

**A measurement of $Wb\bar{b}$ production and a search for monophoton signals of dark matter
using the CMS detector at the CERN LHC**

by

Thomas Mastrianni Perry

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Physics)

at the

UNIVERSITY OF WISCONSIN-MADISON

2016

Date of final oral examination: 3 August 2016

The dissertation is approved by the following members of the Final Oral Committee:

Wesley Smith (Advisor), Professor, Physics

Sridhara Dasu, Professor, Physics

Matt Herndon, Professor, Physics

Yang Bai, Professor, Physics

David C. Schwartz, Professor, Chemistry

CONTENTS

	Page
List of Figures	iv
List of Tables	v
1 Quantum Field Theory and the Standard Model	1
1.1 Local Quantum Field Theory	2
1.1.1 Representations of $SU(2)$	2
1.1.2 Yang-Mills Theory	5
1.2 The Standard Model	8
1.2.1 Gauge symmetry in $SU(3) \times SU(2) \times U(1)$	8
1.2.2 Symmetry breaking in $SU(2) \times U(1) \rightarrow U(1)$	10
1.2.3 Yukawa couplings and the CKM matrix	12
1.3 Interpreting the Standard Model	15
1.3.1 Scattering amplitude and propagators	15
1.3.2 Path integrals and Feynman diagrams	16
1.3.3 Renormalization	18
1.3.4 Cross sections and decay rates	20
1.3.5 QCD and Proton Structure	23
1.4 Dark Matter	25
1.4.1 Experimental motivations	25
1.4.2 Simplified theoretical models	26
2 Phenomenology of Processes	27
2.1 The Standard Model process $pp \rightarrow Wb\bar{b} \rightarrow \ell\nu b\bar{b}$	27
2.1.1 $pp \rightarrow W$	27
2.1.2 $W \rightarrow \ell\nu$	28
2.1.3 $g \rightarrow b\bar{b}$	29
2.2 The Standard Model process $pp \rightarrow Z\gamma \rightarrow \nu\bar{\nu}\gamma$	30

2.2.1	$pp \rightarrow Z/\gamma$	30
2.2.2	$Z \rightarrow \nu\bar{\nu}$	31
2.2.3	$pp \rightarrow Z\gamma$	31
2.3	Dark matter models of $pp \rightarrow \gamma + \text{invisible}$	32
3	The LHC and CMS	34
3.1	The Large Hadron Collider	34
3.1.1	<i>LHC pre-acceleration</i>	35
3.1.2	<i>LHC acceleration</i>	35
3.2	The Compact Muon Solenoid Detector	39
3.2.1	<i>Geometry</i>	40
3.2.2	<i>Magnet</i>	41
3.2.3	<i>Tracking System</i>	42
3.2.4	<i>Electronic Calorimeter</i>	43
3.2.5	<i>Hadronic Calorimeter</i>	45
3.2.6	<i>Muon System</i>	47
4	Event Simulation and Reconstruction	50
4.1	Simulation of Events	50
4.1.1	<i>Monte Carlo Event Generation</i>	50
4.1.2	<i>Monte Carlo Generators</i>	52
4.1.3	<i>Detector Simulation</i>	53
4.2	Reconstruction of Events	54
4.2.1	<i>Track and Primary Vertex Reconstruction</i>	54
4.2.2	<i>Electron ID and Reconstruction</i>	55
4.2.3	<i>Photon ID and Reconstruction</i>	56
4.2.4	<i>Muon ID and Reconstruction</i>	56
4.2.5	<i>Jet ID and Secondary Vertices</i>	57
5	W+bb Cross Section Measurement [TODO - WESLEY COMMENTS / CUTFLOW]	59
5.1	Previous Measurements	59

5.2	Event Selection	59
5.3	Background Estimation	62
5.3.1	<i>QCD</i>	62
5.3.2	<i>W+jets: light and charm component</i>	64
5.3.3	<i>Top backgrounds</i>	64
5.3.4	<i>Z$\ell\bar{\ell}$ backgrounds</i>	67
5.4	Analysis Strategy	70
5.5	Systematic Uncertainties	73
5.6	Signal Extraction	74
5.7	Cross Section and Comparisons	77
6	Monophoton Analysis [TODO WESLEY COMMENTS]	85
6.1	Previous Measurements	85
6.2	Event Selection	86
6.3	Estimation of Contributions	88
6.3.1	<i>Reweighting</i>	89
6.3.2	<i>Vγ Estimates</i>	89
6.3.3	<i>Elecron Mis-ID</i>	90
6.3.4	<i>Non-collision Backgrounds</i>	91
6.3.5	<i>Minor SM Processes</i>	91
6.4	Results	92
6.4.1	<i>pp $\rightarrow Z\gamma \rightarrow \nu\bar{\nu}\gamma$ Cross Section Measurement</i>	92
6.4.2	<i>Limits on Dark Matter</i>	94

LIST OF FIGURES

1.1	Illustrations of SM couplings	19
1.2	Feynman diagrams for $2 \rightarrow 2$ scattering	19
1.3	One-loop corrections to vertices and propagator	20
2.1	Feynman diagrams for $pp \rightarrow Wb\bar{b} \rightarrow \ell\nu b\bar{b}$	27
2.2	Feynman diagrams for $pp \rightarrow Z\gamma \rightarrow \nu\bar{\nu}\gamma$	30
2.3	Feynman diagrams for dark matter monophoton	33
3.1	The LCH accelerator complex at CERN	36
3.2	The LHC dipole	37
3.3	The CMS Detector	40
3.4	The CMS Tracking System	43
3.5	The CMS Electromagnetic Calorimeter	45
3.6	The CMS Hadronic Calorimeter	47
3.7	The CMS Muon System	49
4.1	Example Parton Distribution Function from CTEQ6M	51
5.1	QCD shape for $Wb\bar{b}$ analysis	63
5.2	Wjj control region for the $Wb\bar{b}$ measurement	65
5.3	$t\bar{t}$ -multijet control region for the $Wb\bar{b}$ measurement	66
5.4	$t\bar{t}$ -multilepton control region for the $Wb\bar{b}$ measurement	68
5.5	Single-top control region for the $Wb\bar{b}$ measurement	69
5.6	$Zb\bar{b}$ control region for the $Wb\bar{b}$ analysis	70
5.7	Zjj control region for the $Wb\bar{b}$ measurement	71
5.8	$t\bar{t}$ -multijet control region after fitting for b-tag scale	75
5.9	$t\bar{t}$ -multilepton control region after fitting for JES	76
5.10	Fitted m_T in the $Wb\bar{b}$ signal region	77
5.11	Fitted $\Delta R(b\bar{b})$ and p_T^ℓ in the $Wb\bar{b}$ signal region	78
5.12	Cross section comparison for $Wb\bar{b}$ and generators	82

6.1	Distributions of p_T and E_T^{miss} in the $pp \rightarrow \gamma + \text{invisible}$ analysis	96
6.2	Distributions of p_T/E_T^{miss} and n_{jets} in the $pp \rightarrow \gamma + \text{invisible}$ analysis	97

LIST OF TABLES

1.1	Representations of $SU(2) \times SU(2)$	3
1.2	Standard Model quarks and leptons	9
1.3	Covariant derivatives and fermion couplings	9
1.4	Fundamental particle decay channels and rates	23
5.1	Systematic uncertainties in $Wb\bar{b}$	83
5.2	Initial and final yields in the $Wb\bar{b}$ signal region	84
5.3	Measured cross sections for $pp \rightarrow Wb\bar{b} \rightarrow \ell\nu b\bar{b}$	84
6.1	Estimated yields for $pp \rightarrow \gamma + \text{invisible}$	92
6.2	Summary of systematic uncertainties for signal and different background sources.	93
6.3	Observed (expected) 90%CL upper limits on the dark-matter production cross section σ for mediator mass 10 TeV.	94
6.4	Observed (expected) 90%CL upper limits on the dark-matter production cross section σ for fixed mediator mass 10 TeV	95

1 QUANTUM FIELD THEORY AND THE STANDARD MODEL

The Standard Model of particle physics (SM) is useful. It is a local Quantum Field Theory (QFT) representing the forefront of contemporary understanding of nature on its finest level and is simultaneously the most quantitatively verified physical model of the constituent elements of the universe and known to be an incomplete description. It is therefore one of the goals of modern society to experimentally investigate particles and the interactions between particles within the context of the SM to validate the theory where possible and to guide directions for its extension where necessary. To achieve this goal, the governments from nearly 100 different countries, states and territories have funded tens of thousands of scientists, engineers and technicians to build, operate, maintain and analyze data from the Large Hadron Collider (LHC) at the European Center for Nuclear Research (CERN). This thesis presents analyses of data taken with the Compact Muon Solenoid (CMS) detector using proton-proton collisions provided by the LHC during its operation in 2012 and 2015.

1.1 Local Quantum Field Theory

1.1.1 Representations of $SU(2)$

One of the key underlying principles behind any QFT is that of symmetry. In particular, QFTs arise from the combination of quantum mechanics with Lorentz symmetry which ensures that the equations used to describe the laws of physics remain equivalently valid in all inertial reference frames. Local fields are therefore required to transform as representations of the Lorentz group, namely rotations, J_a , and boosts, K_a where $a \in \{1, 2, 3\}$ for the three spatial dimensions. Boosts transform as vectors under rotation and the two obey the Lie algebras given in Equation 1.1, where ϵ_{abc} is the Levi-Civita symbol.

$$[J_a, J_b] = i\epsilon_{abc}J_c, \quad [K_a, K_b] = -i\epsilon_{abc}J_c, \quad [J_a, K_b] = i\epsilon_{abc}K_c . \quad (1.1)$$

Both J_a and K_a are hermitian, but it is natural to define the non-hermitian objects

$$L_a = \frac{1}{\sqrt{2}}(J_a + iK_a), \quad R_a = \frac{1}{\sqrt{2}}(J_a - iK_a) \quad (1.2)$$

which commute with each other and each independently obey the commutation relations of $SU(2)$,

$$[L_a, L_b] = i\epsilon_{abc}L_c, \quad [R_a, R_b] = i\epsilon_{abc}R_c, \quad [L_a, R_b] = 0 . \quad (1.3)$$

Table 1.1: Selected representations of the Lorentz group are formed by combining representations of $SU(2)$ in the structure $SU(2)_L \times SU(2)_R$. The first column labels the quantum numbers, λ , and the second column translates this into dimensionality. The third column indicates the kind of object which meets the symmetry requirements.

(λ_L, λ_R)	(d_L, d_R)	Name	
(0, 0)	(1, 1)	ϕ	Scalar
$(1/2, 0)$	(2, 1)	ψ_L	Left-handed Weyl spinor
$(0, 1/2)$	(1, 2)	ψ_R	Right-handed Weyl spinor
$(1/2, 1/2)$	(2, 2)	A_μ	Gauge potential

Because of this, the Lorentz group may be expressed in terms of representations of $SU(2)_L \times SU(2)_R$.

The group $SU(2)$ can be thought of as the set of 2×2 complex matrices with unit determinant under the operation of matrix multiplication, and the $(2^2 - 1)$ generators of the group are proportional to the Pauli matrices,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (1.4)$$

which have been diagonalized along the 3 direction and combine with the unit matrix to form $\sigma^\mu = (\mathbf{1}, \vec{\sigma})$. Representations of $SU(2)$ are labelled by a quantum number, $\lambda \in \{0, 1/2, 1, 3/2, \dots\}$, with the fundamental representation being $\lambda = 1/2$ and the dimensionality, d , of a representation being set by $d = 2\lambda + 1$. Representations of the full symmetry are therefore constructed by combining representations from the L and R components, and Table 1.1 illustrates those combinations of lowest dimensionality.

Rotations take spatial coordinates into spatial coordinates, but boosts mix spatial

coordinates with time, which have opposite signs in the spacetime metric. Therefore, under the parity operation, P , which inverts the signs of spatial coordinates,

$$P : J \rightarrow J, \quad P : K \rightarrow -K , \quad (1.5)$$

while, under conjugation, C , because J and K are hermitian,

$$C : J \rightarrow J, \quad C : K \rightarrow K . \quad (1.6)$$

Using Equation 1.2 then, the $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$ representations transform as

$$P : \psi_{L(R)} \rightarrow \psi_{R(L)}, \quad C : \psi_{L(R)} \rightarrow \sigma_2 \psi_{R(L)}^*, \quad CP : \psi_{L(R)} \rightarrow \sigma_2 \psi_{L(R)}^* . \quad (1.7)$$

which illustrates the point that because ψ_L and ψ_R can be interchanged via these discrete transformations, they are not independent. The full $SU(2)_L \times SU(2)_R$ symmetry can thus be maintained by considering only one of the two, and it is customary to work with ψ_L , the ‘left’ component, henceforth simply denoted as ψ , with $\sigma_2 \psi_R^*$ denoted as $\bar{\psi}$. An object which transforms as a vector can be made from spinors by noting¹ that $(\frac{1}{2}, 0) \times (0, \frac{1}{2}) = (\frac{1}{2}, \frac{1}{2})$. Therefore $\psi^\dagger \sigma^\mu \psi$ transforms as a Lorentz vector, and can be used to make a Lorentz-invariant object by contracting the spacetime index, μ . The derivative operator $\partial_\mu = (\partial/\partial t, \partial/\partial \vec{x})$ is translation-invariant and does not give a surface contribution if inserted between the two instances of ψ ,

¹ $d = 2 \times 2(\frac{1}{2}) + 1 = 4$ for (t, x, y, z)

motivating the canonical kinetic term for spinors,

$$i\psi^\dagger \sigma^\mu \partial_\mu \psi . \quad (1.8)$$

1.1.2 Yang-Mills Theory

The above construction for canonical kinetic terms in the Lagrangian was generalized by Chen Ning Yang and Robert Mills for N spinor fields as

$$i \sum_{a=1}^N \psi^{a\dagger} \sigma^\mu \partial_\mu \psi_a = i\Psi^\dagger \sigma^\mu \partial_\mu \Psi , \quad (1.9)$$

and while the existence of this object is imposed by the Lorentz symmetry, further symmetries can be made apparent. A simple example is global phase invariance. For λ which is not a function of x , the transformation $\Psi \rightarrow e^{i\lambda} \Psi$ leaves the kinetic term in Equation 1.9 unchanged. The derivative passes through $e^{i\lambda}$ which combines with $e^{-i\lambda}$ from the transform on Ψ^\dagger to make the unit.

A more complicated example is gauge invariance. The symmetry of global phase invariance can be extended by introducing $(N^2 - 1)$ traceless hermitian matrices, λ_A which satisfy the Lie algebra,

$$[\lambda_A, \lambda_B] = if_{ABC}\lambda_C, \quad \text{Tr}(\lambda_A \lambda_B) = \frac{1}{2}\delta_{AB} , \quad (1.10)$$

where f_{ABC} are the structure functions for $SU(N)$ and δ_{AB} is the Kronecker delta function. For $N = 2$, $f_{ABC} = \epsilon_{abc}$ used in Equations 1.1 and λ_A are the Pauli matrices from Equations 1.4. For $N = 3$, f_{ABC} is also a totally anti-symmetric operator and

λ_A are the Gell-Mann matrices.

These λ_A can each be used to construct a hermitian ($N \times N$) matrix. One degree of freedom is factored out as the overall phase of the trace, demonstrated to be a symmetry of Ψ . This leaves $(N - 1)$ degrees from the real diagonal elements and $2 \times (N^2 - N)/2$ from the unique complex off-diagonal elements. These $(N^2 - 1)$ overall degrees of freedom correspond to the λ_A which are used to construct

$$\mathbf{H} = \frac{1}{2} \sum_{A=1}^{N^2-1} \omega_A \lambda_A \quad (1.11)$$

using ω_A as the parameter of expansion.

\mathbf{H} can then be used to produce a unitary matrix

$$\mathbf{U} = e^{i\mathbf{H}} \quad (1.12)$$

whose unitarity is evident by the hermiticity of \mathbf{H} . Unitarity is crucial for the final cancellation to leave Equation 1.9 invariant under the gauge transformation $\Psi \rightarrow \mathbf{U}\Psi$, but to ensure locality, the particular expansion, ω_a , could have dependence on position and thus interact with ∂_μ .

To accommodate for this, the derivative is generalized into an ($N \times N$) covariant derivative matrix, \mathbf{D}_μ , which has the property

$$\mathbf{D}'_\mu \mathbf{U} \Psi = \mathbf{U} \mathbf{D}_\mu \Psi \quad (1.13)$$

and thus behaves under gauge transformations as

$$\Psi \rightarrow \mathbf{U}\Psi, \quad \mathbf{D}_\mu \rightarrow \mathbf{D}'_\mu = \mathbf{U}\mathbf{D}_\mu\mathbf{U}^\dagger \quad . \quad (1.14)$$

Explicitly, \mathbf{D}_μ is constructed using $(N^2 - 1)$ gauge potentials, A_μ , in the form of a hermitian matrix,

$$\mathbf{A}_\mu = \frac{1}{2} \sum_{B=1}^{N^2-1} A_\mu^B \boldsymbol{\lambda}_B \quad (1.15)$$

which transforms under gauge transformations as

$$\mathbf{A}_\mu \rightarrow \mathbf{A}'_\mu = -i\mathbf{U}\partial_\mu\mathbf{U}^\dagger - i\mathbf{U}\mathbf{A}_\mu\mathbf{U}^\dagger \quad . \quad (1.16)$$

The field strength matrix is formed by taking the commutator

$$\mathbf{F}_{\mu\nu} = -i[\mathbf{D}_\mu, \mathbf{D}_\nu] \quad (1.17)$$

which leads to a kinetic term that is invariant under both gauge and Lorentz transformations, $-1/2g^2\text{Tr}(\mathbf{F}_{\mu\nu}\mathbf{F}^{\mu\nu})$ where g is the coupling strength. This leads to the Yang-Mills Lagrange density of

$$\mathcal{L}_{\text{YM}} = -\frac{1}{2}\text{Tr}(\mathbf{F}_{\mu\nu}\mathbf{F}^{\mu\nu}) + i\Psi^\dagger\sigma^\mu\mathbf{D}_\mu\Psi + i\bar{\Psi}^\dagger\sigma^\mu\overline{\mathbf{D}}_\mu\bar{\Psi} \quad . \quad (1.18)$$

For $N = 2$, this is the Quantum ElectroDynamics (QED) Lagrangian of electromagnetic interactions with a massless electron, and for $N = 3$, this is Quantum ChromoDynamics (QCD).

1.2 The Standard Model

1.2.1 Gauge symmetry in $SU(3) \times SU(2) \times U(1)$

The Standard Model (SM) Lagrangian uses the Yang-Mills construction on the group $SU(3) \times SU(2) \times U(1)$. Strong interactions are described by $SU(3)$ (color) which has eight gluons, G_μ^A , $A \in 1, 2, \dots, 8$. Electroweak interactions incorporate the mixing of $SU(2) \times U(1)$ via the Higgs mechanism with $SU(2)$ (weak charge) having three weak bosons, W_μ^a , $a \in 1, 2, 3$ and $U(1)$ (hypercharge) having one hyperon B_μ . The field strengths are independent, with independent coupling constants g_3, g_2, g_1 , each calculated using the appropriate covariant derivative in Equation 1.17. This is part of the SM Lagrangian describing the gauge bosons,

$$\mathcal{L}_{\text{bosons}} = -\frac{1}{4g_3^2}G_{\mu\nu}^AG^{\mu\nu A} - \frac{1}{4g_2^2}W_{\mu\nu}^aW^{\mu\nu a} - \frac{1}{4g_1^2}B_{\mu\nu}B^{\mu\nu} . \quad (1.19)$$

The matter (antimatter) components are the spin-1/2 fermions which are built from the L representation of Weyl spinors and are listed in Table 1.2. The portion of the SM Lagrangian coming from the fermion kinetic terms is found by exchanging the derivative for the appropriate covariant derivative in Equation 1.9 to be

$$\mathcal{L}_{\text{fermions}} = i \sum_{i=1}^3 \left(\mathbf{Q}_i^\dagger \sigma^\mu \mathbf{D}_\mu \mathbf{Q}_i + L_i^\dagger \sigma^\mu \mathbf{D}_\mu L_i + \bar{\mathbf{u}}_i^\dagger \sigma^\mu \mathbf{D}_\mu \bar{\mathbf{u}}_i + \bar{\mathbf{d}}_i^\dagger \sigma^\mu \mathbf{D}_\mu \bar{\mathbf{d}}_i \bar{e}_i^\dagger \sigma^\mu \mathbf{D}_\mu \bar{e}_i \right) \quad (1.20)$$

with the covariant derivatives described in Table 1.3.

Both quarks and leptons come in three families, known as generations for the quarks and flavors for the leptons. The leptons are colorless and therefore do not

Table 1.2: Fermions in the standard model consist of three families ($i \in 1, 2, 3$) of leptons and quarks in the singlet and doublet configurations. The symmetries are noted as $(SU(3)^c, SU(2))_Y$ where Y is chosen to satisfy the Gell-Mann-Nishijima formula, Equation 1.21.

Name	Symbol	I_3	Symmetry
Quark doublet	$\mathbf{Q}_i = \begin{pmatrix} u_i \\ d_i \end{pmatrix}$	$\begin{pmatrix} 1/2 \\ -1/2 \end{pmatrix}$	$(3^c, 2)_{1/3}$
Lepton doublet	$L_i = \begin{pmatrix} \nu_i \\ e_i \end{pmatrix}$	$\begin{pmatrix} 1/2 \\ -1/2 \end{pmatrix}$	$(1^c, 2)_{-1}$
Quark singlet	$\bar{\mathbf{u}}_i$	0	$(3^c, 1)_{-4/3}$
Quark singlet	$\bar{\mathbf{d}}_i$	0	$(3^c, 1)_{2/3}$
Lepton singlet	\bar{e}_i	0	$(1^c, 2)_2$

Table 1.3: The exact form of the covariant derivative used is determined by the couplings to gauge bosons the fermion has. Below, the fermions are listed, along with the associated derivative and interactions. On the bottom are the vectors of gauge potentials as defined in Equation 1.15.

Fermion type	Covariant derivative	Interactions		
		$SU(3)$	$SU(2)$	$U(1)$
All quarks (doublet)	$\mathbf{D}_\mu \mathbf{Q}_i = \left(\partial_\mu + i\mathbf{G}_\mu + i\mathbf{W}_\mu + \frac{i}{3}B_\mu \right) \mathbf{Q}_i$	Yes	Yes	Yes
All leptons (doublet)	$\mathbf{D}_\mu L_i = \left(\partial_\mu + i\mathbf{W}_\mu + \frac{i}{2}B_\mu \right) L_i$	No	Yes	Yes
u -type quarks (singlet)	$\mathbf{D}_\mu \bar{\mathbf{u}}_i = \left(\partial_\mu - i\mathbf{G}_\mu^* - \frac{2i}{3}B_\mu \right) \bar{\mathbf{u}}_i$	Yes	No	Yes
d -type quarks (singlet)	$\mathbf{D}_\mu \bar{\mathbf{d}}_i = \left(\partial_\mu - i\mathbf{G}_\mu^* + \frac{i}{3}B_\mu \right) \bar{\mathbf{d}}_i$	Yes	No	Yes
Charged leptons (singlet)	$\mathbf{D}_\mu \bar{e}_i = \left(\partial_\mu + iB_\mu \right) \bar{e}_i$	No	No	Yes

$\mathbf{G}_\mu = G_\mu^A \lambda^A / 2, \quad \mathbf{G}_\mu^* = G_\mu^A \lambda^{A*} / 2, \quad \mathbf{W}_\mu = W_\mu^a \sigma^a / 2$

couple to gluons, but the quarks do and additionally come in three color varieties for each generation. The singlet configurations contain only electrically charged fermions,

and all fermions have their charge set by the Gell-Mann-Nishijima formula,

$$Q = I_3 + \frac{Y}{2} \quad , \quad (1.21)$$

where I_3 is the component of weak isospin, $SU(2)$, along the direction in which the Pauli matrices in Equations 1.4 are diagonalized.

