters of the model to be determined experimentally.

The interaction part of the Yukawa Lagrangian becomes:

$$\begin{aligned}\mathcal{L}_{\text{Yukawa}}^{\text{interactions}} = \frac{-h}{v}\Big[& m_u\,\overline{u_L}\,u_R + m_c\,\overline{c_L}\,c_R + m_t\,\overline{t_L}\,t_R + \\ & m_d\,\overline{d_L}\,d_R + m_s\,\overline{s_L}\,s_R + m_b\,\overline{b_L}\,b_R + \\ & m_e\,\overline{e_L}\,e_R + m_\mu\,\overline{\mu_L}\,\mu_R + m_\tau\,\overline{\tau_L}\,\tau_R\Big]\,.\end{aligned} \tag{3.27}$$

The coupling between each fermion and the Higgs boson h is proportional to the fermion mass itself and inversely proportional to the VEV. This is a special feature of the Higgs mechanism which is amenable to experimental verification.

3.1.6 Summary of the SM

The SM is a phenomenological model consisting of several relativistic quantum fields and representing our best understanding of elementary particle physics to date. The SM is built on top of local (or gauge) symmetry requirements, which by their structure give rise to various gauge bosons. Several fermionic fields are also added to the model based on available experimental evidence. Finally, the inclusion of a Higgs scalar field completes the model and provides mass for fermions and some gauge bosons via the Higgs Mechanism—an instance of spontaneous symmetry breaking. The Higgs Mechanism is

necessary because explicit fermion and gauge boson mass terms are not compatible with the local symmetry structure of the SM. The field content of the SM is summarized in Table 3.1.

Table 3.1: Field content of the SM. The notation for the field representations under the SM gauge groups is $(A, B)_C$, where A is the representation under $SU(3)_C$, B the representation under $SU(2)_L$, and C the charge under $U(1)_Y$.

	Field notation	$SU(3)_C \times SU(2)_L \times U(1)_Y$ representation
Higgs	ϕ	$(1,2)_{+1/2}$
Gauge	G^i	$(8,1)_0$
	W^i	$(1,3)_0$
	B	$(1,1)_0$
Fermions	Q_{Li}	$(3,2)_{+1/6}$
	U_{Ri}	$(3,1)_{+2/3}$
	D_{Ri}	$(3,1)_{-1/3}$
	L_{Li}	$(1,2)_{-1/2}$
	E_{Ri}	$(1,1)_{-1}$

The SM has been extremely successful at predicting a large array of experimental results, often to unprecedented precision. At its core, it incorporates 19 free parameters which must be measured experimentally to completely define it:

- 6 quark masses: m_u, m_d, m_c, m_s, m_t, m_b;
- 3 charged lepton masses: m_e, m_μ, m_τ;
- 3 quark weak-eigenstate mixing angles: $\sin\theta_{12}$, $\sin\theta_{23}$, $\sin\theta_{13}$;
- 1 CP-violating phase from the CKM matrix: δ;
- 1 strong CP-violating angle: θ_{QCD};
- 3 gauge couplings: g, g', g_s;
- 2 Higgs parameters: v, λ.

The particle content of the SM after spontaneous symmetry breaking is depicted in Fig. 3.2. Fermions are shown on the left, with quarks in purple and leptons in green, while bosons are on the right, with gauge bosons in orange and the remaining Higgs boson in yellow.

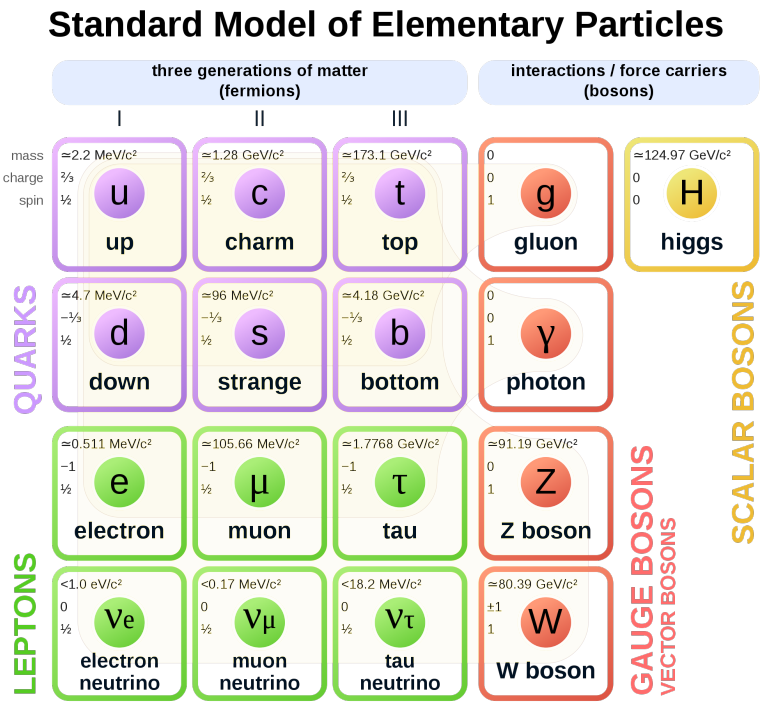


Figure 3.2: Particle content of the SM (public domain).

While the SM successfully predicts and explains a large class of phenomena, it still leaves gaps in our understanding of elementary particle physics and astrophysics. Several problems remain open which the SM as it stands cannot resolve. Here I list some of them with relevant references, though my focus will be on the dark matter problem:

- The hierarchy problem [20];
- Neutrino oscillations and neutrino masses [21];
- Matter-antimatter asymmetry [22];
- The fermion mass hierarchy problem [23];

- Dark energy [24];
- Dark matter.

Dark matter cannot currently be explained by the SM. There is no particle in the SM capable of constituting dark matter without violating several tight experimental constraints. For instance, neutrinos were at one point considered a potential candidate since they are so elusive, but due to their lightness they are also too relativistic to explain the observed clumpy dark matter halo distribution around galaxies [25].

Instead, most promising solutions to the dark matter problem involve a new particle or set of particles. In general, proposed particles must be heavy, slow, and interact weakly with SM particles to constitute a potential candidate for dark matter. In the following section, I discuss some of the interesting theoretical hypotheses to explain dark matter.

3.2 WIMPs and a complex dark sector

This section introduces historically favorable theories for dark matter, starting with weakly interacting massive particles (WIMPs) before moving on to more complex models captured in the concept of a rich dark sector.

The WIMP hypothesis proposes that dark matter consists of a single additional particle beyond the SM. WIMPs could constitute dark matter if they are heavy and interact little. The simplest possibility is that WIMPs only interact via gravity and via the weak force. Since we know the weak force to be a feeble SM interaction (as suggested by its name), the hypothesis suggests that a single particle that couples to the SM via this force alone could solve the dark matter problem.

The WIMP hypothesis has been popular in part because of its simplicity. Occam's razor says that, all other things considered, the simplest theory able to explain a given phenomenon is to be given preference over all others. In this sense, a single particle addition to the SM and zero new force requirements is arguably the simplest extension of the SM. If WIMPs are able to account for dark matter, then they are naturally a leading contender.

One way to extend the SM in Eq. (3.4) with WIMPs is to define a χSM Lagrangian with a new fermion field [15]:

$$\mathcal{L}_{\chi\mathrm{SM}} = \mathcal{L}_{\mathrm{SM}} + \mathcal{L}_{\chi}, \tag{3.28}$$

where the χ Lagrangian is:

$$\mathcal{L}_{\chi} = i\,\bar{\chi}\gamma^{\mu}\partial_{\mu}\chi - \frac{1}{2}\left(m_{\chi}\,\chi\chi + \mathrm{h.c.}\right). \tag{3.29}$$

The fermion content addition is accompanied by a new symmetry, Z_2, under which all SM particles are even. The new fermion is the only odd particle under Z_2 transformations, so its quantum numbers are $\chi(1,1)_{0^-}$, where we now added a '-' to indicate odd parity under the new symmetry. χ is a dark matter candidate because it is stable (being the lightest particle with odd-Z_2 parity) and carries no charge, electromagnetic or color.

The interactions between the new fermion and the SM can only be realized at a non-renormalizable level via higher-dimension operators:

$$\mathcal{L}^{\mathrm{int},d=5,6}_{\chi\mathrm{SM}} \supset \frac{Z_{\chi\phi}}{\Lambda}\chi\chi\phi^{\dagger}\phi + \frac{Z_{\chi f}}{\Lambda^2}\chi\chi\bar{f}f, \tag{3.30}$$

where Λ is some cut-off energy scale above which the effective theory is no longer applicable. The relevant question to ask in assessing this model as a dark matter explanation is whether there are values of $Z_{\chi\phi}/\Lambda$ or $Z_{\chi f}/\Lambda^2$ compatible with the observed dark matter thermal-relic abundance. The thermal-relic abundance is the amount of dark matter leftover from the early universe. It corresponds to the approximately 25% dark matter density observed today.

Fig. 3.3 plots the dark matter co-moving number density in the universe as a function of time or equivalently of temperature. There are three noteworthy regions in this plot. On the left, in the very early universe, the temperature is extremely high and the universe is energetic. Therefore dark matter is produced and annihilated at roughly the same rate, leading to equilibrium abundance. As the universe expands and cools, eventually there is no longer enough energy to produce dark matter particles, so its density begins to fall precipitously as dark matter particles find each other and annihilate. This would continue until all dark matter was gone, except that at some point the universe expands so much that it becomes unlikely for a dark matter particle to find another to annihilate. At that point, the density stabilizes again, which is shown on the right. This is called the freeze-out scenario for dark matter evolution and in this scenario, the observed abundance today corresponds to the freeze-out abundance.

The point at which freeze-out occurs depends on the dark matter scattering probability, which in turn is a function of the annihilation cross section and of the velocity distribution of dark matter particles. Several hypotheses are illustrated in Fig. 3.3 with different colors. The velocity-averaged dark matter annihilation cross section $\langle \sigma v \rangle$ that roughly corresponds to today's 25% dark matter abundance turns out to follow the simple relation [15]:

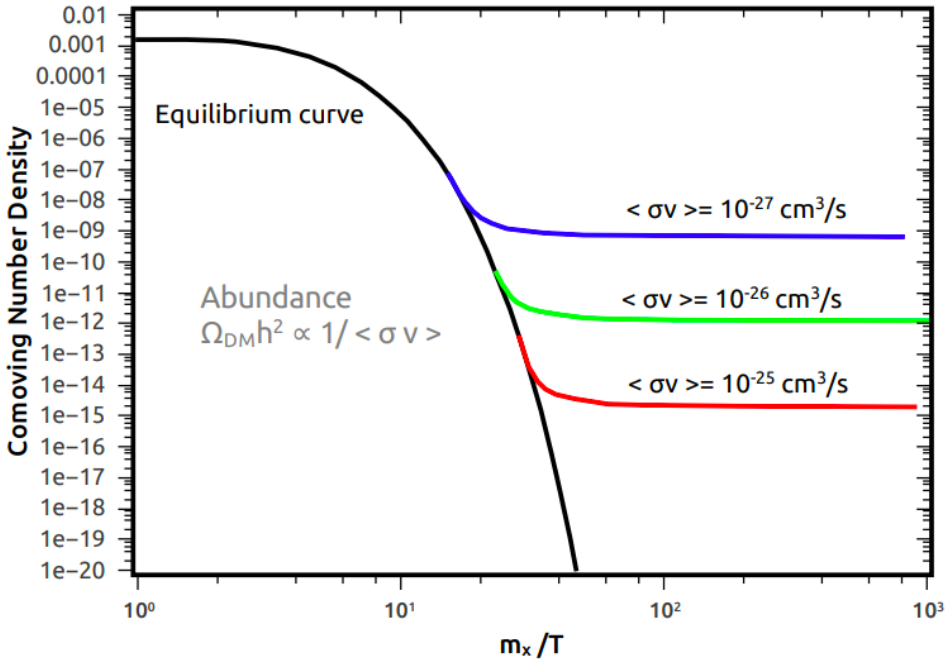


Figure 3.3: Evolution of dark matter abundance in the universe in the freeze-out scenario and resulting thermal-relic abundance [26].

$$\frac{\Omega_c h^2}{25\%} \approx \frac{\text{pb} \cdot c}{\langle \sigma v \rangle}, \tag{3.31}$$

where pb stands for a picobarn. Assuming the 2-to-2 dark matter annihilation term dominates in Eq. (3.30), the allowed dark matter mass range compatible with the thermal-relic abundance is bounded both from above and from below. If $m_{DM} \ll m_Z$, the interaction approaches the Fermi limit where the Z boson can be integrated out. This leads to a cross-section dependence of the form $\langle \sigma v \rangle \sim G_F^2 m_{DM}^2$ where G_F is the Fermi coupling. In this limit, the dark matter mass must be above a few GeV in order to close the thermal-relic abundance. This is also known as the Lee-Weinberg bound [27]. If instead $m_{DM} \gg m_Z$, the Z propagator in the 2-to-2 scattering becomes effectively "massless" and the cross section takes the form $\langle \sigma v \rangle \sim g^4/m_{DM}^2$, where g is the weak coupling. In this case, the dark matter mass is constrained to be below a few TeV in order to close the universe [28]. The fact that the range of allowed dark matter masses to explain the

observed dark matter abundance is compatible with weak-scale interactions did not go unnoticed and led to the suggestion that perhaps dark matter also interacts with the SM via the weak force (the WIMP hypothesis). This coincidence was dubbed the *WIMP miracle*.

Over the past four decades, such reasoning led to the development of an extensive experimental search program for WIMPs that is still very active. Experimental searches for WIMPs and other forms of dark matter are discussed in Chapter 4. While they remain a theoretically attractive dark matter solution, null results from increasingly sensitive experiments have steadily broadened the landscape of dark matter phenomenology towards more complex models that hypothesize additional states instead of a single WIMP.

Fig. 3.4 highlights the diversity of available complex theories of dark matter. A common feature of these theories is to postulate the existence of a *dark sector* which encompasses dark matter but also other states. Dark sector constituents are secluded from the SM because they are not charged under its symmetries (strong, weak, or electromagnetic).

In dark sector theories, the seclusion between hidden and SM particles necessitates some kind of mediator particle to connect the two sectors. Without such a bridge, it would be difficult to experimentally probe these models because dark matter would only interact gravitationally. These mediator particles provide what are known as "portal" interactions between sectors. There are several types of proposed portals [30]:

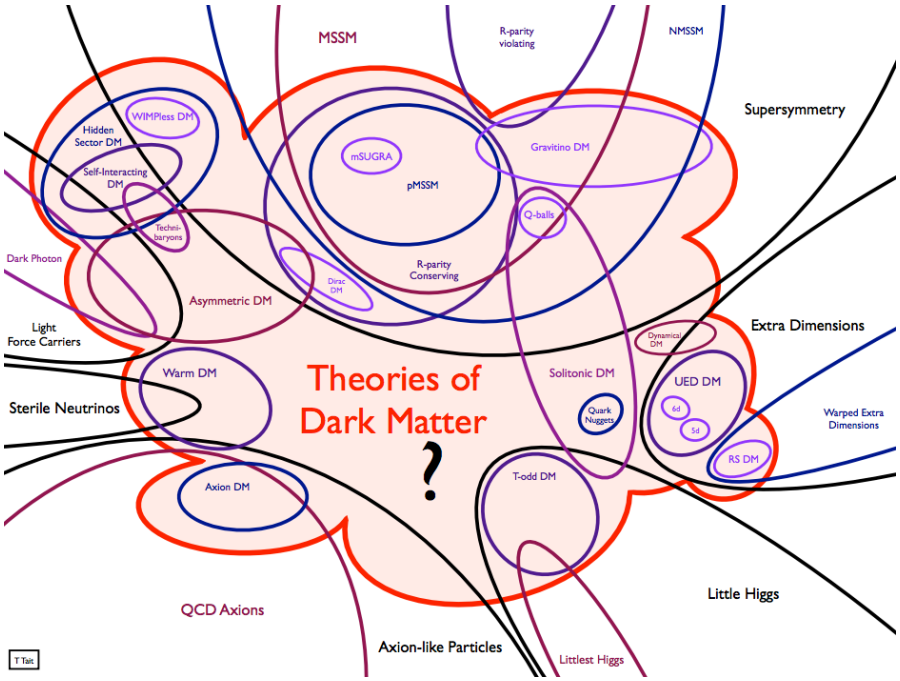


Figure 3.4: The many theories of dark matter [29].

$$\mathcal{L} \supset \begin{cases} -\frac{\epsilon}{2\cos\theta_w} B_{\mu\nu} F'^{\mu\nu}, & \text{vector portal} \\ (\mu\phi + \lambda\phi^2)\, H^\dagger H, & \text{Higgs portal} \\ y_n\, LHN, & \text{neutrino portal} \\ (a/f_a)\, F_{\mu\nu}\tilde{F}^{\mu\nu}, & \text{axion portal.} \end{cases} \tag{3.32}$$

In the vector portal, $B_{\mu\nu}$ is the SM hypercharge field strength tensor, and $F'_{\mu\nu}$ is the A'_μ strength tensor, a gauge boson of a new $U(1)_D$ symmetry in the dark sector; ϵ is the kinetic mixing connecting the field strength tensors (with θ_w the weak mixing angle). In the Higgs portal, H is the SM Higgs scalar field and ϕ is a new scalar field in the dark sector similar to the SM Higgs; μ and λ are coupling constants. In the neutrino portal, L is a SM $SU(2)$ lepton doublet and N is a new heavy neutrino living in the dark sector; y_n is the coupling constant between them. Finally, in the axion portal, $F_{\mu\nu}$ is the field strength tensor of the SM photon field and $\tilde{F}_{\mu\nu}$ is its dual; a is the pseudoscalar axion

field suppressed by a high mass scale f_a (and hence non-renormalizable).

All portals offer interesting scenarios of physics beyond the SM, and all but the axion portal are renormalizable extensions of the SM Lagrangian. Here I will focus on the vector portal because it offers a compelling case for explaining the dark matter thermal relic [30]. The vector portal connects the two sectors via a mixing between kinetic terms of two gauge boson fields: the SM hypercharge and the gauge boson of a new dark $U(1)_D$ symmetry, popularly known as the *dark photon* (since it behaves similarly to ordinary photons but in the dark sector) [31].

The kinetic mixing portal offers an interesting way to connect the SM and dark sectors. It also provides a concrete and exciting experimental program in the form of direct searches for dark photons and for physics involving dark photon interactions. For this reason, I dedicate a section to the physics of kinetic mixing, and the phenomenological implications of a vector portal.

3.3 Kinetic mixing

Kinetic mixing between the dark photon field and neutral gauge bosons of the SM is the main theoretical construct enabling the "bridge" between a complex dark sector and ordinary particles, as mentioned in Section 3.2. This section provides an overview of kinetic mixing and derives some phenomenological implications arising from the addition of dark photons to the SM.[3]

In the dark photon scenario, a new local $U(1)_D$ symmetry is added to the SM. The gauge boson of this symmetry is the dark photon, denoted by A'_μ. The gauge sector of

[3]Unless otherwise noted, the derivation presented here is adapted and expanded from Ref. [32].

the SM Lagrangian is extended with new A'_μ-related terms:

$$\begin{aligned}\mathcal{L}_{\text{gauge}} \supset & -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}A'_{\mu\nu}A'^{\mu\nu} \\ & + \frac{1}{2}\overline{m}_Z^2 Z_\mu Z^\mu + \frac{1}{2}\overline{m}_{A'}^2 A'_\mu A'^\mu \\ & + \frac{\epsilon}{2\cos\theta_w}A'_{\mu\nu}B^{\mu\nu},\end{aligned} \tag{3.33}$$

where $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$ and $A'_{\mu\nu} = \partial_\mu A'_\nu - \partial_\nu A'_\nu$ are the field strength tensors of the hypercharge and dark photon fields, respectively, θ_w is the weak-mixing angle and ϵ is the kinetic mixing coefficient. This coefficient effectively determines the strength of the coupling between dark photons and electromagnetism (after rotation of the weak gauge boson basis). Finally, $\overline{m}_Z$ and $\overline{m}_{A'}$ denote the bare masses before kinetic mixing to differentiate them from their final values after several basis rotations. The Z bare mass is acquired from the Higgs mechanism as discussed in Section 3.1.5, while the dark photon mass can arise from a similar SSB process in the dark sector or from some other mechanism. Unlike the photon, a dark photon is allowed to have mass and hence the dark photon force can have a finite range.

In order to arrive at the final form of the Lagrangian with diagonal kinetic and mass terms, a series of field redefinitions is required. First, we redefine the B_μ field to absorb the kinetic mixing term:

$$B_\mu \to B_\mu + \frac{\epsilon}{\cos\theta_w}A'_\mu, \tag{3.34}$$

which tranforms the hypercharge kinetic term as:

$$-\frac{1}{4}B_{\mu\nu}B^{\mu\nu} \rightarrow -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{\epsilon^2}{4\cos^2\theta_w}A'_{\mu\nu}A'^{\mu\nu} + \frac{\epsilon}{2\cos\theta_w}A'_{\mu\nu}B^{\mu\nu}. \tag{3.35}$$

The last term in Eq. (3.35) cancels the last term in Eq. (3.33), and the redefined Lagrangian becomes:

$$\begin{aligned}\mathcal{L}_{\text{gauge}} \supset & -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}\left(1 - \frac{\epsilon^2}{\cos^2\theta_w}\right)A'_{\mu\nu}A'^{\mu\nu} \\ & + \frac{1}{2}\overline{m}_Z^2 Z_\mu Z^\mu + \frac{1}{2}\overline{m}_{A'}^2 A'_\mu A'^\mu.\end{aligned} \tag{3.36}$$

The kinetic term of the A'_μ field is no longer canonical, but an additional field redefinition can restore this property:

$$A'_\mu \rightarrow \frac{1}{(1-\epsilon^2/\cos^2\theta_w)^{1/2}}A'_\mu, \tag{3.37}$$

turning the Lagrangian into:

$$\begin{aligned}\mathcal{L}_{\text{gauge}} \supset & -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}A'_{\mu\nu}A'^{\mu\nu} \\ & + \frac{1}{2}\overline{m}_Z^2 Z_\mu Z^\mu + \frac{1}{2}\overline{m}_{A'}^2\left(\frac{1}{1-\epsilon^2/\cos^2\theta_w}\right)A'_\mu A'^\mu.\end{aligned} \tag{3.38}$$

The kinetic terms for both fields are canonical again, at the expense of the dark photon mass term which now includes the redefinition constant. The two field redefinitions up to this point can be written as a single basis transformation on the B_μ and A'_μ fields (and making the definition $\epsilon' \equiv \epsilon/\cos\theta_w$):

$$\begin{pmatrix} B_\mu \\ A'_\mu \end{pmatrix} \to \begin{pmatrix} 1 & \frac{\epsilon'}{\sqrt{1-\epsilon'^2}} \\ 0 & \frac{1}{\sqrt{1-\epsilon'^2}} \end{pmatrix} \begin{pmatrix} B_\mu \\ A'_\mu \end{pmatrix}. \tag{3.39}$$

This transformation can be recast in terms of the usual SM post-SSB fields Z_μ and A_μ by employing the weak mixing rotation:

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_w & \sin\theta_w \\ -\sin\theta_w & \cos\theta_w \end{pmatrix} \begin{pmatrix} B_\mu \\ W^0_\mu \end{pmatrix}. \tag{3.40}$$

The basis transformation Eq. (3.39) acting on the post-SSB fields can then be written as a function of A'_μ:

$$\begin{pmatrix} A_\mu \\ A'_\mu \end{pmatrix} \to \begin{pmatrix} 1 & \frac{\epsilon}{\sqrt{1-\epsilon'^2}} \\ 0 & \frac{1}{\sqrt{1-\epsilon'^2}} \end{pmatrix} \begin{pmatrix} A_\mu \\ A'_\mu \end{pmatrix}, \qquad \begin{pmatrix} Z_\mu \\ A'_\mu \end{pmatrix} \to \begin{pmatrix} 1 & \frac{\epsilon\tan\theta_w}{\sqrt{1-\epsilon'^2}} \\ 0 & \frac{1}{\sqrt{1-\epsilon'^2}} \end{pmatrix} \begin{pmatrix} Z_\mu \\ A'_\mu \end{pmatrix}. \tag{3.41}$$

Next, the mass terms in the Lagrangian can be diagonalized by an additional suitable rotation. Since photons are protected from acquiring mass in the SM, we need only look at the mass matrix between Z_μ and A'_μ:

$$M^2_{ZA'} = \overline{m}^2_Z \begin{pmatrix} 1 & -\eta \\ -\eta & \eta^2 + \delta^2 \end{pmatrix}, \tag{3.42}$$

with the definitions:

$$\eta \equiv \frac{\epsilon\tan\theta_w}{\sqrt{1-\epsilon'^2}}, \qquad \delta \equiv \frac{\overline{m}_{A'}/\overline{m}_Z}{\sqrt{1-\epsilon'^2}}. \tag{3.43}$$

The mass matrix Eq. (3.42) can be diagonalized with a rotation by an angle α:

$$\begin{pmatrix} Z_\mu \\ A'_\mu \end{pmatrix} \to \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} Z_\mu \\ A'_\mu \end{pmatrix}, \tag{3.44}$$

where α is defined by:

$$\alpha \equiv \frac{1}{2\eta}\left(1-\eta^2-\delta^2-\text{sign}\left(1-\delta^2\right)\sqrt{4\eta^2+\left(1-\eta^2-\delta^2\right)^2}\right). \tag{3.45}$$

Note the choice of the rotation direction in Eq. (3.44), which is opposite the one defining the weak mixing rotation in Eq. (3.40). With this rotation, the diagonal masses of the Z_μ and A'_μ become:

$$m^2_{Z,A'} = \frac{\overline{m}^2_Z}{2}\left(1+\eta^2+\delta^2 \pm \text{sign}\left(1-\delta^2\right)\sqrt{\left(1+\eta^2+\delta^2\right)^2-4\delta^2}\right). \tag{3.46}$$

All field redefinitions and rotations so far can be summarized in a single matrix transformation between the fields A'_μ, Z_μ, and A_μ:

$$\begin{pmatrix} A'_\mu \\ Z_\mu \\ A_\mu \end{pmatrix} \to \underbrace{\begin{pmatrix} (\eta/\epsilon)\cos\alpha\cot\theta_w & (\eta/\epsilon)\sin\alpha\cot\theta_w & 0 \\ -\sin\alpha-\eta\cos\alpha & \cos\alpha-\eta\sin\alpha & 0 \\ \eta\cos\alpha\cot\theta_w & \eta\sin\alpha\cot\theta_w & 1 \end{pmatrix}}_{\equiv C} \begin{pmatrix} A'_\mu \\ Z_\mu \\ A_\mu \end{pmatrix}, \tag{3.47}$$

where the single matrix is labeled C. This transformation diagonalizes both the kinetic and mass terms of the original Lagrangian, Eq. (3.33).

Beyond shifts in the mass of the Z_μ, the addition of a dark photon symmetry also has consequences to interactions between the gauge bosons in the SM. The relevant

interaction Lagrangian is extracted from the fermion sector, Eq. (3.9), and extended with A'_μ interaction terms:

$$\mathcal{L}_{\text{int}} \supset A'_\mu \mathcal{J}_D^\mu + \sum_f \left[Z_\mu \, \bar{f} \, \gamma^\mu \left(g_v + g_a \gamma^5 \right) f + e \, Q_f \, A_\mu \, \bar{f} \, \gamma^\mu f \right], \tag{3.48}$$

where $\mathcal{J}_D^\mu$ is the dark $U(1)_D$ current, and the couplings to Z_μ and A_μ are summed over all fermion species. With the modifications introduced in Eq. (3.47), the interaction Lagrangian in the mass basis becomes:

$$\begin{aligned} \mathcal{L}_{\text{int}} \rightarrow & \left(C_{A'A'} \, A'_\mu + C_{A'Z} \, Z_\mu \right) \mathcal{J}_D^\mu \\ & + \sum_f \Bigg\{ A'_\mu \, \bar{f} \, \gamma^\mu \left[\left(g_v \, C_{ZA'} + e \, Q_f \, C_{AA'} \right) + g_a \, C_{ZA'} \, \gamma^5 \right] f \\ & + Z_\mu \, \bar{f} \, \gamma^\mu \left[\left(g_v \, C_{ZZ} + e \, Q_f \, C_{AZ} \right) + g_a \, C_{ZZ} \, \gamma^5 \right] f \\ & + \mathrm{A}_\mu \text{ terms} \Bigg\}, \end{aligned} \tag{3.49}$$

where C_{XX} denotes a particular element of the transformation matrix Eq. (3.47), and we have omitted couplings to A_μ since it is massless and therefore does not couple to the $U(1)_D$ current.

There are three noteworthy aspects of this Lagrangian. First, the Z boson couples to the dark sector current $\mathcal{J}_D$ in the first line of Eq. (3.49). This can have ramifications with respect to the total width of the Z. Second, the Z coupling to fermions is modified by the addition of A'_μ in the third line. This feature alone already permits us to place some experimental bounds on dark sector kinetic mixing arising from electroweak precision tests. And third, the dark photon couples to fermions in the SM with a coupling parameter that is a slight modification from the Z coupling, in the second line. This observation is

useful to understand how the original kinetic mixing between the hypercharge and dark photon fields is propagated through the SM Lagrangian. The end result is an effective current between SM fermions and A'_μ that can be probed experimentally.

The total rotation matrix C is a bit opaque. We can gain some more insight by taking two limits: when $\epsilon \ll 1$ and $m_{A'} \ll m_Z$, and when $\epsilon \ll 1$ and $m_{A'} \approx m_Z$. These limits are also useful since experimental searches for dark photons usually operate with one of these two assumptions.

Taking the limit $\epsilon \ll 1$ and $m_{A'} \ll m_Z$ first, matrix C simplifies to:

$$C(\epsilon \ll 1,\ m_{A'} \ll m_Z) \simeq \begin{pmatrix} 1 & -\epsilon \tan\theta_w & 0 \\ 0 & 1 & 0 \\ \epsilon & 0 & 1 \end{pmatrix}. \tag{3.50}$$

The corresponding simplification in the interaction Lagrangian is:

$$\begin{aligned} \mathcal{L}_{\text{int}}(\epsilon \ll 1,\ m_{A'} \ll m_Z) &\simeq \left(A'_\mu - \epsilon \tan\theta_w\, Z_\mu\right) \mathcal{J}_D^\mu + \sum_f \epsilon\, e\, Q_f\, A'_\mu \bar{f}\, \gamma^\mu\, f + \ldots \\ &= \left(A'_\mu - \epsilon' \sin\theta_w\, Z_\mu\right) \mathcal{J}_D^\mu + \sum_f (\epsilon' \cos\theta_w)\, e\, Q_f\, A'_\mu \bar{f}\, \gamma^\mu\, f + \ldots, \end{aligned} \tag{3.51}$$

where we have written the expression in the two equivalent forms typically seen in the literature. This limit is the more well-known version of the modified SM interaction Lagrangian due to the addition of a dark photon.

The second limit, $\epsilon \ll 1$ and $m_{A'} \approx m_Z$, yields:

$$C(\epsilon \ll 1,\ m_{A'} \approx m_Z) \simeq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ \epsilon & \epsilon & \sqrt{2} \end{pmatrix}. \tag{3.52}$$

The interaction Lagrangian would seem to be correspondingly modified to:

$$\mathcal{L}_{\text{int}}(\epsilon \ll 1,\ m_{A'} \approx m_Z) \simeq \frac{1}{\sqrt{2}} \left\{ \left(A'_\mu + Z_\mu \right) \mathcal{J}_D^\mu + \sum_f \left[A'_\mu \bar{f}\, \gamma^\mu \left(-g_v + \epsilon\, e\, Q_f \right) - g_a \gamma^5 \right] f \right\}. \tag{3.53}$$

The couplings between Z_μ and $\mathcal{J}_D^\mu$ and between A'_μ and fermions do not vanish when $\epsilon \to 0$, which may seem troubling since it implies an incomplete decoupling between the dark sector and the SM in the limit of no kinetic mixing. However, as pointed out in Ref. [32], this is an artifact of the choice of rotation angle α in Eq. (3.44). To see this, the authors present a useful alternative formulation of the interaction Lagrangian with a parameterization in terms of different currents associated to each gauge boson:

$$\mathcal{L}_{\text{int}} = A'_\mu\, \mathcal{J}_D^\mu + Z_\mu\, \mathcal{J}_Z^\mu + A_\mu\, \mathcal{J}_{\text{EM}}^\mu. \tag{3.54}$$

If the limit C matrix in Eq. (3.52) is applied to this Lagrangian instead of the one in Eq. (3.49), it becomes:

$$\mathcal{L}_{\text{int}} \simeq \frac{1}{\sqrt{2}} \left[A'_\mu \left(\mathcal{J}_D^\mu - \mathcal{J}_Z^\mu + \epsilon \mathcal{J}_{\text{EM}}^\mu \right) + Z_\mu \left(\mathcal{J}_D^\mu + \mathcal{J}_Z^\mu + \epsilon \mathcal{J}_{\text{EM}}^\mu \right) + \sqrt{2}\, A_\mu\, \mathcal{J}_{\text{EM}}^\mu \right]. \tag{3.55}$$

To understand the behavior of the interaction between the dark sector and the SM, we can look at the coupling between the $\mathcal{J}_D^\mu$ and $\mathcal{J}_{\text{EM}}^\mu$ currents and between $\mathcal{J}_D^\mu$ and $\mathcal{J}_Z^\mu$ currents. Starting with the latter:

$$\begin{aligned}\text{Z-D coupling} &= \frac{1}{2}\left[\mathcal{J}_D^\mu\, Z_\mu\, Z_\nu\, \mathcal{J}_Z^\nu - \mathcal{J}_D^\mu\, A'_\mu\, A'_\nu\, \mathcal{J}_Z^\nu\right] \\ &= \frac{1}{2}\left[\mathcal{J}_D^\mu \left(\frac{g_{\mu\nu}}{q^2 - m_Z^2}\right)\mathcal{J}_Z^\nu - \mathcal{J}_D^\mu \left(\frac{g_{\mu\nu}}{q^2 - {m_{A'}}^2}\right)\mathcal{J}_Z^\nu\right] \\ &= \frac{1}{2\,m_Z^2}\left(\frac{\mathcal{J}_D^\mu\, \mathcal{J}_{\mu Z}}{q^2/m_Z^2 - 1}\right) - \frac{1}{2\,{m_{A'}}^2}\left(\frac{\mathcal{J}_D^\mu\, \mathcal{J}_{\mu Z}}{q^2/{m_{A'}}^2 - 1}\right).\end{aligned} \tag{3.56}$$

At low energy transfers compared to the mass of the A' and the Z ($q^2 \ll {m_{A'}}^2, m_Z^2$), this reduces to a Fermi four-point coupling effective theory:

$$\text{Z-D coupling} \approx \frac{1}{2}\left(\frac{1}{m_Z^2} - \frac{1}{{m_{A'}}^2}\right)\mathcal{J}_D^\mu\, \mathcal{J}_{\mu Z}. \tag{3.57}$$

The effective Fermi theory can be represented schematically with the Feynman diagrams in Fig. 3.5, where the upper row illustrates the connection between the dark current $\mathcal{J}_D^\mu$ and the Z current $\mathcal{J}_Z^\mu$ via either an A' or a Z, according to Eq. (3.56). In the low-energy approximation, only a single effective blob is visible, as per Eq. (3.57).

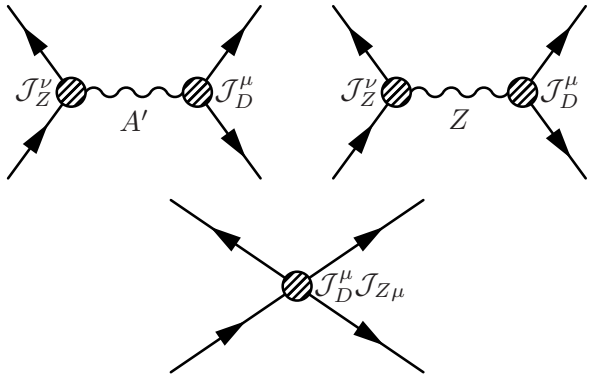


Figure 3.5: Feynman diagrams of the coupling between the dark sector current $\mathcal{J}_D^\mu$ and the Z current $\mathcal{J}_Z^\mu$. The processes are mediated by both the Z (upper right) and the A' (upper left), but in the low-energy Fermi theory the details of the coupling are hidden and only a single effective blob is assumed (center bottom).

We obtain similar results for the currents $\mathcal{J}_D^\mu$ and $\mathcal{J}_{\text{EM}}^\mu$:

$$\text{EM-D coupling} \approx \frac{\epsilon}{2}\left(\frac{1}{m_Z^2}+\frac{1}{{m_{A'}}^2}\right)\mathcal{J}_D^{\mu}\,\mathcal{J}_{\mu\text{EM}}. \tag{3.58}$$

The equivalent schematic Feynman diagrams for this coupling are shown in Fig. 3.6.

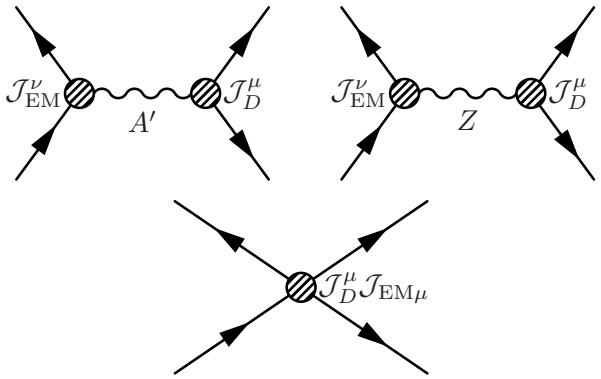


Figure 3.6: Feynman diagrams of the coupling between the dark sector current $\mathcal{J}_D^{\mu}$ and the electromagnetic current $\mathcal{J}_{\text{EM}}^{\mu}$. The processes are also mediated by both the Z (upper right) and the A' (upper left), but in the low-energy Fermi theory the details of the coupling are hidden and only a single effective blob is assumed (center bottom). We ignore photon couplings since it is massless.

Finally, the effective four-Fermi interaction Lagrangian in the approximation $m_{A'} \approx m_Z$ with this parameterization can be written as:

$$\begin{aligned} -\mathcal{L}_{\text{int}} &\approx \frac{1}{2}\left(\frac{1}{{m_{A'}}^2} - \frac{1}{m_Z^2}\right) \mathcal{J}_D^\mu \, \mathcal{J}_{\mu Z} \\ &+ \frac{\epsilon}{2}\left(\frac{1}{{m_{A'}}^2} + \frac{1}{m_Z^2}\right) \mathcal{J}_D^\mu \, \mathcal{J}_{\mu \text{EM}} \\ &\approx \frac{\epsilon}{2}\left(\frac{1}{{m_{A'}}^2} + \frac{1}{m_Z^2}\right) \mathcal{J}_D^\mu \, \mathcal{J}_{\mu \text{EM}}. \end{aligned} \tag{3.59}$$

In this form, it becomes clear that in the limit $\epsilon \to 0$ the interaction Lagrangian indeed vanishes and the dark sector decouples from the SM as expected, providing further evidence for the self-consistency of the theory. For ϵ small but not zero, it is also seen that there is no particular enhancement in the coupling between $\mathcal{J}_D^\mu$ and $\mathcal{J}_{\text{EM}}^\mu$. The coupling converges instead to ϵ/m_Z^2, and therefore no diverging behavior is expected in this limit.[4] This fact is useful when experimentally probing regions of dark photon parameter space close to the Z peak at the LHC, which will be discussed in Part II.

[4]Note, however, that there is still an enhancement to processes involving dark photons in this limit due to contributions from the Z boson itself.

CHAPTER 4

EXPERIMENTAL SEARCHES FOR DARK MATTER

This chapter presents a brief survey of past and current experimental strategies applied in the search for dark matter. These fall broadly into three categories: indirect detection of dark matter annihilation or decay; direct detection of scattering between dark matter and ordinary matter; and production of dark matter in accelerators. The categories differ in the type of interaction between dark matter and ordinary matter that is probed. Fig. 4.1 summarizes the complementarity between the three approaches.

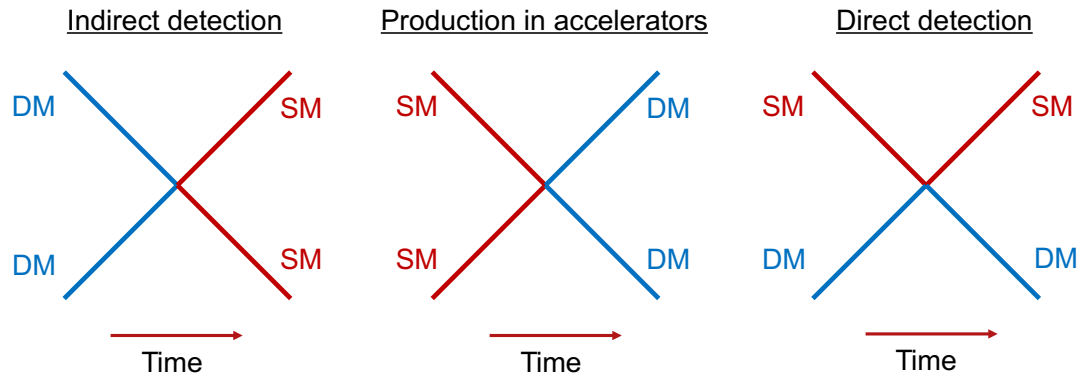


Figure 4.1: Dark matter detection channels. Indirect detection probes self-annihilation of dark matter in space, accelerators attempt to produce dark matter in the lab, and direct detection seeks evidence of scattering between dark matter and nucleons on the Earth.

In general terms, when looking for dark matter interactions, one can try to detect evidence of dark matter self-annihilation resulting in the production of ordinary SM particles, which is known as indirect detection. This would happen most often in space since the dark matter density on Earth is predicted to be small. Most experiments searching for dark matter annihilation are therefore telescope-based and reliant on accurate observations of the night sky.

One can also look for the scattering between dark matter and ordinary matter. Unlike dark matter self-annihilation, this has a non-negligible chance of occurring on Earth

when dark matter particles cross the planet. The usual assumption in this case is that dark matter interacts with the SM via the weak force (the WIMP hypothesis discussed in Chapter 3). The direct detection of this scattering therefore employs large and dense materials in extremely sensitive underground experiments.

The final approach is to identify production of dark matter from the collision of SM particles in a laboratory setting, typically done with a particle collider or fixed-target experiment. This has the advantage of providing a controlled environment in which dark matter could be identified and characterized. The challenge in this approach lies in the effort to ensure the environment is sufficiently under control and that all potential backgrounds are minimized and understood to the extent possible.

In the next section, I briefly discuss each of these three approaches separately with a special focus on the production of dark matter in accelerators, which is the main theme of this dissertation.

4.1 Indirect detection

The main goal of indirect dark matter detection is to identify in space the residues of dark matter self-interaction processes. These processes can typically happen in two ways: self-annihilation of dark matter particles, or decay of a (semi-stable) dark matter state [33]. In the former case, dark matter (DM) particles annihilate and produce SM particles via the weak force if they are WIMPs or via a mediator particle in scenarios with complex dark sectors. In the latter case, the lifetime of WIMP DM must be sufficiently large to account for the observed dark matter abundance in today's universe or else an excited state in the dark sector must be the source of the decay.

The reactions that deplete DM and produce SM particles can involve long chains of loop-induced decays or simple tree-level processes. Either way, the final products should be stable SM particles such as photons, neutrinos, electrons and positrons, or neutrons and protons. By looking for excess amounts of these particles in the sky, it is possible to search for particular DM depletion signatures.

Several studies probe such DM signatures using light in various ranges of the electromagnetic spectrum and a few have reported unexpected excesses in the past decade. For example, analysis of the stacked spectral emission of several galaxy clusters ([34]) and a separate analysis ([35]) of the emission of the Andromeda galaxy and Perseus galaxy cluster—the brightest object in the sky in the X-ray band—with the XMM-Newton X-ray space observatory have uncovered an unexpected emission line at 3.5 keV. The intensity appears to increase closer to the center of the objects. This excess emission is consistent with a sterile neutrino of mass around 7 keV (decaying to a photon and an active neutrino) and also with warm dark matter of mass 3.5 keV [34, 35]. However, current observations are still too faint to completely rule out other explanations, and more studies are needed to further probe this emission line.

An additional observed electromagnetic radiation excess comes from the center of our own Milky Way galaxy. A surplus of gamma-ray radiation in the GeV band is not well understood and could be consistent with dark matter annihilation [36, 37]. This galactic center GeV excess (GCE) was observed in data collected with the Fermi Large Area Telescope, a satellite-based observatory. The excess peaks between 1 and 3 GeV and is consistent with light thermal-relic dark matter annihilating to quarks [33]. However, the evidence is still not conclusive as it is also plausible that the excess is due to processes such as inverse Compton emission from high-energy electrons in earlier galactic history [38].

Indirect detection probes also study cosmic rays (energetic charged particles from the Sun and from outside the solar system) incident on Earth and search for hints of dark matter. One such detector is Payload for Antimatter Matter Exploration and Light-nuclei Astrophysics (PAMELA), mounted on the European Resurs-DK1 satellite with the goal of measuring the abundance of anti-matter (positrons and anti-protons) in cosmic rays. It first unveiled hints of an excess in the energy spectrum of positrons, particularly in the 10 GeV to 100 GeV range [39]. This is consistent with some models of dark matter annihilation [40].

The Alpha Magnetic Spectrometer (AMS-02) is another cosmic ray antimatter detector, launched in 2011 aboard the International Space Station (ISS). AMS-02 confirmed the excess high-energy positrons first seen by PAMELA while increasing the sensitivity of the observation and the range of positron energies probed [41]. Fig. 4.2 shows the positron fraction in cosmic rays as a function of energy observed by both AMS-02 and PAMELA.

However, a recent study found evidence of gamma-ray emission in large areas of the galactic plane due to the Geminga and Monogem [42]. These high-energetic gamma-rays (energies above 5 TeV) could be produced by energetic electrons and positrons via inverse Compton scattering with photon fields near the pulsars. The study found that the same positrons inducing this scattering could also account for almost 20% of the excess seen by PAMELA and AMS-02. It is therefore still unclear whether high-energy positrons from cosmic rays offer conclusive and definite evidence of dark matter annihilation in space.

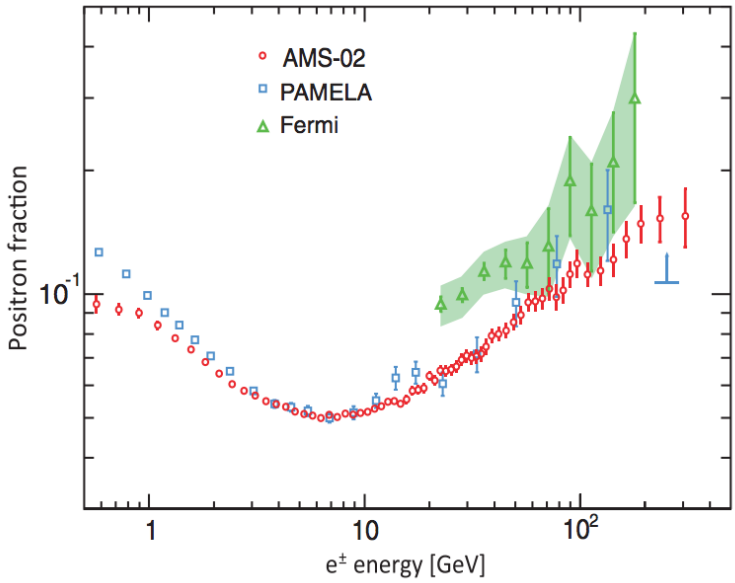


Figure 4.2: Positron flux energy spectrum from cosmic sources measured by AMS-02 and PAMELA [41]. Significant increases in flux are observed for energies higher than about 10 GeV. Most theoretical models of secondary positron production from the propagation of cosmic rays predict a *lower* flux at higher energies.

4.2 Direct detection

Direct detection seeks to directly probe the scattering between ordinary matter and dark matter from the Milky Way halo when it traverses the Earth. Unlike indirect detection, it does not require dark matter particles to meet and annihilate, which happens infrequently in space due to the low dark matter density. The major requirement for direct detection to be useful is that dark matter has a direct coupling to the SM. This is true in the case of WIMP models, discussed in Section 3.2, where this coupling is due to the weak force. Therefore, when a dark matter particle crosses a detector on its journey through the Earth, there is opportunity for scattering with nucleons or accompanying orbital electrons in the detector. The nuclear or electronic recoil from the scattering can be measured by exquisitely sensitive and sophisticated devices, revealing the interaction.

Direct detection experiments typically sit underground, covered by thick layers of

rock. This helps block cosmic rays from reaching the detector and causing background noise. Since the recoil signal generated by the dark matter interaction is faint, it is imperative that the background contamination be minimized to the extent possible. In addition to rock, experiments are surrounded by passive shields of heavy substances like lead and water or active shields with tagging capability to reduce contamination from the underground cavern that houses the setup.

Fig. 4.3 shows one example of a direct detection experiment: the LUX-ZEPLIN (LZ) detector, under construction at the Sanford Underground Research Facility in Lead, South Dakota. The LZ detector uses liquid xenon as the active material that receives a recoil from dark matter scattering. The recoil excites xenon atoms into higher states, which then combine pairwise into excited molecules. This is followed by de-excitation and dissociation of the xenon molecule with the emission of 175 nm light—in the UV spectrum. The light is then collected by hundreds of photomultiplier tubes (PMTs) placed above and below the xenon container. Sometimes xenon atoms can be ionized as well, releasing electrons in the medium. By applying a vertical electric field, the charge released drifts upward, producing additional, delayed electroluminescence light at the top of the detector. These two light signals are denoted S1 and S2 and together enhance the sensitivity of the LZ detector to dark matter nuclear recoils.

Cryogenic detectors are another important class of direct detection techniques. Fig. 4.4 shows one such detector under construction: the SuperCDMS at SNOLAB, in Sudbury, Canada. Here, the active material is a solid crystal lattice made of germanium or silicon atoms. When the dark matter particle strikes one of the atoms, the initial recoil generates a wave of atomic recoils that spreads isotropically throughout the lattice. This phenomenon is called a lattice phonon excitation. To ensure sensitivity to faint dark matter-induced phonons, the entire apparatus has to be cooled cryogenically.

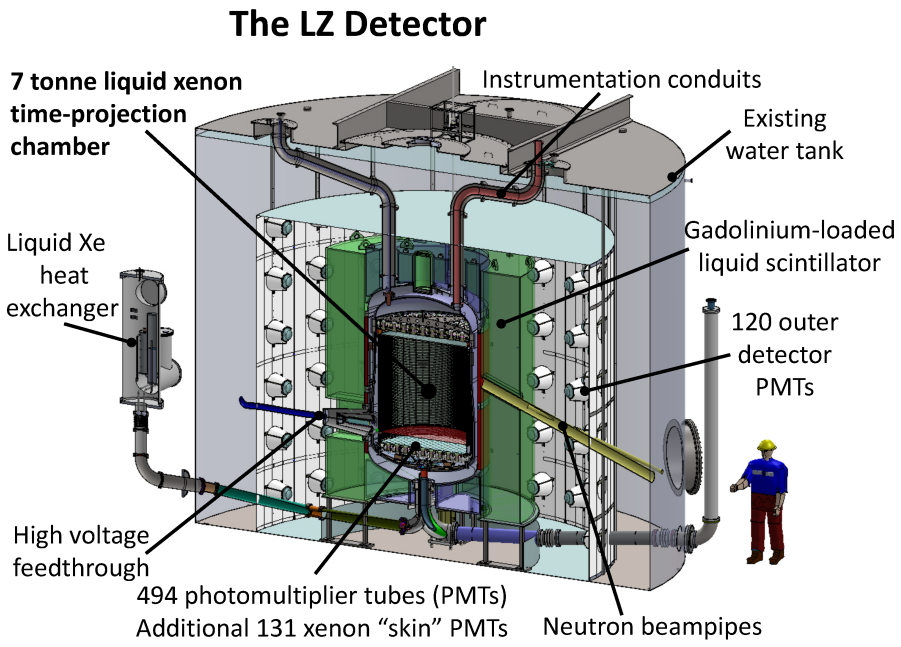


Figure 4.3: The LZ direct detection experiment under construction at the Sanford Underground Research Facility in Lead, South Dakota [43]. The detector uses liquid xenon as active medium.

Ultra-sensitive transition-edge sensors are placed on the surfaces of the lattice to detect arriving phonons. These sensors are also cooled to cryogenic temperatures right at the critical temperature in which they become superconducting (the transition edge). As soon as a heat phonon arrives, the temperature increases slightly, but this has a dramatic effect on the resistance of the sensor. This abrupt change in conductivity effectively amplifies the response to a single phonon manifold and is the key to this type of direct detection technique.

Direct detection experiments have made remarkable progress in advancing the sensitivity to WIMP dark matter in recent decades. Figs. 4.5 and 4.6 highlight this progress in terms of summary plots of several direct detection exclusion results: spin-independent and spin-dependent nucleon–dark matter cross section exclusions are plotted as a function of dark matter mass, respectively. Spin-independent and spin-dependent results are

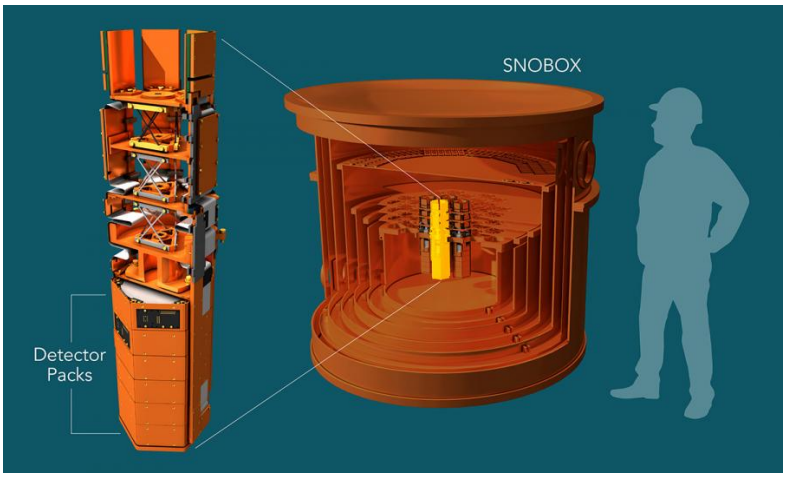


Figure 4.4: Detector setup of the SuperCDMS direct detection experiment at SNOLAB, in Sudbury, Canada [44]. The experiment uses cryogenic bolometer technology for dark matter detection.

typically quoted separately because they probe different types of couplings between nucleons and dark matter. The former interaction arises from scalar or vector couplings, while the latter from axial-vector ones. These couplings lead to varying sensitivity among direct detection experiments. In simplistic terms, spin-dependent interactions vary with the total nuclear spin of the target nucleus, whereas spin-independent interactions are more sensitive to the size of the nucleus compared to the de Broglie wavelength of the DM particle [45].

Although the vast majority of the direct detection community reports negative results in the search for dark matter, a notable exception actually claims discovery: the DAMA/NaI and subsequent DAMA/LIBRA experiments. They have reported a strong seasonal modulation in the observed signal, compatible with the expected modulation if Earth passes through a galactic DM halo in different directions throughout the year [47, 2]. The preferred dark matter hypothesis is shown in Fig. 4.5 with small filled contours.

The DAMA/LIBRA experiment sits in the Laboratori Nazionali del Gran Sasso in Italy and consists of a 5×5 array of sodium iodide crystals enriched with thallium

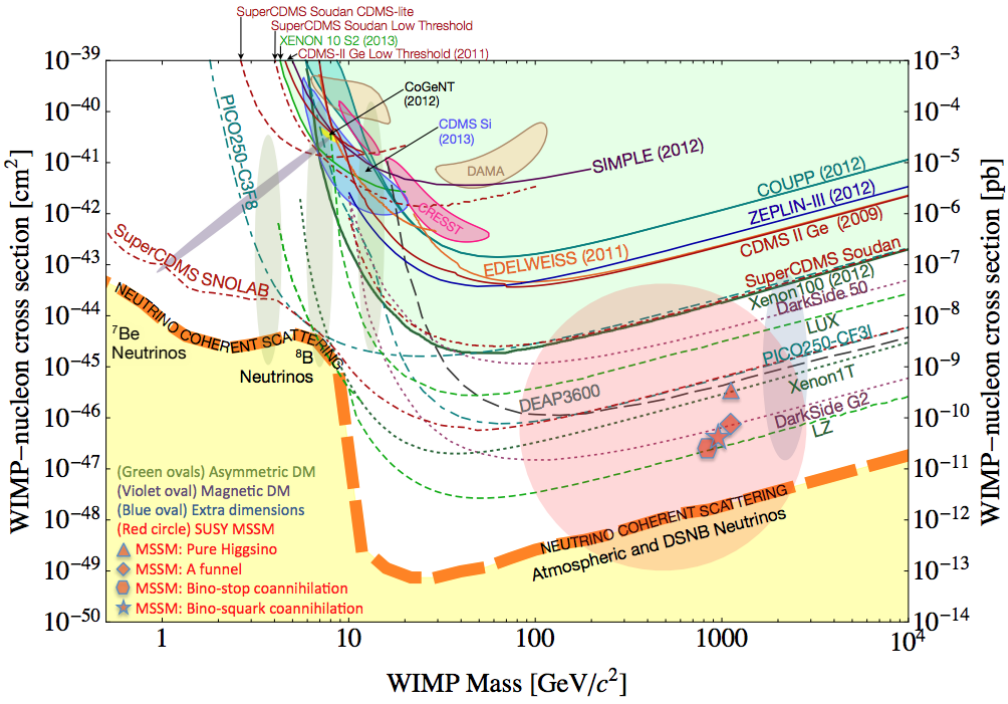


Figure 4.5: Observed (solid lines) and projected (dashed and dotted lines) limits on spin-independent WIMP-nucleon cross section as a function of WIMP mass from several direct detection experiments [46]. Claims of signal compatible with dark matter are shown in small filled contours.

(NaI(Ti)) and coupled to two PMTs on each crystal end [48]. The total NaI active material is roughly 250 kg. The array is surrounded by a sealed copper box and flushed with high-purity nitrogen to remove other gases and humidity. The copper box in turn is enclosed by a low-background heavy shield that is further surrounded by concrete made from local rock [48]. The expected background is extremely low according to the Collaboration.

Fig. 4.7 shows the modulated signal observed by DAMA/LIBRA in the course of roughly 20 years of data-taking. The modulation is undeniable and consistent with seasonal variation in phase with the Earth's orbit around the Sun; the remaining question is whether it is due to a real dark matter signal or to some unexplained systematic effect in the experimental setup. None of the other sensitive direct detection experiments have

identified a signal in the region of phase space where DAMA says dark matter should be, as suggested by the several exclusion plots in Fig. 4.5.

Many new experiments have been developed with the specific purpose of confirming or refuting DAMA's result. Some have released preliminary results, such as COSINE-100 [49] and ANAIS-112 [50], though they are not yet sensitive to the magnitude of the signal seen by DAMA since it has been running for many years. Others, such as SABRE [51], have planned versions in both northern and southern hemispheres, which could provide an interesting verification of the signal because the phases due to dark matter should be the same in both experiments, whereas a systematic effect due to seasonal variations would show inverted phases. And still others are being built with advanced cryogenic technology to enable the measurement of both a scintillation channel and a thermal phonon channel to gain more sensitivity to the faint dark matter signal (i.e. COSINUS [52]).

Apart from additional experiments to probe the DAMA signal, several studies have cast doubt on the claim of dark matter discovery by proposing that the signal is due instead to a variety of systematic effects unaccounted for by the experiment. Some interesting recent examples of alternative explanations include a distortion induced by the data analysis procedure, which removes the average background year-over-year [53], and a claim that the modulation may be due to contamination of the PMTs with surrounding Helium, which could cause a coincidence signal in the two PMTs with about the same rate seen by the DAMA/LIBRA experiment [54]. Thus the complete picture of DAMA's claims of dark matter discovery remains far from elucidated, and no smoking-gun evidence of dark matter has yet been achieved with direct detection techniques.

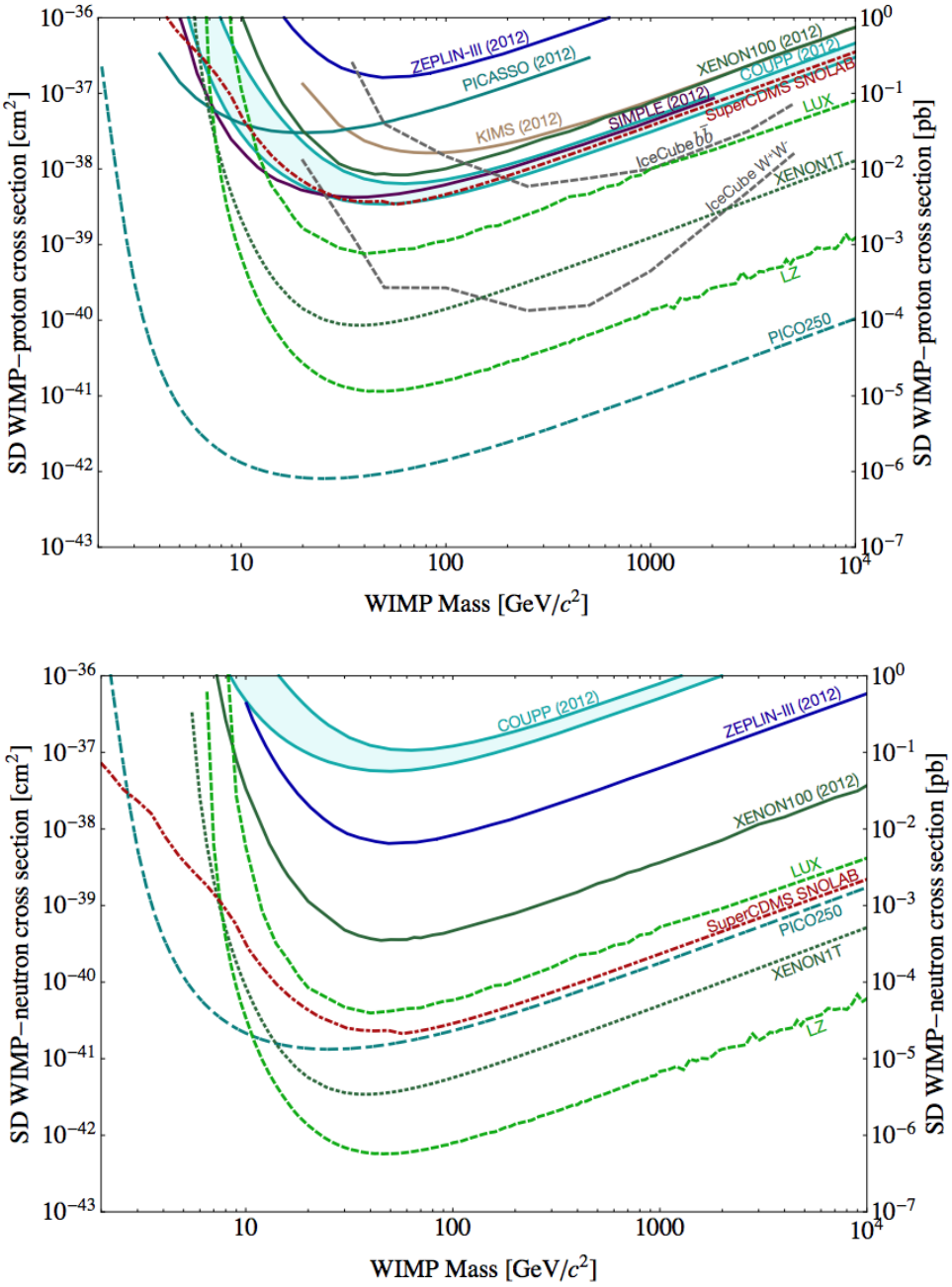


Figure 4.6: Observed (solid lines) and projected (dashed and dotted lines) limits on spin-dependent WIMP-proton (top) and WIMP-nucleon (bottom) cross sections as a function of WIMP mass from several direct detection experiments [46]. Separate plots are required because the spin-dependent axial-vector current probed here couples differently to each nucleon, with varying sensitivities across experiments.

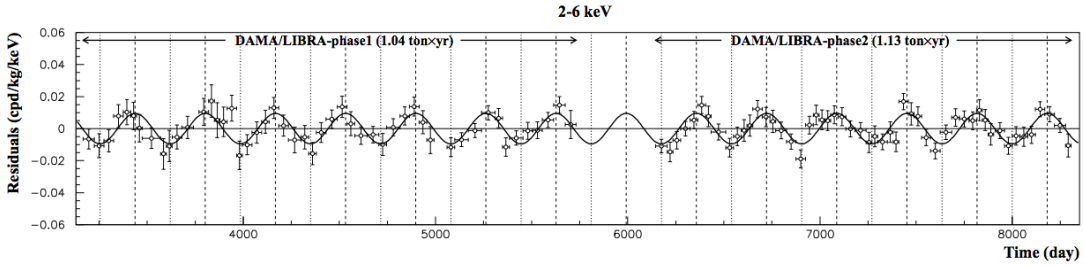


Figure 4.7: Seasonal modulation observed in the DAMA/LIBRA data [2]. The plot cover about 20 years of data-taking, displaying strong evidence of seasonal variation.

4.3 Production in accelerators

In addition to detecting signs of existing dark matter in the galactic halo, we can try to produce dark matter in the lab under a controlled environment. One benefit of this method is that it allows probing different types of couplings between dark matter and SM particles. Direct and indirect detection, in contrast, assume particular forms of this interaction (direct detection, for example, assumes a weak force coupling). The variety of potential couplings manifests itself both in the different available dark matter production modes as well as the varying subsequent dark matter decay channels that detectors can try to identify.

Another advantage of producing dark matter in the lab is that, since the properties of the accelerated particle beam are under control, backgrounds are much better understood. This gives accelerators high sensitivity even to theories of dark matter with small couplings to the SM by virtue of a good modeling of the background processes that could mimic the expected signature.

The central tenet of this approach is to smash SM particles against one another (such as protons-protons, protons-antiprotons, or electrons-positrons), and place detectors all around the collision point in the hopes of capturing and measuring the outgoing collision products. The actual collision can happen in two primary ways: beam versus beam (i.e. colliders), or beam versus static materials (i.e. fixed targets or beam dumps). These two approaches provide complementarity across several metrics, which I explore in Sections 4.3.1 and 4.3.2. The remainder of this dissertation is dedicated to an example of each such approach: a search for dark matter with the CMS detector at the LHC (a collider experiment), and a search for dark photons with the PADME experiment at Laboratori Nazionali di Frascati (a fixed-target experiment). Each experiment is discussed in detail

in Part II and Part III; here I give a general introduction to the techniques and a sense of the complementarity and interplay between them.

4.3.1 Fixed-target geometry

In the fixed-target geometry, a single beam is accelerated and steered towards a stationary thin target. Detectors are placed downstream of the target with the goal of measuring the products of the beam-target collision. A beam dump is also present to capture the non-interacting remnants of the beam. Fig. 4.8 sketches the typical setup of a fixed-target experiment.

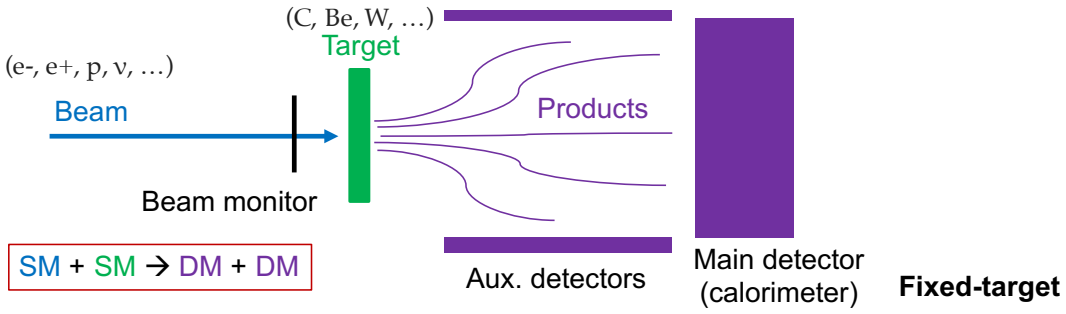


Figure 4.8: Layout of a generic fixed-target experiment. The beam strikes a target and the collision products are measured by detectors placed further downstream.

In the context of dark matter searches, fixed-target experiments can probe different scenarios of dark matter production. A particularly interesting possibility is the investigation of dark photon models with positron and electron beams. According to these models (discussed in Section 3.2), dark photons could be produced in positron-electron annihilation or in electron bremsstrahlung via a kinetic mixing between the dark and ordinary photon fields. Fixed-target experiments are especially suited to investigate the light dark matter parameter space because of the low energy requirements and the relative lack of backgrounds processes compared to a higher-energy collider environments.

Another advantage of fixed-target experiments is that they tend to have compact form factors and relatively low price tags. The scale is small enough that modest-sized collaborations (e.g. dozens of people) are able to build and operate the machine in a fairly short span of time of (e.g. a few years). In contrast, collider experiments typically cost billions of dollars to build and operate and can take a decade or longer from initial design to data-taking.

Fig. 4.9 summarizes recent results from fixed-target (and beam dump) experiments probing the light dark matter space. The excluded limits are given as a function of dark photon mass and kinetic mixing between dark photon and ordinary photon.

