W Boson Charge Asymmetry in the Muonic Decay Channel at the CMS Experiment

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Grayson Laurent Williams, B.S., M.S.

Graduate Program in Physics

The Ohio State University

2011

Dissertation Committee:

Professor Lloyd Stanley DURKIN, Advisor

Christopher HILL

Junko SHIGEMITSU

Ratnasingham SOORYAKUMAR

Abstract

The Large Hadron Collider is a new circular particle accelerator designed to collide bunches of protons at a beam energy of $\sqrt{s} = 7$ TeV. One of the two large general-purpose detectors is the Compact Muon Solenoid. This paper will provide a measurement of the W boson cross section and charge asymmetry in the decay channel W$\rightarrow \mu\nu_\mu$ over the entire 2010 ($\int \mathcal{L}dt =$ 35.9 pb^{-1}) dataset, as a means of testing detector performance and of constraining Parton Distribution Functions, as well as testing theoretical predictions in the new energy regime. The cross section and charge asymmetry were measured by first imposing a series of cuts, then using sideband subtraction to remove background process contributions. Systematic and statistical uncertainties were also calculated. The cross section was found to be

$$\sigma_{PP\rightarrow WX} \times \mathrm{BR}\,(\mathrm{W} \rightarrow \mu\nu_\mu) = 9570 \pm 27 \text{ (stat.)} \pm 220 \text{ (syst.)} \pm 965 \text{ (lum.) pb.},$$

consistent with theoretical predictions, and the overall charge asymmetry was found to be

$$\mathrm{W}^+/\mathrm{W}^- = 1.489 \pm 0.009 \text{ (stat.)} \pm 0.032 \text{ (syst.)},$$

also consistent with theoretical predictions.

I'd like to dedicate this work to my mother, Eléonore Van den Berghe. If it weren't for her, I wouldn't be who or where I am today. Thanks mom for all your support and love throughout the years!

Acknowledgments

This work would not have been possible without the help and support of many people, and it is an honor to be able to thank them here.

To begin, I'd like to thank first and foremost my advisor, Stan Durkin. He helped mold me into something remotely resembling a scientist, and taught me much along the way; both about high-energy physics and about life in general. Many thanks also to Drs. Ratnasingham Sooryakumar, Junko Shigemitsu, Chris Hill, and Ta-Yung Ling for agreeing to serve on my committee and supervising my work along the way. Thanks in addition to Eric Braaten, David Stroud, Mohit Randeria, Samir Mathur, and Junko Shigemitsu (again) for teaching me my core classes and imparting a theoretical foundation in the early years of my graduate career.

I'd also like to thank Jason Gilmore, Jianhui Gu, Ben Bylsma, Gregory Rakness, Jay Hauser, Mikhail "Misha" Ignatenko, Károly Banicz, Frank Geurts, Armando Lanaro, Dick Loveless, Fred Borcherding, Igor Vorobiev, Laria Redjimi Boulahouache, Oana Boeriu, Ronny Remington, Ping Tan, Fan Yang, Slawomir Tkaczyk, Michael Schmitt, Darin Acosta, Robert Clare, and Tim Cox for helping me navigate my way through experimental physics and showing me how the detector works and how to be a successful experimental particle physicist. Thanks in addition to Brenda Mellett, Tom Humanic, Karen Kitts, and Michael Wells for making the many trips to and from Geneva, Switzerland painless and easy.

Next, I'd like to thank the many friends I've met during my graduate career who helped me along the way, both by providing me with their company and with their support. In rough chronological order: Kevin Driver, Mark Murphey, Sarah Parks, Michael Fellinger, Rob Guidry, Adam Hauser, Anthony Link, Louis Nemzer, Kerry Highbarger, Rebecca

Daskalova (née Weber), Lauren Welker, Kevin Knobbe, Phillip Killewald, Sheldon Bailey, Eric King, Christina Kullberg, Jim Davis, James and Veronica Stapleton, Greg Vieira, Catherine Sundt, Chris Murphy, Nicole Ippolito, Ryan Carroll, Nicholas and Mary Kypreos, Keith Rose, Jetendr Shamdasani, Jonathan Efron, Andrei Loginov, Mikhail Borodin, Ruslan Mashinistov, Mark Schreiber, Rachel Wilken (née Bartek), Emily Thompson, Ryan Reece, Josh Kunkle, John Alison, John Penwell, Gabriel Ybeles Smit, Alisha Bruderly, Geoff Smith, William Parker, and Lena Lupia. My apologies to anyone I forgot in this list.

Finally, I'd like to thank my boyfriend, Shane Tan. Thanks for waiting patiently for me for so long and always encouraging me to believe in myself and never give up, and push through to the end. I love you Shane!

Table of Contents

Appendices

List of Figures

List of Tables

Chapter 1
Introduction

The Standard Model is one of the most fundamental theories of physics–it explains the basic subatomic particles which comprise matter and the interactions between them, and aptly describes all fundamental forces of nature except gravity. It has been experimentally verified, and all predicted particles have been found, with the sole exception of the Higgs boson. However, the Standard Model has some limitations–it is only valid up to the electroweak energy scale, and it suffers from what is known as the Hierarchy Problem, wherein values of quantities such as couplings and masses are vastly different as predicted by theory and as measured by experiment. An example is that the weak force is 10^{32} times stronger than gravity.

The Large Hadron Collider (LHC) was built at CERN to resolve the question of the existence (and physical properties) of the Higgs boson, probe the internal structure of the proton, and investigate possible solutions to the Hierarchy Problem such as Supersymmetry (SUSY). The LHC is primarily designed to collide bunches of protons with each other, and follows the ISR (proton-proton) and SPS[1] (proton-antiproton), both at CERN, and the Tevatron (proton-antiproton), at Fermilab just outside Chicago, IL, as the latest and most powerful hadron collider. Although the LHC is conceptually similar to hadron colliders before it, the center-of-mass energy of $\sqrt{s} = 7$ TeV, the instantaneous luminosity and the bunch crossing rate are all many times higher than at previous hadron colliders. Thus, before new physics can be explored, it is vitally important to gain a solid understanding of

[1] The SPS proton-antiproton collider was where the W and Z bosons were discovered, in 1983.

physics at the new energy frontier.

The proton, one of the fundamental buildings blocks of nature and an integral component of every atom, is composed of constituent point particles called quarks, virtual quark-antiquark pairs, and gluons (collectively referred to as partons). The internal structure of the proton is described by so-called Parton Distribution Functions, which must be measured experimentally. One way to do so which is accessible in early running is to measure the charge asymmetry of the W boson. According to predictions of the structure functions, asymmetry is expected to have a dependence on the angle relative to the beam line, so the principal measurement will be made as a function of a quantity called pseudorapidity which encompasses the beam angle. One important source of systematic uncertainty in early running is in modeling of jets emanating from quarks and gluons, due to limitations in the perturbative QCD method used to predict them.

The primary goal of this paper will be to measure the charge asymmetry of the W boson in the muonic decay channel as a function of muon pseudorapidity, as a means of probing the internal structure of the proton and gleaning a better understanding of Parton Distribution Functions. A further investigation into charge asymmetry as a function of jet multiplicity will be undertaken, in order to better understand QCD-related uncertainties. Finally, a measurement of the overall cross section of the W boson in the muonic decay channel will be made, as a cross check on the other measurements and as a means of verifying the understanding of the operation of the CMS detector and the LHC. The data used for this paper is the entire 2010 data run at CMS.

The layout of this paper is as follows: Chapter 2 will provide first a theoretical overview of the Standard Model, including its constituent quantum field theories and fundamental particles. Then, an overview of the proton will be given, as will an overview of Parton Distribution Functions and Deep Inelastic Scattering, a method developed to probe the proton's inner structure. Finally, an overview of W production and decay at the LHC, and associated jet production, will be given. Chapter 3 will discuss the physical setup of the LHC and the CMS experiment, and provide an overview of event triggering and data collection, storage and distribution. Luminosity measurement will also be discussed. Chapter 4 will begin by defining important quantities to be used in this analysis, then mention the Monte Carlo method of generating simulated data as a means of estimating signal and background yields and testing detector acceptance and efficiency, both of which will subsequently be explored in this chapter after an overview of event reconstruction. Then, the important background physics processes will be described, and a description of the selection cuts used to reduce background will be made. Finally, a description of the analysis technique will be presented. Chapter 5 will present the results of the analysis, as well as a description of the sources of systematic error and their overall contribution, and Chapter 6 concludes with a summary of the results and a comparison to past experiments and theoretical predictions.

Chapter 2
Theory and Phenomenology

The purpose of this chapter is to provide the theoretical and phenomenological background, as well as the motivation, for the measurement of W asymmetry at CMS. First, an overview of the Standard Model will be presented, followed by a discussion of the production and decay of W bosons and jets. The structure of the proton and the measurement thereof will be discussed, and an overview of parton distribution functions will be presented.

2.1 The Standard Model

The Standard Model (SM) is one of the biggest triumphs of physics in the past century. It consists of three of the four fundamental forces of nature: the strong force, the weak force, and the electromagnetic force (only gravity is excluded). It has been repeatedly verified, and all particles predicted by the SM have been discovered in experiments, with the notable exception of the Higgs Boson.

The SM encompasses two fundamental types of particle which are the building blocks of matter, and gauge particles which are the force carriers between particles. The fundamental particles are called quarks and leptons, of which there are three generations each. The gauge particles are the photon (for the electromagnetic force), the gluon (for the strong force) and W/Z bosons (for the weak force). Each fundamental force has a charge associated with it, whose value determines how a given particle behaves when acted upon by each force.

The Standard Model is based upon the idea of local gauge invariance, with the (non-

Abelian[2]) gauge symmetry group of the Standard Model being $SU(3)_C \times SU(2)_L \times U(1)_Y$[13]. SU(n) is the n-dimensional *special unitary* group, and is isomorphic to the group of $n \times n$ unitary matrices of determinant 1. U(n) is the n-dimensional *unitary* group consisting of all $n \times n$ unitary matrices. SU(n) has dimensionality $n^2 - 1$, and U(n) has dimensionality n^2. Both SU(n) and U(n) are Lie, or continuous, groups.

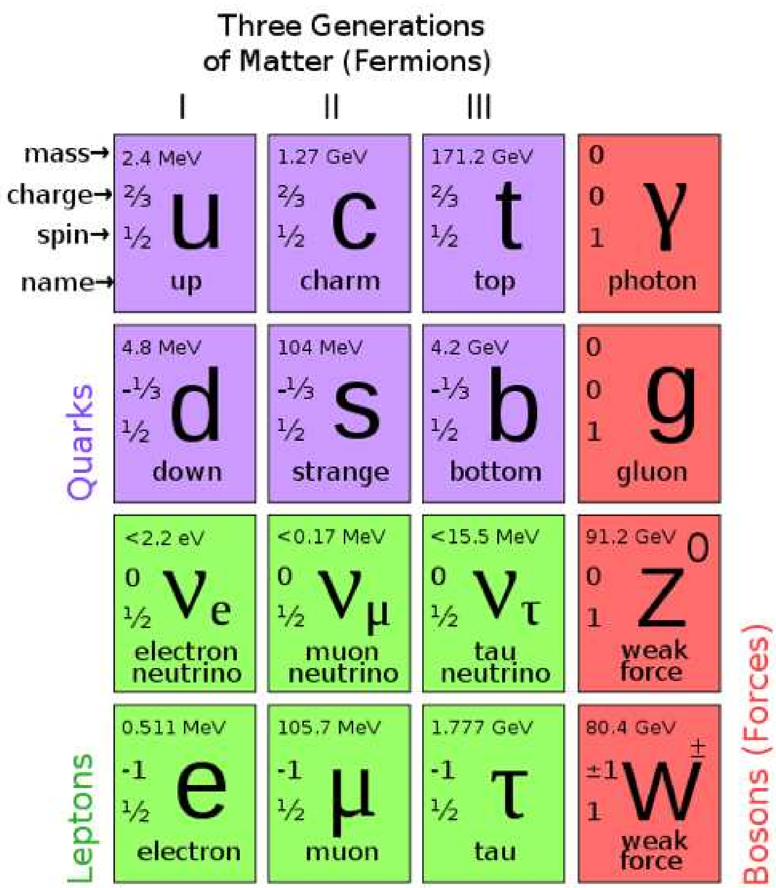


Figure 2.1: Table of the fundamental particles of the Standard Model.[1]

[2] An Abelian, or commutative, group is a group A with a binary operation $\cdot$ such that for all elements a, b in A, $a \cdot b = b \cdot a$.

2.1.1 Quarks

Quarks are fermions; that is, their spin takes on half-integer values and they behave according to Fermi-Dirac statistics. There are three generations of quarks, which are distinguished from one another by different masses. Each generation consists of an up-type quark with charge $+\frac{2}{3}e$ and a down-type quark with charge $-\frac{1}{3}e$. In order of increasing mass, the first generation consists of the up and down quarks, the second generation the charm and strange, and the third generation the top and bottom quarks. Quarks are *confined*; that is, they exist only within bound states. A bound state consisting of three quarks is called a baryon, and a bound state consisting of a quark and an antiquark is called a meson. All bound states must be color-neutral, and thus the three quarks in a baryon must each have a different color, and for mesons, the antiquark must have the anti-color of the quark.

2.1.2 Leptons

Like quarks, leptons exist in three generations, with each generation distinguished from the other by mass. Each generation consists of a massive particle and an associated (near-massless) neutrino, and the generations are comprised of the electron, the muon, the tau and their associated neutrinos. The massive particle carries an electric charge, whereas the neutrino ("little neutral one") is neutral. All neutrinos are left-handed in terms of chirality[3], and likewise all anti-neutrinos are right-handed.

2.1.3 Quantum Electrodynamics

The electromagnetic force has as its gauge field the electromagnetic field, its gauge group U(1) and its gauge particle the photon, a spin-1 boson. U(1) is an abelian gauge group, and local gauge invariance precludes the inclusion of a mass term in the electromagnetic Lagrangian; hence the gauge particle (the photon) is massless. The charge of the electromagnetic force is electric charge, and hence the EM force affects all charged particles[14].

The weak force's gauge group is SU(2). It has three gauge particles (as necessitated

[3]Chirality is defined as follows: a right-handed massless particle has eigenvalue $\lambda = +1$ for the operator γ^5, and a left-handed massless particle has eigenvalue $\lambda = -1$.

by its dimensionality $2^2 - 1 = 3$): the positive W boson, the negative W boson and the neutral Z. The W and Z bosons possess rather high masses and hence have a short lifetime, and as a result cannot be measured directly in particle accelerators. The weak force is unique in several aspects: parity is maximally violated (first observed in Co-60 decay[15]), the combination of charge symmetry and parity symmetry is slightly violated, it is the only force capable of changing the flavor of quarks, and it mediates decay of particles, rather than acting as a binding force. The charge is known as weak isospin, and both quarks and leptons interact via the weak force. The weak force and the electromagnetic force were unified into a single force with gauge group SU(2)$_L\times$U(1), which is described theoretically by Quantum Electrodynamics, or QED. As a consequence of the fact that QED is a parity-conserving theory whereas the weak force is not, quarks and leptons are divided into left-handed SU(2) doublets and right-handed SU(2) singlet states.

Applying the Euler-Lagrange equations to the Lagrangian for a free fermion and demanding local gauge invariance yields the electroweak Lagrangian

$$\mathcal{L} = \bar{\psi}\left(i\gamma^\mu \partial_\mu - m\right)\psi + e\bar{\psi}\gamma^\mu A_\mu \psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{2.1}$$

where A_μ is a vector gauge field and $F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu$ is the electromagnetic field strength tensor. The charged weak currents are

$$\begin{aligned} J_\mu^- &= \bar{\chi}_L \gamma_\mu \tau^- \chi_L \\ J_\mu^+ &= \bar{\chi}_L \gamma_\mu \tau^+ \chi_L \end{aligned} \tag{2.2}$$

and the neutral current is

$$J_\mu^3(x) = \bar{\chi}_L \gamma_\mu \frac{1}{2}\tau_3 \chi_L = \frac{1}{2}\bar{\nu}_L \gamma_\mu \nu_L - \frac{1}{2}\bar{e}_L \gamma_\mu e_L, \tag{2.3}$$

where $\chi_L = \binom{\nu_e}{e^-}_L$ is the left-handed doublet, $\tau_\pm = \frac{1}{2}\left(\tau_1 \pm i\tau_2\right)$ and τ_i denotes the ith Pauli spin matrix. The electromagnetic current is $j_\mu \equiv e j_\mu^{em} = e\bar{\psi}\gamma_\mu Q\psi$, where Q is the charge operator (and the generator for U(1)$_{em}$). By analogy, Y (defined as $Q = T^3 + \frac{Y}{2}$) is the hypercharge operator and generates $SU(2)_L$. Then, the electroweak interaction becomes

$$-ig\left(J^i\right)^\mu W^i_\mu - i\frac{g'}{2}\left(j^Y\right)^\mu B_\mu, \tag{2.4}$$

with

$$W^\pm_\mu = \sqrt{\frac{1}{2}}\left(W^1_\mu \mp iW^2_\mu\right) \tag{2.5}$$

describing the W bosons and the coupling constants g and g' related by

$$\frac{g'}{g} = \tan\left(\theta_W\right), \tag{2.6}$$

with θ_W known as the *Weinberg angle* or weak mixing angle[4]. Finally, the full electroweak Lagrangian is

$$\begin{aligned}\mathcal{L} = &\bar{\chi}_L\gamma^\mu\left[i\partial_\mu - g\frac{1}{2}\tau\cdot\mathbf{W}_\mu - g'\left(\frac{Y}{2}\right)B_\mu\right]\chi_L \\ &+ \bar{\psi}_R\gamma^\mu\left[i\partial_\mu - g'\left(\frac{y}{2}\right)B_\mu\right]\psi_R \\ &- \frac{1}{4}\mathbf{W}_{\mu\nu}\cdot\mathbf{W}^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}\end{aligned} \tag{2.7}$$

where $\mathbf{W}_{\mu\nu} = \partial_\mu\mathbf{W}_\nu - \partial_\nu\mathbf{W}_\mu - g\mathbf{W}_\nu\times\mathbf{W}_\nu$ (the cross product occurs because of the non-Abelian nature of $SU(2)_L$; that is, $[\tau_i,\tau_j]\neq 0$) and $B_{\mu\nu}\equiv\partial_\mu B_\nu - \partial_\nu B_\mu$ are the kinetic energy and self-coupling terms.

The Higgs mechanism uses spontaneous symmetry breaking to convert the scalar terms ψ into an additional degree of freedom each for the W and Z bosons (namely longitudinal polarization, which directly corresponds to the W and Z bosons being massive) and a fourth particle, the as-yet-undiscovered Higgs boson. Using a similar technique, and incorporating the Yukawa coupling

$$\mathcal{L}_{Yukawa} = \mathcal{L}_{Dirac} + \mathcal{L}_{Klein-Gordon} - g\bar{\psi}\psi\psi, \tag{2.8}$$

the masses for fermions can also be generated by spontaneously breaking the symmetry.

[4]Most recently measured to be $\sin^2\theta_W = 0.23122$.

Weak Flavor Mixing

The weak interaction is unique in the fact that the flavor of a quark can change while undergoing a weak decay. This was first noticed by Nicola Cabibbo in 1963[16], and later extended (upon discovery of the three generations of quark) by Makoto Kobayashi and Toshihide Maskawa, after whom the CKM matrix is named. The CKM matrix is a 3×3 unitary matrix whose elements V_{ij} contain the relative probability that a j quark decays into an i quark. Due to the fact that the CKM matrix is a 3×3 complex unitary matrix, it is specified by four parameters: three quark mixing angles and a complex phase, which causes CP violation.

The most recent experimentally measured values of the CKM matrix elements are as follows:

$$\begin{bmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{bmatrix} = \begin{bmatrix} 0.97428 \pm 0.00015 & 0.2253 \pm 0.0007 & 0.00347^{+0.00016}_{-0.00012} \\ 0.2252 \pm 0.0007 & 0.97345^{+0.00015}_{-0.00016} & 0.0410^{+0.0011}_{-0.0007} \\ 0.00862^{+0.00026}_{-0.00020} & 0.0403^{+0.0011}_{-0.0007} & 0.999152^{+0.000030}_{-0.000045} \end{bmatrix}$$

Mandelstam Variables

The Mandelstam variables are numerical quantities that encode the energy, momentum, and angles of particles in a 2-body $\rightarrow$ 2-body scattering process. Consider such a 2-body $\rightarrow$ 2-body process, and label the incoming momenta p and p' and the outgoing momenta k and k'. Then, the Mandelstam variables are defined as

$$s = (p + p')^2 = (k + k')^2 \tag{2.9}$$

$$t = (k - p)^2 = (k' - p')^2 \tag{2.10}$$

$$u = (k' - p)^2 = (k - p')^2. \tag{2.11}$$

For any process, s is the square of the total initial 4-momentum. If a 2-body $\rightarrow$ 2-body diagram contains only one virtual particle, then it can be categorized as belonging to one of three channels, with each channel possessing a characteristic angular dependence of the cross section.

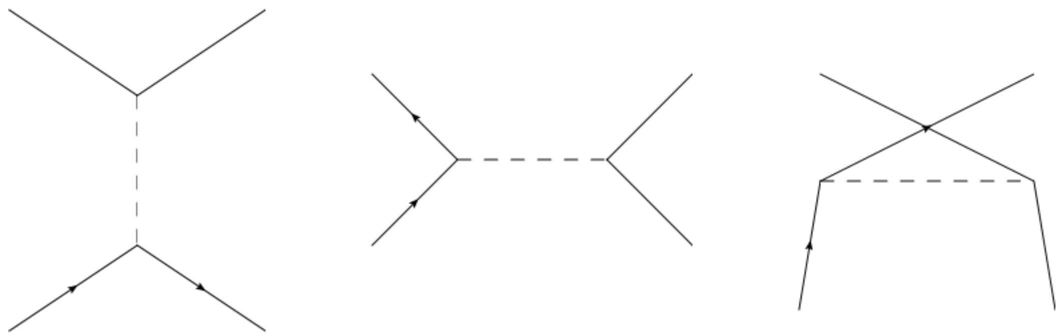


Figure 2.2: Channels of a 2-body → 2-body scattering process. From left to right: s-channel, with $\mathcal{M} \propto \frac{1}{s-m_\phi^2}$; t-channel, with $\mathcal{M} \propto \frac{1}{t-m_\phi^2}$, and u-channel, with $\mathcal{M} \propto \frac{1}{u-m_\phi^2}$.

2.1.4 Quantum Chromodynamics

The strong force's gauge particle is called the gluon, and as the strong force has dimensionality $3^2 - 1 = 8$ (its gauge group is SU(3)), there are 8 gluons. Its charge is the color charge of quarks, and takes values of red, green and blue. The strong force is governed by the theory of Quantum Chromodynamics (QCD). It affects quarks and gluons, and is the force that binds quarks together to form hadrons (such as the proton), as well as binding protons and neutrons together to form atoms. Gluons are unique in that they can bind to other gluons, unlike the photon or massive vector bosons.

An SU(3) local transformation takes the form

$$\psi \to \psi' = e^{-\frac{i}{2} g_s \theta_a(x) T_a} \psi, \tag{2.12}$$

with T_a being the SU(3) generators (namely the eight Gell-Mann matrices), $g_s = \sqrt{4\pi\alpha_s}$, and α_s being the coupling constant for strong interactions. Then the covariant derivative D is

$$D^\mu_{jk} = \delta_{jk}\partial^\mu + ig\left(T_a\right)_{jk} G^\mu_a, \tag{2.13}$$

with G_a^μ referring to the gluon fields, g being the strong coupling, and M_{jk} being the quark mass matrix. Then the gluon field tensor is

$$F_a^{\mu\nu} = \partial^\mu G_a^\nu - \partial^\nu G_a^\mu - g f_{abc} G_b^\mu G_c^\nu, \tag{2.14}$$

with f_{abc} being the structure constants of SU(3) and defined as $[T_a, T_b] = i f_{abc} T_c$. Then, the QCD Lagrangian resulting from the extension of Yang-Mills gauge theory to SU(3) is

$$\mathcal{L} = -\frac{1}{4} F_a^{\mu\nu} F_{\mu\nu a} + \bar{\psi}_j \left(i\gamma_\mu D_{jk}^\mu - M_j \delta_{jk} \right) \psi_k, \tag{2.15}$$

where the indices a, j and k are summed over color and assume the ranges $a = 1, \ldots, 8$ and $j, k = 1, 2, 3$.

Asymptotic freedom is a peculiar feature of QCD, and is the property that the interaction strength between particles becomes arbitrarily weak at extremely small and extremely large distances. This necessitates a scale-dependent or running coupling constant[17], which can be expressed as

$$\alpha_s \left(Q^2 \right) = \frac{2\pi}{\left(11 - \frac{2}{3} n_f \right) \log \left(Q / \Lambda_{QCD} \right)}, \tag{2.16}$$

where Q^2 is the energy scale, n_f is the number of active flavors (or number of quarks with energy less than the energy scale), and Λ_{QCD} is the scale at which the coupling becomes large, satisfying

$$1 = g^2 \left[\left(11 - \frac{2}{3} n_f \right) / 8\pi^2 \right] \log \left(M / \Lambda \right). \tag{2.17}$$

2.2 Proton Structure

2.2.1 Proton Overview

A proton is a bound hadronic state consisting of two up and one down "valence" quarks, and "sea" quark-antiquark pairs and gluons. It has a net charge of $+1e$, and is a spin-1/2 fermion. It is a stable particle which has never been observed to decay, and experimental searches for its decay have established the proton's lifetime as being orders of magnitude larger than the age of the universe. Electron-proton scattering experiments done by Robert Hofstadter in the 1950s showed that the proton had a finite charge radius, and Deep Inelastic

Scattering experiments at the then-new SLAC linear accelerator in the 1960s confirmed the presence of three quarks inside the proton. Although the exact momentum distribution of the constituent quarks and gluons cannot be calculated theoretically, they can be measured experimentally, and described by so-called Parton Distribution Functions, which were first measured at SLAC and later at HERA.

2.2.2 Parton Distribution Functions

Quarks and gluons can only exist in bound, colorless states called hadrons. Baryons are fermionic hadrons and consist of three quarks. The proton is an example of a baryon, and consists of two u quarks and a d quark, with each quark being a different color: red, blue or green. Mesons are bosonic hadrons which consist of a quark and an antiquark. An example of a meson is the J/ψ, which consists of a charm quark and a charm antiquark. In addition to the component valence quarks (which give hadrons their quantum numbers, and to which category the two u and one d quark in the proton belong) there are also virtual quark-antiquark pairs inside the hadron, which are known as sea quarks. Together, quarks and gluons are collectively known as partons.

Consider a process at a hadron collider of total energy $\sqrt{s}$ where a parton of type a comes from a hadron of type A, and a parton of type b comes from a hadron of type B. Then, a carries a fraction x_a of the A hadron's momentum, and b carries a fraction x_b of the B hadron's momentum. As a relevant example, consider W boson production at the LHC. Then, the probability of a hadron A carrying a parton a with momentum fraction x_a is $f_{a/A}(x_A)dx_A$, and the probability of a hadron B carrying a parton b with momentum fraction x_b is $f_{b/B}(x_B)dx_B$. The functions $f_{a/A}(x_A)$ and $f_{b/B}(x_B)$ are known as parton distribution functions. The cross section is then given by

$$\sigma_{p_A+p_B\to W} = \sum_{a,b} \int_0^1 \int_0^1 f_{a/p}(x_A) f_{b/p}(x_B) \hat{\sigma}\left(s, x_A, x_B\right) \mathrm{d}x_a \mathrm{d}x_b \tag{2.18}$$

Such an interaction thus depends on three variables: $f_{a/p}(x_A)$, $f_{b/p}(x_B)$, and the hadronic cross-section σ_{AB}. The latter can be calculated by using perturbation theory and expanding around α_S, the coupling constant for the strong force. However, the parton distribution functions (which are normalized probabilities and hence have integral values in the range $[0, 1]$) cannot be calculated using such a technique (they exist in a regime where α_S is too high for perturbative expansions[18]) and must be measured experimentally. The first measurements of PDFs were made at fixed-target deep inelastic scattering experiments at SLAC in the 1960s[19].

2.2.3 Deep Inelastic Scattering

Deep inelastic scattering (DIS) is an experimental method which is used to probe the inner structure of hadrons, and is conceptually similar to Rutherford scattering. The DIS experiments which were used to investigate the proton were performed by imparting an electron with very high energy, and colliding it with a proton. DIS is so named because the high energies of the electron allow it to probe very deep within the proton (more precisely, at a distance scale which is small compared to the size of the proton), and the high energies result in new hadron production (a process which disintegrates the proton, hence the "inelastic" moniker). The physical quantities which are measured in DIS are the final electron energy and scattering angle. Consider a deep-inelastic scattering process consisting of an electron scattering from a constituent quark inside a proton. Then, in the massless limit, the square of the invariant matrix element is given by

$$\frac{1}{4}\sum_{spins}|\mathcal{M}|^2 = \frac{8e^4Q_i^2}{\hat{t}^2}\left(\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2}\right), \tag{2.19}$$

where $\hat{s}$, $\hat{t}$, $\hat{u}$ are the Mandelstam variables for the electron-quark collision and Q_i is the electric charge of the ith quark in units of e. In the massless limit, $\hat{s}+\hat{t}+\hat{u}=0$ and hence the differential cross section in the center-of-mass system is

$$\frac{d\sigma}{d\cos\theta_{CM}} = \frac{1}{2\hat{s}}\frac{1}{16\pi}\frac{8e^4Q_i^2}{\hat{t}^2}\left(\frac{\hat{s}^2+\hat{u}^2}{4}\right) = \frac{\pi\alpha^2Q_i^2}{\hat{s}}\left(\frac{\hat{s}^2+(\hat{s}+\hat{t})^2}{\hat{t}^2}\right). \tag{2.20}$$

Then, since $\hat{t} = -\hat{s}(1-\cos\theta_{CM})/2$,

$$\frac{d\sigma}{d\hat{t}} = \frac{2\pi\alpha^2Q_i^2}{\hat{s}^2}\left(\frac{\hat{s}^2+(\hat{s}+\hat{t})^2}{\hat{t}^2}\right). \tag{2.21}$$

Then, the invariants $\hat{s}$ and $\hat{t}$ must be relatived to experimental observables. First, it is conventional to specify the invariant square of the momentum transfer q as a positive quantity:

$$Q^2 \equiv -q^2, \tag{2.22}$$

and the invariant $\hat{t}$ is simply $-Q^2$. To relate $\hat{s}$, first let x be the longitudinal fraction of the total momentum carried by each component parton of the proton, such that $0 < x < 1$. Then, for each species of parton i, there exists a function $f_i(x)$ that expresses the probability that the proton contains a parton of type i and momentum fraction x. The parton momentum vector is then $p = xP$, where P is the total proton momentum. Then if k is the initial electron momentum,

$$\hat{s} = (p + k)^2 = 2p \cdot k = 2xP \cdot k = xs, \tag{2.23}$$

where s is the square of the electron-proton center of mass energy. If one makes the further assumption that the electron-parton scattering is elastic, and since the scattered parton has a small mass compared to s and Q^2,

$$0 \approx (p + q)^2 = 2p \cdot q + q^2 = 2xP \cdot q - Q^2, \tag{2.24}$$

and so

$$x = \frac{Q^2}{2P \cdot q}. \tag{2.25}$$

Then from each scattered electron, the values of Q^2 and x for the scattering process can be determined. Finally, using Equation 2.21, the distribution in the x-Q^2 plane is given by

$$\frac{d^2\sigma}{dxdQ^2} = \sum_i f_i(x) Q_i^2 \cdot \frac{2\pi\alpha^2}{Q^4} \left[1 + \left(1 - \frac{Q^2}{xs}\right)\right]^2. \tag{2.26}$$

If the kinematic dependence of the QED cross section is removed by dividing Equation 2.26 by the factor

$$\frac{1 + (1 - Q^2/xs)}{Q^4}, \tag{2.27}$$

the deep inelastic scattering cross section depends only on x and not on Q^2. This feature is known as Bjorken scaling, and was experimentally verified at SLAC-MIT DIS experiments to an accuracy of 10% for values of Q above 1 GeV[20]. However, when QCD effects such as gluon radiation or gluon pair production are taken into account, Bjorken scaling is violated, and the PDFs depend on Q^2 logarithmically. Although QCD PDFs cannot be calculated theoretically from first principle (since the initial conditions required to integrate the equations are determined by the strong-coupling region of QCD and thus are not known *a priori*), their evolution can be described by the DokshitzerGribovLipatovAltarelliParisi, or DGLAP, equation

$$\frac{\partial}{\partial \ln Q^2} f_a(x, Q^2) = \sum_j P_{ab}(x, Q^2) \otimes f_b(x, Q^2), \tag{2.28}$$

where a and b are the different partons, $\otimes$ is defined as

$$A(x) \otimes B(x) = \int_0^1 dy \int_0^1 dz \delta(x - yz) A(y) B(z), \tag{2.29}$$

j is the summation index and runs over all species of quark and antiquark, and P_{ab} are the four splitting functions giving the probability densities for the lowest-order QCD corrections (or, $P_{ab}(x)$ is the probability that a parton b will radiate a parton a with z being the fraction of the original momentum carried by b), as seen in Figure 2.3. Then, the parton distribution functions are extracted by means of fitting to the DGLAP equation.

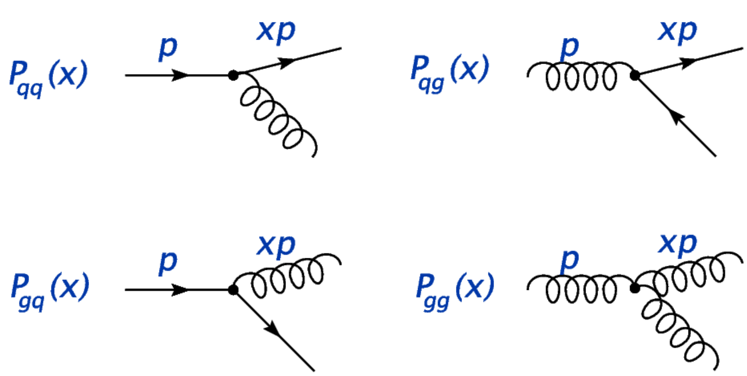


Figure 2.3: DGLAP splitting functions.

2.3 Production and Decay at the LHC

2.3.1 W Production and Decay

The W boson couples bottom states in the aforementioned SU(2) doublets to top states (e.g. e^- to $\bar{\nu}_e$, or u to d). As the W boson carries a charge of ± 1, charge is conserved in any allowable decay mode, and as the W boson is very massive (80.399 ± 0.023 GeV[11]), it decays quickly via many avenues. At the LHC, the W^+ is produced by the annihilation of an u quark and a $\bar{d}$ quark, and the W^- is produced by the annihilation of a d quark and a $\bar{u}$ quark. The W boson decays hadronically approximately 70% of the time, decaying into an up-type quark and a down-type quark (the possible decay channels are $u\bar{d}$, $u\bar{s}$, $u\bar{b}$, $c\bar{d}$, $c\bar{s}$, and $c\bar{b}$). The W is kinematically forbidden from decaying into the top quark, which has a greater rest mass than the W boson.

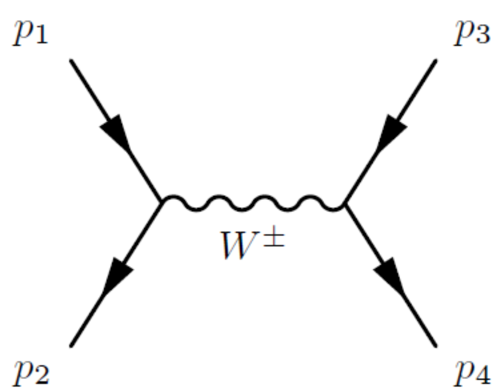


Figure 2.4: Tree-level Feynman diagram of W production at the LHC. Here, p_1 denotes an up-type quark, p_2 denotes a down-type quark, p_3 a lepton, and p_4 its associated neutrino.

The W production cross section is given by

$$\hat{\sigma}\left(q\bar{q}' \to W^+\right) = 2\pi \left|V_{qq'}\right|^2 \frac{G_F}{\sqrt{2}M_W^2}\delta\left(\hat{s} - M_W^2\right), \tag{2.30}$$

where $\hat{s} = (p_1 + p_2)^2$ and $\left|V_{qq'}\right|$ is the CKM matrix element corresponding to q and q'. Then, taking into account quark densities and a color factor,

$$\sigma\left(AB \to W^{+}X\right) = \frac{K}{3}\int_{0}^{1} dx_a \int_{0}^{1} dx_b \sum_{q} q\left(x_a, M_W^2\right)\bar{q}'\left(x_b, M_W^2\right)\hat{\sigma}, \tag{2.31}$$

with the assumption that $q^2 = \hat{s} = M_W^2$ is the appropriate scale of the quark distributions $x_{a,b}$, and with

$$K \simeq 1 + \frac{8\pi}{9}\alpha_s(M_W^2) \tag{2.32}$$

to first order in QCD. Then, define a new quantity *rapidity* as follows:

$$y = \frac{1}{2}\ln\left(\frac{E + p_z c}{E - p_z c}\right), \tag{2.33}$$

where p_z is the momentum component parallel to the beam axis.

