

A Collider Search for Dark Matter Produced in Association
with a Higgs Boson with the CMS Detector at the 13 TeV LHC

By

DUSTIN RAY BURNS

B.S. (Georgia Institute of Technology) 2011

M.S. (University of California, Davis) 2012

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Physics

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Chair Michael Mulhearn

Robin Erbacher

Albert De Roeck

Committee in Charge

2017

Copyright © 2017 by
Dustin Ray Burns
All rights reserved.

To my teachers.

CONTENTS

List of Figures	vi
List of Tables	xii
Abstract	xv
Acknowledgments	xvi
1 Motivation and theoretical background	1
1.1 The Standard Model	1
1.1.1 Symmetries of nature	1
1.1.2 Particle content	8
1.2 Dark matter	11
1.2.1 Background	11
1.3 Beyond the Standard Model	17
1.3.1 Collider searches for DM	17
1.3.2 Signal models	19
2 Experimental apparatus	30
2.1 Large Hadron Collider	30
2.2 Compact Muon Solenoid	33
2.2.1 Tracking detectors	35
2.2.2 Electromagnetic calorimeter	37
2.2.3 Hadronic calorimeter	40
2.2.4 Muon detectors	42
2.2.5 Trigger	45
3 Event reconstruction and simulation	49
3.1 Particle reconstruction	49
3.2 Monte Carlo event simulation	52
3.3 Datasets	54
3.3.1 Data	54

3.3.2	Simulation	57
4	Physics objects	61
4.1	Electrons	61
4.1.1	Electron reconstruction	61
4.1.2	Electron identification	61
4.1.3	Electron isolation	65
4.1.4	Electron energy calibrations	65
4.1.5	Electron efficiency measurements	66
4.2	Muons	70
4.2.1	Muon reconstruction and identification	70
4.2.2	Muon isolation	71
4.2.3	Muon efficiency measurements	71
4.3	Photons for FSR recovery	73
4.4	Jets	76
4.5	MET	77
5	Signal and control regions	82
5.1	Event selection	82
5.1.1	Trigger selection	82
5.1.2	Vertex selection	82
5.1.3	ZZ candidate selection	82
5.1.4	Choice of the best ZZ candidate	85
5.2	Signal region and blinding	86
5.2.1	Cut-and-count based signal region	87
5.2.2	MVA-based signal region	88
5.3	Background modeling	88
5.3.1	Irreducible backgrounds	88
5.3.2	Reducible background estimation	92
5.4	Event yields	98

6 Statistical analysis results	100
6.1 Systematic uncertainties	100
6.1.1 Uncertainties on the reducible background estimation	100
6.1.2 MET systematics	101
6.1.3 Additional systematics	102
6.2 Limit setting	104
6.2.1 Cross section limits	107
7 Conclusions	111
Appendices	113
A Production cross sections for benchmark signal models	114

LIST OF FIGURES

1.1 Particles of the Standard Model.	9
1.2 Four-lepton invariant mass distribution showing H discovery in the ZZ^* decay channel. The red line shows the signal distribution for $m_H = 126$ GeV [1].	12
1.3 (a) Mono- X production topology for signatures with X emitted as ISR. (b) Mono- X production topology for signatures with X emitted from new physics vertex.	18
1.4 Collider production diagram for mono-Higgs effective field theories.	20
1.5 Collider production diagram for Zp2HDM.	22
1.6 Collider production diagram for Z' models.	24
1.7 Collider production diagram for Scalar model.	27
2.1 Deconstructed view of the CMS subdetectors, with human figure for scale. From inside to out, the colored segments correspond to the following systems: light brown is the pixel tracker, cream is the strip tracker, green is the EM calorimeter, orange is the hadronic calorimeter, grey is the solenoid, red is the yoke with white muon chambers [2].	35
2.2 Schematic diagram of tracking detectors with radial distance of modules, shown as black lines, from center on the left axis, z -dimension on the bottom axis, and η across the top [2].	36
2.3 Deconstructed barrel pixel module showing module components [2].	37
2.4 Average hit efficiencies of the strip tracker layers [2].	38
2.5 Schematic layout of the ECAL crystal modules [2].	39
2.6 ECAL barrel crystal with one depolished face attached to avalanche photodiode photodetector (left) and ECAL end cap crystal attached to vacuum phototriode (right) [2].	40
2.7 Cross sectional view of one quadrant of CMS. The labeled sections are the subsystems of the hadronic calorimeter [2].	41

2.8	Energy resolution of HCAL systems for test beam pions [2].	42
2.9	Schematic diagram showing DT chambers in light blue [2].	44
2.10	Schematic diagram showing CSC locations in red in a quadrant cutout of CMS [2].	44
2.11	Schematic diagram showing barrel RPC locations (left) and end cap RPCs (right) in a cross sectional cutout of CMS [2].	45
2.12	Transverse momentum resolution versus pseudorapidity for muons from Z decays [3].	46
2.13	CMS L1 trigger system schematic diagram [2].	47
2.14	CMS data acquisition system schematic diagram [2].	47
3.1	Event display of hadronic jet in the $x - y$ plane, with solid arcs at the ECAL and HCAL surfaces. The locations of clusters are given by the solid dots [4].	51
3.2	Event display of hadronic jet in the $\eta - \phi$ plane for the ECAL (a) and HCAL (b). The locations of clusters are given by the solid dots [4]. . . .	52
3.3	Missing transverse energy reconstruction resolution using the particle flow algorithm [4].	52
3.4	Trigger efficiency measured in data using 4ℓ events collected by single lepton triggers for the $4e$ (top left), 4μ (top right), $2e2\mu$ (bottom left) and 4ℓ (bottom right) final states.	57
3.5	Number of pileup vertices before (a) and after (b) reweighting is applied.	60
4.1	Electron reconstruction efficiencies in data versus η and data/MC scale factors as provided by the EGM POG.	62
4.2	Performance comparison of the MVA trained for the 2015 analysis and the retraining for 2016 conditions. The respective working points are indicated by the markers.	63
4.3	Boosted decision tree output for the training and testing sample for true and fake electrons in the high- p_T end cap training bins.	64

4.4	(a): Electron energy scale measured in the $Z + \ell$ control region for EB and EE electrons. The results of the Crystall-ball fit are reported in the figure. (b): lepton energy scales per 500 pb luminosity block.	67
4.5	Electron selection efficiencies measured using the tag-and-probe technique described in the text, non-gap electrons (a) and gap electrons (b).	68
4.6	Electron selection efficiencies measured using the tag-and-probe technique described in the text, non-gap electrons (a) and gap electrons (b)	69
4.7	Muon reconstruction and identification efficiency at low p_T , measured with the tag-and-probe method on J/Ψ events, as function of p_T in the barrel (a) and end caps (b), and as function of η for $p_T > 7$ GeV (c). In the upper panel of each graph, the larger error bars include also the systematical uncertainties, while the smaller ones are purely statistical. Each graph's lower panel shows the ratio of the two efficiencies, the black error bars are for the statistical uncertainty, the orange rectangles for the systematic uncertainty, and the violet rectangles include both uncertainties.	72
4.8	Efficiency of the muon impact parameter requirements, measured with the tag-and-probe method on Z events, as function of p_T in the barrel (a) and end caps (b), and as function of η for $p_T > 20$ GeV (c). In the upper panel of each graph, the larger error bars include also the systematical uncertainties, while the smaller ones are purely statistical. Each graph's lower panel shows the ratio of the two efficiencies, the black error bars are for the statistical uncertainty, the orange rectangles for the systematical uncertainty, and the violet rectangles include both uncertainties.	73

4.9	Efficiency of the muon isolation requirement, measured with the tag-and-probe method on Z events, as function of p_T in the barrel (left) and end caps (right). In the upper panel of each graph, the larger error bars include also the systematical uncertainties, while the smaller ones are purely statistical. The lower panel of each graph shows the ratio of the two efficiencies, the black error bars are for the statistical uncertainty, the orange rectangles for the systematical uncertainty, and the violet rectangles include both uncertainties.	74
4.10	Tracking efficiency in data and simulation as a function of η for muon $p_T < 10$ GeV(a) and $p_T > 10$ GeV(b) with ReReco data.	75
4.11	(a) Overall data to simulation scale factors for muons, as function of p_T and η . (b) Uncertainties on data to simulation scale factors for muons, as function of p_T and η	75
4.12	Missing transverse energy (MET) after the standard model selection before removal of fake MET.	79
4.13	Maximum azimuthal angle difference between MET and jets.	80
4.14	Minimum azimuthal angle difference between MET and jets.	81
5.1	Dimuon invariant mass distributions for Z_1	83
5.2	Dimuon invariant mass distributions for Z_2	84
5.3	Four-muon invariant mass distribution in log scale.	85
5.4	Four-muon invariant mass distribution in linear scale.	86
5.5	MVA response trained on a weighted background sample and admixture of signal samples with input variables m_{4l} and E_T^{MISS} . Simulated backgrounds are in pink and signals are in blue.	89
5.6	Left: NNLO/NLO QCD k -factor for the $q\bar{q} \rightarrow ZZ$ background as a function of $m(ZZ)$ for the 4ℓ and $2\ell 2\ell'$ final states. Right: NLO/NLO electroweak k -factor for the $q\bar{q} \rightarrow ZZ$ background as a function of $m(ZZ)$	90

5.7	$gg \rightarrow H \rightarrow 2\ell 2\ell'$ cross sections at NNLO, NLO and LO at each H pole mass using the SM H decay width (top left) or at the fixed and small decay width of 4.07 MeV (top right). The cross sections using the fixed value are used to obtain the k -factor for both the signal and the continuum background contributions as a function of $m_{4\ell}$ (bottom).	91
5.8	Fake rates as a function of the probe p_T for electrons (a) and muons (b) which satisfy the loose selection criteria. The fake rates are measured in a $Z(\ell\ell) + \ell$ sample in the 13 TeV data. The barrel selection includes electrons (muons) up to $ \eta = 1.479$ (1.2).	93
5.9	Invariant mass distribution of the events selected in the 2P+2F control sample in the 13 TeV dataset, (top left) 4μ , (top right) $4e$, (bottom left) $2\mu 2e$ and (bottom right) $2e 2\mu$ channels.	95
5.10	Invariant mass distribution of the events selected in the 3P+1F control sample in the 13 TeV dataset, (top left) 4μ , (top right) $4e$, (bottom left) $2\mu 2e$ and (bottom right) $2e 2\mu$ channels.	96
6.1	Difference between the $Z \rightarrow \ell\ell$ mass peak positions in data and simulation normalized by the nominal Z boson mass obtained as a function of the p_T and $ \eta $ of one of the leptons, regardless of the second, for muons (left) and electrons (right).	103
6.2	Different $m_{4\ell}$ distributions after propagating the biases in Fig. 6.1 to Higgs boson events. The change in the mean of the double Crystal-Ball is used to determine the systematic uncertainty due to the lepton momentum scale. The middle plot shows the nominal distribution, while the left (right) plots show the down (up) systematic variations. The 4μ channel is shown in the top row and the $4e$ channel is shown in the bottom row.	104
6.3	One-dimensional cross section times branching fraction limits for the Zp2HDM simplified model using the cut-and-count based event selection strategy. .	108

6.4	One-dimensional cross section times branching fraction limits for the Zp-Baryonic simplified model using the cut-and-count based event selection strategy.	109
6.5	One-dimensional cross section times branching fraction limits for the Zp2HDM simplified model using the MVA-based event selection strategy.	110
6.6	One-dimensional cross section times branching fraction limits for the Zp-Baryonic simplified model using the MVA-based event selection strategy.	110

LIST OF TABLES

1.1	Effective Field Theory Models [5]	20
1.2	Constraints on effective field theory model parameters [5].	21
1.3	New particles and parameters of the Zp2HDM simplified model [5]	23
1.4	New particles and parameters of the ZpBaryonic simplified model [5]	24
1.5	New particles and parameters of the ZpHS simplified model [5]	25
1.6	New particles and parameters of the Scalar simplified model [5]	27
1.7	Benchmark parameter choices for the simplified models [5, 6].	28
1.8	Benchmark mass points for models with a vector mediator [6].	28
1.9	Theoretical and experimental constraints on simplified model parameters [5, 6].	29
1.10	Benchmark mass points for the Zp2HDM simplified model [6].	29
1.11	Benchmark mass points for models with a scalar mediator [6].	29
3.1	Run 2 datasets used in the $H \rightarrow ZZ$ analysis. The first column specifies the range of run ids, indices counting up from the first CMS run, for each dataset. The middle column gives the CMS database file path where the data files are stored. The third column gives the total integrated luminosities for each collection of runs.	55
3.2	Trigger paths used in 2016 collision data.	56
3.3	Trigger efficiencies measured using 4ℓ events.	58
3.4	Benchmark signal samples analyzed.	58
3.5	Higgs signal samples and production cross sections times branching fractions to four leptons times filter efficiencies.	59
3.6	Background Monte Carlo samples and cross sections.	59
4.1	Overview of input variables to the identification classifier. Variables not used in the Run 1 MVA are marked with (\cdot)	64

4.2	Minimum boosted decision tree score required for passing the electron identification.	65
4.3	The requirements for a muon to pass the Tracker High- p_{T} ID. Note that these are equivalent to the Muon POG High- p_{T} ID with the global track requirements removed.	70
5.1	Cross sections for $q\bar{q} \rightarrow ZZ$ production at 13 TeV	90
5.2	The contribution of reducible background processes in the four-lepton signal region predicted from measurements in data using the Opposite-Sign Leptons method. The predictions correspond to 35.9 fb^{-1} of data at 13 TeV.	97
5.3	Cutflow table for the 4μ channel. The benchmark signal sample shown is Zp2HDM with $m_{Z'} = 600 \text{ GeV}$	98
5.4	Final signal region yields using the cut-and-count selection strategy for the 4μ channel. The benchmark signal sample shown is Zp2HDM with $m_{Z'} = 600 \text{ GeV}$	98
5.5	Final signal region yields using the MVA selection strategy for the 4μ channel. The benchmark signal sample shown is Zp2HDM with $m_{Z'} = 600 \text{ GeV}$	99
6.1	Summary of the experimental systematic uncertainties in the $H \rightarrow 4\ell$ measurements.	104
6.2	Summary of the theory systematic uncertainties in the $H \rightarrow 4\ell$ measurements for the inclusive analysis.	105
A.1	Effective field theory model production cross sections [pb]	115
A.2	Scalar simplified model production cross sections [pb] corresponding to mass points in Table 1.11	116
A.3	ZpBaryonic simplified model production cross sections [pb] corresponding to mass points in Table 1.8	117
A.4	ZpHS simplified model production cross sections [pb] corresponding to mass points in Table 1.8	118

A.5 Zp2HDM simplified model production cross sections [pb] corresponding to mass points in Table 1.10	119
---	-----

ABSTRACT

A Collider Search for Dark Matter Produced in Association with a Higgs Boson with the CMS Detector at the 13 TeV LHC

The study presented in this dissertation is a search for dark matter produced in 13 TeV proton-proton collisions at the Large Hadron Collider (LHC) using 35.9 fb^{-1} of data collected in 2016 with the Compact Muon Solenoid (CMS) detector. Dark matter escapes the detector without interacting, resulting in a large imbalance of transverse momentum, which can be observed when a Higgs boson is tagged in the opposite direction. A variety of models which motivate a dark matter and Higgs interaction are discussed. The experimental signature of these models is called mono-Higgs.

In this search, the Higgs is produced primarily from gluon fusion and decays to four leptons via two Z bosons ($H \rightarrow ZZ \rightarrow 4\ell$). In addition to rediscovering the Higgs in the four-lepton final state, cross-validating the analysis strategy with other teams, an extensive study of missing transverse energy (MET) is required to search for the mono-Higgs signature.

Once a background model is developed for the Standard Model processes that can mimic the signal, a counting experiment is performed in an optimized signal region. There is no evidence for an excess of events in the signal region above the backgrounds, so cross section limits are set for two kinematically distinct signal models.

ACKNOWLEDGMENTS

I would like to thank the numerous groups and individuals who have facilitated my academic journey.

First, the teachers and professors who made a particularly strong impact on my education or personal and professional development: my graduate adviser Michael Mulhearn, Professors Mani Tripathi and Robin Erbacher from UC Davis, Professors David Finkelstein, Dierdre Shoemaker, and Yuri Bakhtin from Georgia Tech, Professors Yosi Balas, Catherine Davis, and Imad El-Jeaid from Middle Georgia College, and Scott Hammond, Kate Hoppenrath, Cherie McAdams, Terasa Martin, Vicki Dobbs, and Bonnie Rary from Dacula High School. Lastly, Peg Alton from Dacula Middle School, who convinced me to join the gifted program, the first step on my journey of academic accomplishment. Even a small conversation can make a very large impact.

Second, I would like to thank my primary collaborators in CMS, Giorgia Miniello and Nicola De Filippis, who helped make the completion of this analysis possible. I am honored to be a member of the same team that discovered the Higgs boson in the four-lepton final state, and I am grateful for having access to their expertise, feedback, and documentation. It has been a very rewarding process to watch the different mono-Higgs analysis teams grow and succeed, after having organized the initial effort in CMS myself. I am grateful for having the chance to work with the analyzers on all of the mono-Higgs teams. Harrison Prosper, Shalhout Shalhout, and Eiko Yu have always made themselves available to help with technical questions for my analysis. I am also appreciative of the research and personal support I have received from Francesca Ricci-Tam, Wells Wulsin, Rachel Yohay, and Justin Pilot.

I am lucky to have had a great group of friends in my cohort, including Karen Ng, Henry Stoltenberg, Dusty Stolp, Chad Flores, and James Morad, who have encouraged and supported me throughout graduate school. Kyle Tos, Christine McLean, and Mengyao Shi have been exceptional friends and office-mates. I have relied heavily on the help and friendship from students in my lab, including Zhangqier Wang, Chris Brainerd, and Brandon Buonacorsi.

Finally, I would like to thank my family: my parents David Burns and Tina Liston, for giving me the opportunity to succeed, and my wife, Kayla Burns, for providing me with stability and support throughout college, and for thoroughly editing this document. Appa and Dax Burns have helped me work through many research issues on our evening walks.

Chapter 1

Motivation and theoretical background

The first section of this chapter lays the theoretical framework of the Standard Model (SM) of particle physics, including the historical discovery timeline for symmetries of nature, each corresponding to a leap forward in our physical understanding. The second section gives the SM particle content and their properties. The last section extends the symmetry group of the SM to include candidates for the observed cosmic dark matter (DM) and the particles that could mediate the DM–SM interactions.

1.1 The Standard Model

1.1.1 Symmetries of nature

Each major advance in the history of physics has corresponded to the discovery and mathematical implementation of a new global space-time, global discrete, or local gauge symmetry. In this section, I will review these discoveries, informally developing the mathematical framework needed to understand the symmetry groups of the SM. Along the way, two important subplots will play out: (1) the development of our understanding of physics at smaller distance scales and higher energies and (2) the unification of previously separate physical sectors.

Our ancient ancestors were aware that certain geometrical shapes, e.g. the Platonic solids, possessed the quality of symmetry, and they were driven to understand the compo-

sition of physical substances by breaking them down into fundamental, indivisible units. These units, known as atoms by the ancient Greeks, interact and rearrange themselves according to physical laws to account for the variety of substances and physical phenomena we observe. The attraction to applying symmetry to nature is evidenced by the centuries-long belief that the Earth lie at the center of the universe, with the celestial bodies orbiting in perfect, divine circles.

The end of the scientifically repressive Middle Ages brought along an improvement in astronomical observations and the growth of the pseudo-scientific field of alchemy, which attempted to reduce, understand, and manipulate the fundamental elements of physical substances. When Johannes Kepler discovered his three laws of planetary motion, he unified the description of the motion of the planets. For the first time, the conservation of a physical quantity, what we now know as angular momentum, was associated with a general set of physical objects. At the burgeoning of the scientific revolution, Isaac Newton championed the idea that the same physical laws can be applied to all physical events and that properties of these laws can be abstracted to apply to nature at a fundamental level.

Newton defined an inertial reference frame implicitly as one where his first law holds, that is, that an object remains at constant motion unless acted on by an outside force. Since his laws were the same in all inertial reference frames, a new symmetry of nature was discovered, now called symmetry under Euclidean transformations. Since Euclidean transformations form a mathematical group, it is said that classical mechanics is invariant under the Euclidean group. The invariance under the Euclidean transformations can be used to derive conservation laws: Newton's laws do not depend on spatial translations or rotations, implying the conservation of linear and angular momentum, respectively. These relationships foreshadow Emmy Noether's abstraction of the connection between continuous symmetries and conserved quantities. She proved that there is a conserved quantity, or current, associated with every symmetry of a physical system [7]. This famous theorem facilitates the derivation of conserved quantities and will be used extensively in the theories that follow. The extension of the Euclidean transformations to include time

translations and motion at constant velocity (boosts) forms the Galilean group.

The next symmetry of nature to be discovered came when Hendrik Lorentz found that James Clerk Maxwell's equations, which unified the classical electricity and magnetism sectors [8], were invariant under a set of transformations that generalized the classical Galilean translations, now called Lorentz transformations, which form the Lorentz group [9]. The Lorentz symmetry corresponds to the conservation law of total electric charge [10]. Soon after, Albert Einstein derived the Lorentz symmetry as a property of space-time itself through his special theory of relativity, assuming two simple postulates: the laws of physics and the speed of light are the same in all inertial reference frames [11]. From these simple assumptions, Einstein was able to extend the classical laws of physics to the high velocity, high energy sector. Before Einstein, the symmetries of nature were thought to be consequences of the physical laws themselves, but Einstein's major paradigm shift was to view the symmetry itself as the more fundamental property, an insight key to his formulation of the general theory of relativity for the gravitational interaction [12]. Combining the Lorentz and Euclidean transformations yields the Poincare group, the final global space-time symmetry of the SM.

Just as relativists were probing physics at higher speeds and energies, other physicists were investigating the behavior of systems at smaller distance scales, conducting experiments to explore phenomena such as the Compton effect [13] and photoelectric effect [14], which shows the quantized, particle-like behavior of light, and electron beam diffraction [15], which shows the wave-like behavior of electrons. Quantum theory was developed to consolidate the wave-like and particle-like behaviors of systems, and it extended the validity of classical physics to microscopic scales. Where in classical physics particles' states were described by their absolute position and momentum, quantum theory describes the state of a system by a probabilistic wavefunction. The classical symmetries and conservation laws carried over to quantum theory, with the concept of angular momentum generalized to include spin. Particles with zero or integer spin are called bosons, and were shown to obey Bose-Einstein statistics, where the wavefunction is symmetric under all permutations of particles. Particles with half-integer spin are called fermions, and

they obey Fermi-Dirac statistics, where the wavefunction is symmetric under even permutations and changes sign under odd permutations, implying that no two particles can occupy the same state. These many-particle symmetries were applied to derive properties of materials and states of electrons in atomic and periodic systems, blackbody radiation, and many other properties of matter and radiation [16].

Quantum theory was successfully applied to a myriad of low-energy systems. Paul Dirac extended quantum theory and the Schrödinger equation, which describes the evolution of a quantum system in time, to the relativistic regime [17]. His formulation of the wave equation for a spin- $\frac{1}{2}$ particle with mass m is inherently Lorentz invariant:

$$(i\gamma^\mu \partial_\mu - m)\psi = 0 \quad (1.1)$$

The solutions to the Dirac equation were found to have both positive and negative energy solutions, to the surprise of Dirac. His explanation was that the vacuum consisted of a “sea” of negative energy solutions, each in a distinct state due to the Pauli exclusion principle, and when a pair of electrons was produced, a positive energy state was filled and a negative energy state was vacated, creating a “hole” in the sea. The particle corresponding to this hole would have the same energy as the electron but, in order to conserve total charge, must have the oppositely signed charge. This positively charged electron was not known to exist at the time, but it was soon discovered in cosmic ray experiments and named the positron [18]. Dirac’s theoretical prediction of the positron and its subsequent discovery opened the door for the discovery that every particle has an oppositely charged antimatter partner.

In an attempt to apply relativistic quantum theory to the electromagnetic (EM) field and the spin-0 photon, Dirac formulated the theory of quantum electrodynamics (QED) [19]. QED is the first example of a quantum field theory (QFT), where the physical dynamics apply to the quantum field associated with a particle rather than the particle itself, and particles and antiparticles interact as excitations of the field. This formulation of creation and annihilation of particles and antiparticles gave a more physically intuitive explanation for the negative energy solutions of the Dirac equation than the Dirac

“sea”. QED became the prototype for developing future relativistic quantum theories, and Dirac’s procedure became the template for quantizing a general field theory.

With the discovery of antimatter and the apparatus of QFT in place, physicists continued searching for additional symmetries. If a particle state is an eigenvector of a symmetry operator, then the eigenvalue is an important quantum number of the state, since if the symmetry is conserved, the quantum number can be used to determine if a decay or interaction of this state is allowed. Three important discrete symmetries operations, and their products, have had significant importance: space coordinate parity inversion (P), particle-antiparticle charge conjugation (C), and time reversal (T) [20]. Violation of P was observed for the weak interactions in the decays of Cobalt-60 atoms [21]. Although it was long believed to be a symmetry of all the interactions, CP , the product of C and P was found to be violated by the weak decays of K mesons [22]. It is only the combination of P , C , and T that is now thought to be an exact symmetry of nature, obeyed by all interactions. The CPT symmetry was proven to be conserved in all relativistic QFTs [23, 24].

The mid-twentieth century saw an explosion of discovery in particle physics and was a golden age for the feedback between theoretical and experimental work. The experimental discovery of additional particles such as the muon [25], pion [26], neutrino [27], and many others inspired the theoretical development of the quark model [28], the refinement of QED [29, 30, 31, 32, 33, 34, 35, 36], and the formulation of theories to explain the strong [28, 37, 38, 39, 40] and weak [41] nuclear forces. Conversely, these new theories led to the prediction and subsequent experimental discovery of new fundamental particles, such as the charm [42, 43] and bottom [44] quarks, and composite particles, such as the Ω^- [45].

The development of the theories of the strong and weak forces unveiled a new set of symmetries: local invariance under unitary gauge transformations. In keeping with the trend of the discovery of new symmetries being associated with the unification of physical sectors, the electromagnetic and weak nuclear interactions were found to be components of a single force, called the electroweak force [46, 47]. Again, with QED as the prototype, the Lagrangians for the electroweak and strong forces were constructed to be invariant

under unitary groups. Early work by Hermann Weyl [48], showing the gauge invariance of EM, was extended to QED and lay the mathematical framework for describing the gauge invariance of the other forces. To demonstrate gauge invariance and how the gauge group generators are associated with the force carriers, take the QED terms of the SM Lagrangian as an example case [49]:

$$\mathcal{L} \supset \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \quad (1.2)$$

where $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ is the electromagnetic field tensor, A_μ is the EM four-potential and field corresponding to the photon, and $D_\mu = \partial_\mu + ieA_\mu$ is the gauge covariant derivative. These terms are invariant under the U(1) transformations:

$$\psi \rightarrow e^{i\theta(x)}\psi. \quad (1.3)$$

In general, the generators of an interaction's continuous symmetry group correspond to the gauge fields whose excitations are the gauge bosons that mediate that interaction. Just how A_μ is the field corresponding to the photon in QED and generator of the U(1) symmetry group of EM, the W and Z bosons and gluons are formed from the generators of the symmetry groups for the electroweak and strong interactions, SU(2) \times U(1) and SU(3), respectively. Additionally, the covariant derivative is transformed in a way analogous to the QED covariant derivative, adding a factor for each generator with the corresponding charge in the coefficient. These charges (electric, color, weak isospin and hypercharge) are conserved for all of the forces in all interactions [20].

Although the mathematical descriptions of their symmetries are similar, the EM, weak, and strong nuclear forces are quite different qualitatively. The EM force is the most familiar, being responsible for the interactions of matter and radiation at macroscopic scales and having only one type of charge which is either positive or negative. The strong force is mediated by eight gluons, corresponding to the eight generators of SU(3), and carries three types of charge, known as red(R), green(G), and blue(B). The “negatives” of these charges are called anticolors: antired(\bar{R}), antigreen(\bar{G}), and antiblue(\bar{B}). Keeping with the visible color analogy, the theory of the strong force is called quantum chromodynamics (QCD).

At low energies, the quarks and antiquarks are confined to form only color-neutral states. This is known as color confinement and implies no free quarks exist. Instead, quarks only come in neutral combinations called hadrons. Since hadrons are strictly color-neutral, their states transform as the singlet representation under SU(3). At high energies, the strong coupling decreases, and the quarks can be treated as effectively free particles with perturbation theory using Feynman diagrams. This property of the strong force is known as asymptotic freedom. The weak force is qualitatively different than either the EM or strong forces, as particles do not exchange its mediators to form bound states of any kind. The low-energy limit of the weak theory was developed by Enrico Fermi to explain beta decays [41], but the full description was not developed until the electroweak force was formulated.

The unification of the electromagnetic and weak interactions into the electroweak force at about 100 GeV had one major shortcoming: the invariance of the Lagrangian required the gauge bosons, B from U(1) and W^1, W^2, W^3 from SU(2), to be massless. While the photon is indeed massless, the W and Z bosons have a nonzero mass. Higgs and others [50, 51, 52, 53, 54, 55] proposed a solution by introducing a new scalar field whose excitations were called the Higgs boson (H). The scalar field is a complex doublet, meaning it has four total real components, and its vacuum expectation value (VEV), or value throughout all of space, is one of many nonzero values in the bottom of its “Mexican hat” potential. This spontaneous symmetry can be broken by choosing one of the values for the VEV: $H = \frac{1}{\sqrt{2}}(0, \nu)$, where $\nu = 246$ GeV is called the H VEV. At energies above $O(100)$ GeV, the electroweak symmetry is obeyed, the gauge bosons are massless, and the Higgs field has one of many values along the circle at the base of its potential. When the H field is expressed as a perturbation about this VEV, the electroweak symmetry is broken into the weak and EM forces, and mass terms are generated for the weak force bosons, which are expressed as linear combinations of B, W^1, W^2, W^3 . The process of the Higgs acquiring a VEV, spontaneously breaking the symmetry in its potential, and generating the masses of the weak bosons is known as the Higgs mechanism, and it results in the breaking of the electroweak symmetry. This mechanism can be carried out for a

general operator depending on H with the substitution

$$H \rightarrow \frac{1}{\sqrt{2}}(\nu + h) \quad (1.4)$$

where h is the physical Higgs, corresponding to the leftover degree of freedom of the H doublet that is not absorbed by the three massive gauge bosons. H couples to the SM fermions, detailed in Section 1.1.2, not by the same mixing as described for the gauge bosons, but via Yukawa interactions, direct couplings whose coefficients are related to the fermion masses. The Higgs mechanism was the final piece of the puzzle to understanding the fundamental laws of the known particles.

The final result in this saga is the standard model of particle physics, a relativistic gauge quantum field theory, globally invariant under the Poincare group, and locally invariant under the product of the strong and electroweak symmetry groups, which describes all of the fundamental particles and their interactions. The next section details the particle content of the SM.

1.1.2 Particle content

The particle content of the SM is displayed in Figure 1.1

The particles of the standard model consist of the spin- $\frac{1}{2}$ fermions, which interact to form regular matter, the spin-1 bosons, which mediate the interactions of the fermions, and the spin-0 scalar Higgs boson, which generates the masses of the bosons and fermions via the Higgs mechanism [20].

The fermions come in three sets of increasing masses, called generations, corresponding to the first three columns of Figure 1.1. Across the rows, the particles have similar properties and are abbreviated u^i, d^i, e^i, ν_e^i from top to bottom, where the generation index $i = 1, 2, 3$. Within each generation, the fermions are divided into two categories: (1) the quarks, which are charged under the EM, weak, and strong forces, and (2) the leptons, which are charged under the EM and weak forces. Each fermion has a corresponding antiparticle, whose mass is the same, but whose charges are opposite in sign. The quarks are SU(3) triplets, having a color charge of either R, G, or B. Since the weak force violates P , the fermions can be distinguished by their chirality, labeled as either right-

Quarks	u up	c charm	t top	γ photon	H Higgs boson
	d down	s strange	b bottom	g gluon	
Leptons	e electron	μ muon	τ tau	W W boson	Gauge Bosons
	ν_e e neutrino	ν_μ μ neutrino	ν_τ τ neutrino	Z Z boson	

Figure 1.1: Particles of the Standard Model.

handed (e_R^i) or left-handed (e_L^i). The $(u^i, d^i)_L$ and $(e^i, \nu_e^i)_L$ pairs and their right handed antiparticle pairs are SU(2) doublets and interact via the weak force. The right-handed fermions (left-handed antifermions) are SU(2) singlets and do not interact via the weak force. The weak isospins (T_3) for the left handed fermions are: $(u^i, d^i)_L = (\frac{1}{2}, -\frac{1}{2})$ and $(e^i, \nu_e^i)_L = (-\frac{1}{2}, \frac{1}{2})$. The EM charges are $(u^i, d^i)_L = (\frac{2}{3}e, -\frac{1}{3}e)$ and $(e^i, \nu_e^i)_L = (-1e, 0)$, and finally, the weak hypercharges are, $(u^i, d^i)_L = (\frac{1}{3}, \frac{1}{3})$ and $(e^i, \nu_e^i)_L = (-1, -1)$, from the relation $Y_W = 2(Q - T_3)$.

The fourth column of Figure 1.1 lists the force mediators, or gauge bosons. The 8 gluons (g) of QCD that mediate the strong nuclear force are octets under SU(3) and correspond to linear combinations of the generator gauge fields, G_μ^a . Gluons carry color charge themselves but are electrically neutral. They are massless, consistent with the fact that they correspond to generators of a conserved symmetry, and may be represented

using the Gell-Mann matrices as the linearly independent set of states [56]:

$$\begin{aligned}
& \frac{1}{\sqrt{2}}(r\bar{b} + b\bar{r}) \\
& \frac{1}{\sqrt{2}}(r\bar{g} + g\bar{r}) \\
& \frac{1}{\sqrt{2}}(b\bar{g} + g\bar{b}) \\
& \frac{1}{\sqrt{2}}(r\bar{r} - b\bar{b}) \\
& -i\frac{1}{\sqrt{2}}(r\bar{b} - b\bar{r}) \\
& -i\frac{1}{\sqrt{2}}(r\bar{g} - g\bar{r}) \\
& -i\frac{1}{\sqrt{2}}(b\bar{g} - g\bar{b}) \\
& \frac{1}{\sqrt{6}}(r\bar{r} + b\bar{b} - 2g\bar{g}),
\end{aligned} \tag{1.5}$$

while the color singlet state that the colorless hadrons are in is:

$$\frac{1}{\sqrt{3}}(r\bar{r} + g\bar{g} + b\bar{b}). \tag{1.6}$$

The remaining gauge bosons, γ , W , and Z , mediate the electroweak force. Before electroweak symmetry breaking, the generators of $SU(2) \times U(1)$ correspond to the gauge fields

B , W^1 , W^2 , W^3 , whose excitations are massless gauge bosons. After symmetry breaking via the Higgs mechanism, three of the bosons acquire mass, and the electroweak gauge bosons are reparametrized as:

$$\begin{aligned}
W^\pm &= \frac{1}{\sqrt{2}}(W^1 \pm iW^2) \\
Z &= \cos \theta_w W^3 - \sin \theta_w B \\
\gamma &= \sin \theta_w W^3 + \cos \theta_w B
\end{aligned} \tag{1.7}$$

where θ_w is the weak mixing angle, the massive W^\pm and Z bosons mediate the weak force, and the massless γ is the photon that mediates EM. the W^\pm bosons have an electric charge

of $\pm 1e$, while the Z boson and γ are neutral. The isospin of W^\pm are ± 1 and 0 for Z and γ , giving hypercharges of 0 for W^\pm and 0 for Z and γ .

The final particle of the SM is the scalar H . H is electrically neutral and constructed to be an SU(2) doublet before electroweak symmetry breaking, with one component having weak isospin $\frac{1}{2}$ (hypercharge -1), and the neutral component having isospin $-\frac{1}{2}$ (hypercharge 1), which includes the physical h . After electroweak symmetry breaking, three of the H components are absorbed by the gauge bosons, and the remaining physical h remains neutral. The parity of h is 1. Although H couples to all massive fermions and bosons, the decay channels that are relevant for collider searches are: $ZZ^* \rightarrow 4l$, $WW^* \rightarrow 2l2\nu$, $\gamma\gamma$, $\tau\bar{\tau}$, and $b\bar{b}$. After a decades long search, the discovery and verification of quantum numbers of H was announced in 2012 [57, 58]. The four-lepton invariant mass distribution, showing the H peak at its observed mass of $m_H = 126$ GeV, is shown in Figure 1.2 [1].

1.2 Dark matter

1.2.1 Background

This section gives an overview of the most compelling sources of observational evidence for the existence of dark matter (DM), the potential particles candidates for DM, and the potential methods for detecting them.

1.2.1.1 Observational evidence

The earliest indication that there may be matter in the universe that cannot be detected by conventional optical observations, so called dark matter (DM), came from measurements of the orbital velocities of astronomical bodies in galaxy clusters [59, 60] and galaxies [61, 62]. Classical Newtonian gravity gives a galactic rotation curve, which shows how the velocity v of a massive object depends on its distance r from the center of a galaxy, as $v(r) \propto \sqrt{M(r)/r}$, where $M(r)$ is the total mass in the galaxy within a radius r . However, measurements of the orbital velocities of objects outside of the visible part of galaxies, where $v(r) \propto 1/\sqrt{r}$, show instead $v(r) \propto \text{constant}$, i.e. that the mass $M(r) \propto r$ instead of being constant [63]. The larger than expected velocities imply the existence of this

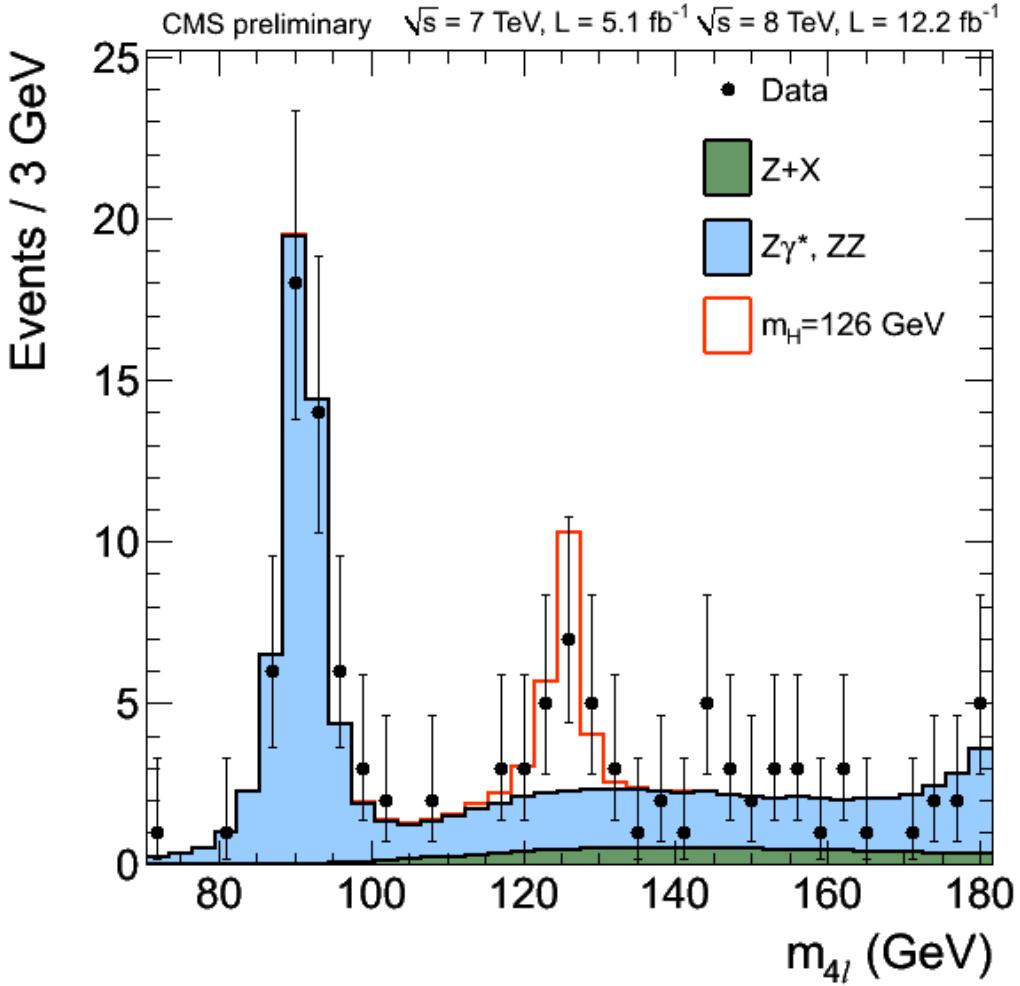


Figure 1.2: Four-lepton invariant mass distribution showing H discovery in the ZZ^* decay channel. The red line shows the signal distribution for $m_H = 126 \text{ GeV}$ [1].

spherically symmetric, dark halo of non-luminous matter in galaxies.

A compelling example showing direct evidence for DM is galaxy cluster 1E0657-558, often referred to as the “bullet cluster” from its bullet-like shape [64]. The bullet cluster passed through another galaxy cluster at some point in the recent cosmological past. The luminous matter, observed by traditional optical telescopes, is seen to lag behind the total mass of the clusters, observed by studying the weak gravitational lensing of objects in the background of the two clusters. The luminous matter in each cluster lags due to its EM interaction with the luminous matter in the opposing cluster, while what is inferred to be the DM continues on a ballistic trajectory, not experiencing lag from EM interactions

[63].

Many other cosmological observations and theories, including observations from strong gravitational lensing in elliptical galaxies [65], observations from weak lensing of distant galaxies by foreground matter [66], and modeling of anisotropies in the cosmic microwave background (CMB) [67, 68], strongly support the existence of DM, and other than a handful of competing theories which modify the laws of gravity instead of adding new matter, the existence of DM is widely accepted [69] and is believed to account for 20–40% of the mass density of the universe [70].

