Search for the Pair Production of First Generation Scalar Leptoquarks with the CMS Detector

Edmund A. Berry

A DISSERTATION PRESENTED TO THE
FACULTY OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF PHYSICS

Advisor: Christopher G. Tully

April 2014

ABSTRACT

A search for the pair-production of first generation scalar leptoquarks is performed in final states consisting of two electrons and at least two jets or an electron, a neutrino, and at least two jets using proton-proton collision data at $\sqrt{s} = 7$ TeV. The data were collected by the Compact Muon Solenoid detector at the Large Hadron Collider in 2011 and correspond to an integrated luminosity of 4.95 fb^{-1}. The number of observed events in both final states is observed to be in good agreement with the predictions for Standard Model processes. A 95% confidence level combined lower limit is set on the mass of a first generation scalar leptoquark at 830 and 640 GeV for $\beta = 1$ and 0.5, respectively, where β is the branching fraction of the leptoquark to an electron and a quark.

The work described in this dissertation
has been published in the following combined paper:

Phys. Rev. D **86**, 052013 (2012)

CONTENTS

ACKNOWLEDGMENTS

The work described in this thesis is a product of the help and support that I have received from many people over many years.

First among these is my advisor, Christopher Tully, whom I would like to thank for taking me onto his team. Chris is one of the most knowledgable and talented physicists I have ever met. His support and insight have been invaluable.

I would also like to thank the members of the lepton-jet exotica working group at CMS, who reviewed this analysis and helped move it towards publication. The conveners have been extremely helpful and informative, and my fellow analysts have been the best colleagues I could have asked for. This is especially true of Paolo Rumerio and Francesco Santanastasio, who have been like second advisors to me. Special thanks is also due to Jim Hirschauer, whose careful review and watchful eye were essential for getting this analysis published. All three are teachers, leaders, and scientists of the highest caliber.

This thesis and many other analyses at CMS would have been unthinkable without a hadronic calorimeter and a team of experts to keep it running. CMS is a successful experiment thanks to the hard work of people like Pawel de Barbaro, Phil Dudero, Yossof Eshaq, Richard Kellogg, Ted Laird, Phil Lawson, Dragoslav Lazic, Jeremy Mans, German Martinez, Sudan Paramesvaran, Halil Saka, Jared Sturdy, Taylan Yetkin, and too many others to name. I would like to thank them for their work and for teaching me so much about the detector.

My colleagues in the Princeton CMS office shared a long and difficult road with me, and it is impossible to imagine life at CERN without them. It has been a privilege to work alongside Davide Gerbaudo, Phil Hebda, Adam Hunt, Paul Lujan, Tatiana Medvedeva, Michael Mooney, Xiaohang Quan, and Andrzej Zuranski.

I have been blessed to have studied under too many excellent teachers to list here. David Dannels showed me how to write my first line of code and how to write my first physics lab report. His presence in the classroom is sorely missed. Mark Vondracek was the first high energy physicist I ever met, and the day he showed me a picture of the CDF detector changed my life. Ivan Furic and Young-Kee Kim were kind enough to welcome me into the HEP group at the University of Chicago and to guide me to my first published result. My sincere thanks go to all of them, and the debt I owe them can never be fully repaid.

My family has been a source of unwavering support, and no words in these acknowledgements can adequately thank them. My mother and father, Carrie Jeffers and John Berry,

are both models of love and compassion. They are my role models in everything I do. Tom Berry and Jim Berry, are smart, driven, and kind in addition to being the best brothers I could imagine. Our time together is too rare and always too short.

Finally, I want to thank my wife and best friend, Colette Berry née Borden. I came to Europe to search for new physics, and instead I found someone who means more to me than any kind of physics ever could. She has worked tirelessly to forge her own career while simultaneously supporting me through mine, and this thesis would not have been possible without her. Thank you for everything.

LIST OF ABBREVIATIONS

ep	electron-proton
$p\overline{p}$	proton-antiproton
pp	proton-proton
ALICE	A large ion collider experiment
APD	Avalanche photodiode
ASIC	Application-specific integrated circuit
ATLAS	A toroidal LHC apparatus
BPix	Barrel pixel detector
BPTX	Beam pickup timing for the experiments
BRW	Buchmüller, Rückl, Wyler model of leptoquarks
CERN	European organization for nuclear research
CMS	Compact muon solenoid
CMSSW	CMS software
CSC	Cathode strip chambers
CTF	Combinatorial track finder
DAQ	Data acquisition framework
DT	Drift tube chambers
EB	Electromagnetic calorimeter barrel

ECAL	Electromagnetic calorimeter
EE	Electromagnetic calorimeter endcap
ES	Electromagnetic calorimeter preshower
FCNC	Flavor-changing neutral current
FPGA	Field programmable gate array
FPix	Forward pixel detector
FSR	Final state radiation
GCT	Global calorimeter trigger
GCT	Global muon trigger
GSF	Gaussian sum filter
GT	Global trigger
GUT	Grand unified theory
HB	Hadronic calorimeter barrel
HCAL	Hadronic calorimeter
HE	Hadronic calorimeter endcap
HEEP	High energy electron pair identification
HF	Hadronic calorimeter forward
HLT	High level trigger
HO	Hadronic calorimeter outer
HPD	Hybrid photodiode
IP	Interaction point
ISR	Initial state radiation
L1	Level 1 trigger
L1A	Level 1 accept
LEP	Large electron-positron collider
LHC	Large hadron collider
LHCb	Large hadron collider beauty

LO	Leading order
LQ	Leptoquarks
LUT	Look up table
mBRW	minimal Buchmüller-Rückl-Wyler model of leptoquarks
MC	Monte Carlo
MET	Missing transverse energy
MIP	Minimum ionizing particle
NLO	Next to leading order
NNLL	Next-to-next-to leading logarithm
PCC	Pixel cluster counting
PD	Primary dataset
PDF	Parton density function
PF	Particle flow
PMT	Photomultiplier tubes
PS	Proton synchrotron
PSB	Proton synchrotron booster
QCD	Quantum chromodynamics
QED	Quantum electrodynamics
RBX	Readout box
RF	Radio frequency
RPC	Resistive plate chambers
SM	Standard Model
SPS	Super proton synchrotron
TCS	Trigger control system
TEC	Tracker endcap
TIB	Tracker inner barrel
TIB	Tracker inner disk

TOB Tracker outer barrel

TPG Trigger primitive generator

TTC Timing, trigger, and control system

VPT Vacuum phototriode

WLS Wavelength shifting fibers

FOREWORD

The Standard Model has been remarkably successful in predicting the interactions of fundamental particles for over 70 years. Despite this success, the Standard Model leaves several important questions unanswered. In particular, the Standard Model does not explain the presence of three generations of both leptons and quarks, nor does it explain the similarity of the arrangements of leptons and quarks under the electroweak interaction.

One possible motivation for these features stems from the idea that the Standard Model is a low energy remnant of a larger, more elegant theory, which includes leptoquarks: heavy bosons that carry both lepton and baryon number and mediate a fundamental interaction between leptons and quarks. Leptoquarks have been a topic of interest in the particle physics community for some time. This interest was perhaps never more intense than in 1997, when the ZEUS and H1 collaborations at the HERA electron-proton (ep) collider observed an excess of $e+p \to e+p+X$ events corresponding to an electron-quark invariant mass of about 200 GeV [1, 2]. This excess sparked a flurry of both theoretical and experimental investigation into leptoquarks. While leptoquarks of this description would have been produced in large numbers at the Tevatron proton-antiproton ($p\overline{p}$) collider, results from the CDF and D0 experiments definitively ruled out any possibility for leptoquarks at a mass scale of 200 GeV in 1998 [3]. Even so, leptoquarks remain an active area of theoretical and experimental research and are an important benchmark in the search for new physics beyond the Standard

Model.

This thesis describes a search for the pair production of scalar leptoquarks decaying to fermions of the first generation. This search is performed using $\sqrt{s} = 7$ TeV proton-proton (pp) collisions at the Large Hadron Collider as measured by the Compact Muon Solenoid detector. The data used were collected in 2011 and correspond to an integrated luminosity of 4.95 fb^{-1}. Chapter 1 gives a brief overview of the Standard Model and an effective model of leptoquarks. Chapters 2 and 3 describe the experimental apparatus used to conduct the search: the Large Hadron Collider and the Compact Muon Solenoid detector. Chapters 4-10 describe the search itself. Chapter 11 presents a conclusion and lists future prospects for the search.

CHAPTER 1

THEORETICAL BACKGROUND

1.1 The Standard Model

The Standard Model (SM) is a theoretical framework in which the known elementary particles are expressed as excitations of relativistic quantum fields. Particle interactions are determined by the Standard Model Lagrangian, which is a function of these fields. All of these fields are bosonic (having integer spin) or fermionic (having half-integer spin), and the corresponding particles are bosons and fermions. This section briefly summarizes the Standard Model. A more detailed explanation may be found elsewhere [4–6].

1.1.1 Matter and forces

The Standard Model expresses particles of matter as spin-1/2 fermions known as "leptons" and "quarks." These particles and their properties are listed in Table 1.1. Leptons include the massive and electrically charged electron (e), muon (μ), and tau (τ) and their corresponding massless and electrically neutral neutrinos: ν_e, ν_μ, and ν_τ. The extension of the Standard Model to include massive neutrinos after the observation of neutrino flavor oscillations is discussed elsewhere [7]. Leptons are arranged into three "generations." Each generation forms

Fermion type	Generation 1st	2nd	3rd	T	T_3	Y	Q
Leptons	$\begin{pmatrix}\nu_e\\ e\end{pmatrix}_L$	$\begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix}_L$	$\begin{pmatrix}\nu_\tau\\ \tau\end{pmatrix}_L$	$\frac{1}{2}$	$\begin{pmatrix}\frac{1}{2}\\ -\frac{1}{2}\end{pmatrix}$	-1	$\begin{pmatrix}0\\ -1\end{pmatrix}$
	e_R	μ_R	τ_R	0	0	-2	-1
Quarks	$\begin{pmatrix}u\\ d\end{pmatrix}_L$	$\begin{pmatrix}c\\ s\end{pmatrix}_L$	$\begin{pmatrix}t\\ b\end{pmatrix}_L$	$\frac{1}{2}$	$\begin{pmatrix}\frac{1}{2}\\ -\frac{1}{2}\end{pmatrix}$	$\frac{1}{3}$	$\begin{pmatrix}\frac{2}{3}\\ -\frac{1}{3}\end{pmatrix}$
	u_R	c_R	t_R	0	0	$\frac{4}{3}$	$\frac{2}{3}$
	d_R	s_R	b_R	0	0	$-\frac{2}{3}$	$-\frac{1}{3}$

Table 1.1: The three generations of fermions and their quantum numbers. T and T_3 represent total weak isospin and the third component of weak isospin, respectively. Y represents hypercharge. Q represents electric charge.

a doublet of left-handed states with non-zero weak isospin $\binom{\nu_\ell}{\ell}_L$ and a singlet of right-handed states ℓ_R with zero weak isospin. In this context, right-handed fermions have parallel spin and momentum, while left-handed fermions have anti-parallel spin and momentum. Quarks include the up (u), down (d), charm (c), strange (s), bottom (b), and top (t). Like leptons, quarks are arranged into three generations. In the case of quarks, each generation forms a doublet of left-handed states with non-zero weak isospin and two singlets of right-handed states with zero weak isospin. Unlike leptons, quarks carry an additonal form of charge known as "color", which has three values: red, green, and blue.

Leptons and quarks interact with each other via the exchange of spin-1 bosons. These bosons include the gluon (g), the W and Z bosons, and the photon (γ). They are responsible for the three forces described by the Standard Model: the strong force, the weak force, and electromagnetism, respectively. A fourth force, gravity, is not described by the Standard Model, but its effects are negligable when compared to the effects of the other three forces on the sub-atomic particle mass scale. The interactions between these spin-1 bosons and the spin-1/2 fermions (along with boson-boson interactions) are shown in the schematic in Figure 1.1. The three forces described by the Standard Model are discussed in greater detail below.

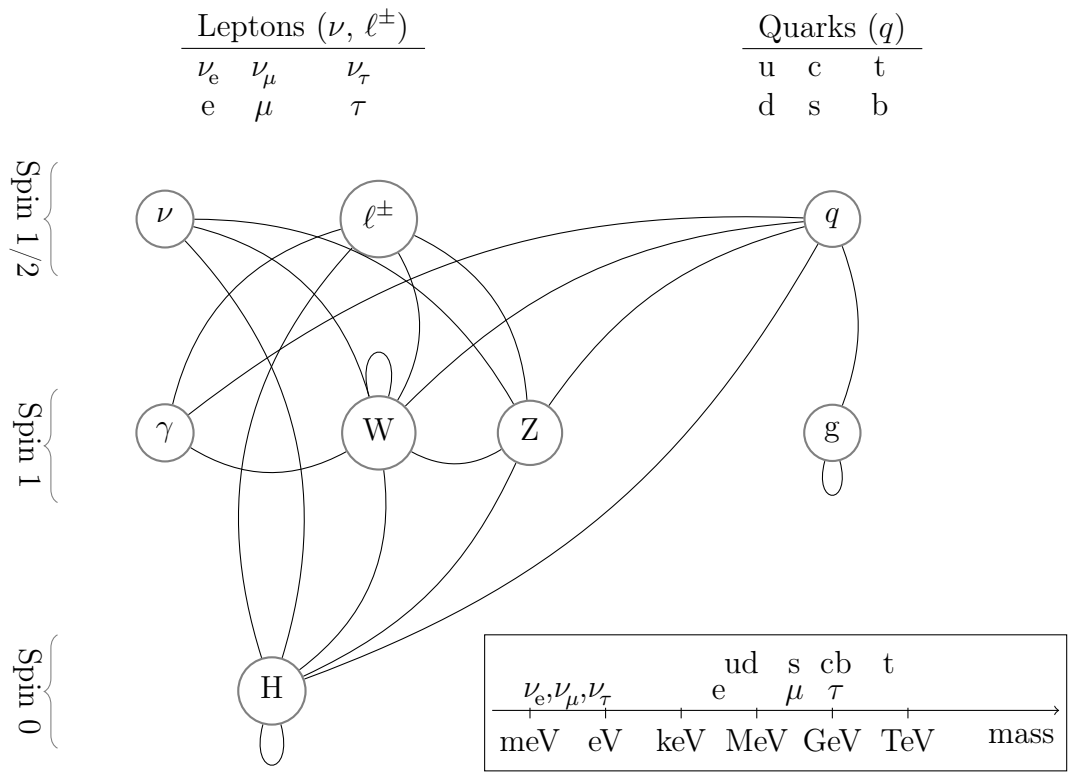


Figure 1.1: Schematic view of the Standard Model particles and of their interactions. Connecting lines represent interactions between particles. The masses of the fundamental constituents are indicated in the bottom-right inset.

Electromagnetism influences particles that carry electric charge. This influence is manifested via the exchange of a photon, and the theory describing it is known as Quantum Electrodynamics (QED). Electromagnetism was the first of the forces in the Standard Model to be discovered, and the strong and weak forces are named for their coupling strengths relative to electromagnetism. Because the photon is massless, the range of electromagnetism is infinite.

The weak force influences all fermions via the exchange of W and Z bosons. The strength of the weak force is roughly 10^{-5} that of electromagnetism, and it has several distinctive properties. First, the weak force behaves differently for fermions with different chiralities. Namely, the charged W boson only interacts with fermions that have left-handed chirality (i.e. leptons and quarks having non-zero weak isospin) and their right-handed anti-particles.

The neutral Z boson interacts with fermions and anti-fermions of both right and left-handed chiralities. Second, because the W and Z bosons are massive ($M \sim 100$ GeV), the weak force has a short range ($R \sim 10^{-18}$ m). The mass associated with these bosons is explained by the Higgs mechanism. Finally, the weak force is the only force with the ability to change the flavor of quarks.

The strong force influences particles that carry a color charge. This influence is manifested via the exchange of gluons: electrically neutral and massless spin-1 bosons that carry a color charge themselves. Only quarks and gluons themselves are known to carry color charge, so the strong force influences only these particles. The theory describing the interaction between quarks and gluons is known as Quantum Chromodynamics (QCD). The strength of the strong force is roughly 100 times that of electromagnetism, and it is the only force that does not decrease in strength as the distance between the two interacting particles increases. As a result, it is energetically favorable for a quark-antiquark pair to be created from the vacuum rather than for isolated quarks to interact over long distances. This means that isolated quarks are rarely observed. Instead, quarks are nearly always observed as a constituent of other non-fundamental, color charge-neutral particles called "hadrons." [1] A similar effect is observed in the case of gluons, which also hadronize and are never observed in isolation. This property is called "confinement." It should be noted that the residual strong force interaction between color charge-neutral hadrons is not zero, and it has a short range ($R \sim 10^{-15}$ m).

In the context of collider physics, hadronizing quarks and gluons produced in particle collisions result in a collimated spray of hadrons, which are called "jets". The reconstruction of the original quark and gluon energy and momentum from jets is a critical challenge of collider physics, and it is discussed in further detail in Section 3.6.

[1] The exception is the top quark, which has a mean lifetime 20 times shorter than the timescale for strong force interactions and does not hadronize.

1.1.2 Gauge invariance and the Higgs mechanism

In general, a Lagrangian is said to be gauge invariant (or to have a gauge symmetry) if it remains invariant under a continuous gauge transformation. A central postulate of the Standard Model Lagrangian in particular is that the dynamics of all particles are determined by specifying underlying local gauge symmetries. This implies that the forces described in the Standard Model are associated with symmetry groups. The strong force is associated with the $\mathrm{SU}(3)_C$ symmetry group, where C corresponds to color. Gauge symmetries allow for the electromagnetic and weak forces to be unified into a single "electroweak" interaction associated with the direct product of the $\mathrm{SU}(2)_L$ and $\mathrm{U}(1)_Y$ symmetry groups [8–10], where L corresponds to left-handed chirality and Y corresponds to weak hypercharge. The combined symmmetry group of the Standard Model Lagrangian is therefore $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$.

The concept of local gauge invariance allows for the unification of the electromagnetic and weak forces and predicts the existence of the W and Z bosons, which had not been observed when the theory was originally described. However, local gauge invariance also requires all fermions and gauge bosons to be massless, which is clearly at odds with various experimental results. This shortcoming was addressed independently by multiple groups in 1964 via what is now called the "Higgs mechanism" [11–13]. The Higgs mechanism introduces a fundamental, self-interacting, spin-0 field that permeates all of space. This field, the Higgs field, interacts with the W and Z bosons associated with the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetry group and all fermions. The quantum of the Higgs field is the Higgs boson, H. The Higgs field is given a non-zero vacuum expectation value, v, the value of which is determined from experiment. The non-zero vacuum expectation value results in a spontaneous breaking of the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetry and imparts mass to the W and Z bosons and the charged fermions (neutrinos, the photon, and the gluon remain massless). The Higgs field, the Higgs boson, and the Higgs mechanism are therefore all essential components of the Standard Model.

1.1.3 Limitations of the Standard Model

The success of the Standard Model has been remarkable. All of the particles described in the previous sections have been observed. Their properties and interactions behave in accordance with Standard Model predictions down to a length scale of approximately 10^{-18} m. In spite of this success, however, the Standard Model is incomplete.

The Standard Model's failure to address gravity as a fundamental force means that it is unable to describe interactions at the Planck scale ($M_{\text{Planck}} \approx 1.22 \times 10^{19}$), at which point the quantum effects of gravity become too large to ignore. The Standard Model also fails to explain why the gap between the Planck scale and the electroweak scale is so large: the weak force is roughly 10^{32} times stronger than gravity (an issue known as the "hierarchy problem"). Other unanswered questions trouble the Standard Model at lower energy scales than the Planck scale. Experimental results from astrophysics suggest that roughly one quarter of the energy density of the universe is composed of "dark matter" while nearly three quarters is composed of "dark energy". However, the Standard Model makes no predictions as to the composition of dark matter, nor does it contain any dark matter particle candidates (i.e. massive, neutral, stable, fundamental particles). In addition, the precise nature of neutrino oscillation and neutrino masses is unknown. Most importantly for this thesis, the Standard Model does not explain the presence of three generations of fermions (see Table 1.1), nor does it explain the similarity of the arrangements of quarks and leptons under the electroweak interaction.

The presence of these unanswered questions in spite of undeniable experimental success leads many to suspect that the Standard Model is a low energy remnant of a larger theory. The nature of this larger theory and various "extensions" to the Standard Model are the subject of intense speculation in the particle physics community. Leptoquarks, the subject of this thesis, form the basis of one such extension of the Standard Model. They are discussed in the following section.

1.2 Leptoquarks

In the Standard Model, leptons and quarks are formally independent from each other. However, both leptons and quarks come in three generations, and they are similarly arranged in multiplets under the electroweak interaction (see Table 1.1). This arrangement is necessary for the cancellation of triangle anomalies, which in turn is necessary for the Standard Model to be a consistent quantum field theory [14].

A possible motivation for this similarity comes from considering a hypothetical interaction between leptons and quarks. This interaction would include one or more bosonic fields interacting between leptons and quarks. Such fields, known as leptoquarks (LQ), would form as color triplets under $\mathrm{SU}(3)_C$, carry both lepton number (L) and baryon number (B), and carry fractional charge (Q). Other properties, including spin, weak isospin, specific electric charge, and fermion number ($F = 3B + L$) are model-dependent.

Leptoquarks are motivated by many theories beyond the Standard Model. Some examples include Grand Unified Theories (GUTs) based on gauge groups including Pati-Salam SU(4) color symmetry [15, 16], SU(5) [17, 18], SO(10) [19, 20], and SU(15) [21, 22]. Other examples include superstring-inspired E_6 models [23], extended technicolor models [24–26], composite models [27, 28], horizontal symmetry theories [29], strongly coupled weak-interaction models [30], and R-parity violating supersymmetric models [31–33].

In practice, searches for leptoquarks at colliders are performed using general effective models which avoid these inconsistencies between various models. The most popular of these effective models is the Buchmüller-Rückl-Wyler effective leptoquark model, which is described in the following section.

1.2.1 Buchmüller-Rückl-Wyler effective leptoquark model

The general effective model used in this thesis was proposed in 1987 by Buchmüller, Rückl, and Wyler [34] (BRW model). The four underlying assumptions of the BRW model are: **1)**

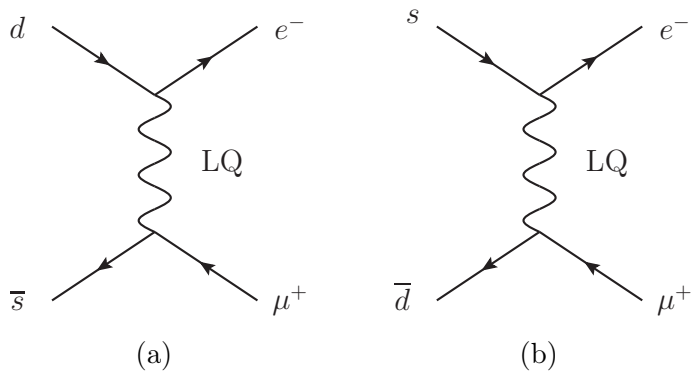


Figure 1.2: Examples of a tree-level FCNC and lepton flavor violating decay $K_L^0(d\overline{s} + s\overline{d}) \to e^-\mu^+$, mediated by a vector leptoquark. This decay mode is permitted by some leptoquark models [15, 16] but forbidden by the Standard Model.

leptoquarks have renormalizable interactions (their couplings to Standard Model lepton-quark pairs are dimensionless), **2)** leptoquark interactions are invariant under the Standard Model gauge groups: $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$, **3)** leptoquarks couple *only* to Standard Model fermions and gauge bosons, and **4)** leptoquarks conserve lepton number (L) and baryon number (B) separately in order to protect against rapid proton decay.

Under these four initial assumptions underlying the BRW model, leptoquarks are permitted to decay to any combination of leptons and quarks. This can lead to tree-level flavor-changing neutral current (FCNC) decays. An example of such a tree-level FCNC decay is shown in Figure 1.2. To protect against these decays, the BRW model may be augmented with the additional requirement **5)** that leptoquarks may only couple to fermions of a single generation[2]. This thesis is focused on first generation leptoquarks (leptoquarks coupling to electrons, electron neutrinos, up quarks, and down quarks) specifically. Other low energy constraints, such as chirally suppressed meson decays (for example, $\pi \to e\nu$ as shown in Figure 1.3) make it necessary to introduce a requirement **6)** that leptoquarks have purely chiral couplings. The addition of requirements **5** and **6** to the BRW model results in the minimal Buchmüller-Rückl-Wyler effective model (mBRW).

[2]While the requirement that leptoquarks only couple to a single fermion generation is traditional, requiring that leptoquarks do not couple to more than one generation of leptons and more than one generation of quarks is enough to protect against tree-level FCNC decays like the one shown in Figure 1.2. For example, a leptoquark could be allowed to couple to first generation leptons and third generation quarks.

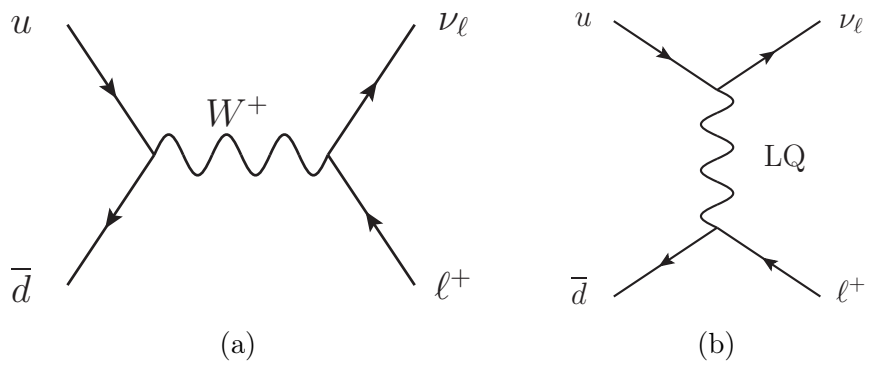


Figure 1.3: Example of a pion decay, $\pi^+ \to \ell^+\nu$, mediated by a Standard Model W^+ boson (a) and a vector leptoquark (b) [15, 16]. The decay $\pi^+ \to e^+\nu$ is chirally suppressed by the Standard Model but not by some leptoquark models.

These requirements lead to seven scalar and seven vector leptoquarks. These leptoquarks may be classified by fermionic number, $F = 3B + L$, such that $|F| = 0$ or 2. The interactions between mBRW leptoquarks and Standard Model fermions are described by the following Lagrangian:

$$\mathcal{L} = \mathcal{L}_{|F|=2} + \mathcal{L}_{F=0} \tag{1.1}$$

The terms $\mathcal{L}_{|F|=2}$ and $\mathcal{L}_{F=0}$ given in Equation 1.1 may be defined as follows:

$$\begin{aligned}
\mathcal{L}_{|F|=2} = &(g_{1L}\overline{q}^c_L i\tau_2 \ell_L + g_{1R}\overline{u}^c_R e_R)S_0 \\
&+ (\tilde{g}_{1R}\overline{d}^c_R e_R)\tilde{S}_0 \\
&+ (g_{3L}\overline{q}^c_L i\tau_2 \vec{\tau} \ell_L)\vec{S}_1 \\
&+ (g_{2L}\overline{d}^c_R \gamma^\mu \ell_L + g_{2R}\overline{q}^c_L \gamma^\mu e_R)V_{\frac{1}{2}\mu} \\
&+ (\tilde{g}_{2L}\overline{u}^c_R \gamma^\mu \ell_L)\tilde{V}_{\frac{1}{2}\mu} \\
&+ \text{h.c.}
\end{aligned} \tag{1.2}$$

$$
\begin{aligned}
\mathcal{L}_{F=0} = &(h_{2L}\overline{u}_R\ell_L + h_{2R}\overline{q}_L i\tau_2 e_R)S_{\frac{1}{2}} \\
&+ (\tilde{h}_{2L}\overline{d}_R\ell_L)\tilde{S}_{\frac{1}{2}} \\
&+ (h_{1L}\overline{q}_L\gamma^\mu\ell_L + h_{1R}\overline{d}_R\gamma^\mu e_R)V_{0\mu} \\
&+ (\tilde{h}_{1R}\overline{u}_R\gamma^\mu e_R)\tilde{V}_{0\mu} \\
&+ (h_{3L}\overline{q}_L\vec{\tau}\gamma^\mu\ell_L)V_{1\mu} \\
&+ \text{h.c.}
\end{aligned}
\tag{1.3}
$$

Equations 1.2 and 1.3 are given in the *Aachen* notation [35]. Scalar leptoquarks are denoted with an S and vector leptoquarks with a V. The subscripts on S and V denote the $\mathrm{SU}(2)_L$ isospin (0, 1/2, 1). The $\sim$ symbol over $\tilde{S}$ and $\tilde{V}$ is purely a label for distinguishing between two distinct leptoquarks with identical spin and weak isospin but different hypercharge (for example S_0 and $\tilde{S}_0$). q_L and ℓ_L represent the $SU(2)_L$ left-handed doublets for quarks and leptons. e_R, u_R, and d_R represent the right-handed lepton, up-type quark, and down-type quark singlets, respectively. τ_i represent the Pauli matrices. γ^μ represent the Dirac matrices. g, h, $\tilde{g}$, and $\tilde{h}$ are coupling constants: the subscripts denote the dimension of the $\mathrm{SU}(2)_L$ representation (1, 2, 3) of the coupled leptoquark and the chirality of the coupled lepton (L, R). It is traditional to write these couplings as $\lambda_R = g_R, h_R$ and $\lambda_L = g_L, h_L$. The ψ^c operators are the charge conjugate of the fermion fields, such that $\psi^c \equiv C\overline{\psi}^T$. For simplicity, color, weak isospin, and generation (flavor) indices are omitted.[3]

Some of the properties of the leptoquarks given by this Lagrangian are listed in Table 1.2, including the leptoquarks' total weak isospin, third component of weak isospin, hypercharge, electric charge, fermionic number, coupling constants, and decay modes. The leptoquark symbols (S and V) in Table 1.2 have an additional subscript beyond those used in the Lagrangian. This subscript corresponds to the chirality (L, R) of the lepton to which the

[3] In the notation used by the original BRW paper [34], leptoquarks with different fermionic number or spin are given distinct symbols (R, S, U, V). The subscript of those symbols corresponds to the dimension of the leptoquark's $\mathrm{SU}(2)_L$ representation (1, 2, 3). In the *Aachen* notation [35], leptoquarks with different spin are given distinct symbols (S, V). The subscript of those symbols corresponds to the leptoquark's $\mathrm{SU}(2)_L$ isospin (0, 1/2, 1). The two notations have the following correspondence: $S_0 \leftrightarrow S_1^{\mathrm{BRW}}$; $\tilde{S}_0 \leftrightarrow \tilde{S}_1^{\mathrm{BRW}}$; $S_{1/2} \leftrightarrow R_2^{\mathrm{BRW}}$; $\tilde{S}_{1/2} \leftrightarrow \tilde{R}_2^{\mathrm{BRW}}$; $S_1 \leftrightarrow S_3^{\mathrm{BRW}}$; $V_0 \leftrightarrow U_1^{\mathrm{BRW}}$; $\tilde{V}_0 \leftrightarrow \tilde{U}_1^{\mathrm{BRW}}$; $V_{1/2} \leftrightarrow V_2^{\mathrm{BRW}}$; $\tilde{V}_{1/2} \leftrightarrow \tilde{V}_2^{\mathrm{BRW}}$; $V_1 \leftrightarrow U_3^{\mathrm{BRW}}$.

leptoquark is coupled. For most experimental searches (including this one), mass degeneracy is assumed within each isospin family. The theoretical motivation behind this assumption is that one would expect all leptoquarks within a given $SU(2)_L$ representation to be degenerate, ignoring loop corrections. For simplicity, therefore, each leptoquark symbol corresponds to a single isospin family, including all of the electric charge possibilities within that family. For example, the symbol $V_{1,L}$ corresponds to any of the three vector leptoquarks that have total weak isospin $T = 1$, couple to a left handed lepton, and have electric charge $Q = -5/3$, $-2/3$, or $1/3$. By construction, the decay branching ratios $\beta = \mathrm{BR}(\mathrm{LQ} \to \ell^{\pm} q)$ of each of these leptoquarks into a final state containing a charged lepton and a quark are fixed to 0, 1/2, or 1 [36].

In addition to interacting with lepton-quark pairs as described in Table 1.2 and in the Lagrangian laid out in Equations 1.1-1.3, leptoquarks may also interact with the Standard Model gauge bosons. In principle, since this Lagrangian was chosen to preserve the gauge symmetries of the Standard Model, the interactions between leptoquarks and Standard Model gauge bosons are completely determined. This is true for scalar leptoquarks. However, for vector leptoquarks interacting with Standard Model gauge bosons (A), cross sections that depend on trilinear ($A\mathrm{LQ}\overline{\mathrm{LQ}}$) and quartic ($AA\mathrm{LQ}\overline{\mathrm{LQ}}$) couplings may require damping via the introduction of anomalous couplings. This would be necessary, for instance, if the vector leptoquarks were composite low energy manifestations of a more fundamental theory at higher energy scales. These anomalous couplings include four couplings for the electroweak interaction: κ_γ, κ_Z, λ_γ, and λ_Z and two couplings for the strong interaction: κ_g and λ_g [37, 38]. An effective Lagrangian for leptoquark interactions with the γ and Z bosons [14, 39] and with gluons [40] may be found elsewhere.

Typically, models including leptoquarks contain a subset of the leptoquarks described in the BRW model. For example, $V_{0,L(R)}$ appears in the Pati-Salam GUT model [15, 16], anti-$\tilde{S}_{1/2,L}$ appears in a refined SU(5) model [17, 18], and $S_{1,L(R)}$ appears in E_6 models [23]. All fourteen leptoquarks described in the BRW model appear in the GUT theory based on

LQ	squark	T	T_3	Y	Q	F	$\lambda_L(lq)$	$\lambda_R(lq)$	$\lambda_L(\nu q)$	Decay mode
$S_{0,L}$	$\bar{\tilde{d}}_R$	0	0	2/3	1/3	-2	g_{1L}	0	$-g_{1L}$	$e_L^+\overline{u}_L$, $\overline{\nu}_L\overline{d}_L$
$S_{0,R}$	-	0	0	2/3	1/3	-2	0	g_{1R}	0	$e_R^+\overline{u}_R$
$\tilde{S}_{0,R}$	-	0	0	8/3	4/3	-2	0	$\tilde{g}_{1R}$	0	$e_R^+\overline{d}_R$
$S_{\frac{1}{2},L}$	-	1/2	1/2	7/3	5/3	0	h_{2L}	0	0	$e_L^+u_L$
	-		$-1/2$		2/3		0	0	h_{2L}	$\overline{\nu}_L u_L$
$S_{\frac{1}{2},R}$	-	1/2	1/2	7/3	5/3	0	0	h_{2R}	0	$e_R^+u_R$
	-		$-1/2$		2/3		0	$-h_{2R}$	0	$e_R^+d_R$
$\tilde{S}_{\frac{1}{2},L}$	$\tilde{u}_L$	1/2	1/2	1/3	2/3	0	$\tilde{h}_{2L}$	0	0	$e_L^+d_L$
	$\tilde{d}_L$		$-1/2$		$-1/3$		0	0	$\tilde{h}_{2L}$	$\overline{\nu}_L d_L$
$S_{1,L}$	-	1	1	2/3	4/3	-2	$-\sqrt{2}g_{3L}$	0	0	$e_L^+\overline{d}_L$
	-		0		1/3		$-g_{3L}$	0	$-g_{3L}$	$e_L^+\overline{u}_L$, $\overline{\nu}_L\overline{d}_L$
	-		-1		$-2/3$		0	0	$\sqrt{2}g_{3L}$	$\overline{\nu}_L\overline{u}_L$
$V_{0,L}$	-	0	0	4/3	2/3	0	h_{1L}	0	h_{1L}	$e_L^+d_R$, $\overline{\nu}_L u_L$
$V_{0,R}$	-	0	0	4/3	2/3	0	0	h_{1R}	0	$e_R^+d_L$
$\tilde{V}_{0,R}$	-	0	0	10/3	5/3	0	0	$\tilde{h}_{1R}$	0	$e_R^+u_L$
$V_{\frac{1}{2},L}$	-	1/2	1/2	5/3	4/3	-2	g_{2L}	0	0	$e_L^+\overline{d}_R$
	-		$-1/2$		1/3		0	0	g_{2L}	$\overline{\nu}_L\overline{d}_R$
$V_{\frac{1}{2},R}$	-	1/2	1/2	5/3	4/3	-2	0	g_{2R}	0	$e_R^+\overline{d}_L$
	-		$-1/2$		1/3		0	g_{2R}	0	$e_R^+\overline{u}_L$
$\tilde{V}_{\frac{1}{2},L}$	-	1/2	1/2	$-1/3$	1/3	-2	$\tilde{g}_{2L}$	0	0	$e_L^+\overline{u}_R$
	-		$-1/2$		$-2/3$		0	0	$\tilde{g}_{2L}$	$\overline{\nu}_L\overline{u}_R$
$V_{1,L}$	-	1	1	4/3	5/3	0	$\sqrt{2}h_{3L}$	0	0	$e_L^+u_R$
	-		0		2/3		$-h_{3L}$	0	h_{3L}	$e_L^+\overline{d}_R$, $\overline{\nu}_L u_R$
	-		-1		$-1/3$		0	0	$\sqrt{2}h_{3L}$	$\overline{\nu}_L d_R$

Table 1.2: The 14 leptoquarks described by the BRW model and their corresponding RPV squarks (where applicable). In the "squark" column, $\tilde{u}$ corresponds to up-type squarks, and $\tilde{d}$ corresponds to down-type squarks. $\lambda_L(lq)$, $\lambda_R(lq)$, and $\lambda_L(\nu q)$ are the coupling constants between BRW lepqtoquarks, leptons, and quarks. In the "Decay mode" column, e^+ corresponds to positively charged leptons, ν corresponds to neutrinos, u corresponds to up-type quarks, and d corresponds to down-type quarks. Antiparticles are not shown. The particle-antiparticle convention is such that $\overline{\text{LQ}}_{F=2} \to lq$ and $\overline{\text{LQ}}_{F=0} \to l\overline{q}$ [14].

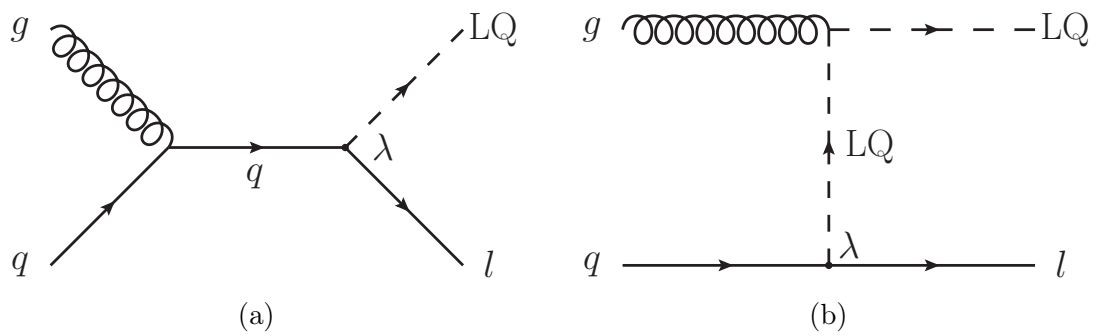


Figure 1.4: Feynman diagrams for the single production of scalar leptoquarks.

the SU(15) gauge group [21, 22].

1.2.2 Leptoquark production in pp collisions

Leptoquarks may be produced in pp collisions via single production and pair production. In the case of leading order (LO) leptoquark single production, the scattering cross sections are proportional to the model-dependent couplings ($\lambda_{L,R}$). For leptoquarks of mass on the order of 1 TeV, these couplings are constrained to be smaller than the electromagnetic coupling ($\lambda_{\text{em}} = \sqrt{4\pi\alpha_{\text{em}}} \sim 0.3$) by low energy processes [41–44]. As a result, search limits on single-production leptoquarks are always given as combined limits on leptoquark mass and on $\lambda_{L,R}$ [45–47]. Figure 1.4 shows two examples of leading order leptoquark single production Feynman diagrams. A generic coupling constant λ is shown on the vertices corresponding to the leptoquark-quark-lepton interactions.

Because leptoquarks are $\text{SU}(3)_C$ color triplets, leptoquark pair production in pp collisions occurs primarily through quark-antiquark annihilation and gluon-gluon fusion [40, 48], with gluon-gluon fusion dominating [49]. These contributions from bosonic couplings ensure that low energy constraints on the fermionic couplings do not constrain scattering cross sections. As a result, searches for the pair production of scalar leptoquarks are given as limits on leptoquark mass only, independent of any coupling [50–53]. In addition, leptoquark pair production cross sections are predicted to be significantly larger than those of leptoquark single production for leptoquarks with mass less than 1.2 TeV [49]. Figure 1.5 shows six

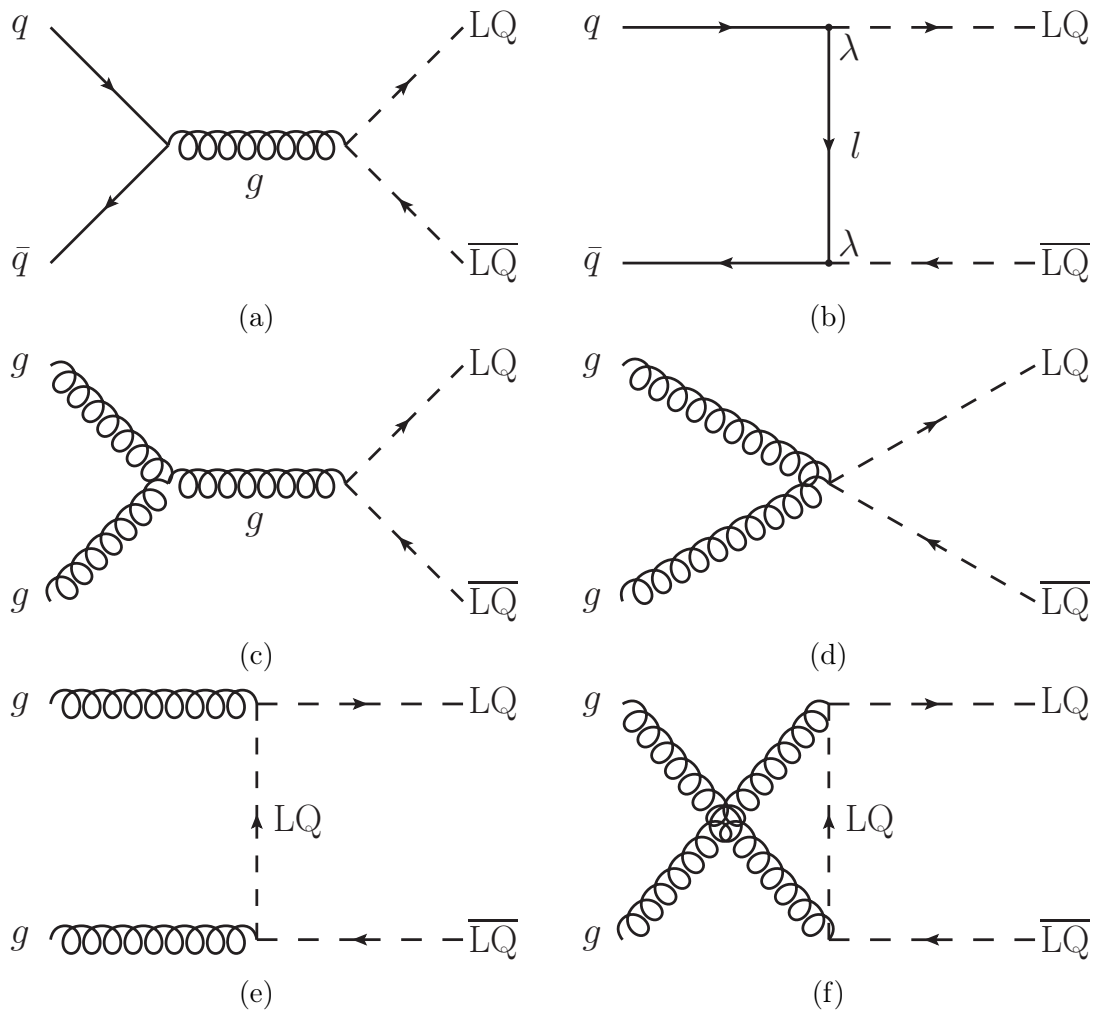


Figure 1.5: Feynman diagrams for the pair production of scalar leptoquarks via quark-antiquark annihilation (a,b) and gluon-gluon fusion (c-e). The t-channel lepton exchange process shown in figure (b) has a cross section proportional to λ^2, which makes it model-dependent.

examples of leading order leptoquark single production Feynman diagrams. It is notable that Figure 1.5b in particular features a t-channel lepton exchange, which implies that its cross section will be proportional to λ^2.

Vector leptoquark pair production cross sections may be greater than those of their scalar counterparts, but the values vary by as much as two orders of magnitude depending on the model and on the anomalous coupling parameters [38]. Scalar leptoquark pair production cross sections are generally model independent. Michael Krămer et al. [54] have written a tool to calculate the leading order and next to leading order (NLO) cross sections for scalar

	$\sqrt{s} = 7$ TeV		$\sqrt{s} = 8$ TeV	
M(LQ) [GeV]	σ(NLO) [pb]	δ_{PDF} [pb]	σ(NLO) [pb]	δ_{PDF} [pb]
200	11.9	0.986	17.4	1.24
250	3.47	0.372	5.26	0.487
300	1.21	0.158	1.89	0.214
350	0.477	$0.728 \cdot 10^{-1}$	0.770	0.102
400	0.205	$0.357 \cdot 10^{-1}$	0.342	$0.520 \cdot 10^{-1}$
450	$0.948 \cdot 10^{-1}$	$0.185 \cdot 10^{-1}$	0.163	$0.278 \cdot 10^{-1}$
500	$0.463 \cdot 10^{-1}$	$0.996 \cdot 10^{-2}$	$0.820 \cdot 10^{-1}$	$0.115 \cdot 10^{-1}$
550	$0.236 \cdot 10^{-1}$	$0.558 \cdot 10^{-2}$	$0.431 \cdot 10^{-1}$	$0.893 \cdot 10^{-2}$
600	$0.124 \cdot 10^{-1}$	$0.321 \cdot 10^{-2}$	$0.235 \cdot 10^{-1}$	$0.530 \cdot 10^{-2}$
650	$0.676 \cdot 10^{-2}$	$0.190 \cdot 10^{-2}$	$0.132 \cdot 10^{-1}$	$0.322 \cdot 10^{-2}$
700	$0.377 \cdot 10^{-2}$	$0.114 \cdot 10^{-2}$	$0.761 \cdot 10^{-2}$	$0.200 \cdot 10^{-2}$
750	$0.214 \cdot 10^{-2}$	$0.700 \cdot 10^{-3}$	$0.448 \cdot 10^{-2}$	$0.126 \cdot 10^{-2}$
800	$0.124 \cdot 10^{-2}$	$0.437 \cdot 10^{-3}$	$0.269 \cdot 10^{-2}$	$0.810 \cdot 10^{-3}$
850	$0.732 \cdot 10^{-3}$	$0.276 \cdot 10^{-3}$	$0.164 \cdot 10^{-2}$	$0.527 \cdot 10^{-3}$
900	$0.436 \cdot 10^{-3}$	$0.176 \cdot 10^{-3}$	$0.101 \cdot 10^{-2}$	$0.347 \cdot 10^{-3}$
950	$0.263 \cdot 10^{-3}$	$0.113 \cdot 10^{-3}$	$0.634 \cdot 10^{-3}$	$0.231 \cdot 10^{-3}$
1000	$0.160 \cdot 10^{-3}$	$0.737 \cdot 10^{-4}$	$0.401 \cdot 10^{-3}$	$0.155 \cdot 10^{-3}$
1050	$0.982 \cdot 10^{-4}$	$0.483 \cdot 10^{-4}$	$0.256 \cdot 10^{-3}$	$0.105 \cdot 10^{-3}$
1100	$0.606 \cdot 10^{-4}$	$0.318 \cdot 10^{-4}$	$0.165 \cdot 10^{-3}$	$0.718 \cdot 10^{-4}$
1150	$0.377 \cdot 10^{-4}$	$0.210 \cdot 10^{-4}$	$0.107 \cdot 10^{-3}$	$0.492 \cdot 10^{-4}$
1200	$0.235 \cdot 10^{-4}$	$0.140 \cdot 10^{-4}$	$0.696 \cdot 10^{-4}$	$0.340 \cdot 10^{-4}$

Table 1.3: Next-to-leading order (NLO) pair production cross sections and PDF uncertainties for scalar leptoquarks in pp collisions at $\sqrt{s} = 7$ TeV (the LHC collision energy in 2010 and 2011) and $\sqrt{s} = 8$ TeV (the LHC collision energy in 2012). Cross sections are given in units of pb (1 b $= 10^{-28}\mathrm{m}^2 = 10^{-24}\mathrm{cm}^2$). In all cases, the renormalization and factorization scale is set to be equal to the leptoquark mass.

leptoquarks in pp collisions for a given $\sqrt{s}$. The cross sections provided by this tool at at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV are shown in Table 1.3 as calculated using the `CTEQ6L1` parton density function (PDF) [55].

