



Instituto de Física Teórica  
Universidade Estadual Paulista

---

---

DOCTORAL THESIS

IFT-T.002/17

# **Search for New Resonances in the Merged Jet plus Dilepton Final State in CMS**

*José Cupertino Ruiz Vargas*

**Advisor**

Prof. Sérgio F. Novaes

**Co-advisor**

Dr. Thiago R. F. P. Tomei

May 2017

TESE DE DOUTORAMENTO

IFT-T.002/17

**Busca de Novas Ressonâncias Decaindo em  
Jato Hadronico e Dois Leptons no Experimento CMS**

*José Cupertino Ruiz Vargas*

**Orientador**

Prof. Sérgio F. Novaes

**Co-Orientador**

Dr. Thiago R. F. P. Tomei

Maio de 2017

# Acknowledgements

I would like to thank my supervisor Prof. Sérgio Novaes, for his dedication, patience, and for introducing me to the CMS Collaboration.

I would like to thank my co-advisor, Dr. Thiago Tomei, for the guidance, expertise, and bright ideas.

To my family: father, mother and brothers, for their encouragement and especially because they always trusted me.

I would like to thank all SPRACE team: Angelo, Cesar, Chang-Seong, David, Eduardo, Pedro, Sandra, Sudha, and Sunil, for their continuous support.

Thanks to my friends at IFT: Carlisson, José David, Fernando, and Prieslei, for their camaraderie.

This project was supported by the grant n° 2012/24593-8, São Paulo Research Foundation (FAPESP). The one year intership at CERN was supported by the European Particle Physics Latin American Network (EPLANET).

# Resumo

Na Organização Europeia para a Pesquisa Nuclear (CERN), o Large Hadron Collider (LHC) colide grupos de prótons 40 milhões de vezes por segundo a uma energia de 13 TeV. Operando junto ao LHC, o Compact Muon Solenoid (CMS) é um detector projetado para identificar uma ampla gama de partículas produzidas nessas colisões. As partículas produzidas em cada colisão são observadas nos subdetectores na busca de pistas sobre a Natureza no seu nível mais fundamental.

Apesar do modelo padrão das partículas elementares ter sido testado em uma variedade de experimentos de altas energias, um dos principais objetivos do LHC é a busca de uma nova física além daquela prevista pela teoria existente.

Nesse trabalho analisamos os dados de colisões próton-próton produzidos pelo LHC operando com energia de centro de massa de 13 TeV e coletados pelo CMS em 2015. O presente estudo envolve a busca de uma ressonância  $X$  não observada previamente, decaindo em um par de bósons vetoriais.

Os resultados são interpretados no contexto do modelo de dimensões extras deformadas de Randall-Sundrum, distinguindo as hipóteses de fundo (modelo padrão) e fundo mais sinal (modelo padrão + graviton). Nenhuma evidência da existência de uma partícula com as características do graviton de Randall-Sundrum foi encontrada.

Limites superiores com nível de confiança de 95% são estabelecidos para a seção de choque de produção da ressonância  $X$  decaindo em um par de bosons  $Z$ , para valores de massa  $M_X$  no intervalo 800 – 3000 GeV. A exclusão esperada (observada) no valor de massa  $M_X = 800$  GeV é 154 fb (93 fb); para o valor de massa  $M_X = 3000$  GeV o valor de exclusão esperado (observado) diminui para 19 fb (15 fb). Um excesso com significância global igual a  $2\sigma$  é observado no ponto de massa  $M_X = 1000$  GeV.

**Palavras Chaves:** Física de Altas Energias; Física de Partículas; Colisões Hadrônicas; Física Além do Modelo Padrão.

**Áreas do conhecimento:** Física de Altas Energias.

# Abstract

At the European Organization for Nuclear Research (CERN), the Large Hadron Collider (LHC) smashes groups of protons 40 million times per second at an energy of 13 TeV. Operating at the LHC, the Compact Muon Solenoid (CMS) is a multipurpose detector conceived to identify a large variety of particles produced in such collisions. The produced particles are observed at the sub-detectors in search of clues about Nature at the most fundamental level.

In spite of the impressive agreement of the standard model with all the experimental results obtained so far, one of the main aims of the LHC is the search of new physics beyond the one foreseen by this theoretical model.

In this work, we analyze proton–proton collisions delivered by the LHC operating at centre-of-mass energy of 13 TeV and collected by CMS during 2015. The channel under study involves the search for an unknown resonance  $X$  decaying into a pair of vector bosons.

The results are interpreted in the context of the Randall-Sundrum warped extra-dimensional model, distinguishing between the hypotheses of background only (standard model) and background plus signal (standard model + graviton). No evidence of the existence of a graviton-like particle was found.

Upper limits at 95% confidence level are set for the production cross section of a resonance  $X$  decaying to a pair of  $Z$  bosons, for values of mass  $M_X$  in the range between 800 – 3000 GeV. The limit expected (observed) at  $M_X = 800$  GeV is 154 fb (93 fb); for the mass point  $M_X = 3000$  GeV the limit expected (observed) decreases to 19 fb (15 fb). A localised excess with global significance equal to  $2\sigma$  is observed in the mass point  $M_X = 1000$  GeV.

**Key Words:** High Energy Physics; Particle Physics; Hadron Colliders, Physics Beyond the Standard Model.

**Research area:** High Energy Physics.

# Contents

<b>List of Tables</b>	<b>2</b>
<b>List of Figures</b>	<b>3</b>
<b>1 Introduction</b>	<b>6</b>
<b>2 The Standard Model and Beyond</b>	<b>10</b>
2.1 The Standard Model . . . . .	10
2.1.1 Strong Sector . . . . .	12
2.1.2 Electroweak Sector . . . . .	13
2.1.3 The Higgs Mechanism . . . . .	16
2.2 Beyond the Standard Model . . . . .	20
2.2.1 Extra Dimensions . . . . .	21
2.2.2 The Randall-Sundrum Model . . . . .	23
2.2.3 Bulk Graviton Model . . . . .	25
<b>3 The LHC and the CMS Experiment</b>	<b>29</b>
3.1 The Large Hadron Collider . . . . .	29
3.2 The CMS Experiment . . . . .	32
3.3 Main Detector Components . . . . .	34
3.3.1 Tracker System . . . . .	34
3.3.2 Electromagnetic Calorimeter . . . . .	34
3.3.3 Hadronic Calorimeter . . . . .	35
3.3.4 Solenoid . . . . .	35
3.3.5 Muon System . . . . .	36
3.3.6 Trigger and Data Acquisition . . . . .	36
3.4 Identification of Physical Objects . . . . .	38
3.4.1 Track Reconstruction . . . . .	38

3.4.2	Primary Vertex Reconstruction . . . . .	38
3.4.3	Jet Reconstruction . . . . .	40
3.4.4	Missing Energy . . . . .	42
3.5	Software and Computing . . . . .	43
3.5.1	Event Reconstruction . . . . .	44
3.5.2	Data Analysis with CMSSW . . . . .	45
3.5.3	Computing Infrastructure . . . . .	46
<b>4</b>	<b>Definition of the ZZ Semi-leptonic Channel</b>	<b>47</b>
4.1	General Overview . . . . .	47
4.2	Decay Chain and Final State . . . . .	48
4.3	Analysis Strategy . . . . .	50
4.4	Event Selection . . . . .	51
4.5	Simulated Samples . . . . .	55
4.6	Signal Characterization . . . . .	57
<b>5</b>	<b>Real Data from pp Collisions</b>	<b>63</b>
5.1	Primary Datasets . . . . .	63
5.2	Kinematic Distributions in the Control Region . . . . .	65
5.2.1	Leptonic Z Selection . . . . .	68
5.2.2	Hadronic Jet Selection . . . . .	71
<b>6</b>	<b>Background Estimation and Systematic Uncertainties</b>	<b>86</b>
6.1	Alpha Method . . . . .	86
6.2	Unblind of the Signal Region . . . . .	93
6.3	Systematic Uncertainties . . . . .	97
<b>7</b>	<b>Final Results and Limits</b>	<b>101</b>
7.1	Statistical Method . . . . .	101
7.2	Confidence Limits . . . . .	102
7.3	Interpretation of Results . . . . .	105
<b>8</b>	<b>Conclusions</b>	<b>107</b>
<b>A</b>	<b>Research Internship, Presentations and Service Work</b>	<b>109</b>
A.1	Internship at CERN . . . . .	109
A.2	Service Work to the Collaboration . . . . .	111

A.3 National and International Presentations . . . . .	112
<b>Bibliography</b>	<b>114</b>

# List of Tables

2.1	Fermion fields, representation and corresponding hypercharge.	14
4.1	Branching ratios	49
4.2	Summary of the event selection.	54
4.3	Description of the simulated samples.	56
4.4	Cut-flow table muon channel.	61
4.5	Cut-flow table electron channel.	62
5.1	Description of the real data samples.	64
5.2	Data/MC ratios	66
6.1	Data/MC ratios after unblind	93
6.2	Estimated background in the signal region for different models.	95
6.3	Background estimation	95
6.4	Systematic Uncertainties	98
A.1	Presentations at CMS internal meeting.	110
A.2	Trigger validation campaigns in 2016	112

# List of Figures

2.1	Elementary particles in the standard model . . . . .	11
2.2	Radiative corrections to the Higgs mass . . . . .	18
2.3	Relative contributions to $\Delta M_H^2$ . . . . .	19
2.4	B2G Summary . . . . .	22
2.5	Circle compactification . . . . .	23
2.6	Orbifold process . . . . .	24
2.7	Gravity and weak branes . . . . .	24
2.8	Bulk Graviton cross section . . . . .	27
2.9	Latest limits . . . . .	28
3.1	CERN’s Accelerator Complex . . . . .	29
3.2	Dipole superconducting magnet . . . . .	30
3.3	Performance of LHC over the Run 2 . . . . .	31
3.4	The CMS detector and its components . . . . .	33
3.5	CMS ECAL . . . . .	35
3.6	Pseudorapidity coverage of the muon system . . . . .	37
3.7	Characteristic signatures of different particles in CMS . . . . .	39
3.8	Tracks in a typical collision . . . . .	40
3.9	Collimated spray of particles . . . . .	41
3.10	Intuition of Tau21 . . . . .	42
3.11	Missing transverse momentum and its magnitude . . . . .	43
3.12	The Event Data Model . . . . .	45
4.1	Combine limit at Run 1 . . . . .	48
4.2	Production and decay channel . . . . .	49
4.3	Run 1 analysis . . . . .	53
4.4	Signal characterization . . . . .	58

4.5	Signal shapes . . . . .	59
4.6	Selection efficiency . . . . .	60
5.1	Trigger efficiency . . . . .	65
5.2	Distribution of number of primary vertices . . . . .	67
5.3	Electron identification efficiency . . . . .	69
5.4	Muon identification efficiency . . . . .	70
5.5	Distribution of $p_T$ for the leading lepton . . . . .	73
5.6	Distribution of $p_T$ for the second lepton . . . . .	74
5.7	Distribution of $\eta$ for the leading lepton . . . . .	75
5.8	Distribution of $\eta$ for the second lepton . . . . .	76
5.9	Distribution of $p_T$ of the leptonic Z candidate . . . . .	77
5.10	Distribution of mass of the leptonic Z candidate . . . . .	78
5.11	Distribution of $\Delta R$ between leptons . . . . .	79
5.12	Distribution of jet $p_T$ . . . . .	80
5.13	Distribution of jet $\eta$ . . . . .	81
5.14	Distribution of $\Delta R$ between the jet and the closest lepton . . . . .	82
5.15	Distribution of jet pruned mass . . . . .	83
5.16	Distribution of $\tau_2/\tau_1 = \tau_{21}$ . . . . .	84
5.17	Distribution of invariant mass . . . . .	85
6.1	Alpha transfer factor . . . . .	89
6.2	Invariant mass distributions in control region . . . . .	90
6.3	Jet mass fits in MC . . . . .	91
6.4	Jet mass fits in data . . . . .	92
6.5	Jet mass distribution . . . . .	94
6.6	Invariant mass distributions . . . . .	96
6.7	Signal characterization . . . . .	97
6.8	Uncertainties due to jet energy correction . . . . .	99
6.9	Uncertainties due to lepton selection . . . . .	100
7.1	Upper limits per category . . . . .	103
7.2	Upper limits combination . . . . .	104
7.3	Significance . . . . .	105
7.4	Nuisance parameters . . . . .	106

8.1 Luminosity plans of the HL-LHC . . . . .	108
A.1 Electron efficiency as function of deltaR . . . . .	109
A.2 Efficiency of the high energy electron trigger $E_T > 105$ GeV as function of pseudorapidity $\eta$ . The comparison of two consecutive pre-releases of the CMSSW is important to reveal either an expected change in the configuration of the trigger or the diagnostic of sys- tematic problems. . . . .	111

# Chapter 1

## Introduction

Understanding Nature from the study of its microscopic constituents is the goal of particle physics. The knowledge about elementary particles obtained through decades of experimental discoveries, accompanied by theoretical and technological developments, allowed to establish the standard model [1] as the main theory of particle physics. The model embeds two quantum field theories: the quantum chromodynamics [2] for strong interactions and the Glashow–Weinberg–Salam theory [3] of electroweak interactions.

The quantum electrodynamics [4] that describes the electromagnetic interaction with astonishing precision was considered a prototype gauge theory for the other interactions. After the discovery of the asymptotic freedom of the non-abelian gauge theories, a reliable theory for the strong interactions at short distance became possible. Based on the  $SU(3)_C$  color group, the quantum chromodynamics (QCD) was very successful describing the strong phenomena at high energies. On the other hand, the weak force has a significant influence only at distances of hundredths of the radius of a proton, and its short range implies that the virtual particles exchanged in weak interactions must be very massive. In fact, the weak bosons are around ninety times heavier than a proton.

In 1967 Abdus Salam and Steven Weinberg proposed a model of the weak interaction in which the gauge bosons acquire mass through the Higgs mechanism [5,6]. They explored the same gauge group proposed in 1961 by Sheldon Glashow [7], the  $SU(2)_L \times U(1)_Y$ . In order to preserve local symmetries, four gauge bosons are introduced, at the outset, all of them massless. The spontaneous symmetry breaking is implemented by a potential constructed out of a Higgs complex scalar doublet, represented by four scalar particles. Three of the Higgs fields are absorbed

by the gauge bosons  $W^+$ ,  $W^-$ , and  $Z$ . The fourth gauge boson remains massless, being associated with the photon, the mediator of the electromagnetic interaction. The three scalar particles that lend mass to the gauge bosons disappear from the physical spectrum, but one neutral scalar, the Higgs particle, remains in the physical spectrum and should be observed in experiments.

The first decisive test of the electroweak model was the experimental confirmation of the existence of weak interactions without exchange of electrical charge. Events mediated by neutral currents represented by the exchange of the  $Z$  boson, were observed in 1973 at CERN's Gargamelle neutrino experiment [8]. Without the  $Z$  contribution, any weak interaction would necessarily entail an exchange of electric charge by the  $W^\pm$  bosons. However, particles interacting through neutral weak currents keep their original identities, therefore, the  $Z$  interaction is flavor diagonal.

In the next decade, the four collaborations of the large Electron–Positron Collider (LEP) were able to deeply scrutinize several aspects of the standard model. They measured almost 20 parameters related to the standard model and arrived to an impressive agreement between experimental data and theoretical predictions.

Another key prediction of the standard model is the Higgs boson, whose mass is not predicted by the theory. After 45 years of intensive search for this particle, in 2012, both the CMS [9] and ATLAS [10] collaborations of the Large Hadron Collider (LHC) presented very strong evidences of the Higgs boson discovery. A boson with mass around 125 GeV sharing much of the characteristics of the Higgs showed up in the  $\gamma\gamma$  and four leptons final states.

Despite its great success describing all the experimental data at present, the standard model has some intrinsic problems. For instance, it does not predict the value of the masses of the elementary particles and does not furnish an explanation for the huge amplitude of the mass spectrum. The model contains a large number of parameters, does not explain the generation structure, and has difficulties associated with hierarchy and fine-tuning problems. It does not incorporate gravity and fails to unify the electroweak and strong forces. The standard model does not have any candidate that could explain the existence of the dark matter either.

There are alternative theories that could solve some of the standard model problems. Much of the current theoretical efforts involve new physics at the TeV scale, such as supersymmetry [11], compact extra dimensions [12], or composite

Higgs models [13] for spontaneous symmetry breaking.

Another example of beyond the standard model (BSM) theory is the Randall-Sundrum (RS) warped extra-dimensional model [14, 15], which offers an unification approach of all interactions including gravity. An extension of the original RS model, the bulk graviton [16], is particularly important from the experimental point of view because it predicts the existence of heavy resonances coupling to pairs of vector bosons (ZZ, WZ, and WW), generically called diboson (VV) resonances.

The experimental approach to the searches is independent of the details of the specific production or decay patterns, allowing the LHC to explore different hypotheses and set constraints to several BSM theories. The direct observation of TeV resonances is technically possible due to the excellent detection capability and general purpose of the CMS experiment. CMS was conceived to discover not only the Higgs boson, by measuring essential observables to identify leptons, photons, hadrons, and missing transverse energy, in a range of energy surpassing the TeV scale.

Given a massive boson  $V$  decaying through a hadronic channel,  $V \rightarrow qq$ , if the mother particle  $V$  has high enough boost the final state particles may often be reconstructed as a single hadronic jet. Separating jets originating from boson decays from those that come from standard QCD processes requires studying the inner structure of the jet in question. The jet mass can be used to discriminate vector bosons as long as the soft radiation, underlying event, and pileup contributions, are carefully removed from the radius of the jet. The procedure to remove the undesirable contamination is called *pruning* and consists on rerunning the jet algorithm and vetoing on these components, resulting in a pruned jet [17].

The extended capabilities of the CMS experiment allow us to perform BSM searches. In particular, we explore the production of a unknown resonance  $X$  that decays to a pair of vector bosons, one being a Z boson, with two leptons and one jet in the final state,  $X \rightarrow ZV \rightarrow \ell\ell + \text{jet}$ . The leptons can be electrons or muons, and their signature corresponds to the decay of a Z boson,  $Z \rightarrow \ell\ell$ . On the other hand, there is a pruned jet associated to the hadronic decay of the vector boson V (W or Z), characterized by a two-prong jet substructure produced by the initial quarks,  $V \rightarrow qq \rightarrow \text{jet}$ .

Using pruned jets and good identified leptons in the final state, we construct

the diboson invariant mass  $m_{ZV}$  to serve as experimental signature of the unknown resonance mass  $M_X$ . The analysis targets the search for bumps or excesses in the distribution of the diboson invariant mass, considering standard model processes  $Z+jets$ ,  $VV$ , and  $t\bar{t}$ , as background, and a bulk graviton as signal,  $X = G_{\text{bulk}}$ . The adoption of other signal models involving charged resonances ( $X = W'$ ) is still possible given the general treatment of the hadronic V boson reconstruction; however, our results are specialized to the neutral resonance scenario.

In Chapter 2, we present a synthetic review of the standard model, quantum chromodynamics, the electroweak theory, physics beyond the standard model, and the bulk graviton model. Chapter 3 gives an overview of the LHC and outlines the CMS experiment. Chapter 4 contains the definition of the ZZ semi-leptonic channel and results from previous searches. Control distributions using real data are presented in Chapter 5, followed by a detailed discussion about background estimation in Chapter 6. Results including the computation of confidence limits are presented in Chapter 7, and finally, the conclusions are given in Chapter 8.

# Chapter 2

## The Standard Model and Beyond

### 2.1 The Standard Model

Four basic forces mediate all known interactions among the particles of matter [18]. Electromagnetic and gravitational are infinite range forces, so they are familiar to everyone for their macroscopic effects. The two remaining forces, which are called the weak and the strong force, cannot be perceived directly because their influence extends no larger than the radius of an atomic nucleus. The strong force binds together the protons and neutrons in the nucleus, and glues the quarks to constitute hadrons. The weak force is mainly responsible for decay processes, characterized by a long lifetime and the large mass of the weak bosons that carry the interaction; the typical lifetime of particles decaying via strong interaction is  $\sim 10^{-23}$  seconds, while decays via weak interaction can take  $\sim 10^{-12}$  seconds or longer [19].

The standard model (SM) [1] of elementary particles describes the nature of the forces by means of non-abelian gauge theories [20]. Electromagnetic and weak forces are mediated by the gauge particles of the Glashow–Weinberg–Salam model [7], namely the massless photon and a triplet of massive vector bosons, the  $W^+$ ,  $W^-$ , and  $Z^0$ . The strong force is attributed to the eight massless gluons of quantum chromodynamics. In addition, there is one scalar Higgs boson, which is massive and electrically neutral. These quantum field theories incorporate quantum mechanics and special relativity to describe three of the four fundamental interactions: strong, electromagnetic, and weak interactions. The gravitational interaction is far weaker and is not expected to contribute significantly to any processes at subatomic level.

The fundamental particles of matter are six leptons and six flavors of quarks, each of the quarks being present in three colors (Fig. 2.1). There exist 12 fermions (6 quarks plus 6 leptons), 4 vector bosons ( $W$ ,  $Z$ , photon, gluon), and one scalar field (Higgs) responsible for the mechanism of electroweak symmetry breaking (EWSB). Each particle has its own anti-particle related by charge conjugation.

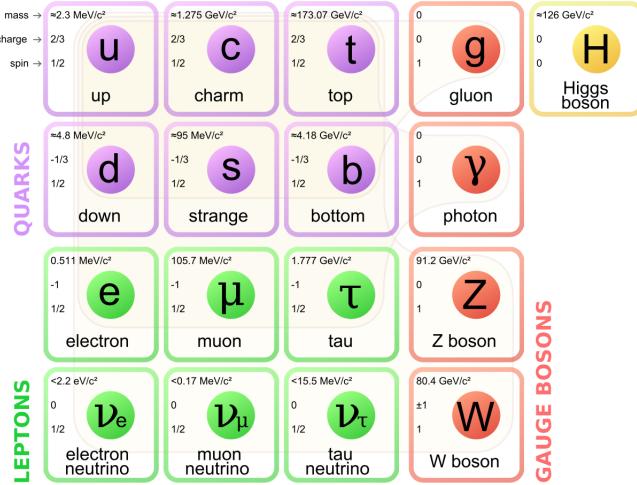


Figure 2.1: The standard model of elementary particles with three generations of quarks and leptons, gauge bosons, and the Higgs boson. [public domain]

Quarks are the constituents of composite particles called hadrons, which are bound states of quark–antiquark for the case of mesons (e.g.,  $\pi^+$ ,  $\pi^-$ ,  $\pi^0$ , etc.), or three–quark states for the case of baryons (e.g., protons, neutrons,  $\Delta^{++}$ , etc.). Single quark states cannot be observed as free particles, have fractional electric charge, and a quantum number called “color”, which is the source of the strong interaction. Gluons, the mediators of color flow in the strong interactions, carry eight combinations of color and anti-color.

Leptons, having the electron as the best known member, do not undergo strong interactions because they do not carry color. Besides the charged leptons, there are neutral leptons: the neutrinos. Neutrinos only interact via the weak force, consequently, their detection in collider experiments is usually inferred by energy imbalance in a specific reaction.

### Gauge Symmetry

Gauge invariance is the mechanism that determines the dynamical forces among the fundamental constituents of matter [21]. The fields associated with the

elementary particles are representations of a symmetry group; for the standard model, the governing symmetry group is

$$SU(3)_C \times SU(2)_L \times U(1)_Y. \quad (2.1)$$

The factor  $SU(3)_C$  corresponds to the strong sector, carries the color charge, and generates  $(3^2 - 1)$  gauge fields associated with the eight gluons. The factor  $SU(2)_L \times U(1)_Y$  corresponds to the electroweak sector, carries isospin and hypercharge, and generates  $(3 + 1)$  gauge fields associated with the weak bosons and the photon.

### 2.1.1 Strong Sector

Quantum chromodynamics (QCD) is the theory of the strong interactions between quarks and gluons, governed by the symmetry group  $SU(3)_C$ . The fundamental representation of  $SU(3)_C$  is a triplet, so the three quark colors red, green, and blue, or  $(r, g, b)$ , form the fundamental representation:

$$q = \begin{pmatrix} q^r \\ q^g \\ q^b \end{pmatrix}. \quad (2.2)$$

In this representation, the  $SU(3)_C$  generators are the Gell-Mann matrices denoted  $\lambda^a$ , with  $a = 1, \dots, 8$ . The QCD Lagrangian is given by

$$\mathcal{L}_{QCD} = \sum_{q=u,d,s,\dots} \bar{q}(i \not{D} - m_q)q - \frac{1}{4}G_{\mu\nu}^a G^{a\mu\nu} \quad (2.3)$$

where the sum runs over the quark flavors up ( $u$ ), down ( $d$ ), strange ( $s$ ), charm ( $c$ ), bottom ( $b$ ), and top ( $t$ ). The strength tensor of the gluon field  $G_\mu^a$  is given by

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_s f^{abc} G_\mu^b G_\nu^c \quad (2.4)$$

where the structure constant  $f^{abc}$  of the  $SU(3)_C$  group is defined through

$$[\lambda^a, \lambda^b] = 2if^{abc}\lambda^c. \quad (2.5)$$

The covariant derivative is defined as

$$D_\mu q = \left( \partial_\mu + ig_s \frac{\lambda^a}{2} G_\mu^a \right) q \quad (2.6)$$

where  $g_s$  is the coupling of the strong force.

The strong interactions have a characteristic energy scale  $\Lambda \sim 200$  MeV interpreted as the energy at which the coupling constant diverges. The running of the coupling constant obtained by the renormalization group equation at leading order of perturbation theory is given by

$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 2n_f) \log(Q^2/\Lambda^2)} \quad (2.7)$$

where  $n_f$  is the number of quark flavors and  $Q^2$  the probed energy. At very large  $Q^2$ , corresponding to small distances,  $\alpha_s = g_s/4\pi$  becomes increasingly small. This phenomena is known as asymptotic freedom, property that allows perturbative expansion at small distances.

### 2.1.2 Electroweak Sector

The 12 fundamental fermions are grouped in three generations:

$$\left\{ \begin{array}{cc} \nu_e & u \\ e & d \end{array} \right\}, \quad \left\{ \begin{array}{cc} \nu_\mu & c \\ \mu & s \end{array} \right\}, \quad \left\{ \begin{array}{cc} \nu_\tau & t \\ \tau & b \end{array} \right\}$$

The three generations differ only in the mass and the flavor quantum number, but are representations of the same symmetry group. Each generation is separated in two doublets and three singlets of  $SU(2)_L$ .

For instance, for the first generation we have

$$\left\{ \begin{array}{cc} \nu_e & u \\ e & d \end{array} \right\} \equiv \left( \begin{array}{c} \nu_e \\ e \end{array} \right)_L, \quad \left( \begin{array}{c} u \\ d \end{array} \right)_L, \quad e_R, \quad u_R, \quad d_R$$

since in the model the neutrinos are considered massless, there is no right-handed component for the neutrino.

The subscripts  $L$  and  $R$  stand for left and right chiral component. For a Dirac

spinor  $f$  its chiral decomposition is

$$f = f_L + f_R = P_L f + P_R f \quad (2.8)$$

where  $P_{L,R} = (1 \mp \gamma_5)/2$  are the left and right chiral projectors. Left and right fields belong to different representations of  $SU(2)$  and exhibit different values of  $U(1)_Y$  hypercharge, property summarized in Table 2.1.

Generation			$SU(2)_L$	$U(1)_Y$
I	II	III		
$\begin{pmatrix} u \\ d \end{pmatrix}_L$	$\begin{pmatrix} c \\ s \end{pmatrix}_L$	$\begin{pmatrix} t \\ b \end{pmatrix}_L$	doublet	$+\frac{1}{6}$
$u_R$	$c_R$	$t_R$	singlet	$+\frac{2}{3}$
$d_R$	$s_R$	$b_R$	singlet	$-\frac{1}{3}$
$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$	doublet	$-\frac{1}{2}$
$e_R$	$\mu_R$	$\tau_R$	singlet	-1

Table 2.1: Fermion fields, representation and corresponding hypercharge.

In the following we use the notation

$$\psi_1 = \begin{pmatrix} u \\ d \end{pmatrix}_L, \quad \psi_2 = u_R, \quad \psi_3 = d_R. \quad (2.9)$$

Here we consider only quarks of the first generation to derive the charge current, though the result can be generalized to include the other generations as well as the leptons. The Lagrangian for fermions can be written as

$$\mathcal{L}_{\text{fermion}} = \sum_{j=1}^3 i\bar{\psi}_j \gamma^\mu D_\mu \psi_j \quad (2.10)$$

with the covariant derivatives defined as

$$\begin{aligned} D_\mu \psi_1 &\equiv \left[ \partial_\mu + ig \frac{\sigma_i}{2} W_\mu^i + ig' y_1 B_\mu \right] \psi_1 \\ D_\mu \psi_2 &\equiv \left[ \partial_\mu + ig' y_2 B_\mu \right] \psi_2 \\ D_\mu \psi_3 &\equiv \left[ \partial_\mu + ig' y_3 B_\mu \right] \psi_3. \end{aligned}$$

$g$  and  $g'$  are the  $SU(2)_L$  and  $U(1)_Y$  gauge couplings,  $W_\mu^i$  and  $B_\mu$  are the respective gauge bosons,  $y_j$  are the hypercharges, and  $\sigma^i$  the Pauli matrices.

The charge current Lagrangian obtained from 2.10 corresponds to

$$\begin{aligned} -\mathcal{L}_{CC} &= g \bar{\psi}_1 \gamma^\mu \frac{\sigma^1}{2} W_\mu^1 \psi_1 + g \bar{\psi}_1 \gamma^\mu \frac{\sigma^2}{2} W_\mu^2 \psi_1 \\ &= \frac{g}{\sqrt{2}} \left[ \bar{u}_L \gamma^\mu d_L W_\mu^+ + \bar{d}_L \gamma^\mu u_L W_\mu^- \right] \end{aligned} \quad (2.11)$$

where

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (W_\mu^1 \mp i W_\mu^2).$$

The neutral current Lagrangian is given by

$$-\mathcal{L}_{NC} = g J_3^\mu W_\mu^3 + g' J_Y^\mu B_\mu = e J_{em}^\mu A_\mu + g_1 J_1^\mu Z_{1\mu}^0. \quad (2.12)$$

The currents  $J_3^\mu$  and  $J_Y^\mu$  are

$$J_3^\mu = \sum_f \bar{f} \gamma^\mu [t_{f_L}^3 P_L + t_{f_R}^3 P_R] f \quad (2.13)$$

$$J_Y^\mu = \sum_f \bar{f} \gamma^\mu [y_{f_L} P_L + y_{f_R} P_R] f \quad (2.14)$$

where  $t_{f_L}^3$  ( $t_{f_R}^3$ ) is the third component of weak isospin for the left (right) chiral component of fermion  $f$ ; for quarks of the first generation,

$$t_{u_L}^3 = +\frac{1}{2}, \quad t_{d_L}^3 = -\frac{1}{2}, \quad \text{and} \quad t_{u_R}^3 = t_{d_R}^3 = 0. \quad (2.15)$$

The weak hypercharges  $y_{f_{L,R}}$  are chosen to yield the correct electric charges,

$$t_{f_L}^3 + y_{f_L} = t_{f_R}^3 + y_{f_R} = q_f \quad (2.16)$$

where  $q_f$  is the electric charge of  $f$  in units of the positron charge.

The mass eigenstates in Eq. 2.12 are the massless photon  $A_\mu$  and the massive  $Z_{1\mu}^0 \equiv Z_\mu$ , where

$$A_\mu = \sin \theta_W W_\mu^3 + \cos \theta_W B_\mu \quad (2.17)$$

$$Z_\mu = \cos \theta_W W_\mu^3 - \sin \theta_W B_\mu \quad (2.18)$$

and the weak angle is  $\theta_W \equiv \tan^{-1}(g/g')$ . The new gauge couplings are

$$e \equiv g \sin \theta_W, \quad g_1^2 \equiv g^2 + g'^2 = \frac{g^2}{\cos \theta_W}. \quad (2.19)$$

The currents in the new basis are

$$J_{em}^\mu = \sum_f q_f \bar{f} \gamma^\mu f \quad (2.20)$$

$$J_1^\mu = \sum_f \bar{f} \gamma^\mu [\epsilon_L^1(f) P_L + \epsilon_R^1(f) P_R] f \quad (2.21)$$

with the chiral couplings

$$\epsilon_L^1(f) = t_{f_L}^3 - q_f \sin^2 \theta_W, \quad \epsilon_R^1(f) = t_{f_R}^3 - q_f \sin^2 \theta_W. \quad (2.22)$$

According to the charge-parity (CP) symmetry, the laws of physics should be the same if a particle is interchanged with its antiparticle (C symmetry) while its spatial coordinates are inverted (P symmetry). CP violation was first observed in 1964 [22], and became a necessary mechanism to explain  $K$  and  $B$  meson systems. The incorporation of CP violation effects in the standard model motivated the introduction of three fermion families by Kobayashi and Maskawa [23]. Further details on the CKM matrix and the mixing effects can be found elsewhere [24].

### 2.1.3 The Higgs Mechanism

The weak bosons of the SM acquire mass through a spontaneous symmetry breaking in which 3 of the 4 generators of the electroweak sector are broken

$$SU(2)_L \times U(1)_Y \rightarrow U(1)_{EM}.$$

To illustrate the mechanism, consider the gauge group  $G = SU(N)$ , with  $N - 1$  diagonal generators.  $G$  can be broken by the vacuum expectation value (VEV) of

a real adjoint Higgs representation  $\Phi$ , which can be represented by a Hermitian traceless  $N \times N$  matrix

$$\Phi = \sum_{i=1}^{N^2-1} \varphi^i T_i \quad (2.23)$$

where  $\varphi^i$  are the real components of  $\Phi$  and the  $T_i$  are the fundamental ( $N \times N$ ) representation matrices. When  $\Phi$  acquires a VEV,  $\langle \Phi \rangle$ ,  $G$  is broken to a subgroup associated with those generators that commute with  $\langle \Phi \rangle$ . The VEV  $\langle \Phi \rangle$  can be diagonalized by an  $SU(N)$  transformation, so that the  $N - 1$  diagonal generators remain unbroken.

The SM introduces a scalar Higgs doublet

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad (2.24)$$

through the scalar Lagrangian

$$\mathcal{L}_{\text{scalar}} = (D^\mu \phi)^\dagger D_\mu \phi - \mu^2 \phi^\dagger \phi - h(\phi^\dagger \phi)^2 \quad , \quad (\mu^2 < 0 < h) \quad (2.25)$$

with the covariant derivative given by

$$D_\mu \phi = \left( \partial_\mu + ig W_\mu^i \frac{\sigma^i}{2} + ig' y_\phi B_\mu \right) \phi. \quad (2.26)$$

The neutral component  $\phi^0$  has weak isospin and weak hypercharge

$$t_{\phi^0} = -t_{3\phi^0} = y_{\phi^0} = \frac{1}{2} \quad (2.27)$$

where  $t_3$  is the third component of weak isospin,  $y_i = q_i - t_{3i}$  the weak hypercharge. When the neutral scalar field  $\phi^0$  acquire a VEV, the photon  $A_\mu$  (Eq. 2.17) remain massless, while the  $Z_\mu$  field develops a mass term

$$M_{Z^0}^2 \equiv \frac{1}{2} g^2 |\langle \phi^0 \rangle|^2 = \frac{1}{4} g^2 v^2 = \frac{M_W^2}{\cos^2 \theta_W} \quad (2.28)$$

where

$$v^2 = 2 |\langle \phi^0 \rangle|^2 \sim (\sqrt{2} G_F)^{-1} \sim (246 \text{ GeV})^2 \quad (2.29)$$

is the square of the weak scale and  $G_F$  is the Fermi constant.

The Higgs field also enters in the Yukawa Lagrangian (ignoring family indices)

$$-\mathcal{L}_{Yuk} = h_d \bar{Q}_L \phi d_R + h_u \bar{Q}_L \tilde{\phi} u_R + h_e \bar{L}_L \phi e_R + h.c. \quad (2.30)$$

where

$$Q_L \equiv \begin{pmatrix} u_L \\ d_L \end{pmatrix} \quad \text{and} \quad L_L \equiv \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}. \quad (2.31)$$

The parameters  $h_d$ ,  $h_u$ , and  $h_e$  are the Yukawa constants which are directly related with the mass of the fermions. The tilde field is defined by

$$\tilde{\phi} \equiv i\sigma^2 \phi^* = \begin{pmatrix} \phi^{0*} \\ -\phi^- \end{pmatrix} \quad (2.32)$$

where  $\sigma^2$  is the second Pauli matrix. After the spontaneous symmetry breaking the mass of the fermions springs

$$-\mathcal{L}_{Yuk} = \frac{1}{\sqrt{2}}(\nu + H)(h_d \bar{d}d + h_u \bar{u}u + h_e \bar{e}e) \quad (2.33)$$

where  $H$  is the scalar field of the Higgs remaining. A single Higgs doublet suffices for the SM, but in many extensions including some  $U(1)'$  models [24], a second doublet may be introduced.

### Fine-tunning on the Higgs Mass

One important issue is associated with the radiative corrections to the Higgs boson mass represented by the loop diagrams depicted in Fig. 2.2.

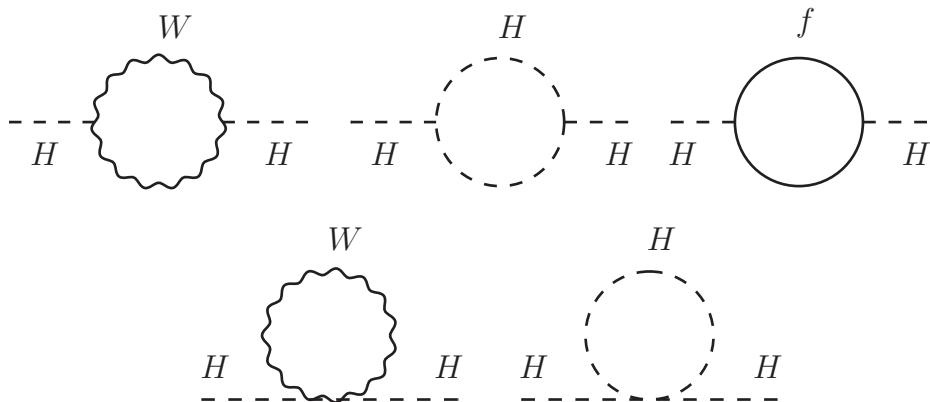


Figure 2.2: Radiative corrections to the Higgs mass.

The Higgs mass receive quantum corrections from loops that contain bosons ( $W, Z, \text{Higgs}$ ) and fermions (quarks and leptons), with the last one being dominated by the top quark since its contribution goes with the square of the fermion mass. Taking into account those contributions we can write the renormalized Higgs mass as, [25]

$$\underbrace{M_H^2}_{\text{physical}} = \underbrace{M_{0,H}^2}_{\text{bare}} + \Lambda^2 \underbrace{\left( 6M_W^2 + 3M_Z^2 + M_H^2 - 12M_{\text{top}}^2 \right)}_{\text{loop corrections}} \frac{G_F}{4\pi^2\sqrt{2}} \quad (2.34)$$

where  $\Lambda$  is the maximum energy for which the SM applies, or in other words, for energies larger than  $\Lambda$  a new theory should be taken into account.