Thus, the integration in (2.31) transforms as:

$$dx_a dx_b = \frac{d\hat{s}dy}{s} \tag{2.34}$$

and hence

$$\frac{d\sigma}{dy}\left(W^{+}\right) = K\frac{2\pi G_F}{3\sqrt{2}}\sum_{q,\bar{q}'}\left|V_{qq'}\right|^2 x_a x_b q\left(x_a, M_W^2\right)\bar{q}'\left(x_b, M_W^2\right), \tag{2.35}$$

with G_F the experimentally-determined weak coupling constant, and x_a and x_b evaluated at

$$x_{a,b} = \frac{M_W}{\sqrt{s}}e^{\pm y}. \tag{2.36}$$

Finally, for proton-proton scattering, the differential cross section in the Cabibbo mixing approximation (that is, a restriction to two generations of quark) is

$$\begin{aligned}\frac{d\sigma}{dy}\left(pp \to W^{+}X\right) =& K\frac{2\pi G_F}{3\sqrt{2}}x_a x_b \left\{\cos^2\theta_C\left[u\left(x_a\right)\bar{d}\left(x_b\right) + \bar{d}\left(x_a\right)u\left(x_b\right)\right]\right.\\ &\left. + \sin^2\theta_C\left[u\left(x_a\right)\bar{s}\left(x_b\right) + \bar{s}\left(x_a\right)u\left(x_b\right)\right]\right\},\end{aligned} \tag{2.37}$$

where $\theta_C \simeq 0.13\,^{\circ}$ is the Cabibbo angle, or mixing angle for quark interactions. In the further approximation of an $SU(3)$ symmetric sea this can be simplified to

$$\frac{d\sigma}{dy}\left(pp \to W^{+}X\right) = K\frac{2\pi G_F}{3\sqrt{2}}x_a x_b\left[u\left(x_a\right)\bar{d}\left(x_b\right) + \bar{d}\left(x_a\right)u\left(x_b\right)\right] \text{[21]}. \tag{2.38}$$

An unstable particle such as the W boson will decay exponentially, with the square of the time-dependent amplitude taking the form

$$|\psi(t)|^2 = |\psi(0)|^2 e^{-\Gamma t}, \tag{2.39}$$

where $\tau = 1/\Gamma$ is the lifetime of the particle, and Γ is known as the width. Then,

$$\psi(t)\ e^{-iE_0 t} e^{-\Gamma t/2}, \tag{2.40}$$

where E_0 is the rest mass energy of the state. Then, using a standard Fourier transform to convert the system from time-based to energy-based yields

$$\psi(E) = \int \psi(t) e^{iEt} dt\ \frac{1}{E - E_0 + (i\Gamma/2)}, \tag{2.41}$$

which in turn implies a cross-section proportional to

$$\sigma \propto \frac{1}{(E - E_0)^2 + \Gamma^2/4}. \tag{2.42}$$

This is known as the non-relativistic Breit-Wigner formula.

W bosons decay in one of two ways: leptonically (that is, $W \to \ell\nu$, where ℓ is an electron, muon or tau and ν is the associated neutrino) or hadronically (that is, to a quark and an antiquark). However, it is practically impossible to distinguish hadronic W decay from background; therefore, all W boson measurements to date have been made using the leptonic decay channel. The first detected decay channel of the W was $W \to e\nu$, at CERN. The decay widths are

$$\Gamma\left(W \to \ell_i \bar{\nu}_i\right) = \frac{\sqrt{2} G_F m_W^3}{12\pi} \tag{2.43}$$

$$\Gamma\left(W \to q_i \bar{q}_j\right) = 3 \frac{\sqrt{2} G_F |V_{ij}|^2 m_W^3}{12\pi} \tag{2.44}$$

for leptonic and hadronic decay, respectively[11]. The total decay width is 2.141 ± 0.041 GeV, with the relative fraction of the decay modes given in Table 2.1.

$\ell^+\nu$	$(10.80 \pm 0.09)\%$
$e^+\nu$	$(10.75 \pm 0.13)\%$
$\mu^+\nu$	$(10.57 \pm 0.15)\%$
$\tau^+\nu$	$(11.25 \pm 0.20)\%$
hadrons	$(67.60 \pm 0.27)\%$
$\pi^+\gamma$	$< 8 \times 10^{-5}$
$D_s^+\gamma$	$< 1.3 \times 10^{-3}$
cX	$(33.4 \pm 2.6)\%$
$c\bar{s}$	$(31 + 13/ - 11)\%$
invisible	$(1.4 \pm 2.8)\%$

Table 2.1: W^+ decay modes[11]. W^- modes are charge conjugates of the modes above.

2.3.2 Jets

A jet is defined as a narrow cone of hadrons produced by the hadronization of a quark or gluon in a hadron collider, and is primarily produced during hard scattering. It is very difficult to distinguish hadronic jets initiated by gluons from those initiated by quarks, and it is even more difficult to determine whether the initial partons in a hard-scattering process were quarks or gluons. Most hadron-hadron collisions at high energy will involve only soft interactions of the constituent partons, which are not calculable using perturbative QCD. For some collisions, however, two partons will exchange a large p_T relative to α_s, allowing for a perturbative treatment to lowest order.

Consider without loss of generality[5] a hard parton-level process involving scattering of a quark and an antiquark into a final state Y. The cross-section of the leading-order QCD prediction then takes the form

$$\sigma(p(P_1) + p(P_2) \to Y + X) = \int_0^1 dx_1 \int_0^1 dx_2 \sum_f f_f(x_1) f_{\bar{f}}(x_2) \cdot \sigma(q_f(x_1 P) + \bar{q}_f(x_2 P) \to Y), \tag{2.45}$$

summing over all relevant species of quarks and antiquarks. Jet pair production, however, is of order α_s^2, with α_s evaluated at a typical momentum transfer of the reaction. The

[5]The same principle applies to any other hadron-hadron collision, with appropriately modified PDFs.

Subprocess	$\mathcal{M}^2/g_s^4$
$\left.\begin{array}{l} qq' \to qq' \\ q\bar{q}' \to q\bar{q}' \end{array}\right\}$	$\frac{4}{9}\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2}$
$qq \to qq$	$\frac{4}{9}\left(\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2}+\frac{\hat{s}^2+\hat{t}^2}{\hat{u}^2}\right)-\frac{8}{27}\frac{\hat{s}^2}{\hat{u}\hat{t}}$
$q\bar{q} \to q'\bar{q}'$	$\frac{4}{9}\frac{\hat{t}^2+\hat{u}^2}{\hat{s}^2}$
$q\bar{q} \to q\bar{q}$	$\frac{4}{9}\left(\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2}+\frac{\hat{t}^2+\hat{u}^2}{\hat{s}^2}\right)-\frac{8}{27}\frac{\hat{u}^2}{\hat{s}\hat{t}}$
$q\bar{q} \to gg$	$\frac{32}{27}\frac{\hat{u}^2+\hat{t}^2}{\hat{u}\hat{t}}-\frac{8}{3}\frac{\hat{u}^2+\hat{t}^2}{\hat{s}^2}$
$gg \to q\bar{q}$	$\frac{1}{6}\frac{\hat{u}^2+\hat{t}^2}{\hat{u}\hat{t}}-\frac{3}{8}\frac{\hat{u}^2+\hat{t}^2}{\hat{s}^2}$
$qg \to qg$	$\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2}-\frac{4}{9}\frac{\hat{s}^2+\hat{u}^2}{\hat{u}\hat{s}}$
$gg \to gg$	$\frac{9}{4}\left(\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2}+\frac{\hat{s}^2+\hat{t}^2}{\hat{u}^2}+\frac{\hat{u}^2+\hat{t}^2}{\hat{s}^2}+3\right)$

Table 2.2: Matrix elements for $2 \to 2$ processes in QCD, averaged over spin and color. q and q' denote separate quarks, and $g_s^2 = 4\pi\alpha_s$ is the square of the QCD coupling constant.

cross-sections take the form

$$d\hat{\sigma}/d\hat{t}(ab \to cd) = |\mathcal{M}|^2/(16\pi\hat{s}^2), \tag{2.46}$$

where $\hat{s} = (a+b)^2$ is the subprocess center-of-mass energy squared, $\hat{t} = (a-c)^2$ and $\hat{u} = (a-d)^2$ are the Mandelstam variables, with the assumption of massless quarks (which implies $s+t+u=0$). The cross-sections for the various sub-processes are given by Table 2.2.

Computation of cross-sections for three or more jets becomes comparatively more complicated. There are four basic cases for the three-jet case at tree level, namely $qq' \to qq'g$, $qq \to qqg$, $q\bar{q} \to ggg$, and $gg \to ggg$, to which all other cases are related by crossing.

2.3.3 Drell-Yan

The Drell-Yan process is the reaction in which a high-mass lepton pair l^+l^- emerges from $q\bar{q}$ annihilation in proton-proton collisions, and is the source of much of the background processes in V+jets analyses[22]. The cross-section is very similar to the fundamental $e^+e^- \to \mu^+\mu^-$ cross section, with additional terms for color and charge factors:

$$\sigma\left(q\bar{q} \to e^+e^-\right) = \frac{4\pi\alpha^2}{3\hat{s}}\frac{1}{3}e_q^2, \tag{2.47}$$

where e_q is the quark charge in units of e, and the $1/3$ factor is due to the fact that q and $\bar{q}$ must have matching color in order to annihilate. Then, the differential mass distribution is

$$\frac{d\hat{\sigma}}{dm^2} = \frac{\hat{\sigma}_0}{3} e_q^2 \delta(\hat{s} - m^2), \tag{2.48}$$

where m is the mass of the lepton pair, and

$$\hat{\sigma}_0 = \frac{4\pi\alpha^2}{3m^2}. \tag{2.49}$$

Now, let $M^2 = q^2$. Then, in the center-of-mass frame of the two protons, the proton momenta take the forms

$$P_1 = (E, 0, 0, E) \tag{2.50}$$

$$P_2 = (E, 0, 0, -E), \tag{2.51}$$

with E satisfying the relation $s = 4E^2$, where s is the proton center-of-mass energy. Then, it is possible to write the quark and antiquark momenta as functions of the vectors P_1, P_2, x_1 and x_2, like so:

$$q = x_1 P_1 + x_2 P_2 = \left((x_1 + x_2) E, 0, 0, (x_1 - x_2) E \right). \tag{2.52}$$

From here, it follows that $M^2 = x_1 x_2 s$, and

$$\cosh y = \frac{x_1 + x_2}{2\sqrt{x_1 x_2}} = \frac{1}{2} \left(\sqrt{\frac{x_1}{x_2}} + \sqrt{\frac{x_2}{x_1}} \right), \tag{2.53}$$

where y is the rapidity of the virtual photon. This in turn implies

$$e^y = \sqrt{\frac{x_1}{x_2}}, \tag{2.54}$$

which finally yields

$$x_1 = \frac{M}{\sqrt{s}} e^y \tag{2.55}$$

$$x_2 = \frac{M}{\sqrt{s}} e^{-y}. \tag{2.56}$$

Thus, it is possible to convert Equation 2.45 into an integral over the variables M^2 and

y, yielding the final cross-section for lepton pair production[23]:

$$\frac{d^2\sigma}{dM^2dy}\left(pp \to l^+l^- + X\right) = \sum_f x_1 f_f(x_1) x_2 f_{\bar{f}}(x_2) \cdot \frac{1}{3} e_f^2 \cdot \frac{4\pi\alpha^2}{3M^4}. \tag{2.57}$$

2.4 W Asymmetry

2.4.1 Momentum Fraction

Rapidity, defined in Equation 2.33, is a Lorentz-invariant quantity, which is important due to the highly relativistic nature of hadrons at the LHC. Consider a process $q\bar{q} \to W$. Then, since energy is conserved, $E_W = E_q + E_{\bar{q}}$. Since the particles are highly relativistic, $E_q = p_q$ and

$$E_W = p_q + p_{\bar{q}}. \tag{2.58}$$

Then, from the conservation of momentum, $\overrightarrow{p}_W = \overrightarrow{p}_q + \overrightarrow{p}_{\bar{q}}$, which reduces to the scalar equation

$$p_W = p_q + p_{\bar{q}} \tag{2.59}$$

for all momenta along the z-axis, since the beam is one-dimensional. Then, combining Equations 2.58 and 2.59 with Equation 2.33 yields

$$y = \frac{1}{2} \log \frac{p_q}{p_{\bar{q}}}, \tag{2.60}$$

which can be written in terms of the momentum fraction x as

$$y = \frac{1}{2} \log \frac{x_q}{x_{\bar{q}}}, \tag{2.61}$$

or

$$e^{2y} = \frac{x_q}{x_{\bar{q}}}. \tag{2.62}$$

Then, using the relation

$$E_W^2 = p_W^2 + m_W^2 \to x_q x_{\bar{q}} = \frac{m_W^2}{s}, \tag{2.63}$$

where $s = 2E_{beam}$ thus implies

$$
\begin{aligned}
x_q &= \frac{m_W}{\sqrt{s}} e^{y} \\
x_{\bar{q}} &= \frac{m_W}{\sqrt{s}} e^{-y}
\end{aligned}
\tag{2.64}
$$

Thus, at the LHC, it is expected for the charge asymmetry (proportional to the momentum fraction carried by the valence quark) to increase with rapidity. However, since rapidity is not directly accessible for the W boson, the quantity pseudorapidity is used instead.

At lower pseudorapidity ranges, the measurement of the W charge asymmetry is very closely related to the W production asymmetry. However, leptonic decays of the W are governed by vector-axial (V-A) coupling, imposing helicity requirements. Helicity is defined as the relationship between a particle's momentum and its spin, and is given by

$$
H = \frac{\mathbf{s} \cdot \mathbf{p}}{|\mathbf{s} \cdot \mathbf{p}|}, \tag{2.65}
$$

where $\mathbf{s}$ is the spin vector (helicity is clearly undefined when $\mathbf{p} = 0$). Then, $H = -1$ for a left-handed particle, and $H = 1$ for a right-handed particle[6]. The nature of the electroweak theory requires that W bosons couple exclusively to left-handed quarks and leptons or right-handed antiquarks and antileptons. Thus, when a W^+ boson is produced, the u quark is left-handed and the $\bar{d}$ quark, being an antiparticle, is right-handed, and when a W^+ decays muonically, the μ^+ (an antiparticle) must also be right-handed, and the neutrino left-handed. Therefore, the direction of motion of the μ^+ must be antiparallel to the proton from whence the u quark came; thus making the decay antisymmetric. The measured muon asymmetry is a convolution of the decay asymmetry and the production asymmetry.

2.4.2 W Asymmetry at Hadron Colliders

The W asymmetry has previously been measured at the Tevatron, which is a proton-antiproton collider. On average, the u quark in a proton carries a greater fraction of the momentum than the d quark, and as the direction of the proton beam defines the positive

[6] Helicity and chirality are related, although there is a subtle distinction. Helicity and chirality are the same for massless particles, whereas in massive particles if the observer is in a reference frame moving faster than a particle, its helicity is reversed compared to a slower reference frame, whereas chirality is invariant.

rapidity direction, it is expected that W^+ boson production will preferentially occur in the $+y$ direction, and W^- production in the $-y$ direction. However, as the overall net charge of a $p\bar{p}$ collision is zero, the overall asymmetry should be zero, with a positive/negative asymmetry in the positive/negative y direction, respectively. The LHC, on the other hand, is a proton-proton collider, with an expected net positive charge asymmetry increasing with increasing absolute rapidity.

Chapter 3
The CMS Experiment

3.1 LHC Overview

The LHC is a proton-proton collider located at the Organisation Européene pour la Recherche Nucléaire (CERN), just outside Geneva, Switzerland. The LHC is housed in the tunnel formerly used for the Large Electron Positron collider, which has a circumference of 26.7 km, a depth of approximately 100 meters below the surface, and consists of 8 straight sections and 8 arcs. It is designed to have a center-of-mass energy of $\sqrt{s} = 14$ TeV, and is fed from an existing accelerator chain (Linac/Booster/PS/SPS). The protons are arranged in a bunch structure, with various gaps introduced for the purposes of synchronization, acquiring calibration data and providing resets to front-end electronics[3]. The bunches are formed in the PS, a 26 GeV accelerator, with the correct 25 ns spacing, and are subsequently accelerated up to 450 GeV in the SPS and injected into the LHC, an operation which is repeated 12 times per counter-rotating beam.

There are eight specific access points in total situated around the LHC ring. Of these, four are collision points, with two new experimental sites, Point 1 (ATLAS) and Point 5 (CMS), and two sites inherited from LEP, Point 2 (ALICE) and Point 8 (LHCb). The other four points are specific to beam operations and access. The protons are accelerated through the ring via electric fields inside radio frequency cavities, which transfer approximately 0.5 MeV per pass through each RF cavity from the radio waves to the protons. A near-circular layout for the LHC is used so that the protons may pass through the RF cavities multiple times, so as to attain the necessary energy. There are 2 independent RF systems, each with

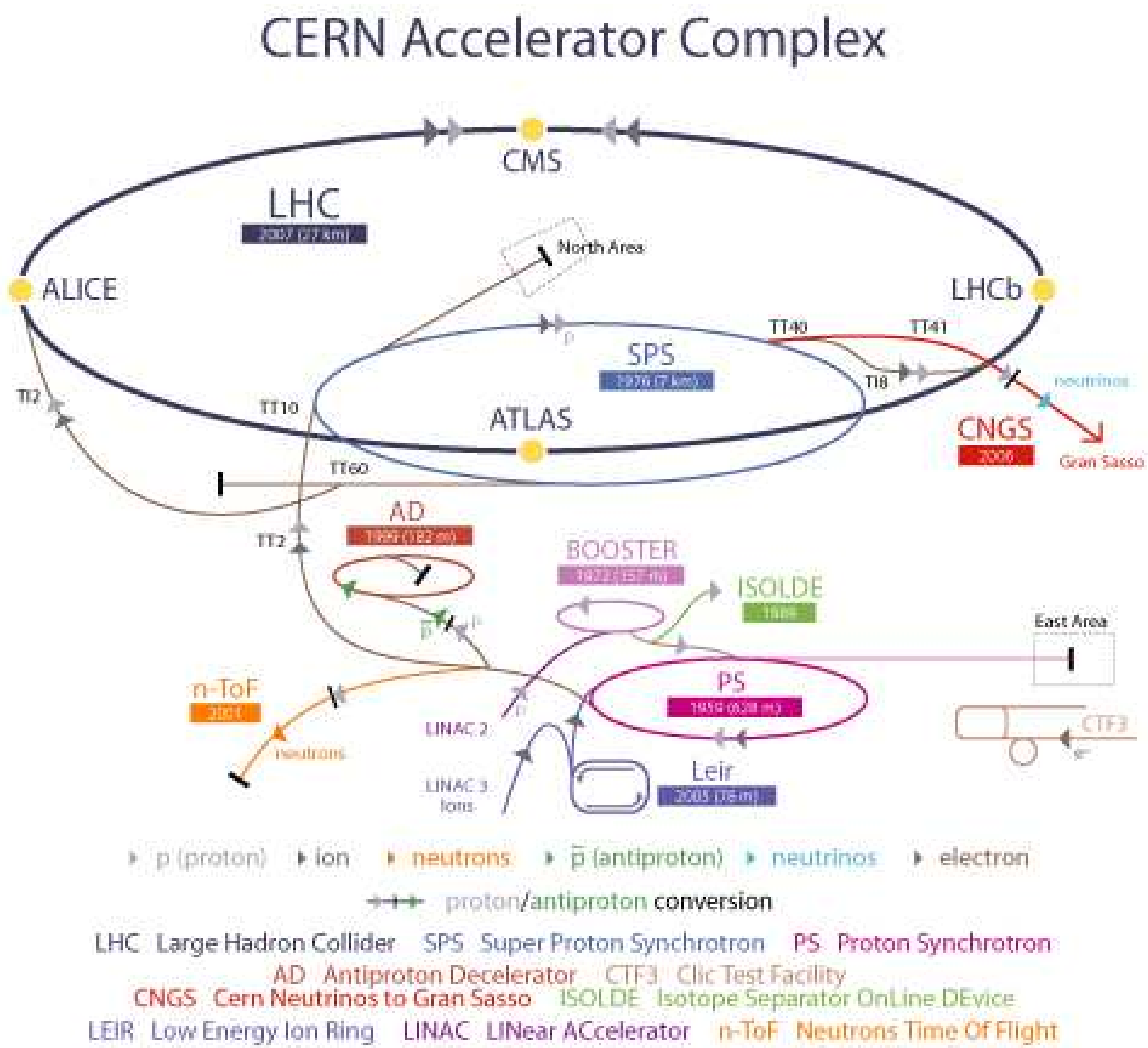


Figure 3.1: View of the LHC accelerator complex[1]

8 RF cavities at a frequency of 400 MHz. There is a loss of approximately 6.7 keV per revolution due to synchrotron radiation. The protons are kept on a circular trajectory by a peak magnetic field of 8.33 T, which is provided by 1,232 superconducting dipole magnets. There are 8 focusing stations (2 per interaction region), with each containing 4 quadrupole magnets to focus the antiparallel beams and cause them to intersect at the interaction points. Overall, there are 392 quadrupole magnets at the LHC[24]. The dipole magnets are cooled via liquid helium to -271 C.

Luminosity is defined as the number of particles per unit area per unit time times the opacity of the target, and integrated luminosity is the luminosity integrated with respect to time. Thus, the number of events per second generated in LHC collisions is given by

$$N_{event} = L\sigma_{event}, \tag{3.1}$$

where σ_{event} is the cross-section of the event, measured in barns[7]. L, the machine luminosity at the LHC, is measured in units of inverse barns per second and given by

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi\epsilon_n \beta_*} F, \tag{3.2}$$

where N_b is the number of particles per bunch, n_b is the number of bunches per beam, f_{rev} the revolution frequency, γ_r the relativistic gamma factor, ϵ_n the normalized transverse beam emittance, β_* the beta function at the collision point, and F the geometric luminosity reduction factor due to the crossing angle at the interaction point[26]:

$$F = \left(1 + \left(\frac{\theta_c \sigma_z}{2\sigma_*}\right)^2\right)^{-1/2} \tag{3.3}$$

The cross section σ is the effective cross-sectional area of the target seen by the incoming particles, and the differential cross section $d\sigma/d\Omega$ is the cross-sectional area per unit of solid angle, and thus the relationship between $d\sigma/d\Omega$ and σ is given by $\sigma = \int d\Omega(d\sigma/d\Omega)$. Then,

[7] A barn is a unit of cross-sectional area, and is defined to be $10^{-28}m^2$. It roughly corresponds to the cross-section of a uranium nucleus, and was coined during the Manhattan Project by Purdue scientists, who described the uranium nucleus as being "as big as a barn"[25].

it follows that $dN/dt = L\sigma$, where N is the number of interactions, and hence

$$\frac{d\sigma}{d\Omega} = \frac{1}{L}\frac{d^2N}{d\Omega dt}. \tag{3.4}$$

The luminosity upon startup in November 2009 was much lower than the design luminosity (corresponding to 8×10^{29} cm^{-2} s^{-1}, for two beams of 450 GeV each). After the Winter 2009 shutdown, collisions at $\sqrt{s} = 7$ TeV were performed, and successive increases in luminosity were achieved, with a total instantaneous luminosity of 10^{33} cm^{-2} s^{-1}, leading to an integrated total luminosity for 2010 of 35.9 recorded inverse picobarns, where one picobarn is equivalent to 10^{-40} m^2.

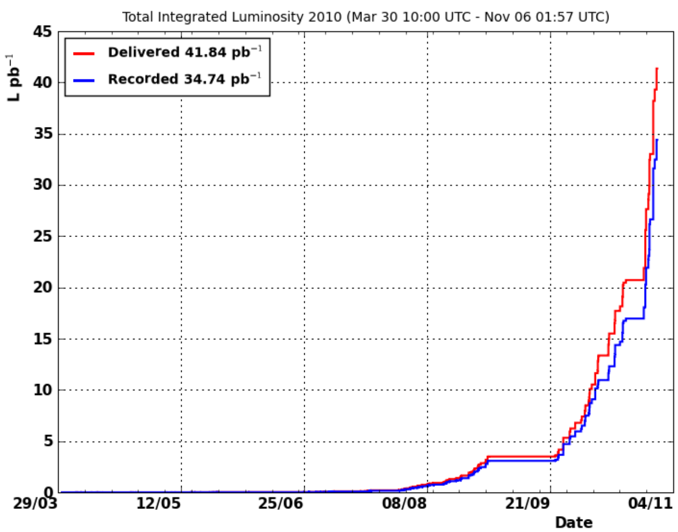


Figure 3.2: Total luminosity delivered/recorded during 2010[2]

3.2 The Compact Muon Solenoid

CMS is a multi-purpose detector located at Point 5 of the LHC, symmetric about the I.P. It is designed with the following considerations in mind:

Energy	[TeV]	7.0
Dipole field	[T]	8.33
Magnet temperature	[K]	1.9
Coil aperture	[mm]	56
Distance between apertures	[mm]	194
Luminosity	[$cm^{-2}s^{-1}$]	10^{34}
Injection energy	[GeV]	450
Circulating current/beam	[A]	0.54
Bunch spacing	[ns]	25
Bunch spacing	[m]	7.48
Particles per bunch		10^{11}
Stored beam energy	[MJ]	334
RMS bunch length	[m]	0.075
Full crossing angle	[μrad]	200
Beam lifetime	[h]	22
Luminosity lifetime	[h]	15
Energy loss per turn	[keV]	6.7
Total radiated power per beam	[kW]	3.6
Number of bunches		$k_B = 2808$
Number of particles per bunch		$N_p = 1.15 \times 10^{11}$
RMS beam radius at interaction point	[μm]	16.7
Number of collisions/crossing		$n_c \approx 20$

Table 3.1: LHC performance parameters[12],[3]

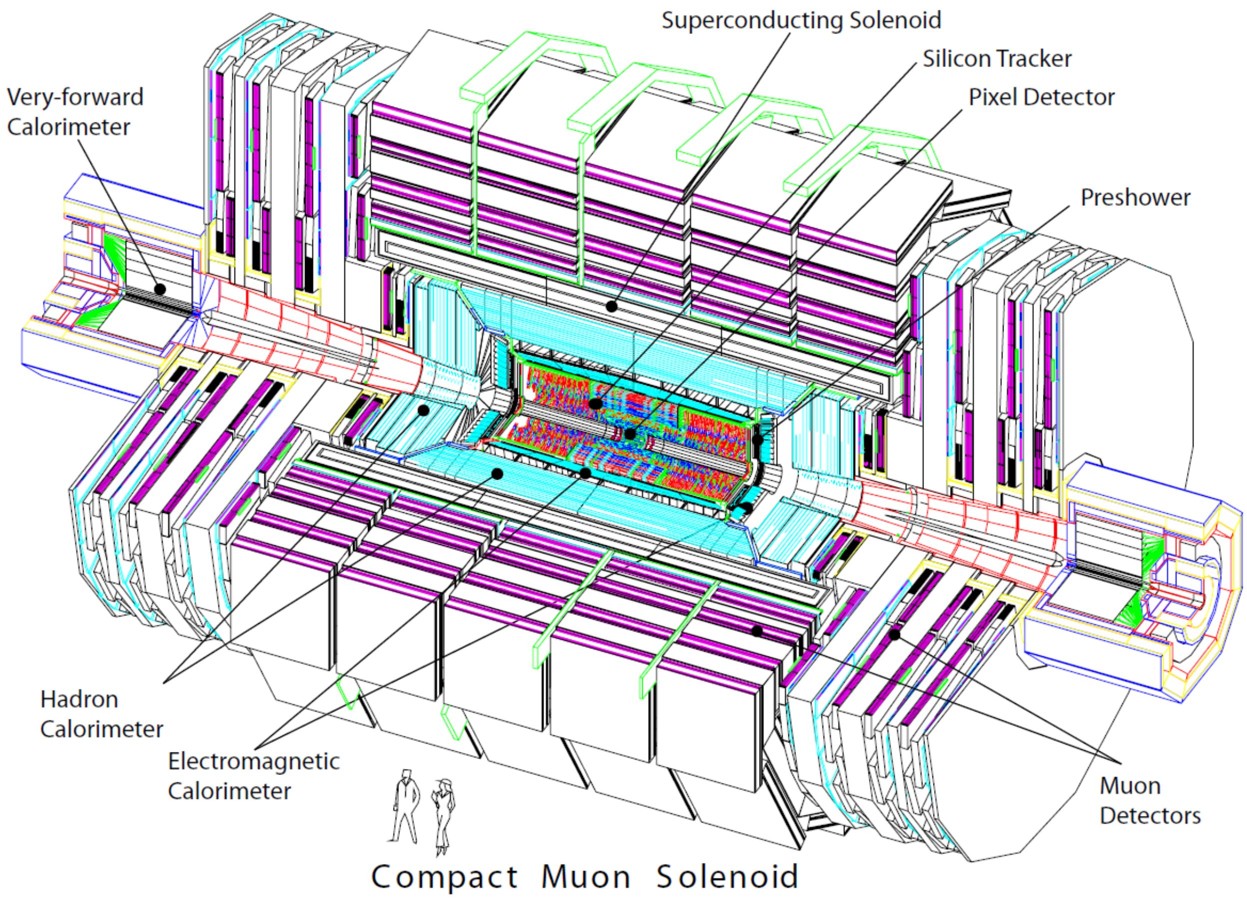


Figure 3.3: Blown-up view of CMS[3]

- Good muon identification and momentum resolution over a wide range of momenta and angles, good dimuon mass resolution, and unambiguous charge resolution;
- Good charged-particle momentum resolution and reconstruction, as well as good triggering and tagging of τ particles and b-jets;
- Good photon and electron energy resolution, diphoton and dielectron mass resolution, and efficient photon and lepton isolation;
- Good missing transverse energy, jet resolution and dijet mass resolution[5].

The centerpiece of CMS is the superconducting solenoid magnet, capable of a magnetic field of 3.8 Tesla. Inside the bore of the magnet coil are situated the pixel detector, the silicon tracker, the hadronic and electromagnetic calorimeters, and muon drift tubes and resistive plate chambers. Situated at either end of CMS are the two muon endcaps. There is also a very-forward calorimeter situated near the beampipe on either end of CMS.

3.2.1 Coordinate Conventions

Per the convention of CMS, the origin is the collision point at the center of the detector. The x-axis points radially inward towards the center of the LHC, the y-axis points vertically upward towards the surface, and the z-axis points along the beam direction, towards the Jura mountains. The azimuthal angle ϕ is measured from the x-axis in the x-y plane, the polar angle θ is measured from the z-axis, and the radial distance ρ is defined as $\rho = \sqrt{x^2 + y^2}$. Then, the pseudorapidity is the angle of a particle relative to the beam axis, and is equal to 0 at an angle perpendicular to the beam axis and ∞ at $0\,^\circ$. Mathematically, the pseudorapidity is defined as $\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right)$, and is a highly useful quantity in terms of energy corrections and alignment (a histogram of number of events versus pseudorapidity should ideally be flat, in the absence of detector biases). The transverse energy E_T and momentum p_T are computed from the x and y components.

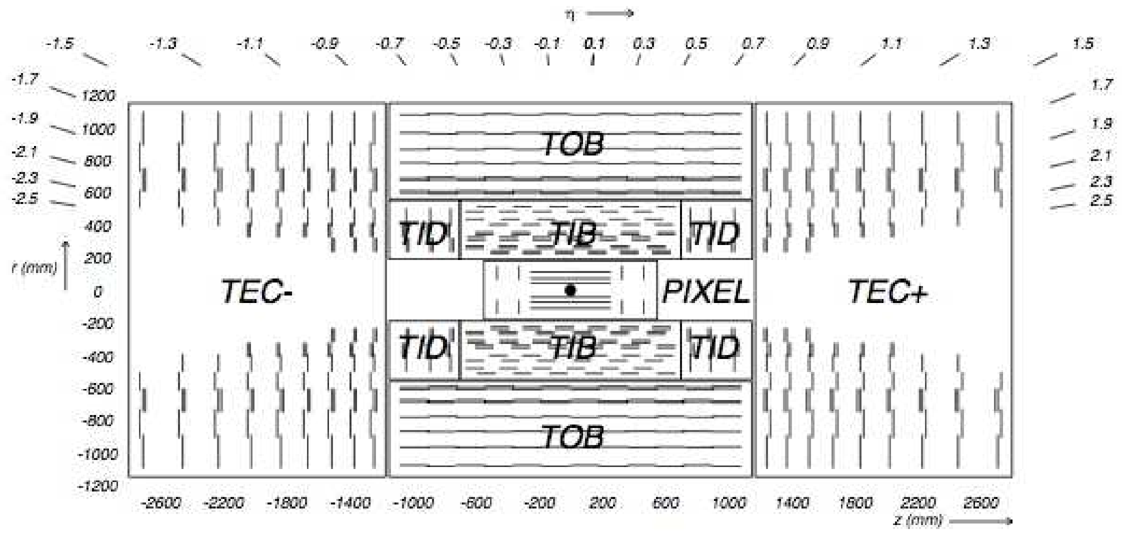


Figure 3.4: Cross-section of the CMS tracker[4].

3.2.2 Magnet

The magnet is based on a similar design to the solenoids used in ALEPH and DELPHI at LEP and H1 at HERA, and comprises 6 meters in diameter and 12.5 meters in length. It has 2,168 turns with a current of 19.5 kA. It is composed of three main parts: a superconducting coil, a vacuum tank and the magnet yoke. It uses a high-purity aluminium-stabilized conductor. The solenoid is 13 m long, with a 5.9 m inner diameter and a design field strength of 4 T, which requires 4 layers of windings. The yoke is responsible for the return of magnetic flux and to help ensure relative uniformity of the magnetic field within the inner radius of the solenoid. It consists of 5 wheels and 3 endcaps, composed of 3 disks each. The vacuum tank contains two double-primary pumps providing a vacuum to the oil diffusion pumps and the coil cryostat, with a cooling capacity of 800 W at 4.45 K. The favorable dimensional ratio (length/radius) of the solenoid and the high field allow efficient muon detection and measurement up to a pseudorapidity of $\eta = 2.4$[27]. In addition, the high magnetic field allows for good momentum resolution, due to the more pronounced bending in the field.

3.2.3 Tracker

The purpose of the tracker is to identify and measure the momentum of charged particles, as well as reconstruct secondary vertices. The tracker, whose components are housed in a cylindrical volume with a length of 5.4 m and a diameter of 2.4 m, is subdivided into four silicon strip subdetectors (the Tracker Outer Barrel, the Tracker Inner Barrel, the Tracker Inner Disk and the Tracker Endcap), and two silicon pixel subdetectors, the pixel barrel and the pixel discs. The tracker is wholly contained within the homogeneous magnetic field of 3.8 T. At the center, closest to the I.P. where the particle flux is the highest ($\approx 10^7$ /s at $r \approx 10$ cm), pixel detectors are used. In the outer region, silicon microstrip detectors are placed. The tracker is designed to provide coverage up to pseudorapidity $|\eta| \leq 2.5$ [28]. The pixel detector is housed in a volume of 1 m length and 30 cm diameter, centered around the I.P. It consists of 6.6×10^7 pixels of size 100 μm $\times$ 150 μm, and is distributed over three barrel layers (at $r = 4.4$ cm, 7.3 cm and 10.2 cm). The endcap disks cover the range 6 cm $\leq r \leq 15$ cm and are assembled in a turbine wheel-like geometry on blades[29].

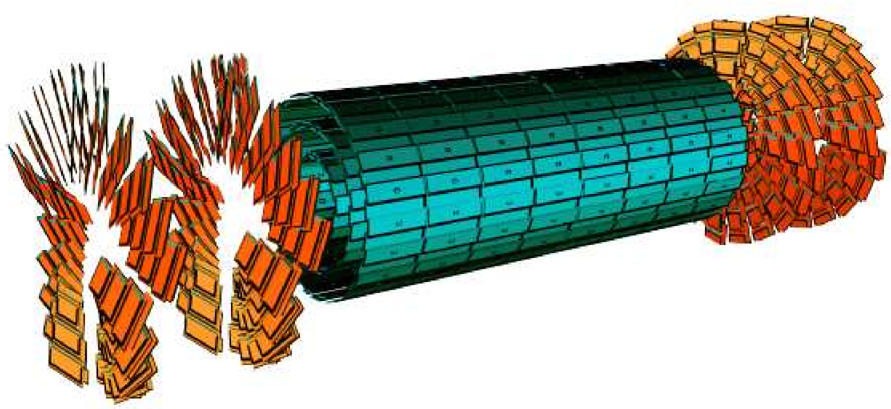


Figure 3.5: Expanded view of the pixel subdetector, showing the three layers of the barrel pixel detector and the four separate endcap disks[30].

Pixels were chosen due to the high rate which results from being located so close to the I.P., and also due to the fact that they have a response time shorter than the 25 ns bunch crossing interval. The Silicon Strip Tracker consists of 15,148 silicon strip modules distributed over 10 barrel layers, and 3 TID and 9 TEC disks. Stereo geometries are used in the first two double-sided barrel layers (TIB and TOB), TID rings 1 and 2 and TEC rings 1, 2, and 5. The usage of stereo geometry helps to filter out "ghost" hits, which are false hits that occur when an individual particle is counted two or more times due to geometry ambiguities[31]. The spatial resolution of the pixel detectors is about 10 μm for the r–ϕ measurement and about 20 μm for the z measurement, and the spatial resolution of the silicon strips is 23–34 μm in the r–ϕ direction and 230 μm in z[3].

Track reconstruction is done in four stages: trajectory seeding, pattern recognition, trajectory cleaning and track fitting and smoothing[32]. Trajectory seeding provides seeds, or initial trajectory candidates, for further reconstruction, based on pairs of hits which are selected to be compatible with the interaction region and a lower p_T limit, taking into account multiple scattering. In general, pixel hits provide the best track seeding, due to their lower occupancy and three-dimensional position information, and so they are used, subject to a vertex constraint. Their seeding efficiency drops in the high-η forward region ($2 < |\eta| < 2.5$). There, to achieve a more efficient track finding, a mixed seeding of hits from pixels and inner strips is required.

Pattern recognition and track fitting is based on a standard combinatorial Kalman filter pattern recognition algorithm. It begins with a first estimate of the track parameters, calculated from the seed, and then proceeds iteratively to collect the full set of hits for a charged particle track. Starting from the current parameters the trajectory is extrapolated to the next layer and compatible hits are selected based on the χ^2 value between the predicted and measured positions, taking into account energy loss and multiple scattering. Ambiguities which may result in double counting are resolved via a cut on the fraction of hits which are shared between two different trajectories. Such a cut is applied twice: first on all trajectories from a single seed, and again on all tracks from all seeds.