All quarks and leptons can exist in $SU(2)$ doublet configurations, and the different I_3 values further break quarks and leptons into types. For quarks, there are u -type (u, c, t) = (up, charm, top) which have charge $+2/3$, and there are d -type (d, s, b) = (down, strange, bottom) which have charge $-1/3$. Leptons are either charged (e, μ, τ) = (electron, muon, tauon) or neutral (ν_e, ν_μ, ν_τ) = (electron-, mu-, tau-neutrino), and all charged fermions are arranged such that mass increases with successive generations within a given type. In units where $\hbar = c = 1$, the top ($m_t = 173$ GeV) and bottom ($m_b = 4$ GeV) are the heaviest quarks of their respective types, with the bottom weighing three orders of magnitude greater than the lightest quark, up ($m_u = 2$ MeV). The mass separation for the charged leptons ($m_\tau = 1.7$ GeV, $m_e = 0.5$ MeV) also spans multiple orders of magnitude, but while the observations of neutrino oscillations indicate that neutrinos have mass, only upper limits on the values they may have have been set on the order of MeV.

1.2.2 Symmetry breaking in $SU(2) \times U(1) \rightarrow U(1)$

Fermions and gauge bosons are massless as written in $\mathcal{L}_{\text{fermions}}$ and $\mathcal{L}_{\text{bosons}}$, but are observed to be massive in nature and can acquire mass through the Brout-Englert-

Higgs mechanism. A scalar (spin-zero) field as given in the top row of Table 1.1 is introduced in a doublet configuration. This adds a Higgs term to the SM Lagrangian,

$$\mathcal{L}_{\text{Higgs}} = (\mathbf{D}_\mu H)^\dagger \mathbf{D}_\mu H + m_H H^\dagger H - \lambda (H^\dagger H)^2 , \quad (1.22)$$

where the first term is the canonical kinetic term for a scalar, the second term generates the mass of the Higgs boson (m_H) and the third term is a potential. This is minimized when

$$H_0^\dagger H_0 = \frac{m^2}{2\lambda} = \frac{v^2}{2}, \quad H_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} . \quad (1.23)$$

Taking perturbations about the minimum vacuum expectation value using the Kibble parameterization, $v \rightarrow v + h(x)$ and the unitary matrix containing three Nambu-Goldstone bosons from the symmetry breaking is factored out, yielding

$$H = \frac{1}{\sqrt{2}} \mathbf{U}(x) \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} . \quad (1.24)$$

In this gauge, the weak bosons, $W_{\mu\nu}^a$, and the hyperion, B_μ appear in Equation 1.25 as linear combinations as the massive W^\pm and Z bosons, and as the massless photon respectively,

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}} (W_\mu^1 \mp iW_\mu^2), & Z_\mu &= \cos \theta_W W_\mu^3 - \sin \theta_W B_\mu, & A_\mu &= \sin \theta_W W_\mu^3 + \cos \theta_W B_\mu \\ M_Z &= \frac{v}{2} \sqrt{g_1^2 + g_2^2}, & M_W &= \cos \theta_W M_Z \end{aligned} . \quad (1.25)$$

Here, the Weinberg angle, $\theta_W = g_1/g_2$, is the ratio between the $U(1)$ and $SU(2)$. The \pm in W_μ^\pm aligns with the electric charge of the W boson, and the Z_μ and A_μ are both electrically neutral and orthogonal.

The vector boson couplings to the fermions are then expressed as three types of currents. Electromagnetic interactions, J_μ^γ , involve couplings between the photon, γ , and charged particles, and any particle which interacts with the photon can also interact with the Z boson. The neutral current interactions, J_μ^Z , have additional couplings to the lepton and quark doublets and the charged current interactions, J_μ^\pm , are between the doublets and the W bosons. Couplings to each of the bosons are scaled by a coupling factor which decides the relative strengths of the interactions, and the Lagrangian for the currents is given in Equation 1.26 where i indicates the generation.

$$\begin{aligned}
\mathcal{L}_{\text{currents}} &= eA^\mu J_\mu^\gamma + g_Z Z^\mu J_\mu^Z + g_W W^{+\mu} J_\mu^- + G_W W^{-\mu} J_\mu^+ \\
J_\mu^\gamma &= \bar{e}_i^\dagger \sigma_\mu e_i + e_i^\dagger \sigma_\mu e_i + \frac{2}{3} \mathbf{Q}_{i1}^\dagger \sigma_\mu \mathbf{Q}_{i1} - \frac{1}{3} \mathbf{Q}_{i2}^\dagger \sigma_\mu \mathbf{Q}_{i2} - \frac{2}{3} \bar{\mathbf{u}}_i^\dagger \sigma_\mu \bar{\mathbf{u}}_i + \frac{1}{3} \bar{\mathbf{d}}_i^\dagger \sigma_\mu \bar{\mathbf{d}}_i \\
J_\mu^Z &= L_i^\dagger \frac{\tau_3}{2} \sigma_\mu L_i + \mathbf{Q}_i^\dagger \frac{\tau_3}{2} \sigma_\mu \mathbf{Q}_i - \sin^2 \theta_W J_\mu^\gamma \\
J_\mu^\pm &= L_i^\dagger \tau_\pm \sigma_\mu L_i + \mathbf{Q}_i^\dagger \tau_\pm \sigma_\mu \mathbf{Q}_i \\
e &= \frac{g_1 g_2}{\sqrt{g_1^2 + g_2^2}}, \quad g_Z = \frac{e}{\cos \theta_W \sin \theta_W}, \quad g_W = \frac{e}{\sqrt{2} \sin \theta_W}
\end{aligned} \tag{1.26}$$

1.2.3 Yukawa couplings and the CKM matrix

In addition to generating mass terms for the bosons, the Higgs field gives rise to fermion mass via Yukawa couplings between the Higgs doublet, H , from Equation 1.22 and the fermion doublets. A singlet configuration is possible to construct from

two $SU(2)$ doublets in their antisymmetric combination², of the general form (MN) where M and N are both L fermion fields. Each term has an associated Yukawa coupling, $\mathbf{Y}^f = y_{ij}$, where $f \in \{u, d, e\}$ for each of the charged fermion types, and this piece of the SM the Lagrangian is

$$\begin{aligned}\mathcal{L}_{\text{Yukakwa}} &= y_{ij}^u \mathbf{Q}_i \bar{\mathbf{u}}_j H + y_{ij}^d \mathbf{Q}_i \bar{\mathbf{d}}_j \bar{H} + y_{ij}^e L_i \bar{e}_j \bar{H} \\ &= \mathcal{L}_{\text{Yuk}}^u + \mathcal{L}_{\text{Yuk}}^d + \mathcal{L}_{\text{Yuk}}^e\end{aligned}. \quad (1.27)$$

In general, any (3×3) matrix can written in terms of a diagonal matrix D and two unitary matrices U and V , and in particular, the coupling matrices can be expressed as $\mathbf{Y}^f = \mathbf{U}^f \mathbf{D}^f \mathbf{V}^f$. The fermion Yukawa couplings are all of the form $\mathbf{Y}^f \mathbf{F} \bar{\mathbf{f}}$ and by incorporating \mathbf{V}^f into $\bar{\mathbf{f}}$, the may each be individually diagonalized as

$$\begin{aligned}\mathcal{L}_{\text{Yuk}}^e &= (m_e(\bar{e}e) + m_\mu(\bar{\mu}\mu) + m_\tau(\bar{\tau}\tau)) \beta(h) \\ \mathcal{L}_{\text{Yuk}}^d &= (m_d(\bar{d}d) + m_\mu(\bar{s}s) + m_\tau(\bar{b}b)) \beta(h) \\ \mathcal{L}_{\text{Yuk}}^u &= (m_u(\bar{u}u) + m_\mu(\bar{c}c) + m_\tau(\bar{t}t)) \beta(h)\end{aligned} \quad (1.28)$$

where β is a function of the Higgs coupling, h . The mass terms for the gauge bosons are also proportional to h . It is because the mass of the Higgs boson is tied to the masses of the gauge bosons and charged fermions that the discovery of the Higgs boson in 2012 was of such importance. The result was announced just a few months after I moved to Geneva to do research at CERN, and it provided the first measurement of a parameter in the SM that had been previously unknown, $m_H = 125\text{GeV}$.

Unlike the leptons, which contain only one charged type, the quarks have two

² The second term in $(1/2, 0) \times (1/2, 0) = (1, 0) + (0, 0)$

different spinor fields, **u** and **d**, which are different but both have Yukawa couplings to the Higgs with the same **Q**. This means that while the lepton sector $\mathcal{L}_{\text{Yuk}}^e$ can be diagonalized in mass simultaneously with either of $\mathcal{L}_{\text{Yuk}}^d$ or $\mathcal{L}_{\text{Yuk}}^u$, the quarks can not be simultaneously diagonalized. The mismatch between the two types of quarks is characterized by the unitary Cabibbo-Kobayashi-Maskawa (CKM) matrix,

$$\mathcal{U}_{\text{CKM}} = \mathbf{U}^d \overline{\mathbf{U}}^{u\dagger} = \begin{pmatrix} c_1 & +s_1c_3 & +s_1s_3 \\ -s_1c_2 & c_1c_2c_3 - s_2s_3e^{i\delta} & c_1c_2s_3 + s_2c_3e^{i\delta} \\ -s_1s_2 & c_1s_2c_3 + c_2s_3e^{i\delta} & c_1s_2s_3 - c_2c_3e^{i\delta} \end{pmatrix} \quad (1.29)$$

where s_a and c_a stand for $\sin \theta_a$ and $\cos \theta_a$. In this choice of basis, the charged current interactions from Equation 1.26 are

$$J_\mu^\pm = L_i^\dagger \tau_\pm \sigma_\mu L_i + \mathbf{d}_i^\dagger (\mathcal{U}_{\text{CKM}})_{ij} \tau_\pm \sigma_\mu \mathbf{u}_j \quad . \quad (1.30)$$

The CKM matrix thus is a matrix of coefficients for transforming between generations and types of quarks for charged current interactions. It also is the location in the SM where parity invariance is violated.

The laws of nature are postulated to be invariant under the combined operation CPT , where C and P are defined in Equations 1.5 and 1.6, and T is the antiunitary time-reversal operator,

$$T^{-1}iT = -i \quad (1.31)$$

Because of the complex phase, δ , in the CKM matrix, the couplings between the quarks and W bosons are not T invariant. To maintain CPT invariance, $\delta \neq 0$

implies that CP must be violated in such interactions.

1.3 Interpreting the Standard Model

1.3.1 Scattering amplitude and propagators

Initial and final states are built by successive applications of the raising (or particle creation) operator, a^\dagger , on the vacuum, $|0\rangle$, at the limit of initial states being measured infinitely far in the past and final states being measured infinitely far in the future.

$$|i\rangle = \lim_{t \rightarrow -\infty} \sum_i a_i^\dagger |0\rangle, \quad |f\rangle = \lim_{t \rightarrow +\infty} \sum_{i'} a_{i'}^\dagger |0\rangle \quad (1.32)$$

The overlap or scattering amplitude between these states, $\langle f|i\rangle$, can be expressed using the Lehmann-Symanzik-Zimmermann (LSZ) reduction formula in Equation 1.33 as factored into two pieces,

$$\langle f|i\rangle = i^{n_j+n_{j'}} \int \prod_j d^4x_j e^{ik_j x_j} (-\partial_j^2 + m^2) \times \langle 0|T \prod_j \phi(x_j)|0\rangle \quad (1.33)$$

The first piece is made of integrals over the momenta, k_j for $j \in j, j'$, of the n_j incoming and $n_{j'}$ outgoing particles. In these integrals, each plane wave multiplies the function $(-\partial_j^2 + m^2)$ where $\partial_j^2 = \partial_j^\mu \partial_{\mu j}$ is the contraction of the Lorentz-invariant derivative operator. The second piece of the LSZ formula is the vacuum expectation value of the time-ordered product of fields, $\langle 0|T \prod_j \phi(x_j)|0\rangle$.

By Wick's theorem, the time-ordered interaction of n fields is identically zero for

odd numbers of fields, and for even numbers of fields is

$$\langle 0 | T\phi(x_1) \dots \phi(x_{2n}) | 0 \rangle = \frac{1}{i^n} \sum_{\text{pairs}} \Delta(x_{i1} - x_{i2}) \dots \Delta(x_{i_{2n-1}} - x_{i_{2n}}). \quad (1.34)$$

In this equivalence, $\Delta(x - x')$ is the Feynman propagator between x and x' and is Green's function for the Klein-Gordan equation which contains the same $(-\partial_j^2 + m^2)$ term as multiplies in the integrals,

$$(-\partial_x^2 + m^2) \Delta(x - x') = \delta^4(x - x') . \quad (1.35)$$

In this way the two components of the LSZ formulation interact. The integrals contain factors of $(-\partial^2 + m^2)$ and the propagators contain factors of $1/(-\partial^2 + m^2)$. The on-shell requirement, that $|f\rangle$ and $|i\rangle$ have fixed masses causes $(-\partial^2 + m^2) \rightarrow 0$. This term in the integral cancels exactly with the corresponding diverging term in the appropriate propagator, but sets to zero the terms which do not have this divergence. This is the mechanism by which final state particles are observed to be on their mass shell, or on-shell, while allowing for interactions to happen between particles at varying masses, off-shell, with varying probabilities.

1.3.2 Path integrals and Feynman diagrams

To calculate the expression in 1.34, an integral over paths through the space of field configurations is used. This is called the path integral, $Z(J)$, and is a function of the

Lagrangian, \mathcal{L} , current sources, J , and fields, ϕ ,

$$Z(J) = \langle 0|0 \rangle_J = \int \mathcal{D}\phi e^{i \int d^4x (\mathcal{L} + J\phi)} . \quad (1.36)$$

The fundamental interactions between particles as described by $Z(J)$ can be visually represented using the Feynman diagram notation. On a 2-D spacetime plane, (x, t) , lines representing particles meet at a vertex, which is the interaction point. Fermions are represented using solid lines with an arrow indicating a direction. If the component of the arrow in the t dimension is positive, the particle fermion is indicated, and if it is negative, the antiparticle is meant. In this way, by rotating $(x, t) \rightarrow (x', t')$ a single diagram can be interpreted as depicting possibly more than one interaction.

To describe the scattering of fermions and gauge bosons in the SM, diagrams with the appropriate initial and final state particles are constructed from the vertices illustrated in Figure 1.1. Three-point vertices are made from the intersection of three line segments at a point, and four-point vertices are made from the intersection of four line segments. External lines correspond to real particles which are observed in the final state and are on-shell. Internal lines correspond to virtual particles that are not observed in the final state and therefore may or may not be on-shell.

A diagram is tree-level if it uses the smallest number of vertices possible to depict an interaction having the correct initial and final states. For the case of two fermions, $\phi(x_1), \phi(x_2)$, which scatter via a gauge boson that subsequently decays to two possibly different fermions, $\phi(x'_1), \phi(x'_2)$, the three unique tree-level diagrams for a given configuration of $\phi(x_1), \phi(x_2), \phi(x'_1), \phi(x'_2)$, are illustrated in Figure 1.2.

These specific diagrams are depictions of the s, t, u channels respectively, named after the Mandelstam variables, which are Lorentz scalar combinations of incoming and outgoing momenta,

$$\begin{aligned} s &= -(k_1 + k_2)^2 = -(k'_1 + k'_2)^2 , \\ t &= -(k_1 + k'_1)^2 = -(k'_2 - k'_2)^2 , \\ u &= -(k_1 + k'_2)^2 = -(k_2 - k'_1)^2 . \end{aligned} \quad (1.37)$$

The Mandelstam variable s is the square of the momenta of the incoming particles in the center-of-mass (CM) frame and \sqrt{s} is the typical variable used by experimentalists to discuss the energy of colliding beams of particles.

Combining contributions from all valid Feynman diagrams according to the statistical rules appropriate for the particles which participate in the interaction, gives $W(J)$ where

$$Z(J) = e^{iW(J)} \quad (1.38)$$

and the time ordered product of the fields can be found by taking the appropriate functional derivatives on W ,

$$\langle 0 | T\phi(x_1) \dots \phi(x_N) | 0 \rangle = \delta_1 \dots \delta_N iW(J)|_{J=0} . \quad (1.39)$$

1.3.3 Renormalization

Crucial for the LSZ formulation are two features of quantum fields, that they are fully separable in the infinite limit, $\langle 0 | \phi(x) | 0 \rangle = 0$, and that they are states of definite

Figure 1.1: The couplings for the SM gauge interactions.

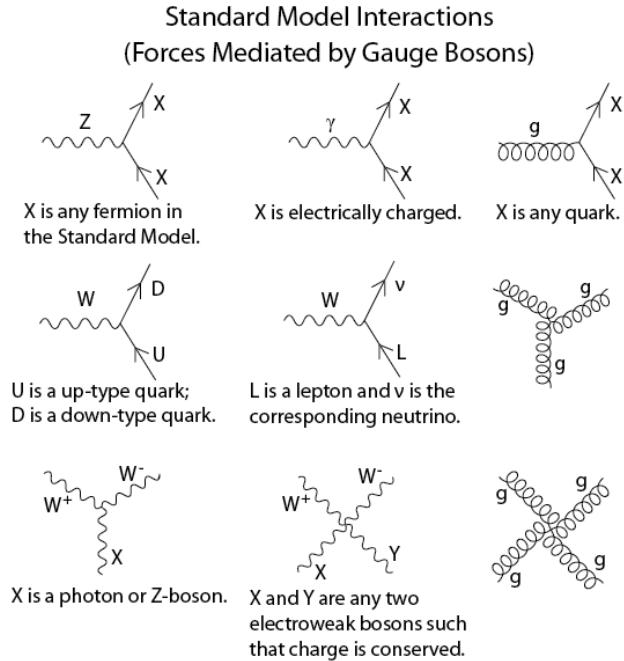
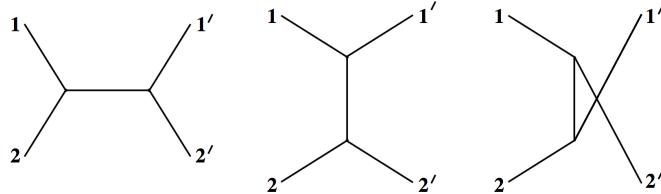


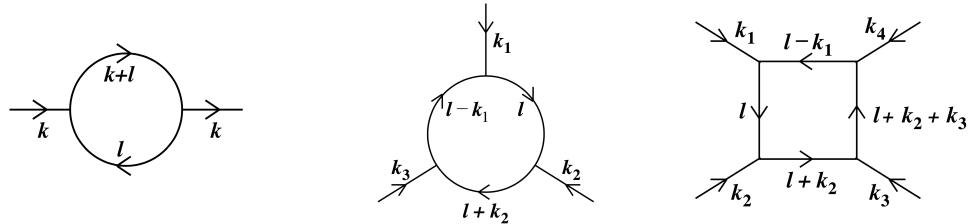
Figure 1.2: Below are Feynman diagrams for $2 \rightarrow 2$ scattering. These are the tree-level diagrams of the s, t, u channels respectively.



momentum, $\langle k|\phi(x)|0\rangle = e^{-ikx}$. In the SM, which allows for fields to interact, these conditions are guaranteed by adjusting the strengths of the fundamental couplings, g_1, g_2, g_3 . This ensures that the quantum states remain properly normalized, and is thus called renormalization.

Renormalization is necessary to account for corrections to the tree-level propagators

Figure 1.3: Renormalization takes place via a modification of the coupling parameters g_1, g_2, g_3 to account for the contributions that loops of virtual particles make on the propagators and vertices. Below are diagrams keeping track of the flow of momentum for one-loop corrections to the propagator and three and four point vertices.



and vertices which arise from contributions of virtual particles connecting to form closed internal loops. The lowest order, one-loop, diagrams are illustrated in Figure 1.3. Renormalization is accomplished by introducing an energy scale, and assuming that the couplings are small compared to this scale. The coupling constants are therefore functions of energy and are quoted at a particular renormalization scale, μ_R .

1.3.4 Cross sections and decay rates

The scattering amplitude $\langle f|i \rangle$ is not a directly measurable observable. What can be observed is some finite distribution of data which may be analyzed to reveal information about the scattering amplitude. Quantum mechanics dictates that only predictions of probability are possible, and the final probability of observing a particular interaction is dependent on many variables, including the energies, types and angular momenta of the incoming and outgoing particles as well as the masses of the propagators and the orientation and efficiency of the detector.

A quantity typically measured is therefore the interaction cross section, σ , and

for the scattering of two incoming particles going to n' particles, $2 \rightarrow n'$, in the CM frame, the differential is

$$d\sigma = \frac{1}{4|\mathbf{k}_1|_{\text{CM}}} |\mathcal{T}|^2 d\text{LIPS}_{n'}(k_1 + k_2) \quad (1.40)$$

where the scattering matrix element, \mathcal{T} , is defined using Equation 1.33, as

$$\langle f | i \rangle = (2\pi)^4 \delta^4 \left(\sum k_{\text{in}} - \sum k_{\text{out}} \right) i \mathcal{T} \quad (1.41)$$

and the Lorentz-invariant measure of the phase space for the n' outgoing particles is

$$d\text{LIPS}_{n'}(k) = (2\pi)^4 \delta^4 \left(k - \sum_{j=1}^{n'} k'_j \right) \prod_{j=1}^{n'} dq'_j \quad (1.42)$$

with the Lorentz-invariant differential dq .

The cross section is used to calculate the rate at which a process occurs, but is not the only relevant factor in determining the overall production rate. The production rate of a given final state is also dependent on the incoming rate of possible interactions and is known as luminosity, \mathcal{L} . Luminosity has the units of inverse area per unit time and the total number of events produced is therefore proportional to $\int \mathcal{L} dt$. In any real detector, final state particles are collected only within a finite solid angle and the number of particles scattered into a given solid angle, Ω , is given by

$$\frac{dN}{d\Omega} = \mathcal{L} \frac{d\sigma}{d\Omega} . \quad (1.43)$$

It is also possible for particles to decay as $1 \rightarrow n'$. Massive particles decay to

lighter ones in both the fermion and boson sectors, with all massive bosons able to spontaneously decay via the diagrams in Figure 1.1. Of the charged fermions, only the first generation is stable for each type, and neutrinos are not known to spontaneously decay, but oscillate between flavors while propagating in free space. Like the differential cross section, the differential decay rate is a function of the scattering amplitude and has integration measure $d\text{LIPS}$,

$$d\Gamma = \frac{1}{2E} |\mathcal{T}|^2 d\text{LIPS}_{n'}(k) . \quad (1.44)$$

The differential decay rate is inversely proportional to the energy of the particle, $E = \sqrt{m^2 + p^2}$. This means that comparatively heavy particles will decay faster than comparatively light ones and that energetic particles will appear to live longer for a stationary observer due to relativistic time dilation effects. The total decay rate of a given particle is found by summing the decay rates from each of the contributing processes, and the primary decay channels and rates for the fundamental particles are given in Table 1.4.

