SEARCH FOR CONTACT INTERACTIONS IN THE DIMUON CHANNEL IN P-P COLLISIONS AT $\sqrt{s} = 7$ TeV AT CMS

by

SOWJANYA GOLLAPINNI

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2012

MAJOR: PHYSICS

Approved By:

Advisor Date

DEDICATION

To

the love of my life, Kranti

and

my precious babies, Samatha and Viplav

ACKNOWLEDGEMENTS

It is a pleasure to thank the many people who made this thesis possible. First and foremost, I would like to acknowledge my advisor, Paul E. Karchin and owe my most sincere gratitude to him. His wide knowledge and logical way of thinking have been of great value to me. His great efforts to explain things clearly and simply made physics fun for me. He is a fabulous advisor and a wonderful human being. I immensely appreciate his patience and kindness towards students. He always made sure of my personal well being and supported me at every step during these years. I count myself very fortunate to have him as my advisor. On the personal side, he treated me and my family as an extended part of his family and it was a joy to have spent time with his family. I especially thank his wife Venus for being so warm and welcoming and making us feel like home.

I am very grateful to Lenny Spiegel for mentoring me at Fermilab. I can't thank him enough for his detailed and constructive comments, and for his important support throughout this work. I had many productive scientific discussions with him which greatly contributed towards my education. I would like to thank the rest of the Wayne State CMS group for their invaluable help and collaboration throughout my research career. I also owe my thanks to all those wonderful people at Fermilab who readily helped me with their expertise. In particular, I would like to thank Patrick Gartung and Catalin Dumitrescu for their help with the technical details of the research. I am also very grateful to Gena Kukartsev for his help with the statistical tools needed for this thesis.

I would also like to take this opportunity to acknowledge all of my CMS collaborators whose indirect support and continuous help made this thesis possible. In particular, I would like to thank the Exotica group for their constant support and feedback. There are countless discussions and presentations that all have contributed to my education. I am very thankful to my thesis dissertation committee members whose comments and suggestions have greatly improved the quality of this thesis.

I would like to thank all the great people in the WSU Physics and Astronomy department. My special thanks to the department chair, Ratna Naik and graduate advisor, Jo Wadehra for always being so supportive and helpful to students. I am grateful to the administrative staff of the physics department, especially Lashara, Shere, De and Doris for assisting me in many different ways. I would also like to thank Scott Payson for making undergraduate lab teaching so much fun.

I would like to thank my parents, Usha Kiran and Subramanyam, for always being proud of me. I am greatly indebted to my father who always took special interest in my education and encouraged all my career decisions. I can't thank my mother enough for all she has done for me, she was always there for me and loved me no matter what. I am very grateful to my father-in-law, Rama Murthy, who unhesitatingly and willingly supported my trip to U.S.A. I thank him for helping me during the financially difficult times.

I am deeply grateful to my wonderful life partner, *Kranti*, who believed in me more than I believed myself and always encouraged me to go beyond what I think I can do. I have always been motivated by him and thank him for all our conversations. Without his unconditional love, support and encouragement none of this would have been possible. From the moment they were born, my little sweethearts, *Samatha* and *Viplav*, are nothing but bundles of fun and joy. Nothing could make my tough days easier to bear if its not for their innocent acts and sweet smiles. My loving thanks to them for being so good and cooperative during the busy days.

TABLE OF CONTENTS

Dedication . . . . . **ii**

Acknowledgements . . . . . **iii**

List of Figures . . . . . **xi**

List of Tables . . . . . **xii**

1 Introduction **1**

2 Theory **4**

2.1 The standard model 4

2.1.1 Introduction to the standard model 5

2.1.2 Origin of the weak force 9

2.1.3 The Electro-weak model 10

2.2 Physics beyond the standard model 11

2.2.1 Contact Interactions 12

2.2.2 Quark and Lepton Compositeness 14

2.2.3 Previous results on Λ 14

2.3 Physics of p-p collisions 15

3 Experimental setup **21**

3.1 The Large Hadron Collider 21

3.1.1 Layout 22

3.1.2 Experiments at the LHC 24

3.1.3 Operation and performance 25

3.2 The CMS detector 28

3.2.1 The CMS coordinate system 29

3.2.2 Layout 30
3.2.3 Inner tracking system 32
3.2.4 Calorimetry 35
3.2.5 Muon system 38
3.2.6 Trigger and Data Acquisition 45

4 The PYTHIA Monte Carlo 49
4.1 Compositeness models in PYTHIA 49
4.1.1 PYTHIA Job details 51
4.2 The left-left iso-scalar model 52

5 Data and event selection 55
5.1 The 2011 dataset 55
5.1.1 Trigger requirements 56
5.2 Dimuon selection criteria 58
5.3 Event pileup 65

6 Simulation of signal and background events 67
6.1 Expected signal and corrections 67
6.1.1 Detector acceptance times mass migration 69
6.1.2 K factors 74
6.1.3 SM backgrounds 76
6.2 Predicted event yields 77

7 Systematic uncertainties 80
7.1 Theoretical sources 80
7.2 Experimental sources 83
7.3 Summary 83

8 Statistical analysis method **85**

8.1 The modified-frequentist method 85

8.2 Agreement of data with SM predictions 90

8.3 Exclusion lower limits on Λ 93

8.3.1 Effect of systematics 96

9 Results and Conclusion **98**

Bibliography **100**

Abstract **110**

Autobiographical Statement **111**

LIST OF FIGURES

Figure 2.1 Fundamental particles in the SM [1]. Note that this figure does not show the SM Higgs boson. 6

Figure 2.2 Fermi's β-decay, before (left) and now (right) [2]. 9

Figure 2.3 Interference between DY and CI amplitudes resulting in dimuon final states [13]. 13

Figure 2.4 Simplified parton model of the proton [4]. 15

Figure 2.5 Detailed partonic structure of the proton [5]. 15

Figure 2.6 MSTW 2008 NLO PDFs (at 68% C.L.) of the proton for a momentum transfer of 10 $(\mathrm{GeV/c})^2$ (left) and 10^3 $(\mathrm{GeV/c})^2$ (right). . . . 17

Figure 2.7 MSTW 2008 NNLO PDFs (at 68% C.L.) of the proton for a momentum transfer of 10 $(\mathrm{GeV/c})^2$ (left) and 10^3 $(\mathrm{GeV/c})^2$ (right). . 17

Figure 2.8 p-p collision event illustrating the hard scattering, underlying event and the initial and final state radiations [45]. 18

Figure 3.1 Schematic layout of the LHC collider tunnel [53]. 22

Figure 3.2 Layout of the CERN accelerator complex [60]. 23

Figure 3.3 Maximum instantaneous luminosity per day delivered to CMS in 2011 for p–p runs at $\sqrt{s} = 7$ TeV. 26

Figure 3.4 Integrated luminosity vs. fill number for 2011 p–p runs. 26

Figure 3.5 Integrated luminosity delivered to (yellow), and recorded by CMS (red) for 2010 and 2011 p–p runs at $\sqrt{s} = 7$ TeV. 27

Figure 3.6 As polar angle increases, pseudorapidity decreases. The "forward region" of a a collider detector corresponds to regions with high η values, typically $\eta > 3$. 30

Figure 3.7 Diagram of the CMS detector with a quadrant cut away to show the interior [55]. 31

Figure 3.8 Particle identification at CMS [52]. 31

Figure 3.9 Layout of the CMS inner tracking system [55]. 33

Figure 3.10 Schematic layout of the CMS tracker showing detailed coverage in the r–z plane [55]. Each line represents a detector module. Double lines indicate back-to-back modules which deliver stereo hits. . . . 34

Figure 3.11 Schematic of an electromagnetic shower. 36

Figure 3.12 Layout of the CMS ECAL showing the arrangement of $PbWO_4$ crystals in the barrel and endcap regions, along with the preshower detector [55] 37

Figure 3.13 Layout of the hadron calorimeter [55]. 38

Figure 3.14 Layout of the CMS muon system shown for one quadrant of the CMS detector [55]. The four barrel DT stations (MB1–MB4), four endcap CSC stations (ME1–ME4), and RPC stations are shown. 39

Figure 3.15 Layout of the barrel muon (MB) DT chambers in one of the 5 CMS barrel wheels [55]. Each DT chamber is denoted by MB/Z/a/b, where a refers to the DT station number ranging from 1 to 4 and b refers to the chamber number (within a DT station) ranging from 1 to 12. In chambers 4 (top) and 10 (bottom), the MB4 chambers are cut in half to simplify the mechanical assembly. YB refers to the yoke barrel regions with same numbering scheme as used for the MB DT stations. 40

Figure 3.16 Muon drift tube operation. The red line with arrow shows a muon traversing the drift tube. The anode wires (shown by red dots) are perpendicular to the page. The horizontal blue lines with arrows show the muon's distance from the anode wire (obtained by multiplying the speed of an electron in the tube by the drift time). 41

Figure 3.17 Quarter-view of the CMS detector [55]. Shown in red color are the CSCs. The ME4/2 chambers will eventually be built and installed as part of the CMS upgrade plans. 42

Figure 3.18 CSC chamber design and operation [55]. Left panel: Layout of a CSC chamber made of 7 trapezoidal planes. The exposed portion shows anode wires and cathode strips. Right panel: A schematic view of a single gas gap illustrating the principle of CSC operation. The muon coordinate along wires is obtained by interpolating charges induced on cathode strips. 42

Figure 3.19 Schematic layout of one of the CMS barrel wheels, labeled W+2. 44

Figure 3.20 Layout of a double-gap RPC [55]. 44

Figure 3.21 The muon p_T resolution as a function of p_T using simulated data for the muon system only, the inner tracking system only, and both [55]. Left panel: $|\eta| < 0.8$, right panel: $1.2 < |\eta| < 2.4$. 46

Figure 4.1 Constructive and destructive LLIM cross sections for $M_{\mu^+\mu^-} > 120$ GeV/c^2. The functional fits differ only in the sign of the interference term in Eq. 2.5. The solid black line corresponds to the DY asymptotic limit. 52

Figure 4.2 Dimuon event yields at different values of Λ for (a) destructive interference and (b) constructive interference. Values are shown for M >120 GeV/c^2. As Λ increases, the dimuon mass distribution tends toward pure DY. The model predictions are shown over the full mass range, although the model is not valid for $M_{\mu\mu}^{Low}c^2 \geq \Lambda$. 53

Figure 5.1 Opposite sign dimuon invariant mass spectra showing the resonant peaks of the SM particles. Dimuon mass is given in units of $c = 1$. 56

Figure 5.2 Single-muon trigger efficiencies for $p_T > 35$ GeV/c as a function of η: the efficiency of the HLT with p_T threshold at 30 GeV/c with respect to L1 trigger with p_T threshold of 12 GeV/c (left), and the combined efficiency of L1 and HLT (right). The efficiencies obtained using $Z \to \mu^+\mu^-$ events are compared with predictions from the MC simulation. 58

Figure 5.3 Dimuon mass distribution using the full 2011 dataset after imposing all the selection requirements. 62

Figure 5.4 Event displays for the two highest-mass $\mu^+\mu^-$ events. Left (right) side of figures (a) and (b) show the transverse (longitudinal) view. (a) This display corresponds to the highest mass (1379 GeV/c^2) dimuon event with muon kinematic variables: p_T = (686, 622) GeV/c, η = (-0.05, +0.63), and ϕ = (-0.46, +2.82). (b) This display corresponds to the next highest mass (1256 GeV/c^2) dimuon event with muon kinematic variables: p_T = (196, 299) GeV/c, η = (-1.46, +1.80), and ϕ = (-0.72, +2.30). 63

Figure 5.5 Dimuon mass spectra for same sign muon candidate events. . . . 64

Figure 5.6 Distribution of the number of primary vertices per event from data for the 2011A and 2011B data taking periods. 65

Figure 6.1 Acceptance times migration for $M > M_{\mu\mu}^{Low}$. 70

Figure 6.2 Reconstructed dimuon mass distributions associated with a particular generator mass window for 100 GeV/c^2 bins. All distributions are normalized to correspond to the luminosity of the 200–300 GeV/c^2 bin. The figure illustrates the gradually degrading mass resolution with increasing dimuon mass. 72

Figure 6.3 Comparison of $A \times M$ values in 100 GeV/c^2 bins for CI and DY production. The plot starts at 600 GeV/c^2 on the x-axis. 73

Figure 6.4 QCD K factors for $M > M_{\mu\mu}^{Low}$. The values are plotted at the threshold points. 75

Figure 6.5 QED K factors for $M > M_{\mu\mu}^{Low}$. The values are plotted at the threshold points. 75

Figure 7.1 Maximal positive and negative PDF uncertainties as a function of mass. 81

Figure 8.1 Observed spectrum of $M_{\mu\mu}$ and predictions from the SM and LLIM (CI/DY) for $\Lambda = 4$ and 5 TeV, for constructive and destructive interference. 92

Figure 8.2 Same as Fig. 8.1 except that the x-axis is also plotted on a logarithmic scale. 92

Figure 8.3 Observed and expected limits as a function of $M_{\mu\mu}^{Low}$ for destructive interference. 95

Figure 8.4 Same as Fig. 8.3, for constructive interference. 95

LIST OF TABLES

Table 4.1 Compositeness model options within PYTHIA for dilepton final states. 50

Table 5.1 The 2011A and 2011B datasets used, with run ranges, HLT paths, and corresponding luminosities. 57

Table 5.2 Effect of various selection cuts on muon event statistics. Here 'initial sample (I)' corresponds to a subset of events which had passed the baseline selection, trigger requirement and contains at least one muon that is both tracker and global. N_{events} refer to the number of events passing a specific cut relative to the previous selection step. The cumulative effect of cuts on the initial sample is given in the last column. 61

Table 5.3 List of dimuon events in data with masses above 800 GeV/c^2, arranged in descending order of their masses. 62

Table 6.1 Description of event samples with detector simulation. For the DY process Z/$\gamma^*\rightarrow \mu\mu$, the minimum dimuon mass is indicated as part of the PYTHIA sample name. The cross section σ and integrated luminosity L are given for each sample generated. 69

Table 6.2 Multiplicative factors for $M > M_{\mu\mu}^{Low}$. The uncertainties shown are statistical. The systematic uncertainties are described in Chapter 7. 70

Table 6.3 Expected event yields for non-DY SM backgrounds. DY event yields are shown for comparison. The uncertainties shown are statistical. . 77

Table 6.4 Observed and expected number of events for $M_{\mu\mu} > M_{\mu\mu}^{Low}$. The expected yields are shown for DY production and CI/DY production for destructive interference with given Λ values. Expected yields include contributions from non-DY backgrounds. 79

Table 6.5 Observed and expected number of events as in Table 6.4. Here CI/DY predictions are for constructive interference. 79

Table 7.1 Maximal positive and negative PDF uncertainties evaluated with CTEQ66 NLO PDF set using the modified tolerance method. . . . 82

Table 7.2 Sources of systematic uncertainty. Where appropriate, the values are quoted for $M_{\mu\mu} > 700$ GeV/c^2, $\Lambda = 13$ TeV, and for constructive interference. 83

Table 8.1 Dimuon event statistics for the data and SM production corresponding to 5.3 fb^{-1}. The numbers shown for SM predictions are after normalization to the Z peak. 91

Chapter 1

Introduction

For decades, particle physicists have always tried to answer the one question, 'what is the Universe made of on the smallest scale of size?' While the question may look simple, the answer is much more complicated even to imagine. Over the years, the standard model (SM) of electro-weak interactions has shown impressive predictive power and consistent agreement with experiment. While the SM successfully explains many of the elementary processes that we observe in nature, there are some predictions that still remain unconfirmed by experiments. The Higgs mechanism that explains the origin of electro-weak symmetry breaking necessary to give masses to the observed gauge bosons is yet to be confirmed by experiments (On July 4, the Large Hadron Collider (LHC) at the Center for European Nuclear Research (CERN) announced the discovery of a new particle, bosonic in nature, that is consistent with the SM Higgs boson, but stressed that further testing is needed to confirm). The SM fails to explain phenomena such as gravity, dark matter, matter-antimatter asymmetry, and neutrino masses[1]. There are also some problems within the theoretical framework of the SM such as the origin of the 19 arbitrary parame-

[1]Within the theoretical framework of the SM, neutrinos are assumed to be massless particles. However, recent neutrino oscillation experiments have shown that neutrinos have non-zero masses although their absolute values are still unknown.

ters, the number of quark/lepton families, and the hierarchy problem[2]. These important shortcomings motivated particle physicists to develop theories and design experiments to explore physics beyond the standard model (BSM). One important possibility is that the SM is a low-energy effective theory, while the true high-energy theory may be described by other degrees of freedom, which would manifest themselves as constituents of quarks and leptons.

The LHC particle accelerator located at CERN bordering France and Switzerland was built to explore new physics at the TeV scale. At the LHC, proton-proton collisions occur at the highest center-of-mass energy ever achieved. The Compact Muon Solenoid (CMS) experiment which is the experimental apparatus for this analysis, is a general-purpose particle detector at the LHC built to probe the electroweak scale. The analysis presented in this thesis is based on the LHC data collected by the CMS detector in 2011. A search for new physics signatures in the form of four-fermion contact interactions in conjunction with an assumed substructure of quarks and leptons, is performed in events with two isolated muons. Even though the exact intermediate particle exchange mechanism for this new physics model is yet to be discovered, the presence of a new interaction will be manifest as a deviation from the SM predictions in the high-mass tail of the invariant mass distribution of the opposite-sign dimuon pairs. The choice to look for contact interaction signals in the dimuon final state is motivated from the fact that muons provide a clean signature in the detector due to low backgrounds.

This thesis is organized as follows. Chapter 2 gives a description of the SM. The theoretical motivation to look for contact interactions is also given here followed by a description of the contact interaction models and a summary of previous experimental searches. The last section of this chapter discusses in detail the hadron collider physics

[2]The hierarchy problem refers to the problem of Higgs boson mass becoming quadratically divergent due to some very large quantum corrections. Within the SM, the solution would be to fine tune the Higgs mass so that it almost completely cancels the quantum fluctuations, but this is deemed unnatural by a majority of particle physicists.

and defines important terminology used throughout this thesis. Chapter 3 introduces the LHC and CMS. All sub-detectors of the CMS experiment are explained briefly with an emphasis on the Muon spectrometers. The muon performance and reconstruction efficiency of the CMS detector in 2011 is also briefly discussed. Chapter 4 focuses on how various compositeness models are implemented in the PYTHIA Monte Carlo event generator and gives important technical details where necessary. Chapter 5 gives details about the full 2011 CMS data sample used in this analysis. This chapter also discusses the muon selection criteria adopted in this analysis to reduce the number of mis-reconstructed muons, cosmic-ray muons, and muons from hadronic decays.

Chapter 6 discusses how the expected signal and SM dimuon background samples are simulated. This chapter also discusses how the next-to-leading order quantum corrections are accounted for in the simulated samples. A detailed description of the SM Drell-Yan (DY) process is given here as it forms the basic physics process for this analysis. A brief summary of how the SM DY, SM non-DY and cosmic backgrounds are dealt with is also given. Chapter 7 gives details about systematic uncertainties that affect the contact interaction analysis. Uncertainties coming from experimental and theoretical sources are summarized here. Chapter 8 gives the statistical analysis technique that is used to derive results for the contact interaction search. First, the agreement of data with SM predictions is tested. As no significant deviation is observed, the rest of the chapter discusses how the 95% confidence level exclusion lower limits are set on the new physics energy scale Λ using a modified-frequentist approach. Finally, Chapter 9 provides a summary of the results and concludes the thesis.

Chapter 2

Theory

This chapter starts with a review of the standard model of particle physics. Some of the important open questions which the standard model (SM) fails to answer are addressed in Section 2.2 as a motivation for searching for new physics that lies beyond the SM. Section 2.2.1 introduces contact interactions as an effective framework for new physics searches along with a review of previous searches. Finally, Section 2.3 discusses the physics of proton-proton collisions.

2.1 The standard model

The standard model of particle physics is the most successful gauge theory to date that incorporates all of the known fundamental particles, namely, the quarks, leptons, and the gauge bosons and describes the electromagnetic (EM), weak and strong interactions between them. The fourth basic force of nature, gravity, is not yet integrated into the theoretical framework of the standard model. The SM is based on the fundamental concepts of relativistic quantum field theory (QFT), in which particles are represented by relativistic fields, and evolved over the years following various experimental observations. The electroweak and QCD theory of the SM is represented by the gauge symmetry group

$SU(3)_C \times SU(2)_L \times U(1)_Y$, conserving the color charge, weak isospin, electric charge and weak hypercharge.

2.1.1 Introduction to the standard model

All fundamental particles of the SM are assumed to be elementary, or point-like, with no internal structure. All observed particles in nature are categorized either as *fermions* (with half-integer spin: $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots$) or *bosons* (with integer spin: 0, 1, 2, . . .). While fermions form the basic building blocks of matter around us, the interactions between them are mediated by gauge bosons[1]. For each particle, there is an associated antiparticle with the same mass and spin but opposite charge.

Within the SM, quarks and leptons are fundamental fermions. There are six quarks (u, c, t, d, s, b) and six leptons $(e^-, \mu^-, \tau^-, \nu_e, \nu_\mu, \nu_\tau)$ with corresponding antiquarks $(\bar{u}, \bar{c}, \bar{t}, \bar{d}, \bar{s}, \bar{b})$ and antileptons $(e^+, \mu^+, \tau^+, \bar{\nu_e}, \bar{\nu_\mu}, \bar{\nu_\tau})$. Quarks and leptons are generally grouped into three generations (or "families") with each family consisting of two quarks and two leptons (charged lepton and its own neutrino). Figure 2.1 shows the three generations of quarks and leptons along with their masses, spin and electric charge. Note that quarks carry fractional charges and the top quark is the heaviest of all the quarks, having a mass of 172.9 ± 1.5 GeV/c^2 [19]. Due to *color confinement* [3], which is explained later in this section, quarks cannot exist freely and tend to group together to form *hadrons*. Hadrons are strongly interacting particles which are further categorized into baryons (formed from the three quark combination, qqq) and mesons (formed from the quark-antiquark combination, $q\bar{q}$).

Unlike classical scattering processes where the interaction at a distance occurs via *force fields*, in quantum theory, collisions occur via exchange of *quanta* associated with the particular type of interaction. In the SM, there are *four* types of force carriers that

[1] A gauge boson is a bosonic particle that carries any of the fundamental interactions of nature. In other words, a force carrier.

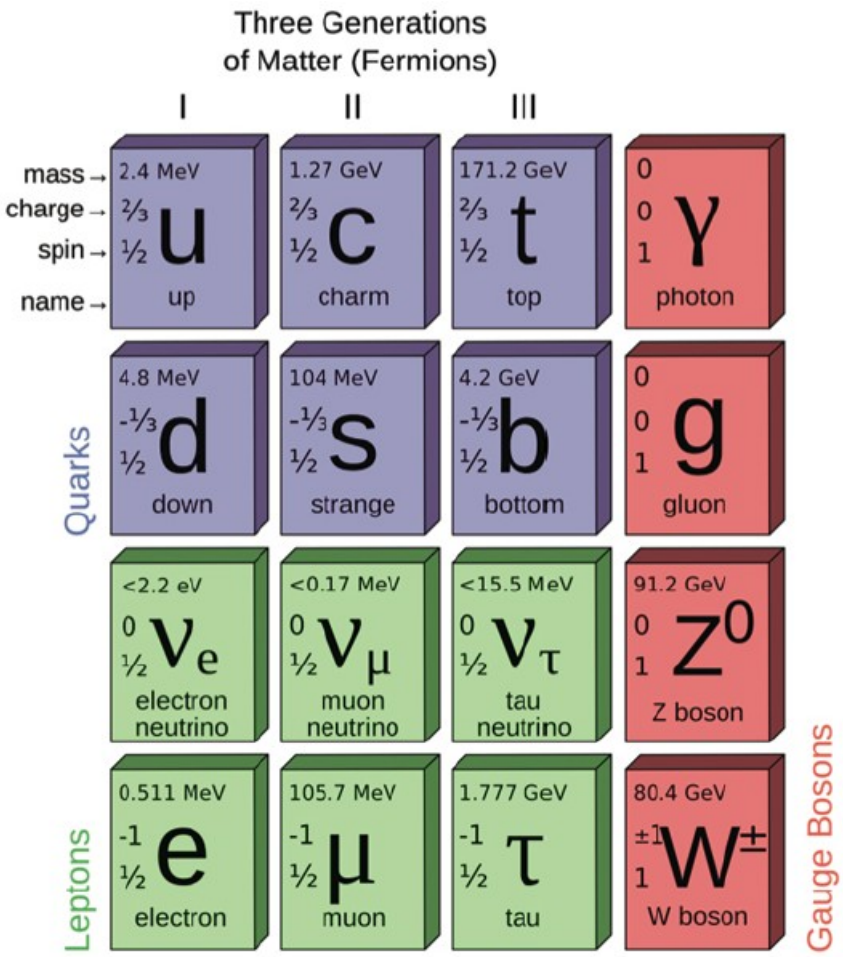


FIG. 2.1. Fundamental particles in the SM [1]. Note that this figure does not show the SM Higgs boson.

mediate the fundamental interactions: photon, gluon, $W^{\pm}$ and Z^o bosons (see Figure 2.1) which mediate electromagnetic, strong and weak interactions respectively. The force carriers act as propagators transmitting information from one fermion to another and are represented as internal lines in Feynman diagrams[2]. In general, each boson exchange can be characterized by a *propagator* term proportional to $(q^2 \pm m^2)^{-1}$ where q^2 is the scalar product of the interaction 4-momentum $(E/c, p_x, p_y, p_z)$ with itself and m is the mass of the exchanged boson.

Electromagnetic interactions

All charged particles undergo EM interactions mediated by photons, which themselves carry no electrical charge. Furthermore, photons have zero mass. As a result, the electromagnetic force is effective over long distances, or long-range. The EM coupling constant specifying the strength of the interaction between charged particles and photons is given by the fine structure constant α (in units of $\hbar = c = 1$),

$$\alpha = \frac{e^2}{4\pi} \simeq \frac{1}{137} \tag{2.1}$$

The EM interaction is governed by the theory of quantum electro-dynamics (QED) and is responsible for holding the electrons and protons together inside the atom.

Strong interactions

Strong interactions are mediated by gluons which are neutral massless particles like the photons. The strong coupling constant α_s is given by (in units of $\hbar = c = 1$),

$$\alpha_s = \frac{g_s^2}{4\pi} \simeq 1 \tag{2.2}$$

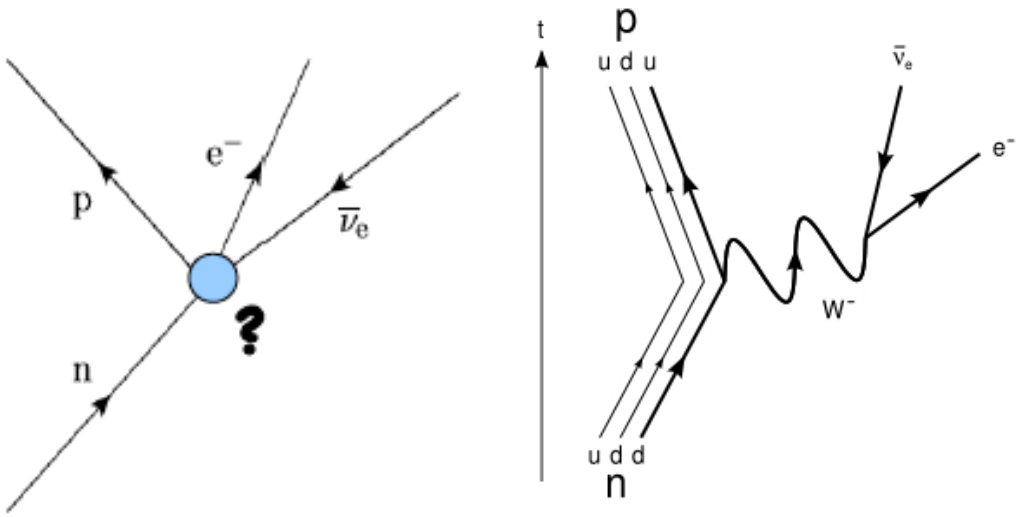
When compared with the EM coupling constant α, α_s is approximately 100 times stronger. Strong forces are responsible for binding the nucleons together to form the nucleus, while

[2]Please note that it is possible to have Ws, Zs and photons represented as external lines in Feynman diagrams as part of initial and final state processes.

inside the nucleons (and other hadrons) they hold the quarks together. Strong interactions take place between constituent quarks that make up the hadrons and are described by the theory of quantum chromodynamics (QCD). In this theory, each quark is assigned a "color charge" (here, color is just a name for an internal degree of freedom and has nothing to do with the real life colors). There are three color charges, blue (B), red (R) and green (G). Quarks carry one of these colors and anti-quarks, the corresponding anticolors (antiblue ($\bar{B}$), antired ($\bar{R}$) and antigreen ($\bar{G}$)). Gluons also carry color and can interact with themselves. Due to the gluon self-interaction, at low q^2, with the addition of higher and higher-order gluon loops, the theory becomes uncalculable. At small distances, or equivalently high q^2, the chromodynamic binding force between quarks weakens. This is called *asymptotic freedom* [6]. Conversely, at large distances (or low q^2), the binding force strengthens as the color field becomes stressed and more and more gluons are spontaneously created. This is known as *color confinement* and is the reason why quarks cannot exist independently and cluster together to form hadrons. Experimentally, there is no color asymmetry and all observed particle states are *colorless*.

Weak interactions

Weak interactions take place between all known fermions and are responsible for the radioactive decay of sub-atomic particles. Weak interactions are different from other interactions in many respects. Unlike the force carriers of EM and strong interactions which are massless, the mediators of weak interactions, $W^{\pm}$ (charged) and Z^o (neutral) bosons, are significantly massive. The masses of $W^{\pm}$ and Z^o are around 81 GeV/c^2 and 91 GeV/c^2, respectively. The large masses of W and Z bosons account for the very short-range of the weak interaction. Another unique feature of weak interactions is that the exchange of $W^{\pm}$ bosons results in change of flavor of quarks or leptons. This is called a "charged-current" interaction while Z boson exchange does not cause this and hence is called a "neutral-current" interaction. Also, the weak interaction is the only interaction



(a) Fermi 4-point CI-beta decay

(b) Beta decay (now)

FIG. 2.2. Fermi's β-decay, before (left) and now (right) [2].

that is known to violate *parity* symmetry (a transformation reversing the spatial coordinates leaves the interaction unchanged) and *CP* symmetry, a combination of charge conjugation, switching particles with antiparticles, and spatial inversion. Furthermore, $W^{\pm}$ bosons couple only to left-handed fermions (fermions whose direction of motion is opposite to the direction of their spin).