1.2.3 Leptoquark decays

Under the mBRW model, scalar and vector leptoquarks decay to a lepton-quark pair with total decay widths given by Equation 1.4 [34, 49].

$$\begin{aligned} \Gamma_{\text{Scalar}} &= \sum_i \frac{\lambda_i^2}{16\pi} M_{\text{LQ}} \\ \Gamma_{\text{Vector}} &= \sum_i \frac{\lambda_i^2}{24\pi} M_{\text{LQ}} \end{aligned} \tag{1.4}$$

In Equation 1.4, M_{LQ} corresponds to the mass of the leptoquark, and λ_i corresponds to the coupling between the leptoquark, lepton, and quark for a given decay mode (i). The sum is taken over all leptoquark decay modes. For a 1 TeV scalar leptoquark having a single decay mode with a coupling constant equal to that of the electromagnetic interaction ($\lambda_{\text{em}} = 0.3$), the decay width is approximately 1.8 GeV. For a similar vector leptoquark, the width is approximately 1.2 GeV. These widths are significantly less than experimental resolutions.

It should be noted that according to Equation 1.4, leptoquark decay width is directly proportional to the coupling constant (λ), and a small coupling constant corresponds to a long lived (narrow width) leptoquark. While searches for the pair-production of scalar leptoquarks are generally treated independently of the coupling constant, the coupling constant is generally assumed to be large enough so that the leptoquark decays relatively promptly. For example, a scalar leptoquark with $\lambda \sim 3 \cdot 10^{-8}$ has a decay length of approximately 1 meter.

Under the mBRW model, the branching fraction $\beta = \text{BR}(\text{LQ} \to \ell^{\pm} q)$ is required to have a value of either 0, 1/2, or 1. This can be seen in Table 1.2. β may be treated with less rigidity if the assumption **3** of the BRW model is relaxed ("leptoquarks couple *only* to Standard Model fermions and gauge bosons"). For example, in the case of RPV SUSY, leptoquark-like squarks may decay to leptons, quarks, and other SUSY partners. Many other examples exist for other models [56–58]. In most experimental searches, β is treated as a free parameter with the additional assumption that $\text{BR}(\text{LQ} \to \ell^{\pm} q) + \text{BR}(\text{LQ} \to \nu q) = 1$.

This additional assumption leads to three separate final states for leptoquark pair pro-

Decay mode	Branching fraction
$LQ\overline{LQ} \to \ell^- q \ell^+ \overline{q}$	β^2
$LQ\overline{LQ} \to \ell^- q \overline{\nu}_\ell \overline{q}', \nu_\ell q' \ell^+ \overline{q}$	$2\beta(1-\beta)$
$LQ\overline{LQ} \to \nu_\ell q \overline{\nu}_\ell \overline{q}$	$(1-\beta)^2$

Table 1.4: List of the three distinct final states for pair produced leptoquarks and their associated branching fractions. β is defined such that $\beta = \mathrm{BR}(LQ \to \ell^\pm q)$.

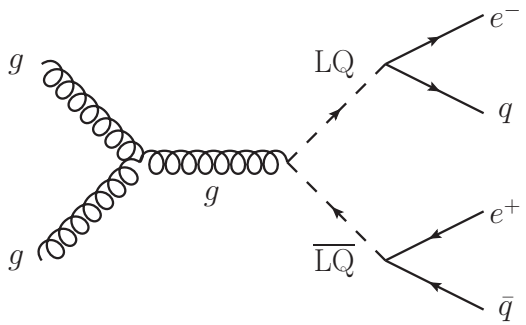
duction: two charged leptons and two quarks; one charged lepton, one neutrino, and two quarks; and two neutrinos and two quarks. For experimental searches, this corresponds to three different analysis final states: two charged leptons and two jets; one charged lepton, two jets, and missing transverse energy; and two jets and missing transverse energy. These final states and their branching ratios are listed in Table 1.4. The leading order Feynman diagrams for the case when first generation leptoquarks decay to an electron or electron neutrino are shown in Figure 1.6.

1.2.4 Current limits

Experimental limits on leptoquark states are obtained by indirect searches and direct searches. The indirect limits are calculated from bounds on leptoquark-induced four-fermion interactions, which come from low energy experiments, or from collider experiments below threshold. For a more detailed overview of indirect limits on leptoquarks, the reader is referred to Reference [59]. The direct limits on leptoquarks are calculated from their production cross section at colliders. Direct limits on first generation leptoquarks are discussed below.

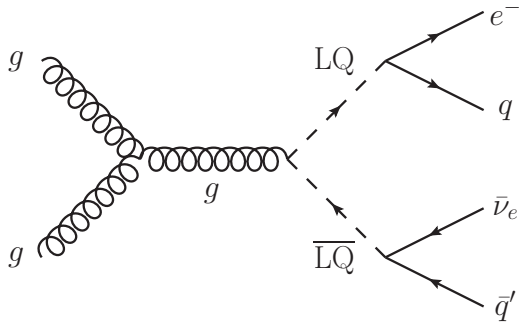
The most stringest limits on the mass of first generation leptoquarks from single production searches are produced by HERA experiments: ZEUS and H1. The most recent limits for all 14 of the first generation BRW leptoquarks are shown in Figure 1.7 [60]. These λ-dependent and model-dependent limits were obtained using ep collisions recorded with the ZEUS detector. Comparable limits were obtained by H1 [61].

Prior to the release of the results contained in this thesis, the most stringent limits on the



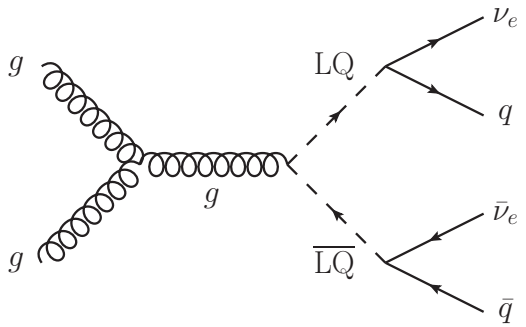
$\text{BR}(\text{LQ}\overline{\text{LQ}} \to e^- q e^+ \bar{q}) = \beta^2$

(a) $eejj$ channel



$\text{BR}(\text{LQ}\overline{\text{LQ}} \to e^- q \bar{\nu}_e \bar{q}', \nu_e q' e^+ \bar{q}) = 2\beta(1-\beta)$

(b) $e\nu jj$ channel



$\text{BR}(\text{LQ}\overline{\text{LQ}} \to \nu_e q \bar{\nu}_e \bar{q}) = (1-\beta)^2$

(c) $\nu\nu jj$ channel

Figure 1.6: Possible final states for a decay of a pair of first-generation scalar leptoquarks. q denotes either an up or down quark.

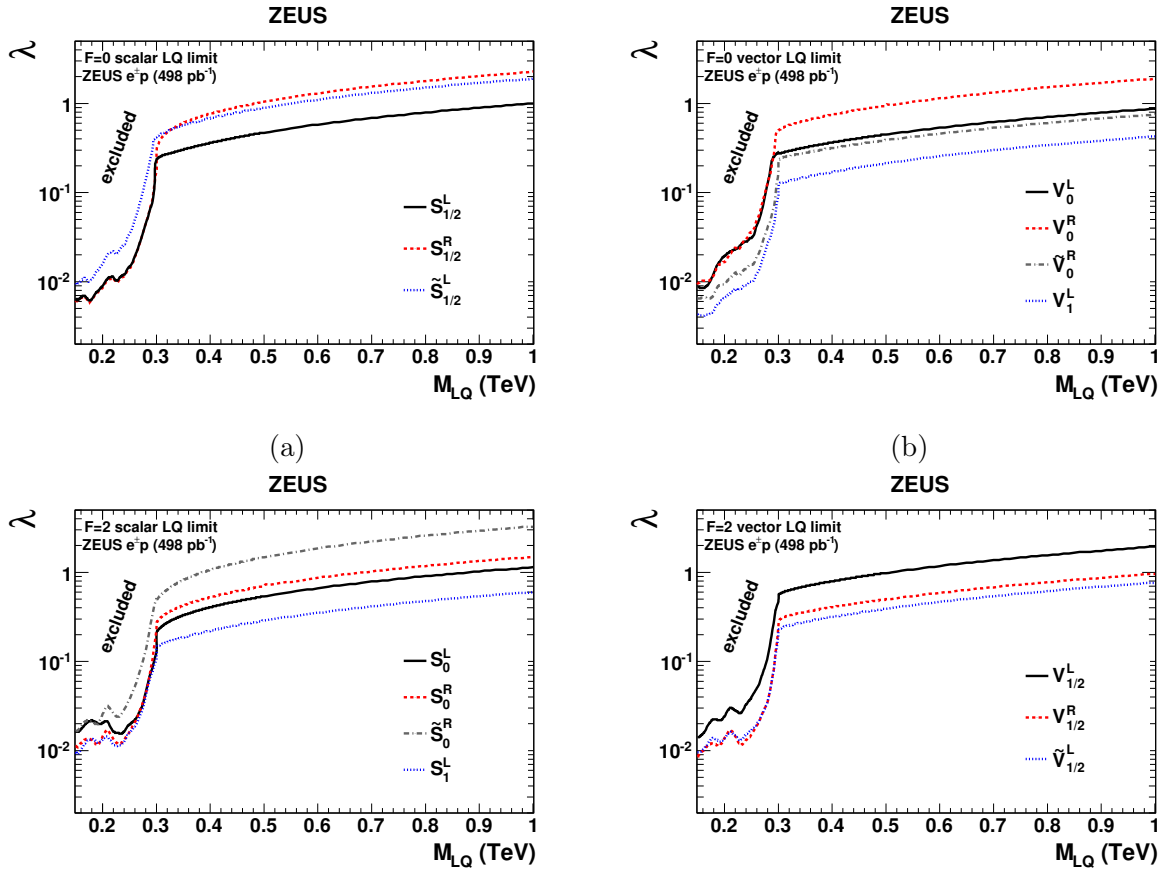


Figure 1.7: 95% confidence level exclusion limits for all 14 of the first generation leptoquarks described by the BRW model. These limits were obtained using electron-proton and positron-proton collisions recorded with the ZEUS detector. In each plot, the $x-$axis corresponds to leptoquark mass, and the $y-$axis corresponds to the leptoquark coupling λ. Plot (a) shows results for the $F = 0$ scalar leptoquarks. Plot (b) shows results for the $F = 0$ vector leptoquarks. Plot (c) shows results for the $F = 2$ scalar leptoquarks. Plot (d) shows results for the $F = 2$ vector leptoquarks [60].

mass of first generation leptoquarks from pair production searches came from the ATLAS experiment at the CERN LHC [51]. Using roughly the earliest 20% of the full 2011 LHC dataset, ATLAS excluded the production of first generation scalar leptoquarks with masses less than 660 (607) GeV, when assuming the branching fraction of a leptoquark to an electron or positron (β) was equal to 1.0 (0.5). The CMS experiment (also at the CERN LHC) also produced limits on first generation leptoquarks in a prior version of the analysis described by this thesis [62, 63]. Using the 2010 LHC dataset, CMS excluded the production of first generation scalar leptoquarks with masses less than 384 (340) GeV, when assuming the branching fraction of a leptoquark to an electron or positron (β) was equal to 1.0 (0.5). The exclusion limits on first generation leptoquarks in the leptoquark mass vs. β plane from the ATLAS search are shown in Figure 1.8. The same results from the CMS search are shown in Figure 1.9. The ATLAS limit is more stringent than the CMS limit in part because the 2011 dataset used by the ATLAS search was roughly 25 times the size of the 2010 dataset used by the CMS search.

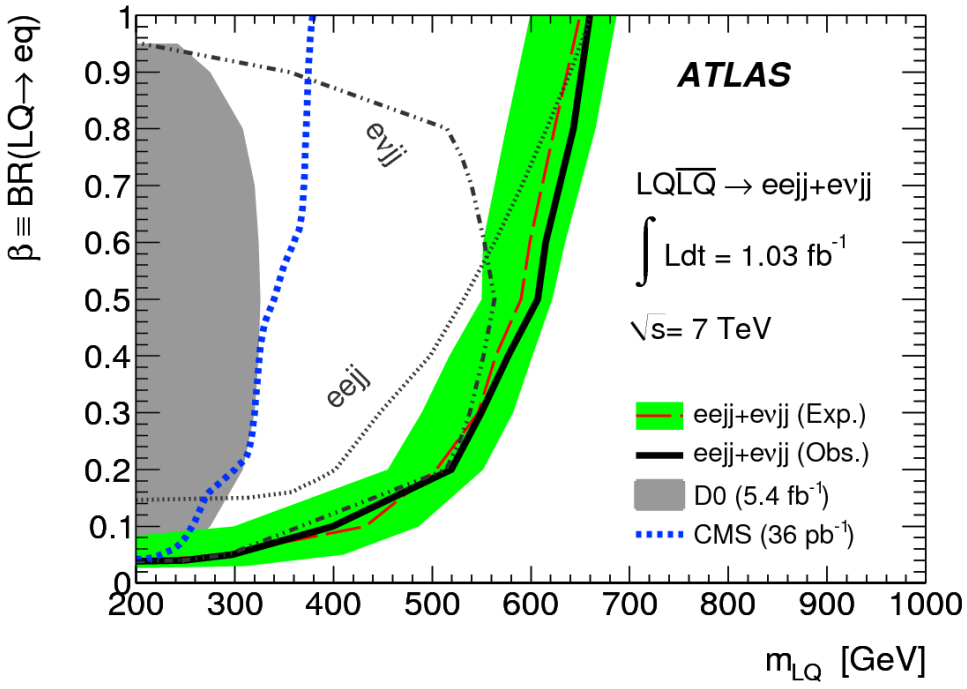


Figure 1.8: The expected and observed exclusion limits at 95% CL on the first generation leptoquark hypothesis in the β vs. mass plane using the central value of signal cross section for the individual $eejj$ and $e\nu jj$ channels and their combination, from the ATLAS experiment in 2011. The green expected limit uncertainty band represents the 95% confidence interval. The black solid line represents the observed combined exclusion limit. The blue dashed line represents the observed exclusion limit from the CMS experiment in 2010 [51].

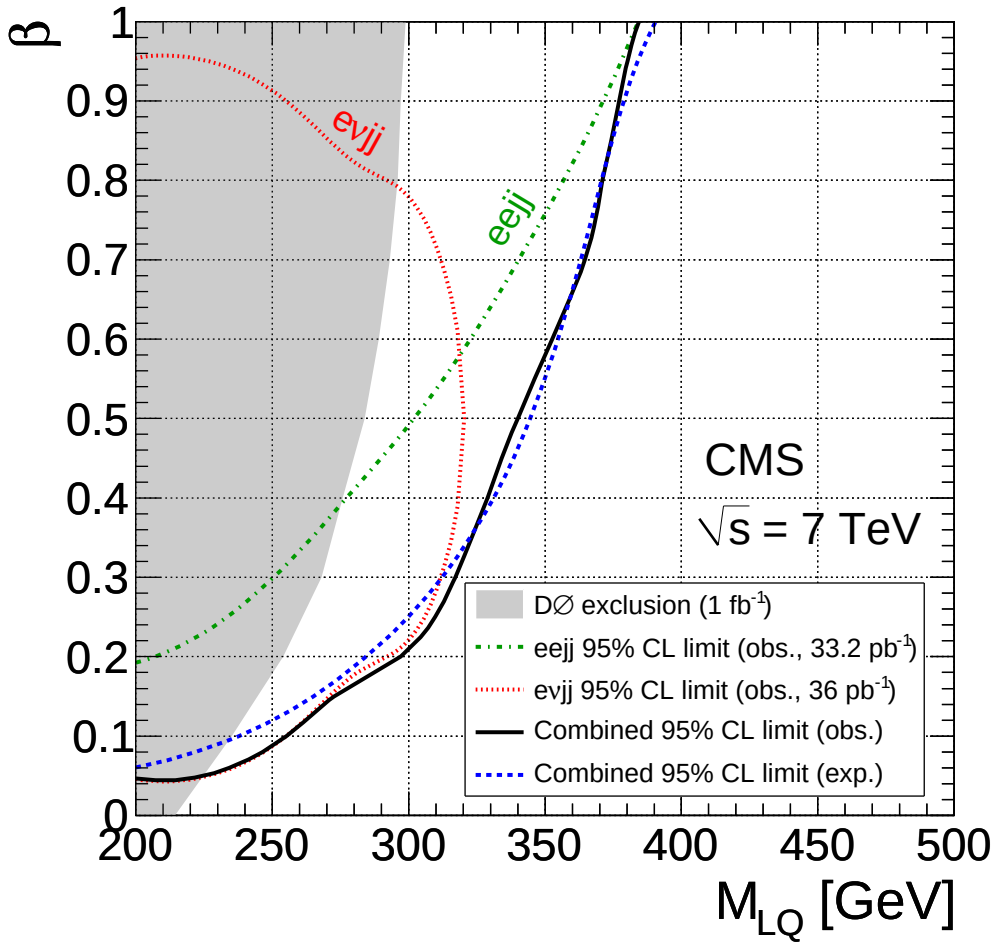


Figure 1.9: The expected and observed exclusion limits at 95% CL on the first generation leptoquark hypothesis in the β vs. mass plane using the central value of signal cross section for the individual $eejj$ and $e\nu jj$ channels and their combination, from the CMS experiment in 2010. The black solid line represents the observed combined exclusion limit, while the blue dashed line represents the expected exclusion limit. The gray shaded area represents the observed exclusion limit from the D0 experiment at the Tevatron [62, 63].

CHAPTER 2

EXPERIMENTAL APPARATUS

This chapter describes the experimental apparatus used to search for leptoquarks in high energy pp collisions. Section 2.1 describes the Large Hadron Collider, which produces high energy pp collisions. Section 2.2 describes the Compact Muon Solenoid detector, which records the results of those collisions.

2.1 Large Hadron Collider

This chapter describes the Large Hadron Collider (LHC) [64]. The LHC is a two-ring superconducting synchrotron, built and operated by the European Organization for Nuclear Research (CERN). The LHC's main purpose is to accelerate two beams of protons in opposite directions up to a nominal beam energy of 7 TeV and to collide them with an instantaneous luminosity[1] of up to 10^{34} cm^{-2}s^{-1}. The LHC is also capable of colliding beams of lead ions ($^{208}_{82}\mathrm{Pb}^{82+}$) with a beam energy of up to 2.76 TeV per nucleon and an instantaneous luminosity of up to 10^{27} cm^{-2}s^{-1}. The remainder of this chapter will discuss the LHC as it pertains to accelerating and colliding protons. Section 2.1.1 of this chapter describes the various

[1] Instantaneous luminosity ($\mathcal{L}$) is defined as the ratio between the production rate in Hz for given physics process in a collider and the production cross section for that process. Instantaneous luminosity depends only on beam parameters, and it is expressed in units of cm^{-2}s^{-1}. It is discussed in further detail in Section 2.1.2.

components of the LHC and its injector complex, while Section 2.1.2 of this chapter lists various parameters associated with LHC operations.

2.1.1 LHC structure

The LHC lies in a 26.7 km long tunnel about 100 m beneath the suburbs of Geneva, Switzerland, as shown in Figure 2.1. The tunnel was dug between 1984 and 1989 and was originally used for the Large Electron-Positron Collider (LEP). Both the LHC and its tunnel are roughly circular, consisting of eight 2.45 km long arcs and eight 545 m long straight sections called "insertions" that connect the arcs.

Each of the LHC's eight arcs contains 154 superconducting dipole magnets (1232 in total) designed to steer two proton beams in opposite directions around the LHC. The dipole magnets contain two separate bores, each containing one of the proton rings. Each dipole magnet is 14.3 m long and weighs 35 tons. The dipoles produce a field of up to 8.3 T using a current of 11,850 A, and their superconductivity is maintained using liquid helium at 1.9 K. A cross section of an LHC dipole magnet is shown in Figure 2.2.

The contents of the LHC's eight insertions vary. The location of each insertion is described as a Point, and each insertion-Point pair is given a number from 1 to 8. The insertions at Points 1, 2, 5, and 8 serve as interaction points, where proton beams cross from one magnet bore to the other and collide within a 130 m long common section. These four interaction points are the sites of four particle detectors, each designed to record the results of high energy particle collisions. Two of these detectors, "A Toroidal LHC ApparatuS" (ATLAS, Point 1) [66] and the "Compact Muon Solenoid" (CMS, Point 5) [67] are general purpose detectors designed to study high luminosity collisions. The "Large Hadron Collider beauty" detector (LHCb, Point 8) [68] is designed to study the physics of bottom quarks. "A Large Ion Collider Experiment" (ALICE, Point 2) [69] is a detector designed to study heavy ion collisions. The insertions at Points 3 and 7 contain beam collimation systems. The insertion at Point 4 contains two radio frequency (RF) cavity systems, designed to accelerate the

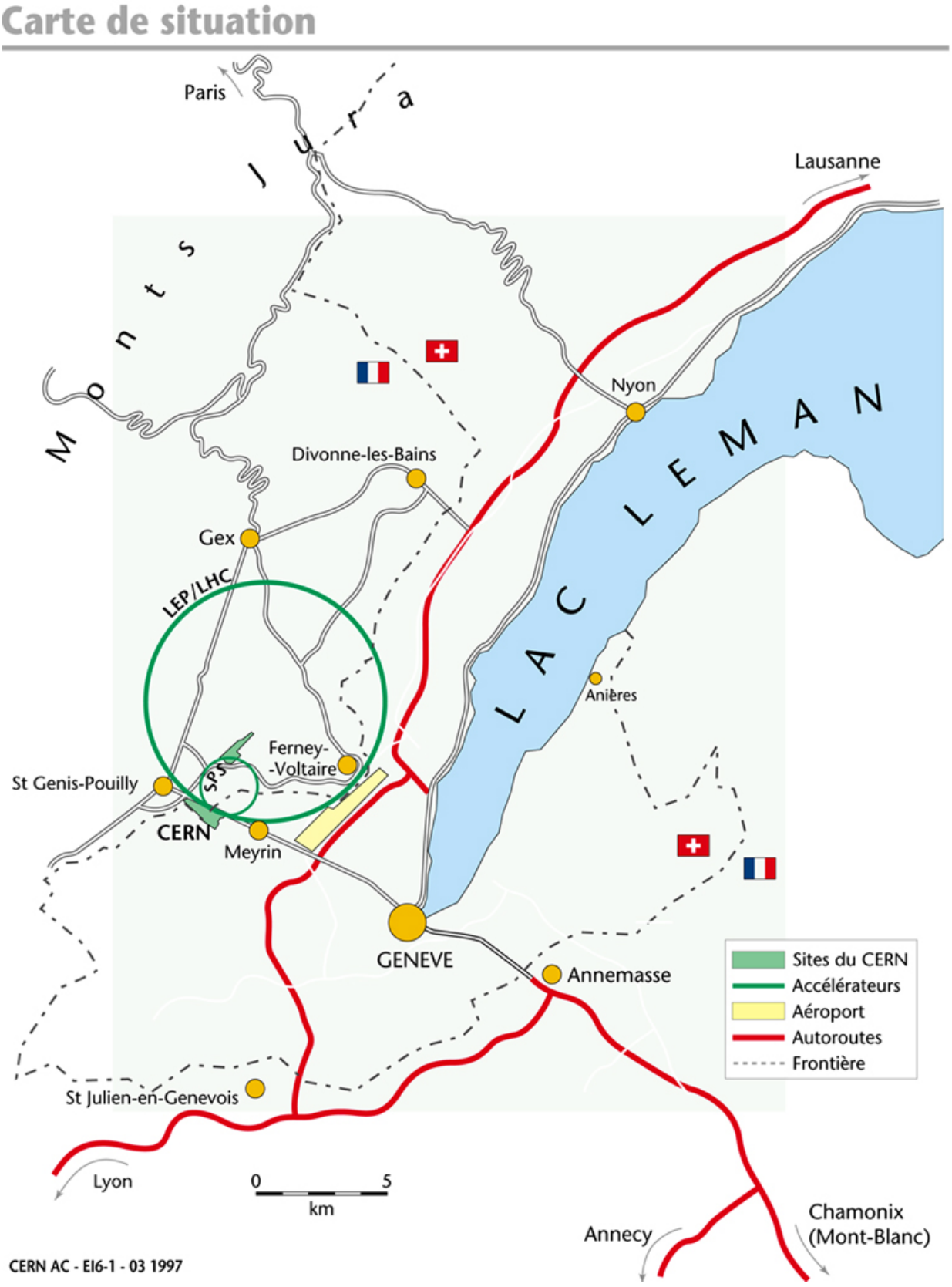


Figure 2.1: Map of CERN, the LHC, the SPS, and the surrounding area [65].

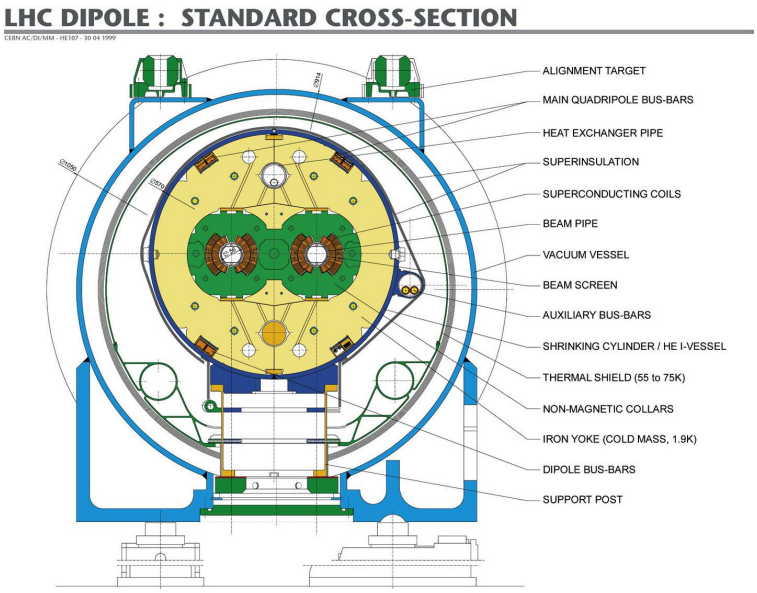


Figure 2.2: Cross section of one of the 1232 LHC dipole magnets [64].

proton beams to higher energies. Each RF cavity system contains eight RF cavities, and each RF cavity delivers 2 megavolts at 400 MHz. The insertion at Point 6 contains systems for dumping the beam from the LHC. Each beam has its own dumping system. A section of the LHC beginning in the middle of one arc and ending in the middle of an adjacent arc (and thereby containing one entire insertion) is called an octant. Octants are labeled with the same number as the interaction point they contain. A schematic of the various octants can be seen in Figure 2.3.

The accelerator complex at CERN serves as the injector for the LHC. The complex is responsible for extracting protons from a bottle of hydrogen gas, accelerating them in stages to 450 GeV, and injecting them into the LHC. Protons are extracted from hydrogen gas by means of a duoplasmatron and injected into a linear accelerator (Linac2) where they are accelerated to 50 MeV. The protons are then passed through a chain of three synchrotron accelerators. First, the protons are transfered to the Proton Synchrotron Booster (PSB), where they are accelerated to 1.4 GeV. Next, the protons are transfered to the Proton Synchrotron (PS), where they arranged in bunches and accelerated to 25 GeV. Finally, the protons are transfered to the Super Proton Synchrotron (SPS), where they are accelerated to

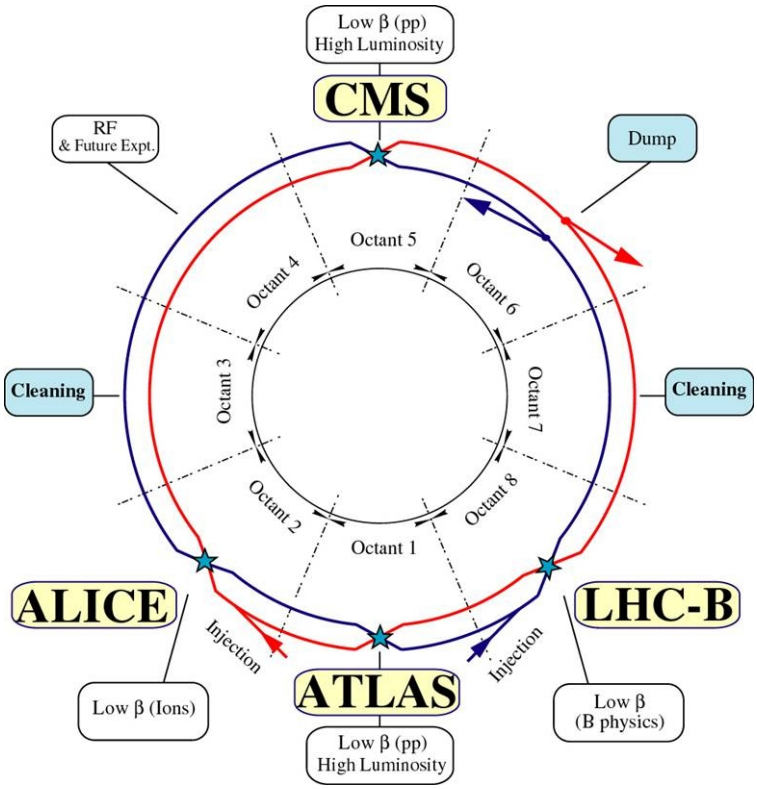


Figure 2.3: Schematic of the LHC including the octants, the role of the insertion within each octant, and the major experiments [64]

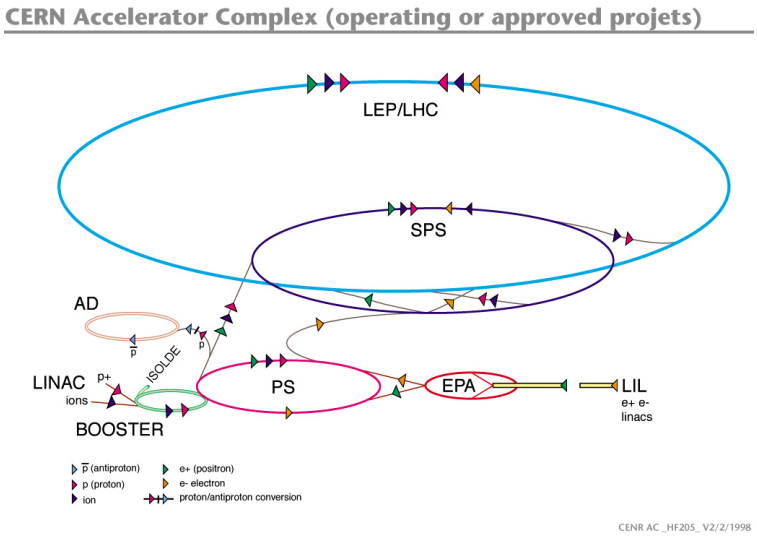


Figure 2.4: Schematic of the CERN accelerator complex [70]

450 GeV before being injected into the LHC. A schematic of the CERN accelerator complex is shown in Figure 2.4.

The arrangement of the beam into bunches as described in the previous paragraph is an important concept. Protons are injected into the LHC in bunches spaced 25 ns apart. Collisions within the LHC primarily take place when bunches cross path at the interaction points. These 25 ns "bunch crossings" form a window in time during which the detectors described above record as much information about a *pp* interaction as they can before the next bunch crossing arrives.

2.1.2 LHC operating parameters

The LHC was designed to search for rare physics processes (i.e. processes with a small production cross section). A critical element involved in searching for these processes is an estimate of the rate (in Hz) with which they will be produced. The ratio between the production rate of a physics process in a collider and the production cross section of that process is known as "instantaneous luminosity", $\mathcal{L}$, and it is measured in units of $\mathrm{cm}^{-2}\mathrm{s}^{-1}$. "Integrated luminosity", $\mathcal{L}_{int}$, refers to an integral of instantanenous luminosity over a period of time, and it is expressed in units of cm^{-2}. The relationship between instantaneous luminosity,

$\mathcal{L}$, the production cross section for a given physics process, σ_{event}, and the number of events produced, N_{event}, is given by Equation 2.1:

$$N_{\text{event}} = \int \sigma_{\text{event}} \mathcal{L} dt \tag{2.1}$$

Luminosity depends only on beam parameters. In the case of the LHC, instantaneous luminosity may be calculated using Equation 2.2:

$$\mathcal{L} = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \varepsilon_n \beta^*} F \tag{2.2}$$

where N_b is the number of protons in a bunch, n_b is the number of bunches in a beam, f_{rev} is the revolution frequency, γ_r is the relativistic gamma factor, ε_n is the transverse beam emmitence, β^* is the value of the beta function at the interaction point (IP), and F is the geometric luminosity reduction factor due to the crossing angle between the two beams at the IP. F may be calculated using Equation 2.3:

$$F = \left(1 + \frac{\theta_c \sigma_z}{2\sigma^*}\right)^{-1/2} \tag{2.3}$$

where θ_c is the crossing angle at the IP, σ^* is the RMS of the transverse beam size at the IP, and σ_z is the RMS of the bunch length. This relationship assumes round beam profiles and equal parameters for both beams. The nominal values for each of these parameters at the LHC may be found in Table 2.1.

The nominal conditions described in Table 2.1 differ somewhat from the conditions under which the LHC operated during its first three years of running. The peak instantaneous luminosity achieved during stable pp collisions as of December 2012 was $0.767 \cdot 10^{34} \text{cm}^{-2}\text{s}^{-1}$ (77% of nominal). This value was achieved with 1368 colliding bunches (49% of nominal). In addition, the LHC has only run with a maximum beam energy of 4 TeV (57% of nominal). The maximum instantaneous luminosity and beam energy for each year between 2010 and

Parameter		Unit	Nominal value
N_b	number of particles per bunch		$1.15 \cdot 10^{11}$
n_b	number of bunches per beam		2808
f_{rev}	revolution frequency	kHz	11.246
γ_r	relativistic γ factor		7461
ε_n	transverse beam emittance	μm	3.75
β^*	beta function at the IP	m	0.55
F	geometric $\mathcal{L}$ reduction factor		0.836
θ_c	crossing action at IP	μrad	285
σ_z	RMS bunch length	cm	7.55
σ^*	RMS beam size at the IP	μm	16.7
$\mathcal{L}$	Instantaneous luminosity	$\text{cm}^{-2}\text{s}^{-1}$	10^{34}

Table 2.1: Nominal LHC beam parameters contributing to the luminosity, $\mathcal{L}$, as defined in Equation 2.2. [64]

Year	Beam energy (TeV)	Maximum instantaneous $\mathcal{L}$ ($10^{34}\text{cm}^{-2}\text{s}^{-1}$)
2010	3.5	0.021
2011	3.5	0.354
2012	4.0	0.767
Nominal	7.0	1.0

Table 2.2: The LHC beam energy and maximum instantaneous luminosity at Point 5 (CMS) for the years 2010-2012, as compared with nominal design values [71]

2012 is shown in Table 2.2. The integrated luminosity that the LHC delivered to Point 5 for the years 2010 through 2012 is shown as a function of time in Figure 2.5.

2.2 Compact Muon Solenoid

The Compact Muon Solenoid (CMS) detector is designed to measure pp collisions from the LHC. CMS is installed 100 m underground near the French village of Cessy. CMS is barrel-shaped and stretches roughly 25 m long with a diameter of roughly 15 m. It weighs 12,500 tons. A schematic of CMS is shown in Figure 2.6.

CMS is designed to operate in a very challenging environment. The total pp cross section for collisions with beam energy of 7 TeV is expected to be about 100 mb. The LHC design luminosity is $10^{34}\text{cm}^{-2}\text{s}^{-1}$, which corresponds to a collision rate of 10^9 collisions per second.

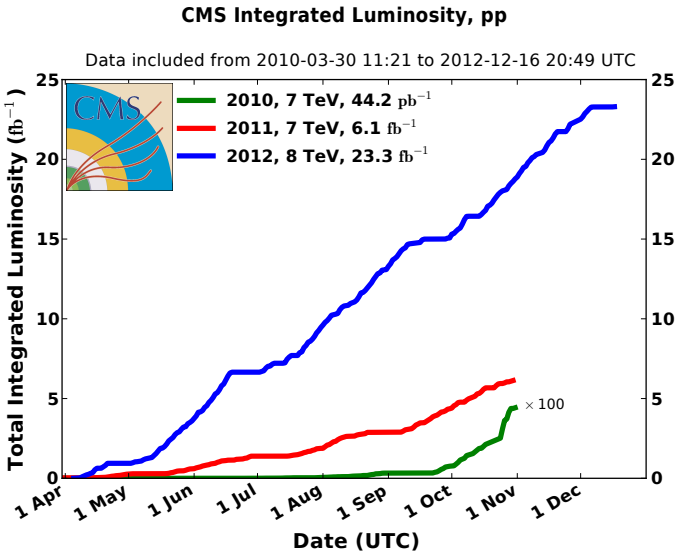


Figure 2.5: Total integrated luminosity delivered by the LHC to Point 5 (CMS) as a function of time, for the years 2010-2012 [71]

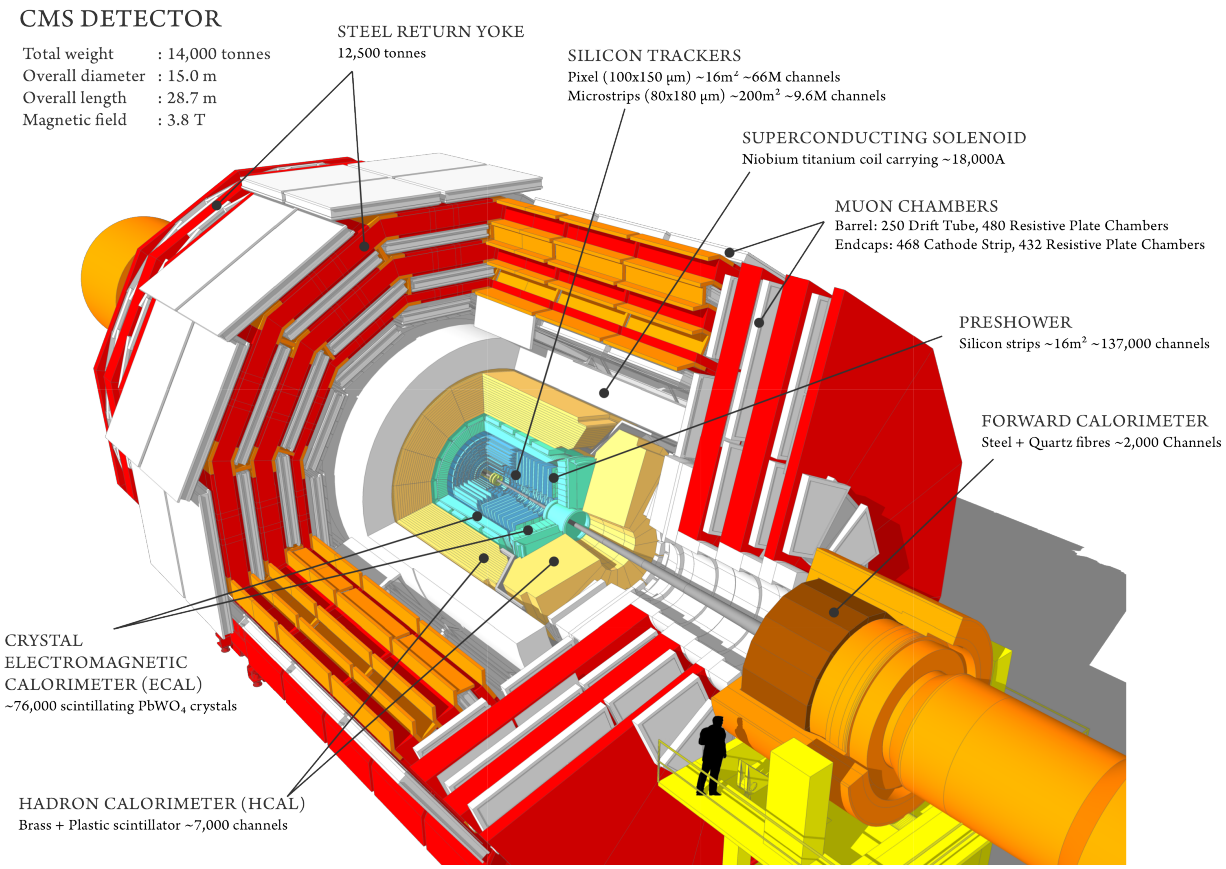


Figure 2.6: A cartoon schematic of the CMS detector

Furthermore, in addition to the pp interaction of interest, 20 inelastic pp collisions will take place in the same bunch crossing (an effect known as "pile-up"). In total, at design luminosity, roughly 1000 charged particles will be produced at the interaction region every 25 ns. This large flux of particles from the interaction region leads to high radiation levels, which can impact the performance of the detector and front-end electronics. In addition, charged particles from pile-up interactions may be confused with charged particles from the interaction of interest. This effect can be ameliorated somewhat by using high granularity detectors with good timing resolution (and therefore low occupancy).

The detector requirements for CMS may be summarized as follows:

- Good muon identification and momentum resolution over a wide range of momenta and angles, good dimuon mass resolution ($\approx 1\%$ at 100 GeV), and the ability to determine unambiguously the charge of muons with $p < 1$ TeV.

- Good charged particle momentum resolution and reconstruction efficiency in the inner tracker. Efficient triggering and offline tagging of τ's and b-jets, requiring pixel detectors close to the interaction region.

- Good electromagnetic energy resolution, good diphoton and dielectron mass resolution ($\approx 1\%$ at 100 GeV), wide geometric coverage, π^0 rejection, and efficient photon and lepton isolation at high luminosities.

- Good resolution of the imbalance of the total energy measured in the transverse plane and good dijet-mass resolution, requiring hadron calorimeters with a large hermetic geometric coverage and fine lateral segmentation.

The design of the CMS detector is described in this section, and it meets these requirements. Unless otherwise stated, all of the information in this section comes from References [67] and [72].

2.2.1 Coordinate system

The CMS detector is described using a coordinate system with its origin at the nominal collision point. The x-axis points radially inward, towards the center of the LHC; the y-axis points vertically upward; and the z-axis points along the clockwise beam direction, towards the Jura Mountains. Because CMS is cyllindrical in shape, it is often described using a cyllindrical coordinate system. The radial distance, r, is measured from the origin. The longitudinal direction is the same as the z-axis, and the transverse direction is in the $x-y$-plane. The azimuthal angle, ϕ, is measured in the $x-y$-plane, and the x-axis is taken to be $\phi = 0$. The polar angle, θ, is measured from the z-axis. When referring to particle trajectory, rapidity, y, is often used instead of θ because rapidity difference is Lorentz-invariant against a boost in the z-axis, while θ is not. A particle's rapidity is defined according to Equation 2.4:

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right) \tag{2.4}$$

where E is particle energy and p_z is particle momentum in the z-direction. For relativistic particles (such that $E \gg m$) rapidity may be approximated as "pseudorapidity", η. Pseudorapidity may be expressed in terms of the polar angle, θ, using Equation 2.5:

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \tag{2.5}$$

A particle's momentum and energy in the transverse direction are denoted p_T and E_T, the values for which are calculated from their x and y components. The imbalance of the total energy measured in the transverse plane is refered to as a vector: $\vec{\not{E}}_T$. The magnitude of $\vec{\not{E}}_T$ is refered to as $\not{E}_T$ or "missing transverse energy" (MET).

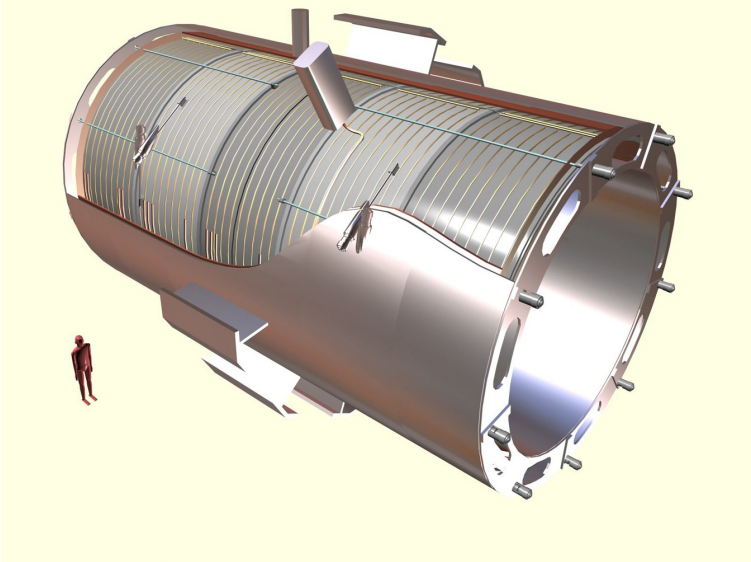


Figure 2.7: Artistic representation of the CMS solenoid. The outer layer (cryostat) has been cut away to show the solenoid's five individual sections. [67].

2.2.2 Solenoid

As its name suggests, a central feature of the CMS detector is a large, superconducting solenoid. Large bending power (i.e. a large magnetic field) is necessary in order to unambiguously determine the sign of the electric charge of muons with momentum greater than 1 TeV. This requirement forced the choice of a superconducting technology for the solenoid. In addition, the solenoid was designed to contain both the inner tracking and calorimetry detector subsystems, while maintaining a favorable length-to-diameter ratio in order to ensure good track momentum resolution in the forward region. The nominal specifications are for a 4 T field with 2.6 GJ of stored energy at full current in a free bore of 6 m diameter and a length of 12.5 m. In practice, the solenoid produces a magnetic field of 3.8 T with 2.35 GJ of stored energy. The flux is returned via a 10,000 ton iron return yoke made up of 5 wheels and two endcaps. Each endcap is composed of three disks. The solenoid is made up of five separate sections and placed within a cryostat, as shown in Figure 2.7. The cryostat is cooled to 4.5 K in order to maintain superconductivity. Other design parameters for the solenoid are given in Table 2.3.