Track fitting and smoothing consists of a least-squares fit for the final estimation of the

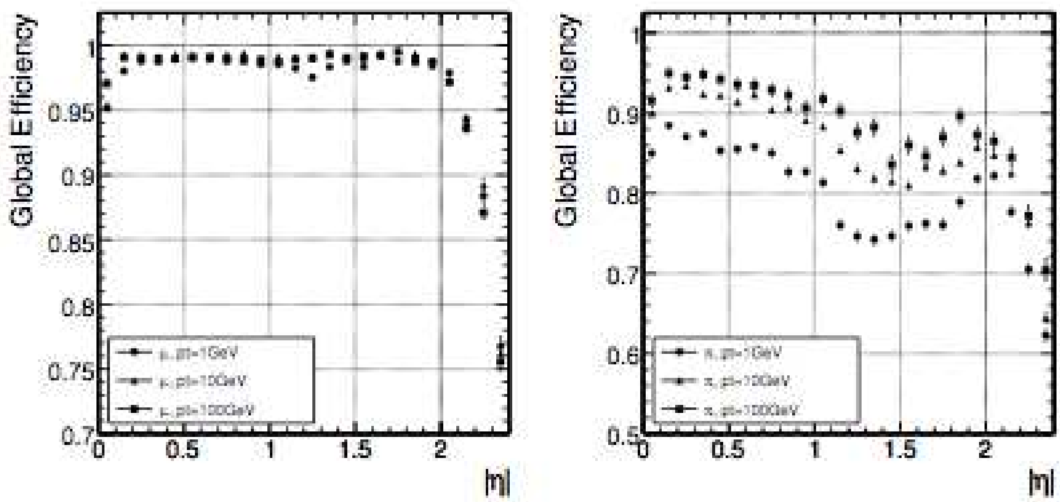


Figure 3.6: Global track reconstruction efficiency of the CMS tracker, for p_T = 1, 10 and 100 GeV. Left plot is muons, right plot is pions[5].

track parameters. A "forward" fit proceeding outwards from the interaction region removes the approximations used in the track finding stage and provides an optimal estimate of the track parameters at the outside of the tracker. A "backward" fit in the opposite direction yields the estimate of the track parameters in the interaction region and–in combination with the forward fit–at each of the intermediate layers[33]. For both forward and backward fits, the track parameters are scaled to reduce biases. At each hit the updated parameters of the smoothing filter are combined with the predicted parameters of the first filter[32].

3.2.4 Calorimeters

The combined CMS calorimeter system will measure the energy and direction of particle jets and of the missing transverse energy flow[34].

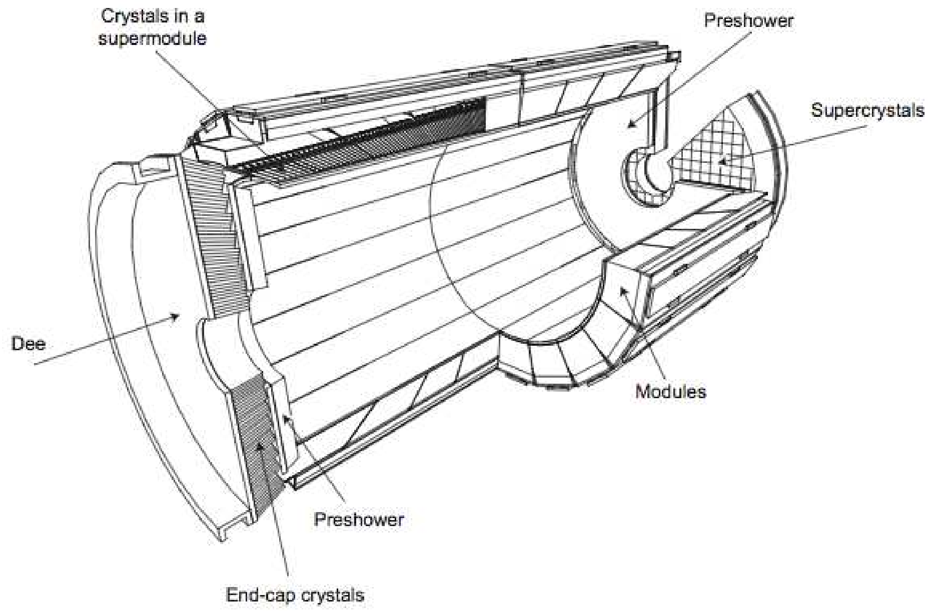


Figure 3.7: Layout of the CMS Electromagnetic Calorimeter[5].

Electromagnetic Calorimeter

The purpose of the electromagnetic calorimeter (ECAL) is to precisely measure the energy and direction of electrons and photons. Lead tungstate ($PbWO_4$) crystals were chosen as the basis for the ECAL, due to short radiation length ($X_0 \simeq 0.89$ cm), small Molière radius ($R_M \simeq 2.2$ cm), fast scintillation (same order of magnitude as LHC bunch crossing time) and ease and cost-effectiveness of production. The geometrical coverage of the ECAL extends to pseudorapidity $|\eta| = 3$, however precision energy measurements of photons and electrons will be carried out to $|\eta| = 2.6$ due to considerations of the radiation dose and compatibility with the inner tracker[35]. The crystal front face is 22×22 mm^2, and the crystal length is 23 cm in the barrel region, and 22 cm in the endcap region (the presence of the preshower allows for a shorter length). In total, the ECAL comprises 83,000 crystals. The energy resolution is not significantly decreased upon exposure to radiation[35].

The physical layout of the ECAL can be divided into segments, namely barrel and endcap. The barrel part of the ECAL covers the pseudorapidity range $|\eta| < 1.479$. The truncated pyramid-shaped crystals are mounted in a geometry which is off-pointing with respect to the mean position of the primary interaction vertex, with a 3° tilt in both ϕ and in η. The barrel granularity is 360-fold in ϕ and (2×85)-fold in η, resulting in a total number of 61,200 crystals[36]. The crystal volume in the barrel amounts to 8.14 m^3, with crystals for each half-barrel grouped into 18 supermodules each subtending 20° in ϕ[36], for a total of 7,324 endcap crystals. The total radiation depth is approximately $25.8X_0$ in the barrel and $24.7X_0$ in each endcap[37].

There are two identical endcap parts of the crystal calorimeter, each comprised using two dee-shaped sections, and covering the range $1.48 < |\eta| < 3.0$, with precision energy measurement to $|\eta| = 2.6$. However, crystals will be installed up to $|\eta| = 3$, in order to augment the energy-flow measurement in the forward direction. The endcap calorimeter uses tapered crystals of the same shape and dimensions ($24.7 \times 24.7 \times 220$ mm^3) grouped together into units of 36, referred to as supercrystals. A total of 268 identical supercrystals will be used to cover each endcap with a further 64 sectioned supercrystals used to complete the inner and outer perimeter.

The endcap preshower covers a pseudorapidity range from $1.65 < |\eta| < 2.61$, and is located in front of the endcap crystals. Its main function is to provide $\pi^0 - \gamma$ separation. In the barrel, an optional preshower covers the pseudorapidity range up to $|\eta| = 0.9$ to enable measurement of the photon angle to an accuracy of about 45 mrad/$\sqrt{E}$ in the η direction. This detector will be built and installed only for the high-luminosity operation, if the activity of the minimum-bias events seen at LHC startup shows that additional angular determination is necessary[36].

Upon entry into the ECAL region, high-energy photons will interact with matter via pair-production, forming an electron-positron pair. High energy electrons and positrons will decelerate, releasing photons via bremsstrahlung. These electromagnetic showers cause energy to be deposited in the crystals, causing scintillation which is then read by avalanche photodiodes in the barrel and vacuum photodiodes in the endcap.

The energy resolution of the ECAL is given by

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + (C)^2, \tag{3.5}$$

where E is in GeV, and was measured during several electron test beams, with momentum between 20 and 250 GeV/c. First, in 2004 a fully equipped barrel supermodule was tested in the CERN H4 beam, with the results conforming to design parameters. The energy was reconstructed by summing all 3×3 crystals. Energy resolution was found to be around 0.5% for 120 GeV electrons, and 0.45% for 120 GeV electrons after correction for containment.

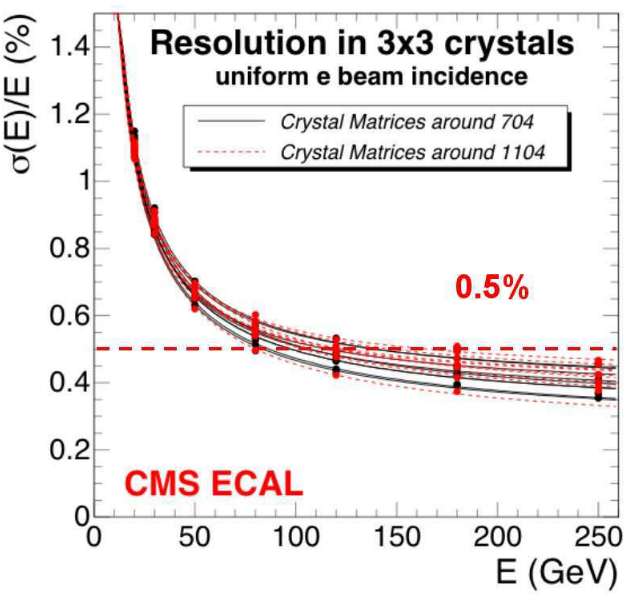


Figure 3.8: Resolution for 18 different central crystals as a function of the reconstructed energy[6]

Hadronic Calorimeter

Unlike the ECAL, the Hadron Calorimeter (HCAL) functions by entering hadrons interacting with nuclei via the strong interaction, producing several lower energy hadrons in the process, which repeats until all hadrons have been stopped. This entire process is called a

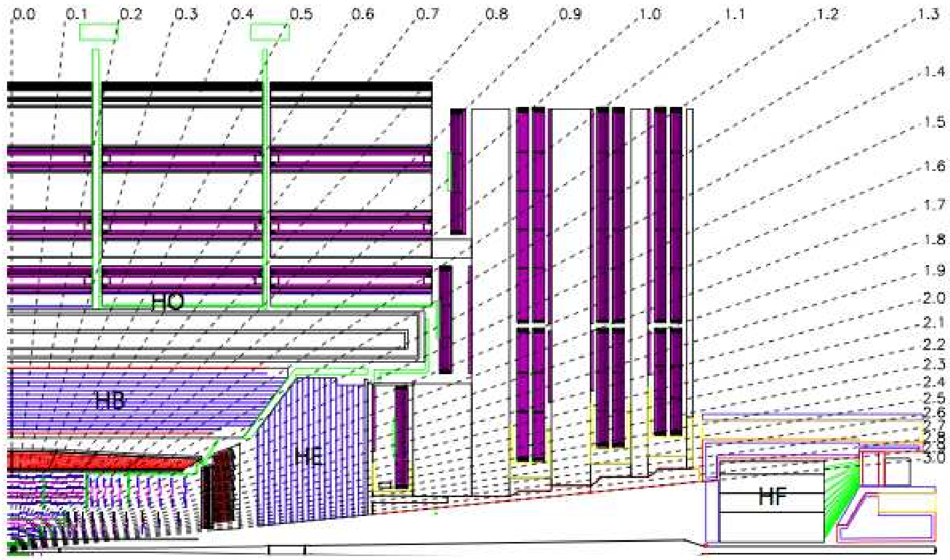


Figure 3.9: Cross-section of the CMS Hadronic Calorimeter, showing the location of the barrel, endcap, outer and forward calorimeters[5].

hadronic shower, and it is how energy is deposited into the calorimeter, although electromagnetic interactions do play a role. The design of the HCAL is severely constrained by the use of crystals in the ECAL, and by the thinness of the barrel calorimetry. As with the ECAL, the HCAL covers the central pseudorapidity range, $|\eta| < 3.0$, and consists of both barrel (which is in turn divided into two half sections, each consisting of 18 identical wedges) and endcap components. Due to the fact that both the barrel and endcap are subject to the full 4 T magnetic field, they are fashioned from non-magnetic materials, namely a copper alloy and stainless steel. The central HCAL is a sampling calorimeter, which means that it consists of active material inserted between copper absorber plates. The innermost and outermost plates are made of stainless steel for structural strength. The active elements of the entire central hadron calorimeter are 4 mm thick plastic scintillator tiles read out using wavelength-shifting plastic fibers. The barrel hadron calorimeter is about 79 cm deep [34].

To extend the reach of the HCAL TO $3 < |\eta| < 5$, two separate forward calorimeters are employed, located on either end of CMS 6 m downstream from the HE endcaps. The forward calorimeter employes quartz fibers as the active medium, embedded in a copper absorber matrix[34]. It is located in a high-radiation and high-rate environment, and is predominantly sensitive to Cerenkov light from neutral pions, leading to it having a very localized response to hadronic showers. The energy of jets is measured from the Cerenkov light signals produced as charged particles pass through the quartz fibers. These signals result principally from the electromagnetic component of showers, which results in excellent directional information for jet reconstruction. Fiber optics convey the Cerenkov signals to photomultiplier tubes, which are located in radiation shielded zones to the side and behind each calorimeter. The forward calorimeter is also part of the CMS luminosity monitor, which consists of the forward calorimeter as well as Roman Pots 300–400 m upstream.

To maximize shower energy resolution (after the crystal ECAL), the inner barrel hadron calorimeter is segmented radially (in depth) into two different sampling hadron compartments (HB1 and HB2). There is an initial layer of sampling immediately following the ECAL electronics, and 17 layers of sampling grouped together into a single tower readout. The two layers of scintillator of the Outer Calorimeter are divided into the same granularity as the barrel and envelop the entire first layer of the CMS muon iron absorber[34].

The Endcap Calorimeter (HE) is of monolithic construction (with each monolith weighing about 300 tons), consisting of staggered copper plates bolted together into 10 degree sectors. The HE outer radial perimeter is polygonal, corresponding to the 18 fold wedge structure of the barrel. The endcap hadron calorimeter is also segmented in depth into two different sampling compartments (HE1 and HE2) with 80 mm copper absorber thickness. The Endcap HCAL has two special regions, with one at high eta ($2 < |\eta| < 3$), which is a moderately high radiation area, and hence the scintillator response decreases. This necessitates a division into three readout sections (HE1, HE2 and HE3) consisting of $(1 + 4 + 14)$ sampling layers[34].

3.2.5 Muon System

As the very name suggests, muon detection and measurement is central to the concept of CMS. The muon system of CMS consists of drift tubes (DTs) in the barrel region, cathode

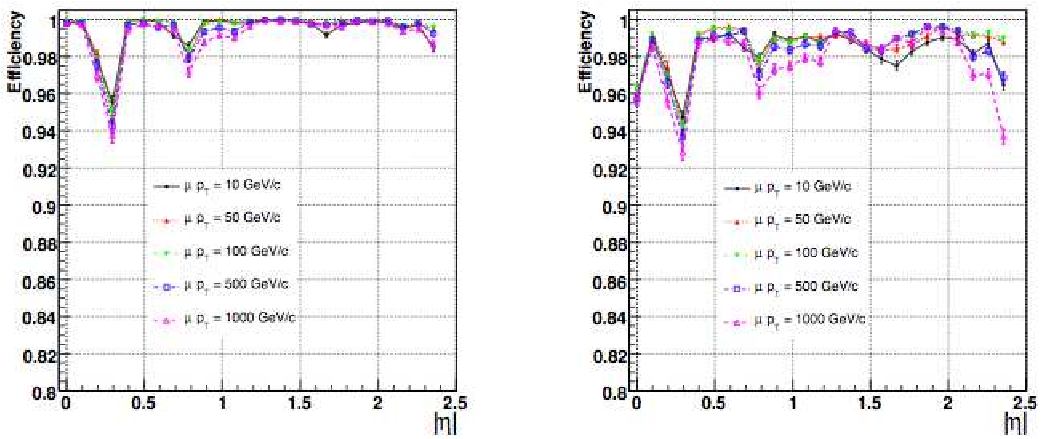


Figure 3.10: Muon reconstruction efficiency as a function of transverse momentum. On the left is standalone reconstruction using only the muon system, on the right is global reconstruction using the muon system in conjunction with the tracker[5].

strip chambers (CSCs) in the endcap region, and resistive plate chambers (RPCs) in both the barrel and endcap regions. The physics performance requirements demand coverage up to $|\eta| = 2.4$, as well as adjustable p_T coverage up to 100 GeV and time resolution of less than 25 ns. The standalone momentum resolution is from 8–15% $\delta p_T/p_T$ at 10 GeV and 20–40% at 1 TeV, and the global momentum resolution (after matching with the tracker) is 1.0–1.5% at 10 GeV and 6–17% at 1 TeV. Spatial position matching is 150–350 μm in the barrel and 75–200 μm in the endcaps[38].

Drift Tubes

Drift tubes were chosen for the barrel region due to the low expected rate per channel (corresponding to DT coverage of the region $0 < |\eta| < 1.3$), relatively low intensity of the local magnetic field, and large coverage at relatively low cost. Tubes were also used to have natural protection against broken wires, and to partially decouple contiguous cells in the presence of electromagnetic debris accompanying the muon itself[38]. The walls of the drift tubes are 2 mm thick, and are arranged in groups of three consecutive layers of tubes, staggered by half a tube, in a single rigid structure known as a drift tube chamber.

Four electrodes shape the drift field: two on the side walls, and two above and below the wires on the ground planes between layers. The incoming muons ionize the gas molecules, which causes electrons to drift towards the wires. The staggered arrangement is so that a mean-timing algorithm can be used (as the name implies, the mean of the electron drift time to the anode wires is used to determine the trajectory of the muon). This arrangement ensures performance even in the presence of stray magnetic fields, and provides 250 micron resolution per layer. The tubes are operated at atmospheric pressure with a binary Ar/CO_2 gas mixture.

In the barrel, four stations are integrated in the return yoke of the magnet, with two stations mounted on the inner and outer face of the yoke, and two located in slots inside. Each station is segmented longitudinally into five rings, each 2.5 m long, and by azimuthal ribs. Each of the three inner stations is composed of 60 chambers, and the outer station is composed of 70 chambers. The basic element is a drift cell of approximately 400 ns maximum drift, a choice which reduces the number of wires to less than 200,000. There are twelve planes of drift tubes in every chamber, and they are organized into three independent subunits called Super Layers (SLs) made up of four planes, staggered by half a cell, with parallel wires. Two of the SLs measure the ϕ coordinate, and the other ones measures the z coordinate. The ϕ SLs have a separation of 23 cm, the maximum allowed. Between them sits the z SL and a honeycomb spacer[38].

Cathode Strip Chambers

The Cathode Strip Chambers (CSCs) are comprised of six planes of anode wires interleaved between seven trapezoidal cathode panels, with a gas gap of about 1 cm. When a muon traverses a gas gap, it leaves an ionization trail of approximately 100 electron-ion pairs. The electrons drift towards the anode wires (+4000 V), and avalanche when they approach a distance of 2 radii from the wire. Drifting ions from the avalanche induce charge on the cathode strips and anode wires which are amplified by the front end electronics. Then, readout electronics decode the signals from wires and strips to measure two muon coordinates in each of the six planes. CSCs were chosen for CMS because of their capability of

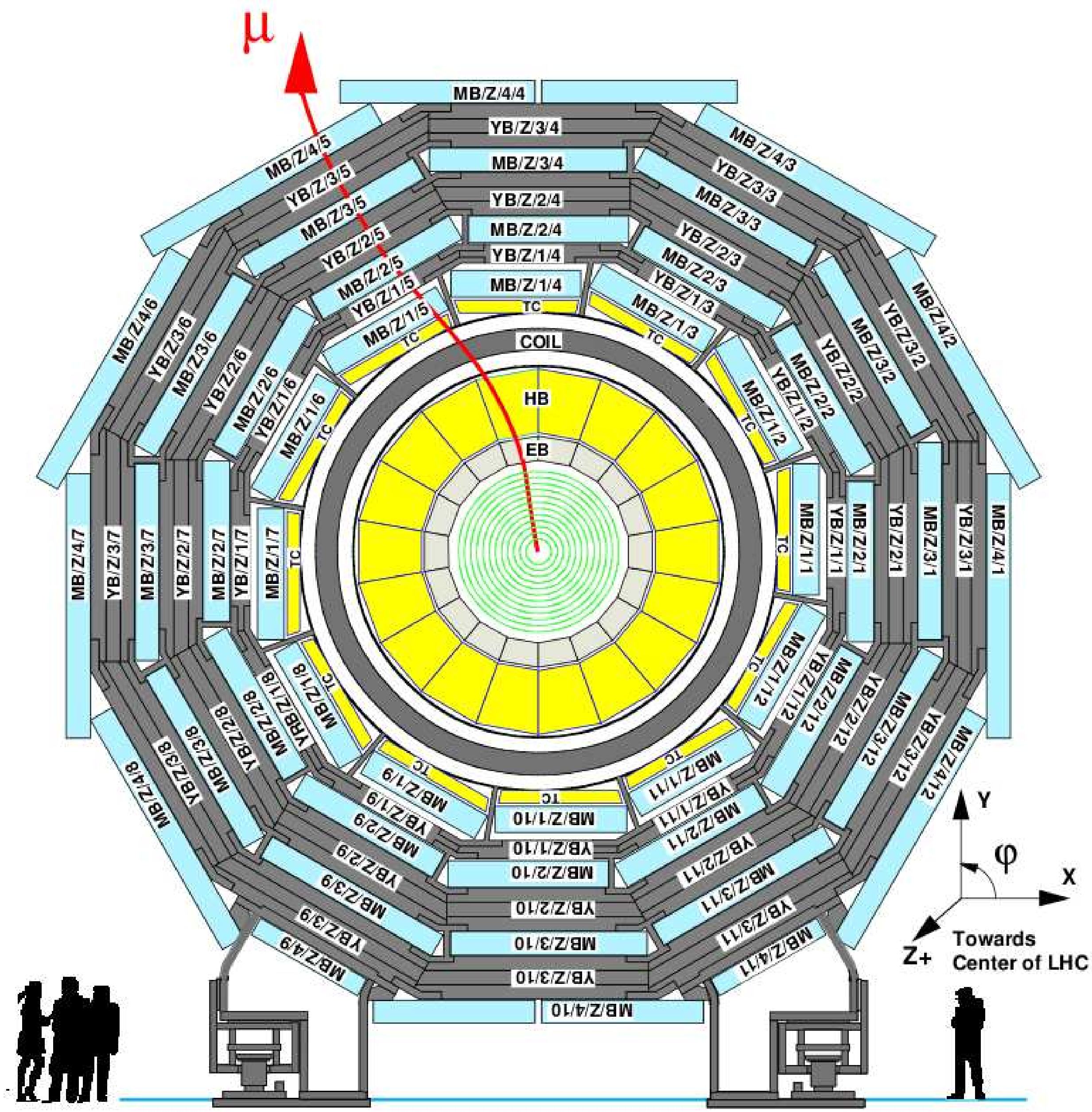


Figure 3.11: Transverse view of the layout of one of the five “wheels” showing the CMS barrel drift tube chambers. The top and bottom chambers are cut so as to simplify assembly[5].

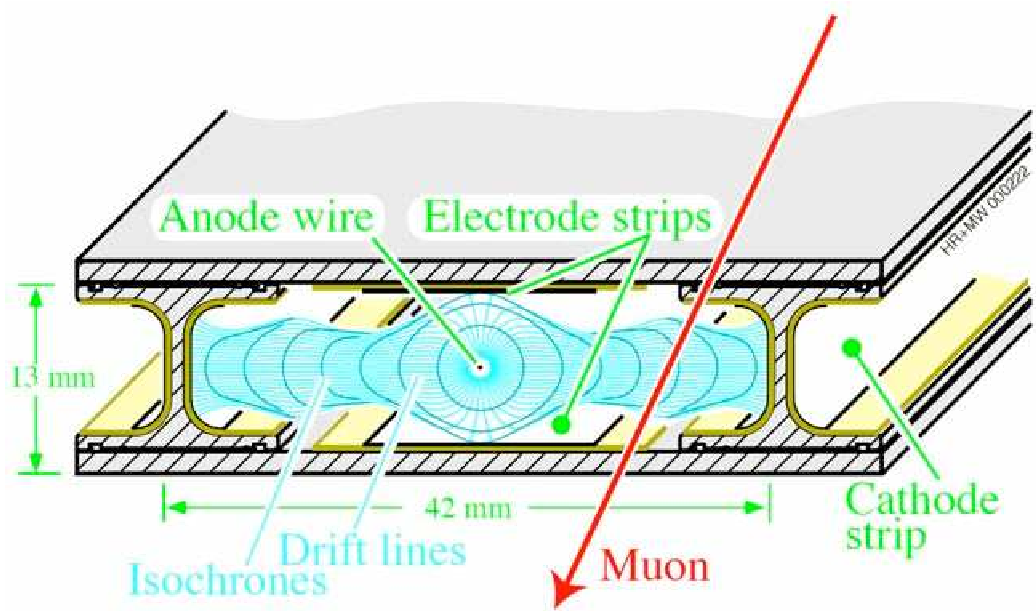


Figure 3.12: Cross-section of a drift tube cell showing drift lines. The plates at the top and bottom are at ground potential, with the voltages applied to the wires +3600V, the strips +1800V, and the cathodes -1200V[5].

providing precise time and space measurements in the presence of a high magnetic field and particle rate, and robust pattern recognition of non-muon backgrounds and efficient matching of external muon tracks to internal track segments[38]. The wider strip width and wire spacing were chosen to limit the number of channels, however the timing resolution is still sufficient to meet the physics requirements. Also, the usage of CSCs enables chambers to be installed radially around the disk structure of the endcaps.

Each of the two endcap regions (denoted plus and minus) has four muon stations, denoted ME1 through ME4, with trapezoidal chambers which are concentrically arranged around the beam line. The stations are separated by iron disks of the flux return yoke, denoted YE1, YE2 and YE3. Both YE1 and YE2 are 600 mm thick, and YE3 is 250 mm thick, and was put in place in order to interact with the forward calorimeter, beam pipe, quadrupole magnets, etc. Thus, ME4 chambers are mounted on the beam side of ME3. ME1 has three rings of chambers, ME2 and ME3 have two rings each, and ME4 has one

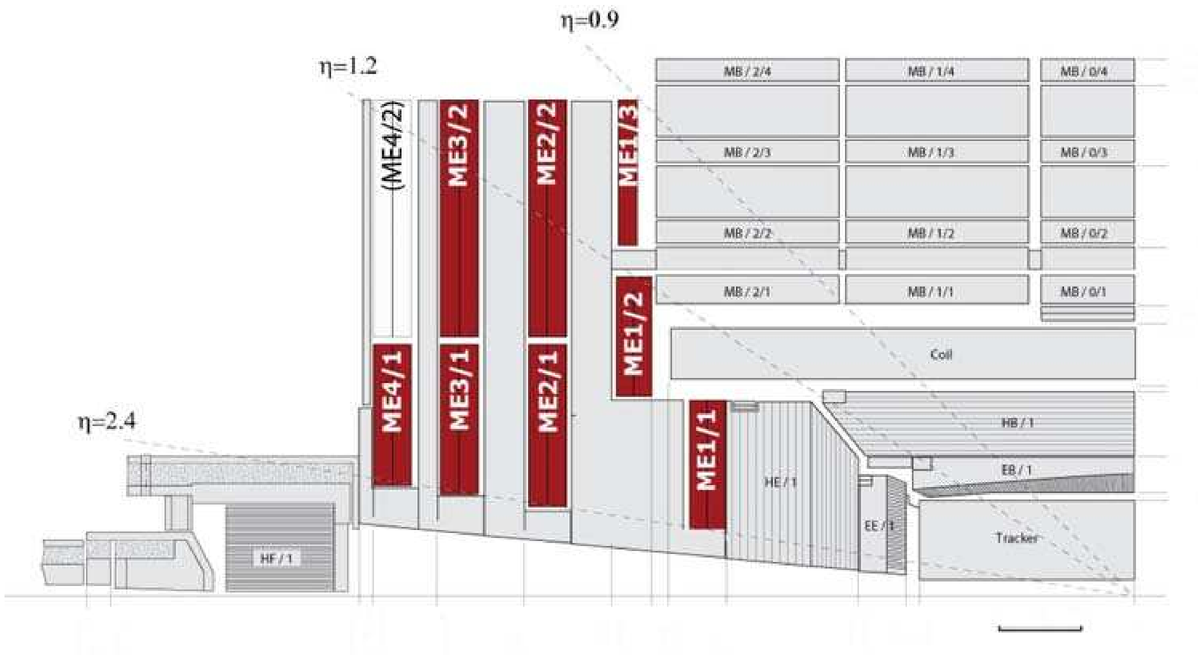


Figure 3.13: Quarter view of the CMS detector, with the CSCs highlighted[5].

ring, with 5 chambers added to ME+4/2 as of May 2010. All of the stations, with the exception of ME1/3, overlap in ϕ for full coverage, and in each of rings 2–4 there are 36 chambers each covering 10° in ϕ at the outer radius, and 18 chambers covering 20° at the inner radius, yielding close to 100% coverage down to $\theta = 10^\circ$ or $\eta = 2.4$. ME1/1 chambers operate in a magnetic field of strength of greater than 3 T, whereas ME1/2 chambers are in a weaker (1 T), non-uniform field, and the other chambers are in much lower fields. Thus, muon measurement is necessarily required to be much more accurate in the first station than in successive ones. Another issue with the iron disks is multiple scattering, which is corrected for by comparing muon hits in the endcaps with extrapolated tracks from the tracker and calorimetry system.

Each CSC consists of six layers of wires between cathode panels, with each cathode panel consisting of six planes of strips running radially, providing six measurements of ϕ (strips) and six measurements of r (wires). The strips are anywhere from 3 to 16 mm wide,

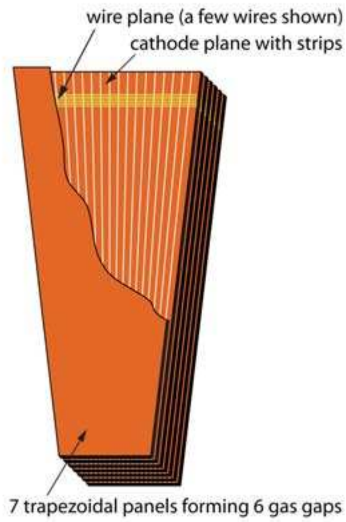


Figure 3.14: Layout of a cathode strip chamber, with the top panel cut away to show the anode wires and cathode strips. Note that only some wires are shown for clarity[5].

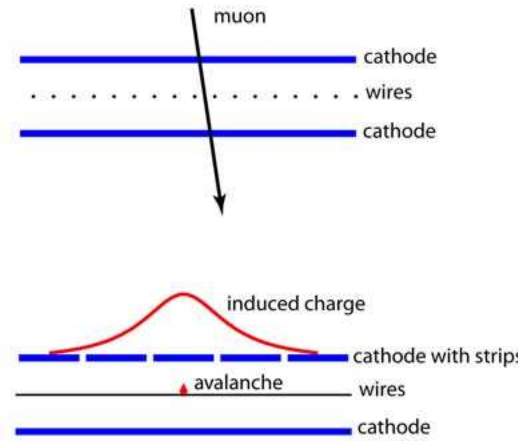


Figure 3.15: A schematic view of CSC operation. Interpolation of charges induced on cathode strips by the avalanched positive ions near the wire allows one to precisely measure the muon trajectory[5].

corresponding to an angle of 2 to 5 mrad in ϕ, and measurement of track coordinates is best suited to measure muon momentum. Overall, the endcap muon system consists of 540 chambers, with about 2.5 million wires, 210,816 anode channels and 273,024 cathode channels, with a typical chamber possessing about 1,000 readout channels[38]. The per-layer spatial resolution is between 150 μm and 700 μm, and the per-CSC spatial resolution is around 75 μ. Timing studies undertaken showed the peaking time resolution to be 4 ns, which is comparable to what was measured during the 2003 test beam.

Resistive Plate Chambers

Resistive Plate Chambers (RPCs) are gaseous parallel-plate chambers. They are characterized by reasonable spatial resolution and a fast time response, comparable to scintillators, and hence can provide a very accurate assignment of the bunch crossing, and can do so in a cost-effective manner due to the fact that they do not require an expensive readout device. Thus, RPCs are a fast dedicated trigger for identifying candidate muon tracks and

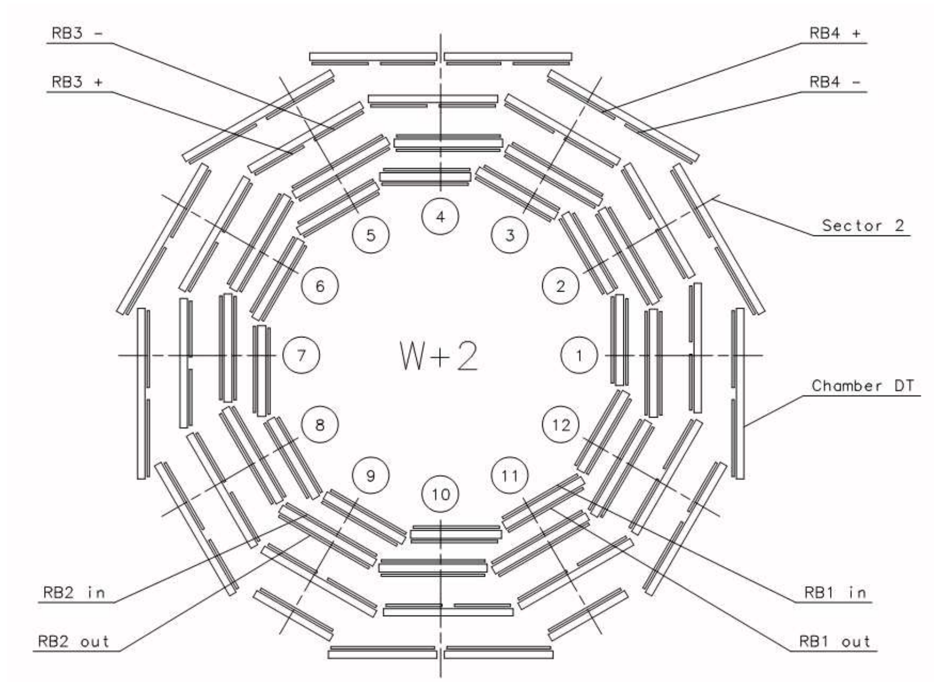


Figure 3.16: Schematic layout of one of the five barrel wheels, with RPC placement shown[5].

assigning the bunch crossing.

RPCs can be found in both the barrel and endcap regions. Each plate is made of phenolic resin, with good surface flatness and high bulk resistivity. The resin is covered with a conductive graphite paint to form electrodes, and readout is performed by aluminium strips located on top of (and insulated from) the graphite paint. The separation of the plates is on the order of a few millimeters, and in normal construction two such assemblies are placed back to back, with the readout strips in the center and the entire assembly being gas-tight. The chambers are run in so-called avalanche mode, which has lower gas amplification and smaller pulses than streamer mode, which is inadequate for a high-rate environment like the LHC but has the advantage that it does not require an amplification stage.

Six layers of RPCs in total are mounted in the barrel chambers, with two layers in each of MB1 and MB2 and one each in the outer stations. In the endcap region, each of the four layers of CSCs will have a layer of RPCs in conjunction, with shape and mounting

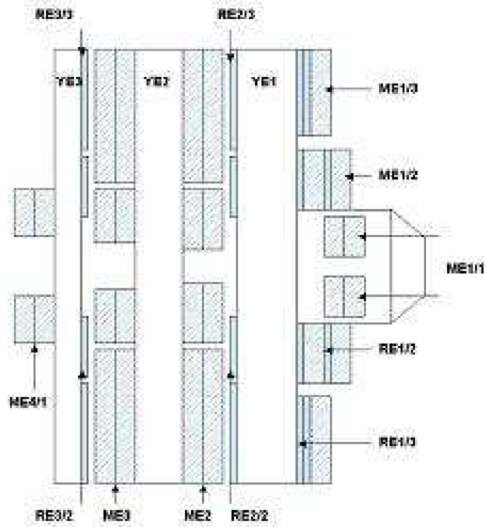


Figure 3.17: Longitudinal cutaway of the endcap muon system, showing location of RPCs and CSCs[5].

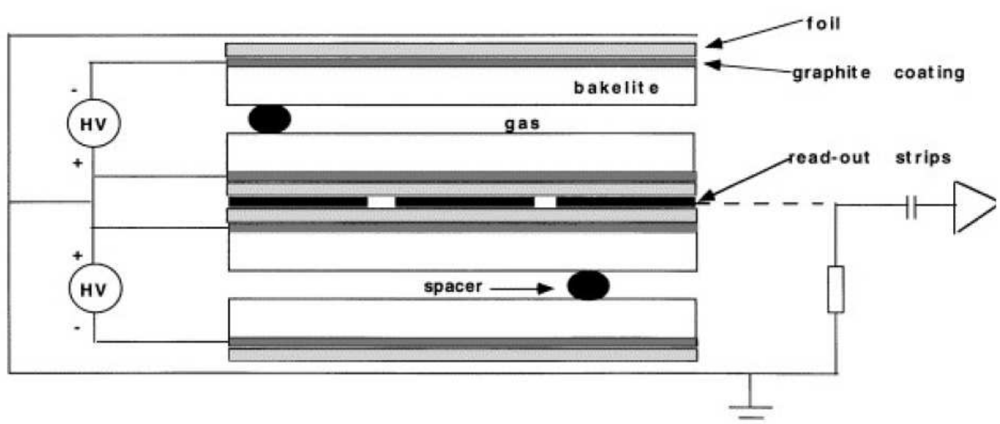


Figure 3.18: Schematic view of the RPC double-gap structure. The readout strips are situated parallel to the beam direction[38].

determined by η segmentation. Coverage up to $|\eta| = 2.1$ is provided[38].

3.2.6 Trigger and Data Acquisition

At the LHC, the proton beams cross each other at a rate of 40 MHz, with roughly 20 inelastic pp collisions per crossing. Thus, the number of events must be pared down drastically so as not to overwhelm the hardware and software resources. To that end, a two-stage trigger is implemented. The first stage is referred to as the Level-1 Trigger (hereafter L1 trigger), which is primarily comprised of field-programmable custom electronics. The second stage, or High-Level Trigger (HLT), is a software trigger which is very similar in design and operation to the offline analysis code. The L1 trigger reduces the initial collision rate to less than $\mathcal{O}(10^6)$ Hz, and the HLT is designed to reduce this further to approximately 100 Hz[39].

The L1 trigger hardware is implemented in field-programmable gate arrays (FPGAs), as well as application-specific integrated circuits (ASICs) and programmable memory lookup tables (LUTs). The L1 trigger is comprised of local, regional and global components. Local triggers (also known as trigger primitive generators, or TPGs) are based on energy deposits in calorimeter trigger towers and track segments or hit patterns in muon chambers. Regional triggers combine their information and rank and sort trigger objects and transfer them to the global trigger, which makes the decision to reject an event or accept it and pass it along to the Data Acquisition (DAQ) system with a latency of 3.2 μs, where it is further filtered by the HLT.