2.1.2 Origin of the weak force

In the 1930's, Enrico Fermi first introduced the weak force as a *four-fermion contact* interaction (CI) (a force with no range) to describe nuclear β-decay, $n \rightarrow pe^{-}\bar{\nu}_e$ [7]. Fermi's idea of β-decay is shown on the left side of Figure 2.2. However, now the weak force is understood to have a very short range, mediated by massive gauge bosons. It was also recognized later that the fundamental interaction in β-decay was not between the proton and neutron themselves, but between their constituents and that it is mediated by the charged W^{-} boson (see Figure 2.2(b)).

Fermi's idea of *pointlike* β-decay can also be understood by looking at the weak interaction scattering amplitude, $f(q^2)$, given by,

$$f(q^2) \propto \frac{g_w^2}{(q^2 + M_W^2)} \tag{2.3}$$

where g_w refers to the weak interaction coupling constant and q^2 is the scalar product of the interaction 4-momentum with itself. For $q^2 \ll M_W^2$, the scattering amplitude is independent of q^2, meaning, the weak interaction is point-like, which is what Fermi postulated.

2.1.3 The Electro-weak model

The first attempt towards the SM was made by Glashow in 1961, who discovered a way to unify the electromagnetic and weak interactions [8]. He postulated that the EM and weak forces are manifestations of a single force, the electroweak force. Mathematically, the EM interaction is represented by the $U(1)_Y$ gauge group, with *weak hypercharge* Y as the generator of the group. The weak interaction is described by the $SU(2)_L$ group, generated by *weak isospin* I and interacting only with left-handed particles (as indicated by the subscript L). Glashow's electroweak theory unifies the above into a single gauge group $SU(2)_L \times U(1)_Y$. While the electroweak symmetry is well behaved above the electroweak scale, where all particles are predicted to be massless, below this scale, some other mechanism is needed to give the weak gauge bosons mass by breaking the $SU(2)_L \times U(1)_Y$ symmetry. This was achieved by incorporating the *Higgs mechanism* into Glashow's electroweak theory giving rise to what is known as the standard model today. The spontaneous symmetry breaking by the Higgs mechanism, postulated by Steven Weinberg [9] and Abdus Salam [10] in 1967, gives masses to the gauge bosons (and all other massive particles in the SM). The experimental evidence for electroweak interactions was first established in 1973 with the discovery of neutral currents by the

Gargamelle collaboration [11, 12] and later in 1983 with the discovery of the W and Z gauge bosons by the UA1 [20, 21] and UA2 [22, 23] collaborations at the Super Proton Synchrotron (SPS) at CERN. While the SM may look like the whole story, its key building block, the Higgs boson, is yet to be confirmed by experiments. It is anticipated that its existence will be confirmed in the near future by the experiments at the Large Hadron Collider at CERN.

2.2 Physics beyond the standard model

Despite its tremendous success in explaining particle phenomena to date, the SM suffers from many deficiencies. There are many theoretical and observational motivations for extensions to the SM. Although general relativity provides insights into the nature of gravity, there is no field theory extension to the SM that incorporates gravity. There are cosmological evidences that 96% of the gravitational matter is made up of dark energy and dark matter, while the known hadronic matter, described by the SM, represents only 4% of it. The SM does not give any explanation for this. The SM also cannot explain the observed magnitude of matter-antimatter asymmetry. Also, neutrino oscillation experiments have shown that neutrinos do have mass while the SM assumes them to be massless. There are also some features within the SM that remain arbitrary, implying a lack of understanding, suggesting the existence of a more comprehensive model. The reason why there are exactly 3 generations of fundamental particles is not explained by the SM. Also, the mass differences between particles and particle families are not explained. For example, why the top quark mass is significantly higher than those of the other fermions is not understood. The SM depends on 19 free parameters (the 3 coupling constants, 9 charged fermion masses, 4 CKM matrix parameters, QCD vacuum angle θ_{QCD} and 2 Higgs model parameters). The SM cannot predict the values of these parameters and they must be inferred from experiments. All particles predicted by the

SM have been experimentally observed except for the Higgs boson (which is yet to be confirmed by experiments at the LHC), through which the fundamental particles obtain mass. Even if the LHC confirms the existence of Higgs boson, its mass suffers from the hierarchy problem arising from radiative corrections. If the new boson discovered at the LHC is not a SM Higgs boson, then the electroweak symmetry breaking and subsequent mass generation must involve physics beyond the standard model.

Physics beyond the SM is one of the most active areas of research pursued both in theoretical and experimental particle physics. Various extensions of the SM are being widely tested at the particle experiments. The popular theories that lie beyond the SM include supersymmetry (SUSY), string theory and extra dimensions. Although less popular, but interesting are the quark and lepton compositeness searches.

2.2.1 Contact Interactions

Contact interactions (CI) are considered as an effective framework for new physics searches. Contact interactions have a long history, dating back to 1930's when Fermi first used it to explain β-decay [7] long before the discovery of the W boson. Similarly, one can write a Lagrangian describing a new vector interaction occurring at an energy scale Λ, without knowing the exact intermediate process. Λ can be much higher than the achievable center-of-mass energy at the LHC, nonetheless, its effects can be detected at energies well below Λ. Experimentally, contact interactions appear as a "non-resonant" enhancement of the expected dilepton (or diquark) events at high invariant masses. The Lagrangian density for four-fermion contact interactions with dimuons in the final state is given by [16],

$$\begin{aligned} L_{ql} = (g_0^2/\Lambda^2)\{\eta_{LL}(\bar{q}_L\gamma^\mu q_L)(\bar{\mu}_L\gamma_\mu\mu_L) &+ \eta_{LR}(\bar{q}_L\gamma^\mu q_L)(\bar{\mu}_R\gamma_\mu\mu_R) \\ +\eta_{RL}(\bar{u}_R\gamma^\mu u_R)(\bar{\mu}_L\gamma_\mu\mu_L) &+ \eta_{RL}(\bar{d}_R\gamma^\mu d_R)(\bar{\mu}_L\gamma_\mu\mu_L) \\ +\eta_{RR}(\bar{u}_R\gamma^\mu u_R)(\bar{\mu}_R\gamma_\mu\mu_R) &+ \eta_{RR}(\bar{d}_R\gamma^\mu d_R)(\bar{\mu}_R\gamma_\mu\mu_R)\} \end{aligned} \tag{2.4}$$

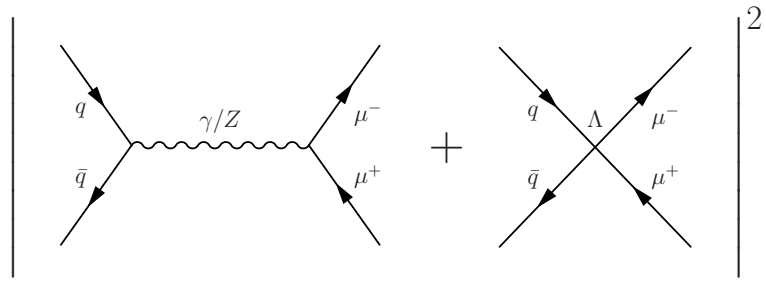


FIG. 2.3. Interference between DY and CI amplitudes resulting in dimuon final states [13].

where, $q_L = (u, d)_L$ is a left-handed quark doublet, u_R and d_R are right-handed quark singlets, and μ_L and μ_R are the left- and right-handed muons. By convention, $g_0^2/4\pi = 1$. The value of η gives the sign of the interference of new physics with the SM Drell-Yan (DY) process [17] with destructive ($\eta = +1$) and constructive ($\eta = -1$) interference possibilities. The parameters η_{ij}, where i and j are left (L) or right (R) define the helicity structure of the new interaction. Lambda represents the compositeness energy scale and is potentially different for each of the individual terms in the Lagrangian, so lower limits on Λ are set separately for the individual currents in Eq. 2.4. As illustrated in Fig. 2.3, standard model DY dimuon production and CI dimuon production have the same final state, so their scattering amplitudes add. The observed differential cross section can be described as

$$\frac{d\sigma}{dm}(\Lambda) = \frac{d\sigma}{dm}(DY) - \eta\frac{I(m)}{\Lambda^2} + \eta^2\frac{C(m)}{\Lambda^4} \tag{2.5}$$

where m is the invariant dimuon mass, $I(m)$ corresponds to the product of DY and contact interaction amplitudes, and $C(m)$ corresponds to a pure contact term.

There can be several different interpretations of the energy scale, Λ, depending on the new physics model. For this thesis, the new physics model is chosen to be the quark and lepton compositeness with left-handed currents, more details of which are given in the next section.

2.2.2 Quark and Lepton Compositeness

The fact that SM fails to explain the variety of observed quark and lepton flavors and their masses suggest that they may in fact be composed of more fundamental constituents, often referred to as "preons" [14, 15], interacting through a new strong gauge interaction called *metacolor*. Below a given interaction scale Λ, the effect of the metacolor interaction is to bind the preons into metacolor-singlet states. As mentioned in the previous section, for parton interaction $\hat{s}$ values that are much less than the Λ scale, the metacolor force will manifest itself in the form of a flavor-diagonal contact interaction [18].

For this analysis, the left-left isoscalar model (LLIM) of quark compositeness, where all the initial state quarks are assumed to be composite, is chosen. This model corresponds to the first term of L_{ql} in Eq. 2.4 ($\eta_{LL} = \pm 1; \eta_{LR} = \eta_{RL} = \eta_{RR} = 0$) and is the conventional benchmark for CI studies in the dilepton channel [19]. More details on the implementation of the LLIM within the PYTHIA Monte Carlo program will be given in Chapter 4.

2.2.3 Previous results on Λ

Previous searches for quark and lepton compositeness, in dilepton and dijet final states, have all resulted in exclusion lower limits on the compositeness energy scale Λ. These include studies from Large Electron-Positron Collider (LEP) [24–28], Hadron Electron Ring Accelerator (HERA) [29,30], the Tevatron [31–36], and recently from the A Toroidal LHC ApparatuS (ATLAS) [37–39] and CMS [55] experiments at the CERN Large Hadron Collider. The best limits in the left-left isoscalar model for dimuon final states are currently $\Lambda > 4.9$ TeV for constructive interference and $\Lambda > 4.5$ TeV for destructive interference at 95% C.L. [38].

2.3 Physics of p-p collisions

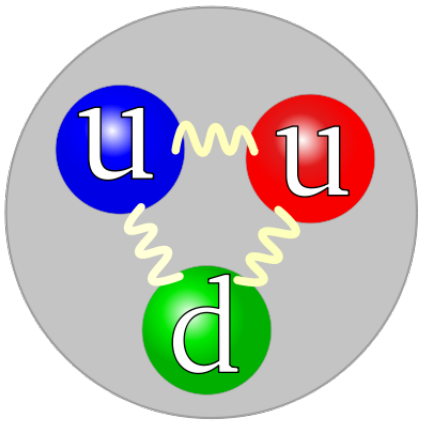


FIG. 2.4. Simplified parton model of the proton [4].

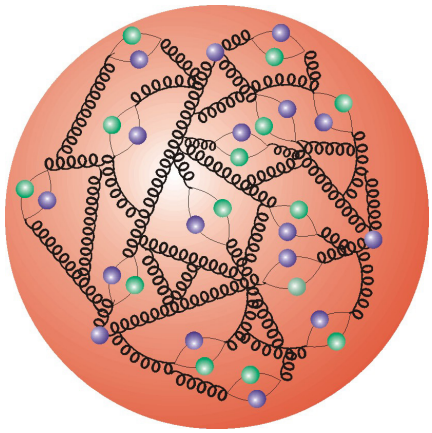


FIG. 2.5. Detailed partonic structure of the proton [5].

At the LHC, hadron collisions occur at very high energies of the order of TeV. In order to understand the scattering cross section of the p-p collisions, it is important to understand the internal structure of the protons. The parton model of hadrons was proposed by Richard Feynman as a way to analyze hadron collision data [42] [43]. Here,

parton refers to quarks and gluons. In 1969, the composite structure of hadrons was revealed in the e-p deep inelastic scattering (DIS) experiments [41]. In a simplified picture, protons are made up of three quarks (uud) referred to as valence quarks (Figure 2.4), but the actual structure of the proton is far more complex than this. In addition to the valence quarks (that make up the proton's quantum numbers), the proton also contains a large number of *virtual* quark-antiquark pairs (commonly known as *sea quarks*) and gluons[3] (Figure 2.5). So, in the framework of the parton model, the p-p collisions are actually interactions occurring between two bags of partons. It is important to note here that the only way Drell-Yan ($q\bar{q} \rightarrow Z/\gamma^* \rightarrow l^+l^-$) [17] processes can occur at the LHC is by utilizing an anti-quark from the available sea quarks. To predict the rates of various processes occurring via the partonic constituents of the proton, a set of parton distribution functions (PDFs) $F_a(x_A, Q^2)$ is defined. PDFs represent the probability densities that a parton 'a' carries a fraction x_A of proton A's longitudinal momentum when probed at a momentum transfer scale Q^2. Due to the non-perturbative nature of QCD at low Q^2, the PDFs are best determined by global fits to the DIS e-nucleon experimental data (*e.g.,* HERA) and also from $p\bar{p}$ collision data (*e.g.,* Tevatron). The PDFs are available for leading order (LO), next-to-leading order (NLO) and next-to-next-to-leading order (NNLO) calculations in the strong coupling constant, α_s.

For the LHC, the two main groups that produce the PDF sets are CTEQ [51] (from the CTEQ Collaboration) and MSTW[4] [44]. Figures 2.6 and 2.7 illustrate the MSTW NLO and NNLO PDFs of the proton, respectively [44].

The presence of the so-called 'proton sea' (sea quarks and gluons) makes it very difficult to calculate the full production cross section, but the QCD factorization theorem [46] plays a remarkable role in simplifying the cross section calculations for hadron-hadron

[3]Sea quarks are formed when gluons of the proton color field split into $q\bar{q}$ pairs. The number of sea quarks in a proton are predicted based on this and also on $q\bar{q}$ pairs annihilating to produce gluons.

[4]Previously called MRST which stands for A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne

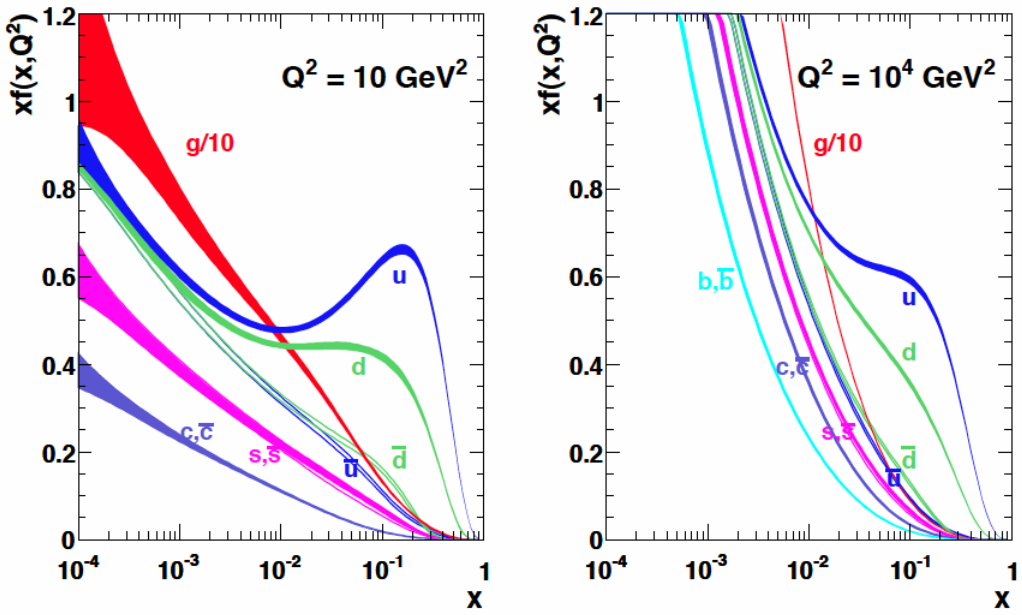


FIG. 2.6. MSTW 2008 NLO PDFs (at 68% C.L.) of the proton for a momentum transfer of 10 $(\mathrm{GeV/c})^2$ (left) and 10^3 $(\mathrm{GeV/c})^2$ (right).

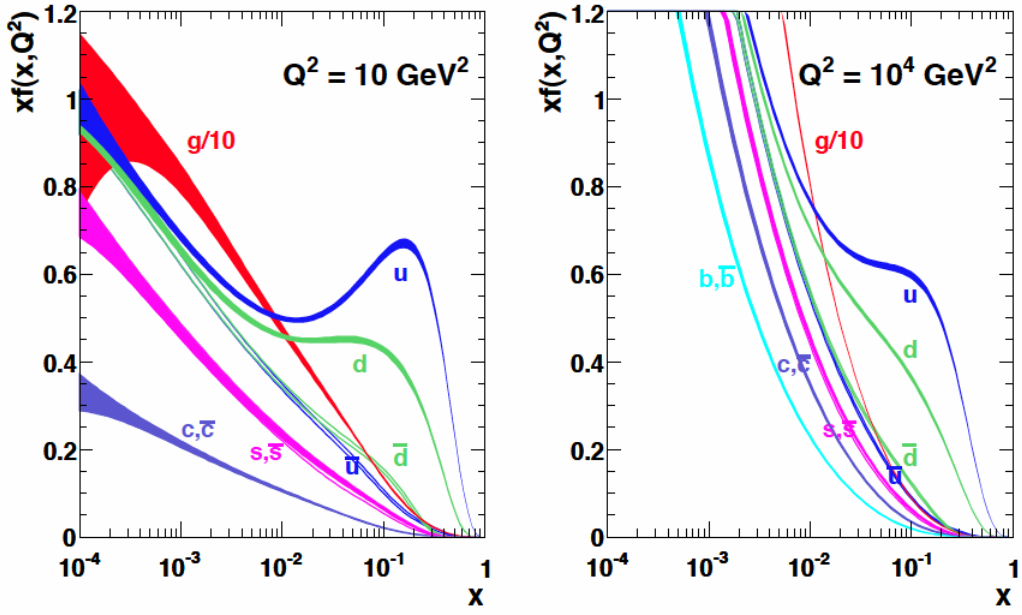


FIG. 2.7. MSTW 2008 NNLO PDFs (at 68% C.L.) of the proton for a momentum transfer of 10 $(\mathrm{GeV/c})^2$ (left) and 10^3 $(\mathrm{GeV/c})^2$ (right).

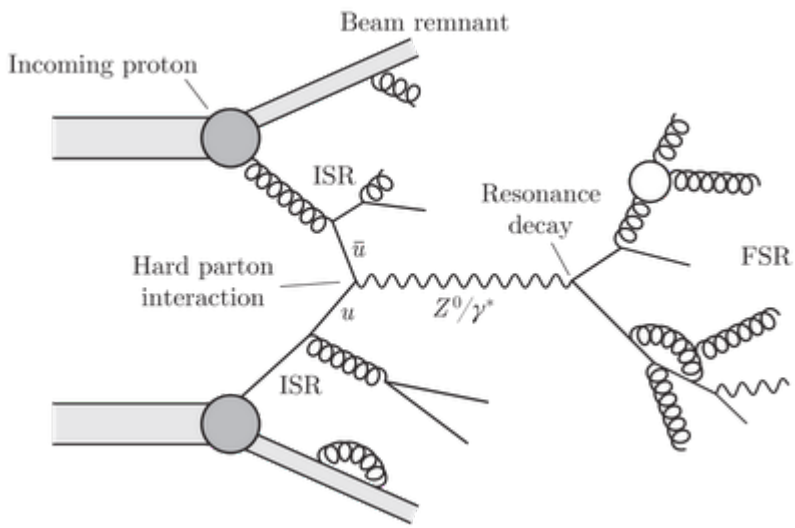


FIG. 2.8. p-p collision event illustrating the hard scattering, underlying event and the initial and final state radiations [45].

collisions. The theorem states that the total cross section can be separated into two parts: the hard scattering process between the two colliding partons and the PDFs for those partons. As an example, the Drell-Yan production cross section in a collision between two hadrons A and B at a given center-of-mass energy, $A + B \rightarrow Z^{\circ}/\gamma^{*} \rightarrow l^{+}l^{-} + X$, can be factorized as,

$$\sigma(A + B \rightarrow l^{+}l^{-} + X) = \sum_{a} \int dx_A dx_B F_a(x_A, Q^2) F_{\bar{a}}(x_B, Q^2) \sigma_{a\bar{a}\rightarrow l^{+}l^{-}}(Q^2) \qquad (2.6)$$

where the sum runs over all the possible types of parton, a $(= gluon, u, \bar{u}, d, \bar{d}, ...)$, x_A and x_B correspond to the proton's momentum fraction as carried by the partons, $\sigma_{a\bar{a}\rightarrow l^{+}l^{-}}$ gives the hard scattering cross section, and X corresponds to everything else that exists in the event from left over proton fragments. This is usually referred to as 'soft QCD scattering' or 'QCD underlying event'. In addition to the hard scattering, gluon and/or photon radiation is expected as a natural by-product due to the presence of color and/or charged objects in the initial and final states of a hadron collision. This is referred to

as initial and final state radiation (ISR and FSR). Figure 2.8 illustrates a typical hadron collision event. Specifically, it shows the resonance decay of a Z boson into $q\bar{q}$ pairs which are capable of radiating gluons and photons in their final state due to their color and charge fields. However, in the case of the DY process, while both photons and gluons can be radiated in the initial state, only photons are present in the FSR because of the absence of color field in the leptonic final state.

QCD underlying event, multiple interactions and pileup

The valence and sea quarks of the proton not actively participating in the initial state radiation or the hard scattering are often referred to as beam remnants or proton remnants (see Figure 2.8). These remnants are color connected to the hard scattering, roughly traveling in the same direction as their parent proton, and either undergo *hadronization* or form a *parton shower* due to the inherent color confinement. Here, hadronization refers to the process in which quarks cluster together into colorless states and form hadrons which may decay further. Hadronization usually produces what is called a low-energy jet, which is a cone of hadrons and other particles. Alternatively, the beam remnants can combine to form a strong color field, eventually decaying into $q\bar{q}$ pairs and gluons which in turn may radiate gluons and decay into more $q\bar{q}$ pairs and gluons and so on. This results in what is called a parton shower (PS). These other softer QCD processes along with the ISR and FSR are collectively referred to as the 'underlying event' in the collider jargon. Another interesting possibility is that several pairs of remnant partons can enter into simultaneous scatterings within a single p-p collision causing multiple partonic interactions (MPI). To further complicate this, in high-energy colliders like the LHC there is an increased probability that more than one hadron-hadron collision can occur in a single bunch crossing as the luminosity is increased, further building up the particle production activity. In the collider jargon, this is referred to as 'pileup'. Effects of pileup need to be carefully considered in any analysis that uses hadron collision data.

Given this rather complex particle production happening in a hadron collision, it is important that the event generators used to simulate signal and background physics processes handle this efficiently. The particle generators used in this analysis are PYTHIA [47] [48] and MC@NLO [49]. Although different, both Monte Carlo generators can handle PS and hadronization along with ISR, FSR and all hard scattering processes. PYTHIA predicts QCD hard scattering rates with LO[5] calculations while MC@NLO, as the name suggests, handles scattering at NLO accuracy. Event generation with MC@NLO consists of two steps. Step 1 generates the hard scattered kinematic information at NLO. In the next step, this information is passed on to another event generator HERWIG [50] to handle PS and hadronization. HERWIG stands for Hadron Emission Reactions With Interfering Gluons.

Not all BSM hard processes are implemented in all Monte Carlo generators. For example, PYTHIA is the only Monte Carlo generator that has compositeness models built into it. In this analysis, PYTHIA is used to make contact interaction signal predictions at LO accuracy (as PYTHIA is a LO generator). To include QCD NLO corrections in the signal samples, DY predictions from PYTHIA and MC@NLO are compared to deduce NLO corrections with the underlying assumption that they can be applied to contact interaction samples as well, as DY and CI processes have the same initial states. In addition to QCD NLO corrections, there are also NLO electroweak (QED) corrections that need to be applied. A detailed discussion on how the signal and background samples are predicted is given in Chapter 6.

PYTHIA (version 6.4 as used in this analysis) is run with PDF set CTEQ6L1 [51] which is at LO. MC@NLO (version 3.4 as used in this analysis) uses NLO PDF set CTEQ6M [51]. HERWIG version 6.4 is used with MC@NLO to model PS and hadronization.

[5]In QCD, partonic cross sections can be evaluated perturbatively and expressed as power series in the strong coupling constant, α_s (Eq. 2.2). The zeroth power of α_s corresponds to LO, first power in α_s corresponds to NLO and so on. Including higher powers of α_s in a cross section calculation increases the accuracy of the prediction.

Chapter 3

Experimental setup

This chapter introduces the large hadron collider (LHC) accelerator and compact muon solenoid (CMS) detector experiments located at CERN. A detailed description of various sub-detectors of the CMS detector is given in Section 3.2 with an emphasis on the muon spectrometer. The CMS trigger and data acquisition system is also discussed briefly.

3.1 The Large Hadron Collider

The LHC [53], which is located at CERN spanning the borders of Switzerland and France, is the most powerful particle accelerator and collider in the world. It is located about 100 m underground in a circular tunnel of circumference 26.7 km. The tunnel was originally constructed for the Large Electron–Positron (LEP) collider, that stopped running in November 2000 after 11 years of successful operation. The LHC was designed to collide proton beams at a center-of-mass energy of 14 TeV at a peak luminosity of 10^{34} cm^{-2}s^{-1} and lead ions at 2.76 TeV at a peak luminosity of 10^{27} cm^{-2}s^{-1}.

3.1.1 Layout

The LHC collider consists of eight arcs and two adjacent parallel beam pipes that intersect at four beam crossing points 1, 2, 5 and 8, which contain four experiments (see Figure 3.1). The LHC is a synchrotron machine that accelerates two counter-rotating proton beams in separate beam pipes. Protons are grouped into ellipsoidal *bunches* to form a *proton beam*. The maximum number of bunches per beam that can be stored in the LHC ring is 2808. During nominal operation, proton bunches are 25 ns apart providing a collision rate of 40 MHz within a given bunch crossing. This corresponds to a design luminosity of 10^{34} $cm^{-2}s^{-1}$. The LHC relies on superconducting magnet technology for beam circulation and collimation. The LHC magnet system consists of 1232 superconducting dipole magnets to keep proton beams in a circular trajectory and 392 quadrupole magnets to keep the beams focused. Superfluid helium is used to cool the superconducting magnets to a temperature below 2° K.

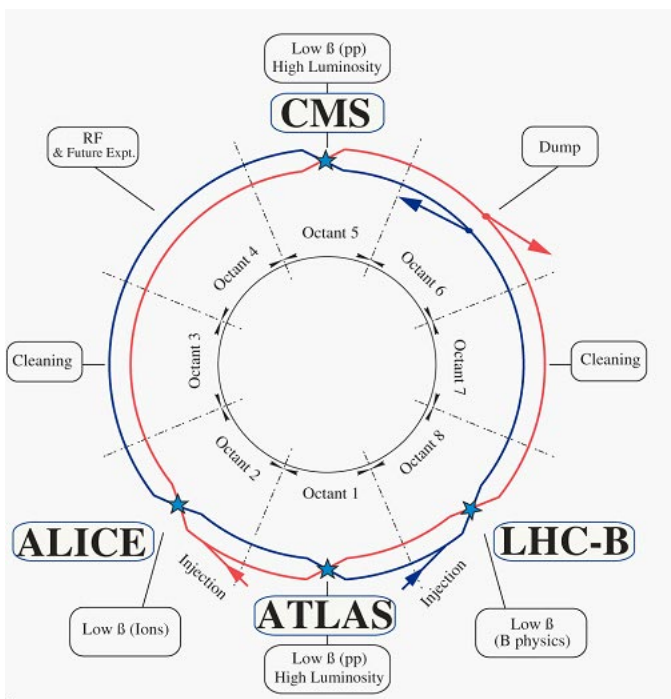


FIG. 3.1. Schematic layout of the LHC collider tunnel [53].

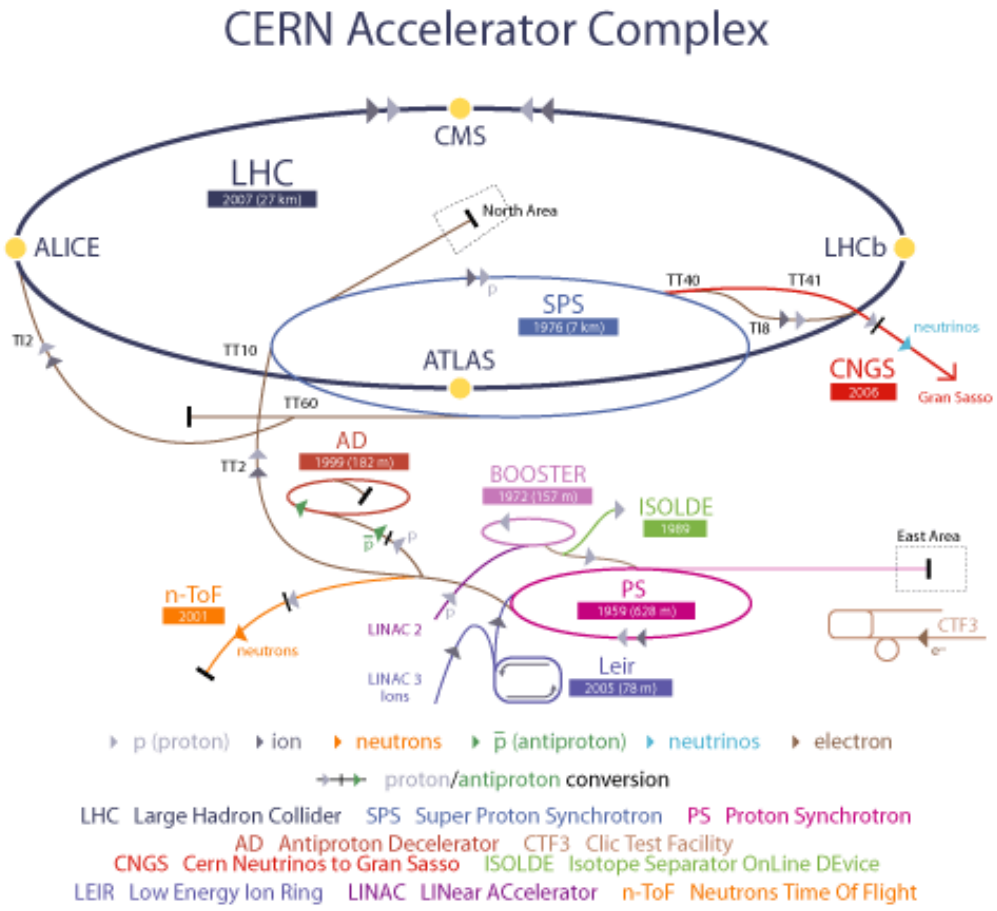


FIG. 3.2. Layout of the CERN accelerator complex [60].