1.2.1.2 Particle candidates

From a variety of searches for different types of new dark particles, much more is known about what DM is not than what it is. Surveys are made to detect gravitational microlensing from massive compact halo objects (MACHOs), such as black holes, dwarf stars, and neutron stars, that could be baryonic matter faking DM. These objects cannot account for the majority of DM [71, 72]. In fact, Big Bang nucleosynthesis, the theory of how light nuclei were produced in the early universe, shows that measurements of the abundances of elements today suggest that most of DM is non-baryonic [73]. Measurements of CMB anisotropies determine the density of non-baryonic matter, an important constraint on potential DM candidates [63]. For example, the only SM particles that could potentially account for DM are neutrinos, but these are excluded because they are not abundant enough to account for the DM density [69].

For all that is unknown about the particle content of the dark sector, there are several properties of DM that are known with high confidence: (1) it does not interact via the EM force, or this interaction is highly suppressed, (2) it is stable over long time scales, (3) it has a relic density consistent with cosmological observations, and (4) it is “cold”, meaning it was non-relativistic by the time galaxies were beginning to form [69]. A plethora of candidates satisfying these properties has been developed, including sterile neutrinos, like SM neutrinos but that do not interact via the weak force [74], axions, theoretical particles developed to address CP violation [75], and particles from “little Higgs” models [76, 77], just to name a few.

The most widely studied candidates, however, are weakly interacting massive particles (WIMPs), with masses in the range of 10 GeV to a few TeV, and whose self-annihilation cross section is similar in scale to the weak strength [69]. The two best motivated WIMPs are the lightest superparticle (LSP) of supersymmetric (SUSY) models [78] and the lightest Kaluza-Klein particle (LKP) of models with extra dimensions [79]. The SUSY model should obey R-parity to guarantee the stability of the LSP. The LSP should be neutral to satisfy constraints from searches for exotic isotopes and is unlikely to be an ordinary sneutrino, which would have been observed in previous WIMP searches [63]. This leaves the lightest neutralino, a mixture of the gauge boson superpartner gauginos, as the best DM candidate from SUSY models. Alternatively, models introducing extra spatial dimensions, such as those of Arkani-Hamed, Dimopoulos, and Dvali (ADD) [80] and Randall and Sundrum (RS)[81], predict a “tower” of excited states of SM particles, called Kaluza-Klein (KK) states, with increasing mass proportional to the inverse of the scale of the extra dimension, and having the same quantum numbers as their corresponding SM particles. Models where all SM fields can propagate in the extra dimensions, as opposed to models where only gravity can, are called universal extra dimensions (UED) [82]. The best motivated LKP is the first KK excitation in UED of the SM U(1) gauge boson [83].

1.2.1.3 Potential detection methods

The methods of detecting DM fall into three categories: (1) indirect searches, where the products of DM annihilations or decays are observed, (2) direct searches, where the recoil of SM particles is measured after scattering with an incident DM particle, and (3) collider searches, where the DM candidate is produced directly from interactions of SM particles. These methods complement one another since they each approach the problem in a different way and have different relative strengths and weaknesses. For example, direct searches are limited in their sensitivities at low DM masses by the inability to measure such smaller nuclear recoils, while colliders excel in this region since the production of low mass particles is uninhibited by kinematic restrictions. Conversely, collider searches are less sensitive than direct and indirect searches at high mass, being limited by the energy scale of the collisions.

Indirect DM searches are performed at experiments designed to detect the SM products of the decays or self-annihilations of DM particles. Since DM is attracted by the gravitational force, it could collect in the centers of massive bodies such as the Sun or Earth, where they would be more likely to annihilate in higher densities. The IceCube detector sets the best upper limits on the high-energy muon flux from DM annihilations within the Sun to $103 \text{ muons/km}^2/\text{yr}$ [84], while the SuperKamiokande telescope has the best upper limits for softer muons at about $1500 \text{ muons/km}^2/\text{yr}$ [85]. Searches for photons from DM annihilating in the galactic halo can produce monoenergetic photon spectra, but these signals are particularly difficult to isolate from photons of regular astrophysical origin. The FERMI/LAT collaboration has found a small signal in a region around the galactic center, with known point sources removed from the data, but the result is not strong enough to be considered a discovery [86]. Finally, DM can produce an excess in the spectra of antiparticles such as positrons. Experiments find small excesses with these signatures, but they may be explained by astrophysical sources, and they predict a DM cross section too high to be consistent with a thermal WIMP [63].

Direct DM searches measure the interaction of DM with regular matter through either elastic or inelastic collisions, in either a spin-dependent or spin-independent manner in terrestrial laboratory detectors [69]. In an elastic scattering experiment, WIMPs interact with the nuclei in the detector as a whole, and the recoil energy spectrum is measured, typically in the range 1-100 keV. In inelastic scattering experiments, the WIMP either excites or ionizes orbital electrons, or the WIMP leaves the nuclei in the detector in an excited state, yielding an energy recoil of the nucleus plus an emitted photon a short time later. These target interactions are also distinguished by whether the DM-nucleon interactions involve the spin degree of freedom of the nuclei. Spin-independent detectors benefit from an increase in the DM-nucleon interaction cross section by increasing the mass of the detector nuclei, while the mass of the detector material does not benefit the spin-independent measurements to the same degree. The best cross section lower limits for spin-independent and spin-dependent neutron interactions come from the Large Underground Xenon (LUX) detector, a time-projection chamber filled with 368 kg of scintillating liquid

xenon, surrounded by highly sensitive light detectors to search for the signature of DM scattering with a xenon atom, and shielded by a large water tank and a mile of Earth overburden [87]. For a WIMP mass of 33 GeV, the cross section lower limits from LUX are on the order of 10^{-45}cm^2 for spin-independent, and $10^{-41}(10^{-39})\text{cm}^2$ for spin-independent neutron(proton) interactions.

Finally, DM can be produced directly in particle colliders, and searches for signatures of high missing energy from DM escaping the detector opposite a tagged SM particle can be explored. Such signatures are referred to as mono- X , where X is the single SM particles observed in the detector. Note that since the DM is produced and is not from cosmic origins, collider searches are not traditional DM searches, where the target signal originates from the cosmic dark matter. Consequently, a DM-like particle produced and detected at a collider experiment, displaying some of the expected properties, may mimic cosmic DM, but may not be stable on cosmological time scales [88]. The methods of collider searches are covered in detail in the next section, including the current statuses of these searches.

When the observations of either of the three detection methods are consistent with the backgrounds only and no signal is observed, the results are cast in the form of exclusion limits, and special care must be taken to compare these limits between the different methods. Of particular interest is the comparison of DM cross section upper limits between direct and collider searches. In order to compare the DM-nucleon cross sections from direct and indirect searches to the mono- X production cross sections from collider searches, a model for how DM couples to nucleons must be specified. For comparisons to the spin-independent cross section upper limits for DM scattering off a nucleus N found by LUX, the following relation will be used:

$$\sigma_{\chi N}^{SI} = \frac{\mu_{\chi N}^2}{\pi} [Z f_p + (A - Z) f_n]^2 \quad (1.8)$$

where $\mu_{\chi N}$ = is the $\chi - N$ reduced mass, A and Z are the atomic mass numbers of N , and $f_{p/n}$ are the model-dependent couplings of DM to protons/neutrons [5]. A set of models describing the explicit coupling of DM to SM particles is detailed in the next section.

1.3 Beyond the Standard Model

1.3.1 Collider searches for DM

Previous DM searches at the Large Hadron Collider (LHC) include analyses with mono- X signatures: X produced in association with large missing transverse momentum (MET) from the DM escaping the detector, where X is a jet [89, 90], t/b quark [91, 92, 93], photon [94, 95, 96], lepton [97, 98], or W/Z boson [99, 100, 101]. The discovery of the Higgs boson has opened a new portal to searching for DM at the LHC through the mono- H signature [5, 102].

Mono- H is purely a discovery mode for DM. Due to the distinct production topologies and suppressed couplings, mono- H analyses cannot contribute to the combination of other mono- X analyses. In contrast with other mono- X signatures, in which X is emitted as initial state radiation (ISR)(see Figure 1.3a), ISR of a H is highly suppressed due to the small H -quark coupling. Therefore, the H is radiated preferentially from the new physics vertex (see Figure 1.3b), directly probing the effective DM-SM coupling. The models describing the effective vertex for the case where X comes from ISR, referred to as ISR models, couple DM to quarks either through effective field theory (EFT) operators, or explicitly with a scalar or vector mediator [6]. Since the effective vertex does not explicitly involve X , the different mono- X searches can be combined, each carrying a weight proportional to the quark- X coupling. Since the quark- H coupling is small compared to the other quark- X couplings, mono- H cannot make a strong contribution to the combination, and is therefore not included. The models that have X emitted directly from the effective vertex, referred to as discovery models, do have a well-motivated mono- H signature. These models couple DM to X directly with EFT operators or a new mediator particle, so not all mono- X analyses are combined as in the ISR case. Each mono- X signature has discovery models that motivate an enhanced DM- X coupling, so although some signatures can be combined for these models, with comparable contributions, they are usually studied independently for the different signatures. These models are called discovery models because they each allow for the detection of DM for each mono- X signature, independently from the others. Therefore, even though the signatures contribute

different amounts to the ISR model combinations, it is of critical importance to look at each signature's discovery models. This dissertation will consist of the study of the discovery models for mono- H .

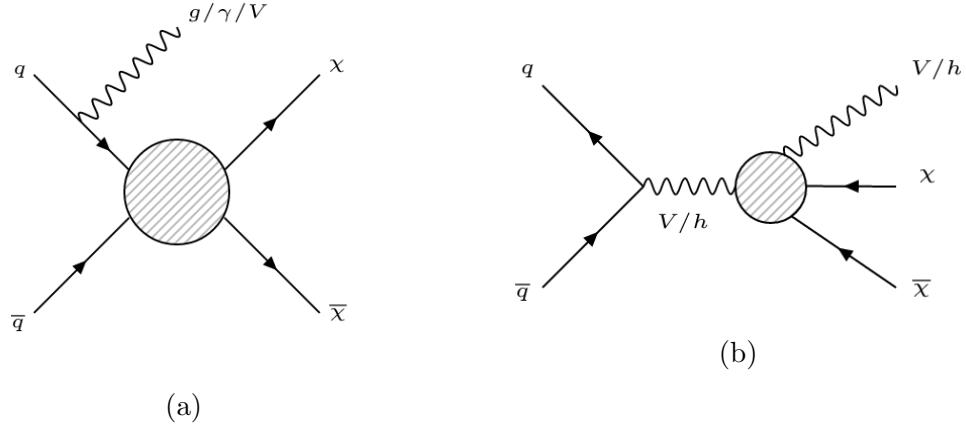


Figure 1.3: (a) Mono- X production topology for signatures with X emitted as ISR. (b) Mono- X production topology for signatures with X emitted from new physics vertex.

Mono- H searches have been done at the 8 TeV LHC for H decaying to two photons [103] and two bottom quarks [104] at ATLAS, with results consistent with SM predictions and limits set on various model parameters. The $b\bar{b}$ final state shows a higher sensitivity to limit setting for the models used in the data interpretation at 8 TeV. For 13 TeV LHC data, mono- H searches are being done at ATLAS for H to $b\bar{b}$ [105] and at CMS for the five H decay modes: ZZ , WW [106], $\gamma\gamma$ [107], $b\bar{b}$ [108], and $\tau\tau$. Because each final state will have various benefits and drawbacks, each has a dedicated analysis, exploring the sensitivity to different models in different regions of parameter space.

The $H \rightarrow ZZ$ decay mode, where the two Z bosons decay to four leptons, is studied in this dissertation. The mode where the Z bosons decay to two leptons and two neutrinos is open for investigation. Over other H decay modes, the four-lepton final state has the advantage of easily reducible backgrounds and a clean reconstruction of final state particles. The lepton combinations of four electrons, four muons, and two electrons two muons are treated individually, then combined in the final results. This channel was key to the discovery of H at 7 and 8 TeV [109, 110], and this analysis is an extension of these studies and their continuation [111, 112]. In particular, the baseline event selection used

here is chosen to match that of the other $H \rightarrow ZZ$ analysis groups through an event by event synchronization exercise [113].

Nine kinematically distinct models with DM are used in the data interpretation, including five effective field theory (EFT) models and four simplified models. The EFT models couple DM to H via an n -dimensional contact operator, with operators of dimension four, five, six, and eight [5]. They have the benefit of being independent of the details of new physics models and having a one-dimensional parameter space. The EFTs have the drawback of limited ranges of validity, being constrained by perturbativity and H and Z to invisible decay limits. The EFT parameter choices are discussed in the next section. The simplified models introduce an additional massive particle to mediate the DM-SM coupling. This mediator particle is a vector, scalar, or pseudoscalar [5, 102]. Although they are better motivated by the addition of new physics, the simplified models have the drawback of more complex parameter spaces, including parameters which affect the kinematics of the final state particles and must be scanned over. The simplified model parameter choices discussed in Section 1.3.2 are chosen to be consistent with other LHC DM searches [6].

1.3.2 Signal models

The signal models are divided into two categories: effective field theories (EFTs) and simplified models.

1.3.2.1 Effective field theory models and benchmarks

The five EFTs are summarized in Table 1.1. The models have Lagrangians with effective operators ranging from dimension four to eight with either scalar or fermionic DM [114, 115], producing mono- H signatures shown in Figure 1.4. The models have two parameters each, the DM mass and the coupling or mass cutoff scale. Although there are regions of parameter space where the kinematics are independent of the coupling or mass cutoff scale, the kinematics generally depend on the choice of both parameters. The DM mass values are the same for all models: 1, 10, 50, 65, 100, 200, 400, 800, 1000, 1300 GeV, as recommended by the LHC DM Working Group (DMWG) [6]. These mass values are chosen with a fine enough grid spacing to cover the range of variation in kinematics. The

additional value of 65 GeV, around half the Higgs' mass, is added where the cross sections begin to drop significantly and assists in producing smooth limit curves.

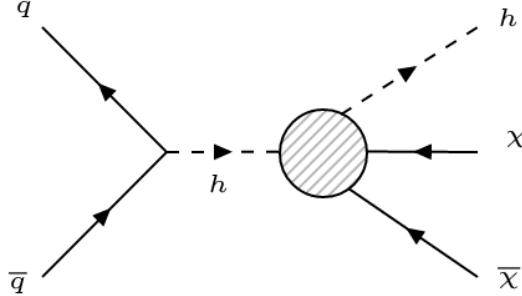


Figure 1.4: Collider production diagram for mono-Higgs effective field theories.

The value of the coupling must be set for each model individually so as to ensure the value is in a range where the kinematics are independent of the coupling. In these regions of parameter space, cross section limits can be reliably scaled to coupling limits. The additional value of $\Lambda = 1000\text{GeV}$ for EFT_xdxHDHc is included to compare with previous limits, even though it is in the region where the kinematics depend on the coupling. Existing constraints on the couplings from perturbativity [5] and invisible branching ratio limits [116, 117] are shown in Table 1.2. The constraints from invisible branching ratio limits only apply when the DM mass is less than half the mediator mass, allowing this decay to be kinematically open. The production cross sections for these benchmark choices are given in Appendix A.

Name	Operator	Param.	Dim.	S_χ
EFT_HHxx_scalar	$\lambda H ^2\chi^2$	$m_\chi, \lambda = 0.1$	4	0
EFT_HHxx_combined	$\frac{1}{\Lambda} H ^2\bar{\chi}\chi$	$m_\chi, \Lambda = 1000 \text{ GeV}$	5	1/2
EFT_HHxxg5x	$\frac{1}{\Lambda} H ^2\bar{\chi}i\gamma_5\chi$	$m_\chi, \Lambda = 100 \text{ GeV}$	5	1/2
EFT_xdxHDHc	$\frac{1}{\Lambda^2}\chi^\dagger i\partial^\mu\chi H^\dagger iD_\mu H$	$m_\chi, \Lambda = 100, 1000 \text{ GeV}$	6	0
EFT_xgxFHDH	$\frac{1}{\Lambda^4}\bar{\chi}\gamma^\mu\chi B_{\mu\nu}H^\dagger D^\nu H$	$m_\chi, \Lambda = 200 \text{ GeV}$	8	1/2

Table 1.1: Effective Field Theory Models [5].

Name	Perturbativity	BR_{inv} Limits
EFT_HHxx_scalar	$\lambda < 4\pi$	$\lambda < 0.016$ ($m_\chi < m_h/2$)
EFT_HHxx_combined	$\Lambda > v/4\pi$	$\Lambda > 10$ TeV ($m_\chi < m_h/2$)
EFT_HHxxg5x	$\Lambda > v/4\pi$	$\Lambda > 10$ TeV ($m_\chi < m_h/2$)
EFT_xdxHDHc	$g_z < 4\pi$, $\Lambda > 30$ GeV	$\Lambda > 400$ GeV ($m_\chi < m_Z/2$)

Table 1.2: Constraints on effective field theory model parameters [5].

1.3.2.2 Simplified models

There are four simplified models, each including one or more new massive particles that mediate the H -DM interaction. The models with vector mediators are motivated by the addition of new symmetries to the Standard Model, with the mediator corresponding to the gauge boson of the new symmetry. Additional models are motivated by the addition of scalar or pseudoscalar mediators as a portal into the dark sector.

1.3.2.3 Z' - Two Higgs doublet model

The Z' - Two Higgs Doublet Model (Zp2HDM) simplified model extends the gauge group of the SM to include a new symmetry, $U(1)_{Z'}$, with Z' the gauge boson [102]. This symmetry is spontaneously broken by a scalar singlet ϕ , generating a Z' mass above the electroweak (EW) symmetry breaking scale. The right-handed quarks are charged under $U(1)_{Z'}$ and all other SM particles are neutral. The Z' coupling to quarks, g_z , is constrained by EW global fits [117] and dijet resonance searches [118, 119] to be:

$$g_z < 0.03 \frac{g}{\cos \theta_w \sin^2 \beta} \frac{\sqrt{M_{Z'}^2 - M_Z^2}}{M_Z}. \quad (1.9)$$

Additionally, a second Higgs doublet is added with a Type 2 two-Higgs doublet model, introducing states Φ_u and Φ_d , which couple to up and down type quarks, respectively, as:

$$\mathcal{L} \supset -y_u Q \bar{\Phi}_u \bar{u} - y_d Q \Phi_d \bar{d} + y_e L \Phi_d \bar{e} + h.c.. \quad (1.10)$$

Φ_u is chosen to be charged under $U(1)_{Z'}$, while Φ_d is neutral. The two Higgs doublets

obtain VEVs ν_u and ν_d after EW symmetry breaking and can be parametrized as:

$$\begin{aligned}\Phi_d &= \frac{1}{\sqrt{2}} \begin{pmatrix} -\sin(\beta)H^+ \\ \nu_d - \sin(\alpha)h + \cos(\alpha)H - i \sin(\beta)A^0 \end{pmatrix} \\ \Phi_u &= \frac{1}{\sqrt{2}} \begin{pmatrix} \cos(\beta)H^+ \\ \nu_u + \cos(\alpha)h + \sin(\alpha)H + i \cos(\beta)A^0 \end{pmatrix}\end{aligned}\quad (1.11)$$

where h and H are neutral CP -even scalars, and A^0 is CP -odd. The angle α is defined as the angle that diagonalizes the $h - H$ mass mixing matrix, and the angle β is defined as $\tan(\beta) = \nu_u/\nu_d$. The h is assumed to be the SM Higgs boson with $m_h = 125$ GeV, while the other scalars have masses > 300 GeV. Due to perturbativity and previous constraints [120], α and β are chosen such that $\tan(\beta) > 0.3$ and $\alpha = \beta - \pi/2$.

Mono- H signals arise when the pseudoscalar A^0 has a large branching ratio to DM, as shown in Figure 1.5. The new particles and parameters of the Zp2HDM model are summarized in Table 1.3. The values of the parameters chosen for various benchmark scenarios are given Section 1.3.2.7.

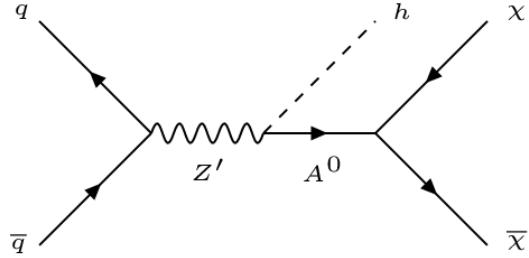


Figure 1.5: Collider production diagram for Zp2HDM.

1.3.2.4 Baryonic Z' model

The Baryonic Z' (ZpBaryonic) simplified model extends the gauge group of the SM to include a new symmetry, $U(1)_B$ for the baryon number B , with the Z' being the gauge boson of $U(1)_B$ [121, 122, 123]. Z' couples to quarks and fermionic DM as:

$$\mathcal{L} \supset g_q \bar{q} \gamma^\mu q Z'_\mu + g_\chi \bar{\chi} \gamma^\mu \chi Z'_\mu \quad (1.12)$$

Particle	Description
χ	Fermionic DM particle
Z'	$U(1)_{Z'}$ gauge boson
ϕ	Z' sector scalar
Φ_u, Φ_d	Two Higgs doublets
h, H	Neutral CP-even scalars
H^\pm	Charged heavy Higgs
A^0	Neutral CP-odd pseudoscalar
Param.	Description
m_χ	DM mass
$m_{Z'}$	Z' mass
g_z	Z' -quark coupling
$y_{u/d/e}$	Φ -up-quark/down-quark/lepton coupling
$\nu_{u/d}$	$\Phi_{u/d}$ VEV
α	$h-H$ mixing angle
β	$\Phi_{u/d}$ VEV angle

Table 1.3: New particles and parameters of the Zp2HDM simplified model [5]

To derive a mono-Higgs signature, $U(1)_B$ is spontaneously broken by a “baryonic Higgs” scalar h_B , which mixes with the SM Higgs via a mixing angle θ . This mixing induces an $h - Z'$ interaction $-g_{hZ'Z'}hZ'_\mu Z'^\mu$, with coupling:

$$g_{hZ'Z'} = \frac{m_{Z'}^2 \sin(\theta)}{\nu_B} \quad (1.13)$$

where $m_{Z'}$ is the mass of the Z' and ν_B is the VEV of h_B . At energies less than $m_{Z'}$, these operators combine to yield an effective Lagrangian:

$$\mathcal{L}_{eff} = -\frac{g_q g_\chi}{m_{Z'}^2} \bar{q} \gamma^\mu q \bar{\chi} \gamma_\mu \chi (1 + \frac{g_{hZ'Z'}}{m_{Z'}^2} h) \quad (1.14)$$

The first term gives rise to mono-jet and mono-EW boson signals, while the second yields

the mono-Higgs signal shown in Figure 1.6.

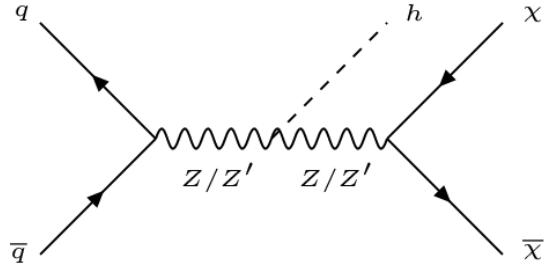


Figure 1.6: Collider production diagram for Z' models.

The new particles and parameters of the ZpBaryonic model are summarized in Table 1.4. Perturbativity arguments require the Z' -quark coupling to be less than 4π [5]. The values of the parameters chosen for various benchmark scenarios are given in Section 1.3.2.7.

Particle	Description
χ	Fermionic DM particle
Z'	$U(1)_B$ gauge boson
h_B	Baryonic Higgs
Param.	Description
m_χ	DM mass
$m_{Z'}$	Z' mass
g_q	Z' -quark coupling
g_χ	Z' -DM coupling
$g_{hZ'Z'}$	Z' -h coupling
θ	h - h_B mixing angle
ν_B	h_B VEV

Table 1.4: New particles and parameters of the ZpBaryonic simplified model [5]

1.3.2.5 Hidden sector Z' model

The Hidden Sector Z' (ZpHS) simplified model mixes the SM and a hidden sector with a $U(1)'$ symmetry [124, 125, 126, 127, 128]. SM particles are neutral under $U(1)'$, while DM is charged. The relevant terms in the Lagrangian are:

$$\mathcal{L}_{eff} \supset \frac{g_2}{2c_W} J_{NC}^\mu Z_\mu + g_\chi \bar{\chi} \gamma^\mu \chi Z'_\mu \quad (1.15)$$

where J_{NC} is the SM neutral current. A mono-Higgs signature arises from a $Z - Z'$ mass mixing, which is diagonalized by the rotation:

$$Z \rightarrow \cos(\theta)Z - \sin(\theta)Z', Z' \rightarrow \cos(\theta)Z' + \sin(\theta)Z \quad (1.16)$$

This mixing yields the mono-Higgs signatures shown in Figure 1.6 through the $h - Z - Z'$ interaction:

$$\mathcal{L}_{eff} \supset \frac{m_Z^2 \sin(\theta)}{\nu} h Z'_\mu Z^\mu \quad (1.17)$$

The new particles and parameters of the ZpHS model are summarized in Table 1.5. In order to be consistent with the invisible Z width of $\Lambda(Z \rightarrow \chi\bar{\chi}) < 3$ MeV, the value of θ is constrained by $\sin \theta < 0.03$ for $m_\chi < m_z/2$ [117]. The values of the parameters chosen for various benchmark scenarios are given in Section 1.3.2.7.

Particle	Description
χ	Fermionic DM particle
Z'	$U(1)'$ gauge boson
Param.	Description
m_χ	DM mass
$m_{Z'}$	Z' mass
g_χ	Z' -DM coupling
θ	$Z-Z'$ mixing angle

Table 1.5: New particles and parameters of the ZpHS simplified model [5]

1.3.2.6 Scalar mediator model

The Scalar Mediator (Scalar) simplified model introduces a real scalar singlet S as a portal into the dark sector [129]. The quark and DM Yukawa coupling terms are:

$$\mathcal{L} \supset -y_\chi \bar{\chi} \chi S - \frac{m_q}{\nu} \bar{q} q h \quad (1.18)$$

and the relevant terms of the scalar potential are:

$$V \supset a|H|^2 + b|H|^2S^2 + \lambda_h|H|^4 \rightarrow \frac{1}{2}a(h+\nu)^2S + \frac{1}{2}b(h+\nu)^2S^2 + \frac{\lambda_h}{4}(h+\nu)^4. \quad (1.19)$$

The expression after the arrow follows when the Higgs acquires a VEV, as $H \rightarrow \frac{1}{\sqrt{2}}(h+\nu)$. After expanding the expression, there is an $h - S$ mixing term $a\nu h S$. The system can be diagonalized with the rotation:

$$h \rightarrow \cos(\theta)h + \sin(\theta)S, S \rightarrow \cos(\theta)S - \sin(\theta)h \quad (1.20)$$

where θ is defined by $\sin(2\theta) = 2a\nu/(m_S^2 - m_h^2)$. Following this rotation, the Yukawa terms become:

$$\mathcal{L} \supset -y_\chi \bar{\chi} \chi (\cos(\theta)S - \sin(\theta)h) - \frac{m_q}{\nu} \bar{q} q (\cos(\theta)h + \sin(\theta)S) \quad (1.21)$$

and the relevant terms in the scalar potential, at first order in $\sin(\theta)$, are:

$$V \supset \frac{\sin(\theta)}{\nu} (2m_h^2 + m_S^2)h^2S + b\nu h S^2. \quad (1.22)$$

These terms give rise to the mono-Higgs interactions shown in Figure 1.7, as well as additional loop diagrams for gluon fusion production. The new particles and parameters of the Scalar model are summarized in Table 1.6. Perturbativity arguments require $\sin \theta < 4\pi$ [5], while 8 TeV H data is consistent with $\cos \theta = 1$, requiring $\sin \theta < 0.4$ [130, 131, 132, 133]. The values of the parameters chosen for various benchmark scenarios are given in Section 1.3.2.7.

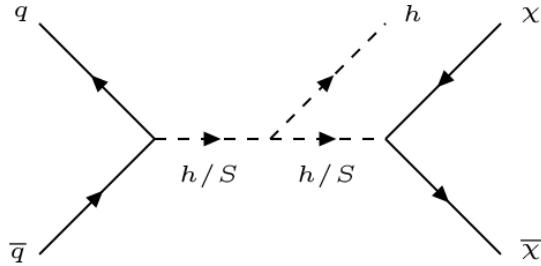


Figure 1.7: Collider production diagram for Scalar model.

Particle	Description
χ	Fermionic DM particle
S	Scalar particle
Param.	Description
m_χ	DM mass
m_S	S mass
y_χ	S-DM coupling
a, b	h - S scalar couplings
θ	h - S mixing angle

Table 1.6: New particles and parameters of the Scalar simplified model [5]

1.3.2.7 Benchmarks

The simplified models are summarized in Table 1.7. Although these models have better physical motivations, the parameter spaces are vastly more complex than the EFTs, making it difficult to understand the cross section scaling rules and kinematic dependence on parameters across the entire parameter spaces. The parameter values are chosen to match the recommendations of the DMWG [6]. For the ZpHS model, which is not included in the DMWG report, the coupling parameters are chosen to match the benchmarks of Carpenter et al. [5] while the DM and mediator mass scan matches the DMWG recommendations for a vector mediator, shown in Table 1.8. The DM and mediator mass scans for the Zp2HDM and Scalar models are shown in Table 1.10 and Table 1.11, respec-

tively. Existing constraints on the couplings from theory and invisible branching ratio limits given in are shown in Table 1.9 [5]. The production cross sections for the simplified models are given in Appendix A.

Name	Mediator(S_{Mediator})	Fixed Param.	Scanned Param.
Zp2HDM	$Z'(1/2)$ $A^0(0)$	$m_\chi = 100 \text{ GeV}$ $g_Z = 0.8$ $\tan \beta = 1$	$(m_{Z'}, m_{A^0}) = \text{Table 1.10}$
ZpBaryonic	$Z'(1/2)$	$g_\chi = 1$ $g_q = 1/3$ $g_{hZ'Z'} = m_{Z'}$	$(m_{Z'}, m_\chi) = \text{Table 1.8}$
ZpHS	$Z'(1/2)$	$g_\chi = 1$ $\sin(\theta) = 0.1$	$(m_{Z'}, m_\chi) = \text{Table 1.8}$
Scalar	$S(0)$	$g_\chi = 1$ $\sin(\theta) = 0.3$ $b = 3$	$(m_S, m_\chi) = \text{Table 1.11}$

Table 1.7: Benchmark parameter choices for the simplified models [5, 6].

m_χ [GeV]	$m_{Z'}$ [GeV]									
	10	20	50	100	200	300	500	1000	2000	10000
1	10	15	50	100						10000
10	10									10000
50	10		50	95	200	300				10000
150	10				200	295	500	1000		10000
500	10						500	995	2000	10000
1000	10							1000	1995	10000

Table 1.8: Benchmark mass points for models with a vector mediator [6].

Name	Perturbativity	BR_{inv} Limits
ZpBaryonic	$g_q < 4\pi$	N/A
ZpHS	N/A	$\sin(\theta) < 0.03$ ($m_\chi < m_Z/2$)
Scalar	$\sin 2\theta < 4\pi$	$\sin \theta < 0.4$

Table 1.9: Theoretical and experimental constraints on simplified model parameters [5, 6].

m_{A^0} [GeV]	$m_{Z'}$ [GeV]							
	600	800	1000	1200	1400	1700	2000	2500
300	600	800	1000	1200	1400	1700	2000	2500
400	600	800	1000	1200	1400	1700	2000	2500
500		800	1000	1200	1400	1700	2000	2500
600		800	1000	1200	1400	1700	2000	2500
700			1000	1200	1400	1700	2000	2500
800			1000	1200	1400	1700	2000	2500

Table 1.10: Benchmark mass points for the Zp2HDM simplified model [6].

m_χ [GeV]	$m_{Z'}$ [GeV]							
	10	20	50	100	200	300	500	1000
1	10	20	50	100	200	300	500	10000
10	10	15	50	100				10000
50	10		50	95	200	300		10000
150	10				200	295	500	1000
500	10						500	995
1000	10							1000

Table 1.11: Benchmark mass points for models with a scalar mediator [6].

Chapter 2

Experimental apparatus

This chapter gives an overview of the experimental apparatus, the Large Hadron Collider (LHC) and Compact Muon Solenoid (CMS) detector, used to collect the data analyzed in this dissertation. The first section reviews the design and performance of the LHC. The second section reviews the design of CMS, its component subdetectors, and data acquisition system.

2.1 Large Hadron Collider

This section reviews the construction and original design specifications of the LHC [134], leading up to the 7 – 8 TeV center of mass (COM) energy collisions recorded from March 2010 to February 2013 (Run 1), the upgrades and repairs made to the LHC and its pre-accelerators during Long Shutdown 1 (LS1) from February 2013 to April 2015, and finally, the 13 TeV COM energy collisions recorded from April 2015 through 2016.

Due to budgetary and logistical concerns, the LHC is located in the repurposed Large Electron-Positron (LEP) collider tunnel, constructed in the 1980s by the European Organization for Nuclear Research (CERN). CERN continues to operate the LHC accelerator facilities, and its laboratory complex hosts the staff, scientists, and engineers who run the machinery and detectors associated with it. Located beneath the border of Switzerland and France near Geneva, the LEP tunnel consists of eight straight sections and eight arched sections, totaling 26.7 km, at depths varying from 45 m to 170 m beneath the surface. Two 2.5 km transfer tunnels connect the main LEP tunnel to the rest of

the CERN complex. A series of pre-accelerators increase the energy of ionized hydrogen gas protons to 450 GeV before they are injected into the LHC. The underground caverns at Points 2 and 8, which were built for LEP, were repurposed for the ALICE and LHCb experiments, which are the two special-purpose LHC experiments, designed to study quark-gluon plasma in heavy ion collisions and the matter-antimatter imbalance, respectively. The facilities at Points 1 and 5 were built new for the general-purpose CMS and ATLAS experiments.

Although the length of the LEP tunnel is sufficient for the LHC, the diameter of the tunnel and the geometry of the straight and arched sections are suboptimal for a proton-proton accelerator. Since synchrotron radiation emission is not as much of a problem for protons, the LHC would ideally have longer arched sections. The two counter circulating particle-antiparticle beams of LEP could occupy the same pipe, being curved by the same magnets, but with an inside diameter of only 3.7 m, the tunnel is too narrow to accommodate the two pipes needed for counter circulating proton-proton beams, necessitating the use of the “two-in-one” super-conducting twin bore magnet design. The LHC beam is steered by 1,232 8 T, superconducting dipole twin bore magnets, which are cooled by a system of NbTi Rutherford cables to a temperature below 2 K. This technology is essential to the LHC operation, but comes at the cost of a higher sensitivity to instabilities in the operation temperature, which may cause the magnet to quench, or lose its superconductivity and current.

The LHC was designed to explore physics at the EW symmetry breaking scale, with a nominal COM energy for collisions of 14 TeV, and to search for rare events produced by physics beyond the SM, with a target luminosity of $10^{34} \text{cm}^{-2}\text{s}^{-1}$. This energy and luminosity are both the highest ever produced. For a general physics process, the rate of event production is given by

$$N = \sigma \times L \propto \sigma \times n_b N_b^2 f_{rev} \gamma \quad (2.1)$$

where σ is the process cross section, L is the LHC luminosity, which is proportional to n_B , the number of bunches per beam, N_b^2 , the number of particles per bunch, f_{rev} , the beam revolution frequency, and γ , the relativistic gamma factor. Consequently, to achieve higher

event rates for rare processes, both high beam intensities and high beam energies are required. To search for rare events, such as H production, the basic strategy for designing the LHC was to maximize these luminosity parameters within the budgetary, engineering, and physical limitations, of which there are many. Combining these constraints yields nominal values of 2808 bunches per beam, 1.2×10^{11} protons per bunch, and a revolution frequency of 11245 turns per second. The luminosity decays over a given run with a lifetime of $\tau \approx 15$ hours, due primarily to losses in particle intensity from collisions, and must periodically be dumped and refilled with an average turnaround time of around 7 hours. The integrated luminosity is the integral of the luminosity as a function of time $L(t) = L_0/(1 + t/\tau)^2$ over a run of length T_{run} given by

$$L_{int} = L_0\tau(1 - e^{-T_{run}/\tau}) \quad (2.2)$$

where L_0 is the initial luminosity. If the LHC runs for 200 days per year with a peak luminosity of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$, the maximum total integrated luminosity, or sum of the integrated luminosity of all runs is about 80 fb^{-1} per year. Due to unforeseen setbacks and inefficiencies in collecting data at the detectors, the total integrated luminosity collected by the experiments is far less than the maximum, totaling around 20 fb^{-1} each from ATLAS and CMS in the entire Run 1, and about 2 fb^{-1} each in 2015.

The LHC machine was designed to attain a per beam energy of 7 TeV, resulting in COM collisions of 14 TeV, but an accident during beam energy ramp-up in September 2008, caused by a faulty electrical connection between two magnets damaging numerous magnets, resulted in delays [135]. As a result, the Run 1 beam energy was set to 3.5 TeV and later increased to 4 TeV, for 7 and 8 TeV collisions. LS1 began at the conclusion of Run 1, and consisted of a two-year period of maintenance and upgrades, including consolidating and repairing interconnections between about 500 magnet cryostats, adding shielding and relocating various electronic equipment, and upgrading to the LHC's ramp-up accelerators [136]. It was decided that Run 2 would proceed with beam energies of 6.5 TeV instead of the originally planned 7 TeV in the interest of time, since it would have taken longer to retrain the magnets to not quench below currents required for 14 TeV than it would to retrain them for 13 TeV [137]. Overall, the LHC has performed and

continues to perform at a very high level, supplying the experiments with beam collisions within the desired luminosity ranges.

2.2 Compact Muon Solenoid

This section reviews the design and performance of the CMS detector [2], including its general layout, subdetector systems, and trigger and data acquisition (DAQ) systems. CMS was designed to explore physics at the TeV scale, recording collisions from the LHC proton beams at their crossing place at Point 5, near Cessy, France. The detector is multi-purpose, in that it is sensitive to detecting a wide array of new physics signatures, but its primary purpose was to validate or refute the Higgs mechanism as being responsible for EW symmetry breaking. Since this goal was accomplished in Run 1, Run 2 looks forward to searching for physics beyond the SM, including signatures from new symmetries such as SUSY, extra dimensions, and DM. Additionally, CMS is designed to record collisions of heavy ion beams to study QCD at this energy scale. CMS is distinguished from other general-purpose detectors by its high magnetic field solenoidal structure, silicon-based inner tracker, and crystal scintillator EM calorimeter.

The primary challenges in designing CMS include: (1) accounting for the pileup of inelastic collisions on each event with both sufficiently high granularity detectors and small timing resolution, (2) ensuring all electronics and detector components can withstand the high radiation exposure, and (3) triggering on the roughly 10^9 events per second to filter out interesting events to a rate manageable by the readout and computing systems. The design requirements can be summarized as follows: (1) good muon identification and charge determination, (2) good charged-particle momentum resolution in the inner tracker, (3) good EM energy resolution, (4) good diphoton, dimuon, dijet, and dielectron mass resolutions, (5) efficient photon and lepton isolation, and (6) good missing energy measurement. All of these requirements will be addressed in the remainder of this chapter.

The cylindrical shape of CMS, with an overall length of 21.6 m and an outer diameter of 14.6 m, is divided into two regions, the barrel and end caps, with the coordinate system centered at the collision point near the center of the cylinder. The standard coordinate

definitions have the x -axis pointing inward toward the center of the LHC, the y -axis pointing upward, and the z -axis in the beam direction in a right-handed manner. The polar coordinates r and ϕ are measured in the $x - y$ plane, transverse to the beam, where the transverse momentum quantity p_T is defined. The missing energy E_T^{miss} (MET) is defined as the imbalance in measured p_T . The polar angle θ is measured from the z -axis. A convenient coordinate for relativistic measurements is the pseudorapidity, defined as $\eta = -\ln \tan(\theta/2)$.

The dominant feature of CMS is the superconducting solenoid, 13 m long and 6 m in diameter, supplying a field of 4 T required to bend charged particles at the energies produced in up to 14 TeV collisions for the momentum and charge measurements. Within and surrounding the solenoid is a series of layered detectors and support structure, a cutout of which is shown in Figure 2.1. At the center of CMS, surrounding the beam interaction point, is the inner tracker, a combination of ten layers of silicon microstrip detectors and three layers of silicon pixel detectors, which provide the required granularity for high occupancy collisions. The next layer, still within the solenoid bore, contains the calorimeters, first the electromagnetic calorimeter (ECAL), surrounded by the hadronic calorimeter (HCAL). The ECAL uses avalanche photodiodes in the barrel and vacuum photodiodes in the end caps to read out scintillation light produced by charged particle interactions in the lead tungstate crystals. The HCAL in the barrel uses hybrid photodetectors to read scintillation light from hadronic interactions with the brass and scintillator detector material. The scintillation light is carried to the photodetectors with clear fibers, from wavelength shifting fibers embedded in the scintillator material. The various end cap HCAL systems ensure full coverage for measuring the missing energy. Finally, muon detecting stations are incorporated into and surround the solenoid support structure where the return field is present, including aluminum drift tubes (DTs) in the barrel and cathode strip chambers (CSCs) in the end caps. These subdetector systems of CMS are covered in greater detail in the remainder of this chapter.

