STUDY OF Z+JETS WITH CMS DETECTOR AT LHC

A thesis

submitted to the

FACULTY OF SCIENCE

PANJAB UNIVERSITY, CHANDIGARH

for the degree of

DOCTOR OF PHILOSOPHY

2012

Lovedeep Kaur

CENTRE OF ADVANCED STUDY IN PHYSICS

PANJAB UNIVERSITY, CHANDIGARH

To my family ...

Acknowledgement

Current thesis is an outcome of my work within the CMS collaboration for past five years. I have thoroughly enjoyed going through all the intricacies of acquisition, retrieval, analysis and presentation of experimental data recorded. First of all, I would like to thank my Ph.D. supervisor Prof. Suman B. Beri for introducing me to the experimental high energy physics. It was her fruitful suggestions and guidance, which enabled me to work with such a stimulating environment like CMS collaboration at CERN. Her continuous encouragement, guidance and support helped me to finish this thesis work.

I also thank the Chairman, Department of Physics for providing the infrastructural support and allowing me to use facilities.

My heartiest thanks are also due to European Center for Nuclear Physics (CERN), Fermi National Lab (FNAL) and University of Wisconsin (UW) for providing major part of the computing resources and also hosting me at different points of time. I am greatly indebted to CMS collaboration for the great culture of performance and mutual cooperation. The work presented in this thesis would not have been possible without the valuable support and help from scientists Dr. Bryan Dahmes and Dr. Kalanand Mishra. They are the ones who taught me the basics of the collider physics and data analysis. I sincerely express my thanks and gratitude to Dr. Sridhara Rao Dasu, Dr. Shin-Shan Eiko Yu, Dr. Anne-Marie, Dr. Darko Mekterovic and Dr. Vitaliano Ciulli for their valuable suggestions during my thesis analysis and approval. I also thank all the fellow students in CMS in general and Panjab University's experimental high energy group in particular for helping me at various moments. I also appreciate the help and positive attitude shown by other members of faculty and engineering team at the group. I thank to Prof. J. B. Singh, Prof. M. Kaur and Dr. V. Bhatnagar for their continuous support. I would like to thank all members of INDIA-CMS collaboration where regular collaboration meetings helped for providing a platform to discuss the analysis work.

I am greatly indebted to my mother, siblings and extended family of my in-laws. Without their help, encouragement and affection I could have never completed this thesis. I also thank Dr. Anil P. Singh, and our daughter Ershleen Thakur for their charming and reassuring presence. Finally I thank all my friends and colleagues for these five years at University.

The financial support by University Grants Commission (U.G.C., India), Depart-

ment of Science and Technology (D.S.T., India), Department of Atomic Energy (D.A.E., India), and American Physical Society (A.P.S., USA), is highly acknowledged without which it would not have been possible to finish this work.

Dated: Lovedeep Kaur

Abstract

A study of jet production in association with a Z boson in proton-proton collisions at Large Hadron Collider (Geneva, Switzerland) at a 7 TeV center-of-mass energy using an integrated luminosity of 4.89 fb^{-1} collected by CMS detector, has been presented. The measured jet multiplicity distributions are corrected for signal selection efficiency and unfolded for detector effects and compared with theoretical predictions. The results are important for understanding and testing the predictions from perturbative Quantum Chromodynamics (pQCD) which describes the strong interactions between quarks and gluons. Furthermore, some of new physics processes at hadron colliders, such as the production of Higgs bosons and super-symmetric particles, can be mimicked by the production of vector bosons in association with jets that constitute irreducible backgrounds to these searches. Therefore, the study and understanding of Z+jets processes are very important in order to estimate the significance of such new signals. We found no major irregularity or disagreement between theory and data for the part of parameter-space under investigations.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The Standard Model of Particle Physics

Elementary particle physics deals basically with the study of the ultimate constituents of matter and the nature of the interactions between them. It is well known that the everyday life is properly described by Newtons laws of classical mechanics. But for objects, that travel at speeds comparable to the speed of light c, the classical rules need to be modified by special relativity; furthermore, for objects that are very small (roughly at the subatomic level), classical mechanics is superseded by quantum mechanics. As elementary particles are both fast and small, their description falls under the domain of the quantum field theory (QFT) [1]. The Standard Model of Particle Physics (SM) [2, 3, 4, 5] is a QFT description of the fundamental building blocks of matter and the interactions between them. It was formulated in 1960s and 1970s and it successfully explains the results of all particle physics experiments performed to date. In this chapter, we present a very brief summary of standard model and also the challenges that it face.

1.1.1 The Fundamental Particles and Interactions

The Fermions

Today, the so called fundamental particles can be divided into two categories namely the force careers and the matter particles. Former are the mediators of interactions, and latter are thought of as the building block of the matter called fermions. The fermions can be subdivided into two families; *leptons and quarks* [6]. The properties of the six leptons (the electron, muon, tau and associated neutrinos) and the six quarks (up, down, charm, strange, beauty and top) are given in Table 1.1.

Quarks		
Fermion	mass	charge
up (u)	3 MeV	+2/3
down (d)	6 MeV	-1/3
charm (c)	1.3 GeV	+2/3
strange (s)	100 MeV	-1/3
top (t)	172 GeV	+2/3
bottom (b)	4.2 GeV	-1/3

Leptons		
Fermion	mass	charge
electron	115 keV	-1
ν_e	< 2 eV	0
muon	106 MeV	-1
ν_μ	< 0.170 MeV	0
tau	1.78 GeV	-1
ν_τ	< 15.5 MeV	0

Table 1.1: Tables of the fermions and their properties, including charge and mass (approximate for quarks, as they cannot be isolated). Antimatter [7] has opposite charges.

Gauge Symmetries and the Bosons

In the quantum field formalism the fermions are states of a quantum field ϕ. In an analogue to classical mechanics the properties of this field, including the fermion dynamics, are completely described by the Lagrangian density. Interactions between the matter particles are introduced into the theory by imposing gauge symmetry conditions on the quantum fields [8, 9, 10]. This symmetry requires that the Lagrangian remains invariant under a Lie group of local space-time transformations of the field ϕ. The most important groups in this context are $U(n)$, the group of $n \times n$ Unitary matrices, and $SU(n)$, the group of $n \times n$ Special Unitary matrices (required to have determinant equal to one). In quantum theory these transformation matrices are built up from quantum mechanical

generators which correspond to a particular observable, such as charge. An $SU(n)$ group of transformations has $n^2 - 1$ generators. The gauge invariance condition requires that the observable associated with the generator of the transformation is conserved. In order for this condition to be met, it is found that one has to introduce additional gauge field terms to the Lagrangian, one for each generator, and these gauge fields describe the propagation of fundamental spin 1 force carrying particles, the *bosons*. The properties of the SM bosons are described in Table 1.2 [6].

Table 1.2: Properties of bosons. Charges refer to electric charge and are in units of electron charge.

Boson	Charge	Mass	Force Mediated
photon	0	0	EM
W	+1,-1	80.4	weak
Z	0	91.2	weak
gluon	0	0	strong

Each boson couples to a certain "charge" and is thus responsible for the mediation of a certain force, or interaction, in which the total charge must always be conserved. The SM describes three of the four fundamental forces of nature; the electromagnetic, weak and strong forces. Currently, there exists no quantum field formalism for gravitational interactions. However, the force of gravity is sufficiently weak that it can be ignored at energies less than the Planck scale $E < M_{Planck} \approx 10^{19} GeV$.

Quantum Electrodynamics (QED)

The theory of Quantum Electrodynamics (QED) [11, 12, 13] describes electromagnetic interactions between electrically charged particles. QED requires gauge invariance under the $U(1)$ group of transformations. The $U(1)$ group has one generator that is associated with the observable of electric charge, and hence one boson, the photon, which mediates the electromagnetic interaction between charged particles, in which charge must always be conserved. Since there is only one generator the $U(1)$ transformations are commutative (the group Abelian), and this means that the associated bosons cannot interact with

one another i.e. the photon is electrically neutral. The coupling strength in QED is proportional to the electric charge of the particle.

Electroweak Theory

Electroweak theory provides a unified description of both the electromagnetic and weak interactions [14, 15, 16, 17]. It demands invariance under the $SU(2) \times U(1)$ group of transformations. This requires the introduction of four gauge bosons. The $SU(2)$ group has $2^2 - 1 = 3$ bosons, which we label W^-, W^+ and W^0. We have already seen that $U(1)$ symmetry introduces one boson into the theory. In this context we label this boson B^0. In contrast to the $U(1)$ transformations, the $SU(2)$ matrices are non-commutative, making electroweak theory non-Abelian. This leads to the important result that the gauge bosons of the theory are self-interacting. The W^0 and B^0 states mix quantum mechanically, producing the physical states of the Z^0 boson (associated with the weak interaction) and the photon (associated with the electromagnetic interaction).

In principle electroweak bosons can couple to any fermion, but interactions must conserve electric charge and also lepton (quantum) number. This restricts the basic interaction vertices in the lepton sector; couplings only occur between leptons within the same generation or doublet. Some examples of basic electroweak processes in the lepton sector are given in Figure 1.1. The coupling strengths g_W and g_z of the W and Z bosons respectively are independent of the particular lepton types (e, μ, τ) involved, a property known as *lepton universality.*

The theory also requires that there can be no flavour changing neutral currents. However, flavour changing charged currents are permitted, and hence the $W^\pm$ bosons "convert" quarks from one flavour to another. It is this feature of electroweak inter actions which is responsible for radioactive β decay, in which a neutron is converted into a proton via a $d \to W^- u$ electroweak vertex, illustrated in Figure 1.2. For W interaction vertices involving the *ud, cs* and *tb* doublets the coupling strength is the same as for

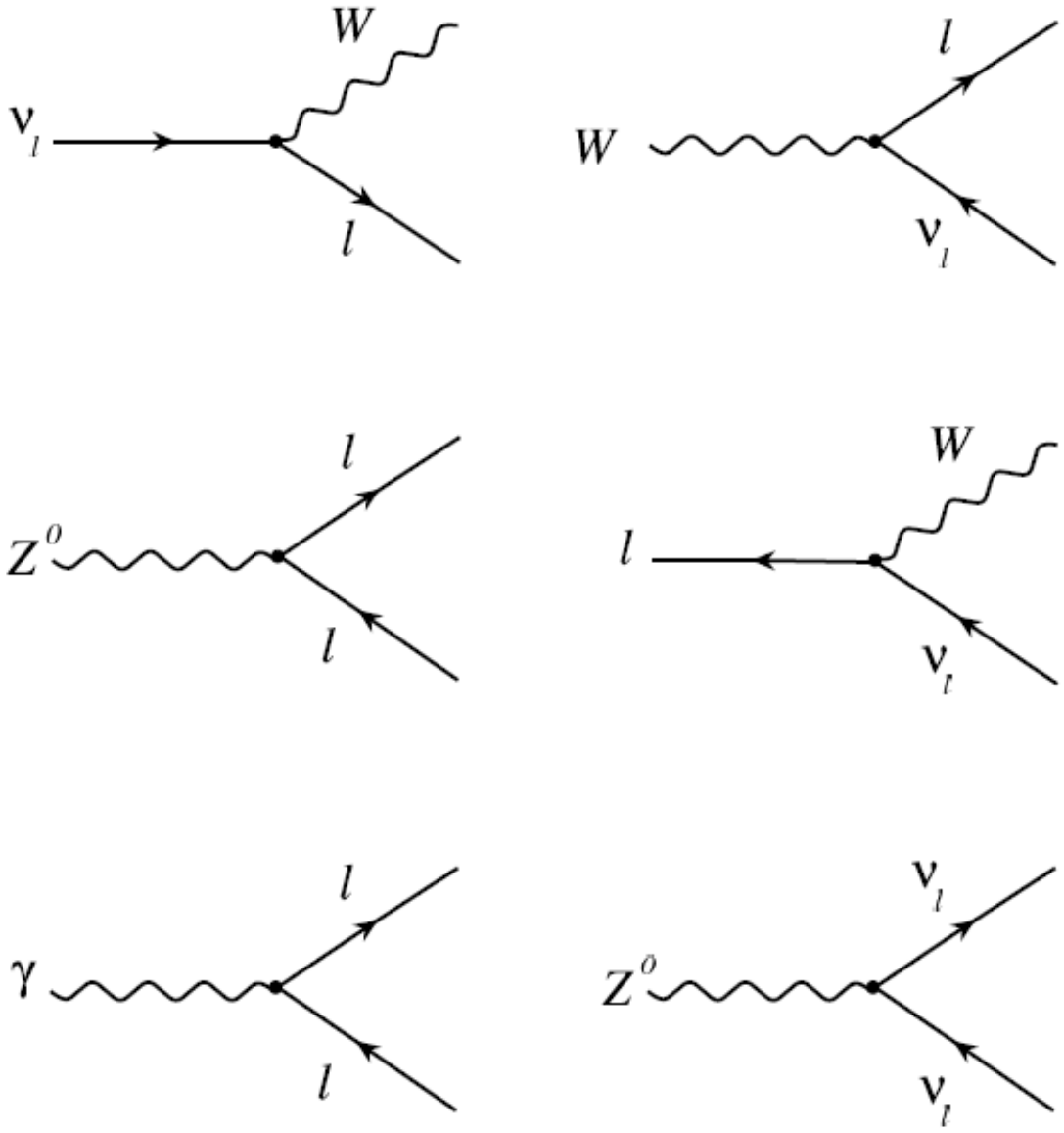


Figure 1.1: Some of the possible processes involving electroweak interaction vertices in the lepton sector. The lepton on each line must be the same type l to conserve lepton number.

the lepton sector W vertices, g_W. However, unlike the leptonic interactions, the W can interact with quarks of differing generations, albeit with a reduced coupling strength. The coupling of the W boson to a particular flavour changing vertex is modified according to the experimentally determined coefficients of the *Cabbibo Kobayashi Maskawa* (CKM) matrix, shown in Equation (1.1) below [18], such that the modified coupling strength

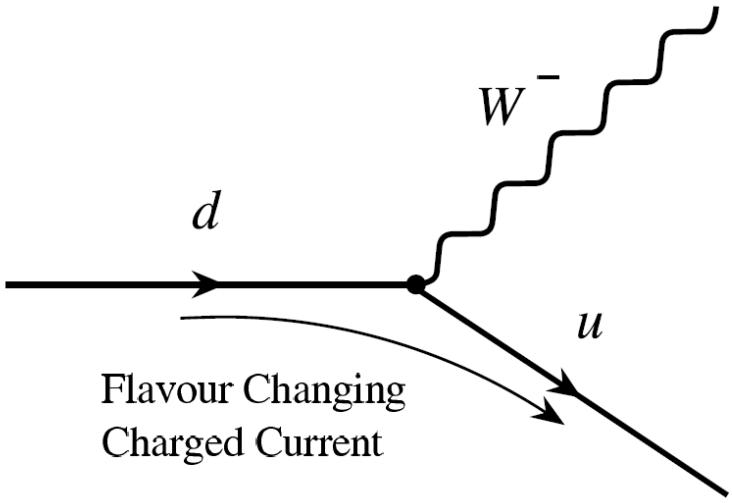


Figure 1.2: Diagram of the $d \rightarrow W^{-}u$ electroweak vertex.

between quarks of flavours x and y is $g_W^{'} = V_{xy} g_W$.

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \approx \begin{pmatrix} 0.964\ to\ 0.975 & 0.221\ to\ 0.227 & 0.005\ to\ 0.014 \\ 0.221\ to\ 0.227 & 0.973\ to\ 0.974 & 0.037\ to\ 0.043 \\ 0.003\ to\ 0.005 & 0.039\ to\ 0.044 & 0.999\ to\ 0.999 \end{pmatrix} \tag{1.1}$$

See that for *ud, cs* and *tb* vertices of the CKM coefficient is essentially unity. The coupling between quarks of different generations is said to be *Cabbibo suppressed.*

The Higgs Mechanism

When electroweak theory was first postulated there was one major obstacle to its acceptance. All four bosons are predicted by the theory to be mass-less and this is in direct conflict with what is observed in nature. If mass terms for the particles were simply added to the Lagrangian the gauge symmetry would be broken. An alternative solution is required. In the so-called *Higgs mechanism* [19, 20] an additional scalar field, the Higgs field, is introduced into the theory, which has a potential function of a form allowing degenerate vacuum solutions with a non-zero vacuum expectation value. The vacuum states are invariant only to the $U(1)$ transformations of the electromagnetic theory, and thus the $SU(2)U(1)$ symmetry is *"spontaneously broken"*. Interaction with the Higgs field

allows the W and Z bosons, as well as other particles, to acquire mass, with the value of the mass dependent on the strength of the coupling. However, the mediator of this field, the spin-0 Higgs boson, has yet to be observed experimentally (see section 1.1.6).

Quantum Chromodynamics (QCD)

The theory of Quantum Chromodynamics (QCD) [21, 22, 23] describes the strong interaction of quarks and requires invariance under $SU(3)$ transformations. In this theory a quark can be in one of three colour states (red, green or blue) which form the fundamental representation of the group. The group has $3^2 - 1 = 8$ generators and thus gauge invariance requires the introduction of 8 mass-less *gluon fields*. These gluons couple to colour charge and since the $SU(3)$ group is non-Abelian the gluons are self-interacting, carrying colour charge themselves. As will be explained in following sections, this has far reaching consequences for the predictive power of QCD theory.

1.1.2 Calculation of Observables in QFT

For any theory to be truly successful it has to have the power to predict observables; quantities that can be measured experimentally and hence allow the theoretical model to be tested. The cross-section for a particular well-defined process is a very useful observable to calculate as it can be directly measured at particle collider experiments. It is the effective area over which the particles in the initial state interact to produce the final state, and is directly proportional to the rate of a process.

QED processes involve the coupling of mass-less photon fields to charged particles, and QCD similarly involves the coupling of mass-less gluon fields to coloured particles. Although we shall see that the strength of this coupling, and particularly its dependency on the energy scale of the interaction, are crucially different in QED vs QCD, formally the procedure for the calculation of cross-sections in QED and QCD processes is exactly the same. In this section we shall outline the calculation of the cross-section for a $2 \rightarrow 2$ QED

process; $e^+e^- \rightarrow \gamma* \rightarrow \mu^+\mu^-$ scattering [2]. Figure 1.3 shows a possible Feynman diagram for this process, involving two QED coupling vertices. An analogous QCD process would be $q\bar{q} \rightarrow g* \rightarrow q\bar{q}$, a possible Feynman diagram of which is shown in Figure 1.4, involving two QCD coupling vertices.

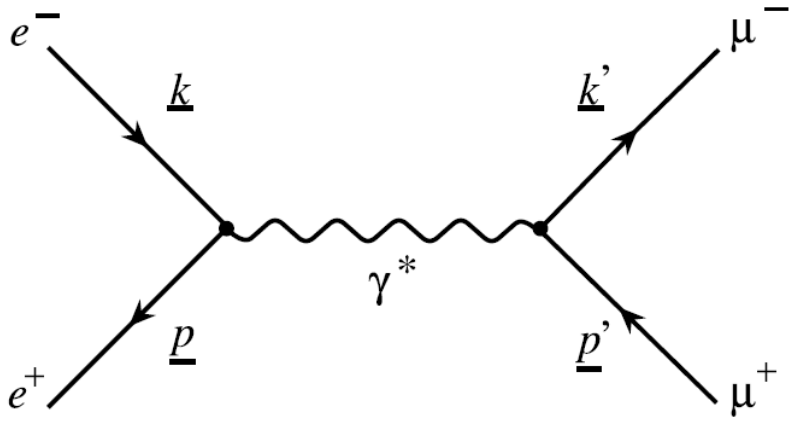


Figure 1.3: The Lowest order Feynman diagram for the process $e^+e^- \rightarrow \gamma* \rightarrow \mu^+\mu^-$

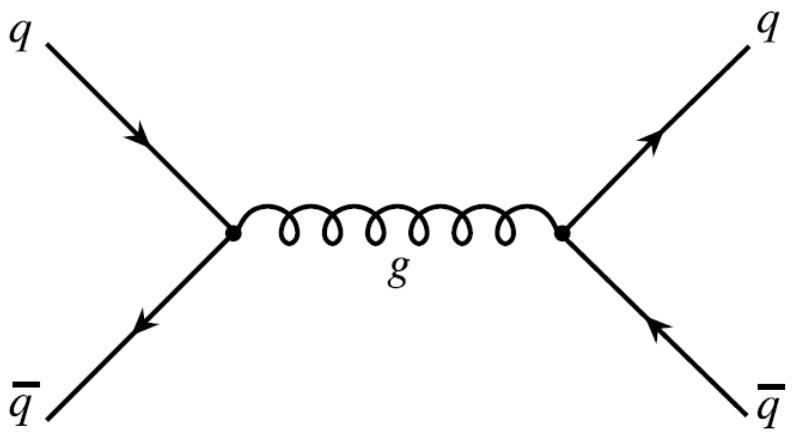


Figure 1.4: The Lowest order Feynman diagram for the process $q\bar{q} \rightarrow g* \rightarrow q\bar{q}$

Shown on the diagram in Figure 1.3 are the four-vectors $\underline{k}, \underline{k'}, \underline{p}$ and $\underline{p'}$ of the initial and final state particles. From such a diagram one can calculate a probability amplitude Γ. Each line in the Feynman diagram can be related to a propagator term in the amplitude calculation, and each interaction vertex introduces a factor $\sqrt{\alpha}$. In this case α is the QED coupling constant, for a QCD process it would be the QCD coupling

constant α_s. The squared modulus of the amplitude for the diagram in Figure 1.3, $|\Gamma|^2$, is given by:

$$|\Gamma|^2 = 32\pi^2\alpha^2(\frac{t^2+u^2}{s^2}) \tag{1.2}$$

where the variables s, t and u are frame invariant scalar quantities formed from the initial and final state particle four-vectors as shown below:

$$s = (\underline{k} + \underline{p})^2$$

$$t = (\underline{k} - \underline{k'})^2$$

$$u = (\underline{k} - \underline{p'})^2$$

Note that $\sqrt{s}$ is the centre-of-mass energy of the interaction. To obtain a cross-section the squared modulus of the amplitude, known as the interaction probability, must be integrated over all possible final and initial state phase space i.e. all kinematically allowed values of $\underline{k}, \underline{k'}, \underline{p}$ and $\underline{p'}$. Performing such an integration yields the following cross-section for $e^+e^- \rightarrow \gamma* \rightarrow \mu^+\mu^-$ scattering:

$$\sigma(e^+e^- \rightarrow \mu^+\mu^-) = \frac{4\pi\alpha^2}{3s} \tag{1.3}$$

Consider this process in the centre-of-mass frame. The total energy in the initial state is $\sqrt{s}$, and the total momentum zero. If the photon propagator is to obey energy and momentum conservation it must acquire a non-zero rest mass equal to $\sqrt{s}$ via a quantum fluctuation permitted by the Heisenberg uncertainty principle. The photon is described as being *off-mass-shell* or *virtual*, and the *virtuality* of the photon Q is in this case the effective rest mass of the photon, or viewed another way, the energy scale at which the quantum fluctuations which permit the interaction occur. Via the uncertainty principle this energy scale Q can be related to a distance scale of order $1/Q$, and it is common also to talk of the interaction occurring at short or large distance scales, depending on the size of Q. Note that since in this case $Q = \sqrt{s}$, the $e^+e^- \rightarrow \gamma* \rightarrow \mu^+\mu^-$ scattering

cross-section given in Equation (1.3) depends on $1/Q^2$: as the required virtuality of the photon propagator increases the cross-section decreases. It is also proportional to α^2, a consequence of there being two interaction vertices in the Feynman diagram considered.

However, Figure 1.3 is far from the only diagram that can be drawn for the $e^+e^- \rightarrow \gamma* \rightarrow \mu^+\mu^-$ process. Just a few of the possible higher order diagrams are shown in Figure 1.5. These diagrams involve two additional interaction vertices, and

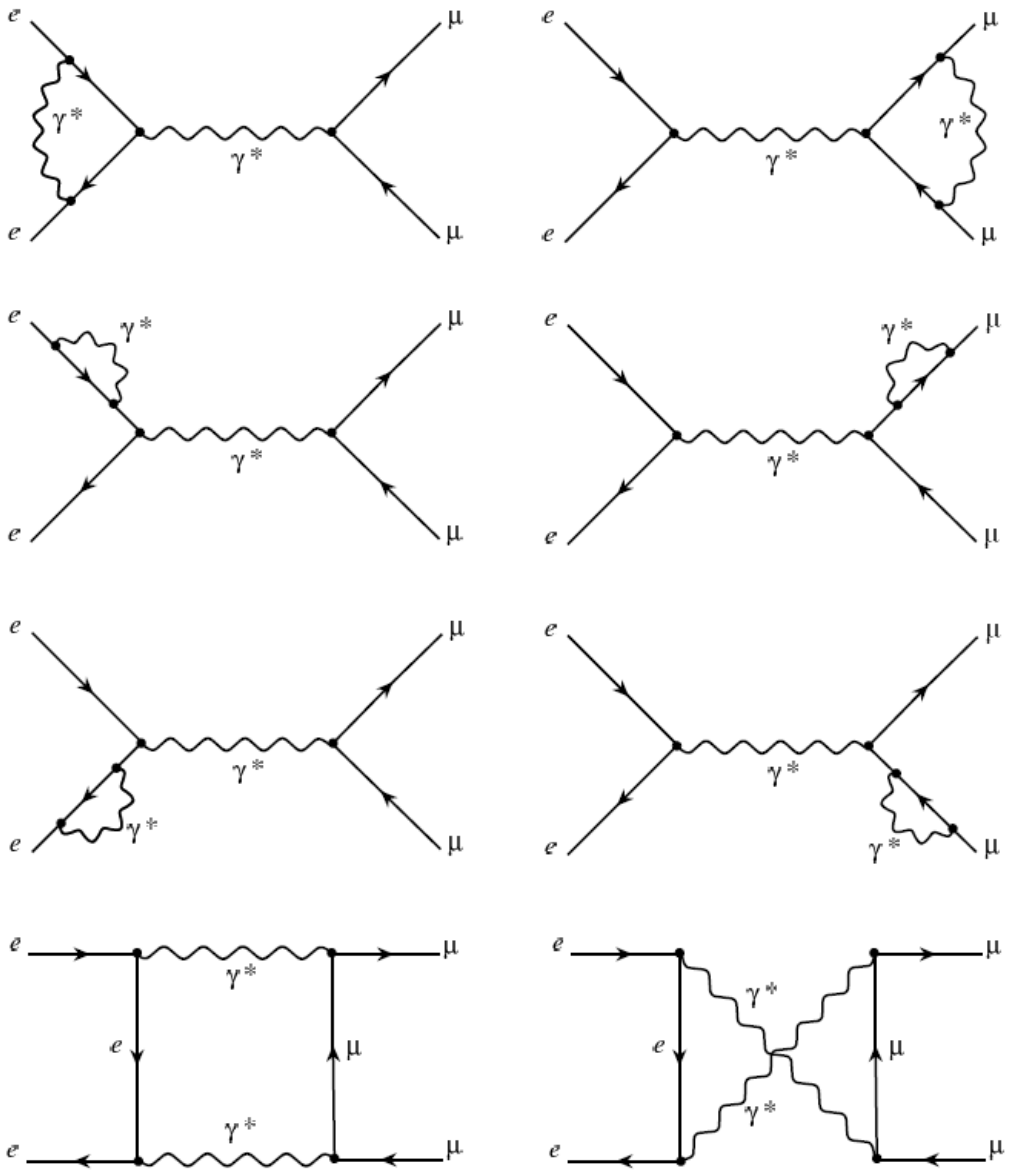


Figure 1.5: Higher order diagrams for the process $e^+e^- \rightarrow \gamma* \rightarrow \mu^+\mu^-$

thus have an associated amplitude involving a factor α^2. The first six are referred to as loop diagrams, because they contain virtual photon loops as shown, and since there is essentially no restriction on the number of photon loops which are permitted in the theory an infinite number of such diagrams can be drawn. A complete evaluation of the cross-section would involve calculating the amplitude for every possible diagram, summing all these amplitudes to get the total amplitude Γ_{tot}, squaring Γ_{tot} to get the interaction probability and then integrating over phase space. The total amplitude will be of the form:

$$\Gamma_{tot} = A_1\alpha + A_2\alpha^2 + \ higher\ orders\ ... \tag{1.4}$$

where A_1 is Γ/α and similarly A_2 is Γ/α^2, where Γ_2 is the total amplitude of the diagrams in Figure 1.5. The interaction probability, directly proportional to the cross-section, is then given by:

$$|\Gamma_{tot}|^2 = A_1^2\alpha^2 + 2Re(A_1.A_2)\alpha^3 + A_2^2\alpha^4 + \ higher\ orders\ ... \tag{1.5}$$

This is known in QFT as a *perturbation series*. Here we have only defined three terms, but since the number of loop diagrams one can draw is infinite, the series is infinite also. Thus in practice when performing a cross-section calculation one has to terminate the perturbation series at some fixed order. Terminating the series at $O(\alpha^2)$ would in this case be defined as the *leading order* (LO) cross-section calculation, and only involve evaluating the amplitude of the diagram in Figure 1.3. Terminating the series at $O(\alpha^3)$ would be defined as the *next-to-leading-order* (NLO) calculation, and would involve evaluating the amplitudes of the diagrams in Figure 1.5, in particular the complex interference of these amplitudes with that of Figure 1.3. Similarly terminating the series at $O(\alpha^4)$ would be the *next-to-next-to-leading-order* (NNLO) calculation, and so on.

Although theoretically the complete perturbation series is an exact expression for the total cross-section of a process, several issues remain if it is to be practically useful:

- Unless the coupling constant α is sufficiently small the series will not converge to

a finite cross-section and termination of the series will be subject to large higher order corrections.

- When integrating the amplitudes over all phase space to obtain the cross-section we encounter a problem with loop diagrams. Within the loop the momentum of the photon propagator is unconstrained and thus the integral is divergent, known as *ultraviolet divergence*. These issues will be addressed in the context of QED and QCD in the sections that follow.

1.1.3 Coupling Strength in QED

Recall that in the gauge theory of QED the couping α is proportional to the charge of the particle. The inherent uncertainties of quantum mechanics mean that the charge of a particle is not as well defined as you might expect. In QED quantum fluctuations allowed for by the Heisenberg uncertainty principle permit an electron to emit a photon with virtuality Q. This fluctuation occurs on a distance scale of order $1/Q$. The virtual photon can in turn produce a virtual e^+e^- pair, which in turn can radiate a photon, and so on, such that at sufficiently large distance scales the electron is surrounded by a virtual cloud of electron-positron pairs as illustrated in Figure 1.6. This cloud becomes polarized, with the virtual particles of opposing charge being attracted to the electron charge, and the virtual particles of like charge being repelled, as illustrated in Figure 1.7. The net effect is that the bare electron charge is screened and reduced. The implication of this is that the QED coupling strength α is dependent on the energy scale Q of the interaction. At large Q scales, short distance scales, less of the virtual cloud is resolvable and the coupling is larger than at small Q scales.

The theory of QED is renormalised (see Section 1.1.5) such that the coupling constant α is proportional to the experimentally measured charge e as follows:

$$\alpha = e^2/4\pi \tag{1.6}$$

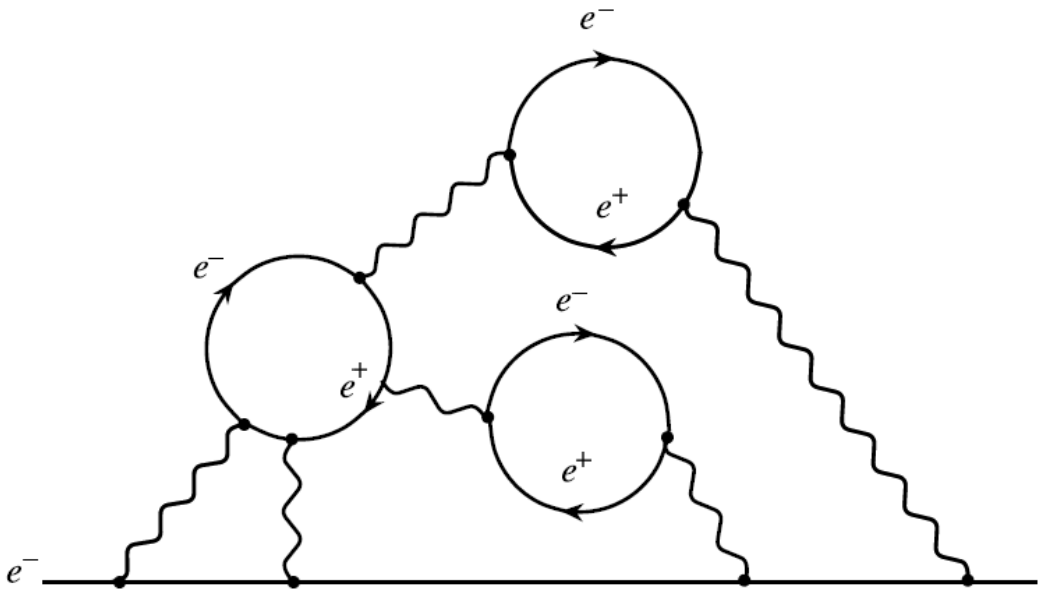


Figure 1.6: Diagram illustrating the formation of a virtual cloud of electron-positron pairs around a bare electron charge

Effectively we have evaluated the coupling at $Q = 0$; this is the fine structure constant $\alpha^{-1} \approx 137$. As the energy scale Q decreases the coupling strength decreases asymptotically to the fine structure constant. Thus in the context of QED all energy scales currently achievable in experiments can be considered small, that is, the interactions occur on a large distance scale such that the coupling α is $<< 1$. This makes perturbation theory a very effective tool for QED predictions, as α is always small enough that a perturbation series will converge after only a few terms.

1.1.4 Asymptotic Freedom and Confinement in QCD

In QCD the coupling strength is proportional to the colour charge of the particle. Similarly to QED, a bare colour charge such as a quark or gluon will be surrounded by a virtual cloud of quark-antiquark pairs, illustrated in Figure 1.8, that form on a distance scale $1/Q$, where Q is the energy of the interaction. However, since gluons are self-interacting, the cloud will also contain gluon-gluon pairs which themselves carry colour charge, also shown in Figure 1.8. This is a crucial difference with respect to QED. It means that the

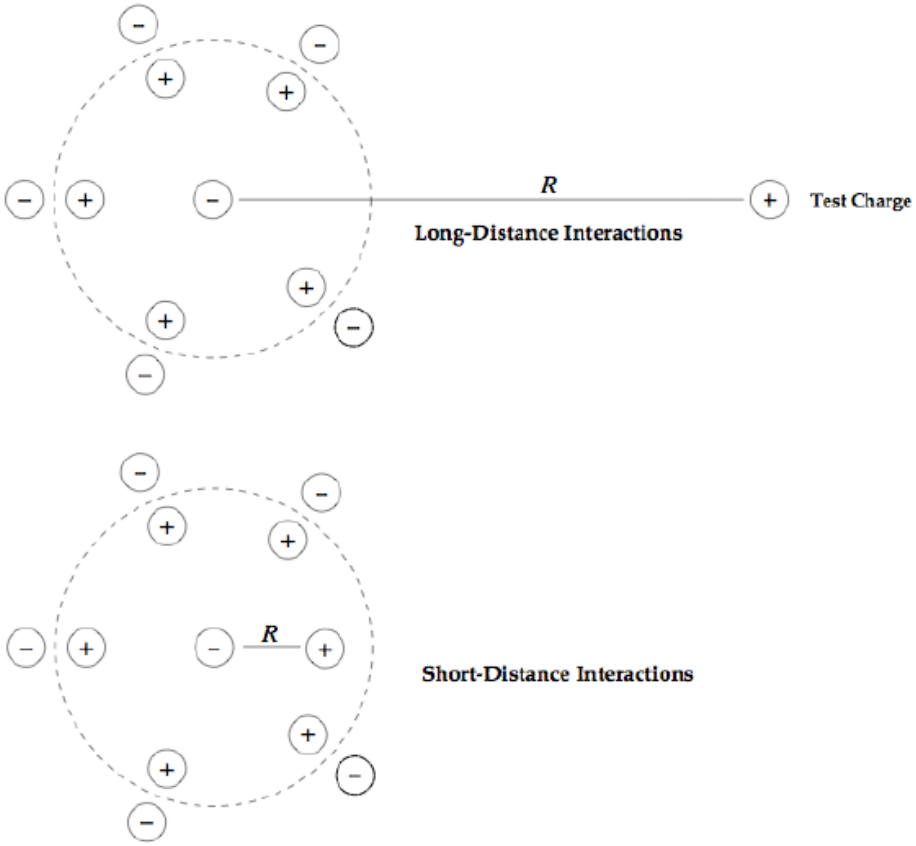


Figure 1.7: Diagram illustrating the effect of QED charge screening by virtual electron-positron pairs

net effect of the virtual cloud in QCD is not to screen and reduce the colour charge, but to increase it. Thus in QCD as the energy scale Q decreases, the resolved colour charge, and hence the strength of the strong coupling, increases. Taken to its limit, over an infinitely large distance scale the QCD coupling strength, and hence potential energy, would itself be infinite. The mass-less property of the gluon means that the strong force has infinite range, and thus it is always preferential to create lower energy colourless bound states of quarks. This is the property of colour confinement; that the physical states in QCD are not the coloured quarks and gluons but colourless bound states of three quarks, the hadrons (rgb or rgb), or two quarks, the mesons (rr,gg or bb). Final-state quarks and gluons created in hard-scatter interactions at hadron colliders undergo a process called hadronization or fragmentation [21] and are observed as jets of hadrons in the detector.

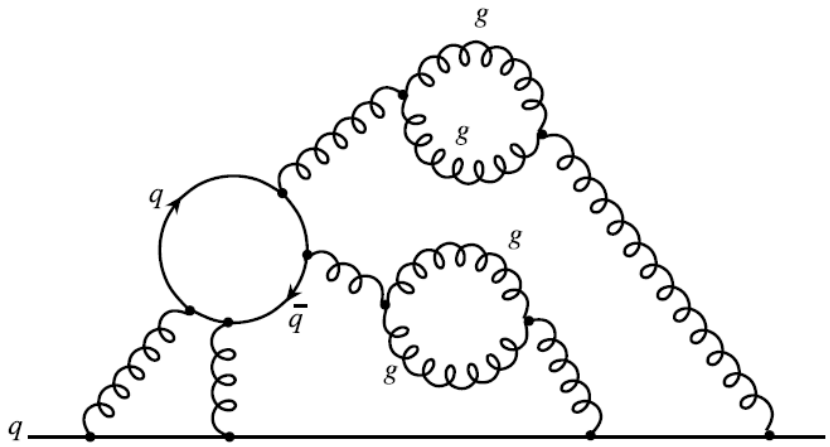


Figure 1.8: Diagram illustrating the formation of a virtual cloud of quark-antiquark pairs around a bare electron charge

Hadronization is a soft QCD process, meaning that it occurs at a low Q scale where the strong coupling constant α_s is too large for perturbation theory to be possible. The perturbation series for the process will not converge to a finite result. At higher Q scales the coupling constant becomes small enough for the perturbation series to converge and perturbative predictions of hard processes can be made. This is known as asymptotic freedom. Precisely what scales we are referring to when we talk of hard and soft processes is detailed in the next section. Figure 1.9 shows the running of α_s with scale $Q = \mu$. As Q increases the coupling decreases. One can see that at the energy scale of the Z boson mass $\alpha_s \approx 0.12$, and that this is not going to reduce significantly at higher energies. This is roughly one order of magnitude larger than the QED coupling constant, and as a consequence the termination of series in a QCD perturbative calculation is subject to more significant higher order corrections than in QED.

1.1.5 Renormalisation

In theory the charge screening effects of QED and QCD could be calculated using an infinite series of Feynman diagrams containing photon propagator loops like those of Figure 1.5. Unfortunately, as we have already noted, these diagrams themselves contain

ultraviolet divergences. Renormalisation is the introduction of one or more dimensionless physical parameters into the theory, dependent on some unphysical renormalisation scale, which effectively absorbs these ultraviolet divergences and thus retrieves the ability to make perturbative calculations [2, 21]. In QED, the renormalisation parameter is the experimentally measured charge of the electron e. We know that this quantity must take into account the ultraviolet divergences, as effectively it measures the electron charge at $Q = 0$. Therefore, if we introduce the measured charge e into the theory as a parameter we can effectively absorb these ultraviolet divergences. This is done by defining the QED coupling constant at $Q = 0$ as the fine structure constant (Equation(1.6)).

It is important to note that the choice of energy scale at which the theory of QED can be renormalised is not unique. Choosing $Q = 0$ is simply convenient since the fine structure constant can be measured very precisely at this scale. If we were to evaluate the electric charge at a higher scale $Q > 0$, we would, as we have already seen, resolve a larger electric charge and hence have a larger QED coupling constant. This implies that by introducing a renormalisation parameter (evaluated at a particular scale) into a quantum field theory one must also necessarily introduce a dependence of this parameter on a renormalisation energy scale Q. As we have already noted, the coupling strength varies with Q, the so-called running coupling constant. The dependence of a renormalised coupling constant on Q is determined by the renormalisation group equation (RGE) [2]:

$$Q^2 \frac{\delta \alpha}{\delta Q^2} = \beta(\alpha) \tag{1.7}$$

Where the β function can be determined to a certain order using perturbation theory, thus allowing for a solution $\alpha(Q^2)$ to a certain order.

In QCD, the property of confinement means that we cannot renormalise the theory at $Q = 0$. Instead, we have to measure the strong coupling constant in processes that occur at large Q scales, where perturbation theory can be applied. The leading order solution to the RGE (Equation(1.8)) gives the dependence of the strong coupling α_s on

the renormalisation scale Q^2 as:

$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 2n_f)ln(Q^2/\Lambda^2_{QCD})} \tag{1.8}$$

where n_f is the number of active fermion flavours and Λ_{QCD} is a scale parameter that must be determined from experimental measurement of α_s at known Q scale. Such measurements yield $\Lambda_{QCD} = 0.2 \pm 0.1$GeV. Physically Λ_{QCD} may be interpreted as the energy scale at which perturbation theory breaks down, as $\alpha_s \to 1$. Perturbative evaluation of QCD processes can only be made if $Q >> \Lambda_{QCD}$. Thus in some sense Λ_{QCD} is the boundary between soft and hard QCD processes. Figure 1.9 shows the running of α_s with scale $Q = \mu$. The points are experimental measurements of α_s at particular renormalisation scales, and the curve is a NNLO evaluation of the running of α_s using the average Λ_{QCD} determination of these results [6].

1.1.6 Unanswered Questions

The Standard Model is a hugely successful description of the fundamental particles that make up the Universe and the interactions between them. However, a number of important experimental observations are not explained by the SM. Additionally, at the present time not all of the components of the SM have been verified experimentally. Currently there exists no direct experimental evidence for the final fundamental particle predicted by the SM, the Higgs boson. However, precise measurement of SM parameters constrain the allowed Higgs mass in the SM framework. Figure 1.10 shows how a Higgs decays as a function of its mass. Because of quantum mechanics, we do not know exactly what will happen. We only know the probability of a Higgs decay For example, Large Electron Positron (LEP) Collider experiments set a lower bound for the Higgs boson of 114.4 GeV at the 95% confidence level (CL) [24] and later on, the Tevatron experiments exclude the Higgs boson in two regions: $100 < m_H < 106\ GeV/c^2$, and $147 < m_H < 179\ GeV/c^2$ at the 95% confidence level (CL) [25, 26]. Recently some preliminary results from LHC excluded the Standard Model Higgs boson in the mass ranges: 112.9 - 115.5 GeV, 131-

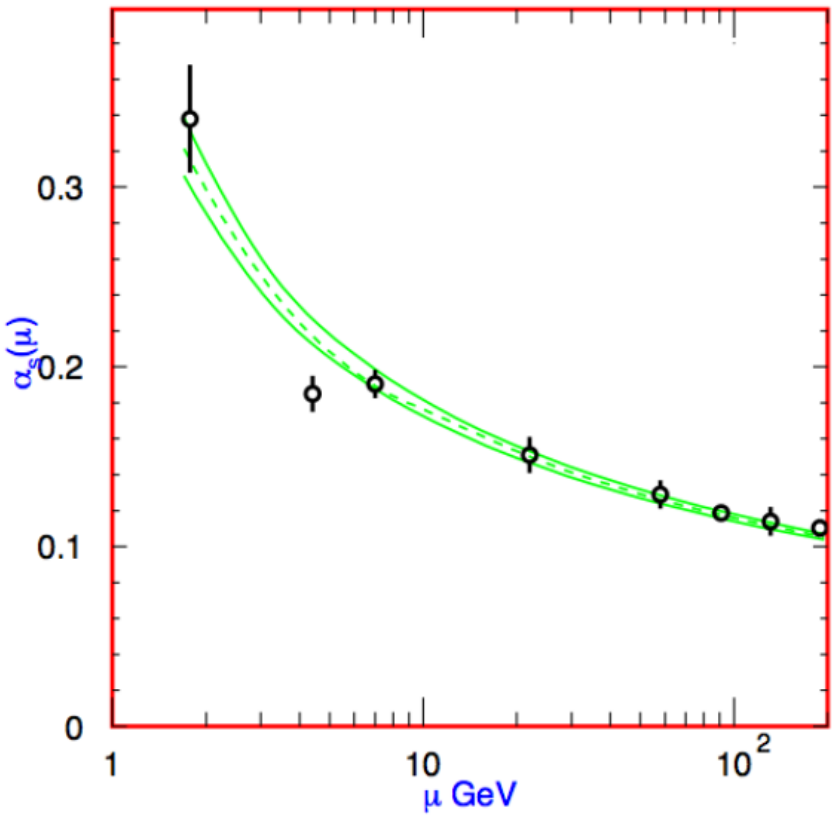


Figure 1.9: Graph showing the running of the strong coupling constant with renormalisation scale $Q = \mu$. The points show the experimental determination of α_s at a particular Q^2 scale, and the curve is a NNLO evaluation of the running of α_s using the average Λ_{QCD} determination of these results. Taken from [6]

238 GeV and 251-466 GeV [27] by the ATLAS experiment with 95% confidence level (CL) and 127-600 GeV by the CMS experiment at the 95% confidence level (CL) and in the mass range 129-525 GeV at 99% CL [28].