The full acceleration process at the LHC occurs in stages, where at each stage, particles are prepared by a series of operations that successively increase their energy prior to being fed to the main accelerator (see Figure 3.2). Protons, first produced in a Duoplasmatron source [61], are injected into the linear particle accelerator (Linac2). There they are accelerated to 50 MeV and then transferred to the proton synchrotron booster (PSB). The PSB accelerates the protons to 1.4 GeV. From there protons are injected into the proton synchrotron (PS) where they are accelerated to 26 GeV and are then fed to the super proton synchrotron (SPS) which increases their energy to 450 GeV before they are injected into the LHC ring. Here the proton bunches are accumulated and accelerated to their peak energy.

3.1.2 Experiments at the LHC

The LHC accelerator complex hosts six experiments to explore particle phenomena at the new energy frontier. The experiments are located at points where the two beams of the LHC, traveling in opposite directions, collide head on.

- The two large experiments, a toroidal LHC apparatus (ATLAS) [54] and CMS [55], are general-purpose detectors that are built to study both proton (p–p) and lead ion (Pb–Pb) collisions. The main goal of these two experiments is to search for the Higgs boson, to explore physics beyond the standard model and to study the quark-gluon plasma. Having two independently designed detectors is vital for cross-confirmation of any new discoveries made. As shown in Figure 3.1, the ATLAS and CMS detectors are located at beam crossing points 1 and 5, respectively.

- The large hadron collider beauty (LHCb) experiment [56] is specially designed to investigate matter-antimatter asymmetry by studying the properties of the beauty quark (or 'b quark') produced in proton-proton collisions. The LHCb detector is located at beam crossing point 8 (Figure 3.1).

- The a large ion collider experiment (ALICE) [57] is a dedicated heavy-ion detector. It is built mainly to study the quark-gluon plasma, the hadronic matter in extreme temperature and density conditions, which probably existed just after the big bang when the Universe was still extremely hot. The ALICE detector is located at beam crossing point 2 (see Figure 3.1).

- The total elastic and diffractive cross section measurement (TOTEM) [58] and large hadron collider forward (LHCf) [59] experiments are specially designed to study physics processes in the "forward region" (the region very close to the particle beam) of p–p or Pb–Pb collisions. The LHCf experiment, built close to the ATLAS

experiment, explores the origin of high energy cosmic rays whereas the TOTEM experiment, built close to the CMS detector, is dedicated to precisely measure the p–p interaction cross section and for in-depth study of the proton structure.

3.1.3 Operation and performance

After an unsuccessful attempt to run the machine in September 2008, LHC resumed its operation in November 2009 by successfully circulating the beams. On 20 November 2009, for the first time low–energy proton beams (with 450 GeV energy per beam) circulated in the LHC tunnel. Soon after, on 30 November, LHC became the world's highest energy particle accelerator, achieving collisions with 1.18 TeV energy per beam. Later on 30 March 2010, LHC broke its own record, by colliding proton beams, with 3.5 TeV energy per beam, at a center-of-mass energy of 7 TeV. The 2010 proton run ended on 4 November 2010 (followed by the heavy-ion runs) and LHC resumed its operation for proton collisions in January 2011. The 2011 run ended in October 2011 followed by a short technical stop. In 2012, LHC has been colliding proton beams at an increased center-of-mass energy of 8 TeV (with 4 TeV energy per beam) making the new physics searches even more sensitive.

The LHC has shown impressive performance during 2010 and 2011. The instantaneous luminosity delivered depends on the LHC *filling* scheme which corresponds to a specific mode of operation for the machine. A *fill* is characterized by many variants, the important ones being the total number of proton bunches per beam, bunch spacing, and expected number of colliding bunches at various interaction points. In 2010, the peak instantaneous luminosity reached $L = 2 \times 10^{32}$ cm^{-2}s^{-1}, with 368 bunches per proton beam and accumulated about 40 pb^{-1} of collision data. In 2011, LHC recorded a peak luminosity of $L = 3.5 \times 10^{33}$ cm^{-2}s^{-1} (which is a ten fold increase compared to 2010), with 1380 bunches of protons per beam (see Figure 3.3). The integrated luminosity delivered to various experiments as a function of the *fill* number for the 2011 data taking is shown

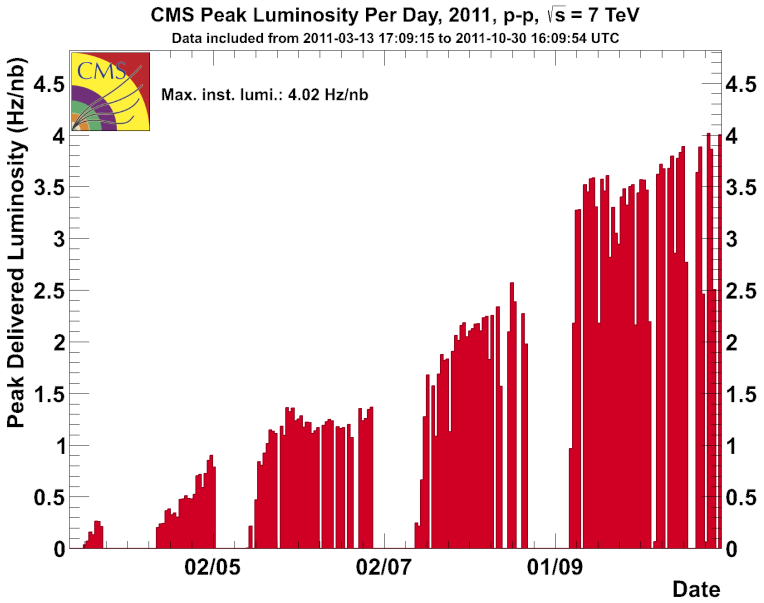


FIG. 3.3. Maximum instantaneous luminosity per day delivered to CMS in 2011 for p–p runs at $\sqrt{s} = 7$ TeV.

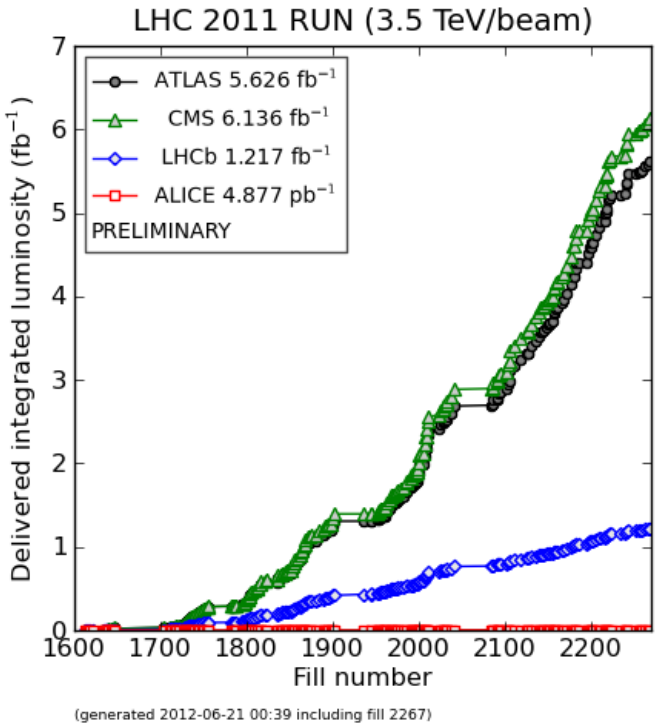
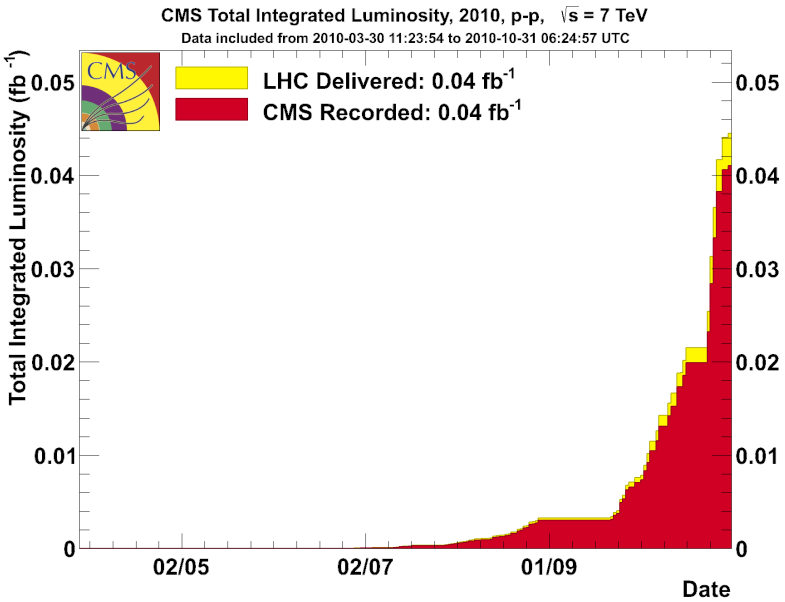
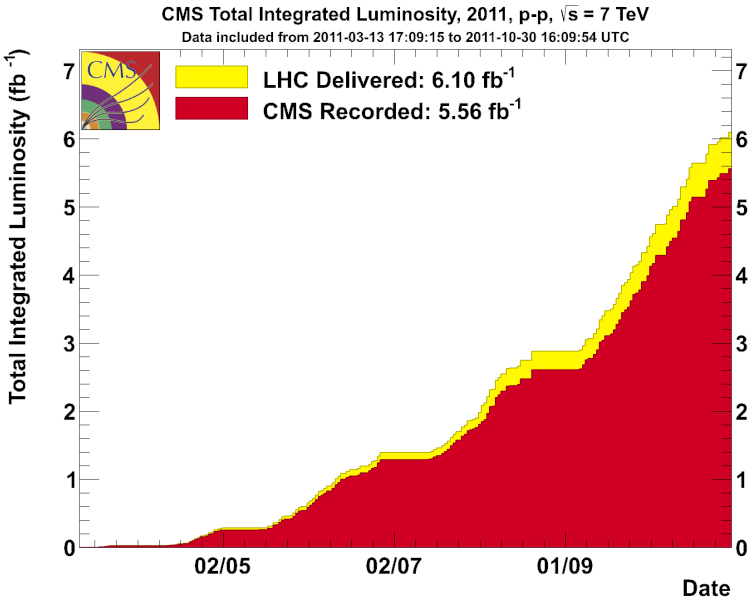


FIG. 3.4. Integrated luminosity vs. fill number for 2011 p–p runs.



(a) 2010 data taking



(b) 2011 data taking

FIG. 3.5. Integrated luminosity delivered to (yellow), and recorded by CMS (red) for 2010 and 2011 p–p runs at $\sqrt{s} = 7$ TeV.

in Figure 3.4. The data accumulated in 2011 is about 6 fb^{-1}, which is roughly a hundred times more than what was accumulated in 2010 (see Figure 3.5 [62]). This analysis uses the full 2011 proton collision data as recorded by the CMS experiment.

3.2 The CMS detector

The CMS detector is a multi-purpose detector located at the LHC at CERN and capable of studying both proton-proton and heavy-ion collisions. It sits about 100 m underground close to the French village of Cessy, between Lake Geneva and the Jura mountains. This section provides a brief description of the design and construction of the CMS detector. A detailed description of the CMS experiment can be found elsewhere [55]. Since one of the primary motivations of the LHC is to search for the Higgs boson and explore physics at the TeV energy scale, the beam energy and design luminosity of the LHC have been chosen accordingly. This goal also requires a very careful design of the detector and relies on the detector's capability to reconstruct certain physics objects with precision. At the design energy, $\sqrt{s} = 14$ TeV, the total proton-proton cross section is expected to be roughly 100 mb. At design luminosity, this yields an event rate of approximately 10^9 collisions per second. The high event rate poses a formidable challenge to the detector design in terms of event selection, data storage and fast electronics (which require very good synchronization). The CMS detector is designed to meet the goals of the LHC physics program offering good particle identification and momentum resolution over a wide range of momenta along with good reconstruction efficiency. The distinguishing features of the CMS detector design are its compactness and the magnetic field configuration which provides large bending power resulting in precise momentum measurements of charged particles.

3.2.1 The CMS coordinate system

CMS uses a right-handed coordinate system, with the origin at the nominal collision point (the geometrical center of the detector), the y-axis pointing upwards (perpendicular to the LHC plane), the x-axis pointing radially inward (towards the center of the LHC ring), and the z-axis along the anti-clockwise beam direction (geographically towards the Jura mountains). In terms of polar coordinates, the azimuthal angle ϕ is measured from the x-axis in the $x - y$ plane and the polar angle θ is measured from the positive z-axis. In experimental particle physics, the Lorentz-invariant quantity *rapidity* (y) is commonly used to describe the geometrical coverage of the detector. In the relativistic limit, rapidity can be approximated by *pseudorapidity* (η). Rapidity and pseudorapidity are defined as follows,

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right); \qquad \eta = \frac{1}{2}\ln\left(\frac{|\mathbf{P}| + p_z}{|\mathbf{P}| - p_z}\right) \tag{3.1}$$

where E is the relativistic energy of the particle, $\boldsymbol{P}$ is the momentum vector of the particle and p_z is the particle's momentum along the beam direction (z axis). Psuedorapidity is commonly expressed in terms of the polar angle, θ, given by,

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]. \tag{3.2}$$

Figure 3.6 shows how the distributions of η and θ compare ($\theta = 0$ and $\eta = \infty$ correspond to the beam axis).

The kinematic variables measured by CMS are transverse momentum (p_T), η and ϕ which cover all of the phase space and can in turn be used to derive all other kinematic variables.

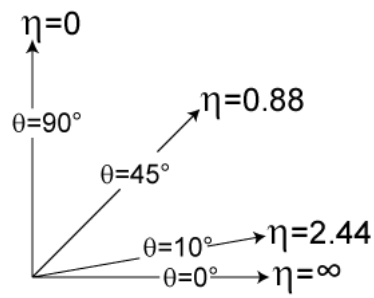


FIG. 3.6. As polar angle increases, pseudorapidity decreases. The "forward region" of a a collider detector corresponds to regions with high η values, typically $\eta > 3$.

3.2.2 Layout

The CMS detector is 21.6 m long with a diameter of 14.6 m and weighs 12 500 tons. A 4 tesla (T) superconducting magnet (13 m long and 6 m inner diameter) sits at the heart of the CMS detector providing a uniform magnetic field of 3.8 T (during data taking). The CMS detector is designed around the solenoidal magnet and is comprised of multiple sub-detectors: a pixel detector, a silicon tracker, electromagnetic and hadronic calorimeters and three types of muon detectors. The bore of the magnet is large enough to accommodate both the inner tracker and the calorimeters inside it. The magnetic field is closed by a 10 000 ton iron return yoke comprising 5 barrel wheels (to cover the length of the solenoid) and two endcaps (one on each end), composed of three layers each. Figure 3.7 shows the overall layout of the CMS detector. Integrated into the return yoke are the four muon stations to ensure full geometric coverage. In the barrel region, each muon station consists of several layers of drift tubes (DT) and resistive plate chambers (RPC), while in the endcap region, muon stations are comprised of cathode strip chambers (CSC) and RPCs.

Figure 3.8 shows a transverse slice of the CMS detector demonstrating how different particles interact with various layers of the detector leaving behind characteristic patterns, or 'signatures', allowing them to be identified. The detector layers are designed in such

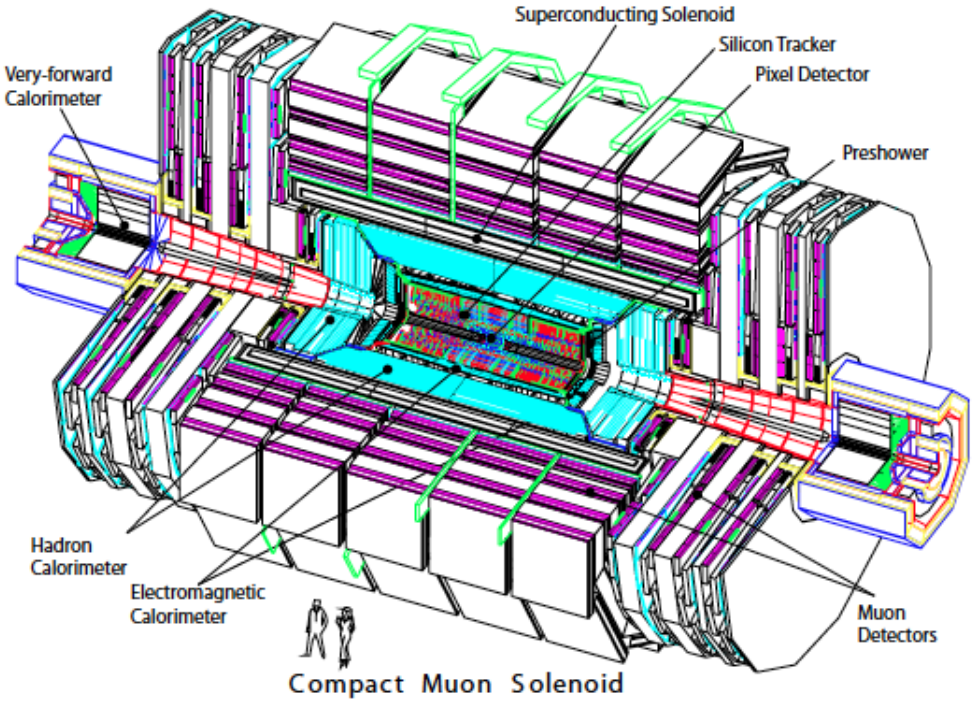


FIG. 3.7. Diagram of the CMS detector with a quadrant cut away to show the interior [55].

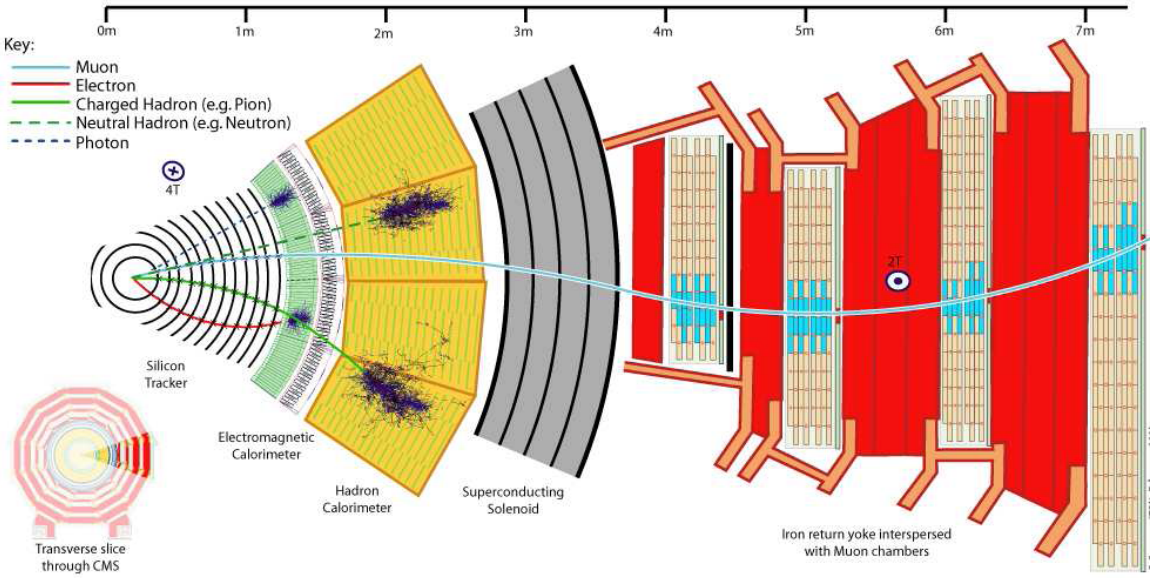


FIG. 3.8. Particle identification at CMS [52].

a way as to exploit the different properties of particles and measure their energy or momentum. The innermost layer (closest to the interaction point) is the silicon tracker comprised of silicon pixel and strip detectors which identify charged particle tracks by recording their positions. The magnetic field bends charged particles allowing one to measure the charge and momentum of a particle based on the curvature of its tracks. The next layer of the detector, the electromagnetic calorimeter (ECAL) measures the energy of electrons and photons with high precision. Surrounding the ECAL is the hadron calorimeter (HCAL) which measures the energy of hadrons and also the missing transverse energy (E_T) in an event. The muon detectors, which form the final layer of the CMS detector, are dedicated to identify particles that penetrate beyond the HCAL. In general, these are muons and weakly interacting particles such as neutrinos. Since muons are charged particles, their charge and momentum can be measured based on the curvature of their tracks. However, neutrinos are neutral particles and hardly interact with the detector. Nevertheless their presence can be inferred by adding up all the transverse momenta of the detected particles, and assigning the missing transverse momentum to the neutrinos. More details on the inner tracker and calorimeters are given in Sections 3.2.3 and 3.2.4, respectively. Since this analysis uses muon data, the muon spectrometer is discussed with greater emphasis in Section 3.2.5.

3.2.3 Inner tracking system

The CMS inner tracking system [63] [64] is built to withstand intense particle fluxes expected at the LHC, featuring high granularity, fast response and radiation hardness along with efficient cooling. The CMS tracker design is entirely based on silicon detector technology with about 200 m^2 of active silicon area making it the largest silicon tracker ever built. Figure 3.9 shows the layout of the tracking system. The tracking system surrounds the interaction point and has a length of 5.8 m and a diameter of 2.5 m. It

is composed of a pixel detector with 1440 pixel modules containing 66 million pixels and a silicon strip tracker with 15 148 strip detector modules containing 9.6 million silicon strips. Each system is completed by endcaps extending the geometrical acceptance to a pseudorapidity of $|\eta| < 2.5$. The CMS tracking system provides a precise and efficient measurement of trajectories of charged particles, as well as precise reconstruction of primary and secondary vertices. It can also precisely measure track impact parameter variables such as the longitudinal and transverse distances from the vertex.

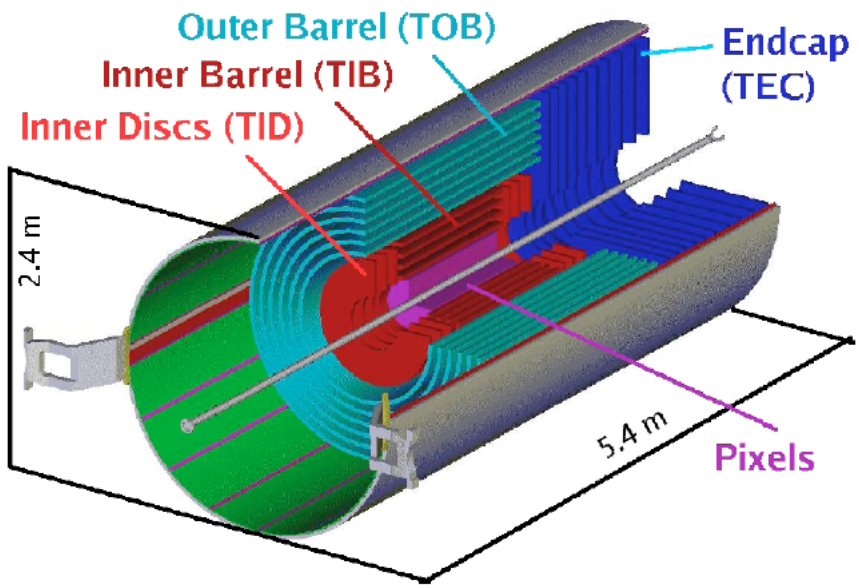


FIG. 3.9. Layout of the CMS inner tracking system [55].

The pixel system sits closest to the interaction point and contains three barrel layers (referred to as pixel barrels (BPix)) located at radii 4.4, 7.3 and 10.2 cm, with two endcap disks (referred to as forward pixels (FPix)) on each side of them (see Figure 3.9). There are altogether 768 pixel modules in the BPix and 672 modules in the FPix. Each module is made of several pixel cells, with a cell size of $100 \times 150\ \mu\text{m}^2$. The pixel layers provide a two-dimensional measurement (in r–ϕ and z) of the particle position, whereas half of the modules of silicon strip detector provide a single coordinate measurement while the

other half (3 out of 7 end cap rings and 4 out of 10 barrel layers are stereo (back-to-back) modules)) provides a two-dimensional measurement .

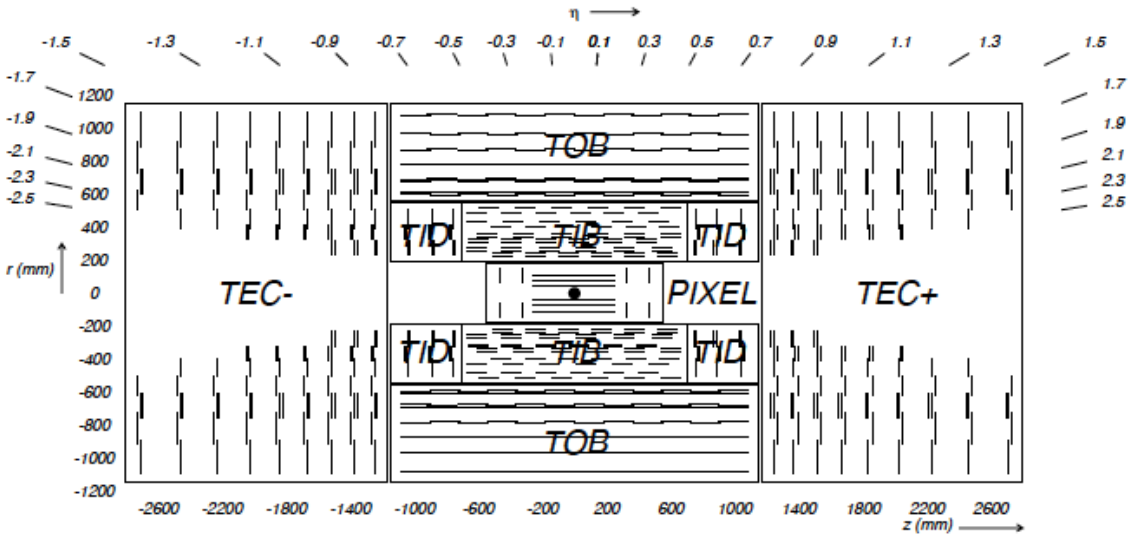


FIG. 3.10. Schematic layout of the CMS tracker showing detailed coverage in the r–z plane [55]. Each line represents a detector module. Double lines indicate back-to-back modules which deliver stereo hits.

The silicon strip tracker is composed of three different subsystems. The tracker inner barrel and disks (TIB/TID), the tracker outer barrel (TOB) and the tracker endcaps (TEC). As shown in Figure 3.10, the radial region between 20 cm and 116 cm is occupied by the silicon strip tracker. The TIB/TID consists of 4 barrel layers complemented by 3 disks at each end and uses 320 μm thick silicon micro-strip sensors with strip pitches ranging from 80 μm to 141 μm providing up to 4 r–ϕ measurements on a trajectory. The TIB/TID are surrounded by the TOB. The TOB consists of 6 barrel layers (providing 6 r–ϕ measurements) of 500 μm thick micro-strip sensors with strip pitches ranging between 122 μm and 183 μm. The TOB extends in z between $\pm$118 mm. Beyond this range, the TECs (TEC– and TEC+ along the z axis) cover the rest of the region as shown in Figure 3.10. Each TEC is composed of 9 disks providing 9 ϕ measurements of the trajectory. Furthermore, some layers of the silicon strip tracker (as shown by double lines

in Figure 3.10) carry a second micro-strip detector module which is mounted back-to-back with a stereo angle of 100 mrad in order to provide a measurement of the second co-ordinate, z in the barrel and r in the disks. This tracker design ensures at least 9 particle hits in the silicon strip tracker in the full range of $|\eta| < 2.4$ with at least 4 of them being two-dimensional measurements.

3.2.4 Calorimetry

A calorimeter measures the energy lost by a particle that enters it. It is designed to 'stop' or 'absorb' particles by forcing them to deposit all of their energy within the detector. The CMS calorimetry is comprised of two types of calorimeters: the electromagnetic calorimeter (ECAL) and hadronic calorimeter (HCAL). Most particles upon interacting with the dense material of a calorimeter lose most of their energy via particle showers which are a cascade of secondary particles produced when a high energy particle interacts with dense matter. High energy electrons, positrons, and photons are capable of initiating electromagnetic showers upon encountering the ECAL whereas in the HCAL hadrons produce hadron showers. A high energy photon interacts with matter mainly via pair production, a process in which an emitted photon produces an e^-e^+ pair each of which in turn radiate more photons by bremsstrahlung which decay into more e^-e^+ pairs and so on (see Figure 3.11). High energy electrons or positrons, on the other hand, lose their energy via bremsstrahlung radiation (means "breaking radiation") which is produced when they are decelerated in matter due to atomic electric fields.

The ECAL [65] of CMS is a hermetic homogeneous calorimeter with a total coverage in pseudorapidity $|\eta| < 3.0$. In a homogenous calorimeter the absorber material (or "passive" material) which produces particle showers is same as the sensitive material (or "active" material) that measures the energy deposited. The CMS ECAL is made of 61 200 scintillating lead tungstate ($PbWO_4$) crystals in the central barrel (EB) and

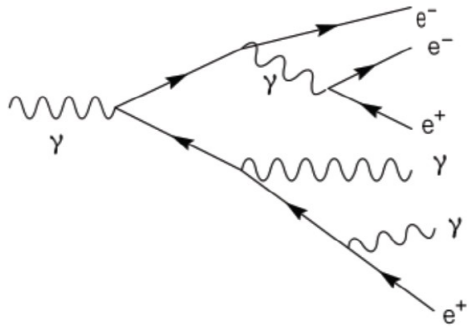


FIG. 3.11. Schematic of an electromagnetic shower.