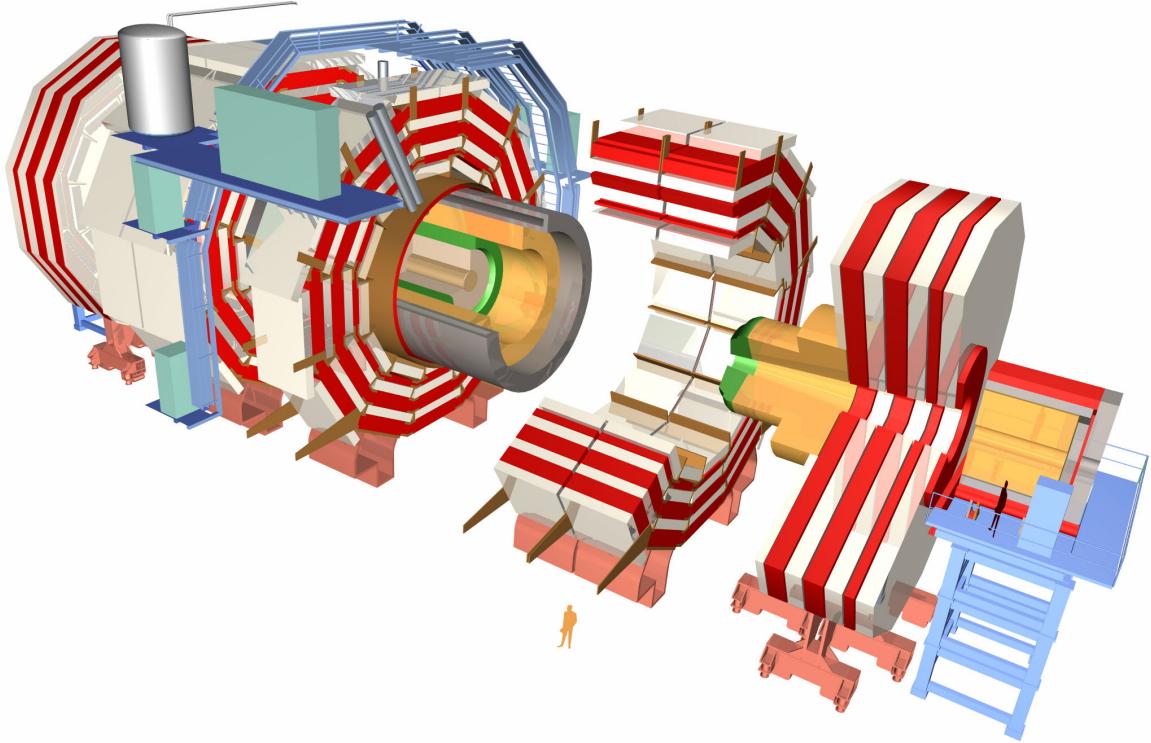


Figure 2.1: Deconstructed view of the CMS subdetectors, with human figure for scale. From inside to out, the colored segments correspond to the following systems: light brown is the pixel tracker, cream is the strip tracker, green is the EM calorimeter, orange is the hadronic calorimeter, grey is the solenoid, red is the yoke with white muon chambers [2].

2.2.1 Tracking detectors

The inner tracking detectors of CMS is supported by a 5.30 m long tube with an inner diameter of 2.38 m suspended from the HCAL barrel. The trackers contain 1,440 pixel and 15,148 strip detector modules, composing the pixel detector and silicon strip tracker, respectively. The detectors are responsible for measuring the trajectories of charged particles, essential to measuring the momenta of particles with energy > 1 GeV in the range $|\eta| < 2.5$, and to reconstructing secondary vertices and impact parameters, needed to identify heavy flavor particles. Being closest to the beam interaction point (IP), the tracking detectors are subjected to the highest radiation doses, and their material may interfere with the trajectories of primary particles through multiple scattering, bremsstrahlung, photon conversion, or nuclear interactions, necessitating the use of silicon technology. Additionally, due to the high particle flux of around 1,000 particles per 25 ns bunch cross-

ing, the detectors must have both high granularity to resolve the trajectories of particles reliably and fast readout times to reduce occupancy from high flux and pileup conditions.

The detector modules of the tracking detector are shown schematically in Figure 2.2. The innermost section, labeled PIXEL, is the pixel detector, composed of 66 million $100\text{ }\mu\text{m} \times 150\text{ }\mu\text{m}$ pixels on modules layered in three barrels at radii 4.4, 7.3, and 10.2 cm and two disks on each end at $z = \pm 34.5, \pm 46.5$ cm. A deconstructed barrel pixel module is shown in Figure 2.3, with the sensor bump-bonded onto readout chips (ROCs) controlled and powered by high-density interconnect (HDI) boards. When a charged particle passes through a pixel sensor, consisting of n -type pixels implanted on a high-resistance n -type substrate, charge carriers are induced in the conduction band of the substrate. These charge carriers then drift in the 4 T magnetic field to the nearby pixels, where an analog signal is read out, amplified, and digitized by the ROC. This drift is called charge sharing. The end cap pixel modules have a similar construction but with different pixel sensor geometries, called plaquettes. The pixel detector has a resolution of $10 - 40\text{ }\mu\text{m}$, sufficient for the imposed design requirements.

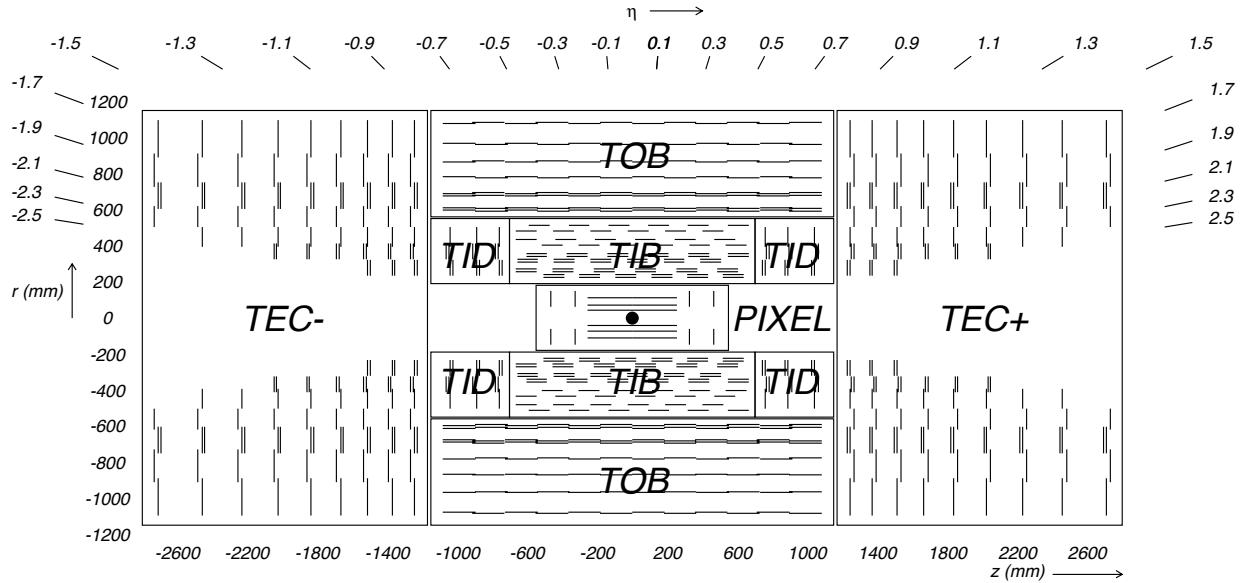


Figure 2.2: Schematic diagram of tracking detectors with radial distance of modules, shown as black lines, from center on the left axis, z -dimension on the bottom axis, and η across the top [2].

The remaining modules of the tracking detector form the strip tracker, which fills the

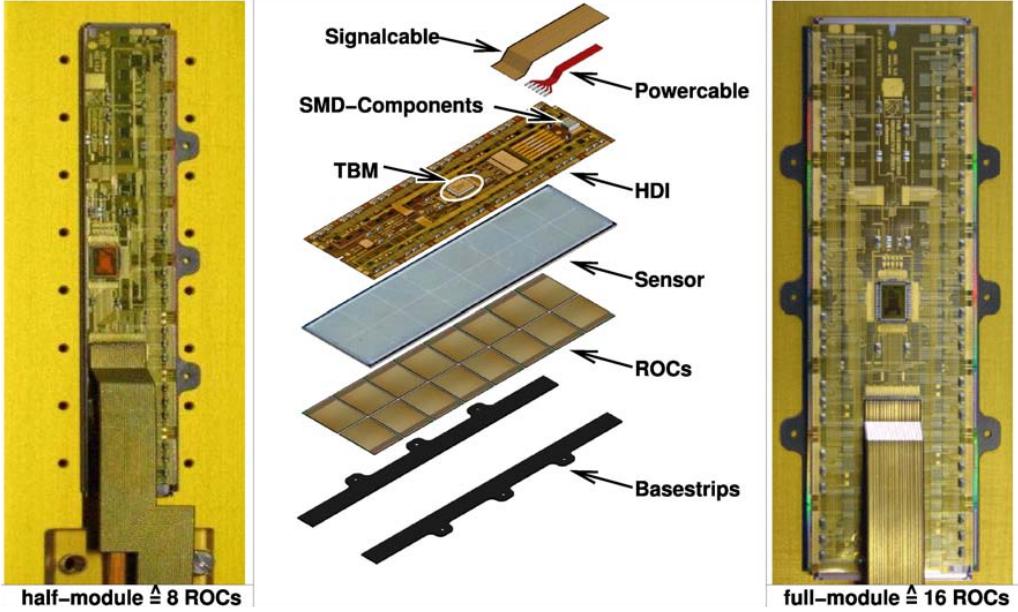


Figure 2.3: Deconstructed barrel pixel module showing module components [2].

volume between $20 - 116$ cm radially, and 118 cm in z . These 15,148 modules are divided into the following sections: tracker inner barrel (TIB), with four layers, tracker outer barrel (TOB) with six layers, tracker inner disks (TID) with six layers, and tracker end caps (TEC) with nine layers. The average particle occupancy at distances greater than 20 cm from the IP is low enough compared to regions closer to the beam line that the strip tracker is not required to have the same granularity as the pixel detector, thus, twenty nine different strip module designs, of different sizes and orientations, are used. The physical principles behind the strip tracker are the same as the pixel detector: charged particles liberate conduction band electrons, which drift toward readout sensors. To enhance the effect of the charge carriers' drift in the magnetic field, the strip detectors are tilted, yielding a resolution of approximately $30 \mu\text{m}$. Excluding defective modules, the detection efficiency of the strip tracker is nearly 100% (Figure 2.4) [138].

2.2.2 Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) consists of 61,200 lead tungstate crystal modules in the ECAL barrel (EB) region, covering $|\eta| < 1.479$, and 7,324 modules in each ECAL end cap (EE), covering $1.479 < |\eta| < 3$. The crystals are oriented radially, as shown

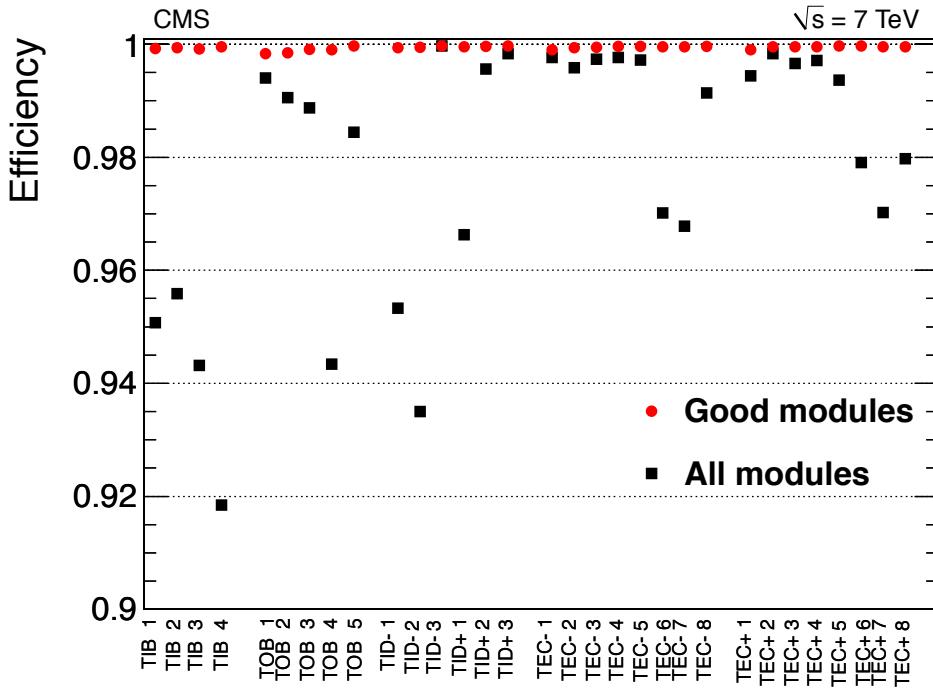


Figure 2.4: Average hit efficiencies of the strip tracker layers [2].

in Figure 2.5, at angles around three degrees in the EB and two to eight degrees in the EE from the vector to the IP to avoid cracks where particles could escape. The various supercrystal geometries are combined to form supermodules in the EB and two D-electrodes (Dees) on each end cap. The front faces of the EB crystals are at $r = 1.29$ m, and present a cross section around $22 \text{ mm} \times 22 \text{ mm}$. The EB extends to an outer radius of 1.77 m. The end cap envelopes are at ± 315.4 cm relative to the IP in z . The crystals themselves have a truncated pyramidal shape. Except for one face on the EB crystals that is depolished to account for nonuniformity in light production from the crystal shape, the crystals are polished on all sides to increase internal reflection.

The ECAL is responsible for recovering the energy of electrons and photons from the showers of scintillation light produced in the crystals. The accuracy of this measurement is of particular importance in the design of CMS, since the Higgs decay to photons and leptons are key channels in the Higgs search, one of the primary purposes of CMS. In front of each set of endcap crystals, the ECAL contains preshower detectors, consisting of a thin layer of lead followed by a thin layer of silicon strip sensors to create and detect showers

from minimum ionizing particles. The primary purpose of the preshower detectors is to identify and veto neutral pion production, in addition to improving the overall position resolution of the ECAL.

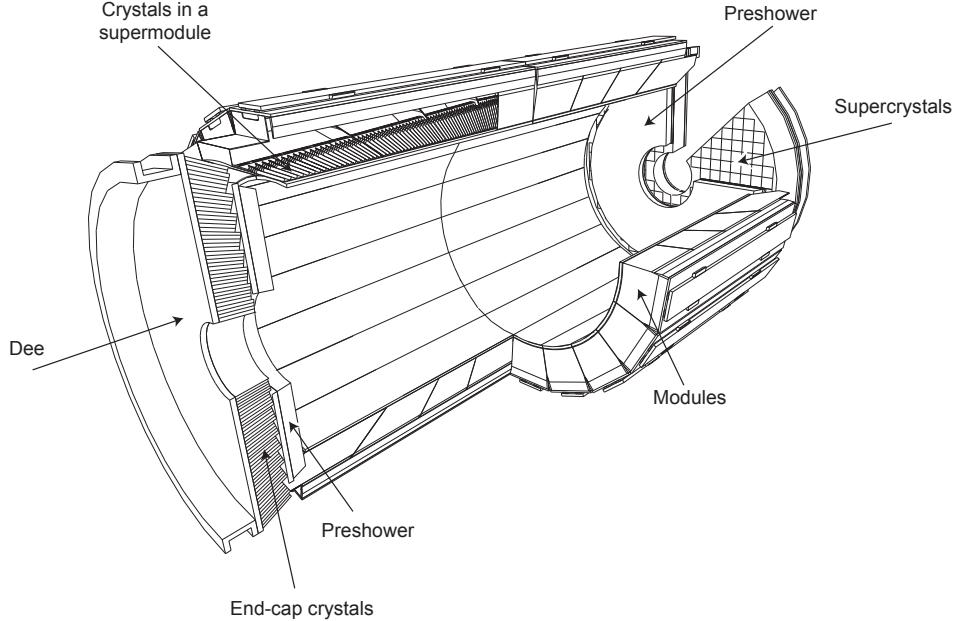


Figure 2.5: Schematic layout of the ECAL crystal modules [2].

Lead tungstate crystals were chosen for the construction of the ECAL because of their radiation hardness. The following properties enable a compact detector with sufficiently high granularity: a high density of 8.28 g/cm^3 , a short radiation length of 0.89 cm , and a small Moliere radius of 2.2 cm . Additionally, the crystals have a fast scintillation decay time of about 25 ns , the same time between bunch crossings, enabling fast response and read out. At operating temperature, about 4.5 photoelectrons are collected in the attached photodetectors per MeV of incident particle energy from the blue-green scintillation light produced in the crystals. EB crystals are glued to avalanche photodiodes (APDs) while EE crystals are read by vacuum phototriodes (VPTs), each specially designed for the CMS ECAL, as shown in Figure 2.6.

The energy resolution of the ECAL depends on the energy of the incident particle, and can be parametrized as the sum of three terms:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \quad (2.3)$$

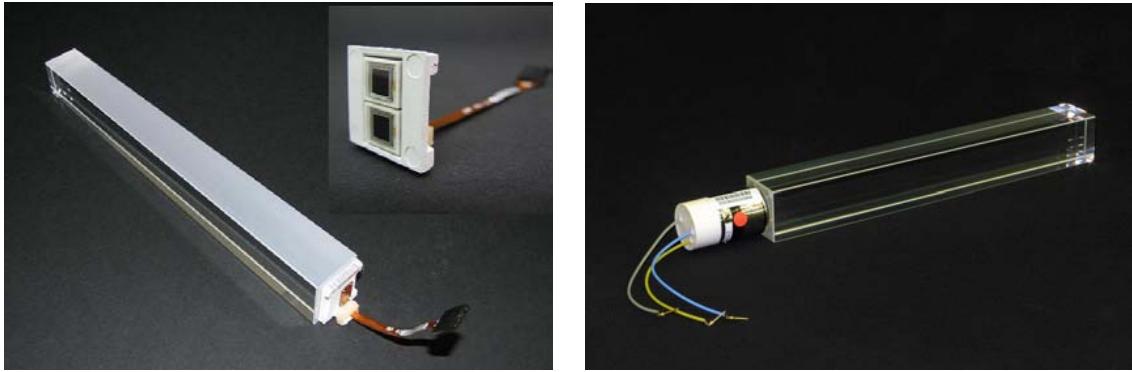


Figure 2.6: ECAL barrel crystal with one depolished face attached to avalanche photodiode photodetector (left) and ECAL end cap crystal attached to vacuum phototriode (right) [2].

The first term, called the stochastic term, arises from fluctuations in the lateral shower containment, photostatistics, and fluctuations in energy deposited in the preshower detector. The second term, called the noise term, arises from electronic, digitization, and pileup noise. The third term, called the constant term, arises from nonuniformity in light collection, calibration errors, and leakage from the back of the crystals. Test beam experiments using electron beams with momenta $20 - 250$ GeV found approximate values for the parameters: $S = 0.028$, $N = 0.12$, and $C = 0.003$.

2.2.3 Hadronic calorimeter

The hadronic calorimeter (HCAL) lies primarily within the bore of the CMS solenoid, surrounding the ECAL, between 1.77 m and 2.95 m from the beam line, covering up to $|\eta| < 5.2$. The HCAL is divided into four subsystems, shown schematically in Figure 2.7: the HCAL barrel (HB) region, covering $|\eta| < 1.3$, the HCAL end caps (HE), covering $1.3 < |\eta| < 3$, the HVAL outer (HO) calorimeter, or tail catcher, double covering the EB and HB regions to ensure complete shower absorption, outside the solenoid, and the HCAL forward (HF) calorimeter, covering up to $|\eta| < 5.2$ at 11.2 m from the IP. The primary function of the HCAL is to measure the energy and direction of hadronic jets, showers of particles produced from particles composed of quarks and gluons interacting with the detector material. Another important function of the HCAL is to contribute to the measurement of MET, which is a key variable in this analysis.

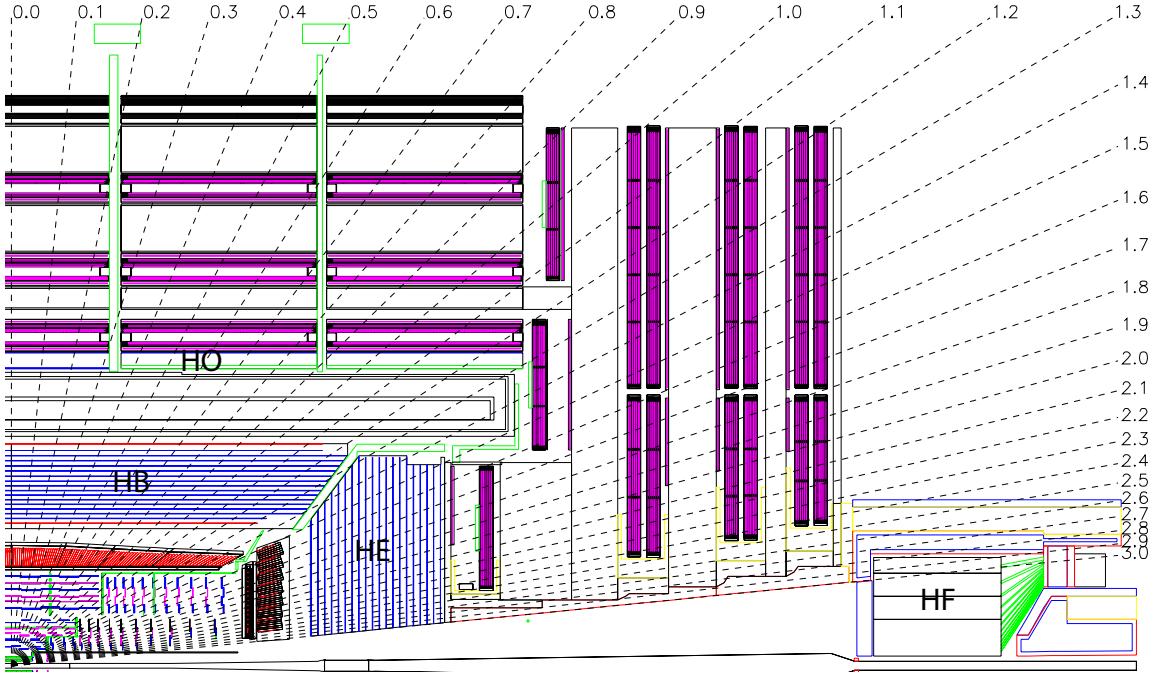


Figure 2.7: Cross sectional view of one quadrant of CMS. The labeled sections are the subsystems of the hadronic calorimeter [2].

The HCAL subsystems consist of layers of absorber material and scintillator material. The absorber material causes incident hadrons to shower into quarks and gluons, whose energy is deposited and read out from the scintillator layers. Tiles of scintillator material are organized into units called trays. The HB consists of two half-barrels, each with eighteen identical ϕ -wedges, each consisting of seventeen layers of scintillator material, with each layer containing 108 trays. The scintillator layers are separated by eight 50.5 mm thick and six 56.5 mm thick brass plates and surrounded by an inner 40 mm thick and outer 75 mm thick steel plate. The HE disks consist of 36 identical ϕ -wedges, with a total of 1,368 trays of 20,916 trapezoidal scintillator tiles divided into 18 layers, alternating with 79 mm thick brass plates. The HO layers are each divided into 12 ϕ -sectors, consist of 40 mm thick detector layers of scintillator and aluminum supports, layered into the 75 mm thick steel beams of the return yoke. The HB, HE, and HO tiles are arranged to attain a granularity of $(0.087, 0.087)$ in (η, ϕ) . The scintillation light produced in the HB, HE, and HO tiles is collected by wavelength-shifting (WLS) fibers, grouped by tray in clear fibers leading to optical decoders which arrange the clear fibers into readout

towers, transmitted to hybrid photodiodes (HPD) for amplification and readout. The energy resolution of HB+HE versus HB+HE+HO systems for test beam pions is shown in Figure 2.8, with a clear improvement when including the HO.

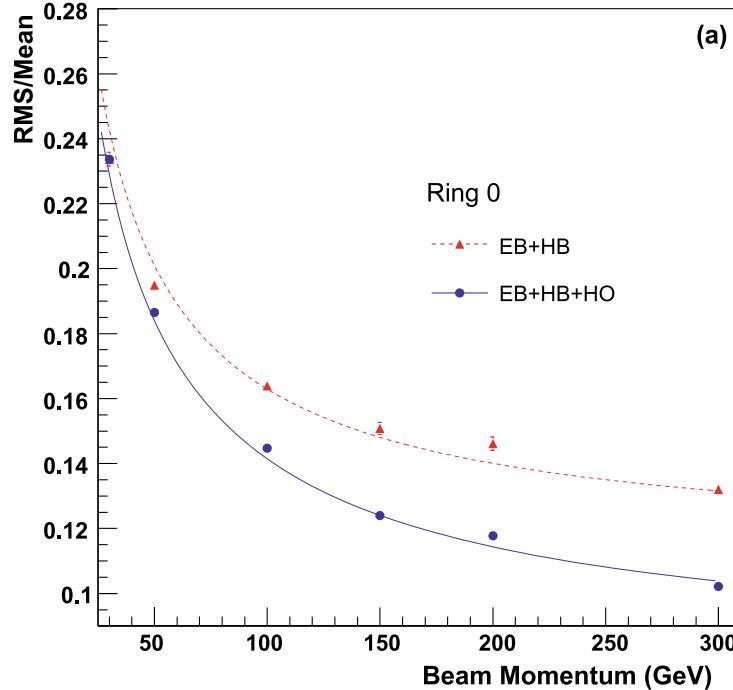


Figure 2.8: Energy resolution of HCAL systems for test beam pions [2].

The HF barrels consist of eighteen identical ϕ -wedges, with layers of 5 mm thick steel plates. Due to higher radiation doses, quartz fibers are used instead of plastic as in the other HCAL subsystems to produce scintillation light. The quartz fibers lie in grooves cut in the steel plates, with half running the full length of the absorber and half starting 22 cm, from the front of the HF barrel. This can be used to classify showers as EM or hadronic, as EM showers deposit most of their energy in the first 22 cm while hadronic showers deposit roughly the same energy throughout the material. Fibers are bundled into towers with a granularity of (0.175, 0.175). Additionally, the HF is used to measure the luminosity of the LHC beam.

2.2.4 Muon detectors

As muons are able to pass through the inner detector material with little radiative losses, the CMS muon system is the outermost subdetector system, consisting of end caps and

a barrel region divided into four layers called stations. The barrel region is segmented into regions based on the five 2.536 m yoke rings at $z = 0, \pm 5.342, \pm 2.686$ and the iron ribs of the yoke support structure. With the goal of reconstructing the momenta and charge of muons over a wide angular and kinematic range, three types of gas-ionization detection mechanisms are employed: drift tubes (DTs), cathode strip chambers (CSCs), and resistive plate chambers (RPCs).

The barrel DT system consists of drift cells, $13\text{ mm} \times 42\text{ mm} \times 2.4\text{ m}$ chambers filled with 85% Ar and 15% CO₂ gas, with outer cathode strips and an inner anode wire to read out charge carriers ionized when a charged particle passes through the gas. Four drift cells are stacked, staggered by half a cell, to form superlayers (SLs). SLs are combined in groups of two or three to form drift chambers. The inner three stations have sixty DT chambers, with cells having anode wires running in the $r - \phi$ and z directions. The outer station has seventy DT chambers with cells having anode wires running only in the $r - \phi$ direction. A schematic diagram of the layout of the DT chambers, layered in the iron yoke, is shown in Figure 2.9.

The end cap CSC system consists of 468 trapezoidal CSC modules, proportional counters with six azimuthal anode wires running perpendicular to seven radial cathode strips, providing measurements of r and ϕ with $80\mu\text{m}$ resolution with pseudorapidity coverage $0.9 < |\eta| < 2.4$. Muons passing though the chambers' 40%:50%:10%, Ar:CO₂:CF₄ gas mixture produce an avalanche of positively charged carriers, whose signals are interpolated across multiple cathode strips along the anode wires to reconstruct the avalanche position. The position of the CSCs in a cutout quadrant of CMS is shown in Figure 2.10.

The final muon subsystem, the 480 rectangular barrel and trapezoidal end cap RPCs, are arranged in six barrel layers, two in stations one and two, and one each in stations three and four, and in three layers in each end cap. The RPC modules contain parallel-plate detectors with 2 – 3 double-gap modules of up to 96 (32) strips, parallel (radial) to the beam line in the barrel (end cap) sections, which collect ionized charge carriers from the 96.2%:3.5%:0.3%, C₂H₂F₄:C₄H₁₀:SF₆ gas. A schematic of the module layout is shown in Figure 2.11. The timescale in which RPCs can tag events is faster than the 25

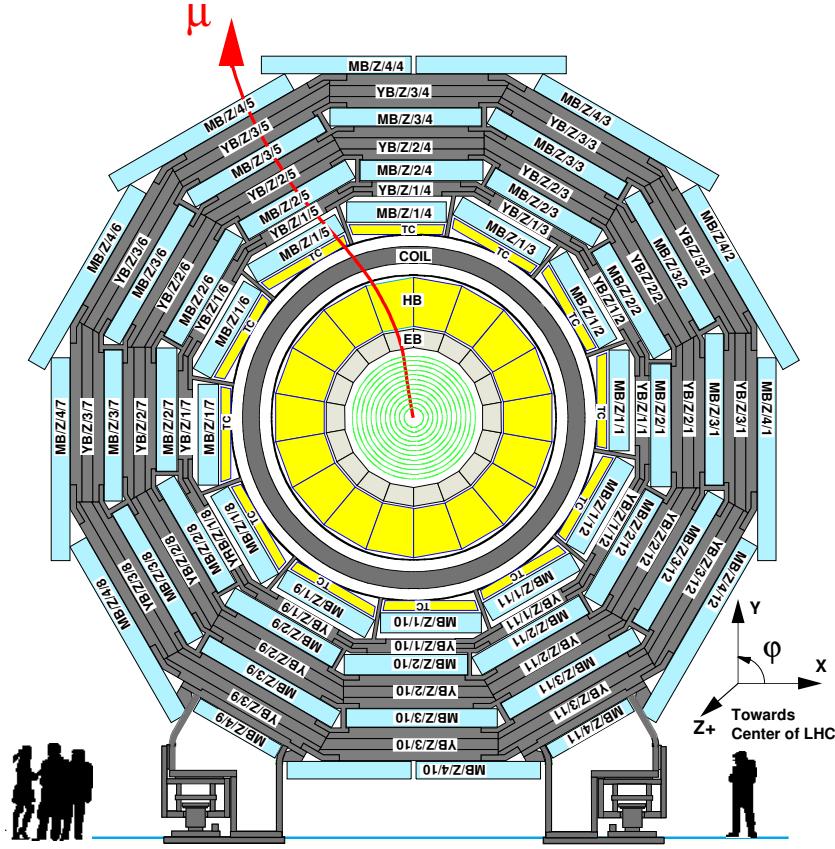


Figure 2.9: Schematic diagram showing DT chambers in light blue [2].

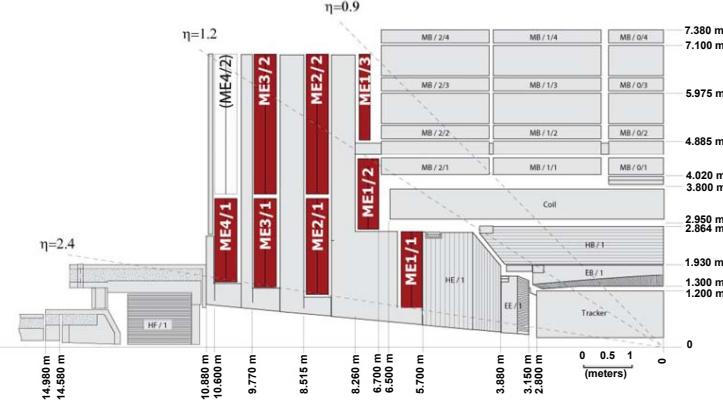


Figure 2.10: Schematic diagram showing CSC locations in red in a quadrant cutout of CMS [2].

ns bunch crossing time of the LHC, which in combination with the other muon systems, allows efficient triggering on muon events, discussed further in the next section.

The objective of the muon system to accurately measure the momenta of muons over

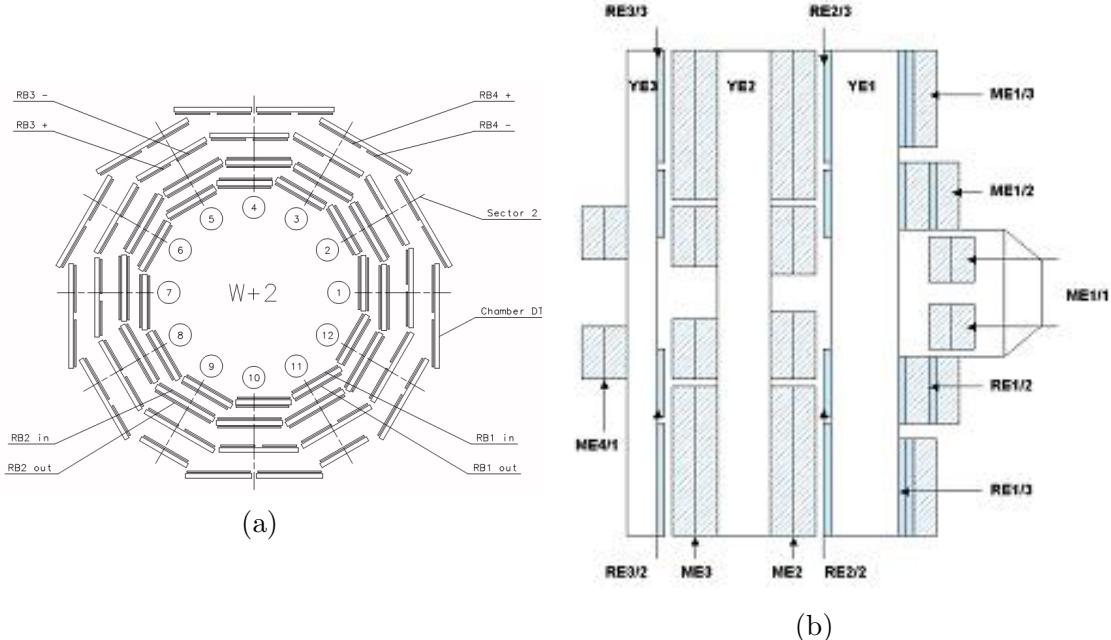


Figure 2.11: Schematic diagram showing barrel RPC locations (left) and end cap RPCs (right) in a cross sectional cutout of CMS [2].

a wide pseudorapidity range is accomplished, as shown by the less than 6% transverse momentum resolution shown in Figure 2.12 [3].

2.2.5 Trigger

In order to reduce the $O(100)$ MHz interaction rate from LHC proton collisions to a computationally manageable $O(100)$ kHz rate, a trigger system consisting of an Level-1 Trigger (L1T) and a High-Level Trigger (HLT) is employed. The L1T, summarized in Figure 2.13, uses mainly field programmable gate array (FPGA) technology on the front-end electronics to compile local trigger primitives from calorimeter towers and muon tracks into regional triggers, which use pattern logic to identify objects like electrons and muons. The highest quality objects are piped to the global trigger, which decides, based on further calculations and input on the status of the subdetectors, whether an event is rejected or an L1 Accept signal is sent to the Timing, Trigger, and Control (TTC) system for reading out the front-end data buffers, with a latency of $3.2 \mu\text{s}$.

After an event passes the L1T, the entire event data is read out by the data acquisition (DAQ) system, summarized in Figure 2.14. First, data is read out from the subdetector

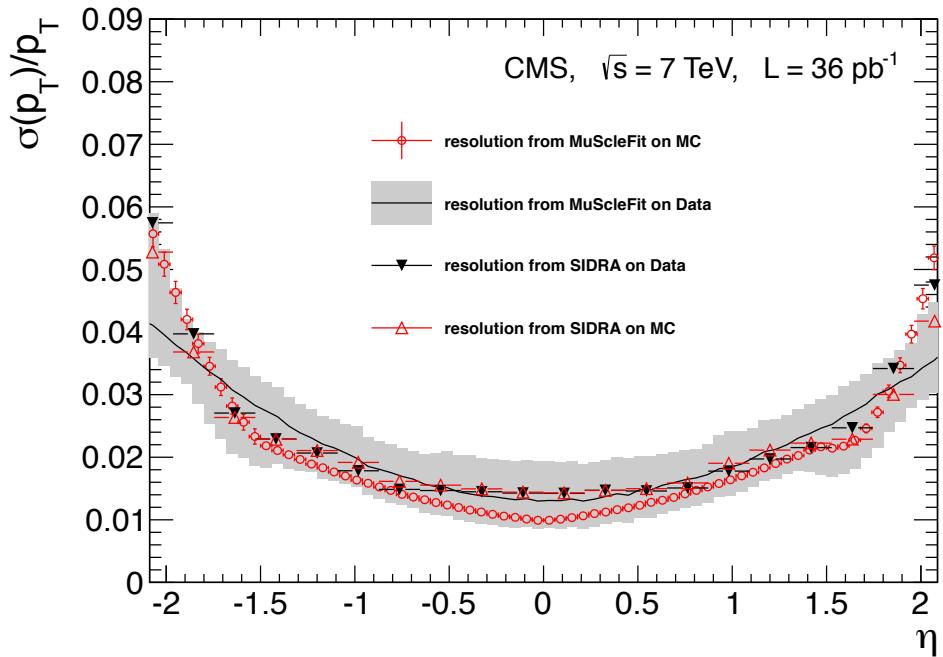


Figure 2.12: Transverse momentum resolution versus pseudorapidity for muons from Z decays [3].

front-end buffers to the front-end drivers (FEDs), followed by the merging of FED fragments by the Event Builder and the submission of the complete event data to the HLT by the Event Filter. The HLT is a software system, which uses filtering and reconstruction algorithms to select events based on their physics object content, with different paths based on the different combinations of objects used in physics analyses. The HLT paths used in this analysis are based on combinations of high quality lepton objects, such as two muons or two electrons, and are discussed further in the next chapter. Data quality monitoring (DQM), is also carried out at this step, to ensure all subsystems are behaving properly and the data being collected is usable for analysis. Data passing the HLT is saved to storage and processed in the offline software system before being provided to analyzers for physics searches.

In order to verify that the live trigger system is performing as expected and to emulate the trigger system for the production of simulated events, an offline trigger software framework is developed and maintained in parallel with the hardware system. For each calorimeter and muon trigger subsystem, there is a corresponding software emulator that

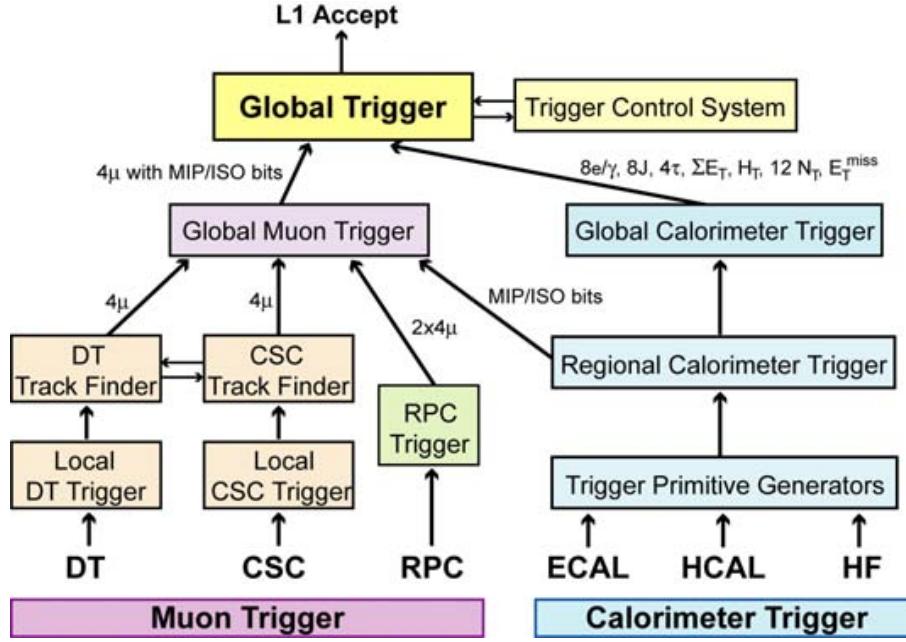


Figure 2.13: CMS L1 trigger system schematic diagram [2].

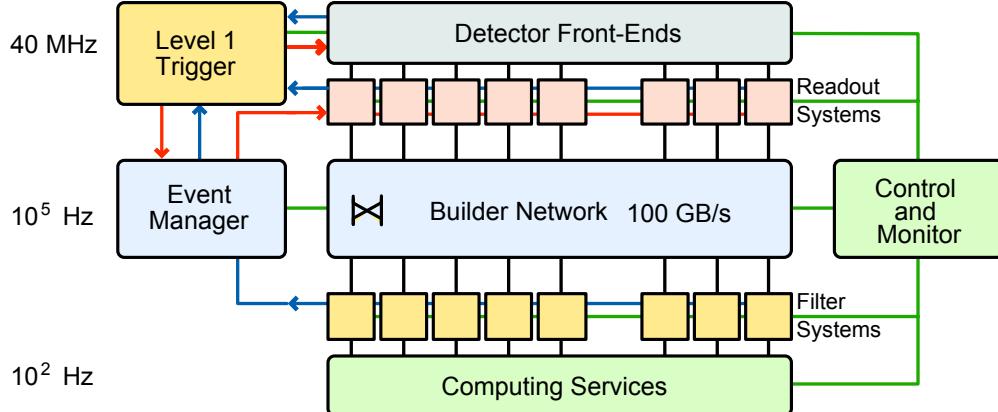


Figure 2.14: CMS data acquisition system schematic diagram [2].

simulates the hardware decisions. The live DQM system used by trigger shifters compares the hardware output to the output of the emulators. In addition to DQM, the trigger software framework is used to calculate the L1T efficiencies, or the efficiency of detection for various reconstructed physics objects. The service work conducted in the process of completing this thesis included contributing to the upgrade of the offline software to be based on calorimeter Layers instead of the regional (RCT) and global trigger (GCT) subsystems of the Run 1 legacy format described above, and the development of automated

workflows for trigger DQM. The Run 2 calorimeter trigger system is based on two layers: Layer1, responsible for constructing towers from the energy deposits in ECAL and HCAL modules, and Layer2, which algorithmically builds physics objects from Layer1 towers. A staggered upgrade plan was carried out with the GCT replaced by Layer2 in 2015 (called Stage1) and the RCT to Layer1 in 2016 (called Stage2). During partial upgrades, data formats would be upconverted or downconverted before or after an upgrade stage, where applicable, to keep the entire workflow functioning.

Chapter 3

Event reconstruction and simulation

This chapter overviews the algorithms used to reconstruct the trajectories and identify (ID) types of particles produced in proton-proton collisions in CMS, collectively known as Particle Flow (PF) [4], and how these collisions are simulated.

3.1 Particle reconstruction

The PF algorithms combine information from all of the CMS subdetectors discussed in the previous chapter to reconstruct the particles produced in the collision event. Since many of the particles produced initially in the collision are unstable, decaying before they have time to interact with the subdetectors, PF only reconstructs the stable particles: electrons, muons, photons, and hadrons. The remaining physics objects of interest, jets, missing energy, taus, etc, can be determined from the information provided by the stable PF identified particles.