Figure 1.11 shows the excluded mass regions for the Higgs boson as of Feb. 2012. Failure to discover the Higgs boson within this mass range would imply the existence of physics outside the Standard Model framework. A key goal of the LHC experiment is to further constrain the predicted Higgs mass with increased precision measurements of the other SM particles. Most recently, immense interest has been re-ignited in Higgs related searches with announcement of an observation of an Higgs like boson at the CMS and ATLAS detectors of LHC. Both the collaborations reported 5 σ signal at around 125 GeV, but further studies are required before we could definitely conclude whether or

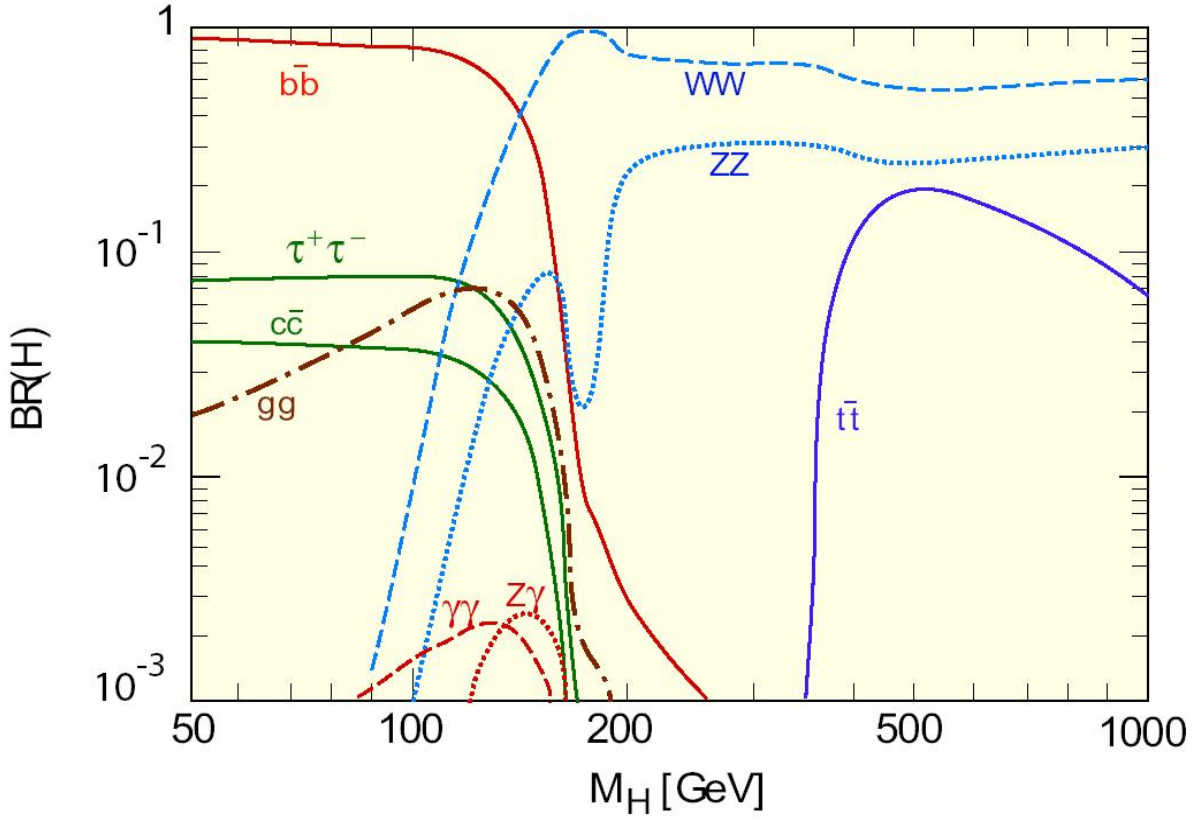


Figure 1.10: Standard Model Higgs branching ratios as a function of Higgs boson mass.

not this is actually the Higgs boson [29], [30]. Adding to this Higgs-Like discovery, the Fermilab experiments also claim for the Discovery of the Higgs boson would undoubtedly be another huge triumph for the Standard Model, but it would not complete the story. Ever since the existence of the Higgs boson has been postulated it has been known that loops of virtual SM particles will introduce quantum corrections to the Higgs mass that could be as large as the Planck scale $M_{Planck} \approx 10^{19} GeV$. How then do such contributions cancel so completely such that the Higgs mass is of the order 100 GeV? This is known as the hierarchy problem. Solutions to this problem require the introduction of new physics beyond the SM framework. The most popular extension to the SM is the theory of supersymmetry [31], in which each SM particle has a supersymmetric partner with opposite spin statistics allowing precise cancellation. Others include the recently proposed "Little Higgs" models [32], which introduce new particles which achieve the cancellation with

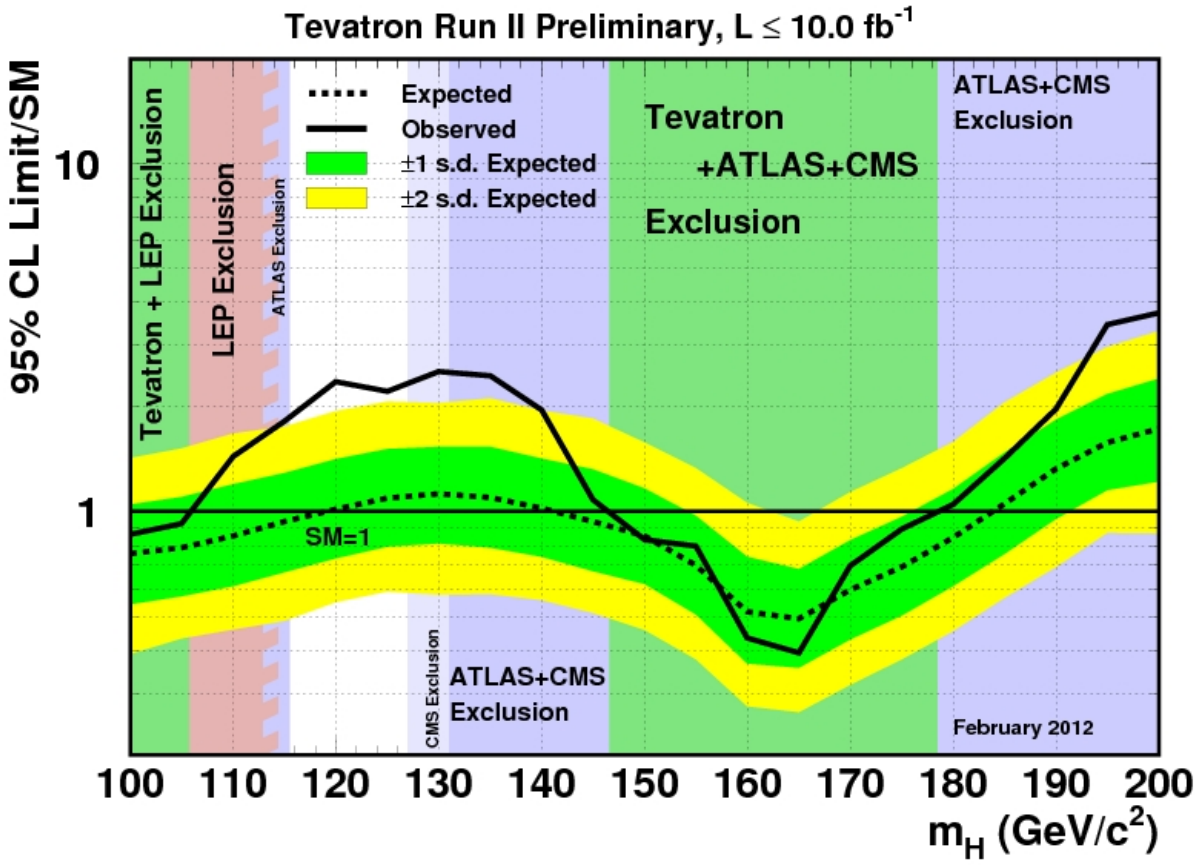


Figure 1.11: Excluded mass regions for the Higgs boson as of Feb. 2012.

same sign spin statistics contributions, and theories of extra dimensions [33, 34]. Closely linked to the hierarchy problem is how the fundamental force of gravity fits into the SM. In particular why is the Planck scale so much larger than the electroweak scale, or put another way, why is gravity so weak? In fact the SM does not predict the coupling strengths of any of the fundamental forces, just as it does not predict the masses of the fundamental particles, nor the existence of three generations of fermions. These are all parameters of the model that have to be established through experimental observation. Is there some theory underlying the SM within which these parameters can be predicted? There are also recent experimental observations which challenge the SM picture of the Universe. A huge wealth of evidence, from the rotation of spiral galaxies to the cosmic microwave background measurements of the WMAP satellite, point to a Universe dominated by cold

dark matter, that is, non-luminous, non-baryonic matter. Additionally observation of neutrino flavour change [35, 36] imply crucially that neutrinos must have mass, in direct contradiction to the massless SM prediction.

1.2 Probing SM with Z+Jets

The production of a vector boson accompanied by hadronic jets is one of the most important Standard Model processes in high energy hadron-hadron collisions.

1.2.1 Z boson production and decays

At proton-proton collisions, the main Z production channel is the Drell-Yan [37] process during which a quark and a antiquark annihilate to create a pair of leptons through the exchange of a virtual photon or Z boson: $q\bar{q} \rightarrow Z/\gamma* \rightarrow l^+l^- + X$. A tree-level and a higher order Feynman diagram for the Drell-Yan process are shown in Figure 1.12. Subleading processes contributing to Z production are quark-gluon scattering $qg \rightarrow qZ$

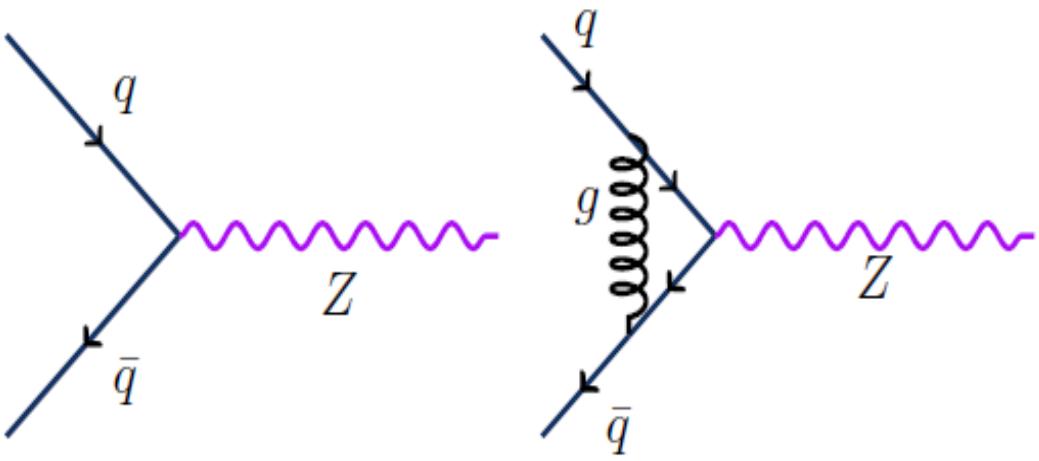


Figure 1.12: The Drell-Yan process

and gluon-gluon fusion $gg \rightarrow q\bar{q}Z$. The Z boson prefers to decay hadronically, with almost 70% of the decays being to strong-interacting particles. "Invisible" decays to neutrinos

Table 1.3: Summary of Z decay modes

Mode	Fraction
e^+e^-	3.363± 0.004 %
$\mu^+\mu^-$	(3.366 ± 0.004) %
$\tau^+\tau^-$	(3.370 ± 0.008) %
invisible	20.00 ± 0.06 %
hadrons (total)	69.91 ± 0.06 %

account for 20% of the decays and the remaining are leptonic decays to electrons, muons and taus in almost equal amounts. A detailed listing given in Table 1.3 [38].

As a result of the Drell-Yan process, the majority of Z bosons are produced at rest or with very little momentum. The lepton pair are identified by the detector and by measuring their energies and momenta, we can reconstruct the mass of the originating particle:

$$M_{ee} = \sqrt{(E_1 + E_2)^2 + (\vec{p}_1 + \vec{p}_2)^2}. \tag{1.9}$$

1.2.2 Jet production

Along with the Z boson, outgoing partons can be produced; for example, in the Drell-Yan case in Figure 1.12 one (or both) of the quarks participating in the strong interaction can radiate a gluon. Due to color confinement, these outgoing partons cannot exist individually; they go through a process called *hadronization*, as a result of which they turn into combinations of colorless hadrons. The number of hadrons produced by the hadronization of a single initial parton is very large and the tight cone of outgoing particles that is formed, is what we call a jet. A convenient dictionary of the terms described above is summarized in Figure 1.13. Production of hard outgoing partons (jets) in association with a Z boson results in more complex and interesting events. An additional factor of interest is that Z+jets final states are common to many rare signals like top decay, or new physics such as associated production of the Higgs boson [39, 40].

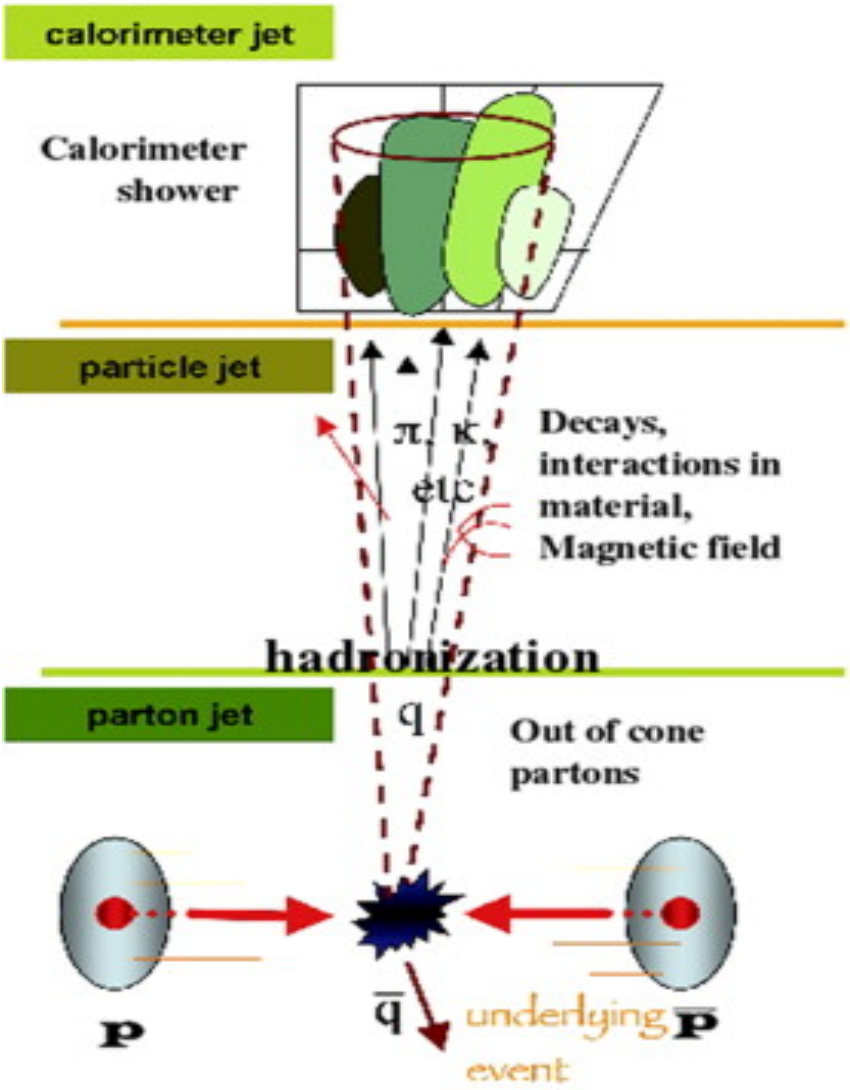


Figure 1.13: A representation of the stages of jet production.

1.2.3 Z+Jets pQCD predictions

The presence of a vector boson in the final state, provides an interaction scale hard enough to allow a perturbative approach to QCD. Therefore $Z/\gamma * + Jets$ production provides adequate conditions to test pQCD. Its cross section can be expressed as the sum of terms with different jet multiplicity:

$$\sigma_{Z/\gamma*+jets} = \sigma_{Z/\gamma*+1jet} + \sigma_{Z/\gamma*+2jets} + \sigma_{Z/\gamma*+3jets} + ... \tag{1.10}$$

and each jet multiplicity can be expanded perturbatively in orders of α_s as:

$$\sigma_{Z/\gamma*+1jet} = a_1\alpha_s + a_2\alpha_s^2 + a_3\alpha_s^3 + a_4\alpha_S^4 + ...$$
$$\sigma_{Z/\gamma*+2jets} = b_2\alpha_s^2 + b_3\alpha_s^3 + b_4\alpha_S^4 + ...$$
$$\sigma_{Z/\gamma*+3jets} = c_3\alpha_s^3 + c_4\alpha_S^4 + ...$$
$$..... \tag{1.11}$$

where a_i, b_i, c_i coefficients in these expansions are in general functions of the jet definition, in particular the cone size used to cluster the partons into jets, and the transverse momentum, rapidity and separation cuts imposed on the jets. To avoid divergences in the calculation, the final state definition needs to be infrared and collinear safe. Therefore jets need to be defined above some energy threshold, and well separated from each other. Leading order contributions (lowest order in α_s) a_1, b_2, c_3, can be calculated from the tree level matrix element. Some of the Feynman diagrams that contribute to these processes at LO for 1 and 2 jets are shown in Figure 1.14 and Figure 1.15, respectively.

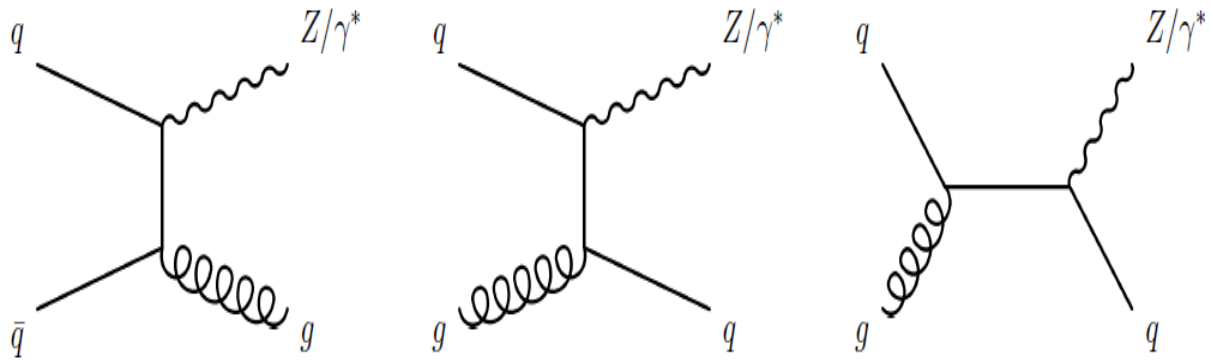


Figure 1.14: Feynman diagrams for $Z/\gamma * +1jet$ production at LO in a hadron collider.

Next-to-leading order (NLO) corrections (a_2, b_3, c_4, etc) are known for up to 2 jets by MCFM [41] and for up to 4 jets by BlackHat [42]. NLO contributions correct the tree-level with virtual diagrams (one-loop corrections, such as those shown in Figure 1.16) and real diagrams (radiation of an extra parton in the final state).

Predictions at the NLO add to the LO n-particle result, the LO n+1- particle

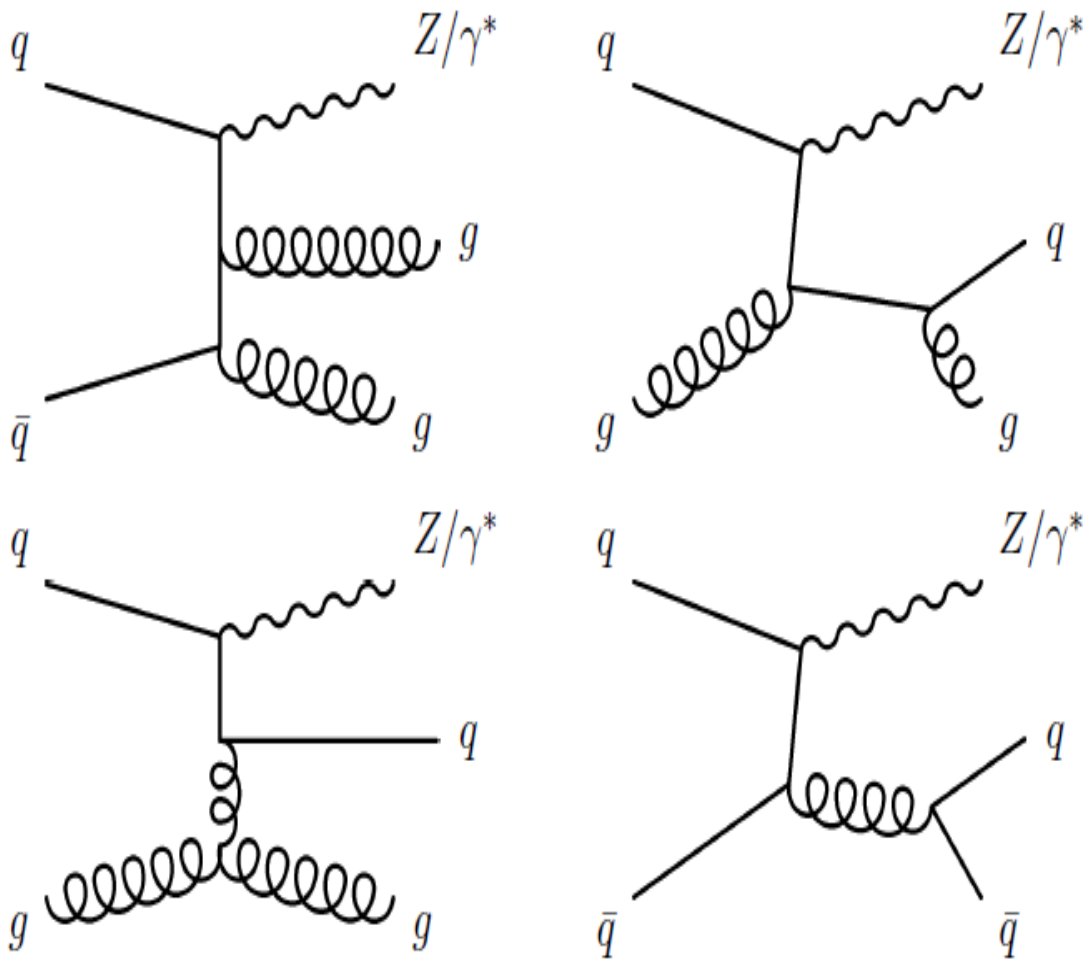


Figure 1.15: Some of the Feynman diagrams for $Z/\gamma * +2jets$ production at LO in a hadron collider.

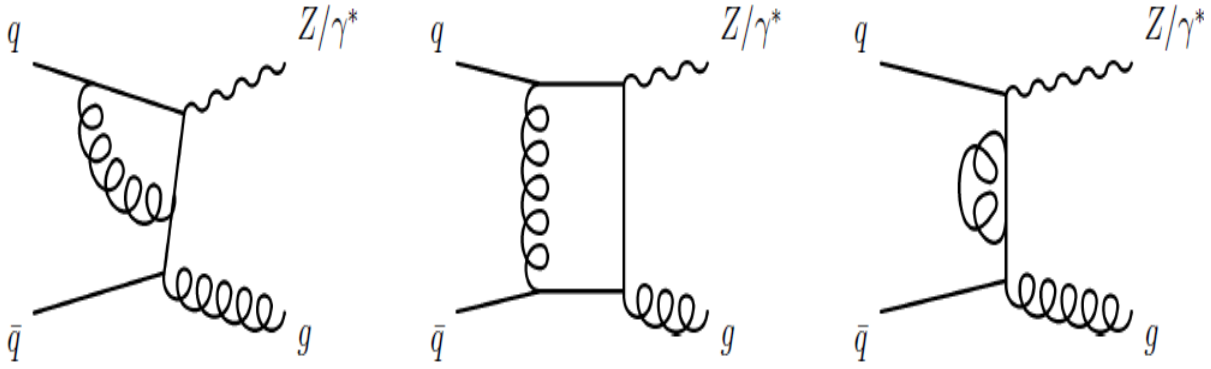


Figure 1.16: Some of the Feynman diagrams contributing to virtual NLO corrections for $Z/\gamma * +1jet$ production in a hadron collider.

contribution and the n-particle 1-loop calculation. Loop amplitudes are infinite, while tree-level amplitudes are finite but with infinite integrals. These two divergences have the same source, and cancel out after integration only if the observable is infrared and collinear safe. Equivalently, the Next-to-next-to-leading order (NNLO) contributions include 2-loop corrections to the LO, one-loop corrections to the NLO, and extra parton radiation. As the number of diagrams increases rapidly with the order of the calculation, NNLO predictions are not available for $Z/\gamma* + jets$. Due to the difficulty in calculating the extra loops involved in higher-order calculations, it is very difficult to calculate cross sections even at next-to-leading order. Current state-of-the-art calculations go up to Z+3 jets at NLO [41, 42].

This thesis deals with a study of Z gauge boson production and tests the reliability of perturbative QCD predictions over a range of jet energies and jet multiplicities. This analysis also provides a contribution to the study of background processes that are relevant for Higgs boson searches [39, 40]. One of the dominant Higgs production modes at the LHC involves the generation of a Higgs boson in association with a Z boson (Figure 1.17).

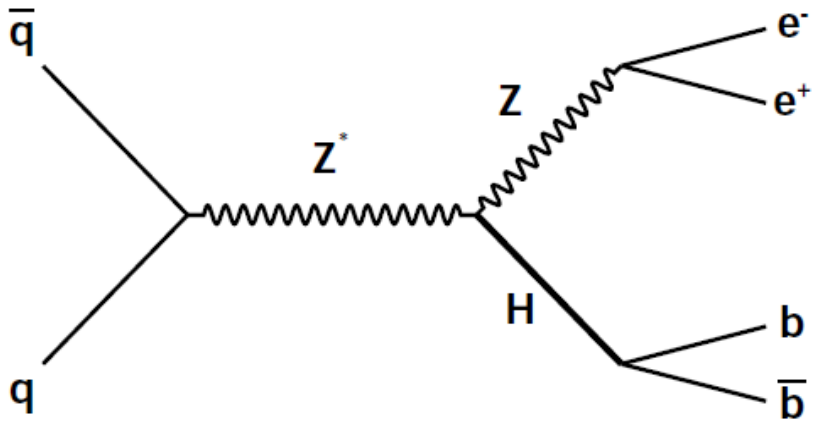


Figure 1.17: SM Higgs boson production in association with a Z boson via $q\bar{q} \rightarrow HZ$ (Higgsstrahlung).

The study of jet properties and cross sections for processes that generate similar final state particles (Figure 1.18) results in a more precise distinction between a possible

Higgs signal and QCD background.

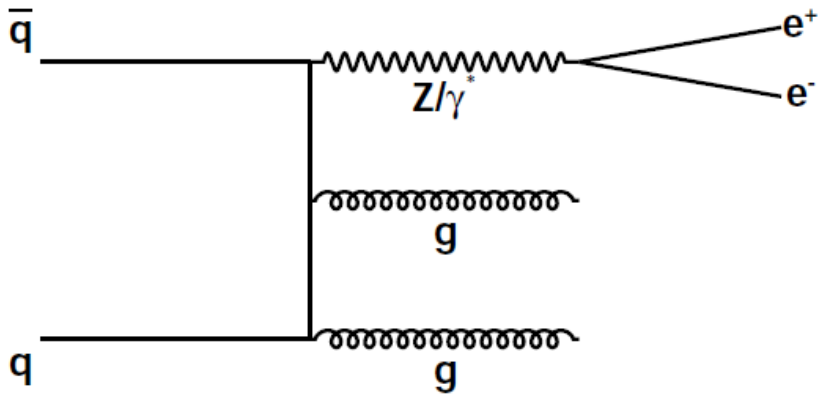


Figure 1.18: Feynman diagram for $Z/\gamma*$ production in association with two gluons via $q\bar{q} \rightarrow Z/\gamma * gg$

The measurement presented in this analysis was performed at a center of mass energy of 7 TeV, with $4.89 fb^{-1}$ of luminosity, and includes jet multiplicities up to four jets. This thesis is organized as follows: The following Chapter 2 gives an overview of the experimental setup, including the LHC's accelerating chains and its commissioning phases. This Chapter gives a brief description of CMS experiment and its sub-detectors. In Chapter 3 a brief description of physics objects reconstruction, using particle flow algorithm from raw detector data, is given. Chapter 4 discusses the Monte Carlo event generators and detector simulation steps. Also the Data and MC samples are described which are used for the analysis. Event selection requirements, including comparisons between data and MC after applying all the selection cuts, are given in Chapter 5. It also illustrates the methods for calculating the event selection efficiencies. Chapter 6 outlines the fit procedure used for extracting signal yields. In order to compare final cross section measurements with theoretical predictions, it is required to correct for the detector effects which has been done using unfolding techniques explained in this Chapter. Chapter 7 describes the main sources for systematic uncertainties as well as the final results. Absolute as well as differential cross-sections as a function of jet p_T and Y are given in this Chapter. A summary of the results is presented in Chapter 8.

This dissertation work is done in collaboration with CMS experiment at CERN (our group at Panjab University being a part of this International Research Collaboration). The results presented here are outcome of the study undertaken for this Ph.D. thesis. CERN [46] stands for "Conseil Europeen pour la Recherche Nucleaire" (in French) which translates to the English "European Council for Nuclear Research", is a high-energy particle physics organization based in Geneva (Switzerland).

Chapter 2

Experimental Setup

This chapter briefly introduces CERN's accelerator complex, describes the Large Hadron Collider (LHC) [47] design (Section 2.1) and commissioning in 2011. Also the Compact Muon Solenoid (CMS) detector layout and its sub-systems are presented briefly in Section 2.2.

2.1 Large Hadron Collider

The Large Hadron Collider at CERN is the world's highest energy particle accelerator and hadron collider. It is installed in the 27 km long underground tunnel that hosted the CERN's Large Electron Positron (LEP) machine [48]. It lies between 45-170 m below the surface of earth.

2.1.1 The accelerator and the experiments

The underground tunnel hosting this LHC ring, has eight straight sections and eight arcs, containing a total of 1624 superconducting magnets (1232 bending dipoles and 392 focusing quadrupoles). The two proton beams circulate in opposite directions with some nominal energy and are allowed to cross at several points along the circumference of the ring and at each such point large particle detectors are hosted. There are four interaction points (IPs) along the tunnel, each housing an experiment: two high luminosity IPs

house the CMS [49] and ATLAS [50] detectors, while two lower luminosity IPs house the LHCb [51] and ALICE [52] detectors. A machine schematic is shown in Figure 2.1. Each

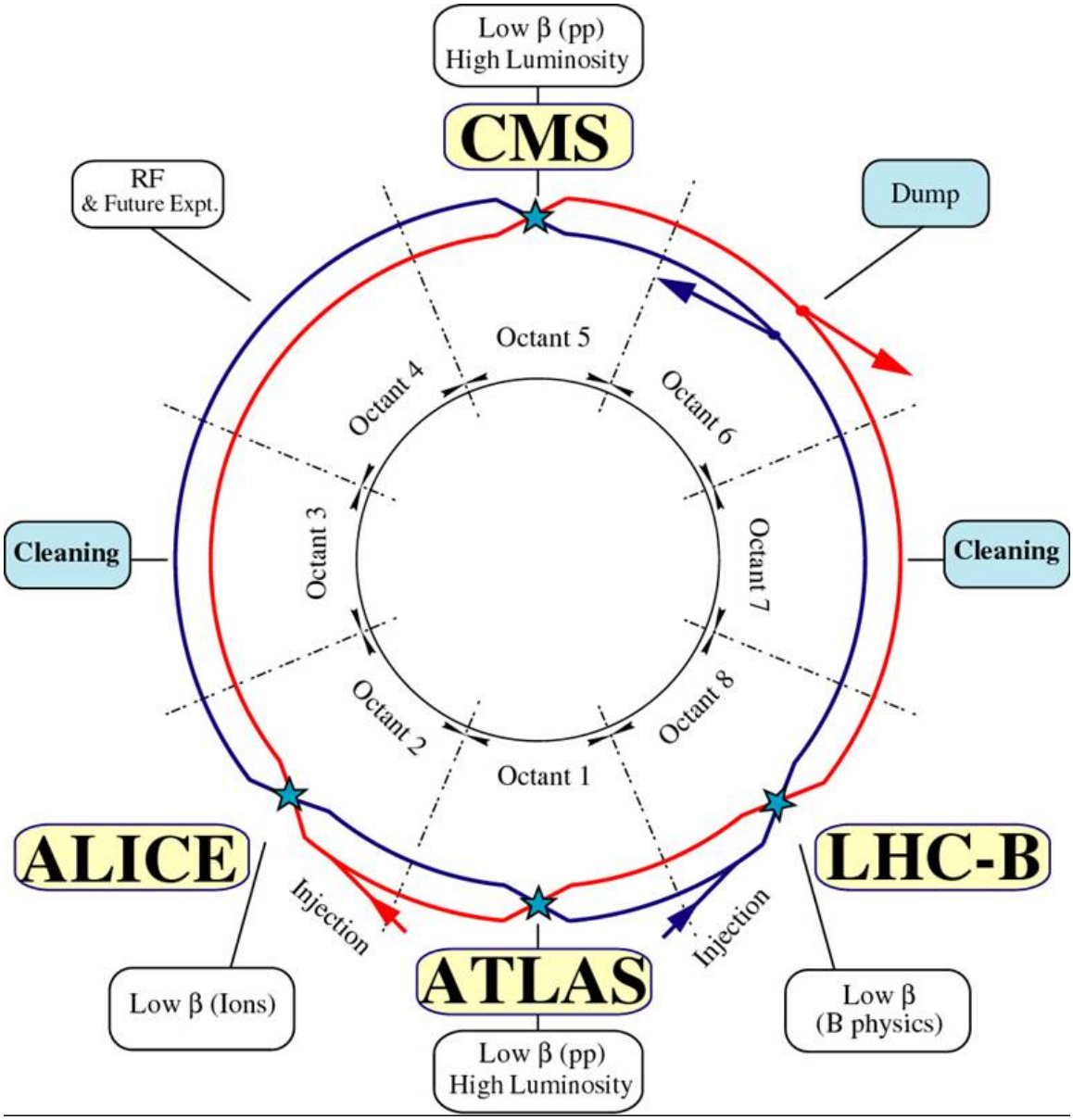


Figure 2.1: Schematic of LHC machine

beam consists of many bunches, whose crossing rate is 40 MHz. The radio frequency cavities accelerate the injected protons by providing boost in each circulation till the collision energy is attained.

The beams are steered around the ring by 8-Tesla magnetic fields produced in 15-

meter-long superconducting niobium-titanium dipole magnets, and focused by quadrupole magnets, 5-7 m long. The LHC uses a design in which both proton beam pipes are contained in the same housing, allowing the same liquid helium cooling system to serve both. Protons are supplied to the ring from an accelerator chain which gradually increases their energy (Figure 2.2). First the protons are accelerated in a linear accelerator (LINAC2) to an energy of 50 MeV and then passed on to Proton Synchrotron Booster (PSB) which increases their energy to 1.4 GeV. After this the beam is then fed to the Proton Synchrotron (PS) where it is accelerated to 25 GeV before they reach the Super Proton Synchrotron (SPS). Here they are accelerated to 450 GeV, which is the typical energy of injection into the LHC.

2.1.2 LHC commissioning phase

The first pp collisions at the LHC were planned to be delivered on September 10th 2008 but the subsequent incident of September 19th caused a full additional year of delay in the schedule of the machine [53]. So, it was decided to start working at the LHC experiments with the exploitation of data coming first from low-luminosity, low-energy collisions, and then gradually higher energy and beam intensity. Hence real first pp collisions at the LHC were delivered on November, 2009 at $\sqrt{s}$ = 900 GeV and then later on centre of mass energy increased to 2.136 TeV which just exceeded that of the Tevatron (1.96 TeV) [54, 55] and this makes the LHC to be the world's highest energy hadron collider. The very early, $\sqrt{s}$ = 900 GeV, data delivered by the LHC in 2009 was insufficient to produce significant advancements in our knowledge of fundamental physics, but it allowed important calibrations of the detectors and data acquisition chain through the re-discovery of known standard model (SM) signals. In year 2010, the first extended period of data taking, beginning in March and ending in November, started. During this time the machine operated at $\sqrt{s}$ = 7 TeV and delivered approximately 47 pb^{-1} of integrated luminosity to both the ATLAS and CMS experiments, out of which 36 pb^{-1}

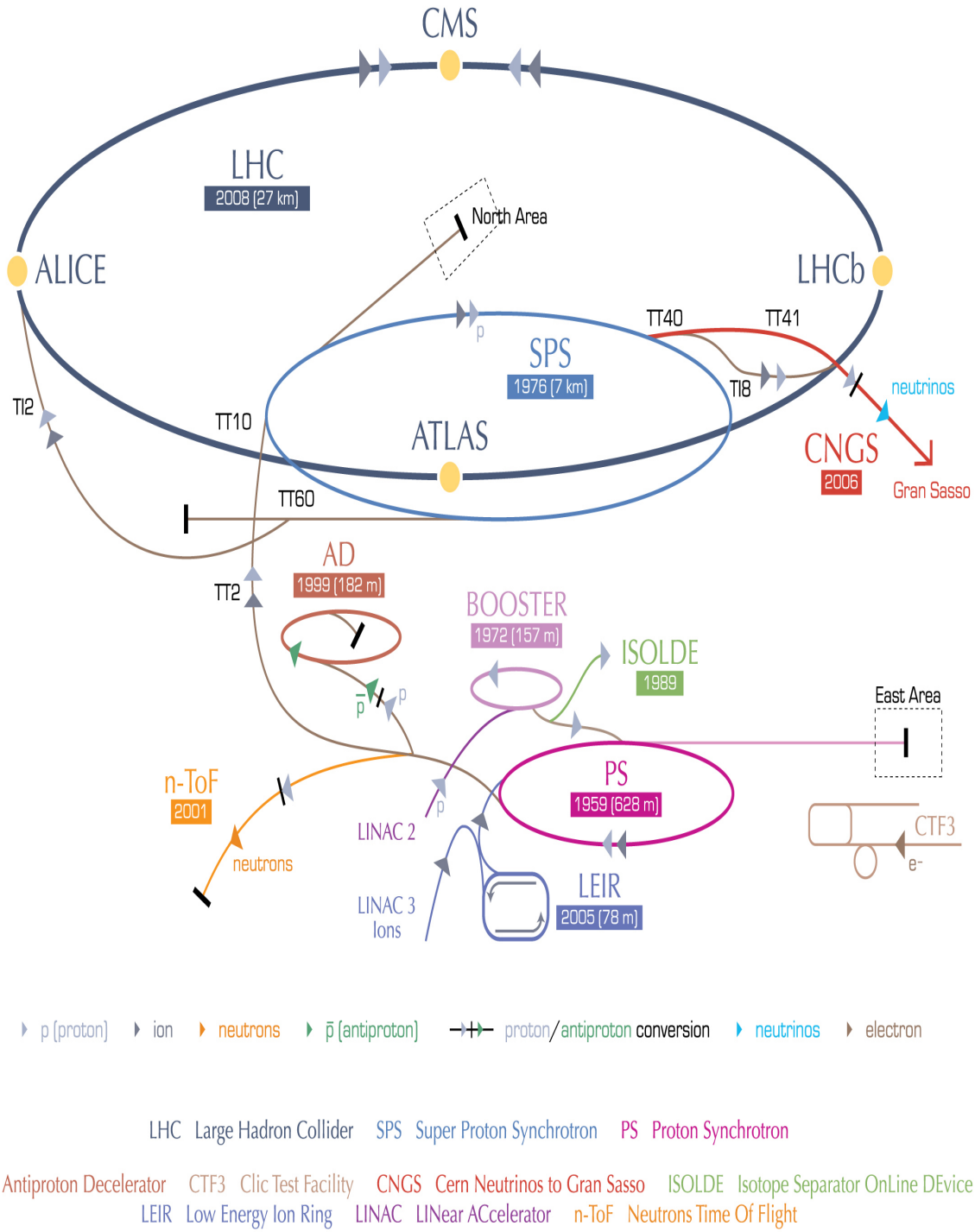


Figure 2.2: The LHC is the last ring in a complex chain of particle accelerators. It is supplied with protons from the injector chain Linac2 → Booster → PS → SPS.

was commissioned for actual analysis in CMS (Figure 2.3). This data was successfully used by the CMS collaboration to tighten most of the existing bounds on supersymmetric

(SUSY) signatures [56, 57, 58, 59, 60], as well as to discover new features of low-energy quantum chromodynamics (QCD) production [61], and to add the 7-TeV measurement point to several cross section graphs [62, 63, 64, 65]. However, it was the year 2011 which was not only crucial for the Large Hadron Collider (LHC) experiments in particular but brought a qualitatively different amount of information on many new physics models.

The proton-proton run of 2011 at LHC started on March 14th and terminated on October 30th. During these seven months of data-taking the LHC delivered $\sim$6fb^{-1} of collision data. The CMS experiment recorded $\sim$5.5 fb^{-1} of these, and after quality checks approximately 5 fb^{-1} of collision data were approved for analyses, as shown in Figure 2.3. The analysis and results discussed in this thesis have been produced by analyzing the corresponding datasets of 2011 collisions.

The design energy of LHC is 14 TeV (7 TeV per beam), and its design luminosity is 10^{34} cm^{-2}s^{-1}, whereas during the 2011 run the instantaneous luminosity of proton-proton collisions delivered by the LHC reached up to 3.5 $\times 10^{33} cm^{-2} s^{-1}$.

2.2 Compact Muon Solenoid

The Compact Muon Solenoid (CMS) experiment is a large general-purpose particle physics detector on the LHC ring located in town of Cessy, inside France. Its primary aim is to elucidate the nature of electroweak symmetry breaking and search for new phenomena at the TeV scale. Figure 2.4 shows the CMS design. Going outwards from the the interaction point, CMS consists of a series of sub-detecting layers namely pixel detector, tracker, electromagnetic calorimeter, hadron calorimeter and finally the muons chambers [49].

2.2.1 Coordinate Conventions

The convention regarding the axes has been fixed using geographic landmarks and is as follows:

- Origin is located in the center of the detector at the collision point.

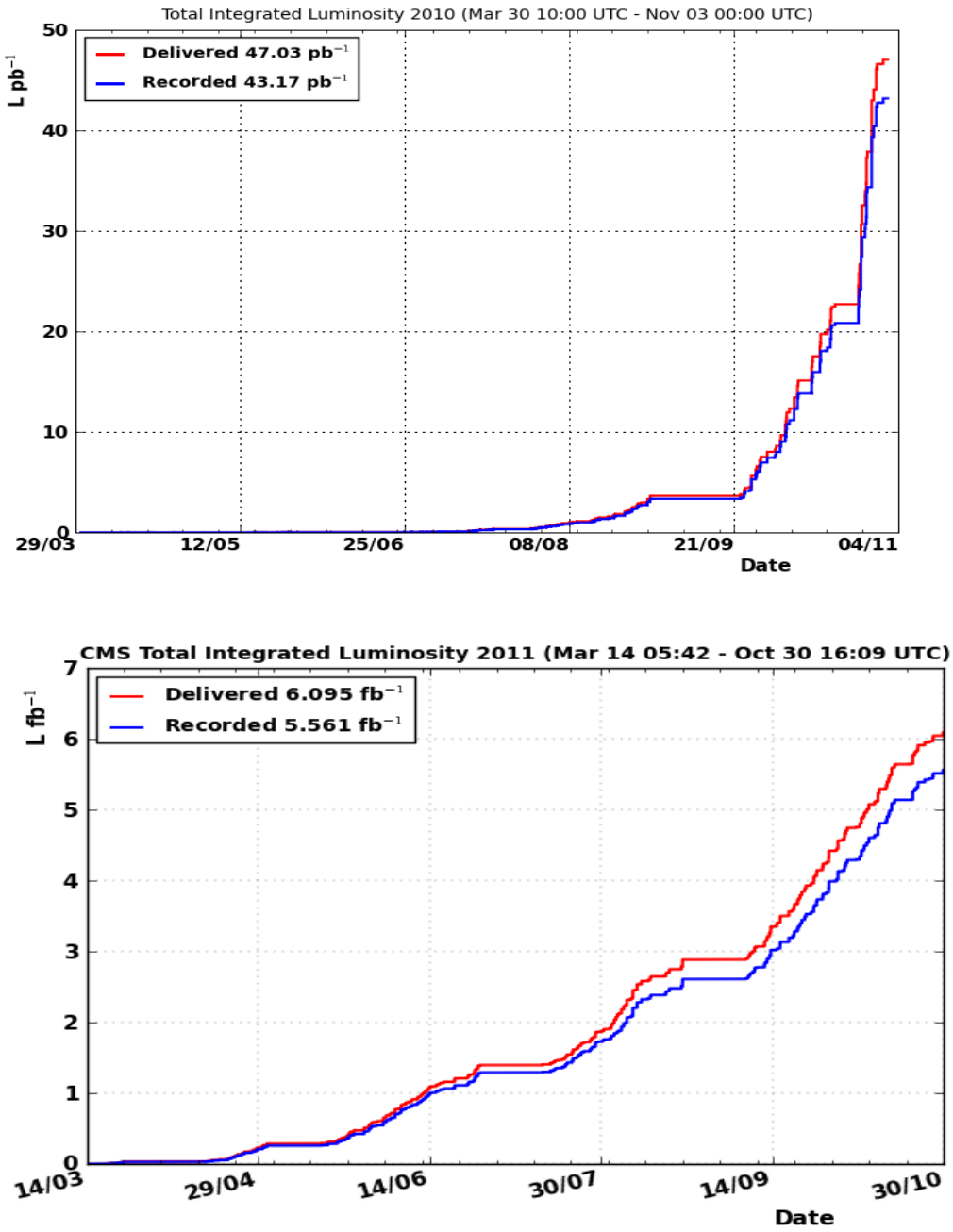


Figure 2.3: Top: Integrated luminosity delivered by the LHC (red) and recorded by the CMS experiment (blue) during the 2010 7-TeV proton-proton run, as a function of date of the year. Bottom: Integrated luminosity delivered by the LHC (red) and recorded by the CMS experiment (blue) during the 2011 7-TeV proton-proton run, as a function of date of the year.

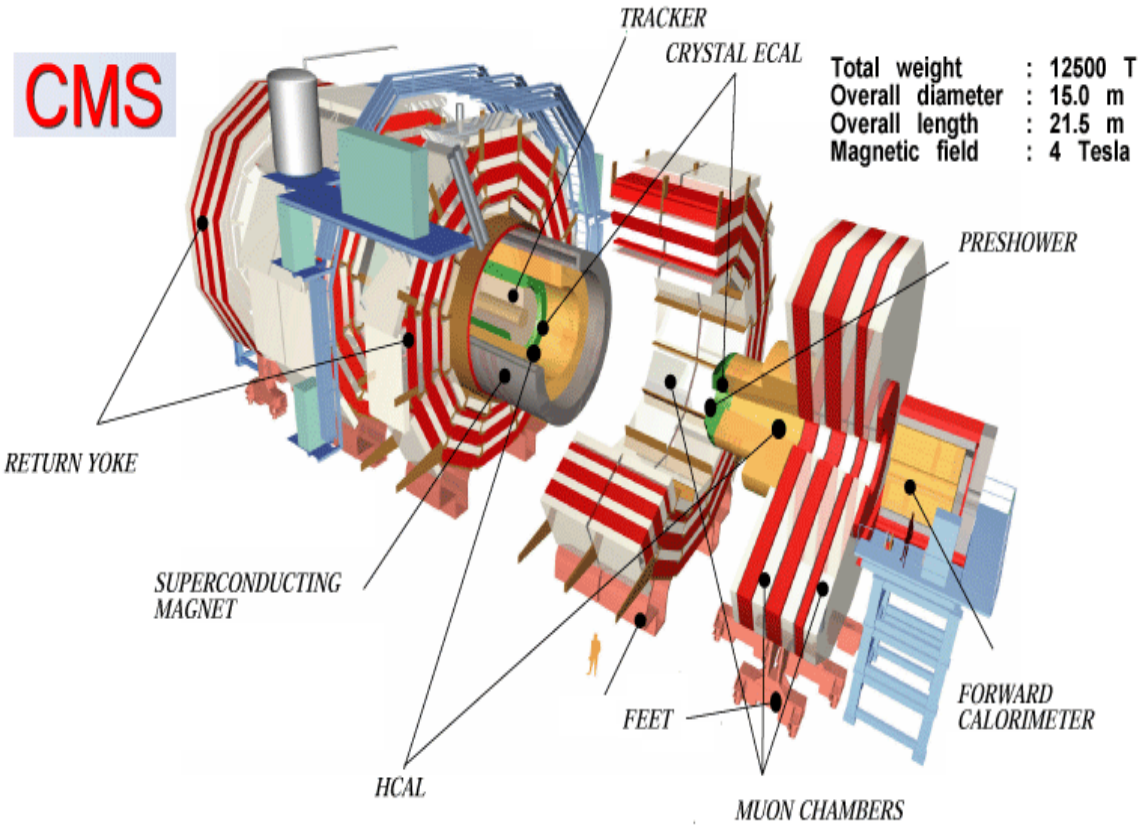


Figure 2.4: Illustration of the CMS detector.

- The y-axis pointing vertically upward, and the x-axis pointing radially inward toward the center of the LHC.

- The z-axis points along the beam direction.

The polar angle θ is then defined with respect to the positive z-axis, and azimuthal angle ϕ is defined in the xy-plane so that $\phi = 0$ is along the positive x-axis and $\phi = \pi/2$ is along the positive y-axis (Figure 2.5). More commonly the positions inside the detector are generally defined in terms of the *pseudorapidity* η, which is defined as:

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \tag{2.1}$$

The pseudorapidity is particularly useful in particle physics because particle production is roughly constant as a function of η. The pseudorapidity approximates the true rapidity,

$$Y = \frac{1}{2}\ln\left(\frac{E + p_z}{E + p_z}\right) \tag{2.2}$$

in the limit of m << E (where m is the invariant mass $m^2 = E^2 - p^2$). Thus, the momentum and energy measured transverse to the beam direction, denoted by p_T and E_T, respectively, are computed from the x and y components. The imbalance of energy measured in the transverse plane is denoted by E_T^{miss}.

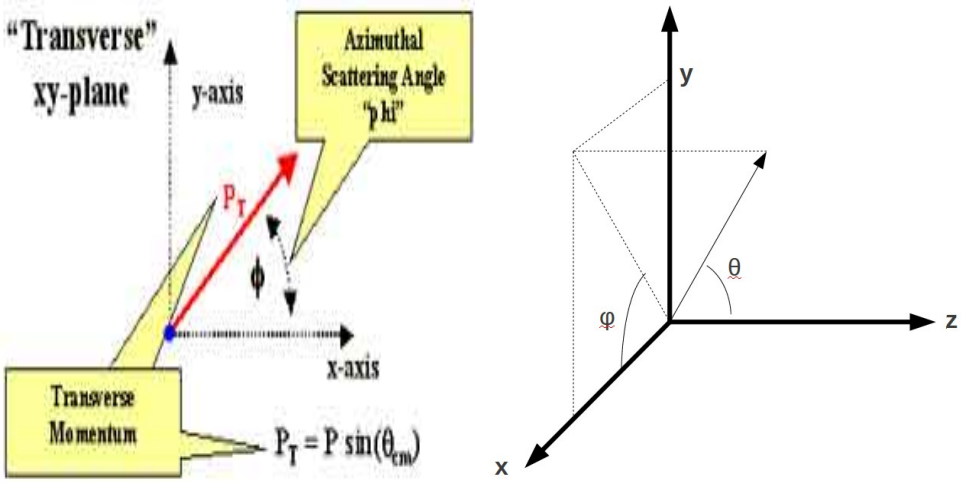


Figure 2.5: Coordinate conventions

The different subdetectors of CMS are briefly described in following sections.

2.2.2 Inner tracking system

The tracking system shown in Figure 2.6 is the innermost part of CMS, and lies closest to the interaction point.