Calorimeter Trigger

The calorimeters are divided into "towers" for triggering purposes, with each tower in the region $|\eta| \leq 1.74$ having an (η, ϕ) coverage of 0.087×0.087, and higher in the region $|\eta| > 1.74$. The TPGs sum the transverse energies measured in ECAL crystals or HCAL readout towers to obtain the trigger tower E_T and attach the correct bunch crossing number, after which the TPGs are transmitted to the Regional Calorimeter Trigger (RCT). The RCT determines electron/photon candidates and E_T sums per calorimeter region, as well as information relevant for muons pertaining to isolation. A region consists of 4×4 trigger

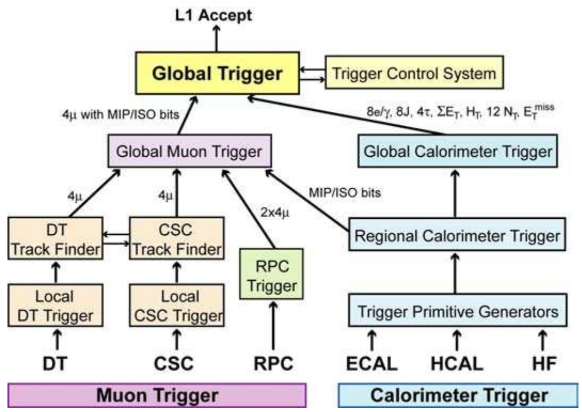


Figure 3.19: Architecture of the Level-1 Trigger[5].

towers except in HF, where a region is 1 trigger tower. The e/γ trigger algorithm passes 4 isolated and 4 non-isolated e/γ candidates per region to the Global Calorimeter Trigger (GCT). Furthermore, the RCT also sums E_T in a given region and determines τ-veto bits for identifying τ-decays (τ-jets are narrower than ordinary quark/gluon jets.) The GCT determines jets, the total E_T, the $\not{E}_T$, and the scalar transverse energy sum of all jets above a programmable threshold (H_T). Finally, up to four jets and four τ jets from the central HCAL and four jets from HF are forwarded to the GT after sorting[5].

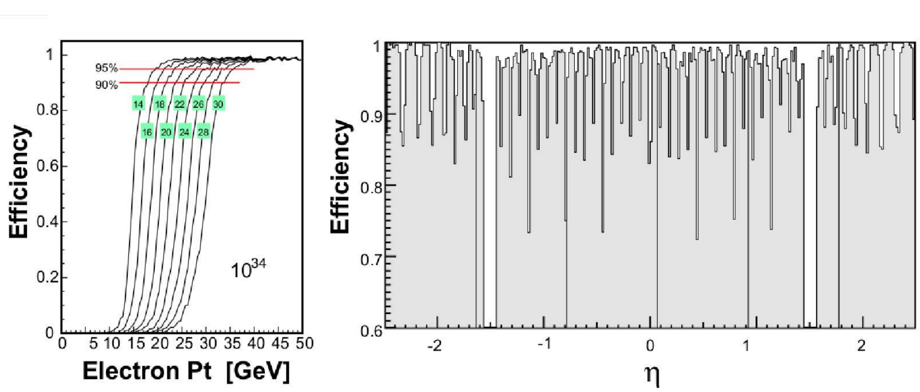


Figure 3.20: Efficiency of the Level-1 Trigger for single electrons as a function of transverse momentum. On the right is electron efficiency as a function of η for electrons with $p_T > 35$ GeV[39].

Muon Trigger

All three muon subsystems take part in the trigger. The DT local trigger information is provided by the barrel chambers, in the form of track segments in the ϕ-projection and hit patterns in the η-projection, and the CSC local trigger information is provided by three-dimensional track segments. The trigger information from the DTs, namely the position, transverse momentum and track quality, is encoded and transmitted to the DT Track Finder (DTTF). The best two Local Charged Tracks (three-dimensional muon tracks from the CSCs, consisting of ϕ, the bending angle ϕ_b, η and the bunch crossing number) are sent to the CSC Track Finder (CSCTF).

The Regional Muon Trigger is where the DTTF and CSCTF join segments to complete tracks as well as their associated physical parameters. The Track Finders serve the purpose of identifying muons, determining their transverse momenta, precise locations, and quality, which in turn is passed on to the Global Muon Trigger (up to four muons each, sorted by p_T and quality). Furthermore, the CSCTF and DTTF compare data to determine the properties of muons passing between the barrel and endcap regions. The RPCs (noted for their excellent timing resolution of approximately 1 ns, useful for ensuring exact bunch crossing identification) use a Pattern Comparator Trigger to assign p_T and electric charge, and delivers the four best track candidates in both barrel and endcap regions based on regional hit patterns to the Global Muon Trigger (GMT).

Finally, the GMT combines information from the DTs, CSCs and RPCs. Its purpose is to improve trigger efficiency, reduce rates and suppress background by combining information from all three systems. For every bunch crossing, the GMT receives up to four muons each for the DTs, barrel RPCs, CSCs, and endcap RPCs (containing p_T, charge, η, ϕ and quality), as well as isolation and minimally ionizing particle bits from the GCT. After filtering out possible duplicates reported by the DT and SC triggers, the muons are then sorted by p_T and quality, and passed to the Global Trigger (GT).

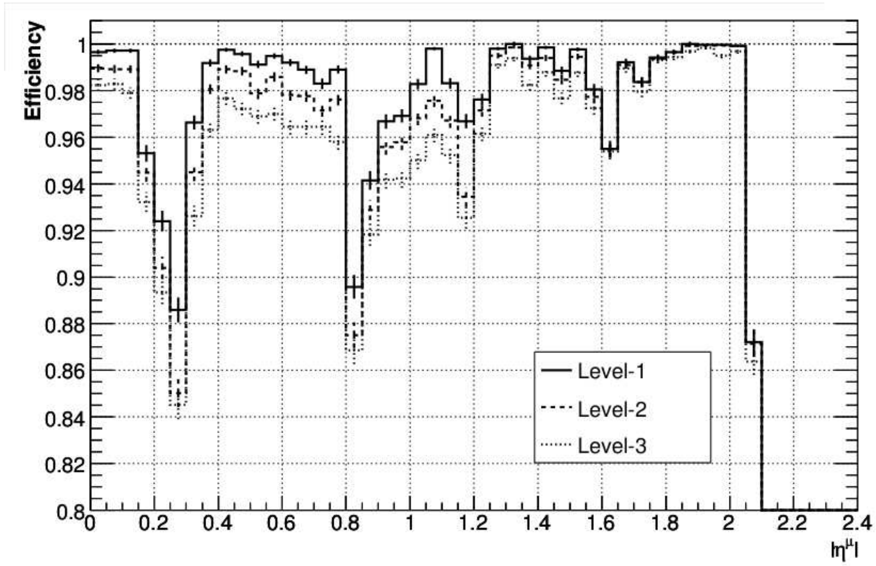


Figure 3.21: Overall combined muon efficiency of the Level-1, Level-2 and Level-3 Trigger as a function of generated pseudo-rapidity.[3].

Global Trigger

The GT has five basic stages: input, logic, decision, distribution, and read-out. It takes input from the GCT and GMT, receiving trigger objects such as e/γ (both isolated and non-isolated), muons, central and forward hadronic jets, τ jets, E_T, $\not{E}_T$, H_T, and jet multiplicities. From there, objects are ranked and sorted (based on p_T or E_T, (η, ϕ)-coordinates, and quality), and the decision is made whether or not to accept or reject an event. If the event is accepted, it is passed along to the Data Acquisition system.

Data Acquisition and High-Level Trigger

The CMS DAQ is a collection of hardware and software components designed to take accepted events from the L1 Trigger (L1As), perform further filtering via the HLT, and output events accepted by the HLT to disk for offline processing and analysis. The input rate is on the order of 100 kHz, which corresponds to a data flow of $\approx$ 100 GBytes/s, which is reduced by the HLT by a factor of 1000[5]. The various subdetector Front-End Systems store data continuously in 40 MHz pipelined buffers. Upon receipt of a L1 trigger via the Trigger, Timing and Control (TTC) system, the data are extracted by the 626 Front-End Drivers (FEDs) and pushed into the DAQ via the 512 Front-end Read-out Links (FRLs) which are capable of merging data between two FEDs.

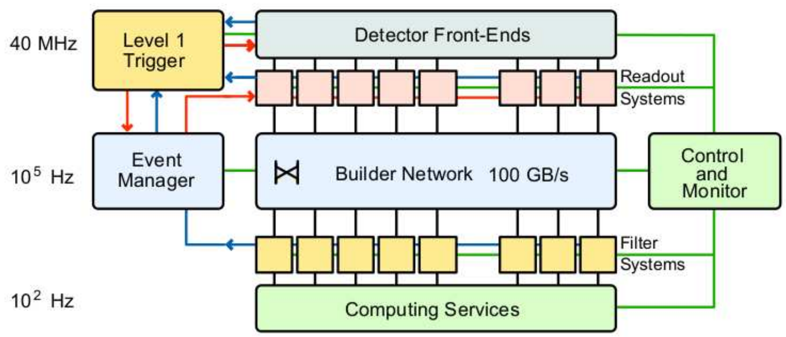


Figure 3.22: Architecture of the CMS Data Acquisition system.

The event builder collects event data for the same L1 from all FEDs and assembles them into a complete event, transferring the event to the Event Filter for further processing. The Trigger-Throttling System (TTS) protects against buffer overflows due to back-pressure from downstream event processing. In the event of back-pressure occurring, any FED can provide feedback to the TTS to throttle the trigger and halt L1As until the buffers are clear. During collisions, trigger thresholds and pre-scales will be optimized; however, instantaneous fluctuations might lead to L1 trigger throttling. CMS defines a *luminosity section*, consisting of 2^{20} LHC orbits or 93 seconds, during which trigger thresholds and pre-scales are not changed.

HLT reconstruction and selection is performed using the same software framework used for offline analysis and reconstruction. The event filter farm consists of 720 commodity server PCs. The HLT takes the place of the traditional Level-2 and Level-3 triggers, and the data given to the HLT are the full raw data contained in the front-end electronics.

Various software filters, known as HLT paths, are applied to incoming events. Events accepted by one of the HLT paths are broken down into reconstructed physics objects in the ROOT data format and written to the Tier-0 data storage site at CERN[40].

3.2.7 Offline Computing

For offline processing and analysis, data from CMS are made available in several formats. Of these, RAW is the largest and contains the full recorded information of the detector, trigger decision, and other data. RECO is created from RAW data by applying cluster and track finding, vertex reconstruction and compression algorithms. Finally, AOD (Analysis Object Data) is more compact, primarily containing the parameters of high-level physics objects.

There are several levels of computing centers, following a tiered hierarchy. Of these, the Tier-0 center (of which there is only one) is hosted at CERN, and its primary tasks are to:

- Accept data from the online system and copy them to permanent tape storage;
- Reconstruct RAW data into RECO datasets at a rate comparable to the average rate of data recording;
- Export RAW and RECO datasets to Tier-1 facilities.

Tier-1 centers are hosted at national laboratories and computing sites around the world (currently 11 sites). Each one is expected to provide high uptime, and is maintained around the clock by a team of experts. Each Tier-1 site holds unique RAW and RECO datasets, and a complete copy of the AOD data. The primary functions of the Tier-1 are:

- Provide long-term storage of RAW data from CMS for redundancy outside CERN;
- Transfer stored data to any Tier-2 center for analysis;
- Carry out second-pass reconstruction from RAW data using improved algorithms;
- Provide quick access to data samples for skimming and intensive analysis.

Tier-2 centers (currently over 160 in number) are typically found at CMS institutes and universities. They divide their resources between the local userbase and CMS as a whole, and require less stringent uptime and reliability than a Tier-1. Their primary functions are:

- Support local analysis activities, including local data storage;
- Support of tasks such as offline calibration, alignment, and detector studies;
- Production of Monte Carlo datasets for local use and transfer to Tier-1 facilities.

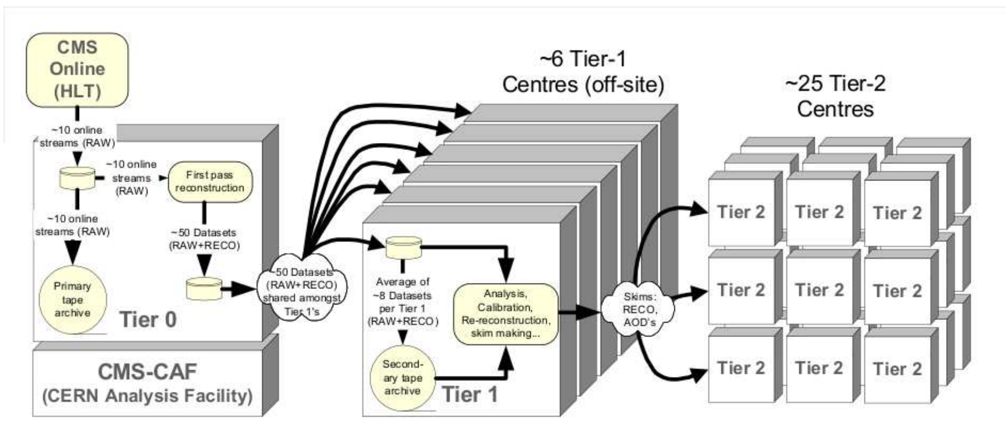


Figure 3.23: Dataflow of CMS offline computing centers.

Tier-3 centers perform a similar function to Tier-2 centers, with no constraints on uptime or available system resources. Finally, CERN also hosts on-site a CERN Analysis Facility, which provides flexible CPU resources as well as quick access to the entire CMS dataset. It effectively combines the flexibility of a Tier-2 with the rapid data access of a Tier-1[3].

3.3 Luminosity

One of the most crucial components of any measurement made at CMS (or any LHC experiment, for that matter) is determining luminosity, both integrated over time and instantaneous. There are two methods in place at CMS to determine relative luminosity, both using the forward hadronic (HF) calorimeters to determine the instantaneous luminosity in real time (information which is also logged by a dedicated DAQ system for further offline analysis). For determining an absolute calibration of luminosity, the physical properties of the beams themselves were measured, a procedure which does not require HF counting. Other methods of determining luminosity, such as using a golden sample in data of a well-known process (such as Z$\rightarrow \mu^+\mu^-$) and comparing to the known cross section, were not feasible at CMS (due to low statistics and a need to validate such processes at 7 TeV collisions before use as a "standard candle") during the 2010 data-taking era, and so are excluded from consideration as a method of luminosity measurement.

HF Calorimeters

Two methods are employed using dedicated hardware in the CMS online luminosity measurement system for online measurement of instantaneous luminosity. The first uses a method called "zero counting", in which the average fraction of empty HF towers is used to estimate the mean number of interactions per bunch crossing. If the mean number of interactions per bunch crossing μ is sufficiently small ($\mu << 1$), then measuring luminosity is straightforward: since the probability of two events is of order $\mathcal{O}(\mu^2)$, simple hit counting is sufficient. If, however, $\mu \approx 1$, then "zero counting" must be used. The probability of n interactions given μ is distributed according to the Poisson formula

$$p(n;\mu) = \frac{\mu^n e^{-\mu}}{n!}, \tag{3.6}$$

which in turn implies $p(0;\mu) = e^{-\mu}$, or $\mu = -\log[p(0)]$.

The second method uses the linear relationship between the average transverse energy per tower and luminosity. Although the forward hadronic calorimeters cover the pseudorapidity range $3 < |\eta| < 5$, the requirement of linearity necessitates the limiting of coverage to four azimuthal rings in the pseudorapidity range $3.5 < |\eta| < 4.2$; else the average fraction of empty towers becomes a sum over exponentials, and hence non-linear.

As a cross-check of online luminosity measurement, two offline algorithms are used. The first is based on energy deposition in the HF, while the second uses tracker and vertex finding. While the offline methods have longer latency (typically on the order of 24 hours), they allow for better background rejection, employ a different data-handling path, and involve a completely different set of systematic uncertainties (in the case of vertex counting). The offline HF method sums E_T depositions over all towers, and uses the coincidence of depositions of at least 1 GeV in the forward and backward HF arrays. Timing cuts are imposed, as a means of further eliminating non-collision backgrounds. The second offline method uses vertex counting, requiring that at least one vertex with at least two associated tracks is present in each event. Furthermore, the z-position of the vertex is required to lie within 150 mm of the IP.

Beam Measurement

Given two Gaussian beams, the instantaneous luminosity is given by

$$\mathcal{L}_0 = \frac{N_1 N_2 f N_b}{2\pi\sqrt{\left(\sigma_{1x}^2 + \sigma_{2x}^2\right)\left(\sigma_{1y}^2 + \sigma_{2y}^2\right)}}, \tag{3.7}$$

where N_1 and N_2 are the bunch intensities, f is the revolution frequency, N_b is the number of bunches per beam and $\sigma_x = \sqrt{\sigma_{1x}^2 + \sigma_{2x}^2}$ and $\sigma_y = \sqrt{\sigma_{1x}^2 + \sigma_{2x}^2}$ are the effective beam sizes in the x and y planes, respectively. Then, the size and shape of the interaction region is measured by recording the relative interaction rate as a function of the transverse beam separations. The Van der Meer scan, first developed by Simon Van der Meer for use at the ISR[41], is used to measure the transverse size of the beams. One beam is held fixed and the central transverse position of the other is shifted slowly relative to the other, then the process is repeated, holding the second beam fixed. As events are recorded by CMS while the beams are being moved, the width of the beams is determined by the event rate as a function of beam offset.

Systematic Errors and Result

As the beam measurement method provides an absolute calibration of overall luminosity, its results were used in determining the integrated luminosity of the 2010 data run, as were

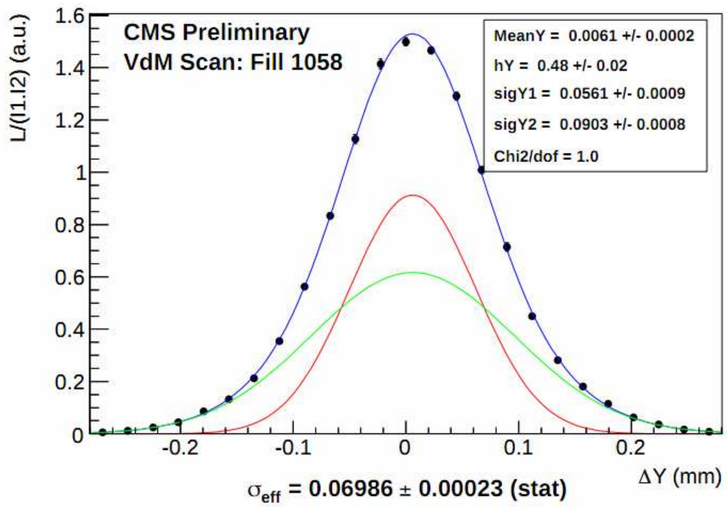


Figure 3.24: Sample van der Meer scan result in y, from LHC fill 1058. The blue curve is the total double Gaussian, the red curve is the core Gaussian (σ_1), and the green curve is the Gaussian for the tails (σ_2).

its systematic uncertainties.

The systematic uncertainty in luminosity is dominated by inaccuracy in measuring the beam currents, which have an RMS measurement error of 5% per beam, or 10% total (using the conservative estimate that the beam intensity measurements are completely correlated, and hence the errors are added linearly). Another source of systematic error is uncertainty in the beam shape, due to the possibility that a double gaussian is not a perfect description of the actual beam shape. Thus, a systematic uncertainty of $\pm$ 3% was derived from replacing the double-gaussian fit with a spline fit.

Another source of systematic error is the fit systematics term, in which variations of the fit parameters are observed in offline distribution fits. Further systematics include scale-calibration errors associated with the methods used to determine beam offsets, a zero-point uncertainty associated with variations in the beam size during the scans, and beam background, such as beam halo. The results are summarized in Table 3.2, with an overall systematic error of 11%. Thus, the measured luminosity at CMS during the 2010 runs was $35.9 \pm 3.9\text{pb}^{-1}$.

Error	Value %
Beam Background	0.1
Fit Systematics	1.0
Beam Shape	3.0
Scale Calibration	2.0
Zero Point Uncertainty	2.0
Beam Current Measurement	10.0
Total	11.0

Table 3.2: Summary of sources of systematic error contributing to uncertainty in luminosity measurement at CMS.

Chapter 4
Analysis

In this analysis, the quantities which will be measured are the overall W cross section and the charge asymmetry of the W boson in the muonic decay channel as a function of muon pseudorapidity, with a further investigation into asymmetry with associated jet multiplicity. The major motivation for measuring W boson asymmetry is to glean a better understanding of the internal structure of the proton, and to measure so-called parton distribution functions at heretofore unseen energy scales. A similar measurement made at the Tevatron was not in excellent agreement with current PDF results, so a measurement made at CMS is an important step to reconciling theoretical predictions and current experimental results. Charge asymmetry of the W boson is expected to rise at higher rapidity $|y|$ due to the selection preference of partons with higher Bjorken scaling variable x. However, due to the presence of a neutrino in the final decay product, rapidity is not directly accessible for the W boson. Therefore, a measurement was made as a function of the pseudorapidity y.

The idea behind measuring charge asymmetry with associated jet production is to provide a clean test of perturbative QCD, as well as to investigate modelling and reconstruction of jets at CMS. Experiments at the Tevatron have indicated that linear scaling of the production cross section with the number of jets $\left(\frac{Njets+1}{Njets} \propto \alpha_s\right)$ is expected to hold only at leading order. Next-to-leading order (NLO) perturbative calculations are already available for $N_{jets} \leq 3$, with improvements expected due to the significantly higher cross-sections available at 7 TeV. A further motivation for studying properties of W boson production in association with jets, such as charge asymmetry, is that at the LHC, the high luminosity and center-of-mass energy means that vector bosons (and associated jets) will be produced in abundance. As such, they form a significant (and irreducible) background for new-physics processes, such as Higgs or exotica searches. Therefore, a good understanding of the properties of W+jets processes is essential for new-physics searches, as well as the scaling into higher jet-multiplicity regimes.

As a starting point, the relevant quantities for this analysis will be defined. Next, an overview of detector acceptance and the identification and reconstruction of detector signals as physics objects relevant to this analysis, namely muons, neutrinos and jets will be given. Then, an overview of the Monte Carlo method for numerically approximating solutions to definite integrals will be provided, as well as the generators used for using Monte Carlo to generate signal and background computer-simulated data. Finally, an overview of the cuts performed on the data as well as the relative efficiency of each successive cut will be presented, as well as a description of the background sources encountered during the extraction of the W$\rightarrow \mu\nu_\mu$ signal.

An event display for a W$\rightarrow \mu\nu_\mu$ candidate event is shown in Figure 4.1.

4.1 Definitions

4.1.1 Asymmetry

The muon charge asymmetry for the W boson is defined to be

$$A = \frac{\sigma\left(W^+ \rightarrow \mu^+\nu_\mu\right) - \sigma\left(W^- \rightarrow \mu^-\nu_\mu\right)}{\sigma\left(W^+ \rightarrow \mu^+\nu_\mu\right) + \sigma\left(W^- \rightarrow \mu^-\nu_\mu\right)}. \tag{4.1}$$

Theory predictions from existing PDF sets predict an average asymmetry of approximately 0.2, due to the fact that the proton carries a net positive charge. It has been previously estimated at CMS in the W$\rightarrow \mu\nu_\mu$ channel by using Monte Carlo. The charge asymmetry of cosmic muons penetrating the detector was measured in 2008.

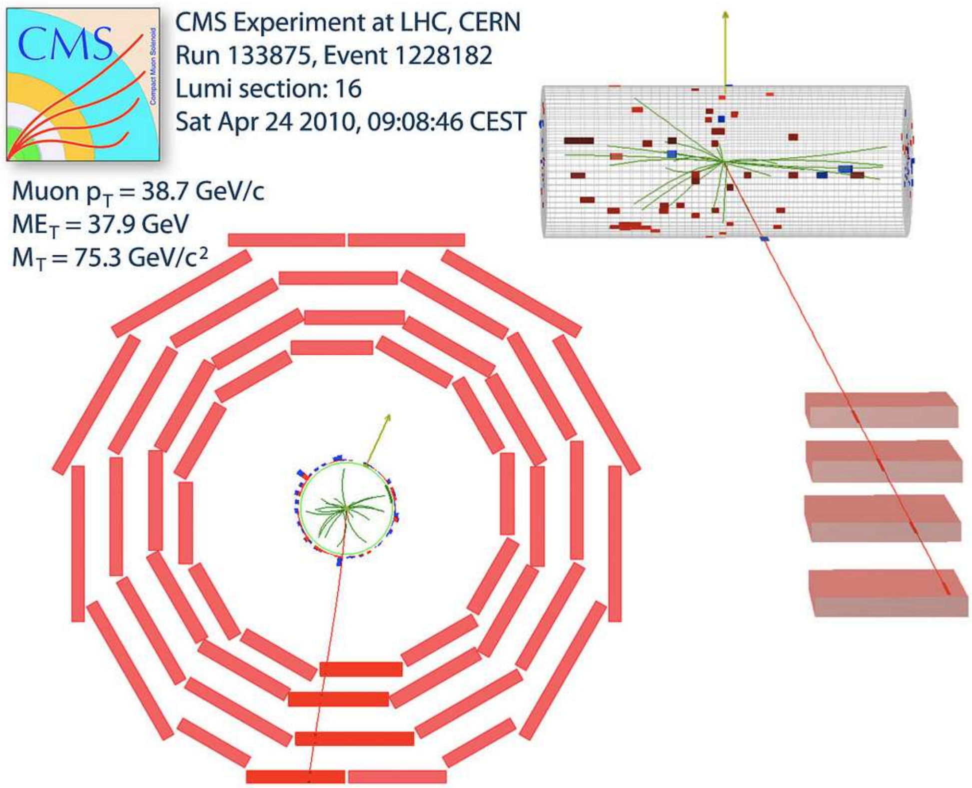


Figure 4.1: Event display for a W→ $\mu\nu_\mu$ candidate event. The figure on the left shows a cutaway transverse view of the barrel of CMS. The concentric circles comprised of red rectangles along the outside represent the barrel muon chambers, and the green circle in the middle represents the tracker. The blue and red blocks along the outside of the green circle represent the energy deposited in the calorimeter, with the size of each block corresponding to the amount of energy deposited. Red represents electromagnetic energy, and blue represents hadronic. The green curves contained within the tracker represent reconstructed tracks, the red dashed line travelling from the green circle through the muon chambers represents the reconstructed muon, and the yellow dashed line represents the missing transverse energy, or neutrino. The image in the upper right corner is from the line of sight of the reconstructed muon, and represents a longitudinal perspective. The muon p_T of 38.7 GeV/c was measured by the bending of the tracks in the muon chambers and tracker due to the magnetic field.

4.1.2 (Missing) Transverse Energy

If the total energy of a particle is given by E, then $E_T = E\cos\theta$ is the transverse energy of the particle, or the component of the energy which is in the x-y plane perpendicular to the beam line. At hadron colliders, a significant proportion of the energy of the incoming hadrons escapes down the beam pipe, and is impossible to measure. Therefore, for particles which are likely to escape the detector without being detected (namely neutrinos and as-yet-undetected Weakly Interacting Massive Particles), the only possible constraint is on energy in the transverse plane. Then the net momentum, or missing transverse energy, is given by

$$\not{E}_T = -\sum_i p_T\left(i\right), \tag{4.2}$$

where the sum is over the transverse momenta of all visible final-state particles.

4.1.3 Invariant and Transverse Mass

The invariant mass of two particles is defined as

$$m^2\left(1,2\right) = \left(|p_1| + |p_2|\right)^2 - \left(p_1 + p_2\right)^2 = \left(E_1 + E_2\right)^2 - |p_1 + p_2|^2, \tag{4.3}$$

and is distinguished by the fact that it is the same in all frames of reference, and is equal to the mass in the rest frame. It is analogous with a quantity known as transverse mass:

$$m_T^2\left(1,2\right) = \left(|p_{T,1}| + |p_{T,2}|\right)^2 - \left(p_{T,1} + p_{T,2}\right)^2 = \left(E_{T,1} + E_{T,2}\right)^2 - |p_{T,1} + p_{T,2}|^2. \tag{4.4}$$

Transverse mass can also be defined in terms of $\not{E}_T$:

$$m_T = \sqrt{2p_T \not{E}_T\left[1 - \cos(\Delta\phi)\right]}, \tag{4.5}$$

where $\Delta\phi$ is the angle in the x-y plane between the lepton and the $\not{E}_T$ direction. Transverse mass is one of the distinguishing variables in the W$\rightarrow \mu\nu_\mu$ decay, and hence a fit to m_T is a common method of extracting signal yields.

4.1.4 Scale and Resolution

Biases in the detector might lead to inaccuracies in measuring quantities such as muon momentum. Momentum scale is defined as p_{true}/p, and is corrected for by measuring known dimuon resonances such as J/Ψ, the Z boson, and Υ, whose masses are well-known due to previous experiments. Momentum resolution is defined as $\sigma\left(p\right)/p$, where $\sigma(p)$ is the width of a di-muon invariant mass peak from a known resonance.

4.1.5 Efficiency

Given a criterion of detector performance such that N_{pass} is the number of events passing the threshold of acceptable performance, and N_{fail} is the number of events failing the threshold, the efficiency is defined as

$$\epsilon = \frac{N_{pass} - N_{fail}}{N_{pass}}. \tag{4.6}$$

Efficiencies are an important quantity in terms of determining the contribution of detector effects to the overall systematic error.

4.2 Monte Carlo

Monte Carlo is an invaluable tool in high energy physics for simulating experimental data using known parameters. By simulating physics events and running the output through a detector simulator, one can vary different parameters and conditions to understand such systematical quantities as resolution and efficiencies. Furthermore, by parlaying already-known properties of various processes, sources of background contamination can be modelled and hence removed from the signal source.

4.2.1 Overview of Monte Carlo

At its core, Monte Carlo is a method for computing the value of a definite integral. First, recall the fundamental theorem of calculus:

$$I = \int_a^b f(x)dx = F(b) - F(a), \tag{4.7}$$

where $F(x)$ is the antiderivative of $f(x)$. Then, the Mean Value Theorem states:

Mean Value Theorem for Integrals 4.2.1 *Given a function $f(x)$ which is continuous on the interval $[a, b]$ and differentiable on the interval (a, b), then there exists a point c in (a, b) such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}. \tag{4.8}$$

Consider N randomly-chosen points $x_i, i = 1, \ldots, N$ in the closed interval $[a, b]$. Then

$$I \approx I_N \equiv (b - a) \times \frac{1}{N} \sum_{i=1}^{N} f(x_i). \tag{4.9}$$

The Monte Carlo method thus depends on choosing N points randomly to approximate I_N, and hence a different result is obtained for each unique set of points x_i. Then, the variance of the function is given by

$$var(f) \equiv \sigma^2 = \frac{1}{N-1} \sum_{i=1}^{N} (f(x_i) - \langle f \rangle)^2. \tag{4.10}$$

Since for any independent stochastic variables Y_i

$$var\left(\sum_{i=1}^{N} Y_i\right) = \sum_{i=1}^{N} var(Y_i), \tag{4.11}$$

it follows that

$$var(I_N) = (b - a)^2 \frac{var(f)}{N} = (b - a)^2 \frac{\sigma^2}{N}. \tag{4.12}$$

At large N, the variance decreases asymptotically as $1/N$. Thus, the error estimate

$$\delta I_N \approx \sqrt{var(I_N)} = (b - a) \frac{\sigma}{\sqrt{N}} \tag{4.13}$$

decreases as $1/\sqrt{N}$. The real advantage of Monte Carlo is apparent in higher dimensions.

For a given dimension d, both the trapezium method and Simpson's method converge as $\propto N^{2/d}$ and $\propto N^{4/d}$, respectively, whereas for any dimension d the Monte Carlo method still converges as $\propto 1/\sqrt{N}$. As a result, the easiest way to get a N-fold increase in accuracy is to input N^2 times as many points into the algorithm. For a given collision event at the LHC, the number of produced particles is on the order of 1000, and hence the dimensionality would be approximately $d = 3^{1000} - 4$, where 4 is subtracted to account for conservation of energy and momentum. For experimental particle physics, the integral to be evaluated is the cross section of a physics process, which can be written as:

$$\int_{3^{1000}-4} \left(\frac{d\sigma}{d\Omega}\right) d\Omega. \tag{4.14}$$

4.2.2 Generators

Several generators are in use at CMS which use Monte Carlo techniques to generate high-energy physics events in a way that is compatible with CMSSW. Among these are PYTHIA, currently at version 8.1. Recently rewritten in C++ from Fortran-77, PYTHIA generates collisions at high energies between elementary particles such as e^+, e^-, p and $\bar{p}$ in various combinations. It contains theory and models for a number of physics aspects, including hard and soft interactions, parton distributions, initial- and final-state parton showers, multiple interactions, fragmentation and decay. It is largely based on original research, but also borrows many formulae and other knowledge from existing literature[42]. It is capable of generating initial- and final-state showers, multiple parton-parton interactions and beam remnants, and an extensive library of SM and BSM subprocesses, for providing a complete description of event structure in high-energy collisions between hadrons or leptons.

Herwig is a Monte Carlo package for simulating Hadron Emission Reactions With Interfering Gluons, and is written in Fortran (currently at version 6.510). It is being superseded by HERWIG++, written in C++ and currently at 1.0. It is closely linked to version 7 of PYTHIA, but uses completely independent physics implementations. Its major foci are careful modeling of parton showers and hadron formation. It is fully operational for e^+-e^- annihilation, with work in progress for proton-proton and electron-proton collisions.

Sherpa, unlike the other multi-purpose simulation programs mentioned above, has no predecessor written in Fortran, but rather was written from the beginning in C++. It is comprised of a highly modular framework, and can generate automated hard matrix elements for tree-level processes, parton showers, simple multi-parton interactions, and hadron decays with special emphasis on tau and heavy quark decays. Planned for the near future are a new matrix element generator allowing for higher multiplicities, as well as further parton shower modules.

For automated computation of matrix elements in Standard Model processes, a software package called MadGraph is used, which allows one to generate amplitudes and events for any process with up to nine external particles, using any model (such as Standard Model, Higgs effective couplings, MSSM, etc). MadGraph is not integrated directly into CMSSW, but rather passes its output along in the form of LHE files[8].

Another framework used at CMS is POWHEG, which is a general-purpose framework for implementing NLO calculations in shower Monte Carlo programs. Typically, POWHEG is used in conjunction with one of the aforementioned generators, which usually only carry perturbative QCD expansions to leading-order (LO). The available processes are single vector boson production with decay, vector boson plus one jet production with decay, single top production in the s- and t- channels and associated with a W boson, Higgs boson production in gluon and vector-boson fusion, jet pair production, heavy-quark pair production, ZZ production, W in association with $b\bar{b}$ production, and W^+W^- plus dijet production[43].

4.2.3 Detector Simulation with Geant4

Finally, generated events and datasets are made to resemble actual conditions at the detector by using a detector simulator package called Geant4 ("GEometry ANd Tracking"). Geant4 is a toolkit for simulating the passage of particles through matter, and provides simulation of tracking, geometry, physics models, and hits, covering electromagnetic, hadronic and optical processes. Written in C++, it is a modular toolkit with modules written for geom-

[8]LHE files are named after the Les Houches Accord, which was an agreement reached between particle physicists standardizing the interface between the matrix element programs and the event generators at a conference in Les Houches, France in 2001.

etry and materials, particle interaction in matter, tracking management, detector response, digitization and hit management, event and track management, a visualization framework and an end-user interface.

4.2.4 Parton Distribution Function Sets

Several different sets of parton distribution functions have been compiled and made available for use in conjunction with the various modular Monte Carlo frameworks. Of these, three PDF sets are truly global; that is, they use results from the Tevatron as well as from DIS experiments at HERA. These are CTEQ, MSTW and NNPDF. These PDF sets differ from each other in ways such as different α_s values (0.118 for CTEQ6.6, 0.119 for NNPDF and 0.120 for MSTW). Other differences include MSTW being the only PDF set at NNLO derived from a fully global fit, as well as different subselections of data, different treatment of systematic errors, treatment of heavy flavors, etc[44].

4.2.5 Truth

In reconstructed MC datasets, additional information is available which contains the generator-level particle information, such as physical and kinematic properties of the particle and the particle's parent and daugher particles (if applicable). These collections, known as MC Truth, can then be compared to the equivalent reconstructed particle which has been fed through the detector simulator and standard reconstruction algorithms for such purposes as determining scale and resolution, and examining the efficacy of identification and reconstruction in CMS.

4.3 Reconstruction

Reconstruction is the process of converting the morass of electrical signals from the detector into events [9] and physics objects, as well as their associated properties (charge, momentum, etc.), into well-defined and easily-accessible objects for offline analysis. To that end, a unified modular software framework was developed, which is capable of both recording data

[9] An event is defined as a single triggered beam crossing.

and analyzing them offline. Data are stored and read from a common file format compatible with the ROOT data-analysis system.

4.3.1 CMSSW

The software framework for analysis is known as CMSSW, and is at version 4_2_3 as of this writing (4_2_3 and 3_8_7 were used in this analysis). In addition to source and output modules, there are several different modules which are relevant for analysis. A producer module, or `EDProducer`, builds and inserts into existing events new user-defined collections. For example, one might run an algorithm to compute a quantity such as missing transverse energy $\not{E}_T$ and store it in the event for further upstream processing. A filter module, or `EDFilter`, works in much the same way as a trigger, filtering out events which do not meet certain criteria. It is frequently used for skimming, or producing compact data sets which meet certain criteria for further analysis. An analysis module, or `EDAnalyzer`, is the default-use case. It is used for making histograms (or some other output of the user's choosing), and cannot modify the data set. Finally, the `EventSetup` module is an external service not bound to the event structure. It is used for such tasks as connecting to databases, reading detector geometry, extracting the magnetic field, and so on.

4.3.2 Muon Reconstruction

Muon reconstruction is divided into three steps. First, hits and segments are reconstructed inside a single chamber. Then, tracks are reconstructed inside the muon system alone. Finally, reconstruction of the final track takes place using the entire CMS tracking system[45]. The muon reconstruction process begins with local reconstruction. To begin, hits in DTs and CSCs are reconstructed from digitized electronics signals, and then matched to form segments, or track stubs. Simultaneously, tracks are reconstructed in the central tracker.

In offline reconstruction, the reconstructed segments from the muon chamber are used to generate seeds consisting of position and direction vectors and an estimate of the muon transverse momentum, which are then used for track fits in the muon system. This is accomplished by using a Kalman filter technique. A Kalman filter technique is an iterative algorithm which uses a dynamical model of a system, a set of known initial conditions and repeated measurements to fit the model to the data. In this case, the fit is to the track, and is based on rejecting hits lying outside of a cone and repeating the algorithm, reducing the cone size with each iteration. The result is a collection of track objects in the muon spectrometer, which are known as standalone muons. To improve muon resolution, an optional beam-spot constraint can be applied to the fit.

