

UNIVERSITÉ LIBRE DE BRUXELLES Faculté des Sciences

Département de Physique

Measurement of Z boson production in association with jets at the LHC and study of a DAQ system for the Triple-GEM detector in view of the CMS upgrade

Thèse présentée par

Alexandre LÉONARD

En vue de l'obtention du grade de

Docteur en Sciences

Bruxelles, 2015

Research funded by a ASP - F.R.S.-FNRS Research Fellow grant of the *Fund for Scientific Research - FNRS*. 2015 Alexandre Léonard Doctoral examination commission:

- Dr. Ulla BLUMENSCHEIN (GAU Göttingen)
- Prof. Dr. Gilles DE LENTDECKER (ULB) (president)
- Prof. Dr. Laurent FAVART (ULB) (supervisor)
- Dr. Philippe GRAS (IRFU CEA-Saclay)
- Prof. Dr. Fabio MALTONI (UCL)
- Prof. Dr. Serge MASSAR (ULB) (secretary)

Abstract:

This PhD thesis presents the measurement of the differential cross section for the production of a Z boson in association with jets in proton-proton collisions taking place at the Large Hadron Collider (LHC) at CERN, at a centre-of-mass energy of 8 TeV. A development of a data acquisition (DAQ) system for the Triple-Gas Electron Multiplier (GEM) detector in view of the Compact Muon Solenoid (CMS) detector upgrade is also presented.

The events used for the data analysis were collected by the CMS detector during the year 2012 and constitute a sample of $19.6 \,\mathrm{fb^{-1}}$ of integrated luminosity. The cross section measurements are performed as a function of the jet multiplicity, the jet transverse momentum and pseudorapidity, and the scalar sum of the jet transverse momenta. The results were obtained by correcting the observed distributions for detector effects. The measured differential cross sections are compared to some state of the art Monte Carlo predictions MADGRAPH 5, SHERPA 2 and MADGRAPH5_AMC@NLO.

These measurements provide an important contribution to the understanding of the perturbative quantum chromodynamics theory. Additionally the highest energies ever obtained in laboratories and the exceptional proper functioning of the LHC make possible to explore regions of the phase space never reached so far.

Following the LHC machine development plan, the accelerator is expected to provide a higher and higher instantaneous luminosity for the near future and new detector technologies must be studied to handle the large particle rates that are expected. A chapter presenting the development of a DAQ system of the CMS muon detection and triggering system upgrade with Triple-GEM detectors is also included in this thesis. A full experimental set-up has been built at the Interuniversity Institute for High Energies (IIHE) (ULB-VUB) for testing of Triple-GEM prototypes and to show the feasibility of the project.

First results were obtained with cosmic muons at the IIHE as well as muons and pions in test beams at CERN and confirm the validity and feasibility of the Triple-GEM project for the CMS upgrade. ii

Acknowledgement:

First of all I would like to express my profound gratitude to my supervisor Prof. Laurent FAVART. I have known him since my first bachelor year in physics where he was teaching the fundamentals of experimental physics through the laboratory exercises. A few years later he made me discover data analysis by studying the last *ep* collisions recorded at HERA which became afterwards the subject of my master thesis that I realised under his supervision. I then started my PhD and I am now finally expressing my gratefulness for his support, his expert guidance, his understanding and his encouragement throughout my PhD but also throughout my bachelor and master studies. I also want to thank Prof. Gilles DE LENTDECKER, for having given me the opportunity to work on the GEM project and to be a bit closer to the CMS detector.

This PhD would not be what it is without my post-doctoral researcher Dr. Tomislav ŠEVA. He has started the data analysis presented in my thesis, and I have to warmly thank him for the work we did together. Among the other people who brought uncountable effort for this analysis and with whom I had great time debugging the code between two coffee breaks, I would like to thank Dr. Philippe GRAS and my workmate Bugra BILIN.

Now comes the place to cite the people I spent my time with along the four years of my thesis. I would like to improvise something for the great Thierry MAERSCHALK but I am not as good as him in this field so it will just be a big thank you for the flying disc. To Florian ZENONI, my office mate for some time with whom I spent unforgettable moments, I would like to tell him grazie mille for the good company he is. I am also grateful to Thomas LENZI, another PhD student who surprised me with his wide knowledge and with whom I spend a lot of time when working on the Triple-GEM DAQ system as well as drinking tea. I cannot forget Dr. Erik VERHAGEN who spent many hours explaining me different aspects of the electronics and with whom I realised my first weldings. I hope we can dance JAVA together again. I would like to thank Dr. Laurent THOMAS with whom I spend some crazy hunting nights as well as morning races and who was an excellent, if not the best, coffee provider.

I am also thankful to many other people of the lab but I am afraid the list would be too long and non-exhaustive if I was to write it. I am thinking about all my office mates, the IT team and the secretary, and all the other physicists of the lab.

Last but not least, I would like to thank my lovely Kamila KOZIRÓG who arrived in my life right before I started my PhD and who had to endure a great amount of complaints but who has always supported me to achieve my goals. I would also like to thank my family, my parents and my sisters and brother for having given me the opportunity to be who I am and for all the love they have shown me up to now.

Contents

Introduction			
The	Theory Elements		
2.1	Partic	le Content	5
2.2	2 The Equations Governing the Interactions		
	2.2.1	The Lagrangian of the Electroweak Interaction	8
	2.2.2	The Lagrangian of Quantum Chromodynamics	11
	2.2.3	The Mass Terms and the Brout-Englert-Higgs Mechanism $\ . \ . \ .$	13
2.3	.3 The Running of the Coupling Constants		
	2.3.1	The Beta Function	18
2.4	$Z + Jets$ in Proton-Proton Collisions $\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$		
	2.4.1	The QCD Factorisation Theorem	21
	2.4.2	Application to the Drell-Yan Process	22
	2.4.3	Parton Distribution Functions	30
3 MC Simulations and Recent Results		ations and Recent Results	35
3.1	Monte	Carlo Generators	36
	3.1.1	Pythia	38
	3.1.2	MadGraph 5	45
	Intr The 2.1 2.2 2.3 2.4 MC 3.1	Introduction The ory El 2.1 Particion 2.2 The E 2.2 2.2.1 2.2.2 2.2.3 2.3 The R 2.3 The R 2.3.1 2.4.1 2.4.2 2.4.2 2.4.3 Simulation 3.1 Montee 3.1.1 3.1.2	Introduction Theory Elements 2.1 Particle Content

CONTENTS

		3.1.3	Sherpa	48
		3.1.4	MadGraph5_aMC@NLO	49
		3.1.5	Matching of Matrix Element and Parton Shower	49
		3.1.6	Pile-up	53
		3.1.7	The Use of Different MC	53
	3.2	LHC a	and Tevatron Measurements	53
		3.2.1	Main Motivation	54
		3.2.2	LHC and Tevatron	55
		3.2.3	Recent Results	57
4	Experimental Set-up			65
	4.1	The L	arge Hadron Collider	65
		4.1.1	History and Goals of the Collider	65
		4.1.2	Technical Description	67
		4.1.3	The LHC beams	68
		4.1.4	Instantaneous and Integrated Luminosities	68
		4.1.5	The Injector Complex	69
	4.2	The C	MS Detector	74
		4.2.1	The Geometry of CMS and its Principles	75
		4.2.2	The Superconducting Solenoid Magnet	77
		4.2.3	The Tracker	79
		4.2.4	The Calorimeters	82
		4.2.5	The Muon Detectors	89
		4.2.6	The Trigger System	93

		4.2.7	Treatment of the Data	97	
5	Eve	ents Reconstruction			
	5.1	Track	and Vertex Reconstruction	100	
	5.2	Partic	le Flow	101	
		5.2.1	Iterative Tracking	102	
		5.2.2	Calorimeter Clustering	102	
		5.2.3	Link Algorithm	103	
	5.3	Leptor	ns	104	
		5.3.1	Muons	104	
		5.3.2	Electrons	106	
		5.3.3	Taus	108	
		5.3.4	Neutrinos	110	
	5.4	Photo	ns	110	
	5.5	Jets		111	
		5.5.1	Jet Algorithm	111	
		5.5.2	Jet Energy Correction	113	
6	$Z \mathbf{B}$	5.5.2 Jet Energy Correction 11 Z Boson Production in Association With Jets 11	117		
	6.1	Introd	uction	117	
		6.1.1	Why Study this Process ?	117	
		6.1.2	Measured Observables	118	
	6.2	Data a	and Simulation Samples	119	
	6.3	Event	Selection	121	
		6.3.1	Trigger	121	

CONTENTS

		6.3.2	Leptons	123
		6.3.3	Z bosons	128
		6.3.4	Jets	128
		6.3.5	Pile-up	130
	6.4	Detect	or Level Results	132
		6.4.1	Data-Simulation Comparison	133
		6.4.2	Background Subtraction	153
	6.5	Detect	or Effect Corrections	154
		6.5.1	Phase Space at Generator Level	154
		6.5.2	The Unfolding Method	156
		6.5.3	Response Matrices Overview	158
		6.5.4	Baye's Unfolding	163
	6.6	Uncert	ainties	164
6.7 Single Channel Differential Cross Section			Channel Differential Cross Section	171
		6.7.1	Channels Comparison	171
		6.7.2	Full Single Channel Results	173
	6.8	Combi	ned Differential Cross Sections	177
		6.8.1	Combination	177
		6.8.2	Results	178
	6.9	Conclu	isions	190
7	ጥዜ	Tu:1-	CEM Detectors	101
1	1 ne	Tuble	-GEM Detectors	191
	7.1	Motiva	ation	191
	7.2	Gas El	lectron Multiplier Detector	193

		7.2.1	Gas Detectors	193	
		7.2.2	Triple-GEM Detector Geometry	195	
		7.2.3	Triple-GEM Front-End Electronic	196	
	7.3	Impler	nentation of a DAQ System	199	
		7.3.1	DAQ Design	199	
		7.3.2	Protocols	202	
		7.3.3	Data Readout of the Cosmic Muons Test Bench	207	
		7.3.4	Cosmic Muons Test Bench Results	214	
	7.4	Tests I	Beam	219	
		7.4.1	Set-up	219	
		7.4.2	Results	219	
	7.5	Conclu	isions	221	
8	Con	clusio	ns	225	
Appendices 23					
A	Dre	ll-Yan	n Cross Section		
	A.1	Virtua	l Photon Production	241	
	A.2	Z Bos	on Production	244	
	A.3	Interfe	erence	247	
	A.4	Drell-	Yan Cross Section	249	
в	Sing	gle Cha	annel Differential Cross Sections	251	

CONTENTS

Chapter 1

Introduction

The Universe has existed for 14 billion years, yet very little is known about it even though our knowledge regarding the laws governing the interaction of its fundamental constituents as well as its behaviour as a whole keeps increasing. In the 5th century B.C. Greek philosophers first suggested a Universe made of indivisible constituents. Although great thinkers, the Greek philosophers do not exactly belong to what we call scientists as the only way they proved or disproved their hypotheses was by relying on arguments. In the 16th century A.D., Galileo was probably one of the first modern scientists in the sense that he performed experiments to test his assumptions. The scientific method adopted by Galileo and still in use nowadays relies on three main axles: establishing a theory founded on hypotheses, deriving predictions from it as logical consequences and making relevant experiments to test those predictions. This results in an ongoing process for which hypotheses, predictions and experiments are all needed to develop each other.

Following the scientific method, scientists finally demystified the atomic assumption from the Greek philosophers. The atom and its smallest (up to now) constituents were discovered along the 19^{th} and 20^{th} centuries and conducted to the elaboration of a theory called the Standard Model (SM) of particle physics in the middle of the 1970s. This theory proposes an approach of the elementary particles based on gauge groups and gauge invariance. All parameters, such as couplings and basic ingredients that are the existing particles have to be determined experimentally. This approach has survived several decades of precision measurements and had to be extended for the existence of new elementary particles (heavy leptons, heavy quarks) but its predictions were never in real conflict with observations and recently one of the last SM predictions, namely the existence of the Higgs boson, was confirmed by two experiments: A Toroidal LHC ApparatuS (ATLAS) and Compact Muon Solenoid (CMS), at the proton-proton (*pp*) Large Hadron Collider (LHC) at CERN. The LHC has offered for the first time pp collisions at a centre-of-mass energy of 8 TeV, allowing for testing of theory predictions in phase space never reached before. In this thesis, a detailed study of the production rate of Z boson in association with jets, in pp collisions at the LHC and detected by the CMS detector is performed. At the LHC the hadronic nature of the collisions and the high energy of the beams are such that jets are present in almost every studied processes with multiplicities higher than ever observed in previous experiments. Studying the production of jets is thus important and can be performed by analysing their production with a Z boson easily detectable and whose characteristics are well known. For high energy, if a high precision needs to be achieved the calculations are getting so difficult that the present question is more "how to make reliable precise predictions from the SM" than to test the validity of the SM. The measurements are thus compared to SM predictions obtained by state of the art Monte Carlo (MC) programs. The scientific method is therefore completed.

Scientists are curious by nature and the last LHC results are of course not the end of the story. The LHC has only started and the energy of its collisions is expected to increase for the next runs to 13 TeV in the centre-of-mass extending the region of exploration. An increase in the rate of collisions is also scheduled allowing for a larger amount of data. To handle such high energies and rates of collisions the CMS detector has to be upgraded too. New technologies must be used to this aim. A second subject treated in this thesis concerns the upgrade of the muon detection system in view of the large luminosity expected in few years at the LHC. A full experimental set-up to study the Triple-GEM detectors to be installed inside the CMS detector has been designed to this aim. My contribution mainly concerns the data acquisition (DAQ) system used for this project.

This thesis is subdivided into eight chapters the first being the present introduction. The second chapter introduces the theory elements related to the SM mentioned above, and in particular to its application to the production of Z bosons in association with jets in pp collisions studied in the data analysis chapter of this thesis. The third chapter discusses the way MC predictions are obtained from different generators as well as the recent results in the field of vector boson associated with jets. The experimental set-up is introduced in chapter four which describes the LHC machine and the CMS detector used to collect data. The way raw data recorded by the detector are translated to higher-level objects such as particle candidates usable for an analysis purposes is explained in chapter five. The sixth chapter presents in details the data analysis of the production of Z bosons in association with jets in pp collisions recorded by the CMS detector in year 2012 at a centre-of-mass energy of 8 TeV. Chapter seven concerns the DAQ development for the study of the Triple-GEM detector taking part in the GEM project for the upgrade of the muon detection system of CMS. Finally, conclusions are drawn in chapter eight. A

list of reference is available after the conclusion chapter. Note that some references are accessible to CMS members only providing details on technical aspects of data treatment and on data analysis. Those are links to CMS web pages and CMS notes. Though not public information those contain valuable informations for CMS readers. Two appendices can be found at the end of this thesis. The first one contains the complete leading-order calculation for the Drell-Yan (DY) process. The second appendix lists the results of the differential cross section measurements and comparisons to MC simulations separately for the muon and electron Z decay channels used for the combined results presented in chapter six.

CHAPTER 1. INTRODUCTION

Chapter 2

Theory Elements

The hypothesis that everything is made of a small number of fundamental constituents, that we refer to as elementary particles, dates from the time of the Greek philosophers such as Leucippe, Democritus and Aristotle. The former is found to be at the very beginning of a theory in which the whole universe is made of atoms (from the Greek $\alpha \tau o \mu o \varsigma$: atomos, meaning indivisible). Even though our current usage of the word atom refers to the smallest unit that defines the chemical elements, which we now know are made of electrons orbiting around a nucleus composed of protons and neutrons themselves built up of even smaller constituents called quarks, this theory of elementary particles has been studied for centuries and is the essence of the SM, a theory with unprecedented success in describing to very high level of accuracy as well as predicting various experimental observations.

This chapter is inspired by many excellent textbooks that treat this subject in much more details. Among these references, the main ones are [1–4]. More focussed references will be cited in the text when appropriate.

2.1 Particle Content

The SM is the theory that provides the best description of the interactions between elementary particles constituting our universe. This quantum field theory (a consistent theory based on both quantum mechanics and special relativity) has been developed since the second half of the 20th century and is still studied across the world. This model is successful in describing three of the four known fundamental forces, namely the electromagnetic force, the weak force (these two forces are actually unified and described by the electroweak interaction in which they are seen as two different aspects of the



Figure 2.1: Overview of the fundamental particles of the SM. Matter is made of quarks (in purple) and leptons (in green). Gauge bosons (in orange) are the force carriers. The Higgs boson (in yellow) is responsible for the mass of each of these fundamental particles. The mass of the Higgs boson refers to the one of the recently discovered Higgs-like boson [6].

same force) and the strong force. The fourth force is gravity. From the work done by Albert Einstein in the 1910's [5] a proper description of the gravitation must be done within the framework of the general theory of relativity that generalises special relativity and Newton's law of universal gravitation. The SM does not include general relativity and therefore can not describe gravitation effects. However, such effects are orders of magnitude smaller than the weakest effects of the weak interactions and ignorance of gravitation in high energy physics measurements is a perfectly valid approximation.

According to the SM, matter is constituted of two types of elementary particles: quarks and leptons, both being spin- $\frac{1}{2}$ fermions. Quarks and leptons are divided into three different generations as illustrated in figure 2.1, each generation being a heavier copy of the first one.

The six quarks are grouped into three pairs, one for each generation, and each quark comes in three colour states. Quarks are therefore electrically charged and coloured particles. It turns out that they also have a weak charge. As a consequence, quarks participate to all fundamental interactions. From the six leptons, three are charged and are subject to both weak and electromagnetic interactions, whilst the three neutral leptons called neutrinos interact only through weak interactions and makes them very difficult to be observed. Not represented in figure 2.1, each fermion comes with its corresponding antiparticle having the same mass but opposite charges.

The SM being a gauge theory, the interactions between fundamental particles are described by the exchange of gauge bosons that mediate the forces. Those bosons are also meaningfully referred to as force carriers. The gauge bosons are, for the electromagnetic interaction the photon, for the weak interactions the W^{\pm} and the Z, and the gluons (there are eight of them) for the strong interactions.

Finally, the Higgs boson is the quantum manifestation of the Higgs field with which almost all particles interact more or less intensively and acquire their mass. The Higgs particle is an unstable boson with no spin, electric charge nor colour charge. It was predicted 50 years ago and its discovery announced in July 2012 has given further credence to the SM.

2.2 The Equations Governing the Interactions

As mentioned in the previous section the SM is a gauge invariant quantum field theory that groups the electroweak theory (EWK) together with the theory of the strong interactions called quantum chromodynamics (QCD). Its fundamental objects were described phenomenologically above and their associated quantum fields are listed here as the mathematical objects of the theory:

- ψ denotes a fermion field.
- W^1, W^2, W^3 and B represents the electroweak boson fields.¹
- \mathcal{A}^A for the gluon fields, where A runs from 1 to 8.
- φ stands for the Higgs field.

The gauge group of the SM is $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$. The first term $SU(3)_C$ is the group of colour transformations for which the quarks transform according to the fundamental representation and the antiquarks according to its complex conjugate representation. The

¹The W^{\pm} , Z and γ bosons are not the quanta directly corresponding to these fields but are the quanta of the mass eigenstates fields, result of the Brout-Englert-Higgs mechanism (see section 2.2.3).

second and third terms $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ form the group of symmetry for the electroweak interaction. The subscript Y indicates that the corresponding U(1) symmetry acts on the weak hypercharge, and the subscript L refers to the fact that the SU(2) symmetry operates on the left-handed states only.

Having all the fields of the SM one may write down the Lagrangian that satisfies causality, gauge invariance and Lorentz invariance. From this Lagrangian, it is theoretically possible to derive all kinds of high-energy calculus involving the fundamental particles of the theory. In particular, when a perturbative approach can be used (this is not always possible), the Feynman rules can be derived from the Lagrangian and subsequently used to compute cross section amplitudes from the corresponding Feynman diagrams.

2.2.1 The Lagrangian of the Electroweak Interaction

As mentioned in the previous section, the electroweak sector of the SM has its foundations in the symmetry group $SU(2)_L \otimes U(1)_Y$. As a consequence of this gauge invariance requirement, the electroweak Lagrangian has to contain three massless gauge bosons, W^i (i = 1, 2, 3), for the $SU(2)_L$ factor, and one additional massless boson, B, coming from the $U(1)_Y$ gauge group, such that their transformation laws compensate the terms arising from the fermion field transformation.

The Lagrangian for a massless free fermion field ψ belonging to a general $SU(2)_L$ representation, with $U(1)_Y$ weak hypercharge Y, can be written as:

$$\mathcal{L} = i \,\bar{\psi} \,\gamma^{\mu} \partial_{\mu} \,\psi. \tag{2.1}$$

The expression (2.1) is clearly not invariant under a $SU(2)_L \otimes U(1)_Y$ local transformation of the fields given by

$$\psi \to \psi' = \exp\left(i \, \frac{1}{2} \, g_W \, T^k \Lambda^k(x)\right) \exp\left(i \, \frac{1}{2} \, g'_W \, Y \, \alpha(x)\right) \psi, \tag{2.2}$$

where g_W is the $\mathrm{SU}(2)_L$ gauge coupling and g'_W is the $\mathrm{U}(1)_Y$ gauge coupling, the matrices T^k form a representation of the $\mathrm{SU}(2)_L$ weak isospin algebra and Y is a scalar forming a representation of the trivial $\mathrm{U}(1)_Y$ algebra. In equation (2.2) the first term therefore represents an $\mathrm{SU}(2)_L$ local rotation around the axis $\Lambda^k(x)$ in the space of the weak isospin T, while the second exponential translates a local phase shift $\alpha(x)$ acting in the space of the weak hypercharge Y.

It can be shown that changing the partial derivatives ∂_{μ} to their covariant equivalent D_{μ} given by:

$$(D_{\mu})_{ij} = \delta_{ij}\partial_{\mu} + ig_W \left(T^k W^k_{\mu}\right)_{ij} + ig'_W Y \delta_{ij} B_{\mu}, \qquad (2.3)$$

makes the Lagrangian gauge invariant with the condition that the gauge fields appearing inside the new derivative belong to the adjoint representation of the symmetry group. Furthermore, this step is also introducing interaction between the gauge fields and the fermionic fields as indicated by the presence of terms involving both the fermionic fields and the gauge fields.

The Lagrangian becomes gauge invariant and is no more describing a free fermion because of the interaction terms contained in the covariant derivative given by equation (2.3). The gauge fields introduced by the gauge invariant requirement are also propagating. It is hence necessary to add terms to the Lagrangian to describe their kinematic. The gauge invariant Lagrangian including these terms can be written

$$\mathcal{L} = i \,\bar{\psi} \,\gamma^{\mu} D_{\mu} \,\psi - \frac{1}{4} W^{i}_{\mu\nu} W^{i\,\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \qquad (2.4)$$

where $W^i_{\mu\nu}$ and $B_{\mu\nu}$ are the field strength tensors of the SU(2)_L gauge fields W^i and U(1)_Y gauge field B, respectively, and are given by

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g_W \epsilon^{ijk} W^j_\mu W^k_\nu, \qquad (2.5)$$

$$B_{\mu\nu} = \partial_{\mu}B_{\nu} - \partial_{\nu}B_{\mu}. \tag{2.6}$$

The totally antisymmetric tensor ϵ^{ijk} appearing in the definition of $W^i_{\mu\nu}$ comes from the structure constants of the group $SU(2)_L$ and verifies the following equations:

$$\left[T^{i}, T^{j}\right] = i\epsilon^{ijk}T^{k}, \quad \epsilon^{123} = 1.$$

$$(2.7)$$

The third term in the definition of $W^i_{\mu\nu}$ indicates the self-interaction of the W^i bosons and is a direct consequence of the non-Abelian nature of SU(2). As explained later on, an analogous term is present for the case of the QCD theory based on the non-Abelian group SU(3), and will lead to the gluon self-interaction.²

The only thing left to determine the coupling of the real fermions, *i.e.* the known leptons and quarks, to the gauge bosons, is to choose the $SU(2)_L$ representation T^k to which they belong as well as to assign the weak hypercharge to each of them. In the SM, it is necessary to make the distinction between left- and right-handed fermions because, as a matter of fact, they belong to different representations of the electroweak group.

We can decompose a fermion field by using the γ_R and γ_L projection operators:

$$\psi = \psi_R + \psi_L \equiv \gamma_R \psi + \gamma_L \psi, \qquad (2.8)$$

²For an Abelian group of symmetry, all the T^i commute with each other and the structure constants therefore vanish. The consequence is the absence of self-interacting terms of the gauge fields. This is the case for the U(1) symmetry of electromagnetism in which the photon does not interact with itself. On the opposite, the SU(2) and SU(3) symmetry of the weak and strong interactions are non-Abelian. The non-vanishing third term in the field strength tensor will contain the self interactions of the gauge bosons such as the coupling of the W^{\pm} and the Z or the gluon self interactions.

with $\gamma_R = \frac{1}{2}(1+\gamma^5)$ and $\gamma_L = \frac{1}{2}(1-\gamma^5)$, where $\gamma^5 = \frac{i}{4!}\varepsilon_{\mu\nu\rho\sigma}\gamma^{\mu}\gamma^{\nu}\gamma^{\rho}\gamma^{\sigma}$ with γ^{μ} the four Dirac matrices. The U(1)_Y weak hypercharges assigned to left- and right-handed fermions, Y_L and Y_R must satisfy the relation $Q = T^3 + Y$ as explained later by the Brout-Englert-Higgs mechanism.

Considering the leptons, it is an observation that all right-handed lepton fields are described by $SU(2)_L$ singlets with the consequence that they do not transform under transformations of this group. On the other hand, the left-handed lepton fields are grouped by pairs of electrically neutral and charged leptons into $SU(2)_L$ doublets. Let us note that there are no right-handed neutrino. Such a neutrino would have both electric charge and weak hypercharge equal to zero and would not couple to the photon nor to the Z or W^{\pm} bosons. In other words, it would not have any kind of interaction.

As far as the quarks are concerned, the right-handed quarks, similarly to the right-handed leptons, do not transform under $SU(2)_L$ and are singlets. While the left-handed quarks are put into $SU(2)_L$ doublets.

• leptons:

$$\psi_R = \gamma_R e^-, \ \gamma_R \mu^-, \ \gamma_R \tau^-, \tag{2.9}$$

$$\psi_L = \gamma_L \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}, \ \gamma_L \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}, \ \gamma_L \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix},$$
(2.10)

• quarks:

$$\psi_R = \gamma_R u, \ \gamma_R c, \ \gamma_R t, \tag{2.11}$$

$$\gamma_R d', \ \gamma_R s', \ \gamma_R b', \tag{2.12}$$

$$\psi_L = \gamma_L \begin{pmatrix} u \\ d' \end{pmatrix}, \ \gamma_L \begin{pmatrix} c \\ s' \end{pmatrix}, \ \gamma_L \begin{pmatrix} t \\ b' \end{pmatrix}.$$
(2.13)

Table 2.1 gives a summary of the electroweak quantum numbers for the quarks and leptons. As a result, the $SU(2)_L$ representation to which the quarks and leptons belong is now chosen. Left-handed leptons and quarks live in its fundamental representation for which the generators T^i are given by the Pauli spin matrices (multiplied by a factor $\frac{1}{2}$), while right-handed fermions belong to the trivial representation because they do not transform under $SU(2)_L$.

In summary, the electroweak sector of the SM Lagrangian (before symmetry breaking by the Brout-Englert-Higgs mechanism) is given by

$$\mathcal{L}_{\rm EWK} = \psi_R i \left(\partial + i g'_W Y_R B \right) \psi_R + \bar{\psi}_L i \left(\partial + i g_W T^i W^i + i g'_W Y_L B \right) \psi_L - \frac{1}{4} W^i_{\mu\nu} W^{i\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu},$$
(2.14)

where the Feynman slash notation $A = \gamma_{\mu} A^{\mu}$ for any four-vector A has been used, where $T^i = \frac{1}{2}\sigma^i$ and with an implicit sum on the right- and left-handed fermions. Let us note that introducing a mass term for the gauge bosons would result in a non-invariant Lagrangian under local gauge transformations. The gauge bosons are hence massless at this point of the theory. Trying to add fermionic mass terms would also not leave the Lagrangian invariant under electroweak symmetry transformations because left-handed and right-handed components of the fermionic fields have different associated quantum numbers. In other words, a mass term of the form $m\bar{\psi}\psi$, which is coupled with both left and right components, also causes a rupture of gauge invariance and is ruled out of the Lagrangian. The Brout-Englert-Higgs mechanism, in addition to providing the scalar field necessary for Yukawa couplings for fermion masses, will provide the mass to the familiar bosons W^{\pm} and Z, while one boson, namely the photon, will remain massless as explained in section 2.2.3.

2.2.2 The Lagrangian of Quantum Chromodynamics

Along with the electroweak sector and its gauge group $SU(2)_L \otimes U(1)_Y$, there also exists the quantum chromodynamics sector that comes with its own symmetry group of gauge transformations $SU(3)_C$. The colour charge of quarks had to be introduced after the observation of spin- $\frac{3}{2}$ baryons such as the Δ^{++} , believed to be constituted of three quarks of the same flavour in a symmetrical state of space and spin degrees of freedom. Such a symmetric configuration for a half-integer spin particle would go against the Fermi-Dirac statistics. The solution to this *a priori* anomaly was the introduction of another degree of

Table 2.1: Electroweak quantum numbers of the fundamental fermions in the SM. The primes denote the fact that the corresponding quarks are combination of the mass eigenstates quarks d, s and b obtained by the Cabibbo - Kobayashi - Maskawa matrix (CKM matrix).

		Q	T_3	Y
	u_L, c_L, t_L	$\frac{2}{3}$	$\frac{1}{2}$	$\frac{1}{6}$
Quarka	$\mathbf{d}_L',\mathbf{s}_L',\mathbf{b}_L'$	$-\frac{1}{3}$	$-\frac{1}{2}$	$\frac{1}{6}$
Quarks	$\mathbf{u}_R,\mathbf{c}_R,\mathbf{t}_R$	$\frac{2}{3}$	0	$\frac{2}{3}$
	$\mathbf{d}_R',\mathbf{s}_R',\mathbf{b}_R'$	$-\frac{1}{3}$	0	$-\frac{1}{3}$
	$ u_e, u_\mu, u_ au$	0	$\frac{1}{2}$	$-\frac{1}{2}$
Leptons	$\mathbf{e}_L^-,\mu_L^-,\tau_L^-$	-1	$-\frac{1}{2}$	$-\frac{1}{2}$
	$\mathbf{e}_R^-,\mu_R^-,\tau_R^-$	-1	0	-1

freedom, namely the colour of quarks: a charge with 3 possible values (one could use the red, green and blue colours) that would be carried by each quark. With this new degree, the baryon wave functions are totally antisymmetric and the Fermi-Dirac statistics is recovered.

However, another postulate had to come along with the colour charge hypothesis in order to avoid the prediction of many new coloured states that were not observed. This additional requirement states that only colour singlet state configurations can exist in nature. With the quarks belonging to the fundamental representation of the colour group, and the antiquarks in its complex conjugate, the firmly established colour singlets are the mesons made of a quark (with its colour) and its antiquark (with the anti-colour of the quark), and the baryons composed of three quarks or three antiquarks of different colours (the combination of the three colours being white or colourless).

The gauge group of QCD being $SU(3)_C$, the Lagrangian of the QCD sector must remain the same under transformations of the quark fields given by

$$\psi^c \to \psi^c \,' = \exp\left(i\frac{1}{2}g_s t^A \mathcal{A}^A(x)\right)\psi^c,\tag{2.15}$$

where g_s is the SU(3)_C gauge coupling and the matrices t^A form the fundamental representation of the SU(3)_C colour algebra³, with the index A running over the eight colour degrees of freedom of the gauge group. The exponential represents therefore a local rotation around the axis $\mathcal{A}^A(x)$ in the colour space. For the Lagrangian of a massless and free quark of colour c described by the Dirac spinor ψ^c ,

$$\mathcal{L} = i \,\bar{\psi}^c \,\gamma^\mu \partial_\mu \,\psi^c, \tag{2.16}$$

this gauge invariance requirement implies the usage of the covariant derivative, in exactly the same way as for the electroweak sector,

$$(D_{\mu})_{ab} = \delta_{ab}\partial_{\mu} + ig_s \left(t^A \mathcal{A}^A_{\mu}(x)\right)_{ab}, \qquad (2.17)$$

where the indices a and b run over the three colours of the triplet representation of the group. As a result of this change, eight massless gauge fields, called the gluon fields, for

$$t^A = \frac{1}{2}\lambda^A$$

$$\begin{split} \lambda^{1} &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda^{2} &= \begin{pmatrix} 0 & -i & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda^{3} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda^{5} &= \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda^{6} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\ \lambda^{7} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda^{8} &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \end{split}$$

³The t^A matrices generators of SU(3) are usually given by the height Gell-Mann matrices

the eight different gauge parameters are introduced. Again, like for the electroweak sector, it is necessary to consider the kinematic of the gluons. This is achieved by introducing the trace term

$$\mathcal{L}_{\text{gluon}} = -\frac{1}{4} F^A_{\mu\nu} F^{A\,\mu\nu}, \qquad (2.18)$$

where the field strength tensor $F^A_{\mu\nu}$ is derived from the gluon field \mathcal{A}_{μ} ,

$$F^{A}_{\mu\nu} = \partial_{\mu}\mathcal{A}^{A}_{\nu} - \partial_{\nu}\mathcal{A}^{A}_{\mu} - g_{s}f^{ABC}\mathcal{A}^{B}_{\mu}\mathcal{A}^{C}_{\nu}, \qquad (2.19)$$

where f^{ABC} are the structure constants of the SU(3)_C group. Given the non-Abelian nature of this group, the f^{ABC} are non-zero and lead to the gluon self-interaction analogous to the third term of equation (2.5) of the electroweak interaction.

In summary, the QCD Lagrangian part of the SM is given by

$$\mathcal{L}_{\text{QCD}} = \bar{\psi}^{c}_{q} i \left(\partial \!\!\!/ + i g_{s} t^{A} \mathcal{A}^{A} \right) \psi^{c}_{q} - \frac{1}{4} F^{A}_{\mu\nu} F^{A \mu\nu}.$$
(2.20)

For the same reason as in the electroweak sector, a mass term for the gluon is ruled out because it would not be invariant under gauge transformations. Regarding the quark mass, the same argument is maintained because the SM gauge group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ applies as a whole.

2.2.3 The Mass Terms and the Brout-Englert-Higgs Mechanism

As mentioned in the previous sections, at this level all particles and in particular the gauge bosons of the theory are massless. It is exactly what we want for the QCD sector, since gluons are indeed massless, but that is not the case of the EWK sector for which the known W^{\pm} and Z bosons are observed to be massive particles, while the photon is massless. As explained earlier, adding a mass term for any of the gauge bosons would violate the gauge invariance principle. It is hence necessary to find a mechanism through which the required bosons, and only those that we know are massive, acquire mass, without breaking the gauge invariance. Furthermore, because left-handed and right-handed components of the fermionic fields have different associated quantum numbers, a mass term of the form $m\bar{\psi}\psi$, which is coupled to both left and right components, also causes a rupture of gauge invariance, while we know that the fermions are massive particles.

To face the weak bosons mass issue, a mechanism has been developed simultaneously by Brout, Englert and Higgs in 1964 [7,8]. In their model, a scalar field is introduced, the Higgs field, and induces a spontaneous symmetry breaking of the electroweak sector breaking the gauge group $SU(2)_L \otimes U(1)_Y$ into the subgroup $U(1)_{em}$, a dimension one symmetry corresponding to a massless boson. The three other dimensions of the group are broken and result in three massive gauge bosons. Additionally, as the theory now contains a scalar field, it is therefore possible to add Yukawa terms to the Lagrangian and in that way give mass to the fermions.

Let us first detail the spontaneous symmetry breaking. The Higgs field introduced in the theory is a non-coloured SU(2) doublet with a weak hypercharge of $\frac{1}{2}$,

$$\varphi = \begin{pmatrix} \phi^+\\ \phi^0 \end{pmatrix}, \tag{2.21}$$

where the superscript denotes the doublet components electric charge.

A $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ gauge invariant Lagrangian can accordingly be written as:

$$\mathcal{L}_{\rm H} = (D_{\mu}\varphi)^{\dagger} (D^{\mu}\varphi) - V(\varphi^{\dagger}\varphi), \qquad (2.22)$$

the covariant derivative of the Higgs field being given by

$$D_{\mu}\varphi = \left(\partial_{\mu} + ig_W T^i W^i_{\mu} + i\frac{1}{2}g'_W B_{\mu}\right)\varphi, \qquad (2.23)$$

and the Higgs potential $V(\varphi^{\dagger}\varphi)$ is chosen to be

$$V(\varphi^{\dagger}\varphi) = \lambda \left(\varphi^{\dagger}\varphi\right)^{2} - \mu^{2}\varphi^{\dagger}\varphi.$$
(2.24)

This form is chosen for its invariance under rotations in the four-dimensional space of the φ components and more importantly because its lowest-energy classical configuration is not found to be at $\varphi = 0$ but instead at all fields satisfying

$$|\varphi| = \sqrt{\frac{\mu^2}{2\lambda}}.$$
(2.25)

There is therefore a dimension-3 circle of degenerate minima and a displacement along this circle can be completed without increasing the potential energy and corresponds to massless excitations.

For now, let us ask the following question: what would happen if we took away all the energy available from the field. The answer would be that the field would find itself in the ground state, which ground state is a matter of realisation. The important thing is that once a particular minimum energy configuration has been chosen, corresponding to some specific direction in the internal SU(2) of the minima of φ , the state itself does not reflect the original symmetry of the Lagrangian it belongs to. In that case, the symmetry is said to be spontaneously broken.

As a result, we are free to choose a particular configuration, *i.e.* a particular realisation, of φ for which the potential is at its minimum. We choose it to be

$$\varphi_v = \frac{1}{\sqrt{2}} \begin{pmatrix} 0\\v \end{pmatrix},\tag{2.26}$$

with $v \equiv \sqrt{\frac{\mu^2}{\lambda}}$. An additional characteristic of this symmetry breaking is that $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ is not completely broken. Indeed, there remains one unbroken generator that we call Q. To find it (see [9]), we switch off the gauge fields and look for Hermitian matrices such that

$$Q\varphi_v = 0. \tag{2.27}$$

If this equation is satisfied, then the subgroup generated by the generators Q will be a symmetry of the vacuum state. For the choice (2.26), it is clear that Q must be of the form

$$\begin{pmatrix} a & 0 \\ b & 0 \end{pmatrix}, \tag{2.28}$$

but Hermiticity of the generator requires b = 0 and a = 1, *i.e.*,

$$Q = T^3 + Y = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$
 (2.29)

There is, by consequence, a single unbroken generator that corresponds to the $U(1)_{em}$ subgroup of $SU(2)_L \otimes U(1)_Y$. From this we find out the relation between electroweak charges, and we can expect a massless boson associated to the unbroken generator to be a linear combination of W^3_{μ} and B_{μ} .

To find out the gauge boson mass terms that will appear from the interaction terms with the Higgs field it is necessary to consider small perturbations of the Higgs field around its minimum and examine the terms quadratic in the vector boson fields appearing in the Lagrangian. For this, it is convenient to introduce the complex fields

$$W^{\pm}_{\mu} = \frac{1}{\sqrt{2}} \left(W^{1}_{\mu} \mp i W^{2}_{\mu} \right), \qquad (2.30)$$

with $(W^{-}_{\mu})^{*} = W^{+}_{\mu}$, as well as the two real fields

$$Z_{\mu} = \frac{1}{\sqrt{g_W^2 + g_W'^2}} \left(g_W W_{\mu}^3 - g_W' B_{\mu} \right), \qquad (2.31)$$

$$A_{\mu} = \frac{1}{\sqrt{g_W^2 + g_W'^2}} \left(g_W' W_{\mu}^3 + g_W B_{\mu} \right).$$
 (2.32)

By computing the contribution of the covariant derivative to the quadratic part of the Lagrangian with

$$\varphi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0\\ v+\chi \end{pmatrix},\tag{2.33}$$

we obtain

$$\left[\left(D_{\mu}\varphi \right)^{\dagger} \left(D^{\mu}\varphi \right) \right]^{(2)} = \frac{1}{2} \partial_{\mu}\chi \partial^{\mu}\chi + \frac{g_{W}^{2}v^{2}}{4} W_{\mu}^{+}W^{-\mu} + \frac{1}{2} \frac{\left(g_{W}^{2} + g_{W}^{\prime 2} \right)v^{2}}{4} Z_{\mu}Z^{\mu}.$$
(2.34)

The result of the spontaneous symmetry breaking of the $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ to a subgroup $\mathrm{U}(1)_{em}$ is a set of three massive vector bosons as indicated by the above quadratic terms. The W^{\pm}_{μ} vector bosons have a mass given by $m_W = g_W \frac{v}{2}$, and ± 1 electric charge. They give rise to the observed charged current interaction. The third massive boson is the Z_{μ} boson with mass $m_Z = \sqrt{g_W^2 + g_W'^2} \frac{v}{2}$ and zero electric charge. Together with the photon A_{μ} remaining massless, they are responsible for the neutral current interactions (electromagnetic and weak interactions). By analysing the quadratic part of the Higgs potential we can get a mass term for the Higgs field quanta, namely the Higgs boson with mass $m_{\chi} = \sqrt{2\lambda}v$.

With little additional work in rearranging the terms in the EWK Lagrangian, it is possible to extract several relations between the masses and the gauge coupling constants. For example, we can identify the coefficient of the coupling to the photon to be the positron charge:

$$e = \frac{g_W g'_W}{\sqrt{g_W^2 + g'_W^2}}.$$
(2.35)

The electroweak mixing angle θ_W can also be introduced as the angle appearing in the linear combination defining the Z_{μ} and A_{μ} bosons and is fixed by the relative strengths of the two electroweak coupling constants:

$$\sin^2 \theta_W = \frac{g_W^{\prime 2}}{g_W^2 + g_W^{\prime 2}}.$$
(2.36)

With this definition, we can write the following relations

$$e = g_W \sin \theta_W, \qquad (2.37)$$

$$m_W = m_Z \cos \theta_W. \tag{2.38}$$

As stated at the beginning of this section, since we have introduced the scalar doublet Higgs field into the model, it is now possible to write a mass term for the fermion fields from the gauge-invariant Yukawa fermion-scalar coupling. For the charged leptons⁴ such terms can be written as:

$$\mathcal{L}_{\text{Yukawa}} = -\lambda_l \bar{\psi}_{l\,L} \varphi \psi_{l\,R} + h.c., \qquad (2.39)$$

which after replacing φ with its vacuum expectation value would lead to a fermion mass given by

$$m_l = \frac{1}{\sqrt{2}} \lambda_l v. \tag{2.40}$$

For the quarks, the Yukawa terms become :

$$\mathcal{L}_{\text{Yukawa}} = -\lambda_d \bar{\psi}_{qL} \varphi \psi_{dR} - \lambda_u \epsilon^{ij} \bar{\psi}_{qLi} \varphi_j^{\dagger} \psi_{uR} + h.c., \qquad (2.41)$$

⁴It would also be possible to give the neutrinos a mass by adding right-handed neutrino to the theory. These last ones would be completely neutral under electroweak transformations, but would couple to the Higgs field that would give the neutrino a mass.

and similarly for the other two generations, leading to quark masses given by

$$m_q = \frac{1}{\sqrt{2}} \lambda_q v. \tag{2.42}$$

2.3 The Running of the Coupling Constants

As surprising as it can be, the coupling constants that appear as the strength of the different interactions are not constant. This leads to very important effects that make dynamics of the fundamental interactions more complex. It is in particular tightly related to confinement, which is the fact that quarks and gluons are not observed freely as states propagating over macroscopic distances.

For this section we introduce the following notation for the electromagnetic and strong coupling constants:

$$\alpha_{em} = \frac{e^2}{4\pi}, \quad \alpha_s = \frac{g_s^2}{4\pi}.$$
(2.43)

To give a feeling of why the coupling constants are said to run, let us take the example of QED and isolate an electron in the vacuum. However, an electron is never alone in the vacuum since pair creations can appear in the vacuum according to Heisenberg's uncertainty principle. An electron will always be surrounded by such pairs forming electric dipoles that will align in a specific direction due to the electron's electric field as illustrated in figure 2.2. The net effect is a screening of the electron electric charge. Far away from the electron, we observe some effective coupling, which is numerically equal to that of classical electrodynamics: $\alpha_{em}(r \to \infty) \simeq \frac{1}{137}$, while approaching the electron and penetrating the screening cloud, we are more and more sensitive to the bare charge of the electron. As a consequence, the effective coupling constant will depend on the distance of the observer given by $\lambda = \hbar/Q$, inversely proportional to the scale Q of the probe used during the interaction.

In the case of QCD, the non-Abelian character of the interactions leads to the additional effect that the vacuum can not only produce quark-antiquark pairs resulting in colour dipoles, but can also create loops of gluons strengthening the colour field in the neighbourhood of a quark, as illustrated in figure 2.3. It turns out that the net effect is a mixture of the two opposite effects: the creation of quark-antiquark pairs screening the colour field and the creations of gluon loops reinforcing the colour field. The result therefore depends on the number of quark flavours and colours. As a matter of fact, the gluon loops dominate and the effective strong coupling constant is found to increase with the distance, meaning that using large scale Q probes the quarks will appear more and more free. On the opposite, large distance effects will result in strong coupling constant



Figure 2.2: Vacuum polarisation induced by the presence of an electron (left draw) and effect on the effective electric charge seen by a photon (middle and right panels). The photon is probing the electron and the screening cloud as a whole when it has a large wavelength but can resolve the inside of the cloud and see the details around the electron at small enough wavelength.



Figure 2.3: Pair creations surrounding an electron (left). Quark-antiquark creations and gluon loops surrounding a quark (right).

resulting in the confinement of the quarks inside hadrons. The scale at which the strong running coupling constant becomes close to unity is defined as $\Lambda_{\rm QCD}$. This parameter can be understood as the scale at which non-perturbative effects (due to the large value of the strong coupling constant at this scale) start to be dominant. Its value is found to be around 200 MeV.

2.3.1 The Beta Function

The running of the coupling constants can be obtained formally. It is determined by the renormalisation group equation that expresses the variation in the effective strength of the interactions with the energy scale at which the process is taking place:

$$Q^2 \frac{\partial \alpha}{\partial Q^2} = \beta(\alpha). \tag{2.44}$$

The β function appearing on the right hand side of the above equation encodes the energy scale dependence of the coupling constant α . It is obtained from the loop corrections to the bare vertices of the theory. For QED, the first non-vanishing term is found to be

$$\beta_{em}(\alpha_{em}) = \frac{1}{3\pi} \alpha_{em}^2 + \dots,$$
 (2.45)

which is positive and therefore in accordance with the fact that α_{em} increases with the scale Q^2 . The first-order dependence of the coupling constant is given by:

$$\alpha_{em}(Q^2) = \frac{\alpha_{em}(\mu_R^2)}{1 - \frac{\alpha_{em}(\mu_R^2)}{3\pi} \ln \frac{Q^2}{\mu_R^2}}.$$
(2.46)

This relation describes the coupling constant evolution between an arbitrary scale μ_R , the renomarlisation scale, and the physical scale Q.

On the other hand, the expansion of the QCD β function in powers of α_s is shown to be, at two-loop order [10]:

$$\beta_{QCD}(\alpha_s) = -\frac{\alpha_s^2}{4\pi} \left(\frac{11}{3} C_A - \frac{2}{3} n_f \right) -\frac{\alpha_s^3}{(4\pi)^2} \left(\frac{34}{3} C_A^2 - 2C_F n_f - \frac{10}{3} C_A n_f \right),$$
(2.47)

where $C_F = (n_c^2 - 1)/2n_c$ and $C_A = n_c$ are the Casimir operators of the fundamental and adjoint representations of the colour group $SU(n_c)$ and n_f is the number of active quark flavours. In this equation, the first term multiple of n_f comes from the quarks contribution while the term in $-C_A$, increasing the colour charge, comes directly from gluon contributions. The leading term is therefore negative in the SM for which $n_c = 3$ and n_f can reach a maximum of 6. The strong coupling constant $\alpha_s(Q^2)$ is found to decrease when increasing the energy scale Q^2 , an effect called asymptotic freedom. As for QED, the explicit Q^2 dependence of the strong coupling constant can be derived from the above equation. Considering the first non-vanishing term, this leads to:

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu_R^2)}{1 + \left(\frac{11}{3}C_A - \frac{2}{3}n_f\right)\frac{\alpha_s(\mu_R^2)}{4\pi}\ln\frac{Q^2}{4\pi}\mu_R^2}$$
(2.48)

Including the two-loop term, the solution becomes:

$$\frac{1}{\alpha_s(Q^2)} + b' \ln\left(\frac{\alpha_s(Q^2)}{1 + b'\alpha_s(Q^2)}\right) - \frac{1}{\alpha_s(\mu_R^2)} - b' \ln\left(\frac{\alpha_s(\mu_R^2)}{1 + b'\alpha_s(\mu_R^2)}\right) \\ = \left(\frac{11}{3}C_A - \frac{2}{3}n_f\right) \frac{1}{4\pi} \ln\frac{Q^2}{\mu_R^2},$$
(2.49)



Figure 2.4: The strong coupling $\alpha_s(Q)$ (solid black line) and its total uncertainty (yellow band) as determined using a two-loop solution to the renormalisation group equation as a function of the scale Q. The extractions of $\alpha_s(Q)$ from several CMS measurements are shown together with results from the H1 and ZEUS experiments at the HERA collider, and from the D0 experiment at the Tevatron collider. Figure extracted from [11].

where

$$b' = \frac{1}{4\pi} \frac{\left(\frac{34}{3}C_A^2 - 2C_F n_f - \frac{10}{3}C_A n_f\right)}{\left(\frac{11}{3}C_A - \frac{2}{3}n_f\right)}.$$
(2.50)

Equation (2.49) is now an implicit equation for $\alpha_s(Q^2)$ and has to be solved numerically. It was used to fit the data from the CMS experiment as shown in figure 2.4 on page 20.

2.4 Z + Jets in Proton-Proton Collisions

The process studied in the data analysis part of this thesis being the production of a Z boson with possible additional jets at the LHC, related elements of the theory are discussed hereunder. The factorisation of the Z prediction from pp interaction to partonparton interaction is first explained and then used to compute the leading order cross section. Perturbative QCD corrections are then introduced together with the parton distribution functions (PDF).

Quarks are coloured particles and as a matter of fact, cannot be observed directly.



Figure 2.5: The parton model description of a hard scattering process as considered by the QCD factorisation theorem.

As explained in section 2.2.2, only mesons and baryons, collectively called hadrons, can actually be observed. All we can do is create beams of hadrons that will provide us with "beams" of partons that constitute them. However, it is not possible to control the momentum of the partons inside the hadrons. We are by consequence brought to study hadron-hadron collisions and their cross sections in order to study parton interactions.

2.4.1 The QCD Factorisation Theorem

The Lagrangian discussed in the previous sections allows us to calculate so-called partonic cross sections, *i.e.* cross sections for direct parton-parton interactions. As depicted in figure 2.5, the QCD factorisation theorem makes the link between this parton-level cross section and the hadronic cross section, *i.e.* interactions between the hadrons to which the partons belong.

More concretely, the theorem claims that the cross section $\sigma(P_1, P_2)$ for a hard scattering process between two hadrons with four-momentum P_1 and P_2 , can be computed by weighting the corresponding partonic cross sections $\hat{\sigma}_{ij}$ with the PDF $f_i(x, \mu_F^2)$ defined as the probability density for finding a parton *i* with a certain longitudinal momentum fraction *x* at resolution scale μ_F^2 inside the incoming hadron⁵. The cross section can thus

⁵The PDF are extracted mainly from deep inelastic scattering measurements. The domain of validity for the QCD factorisation theorem is given by the analogous of Bjorken regime in deep inelastic scattering that becomes $s \to \infty$, $M_{l^+l^-}^2/s$ fixed.

be written as

$$\sigma(P_1, P_2) = \sum_{i,j} \int dx_1 dx_2 \, \left\{ f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) \times \hat{\sigma}_{ij}(p_1, p_2, \alpha_s(\mu_R^2), Q^2/\mu_R^2, \mu_F^2) \right\},\tag{2.51}$$

where the μ_F^2 is the factorisation scale separating long- and short-distance physics, the sum runs over all partons, $p_1 = x_1 P_1$ and $p_2 = x_2 P_2$. Q^2 denotes the characteristic scale of the hard scattering and μ_R^2 is the renormalisation scale for the QCD running constant.