It has been designed to enable precise measurement of the trajectories of charged particles as well as reconstruction of secondary vertices inside detector. The first three layers of the tracker are called silicon pixel detectors, which provide the highest precision and granularity of tracker channels. The pixel detectors extend upto a radius of 11 cm and

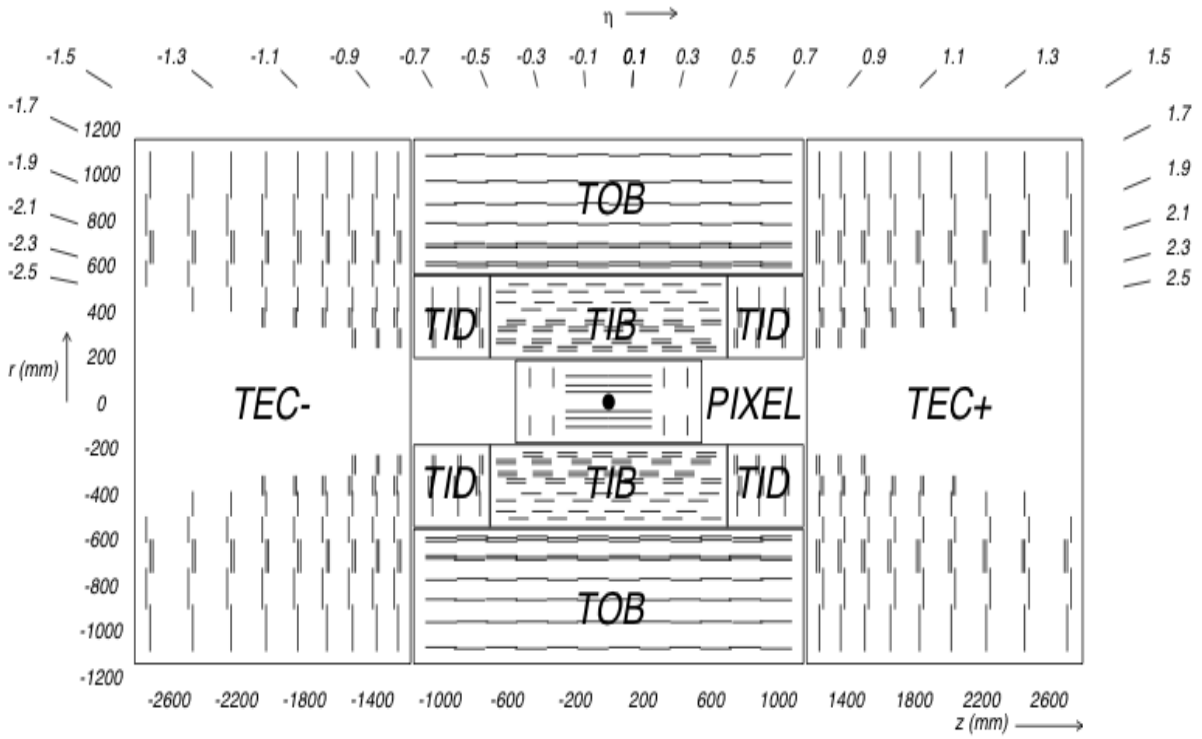


Figure 2.6: Schematic cross section through the tracker. Each line represents a detector module.

comprise of 66 million pixels in total, each with an area of 100×150 μm. The resolution is ~ 10 μm in the r-ϕ plane and ~ 20 μm in the r-z plane. The pixel detector has been designed to provide very high precision and granularity of tracker channels in a fiducial range of $|\eta| < 2.5$ [66]. The next 10 layers of the detector, going upto a radius of 1.1 m, are made of silicon microstrip detectors. They consist of 11.4 million long strips, 10 or 25 cm in length, but only 180 μm in height. The barrel region can be divided into a tracker inner barrel (TIB) for $r < 65$ cm, with a resolution of 230 μm, and a tracker outer barrel (TOB) for $65 < r < 110$ cm, with resolution 530 μm [66]. The tracker endcaps (TEC) cover the region out to $|\eta| < 2.5$ and contain strips oriented radially to the beam. As the innermost layer of CMS, the tracker is subjected to the strong magnetic field generated by the surrounding solenoid. It is therefore possible to measure the transverse momentum

of any charged particle from the curvature of its path in the tracker, via

$$p_T = qrB \tag{2.3}$$

where p_T is the transverse momentum, q is the charge of the particle, r is the radius of the track made by the particle, and B is the strength of the magnetic field (3.8 T). The large magnetic field and the fine granularity of the tracker elements yield excellent momentum resolution:

$$\frac{\sigma p_T}{p_T} = (15 p_T \oplus 0.5)\,\% p_T\,(TeV) \qquad |\eta| < 1.6 \tag{2.4}$$

$$\frac{\sigma p_T}{p_T} = (60 p_T \oplus 0.5)\,\% p_T\,(TeV) \qquad |\eta| = 2.5 \tag{2.5}$$

More energetic particles curve less in the magnetic field, which makes their momentum more difficult to determine. Thus, the uncertainty increases with increasing p_T. The tracker is an important component in the identification of both electrons and muons, which are used in this analysis.

2.2.3 Electromagnetic calorimeter

The next detector component, moving radially outward from the interaction point, is the electromagnetic calorimeter (ECAL), designed to accurately measure the energies of electrons and photons (Figure 2.7). The ECAL is a homogeneous calorimeter composed of crystals of lead tungstate ($PbWO_4$), a dense but optically transparent material which is ideal for stopping high energy particles [67, 68]. In total there are 61200 crystals in the central barrel (EB), $|\eta| < 1.4442$, and 7324 in each of the two endcaps (EE), $1.566 < |\eta| < 3.0$. The crystals in the barrel are organized into supermodules consisting of 1700 crystals, 85 in η and 20 in ϕ. There are 36 supermodules in total, 18 in each half barrel.

To understand the performance of the ECAL, it is necessary to briefly discuss the interaction of electromagnetic particles with matter. High-energy electrons predominantly loose energy via bremsstrahlung, or "breaking radiation", in which electrons in the field

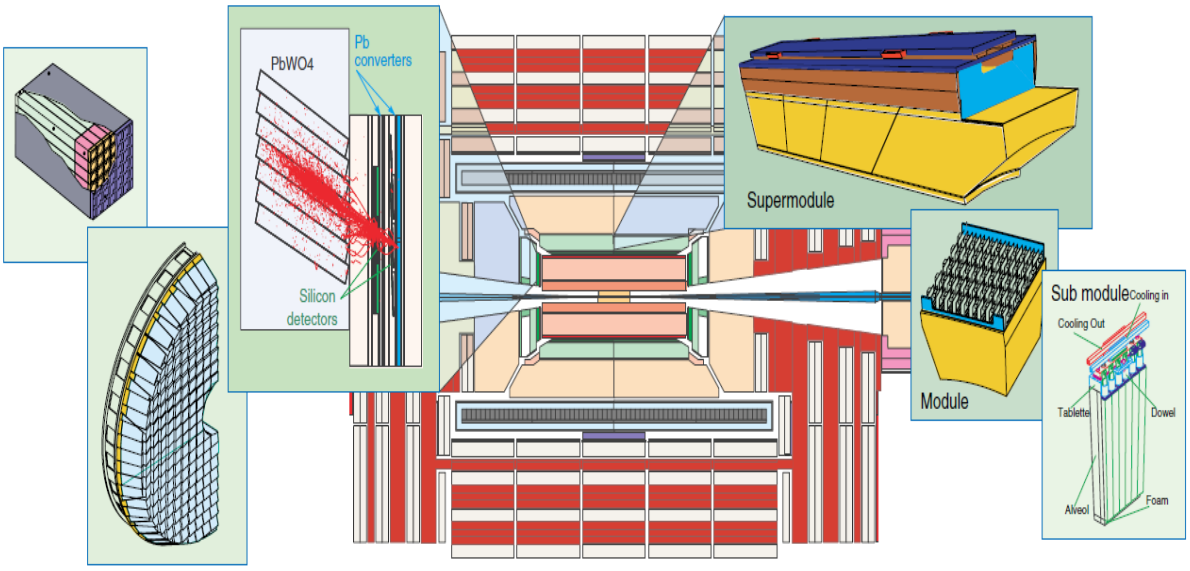


Figure 2.7: Components of the ECAL showing the organization of the calorimeter crystals.

of an atomic nucleus undergo large acceleration, causing the emission of a photon. As the electron passes through the material, it loses energy via bremsstrahlung as

$$\frac{dE}{dx} = -\frac{E}{X_0} \tag{2.6}$$

where X_0 is the characteristic radiation length of the material that the electron is moving through. In the case of the lead tungstate used in the ECAL, $X_0 = 0.89$ cm; the length of each crystal is 23 cm, or $25.8X_0$, ensuring that the incident electrons will undergo bremsstrahlung many times.

Photons tend to interact with matter primarily via pair-production, wherein a photon converts to a e^+e^- lepton pair as it traverses the medium. The combination of this process with bremsstrahlung leads to the creation of an electromagnetic shower. An incident electron will radiate a photon, which then pair-produces two additional electrons. These electrons will subsequently radiate photons, and so forth, until all of the energy of the original electron is contained in the resulting cascade of photons, which can then be measured by the scintillation counter at the base of the crystal. The same process occurs with an incident photon, but the chain begins with a pair-production.

Electromagnetic showers are known empirically to be well collimated in the transverse direction. Showers are generally characterized in the lateral direction by the Moliere radius R_M, defined as the radius of a cylinder containing 90% of the shower energy. In the case of $PbWO_4$, $R_M = 2.2$ cm, the same as the front size of the crystals, ensuring that the ECAL has fine granularity.

The energy resolution of the ECAL [69, 70, 71], given by

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{5.4\%}{\sqrt{E}}\right)^2 + \left(\frac{124MeV}{E}\right)^2 + (0.4\%)^2 \tag{2.7}$$

is parametrized into multiple independent terms. The first term is stochastic, including statistical fluctuations and intrinsic shower-size fluctuations. The second term accounts for electronics noise and pile-up energy, and the third term is a constant covering detector non-uniformity and calibration uncertainty. The overall energy resolution is very good, and is optimal for energies between $1GeV$ and $1TeV$. Additional corrections are required over time, because exposure to radiation causes the crystals to lose transparency, so they must be re-calibrated using a laser system.

2.2.4 Hadron calorimeter

Beyond the ECAL is the hadron calorimeter (HCAL), whose purpose is both to measure the energy of hadron jets and to provide a reliable measure of the missing energy in an event [72, 73]. It is divided up into three components by η, with the HCAL barrel (HB) covering the region $|\eta| < 1.305$, the HCAL endcap (HE) extending from $1.305 < |\eta| < 3.0$, and the HCAL forward (HF) region going from $3.0 < |\eta| < 5.0$. The total amount of material available in the barrel is restricted radially by the outer edge of the ECAL and the inner edge of the magnet [74, 75]. This is filled with the interleaved layers of of dense material, in this case brass, and tiles of plastic scintillator. These layers extend outwards upto 0.96 m, and are divided into 36 azimuthal wedges which compose the two half-barrels of the HB. In the HF steel is used in place of brass, and the scintillator layer is composed of quartz fiber (Figure 2.8) [76, 77, 78].

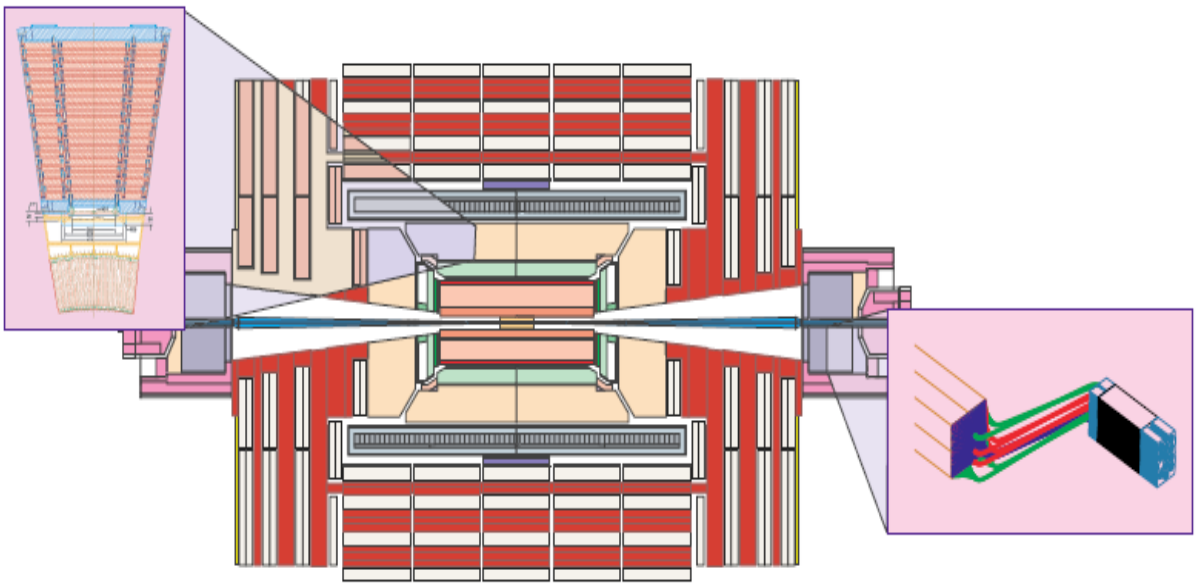


Figure 2.8: Components of the HCAL showing the brass and scintillator of the HB/HE and the steel and quartz of the HF.

The function of the metal layers is to provide a dense material with heavy nuclei to cause the hadron to interact, while the scintillator layers measure the visible energy from those interactions. Hadronic interactions with matter can proceed through a variety of processes, including particle production (dominated by pion production) and π^0 production, with the subsequent decay $\pi^0 \rightarrow \gamma\gamma$. Because of the photon production, hadronic showers will also have an electromagnetic component. In addition, decays to neutral particles will fail to produce energy observed by the scintillator layers. As a result of fluctuations in the number of neutral particles and π^0s produced, there can be large variations in the size and visible energy of hadronic jets, though they are generally much larger than electromagnetic jets, both longitudinally and transversely.

The HCAL extends 6-8 interaction lengths, with the energy of a hadron decreasing by a factor of $1/e$ in a single interaction length. Brass is optimal in the barrel, as it is non-magnetic and possesses a short interaction length, while steel and quartz are used in the endcaps because they are better able to withstand the high rate of radiation from the

beams. The resolution of the HCAL is given by

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{90\%}{\sqrt{E}}\right)^2 + (4.5\%)^2 \qquad |\eta| < 3.0 \tag{2.8}$$

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{172\%}{\sqrt{E}}\right)^2 + (9.0\%)^2 \qquad 3.0 < |\eta| < 5.0 \tag{2.9}$$

The first term is stochastic, including statistical fluctuations and intrinsic shower-size fluctuations. The second term is a constant accounting for detector non-uniformity and calibration uncertainty. Comparison of (2.8) to (2.7) shows that HCAL resolution is worse than ECAL resolution, both because the HCAL is a sampling calorimeter and because hadronic showers tend to have larger statistical fluctuations.

2.2.5 Superconducting magnet

The main distinguishing feature of CMS is a very large superconducting solenoid magnet, 13 m long and 6 m in diameter, which is capable of providing a magnetic field of 4 T in the interior of the solenoid (Figure 2.9). For the collection of data in 2011, the actual operating field of the magnet was 3.8 T. This field will bend the paths of charged particles in the inner layers, allowing their momenta to be determined, and so a high magnetic field is essential for both good momentum resolution and unambiguous determination of particle charge.

Because the magnet itself is so large, it is possible to accommodate not only a tracking system, but also all of the calorimetry, entirely in its interior. This allows better electromagnetic energy resolution as well as missing-transverse-energy and jet energy resolution, as these can be measured before the particles pass through the high-density magnet. The magnetic flux through the interior of the solenoid is returned through a supporting iron yoke 1.5 m thick, weighing 10,000 t and comprising 5 “wheels” and 2 “endcaps” with 3 disks each.

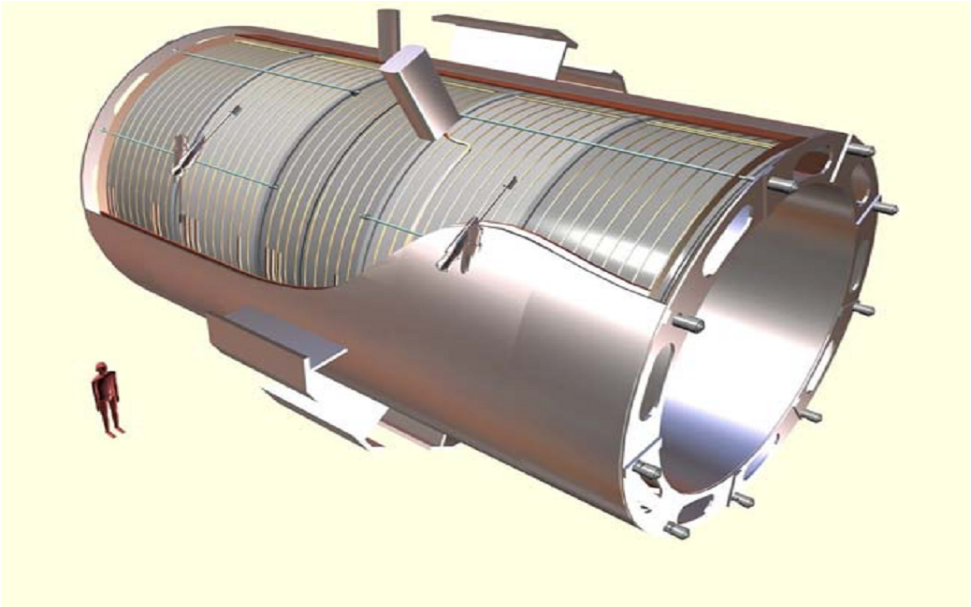


Figure 2.9: A perspective view of the solenoid.

2.2.6 Muon system

As is implied by the experiment's middle name, the detection of muons is of central importance to CMS. The muon system is designed to reconstruct the momentum and charge of muons over the entire kinematic range of the LHC. Due to their particular mass, muons have minimal interactions with matter and are consequently very penetrating. Because of this, the muon system can be placed beyond the solenoid in the iron return yoke. Three types of gaseous particle detectors are used for muon identification [79]

There are three separate detector systems for muons (Figure 2.10): drift tubes (DT) are employed in the central region (out to $|\eta| < 1.2$) for precise trajectory measurements and cathode strip chambers (CSC) are used in the endcaps ($0.9 < |\eta| < 2.4$); resistive plate chambers (RPC), which provide precise timing of muons passing through the detector, cover the barrel and much of the endcaps, out to an $|\eta|$ of 1.6.

The DTs are a type of wire chamber, in which an electric field is set up around a thin wire that is immersed in a gas (in this case, a mixture of argon and carbon dioxide). The passage of the charged muon ionizes this gas, and the electric field causes the freed

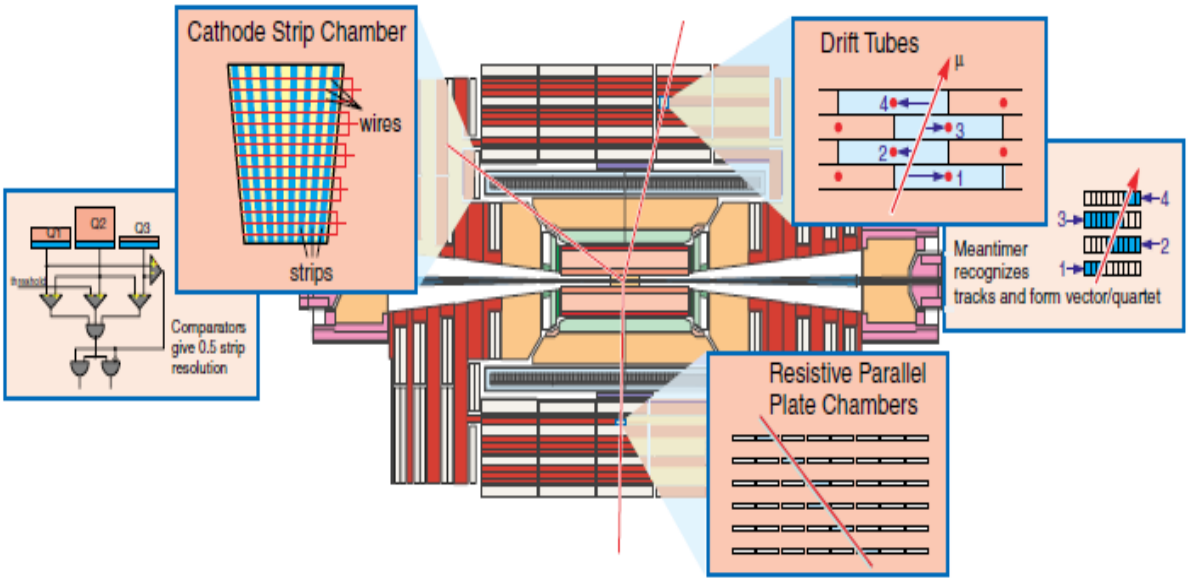


Figure 2.10: Muon system configuration showing the three separate muon detectors

atomic electrons to drift toward the sensing wire. These electrons, accelerated by the field, produce more ions, and the resulting avalanche produces a current in the wire, such that the signal can be recorded. The DTs are organized into 4 stations moving radially outward, and the drift cells of each chamber are offset by half of the cell width with respect to their neighbors. This configuration makes it possible to link the signals in each chamber into a single muon track via the mean timing of the hits on each wire. It also eliminates dead spots in the efficiency and improves rejection of background hits.

The CSCs are multi-wire proportional chambers comprised of six anode wire planes interleaved among seven cathode panels. Wires run azimuthally and define a track's radial coordinate. Following the original CSC concept [80], the muon coordinate along the wires is obtained by interpolating charges induced on strips. The muon production rates and neutron-induced backgrounds are higher closer to the beam line, so CSCs, which have a faster response time, fine segmentation, and good resistance to radiation, are preferred for the endcaps. They are composed of cathode strips that run radially outward and anode wires that run perpendicular to the strips. The anode wires detect the ionized electrons

from the passage of the muon, and the positive ions induce a pulse in the cathode strips. Thus, the strips provide a precision measurement of the path of the muon in the r-ϕ plane, while the wires measure the η of the muon.

Resistive Plate Chambers (RPCs) are gaseous parallel-plate detectors that combine adequate spatial resolution with a time resolution comparable to that of scintillators [81, 82]. The RPCs comprise a third, complementary muon detection system, designed to make fast and independent muon measurements over a wide η range. The parallel resistive plates create a constant electric field in the gas filled interior, such that when an incident muon ionizes the gas, the electrons are pulled to the plates. The electrodes themselves are transparent to the signal, which is instead picked up by a series of conductive detecting strips on the other side. RPCs have a very fast response (1 ns), but coarser spatial resolution than the DTs or CSCs, making them optimal for triggering on muons.

2.3 Trigger

For pp collisions, the LHC has a nominal beam crossing interval of 25 ns at peak luminosity, which corresponds to a crossing frequency of 40 MHz. The actual number of collisions for each proton bunch crossing depends on luminosity, but is approximately 20-25 at the machine's nominal design luminosity of $10^{34} cm^{-1} s^{-2}$. Since it is impossible to store and process the large amount of data associated with the resulting high number of events, a drastic rate reduction has to be achieved. It is therefore necessary to perform an initial event selection, both to reduce the overall event rate and to increase the proportion of relevant physics events in the stored data. The rate is reduced in two steps corresponding to the the Level-1 Trigger (L1) [83] and High-Level Trigger (HLT) [84]. The rate reduction capability is designed to be at least a factor of 106 for the combined L1 and HLT.

2.3.1 Level-1 trigger

The L1 consists of custom designed, largely programmable electronics and has a maximum output rate of 100 kHz. The L1 trigger uses coarsely segmented data from the calorimeters and the muon system, while holding the high-resolution data in pipe-lined memories in the front-end electronics. It is divided into local, regional, and global components, as shown in Figure 2.11.

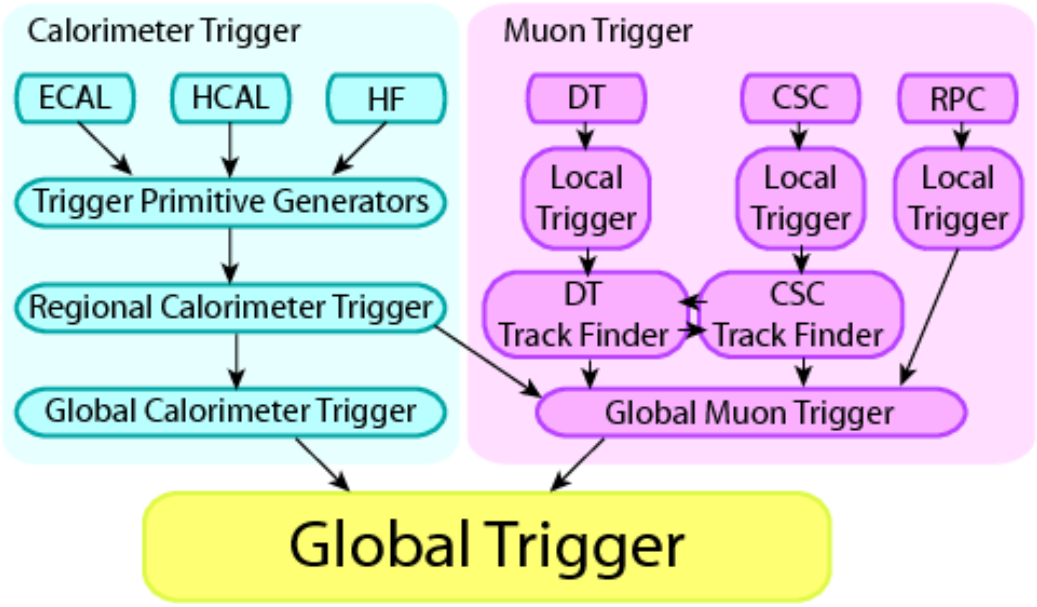


Figure 2.11: Architecture of the L1 trigger.

Local trigger information is produced by so-called Trigger Primitive Generators (TPG), which come from both the calorimeter and muon triggers; in the former case they are based on energy deposits in the calorimeter towers, and in the latter case they use track segments or hit patterns in the muon chambers. The information from the TPGs is passed to the regional triggers, where it is used to determine a sorted list of trigger objects like electron or muon candidates in specific regions, and to rank them by energy. These candidates are then passed to the global calorimeter and global muon triggers, which determine the overall highest-rank calorimeter and muon objects to send to the

global trigger. It is the global trigger which makes the final L1 decision to reject an event or to accept it for further evaluation by the HLT.

2.3.2 High-level trigger

The HLT is a software system implemented in a filter farm of about 1000 commercial processors with a latency of approximately 30 ms. The job of the HLT is to reduce the event rate to a final output rate of approximately 300 Hz, roughly the maximum rate that can be written to a storage disk. While the L1 trigger uses coarse-binned data, the HLT has access to the complete read-out data of the event, and is able to perform complex calculations similar to those performed in an offline analysis. Consequently, it is possible to construct more sophisticated trigger requirements for event selection.

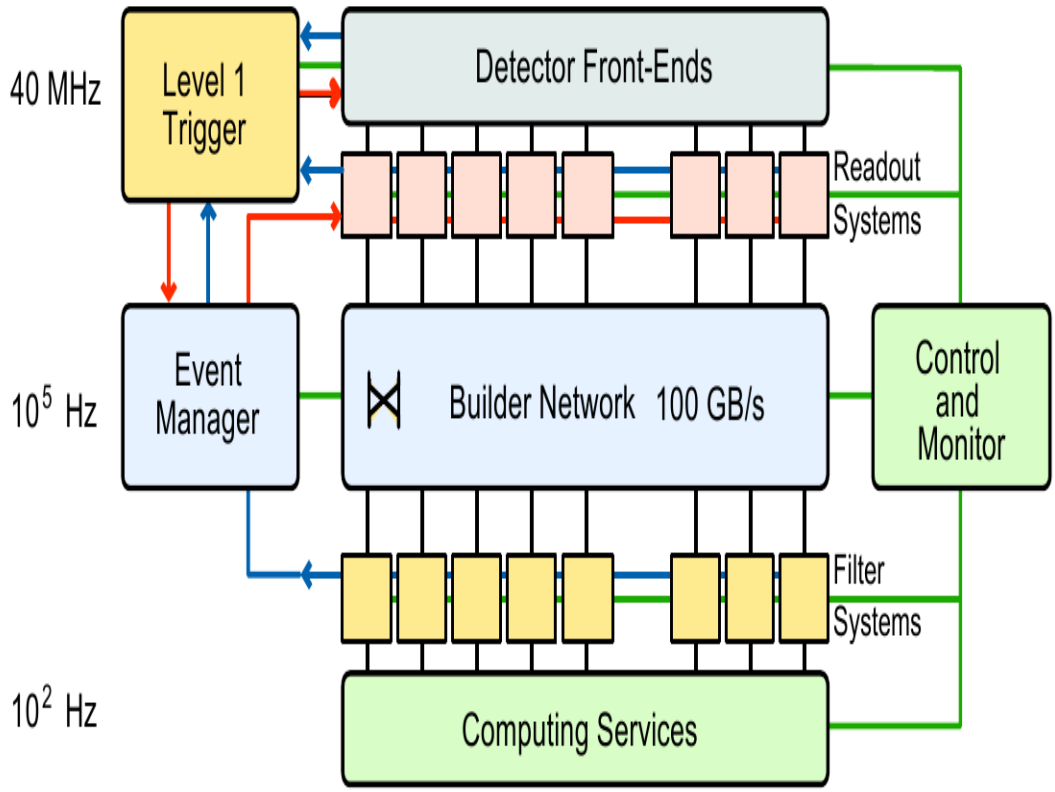


Figure 2.12: Architecture of the CMS Data Acquisition system.

In order to make efficient use of the allocated trigger decision time, the HLT

code is divided in 3 virtual layers: *Level-2* where muon and calorimeter information is used, reconstructing more refined objects starting from the L1-accepted objects, *Level-2.5* additionally uses hit information from the pixel detector and *Level-3* makes use of the full detector information. This way, algorithms that are computationally expensive are only executed on interesting events.

Chapter 3

Offline Event Reconstruction

This chapter describes how candidate electrons, muons jets, tracks and vertices are reconstructed from the raw detector data. A collection of complex software algorithms written in C++/python called CMSSW [85] is used for this reconstruction process. As described in the previous chapter, a stream of digital readout signals from the sub-detectors is recorded for each event that passes all levels of triggering. Powerful PC computing farms are deployed to analyze the data and reconstruct candidates of physical objects (electrons, jets, tracks, etc.) that can then be used in the final analysis.

Particle Flow Event Reconstruction: For this thesis analysis we used particle flow event reconstruction [86, 87] which basically aims at reconstructing and identifying all stable particles in the event, i.e., electrons, muons, photons, charged hadrons and neutral hadrons, with a thorough combination of all CMS sub-detectors towards an optimal determination of their direction, energy and type (Figure 3.1). Then the list of individual particles is used to build jets, to determine the missing transverse energy [88, 89, 90], to reconstruct and identify taus from their decay products, to quantify charged lepton isolation with respect to other particles, to tag b jets, etc.

Since this technique ensures precise measurement of the energy and direction for each particle, as it is combining the capabilities of each sub-detector, therefore a high efficiency and low fake rate is essential.

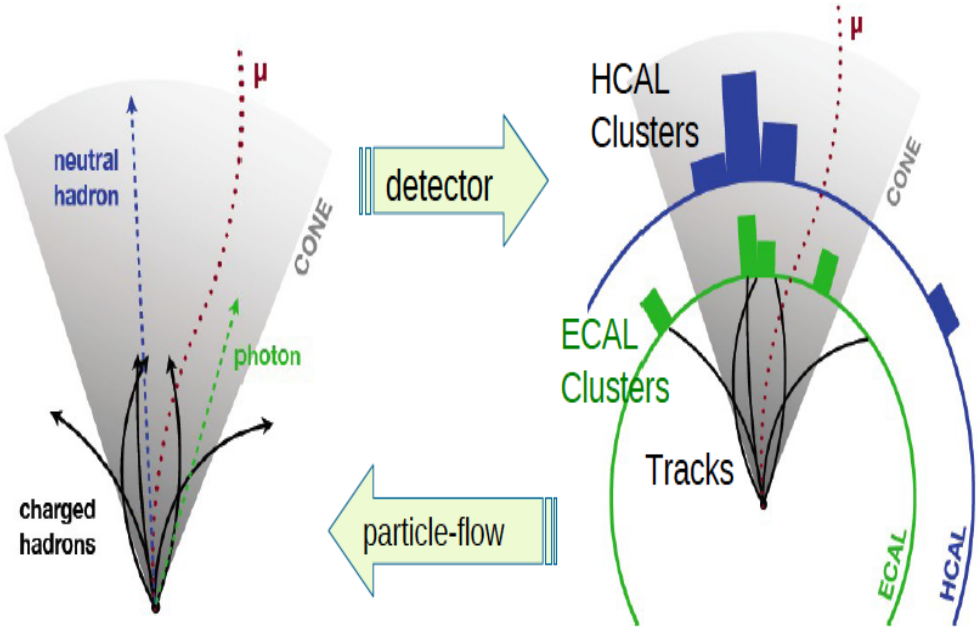


Figure 3.1: The Particle Flow algorithm. Particles in the CMS detector are seen as tracks and energy depositions. The PF algorithm attempts to fully reconstruct an event by combining information for all CMS sub-detectors.

3.1 Iterative Tracking

An iterative tracking [91] approach is used to achieve both high efficiency and low fake rate. It starts with tight track seeding criteria leading to moderate tracking efficiency and negligibly small fake rate. Next steps remove the hits unambiguously assigned to the tracks found in the previous iteration and loosens the seeding criteria in each step to gain efficiency. With the full five-step iterative approach charged particles with as few as three hits up to 50 cm away from the beam pipe and momenta down to 150 MeV can be reconstructed.

3.2 Calorimeter Clustering

The main purpose of clustering algorithm in the calorimeters is to detect and measure the energy and direction of stable neutral particles and separate these neutral particles from

energy deposits of charged hadrons, reconstruct and identify electrons and all accompanying bremsstrahlung photons and energy measurement of charged hadrons for high-pT tracks. This clustering algorithm aims for high detection efficiency for low-energy particles and towards a separation of close energy deposits. The Clustering is performed in each sub-detector separately: ECAL barrel, ECAL endcap, HCAL barrel, HCAL endcap, PS first layer and PS second layer. In the HF, no clustering is performed so far, so that each cell gives rise to one cluster. Clustering algorithm starts from identification of cluster seeds above the noise threshold and then forming topological clusters by a combination of geometrically adjacent calorimeter cells, which give rise to as many particle flow clusters as seeds.

3.3 Link Algorithm

For a given particle there could be many particle flow elements in various sub-detectors i.e. a charged particle track, several calorimeter clusters, muon track etc. The event display of Figure 3.2 shows a similar example of such case. Link algorithm connects

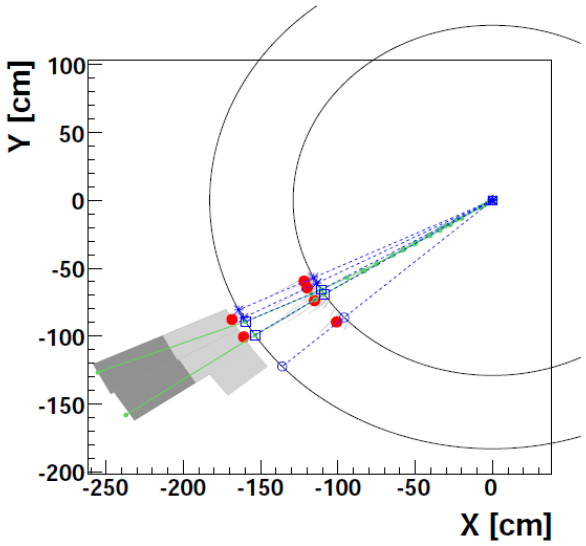


Figure 3.2: An event display of a simple hadronic jet in the (x, y) view

these elements to fully reconstruct each single particle and avoid double counting from different detectors. It tentatively performs for each pair of elements in an event and defines distance between them to quantify the quality of the link and then produces blocks of elements linked directly or indirectly. Due to fine granularity of CMS detectors, these blocks typically contain only one, two or three elements which constitute simple inputs for particle reconstruction and identification algorithm.

For a link between charged-particle track and calorimeter cluster, the track is extrapolated into the ECAL and HCAL, taking into account the typical shower shapes in all sub-detectors. A link to any given cluster is established if the extrapolated position is within cluster boundaries. To collect the energy of all bremsstrahlung photons emitted by electrons, tangents to the tracks are extrapolated to the ECAL from the intersection points between the track and each of the tracker layers. A cluster is linked to the track as a potential bremsstrahlung photon if the extrapolated tangent position is within the boundaries of the cluster.

For a link between two calorimeter clusters (between ECAL and HCAL cluster or between ECAL and PS cluster), the cluster position of more granular calorimeter should be within the cluster envelope in the less granular calorimeter. And for charged-particle track in tracker and a muon track in muon system, link is established when a global fit between the two tracks returns an acceptable χ^2 [49]. When several global muons can be fit with a given muon track and several tracker tracks, only the global muon that returns the smallest χ^2 is retained.

3.4 Particle reconstruction and identification

The reconstruction and identification of a set of particles from each block of elements is finally performed using the particle-flow algorithm. The resulting list of reconstructed particles constitutes a global description of each event, available for subsequent physics analysis.

3.4.1 Muon Reconstruction

Each global muon [92] gives rise to a "particle-flow muon" [86], if its combined momentum is compatible with that determined from the sole tracker within three standard deviations. The muon identification algorithm accepts tracks reconstructed in the muon system and attempts to associate these with tracks in the inner tracking detector. This algorithm reconstructs muons in two ways: outside-in and inside-out. In the outside-in approach a track reconstructed in the muon system is extrapolated to the outermost tracker layer, where a region of interest (compatible with the muon trajectory) is defined. If a matching track is found in the inner tracker, a refit of the whole muon track using the Kalman filter technique is carried out. If there are ambiguities then the global muon track is selected according to the χ^2 value of the track fit. This makes global muons, which provide high reconstruction efficiency at a moderate probability of misidentification A second way to reconstruct muons starts from a muon compatible track in the inner tracker and seeks for hits in the muon system taking into account the energy loss in the material and effects of multiple scattering in the calorimeters and the solenoid coil. At least one hit needs to be found in the muon system to reconstruct the muon as a tracker-muon. This procedure is slightly more efficient for a muon track momentum below 5 GeV. Within the particle flow algorithm no further requirements are imposed on isolated muons, while for non-isolated muons slightly tighter criteria on the track-quality are applied to remove imbalances in the jet response, which can occur if neutral energy is wrongly associated to a muon (i.e. lost in the jets energy).

3.4.2 Electron Reconstruction

Electron reconstruction is performed by combining tracking and ECAL information [87]. Seeds are formed from ECAL clusters as well as from tracks. For high electron momentum the seeding from the ECAL is very efficient (ECAL driven), while for low momentum electrons the seeding from tracks is more efficient (track driven). Each track of the block

is submitted to a pre-identification stage which exploits the tracker as a pre-shower: electrons tend to give rise to short tracks, and to lose energy by bremsstrahlung in the tracker layers on their way to the calorimeter. Pre-identified electron tracks are refit with a Gaussian-Sum Filter [93] in an attempt to follow their trajectories all the way to the ECAL. A final identification is performed with a combination of a number of tracking and calorimetric variables. Each identified electron gives rise to a particle-flow electron. The corresponding track and ECAL clusters (including all ECAL clusters identified as bremsstrahlung photons) are removed from further processing of the block. In case of electrons within a jet the energy is collected by summation of the ECAL deposits of geometrically adjacent clusters. For isolated electrons the energy is measured in super-clusters (clusters of clusters) in the ECAL, which collect bremsstrahlung photons that are emitted along the electron track in the tracker volume.

Electron Identification parameters

To differentiate genuine electrons from other particles that may mimic electron behavior, several "electron identification" variables are used. Since an electron is reconstructed from a ECAL supercluster matched to a track, so it is required that the match is good one. These identification variables will be helpful in distinguishing the signal electrons (from Z decay) from the other faking objects.

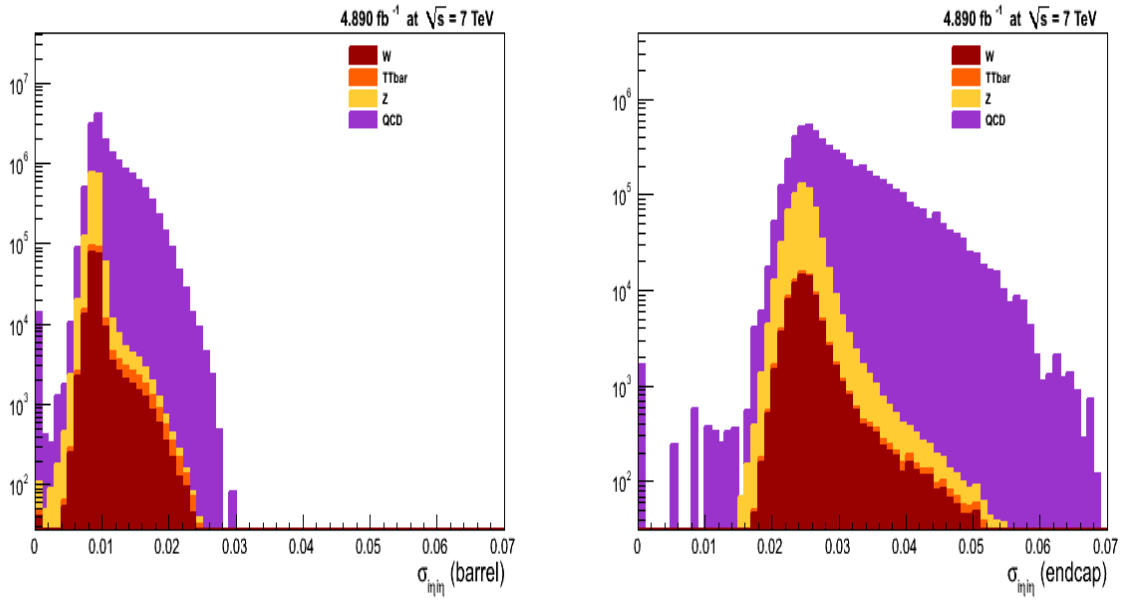
$\sigma_{i\eta i\eta}$ This is the ECAL shower shape variable which describes the width of the ECAL cluster along the η direction computed for all the crystals in the 5×5 block of crystals centered on the highest energy crystal of the seed cluster. Electrons coming from a Z boson have a small spread in η.

$\Delta\phi_{in}/\Delta\eta_{in}$ These variables describe the absolute difference in the $\phi(\eta)$ direction between the supercluster and the associated track as extrapolated to its vertex.

(H/E) This is the ratio of hadronic energy deposits in the Hcal to electromagnetic energy

deposits in the Ecal. The fraction of the electron energy deposited in the hadronic calorimeter is expected to be small.

Figures 3.3-3.6 show the behaviour of the identification variables [94, 95, 96] for signal and background samples.

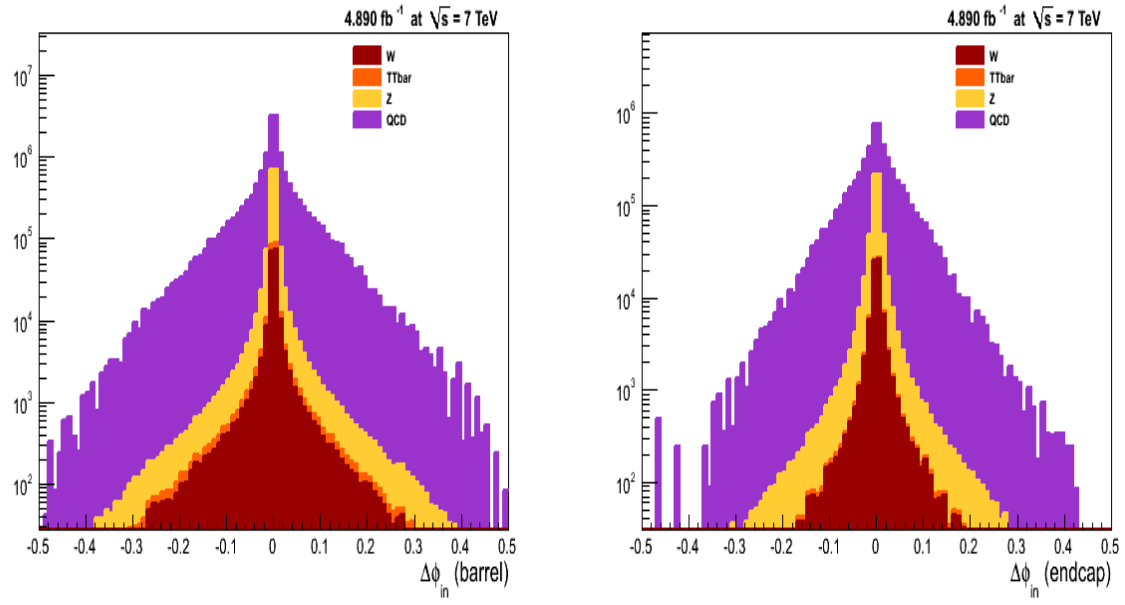


(a) Barrel (b) Endcap

Figure 3.3: $\sigma_{i\eta i\eta}$ electron identification variable distribution for (a) Barrel and (b) Endcap

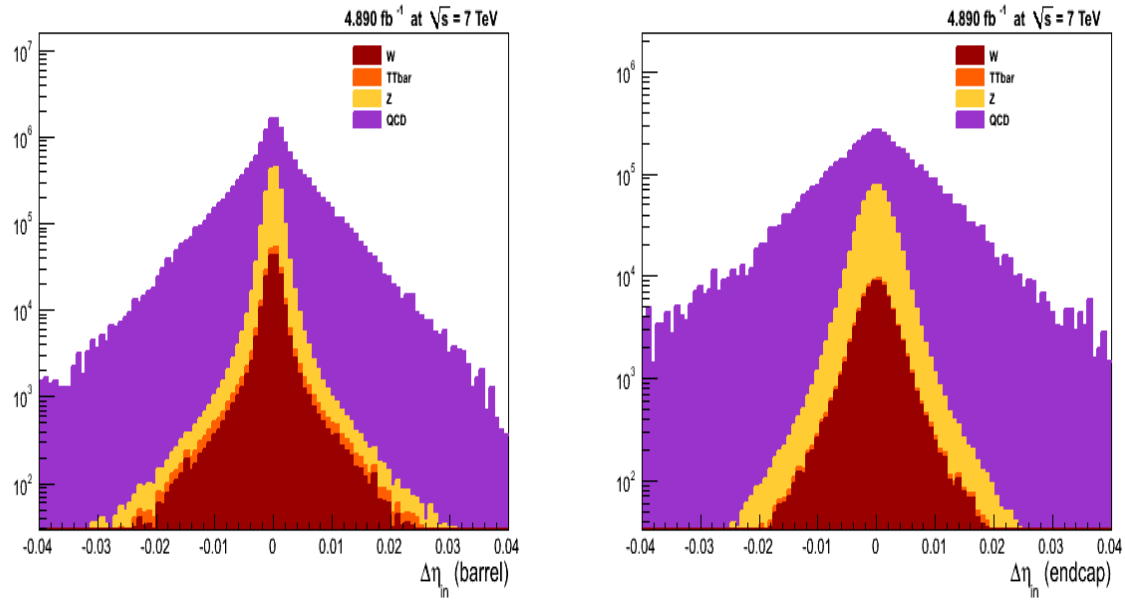
Charged and Neutral Hadrons

Charged hadron energy is obtained from a combination of the tracker and calorimeter measurements, when the two values are found to be compatible. Otherwise, if the energy measured in the calorimeters is small compared to the track momentum, a cleaning procedure to remove potential spurious or mis-reconstructed tracks is invoked [86]. Otherwise If calorimeter response is too large, the particle-flow algorithm assigns the energy excess to a photon and possibly a neutral hadron [86]. If the response of the calorimeter is not correctly accounted for, these effects would lead to a systematic mis-estimation of both the event energy and the particle multiplicity.