At CMS, the heaviest quark and the heaviest lepton both decay before reaching the detector volume. This makes b quarks the heaviest fundamental particles which can be seen to decay inside the detector, and therefore an object of interest. Additionally, their heavy mass means that they couple strongly with the Higgs boson which still has many properties that are under investigation. The W and Z bosons are both so massive that they decay before reaching the innermost layers of the detector and are often identified by their decay products pointing back to a common vertex.

Table 1.4: Below are listed the decay channels and rates for each of the unstable fundamental particles. At CMS, with the detection apparatus located a finite distance away from the interaction vertex, particles such as the W , Z and Higgs bosons, as well as the t and tau, decay before reaching the first layer of the detector.

Particle	Primary decay modes(s)	Total rest-frame $d\Gamma$	Typical decay location
W	$W \rightarrow \ell\nu$	X	Before reaching CMS
Z	$Z \rightarrow f\bar{f}$ (for $2M_f < M_Z$)	X	Before reaching CMS
τ	$\tau \rightarrow W\nu_\tau$	X	Before reaching CMS
μ	$\mu \rightarrow W\nu_\mu$	X	After leaving CMS
t	$t \rightarrow W^+b$	X	Before reaching CMS
b	$b \rightarrow W^-c$	X	Inside CMS
c	$c \rightarrow W^+s$	X	Inside CMS
s	$s \rightarrow W^-u$	X	Inside CMS

1.3.5 QCD and Proton Structure

The Feynman diagrams introduced in Section 1.3.2 describe the interactions between fundamental particles, but at the LHC, collisions take place between protons, which are composite.

One feature of the $SU(3)$ symmetry of the strong force is that gluons carry one unit of color and one unit of anticolor while the quarks carry one unit of color charge. This is what allows gluons to interact with each other as well as with quarks. That quark confinement is necessitated by the $SU(3)$ structure has not been conclusively determined, but observationally, a free gluon or quark has never been observed. Instead, quarks appear as bound in colorless (singlet) combinations called hadrons which are further classified as mesons ($q\bar{q}$) or as baryons (qqq or $\bar{q}q\bar{q}$), and are held together by gluons. Evidently, the binding energy of the quarks has a form such

that after a distance of roughly 10^{-15} meters, the energy stored in the gluon field is greater than the energy needed to create a quark-antiquark pair, bringing the pair into existence. This process of energetic quarks creating particles as they separate is called hadronization and is an important effect at the LHC.

Protons are a type of baryon and at low energy, may combine with a single electron to form a neutral hydrogen atom. At higher energies, the internal structure of the proton becomes more evident, and it contains three valence quarks, uud , which are constantly exchanging gluons. When probed at high enough energy, or equivalently, at short enough length scales, these gluons can also each split into a $q\bar{q}$ pair which typically reannihilate with each other. With gluons inside the proton splitting into quarks and coupling with other gluons, this forms a ‘sea’ of quarks and gluons, and as protons are accelerated to energies of GeV or TeV as is the case at the LHC, the fraction of the momentum of the proton attributed to the gluons becomes higher than that attributed to the valence quarks.

A proton-proton collider was therefore a sensible choice for the LHC. The physics goals of the project are to measure quantities associated with a wide range SM processes and to continue the search for evidence of new physics. Quarks interact with all of the SM gauge bosons as well as with the Higgs boson and the proton contains the lightest quarks of each type in addition to the gluons and sea. Colliding proton beams thus allow for the interactions between many different initial particle configurations to be explored, and with the exception of the neutrinos which interact only via the weak exchange of the Z boson and escape the detectors, all other fundamental SM particles have been directly observed at CERN.

1.4 Dark Matter

1.4.1 Experimental motivations

Albert Einstein's theory of general relativity, GR, has many experimental predictions which run counter to human intuition. GR predicts that massive objects warp a four-dimensional spacetime and thus feel mutual attraction. This has as a consequence, the prediction that even massless objects such as photons will experience a net deviation in their path near a massive object as a result of gravity and this effect was famously verified by Arthur Eddington through the observation of stars around the sun during a full solar eclipse. More recently, the direct detection of gravitational waves by the LIGO Collaboration also aligns with the GR predictions of distorted spacetime around colliding black holes. The time distortion effects due to the varying strengths of Earth's gravitational field on the surface and at the GPS satellites, provide precise tests of the quantitative predictions of GR. However, though these tests and others provide evidence that GR is an accurate theory of gravity, some basic predictions related to gravitational interactions do not agree with observations, motivating the concept for DM.

The first observational evidence for DM came from an analysis of the speeds of galaxies in the Coma cluster by Fritz Zwicky. The magnitude of the angular velocities of the galaxies was too great to be explained by the visible matter alone and DM is now believed to outweigh visible matter in a ratio of 5 : 1 throughout the universe and 10 : 1 throughout the Milky Way galaxy.

1.4.2 Simplified theoretical models

The defining features of DM are that it is massive and appears to interact on large scales only via the gravitational force. On the galactic and supergalactic scales, DM is distributed along similar structures as is visible matter, and it surrounds visible matter in extended halos.

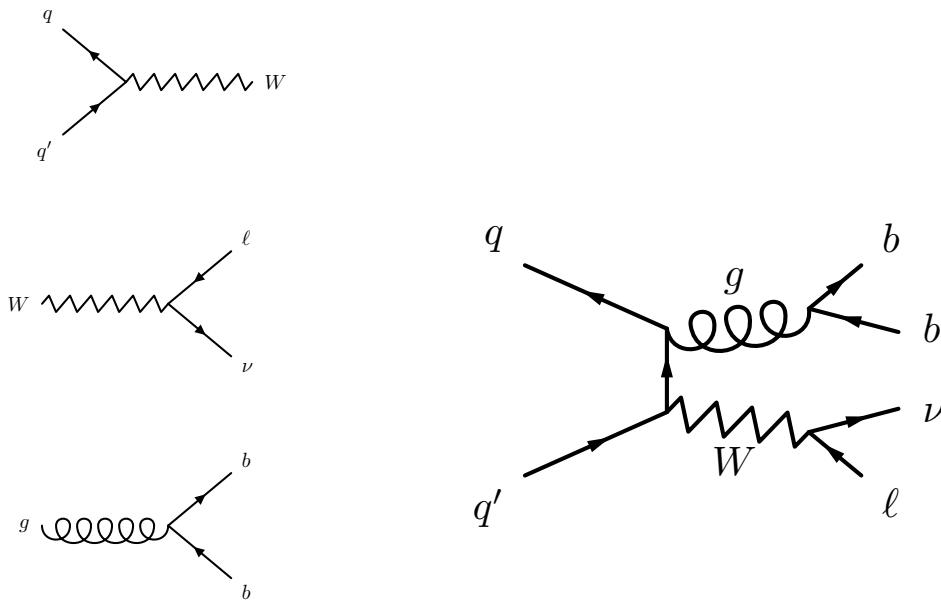
Because DM has not yet been observed, the models of DM being considered in this thesis are simplified and based on minimal assumptions, the first being that DM is even capable of interacting with hadrons and is thus possible to produce at the LHC. While the visible sector of particles is diverse, the models used in this analysis consist of a single DM particle, χ , which is assumed to be a fermion and may be different from $\bar{\chi}$.

One way DM could couple to the SM is via the addition of a $U(1)$ symmetry that gives rise to a vector gauge mediator, M . If some quarks are also charged under $U(1)$, then DM may be produced in the s channel as $f\bar{f} \rightarrow M \rightarrow \chi\bar{\chi}$. If M conserves parity in $f\bar{f} \rightarrow M$, it is said to have a vector coupling, and if it violates parity, it is termed axial-vector. In these models, M is assumed not to couple to leptons, but an effective field theory (EFT) model is also considered in this analysis which estimates a direct interaction between DM and photons. This coupling is mediated by a vertex $\gamma\gamma\chi\bar{\chi}$ and allows for DM production via the channel $pp \rightarrow \gamma \rightarrow \gamma\chi\bar{\chi}$.

2 PHENOMENOLOGY OF PROCESSES

2.1 The Standard Model process $pp \rightarrow Wb\bar{b} \rightarrow \ell\nu b\bar{b}$

Figure 2.1: The Feynman diagram for the process $pp \rightarrow Wb\bar{b} \rightarrow \ell\nu b\bar{b}$ is illustrated below, and is composed from the individual vertices illustrated on the left, each of which is described in Section 2.1.



2.1.1 $pp \rightarrow W$

The W boson couples to all charged fermions and can be created during the collision of a quark-antiquark pair with a relative charge difference of e . In the proton are quarks and the most prevalent valence quark is the u . Therefore in a pp collision, the channel by which most W bosons are produced is via a the annihilation of a valence

u quark from one proton with a \bar{d} from the sea of the other, $u\bar{d} \rightarrow W^+$. Quarks of higher generation can also be found inside the sea as the result of gluons splitting into $q\bar{q}$ pairs, but all interactions are modified by a coefficient in the CKM matrix and higher generation mixing is thus suppressed. In this thesis, all modes of $pp \rightarrow W^\pm$ production are considered.

2.1.2 $W \rightarrow \ell\nu$

Just as the W boson can be created by the collision $q\bar{q}' \rightarrow W$, it can also decay as $W \rightarrow q\bar{q}'$. This is known as hadronic W decay and can be a useful analysis channel for experimentalists, especially for decay products with energies approaching the TeV scale. Leptonic W decay, $W \rightarrow \ell\nu$, is also an important channel for experimentalists and is the one considered in this analysis. Because leptons constitute a negligible fraction of the sea, the detection of leptons at high energy after a pp collision is often a good indicator of the decay of a massive gauge boson, $W \rightarrow \ell\nu$ or $Z \rightarrow \ell\bar{\ell}$.

The W boson is much heavier than any of the leptons and therefore decays with roughly equal probability to any of $e\nu_e, \mu\nu_\mu, \tau\nu_\tau$. From Table 1.4, tauons created at CMS subsequently decay before reaching the detector, so for this analysis, the decay channel of the W investigated is $W \rightarrow \ell\nu$ where $\ell \in e, \mu$.

To reconstruct muons from the decay products, the transverse mass, m_T variable is often used. It is defined by $m_T^2 = m^2 + p_x^2 + p_y^2$ where p_i is the component of the momentum along the i axis, and in the case of a massive particle decaying to two

massless particles, can be rewritten as

$$m_T^2 = 2p_{T,1}p_{T,2}(1 - \cos\phi) \quad (2.1)$$

where ϕ is the angle between the particles and $p_{T,j}$ is the component of the momentum of the particles in the transverse plane.

In the decay of a W boson, a neutrino is produced, but can not be detected. The CMS detector is designed to capture the energy from all of the other particles produced in a collision, so neutrinos are accounted for as E_T^{miss} . The variable E_T^{miss} is the transverse component of the negative vector sum of all of the energy identified as having come from a particular interaction vertex, known as the primary vertex. Therefore the transverse mass of the W boson is

$$(m_T^W)^2 = 2p_T^\ell E_T^{\text{miss}}(1 - \cos\phi) \quad (2.2)$$

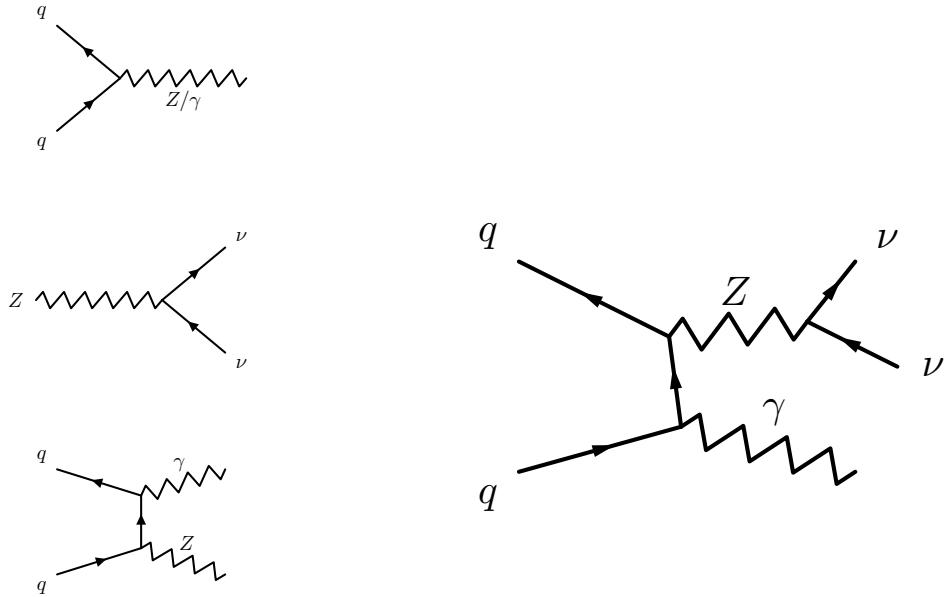
where ϕ is the angle between the lepton and E_T^{miss} .

2.1.3 $g \rightarrow b\bar{b}$

Because quarks couple strongly to gluons and $q\bar{q}' \rightarrow W$ has been shown to be an important production channel in pp collisions, it is possible for one of the initial state quarks to radiate a gluon. This is called initial state radiation, ISR, and if the gluon is produced with enough energy, it is capable of splitting to a quark-antiquark pair. In particular, a $g \rightarrow b\bar{b}$ vertex can be added to either of the incoming quarks to form $pp \rightarrow Wb\bar{b} \rightarrow \ell\nu b\bar{b}$.

2.2 The Standard Model process $pp \rightarrow Z\gamma \rightarrow \nu\bar{\nu}\gamma$

Figure 2.2: The Feynman diagram for the process $pp \rightarrow Z\gamma \rightarrow \nu\bar{\nu}\gamma$ is illustrated below, and is composed from the diagrams illustrated on the left, each of which is described in Section 2.2.



2.2.1 $pp \rightarrow Z/\gamma$

Similar to the W boson, the Z boson and the photon can also each be produced via the collision of quarks in the process $q\bar{q} \rightarrow Z/\gamma$. Unlike interactions with the W boson, interactions with Z/γ conserve parity invariance and do not transport charge. Any interaction which can happen as mediated by a photon can also happen with the exchange of a Z boson, but for collisions at $\sqrt{s} < M_Z = 90$ GeV, the Z can not be made on-shell. In this low energy regime γ exchange dominates, but in 2015, the

LHC ran at $\sqrt{s} = 13$ TeV and the relative mass difference between the Z and the γ played a negligible role in their relative rates of production.

2.2.2 $Z \rightarrow \nu\bar{\nu}$

The only particle which the Z boson can couple to but the photon can not is the neutrino. At $\sqrt{s} = 13$ TeV, the mass differences between the five lightest flavor of quark and the six leptons are negligible and the Z boson can decay into any kinematically allowed pairs, $Z \rightarrow f\bar{f}$. Including the three color possibilities for each quark, these are $3 \times 5 + 6 = 21$ final states, each of which has roughly the same branching fraction. Therefore only approximately $2/7$ of Z decays happen in the leptonic channel $Z \rightarrow \ell\bar{\ell}$. and of these decays, approximately $2/3$ happen as $Z \rightarrow \nu\bar{\nu}$.

2.2.3 $pp \rightarrow Z\gamma$

With the photon coupling only to electrically charged particles, the only place where a vertex containing a photon could be attached to either of the upper two left diagrams in Figure 2.2 is on one of the quarks. Photons are massless and therefore stable and so are a final state observable. Like the gluon in $pp \rightarrow Wb\bar{b} \rightarrow \ell\nu b\bar{b}$, the photon is an example of ISR. In the CM frame of the colliding $q\bar{q}$, which is approximately the lab rest frame for colliding beams of equal energy as is the case at the LHC, conservation of momentum dictates that the Z boson and photon should have equal and opposite momenta.

Unlike any of the other fermions, neutrinos are electrically neutral and therefore only interact via the weak force. So while the cross sections for most fermion-fermion

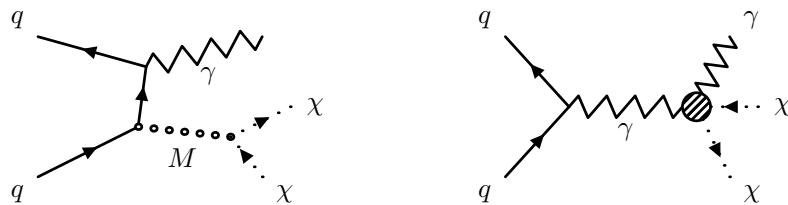
interactions involve contributions from the comparatively stronger electromagnetic and strong forces, the neutrino cross section contains contributions from only the W and Z bosons at tree-level and is much smaller than that of the charged fermions. This makes the detection of neutrinos very difficult in general, and impossible to do with present technology given the extreme backgrounds present in a collider setting.

In the case where the ISR photon is recoiling against a Z boson which decays to neutrinos, no direct detection of the Z boson or of its decay products is possible, leaving only the photon visible in the final state. This is called the monophoton signature, where a photon is observed recoiling against apparently nothing, and while the monophoton signature is predicted to be observed as a result of SM process as in $pp \rightarrow Z\gamma \rightarrow \nu\bar{\nu}\gamma$, if the observed monophoton cross section is measured to be higher than predicted, it could also be an indicator of physics beyond the SM (BSM). Specifically, the monophoton signature used in searches for dark matter.

2.3 $pp \rightarrow \gamma + \text{invisible}$

The existence of particle dark matter is well motivated, and the simplified model theories of DM used in this thesis allow for interactions which can result in the monophoton signature. One of the classes of models considered is a $U(1)$ gauge theory in which $\chi\bar{\chi}$ is produced via a vector or axial-vector mediator M which couples to quarks. The tree-level process in this model which leaves a monophoton signature is illustrated on the left of Figure 2.3, and the relevant parameters governing the cross section of this interaction are the masses of the two particles, m_χ and m_M ,

Figure 2.3: Feynman diagrams for the DM process $pp \rightarrow \gamma + \text{invisible}$ using simplified models are illustrated below. On the left is the $U(1)$ gauge model in which DM production is mediated by M which can be either vector or axial-vector. On the right is the diagram for DM production using an EFT model of the $\gamma\gamma\chi\bar{\chi}$ coupling.



and the strengths of the couplings between M and quarks, g_{Mq} , and between M and DM, $g_{m\chi}$. An EFT describing the four-point interaction vertex $\gamma\gamma\chi\bar{\chi}$ is also considered and illustrated on the right side of Figure 2.3. In this theory, the coupling is a function of two parameters, k_1 and k_2 , and is moderated by a mass scale, Λ . The other parameter in the EFT is m_χ , and by measuring the cross section for $pp \rightarrow \gamma + \text{invisible}$ in comparison with the SM prediction, estimations or limits can be set on the parameters used in either of these two models.

3 THE LHC AND CMS

The measurements presented in this thesis are performed on data of proton-proton collisions collected by the Compact Muon Solenoid (CMS) detector and provided by the Large Hadron Collider (LHC) at the European Center for Nuclear Research (CERN). The LHC was designed to probe physics at the scale of TeV and is capable of operating at multiple energy scales. As measured in the CM frame of protons colliding inside CMS, the LHC operated at $\sqrt{s} = 8$ TeV in 2012 and $\sqrt{s} = 13$ TeV in 2015. The measurement of the $pp \rightarrow Wb\bar{b}$ cross section is performed using 19.8 fb^{-1} of integrated luminosity collected at $\sqrt{s} = 8$ TeV and the measurement of the $pp \rightarrow Z\gamma \rightarrow \nu\bar{\nu}\gamma$ cross section and the extensions to set limits on DM models uses 2.3 fb^{-1} of data collected at $\sqrt{s} = 13$ TeV.

3.1 The Large Hadron Collider

The LHC is a single-ring, double-bore particle accelerator and collider located on the border of France and Switzerland outside Geneva. It was built using the existing 26.7 km of tunnels from the Large Electron Positron collider and hosts four primary experiments, located at four interaction points where beams of hadrons are made to cross. Of the four experiments, two (CMS and ATLAS) are built for studying SM processes and searching for new physics in general, one (ALICE) is designed to investigate quark-gluon plasma resulting from the high energy collisions of heavy ions such as lead, and one (LHCb) was built for the study of b-mesons and CP violation.

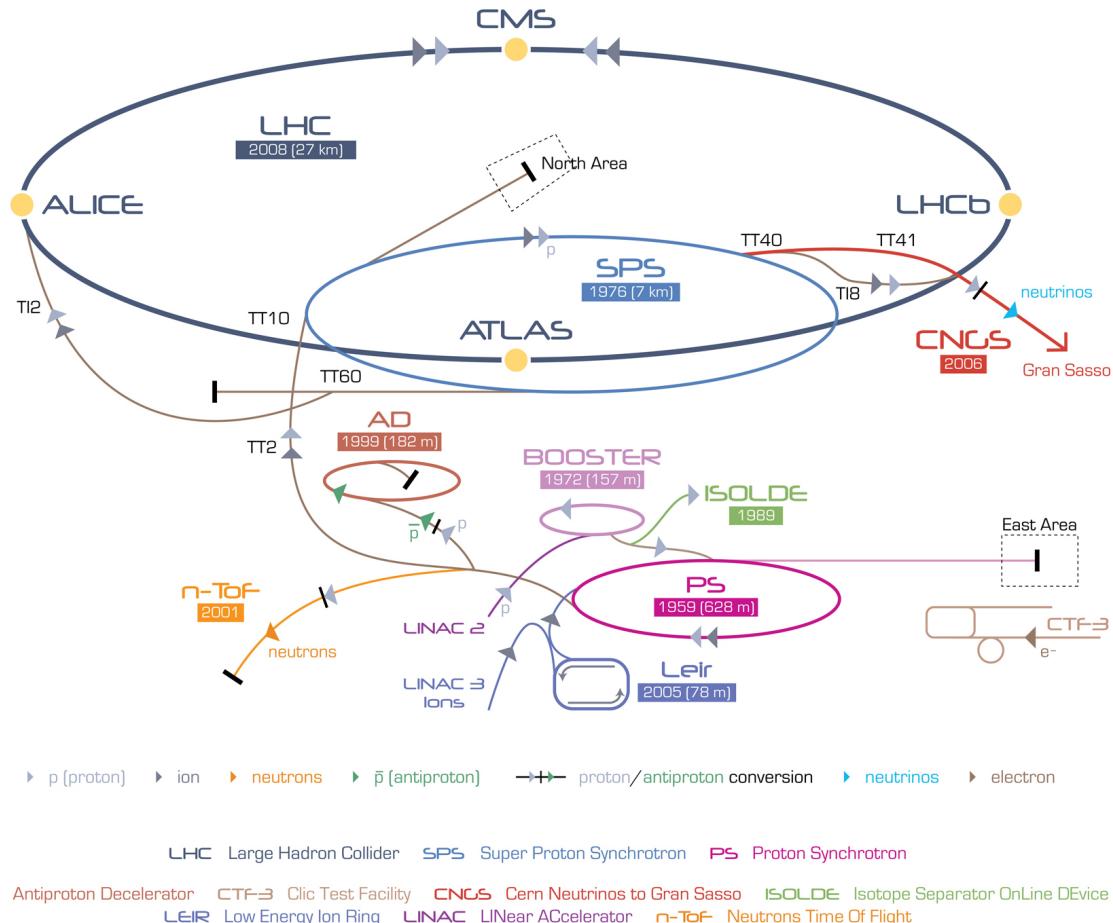
3.1.1 LHC pre-acceleration

To accelerate protons to their collision energy, a multi-stage procedure is used and the major components of the accelerator infrastructure are illustrated in Figure 3.1. First, protons are separated from electrons in neutral hydrogen gas before entering the linear accelerator (LINAC2) which brings them up to an energy of 50 MeV using a series of oscillating electric potentials. In this process, rather than having a continuous stream of accelerating protons, the protons are grouped into bunches, and the beam retains this structure of distinct groups of protons separated by gaps throughout the acceleration procedure. After the LINAC2, protons enter the Proton Synchrotron Booster (BOOSTER) where they are accelerated to 1.4 GeV and prepared for injection into the Proton Synchrotron (PS). Inside the PS, bunches are accelerated to 26 GeV before being injected into the Super Proton Synchrotron (SPS) where they are further accelerated to 450 GeV. After the SPS, bunches of protons are sent into the LHC.