In principle, this scale can be as large as the Planck scale , that is,

$$\Lambda \sim M_{\text{Planck}} = \left( \frac{\hbar c}{G_{\text{Newton}}} \right)^{1/2} \approx 1.2 \times 10^{19} \text{ GeV} .$$

According to Eq. 2.34, the quantum correction goes like  $\Lambda^2$  and it contains the sum of the effect of the bosons loops minus the sum of the effect of the fermions loop. It is important to notice that the quantum correction is not proportional to the Higgs mass itself. Therefore, the correction is present even for a Higgs with zero bare mass.

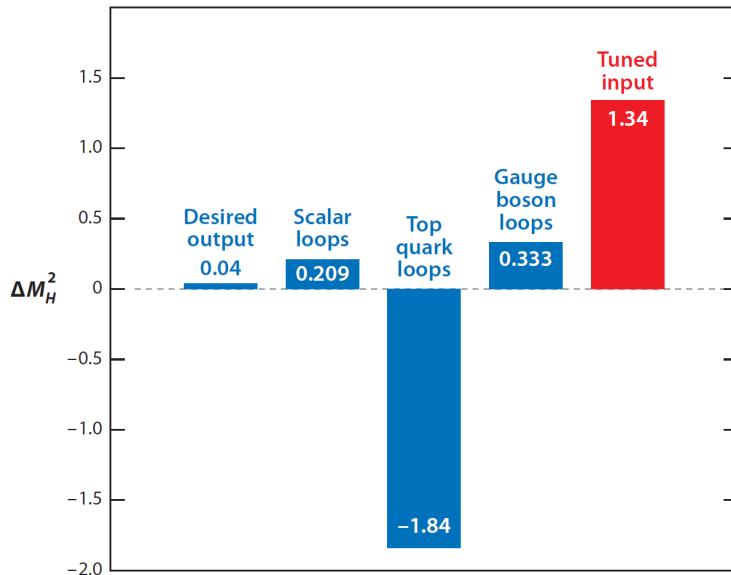


Figure 2.3: Relative contributions to  $\Delta M_H^2$  for a value of  $\Lambda = 5 \text{ TeV}$  [25].

The physical Higgs mass is equal to the bare mass plus a very large number ( $\Lambda^2$ ) multiplied by a (negative) loop factor. Therefore, in order to obtain a reasonable value for the physical Higgs mass (125 GeV), the theory must have a very fine-tuning cancellation between the radiative correction and the bare mass. Even when the scale is not extremely large ( $\Lambda \simeq 5$  TeV), a careful balancing is required to maintain a small Higgs mass (Fig 2.3).

Unless we suppose that the bare mass and the quantum corrections are finely tuned to yield  $M_H \sim 125$  GeV, some new physics must intervene. Such precise balancing is utterly unnatural in physics theories, leading the physicists to propose a series of ways in which this cancellation could occur naturally. For instance, supersymmetry [11] exploits the fact that fermion loops contribute with an overall minus sign relative to the boson loops (because of Fermi statistics), balancing the contributions of fermion and boson loops. In unbroken supersymmetry, the masses of bosons are degenerate with those of their fermion counterparts, so the cancellation is exact.

## 2.2 Beyond the Standard Model

The standard model is in very good agreement with all experimental data obtained so far. Some of the remarkable evidences include:

- Existence of quarks with spin 1/2 and gluons with spin 1, from deep-inelastic scattering experiments at SLAC (1970) and tree-jet events at DESY (1979);
- Discovery of  $J/\psi$  at BNL and SLAC (1974);
- Weak neutral current process mediated by Z bosons discovered at CERN (1973) and Fermilab (1974);
- Existence of the color quantum number from different measurements:  $R$ ,  $\pi^0$  decay into photons, anomaly cancelation, etc.;
- Discovery of  $W^\pm$  and  $Z^0$  at CERN (1983);
- Discovery of the  $\tau$  at SLAC (1975), Y (b quark) at Fermilab (1977) and the top quark at Fermilab (1995);
- Precise measurement of the  $Z^0$ , triple vector boson interaction, and the establishment of 3-family scenario at LEP (90's);
- Discovery of the Higgs boson, CERN (2012).

Despite this great success, some versions of the so-called physics beyond the standard model (BSM) theories [26] have been proposed with the motivation of solving the fine-tuning associated with the quadratic divergence in the Higgs mass [27]. These proposals involve, for instance, the supersymmetric models (SUSY) [11] and various forms of dynamical symmetry breaking and little Higgs models [17]. Some versions of theories with large extra dimensions [12], which allow the  $W/Z$  bosons to propagate freely in the extra dimensions, giving rise to Kaluza-Klein [28] excitations. Such excitations can also occur in the Randall-Sundrum models [14] and have motivated several experimental searches [29–34].

Several aspects of the physics beyond the standard model are being explored by the LHC at the TeV scale. The B2G (Beyond 2 Generations) analysis group in CMS covers models of new physics featuring the decay of new resonances to heavy standard model objects such as top,  $W$ ,  $Z$ , or Higgs bosons. The B2G group maintains synergy with other analyses in TOP, SUSY, and EXO (exotica) physics groups. A summary of the main B2G searches is presented in Fig. 2.4 which shows the observed limits at 95% C.L. for the production of vector-like quarks, resonances decaying into heavy quarks and into a pair of vector bosons, and the search for excited quarks. The search for new physics phenomena is the one of the most important items of the LHC agenda during the next years of scheduled operation.

### 2.2.1 Extra Dimensions

An increasing number of experimentalists are actively exploring the possibility that extra spacetime dimensions might be discovered at the LHC [12]. Despite the energy scales associated to many models are considered unreachable, the examination of curled-up spacetime dimensions is directly testable. We explore here some general properties of extra spacetime dimensions starting from the original Kaluza-Klein theory formulated back in the 1920's [36].

The Kaluza-Klein argument provides a way of unifying particles with different spins. Certain combinations of different particles with different spins in four dimensions can be viewed as different components of a single particle in higher dimensions, with a spin associated with the higher-dimensional Lorentz group. As an example consider a vector field in five dimensions  $V^{(5)} = (V_0, V_1, V_2, V_3, V_4)$ . The first four dimensions can be identified as the usual four-dimensional spacetime

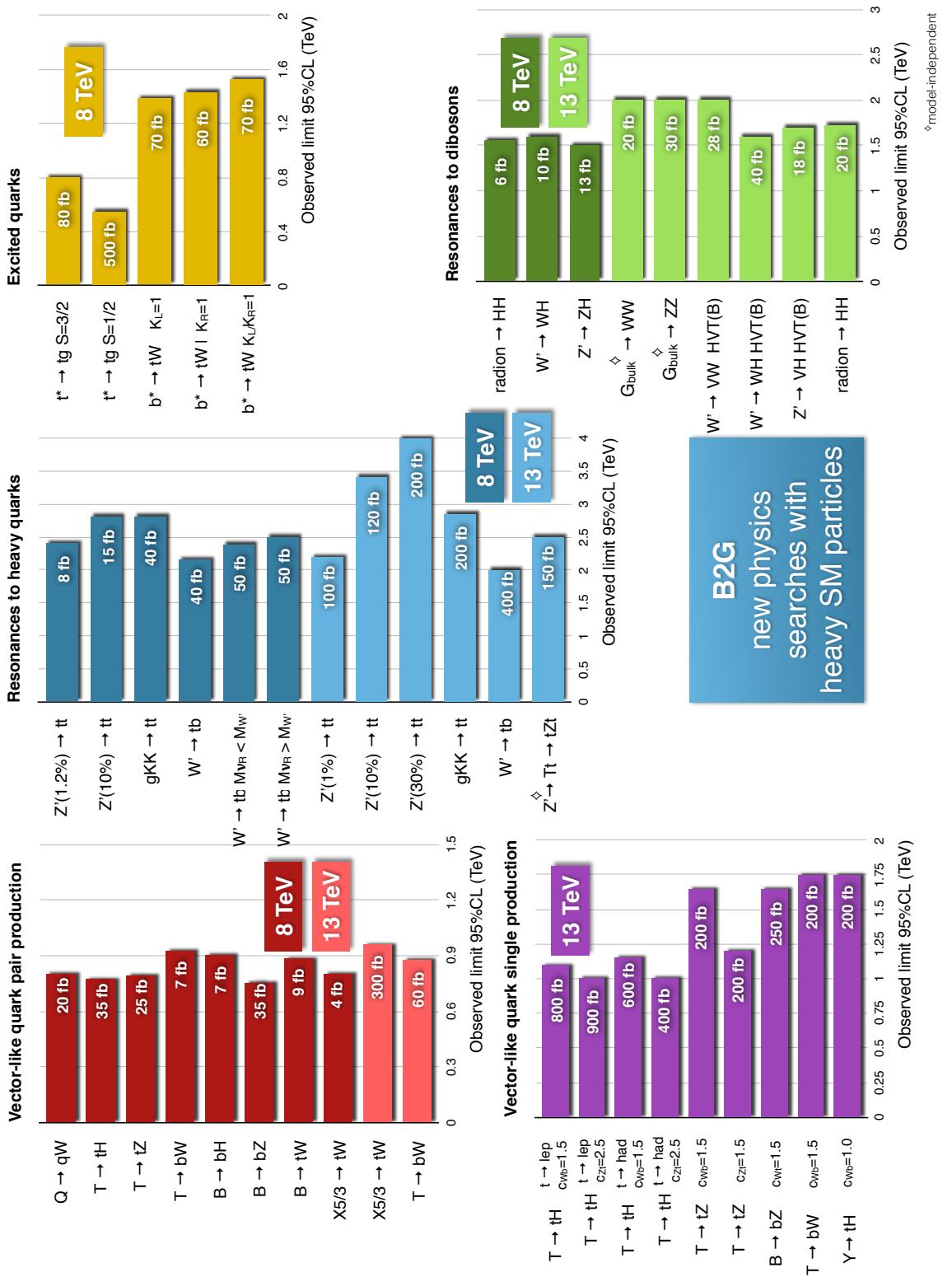


Figure 2.4: Summary of B2G group public physics results [35].

—  $V^{(4)} = (V_0, V_1, V_2, V_3)$  — and the remaining component as a scalar  $\phi = V_4$ . Thus, a single Lorentz vector representation in five dimensions has yielded both a spin-1 particle and a spin-0 particle in four dimensions.

Four-dimensional gravity can indeed be unified with electromagnetism into five-dimensional gravity. While this approach succeeds in unifying both fundamental forces known in the 1920's, now we know that there are at least four fundamental interactions. Therefore, unification along these lines would require even more extra dimensions.

If any extra space dimensions exist they must be sufficiently small in order not to be observable. In other words, they must be compactified down to a small length scale. For instance, consider at every point in spacetime an additional circle of radius  $R$  orthogonal to all of the known dimensions as it is illustrated in Fig. 2.5. If  $R$  is sufficiently small, the extra dimension is essentially unobservable.

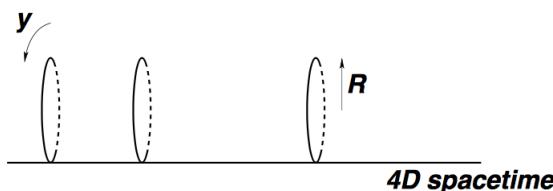


Figure 2.5: Extra dimension compactified on a circle of radius  $R$ .

More generally, the compactification of  $\delta$  extra spacetime dimensions can be achieved by a theory on  $\mathcal{M}_4 \times K$  where  $\mathcal{M}_4$  is the four-dimensional Minkowski spacetime and  $K$  is any  $\delta$ -dimensional compact manifold. For  $\delta = 2$ ,  $K = T_2$  — a two-torus — or  $K = S_1 \times S_1$ , two orthogonal circles, one for each compactified dimension. Many other topologies include manifolds with boundaries, for instance, a two-dimensional rectangle and the surface of a cylinder. A manifold with special points such as endpoints or boundaries is called orbifold.

## 2.2.2 The Randall-Sundrum Model

Beyond the basic idea of flat extra dimensions, deep phenomenological implications are achieved by expanding the compactification procedure to orbifolds. The proper construction of an orbifold begins with a manifold  $K$  and a discrete symmetry  $\Gamma$ . The resulting quotient space  $K/\Gamma$  is the orbifold.

As an example, consider the circle  $S^1$  satisfying the periodic boundary condition  $y \leftrightarrow y + 2\pi R$ . When the circle is restricted by the discrete  $Z_2$  symmetry

$\Gamma : y \leftrightarrow -y$ , the result is a line segment identified as the orbifold  $S^1/Z_2$  (Fig. 2.6).

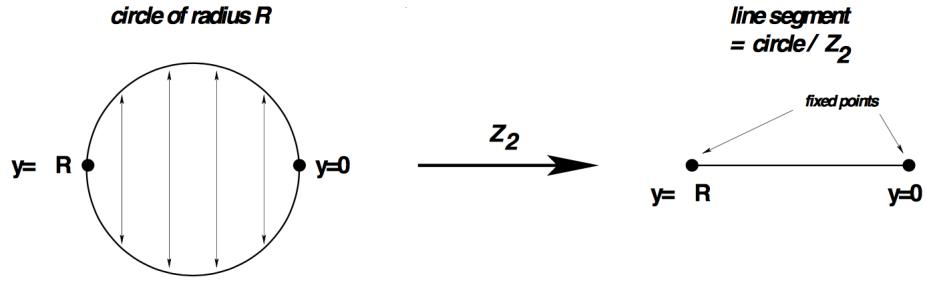


Figure 2.6: The orbifolding process of the circle  $S^1$  subjected to  $Z_2$ , resulting in the line segment  $S^1/Z_2$ .

The compactification scheme allows to describe the extra dimension as a line segment between two four dimensional branes, known as Planck and TeV brane (Fig. 2.7).

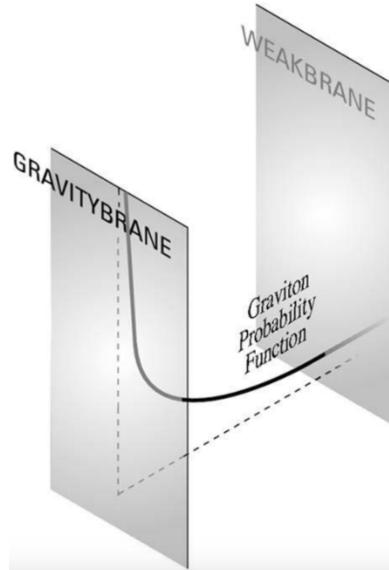


Figure 2.7: The Gravity (Planck) and Weak (TeV) branes are 4-dimensional boundaries of the extra dimension  $y$ .

The Randall-Sundrum (RS) model [14, 15] is a concrete framework of warped extra dimensions. As for the extra dimension, it considers the orbifold  $S^1/Z_2$ , and at every point along the extra dimension, it considers the ordinary flat 4-dimensional Minkowski metric plus a fifth coordinate  $y$ .

The metric satisfying these properties can be written as

$$ds^2 = e^{-A(y)} \eta_{\mu\nu} dx^\mu dx^\nu - dy^2. \quad (2.35)$$

The amount of curvature along the extra dimension depends on the function  $e^{-A(y)}$ , which is the warping factor. The procedure to find the solution for the function  $A(y)$  involves the introduction of the curvature factor

$$k^2 = -\frac{\Lambda_5}{12M_*^3}, \quad (2.36)$$

relating the bulk cosmological constant  $\Lambda_5$  and the 5-dimensional Planck scale  $M_*$ . The bulk cosmological constant [12] comes from the 5-dimensional Einstein action for gravity and it has units of mass to the fifth power,  $[\Lambda_5] = M^5$ ; therefore, the curvature factor has units of mass,  $[k] = (M^5/M^3)^{1/2} = M$ . Furthermore, a solution of the Einstein equation can only exist if the bulk cosmological constant is negative  $\Lambda_5 < 0$ .

According to the RS model, the effective scale of gravity — that is the 4-dimensional Planck scale — is found to be

$$M_{Pl}^2 = M_*^3 \int_{y=-r_c}^{y=r_c} e^{-2k|y|} dy = \frac{M_*^3}{k} \left( 1 - e^{-2kr_c} \right). \quad (2.37)$$

For large values of the compactification radius  $r_c$ , the 4-dimensional Planck scale barely depends on the size of the extra dimension. This means that an exponential hierarchy between the weak and the Planck scales is naturally introduced, providing a solution to the hierarchy problem of the standard model.

### 2.2.3 Bulk Graviton Model

The solution to the hierarchy problem of the standard model through the Randall-Sundrum model with a warped extra dimension invokes the existence of new particles at the TeV scale. A key feature of the bulk graviton model [16] is the propagation of the standard model fields in the extra dimension. As such, the standard model particles are identified with the zero-modes of the 5-dimensional fields. Provided that the new resonances have non-negligible coupling to the standard model particles, new signals might be observed at the LHC. Spin-2 gravitons, whose masses and couplings are set by the TeV scale, would appear in experiments as widely separated resonances.

The bulk graviton extension of the original Randall-Sundrum model predicts a highly enhanced branching ratio of gravitons decaying to vector boson while

suppressing the light fermions and photons channels. The production of bulk gravitons from gluon fusion and their decay into vector bosons  $W/Z$  can be significant.

The golden channel in the bulk graviton model is  $G \rightarrow Z_L Z_L$ , where  $Z_L$  indicates the longitudinal component of the vector boson. The partial decay width of this channel is given by

$$\Gamma(G \rightarrow Z_L Z_L) \approx \frac{(c x_n^G)^2 m_n^G}{960\pi}, \quad (2.38)$$

where  $c \equiv k/M_{Pl}$ , and the roots of the Bessel function of order 1,  $x_n^G = 3.83, 7.02, 10.17, 13.32$ , give the masses of the first 4 Kaluza-Klein gravitons:

$$m_n^G = k e^{-k\pi R} x_n^G, \quad (2.39)$$

where  $k$  is the curvature scale and  $R$  is the proper size of the extra dimension.

The production cross section for the bulk graviton with  $k/M_{Pl} = 0.1$  is shown in Fig. 2.8. These calculations were obtained with MG5\_aMC@NLO [37] as reported in Ref. [38]. The natural width of a bulk graviton with these characteristics is less than 0.1% of its mass, which is negligible with respect to the experimental detector resolution. For the experimental analysis we will use simulated samples with  $k/M_{Pl} = 0.5$ , increasing the graviton width to 1.5%.

The most stringent limits to date on the bulk graviton signal strength are shown in Fig. 2.9, in the mass range 0.6 to 4.0 TeV. As observed in the figure, the sensitivity arises mainly from the 4q and 2q $\ell\nu$  channels.

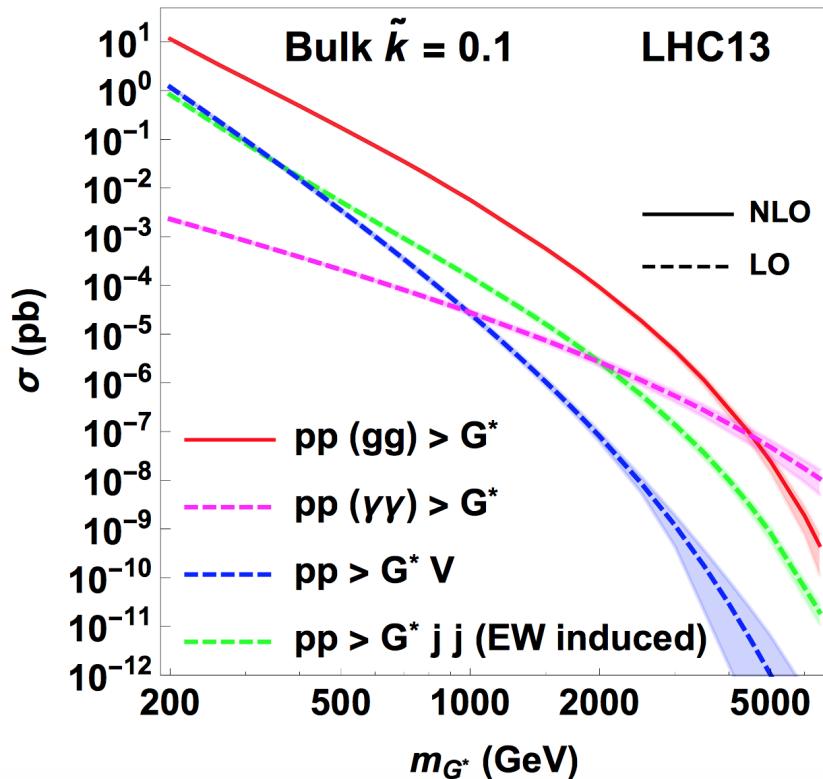


Figure 2.8: Production cross section for the bulk graviton scenario [38]. The red curve corresponds to the inclusive production, the green curve to the associated production with two jets, the blue curve to the associated production with a vector boson, and the magenta curve to the photon-fusion production. NLO (LO) calculations are shown in continuous (dashed) lines.

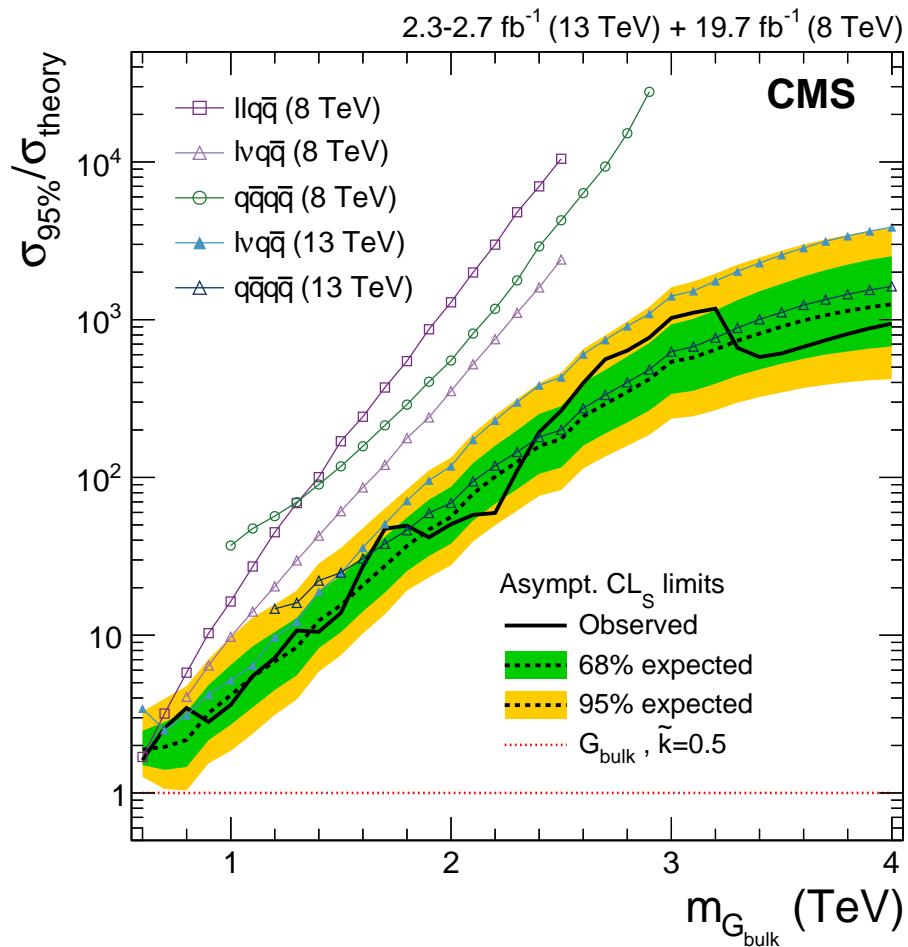


Figure 2.9: Exclusion limits at 95% CL on the signal strength in the bulk graviton model with  $\tilde{k} = 0.5$ , as function of the resonance mass, obtained by combining 8 and 13 TeV diboson searches [39].

# Chapter 3

## The LHC and the CMS Experiment

### 3.1 The Large Hadron Collider

The state-of-the-art experiments in the field of high energy physics are at the European Organization for Nuclear Research (CERN), where the Large Hadron Collider (LHC) and its complex of accelerators is hosted (Fig. 3.1).

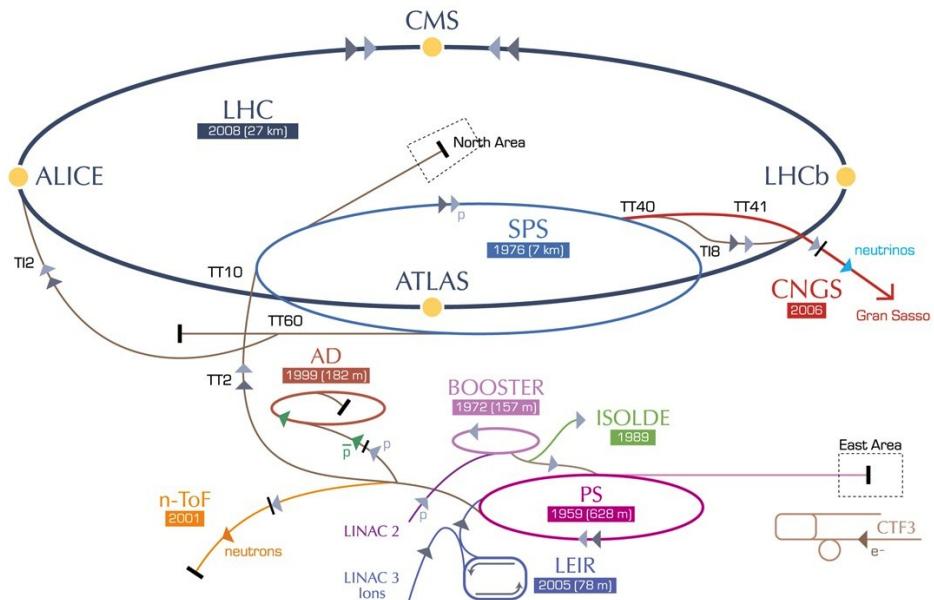


Figure 3.1: Overview of the CERN's accelerator complex [40].

The LHC is a circular accelerator with 27 km of circumference installed in a tunnel 50 to 175 meters underground. It was originally designed to collide protons at a centre-of-mass energy of 14 TeV with a design luminosity of  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , with the possibility of colliding also heavy ions in different configurations.

In a circular accelerator such as the LHC, radio-frequency cavities provide the boost while dipole superconducting magnets (Fig. 3.2) supply the bending magnetic field that keeps the protons in orbit.

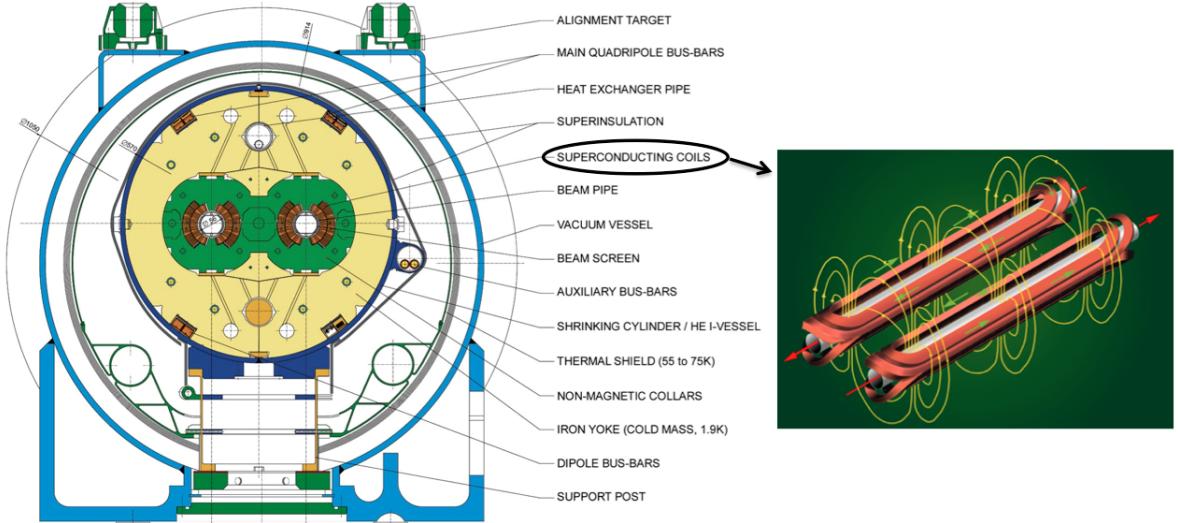


Figure 3.2: Transverse section of the dipole superconducting coil and magnetic field induced by it [41].

The LHC circulates protons inside its beam-pipes in several very closely packed bunches, with each bunch made out of approximately  $10^{11}$  particles. To maximize the probability of the protons colliding with one another, the LHC squeezes the beam within a transverse size of  $\sigma_x \approx \sigma_y \approx 15 \mu\text{m}$ . Every 25 ns these bunches cross one another and several proton-proton collisions may take place, resulting in more than one interaction point per bunch crossing; this phenomenon is known as pileup. For instance, during the 2015 run at 13 TeV, there was in average 20 interactions per bunch crossing.

The event rate (events/s) for a given process generated at the LHC is given by  $\sigma_i \times \mathcal{L}$ , where  $\sigma_i$  is the cross section for the process under study and  $\mathcal{L}$  is the instantaneous luminosity which depends only on beam parameters such as the number of particles per bunch ( $n_1, n_2$ ), the revolution frequency ( $f$ ) and the Gaussian widths of the beam profile in the horizontal and vertical plane ( $\sigma_x, \sigma_y$ ):

$$\mathcal{L} = f \frac{n_1 n_2}{4\pi \sigma_x \sigma_y}. \quad (3.1)$$

The total number of events  $N_i$  produced in a time  $T$  is basically the integral of the event rate for the specific process,  $\sigma_i \int_0^T \mathcal{L} dt$ .

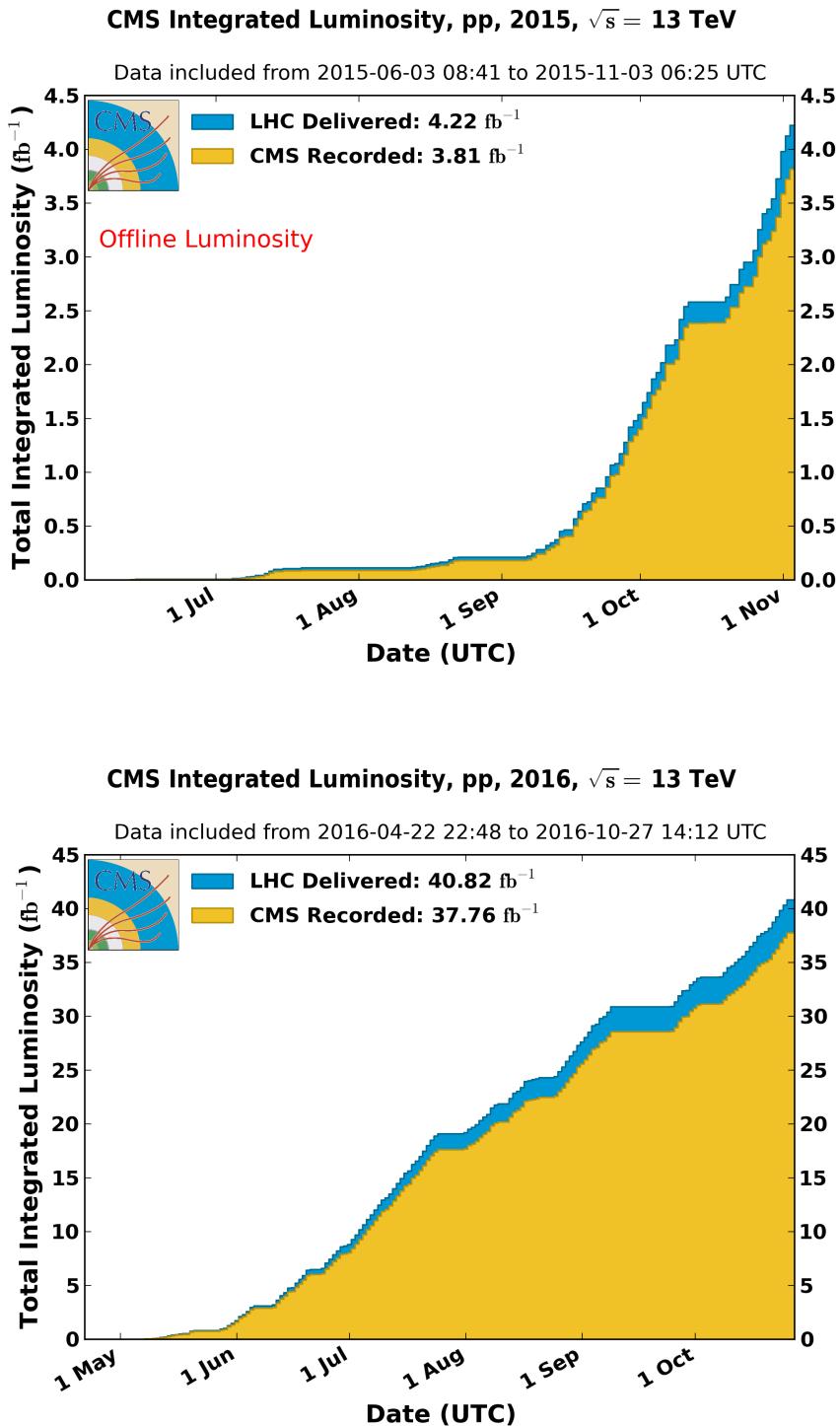


Figure 3.3: Cumulative offline luminosity delivered by the LHC (blue) and recorded by CMS (orange) during stable beams and for p-p collisions at 13 TeV center-of-mass energy in 2015 (top) and 2016 (bottom) [42].

The first collisions at the LHC took place in 2010 using proton beams with a center-of-mass energy of 7 TeV. In 2012 the center of mass energy was increased to 8 TeV and a further  $23 \text{ fb}^{-1}$  of data were delivered. The next two years were dedicated to preparations for the LHC Run 2, ramping the center-of-mass energy up to 13 TeV. In 2015 the LHC delivered  $4.2 \text{ fb}^{-1}$ , and fantastically achieved  $41.4 \text{ fb}^{-1}$  in 2016. The performance of LHC over the Run 2 is summarised in Fig. 3.3.

## 3.2 The CMS Experiment

The Compact Muon Solenoid (CMS) [43, 44] is one of the main experiments at the LHC. It is build around a solenoid magnet that takes the form of a cylindrical coil of superconducting cable that generates a field of 3.8 T. The bulk of the detector weights 12,500 tons and it is 21.6 meters long, 15 meters of diameter and 15 meters high.

CMS has a very compact design that emphasizes good muon identification. It provides good charge and momentum resolution including efficient  $b$  and  $\tau$  tagging capability as well as a good electromagnetic energy resolution and good missing transverse energy resolution. Contained within the superconducting solenoid volume are the silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL). Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid. These sub-detectors are situated in an shell arrangement, and all the barrel systems have a forward equivalent to guarantee a  $4\pi$  solid angle coverage. Studying the layout in Fig. 3.4 more closely, one can observe additional very forward structures (muon detectors and a forward sampling calorimeter) to cover a high  $|\eta|$  range.

The identification of individual particle takes information from the various elements of the CMS detector. The energy of photons is directly obtained from the ECAL measurements, while the energy of electrons is determined from a combination of the electron momentum at the primary interaction vertex as determined by the tracker, the energy of the corresponding ECAL cluster, and the energy sum of the associated bremsstrahlung photons. The energy of muons is obtained from the curvature of the muon of the corresponding track. The energy of charged hadrons

is determined from a combination of their momentum measured in the tracker and the matching ECAL and HCAL energy deposits. Finally, the energy of neutral hadrons is obtained from the corresponding ECAL and HCAL energy deposition.

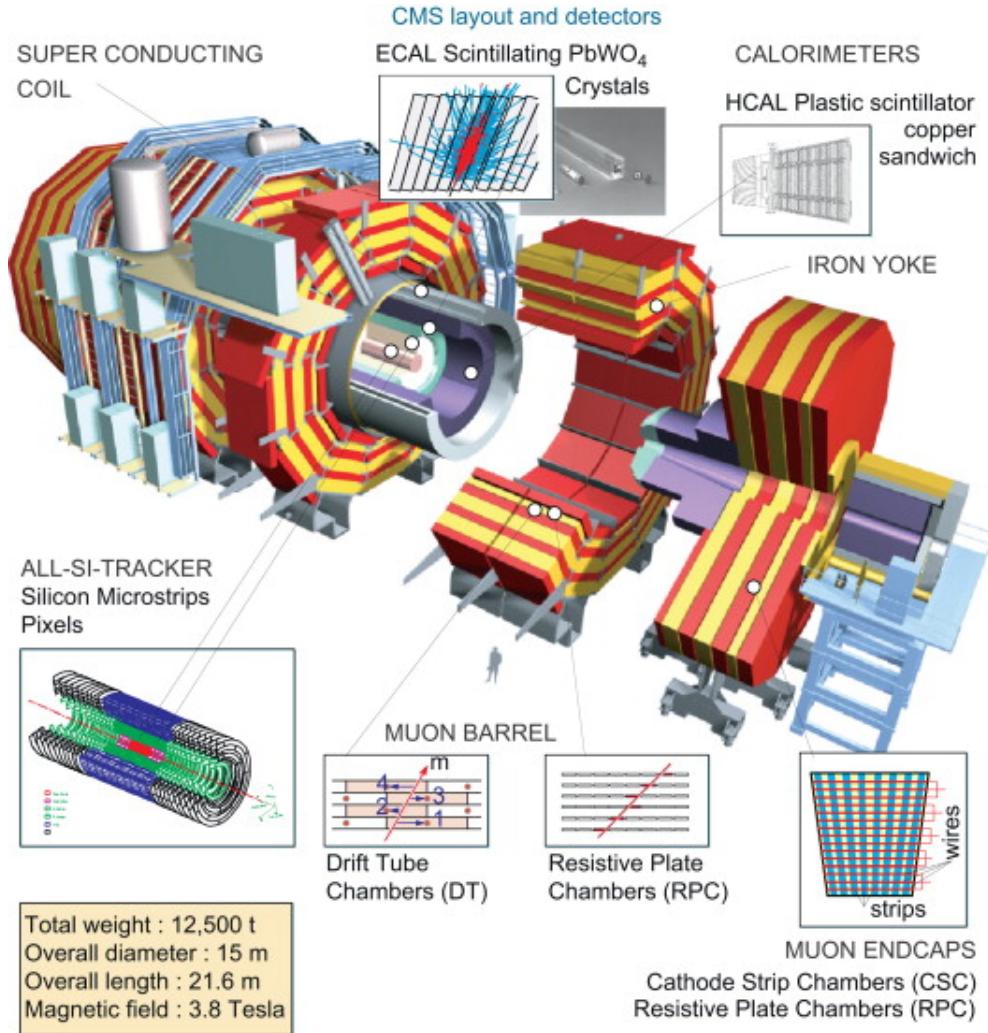


Figure 3.4: The CMS detector and its components [45].

CMS uses a right handed coordinate system centered at the nominal collision point in which the  $+x$  axis points towards the center of the LHC ring,  $y$  points up, and  $+z$  points along the direction of the beam. In a cylindrical system  $r$  is the radius from the nominal beam line and  $\phi$  is the azimuthal angle. The pseudorapidity is defined as  $\eta = -\ln[\tan(\theta/2)]$ , where  $\theta$  is the polar angle measured from the  $+z$  axis. The transverse momentum is defined as  $p_T = p \cdot \sin \theta$ , with  $p$  being the particle momentum.