7 234 crystals in each of the two endcap (EE) regions. The crystals have high density (8.28 g/cm^3), short radiation length[1] (0.89 cm) and small Moliére radius[2] (2.2 cm) which makes them an appropriate choice for operation at the LHC. These properties also result in fine granularity and a compact calorimeter. Furthermore, the scintillation and optical properties of these crystals make them fast and radiation hard. Avalanche photodiodes (APDs) and vacuum phototriodes (VPTs) are used as scintillating light detectors in the barrel and endcap regions, respectively. The EB covers the pseudorapidity range $|\eta| < 1.479$ and EE covers the range $1.479 < |\eta| < 3.0$. A basic layout of the ECAL is shown in Figure 3.12. The cross section of the crystals varies from 22×22 mm^2 to 26×26 mm^2 with a length of 230 mm. In the barrel region, crystals are mounted on thin-walled alveolar structures called submodules, of thickness 0.1 mm, and are further arranged into modules and supermodules. Each EE is divided into two halves, or "Dees" and the crystals are arranged in bigger units (called "supercrystals") on the inner and outer circumference of the Dees. Each Dee holds 3 662 crystals. Also the endcaps are equipped with a preshower detector with the aim of identifying neutral pions in the region $1.653 < |\eta| < 2.6$ and improving the identification of electrons and photons with high granularity.

[1]The radiation length of a material is the mean length (in cm) to reduce the energy of an electron by the factor 1/e via bremsstrahlung radiation.

[2]A characteristic constant of a material describing the transverse dimension of an electromagnetic shower. A smaller Moliére radius means better shower position resolution, and better shower separation due to a smaller degree of shower overlaps.

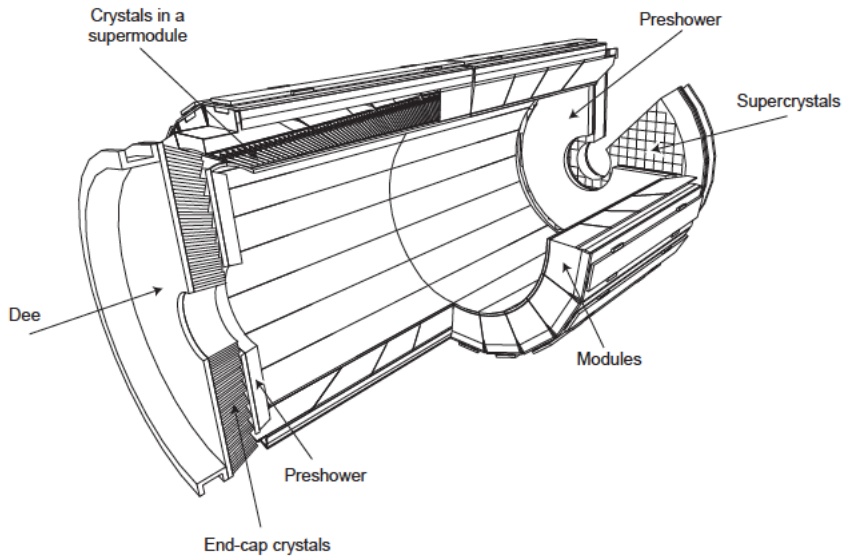


FIG. 3.12. Layout of the CMS ECAL showing the arrangement of $PbWO_4$ crystals in the barrel and endcap regions, along with the preshower detector [55]

Unlike the ECAL, the HCAL [66] is a sampling calorimeter where different materials are used for active and passive regions. It is designed for measuring hadronic jets and other exotic particles that leave a signature in the form of missing transverse energy, E_T. The HCAL surrounds the ECAL and is radially restricted between the outer extent of the ECAL ($R = 1.77$ m) and the inner extent of the magnet ($R = 2.95$ m) which limits the total amount of the absorber material that can be put in the calorimeter. In order to overcome this limitation, an outer hadron calorimeter (HO) or "tail catcher" is placed just outside the solenoidal magnet. Besides the HO, the HCAL consists of a hadron barrel (HB), hadron endcaps (HE) and a hadron forward (HF) calorimeter. Figure 3.13 shows the longitudinal view of the CMS detector with the HCAL coverage in η.

The HB is divided into two half-barrel sections (HB+ and HB–) and covers the pseudorapidity region $|\eta| < 1.3$ and HE covers the region $1.3 < |\eta| < 3.0$. The HB and HE are made of alternating layers of plastic scintillating tiles (the active material) and flat brass absorber plates enclosed between steel plates (the passive material). The steel plates

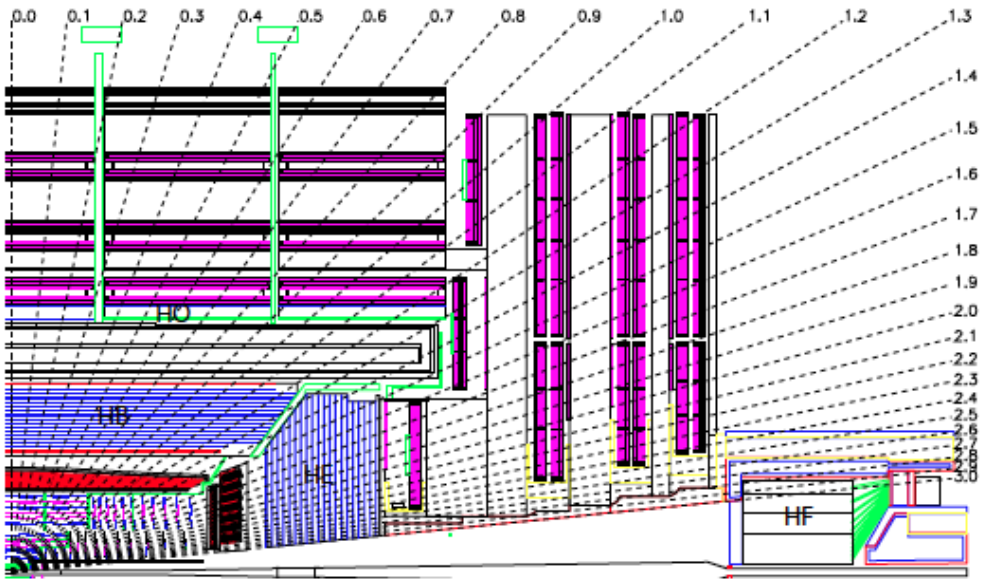


FIG. 3.13. Layout of the hadron calorimeter [55].

are used for structural support. When a hadronic shower passes through a scintillating tile, light is emitted which is collected by a wavelength-shifting fibre and is then fed to a hybrid photodiode (HPD) for readout. The HF uses a Cherenkov-based radiation hard technology and is situated at 11.2 m from the collision point covering the pseudorapidity region $3.0 < |\eta| < 5.2$. Although called a hadron forward calorimeter, it actually measures the energy of both hadronic and electromagnetic showers in the forward region. It is constructed using steel absorbers and quartz fibers. As charged particles pass through the quartz fibers with a velocity greater than the speed of light, they emit Cherenkov radiation, which is then detected by photomultipliers.

3.2.5 Muon system

Muons are potential candidates for recognizing signatures of new interesting processes[3], therefore, the detection of muons is of great importance to CMS. The main function of the

[3]One such signature is the predicted decay of the Higgs boson into four muons via ZZ.

CMS muon system [67] is to efficiently identify, trigger and measure muon momenta over the pseudorapidity range $|\eta| < 2.4$. Three types of gaseous particle detectors are used for muon identification: drift tubes (DT) in the barrel region, cathode strip chambers (CSC) in the endcap region and resistive plate chambers (RPC) in both the barrel and endcap regions. DTs and CSCs cover the region of $|\eta| < 1.2$ and $0.9 < |\eta| < 2.4$, respectively. RPCs cover the region with $|\eta| < 1.6$. Figure 3.14 shows the muon system elements in

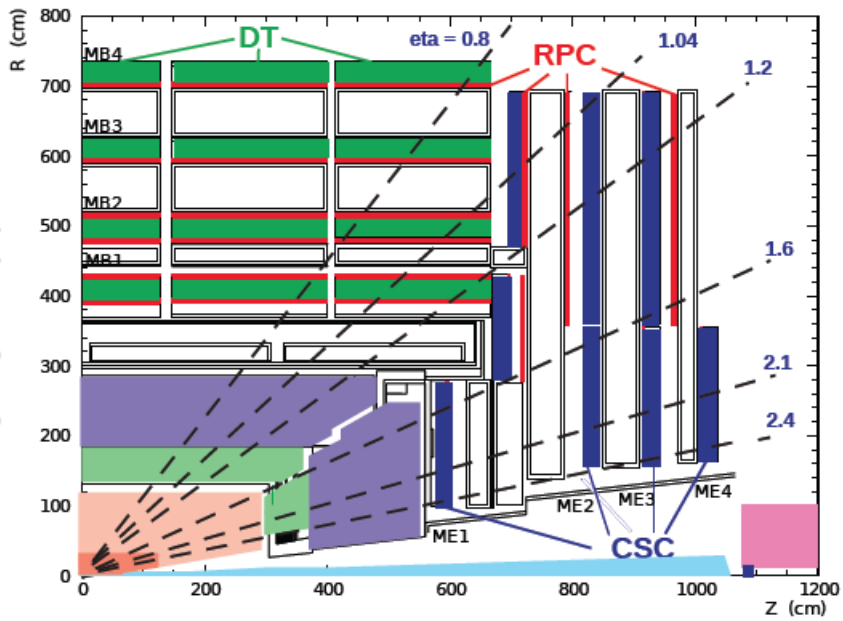


FIG. 3.14. Layout of the CMS muon system shown for one quadrant of the CMS detector [55]. The four barrel DT stations (MB1–MB4), four endcap CSC stations (ME1–ME4), and RPC stations are shown.

one quadrant of the CMS detector. In total there are 1400 muon chambers: 250 DTs, 540 CSCs, and 610 RPCs.

Drift tube chambers made of rectangular drift cells are used in the barrel region where the muon rate is low and magnetic field is uniform. The DT chambers are organized into four stations interspersed among the layers of the flux-return yoke. The first three stations each contain 12 drift tube chambers, of which 8 of them provide a measurement in the $r - \phi$ plane and the rest provide a measurement in the z direction. The fourth

DT station does not measure the z direction. A schematic layout of barrel muon DT chambers is shown in Figure 3.15. A typical drift cell is made of two electrode plates

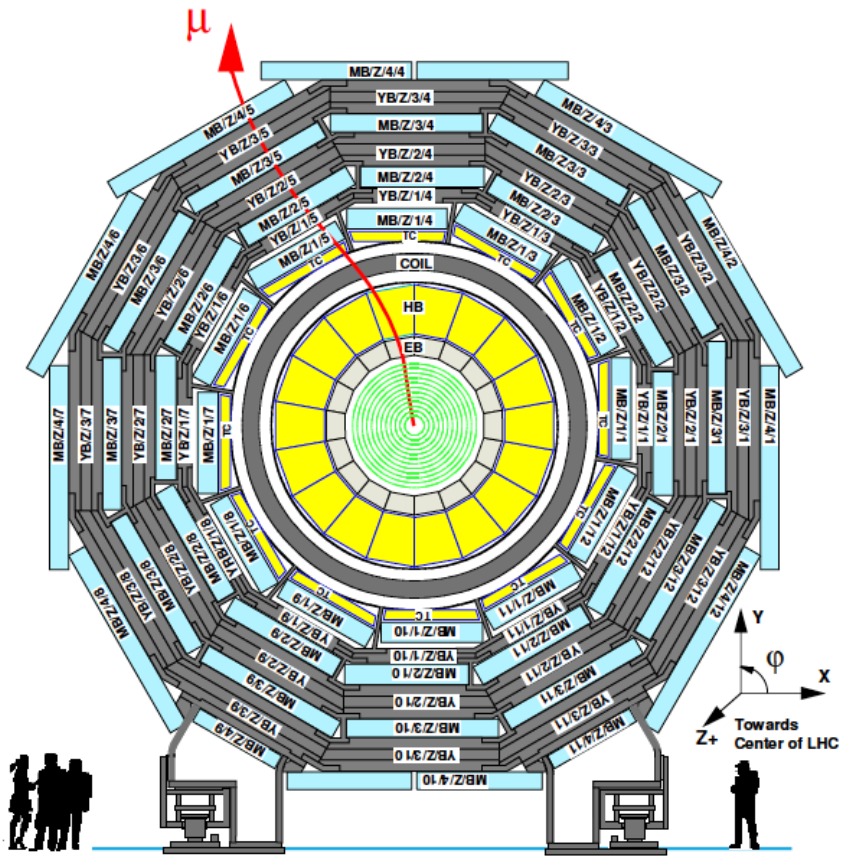


FIG. 3.15. Layout of the barrel muon (MB) DT chambers in one of the 5 CMS barrel wheels [55]. Each DT chamber is denoted by MB/Z/a/b, where a refers to the DT station number ranging from 1 to 4 and b refers to the chamber number (within a DT station) ranging from 1 to 12. In chambers 4 (top) and 10 (bottom), the MB4 chambers are cut in half to simplify the mechanical assembly. YB refers to the yoke barrel regions with same numbering scheme as used for the MB DT stations.

(operated at high voltage) enclosing argon–carbon dioxide gas mixture (Ar-CO_2) and a stretched anode wire. When a muon or any charged particle passes through the gas, it liberates electrons from atoms through ionization of the gas. The resulting electrons drift to the nearest anode wire. By recording where along the wire electrons hit as well as by calculating the muon's original distance from the anode wire, DTs give a muon's position

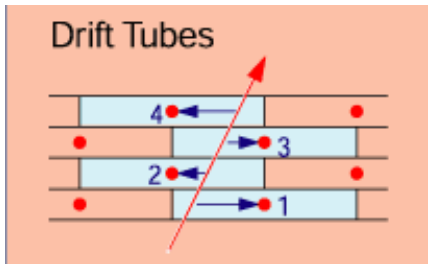


FIG. 3.16. Muon drift tube operation. The red line with arrow shows a muon traversing the drift tube. The anode wires (shown by red dots) are perpendicular to the page. The horizontal blue lines with arrows show the muon's distance from the anode wire (obtained by multiplying the speed of an electron in the tube by the drift time).

measurement in two coordinates. Figure 3.16 demonstrates the basic operation of a drift cell.

Trapezoidal shaped CSC chambers are used as muon detectors in the endcap region where the muon and background rates are high and the magnetic field is large and non-uniform. The CSCs cover the far-forward region of the detector with $0.9 < |\eta| < 2.4$. The CSC features fast response time, fine segmentation and radiation resistance and provides efficient identification of muons. In each endcap, CSCs consists of 4 muon endcap stations (ME1–ME4) arranged in groups as follows: 72 ME1/1, 72 ME1/2, 72 ME1/3, 36 ME2/1, 72 ME2/2, 36 ME3/1, 72 ME3/2, and 36 ME4/1 (see Figure 3.17).

CSCs are multi-wire chambers made of 7 cathode panels forming 6 gas gaps with planes of anode wires. The cathode strips run radially (at constant $\Delta\phi$ width) and anode wires run azimuthally (defining the track's radial coordinate). The basic layout of a CSC chamber is shown in the left panel of Figure 3.18. The right panel shows the basic principle of CSC operation.

RPCs complement the DTs and CSCs in the barrel and endcap regions, respectively. In the barrel region, the barrel RPC (RB) chambers form 6 coaxial cylindrical layers (all around the beam axis), arranged with respect to the four muon stations (see Figure 3.19 [55]). The first and second muon stations each contain two layers of RPC

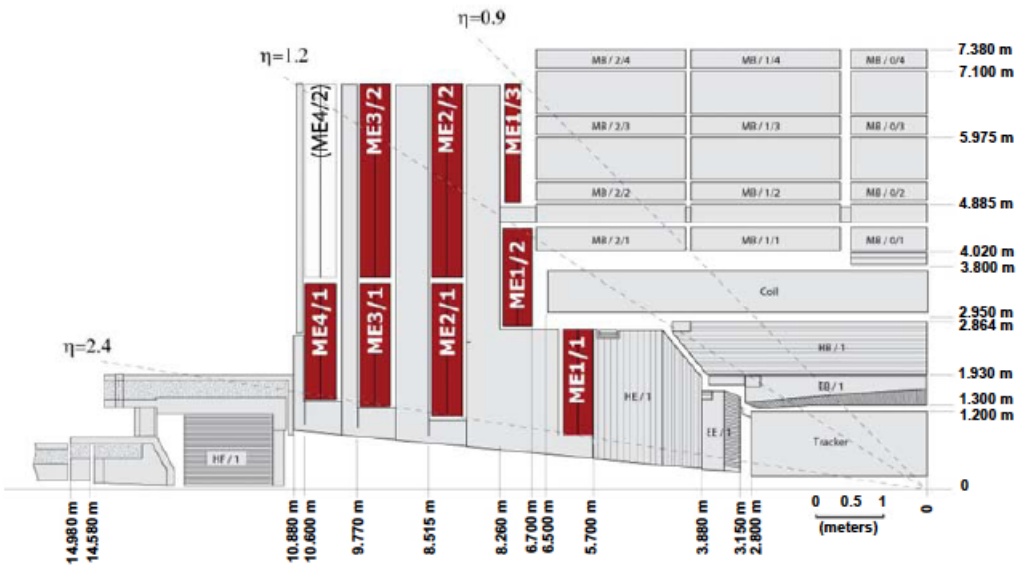


FIG. 3.17. Quarter-view of the CMS detector [55]. Shown in red color are the CSCs. The ME4/2 chambers will eventually be built and installed as part of the CMS upgrade plans.

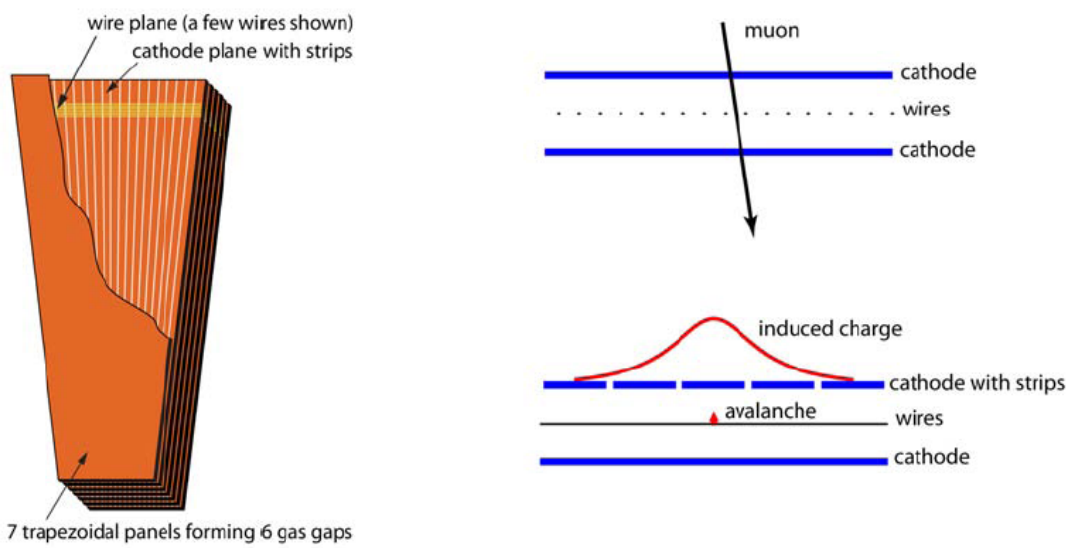


FIG. 3.18. CSC chamber design and operation [55]. Left panel: Layout of a CSC chamber made of 7 trapezoidal planes. The exposed portion shows anode wires and cathode strips. Right panel: A schematic view of a single gas gap illustrating the principle of CSC operation. The muon coordinate along wires is obtained by interpolating charges induced on cathode strips.

chambers located inside (RB1 in and RB2 in) and outside (RB1 out and RB2 out) with respect to the Drift Tube (DT) chambers. The third and fourth stations each contain two RPC chambers, both located on the inner side of a DT layer (named RB3+ and RB3, RB4+ and RB4). Special cases are RB4 in sector 4 (which contains 4 chambers: RB4++, RB4+, RB4-, and RB4- -) and sectors 9 and 11 (which contain only 1 RB4 chamber). The CMS forward and backward regions contain 3 iron disks that constitute the endcap yokes (YE 1–3). The four CSC stations (ME 1–4) and endcap RPC stations (RE 1–3) are mounted on either side of the YE layers. RE1 is located on the interaction point side of YE1, underneath the CSC chambers of ME1. The RE2 chambers are mounted on the back side of the YE1 and RE3 on the interaction point side of YE3. RPCs are gaseous parallel-plate detectors which provide fast response with good time resolution for muon triggering and identification. However, the position resolution of RPCs is coarser than that of DTs or CSCs. The CMS RPC basic double-gap module consists of two gaps (referred to as *up* and *down* gaps) operated in avalanche mode with common read-out strips in between (see Figure 3.20). When a muon passes through the chamber, gas atoms are ionized leading to an avalanche of electrons. The signal (the electrons) is picked up by external metallic detector strips for readout.

Muon reconstruction

The CMS muon system is designed to identify and reconstruct muons over the pseudo-rapidity interval $|\eta| < 2.4$. This results in high reconstruction efficiencies ($> 99\%$) for muons that are within the η range and have sufficient p_T [69]. Also, the high magnetic field makes possible excellent muon momentum resolution. In the standard CMS reconstruction procedure, particle tracks are first reconstructed independently in the inner tracker (referred to as *tracker tracks*) and in the muon system (referred to as *standalone-muon tracks*). Tracks are also reconstructed locally within the muon system using available DT

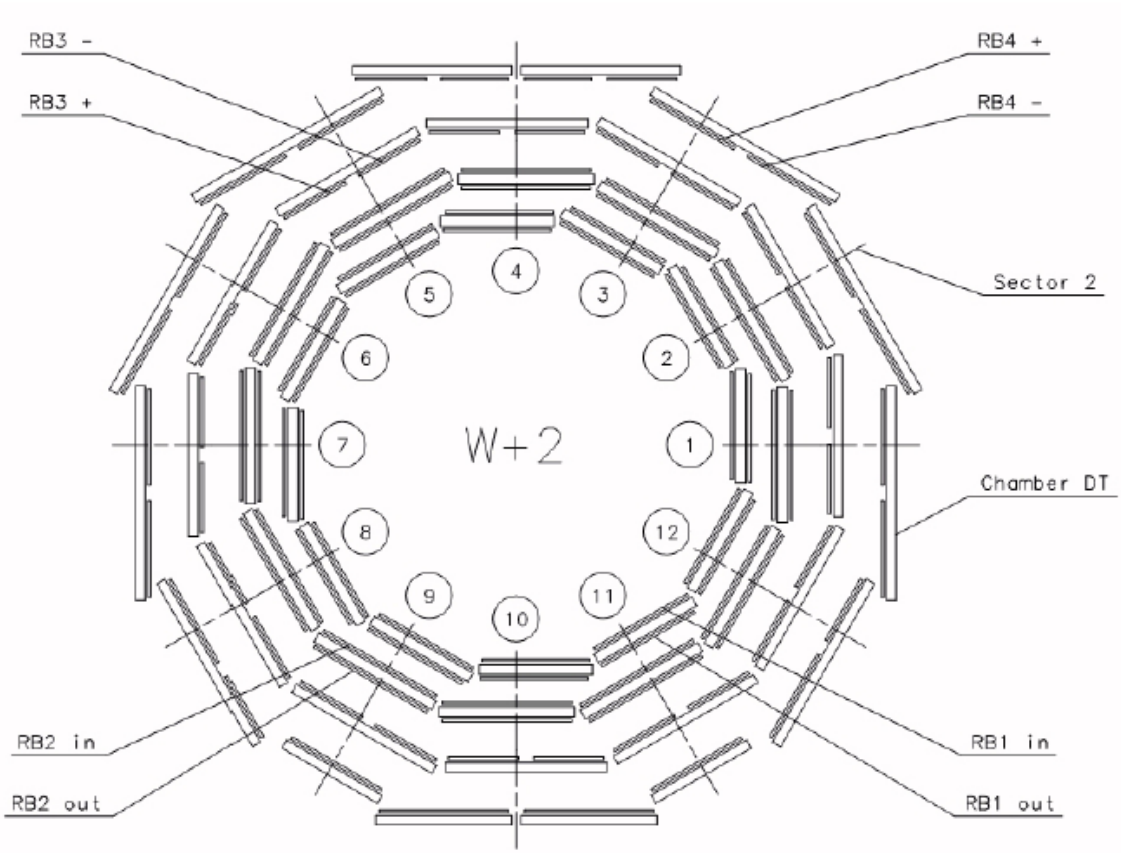


FIG. 3.19. Schematic layout of one of the CMS barrel wheels, labeled W+2.

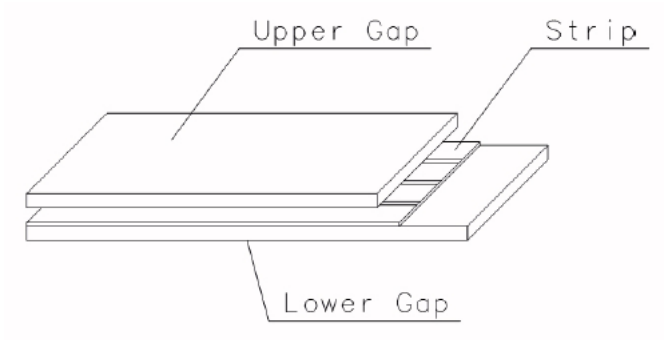


FIG. 3.20. Layout of a double-gap RPC [55].

or CSC particle hits (referred to as *muon segments*). Based on these, muon reconstruction is done following two approaches:

- *Tracker muon reconstruction (inside-out):* In this approach, all tracker tracks above a certain p_T threshold are considered and are extrapolated to the muon system for matching muon segments. If at least one muon segment matches the extrapolated tracker track, the corresponding tracker track qualifies as a *tracker muon track*.

- *Global muon reconstruction (outside-in):* In this approach, standalone muon tracks are extrapolated for matching tracker tracks. If one is found, hits from the tracker track and standalone muon track are combined and used to determine a *global muon track*.

In the specific case when tracker tracks are not available in a given collision event, only a *standalone muon track* is reconstructed. However, this is very rare (occurring in only about 1% of the events) due to high tracker efficiency.

At low momentum ($p_T < \sim$20 GeV/c), the *tracker muon reconstruction* is more efficient than the *global muon reconstruction* as it requires only one matching muon segment. At high momentum ($p_T > \sim$200 GeV/c), global reconstruction gives high efficiency (as muons penetrate more than one station at high p_T) and significantly improves muon momentum resolution compared to the tracker only fit [68, 69]. Figure 3.21 shows a comparison of momentum resolution when reconstructed with different approaches.

3.2.6 Trigger and Data Acquisition

As mentioned in Section 3.2, at nominal running, CMS will observe an event rate of about 10^9 events per second with a single event size of $\sim$1.5 MB. This makes it impossible to store and process all of the observed events, hence a dramatic rejection rate is required. This is achieved by the online event selection process called "trigger", which is designed

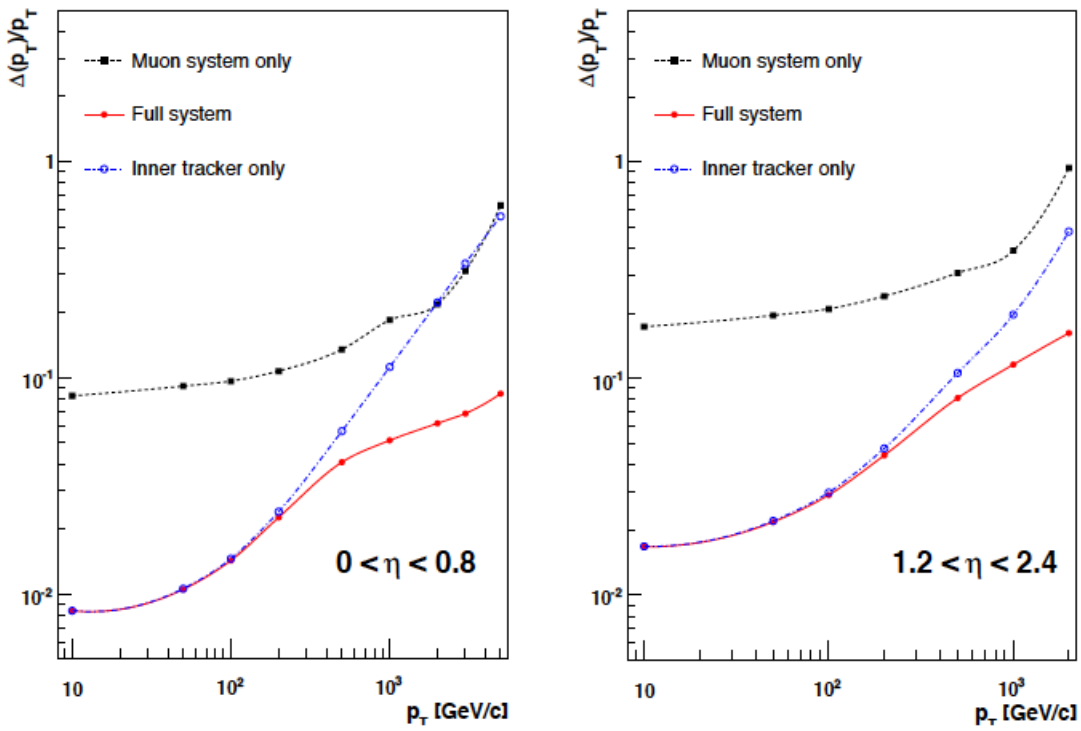


FIG. 3.21. The muon p_T resolution as a function of p_T using simulated data for the muon system only, the inner tracking system only, and both [55]. Left panel: $|\eta| < 0.8$, right panel: $1.2 < |\eta| < 2.4$.

to select potentially interesting events and reduce the huge event rate to a manageable 100 events per second for storage and subsequent analysis. The high event rates and short time between bunch crossings (25 ns) makes it challenging to design the CMS trigger and data acquisition (DAQ) system.

The CMS trigger system performs the event selection in two steps: level-1 (L1) trigger [70] and high-level trigger (HLT) [71]. The L1 trigger consists of custom-designed hardware processors and is extremely fast in decision making. The design output rate of L1 trigger is 100 kHz and produces a trigger decision in less than 1 μs. The decision is based on fast trigger inputs coming from the muon detectors[4] and calorimeters. All muon detectors of CMS are equipped with trigger electronics in their read-out system and provide required information to the L1 trigger. To select an event, the L1 trigger looks for simple signs of interesting physics such as particles with a large amount of energy or p_T, particles with missing E_T, etc. Events accepted by the L1 trigger are sent for further evaluation by the HLT. In the next step, the HLT reduces the L1 output from 100 kHz to 100 Hz. The HLT is a software system implemented in a computer farm consisting of about 1000 processors which run very complex physics tests searching for specific signatures by partially reconstructing the event. The HLT algorithms are a faster version of the offline reconstruction and physics software, looking for specifically interesting events.