(a) Barrel (b) Endcap

Figure 3.4: $\Delta\phi_{in}$ electron identification variable distribution for (a) Barrel and (b) Endcap

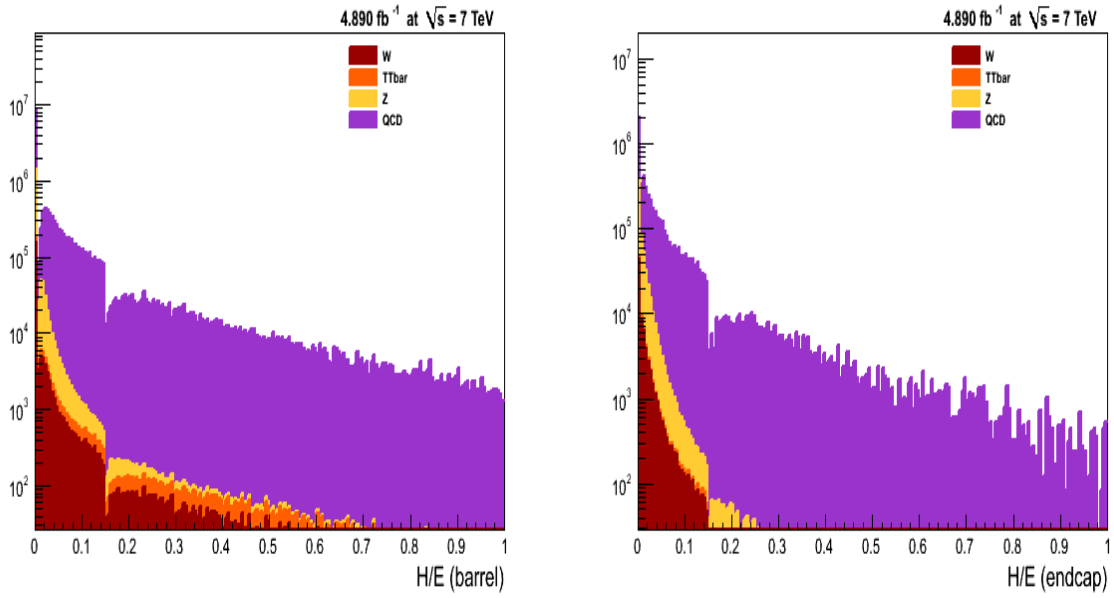


(a) Barrel (b) Endcap

Figure 3.5: $\Delta\eta_{in}$ electron identification variable distribution for (a) Barrel and (b) Endcap

Electron isolation

Real electrons (coming from a Z boson decay) are expected to be fairly isolated. Electron isolation variables are very powerful against QCD background. We compute the en-

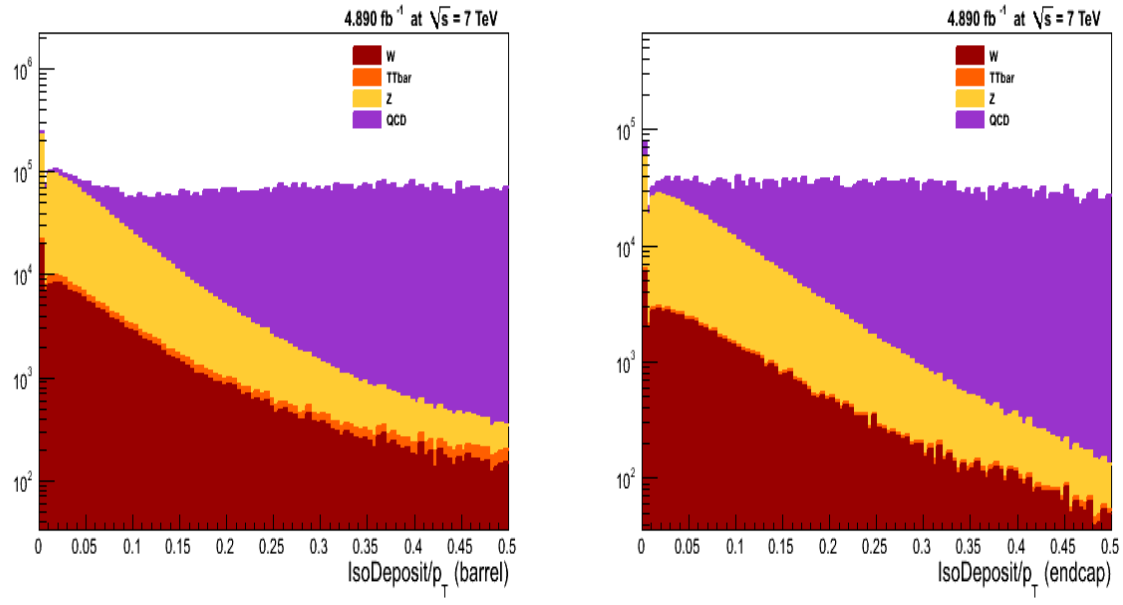


(a) Barrel

(b) Endcap

Figure 3.6: (H/E) electron identification variable distribution for (a) Barrel and (b) Endcap

ergy deposits of stable particles of any type (e.g. photon, neutral or charged hadron, etc). The momentum around the electrons candidate is summed within a cone of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.4$ We then define the relative combined isolation as $\Sigma p_T(charged\ hadrons) + \Sigma p_T(neutral\ hadrons) + \Sigma p_T(\ photons)/p_T$, where first sum runs over the transverse momenta of all particle flow charged hadrons and second is over the transverse momenta of all neutral hadrons and third over all the photons within the cone. Lepton itself is excluded from the summation. Figures 3.7 show the behaviour of the isolation variable for signal and QCD background samples. All prompt electrons are required to have an isolation of less than 0.2. Electron candidates tend to deposit most of their energy in a narrow region of the EM layers, while hadrons deposit their energies in the hadronic layers in a much wider radius.



(a) Barrel

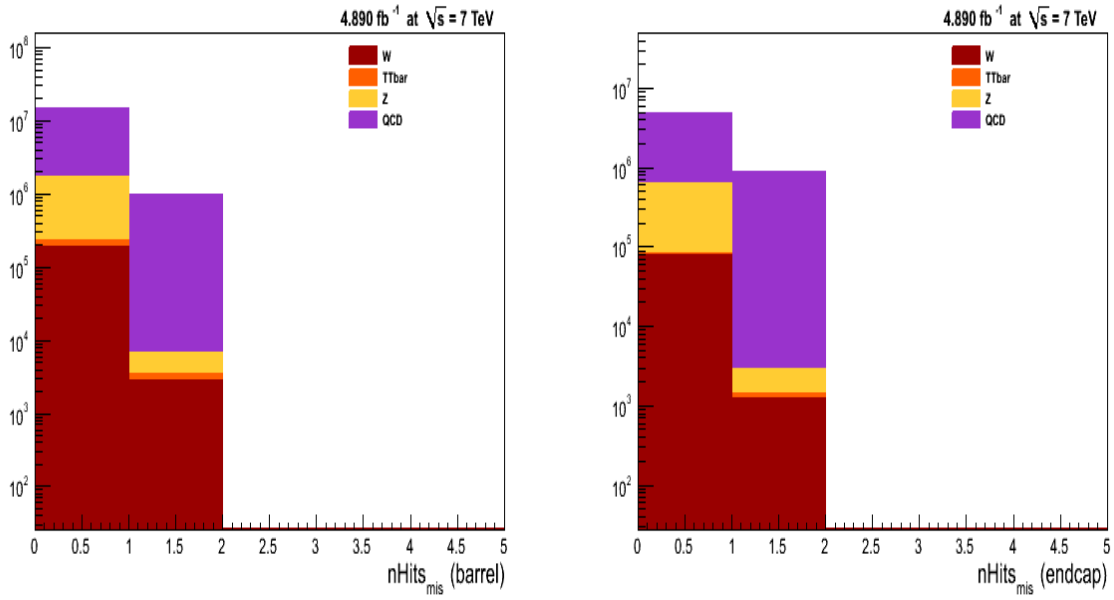
(b) Endcap

Figure 3.7: Combined relative particle flow electron isolation variable distribution shown for full ECAL acceptance

Photon Conversion rejection variables

Electrons from converted photons are a non-negligible background to prompt electrons coming from the Z boson decays [97]. In order to reject such backgrounds, electrons with missing expected hits in front of the innermost valid track hit are rejected as originating from a conversion that occurred in the tracker material. The general track collection is also inspected to locate possible conversion partner tracks. To be identified as a conversion partner, the track must:

- have opposite sign as the electron track
- approximately the same $\Delta cot\,(\theta)$ as the electron track
- small distance ("Dist") in the $R-\phi$ plane, where "Dist" is defined as the distance in the x-y plane between the two tracks when the track in question and the electron GSF track would be parallel when extrapolated.



(a) Barrel

(b) Endcap

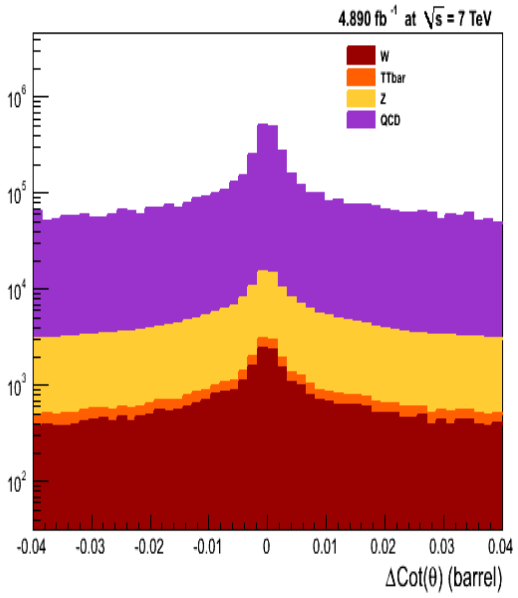
Figure 3.8: Conversion Rejection variable: the number of missing hits in the tracker (a) Barrel and (b) Endcap

3.4.3 Jet Reconstruction

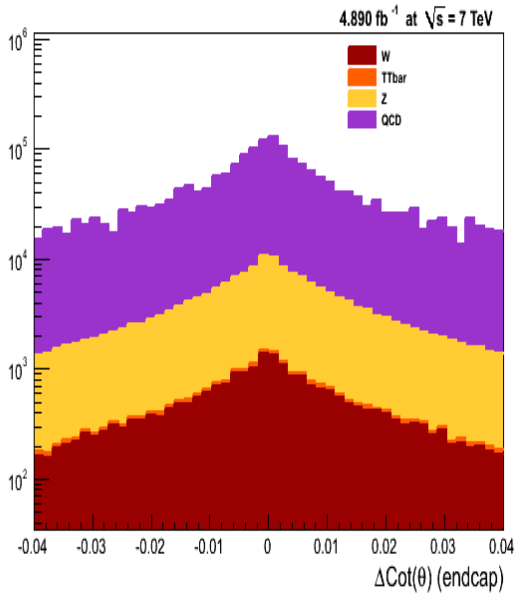
As discussed in Chapter-1, due to QCD confinement the emitted partons produced in hard collisions form colorless hadrons which result in collimated particle showers moving along the direction of the originating parton, called jets. These hadrons deposit their energy in the CMS calorimeters through electromagnetic and hadronic interactions; the collected information is then used to produce detector-level jets. In this section the techniques that are used to reconstruct and identify jets are discussed.

Jet Clustering Algorithm

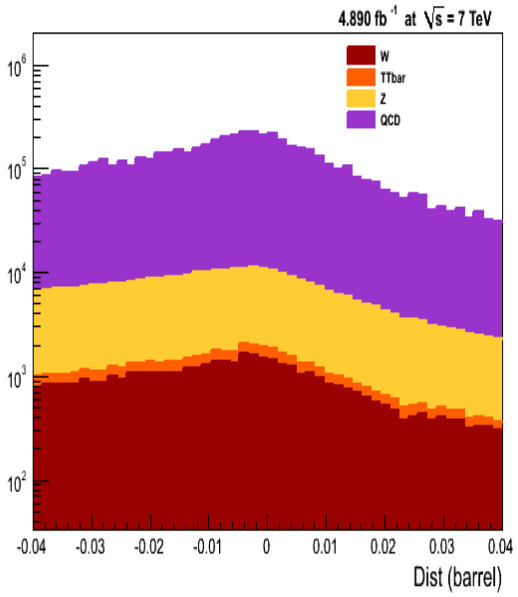
Depending on the reconstruction approach, jet clustering algorithms can be classified into two major groups. In the first group, cone-based algorithms take particles or calorimeter towers as seeds in an iterative search for stable cones in an event. In the second group, a distance is defined between pairs of particles and successive re-combinations of the pair of closest particles are performed, stopping when all resulting objects are too far apart. The



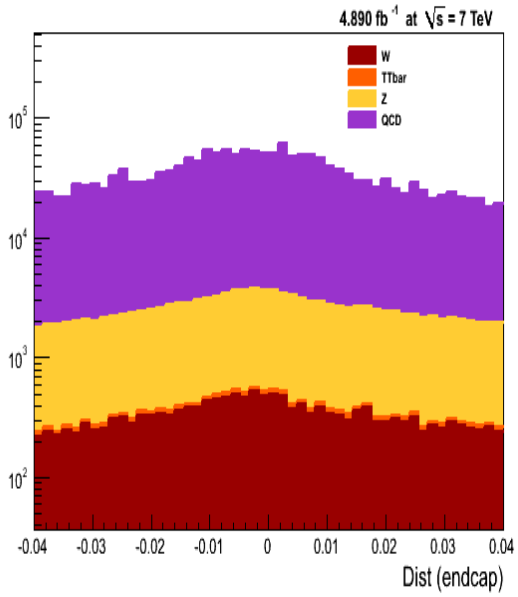
(a) $\Delta cot\,(\theta)$ in Barrel



(b) $\Delta cot\,(\theta)$ in Endcap



(c) "Dist" in Barrel



(d) "Dist" in Endcap

Figure 3.9: Conversion Rejection variable: $\Delta cot\,(\theta)$ in (a) Barrel and (b) Endcap and "Dist" in (c) Barrel and (d) Endcap

jet algorithm of the CMS-choice is the anti-k_t with $\Delta R = 0.5$. In the anti-k_t algorithm [98], one introduces distances d_{ij} between entities (particles, pseudojets) i and j and d_{iB} between object i and the beam (B). For clustering, smallest of the distances is calculated; if it is a d_{ij}, the objects i and j are recombined by adding their four-momenta, while if it

is d_{iB} it is labeled as a jet and is removed from the list of objects. Distance measures are defined as:

$$d_{ij} = \min\left(k_{ti}^{2p}, k_{tj}^{2p}\right) \frac{\Delta_{ij}^2}{R^2} \tag{3.1}$$

$$d_{iB} = k_{ti}^{2p} \tag{3.2}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and k_{ti}, y_i and ϕ_i are the transverse momentum, rapidity and azimuth of particle i respectively. In addition to the usual radius parameter R, a parameter p has been added to govern the relative power of the energy versus geometrical (Δ_{ij}) scales. The anti-k_t algorithm is a sub case of the above set of equations where $p = -1$, while for $p = 0$ and $p = 1$ we get the so-called k_T [99, 100] and Cambridge/Aachen [101] algorithms respectively.

Despite their appealing simplicity and long list of advantages, sequential clustering algorithms are not widely used at current hadron collider experiments, historically due to prohibitively growing CPU requirements for large numbers of input particles to be clustered per event. The CMS software however interfaces to the FastJet [102] package, which provides novel implementations of these algorithms, with reduced processing times.

Jet Energy Corrections

Since these jets are clusters of particle flow candidates, where certain thresholds are applied, the jets underestimate the total contained energy. Also the reconstructed jet has the energy usually different than that of the corresponding particle jet because of the non-uniform and non-linear response of the CMS calorimeters. Moreover, electronic noise and multiple proton-proton interactions in the same bunch crossing (pile-up events) can cause this extra amount of energy. The goal of the jet energy calibration is to find the relation between the energy measured in the detector jet and that of the corresponding particle jet. So, it is required to implement the additional corrections to the jet energies

as explained here.

CMS has adopted a factorized solution to the problem of jet energy corrections [103], where each level of correction takes care of a different effect. The set of corrections used are as follows:

L1 Offset Correction : The goal of the L1 correction is to remove the energy from pileup events, electronic noise and jet energy lost by thresholds. In principle this will remove any dataset dependence on luminosity so that the following corrections are applied upon a luminosity independent sample.

L2 Relative Jet Correction : The goal of the L2 Relative correction is to make the jet response flat versus η. Essentially, the uniformity in pseudorapidity is achieved by correcting a jet in arbitrary η relative to a jet in the central region ($|\eta| < 1.3$).

L3 Absolute Jet Correction : The goal of the L3 Absolute correction is to make the jet response flat versus p_T. Once a jet has been corrected for η dependence, it is corrected back to particle level, i.e. the corrected jet p_T is equal on average to the generator-level jet p_T.

L4 EMF (electromagnetic energy fraction) Jet Correction: The goal of the optional L4 EMF jet correction is to make the jet response uniform vs the electromagnetic energy fraction (EMF).

L5 Jet Flavor Correction: The goal of the optional L5 Flavor jet correction is to correct for the jet flavor dependence.

L6 Underlying Event: The goal of the optional L6 correction is to remove the luminosity independent underlying event energy from jet.

L7 Parton Jet Correction: The optional L7 parton correction is applied to correct back to the parton level , which means that the corrected RecoJet pt is equal to the originating parton pt on average.

L2L3Residuals: The Residual Calibration of Data: After the first collision data at $\sqrt{s}$ = 7TeV, it was realized that the CMS jet energy response simulation is very successful. However, the comparison between data and MC is not perfect but there are small differences depending on eta. Therefore, the JEC plan was adopted accordingly as described below:

- first, the MCtruth L2L3 JEC calibration is applied, which takes care of the bulk of the energy response.
- second, a small residual calibration (eta and pt dependent) is applied which fixes the small differences between data and MC.

By definition, the residual calibration is applied to data only. The MCtruth calibration is good enough for the MC. The need for a residual calibration will remain until CMS develops a perfectly tuned simulation that will reproduce the data features out of the box. The approach of a residual calibration, instead of a full-blown data driven one was found to be more effective in reducing the systematic uncertainties. Currently there is a need for a residual calibration only for the relative energy scale. The absolute energy scale seems to be modeled very well in the simulation.

Jet Identification parameters

A set of quality cuts is applied to every reconstructed jet in order to reduce fake jets from calorimeter noise [104]:

- Charged hadron fraction (chfJet)
- Neutral hadron fraction (nhfJet)
- Number of charged particles (charged jet constituents) (cmultiJet)
- Number of neutral particles (neutral jet constituents) (nmultiJet)
- Number of particles (jet constituents) (ncr)

- Fraction of energy deposited in the Electromagnetic (EM) by charged constituents : charged EM fraction (cemfJet)
- Fraction of energy deposited in the Electromagnetic (EM) by neutral constituents : neutral EM fraction (nemfJet)

Chapter 4

Monte Carlo Event Simulation and Data Samples

Event generator programs are an essential tool in particle physics with the goal to simulate nature as accurately as possible [105]. These generators of particle physics, aim to include all stages of interactions that occur at high energy collisions. Starting from the initial creation of partons in hadronic collisions (parton level), followed by parton showering from initial and final state partons and subsequent hadronization (particle or hadron level), and simulations of all detector elements (detector level) are some of main features of these generators. The output of Monte-Carlo simulations can be used to test theoretical predictions, to estimate efficiencies and acceptances, to study background processes and to correct data for detector effects so that direct comparisons to theoretical predictions can be made. This chapter is a brief description of the Monte Carlo (MC) simulation tools and techniques that have been used in this thesis work. It also gives the description of the data and MC samples (in Sections 4.5 and 4.6) used for this analysis.

4.1 The Pythia Event Generator

PYTHIA is a general purpose MC event generator [106] which describes many physical aspects of a typical high-energy event:

- The initial beam particles coming towards each other, such as protons in LHC, are characterized by a set of parton distribution functions (PDF) which defines the partonic substructure.

- One parton from each initial particle radiates through a Bremsstrahlung - like process, such as $q \to qg$, to initiate a sequence of branchings which build up an initial-state "shower" (Initial State Radiation or ISR).

- One incoming parton from each of the two showers enters the hard process, which produces a number of outgoing particles. PYTHIA is optimized for hard processes that have two particles in the initial and final states, such as $q\bar{q} \to Z/\gamma* \to e^+e^-(2 \to 2process)$.

- Final state particles can also radiate to initiate a final-state shower (Final State Radiation or FSR).

- Initial and final-state color coherence effects are incorporated via the Angular Ordering approximation.

- The remaining partons (spectators) in the two incoming hadrons may interact.

- Beam remnant interactions are taken into account.

- Outgoing quarks and gluons form color neutral hadrons following the string hadronization model.

- Many of the produced hadrons are unstable and decay further.

To describe the hadronization process, the Lund string model [107] is used. Along with the hard scatter, the underlying event contributes to what will be observed in main process and constitutes a background to our measurement. It originates from the remaining partons of the incoming hard interaction protons, as well as additional pp interactions in the same beam crossing. For the description of underlying events, several

phenomenological models exists, with various degrees of sophistication and due to the inherent uncertainties, a different number of "tunes" have been developed. For CMS, Z2 tune [108] agree better with some aspects of data. A schematic representation of a generated event is shown in Figure 4.1, showing the hard collision which would create our signal event, along with the secondary interactions which affect our measurement as well as the hadronization of quarks and gluons and the decays of short-lived particles.

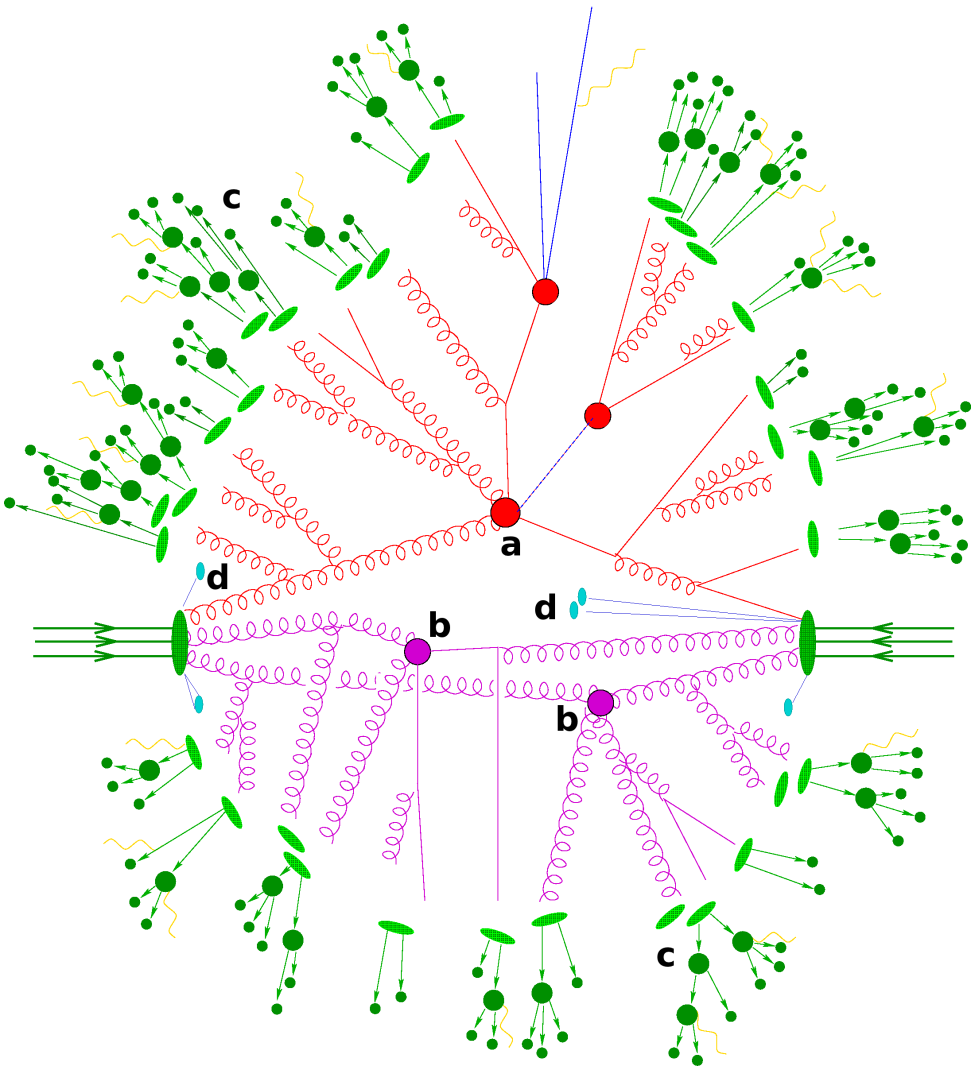


Figure 4.1: Schematic representation of a generated event. Apart from the hard collision (a), softer multiple interactions also take place (b) between the two incoming protons. Resulting quarks and gluons hadronize and decay (c). Beam remnants are marked with (d).

4.2 Combining Matrix Elements with Showering

PYTHIA produces multijet events by incorporating 2→2 hard processes with parton showering that is subsequently interfaced with the string hadronization model. Since parton showers are only approximations of soft parton radiation in the limits of low p_T and small emission angles (soft and collinear limits), final states with several high p_T and well separated jets are not properly described. Such final states occur frequently at LHC due to the large momentum transfer, Q^2. In order to give a more accurate description of such processes, matrix element generators are combined with parton shower simulations by dividing phase space into two regions, separated by a p_T matching threshold. In the high p_T region, jet production relies on matrix elements; the low p_T region is described by the parton showering simulation. Exact book-keeping is necessary when combining the two regions in order to avoid the double counting of parton configurations [109, 110].

4.2.1 MadGraph combined with PYTHIA

MADGRAPH [111] is a tree-level matrix element generator, and it is based on specifying initial and final state particles for any tree level Standard Model (SM) process. It creates a list of all relevant Feynman diagrams and calculates the corresponding matrix elements. The program is able to calculate matrix elements for any SM process. The MADGRAPH output is interfaced with PYTHIA for showering and hadronization. The matching between MADGRAPH and PYTHIA, to avoid double counting when combining different multiplicity final states, is done following a modified CKKW prescription [112, 113, 114]. Given a process, it automatically creates the amplitudes for all the relevant sub-processes and produces the mappings for the integration over the phase space. This process-dependent information is passed to MadEvent which allows the calculation of event cross sections and the generation of unweighted events. The event information (particle IDs, momenta, spin etc.) is stored in the "Les Houches" format [115] and then interfaced with PYTHIA which handles the rest of the generation steps (involving parton

showering, hadronization etc.).

4.2.2 FEWZ

FEWZ (Fully Exclusive W and Z Production) [116] is a simulation code for production of lepton pairs at hadron colliders through the Drell-Yan process. It can calculate cross-section at next-to-next-to-leading-order (NNLO) in the strong coupling constant.

4.2.3 POWHEG

POWHEG (Positive Weight Hardest Emission Generator) [117] is a method to implement NLO calculations with shower Monte Carlo's. It generates the hardest emission first, with NLO accuracy, independently of the shower Monte Carlo and with positive weights.

4.3 Detector Simulation and Digitization

After all final state particles are produced, the full detector effects can be simulated with the detailed simulation of the CMS detector which is based on the GEANT4 [118] toolkit. This is done in two steps. First step involves the simulation of interactions of all visible, stable particles with the detector. GEANT4 takes into account all these different interactions (bremsstrahlung, multiple scattering etc.) of all particles with various materials in the detector. And thus the simulated hits in the detector are obtained. In second step, these simulated hits are used as input to emulators which model the response of the detector readout and trigger electronics and digitize this information by also taking into account noise and other factors. This step is known as digitization. After this the full detector response is available (as in real operation) and the event can be reconstructed.

4.3.1 Pileup simulation

For a bunch crossing rate of 50 ns, the approximate average number of pp interactions per bunch crossing is 10. In such conditions, the rare hard collision which produces the physics

objects recognized by the trigger system and fulfilling the criteria for data acquisition, is usually accompanied by several additional pp interactions overlapping with it in the same bunch crossing. These additional collisions, which are typically of low energy but may still produce significant contributions to global event characteristics such as total visible energy or charged particle multiplicity, are denoted as pile-up events. The analysis of the hard collision properly includes the effect of pile - up, which is also modeled in all the necessary Monte Carlo (MC) simulated samples (will be discussed in Section 4.6.3).

4.4 Full Simulation

As discussed in this Chapter, the full simulation of a collision event at CMS, consists of three phases. During *event generation*, a collision is simulated starting from the proton-proton interaction until the production of the final decay products to be observed with the CMS detector. The *detector simulation* models the interactions of the generated final-state particles as they pass through the CMS detector materials and the responses of the different sub-detectors and the *digitization* step emulates the response of the detector electronics to the detector hits.

All these steps are performed efficiently by a C++ based software framework called CMSSW which is developed by CMS collaboration [85]. It takes care of interfacing with different event generators, passing the information to the detector simulation and digitization software to produce an output file in a format similar to the actual detector readout. On this information the event reconstruction code is executed and produces containers of electrons, muons, jets and other event information to be used for analysis. This chain is referred-to as "full simulation" with the detector simulation being the most time consuming and CPU-intensive step. Finally, full information in stored in ROOT format (a C++ based software) [119].

4.5 Recorded Data

The data analyzed in this thesis corresponds to the full 2011 CMS dataset collected between March - October, 2011 pp collisions of LHC $\sqrt{s} = 7$ TeV running. It corresponds to an integrated luminosity of

$$(4.890 \pm 0.11) fb^{-1}. \tag{4.1}$$

The luminosity measurement has an uncertainty of 2.2% [120]. The luminosity profile rapidly increased throughout data taking period. Hence the double-electron triggers used in this thesis analysis have been changed multiple times (mentioned in Section 4.5.1). The changes in these trigger paths are documented in Section 4.5.1; the average electron trigger efficiency throughout data taking was about 99% (Section 5.7.5). All data analyzed has been put through a quality validation chain requiring all detector subsystems to be flagged as good. The raw data was processed with $CMSSW\_4\_4\_X$ version of the CMS reconstruction software and contains approximately 65 million events of double-electron triggered data with run number from 160404 to 180252. Table 4.1 gives the names of specific datasets used for the present analysis. Data-taking was split into two running

Table 4.1: Primary datasets used in analysis

Name	Number of events	Run-Range
DoubleElectron-Run2011A-08Nov2011-v1	49853042	160329 - 175770
DoubleElectron-Run2011B-19Nov2011-v1	15618141	175832 - 180296

periods: Run2011A and Run2011B, where Run2011A is the period of low instantaneous luminosity ($\approx 1 \times 10^{32} cm^{-1} s^{-2}$ to $1 \times 10^{33} cm^{-1} s^{-2}$) and Run2011B has instantaneous luminosities above $1 \times 10^{33} cm^{-1} s^{-2}$.

4.5.1 Double Electron Triggers

In order to study the di-electron final states with controlled trigger rate, CMS has designed the di-electron triggers as described in Table 4.2.

As the luminosity of pp collisions was continuously increasing with time, so in order to control the total trigger rate at instantaneous luminosities above $1 \times 10^{33} cm^{-1} s^{-2}$, the trigger paths with loose selection criteria were prescaled and new trigger menus with some tighter selections on electrons designed to have a set of unprescaled triggers. Only the unprescaled double electron triggers were used for this analysis.

The main dielectron triggers (Table 4.2), require the two HLT electron candidates to have loose shower shape and calorimeter isolation requirements on both legs and a match to a Level-1 seed for the leading leg. A $p_T > 17$, 8 GeV is required at the HLT level for the two selected electron legs. This allow with the offline selection of $p_T >$ 20 GeV for the two daughter electrons.Since controlling the total trigger rate is most challenging in the dielectron channel, due to large fake electron background rates; so, additional requirements must be added to the track-to-cluster matching and track isolation. The identification and isolation requirements applied at the HLT level are described in Table 4.3.

4.6 Monte Carlo Datasets

To compare the data to the theoretical predictions, Monte Carlo event generators are used which then simulated with CMS detector geometry and then event reconstruction was done as described in previous chapter.

4.6.1 Signal Dataset

Since this analysis deals with two electrons and jets as final decay products, so the signal is modeled with a MadGraph [111] matrix element generator (interfaced with Pythia for parton showering). Drell-Yan plus up to four jets events are generated where the mass of the generated boson is required to be above 50 GeV/c^2 and it is restricted to leptonic decays only (Table 4.4). The cross-sections for Signal has been calculated to NNLO using FEWZ [121, 122] generator. The set of parton distribution functions used is

CTEQ6L1 [123]. This sample serves as the baseline for comparison with data. We have used POWHEG samples for systematic calculations. This sample generates Drell-Yan process where the electronic decay mode with the generated boson invariant mass above 20 GeV/c^2 has been considered (Table 4.4).

4.6.2 Background Datasets

There are many other processes which have signal like final decay products and mimic as signal. Such processes act as background and are irreducible even after the full event selection criteria. The electroweak backgrounds like W+jets, $t\bar{t}$+jets, are generated by Madgraph matrix element generator (interfaced with Pythia for the parton showering) and are listed in Table 4.5. The QCD samples have been produced from the Pythia-LO generator using an electromagnetic-enriching filter that requires isolated electron in the final state at generator level. The filter increases the probability of a reconstructed electron candidate being present in the event and hence decreases the production time. The quoted cross-sections for W and top backgrounds have been calculated to NNLO using FEWZ [121, 122] NLO generator. For QCD, LO pythia cross-sections are used. All the samples use the **TuneZ2** [108] tuned set of parameters to model the underlying event activity in the final state.

In order to compare with the data distributions, all these simulation samples are normalized to the cross sections times the integrated luminosity, using the next-to-next-to-leading-order (NNLO) cross sections or the leading-order (LO) cross sections from the MC generators, as reported in Tables 4.4 and 4.5.

4.6.3 PileUp Simulation and reweighting

As discussed in Section 4.3.1, in order to account for pile-up, all these MC samples are generated with minimum bias interactions overlaid on top of the hard scattering event. The number of minimum bias interactions follows a Poisson distribution of mean ten,

which is appropriate for the 2011 data since the mean number of interactions per bunch crossing ranges between 0 and 10 in the data period considered (Figure 4.2). Such effect of pileup is considered in the Montecarlo production by simulating in each event an additional amount of vertex (Figure 4.3). However the number of vertex in the LHC machine can change depending on the working conditions, while for the Montecarlo samples this distribution is fixed. In order to fix this problem Monte-Carlo samples are re-weighted by a factor $N_{j,data}/N_{j,MC}$ where N_j is the number of events with number of vertex "j" for data and MC respectively. The corresponding Pile-UP (PU) event weights applied in this thesis analysis are shown in Figure 4.4.

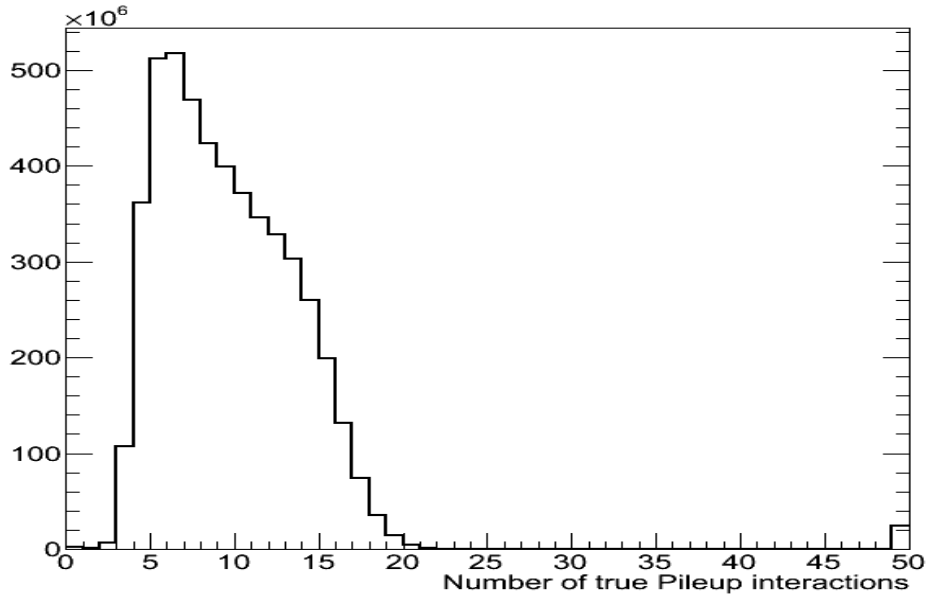


Figure 4.2: Number of true pileup interactions from Data

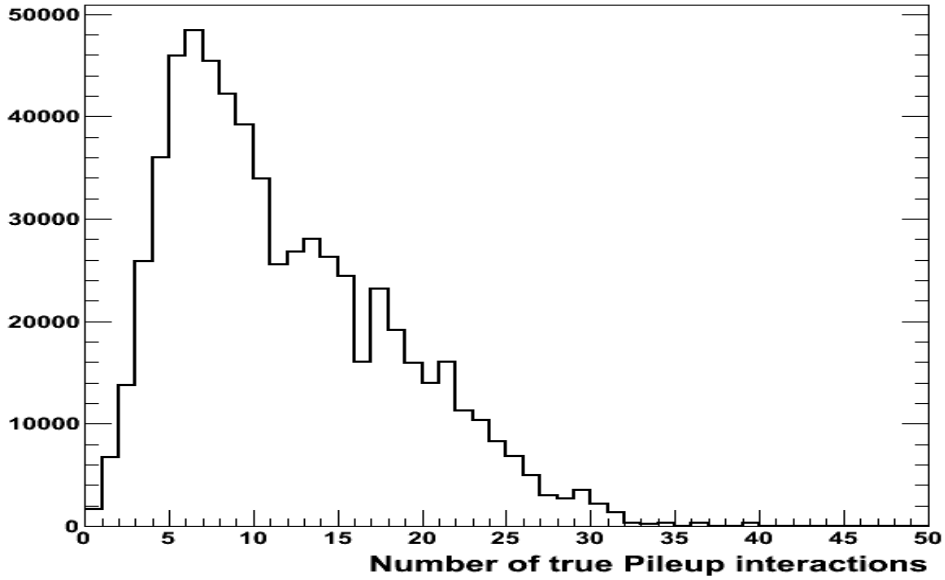


Figure 4.3: Number of true pileup interactions simulated in Monte Carlo samples

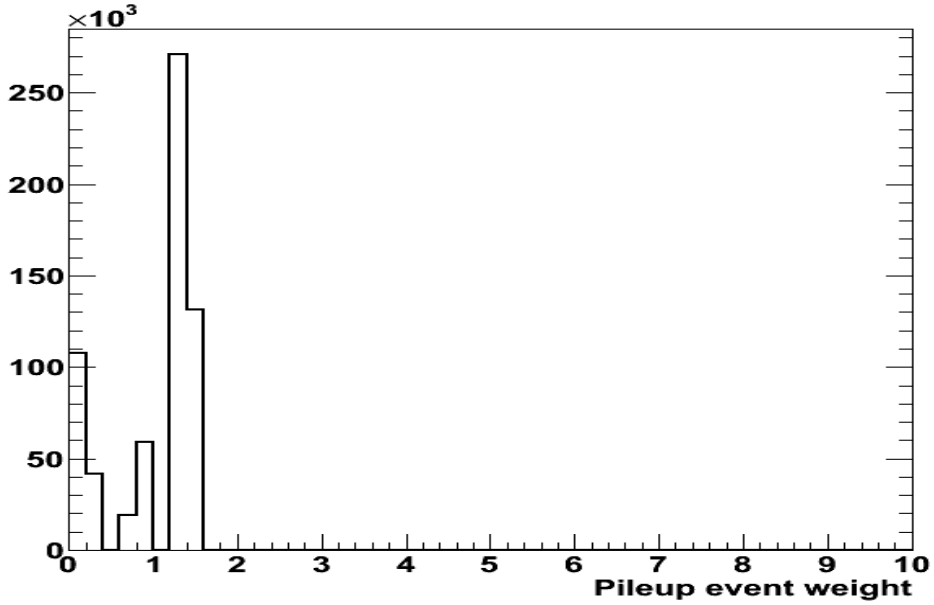


Figure 4.4: Weight factors applied to the Monte Carlo simulated events

Table 4.2: Summary of triggers used in analysis

Trigger Path	Run Range
$HLT\_Ele17\_CaloIdL\_CaloIsoVL\_Ele8\_CaloIdL\_CaloIsoVL\_v1$	160404 -161176
$HLT\_Ele17\_CaloIdL\_CaloIsoVL\_Ele8\_CaloIdL\_CaloIsoVL\_v2$	161217 -163261
$HLT\_Ele17\_CaloIdL\_CaloIsoVL\_Ele8\_CaloIdL\_CaloIsoVL\_v3$	163270 -163869
$HLT\_Ele17\_CaloIdL\_CaloIsoVL\_Ele8\_CaloIdL\_CaloIsoVL\_v4$	165088 -165633
$HLT\_Ele17\_CaloIdL\_CaloIsoVL\_Ele8\_CaloIdL\_CaloIsoVL\_v5$	165970 -166967
$HLT\_Ele17\_CaloIdL\_CaloIsoVL\_Ele8\_CaloIdL\_CaloIsoVL\_v6$	167039- 167913
$HLT\_Ele17\_CaloIdT\_CaloIsoVL\_TrkIdVL\_TrkIsoVL\_Ele8\_CaloIdT\_CaloIsoVL\_TrkIdVL\_TrkIsoVL\_v6$	170722- 170759
$HLT\_Ele17\_CaloIdT\_CaloIsoVL\_TrkIdVL\_TrkIsoVL\_Ele8\_CaloIdT\_CaloIsoVL\_TrkIdVL\_TrkIsoVL\_v7$	170826- 173198
$HLT\_Ele17\_CaloIdT\_CaloIsoVL\_TrkIdVL\_TrkIsoVL\_Ele8\_CaloIdT\_CaloIsoVL\_TrkIdVL\_TrkIsoVL\_v8$	173236- 175875
$HLT\_Ele17\_CaloIdT\_CaloIsoVL\_TrkIdVL\_TrkIsoVL\_Ele8\_CaloIdT\_CaloIsoVL\_TrkIdVL\_TrkIsoVL\_v9$	178420- 179889
$HLT\_Ele17\_CaloIdT\_CaloIsoVL\_TrkIdVL\_TrkIsoVL\_Ele8\_CaloIdT\_CaloIsoVL\_TrkIdVL\_TrkIsoVL\_v10$	179959- 180252

Table 4.3: Summary of requirements applied to electrons in the triggers used for this analysis. The selection requirements are given for electrons in the barrel (endcap). Abbreviations: L=Loose, VL=Very loose, T=Tight, VT=Very Tight.

name	criterion
CaloId_L	H/E $<$ 0.15(0.10) $\sigma_{\eta\eta}<$ 0.014 (0.035)
CaloId_VT	H/E $<$ 0.05(0.05) $\sigma_{\eta\eta}<$ 0.011 (0.031)
TrkId_VL	$\|\Delta\eta\|<$ 0.01 (0.01) $\Delta\phi<$ 0.15 (0.10)
TrkId_T	$\|\Delta\eta\|<$ 0.008 (0.008) $\Delta\phi<$ 0.07 (0.05)
CaloIso_VL	ECalIso/E_T $<$ 0.2 (0.2) HCalIso/E_T $<$ 0.2 (0.2)
CaloIso_T	ECalIso/E_T $<$ 0.15 (0.075) HCalIso/E_T $<$ 0.15 (0.075)
CaloIso_VT	ECalIso/E_T $<$ 0.05 (0.05) HCalIso/E_T $<$ 0.05 (0.05)
TrkIso_VL	TrkIso/E_T $<$ 0.2 (0.2)
TrkIso_T	TrkIso/E_T $<$ 0.15 (0.075)
TrkIso_VT	TrkIso/E_T $<$ 0.05 (0.05)

Table 4.4: The MC Signal dataset

Sample	$\sigma \times B.R$	N_{Evts}
$DYJetsToLL\_TuneZ2\_M-50\_7TeV-madgraph-tauola$	3048	36277961
$DYToEE\_M-20\_CT10\_TuneZ2\_7TeV-powheg-pythia$	3048	29497207

Table 4.5: The MC Samples used

Sample	$\sigma \times B.R \times filter\ efficiency$	N_{Evts}
$WJetsToLNu\_TuneZ2\_7TeV - madgraph - tauola$	31314	81352581
$TTJets\_TuneZ2\_7TeV - madgraph - tauola$	157.5	3701947
$QCD\_Pt - 20to30\_EMEnriched\_TuneZ2\_7TeV - pythia6$	236100000*0.0106	35729669
$QCD\_Pt - 30to80\_EMEnriched\_TuneZ2\_7TeV - pythia$	59440000*0.061	70392060
$QCD\_Pt - 80to170\_EMEnriched\_TuneZ2\_7TeV - pythia6$	898200.0*0.159	8150672
$QCD\_Pt - 20to30\_BCtoE\_TuneZ2\_7TeV - pythia6$	236100000*0.00059	2081560
$QCD\_Pt - 30to80\_BCtoE\_TuneZ2\_7TeV - pythia6$	59440000*0.00242	2030033
$QCD\_Pt - 80to170\_BCtoE\_TuneZ2\_7TeV - pythia$	898200.0*0.0105	1082691

Chapter 5

Event Selection

This chapter deals with the selection of events and reconstructed objects. In order to assure that events with two high p_T electrons contained within the ECAL acceptance and originating from the decay of a Z gauge boson are selected, the selection criteria as mentioned in following sections (5.2, 5.3, 5.4) of this chapter are applied. After identifying the Z candidate events (Section 5.3), the presence of high p_T jets is required. Section 5.7 describes the methods to obtain object reconstruction and selection efficiencies.

5.1 Event Pre-selection

In this Section, we describe the main event filters which are required to remove detector noises, beam backgrounds etc.

- **Beam background removal:** Some general cuts to select good collision events are needed to remove potential beam background or events where no hard collision took place. Events that contain a high number of poorly reconstructed tracks are likely to be produced by beam induced hits in the pixel detector [124]. To remove this background each event containing more than 10 tracks is required to have at least 25% of high purity tracks (i.e. a tighter set of track quality criteria required to be fulfilled). All events with less than 10 tracks are accepted.

- **Primary Vertex Filter:** To select good collision events, each event is required to contain at least one well reconstructed primary vertex. Primary vertices are reconstructed using the so-called Deterministic Annealing (DA) clustering of tracks [125]. In events with more than four tracks having transverse momenta above 0.5 GeV the efficiency of the primary vertex finding is almost 100% [124]. The vertex is required to have at least four degrees of freedom, $N_{DOF} = 2\Sigma_i^{nTracks} w_i - 3$, where w_i is the weight of each track, close to one for tracks compatible with the common vertex. The vertex exhibiting the highest quadratic track momentum sum is chosen to be primary. Reconstructed primary vertices are required to have a z-position within 24 cm of the nominal detector center and a radial position within 2 cm of the beamspot. These cuts remove events in the very tail of the quantities and introduce no additional uncertainty, since the distributions are well simulated [124]. These cuts are applied on the simulated samples as well. Reconstructed leptons will be required to have small impact parameters with respect to this vertex.

- **Anomalous HCAL noise filter:** This filter helps to reject noisy HB/HE (HCAL Barrel/Endcap) events. It is a filter that identifies events containing an readout box (RBX) with bad pulse-shape, timing, hit multiplicity (a high multiplicity of hits in an hybrid photodiode (HPD) or RBX indicates noise), and ADC 0 counts (high multiplicity of ADC 0 counts in an RBX indicates noise). It is designed to reduce noise rate by factor of 100.

5.2 Electron Selection

Electron candidates have to satisfy following basic kinematic requirements:

- Impact parameter $|d_{xy}|$ <0.02cm
- $p_T > 20$ GeV
- $|\eta| < 2.4$, exclude gap region i.e. 1.442 - 1.556

In order to be efficient to get the prompt electron candidates and reduce the background contamination, some selection cuts are applied on idenification variables (described in Section 3.4.2) , which are summarized here in Table 5.1.

Table 5.1: Electron Identification variables

	Barrel	Endcap
$\sigma i\eta i\eta$	< 0.01	< 0.03
$d\phi$	<0.06	< 0.03
$d\eta$	<0.004	< 0.007
H/E	< 0.04	< 0.15
missing hits	<= 0	<=0
abs(dist)	> 0.02	> 0.02
abs(dCot)	> 0.02	> 0.02
Combined Rel.(PF)Isolation (0.4)	<0.2	<0.2

5.3 Z selection

Z candidates are selected based on following criteria:

- Two highest p_T electrons passing electron selection requirements mentioned in Section 5.2.
- Both electrons should have opposite charge.
- Di-electron invariant mass window: 70 GeV $< M_{ee} <$110 GeV.

5.4 Jet Selection

Jets are formed using the anti-kT jet ($\Delta R = 0.5$) clustering algorithm (described in Section 3.4.3) and should pass the jet identification criteria summarized in Table 5.2 (described in Section 3.4.3). These jets are corrected for p_T, η dependence using JEC scheme implemented in CMS (described in Section 3.4.3), summarized here:

- Data: L1FastJet+L2Relative+L3Absolute +L2L3Residual

- MC: L1FastJet+L2Relative+L3Absolute

These corrected jets then should have:

- $p_T > 30$ GeV
- $|\eta| < 2.4$
- Since the jet algorithm can identify fake jets originating from electron energy deposits, all jets overlapping with electrons coming from the $Z \to e^+e^-$ boson within $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\Phi)^2}$ of 0.3 are removed (Figure 5.1).

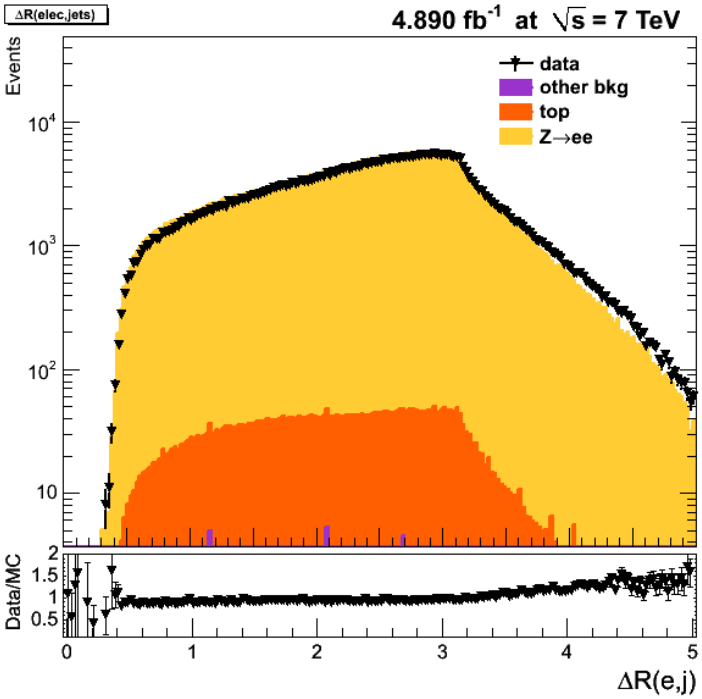


Figure 5.1: Distribution for $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\Phi)^2}$ between the $Z/\gamma*(\to e^+e^-)$ daughters and jets

5.5 Event Statistics

Figure 5.2 shows the distribution for jet multiplicities for data compared to signal and background samples. Table 5.3 summarizes the number of $Z \to e^+e^-$ event candidates for different exclusive jet multiplicities.