The different particles are reconstructed and identified using information from individual subdetectors, or combinations of subdetectors. The direction and momentum of charged particles is measured by the tracker. Electrons are reconstructed using tracks and energy deposits in the ECAL. Muons are reconstructed from a combination of tracker and muon chamber data. Photons are reconstructed from energy deposits in the ECAL. Finally, charged and neutral hadrons are reconstructed from energy deposited primarily in the HCAL, with a contribution from energy deposits in the ECAL. The MET, an observable of particular importance to this analysis, and used to identify DM that does not

interact with the detector material, is the modulus of the sum of transverse momenta of all the PF reconstructed particles.

The basic pieces of information from the subdetectors used by PF are called elements, and consist of charged-particle tracks, muon tracks, and calorimeter clusters. The tracker provides charged-particle track elements. Since the tracker has the best momentum resolution of the subdetectors, it is of critical importance that the tracking efficiency be nearly 100%, with as low a fake rate as possible, to reduce an excess in reconstructed energy. This goal is accomplished using an iterative algorithm: first, tracks are seeded using very tight criteria, yielding a low efficiency, but negligible fake rate, then track seed criteria are loosened and hits that clearly belong to a track are removed, resulting in increasing efficiency. The ECAL and HCAL subsystems (ECAL barrel, HCAL barrel, HCAL end cap, PS first layer, and PS second layer) provide cluster elements. The calorimeter clustering algorithm measures the energy and direction of neutral particles (i.e. photons and neutral hadrons), differentiates energy deposits from neutral and charged hadrons, reconstructs electrons, and contributes to the reconstruction of charged hadrons. The algorithm is summarized as follows: cluster seeds are identified as energy deposit peaks over a given energy, from which topological clusters are grown by appending adjacent cells, and last, topological clusters seed PF clusters. An example is shown in Figure 3.1 and Figure 3.2, where a simple jet is reconstructed into four clusters, shown as dots.

Once the PF elements are determined, they are linked together into blocks, which correspond to the signatures left in the subdetectors of a single particle. Single particles typically leave one to three elements. The linking algorithm determines the quality of the link between all pairwise elements in an event, then forms blocks from the highest quality links, starting from the tracker and proceeding outward through the calorimeters and muon chambers. Once the blocks are formed, PF associates a global event particle with each block. A PF muon is formed from a global muon candidate if its momentum is consistent across all track elements. PF electrons are identified from electron candidates using tracker and ECAL cluster variables, accounting for the bremsstrahlung photons produced when the electron passes through the tracker material. Once the elements asso-

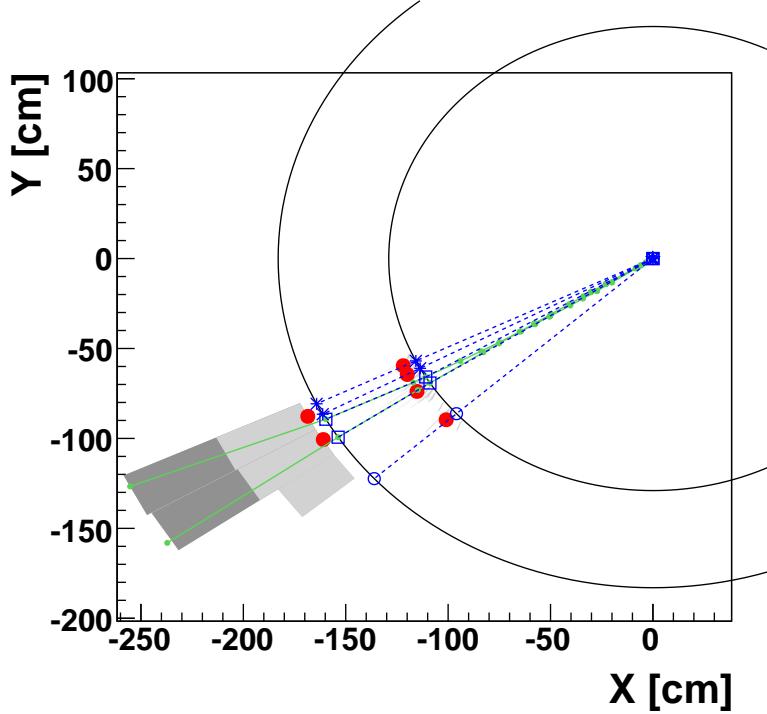


Figure 3.1: Event display of hadronic jet in the $x - y$ plane, with solid arcs at the ECAL and HCAL surfaces. The locations of clusters are given by the solid dots [4].

ciated to PF muons and PF electrons are removed, the remaining elements are analyzed to identify charged hadrons, photons, or neutral hadrons. PF charged hadrons are associated to remaining tracks if the linked clusters are consistent with the measured momenta. If the energy of the linked clusters is much larger than the track momentum, accounting for uncertainties, a PF photon or PF neutral hadron is formed. Any remaining clusters without linked tracks form PF photons or PF neutral hadrons.

As previously discussed, once the PF particles are identified, additional information about the event can be inferred. A quantity of particular importance to this analysis is the missing transverse energy (MET), defined above. The performance of the PF algorithms' determination of the MET is shown in Figure 3.3 by the resolution of PF measured MET as a function of the true MET to be within $\pm 5\%$ above 20 GeV.

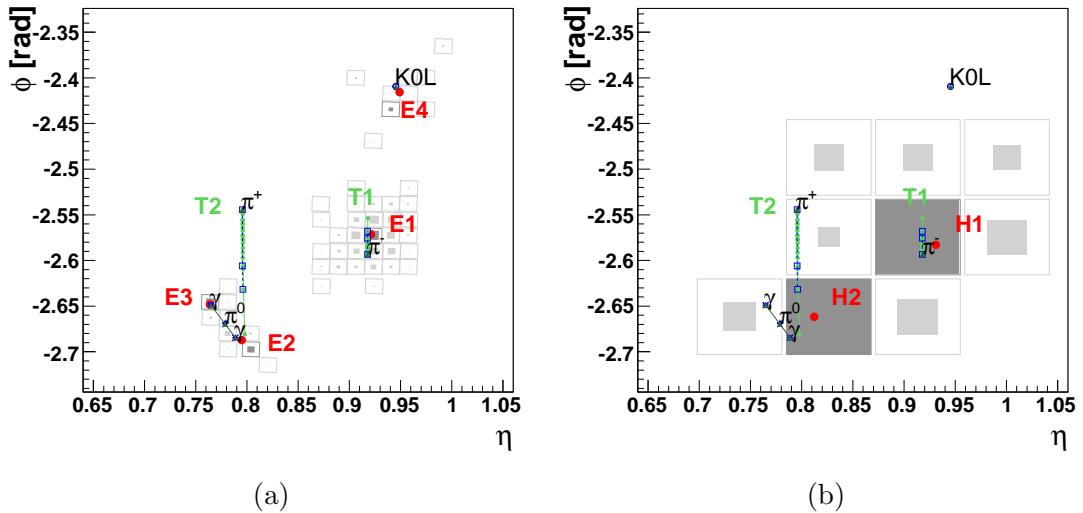


Figure 3.2: Event display of hadronic jet in the $\eta - \phi$ plane for the ECAL (a) and HCAL (b). The locations of clusters are given by the solid dots [4].

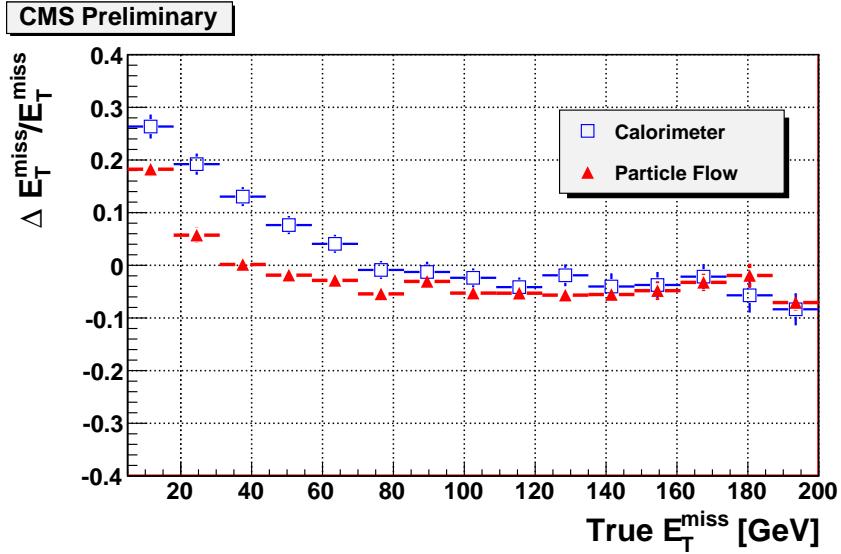


Figure 3.3: Missing transverse energy reconstruction resolution using the particle flow algorithm [4].

3.2 Monte Carlo event simulation

The simulation of proton collision events and their detection using Monte Carlo (MC) techniques is useful for several purposes. In addition to using the simulated events to test the detector hardware and software performance without collecting true data, simulations are used to build background models when searching the data for new physics processes.

New physics signatures usually appear as excesses in data above a SM background. The background model consists of SM processes which produce the same or similar signature as the new signal being searched for. These processes are modeled either using purely simulated events or a combination of simulated events and data-driven techniques. In either case, it is often necessary to weight the background events by correction scale factors measured using data, which account for shortcomings of the simulations, such as the inability to perform perturbative QCD calculations for low momentum transfer processes. MC event generation can be factored into two parts: (1) modeling the initial particles produced in a collision event and (2) modeling how these initial particles evolve in and interact with the detector.

The first part of MC event generation is modeling the proton-proton collision and the initial particles produced at the primary vertex. Several software packages are used by CMS analysts to generate collision events and calculate the cross sections of the simulated processes, including PYTHIA [139], MADGRAPH [140], BlackHat and Sherpa [141], and POWHEG [142]. The packages have different implementations, but the underlying principles are the same. The momenta of the proton partons (i.e. quarks and gluons) that interact in the initial scatter are determined probabilistically by random sampling from the parton distribution functions (PDFs), which give the probability that a parton will carry a fraction x of the proton momentum. This is straightforward for processes with two incoming and one outgoing particle ($2 \rightarrow 1$) and two incoming and two outgoing particles ($2 \rightarrow 2$), in which the outcomes are weighted by their relative cross sections and determined probabilistically. However, for radiative processes, such as ISR and FSR of a photon or gluon, which are generally $1 \rightarrow 2$ processes, higher order matrix elements must be calculated or approximated. Once the initial particles are determined, their fragmentation and decays are simulated in a process called hadronization, until the final stable particles are produced.

The second part of MC event generation is simulating the detector response to the stable particles produced in the first step, including their interaction with the detector material itself, both active elements and structural material. The primary software pack-

age used in this step by CMS is GEANT [143], in which a complete digital representation of the CMS detector is built. GEANT simulates the passage of each stable particle, step-by-step, outward through the detector, probabilistically determining the interaction that occurs at each step depending on the particle’s energy, material it is in, and the EM field present. The detector is not perfectly efficient because the acceptance is less than one and the reconstruction efficiency of the individual detector elements is suboptimal, so calibration values must be measured at CMS and fed back in to the simulations, in order to accurately simulate the performance of the detector. Once the final response of the detector is simulated, the resulting MC may be weighted by scale factors measured using real data, in order to correct for mismodeling of the detector.

3.3 Datasets

3.3.1 Data

3.3.1.1 Triggers and Datasets

This analysis uses a data sample recorded by the CMS experiment during 2016, corresponding to 35.9 fb^{-1} of data. The datasets are listed in Table 3.1, along with their integrated luminosities. The analysis relies on five different primary datasets (PDs) – DoubleEG, DoubleMuon, MuEG, SingleElectron, and SingleMuon – each of which combines a certain collection of HLT paths. To avoid duplicate events from different primary datasets, events are taken:

- from DoubleEG if they pass the diEle or triEle triggers,
- from DoubleMuon if they pass the diMuon or triMuon triggers and fail the diEle and triEle triggers,
- from MuEG if they pass the MuEle or MuDiEle or DiMuEle triggers and fail the diEle, triEle, diMuon and triMuon triggers,
- from SingleElectron if they pass the singleElectron trigger and fail all the above triggers.
- from SingleMuon if they pass the singleMuon trigger and fail all the above triggers.

The HLT paths used for 2016 collision data are listed in Table 3.2, together with their L1 seeds, prescale values and the associated primary datasets.

Run-range	Dataset	Integrated luminosity
273150-275376	/DoubleMuon/Run2016B-23Sep2016-v3/AOD /DoubleEG/Run2016B-23Sep2016-v3/AOD /MuonEG/Run2016B-23Sep2016-v3/AOD /SingleElectron/Run2016B-23Sep2016-v3/AOD /SingleMuon/Run2016B-23Sep2016-v3/AOD	5.892 fb ⁻¹
275656-276283	/DoubleMuon/Run2016C-23Sep2016-v1/AOD /DoubleEG/Run2016C-23Sep2016-v1/AOD /MuonEG/Run2016C-23Sep2016-v1/AOD /SingleElectron/Run2016C-23Sep2016-v1/AOD /SingleMuon/Run2016C-23Sep2016-v1/AOD	2.646 fb ⁻¹
276315-276811	/DoubleMuon/Run2016D-23Sep2016-v1/AOD /DoubleEG/Run2016D-23Sep2016-v1/AOD /MuonEG/Run2016D-23Sep2016-v1/AOD /SingleElectron/Run2016D-23Sep2016-v1/AOD /SingleMuon/Run2016D-23Sep2016-v1/AOD	4.353 fb ⁻¹
276831-277420	/DoubleMuon/Run2016E-23Sep2016-v1/AOD /DoubleEG/Run2016E-23Sep2016-v1/AOD /MuonEG/Run2016E-23Sep2016-v1/AOD /SingleElectron/Run2016E-23Sep2016-v1/AOD /SingleMuon/Run2016E-23Sep2016-v1/AOD	4.117 fb ⁻¹
277932-278808	/DoubleMuon/Run2016F-23Sep2016-v1/AOD /DoubleEG/Run2016F-23Sep2016-v1/AOD /MuonEG/Run2016F-23Sep2016-v1/AOD /SingleElectron/Run2016F-23Sep2016-v1/AOD /SingleMuon/Run2016F-23Sep2016-v1/AOD	3.186 fb ⁻¹
278820-280385	/DoubleMuon/Run2016G-23Sep2016-v1/AOD /DoubleEG/Run2016G-23Sep2016-v1/AOD /MuonEG/Run2016G-23Sep2016-v1/AOD /SingleElectron/Run2016G-23Sep2016-v1/AOD /SingleMuon/Run2016G-23Sep2016-v1/AOD	7.721 fb ⁻¹
281207-284068	/DoubleMuon/Run2016H-PromptReco-v1/AOD /DoubleEG/Run2016H-PromptReco-v1/AOD /MuonEG/Run2016H-PromptReco-v1/AOD /SingleElectron/Run2016H-PromptReco-v1/AOD /SingleMuon/Run2016H-PromptReco-v1/AOD /DoubleMuon/Run2016H-PromptReco-v2/AOD /DoubleEG/Run2016H-PromptReco-v2/AOD /MuonEG/Run2016H-PromptReco-v2/AOD /SingleElectron/Run2016H-PromptReco-v2/AOD /SingleMuon/Run2016H-PromptReco-v2/AOD /DoubleMuon/Run2016H-PromptReco-v3/AOD /DoubleEG/Run2016H-PromptReco-v3/AOD /MuonEG/Run2016H-PromptReco-v3/AOD /SingleElectron/Run2016H-PromptReco-v3/AOD /SingleMuon/Run2016H-PromptReco-v3/AOD	8.857 fb ⁻¹

Table 3.1: Run 2 datasets used in the $H \rightarrow ZZ$ analysis. The first column specifies the range of run ids, indices counting up from the first CMS run, for each dataset. The middle column gives the CMS database file path where the data files are stored. The third column gives the total integrated luminosities for each collection of runs.

HLT path	L1 seed	prescale	primary dataset
HLT_Ele17_Ele12_CaloIdL_TrackIdL_IsoVL_DZ	L1_DoubleEG_15_10	1	DoubleEG
HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ	L1_DoubleEG_22_10	1	DoubleEG
HLT_DoubleEle33_CaloIdL_GsfTrkIdVL	(Multiple)	1	DoubleEG
HLT_Ele16_Ele12_Ele8_CaloIdL_TrackIdL	L1_TripleEG_14_10_8	1	DoubleEG
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL	L1_DoubleMu_11_4	1	DoubleMuon
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL	L1_DoubleMu_11_4	1	DoubleMuon
HLT_TripleMu_12_10_5	L1_TripleMu_5_5_3	1	DoubleMuon
HLT_Mu8_TrkIsoVVL_Ele17_CaloIdL_TrackIdL_IsoVL	L1_Mu5_EG15	1	MuonEG
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL	L1_Mu5_EG20	1	MuonEG
HLT_Mu17_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL	L1_Mu12_EG10	1	MuonEG
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL	L1_Mu20_EG10	1	MuonEG
HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL	L1_SingleMu*	1	MuonEG
HLT_Mu8_DiEle12_CaloIdL_TrackIdL	L1_Mu6_DoubleEG10	1	MuonEG
HLT_DiMu9_Ele9_CaloIdL_TrackIdL	L1_DoubleMu7_EG7	1	MuonEG
HLT_Ele25_eta2p1_WPTight	L1_SingleEG*	1	SingleElectron
HLT_Ele27_WPTight	L1_SingleEG*	1	SingleElectron
HLT_Ele27_eta2p1_WPLoose_Gsf	L1_SingleEG*	1	SingleElectron
HLT_IsoMu20 OR HLT_IsoTkMu20	L1_SingleMu*	1	SingleMuon
HLT_IsoMu22 OR HLT_IsoTkMu22	L1_SingleMu*	1	SingleMuon

Table 3.2: Trigger paths used in 2016 collision data.

3.3.1.2 Trigger Efficiency

The efficiency for data events to pass the combination of triggers with respect to the offline reconstruction and selection is measured by considering four-lepton events triggered by single lepton triggers. One of the four reconstructed leptons, called the “tag,” is geometrically matched to a trigger object passing the final filter of one of the single muon or single electron triggers. The other three leptons are used as “probes.” In each four-lepton event, there are up to four possible tag-probe combinations, and all possible combinations are counted in the denominator of the efficiency. For each of the three probe leptons, all matching trigger filter objects are collected. Then, the matched trigger filter objects of the three probe leptons are combined in attempt to reconstruct any of the triggers used in the analysis. If any of the analysis triggers can be formed using the probe leptons, the set of probes is also counted in the numerator of the efficiency.

This method does not have a perfect closure in MC events due to the fact that the presence of a fourth lepton increases the trigger efficiency, and this effect is not accounted for. Also, in the $2e2\mu$ final state, the three probe leptons cannot be combined to form all possible triggers which can collect events with two electrons and two muons (e.g. if the tag lepton is an electron, the three remaining leptons cannot pass a double electron trigger). Therefore, the method is also exercised on MC, and the difference between data

and MC is used to determine the reliability of the simulation. The efficiency plotted as a function of the minimum transverse momentum (p_T) of the three probe leptons in data and MC using this method can be seen in Figure 3.4. The MC efficiency describes the data within the statistical uncertainties well.

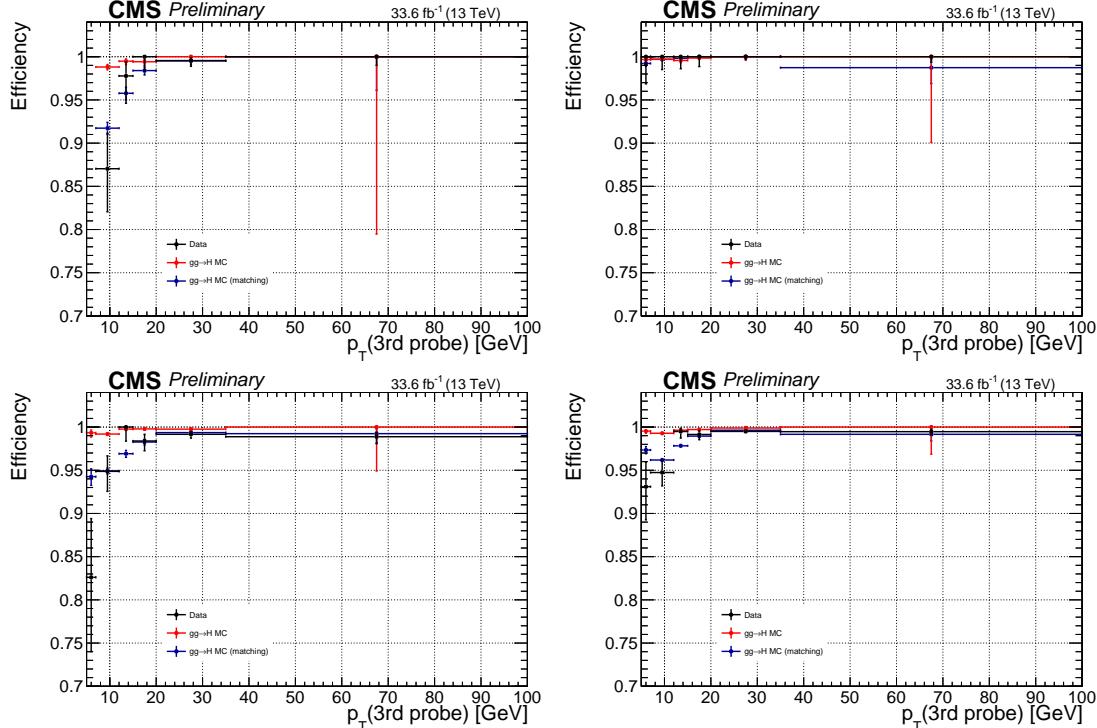


Figure 3.4: Trigger efficiency measured in data using 4ℓ events collected by single lepton triggers for the $4e$ (top left), 4μ (top right), $2e2\mu$ (bottom left) and 4ℓ (bottom right) final states.

A summary of the trigger efficiencies in MC truth, MC, and data using the tag and probe method, is shown in Table 3.3. The trigger efficiency in simulation is found to be $> 99\%$ in each final state.

3.3.2 Simulation

3.3.2.1 Signal Samples

The signal samples used are centrally produced for the benchmarks defined in Section 1.3.2.7 and are summarized in Table 3.4.

Final State	$gg \rightarrow H$ MC	$gg \rightarrow H$ MC (matching)	Data (matching)
$4e$	$0.991^{+.002}_{-.002}$	$0.948^{+.004}_{-.004}$	$0.982^{+.005}_{-.007}$
4μ	$0.997^{+.001}_{-.001}$	$0.997^{+.001}_{-.001}$	$1.000^{+.000}_{-.001}$
$2e2\mu$	$0.995^{+.001}_{-.001}$	$0.964^{+.002}_{-.002}$	$0.983^{+.003}_{-.004}$

Table 3.3: Trigger efficiencies measured using 4ℓ events.

Dataset	Parameters
/ZprimeToA0hToA0chichihZZTo4l_2HDM_MZp-*_MA0-300_13TeV-madgraph-pythia8/[1]	$m_{A^0} = 300$ GeV
/ZprimeToA0hToA0chichihZZTo4l_2HDM_MZp-*_MA0-*_13TeV-madgraph/[2]	$m_{A^0} \neq 300$ GeV
/MonoHZZ4l_ZpBaryonic_MZp-*_MChi-*_13TeV-madgraph/[3]	
[1] RunIISpring16DR80-premix_withHLT_80X_mcRun2_asymptotic_v14-v1/AODSIM	
[2] RunIISpring16reHLT80-PUSpring16RAWAODSIM_reHLT_80X_mcRun2_asymptotic_v14-v1/AODSIM	
[3] RunIISpring16DR80-premix_withHLT_80X_mcRun2_asymptotic_v14-v1/AODSIM	

Table 3.4: Benchmark signal samples analyzed.

3.3.2.2 Background Samples

Descriptions of the SM Higgs boson production are obtained using the POWHEG V2 [144, 145, 146] generator for the five main production modes: gluon fusion ($gg \rightarrow H$) including quark mass effects [147], vector boson fusion (VBF) [148], and associated production (WH , ZH and $t\bar{t}H$ [149]). In the case of WH and ZH , the MiNLO HVJ extension of POWHEG is used [150]. The description of the decay of the Higgs boson to four leptons is obtained using the JHUGEN generator [151]. In the case of WH , ZH and $t\bar{t}H$, the Higgs boson is allowed to decay to $H \rightarrow ZZ \rightarrow 2\ell 2X$, such that four-lepton events where two leptons originate from the decay of associated Z , W bosons or top quarks are also taken into account in the simulation. Showering of parton-level events is done using PYTHIA8.209, and in all cases, matching is performed by allowing QCD emissions at all energies in the shower and vetoing them afterwards according to the POWHEG internal scale. All samples are generated with the NNPDF 3.0 next-to-leading order (NLO) PDFs [152]. The list of Higgs signal samples and their cross sections are shown in Table 3.5.

Production of ZZ via quark-antiquark annihilation is generated at NLO using POWHEG V2 [153] and PYTHIA8, with the same settings as for the Higgs signal. As this simulation covers a large range of ZZ invariant masses, dynamical QCD factorization and renormal-

Process	Dataset	$\sigma \times BR(\times \epsilon_{\text{filter}})$
$gg \rightarrow H \rightarrow ZZ \rightarrow 4\ell$	/GluGluHToZZTo4L_M125_13TeV_powheg2_JHUGenV6_pythia8	12.18 fb
$qq \rightarrow Hqq \rightarrow ZZqq \rightarrow 4\ell qq$	/VBF_HToZZTo4L_M125_13TeV_powheg2_JHUGenV6_pythia8	1.044 fb
$q\bar{q} \rightarrow W^+H \rightarrow W^+ZZ \rightarrow 4\ell + X$	/WplusH_HToZZTo4L_M125_13TeV_powheg2-minlo-HWJ_JHUGenV6_pythia8	0.232 fb
$q\bar{q} \rightarrow W^-H \rightarrow W^-ZZ \rightarrow 4\ell + X$	/WminusH_HToZZTo4L_M125_13TeV_powheg2-minlo-HWJ_JHUGenV6_pythia8	0.147 fb
$q\bar{q} \rightarrow ZH \rightarrow ZZZ \rightarrow 4\ell + X$	/ZH_HToZZ_4LFilter_M125_13TeV_powheg2-minlo-HZJ_JHUGenV6_pythia8	0.668 fb
$gg \rightarrow ttH \rightarrow ttZZ \rightarrow 4\ell + X$	/ttH_HToZZ_4LFilter_M125_13TeV_powheg_JHUGen_pythia8	0.393 fb

Table 3.5: Higgs signal samples and production cross sections times branching fractions to four leptons times filter efficiencies.

Process	Dataset Name	$\sigma \cdot BR$
$qq \rightarrow ZZ \rightarrow 4\ell$	/ZZTo4L_13TeV_powheg_pythia8	1.256pb
$qq \rightarrow ZZ \rightarrow 4\ell$	/ZZTo4L_13TeV-amcatnloFXFX-pythia8	1.212pb
$gg \rightarrow ZZ \rightarrow 4e$	/GluGluToContinToZZTo4e_13TeV_MCFM701	0.00159pb
$gg \rightarrow ZZ \rightarrow 4\mu$	/GluGluToContinToZZTo4mu_13TeV_MCFM701	0.00159pb
$gg \rightarrow ZZ \rightarrow 4\tau$	/GluGluToContinToZZTo4tau_13TeV_MCFM701	0.00159pb
$gg \rightarrow ZZ \rightarrow 2e2\mu$	/GluGluToContinToZZTo2e2mu_13TeV_MCFM701	0.00319pb
$gg \rightarrow ZZ \rightarrow 2e2\tau$	/GluGluToContinToZZTo2e2tau_13TeV_MCFM701	0.00319pb
$gg \rightarrow ZZ \rightarrow 2\mu2\tau$	/GluGluToContinToZZTo2mu2tau_13TeV_MCFM701	0.00319pb
$Z \rightarrow \ell\ell + \text{jets}$	/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	6104pb
$Z \rightarrow \ell\ell + \text{jets}$	/DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	18610pb
$WZ \rightarrow 3\ell\nu$	/WZTo3LNu_TuneCUETP8M1_13TeV-powheg-pythia8	4.430pb
$t\bar{t}$	/TTJets_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	815.96pb
$t\bar{t} \rightarrow 2\ell 2\nu 2b$	/TTTo2L2Nu_13TeV-powheg	87.31pb

Table 3.6: Background Monte Carlo samples and cross sections.

ization scales have been chosen, equal to m_{ZZ} .

The $gg \rightarrow ZZ$ process is simulated at leading order (LO) with MCFM [154, 155]. In order to match the $gg \rightarrow H \rightarrow ZZ$ transverse momentum spectra predicted by POWHEG at NLO, the showering for MCFM samples is performed with different PYTHIA8 settings, allowing only emissions up to the parton-level scale (“wimpy” shower).

Although not directly used to model data observations, additional MC samples of WZ , Drell-Yan+jets, $t\bar{t}$, and tribosons are generated using MADGRAPH5_AMCATHNLO [156] either inclusively or merging several jet multiplicities. Table 3.6 summarizes the MC simulation datasets used for this analysis.

3.3.2.3 Pileup Reweighting

The MC samples are reweighted to match the pileup distribution measured in 2016 data. Scale factors are measured and applied to each event weight before histograms are filled and yields are calculated, based on the number of pileup vertices present in the event. The mean number of pileup vertices for data measured in 2016 is about 20. Figure 3.5 shows the distributions of the numbers of pileup vertices for data and MC before and after the events are reweighted.

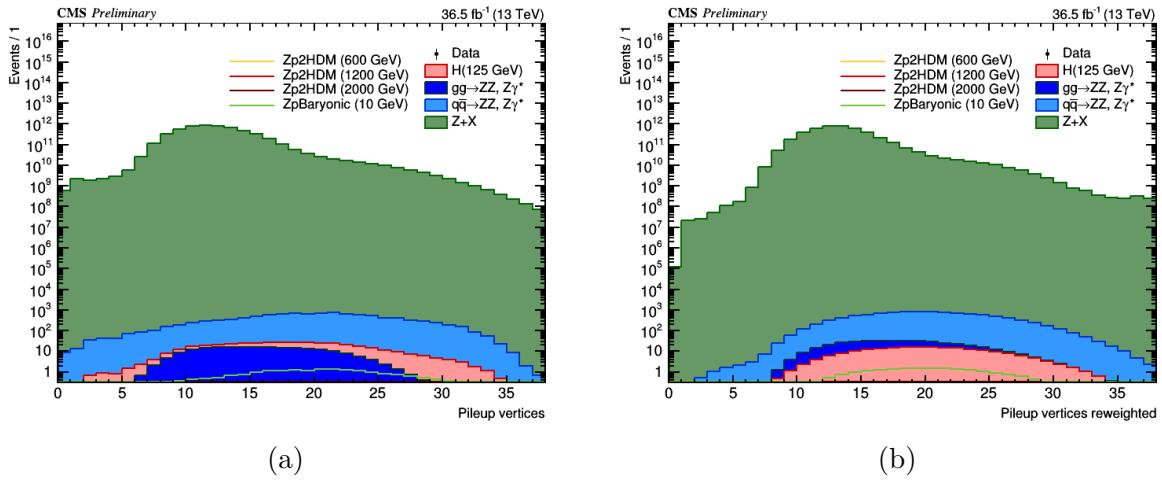


Figure 3.5: Number of pileup vertices before (a) and after (b) reweighting is applied.

Chapter 4

Physics objects

4.1 Electrons

4.1.1 Electron reconstruction

Electron candidates are preselected using loose cuts on track-cluster matching observables, so as to preserve the highest possible efficiency while rejecting part of the QCD background. To be considered for the analysis, electrons are required to have a transverse momentum $p_T^e > 7$ GeV, a reconstructed $|\eta^e| < 2.5$, and to satisfy a loose primary vertex constraint defined as $d_{xy} < 0.5$ and $d_z < 1$. Such electrons are called “loose” electrons.

The early runs in the 2016 data-taking exhibit a tracking inefficiency originating from a reduced hit reconstruction efficiency in the strip detector (“HIP“ effect). The resulting data-MC discrepancy is corrected using scale factors as is done for the electron selection with data efficiencies measured using the same tag-and-probe technique outlined later (see Section 4.1.5). These studies are carried out by the CMS electron and photon (EGM) physics object group (POG) and the results are summarized in this section. The electron reconstruction scale factors are shown Figure 4.1 as a one-dimensional function of the super cluster η only, as it was shown that the p_T dependence of the scale factor is negligible. For more details on electron reconstruction, see Khachatryan et al. [157].

4.1.2 Electron identification

Reconstructed electrons are identified by means of a Gradient Boosted Decision Tree (GBDT) multivariate classifier algorithm, which exploits observables from the EM cluster,

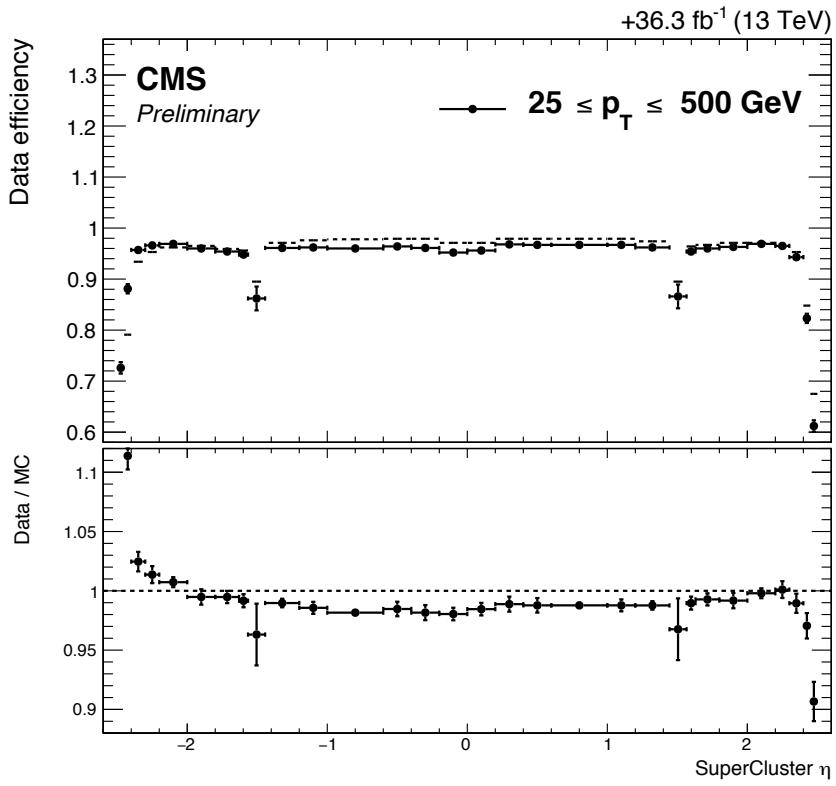


Figure 4.1: Electron reconstruction efficiencies in data versus η and data/MC scale factors as provided by the EGM POG.

the matching between the cluster and the electron track, as well as observables based exclusively on tracking measurements. The BDT has been retrained using CMSSW_8_0_X samples. The classifier is trained on Drell-Yan plus jets MC sample for both signal and background.

The impact of the retraining of the ID for the 2016 conditions is illustrated in the receiver operating characteristic (ROC) curves shown in Figure 4.2. Several studies to improve the performance of the multivariate analysis (MVA) for the harsher 2016 running conditions were performed. One study considered a new splitting of the BDT training bins, where electrons falling into the gap regions of the ECAL, e.g. the EB-EE transition region, were trained separately from the non-gap electrons. However, no improvement for either population was observed, indicating that the current setup is already able to properly take the significantly differing input distributions in those regions into account. Additional

variables were also studied, including more cluster-shape observables. Still, none of these variables helped to improve the performance in the relevant $> 95\%$ signal efficiency regime, though up to a 20% improved background rejection was seen for 80% of working points. Finally, the hyperparameters of the MVA were systematically scanned for their optimal values, but the resulting configuration was found to improve the overall performance only marginally by $< 10\%$ and introduced a significant overtraining effect. Due to the small gains and large overtraining, it was decided to not modify the hyperparameters beyond the interface changes from the latest 4.2.0 version of the TMVA package.

Figure 4.3 shows the output of the BDT on the training and testing samples for true and fake electrons for the high- p_T training bin in the end cap. The good agreement between the training and testing distributions is similar across the six training bins and indicates that the classifier has not been overtrained.

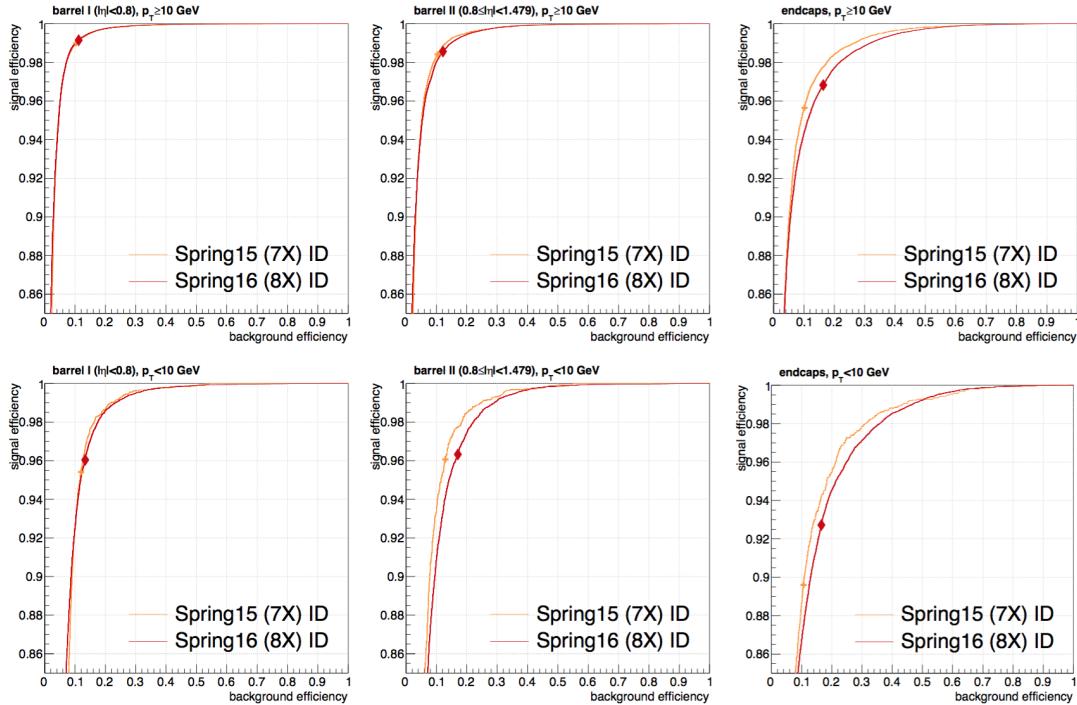


Figure 4.2: Performance comparison of the MVA trained for the 2015 analysis and the retraining for 2016 conditions. The respective working points are indicated by the markers.

Table 4.1 summarizes the full list of observables used as inputs to the classifier and

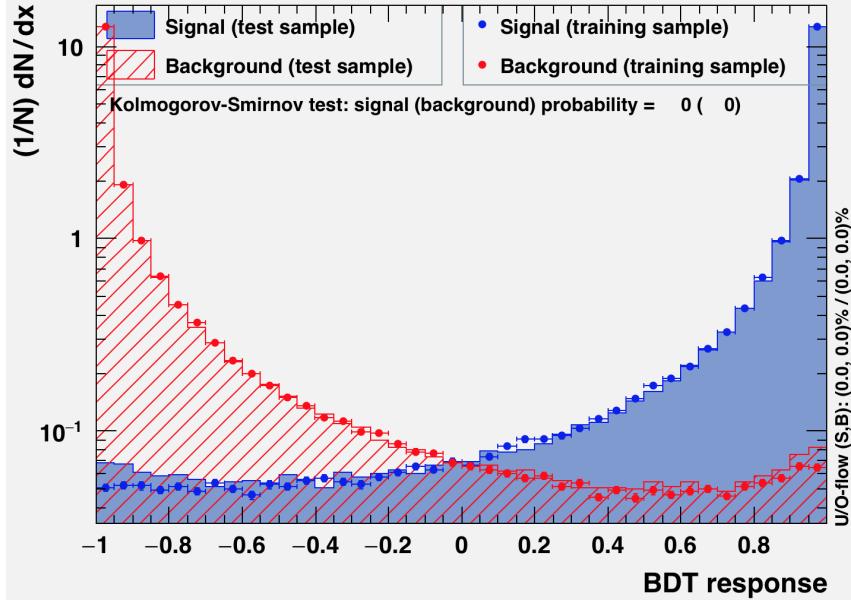


Figure 4.3: Boosted decision tree output for the training and testing sample for true and fake electrons in the high- p_T end cap training bins.

Table 4.2 lists the cut values applied to the BDT score for the chosen working point. For the analysis, we define "tight" electrons as the loose electrons that pass this MVA identification working point.

observable type	observable name
cluster shape	RMS of the energy-crystal number spectrum along η and φ ; $\sigma_{i\eta i\eta}$, $\sigma_{i\varphi i\varphi}$
	supercluster width along η and ϕ
	'ratio of the hadronic energy behind the electron supercluster to the supercluster energy, H/E
	circularity $(E_{5 \times 5} - E_{5 \times 1})/E_{5 \times 5}$
track-cluster matching	sum of the seed and adjacent crystal over the super cluster energy R_9
	for end cap training bins: energy fraction in preshower E_{PS}/E_{raw}
tracking	energy-momentum agreement E_{tot}/p_{in} , E_{ele}/p_{out} , $1/E_{tot} - 1/p_{in}$
	position matching $\Delta\eta_{in}$, $\Delta\varphi_{in}$, $\Delta\eta_{seed}$
	fractional momentum loss $f_{brem} = 1 - p_{out}/p_{in}$
	number of hits of the KF and GSF track N_{KF} , N_{GSF} (·)
	reduced χ^2 of the KF and GSF track χ^2_{KF} , χ^2_{GSF}
	number of expected but missing inner hits (·)
	probability transform of conversion vertex fit χ^2 (·)

Table 4.1: Overview of input variables to the identification classifier. Variables not used in the Run 1 MVA are marked with (·).

minimum BDT score	$ \eta < 0.8$	$0.8 < \eta < 1.479$	$ \eta > 1.479$
$5 < p_T < 10 \text{ GeV}$	-0.211	-0.396	-0.215
$p_T > 10 \text{ GeV}$	-0.870	-0.838	-0.763

Table 4.2: Minimum boosted decision tree score required for passing the electron identification.