At the L1 level, the DAQ system must sustain a maximum input rate of 100 kHz, for a data flow of 100 GB per second coming from approximately 650 data sources, and must provide enough computing power for the HLT to reduce the rate of stored events by a factor of 1000. Before the L1 trigger decision, data is continuously, but, temporarily, stored in the buffers of various sub-detectors. Once a decision is made, corresponding events are extracted and are pushed into the DAQ system where the *event builder* assembles the event fragments into a complete event and transmits them to the

[4]All three muon systems, the DT, the CSC and the RPC, take part in the trigger and can trigger on the p_T of muons independent of the rest of the detector. Also, RPCs are dedicated trigger detectors.

HLT for further processing. During operation, trigger thresholds are optimized in order to fully utilize the available DAQ capacity.

Chapter 4

The PYTHIA Monte Carlo

This chapter introduces the various compositeness models implemented within the PYTHIA Monte Carlo event generator, with an emphasis on the left-left isoscalar model (LLIM). Section 4.1.1 gives the technical details on how to submit a PYTHIA job for the LLIM. Section 4.2 shows the expected cross sections and dimuon mass spectra for the LLIM. Studies shown here are done at PYTHIA generator-level (without the CMS detector simulation), as the main motivation is physics understanding of the LLIM implementation.

4.1 Compositeness models in PYTHIA

PYTHIA Monte Carlo version 6.4 (as described in [47]) is used to simulate contact interaction (CI) signal and standard model (SM) Drell-Yan (DY) samples at the generator-level. Within PYTHIA, subprocess ISUB 165 represents a compositeness scenario where the final states are equivalent to those resulting from γ^*/Z^0 exchange in the SM. The process includes both CI and DY production as well as possible interference between the two. Within subprocess 165 there are two basic compositeness models implemented, the LLIM and a helicity non-conserving model. A choice of the model can be made using the parameter ITCM(5). ITCM(5) ≥ 1 corresponds to a compositeness model, whereas, ITCM(5)

= 0 corresponds to a simplified implementation of the standard model behavior. Table 4.1 summarizes the model choices in detail. These include a distinction as to whether only the first generation quarks are composite or all quarks are considered to be composite.

Table 4.1. Compositeness model options within PYTHIA for dilepton final states.

ITCM(5)	Model	Composite quarks
0	standard model DY	none
1	left-left isoscalar	u,d
2	left-left isoscalar	u,d,c,s,t,b
3	helicity non-conserving	u
4	helicity non-conserving	u,c,t

The standard way of generating DY production in PYTHIA is through subprocess 1 (MSUB1) which is a $2 \rightarrow 1$ process, whereas PYTHIA treats all ISUB 165 processes as $2 \rightarrow 2$ processes by default. This leads to a difference in the default Q^2 scale in parton distributions which is $p_\perp^2$ for ISUB 165 processes and $\hat{s}$ for MSUB1. In order to make the two processes equivalent, following the recommendation of the PYTHIA manual, the MSTP(32) = 4 card is used in ISUB 165 processes to force PYTHIA to use the same $\hat{s}$ Q^2 scale that is used in MSUB1. This corrects for the $2 \rightarrow 2$ implementation of the DY process that is built into the CI interference terms in PYTHIA. With the MSTP(32) = 4 selection, the cross sections as reported by PYTHIA for MSUB1 and ITCM(5) = 0 are identical to within 0.1%.

For this analysis, the LLIM is chosen as it is the conventional benchmark model for compositeness searches in the dilepton channel. In this model, all of the initial state quarks are assumed to be composite (ITCM(5) = 2 selection) and the final state is taken to be $\mu^+\mu^-$. As mentioned in Section 2.2.2, this model corresponds to the first term in Eq. 2.4.

4.1.1 PYTHIA Job details

This subsection gives the technical details on how to submit a PYTHIA job to generate contact interaction events for a given value of Λ. The job options shown below are specifically for a Λ value of 5 TeV and constructive interference.

```
'MSEL         = 0',
'MSUB(165)    = 1',
'MSTP(32)     = 4',
'RTCM(42)     = -1',       ! Constructive interference
'RTCM(41)     = 5000.',    ! Lambda = 5 TeV
'ITCM(5)      = 2',        ! LLIM
'KFPR(165,1)  = 13',
'CKIN(1)      = 120.',
```

The description of the job options is as follows:

- MSEL = 0 lets the user define the physics process.
- MSUB(ISUB) = 1, where ISUB is 165, allows the decay via γ^*/Z^0 exchange.
- MSTP(32) = 4 forces the ISUB 165 processes to use the $\hat{s}$ value for Q^2 scale. In other words, forces a $2 \to 2$ process (CI) to $2 \to 1$ process ($Z \to \mu\mu$).
- RTCM(42) = ± 1 sets the sign of the interference term. Default is +1 (destructive) and setting this to -1 corresponds to the constructive case.
- RTCM(41) sets the Λ value in units of GeV.
- ITCM(5) = 2 selects the LLIM with the selection of all quarks being composite.

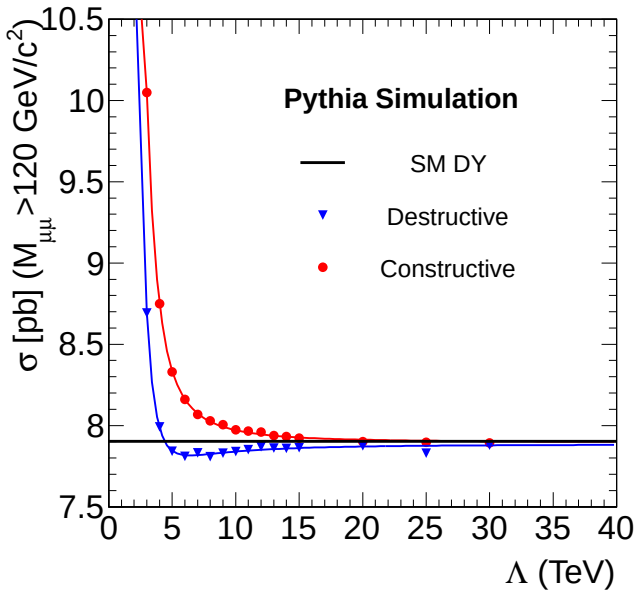
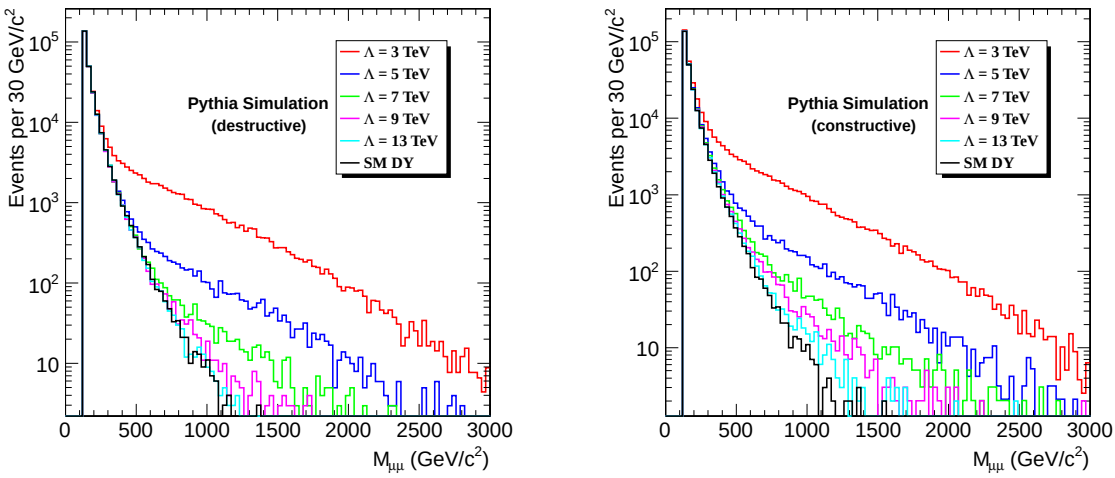


FIG. 4.1. Constructive and destructive LLIM cross sections for $M_{\mu^+\mu^-} > 120$ GeV/c^2. The functional fits differ only in the sign of the interference term in Eq. 2.5. The solid black line corresponds to the DY asymptotic limit.

- KFPR(ISUB, 1) = 13 sets the final state particle type for the chosen ISUB (which is 165 here). A value of 13 corresponds to muons in the final state.
- CKIN(1) sets the lower $\sqrt{\hat{s}}$ value for the interaction in units of GeV.

4.2 The left-left iso-scalar model

The LLIM defines a left-handed current model for contact interactions. This analysis interprets the data in the context of the LLIM which in addition to generating pure CI and DY terms, also allows for interference between the two processes, as shown by Eq. 2.5. Hence the term "CI/DY" or "CI signal" will be used throughout this thesis collectively to refer to the processes contributing to the cross section in Eq. 2.5. Using PYTHIA, the expected cross sections and dimuon mass spectra in the LLIM (at the generator-level) are



(a) Destructive interference

(b) Constructive interference

FIG. 4.2. Dimuon event yields at different values of Λ for (a) destructive interference and (b) constructive interference. Values are shown for M >120 GeV/c^2. As Λ increases, the dimuon mass distribution tends toward pure DY. The model predictions are shown over the full mass range, although the model is not valid for $M_{\mu\mu}^{Low}c^2 \geq \Lambda$.

shown in this section with appropriate kinematic cuts. For all the plots shown here, both muons in the dimuon pair are subject to the requirements $|\eta| < 2.1$ and $p_T > 40$ GeV/c. These requirements are chosen to be less restrictive than the corresponding reconstruction requirements (see Section 5.2), taking into account resolution effects. Also the mass of the dimuon pair is restricted to be above 120 GeV/c^2 as the CI/DY contribution below this value is much smaller than that due to DY.

Figure 4.1 shows a comparison of destructive and constructive differential cross sections for $M_{\mu\mu} > 120$ GeV/c^2 at $\sqrt{s} = 7$ TeV. The fits to the cross section points are based on Eq. 2.5 and differ only in the sign of the interference term (second term). As $\Lambda \to \infty$, the two cross sections approach the DY limit as suggested by Eq. 2.5 and as shown by the solid black line in the plot. Note that in the destructive case, the cross section achieves a minimum around $\Lambda = 6$–7 TeV due to the effect of destructive interference.

Figures 4.2(a) and 4.2(b) show the dimuon event yields in the LLIM at various values of Λ for destructive and constructive interference, respectively. The distributions shown in these figures correspond to a raw sample of 500k events, however, only about half of the total events survive the $p_T > 40$ GeV/c requirement. Comparing the two figures one can see that the cross section for the constructive case is slightly higher than for the destructive case. Also, note that the contact interaction is not visible until above the Z peak which justifies the use of 120 GeV/c^2 dimuon mass requirement in this analysis. The Λ curves in Fig. 4.2 illustrate that the CI/DY leads to a less steeply falling cross section than DY production, with the effect steadily increasing with decreasing Λ. One can also see that as Λ increases, the distribution tends to the pure DY production as expected.

Chapter 5

Data and event selection

This chapter starts with an introduction to the 2011 CMS muon dataset used in this analysis. Section 5.2 discusses in detail the selection criteria used to select muon and dimuon events. The effect of individual selection cuts on dimuon event statistics is also shown. Finally, the event pileup (the occurrence of multiple *pp* interactions recorded by the detector as a single event) during the course of the 2011 data taking is briefly discussed in Section 5.3.

5.1 The 2011 dataset

Proton-proton collision data at LHC center-of-mass energy 7 TeV recorded between January and October 2011 are used in this analysis. Based on the instantaneous luminosity, the 2011 data taking period was divided into two sections, 2011A and 2011B. Collision events are grouped into several *runs* depending on the LHC filling scheme (Section 3.1.3) and expected instantaneous luminosity which further determine the trigger conditions used. Each run has to pass certain validation requirements for the sub-detectors in order to be considered a *good* run. Since this analysis looks for muon final states, validation requirements for calorimeters are not taken into account. The total integrated luminosity

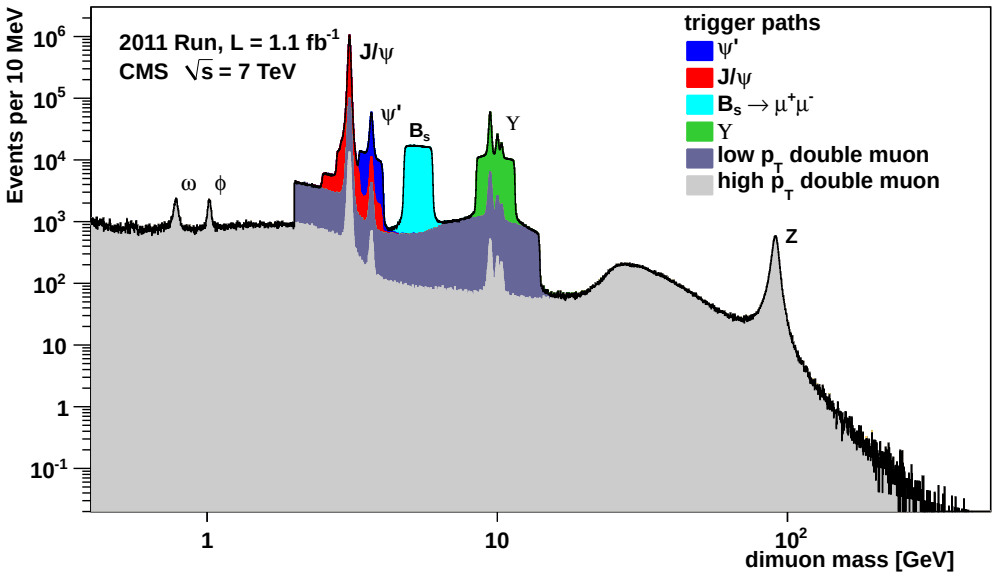


FIG. 5.1. Opposite sign dimuon invariant mass spectra showing the resonant peaks of the SM particles. Dimuon mass is given in units of $c = 1$.

of the 2011A and 2011B muon datasets is 2413 pb^{-1} and 2864 pb^{-1}, respectively. Together, the two datasets yield a total luminosity of 5277 pb^{-1}. The uncertainty in the total luminosity is estimated to be 116 pb^{-1} [76].

One of the important steps to be done during the initial data taking period at any particle physics experiment is to provide confirmation of the previously verified predictions of the standard model (SM). Figure 5.1 [74] shows the invariant mass spectra of opposite-sign muon pairs using a subset of the CMS 2011 dataset corresponding to an integrated luminosity of 1.14 fb^{-1}. The plot shows the Z boson (on the far right), J/ψ and Upsilon mesons, and at lower energy, the lighter resonance decays of rho (ρ), omega (ω), and phi (ϕ) mesons, reaffirming SM production at $\sqrt{s} = 7$ TeV.

5.1.1 Trigger requirements

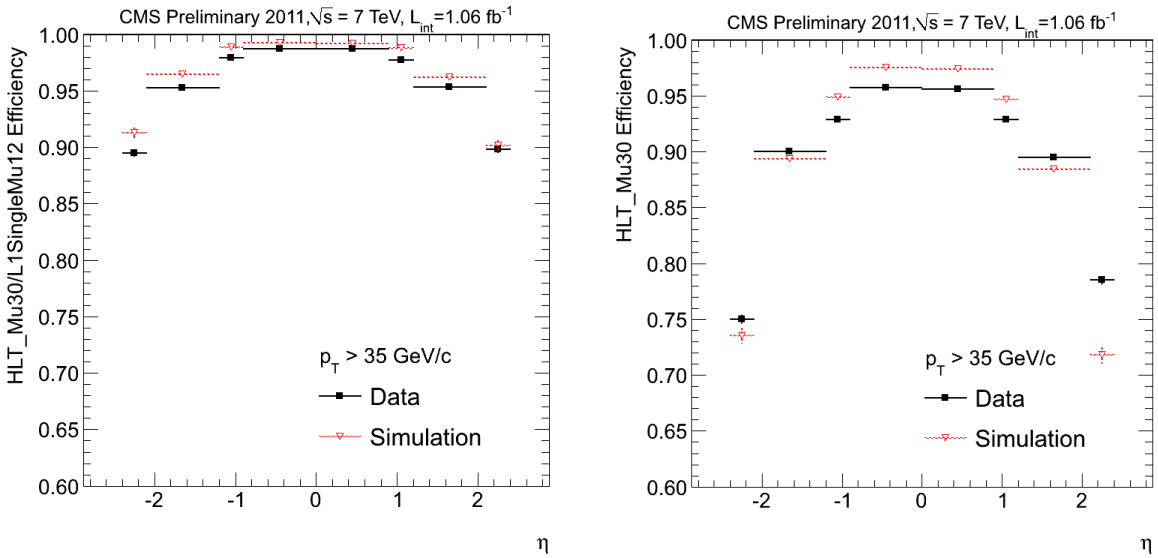
The 2011A and 2011B muon datasets use single muon triggers to select events, both at the level-1 (L1) and high-level trigger (HLT). The relevant single muon HLT trigger

Table 5.1. The 2011A and 2011B datasets used, with run ranges, HLT paths, and corresponding luminosities.

Dataset	Run range	HLT path	Luminosity (pb^{-1})
Run 2011A	160404–163869	HLT_Mu30	233.4
	165088–173198	HLT_Mu40	1906.1
	173236–173692	HLT_Mu40_eta2p1	273.4
		Run2011A total	2412.8
Run 2011B	175832–180252	HLT_Mu40_eta2p1	2864.3
		Run2011A + Run2011B total	**5277.2**

paths for these run periods are HLT_Mu30, HLT_Mu40, and HLT_Mu40_eta2p1. As the names imply, the HLT_Mu30 and HLT_Mu40 triggers select events with muon $p_T > 30$ GeV/c and $p_T > 40$ GeV/c within $|\eta| < 2.1$, respectively. The HLT_Mu40_eta2p1 is same as HLT_Mu40 except that the restriction on η appears also at the L1 level. The corresponding L1 triggers select events at a lower p_T threshold, $p_T > 12$ GeV/c for HLT_Mu30 and $p_T > 16$ GeV/c for HLT_Mu40 and HLT_Mu40_eta2p1 (with an explicit cut of $|\eta| < 2.1$). Table 5.1 summarizes the run ranges and HLT trigger paths used in the analysis. For uniformity, the L1 and HLT conditions of HLT_Mu40_eta2p1 trigger are used for all periods of the 2011 data and signal and background Monte Carlo (MC) sources. In the case when the HLT_Mu40_eta2p1 trigger path does not exist in simulated samples and data (*e.g.,* during the initial period of 2011 data taking, see Table 5.1), it is emulated using HLT_Mu30 or HLT_Mu40, by additional requirements for the HLT object to be in $|\eta| < 2.1$ range and p_T to be greater than 40 GeV/c.

Trigger efficiencies are measured by the CMS muon physics object group (MUON POG) by applying the "Tag and Probe" method [73] to a sample of $Z \to \mu^+\mu^-$ events collected with single muon triggers. Figures 5.2(a) and 5.2(b) show the single muon HLT and L1 + HLT trigger efficiencies for data and MC as a function of η [74]. The overall L1+HLT efficiency of the single muon trigger for 2011 runs (for the event selection criteria discussed in the next section) is about 93.3% for $|\eta| < 2.1$ and $p_T > 35$ GeV/c, whereas



(a) HLT_Mu30 (L1_Mu12) eff. vs. η

(b) L1_Mu12 + HLT_Mu30 eff. vs. η

FIG. 5.2. Single-muon trigger efficiencies for $p_T > 35$ GeV/c as a function of η: the efficiency of the HLT with p_T threshold at 30 GeV/c with respect to L1 trigger with p_T threshold of 12 GeV/c (left), and the combined efficiency of L1 and HLT (right). The efficiencies obtained using $Z \to \mu^+\mu^-$ events are compared with predictions from the MC simulation.

in the barrel ($|\eta| < 0.9$) and endcap ($0.9 < |\eta| < 2.1$) regions, it is around 95.0% and 89.9%, respectively.

5.2 Dimuon selection criteria

To reduce mis-reconstructed muons and muons coming from cosmic rays, a set of reconstruction requirements, as discussed in this section, are used to select dimuon events [73]. The same selection criteria are applied both to data and MC sources. At the generator level, each muon is required to have $p_T > 40$ GeV/c and $|\eta| < 2.6$. These requirements are chosen to be less restrictive than the corresponding reconstruction requirements, taking into account resolution effects.

- *Baseline selection of events*

- In order to avoid events coming from beam backgrounds, events are required to have at least 25% of the silicon tracker tracks (see Section 3.2.5) marked as *high purity*. The reconstructed tracks are filtered to remove tracks that are likely fakes and to provide a means of quantifying the quality of the remaining tracks. The filtering uses information on the number of hits, the normalized χ^2 of the track, and the compatibility of the track originating from a pixel vertex. Tracks that pass the tightest selection are labelled *high purity*. A detailed discussion is given in [72].
- To reject cosmic ray muons triggering in empty bunch crossings, which can produce fake muons when traveling close to the interaction point, events are required to have at least one *good* offline-reconstructed primary vertex. A primary vertex is considered *good* if it is associated with four or more tracks; has $r < 2$ cm, and $|z| < 24$ cm (here r and z correspond to the radial and longitudinal distances from the interaction point, see Figure 3.10).

- *Selection of individual muons*

  In order to reduce the rate of muons from decays-in-flight[1] and punch-through[2] and ensure the quality of muon tracks, both muon candidates are required to pass the following selection criteria:

  - The muon candidate must be reconstructed both as a global and tracker muon (see Section 3.2.6).
  - The muon candidate must have $p_T > 45$ GeV/c, as reconstructed offline.
  - The global muon candidate track is required to have a signal ("hit") in at least one pixel layer, hits in at least one muon detector layer, and hits in at least

[1]Muons from decays-in-flight refer to muons coming from pion and kaon decays.

[2]Secondary hadrons or muons ("fake muons") from a hadronic shower in the calorimeter that survive to reach the muon detector are referred to as hadronic punch-through.

nine strip layers.

- The tracker muon candidate track must have matched *segments* (series of hits) in at least two muon detector stations.
- The muon candidate must have a transverse impact parameter $|d_{xy}| < 0.2$ cm with respect to the beam spot; this selection significantly reduces the rate of muons from decays-in-flight.
- To suppress muons coming from hadronic decays (*e.g., decays of charm and bottom quarks*), a tracker-based isolation is imposed such that the sum of p_T of all tracks, excluding the muon candidate's track, within a cone of radius $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.3$ is less than 10% of the p_T of the muon. Suppressing hadronic decays relies on the feature that hadrons are typically grouped in jets and as such tend to be non-isolated.
- At least one of the reconstructed muon candidates must be matched within $\Delta R < 0.2$ and $\Delta p_T/p_T < 1$ to the HLT muon candidate ("trigger muon"), which has a p_T threshold of 40 GeV/c and is restricted to the $|\eta| < 2.1$ extent of the detector (as described in Section 5.1.1).

- *From muons to dimuons*

  To form a dimuon, the two muons, passing the above selection, must further satisfy the following requirements:

  - must have opposite charge
  - must be consistent with originating from a common vertex. A constrained fit of the muon candidate tracks to a common vertex must satisfy $\chi^2 < 10$.
  - A cosmic ray muon passing close to the detector interaction point can appear as two muon candidates back-to-back, faking a dimuon event. To suppress

Table 5.2. Effect of various selection cuts on muon event statistics. Here 'initial sample (I)' corresponds to a subset of events which had passed the baseline selection, trigger requirement and contains at least one muon that is both tracker and global. N_{events} refer to the number of events passing a specific cut relative to the previous selection step. The cumulative effect of cuts on the initial sample is given in the last column.

Selection criteria	N_{events}	% N_{events} survived	% I survived
Initial sample (I)	13183579	–	–
At least one pixel hit	12417846	94.2	94.2
At least 9 tracker layers	11506203	92.7	87.3
At least one muon hit	11200797	97.3	85.0
Matched muon segments	10490218	93.7	79.6
Impact parameter cut	10449300	99.6	79.3
Offline $p_T > 45$ GeV/c	3604278	34.5	27.3
Tracker based isolation cut	2489304	69.1	18.9
Two oppositely charged muons	146425	5.9	1.1
3D dimuon angle cut	146323	99.9	1.1
Dimuon vertex Chi2 cut	145721	99.6	1.1
$M_{\mu\mu} > 120$ GeV/c^2	21860	15.0	0.2

this, the 3-dimensional angle between the two muon candidates' momenta is required to be less than $\pi - 0.02$ radians.

If an event has more than two reconstructed muon candidates passing all the above requirements, the two highest p_T muon candidates are selected, and the event is retained only if these two muon candidates are oppositely charged. The efficiency to reconstruct and identify a muon, with the above selection, is measured to be greater than 95% over the full pseudorapidity range, $|\eta| < 2.4$ [73]. The dimuon mass resolution, $\sigma_{(\mu\mu)}/M_{\mu\mu}$, is measured to be 6.5% at masses around 1 TeV, rising to 12% at 2 TeV [75].

In the dataset corresponding to 5.3 fb^{-1} (as described in Section 5.1), there are 145 721 $\mu^+\mu^-$ pairs passing all the selection requirements and out of these 21 860 pairs have dimuon invariant mass in excess of 120 GeV/c^2. A total of 96 475 dimuon pairs are found corresponding to the 80–100 GeV/c^2 mass window. Events with more than two reconstructed muons surviving all the cuts are very rare; only 14 such events are found in the

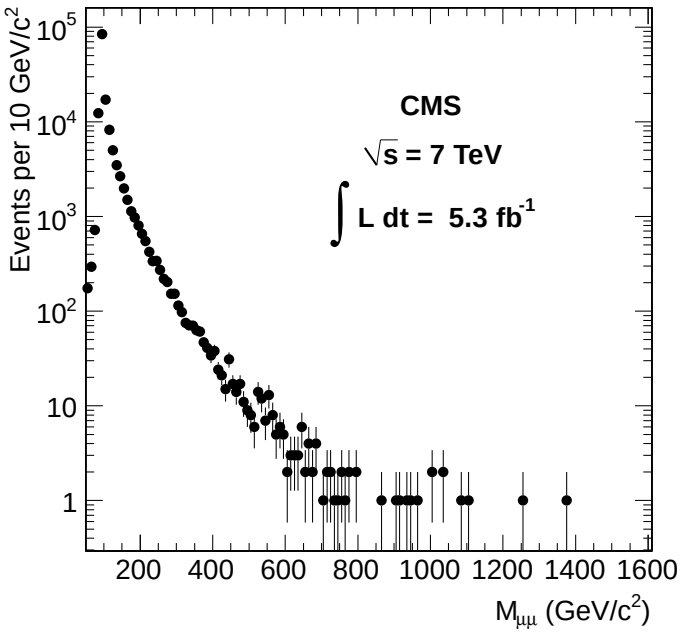


FIG. 5.3. Dimuon mass distribution using the full 2011 dataset after imposing all the selection requirements.

Table 5.3. List of dimuon events in data with masses above 800 GeV/c^2, arranged in descending order of their masses.

Run no.	Lumi section	Event no.	$M_{\mu\mu}$ (GeV/c^2)
179547	319	505584390	1379
177053	506	751646878	1256
167807	428	553377913	1107
171178	112	117297088	1089
167898	1704	1941532760	1034
165617	242	337750503	1030
161312	384	157864904	1009
178738	78	93967757	1001
172163	115	171598658	967
176206	87	102520618	946
176702	334	506891089	933
180093	43	58481079	917
171446	374	452358419	905
171156	565	614885023	870

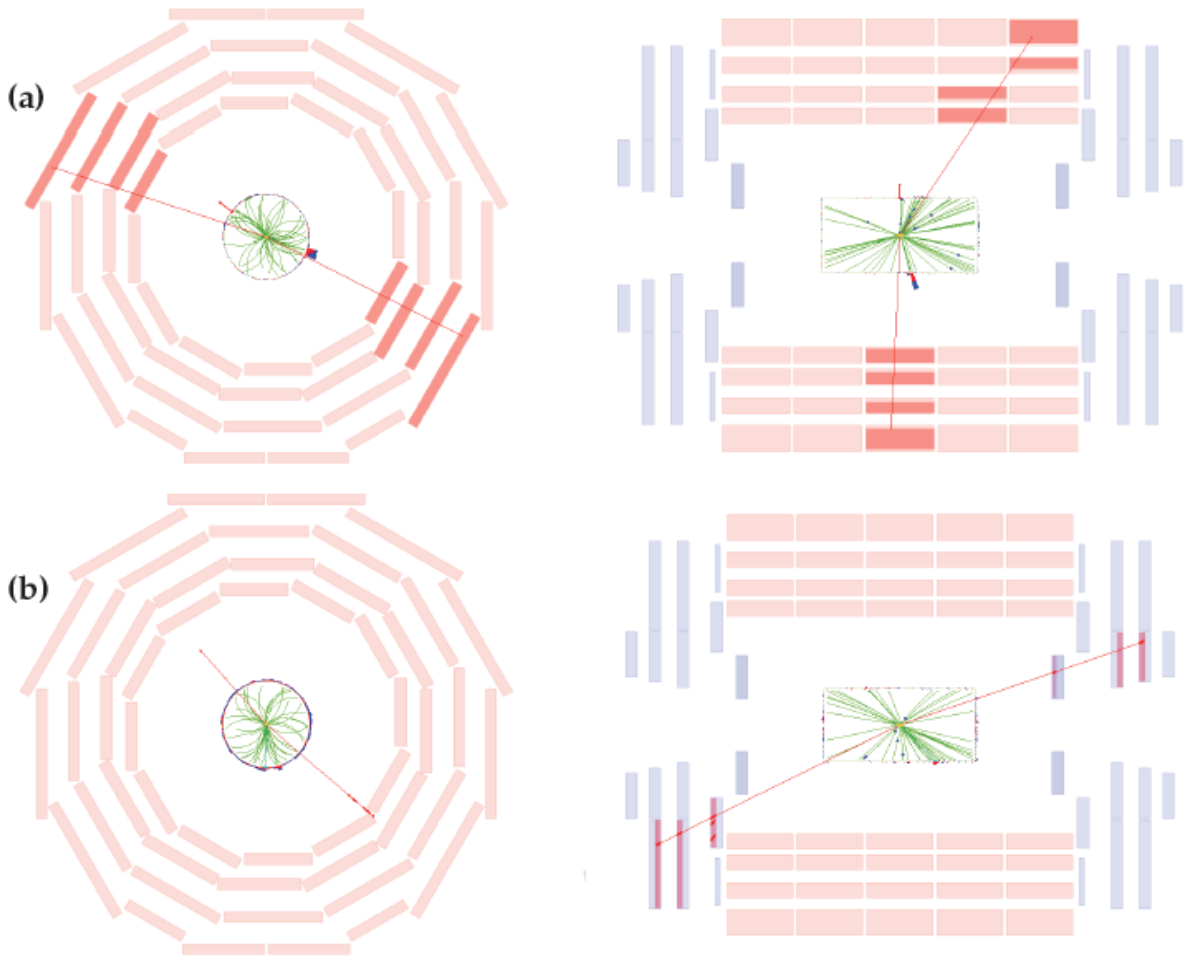


FIG. 5.4. Event displays for the two highest-mass $\mu^+\mu^-$ events. Left (right) side of figures (a) and (b) show the transverse (longitudinal) view. (a) This display corresponds to the highest mass (1379 GeV/c^2) dimuon event with muon kinematic variables: p_T = (686, 622) GeV/c, η = (-0.05, +0.63), and ϕ = (-0.46, +2.82). (b) This display corresponds to the next highest mass (1256 GeV/c^2) dimuon event with muon kinematic variables: p_T = (196, 299) GeV/c, η = (-1.46, +1.80), and ϕ = (-0.72, +2.30).