Table 5.2: Jet Identification

quantity	cut
chf	0
nhf	0.99
cmult	0.0
nConstituent	1
cemf	0.99
nemf	0.99

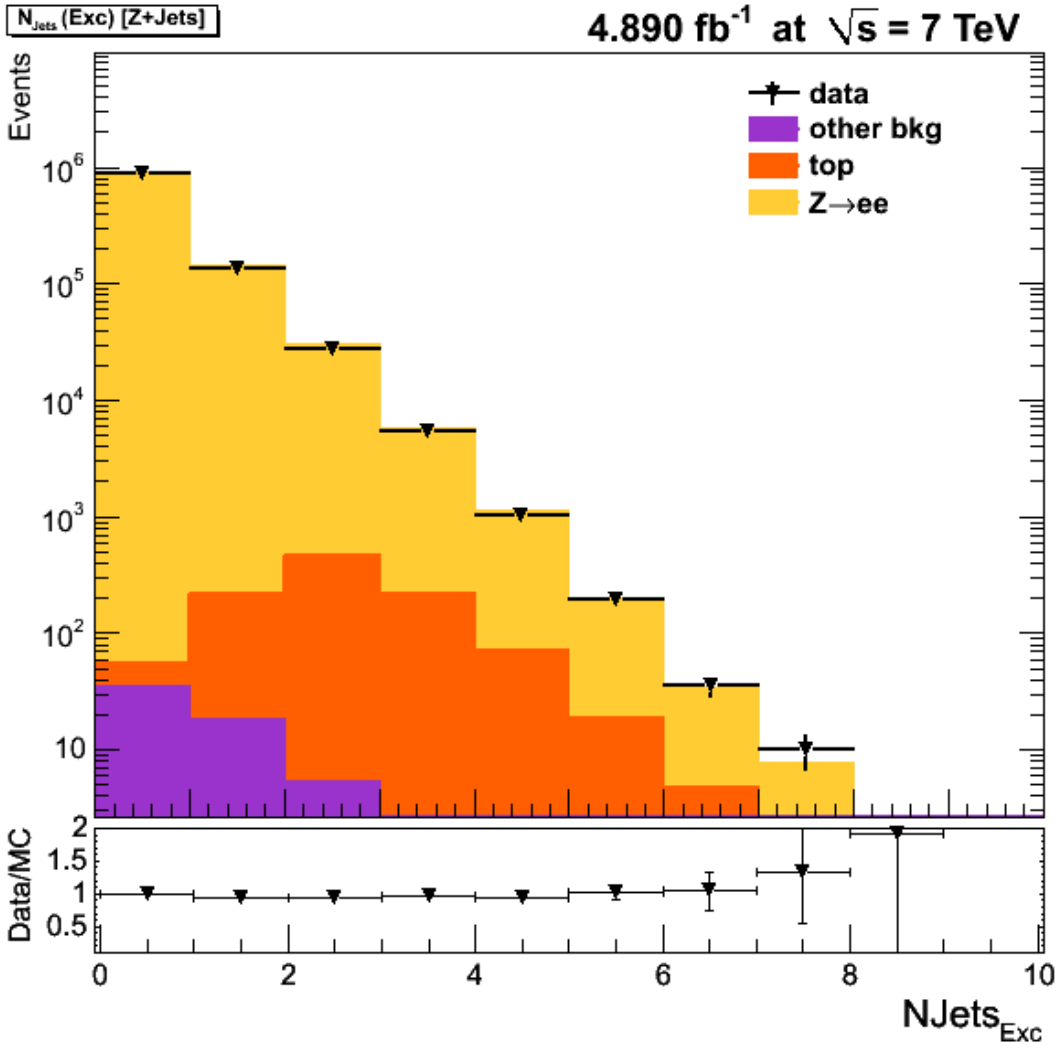


Figure 5.2: Distribution for jet multiplicities

5.6 Data Vs Monte Carlo

This section presents a comparison of basic kinematic distributions for electrons, Z candidates and jets between data and MC simulations (MC distributions are normalized to

Table 5.3: Event breakdown for Data by exclusive jet multiplicities before any background is subtracted or any corrections are applied

sample	N	Fraction
Z/γ*+0jets	887745	0.8405
Z/γ*+1jets	133801	0.1267
Z/γ*+2jets	27956	0.0264
Z/γ*+3jets	5391	0.0051
Z/γ*+4jets	1025	0.0010
Z/γ*+5jets	198	0.0002
Z/γ*+6jets	35	0.0000
Total	1057649	-

the integrated luminosity in data). It is important that the MC distributions describe the data distributions as accurately as possible. The MC simulations also pass all the above selection cuts and hence represent the fraction of events that are left after kinematic and geometric electron cuts, the di-electron invariant mass cut and the primary vertex cut. The event counting as a function of exclusive jet multiplicities is given in Table 5.4.

Table 5.4: Event breakdown by exclusive jet multiplicities for Data and MC (before any background is subtracted or any corrections are applied)

sample	Data	ZJets-MC	TTJets-MC	WJets-MC	QCD-MC
Z/γ*+0jets	887745	917298	21	40	0
Z/γ*+1jets	133801	141149	203	20	0
Z/γ*+2jets	27956	29412	460	5	0
Z/γ*+3jets	5391	5342	220	0	0
Z/γ*+4jets	1025	996	70	0	0
Z/γ*+5jets	198	174	19	0	0
Z/γ*+6jets	35	26	4	0	0

5.6.1 Z/γ*($\rightarrow e^+e^-$) Inclusive Sample

In this section, basic kinematic distributions for electrons and Z candidates are compared. Figure 5.3 shows a comparison of di-electron invariant mass distribution between Data and MC. Figure 5.4 compares basic electron and Z kinematic distributions.

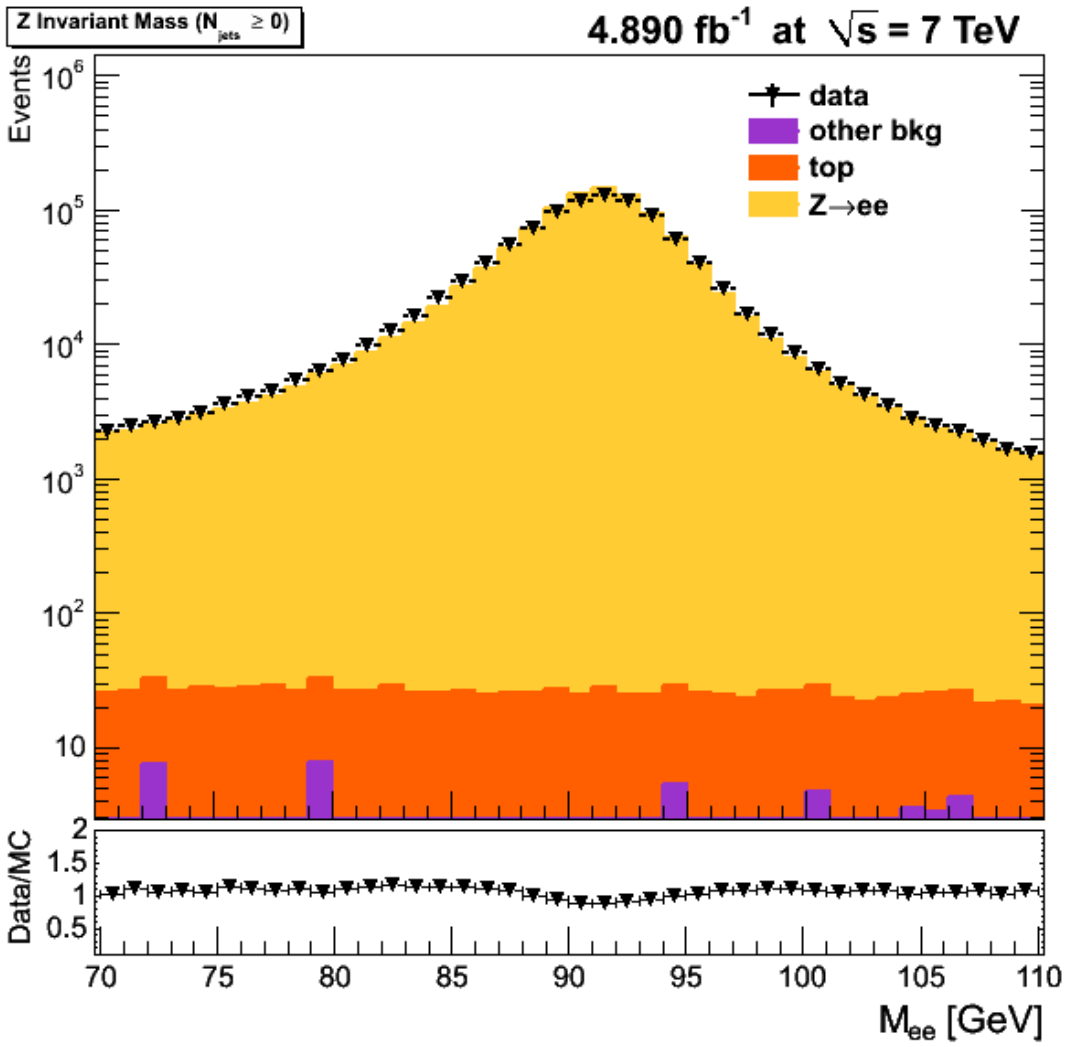
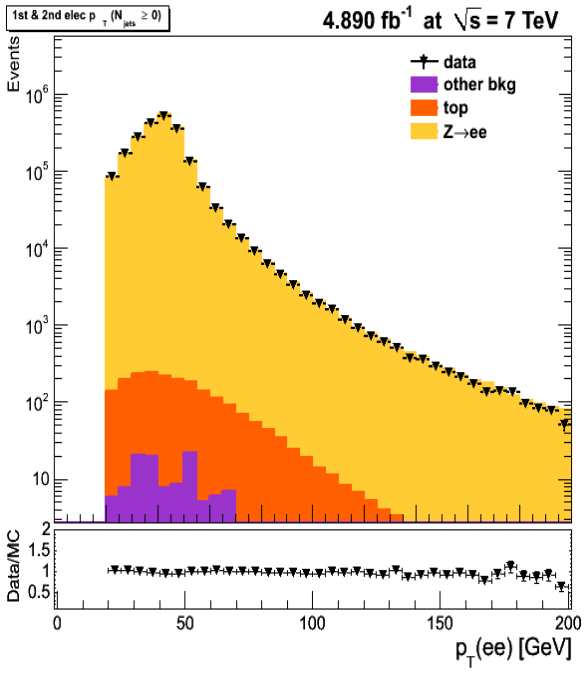


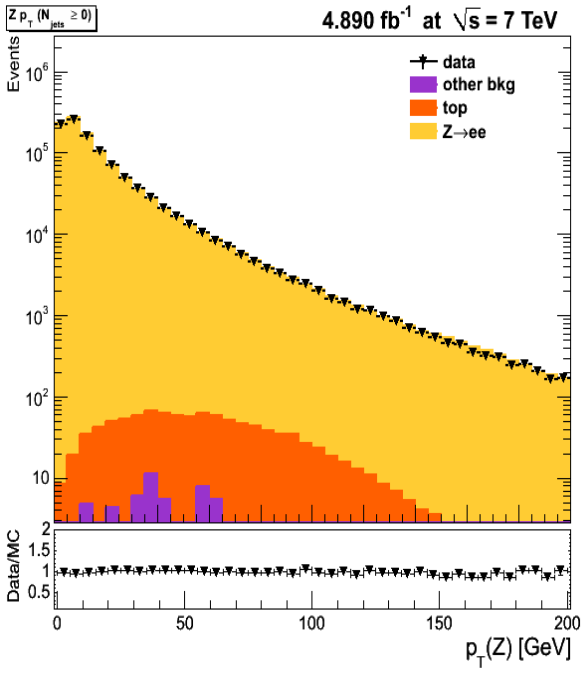
Figure 5.3: Di-electron invariant mass comparison for the $Z/\gamma*(\rightarrow e^+e^-)$ Inclusive sample in data and MC (MadGraph).

5.6.2 $Z/\gamma*(\rightarrow e^+e^-)+\geq 1$ Jet Sample

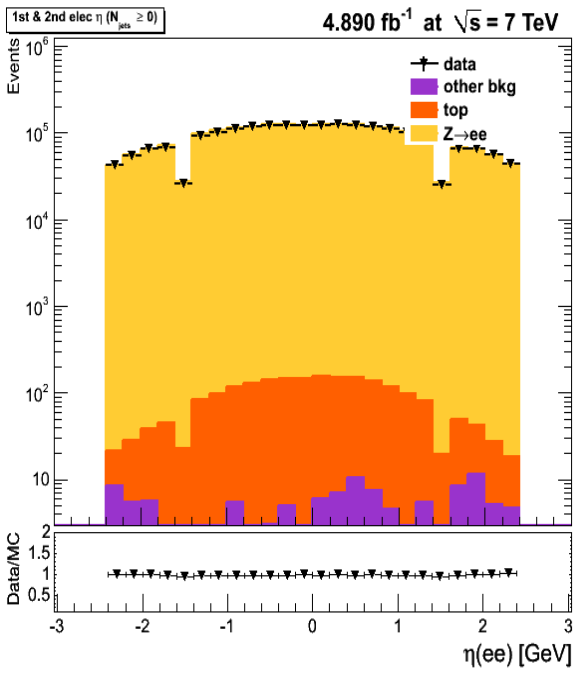
Figure 5.5 shows a comparison of the di-electron invariant mass peak between Data and MC. Figure 5.6 shows comparisons of basic electron and Z distributions. Comparisons of basic kinematic distributions for all and leading jets are shown in Figure 5.7 and Figure 5.8, respectively.



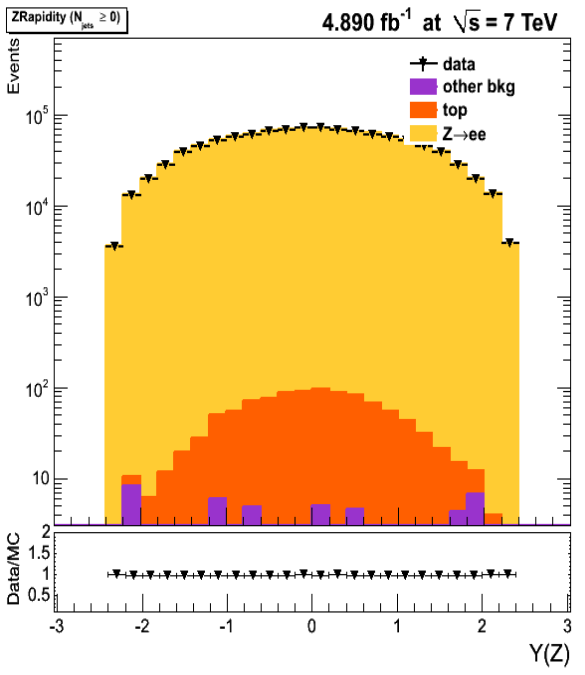
(a) p_T of Z daughters



(b) p_T of Z boson



(c) η of Z daughters



(d) Y of Z boson

Figure 5.4: p_T of both Z electrons (a), η of both Z electrons (b), Z p_T (c), Z rapidity (d) for the $Z/\gamma * (\rightarrow e^+e^-)$ Inclusive sample in data and MC (MadGraph)

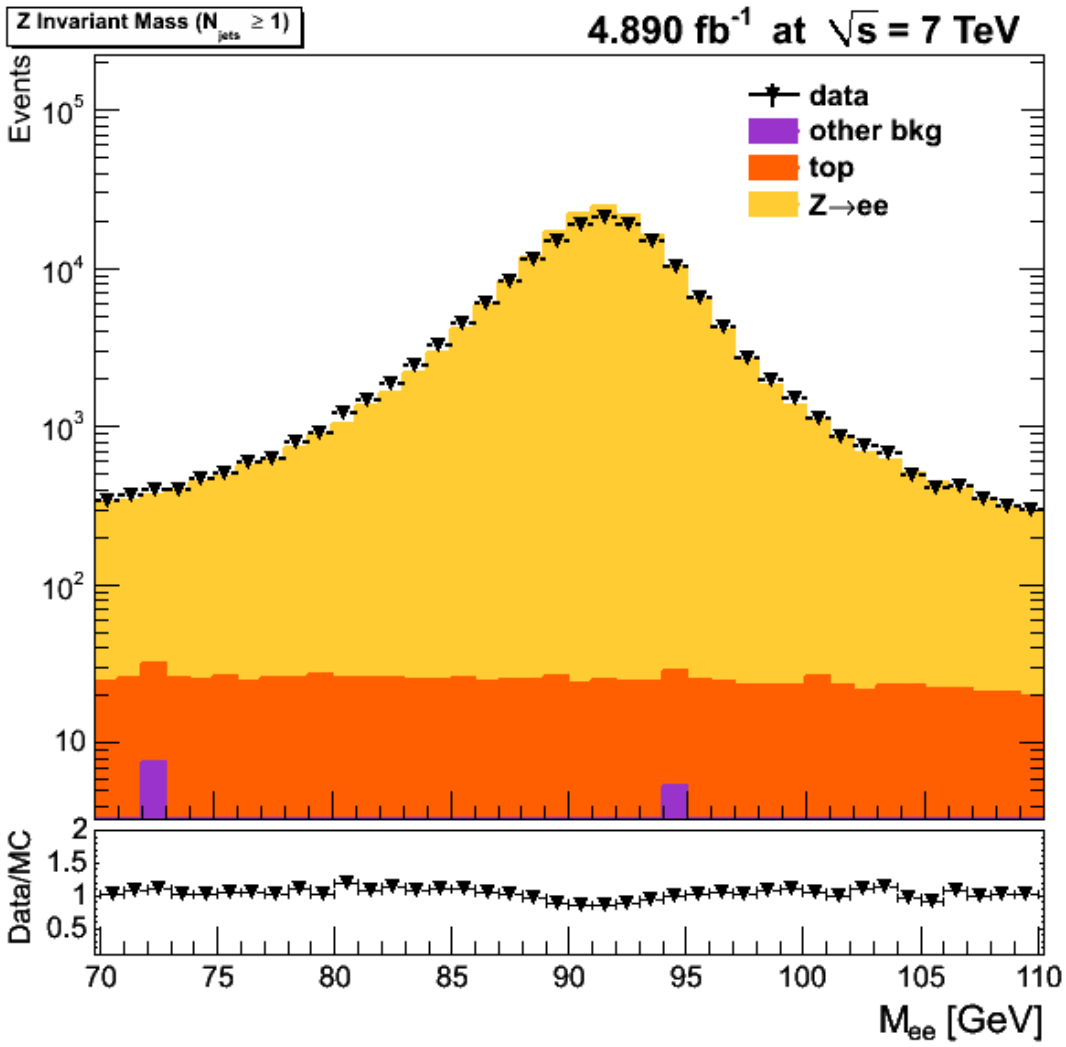
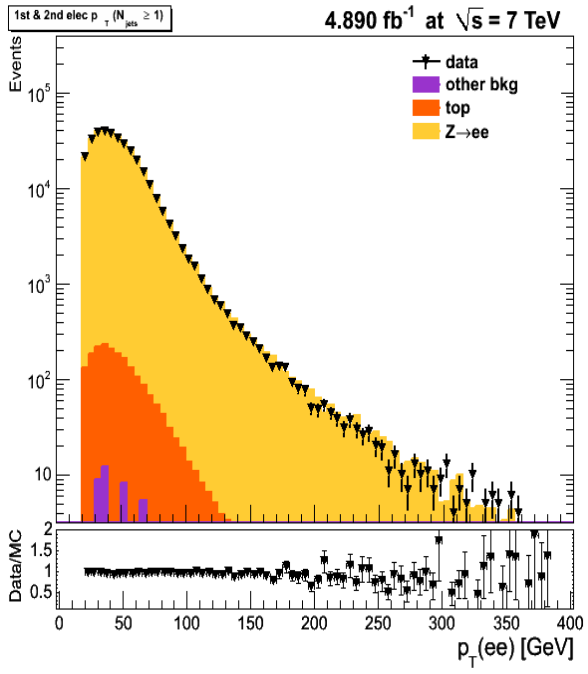


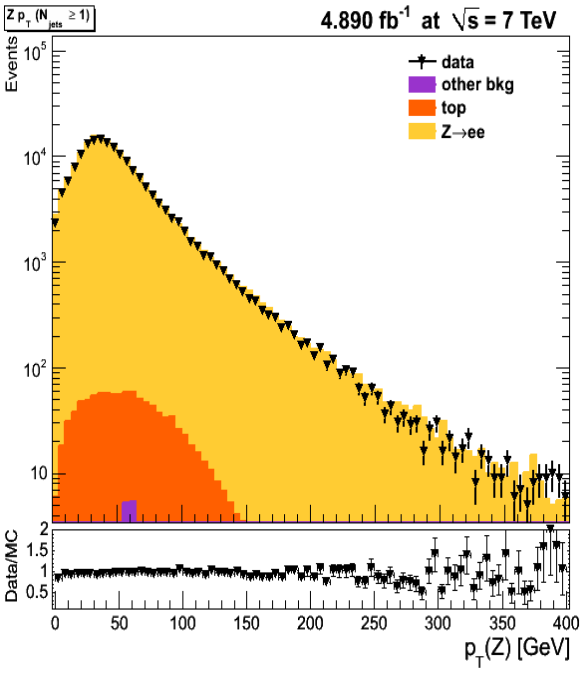
Figure 5.5: Di-electron invariant mass comparison for the Z/$\gamma*(\rightarrow e^+e^-)$+≥1 Jet sample in data and MC (MadGraph).

5.6.3 Z/$\gamma*(\rightarrow e^+e^-)$+≥2 Jet Sample

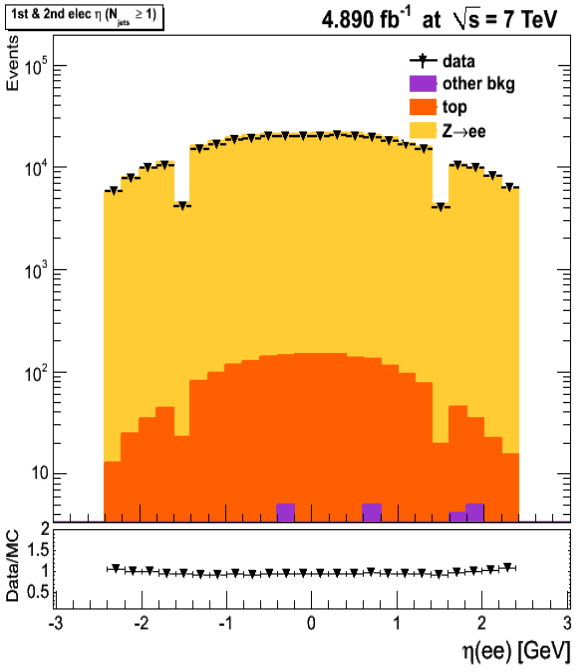
Figure 5.9 shows a comparison of the di-electron invariant mass peak between Data and MC and Figure 5.10 shows comparisons of basic electron and Z distributions. Comparisons of basic kinematic distributions for all and leading jets are shown in Figure 5.11 and Figure 5.12, respectively.



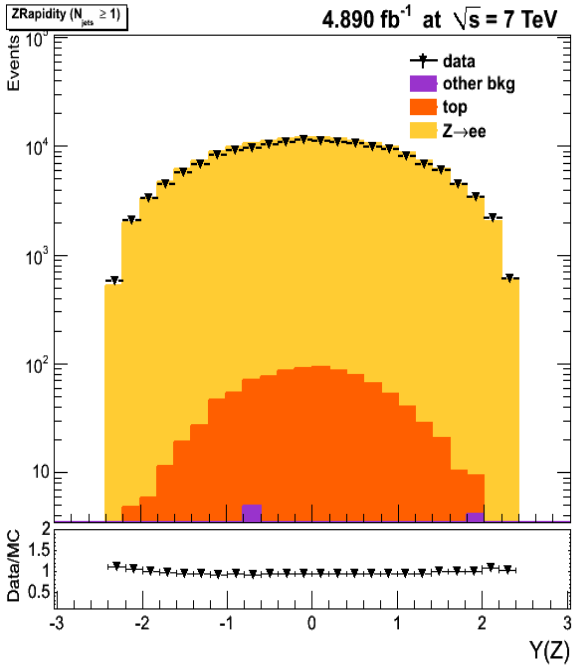
(a) p_T of Z daughters



(b) p_T of Z boson

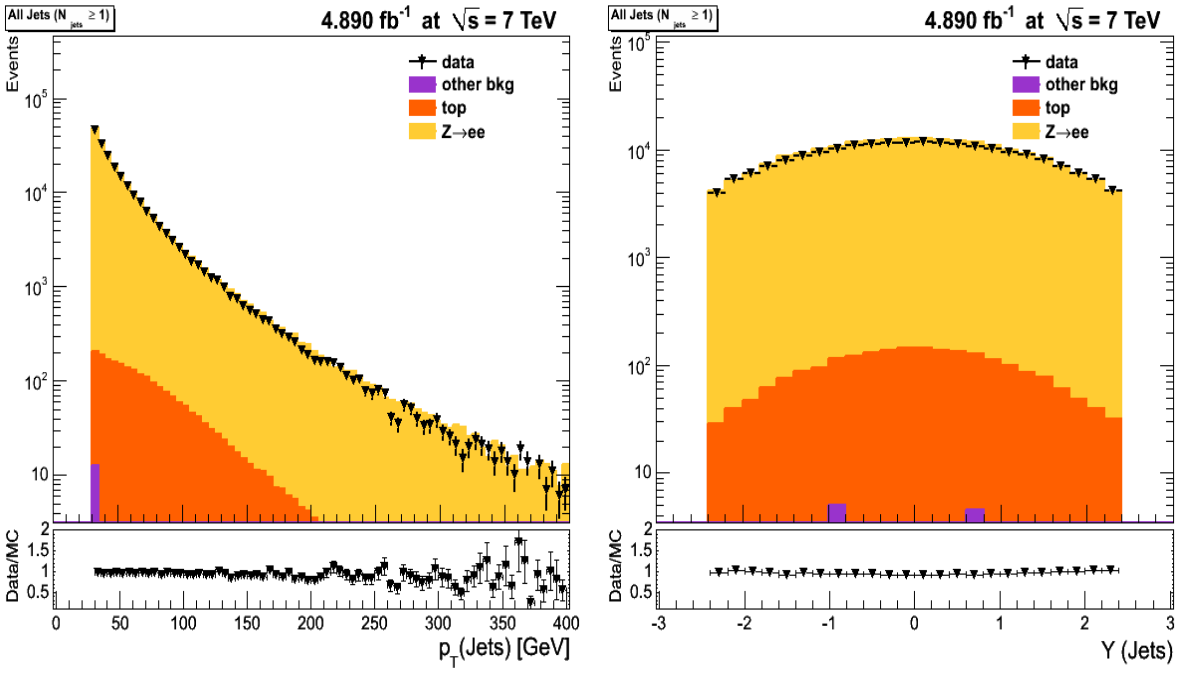


(c) η of Z daughters



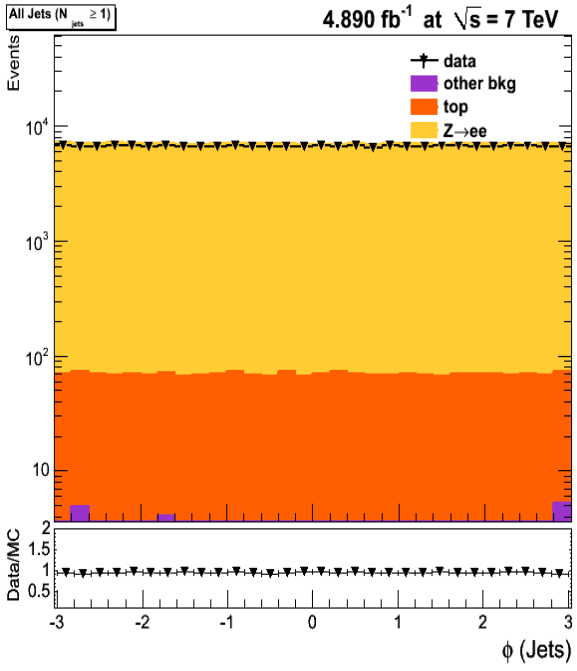
(d) Y of Z boson

Figure 5.6: p_T of both Z electrons (a), η of both Z electrons (b), Z p_T (c), Z rapidity (d) for the $Z/\gamma * (\rightarrow e^+e^-)$+$\geq$1 Jet sample in data and MC (MadGraph)



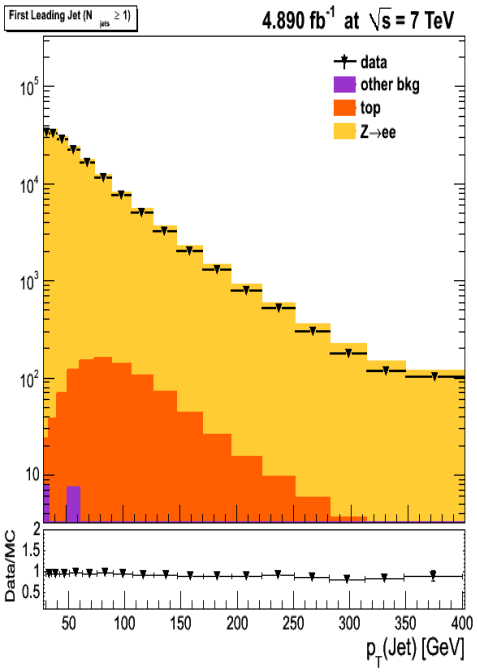
(a) p_T of all jets

(b) Y of all jets

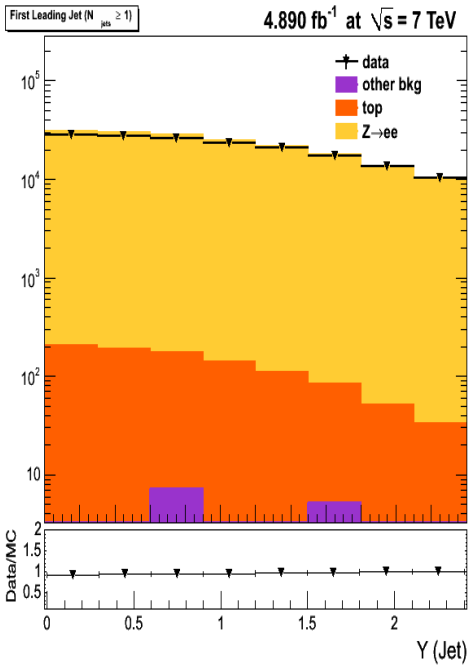


(c) ϕ of all jets

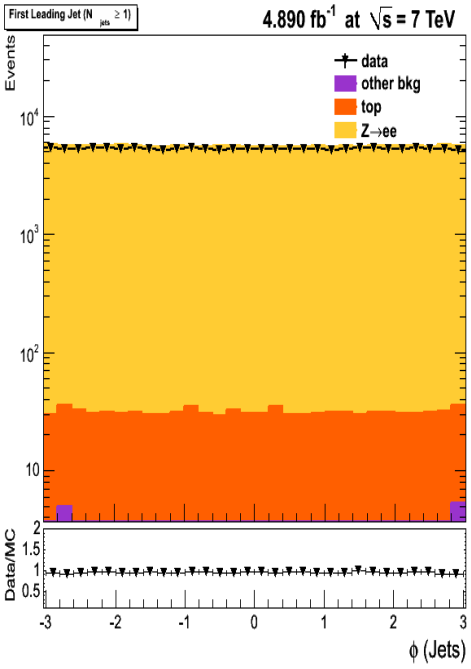
Figure 5.7: p_T(a), Y(b) and ϕ(c) of all jets for the Z/$\gamma*(\rightarrow e^+e^-)$+≥1 Jet sample in data and MC



(a) p_T of first leading jet



(b) Y of first leading jet



(c) ϕ of first leading jet

Figure 5.8: p_T(a), Y(b) and ϕ(c) of the first leading jet for the Z/$\gamma*(\rightarrow e^+e^-)$+≥1 Jet sample in data and MC

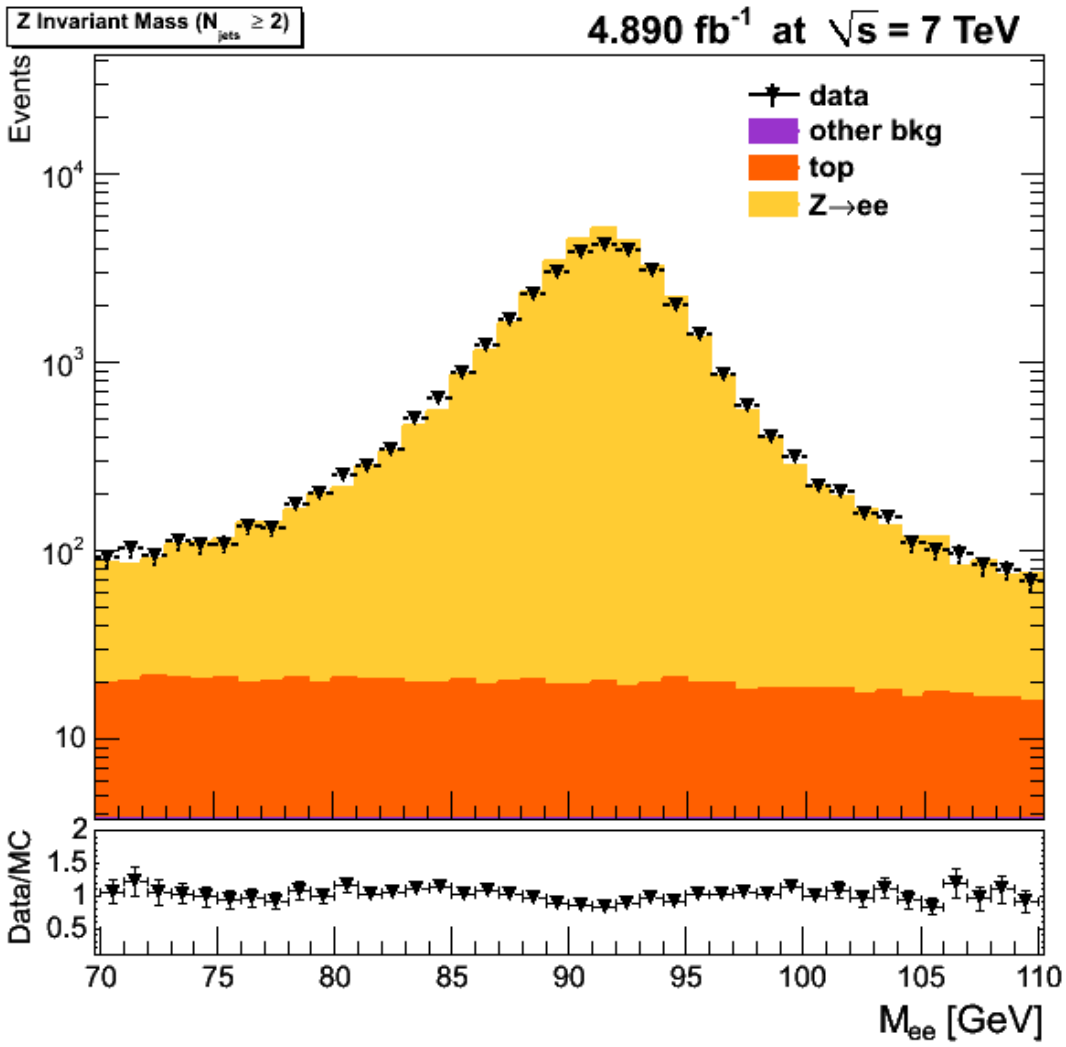
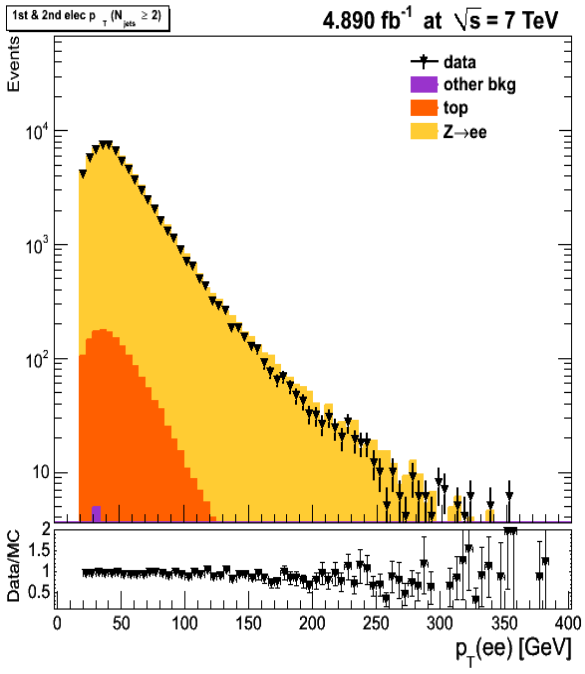


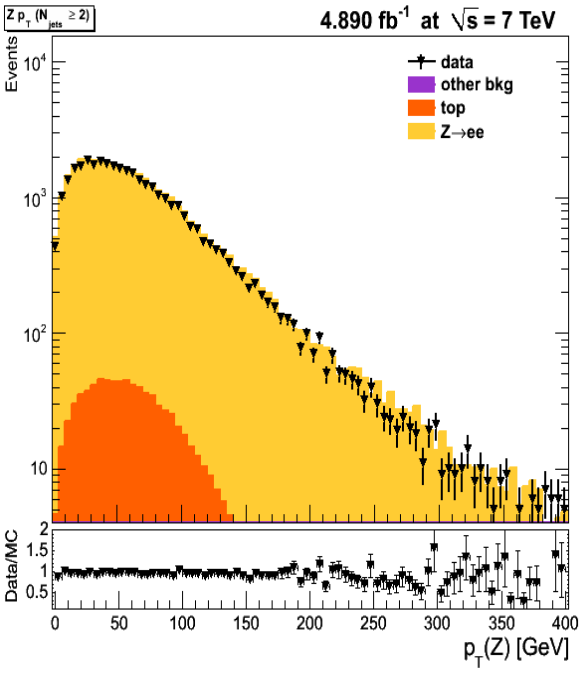
Figure 5.9: Di-electron invariant mass comparison for the Z/$\gamma*(\rightarrow e^+e^-)$+≥2 Jet sample in data and MC (MadGraph).

5.6.4 Z/$\gamma*(\rightarrow e^+e^-)$+≥3 Jet Sample

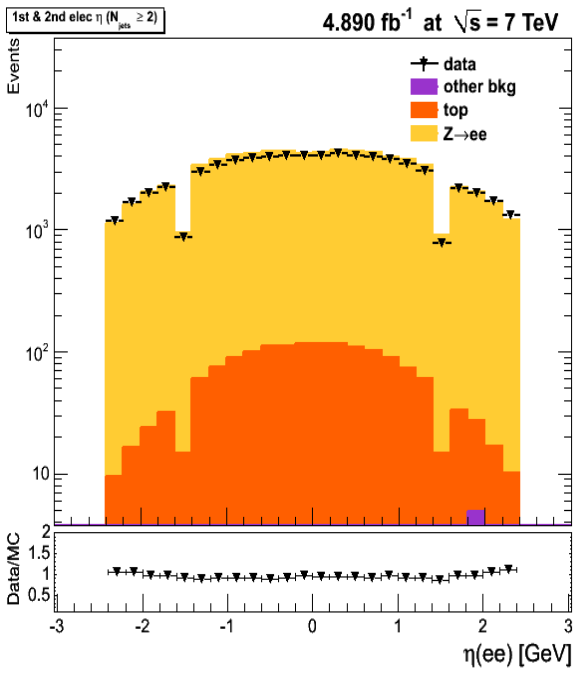
Figure 5.13 shows a comparison of the di-electron invariant mass peak between Data and MC. Figure 5.14 shows comparisons of basic electron and Z distributions. Comparisons of basic kinematic distributions for all and leading jets are shown in Figure 5.15 and Figure 5.16, respectively.



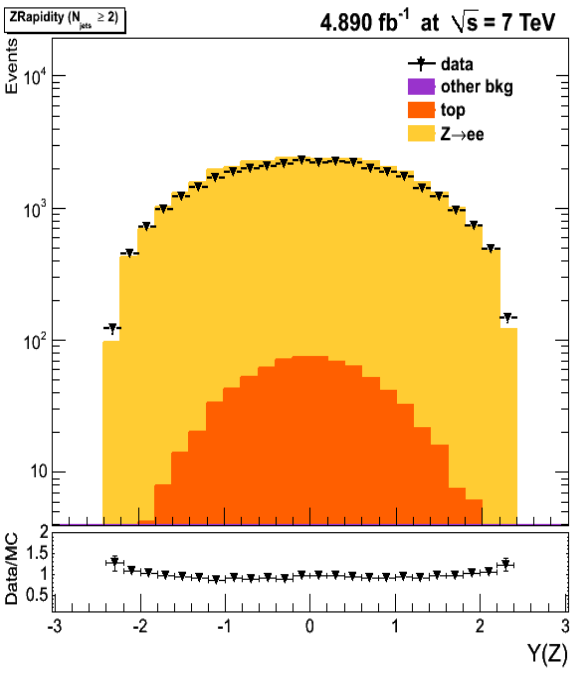
(a) p_T of Z daughters



(b) p_T of Z boson

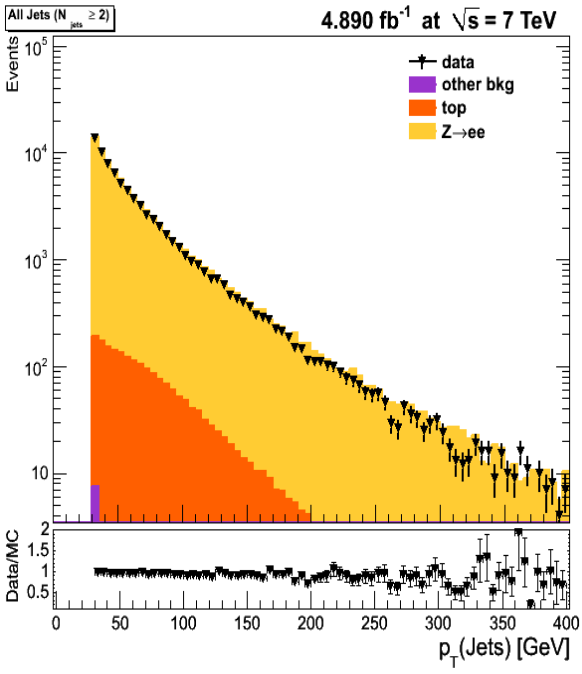


(c) η of Z daughters

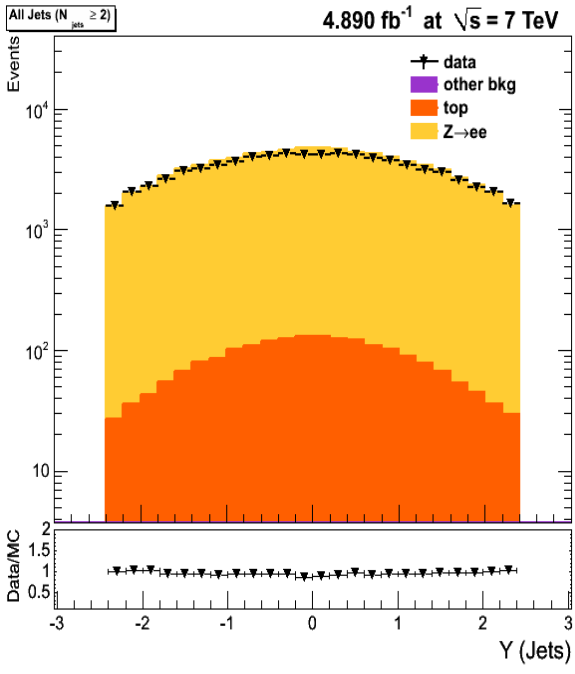


(d) Y of Z boson

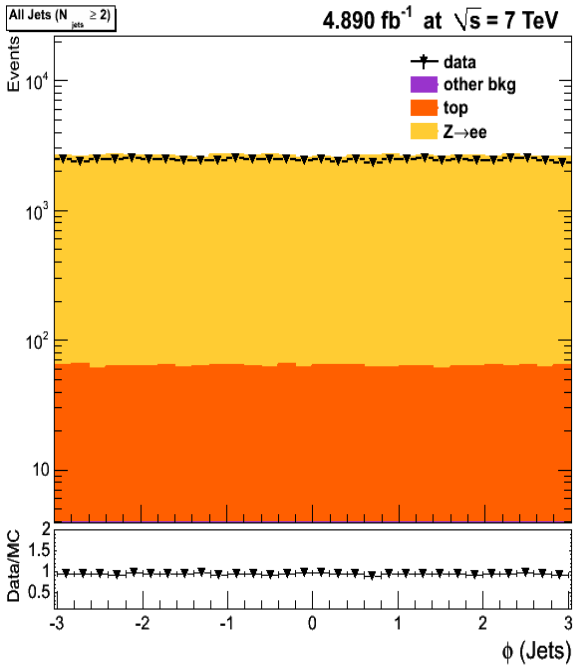
Figure 5.10: p_T of both Z electrons (a), η of both Z electrons (b), Z p_T (c), Z rapidity (d) for the $Z/\gamma * (\rightarrow e^+e^-)+\geq 2$ Jet sample in data and MC (MadGraph)



(a) p_T of all jets

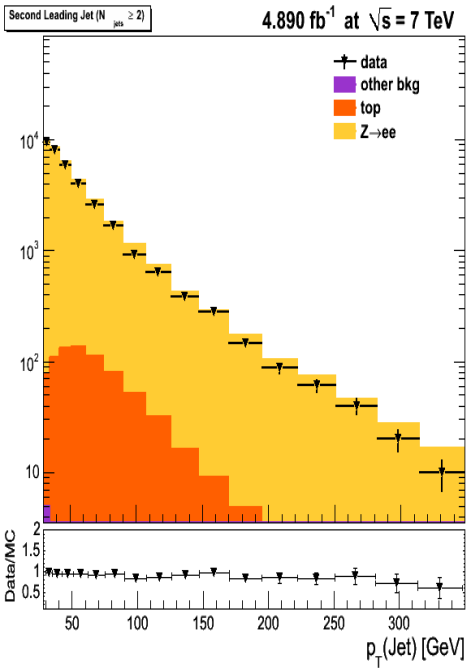


(b) Y of all jets

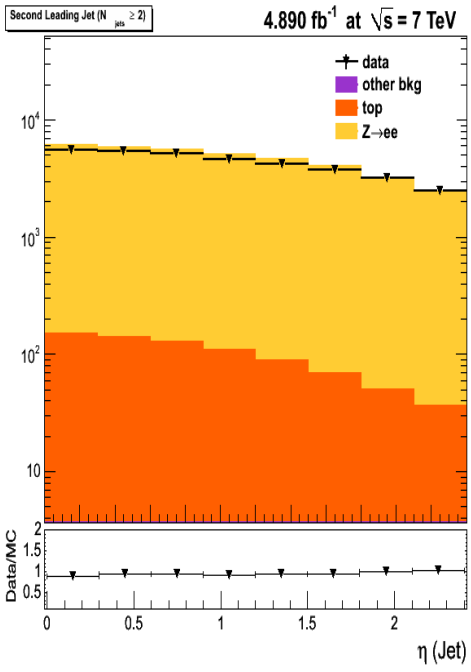


(c) ϕ of all jets

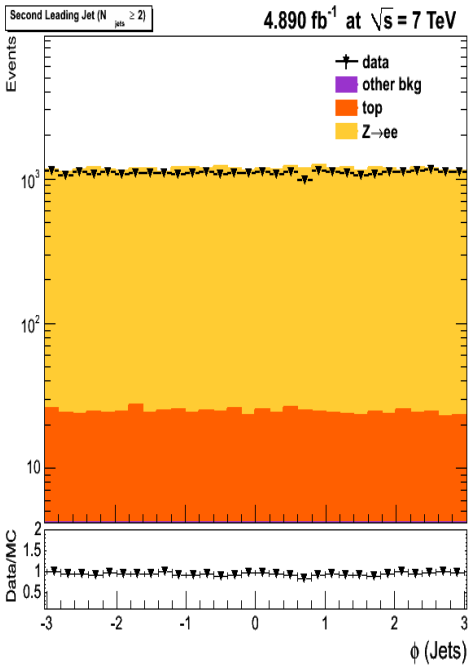
Figure 5.11: p_T(a), Y(b) and ϕ(c) of all jets for the Z/$\gamma*(\rightarrow e^+e^-)$+$\geq$2 Jet sample in data and MC



(a) p_T of first leading jet



(b) Y of first leading jet



(c) ϕ of first leading jet

Figure 5.12: p_T(a), Y(b) and ϕ(c) of the second leading jet for the Z/$\gamma*(\rightarrow e^+e^-)$+$\geq$2 Jet sample in data and MC

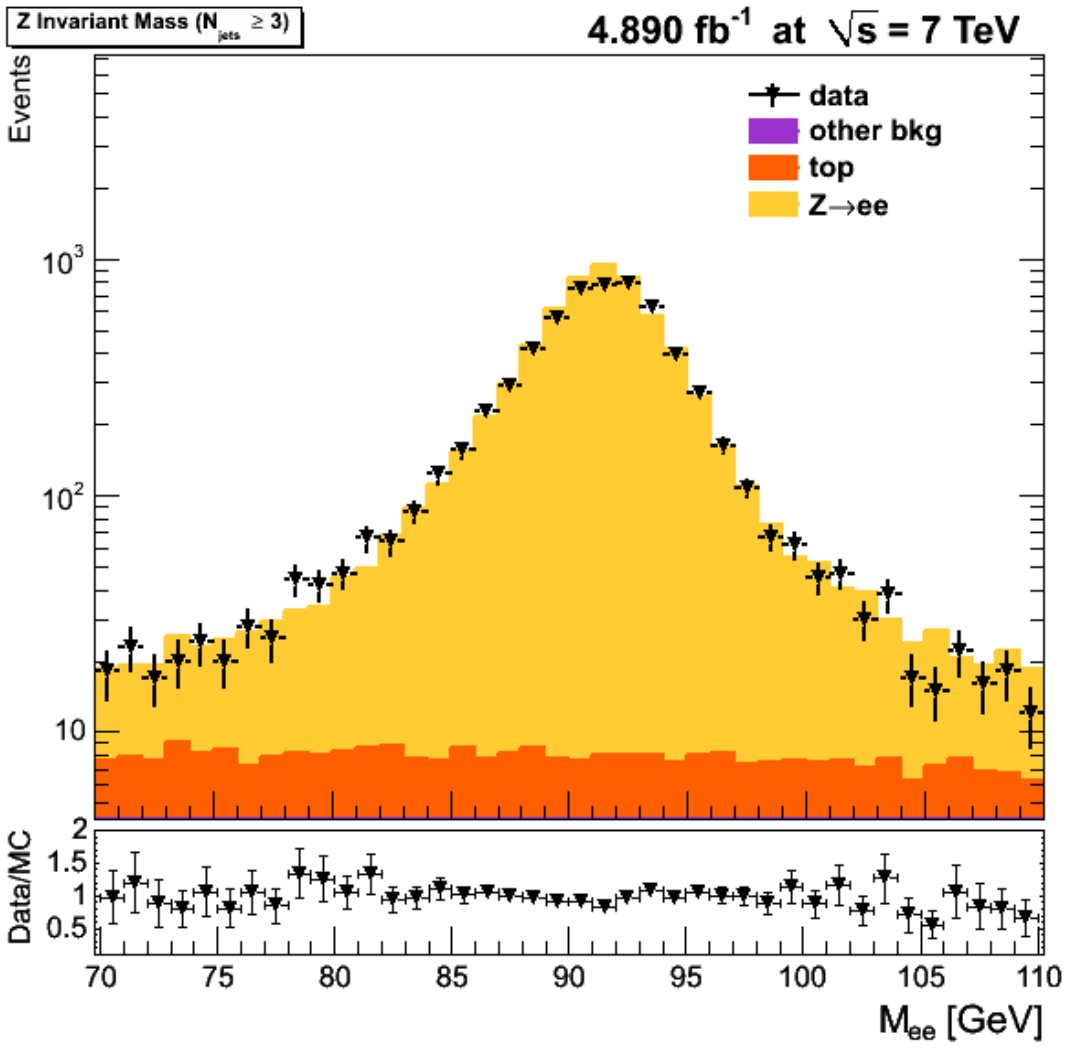
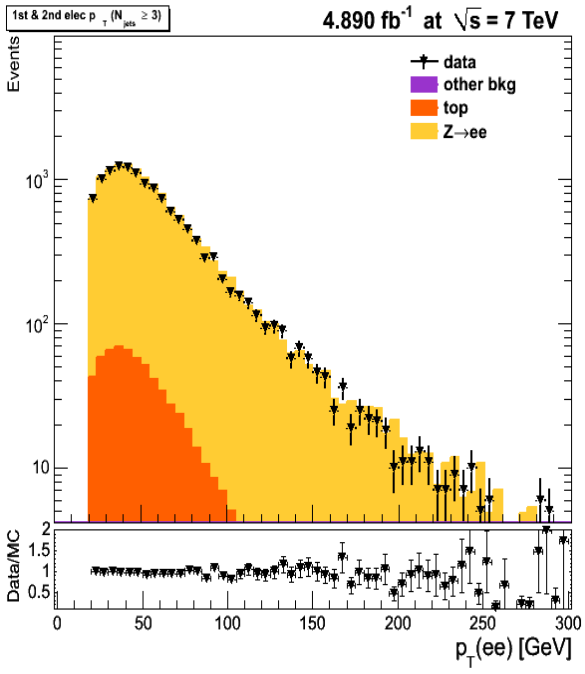


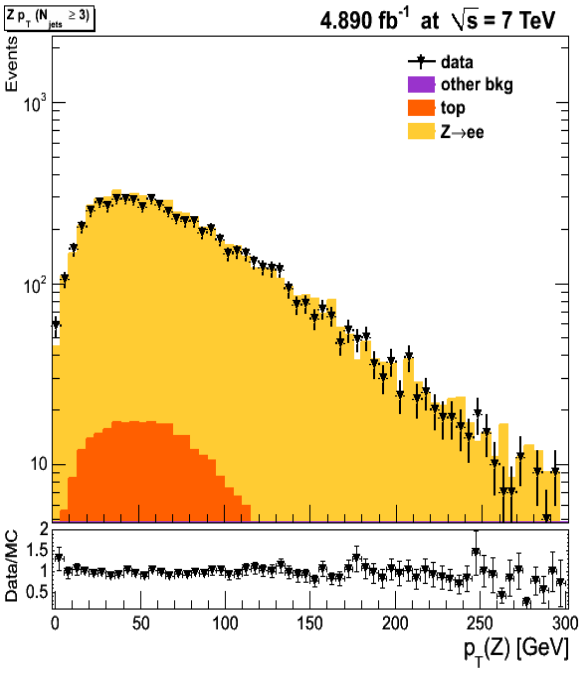
Figure 5.13: Di-electron invariant mass comparison for the Z/$\gamma*(\rightarrow e^+e^-)$+$\geq$3 Jet sample in data and MC (MadGraph).