The CMS solenoid has several features that are distinctive relative to the solenoids used in

General parameters	
Magnetic length	12.5 m
Cold bore diameter	6.3 m
Central magnetic induction	4 T
Total Ampere-turns	41.7 MA-turns
Nominal current Z	19.14 kA
Inductance	14.2 H
Stored energy	2.6 GJ
Cold mass	
Layout	Five modules mechanically and electrically coupled
Radial thickness of cold mass	312 mm
Radiation thickness of cold mass	3.9 X_0
Weight of cold mass	220 t
Maximum induction on conductor	4.6 T
Temperature margin wrt operating temperature	1.8 K
Stored energy/unit cold mass	11.6 kJ/kg
Iron yoke	
Outer diameter of the iron flats	14 m
Length of barrel	13 m
Thickness of the iron layers in barrel	300, 630 and 630 mm
Mass of iron in barrel	6000 t
Thickness of iron disks in endcaps	250, 600, and 600 mm
Mass of iron in each endcap	2000 t
Total mass of iron in return yoke	10000 t

Table 2.3: Design parameters of the CMS solenoid [67]

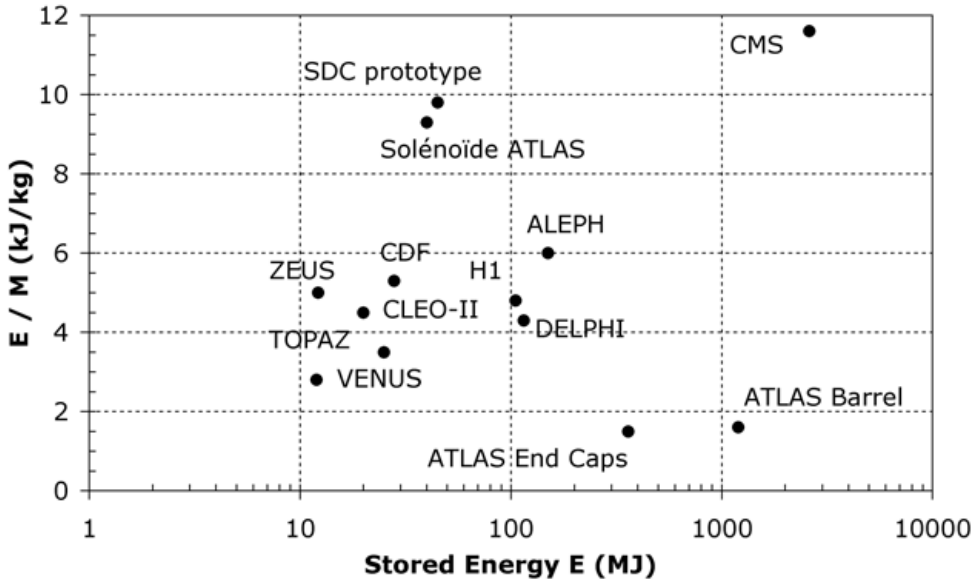


Figure 2.8: Comparison between the CMS solenoid and various other detector magnets. The x-axis (log scale) corresponds to the stored energy, E, of the magnet in units of MJ. The y-axis corresponds to the ratio between nominal stored energy and nominal effective cold coil mass of the magnet, E/M in units of kJ/kg. The CMS solenoid surpasses all of the other magnets displayed here with respect to both variables [67].

previous particle detectors. First, due to the large number of ampere turns needed to generate a 4 T field (41.7 MA-turn), the CMS solenoid winding has four layers instead of the usual single layer. Second, the conductor is made from a Rutherford-type cable of NbTi co-extruded with pure aluminum and mechanically reinforced with an aluminum alloy. Finally, the ratio between the nominal stored energy (2.6 GJ) and the nominal effective cold coil mass (220 t) is significantly higher for the CMS solenoid than for solenoids of previous detectors, as shown in Figure 2.8. This leads to a 0.15% mechanical deformation during energizing [67, 72].

2.2.3 Inner tracker

The purpose of the inner tracker of the CMS detector is to precisely and efficiently measure the trajectories of charged particles produced by LHC collisions. Reconstructed particle trajectories are known as "tracks". The inner tracker is composed of about 200 m^2 of active silicon area, making it the largest silicon tracker ever built. It rests within the bore of

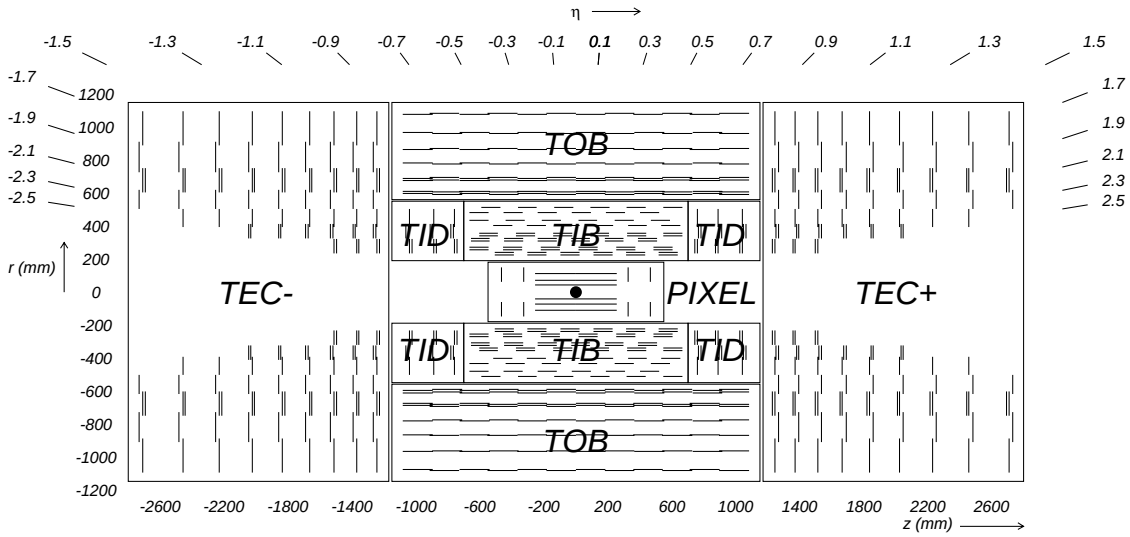


Figure 2.9: A schematic cross section of the inner tracker of the CMS detector. The pixel tracker is labeled near the interaction point at the center of the schematic. The four components of the silicon strip tracker (TIB, TOB, TID, and TEC) are also shown [67].

the CMS solenoid (Section 2.2.2), which provides a homogenous magnetic field of 3.8 T around the full inner tracker volume. By reconstructing trajectories of charged particles with $p_{\mathrm{T}} > 1$ GeV, the inner tracker fulfills several important tasks. First, the inner tracker makes it possible to reconstruct primary decay vertices, secondary decay vertices, and the displacement of tracks from those vertices (known as the track "impact parameter"), which allow for the identification of heavy flavor quarks. Second, along with the electromagnetic calorimeter (Section 2.2.4) and the muon system (Section 2.2.6), the inner tracker allows for the identification of electrons and muons. Third, the inner tracker allows for the identification of tau leptons by identifying various decay topologies. In addition, the inner tracker provides essential input to the high level trigger (Section 2.2.8). As shown in the schematic cross section in Figure 2.9, the inner tracker is made up of two detector subsystems: the pixel tracker and the strip tracker. Both detectors the cover pseudorapidity region $|\eta| < 2.5$. The resolution of several track parameters is shown in Figure 2.10.

The pixel detector consists of three barrel layers (BPix) and two endcap disks (FPix) on either side of the IP. The BPix layers are 53 cm long and have mean radii of 4.4, 7.3, and

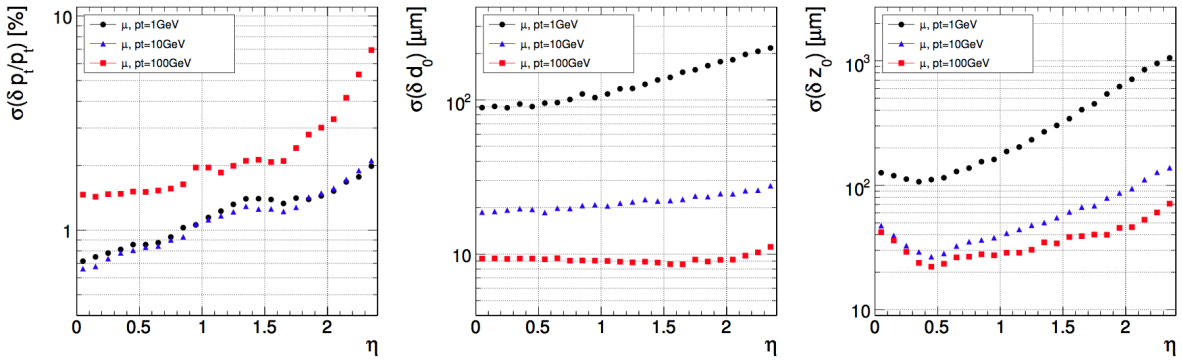


Figure 2.10: The resolution of various parameters of muon tracks as a function of the absolute value of pseudorapidity as reconstructed by the CMS inner tracker. The plot on the left shows transverse momentum, the central plot shows transverse impact parameter, and the plot on the right shows longitudinal impact parameter. Black markers correspond to muons with a transverse momentum of 1 GeV, blue markers correspond to muons with a transverse momentum of 10 GeV, and red markers correspond to muons with a transverse momentum of 100 GeV [67].

10.2 cm. The FPix layers extend from 6 to 15 cm in radius and are placed on either size at $z = \pm 34.5$ cm and $z = \pm 46.6$ cm. A schematic is shown in Figure 2.11. The pixel detector consists of 66 million pixel cells in total (48 million in BPix, 18 million in FPix). Each pixel is 100×150 μm^2, and the total active silicon area for the pixel detector is 1.06 m^2 (0.78 m^2 in BPix and 0.28 m^2 in FPix). Pixels of this size are chosen to ensure an occupancy of 10^{-4} per LHC crossing at design beam energy and instantaneous luminosity. Even during heavy ion collisions, this occupancy is about 10^{-2} in the pixel detector, allowing for efficiency track reconstruction in a high particle density environment. This arrangement of the FPix and BPix detectors ensures that the pixel detector offers three hit coverage over nearly the entire pseudorapidity range.

The strip detector surrounds the pixel detector in the radial region between 20 and 116 cm, and it is made up of four separate detector subsystems. The inner barrel region of the strip detector extends in radius out to 55 cm and is made up of four barrel layers (TIB) with three disks at each end (TID). In this region, the particle flux is low enough to allow the use of silicon microstrip detectors with minimum cell size of 10 cm $\times$ 80 μm, while still maintaining an occupancy of 2-3% per pp LHC crossing at design beam energy and

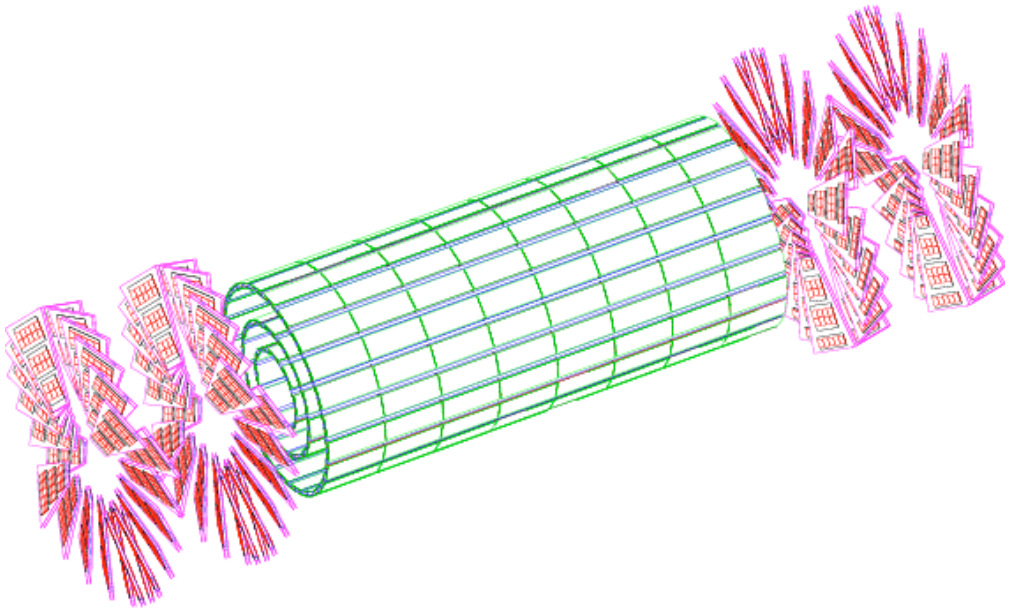


Figure 2.11: A schematic of the pixel tracker of the CMS detector. The three barrel layers (BPix) are shown in green. The four endcap disks (FPix) are shown in pink [72].

instantaneous luminosity. The outer barrel region of the strip detector extends in radius from 55 cm to 116 cm and is made up of an additional 6 barrel layers (TOB) . In this region, the particle flux is lower still, which allows the use of larger-pitch silicon microstrips with a maximum cell size of 25 cm $\times$ 180 μm, while still maintaining an occupancy of 1% per pp LHC crossing at design beam energy and instantaneous luminosity. The TIB, TID, and TOB subsystems collectively extend in out to $z = \pm 118$ cm. Beyond this z range are 9 additional disks on either side (TEC), occupying the region 124 cm $< |z| <$ 282 cm and 22.5 cm $< |r| <$ 113.5 cm. The first two layers in both the TIB and TOB subsystems, the first two disks of the TID subsystem, and disks 1, 2, and 5 of the TEC subsystem are made with modules that are laid in stereo in order to provide a measurement in both $r - \phi$ and $r - z$ coordinates. The pixel detector consists of 9.3 million silicon microstrips in total, and the total active silicon area for the strip detector is 198 m^2. This layout makes it likely that each track has at least nine hits in the region $|\eta| < 2.4$ with at least four two-dimensional measurements.

The CMS inner tracker operates in a high flux of particles, which can cause radiation damage both to the front-end electronics and to the silicon itself. In order to minimize this

Property	Value
Chemical composition	Lead tungstate crystal ($PbWO_4$)
Density	8.28 g/cm^3
Radiation length	0.89 cm
Interaction length	19.5 cm
Molière radius	2.2 cm

Table 2.4: Properties of the EB and EE scintillating crystals [67].

radiation damage and to ensure a high track finding efficiency and a low fake rate, the CMS tracker is operated at a temperature of $-10\,^{\circ}$C. This is made more difficult by two effects: first, the detector itself disappates nearly 60 kW of power; and second, the tracker faces the electromagnetic calorimeter, which must be operated at room temperature and must have good temperature stability. These effeets are overcome by lining the interior of the tracker's support structure with an active, 32-panel thermal screen. On the inside of the screen, cold fluid is circulated in a thin aluminum plate to keep the tracker cool. On the other side of 8mm of Rohacell foam, insulated resistive circuits are powered to keep the outer surface warm enough to avoid condensation. The screen acts in concert with the individual cooling systems of the inner tracker subsystems. It ensures a temperature below $-10\,^{\circ}$C inside the tracker volume and a temperature of $+12\,^{\circ}$C outside the tracker support structure, even when the inner tracker subsystems and their cooling systems are shut off. [67, 72].

2.2.4 Electromagnetic calorimeter

The CMS electromagnetic calorimeter (ECAL) is a hermetic, homogenous calorimeter made of lead tungstate ($PbWO_4$) crystals. Its purpose is to provide a high resolution measurement of the energy of electromagnetic showers. The ECAL is split into two subdetectors consisting of a barrel region (EB) and two endcap regions (EE). In addition, a preshower detector (ES) made of silicon strips is placed in front of the endcaps to identify neutral pions in the high pseudorapidity region. A schematic cross section of the ECAL is shown in Figure 2.12.

Lead tungstate crystals have a high density, a short radiation length, and a small Molière

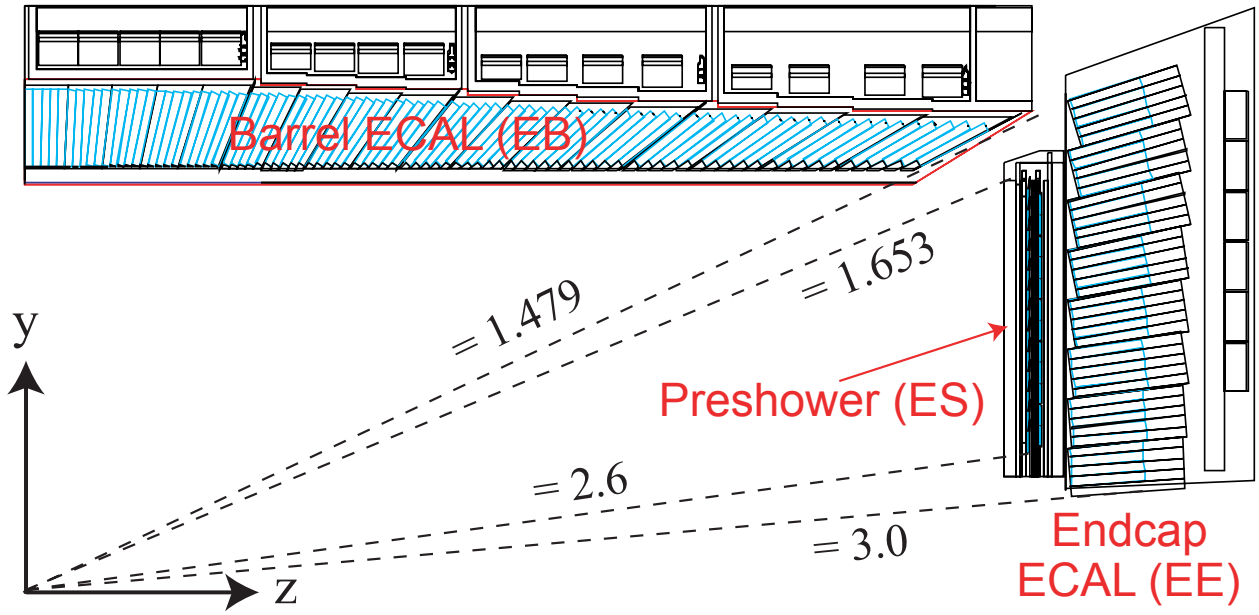
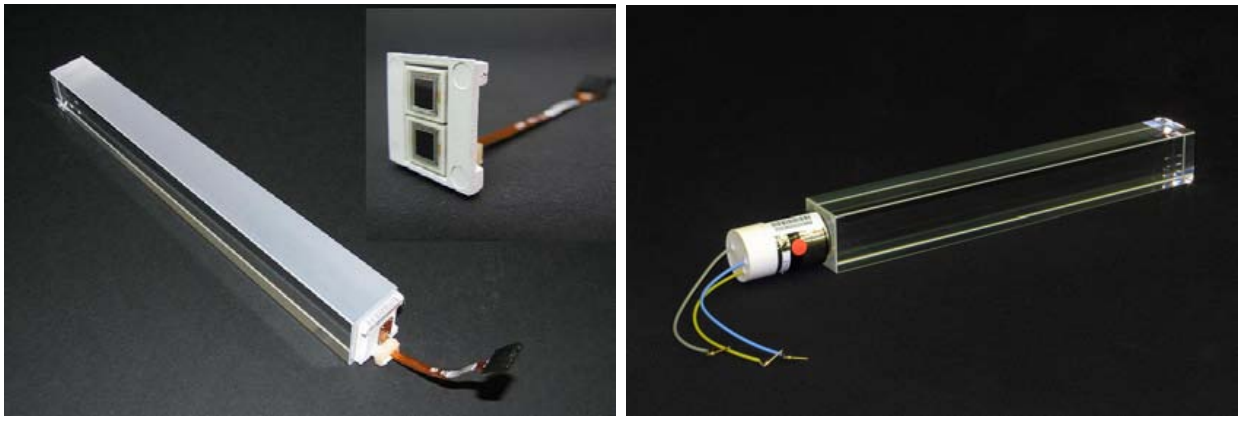


Figure 2.12: A schematic cross section of the CMS ECAL. Only one quarter of the entire ECAL is represented. The three subsystems of the ECAL: the barrel (EB), endcap (EE), and preshower (ES) are all shown. [72]

radius, which allow for a compact calorimeter with fine granularity. The numeric values for these parameters are given in Table 2.4. They are also optically clear and radiation resistant. The crystals scintillate with a relatively low light output that varies with temperature: maximum output is 4.5 photoelectrons per MeV at 18°C, and the output declines from that maximum at −2.1%/°C. The scintillation decay time is roughly the same as the LHC bunch crossing length: 80% of the light is emitted in 25 ns. The light output is blue-green with a broad maximum at 420-430 nm, and it is collected by avalanche photodiodes (APDs) in the EB and vacuum phototriodes (VPTs) in the EE. Photographs of lead tungstate crystals with photodetectors attached are shown in Figure 2.13.

The EB covers the pseudorapidity range $|\eta| < 1.479$. The granularity of the crystals in the EB is 360-fold in ϕ and 2×85-fold in η, resulting in a total of 61,200 crystals in the EB. The crystals themselves are mounted in a quasi-projective geometry in order to avoid cracks with the same trajectory as incident particles from the IP. The crystals have a tapered, pyramidal shape: the front face has a cross section of 22×22 mm^2, the rear face has a cross



(a) (b)

Figure 2.13: Lead tungstate crystals with photodetectors attached. The crystal in Figure (a) is a barrel crystal with one face depolished and an APD capsule attached. An APD capsule with two APDs is shown in the inset. The crystal in Figure (b) is an endcap crystal with all sides polished and a VPT attached [67].

section of 26×26 mm^2, and the total length is 230 mm (25.8 X_0). The crystals are arranged into a thin-walled aveolar structure (submodule). Each submodule is placed into modules of different types, depending on pseudorapidity positioning. Each module contains 400 or 500 crystals. Four modules separated by 4 mm thick aluminum webs make a supermodule, which contains 1700 crystals. The entire EB is composed of 36 supermodules.

The EE covers the pseudorapidity range $1.479 < |\eta| < 3.0$, and the longitudinal distance between the IP and the EE is 3144 mm. All of the crystals are laid out in a rectangular $x - y$ grid with a quasi-projective geometry. As in the EB, the crystals have a tapered, pyramidal shape. In the case of the EE, the front face cross section is 28.62×28.62 mm^2, the read face cross section is 30×30 mm^2, and the total length is 220 mm (24.7 X_0). The EE crystals are all identically shaped and are arranged into groups of 5×5 crystals called supercrystals. Each endcap is made of two halves called Dees, which each have 138 full supercrystals and 18 partial supercrystals on the curved inner and outer edge of the Dee. There are 3662 crystals in each of the four Dees, making a total of 14,648 crystals in the entire EE.

The ES covers the pseudorapidity region $1.653 < |\eta| < 2.6$ and rests in front of the EE. Unlike the EB and EE, the ES is a sampling calorimeter made up of two layers of silicon

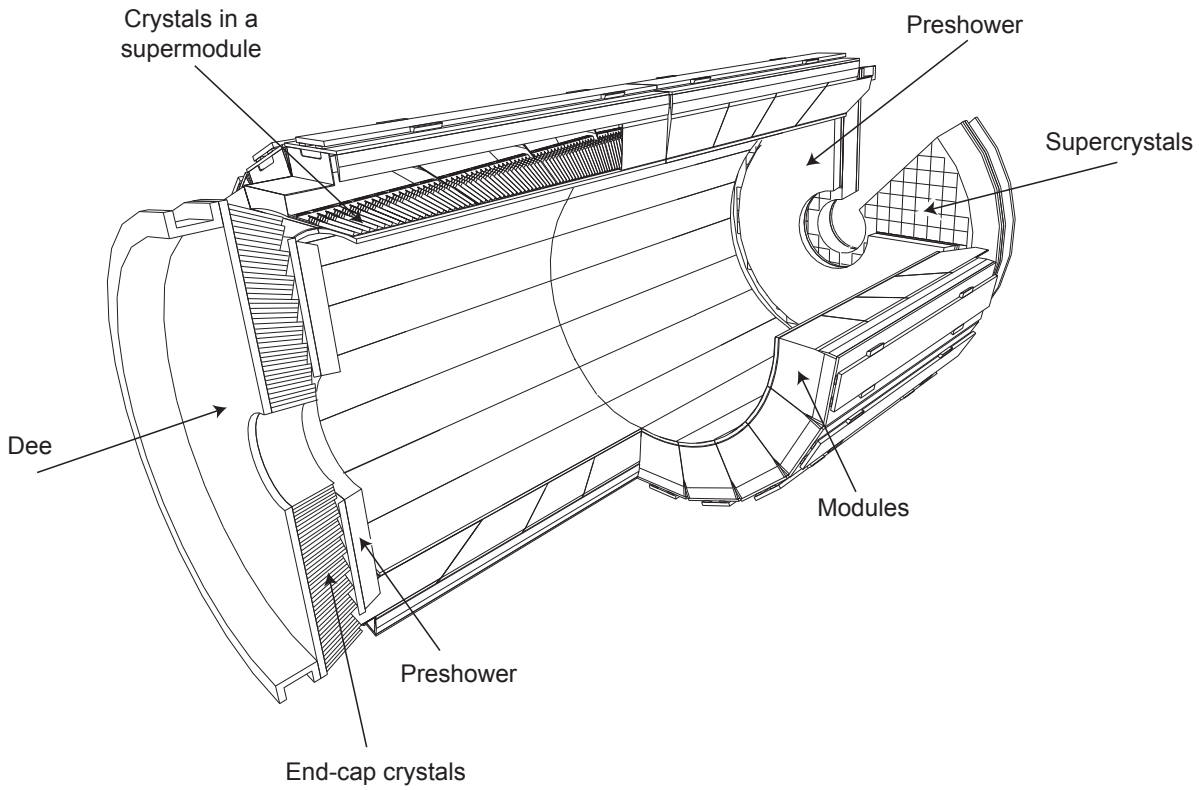


Figure 2.14: A three dimensional schematic of the CMS ECAL, showing the positioning of the modules, supermodules, supercrystals, and Dees as described in Section 2.2.4 [73].

strip sensors alternating with two layers of lead absorber. The total thickness is 20 cm. Each silicon sensor is 63×63 mm^2 with an active area of 61×61 mm^2 divided into 32 strips, each with a 1.9 mm pitch. Because the pitch of the ES sensors is so much thinner than the front face of the EE crystals, the ES has significantly better spatial resolution than the EE alone. This is necessary to distinguish between single photons and closely-spaced photon pairs (e.g. from $\pi^0 \to \gamma\gamma$ decays). The silicon strip sensors in the first layer are set orthogonally to the strips in the second layer. The first and second lead layers have material thicknesses of 2 X_0 and 1 X_0, respectively, so that 95% of the incident photons have begun to shower before reaching the second sensor layer. The ES is divided into four Dees (two Dees per endcap), and the ES Dees mirror the shape and orientation of the EE Dees. A schematic of the entire ECAL, including the modules, supermodules, supercrystals, and Dees as described above may be found in Figure 2.14.

For electromagnetic shower energies below 500 GeV, where shower leakage from the rear of the calorimeter is significant, the ECAL energy resolution (σ/E) for a given energy E may be parameterized by Equation 2.6:

$$\frac{\sigma(E)}{E} = \frac{S}{\sqrt{E}} \oplus \frac{N}{E} \oplus C \tag{2.6}$$

where S is the stochastic term, N is the noise term, and C is the constant term. The stochastic term describes event-to-event fluctuations in the lateral shower containment, photostatic effects, and fluctuations in the energy deposited in the ES absorber (in regions where the ES is present) with respect to the energy measured by the ES silicon. The noise term describes noise from the ECAL electronics, the digitization process, and pile-up. The constant term describes non-uniformity in the longitudinal light collection, intercalibration errors, and the leakage of energy from the back of the crystal. These parameters were determined in a test beam by measuring the energy of electrons from 20 to 250 GeV using a test setup consisting of an array of 3×3 crystals centered on a reference crystal. The ECAL energy resolution

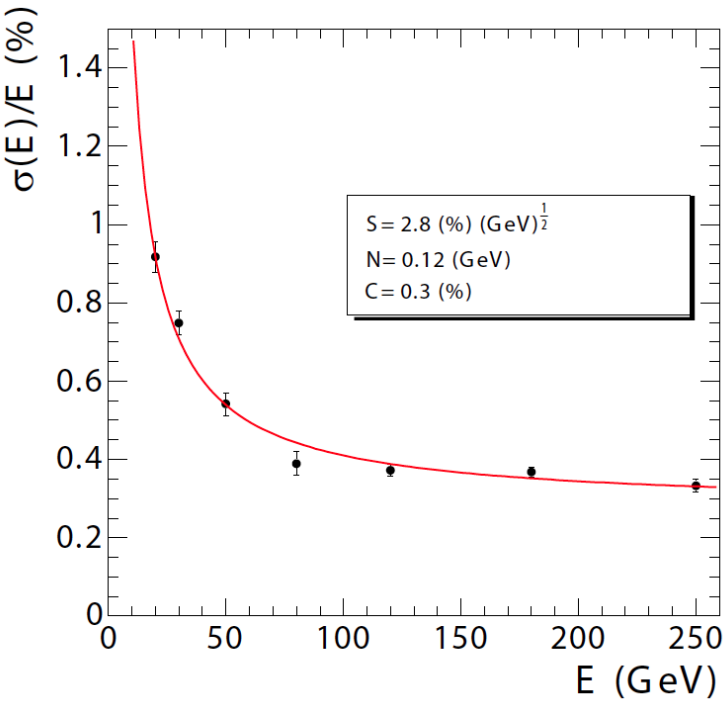


Figure 2.15: Energy resolution of the CMS ECAL as a function of electron energy as measured in a test beam. The fit (red) is parameterized by stochastic (S), noise (N), and constant (S) terms, as described in Equations 2.6 and 2.7 [67].

was found to be:

$$\frac{\sigma(E)}{E} = \frac{2.8\%\ \text{GeV}^{1/2}}{\sqrt{E}} \oplus \frac{0.12\ \text{GeV}}{E} \oplus 0.3\% \tag{2.7}$$

The results are shown in Figure 2.15. For electrons with large energy the dominant contribution comes from the constant term, which is sensitive to intercalibration errors. Calibration of the ECAL's crystals is critical, therefore, to its success. The ECAL energy resolution was measured in 2011 using collision events and found to agree with expectations within 1% in the EB and 3% in the EE [74], and a channel-to-channel calibration precision of 0.6% in the EB and 1.5% in the EE has been achieved [73].

2.2.5 Hadronic calorimeter

The CMS hadronic calorimeter (HCAL) is a hermetic sampling calorimeter made of alternating layers of plastic scintillator and brass absorber. Together with the electromagnetic calorimeter

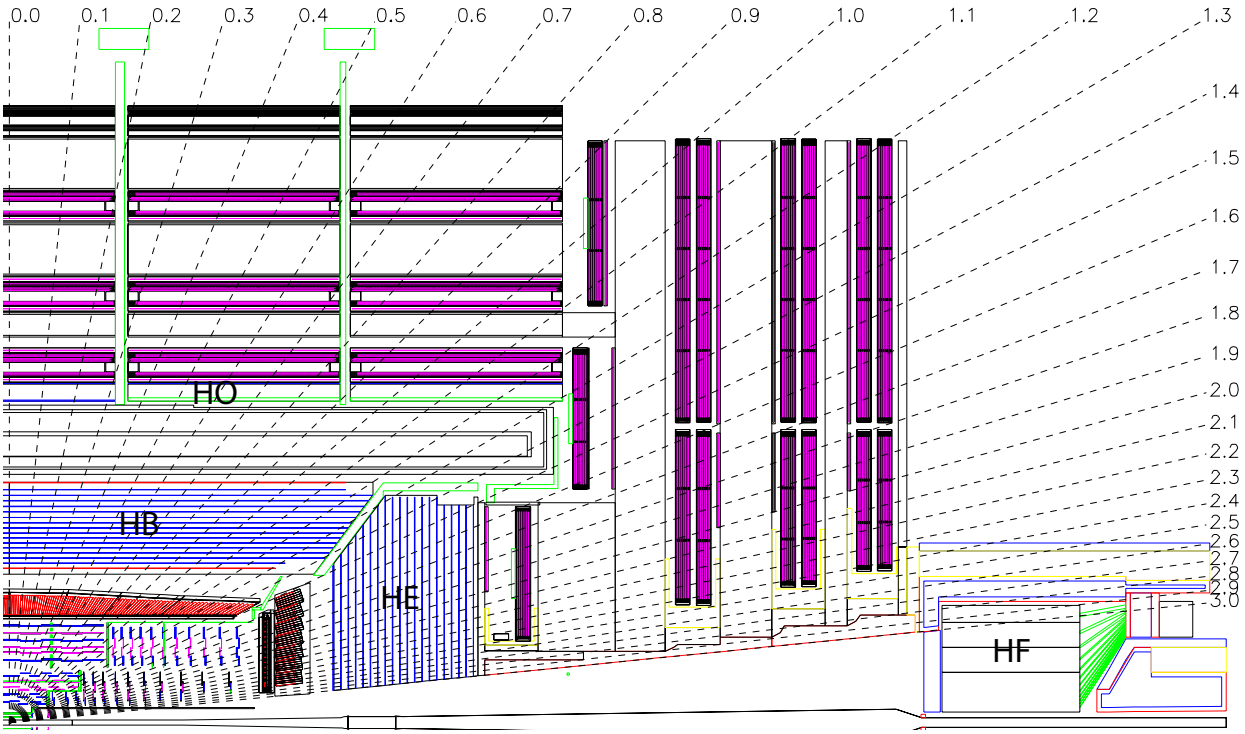


Figure 2.16: A schematic cross section of the CMS HCAL. Only one quarter of the entire HCAL is represented. The four subsystems of the HCAL: the barrel (HB), endcap (HE), outer calorimeter (HO), and forward calorimeter (HF) are all shown [67].

(ECAL), the HCAL forms a complete calorimetry system for the CMS detector, which allows for the reconstruction of hadronic jets and missing transverse energy. Like the ECAL, the HCAL is split into two subdetectors consisting of a barrel region (HB) and two endcap regions (HE). These subdetectors cover the pseudorapdity region $|\eta| < 3.0$. The HCAL is radially restricted by the inner radius of the solenoid (2.95 m from the IP) and the outer radius of the ECAL (1.77 m from the IP). This radial restriction also restricts the number of interaction lengths of material within the HCAL. Therefore, an additional layer of scintillator known as the outer hadronic calorimeter (HO) is placed outside the solenoid in the barrel region to serve as a "tail catcher." In addition, forward hadronic calorimeters (HF) made of radiation-hard quartz fiber and steel absorber are placed 11.2 m from the IP in the pseudorapidity region $3 < |\eta| < 5.2$. A schematic of the HCAL is shown in Figure 2.16.

In order to minimize non-Gaussian tails in the energy resolution, the HCAL must be hermetic and provide good containment of hadronic showers. Hence, the HB and HE absorber

Property	Value
Chemical composition	70% Cu, 30% Zn
Density	8.53 g/cm^3
Radiation length	1.49 cm
Interaction length	16.42 cm

Table 2.5: Properties of the HB and HE brass absorber: C26000 or "cartridge brass" [67].

has been chosen to maximize the number of interaction lengths of material within the CMS solenoid. C26000 "cartridge brass" not only satisfies this requirement, it is also nonmagnetic and composed of relatively low-Z elements (zinc and copper), which means that it will not significantly degrade the muon measurement. Some properties of C26000 brass are shown in Table 2.5.

Plastic scintillator was chosen as the active medium for the HB and HE in part for its minimal thickness, which allowed more material budget to be given to the absorber. The plastic scintillators are composed of tiles embedded with wavelength-shifting (WLS) fibers, which are spliced to high attenuation-length clear plastic fibers that carry the light to the readout electronics. The single innermost plastic scintillator layer of the HB and HE is made of Bicron BC408. This layer samples hadronic showers that begin within the inert material between the ECAL and the HCAL. All other plastic scintillator layers in both detectors are made from Kuraray SCSN81.

The HB covers the pseudorapidity region of $|\eta| < 1.4$. It is composed of 36 identical azimuthal wedges, which form two half-barrels (HB+ and HB-). Each wedge is separated into 16 towers in η and 4 towers in ϕ. This makes a total of 2304 towers in the HB, each of which covers $\Delta\eta \times \Delta\phi = 0.087 \times 5^\circ$. These wedges are composed of flat absorber plates that have been bolted together in layers with a staggered geometry so that there are no gaps with projective dead material, as shown in Figure 2.17. The innermost and outermost absorber layers are made of stainless steel for structural strength; all other layers are made of brass. The innermost stainless steel layer has a thickness of 40 mm. The outermost stainless steel layer has a thickness of 75 mm. The innermost 8 brass layers have a thickness of 50.5 mm,

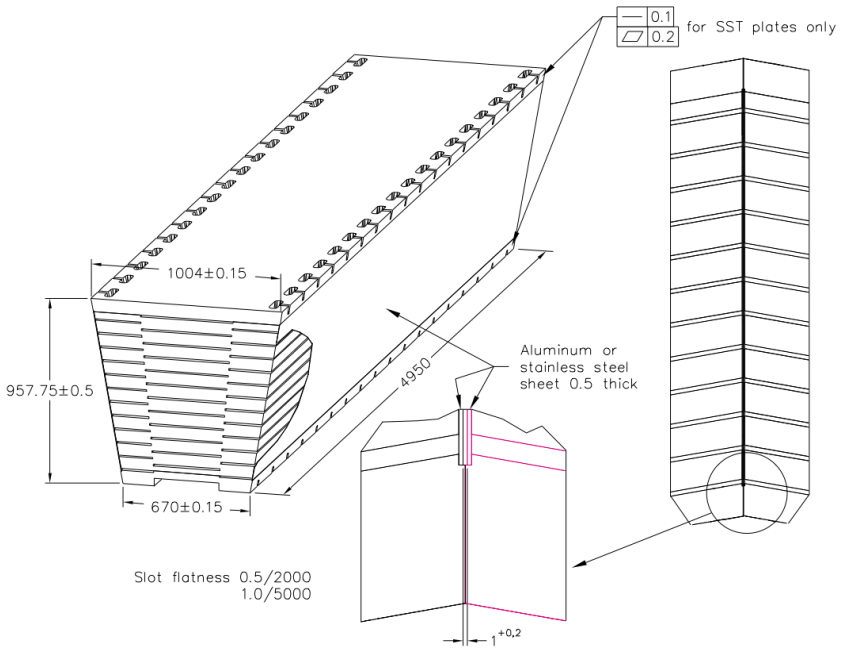


Figure 2.17: A schematic of an HB wedge showing the pattern with which the absorber plates are bolted together to ensure hermiticity. SST refers to the stainless steel plates at the top and bottom of each wedge [67].

and the outermost 14 layers have a thickness of 56.6 mm. The plates are bolted together in such a way as to leave slots for plastic scintillator between them. All of the scintillator for a given layer is grouped together into a single mechanical tray in order to reduce the number of mechanical components to be handled. Each scintillator tile is 3.7 mm thick, except for the innermost and outermost tiles, which are 9 mm thick. A single WLS fiber is embedded in each tray, which interfaces with a clear plastic fiber, which in turn interfaces with an optical cable. The optical cable carries light from the scintillators to an optical decoding unit, which collects the fibers into read-out towers and brings the light to a hybrid photodiode (HPD) . Each wedge has four HPDs (one for each ϕ segmentation), which are located in a readout box (RBX) on the side of the wedge farthest from the IP. All scintillator layers for a single tower are combined into a single longitudinal readout, except in the case of towers 15 and 16 in the most forward region of the HB. A schematic of the HCAL longitudinal segmentation can be seen in Figure 2.18. The material budget of the absorber is less than 6 interaction

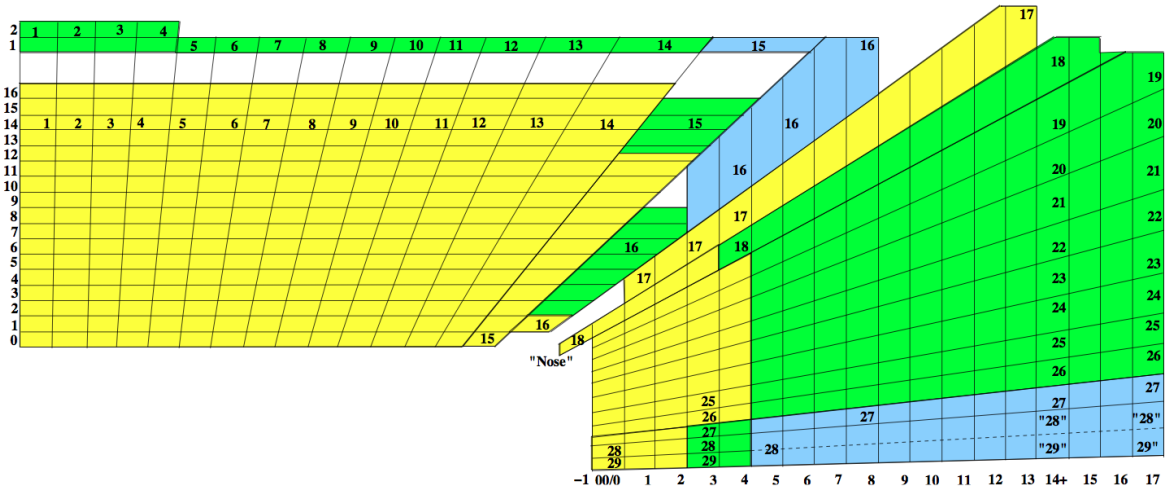


Figure 2.18: A schematic of the CMS HCAL showing the longitudinal depth segmentation of the tower readout. Only one fourth of the detector is shown. Each number outside the colored region on the white edge of the schematic corresponds to a scintillator layer. Each number within the colored area corresponds to a tower. Each tower is read out with either one longitudinal segmentation (only yellow), two longitudinal segmentations (yellow and green), or three longitudinal segmentations (yellow, green, and blue) [75].

lengths at $|\eta| = 0$ and about 10 interaction lengths at $|\eta| = 1.3$. The ECAL adds about 1.1 interaction lengths to this.

The HE covers the pseudorapidity region of $1.3 < |\eta| < 3.0$. It is designed to interlock neatly with the HB. There are a total of 1368 towers in the HE, and the five most central of these towers each covers $\Delta\eta \times \Delta\phi = 0.087 \times 5°$. The rest of the HE towers vary in $\Delta\eta \times \Delta\phi$ coverage, as shown in Table 2.6. The HB and the HE are very similar in design. Like the HB, the HE is made of interlocking C26000 brass absorber plates with plastic scintillator tiles that have been stacked to avoid projective dead material. Also like the HB, the outermost absorber layer is made of stainless steel rather than brass (the innermost layers are brass in the case of the HE, unlike the HB). In the case of the HE, the absorbers are 79 mm thick. In addition, the HE also uses WLS fibers, clear plastic fibers, and optical cable to route light from the plastic scintillators to an optical decoding unit, and HPDs within an RBX are used as the photodetectors. The total length of the HE, including the ECAL, is about 10 interaction lengths.

Tower index	η range Low	High	Detector	Size $\Delta\eta$	$\Delta\phi$	Depth segments
1	0.000	0.087	HB, HO	0.087	5°	HB = 1, HO = 1
2	0.087	0.174	HB, HO	0.087	5°	HB = 1, HO = 1
3	0.174	0.261	HB, HO	0.087	5°	HB = 1, HO = 1
4	0.261	0.348	HB, HO	0.087	5°	HB = 1, HO = 1
5	0.348	0.435	HB, HO	0.087	5°	HB = 1, HO = 1
6	0.435	0.522	HB, HO	0.087	5°	HB = 1, HO = 1
7	0.522	0.609	HB, HO	0.087	5°	HB = 1, HO = 1
8	0.609	0.696	HB, HO	0.087	5°	HB = 1, HO = 1
9	0.696	0.783	HB, HO	0.087	5°	HB = 1, HO = 1
10	0.783	0.879	HB, HO	0.087	5°	HB = 1, HO = 1
11	0.879	0.957	HB, HO	0.087	5°	HB = 1, HO = 1
12	0.957	1.044	HB, HO	0.087	5°	HB = 1, HO = 1
13	1.044	1.131	HB, HO	0.087	5°	HB = 1, HO = 1
14	1.131	1.218	HB, HO	0.087	5°	HB = 1, HO = 1
15	1.218	1.305	HB, HO	0.087	5°	HB = 2, HO = 1
16	1.305	1.392	HB, HE	0.087	5°	HB = 2, HE = 1
17	1.392	1.479	HE	0.087	5°	HE = 1
18	1.479	1.566	HE	0.087	5°	HE = 2
19	1.566	1.653	HE	0.087	5°	HE = 2
20	1.653	1.740	HE	0.087	5°	HE = 2
21	1.740	1.830	HE	0.090	10°	HE = 2
22	1.830	1.930	HE	0.100	10°	HE = 2
23	1.930	2.043	HE	0.113	10°	HE = 2
24	2.043	2.172	HE	0.129	10°	HE = 2
25	2.172	2.322	HE	0.150	10°	HE = 2
26	2.322	2.500	HE	0.178	10°	HE = 2
27	2.500	2.650	HE	0.150	10°	HE = 3
*28	2.650	3.000	HE	0.350	10°	HE = 3
29	2.853	2.964	HF	0.111	10°	HF = 2
30	2.964	3.139	HF	0.175	10°	HF = 2
31	3.139	3.314	HF	0.175	10°	HF = 2
32	3.314	3.489	HF	0.175	10°	HF = 2
33	3.489	3.664	HF	0.175	10°	HF = 2
34	3.664	3.839	HF	0.175	10°	HF = 2
35	3.839	4.013	HF	0.174	10°	HF = 2
36	4.013	4.191	HF	0.178	10°	HF = 2
37	4.191	4.363	HF	0.172	10°	HF = 2
38	4.363	4.538	HF	0.175	10°	HF = 2
39	4.538	4.716	HF	0.178	10°	HF = 2
40	4.716	4.889	HF	0.173	20°	HF = 2
41	4.889	5.191	HF	0.302	20°	HF = 2

Table 2.6: Table showing the segmentation of the CMS HCAL. The pseudorapidity range, detector, coverage in $\Delta\eta \times \Delta\phi$, and longitudinal depth segmentation details are all shown. Longitudinal depth segmentation can also be seen in Figure 2.18 [75].
* The first two depth segmentations have finer η segmentation, namely $2.650 - 2.868$ and $2.868 - 3.000$ and are referred to as tower indices 28 and 29.

The HO covers the pseudorapidity region of $|\eta| < 1.26$. Unlike the HB and HE subdetectors, the HO is composed only of scintillators and is placed outside of the solenoid bore. At $\eta = 0$, the HB and EB together only provide about seven interaction lengths. The HO, therefore, is designed to sample the energy from hadronic showers that penetrate both the HB and HE and leak out of the solenoid. By using the solenoid as an additional absorber, the HO in principle increases the effective thickness of the HCAL in the barrel region and reduces non-Gaussian tails in the energy resolution. The HO is located inside of the barrel muon system (Section 2.2.6), and so it is spatially restricted by that system. The HO is divided into five rings, which are labeled: -2, -1, 0, 1, and 2. Each of these rings is divided into 12 identical ϕ-sectors, and each ϕ-sector has 6 slices in ϕ. Ring 0 of the HO has two scintillator layers on either side of a 19.5 cm piece of iron at 3.82 m and 4.07 m from the IP. The other HO rings have only one HO layer at a distance of 4.07 m from the IP. Each HO layer is 40 mm thick, 16 mm of which is taken up by the scintillator, while the rest is taken up by an aluminum honeycomb support structure, and each ring takes up 2.54 m on the z-axis. Like the HB and HE, the HO scintillator is unified in tiles, and the size and shape of the HO tiles roughly maps to the HB towers. Also like the HB and HE, light from the HO scintillators is collected via WLS fibers, clear plastic fibers, and optical cable. HPDs within an RBX are used as the photodetectors.

The HF covers the pseudorapidity region of $3.0 < |\eta| < 5.0$ and is located 125 cm from the beam. This region is exposed to unprecedented particle fluxes, and successful operation relies strongly on radiation hardness. For this reason, unlike the HB, HE, and HO, the HF uses steel as an absorber and quartz fibers as an active medium (the signal being Cerenkov light emitted within the quartz fibers). In addition, light is collected using photomultiplier tubes (PMTs) rather than HPDs. The HF is essentially cyllindrical and made up of $18 \times 20^\circ$ steel wedges on either side of the IP. The cyllinder's outer radius is 130 cm and the inner radius is 12.5 cm. The steel absorber is 165cm deep, and the quartz fibers are 0.6 mm in diameter. The fibers run parallel to the beamline, and they are bundled together to form a

total of 864 towers (without pointing geometry). Half of the fibers are "long" and run the full length of the HF detector, while the other half are "short" and start 22 cm from the front face of the HF detector. This allows for the discrimination between electromagnetic showers, which are generally absorbed before reaching the short fibers, and hadronic showers, which interact with both long and short fibers.

The HCAL energy resolution (σ/E) for a pion of given energy E may be parameterized by Equation 2.8:

$$\frac{\sigma(E)}{E} = \frac{S}{\sqrt{E}} \oplus C \tag{2.8}$$

where S is the stochastic term and C is the constant term. These terms have been measured in test beams using a 5×5 HB tower array and a pion beam, and the results are given in Equation 2.9:

$$\frac{\sigma(E)}{E} = \frac{115.3\%\ \text{GeV}^{1/2}}{\sqrt{E}} \oplus 5.5\% \tag{2.9}$$

The results are shown graphically in Figure 2.19. Similar results were obtained with the HF using pion and electron test beams, where $S = 198\%$, $C = 9\%$ for electromagnetic shower energy resolution and $S = 280\%$, $C = 11\%$ for hadronic shower energy resolution [67, 72, 76].