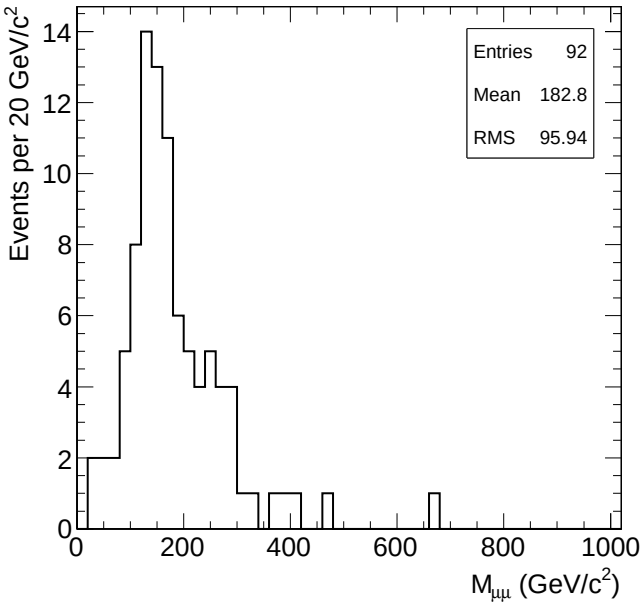


FIG. 5.5. Dimuon mass spectra for same sign muon candidate events.

entire dataset and none of these events have mass above 300 GeV/c^2. Table 5.2 lists the number of events in data that satisfied the successive stages of selection. Figure 5.3 shows the final dimuon mass distribution and Table 5.3 gives the details of dimuon events whose mass is above 800 GeV/c^2. There are 14 dimuon pairs whose mass exceeds 800 GeV/c^2; out of these, 8 events have dimuon mass above 1 TeV. Event displays for the two highest mass $\mu^+\mu^-$ events reconstructed in the CMS detector are shown in Figure 5.4. In the figures, the rectangular pink shapes represent the muon chambers; darker pink chambers contain hits that form part of the reconstructed tracks; the red lines represent muon tracks; the red (blue) trapezoids represent the energy deposits in the electromagnetic calorimeter (hadronic calorimeter); the green lines represent tracks in the inner tracker.

As a control sample, the same sign muon events are also monitored. In the entire dataset, there are only 92 same sign muon candidate events that pass all other selection requirements. Out of the 92 events, 12 events are multiple muon events each consisting of

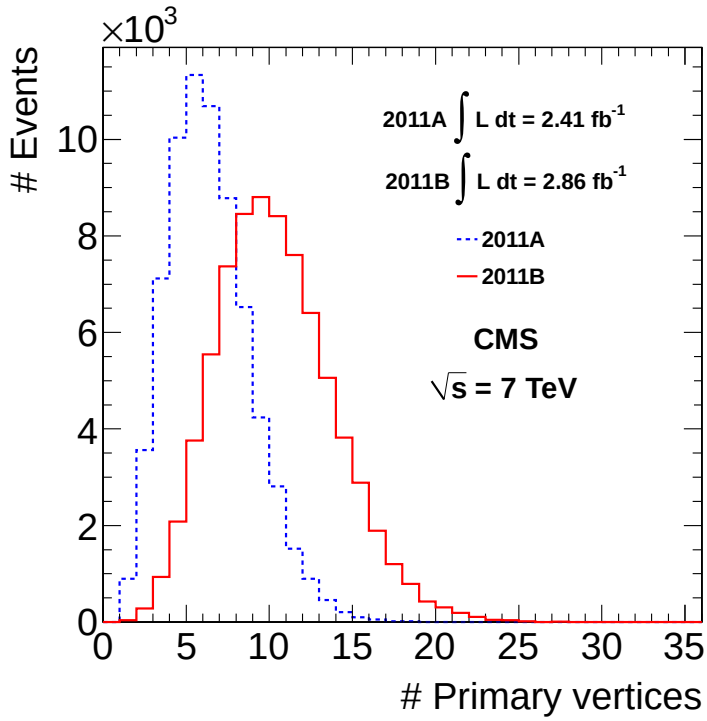


FIG. 5.6. Distribution of the number of primary vertices per event from data for the 2011A and 2011B data taking periods.

3 muons. Figure 5.5 shows the dimuon mass distribution for same sign muon events. It can be seen from the figure that there are no same sign dimuon events with masses above 700 GeV/c^2 showing that the opposite sign requirement is still very efficient at high p_T.

5.3 Event pileup

As mentioned in Section 2.3, as data taking progressed and luminosity increased, the probability of multiple pp interactions occurring within a single proton beam crossing (several hundred bunches of protons form a proton beam) also increased. This is referred to as event pileup. During the course of the 2011 data taking period, an increase in event pileup was observed as the luminosity increased with time. Although not a big effect for

muons, pileup events are in general characterized by a large number of tracks associated with multiple vertices, which, in principle, can lead to a degradation in reconstruction efficiency. Figure 5.6 shows the distribution of the number of primary vertices per event in data, for the 2011A and 2011B data taking periods, in which pileup conditions were quite different. The effect of pileup on reconstruction efficiency is taken into account by weighting simulated events so that the distribution of reconstructed primary vertices per event matches that in data. The weighting factors are determined separately for the two datasets. Simulated event yields for the combined dataset are determined from a luminosity-weighted average. In this analysis, the reconstruction efficiency is found to be insensitive to the variations in event pileup encountered during the data taking period.

Chapter 6

Simulation of signal and background events

In order to test the observed dimuon mass distribution for deviations from the standard model (SM), it is necessary to predict the expected dimuon mass distribution from SM and contact interaction (CI) sources. However, there is no event generator that incorporates all of the required elements: generation of both Drell-Yan (DY) and CI amplitudes, and inclusion of Feynman graphs beyond leading order. For this reason a combination of methods is employed. This chapter discusses in detail how the expected dimuon event yields for CI and relevant SM background processes are predicted.

6.1 Expected signal and corrections

As mentioned in Chapter 4, the Monte Carlo (MC) generator, PYTHIA [47] [48] is used to simulate the left-left isoscalar model (LLIM) CI and SM DY processes. The effect of the detector is determined using a sample of SM DY events that include a simulation of the CMS detector, including the acceptance, effects of the trigger, reconstruction of events, and dimuon mass resolution. The expected number of CI/DY events, ($E_{CI/DY}$), given

by,

$$E_{CI/DY} = [(CI/DY)_{gen} \times K_{QCD} \times K_{QED} \times (A \times M)] + BKG^{non-DY}_{\mu^+\mu^-} \tag{6.1}$$

is the product of the generated number of CI/DY events, $(CI/DY)_{gen}$, a QCD K factor[1] (K_{QCD}), a QED K factor (K_{QED}), to account for higher order corrections, and a factor denoted as "acceptance times migration" ($A \times M$). Contributions of non-DY SM dimuon sources ($BKG^{non-DY}_{\mu^+\mu^-}$) to the expected signal are also taken into account, as described in Section 6.1.3.

All of the multiplicative factors are determined based on a simulated sample of DY events, with the assumption that they can equally be applied to the CI terms. Table 6.1 summarizes the SM DY and non-DY event samples, with CMS detector simulation, used to calculate the $A \times M$ factor and non-DY background rates, respectively. More details on how the $A \times M$ factor is calculated are given in Section 6.1.1. The event generators used are PYTHIA, POWHEG [77–79], and MADGRAPH [80]. As shown in the table, POWHEG and MADGRAPH are next-to-leading order (NLO) generators, whereas PYTHIA is a leading order (LO) generator. Therefore, NLO corrections are applied to all PYTHIA samples by means of a K factor. For non-DY PYTHIA samples, a QCD K factor of 1.3, determined using the NLO generator MC@NLO [49] and LO PYTHIA generator, is applied for the entire mass range considered in this analysis, whereas for PYTHIA DY samples, mass-dependent K factors, as described in Section 6.1.2, are applied.

The CI analysis is limited to a dimuon mass range from 200 to 2000 GeV/c^2. The lower limit is enough above the Z-peak so that a deviation from DY production is observable (as noted in Chapter 4), while the upper limit is chosen large enough to include all events that could be produced for values of Λ accessible with this dataset (see Chapter 5) and not excluded by previous measurements. The minimum mass $M^{Low}_{\mu\mu}$ required in the analysis is varied between the lower and upper limits to optimize the results on Λ, as described in

[1] In general, a K factor is determined from the ratio of the cross section determined using a NLO generator to the cross section determined from a LO generator (σ_{NLO}/σ_{LO}).

Table 6.1. Description of event samples with detector simulation. For the DY process $Z/\gamma^* \rightarrow \mu\mu$, the minimum dimuon mass is indicated as part of the PYTHIA sample name. The cross section σ and integrated luminosity L are given for each sample generated.

Process	Generator	Events	σ(pb)	L(pb^{-1})	Order
$Z/\gamma^* \rightarrow \mu\ \mu$	PYTHIA20	2.15×10^6	1.30×10^3	1.65×10^3	LO
$Z/\gamma^* \rightarrow \mu\ \mu$	PYTHIA120	5.45×10^4	7.90×10^0	6.91×10^3	LO
$Z/\gamma^* \rightarrow \mu\ \mu$	PYTHIA200	5.50×10^4	9.70×10^{-1}	5.67×10^4	LO
$Z/\gamma^* \rightarrow \mu\ \mu$	PYTHIA500	5.50×10^4	2.70×10^{-2}	2.04×10^6	LO
$Z/\gamma^* \rightarrow \mu\ \mu$	PYTHIA800	5.50×10^4	3.10×10^{-3}	1.77×10^7	LO
$Z/\gamma^* \rightarrow \mu\ \mu$	PYTHIA1000	5.50×10^4	9.70×10^{-4}	5.67×10^7	LO
$Z/\gamma^* \rightarrow \tau\ \tau$	PYTHIA	2.03×10^6	1.30×10^3	1.56×10^3	LO
$t\bar{t}$	MADGRAPH	2.40×10^6	1.57×10^2	1.54×10^5	NLO
tW	POWHEG	7.95×10^5	7.90×10^0	1.01×10^5	NLO
$\bar{t}$W	POWHEG	8.02×10^5	7.90×10^0	1.02×10^5	NLO
WW	PYTHIA	4.23×10^6	4.30×10^1	9.83×10^4	LO
WZ	PYTHIA	4.27×10^6	1.80×10^1	2.37×10^5	LO
ZZ	PYTHIA	4.19×10^6	5.90×10^0	7.10×10^5	LO
W+jets	MADGRAPH	2.43×10^7	3.10×10^4	7.82×10^2	NLO
QCD	PYTHIA	1.08×10^6	8.47×10^4	1.28×10^2	LO

Chapter 8. All samples used in this analysis are produced at $\sqrt{s} = 7$ TeV center-of-mass energy using the above mentioned event generators. The samples listed in Table 6.1 use the Geant4 [81] toolkit to simulate the response of the CMS detector.

6.1.1 Detector acceptance times mass migration

For a given value of $M_{\mu\mu}^{Low}$, $A \times M$ is given by the ratio of the number of DY events reconstructed with mass above $M_{\mu\mu}^{Low}$ to the number of DY events generated with mass above $M_{\mu\mu}^{Low}$. Some of the reconstructed events with generator mass below $M_{\mu\mu}^{Low}$ are included due to the mass resolution. Allowing the net migration of dimuons mimics what would be observed in data due to finite mass resolution. The resulting $A \times M$ as a function of $M_{\mu\mu}^{Low}$ is plotted in Figure. 6.1 and values are given in Table 6.2. All uncertainties shown in the table are statistical in nature.

The $A \times M$ values are initially calculated for 100 GeV/c^2 bins (200–300, 300–400, . .

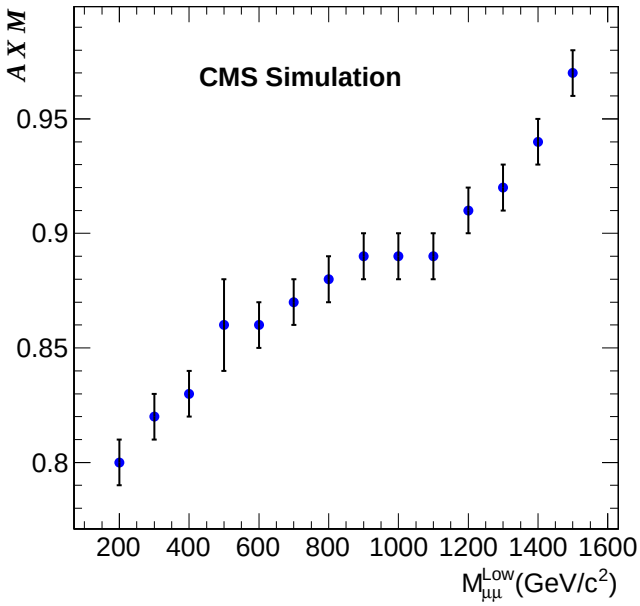


FIG. 6.1. Acceptance times migration for $M > M_{\mu\mu}^{Low}$.

Table 6.2. Multiplicative factors for $M > M_{\mu\mu}^{Low}$. The uncertainties shown are statistical. The systematic uncertainties are described in Chapter 7.

$M_{\mu\mu}^{Low}$ (GeV/c^2)	$A \times M$	QCD K factor	EW K factor
200	0.80±0.01	1.259±0.004	1.01
300	0.82±0.01	1.296±0.004	0.99
400	0.83±0.01	1.321±0.003	0.97
500	0.86±0.02	1.334±0.003	0.95
600	0.86±0.01	1.339±0.003	0.94
700	0.87±0.01	1.343±0.003	0.92
800	0.88±0.01	1.343±0.003	0.91
900	0.88±0.01	1.339±0.002	0.90
1000	0.89±0.01	1.334±0.002	0.89
1100	0.89±0.01	1.324±0.002	0.88
1200	0.91±0.01	1.312±0.002	0.88
1300	0.92±0.01	1.291±0.003	0.87
1400	0.94±0.01	1.264±0.003	0.87
1500	0.97±0.01	1.220±0.004	0.86

. . , 1900–2000 GeV/c^2) based on Z/$\gamma^* \rightarrow \mu\mu$ PYTHIA DY samples with different dimuon mass thresholds, indicated as part of the PYTHIA sample name as listed in Table 6.1. The PYTHIA DY samples are paired in a one-to-one fashion with the 100 GeV/c^2 mass bins (making sure that the chosen PYTHIA sample has a dimuon mass threshold that is lower than the corresponding $M_{\mu\mu}^{Low}$) so as to optimize statistics while at the same time allowing for mass resolution smearing effects. This kind of pairing ensures usage of maximum available statistics for a given mass range and results in minimal statistical errors as shown in Table 6.2. The calculation of $A \times M$ takes into account the reconstruction and generator selection criteria as described in Chapter 5. The effect of pileup on reconstruction efficiency is included following the weighting method described in Section 5.3.

From Figure 6.1, one can see that the boost due to resolution smearing becomes significant for $M_{\mu\mu}^{Low} > 600$ GeV/c^2 and steadily increases with dimuon mass. This behavior can be understood from effects of mass migration as illustrated in Figure 6.2, which shows the reconstructed dimuon mass associated with a particular generator mass window. For illustrative purposes, the distributions are shown only for some of the 100 GeV/c^2 mass bins. From this figure, it can be seen that the reconstructed mass distributions are much broader than the generated ones, with more broadening with increasing mass. This results in an increasing $A \times M$ with increasing mass. The behavior of $A \times M$ at low mass ($M_{\mu\mu}^{Low} = 200$–500 GeV/c^2) can be understood from the 45 GeV/c p_T cut; this effect becomes less significant as one moves to the high mass region where mass migration starts to dominate the $A \times M$.

Dependence on Λ

In order to validate the event yields predicted by the combination of the CI/DY generator and the $A \times M$ factor based on DY events, the event yields are compared with those predicted using only CI/DY events that are simulated in the detector. This

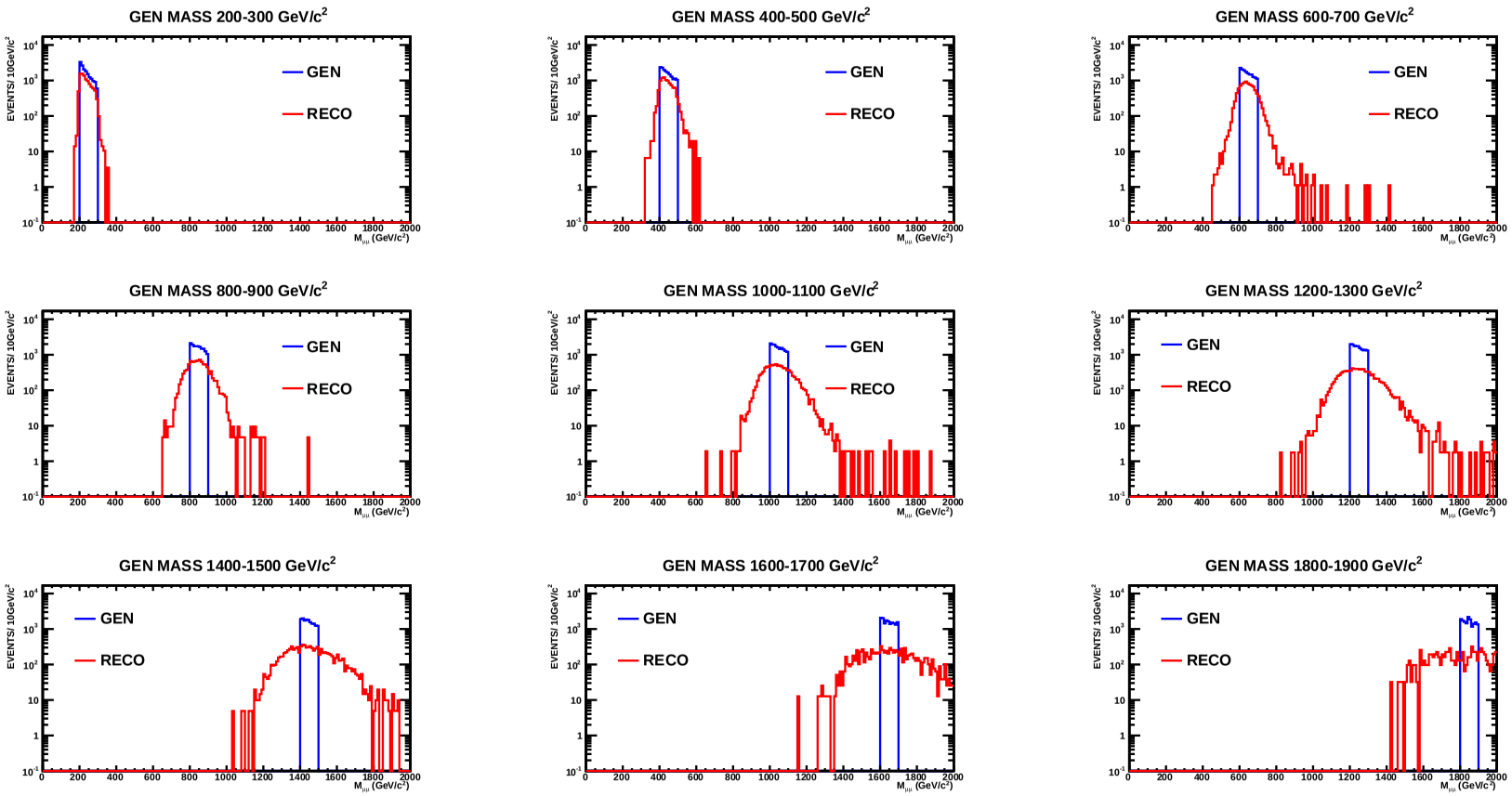


FIG. 6.2. Reconstructed dimuon mass distributions associated with a particular generator mass window for 100 GeV/c^2 bins. All distributions are normalized to correspond to the luminosity of the 200–300 GeV/c^2 bin. The figure illustrates the gradually degrading mass resolution with increasing dimuon mass.

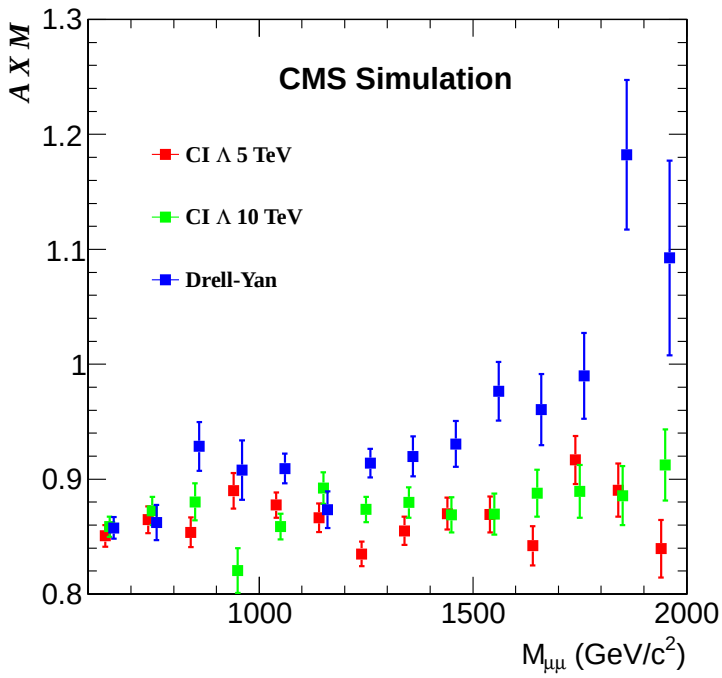


FIG. 6.3. Comparison of $A \times M$ values in 100 GeV/c^2 bins for CI and DY production. The plot starts at 600 GeV/c^2 on the x-axis.

study is performed for the cases of constructive interference with $\Lambda = 5$ and 10 TeV, which represent a wide range of possible CI/DY cross sections. Figure 6.3 shows the comparison of $A \times M$ values for $\Lambda = 5$ and 10 TeV with the values obtained from DY production. It can be seen from the figure that, for dimuon masses above 900 GeV/c^2, the $A \times M$ values for the DY process slightly deviate from those corresponding to the CI/DY Λ values. The deviation is due to resolution smearing, which is sensitive to the shape of the cross section. This can be better understood by looking at the cross section distributions shown in Figures 4.2(a) and 4.2(b) (Chapter 4); one can see from these figures that the cross sections corresponding to constructive interference are significantly higher than the DY cross section even for $\Lambda = 10$ TeV. The results differ by at most 3%, consistent with the statistical precision of the study. The systematic uncertainty on

$A \times M$ is conservatively assigned this value.

6.1.2 K factors

As mentioned in the beginning of the chapter, the SM DY and CI signal samples used in this analysis are generated using the PYTHIA MC program which is a LO event generator. However, considering only LO processes can seriously underestimate the cross sections at a high-energy hadron collider and thus may have serious bearing on the discovery potential of such experiments. Hence it is important to incorporate higher-order terms, at least up to NLO, in the cross section prediction.

Higher-order corrections are divided into two categories: the NLO QCD corrections involving strong interactions and NLO QED corrections involving electroweak (EW) interactions. In the case of DY/CI production, only initial state quarks are affected by higher order strong processes (such as quark/gluon loops or gluon radiation), due to their color charge, whereas, the EW processes affect both initial state quarks and final state leptons as they both carry electric charge. The NLO QED corrections include contributions from processes such as initial and final state photon radiation and virtual EW loop corrections.

QCD K factor

The QCD NLO K factors are calculated from the ratio of MC@NLO to PYTHIA DY event yields, at the generator level. The PYTHIA and MC@NLO samples are subject to the cuts, $|\eta| < 2.6$ and $p_T > 40$ GeV/c. Figure 6.4 and Table 6.2 show the resulting K factors as a function of $M_{\mu\mu}^{Low}$. The large samples of simulated events result in statistical uncertainties less than 0.5%. It is important to note here that the QCD K factor for DY production is always greater than one showing that the NLO processes in this case constructively interfere with LO processes resulting in an enhancement of the overall cross section. The systematic uncertainty is assigned the value 3%, the size of the cor-

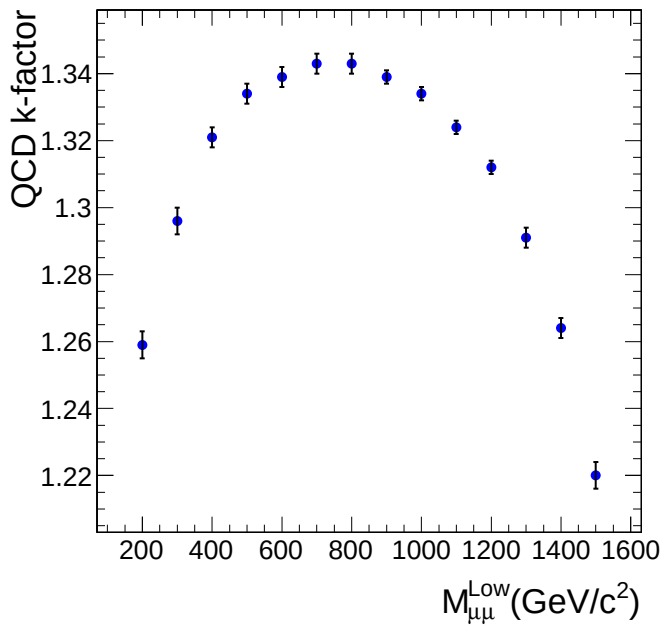


FIG. 6.4. QCD K factors for $M > M_{\mu\mu}^{Low}$. The values are plotted at the threshold points.

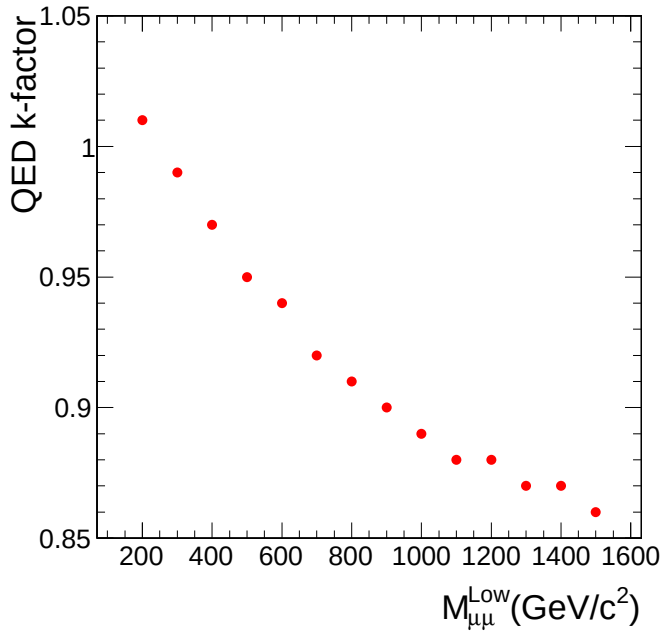


FIG. 6.5. QED K factors for $M > M_{\mu\mu}^{Low}$. The values are plotted at the threshold points.

rection [83] between next-to-next-to-leading order (NNLO) and NLO DY cross sections. For SM processes other than DY production (shown in Table 6.1), the QCD k-factor is found, independent of dimuon mass, from the ratio of the cross section determined using MC@NLO to the cross section determined from PYTHIA.

QED K factor

A mass-dependent QED K factor determined using the HORACE generator as reported in Ref. [82] is used in this analysis. The values of the QED K factor, as a function of $M_{\mu\mu}^{Low}$, are shown in Figure 6.5 and Table 6.2. From the figure, one can see that except for $M_{\mu\mu}^{Low} = 200$ GeV/c^2, the QED K factor is always less than 1, implying that the NLO processes in this case destructively interfere with LO processes decreasing the overall cross section. The systematic uncertainties on the QED K factor are discussed in Chapter 7.

6.1.3 SM backgrounds

The dominant background in this analysis is the SM $Z/\gamma^* \to \mu\mu$ production. Since the dimuon signature for the $Z/\gamma^* \to \mu\mu$ process is indistinguishable from the CI signal process in the detector, the SM DY background is considered an "irreducible" background. PYTHIA simulated $Z/\gamma^* \to \mu\mu$ samples (as shown in Table 6.1) are used to evaluate its contribution to the CI signal. The SM non-DY reducible[2] backgrounds that are relevant to this analysis, in decreasing order of importance, are: $t\bar{t}$, diboson ($WW/WZ/ZZ$), W (including W+jets, $\bar{t}W$), and $Z \to \tau\tau$. Other QCD backgrounds (such as $b\bar{b}$ and $c\bar{c}$), even though relevant, are substantially reduced by applying the tracker based isolation requirement (see Chapter 5).

Using the simulation samples listed in Table 6.1, event yields are predicted for the listed non-DY SM background processes, as shown in Table 6.3. The yields are given as a function of $M_{\mu\mu}^{Low}$ and correspond to the integrated luminosity of the data, 5277

[2]Reducible backgrounds have distinguishable signatures in the detector compared to the signal process and can be reduced by appropriate selection cuts.

Table 6.3. Expected event yields for non-DY SM backgrounds. DY event yields are shown for comparison. The uncertainties shown are statistical.