3.1.2 LHC acceleration

The work of accelerating and containing the protons which form the beam of the LHC is done by superconducting magnets. They are cooled to a temperature of 1.9 K using liquid helium and are housed in the LHC dipole apparatus diagrammed in Figure 3.2. The dipole contains two beam pipes which are each surrounded by superconducting coils of Niobium Titanium (NbTi) which carry oscillating currents when in operation. These constitute RF cavities operating at 400 MHz and having the ability to circulate proton bunches in opposing directions between the two beam pipes with a spacing of 25 ns between bunches. The magnets are capable of reaching a strength of over 8 T,

Figure 3.1: Before protons are released into the LHC for final acceleration and collision, they pass through the LINAC2, BOOSTER, PS, and SPS, undergoing acceleration at each stage.



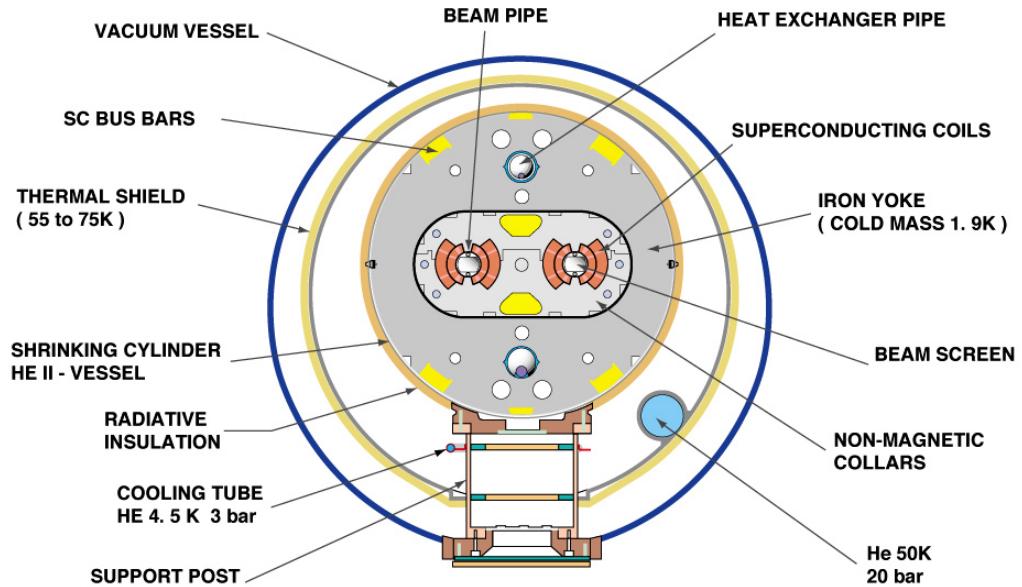
a constraint imposed by the desired energy scale of the accelerator and the radius of the existing LEP tunnels in which the LHC was built.

The Lorentz-invariant magnetic force on a particle of charge q , moving at velocity \mathbf{v} in a magnetic field \mathbf{B} is

$$\mathbf{F}_B = q\mathbf{v} \times \mathbf{B} \quad (3.1)$$

Figure 3.2: Below is a cross section of the LHC dipole apparatus. It contains two beam pipes, each surrounded by superconducting magnetic coils which are held in place by an iron yoke. The system is cooled to a temperature of 1.9 K and is thermally isolated as well as protected from radiation.

CROSS SECTION OF LHC DIPOLE



CERN AC _HE107A_ V02/02/98

and the relativistic version of Newton's second law states

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = \frac{d}{dt} (\gamma m_0 \mathbf{v}) \quad (3.2)$$

where γ is the relativistic correction factor and m_0 is the rest mass of the particle. These can be solved to find that the magnetic field required to hold a proton in planar

circular motion as in the LHC is

$$\mathbf{B} = \frac{p}{qR} \quad (3.3)$$

where $R = 4.3$ km. At individual beam energies of 4 and 6.5 TeV in 2012 and 2015 respectively, the rest mass of the proton constitutes a negligible fraction of the proton momentum and the minimum magnetic field required by the magnets is [todo - unit conversion].

The rate at which a particular collision process occurs at the LHC is proportional to the cross section of that interaction and the luminosity of the colliding beams as given in Equation 1.43. Assuming a Gaussian beam distribution, the machine parameters determine \mathcal{L} as

$$\mathcal{L} = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} \mathcal{F}(\theta) \quad (3.4)$$

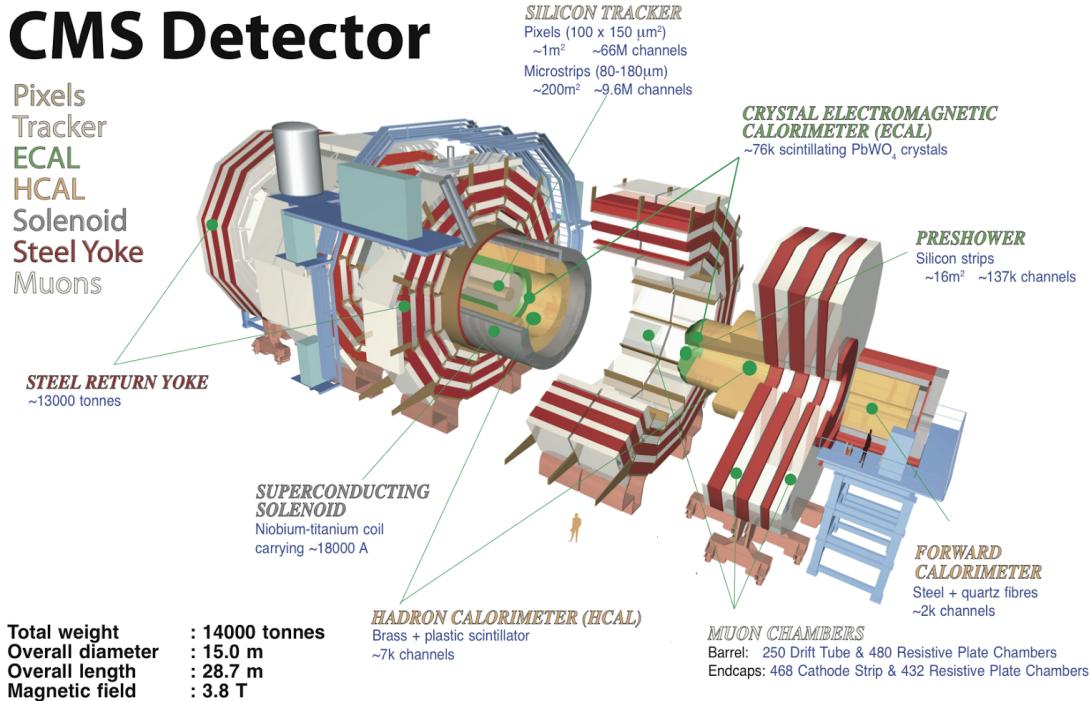
where N_b is the number of particles per bunch, n_b is the number of bunches per beam, f_{rev} is the revolution frequency of the bunches, $\gamma_r = E_p/m_p$ is the relativistic gamma factor for protons at energy E_p , ϵ_n is the normalized emittance which characterizes bunch width, β^* is a measure of the betatron oscillation envelope, and $\mathcal{F}(\theta)$ is a relativistic geometrical correction factor which is a function of the angle at which the beams cross. In addition to pushing the energy frontier, the LHC also has a significantly greater \mathcal{L} than previous hadron colliders. [Reference to Tevatron]

3.2 The Compact Muon Solenoid Detector

The CMS detector was built at Interaction Point 5 on the LHC ring to collect particle collision data exploiting the full physics reach of the LHC. The analysis of these data includes the discovery of the Higgs boson[REF] and high precision measurements of SM processes, as well as searches for physics beyond the standard model. To be able to perform such precision measurements, CMS was designed with four main subdetectors that work in concert and with a superconducting solenoid. The tracking and most of the calorimetric detectors are inside the solenoid while the muon detectors are outside. When running, the solenoid produces a 3.8 T uniform magnetic field in its interior, and has a uniform 2 T field over the bulk of the detector external to the solenoid.

The innermost of the subdetectors is the tracker which uses silicon pixel and strip detectors to record the tracks of charged particles passing through it. The tracks are used in conjunction with the 3.8 T magnetic field to measure the momentum of these particles and this information is used for identifying the pp interaction vertex as well as locating secondary vertices from the decay of heavy flavor quarks such as the b or c. Outside the tracker is the electromagnetic calorimeter (ECAL), which is designed to have good energy resolution in recording the electromagnetic interactions of charged particles such as electrons or photons over a wide range of angles. The hadronic calorimeter (HCAL) is outside the ECAL and is designed to absorb energy which comes in the form of neutral hadrons and provide good resolution in missing transverse energy, E_T^{miss} . Outside the calorimeters is the solenoid and steel return yolk, and the outermost layers of the detector are dedicated to the efficient detection

Figure 3.3: The CMS detector consists primarily of a tracker and electromagnetic and hadronic calorimeters which are mostly located inside a 3.8 T field provided by a superconducting solenoid, as well as a muon detection system located outside the solenoid.



of muons. The overall length of CMS is 21.6 m, with a radius of 7.3 m and a total weight of 12500 tons.

3.2.1 Geometry

The coordinate system used by CMS is one in which the z-axis is aligned with the beam pipe, the y-axis is pointing upward vertically and the x-axis points radially

inward toward the center of the LHC ring. The detector itself is mostly cylindrically symmetric about the beam pipe so cylindrical coordinates are also used. In this system, r is the radial distance as measured from the beam pipe, the azimuthal angle, ϕ , is measured up from the x-axis in the x-y plane, and the polar angle, θ , is measured down from the z-axis. The angle θ is commonly replaced by pseudorapidity,

$$\eta = -\ln(\tan \theta/2) \quad (3.5)$$

since the distribution of particles is roughly constant as a function of η . For the calorimeters, "barrel" refers to the region of $|\eta| < 1.4442$, and "endcap" to the region $3.0 > |\eta| > 1.566$. Instrumentation cables are run through the gap between the barrel and endcap, so this area has detecting components. The HCAL forward region covers $3.0 < |\eta| < 5$ and the tracker extends to $|\eta| < 2.5$.

3.2.2 Magnet

To precisely measure the momentum of a charged particle, it is necessary to measure radius of curvature of that particle as it moves through a magnetic field. The momentum resolution varies as

$$\frac{\delta p}{p} \sim \frac{1}{L^2 B} \quad (3.6)$$

where L is the length of the track of the particle through a magnetic field of strength B . For particles at high energy, this requires a very strong magnetic field which is achieved by the superconducting solenoid in CMS. The solenoid operates at 3.8 T

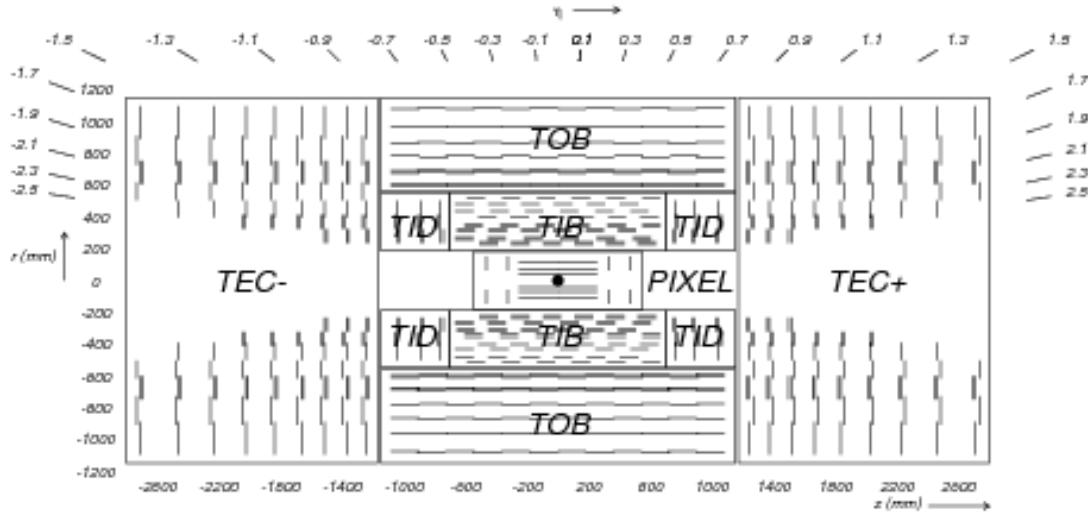
with a bore of 3 m in radius and 12.5 m in length and is constructed from four layers of NbTi superconductor. The steel yoke which provides physical support for the CMS structure and serves as an absorber for the muon system is fully saturated by the fringe magnetic field from the solenoid.

3.2.3 Tracking System

The inner tracking system of CMS is designed to provide precise and efficient measurements of the trajectories of charged particles produced during collisions, as well as a precise reconstruction of secondary vertices. The tracker has a length of 5.8 m and a radius of 1.25 m in a cylindrical structure surrounding the interaction point, as illustrated in Figure 3.4.

At the core of the tracker and closest to the beam line are three concentric cylindrical layers of hybrid pixel detector modules which are complemented by two discs of pixel modules on each end and extend a to a distance of 10 cm from the beam line. [TODO - Explain - silicon] In total, the pixel component of the tracker covers an area of about 1 m² with 66 million pixels. External to the pixel detector are the tracker inner barrel and discs (TIB/TID) which are made from silicon strips and extend out to a distance of 55 cm. There are four layers of strips in the TIB, with 3 discs at each end. The tracker outer barrel (TOB) is composed of 6 layers of micro-strip sensors and extends in z between ± 118 cm and to a radius of 116 cm. At the end of the z range for the TOB are the tracker end caps (TEC) which cover the ranges $124 < |z| < 282$ cm and $22.5 \text{ cm} < r < 113.5$ cm. Each TEC is composed of 9 discs, each carrying up to 7 rings of silicon micro-strip detectors. In total, the

Figure 3.4: Below is a schematic of the CMS tracking system where each line represents a detector module. The system is made from silicon pixels and silicon microstrips distributed into four sections, TIB, TID, TOB, TEC.



tracker contains 9.3 million strips which cover an area of 198 m^2 and extends to an acceptance of $|\eta| < 2.5$. For tracks with momentum on the order of 100 GeV, the momentum resolution is around 1-2% up to $|\eta| < 1.6$ and degrades to around 10% with increasing η .

3.2.4 Electronic Calorimeter

The electronic calorimeter (ECAL) is a homogeneous calorimeter made from nearly 76000 crystals of lead tungstate (PbWO_4) mounted in the barrel and endcap sections with a preshower detector located in front of the endcaps, arranged as shown in Figure 3.5 In the barrel, avalanche photodiodes are used as photodetectors, and in

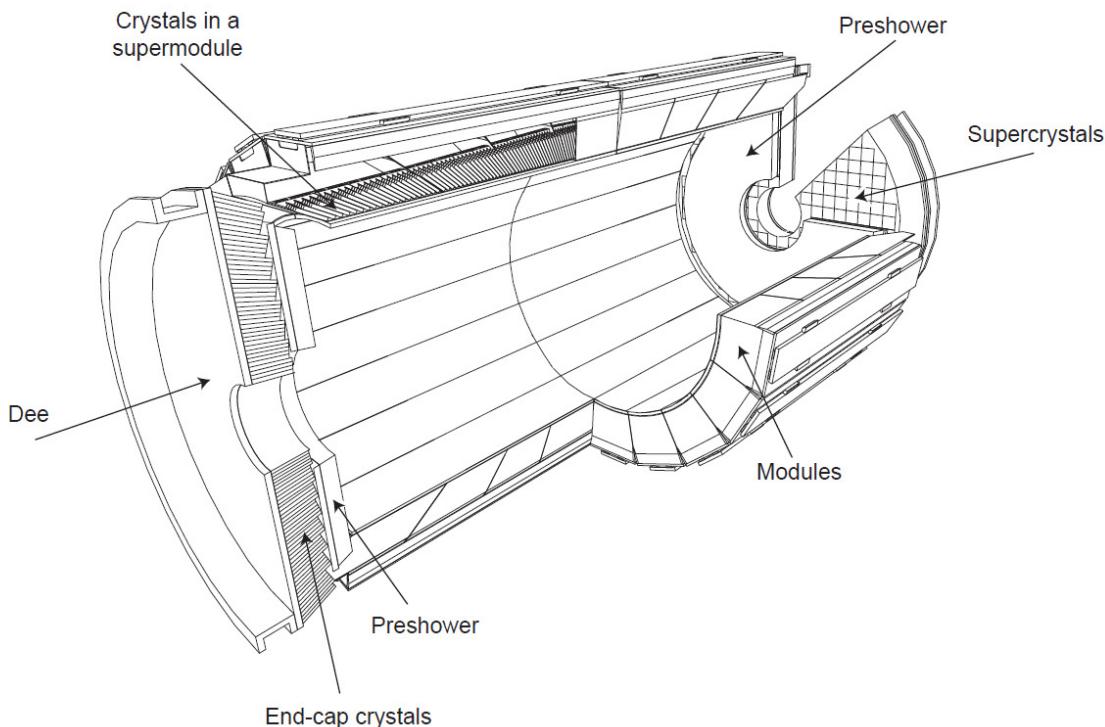
the endcap vacuum phototriodes are used. The material PbWO₄ was chosen for its properties of being dense, optically transparent and radiation hard. The radiation length inside the ECAL is typically less than 1 cm with a Moliere radius of 2.2 cm and about 80% of the light is emitted from a crystal within the first 25 ns. Since the length of a given crystal is on the order of 20 cm, most photons and electrons deposit all of their energy within the ECAL, and do not reach the HCAL.

The use of PbWO₄ crystals allows for excellent position and timing resolution with the energy resolution given by

$$\left(\frac{\delta_E}{E}\right) = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{0.12}{E}\right) + (0.30\%)^2. \quad (3.7)$$

In this expression, the first term comes from the statistical error in the measurement which arises from the stochastic nature of electromagnetic shower evolution and the second term represents the error in the measurement which results from noise in the electronics or energy deposits from additional soft interactions. The ECAL provides stable and accurate measurements of energies over a range from 1 GeV to 1 TeV, with the upper limit set by the energy at which electromagnetic showers penetrate through the ECAL into the HCAL. [REF][TODO - how much?] With time, after undergoing a heavy bombardment of high energy radiation, the PbWO₄ crystals physically deteriorate and develop nonuniform light transmission properties. [REF] This is monitored and corrected for using a laser calibration system that probes for changes in crystal transparency.

Figure 3.5: Below is a diagram of the ECAL, which sits between the tracker and HCAL in CMS. It is made from PbWO₄ crystals throughout the volume with avalanche photodiodes in the barrel and vacuum phototriodes in the endcaps.



3.2.5 Hadronic Calorimeter

Situated mostly between the ECAL and the superconducting solenoid is the hadronic calorimeter (HCAL) which plays a crucial role in the measurement of hadron jets and particles such as neutrinos which escape the detector and result in apparent missing transverse energy. The HCAL is designed to contain the energy of neutral particles which pass through the ECAL and is therefore made from dense materials such as

steel and brass interleaved with scintillating material. Because the HCAL is designed to fit between these two components, it takes the shape of a hollow cylinder of inner radius 1.77 m and outer radius 2.95 m and one half of the HCAL is illustrated in Figure 3.6.

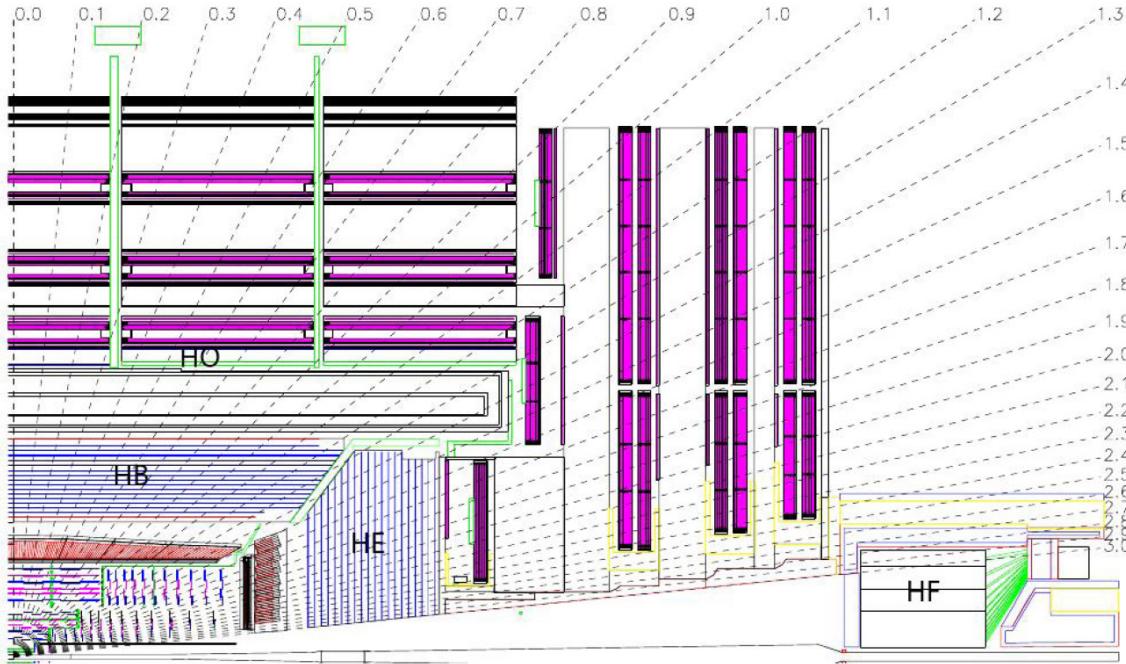
The barrel of the HCAL (HB) extends to $|\eta| < 1.3$ and is constructed from brass absorber plate wedges aligned parallel to the beam axis and mounted in an overlapping configuration, with a smaller amount of steel used in the inner and outermost wedges for structural stability. The endcap of the HCAL (HE) extends this coverage to $|\eta| < 3.0$ and is complemented by the forward hadron calorimeter (HF) which is made from the comparatively radiation-hard steel plates embedded with quartz fibers. Inside the barrel region there is an additional layer of the HCAL, the outer calorimeter (HO), which is located just outside the solenoid and uses it as an absorber for energetic showers which start late in the HB.