For each standalone muon track, a search for corresponding tracks inside the silicon tracker is performed. The best-matching one is chosen, and the track fit is repeated, again using a Kalman filter technique. The combined track from the silicon tracker and the standalone muon is known as a global muon. A complementary approach is to treat every tracker track as a potential muon candidate, and look for compatible signatures in the muon detectors. These are known as tracker muons, and are primarily useful for low-p_T muons, which may not have enough hits in the muon system to reconstruct a standalone muon. Finally, a collection of tracks known as calorimeter muons is reconstructed from the subset of all tracker tracks in the event which have an associated energy deposition in the calorimeter system.

4.3.3 $\not\!\!E_T$ Reconstruction

Neutrinos, by virtue of their negligible mass and lack of an electric charge, pass through CMS without being detected. As a result, the presence of such particles has to be inferred by an imbalance in the total transverse energy of the event. There are three methods in use at CMS for measuring $\not\!\!E_T$: calorimeter MET, track-corrected MET and particle-flow MET.

CaloMET

The traditional method of measuring $\not\!\!E_T$ at hadron colliders is by summing up energy deposits in the calorimeters. This quantity is calculated as the negative vector sum of the transverse energies deposited in the calorimeters, followed by sequential corrections for the presence of identified muons and the underestimation of the hadronic energy deposits in the calorimeter. Correcting for identified muons is performed first, and is accomplished by substituting the transverse momentum of the reconstructed track in the silicon tracker for the minimum ionizing transverse energy expected in the calorimeters. Then, jet energy scale corrections are applied to the transverse energies of the reconstructed jets. The end result is made available for offline analysis as the CaloMET collection.

tcMET

Track-corrected MET, or tcMET, is calculated by taking the uncorrected CaloMET collection in conjunction with muons, electrons and central tracks. Muons are corrected for by subtracting out the muon p_T and adding an offset to the calorimeter deposit consistent with minimum ionization. The transverse energy is further corrected by using all tracks not matched to leptons, and finally by removing the contribution of each good calorimeter track to the E_T and replacing it with the momentum of the track at the vertex.

pfMET

Finally, particle-flow MET, or pfMET, is $\not\!\!E_T$ reconstructed by using particle flow event reconstruction. Particle-flow reconstruction aims to reconstruct every single stable particle in the event, be they muons, electrons, photons, charged or neutral hadrons. It accomplishes this by using data from all CMS subdetectors to make as accurate a determination as possible of kinematic quantities such as direction, energy, charge and type. Then, the p_T vector sum over all reconstructed particles in the event is computed, the opposite of which is the pfMET collection.

4.3.4 Jet Reconstruction and Clustering

Jets are reconstructed in two stages at CMS: a jet reconstruction algorithm provides the inputs to a jet clustering algorithm. Analogous to $\not\!\!E_T$ reconstruction, four different jet reconstruction algorithms are in use at CMS, which combine in different ways information from the various subdetectors into inputs for the jet clustering algorithm. A jet clustering algorithm must meet several criteria:

- Collinear safe: the output of the jet algorithm remains the same if the energy of a particle is distributed among two collinear particles.

- Infrared safe: the output of the jet algorithm remains stable even if soft particles are added.

- Insusceptibility to contamination due to pile-up and underlying event.

There are four jet reconstruction algorithms in use at CMS (Calorimeter jets, Jet-Plus-Tracks jets, Particle Flow jets, and Track jets), as well as five main jet clustering algorithms (iterative cone, midpoint cone, seedless infrared safe cone, k_T, and anti-k_T).

Calorimeter Jets

The calorimeter jet collection, or Calo jets, are reconstructed using energy deposits in the electromagnetic and hadronic calorimeter cells, combined into calorimeter towers. In order to suppress calorimeter readout electronics noise, thresholds are applied to energies of individual towers for reconstruction of jets (and $\not{E}_T$). To suppress pile-up contributions, the contribution from calorimeter towers with $E_T^{towers} < 0.3$ GeV are ignored.

Jet-Plus-Tracks Jets

The Jet-Plus-Tracks algorithm, or JPT, begins with calorimeter jets as described above, then charged particle tracks from the tracking subsystems are associated with each jet based on separation in η-ϕ between the jet axis and the track momentum, as measured at the interaction vertex. Next, the associated tracks are projected onto the surface of the calorimeter, and further classified as either an in-cone track if the projection is within the jet cone on the calorimeter surface, or an out-of-cone track if the CMS magnetic field has bent the track so that it lies outside the jet cone. Finally, the momenta of both in-cone and out-of-cone tracks are added to the energy of the associated Calo jet, with the expected average energy deposition in the calorimeters subtracted out of the momentum of each in-cone track. Furthermore, the direction of the axis of the original Calo jet is corrected.

Particle Flow Jets

The Particle Flow algorithm combines information from all CMS subdetectors to reconstruct all particles in the event, including jets. Charged hadrons are reconstructed from tracks in the central tracker, and photons and neutral hadrons are reconstructed from energy deposits in the calorimeters. Clusters separated from the extrapolated track positions in the calorimeters provide a clear signal of neutral particles, and a neutral particle overlapping with charged particles in the calorimeters is detected as an excess of calorimeter energy with respect to the sum of the momenta of the associated tracks. Then, the resulting list of particles is used to reconstructed PF jets. Jet momentum and spatial resolutions are expected to be better than with Calo jets, due to the excellent performance of the CMS tracker and the granularity of the ECAL.

Track Jets

Track jets are reconstructed from the tracks of charged particles measured in the central tracker, with selection cuts made on the association with the primary vertex and track quality. As track jets are reconstructed with absolutely no input from the calorimeters, it serves as a cross-check for the other types of jets.

Iterative Cone

A cone is defined as a fixed radius in η and ϕ in the direction of the dominant energy flow. For the Iterative Cone algorithm, calorimeter towers or particles with a transverse energy of at least 1 GeV are considered in descending order as seeds for an iterative search, such that all inputs satisfying

$$\sqrt{\Delta\eta^2 + \Delta\phi^2} \leq R \tag{4.15}$$

are associated with the jet, where R is the cone size parameter. Then, a cone is accepted as stable if its center is consistent with the (η, ϕ) location of the vector sum of the constituent 4-vectors. Once a stable cone is found, the iteration stops and the result is declared as a jet. Although this algorithm is neither infrared safe nor collinear safe, it is in use at the

HLT due to its simplicity and speed.

Midpoint Cone

The Midpoint Cone algorithm is quite similar in concept to the Iterative Cone algorithm, in that it is an iterative procedure over cones with the same seed requirements imposed. However, unlike in the Iterative Cone algorithm, infrared safety is addressed by using as additional seeds all midpoints between each pair of jet candidates which are closer than $2R$, and not restricting each input to only one jet candidate. After the algorithm is complete, a further splitting and merging algorithm is applied to ensure that each input is associated with only one jet. Despite being an improvement over the Iterative Cone algorithm, the Midpoint Cone is still not infrared safe for QCD interactions beyond NLO.

Seedless Infrared-Safe Cone

As the name implies, the Seedless Infrared-Safe Cone (SISCone) algorithm is motivated by the lack of infrared safety inherent in the Iterative Cone and Midpoint Cone algorithms. All possible cones are tested: for each two objects located within a distance of $2R$ (the definition of R changes to be a function of the rapidity y rather than pseudorapidity η), the two cones with both objects located on the circumference are tested for stability. Some clusters might not be stable due to nearby jets, in which case they are removed and the algorithm is repeated. The SISCone algorithm is exact, infrared safe and collinear safe; however, it is computationally quite expensive, taking time of $\mathcal{O}\left(N2^N\right)$. Thus, the SISCone method has been repeatedly refined and massaged such that its execution time is comparable to that of Midpoint Cone.

k_T and anti-k_T

Unlike the cone algorithms, k_T and anti-k_T are sequential clustering algorithms; that is, there is no fixed cone size. For each input, the distances to the beam line

$$d_i = \left(E_{T,i}\right)^2 \cdot D^2 \tag{4.16}$$

and to each of the other particles

$$d_{ij} = \min\left(E_{T,i}^2, E_{T,j}^2\right) \cdot R_{ij}^2 \tag{4.17}$$

are calculated. Then, for each i, the smallest d_{ij} is found and compared to d_i. If $d_i > d_{ij}$, then i is moved to the list of final jets; if not, i and j are merged. The k_T algorithm has the advantage of being both infrared safe and collinear safe, and the fact that there is no fixed cone means that there is better clustering of heavy highly boosted decaying particles. However, it also comes with a high computational cost of $\mathcal{O}\left(N^3\right)$. To that end, the Fast k_T algorithm was developed, which only calculates d_{ij} to the nearest neighbor, and carries with it a computational time requirement of $\mathcal{O}\left(N \log N\right)$. The anti-$k_T$ algorithm is similar in principle, but introduces a new factor p, like so:

$$d_{ij} = \min\left(E_{T,i}^{2p}, E_{T,j}^{2p}\right) \cdot R_{ij}^2. \tag{4.18}$$

For $p = 1$, the inclusive k_T algorithm is recovered. The case where $p = -1$ is referred to as the anti-k_T jet clustering algorithm, which is similar in terms of speed and output to the Iterative Cone algorithm, with the exception that it is infrared safe and collinear safe, and possesses a different relative power of energy to geometrical scale ratio than k_T.

4.4 Acceptance and Efficiency

Another important consideration is detector acceptance and efficiency. The detector is not isotropic in transverse momentum or pseudorapidity, as there are regions in which detector coverage necessarily suffers. Specifically, the transition region between barrel and endcap ($0.8 < |\eta| < 1.2$) is dominated by the iron yoke of the magnet, and hence is relatively devoid of muon chambers and calorimetry hardware as compared to the rest of CMS. Furthermore, the region $|\eta| > 2.5$ is not covered by the detector, and is mitigated by the requirement $|\eta| < 2.1$ imposed on the selection cuts. Efficiences are the fraction of actual signal events which are recognized by the detector out of all the actual signal events. Finally, comparing acceptance and efficiency results for data and MC is an effective way to verify the modelling of physics processes and detector performance in MC, as well as testing how well the simulated MC compares to real collision data.

4.4.1 Muon Acceptance

Tag and Probe

The so-called Tag and Probe method is a data-driven method which uses a well-known dimuon mass resonance (such as J/Ψ or Z$\rightarrow$ $\mu^+\mu^-$). Then, one of the muons (the "tag" muon) is required to pass a tight set of selection criteria designed to isolate the required particle type, as well as having a very small fake rate ($\ll$ 1%). The other muon (the "probe" muon) is required to pass a set of different, much looser criteria, and is paired with a "tag" muon such that the invariant mass of the combination is consistent with the resonance mass. Background contributions are removed through techniques such as fitting or sideband subtraction. Then, the ratio of "probe" muons passing the selection cuts (which are defined according to the efficiency to be measured) is the efficiency in question:

$$\epsilon = \frac{P_{\text{pass}}}{P_{\text{all}}}, \tag{4.19}$$

where P_{pass} is the number of probes passing the selection criteria and P_{all} is the total number of probes counted.

Some inherent biases are present in the Tag and Probe method, which causes the efficiency measurement to be lower than the true efficiency, an effect which is more marked at lower efficiencies. This bias arises from the fact that there are two possible scenarios for the probe muon, regardless of whether or not it passes the "passing probe" selection criteria: it can either pass or not pass the tag muon criteria. If it passes, then the efficiency of the tag muon passing the probe criteria is measured. If, however, the probe does not pass the tag criteria then its contribution to the efficiency is not counted, even though it may be a perfectly valid probe. Therefore, a portion of the muon sample which is likely to pass the probe criteria is absent, and thus the overall efficiency measurement is biased low.

Muon Identification

Muon identification efficiency is the efficiency rate of real prompt muons being identified by the detector, as opposed to the misidentification of pions, kaons and protons as muons. For measuring muon identification efficiency, the Tag and Probe method with the Z resonance was used. The dimuon pairs were collected using high-p_T single-muon triggers so as not to bias the trigger selection on the probe muon. The probes are tracks reconstructed using only the inner tracker[10], so as to obviate bias from the muon subdetectors. To subtract out background not originating from the Z resonance, a simultaneous fit is performed to the invariant mass spectra for passing and failing probes with identical signal shapes, with the efficiency computed from the normalizations of the signal shapes in the two spectra.

Muon identification efficiency was also studied using simulations in MC, with Z$\rightarrow \mu^+\mu^-$, W+jets, and muon-enriched QCD samples being generated and studied. In general, data and MC efficiencies are in excellent agreement, with an overall reconstruction efficiency of $99.7 \pm 0.1\%$ using Tag and Probe and a Data/MC ratio of 0.999 ± 0.006.

Muon Trigger

As mentioned in Section 3.2.6, the muon trigger consists of two primary components: the hardware-based Level-1 (L1) Trigger, and the software based High-Level Trigger (HLT). The trigger efficiency for prompt muons was determined using Tag and Probe with Z resonances for the range $20 < p_T < 100$ GeV/c. The probe is matched to a trigger object as follows: first, a muon's tracker track is extrapolated to the muon system, whereupon Level-1 trigger candidates are matched by position. As the HLT includes reconstruction of tracker tracks, the HLT muon is matched to an offline-reconstructed tracker track by comparing directions away from the vertex, and HLT-only efficiences are computed by matching the probe with the L1 candidate, and requiring it to be matched to the HLT candidate as well. The combined L1 and HLT efficiency is obtained by requiring the probe to be matched with the HLT candidate.

[10]The inner tracker efficiency was separately measured to be $> 99\%$[46]

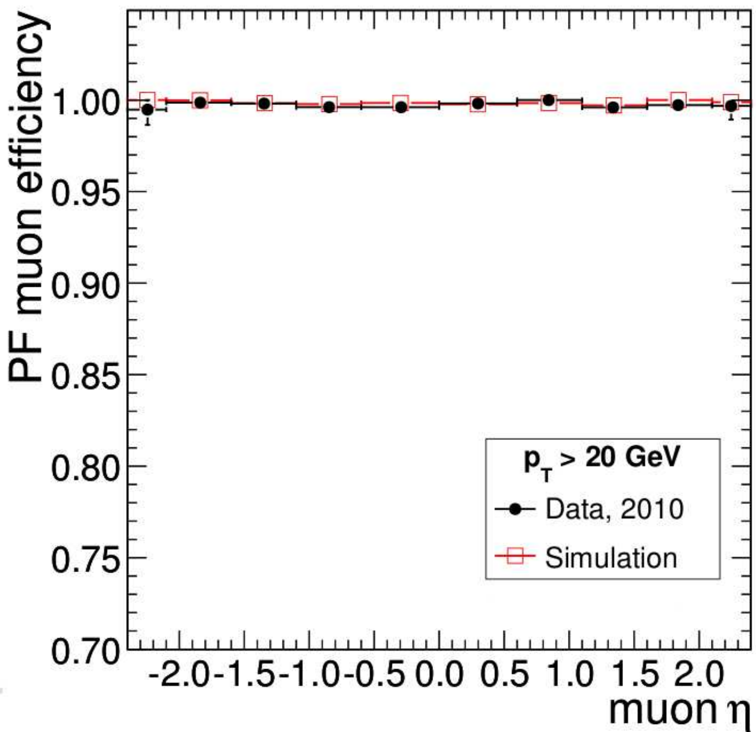


Figure 4.2: Tag and probe results for muon identification efficiency from Z→ $\mu^+\mu^-$. The dataset is the entire 2010 run, and the muon identification efficiency for Particle Flow muons, given that a tracker track exists, is shown as a function of η.

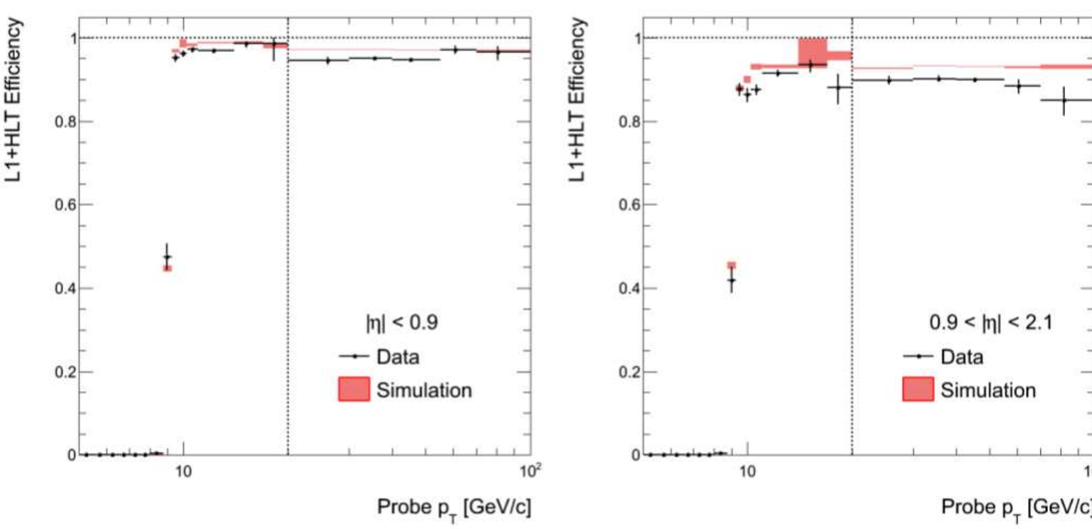


Figure 4.3: Single-muon trigger efficiencies as a function of muon p_T in the barrel (left) and in the overlap-endcap (right) regions. The combined Level-1 and HLT efficiency is shown.

As with the muon identification study, the Z events were selected with high-p_T single-muon triggers, and the muon range $20 < p_T < 100$ GeV/c was considered. A comparison to simulated events in MC was also made, and the overall efficiency rate for combined L1 and HLT triggers was $92.4 \pm 0.3\%$, with a data/MC ratio of 0.971 ± 0.003.

Muon Isolation

Muon isolation is an important variable in discriminating against QCD background. The type of isolation considered for this analysis is the so-called combined relative isolation, in which any track reconstructed in the inner tracker whose distance $\Delta R \equiv \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.3$ is considered. Then, the scalar sum of the p_T of all tracks (excepting the muon track) and the ECAL and HCAL energy depositions is computed, and the ratio of the scalar sum to the muon track p_T is required to be less than a certain threshold[11] for the muon to be considered isolated.

The efficiency of the combined relative muon isolation algorithm was tested using the Tag and Probe method with Z^0 decays, as well as the Lepton Kinematic Template (LKT) method. For the former, the tag muon was required to satisfy the tight muon requirements and have a combined relative isolation < 0.15. Probe muons were also required to satisfy the tight muon identification requirements, and the invariant mass of the combination was required to reside between 70 and 110 GeV/c^2. The LKT method is an improvement upon the "random cone" method, wherein isolation cones are constructed in random directions. The LKT method is predicated upon the assumption that in W/Z events in pp collisions, the kinematics of the muons produced in hard scattering processes are unrelated to the underlying event. Directions in space are drawn from MC-derived kinematical distributions of the muons under consideration. Then, the isolation variables were calculated for each of these random cones, and tested against a threshold in underlying events which are similar to those of the signal under consideration, for example, high-purity golden W$\rightarrow \mu\nu_\mu$ and Z$\rightarrow \mu\mu$ events.

From Figure 4.4, the isolation efficiency for $I^{rel}_{comb} < 0.10$ is 97% or greater, with a

[11] In the case of this analysis, $I^{rel}_{comb} < 0.10$.

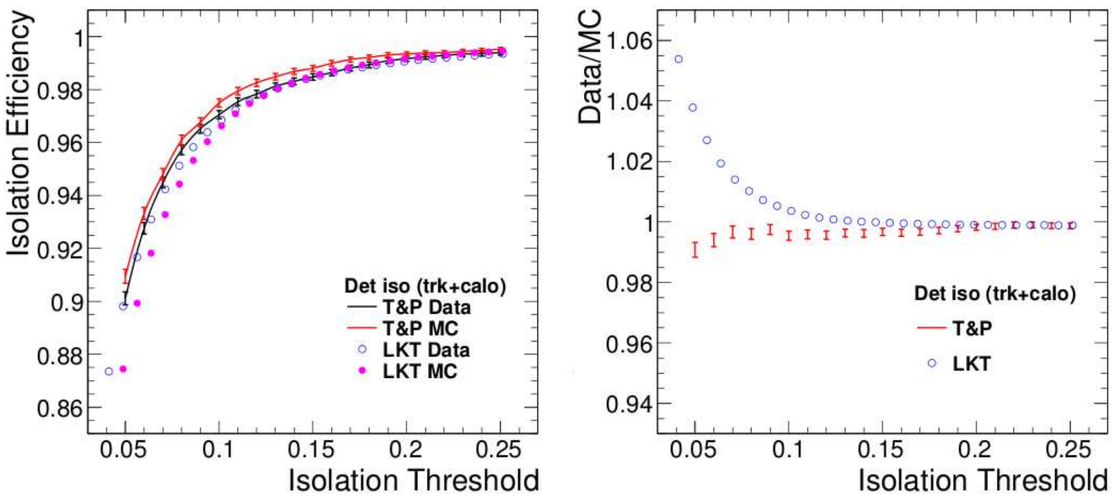


Figure 4.4: (left) efficiency of the combined relative isolation algorithm with muons from Z^0 decays with $20 < p_T < 50$ GeV/c as a function of the isolation variable threshold. Results are shown for both data and MC using both the Tag and Probe and Lepton Kinematic Template methods. (right) Data to MC efficiency ratio.

Data/MC ratio discrepancy of $< 1\%$.

Muon Momentum Scale and Resolution

Uncertainty in the muon transverse momentum can be affected by: alignment of the tracker and muon chambers, material present in the detector which is not properly modeled in MC, and uncertainties in the magnetic field mapping. Depending on the p_T range, one of two methods was used to study muon momentum scale and resolution. In the range $0 < p_T < 100$ GeV/c, muonic decays originating from J/Ψ and Z events are used to calibrate the momentum scale and measure its resolution by employing a mass constraint. In the range $p_T > 100$ GeV/c, muons originating from cosmic rays are used.

In the medium-p_T range, two methods are used to study the muonic p_T measurement. The first is the Muon momentum Scale calibration Fit, or MuScleFit, which fits each resonance of the dimuon spectrum to a Voigtian (Lorentzian convoluted with a Gaussian) profile. For the Lorentzian, Γ and M are taken from accepted PDG values, and the free parameters of the fit are the resulting muon momentum scale and resolution. The second method is the SImulation DRiven Analysis, or SIDRA, which compares the data to the full detector simulation of the Z→ $\mu^+\mu^-$ decay, and hence provides a way to directly modify the simulation so that it better represents actual collision data. As the two methods are complementary, they can be used for comparison, with the differences between the two assigned to systematic uncertainty.

In the high-p_T range, studies are performed by splitting cosmic tracks, with each track being used to measure the resolution of different momentum algorithms. Cosmic rays from the 2008 Cosmic Run At Four Tesla (CRAFT08) were used, with different muon reconstruction algorithms applied to tracks in the upper half and lower half of the detector. To best simulate collision data, muon tracks were selected to be close to the interaction point by requiring at least one pixel hit. However, as a direct consequence of this requirement, this technique is significantly harder to apply to endcap muons than barrel muons.

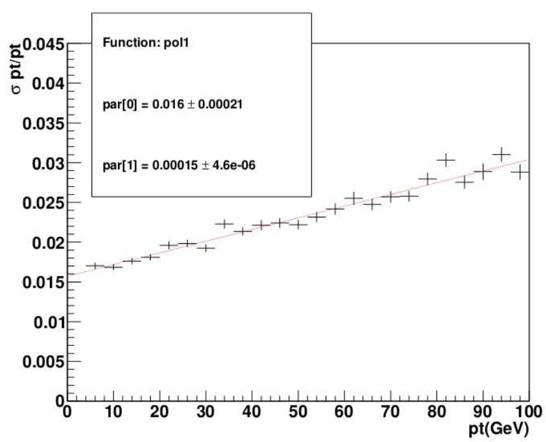


Figure 4.5: Relative transverse momentum resolution as a function of muon transverse momentum for global muons, determined from Gaussian fits to the difference between MC truth and MC reco values. The overlaid curve is the parametrization of the resolution.

Cosmic Ray and Beam Halo Background

A cosmic ray muon typically originates from pion and kaon decays well above the earth's surface, and contributes several types of background. A cosmic ray passing close to the IP can be falsely reconstructed as a collision muon, a muon that is not reconstructed in the tracker can still be reconstructed as a standalone muon in the muon system and matched to a tracker track, and cosmic rays passing through the calorimeters can deposit energy, which can result in fake $\not\!\!E_T$. Cosmic rays were studied using three data samples: MC dimuon events originating from Drell-Yan processes, prompt collision muons selected by requiring a well-constructed primary vertex, and cosmic ray muons selected by requiring zero good primary vertexes and at most two reconstructed tracker tracks.

Due to the nature of cosmic rays (nearly top-down trajectory through the detector, no reconstructable vertex, etc), there are several observables which discriminate well against cosmic rays. Chief among these are the transverse impact parameter with respect to the primary vertex d_{xy}, as well as requiring a good primary vertex and at least one hit in the inner tracker. The histogram of $|d_{xy}|$ is strongly peaked at zero for collision muons, and is flat for cosmic rays. The typical cut on $|d_{xy}|$ (including this analysis) is $|d_{xy}| < 0.2$ cm.

In the event that a cosmic muon passes close to the center of the tracker, it could be reconstructed as a global muon. However, in this scenario, one would expect to find a track with p_T of nearly equal magnitude and opposite direction. Therefore, the distribution of the angle in three-dimensional space between two muon tracks would be more biased towards π radians for cosmic muons than for collision muons. Cosmic muons can also be distinguished by measuring the time each muon would pass the interaction point, under the assumption that the muon is moving at the speed of light from the inside of the detector outward. Collision muons are peaked around zero (with the width of the peak determined by the precision of time measurement), whereas cosmic muons are significantly flatter but also centered around zero (due to trigger and tracker reconstruction being most efficient for in-time particles). Distributions of the shapes of transverse impact parameter, opening angle, muon time at vertex and top-bottom difference of muon time at vertex for cosmic

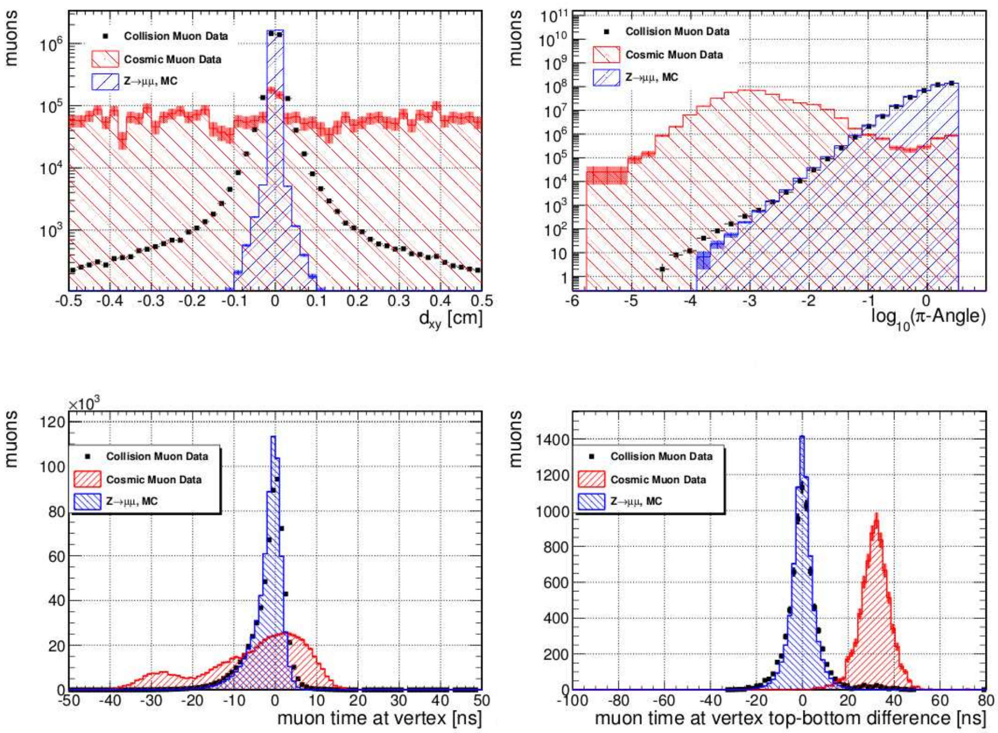


Figure 4.6: Distribution of variables used in the identification of cosmic muons, for collision muon data, cosmic muon data and Z→ $\mu\mu$ MC. (Clockwise from top left) Transverse impact parameter with respect to the primary vertex, distribution of angle between muon tracks, difference between the time the muon passes through the vertex between the tracks in the top and the bottom of the detector, and the time at which the muon passes the vertex.

muons, collision muon data and Z→ $\mu^+\mu^-$ MC can be seen in Figure 4.6.

For the simple cut $|d_{xy}| < 0.2$ cm, the overall efficiency (defined as the fraction of preselected cosmic muon data properly flagged as cosmic muons) is $99.05 \pm 0.04\%$, and the misidentification rate (defined as the fraction of Drell-Yan signal MC misidentified as a cosmic muon) is 0.0045 ± 0.0003 %.

Beam halo (or machine-induced) muons typically follow a trajectory parallel to the beam line. They are identified using a framework consisting of three components: a dedicated L1 beam-halo trigger, the presence of early triggers at the per-chamber level and the reconstruction of a standalone muon track whose trajectory is parallel to the beam line. The "tight" beam halo ID criterion is for any two of the three framework components to be met to identify a muon as beam-halo, and the "loose" beam halo ID criterion is for any one to be met. To measure the performance of the beam-halo algorithms, two datasets were used: 7 TeV MC samples and 2010 collision data. Efficiencies were measured by first selecting for events in which missing transverse energy ($\not\!\!E_T$) was present, motivated by the fact that beam halo muons passing through the endcap calorimeters at constant ϕ can induce a large $\not\!\!E_T$ signal.

Beam halo MC was found to pass the "loose" halo ID 96% of the time, and the "tight" halo ID 65% of the time. By using the requirement of $\not\!\!E_T > 50$ GeV and oriented π radians away from at least one of the reconstructed CSC hits, and vetoing events with reconstructed collision muons, the probability of the event passing the "loose" ("tight") halo ID was found to be 89.3% (73.3%). The probability of misidentifying a collision muon as beam halo in MinBias MC events was found to be on the order of 5×10^{-5} and 5×10^{-7} for "loose" and "tight" beam halo criteria, respectively. For data MinBias events, the probabilities were estimated to be on the order of 2×10^{-4} and 8×10^{-7}, respectively. Finally, the probability for a beam-halo muon to be misidentified as a collision muon was found to be negligible for muons of $p_T > 5$ GeV/c.

4.4.2 Jet Acceptance

Jet Energy Calibration

Particle jet energy is obtained by clustering particles produced during hadronization after the hard interaction. However, particle jet energy is typically different from the jet energy measured in the detector due to the non-uniform and non-linear response of the CMS calorimeters. Furthermore, electronics noise and pile-up can further bias the measurement. To that end, a multi-step jet energy calibration algorithm was developed, whose purpose is to provide a correction factor to normalize the two quantities. Three steps were devised to correct reconstructed jets to the particle jet level: offset (designed to correct jet energy for electronics noise and pile-up), also known as L1; relative (designed to remove variants in jet response relative to η due to non-uniformity of the detector), also known as L2; and absolute corrections (designed to remove variations in jet response versus jet p_T), also known as L3. These three corrections were designed to be applied sequentially, and were derived by using MC truth information, and from using physics processes from actual pp collisions for in-situ calibrations.

The MC truth jet energy corrections were derived by using PYTHIA QCD events at $\sqrt{s} = 7$ TeV. First, calorimeter, JPT, and Particle Flow jets were reconstructed, as were particle jets from the four-momenta of the MC particles[12], upon which the two were matched in the $\eta - \phi$ plane by requiring $\Delta R < 0.25$. Then, for the matched jets, the quantity p_T^{Jet}/p_T^{GenJet} was studied in order to extract jet calibration factors. Uncertainties were estimated by re-deriving the correction factors for the corrected jets, and found to be within 2%. Conservatively, a 2% uncertainty is thus assigned to MC truth jet energy corrections.

The first step in the factorized corrections is the offset correction. Offset components arising from noise, noise plus one pile-up, and the total average are measured separately. The noise-only contribution is measured by using events from random trigger (with the only precondition being the presence of a beam crossing, also known as a ZeroBias trigger), with MinBias trigger events vetoed. Then, the average calorimeter energy summed up inside a

[12]Hereafter referred to as GenJets.

cone of radius $R = 0.5$ at a given η (hereafter referred to as $E_{\mathrm{offset}}(\eta)$) is studied. To study noise plus one pile-up, MinBias trigger events in early runs (where the fraction of events with more than one interaction per bunch crossing is small) were selected, and $E_{\mathrm{offset}}(\eta)$ was investigated and compared to PYTHIA MinBias MC. Finally, the total average offset was determined from inclusive ZeroBias events (with no veto on MinBias triggers). Uncertainties arise from the fact that on average 50% of high-p_T events have additional pile-up, as well as the fact that in some regions noise, pile-up and jet contributions overlap. This yields a systematic uncertainty of 2%, as well as a systematic effect from data/MC differences of 2% for jets at $p_T = 20$ GeV, and decreasing with increasing p_T.

The second step is the relative correction, which is derived using a technique called "dijet p_T balance", and is somewhat similar conceptually to Tag and Probe. In back-to-back dijet events, one "tag" jet satisfying $|\eta| < 1.3$ in the central calorimeter region and one "probe" jet at arbitrary $|\eta|$ are selected, and the p_T balance is measured. Barrel jets are used for the "tag" jet because of the relative uniformity of the detector in that region, and because it has the highest jet p_T reach. Dedicated High-Level triggers fire on the average uncorrected $p_T = (p_{T,1} + p_{T,2})/2$ of the two leading jets above thresholds of 15 GeV/c and 30 GeV/c. Then, the p_T balance

$$B = \frac{p_T^{probe} - p_T^{barrel}}{p_T^{dijet}} \tag{4.20}$$

is examined in bins of η_{probe} and p_T^{dijet} for both data and PYTHIA QCD MC. Then, the relative response is determined from the average value of the B distribution $\langle B \rangle$ in a given η probe and p_T^{dijet} bin:

$$R(\eta^{probe}, p_T^{dijet}) = \frac{2 + \langle B \rangle}{2 - \langle B \rangle}. \tag{4.21}$$

Comparisons between data and MC yield a conservative estimate of uncertainty of $2\% \times |\eta|$.

The third step is the absolute correction, which corrects the calorimeter response as a function of jet p_T. To derive this correction, photon+jet events are selected and two different calibration techniques are applied: p_T balancing and the Missing E_T Projection Fraction (MPF). p_T balancing measures the balance in the transverse plane between the photon and the recoiling jet and uses the photon p_T (measured by the crystal ECAL calorimeter) as a reference object. High Level single photon triggers are collected, then a series of cuts removes the prodigious QCD background. Next, comparisons of $\langle p_T/p_T^{\gamma}\rangle$ versus p_T^{γ} are made, in data and MC. For a one-dimensional linear fit to data/MC points, the parameter value for PF Jets is 0.926 ± 0.017, with a χ^2/ndf of 4/5.

The MPF method begins with the assumption that γ + jet events have no intrinsic $\not\!\!E_T$ and that

$$\overrightarrow{p_T}^{\gamma} + \overrightarrow{p_T}^{recoil} = 0. \tag{4.22}$$

For reconstructed events, Equation 4.22 can be written as

$$R_\gamma \overrightarrow{p_T}^{\gamma} + R_{recoil}\overrightarrow{p_T}^{recoil} = -\overrightarrow{\not\!\!E_T}, \tag{4.23}$$

where R_γ and R_{recoil} are the detector response to the photon and the hadronic recoil[13]. Then, assuming a well calibrated photon ($R_\gamma = 1$), solving for R_{recoil} yields

$$R_{recoil} = 1 + \frac{\overrightarrow{\not\!\!E_T} \cdot \overrightarrow{p_T}^{\gamma}}{(p_T^{\gamma})^2} \equiv R_{MPF} \tag{4.24}$$

The final step in identifying MPF response with true jet response is to set $R_{recoil} = R_{leadjet}$, which is a good approximation in general if particles not clustered into the leading jet have a similar response to the ones inside the jet. For the MPF method applied to PF Jets, a one-dimensional linear fit to data/MC points yields 0.992 ± 0.010 with a χ^2/ndf of 6.66/5.

The final step, applied to collision data only, is the relative correction, which equals the difference between the relative response in data to MC, in terms of jet η. Based on the uncertainty in MC, and comparing the data/MC ratio, an uncertainty of $2\% \times |\eta|$ is applied.

Jet p_T and Position Resolution

Jet p_T resolution information can be extracted from MC truth as well as collision data events. For the former, PYTHIA QCD dijet MC events were used, and CaloJets, JPT and PFlow jets were reconstructed. GenJet information was also collected, and matched unambiguously to a reconstructed jet by requiring $\Delta R < 0.2$. Only the two matched

[13] Recoil is defined as the transverse momentum sum of all particles except the vector boson.

pairs with the highest p_T GenJets were considered. Then, for each pair, the jet response was defined as p_T/p_T^{REF}, where p_T and p_T^{REF} refer to the transverse momentum of the reconstructed and generator level jet, respectively. The MC-truth jet resolution can then be described by a Double Crystal Ball[14] fit.

To measure jet p_T resolution directly from collision data, a method known as dijet asymmetry is employed. MinBias and dijet p_T average triggers (with 15 and 30 GeV thresholds) are employed, events are required to contain at least two barrel ($|\eta| < 1.4$) jets azimuthally separated by $\Delta\phi > 2.7$, and any additional third jet is required to have p_T less than the p_T of the leading jet. Then, the asymmetry is defined as

$$A = \frac{p_T^{jet1} - p_T^{jet2}}{p_T^{jet1} + p_T^{jet2}}, \tag{4.25}$$

where the labels p_T^{jet1} and p_T^{jet2} are the randomly ordered transverse momenta of the two leading jets. Then, the variance of the asymmetry σ_A can be related to the jet p_T resolution like so:

$$\frac{\sigma(p_T)}{p_T} = \sqrt{2}\sigma_A. \tag{4.26}$$

Data distributions are compared to PYTHIA QCD MC expectations, and found to agree to within 10%. Sources of systematic uncertainty include the presence of additional soft radiation, parton showering and hadronization result in some particles emitted outside the jet cones, and part of the event p_T could be underlying event energy.