2.2.6 Muon system

The purpose of the muon system in the CMS experiment is to identify and reconstruct the trajectories of muons. These reconstructed trajectories are used both for triggering (Section 2.2.8) and for offline analyssis. Three types of gaseous detectors are used to measure muons in the CMS detector. Since muons are minimum ionizing particles (MIPs) and do not generally interact with the inner regions of the detector, muon systems are typically placed outside of the tracking and calorimetry systems in most particle detectors. In the case of CMS, the muon system rests outside of the solenoid and covers the pseudorapidity region $|\eta| < 2.4$. A plot showing the number of interaction lengths as a function of pseudorapidity is given in Figure 2.20. This region includes a wide range of incident muon rates, neutron-induced

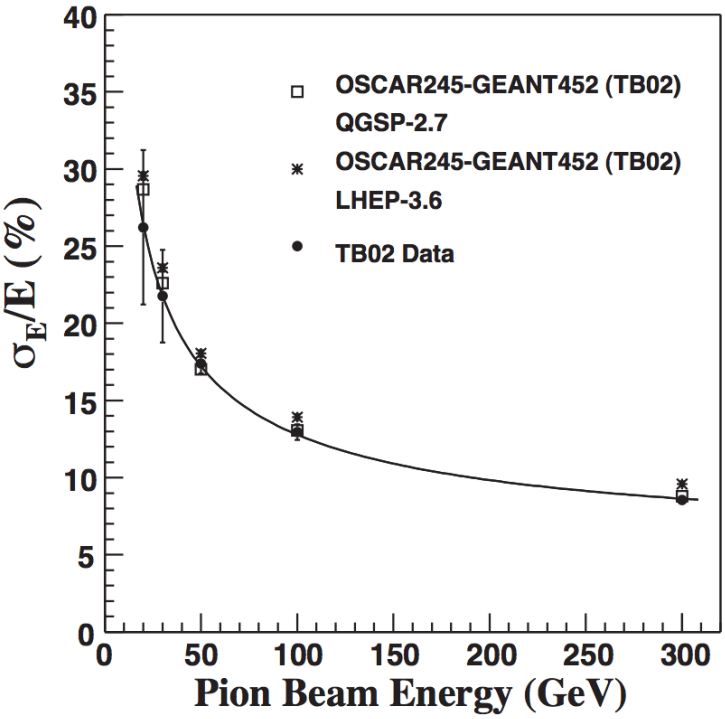


Figure 2.19: Energy resolution of the HB as a function of pion energy as measured in a test beam [77].

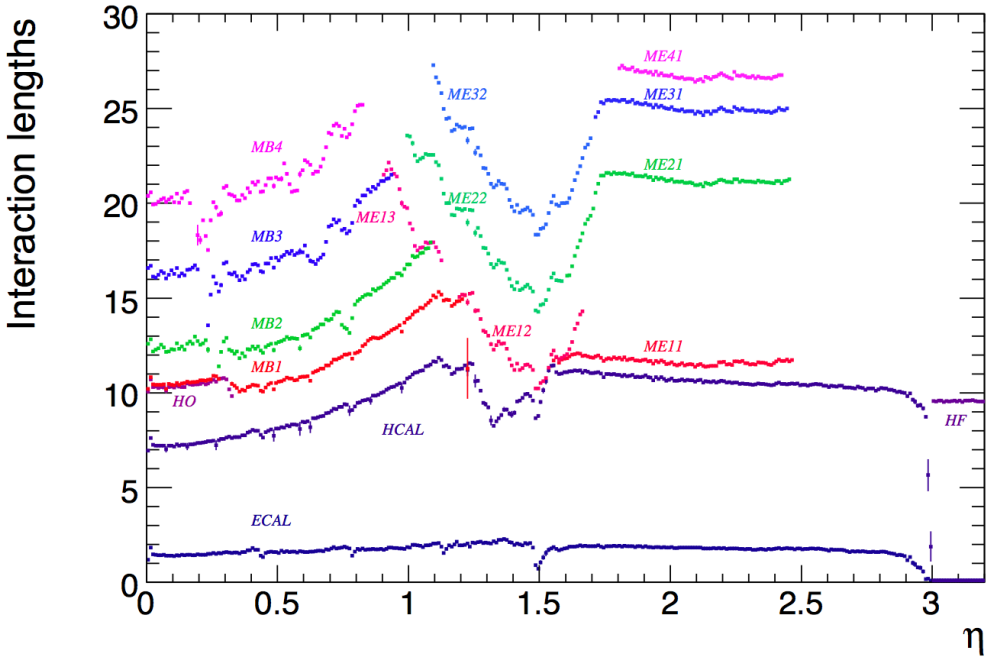


Figure 2.20: Material thickness of the CMS detector in interaction legnths as a function of pseudorapidity. All muon barrel and endcap stations are labeled [67].

.

background rates, and magnetic field strengths. The choice of which detector technologies to use is strongly dependent on these conditions. In the barrel ($|\eta| < 1.2$), where the incident muon rate, the neutron-induced background rate, and the residual magnetic field are relatively low, drift tube chambers (DTs) with rectangular drift cells are used. In the endcaps, where the incident muon rate, the neutron induced background rate, and the residual magnetic field are relatively high, cathode strip chambers (CSCs) are used. In addition, resistive plate chambers (RPCs) are used in both the barrel and the endcap ($|\eta| < 1.6$). RPCs have a faster response and a better time response than the DTs and CSCs, but they have a coarser position resolution. This allows them to unambiguously identify the bunch crossing associated with a reconstructed muon. In total, the muon system is made up of 25,000 m^2 of active detection area and nearly 1 million electronic channels. The layout of the CMS muon system is shown in a schematic in Figure 2.21.

The muon system in the barrel region is organized into four layers or "stations" which are separated by layers of the iron return yoke. These layers are located at radii of approximately 4.0, 4.9, 5.9, and 7.0m from the IP. The stations are labeled MB1-4, with MB1 located closest to the IP. The stations are further divided into five "wheels", which are labeled from YB-2 (farthest in the z-minus direction) to YB+2 (farthest in the z-plus direction), following the segmentation of the iron return yoke. Each of these 5 wheels is split in ϕ into 12 sectors, and each sector covers 30° of azimuthal angle. Sectors in different stations are staggered so that a high-p_{T}muon near a sector boundary would interact with at least 3 of the 4 stations. Stations MB1 and MB2 are made up of packages consisting of one DT chamber placed between two RPCs. Stations MB3 and MB4 are made up of packages consisting of one DT chamber placed with a layer of 1, 2, or 4 RPCs. Stations MB1, MB2, and MB3 provide measurements on muon coordinates in the $r - \phi$ plane and on the $z-$axis, while MB4 only provides measurements in the $r - \phi$ plane. These measurements allow a muon track to be reconstructed with a ϕ positional precision better than 100 μm and a directional precision better than 1 mrad.

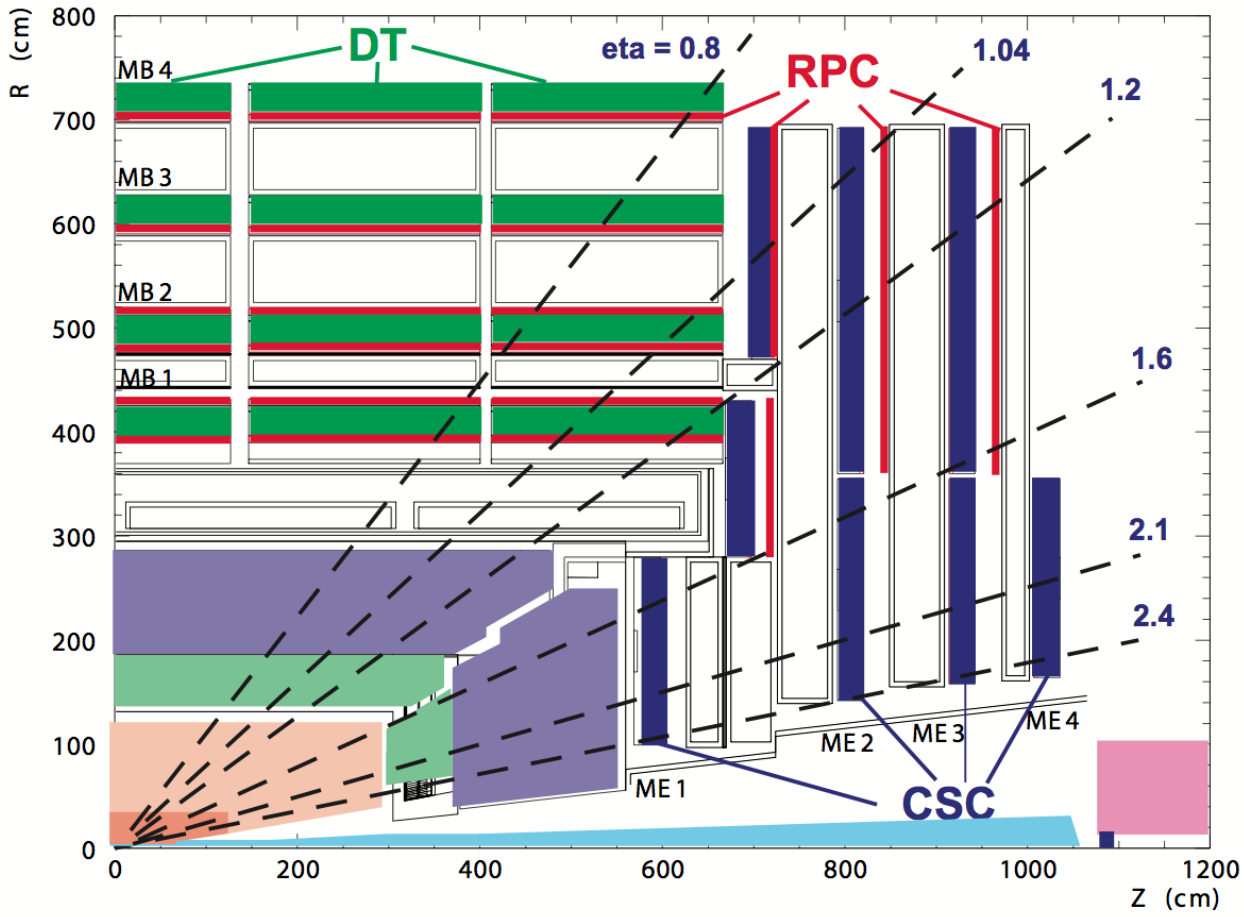


Figure 2.21: A schematic cross section of the CMS muon system. Only one quarter of the entire muon system is represented. The numbering scheme for the various stations (MB1-4 in the barrel, ME1-4 in the endcap) is clearly labeled. The three subsystems: the drift tube chambers (DTs), resistive plate chambers (RPCs), and cathode strip chambers (CSCs) are also shown [72].

The muon system in the endcap region is arranged into 4 stations for each of the 2 endcaps. These stations are labeled ME1-4, where ME1 is the closest to the IP. The stations are each divided into three rings, labeled 1-3, where ring 1 is closest to the beamline. Each individual CSC is trapezoidal in shape and contains 6 gas gaps. Each gap contains radial cathode strips and a set of anode wires that run perpendicularly to the strips. As in the barrel, the CSCs are overlapped to avoid gaps of uninstrumented areas. There are 36 CSCs in each ring of the muon endcap system, except for the innermost ring in stations ME2, ME3, and ME4, where there are 18 CSCs. In addition, only the innermost ring of ME4 has been installed, and 5 spare CSCs were installed on the outermost ring of ME4 on the $z-$plus side. There are a total of 468 CSCs in all of CMS, excluding the extras in ME4/2. Each CSC measures up to 6 spatial coordinates (r, ϕ, z), one for each of its layers. 36 RPCs in total are mounted only in the 2 outer rings of each of the endcap stations. The endcap muon system allows a muon track to be reconstructed with a ϕ positional precision of around 200 μm and a directional precision of about 10 mrad [67, 72].

2.2.7 Luminosity measurement

As discussed in Section 2.1.2, instantaneous luminosity, $\mathcal{L}$, is the ratio between the production rate in Hz for given physics process in a collider and the production cross section for that process. Integrated luminosity, $\mathcal{L}_{int}$, refers to an integral of instantanenous luminosity over a period of time. Instantaneous luminosity depends only on beam parameters, and it is expressed in units of $\mathrm{cm}^{-2}\mathrm{s}^{-1}$. In CMS, instantaneous luminosity is measured with the pixel tracker (see Section 2.2.3), using a method based on pixel cluster counting (PCC). The following description of the PCC method is taken from Reference [78].

In the PCC method, an effective cross section for the formation of pixel clusters is determined using results from a Van der Meer scan. That cross section, σ_{pixel}, is used to determine the instantaneous luminosity for each 23.3 second luminosity section of the 2011 physics sample. The high voltage of the pixel tracker is only turned on during stable collisions,

and pixel data is not available when the CMS data acquisition framework is not running. For these reasons, the PCC method is not useful for online luminosity monitoring. The PCC method is, however, used for all integrated luminosity measurements mentioned in this thesis.

There are roughly 66 million pixels in the pixel tracker. About 200 pixel clusters are formed in each minimum bias interaction, and there are about 5 pixels per pixel cluster. This means that even at high luminosities, the occupation of the pixel detector is on the order of 1/1000, and the probability of a given pixel being hit by two different tracks is very low. As a result, the number of pixel clusters per bunch crossing is expected to be very linear with the number of interactions per bunch cross and therefore a good measure of instantaneous luminosity.

Within the LHC, each bunch crossing produces some number of *pp* interactions, which in turn produce some number of pixel clusters. If these interactions are recorded using zero-bias triggers, then the mean number of pixel clusters per trigger is expressed by Equation 2.10:

$$\langle N_{\text{cluster}} \rangle = \left\langle N_{\text{cluster/interaction}} \right\rangle \langle N_{\text{interactions}} \rangle \equiv \left\langle N_{\text{cluster/interaction}} \right\rangle \mu \tag{2.10}$$

where $\langle N_{\text{cluster}} \rangle$ is the average number of pixel clusters formed per bunch crossing, $\mu = \langle N_{\text{interactions}} \rangle$ is the average number of *pp* interactions in a bunch crossing, and $\left\langle N_{\text{cluster/interaction}} \right\rangle$ is the average number pixel clusters formed per *pp* interaction.

μ, the average number of *pp* interactions in a bunch crossing, may also be expressed as a function of the *pp* interaction cross section, $\sigma_{\text{interaction}}$, the per-bunch instantaneous luminosity, $\mathcal{L}$, and the LHC orbital frequency, f, using Equation 2.11:

$$\mu = \frac{\sigma_{\text{interaction}}}{f} \mathcal{L} \tag{2.11}$$

The LHC orbital frequency, f, is shown in Section 2.1.2 to be 11.246 kHz.

Equations 2.10 and 2.11 may be combined as follows in Equation 2.12:

$$\langle N_{\text{cluster}} \rangle = \left\langle N_{\text{cluster/interaction}} \right\rangle \frac{\sigma_{\text{interaction}}}{f} \mathcal{L} \tag{2.12}$$

If σ_{pixel} is defined to be equal to the product of $\sigma_{\text{interaction}}$ and $\left\langle N_{\text{cluster/interaction}} \right\rangle$, then Equation 2.12 may be rewritten as:

$$\langle N_{\text{cluster}} \rangle = \frac{\sigma_{\text{pixel}}}{f} \mathcal{L} \tag{2.13}$$

This allows σ_{pixel} to be expressed in terms of $\langle N_{\text{cluster}} \rangle$, f, and $\mathcal{L}$: all of which are measureable during a Van der Meer scan, as was documented in Reference [78]. The relevant relationship is given by Equation 2.14:

$$\sigma_{\text{pixel}} = \langle N_{\text{cluster}} \rangle \, f \mathcal{L}^{-1} \tag{2.14}$$

The total systematic uncertainty on the luminosity measurement is 2.2%. The dominant contributions to this uncertainty are "afterglow" (energy from late-arriving particles and from activated detector material) and variations between Van der Meer scans.

2.2.8 Trigger

The LHC beam crossing interval of 25 ns corresponds to a crossing frequency of 40 MHz. In addition, the nominal total size in memory required to store all of the readout for a single CMS event is about 1.5 MB. This corresponds to a data production rate of 40 TB per second, which is far too much to store and process. To resolve this issue, CMS uses an online trigger system to select only approximately one out of every 10^6 events for storage and processing. This selection is divided between two components: the Level 1 Trigger (L1), which reduces the event rate to a maximum of 100 kHz, and the High Level Trigger (HLT) which further reduces the event rate to $\mathcal{O}(100$ Hz). The final rate is meant to allow as inclusive a selection as possible while remaining within the limits of data recording technology. For comparison purposes, the production rates in Hz of various processes (including scalar leptoquarks)

assuming nominal beam energy and instantaneous luminosity are shown in Figure 2.22.

The L1 is composed of custom-designed, largely programmable electronics housed partly on the detector itself and partly in a counting room 90m away from the detector. The hardware is composed of field programmable gate arrays (FPGAs) where possible, but application-specific integrated circuits (ASICs) and largely programmable look up tables (LUTs) are used where speed, density, or radiation hardness is required. The L1 has access to coarsely segmented data from the calorimeters and the muon systems. This data is processed in three components: local, regional, and global. The local component is made up of Trigger Primitive Generators (TPGs) which collect information on energy deposits in the calorimeters and track segments or hit patterns in the muon chambers. The regional component combines the local trigger primitives and uses pattern logic to rank and sort the resulting trigger objects (e.g. muon, electron, and jet candidates in limited regions of the detector). In this context, "rank" is a function of the measured energy or momentum of a trigger object and the uncertainty associated with that measurement. The global component consists of the Global Calorimeter Trigger (GCT) and the Global Muon Trigger (GMT). The GCT and GMT collect the highest rank calorimeter and muon objects and transfer them to the Global Trigger (GT), where the decision to accept the event and pass it on to the HLT is made. This decision is based on detector algorithms and on the readiness of the data acquisition framework (DAQ), as determined by the Trigger Control System (TCS). The L1 Accept decision (L1A) is passed from the GT to the detector subsystems via the Timing, Trigger, and Control system (TTC). The allowed latency for the L1 between a bunch crossing and the distribution of the trigger decision to the detector electronics is 3.2 μs. During this latency period, the full resolution data from the detector is stored in pipeline memories in the front end electronics. A logical schematic of the L1 architecture is shown in Figure 2.23.

The HLT [79] is software system implemented on a farm of commercially available CPUs. Unlike the L1, the HLT has access to the full detector readout and can make complex calculations comparable to those available in a fully offline analysis. Furthermore, the CMS

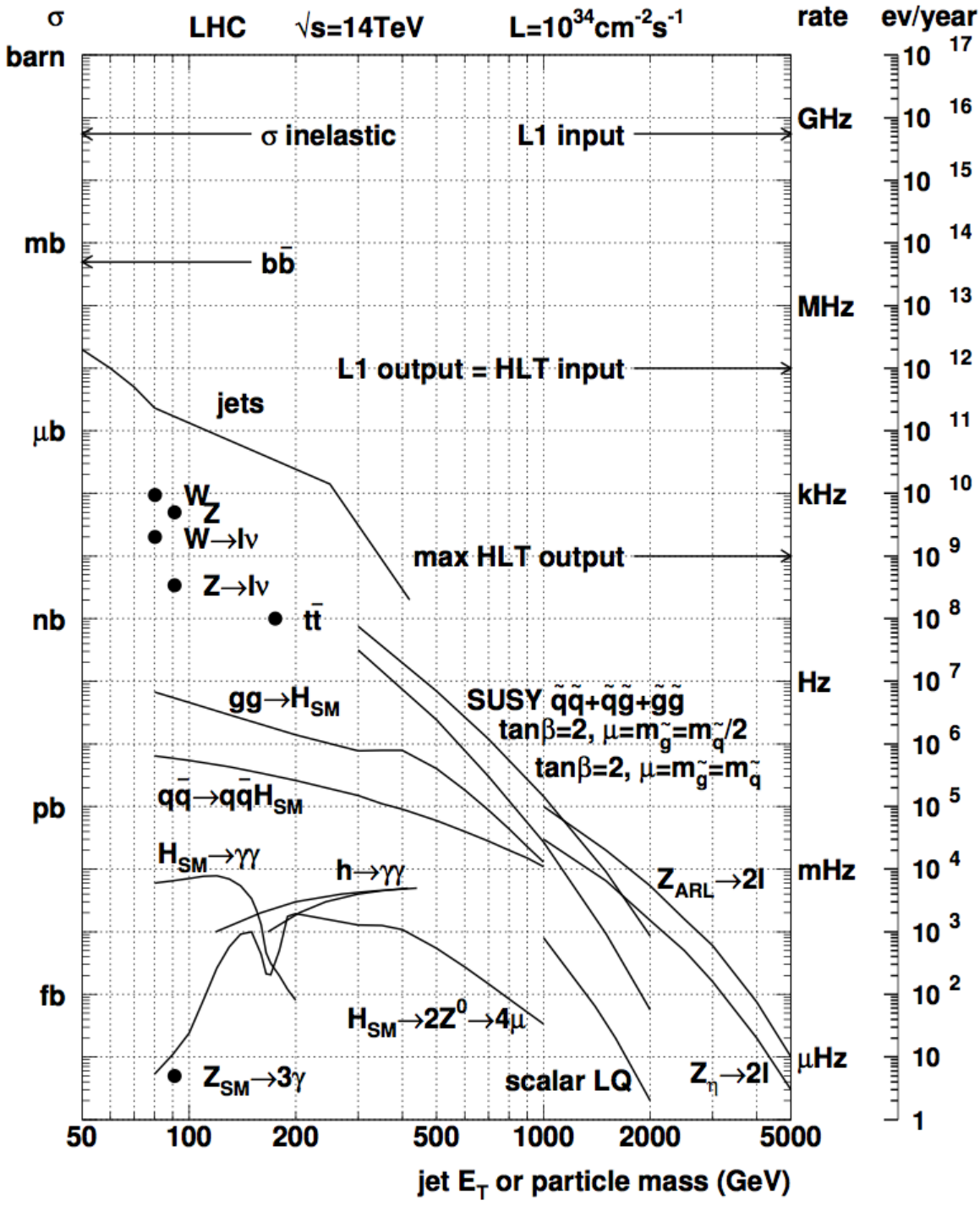


Figure 2.22: Production rates and cross sections associated with various processes (including scalar leptoquarks) in the LHC at design beam energy and luminosity. For comparison purposes, the L1 and HLT input and output rates are shown. It should be noted that the maximum HLT output rate quoted in this diagram (100 Hz) is not a hard maximum. Indeed, the HLT maximum output rate was around 300 Hz during stable beams in 2011 running [67].

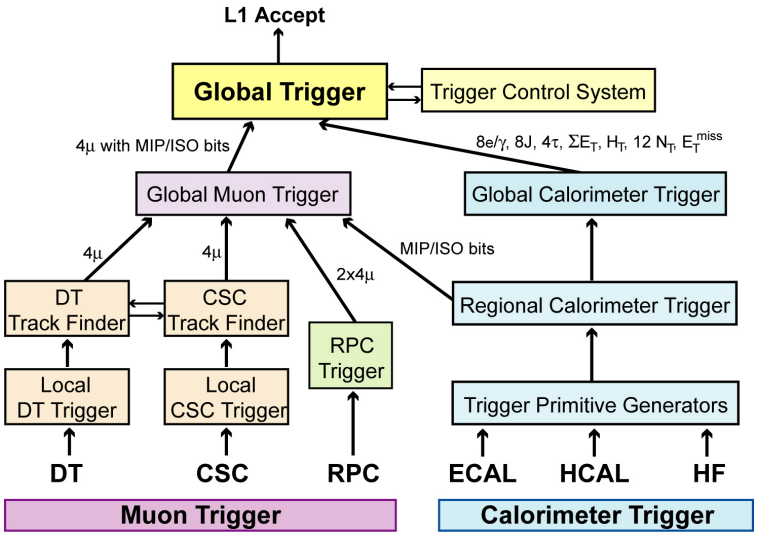


Figure 2.23: Logical schematic of the architecture of the Level 1 Trigger (L1) [67]

software framework (CMSSW) used by the HLT is essentially the same framework used for offline analysis. The HLT selects events using a list of triggers which are written as software algorithms. A common HLT algorithm is to select an event with a trigger object above a given transverse momemntum or energy threshold. These trigger objects may include single objects, for example: electrons, muons, taus, jets, or photons. They may also include composite objects, for example: $\not{E}_{\mathrm{T}}$, or dilepton invariant mass. The HLT may also pass through some L1 decisions for the purpose of detector study or calibration. The full list of triggers is refered to as the HLT "trigger menu". Events accepted by a trigger are sorted into primary datasets (PDs) and written to storage. Some triggers are configured to accept every event that passes their selection requirements. Other triggers only take one out of every N events that pass. The former triggers are known as "unprescaled" triggers. The latter triggers are known as "prescaled" triggers, and N is known as the "trigger prescale". Triggers with similar selection topologies are grouped together and their outputs are written to the same PD, in order to minimize the overlap between various PDs. Because the HLT is entirely software-driven, the trigger menu is extremely flexible and has the ability to change with evolving conditions at the LHC [67, 72].

CHAPTER 3

EVENT RECONSTRUCTION

Event reconstruction refers to the operation of constructing physics objects from the raw data collected by the detector. This operation is essential to all particle physics experiments, including CMS. Different particles interact with various subsystems within the CMS detector in different ways, as shown in the cartoon in Figure 3.1. The event reconstruction operation interprets and combines digitized output from all of these subsystems in order to create a complete picture of the event. Reconstruction in CMS is divided into three separate processes: local reconstruction within a local subdetector module, global reconstruction within a whole subdetector, and higher-level reconstruction which combines reconstructed objects to make higher-level objects.

The local reconstruction process takes digitized electronic signals called "Digis" as its input. These Digis may come from the real detector electronics or from a simulation of those electronics. The same local reconstruction algorithms are used in either case, and the output objects are called "RecHits". The global reconstruction process combines RecHits from various modules of a single subdetector, but it does not combine information from modules from different subdetectors. To use muons as an example: tracker tracks are created from tracker RecHits, and standalone muon tracks are created from muon system RecHits, but

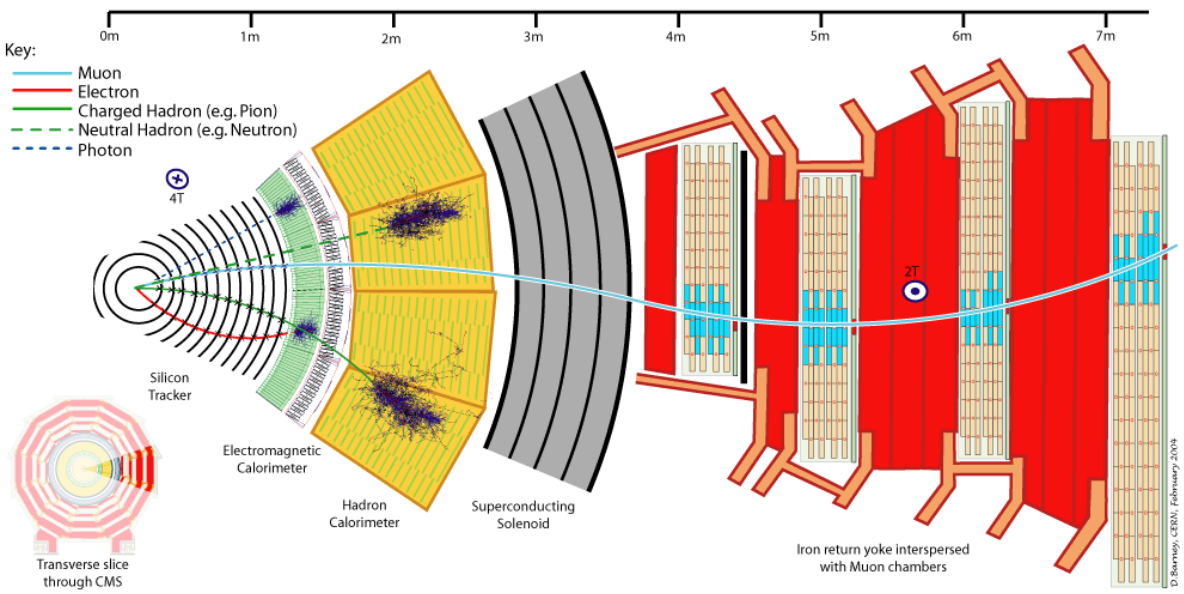


Figure 3.1: A cartoon schematic of the CMS detector, shown in the $r - \phi$ plane. The ideal interactions of muons, electrons, pions, neutrons, and photons with various components of the detector are shown.

global muon candidates are not yet formed. Finally, the higher-level reconstruction process combines RecHits from different subdetectors to produce higher-level objects, including global muon candidates.

All reconstruction processes are performed using the CMS software framework and the ROOT software package [80–82], which are based on the C++ and Python programming languages. As mentioned in Section 2.2.8, this is the same software framework used by the HLT. The reconstruction processes for various physics objects are described in greater detail in the following subsections.

3.1 Track reconstruction

Reconstruted charged particle trajectories or "tracks" are formed using information from the pixel tracker and the silicon strip tracker. Because the track reconstruction process begins before any primary decay vertices have been reconstructed, it is dependent on a precise estimate of the "beamspot", i.e. the location of the interaction point in the transverse plane,

which is calculated using a beamspot fitter [83].

Once the beamspot position has been estimated, track reconstruction proceeds using an iterative process, which is performed by the combinatorial track finder (CTF) [84–86]. Each iteration begins reconstructing tracks using "seeds", which may be thought of as starting points for initial estimates of fully reconstructed particle tracks. Seeds are either triplets of hits in the tracker or pairs of hits with the beamspot or a pixel vertex used as an additional constraint. These seeds yield a preliminary estimate of the particle track with an associated uncertainty, and they are then propagated outward in a search for additional hits to associate with the track. Once found, compatible hits are added to the track, and the track's parameters and associated uncertainties are recalculated. In each iteration, the outward search for hits continues until either the boundary of the tracker has been reached or no additional hits can be found. The search is then repeated by starting from the outermost hits and propagating them inward. The final step of each iteration is to fit the hits to obtain an estimate of the track's parameters. Between iterations, hits that are unambiguously associated with a track are removed from consideration for searches for additional hits. This creates a smaller collection of hits for the next iteration to consider. After each iteration, reconstructed tracks are fitted to remove tracks that are likely fakes and to assess the quality of the remaining tracks. This filtering for likely fakes uses the number of hits, the normalized χ^2 of the track, and the compatibility of the track originating from a pixel vertex. Tracks that pass the tightest filtering selection are labeled "highPurity".

The full track reconstruction process uses a total of six iterations. The difference between each iteration has largely to do with its seeds. Seeds for the first two iterations consist only of triplets of pixel tracker hits and pairs of pixel tracker hits with pixel vertices and the beamspot as an additional constraint. These first two iterations find prompt tracks with $p_{\mathrm{T}} > 0.9$ GeV. The third iteration uses pixel triplets as seeds to reconstruct tracks without a p_{T} requirement. The fourth iteration uses both pixel and strip layers as seeds, which allows the reconstruction algorithm to reconsrtuct displaced tracks. The fifth and sixth iterations

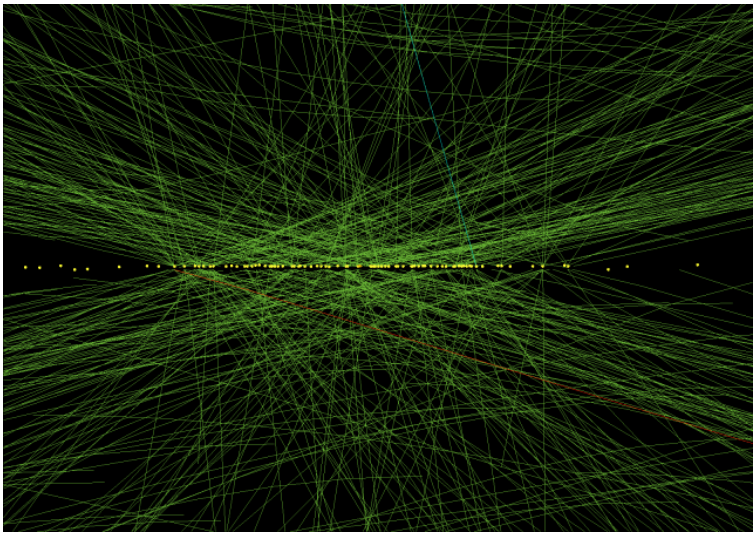


Figure 3.2: An event display showing 78 vertices that were reconstructed using information from the CMS tracker. The yellow dots correspond to reconstructed vertices, and the green lines correspond to reconstructed tracks. The data used to make this figure was collected on July 10, 2012 using *pp* collisions with a center-of-mass energy of 8 TeV [87].

use strip pairs as seeds to reconstruct tracks that do not have pixel hits [84].

3.2 Primary vertex reconstruction

Primary vertex reconstruction takes place once the prompt tracks in the event have been reconstructed. Prompt tracks from the primary interaction region are selected according to their impact parameter significance ($\mathrm{IP}/\sigma_{\mathrm{IP}}$) with respect to the beamspot, number of hits in the pixel and silicon strip trackers, and normalized χ^2 (χ^2/ndof). It is noteworthy that there is no requirement on track p_{T}, in order to ensure a high reconstruction efficiency. Once these tracks have been selected, they are clustered according to the z coordinate at the point of their closest approach to the beamspot. The clustering requires each track to have a separation in the z coordinate (z_{sep}) of less than 1 cm from its nearest neighbor. Primary vertex candidates that have more than one track are fit with an adaptive vertex fit [88]. In this adaptive vertex fit, each track associated with a vertex is given a weight, w, between 0 and 1, based on its compatibility with the vertex. Tracks compatible with a given vertex

have a weight that is close to 1. The variable "number of degrees of freedom" is defined as $n_{\text{dof}} = 2\sum_{\text{tracks}}(w_i - 3)$, where w_i is the weight of the i^{th} track. Number of degrees of freedom is highly correlated with the number of tracks associated with a given vertex, and this varible is used to identify real pp interaction vertices [89]. An event display showing 78 reconstructed vertices in the $r - z$ plane is provided in Figure 3.2. This display was made using data collected on July 10, 2012.

3.3 Electron reconstruction

Electrons in CMS are reconstructed using input from two detectors: the inner tracker, which reconstructs electron tracks, and the ECAL, into which electrons deposit most of their energy. The primary challenge in reconstructing electrons stems from the fact that electrons emit a large fraction of their total energy via Bremsstrahlung radiation while traversing the inner tracker. If an electron undergoes Bremsstrahlung radiation within the tracker, the electron trajectory may be significantly altered. These deviations must be accounted for by both track and energy reconstruction algorithms.

Just as there are two complimentary subdetectors with which to measure electrons in the CMS detector, CMS uses two complimentary algorithms to seed the track reconstruction process: tracker-driven seeding and ECAL-driven seeding. It is possible for an electron to be reconstructed using both algorithms. The tracker-driven seeding algorithm performs best for electrons that have low p_{T} or are buried inside jets. The ECAL-driven seeding algorithm is optimized for isolated electrons down to $p_{\text{T}} \simeq 5$ GeV, which (in principle) includes electrons from leptoquark decays. Electrons radiating within the tracker tend to deposit their energy within several ECAL crystals. Approximately 94% of the incident electron energy is contained in 3×3 crystals, and 97% of the incident electron energy is contained within 5×5 crystals [72]. In addition, electron energy tends to be deposited in the ECAL in a narrow region in pseudorapidity and a longer region in ϕ, due to the electron's trajectory within the magnetic

field. To capture and reconstruct as much of this energy as possible without using input from empty or unnecessary crystals, the electron reconstruction algorithm makes use of a "supercluster" pattern. A supercluster is a collection of one or more clusters of energy deposits in the ECAL within a narrow region in pseudorapidity and a longer region in ϕ. The ECAL-driven seeding algorithm begins by considering superclusters with transverse energy greater than 4 GeV and a ratio of hadronic energy from behind the supercluster over the supercluster energy of $H/E < 0.15$. As a first filtering step, these superclusters are required to be matched to track seeds, which are composed of pairs or triplets of hits in the inner track layers. The electron trajectory is reconstructed using a dedicated modeling of the electron energy loss via Bremsstrahlung radiation and fitted with a Gaussian Sum Filter (GSF) [90]. The final electron energy measurement may be taken from the GSF track, the supercluster energy, or some combination of the two [91–93].

The first filtering step is complimented by a preselection, which varies depending on the seeding algorithm used. In the case of the tracker-driven seeding, the preselection is based on a multivariate analysis [94]. In the case of the ECAL-driven seeding, the preselection is based on matching the supercluster to the GSF track in η and ϕ [91].

3.4 Muon reconstruction

Muons produced in pp collisions in CMS are measured by reconstructing their trajectories to form "tracks". Muon tracks are reconstructed independently by the inner tracker and by the muon subsystem, and those two independent measurements are then combined. The tracks reconstructed by the inner tracker system alone are called "tracker tracks". The tracks reconstructed by the muon subsystem alone are called "standalone muon tracks". There are two reconstruction methods using these objects as inputs: an outside-in method, which produces"global muons", and an inside-out method, which produces "tracker muons".

The global muon reconstruction method begins with the collection of all standalone muons

from the muon subsystem. For each standalone muon, the algorithm searches for a matching tracker track. If a matching tracker track is found, a global muon track is created and refit using all of the hits from both the standalone muon and the tracker track. For muons with large p_{T}, the global muon fit can improve the muon momentum resolution over a tracker-only fit.

The tracker muon reconstruction method begins with the collection of all tracker tracks with $p_{\mathrm{T}} > 0.5$ GeVand $p > 2.5$ GeV. For each of these tracker tracks, the algorithm searches for a matching standalone muon track, taking expected energy loss and uncertainty from multiple scattering into account. If one muon segment (a short track stub made of DT or CSC hits) matches the position predicted by the tracker track, the tracker track qualifies as a tracker muon track. The tracker muon reconstruction method is more efficient for low momentum ($p < 5$ GeV) muons, since it requires only one muon segment in the muon system, while the global muon method generally only becomes efficient with two or more segments.

A third category of muons is added as a catch-all. "Standalone muons" account for only 1% of muons from collisions, and they consist of standalone muon tracks that are rejected by both the global muon and tracker muon reconstruction methods. These muons are uncommon, because the tracker and global muon reconstruction methods have such a high efficiency. They also have a non-negligable fake rate from cosmic muons [95].

3.5 Particle flow event reconstruction

Particle flow (PF) event reconstruction aims to reconstruct and identify all stable particles in an event. In this context, particles are considered stable if they have a mean lifetime comparable to that of the charged pion ($\langle\tau\rangle = 10^{-8}$ s). These stable particles include electrons, muons, photons, charged hadrons, and neutral hadrons. The reconstruction process begins by identifying fundamental "elements": charged particle tracks and calorimetric energy clusters, which are then topologically linked together into "blocks." The PF algorithm then interprets

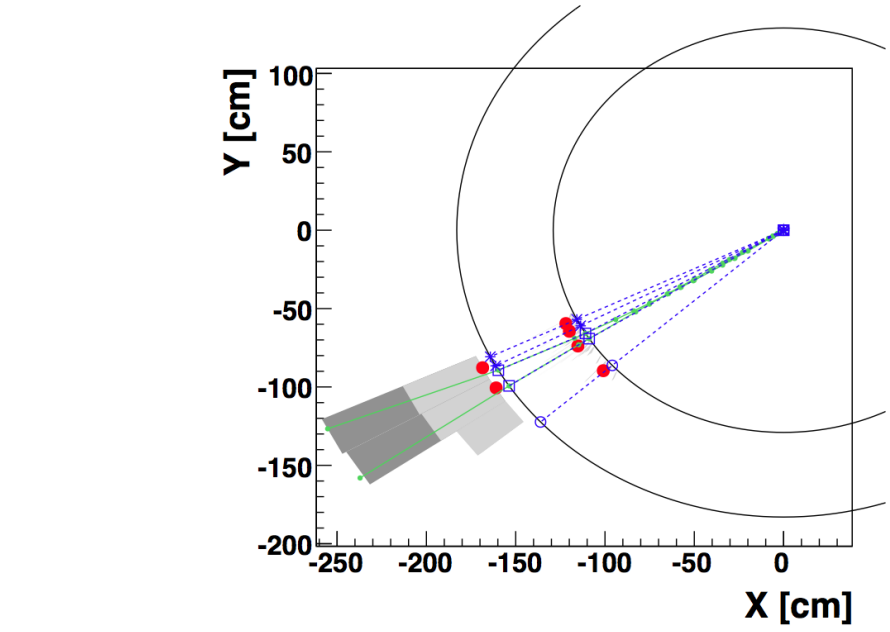
these blocks in terms of particles. Once obtained, this list of interpreted particles is used as if it came from an event generator to build jets and to calculate a $\not\!\!E_{\mathrm{T}}$ measurement. A visualisation of the PF treatment of a hadronic jet with four constituent particles (π^+, π^-, π^0, $\mathrm{K}^0_{\mathrm{L}}$) and a p_{T} of 65 GeV is shown in Figure 3.3 [94].

3.5.1 Fundamental elements

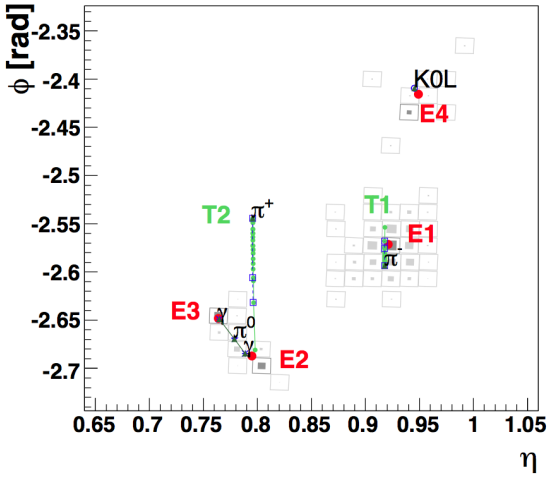
Two kinds of fundamental elements are important to this discussion: charged particle tracks from the inner tracker and calorimetric energy clusters from the ECAL and HCAL.

The reconstruction of charged particle tracks using data from the inner tracker has been discussed in Section 3.1. While the momentum of charged hadrons (e.g. pions) is measured both by reconstructing tracks in the inner tracker and by measuring the energy absorbed by the calorimeters, the resolution of the inner tracker's measurement is vastly superior for charged hadrons with p_{T} up to several hundreds of GeV. This is important, because an average of roughly two thirds of a jet's energy is carried by charged particles. In addition, the tracker provides a precise measurement of the trajectory of charged particles. These two features make charged particle tracks from the inner tracker a cornerstone of the PF reconstruction algorithm. In the simulated simple hadronic jet example of Figure 3.3, reconstructed tracks are represented by solid green lines.

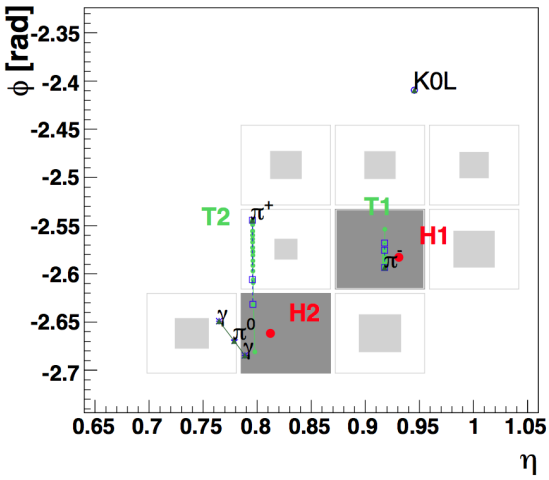
Calorimetric energy clusters are reconstructed using an algorithm specific to PF. This algorithm has been developed with at least four goals in mind: to detect and measure the energy and direction of stable neutral particles (like photons and neutral hadrons); to separate these neutral particles from energy deposits originating from charged hadrons; to reconstruct and identify electrons and the photons associated with Bremsstrahlung radiation; and to ameliorate the energy measurement of charged hadrons for which the track parameters were not determined accurately (as is the case for low-quality or high-p_{T} tracks). The algorithm is performed separately in the EB, EE, PS, HB, and HE. It is not performed in the HF, so each cell in the HF yields at most one cluster. The algorithm consists of three steps. First, the



(a) $r - \phi$ plane



(b) ECAL energy in the $\eta - \phi$ plane



(c) HCAL energy in the $\eta - \phi$ plane

Figure 3.3: An event display of a simulated simple hadronic jet with four constituent particles (π^+, π^-, π^0, K_L^0) and a p_T of 65 GeV interacting with the CMS inner tracker, ECAL, and HCAL. In all figures, the reconstructed charged tracks from the π^+ and π^- are shown as solid green lines, the calorimetric energy cluster positions are shown as closed red markers, the simulated paths of the particles are shown dashed blue lines, and the simulated positions of the particles' impacts with the calorimeters are shown as various open markers. Figure (a) is shown in the $r - \phi$ plane. The ECAL and HCAL surfaces are represented as black circles centered around the interaction point. Figure (b) is a view of the energy deposits in the ECAL as shown in the $\eta - \phi$ plane. The K_L^0, π^-, and photons from the $\pi^0 \to \gamma\gamma$ decay leave four well-separated clusters in the ECAL. The π^+ leaves no energy in the ECAL. Figure (c) is a view of the two energy deposits in the HCAL as shown in the $\eta - \phi$ plane [94].

algorithm identifies seeds for the clusters using local calorimeter cell energy maxima above a given energy. Second, "topological clusters" are formed from these seeds by combining cells with at least one side in common with a cell already included in the cluster and with an energy above a noise threshold (80 MeV in the EB, 300 MeV in the EE, and 800 MeV in the HCAL). Finally, "PF clusters" are formed using topological clusters as seeds. The number of PF clusters identified by the algorithm is equal to the number of cluster seeds identified. In the simulated simple hadronic jet example of Figure 3.3, calorimetric energy clusters in the ECAL and HCAL are represented by closed red markers [94].

3.5.2 Link algorithm

In general, a particle passing through the CMS detector is expected to give rise to more than one PF element. For example, an electron may give rise to a charged particle track and an ECAL energy cluster. These various elements are connected to each other using a linking algorithm in order to reconstruct the original single particle. The linking algorithm is tentatively performed between each pair of elements, and the quality of the link is quantified using a distance (defined below). The algorithm then produces a "block" of linked elements. Each block typically contains 1-3 elements. Three kinds of links are considered: between track and calorimeter cluster, between calorimeter cluster and calorimeter cluster, and between track and track.

In the case of linking between a charged particle track and a calorimeter cluster, the track is first extrapolated from its last measured hit in the tracker to all of the calorimeters to which the track is pointing. These calorimeters include the PS for forward tracks, the EB and EE (corresponding to an electron shower), and the HCAL (corresponding to a hadronic shower). The original track is linked to any cluster with boundaries that include the extrapolated track's position, but the cluster's boundaries may be expanded to account for the presence of gaps between cells, dead cells, or uncertainty on the shower position. The distance associated with the link is defined as the distance in the η-ϕ plane between the

position of the extrapolated track and the position of the cluster.

An additional case of track-to-cluster linking is defined to deal with electrons undergoing Bremsstrahlung radiation. In an attempt to collect all of the energy carried away by Bremsstrahlung photons emitted by electrons in the tracker, the tangents to tracks are extrapolated to the ECAL from the intersection points of the track with an inner tracker layer. A link is created if the extrapolated tangent is within the boundaries of an ECAL cluster, as described in the previous paragraph.

Links may also be created between two calorimeter energy clusters: between an ECAL cluster and an HCAL cluster, or between an ES cluster and an EE cluster, for example. Such a link is created when the cluster in the more granular detector lays within the $\eta - \phi$ envelope of the cluster in the less granular detector. As in the case of the track-to-cluster linking, the envelopes may be slightly enlarged to reflect gaps between cells, dead cells, and positional uncertainties. The distance associated with the link is defined as the distance in the $\eta - \phi$ plane between the two cluster positions.

Finally, a link may be created between a track from the inner tracker and a track in the muon system. Such a link is created for every global muon for which the χ^2 is below a defined maximum. If more than one global muon can be created from several tracker tracks and a given muon track, the algorithm keeps only the global muon with the lowest χ^2. The χ^2 value defines the distance [94].

3.5.3 Particle reconstruction and identification

Once links and blocks have been created by the PF linking algorithm, the PF reconstruction algorithm reconstructs particles from those blocks and identifies them. After the reconstruction algorithm has reconstructed a particle candidate, the corresponding block and its constituent elements are removed from consideration while constructing additional particle candidates. This results in a global event description, which can be used to reconstruct jets and to provide a $\not\!\!E_{\mathrm{T}}$ measurement.

Muons and electrons are first particles to be identified and reconstructed. A global muon is identified as a "PF muon" if its momentum as measured by the combined fit is within three standard deviations of the momentum measured by the tracker only. Electrons are identified from blocks containing a track and an ECAL energy cluster. Electrons are pre-identified by using the tracker as a pre-shower: electrons are expected to have shorter tracks and to lose energy as they proceed radially outward through the tracker. These pre-identified electron tracks are fit with a GSF (see Section 3.3) in order to determine their trajectories. Final identification of "PF electrons" depends on a number of pre-defined ECAL and inner tracker variables.