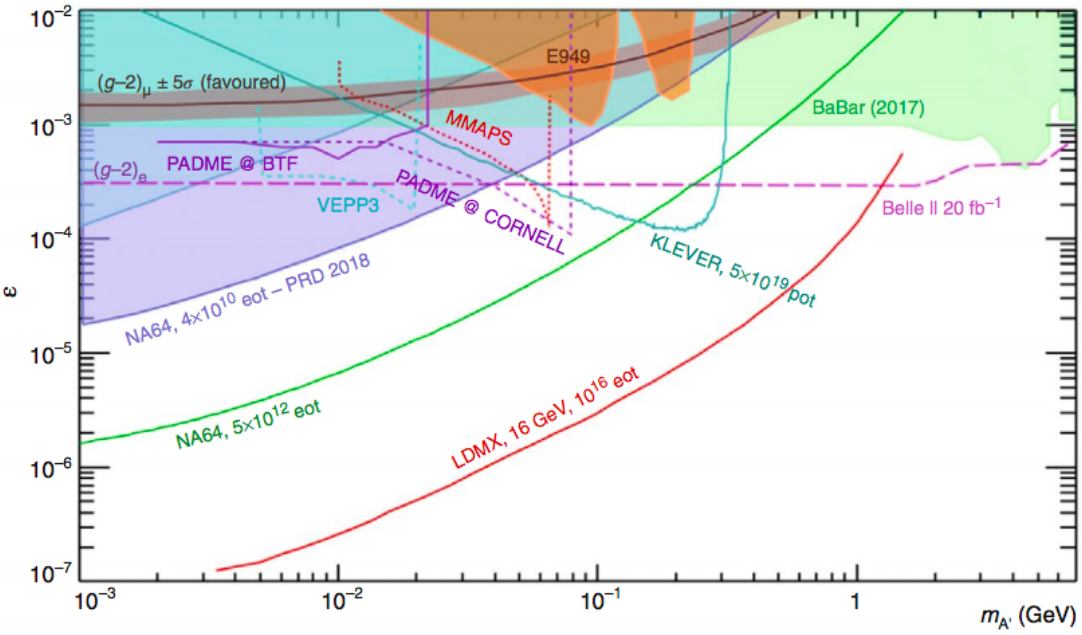


Figure 4.9: Dark photon projected and observed exclusion limits from several accelerator-based experiments [55]. The horizontal axis is the dark photon mass and the vertical axis the kinetic mixing coefficient.

4.3.2 Collider geometry

In the collider geometry, two beams are accelerated and then made to collide at a central location. Detectors are placed all around the interaction point, typically in a cylindrical arrangement and with an almost complete azimuthal coverage. This hermeticity is one of the main advantages of the collider geometry and is difficult to obtain with fixed-target experiments. Fig. 4.10 depicts a sketch of the central features of such experiments.

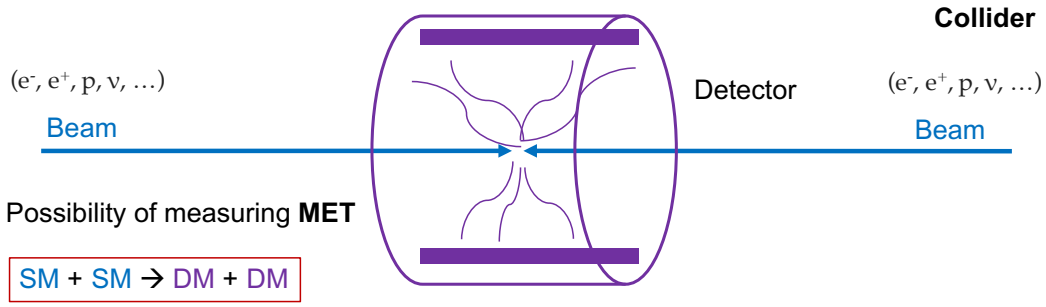


Figure 4.10: Layout of a generic collider experiment. Two beams traveling in opposite directions collide at a central point, and the collision products are measured by a hermetic detector with near full azimuthal coverage.

Having two beams further enables collider experiments to reach the highest energies available. If each beam particle has an energy of, say, 7 TeV in the lab frame, then in the center-of-mass frame, the total energy imparted in the collision is 14 TeV, or twice the single-particle energy. This is naturally unlike fixed-target experiments where only one beam is present and the target has negligible momentum compared to the beam momentum.

The full azimuthal coverage provides an extra benefit in dark matter searches: the ability to measure missing transverse momentum, or $p_{\mathrm{T}}^{\mathrm{miss}}$. $p_{\mathrm{T}}^{\mathrm{miss}}$ is constructed by adding the vector transverse momentum of all visible particle products in the collision and subtracting from zero. Since the two beams are oriented in the longitudinal direction and have no intrinsic transverse momentum, by conservation of momentum, the same should

be true of the sum of all outgoing particles. If there are invisible particles produced that manage to escape the detector, the deviation from zero of the sum of transverse momenta (i.e. the $p_{\mathrm{T}}^{\mathrm{miss}}$) can provide evidence of that. Since dark matter particles have feeble interactions with SM particles, most of the time they will escape the detector, but in collider experiments with full azimuthal coverage their presence can still be identified via $p_{\mathrm{T}}^{\mathrm{miss}}$. Most searches for dark matter in collider experiments rely on the $p_{\mathrm{T}}^{\mathrm{miss}}$ handle.

Figs. 4.11 and 4.12 summarize recent results from traditional searches for dark matter at CMS and ATLAS (respectively), two of the four large experiments at the LHC. These searches either look for dark matter produced in association with one visible particle (such as a W or Z boson) and subsequently escaping the detector, or recoiled against two particles in the form of a dijet signature. These plots assume a vector mediator and specific coupling values between the dark matter and the SM to provide a reference model for which the sensitivity can be compared between several searches.

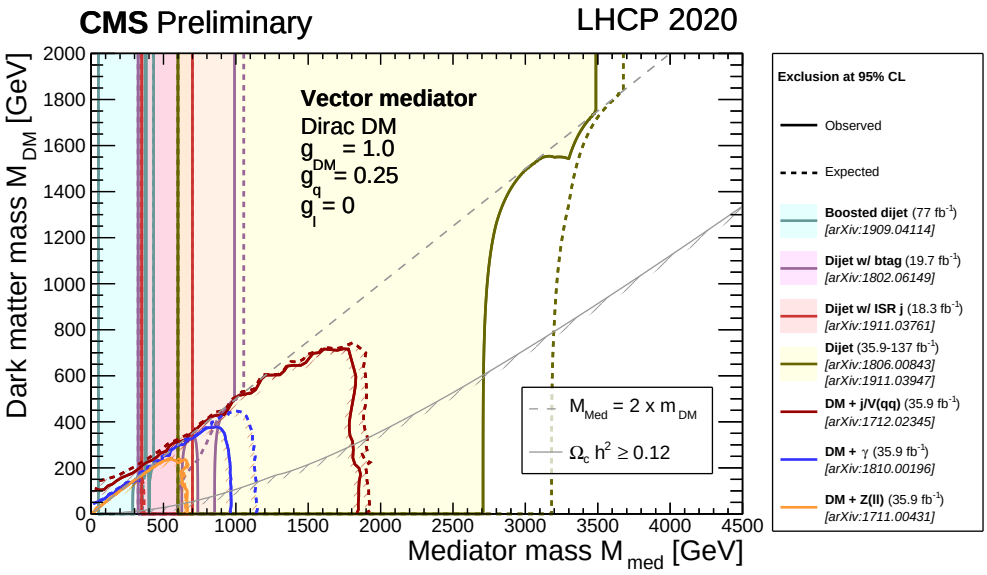


Figure 4.11: CMS exclusion limits with 95% confidence level of vector mediated models of dark matter [56]. The horizontal axis is the mediator mass and the vertical axis the Dirac fermion dark matter mass.

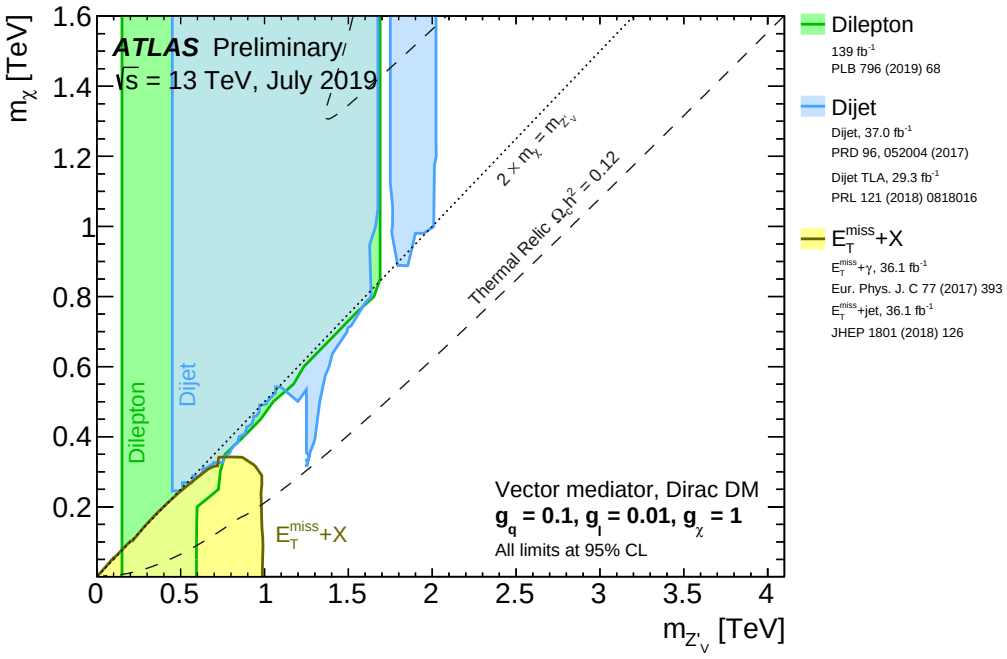


Figure 4.12: ATLAS exclusion limits with 95% confidence level of leptophilic vector mediated models of dark matter [57]. The horizontal axis is the mediator mass and the vertical axis the Dirac fermion dark matter mass.

4.3.3 Long-lived searches for dark matter

A particular subgroup of exotic searches for new physics at colliders that investigates long-lived particles (LLPs) has surged in popularity recently. Searches for LLPs probe alternative and unexplored phase space, and exploit the available experimental apparatus in innovative ways.

Several analyses with displaced signatures have been carried out in Compact Muon Solenoid (CMS): searches for disappearing tracks [58], emerging jets [59], semi-visible jets [60], displaced muons [61], displaced jets [62], and delayed photons [63], among others. These analyses typically set exclusion limits that are parameterized by the lifetime of the LLP in the benchmark model. Fig. 4.13 shows a CMS summary of recent LLP searches, highlighting the comprehensive lifetime coverage.

Several LLP searches use benchmark models based on supersymmetry (SUSY) [64]. SUSY provides a variety of mechanisms that result in long-lived signatures [65], so it has been a natural reference in these types of analysis. This dissertation focuses on a complementary benchmark model that predicts displaced final states but has a direct dark matter motivation instead. It also relies on the dark photon hypothesis to connect SM particles to new physics rather than relying on SUSY. The inelastic dark matter model is extensively discussed in Chapter 5.

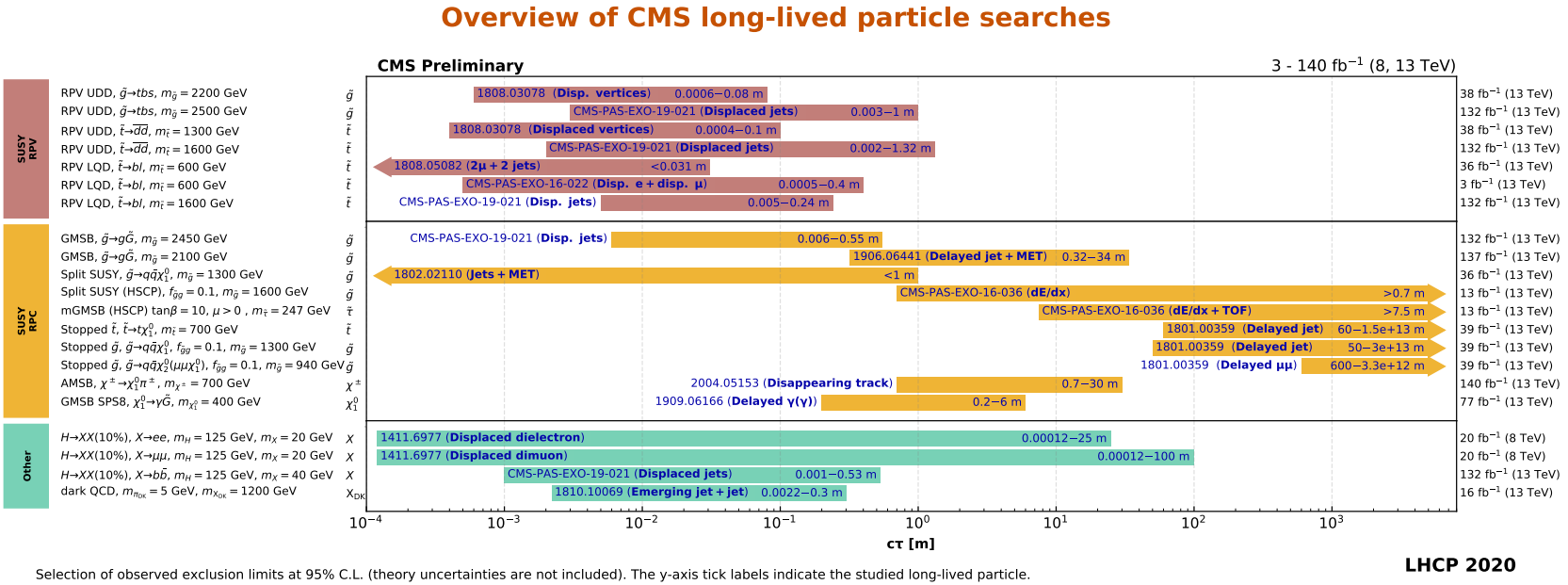


Figure 4.13: CMS exclusion limits with 95% confidence level of several benchmark models with LLP signatures [56]. The horizontal axis represents the lifetime of the LLP.

Part II

Search for Inelastic Dark Matter with CMS

The next eight chapters detail a search for inelastic dark matter with the CMS detector at CERN's Large Hadron Collider (LHC). CMS is a general-purpose detector designed to reconstruct and identify particle products of proton-proton collisions at the LHC. Collisions happen every 25 ns and many different physics processes can occur during each collision. The goal of the inelastic dark matter analysis is to identify production of heavy dark matter via a dark photon mediator in these events, with subsequent decay to visible products such as muons. The expected signature is striking and unique, consisting of significant missing transverse momentum collimated with a pair of displaced, soft, and narrow muons. This is the first search for inelastic dark matter performed at a hadron collider.

Chapter 5 introduces the inelastic dark matter model and the parameters of interest, discussing some of the subtleties involved. Chapter 6 details the experimental apparatus, the LHC complex and CMS detector. Chapters 7 to 10 provide details about the analysis: datasets and triggers used; reconstruction and identification of physics objects; event selection; and background estimation, respectively. Finally, Chapter 11 discusses the systematic uncertainties in the analysis and Chapter 12 concludes with preliminary results and an accompanying discussion.

CHAPTER 5

INELASTIC DARK MATTER AND SEARCH STRATEGY

The inelastic dark matter (iDM) model postulates a more complex dark sector than a single WIMP. There are two dark matter states close in mass which inelastically couple to each other. The lighter particle (denoted here as χ_1) is stable and could lead to the observed thermal-relic equilibrium abundance, while the heavier state (χ_2) is unstable and decays to χ_1 via emission of an off-shell dark photon. As discussed in Chapter 3, the dark photon is the vector mediator of a new $U(1)_D$ gauge symmetry in the dark sector and mixes kinetically with the SM $U(1)_Y$ mediator [66]. The kinetic mixing effectively couples the SM and dark sectors, leading to a detectable signature in proton-proton collisions.

5.1 iDM model

The phenomenology of dark photons and kinetic mixing is discussed extensively in Chapter 3. Here I include an additional description of the dark matter content of the iDM model.[1] Similar to the WIMP case in Section 3.2, we postulate a fermion dark matter field χ. In terms of its Weyl fermion components, the Dirac spinor charged under the dark $U(1)_D$ can be written as:

$$\chi = \begin{pmatrix} \eta \\ \xi^\dagger \end{pmatrix}. \tag{5.1}$$

This leads to the same vector current as in Eq. (3.29):

[1] The theoretical model and search are based on [66], where this search is proposed.

$$\mathcal{L}_\chi = i\,\bar{\chi}\gamma^\mu\partial_\mu\chi - \frac{1}{2}\left(m_\chi\,\chi\chi + \text{h.c.}\right), \tag{5.2}$$

except that we now choose to write it in terms of its Weyl components (and denote $m_\chi \equiv m_D$ to differentiate between Dirac and Majorana masses):

$$\mathcal{L}_\chi = i\left(\eta^\dagger\overline{\sigma}^\mu\partial_\mu\,\eta - \xi^\dagger\overline{\sigma}^\mu\partial_\mu\,\xi\right) - \left(m_D\,\eta\,\xi + \frac{m_\eta}{2}\eta\eta + \frac{m_\xi}{2}\xi\xi + \text{h.c.}\right). \tag{5.3}$$

Majorana masses are only allowed if the $U(1)_D$ symmetry is spontaneously broken in the dark sector. This is a built-in but valid assumption of the model. Unlike the SM photon, the dark photon is not experimentally protected from acquiring mass and, similarly, the gauge symmetry is not protected from spontaneous breaking. In the limit of small Majorana masses compared to the Dirac mass, diagonalization of the mass matrix gives rise to an inelastic coupling between the mass eigenstates. For the case where $m_\eta = m_\xi = m_M$, in particular, the mass eigenstates are:

$$\begin{aligned}\chi_1 &= \frac{i}{\sqrt{2}}\,(\eta - \xi)\\ \chi_2 &= \frac{i}{\sqrt{2}}\,(\eta + \xi),\end{aligned} \tag{5.4}$$

with the associated eigenvalues $m_{1,2} = m_D \mp m_M$. The vector current arising from the Lagrangian in Eq. (5.3) becomes inelastic:

$$\mathcal{J}^\mu = i\left({\chi_1}^\dagger\overline{\sigma}^\mu\chi_2 - {\chi_2}^\dagger\overline{\sigma}^\mu\chi_1\right) \equiv \mathcal{J}^\mu_{\text{iDM}}, \tag{5.5}$$

where, similarly to Ref. [66], we label the current with "iDM" to emphasize its inelastic nature. In the more general case where Majorana masses are not the same, the

diagonalization results in an additional elastic component in the current:

$$\mathcal{J}^{\mu} = \frac{m_D}{\sqrt{m_D^2 + (m_\eta - m_\xi)^2/4}} \mathcal{J}^{\mu}_{\mathrm{iDM}} + \frac{m_\xi - m_\eta}{\sqrt{4m_D^2 + (m_\xi - m_\eta)^2}} \left({\chi_2}^{\dagger} \overline{\sigma}^{\mu} \chi_2 - {\chi_1}^{\dagger} \overline{\sigma}^{\mu} \chi_1 \right) \tag{5.6}$$

as well as more complicated mass eigenvalues:

$$m_{1,2} = \sqrt{m_D^2 + \frac{1}{4}\left(m_\eta - m_\xi\right)^2} \pm \frac{1}{2}\left(m_\eta + m_\xi\right). \tag{5.7}$$

The model used throughout the analysis, based on Ref. [66], assumes a mass splitting between χ_1 and χ_2 that is not much smaller than the Dirac mass, so the elastic coupling effect must be included. In what follows, we take $m_\eta = \Delta$, the mass splitting, and $m_\xi = 0$.

5.2 Cross section studies

In close contact with iDM theorists, a detailed study of signal cross sections was performed with MadGraph5_aMC@NLO [67] (MG) v.2.6.0 to understand the various subtleties of the model. There are several parameters to consider in the context of inelastic dark matter:

- m_1, the mass of light dark matter (χ_1);
- $\Delta \equiv m_2 - m_1$, the absolute mass splitting between heavy (χ_2) and light dark matter;
- $m_{A'}$, the mass of the dark photon;
- Γ and $c\tau = 1/\Gamma$, the decay width (and lifetime) of heavy dark matter;

- α_D, the equivalent to electromagnetic coupling in the dark sector;
- ϵ, the kinetic mixing coefficient between hypercharge and the dark photon.

Due to the many parameters in the model, there is in principle a large freedom in the choice of values that still correspond to allowed physics. Therefore, a standard set of reference values is typically defined. Some care must be taken when comparing results for inelastic dark matter to other models featuring dark photons—it is easy to overstate bounds otherwise [66].

We set the ratio $m_{A'}/m_1 = 3$ constant throughout the analysis. This is also done in the theory paper because it represents the primary region of relevant phase space. The dark photon must be heavier than the sum of dark matter masses ($m_{A'} > m_1 + m_2$) for an on-shell dark photon decay to occur. Furthermore, since the mass splittings between the two dark matter states are small (10 or 40 percent), setting $m_{A'} = 3\,m_1$ ensures the dark photon is more massive in all cases, but not by much. Other corners of the allowed phase space have already been ruled out or lead to reduced cross sections [66].

A general feature of models with long-lived signatures is an inverse correlation between the production cross section times branching ratio of a given process and the displacement of the LLP. This is typically the case because the decay width grows with the relevant coupling constants between new physics and the SM. In the specific case of inelastic dark matter, the decay width of heavy dark matter into light dark matter, in the limit $\Delta \ll m_1, m_{A'}$, is [66]:

$$\Gamma(\chi_2 \to \chi_1\, e^+\, e^-) = \frac{4\,\epsilon^2\,\alpha\,\alpha_D\,\Delta^5}{15\,\pi\,{m_{A'}}^4}. \tag{5.8}$$

Here the decay width is proportional to ϵ^2. Combined with an additional factor of ϵ^2

from dark photon production, very large displacements are extremely suppressed.

Eq. (5.8) defines a relationship (and hence a constraint) between the lifetime, $c\tau$, and ϵ, α_D, $m_{A'}$, and Δ. Moreover, when we define a set of simulated signal samples to produce and analyze, additional freedom is lost in choosing the parameter space. For example, setting m_1 to 5 GeV (and hence $m_{A'}$ to 15 GeV, since we always take $m_{A'}/m_1 = 3$), Δ to 0.5 GeV, and $c\tau$ to 10 mm, there is only a two-dimensional space left to choose from: ϵ and α_D. However, the combination of these two parameters must also match the correct $c\tau$ for each sample. One way to see this is to rearrange Eq. (5.8):

$$\epsilon^2 \alpha_D = \frac{15\,\pi}{4\,\alpha}\left[\frac{(3\,m_1)^4}{c\tau\,\Delta^5}\right]. \tag{5.9}$$

The combination of $\epsilon^2\,\alpha_D$ must match the particular m_1, Δ, and $c\tau$ for each simulated sample. Figs. 5.1 and 5.2 illustrate the relationship between ϵ and α_D for several representative samples produced for the analysis. The vertical red lines define two standard choices of α_D ($\alpha_D = 0.1$ and $\alpha_D = \alpha_{EM}$). For each sample and for each of the two α_D, the corresponding ϵ has to be picked in order to arrive at the correct cross section.

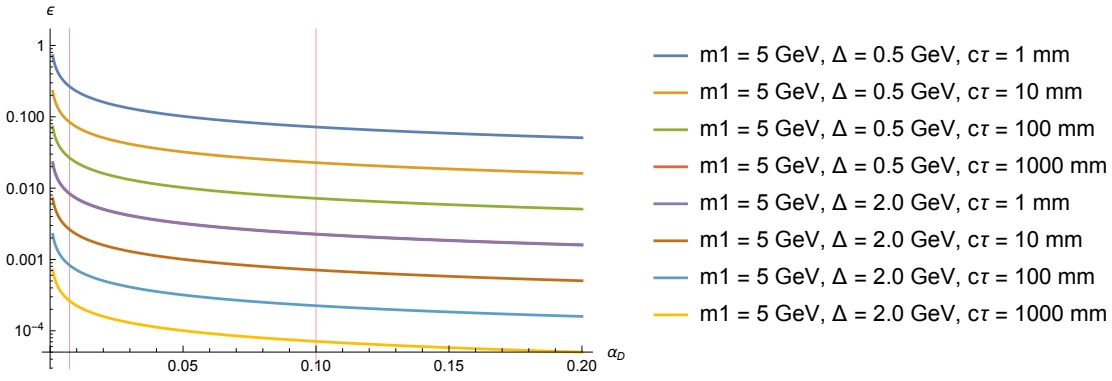


Figure 5.1: Kinetic mixing coefficient ϵ as a function of α_D for different inelastic dark matter parameters, setting $m_1 = 5$ GeV (the mass of the lighter dark matter). The combination of α_D and ϵ has to lie on the lines of constant $c\tau$ for each sample.

Due to the plethora of parameters in the iDM model, observed and projected limits

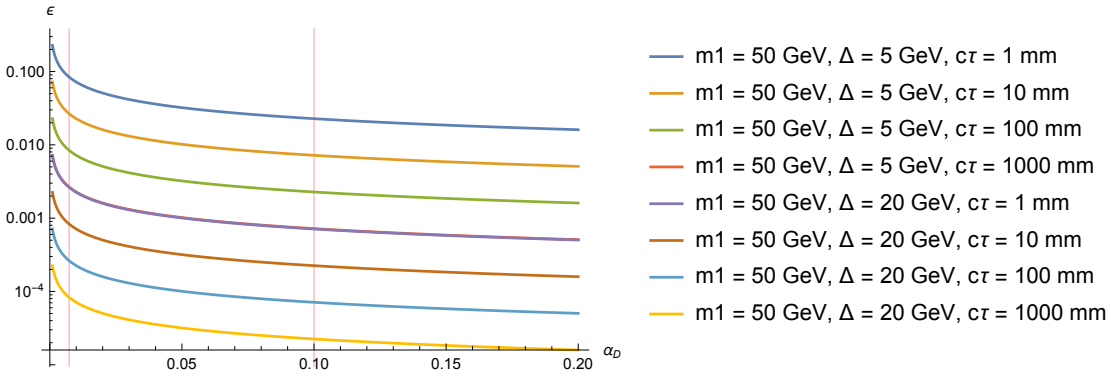


Figure 5.2: Kinetic mixing coefficient ϵ as a function of α_D for different inelastic dark matter parameters, setting $m_1 = 50\,\text{GeV}$ (the mass of the lighter dark matter). The combination of α_D and ϵ have to lie on the lines of constant $c\tau$ for each sample, as discussed in the text.

are usually presented as a 2D plot with the mass of the light dark matter m_1 in the horizontal axis and the quantity $y \equiv \epsilon^2 \alpha_D \left(m_1/m_{A'}\right)^4$ in the vertical axis. These two quantities are useful because in the limit of small mass splitting, $\Delta \ll m_1, m_{A'}$, the dark matter annihilation rate depends predominantly on m_1 and on the specific combination of parameters y. It is possible to maintain the same y and hence the same annihilation rate by varying its constituent parameters in a specific ratio. Furthermore, y is a convenient quantity for thermal-relic dark matter models, as the annihilation rate depends to a good approximation on y and on the dark matter mass alone. This is readily seen by the solid black line in Fig. 5.3. While the ratio $m_1/m_{A'}$ is fixed to 3 in the analysis, it is still beneficial to show limits using the standard quantities in order to directly compare with theoretical expectations. Two examples of projected sensitivity plots are shown in Figs. 5.3 and 5.4.

It is helpful to study the behavior of the expected sensitivity contours. Representing the mass splitting Δ as a fraction of m_1, i.e. $\Delta = x\, m_1$, we can rewrite Eq. (5.8) as:

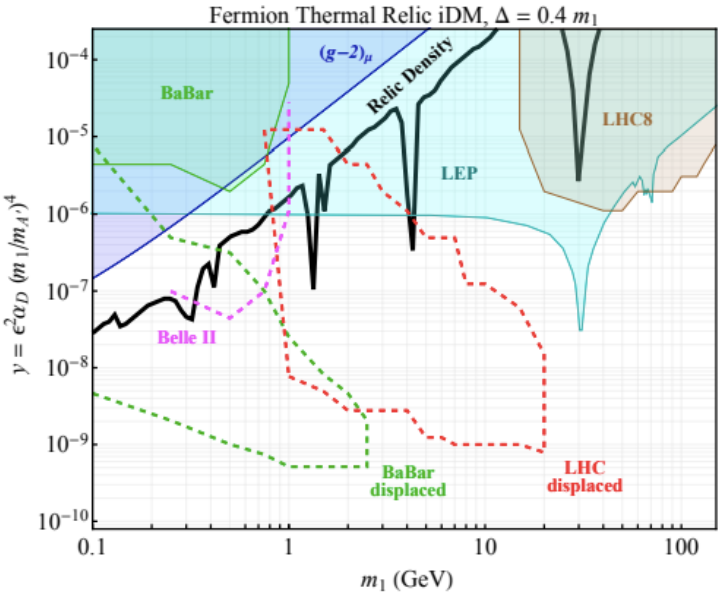


Figure 5.3: Theoretical expected sensitivity plot with the choices $\alpha_D = 0.1$, $m_{A'}/m_1 = 3$, and $\Delta = 0.4\, m_1$ [66]. In this particular example, $y = 0.00123\, \epsilon^2$, and hence $c\tau \approx 2.5$ cm for $\epsilon \approx 0.001$ (or equivalently $y \approx 10^{-9}$) and $m_1 \approx 1$ GeV (see text).

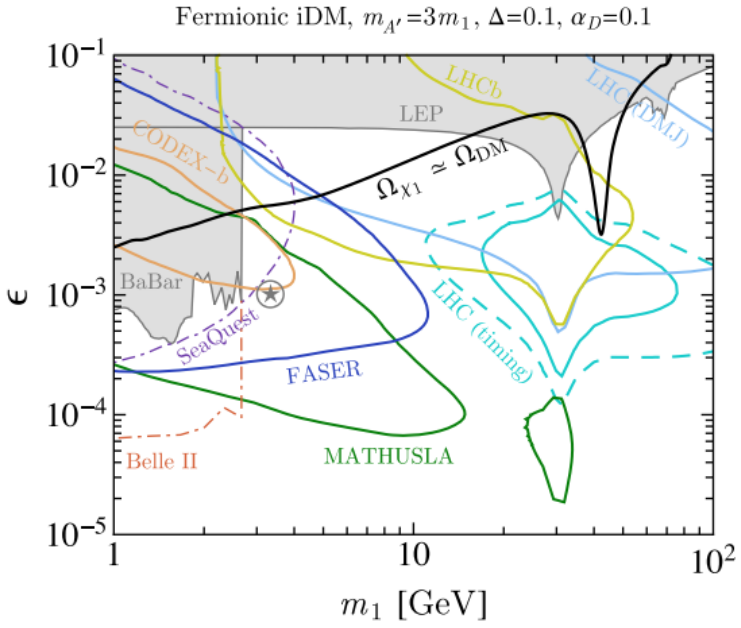


Figure 5.4: Theoretical expected sensitivity plot with the choices $\alpha_D = 0.1$, $m_{A'}/m_1 = 3$, and $\Delta = 0.1\, m_1$ [32]. Note that in this case the sensitivity is plotted as a function of ϵ instead of y (though the relation $y = 0.00123\, \epsilon^2$ still holds since $m_{A'}/m_1 = 3$). The LHC projections are the light-blue contours labeled "*LHC (DMJ)*".

$$\begin{aligned}\Gamma(\chi_2 \to \chi_1\, e^+\, e^-) &= \frac{4\,\alpha}{15\,\pi}\left[\epsilon^2\,\alpha_D\left(\frac{m_1}{m_{A'}}\right)^4\right] x^5\, m_1 \\ &= \frac{4\,\alpha\, x^5}{15\,\pi}\, y\, m_1.\end{aligned} \tag{5.10}$$

The corresponding lifetime is then given by:

$$c\tau = \frac{1}{\Gamma} = \frac{15\,\pi}{4\,\alpha\, x^5}\left(\frac{1}{y\, m_1}\right), \tag{5.11}$$

so it is inversely proportional to both m_1 and y. For the specific case of Fig. 5.3, $x = 0.4$ and the expression for the lifetime as a function of y (or ϵ, given $\alpha_D = 0.1$) and m_1 becomes:

$$c\tau \approx \frac{1.6\times 10^5}{y\, m_1} = \frac{3.1\times 10^{-9}}{y\,(m_1/\,\mathrm{GeV})}\ \mathrm{cm} \approx \frac{2.5}{(\epsilon/0.001)^2\,(m_1/\,\mathrm{GeV})}\,\mathrm{cm}. \tag{5.12}$$

For, say, $m_1 \approx 1\,\mathrm{GeV}$ and $\epsilon \approx 10^{-3} \Rightarrow y \approx 10^{-9}$; this gives $c\tau \approx 2.5\,\mathrm{cm}$. Since we rely on the displacement of the heavy dark matter state to primarily drive the sensitivity of the analysis, lower masses and lower values of the dimensionless coupling y should result in the most sensitive limits. However, this conclusion is dampened by the aforementioned suppression of the cross section with increasing lifetime. Moreover, for lower A'_μ masses, the dark photon recoiling off the initial-state radiation (ISR) jet can be boosted. Taking the boost factor into account, Fig. 5.5 shows the average *lab frame* decay length of χ_2 as a function of its mass [66].

The enhancement of the cross section when the mass of the dark photon is close to the Z peak should also be factored in, which provides some balance to the sensitivity

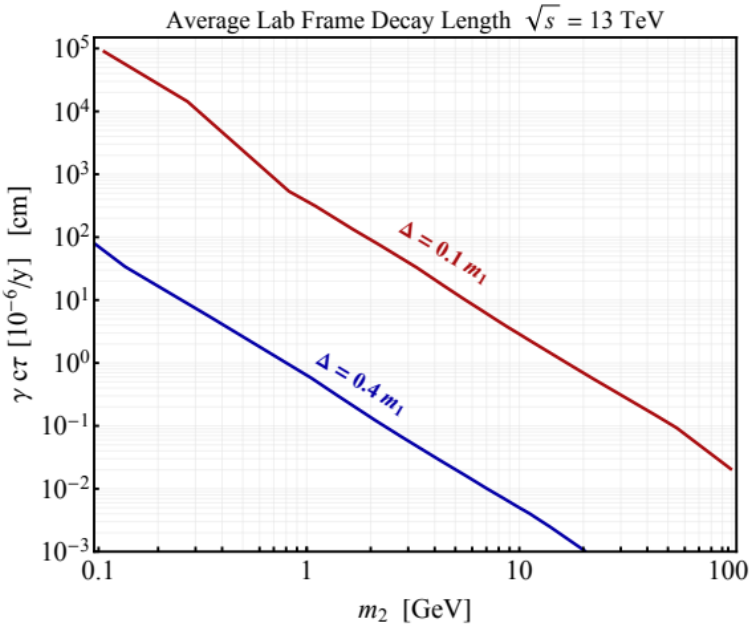


Figure 5.5: Average lab frame decay length of the heavy dark matter χ_2 as a function of its mass for two mass splittings: $\Delta = 0.1\, m_1$ and $\Delta = 0.4\, m_1$. Events are selected with an energetic ISR jet ($p_{\mathrm{T}} > 120\,\mathrm{GeV}$) and the decay length is normalized to $y = 10^{-6}$ for easier comparison with previous plots [66].

at higher masses. This enhancement is due to the coupling between fermions and dark photon, which is inversely proportional to the mass difference between the Z and A' [68]:

$$g_{\bar{f}fZ'} \simeq -\epsilon \left(\frac{\overline{M}_Z^2 \cos\theta_w e Q_f - {m_{A'}}^2 g_y Y_f}{\overline{M}_Z^2 - {m_{A'}}^2} \right). \tag{5.13}$$

To further study this dependence, we calculated in MadGraph the cross section for a range of dark photon masses. The goal was to understand if the observed enhancement is reasonable and in accordance with theoretical expectations. 10,000 events per dark photon mass were generated in the range [3, 240] GeV in steps of 1 GeV. For this particular study, ϵ was set to 0.01 and α_D to 0.1. This neglects the discussion in previous paragraphs since the same couplings are set for all masses, but it is useful to do so here in order to study the effect of Z enhancement by itself. The resulting cross sections are shown in Fig. 5.6.

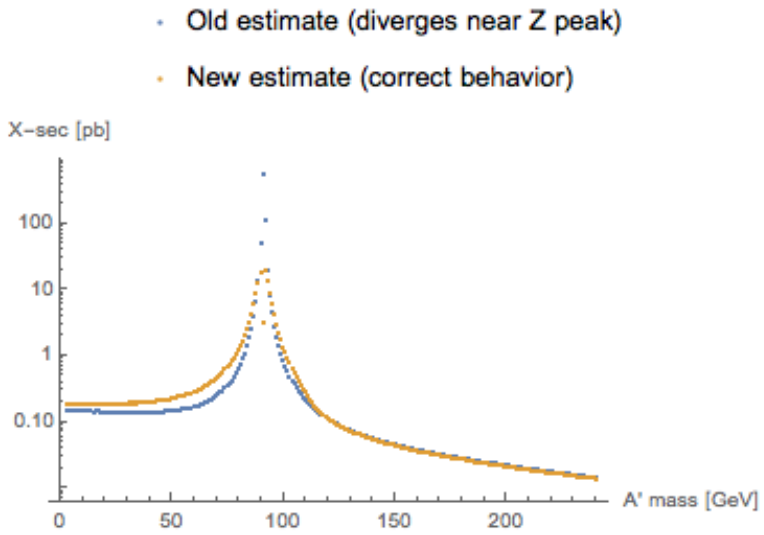


Figure 5.6: Dark photon production cross section as a function of its mass. There is an 100-fold enhancement in the cross section near the Z peak region, due to mixing between the two bosons. The blue curve refers to a previous calculation which did not properly take the mixing into account, resulting in a divergent cross section on the Z mass. The orange curve has the correct behavior and does not diverge, indicating the true enhancement around the Z region. The chosen parameters for this plot are $\alpha_D = 0.1$, $\epsilon = 0.01$, and $\Delta = 0.1\, m_1$.

Indeed, there is an enhancement around the Z peak, with roughly a factor of 100 difference between the cross section at 90 GeV *vs.* 60 GeV. The immediate Z peak region (± 1 GeV) should be avoided, since there is destructive interference between the A' and the Z when their mass difference is smaller than the Z width. This corresponds to the single orange point at $m_{A'} = 91$ GeV in Fig. 5.6. The observed enhancement is also in agreement with other studies of dark photon production, for example in Fig. 5.7.

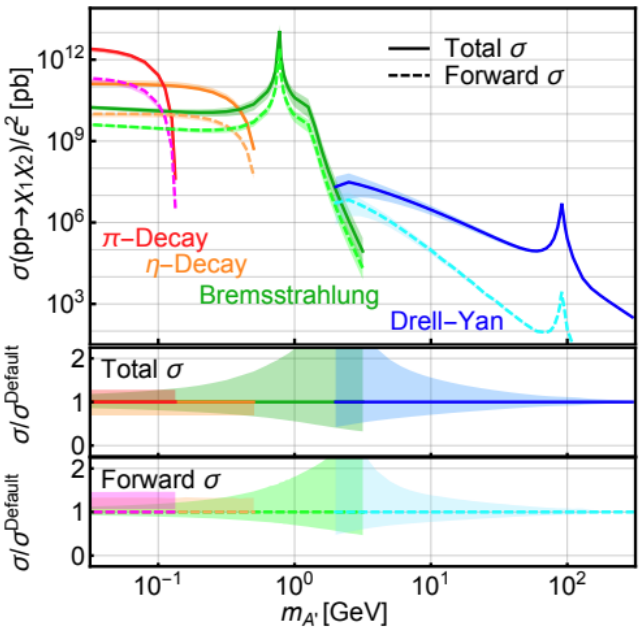


Figure 5.7: Reference dark photon production cross section as a function of its mass [32]. The parameters chosen for this plot are: $\alpha_D = 0.1$, $m_{A'}/m_1 = 3$, $\epsilon = 0.001$, and $\Delta = 0.1\, m_1$. The ≈100 enhancement factor is visible near the Z peak.

5.3 Search strategy

The iDM signature expected in CMS is quite striking. With two dark matter states and a dark photon mediator, processes like the one shown in Fig. 5.8 are possible. The reaction proceeds through the production of a dark photon via the kinetic mixing coupling, which then decays to χ_2 and χ_1. The heavier χ_2, unstable, eventually decays back to χ_1 and to an off-shell dark photon. Finally, the off-shell dark photon immediately decays back to the SM, producing a pair of displaced, collimated, and soft leptons. Unlike traditional dark photon models where the dark photon is the LLP, the lifetime of χ_2 depends on many parameters as discussed in Section 5.2.

Furthermore, the entire dark sector system needs to recoil against an ISR jet for triggering purposes. The chain can be summarized as:

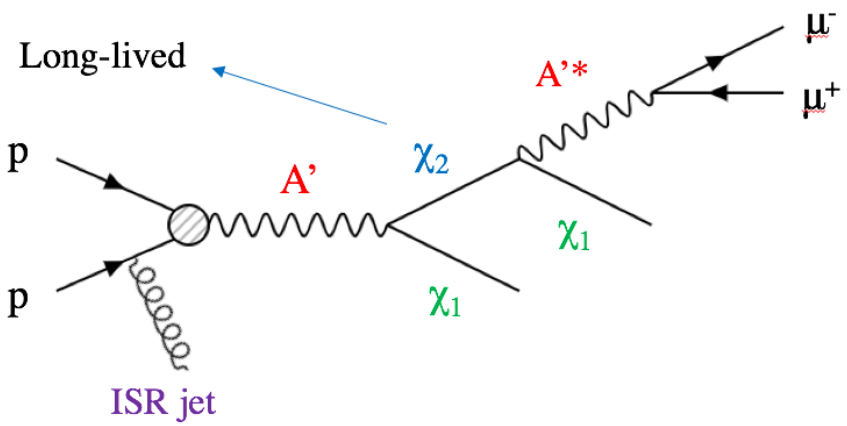


Figure 5.8: Example Feynman diagram of a physics process predicted by inelastic dark matter at the LHC. Protons collide producing an on-shell dark photon, which decays to two different dark matter states. The lighter state (denoted χ_1, in green) escapes the detector, while the heavier one (denoted χ_2, in blue) travels a macroscopic distance before itself decaying. The decay of the heavy χ_2 state produces an additional light χ_1 particle as well as a pair of soft, narrow muons via an off-shell dark photon.

$$\mathrm{pp} \to j\,(A' \to\ \chi_1\,[\chi_2 \to \chi_1\,l^+\,l^-]) \to j\,\chi_1\,\chi_1\,l^+\,l^-, \tag{5.14}$$

where the final state consists of two light and stable χ_1 particles escaping the detector, two leptons, and a hard jet. The escaping χ_1 can be reconstructed in CMS as $p_{\mathrm{T}}^{\mathrm{miss}}$. Since leptons are displaced, collimated, and soft, this first iteration of the analysis is restricted to final-state muons only. Muons are especially well reconstructed by CMS and benefit from the standalone muon chamber detectors, increasing the accessible lifetime. Moreover, the incidence of photon conversion backgrounds in the electron channel might be significant [66]. More specialized techniques for reconstructing soft and displaced electrons are likely needed for a robust sensitivity in this channel. This is left for future studies.

The general search strategy of the analysis is to exploit the unique signature predicted by inelastic dark matter. The main features include significant missing transverse mo-

mentum, $p_{\mathrm{T}}^{\mathrm{miss}}$, collimated with a pair of soft and narrow muons, which are significantly displaced from the primary vertex. The entire $p_{\mathrm{T}}^{\mathrm{miss}}$-dimuon system opposes an ISR jet, exploited for triggering. A schematic diagram of the expected signature is shown in Fig. 5.9.

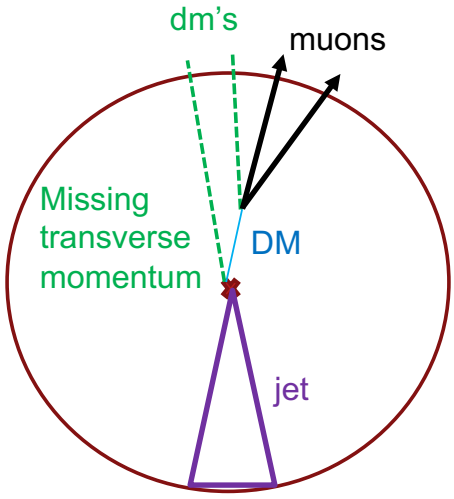


Figure 5.9: Sketch of the expected signature produced by inelastic dark matter in the CMS detector. The main components include a pair of soft, narrow, and displaced muons, collimated with $p_{\mathrm{T}}^{\mathrm{miss}}$ and recoiled against an ISR jet.

Some of the kinematic features of the signal are summarized in Fig. 5.10. The muon p_{T} distribution in iDM events is quite soft due to the narrow mass splitting between dark matter states. The muons also form fairly collimated pairs, particularly for lower mass-splitting samples. In the literature, this signature is commonly referred to as *lepton-jets*. However, this nomenclature is not used here because no special lepton-jet reconstruction algorithm is employed, which could be incorrectly assumed from the "lepton-jet" notation. Since the muons in iDM are collimated but not extremely so, a special lepton-jet reconstruction did not seem necessary or especially beneficial.

Additional signal characteristics seen in Fig. 5.10 include the collimation between the dimuon pair and $p_{\mathrm{T}}^{\mathrm{miss}}$ and a large displacement of the dimuon pair from the interaction point. These two features are exploited later as part of the data-driven background

estimation strategy and largely drive the sensitivity to iDM.

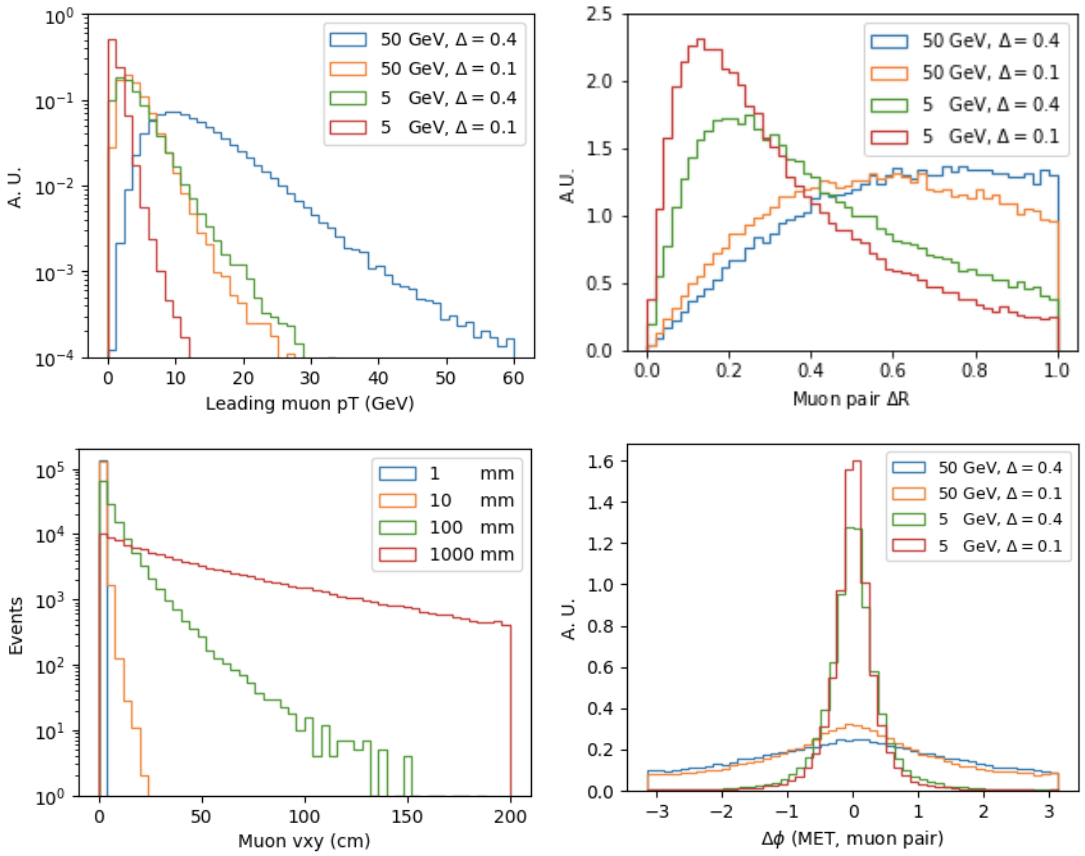


Figure 5.10: Generator-level observables of inelastic dark matter signal. Upper left: leading muon p_{T}. Upper right: dimuon pair dR. Lower left: dimuon vertex v_{xy}. Lower right: azimuthal separation between $p_{\mathrm{T}}^{\mathrm{miss}}$ and muons, $\Delta\phi(p_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{dimuon})$. The legend refers to the mass of the light dark matter state, m_1, and the mass splitting $\Delta \equiv m_2 - m_1$, except for the lower left plot where different lifetimes $c\tau$ are compared. No event selection is applied.

The soft p_{T} spectrum and displaced nature of the final-state muons prevent the use of muon-based triggers for this analysis, several of which were verified to be inefficient. Instead we rely primarily on $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers. These are discussed in Chapter 7. Future specialized triggers based on a combination of $p_{\mathrm{T}}^{\mathrm{miss}}$ and muons may be able to enhance the sensitivity of the search.

To obtain the best possible measurement of muons in this difficult environment (i.e.

soft, collimated, and displaced), two separate muon reconstruction objects are used. A displaced standalone muon reconstruction that employs only muon chamber information is the baseline object, but global muon objects that include tracker information are also considered in a match-and-replace procedure. Global muons offer a better performance overall, though their reconstruction efficiency sharply drops after the end of the tracker region (around $z \approx 60\,\text{cm}$). We found that incorporating both objects to the analysis provides the greatest benefits. This procedure is presented in Chapter 8.

Some of the main event selection requirements include one hard jet with $p_{\text{T}} > 80\,\text{GeV}$ consistent with ISR, $p_{\text{T}}^{\text{miss}} > 200\,\text{GeV}$, and two muons with $p_{\text{T}} > 5\,\text{GeV}$ and a reconstructed dimuon vertex inside the CMS volume ($r < 940\,\text{cm}$, $z < 760\,\text{cm}$). The full event selection is described in Chapter 9.

The unique nature of the iDM signature translates to an expected low-background search. As discussed in Chapter 9, after requiring both a narrow dimuon pair and $p_{\text{T}}^{\text{miss}}$-dimuon collimation, all backgrounds except QCD and Z/W+jets (with misidentified muons) provide negligible contributions to the event yield in the signal region. To further reduce these backgrounds, the long-lived nature of the dimuon pair is exploited. Observables such as dimuon vertex position and $\Delta\phi(p_{\text{T}}^{\text{miss}}, \text{dimuon})$, used in data-driven background estimation, allow the further narrowing of the signal region into a signal-enriched bin where sensitivity to iDM is greatest. The background estimation strategy is described in Chapter 10.

CHAPTER 6

THE LARGE HADRON COLLIDER AND CMS DETECTOR

This chapter describes the apparatus used in the search for iDM. The two main components are the LHC and the CMS detector. The LHC accelerates bunches of protons up to a center-of-mass energy of 14 TeV and then collide them. The CMS detector is used to reconstruct and analyze the collision debris in search of new physics associated to dark matter or other landscapes. Section 6.1 describes the accelerator complex of the LHC and Section 6.2 the CMS detector setup.

6.1 Large Hadron Collider

The LHC is the world's most powerful particle accelerator. It can accelerate protons up to a center-of-mass energy ($\sqrt{s}$) of 14 TeV, although it currently operates at $\sqrt{s} = 13$ TeV. The LHC is located in Geneva, Switzerland, and constructed 100 m underground, inside a 27 km circumference tunnel which previously housed the Large Electron Positron collider (LEP).

Protons are accelerated in bunches spaced out 25 ns in time. Two trains of bunches travel in opposite directions and inside different beam pipes. The two bunches are made to collide at four different points around the LHC ring, which house the four main detectors: CMS [69], ATLAS [70], ALICE [71], and LHCb [72]. Of these, CMS and ATLAS are general-purpose detectors designed to study a diverse range of physics, while LHCb and ALICE are more specialized detectors—the former is a forward spectrometer to study the physics of b quarks and the latter is a heavy-ion machine.

For proton-proton (pp) collisions, the LHC's design peak luminosity is near $1 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The amount of data collected and analyzed by experiments is typ-

ically quoted as the luminosity integrated over time. The integrated luminosity has units of cm^{-2}, or 1/barns (b^{-1}), where $1\,\mathrm{b} = 10^{-34}\,\mathrm{cm}^2$. The total number of pp collisions, or *event rate*, is the product of the luminosity and the total pp cross section, σ_{pp}, given in units of area or barns. This relationship is summarized as:

$$R = \int \sigma_{\mathrm{pp}}\, L\, dt, \tag{6.1}$$

where R is the event rate, L is the instantaneous peak luminosity, and the integral is over time. The total pp cross section at $\sqrt{s} = 13\,\mathrm{TeV}$ is $(110.6 \pm 3.4)\,\mathrm{mb}$ [73]. The total number of events available to analyze depends not only on the particle physics of SM interactions but also on the beam dynamics that accelerator physicists strive to optimize.

The luminosity itself is a function of several parameters. In a central beam collider such as the LHC, it is given by [74]:

$$L = \frac{N_1\, N_2\, n_b\, f_{\mathrm{free}}}{A\, F}, \tag{6.2}$$

where N_1 and N_2 are the number of particles per bunch in the two beams; n_b is the number of bunches per beam; f_{free} is the revolution frequency; A is the overlapping cross sectional area of the beams in position-angle phase space; and F is a geometric correction factor to account for an angular offset between them (it would be impossible to accelerate the beams with a head-on geometry).

Fig. 6.1 reports the total integrated luminosity $L_{\mathrm{int}} = \int L\, dt$ delivered to and recorded by CMS in Run 2, by year (right) and cumulatively (left). The Run 2 campaign started in 2015 and ended in 2018, collecting $150.26\,\mathrm{fb}^{-1}$ of data [**?**]. The amount of data from 2016 to 2018 validated for analysis (the total dataset used in this search) has an integrated

luminosity of 137.19 fb^{-1}.

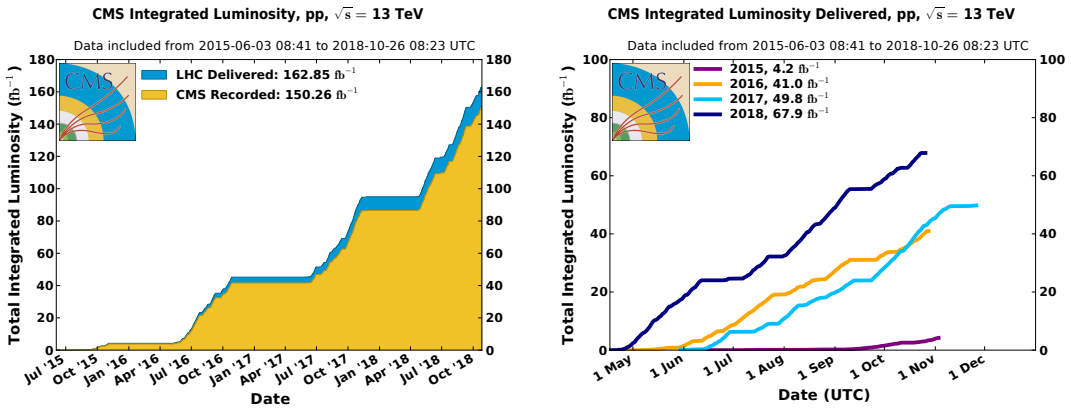


Figure 6.1: Integrated luminosity over time delivered by the LHC and recorded by CMS; total accumulated (left) and separately by year (right) [**?**].

An overview of CERN's accelerator complex is shown in Fig. 6.2. Proton acceleration happens through a series of increasingly energetic boosters, starting from a single bottle of hydrogen gas (H_2). Protons are extracted by dissociating the gas into individual hydrogen atoms and stripping the electrons off with an electric field at LINAC2. Protons reach the proton synchrotron booster (PSB) with an energy of 50 MeV where they are further accelerated to 1.4 GeV. The process continues with the proton synchrotron (PS), where their energy is increased to 25 GeV, and the super proton synchrotron (SPS), to 450 GeV. Finally, protons are delivered to the LHC ring in two beam pipes, where they are further accelerated to the final energy of 6.5 TeV.

The accelerator apparatus consists of several types of electromagnets that perform a variety of functions. Superconducting radio-frequency (RF) cavities are used to accelerate protons by creating standing waves that boost the energy of charged particles passing through these cavities and impart a small force during each passage. The RF electromagnetic fields oscillate at 400 MHz and transfer energy to protons over the course of several turns. This process happens repeatedly inside each booster stage until protons attain sufficient energy to be transferred to the next stage.

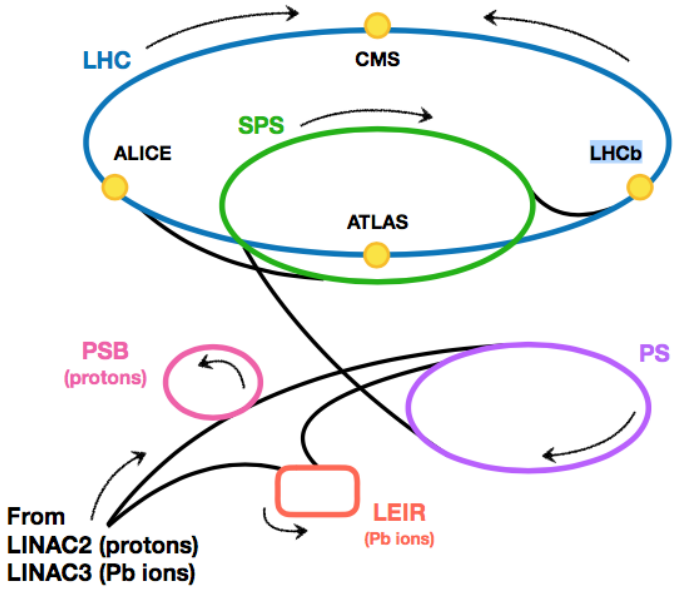


Figure 6.2: CERN's accelerator complex [75]. Protons are accelerated in stages, starting from the LINAC, then passing through the PSB and PS, and finally being delivered to the LHC where they are accelerated to the final design energy of 13 TeV.

Dipole magnets are used to steer the beam around the boosters in circular trajectories. The dipole magnets of the LHC ring are made of Niobium-Titanium (NbTi) superconductors and cooled to a temperature of 1.9 K with superfluid Helium. There are 1,232 dipoles with a strength of 8.33 T and a length of 15 m each. A schematic diagram of the dipole cross section and a map of the generated magnetic field is shown in Fig. 6.3.

An additional noteworthy electromagnet component in the LHC ring is the quadrupole magnet, which focuses (defocuses) the beam along the horizontal (vertical) directions, and vice-versa. There are 392 such magnets and they provide customizable tunings of beam parameters in order to maximize the beam efficiency and hence the luminosity.

The acceleration cycle takes about 45 minutes from injection. During this time, protons circulate in the LHC ring until they reach the design energy, when they become ready for collisions. These take place in one of the four points around the ring corresponding

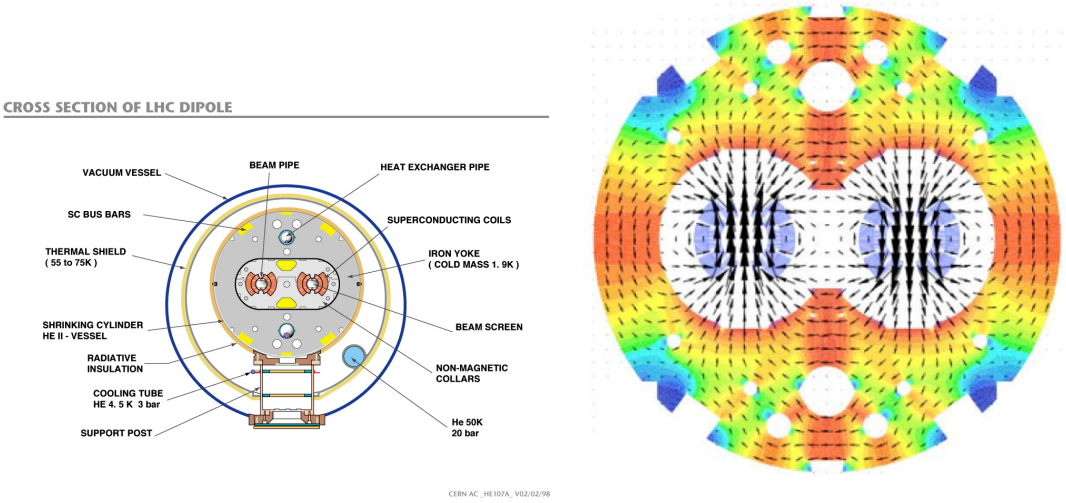


Figure 6.3: Transverse cross section of the LHC's magnetic dipole (left) and corresponding magnetic field map (right) [76, 77].

to the main LHC experiments mentioned earlier. A typical fill of stable proton beams lasts roughly 10 hours after the design energy is reached. The luminosity delivered to the experiments decreases steadily during the course of the fill until it reaches a threshold (ordinarily half of the initial luminosity), at which point the beam must be topped off, and the acceleration cycle repeats.

6.2 Compact Muon Solenoid

The CMS experiment is a solenoid-based detector with multiple sub-components designed to measure different types of particles originating from the central pp collisions. It has a 22 m length and 15 m diameter that is deemed compact when compared to ATLAS, the other general-purpose LHC experiment. CMS also has a higher density than ATLAS due to the instrumentation required to fit in the smaller detector volume with a total mass of 12,500 tons. The other central features of CMS are the superconducting solenoid magnet and the muon spectrometer, further justifying its name. Fig. 6.4 shows

a sketch of the experiment.

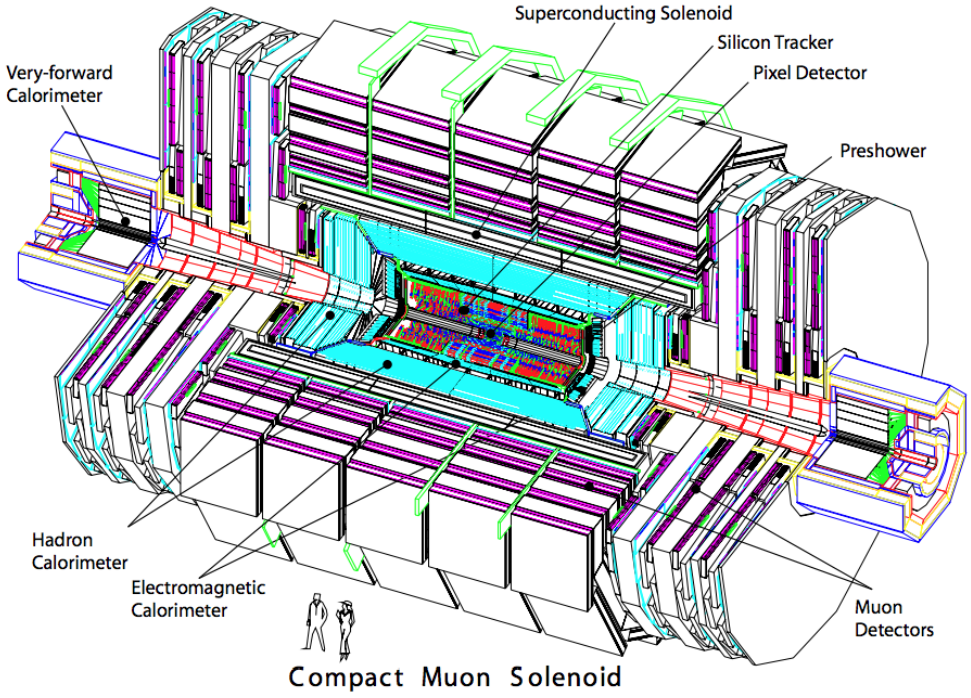


Figure 6.4: Overview of the CMS detector with a cutout of the detector's components, and two humans for scale [69].

The components of CMS are briefly discussed in the next sections, starting with the coordinate system definition in Section 6.2.1, followed by descriptions of the magnet solenoid (Section 6.2.2), silicon trackers (Section 6.2.3), electromagnetic calorimeter (Section 6.2.4), hadronic calorimeter (Section 6.2.5), muon spectrometer (Section 6.2.6), and trigger and data acquisition (Section 6.2.7). A more complete description of the apparatus can be found in [69].

6.2.1 Coordinate system

The coordinate system used in CMS is a right-handed one, with the y axis pointing vertically, the x axis radially inward, and the z axis in the counter-clockwise beam

direction. The origin of the coordinate system is the nominal collision point at the center of the experiment. The most convenient choice of coordinates, however, is cylindrical because of the symmetry of the detector. In these coordinates, r is the radial distance from the origin in the x-y (transverse) plane, and ϕ is the azimuthal angle defined as $\phi \equiv \tan^{-1} y/x$. Furthermore, the polar angle θ is defined as $\theta \equiv \tan^{-1} r/z$, but a more convenient quantity commonly used in high-energy physics experiments is the pseudo-rapidity $\eta \equiv -\ln(\tan\theta/2)$. In the limit of relativistic particles, the pseudo-rapidity approximates the rapidity y given by:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \tag{6.3}$$

The rapidity is invariant under Lorentz boost transformations along the z axis, making it (and hence pseudo-rapidity at LHC energies) an especially suitable quantity. For reference, higher values of η correspond to lower values of θ and the limit $\eta = 0$ refers to the central region of the CMS detector ($\theta = 90°$), while $\eta \to \infty$ refers to very forward activity near the beam pipe ($\theta \to 0°$).

6.2.2 Magnet solenoid

The magnet solenoid of CMS has a strong 3.8 T field along the beam direction and a 12.5 m length. It is made of 4 layers of superconducting NbTi coils reinforced with aluminum conductors, which are cooled down to 4.5 K via liquid Helium. The total mass of the solenoid is 220 tons. The solenoid encompasses both trackers and calorimeters. Fig. 6.5 shows a map of the magnetic field lines inside the solenoid.

The magnetic flux of the solenoid is returned via a set of iron yokes with total mass

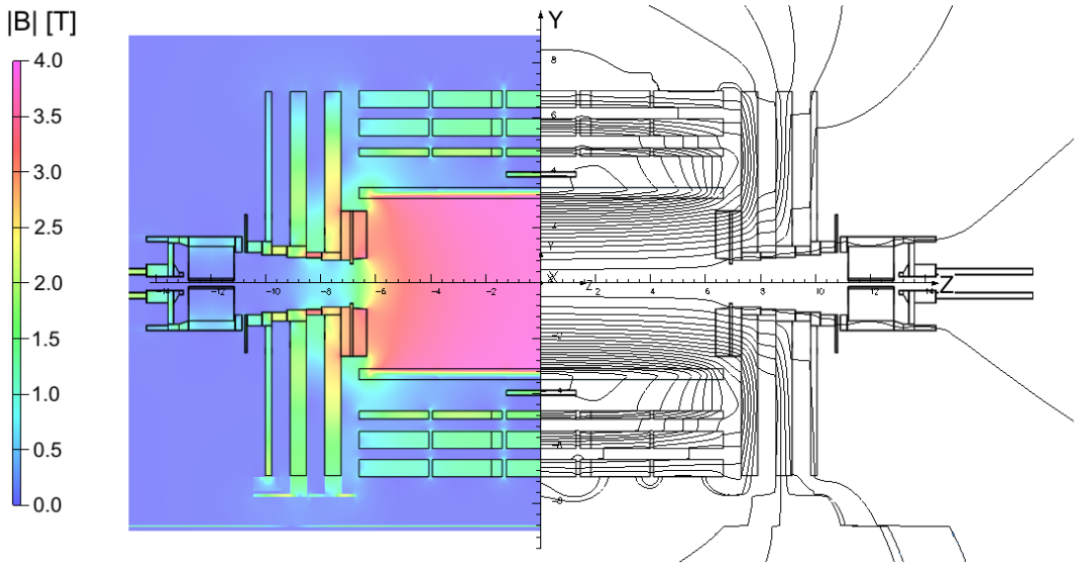


Figure 6.5: Magnetic field map of the CMS detector [78]. The peak uniform field in the center region has 3.8 T strength.

of 10,000 tons embedded in the structure of the detector. The yoke layers are interspersed with the muon chambers, which provide an additional benefit in the form of extra stoppage power for particles other than muons. This ensures a pure sample of muons arriving in the outer muon chambers.

6.2.3 Silicon trackers

The CMS silicon trackers reconstruct the trajectories of charged particles produced at the interaction point. The system is composed of a pixel and a strip silicon tracker, located inside the solenoid magnet [79]. The magnetic field created by the magnet is crucial in measuring the momentum of particles as the bending of charged particles under magnetic fields is proportional to their momentum.

Pixels are the innermost detector in CMS, with the strips located just outside them.

Pixels are able to reconstruct 3D hit positions of incident charged particles since they have 2D segmentation, while strips feature less detailed 2D hit information having only segmentation along one direction. The silicon tracker system has a full azimuthal coverage and a pseudo-rapidity coverage of $-2.5 < \eta < 2.5$. Fig. 6.6 shows a schematic layout of the silicon trackers in CMS.

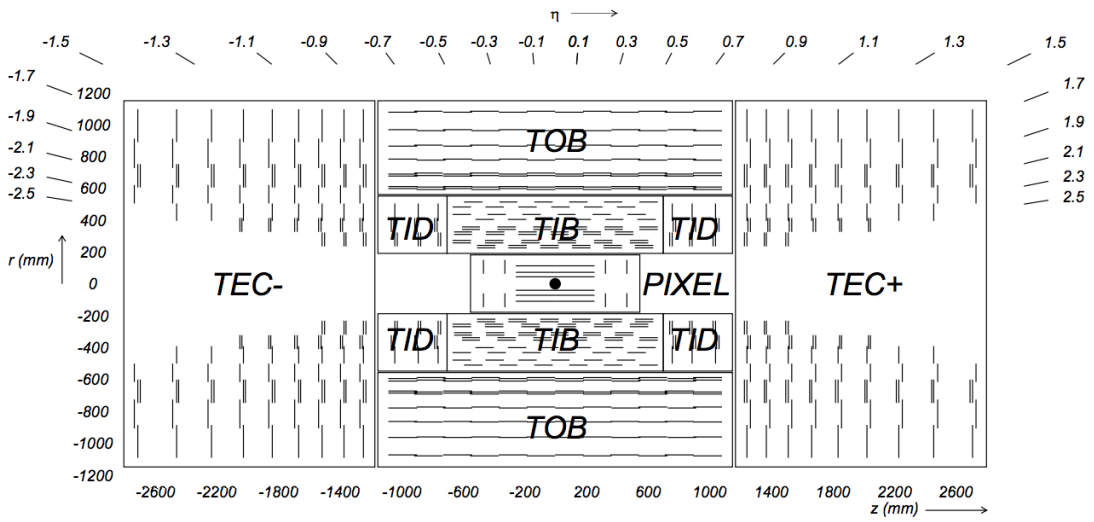


Figure 6.6: Layout of CMS's tracker [69]. Both inner and outer tracker layers are shown in this r-z projection.