4.1.3 Electron isolation

The relative isolation for electrons is defined as:

$$\text{RelPFiso} = \left(\sum_{\text{charged}} p_T + \sum_{\text{neutral}}^{\text{corr}} p_T \right) / p_T^{\text{lepton}} \quad (4.1)$$

where the corrected neutral component of isolation is computed using the formula:

$$\sum_{\text{neutral}}^{\text{corr}} p_T = \max\left(\sum_{\text{neutral}}^{\text{uncorr}} p_T - \rho \times A_{\text{eff}}, 0 \text{GeV}\right) \quad (4.2)$$

and the mean pileup contribution to the isolation cone is obtained as:

$$PU = \rho \times A_{\text{eff}} \quad (4.3)$$

where ρ is the mean energy density in the event, and the effective area A_{eff} is defined as the ratio between the slope of the average isolation and that of ρ as a function of the number of vertices.

The electron isolation working point was optimized and chosen to be $\text{RelPFiso}(\Delta R = 0.3) < 0.35$ [158].

4.1.4 Electron energy calibrations

Electrons in data are corrected for features in ECAL energy scale in bins of p_T and $|\eta|$. Corrections are calculated on a $Z \rightarrow ee$ sample to align the dielectron mass spectrum in the data to that in the MC and to minimize its width.

The $Z \rightarrow ee$ mass resolution in MC is made to match data by applying a pseudorandom Gaussian smearing to electron energies, with Gaussian parameters varying in bins of p_T and $|\eta|$. This has the effect of convoluting the electron energy spectrum with a Gaussian.

The electron energy scale is measured in data by fitting a Crystal-Ball function to the dielectron mass spectrum around the Z peak in the $Z + \ell$ control region. The energy scale for the full 2016 dataset is shown in Figure 4.4(a) and agrees with the MC with 100 MeV. The stability of the energy scale across different run periods is shown in Fig. 4.4(b), where the data is binned into approximately 500 pb luminosity blocks.

4.1.5 Electron efficiency measurements

The tag-and-probe (T&P) study was performed on the single electron primary datasets listed in Table 3.1 using the same golden JSON of 36.8 fb^{-1} as for the main analysis [158].

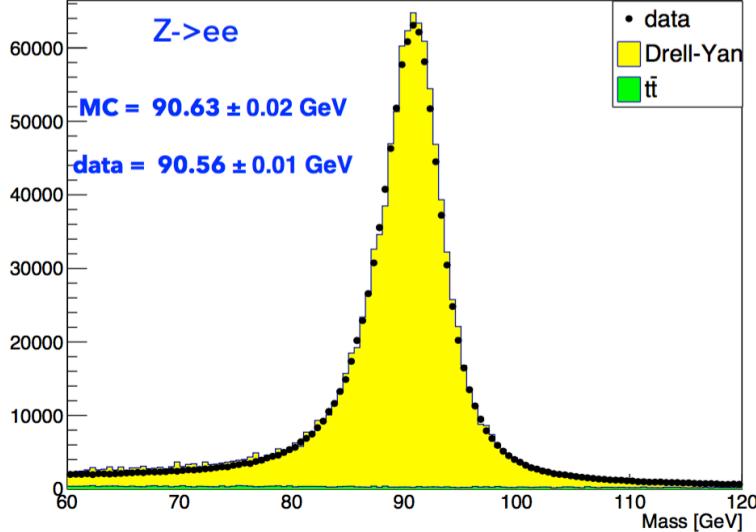
Tag electrons need to satisfy the following quality requirements: (1) trigger matched to `HLT_Ele27_eta2p1_WPTight_Gsf_v*` (2) $p_T > 30 \text{ GeV}$, super cluster (SC) $\eta < 2.1$ but on in EB-EE gap ($1.4442 < |\eta| < 1.566$) and (3) tight working point of the Spring16 cut-based electron ID.

Probe electrons only need to be reconstructed as `GsfElectron`, electrons associated with a `GsfTrack` object. The FSR recovery algorithm used in the main analysis is used consistently throughout the efficiency measurement; the isolations are calculated without any FSR photons matched to electrons and the probe electron p_T as well as the dielectron invariant mass include the FSR photons, if any.

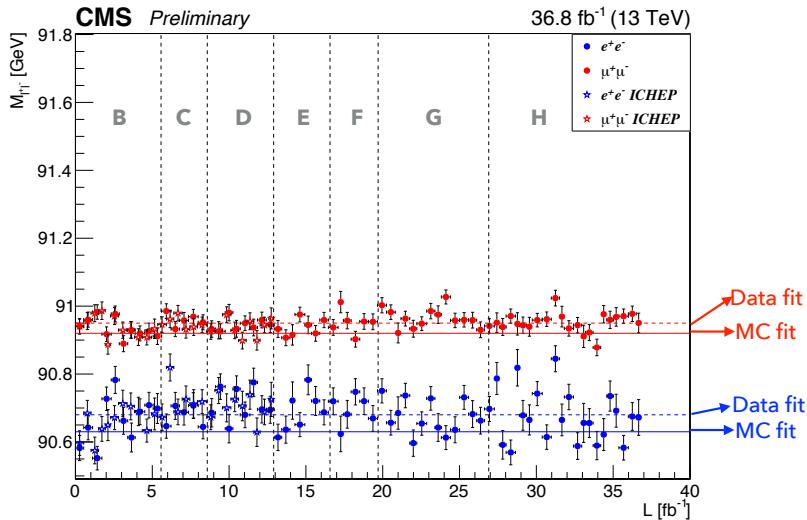
The nominal MC efficiencies are evaluated from the LO MadGraph Drell-Yan sample, while the NLO systematics use the 0,1 jet MadGraph_AMCatNLO sample listed in Table 3.6.

In contrast to previous efficiency measurements, a template fit is used here. The m_{ee} signal shape of the passing and failing probes is taken from MC and convoluted with a Gaussian. The data is then fitted with the convoluted MC template and a CMSShape, an error function with a one-sided exponential tail. This change follows from the usage of the new T&P tool developed by the EGM POG.

The electron selection efficiency is measured as a function of the probe electron p_T and its SC η , and separately for electrons falling in the ECAL gaps. Figure 4.5 shows the p_T turn-on curves measured in data, and the final two-dimensional scale factor is shown in Figure 4.6 together with the systematic uncertainties. These scale factors are very similar



(a)



(b)

Figure 4.4: (a): Electron energy scale measured in the $Z + \ell$ control region for EB and EE electrons. The results of the Crystall-ball fit are reported in the figure. (b): lepton energy scales per 500 pb luminosity block.

to the ICHEP figures, except more homogenous across η and p_T because of the higher statistics and the usage of more stable fitting routines in the new T&P tool.

The EGM recommendations on the evaluation of T&P uncertainties for efficiency measurements are followed. Specifically: (1) Variation of the signal shape from a MC

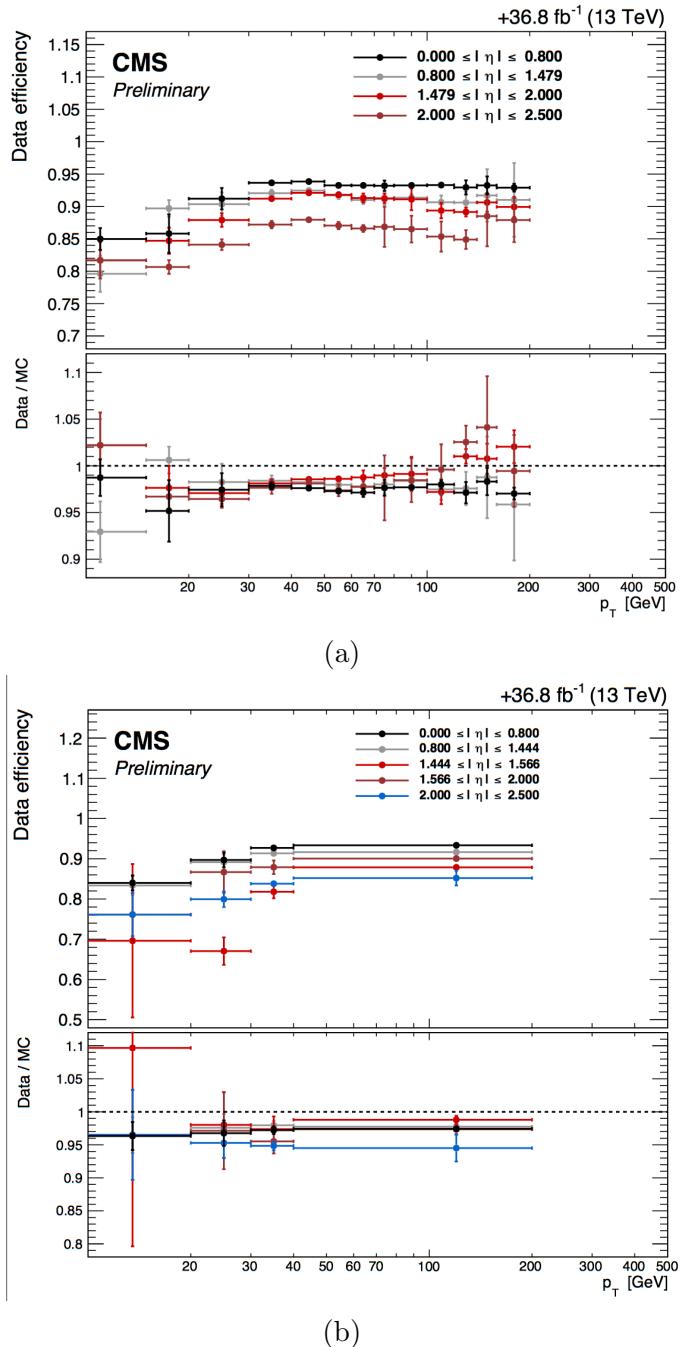
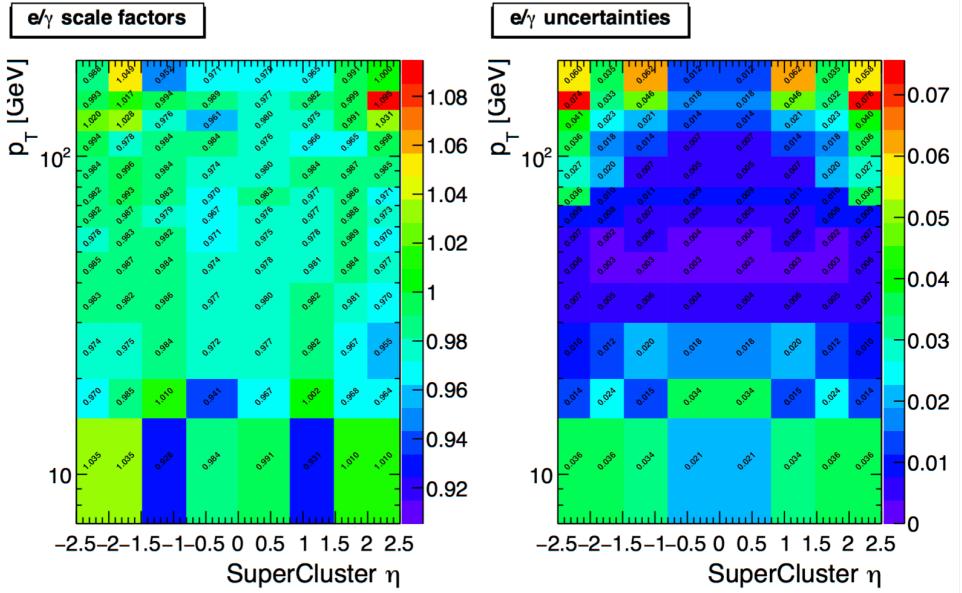
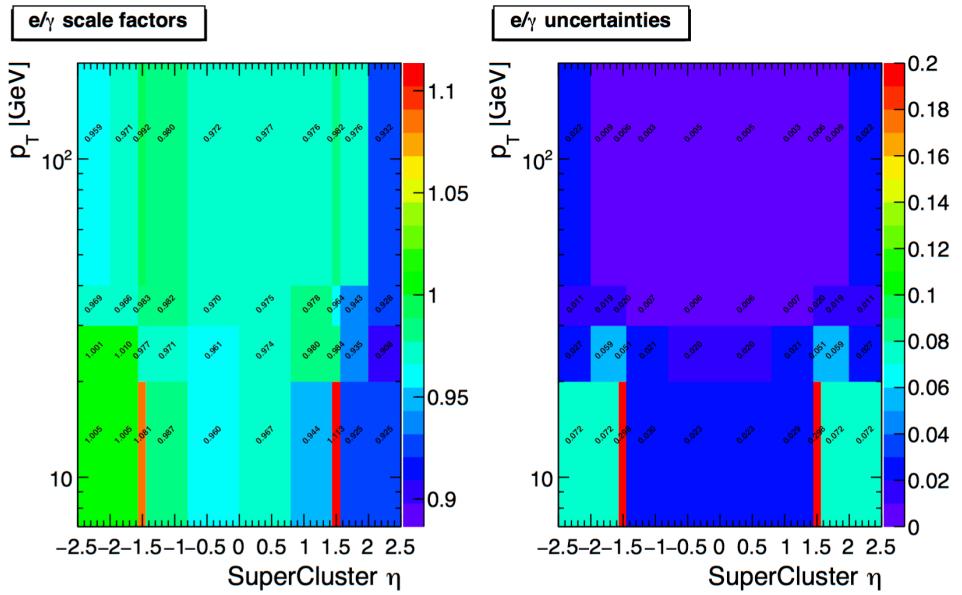


Figure 4.5: Electron selection efficiencies measured using the tag-and-probe technique described in the text, non-gap electrons (a) and gap electrons (b).

shape to an analytic shape (Crystal-Ball) fitted to the MC (2) Variation of the background shape from a CMS-shape to a simple exponential in fits to data (3) Variation of the tag selection: tag $p_T > 35$ GeV and passes MVA-based ID, and (4) Using an NLO MC



(a)



(b)

Figure 4.6: Electron selection efficiencies measured using the tag-and-probe technique described in the text, non-gap electrons (a) and gap electrons (b)

sample for the signal templates. The total uncertainty for the measurement of the scale factors is the quadratic sum of the statistical uncertainties returned from the fit and the aforementioned systematic uncertainties.

4.2 Muons

4.2.1 Muon reconstruction and identification

More details on muon reconstruction can be found in Ref. [158]. "Loose" muons are the muons that satisfy $p_T > 5$, $|\eta| < 2.4$, $d_{xy} < 0.5$, and $d_z < 1$, where d_{xy} and d_z are defined with respect to the primary vertex and using the muonBestTrack. Muons have to be reconstructed by either the Global Muon or Tracker Muon algorithm. Standalone muon tracks that are only reconstructed in the muon system are rejected. Standalone muons are discarded even if they are marked as global or tracker muons.

Loose muons with p_T below 200 GeV are considered "tight" muons if they also pass the PF muon ID. Note that the naming convention used for these IDs differs from the muon POG naming scheme, in which the "tight ID" used here is called the "loose ID". Loose muons with p_T above 200 GeV are considered tight muons if they pass the PF ID or the Tracker High- p_T ID, the definition of which is shown in Table 4.3. This relaxed definition is used to increase signal efficiency for the high-mass search. When a very heavy resonance decays to two Z bosons, both bosons will be very boosted. In the laboratory frame, the leptons coming from the decay of a highly boosted Z will be nearly collinear, and the PF ID loses efficiency for muons separated by approximately $\Delta R < 0.4$, which roughly corresponds to muons originating from Z bosons with $p_T > 500$ GeV.

Plain-text description	Technical description
Muon station matching	Muon is matched to segments in at least two muon stations
Good p_T measurement	$\frac{p_T}{\sigma_{p_T}} < 0.3$
Vertex compatibility ($x - y$)	$d_{xy} < 2$ mm
Vertex compatibility (z)	$d_z < 5$ mm
Pixel hits	At least one pixel hit
Tracker hits	Hits in at least six tracker layers

Table 4.3: The requirements for a muon to pass the Tracker High- p_T ID. Note that these are equivalent to the Muon POG High- p_T ID with the global track requirements removed.

An additional “ghost-cleaning” step is performed to deal with situations when a single muon can be incorrectly reconstructed as two or more muons. In this step, Tracker Muons that are not Global Muons are required to be arbitrated, and if two muons are sharing 50% or more of their segments, then the muon with lower quality is removed.

4.2.2 Muon isolation

PF-based isolation, described for electrons in Section 4.1.3, is also used for the muons. The only difference is the way the pileup contribution is subtracted; for the muons, $\Delta\beta$ correction is applied, whereby $\Delta\beta = \frac{1}{2} \sum_{\text{PU}}^{\text{charged had.}} p_T$ gives an estimate of the energy deposit of neutral particles (i.e. hadrons and photons) from pileup vertices. The relative isolation for muons is then defined as:

$$\text{RelPFiso} = \frac{\sum^{\text{charged had.}} p_T + \max(\sum^{\text{neutral had.}} E_T + \sum^{\text{photon}} E_T - \Delta\beta, 0)}{p_T^{\text{lepton}}} \quad (4.4)$$

The isolation working point for muons was optimized and chosen to be the same as for electrons, $\text{RelPFiso}(\Delta R = 0.3) < 0.35$ [158].

4.2.3 Muon efficiency measurements

Muon efficiencies are measured with the T&P method performed on $Z \rightarrow \mu\mu$ and $J/\Psi \rightarrow \mu\mu$ events in bins of p_T and η [158]. The Z sample is used to measure the muon reconstruction and identification efficiency at high p_T and the efficiency of the isolation and impact parameter requirements at all p_T . The J/Ψ sample is used to measure the reconstruction efficiency at low p_T , as it benefits from a better purity in that kinematic regime. In this case, events are collected using `HLT_Mu7p5_Track2_Jpsi_v*` when probing the reconstruction and identification efficiency in the muon system and using the `HLT_Mu7p5_L2Mu2_Jpsi_v*` when probing the tracking efficiency.

4.2.3.1 Reconstruction and identification

Results for the muon reconstruction and identification efficiency for $p_T > 20$ GeV have been derived by the Muon POG. The probe in this measurement are tracks reconstructed in the inner tracker, and the passing probes are those that are also reconstructed as a global or tracker muon and passing the Muon POG Loose muon identification. Results for low p_T muons were derived using J/Ψ events, with the same definitions of probe and

passing probes. The systematic uncertainties are estimated by varying the analytical signal and background shape models used to fit the dimuon invariant mass [158]. The efficiency and scale factors used for low p_T muons are the ones derived using single muon prompt-reco dataset. The efficiency in data and simulation is shown in Figure 4.7.

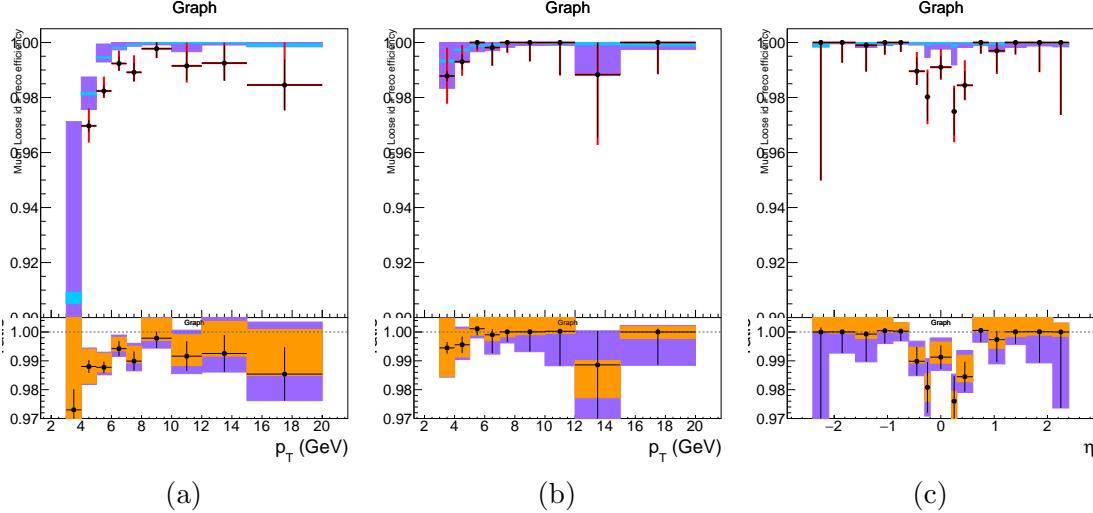


Figure 4.7: Muon reconstruction and identification efficiency at low p_T , measured with the tag-and-probe method on J/Ψ events, as function of p_T in the barrel (a) and end caps (b), and as function of η for $p_T > 7$ GeV (c). In the upper panel of each graph, the larger error bars include also the systematical uncertainties, while the smaller ones are purely statistical. Each graph’s lower panel shows the ratio of the two efficiencies, the black error bars are for the statistical uncertainty, the orange rectangles for the systematic uncertainty, and the violet rectangles include both uncertainties.

4.2.3.2 Impact parameter requirements

The measurement is performed using Z events. Events are selected with HLT_IsoMu20_v* or HLT_IsoMu22_v* triggers. For this measurement, the probe is a muon passing the POG Loose identification criteria, and it is considered a passing probe if it satisfies the SIP3D, d_{xy} , d_z cuts of this analysis. The results are shown in Figure 4.8.

4.2.3.3 Isolation requirements

The isolation efficiency is measured using events from the Z decay for any p_T , selected with either of HLT_IsoMu20_v* or HLT_IsoMu22_v* triggers. The isolation of the muons is calculated after recovery of the FSR photons and subtracting their contribution to the isolation cone of the muons [159]. The results are shown in Figure 4.9.

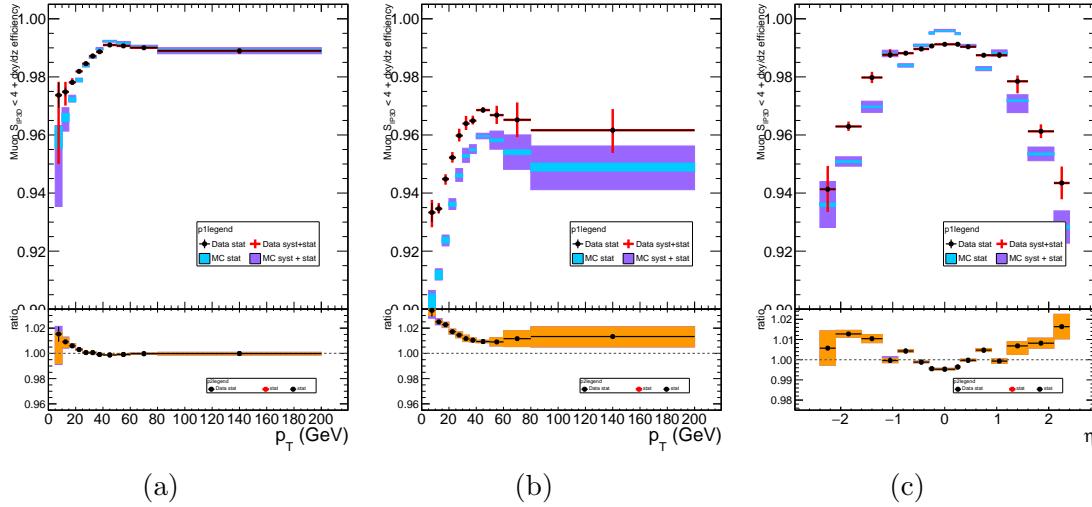


Figure 4.8: Efficiency of the muon impact parameter requirements, measured with the tag-and-probe method on Z events, as function of p_T in the barrel (a) and end caps (b), and as function of η for $p_T > 20$ GeV (c). In the upper panel of each graph, the larger error bars include also the systematical uncertainties, while the smaller ones are purely statistical. Each graph's lower panel shows the ratio of the two efficiencies, the black error bars are for the statistical uncertainty, the orange rectangles for the systematical uncertainty, and the violet rectangles include both uncertainties.

4.2.3.4 Tracking

The efficiency to reconstruct a muon track in the inner detector is measured using probes tracks reconstructed in the muon system [160]. The efficiency and data-to-MC scale factors are measured from Z events as a function of η for $p_T > 10$ GeV and $p_T < 10$ GeV. The values of data-to-MC scale factors used are from the ReReco version of the full dataset collected in 2016. The tracking efficiency in data and simulation as a function of η is shown in Figure 4.10.

4.2.3.5 Overall results

The product of all the data-to-MC scale factors for muon tracking, reconstruction, identification, impact parameter, and isolation requirements is shown in Figure 4.11.

4.3 Photons for FSR recovery

The FSR recovery algorithm was considerably simplified with respect to what was done in Run 1, while maintaining similar performance. The selection of FSR photons is now only

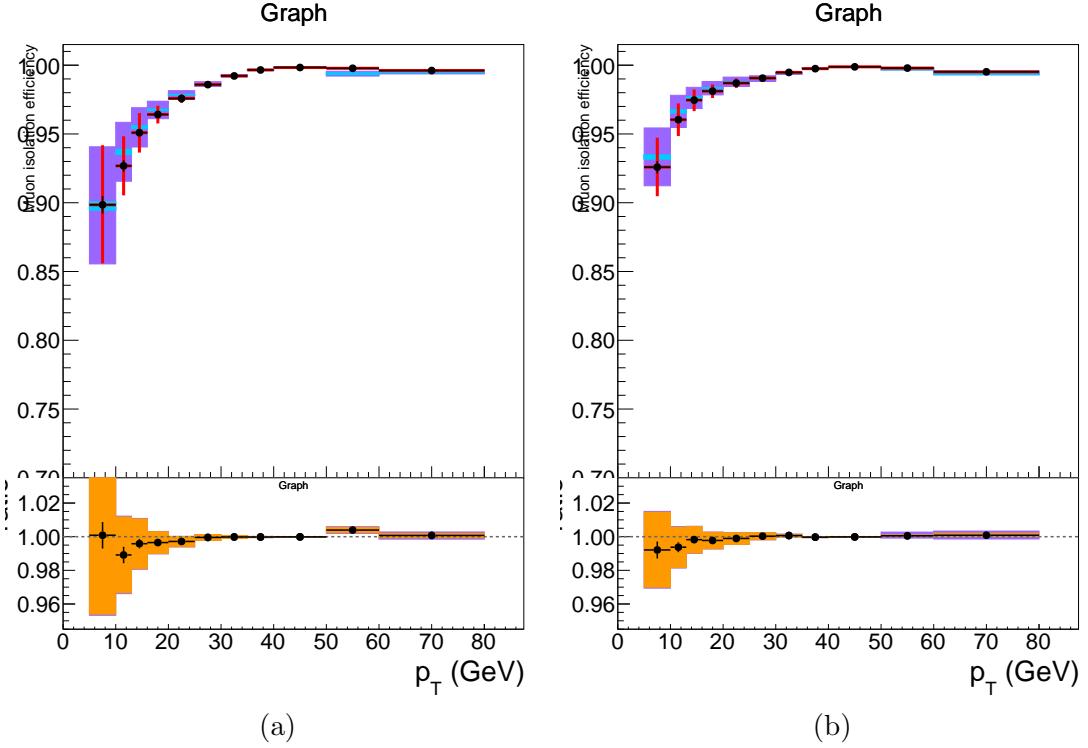
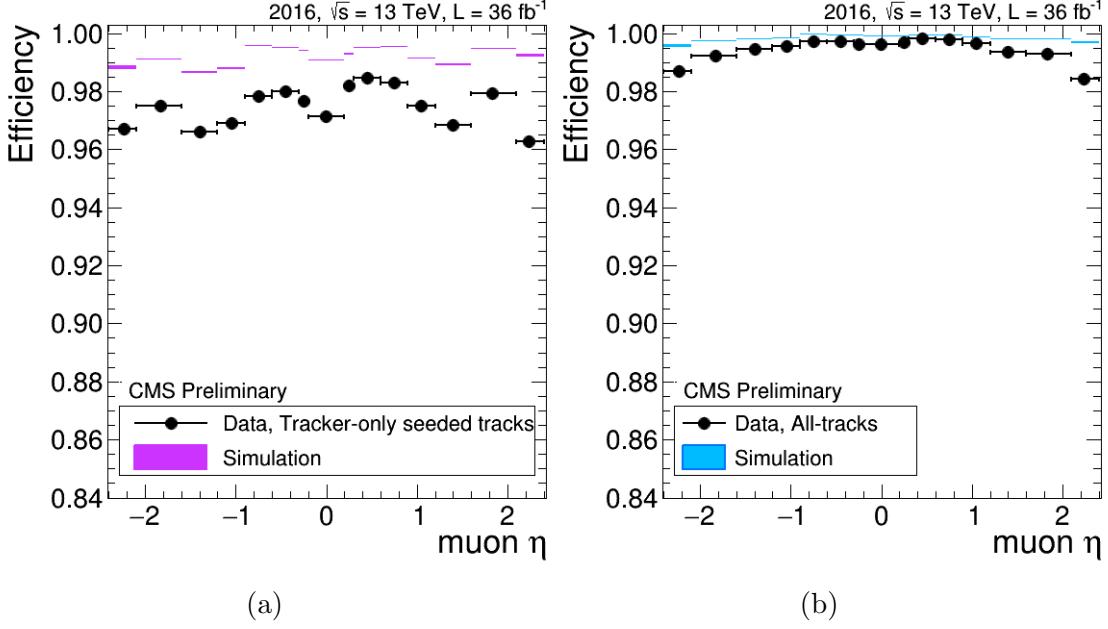


Figure 4.9: Efficiency of the muon isolation requirement, measured with the tag-and-probe method on Z events, as function of p_T in the barrel (left) and end caps (right). In the upper panel of each graph, the larger error bars include also the systematical uncertainties, while the smaller ones are purely statistical. The lower panel of each graph shows the ratio of the two efficiencies, the black error bars are for the statistical uncertainty, the orange rectangles for the systematical uncertainty, and the violet rectangles include both uncertainties.

done per-lepton and no longer depends on any Z mass criterion, simplifying the subsequent ZZ candidate building and selection. For the association of photons with leptons, the rectangular cuts on $\Delta R(\gamma, l)$ and $E_{T,\gamma}$ have been replaced by a cut on $\Delta R(\gamma, l)/E_{T,\gamma}^2$.

Starting from the collection of PF photons provided by the PF algorithm, the selection of photons and their association to a lepton proceeds as follows [158, 159]:

1. The preselection of PF photons is done by requiring $p_{T,\gamma} > 2$ GeV, $|\eta^\gamma| < 2.4$, and a relative PF isolation < 1.8 . The isolation is computed using a cone of radius $R = 0.3$, a threshold of 0.2 GeV on charged hadrons with a veto cone of 0.0001, and 0.5 GeV on neutral hadrons and photons with a veto cone of 0.01, also including the contribution from pileup vertices (with the same radius and threshold as per



(a)

(b)

Figure 4.10: Tracking efficiency in data and simulation as a function of η for muon $p_T < 10$ GeV(a) and $p_T > 10$ GeV(b) with ReReco data.

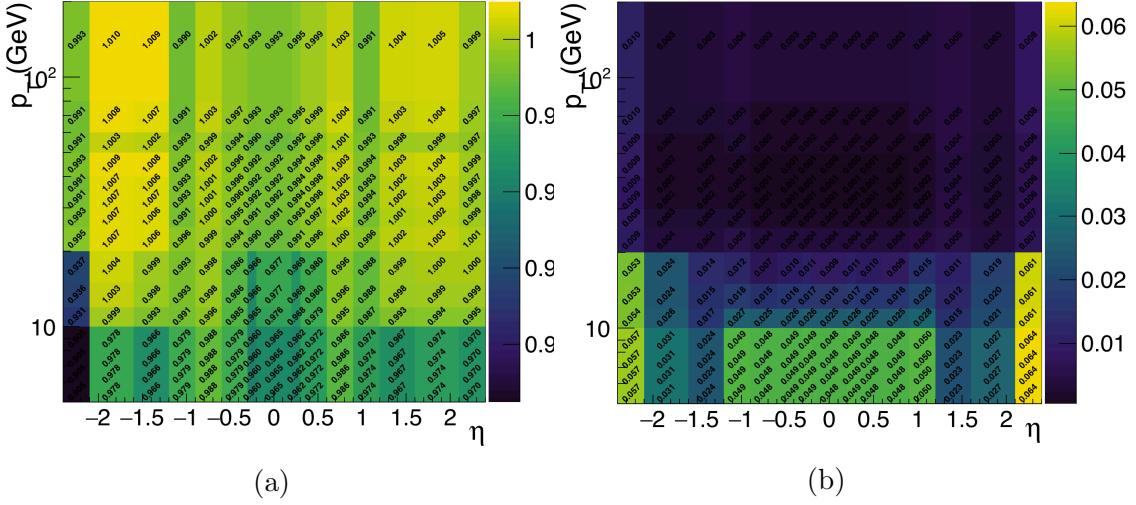


Figure 4.11: (a) Overall data to simulation scale factors for muons, as function of p_T and η . (b) Uncertainties on data to simulation scale factors for muons, as function of p_T and η .

charged isolation).

2. Supercluster veto: all PF photons that match with any electron passing both the loose ID and SIP cuts are removed. The matching is performed by directly associating the two PF candidates.

3. Photons are associated to the closest lepton in the event among all those that pass both the loose ID and SIP cuts.
4. Photons that fail the cuts $\Delta R(\gamma, l)/E_{T,\gamma}^2 < 0.012$, and $\Delta R(\gamma, l) < 0.5$ are discarded.
5. If more than one photon is associated to the same lepton, the lowest- $\Delta R(\gamma, l)/E_{T,\gamma}^2$ is selected.
6. For each FSR photon that was selected, photons from the isolation sum of all the leptons in the event that pass both the loose ID and SIP cuts are excluded. This concerns the photons that are in the isolation cone and outside the isolation veto of said leptons ($\Delta R < 0.4$ AND $\Delta R > 0.01$ for muons and $\Delta R < 0.4$ AND ($\eta^{SC} < 1.479$ OR $\Delta R > 0.08$) for electrons).

4.4 Jets

VBF and other H production mechanisms normally have different jet kinematics. Therefore, jets can be used to categorize events based on the production mechanisms.

4.4.0.6 Jet identification

Jets are reconstructed by using the anti- k_T clustering algorithm out of PF candidates, with a distance parameter $R = 0.4$, after rejecting the charged hadrons that are associated to a pileup primary vertex.

To reduce instrumental background, the loose working point jet ID suggested by the JetMET Physics Object Group is applied. In this analysis, the jets are required to be within $|\eta| < 4.7$ area and have a transverse momentum above 30 GeV. In addition, the jets are cleaned from any of the tight leptons (passing the SIP and isolation cut computed after FSR correction) and FSR photons by a separation criterion: $\Delta R(\text{jet}, \text{lepton}/\text{photon}) > 0.4$.

4.4.0.7 Jet energy corrections

The calorimeter response to particles is not linear, and it is not straightforward to translate the measured jet energy to the true particle or parton energy; therefore, jet energy

corrections (JECs) are needed. In this analysis, standard jet energy corrections are applied to the reconstructed jets, which consist of L1 Pileup, L2 Relative Jet Correction, L3 Absolute Jet Correction for both MC samples and data, and also residual calibration for data.

4.4.0.8 B-tagging

For categorization purposes, whether or not a jet is a b-jet needs to be distinguished. The Combined Secondary Vertex (CSV) algorithm is used as the b-tagging algorithm. It combines information about impact parameter significance, the secondary vertex and jet kinematics. The variables are combined using a likelihood ratio technique to compute the b-tag discriminator. In this analysis, a jet is considered to be b-tagged if it passes the *CSVv2M* working point, i.e. if its `pfCombinedInclusiveSecondaryVertexV2BJetTags` discriminator is greater than 0.8484 [161].

Data-to-MC scale factors for b-tagging efficiency are provided for this working point for the full dataset as a function of jet p_T , η , and flavor. They are applied to simulated jets by downgrading (upgrading) the b-tagging status of a fraction of the b-tagged (untagged) jets that have a scale factor smaller (larger) than one.

4.5 MET

The missing transverse energy, E_T^{MISS} or MET, of an event is defined as the magnitude of the imbalance of momentum in the plane transverse to the beam line. Since momentum is conserved in this plane, any imbalance in momentum is attributed to particles escaping the detector without interacting with the detector material, such as neutrinos or hypothetical dark matter candidates. Raw MET or particle flow MET (PFMET) is defined as the magnitude of the negative vectorial sum of the transverse momentum of all reconstructed particle flow candidates, or

$$\vec{E}_T^{\text{MISS}} = - \sum_{i \in \text{all}} \vec{p}_{T,i} \quad (4.5)$$

The vector quantity that is the negative sum of reconstructed particle momenta is sometimes called the missing transverse momentum, although this term is used inter-

changably with its magnitude, the MET.

An alternative definition of the MET, called the type-1 corrected MET, takes into account the jet energy corrections (JEC), correcting for mismeasurement of MET due to detector inefficiencies and non-linear responses in the calorimeters. The type-1 corrected MET definition is given by:

$$\vec{E}_T^{\text{MISS}}_{\text{Type-1}} = - \sum_{\text{jet}} \vec{p}_{T, \text{jet}}^{\text{JEC}} - \sum_{i \in \text{uncl.}} \vec{p}_{T, i} \quad (4.6)$$

where the total contribution has been split into contributions from jets (first term) and contributions from unclustered objects (second term). The transverse momenta of jets in the first term is then replaced with the JEC transverse momenta.

Systematic uncertainties related to modeling real MET are obtained by varying the JEC and jet energy resolution (JER) and measuring the propagation of these variations to the MET uncertainty. These measurements are described in greater detail in Section 6.1.2.

4.5.0.9 MET filters

Due to detector and instrumental noise, several filters are applied to veto noisy events [162]: HBHENoiseFilter, HBHENoiseIsoFilter, EcalDeadCellTriggerPrimitiveFilter, goodVertices, eeBadScFilter, globalTightHalo2016Filter, BadPFMuonFilter, and BadChargedCandidateFilter.

The first two filters remove noisy events from the HCAL, where the HBHE scintillator produce anomalous signals with pulse shapes and pixel multiplicities discrepant from those from a clean signal. The EcalDeadCellTriggerPrimitiveFilter removes events with down ECAL data links, comparing the sum of energy deposited in each cell of a supercluster to the trigger primitive saturation energy. The goodVertices filter removes events with noisy vertex reconstruction from pileup effects. The eeBadScFilter removes events with noisy ECAL end cap superclusters. The globalTightHalo2016Filter filter removes events with enhanced MET from beam-halo particles which are in time with the beam. The last two filters remove events with mis-reconstructed muon and charged hadron particle flow candidates.

4.5.0.10 Fake MET modeling

Figure 4.12 shows a discrepancy between data and MC in the high- E_T^{MISS} tail. These events typically contain a high p_T object either back-to-back or collinear with the E_T^{MISS} , pointing to artificially high E_T^{MISS} from mismeasurement of the object. These fake events are identified and removed by studying distributions of the transverse angular difference between the MET and various objects in the event [107].

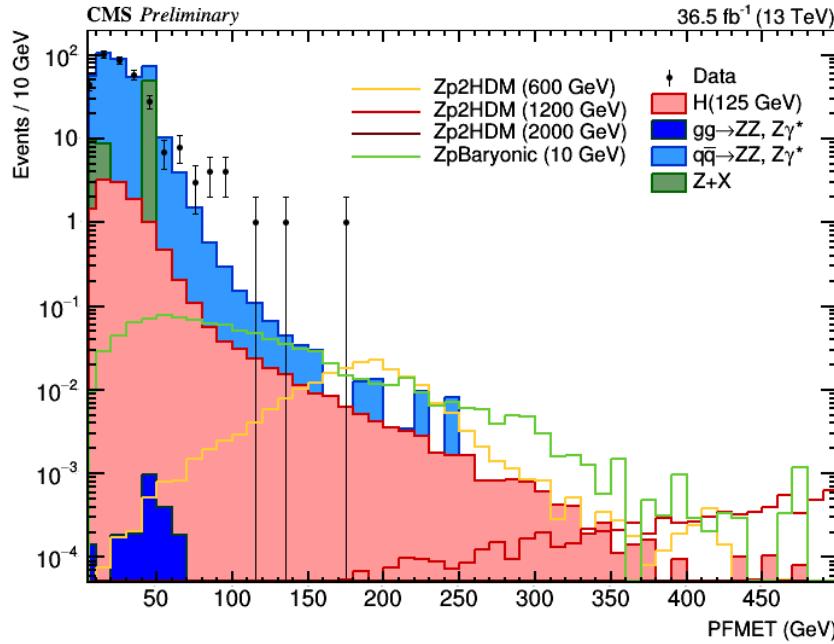


Figure 4.12: Missing transverse energy (MET) after the standard model selection before removal of fake MET.

The following variables are studied in order to understand how to remove fake events from data: (1) $\max|\Delta\phi(\text{jet}, E_T^{\text{MISS}})|$ with the maximum taken over selected jets, Figure 4.13 and (2) $\min|\Delta\phi(\text{jet}, E_T^{\text{MISS}})|$ with the minimum taken over selected jets, Figure 4.14. The maximum variable is to check for the occurrence of objects with mismeasured energy back-to-back with the MET, while the minimum variable is to check for jets with mismeasured energy collinear with the MET.