"PF charged hadrons" are identified and reconstructed using tightly-selected tracks from the inner tracker and calibrated energy clusters from the ECAL and HCAL. The energy clusters are calibrated using Equation 3.1:

$$E_{\text{calib}} = a + b(E,\eta) \cdot E_{\text{ECAL}} + c(E,\eta) \cdot E_{\text{HCAL}} \tag{3.1}$$

where a, $b(E,\eta)$, and $c(E,\eta)$ are coefficients whose values were determined using a simulated sample of simulated single hadrons, η is the pseudorapidity of the HCAL cluster, E_{ECAL} is the uncalibrated energy of the ECAL cluster, E_{HCAL} is the uncalibrated energy of the HCAL cluster, and E is an estimate of the true energy of the hadron (either the total charged particle momentum, or the total uncalibrated calorimeter energy, whichever is larger). Links between the selected tracks and the HCAL clusters are then considered. If an HCAL cluster is linked to several tracks, then the sum of the track momenta is compared to the HCAL cluster. If a single track is linked to several HCAL clusters, only the closest cluster is kept for comparison. Any ECAL clusters associated with the tracks under consideration are ordered according to their distance to the closest track. The list of ECAL clusters is scanned, and the cluster is kept so long as E_{calib} is less than the total charged particle momentum.

If, in the end, E_{calib} is less than the sum of the momenta of the associated tracks, the

link is kept, and each remaining track in the block is reconstructed as a charged hadron, the momentum and energy of which are taken from the track momentum under a charged pion mass hypothesis. If E_{calib} is significantly larger than the charged particle momentum (i.e. outside of the calorimeter energy resolution), then a "PF photon" or a "PF neutral hadron" may be created with the energy associated with the excess. If the excess is larger than the total E_{ECAL}, then a photon is created with the excess ECAL energy, and a neutral hadron is created with the excess HCAL energy. Otherwise, only a photon is created. The preference for photons is justified by the observation that 25% of jet energy is carried by photons, while neutral hadrons leave only 3% of jet energy in the ECAL. Remaining ECAL and HCAL clusters which are not linked to tracks give rise to PF photons and PF neutral hadrons.

The final list of reconstructed PF electrons, muons, taus (not discussed here), charged hadrons, neutral hadrons, and photons are referred to collectively as "PF candidates" [94].

3.6 Jets and $\not{E}_{\mathbf{T}}$

As mentioned at the end of Section 1.1.1, jets are collimated spray of hadrons produced by the hadronization and fragmentation of quarks and gluons (collectively called "partons"). Several algorithms exist [96] for the purpose of reconstructing the energy and momentum of the original parton, using various signatures left by the parton's decay products. In the case of this analysis, both at the trigger level and at the analysis level, the anti-k_{T} algorithm [97] is used with a characteristic radius parameter of $R = 0.5$. At the trigger level, only energy information from the calorimeters is used as an input to the anti-k_{T} algorithm to reconstruct jets. Jets constructed using calorimeter input alone are referred to as "CaloJets". CaloJets provide an acceptable jet energy resolution within relatively little computation time. At the analysis level, the PF candidates described in Section 3.5 are used as an input to the anti-k_{T} algorithm. Jets reconstructed using the PF algorithm are referred to as "PFJets". PFJets have a better energy resolution than CaloJets [94] (especially at low jet p_{T}), since the

PF algorithm measures the momentum of charged hadrons (which carry an average of 90% of the jet energy) using the tracker rather than the calorimeters. However, PFJet reconstruction requires more computation time than CaloJet reconstruction [98].

Some particles, like neutrinos, may be produced in pp collisions without interacting with the detector. Neutrinos are neutral, so they do not leave tracks in the inner tracker or muon system, and they are weakly interacting, so they do not deposit energy in the calorimeters. These particles may only be detected indirectly, via a vector momentum or energy imbalance in the transverse plane. This imbalance (as discussed in Section 2.2.1) is refered to as $\vec{\not{E}}_{\mathrm{T}}$, and the magnetude of $\vec{\not{E}}_{\mathrm{T}}$ is referred to as $\not{E}_{\mathrm{T}}$ (MET). Just as there are many ways of calculating jets, there are many ways of calculating $\not{E}_{\mathrm{T}}$. In the case of this analysis, $\not{E}_{\mathrm{T}}$ is calculated using the negative vector sum of the momenta of the PF candidates discussed in Section 3.5. This is expressed in Equation 3.2:

$$\vec{\not{E}}_{\mathrm{T}} = -\sum_{i} \vec{p}_{\mathrm{T},i} \tag{3.2}$$

where the sum is taken over each PF candidate, i, and $\vec{p}_{\mathrm{T},i}$ refers to the vector of the transverse momentum of the ith PF candidate [99].

PFMHT, a comparable variable to $\not{E}_{\mathrm{T}}$, is calculated using Equation 3.2. However, in the case of PFMHT, PFJets are summed over instead of PF candidates. This variable, therefore, is calculated without input from "unclustered energy" (i.e. without particles from pp collisions that do not form jets). The amount of unclustered energy in an event is heavily dependent on the number of pile-up pp interactions, and PFMHT is occasionally used as a less pile-up-dependent substitute for $\not{E}_{\mathrm{T}}$. In this analysis, PFMHT is used exclusively at the trigger level.

3.7 Event generation and simulation

Simulated events play an essential role in exotic searches like the one presented in this thesis. Simulated events are produced using Monte Carlo methods, and the resulting datasets themselves are often refered to as Monte Carlo (MC) datasets. These MC datasets allow background Standard Model processes to be studied in isolation from each other, and they can also be combined to provide an approximate estimate of the total Standard Model background for the search. In addition, MC datasets allow physicists to study as yet unobserved exotic signals under various hypotheses, including varying masses, branching fractions, and other parameters.

In this thesis, all MC datasets are produced by simulating interacting partons by using a given parton distribution function (PDF). Many different parton distribution functions are available, but this analysis makes use of the `CTEQ6L1` PDFs in all cases [55]. The hard scattering in these parton-parton interactions is modeled using event generators. Different event generators are used to model different processes in this analysis. These include Pythia [100], Sherpa [101], MadGraph [102, 103], and POWHEG [104–108]. For all processes in this thesis, the decay and hadronization of the particles emerging from the hard scatter is modeled by Pythia.

In addition, pile-up pp interactions taking place in the same bunch crossing as the interaction of interest must be modeled in all MC samples. Since the number of pile-up interactions in data depends on frequently changing conditions at the LHC, the pile-up distribution in all MC samples must be reweighted to agree with data. A systematic uncertainty associated with this reweighting is taken into account in all CMS analyses, and it is discussed in the context of this analysis in Section 9.9.

In order to compare simulated events to data, it is also necessary to model the interactions of simulated particles with the CMS detector. This modeling is performed by the Geant4 simulation toolkit [109], which has been built into the CMS software framework. The digital signals from the detector's response to these particle interactions are also simulated. These

simulated digital signals allow the same reconstruction algorithms to be run over simulated events and real events from the CMS detector [67, 72].

CHAPTER 4

ANALYSIS: STRATEGY

Leptoquarks present a very attractive target for searches at the LHC. As discussed in Section 1.2.2, the pair production cross section of scalar leptoquarks at a pp collider like the LHC depends only on the leptoquark mass. As shown in Table 1.3, this cross section is relatively high: mBRW leptoquarks with a mass of 1 TeV have a pair production cross section of approximately $1.6 \cdot 10^{-4}$ pb in pp collisions at $\sqrt{s} = 7$ TeV. In addition, if $\beta = \mathrm{BR}(\mathrm{LQ} \to \ell^{\pm} q)$ is not zero, first generation leptoquarks produced at the LHC would decay to isolated, energetic electrons, which could be easily triggered upon (Section 2.2.8) and reconstructed (Section 3.3) using the CMS detector with relatively little background from Standard Model processes.

This thesis takes β as a free parameter, which means that β is allowed to take any value between 0 and 1, rather than being fixed to a value of 0, 1/2, or 1 as discussed in Section 1.2.1. This leads to a search for first generation leptoquarks in two channels. In the first channel, β is taken to be equal to 1, and the decay of pair-produced leptoquarks to two electrons and two jets ($\mathrm{LQ}\overline{\mathrm{LQ}} \to eejj$) is considered. In the second channel, β is taken to be equal to 1/2, and the decay of pair-produced leptoquarks to one electron, one neutrino, and two jets ($\mathrm{LQ}\overline{\mathrm{LQ}} \to e\nu jj$) is considered. As discussed in Section 1.2.3, a third channel

where β is taken to be zero and pair-produced leptoquarks decay to two neutrinos and two jets ($LQ\overline{LQ} \to \nu\nu jj$) is possible. While the $\beta = 0$ channel is not covered by this analysis, other analyses at CMS are sensitive to decays of massive scalar particles to a final state of $\not{E}_T$ and two or more energetic jets [110]. In the rest of this paper, the search for the pair production of scalar leptoquarks where $\beta = 1$ will be referred to as the "$eejj$ channel", and the search where $\beta = 1/2$ will be referred to as the "$e\nu jj$ channel." Results from these two searches may be combined to search for leptoquarks with values of β anywhere between 0 and 1 with varying sensitivity.

The basic strategy for the $eejj$ channel is to trigger on events with two electrons and then to apply an offline selection requiring exactly two electrons and at least two jets. Neither the $eejj$ channel nor the $e\nu jj$ channel requires exactly two jets, since additional jets are frequently produced in pp collisions at the LHC by initial state and final state gluon radiation (ISR and FSR). Both analyses, however, only use the two leading jets in p_T to reconstruct leptoquark candidates. Several Standard Model processes exist that produce a final state of two electrons and at least two jets. The most significant of these background processes for the $eejj$ channel are Z^0+jets and $t\bar{t}$. Other background processes in the $eejj$ channel include dibosons (WW, WZ, and ZZ), single top, γ+jets (where the photon converts to two electrons), W+jets (where at least one jet is misidentified as an electron) and QCD multijets (where at least two jets are misidentified as electrons). Several variables serve as descriminants to separate leptoquark signal from these Standard Model background processes. Both the $eejj$ channel and the $e\nu jj$ channel define an S_T variable as the scalar sum of the most energetic objects in the final state. In the case of the $eejj$ channel, S_T is defined using Equation 4.1:

$$S_T = p_T(e_1) + p_T(e_2) + p_T(j_1) + p_T(j_2) \tag{4.1}$$

where $p_T(e_1)$ is the scalar value of the p_T of the leading electron in p_T, $p_T(e_2)$ is the scalar value of the p_T of the second leading electron in p_T, $p_T(j_1)$ is the scalar value of the p_T of the leading jet in p_T, and $p_T(j_2)$ is the scalar value of the p_T of the second leading jet in p_T.

The average value of S_{T} is significantly larger for signal events, due to the energetic decay products produced by a heavy boson like a leptoquark. Another effective discriminant in the $eejj$ channel is the invariant mass of the electron pair in the event (m_{ee}). For Z^0+jets events (the Standard Model background in the $eejj$ channel with the highest cross section), the m_{ee} distribution forms a peak at the Z mass of 91.2 GeV, while it is significantly higher on average for leptoquarks. The final discriminant used in the $eejj$ channel comes from reconstructing leptoquark candidates. In the $eejj$ channel, electrons are paired with jets to reconstruct leptoquark candidates. There are two possible ways to combine two electrons (e_1 and e_2) and two jets (j_1 and j_2) to form two LQ candidates: (e_1–j_1,e_2–j_2) or (e_1–j_2, e_2–j_1). The pairing that minimizes the difference between the invariant masses of the resulting reconstructed leptoquark candidates is chosen to form leptoquark candidates. The mass of the least massive reconstructed leptoquark candidate ($m_{\mathrm{ej}}^{\mathrm{min}}$) is used as the final discriminant in the $eejj$ channel.

Similarly, the basic strategy for the $e\nu jj$ channel is to trigger on events with exactly one electron, large $\not\!\!E_{\mathrm{T}}$, and at least two jets and then to apply a tighter offline selection. In the case of the $e\nu jj$ channel, the most significant background processes are W+jets and $t\bar{t}$. Other background processes in the $e\nu jj$ channel include dibosons, single top, Z^0+jets (where one electron is not reconstructed) and QCD multijets (where one jet is misidentified as an electron). A similar S_{T} variable is defined for the $e\nu jj$ channel using Equation 4.2:

$$S_{\mathrm{T}} = p_{\mathrm{T}}(e_1) + \not\!\!E_{\mathrm{T}} + p_{\mathrm{T}}(j_1) + p_{\mathrm{T}}(j_2) \tag{4.2}$$

where $p_{\mathrm{T}}(e_1)$, $p_{\mathrm{T}}(j_1)$, and $p_{\mathrm{T}}(j_2)$ have the same definition as in Equation 4.1, and $\not\!\!E_{\mathrm{T}}$ represents the energy imbalance in the transverse plane as reconstructed by the PF algorithm. As in the $eejj$ channel, the average value of S_{T} is significantly larger for signal events in the $e\nu jj$ channel, due to the energetic decay products produced by a heavy boson like a leptoquark. In addition to the S_{T} variable, another effective discriminant in the $e\nu jj$ channel is the transverse mass

of the electron and $\not\!\!E_{\mathrm{T}}$ in the event ($m_{\mathrm{T,\,e\nu}}$). $m_{\mathrm{T,\,e\nu}}$ is defined by Equation 4.3:

$$m_{\mathrm{T,\,e\nu}} = \sqrt{2 \cdot p_{\mathrm{T}}(e_1) \cdot [1 - \cos(\Delta\phi(e_1, \not\!\!E_{\mathrm{T}}))]} \tag{4.3}$$

where $p_{\mathrm{T}}(e_1)$ and $\not\!\!E_{\mathrm{T}}$ have the same definition as in Equation 4.2, and $\Delta\phi(e_1, \not\!\!E_{\mathrm{T}})$ is the opening angle in ϕ between the leading electron in p_{T} and the $\not\!\!E_{\mathrm{T}}$. For W+jets events (the Standard Model background in the $e\nu jj$ channel with the highest cross section), the $m_{\mathrm{T,\,e\nu}}$ distribution forms a Jacobian peak close to W mass of 80.4 GeV, while it is significantly higher on average for leptoquarks. Another effective discriminant between leptoquarks and $t\bar{t}$ and W+jets background events is $\not\!\!E_{\mathrm{T}}$ itself. $\not\!\!E_{\mathrm{T}}$ is expected to be significantly higher on average for leptoquark events than for $t\bar{t}$ and W+jets background events, due to the energetic neutrino produced by the leptoquark decay. The final discriminant used for the $e\nu jj$ channel comes from reconstructing leptoquark candidates. Similarly to the $eejj$ channel, there are two possible ways to combine one electron (e_1, one neutrino (i.e. $\not\!\!E_{\mathrm{T}}$) and two jets (j_1 and j_2) to form two leptoquark candidates: (e–j_1, $\not\!\!E_{\mathrm{T}}$–j_2) or (e–j_2,$\not\!\!E_{\mathrm{T}}$–j_1). The pairing that minimizes the difference between the transverse masses of the resulting reconstructed leptoquark candidates is chosen to form leptoquark candidates. The invariant mass of the electron-jet pairing (m_{ej}) is used as the final discriminant in the $e\nu jj$ channel.

In both the $eejj$ channel and the $e\nu jj$ channel, a preselection is defined that is dominated by Standard Model background, in order to study the modeling of the various background predictions. Once the backgrounds are well understood, a series of final selections is applied using the discriminants for each channel: S_{T}, m_{ee}, and $m_{\mathrm{ej}}^{\mathrm{min}}$ for the $eejj$ channel and S_{T}, $m_{\mathrm{T,\,e\nu}}$, $\not\!\!E_{\mathrm{T}}$, and m_{ej} for the $e\nu jj$ channel. Each final selection is optimized for a given hypothesis for the leptoquark invariant mass, ranging between 250 and 900 GeV. Further details on the datasets used by this analysis, preselection, final selection optimization, and background modeling are given in the following chapters.

CHAPTER 5

ANALYSIS: SAMPLES

This section describes the datasets used to perform this analysis. Both data from the CMS detector (Section 5.1) and simulation from Monte Carlo event generators (Section 5.2) are described.

5.1 Data samples

The datasets used for this analysis were collected by the CMS detector during 2011, when the LHC was operating with a collision energy of 7 TeV. The entire 2011 dataset corresponds to an integrated luminosity of 4.95 fb^{-1}. Only data that were taken during a period with stable LHC beams and with all CMS subdetectors operating without problems was used in the analysis. Known dead or noisy channels in the calorimeters were masked in the reconstruction. The average number of pile-up interactions for the first period of data taking (*Run2011A*, from March to August, 2011) was approximately 6.5. For the second period of data taking (*Run2011B*, from September to October, 2011) the average number of pile-up interactions was approximately 12. The average number of pile-up interactions over the entire data sample considered is approximately 9.5.

For the $eejj$ channel, the data stored in the *Photon* PD were used as reported in Table 5.1.

Dataset	Run range	$\mathcal{L}_{int}$ [pb^{-1}]
/Photon/Run2011A-May10ReReco-v1/AOD	160329–163869	215.06
/Photon/Run2011A-PromptReco-v4/AOD	165071–168437	948.92
/Photon/Run2011A-05Aug2011-v1/AOD	170053–172619	389.82
/Photon/Run2011A-PromptReco-v6/AOD	172620–175770	705.57
/Photon/Run2011B-PromptReco-v1/AOD	175832–180296	2693
	Total integrated luminosity ($\mathcal{L}_{int}$)	4.95 fb^{-1}

Table 5.1: *Photon* datasets: the dataset name, the run range considered, and the corresponding integrated luminosity $\mathcal{L}_{int}$ of the sample after removing data taking periods with known detector problems.

Dataset	Run range	$\mathcal{L}_{int}$ [pb^{-1}]
/SingleElectron/Run2011A-May10ReReco-v1/AOD	160329–163869	215.63
/ElectronHad/Run2011A-PromptReco-v4/AOD	165071–168437	931.61
/ElectronHad/Run2011A-05Aug2011-v1/AOD	170053–172619	389.88
/ElectronHad/Run2011A-PromptReco-v6/AOD	172620–175770	706.72
/ElectronHad/Run2011B-PromptReco-v1/AOD	175832–180296	2702
	Total integrated luminosity ($\mathcal{L}_{int}$)	4.95 fb^{-1}

Table 5.2: *SingleElectron* and *ElectronHad* datasets: the dataset name, the run range considered, and the corresponding integrated luminosity $\mathcal{L}_{int}$ of the sample after removing data taking periods with known detector problems.

The events used in this channel were collected by unprescaled double-photon or double-electron triggers as described in Section 7.2.1.

For the $e\nu jj$ channel, the data stored in the *SingleElectron* and *ElectronHad* PDs were used as reported in Table 5.2. The events used in this channel were collected by unprescaled single-electron or electron-PFMHT-jet-jet triggers as described in Section 7.3.1.

For the determination of the $t\bar{t}$ background in the $eejj$ channel, the data stored in the *MuEG* PD were used as reported in Table 5.3. The events used in this study were collected by unprescaled muon-plus-photon or muon-plus-electron triggers as described in Section 8.2.1.

Finally, for the determination of the QCD multijet background in both the $eejj$and $e\nu jj$channels, the data stored in the *Photon* PD were used as reported in Table 5.1. The events used in this study were collected by prescaled single-photon triggers as described in Section 8.1.

Dataset	Run range	$\mathcal{L}_{int}$ (pb^{-1})
/MuEG/Run2011A-May10ReReco-v1/AOD	160329–163869	215.63
/MuEG/Run2011A-PromptReco-v4/AOD	165071–168437	946.55
/MuEG/Run2011A-05Aug2011-v1/AOD	170053–172619	389.88
/MuEG/Run2011A-PromptReco-v6/AOD	172620–175770	706.72
/MuEG/Run2011B-PromptReco-v1/AOD	175832–180296	2706
	Total integrated luminosity ($\mathcal{L}_{int}$)	4.95 fb^{-1}

Table 5.3: *MuEG* datasets: the dataset name, the run range considered, and the corresponding integrated luminosity $\mathcal{L}_{int}$ of the sample after removing data taking periods with known detector problems.

5.2 Monte Carlo samples

The collision data described in Section 5.1 are compared to samples of Monte Carlo (MC) generated events. Different processes were modeled using MC generators, as detailed below. For the generation of all the MC samples, the `CTEQ6L1` [55] parton distribution functions (PDFs) were used. The response of the detector was simulated using GEANT4 as described in Section 3.7. The detector geometry description included realistic subsystem conditions such as defunct and noisy channels.

The signal pair-production scalar first generation leptoquark MC samples were generated using PYTHIA [100], version 6.422, tune Z2. Two different sets of leptoquark samples are produced: one where both leptoquarks decay to electrons and up quarks (for the $eejj$ channel), and one where one leptoquark decays to an electron and an up quark and the other leptoquark decays to an electron neutrino and a down quark (for the $e\nu jj$ channel). In both cases, the leptoquarks are produced with a coupling equal to the electromagnetic coupling, $\lambda = \lambda_{\text{em}} = 0.3$. The leptoquarks modeled by the MC have an electric charge of $\pm 1/3$, corresponding to leptoquarks of type $S_{0,L}$ and $S_{1,L}$ as described by the Lagrangian in Equation 1.2 and shown in the top row of Table 1.2. However, since neither the search in the $eejj$ channel nor the search in the $e\nu jj$ channel are sensitive to leptoquark electric charge, hypercharge, or weak isospin, the leptoquarks modeled by the MC are representative of all of the scalar leptoquarks described by the mBRW model. 11 samples were generated

M(LQ) [GeV]	Final state	$\sigma \times$ BR (NLO) [pb]	Events generated	Equivalent $\mathcal{L}_{\text{int}}$ [fb^{-1}]
250	$eejj$	3.47	64656	18.6
350	$eejj$	0.477	67528	141.6
400	$eejj$	0.205	52450	255.9
450	$eejj$	$0.948 \cdot 10^{-1}$	52382	552.6
500	$eejj$	$0.463 \cdot 10^{-1}$	61888	1336.7
550	$eejj$	$0.236 \cdot 10^{-1}$	61910	2623.3
600	$eejj$	$0.124 \cdot 10^{-1}$	49665	4005.2
650	$eejj$	$0.676 \cdot 10^{-2}$	63664	9417.8
750	$eejj$	$0.214 \cdot 10^{-2}$	51745	24179.9
850	$eejj$	$0.732 \cdot 10^{-3}$	52432	71628.4
900	$eejj$	$0.436 \cdot 10^{-3}$	59779	137107.8
250	$e\nu jj$	1.74	60119	34.7
350	$e\nu jj$	0.238	58891	246.9
400	$e\nu jj$	0.102	68538	668.7
450	$e\nu jj$	$0.474 \cdot 10^{-1}$	59746	1260.5
500	$e\nu jj$	$0.232 \cdot 10^{-1}$	52636	2273.7
550	$e\nu jj$	$0.118 \cdot 10^{-1}$	64246	5444.6
600	$e\nu jj$	$0.620 \cdot 10^{-2}$	61474	9915.2
650	$e\nu jj$	$0.338 \cdot 10^{-2}$	57873	17122.2
750	$e\nu jj$	$0.107 \cdot 10^{-2}$	58450	54626.2
850	$e\nu jj$	$0.366 \cdot 10^{-3}$	54181	148035.5

Table 5.4: Leptoquark masses, final states, NLO cross sections, number of events generated, and equivalent luminosities for pair-production scalar leptoquark MC samples. All values correspond to pp collisions at $\sqrt{s} = 7$ TeV (the LHC collision energy in 2010 and 2011). Cross sections are given in units of pb (1 b $= 10^{-28}\text{m}^2 = 10^{-24}\text{cm}^2$). In all cases, the renormalization and factorization scale is set to be equal to the leptoquark mass.

for the $eejj$ channel, and 10 samples were generated for the $e\nu jj$ channel over different ranges of leptoquark masses. The range of leptoquark masses extends from 250 GeV to 900 for the $eejj$ channel and from 250 to 850 for the $e\nu jj$ channel. A broader mass range is needed for the $eejj$ channel because it has lower background and greater sensitivity than the $e\nu jj$ channel. Table 5.4 shows the leptoquark mass value, final state, NLO cross section, number of generated events, and equivalent luminosity for each signal leptoquark sample.

MC samples were also generated to model the Standard Model background. W+jets and Z^0+jets events were generated using SHERPA [101]. The $t\bar{t}$ and γ+jets events were generated using MADGRAPH [102, 103]. The single top events were generated using POWHEG [104–

108]. The diboson (WW, ZZ, WZ) events were generated using PYTHIA. For the MADGRAPH and POWHEG samples, parton showering and hadronization were performed with PYTHIA. The QCD multijet background is estimated from data, as described in Section 8.1.

The total SHERPA cross sections for the W+jets and Z^0+jets samples were rescaled to inclusive next-to-next-to leading order (NNLO) values of 31314 pb (for W+jets) and 3048 pb (for $Z/\gamma \to \ell\ell$ with $M_{\ell\ell} > 50$ GeV). The SHERPA cross sections were calculated using FEWZ [111]. The MADGRAPH $t\bar{t}$ sample was normalized to an inclusive next-to-next-to-leading-logarithm (NNLL) cross section of 163 pb [112]. The γ+jets samples were normalized to LO cross sections calculated with MADGRAPH. The single top samples were generated via three distinct channels: s-channel, t-channel, and tW-channel. An NNLL cross section of 3.6 pb was calculated for the s-channel [113]. NLO cross sections of 54.1 pb and 14.9 pb were calculated for the t-channel and tW-channel, respectively, using MCFM [114]. The WW, WZ, and ZZ samples were normalized to NLO cross sections of 47, 18.2, and 7.4 pb, respectively, calculated with MCFM.

CHAPTER 6

ANALYSIS: OBJECT SELECTION

Electron candidates (Section 3.3), muon candidates (Section 3.4), jets, and $\not{E}_{\mathrm{T}}$ (Section 3.6) are all reconstructed with very loose identification requirements, which are designed to be inclusive for a broad range of analyses. To better reject various backgrounds while still preserving an efficient selection on leptoquarks, additional identification criteria are applied to all of these physics object candidates. These criteria are all detailed in this chapter.

6.1 Electrons

The electron candidates used in this analysis are taken from GSF electron candidates, which are reconstructed according to the process described in Section 3.3. On top of the selection applied during reconstruction, a further electron identification selection is applied to reduce contamination from various backgrounds. This additional identification is referred to as the v3.1 high energy electron pair (HEEP) ID, and it was originally developed for an analysis searching for a Z' resonance decaying to two electrons [115]. Significant backgrounds include electrons produced within jets, electrons from photon conversion, and various physics objects (jets, for example) that are misreconstructed as electrons. These misreconstructed objects are referred to as "fake electrons". Since the Z' search used the HEEP ID to identify high

Variable	Barrel (EB) criterion	Endcap (EE) criterion
E_{T}	> 30 GeV	> 30 GeV
$\lvert\eta_{\mathrm{SC}}\rvert$	$\lvert\eta\rvert < 1.442$	$\lvert\eta\rvert > 1.56$
ecalDrivenSeed()	$= 1$	$= 1$
$\Delta\eta_{\mathrm{in}}$	< 0.005	< 0.007
$\Delta\phi_{\mathrm{in}}$	< 0.09	< 0.09
H/E	< 0.05	< 0.05
$\sigma_{i\eta i\eta}$	-	< 0.03
$E^{2\times5}/E^{5\times5}$	> 0.94 OR $E^{1\times5}/E^{5\times5} > 0.83$	-
$\mathrm{EM_{Iso}} + \mathrm{HAD_{Iso}^{layer1}}$	$< 2 + 0.03 \times E_{\mathrm{T}}$ GeV	< 2.5 GeV (for $E_{\mathrm{T}} < 50$ GeV) $< 2.5 + 0.03 \times (E_{\mathrm{T}} - 50)$ GeV (for $E_{\mathrm{T}} > 50$ GeV)
$\mathrm{HAD_{Iso}^{layer2}}$	-	< 0.5
$\mathrm{TRK_{Iso}}$	< 7.5 GeV	< 15 GeV
Inner Layer Lost Hits	$= 0$	$= 0$

Table 6.1: The "HEEP v3.1" selection criteria for electron ID and isolation.

energy electrons, many searches for heavy, exotic physics processes with a final state of high energy electrons (including first generation leptoquarks) use the HEEP ID also.

Differences in selection aside, the main difference between HEEP electrons and GSF electrons comes from the energy measurement. The energy measurement for GSF electrons is a weighted average of the energy value obtained from the GSF fit and the energy value obtained from the supercluster. The energy measurement for HEEP electrons comes purely from the supercluster. The weighted average gives better performance for low energy electrons ($E_{\mathrm{T}} < 15$ GeV) but for electrons with $E_{\mathrm{T}} > 25$ GeV, the weighted average is dominated by the supercluster measurement and the two methods yield effectively identical results. In certain rare situations, however, it is possible for the weighted average to discard the supercluster measurement completely and use the GSF track measurement alone. This can result in low-energy electrons mistakenly being assigned very high energy measurements. HEEP electrons use only the supercluster energy by default in order to avoid this situation. In addition to passing the HEEP ID, electrons in this analysis are required to have $E_{\mathrm{T}} > 40$ GeV.

The HEEP ID is a cut-based ID, and the contributing variables are described below. The numerical requirements for each of these variables are shown in Table 6.1.

- **E_{T}:** the transverse energy of the electron. E_{T} is defined as the calibrated energy measurement from the ECAL supercluster multiplied by $\sin(\theta_{\mathrm{track}})$, where θ_{track} is the polar angel of the electron GSF track measured at the inner layer of the inner tracker and extrapolated to the pp interaction vertex.

- **η_{SC}:** the pseudorapidity of the supercluster used to reconstruct the electron candidate. Electrons with $|\eta_{\mathrm{SC}}| < 1.442$ are identified as electrons from the EB. Electrons with $1.560 < |\eta_{\mathrm{SC}}| < 2.5$ are identified as electrons from the EE. Electrons with any other value of $|\eta_{\mathrm{SC}}|$ are not considered for this analysis. This requirement excludes electrons that are reconstructed near the gap between the EB and the EE, where the reconstruction efficiency and energy resolution are poor. This gap is illustrated in Figure 2.12.

- **Electron seeding algorithm.** Two electron seeding algorithms are used in CMS: ECAL-driven and tracker-driven. Both seeding algorithms are described in Section 3.3. Only electrons that are purely ECAL-driven or both ECAL-driven and tracker-driven are considered for this analysis. There are two reasons for this requirement. First, ECAL-driven electrons are compatible with the L1 trigger, which uses ECAL-driven measurements to select events with high energy electrons. Second, at the time of this analysis, only ECAL-driven reconstruction had been validated for the high energy electrons used in this analysis.

- **$\Delta\eta_{\mathrm{in}}$:** the difference between the value of η taken from the GSF track measured at the inner layer of the inner tracker and extrapolated to the pp interaction vertex and the value of η taken from the electron's supercluster (η_{SC}). $\Delta\eta_{\mathrm{in}}$ is required to be close to zero.

- **$\Delta\phi_{\mathrm{in}}$:** the difference between the value of ϕ taken from the GSF track measured at the inner layer of the inner tracker and extrapolated to the pp interaction vertex and the value of ϕ taken from the electron's supercluster (ϕ_{SC}). $\Delta\phi_{\mathrm{in}}$ is required to be close to zero.

- H/E**:** ratio between the energy measured by the HCAL within a cone of radius $\sqrt{\Delta\phi^2 \times \Delta\eta^2} = 0.15$ centered on the electron's supercluster position and the energy of the electron's supercluster. H/E is required to be small.

- $\sigma_{i\eta i\eta}$**:** a measure of the spread in pseudorapidity of the electron energy within a 5×5 block of ECAL crystals, centered on the seed crystal. As mentioned in Section 3.3, the spread of the electron energy is expected to be narrow in pseudorapidity, so $\sigma_{i\eta i\eta}$ is required to be small.

- $E^{2\times5}/E^{5\times5}$**:** the fraction of the total electron energy contained within a block of 2 crystals (in η) $\times$ 5 crystals (in ϕ). $E^{2\times5}/E^{5\times5}$ is required to be large in order to remove electrons that are contained within jets.

- $E^{1\times5}/E^{5\times5}$**:** the fraction of the total electron energy contained within a block of 1 crystal (in η) $\times$ 5 crystals (in ϕ). $E^{1\times5}/E^{5\times5}$ is required to be large in order to remove electrons that are contained within jets.

- **ECAL isolation:** the sum of the transverse energy of all EB RecHits with $E > 0.08$ GeV and all EE RecHits with $E > 0.1$ GeV within a cone of radius $\sqrt{\Delta\phi^2 \times \Delta\eta^2} = 0.3$ centered around the electron's position in the ECAL. An inner cone with a radius of three crystals and an η strip of width 3 crystals is excluded from this sum. ECAL isolation is required to be small in order to remove electrons that are contained within jets.

- **HCAL depth 1 isolation:** the sum of the HCAL "depth 1" RecHits within a cone of $\sqrt{\Delta\phi^2 \times \Delta\eta^2} = 0.3$ centered on the electron's position in the ECAL. HCAL RecHits within a cone of $\sqrt{\Delta\phi^2 \times \Delta\eta^2} = 0.15$ are excluded from this sum. In this context, depth 1 is defined as all depths for HCAL towers 1-17, depth 1 for towers 18-29, and depth 2 for towers 27-29. The depth segmentation of the HCAL is described in Section 2.2.5 and shown in Figure 2.18. HCAL depth 1 isolation is required to be small in order

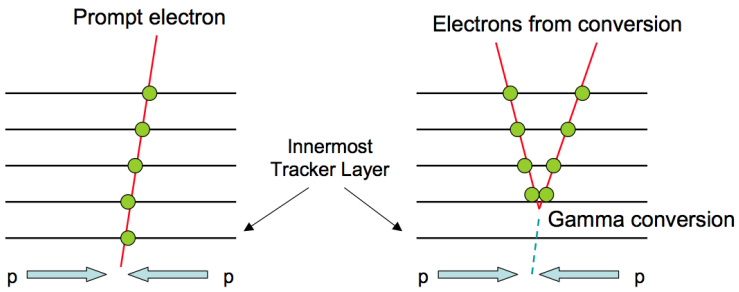


Figure 6.1: Prompt electrons originating from the primary pp collision vertex (left) usually have "hits" on the innermost layer of the inner tracker. Conversely, electrons from converted photons (right) often do not have hits on the innermost layer of the inner tracker.

to remove electrons that are contained within jets.

- **HCAL depth 2 isolation:** the sum of the HCAL "depth 2" RecHits within a cone of $\sqrt{\Delta\phi^2 \times \Delta\eta^2} = 0.3$ centered on the electron's position in the ECAL. HCAL RecHits within a cone of $\sqrt{\Delta\phi^2 \times \Delta\eta^2} = 0.15$ are excluded from this sum. In this context, depth 2 is defined as all depths that are not included in depth 1 for towers 18-29. The depth segmentation of the HCAL is described in Section 2.2.5 and shown in Figure 2.18. HCAL depth 2 isolation is required to be small in order to remove electrons that are contained within jets.

- **Tracker isolation:** the sum of the p_{T} of the tracks within a cone of radius $\sqrt{\Delta\phi^2 \times \Delta\eta^2} = 0.3$ centered around the GSF track at the interaction point having $p_{\mathrm{T}} > 0.7$ GeV and z_0 within 0.2 cm of the z_0 of the electron's GSF track. Tracks within a cone of $\sqrt{\Delta\phi^2 \times \Delta\eta^2} = 0.04$ of the GSF track at the interaction point are not considered. The variable z_0 is defined as the minimum distance in z from the point (0,0,0). Tracker isolation is required to be small in order to remove electrons that are contained within jets.

- **Missing tracker hits:** the number of expected hits in the inner tracker that are missing from a GSF track. As shown in Figure 6.1, electrons that are produced from photon conversions often have missing hits in the first layer of the inner tracker. For

Variable	Criterion
Pixel hits	> 0
Tracker hits	> 10
Muon chamber hits	> 0
Muon stations	> 1
$\lvert d_{xy} \rvert$	< 2 mm
χ^2/ndof	< 10
Tracker isolation	< 3 GeV

Table 6.2: The "tight" selection criteria for muon ID and isolation.

this reason, the HEEP ID requires electrons to have zero missing tracker hits.

6.2 Muons

The muon candidates used in this analysis are taken from global muons, which are reconstructed according to the process described in Section 3.4. On top of the selection applied during reconstruction, a further muon identification selection is applied to reduce contamination from various backgrounds. This additional identification selection is known as the "tight" muon ID, and it was developed by the muon physics object group within CMS. Significant backgrounds include muons produced within jets and hadrons that are incorrectly identified as muons. These misreconstructed objects are refered to as "fake muons", and two important sources of them are hadrons decaying to muons while passing through the detector (e.g. $\pi^{\pm} \to \mu^{\pm}\nu_{\mu}$, "decay in flight") and hadrons that pass through the entire detector and interact with the muon subsystem without showering ("punch through"). The tight muon ID has been studied in detail [95], and it is used by many CMS analyses that include muons in their final stat, including the second generation leptoquark analysis [53].

Like the HEEP ID for electrons, the tight muon ID is a cut-based ID, and the variables contributing to it are described below. The numerical requirements for each of these variables are shown in Table 6.2.

- **Muon reconstruction algorithm.** Several muon reconstruction algorithms are used

by various analyses at CMS. In this analysis, the muon must be a global muon.

- **Pixel hits:** number of hits in the pixel detector. This value must be large in order to suppress muons from decays in flight.

- **Tracker hits:** number of hits in the inner tracker system (including the pixel detector) included in the global muon fit. This number must be large in order to guarantee a good p_{T} measurement and to suppress muons from decays in flight.

- **Muon chamber hits:** number of hits in the muon subsystem included in the global muon fit. This number must be large in order to suppress fake muons from punch through and decays in flight.

- **Muon stations:** number of muon stations with muon segments in them which contribute to the global muon fit. This number must be large in order to suppress fake muons from punch through and accidental matches between unrelated tracks from the inner tracker and segments from the muon system. In addition, requiring at least two muon stations makes this selection consistent with the logic of the muon trigger.

- χ^2**/ndof:** an evaluation of the quality of the global muon fit. This value must be small in order to suppress fake muons from punch through and decays in flight.

- d_{xy}**:** transverse impact parameter with respect to the primary vertex. This number must be small in order to suppress cosmic muons and to suppress fake muons from decays in flight. The cut is loose enough to preserve efficiency for muons from decays of long-lived hadrons containing b and c quarks.

- **Tracker isolation:** the sum of the p_{T} of the tracks within a cone of radius $\sqrt{\Delta\phi^2 \times \Delta\eta^2} = 0.3$ around the muon track, excluding the muon track itself. This requirement is not part of the tight muon ID. It is applied in addition to the tight muon ID in order to suppress muons that are contained within jets.

Variable	Criterion
For all jets	
Neutral hadronic fraction	< 0.99
Neutral electromagnetic fraction	< 0.99
Number of constituents	> 1
For jets with $\|\eta\| < 2.4$	
Charged hadronic fraction	> 0
Charged electromagnetic fraction	< 0.99
Charged multiplicity	> 0

Table 6.3: The loose jet ID selection criteria for PFJets

Additionally for this analysis, muon candidates must have $p_{\mathrm{T}} > 30$ and at least one muon in the event must be reconstructed in the HLT fiducial volume, *i.e.* with $|\eta| < 2.1$.

6.3 Jets

The jets used in this analysis are reconstructed by running the anti-k_{T} jet algorithm with a radius parameter of 0.5 over PF candidates, as described in Section 3.6. On top of the selection applied during reconstruction and detector electronics noise cleaning applied during jet reconstruction, a further jet identification selection is applied to reduce contamination from detector electronics noise or single particles (electrons, for example) that are reconstructed as jets. This additional identification selection is particular to PFJets and is referred to as "loose jet ID".

The loose jet ID is a cut-based ID, and the contributing variables are described below. The numerical requirements for each of these variables are shown in Table 6.3.

- **Neutral hadron fraction:** fraction of the total jet energy associated with hadronic energy deposits that are not linked to tracks
- **Neutral electromagnetic fraction:** fraction of the total jet energy associated with electromagnetic energy deposits that are not linked to tracks
- **Number of constituents:** number of PF candidates included in the jet

- **Charged hadron fraction:** fraction of the total jet energy associated with hadronic energy deposits that are linked to tracks
- **Charged electromagnetic fraction:** fraction of the total jet energy associated with electromagnetic energy deposits that are linked to tracks
- **Charged multiplicity:** number of charged PF candidates included in the jet

Only jets with $p_{\mathrm{T}} > 30$ GeV are considered for the $eejj$ analysis, and only jets with $p_{\mathrm{T}} > 40$ GeV are considered for the $e\nu jj$ analysis. In both analyses, only jets with $|\eta| < 2.4$ are considered. Finally, to address the double-counting of well identified and isolated electrons and muons reconstructed also as jets, for each selected electron or muon the closest selected jet within $\Delta R = 0.3$[1] is removed from the jet collection. In this context, "selected" electrons, muons, and jets are required to pass the p_{T} cut of the event pre-selection described in the next sections, and the ID/isolation and noise cleaning requirements.

[1] ΔR is defined as $\sqrt{\Delta\phi^2 + \Delta\eta^2}$

CHAPTER 7

ANALYSIS: EVENT SELECTION

This chapter describes the process by which events are selected for both channels of this analysis. The event filters which are common to both channels are described in Section 7.1. The $eejj$ channel is described in Section 7.2. The $e\nu jj$ channel is described in Section 7.3. Both channels make use of the datasets described in Chapter 5, but the methods for modeling the various data-driven backgrounds and renormalizing the MC backgrounds are discussed in the following chapter (Chapter 8).

7.1 Event filters

All events in this analysis (both $eejj$ and $e\nu jj$ channels) are required to pass the event filters described in this section. These filters are designed to reject beam background, detector electronics noise, and other detector artifacts. These filters are placed on both data and MC events, unless otherwise indicated.

- **BPTX coincidence:** The beam pickup timing for the experiments (BPTX) is a beam pickup device designed to provide the experiments with precise information on the bunch structure and the timing of the LHC beam. Two are installed for CMS: 175

m in either direction of the IP. A coincidence of signals from both BPTX detectors (indicating a pp crossing at the IP) is required for all data events in this analysis. BPTX coincidence is not required for MC events.

- **Primary vertex requirement:** At least one "good" primary vertex is required to be reconstructed in each event. A good primary vertex is defined to be a primary vertex reconstructed with more than four degrees of freedom (see Section 3.2) and to have position with $|z| < 24$ cm and $\rho < 2$ cm.

- **Beam scraping filter:** This filter requires a fraction of *highPurity* tracks (see Section 3.1) to be great than 25% in events with more than 10 tracks. This requirement removes a rare beam-induced background known as "beam scraping", which arises from interactions between the beam and both residual gas particles and beam collimators. These interactions can produces showers of secondary particles, which the tracker registers as isotropic noise [67].

- **Beam halo filter:** This filter takes information from the endcap muon system (see the CSCs in Section 2.2.6) to reject a beam-induced background known as "beam halo", which manifests as charged particles travelling parallel to the beam at a large radius. Beam halo can be severely detrimental to $\not\!\!E_{\mathrm{T}}$ measurements.

- **HBHE noise filter:** This filter is designed to reject noise from the HB and HE readout electronics, which can be mistakenly reconstructed as hadronic energy deposits.

7.2 Event selection in the $eejj$ channel

This section describes the event selection for the search for leptoquarks in the decay channel $\mathrm{LQ}\overline{\mathrm{LQ}} \to eejj$. The two leading electrons and two leading jets in p_{T} in each event are used for this analysis. Section 7.2.1 describes the trigger requirements. Section 7.2.2 describes the preselection level of the search. Section 7.2.3 describes the optimization of the final selection

HLT path	Run number range
Run2011A	
HLT_DoublePhoton33_v1	160404–161176
HLT_DoublePhoton33_v2	161216–163261
HLT_DoublePhoton33_v3	163269–163869
HLT_DoubleEle33_CaloIdL_v1	165088–165633
HLT_DoubleEle33_CaloIdL_v2	165970–166967
HLT_DoubleEle33_CaloIdL_v3	167039–167913
HLT_DoubleEle33_CaloIdL_v4	170249–173198
HLT_DoubleEle33_CaloIdL_v5	173236–173692
Run2011B	
HLT_DoubleEle33_CaloIdL_v5	175832–178380
HLT_DoubleEle33_CaloIdT_v2	178420–179889
HLT_DoubleEle33_CaloIdT_v3	179959–180252

Table 7.1: HLT paths for $eejj$ channel

level of the search, and Section 7.2.4 describes the results observed at that final selection level.

7.2.1 Trigger

Events in the $eejj$ channel are selected online using unprescaled double photon or double electron triggers, depending on the run period. These triggers require the presence of two clusters in the ECAL, each with a transverse energy $E_{\mathrm{T}} > 33$ GeV and each matched to hits in the pixel detector. The triggers also require the absence of significant energy deposits in the HCAL cells directly behind the ECAL clusters. The exact trigger used depends on the run period. The full list of triggers for the $eejj$ channel and their associated run periods are are listed in Table 7.1.

The efficiency with which these triggers select events with two HEEP electrons (described in Section 6.1) has been studied in the context of the Z'→ ee analysis [115], and it is taken to be 100% with less than 1% uncertainty. A fully efficient trigger is also assumed in the $eejj$ channel. Events used for background estimates are not required to pass these triggers: data-driven background estimates use different triggers, and MC events are not required to

pass any trigger.

7.2.2 Preselection

A sample of events enriched in SM background processes is selected to verify the background estimate in the $eejj$ channel. This $eejj$ preselection proceeds with the following cuts (all the reconstructed objects are required to pass the selection criteria described in Chapter 6):

- Pass the event filters listed in Section 7.1
- Pass the signal triggers listed in Table 7.1 (not required for background estimates)
- Exactly 2 HEEP electrons with $p_{\mathrm{T}} > 40$ GeV
- At least 2 jets with $p_{\mathrm{T}} > 30$ GeVand $|\eta| < 2.4$
- No selected muons passing the ID described in Section 6.2
- $m_{\mathrm{ee}} > 60$ GeV
- $S_{\mathrm{T}} > 250$ GeV

The p_{T} cut on the electrons is chosen to be high enough for the triggers listed in Table 7.1 to be fully efficient. The $|\eta|$ cut on the jets rejects potential fake jets reconstructed from anomalous signals in the HF (which covers the range $|\eta| > 3.0$). This cut has a negligable impact on the leptoquark signal efficiency. The muon veto rejects backgrounds from $t\bar{t}$ events with an $e\mu jj$ final state. The m_{ee} cut is chosen to be compatible with the background MC, which only models $m_{\mathrm{ee}} > 50$ GeV. Both the m_{ee} and S_{T} cuts are looser than the cuts applied in the final selection of the $eejj$ channel.

At this stage of the selection, there is sufficient data to compare with the background predictions for all the observables employed in the final event selection. The distribution of the number of reconstructed primary vertices is shown in Figure 7.1. The p_{T} and η distributions of the two leading electrons and the two leading jets are shown in Figures 7.2, 7.3, 7.4, and

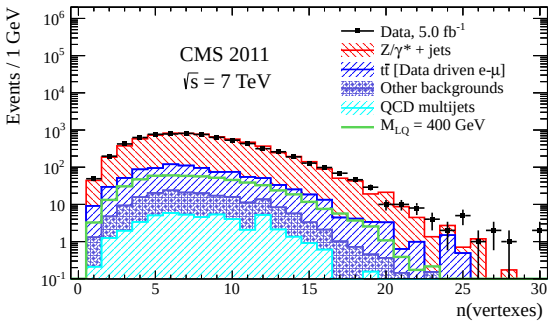


Figure 7.1: The distribution of the number of primary vertices for events passing the $eejj$ preselection.