In the HB, HO and HE, light from particle showers [TODO - not exactly ..] inside scintillators and collected by quartz fibers and then used as an estimate of the total energy of the shower. In the HF, this estimate is made using the Cherenkov radiation from particles with energy above 190 keV collected by the quartz fibers. For the two cases, the energy resolution takes the same functional form

$$\left(\frac{\delta_E}{E} \right) = \left(\frac{A}{\sqrt{E}} \right)^2 + (B)^2 \quad (3.8)$$

where A is 90% (172%) in the HB/HO/HE (HF) and relates to the stochastic uncertainty of shower evolution and B is 4.5% (9.0%) and comes from uncertainties in calibration.

Figure 3.6: A schematic layout of the HCAL, which complements the ECAL in providing a measurement of the total energy produced in a collision. The HCAL is made from brass and steel plates, embedded with quartz fibers.



3.2.6 Muon System

Muons play a central role in the physics program outlined by CMS and the muon detection system is positioned as the outermost layer of the detector. Unlike the other charged leptons, muons typically pass through the ECAL and HCAL and deposit only a fraction of their energy, so a dedicated muon system is necessary in order to determine the momentum of these particles. The muon system is composed of three different kinds of gaseous detectors, drift tubes (DTs), resistive plate chambers

(RPCs) and cathode strip chambers (CSCs) and their layout is illustrated in Figure 3.7.

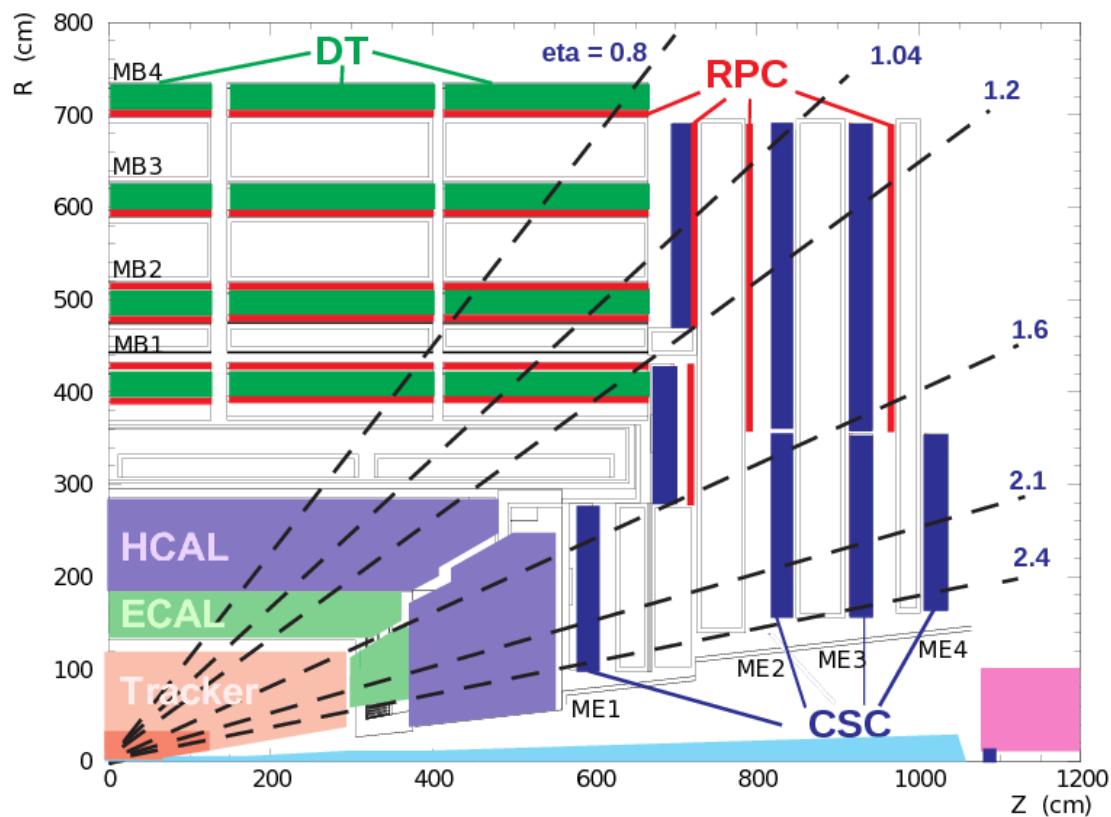
The barrel region of the muon system is covered by DTs in the range $|\eta| < 1.2$ and the endcaps are covered by CSCs in the range $0.9 < |\eta| < 2.4$. The RPCs are located in the range $|\eta| < 1.6$ and provide fast, independent and highly segmented transverse momentum measurements of muons.

The DT system is composed of 4 stations which form concentric cylinders about the beam line and contain 172000 sensitive wires. As charged particles enter the DTs, they ionize the Ar/CO₂ gas mixture, knocking off electrons which then are attracted to the positively charged wires.

The CSCs are less sensitive to uneven magnetic fields and high particle rates so are therefore used in the endcaps. They are made from crossed arrays of positively charged wires and negatively charged strips in gas and are composed of six layers, giving them precise timing as well as positional information. As an upgrade between the 2012 and 2015 data taking periods, a fourth layer of CSCs was added to the CMS detector, adding to the three which were present in 2012.

The RPCs are built from two sheets held at opposite charges and separated by a gas volume. As muons move through the chamber, electrons are ionized from the gas and attracted to small metallic strips which they reach after a small but well known time delay. The timing resolution of RPCs is on the order of 1 ns.

Figure 3.7: The CMS muon system uses DTs, RPCs, and CSCs to provide muon detection up to $\eta < 2.4$. Shown below is the geometrical arrangement of the different muon subsystems and how they fit with the rest of the CMS detector.



4 EVENT SIMULATION AND RECONSTRUCTION

4.1 Simulation of Events

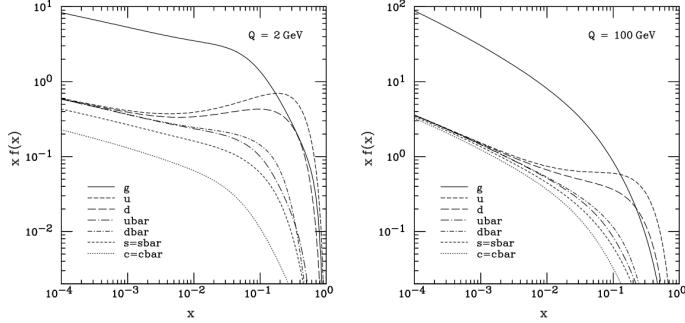
Vital to the analysis of the data gathered using the CMS detector are accompanying predictions to be compared against. Good predictions can be used not only for direct comparison against data as in the case of a cross section measurement, but can also be used in the design of future detectors and experiments, or for optimizations of parameters in blinded analyses. Predictions are made by simulating pp collisions and the subsequent decays and interactions that take place inside the CMS detector volume using Monte Carlo (MC) techniques. The first step in producing these simulations is the generation of the collision event itself, and the second is in simulating the interaction of the collision products with the detector.

4.1.1 Monte Carlo Event Generation

There are two complimentary methods used in producing a simulation of the collision event, the direct calculation of a scattering amplitude (also known as a matrix element, ME), and the showering of particles as they decay, hadronize and radiate.

As discussed in Section 1.3.5, protons are composite objects which form as the result of strong interactions between bound quarks. Protons are therefore modeled using parton distribution functions (PDFs) which describe the probability of finding a given constituent particle, or parton, to contain a given fraction of the momentum of

Figure 4.1: Below are proton PDFs shown at two different values of momentum transfer, Q . The horizontal axis shows the momentum fraction carried by the parton and the vertical axis shows the parton density.



the proton. The PDF is thus actually a set of density functions, one for each parton taken into account, and examples of the CTEQ6M parton distribution function at two values of of momentum transfer, Q , are shown in Figure 4.1. Weighted with probabilities from the PDF set and summed over, MEs are calculated explicitly from Feynman diagrams which have initial state particles found in the PDF and the desired final state.

Radiation is also important to correctly model. At any point in the collision any colored particle can radiate a gluon, and any charged particle can radiate a photon. Gluon radiation from initial state partons is always present in pp collisions at the LHC and results in jets, columnated showers of particles which are the products of quark hadronization and gluon splitting. The effects of radiation and parton showering are simulated using a Markov process in which vertices iteritively are added to partons with probabilities based on the coupling strengths, energies of the participants and

the generation of random numbers. For the constituents of the proton which did not participate in the hard interaction, quarks must be created from the vacuum to enforce confinement. This produces low energy, soft, radiation and is known as the underlying event. The underlying event must be simulated along with the hadronization effects for all colored particles.

4.1.2 Monte Carlo Generators

The two primary generators used in this thesis are MadGraph/MadEvent and Pythia. MadGraph is strictly a ME generator which interfaces with MadEvent for event generation and Pythia is mostly used for hadronization and showering.

For a given $2 \rightarrow n'$ scattering process, the differential cross section is a function of the Lorentz-invariant phase space as in Equation 1.40. To calculate the cross section within a finite phase space, $d\sigma$ is integrated and MadGraph numerically does this through the sampling of random numbers. The phase space can be interpreted as a multidimensional hypercube spanning all degrees of freedom for all final state particles and Equation 1.40 is used to calculate a weight, dw , for each point sampled in the phase space. The average of the weights converges towards $\int dw$.

To produce events with the frequency predicted by the theory being modeled, MadEvent uses the Von Neumann method to unweight events. For each event, a random number, g , is generated between 0 and 1 and compared to the ratio dw/dw_{\max} where dw_{\max} is the largest event weight sampled. If $dw/dw_{\max} > g$ then the event is kept and is otherwise rejected. Accepted events generated with MadGraph/MadEvent in this way have the same frequency and follow the same kinematic distributions as

predicted from the input Equation 1.40.

Pythia performs hadronization using the Lund string model in which quarks are confined to the ends of strings and gluons are represented as kinks on that string. As quarks separate, the string breaks and creates a $q\bar{q}$ pair, thus building confinement directly into the model. The underlying event is modeled in Pythia as a set of $2 \rightarrow 2$ processes which are correlated with each other via the color connections present in the proton, and the set of parameters used by Pythia to perform calculations is referred to as the tune.

4.1.3 Detector Simulation

After events have been produced, they are passed to Geant4 for simulation of the passage of particles through the physical mass of CMS. The Geant4 toolkit includes a full model of CMS, including all of the subdetectors as well as the inert material from the support structure and readout electronics. The magnetic field is emulated using data from measurements on the real field and Geant4 uses all of this information to register hits in the simulated detector as a consequence of the interaction between particles produced in the simulated event and the simulated material of CMS. Additionally, hits are added to the simulated detector taking into account the rates of background noise, and the final output is emulated data which is stored in the same way as would be data as taken from the real detector.

4.2 Reconstruction of Events

Real data collected from the detector and simulated data output from Geant4 consist of time-correlated energy deposits in the various subdetectors of CMS. As a result of the coordinated designs of the subdetectors, the final-state particles which arise from pp collisions at the LHC can be individually identified and reconstructed using the combined information from the entirety of CMS. The associated global event description from this particle-flow (PF) reconstruction provides excellent performance for the identification of electrons and muons, as well as for vertex identification and the evaluation of E_T^{miss} .

4.2.1 Track and Primary Vertex Reconstruction

The subdetector closest to the interaction vertex is the tracker, which records precise information about the trajectories of charged particles as they pass through it. Combined with the magnetic field, this allows for the measurement of the momenta of these particles as well as a means of identifying the location of the primary interaction.

Tracks are identified via an iterative process. The first tracks to be reconstructed are those which pass strict seeding criteria, designed to have a moderate efficiency, but negligibly small fake rate. Then the detector hits associated with these tracks are masked and the remaining hits are used to form track seeds with slightly relaxed criteria. This operation is repeated, with every iteration imposing more complex and time-consuming seeding, filtering and track fitting algorithms.

Because bunches of protons instead of single protons are made to cross in the LHC, multiple collisions can take place during the same bunch crossing. The vertex with the highest scalar sum transverse momentum, p_T , of tracks and passing further quality selections based on the goodness of fit for the tracks and the number of tracks associated with a given vertex is chosen as the primary vertex (PV).

In $pp \rightarrow W b\bar{b} \rightarrow \ell\nu b\bar{b}$ events, the two b quarks and the lepton from the W decay all leave energy deposits in the tracker, thus making the choice of PV unambiguous. However, in the $pp \rightarrow \gamma + \text{invisible}$ events, the only visible final state object is a photon, and photons do not leave hits in the tracker. This makes the identification of the PV in the monophoton analysis difficult and motivates the using of variables that are less sensitive to correct PV identification.

4.2.2 Electron ID and Reconstruction

Electrons are reconstructed using tracker hits and ECAL deposits. The seed of an electron candidate is selected as an energy deposit in the ECAL with $E_T > 4$ GeV having nearby deposits in the tracker. As electrons move in magnetic fields, they emit bremsstrahlung radiation tangential to their flight path and this radiation both appears in the detector, and alters the course of the electron.

The effects of this radiation are taken into account via the Gaussian Sum Filter (GSF) track fitting algorithm. This algorithm uses weighted sums of Gaussian functions to describe electron energy loss and thus allows for non-Gaussian corrections to the fitting of tracks. In the CMS detector, bremsstrahlung from electrons results in the emission of photons in an extended strip in the ϕ direction and electron

superclusters (SCs) are made by including the energy deposits from these photons in the ECAL as part of the candidate electron object.

Further requirements during the reconstruction of the electron improve the purity of selection. The SC and the GSF track are required to be separated by no more than $|\eta| < 0.02$ and $|\phi| < 0.15$ and the fraction of energy deposited in the HCAL directly behind the SC, and the SC is required to be no more than 15%.

4.2.3 Photon ID and Reconstruction

Photons are reconstructed using the same ECAL clustering algorithms as are used for electrons. This allows for the simultaneous reconstruction of photons that have and have not split to $e\bar{e}$ pairs. The size of the SC is determined dynamically and the center is determined to be the barycenter of the distribution, with weights assigned using the logarithm of the fractional energy deposits of the ECAL crystals clustered in the SC.

In an ideal tracker, photons would not interact at all and objects that leave signatures similar to those of photons could be rejected through the rejection of tracks. However, some photons do convert to $e\bar{e}$ pairs inside the tracker volume which leave tracks, so the rejection of tracks is not a perfect way to distinguish between photons and electrons.

4.2.4 Muon ID and Reconstruction

Muon identification is performed using two reconstruction and filtering methods to produce ‘tracker muons’ and ‘standalone muons’ which are combined to form ‘global

muons.’ Tracker muons are identified starting with a track, $p_T > 0.5$ GeV and $p > 2.5$ GeV, which is then extrapolated to the muon system. If the distance between the extrapolated track and the nearest hit in one of the muon chambers is less than 3 cm, a tracker muon is identified. Tracker muons are also identified if the pull between the extrapolated track and the matched station hit is less than four, where pull is defined as the distance between the track and the station hit divided by the uncertainties on both measured quantities. Tracker muons are built from the inside of the detector towards the outside, and standalone muons are built in the other direction. Only hits in the muon stations are used to reconstruct standalone muons, with the additional constraint that the path reconstructed from the hits points back toward the interaction region. Thus, the tracker muon algorithm is well-suited for the identification of low- p_T muons by having low thresholds and requiring only one track and one station hit, while the standalone muon algorithm is aimed at high- p_T muons which have the energy to penetrate multiple layers of muon stations to form tracks which can be traced back to the interaction. Global muons are required to pass the criteria for both standalone muons and tracker muons, and, starting with the standalone muons, the global muon trajectory is refit using information from both the muon stations and the tracker, yielding an improved energy resolution than either one.

4.2.5 Jet ID and Secondary Vertices

The reconstruction of jets is accomplished using the anti- k_t clustering algorithm on particles identified in the PF. The anti- k_t algorithm is both infrared and collinear safe,

meaning that it is stable against soft (low energy) radiation getting clustered into individual jets, and also stable against hard (high energy) jets splitting collinearly and affecting the shape of the jet.

Jets are corrected in simulation and in data to remove energy believed to come from elsewhere than the PV, thus removing the luminosity dependence of the jet. Jets are also corrected to have a response that is independent of η by studying dijet events and calibrating the jets to anti-align. To make the jet response independent of the p_T of the jet, an absolute correction is applied, and in data, one further correction on the relative energy scale is applied. After all of these corrections are applied, simulated jets are observed to have sharper energy resolution than is observed, so jets in MC smeared in energy

Bottom quarks have a relatively long lifetime and are the heaviest fundamental particle that has been seen to decay inside the volume of the CMS detector. A b quark produced in a pp collision at CMS therefore has enough time to hadronize into a jet before decaying, and such jets are called b -jets. The identification, or tagging, of b -jets is focused around the vertex associated with the b -hadron which, since it is not the PV but is still a vertex associated with the event, is called a secondary vertex, SV. The tagging of b -jets is accomplished using a multivariate analysis technique in which information from variables such as the number and energy of tracks that appear to be displaced from the PV, the presence of SVs or of soft leptons is all combined into a single discriminator value.

5 W+BB CROSS SECTION

MEASUREMENT [TODO - WESLEY COMMENTS / CUTFLOW]

We now have the tools in place to examine the first of the two Standard Model (SM) processes investigated in this text. This process is $pp \rightarrow Wb\bar{b}$ at $\sqrt{s} = 8$ TeV with colliding protons provided by the Large Hadron Collider and detected by the Compact Muon Solenoid experiment.

5.1 Previous Measurements

The production of Z bosons **?????** or W bosons **??** in association with b jets has been studied at a center-of-mass energy of 7 TeV using data samples with up to 5 fb^{-1} of integrated luminosity, by the ATLAS and CMS experiments, as well as at the Tevatron **??** at $\sqrt{s} = 1.96$ TeV. This analysis extends previous measurements of the $Wb\bar{b}$ cross section **?**, using data at $\sqrt{s} = 8$ TeV, collected by the CMS detector in 2012 and corresponding to an integrated luminosity of 19.8 fb^{-1} **?**.

5.2 Event Selection

Two decay channels of the W boson are considered, $W \rightarrow \mu\nu_\mu$ and $W \rightarrow e\nu_e$, and events are selected using single-muon (single-electron) triggers with a loosely isolated

muon (electron) with transverse momentum $p_T > 24$ (27) GeV and pseudo-rapidity $|\eta| < 2.1$ (2.5). Individual particles emerging from each collision are then reconstructed with the particle-flow (PF) technique described in Section ?? which ultimately has the effect of dividing them into mutually exclusive categories: charged and neutral hadrons, photons, electrons, and muons.

Both the muon and electron candidates are required to have p_T larger than 30 GeV and $|\eta| < 2.1$ and to originate from the primary vertex of the event, chosen as the vertex with the highest $\sum p_T^2$ of the charged particles associated with it. These leptons must additionally pass a tight ID requirement described in Section ??, and must be isolated where the isolation variable is defined as

$$I = \frac{\sum p_T^{\text{charged}} + \max(0, \sum p_T^\gamma + \sum E_T^{\text{neutral}} - 0.5 \cdot p_T^{\text{PU}})}{p_T^\ell}, \quad (5.1)$$

with the sum running over the PF candidates (hadrons, electrons, photons) in a cone of size $\Delta R < 0.4$ (0.3) around the muon (electron) direction. The isolation includes a correction for pileup effects, which is based on the scalar sum of transverse momenta of charged particles not associated with the primary vertex in the isolation cone (p_T^{PU}). The selected muons (electrons) are required to have $I < 0.12$ (0.10).

Missing transverse energy, E_T^{miss} , is defined in Section ?? as the negative vector sum of the transverse momenta of all reconstructed particle candidates in the event. It is combined with the p_T of a muon or electron passing the identification and isolation requirements to form a W candidate. The transverse mass of the W boson, as defined in Section ??, is a natural discriminator against non-W final states such as QCD multijet events, that have a lepton candidate and E_T^{miss} , but a relatively low value of

m_T . In calculating m_T , the E_T^{miss} is corrected for noise in the electromagnetic and hadron calorimeters ?, and corrections to mitigate the effect of the pileup are also included ?.

Jets are constructed using the anti- k_t clustering algorithm ?, as implemented in the FASTJET package ??, with a distance parameter of 0.5. Jet clustering is performed using individual particle candidates reconstructed with the PF technique. Jets are required to pass identification criteria that eliminate jets originating from noisy channels in the hadron calorimeter ?. Those which originate from pileup interactions are rejected by requiring consistency of the jets with the primary interaction vertex. Small corrections to the relative and absolute jet energy calibrations of the detector are applied as a function of the p_T and η of the jet ?.

The combined secondary vertex (CSV) b-tagging algorithm ?? exploits the long lifetime and relatively large mass of b hadrons to provide optimized b jet discrimination. The CSV algorithm combines information about impact parameter significance, secondary vertex (SV) kinematic properties, and jet kinematic properties in a likelihood-ratio technique. The tagging of a jet is made by imposing a minimum threshold on the CSV discriminator value and in this analysis we require that both jets pass a threshold which has a b-tagging efficiency of about 40% and a misidentification probability of 0.1% for light jets and 1% for charm jets. The scale factors used to correct for the differences in efficiency between data and simulation take into account dependencies on the transverse momentum of the jet.

After all selection requirements for the signal enhanced dataset are applied, the contributing processes to the overall yield are the associated production of a massive

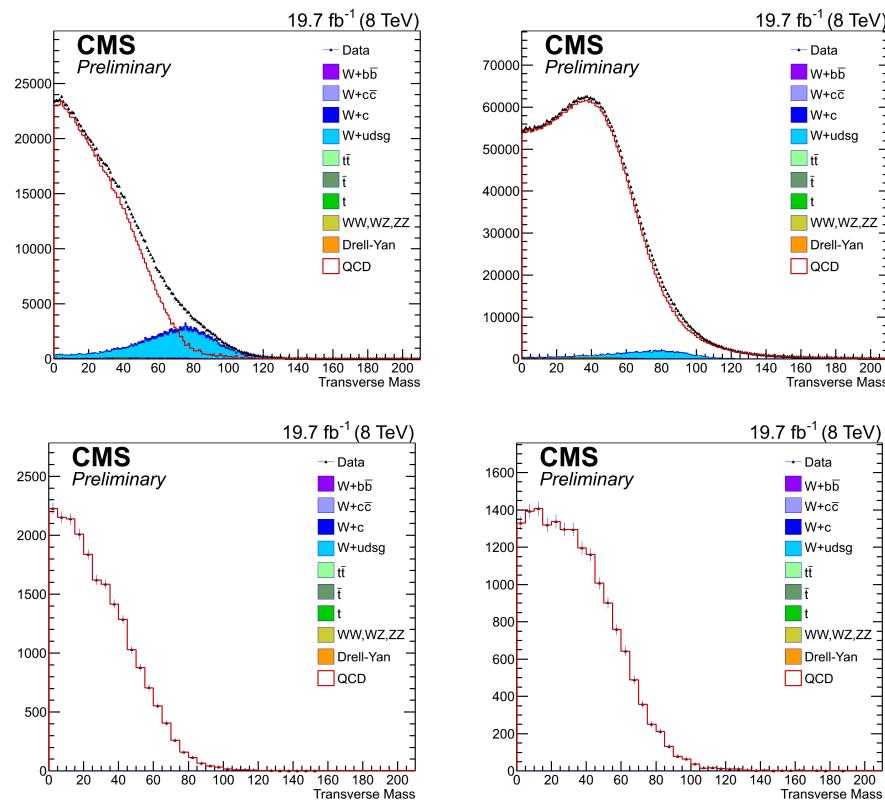
vector boson and jets ($V + \text{jets}$), as well as diboson (WW , WZ , ZZ), $t\bar{t}$, single top, $\gamma + \text{jets}$, and QCD multijet production. The corresponding contributions are estimated from simulation except for QCD, which is estimated from data as described in Section 5.3.1. To prove the validity of the MC shapes and normalizations a set of control regions is provided for the background contributions.