## 3.3 Main Detector Components

### 3.3.1 Tracker System

The tracker system comprises two components: the pixel detector and the silicon strip detector. The pixel detector is the innermost subsystem, closest to the beam pipe, containing 66 million of  $100 \times 150 \mu\text{m}^2$  pixels responsible for track reconstruction and primary vertex identification as well as  $b$  tagging. It is arranged in three barrel layers and two endcap disks at each end. The pixel detector is followed by a silicon strip system, which determines the momentum of electrically charged particles that traverse it. The system is structured in ten barrel layers and twelve endcap disks, composed of 9.6 million strips with pitch between 80 and  $180 \mu\text{m}$ , with a total silicon surface area of  $200 \text{ m}^2$ .

After the first years of CMS operations, there was an upgrade proposal [46] to replace the pixel detector for a new high efficiency and low mass detector including four barrel layer and three forward/backward disks. The new pixel was just installed and the commissioning is currently ongoing, being 2017 the first year to include this important upgrade.

### 3.3.2 Electromagnetic Calorimeter

The tracker is followed by an Electromagnetic Calorimeter (ECAL), which uses 75,000 lead tungstate ( $\text{PbWO}_4$ ) crystals to determine the energy of electrons and photons, producing an amount of light that is proportional to the particle's energy. The ECAL is distributed in a barrel ( $|\eta| < 1.479$ ) and two endcap ( $1.479 < |\eta| < 3.0$ ) regions. Photodetectors placed in the back of each crystal detect the scintillation light and convert it to electrical signals. Electromagnetic showers are very narrow in lead tungstate helping particle identification and the implementation of isolation criteria. Figure 3.5 shows a transverse section of the ECAL.

Pre-shower detectors consisting of two planes of lead followed by silicon sensors are located in front of the endcaps. When a photon passes through the lead layer it causes an electromagnetic shower containing electron-positron pairs, which are detected by the silicon sensors. From this the photon's energy is measured, whilst having two detector layers gives the particle's position. The preshower has a much finer granularity than the ECAL, with detector strips 2 mm wide compared

to the 3 cm-wide ECAL crystals, what allows to distinguish closely-spaced photons.

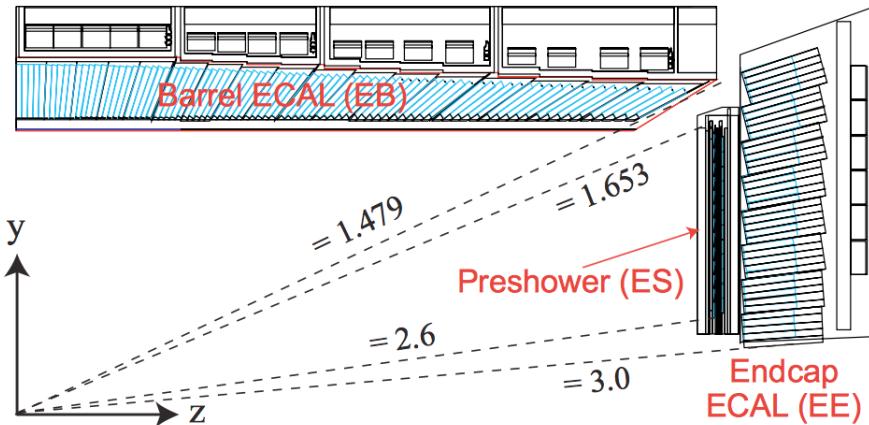


Figure 3.5: Transverse section through the ECAL [43].

### 3.3.3 Hadronic Calorimeter

The ECAL is surrounded by a brass/scintillator sampling Hadron Calorimeter (HCAL) to measure the energy of hadrons. Optical fibers collect up the light produced by the plastic scintillators and feed it into readout boxes where photodetectors amplify the signal. The HCAL covers the region with  $|\eta| < 3.0$ . Their thickness varies from 7 to 10 interaction lengths depending on  $\eta$ ; a scintillator placed outside of the coil at the innermost muon detector extends the instrument thickness to more than 10 interaction lengths everywhere. Quartz fibre and iron forward calorimeters, read out by photomultipliers, extend the calorimeter coverage in the range  $3.0 < |\eta| < 5.0$  comprising the hadron endcap (HE) and hadron forward (HF) calorimeters.

### 3.3.4 Solenoid

The central feature of the CMS apparatus is a 13 m long and 6 m diameter superconducting solenoid, providing a magnetic field of 3.8 T. Within the field volume are the tracker system and the calorimeters. The magnet bending power ensures the unambiguous determination of the sign for muons with a momentum of  $p \approx 1$  TeV, providing a momentum resolution of  $\Delta p/p \approx 10\%$ .

### 3.3.5 Muon System

Outside the solenoid is the muon system, the most visible part of CMS shown in Fig. 3.6. The CMS design relies on the high bending and excellent muon momentum resolution, which uses an iron return yoke interleaved with the muon chambers to increase the magnetic field. With the field parallel to the LHC beam axis, the muon tracks are bent in the transverse plane.

The iron yoke is instrumented with aluminum Drift Tubes (DT) in the barrel ( $|\eta| < 1.2$ ) and Cathode Strip Chambers (CSC) in the end-cap region ( $0.9 < |\eta| < 2.4$ ). Due to the iron yoke, the momentum resolution of the CMS muon system is dominated by the multiple scattering. The standalone muon momentum resolution is  $\sigma(p_T)/p_T = 9\%$  for  $p_T \leq 200 \text{ GeV}/c$  and 15–40% at  $p_T = 1 \text{ TeV}/c$ , depending on  $\eta$ . Including the tracking system improves the result by an order of magnitude for low momenta. At 1 TeV the contribution of both measurements lead to a momentum resolution of about 5%.

The DT and CSC subsystems can each trigger on muons with large transverse momentum in the range  $|\eta| < 2.4$ . However, for the full LHC luminosity, faster trigger chambers are needed to associate the detected muons to the right crossing of proton bunches. Resistive Plate Chambers (RPCs) covering the region  $|\eta| < 1.6$  are used by the muon system for fast trigger.

### 3.3.6 Trigger and Data Acquisition

The CMS online systems need to select around 1 kHz of interesting events out of a rate of 40 MHz. The event selection is done with two trigger levels: the level 1 trigger [47], based on custom electronics, reduces the rate to 100 kHz. The data acquisition (DAQ) system [48, 49] reads out the detector and passes the events to the high-level trigger, a software system based on the full CMS reconstruction software running on a farm of computers. The CMS DAQ system was designed to handle a throughput of 100 GB/s, making it the highest throughput DAQ system in high energy physics to date.

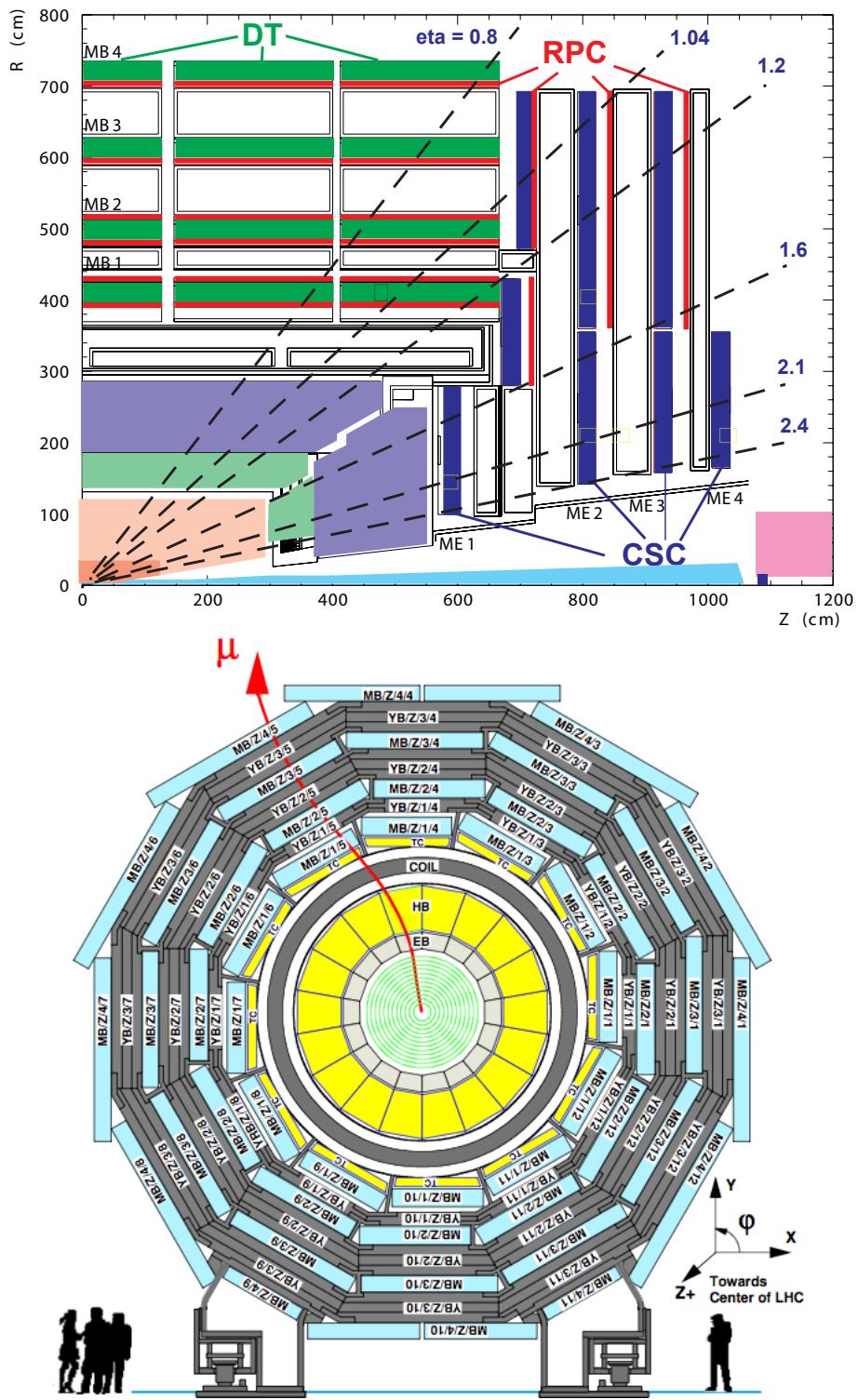


Figure 3.6: Pseudorapidity coverage of the muon system (top) and muon track in the transverse plane (bottom).

## 3.4 Identification of Physical Objects

The identification of stable particles in CMS relies on the different interactions of a particle with the sub-detectors. While photons, electrons and hadrons lose most of their energy and are stopped in the calorimeters, muons deposit only a small fraction of their energy through ionization so they reach the outer part of the detector, where the muon chambers are located. These characteristic signatures (Fig. 3.7) based on tracking and calorimetry are crucial aspects for particle identification.

The sub-detectors in CMS are stacked in radial layers and a particle passes through these layers sequentially from the collision point outward: first the tracking system, then the electromagnetic and hadronic calorimeters, finally the muon system. All layers are embedded in a magnetic field in order to bend the tracks of charged particles for momentum and charge sign determination.

### 3.4.1 Track Reconstruction

The track reconstruction uses an iterative procedure [50] consisting of a number of steps to select the better tracks first; the hits associated with the first tracks are removed, and then, other tracks are reconstructed from the remaining hits.

Each of the tracking steps starts with a collection of seeds formed from 2 (a pair seed) or 3 (a triplet seed) pixel hits consistent with some minimum track  $p_T$ , and coming from some region of the beam spot. The first steps use triplet seeds and higher minimum track  $p_T$ , these are followed by steps using pair seeds and lower  $p_T$ . The later steps use seeds that contain hits from the silicon strip detector to find detached tracks, *e.g.* from decay products of  $K_s^0$  mesons or  $\Lambda^0$  baryons.

### 3.4.2 Primary Vertex Reconstruction

CMS observes the decay products of various particles produced and work backwards to determine which collision interactions produced which particles. In the CMS tracker (made up of silicon pixels and strips), the number of hits grows linearly with pileup. As these hits need to be combined into tracks, the number of possible combinations that make a track grows fast with pileup. Fortunately, the high granularity and efficiency of the tracker provides the means to distinguish the many tracks in an event. An illustration of the CMS tracking capabilities is shown

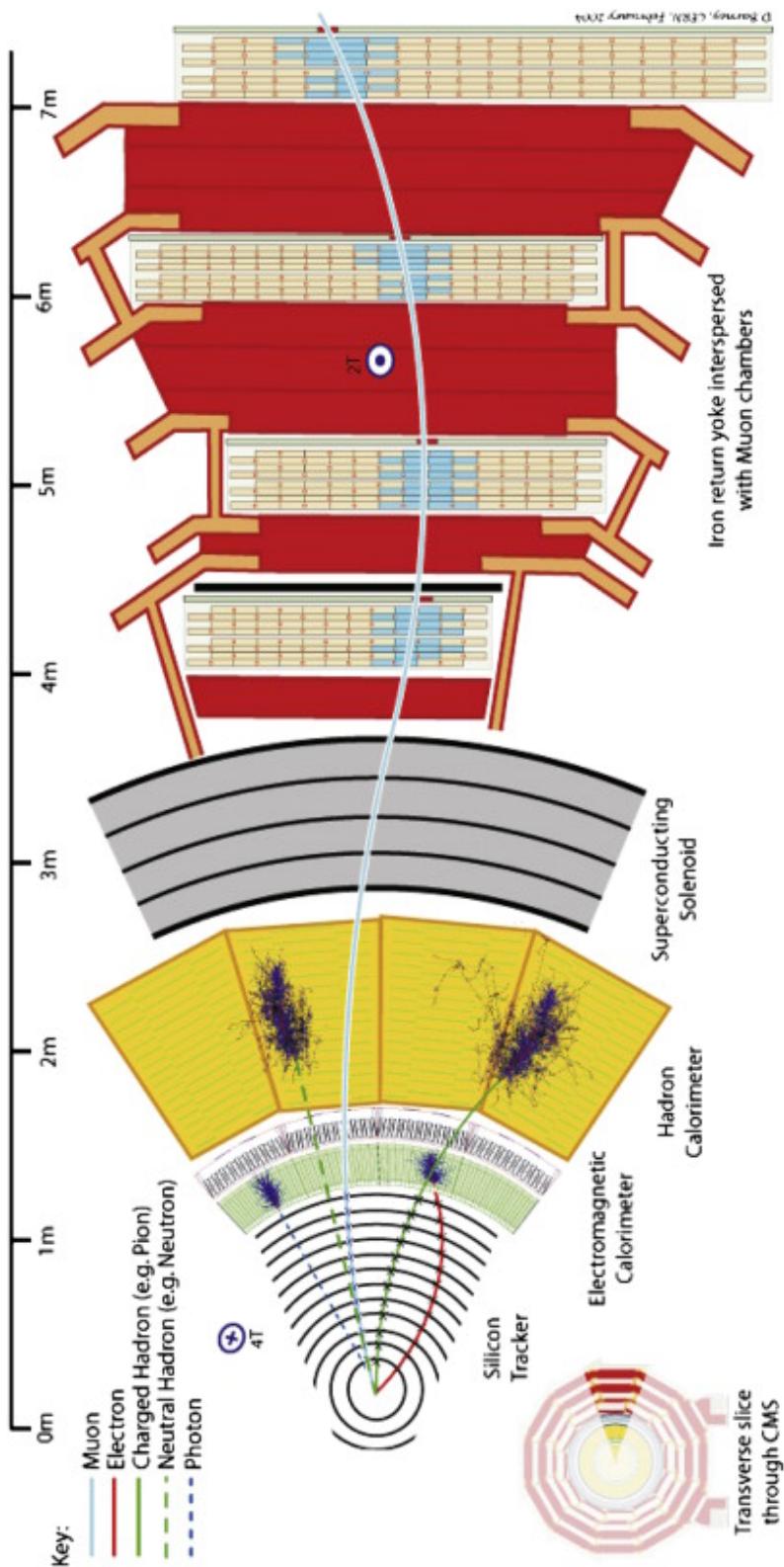


Figure 3.7: Characteristic signatures of different particles in CMS.

in Fig. 3.8; the figure displays the reconstructed tracks and the primary vertices in a real event recorded on 2016-Oct-14, during the proton-proton collisions at 13 TeV delivered by the LHC.

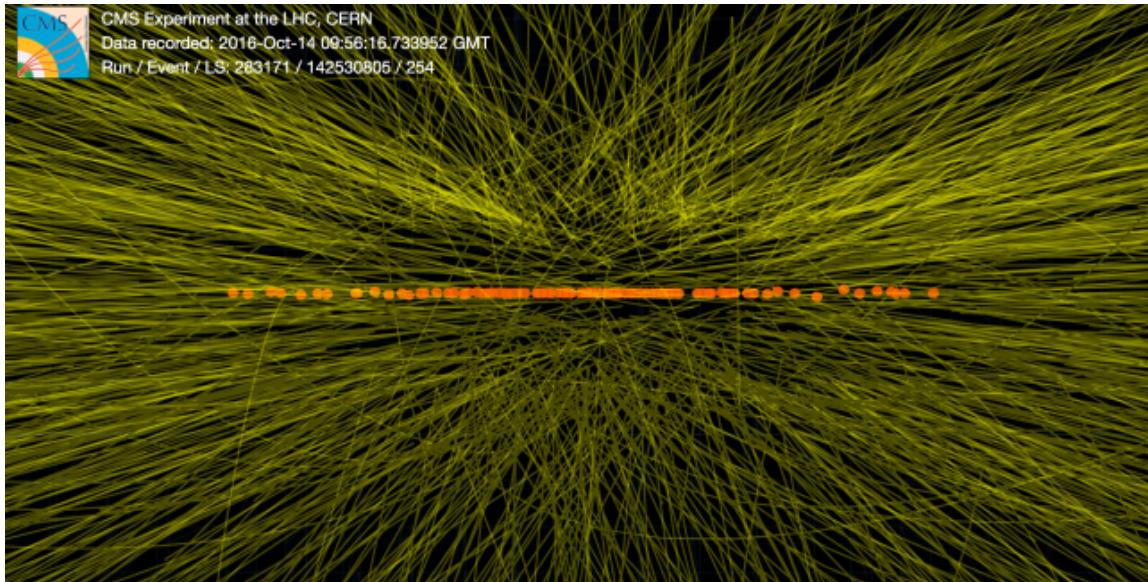


Figure 3.8: Tracks and primary vertices as seen in a typical collision [51].

The vertex reconstruction can be seen as a clustering procedure in which each vertex is a cluster of tracks selected by a fitting algorithm. Among the algorithms used in CMS for vertex fitting are the Kalman filter (KVF) [52], and the Adaptive Vertex Fitter (AVF) [53]. The Kalman filter is a least-squares estimator which minimizes the sum of the squared distances of all tracks from the vertex position. The AVF algorithm is formulated as an iterative re-weighted Kalman filter, associating a weight  $w_i$  interpreted as the probability that track  $i$  belongs to a vertex. The AVF algorithm is a robustification of the Kalman filter to deal with fitting errors, such as mis-associated tracks or mis-measured track errors.

### 3.4.3 Jet Reconstruction

Hadronic jets are the experimental signatures of quarks and gluons (Fig. 3.9). In CMS, jets are clustered with the anti- $k_T$  algorithm [54] starting from a collection of particle-flow (PF) candidates [55, 56]. A correction based on the projected area of the jet on the front face of the calorimeter is used to take into account the extra energy due to neutral particles coming from pileup.

The particle-flow (PF) algorithm integrates measurements from all components of the CMS detector in order to reconstruct a complete list of candidates per event, including muons, electrons, photons, charge and neutral hadrons. The jet clusterization routine loops over the list of PF candidates, and recombines two particles  $i$  and  $j$  based on the condition  $d_{ij} < k_{ti}^{2p}$ , where the distance  $d_{ij}$  is defined as:

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}{R^2}. \quad (3.2)$$

The transverse momentum, pseudo-rapidity, and azimuth angle of particle  $i$ , are  $k_{ti}$ ,  $\eta_i$ , and  $\phi_i$ , respectively. The parameter  $p = -1$  is characteristic of the anti- $k_T$  algorithm [54], and ensures infrared-safe jets. The aperture of the jet is controlled with the parameter  $R$ , which takes the value  $R = 0.8$  for the case of merged jets; the number  $R = 0.8$  is set by the CMS jet study group [57], and it is kept fix along the jet reconstruction routine for both data and simulation.

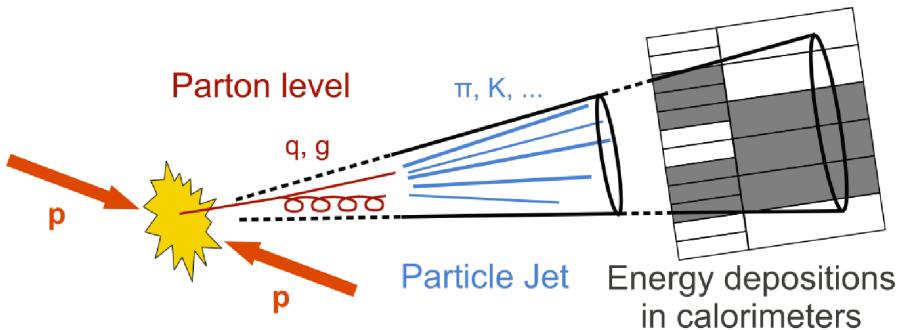


Figure 3.9: Schematic production of a hadronic jet.

Hadronically decaying W and Z bosons are identified as jets with distance parameter  $R = 0.8$ . In order to discriminate W/Z jet candidates against multijet backgrounds, the reconstructed jet mass is required to be close to the W or Z boson mass, in addition to require a two-prong jet substructure produced by the initial quarks. Jets coming from the merged decay products of a single V boson are usually referred as V jets.

Different jet grooming algorithms have been explored in CMS and their performance in multijet processes has been studied in detail [58]. The goal of jet grooming is the elimination of soft, large-angle QCD radiation that increases the V jet mass compared to the initial V boson. Further discrimination is obtained from

the quantity called N-subjettiness [59] defined as

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min(\Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{N,k}), \quad (3.3)$$

where the index  $k$  runs over the jet constituents and the distances  $\Delta R_{n,k}$  are calculated with respect to the axis of the  $n$ th subjet. The normalization factor  $d_0$  is calculated as  $d_0 = \sum_k p_{T,k} R_0$ , setting  $R_0$  to the radius of the original jet.

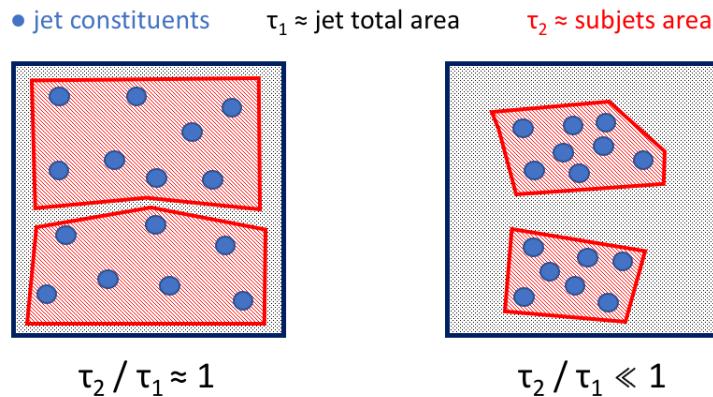


Figure 3.10: Schematic representation of the N-subjettiness discriminant. A two-prong jet substructure produces lower values of the ratio  $\tau_2 / \tau_1$ .

The variable  $\tau_N$  quantifies the capability of clustering the jet constituents in exactly  $N$  subjets. The ratio  $\tau_{21} = \tau_2 / \tau_1$  is actually a powerful discriminant between jets originating from hadronic  $V$  decays and from gluon and single-quark hadronization. Jets coming from hadronic  $W$  or  $Z$  decays are characterized by lower values of  $\tau_{21}$ , given the two-prong substructure adopted by the jet constituents schematically depicted in Fig. 3.10. The visual representation in the above figure has to be taken with a grain of salt, since the definition of  $\tau_N$  (Eq. 3.3) depends on the kinematic properties of the jet constituents, rather than the jet area.

### 3.4.4 Missing Energy

Neutrinos produced in the final state, as well as other hypothetical weakly interactive neutral particles, escape from the detector causing an energy imbalance in the observed event. Momentum conservation is the available way to reveal the presence of neutrinos.

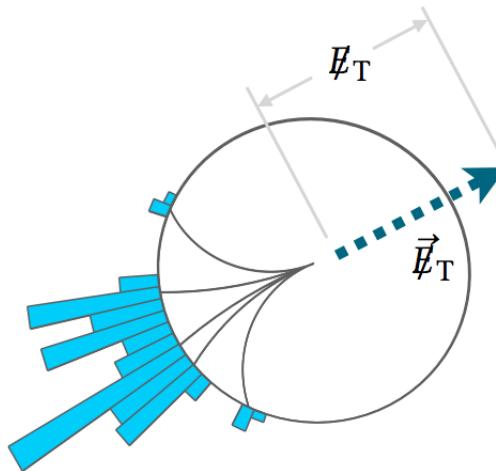


Figure 3.11: Missing transverse momentum and its magnitude.

Since the  $z$ -component of the momentum of the colliding partons is not known, one cannot determine the net missing energy caused by neutrinos. However, the total momentum in the transverse plane is zero to a very good approximation.

One can define the missing transverse momentum as the negative vector sum of the transverse energies of all final-state particles reconstructed in the detector. The missing transverse energy  $E_T$  corresponds to the magnitude of the missing transverse momentum (Fig. 3.11).

$$\vec{E}_T = - \sum_i \vec{E}_T^i \quad (3.4)$$

## 3.5 Software and Computing

The Monte Carlo (MC) method is a numerical technique for calculating probabilities and related quantities by using sequences of random numbers. The MC technique is appropriate to simulate physical events, since the randomness of the MC method allows to capture the uncertainty that is present in most physical measurements.

Of great interest in physics are the MC event generators, they include matrix element calculators like AlpGen [60], MadGraph [61], and multi-purpose MC generator like PYTHIA [62]. Matrix element calculators deliver an event at the parton level, then MC generators can further be used to develop a fully hadronized event.

PYTHIA provides the generator level of what happens in a particle collision: implements models for a number of physics aspects, such as hard and soft interactions, parton distributions, initial and final state radiation, multiple interactions, fragmentation and decay, and hadronization of the parton-level events. However, the generation of events is only the first step in the complete analysis chain.

### 3.5.1 Event Reconstruction

The detailed simulation and reconstruction of physics events is extremely time consuming. The analysis chain is decomposed into four major steps as follows:

1. *Generation of Monte Carlo events*: The Monte Carlo events are created using generators like PYTHIA. These generators produce a list of particles and their four-vectors;
2. *Simulation of material effects*: This is the most time consuming step. The output of this step is called *SimHits*. They contain the information about the energy stored in different detector elements at different times;
3. *Simulation of readout electronics (digitalization)*: The detector converts the energy deposited by the particles into electronic signals that are converted to digital information. Since the simulation of material effects requires large amount of CPU time, the minimum bias events are randomly selected from a large pool of simulated events and combined with the simulated signal events. The combination of minimum bias events with a signal event and the simulation of the detector response to the energy deposition are performed by the reconstruction software. The output created in this step is called *DIGI*;
4. *Reconstruction of physics/analysis objects*: The reconstruction is performed in several sub-steps. First the *DIGI* are combined to reconstructed hits *RecHits*, which for example combine several strips of the silicon tracking detectors. Then *RecHits* are used to find tracks in the inner tracker and the muon chambers and clusters in the calorimeters. The reconstruction can produce more complicated objects like jets or information about the missing energy and finally physical objects like electrons, photons, muons etc.

### 3.5.2 Data Analysis with CMSSW

The huge amount of data collected by CMS requires large resources for storage and dedicated analysis software. The CMS software (CMSSW) consists of over a thousand sub-packages providing an extensive toolkit for users to carry out analysis of data. It also gathers services needed by the reconstruction modules that process the data. The CMSSW executable, called cmsRun, is configured at run time by the user's configuration file. This file tells cmsRun which data to use, which modules to execute, which parameter settings to use for each module, and how the events are filtered.

The data is organised according to the event data model (Fig. 3.12). Each event is a C++ object container for the reconstructed data related to a particular collision, and physical particles are accessed through C++ objects.

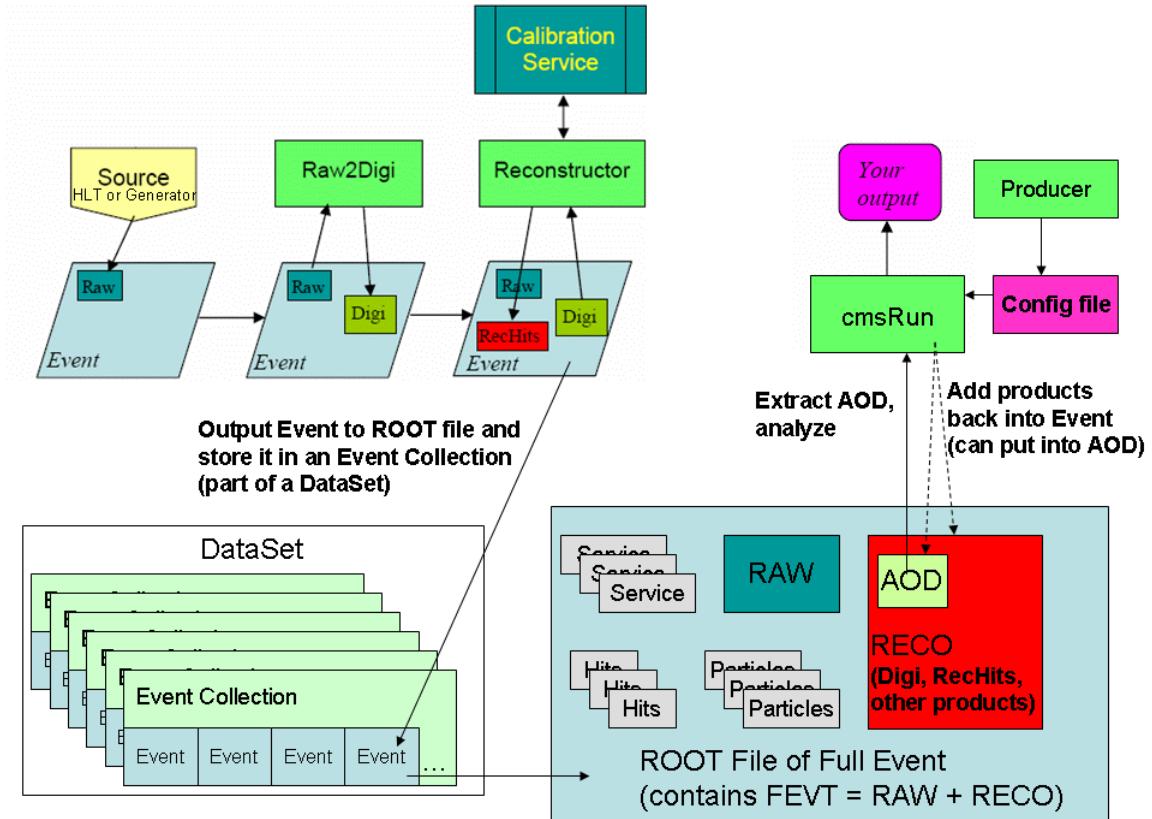


Figure 3.12: The Event Data Model

An event starts as a collection of the raw data. As the event data is processed, products are stored in the event as reconstructed data objects. The event also

contains metadata describing the configuration of the software used for the reconstruction of each object and the conditions for alignment and calibration. Analysis Object Data (AOD) is a subset of the reconstructed data sufficient for most analysis to access the relevant objects.

### 3.5.3 Computing Infrastructure

CMS makes use of a grid of computers connected together in an hierarchical organisation so that users around the world can share data and computational power. The structure called Worldwide LHC Computing Grid (WLCG) is divided in clusters of computers called Tiers which are classified depending on their computational power and storage capacity.

- Tier 0: The Tier T0 is located at CERN, where all raw information coming from the detector is saved, reconstructed, and later on transmitted to the rest of the chain. Recently a mirror of the CERN site was deployed at Budapest and they are connected at 100 Gbps to each other;
- Tier 1: There are fourteen national T1 sites around the world, providing storage and redistribution for MC events generated by the T2's;
- Tier 2: There T2 centers provides capacity for user analysis, calibration studies, and Monte Carlo production for the whole experiment.
- Tier 3: Any small cluster of computers installed at an institute providing local access to the Grid.

The São Paulo Research and Analysis Center (SPRACE) [63] was implemented in 2003 to collaborate with the DØ and CMS experiments. SPRACE hosts a T2 of the CMS computing structure — the BR-SP-SPRACE —, providing processing power of 25,200 HEPSPEC06 and 1,450 TB of storage, all with a redundant 100 Gbps link to the Fermilab T1 in the USA, which is our main connection to the WLCG. During the development of this work we made extensive use of the resources of the local CMS center at SPRACE.

# Chapter 4

## Definition of the ZZ Semi-leptonic Channel

### 4.1 General Overview

Experiments at the energy frontier shift boundaries and open new ways in the exploration of uncharted regions. The possibility of discovering new particles in unexplored regions is one of the motivations to perform data analysis in high energy physics. Although a completely model-independent search can be performed [64], in general we focus on particular benchmark models, based either on theoretical motivations or on experimental hints. From a given benchmark model we can calculate which final states are accessible.

When performing an experimental search for physics beyond the standard model, events dominated by the new physics process are called signal events, while events dominated by standard model processes with the same signature are called background events. In general, standard model processes have much larger cross-sections than the posited signal process; therefore, we have to apply selection requirements that preserve a large fraction of the signal while reducing the standard model background.

The remaining background has to be estimated, both from theoretical predictions and from data-driven methods. Deviations from the background estimation, if unexplained by systematic uncertainties, can be indicative of new physics processes and the compatibility of those observations has to be checked with new signal hypotheses. In the absence of deviations, on the other hand, confidence limits can be used to set constraints on the theoretical models.

## 4.2 Decay Chain and Final State

This thesis project started in 2013, two years in advance of the LHC Run 2. By the time, we defined a general strategy targeting the search for heavy resonances decaying into pairs of boosted Z bosons. Previous CMS analyses performed during the Run 1 [31] suggested an interesting excess around 1.8 TeV in the invariant mass of the diboson pair (Fig. 4.1), so there were some expectations in this particular topology.

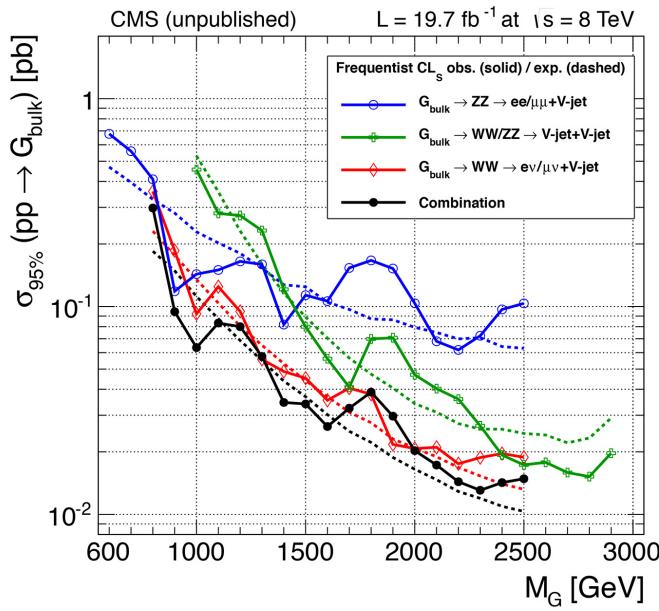


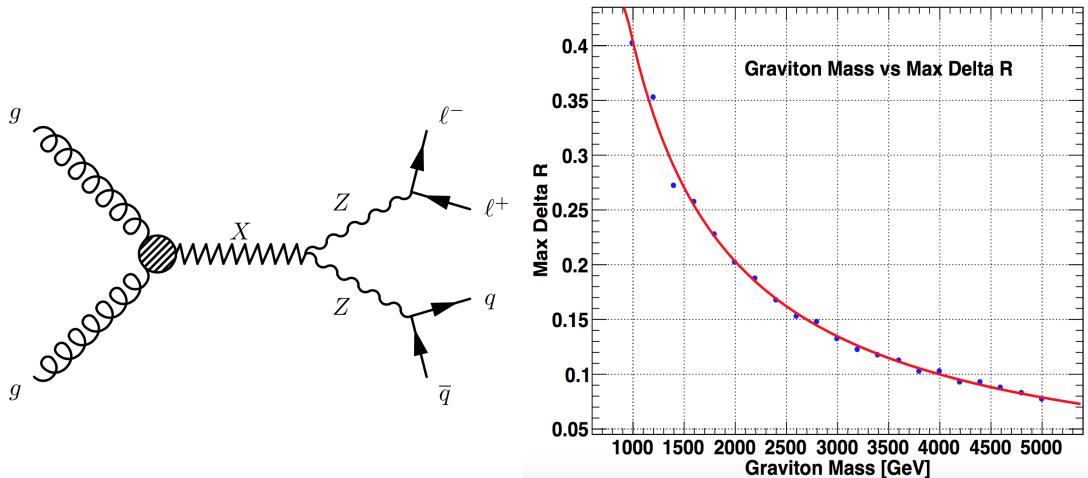
Figure 4.1: Confidence limits on the bulk graviton production cross section obtained by previous CMS analyses.

Table 4.1 shows all possible branching ratios of the ZZ final state. The leptons plus jet channel ( $ZZ \rightarrow 2\ell 2q$ ) presents a good balance in between high event yield with a 14% branching ratio\* and ease of triggering and selection of the signal events. We chose the bulk graviton model as our benchmark due to its large branching ratio into vector bosons, and particularly on the ZZ channel. Figure 4.2 shows the diagram of the signal process, involving the gluon fusion production of the bulk graviton decaying through pair of Z bosons to a  $\ell\ell qq$  state. The products coming from the decay of a boosted Z boson are expected to be close to each other given the high transverse momentum of the parent.

\*This branching ratio refers to decays to the three lepton flavours.

Table 4.1: Branching ratios of the ZZ decay channel into different final states.

Channel	Branching ratio
ZZ $\rightarrow$ 4q	0.49
ZZ $\rightarrow$ 2q2 $\nu$	0.28
ZZ $\rightarrow$ 2 $\ell$ 2q	0.14
ZZ $\rightarrow$ 2 $\ell$ 2 $\nu$	0.04
ZZ $\rightarrow$ 4 $\nu$	0.04
ZZ $\rightarrow$ 4 $\ell$	0.01

Figure 4.2: Production and decay of a resonance  $X$  in the ZZ semi-leptonic channel (left), and separation between the Z boson decay products as function of the graviton mass (right).

The separation between the Z boson decay products, measured in term of  $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ , is shown in Fig. 4.2 as function of the graviton mass. For high mass of the bulk graviton, the hadronization of two boosted quarks coming from one of the Z's will be reconstructed as a single jet. The pair of leptons that come from the decay of the other Z boson are also expected to be close to each other. The final state of interest is therefore a dilepton plus a merged jet.

## 4.3 Analysis Strategy

In order to confront the results obtained in the Run 1, the Run 2 version of the analysis should have used the very same software and analysis framework. However, the different accelerator conditions ( $8 \rightarrow 13$  TeV) and the incorporation of a new CMS data format (AOD  $\rightarrow$  MiniAOD) forced us to develop original analysis tools, giving rise to the SPRACE analysis framework.

The SPRACE framework handles CMS real data in MiniAOD format, as well as data from Monte Carlo simulations. Both real data and simulation need to be split into signal and background. For simulation, the separation of the signal from the background is straightforward because the events are labelled accordingly.