2.4.2 Application to the Drell-Yan Process

The DY process that occurs in high energy hadron-hadron collisions was first suggested by Sidney Drell and Tung-Mow Yan in 1970 [12]. It describes the production of charged lepton-antilepton pairs resulting from the decay of virtual photon created by the annihilation of a quark from one hadron and the corresponding antiquark from another hadron. With current accelerators energies, the creation of a Z boson by the quark-antiquark annihilation is easily reached and dominates at quark-antiquark centre-of-mass energies around the Z mass of 91.187 GeV. For the data analysis performed in this thesis, the DY process is understood to mean the production of charged lepton-antilepton pairs from both virtual photon or Z boson decay.

The hadronic cross section for the DY process $\sigma_{AB\to l^+l^-}$, where A and B are hadrons (for instance A and B are protons at the LHC, but at Tevatron A is a proton, while B is an antiproton), is computed by weighting the corresponding partonic cross sections which at lowest-order can be written as $\hat{\sigma}_{q\bar{q}\to l^+l^-}$, with the PDF $f_{q/A}(x)$:

$$\sigma_{AB\to l^+l^-} = \sum_{q} \int dx_1 dx_2 \, \left\{ f_{q/A}(x_1, \mu_F^2) f_{\bar{q}/B}(x_2, \mu_F^2) \times \hat{\sigma}_{q\bar{q}\to l^+l^-} + (q \leftrightarrow \bar{q}) \right\}.$$
(2.52)

In practice and thanks to asymptotic freedom, a perturbative approach is taken to compute the hard process cross section and the factor $\hat{\sigma}$ is expanded in power of the strong coupling constant α_s . This factor in equation (2.52) therefore becomes:

$$\hat{\sigma}_{q\bar{q}\to l^+l^-} = \left[\hat{\sigma}_0 + \alpha_s(\mu_R^2)\hat{\sigma}_1 + \dots\right]_{q\bar{q}\to l^+l^-}.$$
(2.53)

The μ_F^2 and μ_R^2 dependence of the different terms in (2.52) are in principle exactly compensating each other after integration over the full phase space if all orders are taken into account. The cross section are thus formally invariant under choices of these parameters. However, it is of course not possible to compute the infinity of terms appearing in the α_s power series and specific selection requirements may be introduced to, *e.g.* study jets. Therefore a choice has to be made for μ_F^2 and μ_R^2 that will lead to specific cross section


Figure 2.6: Leading order Feynman diagram for the DY process for virtual photon exchange. The different factors associated to each element of the Feynman diagram according to the Feynman rules are also depicted.

predictions for which the uncertainty related to the choice of scale has to be estimated by comparing the results obtained with different scale choices. In the case of the DY production, the usual choice is given by $\mu_F^2 = \mu_R^2 = M_{l+l-}^2$.

With this approach, it is possible to make cross section predictions for the process of interest in this thesis: $PP \rightarrow Z + X$ by applying the following recipe:

- 1. find out the leading-order partonic contribution: $\hat{\sigma}_{q\bar{q}\to l^+l^-}$,
- 2. compute the corresponding $\hat{\sigma}_{0 q\bar{q} \to l^+ l^-}$,
- 3. convolute with the proton PDF,
- 4. numerically integrate over the momentum fraction x_1 , x_2 as well as phase space variables.

Leading Order Cross Section

In order to apply the above description to the DY process, we consider first the annihilation of a quark-antiquark pair (for some fixed flavour of quark) into a virtual photon or Z boson subsequently decaying into a charged lepton pair limiting ourself to the leading order. Later on, the first order QCD corrections will be included for the simplest case of photon production only. The first elements needed for this calculation are the matrix elements, \mathcal{M}_{γ} and \mathcal{M}_{Z} for the process $q\bar{q} \to \gamma^{*}/Z \to l^{+}l^{-}$. We start by the former and using the EWK Feynman rules depicted in figure 2.6 along with the corresponding



Figure 2.7: Leading order Feynman diagram for the DY process for Z boson exchange. The different factors associated to each element of the Feynman diagram according to the Feynman rules are also depicted.

Feynman diagram, the matrix element is obtained:

$$i\mathcal{M}_{\gamma} = \bar{v}_{c_{\bar{q}}}^{s_{\bar{q}}}(p_{\bar{q}})(-i)Q_{q}e\delta_{c_{\bar{q}}c_{q}}\gamma^{\mu}u_{c_{q}}^{s_{q}}(p_{q}) \times \left(\frac{-ig_{\mu\nu}}{\hat{s}}\right) \times \bar{u}^{s_{l}}(p_{l})(-i)Q_{l}e\gamma^{\nu}v^{s_{\bar{l}}}(p_{\bar{l}})$$
$$= i\frac{Q_{q}Q_{l}e^{2}}{\hat{s}} \times \bar{v}_{c_{\bar{q}}}^{s_{\bar{q}}}(p_{\bar{q}})\delta_{c_{\bar{q}}c_{q}}\gamma^{\mu}u_{c_{q}}^{s_{q}}(p_{q}) \times \bar{u}^{s_{l}}(p_{l})\gamma_{\mu}v^{s_{\bar{l}}}(p_{\bar{l}}), \qquad (2.54)$$

where the indices c_q and $c_{\bar{q}}$ are the quark and antiquark colours that can take n_c different values, and the indices s_q , $s_{\bar{q}}$, s_l and $s_{\bar{l}}$ indicate the quark, antiquark, lepton and antilepton spins, respectively. The symbol \hat{s} represents the quark-antiquark centre-of-mass energy and is linked to the pp centre-of-mass energy s by: $\hat{s} = x_1 x_2 s$, where x_i is the proton momentum fraction brought by the parton i. Q_q and Q_l are the quark and lepton electric charges in units of the positron electric charge e. It is now a matter of calculation, that can be found in section A.1 of appendix A, to obtain the differential and integrated cross sections. The final differential cross section is found to be:

$$\frac{d\sigma_{\gamma}}{d\Omega} = \frac{\alpha^2 Q_q^2 Q_l^2}{8\hat{s}} \frac{1}{n_c} \left\{ (1 + \cos\theta)^2 + (1 - \cos\theta)^2 \right\},\tag{2.55}$$

where θ is defined as the angle of emission of the lepton in the photon rest frame. The cross section is then obtained by integrating the differential cross section over the full solid angle:

$$\sigma_{\gamma} = \frac{4\pi\alpha^2}{3\hat{s}} \frac{1}{n_c} Q_q^2 Q_l^2.$$
(2.56)

Similarly for the production of a Z boson, using the appropriate EWK Feynman rules summarised in figure 2.7 together with the corresponding Feynman diagram, we get the matrix element:

$$i\mathcal{M}_{Z} = \bar{v}_{c_{\bar{q}}}^{s_{\bar{q}}}(p_{\bar{q}}) \frac{(-i)g_{Z}}{2} \delta_{c_{\bar{q}}c_{q}} \gamma^{\mu} (V_{q} - A_{q}\gamma^{5}) u_{c_{q}}^{s_{q}}(p_{q}) \times \left(\frac{-ig_{\mu\nu} + i\frac{q_{\mu}q_{\nu}}{M_{Z}^{2}}}{\hat{s} - M_{Z}^{2} + iM_{Z}\Gamma_{Z}}\right) \\ \times \bar{u}^{s_{l}}(p_{l}) \frac{(-i)g_{Z}}{2} \gamma^{\nu} (V_{l} - A_{l}\gamma^{5}) v^{s_{\bar{l}}}(p_{\bar{l}}) \\ = \left(\frac{g_{Z}}{2}\right)^{2} \frac{i}{\hat{s} - M_{Z}^{2} + iM_{Z}\Gamma_{Z}} \frac{\bar{v}_{c_{\bar{q}}}^{s_{\bar{q}}}(p_{\bar{q}}) \delta_{c_{\bar{q}}c_{q}} \gamma^{\mu} (V_{q} - A_{q}\gamma^{5}) u_{c_{q}}^{s_{q}}(p_{q}) \\ \times \bar{u}^{s_{l}}(p_{l}) \gamma_{\mu} (V_{l} - A_{l}\gamma^{5}) v^{s_{\bar{l}}}(p_{\bar{l}}), \qquad (2.57)$$

where $g_Z = g_W / \cos \theta_W$, $V_f = T_f^3 - 2Q_f \sin^2 \theta_W$, $A_f = T_f^3$ and where we have neglected the fermion masses arising from the $q_\mu q_\nu$ terms.

After similar calculation, slightly more complicated due to the presence of the γ^5 matrix, available in section A.2 of appendix A, the differential cross section for the production of a Z boson only is found to be:

$$\frac{d\sigma_Z}{d\Omega} = \frac{1}{64\pi^2 \hat{s}} \overline{|\mathcal{M}_Z|^2}
= \frac{1}{n_c} \frac{1}{8\hat{s}} \frac{1}{\sin^4 2\theta_W} \alpha^2 \left(\frac{\hat{s}^2}{(\hat{s} - M_Z^2)^2 + M_Z^2 \Gamma_Z^2} \right) \left\{ \\
\left[(V_q^2 + A_q^2) (V_l^2 + A_l^2) + 4A_q V_q A_l V_l \right] (1 + \cos \theta)^2 \\
+ \left[(V_q^2 + A_q^2) (V_l^2 + A_l^2) - 4A_q V_q A_l V_l \right] (1 - \cos \theta)^2 \right\}.$$
(2.58)

Once integrated over the solid angle, this gives:

$$\sigma_Z = \frac{4\pi\alpha^2}{3\hat{s}} \frac{1}{n_c} \frac{1}{\sin^4 2\theta_W} (V_q^2 + A_q^2) (V_l^2 + A_l^2) \frac{\hat{s}^2}{(\hat{s} - M_Z^2)^2 + M_Z^2 \Gamma_Z^2}.$$
 (2.59)

Finally, the interference between the γ^* and the Z production modes can be computed using the previous results. The detailed calculation can be found in section A.3.

The differential cross section of the interference is found to be:

$$\frac{d\sigma_{\text{int.}}}{d\Omega} = \frac{\alpha^2}{n_c} \frac{Q_q Q_l}{4\sin^2 2\theta_W} \frac{\hat{s} - M_Z^2}{(\hat{s} - M_Z^2)^2 + M_Z^2 \Gamma_Z^2} \Big\{ (V_q V_l + A_q A_l)(1 + \cos\theta)^2 + (V_q V_l - A_q A_l)(1 - \cos\theta)^2 \Big\}.$$

Once integrated over the solid angle we get:

$$\sigma_{\rm int.} = \frac{4\pi\alpha^2}{3\hat{s}} \frac{1}{n_c} \frac{2Q_q Q_l}{\sin^2 2\theta_W} V_q V_l \frac{\hat{s}(\hat{s} - M_Z^2)}{(\hat{s} - M_Z^2)^2 + M_Z^2 \Gamma_Z^2}.$$
(2.60)

The total cross section for the DY process at leading order, for some flavour of quark and charged leptons, is^6

$$\hat{\sigma}_{0}(q(p_{1})\bar{q}(p_{2}) \to l^{+}l^{-}) = \sigma_{\gamma^{*}} + \sigma_{\text{int.}} + \sigma_{Z}$$

$$= \frac{4\pi\alpha^{2}}{3\hat{s}} \frac{1}{n_{c}} \Big(Q_{q}^{2}Q_{l}^{2} + 2Q_{q}Q_{l}V_{l}V_{q}\chi_{1}(\hat{s}) + (A_{l}^{2} + V_{l}^{2})(A_{q}^{2} + V_{q}^{2})\chi_{2}(\hat{s}) \Big), \qquad (2.61)$$

where

$$\chi_1 = \kappa \frac{\hat{s}(\hat{s} - M_Z^2)}{(\hat{s} - M_Z^2)^2 + M_Z^2 \Gamma_Z^2}, \qquad \chi_2 = \kappa^2 \frac{\hat{s}^2}{(\hat{s} - M_Z^2)^2 + M_Z^2 \Gamma_Z^2}, \qquad \kappa = \frac{1}{\sin^2 2\theta_W}.$$

This result represents the first term of (2.53). The next section discusses the QCD corrections to this cross section. Since the partons come from hadrons the incoming quark and antiquark have a spectrum of momenta rather than fixed values. It is therefore more appropriate to use the differential cross section as a function of the dilepton mass M^2 . We can obtain the differential expression by substituting M^2 to \hat{s} in (2.61) and multiplying by the appropriate delta distribution:

$$\frac{d\hat{\sigma}_0}{dM^2} = \sigma_B \frac{1}{n_c} \delta(\hat{s} - M^2) \Big(Q_q^2 Q_l^2 + 2Q_q Q_l V_l V_q \chi_1(M^2) \\ + (A_l^2 + V_l^2) (A_q^2 + V_q^2) \chi_2(M^2) \Big)$$
(2.62)

with $\sigma_B = \frac{4\pi\alpha^2}{3M^2}$. In the centre-of-mass of the hadron collision, the relation $\hat{s} = (p_q + p_{\bar{q}})^2 = x_1 x_2 s$ holds and we can finally insert the partonic cross section result into the expression based on the QCD factorisation theorem to obtain the pp cross section for the DY process at leading order:

$$\frac{d\sigma}{dM^2} = \sum_q \int_0^1 dx_1 dx_2 \left\{ f_{q/P}(x_1) f_{\bar{q}/P}(x_2) + (q \leftrightarrow \bar{q}) \right\} \times \frac{d\hat{\sigma}}{dM^2} (q\bar{q} \to l^+ l^-).$$

Perturbative QCD Corrections

The only Feynman diagram present at LO for the DY process is the radiationless quarkantiquark fusion depicted in figure 2.8 (left) which does not contain any QCD vertex interaction. When considering the $O(\alpha_s)$ terms needed to compute $\hat{\sigma}_1$ in equation (2.53) two kinds of diagrams arise concerning the quark-antiquark fusion. One where the incoming quark or antiquark radiates a real gluon and the second type for which the emitted

⁶For the readers who would like to check this result and compare it with the one present in [1] p. 311, they will unfortunately have to do the math again or find another reference because of the mistake present in the expression of κ in this book (the corrected expression can be found in the errata and agrees with the present result).



Figure 2.8: From left to right: Leading-order Feynman diagram for the DY process, $O(\alpha_s)$ Feynman diagram from $q\bar{q}$ with real gluon emission, $O(\alpha_s)$ (when considering the its interference with the LO diagram) diagram with virtual gluon corrections and $O(\alpha_s)$ diagram from qg scattering.

gluon is virtual, so-called virtual gluon corrections. An example of the former is shown in figure 2.8 (second from the left) and the latter is depicted in the same figure (third from the left). However, an additional contribution to the DY cross section has to be taken into account at LO in α_s . Indeed, while the above cross sections were concerning the $q\bar{q}$ annihilation where the quark and antiquark were directly coming from the proton, the hadrons also contain gluons that can be fetched in place of the quark or antiquark. If the gluon subsequently creates a quark-antiquark pair it can also contribute to the DY cross section. Such a diagram is shown in figure 2.8 (rightmost diagram). There are by consequence three types of contribution at $O(\alpha_s)$: virtual gluon correction to the leadingorder contribution (arising through the interference with the LO diagram calculated in the previous section), real gluon corrections and quark(antiquark)-gluon scattering. From now and until the end of this section, only the photon exchange channel is considered for simplicity. The corrections are similar for the Z boson production and their interference.

Concerning the virtual gluon corrections, the loop integrals present when writing down the Feynman rules of the corresponding diagram are divergent for large loop momenta and gives rise to ultra-violet divergencies. These can be regulated and removed by adding a counter term to the QCD Lagrangian where the singularities are absorbed by a redefinition of the quark charge, quark field and gluon field. Infrared divergences also appear from the loop integrals and, following the dimensional regularisation scheme where the space-time dimension is set to $4 - 2\varepsilon$ instead of 4, they add up to:

$$\hat{\sigma}^{q\bar{q},V} = \hat{\sigma}_0 \frac{\alpha_s}{2\pi} C_F D(\epsilon) \delta(1-\tau) \left[-\frac{2}{\varepsilon^2} - \frac{3}{\varepsilon} + \frac{2\pi^2}{3} - 8 + O(\varepsilon) \right], \qquad (2.63)$$

where $\tau = M^2/\hat{s}$. In this expression the $\frac{1}{\varepsilon^2}$ and $\frac{1}{\varepsilon}$ terms come from soft and collinear divergences respectively and

$$D(\varepsilon) = 1 + \varepsilon \left(\ln(4\pi) - \gamma_E - \ln \frac{M^2}{\mu^2} \right) + O(\varepsilon^2),$$

where $\gamma_E = 0.5772$ is the Euler-Mascheroni constant and μ is a scale introduced to preserve the dimensions of physical quantities when working in the dimensional regularisation scheme. Considering the diagrams for the real gluon emission one can obtain⁷:

$$\hat{\sigma}^{q\bar{q},R} = \hat{\sigma}_0 \frac{\alpha_s}{2\pi} C_F D(\epsilon) \left[\frac{2}{\varepsilon^2} \delta(1-\tau) - \frac{2}{\varepsilon} \frac{(1+\tau)^2}{(1-\tau)_+} \right]$$
(2.64)

+
$$4(1+\tau^2)\left(\frac{\ln(1-\tau)}{1-\tau}\right)_+ - 2\left(\frac{1+\tau^2}{1-\tau}\right)\ln\tau\right],$$
 (2.65)

where the $\frac{1}{\varepsilon^2}$ and $\frac{1}{\varepsilon}$ terms also represent soft and collinear divergences respectively.

From the above two expressions it can be seen that the soft divergences exactly cancel when adding the two contributions. The collinear divergencies do not vanish however and the contribution to the cross section is found to be (ignoring terms $O(\varepsilon)$ or higher):

$$\hat{\sigma}^{q\bar{q},V} + \hat{\sigma}^{q\bar{q},R} = \hat{\sigma}_0 \frac{\alpha_s}{2\pi} \left[2\left(-\frac{1}{\varepsilon} - \ln(4\pi) + \gamma_E\right) P_{qq}^{(0)}(\tau) + D_q(\tau,\mu^2) \right],$$

where

$$P_{qq}^{(0)}(\tau) = C_F \left(\frac{1+\tau^2}{(1-\tau)_+} + \frac{3}{2}\delta(1-\tau) \right)$$

and

$$D_q(\tau,\mu^2) = C_F \left[4(1+\tau^2) \left(\frac{\ln(1-\tau) + \frac{1}{2} \ln \frac{M^2}{\mu^2}}{1-\tau} \right)_+ - 2\frac{1+\tau^2}{1-\tau} \ln \tau + \delta(1-\tau) \left(\frac{2\pi^2}{3} - 8 \right) \right].$$

In order to extract a finite result from the above calculation, the remaining divergences are included in a redefinition of the PDF that become now scale-dependent as already announced:

$$f_{q/P}(x,\mu^2) = f_q(x) + \frac{\alpha_s(\mu^2)}{2\pi} \left(-\frac{1}{\varepsilon} - \ln(4\pi) + \gamma_E \right) \int_x^1 \frac{d\xi}{\xi} \left[P_{qq}^{(0)}\left(\frac{x}{\xi}\right) f_{q/P}(\xi) + P_{qg}^{(0)}\left(\frac{x}{\xi}\right) f_g(\xi) \right].$$

The contribution from the quark-gluon scattering can also be computed:

$$\hat{\sigma}^{qg} = \hat{\sigma}_0 \frac{\alpha_s}{2\pi} \left[\left(-\frac{1}{\varepsilon} - \ln(4\pi) + \gamma_E \right) P_{qg}^{(0)}(\tau) + D_g(\tau, \mu^2) \right],$$

⁷The "plus" distribution is defined so that its integral with any sufficiently smooth distribution f is:

$$\int_0^1 dx \frac{f(x)}{(1-x)_+} = \int_0^1 dx \frac{f(x) - f(1)}{1-x},$$

and

$$\frac{1}{(1-x)_{+}} = \frac{1}{1-x} \text{ for } 0 \le x < 1.$$

where

$$P_{qg}^{(0)}(\tau) = \frac{1}{2} \left(\tau^2 + (1-\tau)^2 \right)$$

and

$$D_g(\tau,\mu^2) = \frac{1}{2} \left[\left(\tau^2 + (1-\tau)^2\right) \left(\ln\frac{(1-\tau)^2}{\tau} + \ln\frac{M^2}{\mu^2}\right) + \frac{1}{2} + 3\tau - \frac{7}{2}\tau^2 \right].$$

and leads to similar collinear divergences which are absorbed into the PDF. These NLO results contain one major ingredient, the so-called splitting functions P_{ab} where a and b are partons, and will be discussed in the next section. These translate the probability that a parton has to emit another parton and are directly linked to the gluon radiation or quark emission from the Feynman diagrams used to compute the $O(\alpha_s)$ corrections in this section.

For the sake of completeness, the result to $O(\alpha_s)$ of the DY cross section (limited to the photon exchange) in the modified minimal subtraction scheme ($\overline{\text{MS}}$) used to absorb the collinear divergence with the choice of scale $\mu^2 = M^2$ is given below:

$$\begin{aligned} \frac{d\sigma}{dM^2} = & \frac{4\pi\alpha^2}{3s} \frac{Q_l^2}{n_c} \int_0^1 dx_1 dx_2 dz \delta(x_1 x_2 z s - M^2) \\ & \times \left[\sum_q Q_q^2 \Big\{ f_{q/P}(x_1, M^2) f_{\bar{q}/P}(x_2, M^2) + (q \leftrightarrow \bar{q}) \Big\} \\ & \times \left(\delta(1-z) + \frac{\alpha_s(M)}{2\pi} D_q(z, M^2) \right) \\ & + \sum_q Q_q^2 \Big\{ f_{g/P}(x_1, M^2) \Big(f_{q/P}(x_2, M^2) + f_{\bar{q}/P}(x_2, M^2) \Big) \\ & + (q, \bar{q} \leftrightarrow g) \Big\} \times \frac{\alpha_s(M^2)}{2\pi} D_g(z, M^2) \Big]. \end{aligned}$$

In this equation, the first term in the big square brackets represent the quak-antiquark annihilation with $O(\alpha_s)$ contributions taken into account, and the second term is the new contribution from quark(antiquark)-gluon scattering and is therefore proportional to the gluon PDF $f_{q/P}$ as well as the quark or antiquark PDF $f_{q/P}$ and $f_{\bar{q}/P}$ respectively.

The fact that a new channel is activated when considering $O(\alpha_s)$ corrections together with the fact that gluon densities are much larger (at typical x values) than quarks or antiquarks densities inside the hadrons make the correction to the leading-order cross section significant and one usually provides a K-factor as the ratio of LO+NLO to LO predictions. For the case of the DY process $pp \rightarrow Z/\gamma^* + X$ the K-factor has been



Figure 2.9: The CMS rapidity distribution of an on-shell Z boson at the LHC nominal centre-of-mass energy of 14 TeV. The LO, NLO, and NNLO results have been included. The bands indicate the variation of the renormalisation and factorisation scales in the range $M_Z/2 \leq \mu \leq 2M_Z$. Figure extracted from [13].

computed and reaches values up to 130% affecting both the normalisation and the shape of the distributions as can be seen on picture 2.9 from [13]. Cross sections at NNLO have already been computed [14] including among others, diagrams with two incoming gluons each splitting into a quark-antiquark. The effect is shown on the same figure and shows a decrease of one to two percent with respect to the NLO predictions without however significantly changing the shape of the distribution. Additionally the hard radiations that can take place at beyond leading order can give, as a result, a significant transverse momentum to the virtual photon of Z boson which would not be possible at LO.

2.4.3 Parton Distribution Functions

The PDF appearing in the QCD factorisation formula cannot be computed from first principle. They have to be extracted from measurements at some energy scale μ_0^2 and then evolved using pQCD evolution equations (see next paragraph) to some other desired

energy scale Q^2 . They are mainly extracted from deep inelastic scattering measurements done at the lepton-hadron collider HERA [15,16] and from fixed target experiments which access larger x values. Hadrons colliders also provide some constraints by studying the DY process [17]. The main characteristic of the PDF is that they depend on the hard scale at which the hadron is probed. The number of visible gluons and quarks inside the hadrons depends on the way the hadrons are probed (see section 2.3). At large Q^2 , the detailed structure is seen as resulting of many gluons and quarks carrying a small momentum fraction of the hadron, while a lower scale does not allow such a detailed resolution and results in a smaller number of visible partons, carrying larger longitudinal momentum fractions.

This energy scale dependence is known from QCD calculations. Indeed, given the PDF at some μ_0^2 value, they can be evolved up to another scale Q^2 using the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) [18–21] equations:

$$Q^{2} \frac{\partial}{\partial Q^{2}} \begin{pmatrix} q_{i}(x,Q^{2}) \\ g(x,Q^{2}) \end{pmatrix} = \frac{\alpha_{s}(Q^{2})}{2\pi} \sum_{q_{j},\bar{q}_{j}} \int_{x}^{1} \frac{d\xi}{\xi} \times \begin{pmatrix} P_{q_{i}q_{j}}(\frac{x}{\xi},\alpha_{s}(Q^{2})) & P_{q_{i}g}(\frac{x}{\xi},\alpha_{s}(Q^{2})) \\ P_{gq_{j}}(\frac{x}{\xi},\alpha_{s}(Q^{2})) & P_{gg}(\frac{x}{\xi},\alpha_{s}(Q^{2})) \end{pmatrix} \begin{pmatrix} q_{j}(\xi,Q^{2}) \\ g(\xi,Q^{2}) \end{pmatrix},$$

$$(2.66)$$

where $q_i(x, Q^2)$ and $g(x, Q^2)$ are the PDF associated to the quark or antiquark of flavour *i* and to the gluon respectively, while the functions $P_{ab}(z, \alpha_s(Q^2))$ are the splitting functions which have a perturbative expansion in terms of the running constant α_s :

$$P_{ab}(z,\alpha_s) = P_{ab}^{(0)}(z) + \frac{\alpha_s}{2\pi} P_{ab}^{(1)}(z) + \dots$$
(2.67)

The leading order splitting functions $P_{ab}^{(0)}(z)$ can be interpreted as the probabilities of getting a parton of type *a* from a parton of type *b*, carrying a fraction *z* of the momentum of the parent parton and a transverse momentum squared much less than μ^2 [21].

The leading order, next-to-leading order and even next-to-next-to-leading order terms of the splitting functions have been calculated and can be found in [22–24]. The leading order terms are written here and their corresponding Feynman diagrams are represented in figure 2.10:



Figure 2.10: The processes related to the lowest order QCD splitting functions. From left to right: $P_{qq}^{(0)}(z)$, $P_{qg}^{(0)}(z)$, $P_{gg}^{(0)}(z)$. Each splitting function $P_{ba}(z)$ gives the probability that a parton of type a converts into a parton of type b, carrying a fraction z of the momentum of parton a.

$$P_{qq}^{(0)}(z) = \frac{4}{3} \left(\frac{1+z^2}{(1-z)_+} + \frac{3}{2} \delta(1-z) \right)$$
(2.68)

$$P_{qg}^{(0)}(z) = \frac{1}{2} \left(z^2 + (1-z)^2 \right)$$
(2.69)

$$P_{gq}^{(0)}(z) = \frac{4}{3} \left(\frac{1 + (1 - z)^2}{z} \right)$$
(2.70)

$$P_{gg}^{(0)}(z) = 6\left(\frac{z}{(1-z)_{+}} + \frac{1-z}{z} + z(1-z) + \frac{11C_{A} - 2n_{f}}{6}\delta(1-z)\right). \quad (2.71)$$

In conclusion, with the set of DGLAP equations it is possible to evolve the PDF from a scale μ_0^2 at which they are derived from experimental measurements to an arbitrary scale μ_F^2 called the factorisation scale which fixes the maximal radiation virtuality treated by the evolution of the PDF. Higher virtualities should still be included in the matrix element calculation. In order to illustrate this, an examples of such a PDF evolution is presented in figure 2.11.



Figure 2.11: PDF of the proton for the gluon (dashed green), up quark (solid black) and down quark (dashed blue), at a scale of $Q^2 = 10 \text{ GeV}^2$ (left panel) and $Q^2 = 10000 \text{ GeV}^2$ (right panel). The PDF were obtained using CETQ6ll set.

CHAPTER 2. THEORY ELEMENTS

Chapter 3

MC Simulations and Recent Results

When performing measurements at the LHC or any other particle accelerator, an experimentalist wants to obtain a clear picture of what the Nature produced inside the detector in order to compare its observations to theoretical predictions and that way better understand the physical laws of the fundamental interactions.

Performing a proper measurement is certainly not easy, the raw measurements have to be corrected for detector and other experimental effects before accurate conclusions can be drawn. Predicting an observation is far from being straightforward as well. Starting with the Lagrangian developed in the previous chapter, one possibility is to perform calculations in a perturbative way leading to predictions expressed as a power series in the coupling constant. This approach can be used and makes sense as long as the running coupling constant is small enough. For QCD, as described in section 2.3.1, this is the case for large enough Q^2 but ceases to be valid for scales below the GeV². Since what we really collide at the LHC are hadrons, *i.e.* confined quarks and gluons (similarly for the hadrons observed in the final state), the non-perturbative regime of QCD will inevitably appear somewhere in the prediction. However the hard part of the interaction, taking place at a large Q^2 can be calculated by perturbative QCD (pQCD).

Using phenomenological models for the non-perturbative part, and calculating the harder process in pQCD, one can in principle obtain a prediction for a particular process studied at the LHC. However, all these calculations are still extremely complex. The structure of the events recorded at the LHC is also very complicated. The consequence of these two facts is that numeric simulations are unavoidable to simulate realistic events. Such simulations are obtained by MC programs that effectively divide the production of an event into smaller successive tasks that can be dealt with using both analytic and numeric computations. Finally, once events have been generated, they still need to be

processed to simulate the detector response. This last step also requires MC simulations.

3.1 Monte Carlo Generators

Different MC generators exist and use different ways to simulate high energy processes. Some provide a possibility to generate a complete event, while others provide the hard interaction part only and have to be supplemented by other MC to finalise the event generation. At the end, the same global picture depicted in figure 3.1 is followed. In this conception the generation of a hadron-hadron collision will follow the successive steps listed here:

- Calculation of the fully differential cross section for the process under consideration for the interaction of a pair of incoming partons extracted from the colliding hadrons. This corresponds to the hard process as discussed in section 2.4.1.
- The hadron-hadron differential cross section is calculated from the partonic cross section convoluted with the appropriate PDF.
- The actual particles involved in the process are generated for some phase space configuration according to the cross section computed in the previous step.
- Modelling of initial state radiation (ISR) happening when some charge (in the large sense: electromagnetic, weak or strong) is accelerated, leading to supplementary particles.
- Similar radiations occur for the final state particles, called final state radiation (FSR).
- Eventual short-life particles (such as W/Z bosons, π^0) are decayed producing additional final state particles.
- The colliding hadrons being made of partons, additional parton-parton interactions may take place in parallel with the hard interaction, resulting in multiple parton interactions (MPI). It is therefore necessary to simulate those interactions that contribute to the whole structure of an event.
- Remanent of the colliding hadrons needs to be taken into account for a correct balance in momentum and charge.
- As the final state partons of the hard interactions are moving apart from each other, the strong running coupling constant is increasing which eventually lead



Figure 3.1: Schematic representation of a pp collision generated by a typical event generator. A gluon and a quark (or antiquark) from the incoming hadrons participate in the hard scattering (red blob) from which a gluon, a quark (or antiquark) and a boson subsequently decaying into a quark-antiquark pair, are produced. Softer multiple interactions (purple blobs) are also represented. The fragments of the initial hadrons are treated (cyan). To the incoming parton, additional radiations are attached (initial state radiation). Similarly, additional radiations are attached to final state partons (final state radiation). The quarks and gluons produced during the interactions and the showering are turned into hadrons by hadronisation and then hadrons may decay (green). Sketch obtained from [25].



Figure 3.2: One-loop Feynman diagrams for the Z + jets process.

to confinement effects. Pair productions of quark-antiquark take place and the partons group to form the hadrons observed in the detector. This process is called hadronisation.

• As the final step, long-life particles such as τ leptons or B-hadrons decay while those reaching the detector are left intact.

The different pieces of a hadron-hadron event generation having been listed, different MC generators are now described with an emphasis on what they implement and in what they differ from each other. This will concern only the event generation and not the detector simulation that is described in the next chapter.

The calculation of the hard process can be done by different MC generators among which belong MADGRAPH 5 [26], SHERPA [27], PYTHIA [28,29] and POWHEG BOX [30]. However, not all of them can perform loop calculations present beyond LO diagrams. NLO generators such as SHERPA 2 and POWHEG BOX are able to perform loop calculations and the diagrams shown in figure 3.2 are considered when generating Z + jets events. However, calculations including loop corrections for processes containing a high number of partons are extremely difficult to perform and are only available for a limited number of processes. Tree-level generators are therefore very important for more complex processes for which NLO predictions are not yet available.

3.1.1 **Pythia**

PYTHIA is a general purpose tree-level generator able to compute matrix elements (ME) for a large number of processes. It also implements the necessary tools to simulate ISR and FSR that allow it to fully simulate an event. Two versions of this MC generator, namely PYTHIA 6 [28] and PYTHIA 8 [29], are used for the work performed in this thesis. The former is also the older and is written in FORTRAN and was used successfully for many years, while the latter is more recent and is a complete rewrite from FORTRAN

to C++, which already offers a complete replacement for most applications, notably for LHC physics studies and provides new features such as the additional interleaving of FSR along with ISR and MPI that should improve the description of data. Nevertheless, the concept behind the two versions is the same. Even though PYTHIA was used only to provide parton showering and hadronisation to other generators and not to provide ME predictions for the analysis performed in this thesis, it is interesting to look at how it calculates matrix elements.

Hard Processes

PYTHIA can generate hundreds of different hard processes and is optimised for $2 \rightarrow 1$ and $2 \rightarrow 2$ processes. For each hard partonic process the corresponding matrix element is hard coded within the software.

The hard cross section is calculated as the product of four factors as follows:

- 1. the factor $\frac{\pi}{s}$ giving the overall dimensions of the cross section, in GeV⁻²,
- 2. a Jacobian compensating for possible changes from initial to final phase space volume,
- 3. PDF weights, obtained from PDF internal or external libraries,
- 4. dimensionless cross section $(\hat{s}^2/\pi)d\hat{s}/d\hat{t}$, with $\hat{t} = (p_1 p_3)^2$, which is the factor that has to be hard coded for each process.

This is for the case of $2 \to 2$ processes. However, Z boson production is seen as a $2 \to 1$ process producing a resonance. In that case, the last factor $d\hat{s}/d\hat{t}$ is replaced by $\hat{\sigma}(\hat{s})$. The produced resonance is then made to decay following fixed probabilities.

For the Z + jets process studied in this thesis, PYTHIA can generate the $2 \rightarrow 1$ process $f_i \bar{f}_j \rightarrow \gamma^*/Z$ taking into account the interference between the two production modes at LO as computed in the previous chapter. It can also produce the $2 \rightarrow 2$ processes $f_i \bar{f}_j \rightarrow g \gamma^*/Z$ and $f_i g \rightarrow f_k \gamma^*/Z$. However, the recommendation for inclusive generation of γ^*/Z production is to use the $2 \rightarrow 1$ with specific matrix-element-inspired corrections to the parton shower (see section 3.1.5) allowing for a good description of the full $p_{\rm T}$ spectrum of the gauge boson. When the focus is on high- $p_{\rm T}$ values, then one should use the $2 \rightarrow 2$ processes adding shower on top of it. This approach is however dangerous due to the low- $p_{\rm T}$ divergencies appearing in these ME. It has to be noted that in PYTHIA, Drell-Yan conventionally refers to the $q\bar{q} \rightarrow \gamma^* \rightarrow l^+l^-$ process well below the Z mass.



Figure 3.3: Tree-level Feynman diagrams for the Z + jets. The top Feynman diagram is the lowest order Feynman diagram for the DY process and has zero parton in its final state (no α_s coupling is present in this diagram). The second row lists tree-level diagrams at second order (first order in α_s) leading to one parton in the final state. Important to note is the appearance of an additional production channel for which we have a gluon in the initial state. The third row is another higher order in α_s for which there will be two partons in the final state. The Z boson can be produced by two incoming gluons at this order. Additional diagrams with two partons attached to the same initial parton as well as gluon splitting and quark gluon radiation are missing to complete the set of tree-level diagrams at the NNLO order in α_s .

Parton Shower

The generation of the hard process of an event being only the first step in simulating a full hadron-hadron collision, every generator has to include parton shower (PS) effects as explained above.

As already mentioned in the previous section, two types of radiation are distinguished

when considering parton showering effects. The FSR models the fact that when final state partons, *i.e.*, partons present on the right side of the Feynman diagrams in figure 3.3 are moving apart from the hard process their virtuality is allowed to decrease while softer and softer gluons or quark-antiquark pairs are produced. On the other side, the ISR models the radiation effects happening far before the hard scattering process. As the colliding partons approach, they can emit harder and harder gluons until reaching the virtuality of the hard process.

The next two sections discuss in some more detail how FSR and ISR are modelled in PYTHIA.

Final State Radiation

Final state radiation consists in the transformation of a parton a in a parton b by the emission of a parton c. For example we could have $q \rightarrow gq$, $g \rightarrow q\bar{q}$ or $g \rightarrow gg$. Several splittings can take place successively. FSR is therefore modelled by a series of $a \rightarrow bc$ splittings resulting in a shower of partons. From momentum conservation, each subsequent parton possesses a fraction of the initial parton momentum. The evolution of the shower is therefore parametrised by the energy fraction z carried by one of the two emerging partons, $z = E_b/E_a$. Another parameter is also present in the description of PS, namely the ordering variable t. This variable means that the PS is happening with decreasing t, *i.e.* the first radiation is taking place at a value of t larger than the one of the next radiation, itself larger than the one of the subsequent emissions. Several choices are possible for the ordering variable, one of the most common being the virtuality Q^2 of the parton that is going to split, $t = Q^2 = p_a^2$ and is used in PYTHIA 6¹. Another choice is to use the transverse momentum instead. This last choice is used in PYTHIA 8 as well as in SHERPA.

More specifically, let us consider the emission of a gluon from a quark present in the final state that can be found in Z + jets from the diagrams shown in figure 3.3. These diagrams suffer from collinear divergences that the PS has to take care of. In the collinear approximation the probability $d\mathcal{P}_{a\to b}$ for such an emission can be written in terms of z and $t = \ln(Q^2/\Lambda_{\rm QCD})$ as:

$$d\mathcal{P}_a(z,t) = \sum_b \frac{\alpha_s}{2\pi} P_{ba}(z) dt dz, \qquad (3.1)$$

where P_{ba} are the splitting functions presented in chapter 2, section 2.4.3. This equation presents the collinear divergence in its $1/Q^2$ dependence, while soft divergences arise in the limit $z \to 1$ of the splitting functions.

¹Pythia 6 supports different options for the choice of the ordering variable.

By requiring the conservation of total probability, it is possible to handle these divergences. As a first step we consider the full branching probability between t and t + dt by integrating over every momentum fraction z, $z_{\min}(t) < z < z_{\max}(t)$:

$$d\mathcal{P}_a(t) = \left(\sum_b \int_{z_{\min}(t)}^{z_{\max}(t)} \frac{\alpha_s}{2\pi} P_{ba}(z) dz\right) dt.$$
(3.2)

This probability has to be weighted by another factor in order to conserve the total branching probability. This additional factor, called Sudakov form factor, represents the condition that no emission takes place between the starting scale t_0 and the scale t at which we consider the emission. The branching probability conserving the total probability is therefore given by:

$$d\mathcal{P}_{a}^{FSR}(t) = d\mathcal{P}_{a}(t) \times \exp\left(-\sum_{b} \int_{t_0}^{t} dt' \int_{z_{\min}(t')}^{z_{\max}(t')} \frac{\alpha_s}{2\pi} P_{ba}(z') dz'\right).$$
(3.3)

This formal development makes possible to simulate a realistic cascade of successive parton emissions. In practice, for each branching, we choose the scale t according to the probability (3.3), the type of the branching with probability proportional to the integrated splitting functions, and the z fraction according to the unintegrated corresponding splitting function. This steps are repeated for each splitting until a certain virtuality of order $\Lambda^2_{\rm QCD}$ is reached and where non-perturbative effects emerge. At this stage the hadronisation model (see next sections) is taking care of combining the partons into hadrons.

Initial State Radiation

Some of the developments made for the FSR are still applicable for the case of ISR. However, ISR is slightly more complex and additional elements enter into the formulae. Even though the approach of starting from a parton far before the interaction and evolving it to higher and higher scale until the scale of the hard interaction is reached and, if the evolution allows the hard process to happen, generate the hard interaction and the whole event is correct, this approach is very inefficient especially when the phase space of the hard interaction is small. Another approach has to be used for initial state radiation. The trick is to generate the hard event first and then going back in time tracing the history attached to the partons used as the incoming particles of the hard interaction. This method is known as backward evolution.

The second difference with respect to FSR resides in the presence of the PDF in the equations of ISR. Indeed, in the backward evolution the probability for a parton b to be one of the emitted particles radiated by a parton a in going from scale t to scale t - dt

3.1. MONTE CARLO GENERATORS

(virtuality is now decreasing from hard scale to softer and softer scales) is given by:

$$d\mathcal{P}_b(t) = \sum_a \int \frac{x'}{x} \frac{f_{a/A}(x',t)}{f_{b/A}(x,t)} \frac{\alpha_s}{2\pi} P_{ba}\left(\frac{x}{x'}\right) |dt|, \qquad (3.4)$$

where $f_{a/A}$ is the PDF for a parton of type *a* contained in the colliding hadrons *A*. As was done for the FSR, a non-emission Sudakov form factor has to be added to this expression in order to conserve the total probability. The resulting ISR probability can therefore be expressed as:

$$d\mathcal{P}_b^{ISR}(t) = d\mathcal{P}_b(t) \times \exp\left(-\int_t^{t_{\max}} dt' \sum_a \int \frac{x'}{x} \frac{f_a(x',t')}{f_b(x,t')} \frac{\alpha_s}{2\pi} P_{ba}\left(\frac{x}{x'}\right)\right) dt.$$
(3.5)

To simulate ISR, it is thus necessary to pick up a scale t at which the branching occurs according to (3.5), the type of parton a according to the relative ratios of the integrated splitting functions for different allowed types a, and the energy fraction z according to the unintegrated corresponding splitting function.

Multiparton Interaction

So far in the generation of a hadron-hadron event, only single parton-parton interactions were considered to take place during the collision. However, at high energies the probability that a second parton-parton collision occurs within the same pp collision becomes significant enough to be observed. Such additional parton-parton collisions, referred to as multiparton interactions (MPI), will affect the distributions and therefore need to be taken into account. This subject is not well known and many studies are being conducted at the LHC [31] and within CMS [32] in order to better understand the kinematics and correlations between these MPI. In particular, some studies try to estimate the MPI production in presence of two hard scales, called double parton scattering (DPS) like the Z+ 2 jets process where the 2 jets come from a second parton-parton interaction and not from parton radiations attached to the Z hard event, as depicted in figure 3.4.

Hadronisation

When the energy scale is of the order of $\Lambda_{\rm QCD}$, the strong running coupling constant is no more small with respect to unity and the whole perturbative approach on which Feynman diagrams and by consequence perturbative calculations are based can no longer hold. At this stage, the multiple partons produced by the hard scattering and during the parton showering recombine into hadrons forming colourless states.



Figure 3.4: Single Parton Scattering (left) as opposed to Double Parton Scattering (right) for the production of two jets in association with a Z boson.

As mentioned, this phenomena cannot be described by a perturbative expansion and phenomenological models have to be used. The Lung string model used in PYTHIA is one of them. In this approach, partons moving apart from each other are connected by a string whose energy is proportional to its length and of the order of 1 GeV fm^{-1} . When the partons reach some threshold separation, the energy contained in the string becomes large enough to lead to the creation of a pair of a quark and its antiquark. At this point the string breaks, a quark-antiquark pair is created cancelling out the colour field present between the two initial partons. The two pairs of partons continue to move away one from the other. The process eventually repeats and gives rise to the formation of hadrons. The type of $q\bar{q}$ pairs created from the break up of the string is modelled by empirical functions, called fragmentation functions, tuned to reproduce the observed spectrum from data.

After the hadronisation procedure has been applied to the generated event, and the whole procedure repeated multiple times in order to generate a large enough sample of such events, the simulated events can in principle be used as an input to the simulation of the detector effects (see next chapter) or directly compared to unfolded (corrected for detector effects) data or compared to other MC predictions.

3.1.2 MADGRAPH 5

MADGRAPH 5 [26] is another MC generator which, unlike PYTHIA, provides only ME calculations. Its output has therefore to be interfaced with another MC generator implementing the needed remaining steps such as parton showering and hadronisation in order to get full event predictions. MADGRAPH 5 is a leading order ME generator in the sense that, like PYTHIA, it does not calculate diagrams containing loops. Only tree-level diagrams can be computed in MADGRAPH 5². Still MADGRAPH 5 is able to generate events with up to 4 partons in the final state for the process Z + jets of interest for the analysis performed in this work. Examples of Feynman diagrams taken into account by MADGRAPH for the Z + jets process are presented in figure 3.3 (page 40).

Because of its tree-level nature, MADGRAPH has to deal with one main issue: soft and collinear divergences. Such divergences appear at $O(\alpha_s)$ calculations and would be cancelled by some of the $O(\alpha_s^2)$ terms in the perturbative series. For example, soft divergences appearing in the expressions of the Feynman diagrams with one final state parton in figure 3.3 are cancelled when considering the one-loop diagrams in figure 3.2. To avoid these singularities, the phase space of tree-level generators has to be carefully tailored so that the predictions are performed away from soft and collinear divergences. These excluded regions have therefore to be treated by PS calculations.

Hard Process

MADGRAPH 5 adopts a totally different approach than PYTHIA to compute ME. Indeed, it can generate ME at tree-level for any Lagrangian based model (renormalisable and effective). From a user's input specifying the initial and final state particles, MAD-GRAPH 5 generates all Feynman diagrams for the process, and outputs the computer code necessary to evaluate the ME at a given phase space point. This generated code can subsequently be used by other packages such as MADEVENT for event generation.

The algorithm used to generate the diagrams is based on recursively creating subdiagrams from the diagrams by merging legs, carefully avoiding double-counted diagrams. It is described in [26] and is applied concretely to the specific process $u\bar{u} \rightarrow ge^+e^-$ below with a few explanations to illustrate the procedure.

First iteration:

1. From the Lagrangian describing the model of interest, in this particular case the

²A newer version of MADGRAPH, merged with another generator AMC@NLO is now available and is providing NLO loop calculations for a large number of processes. See section 3.1.4.

SM, create a list of the relevant vertices:

$$(u\bar{u}g), (u\bar{u}Z), (u\bar{u}\gamma), (e^+e^-Z) \text{ and } (e^+e^-\gamma)$$

2. Transform each incoming particle/antiparticle to its equivalent outgoing counterpart. This renders all particles outgoing, which makes the algorithm more general as illustrated in the next figure where arrows entering (leaving) the blob indicate incoming (outgoing) particles. For later use, a flag attached to each external particle is defined, "from group", and set to true.



- 3. If there is a vertex combining all the n = 5 external particles, create the corresponding group [(1, 2, 3, ..., n)] if at least two particles have "from group" to true. This will give a valid diagram. In the treated case, there is of course no vertex to which the five particles can be associated together, so no grouping $(u, \bar{u}, g, e^+, e^-)$ is possible, at the first level of the iteration.
- 4. Create all allowed groupings of particles with at least one particle having "from group" to true. It is for example allowed to group e^+e^- together because they can annihilate, while grouping u and e^- is not allowed in the SM, there are no interaction between these two particles. The list of possible groupings for the case of interest is found to be:

$$\begin{split} &[(u,\bar{u}),g,e^+,e^-],\\ &[(u,g),\bar{u},e^+,e^-],\\ &[u,(\bar{u},g),e^+,e^-],\\ &[u,\bar{u},g,(e^+,e^-)],\\ &[(u,\bar{u}),g,(e^+,e^-)],\\ &[(u,g),\bar{u},(e^+,e^-)],\\ &[u,(\bar{u},g),(e^+,e^-)]. \end{split}$$

3.1. MONTE CARLO GENERATORS

5. Replace the grouped particles with the possible particles resulting from the interactions and set "from group" to true for these new particles and false for any particle that has not been combined in this iteration.

$$\begin{split} [(g),g,e^+,e^-], \ [(\gamma),g,e^+,e^-], \ [(Z),g,e^+,e^-], \\ [(u),\bar{u},e^+,e^-], \\ [u,(\bar{u}),e^+,e^-], \\ [u,\bar{u},g,(\gamma)], \ [u,\bar{u},g,(Z)], \\ [(g),g,(Z)], \ [(\gamma),g,(\gamma)], \ [(\gamma),g,(Z)], \ [(Z),g,(\gamma)], \ [(Z),g,(Z)], \\ [(u),\bar{u},(\gamma)], \ [(u),\bar{u},(Z)], \\ [u,(\bar{u}),(\gamma)], \ [u,(\bar{u}),(Z)]. \end{split}$$

6. Repeat from 3 for the reduced set of external particles as long as at least 3 external particles remain, as follows.

Second iteration:

3. The resulting reduced sets with four particles all have only one particle with a "from group" value of true, for this reason they can not give valid diagrams. From the three particles groups, only the combinations

 $((u), \bar{u}, (\gamma)), ((u), \bar{u}, (Z)), (u, (\bar{u}), (\gamma)), (u, (\bar{u}), (Z))$

are allowed by the interactions.

- 4. Further combination will result in an external state with only one "from group" true particle, from which no diagram can be obtained.
- 5. The iteration stops since no external particles are left.

The four valid groups correspond to the radiation of a gluon from the u quark in the photon mode, radiation of a gluon from the u quark in the Z boson mode, and the corresponding two diagrams for the gluon radiation from the antiquark.

For each of the resulting Feynman diagrams, a code for the ME evaluation is generated in terms of function calls from either HELAS (HELicity Amplitude Subroutines for Feynman Diagram Evaluations) [33] or ALOHA (Automatic Libraries Of Helicity Amplitudes for Feynman Diagram Computations) [34] libraries capable of computing the amplitudes of arbitrary tree-level Feynman diagrams.

3.1.3 Sherpa

Simulation of High Energy Reactions of Particles (SHERPA) is another generator that allows for complete hadronic final states in high-energy particle collisions. Different versions exist among which SHERPA 1.4 and SHERPA 2 are of interest for this thesis. The specificity of SHERPA is its modularity. Indeed, it comes with various modules that the user can decide to activate or not, or even replace by external programs. For example, one could use some specific ME generator inside the SHERPA framework. The main difference between the 1.4 and 2 versions are the accuracy of the ME predictions. The former is a leading-order ME generator able to compute Born-level diagrams with up to 8 partons. The implemented algorithm is different than the one used by MADGRAPH. The version 2 allows for NLO calculations and for the case of Z + jets with up to four jets, it predicts the 0, 1 and 2 jets multiplicities with NLO in QCD accuracy and the additional 3 and 4 jets multiplicities are predicted with LO MEs.

Parton Shower

SHERPA implements its own PS simulation as well as its matching between the ME predictions and PS. The implementation of PS in SHERPA is based on the Catani-Symour dipole factorisation formalism [35]. In this approach, showers are made of splitting dipoles composed of the parton expected to split together with a spectator parton colour-connected to the emitter. Dipoles can therefore be made of two final states partons, two initial state partons or one initial and one final state partons, all configurations being treated equally unlike in PYTHIA where a distinction is made between initial and final state radiation. The ordering parameter of the shower is following the value of the transverse momentum of the final state splitting products with respect to the emitting beam particle.

Hadronisation

Another difference between SHERPA and PYTHIA is the way hadronisation is modelled [36–38]. Indeed, SHERPA assumes a local parton-hadron duality that closely links quantum numbers at hadron and parton levels. Along the hadronisation process particles close to each other in phase space are grouped together into clusters which, if light enough, are converted into stable hadrons, otherwise are forming heavy hadrons subsequently decaying into lighter clusters. In a more systematic way, the hadronisation can be seen as the following sequence:

• first gluons are decayed into quark-antiquark (dipole) or gluon-gluon (dipole pairs)

and other partons are brought on-shell,

- secondly, neighbour particles are merged into clusters,
- heavy clusters decay by emitting a gluon from a quark-antiquark pair, subsequently splitting as in the first step,
- finally, clusters decay into hadrons by mean of various weights for proper flavour, phase space and other dynamical probabilities.