For jets with a high transverse momentum, $> 50 \text{ GeV}$, it is required that $\max|\Delta\phi(\text{jet}, E_T^{\text{MISS}})| < 2.7$ and $\min|\Delta\phi(\text{jet}, E_T^{\text{MISS}})| > 0.5$ to exclude events with large MET from mismeasurement of jet energies. These cuts are based on the Run 1 SM analysis selection, chosen to

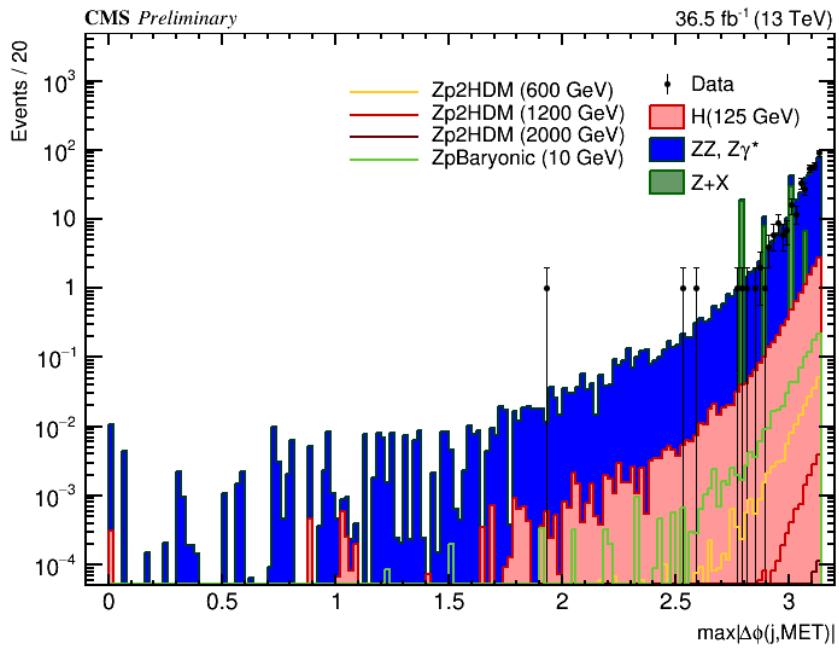


Figure 4.13: Maximum azimuthal angle difference between MET and jets.

balance the small loss in signal efficiency with the increased systematic uncertainty from mismodeling of the MET in background MC simulations (described in Section 6.1.2).

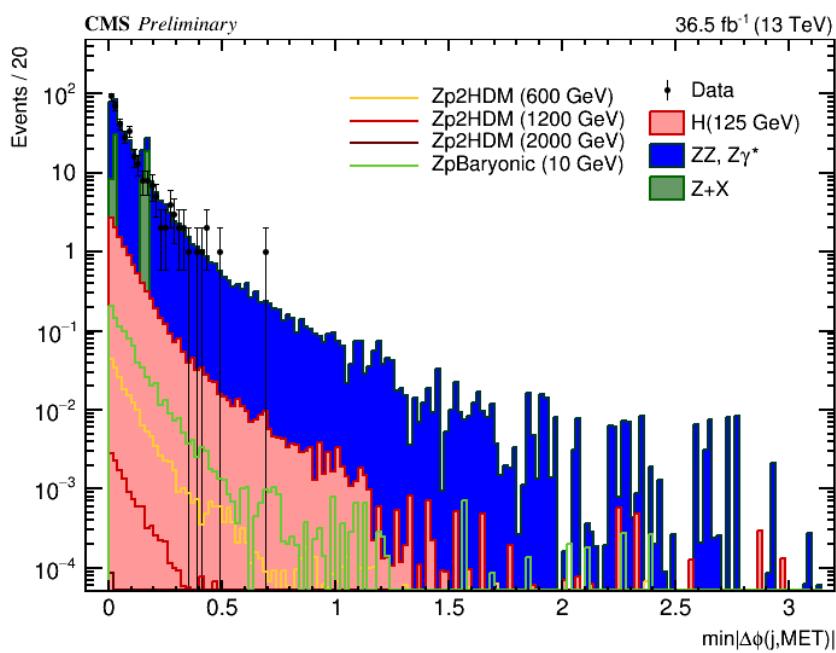


Figure 4.14: Minimum azimuthal angle difference between MET and jets.

Chapter 5

Signal and control regions

5.1 Event selection

5.1.1 Trigger selection

The events are required to have fired the High-Level Trigger (HLT) paths described in section 3.3.1.1. Unlike in the Run 1 analysis, the trigger requirement in the Run 2 analysis does not depend on the selected final state. Instead, it is always the OR of all ten HLT paths, since associated production modes that can have additional leptons will be targeted. This improves the trigger efficiency further.

5.1.2 Vertex selection

The events are required to have at least one good primary vertex (PV) fulfilling all three of the following criteria: a high number of degrees of freedom ($N_{PV} > 4$), collisions restricted along the z -axis ($z_{PV} < 24$ cm), and a small radius of the PV ($r_{PV} < 2$ cm).

5.1.3 ZZ candidate selection

The four-lepton candidates are built from what we call "selected" leptons, which are the tight leptons (defined in Sections 4.1.2 and 4.2.1) that pass the $SIP_{3D} < 4$ vertex constraint and the isolation cuts (defined in Sections 4.1.3 and 4.2.2), where FSR photons are subtracted as described in Section 4.3. A lepton cross-cleaning is applied by discarding electrons that are within $\Delta R < 0.05$ of selected muons.

The construction and selection of four-lepton candidates proceeds according to the

following sequence:

1. **Z candidates** are defined as pairs of selected leptons of opposite charge and matching flavor (e.g. e^+e^- or $\mu^+\mu^-$) that satisfy $12 < m_{\ell\ell(\gamma)} < 120$ GeV, where the Z candidate mass includes the selected FSR photons if any.
2. **ZZ candidates** are defined as pairs of non-overlapping Z candidates. The Z candidate with reconstructed mass $m_{\ell\ell}$ closest to the nominal Z boson mass is denoted as Z_1 , and the second one is denoted as Z_2 . The dimuon distributions for Z_1 and Z_2 are shown in Figure 5.1 and Figure 5.2, respectively. The dielectron distributions are qualitatively the same.

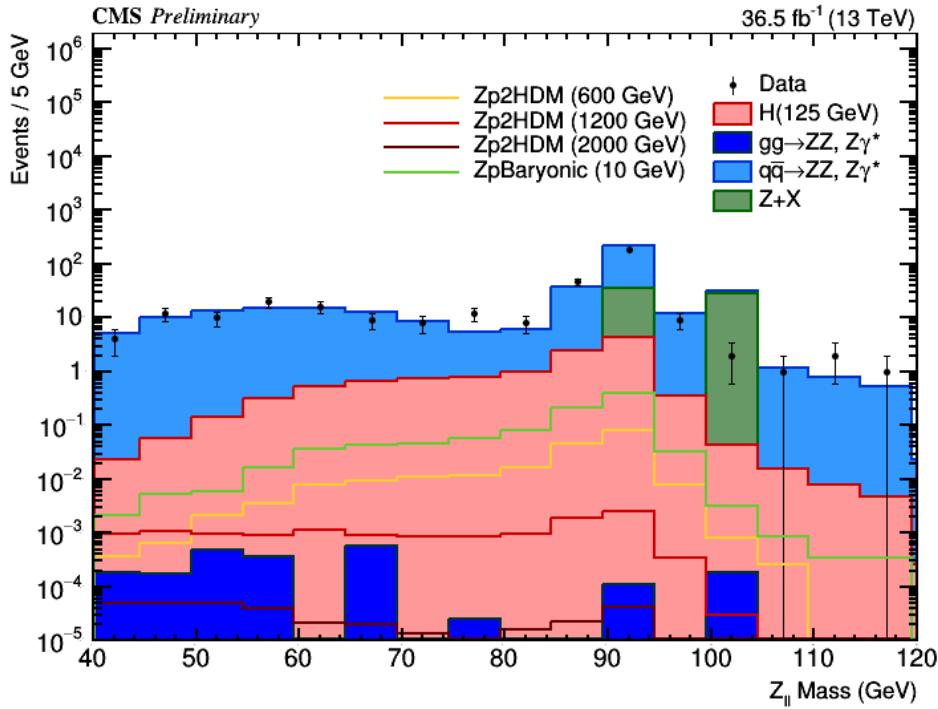


Figure 5.1: Dimuon invariant mass distributions for Z_1 .

ZZ candidates are required to satisfy the following list of requirements:

- **Ghost removal** : $\Delta R(\eta, \phi) > 0.02$ between each of the four leptons
- **Lepton p_T** : two of the four selected leptons should pass $p_{T,i} > 20$ GeV and $p_{T,j} > 10$ GeV

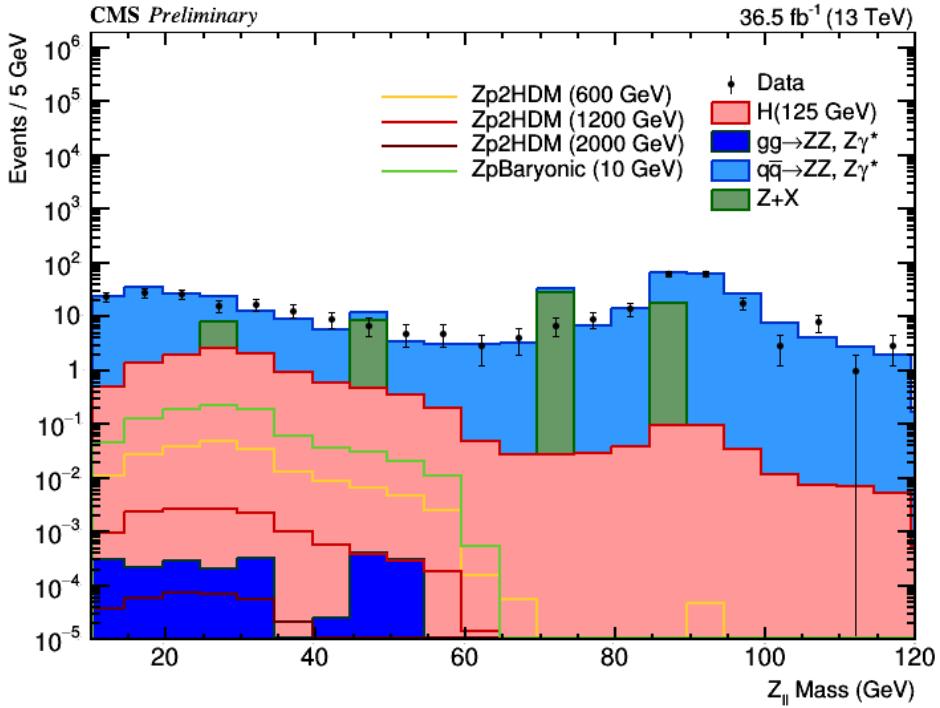


Figure 5.2: Dimuon invariant mass distributions for Z_2 .

- **QCD suppression:** all four oppositely-signed pairs that can be built with the four leptons (regardless of lepton flavor) must satisfy $m_{\ell\ell} > 4$ GeV. Here, selected FSR photons are not used in computing $m_{\ell\ell}$, since a QCD-induced low-mass dilepton (e.g. J/Ψ) may have photons nearby (e.g. from π_0)
- **Z_1 mass:** $m_{Z1} > 40$ GeV
- ‘smart cut’: defining Z_a and Z_b as the mass-sorted alternative pairing Z candidates (Z_a being the one closest to the nominal Z boson mass), require $\text{NOT}(|m_{Za} - m_Z| < |m_{Z1} - m_Z| \text{ AND } m_{Zb} < 12)$. Selected FSR photons are included in m_Z ’s computations. This cut discards 4μ and $4e$ candidates where the alternative pairing looks like an on-shell $Z +$ low-mass $\ell^+\ell^-$. In Run 1, such a situation was avoided by choosing the best ZZ candidate before applying kinematic cuts to it, most precisely before the $m_{Z2} > 12$ GeV cut. The present smart cut allows to choose the best ZZ candidate after all kinematic cuts.
- **Four-lepton invariant mass:** $m_{4\ell} > 70$ GeV. The four-muon invariant mass

distribution after all of the previous selection steps are applied is shown in Figure 5.3, with a close up view of the Z and H peaks shown in Figure 5.4. The four-lepton invariant mass distributions for the other decay channels are qualitatively the same

3. Events containing at least one selected ZZ candidate form the SM signal region.

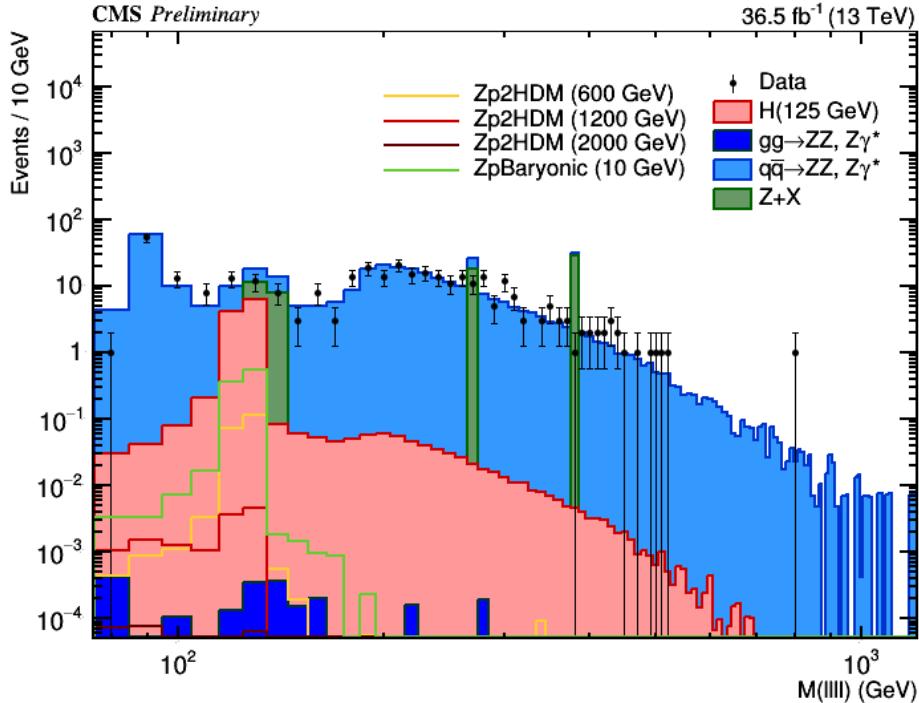


Figure 5.3: Four-muon invariant mass distribution in log scale.

5.1.4 Choice of the best ZZ candidate

Unlike in the Run 1 analysis, the best ZZ candidate in the Run 2 analysis is chosen after all kinematic cuts, a change that allows to test other selection strategies for this candidate choice. This is especially relevant for events with more than four selected leptons, such as VH and $t\bar{t}H$, with associated particles that can decay to additional leptons.

For the current analysis, if more than one ZZ candidate survives the above selection, we choose the one with the highest value of $\mathcal{D}_{\text{bkg}}^{\text{kin}}$, the MELA kinematic discriminant,

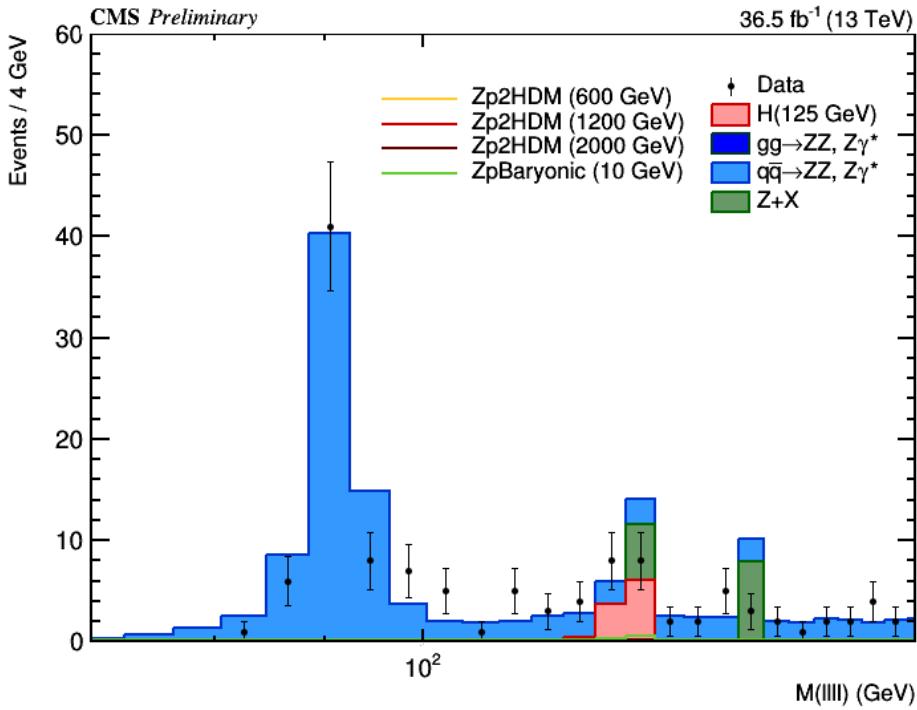


Figure 5.4: Four-muon invariant mass distribution in linear scale.

except if two candidates are composed of four leptons, in which case, the candidate with Z_1 closest in mass to nominal Z boson mass is chosen.

5.2 Signal region and blinding

A one-step SR and optimization process is used in the current analysis, where a near-optimal SR is defined using all of the desired variables, instead of the two-step procedure used in previous analyses. This simplification allows for the use of the same SR for all signal models, including the same variable validation and background model. The tradeoff is a small loss in sensitivity for models with tighter optimal cuts, but the benefit of simplicity outweighs the cost of this loss. In principle, each benchmark signal point can be optimized individually to gain a few percent in sensitivity, but this is left to a future study.

Two strategies are employed to define the signal region: cut-and-count based and multivariate analysis (MVA) based. Several key discriminating variables are studied as inputs

to both SR definitions: the missing transverse energy, E_T^{MISS} or MET; the four-lepton invariant mass, m_{4l} ; the transverse mass of the Higgs and MET, $m_T(4l + E_T^{\text{MISS}}) \equiv m_T$; the difference in the transverse angle ϕ between the Higgs and MET, $|\Delta\phi(4l, E_T^{\text{MISS}})| \equiv |\Delta\phi|$; as well as the lepton and jet multiplicities.

5.2.1 Cut-and-count based signal region

The cut-and-count strategy is the simpler choice and is used as a baseline to measure the performance of the MVA. First, the following selection is applied to isolate events with a Higgs from events with additional prompt particles (e.g. VBF Higgs production):

- Tight-lepton multiplicity = 4
- b-jet multiplicity ≤ 1
- VBF jet multiplicity ≤ 1 .

Next, the event selection is optimized by scanning over a range of cuts for the remaining variables and selecting the set of cuts that maximizes the sensitivity, measured directly by the cross section upper limit. The two variables with the greatest discriminating power between signal and background are m_{4l} and E_T^{MISS} , so we maximize the sensitivity over these two variables. The best sensitivity occurs where the upper limit is minimized, at approximately:

- $|m_{4l} - m_H| < 10 \text{ GeV}$
- $E_T^{\text{MISS}} > 60 \text{ GeV}$

These cuts define the SR, which is applied to all of the signals, losing less than 10% sensitivity for the models with a tighter optimal cut on E_T^{MISS} . Since the signal used to defined the SR has the softest E_T^{MISS} spectrum of all the signals, this SR corresponds to the most modest or, meaning no signal will be cut too hard, while most of the background is still removed.

Note that there is no additional blinding on the MET distribution above a certain threshold, as was the case previously in similar searches. This is due primarily to the need

to understand events that contribute large amounts of fake MET. This is covered thoroughly, including the procedure for removing these events from data, in Section 4.5.0.10. In order to validate the modeling in these SRs, they are split into control regions (CR) based on the MET being above or below 60 GeV, referred to as the high and low MET regions.

5.2.2 MVA-based signal region

This search channel has the advantage of having backgrounds that are easily reduced by applying cuts on the discriminating variables. It was observed in the study of the 2015 data and MC samples that applying additional cuts did not improve the sensitivity since the background levels were already sufficiently low. Applying cuts on additional variables reduced the signal efficiency, which in turn reduced the sensitivity. This is also observed with 2016 MC samples, where applying the $|\Delta\phi(l_{lll}, E_T^{\text{MISS}})|$ does not improve the sensitivity. These observations motivate the use of MVA techniques, which can take all of the desired variables as inputs, but do not reduce the signal efficiency. Although it is simpler to cut on these discriminating variables using the cut-and-count method, there is potential for significant improvement in the sensitivity with an MVA approach.

The SR event selection is optimized for the MVA-based case by training a boosted decision tree (BDT) with the ROOT TMVA package with m_{4l} and E_T^{MISS} the input variables. Including additional input features does not significantly improve the performance of the MVA. Training is done over the weighted set of simulated backgrounds and an admixture of signal models to reduce bias toward a single signal model. The MVA response is shown in Figure 5.5, with the signal peaked toward 1 and the background peaked toward -1 .

5.3 Background modeling

5.3.1 Irreducible backgrounds

5.3.1.1 $q\bar{q} \rightarrow ZZ$ modeling

The $q\bar{q} \rightarrow ZZ$ background is generated at NLO, while the fully differential cross section has been computed at NNLO [163] but is not yet available in a partonic level event generator. Therefore, NNLO/NLO k -factors for the $q\bar{q} \rightarrow ZZ$ background process are

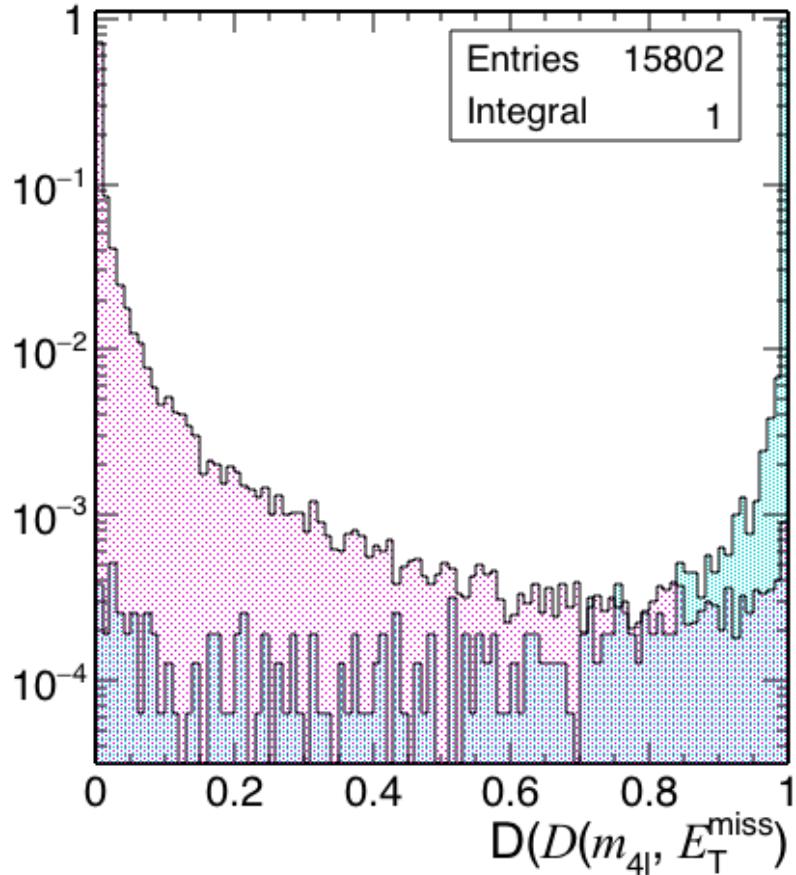


Figure 5.5: MVA response trained on a weighted background sample and admixture of signal samples with input variables m_{4l} and E_T^{MISS} . Simulated backgrounds are in pink and signals are in blue.

applied to the POWHEG sample. The inclusive cross sections obtained using the same PDF and renormalization and factorization scales as the POWHEG sample at LO, NLO, and NNLO are shown in Table 5.1. The NNLO/NLO k -factors are applied in the analysis differentially as a function of $m(ZZ)$.

Additional NLO electroweak corrections, which depend on the initial-state quark flavor and kinematics, are also applied to the $q\bar{q} \rightarrow ZZ$ background process in the region $m(ZZ) > 2m(Z)$. The differential QCD and electroweak k -factors can be seen in Figure 5.6.

QCD Order	$\sigma_{2\ell 2\ell'}(\text{fb})$	$\sigma_{4\ell}(\text{fb})$
LO	$218.5^{+16\%}_{-15\%}$	$98.4^{+13\%}_{-13\%}$
NLO	$290.7^{+5\%}_{-8\%}$	$129.5^{+4\%}_{-6\%}$
NNLO	$324.0^{+2\%}_{-3\%}$	$141.2^{+2\%}_{-2\%}$

Table 5.1: Cross sections for $q\bar{q} \rightarrow ZZ$ production at 13 TeV

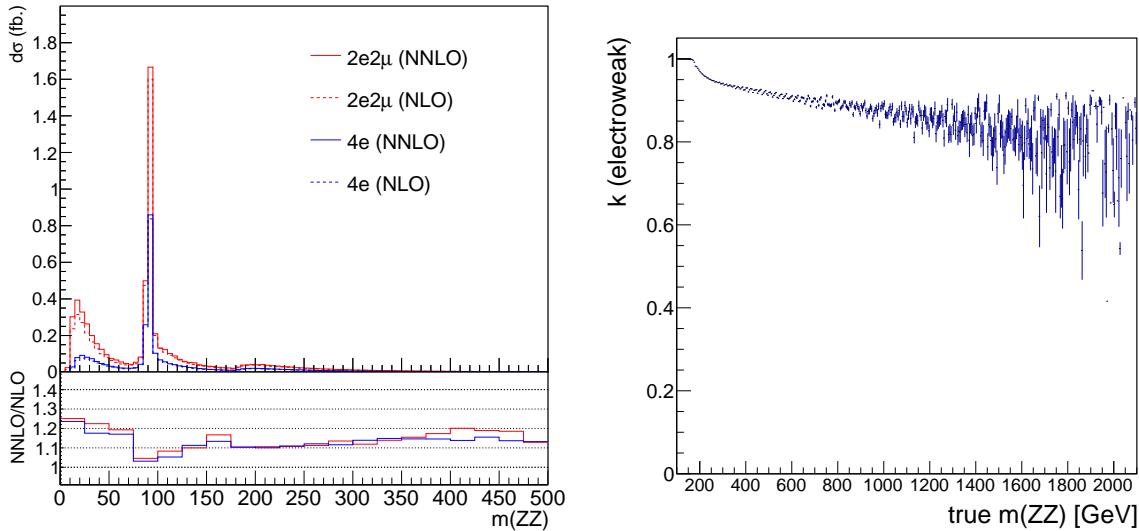


Figure 5.6: Left: NNLO/NLO QCD k -factor for the $q\bar{q} \rightarrow ZZ$ background as a function of $m(ZZ)$ for the 4ℓ and $2\ell 2\ell'$ final states. Right: NLO/NLO electroweak k -factor for the $q\bar{q} \rightarrow ZZ$ background as a function of $m(ZZ)$.

5.3.1.2 $gg \rightarrow ZZ$ modeling

Event simulation for the $gg \rightarrow ZZ$ background is done at LO with the generator MCFM 7.0 [154, 164, 155]. Although no exact calculation exists beyond LO for the $gg \rightarrow ZZ$ background, it has been recently shown that the soft collinear approximation is able to describe the background cross section and the interference term at NNLO [165]. Further calculations also show that the k -factors are very similar at NLO for the signal and background [166] and at NNLO for the signal and interference terms [167]. Therefore, the same k -factor is used for the signal and background [168]. The NNLO k -factor for the signal is obtained as a function of $m_{4\ell}$ using the HNNLO v2 MC program [169, 170, 171] by calculating the NNLO and LO $gg \rightarrow H \rightarrow 2\ell 2\ell'$ cross sections at the small H decay

width of 4.07 MeV and taking their ratios. The NNLO and NLO k -factors and the cross sections from which they are derived are illustrated in Figure 5.7, along with the NNLO, NLO and LO cross sections at the SM H boson decay width [172].

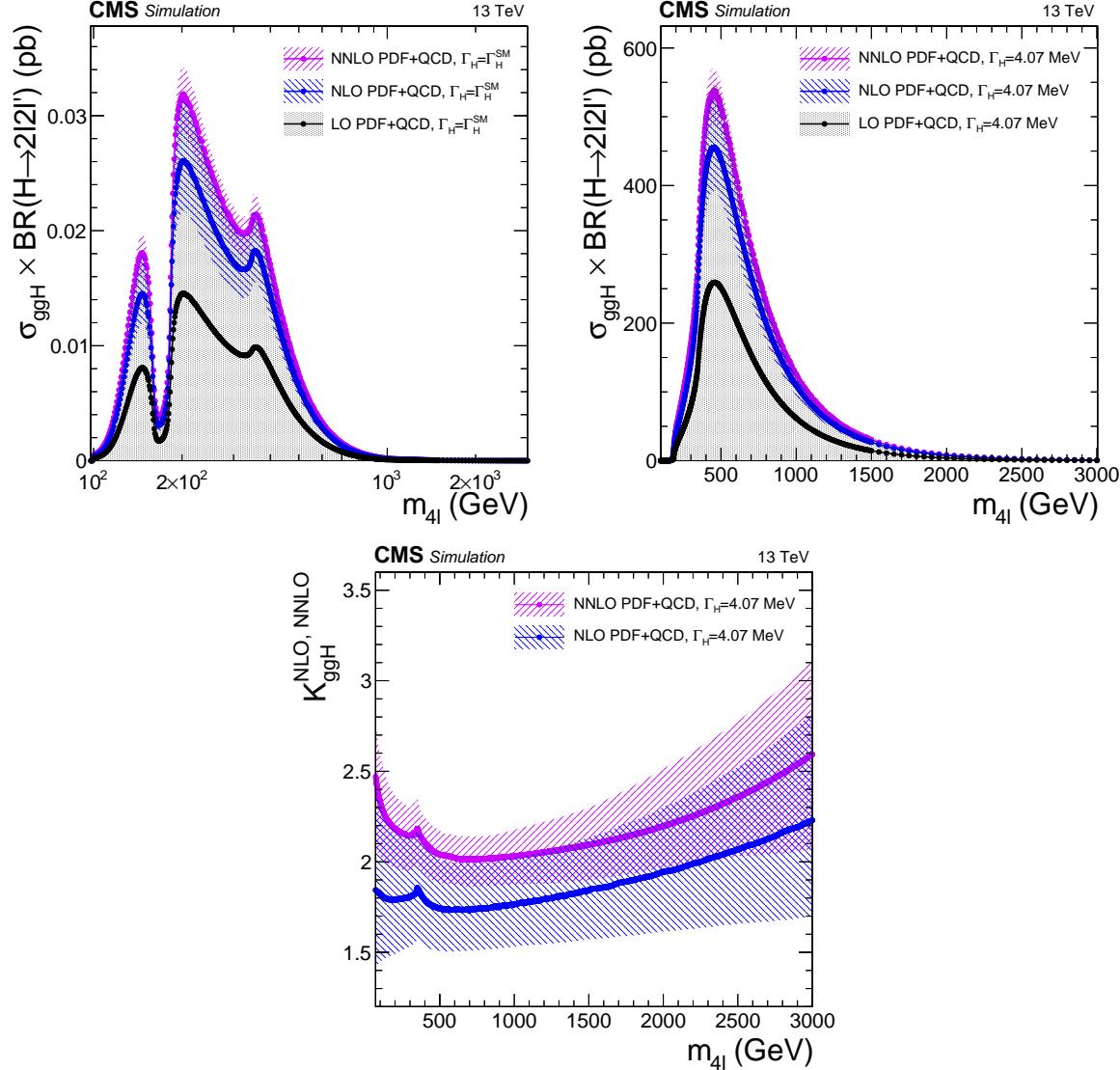


Figure 5.7: $gg \rightarrow H \rightarrow 2\ell 2\ell'$ cross sections at NNLO, NLO and LO at each H pole mass using the SM H decay width (top left) or at the fixed and small decay width of 4.07 MeV (top right). The cross sections using the fixed value are used to obtain the k -factor for both the signal and the continuum background contributions as a function of $m_{4\ell}$ (bottom).

5.3.2 Reducible background estimation

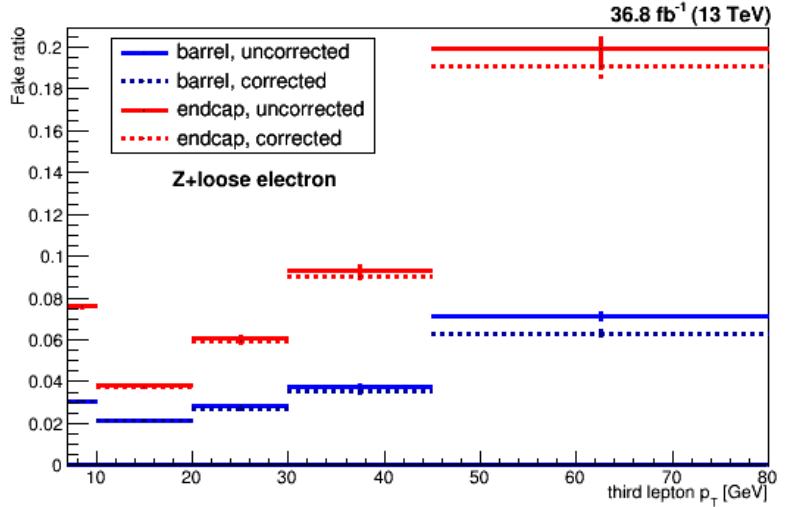
The reducible background for the $H \rightarrow ZZ \rightarrow 4\ell$ analysis, hereafter called $Z + X$, originates from processes which contain one or more non-prompt lepton in the four-lepton final state. The main sources of non-prompt leptons are non-isolated electrons and muons coming from decays of heavy-flavor mesons, mis-reconstructed jets (usually originating from light-flavor quarks), and electrons from γ -conversions. A fake lepton is defined as any jet mis-reconstructed as a lepton and any lepton originating from a heavy meson decay. Similarly, any electron originating from a photon conversion will be considered a fake electron.

The lepton fake rates, f_e and f_μ , are defined as the ratio of the number of electrons/muons passing the tight selection criteria to the number passing the loose criteria. This measures the probability of a lepton passing the loose criteria to also pass the tight criteria. The fake rates are applied in dedicated control samples in order to extract the expected background yield in the SR.

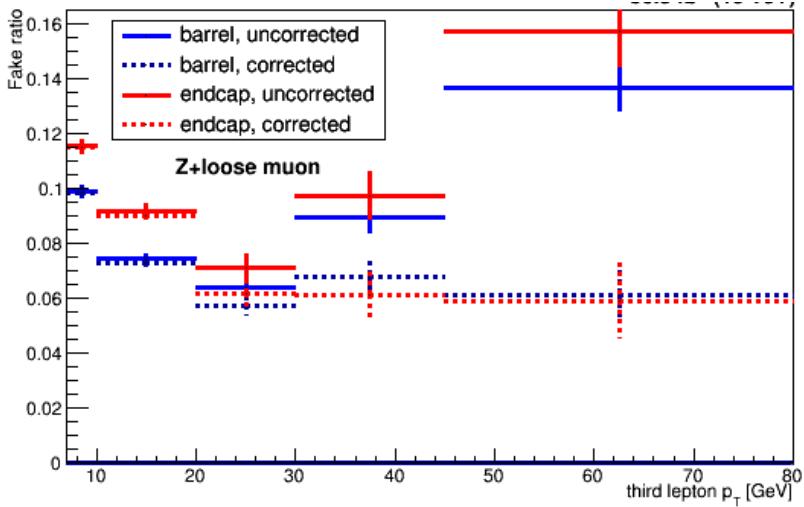
5.3.2.1 Fake rate determination

In order to measure the lepton fake rates f_e and f_μ , samples of $Z(\ell\ell) + e$ and $Z(\ell\ell) + \mu$ events are selected that are expected to be completely dominated by final states that include a Z boson and a fake lepton. These events are required to have two same-flavor, oppositely-charged leptons with $p_T > 20/10$ GeV passing the tight selection criteria, thus forming the Z candidate. In addition, the event must have exactly one lepton passing the loose selection criteria. This lepton is used as the probe lepton for the fake rate measurement. The invariant mass of the probe lepton and the opposite sign lepton from the reconstructed Z candidate must satisfy $m_{2l} > 4$ GeV.

The fake rates are evaluated using the tight requirement $|M_{inv}(\ell_1, \ell_2) - M_Z| < 7$ GeV to reduce the asymmetric contribution from photon conversions populating low masses and using $E_T^{\text{MISS}} < 25$ GeV to reduce contamination from QCD events. The fake rates measured in bins of the transverse momentum of the loose lepton in the barrel and end cap regions are shown in Figure 5.8.



(a)



(b)

Figure 5.8: Fake rates as a function of the probe p_T for electrons (a) and muons (b) which satisfy the loose selection criteria. The fake rates are measured in a $Z(\ell\ell) + \ell$ sample in the 13 TeV data. The barrel selection includes electrons (muons) up to $|\eta| = 1.479$ (1.2).

5.3.2.2 Fake rate application

Two control regions are defined as subsets of four-lepton events which pass the first step of the selection (see Section 5.1.3), requiring an additional pair of same-flavor, oppositely-charged loose leptons, that pass the SIP_{3D} cut. The events must satisfy all kinematic cuts applied for the Higgs phase space selection (see Section 5.1.3). The first control sample

is obtained by requiring that the two loose leptons that do not make the Z_1 candidate do not pass the final identification and isolation criteria. The other two leptons pass the final selection criteria by definition of the Z_1 . This sample is denoted as the “2 Prompt + 2 Fail” (2P+2F) sample. It is expected to be populated with events that intrinsically have only two prompt leptons, mostly DY , with a small fraction of $t\bar{t}$ and $Z\gamma$ events. The second control sample is obtained by requiring one of the four leptons not to pass the final identification and isolation criteria. The other three leptons should pass the final selection criteria. This control sample is denoted as “3 Prompt + 1 Fail” (3P+1F). It is expected to be populated with the type of events that populate the 2P+2F region, albeit with different relative proportions, as well as with WZ events that intrinsically have three prompt leptons.

The control samples obtained in this way, orthogonal by construction to the SR, are enriched with fake leptons and are used to estimate the reducible background in the SR.

The invariant mass distribution of events selected in the 2P+2F control sample is shown in Figure 5.9 for the 13 TeV dataset.

The expected number of reducible background events in the 3P+1F region, N_{3P1F}^{bkg} , can be computed from the number of events observed in the 2P+2F control region, N_{2P2F} , by weighting each event in the region with the factor $(\frac{f_i}{1-f_i} + \frac{f_j}{1-f_j})$, where f_i and f_j correspond to the fake rates of the two loose leptons:

$$N_{3P1F}^{\text{bkg}} = \sum \left(\frac{f_i}{1-f_i} + \frac{f_j}{1-f_j} \right) N_{2P2F} \quad (5.1)$$

Figure 5.10 shows the invariant mass distributions of the events selected in the 3P+1F control sample, together with the expected reducible background estimated from Equation 5.1, stacked on the distribution of WZ and of irreducible background ($ZZ, Z\gamma^* \rightarrow 4\ell$) taken from simulation.

If the fake rates were measured in a sample that had exactly the same background composition as the 2P+2F sample, the difference between the observed number of events in the 3P+1F sample and the expected background predicted from the 2P+2F sample would solely amount to the WZ and $Z\gamma_{\text{conv}}$ contribution, which is small. Large differences arise

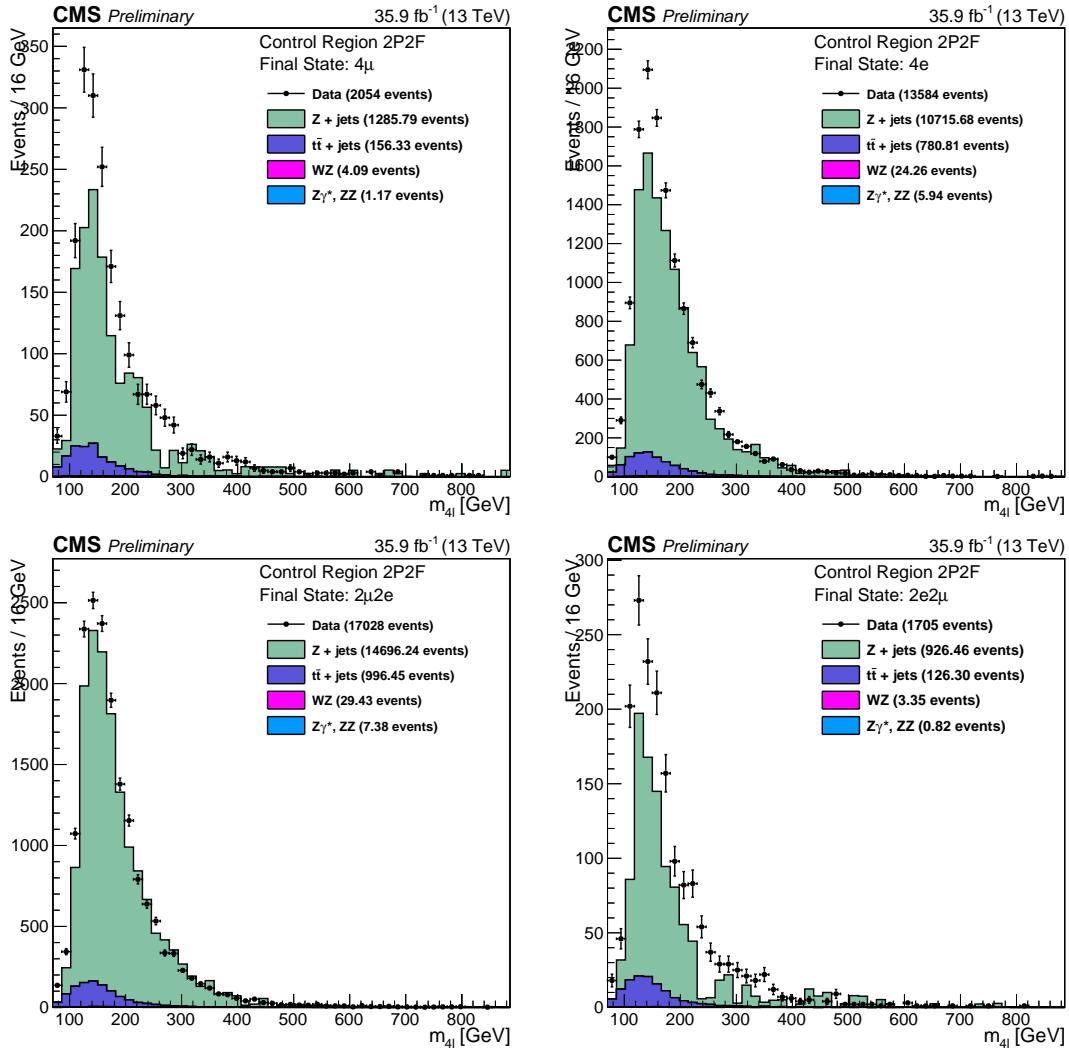


Figure 5.9: Invariant mass distribution of the events selected in the 2P+2F control sample in the 13 TeV dataset, (top left) 4μ , (top right) $4e$, (bottom left) $2\mu 2e$ and (bottom right) $2e 2\mu$ channels.

because the fake rates used in Equation 5.1 do not properly account for the background composition of the 2P+2F control sample.