Pixel tracker

The pixel tracker is located just outside the beam pipe [79]. Owing to the proximity to the proton-proton interaction point, it receives the most radiation of any sub-detector. For this reason, the pixel tracker was replaced after the 2016 run with a new, improved detector. Here we describe both 2016 and 2017/2018 pixel trackers since we use data collected in all Run 2 years. Fig. 6.7 compares the layout of the two setups.

The 2016 pixel tracker consists of 66 million silicon pixel channels arranged in a three-layer radial configuration in the barrel (or central) region and a two-layer configuration in

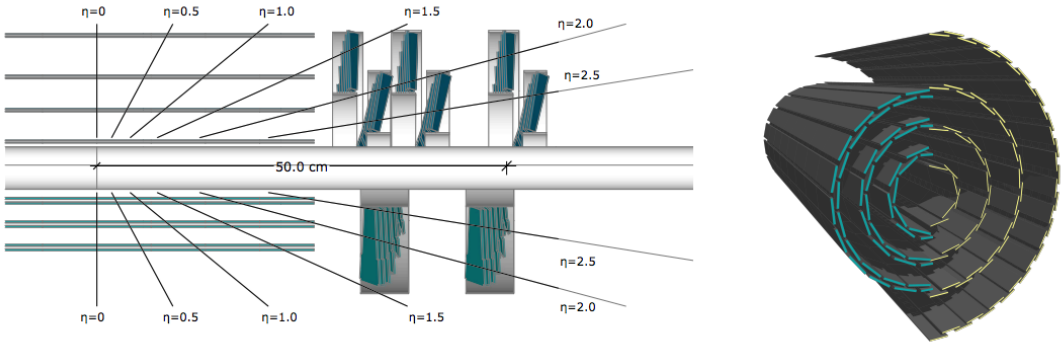


Figure 6.7: Comparison of the CMS pixel system in Phase 0 (below center) and Phase 1 (above center) [80]. The Phase 0 detector (up to 2016) featured only 3 barrel layers and 2 endcap disks, versus 4 barrel layers and 6 endcap disks in Phase 1 (2017 and 2018).

the endcap (or forward) region. The three barrel layers are placed at radii of 4.4, 7.3 and 10.2 cm and the endcap layers at z distances of 34.6 and 46.5 cm. The layers are designed such that a charged particle traversing the tracker has a minimum of three independent hit positions recorded.

The 2017 pixel tracker added a layer to the barrel region, arranging them at radii of 3, 6.8, 10.9 and 16 cm. It also added four layers to the endcap region, organized into three disks with two layers each, at z distances of 29.1, 39.6 and 51.6 cm. The total number of pixel channels increased to 124 million, spread over a 2 m^2 area. The updated tracker layout ensures a minimum of four independent hits per track, increasing the efficiency of track reconstruction and improving momentum and impact parameter resolutions. The forward acceptance of this tracker is given by $|\eta| < 2.5$.

The silicon sensors are themselves composed of several stacked layers. The silicon bulk with a thickness of 285 µm sits at the bottom, interfaced with a readout chip via small bump bonds to create an electrical connection between each pixel and its corresponding readout channel. On top is a flexible wire carrying the signals to the frontend electronics. The silicon bulk is an "n-in-n" material featuring a high-resistance n-substrate doped

with high dose n-implants. The transverse dimensions of each pixel are $100 \times 150\,\mu m^2$ and there are 4,160 pixels per sensor and 16 sensors per module. Fig. 6.8 shows a sketch of the sensor layout.

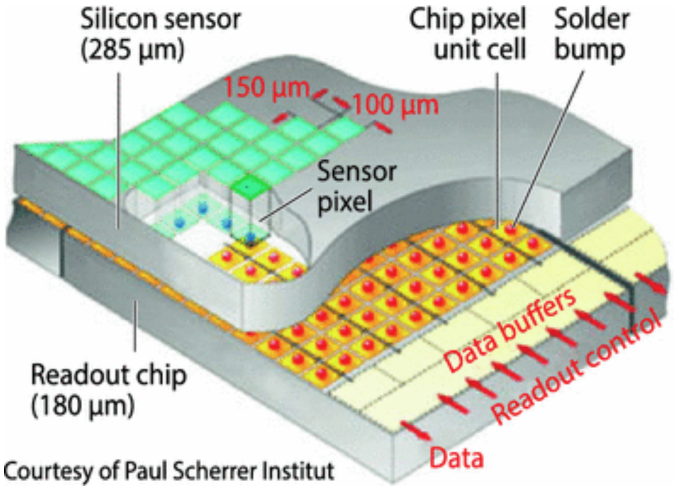


Figure 6.8: Sketch of a typical CMS pixel sensor [81]. Each sensor is composed of a silicon bulk where charge ionization happens, a readout chip bump-bonded to the silicon for charge collection, and a flexible wire to send the signal out to the frontend electronics (not shown).

Strip tracker

The strip silicon tracker surrounds the pixel tracker [79]. With a total silicon area of $198\,m^2$ distributed across 24,244 sensors, this is the largest silicon tracker ever built. While the pixel's goal is to record a limited but accurate subset of three-dimensional information about a particle's trajectory near the interaction point, the strip tracker has the complementary role of recording coarser two-dimensional information about the trajectory across a larger number of measurement surfaces. This allows the strip tracker to obtain a more robust lever arm and hence better measurement of the particle's momentum. The pixel's specialty, in contrast, is the determination of a particle's longitudinal and transverse impact parameter.

The strip tracker is composed of several sub-modules: the tracker inner barrel (TIB), tracker inner disks (TIDs), tracker outer barrel (TOB), and tracker endcaps (TECs). Fig. 6.6 highlights the layout of the strips. The TIB and TIDs immediately surround the pixel tracker. TIB is made of layers of silicon strips oriented parallel to the beam axis, and TIDs feature similar but normally oriented disks of strips. The 1D pitches vary across the different sub-detectors. For the innermost TIB layers 1 and 2, the pitch is 80 μm, and for the outermost layers 3 and 4, 120 μm. The TID disk pitches, in turn, vary from 100 μm to 141 μm. The thickness of the silicon bulk is 320 μm for both sub-detectors.

The TOBs and TECs are the outermost tracker surfaces. TOBs consist of strips with pitch 183 μm (inner 4 layers) and 122 μm (outer 2 layers). The silicon bulk thickness in this case is 500 μm. Finally, the TEC covers the endcap region up to high $|z|$ (124 cm to 282 cm). The strips are oriented radially in rings inside disks—-there are nine disks with up to seven rings per disk. The pitches vary across disks from 97 μm to 184 μm, and the silicon bulk thicknesses are 320 μm in the innermost 4 rings and 500 μm in the outermost 3 rings.

To provide a 2D measurement, the strips are arranged in a "stereo" layout. Pairs of strips are placed at stereo angles of 100 mrad relative to each other. This provides both azimuthal information (due to the one-dimensional strip segmentation) and z (for TIB and TOB) or r (for TID) information depending on the overall orientation of each sub-detector. For the TEC, sensors are oriented radially and rings 1, 2, and 5 have rotated sensors that provide the complementary r information instead.

The strip tracker system is able to record 9 hits on average for a particle with $|\eta| < 2.4$. The local hit resolution is between 23 and 53 μm in the azimuthal direction and 230 and 530 μm in the r or z directions. The ϕ resolutions are better since they only depend on the intrinsic sensor segmentation while the other coordinates rely on the stereo extrapolation.

6.2.4 Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) is a homogenous hermetic calorimeter located outside the silicon trackers [82]. It consists of 61,200 lead tungstate ($PbWO_4$) crystals in the central barrel region ($|\eta| < 1.479$) and a further 14,648 crystals in the two endcap components ($1.479 < |\eta| < 3$). Its main goal is to collect and measure the energy deposited by electromagnetic-interacting particles such as electrons, positrons, and photons.

The hermeticity is achieved by arranging the crystals in a quasi-projective orientation with a 3° angle offset in both η and ϕ directions in the ECAL barrel (EB) and a varying 2° to 8° in the ECAL endcap (EE) which prevents any gaps in the coverage of outgoing particles. The crystals have a tapered format with front- and rear-face areas of $2.2 \times 2.2\,\mathrm{cm}^2$ and $2.6 \times 2.6\,\mathrm{cm}^2$ respectively in the EB and $2.86 \times 2.86\,\mathrm{cm}^2$ and $3.0 \times 3.0\,\mathrm{cm}^2$ in the EE. The crystal lengths are 23 cm EB and 22 cm EE, corresponding to 25.8 and 24.7 radiation lengths. The resulting granularity is roughly 0.02 in both $\Delta\eta$ and $\Delta\phi$. A layout of the ECAL is shown in Fig. 6.9.

The choice of $PbWO_4$ is motivated by its high density ($8.28\,\mathrm{g/cm}^3$), short radiation length (0.89 cm) and Moliére radius (2.2 cm), fast response (80% of the scintillation light is emitted in the first 25 ns after a particle's arrival, which is the LHC bunch spacing), and radiation hardness. This means electromagnetic showers are relatively well-contained, improving the energy resolution. The low and temperature-dependent light yield of roughly 4.5 photoelectrons/MeV at 18 °C, on the other hand, dictates that the gain of photodetectors connected to the crystals must be kept stable and robust. This is accomplished by a laser calibration light system that regularly monitors the gain of the photodetectors. The photodetectors themselves are avalanche photodiodes in the EB (two per crystal) and vacuum phototriodes in the EE (one per crystal). These are glued

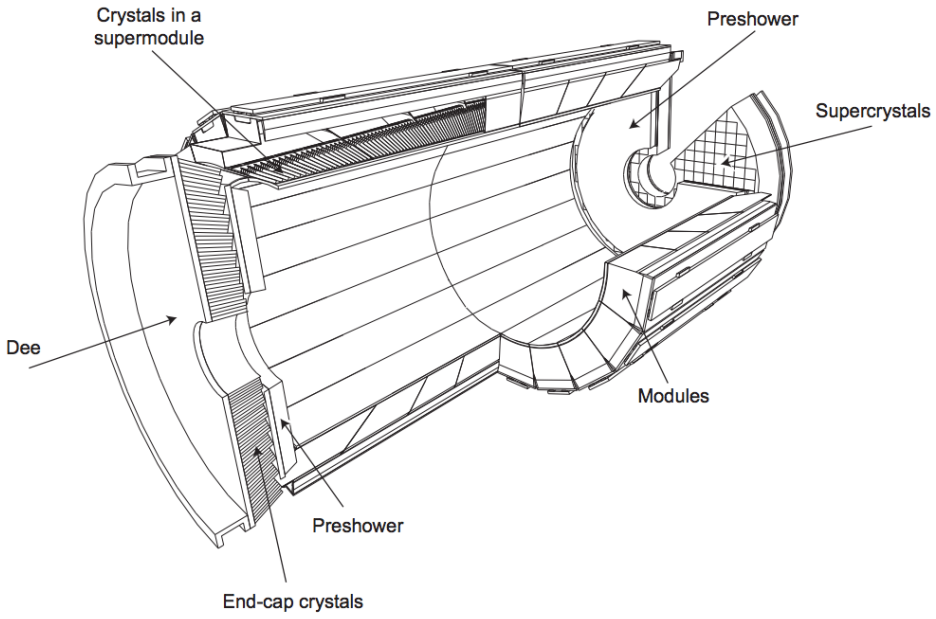


Figure 6.9: Layout of CMS's electromagnetic calorimeter [69]. The barrel, pre-shower, and endcap components together cover a pseudorapidity range of 0 to 4.

to the back side of the crystals.

An additional component of the ECAL is the ECAL pre-shower (ES) detector, which improves the ECAL ability to identify neutral particles such as photons and neutral pions. The ES is a sampling calorimeter with 20 cm thickness and two layers of lead radiators and silicon strips interspersed with one another. This layout initiates the electromagnetic shower before neutral particles reach the EE, thereby improving the total light collection efficiency and therefore the ECAL performance. The ES silicon strips have 1.9 µm pitch and a 61×61 mm^2 active surface.

The energy resolution of the ECAL for particles with energy lower than 500 GeV is given by:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{12\%}{E}\right)^2 + (0.30\%)^2, \tag{6.4}$$

where the first term is the stochastic contribution, the second is a noise term, and the last is a constant term. The stochastic term is due to statistics fluctuations and pre-shower energy containment. The noise term is due primarily to electronics and digitization noise. Finally, the constant term is a manifestation of incomplete and non-uniform longitudinal light collection. For particles with energy greater than about 500 GeV, this lack of containment becomes more severe and the energy resolution parameterization begins to break down. However, only rarely do particles in CMS reach such high energies at the ECAL, so this is not an issue for the experiment.

6.2.5 Hadronic calorimeter

The hadronic calorimeter (HCAL) is a detector situated outside the ECAL, responsible for measuring the energy deposition of hadronic showers [83]. These showers are initiated by charged and neutral hadronic particles such as kaons, pions, and other baryons, as well as heavy quarks. The HCAL is especially important in measurements of jet activity, since jet components are predominantly hadronic processes.

The detector is a sampling calorimeter with the overall design concept consisting of layers of absorber (brass) interspersed with layers of scintillators. The hadronic shower is triggered by interactions between incident particles and the brass layers and the shower particles produce light when interacting with the scintillator material. The amount of light is proportional to the energy of the shower.

Fig. 6.10 illustrates the layout of the HCAL. It is composed of barrel and endcap portions (HCAL barrel (HB) and HCAL endcap (HE)) similarly to the ECAL, as well as a forward detector (HCAL forward (HF)) and a single thick scintillator layer outside the HE (HCAL outer (HO)). The HB covers a pseudo-rapidity range of $0 < |\eta| < 1.3$, and

the HE of $1.3 < |\eta| < 3.0$. The unit blocks for both consist of towers encompassing 18 brass layers and 15 (17) plastic scintillator layers 3.7 mm thick in the HB (HE). There is additionally a 9 mm thick scintillator layer in front of and behind each HB tower and in front of each HE tower. The segmentation in η-ϕ space is about $(0.087, 0.087)$ for the HB and $(0.17, 0.17)$ for the HE.

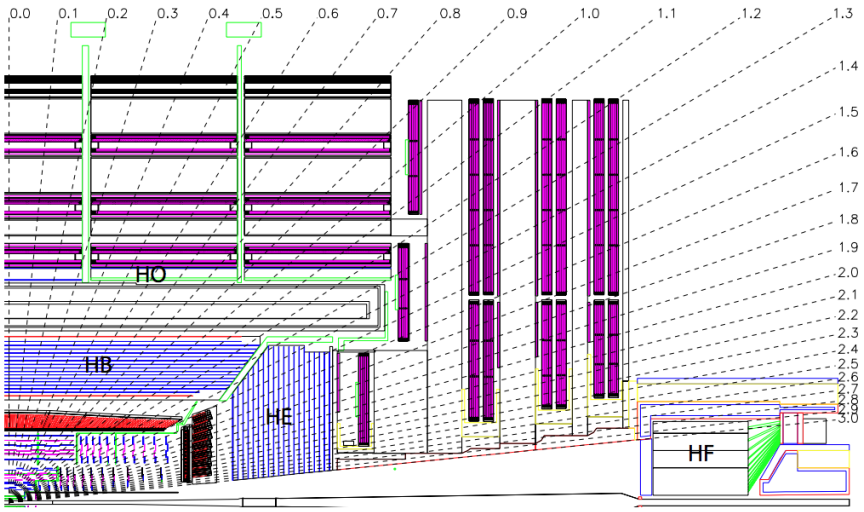


Figure 6.10: Layout of CMS's hadronic calorimeter [69]. All components are shown in this r-z projection. The coverage of the hadronic calorimeter is roughly $|\eta| < 4$.

Finally, the HO and HF provide complementary coverage to the main detectors. The HF extends the range of the HE to $|\eta| \approx 5.0$ and consists of a cylindrical steel absorber with radiation-hard quartz scintillating fibers interspersed in 5 mm intervals. The HO is a single layer of plastic scintillator 10 mm thick sitting outside the first layer of the magnet iron yoke. It ensures any energy leaks from the main HB are accounted for. Overall, the thickness of the HCAL is 5.39 radiation lengths near $|\eta| \approx 0$ and 10.3 near $|\eta| \approx 1.3$.

6.2.6 Muon system

The muon system is the outermost set of detectors in CMS [84]. As mentioned, CMS is especially well-suited to measure muons, which at high energy are minimum-ionizing particles (MIPs). Unlike other charged and most neutral particles, muons are typically not stopped by the inner layers of the detector such as calorimeters and magnet iron return yokes. Therefore the muon system works in a low-background regime. A further benefit is that the magnetic field outside the solenoid is inverted, so muons originating from the interaction point are bent in two directions in their journey outward. This allows for a more precise estimate of muon momentum.

The muon system consists of three main detectors: the drift tubes (DTs), cathode strip chambers (CSCs), and resistive plate chambers (RPCs). These are made with different technologies and generally provide complementary η coverage. Fig. 6.11 illustrates the layout of the muon system in CMS.

The DT chambers cover the barrel region of the detector with $|\eta| < 1.2$. They consist of 4 layers (or stations) oriented parallel to the beam axis at increasing r distances. These stations are interspersed with the flux return yokes of the magnet. The detector works with drift tube technology that relies on the ionization by incident muons of gas inside aluminum tubes (with 4 cm width). The tubes contain a central wire and are subject to an electric potential that sweeps the ionized charges. The negative charges (electrons) eventually drift to the wires where they are collected to construct a signal. Fig. 6.12 (left) depicts the schematic operation of the DT.

The CSC chambers cover the endcap region with $0.9 < |\eta| < 2.4$ and are also composed of 4 interspersed layers, but they are oriented perpendicularly to the beam. Each CSC is made of positively charged wires and negatively charged cathode strips,

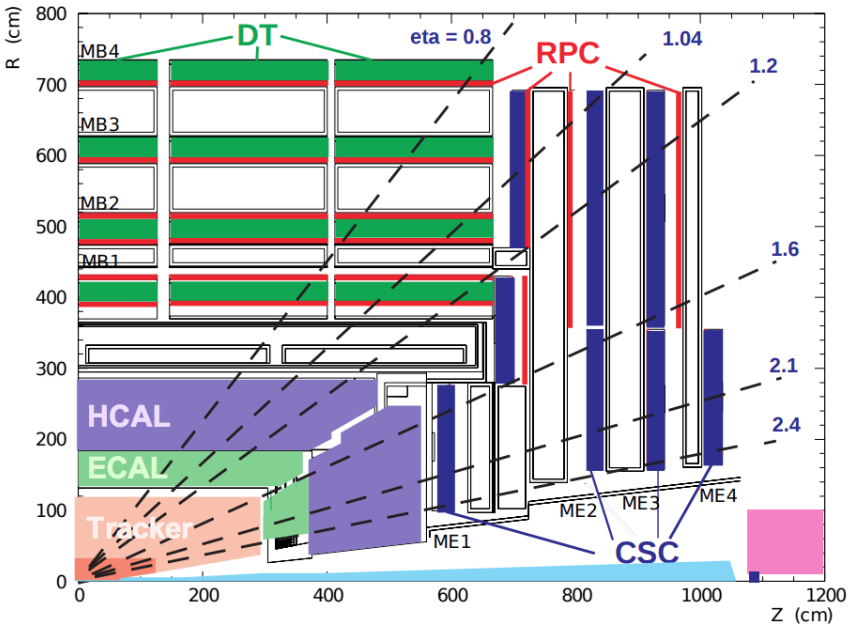


Figure 6.11: Layout of CMS's muon system [85]. The three sub-detectors (DT, RPC, and CSC) are interspersed with the solenoid magnet's iron return yokes. The muon system has a pseudorapidity coverage of $|\eta| < 2.4$.

also oriented perpendicularly to each other. This orientation provides the CSC with two-dimensional hit position measurement in the transverse plane. The technology is similar to DTs where incident muons ionize the gas inside the detector and electrons flow to a central wire while positive ions flow to the cathode strip. Fig. 6.12 (bottom) shows the CSC technology.

Finally, the RPC chambers are interspersed with both DT and RPC layers, providing excellent timing measurements of muons that are used in the trigger system. There are six RPC layers that provide a coverage of $|\eta| < 1.6$. The layers are oriented parallel (perpendicular) to the beam axis in the barrel (endcap) region. Each layer consists of two parallel plates with opposite charge located a fixed distance apart. Gas sits in-between the plates and the ionization creates charges that flow to the oppositely-charged plate. The drift in this case is short since the spacing between the plates is small. Fast detectors then measure the amount of ionized charge and a hit pattern algorithm quickly reconstructs the

likely momentum of the muon and feeds this information to the trigger system. Fig. 6.12 (right) illustrates the concept of the RPC.

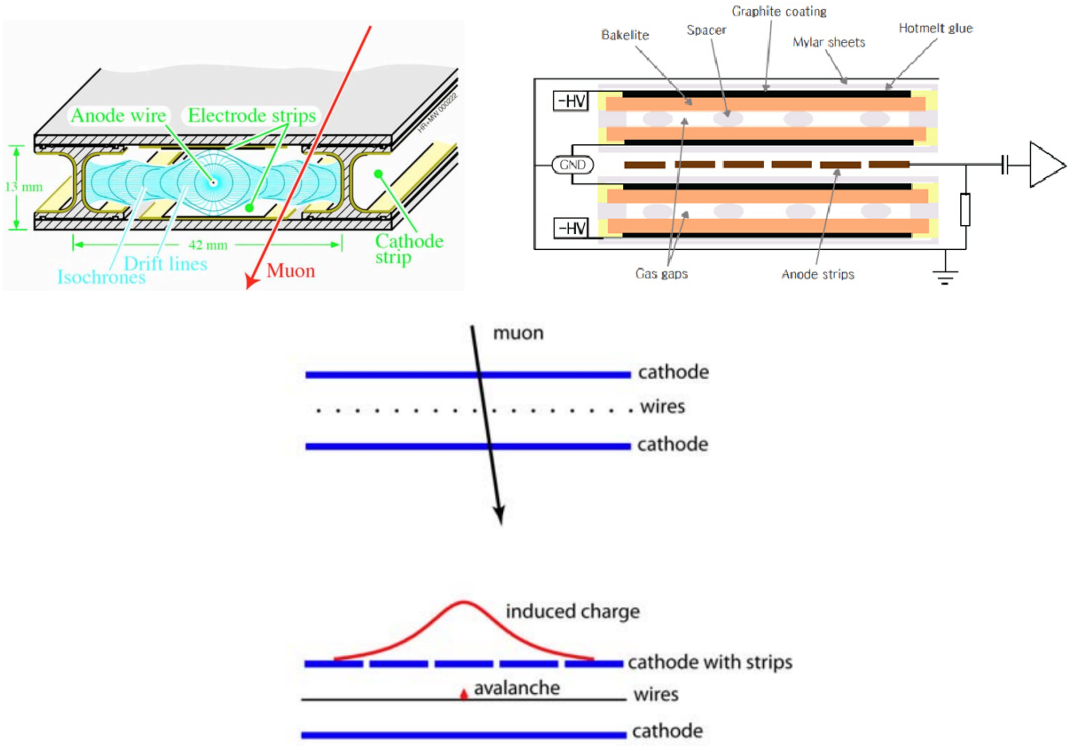


Figure 6.12: Comparison of different muon detector technologies employed at CMS [69, 86]. DT is shown on the upper left, RPC on the upper right, and CSC on the lower center.

6.2.7 Trigger and DAQ systems

The LHC collides proton bunches 40 million times a second. With current detector and network technology, it is impossible to collect and store all these data. Most of the pp collisions at the LHC, however, are uninteresting from a physics perspective since they involve only a glancing interaction between protons. Hard scattering between quarks and gluons in protons happens much more infrequently. CMS has therefore developed a sophisticated trigger system to select and store only the relevant events from the large pool

of pp collisions at the LHC [87]. This system consists of two layers, one hardware-based and one software-based.

Data acquisition system

The data acquisition system (DAQ) of CMS is highly complex owing to the integration of several sub-detectors and the processing of roughly 55 million readout channels overall. In broad terms, the data are collected continuously and stored locally in circular buffers on the electronic frontends of each sub-detector. These buffers hold the data until hardware trigger accept decision is received, thereby triggering the readout of all CMS channels by the DAQ system for high-level trigger processing. If the hardware trigger system rejects the event, the readout is not triggered and the data in the circular buffer is overwritten in the next bunch crossing. Thus, if an event fails the the lower-level trigger, it is permanently discarded. The DAQ output rate is about 100 Gbps.

Level-1 trigger

The Level-1 (L1) trigger is a fast hardware trigger that makes a decision to accept or reject an event in less than 4 µs. This system uses coarse information from a subset of CMS detectors: ECAL, HCAL, and the muon chambers. The result of the L1 trigger is distributed to all CMS detectors and an accept decision triggers the readout of all 55 million channels as mentioned. The L1 trigger system reduces the event rate from 40 MHz to a more manageable 100 kHz.

Fig. 6.13 illustrates the structure of the L1 trigger system. Broadly, the energy information from calorimeters and the hit information from muon chambers are processed regionally first. The regional calorimeter trigger processes the energy deposit measure-

ments and constructs primitive photon/electron candidates as well as local clusters of energy. These objects are then passed to the global calorimeter trigger where clusters of energy are grouped into jets and photons and electrons are distinguished. Identified objects are passed to the global trigger where a final trigger decision will be made once muon information arrives.

In parallel, muon chamber hits (DT and CSC) are subject to several algorithms to reconstruct a coarse representation of the muon path. Pairs of hits are combined into straight segments, which serve as seeds to a track-finding algorithm. RPC hits go instead to a pattern comparator where muon candidates are directly identified. Results from all muon detectors are collected into the global muon trigger where duplicates are removed. The final muon candidate list is then transferred to the global trigger. Finally, the global trigger has all the information to make a global accept/reject decision on the event.

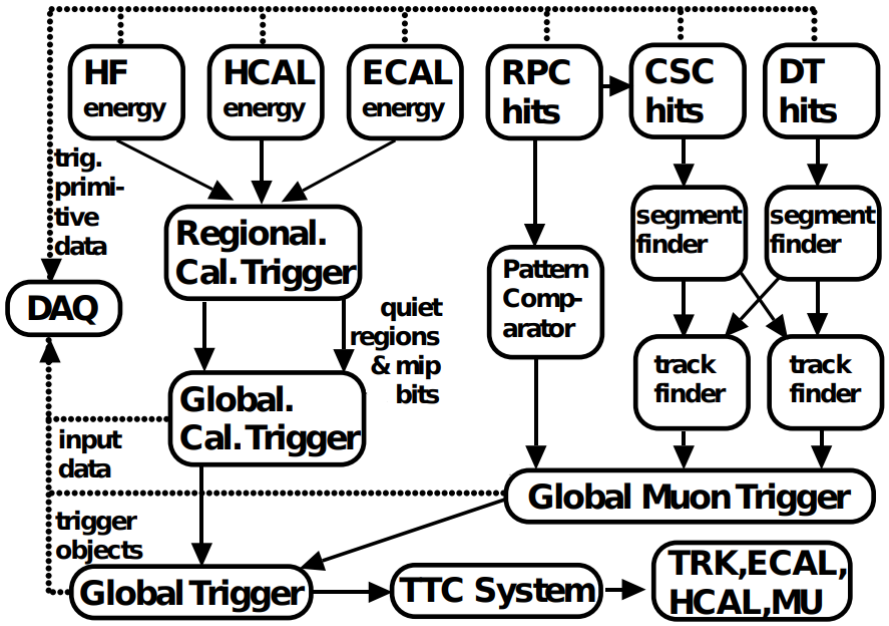


Figure 6.13: CMS's L1 trigger system [87]. The main data processed by the lower-level trigger are calorimeter and muon information only since it must be very fast.

High-level trigger

The high-level trigger (HLT) is a software-based system consisting of a 13,000-core computer farm. It reconstructs physics objects starting from the raw detector data with similar algorithms used later in the offline reconstruction and can make more informed decisions about the relevance of an event. The latency is higher than the L1 system at around 175 ms per event. The HLT has a menu of trigger paths that is configurable at runtime. A trigger path is a set of filters on several physics objects that the event must satisfy. If an event passes at least one full trigger path, it is stored permanently for offline reconstruction. If any filter fails during processing of a trigger path, that path is immediately stopped to save computer resources. The HLT further reduces the total event rate from 100 kHz to 400–1,000 Hz, which is more manageable for storage.

CHAPTER 7

DATA SETS AND TRIGGERS

This chapter describes the collected datasets and triggers applied in the analysis.

7.1 Data sets

The total amount of data collected between between 2015 and 2018 corresponds to 150.26 fb^{-1} at a center-of-mass energy of 13 TeV, 137.19 fb^{-1} of which are used in the analysis. The discrepancy is due to inefficiencies resulting from malfunctioning sub-detectors at different points in time. Events for which one or more sub-detectors report a problem are discarded to ensure a high-quality data set.

Each hard scattering between protons is in general accompanied by several other softer pp interactions in the same bunch crossing. This is dubbed the pileup and the number of reconstructed vertices per event is an indication of the pileup. These softer interactions typically do not contain interesting physics but must be accounted for when reconstructing an event since they produce tracks and calorimeter energy deposits. The pileup varies by year depending on the instantaneous luminosity delivered by the LHC. Fig. 7.1 reports the observed pileup distribution in data for the three years considered in the analysis.

The pileup must also be included in the simulation of physics events to match with the data. However, the actual observed pileup distribution is *a priori* unknown, so the simulated pileup cannot exactly match that of the data. A re-weighting procedure is therefore implemented which consists of dividing the simulated pileup by the observed pileup and deriving a weight from this ratio. Each simulated event is then re-weighted based on the number of vertices reconstructed in that event by using the ratio as a "lookup

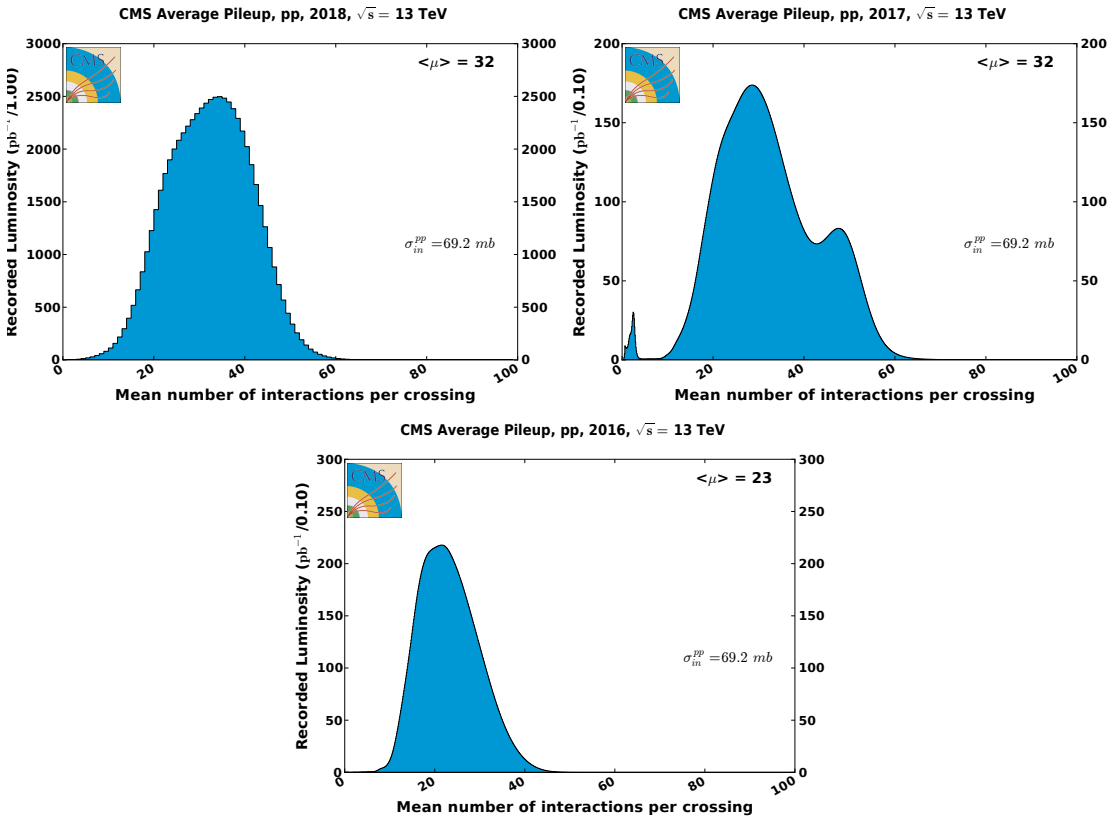


Figure 7.1: True pileup distributions observed in data in 2018 (upper left), 2017 (upper right), and 2016 (lower center) [**?**]. The smaller peak at high pileup in 2017 reflects a roughly 2-month window in October and November when the was running at peak instantaneous luminosity near 1.5×10^{34} cm^{-2} s^{-1}, almost 50% higher than the design luminosity.

table". Fig. 7.2 shows the derived pileup weights used in all three years. The simulated sample employed in the derivation is QCD since it has the largest production cross section at the LHC (at $\sqrt{s}$ = 13 TeV).

The data collected by CMS are split into several overlapping "primary data sets". Primary data sets consist of events passing one or more of a group of related triggers. It is possible for events to fire triggers in more than one group, such as $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers and muon triggers, ending up in both primary data sets. Since we rely on $p_{\mathrm{T}}^{\mathrm{miss}}$-based triggers, the missing transverse momentum (MET) primary data set is used in this analysis. The single-muon primary data set is also used to study the performance of the main $p_{\mathrm{T}}^{\mathrm{miss}}$

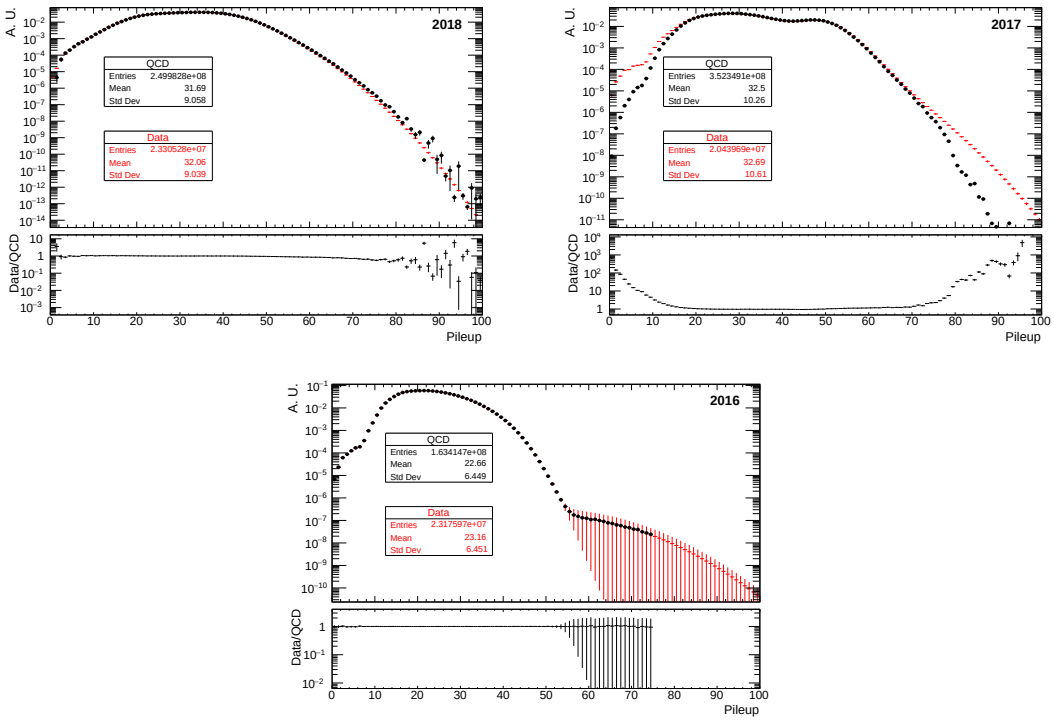


Figure 7.2: True pileup distributions measured in QCD simulation (black) and data (red) in 2018 (upper left), 2017 (upper right), and 2016 (lower center). The correction weights applied to simulated samples are derived from the ratios in the bottom pads.

triggers, as discussed in Section 7.2.

Simulated events were generated with Monte Carlo (MC) techniques for both signal and background processes. Samples were first generated at parton-level to leading order with MadGraph5_aMC@NLO [67] v.2.6.0, using a modified version of the Hidden Abelian Higgs Model (HAHM) [17, 88, 89]. Subsequently, hadronic showers and fragmentation were simulated from the parton objects generated by MadGraph (MG) with Pythia v8.230 [90]. Events produced this way are then mixed with a separate sample of minimum-bias events that contains the pileup interactions described above. This minimum-bias data set is pre-produced and each event in the dataset contains only the pileup interactions without any hard scattering. Events are randomly sampled from this data set and mixed in with the main samples. Finally, the complete events are run through

a detailed simulation of the CMS detector created with the GEANT4 toolkit [91]. This simulation reproduces the material interactions between detector and particles. During this stage, the algorithms used to reconstruct particles from calorimeter energy deposits and tracker hits in data are also applied to simulated events, in an identical way, to emulate the detector response and the overall performance of the apparatus.

The simulated signal samples are listed in Table 7.1 which reports the cross sections computed with MG, taking into consideration the aspects discussed in Chapter 5 involving the parameters of the inelastic dark matter model. The cross sections are calculated for two α_D assumptions per signal point: $\alpha_D = 0.1$ and $\alpha_D = \alpha_{EM}$.

Table 7.1: Inelastic dark matter simulated signal samples. The 32 representative samples used throughout event selection and optimization studies are highlighted in boldface. Cross sections are suppressed with increasing lifetimes as discussed in Section 5.2. To obtain the cross section for a particular lifetime, divide the given cross section by $c\tau$ in mm (e.g. for $c\tau = 100$ mm, divide by 100).

$(m_1, m_{A'})$ [GeV]	$c\tau$ [mm]	Events	$\sigma \times (c\tau/\text{ mm})$ [pb]			
			$\alpha_D = 0.1$		$\alpha_D = \alpha_{EM}$	
			$\Delta = 0.1\, m_1$	$\Delta = 0.4\, m_1$	$\Delta = 0.1\, m_1$	$\Delta = 0.4\, m_1$
1, 3	1, 10, 100, 1000	500k	-	0.045	-	0.500
3, 9	1, 10, 100, 1000	500k	15.8	0.015	167	0.163
5, 15	**1, 10, 100, 1000**	**500k**	9.56	**0.009**	**100**	**0.100**
10, 30	1, 10, 100, 1000	500k	4.84	0.005	50	0.048
20, 60	1, 10, 100, 1000	500k	3.61	0.004	30	0.030
30, 90	1, 10, 100, 1000	500k	150	0.150	400	4.6
40, 120	1, 10, 100, 1000	250k	0.95	0.0007	9.8	0.010
50, 150	**1, 10, 100, 1000**	**250k**	0.22	0.0002	**3.1**	**0.003**
60, 180	1, 10, 100, 1000	250k	0.12	0.0001	1.55	0.001
80, 240	1, 10, 100, 1000	250k	0.04	0.00005	0.60	0.0006

7.2 Triggers

The triggers used in the analysis are primarily METNoMu triggers. The METNoMu designation removes from $p_{\mathrm{T}}^{\mathrm{miss}}$ contributions due to muons. This ensures there is sig-

nificant $p_{\mathrm{T}}^{\mathrm{miss}}$ in the event produced by other physics objects. The full trigger name is `HLT_PFMETNoMuXX_PFMHTNoMuXX_IDTight`, where XX stands for the available METNoMu and MHTNoMu thresholds: 120, 130, or 140 GeV. All three are used for maximum efficiency. At the L1 level, the trigger requires enough energy in the event as measured either by the combination of calorimeter and muon information or by the calorimeter alone. At the HLT level, the trigger requires sufficient jet activity and also removes muon contributions from the $p_{\mathrm{T}}^{\mathrm{miss}}$ and assesses whether the resulting $p_{\mathrm{T}}^{\mathrm{miss}}$ crosses the threshold.

Trigger efficiencies were measured in data by means of a reference trigger, `HLT_IsoMu27`, which belongs to the single-muon primary dataset. The strategy is to use the semi-leptonic decay of the W boson, $W \rightarrow l\nu$ + jets, to measure the efficiency of the $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers. The outgoing lepton provides an unbiased selection of events with which to probe the $p_{\mathrm{T}}^{\mathrm{miss}}$. For this study, the event selection consisted of:

- Leading jet $p_{\mathrm{T}} > 80\,\mathrm{GeV}$, $|\eta| < 2.4$;
- Leading global muon $p_{\mathrm{T}} > 35\,\mathrm{GeV}$, $|\eta| < 2.4$;
- $|\mathrm{MET}_{PF} - \mathrm{MET}_{CALO}|/\mathrm{MET}_{CALO} < 0.5$.

This selection ensures that the muon in the event is energetic enough to sit on the efficiency plateau of the reference trigger, which could otherwise bias the pool of events. The leading jet p_{T} selection ensures enough jet activity such that relevant physics is being probed. Finally, the last requirement reduces the incidence of mis-reconstructed MET by requiring that the $p_{\mathrm{T}}^{\mathrm{miss}}$ computed from the particle-flow algorithm is not drastically different from that computed by calorimeters alone.

Fig. 7.3 shows the METNoMu trigger efficiency for all three years. Black dots correspond to data and red dots to simulated events. In addition to the main $W \rightarrow l\nu$

events, other simulated samples were added for completeness: QCD, multi-boson, and top quark physics. The turn-on curves are fit by a fit function of the form:

$$f(x) = \frac{\epsilon}{2}\left[\mathrm{erf}\left(\frac{x-\mu}{\sigma}\right) + 1\right], \tag{7.1}$$

where the error function erf is given by:

$$\mathrm{erf}(x') = \frac{2}{\sqrt{\pi}} \int_0^{x'} e^{-t^2}\, dt. \tag{7.2}$$

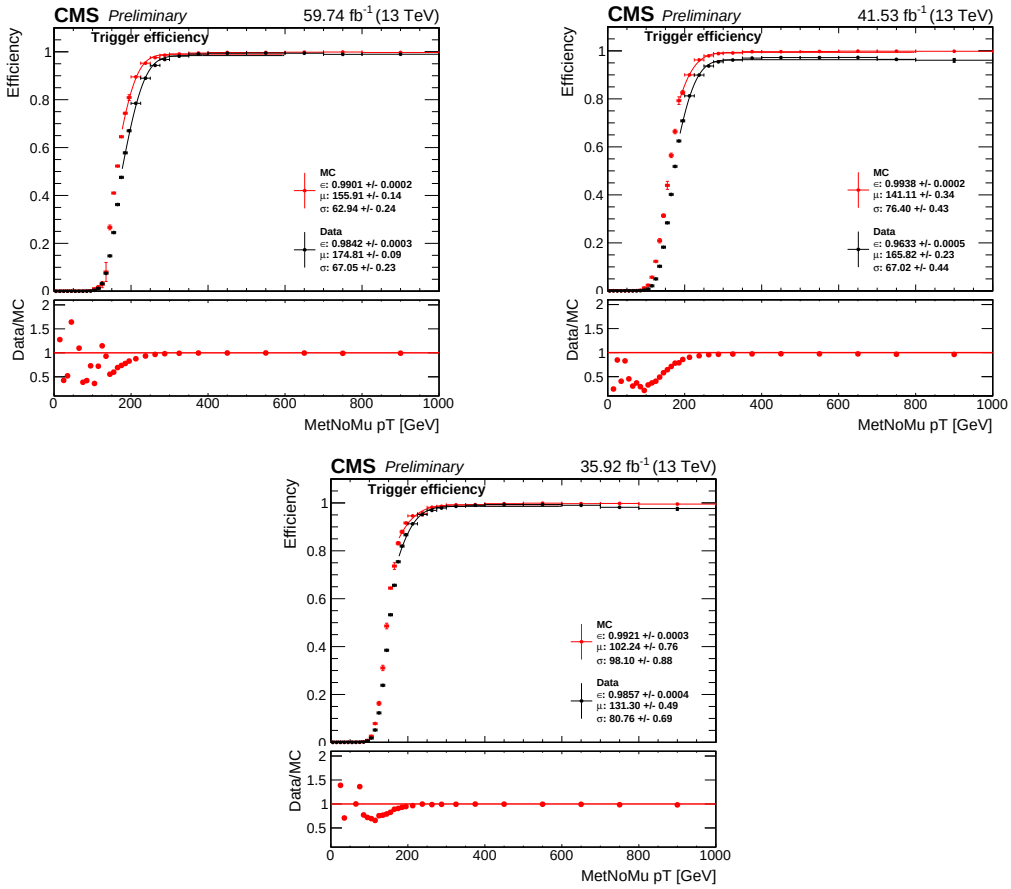


Figure 7.3: Trigger `HLT_PFMETNoMu120_PFMHTNoMu120_IDTight` efficiency *vs.* MET-NoMu p_{T} and accompanying error function fits in 2018 (upper left), 2017 (upper right) and 2016 (lower center).

The plateau of the efficiency is reached at slightly different values depending on the year, ranging from roughly 200 GeV to 300 GeV. We also measured the performance of the trigger in simulated signal events, shown in Fig. 7.4. Here no reference trigger is needed since these events are already preselected. Correspondingly, there is no muon selection and also no jet selection to increase the sample size. From the figure, the plateau of the efficiency curve is reached around 200 GeV for all years.

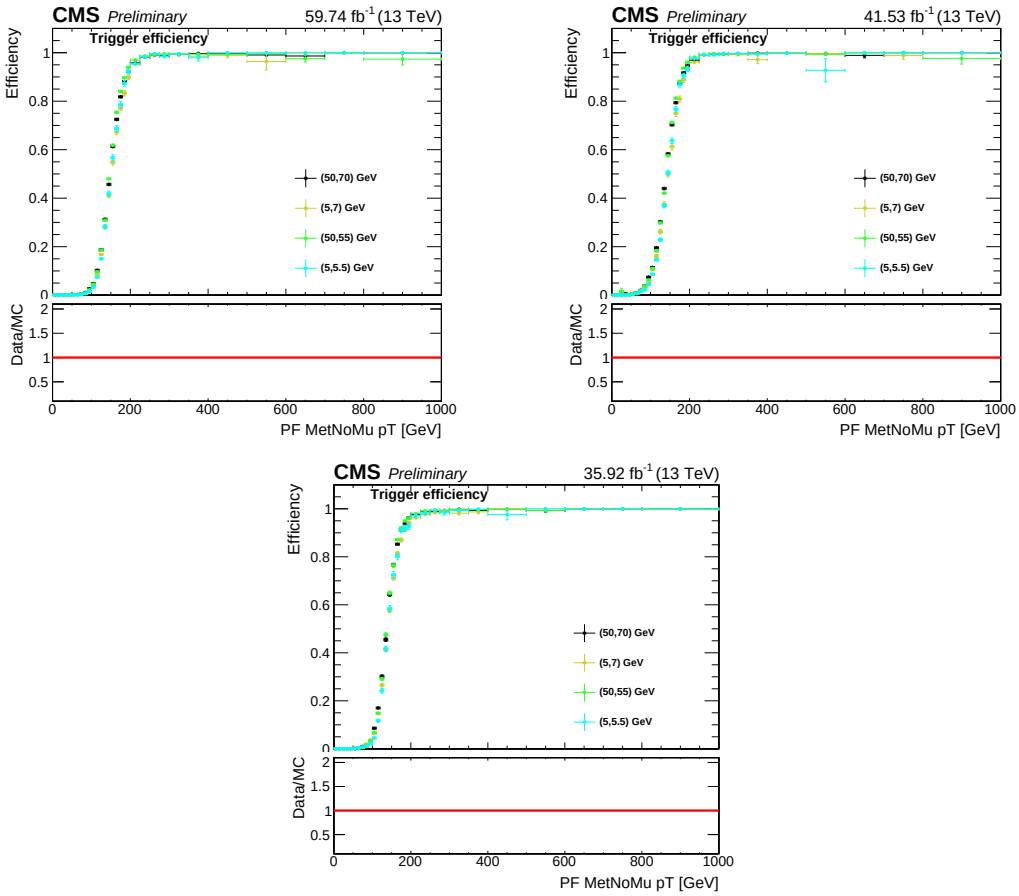


Figure 7.4: Trigger efficiency in signal for the 120 GeV threshold MetNoMu triggers as a function of MetNoMu p_{T} in 2018 (upper left), 2017 (upper right), and 2016 (lower center).

In Fig. 7.3, the efficiency observed in data does not exactly match the simulated efficiency. Therefore the ratio of the two curves, shown below each plot, is taken as an additional weight applied to simulated samples only. This corrects for the slightly

discrepant behavior of the trigger in simulation and in data in a similar manner to the pileup correction described in Section 7.1.

CHAPTER 8

PHYSICS OBJECTS

This chapter describes the reconstruction of physics objects in CMS. I focus on objects relevant to the inelastic dark matter search: muons (Section 8.1), vertices (Section 8.2), jets (Section 8.3), and $p_{\mathrm{T}}^{\mathrm{miss}}$ (Section 8.4). In particular, the reconstruction of muon tracks is extensively discussed, since a complex strategy is employed to ensure good performance for very displaced muons.

Fig. 8.1 illustrates the layout of the CMS detector in a cross sectional view. Physics objects are reconstructed with the particle-flow (PF) algorithm [92] which gathers information from all available sub-detectors to create a global picture of each object. PF is a heuristic algorithm that takes as input energy deposits in the calorimeters and tracks from the trackers and muon chambers; it then outputs a list of stable particles that traversed the detector after each pp collision. The list of reconstructed particles and associated measurements includes photons (using ECAL energy deposits), electrons (ECAL energy plus tracker tracks), neutrons (ECAL plus HCAL), charged hadrons (tracker plus HCAL), neutral hadrons (HCAL), and muons (tracker plus muon chambers).

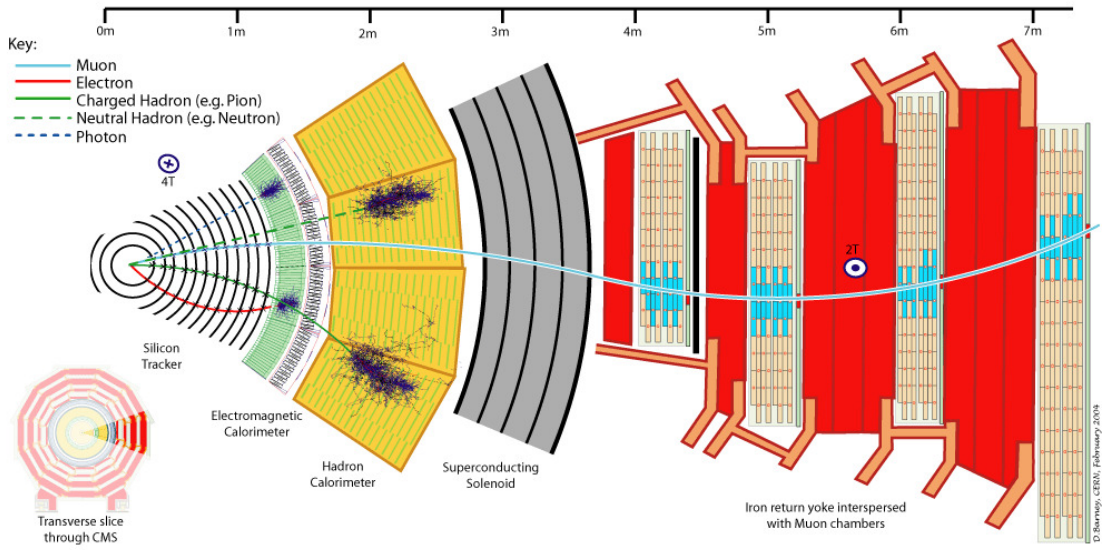


Figure 8.1: Transverse cross section sketch of the CMS detector [93].

8.1 Muons

Muons are crucial objects in the analysis. Muons from inelastic dark matter decays are expected to be considerably displaced from the interaction point, which helps greatly in reducing backgrounds. In particular, the reconstruction of muons relies on data collected from both trackers and muon chambers. I briefly discuss the reconstruction algorithms in each case before presenting the muon strategy used in the analysis.

8.1.1 Track reconstruction

Muon tracks can be constructed with tracker and/or muon chamber data. Naturally, better reconstruction performance is achieved when more information is available. However, tracker hits might not be produced if muons are created far away from the interaction point. For this reason, we study different reconstruction strategies to maximize performance even when muons are displaced. Here we summarize the relevant muon track

reconstruction algorithms.

Tracker tracks

The goal of reconstruction in the trackers is to produce estimates of the momentum vectors of charged particle tracks in each collision. Pileup makes this a challenging endeavor, since there are thousands of tracks per event. Nevertheless, with an iterative approach to track reconstruction it is possible to mitigate this issue to a large extent. A track trajectory is described by five parameters that correspond to the helix formed by charged particles traversing a uniform magnetic field. In the conventional parameterization, these quantities are p_{T}, η, ϕ_0 (angle at point of closest approach), z_0 (longitudinal position at point of closest approach), and d_0 (transverse position at point of closest approach). The reconstruction sequence has 6 major steps [94]: hit reconstruction, clustering, seeding, track finding, track fitting, and track selection. Each is briefly described next.

Charged particles leave hits in the pixel and strip trackers when they cross the silicon layers and ionize silicon atoms. In the pixel detector, hits are created when ionized charge in a pixel rises above a threshold of 3,200 electrons. Pixels where the collected charge does not meet this threshold are zero-suppressed, saving bandwidth and processing time. In the strips, the threshold is relative: a strip is read out if the charge measured exceeds five times its baseline noise or if the charge in two neighboring strips exceed two times the average baseline noise.

Hits are then clustered together. Any adjacent above-threshold pixels form a cluster and similarly for the strips except that cluster seeds must have a measured charge at least three times above the baseline noise. The position of clusters is determined by a charge-weighted average of each cluster hit. They are first calculated in each detector's

local coordinates and later translated to the global coordinate frame.

The next step is track finding. Clusters are assembled into track candidates via an iterative procedure. In each iteration, hits associated to tracks found in the previous iteration are removed from consideration. This decreases the combinatorial complexity of the problem. Track seeds are constructed from three clusters. A rough estimate of the track's trajectory is calculated from this seed and passed on to a Kalman Filter [95] where the full trajectory of the track and associated clusters are found. The Kalman Filter is an iterative algorithm that predicts the evolution of a track's trajectory and estimates the position of the next hit in the detector. The predicted position is compared to the measured hit (if any) and discrepancies are used to update the estimate of the trajectory. This algorithm is efficient and easily accommodates material effects and non-uniform magnetic fields since it operates layer-by-layer.

Finally, tracks found in the track-finding stage are subsequently fitted to estimate the optimal parameters of the trajectory. This fit is also based on the Kalman Filter but the final filtering and smoothing stage uses a Runge-Kutta propagator to extrapolate the trajectory for better performance [96]. Here, too, material effects and magnetic field non-uniformities are taken into account for a precise calculation of the deviation of a track from a perfect helix.

The procedure described is highly efficient in reconstructing real tracks, but it also generates spurious or misidentified tracks. Therefore, selection criteria must be applied to reconstructed tracks to minimize the presence of such tracks. Events with spurious tracks are one of the dominant backgrounds in the analysis, as discussed in Section 9.5.

Tracks reconstructed exclusively with tracker data are referred to as *tracker tracks*.

Standalone muon tracks

Muon chamber track reconstruction closely follows the architecture of the tracker reconstruction described above. A Kalman Filter is also used to find and fit tracks from initial seeds [97]. The hits consist of electric signals generated by the muon chambers (DT, CSC, and RPC) when a muon or other charged particle ionizes the gas inside the chambers. The position of each hit is determined from the electric signals with algorithms that depend on the type of chamber. The DT and CSC chambers are multi-layer detectors while the RPC has a single layer. For this reason, DT and CSC hits are combined into straight-line track "segments." These segments constitute the initial seeds that are fed to the track finding and fitting stage.

Tracks reconstructed exclusively with muon chamber data are referred to as *standalone muon tracks*.

Displaced standalone muon tracks

Standalone tracks, despite using only muon chamber hits, still operate with the assumption that muons are produced at or near the interaction point. The track fit procedure applies a constraint that the muon must be coming from the beam-spot region. This ensures a better fit performance for prompt muons but it also reduces reconstruction efficiency for displaced muons.

A separate reconstruction algorithm, *displaced standalone muon tracks*, removes the beam-spot constraint and additionally uses a modified cosmic muon seed in the track fit stage. The cosmic muon seed is not required to point down, which would be expected of a typical cosmic muon. The collection also benefits from a new muon timing strategy that relies on local information from DT hits to improve reconstruction for out-of-time

muons. These modifications ensure that displaced standalone tracks have a much higher efficiency than global muon tracks for very displaced muons, discussed in Section 8.1.2.

Global muon tracks

Tracker tracks and standalone muon tracks are combined to form *global muon tracks* that use all track information available to produce a better estimate of the muon trajectory. Global muons are built in the "outside-in" direction by matching standalone tracks with tracker tracks. The matching is done by extrapolating the two tracks onto a common surface and comparing the helical parameters there. If two tracks match, all of their hits are used in an additional Kalman Filter iteration to produce a new global muon track.

Global muon tracks provide the best performance for prompt muons. However, they still suffer from a low reconstruction efficiency for very displaced muons such as those expected in inelastic dark matter. This is a direct result of the tracker region ending at $z \approx 60\,\text{cm}$ from the nominal interaction point, after which point the reconstruction efficiency drops to zero. Therefore a strategy that can incorporate both displaced standalone muon tracks and global muon tracks is desired. Such a strategy is adopted in this analysis and described in Section 8.1.5.

8.1.2 Reconstruction efficiency

In this section, I present studies of the reconstruction efficiency of global muon tracks and displaced standalone tracks to argue for the inclusion of both collections in the analysis. Henceforth I refer to the latter collection as displaced standalone (dSA) and to the former as global muons (GM).

The studies are performed on representative simulated signal samples. They are split into barrel ($|\eta| < 1.2$) and endcap ($|\eta| > 1.2$) regions. Results are shown for 2018 conditions, though conclusions remain the same for other years. The efficiency is evaluated by matching a reconstructed object with its truth-level counterpart in dR (where $dR \equiv \sqrt{\Delta\phi^2 + \Delta\eta^2}$) and by requiring charge parity between them. The two collections have different dR thresholds for matching because of varying performance between them. Global muons typically have better resolution than dSA muons, so the dR matching criterion can be tightened. This is illustrated in Fig. 8.2, where the dR between reconstructed and generated object is plotted for both collections. The dR is evaluated at the interaction point since all lifetimes are combined for a given mass point. The requirements are set at $dR < 0.15$ for dSA and $dR < 0.06$ for GM.

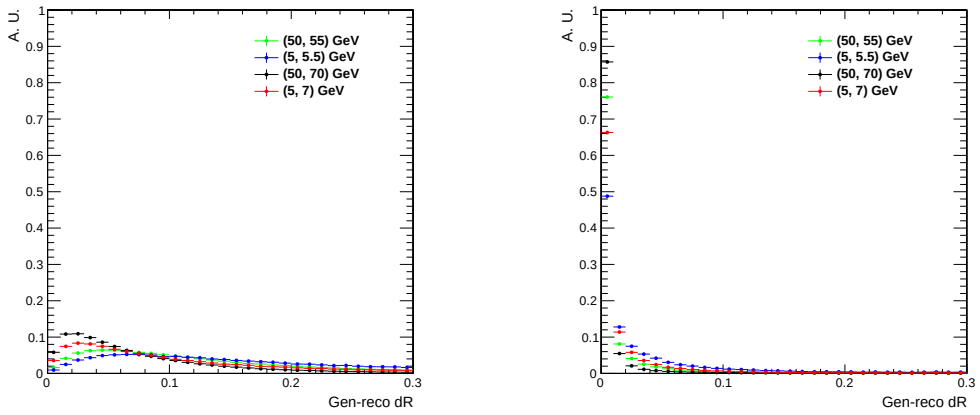


Figure 8.2: Comparison of dR between reconstructed and generated muon for both collections: dSA (left) and GM (right). The horizontal and vertical ranges are identical for easier comparison. The chosen thresholds for matching are 0.15 (dSA) and 0.06 (GM), respectively.

Displaced standalone muons

Figs. 8.3 and 8.4 report the dSA reconstruction efficiency in the barrel and endcap regions, respectively, as a function of several generator-level kinematic observables: p_{T}, η,

ϕ, and dR between the two signal muons. The curves are shown for four representative signal samples with varying dark matter masses and mass splittings. The overall efficiency is roughly dependent on the mass splitting between light and heavy dark matter, since this controls the energy available for decay muons. The efficiency also peaks at lower values of dimuon dR. For inelastic dark matter, the collimation between muons is still loose enough that any loss of efficiency at extremely small values of dR (where distinguishing between two tracks becomes more difficult) is negligible. For analyses with much tighter collimated muons, this could be a concern, motivating the use of specialized reconstruction objects such as lepton jets to recover some of the lost efficiency.

The slightly lower efficiency at higher muon momentum and the more significant loss of efficiency at higher dR between muons is due to a correlation between these two variables in signal. Higher muon momentum corresponds to a boosted decay in which the two muons are very tightly collimated, as seen in Fig. 8.5 where 2D histograms of gen-level dR *vs.* p_{T} are plotted. Conversely, a large dR separation limits the energy available to muons, so p_{T} distributions are quite soft resulting in a lower reconstruction efficiency.

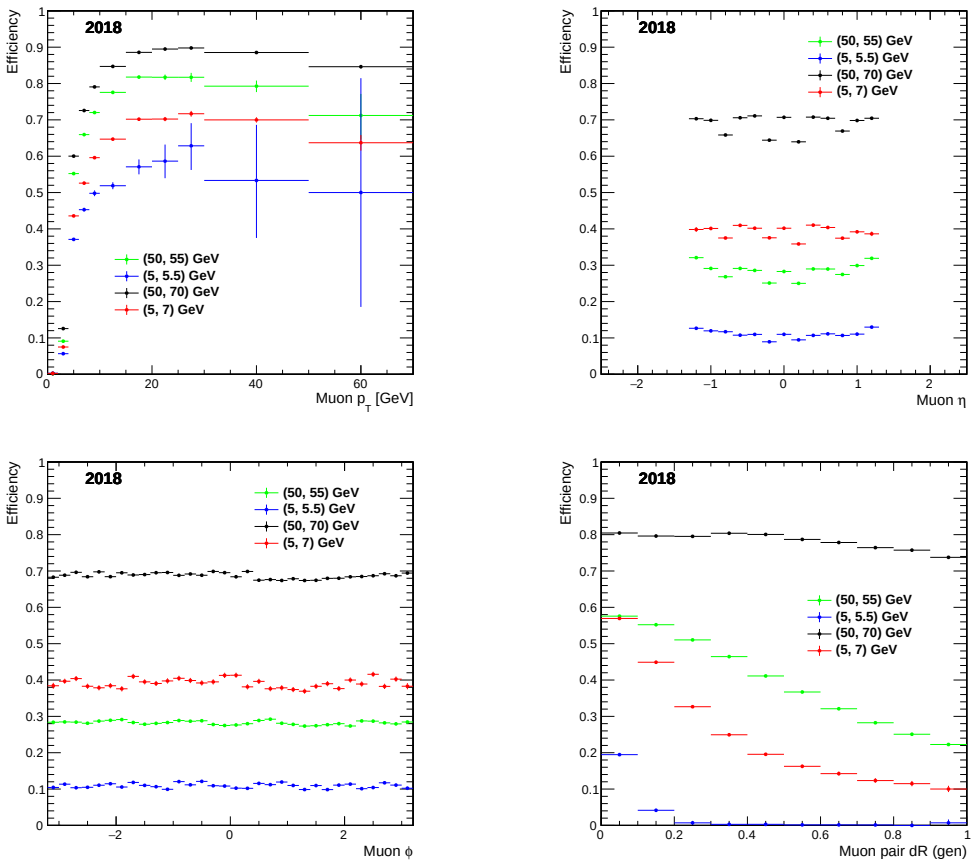


Figure 8.3: (Barrel, 2018) dSA reconstruction efficiency of *vs.* p_{T} (upper left), η (upper right), ϕ (lower left), and dimuon dR (lower right), for different signal samples: green $(m_1, m_2) = (50, 55)\,\mathrm{GeV}$, blue $(m_1, m_2) = (50, 70)\,\mathrm{GeV}$, black $(m_1, m_2) = (50, 70)\,\mathrm{GeV}$, and red $(m_1, m_2) = (5, 7)\,\mathrm{GeV}$. For each mass point, events with all lifetimes are added together.

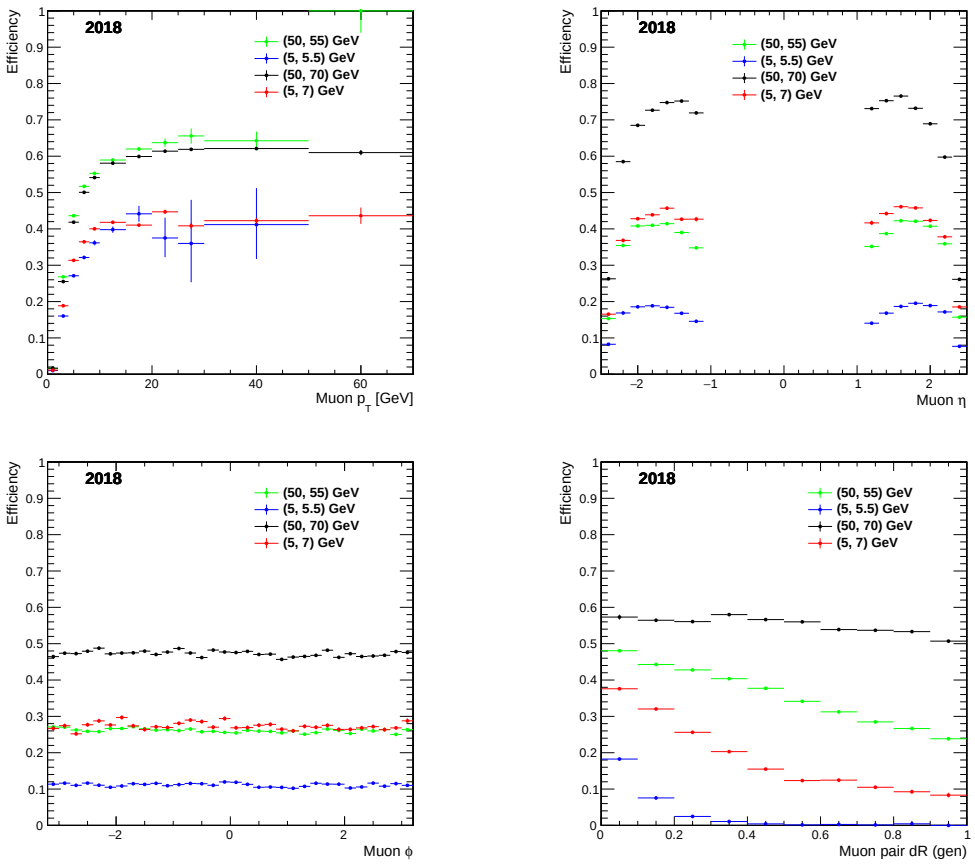


Figure 8.4: (Endcap, 2018) dSA reconstruction efficiency of *vs.* p_{T} (upper left), η (upper right), ϕ (lower left), and dimuon dR (lower right), for different signal samples: green $(m_1, m_2) = (50, 55)\,\mathrm{GeV}$, blue $(m_1, m_2) = (50, 70)\,\mathrm{GeV}$, black $(m_1, m_2) = (50, 70)\,\mathrm{GeV}$, and red $(m_1, m_2) = (5, 7)\,\mathrm{GeV}$. For each mass point, events with all lifetimes are added together.

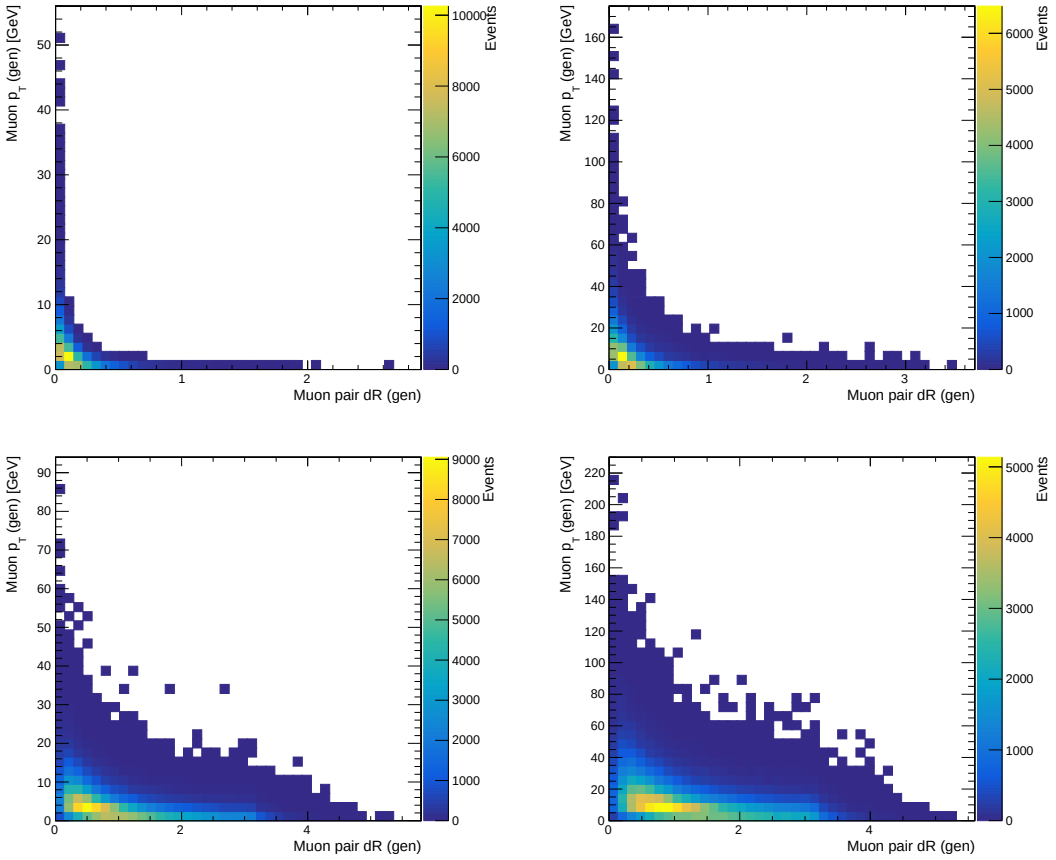


Figure 8.5: Correlation between generated muon p_{T} and dimuon dR for different signal samples: $(m_1, m_2) = (5, 5.5)\,\mathrm{GeV}$ (upper left), $(m_1, m_2) = (5, 7)\,\mathrm{GeV}$ (upper right), $(m_1, m_2) = (50, 55)\,\mathrm{GeV}$ (lower left), and $(m_1, m_2) = (50, 70)\,\mathrm{GeV}$ (lower right). Events with very high p_{T} have very tight dimuon collimation, and vice-versa.