The separation of the signal in real data is a challenging problem that requires dedicated selection criteria in order to enhance the final state, namely, two opposite-sign same-flavour leptons plus one jet. Generally speaking, the analysis strategy follows a series of stages:

1. Get data, both real and simulated samples;
2. Classify into muon and electron channels;
3. Classify into high purity and low purity categories;
4. Classify into signal and control regions;
5. Blind signal region;
6. Estimate background for each channel, each category, and each region;
7. Calculate expected limits for each channel and category;
8. Unblind signal region;
9. Calculate observed limits for each channel and category;
10. Combine individual limits.

The separation between muon and electron channel is a standard choice among CMS analyses. In our case, the muon channel is more competitive than the electron channel providing more sensitivity limits. It is due to the fact that different components of the detector are used in the reconstruction of the objects, translating into different selection criteria and consequently, different efficiencies.

The split into high purity and low purity categories is another enhancement of the analysis. It takes advantage of the N-subjettiness discriminator introduced in Section 3.4.3 to characterize hadronic V jets. By definition, the high purity category contains a larger fraction of signal events than the low purity category. Despite

being less efficient, the low purity category slightly improves the sensitivity of the final limit because it allows to retain some signal events in the situation when the expected background is close to zero.

The classification into signal and control region intends to reduce the bias in the estimation of the background. In a initial stage of the analysis, the kinematic distributions in the signal region cannot be shown, and only information from the control region is used for background estimation purposes. Only at the end of the analysis the signal region is revealed, and the predictions obtained previously are confronted with the real data.

In the next section we introduce the selection criteria applied to select signal events and reject background. Basic distributions at generator level, demonstrating properties of the bulk graviton model in contrast to the background processes, are also presented.

## 4.4 Event Selection

Semi-leptonic events in the boosted regime, characterized by low  $\Delta R$  between the Z boson decay products, can be selected with the set of criteria presented in Table 4.2. The same selection criteria apply to both real data and simulation, and they are intended to select signal-like events and reject the background. Some justifications for these specific requirements are:

- The High Level Trigger is based on the presence of a lepton to minimize purely hadronic backgrounds. Since the online reconstruction is optimized for speed instead of accuracy, we chose single lepton triggers to maximize the probability of selecting an event, even though we eventually require two reconstructed leptons in the offline analysis.
- Electrons reconstructed in the calorimeter can be faked by QCD jets, requiring a high threshold in the transverse momentum, namely  $p_T > 105$  GeV.
- The offline selection of one of the electrons (muons) at  $p_T$  larger than 115 GeV (50 GeV), ensures events in the plateau of the trigger efficiency. The selection on the other lepton at  $p_T$  larger than 20 GeV is a standard CMS threshold applied to particles coming from electroweak processes.

- The dilepton  $p_T$  threshold at 170 GeV intends to select Z candidates in the boosted region. The mass window of the dilepton pair includes the nominal mass of the Z boson (91.2 GeV), and is wide enough to ensure that lepton energy/momentum scale effects can be essentially neglected.
- The remaining offline jet requirements are intended to select V jets, and reject purely hadronic jets. The invariant mass selection reduces the amount of data to run over. Since we are interested only in graviton masses above 800 GeV, a diboson invariant mass selection above 600 GeV is applied.

The V jet selections were gleaned from the Run 1 analysis. Figure 4.3 shows that the jet mass does indeed have a peak at the value of the Z boson mass, while the  $\tau_2/\tau_1$  ratio shows a clear separation between signal and background. Therefore we define two different classification criteria: one based on the jet substructure

- High purity: jets with  $\tau_2/\tau_1 < 0.45$ ;
- Low purity: jets with  $0.45 < \tau_2/\tau_1 < 0.75$ ;

and one based on the jet mass

- Signal region: jets with  $65 < m_j < 105$  GeV;
- Control region: jets with  $20 < m_j < 65$  GeV or  $135 < m_j < 220$  GeV.

The event selection also includes many requirements on the detector status, data quality, object identification, among others. The application of the full event selection on real data is described in Chapter 5.

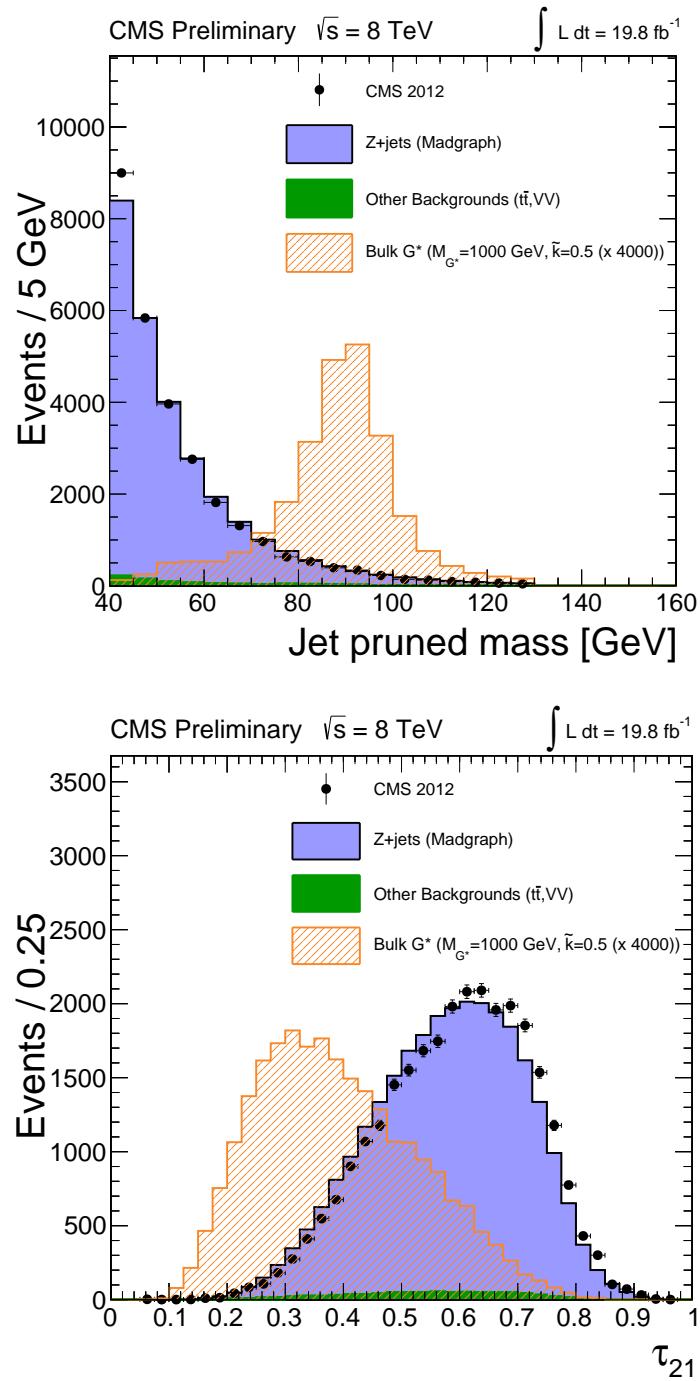


Figure 4.3: Jet pruned mass (top) and  $\tau_2/\tau_1$  ratio (bottom) taken from the Run 1 analysis [65]. The signal region is defined around the peak in the jet mass distribution. The high purity category includes low values of  $\tau_{21}$ , where the signal is concentrated.

Table 4.2: Summary of the event selection.

Selection	Value	Comments
High Level Trigger	$p_T > 105 \text{ GeV}$	
Electron channel	$p_T > 45 \text{ GeV}$ and $ \eta  < 2.1$	
Muon channel		
Electrons		
Electron ID	Loose working point Particle flow isolation	
Isolation	$p_T > 20 \text{ GeV},  \eta  < 2.5$	
Acceptance	$p_T > 115 \text{ GeV},  \eta  < 2.5$	
Dielectron pair seed		leading or second electron
Muons		
Muon ID	High $p_T$ or tracker high $p_T$	
Isolation	Tracker isolation with inner track removal	At least 1 high $p_T$
Acceptance	$p_T > 20 \text{ GeV},  \eta  < 2.4$	
Dimuon pair seed	$p_T > 50 \text{ GeV},  \eta  < 2.1$	
Dilepton selection		leading or second muon
Leptonic Z	$p_T > 170 \text{ GeV}$ $70 < M_{\ell\ell} < 110 \text{ GeV}$	
AK8 Jets		
Jet ID	Loose working point	cleaning w.r.t. good leptons
Acceptance	$p_T > 200 \text{ GeV},  \eta  < 2.4$	
Jet mass (signal region)	$65 < m_j < 105 \text{ GeV}$	
Jet mass (low mass control region)	$20 < m_j < 65 \text{ GeV}$	
Jet mass (high mass control region)	$135 < m_j < 220 \text{ GeV}$	
N-subjettiness (low purity)	$0.45 < \tau_2/\tau_1 < 0.75$	
N-subjettiness (high purity)	$\tau_2/\tau_1 < 0.45$	
Diboson candidate		
Invariant mass	$M_{VZ} > 600 \text{ GeV}$	

## 4.5 Simulated Samples

To optimize and test our selection, we use a set of simulated background and signal samples. For the final state of choice, the standard model processes that constitute the relevant background are Z+jets, diboson (WW + WZ + ZZ), and  $t\bar{t}$ . One advantage of the double lepton channel adopted in this search is the absence of QCD backgrounds due to the fact that leptons are not subject to strong interactions.

For the description of the Z+jets background, a set of samples were generated with the MADGRAPH event generator [66]. The samples are binned according to the variable HT, which is the sum of the transverse energies of partons at matrix-element level. The split in HT bins reduces the computational time specially for high values of this variable, where a re-weighting factor is used to increase the number of available events in the sample. In this way, we can explore this background for high values of the reconstructed diboson invariant mass, since it is correlated with HT. Samples of  $t\bar{t}$  and SM diboson production are generated using POWHEG 2 [67–69] and PYTHIA 8 [70, 71], respectively.

Bulk graviton signal events are generated with MADGRAPH setting  $k/M_{Pl} = 0.5$ ; under this condition, the natural width of the resonance is much smaller than the experimental resolution. Working with narrow resonances is convenient because we can easily extrapolate to lower values of the bulk graviton model parameter, since that would minimally affect the shape of the diboson invariant mass distribution. Parton showering and hadronization processes are simulated by interfacing the event generators to PYTHIA 8 with the CUETP8M1 [72] tune. The NNPDF 3.0 parton distribution functions (PDF) [73] are used to model the momentum distribution of the colliding partons inside the protons. For both signal and background Monte Carlo (MC) samples, events are simulated using a GEANT4-based model [74] of the CMS detector and processed using the same reconstruction algorithms as for real data.

Supplementary minimum bias interactions are added to the generated events in order to match the additional particle production observed in real data from the large number of pileup interactions. In this context, minimum bias events account the inelastic non-diffractive part of the total cross section. Assuming a proton-proton total cross section of  $\sigma_{tot} \sim 100$  mb, the minimum bias component would be close to  $2/3 \sigma_{tot} \sim 70$  mb [75].

Table 4.3: Description of the simulated samples. Z+jets background were generated with the MADGRAPH.  $t\bar{t}$  and SM diboson production are generated using POWHEG 2 and PYTHIA 8, respectively. Bulk graviton signal events are generated with MADGRAPH setting  $k/M_{Pl} = 0.5$ . Detector effects are simulated using a GEANT4-based model.

<b>Sample name</b>	<b>Cross section[pb]</b>	<b>N<sub>events</sub></b>
DYJetsToLL_M-50_HT-100to200	$147.40 \times 1.23$	2655294
DYJetsToLL_M-50_HT-200to400	$40.99 \times 1.23$	962195
DYJetsToLL_M-50_HT-400to600	$5.678 \times 1.23$	1069003
DYJetsToLL_M-50_HT-600toInf	$2.198 \times 1.23$	1031103
WW	118.7	988418
WZ	66.1	1000000
ZZ	15.4	985600
TT	831.76	196937036
BulkGravToZZToZlepZhad_narrow_M-800	41.7E-3	50000
BulkGravToZZToZlepZhad_narrow_M-1000	11.2E-3	50000
BulkGravToZZToZlepZhad_narrow_M-1200	3.74E-3	50000
BulkGravToZZToZlepZhad_narrow_M-1400	1.44E-3	49200
BulkGravToZZToZlepZhad_narrow_M-1600	0.92E-3	50000
BulkGravToZZToZlepZhad_narrow_M-1800	0.76E-3	50000
BulkGravToZZToZlepZhad_narrow_M-2000	0.135E-3	48400
BulkGravToZZToZlepZhad_narrow_M-2500	0.070E-3	50000

## 4.6 Signal Characterization

As we mentioned earlier, the presence of two leptons in the final state reduces considerably the background rate. Moreover, the kinematics of leptons coming from a boosted Z boson, typical signature of signal events, is different than the distribution of regular Z bosons produced by background processes. Figure 4.4 shows a selection of generator level distributions for both signal and background. It can be seen that for signal events both the Z boson and the daughter leptons have, in general, much harder  $p_T$  than in background events.

For modeling the signal invariant mass we use a probability density function consisting of a gaussian peak and a power-law in both tails — double crystal ball — defined below:

$$\begin{aligned} f(x; \alpha_1, \alpha_2, n_1, n_2, \bar{x}, \sigma) &= \exp\left(-\frac{t^2}{2}\right), && \text{for } t > -\alpha_1 \text{ and } t < \alpha_2 \\ &= A_1 \cdot (B_1 - t)^{-n_1}, && \text{for } t \leq -\alpha_1 \\ &= A_2 \cdot (B_2 + t)^{-n_2}, && \text{for } t \geq \alpha_2 \end{aligned}$$

where  $t = \frac{x - \bar{x}}{\sigma}$  and

$$\begin{aligned} A_1 &= \left(\frac{n_1}{|\alpha_1|}\right)^{n_1} \cdot \exp\left(-\frac{\alpha_1^2}{2}\right) , & B_1 &= \frac{n_1}{|\alpha_1|} - |\alpha_1| , \\ A_2 &= \left(\frac{n_2}{|\alpha_2|}\right)^{n_2} \cdot \exp\left(-\frac{\alpha_2^2}{2}\right) , & B_2 &= \frac{n_2}{|\alpha_2|} - |\alpha_2| . \end{aligned}$$

An interpolation procedure is performed to simulate intermediate mass points in order to search for narrow resonances over the continuous background by steps of 100 GeV, starting at 0.8 TeV up to 2.5 TeV. The signal shapes obtained after the interpolation procedure are shown in Fig. 4.5. At generator level the signal invariant mass would look like a narrow delta function so the width of the double crystal ball resembles the detector resolution; therefore, for higher values of invariant mass the detector resolution is systematically worse.

The sensitivity of the analysis to the presence of a bulk graviton signal directly depends on the selection efficiency, and in turn, the efficiency depends on the channel and category under consideration. As shown in Fig. 4.6, the selection efficiency of the muon channel is higher than the electron channel, and the high purity category outperforms the low purity category.

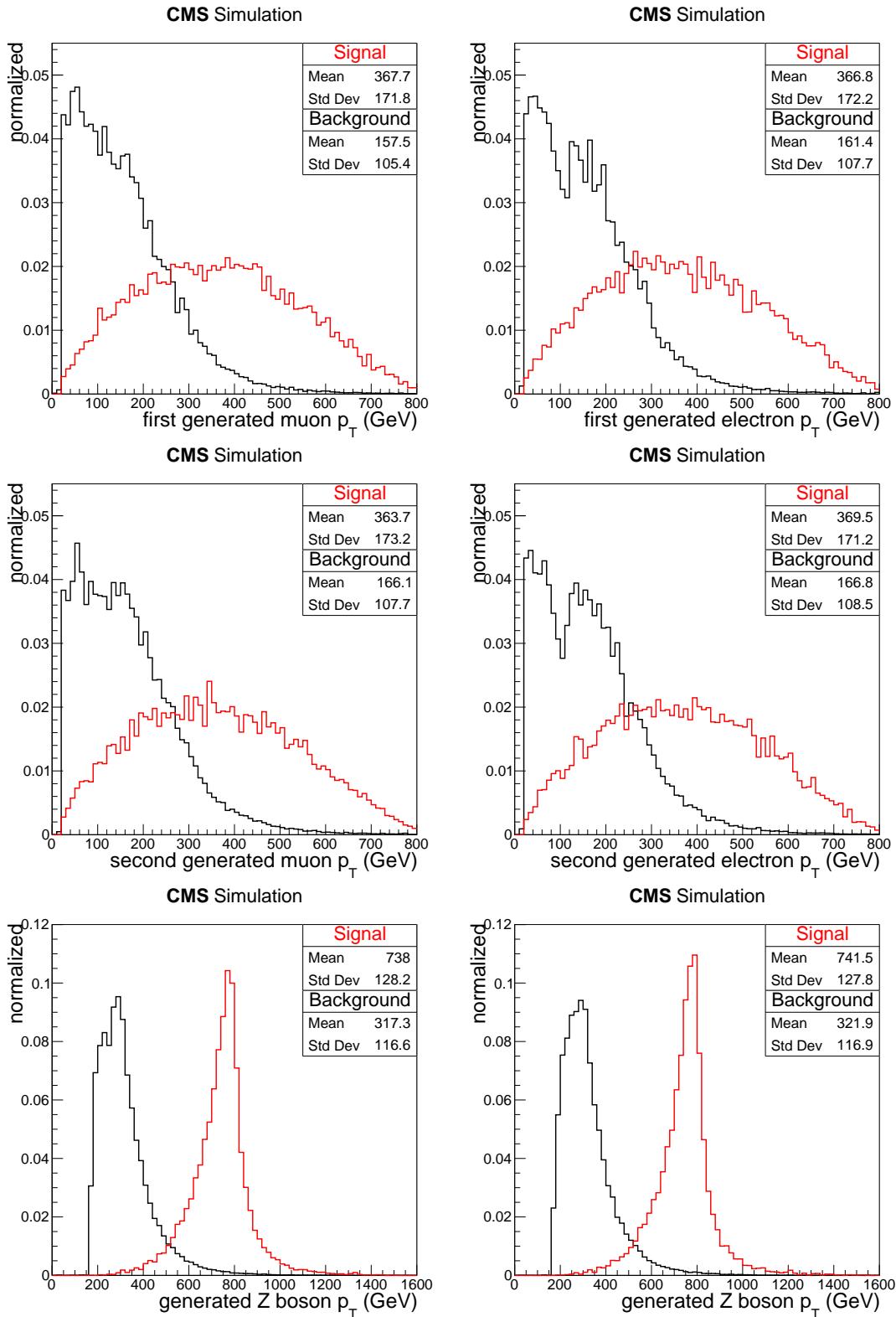


Figure 4.4: Z boson transverse momentum at generator level and leptons kinematics in the muon (left) and electron (right) channels. The signal corresponds to a bulk graviton of mass 1.6 TeV; the background corresponds to standard model processes described in Table 4.3. All distributions are normalized to unity.

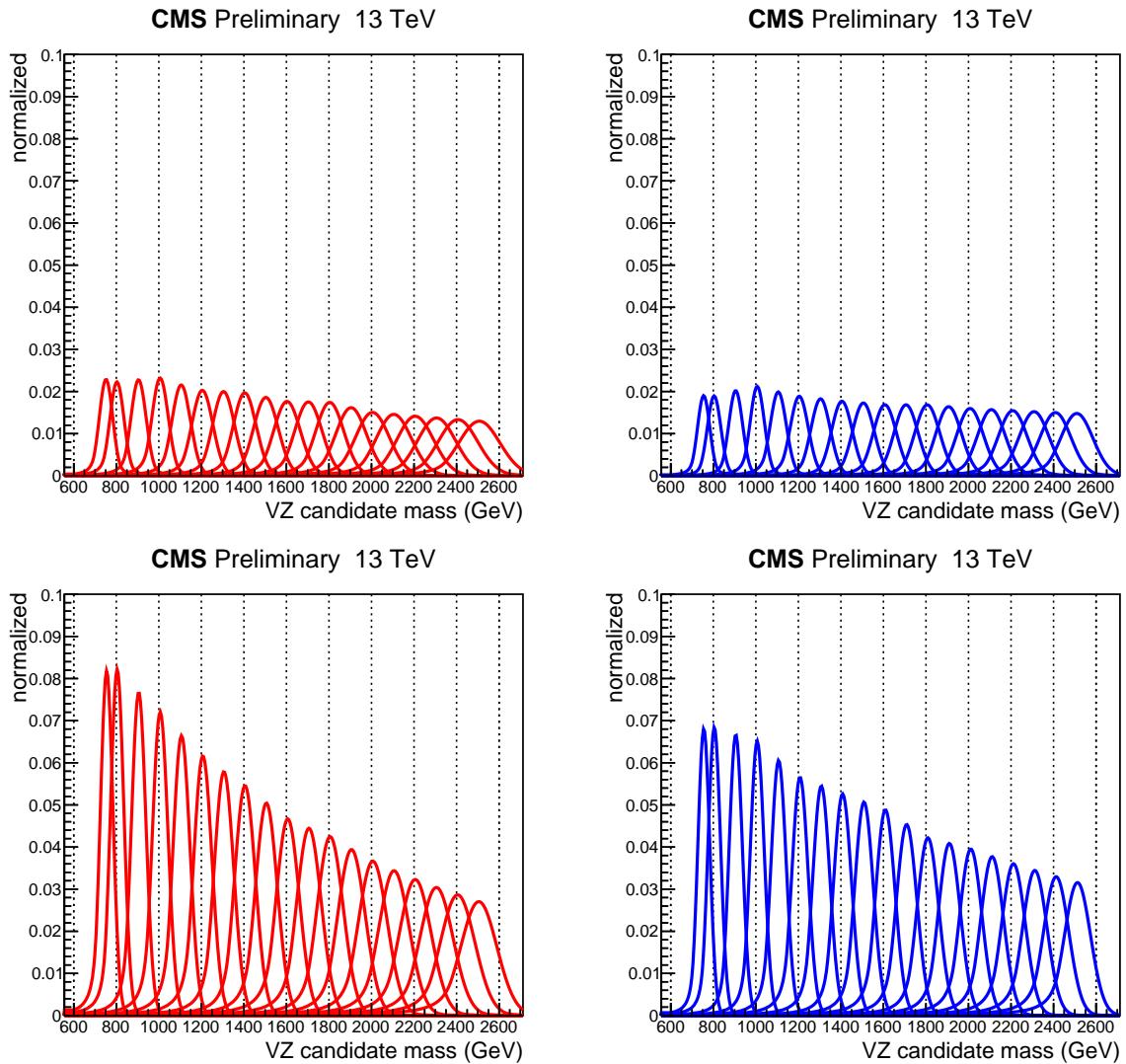


Figure 4.5: Signal shapes for the bulk graviton in the categories: muon low purity (top-left), electron low purity (top-right), muon high purity (bottom-left), and electron high purity (bottom-right).

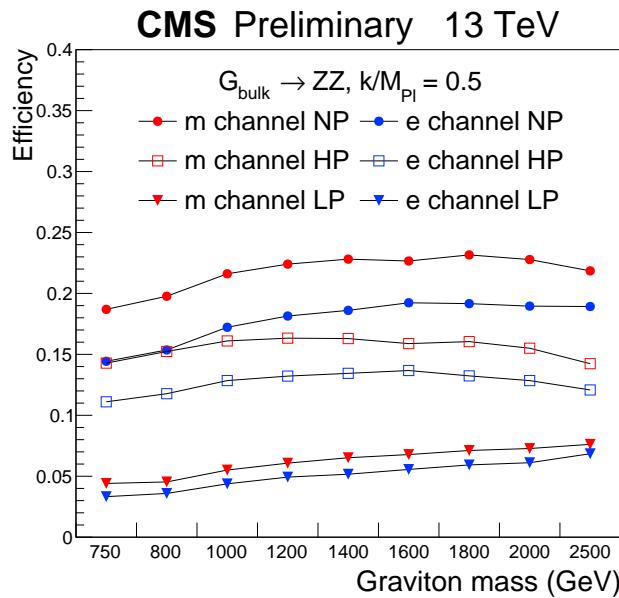


Figure 4.6: Selection efficiency as function of the bulk graviton mass. Muon (electron) channel is shown with red (blue) markers. The notation in the legend stands for LP=Low Purity, HP=High Purity, and NP=No Purity selection. The denominator entering in the calculation of the efficiencies corresponds to ZZ semileptonic decays into the three lepton flavours.

A precise description of the analysis cut-flow in the muon and electron channels is provided in the Tables 4.4 and 4.5, respectively. In the tables, the step named dimuon (dielectron) pair involves the dimuon (dielectron) pair seed plus the dilepton selection defined in the Table 4.2; these two requirements explain the large event drop observed in the electron channel, in comparison to the muon channel.

Besides serving as input for the calculation of the signal efficiency, the numbers in the cut-flow tables were important to cross-check the correct performance of the SPRACE analysis framework. The cross-check consisted in a synchronization with other analysis groups, one at the University of Virginia and other at the University of Taiwan. They independently produced identical cut-flow tables using their own analysis frameworks, and the results of the three groups —Virginia, Taiwan, and SPRACE— perfectly matched.

Table 4.4: Cut-flow table for different graviton mass points in the muon channel.

	<b>M-800</b>	<b>M-1000</b>	<b>M-1200</b>	<b>M-1400</b>	<b>M-1600</b>	<b>M-1800</b>	<b>M-2000</b>	<b>M-2500</b>
All events	50000	50000	49200	50000	50000	48400	50000	48400
Muon events	16384	16268	16402	16227	16315	16534	15874	16237
Muon trigger	15891	15910	16084	15928	15971	16141	15457	15619
Vertex	15891	15910	16084	15928	15971	16141	15457	15619
Muon acceptance	15891	15910	16084	15928	15971	16141	15457	15619
Muon ID	15876	15899	16075	15920	15960	16122	15440	15594
Dimuon pair	14954	15176	15424	15228	15284	15483	14784	14824
Opposite sign	14925	15156	15395	15204	15257	15437	14748	14779
Muon isolation	14437	14689	14950	14800	14834	15022	14364	14397
Z mass window	13667	14028	14289	14155	14212	14386	13773	13723
Jet ID	13449	13836	14103	13974	14014	14170	13610	13530
Jet Cleaning	13373	13800	14092	13970	14014	14170	13610	13529
N-subjettiness	12879	13549	13882	13803	13863	14026	13455	13367
Jet Kinematics	12366	13040	13322	13227	13244	13370	12812	12710
Invariant mass cut	12298	13012	13301	13214	13233	13361	12805	12705

Table 4.5: Cut-flow table for different graviton mass points in the electron channel.

	<b>M-800</b>	<b>M-1000</b>	<b>M-1200</b>	<b>M-1400</b>	<b>M-1600</b>	<b>M-1800</b>	<b>M-2000</b>	<b>M-2500</b>
All events	50000	50000	49200	50000	50000	48400	50000	50000
Electron events	16484	16572	16536	16094	16611	16455	16037	16562
Electron trigger	15439	15897	15996	15650	16204	16063	15678	16137
Vertex	15439	15897	15996	15650	16204	16063	15678	16137
Electron acceptance	15385	15853	15958	15620	16167	16022	15641	16077
Electron ID+Iso.	15044	15433	15304	14716	14987	14646	14162	14575
Dielectron pair	11056	11694	11985	12005	12425	12368	12008	12355
Opposite sign	10880	11500	11792	11801	12206	12152	11793	12054
Z mass window	10600	11240	11517	11562	11954	11889	11558	11805
Jet ID	10428	11079	11370	11386	11787	11733	11401	11640
Jet Cleaning	10365	11066	11363	11383	11786	11733	11399	11640
N-subjettiness	9957	10839	11208	11239	11641	11612	11263	11498
Jet Kinematics	9565	10446	10809	10730	11121	11087	10681	10938
Invariant mass cut	9536	10426	10788	10723	11112	11079	10672	10933

# Chapter 5

## Real Data from pp Collisions

### 5.1 Primary Datasets

The Run 2 of the LHC started in 2015 after a long shutdown of two years, colliding protons at a center-of-mass energy of 13 TeV. After resuming operations, the CMS experiment collected proton-proton collision data equivalent to an integrated luminosity of  $3.8 \text{ fb}^{-1}$ . The fraction of certified data with good magnetic field and without HF — Hadronic Forward detector — was  $2.7 \text{ fb}^{-1}$ . Our analysis only uses central objects in the range  $|\eta| < 2.5$ , then we are allowed to use the whole set of certified data.

We would like to remind the problems faced by the cryogenic system of the CMS magnet which restricted the amount of data collected with the standard 3.8 T field configuration. Despite those unexpected developments, we did our best to incorporate in this thesis the 2016 data but those efforts did not converge in time and, therefore, only results obtained with the 2015 data were included.

CMS primary datasets, the Analysis Object Data (AOD) files, contains the necessary information for the offline analysis. In 2015 a new format called MiniAOD [76] was created, which includes the following information:

- High-level physics objects such as muons, electrons, photons, taus, jets, missing energy;
- Monte Carlo information for simulated samples;
- Primary vertex collection;
- Trigger decision bits.

A feature of the MiniAOD format is the absence of the track collection, which

was dropped purposely in order to keep a small event size. Table 5.1 contains the name of the datasets used in this analysis. SingleElectron (SingleMuon) corresponds to the dataset filled with a stream of data triggered by electron (muon) objects. The second part of the name refers to the era (Run2015 C, D), the separation between bunches (25 ns), and the date of the reconstruction campaign (16Dec2015). The other columns in the table contain the integrated luminosity as well as the number of events. The specific version of the CMS software deployed during the data taken was 7.4.8\_patch1 for Run2015 C, and 7.4.15 for Run2015 D. The last re-reconstruction happened in December 2015, using CMS software version 7.6.3.

Table 5.1: Description of the real data samples.

Dataset Name	$\mathcal{L}$ ( $\text{pb}^{-1}$ )	Events
/SingleElectron/Run2015C_25ns-16Dec2015	17.7	837 k
/SingleElectron/Run2015D-16Dec2015	2672.8	134 M
<hr/>		
/SingleMuon/Run2015C_25ns-16Dec2015	17.7	1.34 M
/SingleMuon/Run2015D-16Dec2015	2672.8	92 M
<hr/>		

Events are recorded by the online selection algorithm requiring a single electron or a single muon. A triggering electron must have  $p_T > 105 \text{ GeV}$ , while a triggering muon must have  $p_T > 45 \text{ GeV}$  and  $|\eta| < 2.1$ . We should recall that the trigger is also simulated in Monte Carlo samples as was discussed in the previous chapter.

The performance of the trigger is analyzed by observing the turn-on curve shown in Fig. 5.1. The numerator in the efficiency calculation corresponds to events passing the trigger, and the denominator includes the full event sample. The turn-on curve is modeled with a sigmoid function because the variable in the x-axis is the offline  $p_T$  rather than the online  $p_T$ ; if the offline  $p_T$  were exactly the same as the online  $p_T$ , the trigger turn-on would be a step function.

By convenience, the offline selection is chosen to accept events in the plateau of the trigger, which is the kinematic region where the efficiency is practically 100%. The simulated samples are corrected for observed differences with the real data

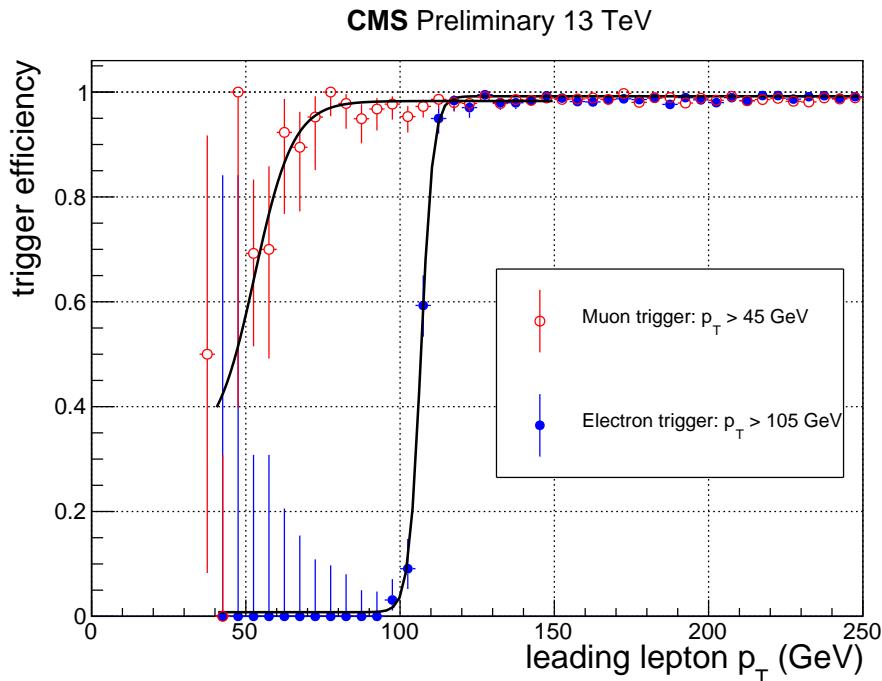


Figure 5.1: Simulation of the trigger efficiency as function of the leading lepton  $p_T$ .

concerning trigger efficiencies; those trigger scale factors are close to unity, and the impact of this correction is accounted in the final result as a systematic uncertainty.

## 5.2 Kinematic Distributions in the Control Region

According to the analysis strategy established in the previous chapter, we initially split the data into signal and control regions with the aim of blinding the signal and restricting the offline analysis to the control regions. For this reason, the remaining part of this chapter will contain distributions excluding the signal region.

The offline analysis starts by selecting events with at least one reconstructed collision vertex within the distance of 24 cm along the beam axis and 2 cm in the transverse plane to the beams from the nominal  $pp$  interaction point. The vertex requirement is standard in CMS analyses [77], and it is fundamental to identify high level objects. For instance, the identification of charged leptons involves a requirement on the impact parameter between the lepton's track and the primary vertex, in order to veto leptons from cosmic rays.

The number of primary vertices per event is a wide distribution with mean

value  $\sim 10$ , as shown in Fig. 5.2. We may recall that in Monte Carlo (MC) samples the pileup is simulated by superimposing minimum bias events on top of the hard process; the inaccuracy of the pileup simulation is therefore compensated with a reweighting procedure. The pileup weights are obtained by dividing the distribution of the true number of interactions in data and MC, and then, these weights are applied to the MC distributions.

In Fig. 5.2, four categories are shown:

- muon low purity (top-left);
- electron low purity (top-right);
- muon high purity (bottom-left);
- electron high purity (bottom-right).

The MC distribution is a stacked histogram with different background components, namely,  $Z+jets$ , diboson, and  $t\bar{t}$ ; the statistical uncertainty due to the limited amount of simulated events is also displayed. From the figure, we observe that  $Z+jets$  is the dominant component with a contribution larger than 95%, while the remaining backgrounds — diboson and  $t\bar{t}$  — correspond to the subdominant component. Moreover, the real data in the figure correspond to the points with the associated error bars. Additionally, the small panel below every distribution contains the Data/MC ratio, which is summarized in the Table 5.2.

Table 5.2: Events yields and Data/MC ratios in the control region.

Category	Data	MC	Data/MC
Muon Low Purity	398	440	0.90
Muon High Purity	68	67	1.01
Electron Low Purity	271	335	0.81
Electron High Purity	44	53	0.82

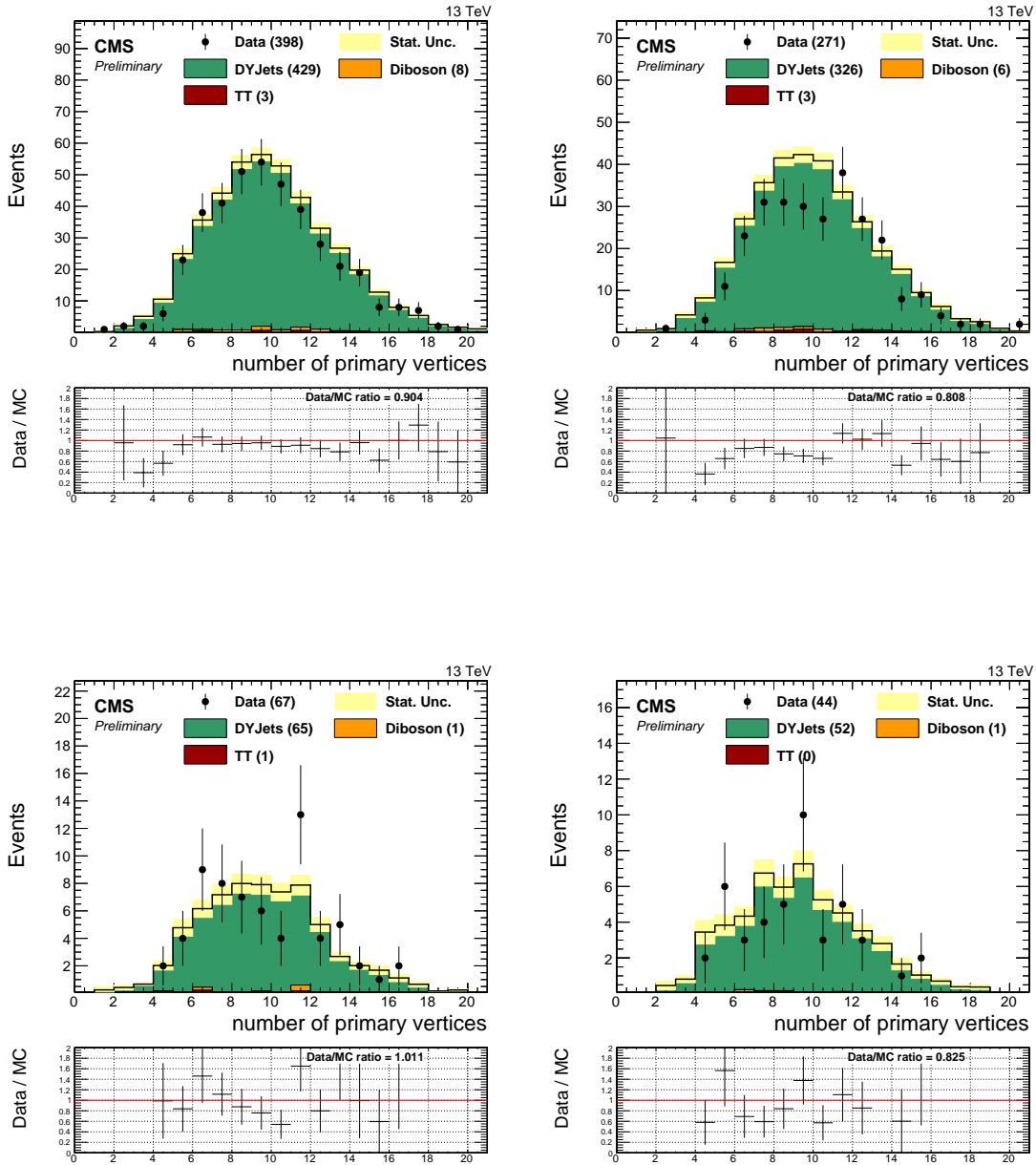


Figure 5.2: Distribution of number of primary vertices. Four categories are shown: muon low purity (top-left), electron low purity (top-right), muon high purity (bottom-left), and electron high purity (bottom-right).