5.6.5 Z/$\gamma*(\rightarrow e^+e^-)$+$\geq$4 Jet Sample

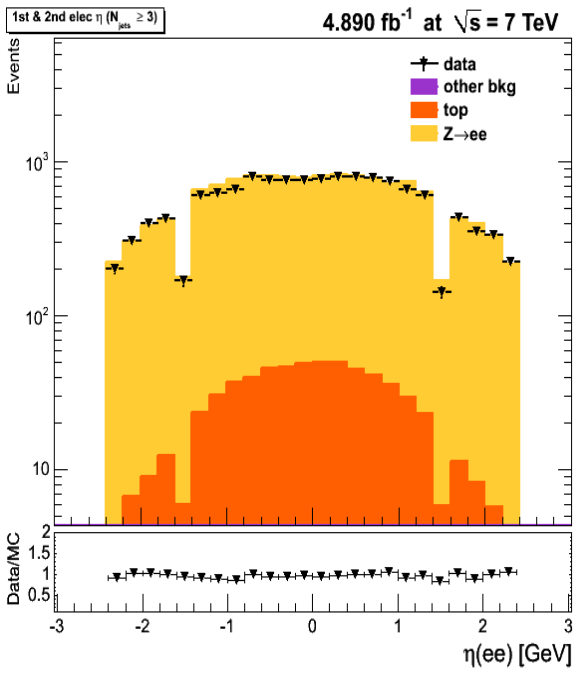
Figure 5.17 shows a comparison of the di-electron invariant mass peak between Data and MC. Figure 5.18 shows comparisons of basic electron and Z distributions. Comparisons of basic kinematic distributions for all and leading jets are shown in Figure 5.19 and Figure 5.20, respectively.



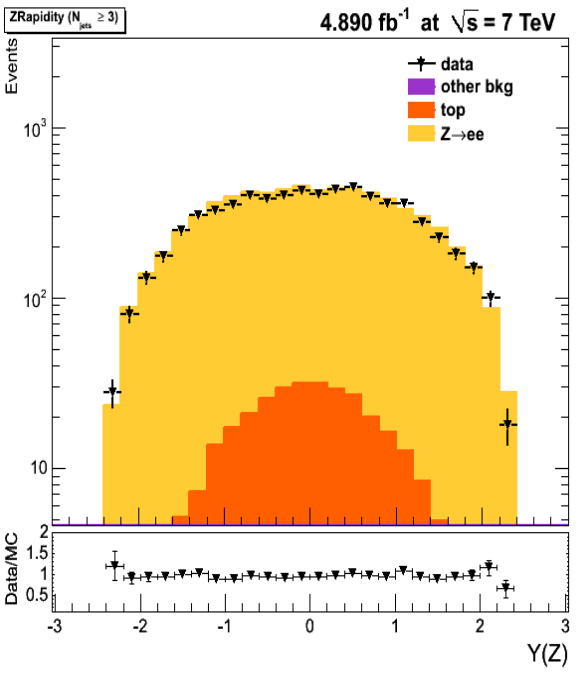
(a) p_T of Z daughters



(b) p_T of Z boson

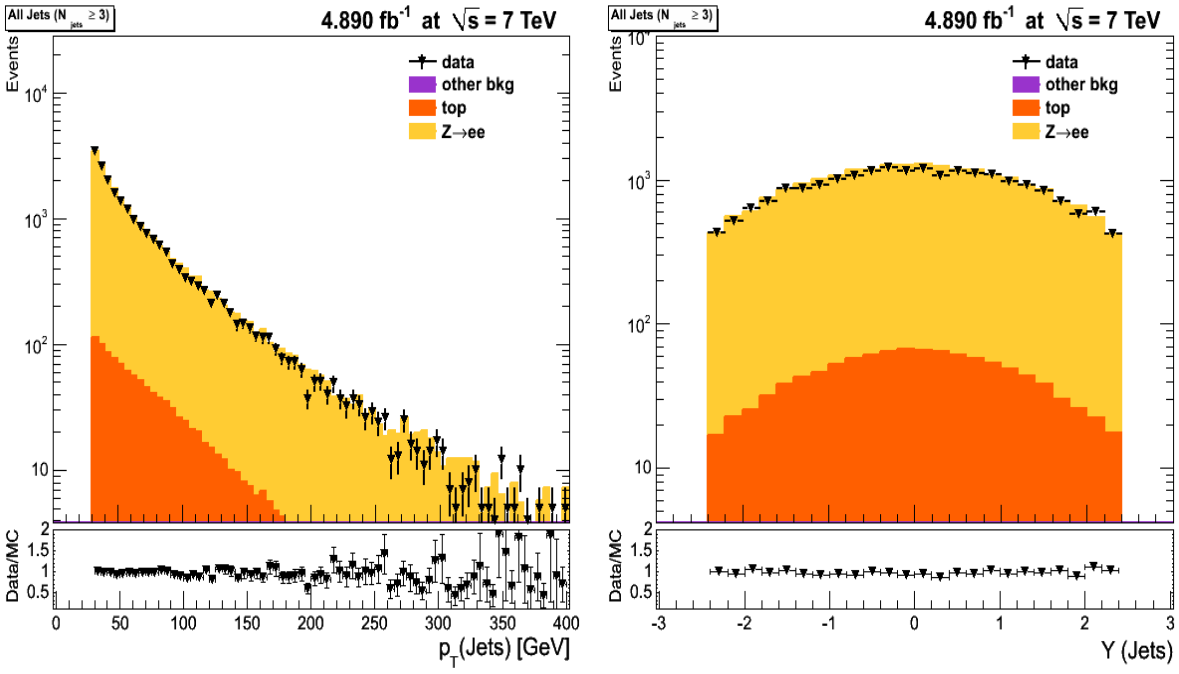


(c) η of Z daughters



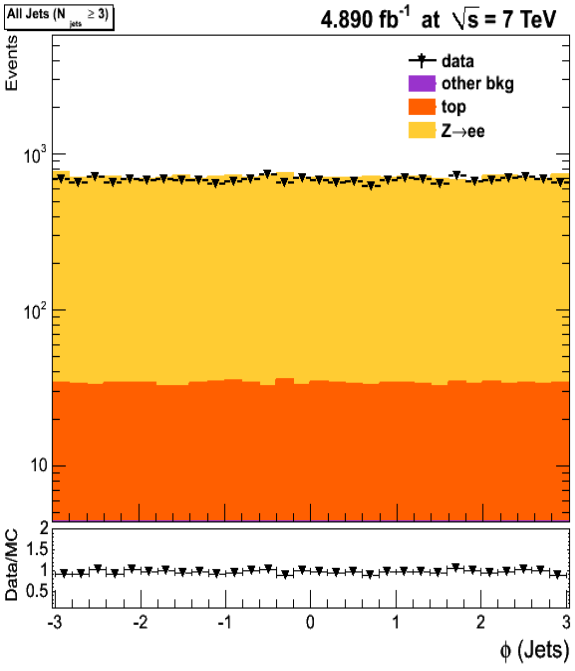
(d) Y of Z boson

Figure 5.14: p_T of both Z electrons (a), η of both Z electrons (b), Z p_T (c), Z rapidity (d) for the Z/$\gamma*(\rightarrow e^+e^-)$+$\geq$3 Jet sample in data and MC (MadGraph)



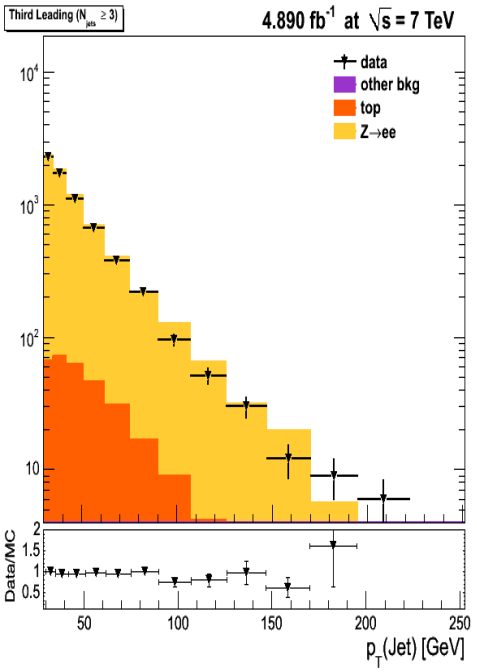
(a) p_T of all jets

(b) Y of all jets

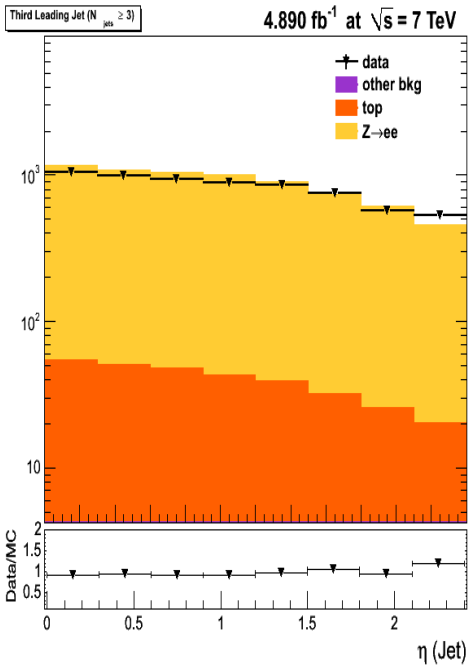


(c) ϕ of all jets

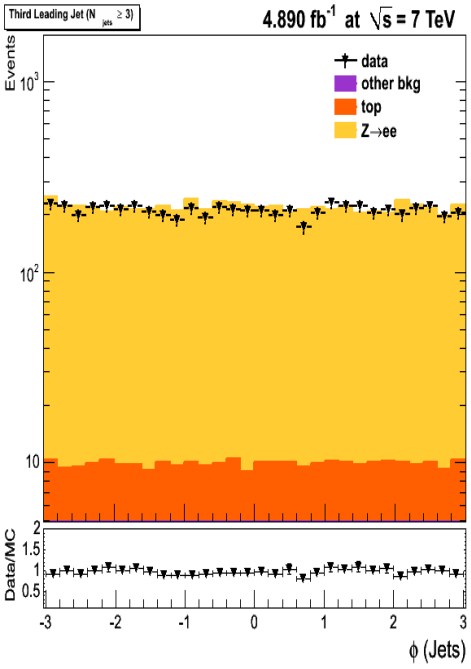
Figure 5.15: p_T(a), Y(b) and ϕ(c) of all jets for the Z/$\gamma*(\rightarrow e^+e^-)$+$\geq$3 Jet sample in data and MC



(a) p_T of first leading jet



(b) Y of first leading jet



(c) ϕ of first leading jet

Figure 5.16: p_T(a), Y(b) and ϕ(c) of the second leading jet for the Z/$\gamma*(\rightarrow e^+e^-)$+≥3 Jet sample in data and MC

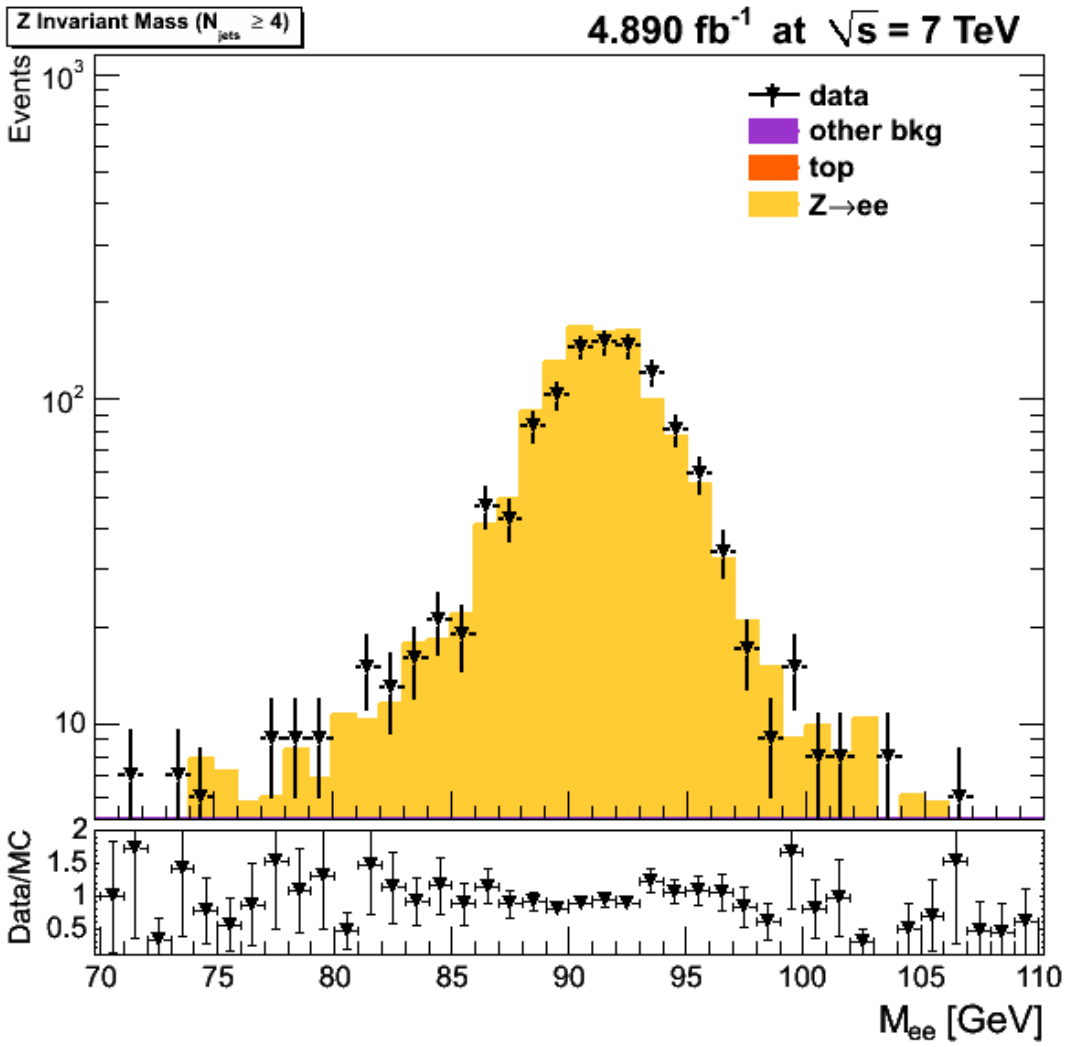
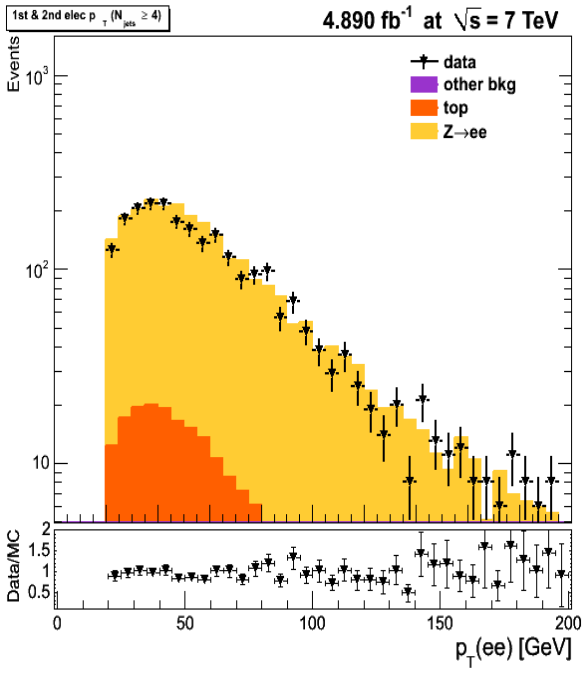


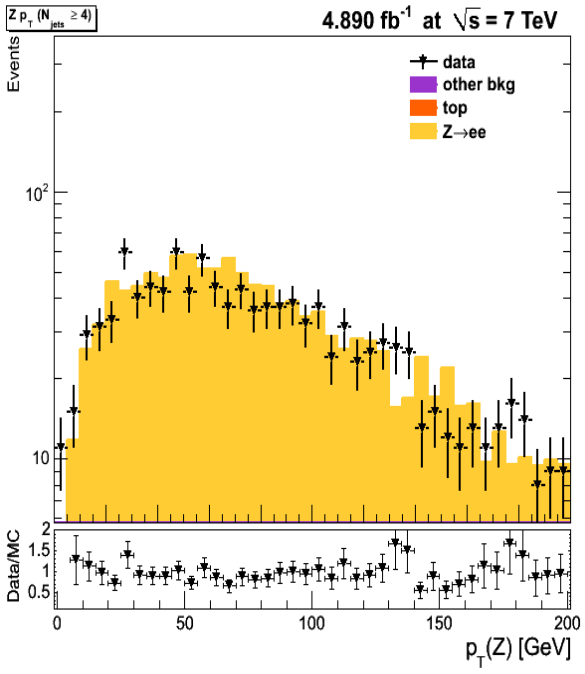
Figure 5.17: Di-electron invariant mass comparison for the Z/$\gamma*(\rightarrow e^+e^-)$+≥4 Jet sample in data and MC (MadGraph).

5.6.6 Z/$\gamma*(\rightarrow e^+e^-)$+≥5,6 Jet Sample

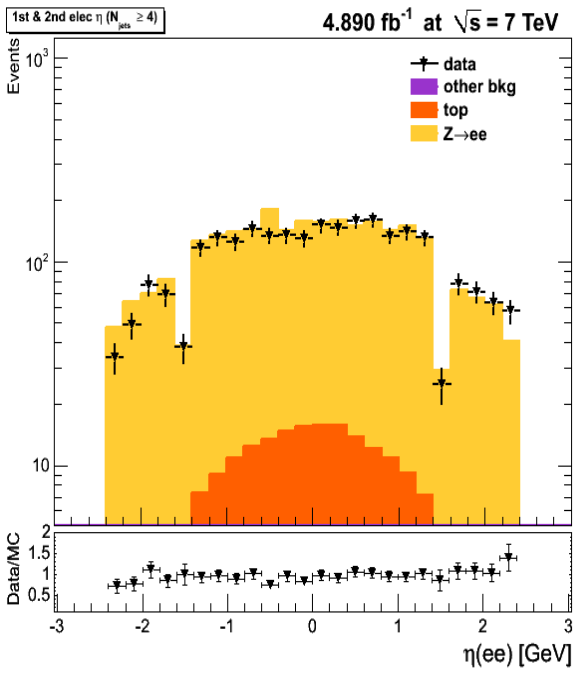
Invariant mass distribution of Z boson for Z/$\gamma*(\rightarrow e^+e^-)$+≥5 and 6 jet events is shown in Figure 5.21. Leading jet kinematic distributions for Z/$\gamma*(\rightarrow e^+e^-)$+≥5 and 6 jet events is shown in Figure 5.22.



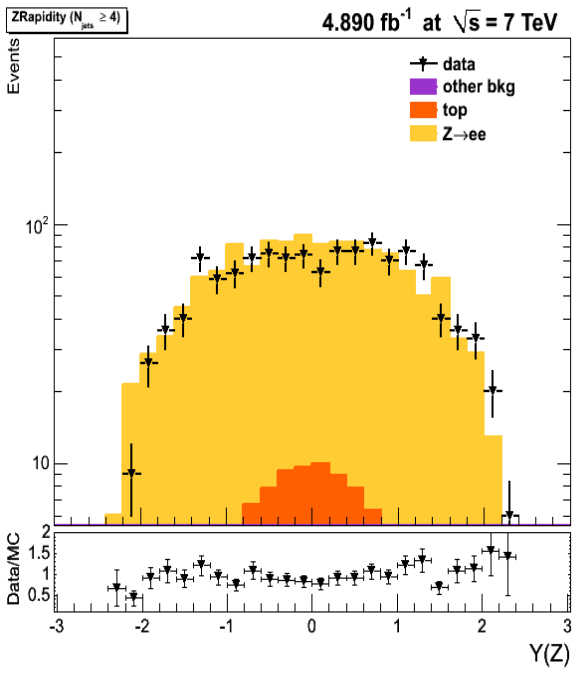
(a) p_T of Z daughters



(b) p_T of Z boson

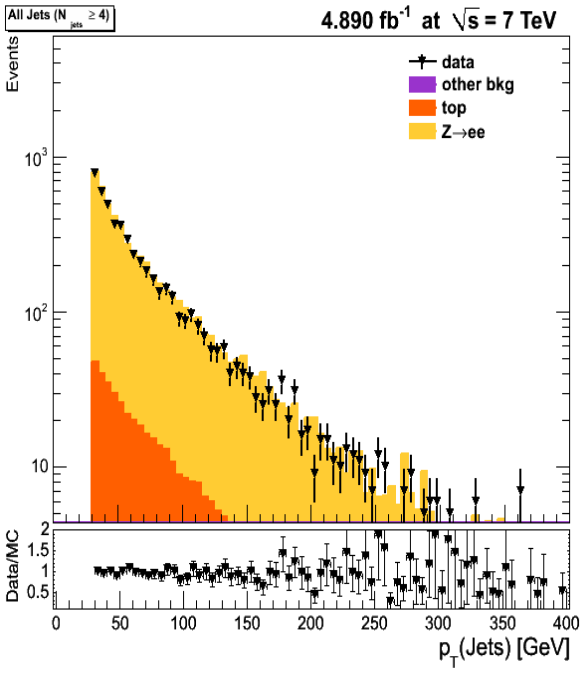


(c) η of Z daughters

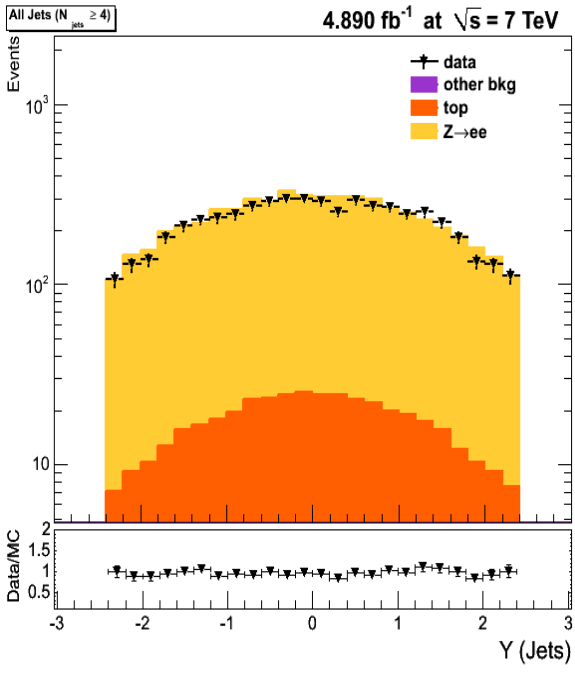


(d) Y of Z boson

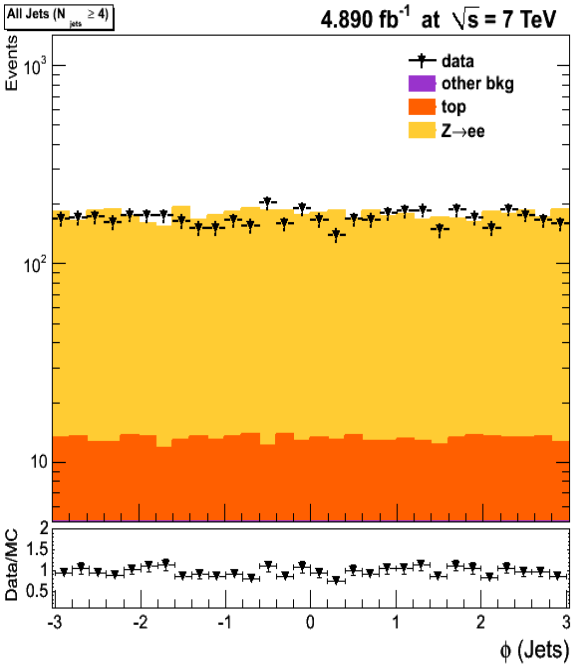
Figure 5.18: p_T of both Z electrons (a), η of both Z electrons (b), Z p_T (c), Z rapidity (d) for the $Z/\gamma * (\rightarrow e^+e^-)$+$\geq$4 Jet sample in data and MC (MadGraph)



(a) p_T of all jets

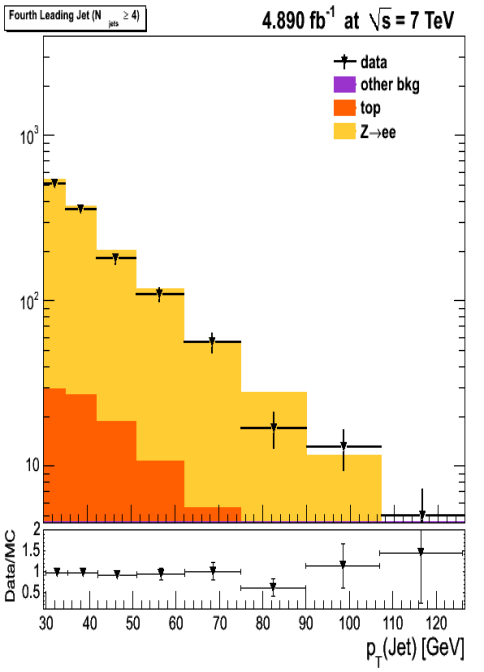


(b) Y of all jets

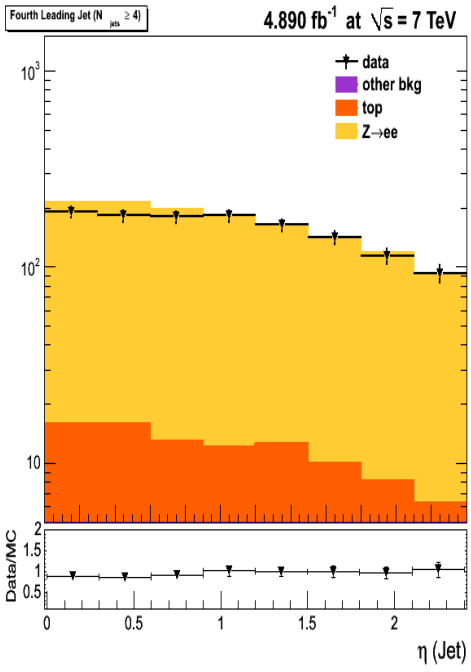


(c) ϕ of all jets

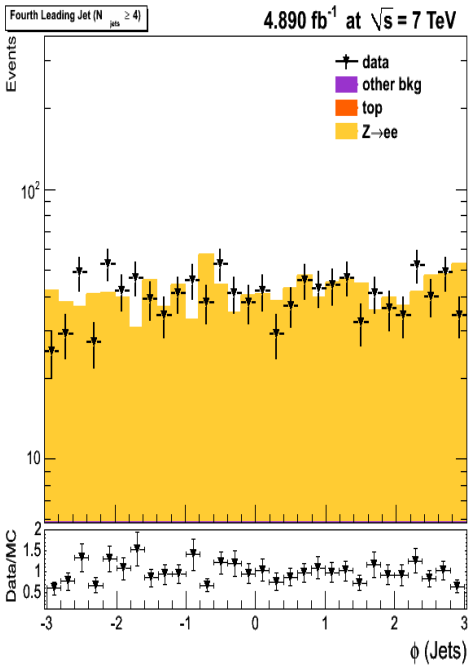
Figure 5.19: p_T(a), Y(b) and ϕ(c) of all jets for the Z/$\gamma*(\rightarrow e^+e^-)$+$\geq$4 Jet sample in data and MC



(a) p_T of first leading jet

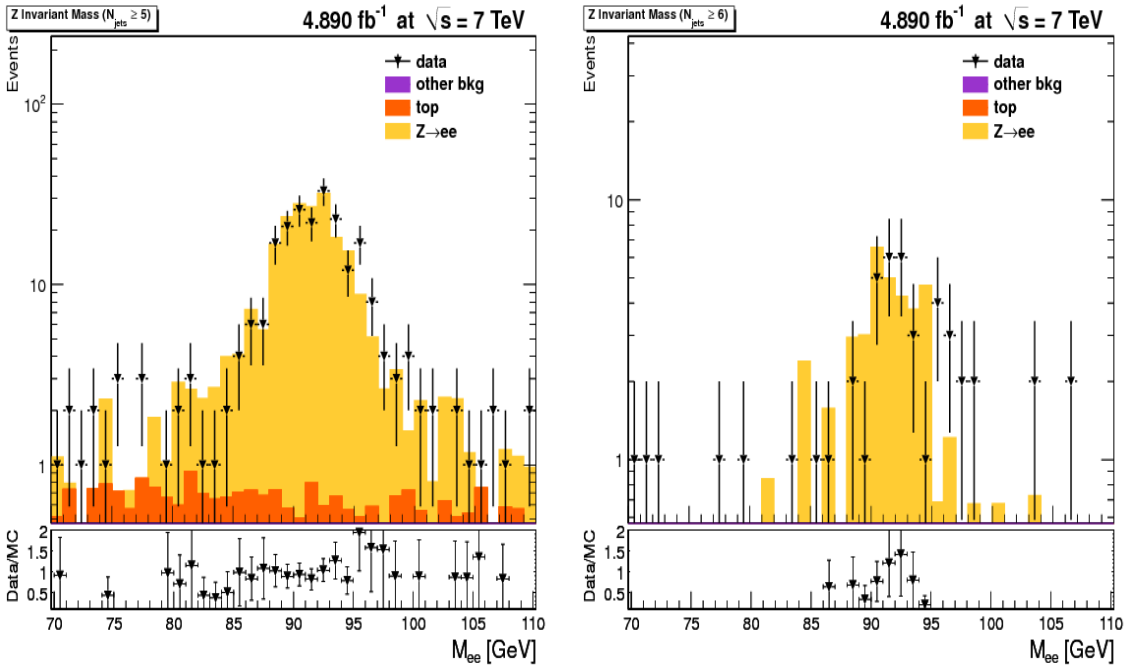


(b) Y of first leading jet



(c) ϕ of first leading jet

Figure 5.20: p_T(a), Y(b) and ϕ(c) of the second leading jet for the Z/$\gamma*(\rightarrow e^+e^-)$+≥4 Jet sample in data and MC



(a) $Z/\gamma*(\rightarrow e^+e^-)+\geq 5$ Jets

(b) $Z/\gamma*(\rightarrow e^+e^-)+\geq 6$ Jets

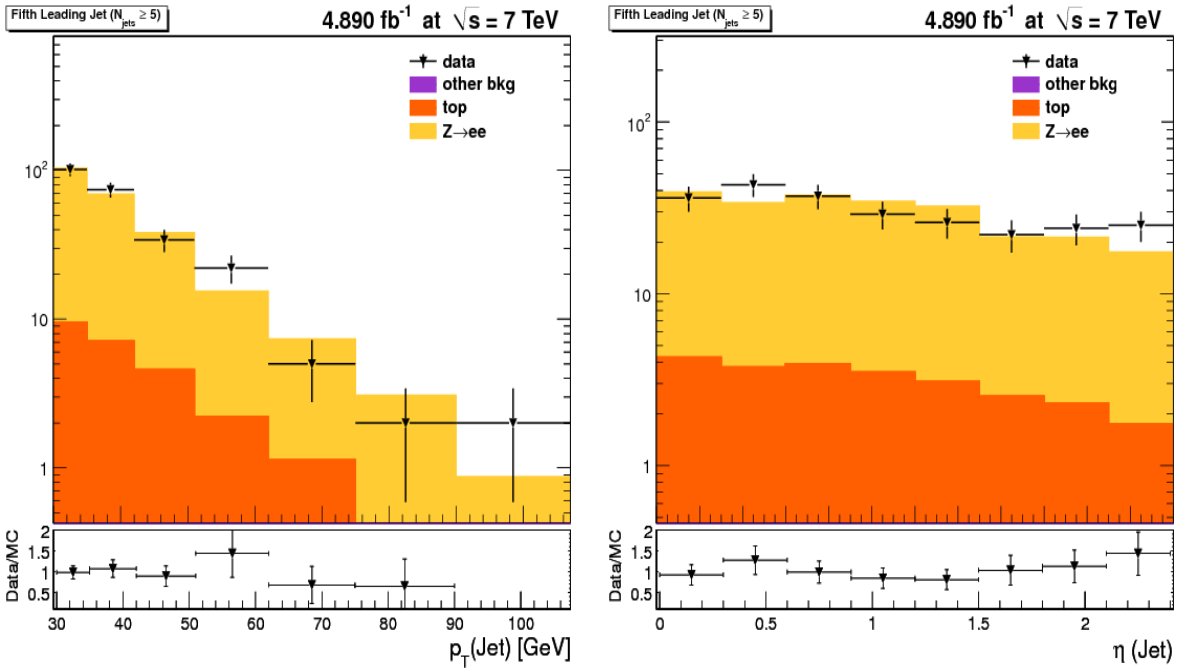
Figure 5.21: Dielectron invariant mass for the $Z/\gamma*(\rightarrow e^+e^-)+\geq 5$ Jet sample in data and MC(a) and for the $Z/\gamma*(\rightarrow e^+e^-)+\geq 6$ Jet sample in data and MC(b)

5.7 Event Selection Efficiency

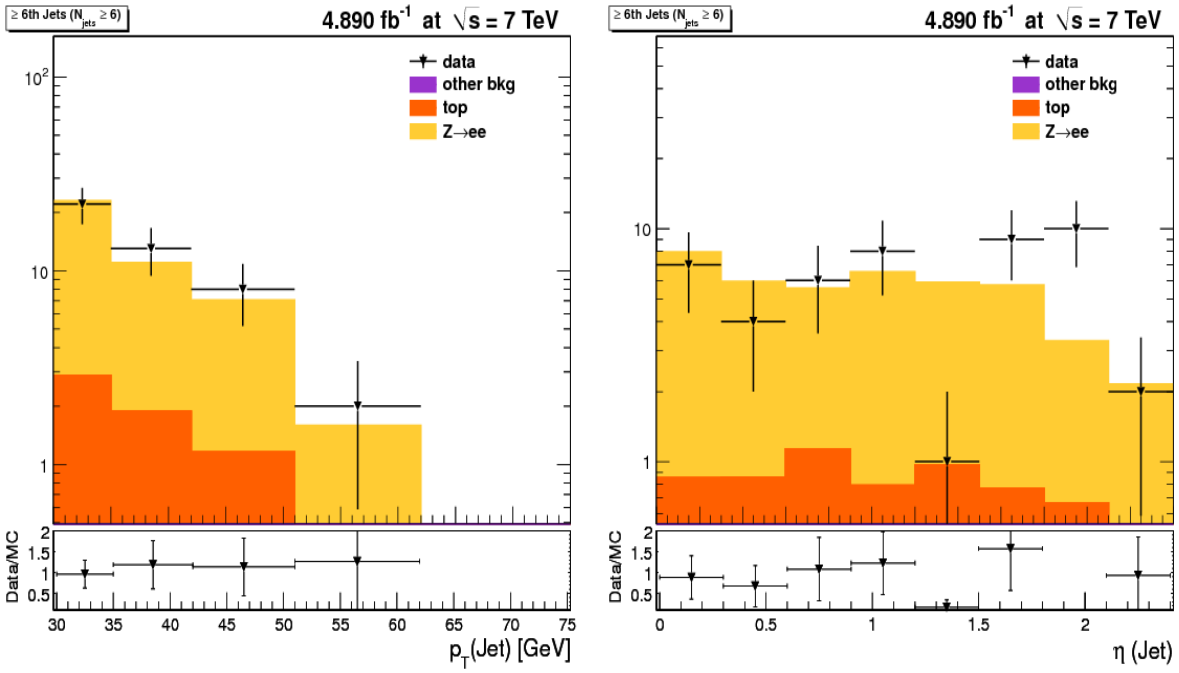
In this section, the methodologies to estimate the electron reconstruction, selection and trigger efficiencies are discussed and their results are presented.

5.7.1 Tag and Probe method

The single electron efficiencies for offline and online selection criteria have been obtained in a data driven way using the Tag and Probe method [126]. This method allows the determination of relative efficiency of a selection (S2) with respect to a loose one (S1) such that S2$\subset$S1. The electrons belonging to the loose selection are called **probes** and are further classified as **passing probes** (S1$\cap$S2) and **failing probes** ($\complement$ (S1$\cap$S2)) depending on whether they make it to tighter selection or not. A third selection usually tighter than passing probes, but **not** exclusive with probe collection, is used to define **tag** electrons.



(a) $Z/\gamma*(\rightarrow e^+e^-)+\geq 5$ Jets



(b) $Z/\gamma*(\rightarrow e^+e^-)+\geq 6$ Jets

Figure 5.22: Fifth Leading jet p_T and Y for the $Z/\gamma*(\rightarrow e^+e^-)+\geq 5$ Jet sample in data and MC(a) , Sixth Leading jet p_T and Y for the $Z/\gamma*(\rightarrow e^+e^-)+\geq 6$ Jet sample in data and MC(b)

Then dielectron candidates are formed by pairing tag and probes, and requiring their invariant mass to be consistent with the Z invariant mass peak (i.e. within 60-120 GeV mass window). At this stage any background component is subtracted via a fit to invariant mass distribution of the tag and probe pairs. This helps to determine the net number of passing and failing di-electron probes which are then used to determine the efficiency of going from selection S1 to selection S2. Since the tag and probe collections are not mutually exclusive, a probe can also act as a tag if it passes the selection requirements and vice-versa. Based on this the events may be classified as TT, TP and TF where T indicates a tag electron, P a passing probe and F indicates a failing probe. The single electron efficiency is determined as follows:

$$\epsilon(S2|S1) = \frac{2N_{TT} + N_{TP}}{2N_{TT} + N_{TP} + N_{TF}} \tag{5.1}$$

where N_{TT}, N_{TP}, and N_{TF} are the number of events observed for each type. The factor of 2 is used in order to avoid double counting of the events where both the electrons pass the tag criteria. Note that in this method, the probe sample is dependent on tag definition which means that *"probe region"* of the phase space may not be exactly same as the one occupied by leptons from Z decays in general. And if the efficiencies differ significantly between two regions, the above calculation may suffer an inherent bias. The effect of such biases is expected to be similar in Data and Monte Carlo and hence can be cancelled out by using ratio (ρ) between the efficiency in data to the efficiency in the MC, both computed using the above method.

$$\rho = \frac{\epsilon_{TnP-Data}}{\epsilon_{TnP-MC}} \tag{5.2}$$

So that,

$$\epsilon_{True-Data} = \rho * \epsilon_{True-MC} \tag{5.3}$$

Once the single electron efficiencies are obtained, they are used to evaluate the event efficiencies in our dielectron selection. Efficiency for an event Z($e^{p_T^1,\eta^1}, e^{p_T^2,\eta^2}$,) can be

factorized into online (ϵ_{HLT}) and offline ($\epsilon_{offline}$) components:

$$\epsilon_{ee} = \epsilon_{offline} * \epsilon_{HLT} \tag{5.4}$$

The $\epsilon_{offline}$ is constructed in terms of single electron efficiencies as follows:

$$\epsilon_{offline} = [\epsilon^1_{rec} * \epsilon^1_{id} * \epsilon^1_{iso}](p^1_T, \eta^1) \times [\epsilon^2_{rec} * \epsilon^2_{id} * \epsilon^2_{iso}](p^2_T, \eta^2) \tag{5.5}$$

Here ϵ^i_{rec}, ϵ^i_{id} and ϵ^i_{iso} represent the single electron efficiencies of reconstruction, identification and isolation for i^{th} electron in the event. The ϵ_{HLT} is constructed as,

$$\epsilon_{HLT} = \epsilon^1_{Ele17} * \epsilon^2_{Ele17} + \epsilon^1_{Ele17} * \epsilon^2_{Ele8!Ele17} + \epsilon^1_{Ele8!Ele17} * \epsilon^2_{Ele17} \tag{5.6}$$

Here ϵ^1_{Ele17}, ϵ^1_{Ele8} are the single electron efficiencies, for the electrons passing all the offline selection criteria, to fire the tight and loose legs respectively, of the relevant double electron trigger paths of Table 4.2. Rest of this section is dedicated to the technical details of the actual calculations. The efficiencies have been obtained in the two dimensional bins of electron supercluster p_T and η. In each bin, the number of passing and failing tag-probe pairs are determined via fitting the tag-probe pair invariant mass distribution in case of data, but for the MC simple event counting has been employed. The overall offline single electron efficiency is factorized into the four consecutive steps and below are the details of calculation and results for each step.

5.7.2 SuperClusters to PF Identified Electron

This step includes the efficiency of the reconstruction and the particle flow (PF) electron identification [127] given that an ECAL supercluster has been reconstructed. The particle flow identification at this stage includes a cut on **mva** variable[1] (i.e. mva>-0.1) [127] which is applied in order to reject the charged pions. The tag and probe electrons are selected as follows:

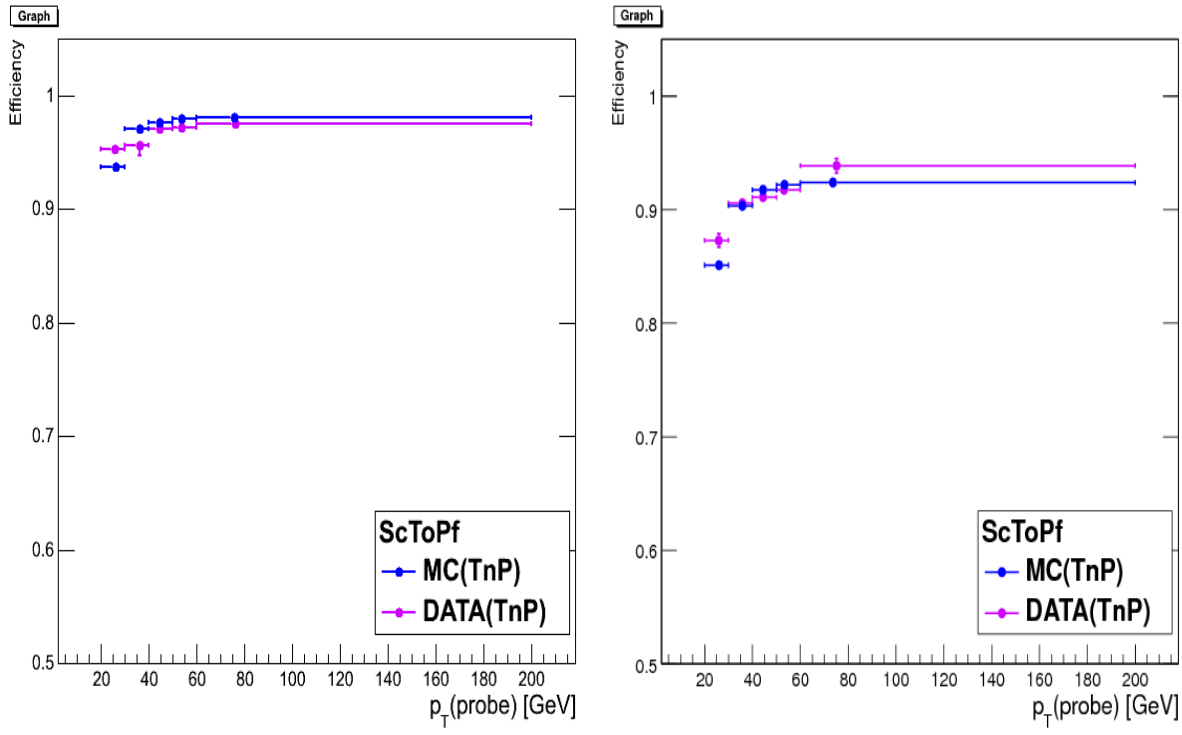
[1] A list of electron identification variables combined into a single discriminator using a multivariate Boosted Decision Trees (BDT) method.(Full definition in [127])

- **Tag** : Electrons passing full selection criteria (Table 5.1) and having $\eta - \phi$ match with the L1 object that fired the tight leg of the relevant trigger path (Table 4.2).

- **Probe**: SuperCluster candidates with $p_T > 20$ GeV, and cleaned off the contamination from hadronic jets.

- **Passing Probe**: Members of probe collection, having a $\eta - \phi$ match with a reconstructed electron ($p_T > 20$ GeV, η in Ecal Acceptance), that also passes the PF identification (mva>-0.1).