7.5. The m_{ee} distribution is shown in Figure 7.6. The S_{T} and m_{ej} distributions are shown in Figure 7.7.

Overall, a good agreement is observed at $eejj$ preselection between data and background predictions for the shapes of all the distributions of observables employed in the final, optimized event selection. An exception is represented by m_{ee} shape, shown in Figure 7.6, for which a discrepancy is visible around the Z^0 mass peak. The data is shifted towards lower values of m_{ee} compared to the simulation, and the distribution of data is wider. This is due to the non-optimal ECAL calibration constants used in the reconstruction of data (the datasets used in this analysis are listed in Section 5.1). The discrepancy is more pronounced around the Z^0 mass peak compared to the m_{ee} tails due to the sharpness of the invariant mass shape in that region, which amplifies any small discrepancy between data and simulation. This effect has a small impact on the final results and it is taken into account in the analysis as a systematic uncertainty on the electron energy scale and resolution, as described in detail in Chapter 9.

7.2.3 Selection optimization

The final event selection criteria are optimized by maximizing the expected signal significance defined as $S/\sqrt{S+B}$, where S (B) is the expected number of signal (background) events

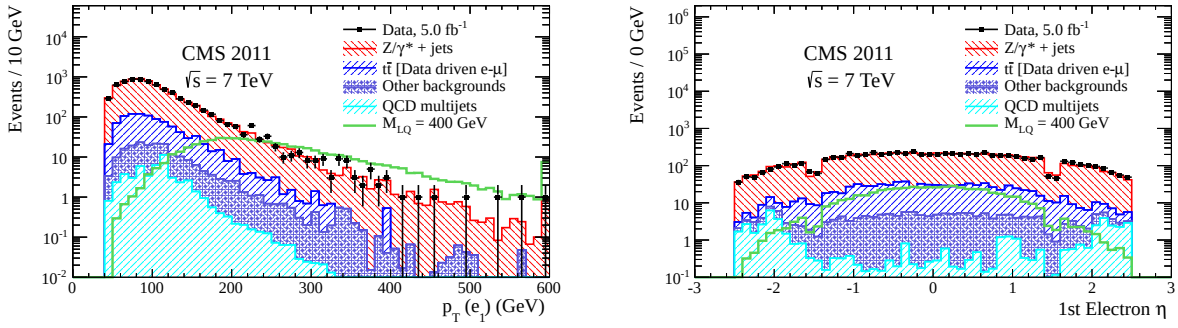


Figure 7.2: The p_{T} (left) and η (right) distributions of the leading (in p_{T}) electron for events passing the $eejj$ preselection.

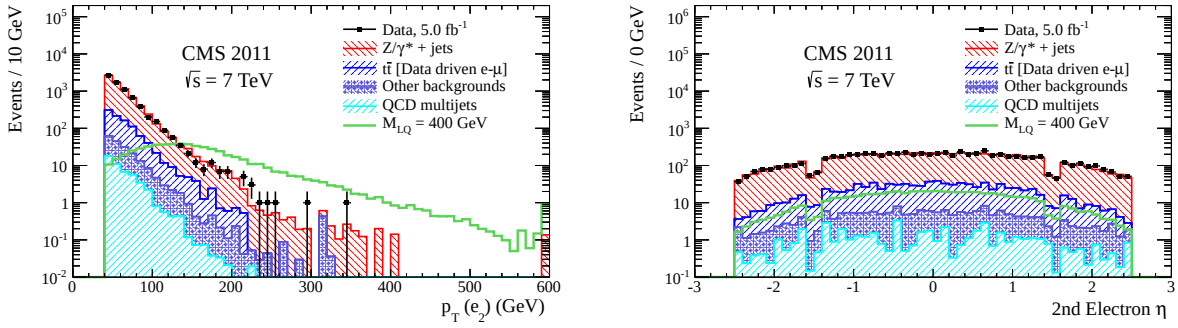


Figure 7.3: The p_{T} (left) and η (right) distributions of the second leading (in p_{T}) electron for events passing the $eejj$ preselection.

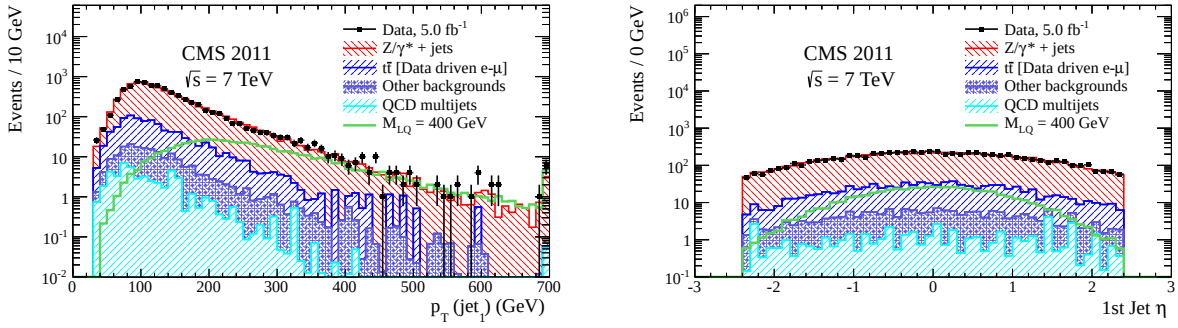


Figure 7.4: The p_{T} (left) and η (right) distributions of the leading (in p_{T}) jet for events passing the $eejj$ preselection.

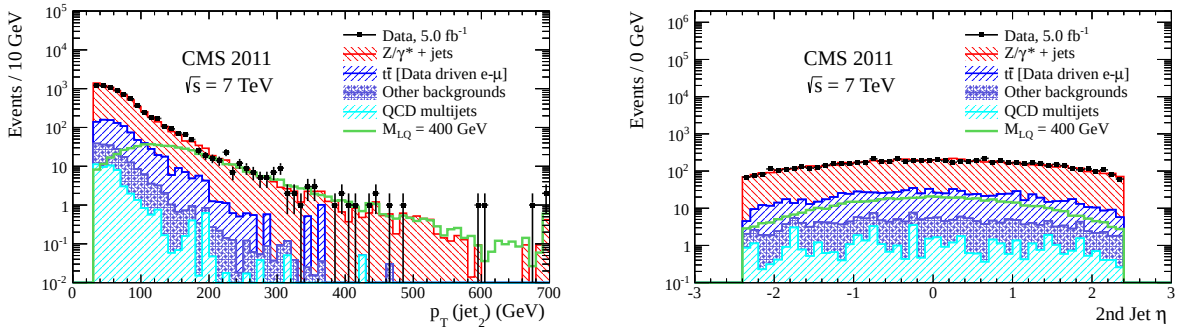


Figure 7.5: The p_{T} (left) and η (right) distributions of the second leading (in p_{T}) jet for events passing the $eejj$ preselection.

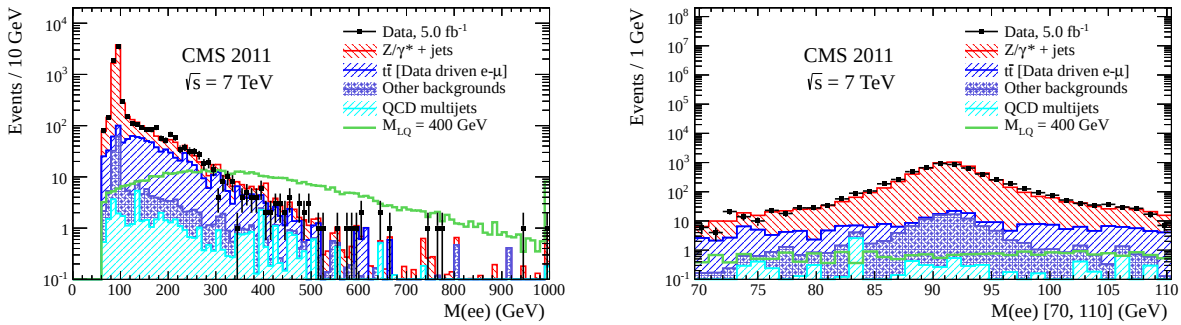


Figure 7.6: The m_{ee} distribution in the full range (left) and in a zoomed region around the Z^0 mass peak (right) passing the $eejj$ preselection.

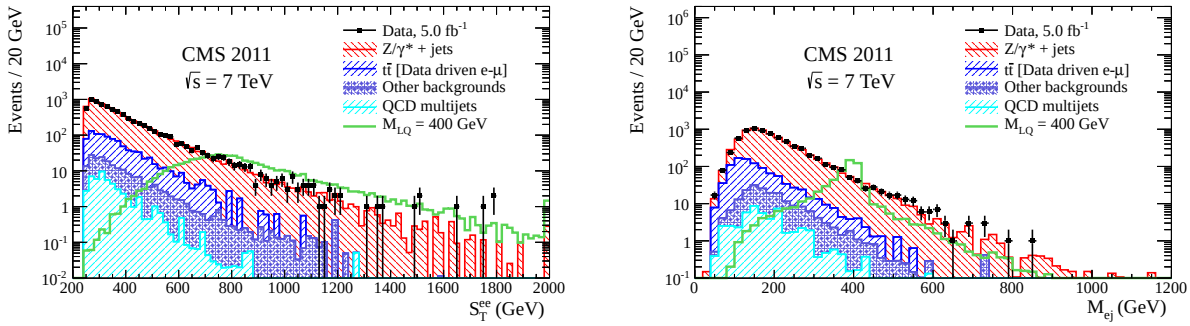


Figure 7.7: The S_{T} (left) and m_{ej} (right) distributions for events passing the $eejj$ preselection.

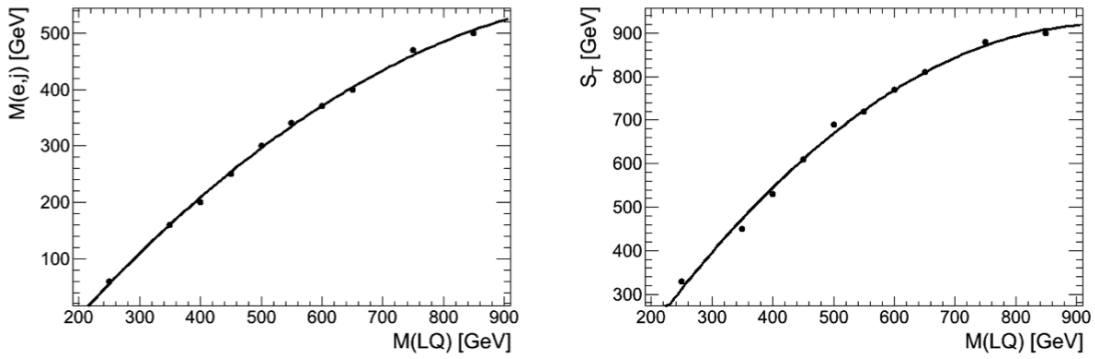


Figure 7.8: Optimized thresholds for for m_{ej} (left) and S_{T} (right) cuts as a function of leptoquark mass in the $eejj$ channel. The leptoquark mass hypothesis under consideration is shown on the x-axis. The optimized cut thresholds are shown on the y-axis. The resulting distributions are fit with a second-degree polynomial.

M_{LQ} [GeV]	250	350	400	450	500	550	600	650	750	850	900
S_{T} [GeV]	330	450	530	610	690	720	770	810	880	900	920
$m_{\mathrm{ej}}^{\mathrm{min}}$ [GeV]	60	160	200	250	300	340	370	400	470	500	520
m_{ee}[GeV]	100	110	120	130	130	130	130	130	140	150	150

Table 7.2: Optimized selection criteria for the $eejj$ channel for different LQ mass hypotheses.

passing the selection requirements. Three variables (S_{T}, $m_{\mathrm{ej}}^{\mathrm{min}}$, and m_{ee}) are optimized simultaneously by scanning appropriate ranges of values. For this study the $t\bar{t}$ MC is used, instead of the data-driven method described in Section 8.2.1. The results of the optimization, obtained for each LQ mass hypothesis, are summarized in Table 7.2 assuming an integrated luminosity of 4.95 fb^{-1}. Figure 7.8 shows the smooth dependence of the optimized m_{ej} and S_{T} cut values on the leptoquark mass hypothesis.

7.2.4 Final selection

Table 7.3 shows the number of events for the data, the backgrounds, and the LQ signal, after applying the final, optimized $eejj$ selection criteria summarized in Table 7.2. Figures 7.9 and 7.10 show the distributions of S_{T} and the electron-jet invariant mass, m_{ej}, for both leptoquark candidates (2 entries in for each event) after the full selection optimized for

M_{LQ}	LQ Signal	Z+Jets	$t\bar{t}$	QCD	Other	Data	Total BG
Presel	-	6234 ± 24	768 ± 19	49.59 ± 0.43	147.6 ± 2.3	7201	7199 ± 31
250	6846.2 ± 32.0	385 ± 6.0	334 ± 13	17.726 ± 0.186	28.3 ± 1.3	770	765 ± 14
350	1119.6 ± 4.5	88.5 ± 2.8	41.2 ± 4.5	1.934 ± 0.034	6.11 ± 0.64	139	138 ± 5.4
400	487.4 ± 2.2	35.7 ± 1.8	19.1 ± 3.1	0.877 ± 0.022	3.12 ± 0.56	55	58.8 ± 3.6
450	225.6 ± 1.0	15.2 ± 1.1	7.8 ± 2.0	0.310 ± 0.013	1.92 ± 0.60	26	25.2 ± 2.3
500	109.30 ± 0.46	6.55 ± 0.70	2.45 ± 1.10	0.192 ± 0.012	1.03 ± 0.42	14	10.2 ± 1.4
550	57.35 ± 0.23	4.65 ± 0.58	0.98 ± 0.69	0.139 ± 0.012	0.84 ± 0.42	11	6.60 ± 0.99
600	30.95 ± 0.14	3.04 ± 0.46	0.49 ± 0.49	0.088 ± 0.011	0.72 ± 0.41	8	4.34 ± 0.79
650	16.998 ± 0.0647	2.14 ± 0.38	0.49 ± 0.49	0.073 ± 0.011	0.48 ± 0.40	6	3.18 ± 0.74
750	5.5264 ± 0.0230	1.04 ± 0.26	$0.00^{+0.56}_{-0.00}$	0.00923 ± 0.00203	0.41 ± 0.40	0	$1.453^{+0.73}_{-0.47}$
850	1.9679 ± 0.0078	0.81 ± 0.23	$0.00^{+0.56}_{-0.00}$	0.00101 ± 0.00022	0.40 ± 0.40	0	$1.21^{+0.72}_{-0.46}$
900	1.1968 ± 0.0044	0.81 ± 0.23	$0.00^{+0.56}_{-0.00}$	0.00101 ± 0.00022	0.40 ± 0.40	0	$1.21^{+0.72}_{-0.46}$

Table 7.3: Number of events after the final $eejj$ selection. Only statistical uncertainties are reported.

M_{LQ}= 400 and 600 GeV, respectively. The dominant background contributions are from $t\bar{t}$ and Z^0+jets events, while the contribution from the other backgrounds is below 10% for the LQ masses within the current reach of this analysis. A good agreement is observed between the data and the background prediction within statistical uncertainties.

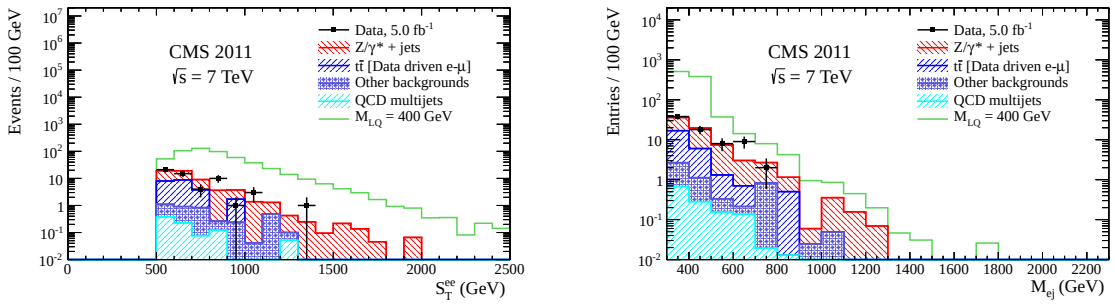


Figure 7.9: The S_{T} (left) and m_{ej} (right) distributions for events passing the full $eejj$ selection optimized for M_{LQ}= 400 GeV. The m_{ej} distribution on the right has two entries per event: one for each leptoquark candidate.

7.3 Event selection in the $e\nu jj$ channel

This section describes the event selection for the search for leptoquarks in the decay channel $\mathrm{LQ}\overline{\mathrm{LQ}} \to e\nu jj$. The leading electron and two leading jets in p_{T} in each event are used for this

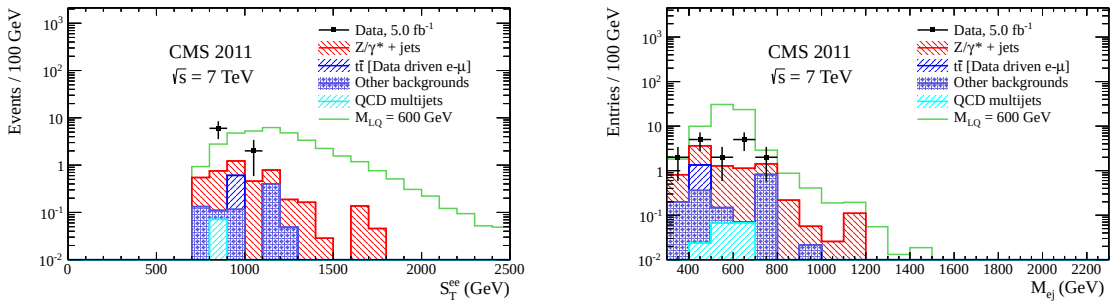


Figure 7.10: The S_{T} (left) and m_{ej} (right) distributions for events passing the full $eejj$ selection optimized for M_{LQ}= 600 GeV. The m_{ej} distribution on the right has two entries per event: one for each leptoquark candidate.

HLT path	Run number range
Run2011A	
HLT_Ele27_CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT_v1	160404–161176
HLT_Ele27_CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT_v2	161216–163261
HLT_Ele27_CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT_v3	163269–163869
HLT_Ele17_CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT_CentralJet30_CentralJet25_PFMHT15_v2	165088–165633
HLT_Ele25_CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT_CentralJet30_CentralJet25_PFMHT20_v4	165970–166967
HLT_Ele22_CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT_CentralJet30_CentralJet25_PFMHT20_v2	167039–167913
HLT_Ele22_CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT_CentralJet30_CentralJet25_PFMHT20_v4	170249–173198
HLT_Ele27_CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT_CentralJet30_CentralJet25_PFMHT20_v2	173236–173692
Run2011B	
HLT_Ele30_CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT_DiCentralJet30_PFMHT25_v1	175832–178380
HLT_Ele27_WP80_DiCentralPFJet25_PFMHT15_v4	178420–179889
HLT_Ele27_WP80_DiCentralPFJet25_PFMHT15_v5	179959–180252

Table 7.4: HLT paths for $e\nu jj$ channel

analysis, along with the $\not\!\!E_{\mathrm{T}}$ measurement. Section 7.3.1 describes the trigger requirements. Section 7.3.2 describes the preselection level of the search. Section 7.3.3 describes the optimization of the final selection level of the search, and Section 7.3.4 describes the results observed at that final selection level.

7.3.1 Trigger

Events in the $e\nu jj$ channel are selected online using unprescaled single electron or electron-$\not\!\!E_{\mathrm{T}}$-jet-jet triggers triggers. Each component of these triggers (e.g. electrons, jets, and PFMHT) is referred to as a "leg". The exact trigger used depends on the run period. The full list of triggers for the $e\nu jj$ channel and their associated run periods are listed in Table 7.4.

ID name	Calorimeter isolation requirements	Calorimeter identification requirements	Track isolation requirements	Track identification requirements
`CaloIdVT`	$H/E<$ 0.05 (0.05) $\sigma_{i\eta i\eta}<$ 0.011 (0.031)			
`CaloIsoT`		$\mathrm{EM_{Iso}}<$ 0.125 (0.075) $\mathrm{HAD_{Iso}}<$ 0.125 (0.075)		
`TrkIdT`			$\Delta\eta_{\mathrm{in}}<$ 0.008 (0.008) $\Delta\phi_{\mathrm{in}}<$ 0.07 (0.05)	
`TrkIsoT`				$\mathrm{TRK_{Iso}}<$ 0.125 (0.075)
`WP80`	$H/E<$ 0.04 (0.025) $\sigma_{i\eta i\eta}<$ 0.01 (0.03)	$\mathrm{EM_{Iso}}<$ 0.07 (0.05) $\mathrm{HAD_{Iso}}<$ 0.10 (0.025)	$\Delta\eta_{\mathrm{in}}<$ 0.004 (0.007) $\Delta\phi_{\mathrm{in}}<$ 0.06 (0.03)	$\mathrm{TRK_{Iso}}<$ 0.09 (0.04)

Table 7.5: Electron identification selections applied in the HLT for electrons in the EB (and EE). The values in the "Electron ID name" column correspond to parts of the HLT trigger names listed in Table 7.4. The variables listed here are defined in Section 6.1.

The names of the triggers listed in Table 7.4 correspond to the algorithms used to reconstruct and identify the objects making up each leg. For the first period of data taking, up to run 178380, electrons with the identification selection `CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT` are used in the trigger with varying p_{T} requirements (the number after the word `Ele` refers to the electron p_{T} threshold in GeV). For the following period, from run 178420 until the end of the data taking, electrons with the identification selection `WP80` are used instead. The explicit cuts used in these electron identification selections are listed in Table 7.5.

During the period of running between run 165088 and run 178380, jets of type `CentralJet` are used in the trigger with varying p_{T} requirements (the number after the words `CentralJet`, `DiCentralJet`, or `DiCentralPFJet` refers to the jet p_{T} threshold in GeV). For the following period, from run 178420 until the end of the data taking, jets of type `CentralPFJet` are used in the trigger. Both types of jets are reconstructed using the anti-k_{T} algorithm, and both types of jets require $|\eta| < 2.6$. The name `CentralJet` refers to jets reconstructed from calorimeter input only. The name `CentralPFJet` refers to jets reconstructed with a PF algorithm that has been simplified from the full offline PF algorithm according to the computing restrictions in the HLT.

The efficiencies for different objects defining the trigger paths are discussed below.

- **Electrons:** For all of the triggers listed in Table 7.4, the electron p_{T} requirements are significantly lower than the electron p_{T} requirements in the $e\nu jj$ preselection, and threshold effects related to electron p_{T} are avoided. However, the ID and isolation requirements are different between online and offline selection and therefore the trigger efficiency should be checked. The efficiency with which the `CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT` and `WP80` electron legs select electrons that are identified offline as HEEP electrons was measured by using $\mathrm{Z}^0 \rightarrow$ ee events. This efficiency averaged over the whole 2011 data taking period is found to be $95.5^{+1.0\%}_{-3.5\%}$. The quoted systematic uncertainties are due to variations observed in the efficiency of the HLT `WP80` identification as a function of electron η which are not taken into account in this analysis; the statistical uncertainties are negligible. The efficiency of the electron leg of the `HLT_Ele27_WP80_DiCentralPFJet25_PFMHT15` trigger as a function of electron pseudorapidity is shown in Figure 7.11. Overall, the efficiency was found to be stable as a function of number of pile-up interactions: the small variations observed are well within the systematic uncertainties already quoted above. This 4.5% overall inefficiency is only included in the signal predictions (to correct for the fact that these trigger requirements are not explicitly applied in the MC). This inefficiency is not propagated to the major backgrounds, since their normalization is already derived from comparison with data in control regions.

- **Jets:** The p_{T} cut on the jets at $e\nu jj$ preselection is 40 GeV, tight enough to avoid trigger threshold bias produced by the jet p_{T} requirements online. The jet legs of the triggers in Table 7.4 are found to be fully efficient.

- **PFMHT**: The $\not\!\!E_{\mathrm{T}}$ cut at $e\nu jj$ preselection is 55 GeV: tight enough to avoid trigger threshold bias produced by the `PFMHT` requirements online. The PFMHT legs of the triggers in Table 7.4 are found to be fully efficient.

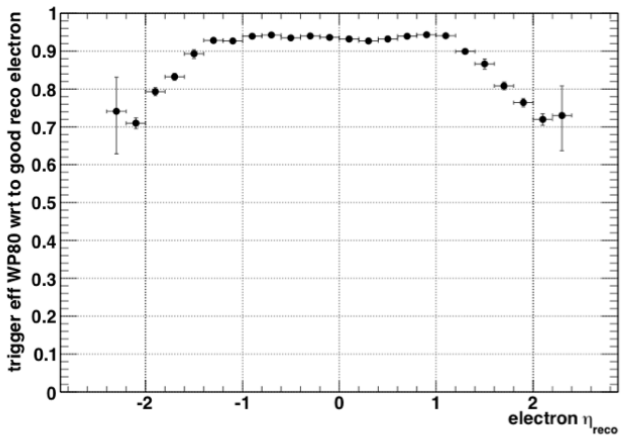


Figure 7.11: Efficiency with which the `Ele27_WP80` leg of the `HLT_Ele27_WP80_DiCentralPFJet25_PFMHT15` trigger identifies an electron with $p_T > 40$, if that same electron is identified offline as a HEEP electron. The efficiency is shown as a function of electron pseudorapidity.

7.3.2 Preselection

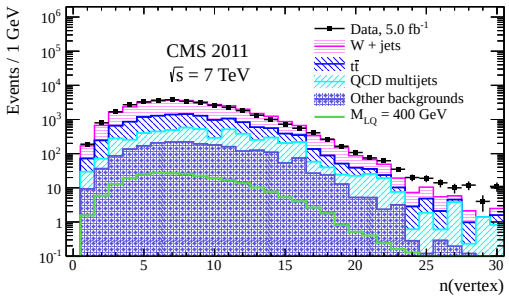


Figure 7.12: The distribution of the number of primary vertices for events passing the $e\nu jj$ preselection.

A sample of events enriched in SM background processes is selected to verify the background estimate in the $e\nu jj$ channel. This $e\nu jj$ preselection proceeds with the following kinematics cuts (all the reconstructed objects are required to pass the selection criteria described in Chapter 6):

- Pass the event filters listed in Section 7.1

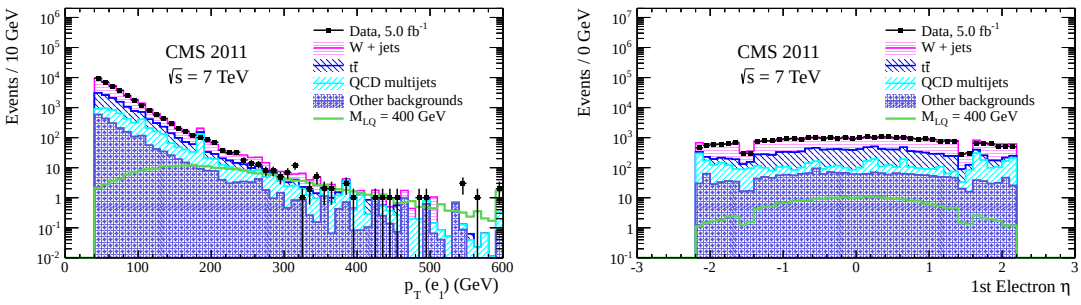


Figure 7.13: The p_{T}(left) and η (right) distributions of the leading (in p_{T}) electron for events passing the $e\nu jj$ preselection.

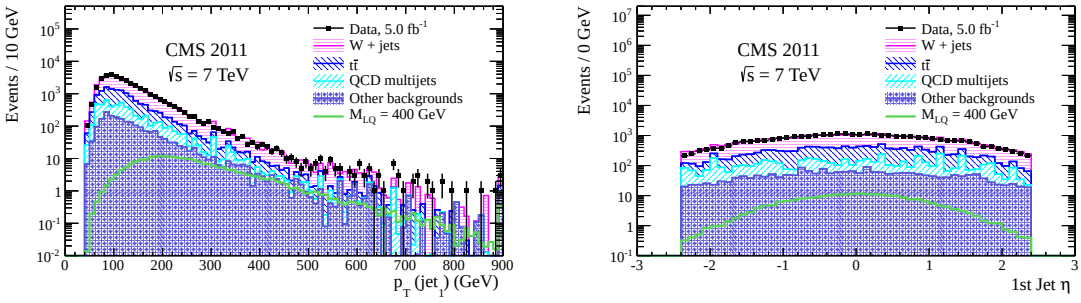


Figure 7.14: The p_{T}(left) and η (right) distributions of the leading (in p_{T}) jet for events passing the $e\nu jj$ preselection.

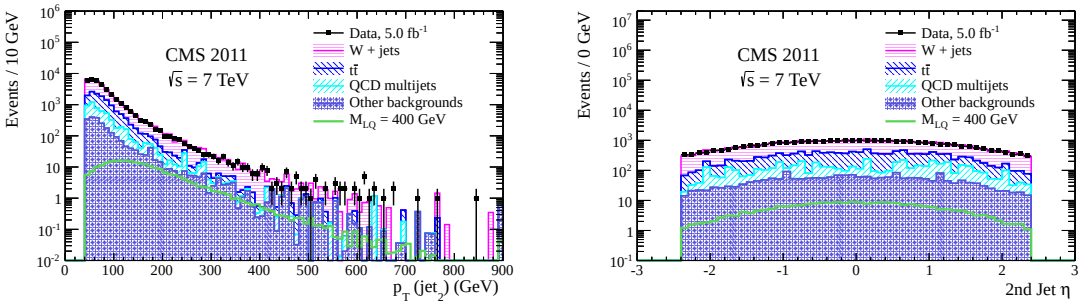


Figure 7.15: The p_{T}(left) and η (right) distributions of the second leading (in p_{T}) jet for events passing the $e\nu jj$ preselection.

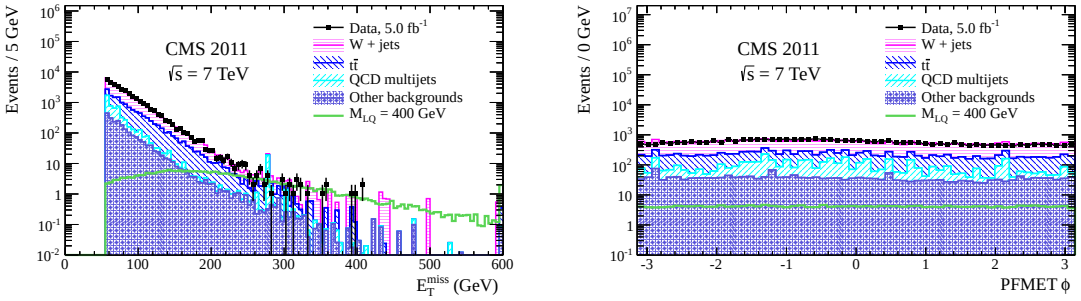


Figure 7.16: The $\not\!\!E_{\mathrm{T}}$(left) and $\phi(\not\!\!E_{\mathrm{T}})$ (right) distributions for events passing the $e\nu jj$ preselection.

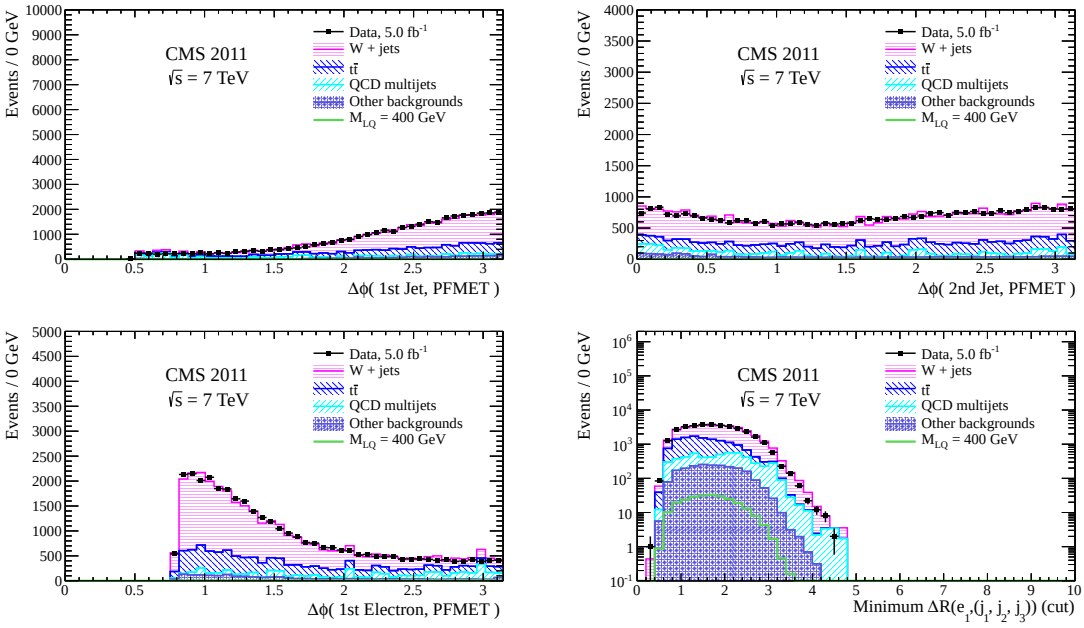


Figure 7.17: The distribution of $\Delta|phi(\not\!\!E_{\mathrm{T}}, \mathrm{j1})$ (top left), $\Delta|phi(\not\!\!E_{\mathrm{T}}, \mathrm{j2})$ (top right), $\Delta|phi(\not\!\!E_{\mathrm{T}}, \mathrm{e})$ (bottom left), and $\min\Delta R(\mathrm{e}, \mathrm{jets})$ (bottom right) for events passing the $e\nu jj$ preselection.

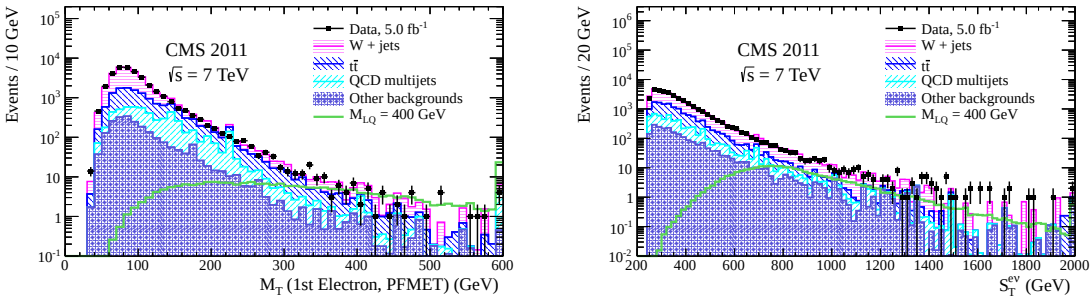


Figure 7.18: The $m_{\mathrm{T,\,e}\nu}$(left) and S_{T}(right) distributions for events passing the $e\nu jj$ preselection.

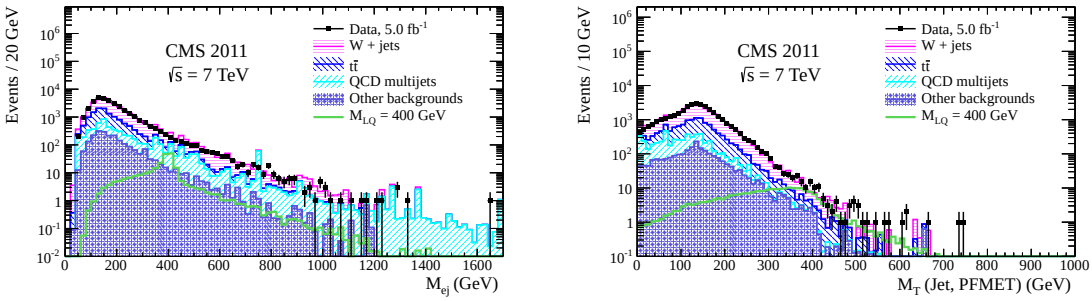


Figure 7.19: The m_{ej}(left) and $m_{\mathrm{T}\nu\mathrm{j}}$ (right) distributions for events passing the $e\nu jj$ preselection.

- Pass the signal triggers listed in Table 7.4
- Exactly 1 HEEP electron with p_{T}> 40 GeVand $|\eta| < 2.1$
- At least 2 jets with p_{T}> 40 GeVand $|\eta| < 2.4$
- $\not\!\!E_{\mathrm{T}} > 55$ GeV
- No selected muons passing the ID described in Section 6.2
- $\Delta\phi(\not\!\!E_{\mathrm{T}}, \mathrm{e})$> 0.8;
- $\Delta\phi(\not\!\!E_{\mathrm{T}}, \mathrm{j1})$> 0.5;
- min$\Delta R(\mathrm{e}, \mathrm{jets})$> 0.7;
- S_{T}> 250 GeV

The p_{T} cut on the electrons is chosen to be high enough for the triggers listed in Table 7.1 to be fully efficient. An additional $|\eta|$ cut is applied to the electrons in order to reduce the contribution of QCD multijet events to the total background; this cut has a negligable effect on the signal efficiency. The $|\eta|$ cut on the jets rejects potential fake jets reconstructed from anomalous signals in the HF ($|\eta| > 3.0$); this cut has a negligable impact on the leptoquark signal efficiency. The value of the $\not\!\!E_{\mathrm{T}}$ cut is set so that the signal triggers described in Table 7.4 are fully efficient. $\Delta\phi(\not\!\!E_{\mathrm{T}}, \mathrm{e})$ and $\Delta\phi(\not\!\!E_{\mathrm{T}}, \mathrm{j1})$ are the opening angles between the $\not\!\!E_{\mathrm{T}}$ and the electron and leading jet in p_{T}, respectively. min$\Delta R(\mathrm{e}, \mathrm{jets})$ is the separation in ΔR between the electron and the leading jet in p_{T}. All three of these cuts were included to reduce the contribution of QCD multijet events to the total background, and the values were chosen to have a negligable impact on the leptoquark signal efficiency. The muon veto rejects backgrounds from $t\bar{t}$ events with an $e\mu jj$ final state. The S_{T} cut is looser than the cuts applied in the final selection of the $e\nu jj$ channel.

At this stage of the selection, there is sufficient data to compare with the background predictions for all the observables employed in the final event selection. The distribution

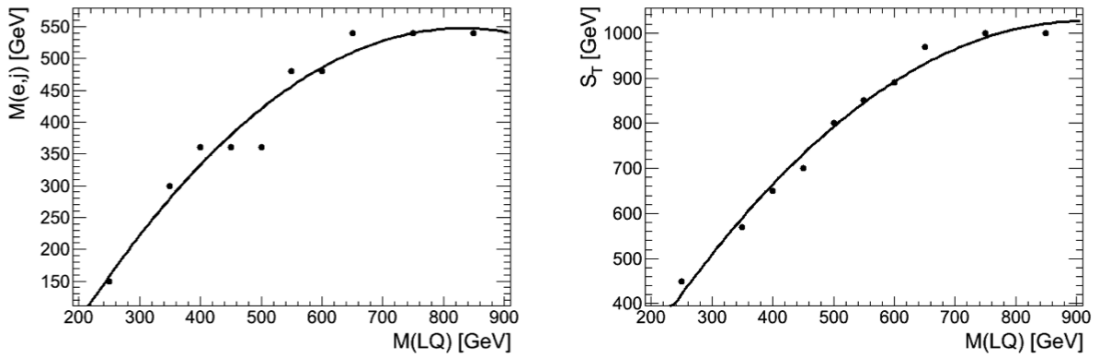


Figure 7.20: Optimized thresholds for for m_{ej} (left) and S_{T} (right) cuts as a function of leptoquark mass in the $e\nu jj$ channel. The leptoquark mass hypothesis under consideration is shown on the x-axis. The optimized cut thresholds are shown on the y-axis. The resulting distributions are fit with a second-degree polynomial.

of the number of reconstructed primary vertices is shown in Figure 7.12. The p_{T} and η distributions of the electron and the two leading jets are shown in Figures 7.13, 7.14, and 7.15. The $\not\!\!E_{\mathrm{T}}$ distribution is shown in Figure 7.16. The $\Delta\phi(\not\!\!E_{\mathrm{T}}, \mathrm{e})$, $\Delta\phi(\not\!\!E_{\mathrm{T}}, \mathrm{j1})$, $\Delta\phi(\not\!\!E_{\mathrm{T}}, \mathrm{j2})$, and $\min\Delta R(\mathrm{e}, \mathrm{jets})$ distributions are shown in Figure 7.17. The $m_{\mathrm{T,\,e\nu}}$ and S_{T} distributions are shown in Figure 7.18. The m_{ej} and $m_{\mathrm{T}\nu\mathrm{j}}$ distributions are shown in Figure 7.19.

7.3.3 Selection optimization

The final event selection criteria are optimized by maximizing the expected signal significance defined as $S/\sqrt{S+B}$, where S (B) is the expected number of signal (background) events passing the selection requirements. An additional requirement of $m_{\mathrm{T,\,e\nu}} > 120$ GeV is imposed to remove background events with on-shell Ws. Then, three variables (S_{T}, $\not\!\!E_{\mathrm{T}}$, and m_{ej}) are optimized simultaneously by scanning appropriate ranges of values. Figure 7.20 shows the smooth dependence of the optimized m_{ej} and S_{T} cut values on the leptoquark mass hypothesis. The results of the optimization, obtained for each LQ mass hypothesis, are summarized in Table 7.6 assuming an integrated luminosity of 4.95 fb^{-1}.

M_{LQ} [GeV]	250	350	400	450	500	550	600	650	750	850
S_{T} [GeV]	450	570	650	700	800	850	890	970	1000	1000
$\not\!\!E_{\mathrm{T}}$ [GeV]	100	120	120	140	160	160	180	180	220	240
m_{ej} [GeV]	150	300	360	360	360	480	480	540	540	540

Table 7.6: Optimized selection criteria for the $e\nu jj$analysis for different LQ mass hypotheses. The requirement $m_{\mathrm{T},\,e\nu} >$ 120 GeVis applied to all mass hypotheses.

M_{LQ}	LQ Signal	W+Jets	$t\bar{t}$	QCD	Other	Data	Total BG
Presel	-	20108 ± 99	9301 ± 42	3267 ± 26	1913 ± 53	34135	34590 ± 120
250	1703.5 ± 13.8	190 ± 9.2	195 ± 6.2	31.7 ± 1.2	43.3 ± 2.19	456	460.4 ± 11.4
350	285.54 ± 2.06	52.1 ± 4.8	34.2 ± 2.6	15.2 ± 0.97	10.9 ± 0.92	89	112.4 ± 5.6
400	126.01 ± 0.82	28.7 ± 3.6	17.5 ± 1.8	6.20 ± 0.46	6.01 ± 0.77	43	58.4 ± 4.1
450	68.38 ± 0.43	19.7 ± 2.9	12.2 ± 1.5	3.01 ± 0.31	4.13 ± 0.44	30	39.1 ± 3.3
500	34.70 ± 0.23	13.3 ± 2.4	6.3 ± 1.1	1.72 ± 0.22	2.80 ± 0.37	20	24.2 ± 2.6
550	16.25 ± 0.10	2.98 ± 0.95	3.38 ± 0.82	0.65 ± 0.10	1.46 ± 0.26	9	8.5 ± 1.3
600	9.442 ± 0.056	2.45 ± 0.87	2.33 ± 0.67	0.57 ± 0.10	1.29 ± 0.25	7	6.6 ± 1.1
650	5.202 ± 0.032	2.03 ± 0.83	1.01 ± 0.41	0.335 ± 0.079	0.76 ± 0.20	5	4.14 ± 0.95
750	1.851 ± 0.010	1.45 ± 0.65	0.62 ± 0.31	0.287 ± 0.080	0.65 ± 0.18	5	3.01 ± 0.75
850	0.6973 ± 0.0037	1.22 ± 0.61	0.62 ± 0.31	0.251 ± 0.078	0.61 ± 0.19	5	2.70 ± 0.71

Table 7.7: Number of events after the final $e\nu jj$ selection. Only statistical uncertainties are reported.

7.3.4 Final selection

Table 7.7 shows the number of events for the data, the backgrounds, and the LQ signal, after applying the final, optimized $e\nu jj$ selection criteria summarized in Table 7.6. Figures 7.21 and 7.22 show the distributions of S_{T} and the selected electron-jet invariant mass, m_{ej}, after the full selection optimized for M_{LQ}= 400 and 600 GeV, respectively. The dominant background contributions are from $t\bar{t}$ and W+jets events, while the contribution from the other backgrounds is below 20% for the LQ masses within the current reach of this analysis. Good agreement is observed between the data and the background prediction, within uncertainties.

7.3.5 Final selection with updated ECAL calibration

One of the 5 events (run 178708, luminosity section 326, event number 532162137) passing the $e\nu jj$ final selection for LQ mass of 850 GeV reported in Table 7.7 is affected by a serious

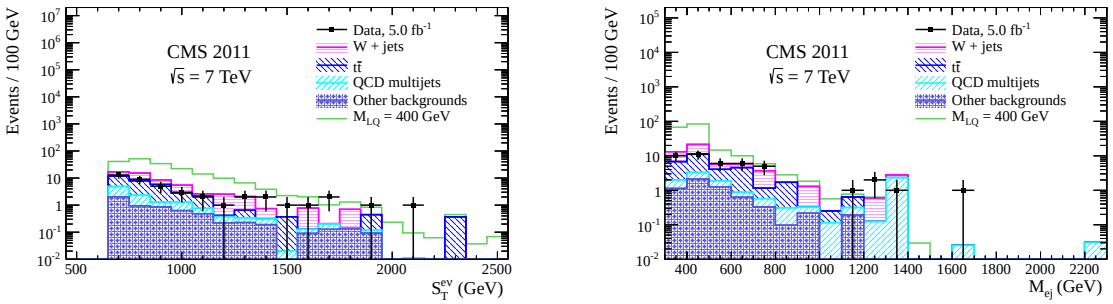


Figure 7.21: The S_{T}(left) and m_{ej}(right) distributions for events passing the full $e\nu jj$selection optimized for M_{LQ}= 400 GeV.

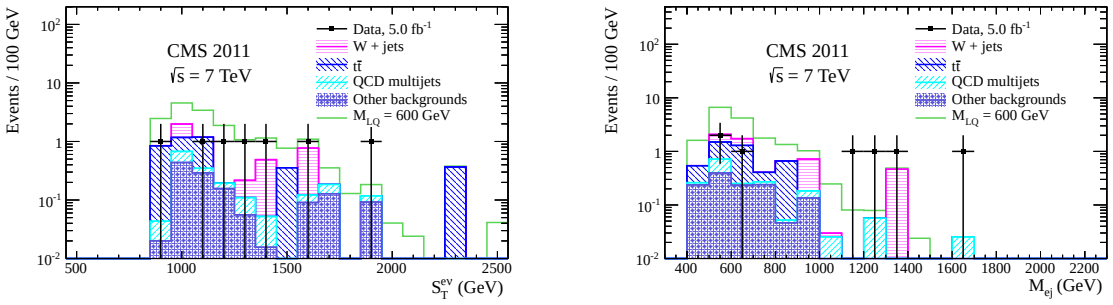


Figure 7.22: The S_{T}(left) and m_{ej}(right) distributions for events passing the full $e\nu jj$selection optimized for M_{LQ}= 600 GeV.

calibration problem in one EE channel. This channel reports a calibration coefficient with a value greater than 3.0, while the expected value is 1.0. The calibration problem is present in the conditions used to reconstruct the dataset /ElectronHad/Run2011B-PromptReco-v1/AOD listed in Table 5.2, which has been used to produce the $e\nu jj$ analysis results shown so far. It is likely that this event contains a real electron with p_{T} of about 200 GeV. The electron in this event is measured by the EE channel affected by the problem. Due to the wrong calibration factor, the electron p_{T} is incorrectly reconstructed at approximately 550 GeV. This mismeasured electron generates mismeasured $\not\!\!E_{\mathrm{T}}$ (approximately 400 GeV) on the opposite side (in ϕ) of the detector. An event display highlighting this event is shown in Figure 7.23 (left).

This calibration problem was identified and fixed in a re-reconstruction of the data that was made available at the end of November 2011, known as the "Nov30" re-reconstruction. An event display highlighting this event under the Nov30 re-reconstruction is shown in Figure 7.23 (right). In the Nov30 re-reconstruction, the electron p_{T} for this event is reduced significantly compared to the value from the original reconstruction, and the $\not\!\!E_{\mathrm{T}}$ is now about 40 GeV. Under the Nov30 re-reconstruction, the event does not even pass the $e\nu jj$ preselection.

Unfortunately, the ElectronHad dataset used $e\nu jj$ channel (see Table 5.2) was not centrally reprocessed under the Nov30 re-reconstruction. Therefore, the 1576 events in data that pass the $e\nu jj$ selection optimized for a LQ with mass of 250 GeV (see Table 7.7) were privately reprocessed using the conditions from the Nov30 re-reconstruction.