3.1.4 MADGRAPH5 AMC@NLO

More recently the MADGRAPH5_AMC@NLO program [39] has been released providing an automated way for the computation of tree-level and next-to-leading order differential cross sections, and their matching to PS simulations. It is a framework grouping the features of the MC generator MADGRAPH 5 presented above and of AMC@NLO, additionally including other components.

The major capability of MADGRAPH5_AMC@NLO is the possibility to compute tree-level and one-loop amplitudes for arbitrary processes. As being partly based on MADGRAPH 5, the central idea behind the ME computations realised in this framework is the same and based on translating Feynman diagrams into computer codes. As explained previously, this can be done at LO given the Lagrangian of a theory. However, at NLO, loop diagrams come with difficulties that cannot (yet) be handled by the automation program FEYNRULES [40] which does not currently compute some NLO-specific terms such as those coming from UV counterterms. Hopefully those are found in a finite and typically small number and can be introduced by hand.

For the present work, a DY + jets sample generated at NLO accuracy with this framework has been used, interfaced with PYTHIA 8 for the PS, for comparison with the data.

3.1.5 Matching of Matrix Element and Parton Shower

The predictions obtained by ME generators (such as PYTHIA, MADGRAPH, SHERPA) are good at describing well separated parton configurations but as already announced, they suffer from divergences in the soft and collinear regions, thus they are not suitable for a proper description of a jet's internal structure. Furthermore, since hadrons are what we observe in nature, fragmentation models need to be applied to the partons.

Parton showering methods on the other hand, as implemented by PYTHIA for example, will produce realistic parton configurations given a basic hard process. They are derived in the collinear limit handling divergences with the use of Sudakov form factors making them well suited to describe jets evolution. The fact that parton showering is obtained in a collinear approximation has the counter effect that largely separated parton configurations will not be properly estimated.

Combining the advantages of the ME and PS seems therefore the proper way of proceeding and is the way ME predictions are used for high energy physics predictions. Doing so, great care must be taken in order to avoid double counting. Indeed, a phase space configuration with n partons coming from the ME can be reproduced by, *e.g.*, the (n-1) ME partons configuration to which PS adds an additional hard emission. Different matching prescriptions exist to combine ME and PS. Three of them are briefly described below.

Parton Shower Reweighing

The idea behind this technique was suggested in [41] in 1986 because of "some serious discrepancies between PETRA/PEP data and ME multijet predictions in e^+e^- annihilations events". The approach was to start from the lowest order hard process, generate the PS and reweigh it in order to mimic a first order ME radiation.

To illustrate this, let us take the original problem of $e^+e^- \rightarrow \text{jets}$ process for which no initial state QCD radiation is present. The lowest order process is simply the creation of a quark-antiquark pair $e^+e^- \rightarrow q\bar{q}$ and the first order QCD correction is found to be describing the radiation of a gluon from the quark or the antiquark $e^+e^- \rightarrow q\bar{q}g$. This radiation can be obtained from ME calculation or from PS. In both cases the cross section for this three jet configuration is given by:

$$\frac{1}{\sigma}\frac{d^2\sigma}{dx_1dx_2} = \frac{2}{3}\frac{\alpha_s}{\pi}\frac{A(x_1, x_2)}{(1 - x_1)(1 - x_2)},\tag{3.6}$$

where $x_i = 2E_i/\sqrt{s}$ with $A(x_1, x_2)$ given by:

$$\begin{aligned} A_{\rm ME}(x_1, x_2) &= x_1^2 + x_2^2, \\ A_{\rm PS}(x_1, x_2) &= 1 + \frac{1 - x_1}{(1 - x_1) + (1 - x_2)} \frac{x_1^2}{(2 - x_2)^2} + \frac{1 - x_2}{(1 - x_1) + (1 - x_2)} \frac{x_2^2}{(2 - x_1)^2}, \end{aligned}$$

from which in the collinear limit $x_1 \to 1$ $(x_2 \to 1)$ we have $A_{\rm ME} = A_{\rm PS} = 1 + x_2^2$ $(= 1 + x_1^2)$, and $A_{\rm ME} \leq A_{\rm PS}$ everywhere. Therefore, by generating the branchings of partons 1 and 2 using PS algorithms but only accepting the branchings with probability



Figure 3.5: Ratio of the three-jet ME factor $A_{ME}(x_1, x_2)$ to the three-jet PS factor $A_{PS}(x_1, x_2)$ for the process $e^+e^- \rightarrow q\bar{q}g$, as a function of x_1 and x_2 .

 $A_{\rm ME}(x_1, x_2)/A_{\rm PS}(x_1, x_2)$ (shown in figure 3.5), it is possible to obtain the result of the three-jet ME calculation³.

The matching prescription can be extended to hadron collisions for which initial state radiations play an important role. The description for the case of W+jets production can be found in [43]. In PYTHIA the matrix-element-inspired correction to the first emission (mentioned in section 3.1.1) is using this technique.

CKKW Matching

The Catani-Krauss-Kuhn-Webber (CKKW) matching prescription [44] was also originally proposed for e^+e^- annihilation into hadronic final states processes and afterwards extended to the case of hadron collisions in [45]. It relies on the separation of the phase space in two regions, delimited by a cutoff y_{cut} defined according to the distance measure of the k_T -clustering algorithm (see section 5.5), such that the PS populates the phase space below the cutoff, and ME filled the region above y_{cut} . In the CKKW prescription, not only the PS is vetoed following some criteria as in the previous method, the ME

³This prescription is not fully consistent and problem can emerge when the first radiation is not the hardest as it could happen in angular ordered shower. However, self-consistent algorithms can still be obtained as described in [42].

prediction itself is also modified.

Indeed, the CKKW prescription aims at giving a refined description of the rate of jets compared to that of PS by replacing collinear approximations appearing in the splitting functions with exact ME, resulting in a weight to be applied to the ME *n*-parton configuration for which all *n* partons are resolvable for the cutoff y_{cut} . The remaining phase space region, below the cutoff, is then populated with plain PS vetoed for hard emission above the y_{cut} threshold.

The SHERPA predictions employed in the analysis of the present work are obtained using this matching prescription to merge ME and PS calculations.

MLM Matching

In the Michelangelo Mangano (MLM) matching prescription, one starts by generating parton-level configurations for a given hard-parton multiplicity, n_{part} , with a p_{T} and distance threshold: $p_{\text{T}} > p_{\text{T}}^{\min}$ and $\Delta R_{ij} > R^{\min}$. From the ME events, one performs plain parton showering using the desired algorithm (such as PYTHIA). Then a jet algorithm is applied on the showered event before hadronisation in order to define the event jet multiplicity n_{jet} . From the n_{part} parton-level jets and the n_{jet} post-shower jets, a matching algorithm is applied resulting in:

- exclusively matched events for which each parton is matched to exactly one jet and vice versa (n_{part} = n_{jets}),
- inclusively matched events for which each parton is matched to a jet but unassociated post-shower jets may exist $(n_{\text{part}} \leq n_{\text{jets}})$.

Finally, the exclusive samples with differing ME multiplicities are combined up to the desire multiplicity for which the inclusive sample is taken instead of the exclusive one. For example, for the case of the inclusive Z + jets with up to 4 partons at the ME, one would construct an inclusive sample as follows:

$$Z + \text{jets} = Z + 0j|_{\text{excl.}} \oplus Z + 1j|_{\text{excl.}} \oplus Z + 2j|_{\text{excl.}} \oplus Z + 3j|_{\text{excl.}} \oplus Z + 4j|_{\text{incl.}}$$

This prescription is used for the MADGRAPH 5 DY + jets sample used in the present analysis.

3.1.6 Pile-up

At hadron colliders and in particular at the LHC the high rate of collisions of bunches of protons is essential to obtain the large statistic needed for the discoveries and measurements of rare processes. However, reaching such a high statistic in a short time has a price. Indeed, many *pp* interactions occur during a single bunch crossing leading to contamination of the rare processes by the other interactions. This is called pile-up and has to be accounted for in the simulation in order to properly describe the superimposed interactions in the detector. This is achieved by adding softer inelastic collisions to the hard process of interest. The average number of interactions being highly dependent on the accelerator conditions during the data taking, the observed vertices multiplicity varies from run to run. On the other hand, MC samples are created with a well defined number of interactions distribution. It is therefore necessary to reweigh the simulated events, on a run by run basis, such that the MC pile-up distribution matches the data one.

3.1.7 The Use of Different MC

The previous sections have described in a general way how the MC simulations are obtained. The whole process is undoubtedly much more complicated than what has been presented here and would easily fill another thesis by itself. The point here is that by comparing different MC generators predictions, using different approaches to predict the observations performed at the LHC, physicists can spot out the potential problems or successes of one or another generator. Typically, one could study several PS models performance using the same ME generator. Doing so, the difference between the predictions would be a direct study of the parton showering model. One could of course perform the study the other way around and use the same PS model with different generators. Another test could consist in using LO PDF for one simulation and NLO PDF for another simulation without changing anything else to the generation of the events. In other words, by switching between different parameters of the simulation one can spot out the pieces of a simulation that need to be reviewed or better that are very successful. Since we know what is implemented in such MC simulations, we can better understand the physics of the fundamental particles interactions.

3.2 LHC and Tevatron Measurements

In this section the latest results anterior to this thesis and concerning the production of jets in association with a vector boson are presented. A particular attention is brought to



Figure 3.6: The measured (circles) and predicted (coloured bands) W, Z and $t\bar{t}$ fiducial cross sections inclusive and for different jet multiplicities, at centre-of-mass energies of $\sqrt{s} = 7 \text{ TeV}$ (empty circles and blue bands) and $\sqrt{s} = 8 \text{ TeV}$ (filled circles and green bands). The 8 TeV CMS measurement is only performed for the inclusive cross section. Corresponding measurements for different jet multiplicities is precisely the subject concerned in this thesis.

Z vector boson produced in association with jets as it constitutes the particular process studied in the present work.

3.2.1 Main Motivation

The study of the production of jets associated with a vector boson W^{\pm} or Z allows to probe different aspects of pQCD calculations with a high level of precision. Figure 3.6 presents the different cross section measurements performed by the CMS Collaboration at the LHC since the first pp collisions taking place at a centre-of-mass energy of 7 TeV, and prove the large activity of this field of study. Among the possible analyses that can be performed, we find the study of the topological properties of the events, the study of the jet multiplicity or the kinematic properties of the jets. These studies can then be used to constrain the PDF. They are also important for searches where many exotic particles are expected to decay into W or Z bosons. Finally, and maybe most important reason during the past couple of years, precision measurement of vector boson production is crucial for Higgs boson studies and Beyond Standard Model (BSM) searches as it is an important background for these processes. Additionally, one of the advantages of vector boson plus jets events is their relatively high production rate and their simple decay signature in the detector that makes them the ideal candidates for checking and tuning MC generators. Not only simulations benefit from boson plus jets studies. The measurements themselves are improved by studying Z + jets events as they constitute ideal candidates for detector calibration response. As mentioned in section 5.5, the jet energy response is calibrated using Z + 1 jet events and the fact that the Z boson energy resolution is very good.

The real beginning of the measurements of vector boson plus jets events properties starts with the UA1 and UA2 experiments at CERN back in the 80's. These detectors were build around the SPPS accelerator (very similar to the SPS see section 4.1.5) to study proton-antiproton collisions and trace the production of the W and Z bosons that were then discovered in 1983. These experiments provided one of the first jet multiplicity measurements associated with a vector boson as well as the jets kinematics distribution as shown in figure 3.7 [46]. At this early time, the statistics and the centre-of-mass energy did not allow for precise results and comparison to theoretical predictions. But that was only the beginning of the vector boson plus jets story.

3.2.2 LHC and Tevatron

The Tevatron and the LHC accelerators provide now much higher energies and statistics necessary to study the details of these processes. The observed cross sections are compared with many different MC theory predictions and are used to tune the subsequent releases of these softwares. The LHC (see chapter 4) has provided pp collisions at centre-of-mass energies of 7 TeV and 8 TeV in years 2011 and 2012 respectively. The corresponding integrated luminosities (*i.e.* the total amount of data, see section 4.1.4 for a definition of luminosity) are 5 fb⁻¹ and 20 fb⁻¹. The Tevatron, located in the United States, at the Fermi National Accelerator Laboratory, produced proton-antiproton collisions at a centre-of-mass energy of 1.96 TeV, between years 2001 and 2011. The total integrated luminosity of Tevatron is of about 10 fb⁻¹. These two accelerators produced about thirteen thousands $Z \rightarrow ee$ associated with at least one jet events for Tevatron and some hundred and ninety thousands such events for the LHC, to be compared with the fifty one $Z \rightarrow ll$ events of the UA1 experiment collected before its shutdown.

An important point is that the LHC is not just a simple rescaling of the Tevatron experiment. Their different energies and different collisions (pp for LHC and $p\bar{p}$ for Tevatron) induce several important consequences. The two accelerators provide different mixtures of quarks and gluons jets in the final state. For example, the leading jet is found to be approximately an equal mixture of quarks and gluons jets at the Tevatron



Figure 3.7: The multiplicity distribution of W + n jets(s) and Z + n jet(s) (top). The transverse energy distribution of the jets (bottom). As can be seen from the top right plot, only 16 Z events were observed at this early time and only two of them had an associated jet with a transverse momentum below 10 GeV. Results extracted from [46].

while at the LHC, it will be mostly quark originated. Additionally, the $q\bar{q}$ contribution is obviously bigger at the Tevatron than at the LHC. Finally, the larger energies of the LHC induce a larger contribution from processes with heavy flavour in the initial state as well as higher transverse momentum and multiplicity values. All these differences lead to different PDF regimes (x values and flavour types) available at the two accelerators and make them partially complementary. Table 3.1 summarises the characteristics of the LHC and Tevatron accelerators.

3.2.3 Recent Results

When comparing the results of different experiments, it is important to notice the phase space covered by each of them, which may be different between experiments. For instance, the lepton $p_{\rm T}$ threshold is 20 GeV for the CMS and ATLAS results, 25 GeV for CDF and 15 GeV for D0. Differences also arise for the pseudorapidity⁴ limits of the leptons as well as for the invariant mass window in the case of Z + jets studies. Additionally, the jet definition also introduces some choice of algorithm. At Tevatron the choice was to use Midpoint cone algorithm with a cone size of 0.7 and 0.5 for CDF and D0 respectively. At the LHC the anti- k_t algorithm is used with cone sizes of 0.4 and 0.5 for ATLAS and CMS respectively. The pseudorapidity range in which jets are considered is also different in the four experiments.

Many measurements have already been performed by the four experiments regarding vector boson associated with jets production. Here a summary of these results is presented. Results associated to the CMS Collaboration are extracted from [47] and [48],

	LHC	Tevatron - Run II
collisions	pp collider	$p\bar{p}$ collider
centre-of-mass energy	$\sqrt{s} = 7 \text{TeV} (2010-2011)$	$\sqrt{s} = 1.96 \mathrm{TeV}$
	$\sqrt{s} = 8 \mathrm{TeV} (2012)$	
main experiments	ATLAS, CMS	CDF, D0
Integrated luminosity	$\sim 5{\rm fb}^{-1}$ (2010-2011)	$\sim 10 {\rm fb}^{-1} \ (2001-2011)$
	$\sim 20 {\rm fb}^{-1} \ (2012)$	

Table 3.1: Summary of LHC and Tevatron characteristics.

⁴The pseudorapidity of a particle is given in term of its polar angle θ by $\eta = -\ln \tan(\theta/2)$. The pseudorapidity is an approximation of the rapidity y given by $y = (1/2) \ln((E + p_z)/(E - p_z))$, where p_z is the component of momentum along the beam axis.

results obtained by the ATLAS Collaboration are taken from [49] and [50], and results from the CDF and D0 Collaborations are from references [51] and [52], respectively. The observed cross sections are compared to different MC generator predictions including the few ones presented in the previous sections of this chapter. For the additional generators, all the relevant information and references may be found in the related experiment publications mentioned above.

We begin with the multiplicity of jets produced in association with a Z boson presented in figure 3.8 (page 60). As can be seen from each of these results, there is an excellent agreement between experiments and theory over more than four orders of magnitude in cross sections. We can also note the important statistic available at the LHC allowing for measurements of events with more than six jets. Interestingly, as shown on the ratio plots of CMS, the MADGRAPH 5 predictions, while being LO with up to four partons in the final state are able to describe correctly even higher multiplicities. This is an achievement in itself and shows that PS generated by PYTHIA6, associated to the Matrix Element predictions of MADGRAPH 5 is leading to very good results. Additionally, from the same picture, the NLO generator POWHEG, while generating only one jet with NLO accuracy, is also able to describe the data when coupled to PYTHIA6 for PS. This is confirmed by the ATLAS measurement extended at large jet rapidities and compared to ALPGEN that has a similar approach to MADGRAPH. Note that the MC@NLO version used in ATLAS results is different (older) than the one used in this work in MADGRAPH5 AMC@NLO. The same can also be seen on CDF results (bottom right plot). Finally, the NLO SHERPA predictions while being globally very good, seem to slightly underestimate the jet multiplicity cross sections as can be observed in the three experiments. In conclusion, we have a good description by fixed order NLO predictions as well as multi-leg ME + PS predictions. Another important observation is that the data measurements tend to be statistically more accurate than the MC predictions thanks to the really high luminosities furnished by the Tevatron and LHC machines.

Figure 3.9 (page 61) summarises the leading jet transverse momentum differential cross section results. The higher energies available at the LHC allow for a wider range in the jet $p_{\rm T}$ spectra. Results from each experiments look compatible with each other and simulations also describe relatively well the data. This said, from the ratio plots we note that SHERPA seems to underestimate the cross section while MADGRAPH and POWHEG are slightly overestimating it. The main region of discrepancy lies in the range between 175 GeV and 300 GeV. For these values CDF provides only two measurements but are already showing the discrepancy of an overestimation also visible in MADGRAPH 5, POWHEG and ALPGEN at the LHC.

Many other cross section measurements have been performed by the four experiments. Cross sections measured as a function of jets rapidity or pseudorapidity, as a function of
jets $H_{\rm T}$ (scalar sums of all jets $p_{\rm T}$), as a function of the vector boson $p_{\rm T}, \ldots$ Furthermore, all these measurements were also done for the case of W + jets. While this process has a cross section approximately ten times larger than the one of Z + jets, its experimental signature is less clean and the background contamination is a bit larger, leading to larger systematic uncertainties but smaller statistical uncertainties. Some of these results are presented below in figures 3.10, 3.11 and 3.12 on pages 62 to 64. From 3.10 one may notice the smaller range available for the third jet $p_{\rm T}$ distribution at Tevatron due to its lower centre-of-mass energy compared to the LHC. In figure 3.11 representing the inclusive jet multiplicity associated to the W boson, one can notice the larger systematics but smaller statistics uncertainties with respect to figure 3.8 for the case of Z boson. The agreement is also very good between data and MC predictions. Finally, figure 3.12 shows the cross section measured as a function of $H_{\rm T}$ for Z and W events with at least 3 jets and the cross section measured as a function of the fourth leading jet $p_{\rm T}$, obtained by the CMS experiment. Apart from the SHERPA $2\beta 2$ predictions, the agreement in the Z boson channel are excellent. This is not however the case for the W channel where the MC predictions show a harder $H_{\rm T}$ spectrum but predict a very good fourth jet transverse momentum distribution.

Most of these measurements have been performed for the case of the Z boson in the analysis presented in this work, with both higher statistics (19.7 fb^{-1}) and higher centreof-mass energy (8 TeV) allowing for higher precision at phase space region covered by the results presented in this chapter and also allowing for measurement at phase space region inaccessible by previous measurements. The results presented in chapter 6, can also be seen as a first step for more sophisticated studies as jet and Z correlation and correlation between jets which are important for BSM searches.



Figure 3.8: The inclusive multiplicity cross sections of Z + n jets(s) as measured by the CMS experiment (top left), the ATLAS experiment (top right) at LHC with $\sqrt{s} = 7$ TeV, and by the CDF experiment (bottom) at Tevatron with $\sqrt{s} = 1.96$ TeV.



Figure 3.9: The leading jet p_T differential cross section for Z + n jet(s) as measured by the CMS experiment (top left), the ATLAS experiment (normalised to the total cross section (top right) at LHC with $\sqrt{s} = 7$ TeV, and by the CDF experiment (bottom) at Tevatron with $\sqrt{s} = 1.96$ TeV.



Figure 3.10: The third leading jet p_T cross section for Z + 3 jets as measured by the CMS experiment (top left), the ATLAS experiment (top right) at LHC with $\sqrt{s} = 7$ TeV, and by the CDF experiment (bottom) at Tevatron with $\sqrt{s} = 1.96$ TeV.



Figure 3.11: The inclusive jet multiplicity cross section in W + jets as measured by the CMS experiment (top left), the ATLAS experiment (bottom) at LHC with $\sqrt{s} = 7 \text{ TeV}$, and by the D0 experiment (top right) at Tevatron with $\sqrt{s} = 1.96 \text{ TeV}$.



Figure 3.12: The jets H_T differential cross section for $Z + \geq 3$ jets (top left) and $W + \geq 3$ jets (top right), and fourth jet pseudorapidity differential cross section for $Z + \geq 4$ jets (bottom left) and $W + \geq 4$ jets (bottom right), as measured by the CMS experiment at LHC with $\sqrt{s} = 7$ TeV.

Chapter 4

Experimental Set-up

This chapter is devoted to the description of the experimental set-up used in the present work. The first part introduces the Large Hadron Collider. The geometry and the physical principles of this enormous accelerator are explained. A fully detailed description can be found in volume 1 of [53] from which the following description is based. The LHC cannot work by itself and a whole set of preaccelerators is necessary to fill it with protons having already a significant energy. These preaccelerators form the injector complex and are quickly reviewed. The second part of this chapter describes the Compact Muon Solenoid detector used to detect and measure the result of billions of pp collisions furnished by the LHC. A detailed description of the CMS detector can be found in volume 2 of [53] on which the present description is based. This detector is the master piece of the experiment called by the same name and allows the CMS Collaboration to track down and analyse the events issued from the pp collisions.

4.1 The Large Hadron Collider

4.1.1 History and Goals of the Collider

The LHC is a particle accelerator which started its operation in September 2008. As can been seen in figure 4.1 the LHC is located around the European Organisation for Nuclear Research (CERN) complex in the Geneva area between the Jura mountains and the Lake Geneva, straddling the Franco-Swiss border. The LHC machine was built by CERN inside a 27 km-long and 3.8 m-wide circular tunnel a hundred meters underground in average. The tunnel itself was built in the late 1980s for the Large Electron-Positron (LEP) collider which was stopped in 2000 to leave the place to the LHC machine.



Figure 4.1: This figure shows the locations of the four main experiments (ALICE, ATLAS, CMS and LHCb) that take place at the LHC in the Geneva area straddling the Franco-Swiss border (red line). Located between 50 m and 150 m underground, huge caverns of the size of a cathedral have been excavated to house the giant detectors. The SPS, the final link in the preacceleration chain, and its connection tunnels to the LHC are also shown.

More precisely the LHC is a pp collider with a nominal beam energy of 7 TeV (centreof-mass energy of 14 TeV) and is also used to accelerate heavy ions (lead) to an energy of 1.38 TeV. At the time of writing, the LHC holds world record for the highest-energy man-made particle collisions.

The goals of the LHC are multiple but it is first of all a discovery machine and makes it possible for the first time to explore phase space regions never reached before. The study of the Higgs boson discovered in 2012 at the LHC is another main subject of the LHC project. Thanks to the LHC and the expected large amount of data, precision SM measurements are also made possible in an unreached phase space or where the previous experiments results lack in statistical precision. Finally, with the highest energy ever obtained in laboratory, heavy ion collisions make possible to study the quark-gluon plasma, a phase of QCD that is believed to be the state in which our universe was at its first few milliseconds after the Big Bang.



Figure 4.2: (Left) Diagram showing the cross-section of an LHC dipole magnet with cold mass and vacuum chamber. (Right) Cross-section of an LHC dipole magnet indicating the magnetic flux. The two circles (in blue) in the centre of the diagram indicate the beam pipes for beams one and two. Notice how the arrows point in opposite directions allowing to control two counter-rotating beams of protons in the same beam pipe.

4.1.2 Technical Description

Two proton beams are accelerated in opposite directions at the LHC. The collider tunnel therefore contains two adjacent and parallel beam pipes which cross at four points around the tunnel where the detectors of the ALICE, ATLAS, CMS and LHCb experiments are located (see figure 4.1). In order to suppress unwanted proton collisions with the gas molecules contained in the beam pipes, a ultrahigh vacuum is made in the whole beam pipe system, reaching a pressure as low as 10^{-13} atmospheres.

The LHC is not exactly circular but instead is made of eight arcs and eight linear sections. The arcs contain the bending dipole magnets necessary to bend the protons trajectories while the linear sections contain both the quadrupole focussing magnets and the accelerating cavities which are responsible to bring the protons to their nominal energy first and then to maintain their speed by compensating for their energy losses.

The dipole magnets are the biggest and more challenging items of the LHC. In order to give the protons their curved trajectories a total of 1232 dipole magnets (see figure 4.2) producing a field of around 8.4 Tesla at a current of about 11700 A are used. The additional set of 392 quadrupole magnets are dedicated to the focalisation of the beams necessary to ensure a good beam profile and the collisions at high luminosity when the beams intersect at the four experiment locations.

Both the dipoles and the quadrupoles are superconducting magnets and necessitate

a sophisticated cooling system to be maintained at their operating temperature of 1.9 K (-271.25 °C). To achieve this temperature superfluid helium 4 is used for the cryogenic system.

4.1.3 The LHC beams

Each proton beam will nominally contain a total of 2808 proton bunches which in turn are made of 115 billions of protons. The size of a bunch of protons is about 30 cm-long with a transverse size of the order of a mm in the LHC pipes but as small as $16 \,\mu\text{m}$ at the collision points. With these numbers, and given the nominal energy of 7 TeV of the protons, one can compute the total beam energy:

enough to melt half a tonne of copper.

The nominal frequency at which bunches cross is 25 ns and at the speed of the protons, this represents about 7.5 m of distance between two bunches. There are however gaps in the bunch structure of the LHC beams among which the longest is a $3 \mu s$ (900 m) time needed to give the beam dump kickers the time to reach to full required voltage. There are also other smaller gaps in the beam arising from similar needs from the SPS (the last preaccelerator of the LHC injector complex, see next sections) and LHC injection kickers.

4.1.4 Instantaneous and Integrated Luminosities

The instantaneous and integrated luminosities, \mathcal{L}_{inst} and \mathcal{L}_{int} respectively, are quantities which characterise the performance of a particle accelerator, the integrated luminosity being simply the integral of the instantaneous luminosity over a certain period of time:

$$\mathcal{L}_{\rm int} = \int_{t_0}^{t_f} \mathcal{L}_{\rm inst}(t) \ dt$$

The instantaneous luminosity is formally given by the ratio of the number of events produced (N) at a certain time (t) during a small time interval dt, to the interaction cross section (σ) :

$$\mathcal{L}_{\text{inst}}(t) = \frac{1}{\sigma} \frac{dN}{dt}.$$

This shows that the dimension of \mathcal{L}_{inst} are events per time per area: $cm^{-2}s^{-1}$.

4.1. THE LARGE HADRON COLLIDER

In practice the luminosity is highly dependent on the particle beam parameters. Indeed, increasing the beams crossing frequency f, the number of protons n_p contained in each protons bunch, the numbers of bunches N_1 and N_2 in each beam and focussing the beams by decreasing their transverse size σ_x , σ_y , will result in a higher rate of interaction:

$$\mathcal{L}_{\text{inst}} = \frac{N_1 N_2 n_p f}{4\pi \sigma_x \sigma_y}$$

The product $\sigma_x \sigma_y$ is often expressed in terms of the Lorentz boost γ , the beam emittance ε (measuring the average spread of the beam particles coordinates in position-momentum phase space) and the β^* beam parameter (measuring the distance between the point where the beams cross and the point where the transverse size of the beams is twice as large):

$$\mathcal{L}_{\text{inst}} = \frac{N_1 N_2 n_p f}{4\pi \sigma_x \sigma_y} = \frac{N_1 N_2 n_p f \gamma}{4\pi \varepsilon \beta^*}.$$

During an LHC run the proton beams get degraded due to their interaction with the other beam and with the gas remanent in the vacuum pipe. This directly results in a decrease of the instantaneous luminosity. An estimation of its evolution with respect to time is given by:

$$\mathcal{L}_{\text{inst}}(t) = \frac{\mathcal{L}_{\text{inst}}(t_0)}{1 + (t_0 - t)/\tau_{\mathcal{L}}}$$

where $\tau_{\mathcal{L}}$ is the luminosity decay time and depends on various factors such as the number of high luminosity experiments, the initial beam intensity, the total cross section of interaction of the accelerated particles, ... At the LHC, a typical estimate of the luminosity lifetime is given by:

$$\tau_{\mathcal{L}} = 14.9 \text{ h.}$$

In general the instantaneous luminosity is made as large as possible (the maximum value being limited by the accelerator machine) in order to maximise the number of produced events. However, sometimes the instantaneous luminosity is lowered on purpose in order to have cleaner events (LHC runs for which the luminosity was kept relatively low are called low pile-up runs). Figure 4.3 shows the evolution of the instantaneous luminosity during the year 2012, as measured in the four experiments, as well as the corresponding integrated luminosities. Table 4.1 summarises the beam and machine parameters for the years 2011 and 2012 together with their nominal values.

4.1.5 The Injector Complex

The LHC is only the final accelerator for the protons. Before injection of the protons into the LHC accelerator, it is necessary to use an entire set of smaller accelerators of different



Figure 4.3: Monthly evolution of the instantaneous luminosity (left) during the year 2012, as measured in the four experiments (ATLAS in black, CMS in green, LHCb in blue and ALICE in red), as well as the corresponding integrated luminosities (right).

Table 4.1: Proton beams and LHC machine main parameters. Nominal values compared to actual values of 2011 and 2012 LHC collisions.

Parametres	2011	2012	nominal
Beam energy (TeV) $$	3.5	4	7
Dipole magnetic field (Tesla)	4.16	4.5	8.33
Number of bunches per beam	1303	1380	2808
Crossing frequency (ns)	50	50	25
Number of protons per bunch	$1.6 \ 10^{11}$	$1.6 \ 10^{11}$	$1.5 \ 10^{11}$
Emittance (µmrad)	3.5	2	3.75
β^* (m)	1.0	0.6	0.55
Max. instantaneous luminosity $(cm^{-2} s^{-1})$	$3.6 \ 10^{33}$	$0.55 \ 10^{34}$	$1.0 \ 10^{34}$

kinds (see figure 4.4 page 71). These preaccelerators will increase, one after the other, the energy of the proton beams. However, the preaccelerators first need to be filled with protons from a source. The very beginning of the proton beams takes place during the protons extraction from the simplest element, the hydrogen.



Figure 4.4: CERN's accelerator complex. The chain used for the LHC experiments is composed of the LINAC 2 (bottom centre) for proton beams and LINAC3 for lead beams, followed by the BOOSTER which then provides the PS with protons, in turn sending the proton bunches to the SPS which finally feeds the LHC.

The Proton Source

Running at almost the speed of light inside the enormous LHC machine, the protons all come from the same surprisingly small bottle of hydrogen (see figure 4.5 page 72).

To get bare protons, the hydrogen gas contained in the bottle is injected into a relatively small cylindrical machine called duoplasmatron (see figure 4.5 page 72). Inside this device a cathode filament is heated and emits electrons that interact with the hydrogen gas which becomes ionised. This plasma is subsequently subjected to different electric fields extracting the positively charged ions. This is the birth of a proton beam. At this stage the protons forming the beam have an energy lying around 100 keV and are pulsed every 100 µs. The beam is then injected into the next accelerator device.



Figure 4.5: Photograph of the hydrogen bottle (in red) from which LHC protons are extracted by the duoplasmatron device (similar to the one located below the hydrogen bottle on the picture).

The Linear Accelerator

The next stage towards the LHC is a straight way through the 30 m-long LINAC 2 (see figure 4.6 page 73). As a linear accelerator, this machine uses radio frequency cavities to charge cylindrical conductors of increasing length that will in turn produce an oscillating electric field between them. As the protons pass through the consecutive empty regions between the cylinders, they get pushed by the conductor behind them and pulled by the one situated just in front of them. The result is a kick of a certain amount of energy. In addition to the radio frequency cavities, the LINAC 2 also contains quadrupoles magnets focussing the proton beam being created. When the protons reach the end of the LINAC 2, they have acquired an energy of 50 MeV and are ready to enter the next accelerating system.

The Booster

The first and smallest circular accelerator of the injector complex is a system made up of four superimposed synchrotron rings with a radius of 25 m, called the Proton Synchrotron Booster and is shown in figure 4.6 page 73. It receives the 50 MeV proton bunches from the LINAC 2 and accelerates them to an energy of 1.4 GeV. Being a synchrotron, the booster is a circular (more precisely a sequence of straight and curved sections) device



Figure 4.6: Photographs of a the LINAC 2 linear accelerator (left) and the proton synchrotron Booster (right).



Figure 4.7: Photographs of a the Proton Synchroton, PS, accelerator (left) and the Super Proton Synchrotron, SPS, accelerator (right).

using a beam-synchronized magnetic field in order to bend the trajectory of the protons inside the pipes of the curved sections and an electric field to accelerate the protons every cycle in the straight sections. The alternative electric field also gives a structure in bunches to the beam. Once the energy of 1.4 GeV is reached and enough bunches are accumulated, the beam is guided to the second circular preaccelerator.

The Proton Synchrotron

The next accelerator is one of the oldest of CERN, the Proton Synchrotron (PS) shown in figure 4.7 page 73. It was build in the 50's with a radius of 100 m and has been upgraded several times significantly improving its performance since its first use in 1959. It uses the same mechanism as the booster to accelerate the protons from an energy of 1.4 GeV

to 25 GeV. This is achieved by using some 277 electromagnets, including 100 dipoles to give the beam its circular trajectory. The beam is then injected to the last preaccelerator.

The Super Proton Synchrotron

Filled with a 25 GeV proton beam, the 6.9 km-long synchrotron called Super Proton Synchrotron (SPS) (see figure 4.7 page 73) takes over the preacceleration process of the protons. The SPS was the main accelerator at the time of the discovery of the W and Z particles where it provided proton-antiproton collisions. It is now used as the last preaccelerator and provides the LHC machine with proton beams of 450 GeV energy. The protons are injected into the LHC via two lines (TI2 and TI8) in order to build up the two clockwise and the anticlockwise beams that the LHC machine will bring to the nominal energy. A total of 1317 electromagnets are present to give the protons their increasing energy, and 744 are used to bend their trajectories along the SPS vacuum pipe.

4.2 The CMS Detector

CMS is a general purpose detector built with different goals in mind among which high precision measurements of different observables of the SM with, in particular, the discovery of its missing piece: the Higgs boson, the exploration of processes taking place at the TeV scale and search for new physics processes beyond the SM.

CMS is also used to analyse heavy ions (lead-lead) collisions and proton-lead collisions taking place in some periods of the LHC runs. This subject, not treated in this thesis, has as main goal to study the quark-gluon plasma: a phase of QCD which is believed to exist at very high densities and temperatures, conditions that were present at the very beginning of our universe.

Below, the CMS detector is described together with its main subdetectors from which the tracker, the calorimeters and the muon detectors are the core of the detection system working in the intense magnetic field present inside the whole detector. A section devoted to the trigger system is present after the description of the CMS detector. Finally, a description concerning the CMS data treatment is given.

4.2.1 The Geometry of CMS and its Principles

We begin this chapter by a general view of the CMS detector. The main subdetectors are described in some more details in the next sections. The data analysis presented in the present work being the study of Z events with associated jets in both the electron and muon decay channel, most parts of the detector are of great interest for it. A particular attention is brought to the muon detection system whose upgrade is related to the second part of this work dedicated to the triple-GEM detectors that will be installed during the long shut down 2.

The global shape of the CMS detector [54] (shown in figure 4.8) is a 21.6 m-long cylinder with a diameter of 15 m and weighting approximately 12 500 Tonnes. The building element that characterises the experiment is a solenoidal superconducting magnet which produces an internal constant and uniform magnetic field of 3.8 Tesla pointing along the direction of the beams. More precisely, the CMS detector has the shape of a dodecagonal-base prism. Its central part, called barrel, is divided into five parts, called wheels, numbered from -2 to +2. The barrel contains several layers of cylindrical detectors coaxial with respect to the direction of the beams. To close the detector at its ends and therefore ensure its (almost) full solid-angle coverage and hermeticity, disks of detectors forming the so called endcaps, are placed at the two extremities. From the inner region to the outer one, the various components of CMS are:

- Silicon tracker (see section 4.2.3) placed in the region r < 1.2 m and $|\eta| < 2.5$. It consists of a silicon pixel detector supplemented by silicon microstrip detectors. It is used to reconstruct charged particle tracks and interaction vertices.
- Electromagnetic calorimeter (ECAL) (see section 4.2.4) located in the region 1.2 < r < 1.8 m and $|\eta| < 3$. The ECAL is composed of thousands of lead tungstate (PbWO₄) scintillating crystals. It is used to measure the energy deposit released by photons and electrons.
- Hadronic calorimeter (HCAL) (see section 4.2.4) placed in the region 1.8 < r < 2.9 m and $|\eta| < 5$. The HCAL made of brass plates alternating with plastic scintillators is used to measure the energy released by the hadrons produced in the interaction, as well as their location with respect to the point of interaction.
- Superconducting solenoid magnet (see section 4.2.2) placed in the region 2.9 < r < 3.8 m and $|\eta| < 1.5$. The magnet generates an internal uniform magnetic field of 3.8 Tesla oriented along the direction of the beams needed to deflect the charged particles in order to measure their transverse momenta through the measured curvature. The flow of the magnetic field is closed by an iron yoke of about 14 m



system whose centre is located at the geometrical centre of the CMS detector. ring, the y-axis pointing upwards and the z-axis aligned with the beam direction in such a way to form a right-handed coordinate Figure 4.8: A perspective view of the CMS detector and its coordinate system with the x-axis point towards the centre of the LHC in diameter and 21.6 m in length in which an average residual magnetic field of 1.8 Tesla pointing in the opposite direction is still present.

• Muon chambers located in the region 4 < r < 7.4 m and $|\eta| < 2.4$. Composed of three distinct types of detectors (drift chambers located in the barrel, cathode strips chambers in the endcaps, and resistive plate chambers in both the barrel and the endcaps), the muon chambers are used to reconstruct the tracks of muons passing through the whole detector.

In figure 4.9 page 78 is shown how the various subdetectors contribute to the identification of the particles that pass through them. For the present analysis, many particles are of importance. The muons (light blue solid line) and the electrons (red solid line) are used to reconstruct the Z boson. Jets, constituted of charged hadrons (green solid line) and neutral hadrons (green dashed line), are accompanying the Z boson in the final state due to higher order QCD interactions. Photons (dark blue dashed line) are also to be considered for the analysis as they may be radiated from final state electrons (and to a much less extend by the muons). Therefore, as shown in the picture, almost all components of CMS are used for a correct identification and measure of these objects. The following sections provide more detailed descriptions of the different subdetectors starting by the magnet.

4.2.2 The Superconducting Solenoid Magnet

The key element of the CMS detector, around which the subdetectors are built, is the superconducting solenoid magnet depicted in figure 4.10. It is 13 m-long and has a diameter of 6 m. It provides a uniform constant magnet field of 3.8 Tesla in its interior, pointing along the beam direction, whose function is to bend the trajectory of electrically charged particles produced from the pp collisions. The CMS magnet is made of coils of niobium-titanium wires, a superconducting material needed to allow electricity flow without resistance (0.1 m Ω) thus allowing high intensity current of up to 19500 A (at designed value of 4 Tesla), giving a total stored energy of 2.66 GJ. The operating current for 3.8 Tesla is 18 160 A, giving a stored energy of 2.3 GJ.

The high field intensity is necessary to bend high- $p_{\rm T}$ particles, since the highest the particle's transverse momentum is, the less curved its trajectory is. The relation between the transverse momentum of a charged particle and its radius of curvature is found to be:

$$p_{\rm T} = 0.3 \, B \, R \tag{4.1}$$

where $p_{\rm T}$ is expressed in GeV, B is the magnetic field intensity in Tesla and R the radius of curvature in m.



the muon chambers together with the return yoke. The magnetic field is indicated and responsible for the curvature of the charged particles. From left to right can be seen the tracker layers, the ECAL crystals, the HCAL, the Superconducting solenoid magnet and Figure 4.9: Schematic view of a slice of the CMS detector with expected behaviour of the subdetectors to the passage of different detector.from charged and neutral hadrons can be seen in the HCAL. The muon only leaves ionisation along its trajectory and escape the particles (solid lines). In the ECAL is present a simulation of electromagnetic deposits of electrons and photons. Hadronic deposit



Figure 4.10: Schematic view of the CMS Superconducting Solenoid Magnet.

Even with such an intense magnetic field, the trajectory of particles with $p_{\rm T}$ above 50 GeV is almost straight ($R = 43 \,\mathrm{m}$) at the scale of the CMS detector. On the other side, particles with less than about 0.75 GeV of $p_{\rm T}$ are so much bent that they never reach the electromagnetic calorimeter whose inner radius is of about 1.3 m, and circle around following a helicoidal trajectory.

From the above relation it is clear that the CMS magnet allows to the tracker to obtain a resolution which is inversely proportional to the transverse momenta of particles. The energy measurement therefore needs to be complemented with other detectors such as the calorimeters for high- $p_{\rm T}$ charged particles, and for neutral particles.

4.2.3 The Tracker

The CMS tracker constitutes the first layer (the closest to the beam pipe) of the detector. Its goal is to provide a precise measurement of charged particle trajectories and interaction vertices. The CMS collaboration has opted for a all-silicon-based tracker using two different technologies. At the most inner region, the particle rate is the highest (10^8 particles per cm² per s at a radius of 4 cm from the interaction point) and requires a high granularity and fast response detector, both requirements achieved by pixel detectors.



Figure 4.11: Schematic cross-section through the CMS tracker. Each black line segment represents a detector module. Double lines indicate back-to-back modules which deliver stereo hits.

Further away from the interaction point, the rate is lower $(6 \times 10^6 \text{ and } 3 \times 10^5 \text{ particles})$ per cm² per s at a radius of 22 and 115 cm from the interaction point, respectively) and micro strips detectors can satisfy the requirements.

The tracker information is also combined with electromagnetic calorimeter and muon system responses for electron and muon identification, respectively, used in the High Level Trigger (HLT) (see section 4.2.6 of CMS and for the complete event reconstruction).

The whole tracker system has a length of 5.8 m and a diameter of 2.5 m. It consists of a central barrel and two endcaps to close the cylinder. The structure is optimised to have in average 12-14 hits per track, ensuring both a high efficiency and a low rate of misidentified tracks.

The tracker geometry is shown in figure 4.11 and its material budget in units of radiation length χ_0 (the radiation length is both the mean distance over which a highenergy electron loses all but 1/e of its energy by bremsstrahlung, and 7/9 of the mean free path for pair production by a high-energy photon, and is therefore an appropriate scale length for describing high-energy electromagnetic cascades) in figure 4.12. It is important to be aware of this in order to take into account photon conversion as well as induced radiation of charged particles passing through the material. The material budget increases from 0.4 χ_0 at η close to 0, to about 1.8 χ_0 at $|\eta|$ around 1.4, beyond which it decreases to about 1 χ_0 at $|\eta|$ around 2.5.



Figure 4.12: Material budget of the CMS tracker in units of radiation length χ_0 (left) and interaction length λ_I (right) as a function of the pseudorapidity η for the latest version of the tracker geometry in the 2012 data taking [55].

The Silicon Pixel Detector

The silicon pixel detector, shown in figure 4.13, is used in CMS as the starting point for the reconstruction of the tracks and is essential for the reconstruction of the primary and secondary vertices.

The silicon pixel detector is placed in the region closest to the collision point just around the beam, where the flow of particles is maximum. The pixel tracker covers the region $|\eta| < 2.5$. It consists of a central part (barrel) composed of three concentric cylindrical 53 cm-long layers (BPix) placed at a distance r from the centre of 4.4 cm, 7.3 cm and 10.2 cm, and two endcaps parts (FPix) placed on each side at $z = \pm 34.5$ cm and $z = \pm 46.5$ cm and extending in the radial direction from an inner radius r_{in} of 6 cm to an outer radius r_{out} of 15 cm, in such a way that each particle emitted from the nominal interaction point will pass through three (two) layers of the whole pixel detector for $|\eta| < 2.1(2.5)$.

The barrel part of the pixel detector contains 768 modules for a total of 48 million pixels covering a total area of 0.78 m^2 . For the forwards regions, each disk of the endcaps is divided into 24 segments, each of which composed of 7 modules summing up to 672 modules for the 4 disks. The endcaps totalise 18 million pixels for a detection surface of 0.28 m^2 . The pixel sensors have a size of $150 \times 100 \text{ µm}^2$ and a thickness of 260 to 300 µm. The spatial resolution achieved in the reconstruction of hits is found to be in the range of 15 to 20 µm.



Figure 4.13: Sketch of the CMS pixel detector [56].

The Silicon Strips

The pixel detector is complemented by the silicon strip tracker (SST) which surrounds it from an inner radius $r_{in} = 20 \text{ cm}$ to an outer radius $r_{out} = 116 \text{ cm}$. In this region the flow of particles is low enough to allow the use of silicon strip detectors. As shown in figure 4.11, the SST is much larger than the Pixel detector. Its central part is divided into the Tracker Inner Barrel (TIB) itself composed of 4 layers extending in radius to 55 cm and Disks (TID) made of 3 disks at each end. The TIB/TID complex provides up to 4 $r - \phi$ measurements and is surrounded by the Tracker Outer Barrel (TOB) consisting of 6 layers with maximal radius extension of 116 cm and providing 6 additional $r - \phi$ measurements. The longitudinal coverage of the TOB reaches $z = \pm 118 \text{ cm}$, beyond which the Tracker Endcaps (TEC \pm) are found. Each TEC is a collection of 9 disks covering the region 124 < |z| < 282 cm and $0.9 < |\eta| < 2.5$.

As it can be seen from figure 4.11, the first two layers of TIB and TOB, the first two rings of TID as well as some rings of TEC are made of multiple back-to-back modules. They are placed with a 100 mrad crossing angle in order to measure the second coordinate (z in the barrel and r on the disks)

Figure 4.14 shows one of the disks of TEC together with a more detailed view of one of its petal.

4.2.4 The Calorimeters

Calorimeters are used in order to measure the energy of the produced particles. While neutrinos pass through the whole detector without interacting and muons just leave ionisation deposits along their trajectory inside the material of CMS, all other produced particles entering the calorimeter initiate a particle shower whose finale stage particles



Figure 4.14: Photograph of a TEC disk (left panel) and zoom on a TEC petal (right panel) [57]. Only detector units on the side facing the interaction point can be seen (rings 1, 3, 5 and 7). Detector units on rings 2, 4 and 6 are located on the back side of the petal.

(electrons and photons, ...) are counted and lead to an estimate of the energy of the initial particle. Unlike the tracker system which can track only charged particles, the calorimeter collects and measures the energy deposit of both charged and neutral particles.

Two kinds of calorimeters can be distinguished, the homogeneous calorimeter and the sampling calorimeter. The former has its entire volume sensitive while the second has material dedicated for the production of the shower that is distinct from the material layers actually measuring the energy deposits.

Additionally, the calorimetry system is often divided into an electromagnetic calorimeter specifically designed to measure the energy of particles interacting primarily via electromagnetic interactions, and a hadronic calorimeter part designed to collect the energy initiated by hadrons. Since the radiation lengths for electromagnetic interactions are significantly shorter (0.89 cm for the CMS ECAL crystals) than the interaction lengths of strong interacting particles (16.2 cm for the CMS HCAL material), the electromagnetic part is always placed first with respect to the hadronic calorimeter.

As it can be seen in figure 4.8 page 76, CMS has both an electromagnetic and a hadronic calorimeter that are described in more detail in the next sections. Figure 4.9 page 78 shows how the different particles interact with the two calorimeter parts, electromagnetic and hadronic. As announced, the electromagnetic calorimeter is placed closer to the interaction point than the hadronic one and is also of smaller size since the longitudinal expansion of the electromagnetic showers are smaller than the hadronic ones.



Figure 4.15: Layout of the CMS electromagnetic calorimeter showing the arrangement of crystal modules, supermodules and endcaps, with preshower in front.

The Electromagnetic Calorimeter

The main function of the CMS electromagnetic calorimeter (ECAL) is to identify electrons and photons and to measure accurately their energies. The ECAL [58, 59] is a homogeneous calorimeter with cylindrical geometry. It is composed of an ECAL Barrel (EB) consisting of 61 200 scintillating lead tungstate (PbWO₄) crystals and two ECAL Endcaps (EE) containing 7324 PbWO₄ crystals each. A layout of the CMS ECAL is schematically presented in figure 4.15.

The EB has an inner radius of 129 cm, a longitudinal length of 630 cm and extends in the region $|\eta| < 1.479$. It is divided in 36 supermodules whose lengths correspond to half of the barrel. Each of these supermodules contains a matrix of 20 × 85 crystals in the plane (η, ϕ) in turn divided into four modules along the η direction which are themselves divided into smaller modules formed by arrays of 5 × 2 crystals. The PbWO₄ crystals (see left picture on figure 4.16) have the shape of a truncated pyramid with a length of 23 cm corresponding to about 24.7 χ_0 , a front area of 22×22 mm² and rear surface of 26×26 mm². The (η, ϕ) coverage of a single crystal corresponds to $\Delta \eta \times \Delta \phi = 0.0175 \times 0.0175$ (*i.e.* approximately 1°). The crystals are grouped into 5 × 5 matrices called trigger towers that provide raw energy measurement information used for the trigger. The axes of the crystals in the calorimeter are inclined by 3° in both η and ϕ directions with respect to the nominal point of interaction in order to minimise spacing and to prevent particles



Figure 4.16: $PbWO_4$ crystals with photodetectors attached. A barrel crystal with the upper face depolished and the APD capsule containing the two APDs (left panel). An endcap crystal with its VPT (right panel).

travelling just between two crystals.

As far as the endcaps are concerned they cover the region $1.479 < |\eta| < 3$ and each of them is formed by two half-disks called Dees. Each of these dees contains 3662 crystals similar to those of the barrel but with a front surface of $28.6 \times 28.6 \text{ mm}^2$ and rear area of $30 \times 30 \text{ mm}^2$ (see right picture of figure 4.16). The crystals are organised in 18.5×5 matrix units. The axis of the crystals are oriented in such a way that they cross the axis of the beams at a distance of 130 cm beyond the nominal point of interaction in order to avoid the situation in which a particle emitted from the interaction passes mainly between two crystals.

The choice of PbWO₄ for the ECAL crystals as scintillating material has been made for the following reasons. First, the high density ($\rho = 8.3 \,\mathrm{g/cm^3}$), short radiation length $(\chi_0 = 0.89 \,\mathrm{cm})$ and the small Molière radius $(\mathrm{RM} = 2.2 \,\mathrm{cm})$ (characterising a material by giving the radius of a cylinder that would contain on average 90% of the shower's energy deposition) allow for a compact and fine granularity calorimeter. Furthermore, the decay time of the scintillation of the order of 15 ns allows to gather about 80% of the light emitted within the 25 ns nominal time between two successive beam bunches crossing. Finally, crystals of lead tungstate are characterised by high resistance to radiations and therefore can operate for several years in the high radiation environment of the LHC, suffering modest deterioration in performance. Only a lost of transparency of the crystals due to the radiation is observed and accounted for by special real time laser monitoring. The main disadvantage of these crystals is the poor light collection (around 10 photoelectrons/MeV) which necessitates an amplification of the light signal. This is achieved through the use of avalanche photodiodes (APD) (see left picture of figure 4.16) in the barrel and photo vacuum triode (VPT) (see right picture of figure 4.16) in the endcaps, resistant to radiation and able to operate under the strong magnetic field of CMS.