Global muons

Figs. 8.6 and 8.7 show the same reconstruction efficiency curves for global muons. As expected, the overall efficiency of global muons is lower than dSA muons because of the displaced nature of muons in signal. This can be seen explicitly by comparing the reconstruction efficiency *vs.* muon displacement across both collections, discussed next.

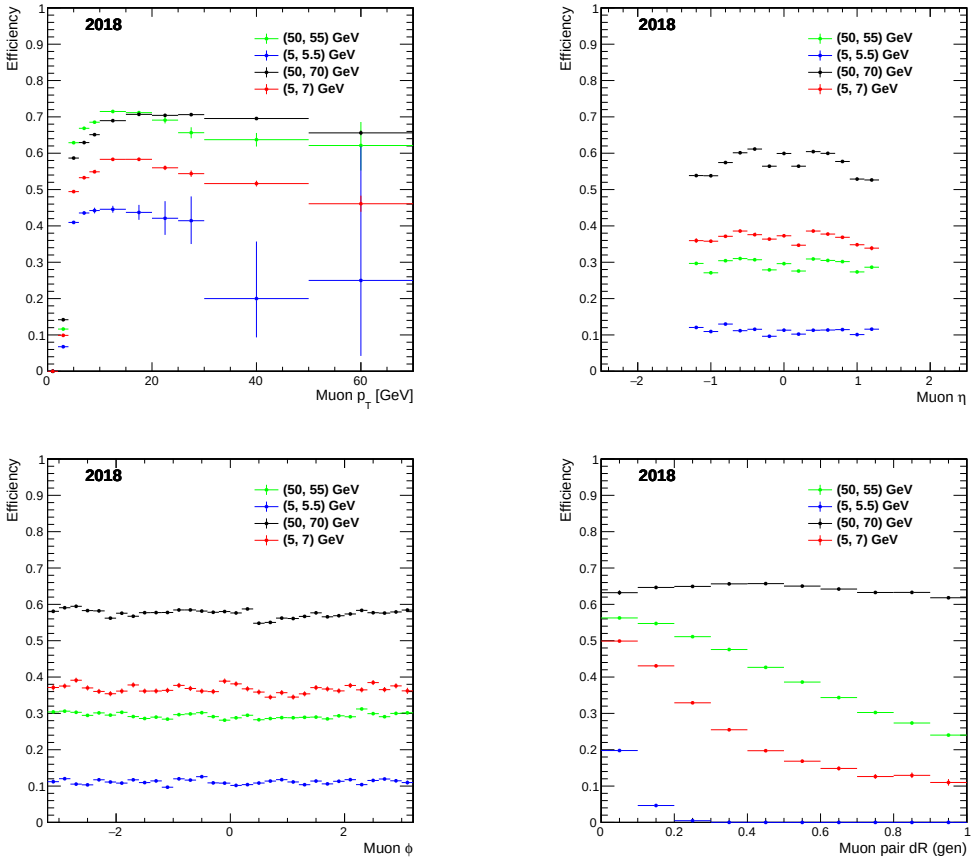


Figure 8.6: (Barrel, 2018) GM reconstruction efficiency of *vs.* p_{T} (upper left), η (upper right), ϕ (lower left), and dimuon dR (lower right), for different signal samples: green $(m_1, m_2) = (50, 55)\,\mathrm{GeV}$, blue $(m_1, m_2) = (50, 70)\,\mathrm{GeV}$, black $(m_1, m_2) = (50, 70)\,\mathrm{GeV}$, and red $(m_1, m_2) = (5, 7)\,\mathrm{GeV}$. For each mass point, events with all lifetimes are added together

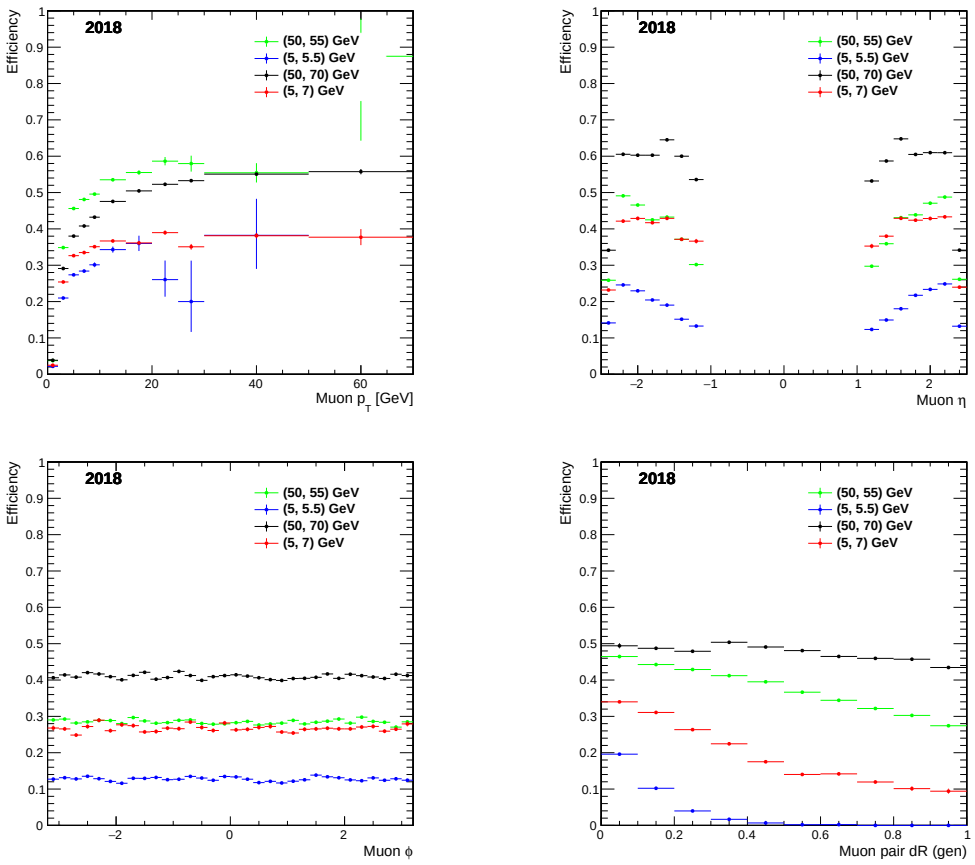


Figure 8.7: (Endcap, 2018) GM reconstruction efficiency of *vs.* p_{T} (upper left), η (upper right), ϕ (lower left), and dimuon dR (lower right), for different signal samples: green $(m_1, m_2) = (50, 55)\,\mathrm{GeV}$, blue $(m_1, m_2) = (50, 70)\,\mathrm{GeV}$, black $(m_1, m_2) = (50, 70)\,\mathrm{GeV}$, and red $(m_1, m_2) = (5, 7)\,\mathrm{GeV}$. For each mass point, events with all lifetimes are added together.

Comparison

Figs. 8.8 and 8.9 compare dSA and GM reconstruction as a function of muon displacement, first for barrel and then for endcap muons. GM efficiency drops sharply after the end of the tracker region ($z \approx 60$ cm from the nominal pp collision point), whereas dSA efficiency remains comparatively high throughout the displacement regime considered in the analysis.

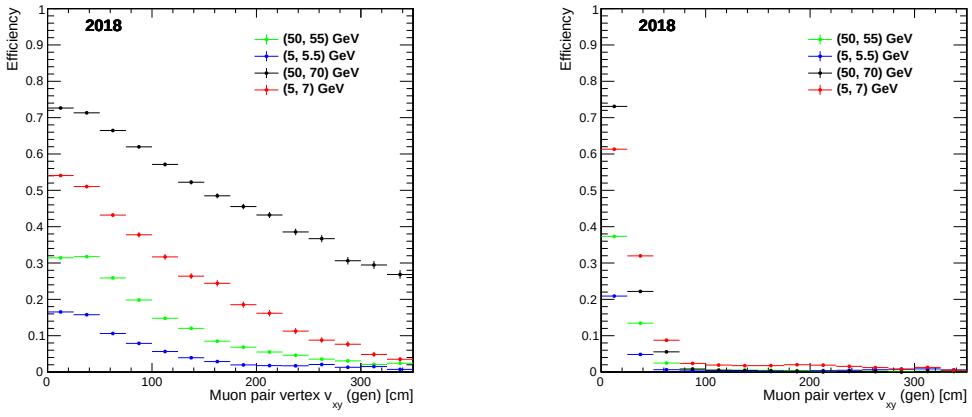


Figure 8.8: (Barrel, 2018) Reconstruction efficiency *vs.* muon true vertex v_{xy} of dSA (left) and GM (right) collections. The GM efficiency drops sharply after the end of the tracker region (≈ 60 cm) whereas dSA efficiency remains high with larger muon displacement.

Finally, Fig. 8.10 shows a "zoomed-in" version of the efficiency curve in the low-p_{T} region for both collections, thereby highlighting the turn-on point of the efficiency. This motivates the p_{T} threshold of 5 GeV in the identification selection for both collections which corresponds roughly to the plateau onset of the efficiency curves.

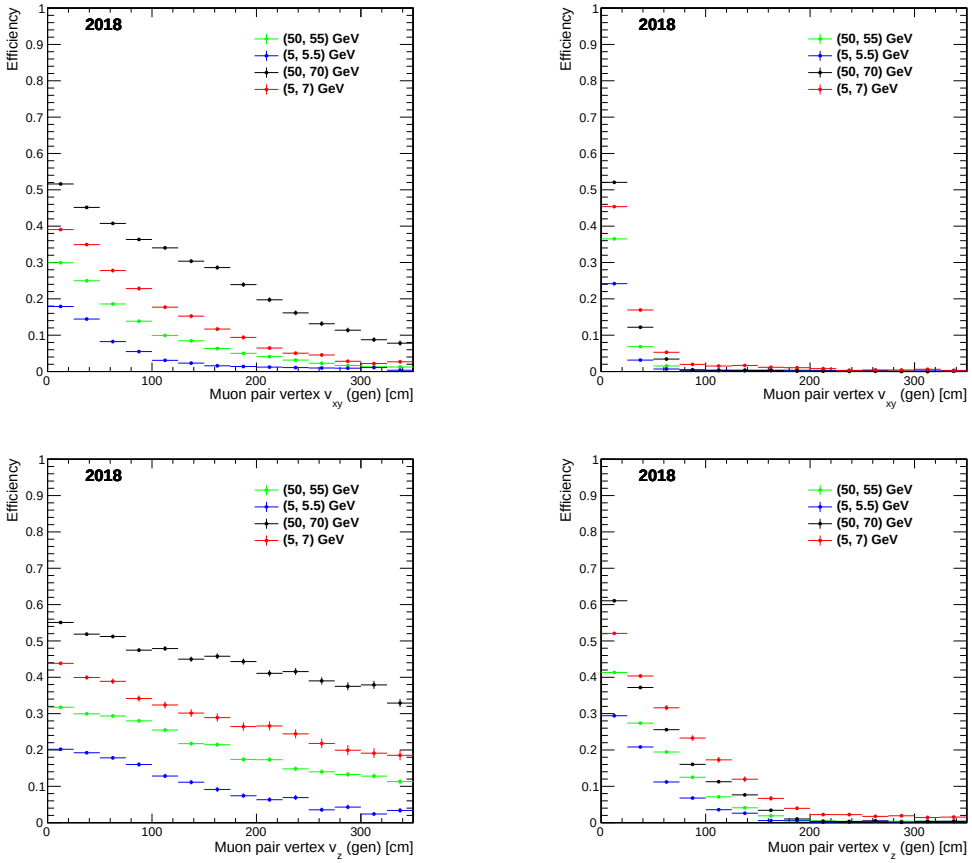


Figure 8.9: (Endcap, 2018) Reconstruction efficiency *vs.* muon true vertex v_{xy} (top) and v_z (bottom) of dSA (left) and GM (right) collections. GM efficiency drops sharply after the end of the tracker region (≈60 cm) whereas dSA efficiency remains high with larger muon displacement.

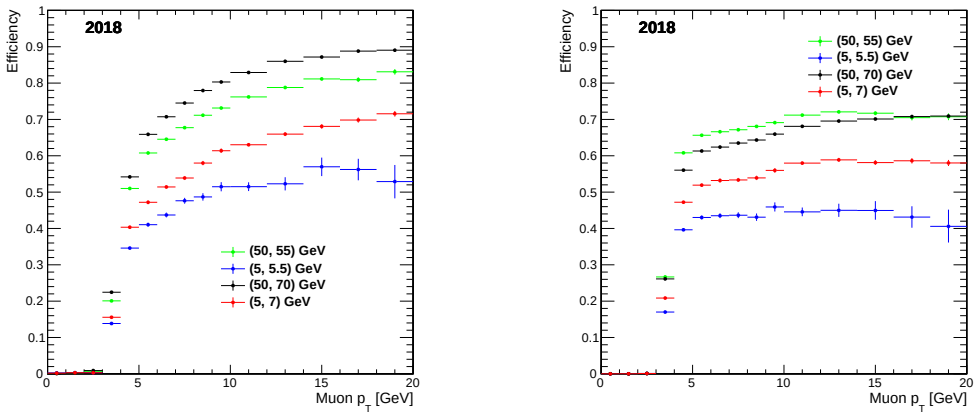


Figure 8.10: (Barrel, 2018) Reconstruction efficiency *vs.* muon p_{T} for both dSA (left) and GM (right) collections, in the low-p_{T} region. We use representative signal samples with 2018 conditions: green $(m_1, m_2) = (50, 55)$ GeV, blue $(m_1, m_2) = (50, 70)$ GeV, black $(m_1, m_2) = (50, 70)$ GeV, and red $(m_1, m_2) = (5, 7)$ GeV. For each mass point, events with all lifetimes are combined.

8.1.3 Joint reconstruction efficiency

In addition to individual reconstruction efficiencies for each collection, we also study the "joint" reconstruction efficiency between dSA muons and GMs. The goal is to validate more formally the assumption that dSA muons are a superset of GMs, in the sense that not much efficiency is lost by *first* requiring 2 dSA muons as a baseline and *only then* trying to further make use of GMs.

Table 8.1 shows a grid of the fraction of GM and dSA muons reconstructed per signal event. Reconstructed muons are matched to generator-level muons via a $dR < 0.3$ requirement. The left (right) table reports numbers for a low (high) signal efficiency sample. Green-colored cells indicate accepted events with a 2-dSA muon baseline, which is independent of the number of GMs in the event. Red-colored cells represent the fraction of events lost with this baseline requirement.

Table 8.1: Fraction of reconstructed muons per event, for dSA and GM collections. Left: low signal efficiency sample. Right: high signal efficiency sample.

$(m_1, m_2) = (5, 5.5)\,\text{GeV},\ 1000\,\text{mm}$

	0 GM	1 GM	2 GM
0 dSA	79.7%	3.5%	0.16%
1 dSA	10.8%	3.08%	0.27%
2 dSA	1.24%	0.68%	0.70%

$(m_1, m_2) = (50, 70)\,\text{GeV},\ 1\,\text{mm}$

	0 GM	1 GM	2 GM
0 dSA	0.74%	1.03%	0.72%
1 dSA	0.34%	18.6%	7.06%
2 dSA	0.05%	4.37%	67.1%

For both low and high signal efficiency samples, the efficiency loss from requiring two dSA muons per event and allowing any number of global muons (last row in each table) tends to be small. We also color orange the middle cells in each table. In principle, these events could be used in the selection if the GM and dSA muon correspond to different physical muons. However, due to the nature of the muon collimation in the iDM signal, there might be a large ambiguity when trying to determine whether the two objects correspond to the same physical muon or whether they are simply two different

physical muons in collimation. Because this ambiguity is both difficult to assess and would likely lead to marginal gains in efficiency, we choose to neglect this scenario in the analysis.

Fig. 8.11 shows the fraction of events with zero, one, and two reconstructed and matched global muons as a function of muon displacement, for events passing the baseline requirement of two reconstructed and matched dSA muons. The representative signal point is $(m_1, m_2) = (50, 55)\,\text{GeV}$ integrated over all lifetimes. Each fraction is dominant at a different region of muon displacement. Global muons can be used when muons are less displaced, but displaced standalone tracks are the only option for more displaced muons.

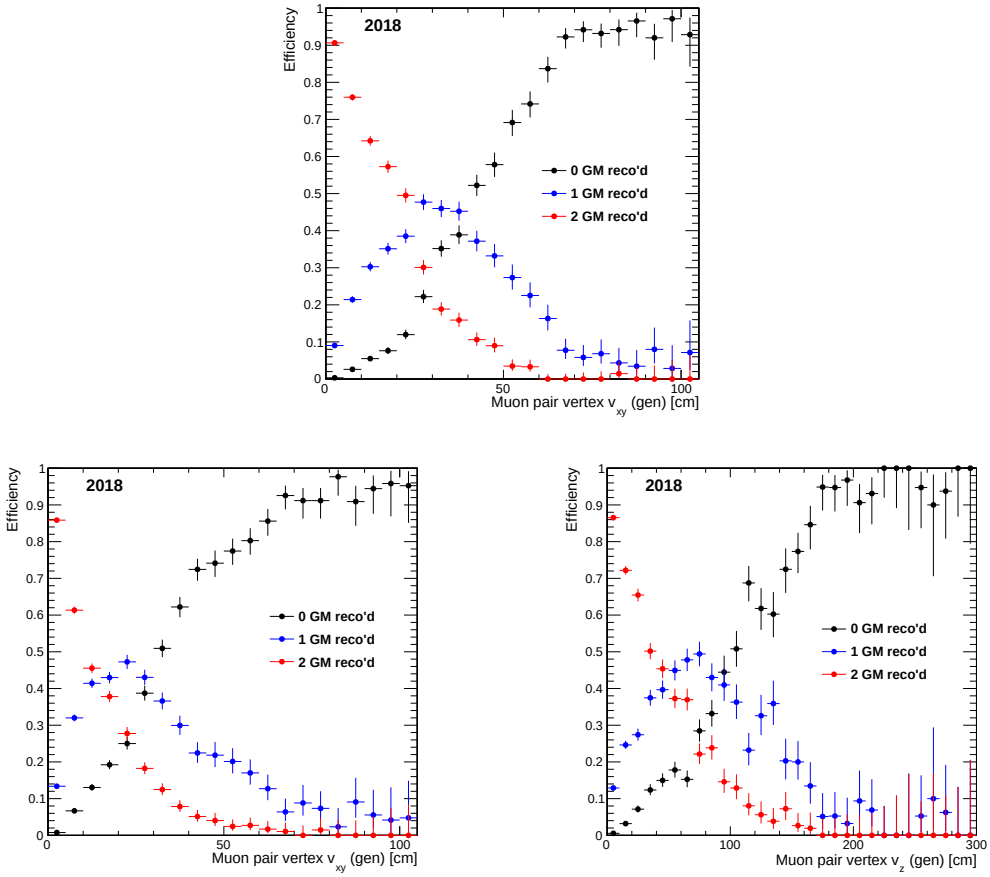


Figure 8.11: Fraction of reconstructed global muons per signal event, as a function of muon displacement. A baseline selection of 2 reconstructed and gen-matched (via *dR* and charge parity) dSA muons is applied. Upper plot: barrel region *vs.* v_{xy}. Lower plots: endcap region *vs.* v_{xy} (left), and endcap region *vs.* v_z (right). The signal sample is $(m_1, m_2) = (50, 55)$ GeV, integrated over all lifetimes. In each case, the three plots add up to 100%.

8.1.4 Resolutions

While the dSA reconstruction efficiency is higher than GM at large displacements, its performance is worse in general. The main reason is that by relaxing some of the fit constraints and using only hits from the muon chamber, more limited information is collected compared to including hits from the tracker as well. This motivates the strategy of replacing dSA muons with GMs when the latter is available which maximizes the amount of useful information.

To illustrate this point, Fig. 8.12 contrasts the p_{T} resolution between dSA and GM for the representative sample $(m_1, m_2) = (50, 55)\,\mathrm{GeV}, c\tau = 100\,\mathrm{mm}$. The GM and dSA resolutions are roughly 1 GeV and 4 GeV, respectively. Furthermore, the dSA tail makes the distribution more asymmetric than its GM counterpart.

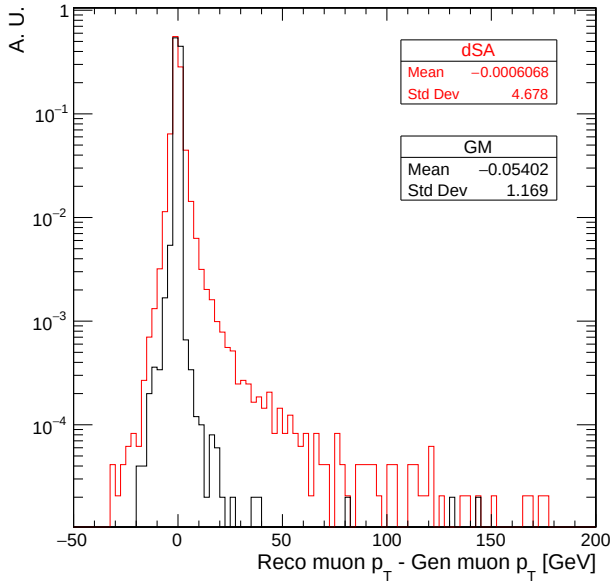


Figure 8.12: Muon p_{T} absolute resolution for dSA and GM collections, for representative sample $(m_1, m_2) = (50, 55)\,\mathrm{GeV}, c\tau = 100\,\mathrm{mm}$. GMs perform better than dSA muons, with roughly 3 times better resolution.

8.1.5 Muon matching and replacement

The interplay between reconstruction efficiency and resolution in the two collections motivates the matching strategy proposed here. The strategy proceeds as follows: after requiring two dSA muons, a matching between GM and dSA muons in the ϕ-η plane is attempted. This is done by extrapolating the trajectories of both objects to the outer edge of the muon chamber, and then matching in dR. The extrapolation step is necessary as the muons are soft and the magnetic field flips direction between the pixel tracker and the muon chamber. The matching criterion is tight ($dR < 0.1$) due to the collimated nature of signal muons. Moreover, because the matching procedure is not a cut, there is no loss in efficiency from a failed match. Fig. 8.13 shows a sketch of the procedure.

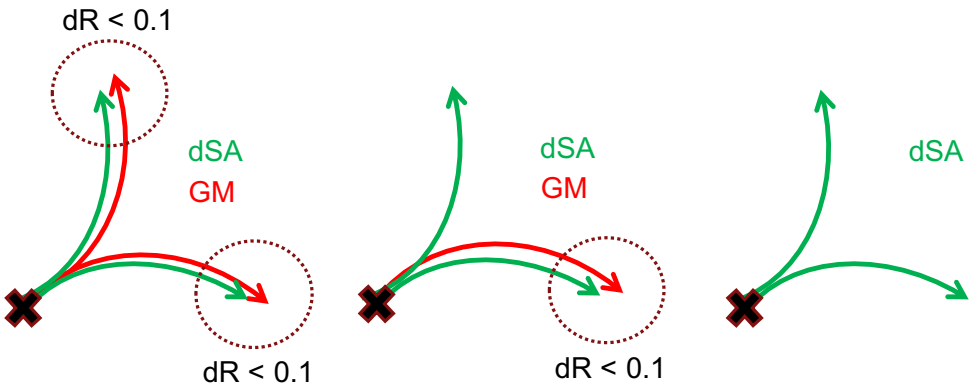


Figure 8.13: Sketch of the muon matching procedure between displaced standalone muons (dSA) and global muons (GM). The criteria for a successful matching is that both objects are within a tight $dR < 0.1$ cone of each other, *at the outermost point in their trajectory* (i.e. comparing them in their outermost muon chamber hits), due to the bending nature of soft muons in the changing magnetic field.

The matching procedure ensures that we maximize the benefits of both reconstruction algorithms in the presence of displaced muons. In each event passing the baseline two-dSA requirement, there are three possible matching outcomes: none, one, or both dSA muons are matched to GMs. This splits events into three categories that have different sources of backgrounds. These categories are used in the event selection to increase the

sensitivity of the analysis, discussed in Chapter 9.

8.2 Primary and secondary vertices

Vertices are reconstructed from charged particle tracks that point to a common z origin. The algorithm to cluster tracks into vertices uses a deterministic annealing strategy [98], in which the free "energy" of the system is minimized by assigning weights to the distance in z between each track and each vertex. The final set of vertices and their associated tracks is the one that minimizes the free energy of this system. The efficiency for the algorithm to reconstruct a vertex is roughly 70%. The primary vertex is chosen to be the vertex with the highest sum of associated jet p_{T}^2. This vertex is assumed to be the one in which the hard scattering happens.

For this analysis, a secondary vertex is constructed from displaced muon tracks. The vertex fit uses a Kalman Filter to determine the common displaced origin (if any) of two muon tracks. The quality of the fit is probed as part of the event selection. This is described in more detail in Section 9.2.3.

8.3 Jets

Quarks cannot exist in isolation because of the nature of the strong force. QCD, a non-Abelian $SU(3)$ theory, has a running coupling constant that is inversely proportional to the distance between two quarks or gluons. This leads to the phenomenon of asymptotic freedom, where quarks feel a relatively weak force once they are bound together in a compact space. Conversely, the strong force prevents bound quarks from dissociating,

since it becomes extremely attractive as the distance between the hadron constituents increases.

Therefore, when quarks and gluons are produced in pp collisions at the LHC, they immediately hadronize by pulling virtual particle pairs from the vacuum. This in turn creates more quarks and gluons, which themselves hadronize, and so on. The ensuing hadronic shower gives rise to a *jet signature* in the detector. Jets consist of a large number of clustered hadrons formed from the hadronization of quarks and gluons.

Jets are reconstructed in CMS by clustering energy deposits in the calorimeters and tracks in the trackers. Several jet reconstruction algorithms exist but a popular one uses PF candidates built from CMS's particle-flow algorithm. This search employs the anti-k_{T} jet reconstruction with a cone radius $R = 0.4$ in η-ϕ space. The distance metric for the anti-k_{T} algorithm is defined by [99]:

$$\begin{aligned} d_{ij} &= \min\left(k_{ti}^{-2}, k_{tj}^{-2}\right)\frac{\Delta_{ij}^2}{R^2}, \\ d_{iB} &= k_{ti}^{-2}, \end{aligned} \tag{8.1}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and k_{ti}, y_i, and ϕ_i are the transverse momentum, rapidity, and azimuth of a given particle i, respectively. The B label denotes the beam direction—the collinearity of a jet with the beam is also factored in. This distance metric is used to determine the best arrangement and clustering of particles into jets per event. FastJet [100] is the implementation of the anti-k_{T} reconstruction used in the analysis.

Jet reconstruction performance differs between simulated samples and data. Furthermore, reconstructed jets in both do not perfectly match the true, physical jets. The mismatch is typically measured with a sample of multi-jet events in CMS and a correc-

tion is derived to calibrate the reconstruction of jets [101]. This calibration is essentially a scaling of the jet momentum. Additionally, jets in simulated samples are smeared to artificially worsen their resolution to match the one observed in data. The first set of corrections are referred to as jet energy scale (JES) and the second set as jet energy resolution (JER) [101]. These corrections are implemented as functions of the jet p_{T} and η.

Jets used in the analysis are required to pass a set of selections. These identification criteria ensure the reconstructed jets correspond to true, physical jets [102]. The standard loose identification criteria consist of:

- Neutral Hadron Fraction < 0.99;
- Neutral EM Fraction < 0.99;
- Number of constituents > 1;
- For jets with $|\eta| < 2.4$:
  - Charged Hadron Fraction > 0;
  - Charged EM Fraction < 0.99;
  - Charged multiplicity > 0.

The analysis uses a slightly tighter version of these loose selections to suppress detector noise and beam backgrounds producing anomalous jets and $p_{\mathrm{T}}^{\mathrm{miss}}$ [103, 104, 105]:

- Charged hadron energy fraction > 0.1;
- Neutral hadron energy fraction < 0.8.

Jets that pass the identification criteria are further subjected to kinematic cuts to ensure they are not caused by pileup:

- $p_{\mathrm{T}} > 30\,\mathrm{GeV}$;
- $|\eta| < 2.5$.

Finally, a b quark identification algorithm (b-tagger) attempts to identify jets originating from heavy flavor physics. Events with b-tagged jets are rejected, since heavy flavor processes do not contribute to the expected iDM signature. Therefore this veto is useful to remove top quark backgrounds.

We use the DeepCSV algorithm [106] for all years with a medium working point, which has a 1% probability of misidentifying a light flavor jet as a b jet. The discriminant thresholds (which range from 0 to 1) for this working point are:

- discriminator > 0.4184 (2018 and 2017);
- discriminator > 0.6324 (2016).

8.4 Missing transverse momentum

CMS has the ability to measure $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$. The hermetic coverage of the detector due to its cylindrical layout makes all particles produced in the collision accessible in principle. Since the incoming beams have zero transverse momentum, the total vector sum of outgoing particle momenta should also have zero $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ by conservation of momentum. However, if invisible particles such as neutrinos (or dark matter) are produced, they will escape the apparatus without detection and leave an unbalanced $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$. Thus $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ is used as an essential handle in the search for feebly-interacting particles in pp collisions at the LHC.

CMS has several algorithms to reconstruct $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ [105]. The one used in this analysis is based on PF candidates in a similar vein to jet reconstruction. Also similarly to jets, $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction performance can be degraded because of detector effects and miscalibration relative to true particle momenta. The corrections applied to jets described in Section 8.3 are also propagated to $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$. The corrected object is:

$$\mathbf{p}_{T}^{\mathrm{miss,corr}} = \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}} - \sum_{\mathrm{jets}} \left(\mathbf{p}_{T,\mathrm{jet}}^{\mathrm{corr}} - \mathbf{p}_{T,\mathrm{jet}} \right). \tag{8.2}$$

An additional correction is applied to $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ due to an undesirable azimuthal sinusoidal modulation present in data. This modulation has several potential causes, among them detector response issues and misalignments between the nominal and actual interaction points. This modulation is removed with a correction, which restores the expected uniform distribution of the ϕ component of $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$.

Occasionally, large anomalous $p_{\mathrm{T}}^{\mathrm{miss}}$ can be produced in events with no real source of invisible particles. This phenomenon can be caused by random detector failures and non-collisional physics backgrounds [105]. Filters were developed to remove such events from consideration. These filters are applied to both data and simulated events, and in the case of inelastic dark matter, result in negligible loss of signal acceptance. The applied filters are:

- Primary vertex filter: removes events without reconstructed primary vertices;
- Beam halo filter: removes beam halo backgrounds;
- ECAL dead cell trigger primitive filter: removes events with dead ECAL channels that fired HLT triggers;
- ECAL bad calibration filter: remove events with poorly calibrated ECAL;

- Bad PF muon filter: removes events with poor PF muon reconstruction;
- Bad muon track filter: removes events with poor track muon reconstruction;
- HBHE noise filter: removes events with poor HCAL performance in general;
- HBHE isolated noise filter: removes events with poor isolated HCAL channels.

8.4.1 Corrections from muon replacement

The particle-flow reconstruction of $p_{\mathrm{T}}^{\mathrm{miss}}$ includes all PF candidates. PF muons with global tracks are therefore included. However, displaced standalone muon reconstruction happens outside of the PF workflow, so dSA muons are not included in $p_{\mathrm{T}}^{\mathrm{miss}}$. One needs to be careful about this subtlety as it relates to muon matching. In general, there are four possible situations encountered for any given physical muon:

1. The muon is reconstructed as a dSA muon but not as a GM. In this case, the $p_{\mathrm{T}}^{\mathrm{miss}}$ calculation does not include the muon, so its momentum needs to be added back to the $p_{\mathrm{T}}^{\mathrm{miss}}$.
2. The muon is reconstructed both as a dSA muon and as a GM and they match. In this case, to use the dSA muon, the GM contribution would need to be subtracted from the $p_{\mathrm{T}}^{\mathrm{miss}}$ and the dSA contribution added in its place. However, to use the GM muon, no correction is needed.
3. The muon is reconstructed both as a dSA muon and as a GM but there is no match. In this case, the dSA object is used since it is the baseline. But if there is no matching, we do not know that the GM object actually refers to the same physical muon. The dSA p_{T} would therefore be added back to the $p_{\mathrm{T}}^{\mathrm{miss}}$ and double-counted. The effect of this mismatch in general is a worsening of the $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ p_{T} and ϕ resolutions, which depends on the level of matching inefficiency.

4. There is a reconstructed GM but not dSA muon. Given the dSA baseline selection, this event would be discarded.

For situation #2, there is an additional benefit of preferring global muons over dSA muons when there is a match: no $p_{\mathrm{T}}^{\mathrm{miss}}$ correction is needed in this case. The only correction needed is to subtract the dSA momentum from $p_{\mathrm{T}}^{\mathrm{miss}}$ in situation #1. These corrections lead to modest improvements in the bias of the $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ p_{T} and ϕ distributions when compared to the equivalent truth-level quantities, shown in Fig. 8.14 for a representative signal sample. The bias after correcting for overlapping dSA and global muons is reduced, though we do not see an equivalent improvement in the resolutions themselves because the muons are generally soft while the $p_{\mathrm{T}}^{\mathrm{miss}}$ is energetic. Fig. 8.15 shows the corrected *vs.* uncorrected $p_{\mathrm{T}}^{\mathrm{miss}}$ distributions. There is a small shift of the distribution towards higher values.

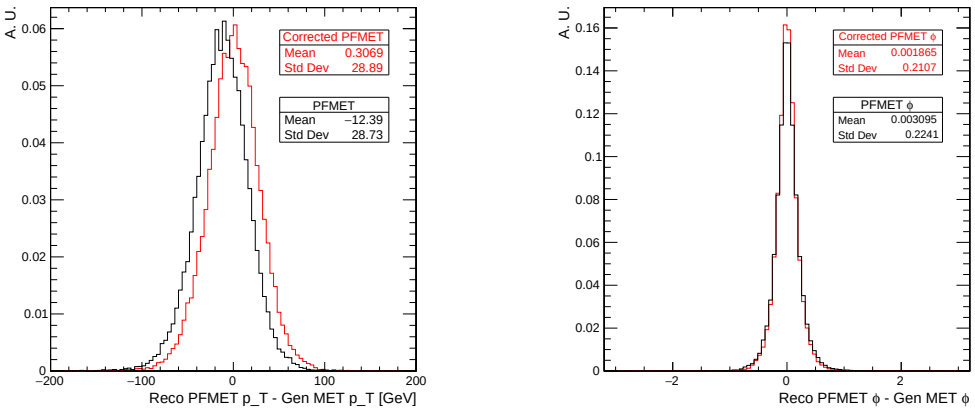


Figure 8.14: $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ p_{T} (left) and ϕ (right) resolutions, before (black curve) and after (red curve) corrections due to overlapping dSA and global muons, for a representative signal sample with $(m_1, m_2) = (50, 55)\,\mathrm{GeV}, c\tau = 100\,\mathrm{mm}$. There is a decreased bias and slightly improved resolutions in both distributions.

Situation #3 is a bit more subtle. To fully understand the effects on the $p_{\mathrm{T}}^{\mathrm{miss}}$, the matching efficiency between GMs and dSA muons needs to be characterized. Since $\phi(\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}})$ is used in the background estimation, the matching efficiency could induce a

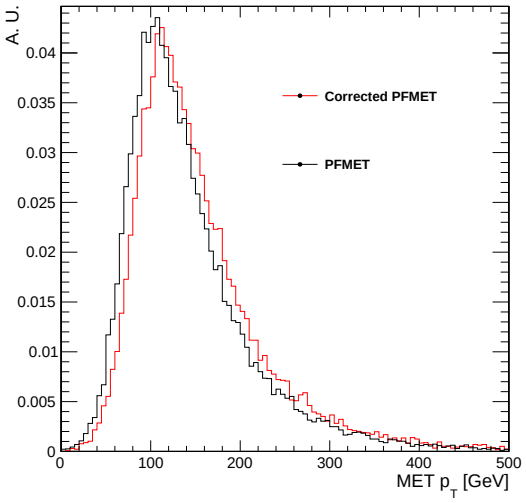


Figure 8.15: $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution before (black curve) and after (red curve) corrections due to overlapping dSA and global muons, for a representative signal sample with $(m_1, m_2) = (50, 55)\,\mathrm{GeV}, c\tau = 100\,\mathrm{mm}$.

systematic error in this estimation. For inelastic dark matter, however, this error is very small due to the contrast between soft muons, on one hand, and the energetic ISR jet on the other, which drives a large $p_{\mathrm{T}}^{\mathrm{miss}}$. To a first approximation, therefore, we can neglect the effects of matching inefficiencies between GMs and dSA muons, though for other analyses targeting signatures with collimated $p_{\mathrm{T}}^{\mathrm{miss}}$ and energetic leptons, this issue might warrant a more thorough investigation.

CHAPTER 9

EVENT SELECTION

This chapter describes the event selection applied to both data and simulated samples. Event selection optimization has the ultimate goal of increasing the relative yield of signal compared to background and therefore obtaining sensitivity to inelastic dark matter. The selection relies on particular features of expected signal kinematics which help to reduce backgrounds that do not share those same features. The selection requirements are largely optimized with “N-1” plots, a technique that consists of plotting the distribution of a particular observable for both signal and background simulated samples after applying all requirements except for the one in the observable under study. This is discussed in more detail in the following sections.

The event selection is split into three groups: jet-MET, muon, and vertex selections. The signal region is defined by applying all selection groups in succession. Two orthogonal validation regions are also defined by inverting some selections in the chain in order to enhance and study particular backgrounds found in the signal region. Validation regions are also used to assess the closure of the data-driven background estimation strategy, discussed in Chapter 10.

9.1 N-1 optimizations

Several selection criteria were optimized with “N-1" studies. In these optimizations, one typically runs both signal and background simulated events through the entire selection except for the requirement in the observable under study. The significance as a function of selection threshold is estimated by looking at the distribution of this observable in signal compared to background. The observable is scanned through a range of thresholds

and for a given threshold, the cumulative distribution is computed in both "forward" and "backward" directions (i.e. integrating from the beginning of the distribution to the threshold, or from the threshold to the end of the distribution). Cumulative distributions are the event yields and a metric is used to determine the significance for each cut. Here, the simple metric $s/\sqrt{b}$ is picked since $s \ll b$ is still satisfied in the early selection stages, making this metric a reasonable approximation of true significance. The forward and backward cumulative distributions are both needed because different observables have different initial and final intervals, e.g. number of jets from 0 to X (backward cumulative distribution), and leading jet p_{T} from Y to ∞ (forward cumulative distribution). The procedure is sketched in Fig. 9.1.

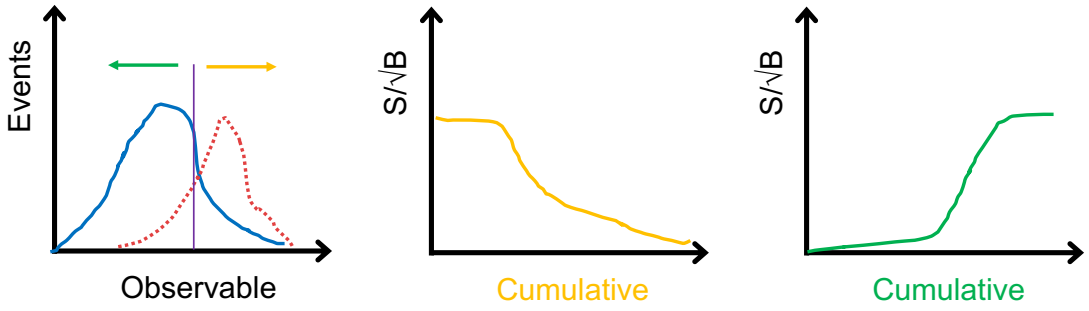


Figure 9.1: Sketch of the optimization procedure with the N-1 method. Both signal and background simulation distributions for a given observable are scanned as a function of selection threshold, in the forward (green) and backward (orange) directions. The cumulative distributions are used with a $S/\sqrt{B}$ metric to determine the optimal threshold. The appropriate distribution to use (forward or backward) depends on the observable under study.

9.2 Event selection

We seek events with two soft, narrow, and displaced muons, collimated with significant missing transverse momentum, and recoiled off an energetic ISR jet. The event selection is accordingly split into three parts: jet-MET, muon, and vertex selections. Jet-MET

selection places requirements on the jet and $p_{\mathrm{T}}^{\mathrm{miss}}$ activity of the event, while muon and vertex selections accept events containing well-identified muons compatible with a common production origin as determined by a Kalman Filter vertex fit. The full set of selections make up the signal region. Inside this signal region, three categories are defined based on the number of GM-dSA muon matches as described in Section 8.1.5. The discussion here focuses on the optimizations leading to the final event selection, with the overall goals of increasing signal efficiency, reducing backgrounds, and maximizing sensitivity to iDM in CMS.

9.2.1 Jet-MET selection

Jet and $p_{\mathrm{T}}^{\mathrm{miss}}$ selections require there to be hard jet activity and significant missing transverse momentum in the event. For example, we require $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers to be fired and at least one energetic jet with p_{T} greater than 80 GeV. There is also a requirement on the minimum azimuthal separation between jets and $p_{\mathrm{T}}^{\mathrm{miss}}$. This has a dual effect of isolating the dark matter system from jets (which could constitute a source of background if there are muons inside) and also of reducing QCD backgrounds. Since the dark matter system (two displaced muons plus $p_{\mathrm{T}}^{\mathrm{miss}}$) is expected to recoil off the leading ISR jet, an especially tight constraint is placed on the azimuthal separation between leading jet and $p_{\mathrm{T}}^{\mathrm{miss}}$ ($|\Delta\phi| > 1.5$), with looser requirements on remaining jets ($|\Delta\phi| > 0.5$) so as not to drastically reduce the signal efficiency. The complete set of jet-MET requirements is:

- $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers fired;
- Offline $p_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV;
- $|\mathrm{MET}_{PF} - \mathrm{MET}_{CALO}|/\mathrm{MET}_{CALO} < 0.5$;
- Leading jet $p_{\mathrm{T}} > 80$ GeV;

- Leading jet $|\eta| < 2.5$;
- $|\Delta\phi(p_{\mathrm{T}}^{\mathrm{miss}}, \text{leading jet})| > 1.5$;
- $|\Delta\phi(p_{\mathrm{T}}^{\mathrm{miss}}, \text{all jets})| > 0.5$;
- nJets < 3 ($p_{\mathrm{T}} > 30\,\mathrm{GeV}$);
- 0 b-tagged jets.

Fig. 9.2 (left) shows N-1 distributions of leading jet p_{T} per event for several signal samples with varying dark matter masses and lifetimes and for background samples. Corresponding scans of forward cumulative distributions are shown on the right. The most sensitive thresholds are located at lower jet p_{T}, specifically around 80 GeV. This informs the choice of leading jet p_{T} requirement in the selection above. This is also in agreement with selections performed in other searches similarly involving $p_{\mathrm{T}}^{\mathrm{miss}}$ and jets (e.g. a CMS search for dark matter using a "monojet" signature set this threshold to 110 GeV [107]). Finally, it is advantageous to keep this threshold as low as possible to maximize signal acceptance, particularly of lower-mass samples.

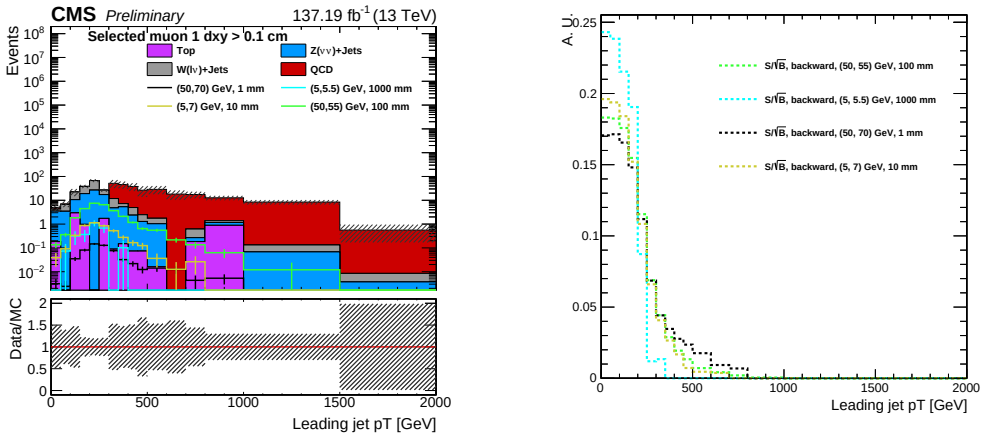


Figure 9.2: N-1 distributions of leading jet p_{T} (left) for signal and background simulated samples, and corresponding backward cumulative distributions (right). The optimal threshold is near lower values of jet p_{T}.

Additional jets in the event are required to have at least 30 GeV p_{T}. This reduces acceptance of soft pileup jets and also agrees with other analyses that use both $p_{\mathrm{T}}^{\mathrm{miss}}$ and jets (e.g. [107]). The offline $p_{\mathrm{T}}^{\mathrm{miss}}$ threshold in turn was chosen based on detailed trigger studies discussed in Chapter 7, which indicated the plateau efficiency of the $p_{\mathrm{T}}^{\mathrm{miss}}$-based triggers is reached near 200 GeV.

Finally, as part of the jet-MET selection for data collected in 2018, a "HCAL endcap minus (HEM) veto" is imposed, which requires there to be no jets in a particular region of ϕ-η space. This requirement addresses an issue in sectors 15 and 16 of the HEM detector that occurred in June of 2018 and made them non-operational for the remainder of 2018 data-taking [108]:

- HEM 15/16 region veto: no jets with $-3.0 < \eta < -1.4$ and $-1.57 < \phi < -0.87$.

9.2.2 Muon selection

As discussed in Section 8.1, muon selection requires a baseline of two well-identified dSA muons. The dSA identification criteria are repeated here for convenience:

- Number of muon chamber stations ≥ 2;
- Total number of hits ≥ 12;
- Track $\chi^2/\mathrm{dof} < 10$;
- $p_T > 5$ GeV;
- $|\eta| < 2.4$;
- $\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}} < 1$;

The chosen p_T threshold lies close to the plateau of dSA reconstruction efficiency as shown earlier in Fig. 8.10. The σ_{p_T}/p_T selection protects against too poorly reconstructed dSA muons with high transverse momentum uncertainty, but is not very restrictive, as shown in Fig. 9.3.

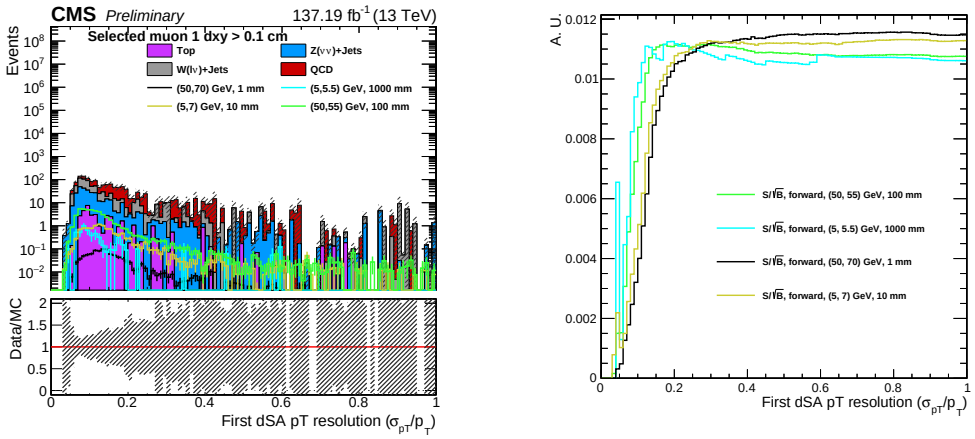


Figure 9.3: N-1 distributions of σ_{p_T}/p_T for signal and background simulated samples. The $\sigma_{p_T}/p_T < 1$ threshold prevents the inclusion of dSA muons with too poor a resolution, but is otherwise efficient and avoids removing too much signal.

For the leading muon, a further requirement is imposed:

- $MT < 300\,\text{GeV}$, where $MT \equiv \sqrt{2 \times p_{T,\mu} \times p_{\text{T}}^{\text{miss}}\ (1 - \cos\Delta\phi)}$.

Values of MT as defined tend to be lower for signal and higher for various backgrounds, providing an extra discrimination handle. This is illustrated in Fig. 9.4, which compares the MT distributions between signal and background simulated samples.

9.2.3 Vertex selection

Events that satisfy the requirement of two well-identified muons are further subjected to a vertex selection. This selection requires the two muons to form a well-reconstructed

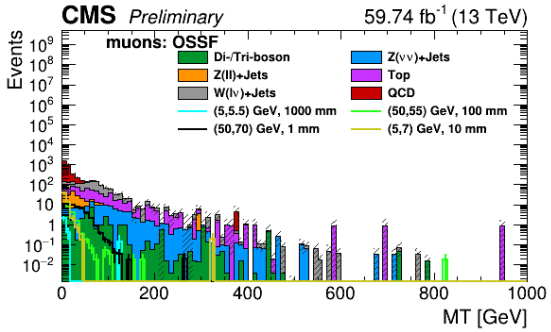


Figure 9.4: N-1 distributions of leading muon MT as defined in the text, for signal and background simulated samples. Due to the soft p_{T} distribution of muons in signal, the MT variable also has a soft distribution, and it is therefore possible to remove SM backgrounds without losing signal efficiency by applying a loose selection on MT. The chosen 300 GeV threshold ensures no signal loss, roughly accounting for the projected MT of higher dark matter mass samples.

vertex which would be expected of a typical iDM event. The vertex is reconstructed with the standard Kalman Filter algorithm used by CMS [109]. Vertex-related requirements are:

- Vertex $\chi^2/\mathrm{dof} < 4$;
- $dR(\mathrm{muons}) < 0.9$;
- charge(muon1) $\neq$ charge(muon2);
- $M_{\mu\mu} < 30\,\mathrm{GeV}$.

These criteria were also optimized with N-1 distributions, shown in Figs. 9.5 to 9.7. The vertex χ^2/dof requirement is important to prevent poorly reconstructed vertices from eventually contaminating the signal region. The dR selection exploits the fact that the dimuon pair is collimated in iDM, as opposed to most SM backgrounds where the distributions are flatter. Finally, the invariant mass of the dimuon pair is limited by the mass splitting between the two dark matter states. Since signal samples produced for the analysis have a maximum mass splitting of 32 GeV (with $(m_1, m_2) = (80, 112)\,\mathrm{GeV}$), we

impose this additional requirement to remove backgrounds with larger dimuon invariant mass. This is shown in Fig. 9.7.

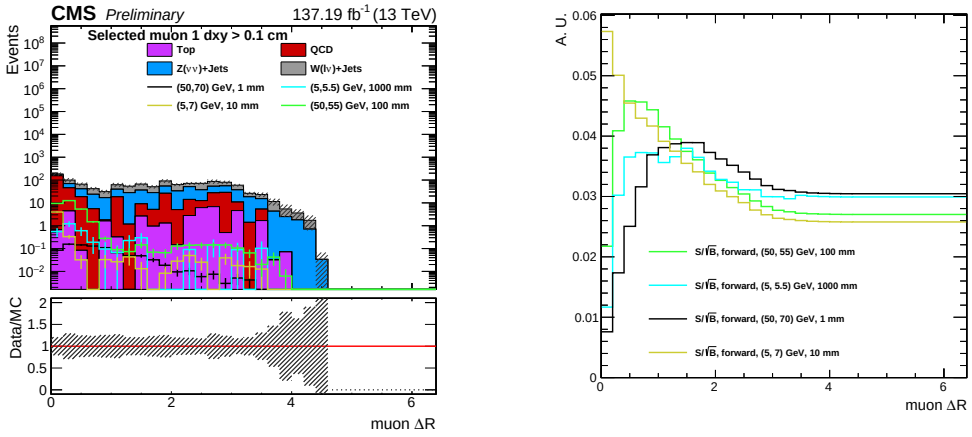


Figure 9.5: N-1 distributions of dimuon dR for signal and background simulated samples, and corresponding forward cumulative distributions (right). The optimal threshold is near $dR < 1$.

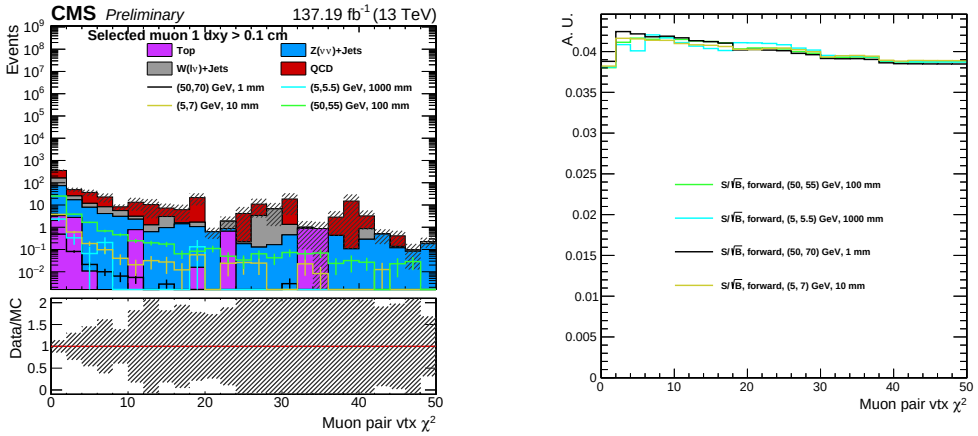


Figure 9.6: N-1 distributions of dimuon vertex χ^2/dof for signal and background simulated samples, and corresponding forward cumulative distributions (right). The optimal threshold is around χ^2/dof < 4.

Finally, both muons are required to have impact parameter d_{xy} greater than 1 mm. This is the first explicit requirement of a displaced signature and removes a large fraction of prompt electroweak backgrounds. N-1 plots of the d_{xy} distribution are shown in Fig. 9.8.

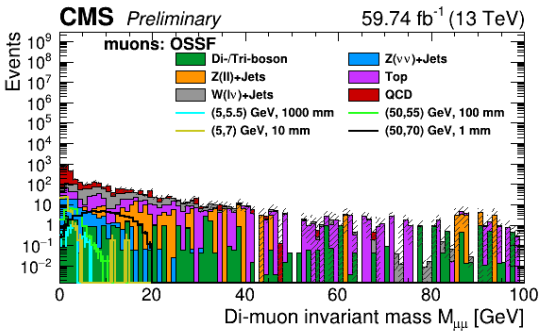


Figure 9.7: N-1 distributions of the dimuon invariant mass $M_{\mu\mu}$ for signal and background simulated samples. Due to the low mass splitting between dark matter states, the invariant mass distribution is signal is quite soft; it is therefore possible to remove SM backgrounds without losing signal efficiency. The chosen threshold is 30 GeV to ensure no signal loss for the highest mass-splitting sample, $(m_1, m_2) = (80, 112)$ GeV.

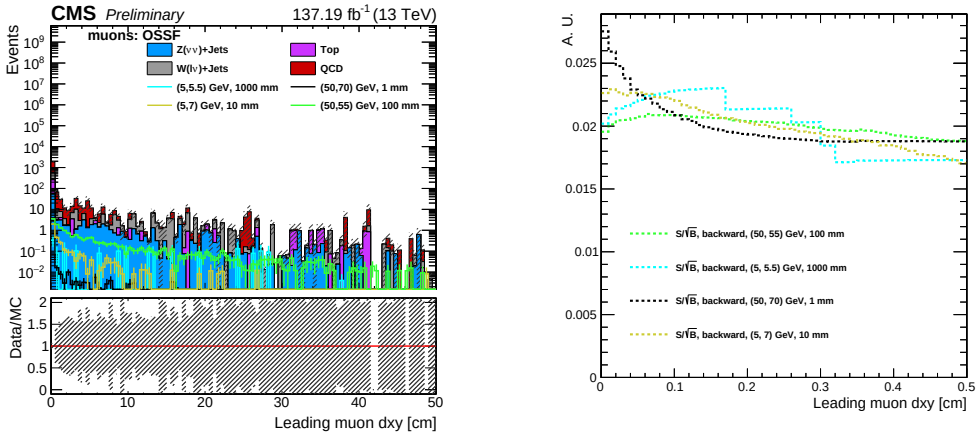


Figure 9.8: N-1 distributions of muon d_{xy} for signal and background simulated samples. The $|d_{xy}| > 1$ mm threshold helps remove prompt electroweak backgrounds and is one of the explicit requirements of a long-lived signature.

9.2.4 Muon replacement

Events passing both muon and vertex selections undergo a muon replacement procedure, where matchings between the selected dSA muons and global muons are attempted. The matching criteria are:

- $dR < 0.1$ between dSA and GM at outer edge of muon chamber;
- Well-reconstructed vertex between new dSA/GM pair or new GM/GM pair.

The dR criterion is tight due to the already natural dimuon collimation of the signal but results in no loss of efficiency since muon replacement is an inclusive procedure (i.e. events are not rejected if matches fail). This strategy is described in detail in Section 8.1.5. The matching procedure produces three possible outcomes: zero, one, or two dSA muons are successfully matched with GMs. For every successful match, the dSA is replaced with a GM and the vertex properties are updated to reflect the replacement.

To reiterate the discussion in Section 8.1.5, this muon-matching procedure has two main benefits: first, it improves the performance of muon and vertex reconstruction where possible since GMs have better performance than dSA muons in general. And second, by splitting the signal region into three categories based on the number of GM-dSA matches, we gain a further ability to discriminate between more displaced signal events and prompt SM backgrounds. Backgrounds with real muons have predominantly one or two GM-dSA matches, while signal events typically have only one or zero GM-dSA matches, depending on displacement.

9.2.5 Photon and electron vetoes

Electrons and photons are vetoed to suppress electroweak backgrounds. The criteria used to identify these objects follow the standard prescriptions developed centrally by CMS.

The working point chosen for electrons was designed for vetoing purposes and has a ≈95% identification efficiency. The criteria for this working point are listed in Tables 9.1 and 9.2, for 2017/2018 and 2016, respectively.

Table 9.1: Cut-based electron ID criteria used in 2017 and 2018 with 'Veto' working point (95% efficiency). The supercluster (SC) η defines the barrel and the endcap regions.

	Veto WP (SC $\lvert\eta\rvert$ < 1.479)	Veto WP (SC $\lvert\eta\rvert$ > 1.479)
$\sigma_{i\eta i\eta}$ (full 5x5) <	0.0126	0.0457
abs(dEtaSeed) <	0.00463	0.00814
abs(dPhiIn) <	0.148	0.19
H/E <	0.05+1.16/E_{SC}+0.0324ρ/E_{SC}	0.05+2.54/E_{SC}+0.183*ρ/E_{SC}
relIsoWithEA <	0.198+0.506/pT	0.203+0.963/pT
abs(1/E-1/ρ) <	0.209	0.132
expected missing inner hits <=	2	3
pass conversion veto	yes	yes

Table 9.2: Cut-based electron ID criteria used in 2016 with 'Veto' working point (95% efficiency). The supercluster (SC) η defines the barrel and the endcap regions.

	Veto WP (SC $\lvert\eta\rvert$ < 1.479)	Veto WP (SC $\lvert\eta\rvert$ > 1.479)
$\sigma_{i\eta i\eta}$ (full 5x5) <	0.0115	0.037
abs(dEtaSeed) <	0.00749	0.00895
abs(dPhiIn) <	0.228	0.213
H/E <	0.356	0.211
relIsoWithEA <	0.175	0.159
abs(1/E-1/ρ) <	0.299	0.15
expected missing inner hits <=	2	3
pass conversion veto	yes	yes

Tables 9.3 and 9.4 list the criteria used for photon ID in 2017/2018 and 2016, respectively, in the 'Loose' working point (roughly 90% efficiency), since an official 'Veto' working point is not available for photons.

Table 9.3: Cut-based photon ID criteria used in 2018 and 2017, with 'Loose' working point (90% efficiency).

	Loose WP (Barrel)	Loose WP (Endcap)
H/E <	0.04596	0.0590
$\sigma_{i\eta i\eta}$ <	0.0106	0.0272
PF charged hadron iso <	1.694	2.089
PF neutral hadron iso <	$24.032 + 0.01512p_{\mathrm{T}} + 0.00002259p_{\mathrm{T}}^2$	$19.722 + 0.0117p_{\mathrm{T}} + 0.000023p_{\mathrm{T}}^2$
PF photon iso <	$2.876 + 0.004017p_{\mathrm{T}}$	$4.162 + 0.0037p_{\mathrm{T}}$

Table 9.4: Cut-based photon ID criteria used in 2016, with 'Loose' working point (90% efficiency).

	Loose WP (Barrel)	Loose WP (Endcap)
H/E <	0.0597	0.0481
$\sigma_{i\eta i\eta}$ <	0.01031	0.03013
PF charged hadron iso <	1.295	1.011
PF neutral hadron iso <	$10.910 + 0.0148p_{\mathrm{T}} + 0.000017p_{\mathrm{T}}^2$	$5.931 + 0.0163p_{\mathrm{T}} + 0.000014p_{\mathrm{T}}^2$
PF photon iso <	$3.630 + 0.0047p_{\mathrm{T}}$	$6.641 + 0.0034p_{\mathrm{T}}$

The PF isolation variables (labeled 'iso' in the tables above) are all ρ-corrected (to account for pileup), computed via:

$$\text{corrected PFIso} = \max(\text{PFIso} - \rho\text{EA}, 0.), \tag{9.1}$$

where EA stands for Effective Area, which are given as a function of photon η in Table 9.5 (2017 and 2018) and Table 9.6 (2016).

Table 9.5: Effective Area (EA) as a function of photon η used for calculation of ρ-corrected PF isolation, in 2017 and 2018.

	EA charged hadrons	EA neutral hadrons	EA photons
Containment	70%	90%	90%
$\lvert\eta\rvert < 1.0$	0.0112	0.0668	0.1113
$1.0 < \lvert\eta\rvert < 1.479$	0.0108	0.1054	0.0953
$1.479 < \lvert\eta\rvert < 2.0$	0.0106	0.0786	0.0619
$2.0 < \lvert\eta\rvert < 2.2$	0.01002	0.0233	0.0837
$2.2 < \lvert\eta\rvert < 2.3$	0.0098	0.0078	0.1070
$2.3 < \lvert\eta\rvert < 2.4$	0.0089	0.0028	0.1212
$2.4 < \lvert\eta\rvert$	0.0087	0.0137	0.1466

Table 9.6: Effective Area (EA) as a function of photon η used for calculation of ρ-corrected PF isolation, in 2016.

	EA charged hadrons	EA neutral hadrons	EA photons
Containment	70%	90%	90%
$\lvert\eta\rvert < 1.0$	0.0360	0.0597	0.1210
$1.0 < \lvert\eta\rvert < 1.479$	0.0377	0.0807	0.1107
$1.479 < \lvert\eta\rvert < 2.0$	0.0306	0.0629	0.06999
$2.0 < \lvert\eta\rvert < 2.2$	0.0283	0.0197	0.1056
$2.2 < \lvert\eta\rvert < 2.3$	0.0254	0.0184	0.1457
$2.3 < \lvert\eta\rvert < 2.4$	0.0217	0.0284	0.1719
$2.4 < \lvert\eta\rvert$	0.0167	0.0591	0.1998

The difference in identification efficiency between simulation and data is also centrally measured. These scale factors are parameterized by the object p_{T} and η. The scale factors are implemented in the analysis with a "veto weight" method which condenses all information about vetoed objects into a single weight. This weight is applied to simulated events to correct for discrepancies between data and simulation inefficiencies. The applied weight is given by:

$$w = \prod_{i \in \{\gamma, e\}} \left(1 - \mathrm{SF}_i\right), \tag{9.2}$$

where the product runs over all identified electrons and photons in the event, and SF_i is the measured scale factor for each such object based on its p_{T} and η. If the scale factors were all identically one, then w would be zero, thereby perfectly vetoing all events containing at least one electron or photon. Since scale factors are close to but not exactly one, the inefficiency is translated into the small weight w which is applied to simulated samples.

Uncertainties on the scale factors are varied independently, and the resulting difference in w is included as a systematic uncertainty in the analysis. This is further discussed in Chapter 11.

9.3 Signal region

The signal region (SR) is defined by the jet-MET, muon, and vertex selections. Inelastic dark matter events produce significant jet and $p_{\mathrm{T}}^{\mathrm{miss}}$ activity in the detector, as well as collimated, soft and displaced muons with low invariant mass. Here, I discuss the physics backgrounds that could mimic this signature. Table 9.7 shows a breakdown of

the simulated background samples found in the signal region.

Table 9.7: (SR) Breakdown of the primary background yields in the signal region before and after imposing the impact parameter displacement requirement and split into the three GM-dSA match categories, according to simulation. Yields are scaled to 2018+2017+2016 luminosity (137.2 fb^{-1}). Top backgrounds include $\mathrm{t\bar{t}}$ and single top quark processes.

	Total	QCD	Top	Z+jets	W+jets
before $\lvert d_{xy}\rvert > 1\,\mathrm{mm}$	1179.9	760.3	149.4	212.2	218.5
after $\lvert d_{xy}\rvert > 1\,\mathrm{mm}$	405.7	218.6	6.1	96.7	84.3
0 GM-dSA	198.7	29.2	0.5	90.4	78.7
1 GM-dSA	168.3	152.1	5.3	6.3	4.6
2 GM-dSA	38.7	37.4	0.3	0.0	1.0

The main SM backgrounds that could in principle end up in the signal region are QCD and top quark processes in the semi-leptonic channel ($\mathrm{t\bar{t}}$ and single top quark). These can have real, displaced muons and significant hadronic activity. Drell–Yan events can also feature muons and jets but due to the displacement requirement they mostly do not survive the event selection and are therefore not a significant source of backgrounds.

In practice, top quark physics constitutes only a small fraction of backgrounds in the signal region, since we veto events with any b-tagged jets. The majority of background events are from QCD. QCD has an extremely large cross section at the LHC and muons can have fairly tight collimation and low invariant mass. A QCD-enriched orthogonal validation region is defined by inverting the requirement on the number of jets with p_{T} greater than 30 GeV (nJets $\geq$ 3). Furthermore, this multijet validation region is very close to the signal region in its background composition and therefore useful in multiple ways.

To elaborate, the selection on number of jets was also initially optimized using N-1 distributions, shown in Fig. 9.9. The most sensitive threshold is actually nJets $<$ 4, though the dependence is weak after nJets $<$ 3. Since the sensitivity does not vary

much with larger values of nJets, a higher threshold choice could increase the signal efficiency without loss in sensitivity. However, this multijet observable plays a key role as a validation region, since it is orthogonal but "physically" close to the signal region, and therefore is a useful proxy to the blinded analysis. This is especially helpful to estimate predicted backgrounds from data, discussed in Chapter 10. Hence we split the nJets distribution such that the validation region and signal region yields are roughly balanced.

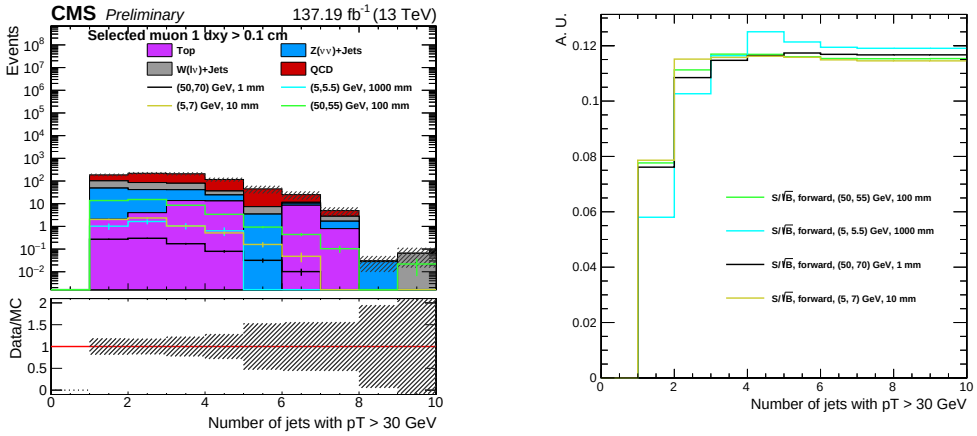


Figure 9.9: N-1 distributions of nJets (number of jets with $p_{\mathrm{T}} > 30$ GeV), for signal and background simulated samples, and corresponding backward cumulative distributions (right). The choice of threshold, nJets < 3, is motivated in the text.

An additional source of backgrounds arises from misidentified muons in electroweak events ($Z \rightarrow \nu\nu$ + jets and $W \rightarrow l\nu$ + jets). These channels have significant $p_{\mathrm{T}}^{\mathrm{miss}}$ and jet activity, and together with the unrelated muons, can survive the event selection. An orthogonal validation region is also defined for these backgrounds, discussed below.

Finally, non-collisional backgrounds such as cosmic and beam halo muons are typically a source of concern in analysis with displaced muons. However, owing to the extensive event selection, these sources of background do not significantly contaminate the signal region. In particular, since we use $p_{\mathrm{T}}^{\mathrm{miss}}$-based triggers, the incidence of cosmic and beam halo muons is already greatly reduced from the outset. It is possible for

these muons to occur simultaneously with $p_{\mathrm{T}}^{\mathrm{miss}}$ and jet activity, but the extra kinematic selections in the analysis (such as several collimation requirements) translate to a low likelihood that such non-collisional muons survive the full event selection.