This section presents the results of the final $e\nu jj$ selection in the same format reported in Section 7.3.4, but using the 1576 events in data that were privately re-reconstructed. The background predictions remain unchanged. The bias introduced by the fact that the whole ElectronHad dataset has not been reprocessed is negligible in this case. It is much more likely that 1 (originally common) background-like event would be promoted to pass the $e\nu jj$ final selection due to a similar calibration problem than it is likely that 1 (rare) signal-like event is removed from the final selection. This is because there are many more background-like

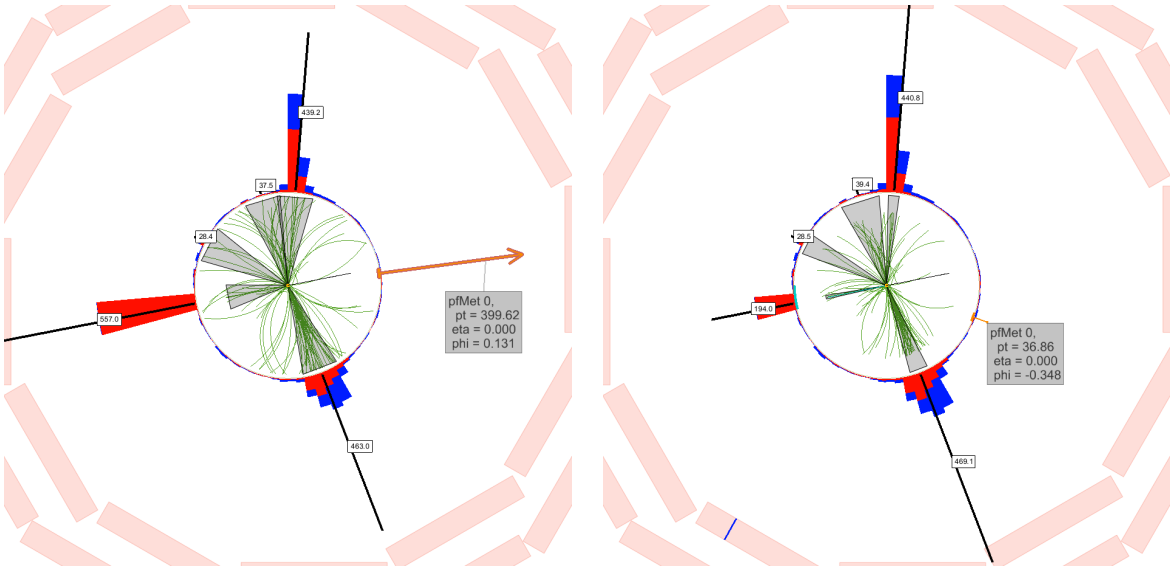


Figure 7.23: Event displays showing the event discussed in this section [run 178708, luminosity section 326, event number 532162137] as reconstructed in the original reconstruction (left) and in the Nov30 re-reconstruction (right). An ECAL calibration problem that affected this event in the original reconstruction was fixed in the Nov30 re-reconstruction. The electron affected by the calibration problem is represented as a red energy cluster in the ECAL on the left side of both event displays. The $\not\!\!E_{\mathrm{T}}$ measurement is represented as a red arrow on the right side of both event displays.

events than signal-like events. It is expected, therefore, that the main effect will be to see fewer events passing the $e\nu jj$ selection criteria in the private re-reconstruction than in the original reconstruction. In addition, because the presence of fake $\not\!\!E_{\rm T}$ is a typical consequence of an electron measurement problem, the $eejj$ channel (where no $\not\!\!E_{\rm T}$ requirement is applied) is believed to be unaffected. As a cross check, the 6 events passing the final $eejj$ selection optimized for LQ mass of 650 GeV (see Table 7.3) were also privately re-reconstructed; no significant difference between the private re-reconstruction and the original reconstruction was observed.

Table 7.8 shows the number of events for the data, the backgrounds, and the LQ signal, after applying the final, optimized $e\nu jj$ selection criteria summarized in Table 7.6. Figures 7.24 and 7.25 show the distributions of $S_{\rm T}$ and the selected electron-jet invariant mass, $m_{\rm ej}$, after the full selection optimized for $M_{\rm LQ}$= 400 and 600 GeV, respectively. The dominant background contributions are from $t\bar{t}$ sand W+jets events, while the contribution from the other backgrounds is below 20% for the leptoquark masses within the current reach of this analysis. A good agreement is observed between the data and the background prediction within statistical uncertainties.

As expected, the event affected by the calibration problem (run 178708, luminosity section 326, event number 532162137) is removed from the final $e\nu jj$ selection with the new ECAL calibration. For the other events, there are only minor changes between the results shown in Section 7.3.4 and those presented here.

M_{LQ}	LQ Signal	W+Jets	$t\bar{t}$	QCD	Other	Data	Total BG
250	1703.5 ± 13.8	190 ± 9.2	195 ± 6.2	31.7 ± 1.2	43.3 ± 2.19	430	460.4 ± 11.4
350	285.54 ± 2.06	52.1 ± 4.8	34.2 ± 2.6	15.2 ± 0.97	10.9 ± 0.92	92	112.4 ± 5.6
400	126.01 ± 0.82	28.7 ± 3.6	17.5 ± 1.8	6.20 ± 0.46	6.01 ± 0.77	43	58.4 ± 4.1
450	68.38 ± 0.43	19.7 ± 2.9	12.2 ± 1.5	3.01 ± 0.31	4.13 ± 0.44	29	39.1 ± 3.3
500	34.70 ± 0.23	13.3 ± 2.4	6.3 ± 1.1	1.72 ± 0.22	2.80 ± 0.37	18	24.2 ± 2.6
550	16.25 ± 0.10	2.98 ± 0.95	3.38 ± 0.82	0.65 ± 0.10	1.46 ± 0.26	10	8.5 ± 1.3
600	9.442 ± 0.056	2.45 ± 0.87	2.33 ± 0.67	0.57 ± 0.10	1.29 ± 0.25	6	6.6 ± 1.1
650	5.202 ± 0.032	2.03 ± 0.83	1.01 ± 0.41	0.335 ± 0.079	0.76 ± 0.20	4	4.14 ± 0.95
750	1.851 ± 0.010	1.45 ± 0.65	0.62 ± 0.31	0.287 ± 0.080	0.65 ± 0.18	4	3.01 ± 0.75
850	0.6973 ± 0.0037	1.22 ± 0.61	0.62 ± 0.31	0.251 ± 0.078	0.61 ± 0.19	4	2.70 ± 0.71

Table 7.8: Number of events after the final $e\nu jj$ selection following a private re-reconstruction. Only statistical uncertainties are reported.

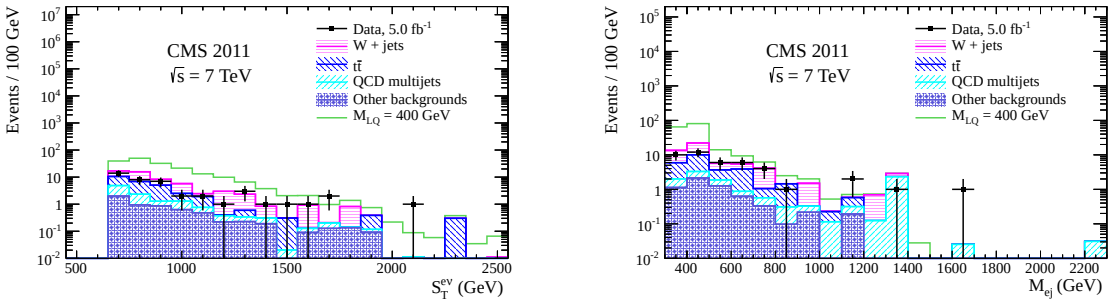


Figure 7.24: The S_{T}(left) and m_{ej}(right) distributions for events passing the full $e\nu jj$selection optimized for M_{LQ}= 400 GeV. Nov30 conditions are used for the data.

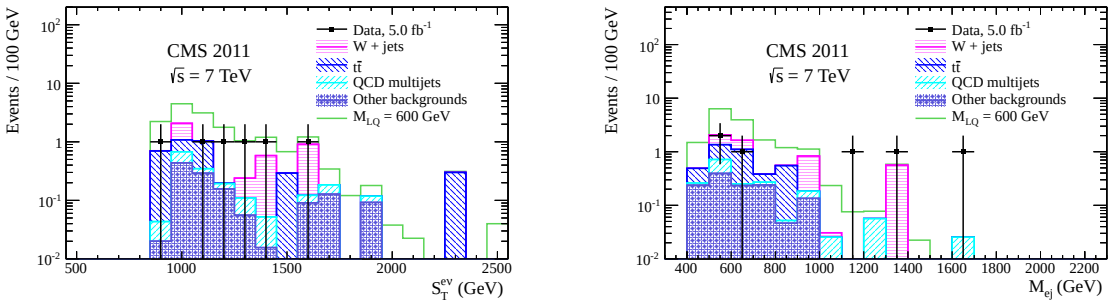


Figure 7.25: The S_{T}(left) and m_{ej}(right) distributions for events passing the full $e\nu jj$selection optimized for M_{LQ}= 600 GeV. Nov30 conditions are used for the data.

CHAPTER 8

ANALYSIS: STANDARD MODEL BACKGROUNDS

This chapter describes the various background estimates that are used throughout this thesis for both the $eejj$ and $e\nu jj$ channels. All of the dominant backgrounds are either purely data driven or are derived from MC samples that have been rescaled to match data within control regions. Section 8.1 describes the data driven QCD multijet background estimate. Section 8.2 describes the $t\bar{t}$ estimates for the $eejj$ channel (purely data driven) and the $e\nu jj$ channel (derived from MC samples rescaled within a control region). Section 8.3 describes the Z^0+jets and W+jets background estimates (also derived from MC samples rescaled within a control region) and other background estimates (derived from MC samples scaled only by cross sections listed in Section 5.2).

8.1 Multijets background

This section describes the data driven estimate of the contribution of QCD multijet events to the background in both the $eejj$ and $e\nu jj$ channels. In both channels, QCD multijet events may only contribute to the background if one or more jets is misidentified as a HEEP electron. Section 8.1.1 gives an overview of the method for this estimate. Section 8.1.2 describes the rates with which jets can be misidentified as a HEEP electron. Section 8.1.3 describes how

these misidentification rates are applied in both the $eejj$ and $e\nu jj$ channels, and Section 8.1.4 describes a closure test used as a cross check for this background estimate.

8.1.1 Method

The QCD multijet background in both $eejj$ and $e\nu jj$ channels is determined from data using a fake rate method. Two data samples, "loose $eejj$" and "loose $e\nu jj$", dominated by QCD multijet events are selected. In both samples, electrons are required to pass "loose" identification requirements instead of the "tight" HEEP selection described in Section 6.1. The other jet and $\not\!\!E_{\mathrm{T}}$ requirements and kinematic selection criteria remain unchanged.

The number of QCD multijet events in the $eejj$ sample, N^{QCD}_{eejj}, at a given stage of the selection is estimated by Equation 8.1:

$$N^{QCD}_{eejj} = \sum_{\text{loose } eejj \text{ events in data}} P(e_{1,\text{ tight}}|e_{1,\text{ loose}} : p_T, \eta) \cdot P(e_{2,\text{ tight}}|e_{2,\text{ loose}} : p_T, \eta) \tag{8.1}$$

Similarly, the number of QCD multijet events in the $e\nu jj$ sample, $N^{QCD}_{e\nu jj}$, at a given stage of the selection is estimated by Equation 8.2:

$$N^{QCD}_{e\nu jj} = \sum_{\text{loose } e\nu jj \text{ events in data}} P(e_{\text{tight}}|e_{\text{loose}} : p_T, \eta) \tag{8.2}$$

where e_{loose} is a GSF electron passing loose electron identification criteria described in Table 8.1, e_{tight} is an electron passing the HEEP ID and isolation criteria described in Section 6.1, and $P(e_{\text{loose}}|e_{\text{tight}} : p_T, \eta)$ is the probability, or fake rate, that a loose electron e_{loose} passes the HEEP ID and isolation requirements. The sum is performed over all of the events in the loose $eejj$ and $e\nu jj$ data samples that pass the event selection.

The events in the loose $e\nu jj$ and $eejj$ samples are selected online by prescaled single photon triggers with different E_{T} thresholds (depending on the run range), as shown in Table 8.2. These triggers require an energy cluster in the ECAL with a reconstructed E_{T} greater

Variable	Barrel (EB) criterion	Endcap (EE) criterion
$\sigma_{i\eta i\eta}$	< 0.013	< 0.034
H/E	< 0.15	< 0.1

Table 8.1: Loose identification criteria for gsf electrons needed for the QCD multijet background estimation. These criteria are used to select the loose $eejj$ and $e\nu jj$ samples. The variables listed in the leftmost column are defined in Section 6.1.

than a threshold in GeV indicated by the number after the word `Photon` in the trigger name. In addition, the triggers in Table 8.2 with the word `CaloIdVL` in their name also require the absence of significant energy deposits in the HCAL cells directly behind the ECAL cluster ($H/E<$ 0.15 for electrons in the EB, $H/E<$ 0.10 for electrons in the EE) and the shape of the ECAL cluster must be consistent with the shape of clusters typically produced by electrons and photons ($\sigma_{i\eta i\eta}<$ 0.024 for electrons in the EB, $\sigma_{i\eta i\eta}<$ 0.040 for electrons in the EE).

Depending on the E_{T} of the triggered photon, each event passing the loose $eejj$ or $e\nu jj$ selection is reweighted in the sum of Equations 8.1 and 8.2 with a weight equal to the lowest trigger prescale among the single photon triggers that fired in that event. The E_{T} threshold of the lowest single photon trigger (`HLT_Photon30_CaloIdVL`) is 30 GeV. This value is 10 GeV below the electron p_{T} used in preselection for both analyses: enough to avoid any trigger threshold bias.

8.1.2 Fake rate calculation

The fake rate, $P(e_{\mathrm{tight}}|e_{\mathrm{loose}} : p_T, \eta)$, is defined as the ratio between the number of HEEP electrons, $N_{e,\mathrm{tight}}$, and the number of loose electrons, $N_{e,\mathrm{loose}}$ identified with the criteria listed in Table 8.1, in a sample of events in data passing the following selection criteria:

1. The event must fire a single photon trigger, as described in Table 8.2
2. There must be exactly one loose electron, selected using the criteria described in Table 8.1

HLT path	Run number range	Effective $\mathcal{L}_{int}(\text{pb}^{-1})$
HLT_Photon30_CaloIdVL_v1	160431–161176	0.145661
HLT_Photon30_CaloIdVL_v2	161217–163261	0.341693
HLT_Photon30_CaloIdVL_v3	163270–163869	0.778178
HLT_Photon30_CaloIdVL_v4	165088–165633	0.441692
HLT_Photon30_CaloIdVL_v5	165970–166967	0.573131
HLT_Photon30_CaloIdVL_v6	167039–173198	0.909814
HLT_Photon30_CaloIdVL_v7	173236–173692	0.537542
HLT_Photon30_CaloIdVL_v8	175860–180252	1.176
HLT_Photon50_CaloIdVL_v1	165088–165633	1.734
HLT_Photon50_CaloIdVL_v2	165970–166967	2.414
HLT_Photon50_CaloIdVL_v3	167039–173198	5.266
HLT_Photon50_CaloIdVL_v4	173236–180252	10.63
HLT_Photon75_CaloIdVL_v1	160431–161176	6.111
HLT_Photon75_CaloIdVL_v2	161217–163261	40.565
HLT_Photon75_CaloIdVL_v3	163270–163869	168.23
HLT_Photon75_CaloIdVL_v4	165088–165633	24.128
HLT_Photon75_CaloIdVL_v5	165970–166967	14.116
HLT_Photon75_CaloIdVL_v6	167039–173198	32.136
HLT_Photon75_CaloIdVL_v7	173236–180252	49.056
HLT_Photon90_CaloIdVL_v1	165088–165633	46.065
HLT_Photon90_CaloIdVL_v2	165970–166967	28.232
HLT_Photon90_CaloIdVL_v3	167039–173198	64.592
HLT_Photon90_CaloIdVL_v4	173236–180252	163.279
HLT_Photon125_v1	165088–165633	136.382
HLT_Photon125_v2	165970–166967	535.17
HLT_Photon135_v1	167039–173198	1107.0
HLT_Photon135_v2	173236–180252	2958.0
HLT_Photon400_v1	167039–173198	1107.0
HLT_Photon400_v2	173236–180252	2958.0

Table 8.2: Single photon HLT paths used for the QCD multijet background estimation.

3. There must be N_{jet} or more jets with $p_T > 40$ GeV, with N_{jet} equal to 0, 1, 2, or 3

4. $\Delta R(e_{\text{loose}}, \text{jets}) > 0.7$

5. $\not{E}_{\text{T}} < 30$ GeV

The requirements of exactly one loose electron and $\not{E}_{\text{T}}$ reduce the contamination of real electrons in the sample from $Z^0 \to ee$ and $W \to e\nu$ events. The fake rates are calculated for the EB and for two separate regions of the EE ($|\eta| < 2.0$ and $|\eta| > 2.0$). In addition, each set of fake rates is obtained for three different jet multiplicity requirements, as described in cut number 3, above. The contamination of real electrons in both $N_{e,\text{tight}}$ and $N_{e,\text{loose}}$ is subtracted from the data using MC predictions. The MC samples are normalized using the cross sections listed in Section 5.2. The MC-corrected fake rate for the $N_{\text{jet}} \geq 2$ case is shown in Figure 8.1 as a function of the loose electron p_{T}, separately for the barrel and the two endcap regions.

Following the approach of the Z'→ ee analysis [115], each fake rate is fit with a first-degree polynomial in the region $p_{\text{T}} < 100$ GeV and a zero-degree polynomial in the region $p_{\text{T}} > 100$ GeV. The fit results are reported in Table 8.3 for the four cases under consideration: $N_{\text{jet}} \geq 0$, $N_{\text{jet}} \geq 1$, $N_{\text{jet}} \geq 2$, and $N_{\text{jet}} \geq 3$. The value of the fake rate decreases with increasing jet multiplicity. The fake rates that are used for the $eejj$ and $e\nu jj$ QCD multijet background estimation are from the $N_{\text{jet}} \geq 2$ case.

8.1.3 Background estimate

At each step of the $eejj$ and $e\nu jj$ selections, the shape of the kinematic distributions and the normalization of the QCD multijet background are estimated with the method described in Section 8.1.1, using Equations 8.1 and 8.2. Tables 7.3 and 7.7 report the QCD multijet contribution at different stages of the event selection for each analysis using the $N_{\text{jet}} \geq 2$ fake rates. After the final selection, the QCD multijet contribution is $\approx 1\%$ ($\approx 8\%$) of the total

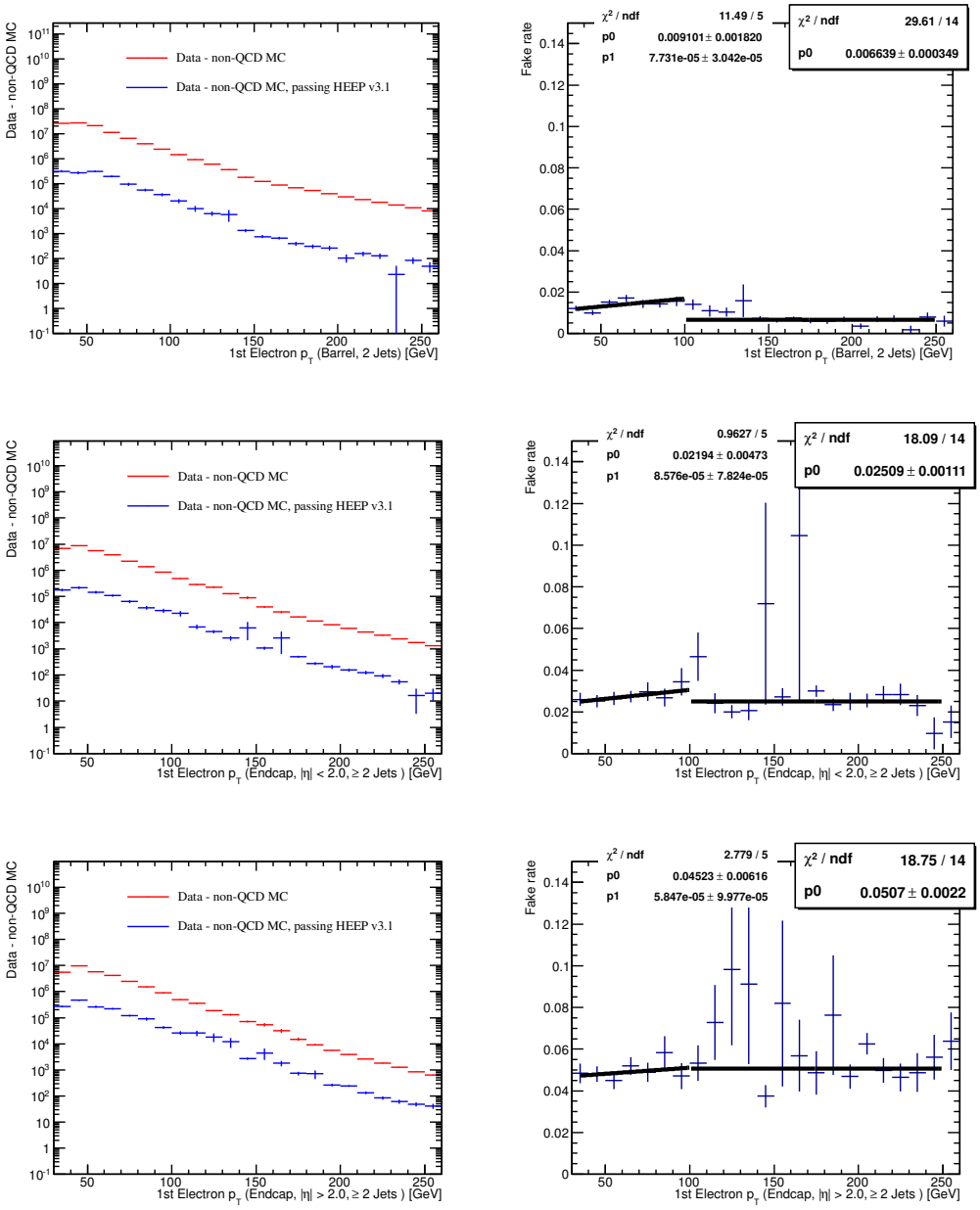


Figure 8.1: The p_{T} distribution of the "loose" and "tight" electrons in the fake rate sample (left) and the corresponding fitted fake rate vs electron p_{T} (right) for selected events with $N_{\mathrm{jet}} \geq 2$. The results are shown for the barrel and the two endcap regions. The non-QCD contribution is subtracted from the electron p_{T} distributions.

Selection	p_0	p_1	p_2
inclusive, barrel	2.301e-02 ± 1.745e-03	-6.570e-05 ± 2.888e-05	7.574e-03 ± 3.228e-04
inclusive, endcap ($\|\eta\| < 2.0$)	4.060e-02 ± 3.991e-03	-1.005e-04 ± 6.649e-05	2.680e-02 ± 1.005e-03
inclusive, endcap ($\|\eta\| > 2.0$)	6.313e-02 ± 5.103e-03	-1.033e-04 ± 8.795e-05	5.270e-02 ± 2.020e-03
1 jet, barrel	1.783e-02 ± 1.595e-03	3.067e-06 ± 2.721e-05	7.585e-03 ± 3.227e-04
1 jet, endcap ($\|\eta\| < 2.0$)	3.673e-02 ± 3.837e-03	-4.838e-05 ± 6.475e-05	2.681e-02 ± 1.005e-03
1 jet, endcap ($\|\eta\| > 2.0$)	6.020e-02 ± 4.995e-03	-6.142e-05 ± 8.666e-05	5.270e-02 ± 2.020e-03
2 jets, barrel	9.101e-03 ± 1.820e-03	7.731e-05 ± 3.042e-05	6.639e-03 ± 3.494e-04
2 jets, endcap ($\|\eta\| < 2.0$)	2.194e-02 ± 4.734e-03	8.576e-05 ± 7.824e-05	2.509e-02 ± 1.109e-03
2 jets, endcap ($\|\eta\| > 2.0$)	4.523e-02 ± 6.162e-03	5.847e-05 ± 9.977e-05	5.070e-02 ± 2.197e-03
3 jets, barrel	1.538e-03 ± 2.476e-03	1.224e-04 ± 4.192e-05	4.753e-03 ± 4.328e-04
3 jets, endcap ($\|\eta\| < 2.0$)	8.950e-04 ± 6.942e-03	3.409e-04 ± 1.170e-04	2.531e-02 ± 1.292e-03
3 jets, endcap ($\|\eta\| > 2.0$)	1.863e-02 ± 8.961e-03	2.997e-04 ± 1.399e-04	4.799e-02 ± 2.416e-03

Table 8.3: Fake rates and relative fit parameters.
The fit functions are: $f(p_{\mathrm{T}}) = p_0 + p_1 \cdot p_{\mathrm{T}}$ for $p_{\mathrm{T}} < 100$, and $f(p_{\mathrm{T}}) = p_2$ for $p_{\mathrm{T}} > 100$.

background in the $eejj$ ($e\nu jj$) channels for leptoquark masses around the current reach of this search.

8.1.4 Closure test

A closure test is performed in order to validate the QCD multijet background estimation using the fake rate method. This closure test uses a control sample of QCD multijet events (with a contamination from non-QCD processes of $\approx 5\%$) selected by the following requirements:

1. the event must fire a single photon trigger, as described in Table 8.2
2. there must be exactly two loose electrons, selected using the criteria described in Table 8.1
3. there must be at least one jet with $p_T > 40$ GeV
4. $\Delta R(e_{\mathrm{loose}}, \mathrm{jets}) > 0.7$
5. $M_{ee} > 110$ GeV, to reduce contamination from $\mathrm{Z}^0 \rightarrow$ ee events

6. $S_{\mathrm{T}} > 200$ GeV, where S_{T} is defined as the scalar sum of the two loose electrons and the leading (in p_{T}) jet

A prediction is made of how many of these events will have one electron passing the HEEP identification and isolation criteria described in Section 6.1, $N^{QCD,pred.}_{\text{loose } e,\text{ tight } e,\ j}$, using the $N_{\mathrm{jet}} \geq 1$ fake rates calculated in Section 8.1.2 and following a similar procedure as the one described in the previous sections:

$$N^{QCD,pred.}_{\text{loose } e,\text{ tight } e,\ j} = \sum_{\text{loose } e,\text{ loose e, } j \text{ events}} P(e_{1,\text{ tight}}|e_{1,\text{ loose}} : p_T, \eta) + P(e_{2,\text{ tight}}|e_{2,\text{ loose}} : p_T, \eta) \tag{8.3}$$

This prediction is then compared with the actual number of events in data where one of the two loose electrons also passes the tight selection criteria, $N^{QCD,actual}_{\text{loose } e,\text{ tight } e,\ j}$. In this comparison, the contribution of $\approx 30\%$ from non-QCD multijet processes contaminating the actual "loose e, tight e, j" data sample is subtracted using an estimate from MC (the samples are normalized using the cross sections listed in Section 5.2).

The predicted and actual values of "loose e, tight e, j" events corresponding to 4.95 fb^{-1} of data are: $N^{QCD,pred.}_{\text{loose } e,\text{ tight } e,\ j} = 10730 \pm 941$ and $N^{QCD,actual}_{\text{loose } e,\text{ tight } e,\ j} = 14790 \pm 465$ (the latter after MC subtraction). Statistical errors derived from the uncertainties of fake rate fit parameters, and the uncertainties on the MC predictions used in the subtraction are included. The ratio between the two values is 0.73 ± 0.06. If an S_{T} [1] cut of 450 GeV is applied (corresponding to the lowest S_{T} cut applied in the $e\nu jj$ final selection), the values become $N^{QCD,pred.}_{\text{loose } e,\text{ tight } e,\ j} = 290 \pm 22$ and $N^{QCD,actual}_{\text{loose } e,\text{ tight } e,\ j} = 232 \pm 40$, and their ratio is 1.25 ± 0.23. Therefore, an uncertainty of 50% (25%) is quoted on this method at final selection level for the $eejj$ ($e\nu jj$) channel. Figure 8.2 shows the distributions of electron and jet reconstructed quantities for "loose e, tight e, j" events. The QCD multijet prediction using the fake rate method is compared with the actual data sample and an acceptable agreement is observed in both the shape and the normalization of the two samples within the quoted uncertainties.

[1] S_{T} is defined as the scalar sum of the two loose electrons and the leading jet

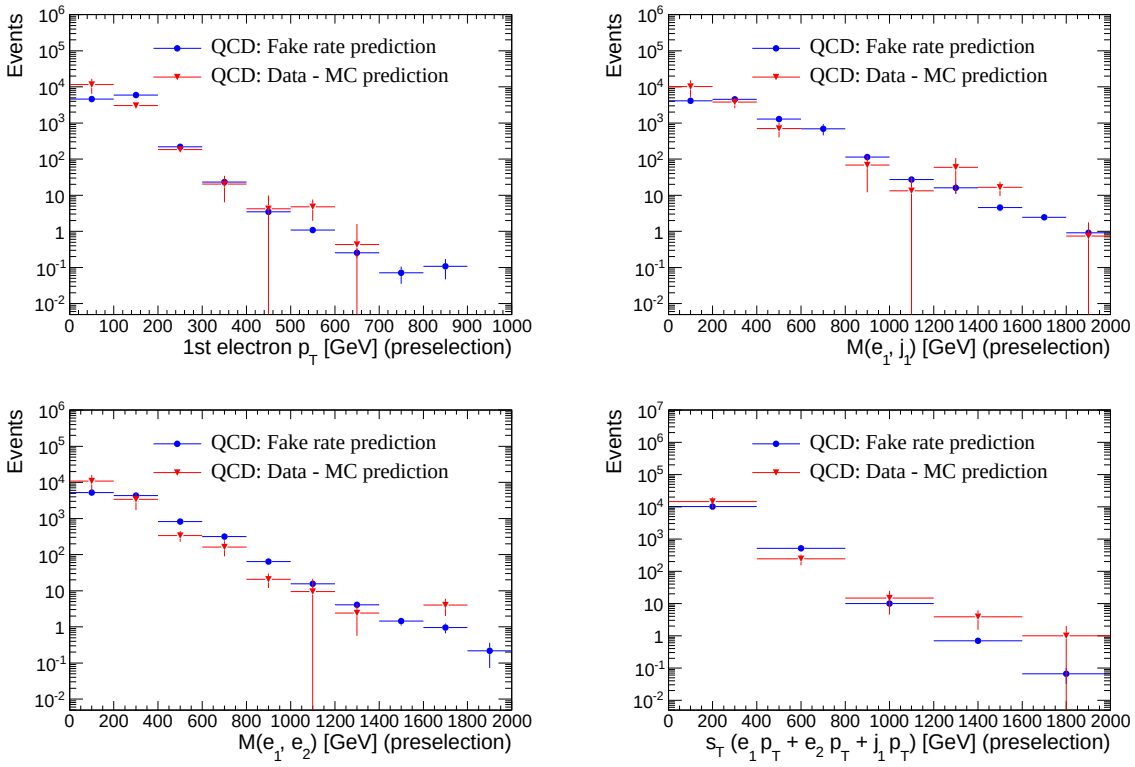


Figure 8.2: The leading electron p_{T} (top left), m_{ej} (top right), m_{ee}(bottom left) and S_{T} (bottom right) distributions for the "loose e, tight e, j" events. The QCD multijet prediction using the fake rate method is compared with the actual data sample (after the MC subtraction of non-QCD events for the latter sample).

8.2 $t\bar{t}$ background

This section describes the estimate of the contribution of $t\bar{t}$ events to the background in both the $eejj$ and $e\nu jj$ channels. In the $eejj$ channel, the $t\bar{t}$ contribution to the total background is estimated from data using $e\mu jj$ events. This process is described in Section 8.2.1 In the $e\nu jj$ channel, the $t\bar{t}$ contribution to the total background is estimated from MC events that have been rescaled within a control region, as described in Section 8.2.2.

8.2.1 $t\bar{t}$ background in the $eejj$ channel

HLT path	Run number range
Run2011A	
HLT_Mu15_Photon20_CaloIdL_v2	160404–161176
HLT_Mu15_Photon20_CaloIdL_v3	161216–163261
HLT_Mu15_Photon20_CaloIdL_v4	163269–163869
HLT_Mu15_Photon20_CaloIdL_v5	165088–165633
HLT_Mu15_Photon20_CaloIdL_v6	165970–166967
HLT_Mu15_Photon20_CaloIdL_v7	167039-167913
HLT_Mu15_Photon20_CaloIdL_v9	170249-173198
HLT_Mu17_Ele8_CaloIdT_CaloIsoVL_v4	173236–173692
Run2011B	
HLT_Mu17_Ele8_CaloIdT_CaloIsoVL_v4	175832–178380
HLT_Mu17_Ele8_CaloIdT_CaloIsoVL_v7	178420–179889
HLT_Mu17_Ele8_CaloIdT_CaloIsoVL_v8	179959–180252

Table 8.4: HLT paths used to select events in the $e\mu jj$ control sample, used for the data driven $t\bar{t}$ background estimate in the $eejj$ channel.

The contribution of $t\bar{t}$ events to the total background in the $eejj$ channel mainly comes from the process in Equation 8.4:

$$t\bar{t} \rightarrow \mathrm{Wb} + \mathrm{Wb} \rightarrow \mathrm{e}\nu\mathrm{b} + \mathrm{e}\nu\mathrm{b} \rightarrow \mathrm{eejj} \tag{8.4}$$

This contribution (both in terms of number of events and shape of kinematic distributions) is estimated from data using a control sample containing one electron, one muon, and at least two jets. It should be noted that this $e\mu jj$ sample is signal-exclusive: mBRW leptoquark

pair production events do not produce $e\mu jj$ final states, under the assumption of no mixing between the generations of leptons and quarks (see assumption **5** in Section 1.2.1). The dominant process contributing to this control sample is identical to the $t\bar{t}$ process described in Equation 8.4, except one of the Ws decays to a muon-neutrino pair instead of an electron-neutrino pair, yielding an $e\mu jj$ final state. In addition to $t\bar{t}$ events, This $e\mu jj$ sample contains a small ($\approx 2\%$ as derived from MC) contamination of non-$t\bar{t}$ events, mainly diboson events.

The events in the $e\mu jj$ control region are selected online using unprescaled muon-photon or muon-electron triggers. The exact trigger used depends on the run period, as reported in Table 8.4. The requirements of these triggers may be broken down as follows:

- Triggers with the word `Mu15` (`Mu17`) in their name require a global muon with $p_{\mathrm{T}} > 15$ (17) GeV in the event.

- Triggers with the word `Photon20` in their name require an energy cluster in the ECAL with a reconstructed E_{T} greater than 20 GeV.

- Triggers with the word `Ele8` in their name require an energy cluster in the ECAL with a reconstructed E_{T} greater than 8 GeV. In addition, this energy cluster must be matched to hits in the pixel detector.

- The word `CaloIdL` in the trigger name implies an identification requirement on the ECAL energy cluster. There must not be significant energy deposits in the HCAL cells directly behind the ECAL cluster ($H/E< 0.15$ for electrons in the EB, $H/E< 0.10$ for electrons in the EE) and the shape of the ECAL cluster must be consistent with the shape of clusters typically produced by electrons and photons ($\sigma_{\mathrm{i}\eta\mathrm{i}\eta}< 0.014$ for electrons in the EB, $\sigma_{\mathrm{i}\eta\mathrm{i}\eta}< 0.035$ for electrons in the EE). Similarly, `CaloIdT` implies a requirement of $H/E< 0.10$ (0.075) for electrons in the EB (EE) and a requirement of $\sigma_{\mathrm{i}\eta\mathrm{i}\eta}< 0.011$ (0.031) for electrons in the EB (EE).

- The word `CaloIsoVL` implies an energy isolation requirement on the ECAL energy cluster. Excluding the energy cluster itself, the sum of the energy in the ECAL cells

surrounding the energy cluster must not be greater than 20% of the electron E_{T}. The same requirement is placed on the energy in the HCAL cells.

In addition to the trigger requirements described above, the $eejj$ preselection and final selection requirements described in Section 7.2 are applied to the $e\mu jj$ sample, where the muon is treated as a second electron.

In MC, before any event reconstruction or selection biases, the $t\bar{t}$ process is expected to produce twice as many $e\mu jj$ events as $eejj$ events with the same kinematic properties. In data, the relation between the number of $eejj$ and $e\mu jj$ events at any given stage of the selection must be corrected for trigger, reconstruction, ID, and isolation efficiencies of leptons. The number of $eejj$ events produced by the $t\bar{t}$ process may be estimated using the number of $e\mu jj$ events via Equation 8.5:

$$\mathrm{N}^{\mathrm{data}}_{eejj} = \mathcal{C} \times \mathrm{N}^{\mathrm{data}}_{e\mu jj} = \frac{1}{2} \times \frac{\epsilon^{trg}_{ee}}{\epsilon^{trg}_{e\mu}} \times \frac{\epsilon^{reco/ID/Iso}_{e}}{\epsilon^{reco/ID/Iso}_{\mu}} \times \mathrm{N}^{\mathrm{data}}_{e\mu jj} \tag{8.5}$$

where :

- ϵ^{trg}_{ee} is the efficiency with which the triggers in Table 7.1 select events with two HEEP electrons, as defined in Section 6.1. This efficiency is taken to be 100%, as described in Section 7.2.1.

- $\epsilon^{trg}_{e\mu}$ is the efficiency with which the triggers in Table 8.4 select events with one HEEP electron and one muon passing the tight muon ID, as described in Section 6.2. This efficiency is also assumed to be 100%

- $\epsilon^{reco/ID/Iso}_{e}$ ($\epsilon^{reco/ID/Iso}_{\mu}$) is the efficiency with which electrons (muons) coming from $e\mu jj$events are reconstructed and pass the identification and isolation requirements. The ratio of these efficiencies is determined from MC by comparing the number of $eejj$ events that pass the $eejj$ preselection with the number of $e\mu jj$ events that pass the $eejj$ preselection if the muon is treated as a second electron. The calculation of

this ratio is described in Equation 8.6:

$$\epsilon_e^{reco/ID/Iso}/\epsilon_\mu^{reco/ID/Iso} = 2 \times \mathrm{N}^{\mathrm{MC}}_{eejj}/\mathrm{N}^{\mathrm{MC}}_{e\mu jj} = 0.98 \pm 0.02\ (\mathrm{stat.}) \pm 0.02\ (\mathrm{syst.}) \quad (8.6)$$

This assumption is justified in light of the fact that both electron and muon efficiencies are measured to have minimal differences between data and MC [115]. The 2% systematic uncertainty on this ratio comes from the uncertainty on the electron reconstruction efficiency (1.5%) and the uncertainty on the muon reconstruction efficiency (1.0%).

The $t\bar{t}$ background prediction is finally obtained by rescaling the $e\mu jj$ data by the factor $\mathcal{C} = 0.49 \pm 0.01\ (\mathrm{stat.}) \pm 0.01\ (\mathrm{syst.})$. The $t\bar{t}$ contributions in the $eejj$ channel shown in the tables and plots of this note are obtained using this method.

The number of $t\bar{t}$ events passing the $eejj$ preselection, estimated using the $e\mu jj$ sample, is 768 ± 19, to be compared with 790 ± 14 events obtained directly from MC (only statistical uncertainties are shown). Figure 8.3 shows the comparison of S_{T} and m_{ej} distributions between the data-driven $t\bar{t}$ background prediction and the $t\bar{t}$ background prediction using MC at $eejj$ preselection level. The agreement is good in both the shape and the normalization.

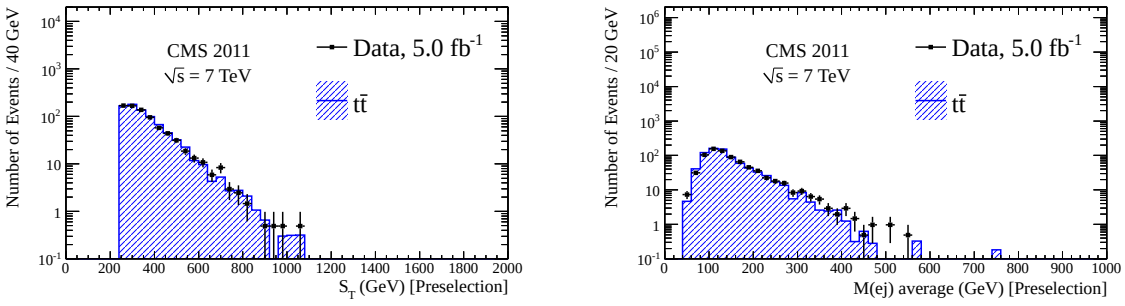


Figure 8.3: Distributions showing the S_{T} (left) and average m_{ej} of the two reconstructed leptoquark candidates (right) for $t\bar{t}$ background predictions using the $e\mu jj$ data sample (black) and MC events passing the $eejj$ preselection (blue)

8.2.2 $t\bar{t}$ (and W+jets) background in the $e\nu jj$ channel

The $t\bar{t}$ and W+jets selection efficiencies in the $e\nu jj$ analysis are derived from MC. The $t\bar{t}$ and W+jets sample normalizations are obtained by comparing data and simulation after two separate selection criteria: one that enriches the samples with $t\bar{t}$ events and one that enriches the samples with W+jets events.

The following selection (selection 1) enriches the samples with W+jets events, with a contamination of $\approx 40\%$, mainly coming from $t\bar{t}$ events. This selection includes events which:

- Pass $e\nu jj$ preselection;
- Have $50 < m_{\mathrm{T,\,e\nu}} < 110$ GeV;
- Have less than 5 jets with $p_{\mathrm{T}} > 40$ GeVand $|\eta| < 3$.

The following selection (selection 2) enriches the samples with $t\bar{t}$ events, with a contamination of $\approx 30\%$, mainly coming from W+jets and QCD multijet events. This selection includes events which:

- Pass $e\nu jj$ preselection
- Have $50 < m_{\mathrm{T,\,e\nu}} < 110$ GeV
- Have at least 5 jets with $p_{\mathrm{T}} > 40$ GeVand $|\eta| < 3$

The results of these two selections can be used to form a system of equations:

$$\begin{cases} \mathrm{N}^1_{\mathrm{data}} = \mathcal{R}_{t\bar{t}}\mathrm{N}^1_{t\bar{t}} + \mathcal{R}_{\mathrm{W}}\mathrm{N}^1_{\mathrm{W}} + \mathrm{N}^1_{\mathrm{QCD}} + \mathrm{N}^1_{\mathrm{Others}} \\ \mathrm{N}^2_{\mathrm{data}} = \mathcal{R}_{t\bar{t}}\mathrm{N}^2_{t\bar{t}} + \mathcal{R}_{\mathrm{W}}\mathrm{N}^2_{\mathrm{W}} + \mathrm{N}^2_{\mathrm{QCD}} + \mathrm{N}^2_{\mathrm{Others}} \end{cases} \tag{8.7}$$

where $\mathrm{N}^{\mathrm{i}}_{\mathrm{data}}$, $\mathrm{N}^{\mathrm{i}}_{\mathrm{W}}$, $\mathrm{N}^{\mathrm{i}}_{\mathrm{Others}}$, $\mathrm{N}^{\mathrm{i}}_{t\bar{t}}$, and $\mathrm{N}^{\mathrm{i}}_{\mathrm{QCD}}$ are, respectively, the number of events in data, the number of events predicted by the W+jets MC sample before rescaling, and the number of events predicted by other MC backgrounds (single-top, diboson, etc..), the

number of events predicted by the $t\bar{t}$ MC sample before rescaling, and the predicted number of QCD multijet events obtained with the method described in Section 8.1, passing selection i. Solving the system yields the following rescaling factors for the MadGraph $t\bar{t}$ and the Sherpa W+jets samples:

$$\mathcal{R}_{t\bar{t}} = 0.72 \pm 0.06 \text{ (stat.)} \pm 0.04 \text{ (syst.)} \tag{8.8}$$

$$\mathcal{R}_{\mathrm{W}} = 1.26 \pm 0.05 \text{ (stat.)} \pm 0.03 \text{ (syst.)}. \tag{8.9}$$

For this study all MC samples are normalized using the cross sections listed in Section 5.2. These factors are already included in the $t\bar{t}$ and W+jets MC predictions for the $e\nu jj$ channel. Figure 8.4 shows the distributions of several reconstructed quantities for events passing the criteria in selection 2 listed above, after the rescaling factors have been applied to the W+jets and $t\bar{t}$ MC. A good agreement between data and background prediction is found in the shape of all these distributions.

8.3 W/Z+jets background

Once the QCD multijets and $t\bar{t}$ background estimates have been determined, the Z^0+jets and W+jets contributions may be extracted. In both the $eejj$ and $e\nu jj$ channels, the Z^0+jets and W+jets contributions to the total background estimates are obtained using MC that has been rescaled within control regions.

8.3.1 Z^0+jets background in the $eejj$ channel

The Z^0+jets selection efficiency in the $eejj$ analysis is derived from MC. The Z^0+jets MC sample normalization is obtained by comparing data and simulation as described below.

After applying the $eejj$ preselection and requiring that the invariant mass of the di-electron pair is close to the Z^0 peak ($70 < m_{\mathrm{ee}} < 110$ GeV), the sample is dominated by

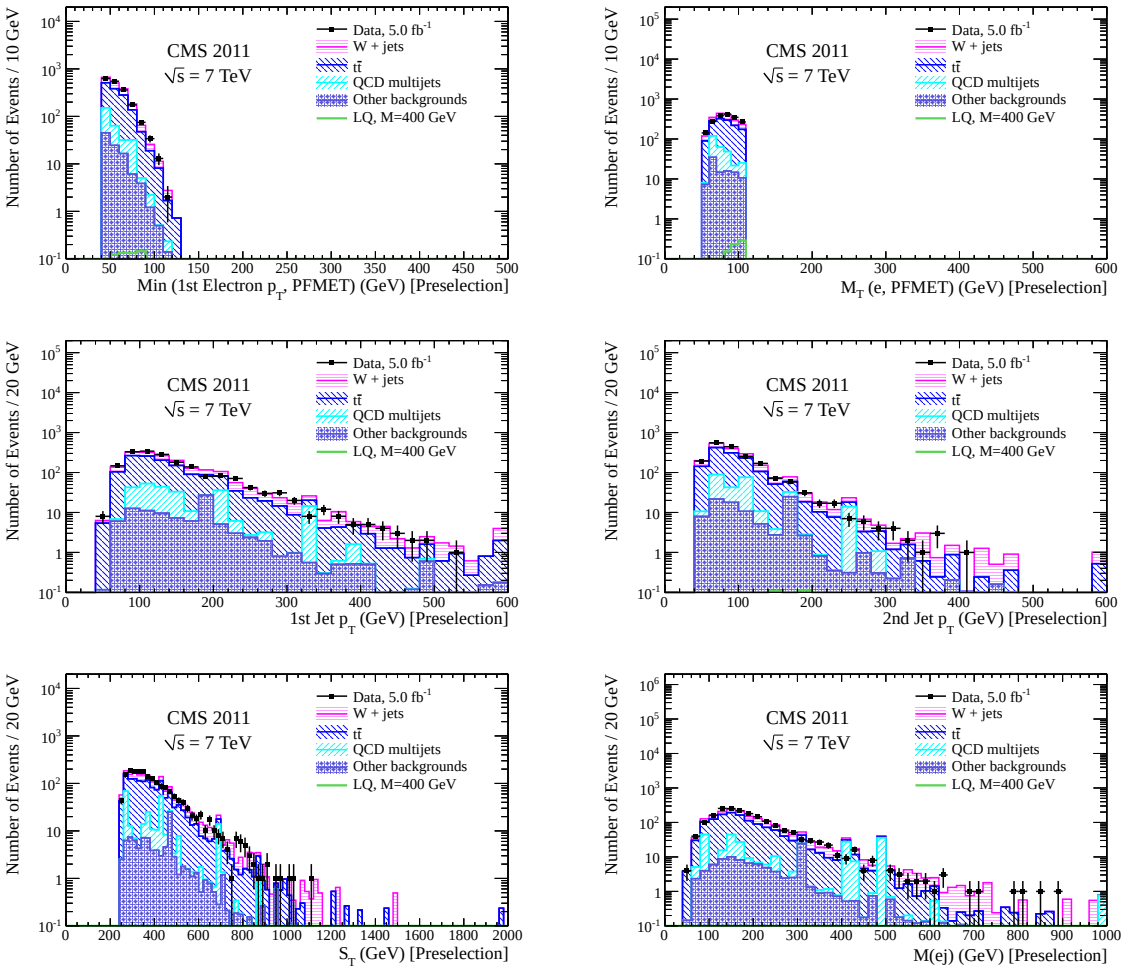


Figure 8.4: The $\min(p_{\mathrm{T}}^{\mathrm{e}}, \not\!\!E_{\mathrm{T}})$ (top left), $m_{\mathrm{T,\,e}\nu}$ (top right), leading jet p_{T} (middle left), second leading jet p_{T} (middle right), S_{T} (bottom left), and m_{ej} (bottom right) distributions for an $e\nu jj$ sample enriched in $t\bar{t}$ events after the rescaling of this MC background simultaneously with the rescaling of the W+jets background.