### 5.2.1 Leptonic Z Selection

Leptonic Z candidates are selected from a pair of same-flavour and opposite-charge leptons  $\ell^+\ell^-$ , where  $\ell = e, \mu$ . The analysis distinguishes electron and muon categories: Z candidates formed by electrons (muons) belong to the electron (muon) channel. Leptons originated from the decay of high  $p_T$  bosons are expected to be closer to each other with small  $\Delta R(\ell^+\ell^-)$ . In this boosted regime, one lepton may spoil the isolation of the other due to the coarse detector resolution in the muon chambers and the calorimeters. For smaller values of  $\Delta R(\ell^+\ell^-)$  a drop in efficiency may occur, unless we apply dedicated identification and isolation requirements. The identification criteria for electrons and muons that optimize the efficiency in the boosted regime are described below.

## Electron Identification

The electron identification is chosen from the recommendations of the EGamma physics object group [78]. We use particle-flow electron candidates with loose requirements defined below:

- Pseudo rapidity  $|\eta|$  of the electron's supercluster:  $< 1.479$  (Barrel), and between 1.479 and 2.5 (Endcap);
- Difference in  $\eta$  between the track position as measured in the inner layer, extrapolated to the interaction vertex and then extrapolated to the calorimeter and the  $\eta$  of the seed cluster of the supercluster  $< 0.0095$  (Barrel), 0.010 (Endcap);
- Difference in  $\phi$  between the track position as measured in the inner layer, extrapolated to the interaction vertex and then extrapolated to the calorimeter and the  $\phi$  of the seed cluster of the supercluster  $< 0.18$  (Barrel), 0.11 (Endcap);
- Hadronic energy over electromagnetic energy  $< 0.082$  (Barrel), 0.10 (Endcap);
- Relative isolation in a cone of aperture  $R = 0.3$  with effective area correction  $< 0.118$  (Barrel), 0.089 (Endcap);
- Number of inner tracker layers lost hits  $< 1$  (Barrel), 2 (Endcap).

The efficiency of this working point is evaluated in Fig. 5.3 using simulated samples of the bulk graviton signal. The mass of the graviton ranges between 0.8 to 2.5 TeV, and all mass points have been added to produce a result with good

statistics. The efficiency is found to be around 84% for  $\Delta R$  separations between 0.15 and 0.5. We studied the impact of the isolation which is embedded by default in the electron identification working point, getting an efficiency around 90% when the isolation is removed. For close-by electrons ( $\Delta R < 0.08$ ) the efficiency drops to 65%.

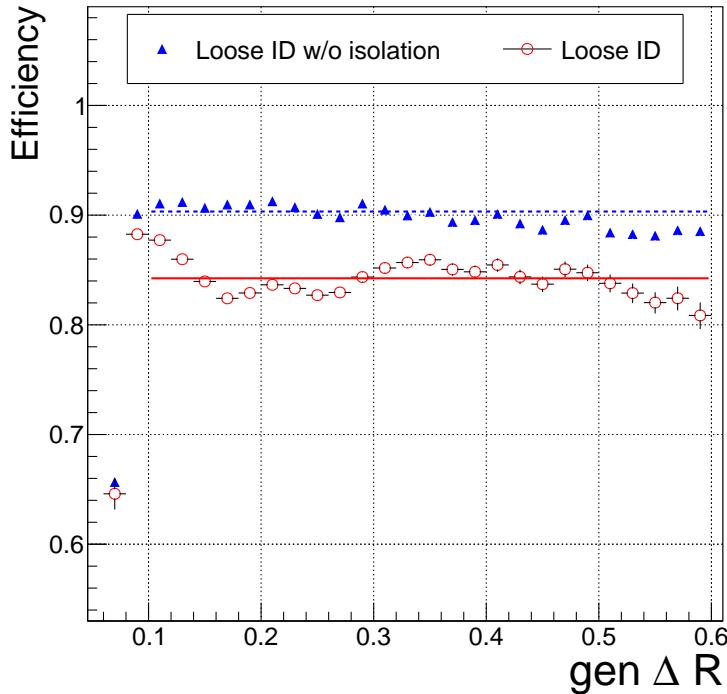


Figure 5.3: Electron identification efficiency for the loose working point (red open markers) as function of the  $\Delta R$  separation between the electrons.

On top of the identification working point we apply additional requirements such that the leading (subleading) electron satisfies  $p_T \geq 115$  (50) GeV and  $|\eta| < 2.4$ ; the pseudorapidity requirement ensures the electrons are inside the geometric acceptance of the detector. The  $p_T$  requirements in the offline selection are tighter than the online trigger, to avoid inefficiencies occurring at values lower than the trigger threshold.

Additionally, the Monte Carlo events are corrected with isolation and identification scale factors that improve the agreement with real data. The computation of the scale factors is done directly by the EGamma physics object group, and the systematic uncertainty associated with this correction was carefully evaluated in our analysis as will be explained in Section 6.3.

## Muon Identification

The muon identification and their momentum assignment follow the recommendation of the muon physics object group [79]. The high  $p_T$  working point is used for the leading candidate, specifically:

- Muon identified as global muon;
- At least one muon chamber hit included in the global-muon track fit;
- Muon segments in at least two muon stations;
- Tracker track transverse impact parameter  $d_{xy} < 2$  mm;
- Longitudinal impact parameter  $d_z < 5$  mm;
- Number of pixel hits larger than zero;
- Number of tracker layers with hits  $> 5$ ;
- Muon track  $dp_T/p_T < 0.3$ .

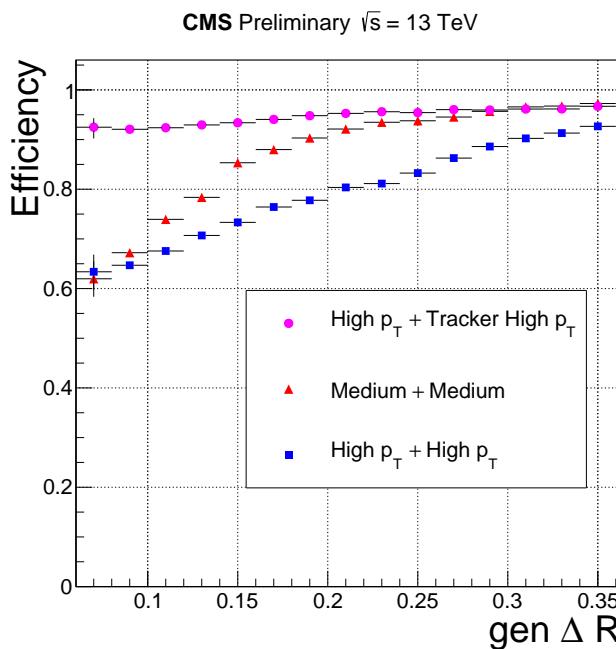


Figure 5.4: Muon identification efficiency as function of the  $\Delta R$  separation. The tracker high  $p_T$  identification recovers the inefficiencies in the case of close-by muons.

The tracker high  $p_T$  is a customized working point used for the sub-leading muon. It was developed to overcome the inefficiencies observed in the case of close-by muons. The tracker high  $p_T$  is essentially the high  $p_T$  working point without the global muon requirement; instead a tracker muon is selected. Fig. 5.4

shows the muon identification efficiency for the bulk graviton signal as function of the  $\Delta R$  separation, comparing the performance of different working points.

On top of the identification working point, muon candidates are required to have relative tracker isolation less than 0.1. For close-by muons ( $\Delta R < 0.3$ ) the  $p_T$  of the inner track of one muon is removed from the isolation cone of the other, to avoid efficiency drops in the boosted regime. Lastly, the leading (sub-leading) muon also satisfies  $p_T \geq 50$  (20) GeV and  $|\eta| < 2.1$  (2.4), in order to avoid inefficiencies occurring at values lower than the trigger threshold.

## Lepton Kinematics

Distributions of the  $p_T$  for the leading and sub-leading lepton are shown in Fig. 5.5 and 5.6, respectively. Similarly, distributions of  $\eta$  for the leading and sub-leading lepton are shown in Fig. 5.7 and 5.8, respectively.

In addition to the lepton identification requirements, dilepton candidates must have specific values of invariant mass  $m_{\ell\ell}$  and transverse momentum  $p_T^{\ell\ell}$ . Namely,  $p_T^{\ell\ell} > 170$  GeV and  $70 < m_{\ell\ell} < 110$  GeV. These requirements intend to select Z candidates in the boosted region of the phase space; when more than one candidate is found, the highest  $p_T$  candidate is chosen.

Distributions of  $p_T$  and mass of the leptonic Z candidate are shown in Fig. 5.9 and 5.10, respectively. Similarly, the  $\Delta R$  separation between the daughters of the leptonic Z candidate is shown in Fig. 5.11.

### 5.2.2 Hadronic Jet Selection

The recommendations from the JetMET physics object group [80] regarding jet identification include the following requirements:

- Neutral Hadron Fraction  $< 0.99$ ;
- Electromagnetic (EM) Fraction  $< 0.99$ ;
- Number of Constituents  $> 1$ ;
- Charged Hadron Fraction  $> 0$ ;
- Charged Multiplicity  $> 0$ ;

The jet identification is a loose requirement intended to reduce calorimeter noise. More important is the acceptance selection  $p_T > 200$  GeV, and  $|\eta| < 2.4$ . Besides

the jet substructure techniques described in Sec. 3.4.3, no further requirements are applied, *e.g.* b tagging \*.

Figures 5.12 and 5.13 show the distributions of jet  $p_T$  and jet  $\eta$  respectively. The  $\Delta R$  separation between the jet and the closest lepton is shown in 5.14. Figure 5.15 shows the distributions of the jet mass containing the gap in the signal region. The n-subjettiness variable  $\tau_{21}$  is shown in Fig. 5.16. Finally, the distribution of the diboson invariant mass is shown in Fig. 5.17. We recall that these distributions are blind, and the yields displayed in the legend of the plots account only the number of events in the control region.

From the analysis of the control region we conclude that there are remarkable discrepancies between the real data and the Monte Carlo simulation, concerning the normalization and the shape of several kinematic distributions. In Chapter 6, we will try a data-driven technique to estimate the invariant mass in the signal region relying on a interpolation of the data from the control regions. The technique known as alpha method [81] will be introduced, and we will obtain final limits based on the results of the data-driven estimation rather than the Monte Carlo simulations.

---

\*The combine secondary vertex for b tagging is usually applied in the low energy regime when the merged (AK8) jet can be resolved into two (AK4) jets.

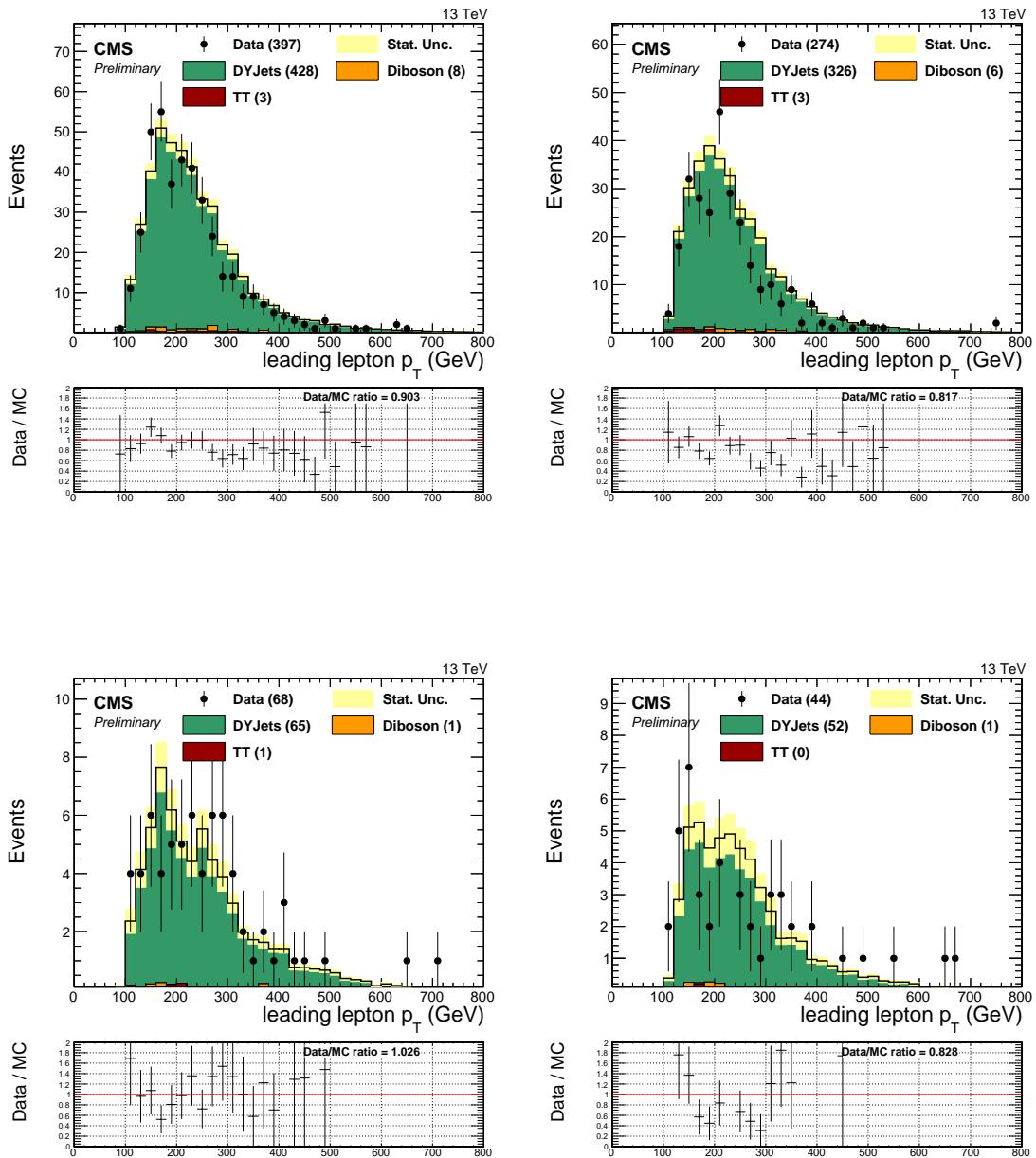


Figure 5.5: Distribution of  $p_T$  for the leading lepton in the category: muon low purity (top-left), electron low purity (top-right), muon high purity (bottom-left), and electron high purity (bottom-right). Simulated backgrounds are displayed as stacked histograms normalized to luminosity ( $2.7 \text{ fb}^{-1}$ ).

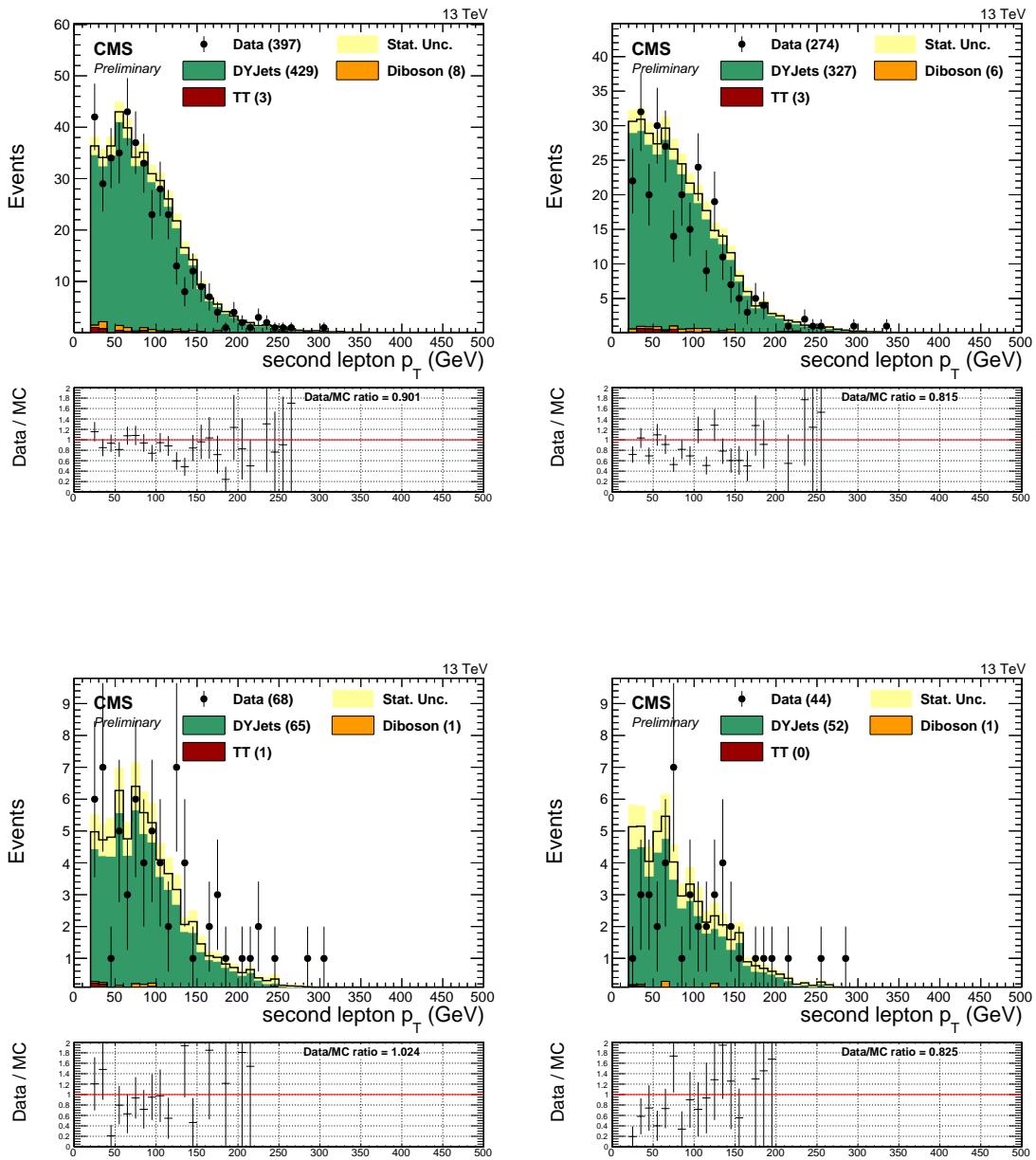


Figure 5.6: Distribution of  $p_T$  for the second lepton in the category: muon low purity (top-left), electron low purity (top-right), muon high purity (bottom-left), and electron high purity (bottom-right). Simulated backgrounds are displayed as stacked histograms normalized to luminosity ( $2.7 \text{ fb}^{-1}$ ).

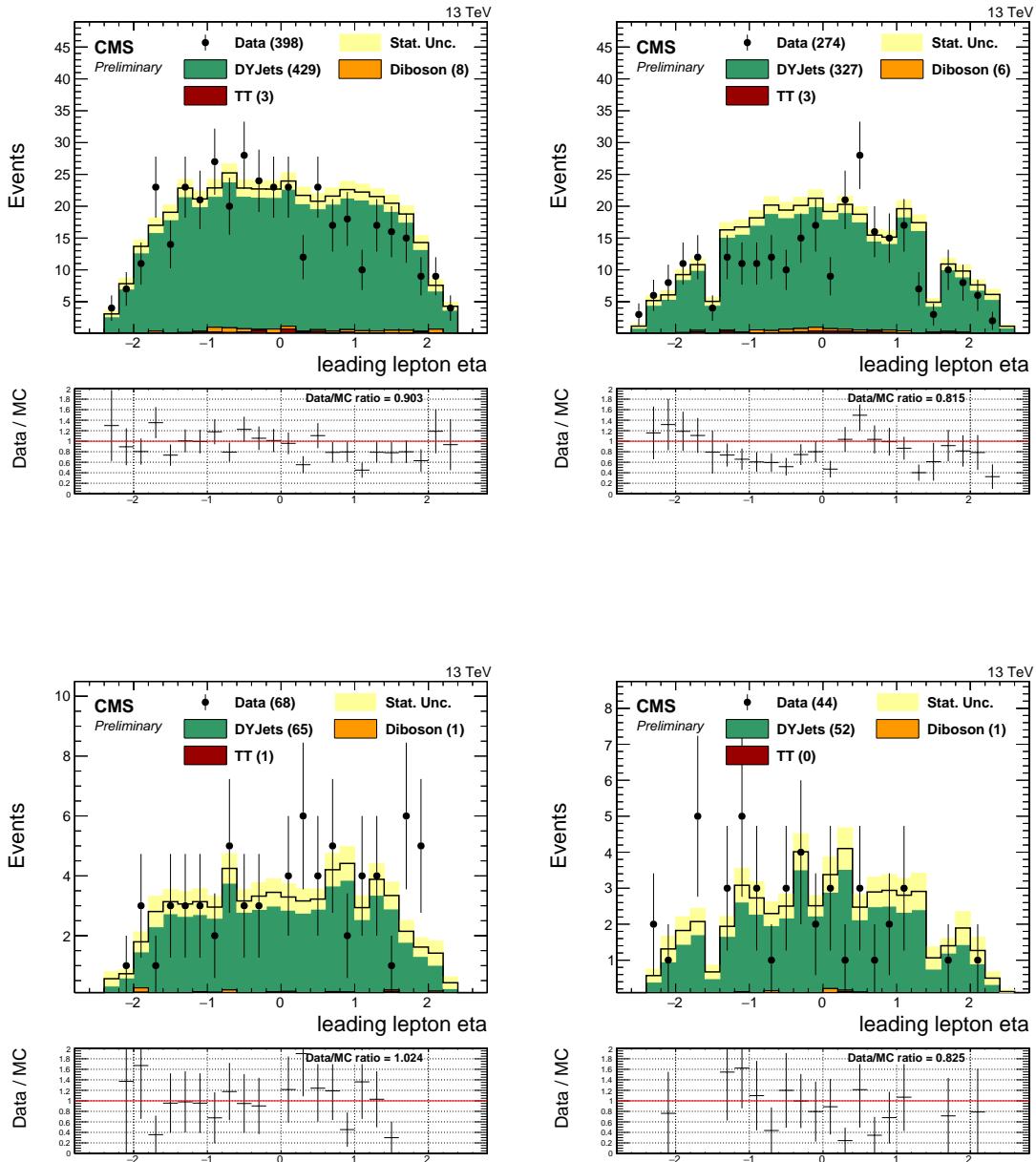


Figure 5.7: Distribution of  $\eta$  for the leading lepton in the category: muon low purity (top-left), electron low purity (top-right), muon high purity (bottom-left), and electron high purity (bottom-right). Simulated backgrounds are displayed as stacked histograms normalized to luminosity ( $2.7 \text{ fb}^{-1}$ ).

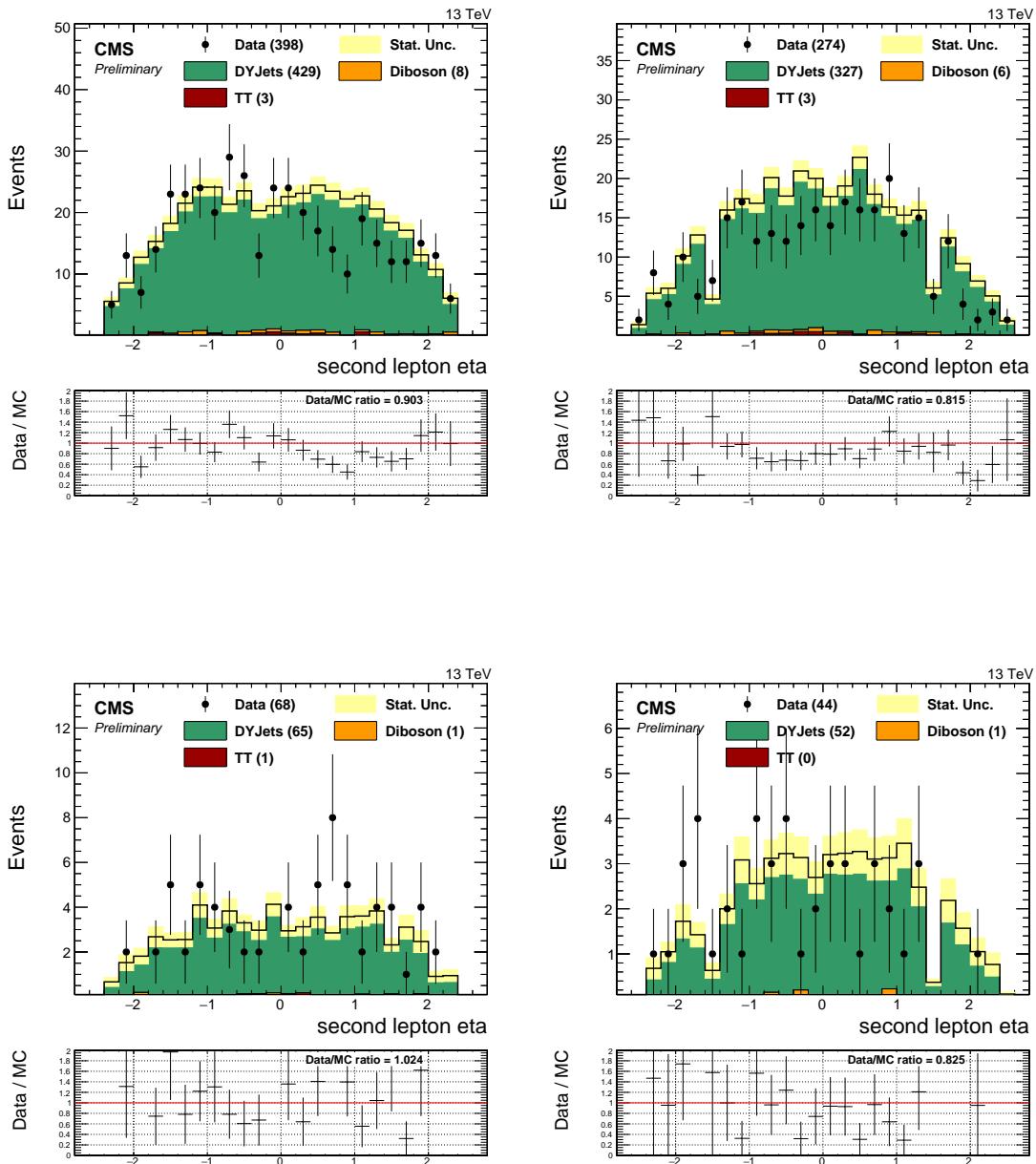


Figure 5.8: Distribution of  $\eta$  for the second lepton in the category: muon low purity (top-left), electron low purity (top-right), muon high purity (bottom-left), and electron high purity (bottom-right). Simulated backgrounds are displayed as stacked histograms normalized to luminosity ( $2.7 \text{ fb}^{-1}$ ).

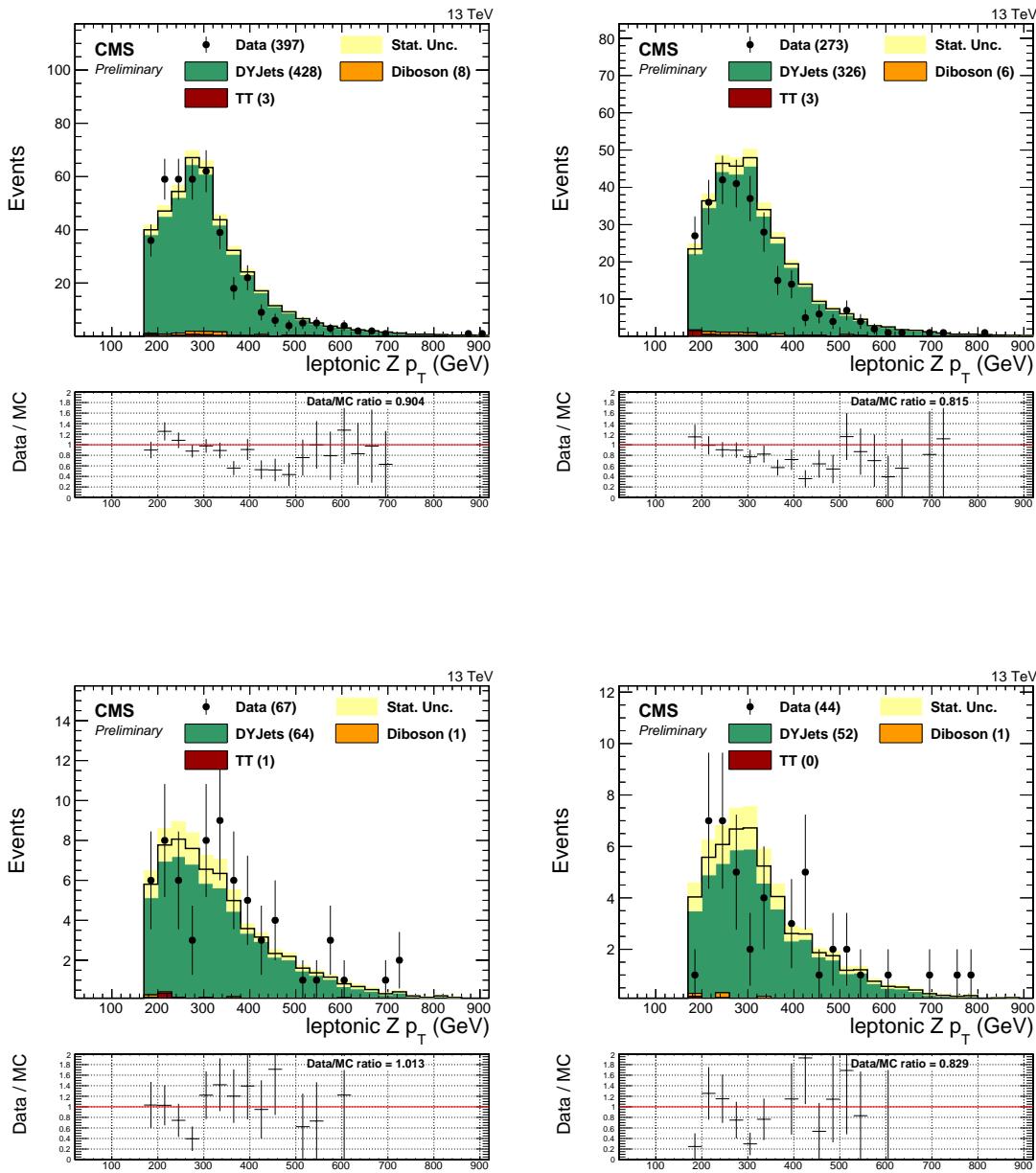


Figure 5.9: Distribution of  $p_T$  of the leptonic Z candidate in the category: muon low purity (top-left), electron low purity (top-right), muon high purity (bottom-left), and electron high purity (bottom-right). Simulated backgrounds are displayed as stacked histograms normalized to luminosity ( $2.7 \text{ fb}^{-1}$ ).

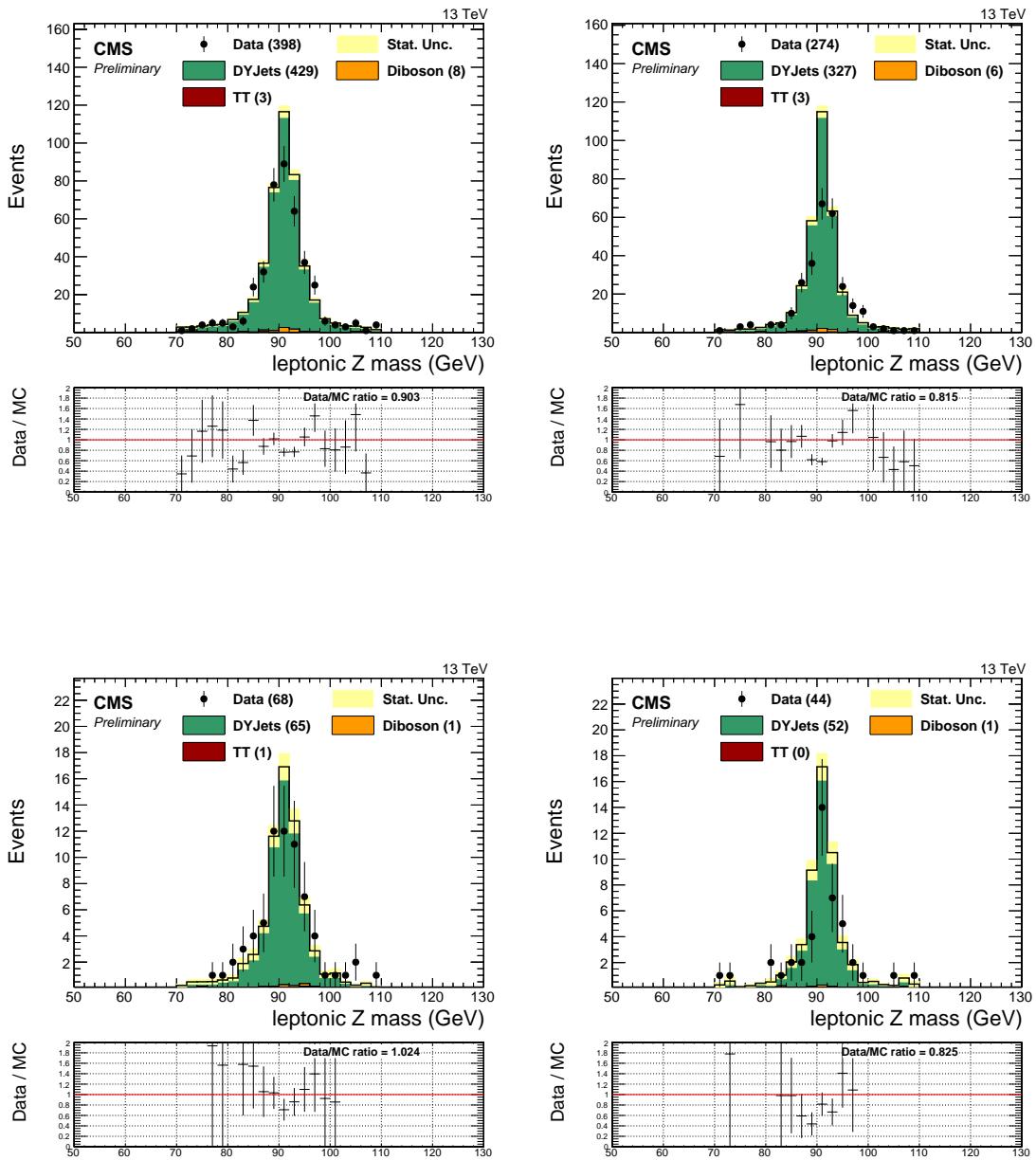


Figure 5.10: Distribution of mass of the leptonic Z candidate in the category: muon low purity (top-left), electron low purity (top-right), muon high purity (bottom-left), and electron high purity (bottom-right). Simulated backgrounds are displayed as stacked histograms normalized to luminosity ( $2.7 \text{ fb}^{-1}$ ).

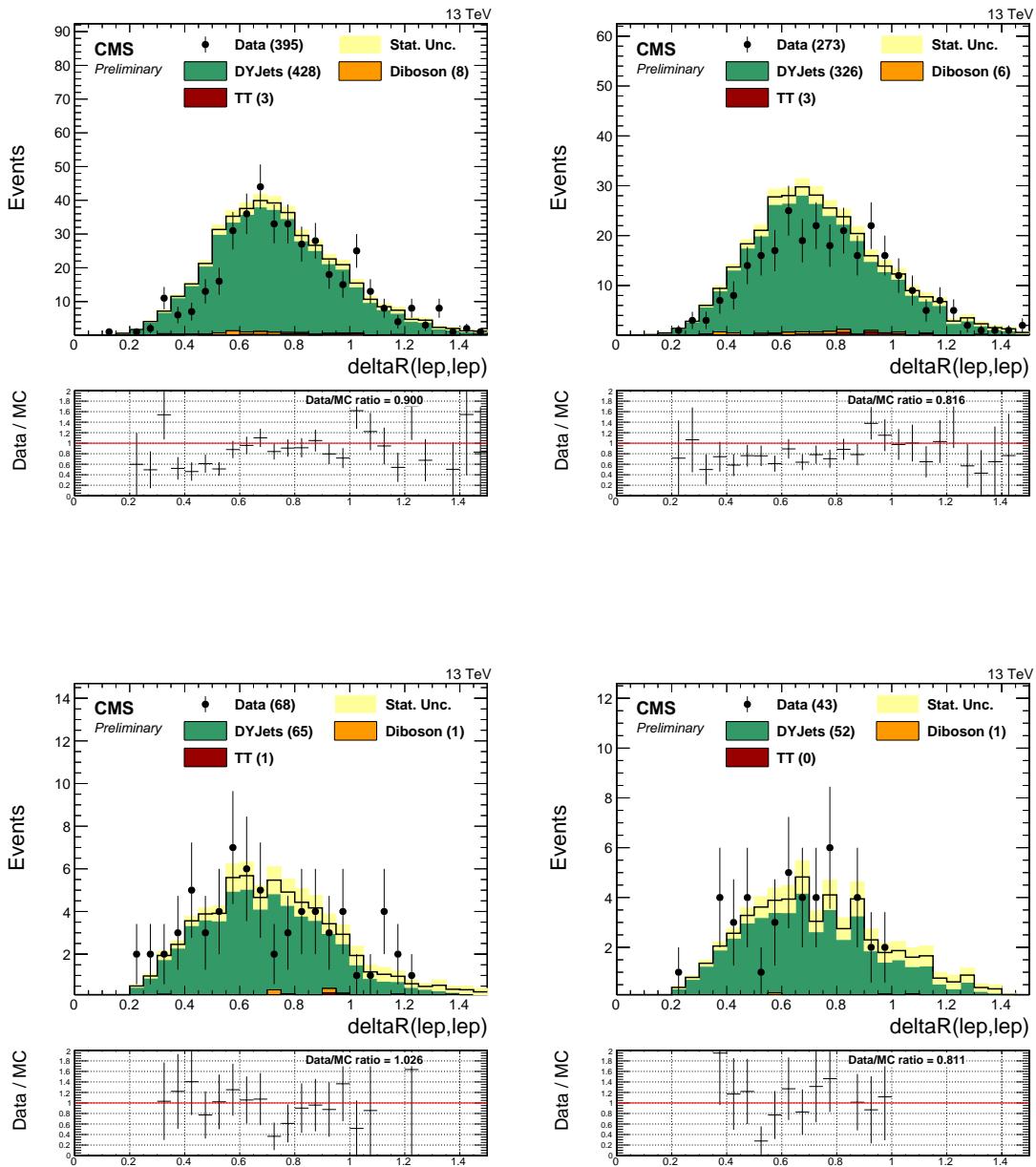


Figure 5.11: Distribution of  $\Delta R$  between leptons in the category: muon low purity (top-left), electron low purity (top-right), muon high purity (bottom-left), and electron high purity (bottom-right). Simulated backgrounds are displayed as stacked histograms normalized to luminosity ( $2.7 \text{ fb}^{-1}$ ).