$M_{\mu\mu}^{Low}$ (GeV/c^2)	DY	$Z \to \tau\tau$	W+Jets+$\bar{t}W$+tW	$t\bar{t}$	Diboson	sum non-DY
200	3509 ± 18	6.96 ± 4.14	47.90 ± 1.35	454 ± 3	123 ± 2	632.3 ± 5.9
300	863.1 ± 9.0	0	12.82 ± 0.70	104 ± 2	38.6 ± 1.2	155.9 ± 2.1
400	306.8 ± 5.4	0	3.32 ± 0.35	26.0 ± 0.8	12.7 ± 0.7	42.0 ± 1.1
500	126.5 ± 3.6	0	1.02 ± 0.20	8.19 ± 0.46	5.07 ± 0.41	14.3 ± 0.64
600	56.98 ± 0.21	0	0.29 ± 0.11	2.92 ± 0.27	2.42 ± 0.28	5.63 ± 0.41
700	28.29 ± 0.14	0	0.07 ± 0.05	1.12 ± 0.17	0.86 ± 0.16	2.06 ± 0.24
800	14.85 ± 0.10	0	0.07 ± 0.05	0.34 ± 0.09	0.51 ± 0.12	0.92 ± 0.16
900	8.08 ± 0.08	0	0.07 ± 0.05	0.05 ± 0.03	0.25 ± 0.08	0.36 ± 0.10
1000	4.51 ± 0.06	0	0.07 ± 0.05	0.05 ± 0.03	0.10 ± 0.05	0.21 ± 0.08
1100	2.55 ± 0.04	0	0.07 ± 0.05	0.05 ± 0.03	0.09 ± 0.05	0.20 ± 0.08
1200	1.53 ± 0.03	0	0.07 ± 0.05	0	0.01 ± 0.01	0.08 ± 0.05
1300	0.93 ± 0.03	0	0.07 ± 0.05	0	0.01 ± 0.01	0.08 ± 0.05
1400	0.56 ± 0.02	0	0.07 ± 0.05	0	0.01 ± 0.01	0.08 ± 0.05
1500	0.32 ± 0.01	0	0.07 ± 0.05	0	0	0.07 ± 0.05

pb^{-1}. Furthermore, they take into account the effects of pileup, and reconstruction and generator selection criteria, as described in Chapter 5. For comparison, Table 6.3 also shows the expected yields for DY events. For $M_{\mu\mu}^{Low} > 1000$ GeV/c^2, the statistical uncertainty in the non-DY background is large, but the absolute yield is much smaller than that for DY.

The other background expected in the detector is the cosmic muon background due to the high-energy atmospheric cosmic rays that manage to penetrate the detector. Sitting about 100 m deep underground, CMS is well shielded from cosmic rays (only about 1% of cosmic rays at the surface of the Earth reach the detector). The three access shafts of CMS cause increased cosmic ray acceptance in some parts of the detector. Hence it is important to significantly reduce this background. Studies [73] show that the one *good* primary vertex and dimuon angle cuts (as discussed in Chapter 5) substantially reduce the cosmic muon background.

6.2 Predicted event yields

A total of 500k events were produced, with $M_{\mu\mu} > 120$ GeV/c^2, for each of 32 signal samples and a SM DY sample. The 32 CI/DY samples (16 samples with constructive

interference and 16 with destructive interference) correspond to Λ values ranging from 3 TeV to 18 TeV in steps of 1 TeV. At the PYTHIA generator-level, all samples are subject to the generator selection criteria, mentioned in Chapter 5, and are produced above an invariant mass of 120 GeV/c^2 to increase the number of events in the signal region. Figure 4.1 shows the corresponding LO production cross sections. Similarly, in order to increase statistics in the high mass region above 600 GeV/c^2, another set of 100k events were produced, with $M_{\mu\mu} > 600$ GeV/c^2, for each of the 32 signal samples and a SM DY sample.

Using the methods described in the previous section (Eq. 6.1), expected event yields corresponding to the integrated luminosity of the data (5277 pb^{-1}) are predicted as a function of Λ and $M_{\mu\mu}^{Low}$. The predictions for destructive and constructive interference are given in Tables 6.4 and 6.5. The DY expected event yields in these tables can be compared with the data values for a given $M_{\mu\mu}^{Low}$.

As seen from Table 6.4, for destructive interference, there are regions of the $M_{\mu\mu}^{Low} - \Lambda$ parameter space where the predicted number of CI/DY events are less than the predicted number of DY events. An example is $M_{\mu\mu}^{Low} = 200$ GeV/c^2, and $\Lambda = 8$ TeV. The statistical method used in this analysis to get results, described in Chapter 7, cannot accommodate an apparent "negative signal. However, for the region of parameter space $M_{\mu\mu}^{Low} > 600$ GeV/c^2 and $\Lambda < 12$ TeV of interest in which the destructive result is optimized, the expected signal is always positive. For constructive interference the expected signal is positive for all values of Λ and $M_{\mu\mu}^{Low}$.

Table 6.4. Observed and expected number of events for $M_{\mu\mu} > M_{\mu\mu}^{Low}$. The expected yields are shown for DY production and CI/DY production for destructive interference with given Λ values. Expected yields include contributions from non-DY backgrounds.

$M_{\mu\mu}^{Low}$ (GeV/c^2)	200	300	400	500	600	700	800	900	1000	1100	1200	1300	1400	1500
data	4320	1013	338.0	141.0	57.0	28.0	14.0	13.0	8.0	3.0	2.0	1.0	0.0	0.0
DY	4141	1019	348.8	140.8	62.6	30.3	15.8	8.4	4.7	2.8	1.6	1.0	0.6	0.4
Λ (TeV)														
18	4092	1006	334.3	136.8	59.6	28.7	14.9	8.0	4.5	2.7	1.6	1.0	0.6	0.4
17	4105	1020	338.5	137.1	59.5	28.5	14.9	8.1	4.6	2.7	1.6	1.0	0.6	0.4
16	4110	1015	337.8	137.5	59.6	28.7	15.1	8.2	4.7	2.8	1.7	1.0	0.7	0.4
15	4101	1012	338.1	138.2	59.9	29.0	15.3	8.4	4.9	3.0	1.8	1.1	0.8	0.5
14	4079	1005	333.7	136.3	59.9	29.2	15.6	8.7	5.2	3.2	2.0	1.3	0.9	0.6
13	4090	1014	336.2	136.7	60.9	30.0	16.3	9.3	5.7	3.6	2.3	1.5	1.0	0.7
12	4078	1007	337.9	141.3	61.7	31.1	17.4	10.3	6.4	4.3	2.8	2.0	1.3	0.9
11	4059	1002	331.9	138.3	64.2	33.1	19.2	11.7	7.6	5.2	3.6	2.5	1.7	1.2
10	4073	1006	342.2	143.8	68.5	36.8	22.1	14.2	9.6	6.8	4.7	3.4	2.4	1.7
9	4064	1017	353.4	151.4	75.9	43.8	28.2	19.2	13.6	9.9	7.1	5.2	3.7	2.6
8	4077	1032	366.6	168.0	90.6	56.2	38.4	27.4	20.1	14.8	10.9	8.0	5.8	4.0
7	4153	1081	407.8	202.1	120.9	82.0	60.0	45.3	32.9	24.9	18.1	13.4	9.0	6.0
6	4195	1165	494.5	283.9	187.6	136.2	104.4	80.2	60.4	46.9	34.1	23.2	17.0	11.0
5	4450	1403	713.0	479.1	348.8	270.6	213.2	165.9	128.5	100.5	77.2	58.1	41.6	28.3
4	5256	2170	1382	1048	807.6	641.4	515.6	401.9	315.6	241.8	180.6	130.2	94.8	66.3
3	8664	5253	4017	3268	2593	2081	1668	1299	1017.8	773.5	595.2	447.9	321.1	224.0

Table 6.5. Observed and expected number of events as in Table 6.4. Here CI/DY predictions are for constructive interference.

$M_{\mu\mu}^{Low}$ (GeV/c^2)	200	300	400	500	600	700	800	900	1000	1100	1200	1300	1400	1500
data	4320	1013	338.0	141.0	57.0	28.0	14.0	13.0	8.0	3.0	2.0	1.0	0.0	0.0
DY	4141	1019	348.8	140.8	62.6	30.3	15.8	8.4	4.7	2.8	1.6	1.0	0.6	0.4
Λ (TeV)														
18	4151	1053	364.6	150.6	69.7	35.2	19.4	11.1	6.7	4.2	2.6	1.6	1.1	0.7
17	4156	1056	364.4	152.2	71.0	36.2	20.1	11.6	7.0	4.4	2.8	1.8	1.2	0.8
16	4194	1070	370.8	156.7	72.3	37.3	21.0	12.2	7.5	4.8	3.1	2.0	1.3	0.8
15	4182	1062	371.2	159.4	73.9	38.4	21.8	12.9	8.0	5.1	3.2	2.1	1.4	0.9
14	4220	1069	374.1	158.0	76.6	40.4	23.3	13.9	8.7	5.7	3.7	2.5	1.7	1.1
13	4189	1094	383.6	167.6	80.1	43.1	25.2	15.3	9.8	6.5	4.3	3.0	2.0	1.3
12	4247	1096	385.0	173.9	84.8	46.7	28.0	17.4	11.4	7.7	5.2	3.5	2.4	1.6
11	4255	1113	394.9	178.1	91.1	51.5	31.7	20.1	13.4	9.1	6.3	4.3	3.0	2.0
10	4288	1137	412.3	188.2	100.7	59.0	37.5	24.7	16.9	11.7	8.2	5.7	4.0	2.7
9	4351	1175	447.2	219.2	116.4	71.1	46.6	31.7	22.1	15.6	11.1	7.9	5.5	3.8
8	4398	1229	477.4	241.9	142.2	90.6	62.2	43.5	31.2	22.6	16.3	11.7	8.2	5.6
7	4551	1342	572.9	313.7	186.3	124.4	90.4	65.0	46.8	32.7	24.1	17.2	12.4	8.5
6	4785	1513	708.1	423.7	277.1	193.7	142.8	106.4	78.6	59.6	43.2	31.5	23.3	15.4
5	5266	1905	1024	689.4	481.8	357.6	279.6	211.5	159.9	121.4	90.3	66.1	47.1	33.1
4	6551	2946	1870	1375	1027	791.4	612.0	471.8	364.7	277.2	207.2	151.3	109.8	75.0
3	10967	6639	4882	3844	2947	2328	1830	1421	1086	821.5	617.7	464.2	339.5	234.4

Chapter 7

Systematic uncertainties

Since the statistical analysis method, described in the next chapter, incorporates the systematic uncertainties in the calculation of final results, a description of the various systematic uncertainties associated with this analysis is given in advance in this chapter. Systematic uncertainties, in general, arise from an imperfect knowledge of the detector, assumptions made by the experimenter, or the model used to make inferences based on the observed data. The main sources of systematic uncertainties on the predicted signal and background event yields can be categorized into *theoretical* and *experimental*. The following sections will give a detailed description of the two types of uncertainties.

7.1 Theoretical sources

In the contact interaction analysis, the theoretical uncertainties come from uncertainties in the proton structure (parton distribution functions (PDFs)) (Section 2.3) and in the higher-order QCD and electroweak K factors (Section 6.1) applied to the contact interaction (CI) signal and Drell-Yan (DY) $Z \to \mu\mu$ samples. A description of these uncertainties is given below.

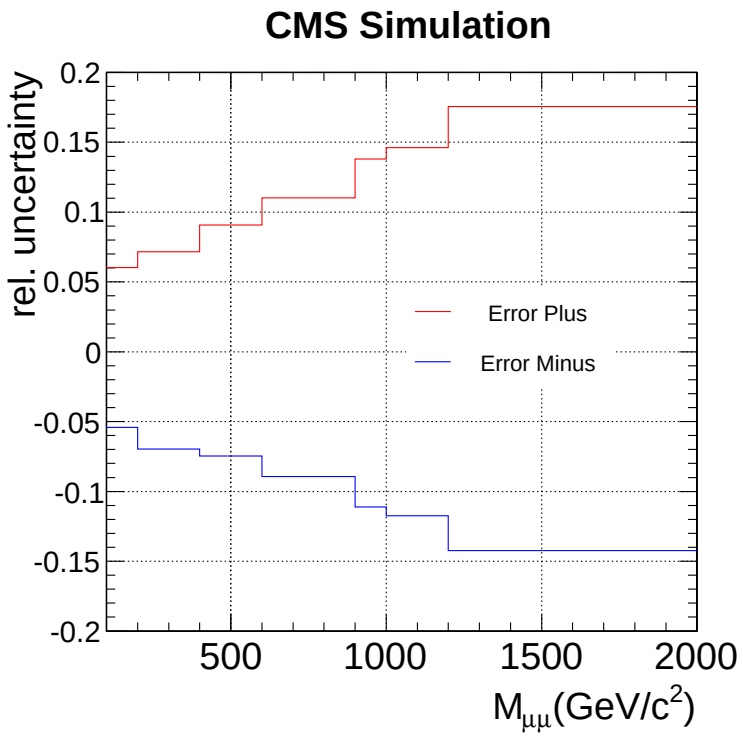


FIG. 7.1. Maximal positive and negative PDF uncertainties as a function of mass.

PDF uncertainties

The uncertainties in the dimuon cross section due to uncertainties in the proton structure are determined using the "modified tolerance method" as described in Ref. [93]. Relying on the PDF4LHC study [94], this analysis also uses the CTEQ66 [95] next-to-leading order (NLO) PDF set (with 90% C.L.) to evaluate systematic uncertainties coming from PDFs. The CTEQ66 PDF set offers a few advantages: it has been calibrated with data from hadron colliders, notably the Tevatron, and the DY simulation samples used in this analysis were generated with the CTEQ66 LO PDF set. The CTEQ66 NLO PDF set consists of 22 orthogonal independent parameters (related to the attributes of the partons), which can be considered as eigenvectors of the PDF parameter space. The uncertainty in the parton structure is determined by fluctuating these parameters

Table 7.1. Maximal positive and negative PDF uncertainties evaluated with CTEQ66 NLO PDF set using the modified tolerance method.

dimuon mass (GeV/c^2)	ΔX^+_{max} (%)	ΔX^-_{max} (%)
80–200	7.2	6.9
200–400	7.2	7.0
400–600	9.0	7.5
600–900	11.9	9.0
900–1000	13.8	11.0
1000–1200	14.6	11.7
1200–2000	17.6	14.2

independently in both positive and negative directions. Table 7.1 and Fig. 7.1 show the plus (ΔX^+_{max}) and minus (ΔX^-_{max}) PDF fluctuations derived for a wide range of dimuon mass. One can see from the figure that the PDF uncertainty grows from about 7% at the Z peak to about 14% at 1 TeV and about 17% for dimuon mass at 2 TeV.

QCD and QED K factors

As mentioned previously, the CI signal and SM DY samples are generated using the PYTHIA Monte Carlo which is basically a leading order (LO) event generator. So, in order to account for higher order diagrams, up to NLO, QCD and QED K factors are applied to the dimuon event yields. The generation of large samples made it possible to keep the statistical errors on QCD K factors to less than 0.5% (as shown in Table 6.2). As mentioned before, the systematic uncertainty on QCD K factor is assigned the value 3%, the size of the correction [83] between next-to-next-to-leading order (NNLO) and NLO DY cross sections. In the case of QED K factors, since the effect of QED corrections on the new physics of CI is unknown, following the conservative approach of [38], the systematic uncertainty is assigned as the size of the correction, $|(\text{QED k-factor}) - 1|$.

Table 7.2. Sources of systematic uncertainty. Where appropriate, the values are quoted for $M_{\mu\mu} > 700$ GeV/c^2, $\Lambda = 13$ TeV, and for constructive interference.

Source	Uncertainty (%)
Integrated luminosity	2.2
Acceptance	3.0
Background	14.8
Maximum PDF variation	12.3
QED K factor	7.7
QCD K factor	3.0
DY event yield	0.5
non-DY event yield	11.7

7.2 Experimental sources

The experimental uncertainties are estimated from a variety of sources: uncertainty in luminosity measurement, SM background estimates (*i.e.,* DY and non-DY event yields), detector acceptance times migration ($A \times M$, see Section 6.1) which includes geometrical acceptance, trigger and reconstruction efficiencies, and dimuon mass resolution. The determination of the uncertainty in integrated luminosity is described in [76] and is measured as 2.2%. With this uncertainty, the total integrated luminosity can be expressed as 5277 ± 116 pb^{-1}. The uncertainty in the factor $A \times M$ is dominated by the difference between acceptances determined using DY and CI/DY simulations, as discussed in Section 6.1.1. The systematic uncertainty on $A \times M$ is conservatively assigned a value of 3%. The statistical errors on DY and non-DY yields, as shown in Table 6.3, are also considered in the calculation of results.

7.3 Summary

Table 7.2 summarizes all the systematic uncertainties mentioned above. Several of the sources of uncertainty vary with dimuon mass or depend on the range of Λ or the choice for the sign of the interference; in these cases the values in Table 7.2 are quoted for

$M_{\mu\mu}^{Low} > 700$ GeV/c^2, $\Lambda = 13$ TeV, and $\eta =$ -1 (constructive interference) as these points correspond very closely to the largest of the final limits (as will be shown in the next chapter). In keeping with the principle of stating the limits as conservatively as possible, the larger of the $+$ or $-$ PDF variations is chosen. The PDF variation is added in quadrature with the uncertainties on the QCD K factor, and QED K factor. By way of contrast, although non-DY event yields have a large relative uncertainty, the effect on the final results is minimal given the reduced contribution of non-DY sources relative to DY production. The systematic uncertainties on integrated luminosity, acceptance, and expected signal and background are included as "nuisance parameters" in the limit setting procedure, described in the next chapter.

Chapter 8

Statistical analysis method

Following the procedures outlined in previous chapters, using the observed and expected yields after all simulation corrections are applied, a search for contact interaction (CI) in the high-mass region of the dimuon mass spectrum is performed. This chapter starts by introducing the frequentist approach for statistical inference and outlines the procedure to test the agreement of the observed dimuon mass spectrum with the predicted distribution including the background contribution. As will be seen in Section 8.2, since the dimuon distribution derived using the CMS 2011 dataset is found to be consistent with the expected contributions from DY and other SM background sources, exclusion lower limits on the compositeness scale Λ are established for the left-left isoscalar model (LLIM). Limits are determined separately for destructive and constructive interference using a modified version of the classical frequentist method described in Section 8.1. Finally the effect of individual systematic uncertainties on the results is discussed.

8.1 The modified-frequentist method

In high energy physics, different statistical approaches are used to characterize the absence of a signal or establish a significant excess of events. The two statistical approaches

commonly used in high energy physics for characterizing the non-observation of a signal are the *Bayesian*[1] and *Frequentist* methods. Given that no significant excess of events is observed in the CI analysis, as will be seen in the next section, exclusion limits are set using a modified version of the classical frequentist method, also known as the CL_s method. This section starts with a general description of the classical frequentist method and the CL_s modification applied to it. A brief explanation of how the expected and observed limits are calculated in the CL_s method is also given. A dedicated discussion of the several input parameters and options used in the limit setting method, specific to the CI analysis, is given in Section 8.3.

In the following subsections, the expected signal will be denoted as 's' and backgrounds as 'b'. The limit setting procedure depends on several parameters which can be categorized as either nuisance parameters or parameters of interest. Parameters of interest are the parameters that are being constrained in a given analysis in the absence of a signal. A nuisance parameter is any parameter that is not under investigation in an experiment but still has an impact on the predictions. Examples of nuisance parameters include detector efficiencies, parton density functions, *etc.* Nuisance parameters will be denoted by θ and parameters of interest by μ in the following subsections. Since the signal and background predictions are subject to multiple uncertainties, handled by nuisance parameters, they generally become functions of nuisance parameters, $s(\theta)$ and $b(\theta)$.

As a convention, it is common to require a 95% confidence level (C.L.) for excluding a signal. *Confidence level* is a measure of the reliability of a result. A 95% C.L. means that in a set of similarly constructed experiments, 95 out of 100 will yield a result that can be expected to be within a specified range.

Classical frequentist approach

The classical frequentist approach begins from defining a test statistic, q_μ, that is

[1] Unlike the Frequentist method, which will be described in the next section, in the Bayesian statistical approach, results are based on a "prior distribution" of the parameter of interest on which limits are set. This makes the results sensitive to the choice of the prior distribution.

designed to discriminate signal-like and background-like events based on their agreement with a set of data. In general there are multiple ways of defining the test statistic, but, by the Neyman-Pearson lemma [84], the ratio of likelihoods is the most powerful discriminator. For many reasons [85], the actual quantity used is a logarithm of the likelihood ratio[2]:

$$q_\mu = -2 \ln \frac{\mathcal{L}(data|s(\mu)+b)}{\mathcal{L}(data|b)} \tag{8.1}$$

where *data* corresponds to the actual observed events or pseudo data, $\mathcal{L}(data|s(\mu)+b)$ is simply a product of Poisson probabilities for the observed number of events, given the expected signal and background rates.

Having constructed the test statistic, the next step is to construct probability density functions[3] (*pdfs*) of q_μ under the *signal* + *background* hypothesis. Since the analytical evaluation of the *pdfs* is generally impossible, especially when nuisance parameters are involved, one way to approximate the *pdfs* is to toss a large number of toy pseudo-observations and evaluate q_μ using the same Poisson probabilities.

Using these *pdfs*, one can then evaluate the probability, $CL_{s+b} = P(q_\mu < q_\mu^{data}|s(\mu)+b)$ of observing a measurement q_μ less than the observed value q_μ^{data}, with the *signal* + *background* hypothesis. In the classical frequentist approach, if $\mathrm{CL}_{s+b} = 0.05$, then one says that the signal is excluded at 95% C.L. However, the classical definition fails to obtain sensible exclusion limits on the signal when an experimental observation appears consistent with little or no signal together with a downward fluctuation of the background [86]. To improve this situation, a number of solutions have been suggested, one of which is the CL_s method. The CL_s method is one of the popular methods used at the LHC and is among the three methods described in the PDG [19].

[2]Please note that this test statistic is not used at the LHC, where the prescribed profile-likelihood test statistic is used.

[3]Any outcome of a measurement is subject to statistical fluctuations and can assume different values in independent measurements. The distribution of these values, assuming that the prediction describes the expected value of the outcome, is referred to as a *pdf*.

The CL_s approach

The CL_s method [86–88] was originally introduced at the time of LEP (large electron-positron collider), for the statistical analysis of Higgs searches to prevent the previously mentioned problem of making *strong* exclusions based on a very weak-signal when downward fluctuations occur. In the CL_s method, in addition to CL_{s+b}, one also calculates CL_b, for *background-only* hypothesis. It is the ratio of these two probabilities, given by the quantity, CL_s,

$$CL_s = \frac{CL_{s+b}}{1 - CL_b} = \frac{P(q_\mu < q_\mu^{data}|s(\mu) + b)}{1 - P(q_\mu \geq q_\mu^{data}|b)} \tag{8.2}$$

that defines the 95% C.L. exclusion, *i.e.,* the value of confidence CL_s is required to be less than or equal to 0.05 to exclude the signal at 95% C.L.

As mentioned earlier, the "LHC–style" CL_s approach uses a profile-likelihood ratio test statistic [89], given by,

$$q_\mu = -2 \ln \frac{\mathcal{L}(data|s(\mu) + b)}{\mathcal{L}(data|s(\hat{\mu}) + b)}, \qquad 0 \leq \hat{\mu} \leq \mu \tag{8.3}$$

where the parameter estimator $\hat{\mu}$ maximizes the likelihood $\mathcal{L}(data|s(\mu) + b)$; the lower constraint $0 \leq \hat{\mu}$ is dictated by physics implying that the signal rate is always positive, whereas the upper constraint $\hat{\mu} \leq \mu$ is imposed to ensure a one-sided confidence interval. The advantage of using this test statistic is that its pdf distribution can be approximated by asymptotic formulae based on the Wilks and Wald theorems [89].

Including systematic uncertainties: LHC–style

The systematic uncertainties on signal and background can be introduced in two ways, via modifications to the test statistic itself or the way pseudo data are generated. In the LHC–style CL_s method, the likelihood definition is modified to include systematic

uncertainties,

$$\mathcal{L}(data|s(\mu,\theta)+b(\theta)) = \mathrm{P}oisson(data|s(\mu,\theta)+b(\theta)) \cdot p(\theta|\tilde{\theta}). \tag{8.4}$$

where θ is the nuisance parameter and $\tilde{\theta}$ is the nominal value of θ. The prior *pdf* of θ is denoted by $p(\theta|\tilde{\theta})$. The test statistic (Eq. 8.3) will then take the form,

$$\tilde{q}_\mu = -2\ln\frac{\mathcal{L}(data|s(\mu)+b,\hat{\theta}_\mu)}{\mathcal{L}(data|s(\hat{\mu})+b,\hat{\theta})}, \qquad \text{with} \quad 0 \le \hat{\mu} \le \mu \tag{8.5}$$

where both numerator and denominator likelihoods are maximized. This allows for constraining nuisance parameters, given μ and *data*. The parameter estimators $\hat{\mu}$ and $\hat{\theta}$ correspond to the global maximum of the likelihood.

Observed and expected limits

After constructing the *pdfs*, $f(\tilde{q_\mu}|s(\mu)+b,\hat{\theta}_\mu^{obs})$ and $f(\tilde{q_\mu}|b,\hat{\theta}_0^{obs})$, for *signal+background* and *background-only* hypotheses, respectively, following the procedure mentioned above, the *observed* lower limit on the parameter of interest, μ, can be evaluated by calculating the probabilities p_μ and p_b as follows:

$$p_\mu = \int_0^{\tilde{q}_\mu^{obs}} f(\tilde{q_\mu}|s(\mu)+b,\hat{\theta}_\mu^{obs})\, d\tilde{q}_\mu, \tag{8.6}$$

$$1-p_b = \int_0^{\tilde{q}_0^{obs}} f(\tilde{q_\mu}|b,\hat{\theta}_0^{obs})\, d\tilde{q}_\mu \tag{8.7}$$

where, $\tilde{q}_\mu^{obs}$ is the observed value of the test statistic using the actual experimental data, for a given signal parameter μ. $\hat{\theta}_\mu^{obs}$ and $\hat{\theta}_0^{obs}$ are values of nuisance parameters, maximizing the likelihood using the experimentally observed data, for *signal + background* and *background-only* hypothesis, respectively. Once p_μ and $1-p_b$ are calculated, then the ratio of these two probabilities, $CL_s(\mu)$ is evaluated. To quote the 95% C.L. limit, μ is adjusted until $CL_s = 0.05$ is reached.

The expected limits are calculated independent of the actual data. The expected median lower limit and $\pm 1\sigma$ and $\pm 2\sigma$ bands for the *background-only* hypothesis are defined by generating a large set of *background-only* pseudo data, *i.e.,* the observed data are replaced with this pseudo data in the above procedure and a 95% C.L. lower limit on μ is evaluated for each of the pseudo data sets. One can then build a cumulative probability distribution of results by integrating from the side corresponding to low values. The point at which the cumulative probability distribution crosses the quantile of 50% is the median expected value. The $\pm 1\sigma$ (68%) band is defined by the crossings of the 16% and 84% quantiles. Crossings at 2.5% and 97.5% define the $\pm 2\sigma$ (95%) band.

8.2 Agreement of data with SM predictions

For the luminosity of 5.3 pb^{-1}, a total of 96 475 events in the data (see Chapter 5) and 92 525 predicted events for SM dimuon production are found corresponding to the 80–100 GeV/c^2 Z mass window. The 80–100 GeV/c^2 mass window is used for normalizing simulation to data. It is important to note here that, although the Z normalization is used in the consistency check of data with the SM production, it is not used in the determination of limits on Λ. Based on the above numbers, the normalization factor is calculated to be 1.043 $\pm$ 0.007. The predicted event yields (see Table 6.3) are then multiplied by the Z normalization factor and are compared with data for each of the $M > M_{\mu\mu}^{Low}$ values in Table 8.1.

Figures 8.1 and 8.2 show the observed dimuon mass spectrum and predictions from the SM DY production for the full mass range considered in this analysis. Also shown in the figure are the LLIM (CI/DY) predictions for $\Lambda = 4$ and 5 TeV with both destructive and constructive possibilities. These Λ values approximately correspond to the best limits previous to this analysis (see Section 2.2.3). From Table 8.1 and Figures. 8.1 and 8.2, one can conclude (by eye) that the observed dimuon distribution is consistent with the

Table 8.1. Dimuon event statistics for the data and SM production corresponding to 5.3 fb^{-1}. The numbers shown for SM predictions are after normalization to the Z peak.

$M_{\mu\mu}^{Low}$ (GeV/c^2)	Data	Predicted DY+BKG (after Z norm.)
200	4320±65.7	4322.4±45.6
300	1013±31.8	1099.5±9.3
400	338±18.4	367.2±5.1
500	141±11.9	155.5±2.8
600	57±7.6	69.9±0.58
700	28 ±5.3	34.6±0.36
800	14±3.7	18.1±0.25
900	13±3.6	9.8±0.15
1000	8 ±2.8	5.5±0.09
1100	3 ±1.7	3.3±0.08
1200	2 ±1.4	1.9±0.06
1300	1	1.2±0.05
1400	0	0.74±0.05
1500	0	0.46±0.05

SM predictions within the statistical errors.

In order to quantify this agreement, the probability (or "p-value") for the background (SM) to fluctuate as large or larger than the observed data is determined. This quantification is based on the *background-only* (SM) hypothesis and requires defining a test statistic and constructing the corresponding sampling distribution ($pdfs$). For the *background-only* hypothesis, the test statistic (Eq. 8.5) will take the form:

$$q_0 = -2\ln\frac{\mathcal{L}(data|b,\hat{\theta}_0)}{\mathcal{L}(data|s(\hat{\mu})+b,\hat{\theta})}, \qquad with \quad \hat{\mu}\geq 0 \tag{8.8}$$

Following the same prescription given in the previous section, pseudo data are generated following Poisson probabilities under the *background-only* hypothesis to construct the sampling distribution $f(q_0|b,\hat{\theta}_0^{obs})$ of the test statistic q_0. From this distribution, the p-value corresponding to a given experimental observation q_0^{obs} can be evaluated as

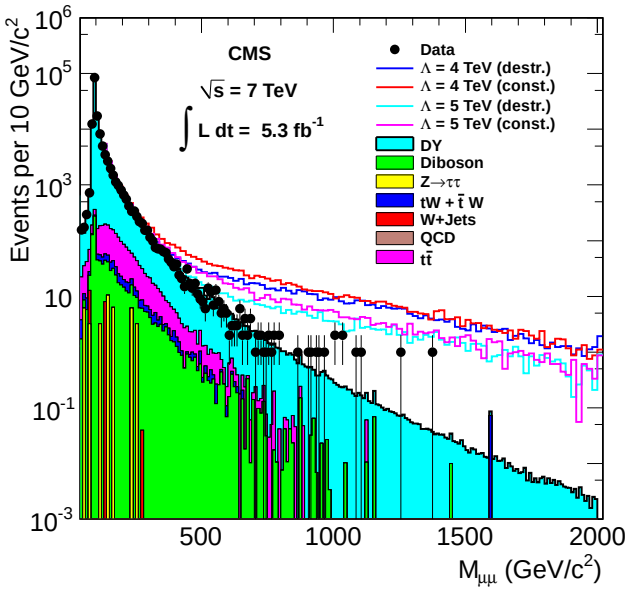


FIG. 8.1. Observed spectrum of $M_{\mu\mu}$ and predictions from the SM and LLIM (CI/DY) for $\Lambda = 4$ and 5 TeV, for constructive and destructive interference.