4.4.3 Missing Transverse Energy Acceptance

Contributions to MET Tails

Missing transverse energy, or the imbalance of momentum in the plane perpendicular to the beam direction, is the method used to infer the presence of neutral weakly interacting particles such as neutrinos. As such, it is a rather complicated component that is sensitive to detector malfunctions, cosmic ray and beam halo particles, and particles impinging poorly instrumented regions of the detector. To study $\not\!\!E_T$ acceptance, the 2010 dataset was

[14] A Double Crystal Ball is a Gaussian, with the high and low tails described by a power law.

used, as were PYTHIA6 and MadGraph simulated datasets. First, a correction to the particle level called the type-I correction is applied, and is functionally identical to the jet energy correction mentioned in the preceding section. It is applied to all PF jets with electromagnetic fraction < 0.9 and corrected $p_T > 10$ GeV.

Some instrumental causes can lead to erroneous $\not\!\!E_T$ measurements. Among these include beam halo, anomalous energy deposits in the calorimeters, detector acceptance, cracks, and dead cells (which affect the tails of the $\not\!\!E_T$ distributions). The contribution of beam halo muons was investigated by applying the CSC-based beam-halo filter to collision muons, as well as applying the collision muon High-Level trigger, and $\not\!\!E_T$ contributions were compared between events passing the collision trigger and the beam-halo trigger. In general, the contribution of beam halo to misidentified $\not\!\!E_T$ is negligible at $\not\!\!E_T$ values typically present in 2010 luminosities.

Anomalous signals in the calorimeters also have the potential to induce uncertainties in $\not\!\!E_T$ measurements. These arise from particles hitting the transducers or rare random discharges, and the basic strategy to correct for them is to look for unphysical charge sharing between neighboring channels, as well as time and pulse shape information. Comparisons of data and MC indicate that these corrections result in a 99.7% efficiency (actual physics results which pass the correction criteria). The Calorimeter $\not\!\!E_T$ distribution of a MinBias data sample before and after removal of anomalous energy deposits can be seen in Figure 4.7.

Cracks, or uninstrumented areas in CMS at the boundary between the barrel and endcap regions, can result in apparent $\not\!\!E_T$. Furthermore, approximately 1% of the ECAL crystals are either inoperational or contain high levels of electronic noise, and hence are masked out. The possible impact of ECAL masked chanels and calorimeter boundaries was explored by using MC events containing at least two reconstructed jets satisfying $p_T^{jet1} > 50$GeV, $p_T^{jet2} > 25$GeV, and checked against data.

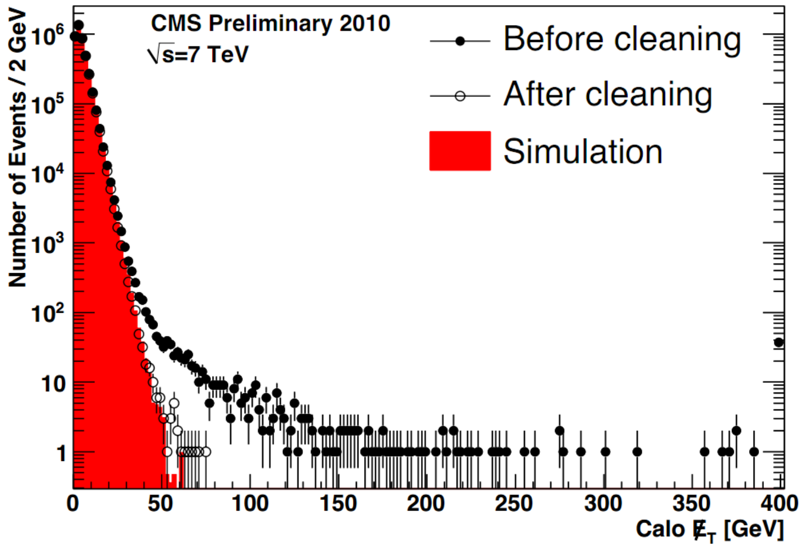


Figure 4.7: Calo $\not{E}_T$ distribution in a minimum bias data sample before (black dots) and after (hollow dots) removal of anomalous calorimeter energy deposits, compared against Monte Carlo simulated data.

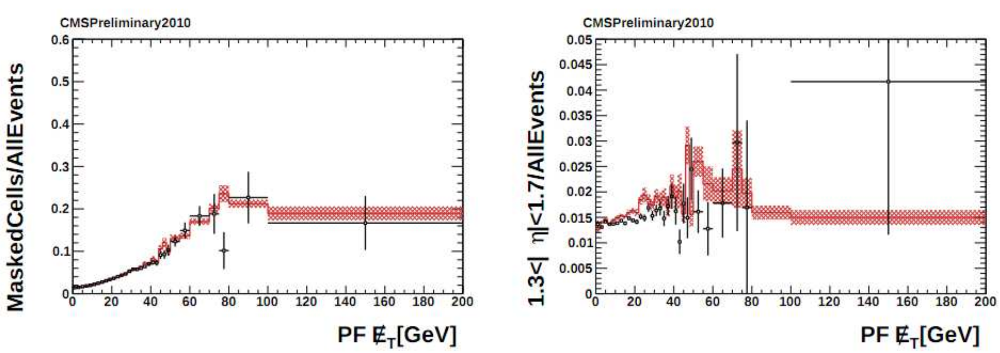


Figure 4.8: Fraction of dijet events with a jet aligned to $\not{E}_T$ via the criterion $\Delta\phi(\not{E}_T, \mathrm{jet}) < 0.2$ and pointing towards (left) an ECAL masked cell and (right) the barrel-endcap boundary.

	Particle Flow $\not\!\!E_T$ (GeV)
No b-tag (data)	11.97 ± 0.02
No b-tag (MC)	12.14 ± 0.01
SSV b-tag (data)	12.10 ± 0.11
SSV b-tag (MC)	12.51 ± 0.05
SMbyPt b-tag (data)	13.67 ± 0.70
SMbyPt b-tag (MC)	13.43 ± 0.28

Table 4.1: Mean $\not\!\!E_T$ values and statistical uncertainties for b-tagged dijet events in data and MC.

Heavy Flavor Contributions to $\not\!\!E_T$ Tails

ECAL masked channels, discussed in the previous section, are one of the two main sources of jet energy underestimation. The second is b jets containing neutrinos, due to the fact that B hadrons have unique fragmentation properties. To study their contribution to $\not\!\!E_T$ tails, an inclusive b-tagged jet sample was examined (b-jet tagging is discussed in Section 5.3.3). $\not\!\!E_T$ is compared in dijet events with and without a secondary vertex (corresponding to a positive `SimpleSecondaryVertex` or `SoftMuonByPt` tag), requiring the leading jet to have $|\eta| < 2.1$ and the two leading jets to have $p_T > 40$ GeV/c. Deviations between data and MC are explained by inadequate b-quark production modelling in MC and slightly better MC $\not\!\!E_T$ resolution. Results are shown in Table 4.1.

Missing Transverse Energy Scale and Resolution

$\not\!\!E_T$ performance is studied by using events in which an identified Z boson or γ is present. The direct photon events used were selected from 2010 data by requiring exactly one reconstructed photon in the barrel portion of the ECAL ($|\eta| < 1.479$) with p_T^V (the transverse momentum of the vector boson) greater than 20 GeV. The Z sample used was selected by requiring two well-identified and isolated electrons ($|\eta| < 2.5$) or muons ($|\eta| < 2.1$) with $p_T > 20$ GeV, and the invariant mass to satisfy $60 < M_{\mu\mu} < 120$ GeV or $70 < M_{ee} < 120$ GeV.

To study the $\not\!\!E_T$ scale and resolution in events with exactly one primary vertex, the recoil is decomposed with respect to the boson (γ or Z) direction in the transverse plane

Dataset Name	Run Range	Trigger Path	Lumi (pb^{-1})
/Mu/Run2010A-Nov4ReReco_v1/RECO	132440-144114	HLT_Mu9	3.18
/Mu/Run2010B-Nov4ReReco_v1/RECO	146240-148058	HLT_Mu11	14.52
/Mu/Run2010B-Nov4ReReco_v1/RECO	148059-149442	HLT_Mu15	18.22
Total			35.92

Table 4.2: Data sets used in this analysis.

(hereafter referred to as $\overrightarrow{u_T}$), and the recoil components parallel and perpendicular to the vector boson axis are studied. Discrepancies between data and background prediction occur due to higher resolution in data. Further sources of systematic uncertainties include residual contamination ($5 \pm 1\%$) from events with more than one interaction, uncertainties due to imperfect knowledge of the true p_T^Z distribution. Overall, the data/MC discrepancy is approximately 10%.

4.5 Signal and Background Processes

In addition to the actual dataset being used, Monte Carlo signal and background datasets must be analyzed so as to determine such quantities as relative yields, the shapes of certain kinematic distributions, and so on.

4.5.1 Dataset

The dataset used for this analysis was the entire 2010 run of proton-proton collisions at 7 TeV at CMS, from 30 March 2010 until October 2010. The run range was 132440 through 149442, corresponding to an integrated luminosity of 35.9 pb^{-1}. The different HLT paths were required due to the trigger being prescaled as a result of increasing luminosity. The November 4 ReReco was done in order to consolidate all 2010 datasets into data files compatible with a specific version of CMSSW; in this case the 3_8_X series.

4.5.2 Monte Carlo Signal W$\rightarrow \mu\nu_\mu$

For this analysis, two sets of signal Monte Carlo were used. Representing the Monte Carlo simulated signal for the W leptonic decay to jets, a W$\rightarrow \mu\nu_\mu$+jets sample was generated using MadGraph, with pile-up simulated. The MC sample contained 15,123,740 events, with a simulated cross-section of 31,314 pb. Then, Equation 3.1 yields an integrated luminosity of 482.97 pb^{-1}, which is then normalized to 35.9 pb^{-1}. The second sample, used for calculating detector efficiencies and as a cross-check for systematics, was generated using MadGraph, and pile-up was not simulated. This sample contained 5,330,940 events, with a simulated cross section of 7,899 pb.

4.5.3 QCD

Inclusive-muon QCD is one of the most prominent backgrounds in the regime of low m_T, and consists primarily of hadronizing b and c quarks with an associated muon being produced. The MC dataset used was the Inclusive Muon sample generated using the PYTHIA6 generator. InclusiveMu15 denotes that the sample only consists of events containing at least one muon with $p_T > 15$ GeV. The QCD sample contains 29,504,866 events, with a cross-section of 296,600,000 pb and a relative trigger efficiency of 0.0002855 (which denotes the total fraction of QCD events which produce at least one muon), yielding an integrated luminosity of 348.43 pb^{-1}.

4.5.4 W$\rightarrow \tau\nu_\tau$

Due to lepton universality, the W boson is approximately as likely to decay into a τ and a ν_τ as it is to a μ and a ν_μ, which presents special problems for a muonic analysis. As the τ decays almost immediately (into an electron and associated neutrinos approximately 17.85% of the time and a muon and associated neutrinos approximately 17.36% of the time, and hadronically in other cases), W$\rightarrow \tau\nu_\tau$ followed by decay of the τ into an electron or muon are experimentally indistinguishable from cases where the W decays immediately into an electron or muon. As such, W$\rightarrow \tau\nu_\tau$ constitutes an irreducible background whose presence must be estimated from Monte Carlo. The samples used in this analysis were generated using PYTHIA6 and POWHEG, with the difference between them assigned to systematic error. The PYTHIA6 sample consists of 5,221,750 events and a cross-section of 7,899 pb, or $\mathcal{L}_{int} = 661.06$ pb^{-1} The POWHEG sample consisted of two datasets, one for $W^+ \to \tau^+\nu_\tau$, and one for $W^- \to \tau^-\nu_\tau$. The W^+ sample has a cross section of 5,775 pb and 1,995,871 events, and the W^-+ sample comprises 3,944 pb and 1,994,870 events, for a total combined luminosity of 410.6 pb$^-1$.

4.5.5 Z$\to \mu^+\mu^-$

Z$\to \mu^+\mu^-$, one of the Drell-Yan background processes, produces two high-p_T muons, one positive and one negative. If one of the muons escapes the detector unseen (for example, if $|\eta| > 2.4$, outside the limits of CMS) or if $\not{E}_T$ is poorly reconstructed, then the signal can mimic that of W$\to \mu\nu_\mu$. The sample used was generated using PYTHIA6, and consists of 2,289,913 events and a cross-section of 1,300 pb, or $\mathcal{L}_{int} = 1761.47$ pb^{-1}.

4.5.6 Z$\to \tau^+\tau^-$

Z$\to \tau^+\tau^-$, the other Drell-Yan background process of note, behaves similarly to Z$\to \mu^+\mu^-$. If one or both of the tau particles decays into a muon, the process behaves functionally identically to Z$\to \mu^+\mu^-$ with one or both of the muons detected. The dataset used is generated using PYTHIA6, and consists of 2,057,446 events and a cross-section of 1,300 pb, yielding $\mathcal{L}_{int} = 1582.65$ pb^{-1}.

4.5.7 $t\bar{t}$

The top quark's only decay channel is into a W boson and a bottom-type quark (d, s or b). Due to its high mass, the top quark decays extremely quickly. Therefore, the production of a top-antitop pair becomes a significant background for W+jets, especially in the regime $N_{jets} \geq 3$. The data sample used in this analysis is generated with Madgraph, with 1,165,716 events, and a cross-section of 157.5 pb. One specific caveat is that the branching fraction of W$\to l\nu_l$ decays used in the MadGraph sample is set to its leading-order value 1/9. The current world average is 0.1080 ± 0.0009 (PDG value). This implies also that the hadronic branching ratio is 0.676 instead of 2/3. To correct for this, a scale factor must be applied. The $t\bar{t}$ sample has to be weighted with $(0.108 \cdot 9) \cdot (0.676 \cdot 1.5) = 0.985608$, and thus the integrated luminosity is $\mathcal{L}_{int}$ =7,294.85 pb^{-1}.

4.5.8 Single Top

For single-top production, there are three separate processes at leading order to consider. The first is quark-antiquark annihilation, or the s-channel, which is similar to Drell-Yan and results from the exchange of a charged W boson. The second is W boson gluon fusion, or the t-channel, and the third, the tW channel, results from top quark production in association with a W. All single-top MC simulated datasets were generated using MadGraph.

s-channel

The s-channel dataset consists of 494,967 events, and a cross-section of 4.21 pb, for an $\mathcal{L}_{int}$ of 117,569 pb^{-1}.

t-channel

The t-channel dataset consists of 484,060 events, and a cross-section of 20.93 pb, for an $\mathcal{L}_{int}$ of 23,127.57 pb^{-1}.

tW-channel

The tW-channel dataset consists of 494,961 events, and a cross-section of 10.56 pb. As with $t\bar{t}$, a MadGraph correction factor of 0.985608 must be applied, for an $\mathcal{L}_{int}$ of 46,196.73 pb^{-1}.

[1]Generated with MadGraph
[2]Generated with PYTHIA

Data Type	N_{events}	σ (pb)	Filter Eff.	MadGraph Corr.	$\mathcal{L}$ (pb^{-1})
[1]W$\rightarrow \mu\nu_{\mu}$+jets (madgraph)	15123740	31314	1.000		483.0
[2]W$\rightarrow \mu\nu_{\mu}$+jets (pythia)	5330940	7899	1.000		674.9
[2]QCD	29504866	296600000	0.0002855		348.4
[2]W$\rightarrow \tau\nu_{\tau}$ (pythia)	5221750	7899	1.000		661.1
[2]W$\rightarrow \tau\nu_{\tau}$ (powheg)	3990741	9719	1.000		410.6
[2]Z$\rightarrow \mu\mu$	2289913	1300	1.000		1761.5
[2]Z$\rightarrow \tau\tau$	2057446	1300	1.000		1582.7
[1]$t\bar{t}$	1165716	157.5	1.000	0.985608	7294.9
[1]s-channel top	494967	4.21	1.000		115569.0
[1]t-channel top	484060	20.93	1.000		23127.6
[1]tW-channel top	494961	10.56	1.000	0.985608	46196.7

Table 4.3: Cross sections and event multiplicities for MC sets used in this analysis.

4.6 Event Selection

Achieving a useful result in any analysis is predicated upon successful estimation and extraction of the signal from any background. Usually, a fit is performed to a distinguishing quantity in data. However, in order to minimize the systematic error associated with fitting a function to the signal and background, a series of cuts must be made on the dataset so that as much background as possible is eliminated before a fit is performed. Since the signal and background PDFs cannot be known *a priori*, their shape and overall yield must be inferred from Monte Carlo samples.

4.6.1 Trigger

Although different high-level trigger paths were used in the data-taking regime, depending on luminosity and HLT prescaling, the signal MC sample relied upon the `HLT_Mu9` trigger path (N.B. the `HLT_Mu9` trigger path accepts events with at least one muon with $p_T \geq 9$ GeV/c and $\eta < 2.1$). Events which have passed the HLT selection criteria are considered to the baseline dataset for this study.

4.6.2 Muon Selection

Despite the overall high performance of the muon detector, it is still possible for spurious muon candidates or muons from other events to be associated with a W$\rightarrow \mu\nu_\mu$ event. To that end, a series of selection cuts are imposed on each muon candidate.

- The muon is required to be identified as both a global muon and a tracker muon. This cut is effective at rejecting muons originating from decays in flight, punch-through, and accidental wrong matchings with the tracker (in the case of global muons) or with noisy segments (in the case of tracker muons). The global muon reconstruction algorithm requires a loose matching with a track in the inner detector, and a tracker muon requires loose matching with segments in the muon chambers. Thus, tracker muon reconstruction is largely immune to alignment issues.

- The number of hits in the pixel detector is required to be greater than zero. This cut further reduces muons originating from decays in flight.

- The number of hits in the tracker is required to be greater than ten. This is because the p_T resolution is very poor with a low number of hits.

- The transverse impact parameter of the muon with respect to the beam spot is required to be less than two millimeters. This is a loose cut designed to reject muons originating from cosmic rays.

- The χ^2 per number of degrees of freedom of the global muon fit is required to be less than ten. This is a simple sanity check on the quality of the fitted track.

- At least one valid hit in the muon chambers used in the global muon fit. This is a further cut discriminating against decays in flight and high-p_T punch-through. Together with the χ^2 requirement, it is known as the GlobalMuonPromptTight requirement.

- At least two muon stations are required. This cut is particularly effective against punch-through and accidental matchings, and is also consistent with the CMS muon trigger logic, which requires at least two muon stations in order to give a meaningful p_T estimate.

- An event is rejected if it contains at least two global muons, with the global muon with the highest transverse momentum satisfying $p_T > 20$ GeV, and the global muon with the second-highest transverse momentum satisfying $p_T > 10$ GeV. This cut is designed to reject Drell-Yan background.

- The muon is required to have a p_T of at least 25 GeV/c as well as $|\eta| < 2.1$ (see Figure 4.13).

4.6.3 Isolation

Heavy quarks in jets can decay into muons and deposit energy in the calorimeters in the process, thus muons inside jets are a strong indication of QCD background processes rather than signal. Thus, isolation is a powerful tool for removing QCD background, as muons originating from the interaction point from an electroweak process tend to be isolated from other tracks and calorimeter deposits, unlike muons originating from hadronic jets (Figure 4.14). Isolation is defined as the sum of all of the transverse energy in a cone around the muon candidate (either by summing the energy deposited in the calorimeter or the tracks in the tracker system). For this analysis, so-called relative combined isolation (defined as the sum of the p_T, the electromagnetic E_T and the hadronic E_T, all divided by the p_T of the muon) was used. The cut imposed was $iso_{rel} = (iso_{trk} + iso_{calo})\,/p_T < 0.10$. The cone specified was $\Delta R < 0.3$, where $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$.

4.6.4 Jet Selection

As mentioned in Section 4.3.4, jet reconstruction is done using four different methods: calorimeter jets, jet-plus-tracks jets, particle flow jets, and track jets, with five separate clustering algorithms: iterative cone, midpoint cone, seedless infrared cone, k_T and anti-k_T. For this analysis, particle-flow jet reconstruction was used, as was the anti-k_T algorithm with $d_{ij} < 0.5$. As with muons, a series of selection cuts and corrections are made to weed out background. First, the so-called L2 and L3 corrections are applied to all jets, where L2 denotes flattening the jet response as a function of pseudorapidity and L3 denotes flattening jet response as a function of transverse momentum. Jets from the actual collision data are further corrected with the "residual" correction, which normalizes the data jet energy response to MC truth. Jets occurring within a cone of $dR < 0.5$ of the muon are removed as well. Jets are required to pass the so-called "loose" identification criteria, which consist of:

- $p_T > 30$ GeV and $|\eta| < 2.4$
- The neutral hadron fraction is required to be less than 0.99.
- The neutral electromagnetic fraction is required to be less than 0.99.
- The number of jet constituents is required to be at least 2.

- The following are required for $|\eta| < 2.4$:
    - The charged hadron fraction is required to be greater than 0.
    - The multiplicity of charged jet components is required to be greater than 0.
    - The charged electromagnetic fraction is required to be less than 0.99.

The "loose" identification criteria, when run on a 7 TeV QCD Dijet MC sample, yielded a selection efficiency of 99.82%, with a fake rate of 7.97%, a low quality rate of 21.62%, and a high quality rate of 70.40%[7].

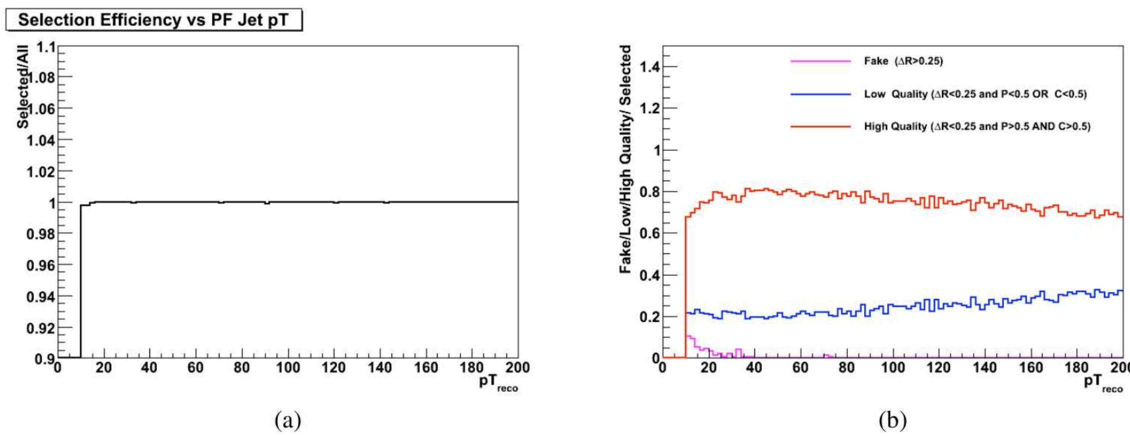


Figure 4.9: 7 TeV MC efficiency of reconstructed jets for "loose" ID. (a) Selection efficiency vs p_T. (b) Fake, low quality, high-quality rate vs p_T.[7]

Another source of background noise in any jet-based analysis is pile-up, which is when jets from a different collision are associated with a W$\rightarrow$ $\mu\nu_\mu$ event. To reduce pile-up, the Fastjet correction is performed, in which a pile-up energy density is calculated as $\rho =$ median $(p_t^{jet}/$ Area(jet)), assumed to be constant, and subtracted as an offset.

4.6.5 Transverse Mass and Missing Transverse Energy

Transverse mass is shown in Figure 4.11, and missing transverse energy is shown in Figure 4.12. For this analysis, the cut $m_T > 50$ GeV was imposed. No cut was imposed on the $\not{E}_T$.

Selection Criterion	Data # of Events	Data Acceptance	Signal MC # of Events	Signal MC Acceptance
Total	4.51113×10^7	100.0%	4.69125×10^6	100.0%
W Cands	4.41285×10^7	$97.8 \pm 0.0\%$	3.43753×10^6	$73.3 \pm 0.0\%$
HLT	7.60416×10^6	$16.9 \pm 0.0\%$	2.85449×10^6	$60.8 \pm 0.0\%$
Z Veto	7.56457×10^6	$16.8 \pm 0.0\%$	2.85223×10^6	$60.8 \pm 0.0\%$
Muon ID	6.36615×10^6	$14.1 \pm 0.0\%$	2.80285×10^6	$59.7 \pm 0.0\%$
Eta	6.07701×10^6	$13.5 \pm 0.0\%$	2.65194×10^6	$56.5 \pm 0.0\%$
Pt	834437	$1.8 \pm 0.0\%$	2.23517×10^6	$47.6 \pm 0.0\%$
Acop	834437	$1.8 \pm 0.0\%$	2.23517×10^6	$47.6 \pm 0.0\%$
Isolation	834437	$1.8 \pm 0.0\%$	2.19899×10^6	$46.9 \pm 0.0\%$
MET/MT	142530	$0.3 \pm 0.0\%$	2.02487×10^6	$38.9 \pm 0.0\%$

Table 4.4: Data acceptance rates and corresponding MC efficiences for W$\rightarrow \mu\nu_\mu$ selection cuts.

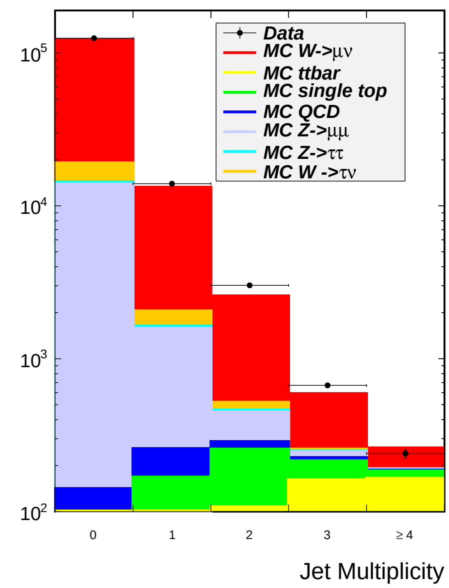


Figure 4.10: Comparison of data and MC signal/background jet multiplicity rates, in bins of 0, 1, 2, 3, and 4 or more jets. The MC datasets have been normalized to the data luminosity and all selection cuts have been made.

4.7 Methodology

For this analysis, the method used to remove lingering background contamination not excised by the aforementioned cuts is known as sideband subtraction, and consists of subtracting the number of luminosity-normalized MC positively-charged and negatively charged events from the total number of data events surviving the selection cuts, then computing the charge asymmetry using Equation 4.1. The number of the dataset and each renormalized MC quantity, including statistical uncertainty in selection, is summed up in Tables 4.5 and 4.6. Statistical errors were computed using the square of the number of events; however, the ROOT `Sumw2()` method was used to compute the statistical errors as the sum of individual weights. In the absence of any further histogram modification such as adding/subtracting histograms, this returns the same error value as simply calculating the square root of the number of events. In the case of combining or rescaling histograms, on the other hand, it

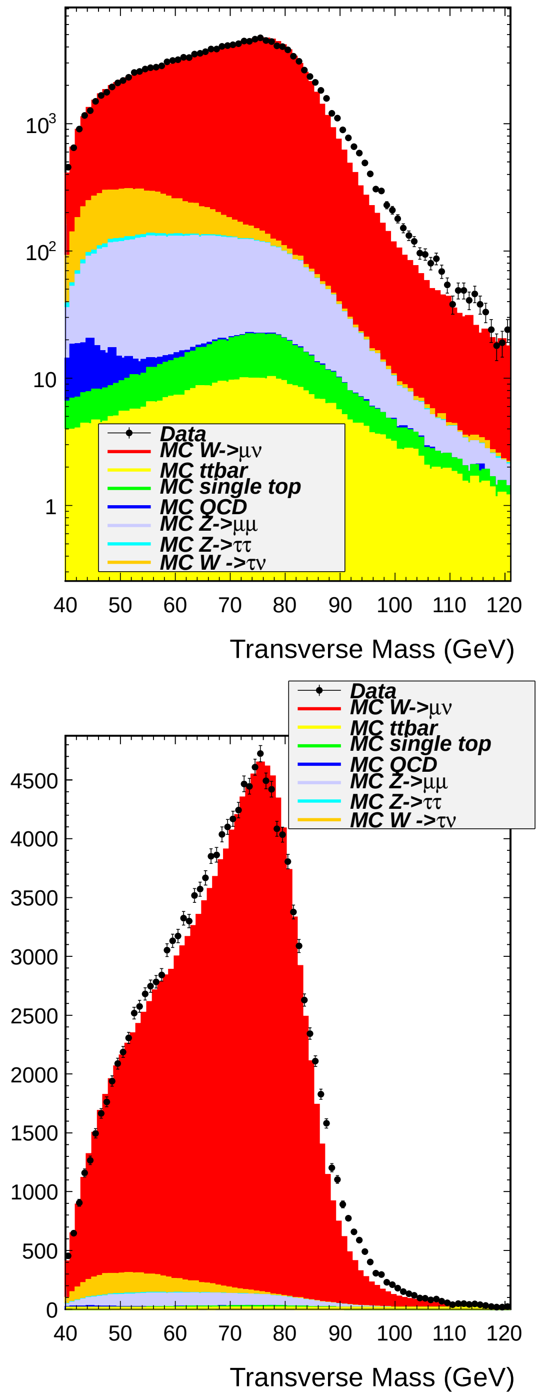


Figure 4.11: Transverse mass for data and normalized MC, after all selection cuts have been made. The plot on the top uses a semi-log scale, and the plot on the bottom is linear.

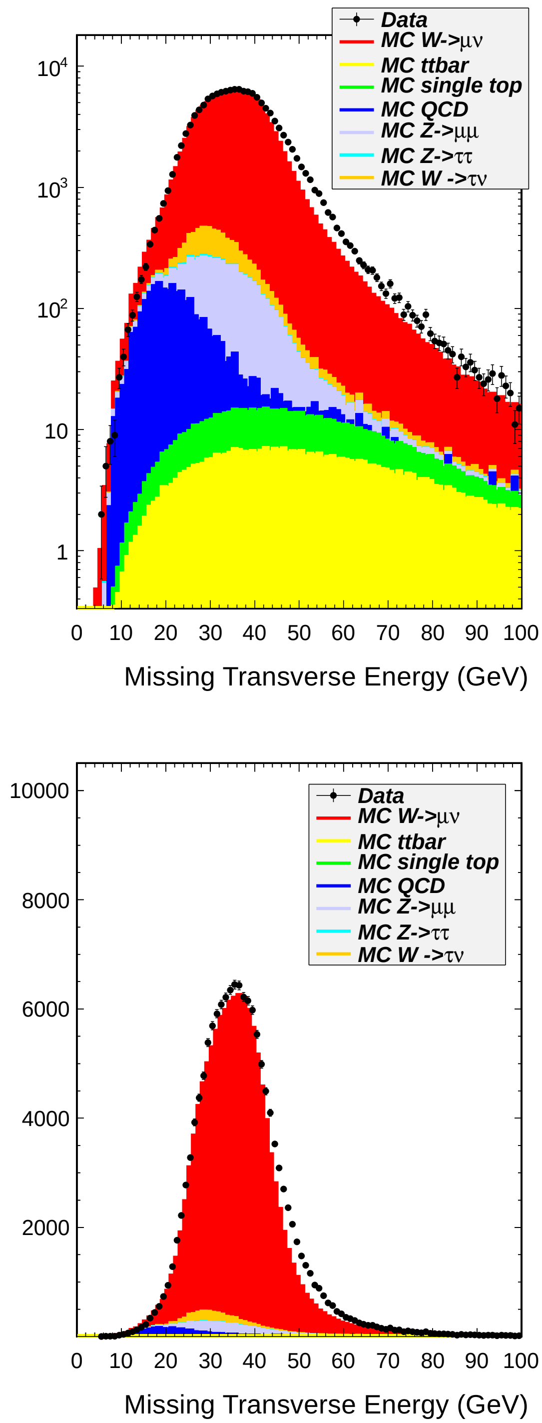


Figure 4.12: Missing transverse energy for data and normalized MC, after all selection cuts have been made. The plot on the top uses a semi-log scale, and the plot on the bottom is linear.

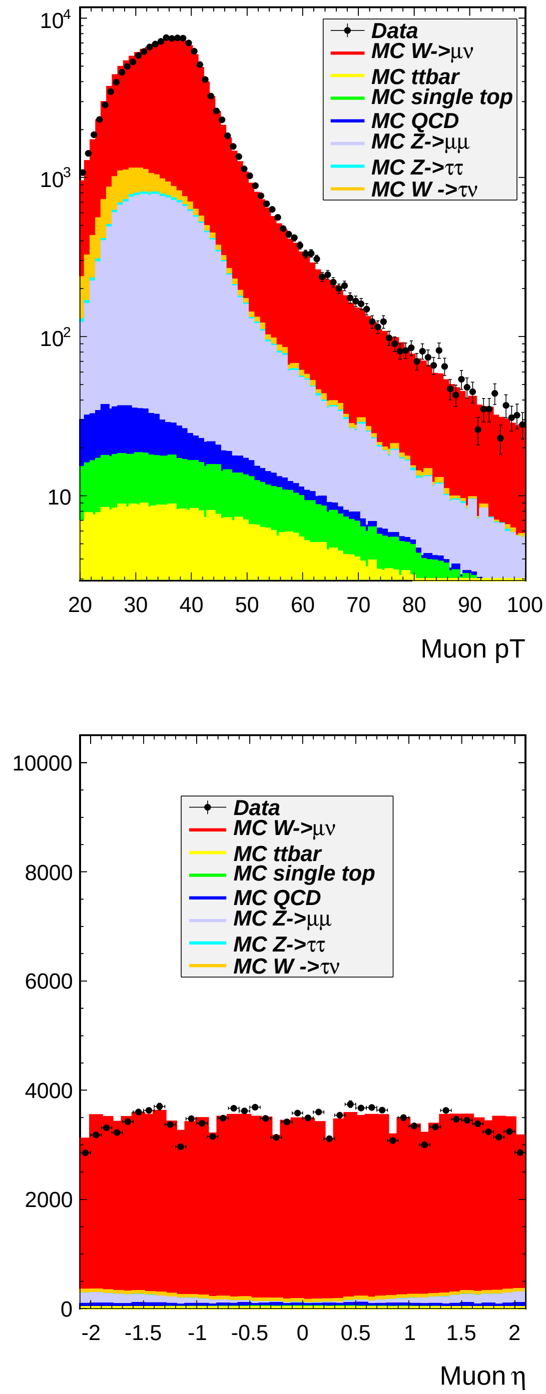


Figure 4.13: Muon transverse momentum (top) and pseudorapidity (bottom) for data and normalized MC. All selection cuts have been made.

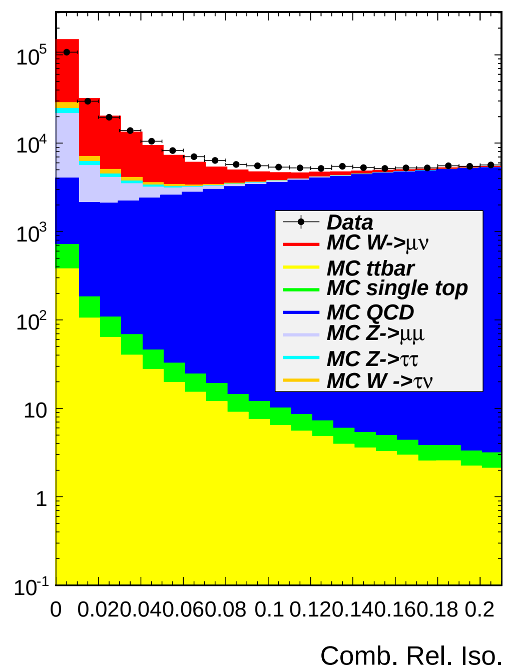


Figure 4.14: Combined relative isolation ($iso_{rel} = (iso_{trk} + iso_{calo})/p_T$) for data and normalized MC. All selection cuts have been made, except for the requirement that $iso_{rel} < 0.10$.

normalizes errors accordingly. For systematic errors, the sources of systematic error were computed separately and added in quadrature. For the cross section measurement, the overall uncertainties were derived using various methods (detailed in the following chapter) and the overall percentage of uncertainty was applied. For the asymmetry calculations, due to the fact that the asymmetry is in effect a cross section ratio measurement, it is expected that some amount of systematic error would cancel out. Therefore, the systematic uncertainty was calculated by varying the relevant quantities (such as luminosity, PDF uncertainty, and so on) by the derived uncertainty, and calculating the asymmetry at the upper and lower boundaries, thus deriving a percentage error to be added in quadrature.

One further source of uncertainty which is interesting in its own right due to its domination in W decays with associated jet production is the uncertainty in the jet energy scale due in large part to initial-state and final-state gluon radiation. To that end, the charge asymmetry was calculated as a function of jet multiplicity, with the uncertainty derived and applied separately from the cross-section and pseudorapidity asymmetry calculations.

Dataset	$n_{jets} = 0$	$n_{jets} = 1$	$n_{jets} = 2$	$n_{jets} = 3$	$n_{jets} \geq 4$
Data	124630.0 $\pm$ 353.0	13971.0 $\pm$ 118.2	3020.0 $\pm$ 55.0	669.0 $\pm$ 25.9	240.0 $\pm$ 15.5
MC W$\to \mu\nu$	118260.0 $\pm$ 88.4	12773.0 $\pm$ 29.1	2349.7 $\pm$ 12.5	384.4 $\pm$ 5.0	80.4 $\pm$ 2.3
$t\bar{t}$	2.8 $\pm$ 0.1	29.9 $\pm$ 0.4	107.4 $\pm$ 0.7	161.6 $\pm$ 0.9	165.4 $\pm$ 0.9
s-channel single top	1.5 $\pm$ 0.0	6.7 $\pm$ 0.0	8.8 $\pm$ 0.1	2.4 $\pm$ 0.0	0.5 $\pm$ 0.0
t-channel single top	29.3 $\pm$ 0.4	125.0 $\pm$ 0.8	126.9 $\pm$ 0.8	41.1 $\pm$ 0.4	14.4 $\pm$ 0.3
tW-channel single top	1.1 $\pm$ 0.0	6.7 $\pm$ 0.1	13.1 $\pm$ 0.1	9.7 $\pm$ 0.1	4.2 $\pm$ 0.1
QCD	122.3 $\pm$ 1.5	175.0 $\pm$ 1.8	63.7 $\pm$ 1.1	18.5 $\pm$ 0.6	6.9 $\pm$ 0.4
Drell-Yan $\mu^+\mu^-$	3586.1 $\pm$ 8.5	342.8 $\pm$ 2.6	42.2 $\pm$ 0.9	5.5 $\pm$ 0.3	0.8 $\pm$ 0.1
Drell-Yan $\tau^+\tau^-$	115.3 $\pm$ 1.6	16.0 $\pm$ 0.6	3.0 $\pm$ 0.3	0.6 $\pm$ 0.1	0.1 $\pm$ 0.1
W$\to \tau\nu$	2548.5 $\pm$ 11.8	225.1 $\pm$ 3.5	31.3 $\pm$ 1.3	3.9 $\pm$ 0.5	0.7 $\pm$ 0.2

Table 4.5: Expected number of events, binned in exclusive jet multiplicity, for relevant subprocesses after normalizing to the 2010 integrated luminosity of 35.9 pb^{-1}. The expected error is statistical only.