To further validate this claim, the event selection was run through a NoBPTX primary dataset, which consist of pure samples of non-collisional muons. The triggers in these datasets ensure that no beam arrives simultaneously with muons, creating therefore a sample enriched with cosmic and beam halo muons. We found zero events surviving the event selection. Nevertheless, events are also vetoed if they contain muons that are back-to-back, a feature characteristic of cosmic muons. This requirement is:

- $\alpha > 2.8$ rad,

where α is the 3D angle between the momentum vectors of two muons.

In the event selection described so far, there has yet to be use of the *displacement* of the dimuon vertex (except for selections on each track d_{xy}) or the *collimation* between the pair of muons and $p_{\mathrm{T}}^{\mathrm{miss}}$. These two key observables define the plane of signal extraction and data-driven background estimation, described in Chapter 10.

Fig. 9.10 depicts a "cutflow" plot of the event selection. Overall, the selection is successful in reducing background rates to a manageable level while keeping as much signal as possible. The sensitivity will be further improved inside the 2D plane formed by vertex displacement and $p_{\mathrm{T}}^{\mathrm{miss}}$-dimuon collimation since these are particularly sensitive and discriminating variables to the iDM signature.

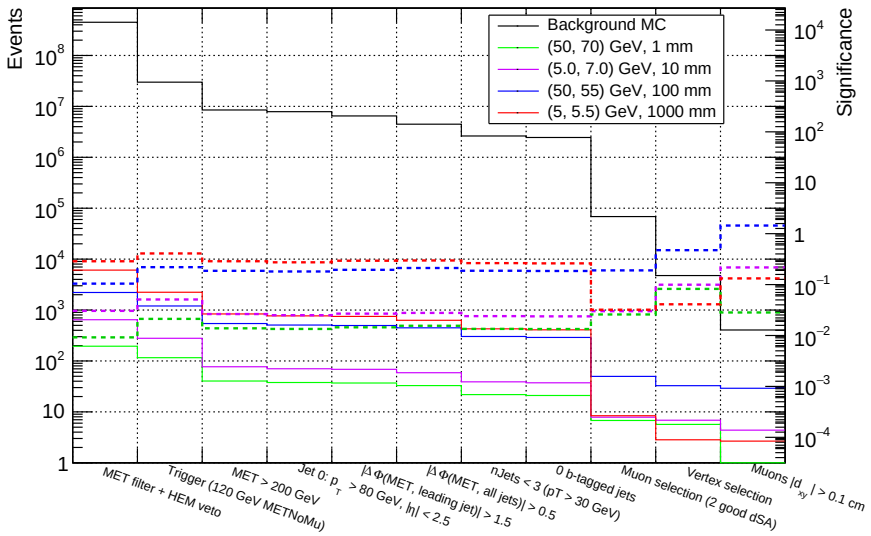


Figure 9.10: "Cutflow" plot of the event selection. Solid curves represent yields (black for the sum of backgrounds and colored for different signal samples) and dashed curves significance $s/\sqrt{b}$.

Other selections

For completeness and reference, I list other selections studied or that have potential for future enhancements:

- trigger acceptance;
- offline $p_{\mathrm{T}}^{\mathrm{miss}}$ cut;
- dSA muons (especially for low mass-splitting samples);
- Muon d_{xy} (especially for low-displacement samples).

It is possible that a custom trigger designed specifically for the iDM signature could help increase signal efficiency in the next run of the LHC. One possibility is a combination of $p_{\mathrm{T}}^{\mathrm{miss}}$ and at least one soft muon in collimation. Detailed studies are warranted to

understand if a reasonable trigger rate is achievable and whether iDM's soft muons would be sensitive to such a trigger.

An additional attempted improvement is the lowering of the 200 GeV offline $p_{\mathrm{T}}^{\mathrm{miss}}$ threshold. This could be advantageous because the distribution of $p_{\mathrm{T}}^{\mathrm{miss}}$ in signal events is overall quite soft. Unfortunately, lowering the threshold also accepts more background events and specifically raises the likelihood of events with mis-reconstructed $p_{\mathrm{T}}^{\mathrm{miss}}$, which are not well modeled in simulation. A further wrinkle is that the agreement between the $p_{\mathrm{T}}^{\mathrm{miss}}$ trigger response in data and in simulation is worsened by lower thresholds (this can be corrected to some extent with a better modeling of the trigger turn-on curve, which was implemented for this analysis). The primary issue with lowering the threshold, however, is the introduction of unwanted correlation between the two key observables used in the background estimation. The composition of events with lower $p_{\mathrm{T}}^{\mathrm{miss}}$ is not well understood and therefore the higher acceptance was not worth the spoiled independence between the observables.

A final optimization attempt was the lowering of the p_{T} threshold in the dSA muon identification criteria. The analysis requires a 5 GeV minimum muon p_{T}, but owing to the soft distribution of iDM muons, this selection removes a non-negligible fraction of signal events (particularly for lower mass-splitting samples, as shown in Fig. 9.10). Lowering this threshold to 3 GeV would increase the signal acceptance two-fold. However, we found that it also leads to a much larger fraction of misidentified muons. Furthermore, the displaced standalone reconstruction has not been carefully studied for such low-momentum muons. The combination of very displaced and very soft muons would need a more thorough assessment before being employed. This might be feasible in the next iteration of the analysis if a dedicated study is done to understand the dSA performance in these conditions.

9.4 Multijet validation region

A multijet validation region (VR) is defined by inverting the requirement on number of jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$. This VR, labeled "nJets", is particularly useful because it contains very similar physics to that in the SR. The main expected backgrounds include real or misidentified muons from jets plus $p_{\mathrm{T}}^{\mathrm{miss}}$ in the event. QCD and Z/W+jets processes with several jets are therefore primary sources of backgrounds (but the angular isolation requirement between jets and $p_{\mathrm{T}}^{\mathrm{miss}}$ mitigates these, as discussed in Section 9.2.1). Moreover, backgrounds are sharply different across match categories, as illustrated in Table 9.7. Since QCD contains real muons, events fall largely in the 1 and 2 GM-dSA match categories. In contrast, the 0 GM-dSA match category is dominated by misidentified muon events from Z/W+jets, which typically have no GM counterpart.

The main selection defining the nJets VR is:

- nJets ($p_{\mathrm{T}} > 30\,\mathrm{GeV}$) ≥ 3.

This VR is further enlarged to boost statistics by loosening some requirements relative to the main event selection. This puts the nJets VR in statistical parity with the SR so that it can appropriately be used to assess independence between the two observables in the data-driven background estimation, and to provide a reliable measure of its closure. The loosened selections are:

- $\left|\Delta\phi(p_{\mathrm{T}}^{\mathrm{miss}}, \text{leading jet})\right| > 1.0$ (*vs.* 1.5 in the SR);
- $dR(\text{muons}) < 1.5$ (*vs.* 0.9 in the SR);
- No dimuon invariant mass requirement (*vs.* $M_{\mu\mu} < 30\,\mathrm{GeV}$ in the SR).

These selections do not change the physics content of the validation region. Their main purpose in the SR is to optimize the signal-to-background sensitivity by removing from consideration regions of phase space where no signal is expected. To further validate this claim, we compare the shapes of two key observables between the enlarged and the original VR. Fig. 9.11 shows the profiles of the vertex v_{xy} significance (left) and $\Delta\phi(p_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{muons})$ (right) for both VRs. The two curves have the same shape in each case.

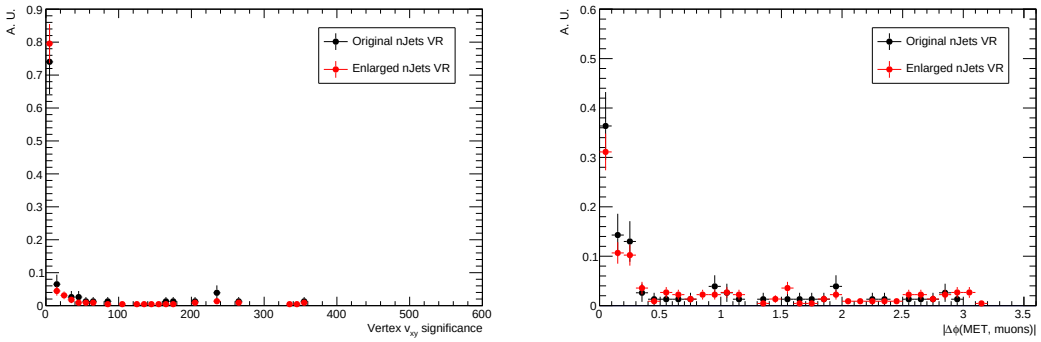


Figure 9.11: Comparison between enlarged and original nJets validation regions across two key observables used in the data-driven background estimation. Left: vertex v_{xy} significance. Right: $\Delta\phi(p_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{muons})$.

Validation regions must have relatively low signal contamination compared to the SR to avoid accidental unblinding of data sensitive to signal. Table 9.8 shows a comparison of total simulated background yields and four representative signal yields in the nJets VR, confirming that this is indeed the case here. Table 9.9 breaks down the background contributions to highlight the similar composition between this validation region and the signal region, reported in Table 9.7. The only difference is a slight increase in the overall top quark yield in the validation region, but since it is spread out fairly evenly between the three match categories, the relative contribution compared to the dominant backgrounds in each category is small.

The negligible contamination indicates that it is safe to look at data. Fig. 9.12 compares simulated distributions with observed data across different observables, before

Table 9.8: (nJets) Simulated yields for background and 4 representative signal samples in orthogonal multijet validation region after imposing d_{xy} requirements, using samples from all three years scaled to each year's luminosity (137.19 /fb in total).

Background MC	(50, 55) GeV, 100 mm	(5, 5.5) GeV, 1000 mm	(50, 70) GeV, 1 mm	(5, 7) GeV, 10 mm
524.2	12.2	1.9	0.3	1.6

Table 9.9: (nJets) Breakdown of the primary background yields in the multijet validation region, before and after imposing the impact parameter displacement selection, and split into the three GM-dSA match categories, according to simulation. Yields are scaled to each year's luminosity for a total of 137.2 fb^{-1}. Top backgrounds include $\text{t}\bar{\text{t}}$ and single top quark processes.

	Total	QCD	Top	Z+jets	W+jets
before $\lvert d_{xy}\rvert > 1$ mm	6846.9	3446.6	2041.6	1169.0	189.7
after $\lvert d_{xy}\rvert > 1$ mm	524.2	300.7	50.5	90.8	82.2
0 GM-dSA	207.1	36.7	16.3	78.8	75.3
1 GM-dSA	241.2	195.0	27.7	11.9	6.7
2 GM-dSA	76.0	69.0	6.5	0.2	0.2

requiring both muons to be displaced. Fig. 9.13 shows additionally the two observables that are later used for the data-driven background estimation: dimuon vertex displacement, v_{xy}, and azimuthal separation between $\mathbf{p}_{\text{T}}^{\text{miss}}$ and muons, $\Delta\phi(p_{\text{T}}^{\text{miss}}, \text{muons})$.

The same observables are shown in Figs. 9.14 and 9.15, respectively, after demanding $d_{xy} > 1$ mm of both muons. The displacement requirement suppresses the majority of electroweak backgrounds and further leads to a slight discrepancy between simulated samples and data in the tails of some kinematic observables. Therefore, as a caveat, we do not expect a perfect agreement between data and simulation, which motivates the data-driven background estimation strategy. Nevertheless, the agreement seen here is robust enough to inspire confidence in a proper understanding of the event selection and collisional backgrounds present in data. It also permits the use of simulated background samples for normalization purposes, a procedure described in Chapter 10.

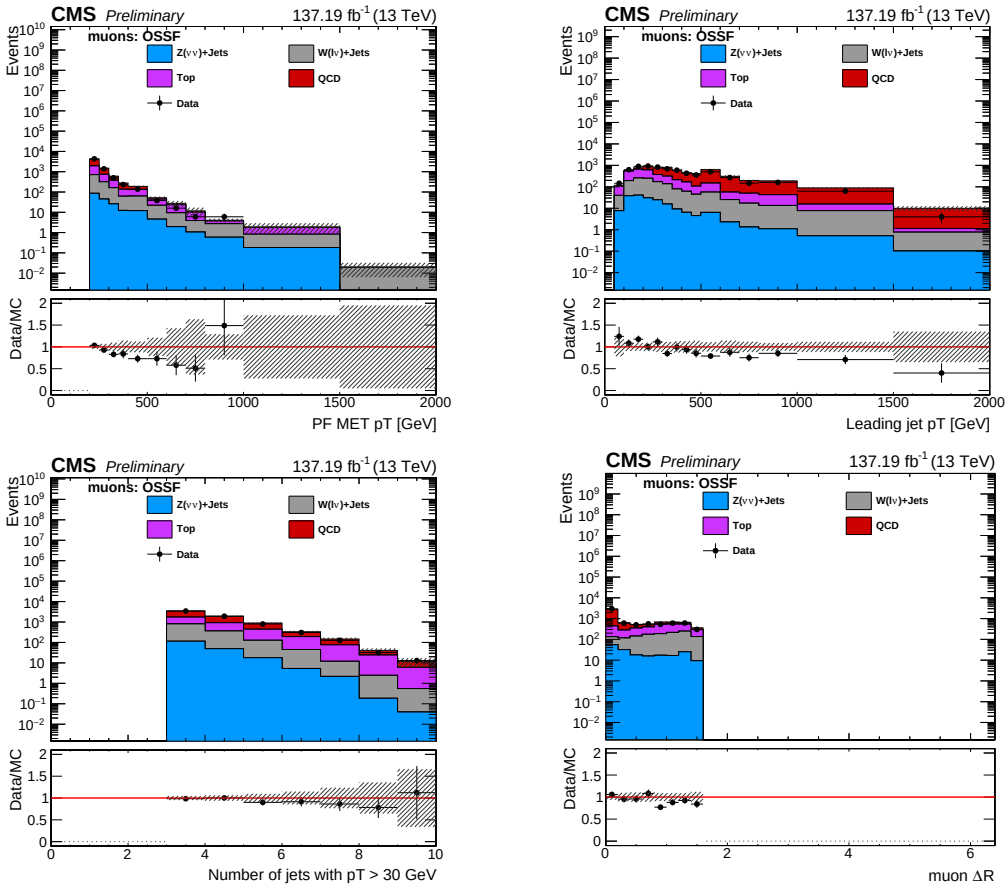


Figure 9.12: (nJets) Comparison of several observables between simulation and data in the multijet validation region after final selection *except for* requirements on displaced muon impact parameters. Upper left: $p_{\mathrm{T}}^{\mathrm{miss}}$; upper right: leading jet p_{T}; lower left: number of jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$; and lower right: dimuon *dR*. Gray bands indicate statistical uncertainties on simulated background samples.

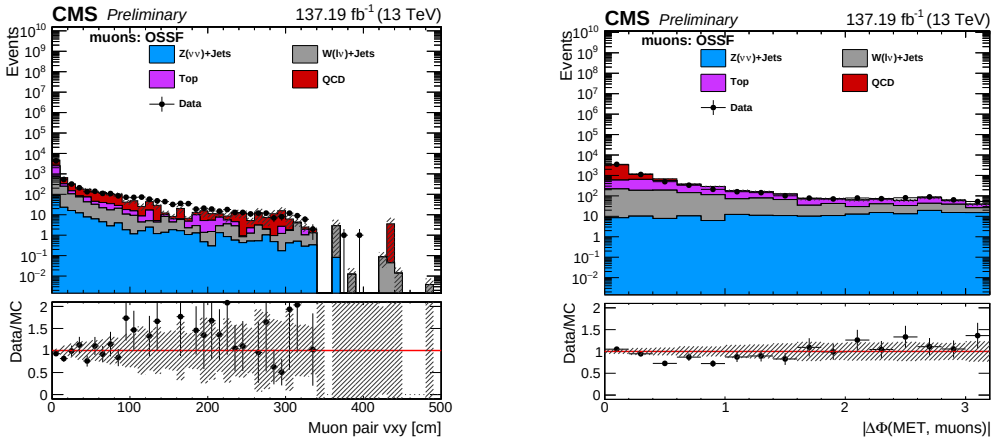


Figure 9.13: (nJets) Comparison of two key observables between simulation and data in the multijet validation region after final selection *except for* requirements on displaced muon impact parameters. Left: muon pair v_{xy}; and right: $|\Delta\phi(p_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{muons})|$. Gray bands indicate statistical uncertainties on simulated background samples. These observables are used later in the data-driven background estimation procedure.

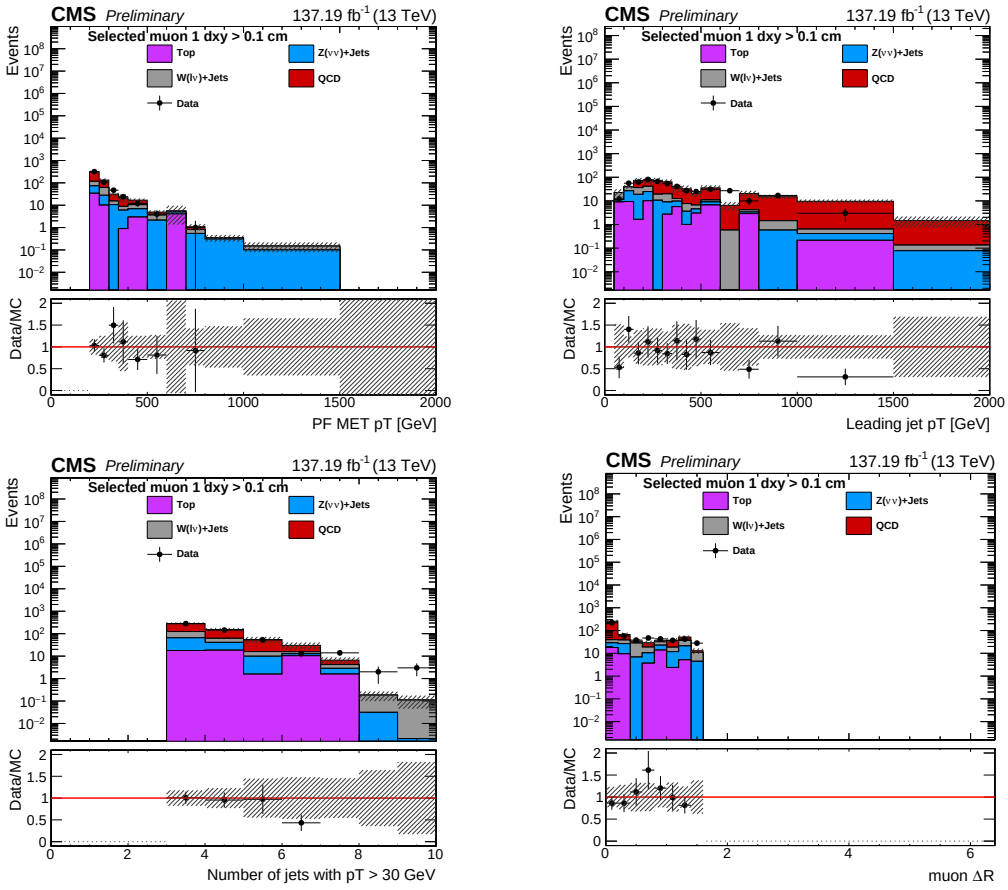


Figure 9.14: (nJets) Comparison of several observables between simulation and data in the multijet validation region after final selection *including* requirements on displaced muon impact parameters. Upper left: $p_{\mathrm{T}}^{\mathrm{miss}}$; upper right: leading jet p_{T}; lower left: number of jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$; and lower right: dimuon *dR*. Gray bands indicate statistical uncertainties on simulated background samples.

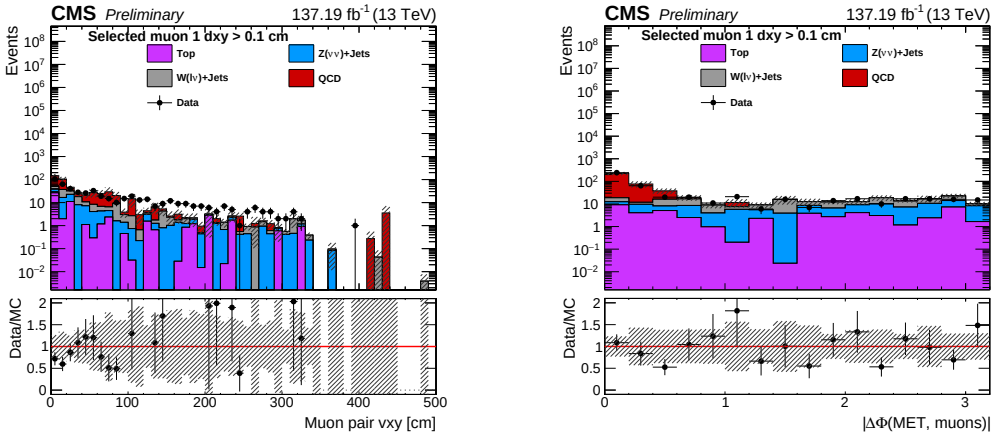


Figure 9.15: (nJets) Comparison of two key observables between simulation and data in the multijet validation region after final selection *including* requirements on displaced muon impact parameters. Left: muon pair v_{xy}; and right: $|\Delta\phi(p_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{muons})|$. Gray bands indicate statistical uncertainties on simulated background samples. These observables are used later in the data-driven background estimation procedure.

9.5 "Fake" muon validation region

We define a "fake" muon validation region by inverting the muon collimation and charge parity requirements:

- $dR(\text{muons}) > 0.9$;
- charge(muon1) = charge(muon2).

Since muons in these background events are not real (in the sense that they do not arise from the hard scattering), there should be no difference whether they have the same charge or not. Similarly, there is no angular correlation between them (and between the muons and $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$, as I show later). Therefore the inversion of these two requirements defines a validation region enriched with Z/W+jets events containing misidentified muons. This VR is labeled "dR".

The dR VR is also further enlarged by loosening the cuts:

- $|\Delta\phi(p_{\mathrm{T}}^{\mathrm{miss}}, \text{leading jet})| > 1.0$ (*vs.* 1.5 in the SR);
- No cut on the dimuon invariant mass (*vs.* $M_{\mu\mu} < 30\,\text{GeV}$ in the SR).

Table 9.10 confirms negligible signal contamination in this validation region by comparing the total simulated background yields to signal yields. This is expected since muons are required *not* to be collimated and to have the same charge. Table 9.11 breaks down the background events into their components, showing how this validation region is dominated by misidentified muons from Z/W+jets events.

Most events with misidentified muons fail the GM-dSA match procedure and therefore the 0 GM-dSA category consists predominantly of such events. Figs. 9.16 and 9.17

Table 9.10: (dR) Simulated yields for background and 4 representative signal samples in orthogonal validation region dR after imposing d_{xy} requirements, using samples from all three years scaled to each year's luminosity (137.2 fb^{-1} in total).

Background MC	(50, 55) GeV, 100 mm	(5, 5.5) GeV, 1000 mm	(50, 70) GeV, 1 mm	(5, 7) GeV, 10 mm
908.6	1.7	0.3	0.0	0.2

Table 9.11: (dR) Breakdown of the primary background simulation yields in the dR validation region before and after imposing the impact parameter displacement requirement and split into the three GM-dSA match categories. Yields are scaled to each year's luminosity for a total of 137.2 fb^{-1}. Top backgrounds include $t\bar{t}$ and single top quark processes.

	Total	QCD	Top	Z+jets	W+jets
before $\lvert d_{xy}\rvert > 1$ mm	11138.7	1876.0	1531.5	7154.5	576.7
after $\lvert d_{xy}\rvert > 1$ mm	908.6	163.1	42.2	379.8	323.5
0 GM-dSA	711.6	38.4	24.9	350.5	297.9
1 GM-dSA	197.0	124.7	17.3	29.3	25.7
2 GM-dSA	0.0	0.0	0.0	0.0	0.0

show a data-simulation comparison of observables in the dR VR before imposing the d_{xy} requirement, and Figs. 9.18 and 9.19 after.

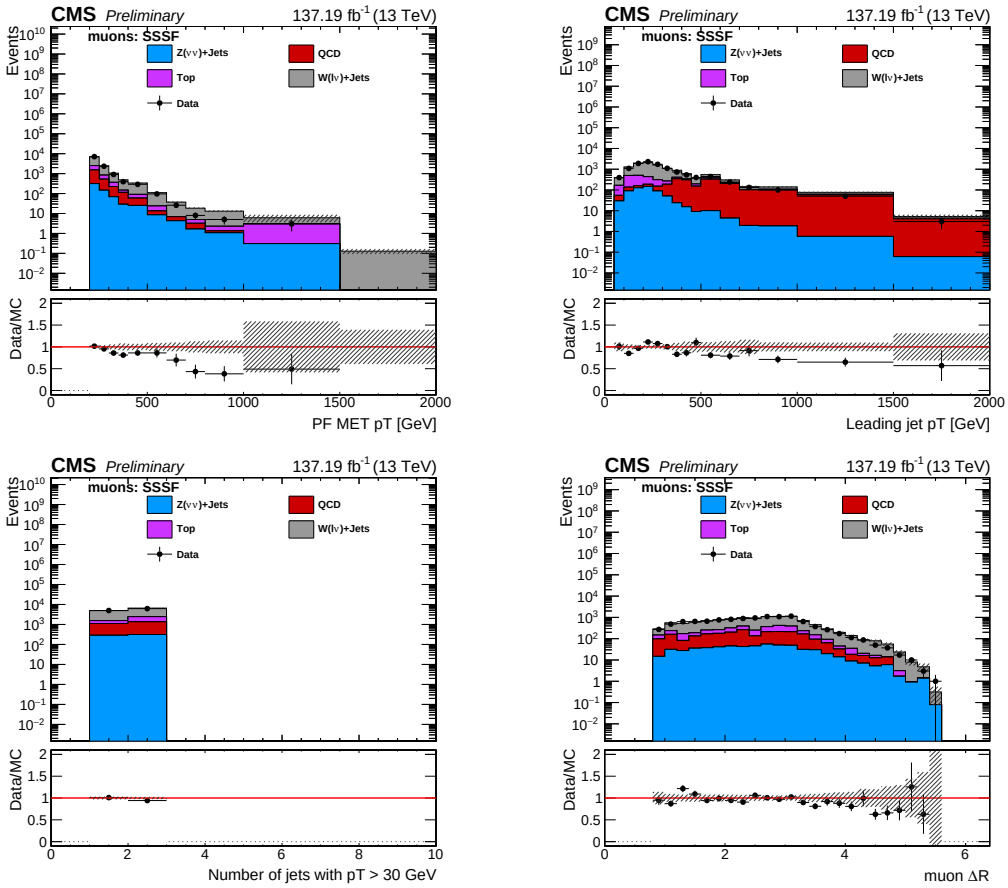


Figure 9.16: (dR) Comparison of several observables between simulation and data in the dR validation region after final selection *except for* requirements on displaced muon impact parameters. Upper left: $p_{\mathrm{T}}^{\mathrm{miss}}$; upper right: leading jet p_{T}; lower left: number of jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$; and lower right: dimuon *dR*. Gray bands indicate statistical uncertainties on simulated background samples.

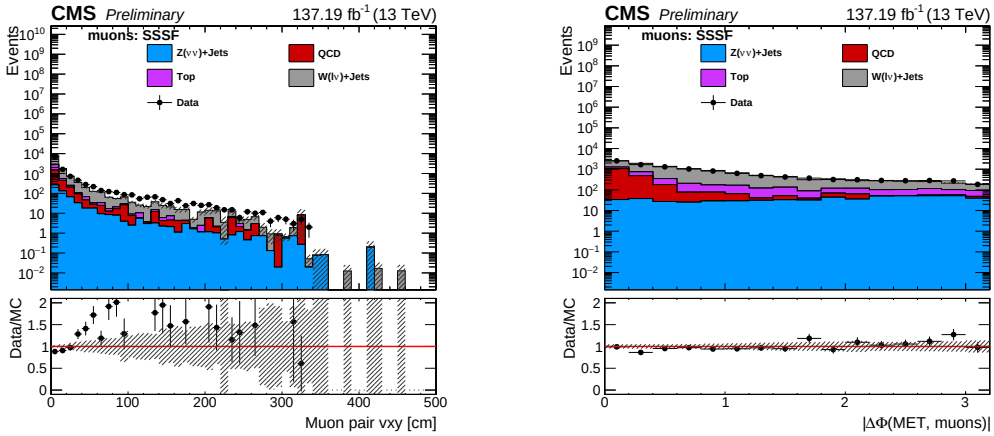


Figure 9.17: (dR) Comparison of two key observables between simulation and data in the dR validation region after final selection *except for* requirements on displaced muon impact parameters. Left: muon pair v_{xy}; and right: $|\Delta\phi(p_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{muons})|$. Gray bands indicate statistical uncertainties on simulated background samples. These observables are used later in the data-driven background estimation procedure.

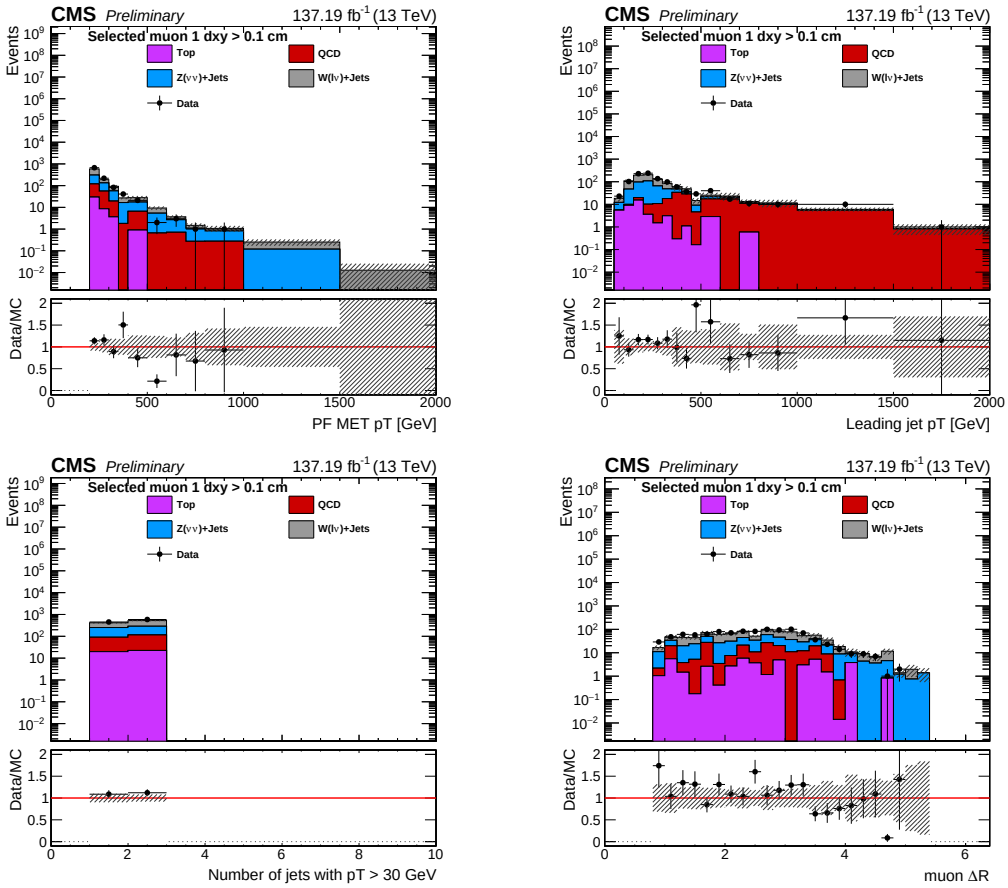


Figure 9.18: (dR) Comparison of several observables between simulation and data in the dR validation region after final selection *including* requirements on displaced muon impact parameters. Upper left: $p_{\mathrm{T}}^{\mathrm{miss}}$; upper right: leading jet p_{T}; lower left: number of jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$; and lower right: dimuon *dR*. Gray bands indicate statistical uncertainties on simulated background samples.

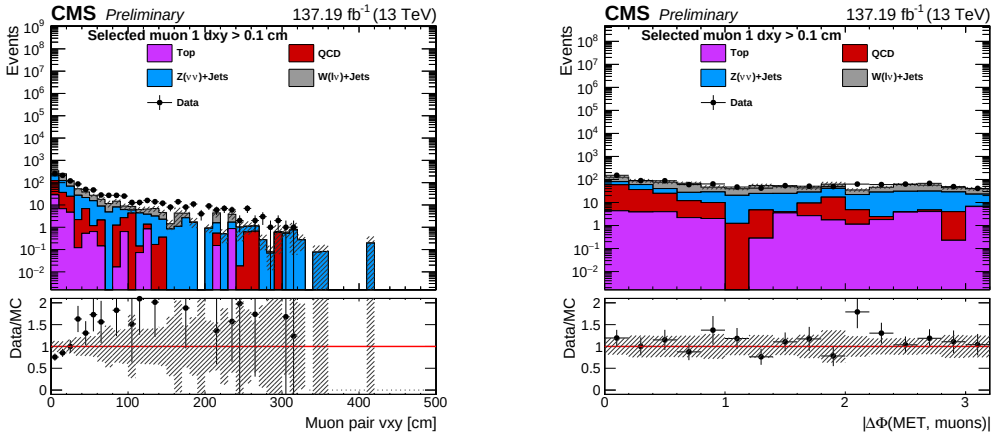


Figure 9.19: (dR) Comparison of two key observables between simulation and data in the dR validation region after final selection *including* requirements on displaced muon impact parameters. Left: muon pair v_{xy}; and right: $|\Delta\phi(p_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{muons})|$. Gray bands indicate statistical uncertainties on simulated background samples. These observables are used later in the data-driven background estimation procedure.

CHAPTER 10

BACKGROUND ESTIMATION

This chapter presents a strategy for a data-driven background estimation with a modified ABCD method. This data-driven approach is needed as a considerable fraction of expected backgrounds are not well modeled in simulation. Furthermore, QCD, the largest SM background, is also difficult to simulate because of its large cross section and high jet multiplicity.

10.1 The ABCD method

The idea behind the classic ABCD method is to pick two uncorrelated observables and plot their 2D distributions in data after the full selection is applied. The 2D plane is split into four bins (A, B, C, and D) and the bin with highest signal yield chosen as the signal bin. In the traditional formulation, other bins have relatively low signal contamination and consist mostly of background events. This situation is sketched in Fig. 10.1.

To estimate backgrounds in the signal bin (region A), provided that the two observables are uncorrelated, it suffices to determine the ratio of observed yields in the other bins:

$$N_A^{\text{pred}} = \frac{N_B^{\text{obs}} \times N_C^{\text{obs}}}{N_D^{\text{obs}}}. \tag{10.1}$$

Since the variables are independent, the ratio between A and B yields is equal to the ratio between C and D, and similarly for the ratios A/C and B/D.

In the case of inelastic dark matter, the dimuon vertex displacement (v_{xy}) and its absolute angular separation from $p_{\text{T}}^{\text{miss}}$ ($|\Delta\phi_{\text{MM}}| \equiv |\Delta\phi(p_{\text{T}}^{\text{miss}}, \text{dimuon})|$) may be used

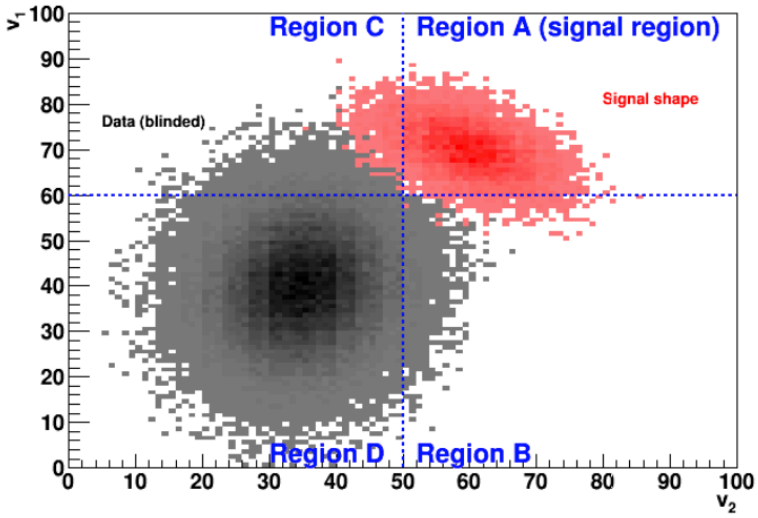


Figure 10.1: Sketch of the classic ABCD method for an arbitrary idealized situation with two independent variables v_1 and v_2. The signal distribution is concentrated in region A, while the background distribution is centered in region D. The independence between v_1 and v_2 allows the estimation of the background in A without unblinding the data there, by using the yields observed in regions B, C, and D [110].

as the two ABCD observables. From a physics standpoint, these two variables should be fully uncorrelated, but the assumption needs to be validated in data. To do so, we checked ABCD closure in the validation regions orthogonal to the signal region defined in Chapter 9. This is done by comparing the predicted yield N_A^{pred} in Eq. (10.1) with the actual observed yield N_A^{obs}.

The ABCD regions in iDM are defined slightly differently, as sketched in Fig. 10.2. The signal-enriched bin is labeled C on the 2D plane. The exact boundaries for each region are variable and depend on the signal hypothesis for maximum sensitivity, as discussed in Section 10.5. Furthermore, we employ a modified version of the ABCD method which fits all 4 bins simultaneously and therefore provides a robust background estimate even in the presence of signal contamination.

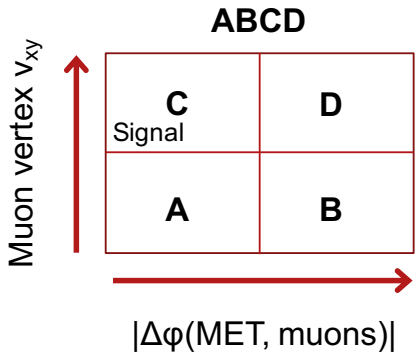


Figure 10.2: ABCD setup in the specific case of iDM. The independent variables are the $p_{\mathrm{T}}^{\mathrm{miss}}$-dimuon collimation $|\Delta\phi_{\mathrm{MM}}|$ and dimuon vertex v_{xy}. The region of highest signal sensitivity is C, at large vertex displacement and narrow collimation between $p_{\mathrm{T}}^{\mathrm{miss}}$ and muons.

10.2 Closure of ABCD in nJets VR

Fig. 10.3 shows 2D histograms of muon v_{xy} *vs.* $|\Delta\phi_{\mathrm{MM}}|$ in data, for the 0, 1, and 2 GM-dSA match categories. At first glance, the two observables do look uncorrelated. The independence is tested by defining several ABCD regions inside the 2D plane and comparing the predicted C yield with the actual observed yield in each case.

Several combinations of bin boundaries were tested for the closure validation. The results are reported in Tables 10.1 to 10.3, which list yields per bin and compare expected and observed yields in the signal-enriched C bin for the three match categories respectively. The closure does seem robust.

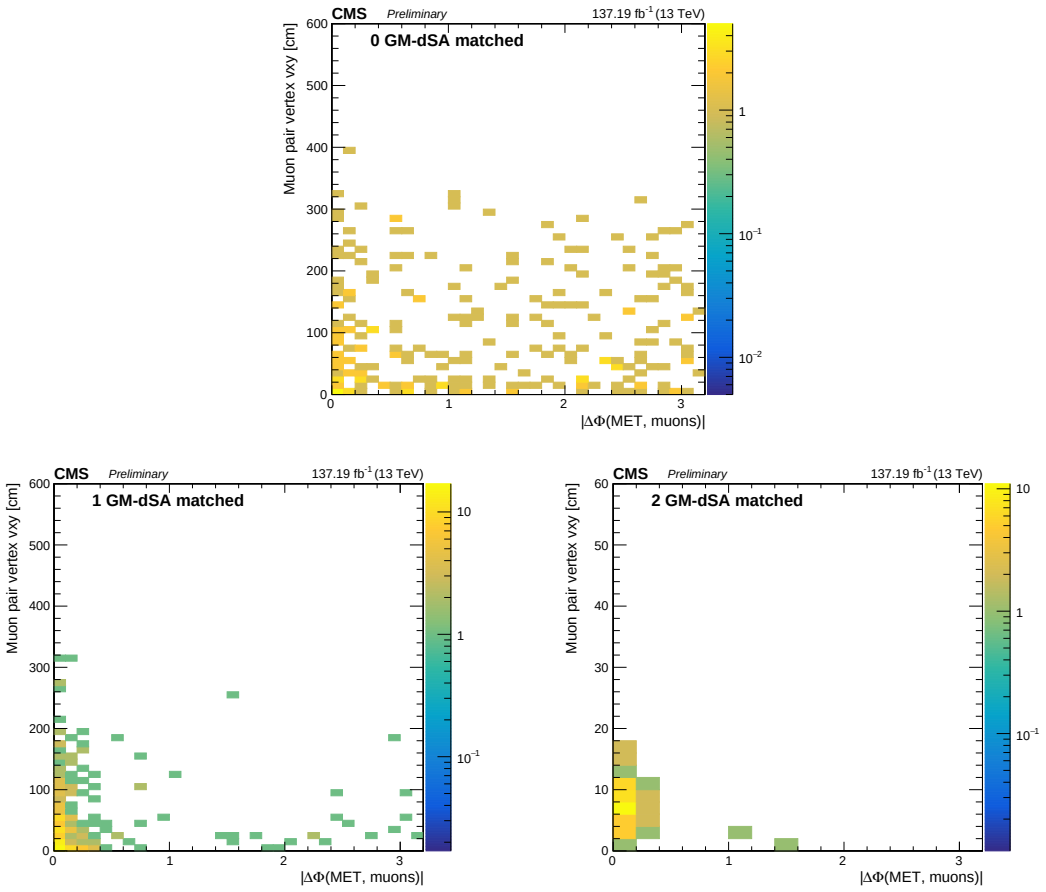


Figure 10.3: (nJets) Yields in 2D plane of v_{xy} *vs.* $|\Delta\phi_{\mathrm{MM}}|$, for the multijet VR, after final selection and split into the three matching categories: (upper center) 0 GM-dSA, (lower left) 1 GM-dSA, and (lower right) 2 GM-dSA. The last plot is zoomed in to the displacement region between 0 and 60 cm, since global muons cannot be very displaced.

Table 10.1: (nJets, 0 GM-dSA) Closure tests of the ABCD method in the multijet VR, for several choices of ABCD bin edges. These tests are performed on the 0 GM-dSA match category (i.e. upper plot in Fig. 10.3). The first and second pair of edges define the $|\Delta\phi_{\mathrm{MM}}|$ and v_{xy} (cm) bins respectively. The uncertainty on the predicted C yield is computed from propagation of the Poisson statistics uncertainties on A, B, and D.

Bin edges	A	B	D	Predicted C	Observed C
{0, 0.2}, {0, 80}	25	102	91	22.3 ± 5.5	27 ± 5.2
{0, 0.2}, {0, 100}	29	109	84	22.3 ± 5.3	23 ± 4.8
{0, 0.2}, {0, 120}	34	120	73	20.7 ± 4.7	18 ± 4.2
{0, 0.1}, {0, 80}	17	110	101	15.6 ± 4.4	17 ± 4.1
{0, 0.1}, {0, 100}	21	117	94	16.9 ± 4.4	13 ± 3.6
{0, 0.1}, {0, 120}	24	130	81	15.0 ± 3.7	10 ± 3.2
{0, 0.2}, {50, 140}	17	60	61	17.3 ± 5.2	17 ± 4.1
{0, 0.1}, {20, 80}	10	64	101	15.8 ± 5.6	17 ± 4.1
{0, 0.1}, {30, 100}	11	57	94	18.1 ± 6.3	13 ± 3.6
{0, 0.1}, {50, 120}	11	53	81	16.8 ± 5.9	10 ± 3.2

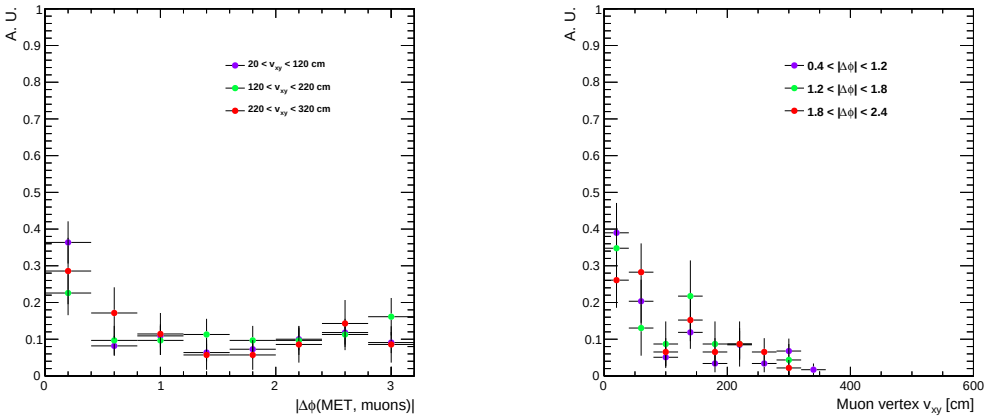


Figure 10.4: (nJets, 0 GM-dSA) 1D horizontal and vertical profiles for the two ABCD variables in the 0 GM-dSA match category for the multijet VR. The profiles are sliced several times to ensure a consistent behavior throughout the entire 2D plane and validate the assumption of independence between the two variables.

Table 10.2: (nJets, 1 GM-dSA) Closure tests of the ABCD method in the multijet VR, for several choices of ABCD bin edges. These tests are performed on the 1 GM-dSA match category (i.e. lower left plot in Fig. 10.3). The first and second pair of edges define the $|\Delta\phi_{\mathrm{MM}}|$ and v_{xy} (cm) bins respectively. The uncertainty on the predicted C yield is computed from propagation of the Poisson statistics uncertainties on A, B, and D. Some identical matching yields when splitting horizontally at $|\Delta\phi_{\mathrm{MM}}| = 0.1$ have been double-checked and verified to be accidental.

Bin edges	A	B	D	Predicted C	Observed C
{0, 0.2}, {0, 80}	105	49	19	40.7 ± 11.7	47 ± 6.9
{0, 0.2}, {0, 100}	115	53	15	32.5 ± 10.0	37 ± 6.1
{0, 0.2}, {0, 120}	127	57	11	24.5 ± 8.4	25 ± 5.0
{0, 0.1}, {0, 80}	77	77	33	33.0 ± 7.8	33 ± 5.7
{0, 0.1}, {0, 100}	84	84	26	26.0 ± 6.5	26 ± 5.1
{0, 0.1}, {0, 120}	92	92	18	18.0 ± 5.0	18 ± 4.2
{0, 0.2}, {50, 140}	53	15	8	28.3 ± 13.0	20 ± 4.5
{0, 0.1}, {20, 80}	46	41	33	37.0 ± 10.2	33 ± 5.7
{0, 0.1}, {30, 100}	45	33	26	35.5 ± 10.7	26 ± 5.1
{0, 0.1}, {50, 120}	35	25	18	25.2 ± 8.9	18 ± 4.2

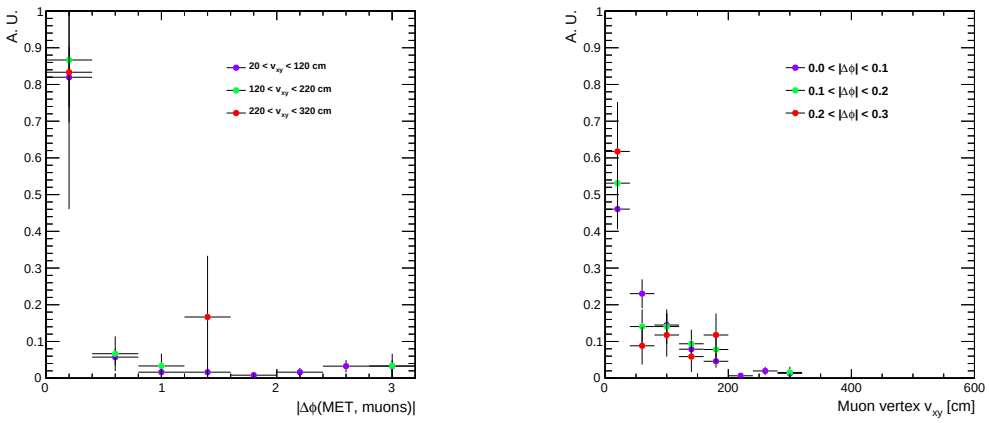


Figure 10.5: (nJets, 1 GM-dSA) 1D horizontal and vertical profiles for the two ABCD variables in the 1 GM-dSA match category for the multijet VR. The profiles are sliced several times to ensure a consistent behavior throughout the entire 2D plane and validate the assumption of independence between the two variables.

Table 10.3: (nJets, 2 GM-dSA) Closure tests of the ABCD method in the multijet VR, for several choices of ABCD bin edges. These tests are performed on the 2 GM-dSA match category (i.e. lower right plot in Fig. 10.3). The first and second pair of edges define the $|\Delta\phi_{\mathrm{MM}}|$ and v_{xy} (cm) bins respectively. The uncertainty on the predicted C yield is computed from propagation of the Poisson statistics uncertainties on A, B, and D.

Bin edges	A	B	D	Predicted C	Observed C
{0, 0.15}, {0, 4}	6	1	7	42.0 ± 48.1	34 ± 5.8
{0, 0.15}, {0, 6}	11	3	5	18.3 ± 14.5	29 ± 5.4
{0, 0.15}, {0, 8}	22	5	3	13.2 ± 10.0	18 ± 4.2
{0, 0.15}, {0, 10}	29	7	1	4.2 ± 4.5	11 ± 3.3

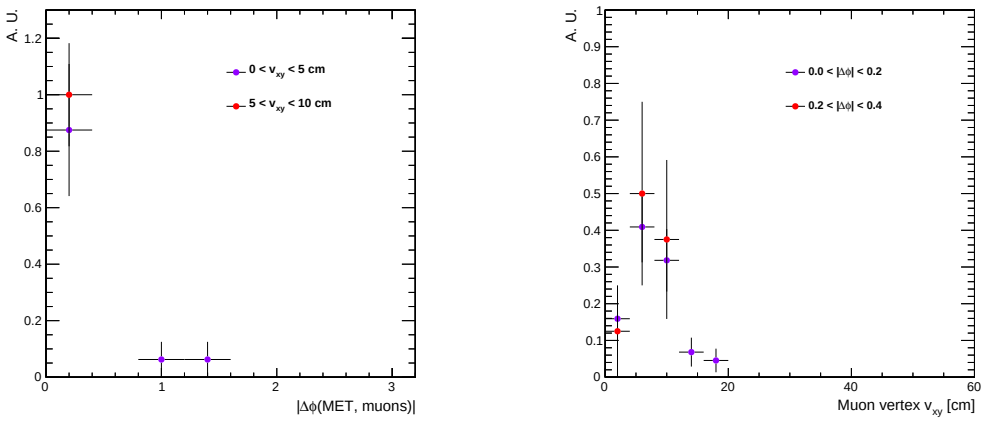


Figure 10.6: (nJets, 2 GM-dSA) 1D horizontal and vertical profiles for the two ABCD variables in the 2 GM-dSA match category for the multijet VR. The profiles are sliced several times to ensure a consistent behavior throughout the entire 2D plane and validate the assumption of independence between the two variables.

10.3 Closure of ABCD in dR VR

We also report closure tests in the dR validation region. Here only the 0 GM-dSA match category (shown in Fig. 10.7) is tested since the vast majority of relevant events fall in this category. This is by design, given that the VR is specifically enriched with events containing misidentified muons in Z/W+jets processes. Table 10.4 and Fig. 10.8 prove the closure of the background estimation in the dR VR as well.

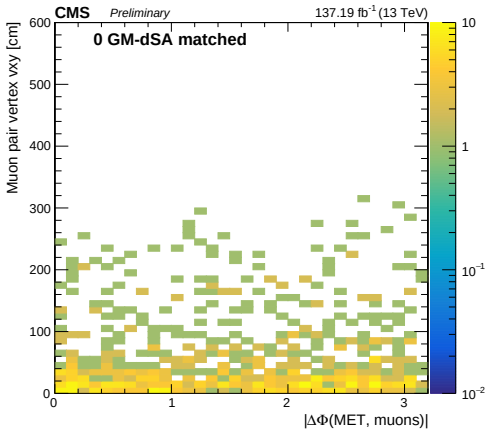


Figure 10.7: (dR, 0 GM-dSA) Yields in 2D plane of v_{xy} *vs.* $|\Delta\phi_{\mathrm{MM}}|$, for the dR VR, after final selection and in the 0 GM-dSA.

Table 10.4: (dR, 0 GM-dSA) Closure tests of the ABCD method in the dR VR, for several choices of ABCD bin edges. These tests are performed on the 0 GM-dSA match category (i.e. Fig. 10.7). The first pair of edges define the $|\Delta\phi_{\mathrm{MM}}|$ bins, and the second the v_{xy} (cm) bins. The uncertainty on the predicted C yield is computed from propagation of the Poisson statistics uncertainties on A, B, and D.

Bin edges	A	B	D	Predicted C	Observed C
{0, 0.5}, {20, 100}	67	292	114	26.2 ± 4.3	24 ± 4.9
{0, 1.5}, {20, 100}	166	193	72	61.9 ± 9.8	66 ± 8.1
{0, 2.5}, {20, 100}	279	80	31	108.1 ± 23.8	107 ± 10.3
{0.2, 1.5}, {0, 50}	205	275	147	109.6 ± 13.6	102 ± 10.1
{0.2, 1.5}, {0, 100}	252	350	72	51.8 ± 7.5	55 ± 7.4
{0.2, 1.5}, {0, 200}	289	401	21	15.1 ± 3.5	18 ± 4.2

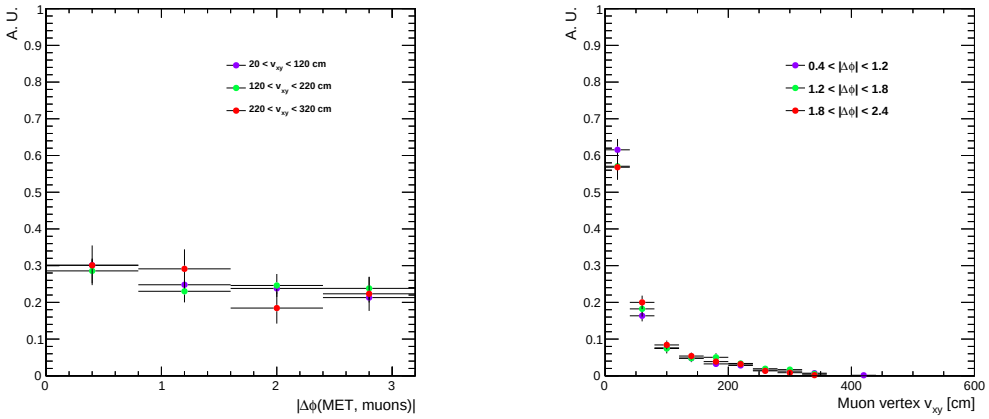


Figure 10.8: (dR, 0 GM-dSA) 1D horizontal and vertical profiles for the two ABCD variables in the 0 GM-dSA match category for the dR VR. The profiles are sliced several times to ensure a consistent behavior throughout the entire 2D plane and validate the assumption of independence between the two variables.

10.4 Initial boundary optimization of ABCD in SR with simulation

The next step after independence validation is to select optimal boundaries for ABCD estimation inside the signal region. In general, the optimal choice will depend on the signal hypothesis, since different dark matter masses and lifetimes lead to different distributions of the two observables. Hence it is advantageous to optimize the ABCD bins separately for each signal point in order to maximize sensitivity.

An initial optimization of ABCD bins with simulation was performed first. This strategy is not well-suited for the final optimization, since simulation cannot be fully relied on for a precise estimation of backgrounds. However, it is still useful to glean approximate information about where the final choice of boundaries should be, as well as validating the need for displacement-dependent bin boundaries.

The procedure consists of scanning across the bins of the 2D histogram formed by v_{xy} and $|\Delta\phi_{\mathrm{MM}}|$, for both signal and background samples, and computing the respective $s/\sqrt{b}$ significances of each boundary. A 2D map of the significance as a function of

boundary is constructed, and the boundaries with highest sensitivity per signal point are chosen. This procedure is sketched out in Fig. 10.9.

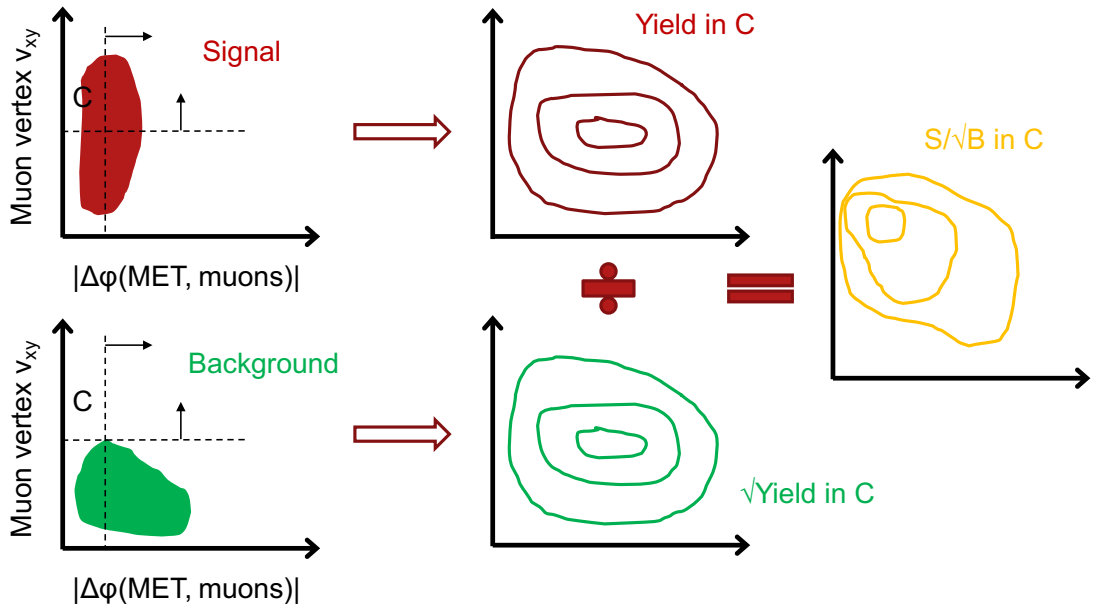


Figure 10.9: Sketch of preliminary sensitivity studies in simulation. As the bin edges are scanned (left), signal yields in the C bin are divided by the square root of background yields in the same region (center), leading to a map of the sensitivity $s/\sqrt{b}$ in C as a function of ABCD bin edges (right).

Fig. 10.10 shows bin optimization scans in simulated samples for four representative signal points. While the absolute scale of the significance is not meaningful since no uncertainties are included, it still illustrates that the optimal v_{xy} boundary has a strong dependence on the lifetime of the sample, while the $|\Delta\phi_{\mathrm{MM}}|$ boundary is more sensitive to the dark matter mass splitting. For larger splittings, a larger azimuthal separation between muons and $p_{\mathrm{T}}^{\mathrm{miss}}$ is expected, though the effect is not as strong. This is also seen from the generator-level plots in Fig. 5.10.

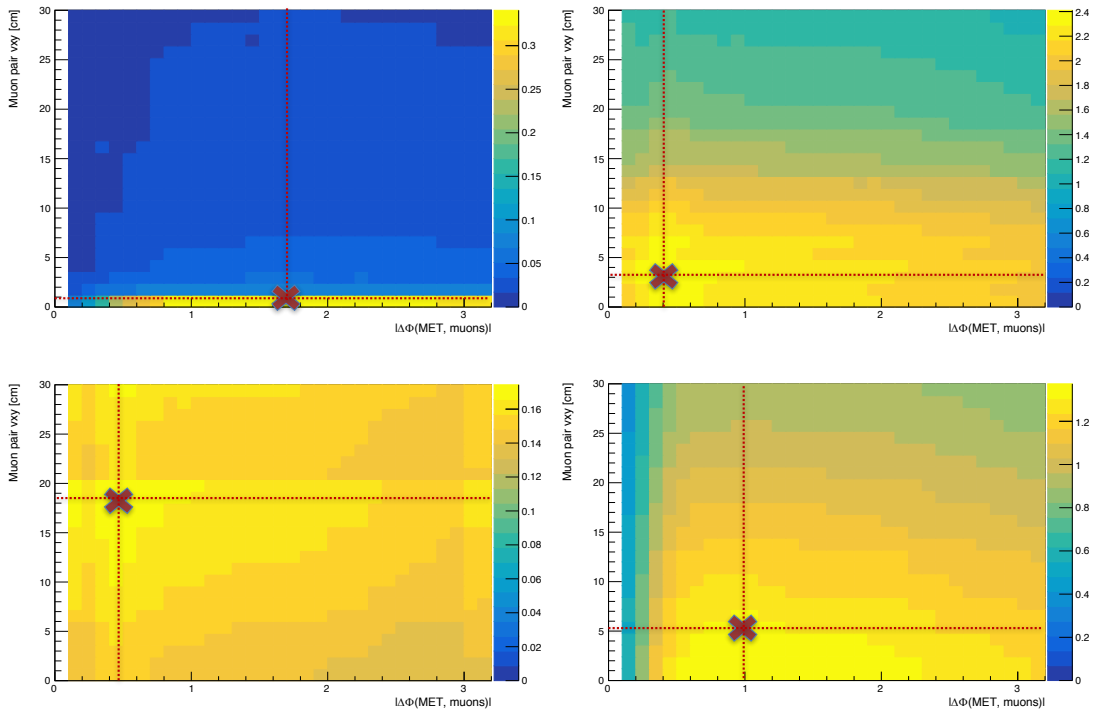


Figure 10.10: Scan of the $s/\sqrt{b}$ significance of ABCD bins for 4 representative signal samples. Clockwise from the upper left: $(m_1, m_2) = (50, 70)\,\text{GeV}$, $c\tau = 1\,\text{mm}$, $(m_1, m_2) = (5, 7)\,\text{GeV}$, $c\tau = 10\,\text{mm}$, $(m_1, m_2) = (50, 55)\,\text{GeV}$, $c\tau = 100\,\text{mm}$, and $(m_1, m_2) = (5, 5.5)\,\text{GeV}$, $c\tau = 1000\,\text{mm}$. The optimal choice of bin edge shifts up in v_{xy} with increased signal displacement, and to the right in $|\Delta\phi_{\text{MM}}|$ with increased absolute mass splitting.

10.5 Boundary optimization of ABCD in SR with data

The ultimate bin boundary optimization uses the data itself for guidance. In broad terms, the strategy is to observe the behavior of the discriminating variables in the multijet validation region, orthogonal to the signal region. As mentioned in Section 9.4, this VR differs from the SR in the number of jets with p_{T} greater than 30 GeV allowed per event. The VR requires at least 3 such jets, while the SR allows only 1 or 2 jets. This difference does not appreciably alter the background composition between the two regions, as evidenced by Tables 9.7 and 9.9. The similar composition enables the derivation of a normalization factor between the two regions by comparing the simulated yields and

applying it to the observed yield in data in the VR. This strategy has the advantage that it does not require unblinding any regions of the SR in data, so a fully blinded analysis can be carried out. This is especially important given the possible signal contamination in the SR for some dark matter mass hypotheses that receive large cross section enhancements near the Z resonance.

Although simulated backgrounds do not perfectly model the kinematics of events in data, they do seem to provide a roughly accurate total event yield based on validation region comparisons shown in Chapter 9. And even if this accuracy is only approximate, it should have little effect on the location of the optimal bin boundaries for each signal point, which is the purpose of this study. The predicted backgrounds here are only used for the bin optimization procedure. The final sensitivity will of course depend on the actual yields observed in all SR bins when the analysis is unblinded.

The motivation for this strategy is the fact that the multijet VR, while still orthogonal to the SR, is particularly "close" to it. This is true in the sense that the only difference between them is the number of jets allowed per event. Apart from events with zero jets (which are vetoed early in the event selection), both regions contain similar physics. This is particularly true after imposing a minimum requirement on the transverse impact parameter of each muon, which removes prompt electroweak backgrounds. Therefore, the background events that survive the final selection form essentially identical distributions in both SR and nJets VR. This is validated by studies of simulated backgrounds in both regions as reported in Chapter 9.

10.5.1 Templates from nJets VR

The nJets VR, being kinematically close to the SR, is used to study background shapes in both observables without fear of signal contamination. We build templates of the 1D background distributions (i.e. projections in both x and y of the 2D ABCD plane) from events in the nJets VR and use them to estimate the expected background in the SR. However, since the overall yields in the SR and VR are different, the simulated background yields are used to set the correct normalization from the template-derived background predictions. The procedure is sketched in Fig. 10.11.

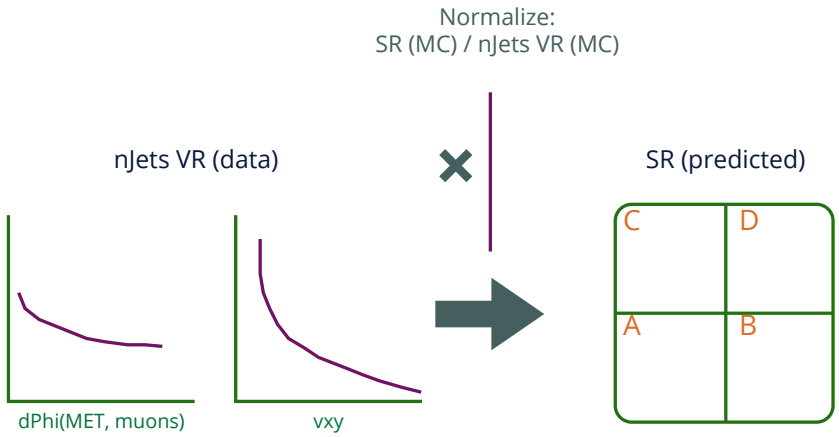


Figure 10.11: Sketch of the background estimation procedure in data using templates built from the nJets VR. The templates are used to estimate the background in the SR without unblinding. The normalization is derived from the ratio of the simulated background yields in the SR and the VR. The simulation has been shown in Chapter 9 to approximate the data in both dR and nJets validation regions sufficiently well for this purpose. Such normalization allows the entire SR to remain blinded, which is necessary to prevent accidental unblinding of signal hypotheses with large cross sections.

To demonstrate that simulated background samples are robust and approximate the shape of the 2D plane sufficiently well, Fig. 10.12 compares the 2D plane formed by v_{xy} and $|\Delta\phi_{\mathrm{MM}}|$ between the SR and nJets VR with simulation. The shapes of the two regions indeed look very similar and also match the shape of the distributions in nJets from data, reported in Fig. 10.3. The similarity is further evidenced by the breakdown of simulated

background yields in Tables 9.7 and 9.9. A final validation is found in Fig. 10.13, where the 1D profiles are plotted for each of the 2D plane variables. The shapes between the nJets VR and SR also seem to agree here.

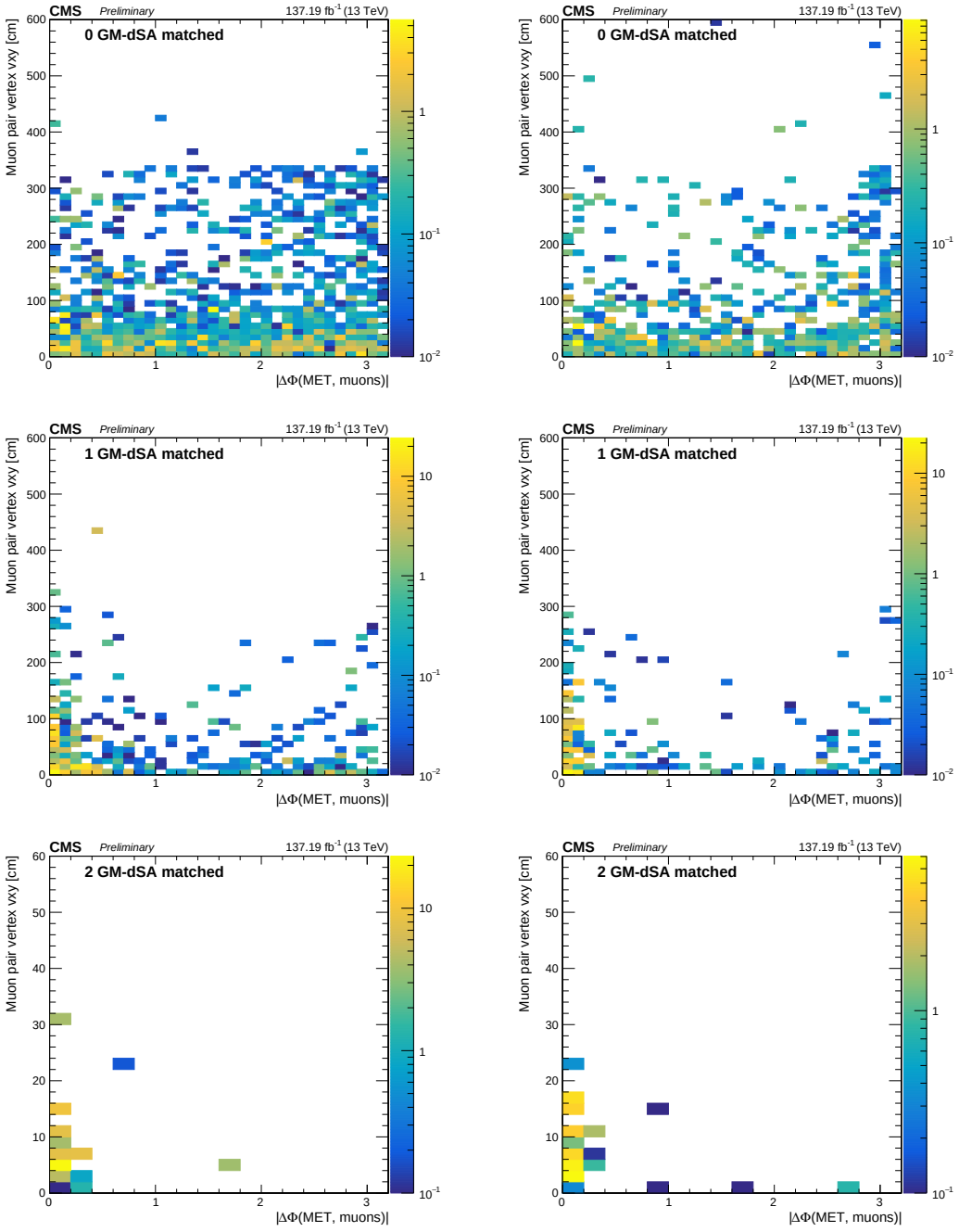


Figure 10.12: Comparison of 2D simulated background yields in nJets VR (left) and SR (right), for the three match categories: 0 GM-dSA (upper), 1 GM-dSA (middle), and 2 GM-dSA (lower). The similar 2D shapes seem to confirm the similar background composition in the two regions. Note also the overall increased statistics of the VR relative to the SR.