The energy resolution of the ECAL can be expressed by the sum of three terms

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c. \tag{4.2}$$

In this formula, the first term a corresponds to the stochastic term and includes the contribution of the random fluctuations in the number of photoelectrons emitted. This term dominates the energy resolution at low energies. The fluctuations follow a Poissonian distribution and the term takes into account the emission of light by the crystal, the efficiency in the collection of light and the quantum efficiency of the photodetector. The noise term b represent the contribution of the electronic noise. The contribution varies depending on the brightness of operation of the LHC. The constant term c, dominant at high energies, takes into account various contributions such as the stability of the operating conditions (in particular temperature and voltage), the presence of inert material in the crystal, the non uniformity of the collection of light along the crystal, inter-calibration errors and damage from radiation.

First estimates of a, b and c are obtained from test measurements and refined with colliding data analysis. A typical resolution was found to be [53]:

$$\frac{\sigma}{E} = \frac{2.8\%}{\sqrt{E}} \oplus \frac{0.12}{E} \oplus 0.30\%. \tag{4.3}$$

where E is given in GeV.

On the inner side of the endcaps two preshower detectors are placed (see figure 4.15 page 84) whose purpose is to facilitate the distinction between a primary γ and a primary π^0 for which the angle between the two emerging photons from its decay is likely to be small. This is achieved thanks to the smaller granularity of the preshower detectors compared to the ECAL granularity. These detectors, which cover the region 1.653 < $|\eta| < 2.6$, are sampling calorimeters formed by two discs of lead (of respective width of 2 χ_0 and 1 χ_0) that initiate the electromagnetic cascade, alternating with two levels of silicon micro strip detectors which also provide the measure of the released energy.

The Hadronic Calorimeter

To complete the ECAL and in order to properly measure hadrons' energies and directions, a hadronic calorimeter (HCAL) [60] is also present in CMS. The HCAL is a hermetic sampling calorimeter covering the entire region within $|\eta| < 5$. As shown in figure 4.17, the HCAL is divided into four subdetectors: hadronic barrel calorimeter (HB), located in the region of the barrel and interior of the magnet; hadronic endcap calorimeter (HE), located in the region of the endcaps, also inside the magnet; hadronic outer calorimeter



Figure 4.17: Longitudinal view of 1/4 of the CMS detector showing the locations of the hadron barrel (HB), end cap (HE), outer (HO) and forward (HF) calorimeters.

(HO), placed along the inner part of the return yoke of the magnetic field, just outside of the magnet; hadronic forward calorimeter (HF), located in the forward region, outside of the endcaps.

The absorbing material of the HCAL is made of brass and its active material is provided by tiles of plastic scintillator. HE and HB are therefore constructed with sheets of absorber interspersed with sheets of scintillating materials in which fluorescent optical fibres (wavelength shifters, WLS) are immersed carrying the light to the photodetectors HPD (Hybrid Photodiodes).

The HCAL detector components are described below:

HB is 9 m-long and extends in a region of radius 178 < r < 288 cm and pseudorapidity |η| < 1.4 surrounding the ECAL. It is composed of two cylinders containing 18 sections each covering an angle Δφ = 20°. Each half-cylinder is divided into 16 sectors along η. The HCAL contains a total of 2304 calorimeter towers of granularity Δη × Δφ = 0.087 × 0.087. Each tower is composed of 15 brass plates arranged parallel to the beams with a thickness of 50 mm and 2 stainless steel at the extremities to ensure the solidity of the structure. Alternatively placed between the brass plates, 17 sheets of 3.7 mm-thick (except the most inner one which is 9 mm-thick) plastic scintillators contain the optical fibres. All the fibres that are relative to the same tower are sent to the same photodetector which provides the integration of the signal.

- HE extends in the region of pseudorapidity $1.3 < |\eta| < 3.0$, partially overlapping the HB. The geometry of HE is similar to that of HB. It is divided into 18 sectors each of which covers an angle $\Delta \phi = 20^{\circ}$. Each endcap is divided radially into 14 rings each divided into 72 sections. Each section covers an angle of 5° in ϕ . For the 8 most inner sections the segmentation in ϕ is 10° to allow the passage of the WLS fibres. The segmentation in η varies from 0.87 for the outer most towers and 0.35 for the towers closest to the beam. The 2304 towers of HE are composed of 19 layers of 3.7 mm-thick plastic scintillators alternating with layers of 78 mm-thick brass absorber.
- HO improves the measurement of the energy of the hadron cascades of higher energy that can overcome the area of HB (whose dimensions are limited by those of the solenoid). HO is placed outside of the solenoid along the inner part of the return yoke of the magnetic flux. It is therefore mounted on the 5 wheels that make the yoke and is divided into 12 sectors each of size 2.5 m in the z direction covering an angle of 30° in ϕ . HO consists of several layers of plastic scintillators increasing the effective thickness of the calorimeter to more than 10 interaction lengths λ_0 : two layers in the central wheel, separated by a layer of 18 cm-thick iron absorber, placed at the radial distance of 3.850 m and 4.097 m. The scintillators have a thickness of 10 mm and have a granularity $\Delta \eta \times \Delta \phi = 0.087 \times 0087$ similar to those of HB so that there is a 1-1 correspondence between a tower of HB and a segment of HO. The light from the scintillators is collected via WLS fibres and transported to the photodetectors located on the return yoke.
- HF ensures a coverage of the hadron calorimeter up to $|\eta| = 5$. It is installed externally to the CMS detector at a distance of 11.2 m from the nominal point of interaction. This calorimeter improves the identification of processes that produce forward jets. The two cylindrical units of HF are 1.65 m-long and have an active radius of 1.4 m. HF is a sampling calorimeter with quartz fibres covered with plastic inserted in sorbing material made of iron. The choice of these materials was dictated by the large dose of radiation present in the forward area which does not allow the use of scintillators. Each unit is composed of 18 sections, each of which covers an angle of 20° in ϕ and contains 24 calorimetric towers in which alternate 1.65 m-long fibres and 1.43 m-long fibres. The quartz fibres emit Čerenkov light to the passage of charged particles. This light is subsequently detected by radioresistant photomultipliers. There are 13 towers in η with segmentation variable from 0.1 to 0.3 depending on the distance from the beam while a coverage in ϕ varying between 10° and 20°. Overall HF is composed of 900 towers sampled via 1800 electronic channels.

4.2. THE CMS DETECTOR

The resolution of the hadron calorimeter was measured by the CMS Collaboration [61,62] and can be parameterised by the following expressions which include a stochastic contribution and a constant term:

barrel + endcap :
$$\frac{\sigma}{E} = \frac{120\%}{\sqrt{E}} \oplus 9.5\%$$
 (4.4)

forward :
$$\frac{\sigma}{E} = \frac{280\%}{\sqrt{E}} \oplus 11\%$$
 (4.5)

where E is given in GeV.

4.2.5 The Muon Detectors

Because muons are the only particles (with the neutrinos) with the ability to cross the whole detector without being stopped by the calorimeters, the muon detection system is for this reason located at the outermost region of the CMS detector and any signal collected from the muon chambers is likely to come from the passage of a muon.

The entire muon detection system is made of a total of 1400 muon chambers. Three different gaseous detection technologies are used for these chambers: 250 of the chambers are composed of drift tubes (DT), 540 are cathode strip chambers (CSC) and the remaining 610 are resistive plate chambers (RPC). The following subsections describe in more detail each of these three types.

The Muon Drift Tubes

The muon drift tubes chambers are located only in the barrel of the CMS detector. The forward region of the detector is not ideal for the drift tube technology because of its intense magnetic field and the high neutron-induced background present at this region.

Each of the five CMS barrel wheels contains 4 cylindrical stations around the beam line. The 3 inner cylinders contain 60 drift chambers each and the outer cylinder has 70 of them. The geometrical layout of the drift tube chambers is shown in figure 4.18. The drift tube chambers are composed of 3 superlayers (except for the outermost stations whose chambers contain 2 superlayers), each of them being divide into 4 layers of rectangular drift cells shifted by half a cell. Inside a chamber, the two outer superlayers contain wires placed parallel to the beam line providing a measurement in the $r - \phi$ plane. When present, the inner superlayer provides a z measurement of the muon track thanks to its wires placed orthogonally to the beam direction.



Figure 4.18: Layout of the CMS barrel muon DT chambers in one of the 5 wheels. The chambers in each wheel are almost identical. Note that in sectors 4 (top) and 10 (bottom) the MB4 chambers are cut in half to simplify the mechanical assembly and the global chamber layout.



Figure 4.19: Section of a CMS muon drift tube cell showing drift lines and isochrones. The voltages applied are +3600 V for wires, +1800 V for electrode strips, and -1200 V for cathode strips.

Inside each of the cells (represented in figure 4.19) of a superlayer, an electric field is set up around a thin wire that is immersed in a mixture of argon and carbon dioxide gas. When a muon passes through a cell it ionises the gas liberating atomic electrons which drift toward the anode wire under the effect of the electric field¹. These electrons being accelerated by the field, can acquire enough energy to in turn ionise the gas, resulting in an avalanche leading to a current in the wire.

The resolution reached by each station is about $100 \,\mu\text{m}$ in position and $1 \,\text{mrad}$ in direction. The DTs are also relatively fast detectors which make them suitable for triggering events containing a muon candidate.

The Cathode Strip Chambers

Cathode strip chambers are used in the endcap disks of the CMS detector and are well suited for the uneven magnetic field and high particle rates present in that region of the detector. Figure 4.20 provides an overview of the CSC geometrical layout inside the end cap disks.

CSC are made of arrays of negatively-charged copper cathode stripes directed radially from the beam direction and positively-charged anode wires orthogonal to both the stripes

¹The electric field in the drift tubes is created by setting 3 different potential respectively to the anode, the electrode stripes and the cathode strips, in such a way that the drift velocity is constant in the whole cell.



Figure 4.20: Longitudinal view of a quarter of the CMS detector. Cathode strip chambers (CSC) of the endcap Muon system are highlighted.

and the beam direction, the whole system laying in a volume of gas. As illustrated in figure 4.21, each CSC module contains 6 layers of gas gaps with wires and stripes. As in the DT, ionising processes occur in the CSC when muons pass through. However, not only the freed electrons induce a signal, the ions resulting from the ionisation induce an electric signal on the cathode stripes, see figure 4.21. The orthogonality of the stripes and the wires provides two position coordinates for each passing particle.

While the CSC resolution is slightly poorer than the one of DT: about 200 µm in position and 10 mrad in direction, its closely spaced wires render the CSCs fast detectors suitable for triggering.

The Resistive Plate Chambers

While the DT and CSC are fast enough to be used for triggering purposes, their time response are however comparable to the time between two consecutive bunch crossing. For this reason, resistive plate chambers (RPC) supplement the whole muon detector. RPCs are gaseous detectors made of parallel plates that combine adequate spatial resolution with a time resolution comparable to that of scintillators. RPCs are placed on each side of the first and second stations of DTs, and on the inner side of the third and fourth DTs, in the barrel section of the CMS detector, as illustrated in figure 4.22. In the endcap disks, the RPCs cover pseudorapidity values up to $|\eta| = 1.6$.

RPCs have a very fast response (1 ns), but coarser spatial resolution than the DTs or



Figure 4.21: Quarter-view of the CMS detector. Cathode strip chambers (CSC) of the Endcap Muon system are highlighted.

CSCs, making them optimal for triggering on muons.

Figure 4.22 summarises the whole muon system detection layout.

4.2.6 The Trigger System

At the LHC, the nominal bunch crossing rate is of 40 MHz, *i.e.* one bunch crossing every 25 ns. Furthermore at every bunch crossing multiple pp collisions generally occur, depending on the configuration of the LHC machine. During the run 1 at 8 TeV an average of 20 collisions per bunch crossing was observed. However, only a small fraction of these events are of real physics interest and the actual technology can not handle the mass storage at such a high speed and also for such an amount of data estimated to be of ~ Mb per event. A trigger system is thus needed in order to select the most interesting events and reduce the rate to the 100 Hz of writing capability.

In CMS a two-level triggering system has been chosen, as illustrated in figure 4.23. The first level, L1, runs on dedicated processors using coarse level granularity information given by the calorimeters and the muon system. It has to be noted that the tracker information is not currently used at this level of the trigger system. The L1 decision to keep processing an event or instead reject it has to be taken for every bunch crossing within $3.2 \,\mu s$. This timing requirement is directly linked to the buffer memory of the subdetectors electronics which can store a maximum of 128 events. After the L1 Trigger decision, the data flow rate is reduced from 40 MHz down to 100 kHz. The events are



Figure 4.22: Schematic longitudinal view of 1/4 of the CMS detector on which the muon detection system and its difference components are highlighted. The DT are present in the barrel part of CMS, the CSC are located in the endcaps, and RPC are found in both the barrel and the endcaps of CMS.



Figure 4.23: Architecture of the CMS Trigger System.
passed to the second level if not rejected by the L1.

The second level, High Level Trigger (HLT), is then investigating deeper the L1selected events in order to further reduce the rate from the 100 kHz down to the targeted 100 Hz for writing on disks. At this level a farm of processors is used to access and treat the full granularity information of all the CMS subdetectors. The rest of this section details the different components of the two levels.

Level-1 Trigger

From the rough information available at L1 the purpose of the executed algorithms is to identify electrons, muons, photons, jets and missing transverse energy. Three main subsystems are used for this task:

- L1 Calorimeter Trigger,
- L1 Muon Trigger,
- L1 Global Trigger.

The last of those is the one making the final decision regarding the selection or rejection of the processed event by combining the L1 Calorimeter Trigger and the L1 Muon Trigger outputs. The schematic representation of the CMS L1 Trigger is shown in figure 4.24.

L1 Calorimeter Trigger

The L1 Calorimeter Trigger uses the HF, HCAL and ECAL energy clusters, reconstructed using high level readout circuits called Trigger Primitive Generators by summing over the transverse energies measured in the crystals. For this purpose, calorimeters are subdivided into trigger towers. This primitive information is then transmitted by high speed serial links to the Regional Calorimeter Trigger whose purpose is to reconstruct low-level regional candidate electrons/photons, transverse energy sums and muons information via Minimum Ionising Particle (MIP) and isolations (ISO) bits. The Global Calorimeter Trigger subsequently sorts the objects (see figure 4.24) according to their transverse energy and forwards the first four to the L1 Global Trigger.

L1 Muon Trigger

The three subsystems, RPC, CSC and DT, of the muon system are used in the L1 Muon Trigger. From the RPC signals, the RPC trigger electronics builds Track Segments and an estimate of the track's transverse momentum. This information is then sent to the Global Muon Trigger. It additionally provides the CSC with useful information to solve possible ambiguities in case more than one muon track cross the same CSC module.



Figure 4.24: Architecture of the L1 Trigger.

The CSC trigger electronics builds track segments made out of the cathode strips alone and associates a transverse momentum estimate as well as a quality flag to each of the segments. These are then passed to the CSC Track Finder module which uses the full CSC information to reconstruct tracks, again assigning each track a $p_{\rm T}$ estimate and a quality flag before forwarding them to the Global Muon Trigger.

Similarly, the DT electronics looks for pattern of aligned hits in order to build Track Segments in each of the four chambers of a superlayer and then performs a recombination of segments from two superlayers leading to tracks candidates which are then sent to the Global Muon Trigger which finally sort the RPC, CSC and DT tracks before trying to recombine them and sending the four highest $p_{\rm T}$ muon candidates to the L1 Global Trigger.

L1 Global Trigger

From the information of the L1 Muon Trigger and the L1 Calorimeter Trigger, the L1 Global Trigger sorts all of the received objects before checking if at least one of the logical condition defined in the physical L1 trigger table is satisfied. If it is the case, the decision to keep processing the event is sent to the Trigger Control System which in turn sends the command to read the corresponding event data from all the remaining CMS subdetectors. Otherwise, the event is abandoned.

4.2. THE CMS DETECTOR

High Level Trigger

For each L1 accepted event, the HLT continues the event processing. For this stage, subdectors data are processed by a farm of computer close to the CMS detector using full granularity detector information and more complex algorithms with respect to L1. The goal of the HLT is to reduce its 100 kHz input to the 100 Hz manageable for writing on disks.

The event reconstruction at HLT can take up to 50 ms. In this period of time, partial tracks and calorimeter clusters reconstructions are performed. To achieve this, the HLT is subdivided into three sublevels executing more and more refined and complex algorithms, starting from simple cluster and supercluster reconstruction from ECAL and HCAL energy deposits taking into account possible bremsstrahlung radiation and building muon tracks from the muon chambers. In a second step, the pixel silicon detector tracker hits are used and combined with the clusters and superclusters reconstructed at the previous stage in order to restrict the full tracker domain to be explored. In this tracker region, track segments are build. Finally, the full tracker information, both pixel and strips, is used to complete the track reconstruction started at the previous stage and the full event analysis is performed and the final decision of keeping or rejecting the event can be made.

At each of the three sublevels, events are rejected if HLT requirements are not fulfilled. Doing so, the event rate is decreased. The subleveled architecture of the HLT allow for a rate smaller and smaller as the algorithm complexity of the consecutive steps increases.

Finally, if the event satisfies one of the HLT table requirements, it is sent to the Storage Manager which saves its entire raw content on disk, typically taking 1 to 3 MB of space per event. At this point, the content of an event is recorded as raw data. All subdectors information are saved and the full off-line reconstruction can start.

4.2.7 Treatment of the Data

CMSSW Framework

Within CMS, a framework has been developed to perform the off-line reconstruction of the selected events. The framework is called CMSSW and is written in C++ language with Python configuration files to simplify the execution setting and avoid recompilation.

The whole framework is based on the idea of an Event Data Model (EDM) in the sense that the access to the data has to be done on an event by event basis and that each event is represented by a C++ class containing all the information of the physical

event (raw detector information as well as fully reconstructed objects such as tracks, lepton candidates, etc ...). It also means that reconstruction algorithms can access raw detector information from the event, and attach reconstructed objects to the event. Finally, collections of events can be stored in ROOT files and accessed, event by event using the ROOT Data Analysis Framework [63].

Event Generation and Reconstruction

Beside the raw real collision data files stored after passing the HLT selection criteria, one can also generate similar pseudo data. Using one of the many available generator programs, physical events can be simulated by running the desired MC program within the CMSSW framework by use of dedicated libraries. The generator has to furnish all the required information regarding the currently generated event before saving it in a ROOT file. At this point, simulated events have been generated but the CMS detector effects has not yet been simulated.

To generate detector effects, the software GEANT 4 [64], a toolkit for the simulation of the passage of particles through matter, is used. First of all, some shifts and smearing is performed on the vertex position which lies exactly at the nominal interaction point in the generated data. This is done following the expected interaction position based on the beams position at every collision. Then the particles of the generated events are extrapolated through the detector by simulating their interaction with the material, the magnetic field present in the whole detector, and the electric field present in some of the subdectors. The result is a simulation of the CMS detector response to the particular generated event. From all subdectors of CMS the response is saved as done with real data events. The only difference being that for the generated envents the additional generator-level information is also stored into the ROOT files.

At this point the reconstruction of events can be performed on both real data files, containing raw detector data information for each recorded event, and on simulated data, using the simulated raw detector information. Doing so, the exact same program is run on both data and simulation to produce the reconstructed higher level objects belonging to a particular event be it a real event or a simulated event. The produced reconstructed level files are saved keeping the EDM format, the simulated ones also have their generated level information, under ROOT files and are accessible to the physicists for the analyses.

Chapter 5

Events Reconstruction

When two protons from the LHC beams collide inelastically a bunch of particles are created from the complex interaction taking place. Some of these final state particles are unstable and decay almost immediately after having been produced leading to secondary more stable particles. We usually distinguish particles that decay after a distance too short to be observable in the detector, the particles decaying at an observable secondary vertex (such as τ decays or B mesons decays), and larger lifetime particles that interact in the detector before decaying or that are stable. Among those the neutral particles do not leave any sign in the tracker while the electrically charged particles will leave a signal in the pixel and strips detectors corresponding to multiple hits on the different layers of the tracking system (see section 4.2.3). Most of the electrons and photons will deposit all their energy into the electromagnetic calorimeter, while hadrons will reach the hadronic calorimeter where they will stop after interacting with the material and releasing their energy. Finally muon particles will traverse the full detector leaving only a small energy deposit and neutrinos will not leave any signal in the detector.

The design of the detector, as explained in chapter 4.2 is deeply based on the different behaviours that different particles have after their production at the interaction point. Thanks to this, it is possible to classify most of the high energy particles that have been produced during the interaction. Electrons, photons, charged hadrons, neutral hadrons, muons, tau leptons and neutrinos, will all have a different signature in the detector. Concerning the neutrinos, their presence can be inferred from the balance of the transverse energy.

In this chapter the general ideas of the CMS particles reconstruction are explained, starting from the tracks and vertices reconstruction, going afterward to the leptons and photons reconstruction and finally reviewing the jets reconstruction which is of particle interest for this thesis.

5.1 Track and Vertex Reconstruction

Being produced close to the nominal point of interaction at the centre of the detector, the emitted particles will first go through the tracker system (see section 4.2.3 page 79). Doing so the electrically charged ones will induce a signal in some pixels of the pixel detector and then on some strips of the silicon strip modules on their way. From these signals it is possible to reconstruct a track corresponding to the trajectory of a charged particle.

The tracking reconstruction starts using the fact that an intense uniform and constant magnetic field (see section 4.2.2 page 77) is present inside the tracker volume and leads to curved trajectories for the charged particles due to the acting Lorentz force. A five-parameters helical trajectory can be used to infer the track of a particle. The first estimation of the parameters is obtained from a fit to the multiple pixel hits and the known beams transverse position.

At this point, the Kalman Filter algorithm [65] is used to refine the estimate of the helix parameters. The information from the silicon strip detector is then used. From this, a clusterisation process is performed and bunches of nearby hit strips are grouped together to form one hit cluster. The weighted strip signal mean is then used to estimate the real position of the particle together with its associated uncertainty. Starting from the first guess, the particle trajectory is inferred taking into account possible random multiple scattering until it reaches the first layer of the strip detector and a position is predicted along with its uncertainty. Combining the two estimates, *i.e.* the newly inferred position and the measured cluster belonging to the first layer, a new one is obtained and the parameters of the helix are updated by performing a least square minimisation. The remaining hits in the subsequent layers of the tracker are added in the same way by the Kalman Filter. This algorithm relies strongly on Gaussian probability densities on the uncertainty measurements. However, this assumption is found to be not optimal for the case of electrons tracks due to their high radiation energy loss. By consequence a slightly different algorithm is used for the electrons (see section 5.3.2 page 106).

Complications can arise when several hits are found to be compatible with the extrapolated point in a specific layer. Quality criteria are therefore applied on the reconstructed track in order to limit their numbers. Among these requirements are found the goodnessof-fit value, the number of valid track hits and also the impact distance with respect to the nominal point of interaction.



Figure 5.1: Magnified view of an event showing that 29 distinct vertices have been reconstructed corresponding to 29 distinct collisions within a single crossing of the LHC beams.

Finally, once all the clusters produced by a charged particle have been associated to a track, the helix parameters are fitted one last time starting this time from the outermost tracker layer. The reconstructed set of tracks is subsequently used together with the pixel hits in order to infer the positions of the interaction vertices. When all vertices have been found, they are sorted by decreasing values of the sum of the squared tracks transverse momenta. The vertex with maximum value being most of the time the one corresponding to the hardest collision.

At the LHC during the 2012 data taking, an average of 20 primary vertices is observed. For illustration purposes, figure 5.1 shows a typical event display for which 29 vertices and hundreds of tracks have been reconstructed.

5.2 Particle Flow

In CMS a particle flow event-reconstruction algorithm [66,67] has been developed aiming at identifying and reconstructing every stable particles, *i.e.* electrons, muons, photons, charged hadrons and neutral hadrons, emitted from the LHC proton-proton interactions by combining the information from all the subdectors. The CMS detector is by its concentric multilayers construction very appropriate for such a way of reconstructing the event particles.

The fundamental elements of the particle flow algorithm are its iterative-tracking strategy and its calorimeter clustering algorithm. High efficiencies and low misidentification rates are achieved thanks to the iterative-tracking, while the clustering algorithm provides high detection efficiencies even for low-energy particles and the capability to separate nearby energy deposits. With these complementary elements, a link algorithm connects the different pieces together to fully reconstruct each particle. The following sections describe in some more details each of these three building blocks of the particle flow algorithm.

5.2.1 Iterative Tracking

First of all both high efficiency and low misidentification rate are essential for this algorithm. Indeed a high efficiency is necessary because without the tracker detecting a charged particle would rely entirely on the calorimeters resulting in a lower efficiency, poorer resolution and biased direction due to the magnetic field. Concerning misidentified tracks, being randomly distributed, they would result in large contamination of the reconstructed energy.

To fulfil such requirements, the iterative tracking algorithm starts by reconstructing tracks with very tight criteria allowing for a negligibly small misidentification rate but on the other hand leading to a moderate efficiency. From the reconstructed tracks the unambiguously assigned tracker hits are removed from the hits collection. The next iteration then tries to reconstruct tracks from the remaining hits but with looser criteria and continuing that way loosening the seeding criteria. This iteration procedure increases the tracking efficiency and keeps the misidentification rate low from the simple fact that by removing the assigned hits the number of potential combination is significantly reduced as well. An efficiency of 99.5% is reached for isolated muons and larger than 90% for charged hadrons in jets, already after 3 iterations. The next iterations allow for the reconstruction of low- $p_{\rm T}$ (few hundreds of MeV), as well as the reconstruction of secondary vertex as far as 50 cm away from the nominal point of interaction, with a misidentification rate kept as low as a few percents.

5.2.2 Calorimeter Clustering

While the tracking provides a high efficiency for the reconstruction of charged particles, nothing can be said about the neutral particles emitted from the interaction. Detecting and measuring the energy of these last, such as photons and neutral hadrons is therefore one of the main goal of the clustering algorithm. Doing so, the algorithm must separate these neutral particles from energy deposits from charged hadrons. Finally, the clustering algorithm must also take care of the special case concerning electrons whose high radiation energy loss by Bremsstrahlung significantly impact the cluster shape. Additionally,

5.2. PARTICLE FLOW

the calorimeter information is important for the electron energy resolution as well as in determining the energy of poorly reconstructed tracks, which is the case for high- $p_{\rm T}$ tracks (due to their large radius of curvature).

The clustering algorithm can be separated in three steps. It starts by locating the multiple local calorimeter-cell energy maxima above a given threshold determining this way the "cluster seeds". In the second step, these clusters seeds are used to build "topological clusters" by adding neighbour calorimeter cells whose energy are above a given threshold and repeating this until no neighbour cell above threshold remains. The threshold values are chosen from the electronic noise distribution and range from 80 MeV in the barrel to 300 MeV in the end-caps for the ECAL and amount to 800 MeV in the HCAL. The obtained particle-flow clusters are then used to compute the distance to each cell in order to accordingly share the corresponding cell energy among all particle-flow clusters.

5.2.3 Link Algorithm

From the tracking and clustering elements obtained in the two previous steps of the particle-flow algorithm, the link algorithm as its name indicates will connect these two different information with each other in order to fully reconstruct each single particle and dismiss the potential double counting effects. For each pair of elements, the following algorithm is performed computing a distance in the (η, φ) plane which characterises the quality and likelihood of the combination.

To link a charged-particle track and a calorimeter cluster, the track extrapolation from its last reconstructed hit to a typical depth inside the calorimeter (one radiation length in the ECAL and one interaction length in the HCAL) passes within the cluster boundaries.

For the specific case of Bremsstrahlung photons, tangents to particle tracks at the different tracker layer positions are extrapolated to the ECAL and added as potential electron radiation if a cluster is found to match.

Similar links are performed between ECAL and HCAL clusters when the ECAL (the more granular calorimeter) cluster is found to be located inside the boundaries of a HCAL (less granular calorimeter) cluster, as well as between a tracker track and a muon track (reconstructed by the muon detection system) if the two tracks are compatible (in case of multiple possible combination sharing a same element, the most likely one, in terms of its χ^2 , is selected).

The algorithm forms block of elements linked to each other from which the particle

reconstruction and identification can be performed. For this, the expected behaviour of each possible particle is used in order to identify one by one the event's particles. The procedure starts by the muon identification which removes the corresponding building elements every time a muon is identified. The algorithm then proceeds similarly for the electrons for which potential electron tracks are refitted using the GSF algorithm (see section 5.3.2). From the remaining tracks, charged hadron are identified and their corresponding tracks and calorimeter clusters are removed form the collections. Among the not yet selected clusters, photon and neutral hadrons are identified using the calorimeter source (ECAL or HCAL) in order to distinguish between them.

5.3 Leptons

Even though the particle flow algorithm can provide all the particles candidates necessary for the analysis performed in this thesis, it has been decided to not use this particle collection for the selection of electrons. This section is therefore describing the more standard way of reconstructing lepton candidates in CMS. Furthermore these methods for reconstructing the candidatesallow one to be aware of the different problems that can arise when reconstructing the event particles. For the sake of completeness, muon reconstruction is also described here but it is the particle flow algorithm that is used for their reconstruction in the analysis.

5.3.1 Muons

The muons produced at the centre of the detector are one of the two types of particles that will most of the time go through the entire detector, as seen on figure 4.9 on page 78 and escape it (the second type being the neutrinos (see section 5.3.4)). Indeed its long mean lifetime (2.2 µs) and its small energy loss when traversing matter ensure a probability for a GeV muon to escape the detector close to unity. Furthermore its mass of 105 MeV (200 times heavier than the electron) makes the radiation loss (key principle of the calorimeters) very small. The muons hence leave only a minimum ionisation deposit along their trajectories and escape the detector.

The muon reconstruction is thus entirely based on the tracker and the muon chambers signals. The process starts with a set of tracks corresponding to the muons trajectories from each subdetector system. When a track and its parameters obtained from the tracker system are found to match another track in the muon chambers (via interpolation by a Kalman Filter) a fit is performed on the complete trajectory and new parameters are



Figure 5.2: Muon identification (left) and isolation (right) efficiencies as measured by the CMS Collaboration using the "tag and probe" method as a function of the probe muon candidate pseudo-rapidity η_{μ} with $20 \leq p_T \leq 500$ GeV. The green and blue circles represent the efficiency measured from data and MC, respectively. The red circles represent the ratio data over MC and is called scale factor. The MC sample are corrected by these scale factors for the analysis. The uncertainties on these efficiencies are statistical only. These results are taken from [68].

extracted.

Once muon candidates have been reconstructed, various quality criteria can by applied on top of the reconstruction, such as a minimum $p_{\rm T}$, a minimum number of hits, ... This selection is analysis dependent and will be discussed in more detail in section 6.3.2 but for the sake of illustration, figure 5.2 shows typical efficiencies obtained from so-called tight criteria [68], evaluated with the data driven "tag and probe" method. This approach consists in selecting two charged lepton candidates (electron or muon) (the same method is used for to measure analogous electron efficiencies) whose invariant mass, M_{ll} , is found to be inside the window mass $|M_{ll} - M_Z| \leq 30 \,\text{GeV}$. This requirement, together with identification and isolation criteria imposed on one of the two leptons (hence highly tagged to be a true lepton, called the tag), enforces the probability that the second lepton candidate be also a true lepton (called the probe). By studying the effect of the trigger or selection criteria on the number of such pairs, we can measure the different efficiencies intervening in the different analysis steps. A global identification efficiency of about 95% is observed for the muon selection in the central region of CMS and for transverse momenta above 20 GeV, which has to be multiplied by the isolation efficiency of about 98%.



Figure 5.3: Illustration of the bremsstrahlung effect on an electron produced during the interaction and of photon conversion happening when a photon passes through the tracker material (left). Spread of the ECAL energy deposit associated to an electron caused by the bremsstrahlung effect (right).

5.3.2 Electrons

As mentioned in the previous section electrons are much lighter than muons and the radiation energy loss, which depends on the inverse square of the mass, becomes significant for the electrons while it was negligible for the muons. By consequence electrons will most of the time deposit their complete energy inside the ECAL calorimeter (almost no energy deposit is expected in the HCAL) where they will create and electromagnetic shower. Additionally the electrons will leave a track in the tracker. These two elements, the track and the energy deposit, form the expected signature of an electron and once a track is found to match an ECAL supercluster energy deposit, an electron candidate is reconstructed.

As mentioned before, electrons' tracks are not reconstructed with the Kalman filter [69]. Indeed, the bremsstrahlung energy loss distribution of electrons propagating in matter is highly non-Gaussian while Kalman filter assumes such Gaussian distributions. A Gaussian-Sum Filter (GSF) algorithm has therefore been developed and implemented for the reconstruction of electrons in the CMS tracker. The idea behind this is to model the radiation loss distribution by a Gaussian mixture rather than by a single Gaussian, which has been shown to improve the electron momentum resolution [70].

Once the GSF tracks have been reconstructed, a GSF electron candidate can be built by associating an ECAL supercluster and a GSF track with compatible η and ϕ angles. The matching criteria are listed in table 5.1. When a valid association is found, a combination of the track and supercluster information provides the electron momentum

Table 5.1: Set of conditions required to build a GSF electron candidate based on a ECAL driven seed.

Seeding:
Two hits in the first layers of the tracker compatible with the tra-
jectory extrapolated using the supercluster information.
$H/E \leq 0.15$ ratio of the HCAL H to ECAL E energy deposits
Track-SC matching:
$ \Delta\eta_{\rm in} \leq 0.02~$ difference between extrapolated η measured at the inner tracker layer and the supercluster η
$ \Delta\phi_{\rm in} \le 0.15~$ difference between extrapolated ϕ measured at the inner tracker layer and the supercluster ϕ

and charge sign measurements. The efficiency of the GSF electron is found to be around 95% for electrons with $p_{\rm T} \geq 35 \,\text{GeV}$.

Additional complications arise for the case of electrons with respect to muons. As a matter of fact, photons emitted by the electrons propagating through the relatively large tracker material (see figure 4.12 page 81) can affect the energy measurement in two ways. Firstly these photons can reach the ECAL and give rise to energy deposit nearby the electron's one. However, since photons are neutral they will not be bent by the magnetic field as illustrated on figure 5.3. The electron and photon trajectories will not coincide. The result is a broader energy deposit in the ϕ direction as shown on figure 5.3, which has to be correctly taken into account when computing the electron's energy. Secondly the radiated photon can also lead to pair creation when passing through the tracker layers. From this, an electron and a positron are created which will be bent in two different directions also resulting in a large spread of the ECAL deposit in the ϕ direction. When integrated along the electron trajectory the total radiated energy can be very large as figure 5.4 shows for different electrons transverse momenta [70]. It was found out that about 35% of the electrons radiate 70% of their initial energy before reaching the ECAL and 95% in 10% of these cases making essential the collect of the bremsstrahlung photons energy in order to reconstruct electrons. To achieve this goal, the super-clustering algorithm uses the transverse properties of the electron shower shape while dynamically searching for separated (bremsstrahlung) energy in the φ direction in order to build cluster seed from collections of 3 to 5 crystals continuous in η in the barrel and by connecting rows of crystals containing energies decreasing monotonically when moving away from a seed crystal together with other such seeds in both x and ydirections in the endcaps [71].



Figure 5.4: Distribution of the fraction, $\frac{\sum E_{brem}^{\gamma}}{E^{e}}$, of the generated electron energy (E^{e}) radiated as bremsstrahlung photons $(\sum E_{brem}^{\gamma})$ for electrons of 10, 30 and 50 GeV. The true emission of bremsstrahlung photons has been integrated up to a radius corresponding to the ECAL inner radius. Figure taken from [70].

As for the muons, additional criteria regarding the shower shape, isolation, ... can be required on the electron candidates and will be discussed in section 6.3.2 but for the sake of illustration, figure 5.5 shows a typical efficiency obtained from so-called medium criteria. A global identification efficiency of about 85% is observed in the $p_{\rm T}$ region around 45 GeV for $|\eta| \leq 0.8$. On the same figure is illustrated the tag and probe method used to extract the efficiencies from data for a particular bin of the efficiency plot.

5.3.3 Taus

Taus are the third and last type of charged leptons. Tau leptons are unstable particles with a lifetime of about 2.9×10^{-13} s. For taus originating from a Z boson decay with a typical transverse momentum of about 45 GeV, they could only travel a distance of a few millimetres given by $\frac{p}{m}c\tau = \frac{45}{1.7} c\tau = 2.2$ mm. They will therefore decay after the interaction before reaching any piece of detector and the only way to reconstruct them is by their decay products. The decay process of taus involves the production of a virtual W boson together with a tau neutrino. In about one third of the cases, the W boson will afterward decay leptonically (into a muon or an electron together with the corresponding neutrino). The identification of tau leptons in this case therefore relies



Figure 5.5: Electron identification efficiencies for $|\eta| \leq 0.8$ (left) and tag and probe method fits for the $30 \leq p_T \leq 40$ GeV bin as measured by the CMS collaboration. The blue line is the fit result of all tag and probe candidates invariant mass distribution. The green line is the fit result of the tag and probe candidates invariant mass distribution for which the probe passes the criteria under study. Finally, the red line is the fit result of the tag and probe candidates invariant mass distribution for which the probe fails to pass one of the criteria under study. In each case, the fit function is the sum of a signal (solid line) and a background (dashed line) contribution. Results taken from [72].

on the reconstructed charged lepton. The rest of the time, the W decay hadronically resulting in jets characterised by a rather small track multiplicity. This specificity of tauoriginated jets makes it possible to implement dedicated algorithms, in order to identify taus. In all cases, the tau neutrino induces the presence of missing energy in the event, which is another sign for the potential production of a tau lepton.

5.3.4 Neutrinos

Neutrinos are neutral leptons therefore interacting only through the weak force. This makes them unobservable in the CMS detector. The neutrinos will escape with their energy not being collected resulting in an imbalance of the transverse energy. The initial transverse energy of the collision being close to zero, from energy-momentum conservation, it must remain so when all final state particles transverse energies are taken into account. But since the neutrinos are not detected, the resulting reconstructed transverse missing energy can be attributed to the neutrinos. Unfortunately, when two neutrinos are emitted back-to-back in the transverse plane they can compensate each other and no missing transverse energy would be seen in such a case.

Several methods exist to build the missing transverse energy. A common one in CMS uses the full set of reconstructed particle flow objects. And from the sum of the corresponding four-momenta estimates the missing transverse energy.

5.4 Photons

Photons as seen by the CMS detector are very close to electron candidates. The main difference is the absence of a track pointing to the ECAL supercluster deposit. A photon is therefore nothing else than an ECAL supercluster energy deposit.

Complications similar to the case of the electron's reconstruction are also present for the photon reconstruction. A photon can convert into an electron-positron pair (as shown in figure 5.3). This happens for half of the photons. In that case, the created electron and positron will, unlike the photon, be bent by the magnetic field. As for the electron, the result is a spread of the energy cluster in the ϕ direction. The photon and electron shower shapes are by consequent very similar to each other and in some cases make the electron/photon disambiguation difficult. Several algorithms have been developed in order to increase the distinction between the primary particles. The main handle to distinguish converted photons and prompt electrons or positron is the pixel detector.



Figure 5.6: A jet, its tracks (green and azure solid lines), its photon candidates (purple dashed lines), its neutral hadrons (red dashed lines), its charged hadrons (azure solid lines, and its energy deposits in CMS's electromagnetic (khaki coloured blocks) and hadronic (teal blue coloured) calorimeter.

5.5 Jets

Jets are made of collimated hadrons and other particles produced by the hadronisation of a quark or a gluon coming from the proton-proton collision. They form the experimental signatures of these partons. The behaviour of a jet inside the CMS detector is by consequent determined by the particles that constitute it, *i.e.* mainly hadrons. The charged hadrons will leave tracks in the CMS tracker pointing directly to the calorimeter deposits formed by the energy release of both neutral and charged hadrons. These basics signals are then combined in order to form jets.

5.5.1 Jet Algorithm

Several approaches exist to reconstruct a jet but the one offering the best results in CMS is based on the particle flow particles. Furthermore given an input collection of particles and their four-momenta p_i various algorithms can be used to group the different particles into jets. CMS has opted for the so-called anti- k_t algorithm [73] whose different steps are listed below:

1. list all the distances, d_{ij} , between particles i and j, given by

$$d_{ij} = \min\left(\frac{1}{k_{t,i}^2}, \frac{1}{k_{t,j}^2}\right) \times \frac{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}{R^2}$$
(5.1)



Figure 5.7: A sample parton-level event, together with many random soft particles, clustered with four different jets algorithms. Clearly visible on the anti- k_t is the disc surface of the jets. Figure taken from [73].

with $k_{t,i}$, y_i and ϕ_i the transverse momentum, rapidity and azimuthal angle of particle *i*, respectively, and *R* the algorithm parameter whose value can be chosen. The main choice in CMS is R = 0.5 while ATLAS uses R = 0.4.

2. list all the distances, d_{iB} , between the particle *i* and the beam, given by

$$d_{iB} = \frac{1}{k_{t,i}^2}$$
(5.2)

- 3. find the minimum of all the distances
- 4. if the minimum is of type d_{iB} the corresponding particle forms a jet and is removed from the list of particles
 - if the minimum is of type d_{ij} the corresponding particles *i* and *j* are grouped together into a new particle by adding the two corresponding four-vectors. The new particle's transverse momentum is therefore $k_{t,i} + k_{t,j}$. This new particle is added to the list and the two particles *i* and *j* are eliminated
- 5. repeat the process from step one until all particles are assigned to a jet.

This algorithm has the advantage of clustering the soft particles with hard ones long before the soft particles cluster among themselves. An isolated hard particle, *i.e.* with no hard neighbours within a distance 2R will simply collect the soft particles lying inside a cone of radius R. The result is a jet conical shape not depending on the low- $p_{\rm T}$ particles but still including them.

When analysing data from the detector or from simulation, three levels can be distinguished. The reconstructed level, reco-level for short, is the one observed after detection. The generated particle level, gen-level for short, contains the physics just before the detector simulation. Finally, the parton level corresponds to the particle distributions before any hadronization. For each of these three levels, the same jet algorithm can be applied, using different input collections of particles. At reco-level, the set of particle flow candidates is used in both data and MC simulation. For the gen-level, the algorithm is run with all the generated particle present before the simulation of the detection. Finally, at parton level, the input collection of particles is made of all particles of the final state before hadronization.

The anti- k_t algorithm described above is one of the different clustering jet algorithms family where the exponent of the transverse momentum terms in the minimum factor of (5.1) is -2. The K_t algorithm's power values are on the opposite 2. The consequence is that, while the anti- k_t first selects a hard particle and its closest neighbour, the K_t will first select a soft particle and its closest neighbour. The anti- k_t therefore tries to group the soft particles around a hard one in priority, while it is not the case for the K_t algorithm. The Cambridge/Aachen jet algorithm drops completely the minimum factor in equation (5.1) and thus relies only on distances between particles to cluster them into jets. Another family of jet algorithms known as cone algorithms exists and the SISCone algorithm is one of them. It defines a jet as an angular cone around some direction of dominant energy flow, the direction being determined by establishing the list of particles in a trial cone, evaluating the sum of their 4-momenta, and use the resulting 4-momentum as a new trial direction for the cone. This procedure is iterated until the cone direction no longer changes, *i.e.* until one has a stable cone. Comparison of these four jet algorithms is shown in figure 5.7 from [73]. CMS has also used the "iterative cone" algorithm in some studies [74]. This algorithm takes the hardest particle in the event, uses it to seed an iterative process of looking for a stable cone, which is then called a jet. Every particle belonging to that jet is then removed from the event and the procedure is repeated with the next hardest available remaining seed until no seed remains.

5.5.2 Jet Energy Correction

The energy of the reconstructed jets has to be corrected for different effects [75–77]. These corrections are factorised in multiple levels each correcting for a different effect

and essentially being a scaling of the jet four-momentum with a scale factor (correction) which depends on various jet related quantities ($p_{\rm T}$, η , flavour, etc.). The levels of correction are applied sequentially. The output of one step is then used as the input to the next level.

Level 1, 2 and 3 Corrections

The corrections start with the first level, L1, whose primary goal is to subtract pileup and electronic noise from the jet reconstructed energy. Indeed, additional protonproton collisions occurring close enough in time can contribute to the same (or close by) calorimeter clusters and the electronic noise can randomly activate a calorimeter's cell belonging to a jet. Both effects are expected to raise up the jet's energy. Their contributions have therefore to be subtracted. After this average offset correction any dataset dependence on luminosity is in principle removed so that the next corrections are applied on a luminosity independent sample.

The second level of correction, L2, deals with the non-uniformity of the CMS detector response to jets in the η direction and aims at making it flat as a function of η . This is achieved by correcting a jet with arbitrary η with respect to a jet in the central region $(|\eta| < 1.3)$. This relative correction can be derived either by using pure MC truth or by employing a data driven method such as the dijet $p_{\rm T}$ balance or with a tag and probe method on a two-body process X+jet, for which the tag object is the well measured X object, it could be a photon or a Z boson, and where the probe is the jet whose response has to be estimated. Figure 5.8 illustrates these methods and the corrections obtained this way.

Once L2 corrections have been applied, the next level, L3, can be used in order to remove jet response variations in the CMS detector as a function of $p_{\rm T}$ which primarily result from calorimeter non-linear response. This absolute correction brings the jet reconstructed energy back to its particle level value in the sense that the corrected jet energy is equal to the particle-level jet energy, on average. Its purpose is to make the response equal to unity at all $p_{\rm T}$ for the control region $|\eta| < 1.3$. Estimates of this correction can be obtained either by using MC truth information or by employing data-driven methods such as γ + jet or Z + jet $p_{\rm T}$ balance. Figure 5.9 illustrates the response and the corrections corresponding to this jet energy level correction. Figure 5.10 shows the combined effect of L2 and L3 corrections.

Other levels of corrections exist such as a correction for variations in jet response with electromagnetic energy fraction or a correction to particle level for different types of jet (light quark, c, b, gluon). These additional levels are found not to be crucial for the



Figure 5.8: *MC* simulation of dijet transverse momenta balance as a function of the the probe jet candidate pseudo-rapidity (left). The relative jet response determined from dijet balance (filled circles) is compared to the relative jet response from MC truth (open boxes). Correction validation as a function of the probe jet candidate pseudo-rapidity (right). The relative jet response determined from dijet transverse momenta balance after Level 2 jet corrections, where the corrections are derived from dijet transverse momenta balance form the correction of the probe is taken from [76].



Figure 5.9: Simulated calorimeter response to jets versus particle jet generated transverse momenta, p_t^{gen} for iterative cone jets with R = 0.5 (left). MC based L3 correction as a function of calorimeter jet p_T for iterative cone jets with R = 0.5 (right). Results taken from [76].



Figure 5.10: CaloJet response $\binom{p_T^{cor}}{p_T^{gen}}$ as a function of pseudorapidity without corrections (filled points), with level 2 corrections (open squares), and with the combined Level 2 & Level 3 correction (open circles) from MC sample. Results taken form [76].

present analysis and for most analyses in CMS and we often limit ourselves with levels 1, 2 and 3 corrections.

L2L3 Residual Calibration of Data

After having applied the previous three levels of correction on both data and MC predictions, a few small discrepancies remain while the MC truth calibration was found to be good enough. An additional η and $p_{\rm T}$ dependent correction, called the L2L3Residual, has therefore been computed and is applied on the data sample only in a way to fix the small differences between data and MC.

Chapter 6

Z Boson Production in Association With Jets

This chapter presents the measurement of the differential cross section of Z boson production in association with jets, in both the muon and electron decay channels, in ppcollisions at a centre-of-mass energy of $\sqrt{s} = 8 \text{ TeV}$ with 19.6 fb⁻¹ of data collected by the CMS detector in 2012. The differential cross section is presented as a function of the jet multiplicity, the transverse momentum of the n^{th} jet, the absolute pseudorapidity of the n^{th} jet and the scalar sum of the jets p_{T} , for $n = 1, \ldots, 5$. The dijet invariant mass spectrum is also measured. The differential cross sections are compared to theoretical predictions from MADGRAPH 5, SHERPA 2 and MADGRAPH5_AMC@NLO. Every step of the analysis is reported in some details. This analysis constitutes the core of the data analysis part of this thesis.

6.1 Introduction

6.1.1 Why Study this Process ?

The large centre-of-mass energy of pp collisions at the LHC allows for the production of events with high jet transverse momenta and high jet multiplicities in association with a Z boson¹. The decay of the Z boson to two oppositely-charged muons or electrons provides a signal that is almost background free. This signature, which has a high reconstruction efficiency thanks to the presence of the charged leptons in the final state, is therefore a

¹For convenience, the Z boson notation actually stands for Z/γ^* unless explicitly mentioned.

"standard candle" well-suited to the validation of the SM (EWK and QCD) calculations. Due to the large centre-of-mass energy at the LHC, Z bosons can be produced with the largest amount of jets ever observed, whose description provides stringent tests of perturbative QCD (pQCD). Furthermore, the production of massive vector bosons with jets is an important background to a number of Standard Model processes (single top, $t\bar{t}$, vector boson fusion, WW scattering, Higgs boson production) as well as in Beyond Standard Model searches such as supersymmetry. Therefore the measurement of the Z + jets cross sections as a function of various kinematic variables that are relevant to this process is crucial at the LHC and its simulation by MC programs has to be pushed to the highest possible precision level.

In the present state, calculations of pQCD can be tested at leading order (LO) with multiparton matrix-element event generators such as MADGRAPH 5 and at beyond leading order accuracy with the recent development of next-to-leading-order (NLO) event generators such as SHERPA 2 and MADGRAPH5_AMC@NLO, which will be used for the comparison with the measurements and were described in chapter 3. Previous measurements of the Z + jets cross section were reported by the CDF and D0 collaborations with proton-antiproton collisions at a centre-of-mass energy $\sqrt{s} = 1.96$ TeV as mentioned in chapter 3. More recent results from proton-proton collisions at centre-of-mass energy $\sqrt{s} = 7$ TeV were published by the ATLAS and CMS collaborations and were also briefly described previously in chapter 3. In this work, thanks to the increased cross section and luminosity compared to previous CMS and ATLAS measurements, distributions of jet $p_{\rm T}$ and $|\eta|$ up to the fifth leading jet can be measured in addition to the extension of the measured phase space for lower multiplicities.

6.1.2 Measured Observables

The differential cross sections are measured as a function of the following jet properties :

- exclusive and inclusive jet multiplicities, N_{jets} ,
- transverse momenta, $p_{\rm T}$, of the $n^{\rm th}$ jet for $N_{\rm jets} \ge n$ with n = 1, ..., 5,
- scalar sum of the jets transverse momentum, $H_{\rm T}$, for jets inclusive² multiplicities from 1 to 5.
- pseudorapidities, η , of the n^{th} jet for $N_{\text{jets}} \ge n$ with n = 1, ..., 5,

²Inclusive quantities are measured for events with a given minimum number of jets passing the selection criteria, *i.e.* $N_{\text{jets}} \ge n$, while exclusive quantities concern events with a specific given number of jets passing the selection criteria, *i.e.* $N_{\text{jets}} = n$.

6.2. DATA AND SIMULATION SAMPLES

The jet multiplicity variable is one of the first test that can be performed on a MC generator to study its jets rate predictions. By comparing different predictions to measurements, one can test different MC techniques, estimate their reliability in different configuration as the number of jets, and therefore the complications of accurately predicting such a high number of radiation, increases.

By measuring the transverse momentum distributions of the leading, subleading, subsubleading, ..., jet, it is possible to quantify the level of agreement between data and theory. Predicting large- $p_{\rm T}$ jets with parton shower techniques alone is not satisfactory, while the prediction of small- $p_{\rm T}$ with matrix elements calculation leads to divergences. The transition region, where the merging of ME and PS occurs, is of particular interest to make statements regarding a particular matching scheme (see section 3.1.5).

The $H_{\rm T}$ variable is of great interest for searches like supersymmetry in all-hadronic events analyses based on an event signature with many jets [78]. These events are expected to be accompanied by large missing transverse momentum and one of their main backgrounds is the irreducible background from Z + jets events, whit the Z boson decaying to $\nu \bar{\nu}$. By studying the $H_{\rm T}$ variable associated to the production of a Z boson, be it in the charged lepton decay channel or not, one can furnish precise background estimation to these supersymmetry searches.

Also related to searches for new physics, the dijet mass distribution is of important interest. Indeed some new physics scenarios are expected to give rise to resonant dijet production [79]. Again, it is important to control the Standard Model background potentially leading to two jets in the final state, as is the case of the present analysis.