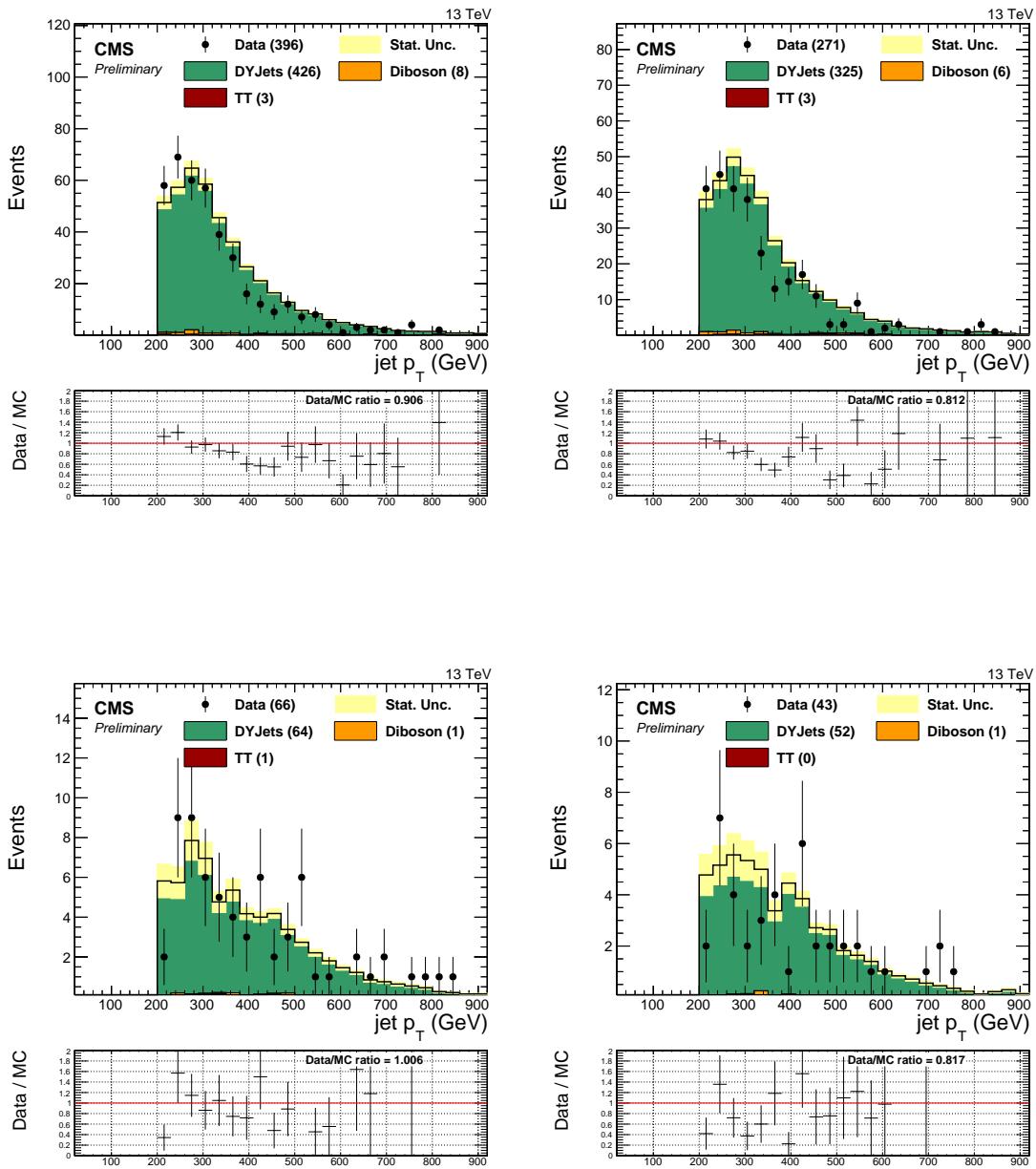


Figure 5.12: Distribution of jet  $p_T$  in the category: muon low purity (top-left), electron low purity (top-right), muon high purity (bottom-left), and electron high purity (bottom-right). Simulated backgrounds are displayed as stacked histograms normalized to luminosity ( $2.7 \text{ fb}^{-1}$ ).

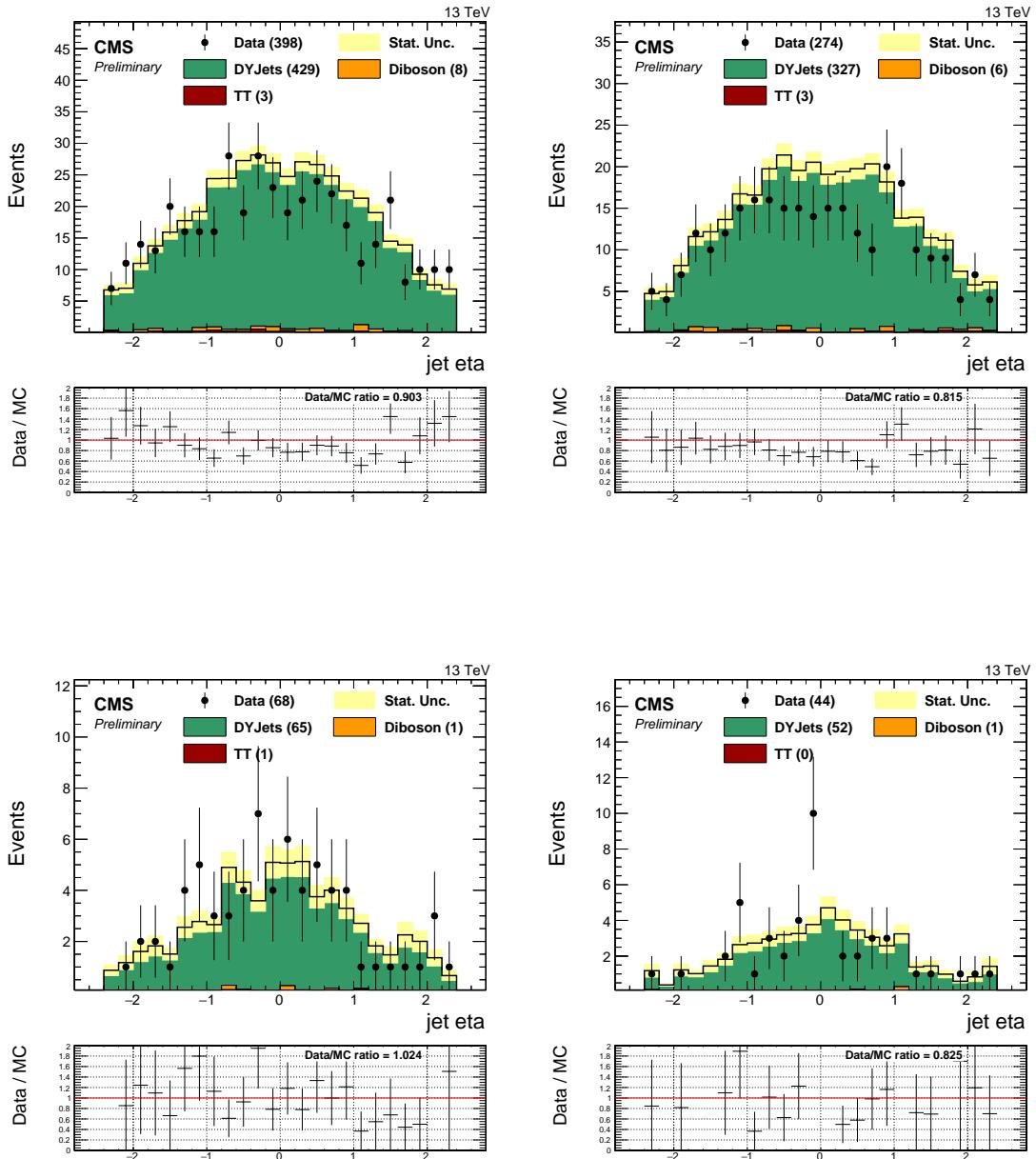


Figure 5.13: Distribution of jet  $\eta$  in the category: muon low purity (top-left), electron low purity (top-right), muon high purity (bottom-left), and electron high purity (bottom-right). Simulated backgrounds are displayed as stacked histograms normalized to luminosity ( $2.7 \text{ fb}^{-1}$ ).

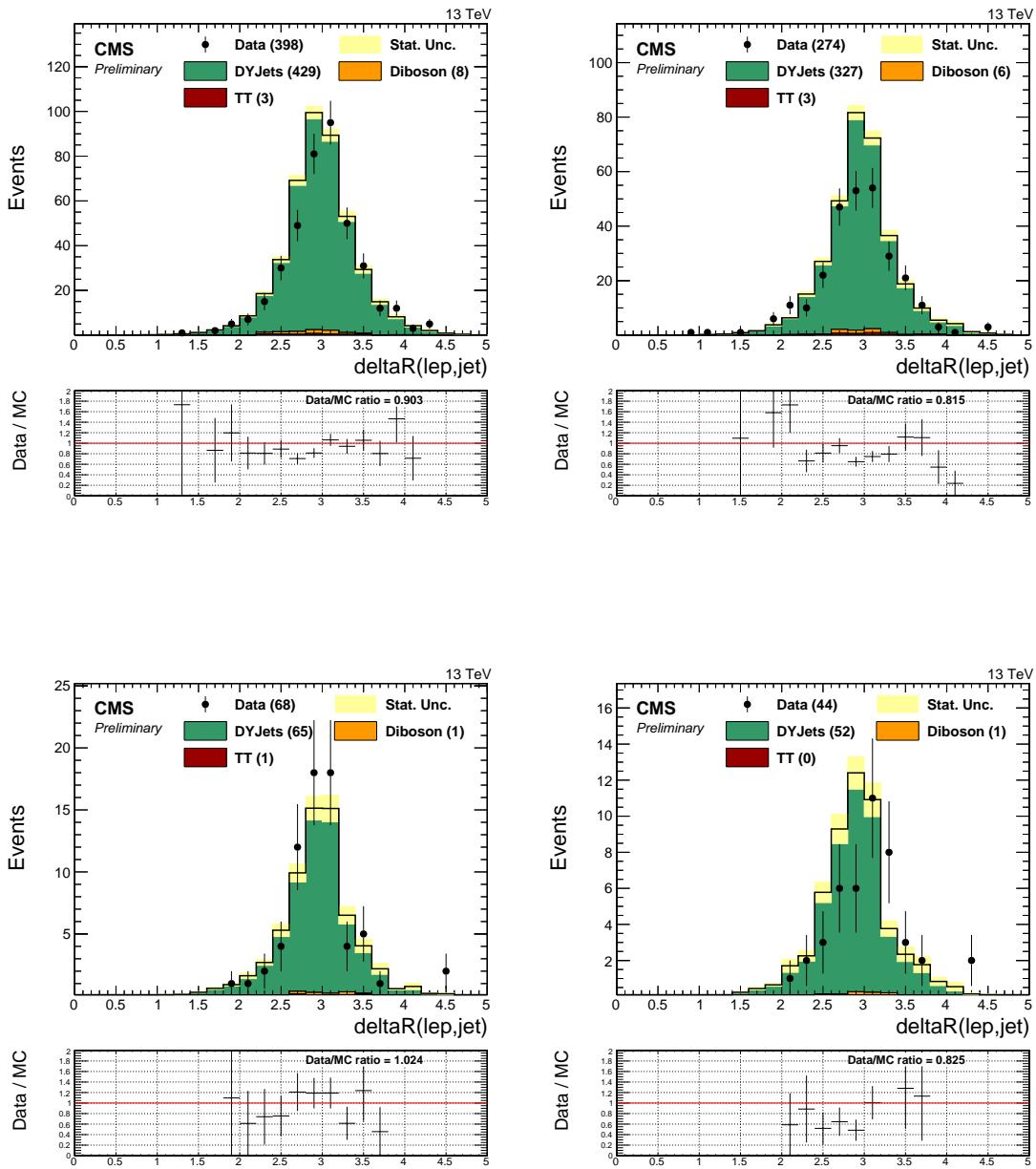


Figure 5.14: Distribution of  $\Delta R$  between the jet and the closest lepton in the category: muon low purity (top-left), electron low purity (top-right), muon high purity (bottom-left), and electron high purity (bottom-right). Simulated backgrounds are displayed as stacked histograms normalized to luminosity ( $2.7 \text{ fb}^{-1}$ ).

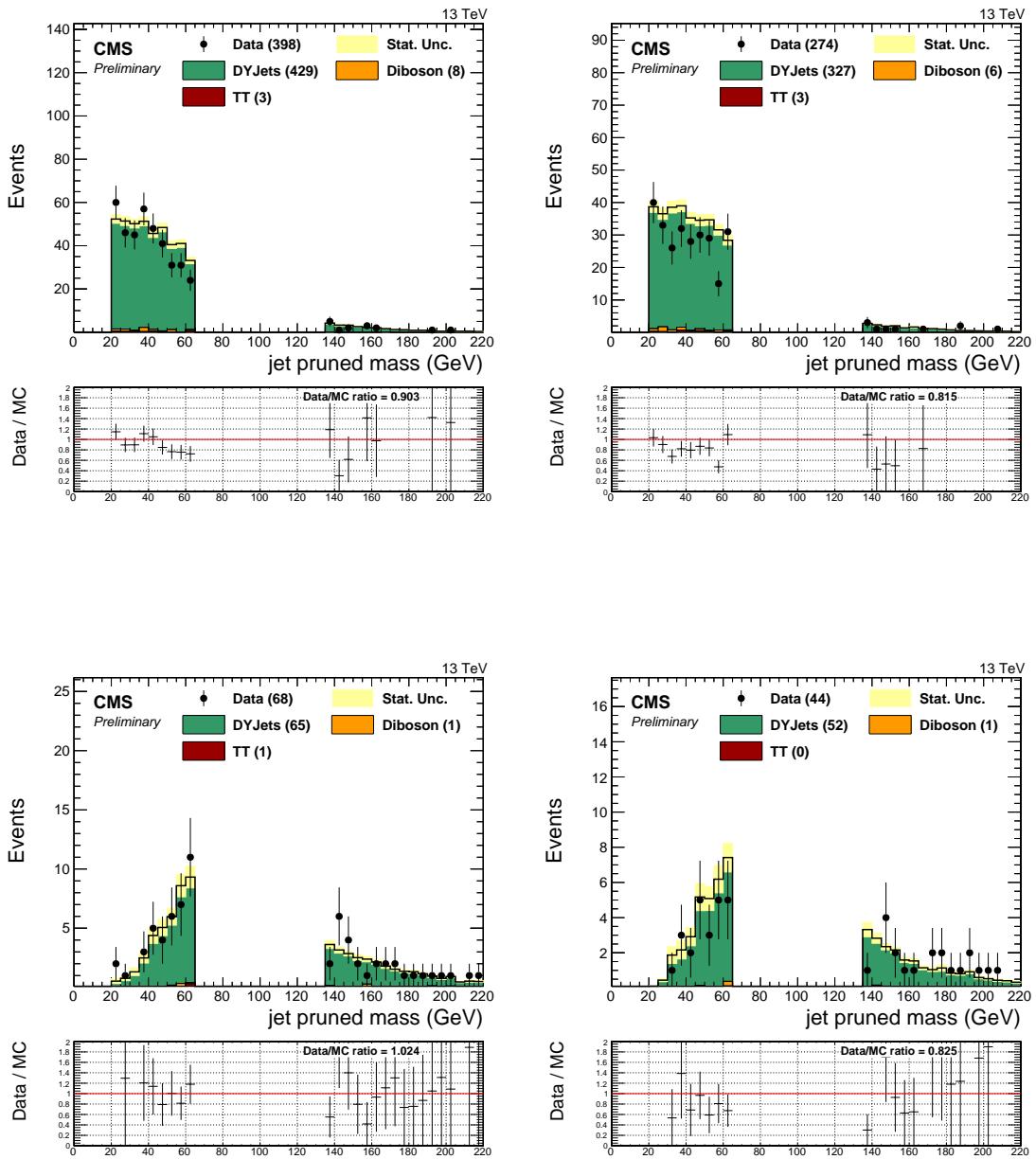


Figure 5.15: Distribution of jet pruned mass in the category: muon low purity (top-left), electron low purity (top-right), muon high purity (bottom-left), and electron high purity (bottom-right). Simulated backgrounds are displayed as stacked histograms normalized to luminosity ( $2.7 \text{ fb}^{-1}$ ).

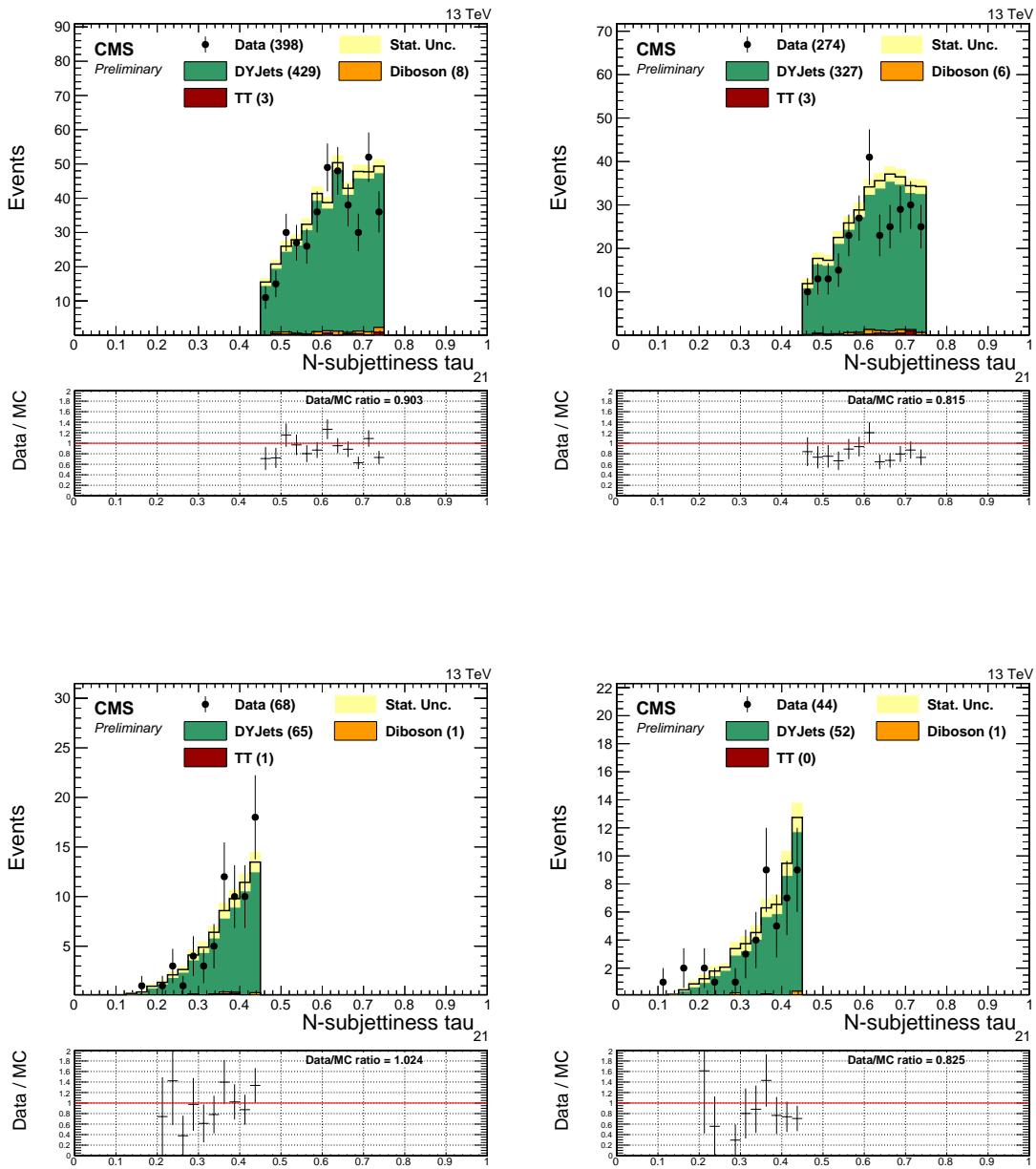


Figure 5.16: Distribution of  $\tau_2/\tau_1 = \tau_{21}$  in the category: muon low purity (top-left), electron low purity (top-right), muon high purity (bottom-left), and electron high purity (bottom-right). Simulated backgrounds are displayed as stacked histograms normalized to luminosity ( $2.7 \text{ fb}^{-1}$ ).

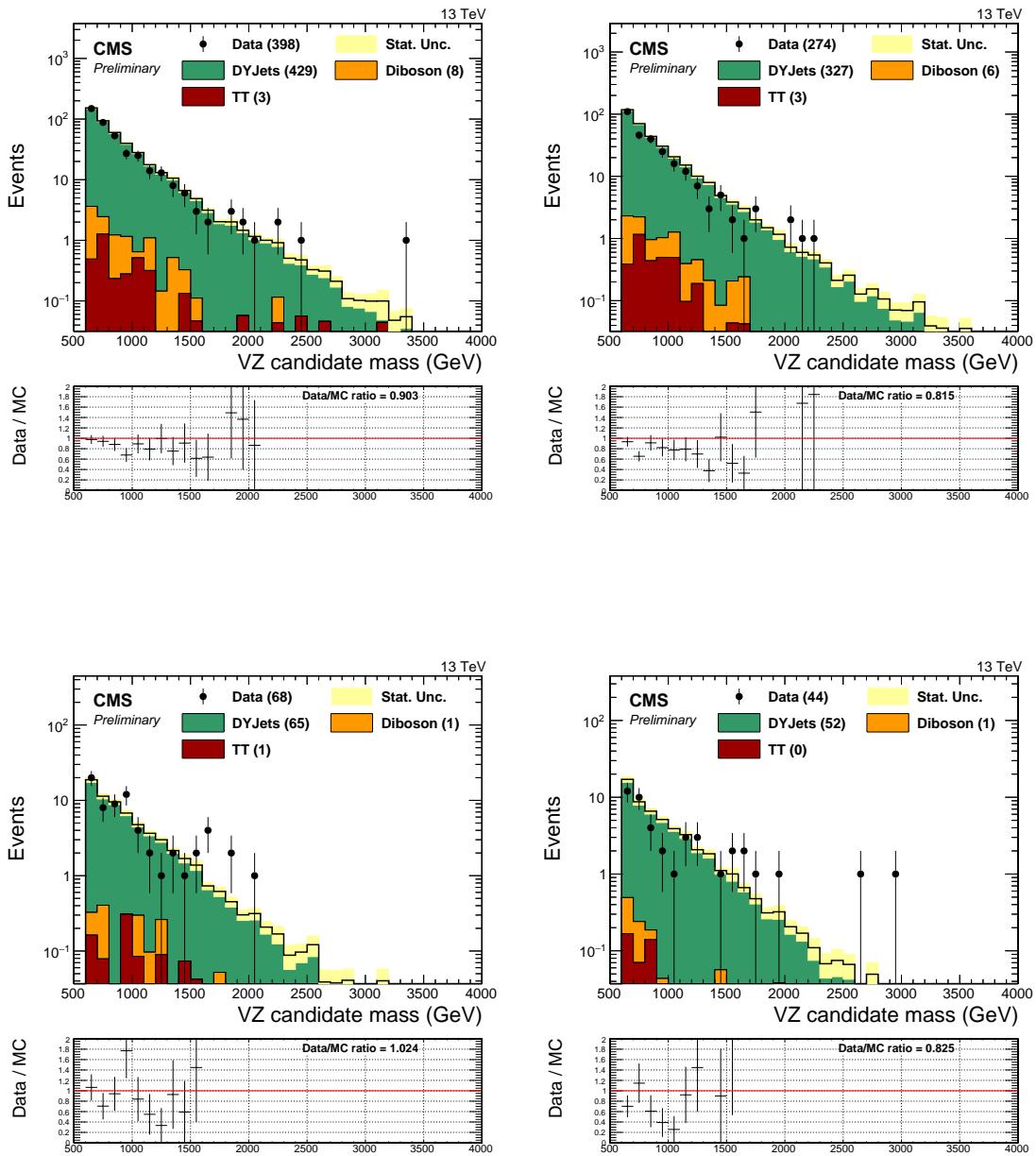


Figure 5.17: Distribution of invariant mass in the category: muon low purity (top-left), electron low purity (top-right), muon high purity (bottom-left), and electron high purity (bottom-right). Simulated backgrounds are displayed as stacked histograms normalized to luminosity ( $2.7 \text{ fb}^{-1}$ ).

# Chapter 6

## Background Estimation and Systematic Uncertainties

### 6.1 Alpha Method

So far we presented distributions of the background based on Monte Carlo (MC) simulations. An alternative data-driven method to estimate the background is preferred because it gets rid of some systematic uncertainties that can make the simulation rather inaccurate. For instance, theoretical uncertainties on the cross section of the background processes can be avoided by applying a normalization based on data. Another example would be the systematic uncertainties related to mismodeling of the process of parton showering and hadronization of the final state quarks. Simulation of the interaction of the particles from the collision with the sensitive volumes of the detector and the readout electronics introduce yet another source of uncertainties.

The data-driven method itself may bring other sources of uncertainties that can be as large as the theoretical uncertainties we want to avoid. Therefore, background estimation techniques cover a broad field of research beyond the material presented in this section. Here, we limit the discussion to one particular technique known as the alpha method [81].

For the estimation of the background in the signal region ( $65 < m_J < 105$  GeV), we use the following inputs:

- Real data distribution of  $m_{ZV}$  in control region:  $(20 < m_J < 65 \text{ GeV}) \cup (135 < m_J < 220 \text{ GeV})$ , which we will call  $f(m_{ZV})_{\text{CR}}^{\text{DATA}}$

- Simulated distribution of the dominant background ( $Z+jets$ ) in both control region and in the signal region, which we will call  $f(m_{ZV})_{\text{CR}}^{\text{Z+jets}}$  and  $f(m_{ZV})_{\text{SR}}^{\text{Z+jets}}$  respectively.
- Simulated distribution of the subdominant background ( $VV$  and  $t\bar{t}$ ) in both control region and in the signal region, which we will call  $f(m_{ZV})_{\text{CR}}^{\text{sub}}$  and  $f(m_{ZV})_{\text{SR}}^{\text{sub}}$  respectively.

The alpha method is applied for the dominant background of our search, which is  $Z+jets$ ; for the subdominant background we use directly the MC estimation. The method exploits the correlation between the jet mass  $m_j$  and the invariant mass  $m_{ZV}$ , by defining a transfer factor as follows:

$$\alpha(m_{ZV}) = \frac{f(m_{ZV})_{\text{SR}}^{\text{Z+jets}}}{f(m_{ZV})_{\text{CR}}^{\text{Z+jets}}} . \quad (6.1)$$

The dominant background estimation in the signal region is obtained by applying that transfer factor to a pure  $Z+jets$  real data sample in the control region, i.e.:

$$\left\langle f(m_{ZV})_{\text{SR}}^{\text{Z+jets}} \right\rangle = \alpha(m_{ZV}) \times \left\langle f(m_{ZV})_{\text{CR}}^{\text{Z+jets}} \right\rangle \quad (6.2)$$

where the angled brackets represent a data estimation of the  $Z+jets$  background. Now the problem has to do with the estimation of the  $Z+jets$  background in the control region. Since that region is essentially signal-free, it is safe to approximate

$$\left\langle f(m_{ZV})_{\text{CR}}^{\text{Z+jets}} \right\rangle = f(m_{ZV})_{\text{CR}}^{\text{data}} - f(m_{ZV})_{\text{CR}}^{\text{sub}} \quad (6.3)$$

whereas the data estimation in the signal region can then be written as

$$\left\langle f(m_{ZV})_{\text{SR}}^{\text{Z+jets}} \right\rangle = \frac{f(m_{ZV})_{\text{SR}}^{\text{Z+jets}}}{f(m_{ZV})_{\text{CR}}^{\text{Z+jets}}} \times \left( f(m_{ZV})_{\text{CR}}^{\text{data}} - f(m_{ZV})_{\text{CR}}^{\text{sub}} \right) . \quad (6.4)$$

In plain words,  $\alpha(m_{ZV})$  is the ratio of the  $m_{ZV}$  distribution in the signal region over the distribution in the control region for the dominant background, and it is shown in Fig. 6.1. This transfer factor is then used, after controlling the presence of the subdominant backgrounds, to correct the  $m_{ZV}$  distribution in the control region that is shown in Fig. 6.2. The resulting prediction in the signal region will be presented in the next section containing the unblind results. We note that the method leads to a prediction of both the shape and normalization of the dominant

background in the signal region; however, we discarded the latter and retain only the shape prediction. For the estimation of the normalization, a more robust method is applied as we will explain in a moment.

In Eq. 6.4, the function  $f(m_{ZV})$  can be modeled by a leveled exponential defined below:

$$f(m_{ZV}) = \exp\left(\frac{-m_{ZV}}{a(m_J) + b(m_J) m_{ZV}}\right), \quad (6.5)$$

where  $a(m_J), b(m_J)$  are parameters that implicitly depend on  $m_J$ . In practice, we assume that these parameters are constant within a given region, i.e. there is a set of constant parameters  $a^{(\text{SR})}, b^{(\text{SR})}$  that leads to a good description of the  $m_{ZV}$  distribution in the signal region, and analogously for the control region.

Our implementation of the alpha method was coded in ROOT [82] enhanced with the RooFit extension [83]. The algorithm starts with the declaration of the probability density functions (PDF), specifically 5 leveled exponentials for the following cases:

- Simulated dominant background in the signal region;
- Simulated subdominant background in the signal region;
- Simulated dominant background in the control region;
- Simulated subdominant background in the control region;
- Real data in the control region.

In total there are 10 correlated parameters that need to be fit. The routine performs a simultaneous fit of the PDFs associated with the dominant background and real data, in order to get the best estimation of the parameters as well as the correlation matrix that allows the propagation of uncertainties to the final result. The modeling of the subdominant background is taken directly from simulation; to facilitate the convergence of the simultaneous fit during the alpha method, the parameters associated to the subdominant component are kept fixed.

The prediction of the background normalization in the signal region is derived by interpolating the data from the control region of the jet mass distribution. The baseline model for the shape of the jet mass was tuned in simulation, as indicated in Fig. 6.3. In order to estimate the adequacy of the model, we made a data-based study on the viability of different alternatives. This translates directly to a systematic uncertainty in the normalization of the background in the signal region. Different model choices are shown in Fig. 6.4, and the estimated yield in

the signal region is reported in Table 6.2 for each model. From these results we estimated a systematic uncertainty ( $\Delta$ ) that ranges from 28% in the EHP category to 42% in the MHP category.

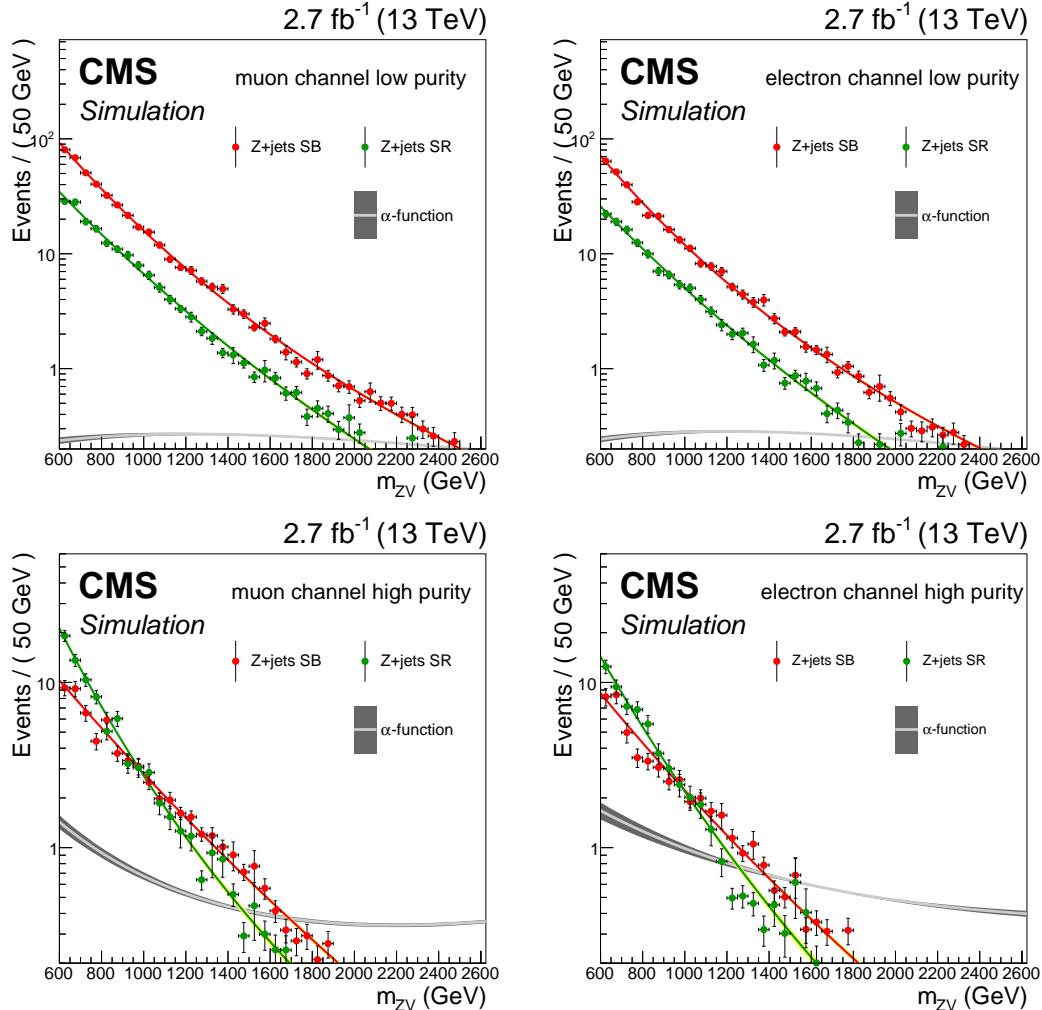


Figure 6.1: Top:  $m_{ZV}$  simulated distributions in the signal (green) and control (red) region for the low purity category, for muons (left) and electrons (right). Bottom:  $m_{ZV}$  simulated distributions in the signal (green) and control (red) region for the high purity category, for muons (left) and electrons (right). The transfer factor  $\alpha(m_{ZV})$  is defined by the ratio  $(m_{ZV} \text{ signal region}) / (m_{ZV} \text{ control region})$ .

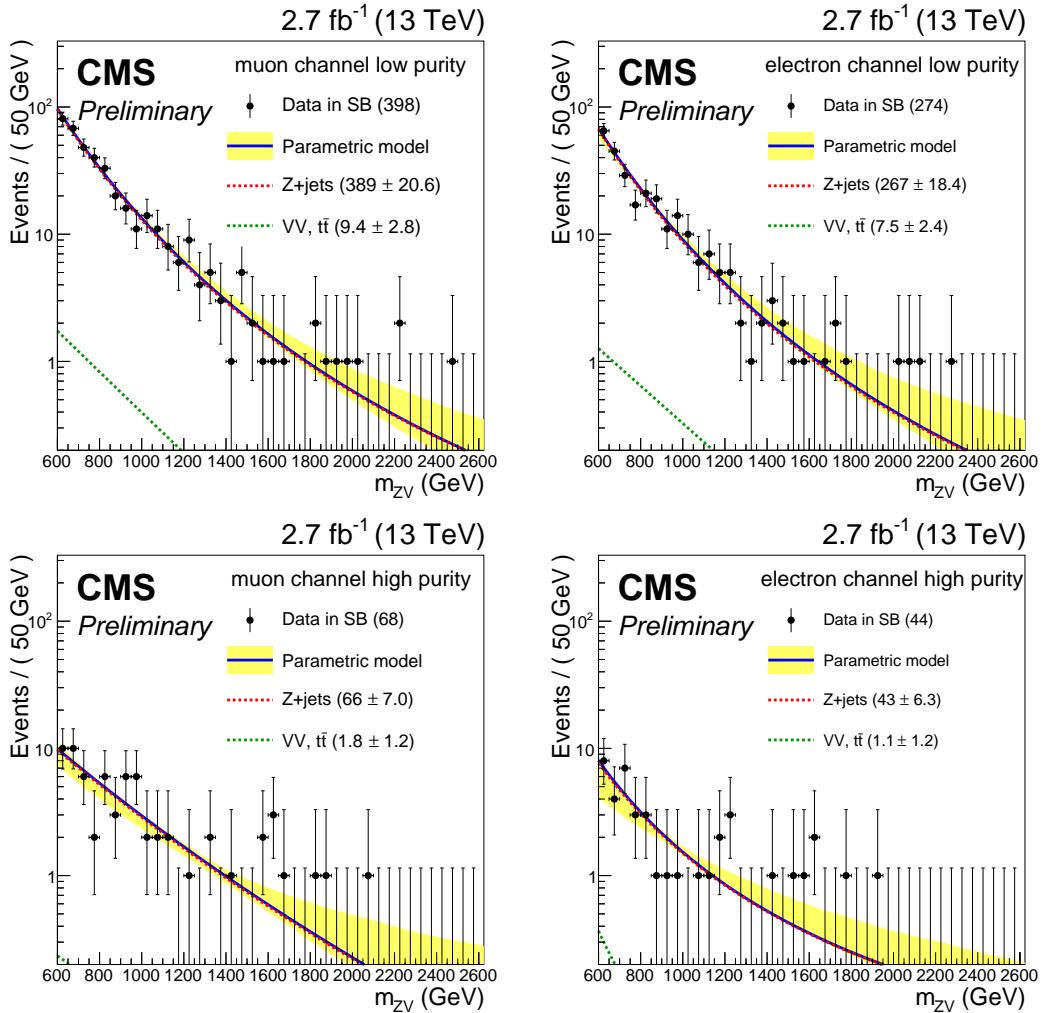


Figure 6.2: Top:  $m_{ZV}$  distributions in the control region for the low purity category, for muons (left) and electrons (right). Bottom:  $m_{ZV}$  distributions in the signal region for the high purity category, for muons (left) and electrons (right).

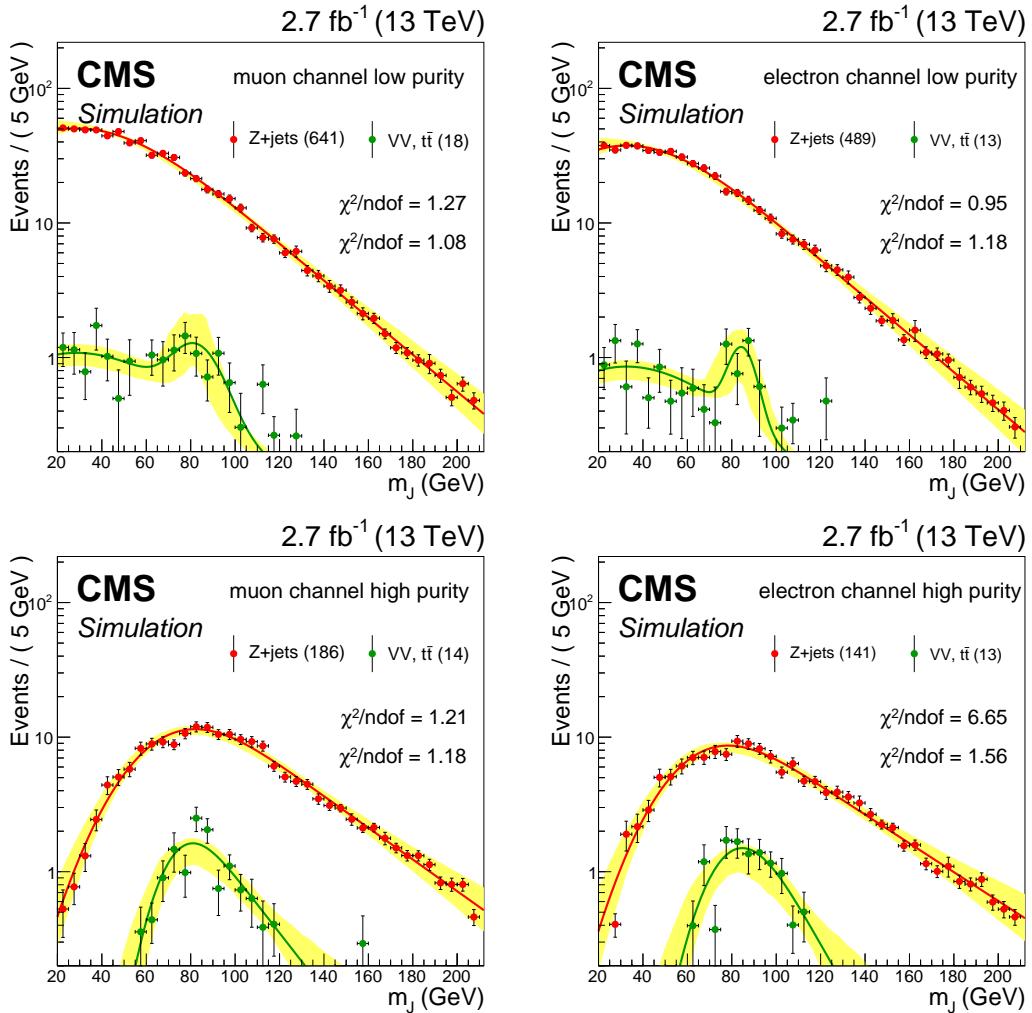


Figure 6.3: Fits of the jet mass for the dominant and subdominant components of the background using MC simulation.