The difference seen in Figure 5.10 between the observed 3P+1F distribution and the expectation from 2P+2F in the channels with loose electrons ($4e$ and $2\mu 2e$), concentrated at low masses, is due to photon conversions. This is confirmed explicitly by simulation. The difference between the 3P+1F observation and the prediction from 2P+2F is corrected to recover the missing contribution from conversions. More precisely, the expected

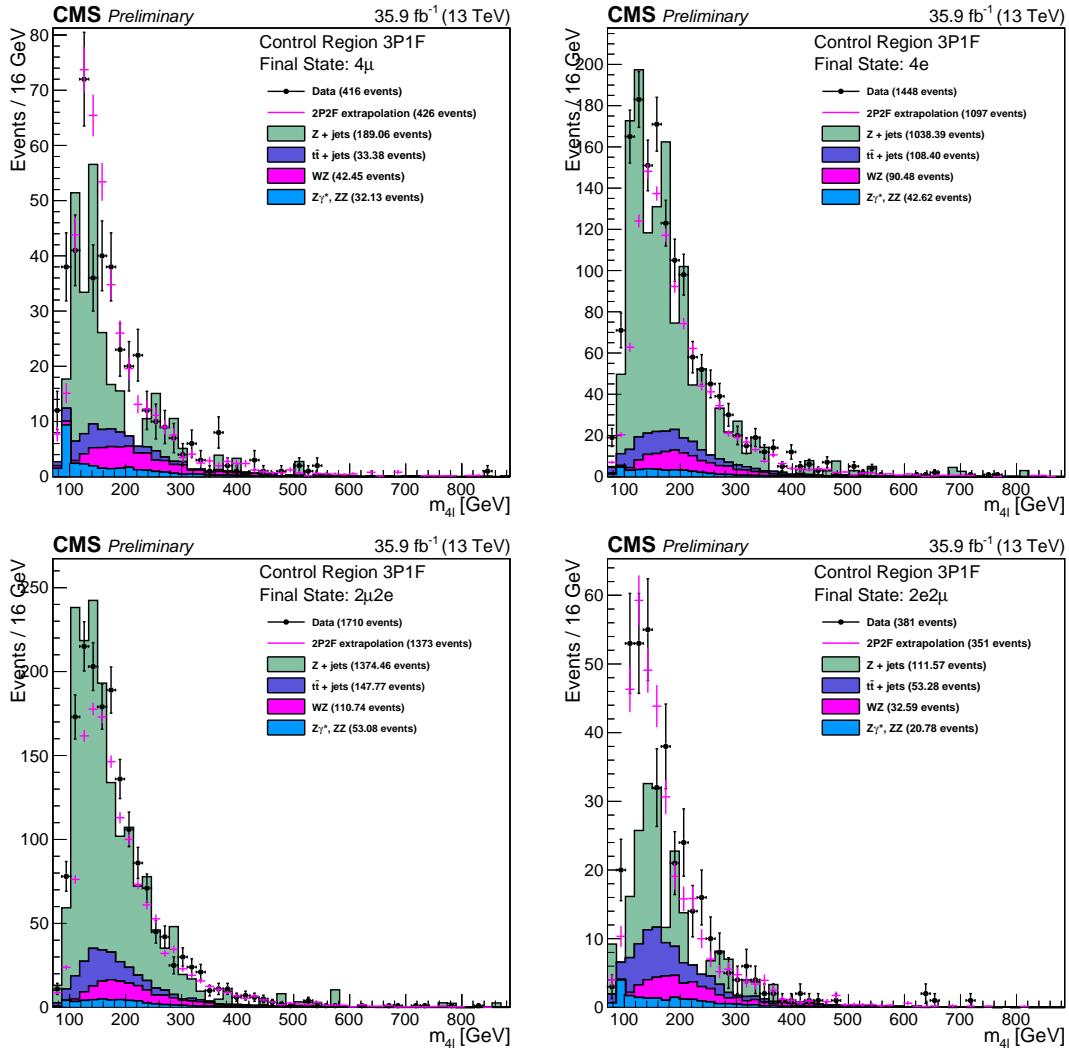


Figure 5.10: Invariant mass distribution of the events selected in the 3P+1F control sample in the 13 TeV dataset, (top left) 4μ , (top right) $4e$, (bottom left) $2\mu 2e$ and (bottom right) $2e 2\mu$ channels.

reducible background in the SR is given by the sum of two terms: a 2P2F component and a 3P1F component.

The 2P2F component is obtained from the number of events observed in the 2P+2F control region, N_{2P2F} , by weighting each event in that region with the factor $\frac{f_i}{1-f_i} \frac{f_j}{1-f_j}$, where f_i and f_j correspond to the fake rates of the two loose leptons. The 3P1F component is obtained from the difference between the number of observed events in the 3P+1F control region, N_{3P1F} , and the expected contribution from the 2P+2F region and ZZ

processes in the signal region, $N_{3P1F}^{ZZ} + N_{3P1F}^{\text{bkg}}$. The N_{3P1F}^{bkg} is given by Equation 5.1 and N_{3P1F}^{ZZ} is the contribution from ZZ which is taken from simulation. The difference $N_{3P1F} - N_{3P1F}^{\text{bkg}} - N_{3P1F}^{ZZ}$, which may be negative, is obtained for each (p_T, η) bin of the probe lepton and is weighted by $\frac{f_i}{1-f_i}$, where f_i denotes the fake rate of this lepton. This 3P1F component accounts for the contribution of reducible background processes with only one fake lepton (like WZ events) and for the contribution of other processes (e.g. photon conversions) that are not properly estimated by the 2P2F component because of the fake rates used.

The full expression for the prediction can be symbolically written as:

$$N_{\text{SR}}^{\text{bkg}} = \sum \frac{f_i}{(1-f_i)} (N_{3P1F} - N_{3P1F}^{\text{bkg}} - N_{3P1F}^{ZZ}) + \sum \frac{f_i}{(1-f_i)} \frac{f_j}{(1-f_j)} N_{2P2F} \quad (5.2)$$

or equivalently:

$$N_{\text{SR}}^{\text{bkg}} = \left(1 - \frac{N_{3P1F}^{ZZ}}{N_{3P1F}}\right) \sum_j^{N_{3P1F}} \frac{f_a^j}{1-f_a^j} - \sum_i^{N_{2P2F}} \frac{f_3^i}{1-f_3^i} \frac{f_4^i}{1-f_4^i} \quad (5.3)$$

For channels where the Z_2 candidate is made from two electrons, the contribution of the 3P1F component is positive and amounts to typically 30% of the total predicted background.

For channels with loose muons (e.g. μ and $2e2\mu$), the 3P+1F sample is rather well described by the prediction from 2P+2F, as seen in Figure 5.10, and the 3P1F component is mainly driven by statistical fluctuations in the 3P+1F sample, which are larger than the expectation from WZ production.

Table 5.2 shows the expected number of events in the four-lepton SRs from the reducible background processes.

	$4e$	4μ	$2e2\mu$	$2\mu2e$
baseline	22.19	32.81	22.48	41.72
13 TeV				

Table 5.2: The contribution of reducible background processes in the four-lepton signal region predicted from measurements in data using the Opposite-Sign Leptons method. The predictions correspond to 35.9 fb^{-1} of data at 13 TeV.

5.4 Event yields

Table 5.3 shows the event yields for the primary backgrounds, a benchmark signal, and the observed events at the major event selection stages. This is called the analysis cutflow. The selection of tight leptons and the formation of the first Z candidate are sufficient to remove most of the background events. After the second Z is formed, the remaining selection steps do not reduce backgrounds significantly but assist in the selection of events with clean ZZ^* candidates.

Sample	$q\bar{q} \rightarrow ZZ$	$gg \rightarrow ZZ$	$Z + X$	ZH	Other H	Total	Signal	Observed
Initial	1.89e+04	866	5.88e+08	17.4	377	5.88e+08	3.57	8.21e+07
HLT	1.89e+04	866	5.88e+08	17.4	377	5.88e+08	3.57	8.21e+07
Z_1 lepton cuts	1.94e+03	144	5.9e+04	2.35	57	6.12e+04	0.516	7.76e+04
m_{Z_1}	1.58e+03	128	4.39e+04	2.13	53.8	4.57e+04	0.484	3.52e+04
At least one Z_2	379	76.4	30.3	0.987	24.3	511	0.228	367
m_{Z_2}	379	76.4	56.7	0.987	24.3	537	0.228	367
$m_{ll} > 4$ for OS-SF	312	33.7	27.7	0.92	11.3	386	0.213	367
$m_{lll} > 70$ GeV	312	33.7	27.7	0.92	11.3	385	0.213	365

Table 5.3: Cutflow table for the 4μ channel. The benchmark signal sample shown is Zp2HDM with $m_{Z'} = 600$ GeV.

The final yields in the cut-and-count based SR are shown in Table 5.4. The final yields in the MVA-based SR, requiring the respond D to be greater than 0.9, are shown in Table 5.5.

Sample	$q\bar{q} \rightarrow ZZ$	$gg \rightarrow ZZ$	$Z + X$	H	Total	Signal	Observed
Yield	0.35	0.034	0.023	0.33	0.74	0.20	1.0

Table 5.4: Final signal region yields using the cut-and-count selection strategy for the 4μ channel. The benchmark signal sample shown is Zp2HDM with $m_{Z'} = 600$ GeV.

Sample	$q\bar{q} \rightarrow ZZ$	$gg \rightarrow ZZ$	$Z + X$	H	Total	Signal	Observed
Yield	117.5	4.54	17.6	11.7	151.3	0.21	151.0

Table 5.5: Final signal region yields using the MVA selection strategy for the 4μ channel. The benchmark signal sample shown is Zp2HDM with $m_{Z'} = 600$ GeV.

Chapter 6

Statistical analysis results

6.1 Systematic uncertainties

This section covers the systematic, including statistical, errors associated with the analysis. The general strategy is to duplicate the systematics applied in the SM $H \rightarrow ZZ$ search, including errors on the background estimation from data, while adding the errors associated with MET modeling, guided by the mono- $H \rightarrow \gamma\gamma$ strategy [107].

6.1.1 Uncertainties on the reducible background estimation

The main source of systematic uncertainty on the background estimation method is due to the different compositions of the reducible background processes (DY , $t\bar{t}$, WZ , $Z\gamma^{(*)}$) in the regions where the fake rates are measured and where they are applied. The OS method corrects for the resulting bias with the 3P1F component of its prediction. The closure tests presented here are used to assess a possible residual bias in the OS method.

The systematic uncertainty due to different compositions of events can be estimated by measuring the fake rates for individual background processes in the $Z + 1L$ region in simulation. The weighted average of these individual fake rates is used for the overall fake rate. The exact composition of the background processes in the 2P+2F region where the fake rates are applied can be determined from simulation. The individual fake rates can then be reweighted according to the 2P+2F composition. The difference between the reweighed fake rate and the average one can be used as a measure of the uncertainty on the measurement of the fake rates. The effect of this systematic uncertainty is propagated

to the final estimates, and it amounts to approximately 32% for $4e$, 33% for $2e2\mu$, and 35% for the 4μ final state.

Additional uncertainties arise from the limited size of the samples in the four-lepton control regions, where the fake rates are measured and applied contributes a statistical uncertainty. The dominate statistical uncertainty is driven by the number of events in the control region and is typically in the range of 1–10%.

In order to estimate the uncertainty on the m_{4l} shape, the differences among the shapes of predicted background distributions for all three channels are studied. The envelope of differences among these distribution shapes is used as an estimate of the shape uncertainty. The uncertainty is estimated to be roughly in the range of 5–15%. Since the difference of the shapes slowly varies with m_{4l} , it is taken as a constant versus m_{4l} and is absorbed in the much larger uncertainty on the predicted yield of backgrounds.

6.1.2 MET systematics

There are two types of systematic uncertainties related to the modeling of MET: (1) those from the measurement of real MET, as from the signal samples or backgrounds with neutrinos, and (2) those from fake MET, due to the mismeasurement of jets and other objects. The fake MET systematics apply to the H signals with no associated W production and to the non-resonant backgrounds.

The uncertainties from the mismodeling of real MET are measured by varying several corrections used to calculate MET, then propagating these variations to the efficiency of MC samples to pass the MET selection [162]. The corrections used in this calculation are: jet energy, jet resolution, muon energy, electron energy, tau energy, photon energy, and unclustered jet energy. Each correction is varied up and down by one standard deviation of the input distribution, with the systematic uncertainty taken as the maximum difference in efficiencies across all correction variations. The efficiencies for the V+H samples to pass the MET selection vary by a few percent and the variation in signal sample efficiencies is less than one percent.

The second systematic uncertainty results from the mismodeling of fake MET, primarily due to the mismeasurement of jets (see Section 5.3.2). This uncertainty is measured

in the sideband CR as the percent difference between the efficiency for the data and total background sample to pass the MET selection. These efficiencies differ by 42%, which is taken as the systematic on backgrounds without real MET.

6.1.3 Additional systematics

Both signal and background samples are affected by several additional systematic uncertainties, including the uncertainty on the integrated luminosity (2.6%) and the uncertainty on the lepton identification and reconstruction efficiency (ranging from 2.5–9%). Experimental uncertainties for the reducible background estimation, described in Section 5.3.1.1, vary between 36% (4μ) and 43% ($4e$). The uncertainty on the lepton energy scale is determined by considering the $Z \rightarrow \ell\ell$ mass distributions in data and simulation. Events are separated into categories based on the p_T and η of one of the two leptons, determined randomly, and integrating over the other. The dilepton mass distributions are then fit to a Breit-Wigner parameterization convolved with a double-sided Crystal-Ball function. The offset in the measured peak position with respect to the nominal Z boson mass in data and simulation are extracted. The results are shown in Figure 6.1. The relative difference between data and simulation is propagated to the reconstructed four-lepton mass from simulated H events. The results of the propagation can be seen in Figure 6.2. In the case of electrons, since the same dataset is used to derive and validate the momentum scale corrections, the size of the corrections are taken into account for the final value of the uncertainty. The uncertainty is determined to be 0.04% (0.3%) for the 4μ ($4e$) channel. The uncertainty on the 4ℓ mass resolution coming from the uncertainty on the per-lepton energy resolution is 20%. The experimental systematic uncertainties are summarized in Table 6.1.

Theoretical uncertainties which affect both the background signal and background estimation include uncertainties from the renormalization and factorization scale and choice of PDF set. The uncertainty from the renormalization and factorization scale is determined by varying these scales between 0.5 and 2 times their nominal value while keeping their ratio between 0.5 and 2. The uncertainty from the PDF set is determined by taking the root mean square of the variation when using different replicas of the

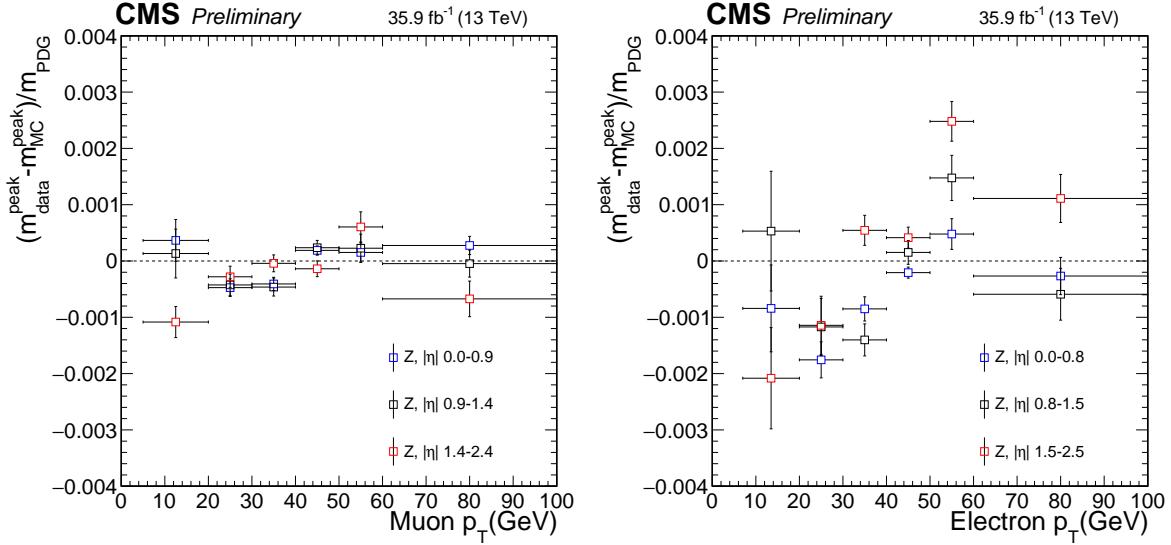


Figure 6.1: Difference between the $Z \rightarrow \ell\ell$ mass peak positions in data and simulation normalized by the nominal Z boson mass obtained as a function of the p_{T} and $|\eta|$ of one of the leptons, regardless of the second, for muons (left) and electrons (right).

default NNPDF set. An additional uncertainty of 10% on the k -factor used for the $gg \rightarrow ZZ$ prediction is applied as described in Section 5.3.1. A systematic uncertainty of 2% on the branching ratio of $H \rightarrow ZZ \rightarrow 4\ell$ only affects the H signal yields. In the case of event categorization, experimental and theoretical uncertainties that account for possible migration of signal and background events between categories are included. The main sources of uncertainty on the event categorization include the QCD scale, PDF set, and the modeling of hadronization and the underlying event. These uncertainties amount to between 4–20% for the signal and 3–20% for the background depending on the category. The lower range corresponds to the VBF and VH processes and the upper range corresponds to the $gg \rightarrow H$ process yield in the VBF-2jet-tagged category. Additional uncertainties come from the imprecise knowledge of the jet energy scale (from 2% for the $gg \rightarrow H$ yield in the untagged category to 15% for $gg \rightarrow H$ yield in the VBF-2jet-tagged category) and b-tagging efficiency and mistag rate (up to 6% in the tagged category). The theoretical systematic uncertainties are summarized in Table 6.2.

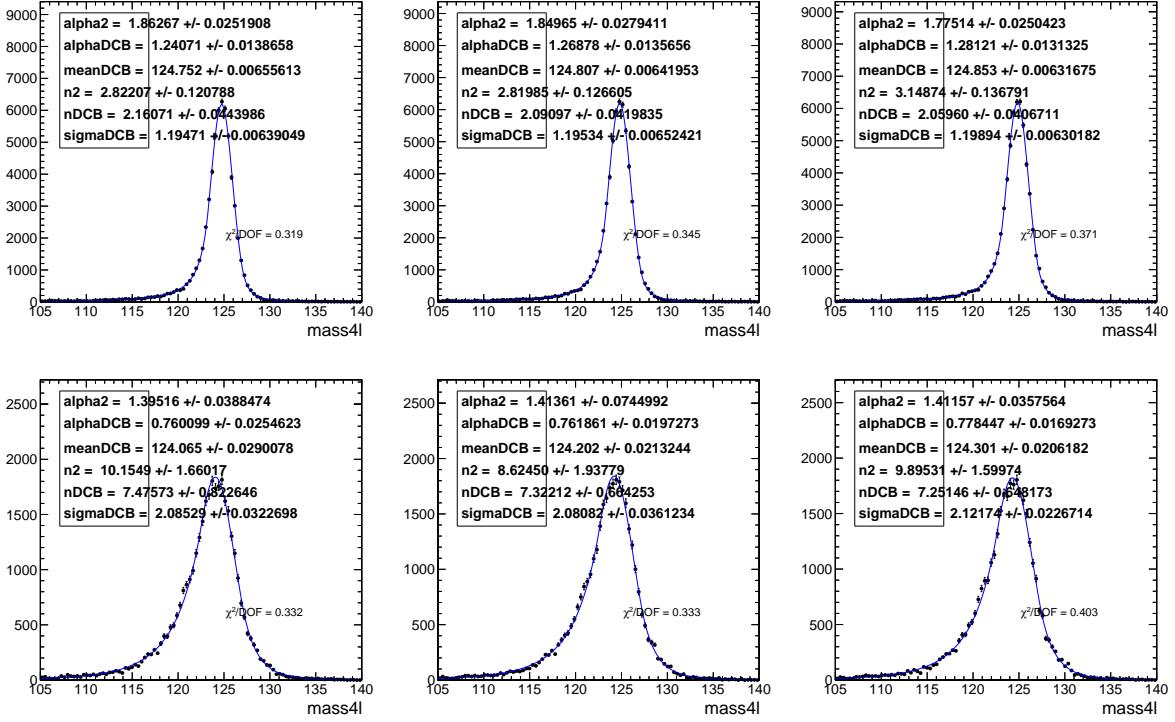


Figure 6.2: Different $m_{4\ell}$ distributions after propagating the biases in Fig. 6.1 to Higgs boson events. The change in the mean of the double Crystal-Ball is used to determine the systematic uncertainty due to the lepton momentum scale. The middle plot shows the nominal distribution, while the left (right) plots show the down (up) systematic variations. The 4μ channel is shown in the top row and the $4e$ channel is shown in the bottom row.

Source of uncertainty	Value
Luminosity	2.6 %
Lepton identification/reconstruction efficiencies	2.5 – 9 %
Reducible background ($Z+X$)	36 – 43 %
Lepton energy scale	0.04 – 0.3 %
Lepton energy resolution	20 %

Table 6.1: Summary of the experimental systematic uncertainties in the $H \rightarrow 4\ell$ measurements.

6.2 Limit setting

The primary tool used to interpret the analysis described above in the context of the signal

Source of uncertainty	Value
QCD scale (gg)	$\pm 3.9 \%$
PDF set (gg)	$\pm 3.2 \%$
Bkg K factor (gg)	$\pm 10 \%$
QCD scale (VBF)	$+0.4/-0.3 \%$
PDF set (VBF)	$\pm 2.1 \%$
QCD scale (WH)	$+0.5/-0.7 \%$
PDF set (WH)	$\pm 1.9 \%$
QCD scale (ZH)	$+3.8/-3.1 \%$
PDF set (ZH)	$\pm 1.6 \%$
QCD scale ($t\bar{t}H$)	$+5.8/-9.2 \%$
PDF set ($t\bar{t}H$)	$\pm 3.6 \%$
$BR(H \rightarrow ZZ \rightarrow 4\ell)$	2 %
QCD scale ($q\bar{q} \rightarrow ZZ$)	$+3.2/-4.2 \%$
PDF set ($q\bar{q} \rightarrow ZZ$)	$+3.1/-3.4 \%$
Electroweak corrections ($q\bar{q} \rightarrow ZZ$)	$\pm 0.1 \%$

Table 6.2: Summary of the theory systematic uncertainties in the $H \rightarrow 4\ell$ measurements for the inclusive analysis.

models is the HiggsAnalysis-CombinedLimit package [173], a collection of RooStats-based software [174] used within the Higgs physics analysis group (PAG) [175], hereafter called the combine tool, or simply "combine". Specifically, one-sided Bayesian credible interval limits are set on the expected and observed signal production cross section times branching ratio (BR) of $H \rightarrow ZZ \rightarrow 4\ell$ ($\sigma \times BR$). Limits are set on the various signal benchmarks using the asymptotic CL_S method [176], an approach to calculating a profile likelihood ratio using an approximation of the LHC test-statistic distributions. The upper limit on the cross section gives the maximum number of events that can be attributed to the signal process, consistent with the data that is observed.

Combine finds the upper limit as the numerical solution to Equation 6.1, which sets the integral of the posterior probability $p(\sigma|D)d\sigma$ equal to the desired confidence level

for the measurement, typically 95%. The posterior probability gives the degree of belief that σ lies in the interval $[\sigma, \sigma + d\sigma]$ and is formed by inverting a multi-Poisson model, $p(D|\sigma, \boldsymbol{\theta})$ using Bayes' theorem (Equation 6.2), after numerically marginalizing priors describing the uncertainties $\pi(\boldsymbol{\theta})$ (Equation 6.3).

$$\int_0^{\sigma_{\text{up}}} p(\sigma|D)d\sigma = 1 - \alpha \quad (6.1)$$

where,

$$p(\sigma|D) = p(D|\sigma)\pi(\sigma)/p(D) \quad (6.2)$$

and,

$$p(D|\sigma) = \int p(D|\sigma, \boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}. \quad (6.3)$$

Combine takes specially formatted text files called data cards as input. The data cards contain the signal, background, and data yields, along with associated systematic uncertainties, and the location of the file containing the shape distributions for each sample. When the shape distributions are included, combine computes the limits across all bins, then combines the results. A non-shape-based limit is equivalent to a shape-based limit with one bin. This can be used as a crosscheck that the shape-based limit is behaving correctly and to analyze the improvement of using the shape-based approach over non-shape-based.

Combine returns the 95% confidence level (CL) expected and observed limits on the signal strength parameter μ , as well as one and two standard deviations of the expected limit, where μ is the signal scale factor in the i^{th} bin in the mean count, $n_i = \mu * s_i + b_i$, where s_i and b_i are the signal and background yields, respectively. The signal strength parameter gives the ratio of the 95% CL expected signal yield to the theoretical yield, or more generally:

$$\mu[\sigma_{\text{norm}} * \text{BR}] = \frac{\sigma_{95\%\text{CL}} * \text{BR}}{\sigma_{\text{norm}} * \text{BR}}, \quad (6.4)$$

where σ_{norm} is the cross section used to normalize the signal yields given as input to combine, typically the theoretical production cross section, or set to 1 pb for comparison

to other H decay channels. This formula is used to calculate $\sigma_{95\%CL} \times \text{BR}$, which is the desired output variable. This limit can then be compared to the theoretical $\sigma \times \text{BR}$. The points where the upper limit is lower than the theory value are interpreted as exclusions.

6.2.1 Cross section limits

After the event selection is optimized and all of the uncertainties are accounted for, limits are set on the signal strength parameter μ , which is scaled according to Equation 6.4 to obtain the cross section limit. For each model, these results are found for both the cut-and-count and MVA-based strategies. For Zp2HDM, the one-dimensional slice of mass points fixing $m_{A^0} = 300$ GeV is selected for the limit plots since this region has the highest cross section and the largest branching fraction to DM production. For ZpBaryonic, the one-dimensional slice of mass points fixing $m_\chi = 1$ GeV is selected for the limit plots since this region has the highest cross section.

6.2.1.1 Cut-and-count based limits

The one-dimensional limits, obtained using the optimized selection given in Section 5.2.1, are presented for $\sigma_{95\%CL} \times \text{BR}$ in Figure 6.3 for Zp2HDM and Figure 6.4 for ZpBaryonic. In both models, the observed limits do not deviate more than one standard deviation from the expected limits, indicating the absence of a large excess of events in the SR above the SM prediction. Since the limit curves lie above the theoretical cross section curves, no benchmark points can be excluded using this analysis alone. However, when the Zp2HDM results are combined with those from the other H decay channels, there is sufficient sensitivity to exclude a large portion of the parameter space, up to mediator masses in the TeV range. These are the first results obtained for ZpBaryonic from any of the mono- H analyses. Since the ZpBaryonic theoretical cross section curve is flat up to about 1 TeV, any small increase in sensitivity by combining with other analyses will allow for the exclusion of these benchmark models.

The limiting factor in improving the sensitivity of the limits obtained with this strategy is the signal efficiency. The cross section limit is directly proportional to the signal selection efficiency, and each additional cut reduces the signal efficiency by a few percent or more. From Table 5.3, the largest fraction of signal events is lost during the tight

lepton and Z selection steps, even before the mono- H selection is applied. One method of improving the sensitivity of this analysis is to loosen the tight lepton selection criteria, but allowing more signal to pass a looser cut would also let more background pass. However, since MET is such a powerful discriminating variable, the additional backgrounds could potentially be reduced by a harder MET cut. Defining a new working point for the selection of tight leptons would require remeasuring lepton systematics and background estimates from data and the validation of modeling in additional CRs. This analysis is left to a future study.

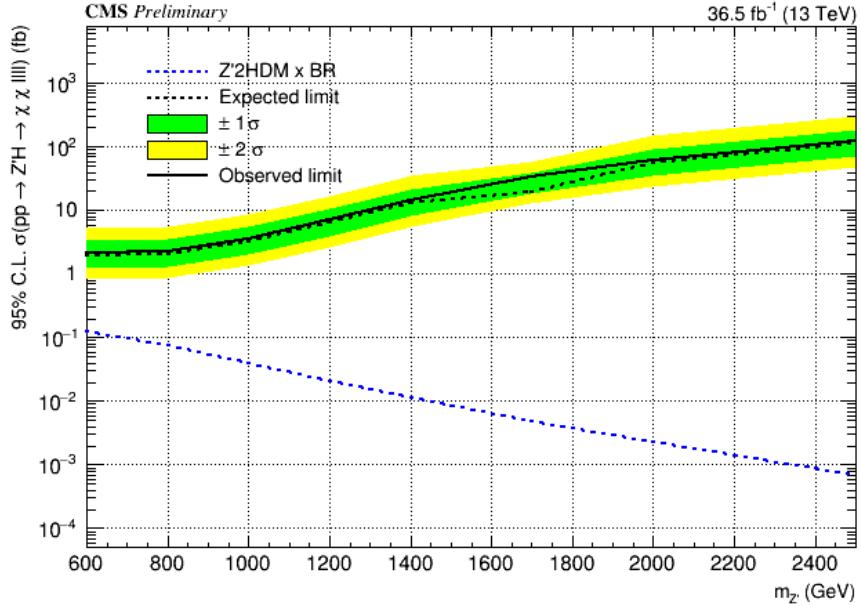


Figure 6.3: One-dimensional cross section times branching fraction limits for the Zp2HDM simplified model using the cut-and-count based event selection strategy.

6.2.1.2 MVA-based limits

The one-dimensional limits, obtained using the optimized selection given in Section 5.2.2, are presented for $\sigma_{95\%CL} \times BR$ in Figure 6.5 for Zp2HDM and Figure 6.6 for ZpBaryonic. The sensitivity obtained using the MVA approach is not as good as for the cut-and-count approach. This reiterates the strong discriminating power of MET used in the cut-and-count strategy, rather than indicating a poor performance of the MVA strategy. Again, the signal efficiency is the limiting factor in improving the sensitivity. Since there are no

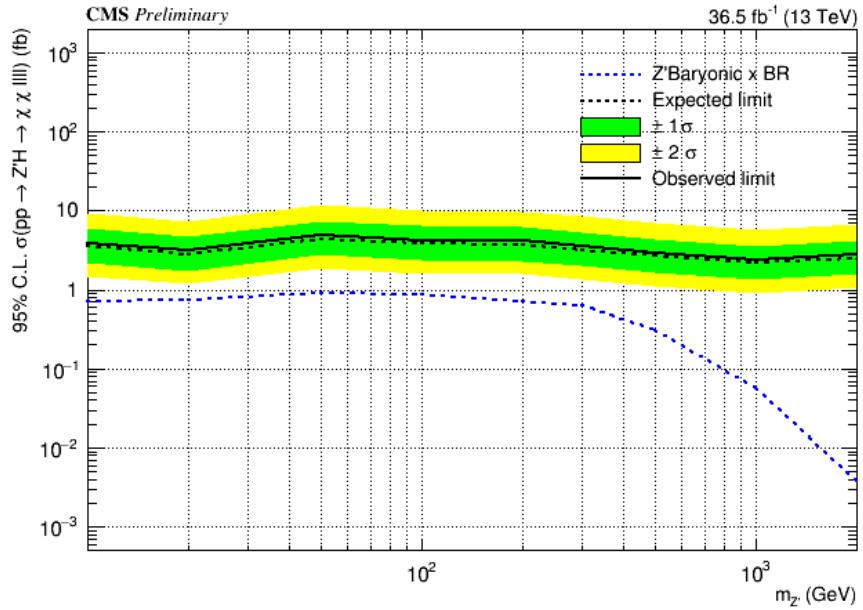


Figure 6.4: One-dimensional cross section times branching fraction limits for the ZpBaryonic simplified model using the cut-and-count based event selection strategy.

additional cuts beyond the SM selection in this selection strategy, the SM selection itself would need to be modified, requiring additional study.

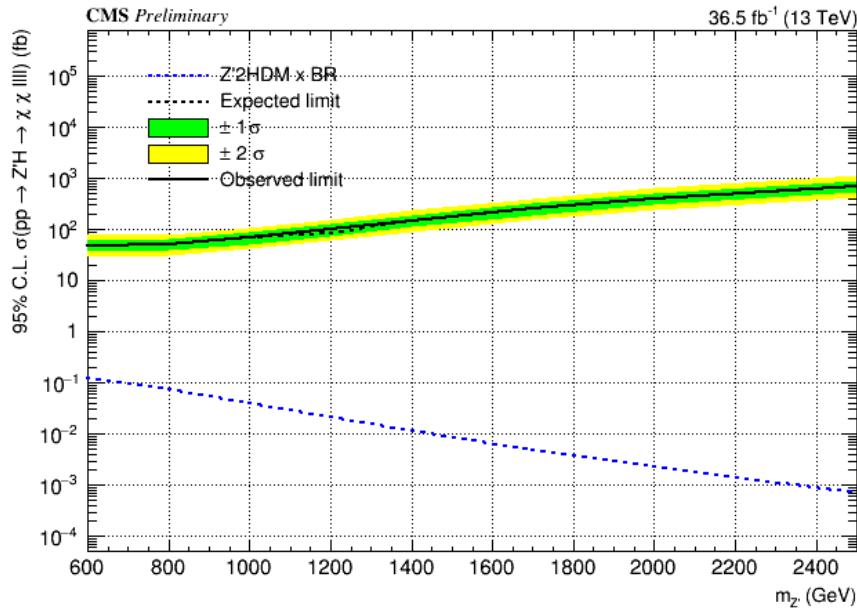


Figure 6.5: One-dimensional cross section times branching fraction limits for the Zp2HDM simplified model using the MVA-based event selection strategy.

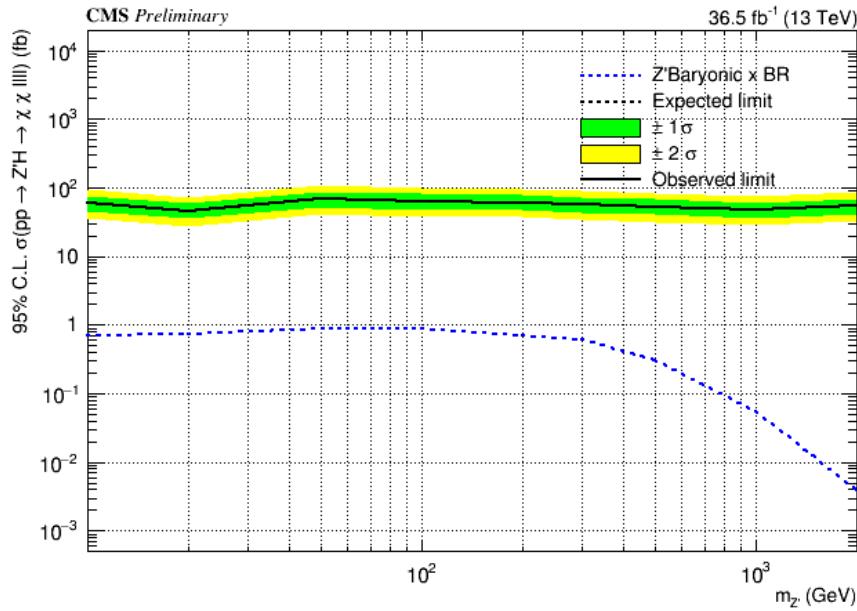


Figure 6.6: One-dimensional cross section times branching fraction limits for the ZpBaryonic simplified model using the MVA-based event selection strategy.

Chapter 7

Conclusions

The discovery of the Higgs boson was the highlight of the physics results of Run 1 at the LHC. In addition to measuring the properties and couplings of the Higgs, efforts were launched in Run 2 to use the Higgs as a probe for new physics, including to search for dark matter. In parallel with these studies, progress was made by complementary, non-collider searches. Although no direct detection has been confirmed using any approach, large areas of parameter space have been excluded. Collider searches, including the search for the mono-Higgs signature, have much better sensitivity to low-mass dark matter than direct searches, as well as the ability to study higher order couplings that are not accessible to direct searches. The main results of this dissertation are (1) the rediscovery of the Higgs boson with Run 2 data and (2) the cross section limits on two dark matter models, contributing to the world-leading low-mass limits and mediator mass exclusions for these models.

The impact of this study is very high, being among the first-ever searches for the mono-Higgs signature. Several key contributions were made to the standard model Higgs search in the four-lepton final state, including (1) a cross-check exercise to validate the event selection with other groups and (2) the definition of a new event category for Higgs candidate events with large missing transverse momentum. The key ingredient that this study added to the Standard Model Higgs search and collaboration documentation is the study of missing energy, in particular, whether the background and statistical modeling techniques remain valid in the high missing energy regime.

Numerous extensions of this analysis would have a large impact on the field. The selection used to identify events with a Higgs boson can be reoptimized for mono-Higgs signals rather than for the Standard Model Higgs signals. This could significantly increase the limit-setting sensitivity. The limits found with the four-lepton final state can be combined with those from the other Higgs decay channel analyses to obtain even stronger results. This work is underway. There are additional models that predict a mono-Higgs signature that can be studied with the datasets currently available. The most commonly studied models are those that predict very hard missing energy spectra, which gives an advantage to other Higgs decay channels. However, the four-lepton channel would have better sensitivity for the unstudied models with softer spectra. Finally, a more detailed study can be done to understand the relationship of the limits found here with those set by other search strategies in order to guide the design of future analyses and experiments.

There is strong motivation for the existence of a coupling between the dark sector and ordinary matter. Studies such as this one shed light on the nature of this interaction. Technology and analysis techniques are constantly advancing. If such an interaction exists, it is only a matter of time before dark matter is observed, expanding our understanding of the mass-energy content of the universe beyond the mere 5% that we currently know.

Appendices

Appendix A

Production cross sections for benchmark signal models

m_χ [GeV]	1	10	50	65	100	200	400	800	1000	1300
EFT_HHxx_scalar										
EFT_HHxx_combined	0.10071E+01	0.99793E+00	0.60671E+00	0.48291E-04	0.22725E-05	0.11059E-06	0.36569E-08	0.40762E-10	0.64956E-11	0.51740E-12
EFT_HHxg5x	0.15731E+01	0.15194E+01	0.34134E+00	0.41039E-04	0.10581E-04	0.16553E-05	0.14628E-06	0.40608E-08	0.85950E-09	0.96480E-10
EFT_xdxHDHc	0.15735E+03	0.15394E+03	0.94804E+02	0.12990E-01	0.23075E-02	0.41820E-03	0.45743E-04	0.16734E-05	0.39327E-06	0.49769E-07
$\Lambda = 100$ GeV	0.29530E+00	0.29067E+00	0.10540E+00	0.89849E-01	0.64959E-01	0.30639E-01	0.88644E-02	0.97986E-03	0.33847E-03	0.68674E-04
$\Lambda = 1000$ GeV	0.16306E-04	0.15508E-04	0.12088E-05	0.88288E-06	0.53312E-06	0.18046E-06	0.34918E-07	0.27514E-08	0.90662E-09	0.19313E-09
EFT_xgxFHDH	0.57027E+00	0.57001E+00	0.56025E+00	0.55337E+00	0.53270E+00	0.45792E+00	0.29777E+00	0.10288E+00	0.57444E-01	0.23260E-01

Table A.1: Effective field theory model production cross sections [pb]

m_χ [GeV]	m_S [GeV]					
1	0.21915E+01	0.20798E+01	0.19192E+01	0.18118E+01	0.16735E+01	0.52244E+01
10	0.17416E+01	0.17420E+01	0.18581E+01	0.17510E+01	0.17397E+00	0.41877E+01
50	0.39053E+00	0.38877E+00	0.38409E+00	0.37097E+00	0.12861E+01	0.28732E+01
150	0.24136E-05			0.38372E-05	0.21922E-04	0.42337E-03
500	0.34099E-08				0.49399E-08	0.57124E-04
1000	0.17012E-10				0.25206E-06	0.36823E-06
					0.55260E-10	0.11067E-07

Table A.2: Scalar simplified model production cross sections [pb] corresponding to mass points in Table 1.11

m_χ [GeV]	$m_{Z'}$ [GeV]					
1	0.26615E+01	0.27802E+01	0.33248E+01	0.32341E+01	0.26566E+01	0.23191E+01
10	0.21182E-01	0.74027E-01	0.32732E+01	0.32250E+01	0.10842E+01	0.18700E+00
50	0.36342E-03	0.12726E-01	0.31337E+00	0.21226E+01	0.20120E+01	0.11728E-01
150	0.55972E-05			0.56526E-02	0.18000E+00	0.17399E-07
500	0.80295E-08				0.36591E-04	0.17380E-07
1000	0.49387E-10				0.98079E-06	0.17340E-07
					0.57596E-03	0.16918E-07
						0.13179E-07
						0.80146E-08

Table A.3: Zp Baryonic simplified model production cross sections [pb] corresponding to mass points in Table 1.8

m_χ [GeV]	$m_{Z'}$ [GeV]						
1	0.61935E-02	0.63192E-02	0.82991E-02	0.11942E-01	0.19171E-01	0.21560E-01	0.16010E-01
10	0.58781E-02	0.58938E-02	0.82944E-02	0.11937E-01			0.64416E-02
50	0.10294E-03		0.77820E-04	0.15258E-04	0.12066E-01	0.12105E-01	
150	0.28382E-06				0.83917E-05	0.72889E-03	0.65401E-02
500	0.34689E-09					0.43355E-06	0.87799E-04
1000	0.20703E-11						0.44782E-07

Table A.4: ZpHS simplified model production cross sections [pb] corresponding to mass points in Table 1.8

m_{A^0} [GeV]	$m_{Z'}$ [GeV]					
300	42.386	45.097	35.444	26.07	18.942	11.778
400	5.8513	14.847	14.534	11.792	9.029	5.851
500	5.9605	8.4961	7.9575	6.5515	4.5063	3.0028
600	1.5853	4.6972	5.4808	4.9946	3.7044	2.5694
700		2.1092	3.4848	3.6766	3.0253	2.2023
800		0.65378	1.9638	2.5511	2.4077	1.8689
						1.0692

Table A.5: Zp2HDM simplified model production cross sections [pb] corresponding to mass points in Table 1.10

BIBLIOGRAPHY

- [1] 2012. URL: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/Hig12041TWiki>.
- [2] The CMS Collaboration. The cms experiment at the cern lhc. *Journal of Instrumentation*, 3(08):S08004, 2008. URL: <http://stacks.iop.org/1748-0221/3/i=08/a=S08004>.
- [3] The CMS collaboration. Performance of cms muon reconstruction in pp collision events at $s = 7$ tev. *Journal of Instrumentation*, 7(10):P10002, 2012. URL: <http://stacks.iop.org/1748-0221/7/i=10/a=P10002>.
- [4] Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. 2009.
- [5] Linda Carpenter, Anthony DiFranzo, Michael Mulhearn, Chase Shimmin, Sean Tulin, and Daniel Whiteson. Mono-Higgs-boson: A new collider probe of dark matter. *Phys. Rev.*, D89(7):075017, 2014. [arXiv:1312.2592](https://arxiv.org/abs/1312.2592), doi:10.1103/PhysRevD.89.075017.
- [6] Daniel Abercrombie et al. Dark Matter Benchmark Models for Early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum. 2015. [arXiv:1507.00966](https://arxiv.org/abs/1507.00966).
- [7] Emmy Noether. Invariant Variation Problems. *Gott. Nachr.*, 1918:235–257, 1918. [Transp. Theory Statist. Phys.1,186(1971)]. [arXiv:physics/0503066](https://arxiv.org/abs/physics/0503066), doi:10.1080/00411457108231446.
- [8] J. Clerk Maxwell. A dynamical theory of the electromagnetic field. *Philosophical Transactions of the Royal Society of London*, 155:459–512, 1865. URL: <http://rstl.royalsocietypublishing.org/content/155/459.short>, [arXiv: http://rstl.royalsocietypublishing.org/content/155/459.full.pdf+html](http://rstl.royalsocietypublishing.org/content/155/459.full.pdf+html), doi:10.1098/rstl.1865.0008.
- [9] H. A. Lorentz. Simplified Theory of Electrical and Optical Phenomena in Moving Systems. *Koninklijke Nederlandse Akademie van Wetenschappen Proceedings Series B Physical Sciences*, 1:427–442, 1898.
- [10] Emmy Noether. Invariant Variation Problems. *Gott. Nachr.*, 1918:235–257, 1918. [Transp. Theory Statist. Phys.1,186(1971)]. [arXiv:physics/0503066](https://arxiv.org/abs/physics/0503066), doi:10.1080/00411457108231446.
- [11] Albert Einstein. On the electrodynamics of moving bodies. *Annalen Phys.*, 17:891–921, 1905. [Annalen Phys.14,194(2005)]. doi:10.1002/andp.200590006.
- [12] Albert Einstein. On the Foundations of the Generalized Theory of Relativity and the Theory of Gravitation. *Phys. Z.*, 15:176–180, 1914.