5.3 Background Estimation

5.3.1 QCD

The QCD multijet sample derived using a data-driven method. The shapes of the distributions for QCD multijet events are taken as the difference between the data sample and the sum of the other simulated backgrounds in a region of phase space enriched in multijets. The shape of the QCD used in the signal region comes from the distributions illustrated in Fig. 5.1. This region is found using the same selection requirements as those in the signal region, but requiring the muon (electron) to be antiisolated: $I > 0.20$ (0.15). In the fiducial regions used in this analysis, minimal correlation is observed between I and m_T , validating the use of an inverted isolation requirement to obtain the QCD shape. This shape is then scaled by $(d_{20} - m_{20})/q_{20}$ where d_{20} is the yield in data in the range $0 < m_T < 20$, m_{20} is the combined yield from the simulated samples in this range, and q_{20} is the corresponding unnormalized yield of QCD multijet. This has the effect of normalizing the QCD sample such that the combination of the QCD and the simulated backgrounds has the same total yield as data in the range $0 < m_T < 20$. If $d_{20} < m_{20}$, the QCD contribution is taken

Figure 5.1: The shape for the QCD is found by inverting the lepton isolation and subtracting MC from the data. Shown above is the data, MC background and extrapolated QCD shape (difference between data and MC backgrounds) in this inverted region for both the muon and electron channels in the $W + jj$ and $W + b\bar{b}$ phase spaces. The requirement of two well-identified b tags essentially eliminates all MC backgrounds in the $W + b\bar{b}$ region, leaving the QCD shape the same as that of the data.



to be negligible. The relative uncertainty in the yield of QCD multijet events is estimated to be $\pm 50\%$, taking into account both the fit result and the extrapolation from $0 < m_T < 20$ to the high- m_T range. This relative uncertainty also covers shape mismodelings of the multijet contribution in the final sample.

5.3.2 W+jets: light and charm component

$W + \text{jets}$ is the dominating background in the $W + jj$ phase space, which is found using identical selections as are used in the signal region with the exception of the b tag requirement; in the $W + jj$ phase space, no requirements on b tags are made. This control region therefore serves as a cross check on the reconstructed objects observed in the signal region before the added complication of b tagging has been introduced. In Figure 5.2 is shown the p_T of the identified lepton along with the E_T^{miss} and m_T in both decay channels. Agreement between simulation and data is on the order of 10%.

The true Wc contribution in the signal region is minimal, and only possible due to mistaging of a second jet in the event. However, the contribution of events with one hard charm and one hard anti-charm originated by gluon splitting is not negligible and moreover these events have kinematics closely related to that of our signal.

5.3.3 Top backgrounds

To validate the description of the $t\bar{t}$ contribution two control regions are defined and referred to as multilepton and multijet $t\bar{t}$ regions. The selections for the multijet region are the same as those for the $Wb\bar{b}$ region except that additional jet activity is required by selecting events with at least three jets in the final state. Because of the loosening of the jet veto in this phase space, it is less sensitive to the effects of jet energy scaling than the signal region (2% in the $t\bar{t}$ multijet phase space, 6% in the $Wb\bar{b}$ phase space.) As can be seen in Figure 5.3, this control region is dominated by $t\bar{t}$, with $t\bar{t}$ accounting for 80% of the data yield.

Figure 5.2: Selecting for a tight ID muon with $p_T > 30$ GeV and exactly two central jets passing loose ID, we recover the distributions shown above. Shown in the upper left (right) is the momentum of the leading lepton in the muon (electron) channel. The missing transverse energy is shown in the center, and transverse mass is given in the bottom two distributions. The shaded band in ratio plots shows statistical uncertainty.

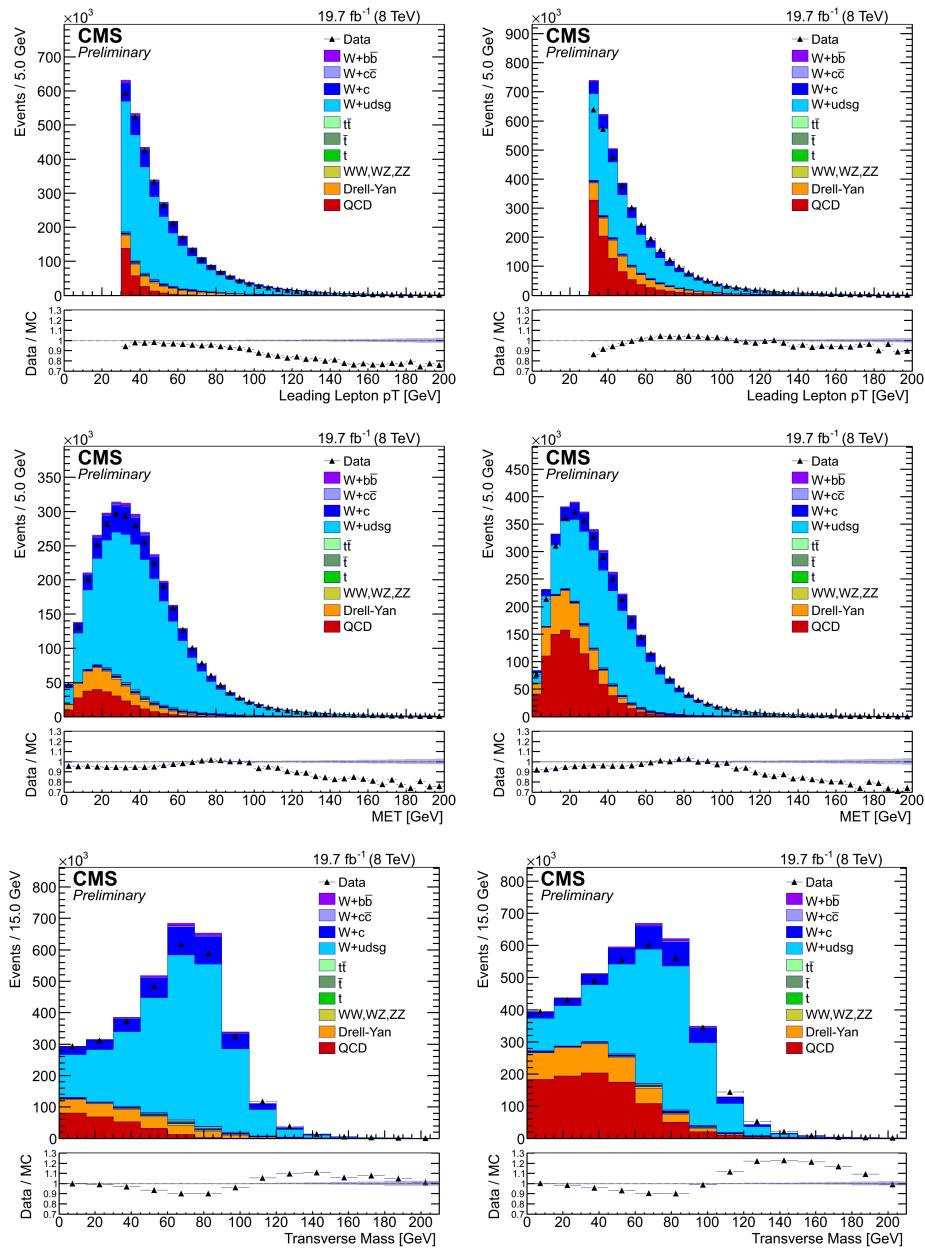
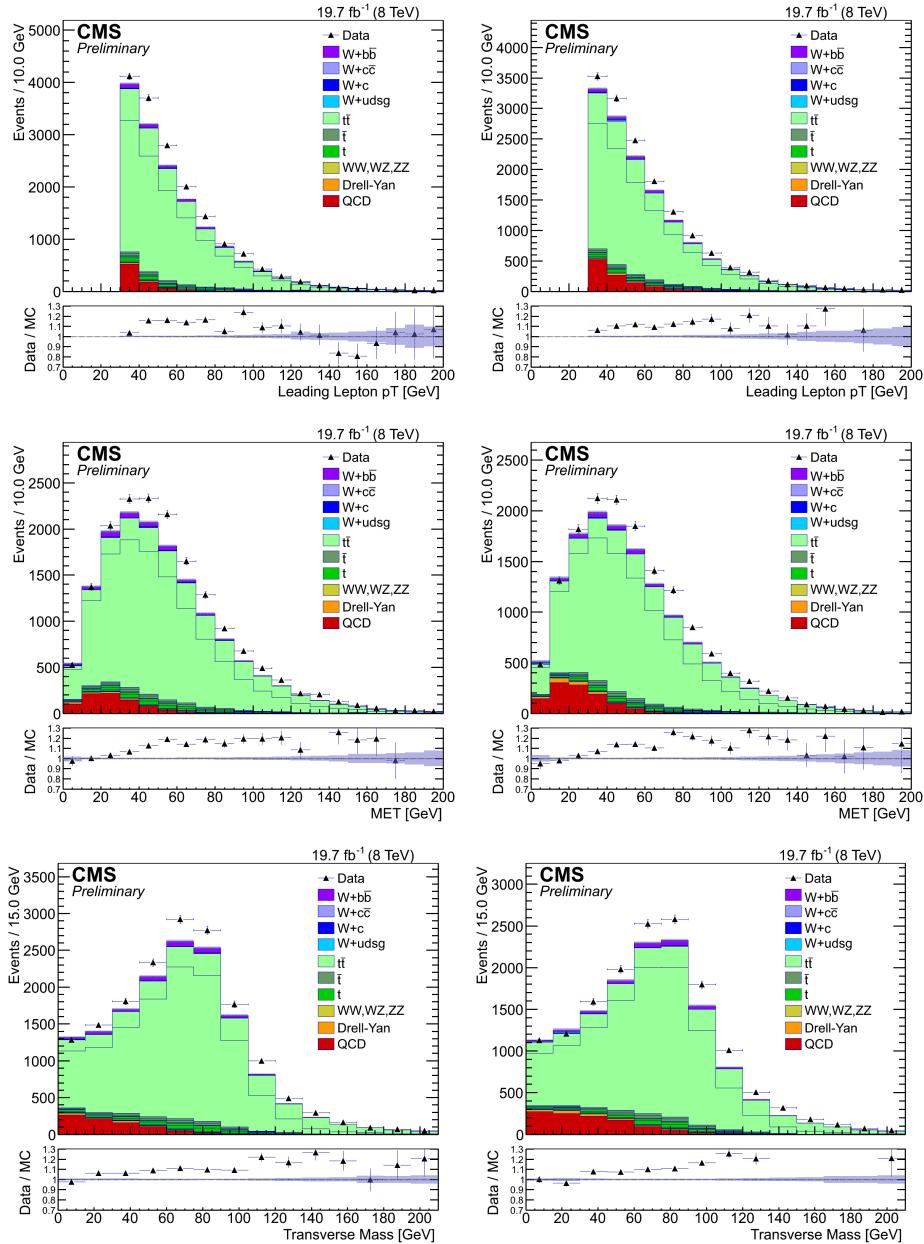


Figure 5.3: Distributions in the $t\bar{t}$ multijet control region are shown here in both channels. These raw distributions are made before any of the scaling outlined in Section 5.4. Left plots are in the muon decay channel and right plots are in the electron decay channel.



The selections for the multilepton region differ from those for the $Wb\bar{b}$ region in that exactly two well-isolated opposite-flavor leptons are required. Figure 5.4 shows representative distributions in this phase space where $t\bar{t}$ accounts for over 95% of the simulated samples.

The single top control region is defined by the signal selection requirements without the third and forward jet vetos, and with the leading jet required to be central ($|\eta| < 2.4$) and tightly b-tagged, while the subleading jet has no b requirement and must fall within $2.4 < |\eta| < 5.0$. As illustrated in Figure 5.5, there are many backgrounds contaminating the purity of this phase space, but agreement between data and simulation is on the order of 5-10%.

5.3.4 $Z\ell\bar{\ell}$ backgrounds

The Drell-Yan background is validated in a control region where the $Wb\bar{b}$ selection requirements are applied, but the lepton veto is inverted, requiring two isolated, same-flavor leptons and the m_T requirement is dropped. This is referred to as the $Zb\bar{b}$ region and distributions of the mass and transverse momentum of the dilepton pair is shown in Figure 5.6. Contamination from $t\bar{t}$ is evident. A cleaner Drell-Yan phase space is found by requiring exactly two jets but placing no b tag requirement and is referred to as Zjj . Figure 5.7 shows the same distributions as Figure 5.6 in this phase space.

Figure 5.4: Distributions in the $t\bar{t}$ multilepton control region are shown here in both channels. These raw distributions are made before any of the scaling outlined in Section 5.4. Left plots are in the muon decay channel and right plots are in the electron decay channel.

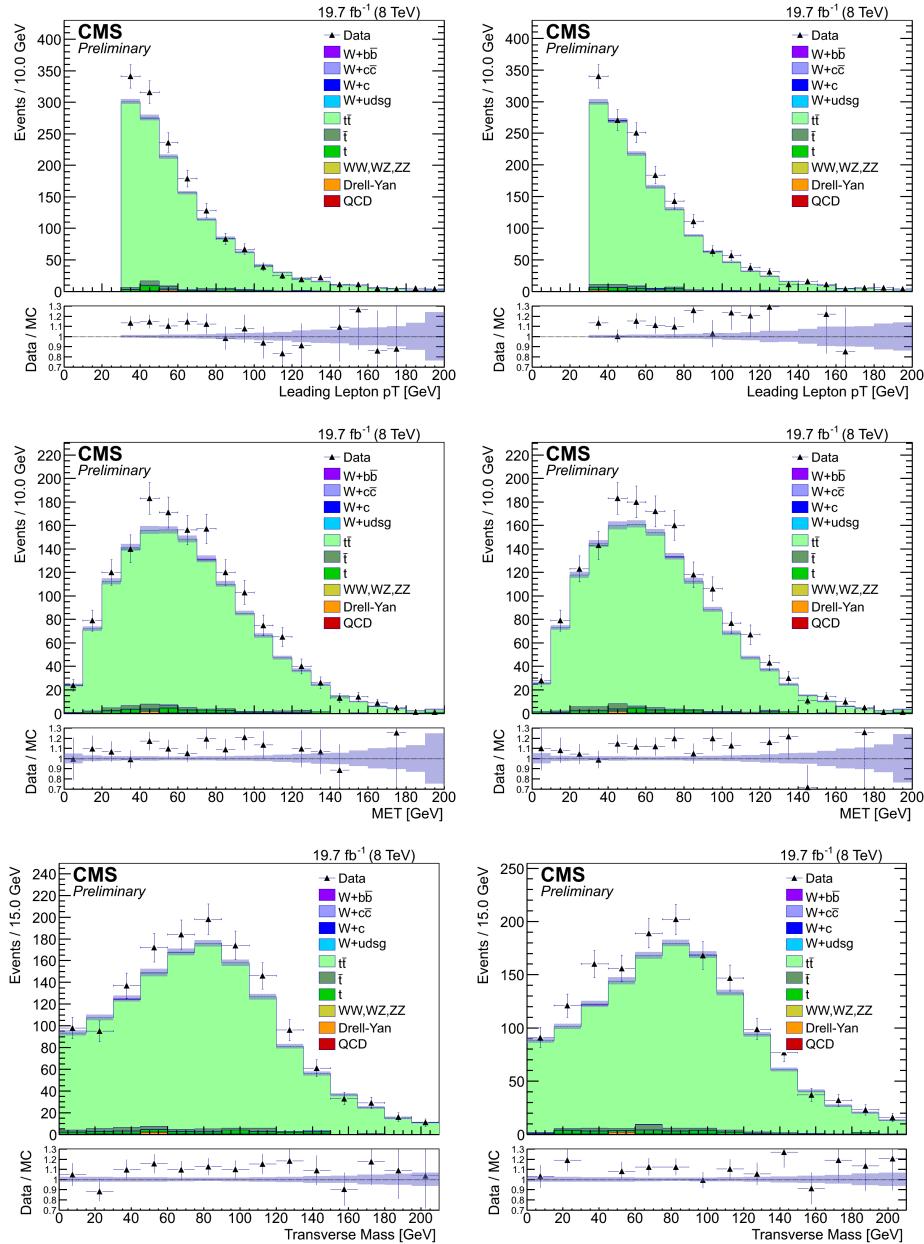


Figure 5.5: The single top control region is defined by one b -tagged central jet and one forward jet. Shown above are distributions in the single top control region. Left plots are in the muon decay channel and right plots are in the electron decay channel.

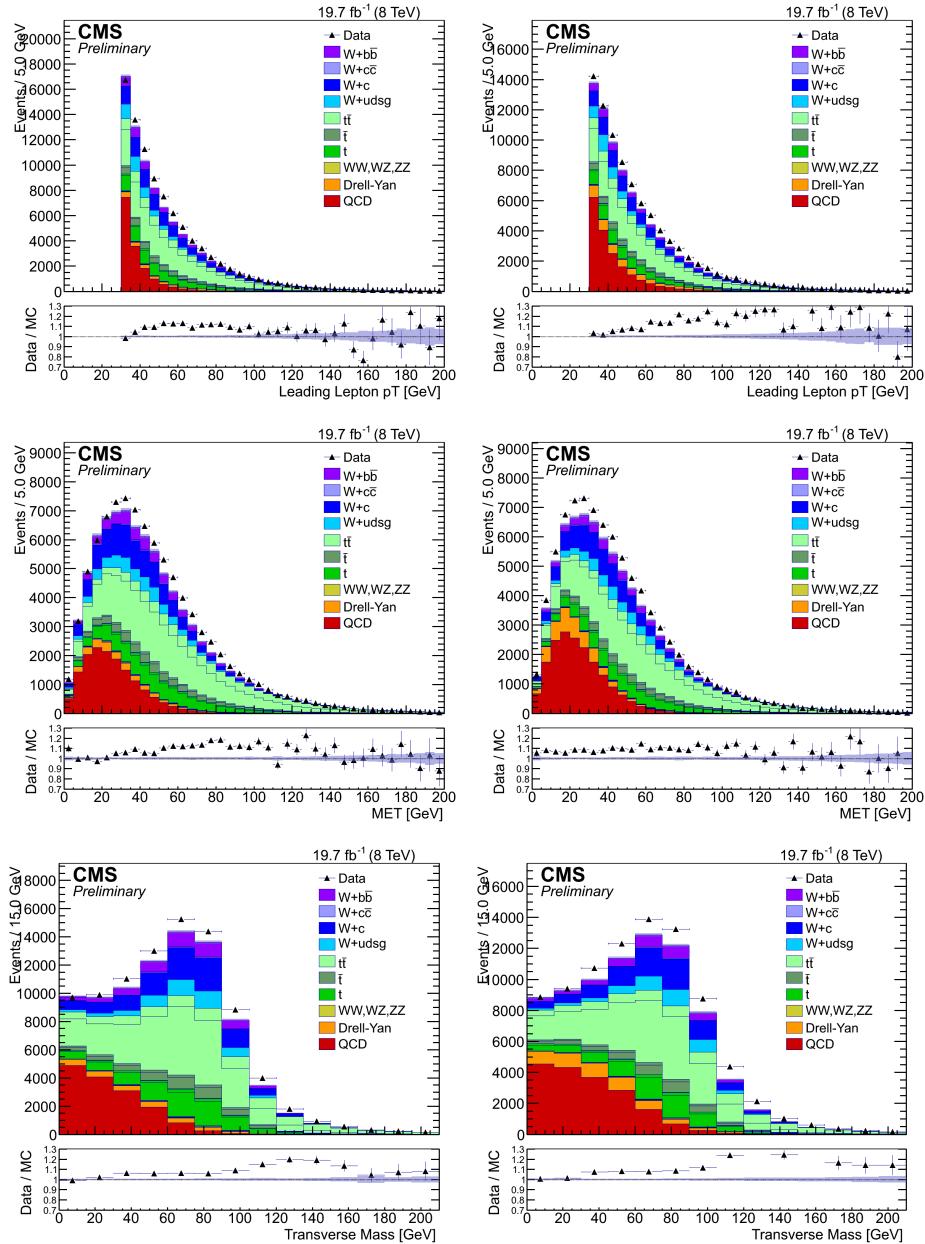
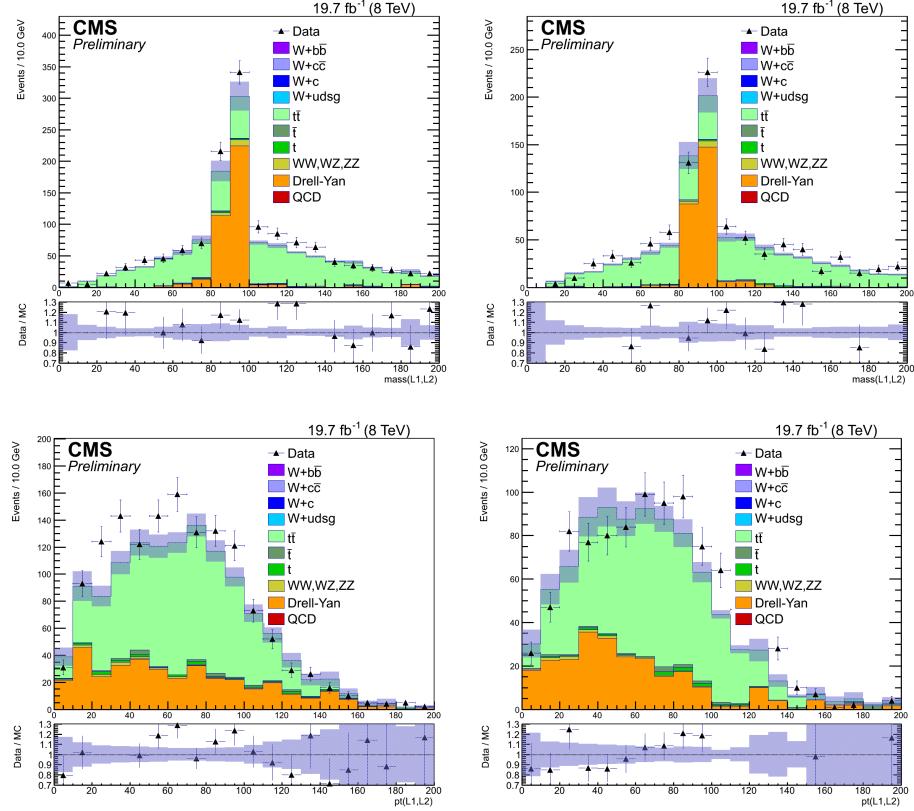


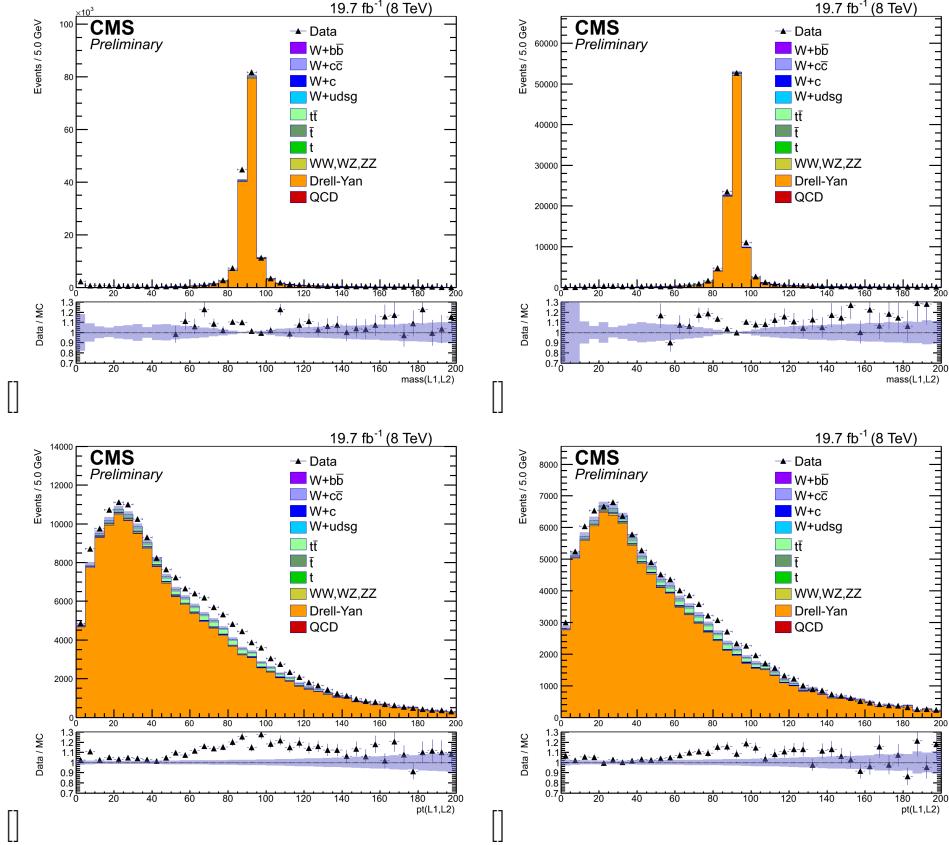
Figure 5.6: Above are distributions of the mass and transverse momentum of the dilepton pair in the $Zb\bar{b}$ phase space. Left plots are in the muon decay channel and right plots are in the electron decay channel.