Dataset	$0 \leq$	η	< 0.8	$0.8 \leq$	η	< 1.6	$1.6 \leq$	η	< 2.1
Data	56567.0	$\pm$	237.8	54098.0	$\pm$	232.6	31865.0	$\pm$	178.5
MC W$\rightarrow \mu\nu$	52349.2	$\pm$	58.8	50600.0	$\pm$	57.8	30898.6	$\pm$	45.2
$t\bar{t}$	252.6	$\pm$	1.1	163.5	$\pm$	0.9	51.0	$\pm$	0.5
s-channel single top	8.7	$\pm$	0.1	7.6	$\pm$	0.0	3.6	$\pm$	0.0
t-channel single top	170.1	$\pm$	0.9	123.3	$\pm$	0.8	43.5	$\pm$	0.5
tW-channel single top	19.2	$\pm$	0.1	12.0	$\pm$	0.1	3.6	$\pm$	0.1
QCD	213.2	$\pm$	2.0	125.4	$\pm$	1.6	47.9	$\pm$	1.0
Drell-Yan $\mu^+\mu^-$	485.3	$\pm$	3.1	1742.8	$\pm$	6.0	1749.3	$\pm$	6.0
Drell-Yan $\tau^+\tau^-$	48.2	$\pm$	1.0	52.1	$\pm$	1.1	34.8	$\pm$	0.9
W$\rightarrow \tau\nu$	1056.1	$\pm$	7.6	1078.6	$\pm$	7.7	674.8	$\pm$	6.1

Table 4.6: Expected number of events for relevant subprocesses, binned in muon pseudorapidity, after normalizing to the 2010 integrated luminosity of 35.9 pb^{-1}. The expected error is statistical only.

Chapter 5
Results and Uncertainties

In this section, the results for the W$\rightarrow$ $\mu\nu_\mu$ charge asymmetry as a function of muon pseudorapidity (in bins of $0 \leq |\eta| < 0.8$, $0.8 \leq |\eta| < 1.6$, and $1.6 \leq |\eta| < 2.1$) and as a function of jet multiplicity (in both exclusive and inclusive bins consisting of zero, one, two, three and four or more jets) will be presented, as will a calculation of the overall cross section of the process $\sigma_{PP\rightarrow WX} \times \mathrm{BR}\,(\mathrm{W} \rightarrow \mu\nu_\mu)$. Also, major sources of systematic uncertainties for each measurement will be discussed.

5.1 W Cross Section

The total yield of W$\rightarrow$ $\mu\nu_\mu$ events present in the 2010 dataset can be inferred from Table 4.5, as can the overall yield of the combined background. Then, out of the 142,530 total events in data, approximately 94.0$\pm$0.3% are actual signal events, yielding a W$\rightarrow$ $\mu\nu_\mu$ signal yield of approximately 133,987 $\pm$ 377.8 events. Then, the true number of W$\rightarrow$ $\mu\nu_\mu$ signal events occurring in the detector (rather than just the events identified) can be estimated by using

$$N_{total} = \frac{N_{detected}}{\epsilon}, \tag{5.1}$$

where ϵ is the overall candidate selection efficiency calculated in Section 4.6. From there, the number of W$\rightarrow$ $l\nu_l$ events analyzed is 4.69125×10^6 and the number passing the final MET/MT cut is 1.82737×10^6, leading to an acceptance figure of 38.9%. Then, using

$\sigma_{PP\rightarrow WX} \times \mathrm{BR}\,(\mathrm{W} \rightarrow \mu\nu)$ Cross Section Uncertainties	
PDF and Theoretical	$\pm 1.1\%$
Muon Momentum Scale/Resolution	$\pm 2.0\%$
Pre-triggering	$\pm 0.5\%$
Systematic Uncertainty	$\pm 2.3\%$

Table 5.1: Summary of systematic uncertainties for the cross section calculation.

Equation 3.1, the cross section (including statistical and systematic uncertainty) is

$$\sigma_{PP\rightarrow WX} \times \mathrm{BR}\,(\mathrm{W} \rightarrow \mu\nu_\mu) = 9580 \pm 27\ (\text{stat.}) \pm 220\ (\text{syst.}) \pm 1054\ (\text{lum.})\ \text{pb}. \tag{5.2}$$

The NNLO calculation of $\sigma_{PP\rightarrow W} \times \mathrm{BR}\,(\mathrm{W} \rightarrow \mu\nu_\mu)$ yields a cross-section of 10440 ± 520 pb. The W^+ and W^- cross sections were calculated independently; their values are

$$\sigma_{PP\rightarrow WX} \times \mathrm{BR}\left(\mathrm{W}^+ \rightarrow \mu^+\nu_\mu\right) = 5730 \pm 21\ (\text{stat.}) \pm 132\ (\text{syst.}) \pm 630\ (\text{lum.})\ \text{pb}. \tag{5.3}$$

and

$$\sigma_{PP\rightarrow WX} \times \mathrm{BR}\left(\mathrm{W}^- \rightarrow \mu^-\nu_\mu\right) = 3850 \pm 17\ (\text{stat.}) \pm 89\ (\text{syst.}) \pm 424\ (\text{lum.})\ \text{pb}., \tag{5.4}$$

respectively. The NNLO predictions are 6150 ± 290 pb for W^+ and 4290 ± 230 pb for W^-.

5.2 W Asymmetry

As mentioned in Section 4.7, the procedure for measuring charge asymmetry consists of subtracting out luminosity-normalized background yields, then calculating asymmetry (defined in Equation 4.1) and associated statistical uncertainties, and finally calculating the contributions of the major sources of systematic uncertainties. To begin, the primary result is W boson charge asymmetry as a function of muon pseudorapidity. As a further check, charge asymmetry as a function of jet multiplicity for both inclusive[15] and exclusive[16] jet

[15]Inclusive in this case is defined as all jet multiplicities greater than or equal to N, where N is one of the bins of jet multiplicity. For this analysis, the binned jet multiplicities are 0 (or more) jets, 1 (or more) jets, 2 (or more) jets, 3 (or more) jets, and 4 or more jets.

[16]Exclusive, as opposed to inclusive, simply means that each bin contains W events with exactly N jets. As with the inclusive case, the bins are exactly 0, 1, 2 or 3 jets; as well as a separate bin for 4 or more. Statistics are low enough for the $N \geq 4$ case that they are grouped together, even in the exclusive case.

	$0 \leq \lvert\eta\rvert < 0.8$	$0.8 \leq \lvert\eta\rvert < 1.6$	$1.6 \leq \lvert\eta\rvert < 2.1$
Luminosity	$\pm 0.2\%$	$\pm 1.2\%$	$^{+2.5}_{-2.4}\%$
PDF and Theoretical	$\pm 3.5\%$	$\pm 2.8\%$	$\pm 2.6\%$
Muon Momentum Scale/Resolution	$\pm 2.0\%$	$\pm 2.0\%$	$\pm 2.0\%$
Muon Efficiency	$\pm 1.9\%$	$\pm 2.9\%$	$\pm 4.0\%$
Systematic Uncertainty	$\pm 4.2\%$	$\pm 4.4\%$	$^{+5.6}_{-5.5}\%$

Table 5.2: Summary of systematic uncertainties for charge asymmetry as a function of muon pseudorapidity.

multiplicities was calculated, with separate sources of systematic uncertainties (see Figure 5.2).

The sources of systematic errors and their relative contributions to the overall charge asymmetry as a function of muon pseudorapidity are summarized in Table 5.2, and the plot of charge asymmetry versus muon $|\eta|$ is shown in Figure 5.1.

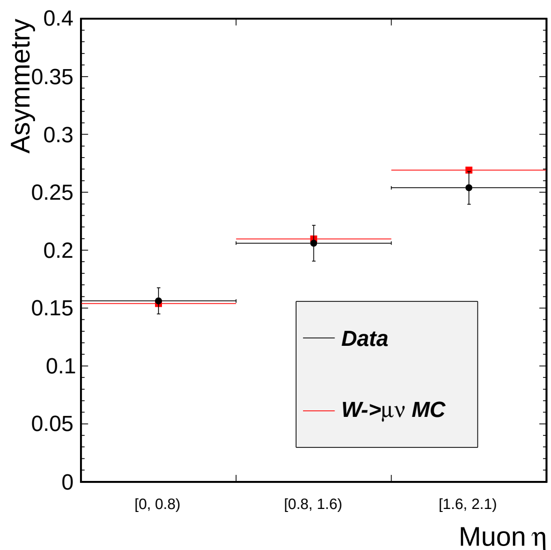
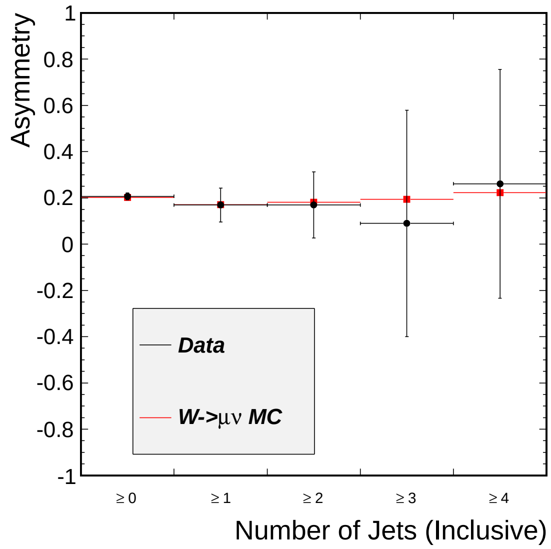


Figure 5.1: Charge asymmetry as a function of muon pseudorapidity for 2010 data and signal W$\rightarrow \mu\nu_\mu$ MC. The error bars represent statistical uncertainty $\oplus$ systematic uncertainty.

The overall systematic uncertainties and their contribution to the exclusive and inclusive



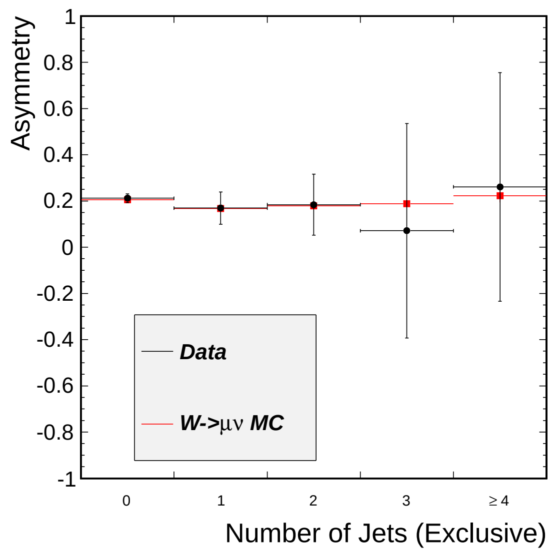


Figure 5.2: Charge asymmetry versus (top) inclusive jet multiplicity and versus (bottom) exclusive jet multiplicity for data and normalized MC. The error bars represent statistical uncertainty $\oplus$ systematic uncertainty.

Exclusive	$N_{jets} = 0$	$N_{jets} = 1$	$N_{jets} = 2$	$N_{jets} = 3$	$N_{jets} \geq 4$
Luminosity	$\pm 1.0\%$	$\pm 0.7\%$	$\pm 0.8\%$	$^{+6.2}_{-5.9}\%$	$^{+70.3}_{-29.5}\%$
PDF and Theoretical	$\pm 2.8\%$	$\pm 2.4\%$	$\pm 2.1\%$	$^{+9.6}_{-9.2}\%$	$^{+113.8}_{-36.0}\%$
Jet Energy Scale/Resolution	$\pm 3.3\%$	$^{+26.2}_{-27.5}\%$	$^{+36.5}_{-33.3}\%$	$^{+123.5}_{-133.0}\%$	$^{+134.6}_{-138.5}\%$
Jet Counting	$\pm 0.3\%$	$\pm 0.6\%$	$\pm 3.2\%$	$\pm 2.8\%$	$\pm 3.3\%$
Systematic Uncertainty	$\pm 4.4\%$	$^{+26.3}_{-27.6}\%$	$^{+36.7}_{-33.5}\%$	$^{+124.1}_{-133.5}\%$	$^{+189.8}_{-146.1}\%$

Inclusive	$N_{jets} \geq 0$	$N_{jets} \geq 1$	$N_{jets} \geq 2$	$N_{jets} \geq 3$	$N_{jets} \geq 4$
Luminosity	$\pm 1.0\%$	$\pm 0.9\%$	$\pm 1.8\%$	$^{+9.6}_{-8.4}\%$	$^{+70.3}_{-29.5}\%$
PDF and Theoretical	$\pm 2.7\%$	$\pm 2.4\%$	$\pm 3.3\%$	$^{+13.1}_{-11.1}\%$	$^{+113.8}_{-36.0}\%$
Jet Energy Scale/Resolution	$\pm 0.0\%$	$^{+29.7}_{-30.3}\%$	$^{+43.9}_{-41.7}\%$	$^{+123.7}_{-133.5}\%$	$^{+134.6}_{-138.5}\%$
Jet Counting	$\pm 0.0\%$	$\pm 0.6\%$	$\pm 3.2\%$	$\pm 2.8\%$	$\pm 3.3\%$
Systematic Uncertainty	$\pm 2.9\%$	$^{+29.8}_{-30.4}\%$	$^{+44.2}_{-42.0}\%$	$^{+124.8}_{-134.3}\%$	$^{+189.8}_{-146.1}\%$

Table 5.3: Overall summary of systematic uncertainties for charge asymmetry as a function of (top) exclusive and (bottom) inclusive jet multiplicity.

jet multiplicity asymmetries are summed up in Table 5.3.

Dataset	$0 \leq \eta < 0.8$	$0.8 \leq \eta < 1.6$	$1.6 \leq \eta < 2.1$
Asymmetry	0.156191	0.206045	0.253892
Statistical Unc.	$\pm 2.8\%$	$\pm 2.3\%$	$\pm 2.5\%$
Systematic Unc.	$\pm 2.8\%$	$\pm 3.9\%$	$\pm 5.3\%$
MC Asymmetry	0.153904	0.209642	0.269184

Table 5.4: Overall summary of systematic uncertainties for charge asymmetry as a function of muon pseudorapidity.

	$N_{jets} = 0$	$N_{jets} = 1$	$N_{jets} = 2$	$N_{jets} = 3$	$N_{jets} \geq 4$
Asymmetry	0.198005	0.161055	0.178672	0.0737830	0.242528
Statistical Unc.	$\pm 1.5\%$	$\pm 5.7\%$	$\pm 11.9\%$	$\pm 82.3\%$	$\pm 140.2\%$
Systematic Unc.	$\pm 4.4\%$	$^{+24.8}_{-25.8}\%$	$^{+35.6}_{-31.6}\%$	$^{+126.0}_{-137.8}\%$	$^{+197.7}_{-142.3}\%$
MC Asymmetry	0.205722	0.167726	0.179115	0.187618	0.222679

	$N_{jets} \geq 0$	$N_{jets} \geq 1$	$N_{jets} \geq 2$	$N_{jets} \geq 3$	$N_{jets} \geq 4$
Asymmetry	0.193663	0.161854	0.165221	0.0905569	0.242528
Statistical Unc.	$\pm 1.5\%$	$\pm 5.2\%$	$\pm 12.4\%$	$\pm 70.8\%$	$\pm 140.2\%$
Systematic Unc.	$\pm 3.1\%$	$^{+28.3}_{-28.6}\%$	$^{+43.0}_{-40.2}\%$	$^{+125.0}_{-136.3}\%$	$^{+197.7}_{-142.3}\%$
MC Asymmetry	0.201587	0.170217	0.181521	0.193686	0.222679

Table 5.5: Summary of charge asymmetry, including systematic and statistical errors and MC prediction, as a function of (top) exclusive and (bottom) inclusive jet multiplicity.

The overall measurement for the charge asymmetry as a function of muon pseudorapidity is presented in Table 5.4, as is the expected value from simulated data. The same measurement with associated jets, both exclusive and inclusive, is presented in Table 5.5. As with the measurement as a function of muon η, values from Monte Carlo are provided for comparison.

5.3 Systematic Uncertainties

In addition to statistical uncertainties in the result; that is, uncertainty resulting from only having finitely many data points, a major source of uncertainty in any experimental measurement is systematic uncertainty; that is, uncertainties stemming from theoretical uncertainties, as well as from inherent inaccuracies in event reconstruction efficiencies, background estimation, luminosity measurement, and energy and momentum scale and resolution.

5.3.1 Luminosity

Luminosity uncertainty is the largest absolute component of uncertainty relevant to this analysis. To measure the contribution of luminosity uncertainty to the asymmetry measurement, the normalization for each of the Monte Carlo samples was varied by 1.11×35.9 pb^{-1} and 0.89×35.9 pb^{-1}. The results are summed up bin-by-bin in Table 5.6, and in Figures 5.3 and 5.4.

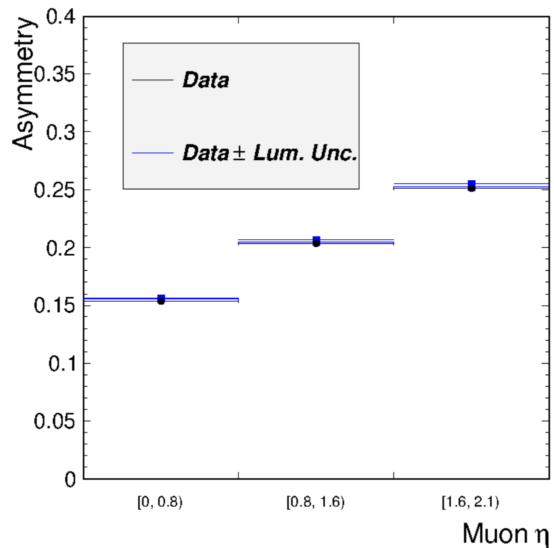


Figure 5.3: Charge asymmetry with luminosity uncertainties for muon η binning.

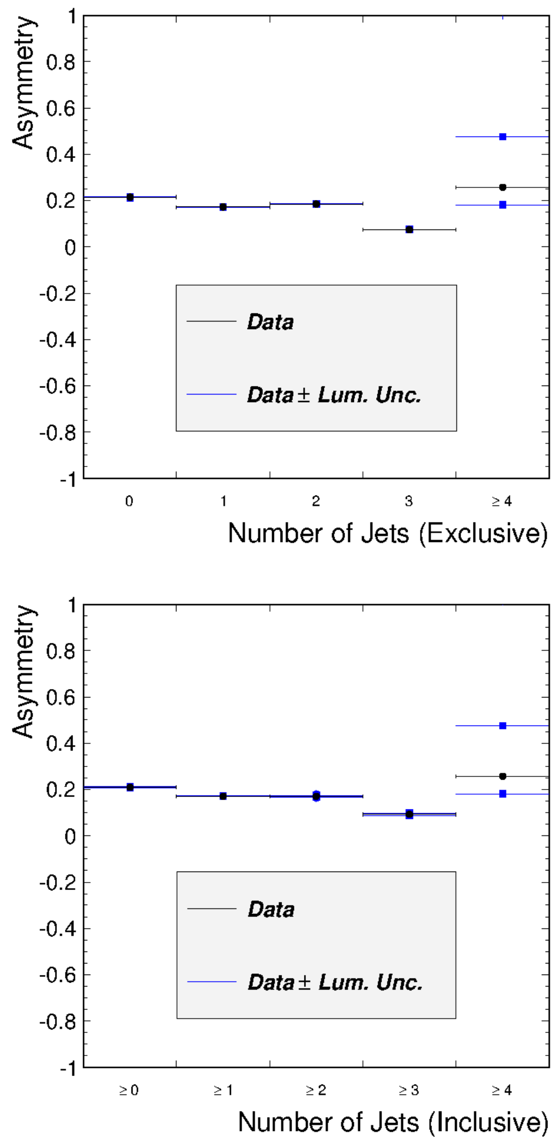


Figure 5.4: Charge asymmetry with luminosity uncertainties for (top) exclusive jet binning and (bottom) inclusive jet binning.

Exclusive	$N_{jets}=0$	$N_{jets}=1$	$N_{jets}=2$	$N_{jets}=3$	$N_{jets}\geq 4$
$t\bar{t}$	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.5\%$	$^{+4.7}_{-4.4}\%$	$^{+63.8}_{-28.2}\%$
s-channel top	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.2\%$	$\pm 0.0\%$
t-channel top	$\pm 0.0\%$	$\pm 0.1\%$	$\pm 0.4\%$	$\pm 3.4\%$	$\pm 1.1\%$
tW-channel top	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.1\%$	$\pm 0.3\%$	$\pm 1.1\%$
QCD	$\pm 0.0\%$	$\pm 0.1\%$	$\pm 0.3\%$	$\pm 0.2\%$	$\pm 1.4\%$
DY $\mu\mu$	$\pm 0.2\%$	$\pm 0.2\%$	$\pm 0.1\%$	$\pm 0.1\%$	$\pm 0.3\%$
DY $\tau\tau$	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.0\%$
W$\rightarrow \tau\nu_\tau$	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.5\%$	$\pm 0.2\%$
Uncertainty	$\pm 0.2\%$	$\pm 0.2\%$	$\pm 0.7\%$	$^{+5.8}_{-5.6}\%$	$^{+63.8}_{-28.3}\%$

Inclusive	$N_{jets}\geq 0$	$N_{jets}\geq 1$	$N_{jets}\geq 2$	$N_{jets}\geq 3$	$N_{jets}\geq 4$
$t\bar{t}$	$\pm 0.0\%$	$\pm 0.3\%$	$\pm 1.6\%$	$^{+8.6}_{-7.4}\%$	$^{+63.8}_{-28.2}\%$
s-channel top	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.2\%$	$\pm 0.0\%$
t-channel top	$\pm 0.0\%$	$\pm 0.2\%$	$\pm 0.5\%$	$^{+3.2}_{-3.1}\%$	$\pm 1.1\%$
tW-channel top	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.1\%$	$\pm 0.4\%$	$\pm 1.1\%$
QCD	$\pm 0.0\%$	$\pm 0.2\%$	$\pm 0.3\%$	$\pm 0.3\%$	$\pm 1.4\%$
DY $\mu\mu$	$\pm 0.2\%$	$\pm 0.2\%$	$\pm 0.1\%$	$\pm 0.0\%$	$\pm 0.3\%$
DY $\tau\tau$	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.0\%$
W$\rightarrow \tau\nu_\tau$	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.1\%$	$\pm 0.4\%$	$\pm 0.2\%$
Uncertainty	$\pm 0.2\%$	$\pm 0.5\%$	$\pm 1.7\%$	$^{+9.2}_{-8.1}\%$	$^{+63.8}_{-28.3}\%$

	$0\leq\lvert\eta\rvert<0.8$	$0.8\leq\lvert\eta\rvert<1.6$	$1.6\leq\lvert\eta\rvert<2.1$
$t\bar{t}$	$\pm 0.1\%$	$\pm 0.0\%$	$\pm 0.0\%$
s-channel top	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.0\%$
t-channel top	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.0\%$
tW-channel top	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.0\%$
QCD	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.0\%$
DY $\mu\mu$	$\pm 0.0\%$	$\pm 0.3\%$	$\pm 0.5\%$
DY $\tau\tau$	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.0\%$
W$\rightarrow \tau\nu_\tau$	$\pm 0.0\%$	$\pm 0.0\%$	$\pm 0.0\%$
Total Uncertainty	$\pm 0.1\%$	$\pm 0.3\%$	$\pm 0.5\%$

Table 5.6: Bin-by-bin luminosity uncertainty for charge asymmetry as a function of (top) exclusive, (middle) inclusive jet multiplicity, and (bottom) muon η.

5.3.2 PDF Uncertainties

Uncertainties in the parton distribution functions used to generate the MC simulated datasets also have a large effect. To generate the MC simulated datasets, the PDF set `CTEQ6M` was used, and the theoretical predictions were computed to NNLO with the `FEWZ` program. Then, three sources of uncertainty are considered: scale uncertainties, determined by varying the factorization and renormalization scales by a factor of 2 up and down, PDF uncertainties, and α_s uncertainties. The PDF uncertainties were determined by combining the NLO PDF and α_s errors from the MSTW08, CT10 and NNPDF2.0 distributions, followed by adding the NNLO scale uncertainties in quadrature.

To determine PDF uncertainties for CT10, the following formulas are used:

$$\begin{aligned} \Delta X^{+} &= \sqrt{\sum_{i=1}^{N}[\max(X_i^{+} - X_0, X_i^{-} - X_0, 0)]^2} \\ \Delta X^{-} &= \sqrt{\sum_{i=1}^{N}[\max(X_0 - X_i^{+}, X_0 - X_i^{-}, 0)]^2}, \end{aligned} \tag{5.5}$$

where $N = 26$ is the number of parameters used by `CTEQ6M` to calculate the cross section to NLO, X_0 is the reference cross-section, and X_i^+ and X_i^- are the induced up and down variations corresponding to the ith parameter, respectively. Uncertainties in α_s were calculated by considering PDFs for $\alpha_s = 0.116, \ldots, 0.120$ (sampling within the 68% confidence interval around the nominal value of $\alpha_s = 0.118$), with Equation 5.5 applied to the resulting five PDFs.

For MSTW, the PDF+α_s uncertainty was calculated by considering the PDF uncertainties from 5 PDF sets (comprised of 40 error PDFs each), corresponding to $\alpha_s = \alpha_s^0$, $\alpha_s^0 \pm 0.5\sigma$, $\alpha_s^0 \pm \sigma$, where σ is the 68% confidence level on α_s^0 uncertainty, and α_s^0 is the nominal MSTW α_s value. Then, Equation 5.5 is applied to each set, and the final PDF+α_s uncertainty is given by:

$$\begin{aligned} \Delta X^{+} &= \max_{\alpha_s}\{X_{\alpha_s} + \Delta X_{\alpha_s}\} - X_0 \\ \Delta X^{-} &= X_0 - \max_{\alpha_s}\{X_{\alpha_s} - \Delta X_{\alpha_s}\} \end{aligned} \tag{5.6}$$

α_s	0.116	0.117	0.118	0.119	0.120	0.121	0.122
N_{rep}	5	27	72	100	72	27	5

Table 5.7: Number of replicas used for each α_s value to compute overall PDF+α uncertainty for NNPDF.

For NNPDF, the uncertainty is obtained by calculating a set of replicas from NNPDF sets with $\alpha_s = 0.116, \ldots, 0.122$, corresponding to a Gaussian centered around the nominal value $\alpha_s = 0.119$, as shown in Table 5.7.

Then the uncertainty is computed with the following formulas:

$$\begin{aligned} \Delta X^+ &= \sqrt{\frac{1}{N^+ - 1}\sum_{i=1}^{N^+}(X_i - X_0)^2} \\ \Delta X^- &= \sqrt{\frac{1}{N^- - 1}\sum_{j=1}^{N^-}(X_0 - X_j)^2}, \end{aligned} \tag{5.7}$$

where as before X_0 is computed with the nominal value $\alpha_s = 0.119$, i runs over the N^+ replicas with $X_i > X_0$, and j runs over the N^- replicas with $X_j < X_0$.

Finally, the overall PDF+α uncertainty is computed using

$$\begin{aligned} \Delta_{\text{max}} &= \max_i(X_i + \Delta X_i^+) \\ \Delta_{\text{min}} &= \max_i(X_i - \Delta X_i^-) \\ \text{PDF+}\alpha_s\text{ uncertainty} &= \frac{1}{2}(\Delta_{\text{max}} - \Delta_{\text{min}}), \end{aligned} \tag{5.8}$$

where i runs over the sets CT10, MSTW2008, and NNPDF2.0. The overall muonic W acceptances are as follows: $0.4067^{+0.0030}_{-0.0023}\%$ for CT10, $0.4080 \pm 0.0021\%$ for MSTW, and $0.4077^{+0.0022}_{-0.0023}\%$ for NNPDF, for an overall systematic of 0.70%.

5.3.3 QCD and Electroweak Uncertainties

Further contributions to the theoretical uncertainty arise from initial and final state radiation (ISR and FSR, respectively), resummation and NNLO QCD effects, factorization and renormalization scale dependance, and missing electroweak corrections. At the energy scale relevant to vector boson production, $\mathcal{O}(\alpha)$ electroweak effects (hereafter referred to as NLO

EWK) are of similar order to $\mathcal{O}(\alpha_s^2)$ QCD effects (hereafter NNLO QCD). Some of the NLO EWK corrections are partially accounted for, namely QED ISR and FSR, which are generated by PYTHIA using a parton shower approximation. Notably missing are one-loop virtual corrections and photon emission. These are quantified by using the HORACE event generator, which implements W/Z production and leptonic decay to LO QCD and NLO EWK, and furthermore uses the parton shower method to account for more-complicated FSR. The uncertainties due to electroweak corrections are therefore calculated by comparing HORACE output after performing the full suite of corrections versus simply performing FSR corrections.

b-tagging

b-tagging is a jet flavor tagging method used to identify jets originating from a bottom quark. Jets originating from a bottom quark have several unique features:

- Hadrons containing bottom quarks have a large enough lifetime that they can travel some distance before decaying, but a small enough lifetime that they decay inside the detector, unlike light quark hadrons. Thus, modern precise silicon detectors such as those in use at CMS can identify particles which originate from a different place than where the bottom quark was formed (i.e. the interaction point).

- The bottom quark is much more massive than its decay products. Thus, the decay products tend to have much higher p_T, causing b-jets to be wider, have higher multiplicities and invariant masses, and to contain low-energy leptons with momentum perpendicular to the jet. Jets with these properties have a greater likelihood to be b-jets.

However, b-jet tagging is by no means foolproof, and is an ongoing area of study. At CMS, several different b-jet tagging algorithms are in use, each of which is available in High Efficiency (high statistics at the expense of a relatively higher mistag rate) and High Purity (lower statistics, with increased b-jet purity) variants, as well as Loose, Medium and Tight (10%, 1% and 0.1% light hadron mistag rate, respectively). Each algorithm produces a quantity known as a 'discriminator' for each jet, upon which one can cut (varying on whether or not Loose, Medium or Tight variants are specified) to distinguish between b-jets and non b-jets.

- **Track Counting** This is a very simple tag, exploiting the long lifetime of B hadrons. It calculates the signed impact parameter significance of all good tracks, and orders them by decreasing significance. Its b tag discriminator is defined as the significance of the Nth track. For the high efficiency variant, $N = 2$, and for the high purity variant, $N = 3$.

- **Jet Probability** This is a more sophisticated algorithm than Track Counting, also exploiting the long lifetime of B hadrons. Its b-tag discriminator is equal to the negative logarithm of the confidence level that all the tracks in the jet are consistent with originating at the primary vertex. This confidence level is calculated from the signed impact parameter significances of all good tracks, with the resolution function read from an online database.

- **Soft Muon and Soft Electron** These two algorithms tag b-jets by searching for the lepton from a semi-leptonic B decay, which typically has a large relative P_t with respect to the jet axis. Their b-tag discriminators are the output of neural nets based on the lepton's relative P_t, impact parameter significance and a few other variables. For each of these taggers, there exists a simple variant (only taking into account the presence and the p_T of the lepton) and a complex variant (which uses jet quantities, in addition to lepton presence and p_T).

- **Simple Secondary Vertex** These algorithms reconstruct the B-decay vertex using an adaptive vertex finder, and then use variables related to it, such as decay length significance to calculate the b-tag discriminator. It has the advantage that is less sensitive to Tracker misalignment than the other lifetime-based tags. Two versions are provided: simpleSecondaryVertexHighEffBJetTags, and simpleSecondaryVertex-HighPurBJetTags (with increased purity due to a cut on the track multiplicity at the secondary vertex).

- **Combined Secondary Vertex** This sophisticated and complex tag exploits all known variables, which can distinguish b from non-b jets. Its goal is to provide optimal b tag performance, by combining information about impact parameter significance, the secondary vertex and jet kinematics. (Currently lepton information is not included). The variables are combined using a likelihood ratio technique to compute the b tag discriminator.

One method that was attempted for the purpose of controlling $t\bar{t}$ background was using a b-jet veto to remove jets that passed the criteria to be considered a b-jet. The algorithms used were Track Counting High Efficiency, loose cuts (TCHEL), Track Counting High Efficiency, medium cuts (TCHEM), Track Counting High Efficiency, tight cuts (TCHET), Track Counting High Purity, loose cuts (TCHPL), Track Counting High Purity, medium cuts (TCHPM), Track Counting High Purity, tight cuts (TCHPT), Simple Secondary Vertex High Efficiency, loose cuts (SSVHEL), Simple Secondary Vertex High Efficiency, medium cuts (SSVHEM), Simple Secondary Vertex High Efficiency, tight cuts (SSVHET), and Simple Secondary Vertex High Purity, tight cuts (SSVHPT). For each algorithm, the discriminator for each jet in a W candidate passing all cuts was computed, and if the value for the discriminant was lower than that of the "reference" discriminant corresponding to the appropriate cut, the jet was thrown out. However, no algorithm proved to sufficiently reduce asymmetry-contaminating background in the higher-jet regimes to be useful. Table 5.8 contains surviving jet multiplicities for each algorithm for the 2010 dataset.

Jet Energy Scale and Resolution

As mentioned in Section 4.4.2, the determination of the jet energy scale and resolution is subject to uncertainties (summarized here). For the MC truth jet energy correction, the uncertainty was estimated by re-deriving the correction factors for the corrected jets, and found to be within 2%. For the offset correction, the uncertainty stems from additional pile-

B-jet Algorithm	$N_{jets} = 0$	$N_{jets} = 1$	$N_{jets} = 2$	$N_{jets} = 3$	$N_{jets} \geq 4$
TCHEL	80359	3564	664	109	32
	54258	2678	497	86	26
TCHEM	80041	3745	772	129	41
	54284	2799	582	114	36
TCHET	79908	3810	818	140	52
	54179	2836	631	122	47
TCHPL	80292	3622	670	110	34
	54488	2678	515	106	28
TCHPM	79992	3775	783	134	44
	54245	2820	595	118	37
TCHPT	79942	3793	803	142	48
	54205	2830	614	125	41
SSVHEL	80059	3741	756	134	38
	54312	2781	574	113	35
SSVHEM	80020	3759	774	136	39
	54276	2803	585	113	38
SSVHET	79906	3811	817	145	49
	54176	2835	634	123	47
SSVHPT	84557	159	8	3	1
	57678	126	9	2	0

Table 5.8: Jet multiplicities for all jet regimes considered after application of b-jet veto algorithms. For each row, the top number is the number of jets for W+ candidates, and the bottom number is the number of jets for W- candidates.

Exclusive	$N_{jets} = 0$	$N_{jets} = 1$	$N_{jets} = 2$	$N_{jets} = 3$	$N_{jets} \geq 4$
Asymmetry + unc.	0.204575	0.203348	0.243955	0.164914	0.568971
Baseline asymmetry	0.198005	0.161055	0.178672	0.0737830	0.242528
Asymmetry - unc.	0.191419	0.116763	0.119251	-0.0243703	-0.0934302
Uncertainty	$\pm 3.3\%$	${}^{+26.2}_{-27.5}\%$	${}^{+36.5}_{-33.3}\%$	${}^{+123.5}_{-133.0}\%$	${}^{+134.6}_{-138.5}\%$

Inclusive	$N_{jets} \geq 0$	$N_{jets} \geq 1$	$N_{jets} \geq 2$	$N_{jets} \geq 3$	$N_{jets} \geq 4$
Asymmetry + unc.	0.193663	0.209957	0.237693	0.2025930	0.568971
Baseline asymmetry	0.193663	0.161854	0.165221	0.0905569	0.242528
Asymmetry - unc.	0.193663	0.112871	0.096332	-0.0303407	-0.0934302
Uncertainty	$\pm 0.0\%$	${}^{+29.7}_{-30.3}\%$	${}^{+43.9}_{-41.7}\%$	${}^{+123.7}_{-133.5}\%$	${}^{+134.6}_{-138.5}\%$

Table 5.9: Bin-by-bin jet energy scale uncertainty for charge asymmetry as a function of (top) exclusive and (bottom) inclusive jet multiplicity.

up, as well as overlapping pile-up and jet contributions, resulting in a systematic uncertainty of 2%, as well as 2% from data/MC differences. For the relative correction, comparisons between data and MC yield a conservative estimate of uncertainty of $2\% \times |\eta|$. For the absolute correction, the dijet p_T balance method yields an uncertainty of $7.4 \pm 1.7\%$, and the MPF method yielded an uncertainty of $0.8 \pm 1.0\%$.

The derived uncertainties are available in the online database as a function of jet p_T and η, and are calculated individually for each reconstructed jet. To determine their effect on the asymmetry as a function of jet multiplicity, each jet was varied up and down in p_T and η, and the "loose" PF jet criteria were applied in each case. Then, the asymmetry was calculated for both the high and low boundaries for the jet energy correction uncertainty. The jet p_T uncertainty was further varied by $\pm 10\%$ of the jet p_T resolution; so for example if the p_T resolution was 0.15, the asymmetry was further calculated for a resolution of 0.135 and 0.165 and compared. The results can be seen in Figure 5.5 and in Table 5.9.