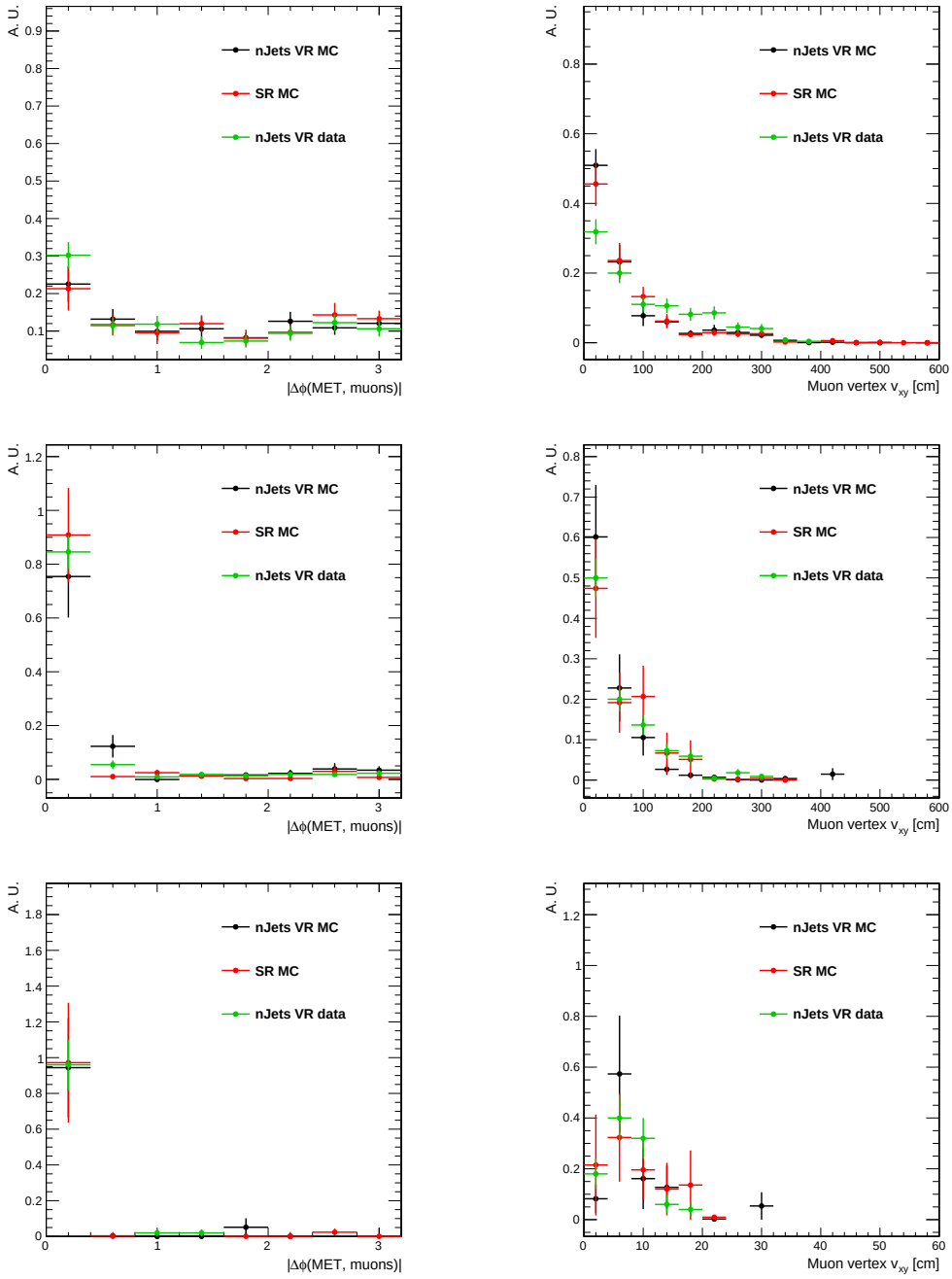


Figure 10.13: Comparison of $|\Delta\phi_{\mathrm{MM}}|$ templates (left) and v_{xy} templates (right) between nJets VR (data and simulation), and SR (simulation). Upper, middle, and lower plots feature 0, 1, and 2 GM-dSA categories respectively. The shapes are consistent throughout, which further confirm the similar background composition between the two regions.

10.5.2 SR background predictions from templates

The derived unit-normalized templates are employed to predict the background yields in the signal region. Defining normalized vertical and horizontal template functions as $F = F(v_{xy})$ and $G = G(\Delta\phi)$, respectively, the predicted SR background yields are:

$$
\begin{aligned}
N^{\mathrm{SR}}_{A,\mathrm{pred}} &= \frac{N^{\mathrm{SR}}_{\mathrm{MC}}}{N^{\mathrm{VR}}_{\mathrm{MC}}} \times \left(\int_0^x G(\Delta\phi)\, d\Delta\phi \right) \times \left(\int_0^y F(v_{xy})\, dv_{xy} \right) \\
N^{\mathrm{SR}}_{B,\mathrm{pred}} &= \frac{N^{\mathrm{SR}}_{\mathrm{MC}}}{N^{\mathrm{VR}}_{\mathrm{MC}}} \times \left(\int_x^\infty G(\Delta\phi)\, d\Delta\phi \right) \times \left(\int_0^y F(v_{xy})\, dv_{xy} \right) \\
N^{\mathrm{SR}}_{C,\mathrm{pred}} &= \frac{N^{\mathrm{SR}}_{\mathrm{MC}}}{N^{\mathrm{VR}}_{\mathrm{MC}}} \times \left(\int_0^x G(\Delta\phi)\, d\Delta\phi \right) \times \left(\int_y^\infty F(v_{xy})\, dv_{xy} \right) \\
N^{\mathrm{SR}}_{D,\mathrm{pred}} &= \frac{N^{\mathrm{SR}}_{\mathrm{MC}}}{N^{\mathrm{VR}}_{\mathrm{MC}}} \times \left(\int_x^\infty G(\Delta\phi)\, d\Delta\phi \right) \times \left(\int_y^\infty F(v_{xy})\, dv_{xy} \right)
\end{aligned}
\tag{10.2}
$$

where x and y edges define the ABCD bins and the template integrals for each bin are independent between the two observables. The data templates and ABCD formulas are built and applied separately to each match category because of the varying background composition.

10.5.3 Bin optimization from asymptotic discovery sensitivity

To optimize the ABCD bins using the template-based predicted backgrounds, the discovery sensitivity in the asymptotic approximation limit, based on Asimov datasets, is used as a proxy of optimal exclusion sensitivity [111]. An Asimov dataset is one in which all observables are set to their expected values, or equivalently in which statistical fluctuations are suppressed. Ref. [111] demonstrates that Asimov datasets can be used to obtain the median significance for both discovery and exclusion hypotheses, without

generating any MC toy experiments. This provides a quick way to estimate sensitivity in the asymptotic approximation limit.

The choice of bin edges that maximizes discovery sensitivity is also the one that maximizes exclusion sensitivity, but the former has the extra advantage of offering a closed analytical form that is faster to compute. Once the optimal bins have been found, the proper expected sensitivity can be determined with a more accurate (but slower) framework, discussed below.

The asymptotic discovery sensitivity based on Asimov datasets in a given bin is expressed as [112]:

$$Z_A = \sqrt{2\left((s+b)\ln\left[\frac{(s+b)(b+\sigma_b^2)}{b^2+(s+b)\sigma_b^2}\right] - \frac{b^2}{\sigma_b^2}\ln\left[1+\frac{\sigma_b^2 s}{b(b+\sigma_b^2)}\right]\right)}, \tag{10.3}$$

where s is the expected signal yield, b is the predicted background yield, and σ_b is the uncertainty on that prediction. It is important that the expression account for the uncertainty on the background prediction since the closure of the ABCD method can have an impact on the actual sensitivity. The usual expression $s/\sqrt{s+b}$ used to estimate sensitivity is not adequate here because it does not include uncertainties, so it can not only overestimate the sensitivity but also shift the optimal bin edges away from their true locations. Furthermore, the approximation $s \ll b$ is not valid in the regime considered here (i.e. after the final event selection is applied). In fact, Eq. (10.3) reduces to its more well-known forms in the limit of zero uncertainty and of small s/b ratios:

$$\sigma_b \to 0 \Rightarrow Z_A = \sqrt{2\left[(s+b)\ln\left(1+\frac{s}{b}\right)-s\right]}, \tag{10.4}$$

$$s \ll b \Rightarrow Z_A = \frac{s}{\sqrt{b}}. \tag{10.5}$$

Since signal and background yields should have similar sizes, we must resort to the full Eq. (10.3) to estimate a reasonable uncertainty on the prediction due to ABCD closure. The error σ_b is taken as the uncertainty on the prediction due to Poisson fluctuations on bins A, B, and D:

$$\sigma_b = N^{VR}_{C,\text{pred}}\sqrt{\frac{1}{N_A}+\frac{1}{N_B}+\frac{1}{N_D}}. \tag{10.6}$$

To validate asymptotic discovery as a proxy, discovery sensitivity is compared against expected exclusion sensitivity as computed by the Higgs Combine framework in the asymptotic approximation for a range of bin edges. This is shown in Table 10.5 for a representative signal sample. The comparison confirms that both methods predict the same region of maximum sensitivity, encouraging the use of asymptotic discovery sensitivity for a fast computation of bin edges as a proxy of optimal exclusion sensitivity.

Table 10.5: Comparison of asymptotic discovery significance and combine exclusion limits using asymptotic approximation, both with Asimov datasets, for the sample $(m_1, m_2) = (50, 55)$ GeV.

S_v \ $\Delta\phi$	0.2	0.6	1.0
400	1.50	2.15	2.18
350	0.87	1.27	1.28
300	0.67	0.97	0.98
250	0.52	0.76	0.76
200	0.41	0.58	0.58
150	0.36	0.51	0.51
100			

S_v \ $\Delta\phi$	0.2	0.6	1.0
400	2.83	1.91	1.51
350	4.05	3.47	2.88
300	4.91	4.77	3.84
250	5.44	5.76	4.86
200	6.79	7.16	6.20
150	7.53	7.88	6.97
100	8.60	9.10	8.28

Figs. 10.14 and 10.15 show scans of bin edges in the 2D plane defined by v_{xy} and $|\Delta\phi_{\mathrm{MM}}|$ for a few signal samples. These plots emphasize that the optimal bin choice depends strongly on the displacement of the signal. Fig. 10.14 represents the ideal scenario, in which the optimal bin shifts upward proportionally with increasing displacement. In other cases such as in Fig. 10.15, lower and/or larger lifetimes (i.e. 1 mm and 1,000 mm) may have limited statistics after the full event selection is applied. The optimal strategy here is to maximize the size of the signal-enriched bin and therefore the algorithm correctly picks lower v_{xy} edges.

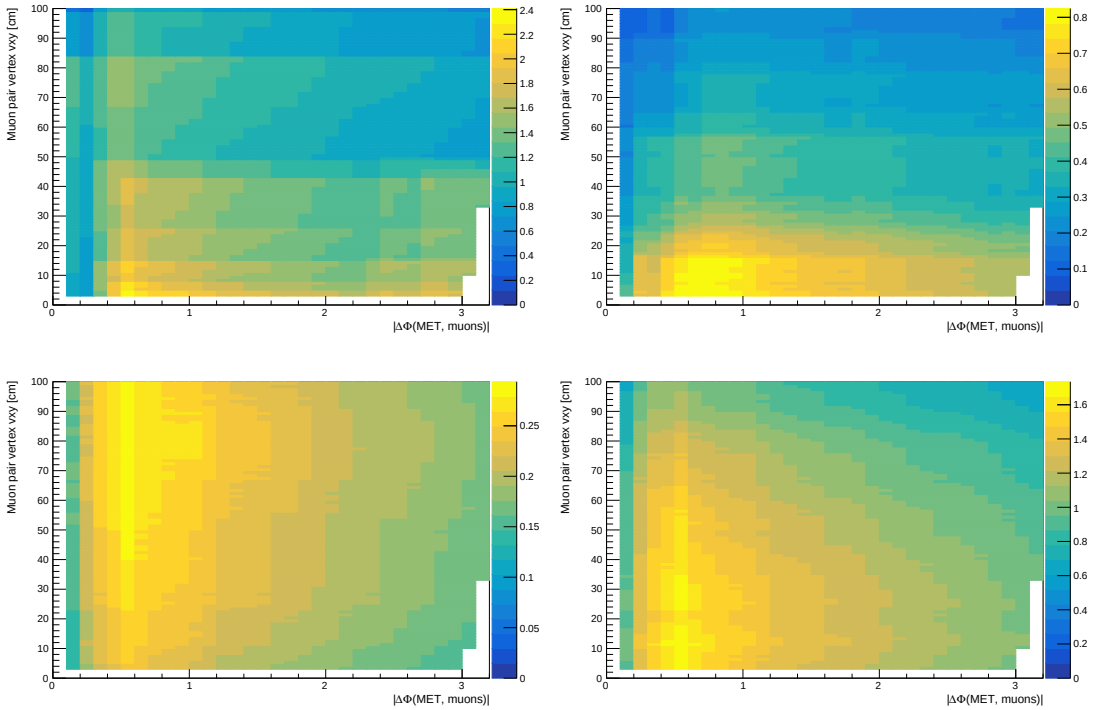


Figure 10.14: (0 GM-dSA) Scan of the asymptotic discovery significance of different ABCD bins for representative signal samples with $(m_1, m_2) = (40, 44)$ GeV, using the method described in the text. Clockwise from the upper left: $c\tau = 1$ mm, $c\tau = 10$ mm, $c\tau = 100$ mm, and $c\tau = 1000$ mm.

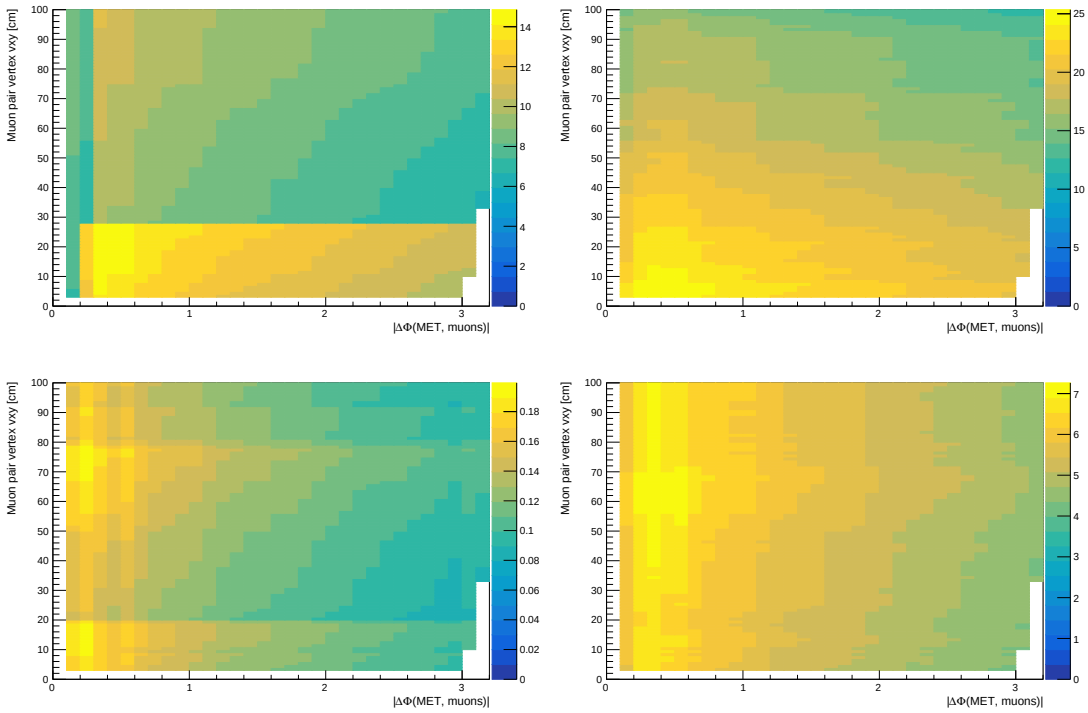


Figure 10.15: (0 GM-dSA) Scan of the asymptotic discovery significance of different ABCD bins for representative signal samples with $(m_1, m_2) = (5, 5.5)$ GeV, using the method described in the text. Clockwise from the upper left: $c\tau = 1$ mm, $c\tau = 10$ mm, $c\tau = 100$ mm, and $c\tau = 1000$ mm. The horizontal bands in the left plots are caused by low event selection yields, which occurs in some samples (in the 1 mm case, due to the $d_{xy} > 1$ mm selection, and in the 1000 mm case, due to low cross sections).

To simplify the bin optimization, boundaries are divided into groups based on the dark matter displacement of signal samples. We found that the distribution of optimal boundaries is roughly the same for each group, except when a sample has low statistics due to low selection efficiency. In this case, shifting the boundaries does not have a strong effect in the sensitivity anyway. Table 10.6 lists the optimal boundaries of all groups for each match category as well as the category that typically yields the best sensitivity. The boundary optimization also ensures that none of the bins have zero events, which would spoil the ABCD method. Boundaries differ by category because of the varying background composition (QCD is dominant in 1 and 2 GM-dSA and misidentified muon events in 0 GM-dSA) and the reduced displacement of GMs compared to dSA muons.

Sample displacement	(x, y) edges [rad, cm]			Most sensitive
	0 GM-dSA	1 GM-dSA	2 GM-dSA	
1 mm	(0.5, 10)	(0.7, 5)	(0.5, 2)	2 GM-dSA
10 mm	(0.5, 10)	(0.7, 5)	(0.5, 2)	2 GM-dSA
100 mm	(0.5, 40)	(0.7, 5)	(0.5, 9)	0 GM-dSA
1000 mm	(0.5, 80)	(0.7, 5)	(0.5, 9)	0 GM-dSA

Table 10.7 reports the preliminary background and signal predictions in each ABCD bin in the most sensitive category based on the templated method. I include the four representative samples considered throughout the analysis plus the sample with the largest cross-section, which receives an enhancement from Z mixing with the dark photon (since $m_{A'} = 3\,m_1 \approx 90\,\text{GeV}$). The large cross section of a few signal points motivates the usage of simulated samples for normalization, so we can remain fully blinded in the SR and avoid signal contamination concerns. The background predictions here are used solely for the bin optimization; the actual background in the C bin is predicted from the unblinded SR itself.

Table 10.7: SR background and signal yields used in the ABCD bin optimization procedure, for representative signal points. These background predictions are used solely for bin optimization; the ultimate prediction comes from the data itself in the SR when it is unblinded. Background yields $N^{\text{bkg}}_{A,B,C,D}$ are based on templates derived from the nJets VR in data and normalized with simulation as described in the text. Signal yields $N^{\text{sig}}_{A,B,C,D}$ are extracted directly from simulation. The edges defining the bins are the ones listed in Table 10.6. For illustrative purposes, we report yields from the most sensitive match category in each case, even though all categories are fit simultaneously.

Sample	N_A^{bkg}	N_A^{sig}	N_B^{bkg}	N_B^{sig}	N_C^{bkg}	N_C^{sig}	N_D^{bkg}	N_D^{sig}	Category
(5, 5.5) GeV, 1 mm	5.8	34.1	2.3	0.0	18.5	568.4	3.5	23.2	2 GM-dSA
(5, 5.5) GeV, 10 mm	5.8	12.9	2.3	0.0	18.5	213.5	3.5	63.6	2 GM-dSA
(5, 5.5) GeV, 100 mm	23.8	13.7	44.7	0.0	43.8	96.6	107.6	0.0	0 GM-dSA
(5, 5.5) GeV, 1000 mm	35.2	0.7	77.1	0.0	32.4	1.3	75.2	0.1	0 GM-dSA
(5, 7) GeV, 1 mm	5.8	2.3	2.3	0.4	18.5	1.5	3.5	0.3	2 GM-dSA
(5, 7) GeV, 10 mm	5.8	0.2	2.3	0.02	18.5	1.6	3.5	0.2	2 GM-dSA
(5, 7) GeV, 100 mm	23.8	0.05	44.7	0.0	43.8	0.4	107.6	0.0	0 GM-dSA
(5, 7) GeV, 1000 mm	35.2	0.0	77.1	0.0	32.4	0.01	75.2	0.0	0 GM-dSA
(50, 55) GeV, 1 mm	5.8	61.0	2.3	45.4	18.5	0.0	3.5	0.0	2 GM-dSA
(50, 55) GeV, 10 mm	5.8	39.3	2.3	40.1	18.5	26.4	3.5	16.8	2 GM-dSA
(50, 55) GeV, 100 mm	23.8	1.4	44.7	0.2	43.8	6.8	107.6	0.5	0 GM-dSA
(50, 55) GeV, 1000 mm	35.2	0.3	77.1	0.1	32.4	1.2	75.2	0.2	0 GM-dSA
(50, 70) GeV, 1 mm	5.8	0.1	2.3	0.3	18.5	0.0	3.5	0.0	2 GM-dSA
(50, 70) GeV, 10 mm	5.8	0.1	2.3	0.2	18.5	0.02	3.5	0.03	2 GM-dSA
(50, 70) GeV, 100 mm	23.8	0.0	44.7	0.0	43.8	0.01	107.6	0.0	0 GM-dSA
(50, 70) GeV, 1000 mm	35.2	0.0	77.1	0.0	32.4	0.01	75.2	0.0	0 GM-dSA
(30, 33) GeV, 1 mm	5.8	5460	2.3	2262	18.5	153.3	3.5	10.6	2 GM-dSA
(30, 33) GeV, 10 mm	5.8	2315	2.3	810.3	18.5	3218	3.5	1078	2 GM-dSA
(30, 33) GeV, 100 mm	23.8	129.4	44.7	1.7	43.8	800.2	107.6	14.1	0 GM-dSA
(30, 33) GeV, 1000 mm	35.2	14.0	77.1	0.4	32.4	75.3	75.2	2.0	0 GM-dSA

CHAPTER 11

UNCERTAINTIES

This chapter discusses the analysis uncertainties. The dominant uncertainty is statistical and arises from limited number of events in the ABCD method. The uncertainty on the four fit parameters of the ABCD calculation are automatically accounted for in the signal extraction.

The dominant systematic uncertainty is associated to ABCD closure and to the extent that background predictions match the observed rate in the orthogonal validation regions. This was described in Chapter 10 but here we discuss the actual closure uncertainties used. We also report systematic uncertainties impacting the signal prediction extracted from simulation, which are however sub-dominant to the closure uncertainties. These signal uncertainties are the same across all ABCD bins, since they affect equally the overall simulation yield estimates. The exception are the global muon scale factor uncertainties which depend on the match category.

The treatment of correlations between systematic effects is done according to the standard CMS procedures whenever available.

11.1 Luminosity

There is an uncertainty associated to the measurement of the integrated luminosity collected by CMS. The uncertainties by year are estimated centrally by CMS and reported in Table 11.1. These uncertainties are applied to simulated samples of signal when scaling the simulation yields to the correct luminosity. They are treated as uncorrelated across years.

Table 11.1: Luminosity uncertainties assigned per year.

Year	Uncertainty
2018	2.5%
2017	2.3%
2016	2.5%

11.2 Trigger scale factor

The simulated trigger efficiency cannot match the behavior in data exactly due to several factors such as limited sample sizes and unexpected inefficiencies arising over the course of data-taking. Therefore, a correction is required to be applied to simulated samples in order to restore the agreement. This trigger scale factor is derived from the efficiency curves of both simulated samples and data as discussed in Section 7.2. Here we study the uncertainties on that measurement.

The scale factor is constructed by taking the ratio of the trigger simulated efficiency to the one observed in data for each year. This ratio is used as a trigger weight for each event, which is a function of the event's MetNoMu, and it is never greater than 1. We estimate the uncertainty on the trigger weight by varying the data and simulation measurements independently based on their statistical uncertainties (up and down by 1σ) and recalculating the weights. These new weights are then fed into the analysis workflow and the new signal yields are compared to the nominal ones. The envelope of the percent discrepancy (i.e. the maximum uncertainty across all signal samples) is conservatively taken as the scale factor uncertainty. Thus two scale factor uncertainties are derived (one for simulation and one for data) which are treated as independent. This procedure is adapted from Ref. [58]. Table 11.2 lists the resulting uncertainties by year.

Table 11.2: Trigger scale factor uncertainties for data and simulation.

Year	Data	Simulation
2018	0.5%	1%
2017	0.5%	1%
2016	0.5%	1%

11.3 JES and JER

Jet energy scale and resolution uncertainties are assessed in a similar manner to the trigger scale factor uncertainties. The $\pm 1\sigma$ variations in each of JES and JER are independently propagated through the event selection and the resulting differences in signal yield after the final selection are taken as uncertainties. The jet variations are also propagated to the $p_{\mathrm{T}}^{\mathrm{miss}}$ as part of this assessment. Table 11.3 shows the JES and JER uncertainties estimated for each year.

Table 11.3: JES and JER uncertainties per year.

Year	JES	JER
2018	2%	2.5%
2017	6%	9%
2016	2%	1%

The noticeably larger uncertainty in 2017 has been investigated and found to be caused by a well-known issue in 2017 data. Jets in the region $2.65 < \eta < 3.3$ have a much larger uncertainty compared to 2018 as a result of increased noise in the endcap calorimeter during the 2017 data-taking period. This is shown in Fig. 11.1. Combined with the fact that more jets are present in this region, the uncertainty directly translates into jets with reduced p_{T} and leads to an increase in the number of events in the signal region, which is defined by requiring fewer than three jets with p_{T} greater than 30 GeV.

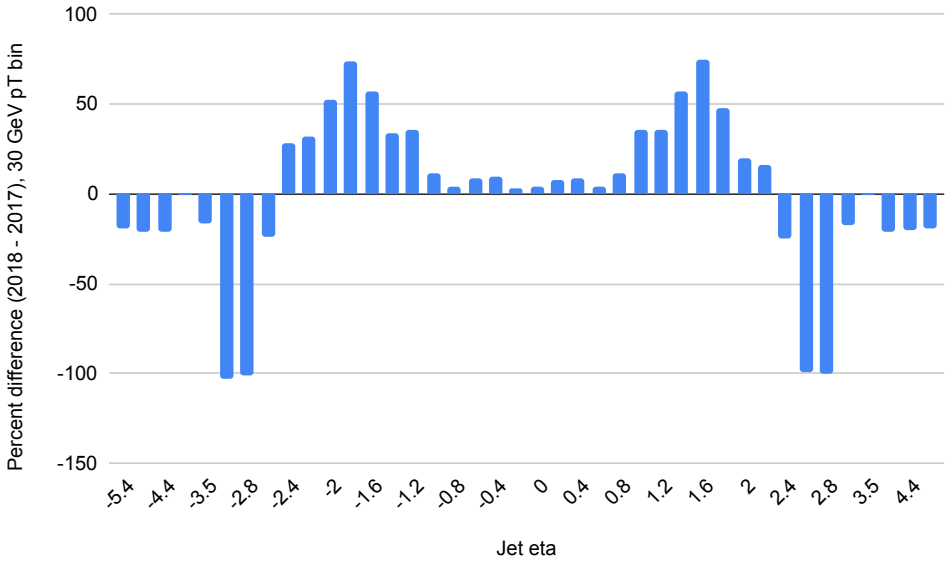


Figure 11.1: JES uncertainty comparison between 2017 and 2018 for jets with 30 GeV p_{T}, as a function of jet η, measured centrally by CMS. The comparison is given as $100\% \times (2018 - 2017)/(2018 + 2017) \times 2$.

11.4 Closure of ABCD in data

One source of uncertainty in data comes from the closure of the ABCD method, caused by slight deviations from perfect independence between the two variables in the ABCD plane. This uncertainty is assessed with the orthogonal nJets and dR validation regions. The first is used to assess closure in the 1 and 2 match categories, which are dominated by QCD, and the second in the 0 match category, dominated by misidentified muons in Z/W+jets events.

Closure is assessed by predicting C bin yields from other bins in the validation regions and comparing against the observed yields. This is done for all combinations of bin boundaries used in the signal extraction. To reduce the influence of statistical fluctuations, the bin edges are varied locally around each boundary and the calculation is repeated each time. The average closure is then taken as a measure of the uncertainty.

Table 11.4 reports the closure uncertainties estimated for the bin boundaries found in Table 10.6. The higher uncertainties in the 2 GM-dSA match category are due to limited statistical power compared to the other categories.

Table 11.4: Closure uncertainties per match category.

Match category	Sample displacement	(x, y) edges [rad, cm]	Uncertainty on C bin
0 GM-dSA	1 mm	(0.5, 10)	25%
	10 mm	(0.5, 10)	25%
	100 mm	(0.5, 40)	20%
	1000 mm	(0.5, 80)	10%
1 GM-dSA	1 mm	(0.7, 5)	5%
	10 mm	(0.7, 5)	5%
	100 mm	(0.7, 5)	5%
	1000 mm	(0.7, 5)	5%
2 GM-dSA	1 mm	(0.5, 2)	80%
	10 mm	(0.5, 2)	80%
	100 mm	(0.5, 9)	20%
	1000 mm	(0.5, 9)	20%

11.5 Veto ID scale factor

The electron and photon veto scale factor weight was discussed in Section 9.2.5. Uncertainties are calculated by varying the individual scale factors by the statistical and systematic up and down variations and re-calculating the weight. The new weight is applied to background simulated samples and the yield variation is taken as a measure of the veto scale factor uncertainty. The resulting veto uncertainties are small and sub-dominant compared to other scale factors. They are rounded up to 0.5% to be conservative. The uncertainties are reported in Table 11.5.

Table 11.5: Photon and electron veto ID scale factor uncertainties assigned per year.

Year	Uncertainty
2018	0.5%
2017	0.5%
2016	0.5%

11.6 Muon scale factors

There are two muon scale factor uncertainties to consider: GM and dSA muons. They are both measured by varying the scale factors up and down according to the systematic and statistical uncertainties on the measured scale factors.

11.6.1 Global muons

The uncertainties measured for global muons are shown in Table 11.6. They were derived with scale factor uncertainties measured centrally by CMS for all years and include both statistical and systematic components. Results are broken down by match category since each category has a different number of global muons. These are assumed 100% correlated between years since the technique is the same.

Table 11.6: GM ID scale factor uncertainties assigned per year.

Year	0 GM-dSA	1 GM-dSA	2 GM-dSA
2018	-	0.5%	0.8%
2017	-	0.2%	0.5%
2016	-	0.6%	1.3%

11.6.2 Displaced standalone muons

The preliminary and conservative assignment for the uncertainty on displaced standalone scale factors is 5% for all years. The ultimate measurement, in progress, will come from a sample of cosmic muons.

11.7 Summary

Table 11.7 presents a summary of all uncertainties in the analysis and their yearly correlations.

Table 11.7: Summary of systematic uncertainties in the analysis (in percent).

Uncertainty	2016	2017	2018	Correlation	Bin(s)	Process
Luminosity	2.5	2.3	2.5	Uncorrelated	All	Signal
Trigger (MC)	1	1	1	100% correlated	All	Signal
Trigger (data)	0.5	0.5	0.5	100% correlated	All	Signal
JES	2	6	2	100% correlated	All	Signal
JER	1	9	2.5	Uncorrelated	All	Signal
Veto ID	0.5	0.5	0.5	100% correlated	All	Signal
GM ID (1 GM-dSA)	0.6	0.2	0.5	100% correlated	All	Signal
GM ID (2 GM-dSA)	1.3	0.5	0.8	100% correlated	All	Signal
dSA ID (prelim.)	5	5	5	100% correlated	All	Signal
dSA reco. (prelim.)	5	5	5	100% correlated	All	Signal
Closure (0 GM-dSA)		10–25		100% correlated	C	Background
Closure (1 GM-dSA)		5		100% correlated	C	Background
Closure (2 GM-dSA)		20–80		100% correlated	C	Background

CHAPTER 12

RESULTS AND CONCLUSIONS

This chapter discusses the signal extraction procedure and presents preliminary results of the analysis, illustrating the projected sensitivity achievable with CMS after the signal region is approved for unblinding.

12.1 Exclusion limit estimation

The predicted background and signal yields from the ABCD bin optimization are input into HIGGS COMBINE for a preliminary estimation of the exclusion sensitivity. In this case, a more accurate (albeit slower) computation of the expected sensitivity is performed. Note that the predicted backgrounds are really only approximations derived from the nJets templates. Nevertheless, this provides a reasonable estimate of projected sensitivity.

The procedure consists of simultaneously fitting all four ABCD bin yields with free parameters N_A^{bkg}, c_1 and c_2 in addition to the signal strength μ. This method has previously been used by e.g. Ref. [63]. It relies on the fact that the two observables in the 2D plane are independent, which affords the use of only 3 parameters (instead of 4) to predict the background in all bins. Combined with the signal strength, there are then 4 unknowns and 4 equations to be fit by HIGGS COMBINE:

$$N_A = N_A^{\text{bkg}} + \mu \times N_A^{\text{sig}}, \tag{12.1}$$

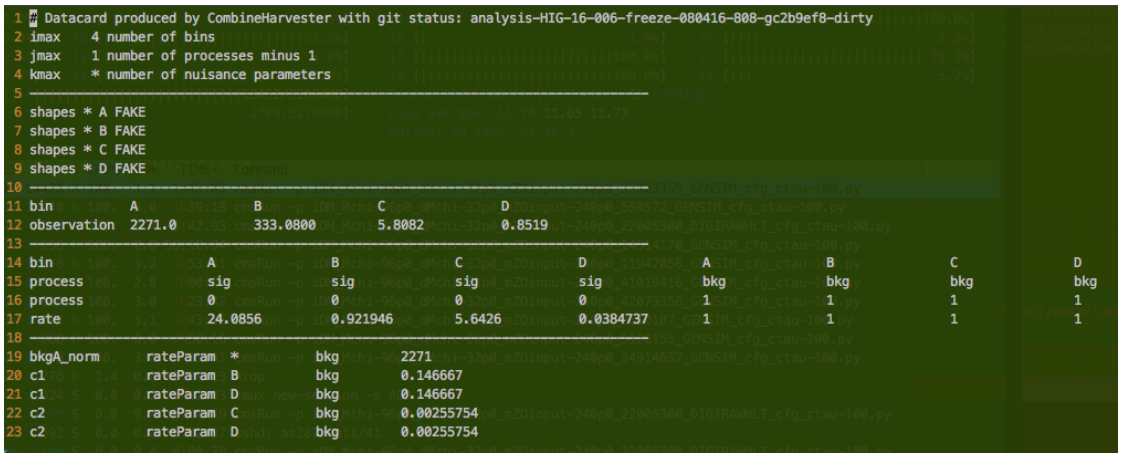
$$N_B = N_A^{\text{bkg}} \times c_1 + \mu \times N_B^{\text{sig}}, \tag{12.2}$$

$$N_C = N_A^{\text{bkg}} \times c_2 + \mu \times N_C^{\text{sig}}, \tag{12.3}$$

$$N_D = N_A^{\text{bkg}} \times c_1 \times c_2 + \mu \times N_D^{\text{sig}}. \tag{12.4}$$

Here, the four unknowns are N_A^{bkg}, which sets the background normalization for all bins, c_1 and c_2, the ratios B/A and C/A, and μ, the signal strength. The expected signal yields are directly extracted from simulation for each signal hypothesis. The initial guesses for c_1 and c_2 are derived directly from the template, while the initial guess for N_A^{bkg} is just that found by normalizing with background simulation. This method is robust against signal contamination in any and all bins since it only relies on the background independence between the two variables in the 2D plane. COMBINE fits all the bins *simultaneously*, which is the key to enable this method.

Fig. 12.1 provides a simplified example of a COMBINE datacard with all the ingredients described. COMBINE datacards are automatically generated using the COMBINE-HARVESTER package, which serves as a bridge between the ABCD bin optimization code and the computation of expected limits. The actual COMBINE datacard used is more complex: the three GM-dSA match categories are simultaneously fit. Additionally, each year is treated separately to allow for separate systematic uncertainties. Thus we essentially perform nine simultaneous ABCD fits, but some fit parameters are common. The coefficients c_1 and c_2 only vary between match categories, but not between years. The signal strength parameter μ is the same for all ABCD fits. The background normalization N_A^{bkg} does vary for each ABCD fit. Therefore, there are a total of 36 equations with 16 unknowns (9 background normalizations N_A^{bkg}, 6 correlation coefficients c_1 and c_2, and 1 signal strength μ). This has the added advantage of automatically considering the inclusive match category as well.



```
# Datacard produced by CombineHarvester with git status: analysis-HIG-16-006-freeze-080416-808-gc2b9ef8-dirty
imax    4 number of bins
jmax    1 number of processes minus 1
kmax    * number of nuisance parameters
--------------------------------------------------------------------------------
shapes * A FAKE
shapes * B FAKE
shapes * C FAKE
shapes * D FAKE
--------------------------------------------------------------------------------
bin          A       B       C       D
observation  2271.0  333.0800  5.8082  0.8519
--------------------------------------------------------------------------------
bin          A       B       C       D       A       B       C       D
process      sig     sig     sig     sig     bkg     bkg     bkg     bkg
process      0       0       0       0       1       1       1       1
rate         24.0856 0.921946 5.6426  0.0384737 1       1       1       1
--------------------------------------------------------------------------------
bkgA_norm    rateParam  *       bkg     2271
c1           rateParam  B       bkg     0.146667
c1           rateParam  D       bkg     0.146667
c2           rateParam  C       bkg     0.00255754
c2           rateParam  D       bkg     0.00255754
```

Figure 12.1: Example of Higgs Combine datacard used in the estimation of expected exclusion limits with only a single match category. The procedure uses a modified ABCD simultaneous fit to all four bins, similar to [63]. The actual datacards feature a separate ABCD for each GM-dSA match category and for each year.

12.2 Preliminary results

Figs. 12.2 to 12.4 report 1D projected sensitivities for representative samples. Figs. 12.2 and 12.3 show the expected sensitivity as a function of χ_2 lifetime for four m_1 and Δ combinations, assuming $\alpha_D = \alpha_{EM}$ and $\alpha_D = 0.1$, respectively. The theoretical cross sections calculated in Section 5.2 and summarized in Table 7.1 are displayed in dashed blue lines and the plots in each figure are sorted by decreasing cross section, starting from the upper left corner. The projected limit is shown by red lines, while the green and yellow bands correspond to the 68% and 95% standard deviation bands, respectively. Since the signal region is still blinded while awaiting approval to unblind from the Collaboration, the "observed" limits shown in black dots and black lines are similar to the projected ones in red.

The behavior of the cross section varies with α_D, m_1, Δ, and $c\tau$. Lower values of α_D translate to higher values of ϵ and hence cross sections. This is evident when comparing Fig. 12.2 and Fig. 12.3. The cross section is also lowered for higher m_1 and higher Δ, as

shown by the plot variations in each figure. In Fig. 12.3, for the largest mass-splittings and lifetimes, the cross section is too small and no events remain after the full selection is applied.

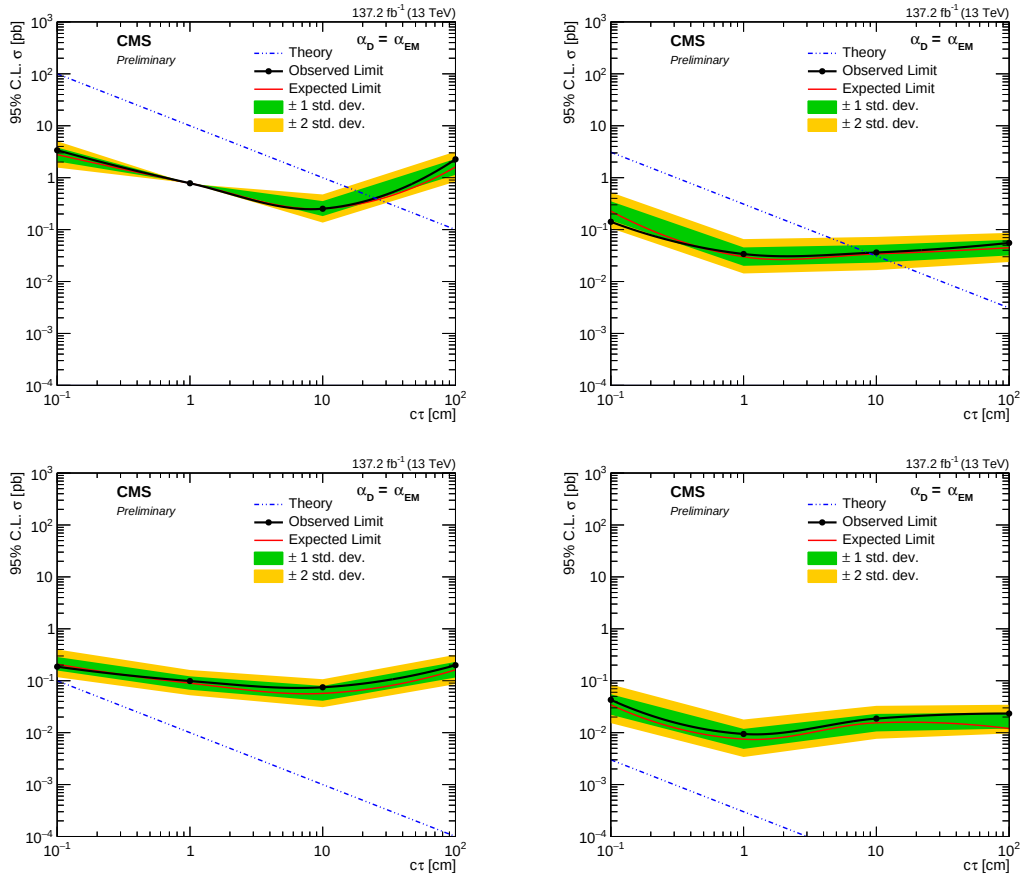


Figure 12.2: Projected 1D expected exclusion limits for four representative mass points as a function of proper lifetime, assuming $\alpha_D = \alpha_{EM}$. Upper left: $m_1 = 5\,\text{GeV}$, $\Delta = 0.1\,m_1$. Upper right: $m_1 = 50\,\text{GeV}$, $\Delta = 0.1\,m_1$. Lower left: $m_1 = 5\,\text{GeV}$, $\Delta = 0.4\,m_1$. Lower right: $m_1 = 50\,\text{GeV}$, $\Delta = 0.4\,m_1$.

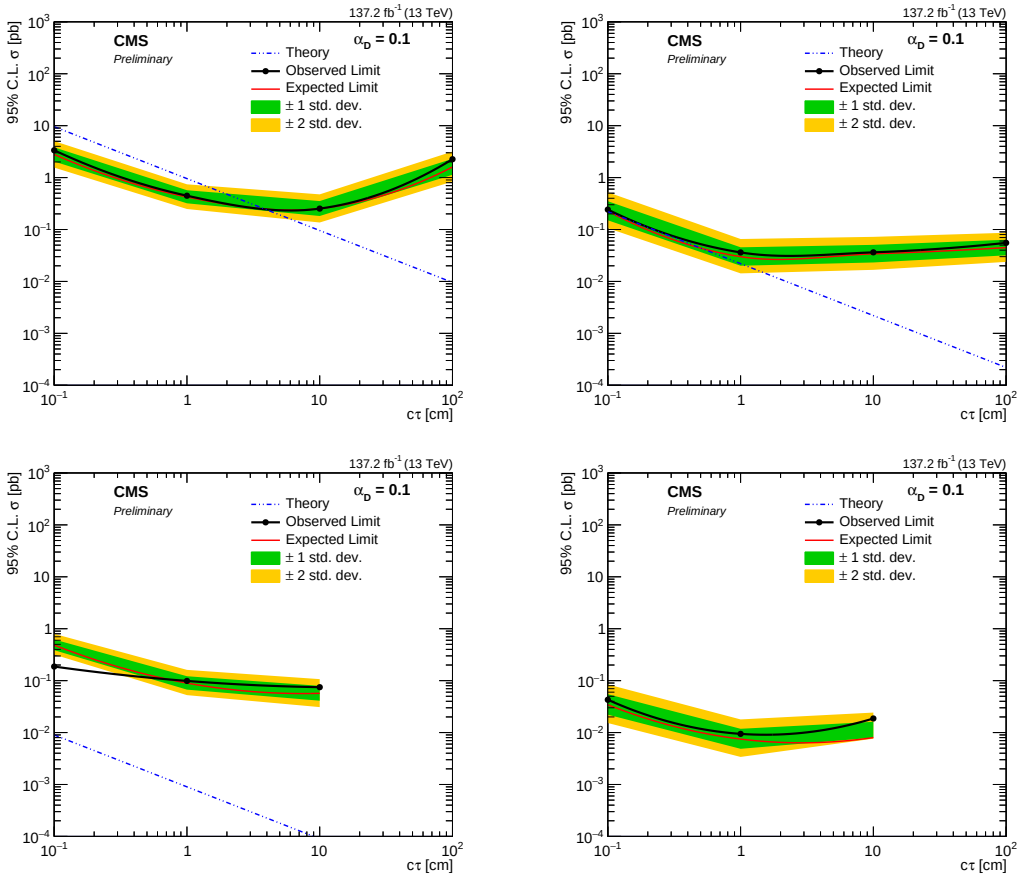


Figure 12.3: Projected 1D expected exclusion limits for four representative mass points as a function of proper lifetime, assuming $\alpha_D = 0.1$. Upper left: $m_1 = 5\,\text{GeV}$, $\Delta = 0.1\,m_1$. Upper right: $m_1 = 50\,\text{GeV}$, $\Delta = 0.1\,m_1$. Lower middle: $m_1 = 5\,\text{GeV}$, $\Delta = 0.4\,m_1$. Some cross sections for the $\Delta = 0.4$ mass splitting case are too small and no events remain after the full selection is applied.

Fig. 12.4 plots the expected sensitivity as a function of light dark matter mass for a combination of lifetimes, mass splittings, and α_D. The sensitivity is highest for moderately displaced signals, since prompt signals must contend with larger background overlap, while highly displaced ones suffer from low yields due to the decreased muon reconstruction and identification efficiency.

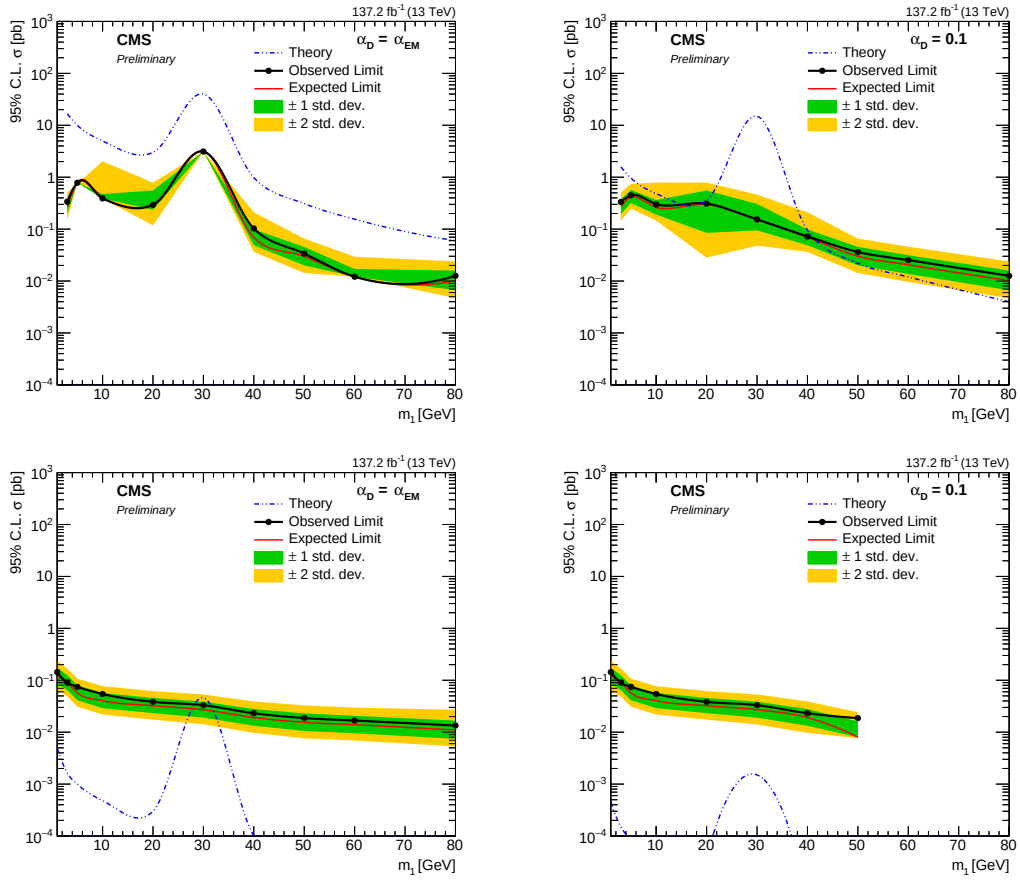


Figure 12.4: Projected 1D expected exclusion limits for a combination of lifetimes, mass-splittings and α_D as a function of light dark matter mass. Upper left: $\alpha_D = \alpha_{EM}$, $\Delta = 0.1\,m_1, c\tau = 10\,\text{mm}$. Upper right: $\alpha_D = 0.1$, $\Delta = 0.1\,m_1, c\tau = 10\,\text{mm}$. Lower left: $\alpha_D = \alpha_{EM}$, $\Delta = 0.4\,m_1, c\tau = 100\,\text{mm}$. Lower right: $\alpha_D = 0.1$, $\Delta = 0.4\,m_1, c\tau = 100\,\text{mm}$. Some cross sections for the $\Delta = 0.4$ mass splitting case are too small and no events remain after the full selection is applied.

Fig. 12.5 displays 2D projected exclusion plots for $\Delta = 0.4\,m_1$ (left) and $\Delta = 0.1\,m_1$ (right), assuming $\alpha_D = 0.1$ (top) and $\alpha_D = \alpha_{EM}$ (bottom). These plots include all processed samples. The 95% CL limits are reported in text overlaid on each bin, and exclusion limit curves (i.e. where $\sigma_{\text{obs}}/\sigma_{\text{th}} = 1$) are also drawn. Overall, these preliminary estimates indicate that CMS has untapped sensitivity to inelastic dark matter. The 10% mass splitting parameter space is particularly sensitive, since the cross section is considerable suppressed for larger mass splittings such as 40%. Similarly, sensitivity seems to be enhanced around dark photon masses near the Z boson, which agrees with the studies discussed in Section 5.2. The largest lifetimes have particularly low cross sections and hence less sensitivity due to the double suppression arising from dark photon production on one hand, and small decay widths on the other. An overall factor of ϵ^4 results from this double suppression. More data will be needed to probe this region of parameter space. The Run 3 of CMS, starting in 2022, will enable an even deeper investigation of inelastic dark matter by collecting an additional $150\,\text{fb}^{-1}$ of data. Moreover, specialized triggers currently under study will increase the selection efficiency to inelastic dark matter events, further enhancing the sensitivity. Nevertheless, even with the data already collected between 2016 and 2018 CMS will likely be able to place the first hadron collider limits on inelastic dark matter with unprecedented dark matter mass reach compared to other accelerator-based experiments.

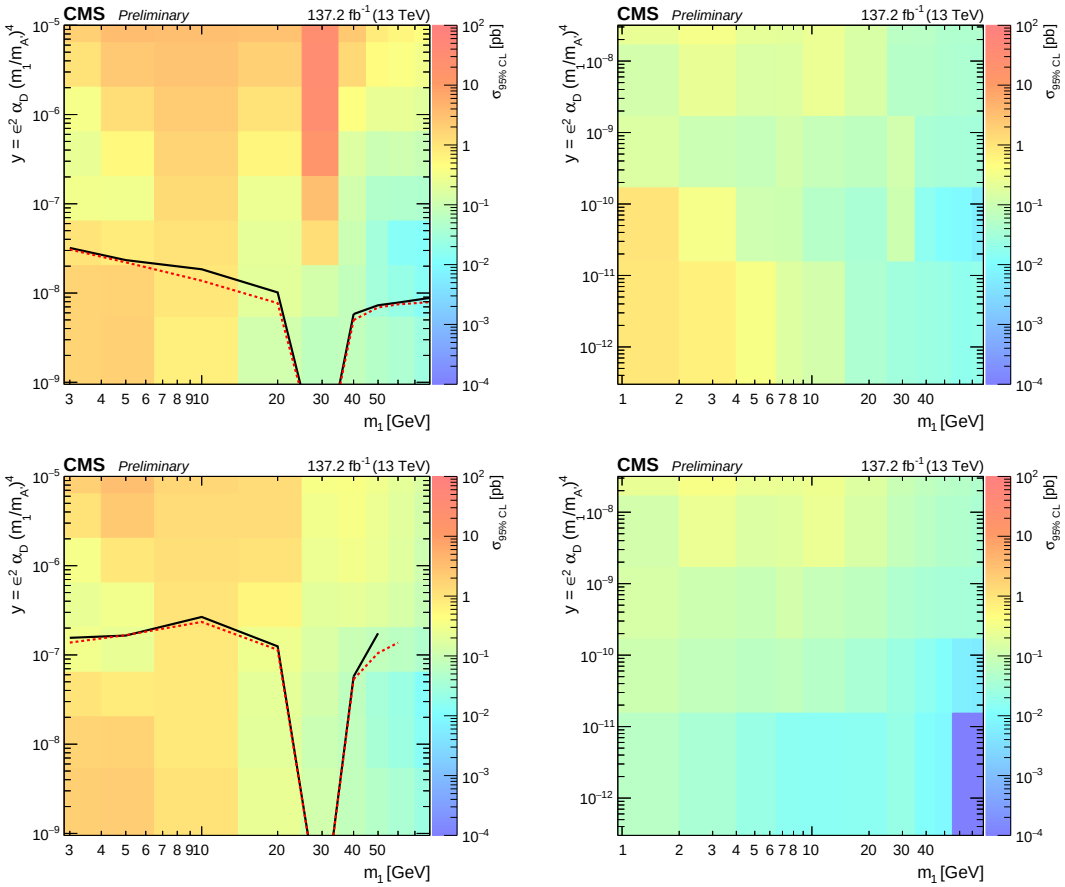


Figure 12.5: Projected 2D expected exclusion limits for all signal samples. Upper left: $\alpha_D = \alpha_{EM}$, $\Delta = 0.1\,m_1$. Upper right: $\alpha_D = \alpha_{EM}$, $\Delta = 0.4\,m_1$. Lower left: $\alpha_D = 0.1$, $\Delta = 0.1\,m_1$. Lower right: $\alpha_D = 0.1$, $\Delta = 0.4\,m_1$. Histograms indicate expected cross section limits and contours (when they exist) indicate expected excluded parameter space (i.e. the region where $\sigma_{obs}/\sigma_{th} < 1$).

Part III

Search for Dark Photons with PADME

The following three chapters describe the Positron Annihilation into Dark Matter Experiment (PADME). PADME seeks to uncover evidence of dark photons by colliding a beam of positrons onto a target made of polycrystalline diamond. It is therefore an example of a fixed-target experiment that features the advantages discussed in Section 4.3. PADME directly probes the kinetic mixing between dark photons and SM bosons and is sensitive to a complex dark sector independent of any particular models. This provides good complementarity to more specific searches such as the inelastic dark matter analysis described in Part II. As mentioned in Section 4.3, fixed-target experiments are also complementary to collider experiments across a number of metrics, illustrating the benefits of pursuing a varied approach in the search for dark matter.

Chapter 13 introduces the search strategy and experimental apparatus, highlighting the expected phenomenology, sub-detectors unique to PADME, and expected physics backgrounds. Chapter 14 describes a performance study of the small-angle calorimeter, a key component of the setup dedicated to fast identification of Bremsstrahlung radiation (a major source of background). Finally, Chapter 15 discusses some prospects, including searches for more specific dark matter models that can eventually be performed with the PADME data. It also discusses the potential to probe the Beryllium and Helium anomalies, a 17 MeV excess recently observed with atomic spectrometry measurements [113, 114].

CHAPTER 13

THE PADME EXPERIMENT

This chapter introduces the PADME experiment, starting with the search strategy of the experiment, followed by a description of the accelerator complex at Laboratori Nazionali di Frascati, and finally the experimental setup of PADME itself. Several sub-detectors critical for the experiment are mentioned. References to more detailed accounts are provided where appropriate.

13.1 Search strategy

PADME seeks to identify the production of dark photons from positron-electron annihilation via the kinetic mixing. If a dark sector is connected to the SM by this coupling, then processes such as the one shown in Fig. 13.1 are possible.

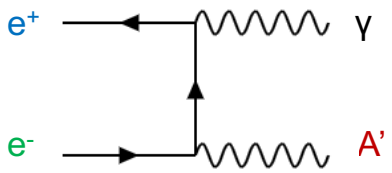


Figure 13.1: Feynman diagram of electron-positron annihilation with dark photon production. This process is enabled by a kinetic mixing coupling between the SM hypercharge and the dark photon, as discussed in Chapter 3.

In this Feynman diagram, a positron and an electron annihilate to produce two photons, but one photon converts into a dark photon via the kinetic mixing. In the experimental apparatus, the equivalent topology is minimal: only one outgoing photon is measured in addition to the incoming electron and positron; the dark photon escapes the detector without leaving a trace. This topology is often referred to as an *invisible* channel. The dark photon production rate is suppressed by a factor of ϵ^2 (where ϵ

is as usual the kinetic mixing coefficient) compared to ordinary positron-annihilation with two-photon production. Nevertheless, PADME is carefully designed to identify and remove backgrounds, such that it still has sensitivity despite low signal production rates.

The annihilation of positrons and electrons is set up with a positron beam and a fixed target made of diamond. The 500 MeV beam positrons strike the target, where interactions between carbon electrons and positrons can occur. The predominant interaction is actually emission of Bremsstrahlung radiation by positrons, which results from being in the vicinity of electric fields generated by nuclei of carbon atoms. Less common is the aforementioned two-gamma production, whereby positron and electron completely annihilate to produce two photons (sometimes a third photon can be radiatively emitted as well). PADME's objective is to identify these background processes with high confidence and remove them from consideration, leaving only signal events where a dark photon is produced.

To identify signal events, PADME measures all legs of the process shown in Fig. 13.1 except for the dark photon. By constraining the momentum of all other particles in the interaction, the *missing mass* of the system can be reconstructed. The missing mass is the momentum carried away by the dark photon:

$$m^2_{\text{miss}} = \left(p_{e^+} + p_{e^-} - p_\gamma\right)^2, \tag{13.1}$$

where p_{e^+}, p_{e^-}, and p_γ are the positron, electron, and photon four-momenta respectively. The missing mass distribution should feature a resonant peak at the location of the dark photon mass. Fig. 13.2 depicts the shape of this distribution for several dark photon mass hypotheses. The strategy of PADME is therefore a "bump-hunt" in the missing mass distribution. It does not depend on any particular dynamics in the dark sector,

aside from the required kinetic mixing between SM and dark photons, making PADME a model-independent experiment.

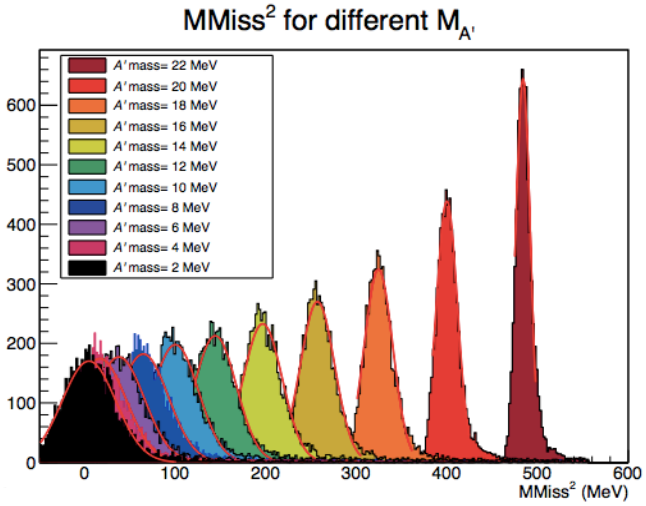


Figure 13.2: Missing mass distribution for several dark photon mass hypotheses.

The other positron-electron interactions—Bremsstrahlung, two- and three-gamma production—have different final-state topologies compared to signal. In Bremsstrahlung radiation, the positron merely emits a photon but is not destroyed in the process. Thus there are two outgoing particles, the positron itself (with lowered energy) and the soft emitted photon. In the other interactions, there are respectively two and three final-state photons. Therefore, in order to tag and remove these backgrounds, it is critical to identify all outgoing particles in the system. PADME features several sub-detectors designed with this goal in mind. These are discussed briefly in Section 13.3. Fig. 13.3 compares the expected missing mass distribution (without any signal) before and after background mitigation as predicted by MC simulations.

The tagging and removal of backgrounds and bump-hunt in the missing mass distribution give PADME sensitivity to investigate the existence of a vector portal between the SM and the dark sector in the invisible channel. The use of a positron beam is also unique in the landscape of fixed-target experiments exploring the dark sector. Typical

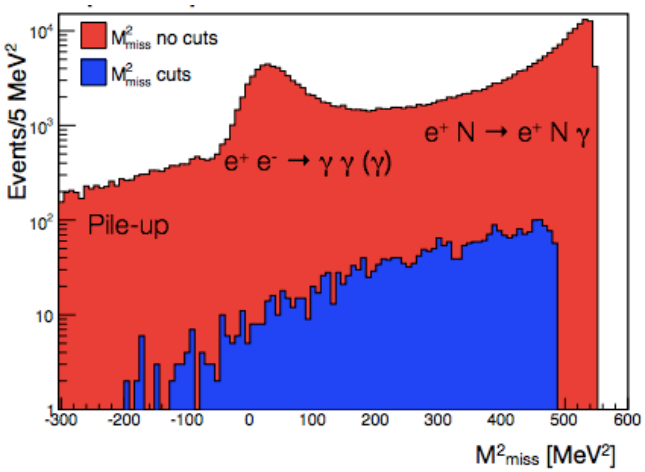


Figure 13.3: Missing mass distribution of expected backgrounds according to MC, before (red) and after (blue) mitigation with auxiliary detectors.

electron-beam experiments rely on the Bremsstrahlung emission of a dark photon ("dark-sstrahlung"), which is an α^3 process, where α is the fine-structure constant. Positron-electron annihilation is instead an α^2 process, a fact that is leveraged by PADME. These differences and the sensitivity of PADME under different data-taking scenarios are explored more fully in Chapter 15, where I also comment on additional parallel searches for specific models of dark matter enabled by the PADME data.

13.2 DAFNE complex

PADME is located at the Laboratori Nazionali di Frascati (LNF), in Frascati, Italy, and is part of the Instituto Nazionali di Fisica Nucleari (INFN). The 500 MeV positron beam is produced and accelerated at the Double Annular Φ Factory for Nice Experiments (DAFNE) complex. DAFNE is an electron-positron collider in operation since 1999 that also houses other past and current experiments, such as KLOE and KLOE-2 (which study Φ meson decays [115]). DAFNE has two accelerating lines for electrons and positrons, roughly 100 m long each. It is capable of accelerating both particles to a center-of-mass

energy of 1.02 GeV, the mass of Φ mesons.

PADME is a small experiment that sits at the end of a test beam line. The Beam Test Facility (BTF) of LNF houses the entire PADME detector. Fig. 13.4 shows the DAFNE complex and the relative location of the PADME experiment.

Figure 13.4: Bird's-eye view of the DAFNE complex and the location of the PADME experiment at the BTF (test beam line) [116].

Positrons are primarily produced at the LINAC by striking a 5.5 A beam of 200 MeV electrons against a 2 χ_0 Tungsten-Rhenium target [117]. The electrons themselves are produced with a thermo-ionic gun. The emitted positrons are further accelerated to a maximum energy of roughly 550 MeV. A 1 nC beam charge is achievable with roughly 10 ns-long pulses.

A secondary way to produce positrons is further downstream at the beginning of the BTF line, instead of at the LINAC. This is done by diverting electron bunches from the LINAC and striking a beam-attenuating target in the BTF, which emits positrons. In this case, there is a greater degree of control of the beam condition, such as the

energy range (which can be varied from 750 MeV down to a few tens of MeVs) and the beam multiplicity (which can be varied from a single particle per bunch up to roughly a million based on energy requirements) [117]. The primary and secondary ways to produce positrons at DAFNE are illustrated in Fig. 13.5.

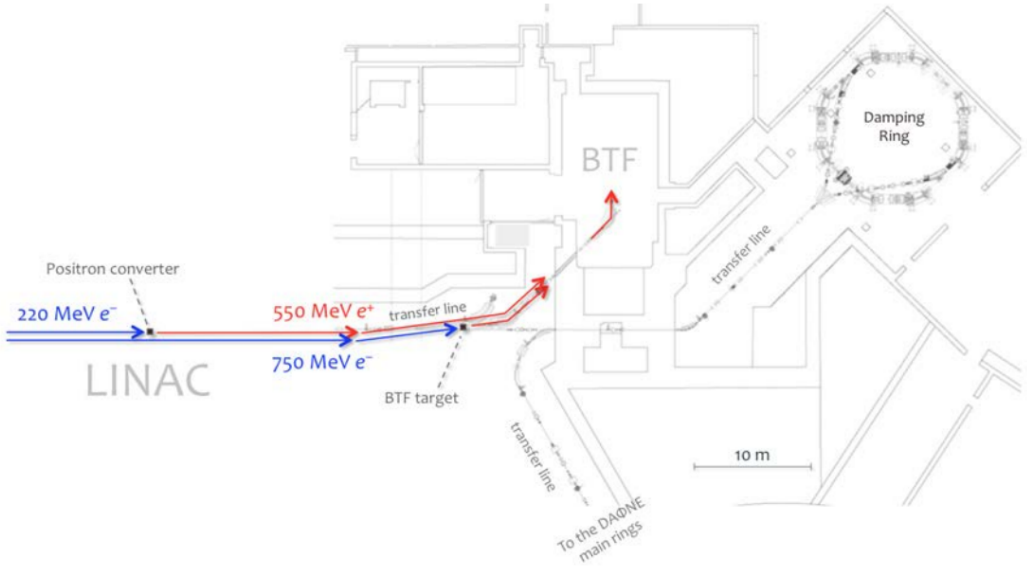


Figure 13.5: Primary and secondary positron beams delivered to the BTF line from the LINAC for injection into PADME [117].

The LINAC is capable of generating 50 electron/positron bunches per second. One bunch is used for diagnostics and a variable number is sent to KLOE-2. The remaining bunches are sent to the BTF line, where they are injected into PADME. Each 550 MeV bunch contains roughly 10^4 to 10^5 positrons, with length 150 ns to 200 ns [117]. PADME expects roughly 1×10^{13} positrons on target (POTs) delivered in a 6-month data-taking window, and proportionally more as time permits.

13.3 PADME detector

Fig. 13.6 shows a schematic of the PADME experimental setup. The positron beam is delivered by DAFNE and strikes the diamond target from the right. The target sits at the entrance of a 0.5 T magnet. The non-interacting beam is diverted by the magnet and conducted to a beam dump adjacent to the electromagnetic calorimeter on the left. The calorimeter is the main PADME detector, placed immediately downstream of the magnet. Any photons produced in the target ignore the magnetic field and travel straight to the calorimeter.

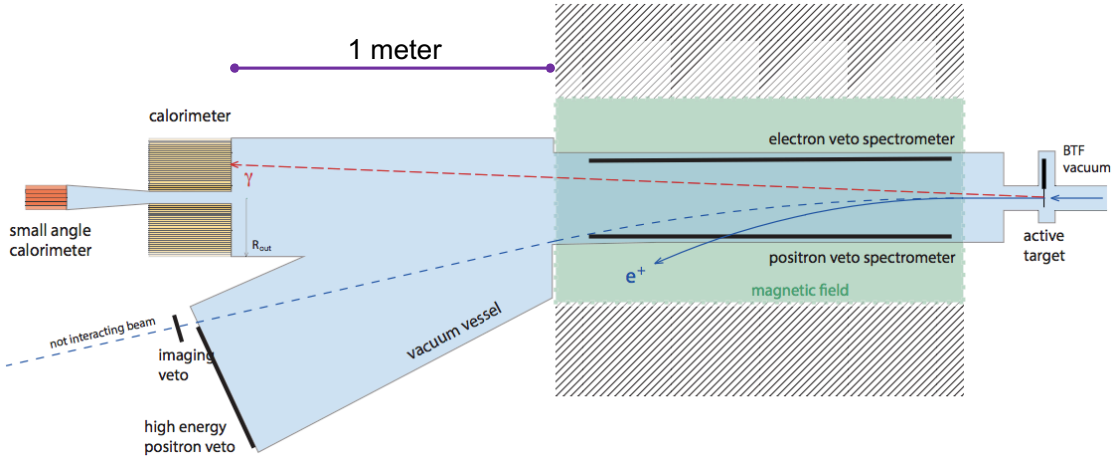


Figure 13.6: Experimental setup of the PADME detector. A 500 MeV positron beam strikes a carbon-based fixed-target from the right. The products of the interaction are measured by several detectors, such as the electromagnetic calorimeter, small-angle calorimeter, and positron vetoes. The entire detector is enclosed by a vacuum chamber.

Some beam-target interactions (such as the aforementioned Bremsstrahlung radiation) result in decreased energy of incident positrons. These positrons bend more sharply under the magnetic field than the non-interacting beam positrons and hit the walls of the magnet. Plastic scintillator vetoes are placed on both magnet walls to identify such events. Finally, the main calorimeter features a central hole that lets the abundant Bremsstrahlung radiation through. Behind the calorimeter and flush with it sits a small

Cherenkov calorimeter designed to be fast enough to tag individual Bremsstrahlung photons. This small-angle calorimeter (SAC) works as an additional, critical veto. The entire PADME setup is enclosed in a vacuum chamber to minimize air interactions.