6.2 Data and Simulation Samples

The present analysis is based on data collected by the CMS Collaboration during 2012 (see table 6.1 page 120). We restrict the analysis to the lumi sections (parts of an LHC run) validated in the JSON file also reported in table 6.1 corresponding to an integrated luminosity of $19.6 \,\mathrm{fb^{-1}}$. This analysis has been done using the CMSSW_5_3_11 software release.

A list of MC samples, given in table 6.2, has been used in the analysis for detector level comparisons, non negligible background contamination estimation and detector effects corrections. The signal is modelled with a set of five subsamples of DY simulated events restricted to opposite charged leptonic Z decays with an invariant mass above 50 GeV. The first signal file in table 6.2 contains $pp \rightarrow Z + X \rightarrow l^+ l^- + X$ events, where $l^+ l^- = e^+ e^-$, $\mu^+ \mu^-$ or $\tau^+ \tau^-$ and X represents 0 to 4 final state partons. This sample is therefore called the inclusive sample and contains about 30 millions of generated events. During the analysis work, this 30 millions events sample was found to be rather limited in number of events for the jet multiplicities above 2 (being an inclusive it is highly dominated by the 0 and 1 final state parton events) which was a problem for the technique used to correct for detector effects. To increase the statistic for the DY sample, four additional exclusive samples (file 2 to 5 in table 6.2) have therefore been used: Z + 1, Z + 2, Z + 3 and Z + 4 final state partons. Each of these samples has been renormalised to the corresponding prediction of the inclusive sample.

The Z boson and the partons are generated with the matrix element generator MAD-GRAPH 5 [26] using the PDF set CTEQ6L1 from [80]. The parton shower and hadronisation effects are simulated by PYTHIA 6 [28] using the Z2 tune [81].

Background processes coming from double electroweak boson production (the 6 last files in the background list of table 6.2) with decay modes leading to at least two charged

Table 6.1: Datasets with the corresponding number of recorded events and integrated luminosity, and JSON file used in the analysis.

Name	# Events	\mathcal{L} (fb ⁻¹)
/DoubleMu/Run2012A-22Jan2013-v1	5636274	0.88
/ DoubleMuParked/Run2012B-22Jan2013-v1	29308627	4.40
/ DoubleMuParked/Run2012C-22Jan2013-v1	36820243	6.95
/ DoubleMuParked/Run2012D-22Jan2013-v1	38006513	7.35
/DoubleElectron/Run2012A-22Jan2013-v1	12964286	0.88
/ Double Electron/Run 2012 B-22 Jan 2013-v1	23571931	4.40
/ Double Electron/Run 2012 C-22 Jan 2013-v1	33843769	6.99
/DoubleElectron/Run2012D-22Jan2013-v1	34526899	7.34
/MuEG/Run2012A-22Jan2013-v1	1939498	0.88
$/{\rm MuEG}/{\rm Run2012B}\text{-}22{\rm Jan2013}\text{-}v1$	13502523	4.41
/MuEG/Run2012C-22Jan2013-v1	20027780	7.03
$/\mathrm{MuEG}/\mathrm{Run2012D}\text{-}22\mathrm{Jan2013}\text{-}v1$	20830686	7.34

JSON	1				
Cert	_190456-208686_	_8TeV_	_22Jan2013ReReco_	_Collisions12_	_JSON.txt

leptons, as well as $t\bar{t}$ production (the second file in the background list of table 6.2) are modelled with MADGRAPH 5 and PYTHIA 6 too. Single top events (files 3 to 8 in the background list of table 6.2) have been produced using the POWHEG MC generator. For all backgrounds except the W + jets, tau decays are handle by interfacing the MC generator to the TAUOLA [82] libraries dedicated to this aim.

The cross section for the signal is normalised to match the next-to-next-to-leading order (NNLO) prediction for inclusive Z production obtained with FEWZ [14] and using CTEQ6M PDF set [83]. The double boson background samples are normalised to the next-to-leading-order (NLO) predictions calculated by the MCFM [84] generator using the CTEQ PDF set. The $t\bar{t}$ cross section is normalised to the next-to-next-to-leading-log (NNLL) calculation from [85]. It has to be noted that background contribution coming from the DY process with Z decaying into a $\tau^+\tau^-$ pair, already included in the signal sample, is also considered in the analysis.

For comparison to data at reconstructed level, each simulated sample has been normalised to the luminosity of the data sample.

6.3 Event Selection

In this section the full event selection is described. After listing the triggers used in this analysis, the requirements that are imposed on the lepton candidates, separately for the muons and electrons, are described. The pairs made of two opposite charged leptons, $\mu^+\mu^-$ or e^+e^- , are required to have an invariant mass around the Z mass peak. Finally the procedure for jet reconstruction is explained. The event should contain at least one such jet to fulfil the selection requirement.

6.3.1 Trigger

The data samples used in this analysis were collected with unprescaled³ Double Muon and Double Electron triggers during the whole 2012 data-taking period. For the muon

³Because the instantaneous luminosity decreases with time during a single LHC run (see section 4.1.4), some trigger are found to be fired too often at the high luminosity of the run while their rate is low enough for the low luminosity period of the run. Therefore, during the high instantaneous luminosity period of a run, a prescale is applied on some trigger in order to artificially decrease their rate.

channel, the events passing the following trigger have been selected:

 $HLT_Mu17_Mu8_v^*,$

asking for a minimum of two HLT muons with $p_{\rm T}$ thresholds of 17 and 8 GeV. For the electron channel, the corresponding trigger is:

HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_ Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*,

requiring an HLT electron with $p_{\rm T}$ threshold of 17 GeV, a tight calorimeter identifica-

Table 6.2: MC files used in the analysis with their respective cross section times branching ratios. For the names, TZ2 stands for TuneZ2, mad is for madgraph, and S12 stands for Summer12_DR53X-PU_S10_START53_V7A.

Name	# Events	σ (pb) × BR
Signal		
/DYJetsToLL_M-50_TZ2Star_8TeV-mad-tarball/S12-v1	30459503	3531.8
$/\rm DY1JetsToLL_M-50_TZ2Star_8TeV-mad-tarball/S12-v1$	24045248	666.30
$/\rm{DY2JetsToLL_M-50_TZ2Star_8TeV-mad-tarball/S12-v1}$	21852156	214.97
$/\rm DY3JetsToLL_M-50_TZ2Star_8TeV-mad-tarball/S12-v1$	11015445	60.69
$/\mathrm{DY4JetsToLL_M-50_TZ2Star_8TeV-mad-tarball/S12-v1}$	6402827	27.36
Background		
$/ WJetsToLNu_TZ2Star_8TeV-mad-tarball/S12-v1,2$	76102995	36864
$/{\rm TTJets_MassiveBinDECAY_TZ2star_8TeV-mad-tauola/S12-v1}$	6923652	245
$/T_s-channel_TZ2star_8TeV-powheg-tauola/S12-v1$	259961	3.79
$/{\rm T_t-channel_TZ2star_8TeV-powheg-tauola/S12-v1}$	3758227	56.4
$/T_tW-channel_TZ2star_8TeV-powheg-tauola/S12-v1$	497658	11.1
$/ Tbar_s\text{-}channel_TZ2star_8TeV\text{-}powheg\text{-}tauola/S12\text{-}v1$	139974	1.76
$/ Tbar_t\text{-}channel_TZ2star_8TeV\text{-}powheg\text{-}tauola/S12\text{-}v1$	1903681	30.7
$/Tbar_tW\text{-}channel_TZ2star_8TeV\text{-}powheg\text{-}tauola/S12\text{-}v1$	493460	11.1
$/ ZZ Jets To 2 L2 Q_TZ 2 star_8 TeV-mad-tauola/S12-v1$	1936727	17.654×0.14
/ ZZ Jets To 2 L2 Nu TZ 2 star 8 TeV - mad - tauola / S12 - v3	944911	17.654×0.04
$/ ZZ Jets To 4 L_TZ 2 star 8 TeV - mad - tauola / S12 - v1$	4807893	17.654×0.01
$/WZJetsTo2L2Q_TZ2star_8TeV-mad-tauola/S12-v1$	3215990	33.21×0.07
$/WZJetsTo3LNu_TZ2_8TeV-mad-tauola/S12-v1$	1995334	33.21×0.03
$/WWJetsTo2L2Nu_TZ2star_8TeV-mad-tauola/S12-v1$	1933235	54.838×0.11

tion, a very loose track identification and isolation, and a second HLT electron with $p_{\rm T}$ threshold of 8 GeV, a very loose calorimeter identification and some track identification and isolation, as listed in table 6.3.

The double muon trigger efficiencies measured by the CMS collaboration using the tag and probe method are shown in figure 6.1. Additionally, the scale factors (ratio of MC to data efficiencies) are also presented. As can be seen from this figure the MC predictions are overestimating the trigger efficiency resulting in scale factors smaller than unity by a few percents. In order to correct for this effect, the scale factors are applied as weight to the MC simulated events. As far as the electron trigger is concerned, no such discrepancy was observed by the CMS collaboration and therefore no trigger scale factor need to be accounted for.

6.3.2 Leptons

This analysis aims to study the jets associated to the production of a Z boson. The selected events are thus required to contain a minimum of two opposite electric charge muons or electrons.

Table 6.3: Identification and isolation criteria using calorimetry and tracker information at trigger level for the double electron trigger. The quantity H/E represents the ratio of the energy deposit in the HCAL to the energy deposit in the ECAL, $\sigma_{i\eta i\eta}$ is a measure of the η covariance of the 5×5 matrix centred on the GSF electron seed crystal, ECalIso/ E_T , $HCalIso/E_T$ and $TrkIso/E_T$ are respectively the ratio of the electromagnetic, hadronic, and tracker isolations over the transverse energy, $d\eta$ and $d\phi$ represent the difference in pseuorapidity and azimuthal angle obtained from tracker and calorimeter informations.

Criteria	Barrel Selection	Endcap Selection	
Calldi	H/E < 0.15	H/E < 0.10	
Callul	$\sigma_{i\eta i\eta} < 0.014$	$\sigma_{i\eta i\eta} < 0.035$	
CalldT	H/E < 0.10	H/E < 0.075	
Canui	$\sigma_{i\eta i\eta} < 0.011$	$\sigma_{i\eta i\eta} < 0.031$	
ECalls ECalls		$p/E_{\rm T} < 0.2$	
	$HCalIso/E_{\rm T} < 0.2$		
TrleIdVI	$d\eta < 0.01$		
	$d\phi < 0.15$	$d\phi < 0.10$	
TrkIdVL	TrkIso /	$E_{\rm T} < 0.2$	



Figure 6.1: Histograms showing the double muon trigger efficiencies as measured by the tag and probe method on the 2012 data sample (top), simulated DY + jets sample (middle), and the corresponding scale factors (ratio of MC to data efficiencies) (bottom) as a function of the leading and subleading muons absolute pseudorapidities, for $p_T > 20$ GeV. Results from [86].

Muons

The muon candidates are selected from the particle flow collection (see section 5.2) on which a matching with the trigger objects is required. The matching is made in $\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2}$ with $\Delta R \leq 0.3$, where $\Delta \phi$ is the difference in azimuthal angle of the muon candidates at trigger and at particle-flow levels and $\Delta \eta$ represents the difference in pseudorapidity of these two particle candidates. Muon candidates are requested to satisfy the *Tight identification* criterion (see table 6.4).

To suppress the contamination of muons contained in jets, an isolation requirement is imposed to all muon candidates:

$$\texttt{PFIsoCorr} = \left[\sum_{\text{Ch.had.}} p_{\text{T}} + \max\left(0, \sum_{\text{T}} p_{\text{T}} + \sum_{\text{T}} p_{\text{T}} - 0.5 \sum_{\text{T}} p_{\text{T}}\right)\right] / p_{\text{T}}^{\mu} \le 0.2$$

where the sums run over the corresponding particles inside a cone of radius $\Delta R \leq 0.4$, for charged hadrons (Ch.had.), neutral hadrons (N.had.), photons (EM) and charged particles from the pile-up (PU). This corresponds to the standard *Loose isolation* criteria known as the *combined relative particle flow isolation*, corrected for pile-up by applying the so called *Delta Beta correction*. The factor 0.5 corresponds to the naive average ratio of neutral to charged particles and has been measured in jets in [88].

Finally, we restrict the muon kinematic to $p_{\rm T} \ge 20 \,\text{GeV}$ and $|\eta| \le 2.4$.

Table 6.4: List of criteria applied for the Tight identification of muons, where dxy is the transverse impact parameter of the muon track and dz the longitudinal distance with respect to the primary vertex. See [87] for a detailed description of the different criteria.

Variable	Criterion
isGlobal	= 1
isPF	= 1
global track χ^2	< 10
nb. muon hits	> 0
nb. matched stations	> 1
nb. pixel hits	> 0
nb. tracker layers with hits	> 5
dxy (cm)	< 0.2
dz (cm)	< 0.5

The muon selection detailed above provides a high purity sample of isolated muons. This of course comes along with an efficiency decrease. Figure 5.2 page 105 shows the muon identification and isolation efficiencies as measured by the CMS collaboration both on data and MC. As for the muon trigger efficiency, though not as large, the MC predictions overestimate the identification selection efficiency and the corresponding scale factors, shown in table 6.5, need to be applied on the simulation. The isolation efficiency is well described by the MC simulation except for the last $|\eta|$ bins. The MC is also reweighed for that effect and the corresponding numbers are shown in table 6.6. The global efficiency loss induced by identification and isolation criteria will be corrected for by the unfolding method (see section 6.5).

Electrons

126

In a similar way the electron candidates are selected from the *gsfElectron* collection (see section 5.3.2) and required to match the trigger objects. They are additionally requested to satisfy the *Medium identification* criterion as defined by the CMS group in charge of the electron and photon reconstruction (see table 6.7).

An isolation requirement is also contained in the *Medium identification* criterion and imposed to all electron candidates:

$$\texttt{PFIsoRhoCorr} = \left[\sum_{T}^{\text{Ch.had.}} p_{\text{T}} + \max\left(0., \sum_{T}^{\text{N.had.}} p_{\text{T}} + \sum_{T}^{\text{EM}} p_{\text{T}} - \rho \text{ EA}\right)\right] / p_{\text{T}}^{e} \le 0.15$$

where the sums run over the corresponding particles inside a cone of radius $\Delta R \leq 0.3$. The term ρ EA represents a correction for pile-up effects, where ρ , the level of diffuse

Scale Factors for Muons Tight ID				
	$0 < \eta \le 0.9$	$0.9 < \eta \le 1.2$	$1.2 < \eta \le 2.1$	$2.1 < \eta \le 2.4$
$20 < p_{\rm T} \le 25$	0.989 ± 0.002	0.987 ± 0.003	1.002 ± 0.001	
$25 < p_{\rm T} \le 30$	0.994 ± 0.001	0.994 ± 0.001	0.999 ± 0.001	
$30 < p_{\rm T} \le 35$	0.994 ± 0.001	0.991 ± 0.001	0.998 ± 0.001	
$35 < p_{\rm T} \le 40$	0.994 ± 0.000	0.990 ± 0.001	0.997 ± 0.001	
$40 < p_{\rm T} \le 50$	0.992 ± 0.000	0.990 ± 0.000	0.997 ± 0.000	0.004 ± 0.001
$50 < p_{\rm T} \le 60$	0.991 ± 0.001	0.991 ± 0.001	0.998 ± 0.001	0.994 ± 0.001
$60 < p_{\rm T} \le 90$	0.990 ± 0.001	0.986 ± 0.002	0.994 ± 0.002	
$90 < p_{\rm T} \le 140$	1.004 ± 0.003	1.012 ± 0.007	1.019 ± 0.006	
$140 < p_{\rm T} \le 300$	1.028 ± 0.019	0.956 ± 0.034	1.016 ± 0.034	
$300 < p_{\rm T}$	1.000 ± 0.162	1.000 ± 0.255	0.609 ± 0.415	

Table 6.5: Muon Identification scale factors (Data/MC) for Tight ID.

noise, corresponds to the amount of transverse momentum added to the event per unit area. The effective areas (EA) used for this correction have been estimated with 2012 data by the CMS group in charge of the electron and photon reconstructions.

Finally, we restrict the electron kinematic to $p_{\rm T}$ > 20 GeV and $|\eta_{SC}| \leq 1.442$ or

Scale Factors for Tight Muons CombRelIsodBeta < 0.2 (R=0.4)				
	$0 < \eta \le 0.9$	$0.9 < \eta \le 1.2$	$1.2 < \eta \le 2.1$	$2.1 < \eta \le 2.4$
$20 < p_{\rm T} \le 25$	0.987 ± 0.001	0.992 ± 0.002	1.000 ± 0.001	
$25 < p_{\rm T} \le 30$	1.001 ± 0.001	1.004 ± 0.001	1.006 ± 0.001	
$30 < p_{\rm T} \le 35$	0.999 ± 0.000	1.002 ± 0.001	1.005 ± 0.000	
$35 < p_{\rm T} \le 40$	0.999 ± 0.000	1.001 ± 0.000	1.002 ± 0.000	
$40 < p_{\rm T} \le 50$	0.998 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.028 ± 0.000
$50 < p_{\rm T} \le 60$	0.999 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.028 ± 0.000
$60 < p_{\rm T} \le 90$	1.001 ± 0.000	1.000 ± 0.000	1.001 ± 0.000	
$90 < p_{\rm T} \le 140$	1.001 ± 0.001	0.999 ± 0.001	1.000 ± 0.001	
$140 < p_{\rm T} \le 300$	1.001 ± 0.001	1.000 ± 0.002	1.002 ± 0.002	
$300 < p_{\rm T}$	1.011 ± 0.000	1.000 ± 0.016	1.000 ± 0.012	

Table 6.6: Muon Isolation scale factors (Data/MC).

Table 6.7: List of criteria applied for the Medium identification of electrons where 1/E - 1/p represents the difference between the inverse of the supercluster energy and the inverse of the track momentum, vtxFit required the track to be used for the vertex fit and mHits concerns the number of missed hit in the electron track. See [89] for a detailed description of the different criteria.

Variable	Barrel Criterion	Endcap Criterion
$\Delta \eta_{\rm in}$	< 0.004	< 0.007
$\Delta \phi_{ m in}$	< 0.060	< 0.030
$\sigma_{i\eta i\eta}$	< 0.010	< 0.030
H/E	< 0.120	< 0.100
$1/E - 1/p \ (1/\text{GeV})$	< 0.050	< 0.050
dxy (cm)	< 0.020	< 0.020
dz (cm)	< 0.100	< 0.100
vtxFit	true	true
mHits	≤ 1	≤ 1

 $1.566 \leq |\eta_{SC}| \leq 2.4$ where η_{SC} is the corresponding supercluster η .

Similarly to the muon case, these identification criteria increasing the purity of the electron sample also decrease the efficiency of the selection. The data and MC efficiencies were measured using the same tag and probe method. The resulting scale factors applied on the MC predictions are listed in table 6.8.

6.3.3 Z bosons

The Z boson candidates are reconstructed from the lepton pair (muons or electrons) with the two highest transverse momenta, the four-momentum vector of the Z boson being obtained as the sum of the two charged lepton four-momenta:

$$p_Z = p_{l_1} + p_{l_2}.$$

The Z boson candidate reconstructed in this way must have an invariant mass between 71 and 111 GeV for the event to be selected.

6.3.4 Jets

The input to the CMS jet clustering algorithm are the four-momentum vectors of the collection of particles reconstructed using the PF technique from which all charged PF hadron candidates associated to a vertex different than the one corresponding to the Z boson are removed. This last step is called *Charged Hadron Subtraction* (CHS) [90] and results in keeping all charged particles that originate from the Z boson vertex or that cannot be associated to any other primary vertex satisfying the following selection criteria:

• requirement on the z component of Primary Vertex to be: z < 24 cm,

Table 6.8: Electron Identification scale factors (Data/MC) for Medium ID and Isolation.

Scale Factors for Medium Electrons ID and Iso				
	$0. < \eta \le 0.8$	$0.8 < \eta \le 1.442$	$1.566 < \eta \le 2.0$	$2.0 < \eta \le 2.5$
$20 < p_{\rm T} \le 30$	1.010 ± 0.003	0.981 ± 0.006	0.992 ± 0.009	1.045 ± 0.005
$30 < p_{\rm T} \le 40$	1.006 ± 0.000	0.987 ± 0.000	0.993 ± 0.000	1.031 ± 0.000
$40 < p_{\rm T} \le 50$	1.009 ± 0.001	0.993 ± 0.001	1.008 ± 0.000	1.019 ± 0.000
$50 < p_{\rm T}$	1.008 ± 0.002	0.995 ± 0.001	1.009 ± 0.000	1.014 ± 0.001

- requirement on the distance from the beam axis to be: $\rho < 2$ cm,
- requirement on the number of degrees of freedom of the vertex fit (number of hits used in order to reconstruct the tracks belonging to the vertex): n.d.f. > 4,

as recommended by CMS group in charge of the jet reconstruction [91]. This CHS procedure accounts for roughly half of the offset energy from pile-up in the tracker-covered region. Only the remaining offset energy has to be subtracted using the jet area method. Jets are reconstructed using the anti- k_t clustering algorithm [73], with a size parameter of R = 0.5, by summing the four-momenta of individual PF particles according to the FASTJET package [92]. Well identified muons happen to satisfy the jets requirements at this point. Therefore, once reconstructed, the jets overlapping (within $\Delta R = 0.5$) with any of the two leptons coming form the decay of the Z boson are removed from the jet collection.

A certain number of selection criteria concerning the jet composition is applied in order to avoid misidentification and to increase noise rejection. This list of criteria (see table 6.9) constitutes the *Loose identification* criterion.

Reconstructed jets transverse momenta are corrected in the simulated DY sample in order to increase the simulation performance with respect to what is observed in data [75]. To this aim a correction depending on $|\eta|$ is applied on the jet energy and transverse momentum using the following formulae:

$$p_{\rm T}^{\rm corrected} = \max\left(0, \ p_{\rm T}^{\rm gen} + C(|\eta|) \times (p_{\rm T}^{\rm reco} - p_{\rm T}^{\rm gen})\right),$$
$$E^{\rm corrected} = \frac{p_{\rm T}^{\rm reco}}{p_{\rm T}^{\rm corrected}} E^{\rm reco},$$

where $p_{\rm T}^{\rm reco}$ and $p_{\rm T}^{\rm gen}$ are respectively the reconstructed and generated transverse momentum of a jet and the correction factors $C(|\eta|)$ are given in table 6.10 (this correction comes together with a systematic uncertainty for which up and down variations are considered).

Variable	Criterion
Neutral Hadron Fraction	< 0.99
Neutral EM Fraction	< 0.99
Number of Constituents	> 1
Charged Hadron Fraction	> 0
Charged Multiplicity	> 0
Charged EM Fraction	< 0.99

Moreover, a threshold of 30 GeV on jets $p_{\rm T}$ is applied to reduce the pile-up contamination as well as large uncertainty on the energy measurement. In order to further reduce the pile-up contamination, we apply a *Loose* selection criterion on the pile-up jet identification variable which uses vertex and shape variables to determine if a jet comes from pile-up or not [93]. To ensure a good quality of the tracking information in the whole area of the jets, jets with $|\eta| \ge 2.4$ are removed from the collection. Finally, the jet collection is ordered by decreasing $p_{\rm T}$ values.

At particle level jets are clustered from MC stable particles after hadronisation and removal of neutrinos as well as the attached FSR radiative photons in a cone of 0.1 opening. Again, the anti- k_t clustering algorithm with cone size of R = 0.5 is used at this level.

6.3.5 Pile-up

Pile-up Reweighing

The simulated samples have been generated with a given distribution of number of pile-up interactions which is meant to roughly cover the conditions expected for each data-taking period. Of course the match between produced simulations and data cannot be exact as the run selection depends on the analysis and that the exact run conditions are not known when producing the simulated samples. To improve the agreement to the data the prescription for pile-up re-weighting described in the reference [94] has been applied. The method uses the number of pile-up interactions from the simulation truth to compute the weights.Naturally the problem is then to determine the target pile-up distribution. This is derived from data by using the measured instantaneous luminosity together with

	$C(\eta)$	$C^{\mathrm{up}}(\eta)$	$C^{\mathrm{down}}(\eta)$
$0.0 \le \eta < 0.5$	1.079	1.105	1.053
$0.5 \le \eta < 1.1$	1.099	1.127	1.071
$1.1 \le \eta < 1.7$	1.121	1.150	1.092
$ 1.7 \le \eta < 2.3$	1.208	1.254	1.162
$2.3 \le \eta < 2.8$	1.254	1.316	1.192
$2.8 \le \eta < 3.2$	1.395	1.458	1.332
$3.2 \le \eta $	1.056	1.247	0.865

Table 6.10: Correction factors for jet transverse momenta and energies.


Figure 6.2: Simulation to data comparison at detector level of the number of reconstructed primary vertices after pile-up reweighing using: 0.95% (left), 100% (middle) and 105% (right) of the minimum bias cross section.

the total *pp* inelastic cross section (also known as the minimum bias cross section) leading to an expected pile-up distribution. An event by event weight depending on the generated number of pile-up interactions is therefore obtained and applied on the simulated samples.

The result is shown in figure 6.2 from which it can be seen that the $\pm 5\%$ pp total inelastic cross section uncertainty completely covers the last bit of disagreement between data and simulation.

Remanent Pile-up Contribution

As discussed in section 6.3.4, the pile-up contribution is reduced using the CHS method. Nevertheless one should inquire a possible residual effect of pile-up on, *e.g.* the jet multiplicity. To quantify this possible effect, figure 6.3 shows the average number of jets per reconstructed Z boson for different pile-up conditions for all the events, for the events containing not more than 10, between 11 and 14, between 15 and 18, and at least 19 reconstructed vertices. As can be seen, the effect is not significant. The check was also done by looking at the jet multiplicity distributions for different number of vertices as shown on figure 6.4. There is no significant dependence on the number of vertices, hence nor on the amount of pile-up.

Additionally, the true pile-up distribution from generated MC information, after the selection, shown in figure 6.5, does not show any dependency in jet multiplicity. Possible additional effects of the pile-up can therefore be ignored in the following.



Figure 6.3: Measured average number of jets per reconstructed Z boson for multiplicity above or equal one for different pile-up conditions in the muon channel. The first bin represents the average over the full samples. The other bins are limited to the given range of reconstructed vertices. (left) Jet p_T threshold is 30 GeV (right) jet p_T threshold is 50 GeV.



Figure 6.4: Measured jet inclusive multiplicity for different ranges of number of vertices in (left) muon and (right) electron decay channels. Each distribution is normalised such that the first bin content value is one. No dependency on the number of vertices, hence on pile-up conditions, is observed.

6.4 Detector Level Results

The events passing the selection requirements, two well identified oppositely charged leptons with a transverse momentum above 20 GeV, a pseudorapidity within 2.4 and an invariant mass compatible with the Z mass peak are analysed by filling histograms for



Figure 6.5: *MC* true pile-up distribution for different jet multiplicities (left). *MC* reconstructed number of vertices distribution for different jet multiplicities (right) in the electron decay channel. Each distribution is normalised to unity. No dependency on the number of jets is observed.

various observables. This is done for the data and all the simulated samples listed in table 6.2, separately for the two decay channels.

6.4.1 Data-Simulation Comparison

For each of the observables, it is interesting to look at the comparison between what is observed in the data and what is predicted by the simulation. To this aim, and for each histogram, the sum of all background simulation predictions and the signal simulation prediction, normalised to the data integrated luminosity is computed and stored into a stack of histograms so that each process simulated contribution is visible. The data histograms are then superimposed on the MC stacked histograms. The errors bars on the shown histograms are data statistical uncertainties only. The full uncertainty treatment is applied for the cross section measurements and explained in section 6.6. For each observable, the MC prediction over data ratio is presented to quantify possible disagreements. The corresponding error bars correspond to statistical uncertainties coming from both the data and the simulated finite number of events.

As shown in the next sections, at this level it can already be noticed that no significant difference is observed between the two decay channels apart from a global efficiency factor (muons are more easily triggered, identified and isolated than electrons which results in a lower global selection inefficiency). Therefore the remarks made in the next paragraphs are valid for both decay channels.

General Control plots

Figure 6.6 page 138 shows the dilepton invariant mass distributions for the selected events without any requirement on the number of reconstructed jet. The shape of the ratio comparison is understood and due to imperfect lepton momentum measurements. Correcting for these effects is important in precise measurements of differential cross sections directly related to the lepton kinematics but for the present work focussed on jet kinematics measurements, this is of no influence as we will integrate over this distribution between 71 and 111 GeV.

However, for the sake of completeness, it has been verified that applying a momentum scale correction to remove a bias in the reconstructed muon momenta due to the differences in the tracker misalignment between data and simulation and the residual magnetic field mismodeling, following the standard CMS procedure described in [95], has no effect on jets related distributions. This correction is therefore not applied for this analysis. Analog ECAL electron energy deposits corrections were developed in [96] but are not applied in this work.

From the simulation histogram different colours it can be seen that, for the inclusive Z + X case, the background contamination is very small: less than percent level around the Z mass peak and a few percents in the tails of the distribution. Additionally, we can distinguish the backgrounds which lead to a real Z boson resonance in the final state from those without resonance such as the $t\bar{t}$ and single top processes.

Figures 6.7, 6.8 and 6.9 respectively show the leptons pseudorapidity, azimuthal angle and transverse momentum for all selected events without any requirement on the number of reconstructed jet. An excellent description is reached by the simulation. The drops in the electron channel pseudorapidity distribution around $|\eta| = 1.5$ are due to the transition region between the barrel and endcap ECAL from which reconstructed electrons are excluded by the criterion $1.442 \leq |\eta| \leq 1.566$. From the azimuthal distributions, the expected symmetry of the produced leptons is indeed observed and well described by the simulation. Finally, the lepton transverse momentum histograms exhibit a peak at a value around 45 GeV related to the Z boson mass. Most of the time the two leptons will carry away half of the Z mass energy in their $p_{\rm T}$. The 20 GeV threshold is necessary in order to reach the trigger efficiency plateau and therefore avoid large trigger efficiency uncertainties. Except for a slight overestimation of the MC yields at larger muon $p_{\rm T}$ which is at the limit of compatibility with the systematic uncertainties related to the muon scale factors, the agreement is very good.

On page 140, figures 6.10 and 6.11 respectively display the Z boson pseudorapidity for events without any requirement on the number of jets and for events with at least one jet above the 30 GeV threshold. The agreement in both cases is excellent and shows the expected symmetry between forward and backward production. From figure 6.11 it can be seen that the $p_{\rm T}$ constrain on the reconstructed jet is directly affecting the Z boson η distribution by making it produced more centrally in the CMS detector.

On page 141, figures 6.12 and 6.13 respectively display the Z boson $p_{\rm T}$ for events without any requirement on the number of jets and for events with at least one jet above the 30 GeV threshold. One more time, from figure 6.13 the reconstructed jet $p_{\rm T}$ constrain leads to a peak in the Z boson transverse momentum distribution located around the jet $p_{\rm T}$ threshold. A globally good agreement is again observed apart from a slight slope in the ratio as the Z $p_{\rm T}$ increases. This pattern, even in the inclusive $N_{jets} \geq 0$ case for which jets with $p_{\rm T}$ values below the threshold of 30 GeV may exist, is directly linked to the disagreement in the jets transverse momentum distribution and will be discussed later in this chapter.

The possible disagreement observed for the next observables will be discussed in section 6.8.2 after detector effect correction, selection inefficiency correction and full treatment of the systematic uncertainties.

Jet Multiplicities

For the control plot relative to the jet multiplicity, the requirement concerning a minimum of one jet to be reconstructed in the event has been released.

An excellent agreement is found between the data and the MC prediction for the exclusive and inclusive jet multiplicities up to 4 jets (see figure 6.14 page 143 to figure 6.15 page 143). This is in accordance with the fact the signal sample was generated by the tree level MADGRAPH 5 MC generator providing Z plus up to four partons, interfaced with PYTHIA 6 for the parton shower and rescaled to the NNLO inclusive cross section. The 5^{th} jet exclusive multiplicity is underestimated by 8% and 14% respectively for the electron and muon decay channel. The distributions show up to 7 jets. The number of observed events with 6 jets is 154 and 224 respectively for the electron and the muon decay channel and 24 and 33 for the 7-jets bin. Number of events in each bin of exclusive jet multiplicity for both data and MC signal and backgrounds are listed in table 6.11 for the muon decay channel and in table 6.12 for the electron decay channel. From those tables background contamination is observed to be above the percent level for jet multiplicities above one. The dominant contribution comes from the $t\bar{t}$ process which varies between 2 and 12% for jet multiplicities between 2 and 5, respectively. Other background contributions are at least three times smaller in presence of at least one jet. When comparing other following observables, one should keep in mind the corresponding

jet multiplicity yields which gives the average ratio of simulation to data. For example, when looking at the fifth leading jet distribution, the mean discrepancy of about 15% directly comes from the total number of events with at least five reconstructed jets already seen in the jet multiplicity distribution.

Jets Transverse Momenta

For these comparisons, the jet $p_{\rm T}$ threshold has been lowered from 30 to 20 GeV in order to provide two additional bins below the threshold used to improve the detector effects corrections by including migrations from low- $p_{\rm T}$ jets. The $p_{\rm T}$ distributions (see figure 6.16 page 144 to figure 6.20 page 146) are falling a bit more rapidly in the data than in the MC for the leading jet in $N_{\rm jets} \geq 1$ and for the second jet for $N_{\rm jets} \geq 2$. The statistical precision is not high enough to say if the tendency is the same for larger multiplicities. Small deviations are observed at $p_{\rm T}$ values below the threshold of 30 GeV applied in the analysis for jet multiplicities above one. As announced, these two lower $p_{\rm T}$ to higher- $p_{\rm T}$ bins and therefore arise as correction factors only. However, it has been carefully checked that this issue does not come from a potential pile-up misestimation by making the same plots for different ranges of number of reconstructed vertices.

Jets Pseudorapidities

The jet η distributions show a global good agreement between the data and the MC (see figure 6.21 page 147 to figure 6.25 page 149). A lack of MC events around 10% is found for $|\eta| > 2.1$ in both decay channels.

Jets Transverse Momenta Scalar Sum

Faster falling distributions in the data with respect to the MC are also observed in the $H_{\rm T}$ distributions (see figure 6.26 page 150 to figure 6.30 page 152) for inclusive jet multiplicities up to 3, but in a less pronounced way than in the $p_{\rm T}$ spectra. For jet multiplicities above 2, at the low $H_{\rm T}$ values close to the threshold of 90 GeV, the MC underestimates the number of events by an amount increasing with the jet multiplicity.

6.4. DETECTOR LEVEL RESULTS

Dijet Mass

The dijet invariant mass distributions exhibit a lack of events in the MC for an invariant dijet mass below 50 GeV whilst a global good agreement is found for higher masses (see figure 6.31 at page 153). It has been checked that the low dijet mass discrepancy was not due to bad pile-up jets contamination by repeating the measurement for three regions in the number of reconstructed primary vertices: the low pile-up region for which $\sharp Vtx \leq 14$, the medium pile-up region with $14 < \sharp Vtx \leq 18$ and the high pile-up region $18 < \sharp Vtx$. No significant dependence was observed.



Figure 6.6: Detector level data - simulation comparison of dilepton invariant mass with $N_{jets} \geq 0$ for (left) muon and (right) electron decay channels.



Figure 6.7: Detector level data - simulation comparison of lepton pseudorapidity with $N_{jets} \ge 0$ for (left) muon and (right) electron decay channels.



Figure 6.8: Detector level data - simulation comparison of lepton azimuthal angle with $N_{jets} \ge 0$ for (left) muon and (right) electron decay channels.



Figure 6.9: Detector level data - simulation comparison of lepton transverse momentum with $N_{jets} \ge 0$ for (left) muon and (right) electron decay channels.



Figure 6.10: Detector level data - simulation comparison of Z boson pseudorapidity with $N_{jets} \geq 0$ for (left) muon and (right) electron decay channels.



Figure 6.11: Detector level data - simulation comparison of Z boson pseudorapidity with $N_{jets} \geq 1$ for (left) muon and (right) electron decay channels.



Figure 6.12: Detector level data - simulation comparison of Z boson transverse momentum with $N_{jets} \ge 0$ for (left) muon and (right) electron decay channels.



Figure 6.13: Detector level data - simulation comparison of Z boson transverse momentum with $N_{jets} \ge 1$ for (left) muon and (right) electron decay channels.

Table 6.11: Number of events in data and MC samples for exclusive jet multiplicity in muon channel after full selection.

	$N_{\rm jets} = 0$	$N_{\rm jets}=1$	$N_{\rm jets}=2$	$N_{\rm jets}=3$	$N_{\rm jets} = 4$	$N_{\rm jets}=5$	$N_{\rm jets}=6$	$N_{\rm jets}=7$
WJets	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DYTauTau	2075.6	300.8	77.0	14.6	0.0	0.0 0.0		0.0
ZZJets4L	130.7	150.4	107.6	29.4	7.0	1.5	0.3	0.0
Single Top	127.2	410.4	196.8	60.1	9.7	0.9	1.3	0.0
ZZJets2L2Nu	1185.5	332.1	89.7	19.9	3.8	0.6	0.2	0.0
ZZJets2L2Q	719.6	2088.2	1803.9	511.2	130.7	26.4	$\begin{array}{c} 5.7 \\ 0.0 \end{array}$	$\begin{array}{c} 0.7 \\ 0.0 \end{array}$
WWJets2L2Nu	1112.0	362.0	107.3	22.6	4.5	0.4		
WZJets2L2Q	712.2	1387.1	1193.0	522.8	167.4	39.2	8.2	1.3
WZJets3LNu	727.2	774.1	315.8	81.8	18.9	3.2	0.8	0.1
TTJets	312.8	2071.4	3933.1	1840.2	570.9	142.7	35.0	9.2
DYJets	5903239.3	908971.9	178530.9	32043.0	5727.1	908.3	146.5	19.8
TOTAL	5910342.1	916848.4	186354.9	35145.6	6640.0	1123.1	198.0	31.0
Data	5902750	913503	186087	35718	6751	1301	224	33
Ratio	1.0013	1.0037	1.0014	0.9840	0.9836	0.8633	0.8838	0.9409

Table 6.12: Number of events in data and MC samples for exclusive jet multiplicity in electron channel after full selection.

	$N_{\rm jets} = 0$	$N_{\rm jets} = 1$	$N_{\rm jets}=2$	$N_{\rm jets}=3$	$N_{\rm jets} = 4$	$N_{\rm jets}=5$	$N_{\rm jets}=6$	$N_{\rm jets}=7$	
WJets	1016.5	801.9	222.7	38.4	0.0	0.0	0.0	0.0	
DYTauTau	1392.2	247.9	49.6	13.8	0.0	0.0	0.0	0.0	
ZZJets4L	86.4	110.1	81.2	22.5	5.5	1.1	0.2	0.0	
Single Top	67.7	301.3	131.0	38.0	9.5	3.2	0.0	0.0	
ZZJets2L2Nu	810.3	239.0	65.0	13.8	2.9	0.5	0.0	0.0	
ZZJets2L2Q	486.0	1488.7	1284.8	371.9	93.1	19.9	3.9	0.7	
WWJets2L2Nu	745.8	246.5	77.0	16.5	3.2	0.9	0.3	0.1	
WZJets2L2Q	480.4	977.0	854.2	381.1	123.3	30.0	6.5	1.2	
WZJets3LNu	490.6	567.6	237.5	62.3	13.7	2.9	0.4	0.1	
TTJets	222.7	1531.3	2772.8	1286.7	447.0	122.2	34.3	2.2	
DYJets	3843135.3	619067.1	124236.0	22649.6	4106.4	675.3	101.8	13. 0	
TOTAL	3848934.0	625578.3	130011.9	24894.4	4804.8	856.1	147.3	17.3	
Data	3882357	635598	132659	25704	4936	929	154	24	
Ratio	0.9914	0.9842	0.9800	0.9685	0.9734	0.9215	0.9565	0.7202	



Figure 6.14: Detector level data - simulation comparison of exclusive jet multiplicity for (left) muon and (right) electron decay channels.



Figure 6.15: Detector level data - simulation comparison of inclusive jet multiplicity for (left) muon and (right) electron decay channels.



Figure 6.16: Detector level data - simulation comparison of the 1st leading jet transverse momentum with $N_{jets} \geq 1$ for (left) muon and (right) electron decay channels.



Figure 6.17: Detector level data - simulation comparison of the 2^{nd} leading jet transverse momentum with $N_{jets} \geq 2$ for (left) muon and (right) electron decay channels.



Figure 6.18: Detector level data - simulation comparison of the 3^{rd} leading jet transverse momentum with $N_{jets} \geq 3$ for (left) muon and (right) electron decay channels.



Figure 6.19: Detector level data - simulation comparison of the 4th leading jet transverse momentum with $N_{jets} \ge 4$ for (left) muon and (right) electron decay channels.



Figure 6.20: Detector level data - simulation comparison of the 5th leading jet transverse momentum with $N_{jets} \geq 5$ for (left) muon and (right) electron decay channels.

146



Figure 6.21: Detector level data - simulation comparison of the 1st leading jet pseudorapidity with $N_{jets} \ge 1$ for (left) muon and (right) electron decay channels.



Figure 6.22: Detector level data - simulation comparison of the 2^{nd} leading jet pseudorapidity with $N_{jets} \ge 2$ for (left) muon and (right) electron decay channels.



Figure 6.23: Detector level data - simulation comparison of the 3^{rd} leading jet pseudorapidity with $N_{jets} \geq 3$ for (left) muon and (right) electron decay channels.



Figure 6.24: Detector level data - simulation comparison of the 4th leading jet pseudorapidity with $N_{jets} \ge 4$ for (left) muon and (right) electron decay channels.



Figure 6.25: Detector level data - simulation comparison of the 5th leading jet pseudorapidity with $N_{jets} \geq 5$ for (left) muon and (right) electron decay channels.



Figure 6.26: Detector level data - simulation comparison of the scalar sum of the jets transverse momentum with $N_{jets} \ge 1$ for (left) muon and (right) electron decay channels.



Figure 6.27: Detector level data - simulation comparison of the scalar sum of the jets transverse momentum with $N_{jets} \ge 2$ for (left) muon and (right) electron decay channels.



Figure 6.28: Detector level data - simulation comparison of the scalar sum of the jets transverse momentum with $N_{jets} \geq 3$ for (left) muon and (right) electron decay channels.



Figure 6.29: Detector level data - simulation comparison of the scalar sum of the jets transverse momentum with $N_{jets} \ge 4$ for (left) muon and (right) electron decay channels.



Figure 6.30: Detector level data - simulation comparison of the scalar sum of the jets transverse momentum with $N_{jets} \ge 5$ for (left) muon and (right) electron decay channels.

152



Figure 6.31: Detector level data - simulation comparison of the two leading jets invariant mass with $N_{jets} \ge 2$ for (left) muon and (right) electron decay channels.

6.4.2 Background Subtraction

At this level, the black data points shown in the previous comparisons correspond to the sum of a signal contribution, *i.e.* real Z boson events decaying into two oppositely charged leptons accompanied with possible jets, and a background contribution coming from different other processes leading to the same reconstructed signature in the CMS detector.

In order to estimate the signal contribution in the data at reconstructed level, the background simulation predictions are subtracted, bin by bin, from the observed data histograms. The subtraction is obviously dependent on the cross section of each background process and a systematic uncertainty will therefore arise from the limited knowledge of the background cross sections (see section 6.6).

From now on, the signal data refers to the observed data (black points in the previous histograms) from which background MC estimated contributions have been subtracted.

6.5 Detector Effect Corrections

The measured observables of interest, presented in the previous section, differ from their true values due to detector effects. For example, a jet could be produced during the interaction with some true transverse momentum $p_{\rm T}^{\rm true}$, but due to finite detector resolution, it could be reconstructed with transverse momentum $p_{\rm T}^{\rm reco} \neq p_{\rm T}^{\rm true}$. The same can happen for all other reconstructed quantities such as lepton transverse momenta, pseudorapidities, ... The global result on the observed distributions is a broadening of the true distributions. To be able to compare measured observables to theoretical predictions detector level distributions have to be corrected for these effects. This is done by an unfolding procedure. The first thing that has to be decided is what we really mean by the generator level *i.e.* what cross section we want to extract and in what phase space.

6.5.1 Phase Space at Generator Level

Using the MADGRAPH 5 generator, generated leptons four-momenta can be accessed before (flag status 3) or after (flag status 1) QED FSR simulation thanks to the corresponding flag provided by the MC. This is however not the case for SHERPA where only the status 1 generated leptons can be retrieved. In order to be able to compare results between different generators, the status 1 generated leptons are used in the unfolding procedure. To reduce the effect of FSR on the Z boson kinematics and invariant mass, leptons four-momenta are corrected for hard photon radiation candidates: $p_{\text{corr.}}^l = p^l + \sum_{\gamma} p^{\gamma}$, where the generated photon must lie inside a cone of radius $\Delta R \leq 0.1$ around the lepton candidate. Dressed generated muons and electrons are found to give the same invariant mass distribution as shown in figure 6.33.

This cross section is, however, not fully equivalent to the one with the kinematics of leptons before FSR (status 3 in MADGRAPH) as shown in figure 6.32 which displays the comparison of the dilpeton invariant mass for dressed leptons with different cone sizes, and pre-FSR status 3 leptons. For the muon channel, the effect only affects the normalisation on jet related measurements and is estimated from the MADGRAPH 5 sample to be of 2% less than with pre-FSR muons as illustrated in figure 6.34 for the leading jet transverse momentum distribution. For the electron channel however, a nontrivial effect is observed at low values in the jet transverse momentum distribution as shown in figure 6.34. The effect ranges from 4% to 2% and is likely to come from the following two facts. First the correction on the electron transverse momenta is smearing the Z boson transverse momentum distribution. Secondly the correlation between the jet $p_{\rm T}$ and the Z boson $p_{\rm T}$ leads and the transverse momentum threshold applied on the jet kinematics result in a lower physics acceptance at low Z $p_{\rm T}$ values that is also visible



Figure 6.32: Z boson mass at generator level for dressed post-FSR leptons for different cone sizes for $\mu^+\mu^-$ (top left) and e^+e^- (top right) decay channels. Comparison to mass distribution from pre-FSR leptons is shown for $\mu^+\mu^-$ (bottom left) and e^+e^- (bottom right) decay channels, by including radiated photons in a cone of different ΔR around the leptons.



Figure 6.33: Z boson mass at generator level for dressed post-FSR muons and electrons with cone sizes $\Delta R = 0.1$ (left). Ratio of Z mass from electrons to muons is also shown (right).

at low jet transverse momenta. The correction being much less for the muon channel, no such effect is observed. For completeness, a table with the ratios as a function of the number of jets is given in table 6.13.

6.5.2 The Unfolding Method

To extract the cross section an unfolding procedure is applied. This general procedure is MC dependent since one has to create a mapping between the true value (generated) of an observable and the reconstructed (measured) value distorted due to detector effects such

$N_{\rm jets}$	muon channel	electron channel
= 0	98.0%	98.1%
= 1	98.0%	97.0%
= 2	98.0%	96.8%
= 3	98.1%	96.8%
=4	98.0%	96.8%
= 5	97.8%	96.8%

Table 6.13: Ratios of event yields for dressed post-FSR leptons over pre-FSR leptons as a function of the number of jets.



Figure 6.34: Ratio of generated jet transverse momentum distribution for (blue) dressed post-FSR muons over pre-FSR muons and (red) dressed post-FSR electrons over pre-FSR electrons.

as finite resolution and limited acceptance. This mapping is contained in the response matrix for which each element (i, j) is related to the probability that the observable, generated in the i^{th} bin, would be measured in the j^{th} bin. Given this response matrix, one can solve the problem by inverting and applying its inverse on measured data distributions, resulting in data distributions at particle level. Simple inversion of the matrix as well as a procedure trying to overcome the instability related to the inversion of the response matrix, based on its Singular Value Decomposition (SVD) and described in [97], can be used via the RooUnfold framework [98]. Another way of treating the unfolding problem which does not require the inversion of the response matrix and thus avoids any trouble encountered with matrix inversion is based on Baye's theorem and described in [99]. Finally, one can use the so called bin-by-bin correction which uses generalised efficiency based on MC simulation to estimate the number of true events from the number of events observed in a particular bin. This last method is however not ideal in situations where correlations between bins cannot be neglected.

In this analysis, the Baye's method⁴ is used to obtain the cross section at particle level. The method has been compared to the simple inversion, bin-by-bin correction and SVD procedures. The choice for the Baye's method has been lead by the drawbacks of each of

⁴In the RooUnfold framework the prior distribution used for the Baye's method corresponds to the MC truth while in the reference [99] a uniform distribution is used. It should also be noted that the expression for the covariance matrix in [99] is wrong and should be replaced according to [98], the correct one being used in RooUnfold.

the alternative methods. Simple inversion is not generally suitable since it suffers from large instabilities with respect to small fluctuations. The bin-by-bin correction cannot be used in every cases since, as mentioned above, correlations between bins are not taken into account by this method. Lastly, the SVD method could in principle be used but would require, as raised in [97], a MC sample with a statistic one or two order of magnitude larger than the one of data, which is not at our disposal at this time. Moreover, due to technical methods used to compute the covariance matrix, it is not possible to use variable bin widths and get a correct covariance matrix from the implementation of the SVD method in the RooUnfold framework.

It has to be noted that the SVD methods depends on the choice of a regularisation parameter which is needed since matrix inversion is sensitive to the statistical fluctuations [98]. The choice of the regularisation parameter is in the range $[1, n_{\text{bins}}]$. The optimal value of the regularisation parameter is the one for which the errors associated to the unfolding procedure are small when compared to statistical ones. This parameter can be seen as a cutoff for quickly oscillating terms corresponding to data statistical fluctuations. By choosing a too small regularisation parameter one gets ride of these spurious fluctuations but the result tends to be biased by the MC truth. On the other hand, a too large regularisation parameter will decrease the MC dependence but give too large importance too data fluctuation which will be interpreted as real shape. The Baye's method using an iterative approach also requires a choice for the maximum number of iteration to be done. As discussed in [99] this should not lead to different results once the convergence has been reached. Details concerning the Baye's method are given later in this chapter.

6.5.3 Response Matrices Overview

The MC sample used for the reconstruction of the response matrix is the MADGRAPH 5 DY + jets sample. The true values refer to the quantities calculated with the generatorlevel objects which in this case are stable leptons (after final state radiation) from the decay of the Z boson dressed with all the photons that are within the cone of radius 0.1 and the generator-level jets. With dressed leptons constrained to the same $p_{\rm T}$ and η selection criteria, one reconstructs the generated Z boson candidate and applies the Z mass requirement as in measured data as described in section 6.3. The generator level hadron jets are clustered with anti- k_t jet clustering algorithm with a cone size of 0.5. Photons and stable leptons that form the dressed leptons and neutrinos are removed from the input collection of the clustering algorithm. With this generator level quantities and the same reconstructed level quantities, it is possible to fill the response matrix event by event and so reconstruct and estimate of the detector response. It has to be noted that the choice of a good binning is important in order to limit migrations from one bin to others and to ensure that the correction applied to the measured quantities at reconstructed level, inside a particular bin, is based on events coming mainly from the same bin at generator level. The binning has been chosen to keep a large enough number of events inside each bin whenever possible, and trying to have more than 50% of the reconstructed events coming from the same bin at generator-level (except for the jet multiplicity where no choice for the binning is possible). This can be seen on the diagonal of each response matrix where each row (generated) has been normalised to unity (taking the underflow and overflow bins into account).

For the response matrices related to the n^{th} jet p_{T} , two additional bins are presented below the threshold used in the analysis in order to consider the migrations coming from p_{T} values below 30 GeV bins to the phase space we select. Therefore, the threshold has been lowered to 20 GeV in order to fill these response matrices. For the other variables the migration being much smaller, such a consideration is not needed. Still, underflow and overflow bins are taken into account when proceeding to the unfolding.

For illustration purposes, figures 6.35 to 6.40 show examples of response matrices for the variables of interest obtained using both the MADGRAPH 5 and the SHERPA 1.4 DY + jets simulated samples, for the muon decay channel only. Two different MC generators are used in order to estimate the potential difference arising due to a specific choice of response matrix (see section 6.6). Similar response matrices are also obtained for the electron decay channel and are extremely close to the one obtain with the muon decay channel. This is expected since, these response matrices are normalised to each generated bin, and therefore only depend on the detector response to jets, which is of course independent of the decay channel under consideration. The difference is only reflected when absolute number of entries are compared. Since muon reconstruction efficiencies are higher than electron reconstruction efficiencies, the muon decay channel response matrices contain more entries. This effect is directly linked to the ϵ_i factor of equation (6.1).