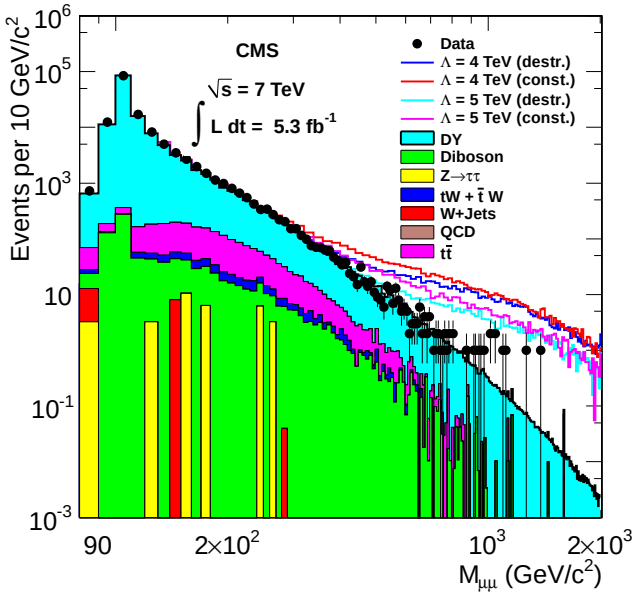


FIG. 8.2. Same as Fig. 8.1 except that the x-axis is also plotted on a logarithmic scale.

follows:

$$\text{p-value} = P(q_0 \geq q_0^{obs}) = \int_{q_0^{obs}}^{\infty} f(q_0|b, \hat{\theta}_0^{obs})\, dq_0. \tag{8.9}$$

Generally, good agreement between experimentally observed data and the SM-only hypothesis is found if the p-value is at least 5% or higher. Following the procedure mentioned above, for the CI analysis, p-values are calculated based on the observed data, SM background predictions and a set of nuisance parameters, as listed in Table 7.2. In the analysis performed with the 2011 CMS data, a p-value of 39% was found corresponding to $M_{\mu\mu}^{Low} = 700$ GeV/c^2, thus showing consistency with the SM. The reason to quote the p-value corresponding to $M_{\mu\mu}^{Low} = 700$ GeV/c^2 is because the sensitivity to Λ is found to be maximal for this value of $M_{\mu\mu}^{Low}$ (see Section 8.3). Since no excess over the SM predictions is observed in this analysis using the full 2011 CMS dataset, the next logical step would be to set exclusion lower limits on the compositeness energy scale Λ which is the subject of the next section.

8.3 Exclusion lower limits on Λ

In the CI analysis, the expected and observed 95% C.L. lower limits on Λ are determined using the CL_s modified frequentist method as described in Section 8.1. The software routine used to evaluate the limits is a part of the package with the standard procedures for statistical inference in CMS physics analyses. The routine estimates observed limits on the process cross section in a counting experiment, and the corresponding median expected limit with 1- and 2-σ quantile bands. The macro is written in the RooStats[4] framework and is compiled in ROOT version 5.32.00 or higher.

The principal arguments to the macro are: (1) The integrated luminosity (in pb^{-1}), (2) The absolute error on the integrated luminosity, (3) A nominal value of acceptance, (4)

[4]RooStats is a standard package within ROOT and is the statistical framework recommended by the CMS collaboration statistics committee

The absolute error on acceptance, (5) The number of observed events, (6) The estimated value for the background, and (7) The absolute error on the background.

The integrated luminosity for this analysis is 5277 pb^{-1} and the absolute error is taken to be 116 pb^{-1}, corresponding to a 2.2% uncertainty in the integrated luminosity (see Chapter 7). The acceptance value (the $A \times M$ factor) is set to 1 as the expected yields already include the acceptance correction. The absolute errors on acceptance are taken to be 3%, as described in Chapter 7. The expected mean for the number of signal events is the number of CI/DY events expected using a given Λ less the number of DY events. The expected mean for the number of background events is the total number of events from the DY process and non-DY SM backgrounds. The observed number of events come from the 2011 CMS data. The observed and expected number of events are given in Tables 6.4 and 6.5. The specific values of systematic uncertainties given as input to the calculation are summarized in Table 7.2.

The systematic uncertainties on integrated luminosity, acceptance, and expected background are treated as nuisance parameters. Other options that are chosen within the calculation are: profile likelihood ratio as a test statistic, following the LHC–style recommendation, Poisson statistics for observed number of events to construct likelihoods as it is a natural choice when there are zero events present in data, and lognormal [91, 92] prior *pdf s* for nuisance parameters.

Based on the previously described inputs and options, cross sections are returned by the calculation for each $M_{\mu\mu}^{Low}$ corresponding to a 95% C.L. fluctuation in the signal level. The cross section values are then converted to event counts by multiplying by the integrated luminosity and the resulting signal events are added to the SM background events (from the DY row of Table 6.4 or 6.5) to give a final signal plus background prediction for each $M_{\mu\mu}^{Low}$. These events are then matched to the expected signal+background events (for a given $M_{\mu\mu}^{Low}$) in Tables 6.4 and 6.5 to find the corresponding Λ value for destructive

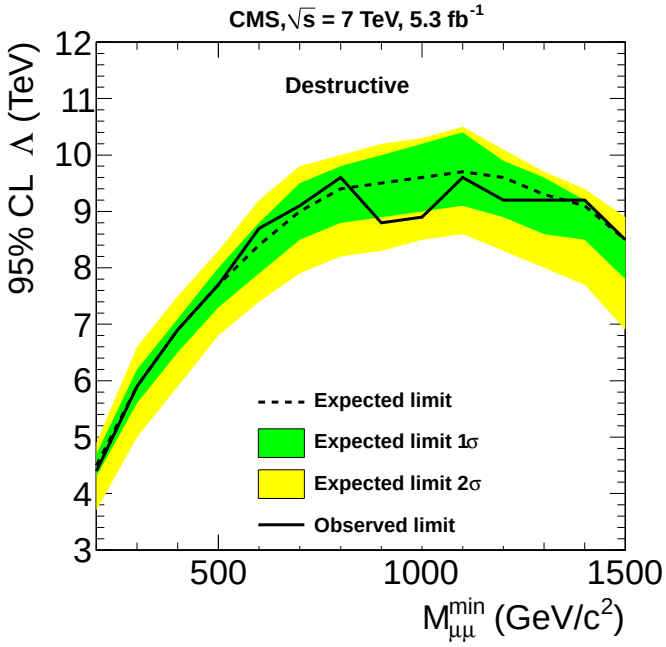


FIG. 8.3. Observed and expected limits as a function of $M_{\mu\mu}^{Low}$ for destructive interference.

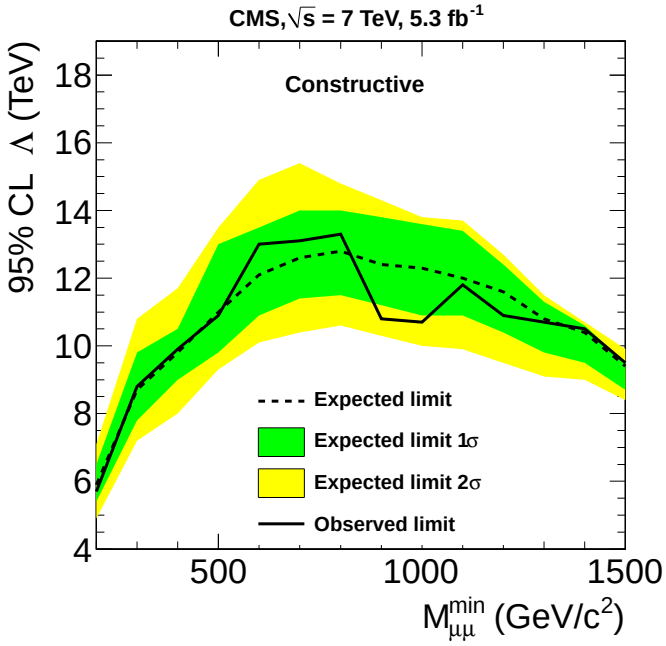


FIG. 8.4. Same as Fig. 8.3, for constructive interference.

and constructive interference, respectively. This value gives the 95% C.L. lower limit on Λ. Linear interpolation is used for signal plus background estimates that fall between two Λ values.

The observed and expected lower limits on Λ at 95% CL as a function of $M_{\mu\mu}^{Low}$ for destructive and constructive interference are shown in Figures. 8.3 and 8.4, respectively. The 1- and 2-σ uncertainties in the expected limits are indicated by the shaded bands. For both types of interference, the sensitivity to Λ is maximal for $M_{\mu\mu}^{Low}$ in the middle of the range studied. In both cases, a minimum mass of 700 GeV/c^2 is selected to quote the exclusion limits on the compositeness scale Λ. This results in an observed (expected) limit of 9.1 TeV (9.0 TeV) for destructive interference and 13.1 TeV (12.6 TeV) for constructive interference [96].

8.3.1 Effect of systematics

The effects of individual systematic uncertainties (given in Table 7.2) are studied by finding the change in limit for an explicit change in each uncertainty, or central value to which it corresponds. The PDF uncertainty of approximately 12% has the largest influence on the limits. For example, if the PDF uncertainty is set to zero, the constructive limit is about 6% higher. All other sources of systematic uncertainties listed in Table 7.2 have a negligible effect on the Λ limits, as shown below.

By way of contrast, although the non-DY background has a large relative uncertainty of 15%, the effect on the limits is minimal given the small contribution of non-DY sources relative to DY production (see Table 6.3). Since the QED K factor may or may not apply to the new physics associated with CI, the effect of removing this correction is studied, which results in an increase in the constructive limit of about 4%. Including the QED K factor for CI gives the more conservative limit.

To further justify the choice of restricting the CI analysis to a 2 TeV upper limit in

dimuon mass (see Section 6.1), the effect of including the CI signal events with dimuon masses above 2 TeV is studied. This results in an increase of the constructive limit by 0.7% which is a very small effect.

Chapter 9

Results and Conclusion

The CMS detector was used to measure the invariant mass distribution of $\mu^+\mu^-$ pairs produced in pp collisions at a center-of-mass energy of 7 TeV, based on an integrated luminosity of 5.3 fb^{-1}. The dimuon invariant mass distribution in the range 200 to 2000 GeV/c^2 was found to be consistent with Drell-Yan and other standard model sources of dimuons. The data were analyzed in the context of the interference of amplitudes from standard model Drell-Yan production and a left-left isoscalar contact interaction model of quark and muon compositeness, with energy scale parameter Λ. Lower limits were set on Λ at the 95% C.L. of 9.1 TeV for destructive interference and 13.1 TeV for constructive interference [96]. These limits represent significant improvements on the current published values of 4.5 TeV and 4.9 TeV.

As mentioned previously, starting this year LHC has been running at an increased center-of-mass energy of 8 TeV and is expected to accumulate around 20 fb^{-1} of data before the scheduled long shut-down in Winter 2013. Initially, inclusion of more data might result in a rapid increase of the expected limit on Λ, but, eventually it is expected to reach a plateau, limited by the attainable center-of-mass energy at the LHC. In that case, one can gain more sensitivity to new physics by analyzing the angular distribution of the dimuon system and including the other compositeness models described in Chapter

4. All of these possibilities will be explored at CMS in the near future. Also, the LHC energy will increase to its design energy, $\sqrt{s} = 14$ TeV in 2015 which will make the physics searches even more interesting. It is a very exciting time for particle physics and the search for new physics at the LHC has just begun !

BIBLIOGRAPHY

[1] http://en.wikipedia.org/wiki/Standard_Model

[2] http://en.wikipedia.org/wiki/Beta_decay

[3] Kenneth G. Wilson, "Confinement of Quarks", *Phys. Rev. D* **10** (1974) 2445–2459.

[4] http://en.wikipedia.org/wiki/Quark

[5] http://www.physics.ox.ac.uk/documents/PUS/dis/hera.htm

[6] G. Altarelli, and G. Parisi, "Asymptotic freedom in parton language", *Nucl. Phys. B* **126** (1977) 298.

[7] E. Fermi, "Versuch einer theorie der β-strahlen / Towards the theory of β-rays", *Z. Phys.* **88** (1934) 161.

[8] S. L. Glashow, "Partial symmetries of weak interactions", *Nucl. Phys.* **22** (1961) 579–588.

[9] S. Weinberg, "A model of leptons", *Phys. Rev. Lett.* **19** (1967) 1264–1266.

[10] A. Salam, "Weak and electromagnetic interactions. Elementary Particle Theory", Proceedings Of The Nobel Symposium Held At Lerum, Sweden (1968) 367–377.

[11] F. J. Hasert *et al.,* "Search for elastic muon-neutrino electron scattering", *Phys. Lett. B* **46** (1973), 121–124.

[12] F. J. Hasert *et al.,* "Observation of neutrino-like interactions without muon or electron in the gargamelle neutrino experiment", *Phys. Lett. B* **46** (1973), 138–140.

[13] E. N. Thompson *et al.,* "Search for Contact Interactions in the Dimuon Final State at ATLAS", Proceedings of the DPF-2009 Conference, Detroit, MI, July 27–31, 2009. arXiv: 0910.3384v2.

[14] J. C. Pati, Abdus Salam, and J. Strathdee, "Are Quarks Composite", *Phys. Lett. B* **59** (1975) 265.

[15] J. C. Pati, Abdus Salam, and J. Strathdee, "Are Quarks Composite", *Int. Centre Theor. Phys.* **IC/75/139**, addendum (1975).

[16] E. Eitchen, I. Hinchliffe, K. Lane *et al.,* "Supercollider Physics", *Rev. Modern Phys.* **56** (1984) 579–707.

[17] S. Drell and T. Yan, "Massive Lepton-Pair Production in Hadron-Hadron Collisions at High energy", *Phys. Rev. Lett.* **25** (1970) 316–320.

[18] E. Eitchen, K. Lane, and M. Peskin, "New Tests for Quark and Lepton Substructure", *Phys. Rev. Lett.* **50**, (1983) 811–814.

[19] K. Nakumara *et al.,* "Particle Data Group", *J. Phys. G37* (2010).

[20] G. Arnison *et al.,* "Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ GeV", *Phys. Lett. B* **122** (1983) 103–116.

[21] M. Banner *et al.,* "Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN anti-p p collider", *Phys. Lett. B* **122** (1983) 476–485.

[22] G. Arnison *et al.,* "Experimental observation of lepton pairs of invariant mass around 95 GeV/c^2 at the CERN SPS collider", *Phys. Lett. B* **126** (1983) 398–410.

[23] P. Bagnaia *et al.,* "Evidence for $Z^o \to e^+e^-$ at the CERN anti-p p collider", *Phys. Lett. B* **129** (1983) 130–140.

[24] ALEPH Collaboration, "Measurement of the High-Mass Drell-Yan Cross Section and Limits on Quark-Electron Compositeness Scales", *Eur. Phys. J. C* **49** (2007) 411–437.

[25] DELPHI Collaboration, "Measurement and interpretation of fermion-pair production at LEP energies above the Z resonance", *Eur. Phys. J. C* **45** (2004) 589–632.

[26] L3 Collaboration, "Search for manifestations of new physics in fermion-pair production at LEP", *Phys. Lett. B* **489** (2000) 81.

[27] OPAL Collaboration, "Production of fermion-pair events in e^+e^- collisions at 161 GeV centre-of-mass energy", *Phys. Lett. B* **391** (1997) 221–234.

[28] OPAL Collaboration, "Tests of the standard model and constraints on new physics from measurements of fermion-pair production at 189-209 GeV at LEP", *Eur. Phys. J. C* **33** (2004) 173–212.

[29] H1 Collaboration, "Search for new physics in eq interactions at HERA", *Phys. Lett. B* **568** (2003) 35–47.

[30] ZEUS Collaboration, "Search for contact interactions, large extra dimensions and finite quark radius in ep collisions at HERA, *Phys. Lett. B* **591** (2004) 23–41.

[31] F. Abe *et al.*, (CDF Collaboration), "Search for New Gauge Bosons in $\bar{p}p$ Collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. Lett.* **68** (1992) 1463–1467.

[32] F. Abe *et al.*, (CDF Collaboration), "Limits on Quark-Lepton Compositeness Scales from Dileptons Produced in 1.8 TeV $p\bar{p}$ Collisions", *Phys. Rev. Lett.* **79** (1997) 2198–2203.

[33] CDF Collaboration, "Search for Quark-Lepton Compositeness and a Heavy W Boson Using the en Channel in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV, *Phys. Rev. Lett.* **82** (2001) 231803.

[34] CDF Collaboration, "Search for $Z' \to e^+e^-$ Using Dielectron Mass Angular Distribution", *Phys. Rev. Lett.* **96** (2006) 211801.

[35] D0 Collaboration, "Measurement of the High-Mass Drell-Yan Cross Section and Limits on Quark-Electron Compositeness Scales", *Phys. Rev. Lett.* **82** (1999) 4769–4774.

[36] D0 Collaboration, "Measurement of Dijet Angular Distribution at $\sqrt{s} = 1.96$ TeV and Searches for Quark Compositeness and Extra Spatial Dimensions", *Phys. Rev. Lett.* **103** (2009) 191803.

[37] ATLAS Collaboration, "Search for quark contact interactions in dijet angular distributions in pp collisions at $\sqrt{s} = 7$ TeV measured with the ATLAS detector", *Phys. Lett. B* **694** (2011) 327–345.

[38] ATLAS Collaboration, "Search for Contact Interactions in Dimuon Events from pp Collision at $\sqrt{s} = 7$ TeV with the ATLAS Detector, *Phys. Rev. D* **84** (2011) 011101.

[39] ATLAS Collaboration, "Search for Contact Interactions in Dimuon Events from pp Collision at $\sqrt{s} = 7$ TeV with the ATLAS Detector, *Phys. Rev. X* arXiv:1112.4462v1 (2011).

[40] CMS Collaboration, "Search for Quark Compositeness with the Dijet Centrality Ratio in pp Collisions at $\sqrt{s} = 7$ TeV, *Phys. Rev. Lett.* **105** (2010) 262001.

[41] E. D. Bloom, D. H. Coward, J. Drees, G. Miller, L. W. Mo, R. E. Taylor, M. Breidenback, H. W. Kendall *et al.*, "High-energy inelastic e-p scattering at 6° and 10°", *Phys. Rev. Lett.* **23**, 16 (1969) 930–939.

[42] R. P. Feynman, "Very-High Energy Collisions of Hadrons", *Phys. Rev. Lett.* **23** (1969) 1415–1417.

[43] R. P. Feynman, "The Behavior of Hadron Collisions at Extreme Energies", in High Energy Collisions: Proceedings of the Third Topical Conference on High Energy Collisions of Hadrons, Stony Brook, New York, edited by C. N. Yang *et al.,* Pages 237–249 (Gordon and Breach, New York, 1969).

[44] A. D. Martin, W. J. Sterling, R. S. Thorne, and Watt, G., "Parton distributions for the LHC", *Eur. Phys. J. C* **63** (2009) 189–285.

[45] http://www.imperialhep.blogspot.com/2011/08/strangeness-at-lhcb.html

[46] J. C. Collins, D. E. Soper, and G. Sterman, "Factorization of hard processes in QCD", *Adv. Ser. Direct. High Energy Phys.* **5** (1988) 1–91.

[47] T. Sjöstrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **05** (2006).

[48] T. Sjöstrand, and P. Skands, "Transverse-momentum-ordered showers and interleaved multiple interactions", *Euro. Phys. Journal C* **39** (2005) 129–154.

[49] S. Frixione and B. Webber, "Matching NLO QCD computations and parton shower simulations", *JHEP* **0206** (2002) 029.

[50] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti *et al.,* Herwig 6.5, *JHEP* **01** (2001) 0101.

[51] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. Nadolsky, W. K. Tung, "New Generation of Parton Distributions with Uncertainties from Global QCD Analysis", *JHEP* **0207** (2002) 012.

[52] http://cms.web.cern.ch/news/how-cms-detects-particles

[53] Lyndon Evans and Philip Bryant, "LHCMachine", *JINST* **3** (2008) S08001.

[54] The ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider", *JINST* **3** (2008) S08003.

[55] The CMS Collaboration, "The CMS Experiment at the CERN LHC", *JINST* **3** (2008) S08004.

[56] The LHCb Collaboration, "The LHCb Detector at the CERN LHC", *JINST* **3** (2008) S08005.

[57] The ALICE Collaboration, "The ALICE Experiment at the CERN LHC", *JINST* **3** (2008) S08002.

[58] The TOTEM Collaboration, "The TOTEM Experiment at the CERN LHC", *JINST* **3** (2008) S08007.

[59] The LHCf Collaboration, "The LHCf Experiment at the CERN LHC", *JINST* **3** (2008) S08006.

[60] CERN, http://public.web.cern.ch/public/en/research/AccelComplex-en.html

[61] A. A. Glazkov, N. R. Lobanov, I. S. Balikoev, V. T. Barchenko, and S. I. Zagranichny, "Duoplasmatron-type ion source with improved technical and operational performance for linear accelerator", *3rd European Particle Accelerator Conference*, Berlin, Germany, 24–28 Mar 1992, pp.993.

[62] The CMS Collaboration, "CMS Luminosity Public Results". [https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults].

[63] The CMS Collaboration, "The CMS tracker system project: technical design report", *CERN-LHCC-98-006*, http://cdsweb.cern.ch/record/368412.

[64] The CMS Collaboration, "The CMS tracker: addendum to the technical design report", *CERN-LHCC-2000-016*, http://cdsweb.cern.ch/record/490194.

[65] The CMS Collaboration, "The CMS Electromagnetic Calorimeter Project: Technical Design Report", Technical Design Report CMS Series, CERN, *CMS-TDR-004*, 1997. [https://cdsweb.cern.ch/record/349375].

[66] The CMS Collaboration, "The CMS Hadron Calorimeter Project: Technical Design Report", Technical Design Report CMS Series, CERN, *CMS-TDR-002*, 1997. [https://cdsweb.cern.ch/record/357153].

[67] The CMS Collaboration, "The CMS Muon Project: Technical Design Report", Technical Design Report CMS Series, CERN, *CMS-TDR-003*, 1997. [https://cdsweb.cern.ch/record/343814].

[68] The CMS Collaboration, "Performance of CMS Muon Reconstruction in Cosmic-Ray Events, *JINST* **5** (2010) T03022, arXiv: 0911.4994.

[69] The CMS Collaboration, "CMS Physics TDR: Volume 1, Detector Performance and Software", *CERN-LHCC-2006-001* (2006).

[70] The CMS Collaboration, "The CMS TriDAS Project: Technical Design Report, Volume 1 – The Trigger System", Technical Design Report CMS Series, CERN, *CMS-TDR-006-1*, 2000. [https://cdsweb.cern.ch/record/706847].

[71] The CMS Collaboration, "The CMS TriDAS Project: Technical Design Report, Volume 2 – Data Acquisition and High Level Trigger", Technical Design Report CMS Series, CERN, *CMS-TDR-006-2*, 2002. [https://cdsweb.cern.ch/record/578006].

[72] The CMS Collaboration, "CMS Tracking Performance Results from early LHC Operation", *Eur. Phys. J. C* **70** (2010) 1165–1192.

[73] The CMS Collaboration, "Performance of CMS muon reconstruction in *pp* collision events at $\sqrt{s} = 7$ TeV", *CMS-MUO-10-004* (2012), arXiv: 1206.4071v1.

[74] CMS muon public results page: https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsMUO#Muon_trigger_efficiency_in_2011

[75] The CMS Collaboration, "Search for narrow resonances in dilepton mass spectra in *pp* collisions at $\sqrt{s} = 7$ TeV", CERN-PH-EP/2012-157 (2012), arXiv: 1206.1849v1.

[76] The CMS Collaboration, "Absolute Calibration of the Luminosity Measurements at CMS: Winter 2012 Update", CMS Physics Analysis Summary (2012) *CMS-PAS-SMP-12-008*.

[77] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **11** (2004) 040.

[78] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method", *JHEP* **11** (2007) 070.

[79] S. Alioli *et al.,* "NLO vector-boson production matched with shower in POWHEG", *JHEP* **07** (2008) 060.

[80] F. Maltoni and T. Stelzer, "MadEvent: Automatic event generation with MadGraph", *JHEP* **02** (2003) 027.

[81] The Geant4 Collaboration, "Geant4–a simulation toolkit", *Nucl. Inst. and Meth. A* **506** (2003) 250–303.

[82] C. Carloni Calame *et al.,* "Precision electroweak calculation of the production of a high transverse-momentum lepton pair at hadron colliders", *JHEP* **10** (2007) 109.

[83] G. Balossini, G. Montagna, C. Carloni *et al.,* "Electroweak QCD corrections to Drell Yan processes, *Acta Phys. Polon. B* **39** (2008) 1675, arXiv: 0805.1129v1.

[84] Jerzy Neyman, Egon Pearson, "On the Problem of the Most Efficient Tests of Statistical Hypotheses", *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* **231** (1933) (694706): 289337.

[85] Wilks, S. S., "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses", *The Annals of Mathematical Statistics* **9** (1938) 60–62.

[86] A. L. Read, "Presentation of search results: the CL_s technique", *J. Phys. G: Nucl. Part. Phys.* **28** (2002).

[87] A. L. Read, "Modified frequentist analysis of search results (the CLs method)", in *Proceedings of the First Workshop on Confidence Limits*, CERN, Geneva, Switzerland, 2000.

[88] Thomas Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth. A* **434** (1999) 435–443.

[89] Tom Junk, "Sensitivity, Exclusion and Discovery with Small Signals, Large Backgrounds, and Large Systematic Uncertainties", CDF/DOC/STATISTICS/PUBLIC/8128, October 2007.

[90] Falko Dulat, Bernhard Mistlberger, "Limit setting procedures and theoretical uncertainties in Higgs boson searches", arXiv: 1204.3851v2 (2012).

[91] W. T. Eadie, D. Drijard, F. E. James, M. Roos, B. Sadoulet, "Statistical Methods in Experimental Physics", (North Holland, Amsterdam, 1971), pp. 76, 79–80.

[92] Frederick James, "Statistical Methods in Experimental Physics", 2nd Edition, (World Scientific, Singapore, 2006), pp. 83–87.

[93] Z. Sullivan, "Fully differential W' production and decay at next-to-leading order in QCD", *Phys. Rev. D* **66** (2002) 075011.

[94] M. Botjel *et al.,* "The PDF4LHC Working Group Interim Recommendations", arXiv: 1101.05-38-v1-hep-ph (2011).

[95] P. M. Nadolsky *et al.,* "Implications of CTEQ global analysis for collider observables", *Phys. Rev. D* **78** (2008) 013004.

[96] The CMS Collaboration, "Search for Evidence of Contact Interactions in Dimuon Mass Spectrum", CMS Physics Analysis Summary (2012) *CMS-PAS-EXO-11-009*.

ABSTRACT

SEARCH FOR CONTACT INTERACTIONS IN THE DIMUON CHANNEL IN P-P COLLISIONS AT $\sqrt{s} = 7$ TeV AT CMS

by

SOWJANYA GOLLAPINNI

August 2012

Advisor: Dr. Paul E. Karchin

Major: Physics

Degree: Doctor of Philosophy

The standard model fails to explain the variety of observed quark and lepton flavors and their masses suggesting that there might exist a more fundamental basis. If quarks and leptons are composite particles made up of more basic constituents, a new physics interaction in the form of a four-fermion contact interaction arises between them. Experimentally the signal is manifest as a deviation from the standard model prediction in the high-mass tail for the invariant mass distribution of the opposite-sign dimuon pairs. The Large Hadron Collider accelerator at the Center for European Nuclear Research is built to explore new physics possibilities from proton-proton collisions occurring at the world's highest center-of-mass energy. This thesis discusses in detail a search strategy for a new physics possibility based on a left-handed current model of contact interactions. Based on 5.3 fb^{-1} of 2011 data as collected by the Compact Muon Solenoid detector, exclusion lower limits at 95% confidence level are set on the compositeness energy scale Λ, for both destructive and constructive interferences of the new physics with the standard model Drell-Yan process. These limits form the most stringent limits to date and exceed the current published limits significantly.

AUTOBIOGRAPHICAL STATEMENT

SOWJANYA GOLLAPINNI

Starting from my college years, I have always chosen physics as my major. The origin of the universe always intrigued me and eventually led me to commit to a career in experimental high-energy physics. I received my Bachelors degree in science in May 2003 from Sri Venkateswara University, India, majoring in physics, mathematics and computer science. Following my interest, I pursued my Masters education in physics with a specialization in particle physics. I received my Masters degree in June 2005 from University of Hyderabad, India. Extremely impressed by the high energy research studies performed at the department of physics, Wayne State University, I started my career as a Ph.D student at Wayne State in Fall 2007 and joined Prof. Paul E. Karchin in Summer 2008 to start my research in the field of particle physics as a participant of the Compact Muon Solenoid (CMS) experiment at the European Center for Nuclear Research (CERN), Switzerland. As part of my Ph.D program, I received my M.S. degree in physics from Wayne State University in September 2009.

I have been stationed in Batavia, Illinois since Summer 2009 to conduct my research at Fermilab where a much wider research community is available with their expertise at close reach. My Ph.D dissertation is based on a search conducted by me for contact interactions in the dimuon channel using 2011 CMS data. The experimental data is interpreted in the context of a compositeness model that assigns sub-structure to quarks and leptons, the elementary particles known to date. My thesis research yielded important results and are well appreciated by the particle physics community. I had an opportunity to present my thesis results at the April meeting of the American Physical Society (APS) in Atlanta, Georgia in April 2012 and at the Phenomenology 2012 Symposium (PHENO 2012) in Pittsburgh, Pennsylvania in May 2012.

During the course of my Ph.D, I have acquired efficient computing skills and improved my understanding of physics processes in general. For my impressive performance in research, I received the Universities Research Association, Inc. (URA) visiting scholar award at Fermilab in Summer 2011. I am also a recipient of a Summer 2012 dissertation fellowship awarded by Wayne State University. As a graduate student at Wayne State University, I also worked as a teaching assistant for undergraduate laboratories from Fall 2007 to Winter 2009. I always enjoyed these interactive sessions and thought of them as a rare opportunity to share my enthusiasm for physics.

I believe this is a unique time to study particle physics with the Large Hadron Collider at CERN setting new energy frontiers to be explored and providing exciting future prospects for researching new physics possibilities. Hence I would like to pursue my future career in the field of experimental particle physics exploring the big world of little particles!