Next, I briefly discuss the main components of the PADME detector.

13.3.1 Active target

Fig. 13.7 shows a photo of the active target prior to installation. The target is made of polycrystalline diamond grown by chemical vapor deposition and cut to transverse dimensions of $2 \times 2\,\mathrm{cm}^2$ with a $100\,\mu\mathrm{m}$ thickness. The front and back surfaces of the diamond film are fitted with 19 graphite strips each, oriented perpendicular to one another. The strips have a 1.9 cm length and 0.85 mm width, with a measured resistance of $2.5\,\mathrm{k}\Omega$ [118]. Charge released via ionization of carbon atoms by beam positrons is collected in these electrodes, providing a complementary measurement of the beam profile along both transverse directions. Furthermore, the target also provides a measurement of the beam multiplicity. The thickness and composition (carbon has $Z = 6$, where Z is the atomic number) are optimized to reduce Bremsstrahlung emission (which grows as Z^2) in relation to the expected signal cross section (which grows as Z).

The target design goals include a beam impact point precision of roughly 1 mm and a measurement of the luminosity at the percent-level resolution. Finally, the active target at PADME is the first fully carbon detector used in a high-energy physics experiment [118].

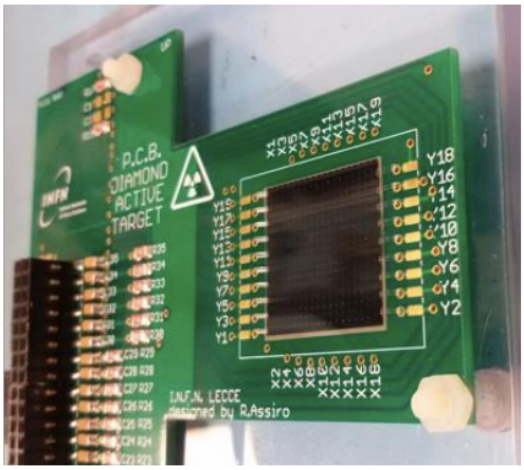


Figure 13.7: Photo of PADME's active target. The target is made of CVD-grown polycrystalline diamond and features several graphite strips to measure the transverse beam profile.

13.3.2 Magnet

The dipole magnet used in PADME is a MPB-S series magnet loaned from CERN's SPS transport line. It provides a 0.55 T magnetic field and has a 1 m length, 52 cm width and 23 cm vertical gap [119]. Photos of the magnet are shown in Fig. 13.8. The dipole aperture corresponds to an angular acceptance relative to the target of up to ≈82 mrad in the PADME geometry.

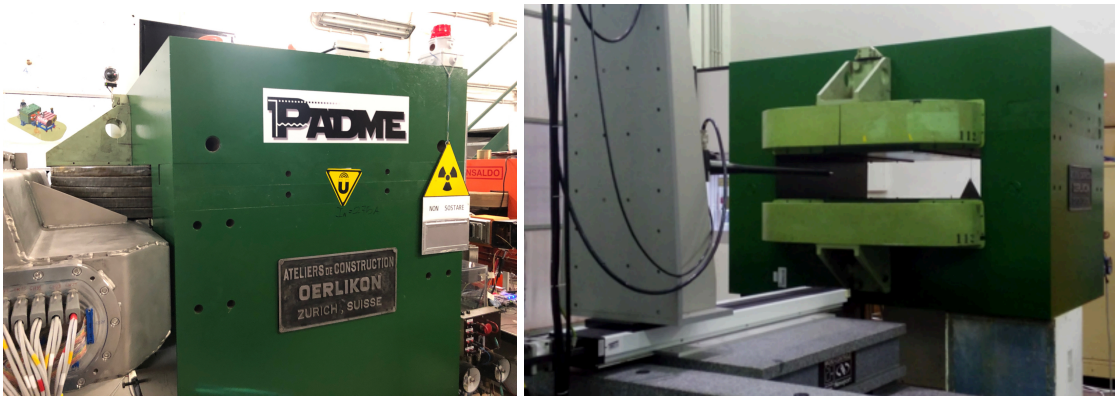


Figure 13.8: PADME's dipole magnet: back (left) and front (right). The magnet has a 0.55 T field and a 23 cm vertical aperture.

As mentioned, the walls of the magnet are instrumented with plastic scintillators, which serve as positron and electron vetoes. These are discussed in Section 13.3.5.

13.3.3 Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) is the main PADME detector. It consists of 616 bismuth germanium oxide (BGO) crystals arranged in a cylindrical configuration. The overall size of the calorimeter is fixed by the vertical aperture of the magnet and by the size of the experimental hall, which limits the target-ECAL distance to 4 m. The ECAL's front face is currently located 3.5 m away from the target. Crystals are refurbished from the old L3 experiment at CERN [120]. Each crystal is polished to have dimensions of $2.1 \times 2.1 \times 23\,\mathrm{cm}^3$. Crystals are coupled to PMTs, which convert the scintillation light into electric signals and ship them out to 12-bit digitizers. Fig. 13.9 shows sketches of the ECAL from the front and back.

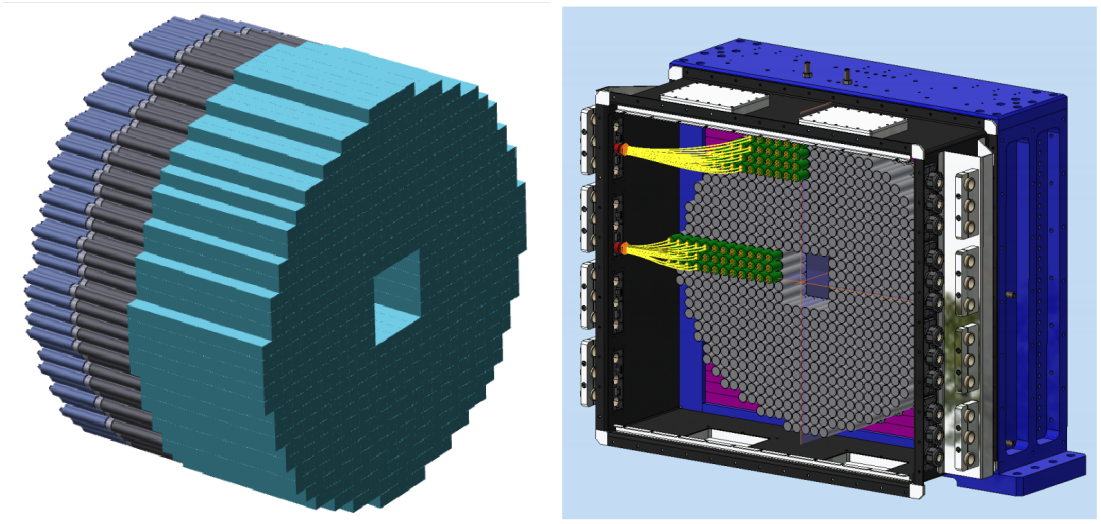


Figure 13.9: PADME's electromagnetic calorimeter (ECAL). CAD layouts are shown from the front (left) and front the back (right). The ECAL consists of 616 BGO crystals with a 2% energy resolution and a central square hole to remove Bremsstrahlung radiation from acceptance.

The ECAL has a central square hole roughly 25 crystals in area. This hole removes from acceptance a large fraction of Bremsstrahlung radiation, which is the largest background to PADME but has a sharp angular distribution around small angles from the beam axis.

The ECAL energy resolution was measured as [120]:

$$\frac{\sigma(E)}{E} = \frac{2.0\%}{\sqrt{E/\,\mathrm{GeV}}} \oplus 1.1\% \tag{13.2}$$

with a test beam campaign at Frascati. Moreover, the timing resolution is dependent on the scintillation properties of BGO. It takes on the order of hundreds of nanoseconds for the shower caused by an incident particle to fade away, making the timing resolution for single particles on the same timescale. This motivates the hole in the center, since it would be impossible for the ECAL to distinguish between individual Bremsstrahlung photons.

The digitized signals constitute hits that are processed by a clustering algorithm to reconstruct the full energy deposition of the incident particle spread between neighboring crystals and to determine the most likely impact position. More information about the ECAL reconstruction and performance can be found in Ref. [120].

13.3.4 Small-angle calorimeter

The SAC is an additional detector placed immediately downstream of the ECAL, flush with its central square hole. The goal of the SAC is to identify individual Bremsstrahlung photons emitted by incident positrons. This requires the SAC to have a fast response, motivating the use of Cherenkov radiation for the medium. Cherenkov radiation is

produced instantaneously as an incident particle penetrates a scintillating crystal.

The SAC consists of 25 lead fluoride (PbF_2) crystals (Cherenkov radiators) arranged in a 5×5 grid. Each crystal has dimensions of $3 \times 3 \times 14\,\mathrm{cm}^3$, which are optimized to the shape of electromagnetic showers developed inside PbF_2 crystals. More details are found in Chapter 14. Fig. 13.10 shows a photo of the SAC.

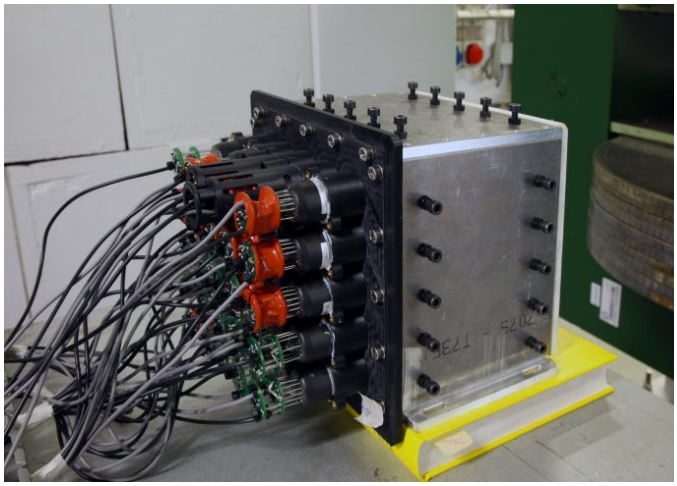


Figure 13.10: PADME's small-angle calorimeter (SAC). The SAC is a Cherenkov detector consisting of a 5×5 grid of PbF_2 crystals with very fast response.

Chapter 14 reports a detailed study of SAC prototype performance done with a test beam campaign in Frascati. This includes measurements of the energy response and resolution and more critically the expected timing capability of the SAC.

13.3.5 Vetoes

The vetoes consist of three sets of plastic scintillators strategically placed inside the PADME setup. Their main goal is to identify positrons that underwent Bremsstrahlung emission in the target and therefore lost energy, bending more sharply under the magnetic field. The positron veto (PV) is placed on the left wall of the magnet and consists of a

string of 96 plastic scintillator bars with dimensions $10 \times 10 \times 180\,\mathrm{mm}^3$ for a total length of 1 m alongside the wall of the magnet. The bars are oriented with their long side parallel to the magnetic field and rotated by roughly 5° to reduce geometric inefficiencies [121]. An additional 16 bars are placed further downstream, between the magnet and the beam dump, to measure higher-energy positrons (the high-energy positron veto (HEPV)). The scintillators are made of polystyrene-based material with 1.5% concentration of POPOP [121]. Optical wavelength shifter fibers are placed on grooves on the bars to match the optimal emission spectrum of POPOP ($\approx$400 nm). The fibers are connected to special PMTs that can operate in magnetic fields and in vacuum. Fig. 13.11 shows a photo of a veto prototype.

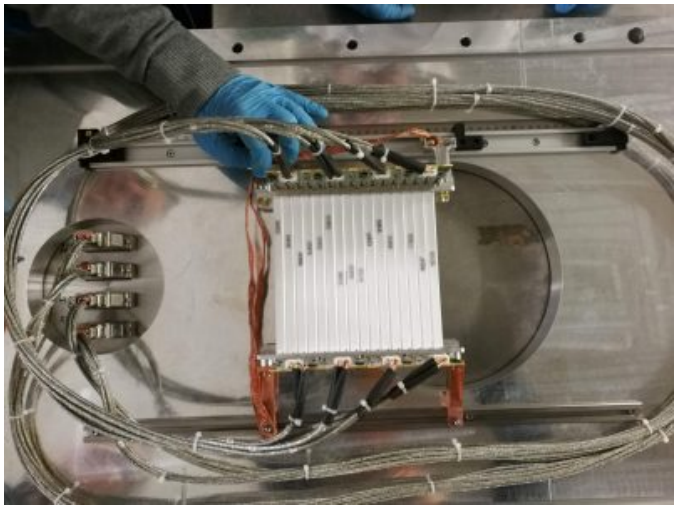


Figure 13.11: Prototype of PADME's plastic scintillator vetoes. The positron and high-energy positron vetoes are essential for tagging positrons undergoing Bremsstrahlung emission and removing them from consideration. They can additionally be used in searches for visible decays of dark photons.

Apart from the main PV and HEPVs for tagging radiating positrons, the right side of the magnet is also instrumented with the same scintillator bars. This electron veto (EV) is not so useful for tagging Bremsstrahlung backgrounds but enables an additional physics program to search for visible decays of dark photons. This is possible by identifying and time-correlating positrons and electrons with the PV and EV. This possibility is further

discussed in Chapter 15.

13.3.6 Timepix3

The Timepix3 integrated circuit is a dedicated beam monitoring device placed at the beam dump downstream of the magnet, next to the calorimeter. It consists of pixel silicon sensors with 256×256 pixels 55 µm in size each. The entire detector has 12 such sensors for a total of 786,432 pixels in a $8.4 \times 2.8\,\mathrm{mm}^2$ surface area. Its purpose is to monitor the non-interacting positron beam that gets diverted away from the calorimeter by the magnet. Precise knowledge of the beam divergence and transverse shape is essential for the missing-mass technique, since the positron four-momentum directly enters the calculation of the invariant mass. This device is currently the largest Timepix3 detector used in a high-energy physics experiment. Fig. 13.12 shows a photo of PADME's Timepix3 attached to the outside wall of the vacuum chamber.

Figure 13.12: Photo of PADME's Timepix3 beam monitoring device.

CHAPTER 14

SMALL-ANGLE CALORIMETER PERFORMANCE

This chapter is adapted from a paper published by the author in Ref. [122].

14.1 Introduction

PADME requires excellent background rejection capability. Among the largest sources of background are 2-gamma ($e^+e^- \to \gamma\gamma$) and 3-gamma ($e^+e^- \to \gamma\gamma\gamma$) events, where 1 or 2 gamma particles (respectively) escape detection via the hole in the ECAL. As mentioned in Chapter 13, to mitigate such backgrounds, a very fast calorimeter, the SAC, is placed behind the main ECAL, flush with the central hole. The in-time correlation of photon events in the SAC and ECAL allows the tagging of all gamma products and hence the efficient vetoing of such background events.

Here, we evaluate the performance of a prototype of the SAC with a test beam done at LNF using fast PbF_2 crystals and the newly developed Hamamatsu R13478UV PMT, optimized for fast response. We demonstrate that this detector meets the requirements for an efficient rejection of 2- and 3-gamma events, namely: (a) a timing resolution less than 200 ps for Cherenkov radiation detection; (b) moderate single-crystal energy resolution better than 10% at close to beam energy; (c) moderate light yield between 0.5 and 2 photo-electrons (p.e.)/MeV; (d) double-peak separation resolution capable of distinguishing several dozen photons in a 200 ns time window; (e) radiation hardness of order 1 Gy per 10^{13} positrons on target; and (f) acceptance of low-wavelength light due to the Cherenkov spectrum [123]. Furthermore, we encourage the investigation of a related (but more expensive) setup which uses two compact ultra-fast PMTs (R9880U-110) coupled to a single PbF_2 crystal in order to provide independent efficiency measurements

and timing references for higher-energy applications.

14.2 The PADME SAC

Fig. 13.6 shows the general layout of the PADME detector, including SAC placement behind the main calorimeter. The SAC consists of 25 PbF_2 crystals, each with transverse dimensions $30 \times 30\,mm^2$ and length 140 mm. The total transverse area is therefore $150 \times 150\,mm^2$, slightly larger than the central square hole of the ECAL. The non-interacting beam is diverted to an off-axis beam dump by means of a 0.5 T magnet. The photon rate in the central crystal due to Bremsstrahlung is expected to reach several hundred MHz, depending on beam intensity.

The lateral surfaces of each crystal are wrapped with 50 μm thick black Tedlar to minimize optical cross-talk and the back surfaces are coupled to Hamamatsu R13478UV PMTs via UV transparent optical grease with matching index of refraction for optimal light transmission. We describe below the investigation that led to the choices of crystal and PMT as it offers some interesting directions for future similar experiments.

14.3 Crystal/glass and PMT choices

Given the requirements for SAC performance outlined in Section 14.1, two options of radiating material were considered: SF57 (lead silicate, used for example in the Large Angle Veto of the NA62 experiment at CERN [124]) and PbF_2 (lead fluoride, used for example in the calorimeters of the Muon g-2 experiment at Fermilab [125, 126] and in the segmented calorimeter of the A4 experiment at MAMI [127]). Both materials are

suitable Cherenkov radiators due to their high refractive index, allowing for good timing resolution. However, SF57 has two (related) main disadvantages compared to PbF_2. The SF57 transparency window cuts off at 450 nm wavelength, whereas the Cherenkov spectrum is peaked at lower wavelengths due to its $1/E_\gamma$ energy dependence. This can be seen in Fig. 14.1, which shows the measured transparency profiles (measured by the Atomki Lab in Debrecen) for SF57 and PbF_2 compared with the Cherenkov spectrum. For this reason, SF57 offers a low light yield: in our preliminary tests, we obtained 0.15 p.e./MeV, which disfavors its use at energies below 1 GeV.[1]

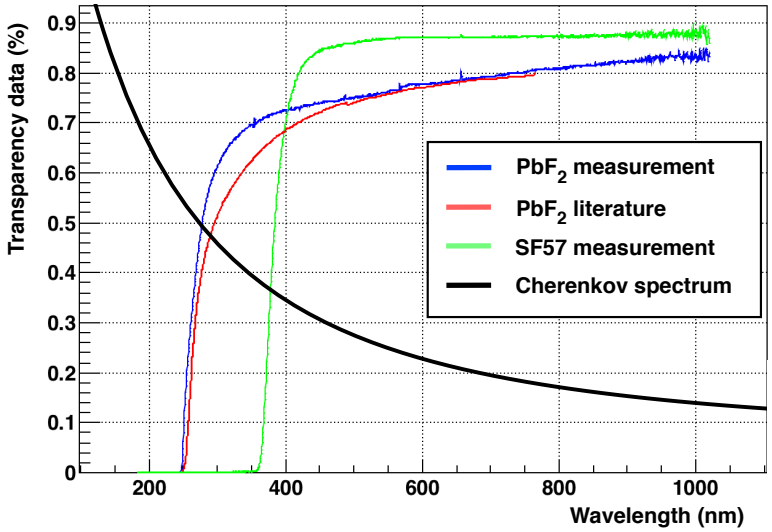


Figure 14.1: PbF_2 (red and blue) and SF57 (green) transparency profiles, compared to the Cherenkov spectrum (black, not to scale). For PbF_2, our own measurement (blue) agrees well with the literature (red) [128].

Furthermore, PbF_2 is denser and hence more compact, and 10x more radiation-hard. Table 14.1 contrasts some PbF_2 and SF57 properties which are relevant for Cherenkov-based calorimetry. PbF_2 has a shorter radiation length and smaller Moliere radius compared to SF57. Smaller electromagnetic showers reduce the detector occupancy, thus enhancing its rate capabilities. Hence, PbF_2 is the preferred solution for PADME's

[1]However, note that this measurement was also limited by the PMT's small dimension (model R9880U-110, described below).

requirements.

Table 14.1: Comparison between PbF_2 and SF57 of some properties relevant to Cherenkov calorimetry.

Property	PbF_2	SF57	Units
Density	7.77	5.51	g/cm^3
Radiation length (X_0)	0.93	1.54	cm
Moliére Radius	2.12	2.61	cm
Interaction Length (λ)	22.1	20.6	cm
λ/X_0	23.65	13.3	-
n	1.8	1.8	-

We also considered two candidates for PMT in the SAC: the ultra-compact Hamamatsu R9880U-110 and the Hamamatsu R13478 (both Q and UV versions). The first option is a compact PMT with only 16 mm in diameter (8 mm sensitive area) and a rise time of 0.57 ns and a transit time of 0.2 ns. The small size of this PMT allows the coupling of two such devices to the back of a single crystal. This could provide several benefits such as improved light yield and an independent time reference between the two PMTs. Furthermore, having two PMTs enables an efficiency measurement of each one separately.

The disadvantage of the R9880U-110 PMT is that its small dimensions limit single-PMT light collection efficiency. Without employing the two-PMT solution, a single R9880U-110 PMT has a cross sectional area of only 5.5% the crystal transverse dimensions. Even coupling two of them provides only 11% of geometric acceptance while doubling the cost. Given PADME's relatively low beam energy, this light yield is unacceptably small. Nevertheless, for higher energy applications, the two-PMT setup could be an interesting solution to explore.

The second candidate considered was the Hamamatsu R13478, Q and UV versions.

This PMT has a diameter of 26 mm (17.3 mm sensitive area) and thus covers a larger area fraction overall (26%). Compared to the R98880U-110, it has a similar rise time of 0.9 ns but a slower transit time of 9.1 ns. These specifications are fast enough for the PADME use case and the improved light collection efficiency enables only one PMT per crystal.

Fig. 14.2 illustrates some key measures of performance underscoring the determination of the optimal PMT. On the left are the quantum efficiency (QE) curves for the different PMTs, as well as the PbF_2 transparency profile and Cherenkov spectrum. The plot on the right shows the convolution of PbF_2, Cherenkov, and QE curves for each PMT. Due to the drop in PbF_2's transparency around a wavelength of 250 nm, there is no significant difference in performance between the more expensive R13478Q model and the more affordable R13478UV. The compact R9880U-110 actually performs better than the R13478's by a factor of roughly 30%, but its reduced light acceptance must also be accounted for as mentioned above.

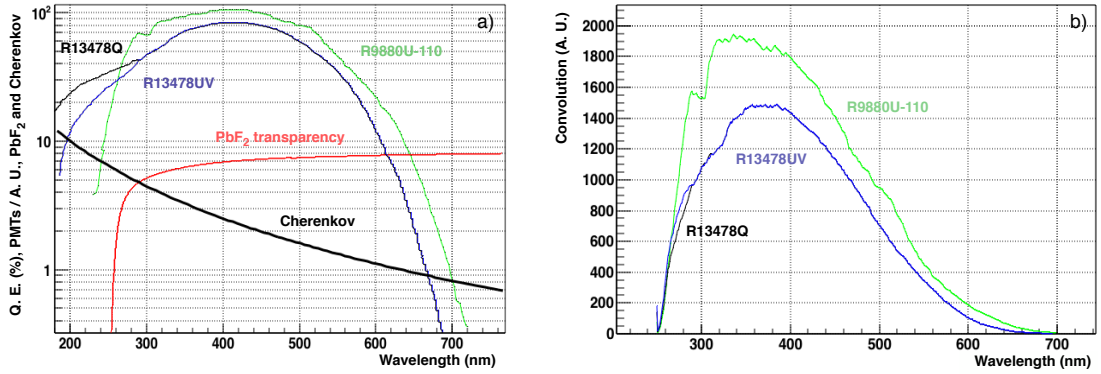


Figure 14.2: (a) Quantum efficiency (QE) curves for three different PMTs: R13478Q (black), R13478UV (purple), and R9880U-110 (green) [129, 130]. The Cherenkov spectrum (bold black) and PbF_2 transparency (red, [128]) are also plotted for comparison. (b) Convolution of each QE curve with PbF_2 transparency and Cherenkov spectrum. Due to the drop in transparency below 250 nm there is effectively no difference between the two R13478 PMT models. The R9880U-110 PMT has about 30% better performance over the entire spectrum.

To decide between the R9880U-110 and the R13478UV PMTs, we completed a study to understand the signal distribution expected for each. We considered two configurations: the R13478UV PMT with a tapered voltage divider, with typical gain G of 3.2×10^5, and the more compact R9880U-110, with a typical gain of 2×10^6. Both PMTs are assumed to be coupled to a crystal producing ≈20 Cherenkov photons reaching the photosensors face (N_{ph}) for each MeV of deposited energy E_{dep}, and to have average 20% QE. To account for differences between the two PMTs, two corrections are added depending on the PMT. The light acceptance due to photocathode size A_C corrects for the different active areas, while QE_{corr} corrects for the different integrated QEs (see Fig. 14.2). First we calculate the number of p.e. produced by each configuration:

$$N_{p.e.} = N_{\text{ph}} \times E_{\text{dep}} \times QE \times A_C \times QE_{\text{corr}}. \tag{14.1}$$

With these assumptions, we estimate the expected charge distribution from an incident particle based on a Gaussian spread model:

$$Q_{\text{tot}} = \text{Gaus}(N_{p.e.}, \sqrt{N_{p.e.}}) \times G \times e. \tag{14.2}$$

The results are shown in Fig. 14.3, with the charge distributions and corresponding estimated resolutions. Despite a higher gain, the R9880-110's small surface area limits light collection efficiency and hence the charge resolution. The R13478UV PMT offers in this configuration roughly a factor of 2 better resolution.

From these considerations, we conclude the larger Hamamatsu R13478UV PMT, coupled to PbF_2 crystals, is the best solution for PADME's requirements. This has acceptable timing resolution while providing sufficient light yield for the low beam energies in the experiment at a reasonable cost.

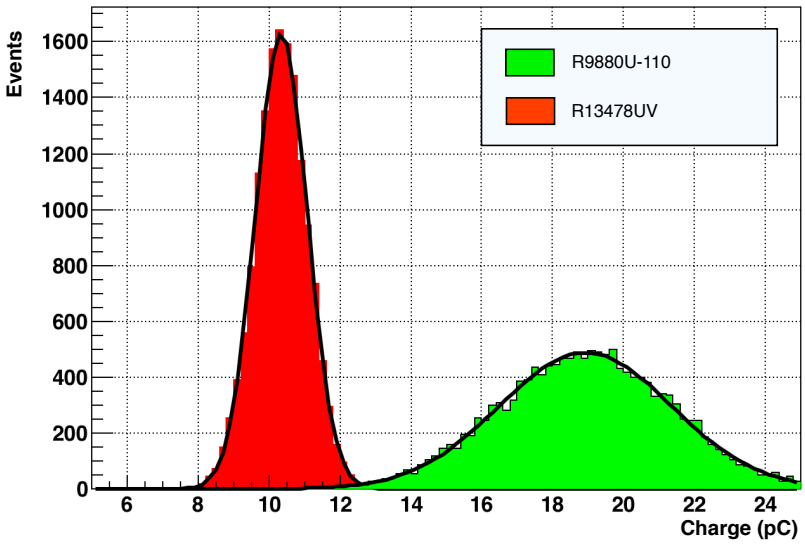


Figure 14.3: Simulation of the charge distribution expected for each PMT for a given deposited energy. R13478UV (red) and R9880U-110 (green) PMTs are compared.

14.4 Monte Carlo simulation and radiation damage

We developed a full MC simulation of the PADME detector with the GEANT4 framework [131] to study different aspects of the experiment. Three major goals discussed in this chapter include assessing the expected levels of absorbed radiation dose in the calorimeter, studying the detector response as a function of crystal properties, and correcting for the average shower energy leakage in our single-crystal test beam. In the first case, we simulate the entire physics, starting from beam positrons. In the second, we directly simulate high-energy photons incident on the crystals. And in the third, we simulate electrons striking a single crystal in order to match accurately the electron test beam results. In this section, we describe our radiation damage study, while the other two questions are addressed in later sections.

The full simulation contains all relevant components of the detector, such as target, magnet, ECAL and SAC, and veto spectrometer. For the radiation dose study, we gener-

ated 400 million events consisting of positrons with energy 550 MeV striking the target. The resulting photons (mostly from positron Bremsstrahlung emission, but also from 2- and 3-gamma production) then strike the calorimeters. The spectrum of Bremsstrahlung radiation is highly peaked at small angles, so the SAC receives most of the radiation. We then re-scaled the statistics of simulated data up to the expected integrated luminosity of 1×10^{13} positrons on target over the course of the first run of the experiment.

Fig. 14.4 shows the total expected radiation dose absorbed by SAC crystals. This estimate is made by assuming that the energy of striking photons is entirely transferred to the crystals, and that the energy deposition is uniform across the crystal. The dose is then obtained by dividing the total energy deposit by the mass of each crystal. The assumption of uniformity is not accurate, as most of the energy deposits happen early on during the shower development inside a crystal [132]. Nevertheless, we are only interested in an order-of-magnitude estimate of radiation dose to show that radiation damage on the crystals will be negligible throughout the experiment and that we can safely mitigate concerns about transparency losses.

There are several studies of transmission loss in PbF_2 crystals due to radiation damage in the literature, e.g. [133] (with protons) and [134] (with neutrons and gamma rays). We show the latter's PbF_2 transparency data in Fig. 14.5, where transmission efficiency as a function of wavelength is plotted before irradiation, and after 4 kGy and 40 kGy of combined irradiation with neutrons and gamma rays. Comparing this data with our SAC crystal doses, we therefore predict negligible transparency loss after the full run period of PADME data-taking.

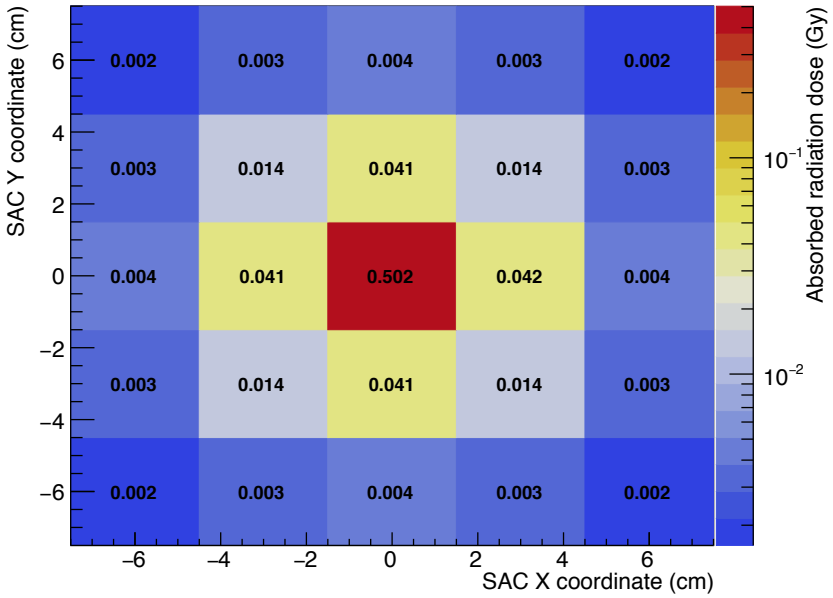


Figure 14.4: Expected radiation dose on SAC crystals from 10^{13} positrons on target, over the course of the experiment, as estimated by a GEANT4 MC simulation. The center crystal (which receives the most radiation due to Bremsstrahlung's sharply peaked angular spectrum) absorbs a dose of ≈1 Gy.

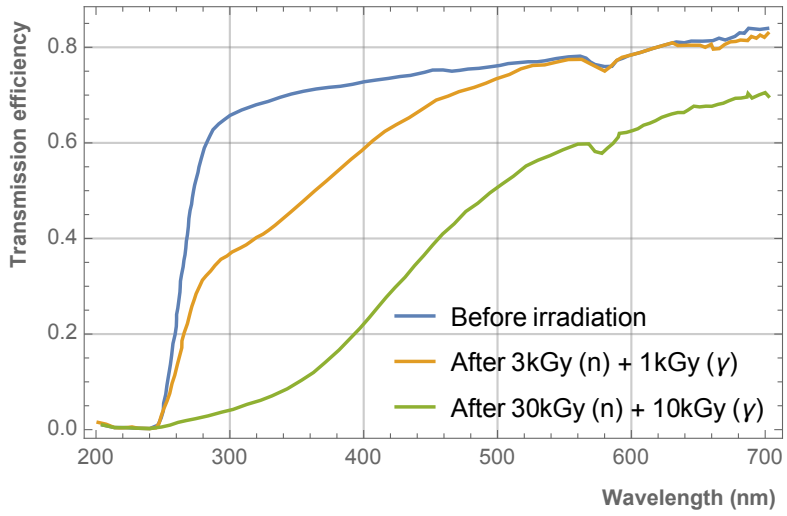


Figure 14.5: Transmission efficiency data before and after 4 kGy and 40 kGy of irradiation with neutrons and gamma rays [134], for a 1 cm^3 PbF_2 cube. Comparing with our expected radiation doses (Fig. 14.4), the transmission loss for PADME crystals should be negligible.

14.5 Optical Monte Carlo simulation

To establish our crystal specifications and to study different aspects of light collection efficiency, timing and energy resolution, we adapted our detector MC simulation to include a detailed optical description of physics inside the crystals. The primary goal is to investigate differences in light yield and photon arrival times as a function of crystal length and determine the optimal length. Manufacturing limitations restrict the crystal's length to 18 cm.

This simulation consists of a PbF_2 crystal with transverse dimensions $30 \times 30\,\mathrm{mm}^2$ and a variable length in the range 10 cm to 20 cm. It can be wrapped with either white millipore or black tedlar materials. The crystal is coupled to a thin layer of grease, modeled as Epoxy EJ-500 (radius 12.7 mm, thickness 1 mm), and then to a sensitive detector layer which models the PMT. The QE of the R13478UV PMT is implemented as a function of energy according to its datasheet [129]. Other optical properties are also implemented as functions of energy, such as absorption length from PbF_2 transparency [128] and refractive index via a parametrized dispersion formula [135]. These were calculated for optical photons with energy ranging from 1.6 eV to 5.0 eV in steps of 0.02 eV. Note that we do not simulate the entire detector here, just a single crystal in the SAC.

A single energetic photon (energy: 200 MeV) is fired at a distance of 1 mm from the crystal's front surface, which produces a few thousand optical photons after showering inside the crystal. The simulation then tracks each individual optical photon, until it either reaches the sensitive area corresponding to the PMT or gets lost along the way. We run 100 events per crystal length.

The arrival time distribution of photons for each length considered is displayed in

Fig. 14.6. Two peaks can be identified in each curve: a narrow, high peak which represents the arrivals of most photons, and a lower, broader peak, which corresponds to photons that underwent back-scattering inside the crystal. A shift and broadening of the narrow peak with increasing length is clearly visible.

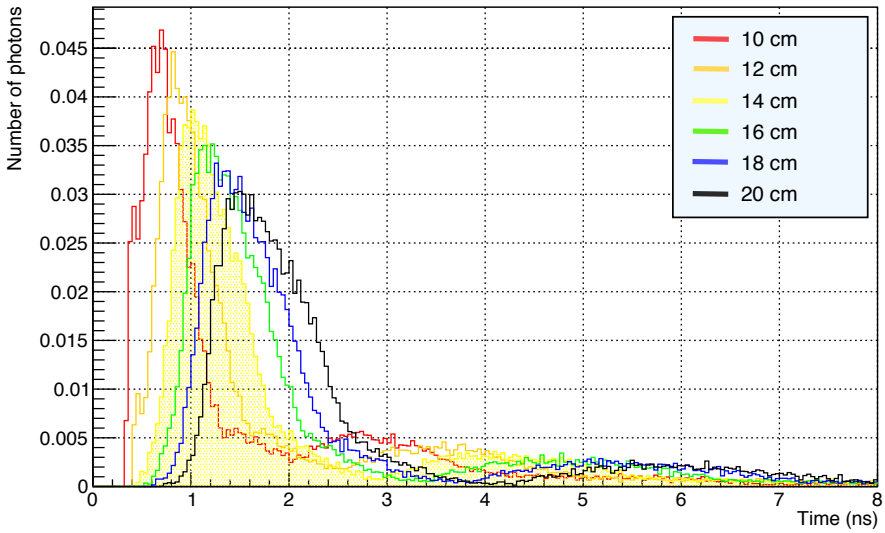


Figure 14.6: Arrival time distribution of Cherenkov photons versus crystal length for a 200 MeV incident photon. The distribution shifts to the right and broadens with increasing length.

The arrival profiles were fitted to a convolution of a Landau (narrow peak) and a Gaussian (broader peak) distribution, and the sigma of the Landau was taken as indicative of the double-particle separation capabilities of the detector in simulation. The choice of a Landau fit was empirically driven and not based on underlying physical processes. We do not account for effects from the PMT itself, but those should be of second order and so this procedure at least allows the relative comparison of different crystal lengths. The simulation Landau spread time is plotted in Fig. 14.7. There is a roughly 16% increase when going from 14 cm to 18 cm.

Fig. 14.7 also shows the light yield as a function of crystal length. The light yield is determined from a convolution of the PMT's QE with the energy distribution of arriving

photons. The light yield decreases with increasing length due to a higher chance of Cherenkov photons getting lost (absorbed or escaping) while traveling towards the PMT. In particular, there is a drop of about 14% in collected light between lengths of 18 cm and 14 cm.

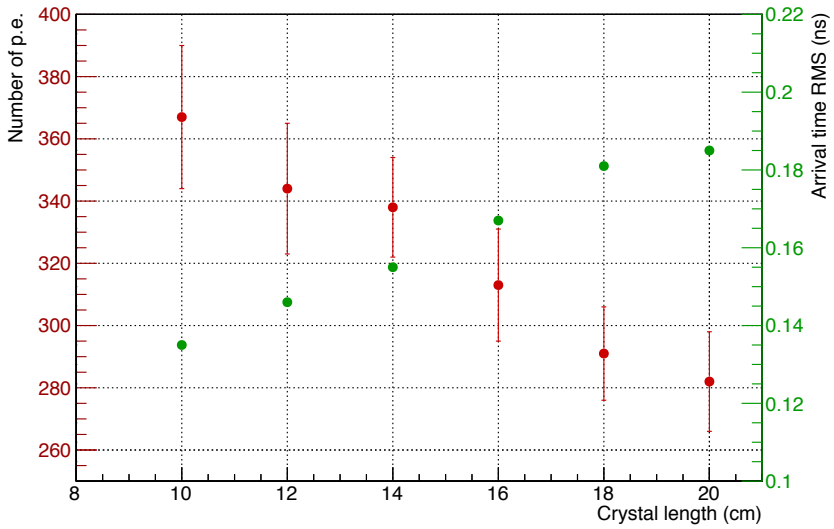


Figure 14.7: Light yield (red) and Landau time spread (green) versus crystal length for a 200 MeV incident photon, based on a detailed optical MC simulation. The light yield is estimated by convolving the energy distribution of arriving photons with the PMT's QE profile, and the timing resolution by a Landau fit of the arrival time distribution.

From the optical MC simulation, it is evident that shorter crystals are better, both for optimal light yield and for timing considerations. Increasing the length causes more Cherenkov photons to be lost along the way. At the same time, in a longer crystal photons have longer travel paths. Consequently, there is more opportunity for scattering, which broadens the arrival time profile of photons. Thus the simulation reveals that increasing the crystal length leads to broader signals and longer tails, affecting the double-peak separation capability, which is a crucial parameter for PADME SAC performance.

Nevertheless, there is a length limit where the crystal is too short to develop a complete Cherenkov shower. From [128], about 50% of Cherenkov photons from a 3 GeV shower are produced within the first 60 mm of a PbF_2 crystal and about 80% are produced within

the first 100 mm. This makes 140 mm a safe length to ensure we can harness most of the Cherenkov photons in a 500 MeV shower.

14.6 Test beam setup

The chosen crystal and PMT options were subjected to a test beam at LNF's BTF, in order to characterize the SAC response and measure energy and timing resolution, including double-peak separation capability. We obtained PbF_2 crystals on loan from Brookhaven National Laboratory for this test with dimensions $30 \times 30 \times 140\,\mathrm{mm}^3$.

The BTF at LNF is part of the DAFNE accelerator complex. A LINAC provides bunches of 10^{10} electrons or 10^9 positrons with energy up to 750 MeV or 550 MeV, respectively. The LINAC has to switch between electron and positron modes on a regular interval, in order to top off the main DAFNE rings, and in this case the energy is fixed to $E_0 = 510$ MeV for both charges. There are 50 bunches per second exiting the LINAC, which are shared between the BTF and the main ring under normal operation. In this regime, the BTF gets 38 bunches/s in electron mode and 18 bunches/s in positron mode. Each mode lasts 180 s and then a 90-second switch mode is activated. During the switch mode, there are 60 s with no beam at all and then 30 s where the BTF gets the entire 49 bunches/s (one per second is used for beam-energy monitoring). The cycle then repeats.

The BTF setup also allows the tuning of the beam intensity delivered to the experimental hall, from $\sim 10^{10}$ particles/bunch down to a single particle/bunch by intercepting the primary beam with a 2 X_0 target. The resulting secondary particles can be further filtered to allow tuning of the beam energy from E_0 down to few tens of MeV. The beam spot and position can be adjusted by means of quadrupoles, dipoles, and correctors in

the BTF line, and is monitored in real-time by silicon pixel hybrid detectors (FITPIX [136]) with active area $14 \times 14\,\mathrm{mm}^2$ and 55 µm pitch.

A schematic of the detector test setup is shown in Fig. 14.8. A R13478UV PMT (voltage: 1,600 V) was coupled to a PbF_2 crystal using optical grease and then connected to a 12-bit, 5 GSPS, 1024-sample digitizer (model CAEN V1742) for data acquisition. A plastic scintillator coupled to two small "finger" PMTs (model R9880U-110) provided a reference signal for comparison with the PbF_2 one.

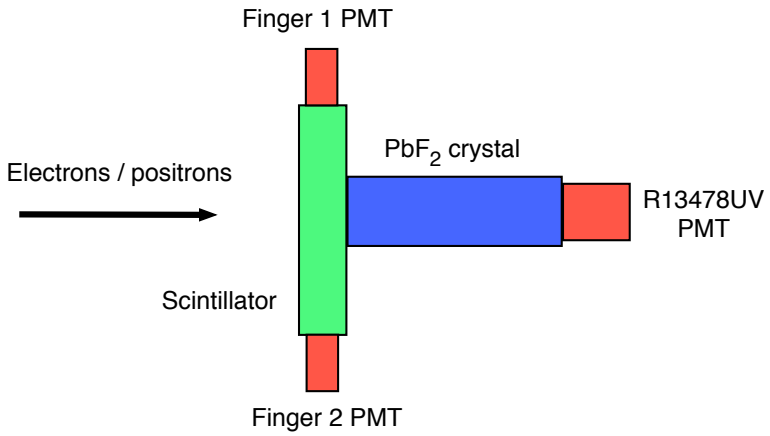


Figure 14.8: Schematic layout of the test beam detector setup. A PbF_2 crystal is coupled to the R13478UV PMT via optical grease. Two compact finger PMTs connect to a plastic scintillator bar and provide a coincidence reference signal. The operating voltage of the PMT is 1,600 V.

A picture of the PMT and crystal setup is shown in Fig. 14.9.

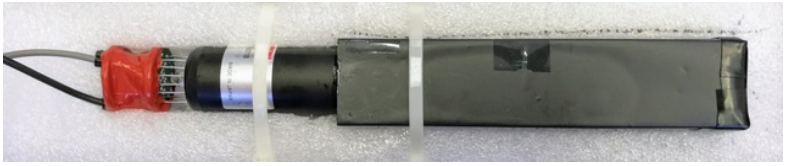


Figure 14.9: Photo of the PMT and crystal setup for the SAC test beam. A PbF_2 crystal is coupled to the R13478UV PMT via optical grease.

For the studies presented here, the beam energy was varied between 100 and 400 MeV in steps of 100 MeV and the average number of particles per pulse delivered to the BTF

experimental hall was set to ≈1 for electrons. The beam was centered on the crystal and the spot was kept within a standard deviation of 3 mm (at 400 MeV) and 5 mm (at 100 MeV) in the transverse plane.

14.6.1 Single-crystal Monte Carlo simulation

Since only one PbF_2 crystal was available for testing, we adapted our dedicated MC simulation to characterize the average energy leakage and provide a correction factor to light yield and energy resolution measurements.

The simulation in this case consists of a single PbF_2 crystal with the same dimensions as the one used at BTF and a beam of electrons with similar energy and multiplicity as the test beam. The primary goal was to estimate the average fraction of incident energy that escapes the crystal, decreasing the energy collection efficiency. This is a purely geometrical correction which can be used to re-scale the light yield and energy resolution obtained from test beam data.

Fig. 14.10 shows the average fraction of deposited energy in the crystal as a function of incident electron energy. For each energy, 10k events were simulated and the distribution of deposited energy was fit to a Crystal-Ball (CB) function [137]. The extracted mean was used as a data point in Fig. 14.10 and the extracted sigma as the uncertainty in that value.

The plot was fit to the empirical function:

$$f(E) = \sqrt{c_0^2 + \frac{c_1^2}{E/\mathrm{GeV}}}, \tag{14.3}$$

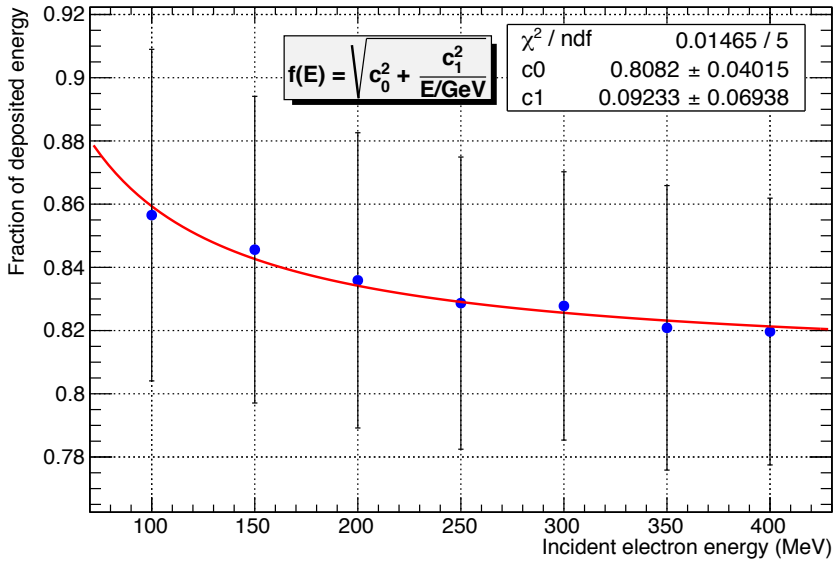


Figure 14.10: Fraction of deposited energy on a $30 \times 30 \times 140\,\mathrm{mm}^3$ PbF_2 crystal with a high-energy incident electron, as determined by a Geant4 MC simulation. Each data point is the mean of a Crystal-Ball fit to the distribution of deposited energies with 10k events.

where $f(E)$ is the average fraction of energy deposited on the crystal.

Note that this procedure can only correct for average energy leakage and does not account for experimental fluctuations in that quantity. Therefore the energy resolution quoted here is only an upper limit to the actual one achievable by PADME, and should be interpreted as a single-crystal SAC energy resolution.

14.7 Charge reconstruction

The results presented here are based on data taken in July of 2017 at the LNF BTF. Electron beams with average multiplicity of 1 particle/pulse and energies of 100, 200, 300 and 400 MeV impinged on the detector setup. Data acquisition triggered on accelerator signals and 1024 samples at 5 GSPS (i.e. 0.2 ns/sample) were collected per trigger [138]. Since PbF_2 has a fast Cherenkov emission of less than a few nanoseconds, the signal is

centered within a small time window of the waveform.

The total energy deposited in the crystal was reconstructed in three steps. First a run-level pedestal was calculated by averaging the analog-to-digital converter (ADC) counts of each 1000-sample empty event in the run (i.e. only noise, no signal peaks). The last 24 samples of each event were not used. The average ADC counts for all empty events form a Gaussian distribution, the mean of which is taken as the pedestal for that run. The sigma in turn informs the noise level, which was around 1.1 pC for all runs. This is the noise considering all 1000 samples, but for a roughly 50-sample signal window the average noise is scaled down to roughly 0.1 pC.

After subtracting the pedestal from each ADC count, the integrated charge was calculated by identifying all signal peaks in a given event, and integrating the area underneath each peak. The peak boundaries were set via a simple threshold ($|\text{ADC}-\text{pedestal}|/4096 > 0.005$). A sample digitized trace, with two electron peaks and thresholds identified, is shown in Fig. 14.11.

The choice of threshold was made to mitigate the after-pulse ringing that can be seen in the figure. This ringing continues with approximately constant magnitude for about 40 ns after the end of a pulse. Immediately following each pulse, the signal shoots below the pedestal-subtracted zero level, which motivates the low threshold in order to exclude such ringing effects from the charge estimate.

Finally, the distribution of integrated charges was plotted. A representative example, for 300 MeV electrons, is shown in Fig. 14.12. The peaks were fit to a sum of CB functions in order to extract the mean and sigma.

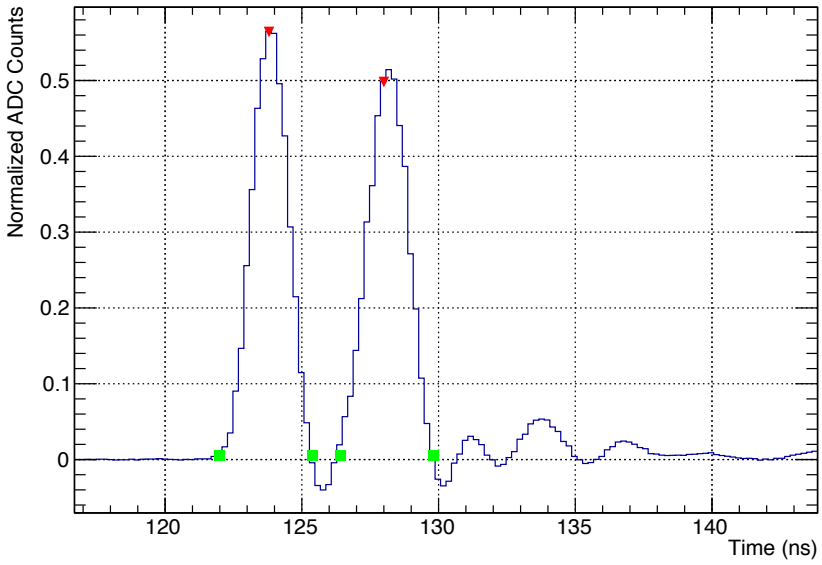


Figure 14.11: Sample digitized trace with two electron peaks. The peaks are identified with a custom algorithm and confirmed with ROOT's *TSpectrum* class. For each peak the thresholds are determined by using the after-pulse ringing and setting a low threshold (but still above the average noise level). The time between samples is 0.2 ns.

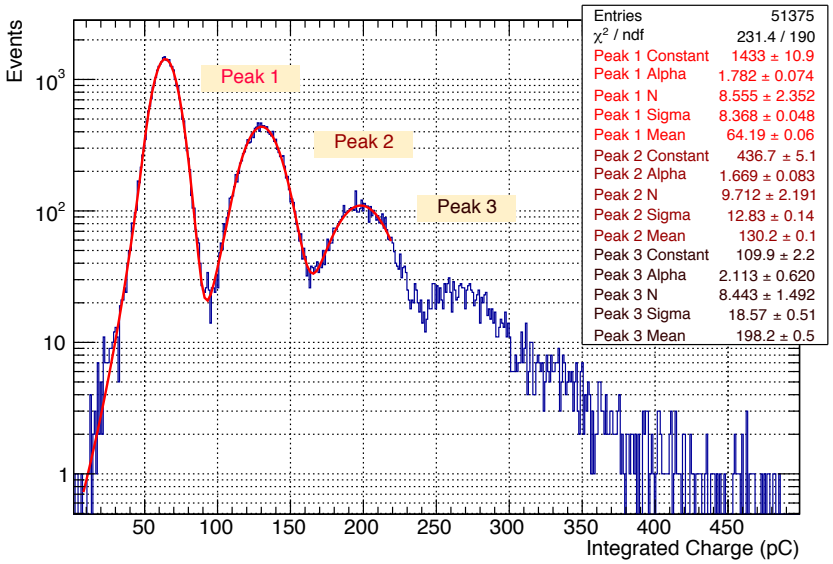


Figure 14.12: Integrated charge distribution for 300 MeV electrons. The first 3 peaks were fit with a sum of Crystal-Ball functions for extraction of means and sigmas.

14.8 Detector performance

To estimate single-crystal performance, we measured linearity and energy resolution using only single-electron events at different incident electron energies. We also determined the timing resolution by using the auxiliary finger PMTs as reference and performed a data-driven study of our estimated double-peak separation capability. These studies are discussed below.

14.8.1 Linearity and light yield

Fig. 14.13 shows the collected charge as a function of deposited energy. Each data point is the mean of the corresponding CB fit. The fit is performed only on single-electron peaks (labeled 'Peak 1' in Fig. 14.12), though we also plot data points corresponding to multiple-electron peaks for completeness ('Peak 2' and 'Peak 3'). Since a typical single-electron pulse has a short duration and given the sampling resolution of 0.2 ns, only rarely do two or more electron pulses exactly overlap in time, and so multiple-electron events are not appropriate indicators of the charge linearity or energy resolution.

The plot in Fig. 14.13 is a function of the actual deposited energy on the crystal, and not of incident beam energy. This is done to account for the fact that a single finite-sized crystal does not provide full energy containment. The correction used is shown in Fig. 14.10, as described in Section 14.6.1.

From the slope of the solid fit line in Fig. 14.13, we find a light yield of 2.05 p.e./MeV. Note that the linearity is valid only for the single-peak regime, which is what we are interested in. For double-electron peaks or higher, the ringing after-pulse is not adequately captured by our pulse-area calculation if the two peaks are close enough in time. This

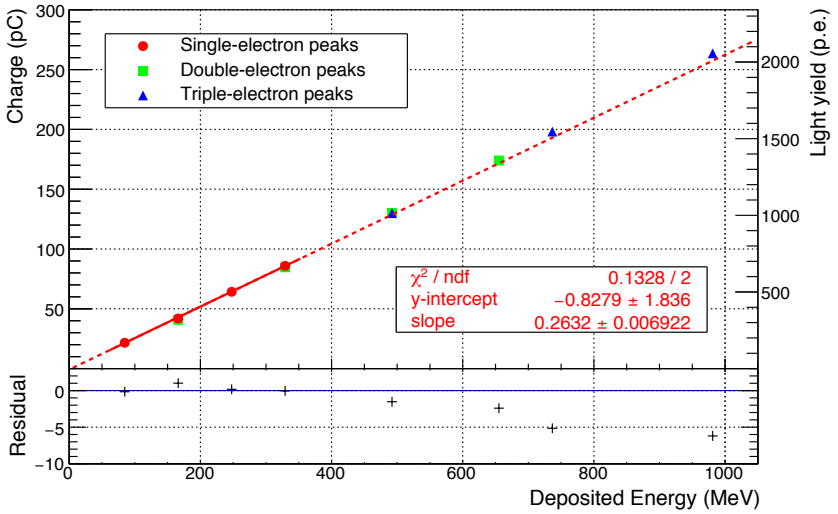


Figure 14.13: Detector linearity as a function of deposited energy (corrected for average leakage energy, see Fig. 14.10). Only single-electron peaks are fit. The residuals are also shown, demonstrating good linearity up to at least 400 MeV. A bias in the linearity is introduced for multiple-electron peaks due to the nature of our threshold setting and the after-pulse ringing, but does not affect the resolution since only single-electron peaks are used. The light yield is shown on the right, assuming a gain of 8×10^5 for a PMT operating voltage of 1,600 V [129]. The obtained light yield is 2.05 p.e./MeV.

overestimates the total charge and introduces a bias in the linearity.

14.8.2 Energy resolution

The energy resolution was calculated as the fitted sigma over mean of single-electron peaks in the charge distributions (e.g. Fig. 14.12). Fig. 14.14 shows the achieved energy resolution as a function of deposited energy on the crystal.

The data were fitted according to the expression:

$$\frac{\sigma_E}{E} = \sqrt{p_0^2 + \frac{p_1^2}{E/\mathrm{GeV}}}, \tag{14.4}$$

where p_1 and p_2 are two fit parameters. Note that the amount of deposited energy on the

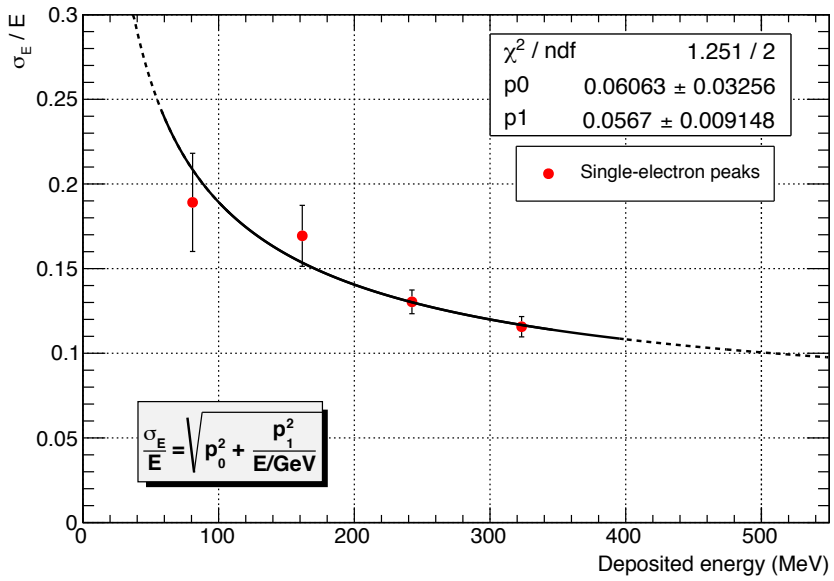


Figure 14.14: SAC energy resolution as a function of deposited energy. The amount of deposited energy is calculated from the incident beam energy using $f(E)$ (fraction of deposited energy) as computed with the MC simulation (Fig. 14.10). The fit yields a resolution of roughly 10% at 550 MeV, which meets PADME's requirements.

crystal (the values on the x-axis) is obtained from the incident beam energy by applying the function $f(E)$, from Fig. 14.10. We emphasize that this procedure can only correct for *average* energy leakage, but not for *fluctuations* in that quantity, which adds a significant contribution to the energy resolution.

The chosen fit function is an ad-hoc attempt to capture the overall behavior of the energy resolution curve, and not specific contributions to it. Major sources of fluctuations in the energy resolution include the aforementioned shower leakage (lateral, longitudinal, and albedo), stochastic fluctuations in the number of photo-electrons, and electronics noise. We estimate that the dominant contribution to shower leakage is the lateral lack of containment, which will be mitigated with a full calorimeter comprised of 5×5 crystals, thereby improving the energy resolution. Nevertheless, we find that with a single crystal, the energy resolution is roughly 10% at 550 MeV, which already meets PADME's SAC requirements.

14.8.3 Timing resolution

The timing resolution of the PbF_2 + R13478UV PMT setup was determined with help from the scintillator bar and finger PMTs. First a cut was imposed on the integrated charge to select events with only one electron in them. For each selected event, the rising edge of the electron pulse was fit to a straight line, using as endpoints the 20% and 80% heights of the pulse amplitude. The location of the fit at 50% height was then taken as the reference time. This procedure was done for all three channels: PbF_2 and the two finger PMTs. The distribution of time differences between each channel was plotted. Two examples can be seen in Fig. 14.15.

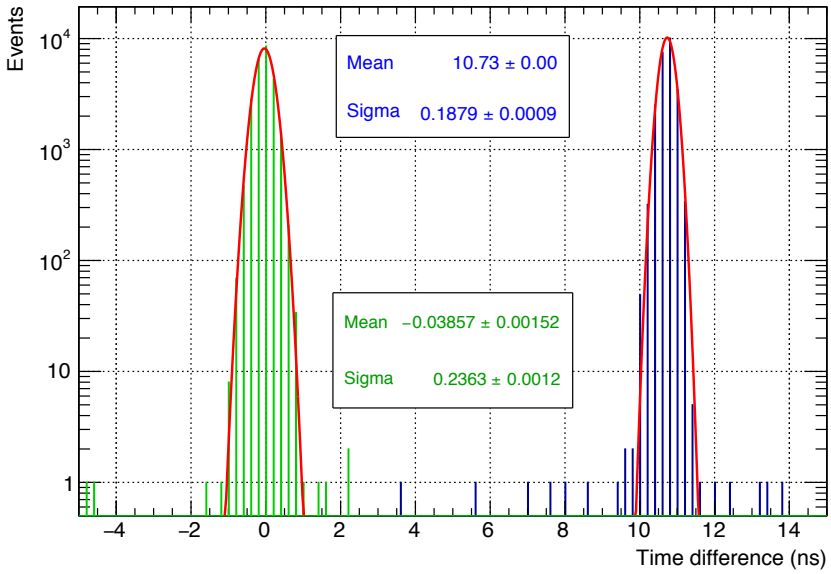


Figure 14.15: Difference in rise time between PbF_2 and finger 1 channels (blue) and finger 1 and finger 2 channels (green), for a 300 MeV beam. The measurement is taken at 50% of the amplitude according to a linear fit in the 20-80% range.

The timing resolution was determined by extracting the sigmas of a Gaussian fit to each such distribution. From the two-finger distributions, the finger resolution was determined to be 174 ps. Then, from the other distributions, the PbF_2 resolution was extracted by inserting the calculated finger resolution and summing in quadrature. The resulting resolutions for all channels and runs are shown in Fig. 14.16. The PbF_2 resolu-

tion was found to be 81 ps, which comfortably meets the 200 ps timing requirement for a successful SAC performance.

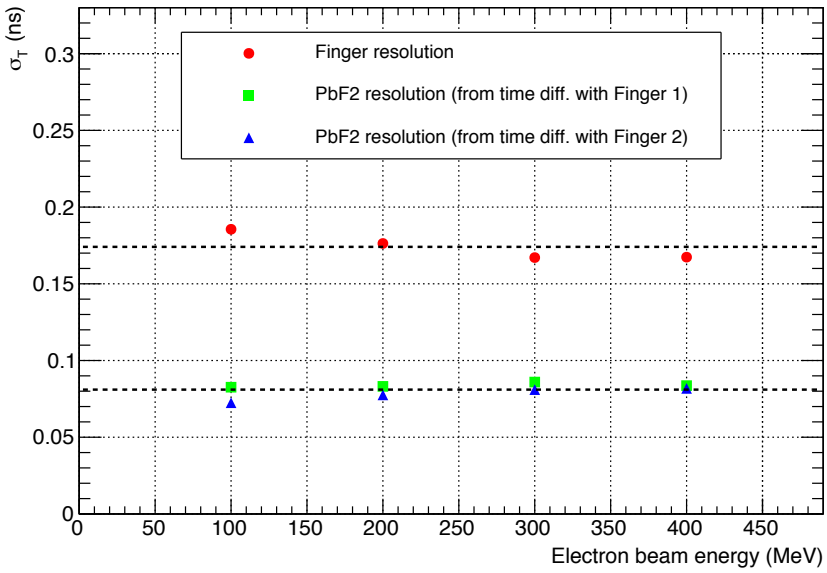


Figure 14.16: PbF_2 and finger timing resolutions, calculated from the difference in rise time between the different channels. The SAC timing resolution of roughly 80 ps meets PADME's requirements. Note that timing resolution is not significantly affected by beam energy due to our method of estimating the arrival time. This method relies on calculating the slope of the rising pulse, whose midpoint does not vary much with amplitude.

14.8.4 Double-peak separation resolution

In addition to timing resolution, we measured the double-peak separation capability of the setup using a purely data-driven method. From the charge distribution, a cut was imposed to select only single-electron events. Traces of selected events were then randomly added in pairs, with an artificially introduced time separation between single-electron peaks τ. By varying τ, we simulated increasingly overlapping pulses. For each value of τ, the peak separation was measured by means of the ratio between the height of the separation trough (h), and the height of the smaller of two peaks (H). This ratio, $r = h/H$, characterizes the degree of separability between the two peaks. A value closer

to 1, for example, implies the two peaks are very close together and distinguishing them is more challenging. On the other hand, a ratio closer to 0 means the peaks are far apart and identifying them is straightforward. A sample trace showing the relevant quantities defined above can be seen in Fig. 14.17.

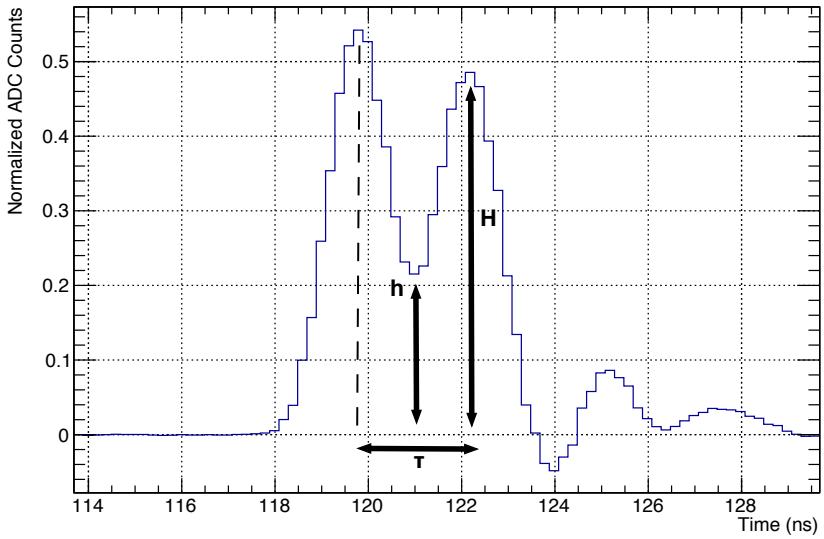


Figure 14.17: Sample trace used in the data-driven estimate of double-peak separation resolution. Two single-electron waveforms are overlapped with an artificially introduced separation τ. The ratio $r = h/H$ between the inter-peak trough height and the second peak height characterizes the separability of the two peaks.

Fig. 14.18 shows the distribution of ratios r for a 300 MeV beam. The white dots correspond to the 90-th percentile of the ratio distribution for each τ. Assuming a minimum separation capability of $r < 0.95$, then roughly 90% of peaks with 1.8 ns separation can be distinguished. As a conservative estimate we take this to be the double-peak separation resolution.

With a more sophisticated algorithm (for example, template fitting), it is likely that this capability can be improved further. Even when there is no local minimum (i.e. the peaks are too close together), in which case this algorithm fails, template fitting might still be able to identify a broader shoulder as a second peak. Nevertheless, this separation resolution already meets the PADME tagging requirements.

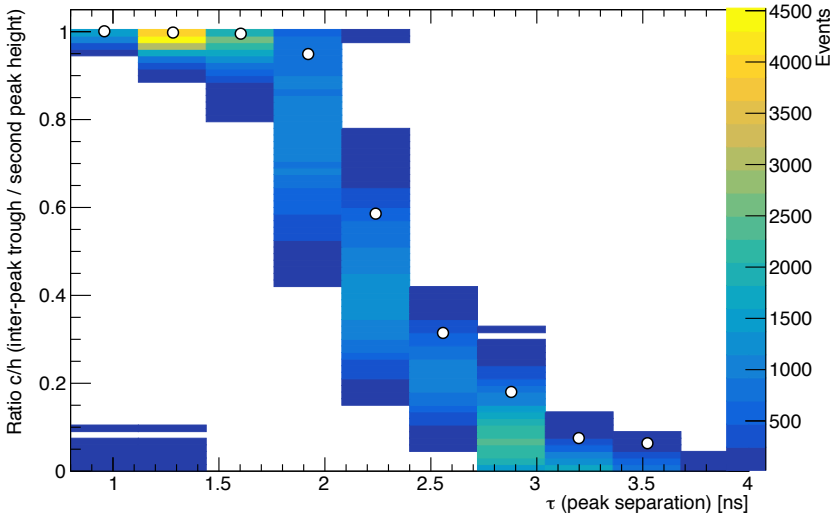


Figure 14.18: Ratio $r = h/H$ for different values of τ. This ratio informs the separability between double-electron peaks, with a value near 1 (0) implying challenging (easier) separation. The white dots represent the 90-th percentile of each distribution.

14.9 Conclusions

We have characterized the performance of a prototype of the small-angle calorimeter in the PADME experiment. Part of that characterization stems from a test beam with a single-crystal prototype of the SAC performed at LNF's BTF facility, in Frascati, Italy. The energy, timing, and double-peak separation resolutions all meet the demanded specifications in order to effectively tag and veto 2- and 3-gamma events from positron-electron interactions. The PbF_2 crystal coupled to the newly developed Hamamatsu R13478UV PMT has been found to provide a light yield of 2.05 p.e./MeV, an energy resolution of 10% at 550 MeV, and a timing resolution of 81 ps, with a double-peak separation capability of 1.8 ns. The robust double-peak resolution achieved suggests that the present setup might be suitable for a variety of high-intensity applications, being able to cope with rates higher than 100 MHz.

We also investigated other possible detector solutions, and in particular we recommend the further consideration of a two-PMT setup using the ultra-compact Hamamatsu

R9880-U110. This would allow an independent measurement of each PMT efficiency and provide a separate timing reference. This solution is anticipated for higher-energy applications, in which light yield is not a crucial factor for a satisfactory performance.

CHAPTER 15

PADME PROSPECTS

The PADME setup offers additional venues for exploration of a complex dark sector that go beyond the main missing-mass technique. There are strong theoretical and experimental reasons to study alternative models of dark matter. This chapter presents the projected sensitivity of the dark photon search and discusses some of the extra capabilities of PADME, laying out ideas for future dark sector searches within the PADME data. It also explores possible upgrades of the detector and accelerator setup that have been proposed within the Collaboration and the larger fixed-target community.

15.1 Projected sensitivity

Fig. 15.1 shows a projected sensitivity contour for PADME under the assumption of 1×10^{13} POT. The horizontal axis displays the dark photon mass and the vertical axis the kinetic mixing coefficient. Several existing and projected constraints from other experiments are also shown. In its first iteration, PADME can constrain regions of phase space down to a kinetic mixing coefficient of $\approx 5 \times 10^{-5}$ and up to a dark photon mass of $\approx$24 MeV. With several proposed upgrades, some of which are described below, the sensitivity may be enhanced substantially.