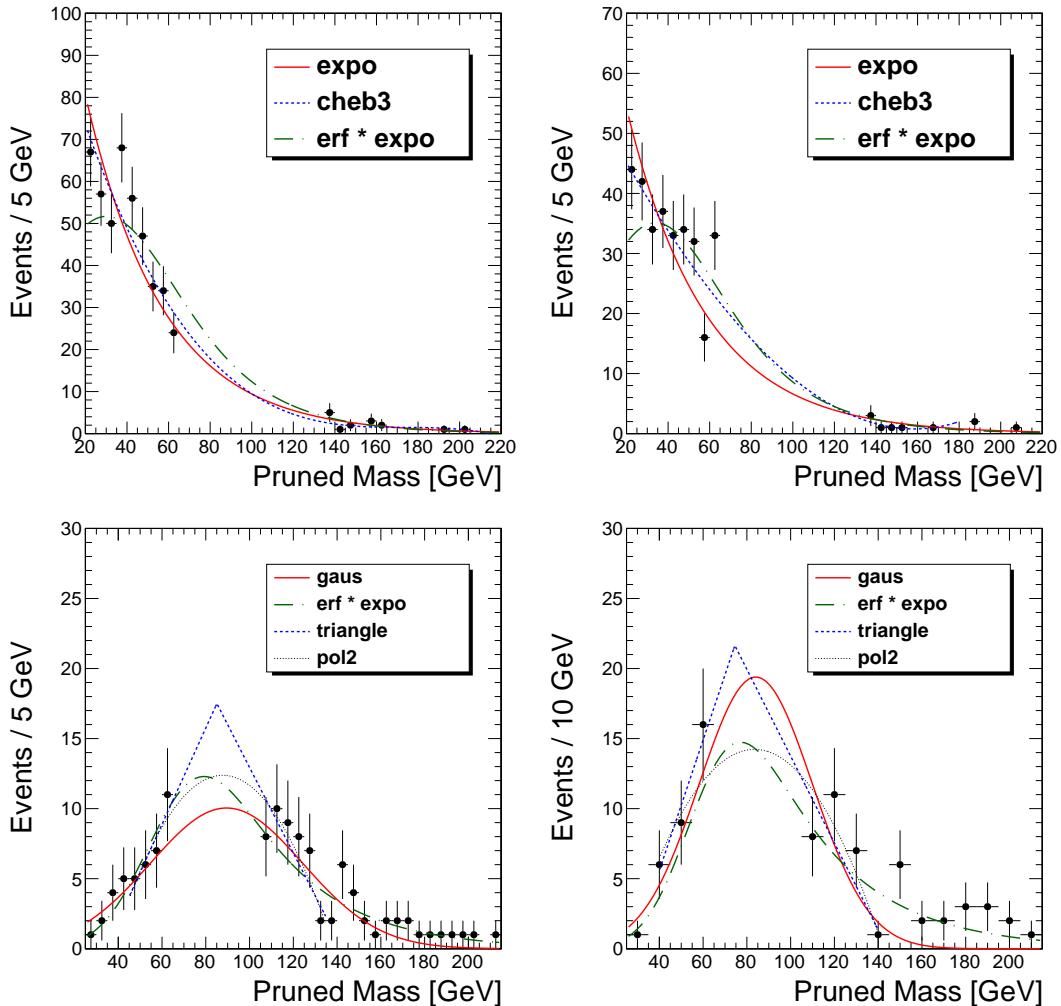


Figure 6.4: Fit of the jet mass using different models to estimate the systematic uncertainty in the background normalization. The categories are: muon low purity (top-left), electron low purity (top-right), muon high purity (bottom-left), and electron high purity (bottom-right).

## 6.2 Unblind of the Signal Region

Fig. 6.5 shows the jet mass distribution in data and the parametric model made out of two components. The dominant component accounts for the Z+jets background, and the subdominant component corresponds to diboson VV and  $t\bar{t}$  background. The parametric model consists of an error function multiplied by an exponential (ErfExp); in the low-purity categories a gaussian on top of the ErfExp is used for modeling the subdominant backgrounds. The goodness of this choice was evaluated first in simulation as demonstrated in Fig. 6.3. The small panel below every plot contains a pull histogram between the data and the adjusted model. Additionally, Table 6.1 contains the event yields and Data/MC ratios including the signal region.

The description of the jet mass with the parametric model was tested in simulation and compared against alternative functions. To account for the mismodeling of the jet mass, an uncertainty was added as a systematic error that ranges between 28-42%. The expected number of events in the signal region is given in Table 6.3 for the categories muon low purity (MLP), electron low purity (ELP), muon high purity (MHP), and electron high purity (EHP). The estimated background is reported in the format A  $\oplus$  B, where A and B represent the dominant and subdominant components, respectively.

Table 6.1: Events yields and Data/MC ratios including the signal region.

Category	Data	MC	Data/MC
Muon Low Purity	592	596	0.993
Muon High Purity	222	208	1.067
Electron Low Purity	412	412	0.999
Electron High Purity	155	130	1.192

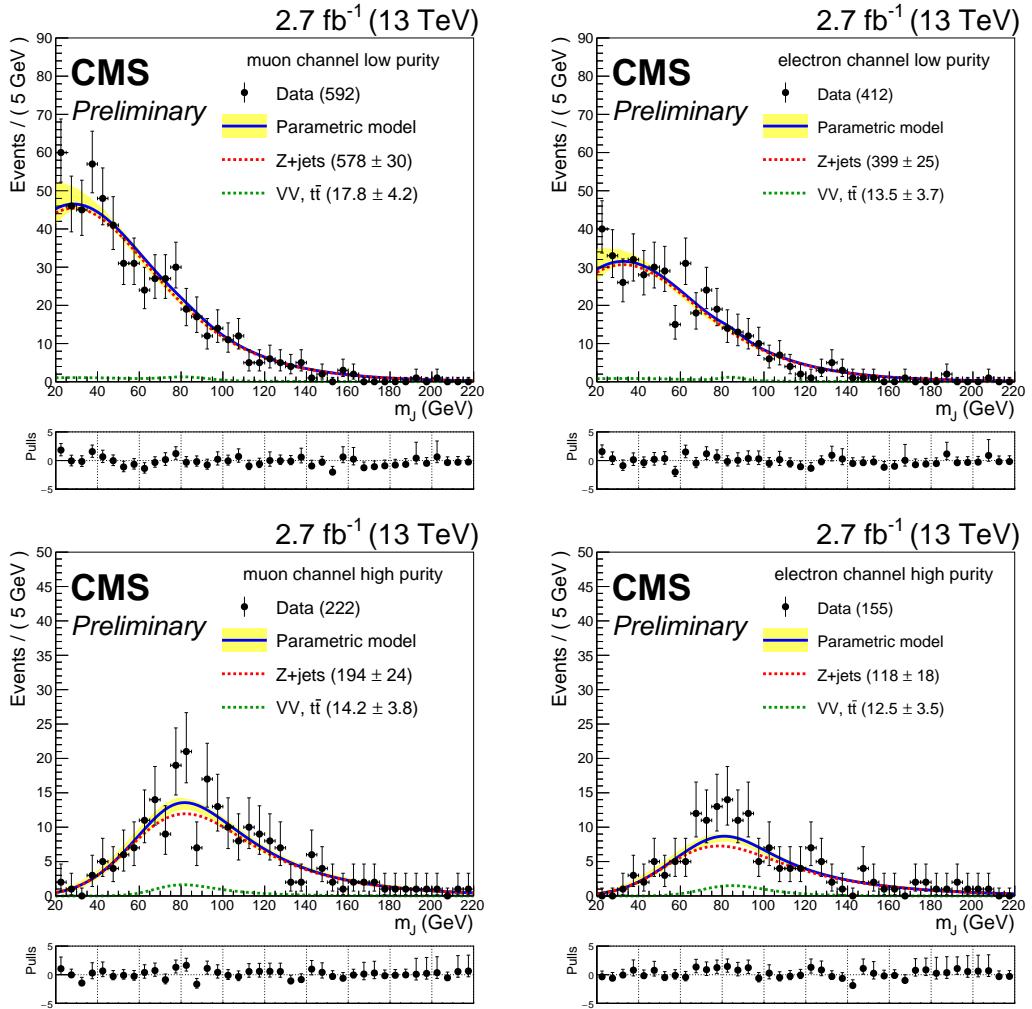


Figure 6.5: Jet mass distribution in data and parametric model (blue line). The dominant (red) and subdominant (green) components are also shown. The yield for each component in the full region ( $20 < m_J < 220$  GeV) is written in the legend.

Table 6.2: Estimated background in the signal region for different models and associated uncertainty.

	Exp	Chebychev 3	ErfExp	Gaus	Polynomial 2	Triangle	$\frac{\text{largest} - \text{smallest}}{\text{smallest}}$	$\frac{\text{largest} - \text{smallest}}{\text{largest}}$	$\Delta$
MLP	118	129	162	-	-	-	37%	27%	32%
ELP	82	100	114	-	-	-	39%	28%	34%
MHP	-	-	89	76	94	114	50%	33%	42%
EHP	-	-	53	70	55	70	32%	24%	28%

Table 6.3: Background estimation obtained by the integral of the parametric model in the signal region ( $65 < m_j < 105$  GeV). The estimated background is reported in the format  $A \oplus B$ , where A and B represent the dominant and subdominant components, respectively. The systematic uncertainty due to mismodeling of the jet mass ranges between 28-42%.

Category	Parametric model	Estimated background	Syst. unc.	Data in SR
MLP	(ErfExp) $\oplus$ (ErfExp + Gaussian)	$(150 \pm 11) \oplus (7.4 \pm 3)$	32 %	157
ELP	(ErfExp) $\oplus$ (ErfExp + Gaussian)	$(105 \pm 10) \oplus (5 \pm 2)$	34 %	116
MHP	(ErfExp) $\oplus$ (ErfExp)	$(88 \pm 8) \oplus (10 \pm 3)$	42 %	110
EHP	(ErfExp) $\oplus$ (ErfExp)	$(53 \pm 7) \oplus (9 \pm 3)$	28 %	85

The  $m_{ZV}$  distribution of data in the signal region and the final background estimation is shown in Fig. 6.6. The decomposition of the parametric model into dominant and subdominant components is also shown. The error band in the parametric model accounts for the shape uncertainties due to the transfer factor and the uncertainties in the fit of the  $m_{ZV}$  distribution in data in the sideband region. The normalization uncertainty (between 28–42 %) due to mismodeling of the jet mass is considered as a nuisance parameter in the statistical framework that calculates the expected limits. An expanded discussion of systematic uncertainties and the statistical treatment is giving in the following sections.

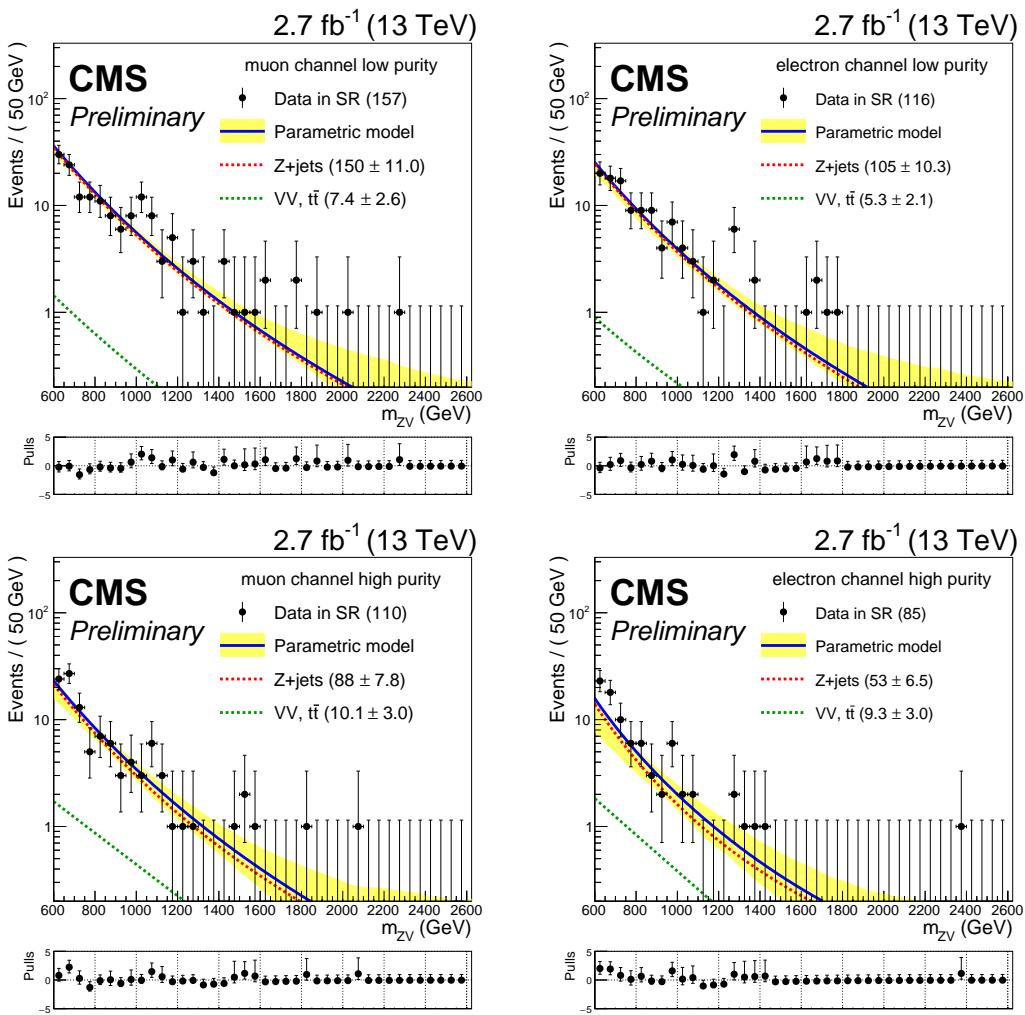


Figure 6.6: Top:  $m_{ZV}$  distributions in the signal region for the low purity category, for muons (left) and electrons (right). Bottom:  $m_{ZV}$  distributions in the signal region for the high purity category, for muons (left) and electrons (right).

## 6.3 Systematic Uncertainties

A graviton with mass  $M_G$  is expected to decay into a pair of Z bosons with  $p_T \sim M_G/2$  each, as verified in the distribution of the Z boson  $p_T$  at generator and reconstructed level shown in Fig. 6.7. The mean value of the Z boson  $p_T$  in Fig. 6.7 is compared with the reconstructed jet  $p_T$  for two scenarios: with and without applying the jet energy corrections recommended by the JetMET physics object group [80]. As expected, the jet energy corrections improve the agreement between reconstruction and generator level, both for the jet  $p_T$  and the jet mass distributions.

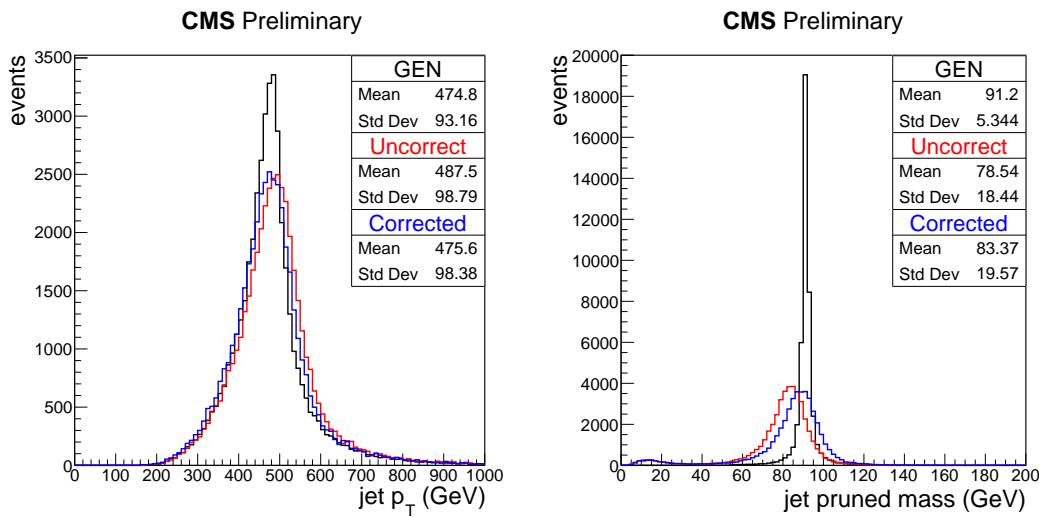


Figure 6.7: Jet  $p_T$  (left) and jet mass (right) corresponding to a bulk graviton of mass 1 TeV. Jet energy corrections improve the agreement between the reconstructed jet and the generated Z boson.

Other systematic uncertainties influence both the normalisation and shape of the background and signal. We consider effects from leptons (trigger, selection), hadrons (jet energy, V tagging), and LHC luminosity [84]. Following the recommendations from the jet physics object group [80], we assign a relative uncertainty of 6.7% (26%) on the signal yield due to the V tagging in the high (low) purity category. Jet energy corrections (JEC) are considered to scale up/down the jet  $p_T$  affecting the position of the signal peak, as shown in Fig. 6.8. The associated systematic uncertainty due to the shift of the signal peak varies between 0.6% and 0.95%, and increases with the mass of the resonance.

Lepton selections, both at trigger level and offline, also contribute to the systematic uncertainties. To estimate this uncertainty we vary the data/MC scale

factors in each  $\eta - p_T$  bin, leading to the results shown in Fig. 6.9. In summary, the systematic uncertainties taken into account are gathered in Table 6.4.

Table 6.4: Summary of systematic uncertainties on the signal normalization.

Source	Value	Comment
Luminosity	2.7%	correlated between all categories
Electron trigger and ID	2.5%	tag-and-probe study
Electron energy scale	0.5%	
Muon trigger and ID	10%	tag-and-probe study
Muon momentum scale	0.5%	
Jet energy scale	1%	correlated between all categories
V tagging	6.7% 26%	high purity low purity

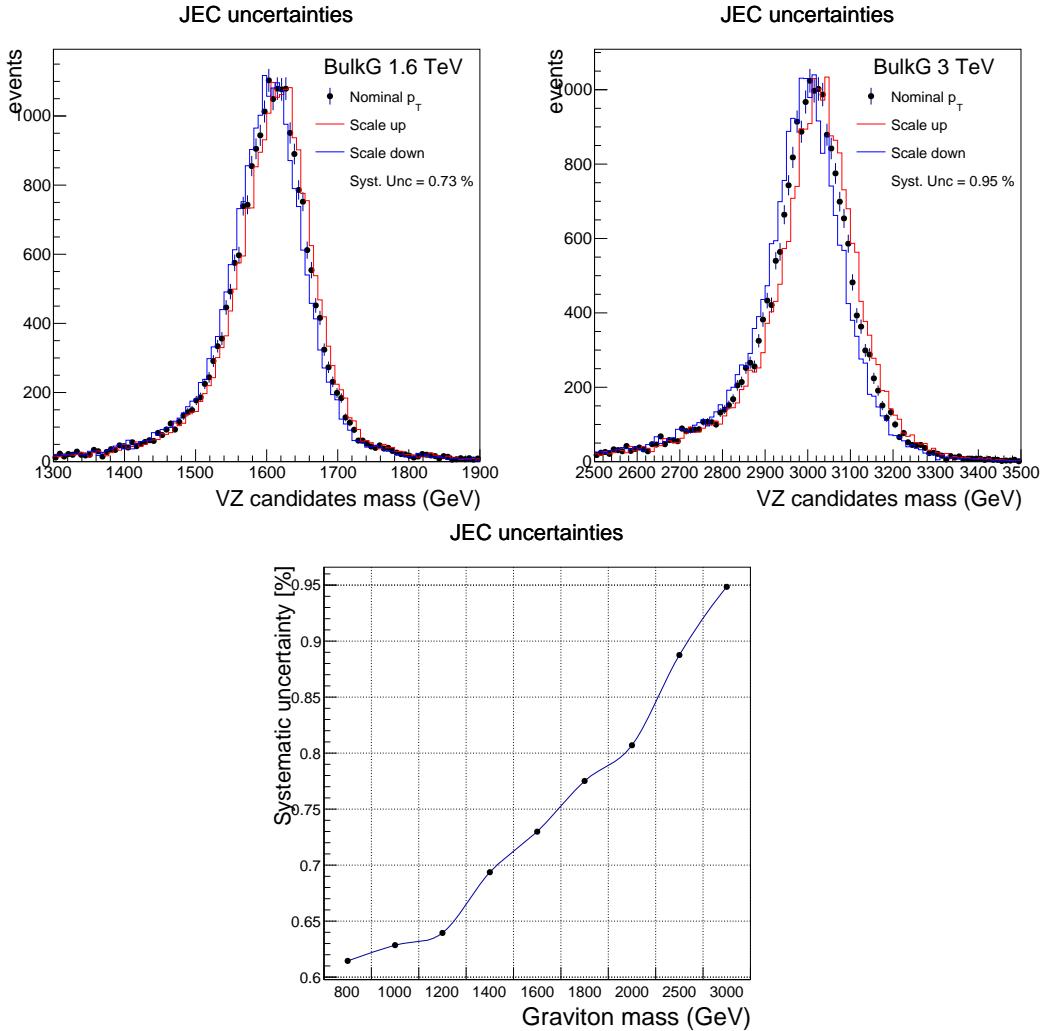


Figure 6.8: Scale up/down variations of jet energy correction (JEC) uncertainties, for a bulk graviton of mass 1.6 TeV (top-left) and 3 TeV (top-right). The associated systematic uncertainty due to the shift of the signal peak varies between 0.6% and 0.95% (bottom plot).

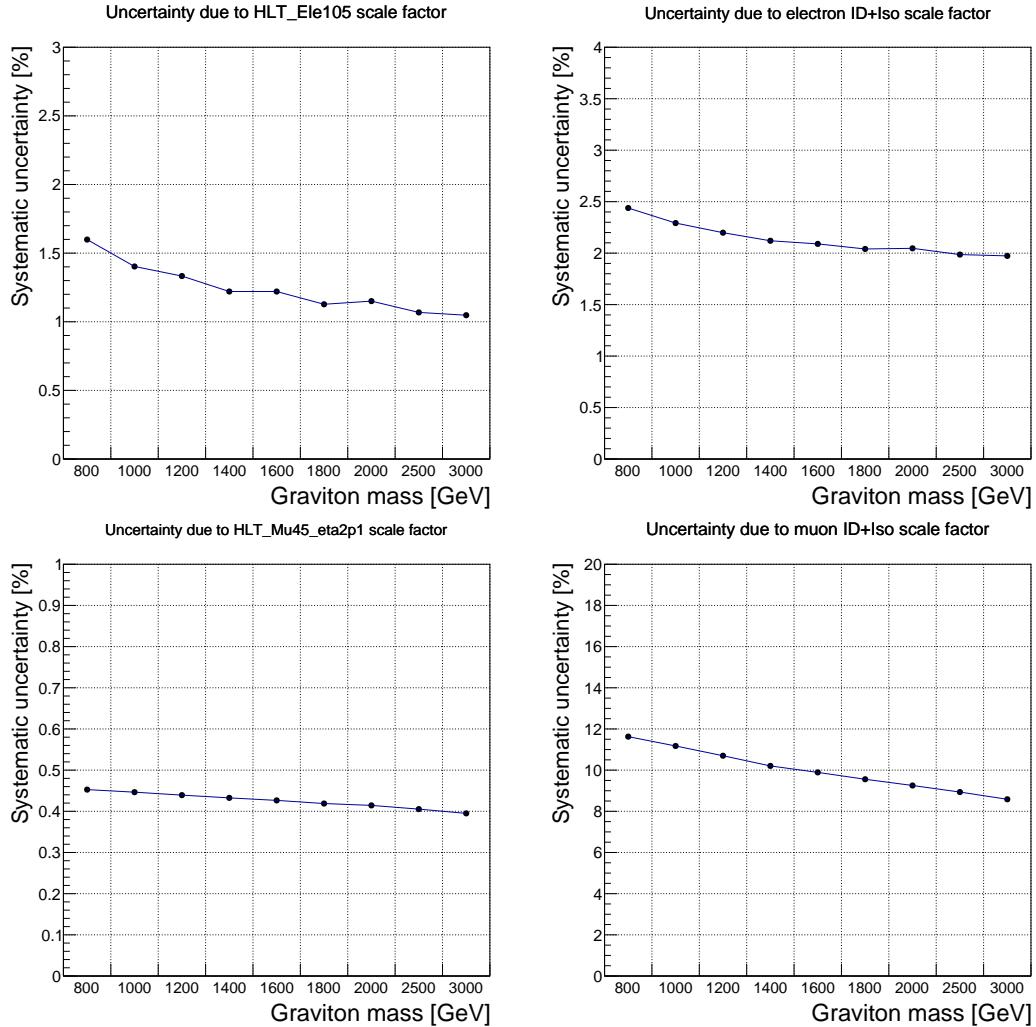


Figure 6.9: Systematic uncertainty due to the scale factors on the electron trigger (top-left), ID+Isolation (top-right). Systematic uncertainty due to the scale factors on the muon trigger (bottom-left), ID+Isolation (bottom-right).

# Chapter 7

## Final Results and Limits

### 7.1 Statistical Method

It is possible to showcase the statistical component of the analysis through a counting experiment model [85]. Consider the observable  $n$  a random variable describing the number events. Rather than a unique value, repeated experiments may yield different number of events independently of the history or previous results. The observed event yield is expected to be distributed according to the Poisson law with mean value

$$\alpha = \mu \cdot s + b. \quad (7.1)$$

The standard model provides an estimation for the background yield  $b$ , and we want to find a confidence interval on the signal yield  $\mu \cdot s$ . The parameter  $\mu$  determines the strength of the signal process, with  $\mu = 0$  corresponding to the background-only hypothesis and  $\mu = 1$  being the nominal signal hypothesis. The probability function for the observable  $n$  can be written as

$$p(n|\alpha) = \frac{\alpha^n e^{-\alpha}}{n!}, \quad (7.2)$$

that is the probability for observing  $n$  events, assuming the parameter  $\alpha$  is fixed.

The likelihood is a function of the model parameters, and is used to quantify the result obtained after throwing one experiment. In case of the counting experiment, the likelihood looks similar to the probability function with the substitution of  $n$

by the actual realization of the event count, denoted here as  $N$ :

$$L(\mu) = \frac{(\mu s + b)^N e^{-(\mu s + b)}}{N!} L(b). \quad (7.3)$$

The term  $L(b)$  describes our knowledge about the background obtained from a subsidiary measurement. For example, a control sample where mainly background events are expected may yield to  $m$  background events, that is

$$L(b) = \frac{b^m}{m!} e^{-b}. \quad (7.4)$$

To test a hypothesis value of  $\mu$  the profile likelihood ratio is considered:

$$\lambda(\mu) = \frac{L(\mu)}{L(\hat{\mu})}, \quad (7.5)$$

where the denominator is the maximized likelihood function. The above definition ensures that  $0 \leq \lambda \leq 1$ , with  $\lambda$  near 1 implying good agreement between the data and the hypothesis value of  $\mu$ .

The procedure to establish a confidence interval for the signal strength  $\mu$  is based on the asymptotic modified frequentist CLs method [86] implemented in a tool provided by the Higgs group in CMS. The method relies on an asymptotic approximation of the distributions of a test-statistic based on the profile likelihood ratio. The asymptotic method is fairly accurate when the event yields are not too small and the systematic uncertainties do not play a major role in the result.

## 7.2 Confidence Limits

For each mass hypothesis a maximum likelihood fit of the data in each category is performed using background-only and signal-plus-background model. A likelihood ratio of the two fits is used as a test statistic for the asymptotic method with systematic uncertainties incorporated as nuisance parameters. The results are expressed as upper limits on the cross section times branching fraction for the process  $X \rightarrow ZZ$  for the bulk graviton model.

Figure 7.1 shows the observed and expected limits for the full dataset recorded in 2015 at 13 TeV, corresponding to a luminosity of  $2.69 \text{ fb}^{-1}$ . The theoretical graviton production cross sections and branching ratios, represented by the red

line, are taken from [38].

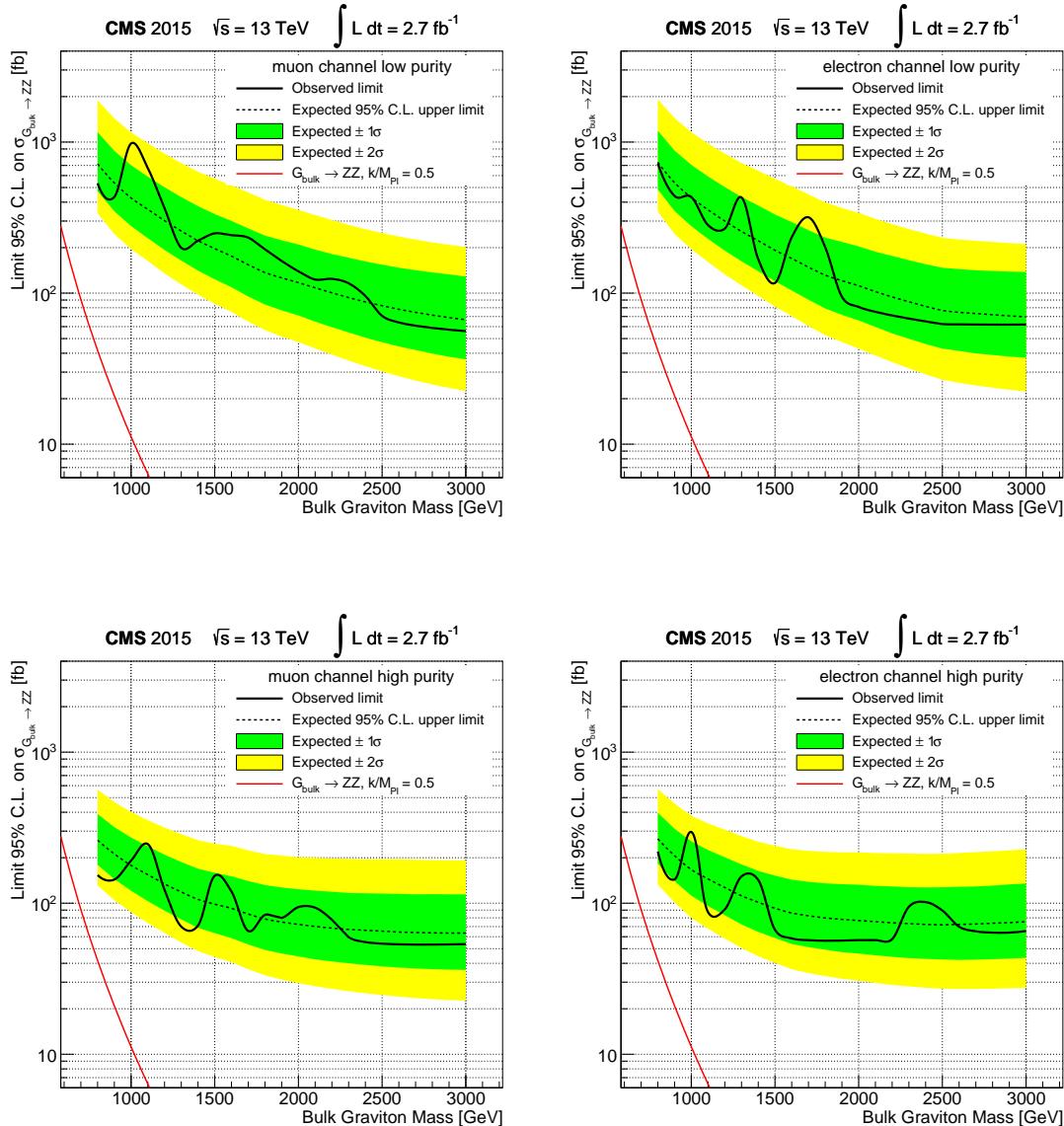


Figure 7.1: Upper limits in the categories: muon low purity (top-left), electron low purity (top-right), muon high purity (bottom-left), and electron high purity (bottom-right.)

Finally, the product of the likelihoods between the different categories is computed to combine the individual limits and obtain a more sensitive result, as shown in Fig. 7.2.

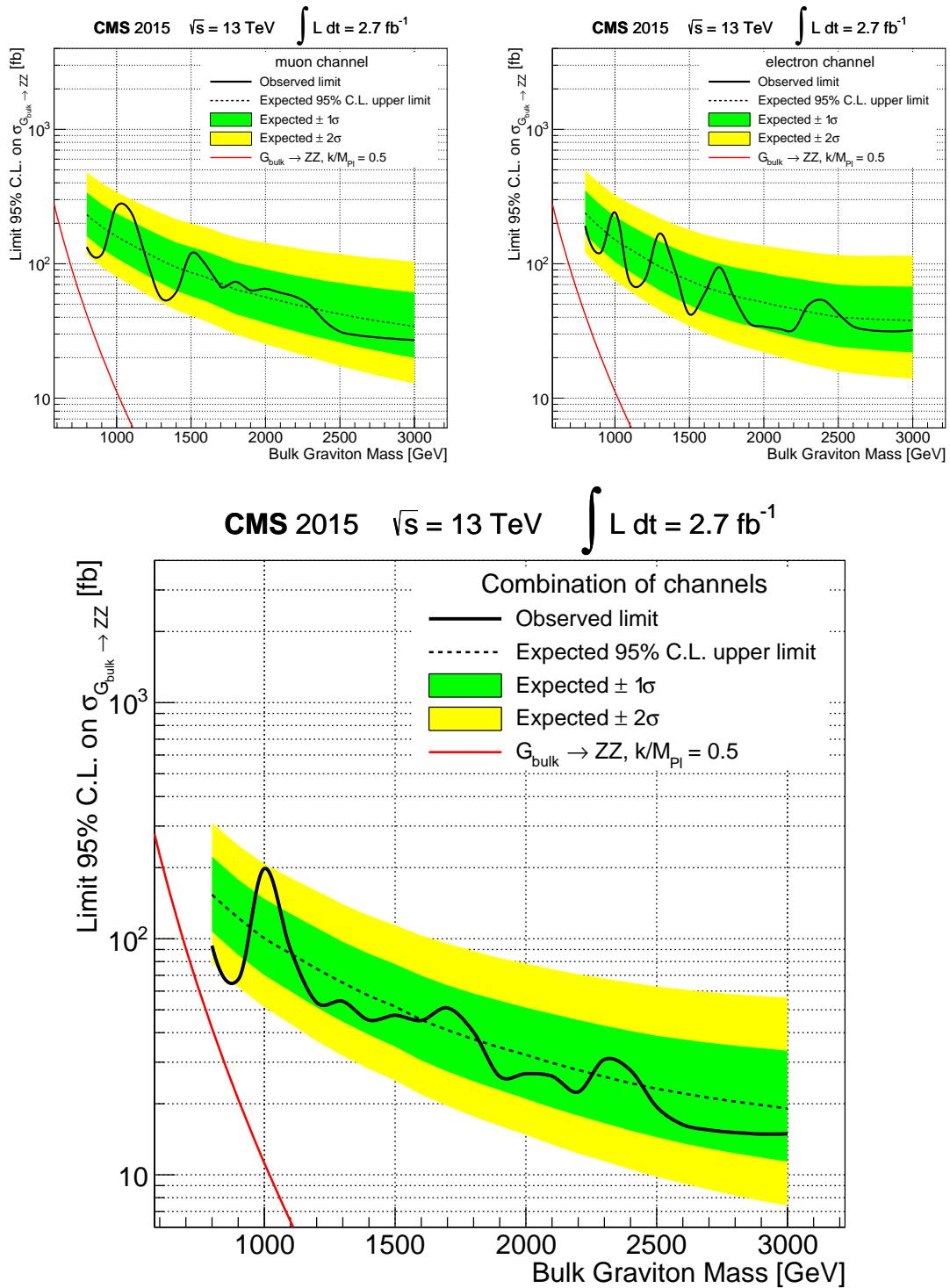


Figure 7.2: Upper limits obtained after combination of the high and low purity categories in the muon channel (top-left), electron electron channel (top-right), and the combination of all categories (bottom).

## 7.3 Interpretation of Results

The final limit in Fig. 7.2 combines all categories: muon low purity, electron low purity, muon high purity, and electron high purity. When the observed limit is above the expected value, it is possible to claim an excess over the standard model background. The significance of an excess is evaluated through the p-value, that is the incompatibility of the background-only hypothesis with the observed data. According to Fig. 7.3, the most significant point is found at 1 TeV corresponding to a p-value = 0.018. The p-value is often converted into a number of standard deviations of a gaussian distribution; if the significance is larger than  $5\sigma$ , an announcement of discovery can be made. Our particular excess in 1 TeV corresponds to a significance of  $2\sigma$ , as the solid line in Fig. 7.2 just touches the upper yellow band.

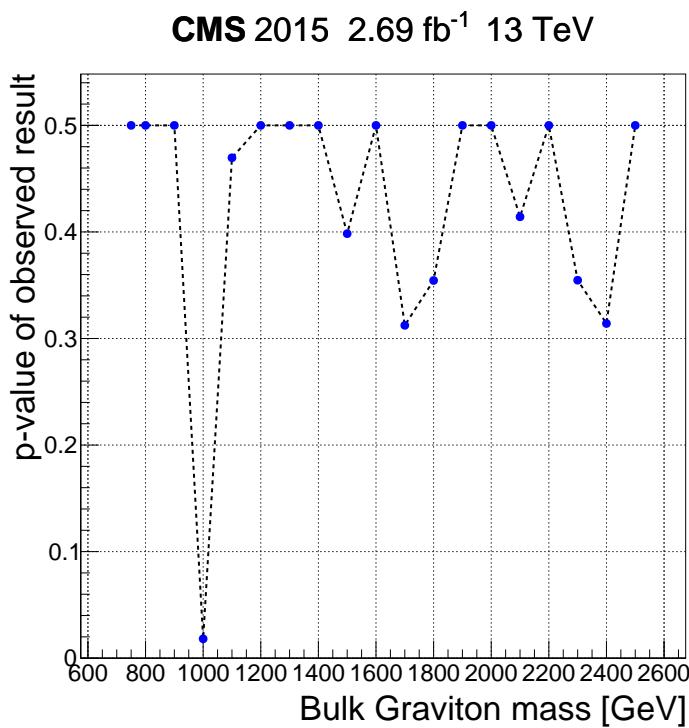


Figure 7.3: Significance of the observed results. The most significant result corresponds to the lowest p-value, observed at 1 TeV.