Z^0+jets events with a contamination from other SM processes of $\approx 5\%$, estimated from MC, mostly coming from $t\bar{t}$ and diboson events. At this stage of the selection, the data are compared to the background predictions and an overall good agreement is found in the shape of the distributions of all the observables employed in the final selection. A data/MC rescaling factor is then computed for the Z^0+jets background, using Equation 8.10:

$$\mathcal{R}_{Z^0} = \frac{N_{data} - (N_{Others} + N_{QCD})}{N_{Z^0}} = 1.27 \pm 0.02 \text{ (stat.)} \pm 0.001 \text{ (syst.)} \tag{8.10}$$

where N_{data}, N_{Others}, N_{Z^0}, and N_{QCD} are, respectively, the number of events in data, the number of events predicted by other MC samples (including $t\bar{t}$, W+jets, diboson, and others), the number of events predicted by the Z^0+jets MC sample before rescaling, and the number of events predicted by the QCD multijet events obtained with the method described in Section 8.1. This factor is already included in the Z^0+jets MC predictions for $eejj$ channel. The systematic uncertainty on the rescaling factor comes from the systematic uncertainty on the data driven QCD multijet and $t\bar{t}$ estimates.

8.3.2 W+jets background in the $e\nu jj$ channel

The W+jets selection efficiency in the $e\nu jj$ analysis is derived from MC. The W+jets MC sample normalization is obtained by comparing data and simulation, simultaneously with the $t\bar{t}$ background, as described in Section 8.2.2.

8.4 Other backgrounds (diboson, single top, and others)

The following SM processes give a small contribution to the total background in both $eejj$ and $e\nu jj$ analyses: single top, diboson, γ + jets, Z^0+jets (only in the $e\nu jj$ channel), and W+jets (only in the $eejj$ channel). The selection efficiencies for these backgrounds are derived from MC, and the samples are normalized to the cross sections listed in Section 5.2.

CHAPTER 9

ANALYSIS: SYSTEMATIC UNCERTAINTIES

The impact of the systematic uncertainties on the number of signal and background events are summarized for a LQ mass hypothesis of 500 GeV in Tables 9.3 and 9.4 for the $eejj$ and $e\nu jj$ channels, respectively. Further details on how these uncertainties are computed are given in this chapter.

9.1 Background normalization

The uncertainties on the normalization factors of the main backgrounds are discussed in Chapter 8 and summarized below:

- overall uncertainty on QCD multijet background in the $eejj$ ($e\nu jj$) channel: 50% (25%)
- scale factor for $e\mu jj$ sample for $t\bar{t}$ background estimate in $eejj$ channel: $\mathcal{C} = 0.49 \pm 0.01$ (stat.) ± 0.01 (syst.)
- scale factor for $t\bar{t}$ MC sample in $e\nu jj$ channel: $\mathcal{R}_{t\bar{t}} = 0.72 \pm 0.06$ (stat.) ± 0.04 (syst.)
- scale factor for Z^0+jets MC sample in $eejj$ channel: $\mathcal{R}_{Z^0} = 1.27 \pm 0.02$ (stat.) ± 0.001 (syst.)
- scale factor for W+jets MC samples in $e\nu jj$ channel: $\mathcal{R}_{W} = 1.26 \pm 0.05$ (stat.) ± 0.03 (syst.)

9.2 $t\bar{t}$, Z^0+jets, and W+jets background shape

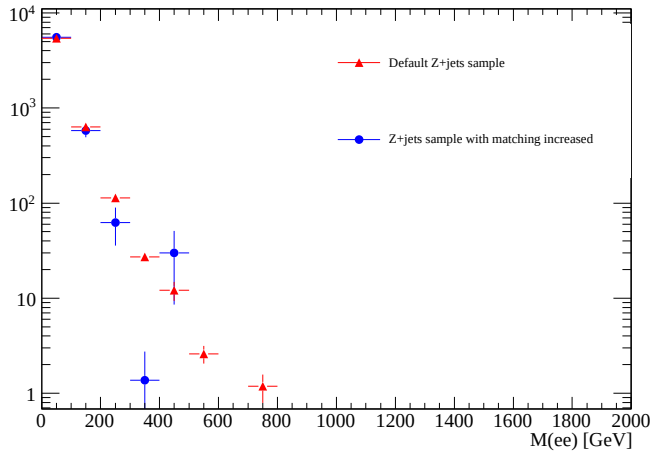
The systematic uncertainties due to the modeling of the shape of the Z^0+jets background (in the $eejj$ channel), and both $t\bar{t}$ and W+jets backgrounds (in the $e\nu jj$ channel) are determined by comparing the background predictions obtained with the default MC samples listed in Section 5.2 with alternative MadGraph MC samples. These alternative MadGraph samples have renormalization/factorization scales and jet matching thresholds varied by a factor of two.

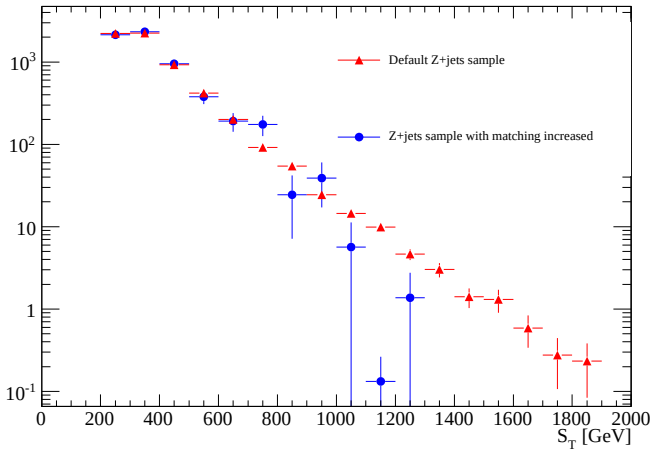
The number of events in the alternative MC samples is limited, and the final $eejj$ and $e\nu jj$ selection is too tight to evaluate these modeling uncertainties. Therefore, instead of applying the full final $eejj$ and $e\nu jj$ selection, looser selection requirements are defined. This looser selection is used to quantify the variation in the shapes of $m_{\rm ee}$, $S_{\rm T}$, and $m_{\rm ej}$ (for the $eejj$ channel) and $\not\!\!E_{\rm T}$, $S_{\rm T}$, and $m_{\rm ej}$ (for the $e\nu jj$ channel) between the default and the alternative samples. The explicit selection critera are listed in Table 9.1.

The alternative samples are first normalized at either $eejj$ or $e\nu jj$ preselection level to have the same number of events of the corresponding default samples. Then the additional selections reported in Table 9.1 are applied. Finally, the most significant discrepancy between the default sample and the alternative samples is calculated for each set of cuts considered. The result of this study are summarized in Table 9.1. The distributions of the studied reconstructed quantities, for the default sample and for the alternative sample showing the most significant discrepancy, are presented in Figures 9.1, 9.2, 9.3.

This study yields the following systematic uncertainties on background shape:

- Z^0+jets in $eejj$ channel : 15%
- W+jets in $e\nu jj$ channel : 20%
- $t\bar{t}$ in $e\nu jj$ channel : 10%





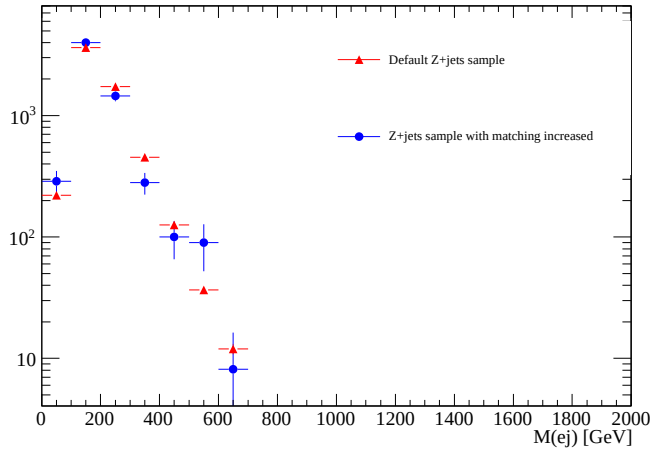
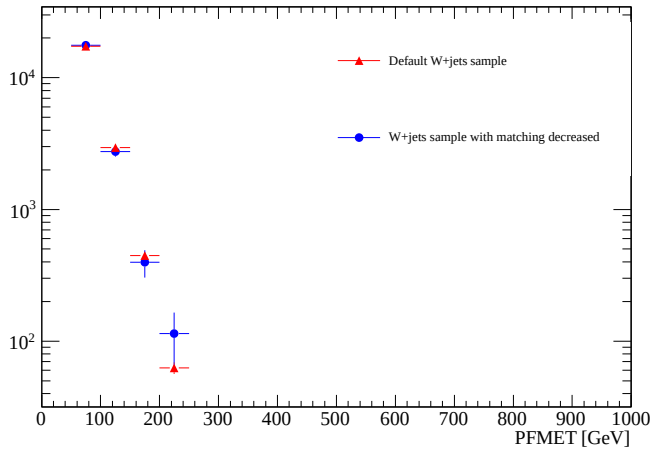
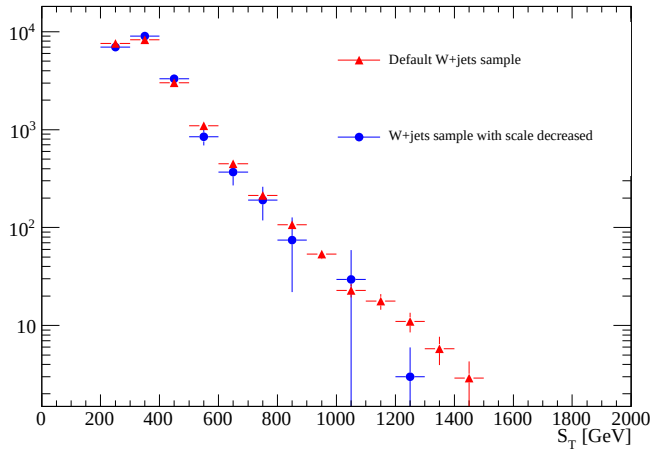


Figure 9.1: Comparison of m_{ee}, S_{T}, and m_{ej} distributions between default and alternative samples for Z^0+jets background shape studies in the $eejj$ channel. The plots refer to the selection criteria reported in Table 9.1.





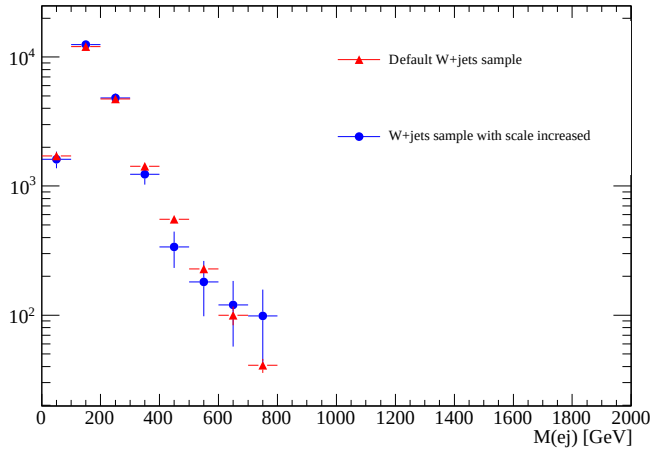
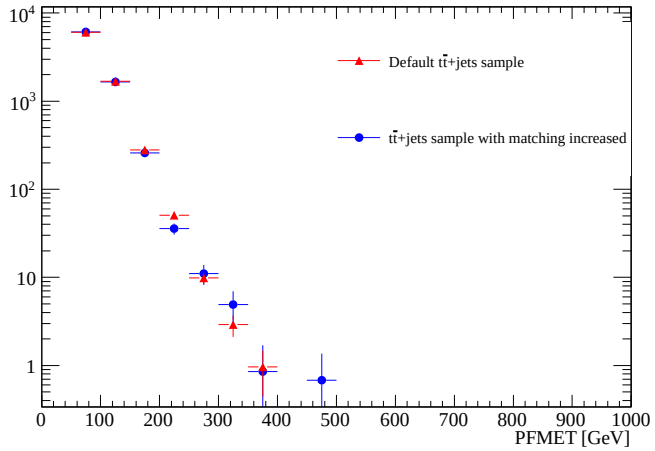
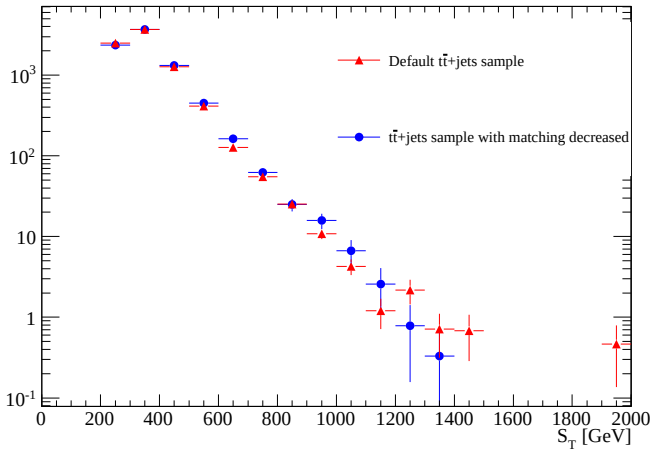


Figure 9.2: Comparison of $\not\!\!E_{\mathrm{T}}$, S_{T}, and m_{ej} distributions between default and alternative samples for W+jets background shape studies in $e\nu jj$ channel. The plots refer to the selection criteria reported in Table 9.1.





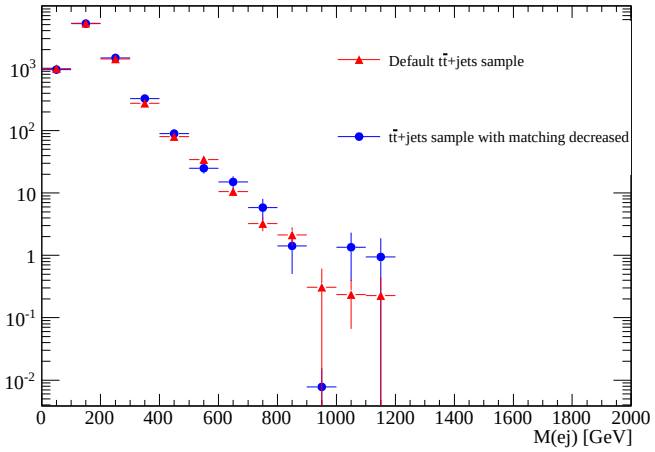


Figure 9.3: Comparison of $\not\!\!E_{\mathrm{T}}$, S_{T}, and m_{ej} distributions between default and alternative samples for $t\bar{t}$ background shape studies in $e\nu jj$ channel. The plots refer to the selection criteria reported in Table 9.1.

Loose Selection	N^A_{events} in Alternative Sample	N^D_{events} in Default Sample	$(N^A_{events}/N^D_{events}) - 1$ [%]	Reference to Figure
Z^0+jets in the $eejj$ Channel				
$eejj$presel. + m_{ee}> 100 GeV	674.05 ± 78.91	789.11 ± 57.68	-14.58 ± 12.43	m_{ee} (Fig. 9.1, top)
$eejj$presel. + S_T> 330 GeV	3498.61 ± 169.77	3090.92 ± 131.94	13.19 ± 6.98	S_T (Fig. 9.1, middle)
$eejj$presel. + m_{ej}> 60 GeV	3498.61 ± 169.77	3090.92 ± 131.94	13.19 ± 6.98	m_{ej} (Fig. 9.1, bottom)
W+jets in the $e\nu jj$ Channel				
$e\nu jj$presel. + $\not\!\!E_T$> 100 GeV	3259.31 ± 214.76	3489.69 ± 283.68	-6.60 ± 10.21	$\not\!\!E_T$ (Fig. 9.2, top)
$e\nu jj$presel. + S_T> 450 GeV	2509.97 ± 224.49	3074.69 ± 263.94	-18.37 ± 11.38	S_T (Fig. 9.2, middle)
$e\nu jj$presel. + m_{ej}> 150 GeV	13270.35 ± 568.15	12937.28 ± 645.48	2.57 ± 6.65	m_{ej}(Fig. 9.2, bottom)
$t\bar{t}$ in the $e\nu jj$ Channel				
$e\nu jj$presel. + $\not\!\!E_T$> 100 GeV	1958.67 ± 32.58	2022.04 ± 22.01	-3.13 ± 1.94	$\not\!\!E_T$ (Fig. 9.3, top)
$e\nu jj$presel. + S_T> 450 GeV	1226.93 ± 25.79	1104.13 ± 14.96	11.12 ± 2.70	S_T (Fig. 9.3, middle)
$e\nu jj$presel. + m_{ej}> 150 GeV	4007.07 ± 46.61	3919.81 ± 35.16	2.23 ± 1.49	m_{ej} (Fig. 9.3, bottom)

Table 9.1: Summary of the study of systematic uncertainties on background shape. The "alternative sample" is the one that presents the largest deviation with respect to the default sample.

9.3 Electron, jet, and $\not\!\!E_T$ energy scales

The electron energy scale uncertainty in the EB (EE) is estimated to be 1% (3%) [74]. A conservative 4% uncertainty on the jet energy scale is used for the entire η and p_T range of the reconstructed jets [116]. For the jet and electron energy scales, the event selection is repeated rescaling by $1 \pm \sigma$ the energy of the reconstructed objects, where σ is the relative uncertainty on their energy scales. While the electron and jet energy scale uncertainties are independent, the $\not\!\!E_T$ scale is directly correlated with both electron and jet energy scales. Therefore, while varying the electron and jet energy scales, a new $\not\!\!E_T$ vector is computed event-by-event using Equation 9.1:

$$\vec{\not\!\!E}_T{}' = \vec{\not\!\!E}_T + \sum_{\text{electrons OR jets}} (\vec{p_T} - \vec{p_T}') \tag{9.1}$$

where $\vec{p_T}$ is the p_T vector of the original electron/jet, $\vec{p_T}'$ is the p_T vector of the electron/jet with modified energy scale, and the sum is performed over all the reconstructed electrons/jets with p_T>30 GeV. The analysis is then repeated using both $\vec{p_T}'$ of electrons/jets and $\vec{\not\!\!E}_T{}'$ at the same time.

For the backgrounds with data driven normalization, the same rescaling procedure adopted

for the default MC sample is also repeated for the samples with varied electron and jet energy scales. This makes it possible to evaluate the impact of electron and jet energy scale systematic uncertainties on the final selection relative to the preselection stage. In order to protect against statistical fluctuations while calculating these systematic uncertainties, a maximum of 10% statistical uncertainty on the number of MC events passing the final selection is required for each systematic uncertainty measurement. When the number of MC events becomes too small (with increasing mass on LQ hypothesis) the impact of the systematic uncertainty for a given selection is estimated to be the same as the impact of the systematic uncertainty corresponding to the selection of the highest mass point satisfying the 10% condition.

At each stage of the final selection, the greatest magnitude of relative change in the number of signal and background events due to the variations of the jet or electron energy scales is used to assess the (symmetric) systematic uncertainties.

These uncertainties are not considered for the QCD multijet background in either the $eejj$ or $e\nu jj$ channels, nor is it considered for the $t\bar{t}$ background in the $eejj$ channel, since for those both the background shape and normalization are already derived from data.

9.4 Electron and jet energy resolution

The effect of the jet energy resolution uncertainty in MC is calculated by matching jets directly from the MC generator to jets that have been reconstructed by the CMS reconstruction algorithms. The p_{T} of the matched reconstructed jet is then reassigned according to Equation 9.2:

$$p^{'}_{T,\mathrm{RECO}} = p_{T,GEN} + c \cdot (p_{T,\mathrm{RECO}} - p_{T,\mathrm{GEN}}) \tag{9.2}$$

where c is the ratio of the jet energy resolution in data vs. MC. This ratio has been found to depend on the reconstructed jet's pseudorapidity, and the values for various pseudorapidity regions are shown in Table 9.2. No change is made to the p_{T} of reconstructed jets for which no

$\|\eta\|$ range	Data/MC Ratio
0.0 - 0.5	1.052 ± 0.012 (stat) ± 0.062 (syst)
0.5 - 1.1	1.057 ± 0.012 (stat) ± 0.056 (syst)
1.1 - 1.7	1.096 ± 0.017 (stat) ± 0.063 (syst)
1.7 - 2.3	1.134 ± 0.035 (stat) ± 0.087 (syst)
2.3 - 5.0	1.288 ± 0.127 (stat) ± 0.155 (syst)

Table 9.2: Ratio between data and MC values for jet energy resolution for different pseudorapity regions of the detector, as measured by the CMS JetMET group

generator match can be made. All of the changes to the reconstructed jet p_{T} are propagated to the $\not\!\!E_{\mathrm{T}}$, using Equation 9.1, as discussed in the previous section.

For the backgrounds with data driven normalization, the same rescaling procedure adopted for the default MC sample is also repeated for the samples with varied electron and jet energy scales. This makes it possible to evaluate the impact of electron and jet energy scale systematic uncertainties on the final selection relative to the preselection stage. In order to protect against statistical fluctuations while calculating these systematic uncertainties, a maximum of 10% statistical uncertainty on the number of MC events passing the final selection is required for each systematic uncertainty measurement. When the number of MC events becomes too small (with increasing mass on LQ hypothesis) the impact of the systematic uncertainty for a given selection is estimated to be the same as the impact of the systematic uncertainty corresponding to the selection of the highest mass point satisfying the 10% condition.

At each stage of the final selection, the greatest magnitude of relative change in the number of signal and background events due to the variations of the jet or electron energy scales is used to assess the (symmetric) systematic uncertainties.

The effect of the electron energy resolution uncertainty in MC is calculated using the same method as for jets, but the value for c is taken to be 1% for electrons in the barrel and 3% for electrons in the endcap [74]. All of the changes to electron p_{T} are propagated to the $\not\!\!E_{\mathrm{T}}$.

These uncertainties are not considered for the QCD multijet background in either the $eejj$ or $e\nu jj$ channels, nor is it considered for the $t\bar{t}$ background in the $eejj$ channel, since for those both the background shape and normalization are already derived from data.

9.5 Integrated luminosity

The uncertainty on the luminosity is 2.2% [78], as discussed in Section 2.2.7.

9.6 MC statistics

The statistical uncertainty on the number of MC events is summarized after the full $eejj$ ($e\nu jj$) event selection in Table 7.3 (Table 7.7 or Table 7.8) for signal and background samples. This is the dominant systematic uncertainty for both the $e\nu jj$ and $eejj$ channels.

9.7 Electron trigger, reconstruction, identification and isolation uncertainties

For the $eejj$ channel, the trigger efficiency is assumed to be 100% with negligible uncertainty, as described in Section 7.2.1. For the $e\nu jj$ channel, a combination of single-electron and electron+$\not\!\!E_{\mathrm{T}}$+dijets triggers are used depending on the data taking period: the efficiency with which the $e\nu jj$ trigger fires on events passing the electron, $\not\!\!E_{\mathrm{T}}$, and jet-related cuts of the $e\nu jj$ preselection is measured to be $95.5^{+1.0\%}_{-3.5\%}$, as described in Section 7.3.1.

The efficiency with which electrons from Z^0 decays are reconstructed and pass the HEEP ID requirements have been studied in the context of the Z'$\rightarrow$ ee analysis [115]. An overall uncertainty of 1.5% on the single electron efficiency is used in this analysis.

The trigger uncertainty and the reconstruction and identification uncertainties are summed in quadrature giving a total 3% ($^{+2.0\%}_{-4\%}$) contribution to the uncertainty on the signal selection

efficiency for the $eejj$ ($e\nu jj$) channel. A conservative value of $\pm 4\%$ is quoted for the $e\nu jj$ channel. For those background estimations where data-driven uncertainties have already been determined, these uncertainties are not considered.

9.8 Parton distribution function (PDF)

Uncertainties due to the choice of parton distribution functions (PDF) of the proton lead to changes in the total cross section and the selection efficiencies for both signal and background processes. The effect of the PDF uncertainties on the signal acceptance amounts to less than 1%. For those background estimations where data-driven uncertainties have already been determined, these uncertainties are not considered.

9.9 Pile-up

All MC events in this analysis are reweighted to represent the pile-up conditions observed in data (see Section 3.7). This reweighting assumes a value of the inelastic pp interaction cross section, which has an associated uncertainty of 8.5%. The MC is reweighted after varying this cross section up and down by 8.5%, and the analysis is repeated. The number of events passing the full selection is calculated for the $\pm 10\%$ cases and the default one. The maximum variation with respect to the default is used to assess a systematic uncertainty. The uncertainty is less than 1% for leptoquark signal at all masses in both $eejj$ and $e\nu jj$ channels. For backgrounds, the uncertainty is less than 1% (3%) in the $eejj$ ($e\nu jj$) channels.

Systematic Uncertainty	Magnitude [%]	Effect on Signal [%]	Effect on Background [%]
Background Normalization	Section 9.1	–	2.5
Background Shape	Section 9.2	–	10.5
Jet Energy Scale $\star$	4	2	1
Electron Energy Scale Barrel (Endcap) $\star$	1(3)	1	6
Jet Energy Resolution	Section 9.4	0.5	0.5
Electron Energy Resolution Barrel (Endcap)	1(3)	0.5	1
Electron Trigger/Reco/ID/Iso Efficiency	Section 9.7	3	–
Pileup	8.5	1	1
MC statistics	Table 7.3	0.5	28
Integrated Luminosity	2.2	2.2	0.5
Total		4.5	30

Table 9.3: Systematic uncertainties for $eejj$ channel relative to the selection optimized for a LQ mass of 500 GeV. Note $\star$: the jet and electron energy scale uncertainties are also propagated to the $\not\!\!E_{\mathrm{T}}$ calculation.

Systematic Uncertainty	Magnitude [%]	Effect on Signal [%]	Effect on Background [%]
Background Normalization	Section 9.1	–	7
Background Shape	Section 9.2	–	11
Jet Energy Scale $\star$	4	6	12
Electron Energy Scale Barrel (Endcap) $\star$	1(3)	1.5	4.5
Jet Energy Resolution	Section 9.4	0.5	9
Electron Energy Resolution Barrel (Endcap)	1(3)	< 0.5	1.5
Electron Trigger/Reco/ID/Iso Efficiency	Section 9.7	4	–
Pileup	8.5	1	2
MC statistics	Table 7.7 or 7.8	1.0	19
Integrated Luminosity	2.2	2.2	–
Total		8	28

Table 9.4: Systematic uncertainties for $e\nu jj$ channel relative to the selection optimized for a LQ mass of 500 GeV. Note $\star$: the jet and electron energy scale uncertainties are also propagated to the $\not\!\!E_{\mathrm{T}}$ calculation.

CHAPTER 10

ANALYSIS: LIMIT CALCULATION

In both the $eejj$ and $e\nu jj$ channels, the number of observed data events passing the full selection criteria is consistent with the standard model background prediction. In the absence of a leptoquark signal, an upper limit on the leptoquark production cross section times branching fraction is set using the asymptotic CL_S modified frequentist approach [117, 118].

The 95% C.L. upper limit on $\sigma \times \beta^2$ (for the $eejj$ channel) and $\sigma \times 2\beta(1-\beta)$ (for the $e\nu jj$ channel) as a function of leptoquark mass are shown together with the NLO predictions for the scalar leptoquark pair production cross section in Figure 10.1. By comparing the observed upper limit with the theoretical cross section values, first generation scalar leptoquarks with mass less than 830 (640) GeV are excluded with the assumption that $\beta = 1(0.5)$. This is to be compared with a median expected limit of 790 (640) GeV.

If a new limit is calculated assuming that the systematical uncertainties on signal and backgrounds are only coming from the MC statistics, the results are very similar to the ones obtained with full systematics. In particular, no change is observed in the observed lower leptoquark mass limit for the $eejj$ channel and $\beta = 1$. The lower leptoquark mass limit for the $e\nu jj$ channel and $\beta = 0.5$ changes from 640 GeV to 642 GeV. This confirms that the limit is dominated by statistical fluctuations while systematics play a very small role.

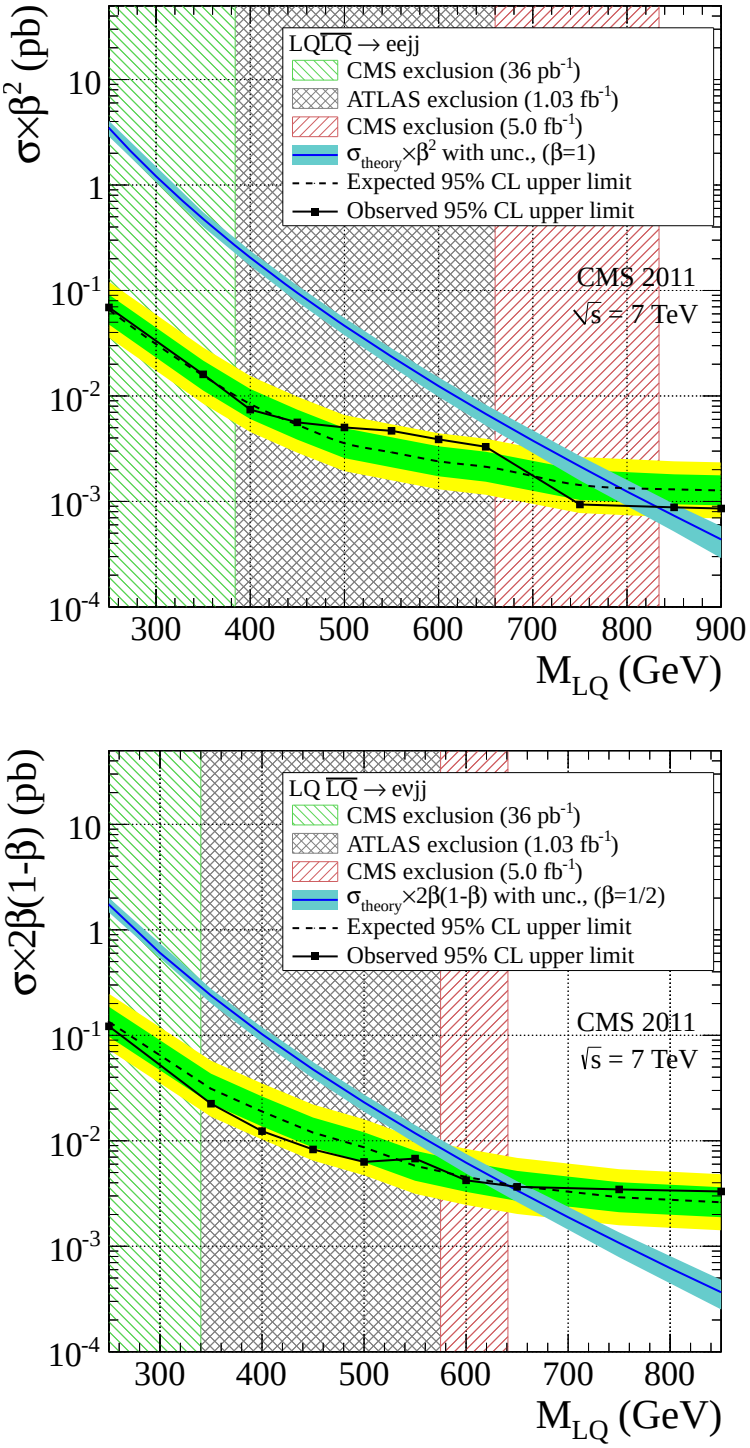


Figure 10.1: The expected and observed upper limit at 95% C.L. on the leptoquark pair production cross section times β^2 in the top plot ($2\beta(1-\beta)$ in the bottom plot) as a function of the leptoquark mass obtained with the $eejj$ ($e\nu jj$) analysis. The systematic uncertainties described in the text are included in the calculation. The dark blue curve and the light blue band represent, respectively, the theoretical leptoquark pair production cross section and the uncertainties due to the choice of PDF and renormalization/factorization scales [54].

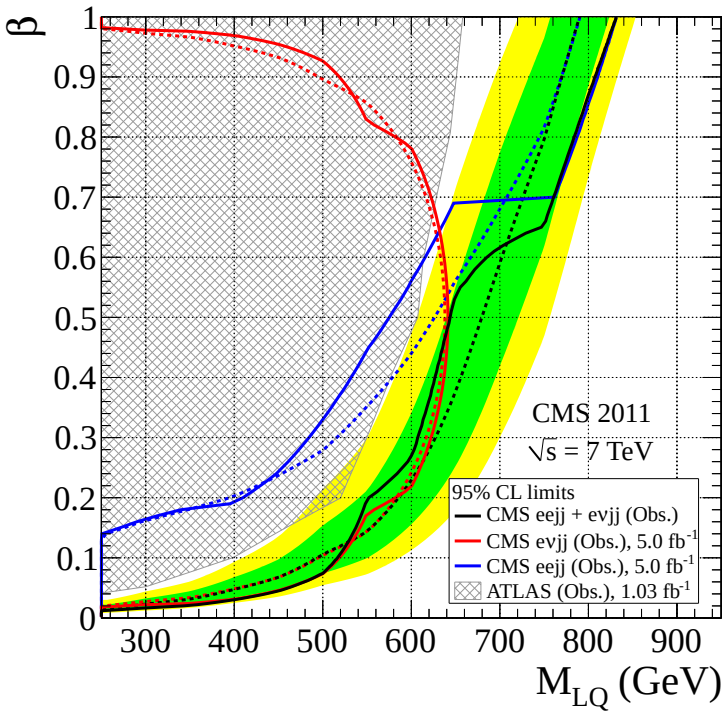


Figure 10.2: The expected and observed exclusion limits at 95% CL on the first generation leptoquark hypothesis in the β vs. mass plane using the central value of signal crosss section for the individual $eejj$ and $e\nu jj$ channels and their combination. The dark green and light yellow expected limit uncertainty bands represent the 68% and 95% confidence intervals. Solid lines represent the observed limits. The systematic uncertainties reported in Chapter 9 are included in the calculation. The shaded region is excluded by the current ATLAS limits [51].

The expected limit on the branching fraction, β, as a function of leptoquark mass may be further improved by combining the $eejj$ and $e\nu jj$ channels, as shown in Figure 10.2. These combinations lead to the exclusion of first generation scalar leptoquarks with masses less than 640 GeV for $\beta = 0.5$, compared with a median expected limit of 680 GeV. These limits are currently the most stringent on first generation scalar leptoquarks [53].

CHAPTER 11

CONCLUSION AND PROSPECTS

This thesis has described a search for the pair production of first generation scalar leptoquarks in $\sqrt{s} = 7$ TeV pp collisions using the CMS detector at the LHC. The data used for this search was collected in 2011 and corresponds to an integrated luminosity of 4.95 fb^{-1}. For each leptoquark final state and mass under consideration, the number of events observed is in good agreement with the Standard Model prediction. In light of this, a 95% confidence level combined lower limit is set on the mass of a first generation scalar leptoquark at 830 and 640 GeV for $\beta = 1$ and 0.5, respectively, where β is the branching fraction of the leptoquark to an electron and a quark. This represents a substantial improvement over the 2010 interation of this analysis, which excluded the production of first generation scalar leptoquarks with masses less than 384 and 340 GeV, when assuming $\beta = 1$ and 0.5, respectively [62, 63].

This analysis is currently being updated for the 2012 CMS dataset, which consists of $\sqrt{s} = 8$ TeV pp collisions and corresponds to an integrated luminosity of 19.6 fb^{-1}. This significant improvement in collision energy and dataset size will extend the sensitivity to leptoquarks masses up to 1 TeV (750 GeV) in the $eejj$ ($e\nu jj$) channel. Even higher collision energies, larger datasets, and stronger sensitivities await this analysis after the end of LS1 and the continuation of the LHC pp physics program.

BIBLIOGRAPHY

[1] **H1** Collaboration, *Z. Phys. C* **74**, 191–206 (1997), arXiv:hep-ex/9702012 [hep-ex].

[2] **ZEUS** Collaboration, *Z. Phys. C* **74**, 207–220 (1997), arXiv:hep-ex/9702015 [hep-ex].

[3] **CDF and D0** Collaboration, arXiv:hep-ex/9810015 [hep-ex].

[4] D. Griffiths, *Introduction to Elementary Particles*. John Wiley & Sons, 1987.

[5] C. Burgess and G. Moore, *The Standard Model: A Primer*. Cambridge University Press, 2012.

[6] M. Peskin and D. Schroeder, *An Introduction To Quantum Field Theory (Frontiers in Physics)*. Westview Press, 1995.

[7] G. Altarelli and K. Winter, *Neutrino Mass*. Springer, 2003.

[8] S. L. Glashow, *Nucl. Phys.* **22**, 579–588 (1961).

[9] S. Weinberg, *Phys. Rev. Lett.* **19**, 1264–1266 (1967).

[10] A. Salam, *Conference Proceedings* **C680519**, 367–377 (1968). Originally printed in "Svartholm: Elementary Particle Theory, Proceedings Of The Nobel Symposium Held 1968 At Lerum, Sweden," Stockholm 1968, 367-377.

[11] P. W. Higgs, *Phys. Lett.* **12**, 132–133 (1964).

[12] F. Englert and R. Brout, *Phys. Rev. Lett.* **13**, 321–322 (1964).

[13] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Phys. Rev. Lett.* **13**, 585–587 (1964).

[14] J. Blümlein and R. Rückl, *Phys. Lett. B* **304**, 337–346 (1993).

[15] J. C. Pati and A. Salam, *Phys. Rev. D* **8**, 1240–1251 (1973).

[16] J. C. Pati and A. Salam, *Phys. Rev. D* **10**, 275 (1974).

[17] H. Georgi and S. Glashow, *Phys. Rev. Lett.* **32**, 438 (1974).

[18] H. Murayama and T. Yanagida, *Mod. Phys. Lett. A* **7**, 147–152 (1992).

[19] H. Fritzsch and P. Minkowski, *Annals of Physics* **93**, 193 – 266 (1975).

[20] G. Senjanović and A. Sokorac, *Z. Phys. C* **20**, 255–257 (1983).

[21] P. H. Frampton and B.-H. Lee, *Phys. Rev. Lett.* **64**, 619–621 (1990).

[22] P. H. Frampton and T. W. Kephart, *Phys. Rev. D* **42**, 3892–3894 (1990).

[23] J. L. Hewett and T. G. Rizzo, *Phys. Rep.* **183**, 193 – 381 (1989).

[24] S. Dimopoulos and L. Susskind, *Nucl. Phys. B* **155**, 237 (1979).

[25] S. Dimopoulos, *Nucl. Phys. B* **168**, 69 (1980).

[26] E. Eichten and K. Lane, *Phys. Lett. B* **90**, 85 (1980).

[27] B. Schrempp and F. Schrempp, *Phys. Lett. B* **153**, 101 (1985).

[28] J. Wudka, *Phys. Lett. B* **167**, 337–342 (1986).

[29] S. Pakvasa, *Int. J. Mod. Phys. A* **2**, 1317–1325 (1987).

[30] L. F. Abbott and E. Farhi, *Phys. Lett. B* **101**, 69–72 (1981).

[31] D. Choudhury and S. Raychaudhuri, *Phys. Lett. B* **401**, 54–61 (1997).

[32] H. K. Dreiner, `arXiv:hep-ph/9707435 [hep-ph]`.

[33] J. A. Evans and Y. Kats, *JHEP* **2013**, 1–40 (2013).

[34] W. Buchmüller, R. Rückl, and D. Wyler, *Phys. Lett. B* **191**, 442–448 (1987).

[35] A. Djouadi, T. Köhler, M. Spira, and J. Tutas, *Z. Phys. C* **46**, 679–685 (1990).

[36] M. Kuze and Y. Sirois, *Progress in Particle and Nuclear Physics* **50**, 1 – 62 (2003).

[37] J. Hewett, T. Rizzo, S. Pakvasa, H. Haber, and A. Pomarol, `arXiv:hep-ph/9310361 [hep-ph]`.

[38] B. Dion, L. Marleau, and G. Simon, *Phys. Rev. D* **59**, 015001 (1998).

[39] J. Blümlein, E. Boos, and A. Kryukov, *Phys. Lett. B* **392**, 150 – 154 (1997).

[40] J. Blümlein, E. Boos, and A. Kryukov, `arXiv:hep-ph/9610408 [hep-ph]`.

[41] W. Buchmüller and D. Wyler, *Phys. Lett. B* **177**, 377 – 382 (1986).

[42] M. Leurer, *Phys. Rev. D* **50**, 536–541 (1994).

[43] M. Leurer, *Phys. Rev. D* **49**, 333–342 (1994).

[44] M. Leurer, *Phys. Rev. Lett.* **71**, 1324–1327 (1993).

[45] J. L. Hewett and S. Pakvasa, *Phys. Lett. B* **227**, 178 – 181 (1989).

[46] G. Bélanger, D. London, and H. Nadeau, *Phys. Rev. D* **49**, 3140–3147 (1994).

[47] V. Ilyin, A. Pukhov, V. Savrin, A. Semenov, and W. von Schlippe, *Phys. Lett. B* **351**, 504 – 509 (1995).

[48] J. L. Hewett and S. Pakvasa, *Phys. Rev. D* **37**, 3165–3171 (1988).

[49] A. Belyaev, C. Leroy, R. Mehdiyev, and A. Pukhov, *JHEP* **2005**, 005 (2005).

[50] **D0** Collaboration, *Phys. Rev. D* **84**, 071104 (2011).

[51] **ATLAS** Collaboration, *Phys. Lett. B* **709**, 158 – 176 (2012).

[52] **ATLAS** Collaboration, *Eur. Phys. J. C* **72**, 1–21 (2012).

[53] **CMS** Collaboration, *Phys. Rev. D* **86**, 052013 (2012).

[54] M. Krämer, T. Plehn, M. Spira, and P. M. Zerwas, *Phys. Rev. D* **71**, 057503 (2005).

[55] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, D. Stump, W.-K. Tung, and C.-P. Yuan, *Phys. Rev. D* **78**, 013004 (2008).

[56] K.S. Babu, C. Kolda, and J. March-Russell, *Phys. Lett. B* **408**, 261 – 267 (1997).

[57] G. Altarelli, G. F. Giudice, and M. L. Mangano, *Nucl. Phys. B* **506**, 29 – 47 (1997).

[58] J. L. Hewett and T. G. Rizzo, *Phys. Rev. D* **58**, 055005 (1998).

[59] **Particle Data Group** Collaboration, *Phys. Rev. D* **86**, 010001 (2012).

[60] **ZEUS** Collaboration, *Phys. Rev. D* **86**, 012005 (2012).

[61] **H1** Collaboration, *Phys. Lett. B* **704**, 388 – 396 (2011).

[62] **CMS** Collaboration, *Phys. Rev. Lett.* **106**, 201802 (2011).

[63] **CMS** Collaboration, *Phys. Lett. B* **703**, 246 – 266 (2011).

[64] L. Evans and P. Bryant, *JINST* **3**, S08001 (2008).

[65] S. Dailler. LHC-PHO-1997-169.

[66] **ATLAS** Collaboration, *JINST* **3**, S08003 (2008).

[67] **CMS** Collaboration, *JINST* **3**, S08004 (2008).

[68] **LHCb** Collaboration, *JINST* **3**, S08005 (2008).

[69] **ALICE** Collaboration, *JINST* **3**, S08002 (2008).

[70] J.-L. Caron. LHC-PHO-1991-001.

[71] **CMS** Collaboration. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults`.

[72] **CMS** Collaboration, *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. CERN, 2006.

[73] P. Adzic, *JINST* **2**, P04004 (2007).

[74] **CMS** Collaboration. CMS-PAS-EGM-10-003.

[75] **CMS** Collaboration. CMS-NOTE-2005-016.

[76] **CMS** Collaboration, *Eur. Phys. J. C* **53**, 139–166 (2008).

[77] **CMS** Collaboration, *Eur. Phys. J. C* **55**, 159–171 (2008).

[78] **CMS** Collaboration. CMS-PAS-SMP-12-008.

[79] **CMS** Collaboration, *Eur. Phys. J. C* **46**, 605–667 (2006).

[80] R. Brun and F. Rademakers, *Nucl. Instrum. Meth. A* **389**, 81 – 86 (1997).

[81] I. Antcheva *et al.*, *Comp. Phys. Comm.* **180**, 2499 – 2512 (2009).

[82] I. Antcheva *et al.*, *Comp. Phys. Comm.* **182**, 1384 – 1385 (2011).

[83] **CMS** Collaboration. CMS-NOTE-2007-021.

[84] **CMS** Collaboration, *Eur. Phys. J. C* **70**, 1165–1192 (2010).

[85] **CMS** Collaboration. CMS-NOTE-2006-041.

[86] **CMS** Collaboration. CMS-PAS-TRK-10-001.

[87] **CMS** Collaboration. CMS-PHO-EVENTS-2012-006.

[88] **CMS** Collaboration. CMS-NOTE-2007-008.

[89] **CMS** Collaboration. CMS-PAS-TRK-10-005.

[90] **CMS** Collaboration. CMS-RN-2003-001.

[91] S. Baffioni, C. Charlot, F. Ferri, D. Futyan, P. Meridiani, I. Puljak, C. Rovelli, R. Salerno, and Y. Sirois, *Eur. Phys. J. C* **49**, 1099–1116 (2007).

[92] **CMS** Collaboration. CMS-PAS-EGM-10-001.

[93] **CMS** Collaboration. CMS-PAS-EGM-10-004.

[94] **CMS** Collaboration. CMS-PAS-PFT-10-001.

[95] **CMS** Collaboration. CMS-PAS-MUO-10-002.

[96] G. P. Salam, *Eur. Phys. J. C* **67**, 637–686 (2010).

[97] M. Cacciari, G. P. Salam, and G. Soyez, *JHEP* **2008**, 063 (2008).

[98] **CMS** Collaboration. CMS-PAS-JME-10-003.

[99] **CMS** Collaboration, *JINST* **6**, P09001 (2011).

[100] T. Sjöstrand, S. Mrenna, and P. Skands, *JHEP* **05**, 026 (2006).

[101] T. Gleisberg, S. Hoeche, F. Krauss, M. Schoenherr, S. Schumann, F. Siegert, and J. Winter, *JHEP* **0902**, 007 (2009).

[102] F. Maltoni and T. Stelzer, *JHEP* **02**, 027 (2003).

[103] J. Alwall *et al.*, *JHEP* **09**, 028 (2007).

[104] S. Alioli, P. Nason, C. Oleari, and E. Re, *JHEP* **0909**, 111 (2009).

[105] Emanuele Re, *Eur. Phys. J. C* **71**, 1547 (2011), arXiv:1009.2450 [hep-ph].

[106] P. Nason, *JHEP* **0411**, 040 (2004), arXiv:hep-ph/0409146 [hep-ph].

[107] S. Frixione, P. Nason, and C. Oleari, *JHEP* **0711**, 070 (2007).

[108] S. Alioli, P. Nason, C. Oleari, and Emanuele Re, *JHEP* **1006**, 043 (2010), arXiv:1002.2581 [hep-ph].

[109] S. Agostinelli *et al.*, *Nucl. Instrum. Meth. A* **506**, 250 – 303 (2003).

[110] **CMS** Collaboration, *Phys. Rev. Lett.* **107**, 221804 (2011).

[111] N. Kidonakis, *Proceedings of the DPF-2009 Conference* (2009), arXiv:0909.0037 [hep-ph].

[112] K. Melnikov and F. Petriello, *Phys. Rev. D* **74**, 114017 (2006).

[113] N. Kidonakis, *Phys. Rev. D* **81**, 054028 (2010).

[114] J. M. Campbell and R. K. Ellis, *Phys. Rev. D* **62**, 114012 (2000).

[115] **CMS** Collaboration, *Phys. Lett. B* **714**, 158 – 179 (2012).

[116] **CMS** Collaboration. CMS-PAS-JME-10-010.

[117] T. Junk, *Nucl. Instrum. Meth. A* **434**, 435 – 443 (1999).

[118] A. L. Read, *J. Phys. G* **28**, 2693 (2002).