Figure 6.35: Response matrix for exclusive jet multiplicity built from (left) MADGRAPH 5 and (right) SHERPA 1.4, DY + Jets MC samples for the muon decay channel.



Figure 6.36: Response matrix for 1^{st} jet p_T built from (left) MADGRAPH 5 and (right) SHERPA 1.4, DY + Jets MC samples for the muon decay channel.



Figure 6.37: Response matrix for 2^{nd} jet p_T built from (left) MADGRAPH 5 and (right) SHERPA 1.4, DY + Jets MC samples for the muon decay channel.



Figure 6.38: Response matrix for 3^{rd} jet p_T built from (left) MADGRAPH 5 and (right) SHERPA 1.4, DY + Jets MC samples for the muon decay channel.



Figure 6.39: Response matrix for 2^{nd} jet η built from (left) MADGRAPH 5 and (right) SHERPA 1.4, DY + Jets MC samples for the muon decay channel.



Figure 6.40: Response matrix for jets H_T with $N_{jets} \ge 3$ built from (left) MADGRAPH 5 and (right) SHERPA 1.4, DY + Jets MC samples for the muon decay channel.

6.5.4 Baye's Unfolding

Once the response matrix has been obtained, the unfolding procedure can be performed. The Baye's unfolding method, used in the present work is described in some detail in the following paragraphs.

The Algorithm

In this method, the best estimate of the true number of events $\hat{n}(i)$ inside one particular bin *i* for some distribution is given by

$$\hat{n}(i) = \frac{1}{\epsilon_i} \sum_{j=1}^{n_{\text{bins}}} n_{\text{obs}}(j) S(i_{\text{gen}}|j_{\text{reco}}).$$
(6.1)

In this formula, the term $0 \leq \epsilon_i \equiv \sum_{j=1}^{n_{\text{bins}}} R(j_{\text{reco}}|i_{\text{gen}}) \leq 1$ gives the efficiency of observing an event generated in bin *i*. The sum in the definition of ϵ_i runs over the n_{bins} elements of the *i*th row of the response matrix $R(i_{\text{reco}}|j_{\text{gen}})$ which represents the probability to observe an event generated in bin *j*, in bin *i*. The matrix $S(i_{\text{gen}}|j_{\text{reco}})$ is the smearing matrix which describes the cell-to-cell migration by giving the probability that an event observed in bin *j* was generated in bin *i*. The smearing matrix is related to the response matrix by the following relation:

$$S(i_{\text{gen}}|j_{\text{reco}}) = \frac{R(j_{\text{reco}}|i_{\text{gen}})P_0(i_{\text{gen}})}{\sum_{l=1}^{n_{\text{bins}}} R(j_{\text{reco}}|l_{\text{gen}})P_0(l_{\text{gen}})},$$

in which $P_0(i_{\text{gen}})$ indicates the initial probability for the event to be generated in bin *i*. The iteration in the Baye's method refers to the estimation of the factors $P_0(i_{\text{gen}})$. Different choices can be made for the first estimation of $P_0(i_{\text{gen}})$ such as a flat distribution or the generated MC distribution. In our analysis the second choice has been made. Finally, $n_{\text{obs}}(j)$ gives the number of signal events observed in bin *i*, *i.e.* the number of data events observed in bin *i*, minus the total estimated number of background events and minus the total estimated number of fake events. The last concerns the events that pass the reconstruction level criteria but that are not present at the generator level (in the considered phase space) and has been estimated using the MC DY jets sample and rescaling the predictions bin by bin to fit the data minus background observations.

Starting with the MC generated distribution for $P_0(i_{\text{gen}})$, the first estimates $\hat{n}(i)$ can be obtained. From the later, a new probability distribution $P_0(i_{\text{gen}})$ can be inferred and in turn be used to estimate for a second time the $\hat{n}(i)$. The iteration is carried on until the desired number of iteration (discussed in the next paragraph) is reached.

Choice of the Number of Iterations

The Baye's method using an iterative approach requires a choice for the maximum number of iterations to be done. As discussed in [99] this should not lead to different results once the convergence has been reached. It has been checked that the result are indeed stable with respect to the chosen number of iteration by comparing the nominal result with distribution obtained with ± 1 additional iteration.

In this analysis, the iteration procedure is stopped as soon as the distribution obtained by folding, *i.e.* applying the raw response matrix, the unfolded distribution is compatible with the initial data signal distribution n_{obs} . This agreement is quantified by computing the $\chi^2/n.d.f.$ of the two distributions at each iteration step. The distributions are taken to be compatible when the $\chi^2/n.d.f.$ is less than 1 (errors on the folded unfolded distribution are not taken into account because there are totally correlated with the initial distribution uncertainties). Additionally, a minimum of 4 iterations is requested to ensure that the result be unbiased with respect to the MC predictions.

6.6 Uncertainties

Statistical uncertainties from measured spectra and response matrices are propagated analytically to the final results by mean of the unfolding procedure which provides the covariance matrices for both data and the response matrices finite number of events.

Systematic uncertainties originate from several sources among which the largest effects come from the uncertainty on the jet energy scale (JES) and resolutions corrections (JER). Another important contribution is the uncertainty on the background cross section, in particular on the cross section of the $t\bar{t}$ process, the main background for our analysis. In addition, the uncertainty coming from the unfolding procedure must be taken into account for the total systematic uncertainty. Finally, pile-up, luminosity and efficiencies scale factors uncertainties are also considered as a contribution to the total systematic uncertainty.

The following contributions, also listed in table 6.14 for the case of the jet exclusive multiplicity in the muon decay channel, are thus considered for the result:

• Jet energy scale (JES)

This uncertainty is calculated by rescaling the jet $p_{\rm T}$ spectrum up and down in data. The uncertainties, shown in figure 6.41, are $p_{\rm T}$ and η dependent and are estimated to be up to 5% [77] for the phase space of interest. For variables with many jets

6.6. UNCERTAINTIES

a cumulative effect is observed leading to uncertainties of up to 20% on the final results.

• Jet energy resolution (JER)

This correction has been estimated for data and MC in [75]. MC slightly overestimates the resolution compared to data. This difference is taken into account as already introduced in table 6.10. The effect is propagated accordingly by smearing the response matrix of the DY sample. This uncertainty is found to range from 0.3% to 1.2% for up to the fifth jet properties.

• Background estimation (XSEC)

Dominant background contributions come from $t\bar{t}$ and diboson processes and are modelled based on simulation. The total uncertainty receives contributions from the cross section and the total integrated luminosity uncertainties.

- Cross section: uncertainty on the background model is estimated by varying the background cross section for each of the background processes $(t\bar{t}, ZZ, WZ \text{ and } WW)$ independently by 10% for $t\bar{t}$ [100] and 3% for VV processes.
- Luminosity: because the subtraction of the background events is proportional to the luminosity, the total contribution is varied within one sigma of the total integrated luminosity.

The resulting uncertainty from the background estimation is found to be up to 1%.

• Pile-up (PU)

The contribution from pile-up correction uncertainty is taken into account by varying the minimum bias cross section by $\pm 5\%$ as shown in figure 6.2 page 131. This affects the background shape but the dominant effect comes from the modification in the response matrix of the DY sample. The resulting uncertainties are found to range from 0.2% to 2.2% depending on the jet multiplicity.

Table 6.14: Differential cross section in Exclusive jet multiplicity and break down of the systematic uncertainties for the muon decay channel.

Exclusive jet multiplicity											
N _{jets}	$\frac{d\sigma}{dN_{\rm jets}} [{\rm pb}]$	Tot. Unc [%]	stat [%]	MC stat. [%]	JES [%]	JER [%]	PU [%]	XSEC [%]	Lumi [%]	Unf [%]	Eff [%]
= 1	58.7	7.9	0.19	0.095	6.5	0.37	0.098	0.040	2.6	2.5	2.5
= 2	12.3	9.9	0.38	0.15	8.6	0.20	0.19	0.33	2.7	3.3	2.6
= 3	2.42	13.	0.85	0.27	11.	0.24	0.18	0.78	2.9	5.2	2.7
= 4	0.461	16.	2.0	0.55	15.	0.062	0.39	1.3	3.1	5.8	2.9
= 5	0.0938	21.	4.4	1.3	18.	0.22	0.049	1.7	3.2	10.	2.7
= 6	0.0148	28.	11.	3.0	24.	0.46	0.70	2.6	3.5	3.5	3.0
= 7	0.00221	38.	32.	8.5	17.	0.25	6.4	4.9	4.1	3.7	3.8

• Unfolding (Unf)

The arbitrary choice of the generator used to model the detector effects leads to a possible bias. An alternative response matrix obtained using the SHERPA 1.4 MC generator which uses a different hadronisation model (see section 3.1.3) has been consider for the unfolding. The resulting unfolded data has been compared to the unfolded distribution obtained using MADGRAPH 5 and the difference has been assigned as the unfolding systematic uncertainty.

• Luminosity (Lumi)

The total integrated luminosity uncertainty of 2.6% [101] is considered.

• Efficiency Correction (Eff)

The uncertainty of the data-to-simulation correction factor on the efficiency of the lepton reconstruction, identification, isolation and trigger is set as global factor on reconstructed distributions in data for each of the channels. These uncertainties amount to a total uncertainty of 2.5% and 0.5% in the muon and electron decay channels.

The estimate of the total systematic uncertainty on the cross section measurements is made by varying independently and in both direction (up and down) each of the contributing factor. The mean difference between the up and down variation is considered to be the corresponding systematic uncertainty. Finally, all these systematic uncertainties are added in quadrature assuming each uncertainty source is independent, yielding to the total systematic uncertainty.

Covariance matrices associated to each source of systematic are also computed following the recommendation of [102]. The diagonal elements are given by the squares of the above mentioned systematic uncertainty:

$$\sigma_{ii} = \sigma_i^2,$$

where $\sigma_i = |X_i^{\text{up}} - X_i^{\text{down}}|/2$. The off-diagonal entries are computed assuming a 100% correlation with the sign of the correlation determined by looking at variation direction when varying the source up and down, as in the $t\bar{t}$ analysis [103]:

$$\sigma_{ij} = \operatorname{sign}(X_i^{\operatorname{up}} - X_i^{\operatorname{down}}) \cdot \operatorname{sign}(X_j^{\operatorname{up}} - X_j^{\operatorname{down}}) \cdot \sigma_i \sigma_j$$

For illustration purposes, the effect of the main systematic uncertainties is presented for the jet multiplicity and 1st jet $p_{\rm T}$ distributions in figure 6.42 to 6.47. The cumulative effect of the JES uncertainty can be seen on the jet multiplicity variable, while not present on the jet $p_{\rm T}$ spectrum. The JER and PU uncertainties have very small effect


JES Uncertainties (for AK5 PF CHS jets)

Figure 6.41: Jet Energy Scale Uncertainties (in percentage) for anti- k_t 0.5 particle flow jets with charged hadron subtractions, as a function of the jet transverse momentum and jet pseudorapidity. To estimate the JES systematic uncertainty, the jet energy is varied by \pm the number in the present table. The dashed line represents the jet kinematics requirements.

on the jet multiplicity as well as on the jet $p_{\rm T}$. Concerning the background cross section uncertainty, while flat in jet $p_{\rm T}$, its increasing effect with the number of jets is directly related to the higher contamination at large number of jets. The same effect is also seen for the luminosity uncertainty. Finally, scale factors uncertainties are flat in each distribution, as expected as affecting the muon variables only.



Figure 6.42: Systematic effects on the exclusive jet multiplicity differential cross section due to up and down variations of (left) Jet Energy Scale and (right) Jet Energy Resolution, for the muon decay channel.



Figure 6.43: Systematic effects on the exclusive jet multiplicity differential cross section due to up and down variations of (left) PU in the MC and (right) Background cross sections, for the muon decay channel.



Figure 6.44: Systematic effects on the exclusive jet multiplicity differential cross section due to up and down variations of (left) the luminosity and (right) muon correction scale factors, for the muon decay channel.



Figure 6.45: Systematic effects on the 1^{st} jet p_T differential cross section due to up and down variations of (left) Jet Energy Scale and (right) Jet Energy Resolution, for the muon decay channel.



Figure 6.46: Systematic effects on the 1^{st} jet p_T differential cross section due to up and down variations of (left) PU in the MC and (right) Background cross sections, for the muon decay channel.



Figure 6.47: Systematic effects on the 1^{st} jet p_T differential cross section due to up and down variations of (left) the luminosity and (right) muon correction scale factors, for the muon decay channel.

6.7 Single Channel Differential Cross Section

From the detector level reconstructed data distributions, background MC estimations, and response matrices estimated with the DY + jets samples, the measured differential cross sections are thus obtained by first subtracting the background estimates from the data, then unfolding the obtained distributions and finally dividing the result by the estimated data luminosity and each bin content by its bin width.

6.7.1 Channels Comparison

The resulting differential cross sections obtained for each of the two Z boson decay channels, together with their statistical uncertainties, are shown for the jet multiplicities, the first leading jet transverse momentum, the second leading jet pseudorapidity and the jets scalar sum of transverse momenta for $N_{\text{jets}} \geq 3$ in figures 6.48 and 6.49. In order to decrease the statistical uncertainty, the two channels will be combined later on. It is therefore necessary to check the compatibility of the two measurements. To this aim, the ratios between the electron and muon channels differential cross sections, taking into account statistical uncertainties alone, are also presented in these figures.

By looking at the differential cross sections as a function of the jet multiplicity, the electron result is about 2% higher for the first two bins than the muon cross section, while for higher multiplicities, the two channels look compatible within statistical uncertainties. The few percent difference is however covered when taking into account the lepton scale factor uncertainties which are of the percent level.

Concerning the differential cross sections as a function of the leading jet transverse momentum, a specific pattern is observed in the ratio between the electron and muon channels with the former being up to 8% larger at $p_{\rm T}$ values around 220 GeV. Considering the total systematic uncertainty that can affect in a different way the two decay channels in top of the statistical error, this difference stands below two sigma. As shown later (figure 6.51 on page 175) each channel is found to be compatible with MAD-GRAPH5_AMC@NLO.

For the second leading jet pseudorapidity and the jets scalar sum of transverse momenta, the ratios do not indicate significant differences between the two channels.



Figure 6.48: Differential cross section as a function of (left) the exclusive jet multiplicity and (right) the 1st leading jet $p_{\rm T}$ with $N_{\rm jets} \ge 1$, for the muon (green) and electron (blue) decay channels. Only statistical uncertainties are taken into account.



Figure 6.49: Differential cross section as a function of (left) the 2nd leading jet $|\eta|$ with $N_{\text{jets}} \geq 2$ and (right) the scalar sum of jets transverse momenta with $N_{\text{jets}} \geq 3$, for the muon (green) and electron (blue) decay channels. Only statistical uncertainties are taken into account.

6.7.2 Full Single Channel Results

The full treatment of systematic uncertainty is performed on each channel separately by repeating the unfolding procedure for each source of systematic uncertainty for both up and down variations. The mean difference of the up and down variations is quoted as the systematic uncertainty on a bin by bin basis.

The resulting differential cross sections, together with their total and statistical uncertainties are then compared to the following MC generator predictions (see chapter 3 for more details):

- MADGRAPH 5, tree-level generator with up to four final state partons, using the CTEQ6L1 [80] pdf and interfaced with PYTHIA 6 Tune Z2 Star for parton shower and hadronisation, normalised to the inclusive NNLO cross section (see section 6.2) computed with FEWZ,
- SHERPA 2, NLO generator with up to four final state partons, with matrix element computation of Z + 0, 1, 2 parton events at NLO accuracy, using the CT10 [104] PDF and implementing its own parton showering and hadronisation model,
- MADGRAPH5_AMC@NLO, NLO generator with up to three final state partons, with matrix element computation of Z + 0, 1, 2 parton events at NLO accuracy, using the NNPDF [105] PDF and interfaced with PYTHIA 8 for parton shower and hadronisation, normalised to the inclusive NNLO cross section.

The uncertainty from the PDF, including the contribution coming from the uncertainty on α_s , has been evaluated with the CT10 PDF set on the tree level calculation (MADGRAPH 5) and goes from 1.5% to 5.5% (68% confidence level uncertainty) depending on the jet multiplicity, the highest values are obtained for the highest multiplicities. Due to lack of time, the renormalisation and factorisation scale uncertainties have not been computed in the present work but will be present in the foreseen paper and estimated using the NLO predictions of MADGRAPH5_AMC@NLO. However, an idea of their contribution can be obtained from the 7 TeV results [47]. From their measurements, the systematic scale uncertainty on the NLO predictions varies from 5% in the first jet multiplicity bin, to about 50% for the 6 jets bin. For the $p_{\rm T}$ distributions, the uncertainty varies from 5% to 40% for the leading jet, and from 20% to 60% for the third and fourth leading jet. Roughly the same values are obtained for $H_{\rm T}$ distributions.

For illustration purposes, in this section the results are shown for each decay channel individually only for the jet multiplicities, the first leading jet transverse momentum, the second leading jet pseudorapidity and the jets scalar sum of transverse momenta for $N_{\text{jets}} \geq 3$. All results can be found in appendix B. Additionally, the corresponding tables with break down of the uncertainties are available in the related CMS Analysis Note [106].

For each observable, the data cross section measurement is drawn on top of the different DY generated distributions. Additionally, the lower panels in each figure show the ratios of the theory predictions to data. Errors bars around the experimental points show the statistical uncertainty, while the crosshatched bands indicate the statistical plus systematic uncertainties added in quadrature. The coloured filled band around the MC prediction represents the statistical uncertainty of the generated sample. The uncertainties from pdf and scale variation are not included in these plots.

From the ratio plots of the jet multiplicity cross section (see figure 6.50), it can already be stated that the three predictions are in remarkably good agreement with data, even for multiplicities above the number of generated partons, achieved thanks to the parton showering modelling.

A small difference is observed between the LO (MADGRAPH 5) predictions and the NLO predictions (SHERPA 2 and MADGRAPH5_AMC@NLO) for the leading jet transverse momentum differential cross section (see figure 6.51). An overestimation of the LO MC is seen in the range $150 \leq p_{\rm T} \leq 450$ GeV, which does not appear in the NLO predictions. Let us note here that the scale variation uncertainty is expected to be larger for the LO predictions than for the NLO predictions (as already seen in figure 2.9 on page 30).

A similar effect is seen in the pseudorapidity tail of the second leading jet differential cross section (see figure 6.52) where the LO predictions are decreasing faster than what is observed in the data, while the NLO predictions show a very good description of the data.

Concerning the scalar sum of the jets transverse momenta for at least 3 jets (see figure 6.53), all three predictions agree with the measurement within uncertainties. For values close to the threshold the MC predictions underestimate the cross section but are still covered by the large systematic uncertainties resulting from the cumulative effect of JES.

All the measurements in each decay channel are combined in the next section.



Figure 6.50: Differential cross section as a function of the exclusive jet multiplicity, for the (left) muon and (right) electron decay channels.



Figure 6.51: Differential cross section as a function of the 1st leading jet $p_{\rm T}$ with $N_{\rm jets} \ge 1$, for the (left) muon and (right) electron decay channels.



Figure 6.52: Differential cross section as a function of the 2nd leading jet $|\eta|$ with $N_{\text{jets}} \geq 2$, for (left) the muon and (right) electron decay channels.



Figure 6.53: Differential cross section as a function of the jets $H_{\rm T}$ with $N_{\rm jets} \ge 3$, for the (left) muon and (right) electron decay channels.

6.8 Combined Differential Cross Sections

6.8.1 Combination

With the single channel muon and electron differential cross section results one can obtain and estimate for the $Z \rightarrow l\bar{l}$ channel, *i.e.* for one of the lepton decay channel, by combining those measurements. Since the coupling of the leptons to gauge bosons are flavourindependent, this lepton universality tells us that we are precisely measuring the same observables in both the muon and electron decay channels. For each bin of each observable we therefore have two measurements, one done with the electron channel, and one using the muon channel. These measurements are however not independent, for example, the same method was used for the unfolding, using the same MC generator to reconstruct the response matrices. The estimated luminosity is of course correlated between the two channels, as well as the background, pile-up, jet energy scale and jet energy resolution uncertainties.

In order to combine the two channels a weighted mean is performed to estimate the cross section values and its covariance matrix. For the combined cross section estimate, one generally writes:

$$x_{\alpha}^{\text{comb.}} = \sum_{i=1}^{2n} \lambda_{\alpha i} y_i = \sum_{i=1}^n \lambda_{\alpha i} x_i^{ee} + \sum_{i=n+1}^{2n} \lambda_{\alpha i} x_{i-n}^{\mu\mu}$$

where α represents the bin number and ranges from 1 to n, the index i is taking value between 1 and 2n and where the superscripts ee and $\mu\mu$ denote the decay channel, the x_{α}^{ll} are the bin α values of the observable x as measured in the ll decay channel. The 2n vector y and the $n \times 2n$ matrix λ respectively contain the measurements and the coefficients of the electron channel for $1 \leq i \leq n$ and of the muon channel for $n + 1 \leq i \leq 2n$. For the weighted mean method the coefficients are simple and are given by:

$$\begin{aligned} x_{\alpha}^{\text{comb.}} &= \lambda_{\alpha\alpha} y_{\alpha} + \lambda_{\alpha\alpha+n} y_{\alpha+n} \\ &= \frac{(\sigma_{\alpha}^{ee})^{-2}}{(\sigma_{\alpha}^{ee})^{-2} + (\sigma_{\alpha}^{\mu\mu})^{-2}} x_{\alpha}^{ee} + \frac{(\sigma_{\alpha}^{\mu\mu})^{-2}}{(\sigma_{\alpha}^{ee})^{-2} + (\sigma_{\alpha}^{\mu\mu})^{-2}} x_{\alpha}^{\mu\mu} \end{aligned}$$

where the σ_{α}^{ll} are the total uncertainty attached to the related measurement.

To estimate the uncertainty on the combination the propagation of the covariance matrix is applied using the coefficients $\lambda_{\alpha i}$ as determined in the previous equation. That is, the covariance matrix of the combined cross section for some observable x, is given by:

$$\sigma_{\alpha\beta}^{\text{comb.}} = \sum_{i=1}^{2n} \sum_{j=1}^{2n} \lambda_{\alpha i} \mathcal{M}_{ij} \lambda_{\beta j}$$

where \mathcal{M}_{ij} is a $2n \times 2n$ matrix representing the full measurement covariance and is obtained by grouping each channel covariance matrix taking correlation into account as follows:

$$\mathcal{M} = \sum_{s \in \text{ sources}} \mathcal{M}^{s}$$
$$= \sum_{s \in \{\text{JES, JER, PU, XSec, Lumi\}}} \left(\frac{(\sigma^{ee})^{s}}{(\sigma^{ee}_{\alpha})^{s}(\sigma^{\mu\mu}_{\beta})^{s}} \left| \frac{(\sigma^{\mu\mu}_{\alpha})^{s}}{(\sigma^{\mu\mu}_{\beta})^{s}} \right| \right)$$
$$+ \sum_{s \in \{\text{MC stat., Lep. SF.}\}} \left(\frac{(\sigma^{ee})^{s}}{0} \frac{(\sigma^{\mu\mu}_{\alpha})^{s}}{(\sigma^{\mu\mu}_{\beta})^{s}} \right)$$

where the $(\sigma^{ll})^s$ are the single channel covariance matrices for the effect *s*, obtained as described in section 6.6, and $(\sigma^{ll}_{\alpha})^s = \sqrt{(\sigma^{ll})^s_{\alpha\alpha}}$. With this result, it is therefore possible to obtain the covariance matrices of the combination for each source of uncertainty as well as the total uncertainties.

6.8.2 Results

The cross section results shown here correspond to pp collision at $\sqrt{s} = 8$ TeV and have been integrated over the phase space $71 \leq M_{l^+l^-} \leq 111$ GeV with leptons satisfying $|\eta^l| \leq 2.4$ and $p_T^l \geq 20$ GeV and corrected for EWK FSR (see section 6.5.1. The jets are considered for $p_T^j \geq 30$ GeV and $|\eta^j| \leq 2.4$ using anti- k_t with a cone size of 0.5 at the hadron level.

Jet Multiplicities

The measured cross sections, for the combination of the two decay channels $(Z \to \mu^+ \mu^$ and $Z \to e^+ e^-)$, as a function of exclusive and inclusive jet multiplicities are shown in figure 6.54 page 182 for up to seven jets in the final state.

The agreement of the predictions with the measurement is very good for jet multiplicities going up to the maximum number of final state partons included in the matrix element calculations, namely 4 in the MC generators used here. It is already good at treelevel (MADGRAPH 5) renormalised to the NNLO inclusive cross section. For larger jet multiplicity the difference between predictions and data is still within the uncertainties.

Jet Transverse Momemta

The measured differential cross sections as a function of jet $p_{\rm T}$ for the first, second, third, fourth and fifth leading jets are presented in figures 6.55, 6.56, and 6.57 from pages 183 to 184. The cross sections are falling rapidly with increasing $p_{\rm T}$ for all the jets in the final state: for the jets with the largest transverse momentum (figure 6.55 left) it decreases over almost two orders of magnitude for $p_{\rm T}$ between 30 and 100 GeV, while the cross section for the 5th jet decreases over 3 orders of magnitude in the same $p_{\rm T}$ range.

For the leading jet, the agreement of the MADGRAPH 5 prediction with the measurement is very good up to ~150 GeV. Discrepancies are observed from ~150 to ~450 GeV. A similar bump on the ratio with the tree level calculation was observed at $\sqrt{s} = 7$ TeV in the CMS measurement [47], using for the prediction the same generators as here, as well as in the ATLAS measurement [49], using ALPGEN [108] interfaced to HERWIG [109] for the prediction (see chapter 3). The SHERPA 2 calculation predicts a slightly more compatible spectrum for the hardest jet. The same level of agreement is observed for the additional jets. MADGRAPH5_AMC@NLO predictions do not show the bump observed in MADGRAPH 5 and are in excellent agreement with the data.

Jet Pseudorapidities

The inclusive jet differential cross sections as a function of jet absolute pseudorapidity for the first, second, third, fourth and fifth leading jets are presented in figure 6.58, 6.59, and 6.60 from pages 185 to 186.

The pseudorapidity distributions tend to become flatter as the number of jets increases. Between the very central region at pseudorapidities around 0 and the edge of the accepted region ($|\eta| = 2.4$), the difference is about a factor of 2 for the first jet multiplicity and decreases to a factor of about 1.5 for the four jets case.

MADGRAPH 5 predicts a more central distribution than what is measured in data for jets 1 to 4. The shape of the distribution is well described for the 5th jet even though the global yield is underestimated as already seen in jet multiplicity distributions. The NLO generators SHERPA 2 and MADGRAPH5_AMC@NLO on the other hand do not exhibit a too central distribution. Their predictions agree outstandingly well with the measurements apart for the slightly too small total yield of the 4th and 5th jets differential cross sections of MADGRAPH5_AMC@NLO as expected from the jet multiplicity ratios.

Jet Transverse Momenta Scalar Sum

Differential cross sections as a function of the scalar sum of jet transverse momenta, $H_{\rm T}$, for inclusive one, two, three, four and five jets production are shown in figures 6.61, 6.62 and 6.63 from pages 187 to 188.

The predictions of the considered generators agree well with the measurements within the experimental uncertainties. The agreement is excellent in the whole range for $N_{\text{jets}} \geq$ 1 and $N_{\text{jets}} \geq$ 2 and all three generators predictions agree. However, for higher jet multiplicities, $N_{\text{jets}} \geq$ 3, $N_{\text{jets}} \geq$ 4 and $N_{\text{jets}} \geq$ 5, all three predictions underestimate the cross section at the beginning of the spectra close to the threshold, but the precision of the measurement does not allow to discard the prediction event though SHERPA 2 seems to better agree than the others.

Dijet Mass

The differential cross section as a function of the two leading jets invariant mass, M_{jj} , for inclusive two jets production is shown in figure 6.64 page 189. Again, all three generators provide an excellent level of agreement with the measurement apart at the lowest mass bin where the cross section is rapidly increasing to reach its maximum at ~ 100 GeV. From the uncertainties it is not really possible to distinguish one best generator in this region even though SHERPA 2 predictions are the closest to the measured cross section.

General Comments

It has to be noted again that the uncertainty due to renormalisation and factorisation scale variations, is not included here. Their consideration will lead to full compatibility. The LO predictions of MADGRAPH 5 are obtained using PYTHIA 6, while this is PYTHIA 8 which is used for the NLO predictions of MADGRAPH5_AMC@NLO. It is therefore difficult to make a clear statement regarding the source of the difference between MADGRAPH 5 and MADGRAPH5_AMC@NLO predictions. A study of MADGRAPH 5 interfaced with PYTHIA 8 should be performed to that aim. Finally, even though the presented variable are not very sensitive to the PDF set used for the prediction (as seen from the relatively low uncertainty values related to the choice of a PDF set from the 7 TeV paper [47]), it should still be kept in mind that each generator in this analysis uses a different PDF set. This remark is more important for observables such as correlation between jets which can be more sensitive to the PDF.

Additionally, all plots and tables with break down of the uncertainties are available

in the related CMS Analysis Note [106] and will be part of the results to be published.



Figure 6.54: Differential cross section as a function of the (left) exclusive and (right) inclusive jet multiplicities, for the combination of the muon and electron decay channels.



Figure 6.55: Differential cross section as a function of the (left) 1st leading jet $p_{\rm T}$ for $N_{\rm jets} \geq 1$ and (right) 2nd leading jet $p_{\rm T}$ for $N_{\rm jets} \geq 2$, for the combination of the muon and electron decay channels.



Figure 6.56: Differential cross section as a function of the (left) $3^{\rm rd}$ leading jet $p_{\rm T}$ for $N_{\rm jets} \geq 3$ and (right) $4^{\rm th}$ leading jet $p_{\rm T}$ for $N_{\rm jets} \geq 4$, for the combination of the muon and electron decay channels.



Figure 6.57: Differential cross section as a function of the 5th leading jet $p_{\rm T}$ for $N_{\rm jets} \geq 5$, for the combination of the muon and electron decay channels.



Figure 6.58: Differential cross section as a function of the (left) 1st leading jet η for $N_{\text{jets}} \geq 1$ and (right) 2nd leading jet η for $N_{\text{jets}} \geq 2$, for the combination of the muon and electron decay channels.



Figure 6.59: Differential cross section as a function of the (left) $3^{\rm rd}$ leading jet η for $N_{\rm jets} \geq 3$ and (right) $4^{\rm th}$ leading jet η for $N_{\rm jets} \geq 4$, for the combination of the muon and electron decay channels.



Figure 6.60: Differential cross section as a function of the 5th leading jet η for $N_{\text{jets}} \geq 5$, for the combination of the muon and electron decay channels.

186



Figure 6.61: Differential cross section as a function of the jets $H_{\rm T}$ for (left) $N_{\rm jets} \ge 1$ and (right) $N_{\rm jets} \ge 2$, for the combination of the muon and electron decay channels.



Figure 6.62: Differential cross section as a function of the jets $H_{\rm T}$ for (left) $N_{\rm jets} \geq 3$ and (right) $N_{\rm jets} \geq 4$, for the combination of the muon and electron decay channels.



Figure 6.63: Differential cross section as a function of the jets $H_{\rm T}$ for $N_{\rm jets} \ge 5$, for the combination of the muon and electron decay channels.



Figure 6.64: Differential cross section as a function of the dijet mass for $N_{\text{jets}} \ge 2$, for the combination of the muon and electron decay channels.

6.9 Conclusions

The results presented in this chapter are the measurement of the differential cross sections for the production of a Z boson in association with up to 5 jets. They provide an important contribution to the understanding of QCD and of its predictions implemented in different MC generators. Additionally, such measurements are of great importance for many other studies for which the Z boson plus jets process is an important background and needs to be controlled accurately.

From the present results, it can be stated that the tree-level generator MADGRAPH 5, when interfaced with PYTHIA 6 for parton shower and hadronisation, is able to describe data to a very good level of agreement apart for little regions of the probed phase space. The NLO generators MADGRAPH5_AMC@NLO interfaced with PYTHIA 8 and SHERPA 2 are in excellent agreement with data and are able to describe some of the regions where the LO generator does not exactly describe the data. Still, some phase space corners. such as the lowest values in the scalar sum of the jets transverse momenta distributions or in dijet invariant distributions are not very well reproduced by the MC even though large uncertainties affect the measurement in these regions. However, a complete comparison of the data to MC generator predictions should include the uncertainties coming from PDF and scale variations.

The data measurements contained in this thesis were preapproved in June 2014 and are available [110]. These results and other Z + jets measurements such as double- and triple-differential cross sections as a function of jet pseudorapidity, transverse momentum and Z transverse momentum [111], azimuthal [112] and longitudinal [113] correlations, are being grouped inside a single paper to be published in a near future. This paper will contain a review of the Z + jets processes at 8 TeV in pp collisions providing reference information.

The next runs of the LHC starting this year are going to provide proton-proton collisions at even higher centre-of-mass energies $\sqrt{s} = 13$ TeV. The LHC will therefore once more allow the exploration of phase space domains unreachable up to now. The same analysis has thus to be performed to provide a "standard candle" in the exploration of this new phase space and also to test and better understand QCD at these extreme energies.

Chapter 7

The Triple-GEM Detectors

This chapter presents the development of a data acquisition (DAQ) system for the Triple-Gas Electron Multiplier (GEM) detectors to be installed in the CMS detector during the second upgrade stage. The present part of this thesis, more hardware oriented, details the development of a cosmic muon test bench at the IIHE, based on the Triple-GEM detectors, which was afterwards tested in muon and pion beams at CERN. A full DAQ system had to be designed from scratch for testing the readout electronic of those detectors. This chapter introduces first the motivation for the GEM technology and the detectors. Then the implementation of the complete DAQ system is discussed with an emphasis on the way the communication between different electronic entities is achieved. Finally, first results obtained with the set-up developed at the IIHE with cosmic muons as well as test beam results are shown and prove the successful development of the full DAQ system. A detailed description of the electronic components used for the test bench concerned in this chapter can be found in the following reference [114].

7.1 Motivation

To understand the motivation of studying the Triple-GEM detectors introduced in the next sections, a closer look at the LHC upgrade plan is necessary. Table 7.1 summarises the past, present and future operation and maintenance periods of the LHC.

The most crucial parameter in that table is the instantaneous luminosity which, as explained in chapter 4, is related to the total rate of beam particles interaction inside the detector. This parameter is scheduled to be increased successively for the different LHC phases and has a significant impact on the detector upgrades. Indeed, the current muon system in the endcaps, for instance, is not appropriate for the high particle rates foreseen after the long shutdown 2, in 2019. The RPCs (see chapter 4) in particular do not allow rates beyond the kHz cm⁻² scale [53], while a rate of a few kHz cm⁻² is expected in the forward region $|\eta| \ge 1.6$. Another technology, able to handle the future high rates, has therefore to be used and the idea of using GEM detectors in place of RPC for high pseudorapidity ranges came up inside the CMS collaboration. A group has been set up in 2010 to study the feasibility of this technology that offers much higher rate capability, up to 10 MHz cm⁻², fulfilling the requirements of LHC phase 2 starting after the long shutdown 2.

The Triple-GEM detectors (described later in this chapter) are planed to be installed during the next CMS upgrade in the detector slots originally dedicated to the RPC modules at pseudorapidity values between 1.6 and 2.4 as shown in figure 7.1. These detectors will improve the muon transverse momentum resolution at trigger level and by consequence maintain the desired trigger rate that would have increased otherwise as illustrated in figure 7.2. A reasonably low muon transverse momentum threshold can thus be applied at trigger level allowing for greater acceptance of rare physics signatures leading to muons in the final state. The Triple-GEM detectors will in addition offer a redundancy in this region of the detector where the CSC alone are present up to now.

For the detector upgrade plan to be fully accepted, the Triple-GEM detectors must be tested in order to show that, indeed, this technology can be used and provides the required characteristics. To this aim, a test bench has been build at the IIHE and tests with cosmic muons in a first step, and then with muons and pions beams at CERN, have been performed using the set-up developed at the IIHE. After introducing the Triple-GEM detectors in some more details, the full development of the IIHE test bench is described with an emphasis on the data acquisition chain that had to be developed.

Period	Energy	Instantaneous Luminosity						
2010 - 2012	7 - 8 TeV	$0.5 \times 10^{34} \mathrm{cm}^{-2} \mathrm{s}^{-1}$						
	Long Shutdov	vn 1 (LS1)						
2015 - 2017	13 - 14 TeV	$1.0 \times 10^{34} \mathrm{cm}^{-2} \mathrm{s}^{-1}$						
	Long Shutdow	vn 2 (LS2)						
2019 - 2021	$14\mathrm{TeV}$	$2.0 \times 10^{34} \mathrm{cm}^{-2} \mathrm{s}^{-1}$						
Long Shutdown 3 (LS3)								
2022 - 2030 (?)	$14\mathrm{TeV}$	$1.0 \times 10^{35} \mathrm{cm}^{-2} \mathrm{s}^{-1}$						

Table 7.1: Summary of the past, present and future LHC periods with their corresponding energy and luminosity.



Figure 7.1: Schematic longitudinal view of 1/4 of the CMS detector on which the locations GE1/1 and GE2/1 of the first to be installed Triple-GEM detectors are shown in red. As can be seen, these locations cover the region $1.6 \leq |\eta| \leq 2.2$ and establish a redundancy with the CSC detectors in this pseudorapidity region.

7.2 Gas Electron Multiplier Detector

This section covers the geometry and principles of functioning of a GEM detector. The general principle is first described and then applied to the specific case of the Triple-GEM detectors. The front-end readout electronics is also reviewed in this section.

7.2.1 Gas Detectors

The GEM detectors are one example of gaseous ionisation detectors. Those are based on the ionisation process taking place when an ionising particle such as a electrically charged particle or a photon traverses a medium filled with gas (or made of condensed matter). Indeed when such a particle penetrates a medium it mainly interacts via electromagnetic interaction with the nuclei of the medium, resulting in both excitation and ionisation of the atoms and leading to an energy loss of the incoming particle. The average differential



Figure 7.2: Level-1 simulated single muon trigger rate in pseudorapidity range 1.6 < $|\eta| < 2.2$ as a function of the threshold on the reconstructed muon transverse momentum for the LHC phase 2 run conditions with 50 pile-up interactions at a centre-of-mass energy $\sqrt{s} = 14$ TeV, for the current CMS detector (blue) and for the current detector supplemented by the GE1/1 GEM detector (purple). For a given L1 Trigger Rate, we see that the Triple-GEM detector in GE1/1 allows a smaller threshold than the current detector, resulting in larger acceptance phase space. Results extracted from [115].

energy loss per unit of length for swift charged particles (protons, alpha particles, atomic ions, muons but not electrons that suffer much larger losses by Bremsstrahlung) can be approximated by the Bethe-Bloch relation [116]:

$$\frac{dE}{dX} = -K\frac{Z}{A}\frac{\rho}{\beta^2} \left(\ln\frac{2mc^2\beta^2 E_M}{I^2(1-\beta^2)} - 2\beta^2\right), \ K = \frac{2\pi N z^2 e^4}{mc^2}$$
(7.1)

where N is the Avogadro number, m and e are the electron mass and charge, Z, A and ρ are the atomic number and mass, and the density of the medium, respectively, and I is its effective ionisation potential, z is the charge and β the velocity of the incoming particle. E_M is the maximum energy transfer allowed in each interaction.

Photons also ionise the medium but while charged particles will ionise the medium along their path, photon interactions will lead to a single localised event.

In both cases, positively charged ions and free electrons are produced from ionised

Without an electric field the electrons and ions would quickly recombined. atoms. Gaseous ionisation detector therefore always come with an electric field created by a voltage difference between a cathode and an anode in order to make electrons drift in one direction (towards the anode), and ions drift in the opposite direction (towards the cathode). Depending on the intensity of the electric field, different behaviours can be observed. For low difference of potential, the freed electrons tend to recombine with ions leading to very small or no signal. By increasing the electric field, the electrons created during ionisation are able to drift towards and reach the anode. At higher electric field intensities the electrons are furthermore accelerated and gain enough energy to in turn ionise the gas, freeing additional electrons that are accelerated by the electric field and also produce further ionisation. This last regime therefore creates a cascade of electrons drifting in the direction of the anode. As explained in the next paragraphs, the electric field inside a Triple-GEM detector is not uniform. In some regions electrons are drifting without having enough energy to ionise the gas while in other regions, the electric field is high enough to allow additional ionisation to occur.

7.2.2 Triple-GEM Detector Geometry

As shown in figure 7.3, a Triple-GEM detector is made of three GEM foils sandwiched between a kapton-covered cathode at the top and an anode at the bottom on which silicon strips have been placed. The whole volume being filled with gas.

A GEM foil, shown in figure 7.4 is composed of a 50 μ m-thick kapton plate inserted between two 5 μ m-thick copper plates, on which microscopic holes with conical sections have been drilled. The inner and outer diameters of the holes are 50 and 70 μ m, respectively and the distance between the centre of a hole and its neighbours of 140 μ m.

The distance between the different elements from the cathode to the anode can be changed and different configurations are studied. Typical configurations being from top to bottom 3-2-2-2 and 3-1-2-1 mm.

To develop and test the DAQ system at the IIHE, a small $10 \times 10 \text{ cm}^2$ Triple-GEM prototype, shown in figure 7.5, was at our disposal. The prototype was in a 3-2-2-2 mm gap configuration. The gas mixture used for our set-up was Ar(70%)-CO₂(30%). The high voltage was provided to each part of the Triple-GEM detector via a voltage divider. Typical values for the tension used during the tests are given in figure 7.3. The resulting electric field of about 2 to 5 kV cm⁻¹ in the drift, transfer and induction regions is not high enough for secondary ionisation processes to take place. On the other hand, the voltage difference of about 350 V, leading to an electric field of 70 kV cm⁻¹ within the holes of the GEM foils, sufficiently accelerates the freed electrons so that further ionisation occur



Figure 7.3: Schematic representation of a Triple-GEM detector in a 3-2-2-2 configuration. When an electrically charged particle passes through the gas filling the detector volume, ionisation processes occur freeing electrons that in turn ionise other atoms when accelerated by the high electric field present in the holes of the GEM foils. The resulting electron cascade drifts towards the readout strips due to the electric field created by different tensions applied on each layer of the Triple-GEM detector.



Figure 7.4: Schematic view of a portion of a GEM foil.

resulting in an electron cascade.

7.2.3 Triple-GEM Front-End Electronic

The prototype has two times 128 strips in one direction and another two times 128 strips in a perpendicular direction. For each set of 128 strips a so-called VFAT2 hybrid, shown in figure 7.6, hosting a readout VFAT2 chip [117], can be connected. The VFAT2 hybrid together with its chip form the electronic entity the closest to the detector. It has to be noted that this is exactly similar to the final design of the Triple-GEM detectors to be installed inside CMS as we will see later.



Figure 7.5: Photograph of the Triple-GEM $10 \times 10 \text{ cm}^2$ prototype detector. The gas input can be seen on the bottom right of the picture and its output on the top left. On the bottom left lies the high voltage input and the voltage divider. The green elements visible on the picture are each connected to 128 strips of the detector and are used to collect the total signal of the 128 strips. The VFAT2 hybrids (see picture 7.6) will be connected in place of one of these green elements for the data taking operation.

The VFAT2 chip is a trigger and tracking front-end Application Specific Integrated Circuit (ASIC) providing fast regional hit information (for triggering purposes) and precise spatial hit information (for tracking purposes) by reading its 128 input channels, which are in our set-up connected to one of the four sets of 128 strips of the Triple-GEM detector. The VFAT2 chip works synchronously with its input clock signal at a frequency of 40 MHz.

At every clock cycle, the VFAT2 chip reads its 128 channels and writes the logical states (hit or not with an adaptable threshold) of each of them, in parallel for all 128 channels, into the first of two RAM memories (SRAM1). At the same time a fast OR of a configurable number of channels is performed to set trigger flags, called the S-bits, that can be used by the trigger system. The SRAM1 can contain up to 256 recorded events, corresponding to 6.4 µs. This is directly dictated by the fact that the CMS L1 trigger response, that can be sent to the VFAT2 when an event needs to be kept, is not expected to exceed 256 LHC clock periods.

When such a L1 accept signal is received by the VFAT2, the corresponding event in SRAM1 is fetched and written together with some additional header information into the



Figure 7.6: Photograph of VFAT2 Hybrids hosting the VFAT2 readout chip, (left) the TOTEM version used with the $10 \times 10 \text{ cm}^2$ prototype detector and (right) the CMS version used on the GEM Electronic Board (GEB) to be installed in CMS. The white connector on the left of each VFAT2 hybrid connects the 128 strips of a Triple-GEM detector to the VFAT2 chip. The white connector on the right of the CMS VFAT2 hybrid and (not visible on this picture) on the back of the TOTEM VFAT2 hybrid links the VFAT2 chip to the next element in the readout chain.

second VFAT2 RAM memory (SRAM2). In order to know which of the 256 stored events in SRAM1 actually corresponds to the triggered event, the CMS L1 latency (the time it takes to build the L1 response) has to be precisely measured¹ so that the right event is transmitted to the SRAM2. The latter is a First In First Out (FIFO) memory able to store up to 128 events. The VFAT2 then transmits events stored in its FIFO over a serial link to the next stage of the DAQ system.

The VFAT2 chip parameters such as the latency and threshold mentioned above, can be configured via its dedicated registers that can be accessed by its I2C (see section 7.3.2) port. Establishing the communication with the front-end VFAT2 chip was one of the first development steps as explained later in this chapter.

¹One thing to keep in mind is that the L1 response is entirely determined by hardware. By consequence, the latency is fully determined by the number of clock cycles it takes to determine the L1 response by the hardware plus the time it takes for the L1 response signal to be sent to the targeted subdetectors. As a result, the latency time is fixed once the hardware and the transmission lines are fixed and must be measured once and for all.



Figure 7.7: High level representation of the DAQ chain from the user interface to the experiment server. Double arrows indicate the control commands and data transfer in both direction. The central node, is the intermediate between the user and the experiment. Not shown are all the detector electronics and supplies.

7.3 Implementation of a DAQ System

In order to test the Triple-GEM prototype detector and the whole readout electronics chain, a complete DAQ system has been implemented. The whole chain, as illustrated in figure 7.7, includes a user interface (web application) connected to the central node (web server and database), itself connected to a server close by the experimental set-up for control and monitoring.

The next sections detail the DAQ architecture and the way the communication is handled between the different entities.

7.3.1 DAQ Design

User Interface

A user-friendly web interface as been developed in order to give a user the ability to control remotely different parameters related to the experimental set-up as well as to launch different data acquisition scripts. The goal was to design a small application for the local set-up that could also be used by other students to perform measurements. This lead to the choice of a web application that can be run on a simple web browser and that communicates with the central node of the DAQ, *i.e.* a web server coupled to a database, by the internet protocol. Figure 7.8 shows screen shots of the web interface giving the options to start a new data acquisition run, modify the high voltage supplied to the Triple-GEM detector and photomultipliers, change the gas flow and mixture, and monitor each of these quantities by giving real time information.

Overv High Vol Current val Channel 0 1	Itage Iues delivered	by the HV m								
High Vol Current val Channel 0 1	Itage lues delivered Description	by the HV m								
High Vol Current val Channel 0 1	ltage lues delivered Description	by the HV m								
Channel 0 1	Description	by the first first	odule			R	uns formation about	the current run		
0		Voltag	aje(V) C	Current (I)	Power	Ru	un art time	447 Mar 24, 2015 ai	13:20:05	
1	GEM HV	0	0)	Off	St	arted by	Gilles De Lento	decker	
2	unused	o	o)	Off		omment	OFF		
	unused	o	o)	Off					
3	PM3 (below (EM) O	o)	Off					
4	PM4 (below 0	EM) O	o)	Off					
5	PM5 (above (EM) O	o)	Off					
Gas Syst Current val	tem Iues delivered	by the Gas n	nodule							
Channel	Descrip	otion	Flow (n	ml/min)						
1	Ar		18							
2	CO2									
3			12							
	CF4		12 0							
ligh Vo	CF4 Mixture oltage m	nonitori	12 0 25							
4 High Vo Current para The table bello	CF4 Mixture Dltage m arneters and r	nonitoring va	12 0 25 NG alues d paramet	ters (by the	e user), set p	parameters (in	iside the HV module	and monitoring para	ameters (produce	id by the module
4 High Vo Current para The table bello Channel Des	CF4 Mixture Dltage m ameters and r pw lists each cha escription Requ Pow	nonitoring va nnet's requeste uested Set er Powe	12 0 25 NG alues d paramet r Voltag	ters (by the sted St e (V) Vo (V)	e user), set p et M oltage Vo)	oarameters (in Ionitoring Oltage (V)	iside the HV moduli Requested Maximum Curre (uA)	ə), and monitoring para Set Maximum nt Current (uA)	ameters (produce Monitoring Current (uA)	d by the module Temperatur (°C)
4 High Vo Current para The table bello Channel Des	CF4 Mixture Ditage m anneters and r pw lists each cha ascription Req Pow	nonitoring v nnet's requeste uested Set er Powe	12 0 25 NG alues d paramet r Reque: r Voltag	ters (by the sted Si e (V) Vi (V 44	e user), set p et M oltage Vo 100 31	Darameters (in Ionitoring Oltage (V) 1833	iside the HV moduli Requested Maximum Curre (UA) 1000	e), and monitoring para Set Maximum nt Current (uA) 2000	meters (produce Monitoring Current (uA) 558.4	d by the module Temperatur ('C)
4 High Vo Current para The table bello Channet Des 0 GEN 1 unu	CF4 Mixture	nonitoring v nnet's requeste uested Set er Powe	12 0 25 NG alues d paramet Reque: r Voltag 4400 0	ters (by the sted So e (V) Vo (V 44 0	e user), set p et M oltage Vo) 100 31	oarameters (in fonitoring oltage (V) 1833	side the HV module Requested Maximum Curre (uA) 1000 1000	e), and monitoring para Set Maximum Current (uA) 1000 1000	Monitoring Current (uA) 558.4 0	ad by the module Temperatur (°C) 26 23 23
4 High Vo Current para The table bello Channel Des 0 GEN 1 unu 2 unu 2 RM	CF4 Mixture	nonitoring va nonet's requeste vested Set Power off off	12 0 25 NG alues d paramet voltag 4400 0 0	ters (by the sted Se le (V) V4 (V 44 0 0	e user), set p et Moltage Vo 0 31 0 0 31 0 0	oarameters (in Ionitoring Oltage (V) 183:3	side the HV moduli Requested Maximum Curre (uA) 1000 1000	a), and monitoring para Set Maximum Current (uA) 1000 1000	Monitoring Current (uA) 558.4 0 0	d by the module Temperatur ('C) 26 23 23
4 High Vo Current para The table bello Channel Des 0 GEN 1 unu 2 unu 3 PM PMM	CF4 Mixture	nonitoring va nnet's requeste uested Set er Powe on off off	12 0 25 25 alues d parameter Voltag 4400 0 0 0 900	ters (by the sted St le (V) VG 44 0 0 0 90	e user), set p et Motage Vo 1000 31 0 0 0 0 000 90	Darameters (in Ionitoring ottage (V) 183.3	side the HV module Requested Maximum Curre (uA) 1000 1000 1000	e), and monitoring para Set Maximum Current (uA) 1000 1000 1000 1000	Monitoring Current (uA) 558.4 0 0 326.28	d by the module Temperatur ('C) 26 23 23 23 25
4 High Vo Current para The table bellor Channel Des 0 GEN 1 unu 2 unu 3 PM 4 PM	CF4 Mixture	nonitoring va nonitoring va nonet's requeste vested Set Power on orri orri on	12 0 25 25 25 25 25 20 25 20 20 20 20 20 20 20 20 20 20 20 20 20	ters (by the sted Se (V) V 44 0 90 90 90	e user), set p et M toltage Vo 1000 31 00 0 00 90 900 86	Darameters (in Ionitoring oltage (V) 183.3 00.1	side the HV module Requested Maximum Curre (uA) 1000 1000 1000	e), and monitoring para Set Maximum Current (uA) 1000 1000 1000 1000	Ameters (produce Monitoring Current (uA) 5584 0 0 32628 32554	ed by the module Temperatur (*C) 26 23 23 23 25 26

Figure 7.8: Screen shots of the web interface from which a user can configure the experimental set-up by providing the desired high voltage and gas mixture, as well as monitoring the each of them.