The NA64 experiment applies a missing-energy technique to search for dark photons [139]. The experiment strikes a 100 GeV electron beam onto an active target that is also an electromagnetic calorimeter. The calorimeter measures the electron energy, which allows the estimation of any missing energy carried away by an invisibly-decaying dark photon produced via Bremsstrahlung emission. Recently released results set a constraint of $\epsilon < 5 \times 10^{-4}$ at 10 MeV [140].

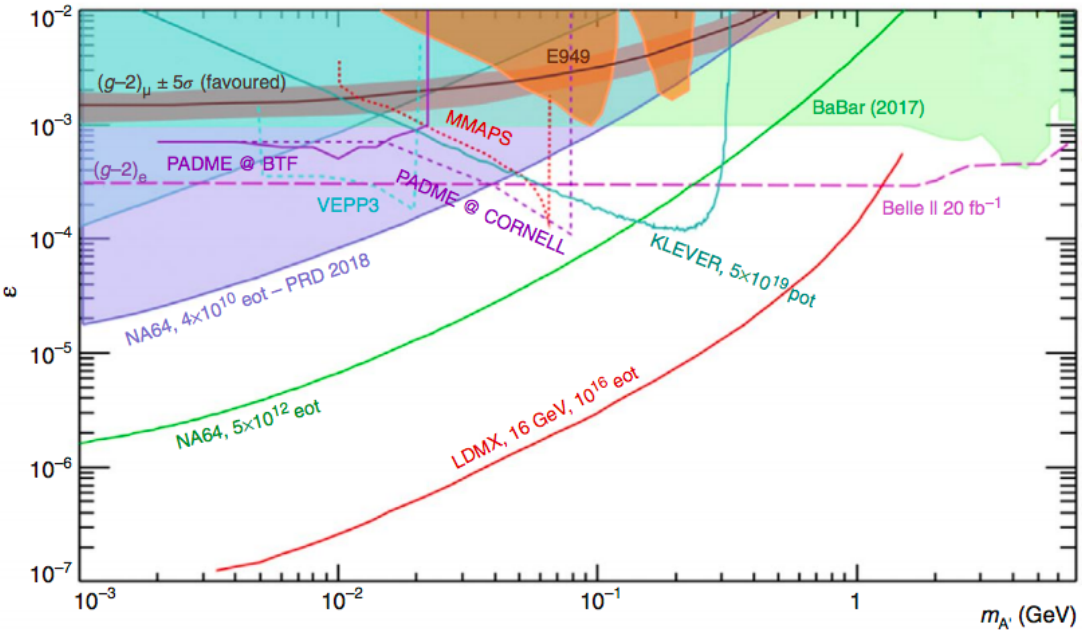


Figure 15.1: Projected sensitivity of PADME's first run and several other accelerator-based experiments [55]. The horizontal axis is the dark photon mass and the vertical axis the kinetic mixing coefficient.

15.2 The Beryllium and Helium anomalies

In 2016, an atomic spectrometry group from Atomki Lab, in Hebrecen, Hungary, first reported an unexpected excess in the angular distribution of electron-positron pairs from internal pair creation (IPC) following the excitation of Beryllium nuclei by proton bombardment [113]. This process is depicted in Fig. 15.2 and the observed angular distribution of the electron-positron pair is shown in Fig. 15.3. According to the group, this anomaly does not have any known physics explanation and all systematic effects considered have been ruled out. Moreover, the invariant mass distribution of the electron-positron pair also reveals a bump around 17 MeV with very high statistical significance, as seen in Fig. 15.4. They argue that this provides evidence of a new boson with mass roughly 17 MeV. The boson has been referred to as the X17 particle in popular press.

More recently, in 2019, the same group announced the observation of a similar excess

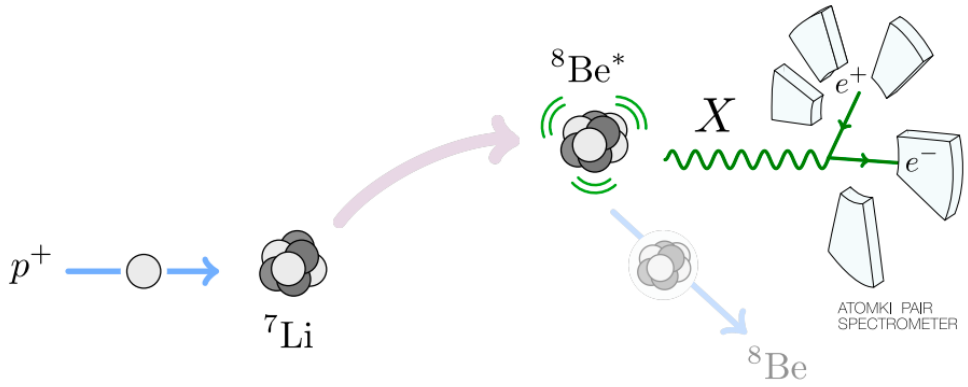


Figure 15.2: Sketch of the experimental strategy of the Atomki group [141]. Lithium nuclei are bombarded with protons and excited to a higher-energy state of Beryllium. The Beryllium subsequently decays back to its ground state via emission of an electron-positron pair in a phenomenon known as internal pair creation. The angular correlation and invariant mass of the pair are measured with a five-leg spectrometer.

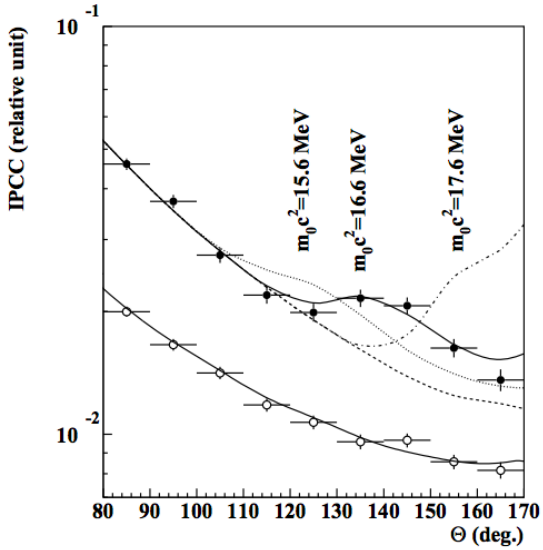


Figure 15.3: Angular correlation of electron-positron pairs in the de-excitation of Beryllium nuclei following proton bombardment of Lithium atoms [113].

in a different system. This time, they bombarded Helium nuclei with protons and observed IPC positron-electron pairs produced via an alternative de-excitation path [114]. There is still an unexplained excess of events at large angles in the pair angular distribution and the invariant mass distribution again exhibits a peak around 17 MeV (the best-fit value is 16.7 MeV). The results for Helium are shown in Fig. 15.5 and Fig. 15.6.

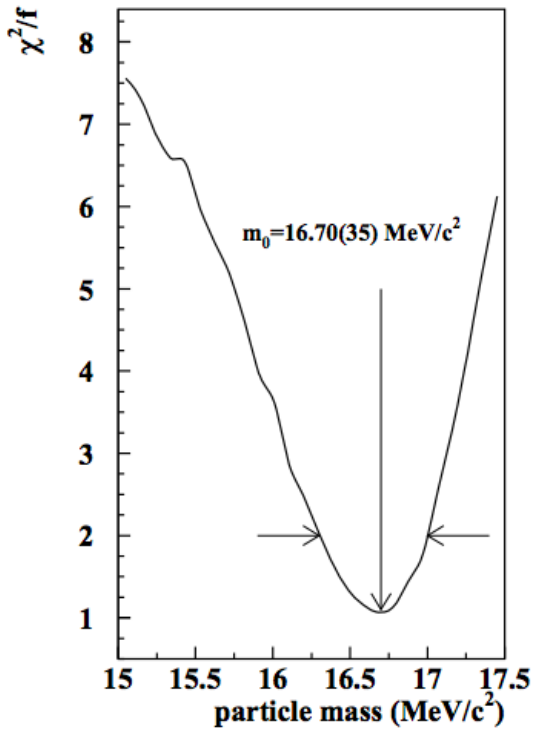


Figure 15.4: Best fit to the invariant mass of electron-positron pairs resulting from Beryllium de-excitation following proton bombardment of Lithium atoms [113].

While the observed excesses in two different systems are indeed quite statistically significant, much is still unclear about the reported results and many in the community remain skeptical about the possibility of a newly discovered elementary particle. One of the main criticisms of the study is that despite probing two different physical systems, the experimental setup was the same (albeit upgraded). It is entirely possible that an unknown systematic effect is the cause of the observed behavior at large angles in both cases. For this claim to gain more solid footing, there needs to be independent verification from an independent group.

Other potentially serious concerns are the less-than-ideal statistical modeling employed in the data analysis and the antiquated Monte Carlo toolkit used in the simulation of expected backgrounds (GEANT3). It is possible that low-energy physics is not being correctly accounted for by the toolkit. In fact, a much more comprehensive and complete

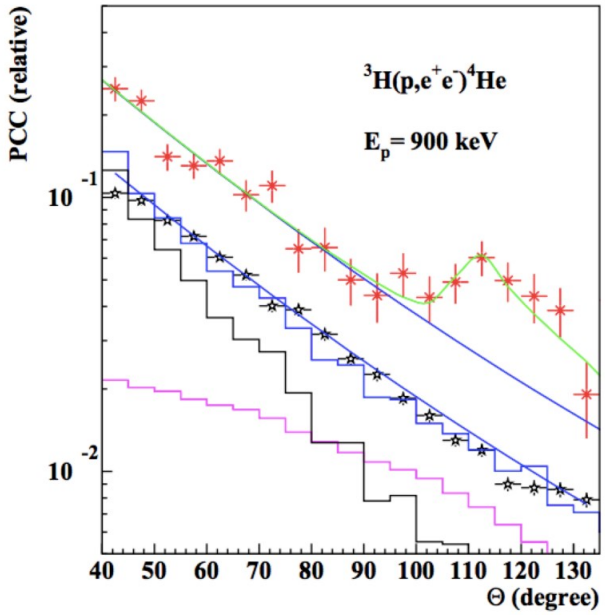


Figure 15.5: Angular correlation of electron-positron pairs in the de-excitation of Helium nuclei following proton bombardment of tritium atoms [114]

package, GEANT4, has been available for more than 20 years, providing a full library of low-energy physics processes [131]. It is not clear why the authors chose to use the much older GEANT3 instead.

Nevertheless, if a 17 MeV particle really exists, then PADME is ideally suited to provide an independent confirmation. The mass sensitivity of PADME is roughly 2 MeV to 24 MeV which makes the X17 boson accessible. If it indeed decays to a positron-electron pair, then PADME should be able to produce it with the inverse reaction when bombarding carbon atoms with a beam of positrons.

PADME has another advantage in probing the X17 excess: since it is one of only a few fixed-target experiments worldwide with a positron beam, knowing the mass hypothesis *a priori* allows us to directly induce resonant production of the particle. By scanning the beam energy in a narrow range near 282.3 MeV, the production cross section would become significantly enhanced. This point is illustrated in Fig. 15.7, which compares

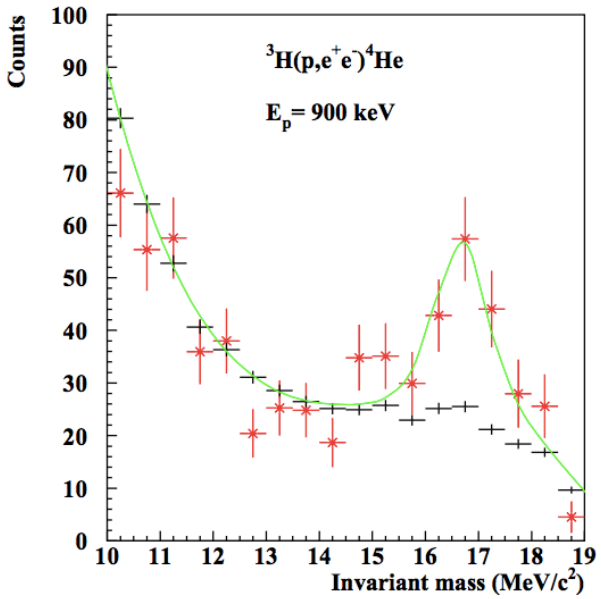


Figure 15.6: Invariant mass of electron-positron pairs in the de-excitation of Helium nuclei following proton bombardment of tritium atoms [114].

the typical PADME operation mode (radiative annihilation) with the proposed mode (resonant annihilation) and also with the typical dark photon production mechanism in electron-beam fixed-target experiments (dark-sstrahlung, or radiative emission of dark photons). Electron beam experiments only have access to an $O(\alpha^3)$ process versus PADME's $O(\alpha^2)$ annihilation. With a narrow beam energy and the ensuing resonant production, we could access an $O(\alpha)$ process and hence probe a significantly larger cross section. This is a marked advantage of positron-type experiments over electron-type experiments.

PADME is currently evaluating the feasibility of performing a dedicated run to probe the X17 excess and hopefully provide an independent verification of its existence. Note, however, that PADME is unable to entirely refute the hypothesis if it fails to see an excess. It is always possible that the coupling to SM particles is smaller than its sensitivity. Nevertheless, a positive signal will be difficult to miss.

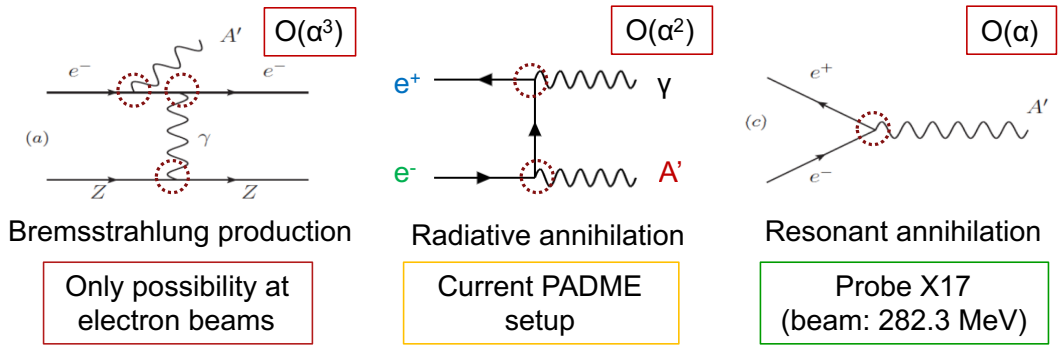


Figure 15.7: Production modes of X17 particle available with electron beams (left) and with positron beams (center and right).

15.3 Axion-like particles

In addition to probing the main invisible channel with a missing-mass technique, PADME is also sensitive to physics with visible final-state signatures. A prominent example is axion-like particles (ALPs), pseudo-scalar bosons that can couple to photons and electrons [142]. These couplings imply that an ALP produced in the electron-positron annihilation could decay to a pair of photons or back to a pair of charged leptons. These processes are depicted in Fig. 15.8.

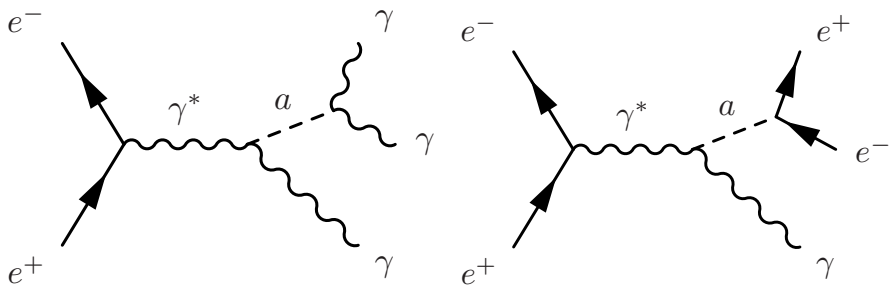


Figure 15.8: Feynman diagrams of ALPs produced in electron-positron annihilation and decaying to a pair of photons (left) and a pair of electron-positron (right).

If the ALP is long-lived and decays outside the detector, the signature would be identical to a dark photon, with one photon in the calorimeter resulting in missing mass. But if the ALP decays promptly, the ECAL could potentially identify the extra photon

pair, and the final-state signature would contain 3 ECAL photons. The feasibility of this search is under investigation. Fig. 15.9 shows a preliminary estimate of the region in ALP mass and coupling phase space that PADME could be sensitive to.

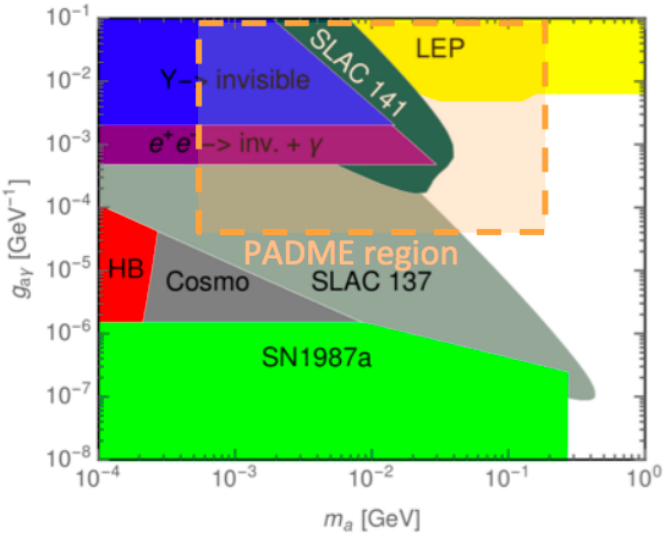


Figure 15.9: Preliminary potential PADME reach in ALP searches.

15.4 Multi-lepton prompt signatures

The PADME vetoes responsible for tagging Bremsstrahlung-emitting positrons can potentially be used for physics searches as well. There are two sets of scintillator bars, one on each wall of the magnet, for positively and negatively charged particles. Therefore, PADME might be sensitive to a range of beyond standard model (BSM) physics with prompt lepton final states. By employing the vetoes, for example, the ALP coupling to leptons can also be explored. A signature containing a single photon in the ECAL plus a positron and an electron in the vetoes correlated in time could be indicative of an ALP decay.

Furthermore, complex models of BSM physics could have multi-lepton final states.

Dark Higgs models in which a dark Higgs is produced from the positron-electron annihilation and subsequently decays to one or more dark photons could leave a cascading chain of leptons in the vetoes. This topology is depicted in Fig. 15.10. There is a limit to how many leptons might be realistically produced with the available beam energy and visible within the Bremsstrahlung background, but preliminary studies indicate that PADME might have at least some sensitivity to such models, making this an intriguing possibility for further searches with the available data.

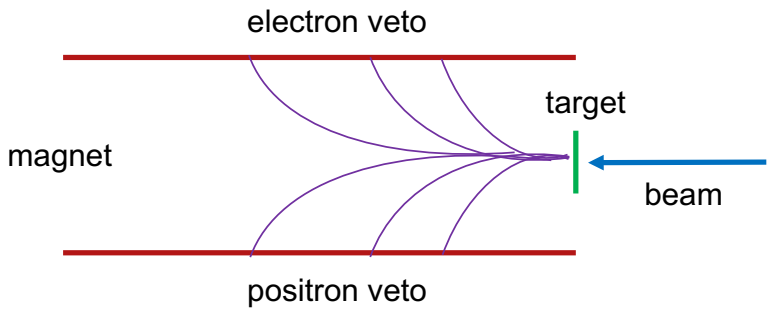


Figure 15.10: Topology of events with multi-lepton prompt signatures in PADME.

15.5 Multi-lepton/photon displaced signatures

Finally, PADME could also be sensitive to displaced signatures where long-lived particles decay to leptons and photons after traveling a macroscopic distance. If a long-lived particle is produced in the target and travels a distance of roughly a meter, its decay products might be visible in the ECAL. This topology is depicted in Fig. 15.11.

One such model is inelastic dark matter, decaying to a collimated lepton or photon pair. This model is discussed extensively in Part II. Initial studies are being performed to understand the feasibility of such long-lived searches with the PADME data. Such analyses could provide complementary coverage to long-lived signatures in a light dark matter phase space, therefore offering compelling motivation for further exploration.

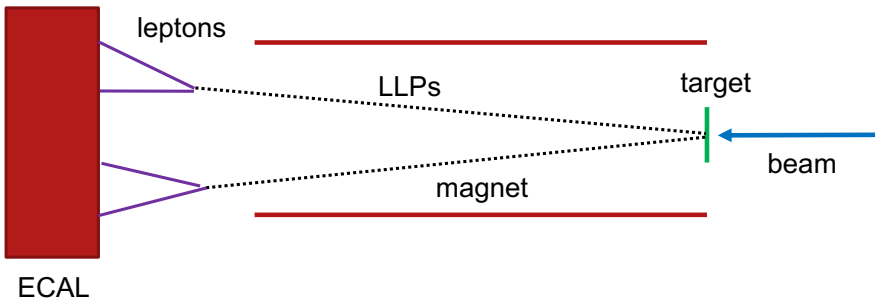


Figure 15.11: Topology of events with multi-lepton/photon displaced signatures in PADME.

15.6 Accelerator upgrades

The current PADME reach is limited by the available beam, both in energy and luminosity. PADME's dark photon mass reach caps around 24 MeV, set by the 550 MeV beam energy, while the luminosity is limited by the instantaneous pileup. Several plans are under discussion for the future of PADME that hinge on a better beam delivered to the detector. In broad terms, there are two possibilities: the PADME detector could be moved to a laboratory with a higher-energy beam, or the acceleration complex of DAFNE could be upgraded.

Two options are considered in moving the PADME detector. The Cornell Electron Storage Ring (CESR) is able to produce and store positron beams with an energy up to 6 GeV, roughly a factor of 10 higher than the DAFNE beam energy. This could extend the mass reach to 80 MeV dark photons. With the resonant extraction technique described below, the sensitivity to the kinetic mixing could also be considerably enhanced. Fig. 15.1 shows projected limits of a potential PADME experiment at Cornell.

Similarly, Thomas Jefferson National Accelerator Facility (JLab) has a 12 GeV positron beam, which could further extend the mass reach to 110 MeV dark photons. There has been notable interest from JLab in bringing the detector on-site once PADME

has finished running at LNF [143]. Fig. 15.12 shows projected sensitivities for an eventual run of PADME at JLab. Two experimental techniques are studied: a thin target (the current approach) and a thick target. The thick-target positron-beam approach could provide a considerable gain in sensitivity by probing dark photon production via resonant annihilation between secondary positrons, emitted by electromagnetic showers developed inside the thick target, and electrons in the target. Research needs to be performed in order to achieve the beam structure required for the thick-target approach at JLab [144].

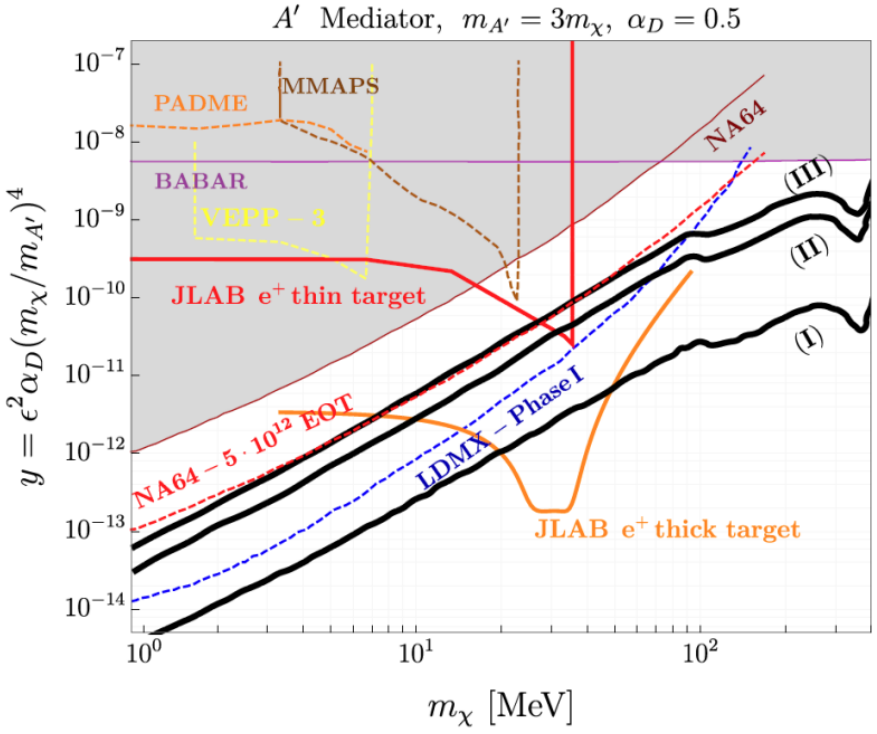


Figure 15.12: Projected sensitivities for PADME at JLab, under thin-target and a thick-target experimental assumptions [144].

15.6.1 Resonant extraction

Strategies to increase the luminosity delivered to the experiment by means of a reduction in instantaneous pileup have also been studied. One proposed idea both for Cornell's

CESR ([145]) and LNF's DAFNE ([117]) is the implementation of a resonant beam extraction scheme. In resonant extraction, the circulating beam at nominal energy is *slowly* extracted over thousands of turns instead of instantaneously over a single turn. This spreads out bunch particles in time, effectively extending the length of a single bunch and increasing the duty cycle of the machine. Therefore, the detector can receive positrons during a longer span and with less instantaneous pileup, which leads to both more positrons collected and less background energy deposition in each event.

The general principle behind resonant extraction is to induce a resonance in beam phase space. Particles are put on unstable orbits but in a controlled manner [117]. With each turn, they are slowly driven towards increasing radii relative to the nominal trajectory. At a certain point, the displacement of the orbit becomes so large that these particles enter the field region of an electrostatic or magnetic kicker, whereby they are extracted from the beam line. The phase space resonance can be induced by different types of magnets; two of them are depicted in Fig. 15.13. The most common is a sextupole-induced resonance with a third-integer tune (Fig. 15.13 on the right).

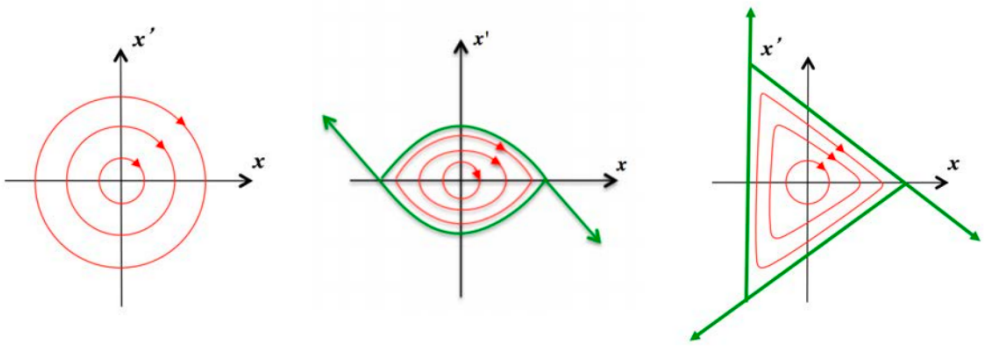


Figure 15.13: Examples of phase-space resonances [117]. Left: linear phase space (no resonance). Center: half-integer resonance (octupole-induced). Right: third-integer resonance (sextupole-induced).

After the resonant phase space has been induced, it can be slowly shrunk by approaching the exact tune of the resonance. This is because at the nominal fractional tune

of the resonance, no stable orbit exists and all particles have unstable orbits. Thus the beam can be slowly extracted by adiabatically approaching the exact fractional tune in a controlled way. This is shown in Fig. 15.14 for the case of third-integer resonance. The figure highlights the large number of turns needed for full extraction of the beam.

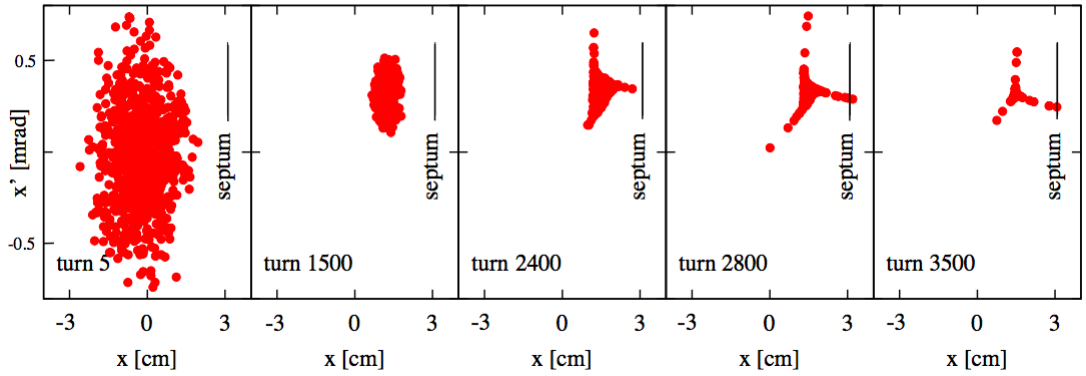


Figure 15.14: Simulation of resonant extraction in CESR with a third-integer resonance over thousands of turns [145].

15.6.2 PADME integrated luminosity

To demonstrate the potential gains in luminosity from a resonant extraction scheme, we compute the PADME instantaneous and integrated luminosity in units of pb^{-1}, which is common in collider studies, instead of POTs, the typical quantity quoted in fixed-target experiments.

The instantaneous luminosity in a fixed-target experiment can be expressed as [146]:

$$\mathcal{L}_{\text{inst}} = \left(\frac{I_b}{e}\right) N_A \left(\frac{\rho\, x}{\mathcal{A}}\right), \tag{15.1}$$

where I_b/e is the beam current in units of the electron charge, $N_A = 6.02 \times 10^{23}$ is Avogadro's number, ρ is the density in g/cm^3, x is the target thickness, and $\mathcal{A}$ is the

gram-molecular weight of the target material, in grams.

For PADME, the relevant quantities are $\rho = 6 \times 3.51\,\mathrm{g/cm^3}$ (density of diamond); $\mathcal{A} = 12.01\,\mathrm{g}$ (carbon's gram-molecular weight); $x = 100\,\mu\mathrm{m} = 10^{-2}\,\mathrm{cm}$; and $I_b/e = 23{,}000 \times 50 = 1.15 \times 10^6$ positrons/second in the first run of PADME, but potentially up to $600{,}000 \times 50 = 3 \times 10^7$ positrons/second in the second run. The first quantity is the number of positrons per bunch, multiplied by the bunch repetition rate. Note that we multiply the density ρ by 6 to account for 6 electrons per carbon atom. The second run could have higher charge per bunch due to ongoing upgrades to DAFNE.

The average instantaneous luminosity for the first and second runs of PADME with these numbers is:

$$\begin{aligned} L_1 &= \left(1.2 \times 10^6\right)\left(6 \times 10^{23}\right)\left(\frac{6 \times 3.5 \times 10^{-2}}{12.01}\right) = 1.2 \times 10^{28}\,\mathrm{cm^{-2}\,s^{-1}} \\ L_2 &= \left(3.0 \times 10^7\right)\left(6 \times 10^{23}\right)\left(\frac{6 \times 3.5 \times 10^{-2}}{12.01}\right) = 3.2 \times 10^{29}\,\mathrm{cm^{-2}\,s^{-1}}. \end{aligned} \tag{15.2}$$

They can also be expressed in units of $\mu\mathrm{b^{-1}\,s^{-1}}$:

$$\begin{aligned} L_1 &= 1.2 \times 10^{28}\,\mathrm{cm^{-2}\,s^{-1}}/10^{30} = 0.012\,\mu\mathrm{b^{-1}\,s^{-1}} \\ L_2 &= 3.2 \times 10^{29}\,\mathrm{cm^{-2}\,s^{-1}}/10^{30} = 0.320\,\mu\mathrm{b^{-1}\,s^{-1}}. \end{aligned} \tag{15.3}$$

Finally, integrating both instantaneous luminosity numbers over one year of data-taking and converting to $\mathrm{pb^{-1}}$ gives the integrated luminosity:

$$\begin{aligned} L_{\mathrm{int},1} &= 0.012 \times 3.2 \times 10^7/10^6 = 0.384\,\mathrm{pb^{-1}} \\ L_{\mathrm{int},2} &= 0.320 \times 3.2 \times 10^7/10^6 = 10.1\,\mathrm{pb^{-1}}. \end{aligned} \tag{15.4}$$

The benefits of resonant extraction become evident when we compute the projected integrated luminosity assuming that the instantaneous pileup is mitigated with this technique. For a back-of-the-envelope estimate, consider the duty cycle currently available with the DAFNE accelerator. With a bunch length of roughly 200 ns, and 50 bunches per second, the duty cycle is only about 10^{-5}. This means that, assuming the *average* instantaneous luminosity to be $1.2 \times 10^{28}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ as computed in Eq. (15.2), the *peak* instantaneous luminosity delivered by DAFNE in a single bunch can reach values of roughly $1.2 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. If the Cornell and JLab accelerators can provide similar instantaneous luminosity benchmarks (a plausible assumption), just increasing the duty cycle with a resonant extraction strategy could push the overall integrated luminosity up by a similar ratio. If resonant extraction can provide a quasi-continuous beam (e.g. a duty cycle of 10% or more), then the integrated luminosity could reach roughly $0.384 \times 10^4 = 3{,}840\,\mathrm{pb}^{-1}$ over one year of data-taking. This is a drastic increase in data that makes the proposal quite compelling. More concrete plans are currently being developed for such upgrades after the end of data-taking at LNF.

Part IV

Conclusions

CHAPTER 16

THE FUTURE OF DARK MATTER

Dark matter is arguably one of the most challenging open problems in physics and its nature remains a mystery almost a hundred years after its initial discovery. Despite sustained and concentrated efforts from the theoretical and experimental physics communities, we have been unable to uncover the truth about this fascinating phenomenon.

Nevertheless, recent experimental findings do give us hints about where to keep looking. The DAMA/LIBRA seasonal modulation, discussed in Chapter 4, is a longstanding result that has yet to be satisfactorily explained. Skepticism abounds, but no systematic effects proposed to date are able to conclusively explain away the observed modulation (though one 2019 study introduces a promising explanation that suggests Helium contamination in the PMTs may be the culprit). Several experiments are currently running or coming online in the next few years to hopefully replicate—or disprove—the modulation, such as COSINE-100, ANAIS-112, SABRE, and COSINUS.

Furthermore, recently, in June of 2020, a new result was released by the XENON Collaboration [1]. This direct-detection experiment claims to observe an excess of electro-recoil events at low recoil energy of less than about 5 keV. Fig. 16.1 shows a plot of the number of events recorded as a function of electronic recoil energy. A significant excess above 2σ is reached around 2-3 keV. The Collaboration suggests this could be due to unexpected tritium contamination, but it also seems compatible with a dark matter explanation in the form of solar axions. Notably, the energy range of the observed excess is identical to the one observed by DAMA/LIBRA [2]. More observation is needed, but if confirmed this could be an interesting pointer to where dark matter is hiding.

In the indirect detection world, astrophysical evidence points to unexplained phe-

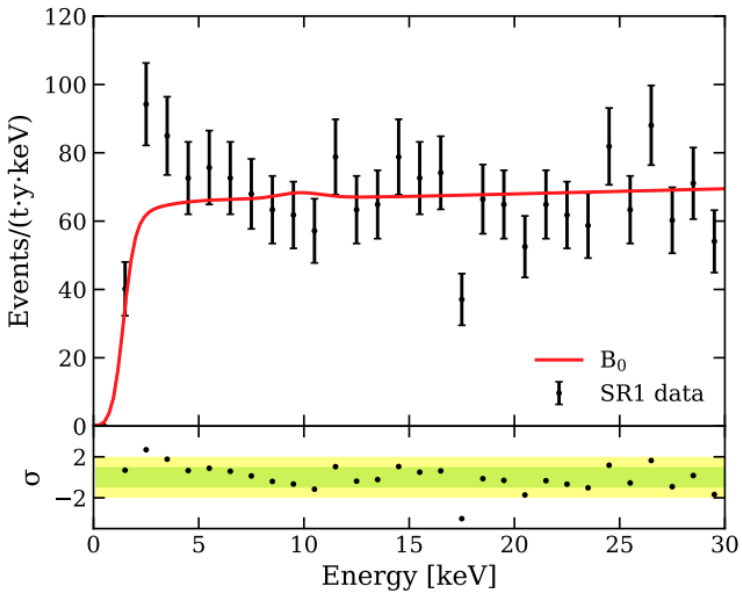


Figure 16.1: Recent observed excess in electronic recoil events by the XENON Collaboration [1]. The excess seems to peak around recoil energies of 2-3 keV and is compatible with the energy range of the DAMA/LIBRA excess [2].

nomena that could be related to dark matter. The 3.5 keV emission line originating from some galaxies, the Milky Way's galactic center GeV excess, and the anomalous positron fraction observed by PAMELA and AMS-02 in cosmic rays, discussed in Chapter 4, are interesting examples. More observation is needed to understand the sources of such deviations and correlate them to possible new physics scenarios.

Atomki Lab's Beryllium and Helium excesses described in Chapter 15 have also caused immense interest lately. Multiple experiments currently taking data or under development will be able to further study these anomalies. In particular, PADME can probe the proposed "X17" boson with mass near 17 MeV that could explain the observed excesses. PADME is already running and should be able to directly investigate the anomaly within the next few years. Providing independent confirmation is critical to establish any claims of dark matter discovery.

Sensitivity to models of light dark matter (in the MeV range) is only one of several

advantages offered by fixed-target and beam dump experiments. To fully investigate the dark matter problem, complementarity between experimental approaches is indispensable. Small-scale experiments offer an exciting search program at a fraction of the cost of collider setups. This point was recently highlighted by the 2020 European Strategy for Particle Physics [147]:

> *Given the challenges faced by CERN in preparing for the future collider, the role of the National Laboratories in advancing the exploration of the lower energy regime cannot be over-emphasised. In addition to the examples already mentioned above, a broad programme of axion searches is proposed at DESY, a search for low-mass dark matter particles with a positron beam is under way at Frascati, and the COSY facility could be used as a demonstrator for measuring the electric dipole moment of the proton at Jülich. These initiatives should be strongly encouraged and supported. Europe has the opportunity to play a leading role in this diverse scientific programme by supporting highimpact projects, which mostly require modest investment and play a crucial role in training and preparing a new generation of versatile scientists to address the challenges of the future.*

This message will almost certainly be emphasized as well in the Snowmass 2021 process currently underway in the United States. In addition to the adjacent phase space coverage, small-scale experiments provide an excellent opportunity for students to learn hands-on and acquire the hardware and software skills needed to succeed in high-energy physics research. As such, it is an invaluable tool in the formation of future physicists who will be designing and building the next generation of dark matter experiments.

In the near term, several dark matter experiments are under design or construction. The Light Dark Matter Experiment (LDMX) [148] aims to uncover evidence of dark

matter, having sensitivity to a wide range of dark matter and mediator models in the sub-GeV range. It will employ a missing-momentum technique with a multi-GeV electron beam at SLAC. This approach is complementary to PADME's positron beam and probes Bremsstrahlung-like emission of dark sector particles when beam electrons strike a target. It benefits from a larger beam luminosity and smaller background contamination because of the experimental setup, resulting in large sensitivity gains. LDMX has been funded and will begin construction soon.

Within the collider environment, the FASER experiment [149] has been approved by CERN and is being built at the LHC with the goal of starting taking data already in 2021. FASER will look for new physics with extremely displaced signatures. The detector sits 480 m away from the ATLAS interaction point and will attempt to detect long-lived particles produced in proton-proton collisions that only feebly interact with other standard model particles (such as dark matter). Other experiments similarly targeting long-lived phenomenology (but not yet approved) include the milliQan experiment [150], a search for milli-charged particles; MATHUSLA [151], a large detector to search for displaced particles originating from the CMS interaction point; and SHiP [152], a general-purpose beam-dump experiment at CERN's SPS to search for a wide array of exotic new physics.

In the longer term, new collider options are being considered. A future electron-positron collider would allow a more precise investigation of Higgs and electroweak properties and could further serve as an intermediate stage to a 100 TeV hadron collider [147]. This is the preferred option as outlined in the 2020 European Strategy for Particle Physics. New physics could manifest itself at these higher energies. Alternatives include a possible multi-TeV muon collider [153] and linear colliders such as the International Linear Collider [154] and the Compact Linear Collider [155].

Whichever path is taken toward the future of high-energy physics, it is imperative to

employ innovative experimental approaches and explore novel theoretical ideas in the search for dark matter. This dissertation showcased two such examples: the first search for inelastic dark matter at a hadron collider with CMS, and a search for dark photons using a positron beam with PADME. The complementarity of experimental techniques discussed here is a small-scale version of the much larger complementarity that will be required to rummage through the myriad places where dark matter could be hiding.

One hundred years after its discovery, we still do not know what dark matter is. A comprehensive search program is paramount to ensure that—a hundred years from now—we will have fully understood dark matter.

BIBLIOGRAPHY

[1] XENON Collaboration, "Observation of Excess Electronic Recoil Events in XENON1T", arXiv:2006.09721.

[2] R. Bernabei et al., "First model independent results from DAMA/LIBRA-phase2", *Nucl. Phys. At. Energy* **19** (2018) 307, doi:10.15407/jnpae2018.04.307, arXiv:1805.10486.

[3] K. G. Begeman, "H I rotation curves of spiral galaxies. I - NGC 3198", *Astron. Astrophys.* **223** (1989) 47.

[4] V. C. Rubin, W. K. Ford, Jr., and N. Thonnard, "Extended rotation curves of high-luminosity spiral galaxies. IV - systematic dynamical properties, SA through SC", *Astrophys. J. Lett.* **225** (1978) L107, doi:10.1086/182804.

[5] J. F. Navarro, C. S. Frenk, and S. D. M. White, "A universal density profile from hierarchical clustering", *ApJ* **490** (1997) 493, doi:10.1086/304888, arXiv:astro-ph/9611107.

[6] S.-H. Oh et al., "Dark and luminous matter in THINGS dwarf galaxies", *ApJ* **141** (2011) 193, doi:10.1088/0004-6256/141/6/193, arXiv:1011.0899.

[7] A. Einstein, "über den Einfluß der Schwerkraft auf die Ausbreitung des Lichtes", *Ann. Phys.* **35** (1911) 898.

[8] G. Squires et al., "The dark matter, gas, and galaxy distributions in Abell 2218: A weak gravitational lensing and X-ray analysis", *ApJ* **461** (1996) 572, doi:10.1086/177085, arXiv:astro-ph/9507008.

[9] D. Clowe et al., "A direct empirical proof of the existence of dark matter", *ApJ* **648** (2006) L109, doi:10.1086/508162, arXiv:astro-ph/0608407.

[10] D. J. Fixsen, "The temperature of the cosmic microwave background", *ApJ* **707** (2009) 916, doi:10.1088/0004-637X/707/2/916, arXiv:0911.1955.

[11] P. J. E. Peebles and B. Ratra, "The cosmological constant and dark energy", *Rev. Mod. Phys.* **75** (2003) 559, doi:10.1103/RevModPhys.75.559.

[12] Planck Collaboration, "Planck 2018 results. VI. Cosmological parameters", arXiv:1807.06209.

[13] A. Fraknoi, D. Morrison, and S. C. Wolff, "The Big Bang", in *Astronomy*. OpenStax, Houston, Texas, 2016.

[14] Planck Collaboration, "Planck 2013 results. XV. CMB power spectra and likelihood", *A&A* **571** (2014) A15, doi:10.1051/0004-6361/201321573, arXiv:1303.5075.

[15] Y. Grossman and Y. Nir, "The standard model", (2020). [Draft Manuscript].

[16] P. Langacker, "The Standard Model and Beyond". Series in High Energy Physics, Cosmology and Gravitation. Taylor & Francis, 2010. doi:10.1201/b22175, ISBN 978-1-4200-7906-7.

[17] D. Curtin et al., "Exotic decays of the 125 GeV Higgs boson", *Phys. Rev. D* **90** (2014) 075004, doi:10.1103/PhysRevD.90.075004, arXiv:1312.4992.

[18] Particle Data Group et al., "Review of Particle Physics", *Phys. Rev. D* **98** (2018) 030001, doi:10.1103/PhysRevD.98.030001.

[19] J. Ellis, "Higgs physics", in *Proceedings, 2013 European School of High-Energy Physics (ESHEP 2013): Paradfurdo, Hungary, June 5-18, 2013*, p. 117. 2015. arXiv:1312.5672. doi:10.5170/CERN-2015-004.117.

[20] C. Csáki, S. Lombardo, and O. Telem, "TASI lectures on non-supersymmetric BSM models", in *Anticipating the Next Discoveries in Particle Physics*, p. 501. World Scientific, Boulder, Colorado, 2018. arXiv:1811.04279. doi:10.1142/9789813233348$_0$007.

[21] S. M. Bilenky, C. Giunti, and W. Grimus, "Phenomenology of neutrino oscillations", *Prog. Part. Nucl. Phys.* **43** (1999) 1, doi:10.1016/S0146-6410(99)00092-7, arXiv:hep-ph/9812360.

[22] M. Dine and A. Kusenko, "The origin of the matter-antimatter asymmetry", *Rev. Mod. Phys.* **76** (2003) 1, doi:10.1103/RevModPhys.76.1, arXiv:hep-ph/0303065.

[23] K. S. Babu and X.-G. He, "Fermion mass hierarchy and the strong CP problem", *Phys. Lett. B* **219** (1989) 342, doi:10.1016/0370-2693(89)90401-2.

[24] D. Huterer and D. L. Shafer, "Dark energy two decades after: Observables, probes, consistency tests", *Rep. Prog. Phys.* **81** (2018) 016901, doi:10.1088/1361-6633/aa997e, arXiv:1709.01091.

[25] O. E. Gerhard and D. N. Spergel, "Dwarf spheroidal galaxies and the mass of the neutrino", *Astrophys. J. Lett.* **389** (1992) L9, doi:10.1086/186336.

[26] G. Arcadi et al., "The waning of the WIMP? A review of models, searches, and constraints", *Eur. Phys. J. C* **78** (2018) 203, doi:10.1140/epjc/s10052-018-5662-y, arXiv:1703.07364.

[27] B. W. Lee and S. Weinberg, "Cosmological lower bound on heavy-neutrino masses", *Phys. Rev. Lett.* **39** (1977) 165, doi:10.1103/PhysRevLett.39.165.

[28] M. Lisanti, "Lectures on dark matter physics", *New Front. Fields Strings* (2017) 399, doi:10.1142/9789813149441$_0$007, arXiv:1603.03797.

[29] T. Tait, "Dark matter: Theoretical overview", (2016). [Conference Presentation].

[30] J. Alexander et al., "Dark sectors 2016 workshop: Community report", arXiv:1608.08632.

[31] B. Holdom, "Two U(1)'s and ϵ charge shifts", *Phys. Lett. B* **166** (1986) 196, doi:10.1016/0370-2693(86)91377-8.

[32] A. Berlin and F. Kling, "Inelastic dark matter at the LHC lifetime frontier: ATLAS, CMS, LHCb, CODEX-b, FASER, and MATHUSLA", *Phys. Rev. D* **99** (2019) 015021, doi:10.1103/PhysRevD.99.015021, arXiv:1810.01879.

[33] T. R. Slatyer, "TASI lectures on indirect detection of dark matter", arXiv:1710.05137. [MIT-CTP/4946].

[34] E. Bulbul et al., "Detection of an unidentified emission line in the stacked X-ray spectrum of galaxy clusters", *ApJ* **789** (2014) 13, doi:10.1088/0004-637X/789/1/13, arXiv:1402.2301.

[35] A. Boyarsky, O. Ruchayskiy, D. Iakubovskyi, and J. Franse, "An unidentified line in X-ray spectra of the Andromeda galaxy and Perseus galaxy cluster", *Phys. Rev. Lett.* **113** (2014) 251301, doi:10.1103/PhysRevLett.113.251301, arXiv:1402.4119.

[36] L. Goodenough and D. Hooper, "Possible evidence for dark matter annihilation In the inner Milky Way from the Fermi Gamma Ray Space Telescope", arXiv:0910.2998. [FERMILAB-PUB-09-494-A].

[37] D. Hooper and L. Goodenough, "Dark Matter annihilation in the Galactic Center as seen by the Fermi Gamma Ray Space Telescope", *Phys. Lett. B* **697** (2011) 412, doi:10.1016/j.physletb.2011.02.029, arXiv:1010.2752.

[38] J. Petrovic, P. D. Serpico, and G. Zaharijas, "Galactic Center gamma-ray "excess" from an active past of the Galactic Centre?", *J. Cosmol. Astropart. Phys.* **2014** (2014) 052, doi:10.1088/1475-7516/2014/10/052, arXiv:1405.7928.

[39] O. Adriani et al., "Observation of an anomalous positron abundance in the cosmic radiation", *Nature* **458** (2009) 607, doi:10.1038/nature07942, arXiv:0810.4995.

[40] I. Cholis, D. P. Finkbeiner, L. Goodenough, and N. Weiner, "The PAMELA positron excess from annihilations into a light boson", *J. Cosmol. Astropart. Phys.* **2009** (2009) 007, doi:10.1088/1475-7516/2009/12/007, arXiv:0810.5344.

[41] AMS Collaboration, "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", *Phys. Rev. Lett.* **110** (2013) 141102, doi:10.1103/PhysRevLett.110.141102.

[42] M. Di Mauro, S. Manconi, and F. Donato, "Detection of a γ-ray halo around Geminga with the Fermi-LAT data and implications for the positron flux", *Phys. Rev. D* **100** (2019) 123015, doi:10.1103/PhysRevD.100.123015, arXiv:1903.05647.

[43] B. J. Mount et al., "LUX-ZEPLIN (LZ) technical design report", Technical Report LBNL-1007256, 2017. arXiv:1703.09144.

[44] SuperCDMS Collaboration, "Projected sensitivity of the SuperCDMS SNOLAB experiment", *Phys. Rev. D* **95** (2017) 082002, doi:10.1103/PhysRevD.95.082002, arXiv:1610.00006.

[45] M. Schumann, "Direct detection of WIMP dark matter: Concepts and status", *J. Phys. G: Nucl. Part. Phys.* **46** (2019) 103003, doi:10.1088/1361-6471/ab2ea5, arXiv:1903.03026.

[46] P. Cushman et al., "Snowmass CF1 summary: WIMP dark matter direct detection", arXiv:1310.8327.

[47] R. Bernabei et al., "Final model independent result of DAMA/LIBRA-phase1",

Eur. Phys. J. C **73** (2013) 2648, doi:10.1140/epjc/s10052-013-2648-7, arXiv:1308.5109.

[48] R. Bernabei et al., "The DAMA/LIBRA apparatus", *Nucl. Instrum. Meth. A* **592** (2008) 297, doi:10.1016/j.nima.2008.04.082, arXiv:0804.2738.

[49] G. Adhikari et al., "An experiment to search for dark-matter interactions using sodium iodide detectors", *Nature* **564** (2018) 83, doi:10.1038/s41586-018-0739-1, arXiv:1906.01791.

[50] J. Amaré et al., "First results on dark matter annual modulation from ANAIS-112 experiment", *Phys. Rev. Lett.* **123** (2019) 031301, doi:10.1103/PhysRevLett.123.031301, arXiv:1903.03973.

[51] M. Antonello et al., "The SABRE project and the SABRE PoP", *Eur. Phys. J. C* **79** (2019) 363, doi:10.1140/epjc/s10052-019-6860-y, arXiv:1806.09340.

[52] G. Angloher et al., "The COSINUS project - perspectives of a NaI scintillating calorimeter for dark matter search", *Eur. Phys. J. C* **76** (2016) 441, doi:10.1140/epjc/s10052-016-4278-3, arXiv:1603.02214.

[53] D. Buttazzo, P. Panci, N. Rossi, and A. Strumia, "Annual modulations from secular variations: Relaxing DAMA?", *JHEP* **2020** (2020) 137, doi:10.1007/JHEP04(2020)137, arXiv:2002.00459.

[54] D. Ferenc, D. F. Šegedin, I. F. Šegedin, and M. Š. Ferenc, "Helium migration through photomultiplier tubes – the probable cause of the DAMA seasonal variation effect", arXiv:1901.02139.

[55] J. Jaeckel, M. Lamont, and C. Vallée, "The quest for new physics with the Physics Beyond Colliders programme", *Nat. Phys.* **16** (2020) 393, doi:10.1038/s41567-020-0838-4.

[56] CMS Collaboration, "CMS Exotica Summary Plots for 13 TeV Data". https://twiki.cern.ch/twiki/bin/view/CMSPublic/SummaryPlotsEXO13TeV, 2020.

[57] ATLAS Collaboration, "Summary Plots from the ATLAS Exotic Physics Group". https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/EXOTICS/, 2020.

[58] CMS Collaboration, "Search for disappearing tracks in proton-proton collisions at $\sqrt{s}$ = 13 TeV", *Phys. Lett. B* **806** (2020) 135502, doi:10.1016/j.physletb.2020.135502, arXiv:2004.05153.

[59] CMS Collaboration, "Search for new particles decaying to a jet and an emerging jet", *JHEP* **2019** (2019) 179, doi:10.1007/JHEP02(2019)179, arXiv:1810.10069.

[60] T. Cohen, M. Lisanti, and H. K. Lou, "Semi-visible jets: Dark matter undercover at the LHC", *Phys. Rev. Lett.* **115** (2015) 171804, doi:10.1103/PhysRevLett.115.171804, arXiv:1503.00009.

[61] CMS Collaboration, "Search for decays of stopped exotic long-lived particles produced in proton-proton collisions at $\sqrt{s}$ = 13 TeV", *JHEP* **2018** (2018) 127, doi:10.1007/JHEP05(2018)127, arXiv:1801.00359.

[62] CMS Collaboration, "Search for long-lived particles decaying into displaced jets in proton-proton collisions at $\sqrt{s}$ = 13 TeV", *Phys. Rev. D* **99** (2019) 032011, doi:10.1103/PhysRevD.99.032011, arXiv:1811.07991.

[63] CMS Collaboration, "Search for long-lived particles using delayed photons in proton-proton collisions at $\sqrt{s}$ = 13 TeV", *Phys. Rev. D* **100** (2019) 112003, doi:10.1103/PhysRevD.100.112003, arXiv:1909.06166.

[64] S. P. Martin, "A supersymmetry primer", doi:10.1142/9789812839657$_0$001, arXiv:hep-ph/9709356.

[65] J. Alimena et al., "Searching for long-lived particles beyond the Standard Model at the Large Hadron Collider", arXiv:1903.04497.

[66] E. Izaguirre, G. Krnjaic, and B. Shuve, "Discovering inelastic thermal-relic dark matter at colliders", *Phys. Rev. D* **93** (2016) 063523, doi:10.1103/PhysRevD.93.063523, arXiv:1508.03050.

[67] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **2014** (2014) 79, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[68] I. Hoenig, G. Samach, and D. Tucker-Smith, "Searching for dilepton resonances below the Z mass at the LHC", *Phys. Rev. D* **90** (2014) 075016, doi:10.1103/PhysRevD.90.075016, arXiv:1408.1075.

[69] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[70] ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider", *JINST* **3** (2008) S08003, doi:10.1088/1748-0221/3/08/S08003.

[71] ALICE Collaboration, "The ALICE experiment at the CERN LHC", *JINST* **3** (2008) S08002, doi:10.1088/1748-0221/3/08/S08002.

[72] LHCb Collaboration, "The LHCb Detector at the LHC", *JINST* **3** (2008) S08005, doi:10.1088/1748-0221/3/08/S08005.

[73] TOTEM Collaboration, "First measurement of elastic, inelastic and total cross-section at $\sqrt{s}$ = 13 TeV by TOTEM and overview of cross-section data at LHC energies", arXiv:1712.06153. [CERN-EP-2017-321-v2].

[74] H. Wiedemann, "Introduction to Accelerator Physics", in *Particle Accelerator Physics*, Graduate Texts in Physics. Springer Cham, fourth edition, 2015. doi:10.1007/978-3-319-18317-6.

[75] M. Zientek, "Search for Dark Matter Produced in Association with a Higgs Boson Decaying to Two Photons in Proton-Proton Collisions at 13 TeV with the CMS Detector". PhD thesis, 2019. doi:10.7298/bjhw-8g62.

[76] J.-L. Caron, "Cross section of LHC dipole.. Dipole LHC: Coupe transversale.", (1998). AC Collection. Legacy of AC. Pictures from 1992 to 2002.

[77] J.-P. Burnet, "Putting it into practice", doi:10.5170/CERN-2015-003.445, arXiv:1607.01596. 13 pages, contribution to the 2014 CAS - CERN Accelerator School: Power Converters, Baden, Switzerland, 7-14 May 2014.

[78] CMS Collaboration, "Precise mapping of the magnetic field in the CMS barrel yoke using cosmic rays", *JINST* **5** (2010) T03021, doi:10.1088/1748-0221/5/03/T03021.

[79] CMS Collaboration, "The CMS tracker: Addendum to the technical design report", Technical Report CERN-LHCC-2000-016, CERN, Geneva, 2000.

[80] T. Sakuma and T. McCauley, "Detector and event visualization with SketchUp at the CMS experiment", *J. Phys.: Conf. Ser.* **513** (2014) 022032, doi:10.1088/1742-6596/513/2/022032.

[81] F. Hartmann, "CMS: Increasing size by 2 orders of magnitude", in *Evolution of Silicon Sensor Technology in Particle Physics*, Springer Tracts in Modern Physics, p. 219. Springer, 2017. doi:10.1007/978-3-319-64436-3$_6$.

[82] CMS Collaboration, "The CMS electromagnetic calorimeter project: Technical design report", Technical Report CERN-LHCC-97-033, CERN, Geneva, 1997.

[83] CMS Collaboration, "The CMS hadron calorimeter project: Technical design report", Technical Report CERN-LHCC-97-031, CERN, Geneva, 1997.

[84] CMS Collaboration, "The CMS muon project: Technical design report", Technical Report CERN-LHCC-97-032, CERN, Geneva, 1997.

[85] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s}$ = 7 TeV", *JINST* **7** (2012) P10002, doi:10.1088/1748-0221/7/10/P10002.

[86] S. K. Park et al., "CMS endcap RPC gas gap production for upgrade", *JINST* **7** (2012) P11013, doi:10.1088/1748-0221/7/11/P11013.

[87] V. Khachatryan et al., "The CMS trigger system", *JINST* **12** (2017) P01020, doi:10.1088/1748-0221/12/01/P01020.

[88] J. D. Wells, "How to find a hidden world at the Large Hadron Collider", arXiv:0803.1243. [MCTP-07-51, CERN-PH-TH-2008-47].

[89] D. Curtin, R. Essig, S. Gori, and J. Shelton, "Illuminating dark photons with high-energy colliders", *JHEP* **2015** (2015) 157, doi:10.1007/JHEP02(2015)157, arXiv:1412.0018.

[90] T. Sjöstrand et al., "An Introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.

[91] J. Allison et al., "Recent developments in Geant4", *Nucl. Instrum. Meth. A* **835** (2016) 186, doi:10.1016/j.nima.2016.06.125.

[92] A. M. Sirunyan et al., "Particle-flow reconstruction and global event description with the CMS detector", *JINST* **12** (2017) P10003, doi:10.1088/1748-0221/12/10/P10003.

[93] S. R. Davis, "Interactive Slice of the CMS detector", 2016. [CMS-OUTREACH-2016-027].

[94] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST* **9** (2014) P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.

[95] R. Frühwirth, "Application of Kalman filtering to track and vertex fitting", *Nucl. Instrum. Meth. A* **262** (1987) 444, doi:10.1016/0168-9002(87)90887-4.

[96] A. Strandlie and W. Wittek, "Propagation of Covariance Matrices of Track Parameters in Homogeneous Magnetic Fields in CMS", Technical Report CMS-NOTE-2006-001, CERN, Geneva, 2006.

[97] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s}$ = 13 TeV", *JINST* **13** (2018) P06015, doi:10.1088/1748-0221/13/06/P06015, arXiv:1804.04528.

[98] K. Rose, "Deterministic annealing for clustering, compression, classification, regression, and related optimization problems", *Proc. IEEE* **86** (1998) 2210, doi:10.1109/5.726788.

[99] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_{T} jet clustering algorithm", *JHEP* **2008** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[100] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet user manual", *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[101] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *JINST* **12** (2017) P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.

[102] CMS Collaboration, "Jet algorithms performance in 13 TeV data", Technical Report CMS-PAS-JME-16-003, CERN, Geneva, 2017.

[103] CMS Collaboration, "Search for dark matter produced with an energetic jet or a hadronically decaying W or Z boson at $\sqrt{s}$ = 13 TeV", *JHEP* **2017** (2017) 14, doi:10.1007/JHEP07(2017)014, arXiv:1703.01651.

[104] CMS Collaboration, "Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector", Technical Report CMS-PAS-JME-16-004, CERN, Geneva, 2016.

[105] CMS Collaboration, "Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s}$ = 13 TeV using the CMS detector", *JINST* **14** (2019) P07004, doi:10.1088/1748-0221/14/07/P07004, arXiv:1903.06078.

[106] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018) P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.

[107] CMS Collaboration, "Search for dark matter, extra dimensions, and unparticles in monojet events in proton-proton collisions at $\sqrt{s}$ = 8 TeV", *Eur. Phys. J. C* **75** (2015) 235, doi:10.1140/epjc/s10052-015-3451-4, arXiv:1408.3583.

[108] J. Duarte, "CMS and LHC report", 2018. [Presented at the All Experimenters Meeting, Fermilab, 9 July 2018].

[109] W. Erdmann, "Vertex reconstruction at the CMS experiment", *J. Phys.: Conf. Ser.* **110** (2008) 092009, doi:10.1088/1742-6596/110/9/092009.

[110] W. Buttinger, "Background estimation with the ABCD method", 2018. [Draft Manuscript].

[111] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727.

[112] G. Cowan, "Statistical methods for particle physics – Lecture 3: Asymptotics I; Asimov data set", 2015. [Presented at the Statistical Inference for Astro and Particle Physics Workshop, at Weizmann Institute, Rehovot, March 8-12, 2015].

[113] A. J. Krasznahorkay et al., "Observation of anomalous internal pair creation in ^{8}Be: A possible signature of a light, neutral boson", *Phys. Rev. Lett.* **116** (2016) 042501, doi:10.1103/PhysRevLett.116.042501, arXiv:1504.01527.

[114] A. J. Krasznahorkay et al., "New evidence supporting the existence of the hypothetic X17 particle", arXiv:1910.10459.

[115] G. Amelino-Camelia et al., "Physics with the KLOE-2 experiment at the upgraded DAΦNE", *Eur. Phys. J. C* **68** (2010) 619, doi:10.1140/epjc/s10052-010-1351-1.

[116] P. Valente, "The PADME experiment at DAFNE linac", *PoS* **ICHEP2016** (2017) 204, doi:10.22323/1.282.0204.

[117] P. Valente, "POSEYDON - Converting the DAFNE collider into a double positron facility: A high duty-cycle pulse stretcher and a storage ring", Technical Report INFN-17-15/LNF, 2017. arXiv:1711.06877.

[118] F. Oliva, "Operation and performance of the active target of PADME", *Nucl. Instrum. Meth. A* **958** (2020) 162354, doi:10.1016/j.nima.2019.162354.

[119] F. Ferrarotto, "The Investigation on the dark sector at the PADME Experiment", *Universe* **5** (2019) 59, doi:10.3390/universe5020059.

[120] M. Raggi et al., "Performance of the PADME calorimeter prototype at the DAϕNE BTF", *Nucl. Instrum. Meth. A* **862** (2017) 31, doi:10.1016/j.nima.2017.05.007, arXiv:1611.05649.

[121] F. Oliva, "The charged particle veto system of the PADME experiment", *Nucl. Instrum. Meth. A* **936** (2019) 259, doi:10.1016/j.nima.2018.10.147.

[122] A. Frankenthal et al., "Characterization and performance of PADME's Cherenkov-based small-angle calorimeter", *Nucl. Instrum. Meth. A* **919** (2019) 89, doi:10.1016/j.nima.2018.12.035, arXiv:1809.10840.

[123] F. Ferrarotto, "The PADME calorimeters for missing mass dark photon searches", *JINST* **13** (2018) C03044, doi:10.1088/1748-0221/13/03/C03044.

[124] NA62 Collaboration, "The beam and detector of the NA62 experiment at CERN", *JINST* **12** (2017) P05025, doi:10.1088/1748-0221/12/05/P05025, arXiv:1703.08501.

[125] Muon g-2 Collaboration, "Muon (g-2) technical design report", Technical Report FERMILAB-FN-0992-E, FERMILAB-DESIGN-2014-02, 2015. arXiv:1501.06858.

[126] A. Anastasi et al., "Electron beam test of key elements of the laser-based calibration system for the muon g-2 experiment", *Nucl. Instrum. Meth. A* **842** (2017) 86, doi:10.1016/j.nima.2016.10.047, arXiv:1610.03210.

[127] S. Baunack et al., "Real-time calibration of the A4 electromagnetic lead fluoride

(PbF2) calorimeter", *Nucl. Instrum. Meth. A* **640** (2011) 58, doi:10.1016/j.nima.2011.02.099.

[128] A. Fienberg et al., "Studies of an array of PbF2 cherenkov crystals with large-area SiPM readout", *Nucl. Instrum. Meth. A* **783** (2015) 12, doi:10.1016/j.nima.2015.02.028, arXiv:1412.5525.

[129] Hamamatsu, "Photomultiplier Tubes R13478, R13449, R13408, R13089", 2017. [Datasheet].

[130] Hamamatsu, "Metal package photomultiplier tube R9880U series", 2017. [Datasheet].

[131] GEANT4 Collaboration, "GEANT4—a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[132] R. Wigmans and M. T. Zeyrek, "On the differences between calorimetric detection of electrons and photons", *Nucl. Instrum. Meth. A* **485** (2002) 385, doi:10.1016/S0168-9002(01)02141-6.

[133] A. Barysevich et al., "Radiation damage of heavy crystalline detector materials by 24 GeV protons", *Nucl Instrum Meth A* **701** (2013) 231, doi:10.1016/j.nima.2012.10.073.

[134] D. Anderson, M. Kobayashi, C. Woody, and Y. Yoshimura, "Lead fluoride: An ultra-compact Cherenkov radiator for EM calorimetry", *Nucl. Instrum. Meth. A* **290** (1990) 385, doi:10.1016/0168-9002(90)90553-I.

[135] I. Malitson and M. Dodge, "Refraction and dispersion of lead fluoride", *J. Opt. Soc. Am.* **59** (1969) 500A.

[136] V. Kraus et al., "FITPix — fast interface for Timepix pixel detectors", *JINST* **6** (2011) C01079, doi:10.1088/1748-0221/6/01/C01079.

[137] J. E. Gaiser, "Charmonium Spectroscopy from Radiative Decays of the J/ψ and ψ". PhD thesis, Stanford University, 1982. SLAC Report SLAC-R-0255.

[138] E. Leonardi, M. Raggi, and P. Valente, "Development and test of a DRS4-Based DAQ system for the PADME experiment at the DAΦNE BTF", *J. Phys.: Conf. Ser.* **898** (2017) 032024, doi:10.1088/1742-6596/898/3/032024.

[139] NA64 Collaboration, "Proposal for an experiment to search for light dark matter at the SPS", arXiv:1312.3309.

[140] NA64 Collaboration, "Search for vector mediator of dark matter production in invisible decay mode", *Phys. Rev. D* **97** (2018) 072002, doi:10.1103/PhysRevD.97.072002, arXiv:1710.00971.

[141] J. L. Feng et al., "Particle physics models for the 17 MeV anomaly in Beryllium nuclear decays", *Phys. Rev. D* **95** (2017) 035017, doi:10.1103/PhysRevD.95.035017, arXiv:1608.03591.

[142] I. G. Irastorza and J. Redondo, "New experimental approaches in the search for axion-like particles", *Prog. Part. Nucl. Phys.* **102** (2018) 89, doi:10.1016/j.ppnp.2018.05.003, arXiv:1801.08127.

[143] L. Marsicano et al., "Novel way to search for light dark matter in lepton beam-dump experiments", *Phys. Rev. Lett.* **121** (2018) 041802, doi:10.1103/PhysRevLett.121.041802, arXiv:1807.05884.

[144] L. Marsicano, "Light dark matter searches with positrons at JLab", 2020. [Presented at Snowmass's Cross-frontier meeting (AF - EF - RF): Dark sectors and Light Long-lived particles, 16 July 2020].

[145] J. Alexander, "MMAPS: Missing-Mass A-Prime Search", *EPJ Web Conf.* **142** (2017) 01001, doi:10.1051/epjconf/201714201001.

[146] J. Spencer, "Internal targets in storage rings", in *12th IEEE Particle Accelerator Conference (PAC1987): Accelerator Engineering and Technology*. 1986. [SLAC-PUB-4089].

[147] CERN, "Deliberation document on the 2020 Update of the European Strategy for Particle Physics", Technical Report CERN-ESU-016, Geneva, 2020.

[148] LDMX Collaboration, "Light Dark Matter eXperiment (LDMX)", Technical Report FERMILAB-PUB-18-324-A, SLAC-PUB-17303, 2018. arXiv:1808.05219.

[149] J. L. Feng, I. Galon, F. Kling, and S. Trojanowski, "ForwArd Search ExpeRiment at the LHC", *Phys. Rev. D* **97** (2018) 035001, doi:10.1103/PhysRevD.97.035001, arXiv:1708.09389.

[150] A. Haas, C. S. Hill, E. Izaguirre, and I. Yavin, "Looking for milli-charged

particles with a new experiment at the LHC", *Phys. Lett. B* **746** (2015) 117, doi:10.1016/j.physletb.2015.04.062, arXiv:1410.6816.

[151] H. Lubatti et al., "MATHUSLA: A detector proposal to explore the lifetime frontier at the HL-LHC", arXiv:1901.04040.

[152] S. Alekhin et al., "A facility to search for hidden particles at the CERN SPS: The SHiP physics case", *Rep. Prog. Phys.* **79** (2016) 124201, doi:10.1088/0034-4885/79/12/124201, arXiv:1504.04855.

[153] J. P. Delahaye et al., "Muon colliders", arXiv:1901.06150.

[154] S. Michizono, "The International Linear Collider", *Nat. Rev. Phys.* **1** (2019), no. 4, 244, doi:10.1038/s42254-019-0044-4.

[155] M. Aicheler et al., "A Multi-TeV Linear Collider Based on CLIC Technology: CLIC Conceptual Design Report". CERN Yellow Reports: Monographs. CERN, Geneva, 2012. doi:10.5170/CERN-2012-007.