The single electron efficiencies for this step calculated in the bins of electron p_T and η both for the MC and Data are shown in the Figure 5.23. The data-MC scale factors so



(a) ECAL Barrel ($|\eta| < 1.442$) (b) ECAL Endcap ($1.556 < |\eta| < 2.4$)

Figure 5.23: Single electron efficiency for SC→PF step in barrel (a) and endcap (b) regions shown as function of probe supercluster p_T.

calculated are shown in Table 5.5.

Table 5.5: Scale Factors for SC→PF step

p_T (GeV)	Barrel	Endcap
20-30	1.017±0.005	1.026±0.007
30-40	0.986±0.005	1.003±0.002
40-50	0.995±0.001	0.993±0.001
50-60	0.992±0.001	0.995±0.001
60-200	0.993±0.002	1.016±0.007

5.7.3 Particle Flow to Identified Electron

This step calculates the efficiencies for the electron identification criteria (Table 5.1) used for this thesis analysis, relative to the electrons who already have passed the PF identification. The tag definition is same as used in previous step, and probe definitions are given below:

- **Probe**: Electrons ($p_T > 20$ GeV, η in Ecal Acceptance) passing PF Identification.
- **Passing Probe**: Members of Probe collection who pass the electron identification criteria (Table 5.1).

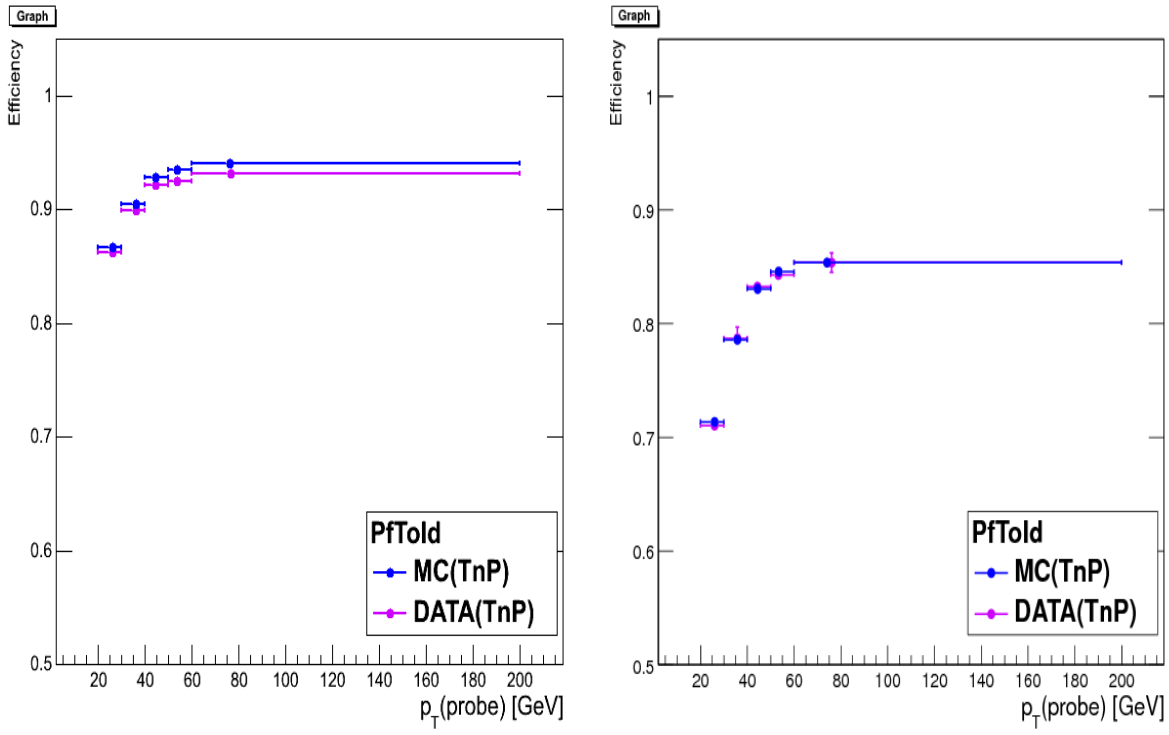
The efficiencies calculated using tag and probe method on data and MC are shown in the Figure 5.24. The data-MC scale factors so calculated are shown in Table 5.6.

Table 5.6: Scale Factors for PF→electron-ID step

p_T (GeV)	Barrel	Endcap
20-30	0.995±0.003	0.996±0.005
30-40	0.994±0.001	1.002±0.007
40-50	0.992±0.001	1.002±0.001
50-60	0.989±0.001	0.997±0.004
60-200	0.991±0.007	1.001±0.010

5.7.4 Identified To Isolated Electron

This step is for calculating the efficiency of isolation criteria (Table 5.1). The tag definition remains same as previous steps and probe definitions are listed below:



(a) ECAL Barrel ($|\eta| < 1.442$)

(b) ECAL Endcap ($1.556 < |\eta| < 2.4$)

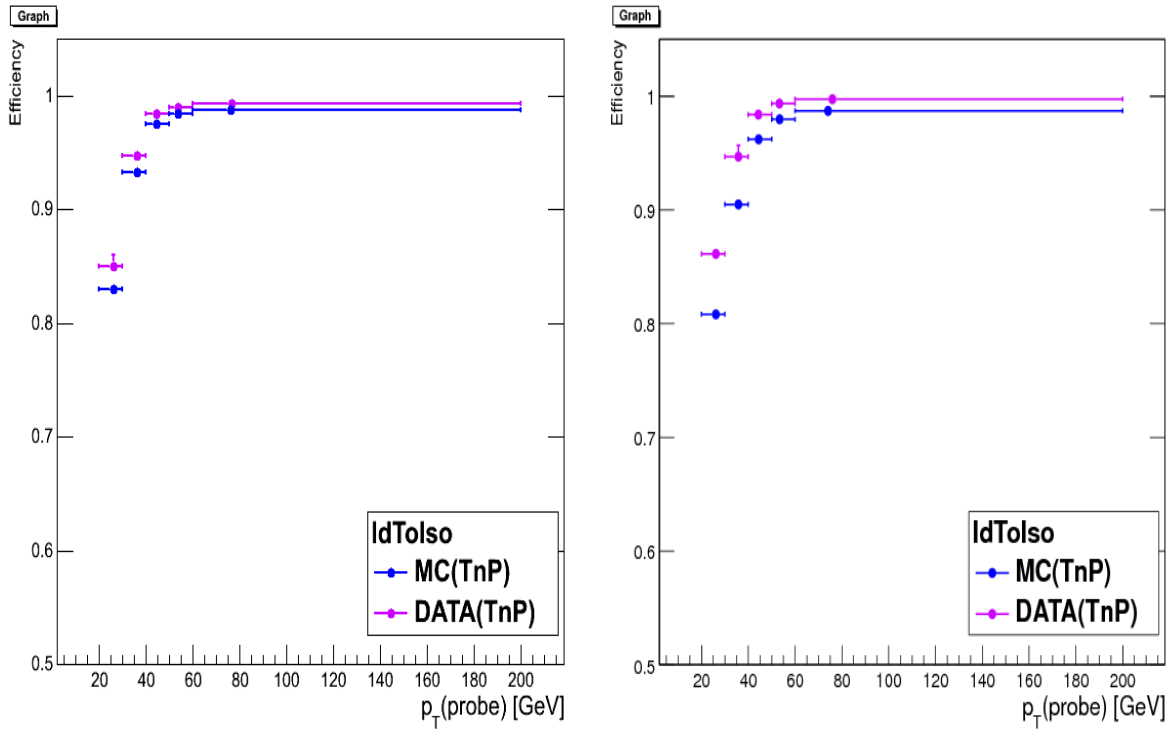
Figure 5.24: Single electron efficiency for PF→electron-ID step in barrel (a) and endcap (b) regions shown as function of probe supercluster p_T.

- **Probe**: Electrons ($p_T > 20$ GeV, η in Ecal Acceptance) passing PF and electron Identification (Table 5.1).
- **Passing Probe**: Members of Probe collection who pass the isolation criteria (Table 5.1).

The efficiencies thus calculated are shown in the Figure 5.25. The data-MC scale factors so calculated are shown in Table 5.7.

5.7.5 Evaluating Trigger Efficiency Terms

As already presented in equation 5.6, the online efficiency for a double electron path is evaluated using terms like ϵ_{Ele17} and ϵ_{Ele8}. For calculating these the probes are defined as follows:



(a) ECAL Barrel ($|\eta| < 1.442$) (b) ECAL Endcap ($1.556 < |\eta| < 2.4$)

Figure 5.25: Single electron efficiency for electron-ID→electron-Iso (isolation) step in barrel (a) and endcap (b) regions shown as function of probe supercluster p_T.

Table 5.7: Scale Factors for electron-ID→electron-Iso (isolation) step

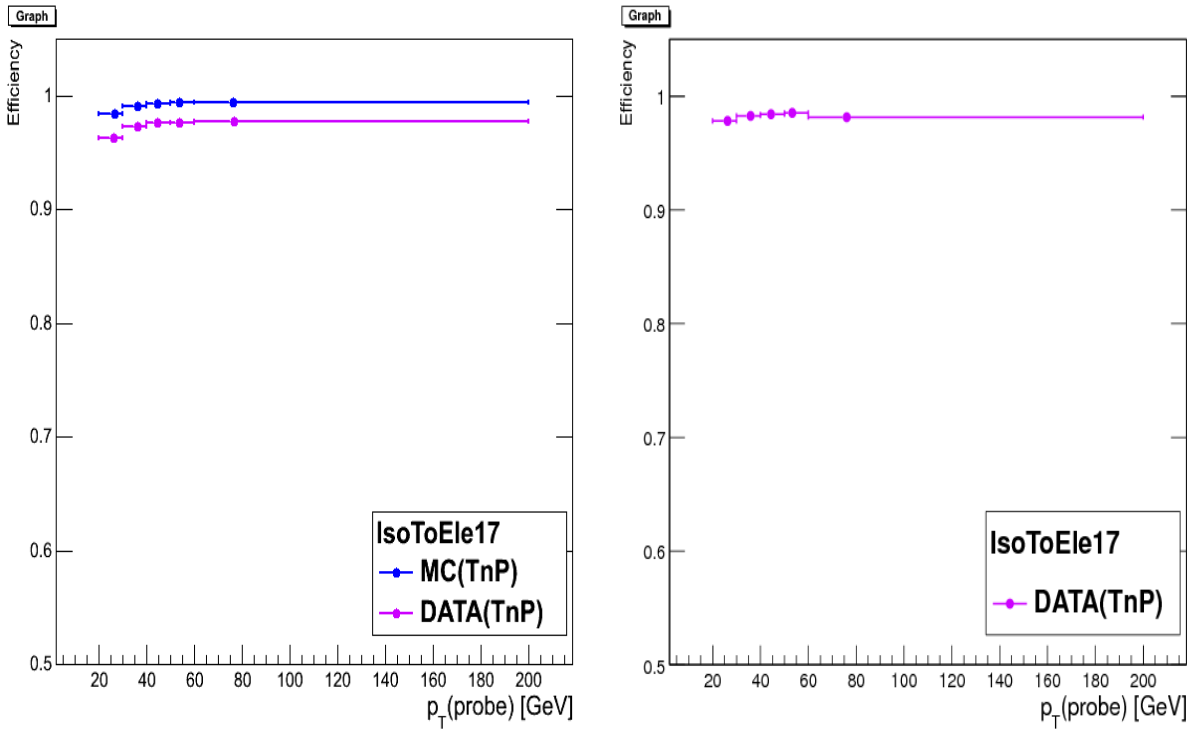
p_T (GeV)	Barrel	Endcap
20-30	1.024±0.007	1.066±0.005
30-40	1.016±0.001	1.046±0.006
40-50	1.010±0.001	1.023±0.001
50-60	1.006±0.001	1.014±0.001
60-200	1.006±0.001	1.010±0.003

- **Probe**: Electrons passing complete offline selection (Section 5.2).

- **Passing Probe for** ϵ_{Ele17}: Members of Probe collection having an $\eta - \phi$ match to the L1 object firing tight leg of relevant double electron trigger.

- **Passing Probe for** ϵ_{Ele8}: Members of Probe collection having an $\eta - \phi$ match to the L1 object firing loose leg of relevant double electron trigger.

Since we do not have a completely faithful trigger emulation in the MC, we do not apply it to our MC samples in this analysis. Hence the efficiencies are calculated on data only, and assumed to be 100% efficient in MC. The results are shown in Figures 5.26 and 5.27.



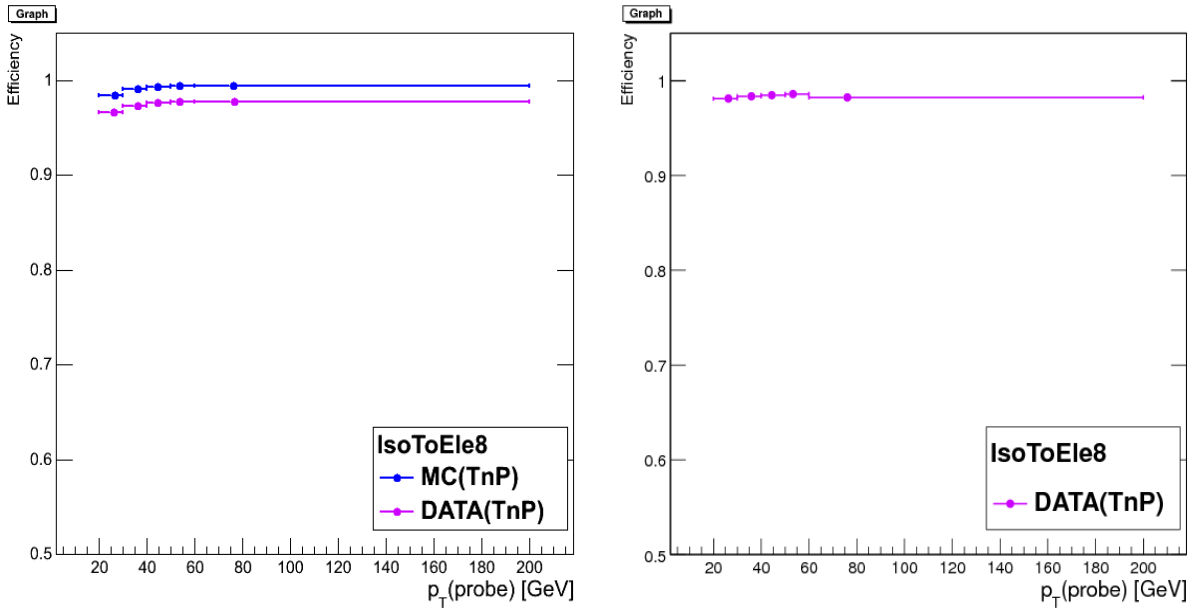
(a) ECAL Barrel ($|\eta| < 1.442$)

(b) ECAL Endcap ($1.556 < |\eta| < 2.4$)

Figure 5.26: Electron efficiency to fire the tight leg (with 17 GeV threshold) of the relevant double electron trigger path in barrel (a) and endcap (b) regions shown as function of probe supercluster p_T.

5.7.6 Uncertainty on Efficiency Calculation

The tag and probe based estimation of the electron efficiencies can be biased due to many sources. Prime among them are the fitting related biases which come from the choice of the shapes used for signal and background. Another source can be the biasing of events used for efficiency studies by requiring the presence of a high quality tag. Since we are using the scale factors (Data/MC ratio) we expect any residual tag and probe related bias



(a) ECAL Barrel ($|\eta| < 1.442$)

(b) ECAL Endcap ($1.556 < |\eta| < 2.4$)

Figure 5.27: Electron efficiency to fire the loose leg (with 8 GeV threshold) of the relevant double electron trigger path in barrel (a) and endcap (b) regions shown as function of probe supercluster p_T.

to cancel out between the denominator and the numerator. We assume a 2% systematic uncertainty on the event efficiencies to incorporate the dependence of efficiencies of any factor.

Scale factors as function of N_{Jet}

This thesis analysis uses the event efficiencies calculated as function of N_{jet}, and the scale factor themselves are calculated as function of electron p_T and η only, so these are checked for any possible dependence on N_{Jet}. This dependence as a function of N_{Jet} are shown in Table 5.8 and we see that all of them are consistent with each other and are also close to unity.

Table 5.8: Data versus MC scale factors as function of N_{Jet}

N_{Jet}	0	1	2	3
ϵ_{ee}	1.001	1.002	1.004	1.001

Chapter 6

Signal Yield Extraction and Unfolding

This chapter has been organized into two main parts: First we discuss about the signal yield extraction from fitting the Z invariant mass distribution in section 6.1 and section 6.2. The background subtracted results are also presented in section 6.3 and in section 6.4 we discuss the multiplicity scaling results. Then in order to dis-entangle the detector effects from these measurements, we have to do unfolding procedure which has been discussed in sections 6.5 and 6.6.

6.1 Signal Yield Extraction

The selected events are expected to have high signal purity. Major irreducible background is $t\bar{t}$ and is most prominent in N_{jet}=2 bin i.e. Z+2jet events. In order to extract the number of pure signal events, we perform extended maximum likelihood fit to Z invariant mass in each jet multiplicity (N_{jet}) bin. While doing the fit we use the invariant mass (M_{ee}) distribution from Z+0jet events as signal shape and use a falling exponential parametrization of background shape.

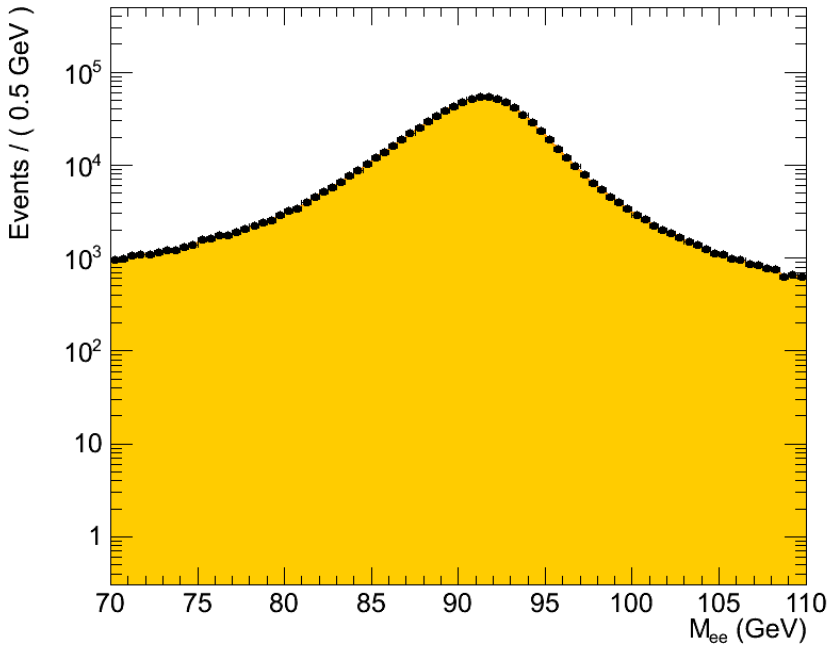


Figure 6.1: Invariant mass (M_{ee}) distribution from Z+0jet events taken as signal shape (data points shown with black markers and signal MC with yellow filled area)

6.2 Fit to obtain Signal Yields

We introduce the differential yields for signal+backgrounds in each jet multiplicity bin using following Probability distribution functions (PDF):

$$\frac{dN^0_{ee}}{dm} = S^0_{ee} f^0_{peak}(m), \tag{6.1}$$

$$\frac{dN^j_{ee}}{dm} = S^j_{ee} f^j_{peak}(m) + B^j_{ee} b(m),\ m\ is\ invariant\ mass. \tag{6.2}$$

Here S^0_{ee} is signal yield for Z+0jet events, S^j_{ee}, B^j_{ee} are the yields in j'th jet multiplicity bin for signal and background respectively and j varies from 1 to 4. The yields have been factorized out so that the shape functions $f_{peak}(m)$ and $b(m)$ are normalized to unity. The $f^j_{peak}(m)$ are constructed by convolving $\frac{dN^0_{ee}}{dm}$ distribution with a Gaussian resolution function. This method has two underlying assumptions, first one being that Z+0jet sample is overwhelmingly rich in signal i.e. residual backgrounds have practically no effect on the shape. The second one comes from assuming that M_{ee} shape dependence

on N_{jet} can be taken care off by Gaussian resolution function. The background shape is modeled using a falling exponential function in each bin. The S^j_{ee}, B^j_{ee} are finally obtained by combined fit over N_{jet} = 1, 2, 3 and 4 and the fitted functions are shown overlaid on the data in Figure 6.2. The signal yields from fit are shown in Table 6.1.

Table 6.1: Signal and Background Yields Obtained from Fit to selected Events

sample	N_{Signal}	N_{Back}
Z/γ*+1jets	133751±365.787	150.0002±30.551
Z/γ*+2jets	27465±167.103	490.578±78.0372
Z/γ*+3jets	5264.96±73.3523	125.969±35.1135
Z/γ*+4jets	1002±31.9911	23.1248±15.3675

6.2.1 Toy MC pseudo-experiments

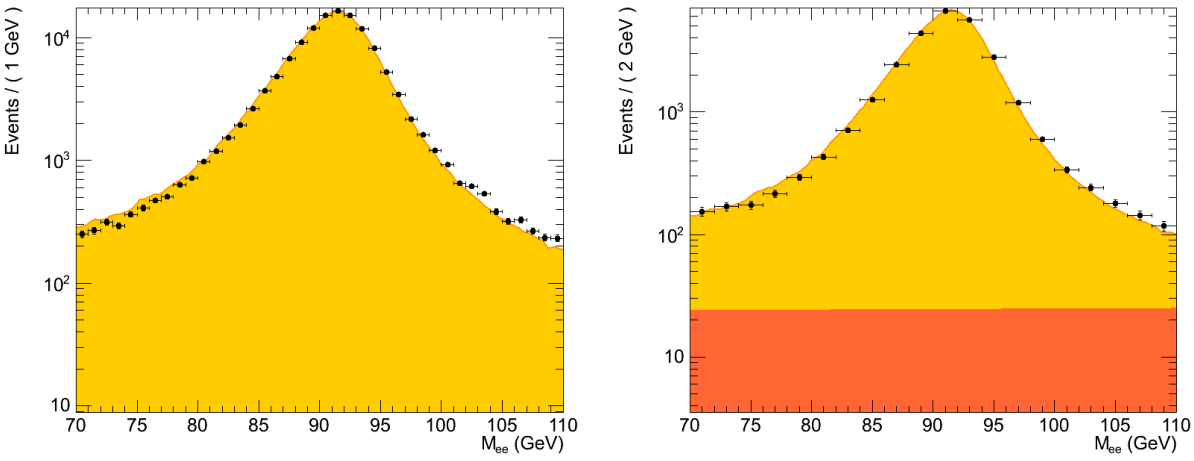
To validate the fitting procedure for signal yield extraction, we performed several tests using large number of toy MC pseudo-experiments. The signal and background yields were fluctuated according to Poisson distributions and the M_{ee} is generated according to PDFs (equation 6.1, 6.2). To validate the fit results, the pull of the signal yield was determined. The pull is the difference between the fit yield and the generated number of events, divided by the error on the fit i.e. for a given parameter P in the fit, the pull is defined as:

$$pull \quad = \quad \frac{P^{fit} - P^{gen}}{P^{err}} \tag{6.3}$$

We show the pull distributions for the signal yields with N_{jet} = 1, 2, 3, and 4 in Figure 6.3. The Gaussian distribution with mean 0 and sigma = 1 for the pull, indicate an intrinsically unbiased fit.

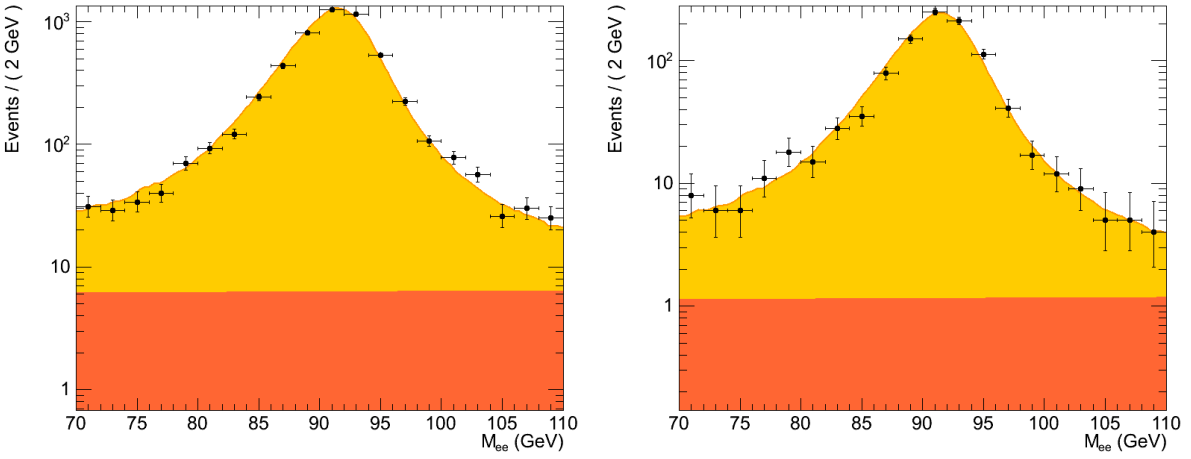
6.2.2 MC closure test

For this MC closure test, fitting procedure was performed same as described in Section 6.1 but by replacing the di-electron mass distributions from data with "pseudo-data" which



(a) Z+1Jet

(b) Z+2Jet

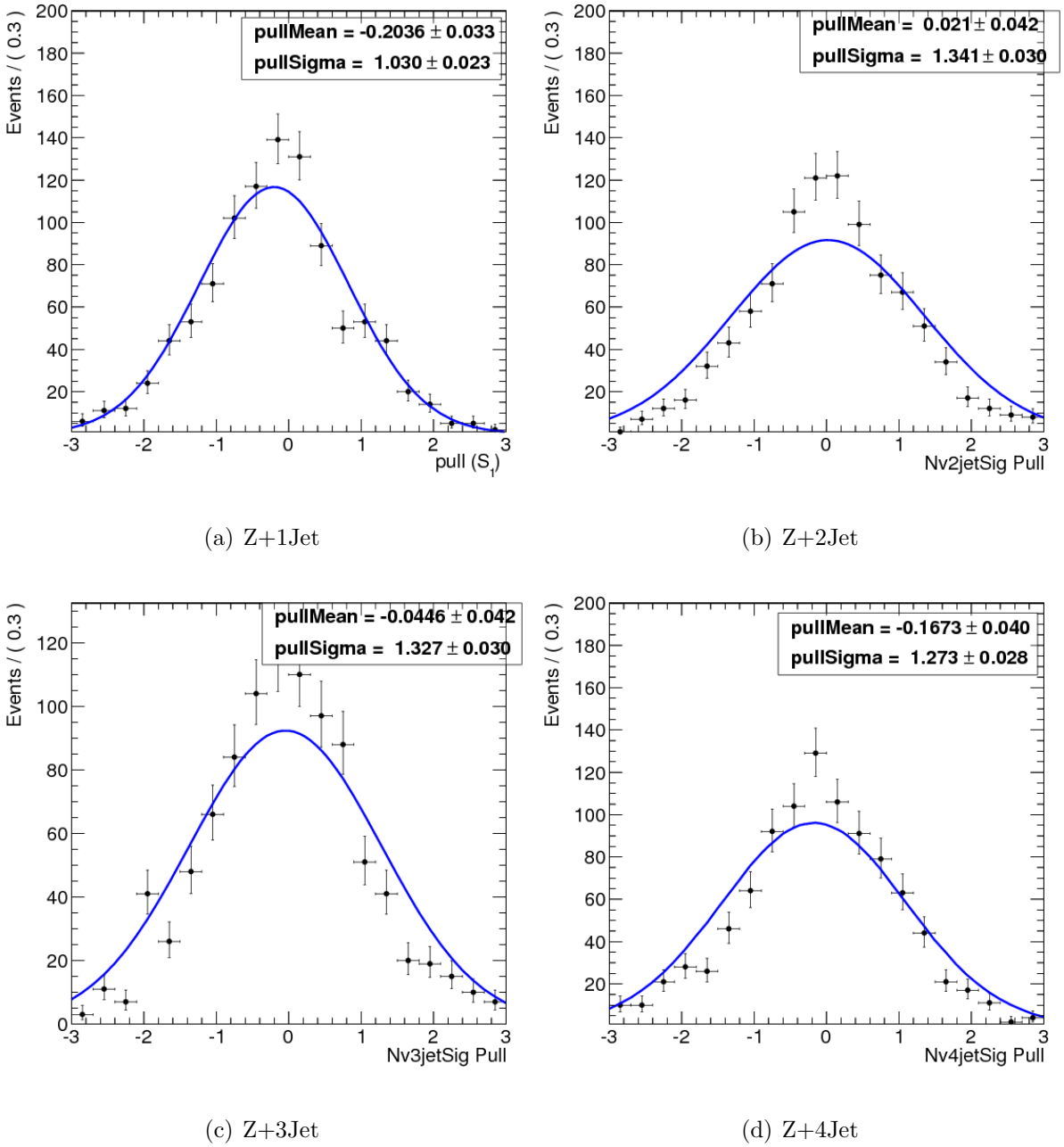


(c) Z+3Jet

(d) Z+4Jet

Figure 6.2: The simultaneous fit to di-electron invariant mass distribution (data points shown with black markers, signal PDF in yellow and background PDF in orange)

is actually merging of all MC events (from signal as well as background Monte Carlo samples) passing the full selection criteria. Fit shapes are shown in Figure 6.4. The corresponding fit yields are given in Table 6.2 which are found to be consistent with Table 5.4.



(a) Z+1Jet

(b) Z+2Jet

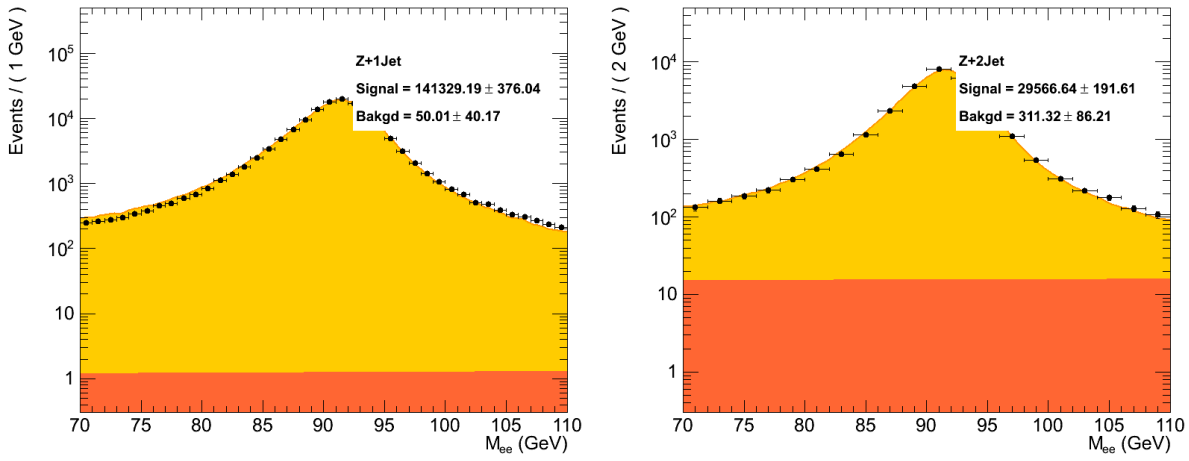
(c) Z+3Jet

(d) Z+4Jet

Figure 6.3: Pull distributions for signal yields obtained from toy MC studies

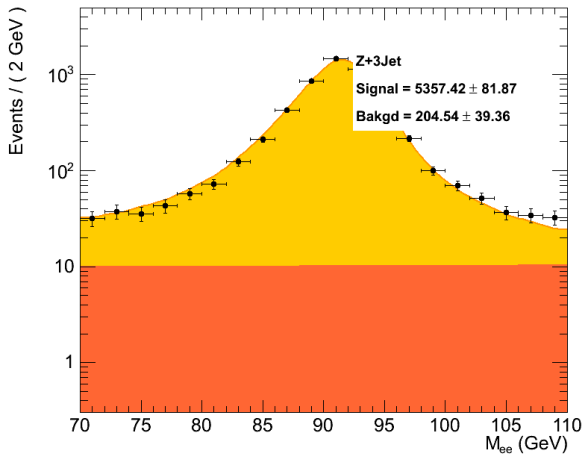
6.3 Background subtraction

Though the residual backgrounds form a very small fraction of the total events in passing full selection criteria as discussed in above section (Table 6.1 and 6.2). Nevertheless we perform background subtraction from the jet kinematic distributions (p_T and Y) using

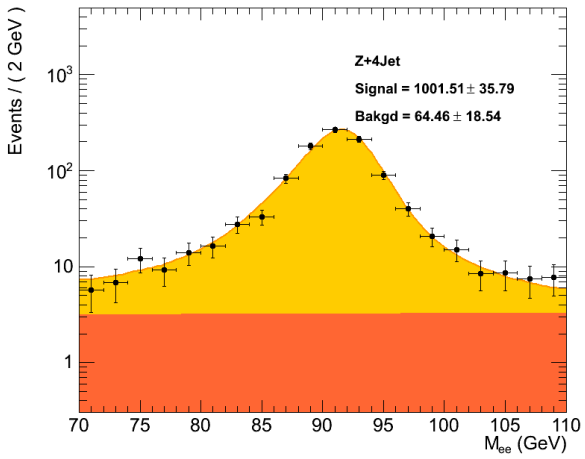


(a) Z+1Jet

(b) Z+2Jet



(c) Z+3Jet



(d) Z+4Jet

Figure 6.4: The simultaneous fit to di-electron invariant mass distribution in different jet multiplicity bins for MC closure test (data points shown with black markers, signal PDF in yellow and background PDF in orange)

the shapes from Monte Carlo samples, normalized to the corresponding yields from the fit.

Table 6.2: MC closure fit yields, "Tru" stands for true yields, "Fit" means yields from MC closure fitting

N_{Jet}	N_{Signal}	N_{Back}
1(Tru)	141149	223
1(Fit)	141329±376	50±40
2(Tru)	29412	465
2(Fit)	29566±191	311±86
3(Tru)	5342	220
3(Fit)	5357±81	204±39
4(Tru)	996	69
4(Fit)	1001±35	64±18

6.4 Fit to obtain Multiplicity Scaling

In this section, we test for scaling [128] and it has already been observed, at CMS with $36pb^{-1}$ of the 2010 data collected at $\sqrt{s} = 7\ TeV$ [129], that the $Z + n\ jets$ yields scale with N_{jet} as follows:

$$C_n \quad = \quad N_{n\ jet}/N_{(n+1)jet} \tag{6.4}$$

To the leading order in strong coupling constant (α_s), we expect that C_n in equation 6.4 is independent of the N_{jet}, although higher order contributions may lead to minor dependence. For this scaling effect we performed another fit with following assumption:

$$C_n = (\alpha + \beta.n) \tag{6.5}$$

Here the β term has been included to test the dependence on jet multiplicity. The fitted values of the parameters are shown in Table 6.3. The fit was validated through a large

Table 6.3: The values of α and β along with the uncertainties as obtained from the fit

Parameter	Value
α	4.4±0.12
β	0.25±0.05

number of MC pseudo-experiments. The pull distributions for α, and β are shown in Figure 6.5. We see that the results show negligible bias in the parameters of interest.

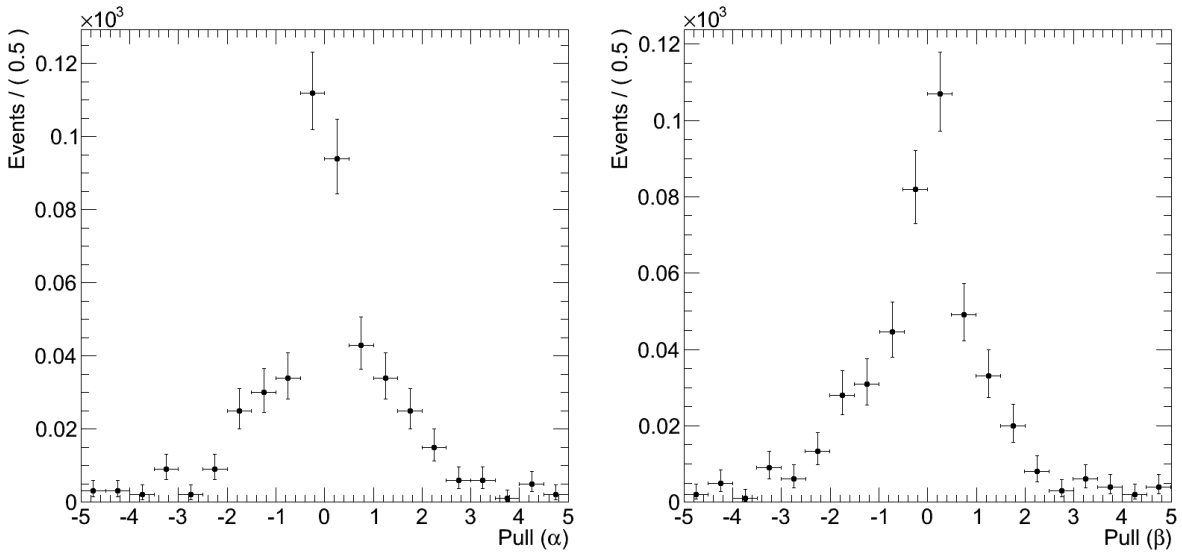


Figure 6.5: Pull distributions for α and β from toy MC studies.

6.5 Correcting for the detector effects

The background subtracted data (Section 6.3), have to be corrected for detector effects in order to be compared with theoretical predictions at the hadron level. This section explains the methods which help to correct for the measured data jet distributions, for event migration due to finite jet energy resolution of the detector, and leads to determine the particle level cross sections. Section 6.5 illustrates the corrections to be applied to the p_T and Y spectrum of jets.

6.6 Correction for jet p_T and Y spectrum

The finite detector resolution along with the steeply falling jet p_T spectrum distorts the measured cross section with respect to the particle level cross section. So the observed number of the events in each p_T bin can be divided into two classes:

- those who were in same bins even at particle level.
- those who migrated from the neighbouring bins at particle level.

Hence in order to be able to perform comparison of the results with theory and corresponding results from other experiments, they need to be *unfolded* from these so called *detector effects*. Here we attempt to retrieve the true data distribution of the measured quantities: the transverse momentum (p_T) and rapidity (Y) of the leading jets and the exclusive number of reconstructed jets (N_{jet}), which are used for the cross section measurements.

6.6.1 Iterative Bayesian Method from RooUnfold package

The RooUnfold [130] package provides a common framework to evaluate and apply several unfolding algorithms. It provides interfaces for the Iterative Bayes, Singular Value Decomposition, TUnfold, bin-by-bin and matrix inversion reference methods. Unfolding procedure uses a migration matrix and two training distributions as an input. The training distributions are the measured M_j and the truth T_i. The matrix contains the fraction R_{ji} of events T_i which end up in M_j. For this analysis Iterative Bayesian Method is used to apply unfolding. The iterative Bayes method (described in [131]) is based on the repeated application of Bayes' theorem to invert the response matrix. Regularization is achieved by stopping iterations before reaching the "true" inverse. The regularization parameter is just the number of iterations which in principle, has to be tuned according to the sample statistics and binning. It determines the allowed amount of smoothing, placing either more weight on the data or on the training sample truth. The response matrices used are derived from Z+Jets MadGraph Tune Z2 Monte Carlo sample (Table 4.4). The regularization parameter is chosen as kBayes = 4, as it gives the best results on MC closure tests. To train the response matrix, same kinematic cuts in electrons and jets are applied at hadron level and detector level. Thus, the cross section measurements refer to hadron level jets with p_T >30 GeV and $|\eta|$ <2.4 in the kinematic range of $Z/\gamma*$ decay products: p_T >20 GeV, $|\eta|$ <2.4 and $70 < M_{ee} < 110$. The truth level distribution is obtained from generated jets with p_T >15 GeV and $|\eta|$ <2.4. The reconstructed as well

as generated jet collection is cleaned of jets that lie within ΔR=0.3 of electrons. For the jet p_T and Y distributions, the reconstructed jets are required to match to generated jets within $\Delta R < 0.3$ cone. The response matrices thus obtained for jet multiplicity, p_T and Y are shown in Figures 6.6, 6.7 and 6.8, respectively. Applying the response matrix of

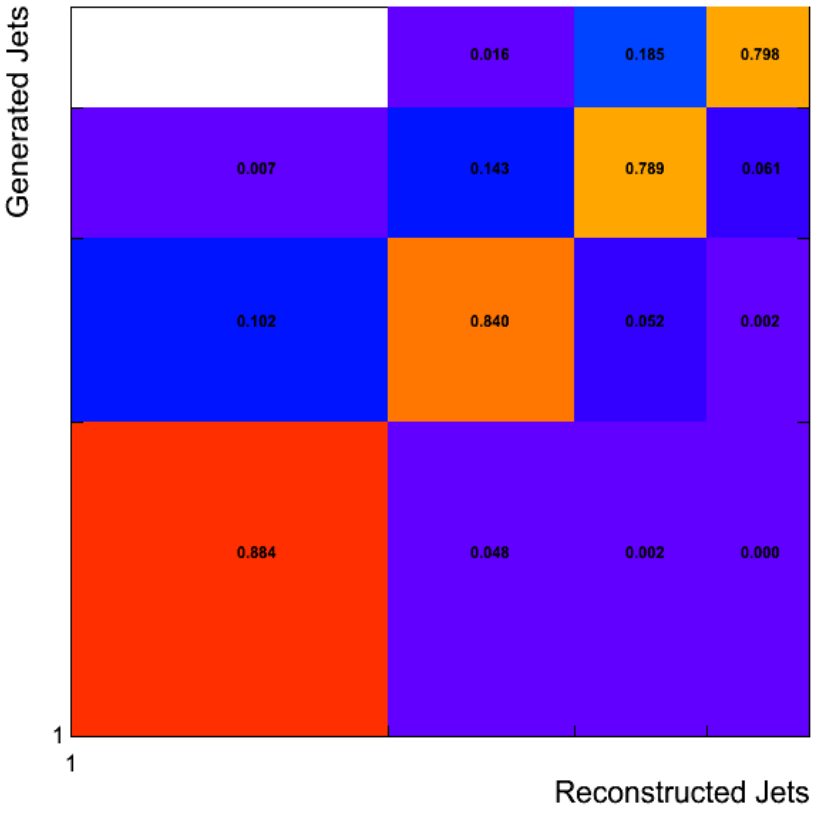
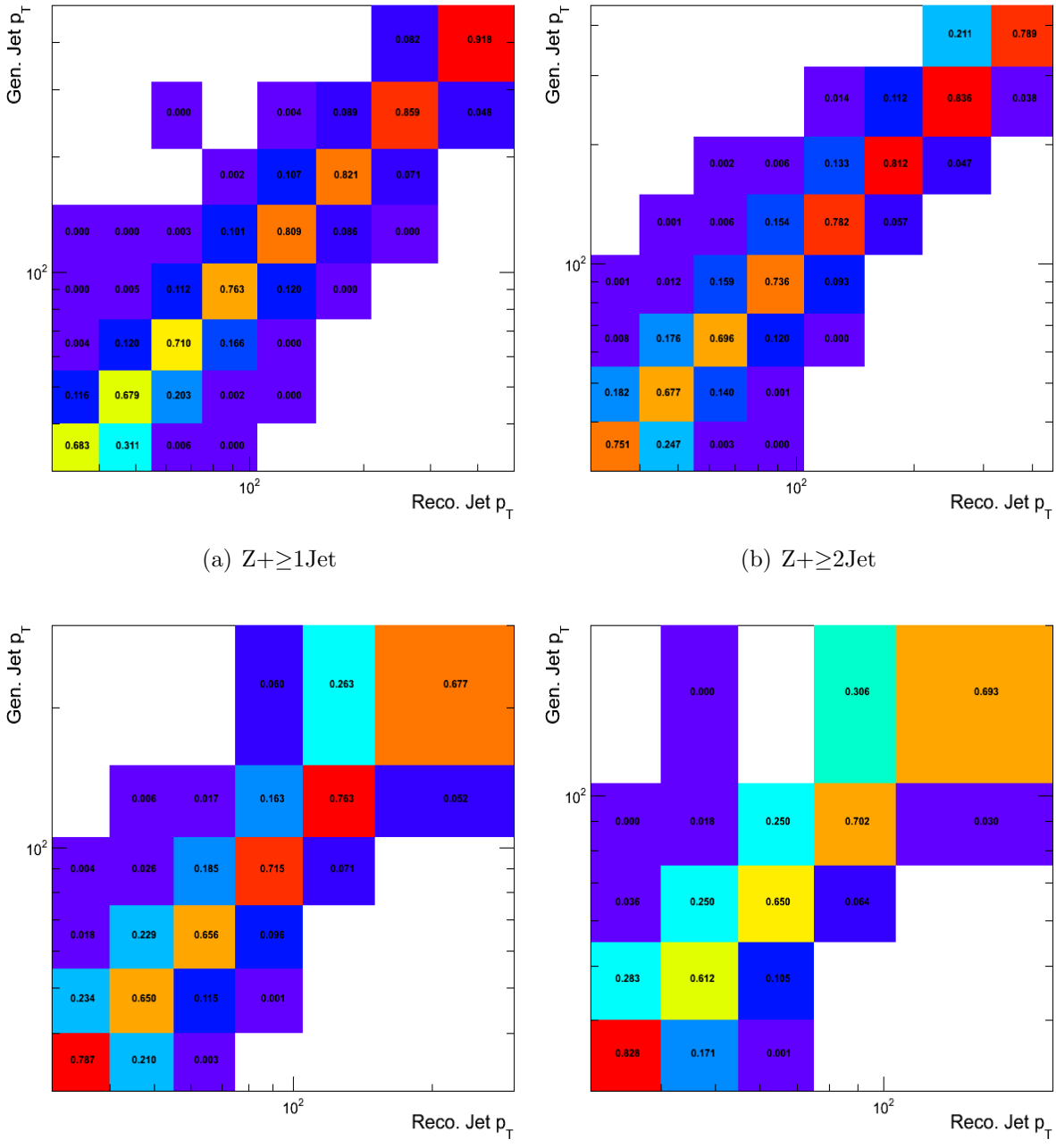


Figure 6.6: The unfolding response matrix for jet multiplicity spectrum.