7.3. IMPLEMENTATION OF A DAQ SYSTEM

The way the client information, *i.e.* user requests, is propagated to the experiment side is done through the intermediate of a database. User requests are formatted by the web application and sent to the database in which they are stored. As explained later, the experiment side softwares are listening for new requests entering the database. When a user request is found, it is processed by the experiment side program.

For real time monitoring information, the communication between the user web application and the web server is kept open and active at all time so that the web server itself can forward information from the experiment side to the client without any action from the user.

Central Node

As already introduced, the central node is a web server handling both the communication with the user web application and with the experiment side softwares. It uses the Node.js javascript technology to store the user requests into a MySQL database, to give the user the requested web page for set-up configuration and to forward to the user the monitoring information from the experiment side.

Experiment Side

On the experiment side, a server with C/C++ softwares is located near the different modules needed for the experimental set-up. Its task is to execute user requests stored in the database as well as to fill the database with monitoring information. Among the modules with which the computer programs communicate is found a VME crate containing the high voltage module able to supply high tension (up to 6 kV) independently to 6 channels. When the program reads a user request concerning the high voltage, the formatted request is decoded and sent to the master module of the VME crate for the request execution. Additionally, the program also sends read requests every 5 seconds to the VME master module in order to read the high voltage state and inform the user about it.

The gas system is another module linked to the experiment side computer. Again, user requests concerning gas flow and mixture are forwarded to the flowmeters for proper action and read requests are also performed in order to get the flowmeter states and provide the user with the monitoring flowmeters data.

Figure 7.9 shows photographs of the experiment side server, VME crate and gas system.



Figure 7.9: Photograph of the (left) experiment side server and the VME crate and (right) gas bottles and flowmeters. The VME crate contains a master module and a high voltage module. The VME crate and the flowmeter system are linked to the server by USB cables.

The experiment side server controls the experimental set-up (high voltage, gas) and also handles the data acquisition scripts. By logging in to the experiment side server, one can launch the data acquisition remotely, and store events recorded by the Triple-GEM detector.

7.3.2 Protocols

Having now a global picture of the DAQ chain from the user side to the experiment side, the communication with the detector electronic system can be described in more details. Before, some details on the protocols used to communicate between different components of the detector readout electronic must be given. Communication protocols are the key element in inter-module communication as they represent the language spoken by each of the entities. The protocols give the rules to be followed when some information has to be exchanged between modules. Communication protocols can be used at different levels. For instance, the layer closer to the hardware need to set the rules for how to send bits of information but does not need to know anything about the meaning of the words formed by the bits it is sending. The meaning of those bits belong to a higher level for which another protocol exists giving the rules regarding how to interpret the words but saying nothing about how these words must be sent. The next sections detail four different communication protocols used in the readout electronic. Two of them, namely
the UART and I2C protocols are close to the hardware in the sense exposed above. The last one, the IPbus protocol, is a higher level protocol for interpreting the sent bits. The third one, the 8b/10b encoding, is a temporary stage of a data to be transmitted and is used for optimisation.

UART

The Universal Asynchronous Receiver/Transmitter (UART) is a simple protocol that takes bytes of data and transmits the bits sequentially on a single line of communication. More precisely, for the transmission of one byte of data, the procedure below is followed. Concerning the transmitter entity, *i.e.* the entity sending the byte, the procedure is as followed:

- 1. one line (an electrical wire on which a tension can be applied for example) is shared between the two communicating entities,
- 2. the line is in idle mode, held high (logic 1) when no data needs to be transferred,
- 3. when the transmitter has a byte of data to transfer to the other one, it issues a start bit (logic 0) on the line,
- 4. each bit of the data byte is transferred on the line (one by one),
- 5. after the last bit has been sent, the emitter sends a stop bit (logic 1)
- 6. stay in idle mode or proceed next byte.

The receiver entity at the other end of the line is following the procedure below:

- 1. the receiver listens to the line for a start bit (logic 0),
- 2. when a start bit is detected, it reads the next 8 bits on the line,
- 3. after the last bit has been read, the stop bit is expected,
- 4. the receiver returns to its listening mode.

Figure 7.10 represents the construction of a UART frame. The rate at which the bits are transmitted (baud rate, one baud corresponds to one bit per second) has been standardised to multiples and submultiples of 9600 baud, and ranges from 110 baud to 3686 400 baud. Of course, the transmitter and receiver must share the same configuration in order to communicate.



Figure 7.10: UART frame as seen on the common line between two communication entities showing the start, data and stop bits.

The UART protocol therefore provides rules telling how to transfer data bytes between two entities, but nothing about how to interpret the data bytes. Formatting rules must thus be used on top of this layer in order to be able to decode what is received. This protocol was used in the first development steps of the readout electronic system as will be explained a bit later in this chapter.

I2C

Another protocol setting the rules for how to transfer data bytes between entities is the I2C protocol for which entities are separated in masters and slaves. One master can address any of the slaves sharing the I2C bus, *i.e.* one data line (SDA) and one clock line (SCL), used for the communication. The SDA line is hold high by default and both the master and the slaves can drive this line. The SCL line on the other hand is controlled only by the master and is also hold high when no transaction is in progress.

When the master desires to send data to a slave it begins by issuing a start sequence on the I2C bus. The start and stop sequences mark the beginning and end of a transaction with the slave and are defined as a passage from high to low for the start and from low to high for the stop when the SCL line is high (see figure 7.11). The start sequence together with the stop sequences are special behaviours. Indeed, they are the only places where the SDA line is allowed to change while the SCL line is high. When transferring data, SDA must remain stable and not change while SCL is high.

Once the start sequence has been emitted, the 8 bits of a data byte are transferred sequentially and synchronously with the clock generated by the master on the SCL line, starting with the most significant bit (MSB). For every 8 bits transferred, the slave sends back an acknowledgement bit (see figure 7.11). If the receiving device sends back a low ACK bit, then it has received the data and is ready to accept another byte. If it sends back a high ACK bit then it is indicating it cannot accept any further data and the



Figure 7.11: (Top) I2C start and stop sequences for which the SDA line is allowed to change while the SCL line is held high. (Bottom) I2C frame as seen on the I2C bus showing the start, data and stop bits on the SDA line synchronised with the clock sent on the SCL line.

master should terminate the transfer by sending a stop sequence.

The I2C protocol clock can reach a speed up to about 400 kHz in the fastest modes. The fact that the slave addressing is done with a byte whose least significant bit (LSB) is the Read/Write bit means that up to $2^7 = 128$ slaves can share the same bus. This protocol of communication is extensively used in electronics and, as already announced, has been uses in our set-up as it is the protocol used by the readout chip VFAT2 (see later in this chapter). Like the UART protocol the I2C provides rules telling how to transfer data bytes between two entities, but a priori nothing about how to interpret the data byte. However, the byte of data exchanged between the master and the slave often have a direct meaning that also became part of the UART protocol.

In practice, a first sequence is sent on the I2C bus with the 7 MSB of the data being the slave address the master wants to communicate with and the less significant bit (LSB) being a Read/Write (low/high) bit indicating if the coming request is a read or write request. If it is a write request, the next transaction will be the register address where to write, and the subsequent bytes will be the data to be written. When the master wants to read from a slave, it must still first write into a specific register the address where to

read from (this is transferring the data contained in the register to be read to the specific register) and only then send a read request. Finally, a stop sequence is issued to inform the slave that the communication is terminated.

Once more, this practical usage is often encountered but not required. With the VFAT2, a slightly simpler meaning is given to the bytes of data exchanged during the communication.

8b/10b

The 8b/10b is more an encoding system than a real protocol. It consists in a mapping of any 8-bit data word to a 10-bit word such that the 10-bit word contains at least four transitions between the two logic states (from low to high or vice versa) and that the resulting flux (when more than one byte is transferred) never contains more than five times the same state (low or high) consecutively.

The 8b/10b was developed in order to bring solutions to problem arising when transmitting bytes. To understand the potential issues, let's consider the byte 11111111b. This byte contains eight ones in a raw, and therefore no transitions. This raises two problems. Firstly, some protocols of communication directly read the clock from the data byte but if no transition appear, it is of course more difficult. Secondly, this series of ones has a side effect of charging the line (for an electrical wire), which can alter the transmission. More generally, direct byte transmission is not always optimised for the transmission medium. The 8b/10b makes the clock recovery easier and avoids any charging effect. For instance, the example byte 1111111b would be mapped to the 10-bit 1010110001b or 0101001110b, according to the previous sent byte, which both have an equal number of ones and zeros leading to many transitions between logic states for easy clock recovery, and no charge effect.

The 8b/10b encoding is used in the set-up for transferring data via optical fibres as explained later.

IPbus

The IPbus protocol (see reference [118] for a complete description of the protocol and its implementation) is another example of a communication protocol that is both simple and reliable. It is based on the transfer of 32-bit data words between a master and a slave. With the IPbus protocol the master entity sends requests to the slave that replies accordingly. This protocol is therefore well suited for writing and reading registers.

When the master wants to send requests (multiple requests can be sent), it has to build an IPbus packet of 32-bit words, which consists of a 32-bit packet header, illustrated in figure 7.12, specifying the protocol version, the packet identification number, a byteorder qualifier, and a packet type, followed by the actual requests, each request being made of a 32-bit transaction header, illustrated in figure 7.12, specifying the protocol version, the request identification number, a number of 32-bit data words necessary to fully specify the request, the type of the request and an information code, followed by as many 32-bit data words needed to fully specify the request. To illustrate this, let's take the example of reading the content of n = READ_SIZE successive registers starting at address BASE_ADDRESS. The master starts by sending (by whatever mean) the IPbus 32-bit packet header with some packet ID, followed by the 32-bit transaction header specifying the request type as READ (type ID = 0), the number of registers to be read and the info code 0xf to specify that it is a request. Then the 32-bit data word BASE_ADDRESS, corresponding to the start register address is sent. The slave, after decoding the full master request, replies by sending first an IPBus 32-bit packet header with the same packet ID, followed by a transaction header, the info code 0x0 to specify that it is a response and the $n = \text{READ}_\text{SIZE}$ 32-bit data words corresponding to the content of the registers at address BASE_ADDRESS up to BASE_ADDRESS + READ_SIZE. Figure 7.13 illustrates the reading request and response.

According to the type ID, different requests can be sent among which read transaction (see example above), non-incrementing read transaction, write transaction and non-incrementing write transaction. A complete reliability mechanism can also be implemented thanks to the ID numbers and info code contained inside the requests and responses so that it is possible to check if things went wrong or not.

The IPbus protocol is used in the first higher layers of communication between the experiment side server and the detector electronics as described later in this chapter. Unlike the UART or I2C protocols described above the IPbus protocol does not give the rules for how the 32-bit words should be transferred, it only says how they should be interpreted. It is therefore necessary to implement the protocol on top of a lower layer of communication that will take care of transmitting the data.

7.3.3 Data Readout of the Cosmic Muons Test Bench

In this section, three set-up differing in their electronic readout chain are presented. The first of them uses a Xilinx development board as the intermediate entity between the experiment side server and the VFAT2. This set-up corresponds to the first development towards the final design architecture. The second one uses two Gigabit Link Interface

	31	24	23	16	15	8	7	0
	Protocol version	Rsvd.		Packet II	D (16 bits)		Byte-order qualifier	Packet type
word 0	0x2	0x0		0x0 -	0xffff		0xf	0x0 - 0x2

IPbus Packet Header

IPbus Transaction Header

	31 28	27 10	15	8	7 4	3 0
	Protocol version	Transaction ID (12 bits)		Words (8 bits)	Type ID	Info Code
word 0	0x2	0x0 - 0xfff			Read / Write	0x0 - 0xf

Figure 7.12: IPbus (top) packet and (bottom) transaction headers format and content.

Request

	31	24	23	16	15 8	7	0
word 0	0x2	0×0	Packet ID			0xf	0×0
word 1	0x2	Transaction ID			Words = READ_SIZE	Type ID = 0	0xf
word 2	BASE_ADDRESS						

Response



Figure 7.13: Construction of a READ_SIZE read transaction with (top) the master request showing the IPbus packet header (word 0), the IPbus transaction header of type ID 0 for read (word 1), and the register address where to start reading from (word 2), and (bottom) the slave response showing the IPbus packet header (word 0), the IPbus transaction header of type ID 0 for read response (word 1) and the $n = READ_SIZE$ registers contents (word 2 to word n).

Board (GLIB) connected with optical fibres, a technology to be used for the final design. The third and last one uses a GLIB inside a μ TCA crate and an Opto-hybrid board connected with optical fibres to the GLIB. This last set-up is very similar to the architecture that is going to be used for the final design.



Figure 7.14: Schematic representation of the first set-up used for the test bench at the IIHE. The set-up regroups the experiment side server linked to a Xilinx SP601 development board, in turn connected to the VFAT2 chip on the Triple-GEM prototype. Two PMs are found on both sides of the Triple-GEM detector and provide the trigger signal after small logic done with a NIM crate. Not shown are the power supply of the SP601 board, the high voltage in the VME crate for the Triple-GEM and PM, and the gas supply to the Triple-GEM.

For each set-up, two photomultipliers (PM) are placed on both sides of the Triple-GEM prototype detector in order to give the trigger signal to the VFAT2. All three architectures allow the proper working of the IIHE test bench with cosmic muons using the Triple-GEM prototype introduced above. Let us already remark that the set-ups can be slightly modified in order to work with a iron-55 radioactive source. In that case, the PMs are not needed and the trigger signal is directly obtained using the raw strips signal, amplified by a preamplifier and an amplifier, and properly shaped before sent to the VFAT2.

First Development Set-up

As can be seen in figure 7.14, the rather simple first set-up consists in four main entities: the experiment side server hosting the C++ softwares, a Xilinx SP601 development board [119] with a Spartan 6 FPGA, the Triple-GEM prototype detector and its VFAT2 frontend electronic, and the triggering system made of two PMs and a NIM crate to handle the trigger logic. The goal of this set-up is to quickly establish the communication with the VFAT2 chip on the detector and be able to send I2C commands through the SP601 board to the VFAT2. To this aim, user requests first need to be transferred to the SP601 FPGA. The IPbus protocol has been implemented through a C++ software on the experiment side server together with a friendly user interface. The IPbus transactions, are then sent by serial link (USB) using the UART protocol to the board.

On the SP601 board, in order to be able to read from the serial input, a soft processor core called MicroBlaze [120] has been implemented directly on the FPGA. The MicroBlaze is a full VHDL entity as any others and can therefore be used in parallel with any other VHDL entity implemented on the FPGA. This allows one to have an operating system such as a linux kernel inside the FPGA with programs written in C to handle the communication coming from the server, decode this information and transmit it to other components of the FPGA. This is exactly what has been realised for the first set-up. IPbus transactions are sent via UART to the MicroBlaze on the FPGA. The C program running on the FPGA, when receiving an IPbus transaction, decodes it, and forwards the decoded request to an I2C core implemented on the same FPGA.

Finally, using a FPGA Mezzanine Card (FMC) developed specially to allow the connection of the VFAT2 hybrid and LEMO connections to the GLIB board, the I2C core on the FPGA, an I2C master device, can in turn forward the request to the VFAT2 chip. In addition, the FMC board also allows the interface with the NIM crate. Indeed, when the NIM crate sends a trigger signal, it is transmitted to the FPGA thanks to LEMO connectors located on the FMC board, and the corresponding trigger I2C command is sent to the VFAT2 again via the FMC.

When the VFAT2 receives an I2C command, its internal logic decodes it and executes the request, which can be a reconfiguration of its internal registers, for instance, to change the latency parameter or the threshold applied on the strips. When the VFAT2 is sent a trigger signal to keep the triggered event, the VFAT2 fetches the corresponding event and places it in its SRAM2 as explained earlier in this chapter. Once some triggered events are found in SRAM2, they are sent to the SP601's FPGA via a dedicated line through the FMC board on the SP601. The FPGA then forwards the triggered events to the server that stores them into files on the disk.

As far as the trigger signal is concerned, it results from the coincidence of the two PM signals (see figure 7.15). This is achieved by sending the PM signals to a NIM crate that outputs a logic one when the two PM signals are low (the response signal of a PM to the passage of a particle is negative) in coincidence.

As already mentioned, this architecture made it possible to establish quickly, thanks to the advantages offered by the MicroBlaze, the communication with the VFAT2 electronic. The UART connection as well as the MicroBlaze however limit the possibilities of the



Figure 7.15: Photograph of the Triple-GEM prototype detector (at the centre) with a small photomultiplier above it (in black), not visible a larger photomultiplier is located beneath the wooden plate holding the Triple-GEM detector. The VFAT2 hybrid connected to the detector and its grey ribbon going to the next element in the readout chain can also be seen.

set-up. Indeed the serial communication does not allow the high transaction rate needed and even though the MicroBlaze has many advantages, being a soft core system, the full hardware time prediction is lost. Still, tests with cosmic muons have been successfully performed with this set-up and gave the green light to change the set-up closer to the desired architecture reusing many of the developments performed in this first step.

Second Development Set-up

In the second set-up, improving the first one presented in the previous section, the IPbus communication is now handled with UDP/IP (Internet Protocol). User requests are transferred to a first Gigabit Link Interface Board (GLIB) [121] on which optical links are plugged and connected on the other side to a second GLIB hosting the FMC mezzanine board interfaced with the VFAT2 on the Triple-GEM detector as in the first set-up. Trigger signal is obtained in the same way by use of two PM, and sent to the second GLIB's FPGA via the FMC.

Now, when an IPbus transaction is sent to the first GLIB, it is decoded by an IPbus core implemented on the first GLIB's FPGA, before being sent through an optical link



Figure 7.16: Schematic representation of the second set-up used for the test bench at the IIHE. The set-up regroups the experiment side server linked to an ethernet switch to which a first GLIB is also connected that is, in turn, connected to a second GLIB via optical links. The second GLIB is itself linked to the VFAT2 chip on the Triple-GEM prototype. Two PMs are found on both sides of the Triple-GEM detector and provide the trigger signal after small logic done with a NIM crate. Not shown are the power supplies of the two GLIB, the high voltage in the VME crate for the Triple-GEM and PM, and the gas supply to the Triple-GEM.

to the FPGA on the second GLIB. To transmit data through the optical link, the 8b/10b encoding is used for better performance. On the second GLIB, the FPGA handles both the communication with the first one by the optical links, and with the VFAT2 chip as done with the SP601 board in the previous set-up.

With the use of the optical links the set-up is closest to the desired design. Indeed, in CMS, GLIB boards (or similar board) are going to be located in a crate inside a cavern close by the CMS detector and linked to the Triple-GEM electronics on CMS by optical links. Furthermore, the use of UDP/IP for IPbus transaction is also what is going to be used in the final design allowing for much faster (of the order of the Gb/s) transaction rates than in the previous set-up.

This second step made it possible to develop the optical link interface on both FPGA as well as the implementation of the IPbus core on the first GLIB. The I2C core used in the first set-up was reused. Same tests were performed with cosmic muons with the same successful results.



Figure 7.17: Schematic representation of the final set-up used for the test bench at the IIHE and for test beams at CERN. The set-up groups the experiment side server linked to an ethernet switch to which a μ TCA crate hosting a GLIB is also connected. The GLIB is connected via optical links to the Opto-Hybrid board plugged into the GEB in turn linked to the VFAT2 hybrid on the Triple-GEM prototype. Two PMs are found on both sides of the Triple-GEM detector and provide the trigger signal after small logic done with a NIM crate. Not shown are the power supplies of the the μ TCA crate, Opto-Hybrid board and GEB, the high voltage in the VME crate for the Triple-GEM and PM, and the gas supply to the Triple-GEM.

Third Set-up

The last set-up, illustrated in figure 7.17, is another upgrade of the previous one. The first GLIB of the previous set-up is now located inside a μ TCA crate connected to the same network as the experiment side server. The second GLIB that was connected to the Triple-GEM prototype is now replaced by the so-called Opto-Hybrid board plugged in a GEM Electronic Board (GEB) to which the VFAT2 hybrid is connected by a ribbon of electrical wires. The rest of the set-up is the same.

The GEB is a large Printed Circuit Board (PCB) of trapezoidal shape with a big base of length 45 cm, a small base of length 28 cm and a hight of 100 cm, dedicated to

root the VFAT2 signals to the Opto-Hybrid board plugged in it. Figure 7.21 shows an exploded schematic view of a CMS Triple-GEM module with the drift cathode, the three large GEM foils and the GEB. The GEB is divided in 24 regions: 3 segmentations in the azimuthal direction and 8 segmentations in the radial direction. Each of the 24 regions contains a total of 128 strips, aligned in the radial direction, to which the CMS or TOTEM VFAT2 hybrids can be connected.

For this last architecture, the developments done for the previous versions were reused with minimal modifications.

Results for cosmic muons have been obtained using the previous set-ups and are presented in the next section.

7.3.4 Cosmic Muons Test Bench Results

Before being able to record cosmic muon events with the set-ups discussed above, the VFAT2 threshold and latency parameters have to be properly configured.

VFAT2 Threshold Scan

The VFAT2 chip is a binary chip in the sense that it will assign a logic zero or one (hit or missed) to each of its 128 channels by mean of its internal comparator. Electronic noise (random fluctuation in an electrical signal) is always present and a minimum value for the electric amplitude signal has to be established. This threshold value, used by the VFAT2 comparator, is one of the VFAT2 configurable parameters. It is thus possible to adjust the threshold value assigned to all the strips in order to reduce the electronic noise contamination.

When scanning the different threshold values, independently of the PM signals random trigger signals are sent to the VFAT2 chip. Bypassing the PM signal is done on purpose to avoid recording only true cosmic muon events that would bias the electronic noise study performed with the threshold scan. A fake trigger signal is therefore sent at regular interval in order to tell the VFAT2 chip to store the state of its 128 channels into its SRAM2 memory from which the data are extracted. A total of 1000 events is stored this way for each threshold value. For each of those events, a logical OR is performed on the 128 strips. An event is then considered hit as soon as one strip has been assigned a logic one by the VFAT2 chip.

At the lowest threshold values, we expect the strips to be seen as hit (logic one) all the

time due to unavoidable electronic noise. As the threshold is increased, the probability for the electronic noise on a strip to lead to a hit signal is decreasing sharply. At higher threshold values, only a small fraction of the strips is expected to be activated by the electronic noise.

This is exactly what has been observed with our set-up, as shown in figure 7.18 showing the ratio of hit events to the total number of events triggered as a function of the threshold value. A plateau very close to unity is observed for the lowest threshold values as expected. After increasing the value of the VFAT2 threshold parameter of a few units, a sharp decrease is observed in accordance with the fact that the electronic noise is less and less likely to overpass the threshold value. A long tail is observed for larger threshold values due to residual higher than average electronic noise fluctuations and real signals.

From the present result, a threshold value of about 15 (in VFAT2 units) would keep electronic noise below 10% while true muon signals are expected to overpass this value in most cases. Increasing the threshold would of course further reduce the electronic noise but would decrease the muon detection efficiency at the same time. On the other hand, a higher muon detection efficiency can be reached by decreasing the threshold value with the price of a larger electronic noise contamination. The choice for the threshold value is by consequence dictated by a compromise between purity and efficiency.

Latency Scan

Once the threshold parameters of the VFAT2 chip has been adjusted to the desired level, and only then, the latency parameter value can be adjusted for the set-up in use. As mentioned, the latency parameter characterises the time delay for the trigger signal to be received by the VFAT2 chip.

To understand this, let us suppose a muon passes through the whole set-up (the first PM, the Triple-GEM prototype and the second PM). The PM1 signal is sent to the NIM crate slightly before the PM2 signal. During this period of time the charge collection is happening on the strip of the Triple-GEM detector that gives rise to the VFAT2 response after a short time. The VFAT2 chip thus stores the 128 channels state to its SRAM1 memory. During the same time, the PM signals are treaded by the NIM logic, from which a trigger accept signal is output and sent by the intermediate board (Opto-Hybrid or second GLIB) to the VFAT2. This process, *i.e.* the treatment of the PM signals took a little time, t_{latency} , during which the VFAT2 chip has carried on filling its SRAM1 with events at a rate given by the external LHC clock of f = 40 MHz. The triggered event therefore lies at position $t_{\text{latency}} \times f$ in SRAM1 and the VFAT2, after receiving the



Figure 7.18: Threshold scan result for the IIHE test bench. For each value of the VFAT2 threshold parameter, a total of 1000 fake trigger signals is sent to the VFAT2. From the 1000 events, the ratio of hit events (logical OR between the 128 channels) is plotted. The error bars represent the statistical uncertainty.

trigger accept signal, has to move this specific event to its SRAM2 memory for further processing. The time t_{latency} even though fixed once the set-up (electrical wires, high tensions, positions, threshold parameters, ...) fixed, can hardly be computed and must be determined using the detector itself.

To this aim, a latency scan consisting in probing the various SRAM1 depths is performed. For every latency value t_{latency} configurable on the VFAT2 chip, a total of 1000 events triggered by the PMs is sent to the VFAT2. For each of those trigger accept signals, the VFAT2 chip fetches the event located at position $t_{\text{latency}} \times f$ in SRAM1 to move it to its SRAM2 memory from which the event is extracted. As for the threshold scan, a logical OR operation is performed on all the 128 channels and the ratio of hit events to the total number of triggered events for the particular latency value is plotted.

The result is shown in figure 7.19. First of all, a peak at a latency value of about 29 clock cycles is observed meaning that the maximum ratio of hit events is obtained for this particular value. On both sides of the peak the low ratio values indicates that the events fetched inside the SRAM1 VFAT2 memory where not hit in most of the cases and proves



Figure 7.19: Latency scan result. For each value of the VFAT2 latency parameter, a total of 1000 real trigger signals obtained by the coincidence of the two PM is sent to the VFAT2 that fetches and move the event determined by the latency parameter from its SRAM1 to its SRAM2 memories. From the 1000 events, the ratio of hit events (logical OR between the 128 channels) is plotted. The error bars represent the statistical uncertainty.

that the selected event was either happening before or after the real event triggered by the PM. From the plot, the time for the PM trigger signal to be received by the VFAT2 chip seems to be of 29×25 ns = 0.725 µs. Secondly, the ratio value at the peak is only of about 26%. This number is directly linked to the efficiency of the Triple-GEM detector since the corresponding events are very likely to be real muon events triggered by the coincidence of two PM. However, the efficiency is already very reduced by a geometrical acceptance factor that can be computed for the given set-up. Indeed, the bottom PM is very large compared to the area spanned by the Triple-GEM 128 channels under consideration. A muon with sufficiently large incoming angle (with respect to the vertical) can cross both PM without necessarily passing through the active strips of the Triple-GEM prototype. Further investigation would be required to evaluate the Triple-GEM detector efficiency.



Figure 7.20: Cluster size result showing the number of events as a function of the number of contiguous hit strips found among the 128 channels, for 140 recorded events (several clusters can occur for each event).

Cluster Size

With the set-up and the VFAT2 chip properly configured, cosmic muon events triggered by the PM placed on top and below the Triple-GEM detector can be recorded. For illustration purposes, figure 7.20 shows the number of contiguous hit strips, called the cluster size, for 140 triggered events. As can be seen a bit more than 50% of the events have only a single hit strip. For the rest of the events, the great majority of the cosmic muons lead to a two-strip wide cluster and some to three or four-strips wide clusters. Only a couple of events have more than 4 hit strips. This distribution is likely to be resulting from the $\cos^2 \theta$ shape distribution of the cosmic muon flux, where θ is the angle with respect to the vertical. In order to make a statement regarding the expected and the observed cluster size, further investigation is required considering, among other, the geometrical acceptance of the detector.

7.4 Tests Beam

The last architecture developed for the cosmic muon test bench was then used in test beam at CERN at the end of year 2014. This allowed to test the developed DAQ in another environment and also to use of the S-bits trigger signals provided by the VFAT2 as explained below.

7.4.1 Set-up

In place of the Triple-GEM 10×10 prototype detector, a full Triple-GEM module similar to what is going to be installed in CMS as the GEM detector as shown in figure 7.21 was used. The trigger signal was furnished to the Triple-GEM front-end electronics thanks to large PM placed at the beginning of the experimental set-up. The readout electronic architecture is however the same than the one used for the test bench at the IIHE.

Additionally to the tracking feature of the VFAT2 chip exploited in the cosmic muon test bench, its triggering capability has also been tested in test beam at CERN. The VFAT2 chip can perform a fast logical **OR** operation on its channels for every event entering its SRAM1 memory, *i.e.* at the LHC rate of 40 MHz. This feature is going to be used for the L1 trigger of CMS. For this purpose, the Opto-Hybrid board sends the VFAT2 trigger signal to the CMS trigger system by mean of the optical links.

The beam used to provide the incident muons and pions are secondary beams obtained by sending bunches of protons produced by the SPS onto a fixed target of heavy material, and guiding by means of magnets the produced particles up to the experiment site. Essentially pions are produced during the collision. The magnet configuration selects particles with a momentum of about 150 GeV. Additionally a thick concrete block can be inserted on the trajectory of the pion beam in order to stop the pions while the muons (from earlier pion decays) pass through the block. This way, a beam of muons can be obtained (due to the presence of this thick wall, the muon beam was found to be approximately twice as large as the pion beam in its transverse size). Both configurations were used during the test beam.

7.4.2 Results

The first step in the test beam before recording data was to characterise the front-end electronics inside the new environment. A threshold scan followed by a latency scan have thus been performed in a same way as explained in above sections. The results are shown



Figure 7.21: (Left) Exploded schematic view of a CMS Triple-GEM module and (right) four similar Triple-GEM modules placed for the test beam. At the bottom of the second module from the right, one can see the orange optical links leaving the Opto-Hybrid board to join the GLIB board inside a μ TCA crate close by the set-up. The beam of muons or pions is coming perpendicularly from the left on this photograph and the set-up position is adjusted by hand in order to irradiate the desire part of the detectors in place.

in figure 7.22 on page 222.

From the threshold scan result (top panel in figure 7.22), it can be seen that the noise level is significantly higher than what was measured for the IIHE cosmic muon test bench. A higher threshold, of value 25 in VFAT2 units, has therefore to be applied to maintain the noise contamination at the same maximum level of 10%. It has to be noted that the noise level is very sensitive to the environment. In order to reduce the noise at the IIHE bench, a Faraday shield was build afterward and a better grounding was realised.

The latency scan result (bottom panel in figure 7.22) exhibits a peak at 21 clock cycles, *i.e.* $0.525 \,\mu$ s, which is smaller than what was obtained for the cosmic muon test bench at the IIHE. The two values are however not comparable because entirely dependent on the external trigger signal delay that was obtained by slightly different means in both cases. Additionally, different high tensions were also applied to the GEM detector that also affect the latency value.

With the characterisation of the front-end electronics completed, triggered events can be recorded. For illustration purposes, a beam profile showing the number of times each channel is hit has been performed on the recorded events and is shown in figure 7.23. From the plot, about 40 bins are empty indicating probable dead channels. A possible explanation for this can be given by the fact that the VFAT2 chips used for the test beam may have been damaged in previous test beams. Apart from these dead channels, some channels also show abnormal high activity. This could also be explained by the fact these channels are more noisy than the other ones. At the time of writing the VFAT2 chips are being investigated further with the IIHE cosmic muon test bench. The bell shape however clearly indicates the position of the beam with respect to the detector.

7.5 Conclusions

Before having the green light to install new detectors inside CMS in view of its upgrade, proof of concept must be established. By designing the DAQ system from scratch for the Triple-GEM detector and by testing it in real situation, such a proof can be obtained.

All the results got with both the IIHE test bench and the CERN test beam set-ups have confirmed the feasibility and validity of the CMS Triple-GEM readout chain made of the GEB, the Opto-Hybrid and the GLIB within the μ TCA. This design for the DAQ system has thus been validated and is now submitted to the CMS collaboration for approval, as part of the Technical Design Review.



Figure 7.22: (Top) Threshold scan result and (bottom) latency scan result, for the CERN test beam set-up.



Figure 7.23: Muon beam profile result showing the number of time a strip is hit for each of the 128 channels.

Chapter 8

Conclusions

The Standard Model is today a firmly established theory. Since about 40 years it has been regularly confirmed in its gauge approach and has gained in accuracy over all these decades. Precision measurements mainly from detectors at accelerators allowed detailed confrontations of the Standard Model predictions on all its aspects. This happened especially a few years ago with the discovery of the Higgs boson, which had been looked for since it was predicted back in the 1960s by Brout and Englert, and Higgs, but was only experimentally accessible with the world records energies of the LHC accelerator.

With the highest energies ever produced by man-made machine, the LHC has opened new horizons where to test the structure of elements and their interactions. The Standard Model is being scrutinised and a large number of models Beyond the Standard Model theories are currently being confronted to measurements.

At the LHC, jets are present in almost all studied processes with rates higher than ever before. To understand and to be able to describe their production up to a high level of accuracy is therefore crucial. One way to achieve that is to study jets production in association with a well known particle produced in abundance and easy to detect. This is what is done in the data analysis presented in this thesis where the production of an electroweak Z boson in association with jets has been measured in phase space regions never reached before (or partially and statistically limited) in previous experiments. The differential cross sections have been measured as a function of the jet multiplicity and jets kinematics variables, transverse momentum, pseudorapidity, scalar sum of all jets transverse momentum and two leading jets invariant mass. These observables are of great interest in many respects. The jet multiplicity is providing a direct result testing the capability of a MC generator in predicting a large number of QCD radiations while limited by the complexity of calculating such radiations with matrix element perturbative techniques. The jet transverse momentum distributions allow for testing the hardness of the QCD radiations as well as the matching scheme performed in MC simulations when combining ME and PS predictions. Similarly for the pseudorapidity observables from which it can be seen whether or not a MC generator correctly populates the pseudorapidity phase space. The scalar sum of jets traverse momenta gives hints for searches like supersymmetry in all-hadronic final states for which events are expected to give rise to many jets accompanied with large missing transverse energy. The present measurements therefore also provide to such analyses an accurate estimation of the irreducible Z + jets background. Finally, the dijet invariant mass distribution is relevant for new physics scenarios in which possible resonant dijet production can take place.

The measurements have been performed using a data sample of total integrated luminosity of $19.6 \,\mathrm{fb}^{-1}$ collected by the CMS detector in year 2012 at the proton-proton collisions from the LHC with a centre-of-mass energy of 8 TeV. The measured differential cross sections have been compared to three MC generators predictions: the LO multi-leg MADGRAPH 5 matrix element generator interfaced with PYTHIA 6 for parton shower and hadronisation, the NLO multi-leg SHERPA 2 prediction implementing its own parton showering and hadronisation models, and the NLO multi-leg MADGRAPH5 AMC@NLO matrix element generator interfaced with PYTHIA 8 for parton shower and hadronisation. All three predictions perform really well when compared to data with the NLO generators providing an even better description at some phase space regions in the jet transverse momentum and pseudorapidity distributions. However some other regions, such as low dijet invariant mass or low scalar sum of jets transverse momenta are found to be underestimated by the simulations. These preliminary results are shown in conferences since 2014 and available in the CMS Physics Analysis Summary SMP-13-007 [110]. A reference CMS publication is in preparation including the present results, more differential distributions and jet and Z angular correlations, and expected to be published in a near future.

The LHC is being upgraded at the time of writing for larger centre-of-mass collisions, 13 TeV, and also for higher instantaneous luminosity with an average number of protonproton interactions greater than a hundred per bunch crossing foreseen at the highest values. Performing the Z + jets analysis in these new conditions is one of the High Priority Analyses of CMS. Providing both higher statistics and access to new phase space regions, the first 13 TeV collision runs are eagerly awaited by scientists.

Such a machine upgrade however also demands a detector upgrade. The CMS detector has therefore been improved essentially to support the higher particle rates but still, the current state of the detector is not good enough to handle the largest particle rates expected in about 4 years from now and new technologies have to be studied to fulfil the requirements. The second part in this thesis, more hardware oriented, describes a complete DAQ system design and implementation to test the Triple-GEM detectors to be installed for the next CMS upgrade. The full readout chain composed of the GEB connected to the Triple-GEM detector, the Opto-Hybrid board collecting the GEB signals and the GLIB board inside a μ TCA crate linked to the detector electronics by optical link, has been tested with a set-up developed to this aim at the IIHE. The same version of the DAQ system has also been tested in beams at CERN. Results show the feasibility of this design are now submitted to the CMS collaboration for approval, as part of the Technical Design Review.

To conclude, the scientific method which has already led to major discoveries in the past is still being used to understand the physics taking place at very high energy. To this aim laboratory experiments are being performed at CERN with the LHC machine and detectors such as CMS. New discoveries as the Higgs boson discovery in 2012 as well as every step forward in our understanding of particle interactions requires a long and complex experimental work. This thesis participated to this large scale effort at two levels, the analysis of data taken in 2012 and the preparation of the CMS upgrade.

Bibliography

- R. K. Ellis, W. J. Stirling, and B. Webber, QCD and collider physics Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 2003.
- [2] F. Halzen and A. Martin, Quarks and leptons: an introductory course in modern particle physics. Wiley, 1984.
- [3] R. Field, *Applications of perturbative QCD*. Frontiers in physics, Addison-Wesley, The Advanced Book Program, 1989.
- [4] M. E. Peskin and D. V. Schroeder, An introduction to quantum field theory. Advanced book program, Westview Press Reading (Mass.), 1995.
- [5] A. Einstein, "Die Grundlagen der Allgemeinene Relativitätstheorie. (German) [The foundations of the Theory of General Relativity]," *j-ANN-PHYS-1900-4*, vol. 354, no. 7, pp. 769–822, 1916.
- [6] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Phys.Lett.*, vol. B716, pp. 30–61, 2012.
- [7] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons," *Phys.Rev.Lett.*, vol. 13, pp. 321–323, 1964.
- [8] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," *Phys. Rev. Lett.*, vol. 13, pp. 508–509, 1964.
- [9] V. Rubakov, Classical theory of gauge fields. Princeton University Press, 2002.
- [10] S. Larin and J. Vermaseren, "The Three loop QCD Beta function and anomalous dimensions," *Phys.Lett.*, vol. B303, pp. 334–336, 1993.
- [11] CMS Collaboration, "Constraints on parton distribution functions and extraction of the strong coupling constant from the inclusive jet cross section in pp collisions at $\sqrt{s} = 7$ TeV," arXiv:1410.6765, 2014.

- [12] S. Drell and T.-M. Yan, "Massive Lepton Pair Production in Hadron-Hadron Collisions at High-Energies," *Phys. Rev. Lett.*, vol. 25, pp. 316–320, 1970.
- [13] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, "High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO," *Phys.Rev.*, vol. D69, p. 094008, 2004.
- [14] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, "FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order," *Comput. Phys. Commun.*, vol. 182, pp. 2388–2403, 2011.
- [15] H1 Collaboration, "Inclusive Deep Inelastic Scattering at High Q² with Longitudinally Polarised Lepton Beams at HERA," JHEP, vol. 1209, p. 061, 2012.
- [16] ZEUS Collaboration, "Measurement of high-Q2 neutral current deep inelastic e^+p scattering cross sections with a longitudinally polarized positron beam at HERA," *Phys.Rev.*, vol. D87, no. 5, p. 052014, 2013.
- [17] CMS Collaboration, "Measurement of the differential and double-differential Drell-Yan cross sections in proton-proton collisions at $\sqrt{s} = 7$ TeV," *JHEP*, vol. 1312, p. 030, 2013.
- [18] Y. L. Dokshitzer, "Calculation of the Structure Functions for Deep Inelastic Scattering and e⁺e⁻ Annihilation by Perturbation Theory in Quantum Chromodynamics.," *Sov.Phys.JETP*, vol. 46, pp. 641–653, 1977.
- [19] V. Gribov and L. Lipatov, "Deep inelastic *ep* scattering in perturbation theory," *Sov.J.Nucl.Phys.*, vol. 15, pp. 438–450, 1972.
- [20] L. Lipatov, "The parton model and perturbation theory," Sov.J.Nucl.Phys., vol. 20, pp. 94–102, 1975.
- [21] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language," Nucl. Phys., vol. B126, p. 298, 1977.
- [22] G. Curci, W. Furmanski, and R. Petronzio, "Evolution of Parton Densities Beyond Leading Order: The Nonsinglet Case," *Nucl. Phys.*, vol. B175, p. 27, 1980.
- [23] W. Furmanski and R. Petronzio, "Singlet Parton Densities Beyond Leading Order," *Phys.Lett.*, vol. B97, p. 437, 1980.
- [24] A. Mitov, S. Moch, and A. Vogt, "NNLO splitting and coefficient functions with time-like kinematics," *Nucl. Phys. Proc. Suppl.*, vol. 160, pp. 51–56, 2006.
- [25] T. Gleisberg, S. Hoeche, F. Krauss, A. Schalicke, S. Schumann, et al., "SHERPA 1. alpha: A Proof of concept version," *JHEP*, vol. 0402, p. 056, 2004.

- [26] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, "MadGraph 5 : Going Beyond," JHEP, vol. 1106, p. 128, 2011.
- [27] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, et al., "Event generation with SHERPA 1.1," JHEP, vol. 0902, p. 007, 2009.
- [28] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual," *JHEP*, vol. 0605, p. 026, 2006.
- [29] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, et al., "An Introduction to PYTHIA 8.2," Comput. Phys. Commun., vol. 191, pp. 159–177, 2015.
- [30] S. Alioli, P. Nason, C. Oleari, and E. Re, "Vector boson plus one jet production in POWHEG," *JHEP*, vol. 1101, p. 095, 2011.
- [31] M. Myska, "Searches for Double Parton Scattering at the LHC," PoS, vol. EPS-HEP2011, p. 466, 2011.
- [32] CMS Collaboration, "Study of double parton scattering using W + 2-jet events in proton-proton collisions at $\sqrt{s} = 7$ TeV," *JHEP*, vol. 1403, p. 032, 2014.
- [33] H. Murayama, I. Watanabe, and K. Hagiwara, *HELAS: HELicity amplitude sub*routines for Feynman diagram evaluations. National Laboratory for High Physics, 1992.
- [34] P. de Aquino, W. Link, F. Maltoni, O. Mattelaer, and T. Stelzer, "ALOHA: Automatic Libraries Of Helicity Amplitudes for Feynman Diagram Computations," *Comput. Phys. Commun.*, vol. 183, pp. 2254–2263, 2012.
- [35] S. Catani, S. Dittmaier, M. H. Seymour, and Z. Trocsanyi, "The Dipole formalism for next-to-leading order QCD calculations with massive partons," *Nucl. Phys.*, vol. B627, pp. 189–265, 2002.
- [36] T. D. Gottschalk, "An Improved Description of Hadronization in the QCD Cluster Model for e⁺e⁻ Annihilation," Nucl. Phys., vol. B239, p. 349, 1984.
- [37] T. D. Gottschalk and D. A. Morris, "A New Model for Hadronization and e^+e^- Annihilation," *Nucl. Phys.*, vol. B288, p. 729, 1987.
- [38] B. Webber, "A QCD Model for Jet Fragmentation Including Soft Gluon Interference," Nucl. Phys., vol. B238, p. 492, 1984.
- [39] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations," *JHEP*, vol. 1407, p. 079, 2014.

- [40] N. D. Christensen and C. Duhr, "FeynRules Feynman rules made easy," Comput. Phys. Commun., vol. 180, pp. 1614–1641, 2009.
- [41] M. Bengtsson and T. Sjostrand, "Coherent Parton Showers Versus Matrix Elements: Implications of PETRA - PEP Data," *Phys.Lett.*, vol. B185, p. 435, 1987.
- [42] M. H. Seymour, "Matrix element corrections to parton shower algorithms," Comput. Phys. Commun., vol. 90, pp. 95–101, 1995.
- [43] G. Miu and T. Sjostrand, "W production in an improved parton shower approach," *Phys.Lett.*, vol. B449, pp. 313–320, 1999.
- [44] S. Catani, F. Krauss, R. Kuhn, and B. Webber, "QCD matrix elements + parton showers," *JHEP*, vol. 0111, p. 063, 2001.
- [45] F. Krauss, "Matrix elements and parton showers in hadronic interactions," JHEP, vol. 0208, p. 015, 2002.
- [46] UA2 Collaboration, "Associated production of W's and Z's with jets in UA2," Conf. Proc., vol. C850313, p. 431, 1985.
- [47] CMS Collaboration, "Measurements of jet multiplicity and differential production cross sections of Z + jets events in proton-proton collisions at $\sqrt{s} = 7$ TeV," *Phys.Rev.*, vol. D91, no. 5, p. 052008, 2015.
- [48] CMS Collaboration, "Differential cross section measurements for the production of a W boson in association with jets in proton-proton collisions at $\sqrt{s} = 7$ TeV," *Phys.Lett.*, vol. B741, p. 12, 2015.
- [49] ATLAS Collaboration, "Measurement of the production cross section of jets in association with a Z boson in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector," *JHEP*, vol. 1307, p. 032, 2013.
- [50] ATLAS Collabortion, "Measurements of the W production cross sections in association with jets with the ATLAS detector," Tech. Rep. ATLAS-CONF-2014-035, CERN, Geneva, Jul 2014.
- [51] CDF Collaboration, "Measurement of differential production cross section for Z/γ^* bosons in association with jets in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," *Phys.Rev.*, vol. D91, no. 1, p. 012002, 2015.
- [52] D0 Collaboration, "Measurements of inclusive W + jets production rates as a function of jet transverse momentum in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," *Phys.Lett.*, vol. B705, pp. 200–207, 2011.

- [53] R. Voss and A. Breskin, "The CERN Large Hadron Collider, accelerator and experiments," 2009.
- [54] CMS Collaboration, "The CMS experiment at the CERN LHC," JINST, vol. 3, p. S08004, 2008.
- [55] CMS Internal, "CMS tracker detector performance results." https://twiki.cern. ch/twiki/bin/view/CMSPublic/DPGResultsTRK.
- [56] CMS Collaboration, "Commissioning and Performance of the CMS Pixel Tracker with Cosmic Ray Muons," *JINST*, vol. 5, p. T03007, 2010.
- [57] R. Bremer, "Integration of the end cap TEC + of the CMS silicon strip tracker," 2008.
- [58] CMS Collaboration, "CMS: The electromagnetic calorimeter. Technical design report," 1997.
- [59] CMS Collaboration, "Addendum to the CMS ECAL technical design report: Changes to the CMS ECAL electronics," 2002.
- [60] CMS Collaboration, The CMS hadron calorimeter project: Technical Design Report. Technical Design Report CMS, Geneva: CERN, 1997.
- [61] G. Baiatian *et al.*, "Energy Response and Longitudinal Shower Profiles Measured in CMS HCAL and Comparison With Geant4," 2007.
- [62] G. Bayatian, A. Sirunian, I. Emelyanchik, V. Massolov, N. Shumeiko, et al., "Design, performance and calibration of the CMS forward calorimeter wedges," *Eur.Phys.J.*, vol. C53, pp. 139–166, 2008.
- [63] R. Brun and F. Rademakers, "ROOT: An object oriented data analysis framework," *Nucl.Instrum.Meth.*, vol. A389, pp. 81–86, 1997.
- [64] S. Agostinelli et al., "GEANT4: A Simulation toolkit," Nucl.Instrum.Meth., vol. A506, pp. 250–303, 2003.
- [65] W. Adam, B. Mangano, T. Speer, and T. Todorov, "Track Reconstruction in the CMS tracker," Tech. Rep. CMS-NOTE-2006-041, CERN, Geneva, Dec 2006.
- [66] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET," CMS Physics Analysis Summary CMS-PAS-PFT-09-001, CERN, 2009.
- [67] CMS Collaboration, "The CMS Particle Flow Algorithm," arXiv:1401.8155, pp. 295–304, 2014.

- [68] CMS Internal, "CMS muon results." https://twiki.cern.ch/twiki/bin/view/ CMSPublic/PhysicsResultsMUO.
- [69] W. Adam, R. Frühwirth, A. Strandlie, and T. Todor, "Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC," Tech. Rep. CMS-NOTE-2005-001, CERN, Geneva, Jan 2005.
- [70] S. Baffioni, C. Charlot, F. Ferri, D. Futyan, P. Meridiani, et al., "Electron reconstruction in CMS," Eur. Phys. J., vol. C49, pp. 1099–1116, 2007.
- [71] CMS Collaboration, "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV," arXiv:1502.02701, 2015.
- [72] CMS Collaboration, "Electron performance with 19.6 fb⁻¹ of data collected at $\sqrt{s} = 8$ TeV with the CMS detector.," CMS Detector Performance Note CMS DP 2013/003, CERN, Mar 2013.
- [73] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k(t) jet clustering algorithm," *JHEP*, vol. 0804, p. 063, 2008.
- [74] CMS Collaboration, "CMS technical design report, volume II: Physics performance," J.Phys., vol. G34, pp. 995–1579, 2007.
- [75] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS," *JINST*, vol. 6, p. P11002, 2011.
- [76] CMS Collaboration, "Plans for Jet Energy Corrections at CMS," CMS Physics Analysis Summary CMS-PAS-JME-07-002, CERN, 2008. Geneva, Jul 2008.
- [77] CMS Internal, "Jet energy corrections: Official software tools for applying JEC corrections and uncertainties." https://twiki.cern.ch/twiki/bin/view/ CMSPublic/WorkBookJetEnergyCorrections.
- [78] CMS Collaboration, "Search for supersymmetry in all-hadronic events with missing energy," CMS Physics Analysis Summary CMS-PAS-SUS-11-004, CERN, Geneva, 2011.
- [79] ATLAS Collaboration, "Search for New Physics in the Dijet Mass Distribution using 1 fb⁻¹ of pp Collision Data at $\sqrt{s} = 7$ TeV collected by the ATLAS Detector," *Phys.Lett.*, vol. B708, pp. 37–54, 2012.
- [80] P. Nadolsky, J. Gao, M. Guzzi, J. Huston, H.-L. Lai, et al., "Progress in CTEQ-TEA PDF Analysis," pp. 417–420, 2012.
- [81] CMS Collaboration, "Charged particle multiplicities in pp interactions at $\sqrt{s} = 0.9$, 2.36, and 7 TeV," *JHEP*, vol. 1101, p. 079, 2011.

- [82] Z. Was, "TAUOLA for simulation of tau decay and production: perspectives for precision low energy and LHC applications," *Nucl. Phys. Proc. Suppl.*, vol. 218, pp. 249– 255, 2011.
- [83] CMS Internal, "Standard model cross sections for CMS at 8 TeV." https://twiki. cern.ch/twiki/bin/viewauth/CMS/StandardModelCrossSectionsat8TeV.
- [84] J. M. Campbell and R. Ellis, "MCFM for the Tevatron and the LHC," Nucl. Phys. Proc. Suppl., vol. 205-206, pp. 10–15, 2010.
- [85] N. Kidonakis, "Differential and total cross sections for top pair and single top production," pp. 831–834, 2012.
- [86] CMS Internal, "Reference muon id and isolation efficiencies." https://twiki. cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffs.
- [87] CMS Internal, "Baseline muon selections." https://twiki.cern.ch/twiki/bin/ view/CMSPublic/SWGuideMuonId.
- [88] CMS Collaboration, "Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV," Tech. Rep. CMS-PAS-PFT-10-002, CERN, 2010.
- [89] CMS Internal, "Cut based electron id." https://twiki.cern.ch/twiki/bin/ view/CMS/EgammaCutBasedIdentification.
- [90] CMS Collaboration, "Jet performance in CMS," Tech. Rep. CMS-CR-2013-325, CERN, Geneva, Oct 2013.
- [91] CMS Internal, "Jet identification." https://twiki.cern.ch/twiki/bin/view/ CMS/JetID.
- [92] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual," *Eur.Phys.J.*, vol. C72, p. 1896, 2012.
- [93] CMS Collaboration, "Pileup Jet Identification," CMS Physics Analysis Summary CMS-PAS-JME-13-005, CERN, 2013.
- [94] CMS Internal, "Pileup reweighting utilities." https://twiki.cern.ch/twiki/bin/ viewauth/CMS/PileupMCReweightingUtilities.
- [95] A. Bodek, A. van Dyne, J. Han, W. Sakumoto, and A. Strelnikov, "Extracting Muon Momentum Scale Corrections for Hadron Collider Experiments," *Eur.Phys.J.*, vol. C72, p. 2194, 2012.