5.4 Analysis Strategy

The $Wb\bar{b}$ yield is ultimately measured using a likelihood fit to the m_T distribution in the signal region, after having rescaled the simulation. Since the dominant background in the signal region arises from the $t\bar{t}$ process, the data and simulation are compared in two $t\bar{t}$ -dominated control regions. The simulation is reweighted to describe the control regions and then is used to predict the m_T distributions in the signal region.

Figure 5.7: Above are distributions of the mass and transverse momentum of the dilepton pair in the Zjj phase space. Left plots are in the muon decay channel and right plots are in the electron decay channel.



The signal region requires a muon (electron) with $p_T > 30$ GeV, pseudorapidity $|\eta| < 2.1$, and satisfying $I < 0.12$ (0.10). Exactly two b-tagged jets with $p_T > 25$ GeV and $|\eta| < 2.4$ are also required. Events with additional leptons with $p_T > 10$ GeV and $|\eta| < 2.4$ or a third jet with $p_T > 25$ GeV and $|\eta| < 5.0$ are rejected. The $t\bar{t}$ -multijet control region is obtained using the same selection criteria as in the signal region, but requiring at least three jets in the event with $p_T > 25$ GeV and $|\eta| < 2.4$ instead

of vetoing events which have more than two. The $t\bar{t}$ -multilepton control region uses similar selection criteria as the signal region, but changing the lepton requirement from vetoing events which contain a second lepton, to requiring two isolated leptons of different flavor, both with $p_T > 30$ GeV and $|\eta| < 2.1$.

The shape of the $Wb\bar{b}$ signal distribution is obtained by separating the $W + \text{jets}$ simulated sample into three substitution samples labeled as $Wb\bar{b}$, $Wc\bar{c}$, and $Wusdcg$. The separation is done at the truth generator level. If an event contains a b jet, from matrix element or parton shower, it falls into the $Wb\bar{b}$ category. A b jet at generator level requires the presence of a b hadron within a cone of radius $R = 0.4$ with respect to the jet axis. The jets are constructed at the generated level using all stable particles in the event (excluding neutrinos). Jets with a distance smaller than $R = 0.5$ with respect to a lepton are removed from the event. If an event contains no b jets but an even, non-zero, number of charm jets, again from matrix element or parton shower, it falls into the $Wc\bar{c}$ category. The remaining events fall into the $Wusdcg$ category. The energy of the selected leptons at the generated level is corrected for the final state radiation (FSR) by summing up the four-momenta of all the photons generated within a cone of radius $R = 0.1$ around the lepton. Generated leptons originating in simulation from the decay of b hadrons or τ leptons are not considered.

Two major parameters in the simulations significantly affect the shape and normalization of the simulated distributions: the b-tagging efficiency and the jet energy scale (JES). Both control and signal regions show similar sensitivity to the b-tagging efficiency, and its adjustment affects all the regions in a correlated manner. The JES affects more the $t\bar{t}$ predictions in the signal and multilepton regions where a veto on

the third jet is applied. The effect on the leading jets is moderate, but because of the multijet nature of the $t\bar{t}$ production, the JES variations lead to significant migration of jets into and out of the veto region. The $t\bar{t}$ -multijet control region, since it has no veto on the third jet, is less sensitive to JES. The same is observed to be true for the $Wb\bar{b}$ LO sample in all regions and may be due in part to the fact that $Wb\bar{b}$ has no additional jets to veto.

The fit procedure thus consists of three steps. First, using the simulated samples detailed above, a fit is performed in the $t\bar{t}$ -multijet region using the m_T variable. The result of this fit gives an estimation of the b-tagging efficiency rescaling factor, which is measured separately in the muon and electron channels and averaged. The reweighted samples are then used in the next step where a fit to the m_T variable in the $t\bar{t}$ -multilepton region is performed and the jet energy scale in simulation is adjusted. As a result of these two steps, the simulation is expected to properly describe the $t\bar{t}$ contribution and the final step is to extract the number of $Wb\bar{b}$ events from a fit in the signal region.

5.5 Systematic Uncertainties

The major sources of the systematic uncertainties are listed in Table 5.1. The size of the variation is presented for each uncertainty source together with its effect on the signal. Some of the uncertainties affect only the normalization of the respective contributions, e.g. the uncertainty on the theoretical cross sections. In other cases, the uncertainties affect also the shape of the corresponding m_T distributions, and these

are listed in the table under “norm. + shape”. In the fit procedure, interpolations are performed following a log normal distribution for uncertainties affecting only normalizations and a quadratic distribution for uncertainties affecting both the shapes and normalizations.

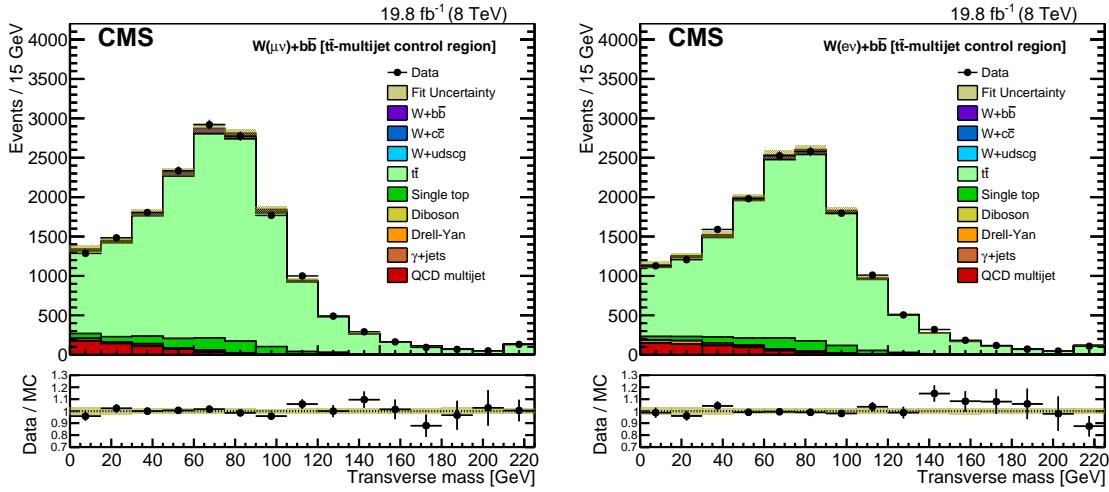
The 50% uncertainty on the QCD background is taken as a conservative estimate from the control region, and increasing this uncertainty does not lead to any change in the result of the fit. The common b-tagging and JES rescaling factor uncertainties are set to $\pm 100\%$ of the factor itself, allowing the fit to remeasure them in the signal region. The scale uncertainties are estimated by changing the renormalization and factorization scales up and down by a factor of two. The PDF uncertainties are estimated from the change in acceptance found by varying the PDF set.

5.6 Signal Extraction

The fit in the $t\bar{t}$ -multijet region is used to obtain rescaling factors for the muon and electron channels separately to better describe the b-tagging efficiency in the simulation, as presented in Section 5.4. The results of the fit are presented in Fig. 5.8. The measured rescaling factors, 1.17 ± 0.12 (muon channel) and 1.13 ± 0.11 (electron channel), are averaged to 1.15 ± 0.14 , where the uncertainty allows the following fits to vary the rescaling factor between 1.01 and 1.29. The simulation is rescaled accordingly for the next fit and the uncertainty on the rescaling is included in the fit procedure.

A fit to the $t\bar{t}$ -multilepton region adjusts the jet energy scale, as described in

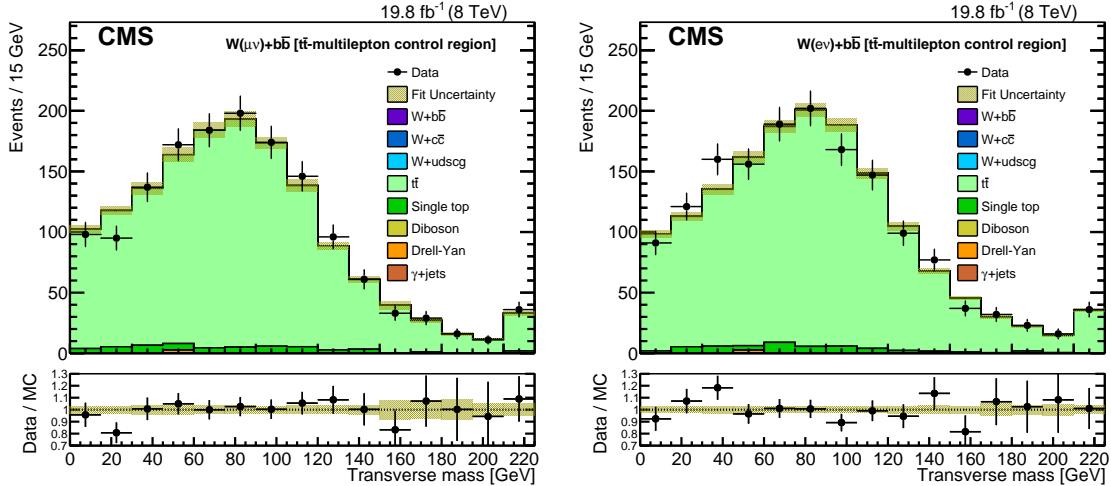
Figure 5.8: The transverse mass distributions in the $t\bar{t}$ -multijet phase space after fitting to obtain the b -tag rescale factors. The lepton channels are shown separately with the muon sample on the left and the electron sample on the right. The highest bin contains overflow events. The shaded area represents the total uncertainty on the simulation as output from the fit.



Section 5.4. As a result, the simulated m_T distributions change both the shape and normalization. The best fit suggests changing the JES by approximately $1.3\sigma_{\text{JES}}$ from its central value. Figure 5.9 shows the results of the fits in the $t\bar{t}$ -multilepton enhanced data set for the muon channel (left) and the electron channel (right). The JES is therefore shifted by $1.3\sigma_{\text{JES}}$ in the simulation with the uncertainty set to $1.3\sigma_{\text{JES}}$. Thus the simulation is tuned to describe the $t\bar{t}$ control regions and is used to extract the signal yield in the signal region.

The results of the fit in the $Wb\bar{b}$ signal region are presented in Fig. 5.10. All background contributions are allowed to vary in the fit within their uncertainties, while the $Wb\bar{b}$ normalization remains a free parameter of the fit. The correlation

Figure 5.9: The transverse mass distributions in the $t\bar{t}$ -multilepton enhanced data set after fitting to find the appropriate jet energy scale. The lepton channels are shown separately with the muon sample on the left and the electron sample on the right. The highest bin contains overflow events. The shaded area represents the total uncertainty on the simulation as output from the fit.

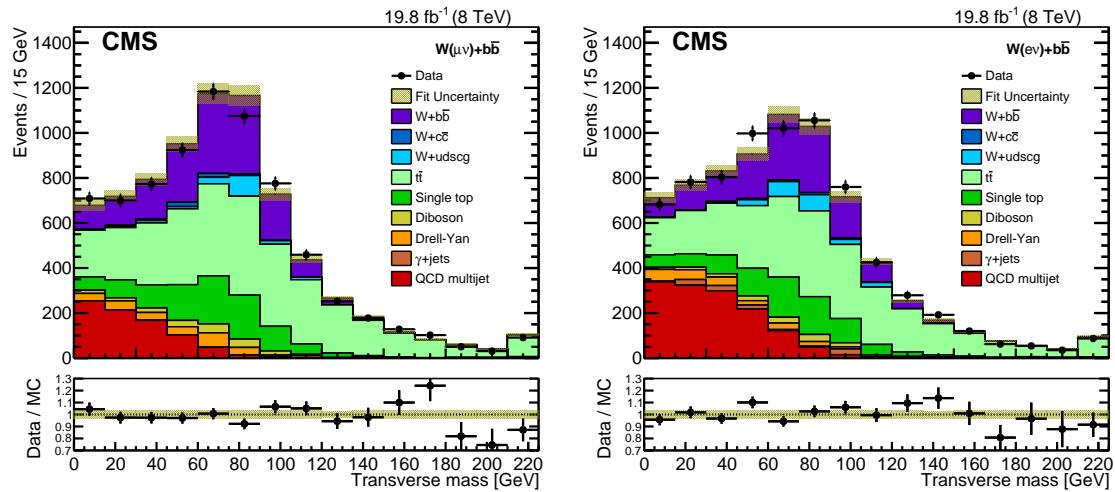


between different sources of uncertainties is taken into account. The composition of the event sample in the signal region is summarized in Table 5.2. Events coming from the production of a Higgs boson in association with a vector boson constitute a negligible fraction of the overall event yield.

Distributions for variables other than those being directly fitted are also produced by applying the results from the three fits to the simulated samples. In the signal phase space, lepton isolation was found to be essentially uncorrelated with the shape of the transverse mass variable for QCD events, but this was not the case for the ΔR distance between the two b-tagged jets, $\Delta R(b\bar{b})$, or the lepton p_T . The shape of the QCD distribution for these variables was therefore taken from an $m_T < 30$ GeV sideband

(as opposed to the inverted isolation sideband used for the m_T variable) and the normalization set to the final fitted normalization given in Table 5.2. Distributions of $\Delta R(b\bar{b})$ and p_T^ℓ in the combined lepton channel are presented in Fig. 5.11. Agreement between data and simulation is observed.

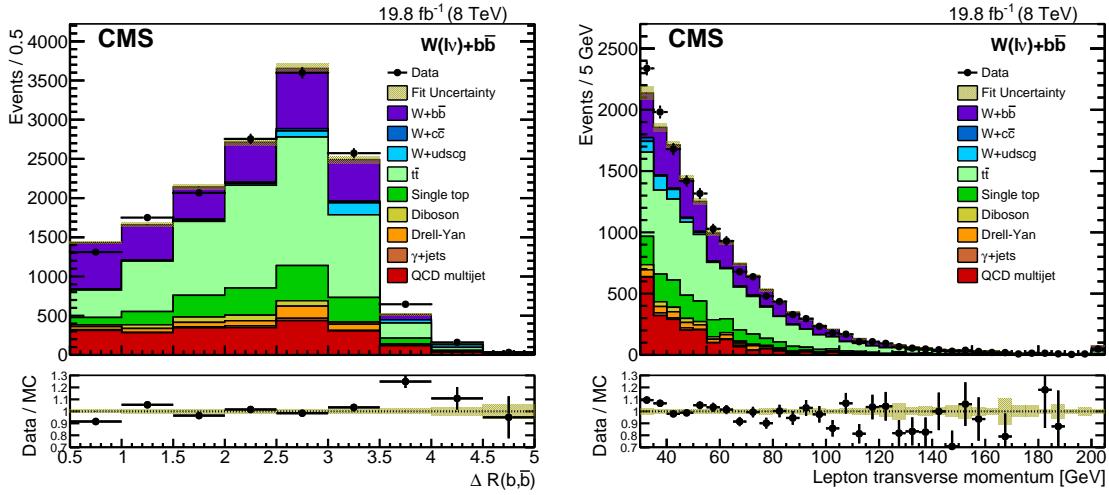
Figure 5.10: Transverse mass distributions in the $Wb\bar{b}$ signal region after fitting simultaneously muon and electron decay channels. The lepton channels are shown separately with the muon sample on the left and the electron sample on the right. The highest bin contains overflow events. The shaded area represents the total uncertainty on the simulation as output from the fit.



5.7 Cross Section and Comparisons

The cross section for the $Wb\bar{b}$ process, $\sigma(pp \rightarrow Wb\bar{b})$, is derived from the signal strength measurement as obtained from the fit. The cross section is written as

Figure 5.11: Distributions of $\Delta R(b\bar{b})$ and p_T^ℓ after applying the results from the fits to the simulation. The QCD shape was taken from an $m_T < 30$ GeV sideband and the electron and muon channels have been combined in these distributions. The highest bin contains overflow events and the shaded area represents the total uncertainty on the simulation as output from the fit.



$$\sigma(pp \rightarrow Wb\bar{b}) = \frac{N_{\text{signal}}^{\text{Data}}}{A \cdot \epsilon \cdot \mathcal{L}} = \frac{N_{\text{signal}}^{\text{Data}}}{(N_{\text{signal}}^{\text{MC}}/N_{\text{generated}}^{\text{MC}}) \cdot \mathcal{L}} = \alpha \sigma_{\text{gen}}$$

where $N_{\text{signal}}^{\text{Data}}$ is the number of observed signal events, $N_{\text{signal}}^{\text{MC}}$ is the number of expected signal events from simulation, $N_{\text{generated}}^{\text{MC}}$ is the number of generated events in the fiducial region, A , ϵ are the acceptance and efficiency correction factors, α is the measured signal strength in the given lepton channel, and σ_{gen} is the simulated fiducial cross section of the signal sample.

In this analysis, the fiducial cross section was calculated in the following manner: Madgraph is used to compute the $Wb\bar{b}$ cross section with fiducial cuts applied. Then

a k-factor for inclusive W production is applied, obtained from the ratio of the inclusive W cross sections calculated with FEWZ (at NNLO using the five-flavour CTEQ 6M PDF set) and with Madgraph. The product $A \cdot \epsilon$ is 11 (13)% in the muon (electron) channels and results from the combined effect of the efficiency from lepton identification requirements (80%), and b tag efficiency (40% per jet). The uncertainty on this product is 10% as listed in the bottom row of Table 5.1, which was calculated by varying the PDF set using the LHAPDF/PDF4LHC **????** prescription considering PDF sets from CTEQ, MSTW, NNPDF, and HERA as well as varying the choice of scales μ_F , μ_R simultaneously up and down by a factor of two.

The $Wb\bar{b}$ cross section is measured within a fiducial volume, which is defined by requiring leptons with $p_T > 30\text{GeV}$ and $|\eta| < 2.1$ and exactly two b-tagged jets of $p_T > 25\text{GeV}$ and $|\eta| < 2.4$. The measured cross sections are presented in Table 5.3. The combination of the muon and electron measurements is done using a simultaneous fit to both channels, taking into account correlations between different sources of uncertainties.

The measured cross sections are compared to theoretical predictions from MCFM **??** with the MSTW 2008 PDF set, as well as from MADGRAPH 5 interfaced with PYTHIA6 in the four- and five-flavour schemes and MADGRAPH 5 with PYTHIA8 **?** in the four-flavour scheme. In the five-flavour scheme, the PDF set CTEQ 6L was used and PYTHIA6 was run using TuneZ2*. The two four-flavour samples were produced using a NNLO PDF set interfaced with PYTHIA (version 6 in one sample, version 8 in the other) in the CUETP8M1 tune.

Comparisons between the results of calculations performed under different assump-

tions provide important feedback on the functioning and validity of the techniques employed. Differences in predictions arising from the modelling of b quarks as massive or massless are possible, as are variations in predictions arising from the use of different showering packages (PYTHIA6 vs. PYTHIA8) or matrix element generators (MADGRAPH vs. MCFM). In the phase space explored here, these predictions are all very close in their central value and agree with each other well within their respective uncertainties.

The MCFM cross section calculation is performed at the level of parton jets and thus requires a hadronization correction. The multiplicative hadronization correction factor 0.81 ± 0.07 is calculated using the MADGRAPH + PYTHIA6 sample and agrees well with a similar factor calculated in the 7TeV Z+b analysis calculated as 0.84 ± 0.03 ?. The correction factor is obtained for jets computed excluding neutrinos from the particle list, as such jets are closer in kinematics to particle jets at the detector level. The uncertainty reflects both the statistics of the MADGRAPH + PYTHIA6 sample as well as a comparison with the MADGRAPH + PYTHIA8 sample.

The MCFM and four-flavour MADGRAPH predictions do not account for $Wb\bar{b}$ production where the $b\bar{b}$ system comes from multiple parton scattering. CMS simulations of MADGRAPH + PYTHIA events that include double parton interaction (DPS) reproduce the W + jets data ?, therefore a MADGRAPH + PYTHIA8 sample of a W boson produced in association with a $b\bar{b}$ pair coming from DPS was generated to study the effect on the fiducial cross section. Using this dedicated sample, an additive correction σ_{DPS} is estimated to be 0.06 ± 0.06 pb, where the uncertainty is

conservatively assigned to be 100% of the value.

The uncertainty in the theoretical cross sections arising from the choice of PDF is also accounted for, using the LHAPDF/PDF4LHC **????** prescription in which PDF sets from CTEQ, MSTW, NNPDF, and HERA are considered. Uncertainties in the theoretical cross section due to the choice of scale are also estimated by varying the scales μ_F , μ_R simultaneously up and down by a factor of two.

The resulting cross section predictions in the fiducial phase space at the hadron level and including the estimated hadronization and DPS corrections when needed are compared in Fig. 5.12 with the measured value. Within one standard deviation the predictions agree with the measured cross section. The results also agree within one standard deviation with previously published $Wb\bar{b}$ measurements at 7 TeV, where data are found to be well described by the same predictions.

Figure 5.12: Comparison between the measured $Wb\bar{b}$ cross section and various QCD predictions. The blue error bars on the predictions represent the uncertainty in the given sample associated with PDF choice and the black bars represent the total uncertainty. In the case of the MADGRAPH + PYTHIA6 (5F) sample, the effects of DPS are already included in the generated sample so the extra DPS factor was not needed and the blue and black error bars overlap perfectly.