- [13] Arthur H. Compton. A quantum theory of the scattering of x-rays by light elements. *Phys. Rev.*, 21:483–502, May 1923. URL: <http://link.aps.org/doi/10.1103/PhysRev.21.483>, doi:10.1103/PhysRev.21.483.
- [14] Albert Einstein. Concerning an heuristic point of view toward the emission and transformation of light. *Annalen Phys.*, 17:132–148, 1905.
- [15] C. Davisson and L. H. Germer. Diffraction of Electrons by a Crystal of Nickel. *Phys. Rev.*, 30:705–740, 1927. doi:10.1103/PhysRev.30.705.
- [16] R. Shankar. *Principles of Quantum Mechanics*. Plenum Press, New York, 2 edition, 1994.
- [17] P.A.M. Dirac. *The Principles of Quantum Mechanics*. Oxford University Press, 4 edition, 1967.
- [18] Carl D. Anderson. The positive electron. *Phys. Rev.*, 43:491–494, Mar 1933. URL: <http://link.aps.org/doi/10.1103/PhysRev.43.491>, doi:10.1103/PhysRev.43.491.
- [19] P. A. M. Dirac. The quantum theory of the emission and absorption of radiation. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 114(767):243–265, 1927. URL: <http://rspa.royalsocietypublishing.org/content/114/767/243>, arXiv:<http://rspa.royalsocietypublishing.org/content/114/767/243.full.pdf>, doi:10.1098/rspa.1927.0039.
- [20] Alessandro Bettini. *Introduction to Elementary Particle Physics*. Cambridge University Press, 2 edition, 2014.
- [21] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Phys. Rev.*, 105:1413–1415, Feb 1957. URL: <http://link.aps.org/doi/10.1103/PhysRev.105.1413>, doi:10.1103/PhysRev.105.1413.
- [22] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2π decay of the k_2^0 meson. *Phys. Rev. Lett.*, 13:138–140, Jul 1964. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.13.138>, doi:10.1103/PhysRevLett.13.138.
- [23] Julian Schwinger. The theory of quantized fields. i. *Phys. Rev.*, 82:914–927, Jun 1951. URL: <http://link.aps.org/doi/10.1103/PhysRev.82.914>, doi:10.1103/PhysRev.82.914.
- [24] Gerhart Luders. On the Equivalence of Invariance under Time Reversal and under Particle-Antiparticle Conjugation for Relativistic Field Theories. *Kong. Dan. Vid. Sel. Mat. Fys. Med.*, 28N5:1–17, 1954.

- [25] Carl D. Anderson and Seth H. Neddermeyer. Cloud chamber observations of cosmic rays at 4300 meters elevation and near sea-level. *Phys. Rev.*, 50:263–271, Aug 1936. URL: <http://link.aps.org/doi/10.1103/PhysRev.50.263>, doi:10.1103/PhysRev.50.263.
- [26] C. M. G. Lattes, G. P. S. Occhialini, and C. F. Powell. Observations on the Tracks of Slow Mesons in Photographic Emulsions. 1. *Nature*, 160:453–456, 1947. doi:10.1038/160453a0.
- [27] C. L. Cowan, Jr., F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire. Detection of the Free Neutrino: A Confirmation. *Science*, 124:103–104, July 1956. doi:10.1126/science.124.3212.103.
- [28] Murray Gell-Mann. A Schematic Model of Baryons and Mesons. *Phys. Lett.*, 8:214–215, 1964. doi:10.1016/S0031-9163(64)92001-3.
- [29] Julian Schwinger. On quantum-electrodynamics and the magnetic moment of the electron. *Phys. Rev.*, 73:416–417, Feb 1948. URL: <http://link.aps.org/doi/10.1103/PhysRev.73.416>, doi:10.1103/PhysRev.73.416.
- [30] Julian Schwinger. Quantum electrodynamics. i. a covariant formulation. *Phys. Rev.*, 74:1439–1461, Nov 1948. URL: <http://link.aps.org/doi/10.1103/PhysRev.74.1439>, doi:10.1103/PhysRev.74.1439.
- [31] R. P. Feynman. The theory of positrons. *Phys. Rev.*, 76:749–759, Sep 1949. URL: <http://link.aps.org/doi/10.1103/PhysRev.76.749>, doi:10.1103/PhysRev.76.749.
- [32] R. P. Feynman. Space-time approach to quantum electrodynamics. *Phys. Rev.*, 76:769–789, Sep 1949. URL: <http://link.aps.org/doi/10.1103/PhysRev.76.769>, doi:10.1103/PhysRev.76.769.
- [33] R. P. Feynman. Mathematical formulation of the quantum theory of electromagnetic interaction. *Phys. Rev.*, 80:440–457, Nov 1950. URL: <http://link.aps.org/doi/10.1103/PhysRev.80.440>, doi:10.1103/PhysRev.80.440.
- [34] S. Tomonaga. On a relativistically invariant formulation of the quantum theory of wave fields. *Prog. Theor. Phys.*, 1:27–42, 1946. doi:10.1143/PTP.1.27.
- [35] F. J. Dyson. The radiation theories of Tomonaga, Schwinger, and Feynman. *Phys. Rev.*, 75:486–502, Feb 1949. URL: <http://link.aps.org/doi/10.1103/PhysRev.75.486>, doi:10.1103/PhysRev.75.486.
- [36] F. J. Dyson. The s matrix in quantum electrodynamics. *Phys. Rev.*, 75:1736–1755, Jun 1949. URL: <http://link.aps.org/doi/10.1103/PhysRev.75.1736>, doi:10.1103/PhysRev.75.1736.

- [37] G. Zweig. An SU(3) model for strong interaction symmetry and its breaking. Version 1. 1964.
- [38] G. Zweig. An SU(3) model for strong interaction symmetry and its breaking. Version 2. In D.B. Lichtenberg and Simon Peter Rosen, editors, *Developments in the quark theory of hadrons. VOL. 1. 1964 - 1978*, pages 22–101. 1964. URL: <https://inspirehep.net/record/4674/files/cern-th-412.pdf>.
- [39] H. David Politzer. Reliable perturbative results for strong interactions? *Phys. Rev. Lett.*, 30:1346–1349, Jun 1973. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.30.1346>, doi:10.1103/PhysRevLett.30.1346.
- [40] David J. Gross and Frank Wilczek. Asymptotically free gauge theories. i. *Phys. Rev. D*, 8:3633–3652, Nov 1973. URL: <http://link.aps.org/doi/10.1103/PhysRevD.8.3633>, doi:10.1103/PhysRevD.8.3633.
- [41] Enrico Fermi. Tentativo di una teoria dell’emissione dei raggi beta. *Ric. Sci.*, 4:491–495, 1933.
- [42] J. et al. Augustin. Discovery of a narrow resonance in e^+e^- annihilation. *Phys. Rev. Lett.*, 33:1406–1408, Dec 1974. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.33.1406>, doi:10.1103/PhysRevLett.33.1406.
- [43] J. et al. Aubert. Experimental observation of a heavy particle j . *Phys. Rev. Lett.*, 33:1404–1406, Dec 1974. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.33.1404>, doi:10.1103/PhysRevLett.33.1404.
- [44] S. W. Herb, D. C. Hom, L. M. Lederman, J. C. Sens, H. D. Snyder, J. K. Yoh, J. A. Appel, B. C. Brown, C. N. Brown, W. R. Innes, K. Ueno, T. Yamanouchi, A. S. Ito, H. Jöstlein, D. M. Kaplan, and R. D. Kephart. Observation of a dimuon resonance at 9.5 gev in 400-gev proton-nucleus collisions. *Phys. Rev. Lett.*, 39:252–255, Aug 1977. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.39.252>, doi:10.1103/PhysRevLett.39.252.
- [45] V. E. Barnes et al. Observation of a Hyperon with Strangeness -3. *Phys. Rev. Lett.*, 12:204–206, 1964. doi:10.1103/PhysRevLett.12.204.
- [46] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961. doi:10.1016/0029-5582(61)90469-2.
- [47] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.19.1264>, doi:10.1103/PhysRevLett.19.1264.
- [48] H. Weyl. Quantum mechanics and group theory. *Z. Phys.*, 46:1, 1927. doi:10.1007/BF02055756.

- [49] Daniel V. Schroeder Michael E. Peskin. *An Introduction to Quantum Field Theory*. Addison-Wesley Publishing Company, 1 edition, 1995.
- [50] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.13.321>, doi:10.1103/PhysRevLett.13.321.
- [51] Peter W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys. Lett.*, 12:132–133, 1964. doi:10.1016/0031-9163(64)91136-9.
- [52] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.13.508>, doi:10.1103/PhysRevLett.13.508.
- [53] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585–587, Nov 1964. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.13.585>, doi:10.1103/PhysRevLett.13.585.
- [54] Peter W. Higgs. Spontaneous symmetry breakdown without massless bosons. *Phys. Rev.*, 145:1156–1163, May 1966. URL: <http://link.aps.org/doi/10.1103/PhysRev.145.1156>, doi:10.1103/PhysRev.145.1156.
- [55] T. W. B. Kibble. Symmetry breaking in nonAbelian gauge theories. *Phys. Rev.*, 155:1554–1561, 1967. doi:10.1103/PhysRev.155.1554.
- [56] David Griffiths. *Introduction to Elementary Particles*. John Wiley and Sons, INC., 1 edition, 1987.
- [57] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012. arXiv:1207.7235, doi:10.1016/j.physletb.2012.08.021.
- [58] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett.*, B716:1–29, 2012. arXiv:1207.7214, doi:10.1016/j.physletb.2012.08.020.
- [59] J. C. Kapteyn. First Attempt at a Theory of the Arrangement and Motion of the Sidereal System. *Astrophys. J.*, 55:302–328, 1922. doi:10.1086/142670.
- [60] F. Zwicky. On the Masses of Nebulae and of Clusters of Nebulae. *Astrophys. J.*, 86:217–246, 1937. doi:10.1086/143864.
- [61] Vera C. Rubin and W. Kent Ford, Jr. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. *Astrophys. J.*, 159:379–403, 1970. doi:10.1086/150317.

- [62] V. C. Rubin, N. Thonnard, and W. K. Ford, Jr. Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 /R = 4kpc/ to UGC 2885 /R = 122 kpc/. *Astrophys. J.*, 238:471, 1980. doi:10.1086/158003.
- [63] K. A. Olive et al. Review of Particle Physics. *Chin. Phys.*, C38:090001, 2014. doi:10.1088/1674-1137/38/9/090001.
- [64] Douglas Clowe, Anthony Gonzalez, and Maxim Markevitch. Weak lensing mass reconstruction of the interacting cluster 1E0657-558: Direct evidence for the existence of dark matter. *Astrophys. J.*, 604:596–603, 2004. arXiv:astro-ph/0312273, doi:10.1086/381970.
- [65] L. V. E. Koopmans and T. Treu. The structure and dynamics of luminous and dark matter in the early-type lens galaxy of 0047-281 at z=0.485. *Astrophys. J.*, 583:606–615, 2003. arXiv:astro-ph/0205281, doi:10.1086/345423.
- [66] Henk Hoekstra, Howard Yee, and Mike Gladders. Current status of weak gravitational lensing. *New Astron. Rev.*, 46:767–781, 2002. arXiv:astro-ph/0205205, doi:10.1016/S1387-6473(02)00245-2.
- [67] Wayne Hu and Scott Dodelson. Cosmic microwave background anisotropies. *Ann. Rev. Astron. Astrophys.*, 40:171–216, 2002. arXiv:astro-ph/0110414, doi:10.1146/annurev.astro.40.060401.093926.
- [68] Wayne Hu, Naoshi Sugiyama, and Joseph Silk. The Physics of microwave background anisotropies. *Nature*, 386:37–43, 1997. arXiv:astro-ph/9604166, doi:10.1038/386037a0.
- [69] Gianfranco Bertone, Dan Hooper, and Joseph Silk. Particle dark matter: Evidence, candidates and constraints. *Phys. Rept.*, 405:279–390, 2005. arXiv:hep-ph/0404175, doi:10.1016/j.physrep.2004.08.031.
- [70] Lars Bergström. Nonbaryonic dark matter: Observational evidence and detection methods. *Rept. Prog. Phys.*, 63:793, 2000. arXiv:hep-ph/0002126, doi:10.1088/0034-4885/63/5/2r3.
- [71] P. Tisserand et al. Limits on the Macho Content of the Galactic Halo from the EROS-2 Survey of the Magellanic Clouds. *Astron. Astrophys.*, 469:387–404, 2007. arXiv:astro-ph/0607207, doi:10.1051/0004-6361:20066017.
- [72] L. Wyrzykowski et al. The OGLE View of Microlensing towards the Magellanic Clouds. IV. OGLE-III SMC Data and Final Conclusions on MACHOs. *Mon. Not. Roy. Astron. Soc.*, 416:2949, 2011. arXiv:1106.2925, doi:10.1111/j.1365-2966.2011.19243.x.
- [73] Craig J. Copi, David N. Schramm, and Michael S. Turner. Big bang nucleosynthesis and the baryon density of the universe. *Science*, 267:192–199, 1995. arXiv:astro-ph/9407006, doi:10.1126/science.7809624.

- [74] Scott Dodelson and Lawrence M. Widrow. Sterile-neutrinos as dark matter. *Phys. Rev. Lett.*, 72:17–20, 1994. [arXiv:hep-ph/9303287](#), doi:10.1103/PhysRevLett.72.17.
- [75] L. J. Rosenberg and K. A. van Bibber. Searches for invisible axions. *Phys. Rept.*, 325:1–39, 2000. doi:10.1016/S0370-1573(99)00045-9.
- [76] Andreas Birkedal-Hansen and Jay G. Wacker. Scalar dark matter from theory space. *Phys. Rev.*, D69:065022, 2004. [arXiv:hep-ph/0306161](#), doi:10.1103/PhysRevD.69.065022.
- [77] Hsin-Chia Cheng and Ian Low. TeV symmetry and the little hierarchy problem. *JHEP*, 09:051, 2003. [arXiv:hep-ph/0308199](#), doi:10.1088/1126-6708/2003/09/051.
- [78] Gerard Jungman, Marc Kamionkowski, and Kim Griest. Supersymmetric dark matter. *Phys. Rept.*, 267:195–373, 1996. [arXiv:hep-ph/9506380](#), doi:10.1016/0370-1573(95)00058-5.
- [79] Edward W. Kolb and Richard Slansky. Dimensional Reduction in the Early Universe: Where Have the Massive Particles Gone? *Phys. Lett.*, B135:378, 1984. doi:10.1016/0370-2693(84)90298-3.
- [80] Nima Arkani-Hamed, Savas Dimopoulos, and G. R. Dvali. The Hierarchy problem and new dimensions at a millimeter. *Phys. Lett.*, B429:263–272, 1998. [arXiv:hep-ph/9803315](#), doi:10.1016/S0370-2693(98)00466-3.
- [81] Lisa Randall and Raman Sundrum. A Large mass hierarchy from a small extra dimension. *Phys. Rev. Lett.*, 83:3370–3373, 1999. [arXiv:hep-ph/9905221](#), doi:10.1103/PhysRevLett.83.3370.
- [82] Thomas Appelquist, Hsin-Chia Cheng, and Bogdan A. Dobrescu. Bounds on universal extra dimensions. *Phys. Rev.*, D64:035002, 2001. [arXiv:hep-ph/0012100](#), doi:10.1103/PhysRevD.64.035002.
- [83] Hsin-Chia Cheng, Konstantin T. Matchev, and Martin Schmaltz. Radiative corrections to Kaluza-Klein masses. *Phys. Rev.*, D66:036005, 2002. [arXiv:hep-ph/0204342](#), doi:10.1103/PhysRevD.66.036005.
- [84] M. et al. Aartsen. Search for dark matter annihilations in the sun with the 79-string IceCube detector. *Phys. Rev. Lett.*, 110:131302, Mar 2013. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.110.131302>, doi:10.1103/PhysRevLett.110.131302.
- [85] T. Tanaka et al. (The Super-Kamiokande Collaboration). An indirect search for weakly interacting massive particles in the sun using 3109.6days of upward-going muons in super-kamiokande. *The Astrophysical Journal*, 742(2):78, 2011. URL: <http://stacks.iop.org/0004-637X/742/i=2/a=78>.

- [86] M. Ackermann et al. Search for gamma-ray spectral lines with the Fermi large area telescope and dark matter implications. *Phys. Rev.*, D88:082002, 2013. arXiv: 1305.5597, doi:10.1103/PhysRevD.88.082002.
- [87] D. S. Akerib et al. Results on the Spin-Dependent Scattering of Weakly Interacting Massive Particles on Nucleons from the Run 3 Data of the LUX Experiment. *Phys. Rev. Lett.*, 116(16):161302, 2016. arXiv:1602.03489, doi:10.1103/PhysRevLett.116.161302.
- [88] Andrew Askew, Sushil Chauhan, Bjorn Penning, William Shepherd, and Mani Tripathi. Searching for Dark Matter at Hadron Colliders. *Int. J. Mod. Phys.*, A29:1430041, 2014. arXiv:1406.5662, doi:10.1142/S0217751X14300415.
- [89] Georges Aad et al. Search for new phenomena in final states with an energetic jet and large missing transverse momentum in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *Eur. Phys. J.*, C75(7):299, 2015. [Erratum: Eur. Phys. J.C75,no.9,408(2015)]. arXiv:1502.01518, doi:10.1140/epjc/s10052-015-3517-3, 10.1140/epjc/s10052-015-3639-7.
- [90] Vardan Khachatryan et al. Search for dark matter, extra dimensions, and unparticles in monojet events in protonproton collisions at $\sqrt{s} = 8$ TeV. *Eur. Phys. J.*, C75(5):235, 2015. arXiv:1408.3583, doi:10.1140/epjc/s10052-015-3451-4.
- [91] Georges Aad et al. Search for dark matter in events with heavy quarks and missing transverse momentum in pp collisions with the ATLAS detector. *Eur. Phys. J.*, C75(2):92, 2015. arXiv:1410.4031, doi:10.1140/epjc/s10052-015-3306-z.
- [92] Vardan Khachatryan et al. Search for Monotop Signatures in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV. *Phys. Rev. Lett.*, 114(10):101801, 2015. arXiv:1410.1149, doi:10.1103/PhysRevLett.114.101801.
- [93] Vardan Khachatryan et al. Search for the production of dark matter in association with top-quark pairs in the single-lepton final state in proton-proton collisions at $\text{sqrt}(s) = 8$ TeV. *JHEP*, 06:121, 2015. arXiv:1504.03198, doi: 10.1007/JHEP06(2015)121.
- [94] Georges Aad et al. Search for new phenomena in events with a photon and missing transverse momentum in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *Phys. Rev.*, D91(1):012008, 2015. [Erratum: Phys. Rev.D92,no.5,059903(2015)]. arXiv:1411.1559, doi:10.1103/PhysRevD.92.059903, 10.1103/PhysRevD.91.012008.
- [95] Serguei Chatrchyan et al. Search for Dark Matter and Large Extra Dimensions in pp Collisions Yielding a Photon and Missing Transverse Energy. *Phys. Rev. Lett.*, 108:261803, 2012. arXiv:1204.0821, doi:10.1103/PhysRevLett.108.261803.

- [96] Vardan Khachatryan et al. Search for new phenomena in monophoton final states in proton-proton collisions at $\sqrt{s} = 8$ TeV. *Phys. Lett.*, B755:102–124, 2016. [arXiv:1410.8812](#), doi:10.1016/j.physletb.2016.01.057.
- [97] Vardan Khachatryan et al. Search for physics beyond the standard model in final states with a lepton and missing transverse energy in proton-proton collisions at $\sqrt{s} = 8$ TeV. *Phys. Rev.*, D91(9):092005, 2015. [arXiv:1408.2745](#), doi:10.1103/PhysRevD.91.092005.
- [98] Georges Aad et al. Search for new particles in events with one lepton and missing transverse momentum in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *JHEP*, 09:037, 2014. [arXiv:1407.7494](#), doi:10.1007/JHEP09(2014)037.
- [99] Georges Aad et al. Search for dark matter in events with a Z boson and missing transverse momentum in pp collisions at $\sqrt{s}=8$ TeV with the ATLAS detector. *Phys. Rev.*, D90(1):012004, 2014. [arXiv:1404.0051](#), doi:10.1103/PhysRevD.90.012004.
- [100] Georges Aad et al. Search for dark matter in events with a hadronically decaying W or Z boson and missing transverse momentum in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *Phys. Rev. Lett.*, 112(4):041802, 2014. [arXiv:1309.4017](#), doi:10.1103/PhysRevLett.112.041802.
- [101] Vardan Khachatryan et al. Search for dark matter and unparticles produced in association with a Z boson in proton-proton collisions at $\sqrt{s} = 8$ TeV. *Phys. Rev.*, D93(5):052011, 2016. [arXiv:1511.09375](#), doi:10.1103/PhysRevD.93.052011.
- [102] Asher Berlin, Tongyan Lin, and Lian-Tao Wang. Mono-Higgs Detection of Dark Matter at the LHC. *JHEP*, 06:078, 2014. [arXiv:1402.7074](#), doi:10.1007/JHEP06(2014)078.
- [103] Georges Aad et al. Search for Dark Matter in Events with Missing Transverse Momentum and a Higgs Boson Decaying to Two Photons in pp Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector. *Phys. Rev. Lett.*, 115(13):131801, 2015. [arXiv:1506.01081](#), doi:10.1103/PhysRevLett.115.131801.
- [104] Georges Aad et al. Search for dark matter produced in association with a Higgs boson decaying to two bottom quarks in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *Phys. Rev.*, D93(7):072007, 2016. [arXiv:1510.06218](#), doi:10.1103/PhysRevD.93.072007.
- [105] The ATLAS collaboration. Search for Dark Matter in association with a Higgs boson decaying to b -quarks in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. 2016.
- [106] CMS collaboration. Search for dark matter produced in association with a higgs boson decaying to a pair of w+w- bosons in pp collisions at 13 tev. CMS Analysis Note 2015/338, 2015.

- [107] CMS collaboration. Search for dark matter produced in association with a higgs boson decaying to two photons. CMS Analysis Note 2015/203, 2015.
- [108] CMS collaboration. Search for dark matter candidate in higgs to bb + missing transverse energy (met) (mono-h) final state. CMS Analysis Note 2015/209, 2015.
- [109] Serguei Chatrchyan et al. Measurement of the properties of a Higgs boson in the four-lepton final state. *Phys. Rev.*, D89(9):092007, 2014. [arXiv:1312.5353](https://arxiv.org/abs/1312.5353), doi: 10.1103/PhysRevD.89.092007.
- [110] ATLAS collaboration. Measurements of the properties of the Higgs-like boson in the four lepton decay channel with the ATLAS detector using 25 fb⁻¹ of proton-proton collision data. 2013.
- [111] CMS Collaboration. Measurement of the properties of the higgs boson in the four-lepton final state at s = 13 tev. CMS Analysis Note 2015/277, 2015.
- [112] CMS Collaboration. Studies of higgs boson production in the four-lepton final state at $\sqrt{s} = 13$ tev. CMS Physics Analysis Summary CMS-PAS-HIG-15-004, 2015. URL: <http://cdsweb.cern.ch/record/1279362>.
- [113] CMS Collaborations. H4l synchronization for 2015. URL: <https://twiki.cern.ch/twiki/bin/view/CMS/HiggsZZ4l2015Sync>.
- [114] John McDonald. Gauge singlet scalars as cold dark matter. *Phys. Rev.*, D50:3637–3649, 1994. [arXiv:hep-ph/0702143](https://arxiv.org/abs/hep-ph/0702143), doi: 10.1103/PhysRevD.50.3637.
- [115] Laura Lopez-Honorez, Thomas Schwetz, and Jure Zupan. Higgs portal, fermionic dark matter, and a Standard Model like Higgs at 125 GeV. *Phys. Lett.*, B716:179–185, 2012. [arXiv:1203.2064](https://arxiv.org/abs/1203.2064), doi: 10.1016/j.physletb.2012.07.017.
- [116] G. Belanger, B. Dumont, U. Ellwanger, J. F. Gunion, and S. Kraml. Status of invisible Higgs decays. *Phys. Lett.*, B723:340–347, 2013. [arXiv:1302.5694](https://arxiv.org/abs/1302.5694), doi: 10.1016/j.physletb.2013.05.024.
- [117] J. et al. Beringer. Review of particle physics*. *Phys. Rev. D*, 86:010001, Jul 2012. URL: <http://link.aps.org/doi/10.1103/PhysRevD.86.010001>, doi: 10.1103/PhysRevD.86.010001.
- [118] T. Aaltonen et al. Search for new particles decaying into dijets in proton-antiproton collisions at s***(1/2) = 1.96-TeV. *Phys. Rev.*, D79:112002, 2009. [arXiv:0812.4036](https://arxiv.org/abs/0812.4036), doi: 10.1103/PhysRevD.79.112002.
- [119] Serguei Chatrchyan et al. Search for narrow resonances using the dijet mass spectrum in pp collisions at $\sqrt{s}=8$ TeV. *Phys. Rev.*, D87(11):114015, 2013. [arXiv:1302.4794](https://arxiv.org/abs/1302.4794), doi: 10.1103/PhysRevD.87.114015.
- [120] Nathaniel Craig, Jamison Galloway, and Scott Thomas. Searching for Signs of the Second Higgs Doublet. 2013. [arXiv:1305.2424](https://arxiv.org/abs/1305.2424).

- [121] Christopher D. Carone and Hitoshi Murayama. Possible light U(1) gauge boson coupled to baryon number. *Phys. Rev. Lett.*, 74:3122–3125, 1995. [arXiv:hep-ph/9411256](#), doi:10.1103/PhysRevLett.74.3122.
- [122] Kaustubh Agashe and Geraldine Servant. Baryon number in warped GUTs: Model building and (dark matter related) phenomenology. *JCAP*, 0502:002, 2005. [arXiv:hep-ph/0411254](#), doi:10.1088/1475-7516/2005/02/002.
- [123] Pavel Fileviez Perez and Mark B. Wise. Baryon and lepton number as local gauge symmetries. *Phys. Rev.*, D82:011901, 2010. [Erratum: *Phys. Rev.*D82,079901(2010)]. [arXiv:1002.1754](#), doi:10.1103/PhysRevD.82.079901, 10.1103/PhysRevD.82.011901.
- [124] We-Fu Chang, John N. Ng, and Jackson M. S. Wu. A Very Narrow Shadow Extra Z-boson at Colliders. *Phys. Rev.*, D74:095005, 2006. [Erratum: *Phys. Rev.*D79,039902(2009)]. [arXiv:hep-ph/0608068](#), doi:10.1103/PhysRevD.74.095005, 10.1103/PhysRevD.79.039902.
- [125] Maxim Pospelov, Adam Ritz, and Mikhail B. Voloshin. Secluded WIMP Dark Matter. *Phys. Lett.*, B662:53–61, 2008. [arXiv:0711.4866](#), doi:10.1016/j.physletb.2008.02.052.
- [126] Daniel Feldman, Zuowei Liu, and Pran Nath. The Stueckelberg Z-prime Extension with Kinetic Mixing and Milli-Charged Dark Matter From the Hidden Sector. *Phys. Rev.*, D75:115001, 2007. [arXiv:hep-ph/0702123](#), doi:10.1103/PhysRevD.75.115001.
- [127] Jonathan L. Feng, Huitzu Tu, and Hai-Bo Yu. Thermal Relics in Hidden Sectors. *JCAP*, 0810:043, 2008. [arXiv:0808.2318](#), doi:10.1088/1475-7516/2008/10/043.
- [128] Shrihari Gopalakrishna, Sunghoon Jung, and James D. Wells. Higgs boson decays to four fermions through an abelian hidden sector. *Phys. Rev.*, D78:055002, 2008. [arXiv:0801.3456](#), doi:10.1103/PhysRevD.78.055002.
- [129] John March-Russell, Stephen M. West, Daniel Cumberbatch, and Dan Hooper. Heavy Dark Matter Through the Higgs Portal. *JHEP*, 07:058, 2008. [arXiv:0801.3440](#), doi:10.1088/1126-6708/2008/07/058.
- [130] Adam Falkowski, Francesco Riva, and Alfredo Urbano. Higgs at last. *JHEP*, 11:111, 2013. [arXiv:1303.1812](#), doi:10.1007/JHEP11(2013)111.
- [131] Abdelhak Djouadi and Grégory Moreau. The couplings of the Higgs boson and its CP properties from fits of the signal strengths and their ratios at the 7+8 TeV LHC. *Eur. Phys. J.*, C73(9):2512, 2013. [arXiv:1303.6591](#), doi:10.1140/epjc/s10052-013-2512-9.

- [132] Pier Paolo Giardino, Kristjan Kannike, Isabella Masina, Martti Raidal, and Alessandro Strumia. The universal Higgs fit. *JHEP*, 05:046, 2014. [arXiv:1303.3570](https://arxiv.org/abs/1303.3570), doi:10.1007/JHEP05(2014)046.
- [133] John Ellis and Tevong You. Updated Global Analysis of Higgs Couplings. *JHEP*, 06:103, 2013. [arXiv:1303.3879](https://arxiv.org/abs/1303.3879), doi:10.1007/JHEP06(2013)103.
- [134] Lyndon Evans and Philip Bryant. Lhc machine. *Journal of Instrumentation*, 3(08):S08001, 2008. URL: <http://stacks.iop.org/1748-0221/3/i=08/a=S08001>.
- [135] 2008. URL: <http://press.cern/press-releases/2008/10/cern-releases-analysis-lhc-incident>.
- [136] 2012. URL: <http://home.cern/about/updates/2013/02/long-shutdown-1-exciting-times-ahead>.
- [137] 2012. URL: <http://home.cern/about/engineering/restarting-lhc-why-13-tev>.
- [138] Serguei Chatrchyan et al. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *JINST*, 9(10):P10009, 2014. [arXiv:1405.6569](https://arxiv.org/abs/1405.6569), doi:10.1088/1748-0221/9/10/P10009.
- [139] Torbjrn Sjstrand, Stephen Mrenna, and Peter Skands. Pythia 6.4 physics and manual. *Journal of High Energy Physics*, 2006(05):026, 2006. URL: <http://stacks.iop.org/1126-6708/2006/i=05/a=026>.
- [140] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. Madgraph 5: going beyond. *Journal of High Energy Physics*, 2011(6):1–40, 2011. URL: [http://dx.doi.org/10.1007/JHEP06\(2011\)128](https://dx.doi.org/10.1007/JHEP06(2011)128), doi:10.1007/JHEP06(2011)128.
- [141] C. F. Berger, Z. Bern, Lance J. Dixon, Fernando Febres Cordero, D. Forde, T. Gleisberg, H. Ita, D. A. Kosower, and D. Maitre. Multi-jet cross sections at NLO with BlackHat and Sherpa. In *QCD and high energy interactions. Proceedings, 44th Rencontres de Moriond, La Thuile, Italy, March 14-21, 2009*, pages 265–268, 2009. URL: <http://www-public.slac.stanford.edu/sciDoc/docMeta.aspx?slacPubNumber=slac-pub-13630>, arXiv:0905.2735.
- [142] Simone Alioli. NLO and Parton Showers: The POWHEG-BOX. In *Physics at the LHC2010. Proceedings, 5th Conference, PLHC2010, Hamburg, Germany, June 7-12, 2010*, pages 204–208, 2010. URL: <https://inspirehep.net/record/867509/files/arXiv:1009.2348.pdf>, arXiv:1009.2348.
- [143] *Geant4 user’s documents*. CERN, Geneva, 1998. URL: <https://cds.cern.ch/record/998155>.

- [144] S. Alioli, P. Nason, C. Oleari, and E. Re. NLO vector-boson production matched with shower in POWHEG. *JHEP*, 07:060, 2008. [arXiv:0805.4802](#), doi:10.1088/1126-6708/2008/07/060.
- [145] Paolo Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004. [arXiv:hep-ph/0409146](#), doi:10.1088/1126-6708/2004/11/040.
- [146] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *JHEP*, 11:070, 2007. [arXiv:0709.2092](#), doi:10.1088/1126-6708/2007/11/070.
- [147] E. Bagnaschi, G. Degrassi, P. Slavich, and A. Vicini. Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM. *JHEP*, 02:088, 2012. [arXiv:1111.2854](#), doi:10.1007/JHEP02(2012)088.
- [148] Paolo Nason and Carlo Oleari. NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG. *JHEP*, 02:037, 2010. [arXiv:0911.5299](#), doi:10.1007/JHEP02(2010)037.
- [149] Heribertus B. Hartanto, Barbara Jager, Laura Reina, and Doreen Wackerlo. Higgs boson production in association with top quarks in the POWHEG BOX. *Phys. Rev.*, D91(9):094003, 2015. [arXiv:1501.04498](#), doi:10.1103/PhysRevD.91.094003.
- [150] Gionata Luisoni, Paolo Nason, Carlo Oleari, and Francesco Tramontano. $HW^\pm/\text{HZ} + 0$ and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO. *JHEP*, 10:083, 2013. [arXiv:1306.2542](#), doi:10.1007/JHEP10(2013)083.
- [151] Yanyan Gao, Andrei V. Gritsan, Zijin Guo, Kirill Melnikov, Markus Schulze, and Nhan V. Tran. Spin determination of single-produced resonances at hadron colliders. *Phys. Rev. D*, 81:075022, 2010. [arXiv:1001.3396](#), doi:10.1103/PhysRevD.81.075022.
- [152] Richard D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015. [arXiv:1410.8849](#), doi:10.1007/JHEP04(2015)040.
- [153] Paolo Nason and Giulia Zanderighi. W^+W^- , WZ and ZZ production in the POWHEG-BOX-V2. *Eur. Phys. J.*, C74(1):2702, 2014. [arXiv:1311.1365](#), doi:10.1140/epjc/s10052-013-2702-5.
- [154] John M. Campbell and R. K. Ellis. MCFM for the Tevatron and the LHC. *Nucl. Phys. Proc. Suppl.*, 205:10, 2010. [arXiv:1007.3492](#), doi:10.1016/j.nuclphysbps.2010.08.011.
- [155] John M. Campbell, R. Keith Ellis, and Ciaran Williams. Bounding the Higgs width at the LHC using full analytic results for $gg \rightarrow e^-e^+\mu^-\mu^+$. *JHEP*, 04:060, 2014. [arXiv:1311.3589](#), doi:10.1007/JHEP04(2014)060.

- [156] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014. [arXiv:1405.0301](https://arxiv.org/abs/1405.0301), doi:10.1007/JHEP07(2014)079.
- [157] Vardan Khachatryan et al. Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8\text{TeV}$. *JINST*, 10:P06005, 2015. doi:10.1088/1748-0221/10/06/P06005.
- [158] Measurement of the properties of the higgs boson in the four-lepton final state at $\sqrt{s} = 13\text{ tev}$. CMS Physics Analysis Note CMS-AN-15-277, 2016.
- [159] New results on the study of higgs boson production in the four-lepton final state at $\sqrt{s} = 13\text{ tev}$. CMS Physics Analysis Note CMS-AN-16-217, 2016.
- [160] E. Brondolin and G. Petrucciani. Measurement of the muon-tracking efficiency using the tag and probe method on 2015 runii data. *CMS AN*, 2015/215, 2015.
- [161] BtagPOG. Usage of b/c tag objects for 13 tev data in 2016 and 80x mc. *CMS Twiki*, 2016. URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation80XReReco>.
- [162] CMS Collaboration. Met filter recommendations for run ii. URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFiltersRun2>.
- [163] Massimiliano Grazzini, Stefan Kallweit, and Dirk Rathlev. ZZ production at the LHC: Fiducial cross sections and distributions in NNLO QCD. *Phys. Lett. B*, 750:407 – 410, 2015. URL: <http://www.sciencedirect.com/science/article/pii/S0370269315007303>, arXiv:1507.06257, doi:<http://dx.doi.org/10.1016/j.physletb.2015.09.055>.
- [164] John M. Campbell, R. Keith Ellis, and Ciaran Williams. Vector boson pair production at the LHC. *JHEP*, 07:018, 2011. [arXiv:1105.0020](https://arxiv.org/abs/1105.0020), doi:10.1007/JHEP07(2011)018.
- [165] Marco Bonvini, Fabrizio Caola, Stefano Forte, Kirill Melnikov, and Giovanni Ridolfi. Signal-background interference effects in $gg \rightarrow H \rightarrow WW$ beyond leading order. *Phys. Rev. D*, 88:034032, 2013. [arXiv:1304.3053](https://arxiv.org/abs/1304.3053), doi:10.1103/PhysRevD.88.034032.
- [166] Kirill Melnikov and Matthew Dowling. Production of two Z-bosons in gluon fusion in the heavy top quark approximation. *Phys. Lett. B*, 744:43, 2015. [arXiv:1503.01274](https://arxiv.org/abs/1503.01274), doi:10.1016/j.physletb.2015.03.030.
- [167] Chong Sheng Li, Hai Tao Li, Ding Yu Shao, and Jian Wang. Soft gluon resummation in the signal-background interference process of $gg(\rightarrow h^*) \rightarrow ZZ$. 2015. [arXiv:1504.02388](https://arxiv.org/abs/1504.02388).

- [168] Giampiero Passarino. Higgs CAT. *Eur. Phys. J. C*, 74:2866, 2014. [arXiv:1312.2397](https://arxiv.org/abs/1312.2397), doi:10.1140/epjc/s10052-014-2866-7.
- [169] Stefano Catani and Massimiliano Grazzini. An NNLO subtraction formalism in hadron collisions and its application to Higgs boson production at the LHC. *Phys. Rev. Lett.*, 98:222002, 2007. [arXiv:hep-ph/0703012](https://arxiv.org/abs/hep-ph/0703012), doi:10.1103/PhysRevLett.98.222002.
- [170] Massimiliano Grazzini. NNLO predictions for the Higgs boson signal in the $H \rightarrow WW \rightarrow l\nu l\nu$ and $H \rightarrow ZZ \rightarrow 4l$ decay channels. *JHEP*, 02:043, 2008. [arXiv:0801.3232](https://arxiv.org/abs/0801.3232), doi:10.1088/1126-6708/2008/02/043.
- [171] Massimiliano Grazzini and Hayk Sargsyan. Heavy-quark mass effects in Higgs boson production at the LHC. *JHEP*, 09:129, 2013. [arXiv:1306.4581](https://arxiv.org/abs/1306.4581), doi:10.1007/JHEP09(2013)129.
- [172] LHC Higgs Cross Section Working Group. Handbook of LHC Higgs Cross Sections: 3. Higgs Properties. CERN Report CERN-2013-004, 2013. [arXiv:1307.1347](https://arxiv.org/abs/1307.1347), doi:10.5170/CERN-2013-004.
- [173] CMS Collaboration. Documentation of the roostats -based statistics tools for higgs pag. URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideHiggsAnalysisCombinedLimit>.
- [174] LHC experiments and ROOT team. Welcome to the roostats wiki. URL: <https://twiki.cern.ch/twiki/bin/view/RooStats/WebHome>.
- [175] CMS Collaboration. Cms higgs physics analysis group (higgs pag). URL: <https://twiki.cern.ch/twiki/bin/view/CMS/HiggsWG>.
- [176] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J.*, C71:1554, 2011. [Erratum: Eur. Phys. J.C73,2501(2013)]. [arXiv:1007.1727](https://arxiv.org/abs/1007.1727), doi:10.1140/epjc/s10052-011-1554-0, 10.1140/epjc/s10052-013-2501-z.