- [96] CMS Collaboration, "Measurement of the properties of a Higgs boson in the fourlepton final state," *Phys.Rev.*, vol. D89, no. 9, p. 092007, 2014.
- [97] A. Hocker and V. Kartvelishvili, "SVD approach to data unfolding," Nucl.Instrum.Meth., vol. A372, pp. 469–481, 1996.
- [98] T. Adye, "Unfolding algorithms and tests using RooUnfold," arXiv:1105.1160, 2011.
- [99] G. D'Agostini, "A Multidimensional unfolding method based on Bayes' theorem," *Nucl.Instrum.Meth.*, vol. A362, pp. 487–498, 1995.
- [100] M. Czakon, P. Fiedler, and A. Mitov, "The total top quark pair production crosssection at hadron colliders through $O(\alpha_S^4)$," 2013.
- [101] CMS Collaboration, "Absolute Calibration of the Luminosity Measurement at CMS: Winter 2012 Update," CMS Physics Analysis Summary CMS-PAS-SMP-12-008, CERN, Geneva, 2012.
- [102] CMS Internal, "Estimation of covariance matrix to combine the two channels of Z + jet analysis with the BLUE method." https://indico.cern.ch/event/298918/ contribution/3/material/slides/0.pdf.
- [103] CMS Collaboration, "Measurement of the $t\bar{t}$ charge asymmetry with lepton + jets events at 8 TeV," CMS Physics Analysis Summary CMS-PAS-TOP-12-033, CERN, Geneva, 2013.
- [104] M. Guzzi, P. Nadolsky, E. Berger, H.-L. Lai, F. Olness, et al., "CT10 parton distributions and other developments in the global QCD analysis," 2011.
- [105] M. Ubiali, "Towards the NNPDF3.0 parton set for the second LHC run," PoS, vol. DIS2014, p. 041, 2014.
- [106] T. Seva, L. Favart, A. Léonard, and P. Gras, "Differential cross section of jets associated to Z boson in Proton-Proton Collisions at $\sqrt{s}=8$ TeV," CMS Analysis Note CMS-AN-13-049, CERN, 2013.
- [107] A. Valassi, "Combining correlated measurements of several different physical quantities," Nucl.Instrum.Meth., vol. A500, pp. 391–405, 2003.
- [108] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, "ALPGEN, a generator for hard multiparton processes in hadronic collisions," *JHEP*, vol. 0307, p. 001, 2003.
- [109] G. Corcella, I. Knowles, G. Marchesini, S. Moretti, K. Odagiri, et al., "HERWIG 6.5 release note," arXiv:hep-ph/0210213, 2002.

- [110] CMS Collaboration, "Z + jets differential cross section at 8 TeV," CMS Physics Analysis Summary CMS-PAS-SMP-13-007, CERN, 2013.
- [111] CMS Collaboration, "Measurement of the double differential cross section of Z bosons produced in association with jets in pp collisions at $\sqrt{s} = 8$ TeV," CMS Physics Analysis Summary CMS-PAS-SMP-14-009, CERN, 2014.
- [112] D. Saltzberg, E. Takasugi, and M. Weber, "Azimuthal correlations in Z + jets events in pp collisions at $\sqrt{s} = 8$ TeV," CMS Analysis Note CMS-AN-14-233, CERN, 2014.
- [113] F. Zhang, L. Favart, and A. Léonard, "Angular correlations of Z + jets events in pp collisions at central mass of 8TeV," CMS Analysis Note CMS-AN-15-033, CERN, 2014.
- [114] E. Verhagen, "Development of the new trigger and data acquisition system for the CMS forward muon spectrometer upgrade," Ph.D. Thesis, 2015.
- [115] GEM for CMS collaboration, "CMS Technical design report for the muon endcap upgrade: GE1/1 - The station 1 GEM project," 2014.
- [116] W. Blum and L. Rolandi, Particle Detection with Drift Chambers. Springer-Verlag, 1994.
- [117] P. Aspell, "VFAT2: A front-end system on chip providing fast trigger information, digitized data storage and formatting for the charge sensitive readout of multichannel silicon and gas particle detectors," p. 63, 2007.
- [118] C. Ghabrous Larrea, K. Harder, D. Newbold, D. Sankey, A. Rose, et al., "IPbus: a flexible Ethernet-based control system for xTCA hardware," JINST, vol. 10, no. 02, p. C02019, 2015.
- [119] "Spartan-6 FPGA SP601 Evaluation Kit." http://www.xilinx.com/products/ boards-and-kits/ek-s6-sp601-g.html.
- [120] "MicroBlaze Soft Processor Core." http://www.xilinx.com/tools/microblaze. htm.
- [121] "GLIB-project public page." https://espace.cern.ch/project-GBLIB/public/ default.aspx.

BIBLIOGRAPHY
Appendices

Appendix A

Drell-Yan Cross Section

In this appendix the detailed computation of the Drell-Yan process cross section is presented. The calculus includes the virtual photon and Z boson production as well as their interference. The result is given in terms of the differential and integrated cross sections.

A.1 Virtual Photon Production

The first elements needed for this calculation are the matrix elements, \mathcal{M}_{γ} and \mathcal{M}_{Z} for the process $q\bar{q} \to \gamma^*/Z \to l^+l^-$. We start with the former using the EWK Feynman rules depicted in figure 2.6 along with the corresponding Feynman diagram. The result is quickly obtained:

$$i\mathcal{M}_{\gamma} = \bar{v}_{c_{\bar{q}}}^{s_{\bar{q}}}(p_{\bar{q}})(-i)Q_{q}e\delta_{c_{\bar{q}}c_{q}}\gamma^{\mu}u_{c_{q}}^{s_{q}}(p_{q}) \times \left(\frac{-ig_{\mu\nu}}{\hat{s}}\right) \times \bar{u}^{s_{l}}(p_{l})(-i)Q_{l}e\gamma^{\nu}v^{s_{\bar{l}}}(p_{\bar{l}})$$
$$= i\frac{Q_{q}Q_{l}e^{2}}{\hat{s}} \times \bar{v}_{c_{\bar{q}}}^{s_{\bar{q}}}(p_{\bar{q}})\delta_{c_{\bar{q}}c_{q}}\gamma^{\mu}u_{c_{q}}^{s_{q}}(p_{q}) \times \bar{u}^{s_{l}}(p_{l})\gamma_{\mu}v^{s_{\bar{l}}}(p_{\bar{l}}).$$
(A.1)

where the indices c_q , $c_{\bar{q}}$, s_q , $s_{\bar{q}}$, s_l and $s_{\bar{l}}$ indicate the quark and antiquark colours and spins and the lepton and antilepton spins, respectively. The symbol \hat{s} represents the quark-antiquark centre-of-mass energy and is linked to the proton-proton centre-of-mass energy s by: $\hat{s} = x_1 x_2 s$ where x_i is the proton momentum fraction carried by the parton i. Q_q and Q_l are the quark and lepton electric charges in units of the positron electric charge e. It is now a matter of calculation to obtain the differential and integrated cross sections. The following relation, for two spinors ψ and χ ,

$$(\bar{\psi}\gamma^{\mu}\chi)^{*} = (\psi^{\dagger}\gamma^{0}\gamma^{\mu}\chi)^{*} = \chi^{\dagger}(\gamma^{\mu})^{\dagger}(\gamma^{0})^{\dagger}\psi = \chi^{\dagger}(\gamma^{\mu})^{\dagger}\gamma^{0}\psi = \chi^{\dagger}\gamma^{0}\gamma^{\mu}\gamma^{0}\gamma^{0}\psi = \bar{\chi}\gamma^{\mu}\psi,$$

allows us to write the complex conjugate of A.1

$$i\mathcal{M}^*_{\gamma} = i\frac{Q_q Q_l e^2}{\hat{s}} \times \bar{u}^{s_q}_{c_q}(p_q)\delta_{c_q c_{\bar{q}}}\gamma^{\nu}v^{s_{\bar{q}}}_{c_{\bar{q}}}(p_{\bar{q}}) \times \bar{v}^{s_{\bar{l}}}(p_{\bar{l}})\gamma_{\nu}u^{s_l}(p_l).$$
(A.2)

Equipped with these two expressions, we can compute the square of the matrix element:

$$\begin{aligned} |\mathcal{M}_{\gamma}|^{2} &= \mathcal{M}_{\gamma}\mathcal{M}_{\gamma}^{*} \\ &= \left(\frac{Q_{q}Q_{l}e^{2}}{\hat{s}}\right)^{2} \times \bar{v}_{c\bar{q}}^{s_{\bar{q}}}(p_{\bar{q}})\delta_{c_{\bar{q}}c_{q}}\gamma^{\mu}u_{c_{q}}^{s_{q}}(p_{q}) \ \bar{u}_{c_{q}}^{s_{q}}(p_{q})\delta_{c_{q}c_{\bar{q}}}\gamma^{\nu}v_{c_{\bar{q}}}^{s_{\bar{q}}}(p_{\bar{q}}) \\ &\times \bar{u}^{s_{l}}(p_{l})\gamma_{\mu}v^{s_{\bar{l}}}(p_{\bar{l}}) \ \bar{v}^{s_{\bar{l}}}(p_{\bar{l}})\gamma_{\nu}u^{s_{l}}(p_{l}). \end{aligned}$$
(A.3)

The next step is to average over the initial spin and colour configurations and sum over the final particle spins:

$$\begin{split} \overline{|\mathcal{M}_{\gamma}|^{2}} &= \sum_{s_{l}} \sum_{s_{\bar{l}}} \frac{1}{2} \sum_{s_{\bar{q}}} \frac{1}{2} \sum_{s_{\bar{q}}} \frac{1}{n_{c}} \sum_{c_{\bar{q}}} \frac{1}{n_{c}} \sum_{c_{\bar{q}}} |\mathcal{M}_{\gamma}|^{2} \\ &= \left(\frac{Q_{q}Q_{l}e^{2}}{\hat{s}} \right)^{2} \frac{1}{4} \frac{1}{n_{c}} \sum_{s_{l},s_{\bar{l}},s_{\bar{q}},s_{\bar{q}}} \left\{ \bar{v}^{s_{\bar{q}}}(p_{\bar{q}})\gamma^{\mu}u^{s_{\bar{q}}}(p_{q}) \ \bar{u}^{s_{\bar{q}}}(p_{\bar{q}})\gamma^{\nu}v^{s_{\bar{q}}}(p_{\bar{q}}) \\ &\times \ \bar{u}^{s_{l}}(p_{l})\gamma_{\mu}v^{s_{\bar{l}}}(p_{\bar{l}}) \ \bar{v}^{s_{\bar{l}}}(p_{\bar{l}})\gamma_{\nu}u^{s_{l}}(p_{l}) \right\} \\ &= \left(\frac{Q_{q}Q_{l}e^{2}}{\hat{s}} \right)^{2} \frac{1}{4} \frac{1}{n_{c}} \sum_{s_{l},s_{\bar{l}},s_{\bar{q}},s_{\bar{q}}} \left\{ \bar{v}^{s_{\bar{q}}}(p_{\bar{q}})_{\alpha} (\gamma^{\mu})_{\alpha\beta} u^{s_{q}}(p_{q})_{\beta} \ \bar{u}^{s_{\bar{q}}}(p_{\bar{q}})_{\delta} (\gamma^{\nu})_{\delta\epsilon} v^{s_{\bar{q}}}(p_{\bar{q}})_{\epsilon} \\ &\times \ \bar{u}^{s_{l}}(p_{l})_{\rho} (\gamma_{\mu})_{\rho\sigma} v^{s_{\bar{l}}}(p_{\bar{l}})_{\sigma} \ \bar{v}^{s_{\bar{l}}}(p_{\bar{l}})_{\tau} (\gamma_{\nu})_{\tau\varphi} u^{s_{l}}(p_{l})_{\varphi} \right\} \\ &= \left(\frac{Q_{q}Q_{l}e^{2}}{\hat{s}} \right)^{2} \frac{1}{4} \frac{1}{n_{c}} \left\{ \left(\not{p}_{q} + m_{q} \right)_{\beta\delta} \left(\not{p}_{\bar{q}} + m_{q} \right)_{\epsilon\alpha} (\gamma^{\mu})_{\alpha\beta} (\gamma^{\nu})_{\delta\epsilon} \\ &\times \left(\not{p}_{l} + m_{l} \right)_{\varphi\rho} \left(\not{p}_{\bar{l}} + m_{l} \right)_{\sigma\tau} (\gamma_{\mu})_{\rho\sigma} (\gamma_{\nu})_{\tau\varphi} \right\} \\ &= \left(\frac{Q_{q}Q_{l}e^{2}}{\hat{s}} \right)^{2} \frac{1}{4} \frac{1}{n_{c}} \left\{ (\gamma^{\mu})_{\alpha\beta} \left(\not{p}_{q} \right)_{\beta\delta} (\gamma^{\nu})_{\delta\epsilon} \left(\not{p}_{\bar{q}} \right)_{\epsilon\alpha} (\gamma_{\nu})_{\tau\varphi} \left(\not{p}_{l} \right)_{\varphi\rho} (\gamma_{\mu})_{\rho\sigma} \left(\not{p}_{\bar{l}} \right)_{\sigma\tau} \right\}. \end{split}$$
(A.4)

To go from the first line to the second line, the sum over $c_{\bar{q}}$ has been trivially done thanks to the $\delta_{c_{\bar{q}}c_q}$ which translates the fact that colour must be conserved. Since the photon does not have a colour charge the interaction can take place only between a quark of some colour (red, green or blue) and the antiquark of the opposite colour (antired, antigreen or antiblue, respectively). The next sum over c_q just brings a factor n_c to the numerator, cancelling one of the two factors n_c in the denominator. At this point we obtain the colour factor $\frac{1}{n_c}$ present in the final result. The passage to the third line is simply adding the implicit summation indices of the spinor and gamma matrices components. From the third to the fourth line, the spin completeness relations $\sum_s u^s(p)\bar{u}^s(p) = \not p + m$ and $\sum_s v^s(p)\bar{v}^s(p) = \not p - m$ have been used. To go from line four to line five, we neglect the fermion masses and rearrange the terms to make two traces appear. The traces can be solved using the anticommutation relations defining the gamma matrices:

Similarly for the second trace related to the leptons (one can directly get the result by replacing the indices $q \to l$ and $\bar{q} \to \bar{l}$ and swapping and lowering the indices μ and ν in the previous expression):

$$(\gamma_{\nu})_{\tau\varphi} \left(\not\!\!p_l \right)_{\varphi\rho} (\gamma_{\mu})_{\rho\sigma} \left(\not\!\!p_{\bar{l}} \right)_{\sigma\tau} = 4 \left(p_{l\nu} p_{\bar{l}\,\mu} - p_l \cdot p_{\bar{l}} g_{\nu\mu} + p_{l\,\mu} p_{\bar{l}\,\nu} \right).$$
(A.6)

Grouping the two results (A.5) and (A.6) into (A.4) and using the relation $g^{\mu\nu}g_{\nu\mu} = 4$ leads to expression:

$$\overline{|\mathcal{M}_{\gamma}|^{2}} = \left(\frac{Q_{q}Q_{l}e^{2}}{\hat{s}}\right)^{2} \frac{1}{4} \frac{1}{n_{c}} 32\left\{(p_{l} \cdot p_{q})(p_{\bar{l}} \cdot p_{\bar{q}}) + (p_{l} \cdot p_{\bar{q}})(p_{\bar{l}} \cdot p_{q})\right\}$$
$$= \left(\frac{Q_{q}Q_{l}e^{2}}{\hat{s}}\right)^{2} \frac{1}{4} \frac{1}{n_{c}} 32\left\{(\frac{\hat{s}}{4})^{2}(1 + \cos\theta)^{2} + (\frac{\hat{s}}{4})^{2}(1 - \cos\theta)^{2}\right\}, \quad (A.7)$$

where θ is defined as the angle of emission of the lepton in the photon rest frame. Inserting this result into the general formula for the differential cross section of a $2 \rightarrow 2$ process

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2 \hat{s}} \overline{|\mathcal{M}|^2},\tag{A.8}$$

yields to the differential cross section

$$\frac{d\sigma_{\gamma}}{d\Omega} = \frac{\alpha^2 Q_q^2 Q_l^2}{8\hat{s}} \frac{1}{n_c} \left\{ (1 + \cos\theta)^2 + (1 - \cos\theta)^2 \right\},\tag{A.9}$$

which when integrated over the full solid angle becomes

$$\sigma_{\gamma} = \frac{4\pi\alpha^2}{3\hat{s}} \frac{1}{n_c} Q_q^2 Q_l^2. \tag{A.10}$$

A.2 Z Boson Production

Similarly for the production of a Z boson, using the appropriate EWK Feynman rules summarised in figure 2.7 together with the corresponding Feynman diagram, we get:

$$i\mathcal{M}_{Z} = \bar{v}_{c\bar{q}}^{s\bar{q}}(p_{\bar{q}}) \frac{(-i)g_{Z}}{2} \delta_{c_{\bar{q}}c_{q}} \gamma^{\mu} (V_{q} - A_{q}\gamma^{5}) u_{c_{q}}^{s_{q}}(p_{q}) \times \left(\frac{-ig_{\mu\nu} + i\frac{q_{\mu}q_{\nu}}{M_{Z}^{2}}}{\hat{s} - M_{Z}^{2} + iM_{Z}\Gamma_{Z}}\right) \\ \times \bar{u}^{s_{l}}(p_{l}) \frac{(-i)g_{Z}}{2} \gamma^{\nu} (V_{l} - A_{l}\gamma^{5}) v^{s_{\bar{l}}}(p_{\bar{l}}) \\ = \left(\frac{g_{Z}}{2}\right)^{2} \frac{i}{\hat{s} - M_{Z}^{2} + iM_{Z}\Gamma_{Z}} \bar{v}_{c_{\bar{q}}}^{s_{\bar{q}}}(p_{\bar{q}}) \delta_{c_{\bar{q}}c_{q}} \gamma^{\mu} (V_{q} - A_{q}\gamma^{5}) u_{c_{q}}^{s_{q}}(p_{q}) \\ \times \bar{u}^{s_{l}}(p_{l}) \gamma_{\mu} (V_{l} - A_{l}\gamma^{5}) v^{s_{\bar{l}}}(p_{\bar{l}})$$
(A.11)

where $g_Z = g_W / \cos \theta_W$, $V_f = T_f^3 - 2Q_f \sin^2 \theta_W$, $A_f = T_f^3$ and where we have neglected the fermion masses arising from the $q_\mu q_\nu$ terms. Getting the complex conjugate expression is similar to the photon case except for the Z propagator and the presence of γ^5 matrices:

$$i\mathcal{M}_{Z}^{*} = \left(\frac{g_{Z}}{2}\right)^{2} \frac{i}{\hat{s} - M_{Z}^{2} - iM_{Z}\Gamma_{Z}} \quad \bar{u}_{c_{q}}^{s_{q}}(p_{q})\delta_{c_{q}c_{\bar{q}}}\gamma^{\nu}(V_{q} - A_{q}\gamma^{5})v_{c_{\bar{q}}}^{s_{\bar{q}}}(p_{\bar{q}}) \\ \times \bar{v}^{s_{\bar{l}}}(p_{\bar{l}})\gamma_{\nu}(V_{l} - A_{l}\gamma^{5})u^{s_{l}}(p_{l}).$$
(A.12)

Equipped with these two expressions, we can compute the square of the matrix element:

$$\begin{aligned} |\mathcal{M}_{Z}|^{2} &= \mathcal{M}_{Z}\mathcal{M}_{Z}^{*} \\ &= \left(\frac{g_{Z}}{2}\right)^{4} \frac{1}{(\hat{s} - M_{Z}^{2})^{2} + M_{Z}^{2}\Gamma_{Z}^{2}} \\ &\times \bar{v}_{c_{\bar{q}}}^{s_{\bar{q}}}(p_{\bar{q}})\delta_{c_{\bar{q}}c_{q}}\gamma^{\mu}(V_{q} - A_{q}\gamma^{5})u_{c_{q}}^{s_{q}}(p_{q}) \ \bar{u}_{c_{q}}^{s_{q}}(p_{q})\delta_{c_{q}c_{\bar{q}}}\gamma^{\nu}(V_{q} - A_{q}\gamma^{5})v_{c_{\bar{q}}}^{s_{\bar{q}}}(p_{\bar{q}}) \\ &\times \bar{u}^{s_{l}}(p_{l})\gamma_{\mu}(V_{l} - A_{l}\gamma^{5})v^{s_{\bar{l}}}(p_{\bar{l}}) \ \bar{v}^{s_{\bar{l}}}(p_{\bar{l}})\gamma_{\nu}(V_{l} - A_{l}\gamma^{5})u^{s_{l}}(p_{l}). \end{aligned}$$
(A.13)

The next step is to average over the initial spin and colour configurations and sum over the final particle spins:

$$\overline{|\mathcal{M}_{Z}|^{2}} = \sum_{s_{l}} \sum_{s_{\bar{l}}} \frac{1}{2} \sum_{s_{\bar{q}}} \frac{1}{2} \sum_{s_{\bar{q}}} \frac{1}{n_{c}} \sum_{c_{\bar{q}}} \frac{1}{n_{c}} \sum_{c_{\bar{q}}} |\mathcal{M}_{Z}|^{2} \\
= \left(\frac{g_{Z}}{2}\right)^{4} \frac{1}{(\hat{s} - M_{Z}^{2})^{2} + M_{Z}^{2} \Gamma_{Z}^{2}} \frac{1}{4} \frac{1}{n_{c}} \sum_{s_{l}, s_{\bar{l}}, s_{\bar{q}}, s_{\bar{q}}} \left\{ \frac{\bar{v}^{s_{\bar{q}}}(p_{\bar{q}}) \gamma^{\mu} (V_{q} - A_{q} \gamma^{5}) u^{s_{q}}(p_{q})}{\bar{u}^{s_{q}}(p_{q}) \gamma^{\nu} (V_{q} - A_{q} \gamma^{5}) v^{s_{\bar{q}}}(p_{\bar{q}})} \\
\times \bar{u}^{s_{l}}(p_{l}) \gamma_{\mu} (V_{l} - A_{l} \gamma^{5}) v^{s_{\bar{l}}}(p_{\bar{l}}) \quad \bar{v}^{s_{\bar{l}}}(p_{\bar{l}}) \gamma_{\nu} (V_{l} - A_{l} \gamma^{5}) u^{s_{l}}(p_{l}) \right\} \\
= \left(\frac{g_{Z}}{2}\right)^{4} \frac{1}{(\hat{s} - M_{Z}^{2})^{2} + M_{Z}^{2} \Gamma_{Z}^{2}} \frac{1}{4} \frac{1}{n_{c}}} \left\{ (\gamma^{\mu} (V_{q} - A_{q} \gamma^{5}))_{\alpha\beta} (\not{p}_{q})_{\beta\delta} (\gamma^{\nu} (V_{q} - A_{q} \gamma^{5}))_{\delta\epsilon} (\not{p}_{\bar{q}})_{\epsilon\alpha} \\
\times (\gamma_{\nu} (V_{l} - A_{l} \gamma^{5}))_{\tau\varphi} (\not{p}_{l})_{\varphi\rho} (\gamma_{\mu} (V_{l} - A_{l} \gamma^{5}))_{\rho\sigma} (\not{p}_{\bar{l}})_{\sigma\tau} \right\}, \quad (A.14)$$

A.2. Z BOSON PRODUCTION

where equivalent operations to the photon case (see equation (A.4)) have been operated.

The traces are a bit longer than for the photon case due to the presence of the γ^5 matrices. The relations $\{\gamma^5, \gamma^{\mu}\} = 0$ and $(\gamma^5)^2 = I$, together with the anticommutation relations defining the gamma matrices are used to solve the traces:

$$\begin{aligned} \left(\gamma^{\mu}(V_{q}-A_{q}\gamma^{5})\right)_{\alpha\beta}(\not{p}_{q})_{\beta\delta}\left(\gamma^{\nu}(V_{q}-A_{q}\gamma^{5})\right)_{\delta\epsilon}(\not{p}_{\bar{q}})_{\epsilon\alpha} \\ &= \operatorname{Tr}\left[\gamma^{\mu}(V_{q}-A_{q}\gamma^{5})p_{q\lambda}\gamma^{\lambda}\gamma^{\nu}(V_{q}-A_{q}\gamma^{5})p_{\bar{q}\omega}\gamma^{\omega}\right] \\ &= p_{q\lambda}p_{\bar{q}\omega}\left\{\operatorname{Tr}\left[\gamma^{\mu}V_{q}\gamma^{\lambda}\gamma^{\nu}V_{q}\gamma^{\omega}\right] + \operatorname{Tr}\left[\gamma^{\mu}V_{q}\gamma^{\lambda}\gamma^{\nu}(-A_{q})\gamma^{5}\gamma^{\omega}\right] \\ &+ \operatorname{Tr}\left[\gamma^{\mu}(-A_{q})\gamma^{5}\gamma^{\lambda}\gamma^{\nu}V_{q}\gamma^{\omega}\right] + \operatorname{Tr}\left[\gamma^{\mu}(-A_{q})\gamma^{5}\gamma^{\lambda}\gamma^{\nu}(-A_{q})\gamma^{5}\gamma^{\omega}\right]\right\} \\ &= p_{q\lambda}p_{\bar{q}\omega}\left\{\left(V_{q}^{2}+A_{q}^{2}\right)\operatorname{Tr}\left[\gamma^{\mu}\gamma^{\lambda}\gamma^{\nu}\gamma^{\omega}\right] + 2A_{q}V_{q}\operatorname{Tr}\left[\gamma^{\mu}\gamma^{\lambda}\gamma^{\nu}\gamma^{\omega}\gamma^{5}\right]\right\} \\ &= p_{q\lambda}p_{\bar{q}\omega}\left\{4\left(V_{q}^{2}+A_{q}^{2}\right)\left(g^{\mu\lambda}g^{\nu\omega}-g^{\mu\nu}g^{\lambda\omega}+g^{\mu\omega}g^{\lambda\nu}\right)-i8A_{q}V_{q}\epsilon^{\mu\lambda\nu\omega}\right\} \\ &= 4\left(V_{q}^{2}+A_{q}^{2}\right)\left(p_{q}^{\mu}p_{\bar{q}}^{\nu}-p_{q}\cdot p_{\bar{q}}g^{\mu\nu}+p_{q}^{\nu}p_{\bar{q}}^{\mu}\right)+i8A_{q}V_{q}p_{q\lambda}p_{\bar{q}\omega}\epsilon^{\mu\nu\lambda\omega}, \end{aligned}$$

$$(A.15)$$

where we have used Tr $[\gamma^{\mu}\gamma^{\lambda}\gamma^{\nu}\gamma^{\omega}\gamma^{5}] = -i4\epsilon^{\mu\lambda\nu\omega}$. Similarly for the trace related to the leptons:

$$\begin{aligned} \left(\gamma_{\nu}(V_{l}-A_{l}\gamma^{5})\right)_{\tau\varphi}\left(\not{p}_{l}\right)_{\varphi\rho}\left(\gamma_{\mu}(V_{l}-A_{l}\gamma^{5})\right)_{\rho\sigma}\left(\not{p}_{\bar{l}}\right)_{\sigma\tau} \\ &= \operatorname{Tr}\left[\gamma_{\nu}(V_{l}-A_{l}\gamma^{5})p_{q}^{\kappa}\gamma_{\kappa}\gamma_{\mu}(V_{l}-A_{l}\gamma^{5})p_{\bar{l}}^{t}\gamma_{\iota}\right] \\ &= p_{l}^{\kappa}p_{\bar{l}}^{\iota}\left\{\operatorname{Tr}\left[\gamma_{\nu}V_{l}\gamma_{\kappa}\gamma_{\mu}V_{l}\gamma_{\iota}\right] + \operatorname{Tr}\left[\gamma_{\nu}V_{l}\gamma_{\kappa}\gamma_{\mu}(-A_{l})\gamma^{5}\gamma_{\iota}\right] \\ &+ \operatorname{Tr}\left[\gamma_{\nu}(-A_{l})\gamma^{5}\gamma_{\kappa}\gamma_{\mu}V_{l}\gamma_{\iota}\right] + \operatorname{Tr}\left[\gamma_{\nu}(-A_{l})\gamma^{5}\gamma_{\kappa}\gamma_{\mu}(-A_{l})\gamma^{5}\gamma_{\iota}\right]\right\} \\ &= p_{l}^{\kappa}p_{\bar{l}}^{\iota}\left\{\left(V_{l}^{2}+A_{l}^{2}\right)\operatorname{Tr}\left[\gamma_{\nu}\gamma_{\kappa}\gamma_{\mu}\gamma_{\iota}\right] + 2A_{l}V_{l}\operatorname{Tr}\left[\gamma_{\nu}\gamma_{\kappa}\gamma_{\mu}\gamma_{\iota}\gamma^{5}\right]\right\} \\ &= p_{l}^{\kappa}p_{\bar{l}}^{\iota}\left\{4(V_{l}^{2}+A_{l}^{2})(g_{\nu\kappa}g_{\mu\iota}-g_{\nu\mu}g_{\kappa\iota}+g_{\nu\iota}g_{\kappa\mu})+i8A_{l}V_{l}\epsilon_{\nu\kappa\mu\iota}\right\} \\ &= 4(V_{l}^{2}+A_{l}^{2})(p_{l\nu}p_{\bar{l}\mu}-p_{l}\cdot p_{\bar{l}}g_{\nu\mu}+p_{l\mu}p_{\bar{l}\nu})+i8A_{l}V_{l}p_{l}^{\kappa}p_{l}^{\iota}\epsilon_{\mu\nu\kappa\iota}, \end{aligned}$$

$$(A.16)$$

where we have used Tr $[\gamma_{\nu}\gamma_{\kappa}\gamma_{\mu}\gamma_{\iota}\gamma^{5}] = i4\epsilon_{\nu\kappa\mu\iota}$. The product of the two results is simpler than what it may look. The $\epsilon^{\nu\lambda\mu\omega}$ being antisymmetric in $\mu \leftrightarrow \nu$ will give 0 when multiplied by the first term of these two results which is symmetric in $\mu \leftrightarrow \nu$. Furthermore, the first term of each of the two traces is totally similar the the one of the photon case. The product of the quark and lepton related traces leads to:

$$\begin{split} \operatorname{Tr} \left[\gamma^{\mu} (V_{q} - A_{q} \gamma^{5}) p_{q \lambda} \gamma^{\lambda} \gamma^{\nu} (V_{q} - A_{q} \gamma^{5}) p_{\bar{q} \omega} \gamma^{\omega} \right] \times \operatorname{Tr} \left[\gamma_{\nu} (V_{l} - A_{l} \gamma^{5}) p_{q}^{\kappa} \gamma_{\kappa} \gamma_{\mu} (V_{l} - A_{l} \gamma^{5}) p_{\bar{l}}^{\iota} \gamma_{l} \right] \\ &= 32 (V_{q}^{2} + A_{q}^{2}) (V_{l}^{2} + A_{l}^{2}) \left[(p_{q} \cdot p_{l}) (p_{\bar{q}} \cdot p_{\bar{l}}) + (p_{q} \cdot p_{\bar{l}}) (p_{\bar{q}} \cdot p_{l}) \right] \\ &- 64 A_{q} V_{q} A_{l} V_{l} p_{q \lambda} p_{\bar{q} \omega} p_{l}^{\kappa} p_{\bar{l}}^{\iota} \epsilon^{\mu \nu \lambda \omega} \epsilon_{\mu \nu \kappa \iota} \\ &= 32 (V_{q}^{2} + A_{q}^{2}) (V_{l}^{2} + A_{l}^{2}) \left[(p_{q} \cdot p_{l}) (p_{\bar{q}} \cdot p_{\bar{l}}) + (p_{q} \cdot p_{\bar{l}}) (p_{\bar{q}} \cdot p_{l}) \right] \\ &+ 128 A_{q} V_{q} A_{l} V_{l} p_{q \lambda} p_{\bar{q} \omega} p_{l}^{\kappa} p_{\bar{l}}^{\iota} \left(\delta_{\kappa}^{\lambda} \delta_{\iota}^{\omega} - \delta_{\iota}^{\lambda} \delta_{\kappa}^{\omega} \right) \\ &= 32 (V_{q}^{2} + A_{q}^{2}) (V_{l}^{2} + A_{l}^{2}) \left[(p_{q} \cdot p_{l}) (p_{\bar{q}} \cdot p_{\bar{l}}) + (p_{q} \cdot p_{\bar{l}}) (p_{\bar{q}} \cdot p_{l}) \right] \\ &+ 128 A_{q} V_{q} A_{l} V_{l} p_{q \lambda} p_{\bar{q} \omega} p_{l}^{\kappa} p_{\bar{l}} \left(\delta_{\kappa}^{\lambda} \delta_{\iota}^{\omega} - \delta_{\iota}^{\lambda} \delta_{\kappa}^{\omega} \right) \\ &= 32 (V_{q}^{2} + A_{q}^{2}) (V_{l}^{2} + A_{l}^{2}) \left[(p_{q} \cdot p_{l}) (p_{\bar{q}} \cdot p_{\bar{l}}) + (p_{q} \cdot p_{\bar{l}}) (p_{\bar{q}} \cdot p_{l}) \right] . \end{split}$$

Combining the different pieces together we get:

$$\overline{|\mathcal{M}_{Z}|^{2}} = \left(\frac{g_{Z}}{2}\right)^{4} \frac{1}{(\hat{s} - M_{Z}^{2})^{2} + M_{Z}^{2}\Gamma_{Z}^{2}} \frac{1}{4} \frac{1}{n_{c}} 32 \left\{ \left(V_{q}^{2} + A_{q}^{2}\right)(V_{l}^{2} + A_{l}^{2}\right) \left[(p_{q} \cdot p_{l})(p_{\bar{q}} \cdot p_{\bar{l}}) + (p_{q} \cdot p_{\bar{l}})(p_{\bar{q}} \cdot p_{l})\right] \\ + 4A_{q}V_{q}A_{l}V_{l} \left[(p_{q} \cdot p_{l})(p_{\bar{q}} \cdot p_{\bar{l}}) - (p_{q} \cdot p_{\bar{l}})(p_{\bar{q}} \cdot p_{l})\right] \right\} \\ = \left(\frac{g_{Z}}{2}\right)^{4} \frac{1}{(\hat{s} - M_{Z}^{2})^{2} + M_{Z}^{2}\Gamma_{Z}^{2}} \frac{1}{4} \frac{1}{n_{c}} 32 \left\{ \left(V_{q}^{2} + A_{q}^{2}\right)(V_{l}^{2} + A_{l}^{2}) \left[(\frac{\hat{s}}{4})^{2}(1 + \cos\theta)^{2} + (\frac{\hat{s}}{4})^{2}(1 - \cos\theta)^{2}\right] \\ + 4A_{q}V_{q}A_{l}V_{l} \left[(\frac{\hat{s}}{4})^{2}(1 + \cos\theta)^{2} - (\frac{\hat{s}}{4})^{2}(1 - \cos\theta)^{2}\right] \right\}.$$
(A.17)

With this result we can now write down the differential cross section for the production of a ${\cal Z}$ boson:

$$\frac{d\sigma_Z}{d\Omega} = \frac{1}{64\pi^2 \hat{s}} \overline{|\mathcal{M}_Z|^2}
= \frac{1}{n_c} \frac{1}{8\hat{s}} \frac{1}{\sin^4 2\theta_W} \alpha^2 \left(\frac{\hat{s}^2}{(\hat{s} - M_Z^2)^2 + M_Z^2 \Gamma_Z^2} \right) \left\{ \\ \left[(V_q^2 + A_q^2) (V_l^2 + A_l^2) + 4A_q V_q A_l V_l \right] (1 + \cos \theta)^2 \\ + \left[(V_q^2 + A_q^2) (V_l^2 + A_l^2) - 4A_q V_q A_l V_l \right] (1 - \cos \theta)^2 \right\}.$$
(A.18)

Once integrated over the solid angle, this gives:

$$\sigma_Z = \frac{4\pi\alpha^2}{3\hat{s}} \frac{1}{n_c} \frac{1}{\sin^4 2\theta_W} (V_q^2 + A_q^2) (V_l^2 + A_l^2) \frac{\hat{s}^2}{(\hat{s} - M_Z^2)^2 + M_Z^2 \Gamma_Z^2}.$$
 (A.19)

A.3 Interference

Finally, the interference between the γ^* and the Z production modes can be computed using the previous results:

$$\begin{split} \overline{\mathcal{M}_{\gamma}\mathcal{M}_{Z}^{*}} &= \sum_{s_{l}} \sum_{s_{\bar{l}}} \frac{1}{2} \sum_{s_{\bar{q}}} \frac{1}{2} \sum_{s_{\bar{q}}} \frac{1}{n_{c}} \sum_{c_{\bar{q}}} \frac{1}{n_{c}} \sum_{c_{\bar{q}}} \mathcal{M}_{\gamma}\mathcal{M}_{Z}^{*} \\ &= \frac{Q_{q}Q_{l}e^{2}}{\hat{s}} \left(\frac{g_{Z}}{2}\right)^{2} \left(\frac{1}{\hat{s} - M_{Z}^{2} - iM_{Z}\Gamma_{Z}}\right) \frac{1}{4} \frac{1}{n_{c}} \left\{ \\ \overline{v}^{s_{\bar{q}}}(p_{\bar{q}})\gamma^{\mu}u^{s_{q}}(p_{q}) \quad \overline{u}^{s_{q}}(p_{q})\gamma^{\nu}(V_{q} - A_{q}\gamma^{5})v^{s_{\bar{q}}}(p_{\bar{q}}) \\ &\times \quad \overline{u}^{s_{l}}(p_{l})\gamma_{\mu}v^{s_{\bar{l}}}(p_{\bar{l}}) \quad \overline{v}^{s_{\bar{l}}}(p_{\bar{l}})\gamma_{\nu}(V_{l} - A_{l}\gamma^{5})u^{s_{l}}(p_{l}) \right\} \\ &= \frac{Q_{q}Q_{l}e^{2}}{\hat{s}} \left(\frac{g_{Z}}{2}\right)^{2} \left(\frac{1}{\hat{s} - M_{Z}^{2} - iM_{Z}\Gamma_{Z}}\right) \frac{1}{4} \frac{1}{n_{c}} \left\{ \\ (\gamma^{\mu})_{\alpha\beta}(p_{q})_{\beta\delta}(\gamma^{\nu}(V_{q} - A_{q}\gamma^{5}))_{\delta\epsilon}(p_{\bar{q}})_{\epsilon\alpha} \\ &\times \quad (\gamma_{\nu}(V_{l} - A_{l}\gamma^{5}))_{\tau\varphi}(p_{l})_{\varphi\rho}(\gamma_{\mu})_{\rho\sigma}(p_{\bar{l}})_{\sigma\tau} \right\}, \end{split}$$

where again the same operations as in equation (A.4) have been performed.

Similarly to the photon and boson cases, we can solve the traces:

$$\begin{split} (\gamma^{\mu})_{\alpha\beta}(p_{q})_{\beta\delta}(\gamma^{\nu}(V_{q}-A_{q}\gamma^{5}))_{\delta\epsilon}(p_{\bar{q}})_{\epsilon\alpha} &= \mathrm{Tr}\left[\gamma^{\mu}p_{q\,\lambda}\gamma^{\lambda}\gamma^{\nu}(V_{q}-A_{q}\gamma^{5})p_{\bar{q}\,\omega}\gamma^{\omega}\right] \\ &= p_{q\,\lambda}p_{\bar{q}\,\omega}\Big\{V_{q}\mathrm{Tr}\left[\gamma^{\mu}\gamma^{\lambda}\gamma^{\nu}\gamma^{\omega}\right] + A_{q}\mathrm{Tr}\left[\gamma^{\mu}\gamma^{\lambda}\gamma^{\nu}\gamma^{\omega}\gamma^{5}\right]\Big\} \\ &= p_{q\,\lambda}p_{\bar{q}\,\omega}\Big\{4V_{q}(g^{\mu\lambda}g^{\nu\omega} - g^{\mu\nu}g^{\lambda\omega} + g^{\mu\omega}g^{\lambda\nu}) - i4A_{q}\epsilon^{\mu\lambda\nu\omega}\Big\} \\ &= 4V_{q}(p_{q}^{\mu}p_{\bar{q}}^{\nu} - p_{q}\cdot p_{\bar{q}}g^{\mu\nu} + p_{q}^{\nu}p_{\bar{q}}^{\mu}) + i4A_{q}p_{q\,\lambda}p_{\bar{q}\,\omega}\epsilon^{\mu\nu\lambda\omega}. \end{split}$$

And for the trace associated to the leptons:

$$(\gamma_{\nu}(V_{l} - A_{l}\gamma^{5}))_{\tau\varphi}(\not p_{l})_{\varphi\rho}(\gamma_{\mu})_{\rho\sigma}(\not p_{\bar{l}})_{\sigma\tau} = \operatorname{Tr}\left[\gamma_{\nu}(V_{l} - A_{l}\gamma^{5})p_{l}^{\kappa}\gamma_{\kappa}\gamma_{\mu}p_{l}^{\iota}\gamma^{\iota}\right]$$
$$= p_{l}^{\kappa}p_{\bar{l}}^{\iota}\left\{V_{l}\operatorname{Tr}\left[\gamma_{\nu}\gamma_{\kappa}\gamma_{\mu}\gamma_{\iota}\right] + A_{l}\operatorname{Tr}\left[\gamma_{\nu}\gamma_{\kappa}\gamma_{\mu}\gamma_{\iota}\gamma^{5}\right]\right\}$$
$$= p_{l}^{\kappa}p_{\bar{l}}^{\iota}\left\{4V_{l}(g_{\nu\kappa}g_{\mu\iota} - g_{\nu\mu}g_{\kappa\iota} + g_{\nu\iota}g_{\kappa\mu}) + i4A_{l}\epsilon_{\nu\kappa\mu\iota}\right\}$$
$$= 4V_{l}(p_{l\nu}p_{\bar{l}\mu} - p_{l}\cdot p_{\bar{l}}g_{\nu\mu} + p_{l\mu}p_{\bar{l}\nu}) + i4A_{l}p_{l}^{\kappa}p_{\bar{l}}^{\iota}\epsilon_{\mu\nu\kappa\iota}.$$

Multiplying these two results we get:

$$\begin{aligned} \operatorname{Tr} \left[\gamma^{\mu} p_{q \lambda} \gamma^{\lambda} \gamma^{\nu} (V_{q} - A_{q} \gamma^{5}) p_{\bar{q} \omega} \gamma^{\omega} \right] &\times \operatorname{Tr} \left[\gamma_{\nu} (V_{l} - A_{l} \gamma^{5}) p_{l}^{\kappa} \gamma_{\kappa} \gamma_{\mu} p_{\bar{l}}^{\iota} \gamma^{\iota} \right] \\ &= 32 V_{q} V_{l} \Big\{ (p_{q} \cdot p_{l}) (p_{\bar{q}} \cdot p_{\bar{l}}) + (p_{q} \cdot p_{\bar{l}}) (p_{\bar{q}} \cdot p_{l}) \Big\} \\ &+ 32 A_{q} A_{l} \Big\{ p_{q \lambda} p_{\bar{q} \omega} p_{l}^{\kappa} p_{\bar{l}}^{\iota} (\delta_{\kappa}^{\lambda} \delta_{\iota}^{\omega} - \delta_{\iota}^{\lambda} \delta_{\kappa}^{\omega} \Big\} \\ &= 32 V_{q} V_{l} \Big\{ (p_{q} \cdot p_{l}) (p_{\bar{q}} \cdot p_{\bar{l}}) + (p_{q} \cdot p_{\bar{l}}) (p_{\bar{q}} \cdot p_{l}) \Big\} \\ &+ 32 A_{q} A_{l} \Big\{ (p_{q} \cdot p_{l}) (p_{\bar{q}} \cdot p_{\bar{l}}) - (p_{q} \cdot p_{\bar{l}}) (p_{\bar{q}} \cdot p_{l}) \Big\}. \end{aligned}$$

This gives us:

$$\begin{split} \overline{\mathcal{M}_{\gamma}\mathcal{M}_{Z}^{*}} &= \frac{Q_{q}Q_{l}e^{2}}{\hat{s}} \left(\frac{g_{Z}}{2}\right)^{2} \left(\frac{1}{\hat{s} - M_{Z}^{2} - iM_{Z}\Gamma_{Z}}\right) \frac{1}{4} \frac{1}{n_{c}} \Big[\\ & 32V_{q}V_{l} \Big\{ (p_{q} \cdot p_{l})(p_{\bar{q}} \cdot p_{\bar{l}}) + (p_{q} \cdot p_{\bar{l}})(p_{\bar{q}} \cdot p_{l}) \Big\} \\ & + 32A_{q}A_{l} \Big\{ (p_{q} \cdot p_{l})(p_{\bar{q}} \cdot p_{\bar{l}}) - (p_{q} \cdot p_{\bar{l}})(p_{\bar{q}} \cdot p_{l}) \Big\} \Big] \\ &= \frac{Q_{q}Q_{l}e^{2}}{\hat{s}} \left(\frac{g_{Z}}{2}\right)^{2} \left(\frac{1}{\hat{s} - M_{Z}^{2} - iM_{Z}\Gamma_{Z}}\right) \frac{1}{4} \frac{1}{n_{c}} 32 \left(\frac{\hat{s}}{4}\right)^{2} \Big\{ \\ & (V_{q}V_{l} + A_{q}A_{l})(1 + \cos\theta)^{2} + (V_{q}V_{l} - A_{q}A_{l})(1 - \cos\theta)^{2} \Big\}. \end{split}$$

The complex conjugate $\overline{\mathcal{M}_{\gamma}^* \mathcal{M}_Z}$ is exactly the same apart from the Z boson propagator whose imaginary part changes sign. The cross section for the interference part can therefore be written as:

$$\begin{split} \frac{d\sigma_{\text{int.}}}{d\Omega} &= \frac{1}{64\pi^2 \hat{s}} \Big(\overline{\mathcal{M}_{\gamma} \mathcal{M}_Z^*} + \overline{\mathcal{M}_{\gamma}^* \mathcal{M}_Z} \Big) \\ &= \frac{1}{64\pi^2 \hat{s}} \frac{Q_q Q_l e^2}{\hat{s}} \left(\frac{g_Z}{2} \right)^2 \left(\frac{2(\hat{s} - M_Z^2)}{(\hat{s} - M_Z^2)^2 + M_Z^2 \Gamma_Z^2} \right) \frac{1}{4} \frac{1}{n_c} 32 \left(\frac{\hat{s}}{4} \right)^2 \Big\{ \\ & (V_q V_l + A_q A_l) (1 + \cos \theta)^2 + (V_q V_l - A_q A_l) (1 - \cos \theta)^2 \Big\} \\ &= \frac{\alpha^2}{n_c} \frac{Q_q Q_l}{4 \sin^2 2\theta_W} \frac{\hat{s} - M_Z^2}{(\hat{s} - M_Z^2)^2 + M_Z^2 \Gamma_Z^2} \Big\{ \\ & (V_q V_l + A_q A_l) (1 + \cos \theta)^2 + (V_q V_l - A_q A_l) (1 - \cos \theta)^2 \Big\}. \end{split}$$

Once integrated over the solid angle we get:

$$\sigma_{\rm int.} = \frac{4\pi\alpha^2}{3\hat{s}} \frac{1}{n_c} \frac{2Q_q Q_l}{\sin^2 2\theta_W} V_q V_l \frac{\hat{s}(\hat{s} - M_Z^2)}{(\hat{s} - M_Z^2)^2 + M_Z^2 \Gamma_Z^2}.$$
 (A.20)

A.4 Drell-Yan Cross Section

The total cross section for the Drell-Yan process at leading order, for a given flavour of quark and charged leptons, is

$$\hat{\sigma}_{0}(q(p_{1})\bar{q}(p_{2}) \to l^{+}l^{-}) = \sigma_{\gamma^{*}} + \sigma_{\text{int.}} + \sigma_{Z}$$

$$= \frac{4\pi\alpha^{2}}{3\hat{s}} \frac{1}{n_{c}} \Big(Q_{q}^{2}Q_{l}^{2} + 2Q_{q}Q_{l}V_{l}V_{q}\chi_{1}(\hat{s}) + (A_{l}^{2} + V_{l}^{2})(A_{q}^{2} + V_{q}^{2})\chi_{2}(\hat{s}) \Big), \quad (A.21)$$

where

$$\chi_1 = \kappa \frac{\hat{s}(\hat{s} - M_Z^2)}{(\hat{s} - M_Z^2)^2 + M_Z^2 \Gamma_Z^2},$$

$$\chi_2 = \kappa^2 \frac{\hat{s}^2}{(\hat{s} - M_Z^2)^2 + M_Z^2 \Gamma_Z^2},$$

$$\kappa = \frac{1}{\sin^2 2\theta_W}.$$

Appendix B

Single Channel Differential Cross Sections

In this appendix, the differential cross sections are shown individually for the two decay channels. For each observable, the data cross section measurement is drawn on top of the Drell-Yan MC generated distributions. Additionally, the lower panels in each figure show the ratios of the theory predictions to data. Errors bars around the experimental points show the statistical uncertainty, while the crosshatched bands indicate the statistical plus systematic uncertainties added in quadrature. The coloured filled band around the MC prediction represents the statistical uncertainty of the generated sample.



Figure B.1: Differential cross section as a function of the exclusive jet multiplicity, for the (left) the muon and (right) the electron decay channels.



Figure B.2: Differential cross section as a function of the inclusive jet multiplicity, for the (left) the muon and (right) the electron decay channels.



Figure B.3: Differential cross section as a function of the 1st leading jet $p_{\rm T}$ with $N_{\rm jets} \ge 1$, for the (left) the muon and (right) the electron decay channels.



Figure B.4: Differential cross section as a function of the 2nd leading jet $p_{\rm T}$ with $N_{\rm jets} \ge 2$, for the (left) the muon and (right) the electron decay channels.



Figure B.5: Differential cross section as a function of the $3^{\rm rd}$ leading jet $p_{\rm T}$ with $N_{\rm jets} \ge 3$, for the (left) the muon and (right) the electron decay channels.



Figure B.6: Differential cross section as a function of the 4th leading jet $p_{\rm T}$ with $N_{\rm jets} \ge 4$, for the (left) the muon and (right) the electron decay channels.



Figure B.7: Differential cross section as a function of the 5th leading jet $p_{\rm T}$ with $N_{\rm jets} \ge 5$, for the (left) the muon and (right) the electron decay channels.



Figure B.8: Differential cross section as a function of the 1st leading jet $|\eta|$ with $N_{\text{jets}} \ge 1$, for the (left) the muon and (right) the electron decay channels.



Figure B.9: Differential cross section as a function of the 2nd leading jet $|\eta|$ with $N_{\text{jets}} \geq 2$, for the (left) the muon and (right) the electron decay channels.



Figure B.10: Differential cross section as a function of the $3^{\rm rd}$ leading jet $|\eta|$ with $N_{\rm jets} \ge 3$, for the (left) the muon and (right) the electron decay channels.



Figure B.11: Differential cross section as a function of the 4th leading jet $|\eta|$ with $N_{\text{jets}} \ge 4$, for the (left) the muon and (right) the electron decay channels.



Figure B.12: Differential cross section as a function of the 5th leading jet $|\eta|$ with $N_{\text{jets}} \ge 5$, for the (left) the muon and (right) the electron decay channels.



Figure B.13: Differential cross section as a function of the jets $H_{\rm T}$ with $N_{\rm jets} \ge 1$, for the (left) the muon and (right) the electron decay channels.



Figure B.14: Differential cross section as a function of the jets $H_{\rm T}$ with $N_{\rm jets} \ge 2$, for the (left) the muon and (right) the electron decay channels.



Figure B.15: Differential cross section as a function of the jets $H_{\rm T}$ with $N_{\rm jets} \geq 3$, for the (left) the muon and (right) the electron decay channels.



Figure B.16: Differential cross section as a function of the jets $H_{\rm T}$ with $N_{\rm jets} \ge 4$, for the (left) the muon and (right) the electron decay channels.



Figure B.17: Differential cross section as a function of the jets $H_{\rm T}$ with $N_{\rm jets} \ge 5$, for the (left) the muon and (right) the electron decay channels.



Figure B.18: Differential cross section as a function of the dijet mass with $N_{\text{jets}} \ge 2$, for the (left) the muon and (right) the electron decay channels.