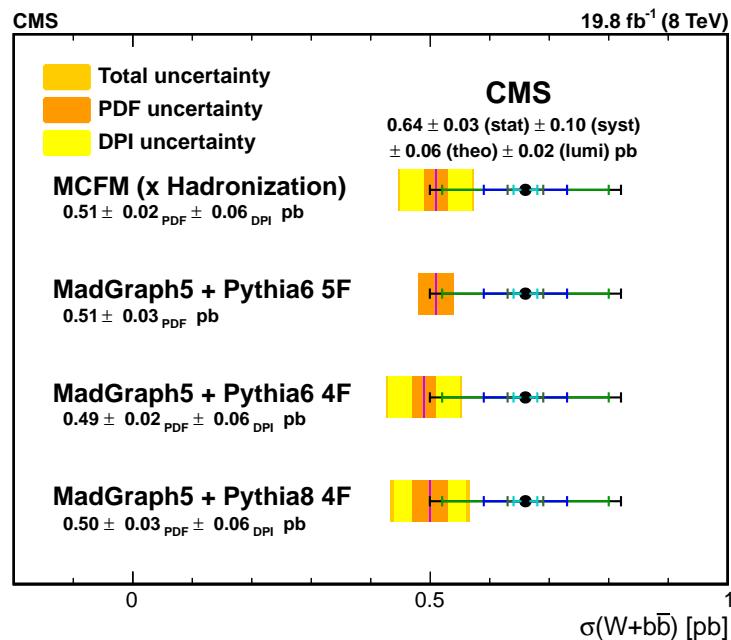


Table 5.1: Breakdown of the major sources of systematic uncertainty in the $Wb\bar{b}$ phase space. The last column indicates the contribution of the given systematic to the overall uncertainty on the measured cross section. The uncertainty labeled "b tag rescale" is the uncertainty associated with the rescaling of the b tag efficiency scale factors. In the "variation" column, the uncertainties which are correlated across all simulated samples and affect both shape and normalization are indicated by σ_X to indicate that an input variation of one standard deviation is set on the uncertainty X in the fitting procedure. UES refers to the energy scale of energy deposits not clustered into jets and MES and EES refer to the muon and electron energy scales. The uncertainty labeled as "Id/Iso/Trg" is the uncertainty associated with the efficiency of the lepton identification, isolation, and triggering. The uncertainty on the luminosity and the uncertainty on the acceptance due to PDF and scale choices are not included in the fit, and are treated separately.

		uncertainty	variation	effect on the measured cross section	
normalization	uncorrelated	$t\bar{t}$	7.4%	10.5%	
		Single Top	5.4%	3.3%	
		$Wusdcg$	13.2%	< 2%	
		$Wc\bar{c}$	8.1%	< 2%	
		Diboson	8.1%	< 2%	
		Drell-Yan	7.9%	< 2%	
		$\gamma+$ jets	10.0%	< 2%	
		QCD	50%	1-4%	
norm. + shape	correlated	b tag rescale	12.9%	13.0%	
		JES rescale	$1.3 \times \sigma_{\text{JES}}$	3.6%	
	correlated	JES	σ_{JES}	5.6%	
		UES	σ_{UES}	2.8%	
		MES	σ_{MES}	3.6%	
		EES	σ_{EES}	< 2%	
		Id/Iso/Trg	$\sigma_{\text{Id/Iso/Trg}}$	< 2%	
luminosity				2.6%	
theory (scale+PDF)				10%	

Table 5.2: Initial and final yields obtained in the $Wb\bar{b}$ signal region. The uncertainties on the signal strength represent the total uncertainty of the measurement.

	muon		electron	
	Initial	Fitted	Initial	Fitted
Data	7432		7357	
$Wb\bar{b}$	1322.7	1731.0	1120.5	1495.3
$Wc\bar{c}$	59.7	57.4	36.0	37.2
$Wusdcg$	181.7	173.8	220.0	214.3
$t\bar{t}$	3048.6	3276.8	2639.6	2891.7
Single Top	958.0	980.1	820.2	855.5
Drell-Yan	261.2	262.3	220.4	225.1
Diboson	175.4	178.9	138.9	144.4
$\gamma+$ jets	0.0	0.0	98.3	104.7
QCD	1109.0	816.9	1653.8	1348.0
Signal strength	1.21 ± 0.24		1.35 ± 0.28	
Combined	1.29 ± 0.22			

Table 5.3: Measured cross sections in the muon, electron, and combined lepton channels.

Channel	$\sigma(pp \rightarrow Wb\bar{b})$, pb
Muon	$0.62 \pm 0.04(\text{stat}) \pm 0.14(\text{syst}) \pm 0.06(\text{theo}) \pm 0.02(\text{lumi})$
Electron	$0.69 \pm 0.05(\text{stat}) \pm 0.19(\text{syst}) \pm 0.07(\text{theo}) \pm 0.02(\text{lumi})$
Combined	$0.66 \pm 0.03(\text{stat}) \pm 0.14(\text{syst}) \pm 0.07(\text{theo}) \pm 0.02(\text{lumi})$

6 MONOPHOTON ANALYSIS [TODO WESLEY COMMENTS]

While studying the Standard Model process $pp \rightarrow Wb\bar{b}$ is a means for probing the physics associated with jets, hadronization and the modeling of the proton parton distribution function, the monophoton analysis, characterized by a final state consisting of $\gamma + E_T^{\text{miss}}$ is also of interest for validating SM and possibly extending it via the search for dark matter (DM). One interpretation presented in this work is the SM process $pp \rightarrow Z\gamma \rightarrow \nu\bar{\nu}\gamma$ in which the E_T^{miss} is interpreted as coming from the invisible decay of the Z boson, $Z \rightarrow \nu\bar{\nu}$. As discussed in Section ??, it is also possible to extend this cross section measurement into a dark matter search under the interpretation that the E_T^{miss} arises from the annihilation of incoming particles into DM and the γ is initial state radiation recoiling against the process. Under this interpretation, limits are set on the cross section of DM as a function of the mediator mass for vector and axial-vector models. This analysis is performed using protons colliding at $\sqrt{s} = 13$ TeV provided by the LHC and detected by the CMS detector.

6.1 Previous Measurements

A similar search was performed by the ATLAS experiment ? using 3.2fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV collected in 2015, and no evidence for new physics was found. The fiducial cross section for the monophoton final state with photon

$E_T^\gamma > 150\text{GeV}$ and $E_T^{\text{miss}} > 150\text{GeV}$ was measured to be below 17.8 fb^{-1} . A prior search ? at the CMS experiment used the LHC run 1 data collected in 2012 at $\sqrt{s} = 8 \text{ TeV}$, and set a cross section upper limit for the monophoton final state with photon $E_T^\gamma > 145\text{GeV}$ and $E_T^{\text{miss}} > 140\text{GeV}$ at 13 fb^{-1} .

6.2 Event Selection

The sample of data analyzed was collected using a trigger that requires at least one photon candidate with $E_T^\gamma > 165\text{GeV}$ and requires at least 90% of the energy deposited in the calorimeters to be deposited in the ECAL to reject jets. This trigger is 98% efficient at selecting photons which pass the other analysis selections. Events passing the trigger are further required to have at least one photon with $E_T^\gamma > 175\text{GeV}$ in the barrel fiducial region ($|\eta| < 1.44$).

To distinguish photons from electrons, which leave a similar signature of energy deposits in the ECAL and HCAL, candidate photons are required to not have any associated track seeds in the pixel detector. To distinguish photons from jets, selections based on calorimetric information and isolation are applied. The fraction of energy deposited in the ECAL compared to the total deposit in the calorimeters is tightened relative to the trigger to be 95% and the shower shape variable in the η direction, $\sigma_{i\eta i\eta}$, described in Section ??, is required to be $\sigma_{i\eta i\eta} < 0.0102$.

For a photon object to be considered isolated, scalar sums of the transverse momenta of PF charged hadrons, neutral hadrons, and photons within a cone of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.3$ around the candidate photon must individually fall

below bounds defined for 80% signal efficiency. Only the PF candidates that do not overlap with the electromagnetic shower of the candidate photon are used in the isolation sums.

Because photon objects are not reconstructed from tracks, there is an ambiguity in identifying the collision vertex that the photon originates from in the presence of pileup collisions.

Association of a vertex to the photon candidate impacts the photon in two ways. First, the photon momentum direction is given by a straight line connecting the ECAL cluster position and the vertex. Second, the isolation sum using the PF charged hadrons include only the candidates whose tracks are associated to the vertex. While the first effect is minor and is not relevant for this analysis, the second will cause photon candidates that are actually non-isolated to appear otherwise, if the vertex is misassigned. In practice, photon momentum is always computed with respect to the primary vertex, defined as the vertex with the highest $\sum p_T^2$ of the associated tracks. However, for the charged hadron isolation sum, all vertices are considered, and the maximum value of the isolation sum is used as a conservative estimate of the true isolation sum (worst isolation).

To reduce the contribution of backgrounds arising from occurrences in the CMS detector which did not originate from collisions, the pulse in the seed crystal of the photon cluster is required to be within $\pm 3\text{ns}$ of the time expected for particles from a collision, and the cluster must not be so narrow that it is consistent with a cluster formed by a single crystal. To reduce contamination from beam halo, the ECAL crystals not associated with the photon candidate are examined for evidence of the

passage of a minimum-ionizing particle roughly parallel to the beam axis (beam halo tag). If at least 4.9 GeV of energy is found deposited along this trajectory, the event is rejected.

The candidate events are required to have $E_T^{\text{miss}} > 170\text{GeV}$. The azimuthal opening angle between the candidate photon and E_T^{miss} is required to be greater than 2 to ensure that the main source of E_T^{miss} is not photon energy mismeasurement.

Because jet energy mismeasurement can also give rise to E_T^{miss} , events are rejected if the minimum azimuthal opening angle between E_T^{miss} and up to four leading jets ($\min\Delta\phi(E_T^{\text{miss}}, j)$) is less than 0.5.

Finally, events are also vetoed if they contain a charged lepton (an electron or a muon) with $p_T > 10\text{GeV}$ that is separated from the photon by $\Delta R > 0.5$.

After applying all of the selection criteria, 77 candidate events are found in data.

6.3 Estimation of Contributions

The dominant SM processes contributing to this phase space are the associated productions of a Z or W boson with a high-energy photon. If the Z boson decays into a neutrino-antineutrino pair, the final state exhibits a high- E_T photon and large missing transverse energy. Similarly, if the W boson decays into a lepton-neutrino pair and the lepton is outside of the detector acceptance or fails reconstruction, the event appears to be $\gamma + E_T^{\text{miss}}$. Together, these two processes account for approximately 75% of the simulated events and are estimated using Monte Carlo (MC) simulations. Hard-scattering events are generated with MADGRAPH5_AMC@NLO version 2.2 ?

at leading order (LO) in QCD, with NNPDF3.0 LO ($\alpha_s = 0.130$) as the parton distribution function. Parton shower and hadronization is performed by PYTHIA8.2 [??](#). Generated particles are processed through the full GEANT-based simulation of the CMS detector [??](#) and event reconstruction used for data. Minimum-bias simulations are overlaid to model pileup interactions.

6.3.1 Reweighting

To account for differences arising from imperfect modeling of the data in the simulation, a total correction factor ρ of 0.99 ± 0.06 is applied to all MC-based estimates. This is the product of individual correction factors taken as the ratios of the efficiencies measured in data and in simulation. They include 0.99 ± 0.016 for photon identification measured using $Z \rightarrow e\bar{e}$ events, 1.00 ± 0.0246 for pixel seed measured using $Z \rightarrow e\bar{e}$ events, and 1.00 ± 0.05 for worst isolation, beam halo tag and lepton veto measured using $Z\gamma \rightarrow \nu\bar{\nu}\gamma$ events. Generated samples are then weighted event by event with a product of two factors. The first factor matches the distribution of the generator-level photon p_T to that calculated at next-to-next-to-leading order (NNLO) in QCD using the DYRES [?](#) calculator. The second factor, taken from Refs. [??](#), further corrects this distribution to account for electroweak next-to-leading order (NLO) effects.

6.3.2 $V\gamma$ Estimates

After accounting for event selection efficiency difference between data and MC, respectively 42.1 ± 6.3 and 10.7 ± 1.5 events are estimated from $Z\gamma \rightarrow \nu\bar{\nu}\gamma$ and $W\gamma \rightarrow \ell\nu\gamma$. Four sources of systematic uncertainty on $Z\gamma$ and $W\gamma$ estimates are

considered: PDF and scale uncertainties, which are 5.37% and 8.9% respectively; electroweak correction uncertainties, where the full correction is conservatively taken as the uncertainty which is 11% for $Z\gamma$ and 7% for $W\gamma$; scale factors, which are 6%; and the systematic uncertainty due to jet/ E_T^{miss}/γ energy scale and pileup, which is 6.2%. As a crosscheck, the total contribution from $Z\gamma \rightarrow \nu\bar{\nu}\gamma$ is estimated in data using a sample of $Z\gamma \rightarrow \ell\bar{\ell}\gamma$ candidates, where the leptons from the decay of the Z boson are removed and considered as E_T^{miss} ?. This provides an estimate of 64.6 ± 17.6 , where the uncertainty is dominated by the size of the sample.

6.3.3 Elecron Mis-ID

The most important SM background comes from events where electrons are misidentified as photons, mainly in the $W \rightarrow e\nu$ process. Seeding efficiency in the pixel detector for electron tracks is $\epsilon = 0.982 \pm 0.004$ for electrons with $p_T > 100\text{GeV}$. This efficiency is measured in data using the tag-and-probe method ? on $Z \rightarrow e\bar{e}$ events, and is verified with MC simulation. Electrons from W boson decay that are not seeded appear as isolated photons accompanied with large E_T^{miss} from the escaping neutrino. This class of events is modeled by an electron proxy event sample selected in data using criteria that are identical to those described in Sec. 6.2, except the photon candidate is required to have a pixel seed. The number of electron proxy events is then scaled by $(1 - \epsilon)/\epsilon$ to yield an estimated contribution of 7.4 ± 1.2 from electron misidentification events. The dominant uncertainty in the estimate is the statistical uncertainty in the tag-and-probe fit, and is assessed by generating a large ensemble of toy dielectron mass distributions on which the fit procedure is repeated.

The standard deviation of the number of $Z \rightarrow e\bar{e}$ events obtained from the fits is then propagated to the uncertainty in ϵ .

6.3.4 Non-collision Backgrounds

Non-collision backgrounds, from things such as detector noise, cosmic rays, and beam halo, are estimated from the time distribution of the cluster seeds since each process exhibits a distinctive time distribution when the cluster is in the ECAL barrel. Templates for anomalous signals, cosmic ray muons, and beam halo events are obtained by inverting the shower shape and beam halo tag requirements, and are fitted to the timing distribution of the candidate sample. The only nonnegligible residual contribution to the candidate sample is found to arise from the beam halo, with an estimated 5.9 ± 4.7 events, where the uncertainty is from the template fit.

6.3.5 Minor SM Processes

The SM processes $W \rightarrow \ell\nu\gamma$, $Z \rightarrow \ell\bar{\ell}\gamma$, $W(l\nu)$ and $\gamma + jets$ are generated with MADGRAPH5_ *aMC@NLO* at LO ? with up to 2 jets and then processed with PYTHIA 6.426 generator ? for showering and hadronization, with the NNPDF3.0 LO($\alpha_s = 0.130$) parton distribution function. The total background expectation from these processes is 3.05 ± 0.67 events, where the uncertainty includes the statistical and systematic uncertainty due to scale factor and jet/ E_T^{miss}/γ energy scale.

6.4 Results

After applying our full selection criteria, we observe 77 events in 2.32 fb^{-1} of data. Table 6.1 shows the estimated number of events and uncertainty from each background for the full 2015 run.

The p_T spectrum and particle flow E_T^{miss} of the full combination of our selected candidate events and estimated backgrounds can be seen in Figure 6.1 and the p_T/E_T^{miss} and number of jets distribution is shown in Figure 6.2.

Table 6.1: Summary of estimated backgrounds and observed total number of candidates for 2.32 fb^{-1} of 2015 data. The category Others includes $W \rightarrow \mu\nu$, $Z \rightarrow \ell\bar{\ell}\gamma$ and $t\bar{t}\gamma$

Process	Estimate
$Z\gamma \rightarrow \nu\bar{\nu}\gamma$	42.10 ± 6.31
$W\gamma \rightarrow \ell\nu\gamma$	10.69 ± 1.49
$W \rightarrow e\nu$	7.80 ± 1.78
$\text{jet} \rightarrow \gamma \text{ fakes}$	3.36 ± 1.13
Beam halo	5.9 ± 4.7
Others	3.05 ± 0.67
Total Expectation	72.9 ± 8.30
Data	77

6.4.1 $pp \rightarrow Z\gamma \rightarrow \nu\bar{\nu}\gamma$ Cross Section Measurement

The $pp \rightarrow Z\gamma \rightarrow \nu\bar{\nu}\gamma$ cross section for $p_T^\gamma > 175 \text{ GeV}$ in the rapidity range $|\eta| < 1.4$ is calculated using the formula

$$\sigma(pp \rightarrow Z\gamma \rightarrow \nu\bar{\nu}\gamma) = \frac{N_{\text{data}} - N_{BG}}{A \times \epsilon \times L} \quad (6.1)$$

where N_{data} is the observed number of events, N_{BG} is the number of estimated background events, A is the geometrical and kinematic acceptance of the selection criteria, ϵ is the selection efficiency within the acceptance, and L is the integrated luminosity. The product of $A \times \epsilon_{MC}$ is estimated from LO MADGRAPH simulation and a correction factor, ρ , described in Section 6.3.1 is applied to account for the difference between the efficiency in the data and Monte Carlo:

$$A \times \epsilon = A \times \epsilon_{MC} \times \rho. \quad (6.2)$$

The product of $A \times \epsilon_{MC}$ is estimated to be 0.314 ± 0.002 (stat) ± 0.048 (syst) and rho is 0.99 ± 0.06 .

The photon energy scale, jet and E_T^{miss} energy scale and resolution, and pileup related contributions are considered as sources of systematic uncertainty in the acceptance calculation. The uncertainty on the photon energy scale is about 1.5% and the uncertainty from variations in the E_T^{miss} energy scales is 5%. Contributions from the jet energy scale are accounted for in the uncertainty on the E_T^{miss} . The uncertainty on the integrated luminosity is 2.7% ?. A summary of the systematic uncertainties are shown in Table 6.2.

Sources	$Z\gamma \rightarrow \nu\bar{\nu}\gamma$ [%]	$W\gamma \rightarrow \ell\nu\gamma$ [%]	Jets faking photon [%]	Electron faking photon	gamma-jet	Other bkgs [%]
Luminosity	2.7	2.7	-	-	2.7	2.7
PDF and Scale	5.37	8.9	-	-	-	-
EWK corrections	11	7	-	-	-	-
Jets faking photon	-	-	30	-	-	-
Electron faking photon	-	-	-	20	-	-
Jet, MET, photon energy scale	6	6	-	-	6	6
Scale Factors	6	6	-	-	6	6

Table 6.2: Summary of systematic unceratinties for signal and different background sources.

The measured cross section for $pp \rightarrow Z\gamma \rightarrow \nu\bar{\nu}\gamma$ for photon $p_T > 175$ GeV within

rapidity range $|\eta_\gamma| < 1.4$ is

$$64.06 \pm 12.14(\text{stat.}) \pm 12.88(\text{sys.}) \pm 1.72(\text{lumi.})\text{fb}. \quad (6.3)$$

The NNLO theoretical cross section is 65.55 ± 0.02 fb where the uncertainty includes only the scale variations. The measured cross section agrees well with the NNLO theoretical cross section and this agreement with the SM prediction constrains possible DM models.

6.4.2 Limits on Dark Matter

Interpreting these results as setting limits on the cross section of a DM particle as a function of DM mass, Tables 6.3 and 6.4 show 90% confidence level (CL) upper limits on the production cross sections provided for the Vector and the Axial-vector model for a mediator mass of 10 TeV.

Mass DM [GeV]	σ [fb]
1	3.821(3.242)
10	3.820(3.244)
50	3.827(3.249)
150	3.826(3.254)
500	3.588(3.052)
1000	3.370(2.862)

Table 6.3: Observed (expected) 90%CL upper limits on the dark-matter production cross section σ for mediator mass 10 TeV.

Mass DM [GeV]	σ [pb]
1	3.782(3.211)
10	3.785(3.213)
50	3.793(3.213)
150	3.754(3.192)
500	3.488(2.961)
1000	3.30(2.814)

Table 6.4: Observed (expected) 90%CL upper limits on the dark-matter production cross section σ for fixed mediator mass 10 TeV

Figure 6.1: The photon p_T and E_T^{miss} distribution for the candidate sample, compared with estimated contributions from SM backgrounds, here QCD γ refers to γ +jet background and the background uncertainty includes statistical and systematic error.

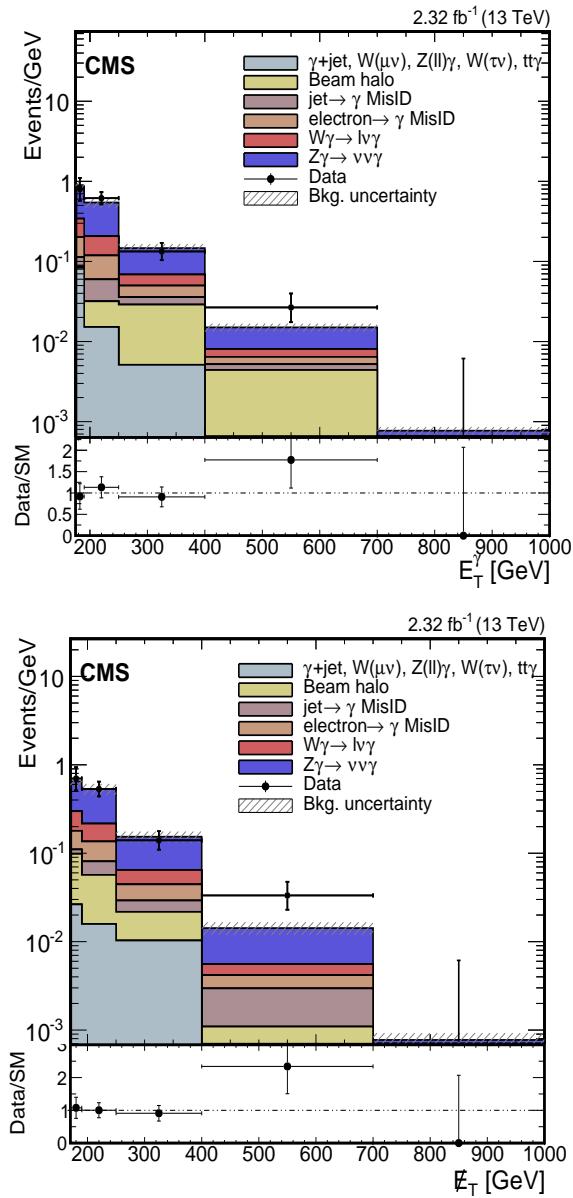


Figure 6.2: The photon p_T/E_T^{miss} and number of jets distribution for the candidate sample, compared with estimated contributions from SM backgrounds. The background uncertainty shown includes statistical and systematic errors.