The impacts of the nuisance parameters on the signal strength for the 1 TeV point are shown in Fig. 7.4. The direction of the  $+1$  sigma and  $-1$  sigma impacts indicates whether the parameter is correlated or anti-correlated with the variation of the signal strength.

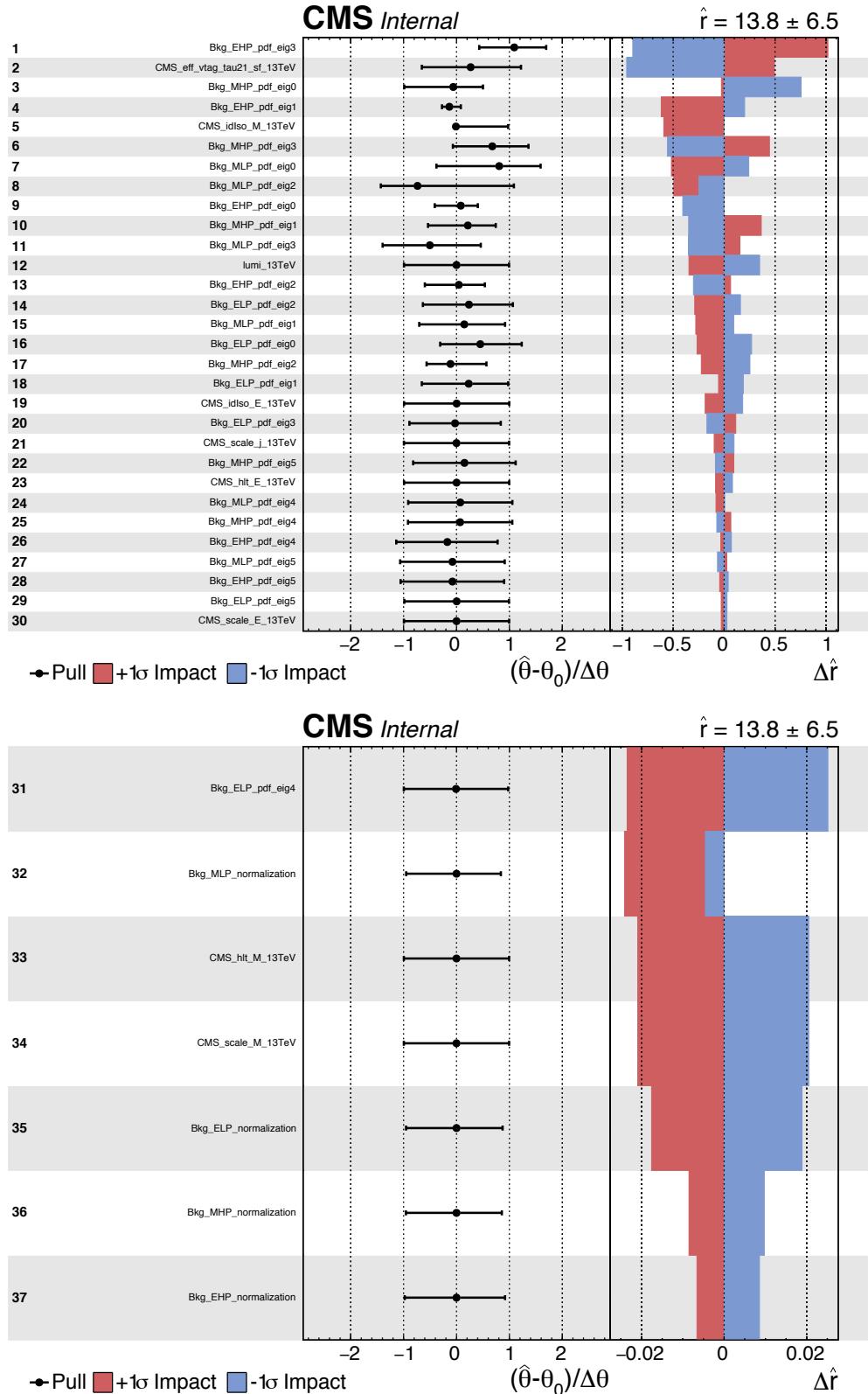


Figure 7.4: Impacts on signal strength for point at 1 TeV. The signal strength is denoted with the symbol  $\hat{r}$ , and a generic uncertainty is denoted as  $\theta$ . Each name in the vertical axis is associated with a certain systematic uncertainty, in decreasing order of relevance.

# Chapter 8

## Conclusions

The knowledge on elementary particles collected along decades of experimental searches accompanied by theoretical and technological developments, established the standard model as the most suitable theory to describe three fundamental forces: strong, weak, and electromagnetic. The gravitational force also plays an important role in the description of the universe, however, its connection with the other fundamental forces is still not well understood.

It is possible to overcome the difficulties of a common description of gravity and the standard model via an unification approach, as long as the predictions of the new theory be susceptible of experimental demonstration. In this thesis, special attention is paid to a model that not only predicts the existence of a previously unobserved particle, but also establishes the existence of extra dimensions. In fact, the bulk graviton arises in the context of warped extra dimensional models, and its existence can be revealed through the analysis of data collected by the CMS experiment.

Our search targets a heavy resonance at the TeV scale producing two leptons and one jet in the final state, coming from the intermediate decay of two vector bosons. Similar searches were performed by CMS using proton-proton collision data collected in 2012 at center-of-mass energy  $\sqrt{s} = 8$  TeV, and more recently, using 2016 data at  $\sqrt{s} = 13$  TeV.

This thesis presents the analysis of the CMS 2015 data corresponding to an integrated luminosity of  $2.7 \text{ fb}^{-1}$  collected at  $\sqrt{s} = 13$  TeV. Our results [87] are compatible with previous [31] and posterior [88, 89] searches in the mass range 0.8 - 3 TeV, showing no excess above the standard model prediction. Given the reduced statistics in data and the low cross section of the benchmark model (bulk

graviton with  $k/M_{Pl} = 0.5$ ), our analysis did not reach the sensitivity required to establish exclusion limits on the theoretical cross section of the model under study.

## Outlook of the LHC

While no new physics has yet been seen at the LHC, many models have been eliminated or have had their parameter space quite limited. The experimental approach to the searches is independent of the details of the specific production or decay patterns, and the sensitivity for observing new physics signals largely depends on the available luminosity.

There are plans to increase the luminosity of the LHC above the original design, extending the physics program for the high luminosity LHC [90]. The second phase of the LHC will provide an additional integrated luminosity of about  $2,500 \text{ fb}^{-1}$  over 20 years of operation (Fig. 8.1), enlarging the discovery potential of new particles.

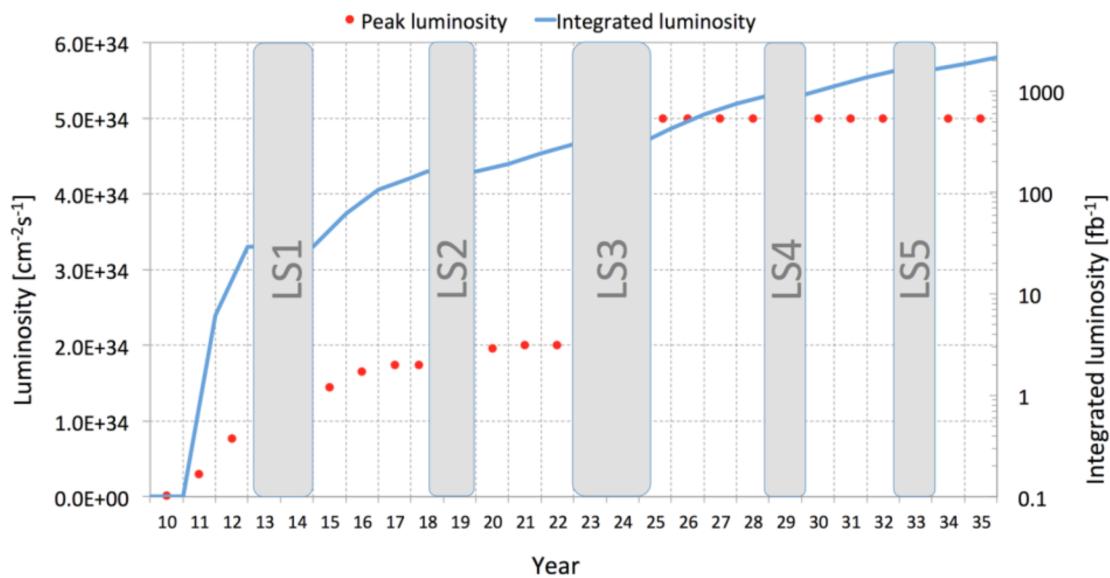


Figure 8.1: Projected LHC performance through 2035, showing preliminary dates for long shutdowns [90].

# Appendix A

## Research Internship, Presentations and Service Work

### A.1 Internship at CERN

Important advances in this thesis project were accomplished during the one-year internship at CERN, working in collaboration with the B2G physics analysis group. This internship occurred between July 2014 and June 2015, and the commissioning of the analysis framework and several optimizations were performed during this period.

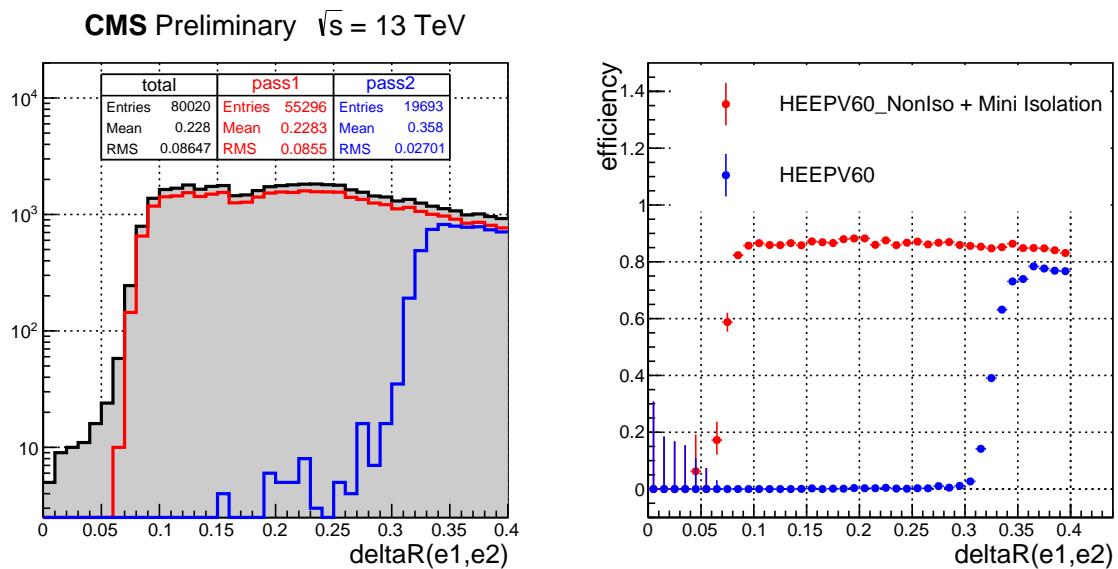


Figure A.1: Electron efficiency as function of  $\Delta R$ .

Another activity during the stay at CERN involved the studies on electron identification and isolation. In hadron colliders like the LHC, a proper isolation of electrons from the hadronic contamination is important to achieve high levels of efficiency and improve the sensitivity of the searches. The isolation of a pair of boosted electrons is particularly difficult because the presence of one electron may spoil the isolation of the other.

This behavior was studied in Monte Carlo simulations, as shown in Fig. A.1. In this context, there were proposals to apply a new technique called mini-isolation characterized with a variable isolation cone. In the beginning the mini-isolation seemed to be an acceptable solution for the problem of the isolation of boosted electrons, but there were complications regarding the performance of the technique in real data. At the end, the mini-isolation technique was not adopted, and the analysis opted for the standard particle-flow based isolation suggested by the electron physics object group.

The list of presentations given at the diboson resonance meeting, which is internal for members of the CMS Collaboration, are shown in Table A.1.

Table A.1: Presentations at CMS internal meeting.

Date	Link to the Agenda
13 Aug 2014	<a href="https://indico.cern.ch/event/334733">https://indico.cern.ch/event/334733</a>
04 Nov 2014	<a href="https://indico.cern.ch/event/348912">https://indico.cern.ch/event/348912</a>
14 Nov 2014	<a href="https://indico.cern.ch/event/338802">https://indico.cern.ch/event/338802</a>
26 Nov 2014	<a href="https://indico.cern.ch/event/349870">https://indico.cern.ch/event/349870</a>
01 Dec 2014	<a href="https://indico.cern.ch/event/355405">https://indico.cern.ch/event/355405</a>
10 Dec 2014	<a href="https://indico.cern.ch/event/357685">https://indico.cern.ch/event/357685</a>
21 Jan 2015	<a href="https://indico.cern.ch/event/367603">https://indico.cern.ch/event/367603</a>
28 Jan 2015	<a href="https://indico.cern.ch/event/369935">https://indico.cern.ch/event/369935</a>
02 Feb 2015	<a href="https://indico.cern.ch/event/369661">https://indico.cern.ch/event/369661</a>
25 Mar 2015	<a href="https://indico.cern.ch/event/383553">https://indico.cern.ch/event/383553</a>
01 Apr 2015	<a href="https://indico.cern.ch/event/384925">https://indico.cern.ch/event/384925</a>
13 Apr 2015	<a href="https://indico.cern.ch/event/387620">https://indico.cern.ch/event/387620</a>
13 May 2015	<a href="https://indico.cern.ch/event/394192">https://indico.cern.ch/event/394192</a>
20 May 2015	<a href="https://indico.cern.ch/event/395573">https://indico.cern.ch/event/395573</a>
27 May 2015	<a href="https://indico.cern.ch/event/396649">https://indico.cern.ch/event/396649</a>
01 Jul 2015	<a href="https://indico.cern.ch/event/405140">https://indico.cern.ch/event/405140</a>
06 Jul 2015	<a href="https://indico.cern.ch/event/405333">https://indico.cern.ch/event/405333</a>
15 Jul 2015	<a href="https://indico.cern.ch/event/433384">https://indico.cern.ch/event/433384</a>

## A.2 Service Work to the Collaboration

In addition to the research activities, every collaborator in CMS is appointed to a service work in order to guarantee the good operation of the experiment. In this section we describe the specific duties that were performed in 2016.

The CMS trigger system is responsible for selecting in real-time those interesting events that should be recorded for offline analysis. Every release of the CMS software (CMSSW) is accompanied with a set of validation samples with the end of monitoring the performance of individual triggers. The responsibilities of the trigger validator include:

- Make systematic comparisons between consecutive CMSSW releases.
- Maintain the validation packages for Susy and Exotica analysis groups.
- Report to the trigger studies group in charge of the strategy for trigger evolution and monitoring.

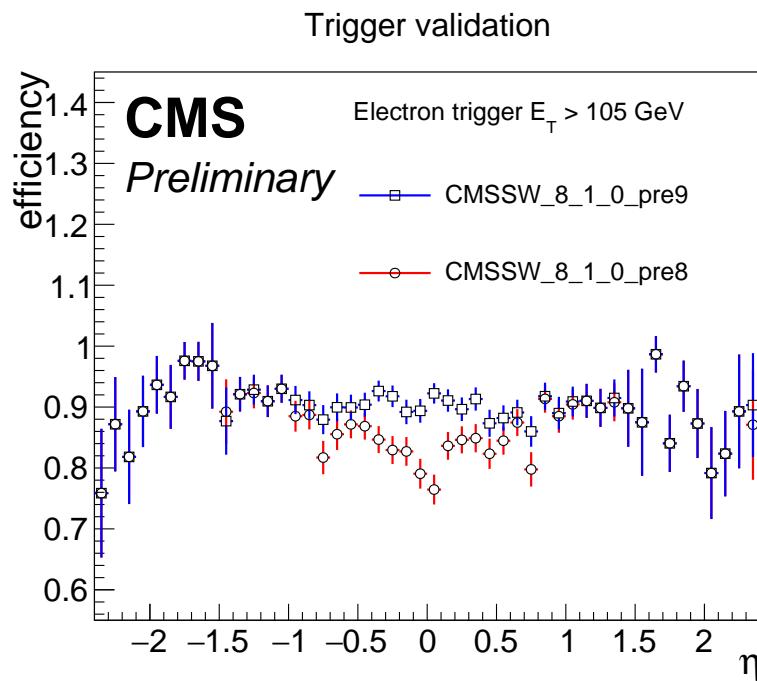


Figure A.2: Efficiency of the high energy electron trigger  $E_T > 105 \text{ GeV}$  as function of pseudorapidity  $\eta$ . The comparison of two consecutive pre-releases of the CMSSW is important to reveal either an expected change in the configuration of the trigger or the diagnostic of systematic problems.

Figure A.2 shows one of the monitor elements for trigger validation with the efficiency to select high energy electrons ( $E_T > 105$  GeV). In this particular example we observe discrepancies in the efficiency in the central region of  $\eta$ , which is an indication of either an expected change in the configuration of the trigger or an actual issue. The validations between consecutive releases of the CMSSW is important to spot as early as possible any problem that can affect the normal behaviour of the trigger. Since the CMSSW is continuously evolving, the validation has to be performed in a regular weekly basis.

Table A.2: Trigger validation campaigns in 2016.

Release Name	Date
8_1_0_PRE4	May 15
8_1_0_PRE5	May 29
8_1_0_PRE6	Jun 15
8_0_10_HLT	Jun 16
8_1_0_PRE8	Jul 15
8_1_0_PRE9	Jul 29
8_0_16	Aug 13
8_0_16_Tranch4GT	Aug 22
8_1_0_PRE10	Sep 1
8_0_19_Tranch4GT	Sep 13
8_1_0_PRE11	Sep 22
8_1_0_PRE12	Oct 11
8_1_0_PRE15	Nov 04
8_1_0_PRE16	Nov 22

### A.3 National and International Presentations

In addition to the CMS internal presentations, the results derived from this project were presented in public conferences for a wider audience.

- **November 2014:** Oral presentation at “CMS Exotica Workshop 2014”, Madrid (Spain).
- **September 2015:** Oral presentation at “XXXVI Encontro Nacional de Física de Partículas e Campos”, Caxambu, MG (Brazil) [91].

- **November 2015:** Oral presentation at II Simpósio de Física, Astronomia e Meteorologia, Bauru, SP (Brazil).
- **May 2016:** Pre-approval presentation of the analysis to the B2G conveners.
- **July 2016:** Approval of the analysis by the CMS Collaboration [87].
- **August 2016:** Poster “Search for new resonances in the merged jet + dilepton final state in CMS” presented on behalf of the CMS collaboration at the 38th International Conference on High Energy Physics (ICHEP 2016) [92], Chicago (USA) — [Proceedings of Science](#).
- **September 2016:** Oral presentation at “Encontro de Física 2016” (ENF 2016) [93], Natal, RN (Brazil).

# Bibliography

- [1] S. F. Novaes, “Standard model: An Introduction,” in *Particles and fields. Proceedings, 10th Jorge Andre Swieca Summer School, Sao Paulo, Brazil, February 6-12, 1999*, pp. 5–102. 1999. [arXiv:hep-ph/0001283](https://arxiv.org/abs/hep-ph/0001283).
- [2] T. Muta, “Foundations of quantum chromodynamics. Second edition,” *World Sci. Lect. Notes Phys.* **57** (1998) 1–409.
- [3] S. M. Bilenky and J. Hosek, “Glashow-Weinberg-Salam Theory of Electroweak Interactions and the Neutral Currents,” *Phys. Rept.* **90** (1982) 73–157.
- [4] R. P. Feynman, *QED: The Strange Theory of Light and Matter*. Princeton University Press, 1988.
- [5] P. W. Higgs, “Spontaneous Symmetry Breakdown without Massless Bosons,” *Phys. Rev.* **145** (1966) 1156–1163.
- [6] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons,” *Phys. Rev. Lett.* **13** (1964) 321–323.
- [7] S. L. Glashow, “Partial Symmetries of Weak Interactions,” *Nucl. Phys.* **22** (1961) 579–588.
- [8] **Gargamelle Neutrino** Collaboration, F. J. Hasert *et al.*, “Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment,” *Phys. Lett.* **46B** (1973) 138–140.
- [9] **CMS** Collaboration, S. Chatrchyan *et al.*, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC,” *Phys. Lett.* **B716** (2012) 30–61, [arXiv:1207.7235 \[hep-ex\]](https://arxiv.org/abs/1207.7235).

- [10] ATLAS Collaboration, G. Aad *et al.*, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC,” *Phys. Lett. B* **716** (2012) 1–29, [arXiv:1207.7214 \[hep-ex\]](https://arxiv.org/abs/1207.7214).
- [11] S. P. Martin, “A Supersymmetry primer,” [arXiv:hep-ph/9709356](https://arxiv.org/abs/hep-ph/9709356). [Adv. Ser. Direct. High Energy Phys.18,1(1998)].
- [12] C. Csaki, “TASI lectures on extra dimensions and branes,” in *From fields to strings: Circumnavigating theoretical physics. Ian Kogan memorial collection (3 volume set)*, pp. 605–698. 2004. [arXiv:hep-ph/0404096](https://arxiv.org/abs/hep-ph/0404096).
- [13] K. Agashe, R. Contino, and A. Pomarol, “The Minimal composite Higgs model,” *Nucl. Phys. B* **719** (2005) 165–187, [arXiv:hep-ph/0412089](https://arxiv.org/abs/hep-ph/0412089).
- [14] L. Randall and R. Sundrum, “A Large mass hierarchy from a small extra dimension,” *Phys. Rev. Lett.* **83** (1999) 3370, [arXiv:hep-ph/9905221](https://arxiv.org/abs/hep-ph/9905221).
- [15] L. Randall and R. Sundrum, “An Alternative to compactification,” *Phys. Rev. Lett.* **83** (1999) 4690, [arXiv:hep-th/9906064](https://arxiv.org/abs/hep-th/9906064).
- [16] K. Agashe, H. Davoudiasl, G. Perez, and A. Soni, “Warped Gravitons at the LHC and Beyond,” *Phys. Rev. D* **76** (2007) 036006, [arXiv:hep-ph/0701186](https://arxiv.org/abs/hep-ph/0701186).
- [17] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, “Techniques for improved heavy particle searches with jet substructure,” *Phys. Rev. D* **80** (2009) 051501, [arXiv:0903.5081 \[hep-ph\]](https://arxiv.org/abs/0903.5081).
- [18] G. t Hooft, “Gauge theories of the forces between elementary particles,” *Scientific American* **242** no. 6, (1980) 104–138.
- [19] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. Wiley, USA, 1984.
- [20] C. Quigg, *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions*. Princeton University Press, USA, 2013.
- [21] A. Pich, “The Standard Model of Electroweak Interactions,” in *Proceedings, High-energy Physics. Proceedings, 18th European School (ESHEP 2010): Raseborg, Finland, June 20 - July 3, 2010*, pp. 1–50. 2012. [arXiv:1201.0537 \[hep-ph\]](https://arxiv.org/abs/1201.0537).

- [22] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, “Evidence for the  $2\pi$  Decay of the  $K_2^0$  Meson,” *Phys. Rev. Lett.* **13** (1964) 138–140.
- [23] M. Kobayashi and T. Maskawa, “CP Violation in the Renormalizable Theory of Weak Interaction,” *Prog. Theor. Phys.* **49** (1973) 652–657.
- [24] P. Langacker, *The standard model and beyond*. CRC Press, USA, 2010.
- [25] C. Quigg, “Unanswered Questions in the Electroweak Theory,” *Ann. Rev. Nucl. Part. Sci.* **59** (2009) 505–555, [arXiv:0905.3187 \[hep-ph\]](https://arxiv.org/abs/0905.3187).
- [26] B. C. Allanach *et al.*, “Les Houches physics at TeV colliders 2005 beyond the standard model working group: Summary report,” in *Physics at TeV colliders. Proceedings, Workshop, Les Houches, France, May 2-20, 2005*. 2006. [arXiv:hep-ph/0602198 \[hep-ph\]](https://arxiv.org/abs/hep-ph/0602198).
- [27] M. Bustamante, L. Cieri, and J. Ellis, “Beyond the Standard Model for Montaneros,” in *High-energy physics. Proceedings, 5th CERN-Latin-American School, Recinto Quirama, Colombia, March 15-28, 2009*. 2009. [arXiv:0911.4409 \[hep-ph\]](https://arxiv.org/abs/0911.4409).
- [28] K. R. Dienes, “New directions for new dimensions: An introduction to Kaluza-Klein theory, large extra dimensions, and the brane world,” in *Particle physics and cosmology: The quest for physics beyond the standard model(s). Proceedings, Theoretical Advanced Study Institute, TASI 2002, Boulder, USA, June 3-28, 2002*, pp. 447–545. 2002.
- [29] **CMS** Collaboration, S. Chatrchyan *et al.*, “Search for exotic resonances decaying into  $WZ/ZZ$  in  $pp$  collisions at  $\sqrt{s} = 7$  TeV,” *JHEP* **02** (2013) 036, [arXiv:1211.5779 \[hep-ex\]](https://arxiv.org/abs/1211.5779).
- [30] **CMS** Collaboration, V. Khachatryan *et al.*, “Search for massive resonances in dijet systems containing jets tagged as W or Z boson decays in  $pp$  collisions at  $\sqrt{s} = 8$  TeV,” *JHEP* **08** (2014) 173, [arXiv:1405.1994 \[hep-ex\]](https://arxiv.org/abs/1405.1994).
- [31] **CMS** Collaboration, V. Khachatryan *et al.*, “Search for massive resonances decaying into pairs of boosted bosons in semi-leptonic final states at  $\sqrt{s} = 8$  TeV,” *JHEP* **08** (2014) 174, [arXiv:1405.3447 \[hep-ex\]](https://arxiv.org/abs/1405.3447).

- [32] CMS Collaboration, V. Khachatryan *et al.*, “Search for Narrow High-Mass Resonances in Proton-Proton Collisions at  $\sqrt{s} = 8$  TeV Decaying to a Z and a Higgs Boson,” *Phys. Lett. B* **748** (2015) 255–277, [arXiv:1502.04994 \[hep-ex\]](https://arxiv.org/abs/1502.04994).
- [33] CMS Collaboration, V. Khachatryan *et al.*, “Search for heavy resonances decaying to two Higgs bosons in final states containing four b quarks,” *Eur. Phys. J. C* **76** no. 7, (2016) 371, [arXiv:1602.08762 \[hep-ex\]](https://arxiv.org/abs/1602.08762).
- [34] CMS Collaboration, V. Khachatryan *et al.*, “Search for two Higgs bosons in final states containing two photons and two bottom quarks in proton-proton collisions at 8 TeV,” *Phys. Rev. D* **94** no. 5, (2016) 052012, [arXiv:1603.06896 \[hep-ex\]](https://arxiv.org/abs/1603.06896).
- [35] “CMS Beyond-two-generations public physics results, <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsB2G.”>
- [36] T. Kaluza, “On the Problem of Unity in Physics,” *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)* **1921** (1921) 966–972.
- [37] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, “MadGraph 5 : going beyond,” *JHEP* **06** (2011) 128, [arXiv:1106.0522 \[hep-ph\]](https://arxiv.org/abs/1106.0522).
- [38] A. Oliveira, “Gravity particles from Warped Extra Dimensions, a review. Part I - KK Graviton,” [arXiv:1404.0102 \[hep-ph\]](https://arxiv.org/abs/1404.0102).
- [39] CMS Collaboration, A. M. Sirunyan *et al.*, “Combination of searches for heavy resonances decaying to WW, WZ, ZZ, WH, and ZH boson pairs in proton-proton collisions at  $\sqrt{s} = 8$  and 13 TeV,” [arXiv:1705.09171 \[hep-ex\]](https://arxiv.org/abs/1705.09171).
- [40] E. A. Mobs, “The CERN accelerator complex. Complexe des accélérateurs du CERN.”. <https://cds.cern.ch/record/2225847>. General Photo.
- [41] O. S. Bruning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report*. CERN, Geneva, 2004.  
<https://cds.cern.ch/record/782076>.

- [42] "Public CMS Luminosity Information,  
<http://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>".
- [43] **CMS** Collaboration, G. L. Bayatian *et al.*, "CMS Physics: Technical Design Report Volume 1: Detector Performance and Software,".  
<http://cds.cern.ch/record/922757>.
- [44] **CMS** Collaboration, G. L. Bayatian *et al.*, "CMS Physics: Technical Design Report Volume 2: Physics Performance," *J. Phys. G* **34** no. CERN-LHCC-2006-021. CMS-TDR-8-2, (2007) 995–1579. 669 p.  
<http://cds.cern.ch/record/942733>. revised version submitted on 2006-09-22 17:44:47.
- [45] "The Compact Muon Solenoid (CMS). Virtual tour of CMS,  
<http://home.cern/about/experiments/cms>".
- [46] A. Dominguez, Abbaneo, *et al.*, "CMS Technical Design Report for the Pixel Detector Upgrade," Tech. Rep. CERN-LHCC-2012-016. CMS-TDR-11, CERN, Sep, 2012. <https://cds.cern.ch/record/1481838>.
- [47] **CMS** Collaboration, S. Dasu *et al.*, "CMS. The TriDAS project. Technical design report, vol. 1: The trigger systems,".  
<https://cds.cern.ch/record/706847>.
- [48] **CMS** Collaboration, P. Sphicas, "CMS: The TriDAS project. Technical design report, Vol. 2: Data acquisition and high-level trigger,".  
<http://cds.cern.ch/record/578006>.
- [49] T. Bawej *et al.*, "The New CMS DAQ System for Run-2 of the LHC," *IEEE Trans. Nucl. Sci.* **62** no. 3, (2015) 1099–1103. [,7097437(2014)].
- [50] **CMS** Collaboration, V. Khachatryan *et al.*, "CMS Tracking Performance Results from early LHC Operation," *Eur. Phys. J. C* **70** (2010) 1165–1192, [arXiv:1007.1988 \[physics.ins-det\]](https://arxiv.org/abs/1007.1988).
- [51] T. Mc Cauley, "Collisions recorded by the CMS detector on 14 Oct 2016 during the high pile-up fill,". <https://cds.cern.ch/record/2231915>.

- [52] R. Fruhwirth, P. Kubinec, W. Mitaroff, and M. Regler, “Vertex reconstruction and track bundling at the LEP collider using robust algorithms,” *Comput. Phys. Commun.* **96** (1996) 189–208.
- [53] R. Fruhwirth, W. Waltenberger, K. Prokofiev, T. Speer, and P. Vanlaer, “New developments in vertex reconstruction for CMS,” *Nucl. Instrum. Meth. A* **502** (2003) 699–701.
- [54] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- $k_t$  jet clustering algorithm,” *JHEP* **04** (2008) 063, [arXiv:0802.1189 \[hep-ex\]](https://arxiv.org/abs/0802.1189).
- [55] **CMS** Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET,” Tech. Rep. CMS-PAS-PFT-09-001, CERN, Geneva, 2009.
- [56] **CMS** Collaboration, “Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector,” Tech. Rep. CMS-PAS-PFT-10-001, CERN, Geneva, 2010.
- [57] **CMS** Collaboration, V. Khachatryan *et al.*, “Identification techniques for highly boosted W bosons that decay into hadrons,” *JHEP* **12** (2014) 017, [arXiv:1410.4227 \[hep-ex\]](https://arxiv.org/abs/1410.4227).
- [58] **CMS** Collaboration, S. Chatrchyan *et al.*, “Studies of jet mass in dijet and W/Z + jet events,” *JHEP* **05** (2013) 090, [arXiv:1303.4811 \[hep-ex\]](https://arxiv.org/abs/1303.4811).
- [59] J. Thaler and K. Van Tilburg, “Maximizing Boosted Top Identification by Minimizing N-subjettiness,” *JHEP* **02** (2012) 093, [arXiv:1108.2701 \[hep-ph\]](https://arxiv.org/abs/1108.2701).
- [60] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, “ALPGEN, a generator for hard multiparton processes in hadronic collisions,” *JHEP* **07** (2003) 001, [arXiv:hep-ph/0206293](https://arxiv.org/abs/hep-ph/0206293).
- [61] F. Maltoni and T. Stelzer, “MadEvent: Automatic event generation with MadGraph,” *JHEP* **02** (2003) 027, [arXiv:hep-ph/0208156](https://arxiv.org/abs/hep-ph/0208156).
- [62] T. Sjostrand, S. Mrenna, and P. Z. Skands, “A Brief Introduction to PYTHIA 8.1,” *Comput. Phys. Commun.* **178** (2008) 852–867, [arXiv:0710.3820 \[hep-ph\]](https://arxiv.org/abs/0710.3820).

- [63] “São Paulo Research and Analysis Center, <http://sprace.org.br>.”
- [64] **CMS** Collaboration, “MUSiC, a Model Unspecific Search for New Physics, in pp Collisions at  $\sqrt{s} = 8$  TeV,” Tech. Rep. CMS-PAS-EXO-14-016, CERN, Geneva, 2017. <https://cds.cern.ch/record/2256653>.
- [65] **CMS** Collaboration, “Search for a narrow spin-2 resonance decaying to Z bosons in the semileptonic final state,” Tech. Rep. CMS-PAS-EXO-12-022, CERN, Geneva, 2013. <https://cds.cern.ch/record/1596494>.
- [66] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations,” *JHEP* **07** (2014) 079, [arXiv:1405.0301 \[hep-ph\]](arXiv:1405.0301 [hep-ph]).
- [67] P. Nason, “A New method for combining NLO QCD with shower Monte Carlo algorithms,” *JHEP* **11** (2004) 040, [arXiv:hep-ph/0409146 \[hep-ph\]](arXiv:hep-ph/0409146 [hep-ph]).
- [68] S. Frixione, P. Nason, and C. Oleari, “Matching nlo qcd computations with parton shower simulations: the powheg method,” *JHEP* **11** (2007) 070.
- [69] S. Alioli, P. Nason, C. Oleari, and E. Re, “Nlo vector-boson production matched with shower in powheg,” *JHEP* **07** (2008) 06.
- [70] T. Sjöstrand, S. Mrenna and P. Z. Skands, “PYTHIA 6.4 Physics and Manual,” *JHEP* **05** (2006) 026, <arXiv:hep-ph/0603175>.
- [71] T. Sjöstrand, S. Mrenna, and P. Z. Skands, “A Brief Introduction to PYTHIA 8.1,” *Comput. Phys. Commun.* **178** (2008) 852–867, [arXiv:0710.3820 \[hep-ph\]](arXiv:0710.3820 [hep-ph]).
- [72] P. Skands, S. Carrazza, and J. Rojo, “Tuning PYTHIA 8.1: the Monash 2013 Tune,” *Eur. Phys. J. C* **74** no. 8, (2014) 3024, [arXiv:1404.5630 \[hep-ph\]](arXiv:1404.5630 [hep-ph]).
- [73] **NNPDF** Collaboration, R. D. Ball *et al.*, “Parton distributions for the LHC Run II,” *JHEP* **04** (2015) 040, [arXiv:1410.8849 \[hep-ph\]](arXiv:1410.8849 [hep-ph]).

- [74] **GEANT4** Collaboration, S. Agostinelli *et al.*, “GEANT4: a simulation toolkit,” *Nucl. Instrum. Meth.* **A506** (2003) 250.
- [75] **CMS** Collaboration, “Measurement of the inelastic proton-proton cross section at  $\sqrt{s} = 13$  TeV,” Tech. Rep. CMS-PAS-FSQ-15-005, CERN, Geneva, 2016. <https://cds.cern.ch/record/2145896>.
- [76] “MiniAOD Analysis Documentation,  
<http://twiki.cern.ch/twiki/bin/view/CMS/WorkBookMiniAOD>.”.
- [77] **CMS** Collaboration, S. Chatrchyan *et al.*, “Description and performance of track and primary-vertex reconstruction with the CMS tracker,” *JINST* **9** no. 10, (2014) P10009, [arXiv:1405.6569 \[physics.ins-det\]](https://arxiv.org/abs/1405.6569).
- [78] “E/gamma Physics Object Group, *Internal Twiki*,  
<http://twiki.cern.ch/twiki/bin/view/CMS/EgammaPOG>.”.
- [79] “Muon Physics Object Group, *Internal Twiki*,  
<http://twiki.cern.ch/twiki/bin/view/CMS/MuonPOG>.”.
- [80] “JetMET Physics Object Group, *Internal Twiki*,  
<http://twiki.cern.ch/twiki/bin/view/CMS/JetMET>.”.
- [81] M. U. Mozer, “Electroweak Physics at the LHC,” *Springer Tracts Mod. Phys.* **267** (2016) 1–115.
- [82] I. Antcheva *et al.*, “ROOT: A C++ framework for petabyte data storage, statistical analysis and visualization,” *Comput. Phys. Commun.* **182** (2011) 1384–1385.
- [83] W. Verkerke and D. P. Kirkby, “The RooFit toolkit for data modeling,” *eConf* **C0303241** (2003) MOLT007, [arXiv:physics/0306116](https://arxiv.org/abs/physics/0306116).
- [84] **CMS** Collaboration, “CMS Luminosity Measurement for the 2015 Data Taking Period,” Tech. Rep. CMS-PAS-LUM-15-001, CERN, Geneva, 2016.
- [85] L. Lista, “Statistical Methods for Data Analysis in Particle Physics,” *Lect. Notes Phys.* **909** (2016) pp.1–172.

- [86] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics,” *Eur. Phys. J.* **C71** (2011) 1554, [arXiv:1007.1727 \[physics.data-an\]](https://arxiv.org/abs/1007.1727). [Erratum: Eur. Phys. J.C73,2501(2013)].
- [87] CMS Collaboration, “Search for diboson resonances in the semileptonic  $X \rightarrow ZV \rightarrow \ell^+\ell^- q\bar{q}$  final state at  $\sqrt{s} = 13$  TeV with CMS,” Tech. Rep. CMS-PAS-B2G-16-010, CERN, Geneva, 2016. <http://cds.cern.ch/record/2199611>.
- [88] CMS Collaboration, “Search for heavy resonances decaying into a Z boson and a W boson in the  $\ell^+\ell^- q\bar{q}$  final state,” Tech. Rep. CMS-PAS-B2G-16-022, CERN, Geneva, 2017. <http://cds.cern.ch/record/2242955>.
- [89] CMS Collaboration, “Search for new diboson resonances in the dilepton + jets final state at  $\sqrt{s} = 13$  TeV with 2016 data,” Tech. Rep. CMS-PAS-HIG-16-034, CERN, Geneva, 2017. <http://cds.cern.ch/record/2243295>.
- [90] CMS Collaboration, “Technical Proposal for the Phase-II Upgrade of the CMS Detector,” Tech. Rep. CERN-LHCC-2015-010. LHCC-P-008. CMS-TDR-15-02, CERN, Geneva, Jun, 2015. <https://cds.cern.ch/record/2020886>.
- [91] J. C. Ruiz Vargas, “Search for Physics Beyond the Standard Model in the VZ channel at the CMS Experiment,” in *XXXVI Encontro Nacional de Física de Partículas e Campos, Caxambu, MG (Brasil)*. Caxambu, 2015.
- [92] J. C. Ruiz Vargas, “Search for new resonances in the merged jet + dilepton final state in CMS,” in *Proceedings, 38th International Conference on High Energy Physics (ICHEP 2016)*, vol. ICHEP2016. 2016.
- [93] J. C. Ruiz Vargas, “Search for high mass resonances in the dilepton + jet final state in CMS,” in *Encontro de Física, Natal, RN (Brasil)*. Natal, 2016.