Jet Counting

To measure the effect on the charge asymmetry on jet miscounting, the charge asymmetry was recalculated using the number of generator-level jets with $p_T^{jet} > 30$ GeV/c and $|\eta| < 2.4$ present in the MC W$\rightarrow \mu\nu_\mu$ sample, and then compared to the asymmetry in the same sample as measured with reconstructed jets. The bin-by-bin contribution to the asymmetry

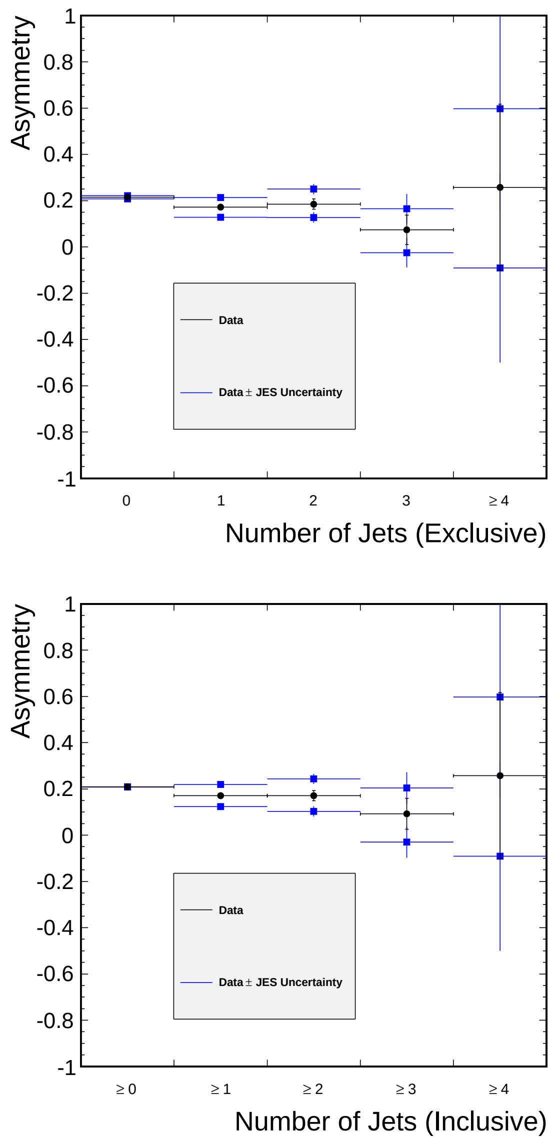


Figure 5.5: Charge asymmetry with jet energy scale and resolution uncertainties for (top) exclusive jet binning and (bottom) inclusive jet binning.

	$N_{jets}=0$	$N_{jets}=1$	$N_{jets}=2$	$N_{jets}=3$	$N_{jets}\geq 4$
Reco n_{jets} Asymmetry	0.205722	0.167726	0.179115	0.187618	0.222679
Truth n_{jets} Asymmetry	0.206326	0.168691	0.184794	0.192814	0.229955
Uncertainty	±0.3%	±0.6%	±3.2%	±2.8%	±3.3%

	$N_{jets}\geq 0$	$N_{jets}\geq 1$	$N_{jets}\geq 2$	$N_{jets}\geq 3$	$N_{jets}\geq 4$
Reco n_{jets} Asymmetry	0.201587	0.170217	0.181521	0.193686	0.222679
Truth n_{jets} Asymmetry	0.201579	0.171946	0.187156	0.198665	0.229955
Uncertainty	±0.0%	±0.6%	±3.2%	±2.8%	±3.3%

Table 5.10: Bin-by-bin jet counting uncertainty for charge asymmetry as a function of (top) exclusive and (bottom) inclusive jet multiplicity.

Quantity	ISR+NNLO	$>$ NNLO	PDF	FSR	EWK	Total
W$\rightarrow\mu\nu_\mu$ Acceptance	0.65%	0.44%	0.70%	0.21%	0.13%	1.08%

Table 5.11: Overall theoretical systematic uncertainties due to initial state radiation and NNLO corrections, higher-order corrections, PDF uncertainties, final state radiation, and electroweak corrections.

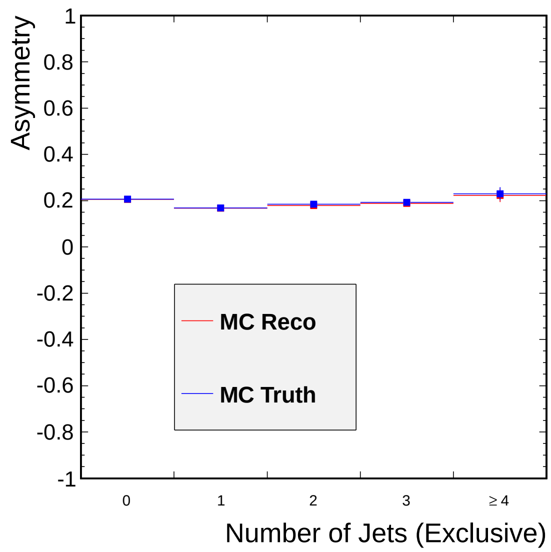
as a function of jet multiplicity is given in Table 5.10, with the plots shown in Figure 5.6.

The overall theoretical uncertainties for the W$\rightarrow\mu\nu_\mu$ process are summed up in Table 5.11. To compute the effect of the theoretical uncertainty on the charge asymmetry, the MC cross sections were varied by adding, then subtracting the appropriate uncertainty, and computing the resultant expected yields, and summing up the overall contribution of each background MC dataset in quadrature. The results are shown in Table 5.12.

5.3.4 Muon Identification and Reconstruction

The overall muon efficiency (including reconstruction, identification, selection, isolation and trigger) is detailed in Sections 4.4.1 and 4.3.2. The single muon efficiencies were determined by using a Tag and Probe method on a golden sample of Z$\rightarrow\mu\mu$ candidates from the entire 2010 dataset. Then the single muon efficiency is decomposed as

$$\epsilon_\mu \equiv \epsilon_{tracker}\cdot\epsilon_{trigger}\cdot\epsilon_{reconstruction}\cdot\epsilon_{isolation} = \epsilon_{tracker}\cdot\epsilon_{global} \tag{5.9}$$



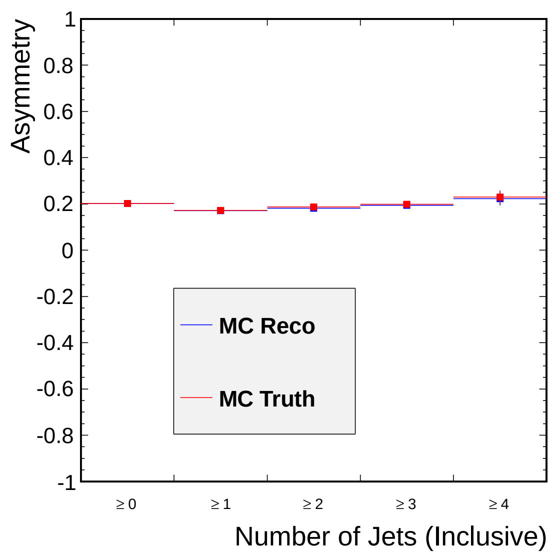


Figure 5.6: Charge asymmetry with jet counting uncertainties for (top) exclusive jet binning and (bottom) inclusive jet binning.

Exclusive	$N_{jets}=0$	$N_{jets}=1$	$N_{jets}=2$	$N_{jets}=3$	$N_{jets}\geq 4$
$t\bar{t}$	$\pm0.0\%$	$\pm0.0\%$	$\pm0.7\%$	$^{+6.6}_{-6.0}\%$	$^{+113.7}_{-35.6}\%$
s-channel top	$\pm0.0\%$	$\pm0.0\%$	$\pm0.0\%$	$\pm0.1\%$	$\pm0.0\%$
t-channel top	$\pm0.0\%$	$\pm0.0\%$	$\pm0.2\%$	$^{+1.6}_{-1.5}\%$	$\pm0.5\%$
tW-channel top	$\pm0.0\%$	$\pm0.0\%$	$\pm0.0\%$	$\pm0.2\%$	$\pm0.8\%$
QCD	$\pm0.0\%$	$\pm0.7\%$	$\pm1.6\%$	$\pm4.5\%$	$\pm4.7\%$
DY $\mu\mu$	$\pm0.4\%$	$\pm0.3\%$	$\pm0.2\%$	$\pm0.1\%$	$\pm0.6\%$
DY $\tau\tau$	$\pm0.0\%$	$\pm0.0\%$	$\pm0.0\%$	$\pm0.0\%$	$\pm0.1\%$
W$\rightarrow\tau\nu_\tau$	$\pm2.8\%$	$\pm2.3\%$	$\pm1.2\%$	$\pm5.1\%$	$\pm2.0\%$
Uncertainty	$\pm2.8\%$	$\pm2.4\%$	$\pm2.1\%$	$^{+9.6}_{-9.2}\%$	$^{+113.8}_{-36.0}\%$

Inclusive	$N_{jets}\geq 0$	$N_{jets}\geq 1$	$N_{jets}\geq 2$	$N_{jets}\geq 3$	$N_{jets}\geq 4$
$t\bar{t}$	$\pm0.1\%$	$\pm0.5\%$	$\pm2.3\%$	$^{+12.2}_{-10.1}\%$	$^{+113.7}_{-35.6}\%$
s-channel top	$\pm0.0\%$	$\pm0.0\%$	$\pm0.0\%$	$\pm0.1\%$	$\pm0.0\%$
t-channel top	$\pm0.0\%$	$\pm0.1\%$	$^{+0.3}_{-0.2}\%$	$^{+1.5}_{-1.4}\%$	$\pm0.5\%$
tW-channel top	$\pm0.0\%$	$\pm0.0\%$	$\pm0.1\%$	$\pm0.3\%$	$\pm0.8\%$
QCD	$\pm0.1\%$	$\pm0.9\%$	$\pm1.8\%$	$\pm2.9\%$	$\pm4.7\%$
DY $\mu\mu$	$\pm0.4\%$	$\pm0.3\%$	$\pm0.2\%$	$\pm0.0\%$	$\pm0.6\%$
DY $\tau\tau$	$\pm0.0\%$	$\pm0.0\%$	$\pm0.0\%$	$\pm0.0\%$	$\pm0.1\%$
W$\rightarrow\tau\nu_\tau$	$\pm2.7\%$	$\pm2.2\%$	$\pm1.4\%$	$\pm3.3\%$	$\pm2.0\%$
Uncertainty	$\pm2.7\%$	$\pm2.4\%$	$\pm3.3\%$	$^{+13.1}_{-11.1}\%$	$^{+113.8}_{-36.0}\%$

	$0\leq\lvert\eta\rvert<0.8$	$0.8\leq\lvert\eta\rvert<1.6$	$1.6\leq\lvert\eta\rvert<2.1$
$t\bar{t}$	$\pm0.1\%$	$\pm0.1\%$	$\pm0.0\%$
s-channel top	$\pm0.0\%$	$\pm0.0\%$	$\pm0.0\%$
t-channel top	$\pm0.0\%$	$\pm0.0\%$	$\pm0.0\%$
tW-channel top	$\pm0.0\%$	$\pm0.0\%$	$\pm0.0\%$
QCD	$\pm0.2\%$	$\pm0.1\%$	$\pm0.0\%$
DY $\mu\mu$	$\pm0.1\%$	$\pm0.5\%$	$\pm1.0\%$
DY $\tau\tau$	$\pm0.0\%$	$\pm0.0\%$	$\pm0.0\%$
W$\rightarrow\tau\nu_\tau$	$\pm3.4\%$	$\pm2.6\%$	$\pm2.1\%$
Total Uncertainty	$\pm3.5\%$	$\pm2.8\%$	$\pm2.6\%$

Table 5.12: Error percentages due to theoretical uncertainty for (top) exclusive, (middle) inclusive jet bins, and (bottom) muon η.

The efficiency numbers are $\epsilon_{data} = 0.8548 \pm 0.0025$ (stat.) $\pm$ 0.0005 (sys.), and $\epsilon_{MC} = 0.8989 \pm 0.0004$ (sys.), with the systematic error source being assigned to fit uncertainties and an overall data/MC ratio of 0.9509 ± 0.0028 (stat.) $\pm$ 0.005 (syst.), yielding an overall contribution of 0.6%. A further source of error is pre-triggering, when the trigger fires 25 ns (one bunch crossing early) in the region $|\eta| < 1.08$. This yields a further discrepancy of 0.5%.

5.3.5 Muon Momentum Scale and Resolution

The procedure for determining muon momentum scale and resolution is described in Section 4.4.1. As mentioned, the differences between the MuScleFit and SIDRA approaches are assigned to systematic error. For the azimuthal correction, the difference is 0.2%, and for the pseudorapidity correction, the difference is on the order of $\mathcal{O}(1\%)$. Theoretical uncertainties in the models used by these two methods can also add an additional uncertainty of 0.5%. The contribution of muon momentum scale and resolution to the charge asymmetry is determined by comparing the measurements made with generator-level muons in signal MC to the measurements made from reconstructed muons smeared with the muon momentum scale and resolution corrections. Conservatively, an error of 2% is assigned to muon momentum scale and resolution uncertainty.

5.3.6 $\not{E}_T$ Scale and Resolution

The effect of $\not{E}_T$ scale and resolution uncertainties was determined in much the same way as for muon momentum scale and resolution. Hadronic recoil distributions were derived from Z$\rightarrow$ $\mu\mu$ data, and then used to smear boson p_T distribution on generated W$\rightarrow$ $\mu\nu_\mu$ events and compared with generator level.

5.3.7 Charge Misidentification

Charge misidentification occurs when a muon's true charge is misidentified. The charge misidentification rate was tested by using a MinBias MC sample, and comparing the generator-level charge to the final reconstructed charge for global and tracker muons. However, the CMS tracker is quite effective at charge identification, and thus the rate of misidentification was negligible ($< 10^{-5}$).

Chapter 6
Conclusion

In this paper, the primary results calculated over the entire 2010 dataset taken at CMS were the W cross section in the muonic decay channel, the W charge asymmetry as a function of muon pseudorapidity, and the W charge asymmetry as a function of jet multiplicity.

6.1 W Cross Section

The W cross section was calculated by first using selection cuts (described in Section 4.6) to reduce background processes, then using selection efficiencies performed on signal Monte Carlo simulated data to determine the true number of signal events, and then accounting for the integrated luminosity of 35.9 pb^{-1} to arrive at the result

$$\sigma_{PP\rightarrow WX} \times \mathrm{BR}\left(\mathrm{W} \rightarrow \mu\nu_{\mu}\right) = 9570 \pm 27 \text{ (stat.)} \pm 220 \text{ (syst.)} \pm 965 \text{ (lum.) pb.} \tag{6.1}$$

This is in good agreement with the theoretical NNLO prediction of

$$\sigma_{PP\rightarrow WX} \times \mathrm{BR}\left(\mathrm{W} \rightarrow \mu\nu_{\mu}\right) = 10440 \pm 520 \text{ pb.} \tag{6.2}$$

Comparison to past experimental results and the theoretical results predicted by FEWZ and MSTW2008 can be seen in Figure 6.1.

The calculated W^{+} and W^{-} cross sections are

$$\sigma_{PP\rightarrow WX} \times \mathrm{BR}\left(\mathrm{W}^{+} \rightarrow \mu^{+}\nu_{\mu}\right) = 5730 \pm 21 \text{ (stat.)} \pm 132 \text{ (syst.)} \pm 630 \text{ (lum.) pb.} \tag{6.3}$$

and

$$\sigma_{PP \to WX} \times \text{BR}\left(\text{W}^- \to \mu^- \nu_\mu\right) = 3850 \pm 17 \text{ (stat.)} \pm 89 \text{ (syst.)} \pm 424 \text{ (lum.) pb.} \quad (6.4)$$

respectively, which when combined with statistical and systematic uncertainties are within the range predicted at NNLO of 6150 ± 290 pb for W^+ and 4290 ± 230 pb for W^-.

The largest uncertainty is the luminosity, at 11%. The relatively large uncertainty in the luminosity is due to the paucity of events necessary for the accurate measurement of well-known cross sections such as Z$\to \mu^+\mu^-$, and the reliance of beam width measurement to measure normalized luminosity. Other sources of systematic error which were taken into account were PDF and theoretical errors (in estimating the background contribution), muon momentum scale and resolution uncertainties, and pre-triggering.

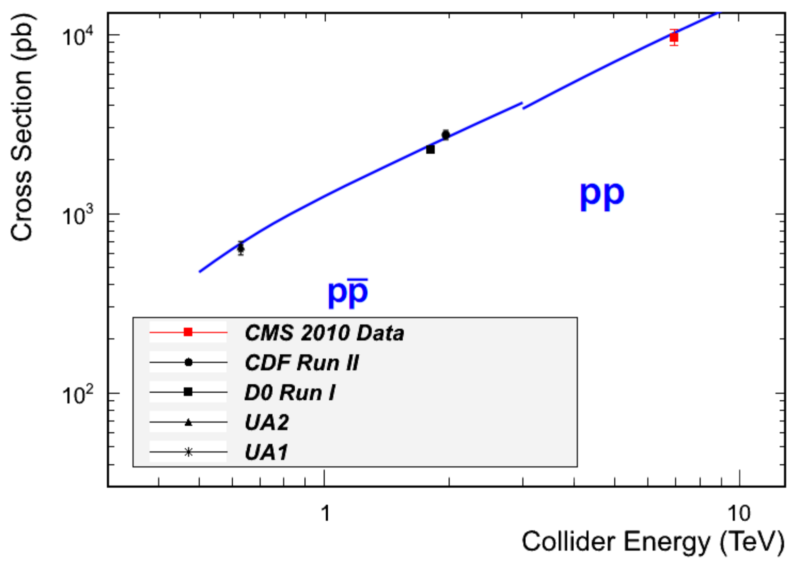


Figure 6.1: Measurement of the inclusive cross section $\sigma_{PP\to W} \times \text{BR}\left(\text{W} \to \mu\nu_\mu\right)$ at CMS using the 2010 dataset and at lower-energy colliders. The error bars shown are statistical $\oplus$ systematic uncertainty (not including luminosity), and the theory curve is the NNLO prediction given by FEWZ and MSTW2008[8][9].

6.2 W Asymmetry

The W asymmetry was calculated by first applying the same selection cuts as for the cross section calculation, normalizing the expected Monte Carlo yields to the luminosity of 35.9 pb^{-1}, and then using sideband subtraction to remove the expected background contamination. The overall results as a function of muon pseudorapidity are stated in Table 5.4, and as a function of exclusive and inclusive jet multiplicity in Table 5.5. The overall $\mathrm{W}^+/\mathrm{W}^-$ ratio is 1.489 ± 0.009 (stat.) $\pm\, 0.032$ (syst.), which compares well with the MSTW2008 NNLO value of 1.429 ± 0.013 at $\sqrt{s} = 7$ TeV and 1.435 ± 0.098 from MadGraph Monte Carlo. Comparisons to predictions made from CT10W and MSTW2008 PDF sets are shown in 6.2, with overall excellent agreement. The overall asymmetry agrees with the prediction that asymmetry should increase with $|\eta|$ due to the e^{+y} dependence of the u quark momentum fraction and the e^{-y} dependence of the $\bar{d}$ fraction (see Equation 2.64), and due to the fact that the up quark carries a larger momentum fraction in the proton than the down quark. The major sources of systematic uncertainty were PDF uncertainties and theoretical errors. Other sources of error include muon efficiency, muon momentum scale and resolution and luminosity, which in principle mostly cancel due to the final measurement effectively being a cross section ratio.

In order to investigate QCD uncertainties, W charge asymmetry was also measured as a function of jet multiplicity. Here, the major systematic uncertainty was jet energy scale and resolution, due largely to initial-state and final-state gluon radiation, and QCD multijet production. Comparisons to predictions from PYTHIA and MadGraph Monte Carlo datasets are shown in 6.3 for both inclusive and exclusive jet binning, and overall better agreement with MadGraph MC (due to better QCD modelling than with the PYTHIA generator). Overall agreement is good except in the $N_{jets} = 3$ bin, although with the large uncertainties due to statistics and jet energy scale and resolution it is still within range.

The charge asymmetry of the W boson in the muonic decay channel has been measured in previous experiments, although as it has only been previously measured at $p\bar{p}$ colliders results are not directly comparable. Concurrent with this measurement, W muon charge asymmetry was also measured at the ATLAS detector using the entire 36 pb^{-1} of the 2010 dataset. The charge asymmetry as a function of W rapidity at ATLAS is shown in Figure 6.4. The kinematic cuts applied by ATLAS were muon $p_T > 20$ GeV, $\not{E}_T > 25$ GeV, and $m_T > 40$ GeV. Although the systematics are different between ATLAS and CMS, the results are in overall very good agreement.

This analysis represents the first measurement of W charge asymmetry in the muonic decay channel at CMS with $\sqrt{s} = 7$ TeV, and demonstrates agreement with theoretical predictions despite high statistical and systematic uncertainties in higher jet multiplicity regimes, and furthermore demonstrates good identification and reconstruction of muons, missing transverse energy and jets. Nevertheless, with increased statistics and greater understanding of the detector, other possibilities avail themselves. One further measurement which could be made is the cross-section ratio of $\sigma(W+N_{jets})/\sigma(W+(N_{jets}-1))$ as a test of perturbative QCD calculations, and hence a cross-check against W asymmetry as a function of jet multiplicity. Another possibility as statistics increase is to use data-driven methods to estimate background, rather than relying on a Monte Carlo cut-and-count method as used here. One further possibility is to measure charge asymmetry with different cuts, such as different p_T, $\not{E}_T$ and m_T cuts. As further studies of electroweak and QCD processes are undertaken at the LHC, further constraints on the theoretical uncertainties will be made, and improvements and refinements will be made to the identification and reconstruction algorithms. Further statistics in the higher jet multiplicity regimes should enable the use of methods such as B jet tagging to distinguish between different types of jets produced in association with a W boson, as well as cut down on the prodigious statistical error. Further studies with increased data will prove invaluable in constraining PDFs and understanding W boson events, which are an important background as the LHC soars into the new physics regime.

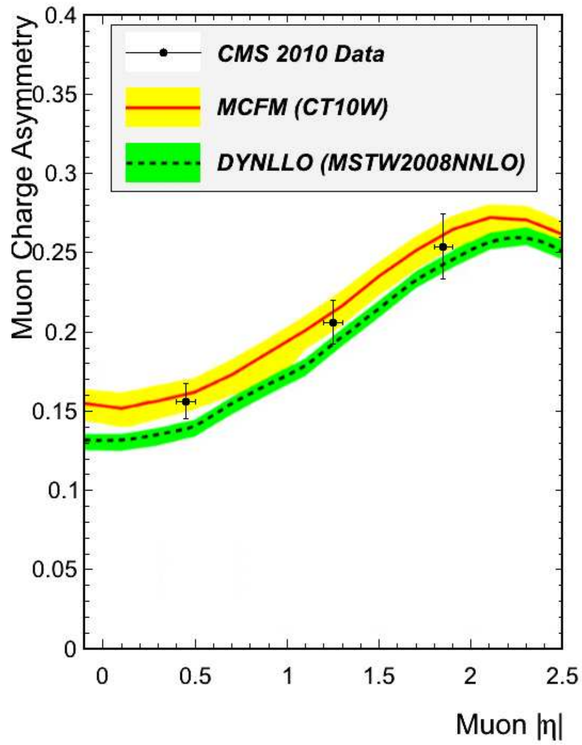


Figure 6.2: The muon charge asymmetry as a function of pseudorapidity, along with predictions from CT10W and MSTW2008NNLO.

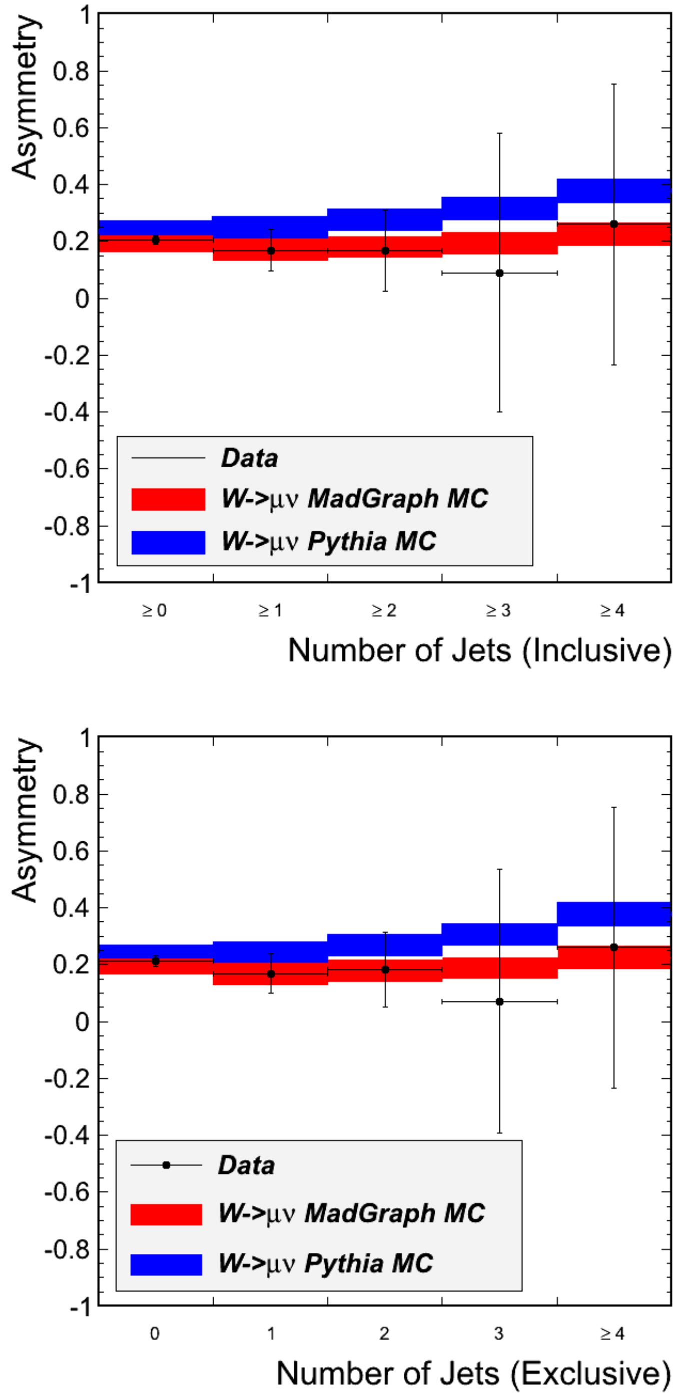


Figure 6.3: The muon charge asymmetry as a function of (top) inclusive jet multiplicity and (bottom) exclusive jet multiplicity, along with predictions from PYTHIA and MadGraph. The error bars on data represent statistical uncertainty $\oplus$ systematic uncertainty.

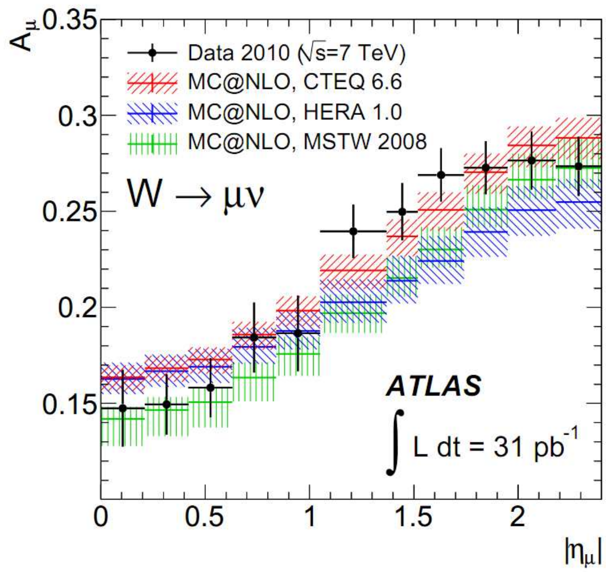


Figure 6.4: The W production charge asymmetry as measured at ATLAS, along with predictions from CTEQ, HERA and MSTW at NLO[10].

Bibliography

[1] Cern: The accelerator complex. http://public.web.cern.ch/public/en/Research/AccelComplex-en.html, May 2010.

[2] Cern: Luminosity overview for proton collisions. http://cms-service-lumi.web.cern.ch/cms-service-lumi/overview.php, May 2011.

[3] G. L. Bayatian et al. CMS physics: Technical design report. CERN-LHCC-2006-001.

[4] Paolo Azzurri. Track Reconstruction Performance in CMS. *Nucl. Phys. Proc. Suppl.*, 197:275–278, 2009.

[5] S. Chatrchyan et al. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(8):S08004, 2008.

[6] Y. Sirois. Inter-calibration and energy measurements performance of the cms pbwo4 electromagnetic calorimeter. *Nuclear Physics B - Proceedings Supplements*, 172:117 – 125, 2007. Proceedings of the 10th Topical Seminar on Innovative Particle and Radiation Detectors, Proceedings of the 10th Topical Seminar on Innovative Particle and Radiation Detectors.

[7] N. Saoulidou. Particle flow jet identification criteria. *CMS Note*, 2010. CMS AN AN-10-003.

[8] Kirill Melnikov and Frank Petriello. Electroweak gauge boson production at hadron colliders through O(alpha(s)**2). *Phys. Rev.*, D74:114017, 2006.

[9] Kirill Melnikov and Frank Petriello. The W boson production cross section at the LHC through O(alpha**2(s). *Phys. Rev. Lett.*, 96:231803, 2006.

[10] Georges Aad et al. Measurement of the Muon Charge Asymmetry from W Bosons Produced in pp Collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Phys. Lett. B*, 701:31–49, 2011. * Temporary entry *.

[11] K. Nakamura et al. Review of particle physics. *J. Phys.*, G37:075021, 2010.

[12] The Large Hadron Collider: Conceptual design. CERN-AC-95-05-LHC.

[13] F. Halzen and Alan D. Martin. QUARKS AND LEPTONS: AN INTRODUCTORY COURSE IN MODERN PARTICLE PHYSICS. New York, Usa: Wiley (1984) 396p.

[14] Michael Edward Peskin and Daniel V. Schroeder. An Introduction to quantum field theory. Reading, USA: Addison-Wesley (1995) 842 p.

[15] T. D. Lee and Chen-Ning Yang. Question of Parity Conservation in Weak Interactions. *Phys. Rev.*, 104:254–258, 1956.

[16] Nicola Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, 10(12):531–533, Jun 1963.

[17] D. J. Gross and Frank Wilczek. Asymptotically Free Gauge Theories. 1. *Phys. Rev.*, D8:3633–3652, 1973.

[18] Davison E. Soper. Parton distribution functions. *Nucl. Phys. Proc. Suppl.*, 53:69–80, 1997.

[19] A. V. Belitsky and A. V. Radyushkin. Unraveling hadron structure with generalized parton distributions. *Phys. Rept.*, 418:1–387, 2005.

[20] A. Bodek et al. Comparisons of Deep Inelastic e p and e n Cross- Sections. *Phys. Rev. Lett.*, 30:1087, 1973.

[21] Vernon D. Barger and R. J. N. Phillips. COLLIDER PHYSICS. REDWOOD CITY, USA: ADDISON-WESLEY (1987) 592 P. (FRONTIERS IN PHYSICS, 71).

[22] Sidney D. Drell and Tung-Mow Yan. Massive lepton-pair production in hadron-hadron collisions at high energies. *Phys. Rev. Lett.*, 25(5):316–320, Aug 1970.

[23] W J Stirling and M R Whalley. A compilation of drell-yan cross sections. *Journal of Physics G: Nuclear and Particle Physics*, 19(D):D1, 1993.

[24] L. Evans. LHC status. Prepared for APAC 2007: Asian Particle Accelerator Conference, Indore, India, 29 Jan - 2 Feb 2007.

[25] Doreen Wackeroth. High energy physics made painless: Cross section.

[26] Lyndon Evans, (ed.) and Philip Bryant, (ed.). LHC Machine. *JINST*, 3:S08001, 2008.

[27] G. Acquistapace et al. CMS, the magnet project: Technical design report. CERN-LHCC-97-10.

[28] A. Tricomi. Performances of the ATLAS and CMS silicon tracker. *Eur. Phys. J.*, C33:s1023–s1025, 2004.

[29] M. Weber. The CMS tracker. *Nucl. Phys. Proc. Suppl.*, 142:430–433, 2005.

[30] Serguei Chatrchyan et al. Commissioning and Performance of the CMS Pixel Tracker with Cosmic Ray Muons. *JINST*, 5:T03007, 2010.

[31] Bruce J. King. An all-solid state central tracker for the proposed DESY electron positron linear collider. 1996.

[32] F. P. Schilling. Track reconstruction and alignment with the CMS silicon tracker. 2006.

[33] Wolfgang Adam. Track and vertex reconstruction in CMS. *Nucl. Instrum. Meth.*, A582:781–784, 2007.

[34] CMS: The hadron calorimeter Technical Design Report. CERN-LHCC-97-31.

[35] CMS: The electromagnetic calorimeter Technical Design report. CERN-LHCC-97-33.

[36] CMS: The tracker Technical Design Report. CERN-LHCC-98-06.

[37] S. Baffioni et al. Electron reconstruction in CMS. *Eur. Phys. J.*, C49:1099–1116, 2007.

[38]

[39] P. Sphicas, (ed.). CMS: The TriDAS project. Technical design report, Vol. 2: Data acquisition and high-level trigger. CERN-LHCC-2002-026.

[40] L. Agostino et al. Commissioning of the CMS High Level Trigger. *JINST*, 4:P10005, 2009.

[41] S. Van der Meer. Calibration of the effective beam height in the ISR. 1968. CERN-ISR-PO/68-3.

[42] Pythia. http://home.thep.lu.se/~torbjorn/Pythia.html, May 2011.

[43] Homepage of the powheg box. http://powhegbox.mib.infn.it/, May 2011.

[44] Pdf4lhc. http://www.hep.ucl.ac.uk/pdf4lhc/index.html, May 2011.

[45] R. Bellan. Muon reconstruction with the cms tracking system. *Nuclear Physics B - Proceedings Supplements*, 177-178:253 – 254, 2008. Proceedings of the Hadron Collider Physics Symposium 2007, Proceedings of the Hadron Collider Physics Symposium 2007.

[46] Measurement of Tracking Efficiency. 2010. CMS PAS TRK-10-002.

[47] A. Dubak. Measurement of the e+ p neutral current DIS cross section and the F2, F(L), xF3 structure functions in the H1 experiment at HERA. MPP-2003-65.

[48] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852–867, 2008.

[49] D. Dobur. Jets and Missing Transverse Energy Reconstruction with CMS. *ArXiv e-prints*, April 2009.

[50] ParticleFlow Event Reconstruction in CMS and Performance for Jets, Taus, and E_T^{miss}. CMS PAS PFT-09-001.

[51] Gavin P. Salam and Gregory Soyez. A practical Seedless Infrared-Safe Cone jet algorithm. *JHEP*, 05:086, 2007.

[52] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008.

[53] Monte carlo network. http://www.montecarlonet.org/, May 2011.

[54] S. Agostinelli et al. G4–a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250 – 303, 2003.

[55] A. Quadt. Top quark physics at hadron colliders. *The European Physical Journal C - Particles and Fields*, 48:835–1000, 2006. 10.1140/epjc/s2006-02631-6.

[56] Serguei Chatrchyan et al. Measurement of the lepton charge asymmetry in inclusive W production in pp collisions at sqrt(s) = 7 TeV. *JHEP*, 04:050, 2011.

[57] Rates of Jets Produced in Association with W and Z Bosons. CMS PAS EWK-10-012.

[58] Brian Robert Martin and Graham Shaw. Particle physics. Chichester, UK: Wiley (2008) 442 p.

[59] Measurement of CMS Luminosity. CMS PAS EWK-10-004.

[60] Simon White, Reyes Alemany-Fernandez, Helmut Burkhardt, and Mike Lamont. First Luminosity Scans in the LHC. 1st International Particle Accelerator Conference: IPAC'10, 23-28 May 2010, Kyoto, Japan.

[61] V. Halyo A. Hunt K. Mishra N. Adam, J. Berryhill. Generic Tag and Probe Tool for Measuring Efficiency at CMS with Early Data. 2009. CMS AN-2009/111.

[62] Performance of muon reconstruction and identification in pp collisions at $\sqrt{s} = 7$ TeV. 2011. CMS PAS MUO-10-004.

[63] Jet Performance in pp Collisions at $\sqrt{s} = 7$ TeV. 2010. CMS PAS JME-10-003.

[64] Measurement of the Jet Energy Resolutions and Jet Reconstruction Efficiency at CMS. CMS PAS JME-09-007.

[65] On Measuring Missing Transverse Energy with the CMS Detector in pp Collisions at $\sqrt{s} = 7$ TeV. 2010. CMS PAS JME-10-009.

[66] John M. Campbell, J. W. Huston, and W. J. Stirling. Hard Interactions of Quarks and Gluons: A Primer for LHC Physics. *Rept. Prog. Phys.*, 70:89, 2007.

[67] S. Catani, G. Ferrera, and M. Grazzini. W boson production at hadron colliders: the lepton charge asymmetry in NNLO QCD. *JHEP*, 05:006, 2010.

[68] T. Aaltonen et al. Direct Measurement of the W Production Charge Asymmetry in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV. *Phys.Rev.Lett.*, 102:181801, 2009.

[69] Mika Vesterinen. A Measurement of the muon charge asymmetry in W boson events. 2010.

Appendix A
Datasets

The following table is comprised of a list of datasets, both real and simulated, used in this analysis, along with cross-sections and number of events. For collision data, The JSON file used (containing the list of DQM-certified good lumisections) was "Cert_132440-149442_7TeV_StreamExpress_Collisions10_JSON.txt".

Dataset	N_{events}	σ (pb)	Filter Eff.	MadGraph Corr.	$\mathcal{L}$ (pb^{-1})
/Mu/Run2010A-Nov4ReReco_v1/RECO	51 860 222				3.18
/Mu/Run2010B-Nov4ReReco_v1/RECO	33 470 281				32.74
/WJetsToLNu_TuneZ2_7TeV-madgraph-tauola	15 123 740	31 314	1.000		482.97
/WToMuNu_TuneZ2_7TeV-pythia6	5 330 940	7 899	1.000		674.89
/QCD_Pt-20_MuEnrichedPt-15_TuneZ2_7TeV-pythia6	29 504 866	296 600 000	0.0002855		348.43
/WToTauNu_TuneZ2_7TeV-pythia6-tauola	5 221 750	7 899	1.000		661.06
/DYToMuMu_M-20_TuneZ2_7TeV-pythia6	2 289 913	1 300	1.000		1761.47
/DYToTauTau_M-20_TuneZ2_7TeV-pythia6	2 057 446	1 300	1.000		1582.65
/TTJets_TuneZ2_7TeV-madgraph-tauola	1 165 716	157.5	1.000	0.985608	7294.85
/TToBLNu_TuneZ2_s-channel_7TeV-madgraph	494 967	4.21	1.000		115 569
/TToBLNu_TuneZ2_t-channel_7TeV-madgraph	484 060	20.93	1.000		23 127.57
/TToBLNu_TuneZ2_tW-channel_7TeV-madgraph	494 961	10.56	1.000	0.985608	46 196.73

Table A.1: Full DBS names, cross sections and event multiplicities for data sets (real and simulated) used in this analysis. The first two rows describe the actual 2010 collision data taken, the rest are Monte Carlo simulated data. For each of the Monte Carlo data sets, the Fall10 production was used, and the GEN-SIM-RECO data format was used.