Figure 6.6 to jet multiplicity spectrum of data, we get the unfolded jet multiplicities as shown in Figure 6.9 and Table 6.4. Finally the response matrices from Figures 6.7 and 6.8 are applied to the background subtracted jet p_T and Y distributions which results into unfolded spectrums. Figure 6.10 and Figure 6.11 show the jet p_T and Y spectrum, respectively, before and after unfolding for N_{Jet} = 1, 2, 3 and 4. Note that since the jet energy resolution is worse in the low p_T region and the spectrum falls steeply, the contamination in a bin from lower p_T neighbours is much higher than from higher p_T. By same logic the

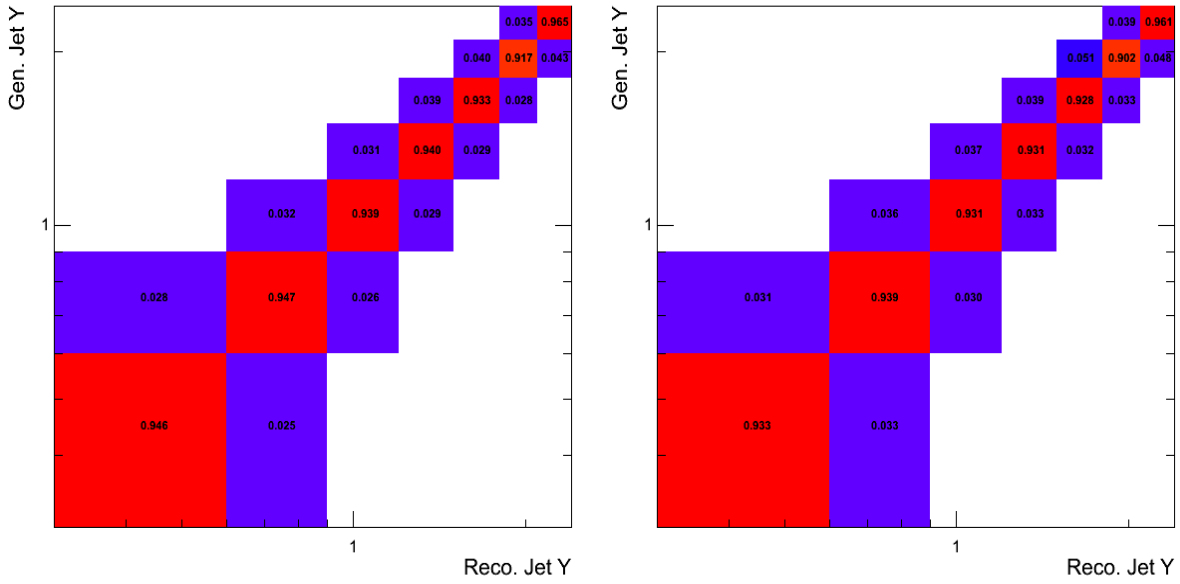


(a) Z+≥1Jet (b) Z+≥2Jet

(c) Z+≥3Jet (d) Z+≥4Jet

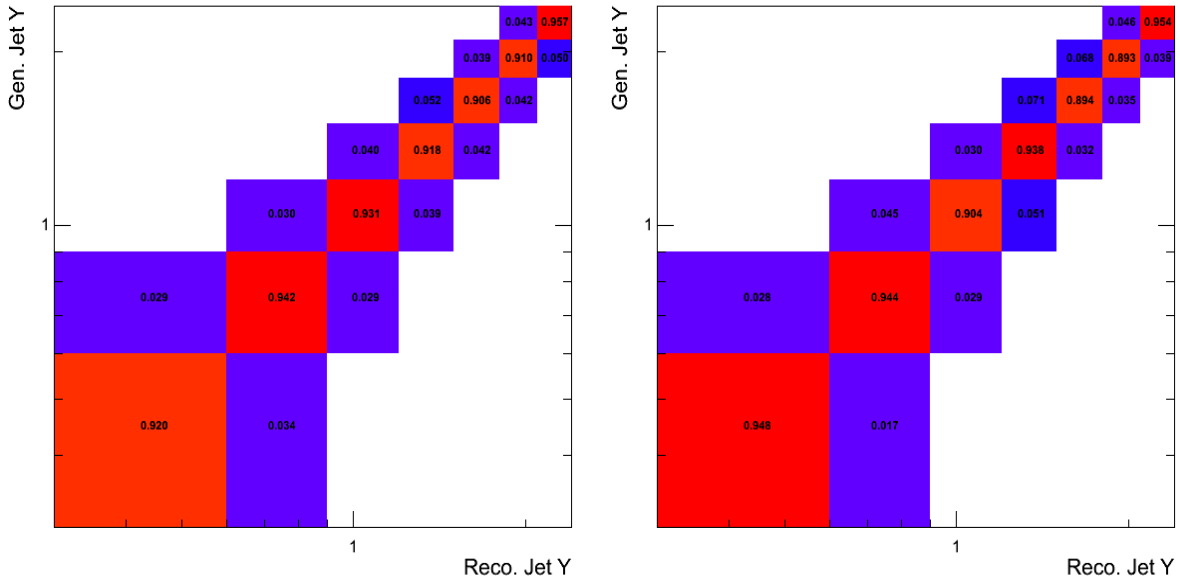
Figure 6.7: The unfolding response matrices for nth leading jet p_T in Z+njet sample.

population of jets passing 30 GeV threshold but having true p_T <30 GeV far outnumbers the ones who could not make it past thresholds but otherwise had true p_T >30 GeV. Since unfolding procedure is expected to correct for this bias, the corresponding correction fac-



(a) Z+≥1Jet

(b) Z+≥2Jet



(c) Z+≥3Jet

(d) Z+≥4Jet

Figure 6.8: The unfolding response matrices for nth leading jet Y in Z+njet sample.

tors on measured data are expected to be less than unity and integral of the unfolded distribution is less than original. We can see in the above referred figures that majority of the loss is concentrated in low p_T bins, and nearly uniformly spread over the jet rapidity

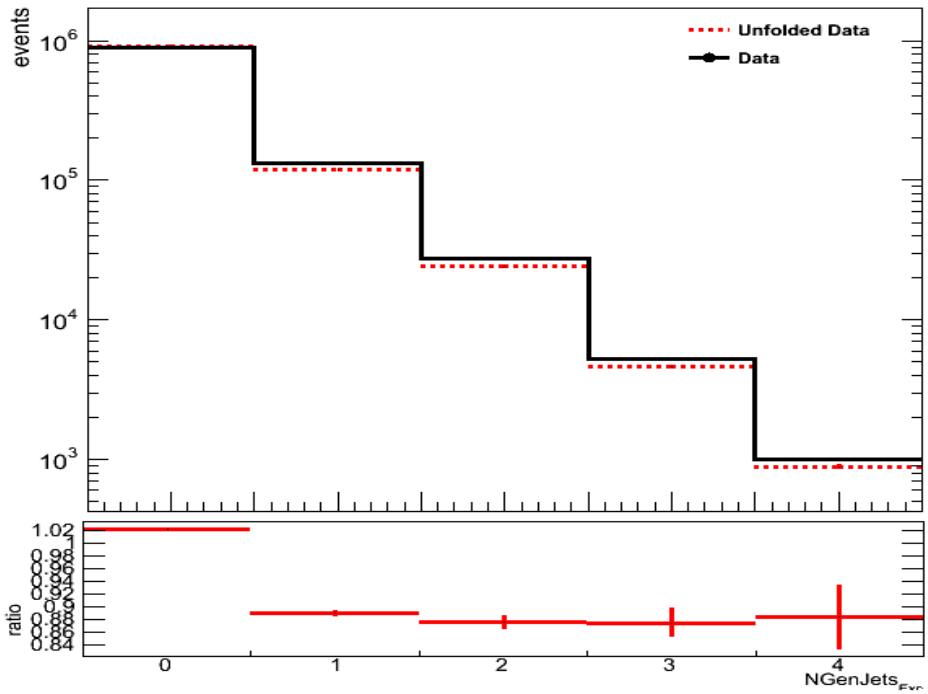


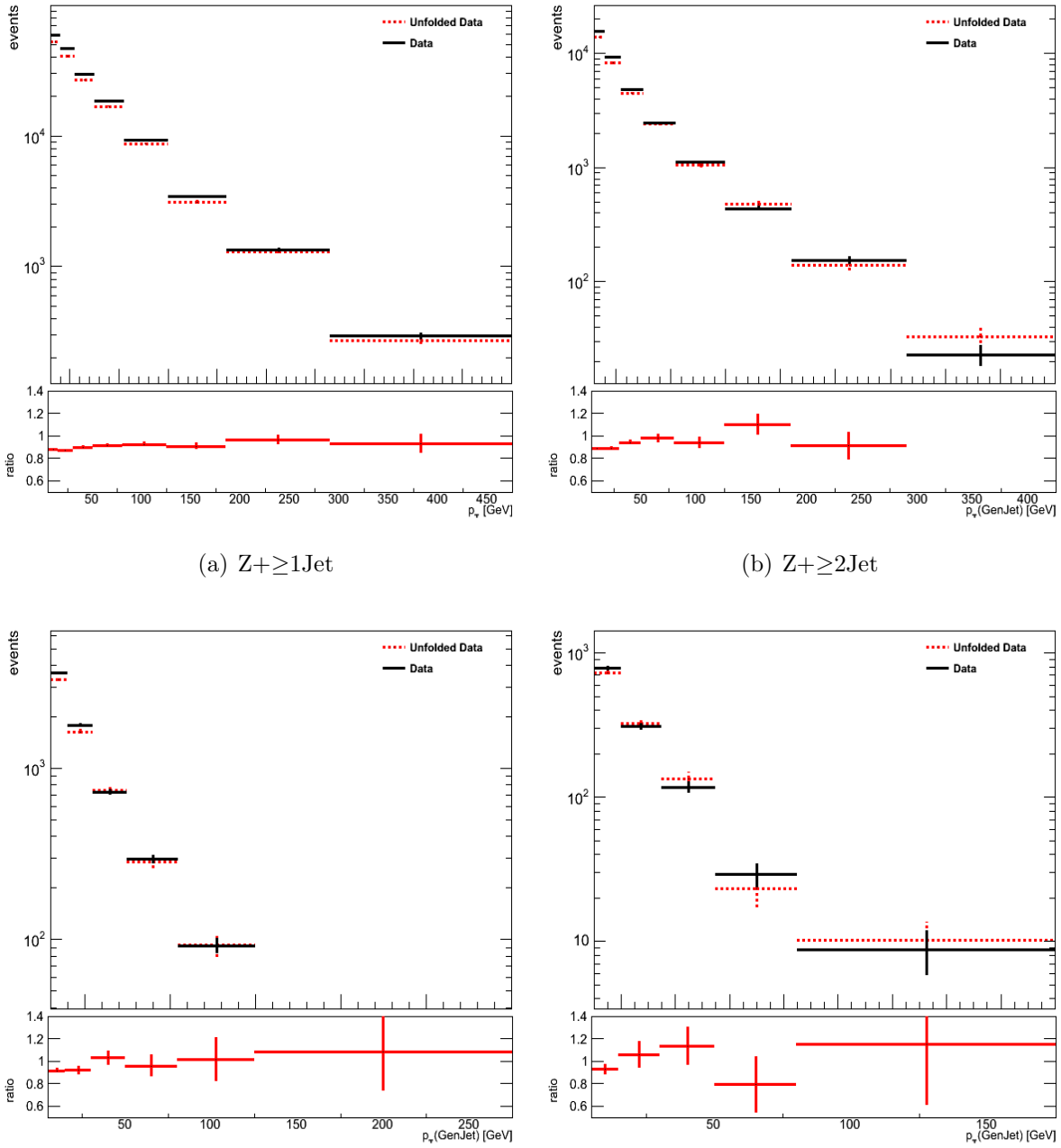
Figure 6.9: Unfolded (dashed line) and folded (solid line) jet multiplicity spectrum.

Table 6.4: Background subtracted folded and unfolded data yields

sample	folded yields	unfolded yields
Z/γ*+0jets	887745	906824
Z/γ*+1jets	133801	118928
Z/γ*+2jets	27956	24059
Z/γ*+3jets	5391	4606
Z/γ*+4jets	1025	885

distribution. We also observe that correction factors are different from unity even for the high p_T bins where the resolution is expected to be very good. This happens due to the fact that even if the resolution improves, the spectrum becomes much more steeper.

The observed difference between unfolded and measured rapidity distributions is primarily due to resolution effects associated with a p_T threshold. In-order to disentangle the effect of p_T resolution, the unfolded jet Y distributions are normalized with integral of jet p_T and the observed shape difference between unfolded and measured distribution



(a) Z+≥1Jet

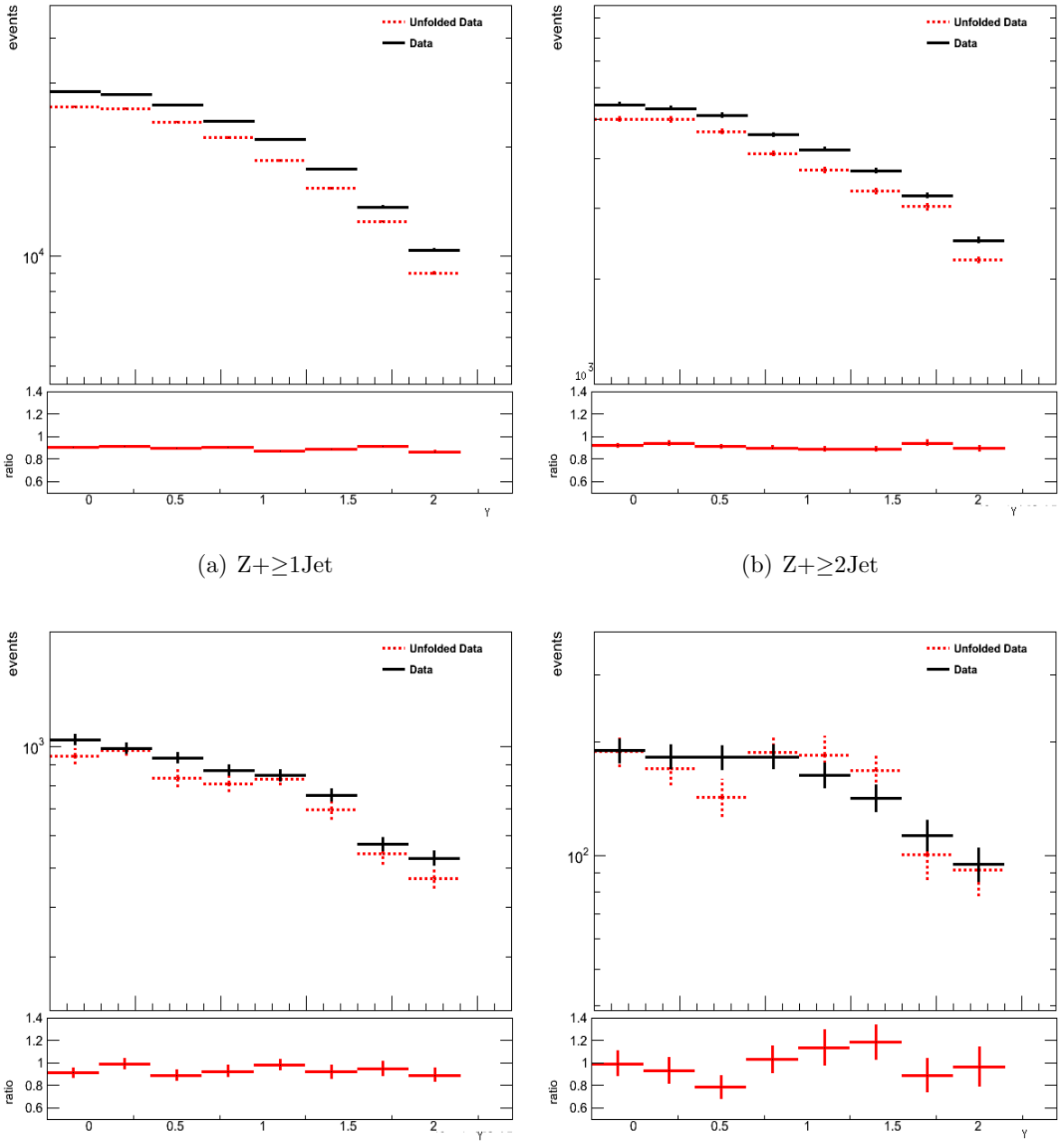
(b) Z+≥2Jet

(c) Z+≥3Jet

(d) Z+≥4Jet

Figure 6.10: The jet p_T spectrum before (solid black) and after (dashed line) unfolding for nth leading jet in Z+njet sample.

was extremely small (Figure 6.12).



(a) Z+≥1Jet

(b) Z+≥2Jet

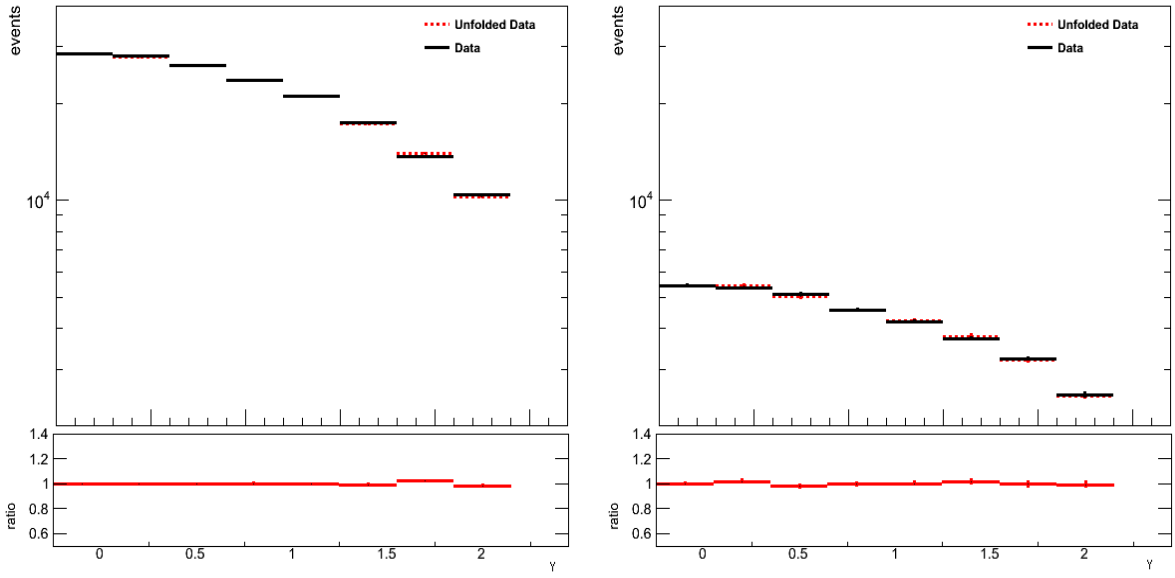
(c) Z+≥3Jet

(d) Z+≥4Jet

Figure 6.11: The jet Y spectrum before (solid black) and after (dashed line) unfolding for nth leading jet in Z+njet sample.

6.6.2 Validating unfolding with Monte Carlo

Before unfolding the data, a series of tests are performed to validate the unfolding procedure. Two different types of checks are performed:



(a) Z+≥1Jet (b) Z+≥2Jet

Figure 6.12: The unfolded and normalized (to jet p_T integral) jet Y distributions for nth leading jet in Z+njet sample.

- Unfolding distributions using the same signal MC as the one used to create the response matrix; this was performed on the MadGraph TuneZ2 MC. The sample was split into two equal parts; the first half was used for creating the response matrix while the second for testing the unfolding (Figures 6.13, 6.14 and 6.15, respectively).

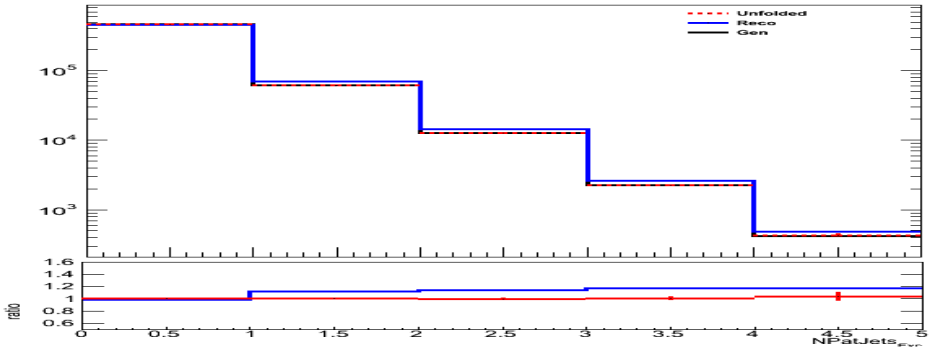
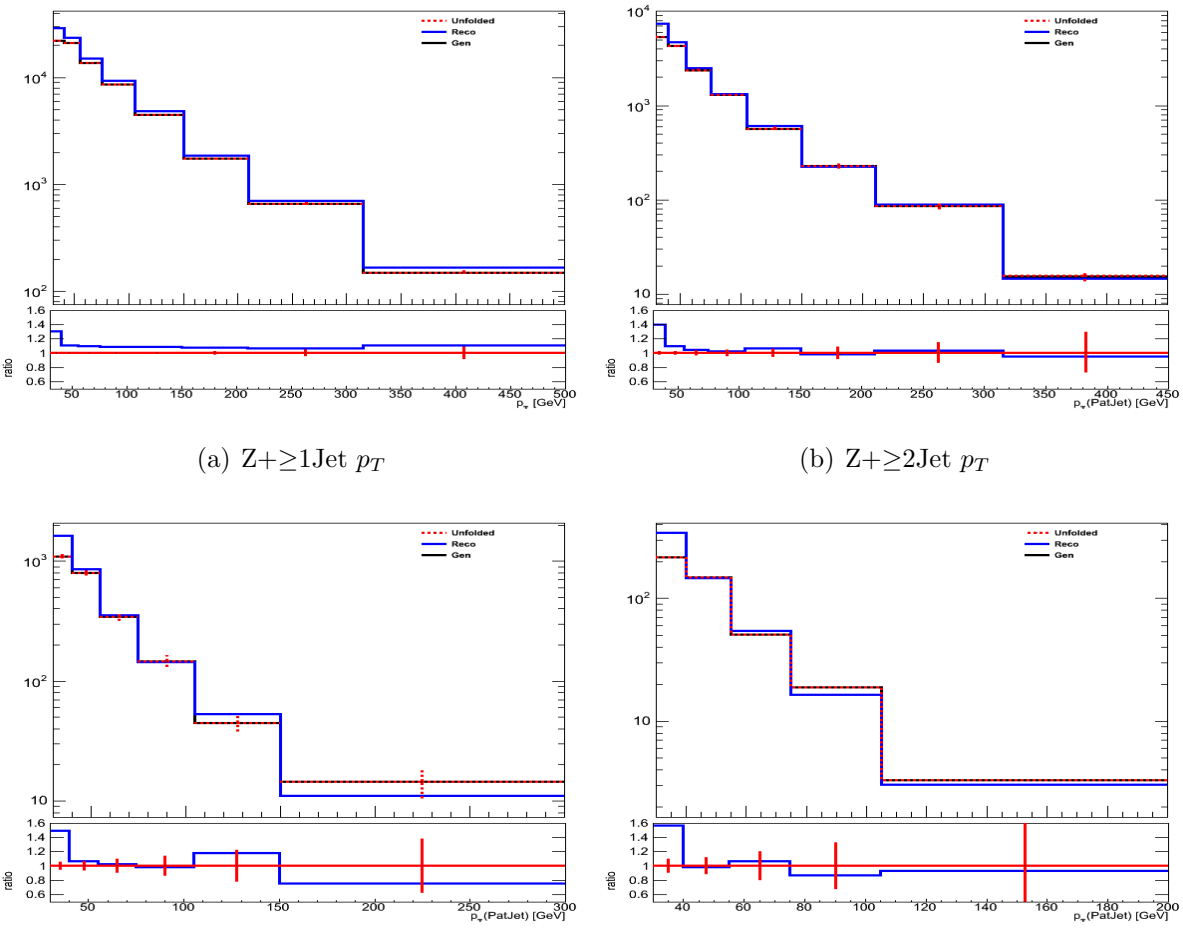


Figure 6.13: Unfolded jet multiplicity spectrum for MC closure test using response matrix from Madgraph sample splitted into two disjoint parts



(a) Z+≥1Jet p_T

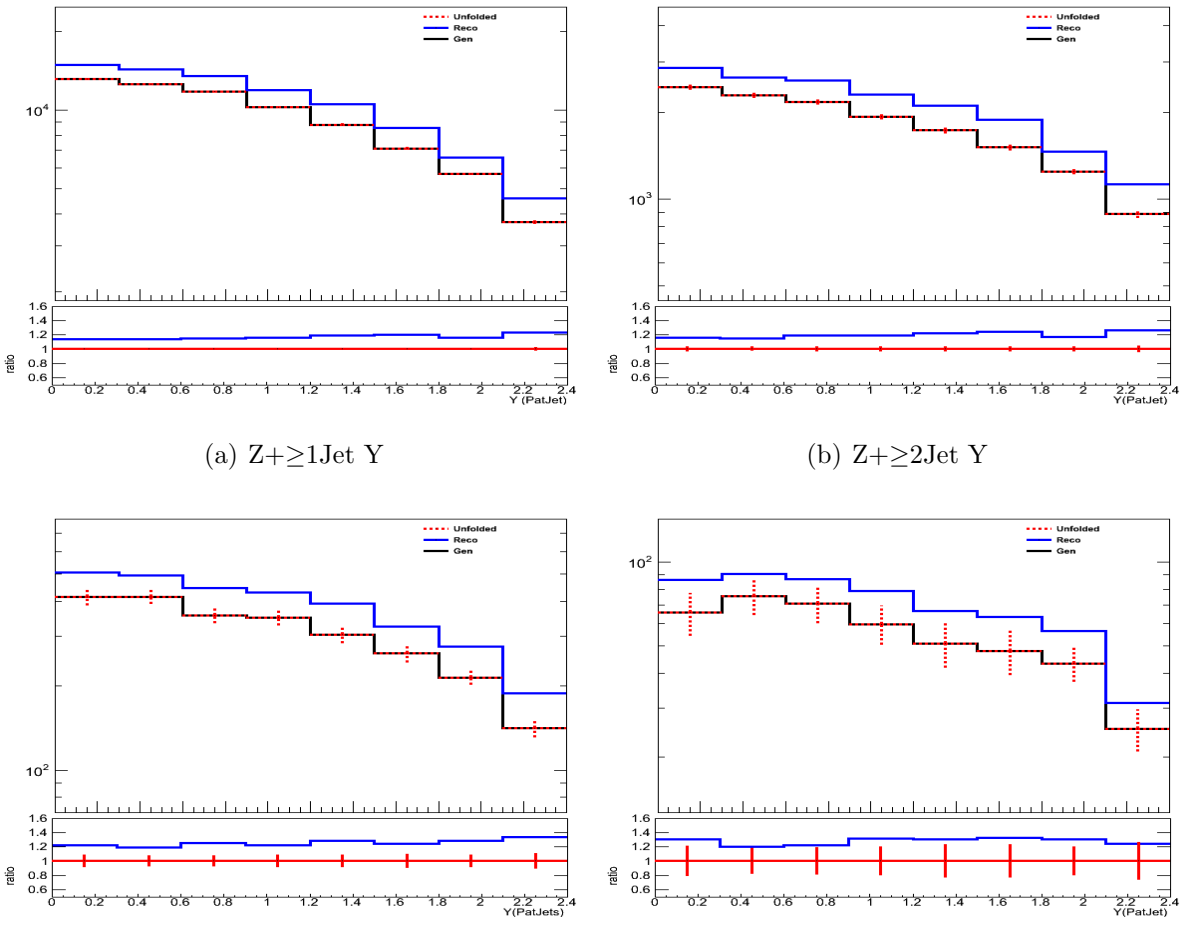
(b) Z+≥2Jet p_T

(c) Z+≥3Jet p_T

(d) Z+≥4Jet p_T

Figure 6.14: The unfolded Jet PT spectrum for MC closure test using response matrix from Madgraph sample splitted into two disjoint parts.

- Unfolding distributions using different MC samples from the ones used to create the response matrix; the MadGraph TuneZ2 response matrix was used to unfold the POWHEG results shown in Figures 6.16, 6.17 and 6.18, respectively.



(a) Z+≥1Jet Y

(b) Z+≥2Jet Y

(c) Z+≥3Jet Y

(d) Z+≥4Jet Y

Figure 6.15: The unfolded Jet Y spectrum for MC closure test using response matrix from Madgraph sample splitted into two disjoint parts.

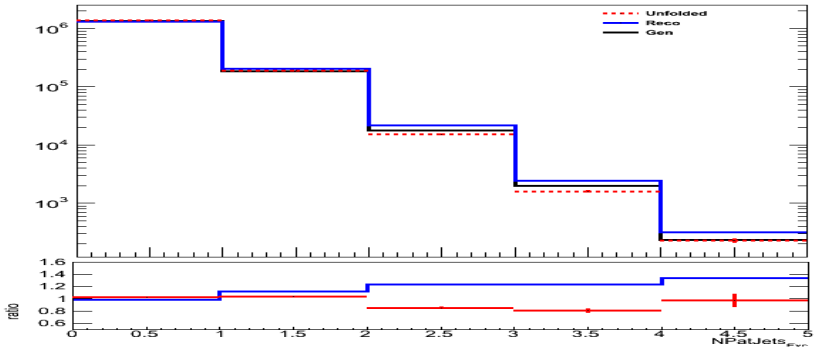
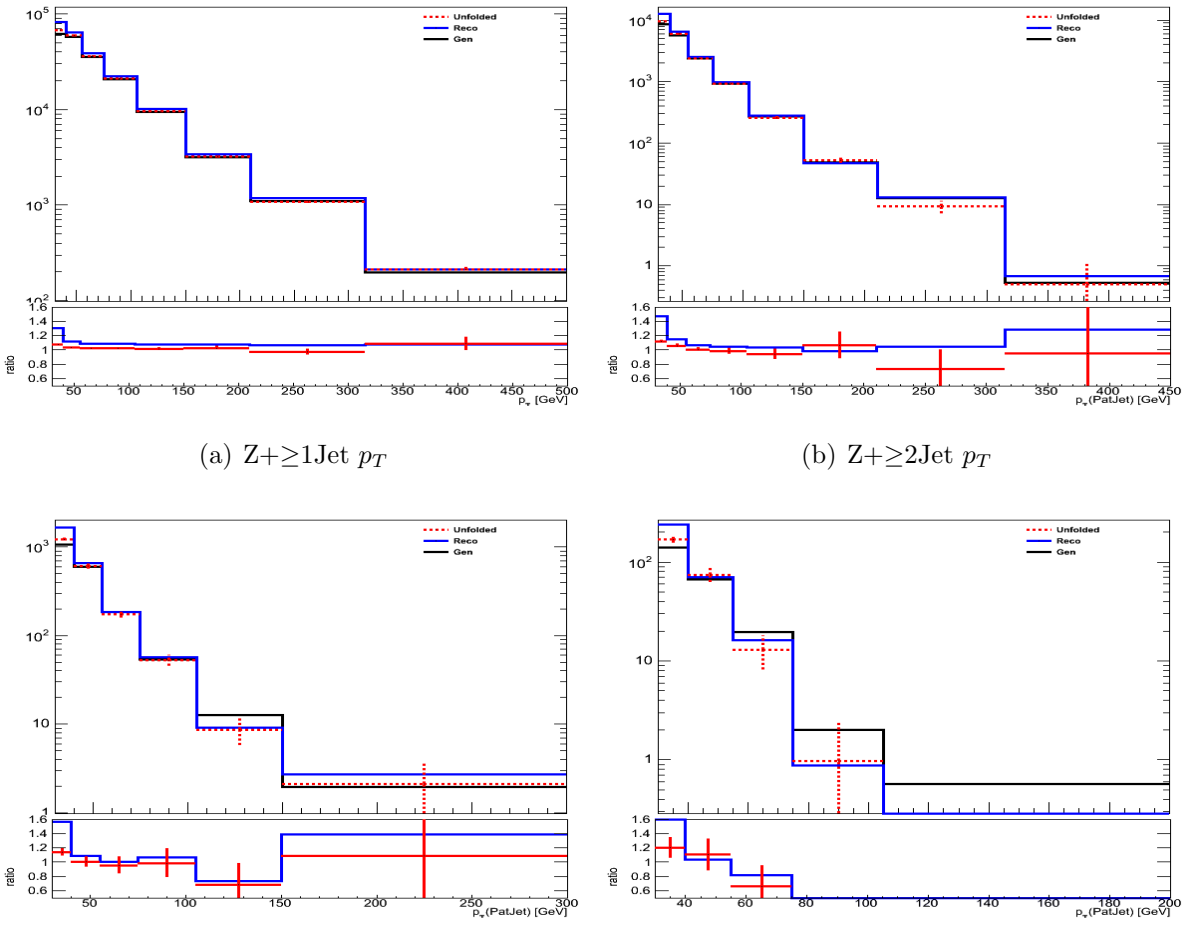


Figure 6.16: Unfolded jet multiplicity spectrum for MC closure test using response matrix from POWHEG sample.



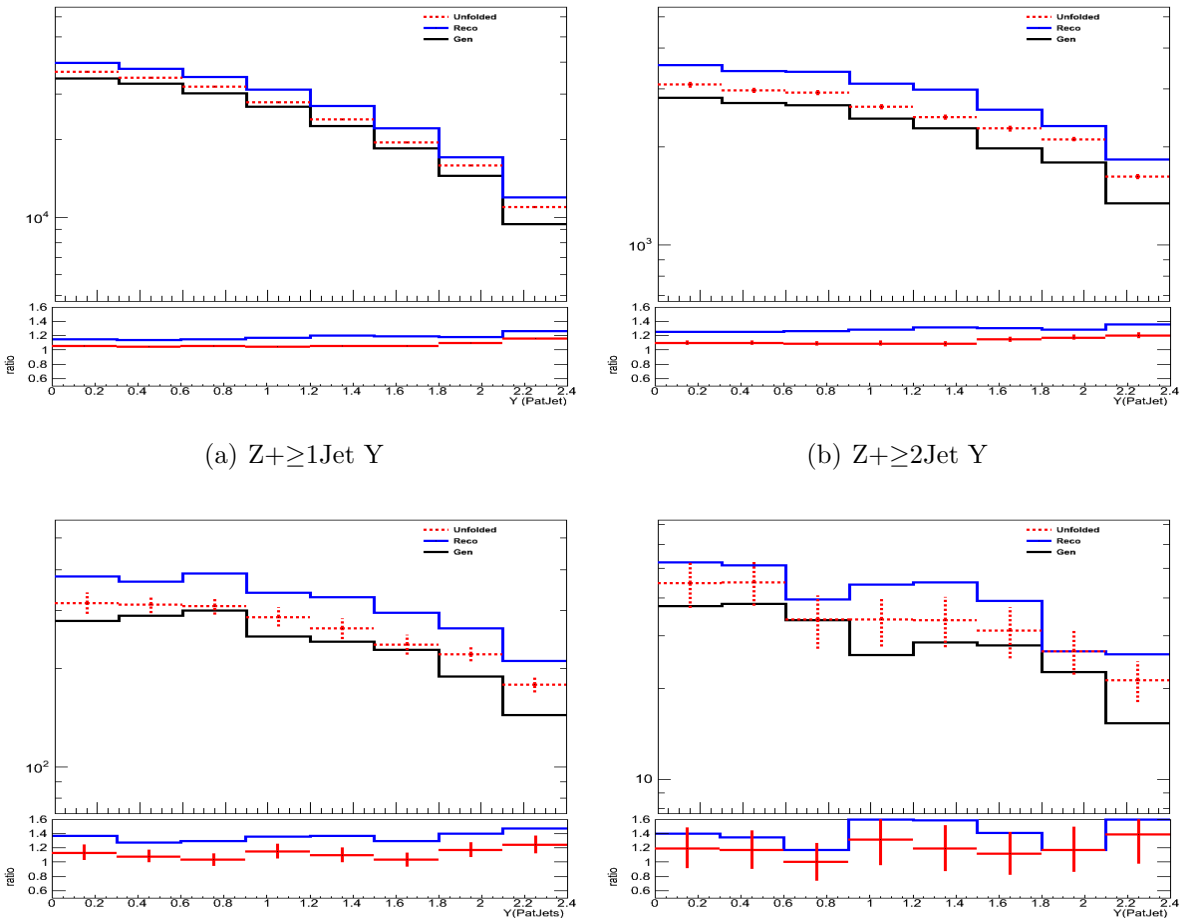
(a) Z+≥1Jet p_T

(b) Z+≥2Jet p_T

(c) Z+≥3Jet p_T

(d) Z+≥4Jet p_T

Figure 6.17: The unfolded Jet PT for MC closure test using response matrix from POWHEG sample.



(a) Z+≥1Jet Y

(b) Z+≥2Jet Y

(c) Z+≥3Jet Y

(d) Z+≥4Jet Y

Figure 6.18: The unfolded Jet Y for MC closure test using response matrix from POWHEG sample.

Chapter 7

Results

In this chapter final measurements and various sources for systematic uncertainties to the differential Z+njet cross-sections are discussed. This chapter has been organized in two parts; Section 7.1 describes the systematic uncertainties and Section 7.2 discusses about the absolute and differential cross-section results in various bins of jet multiplicity obtained after background removal and unfolding procedure discussed in previous chapter.

7.1 Systematic Uncertainties

In the current analysis, we are not very sensitive to the typical uncertainties in lepton energy scale or resolution. So the major part of systematics comes from the jet related variables. Below is list of major sources:

7.1.1 Jet Energy Scale (JES)

The reconstructed energy of jets, all having same true energy, not only show random jet by jet fluctuation, but also a difference between the average of measured jet energy values and the true energy. For example in Z+1 Jet events, there will be a difference between the measured energy of the jet and measured energy of recoiling Z boson. This scale mismatch is corrected by applying the centrally provided jet energy corrections (JEC) as mentioned in the Section 3.4.3. But since the true energy of a jet in data is not a known

priori, such corrections always carry an unavoidable uncertainty that affects the results. We have estimated this affect by re-performing measurement in data with jet energy corrections shifted with $\pm\sigma$, where σ is the total uncertainty on the applied corrections and is centrally provided by CMS [133]. The difference between the measured cross-sections before and after these shifts is taken as a measure of the systematic uncertainty due to the uncertainty on the jet energy scale. The cross-section (Section 7.2) is calculated again to estimate this dependence. Figure 7.1 illustrates the dependence of JES uncertainty on jet multiplicities. The JES is the dominant source of uncertainty in this analysis. Figure 7.2 and Figure 7.3 show the effect of the JES uncertainty on the corrected jet p_T and Y distributions, respectively.

7.1.2 Unfolding

Here we discuss the systematics associated with the unfolding procedure itself. Since the response matrix is derived from MC, there is always scope for uncertainties introduced by imperfect representation of data by the event generator or by imperfections in the simulation. While we make no attempt to evaluate the later, but model dependence on event generators is estimated by comparing the corrections, obtained when we unfold a simulated POWHEG sample with a response matrix trained on the MADGRAPH, with the ones obtained by unfolding a MADGRAPH sample with same matrix. Dependence of this uncertainty along with Jet Energy Resolution (Section 7.1.3) on the corrected jet p_T and Y distributions is shown in Figure 7.2 and Figure 7.3, respectively.

7.1.3 Jet Energy Resolution (JER)

The reconstructed jet energy fluctuates on a jet by jet basis even when the incident jet has always the same energy. The width of this fluctuation is called the jet energy resolution and is a non negligible source of systematic uncertainty if not corrected. In our analysis the unfolding procedure is supposed to take care of this effect, but we still have to be

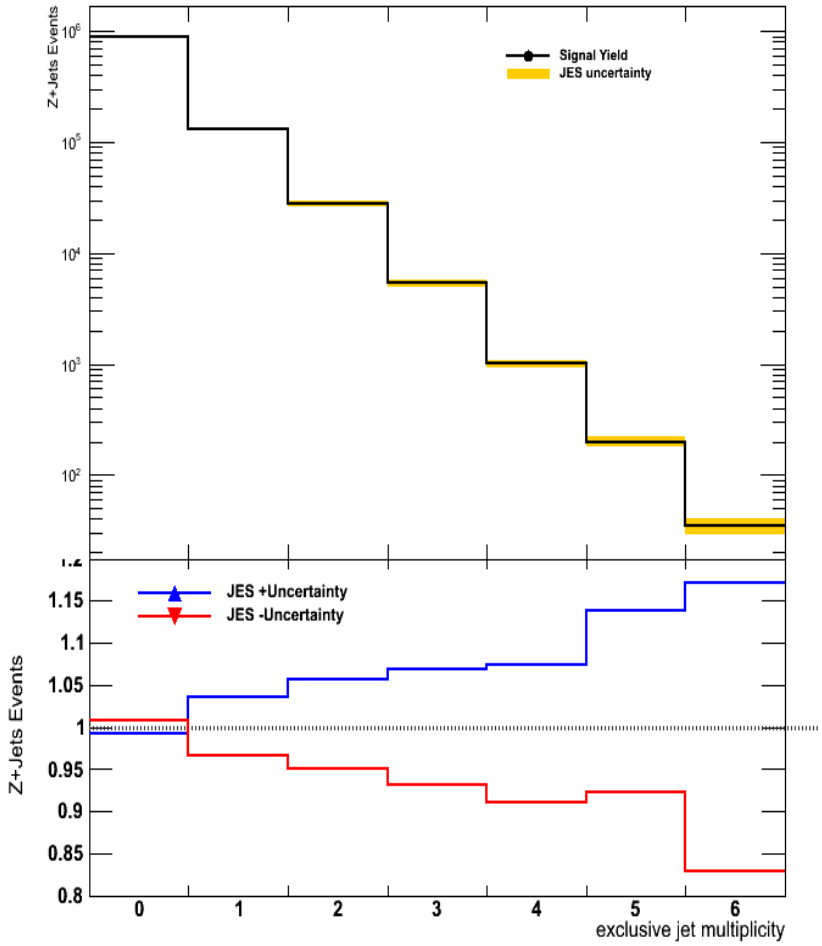


Figure 7.1: JES uncertainties for exclusive jet multiplicity distribution

careful about the fact that jet energy resolution may not be same in data and the MC used to train the response matrix.

JER uncertainty estimation

Jet p_T resolutions and data/MC resolution scaling factors are derived using photon+jet and Z+jet samples by CMS [134]. We estimate the associated systematics by scaling the reconstructed jet p_T based on the p_T difference between matched reconstructed and generated jets as: $p_T \rightarrow max[0., p_{T,gen} + c * (p_T - p_{T,gen})]$ where "c" is the core resolution scaling factor, i.e. the measured data/MC resolution ratio (Table 7.1). And then

compared the unfolded results with those obtained in same sample without this manual smearing. Note that unfolding in both the cases is done using same matrix as the one used for unfolding the data. After this, the cross-section (Section 7.2) is calculated again. Figure 7.2 and Figure 7.3 show the effect of the JER uncertainty on the corrected jet p_T and Y distributions, respectively.

Table 7.1: Data/MC resolution scaling factors [135]

eta bin	Data/MC Ratio
0.0-0.5	1.052
0.5-1.1	1.057
1.1-1.7	1.096
1.7-2.3	1.134
2.3-5.0	1.288

7.1.4 Pileup Uncertainty

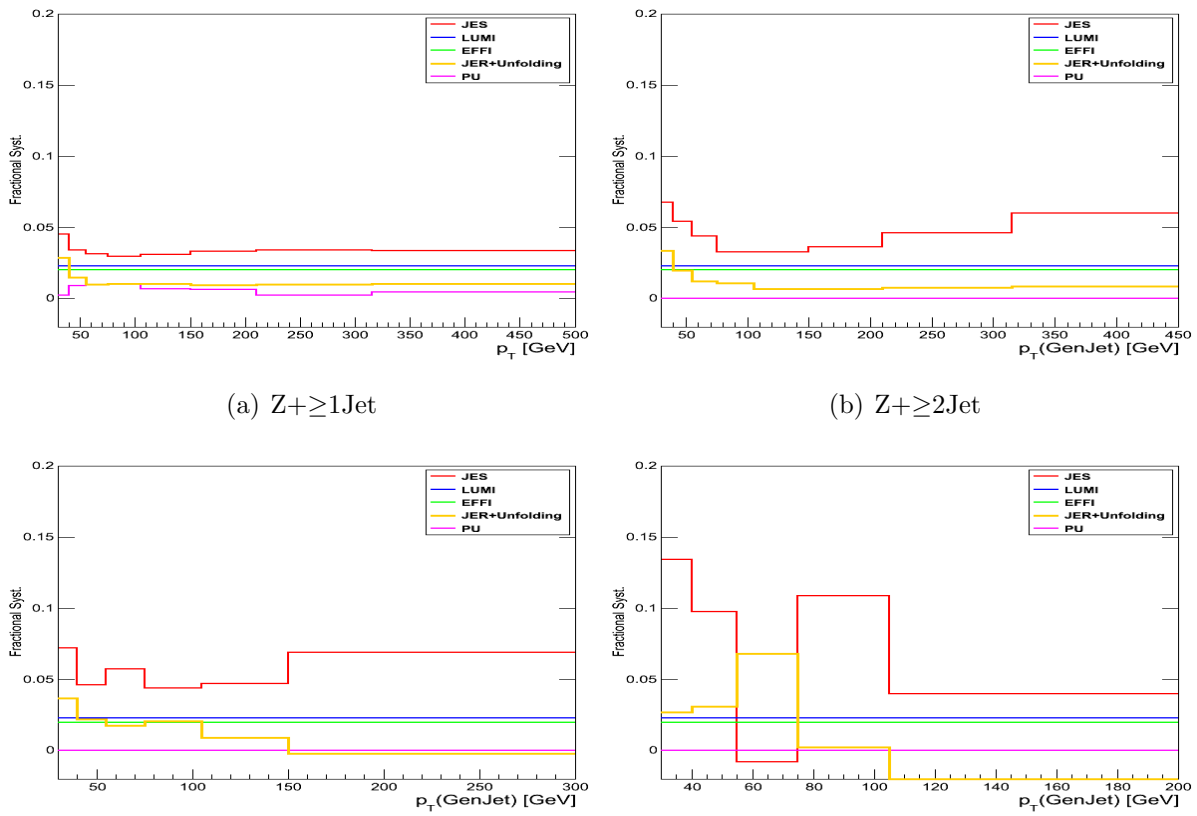
The collected data suffers from the uncertainties related to the pileup effects and the methods used for the mitigation. As the number of interactions in the data are estimated from the measured luminosity in each bunch crossing times an average total inelastic cross section. Thus, resulting into sources of errors for the pileup simulation. The sensitivity of our results on the pileup can be evaluated by $\pm 5\%$ variation in 68mb recommended inelastic cross-section and re-evaluating the Pile-up weight (discussed previously in Section 4.6.3). Finally we do the full analysis again and estimate the dependence of cross-sections (shown in Figure 7.2 and Figure 7.3).

7.1.5 luminosity uncertainty

The uncertainty due to the luminosity measurement is 2.2% [120]. Dependence of this uncertainty on cross-section measurements is shown in Figure 7.2 and Figure 7.3.

7.1.6 Systematic Uncertainty due to Object Selection Efficiencies

The systematic uncertainties of the object based efficiencies as discussed in Chapter 5 are then propagated to the cross sections. In order to consider all the efficiency related errors, the cross-sections are re-evaluated by varying the efficiency numbers by 2% as mentioned earlier in chapter 5. Dependence of this uncertainty on cross-section measurements is shown in Figure 7.2 and Figure 7.3.



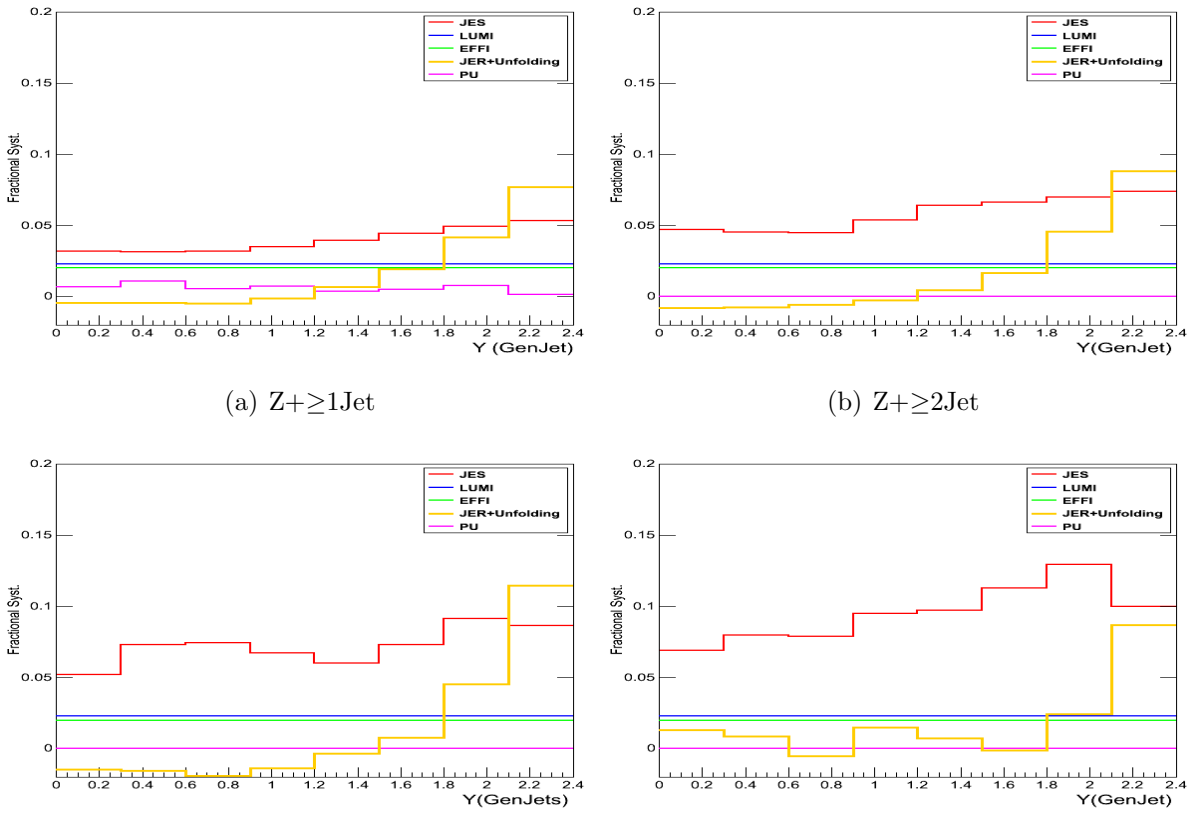
(a) Z+≥1Jet

(b) Z+≥2Jet

(c) Z+≥3Jet

(d) Z+≥4Jet

Figure 7.2: Systematic uncertainties for cross-sections as function of corrected jet p_T



(a) Z+≥1Jet

(b) Z+≥2Jet

(c) Z+≥3Jet

(d) Z+≥4Jet

Figure 7.3: Systematic uncertainties for cross-sections as function of corrected jet Y

7.2 Results

Finally the absolute and differential cross-sections for Z+jets process have been calculated for 4.890 fb^{-1} of data collected by CMS during 2011 at $\sqrt{(}s) = 7\ TeV$. The results so obtained in various bins of jet multiplicity are described in the following sections.

7.2.1 Absolute Cross-sections

The $Z/\gamma* \rightarrow e^+e^-$ inclusive cross section is measured to provide a basic cross check for some of the techniques used in the final measurement of the $Z/\gamma* \rightarrow e^+e^-+ \geq$ n jet cross sections. In order to determine the inclusive cross section times branching fraction into

electrons, the following equation is evaluated:

$$(\sigma \times B.R)^{Z+j} = \frac{S_j}{\langle \epsilon_{ee}(j) \rangle \mathcal{A} \times \mathcal{L}} \tag{7.1}$$

Here $\mathcal{A}$ and $\mathcal{L}$ represent acceptance and integrated luminosity respectively. The $\langle \epsilon_{ee}(j) \rangle$ represents average signal selection efficiency per event in j^{th} bin. The ϵ_{ee} is determined using Madgraph signal MC duly corrected for the differences with respect to the data. This is given in Table 7.2. All the efficiencies are examined for jet multiplicity dependence.

Table 7.2: Corrected MC efficiencies. The errors are obtained by shifting scale factors by $\pm\sigma$

.

N_{Jet}	0	1	2	$\geq$3
ϵ_{ee}	0.55$\pm$0.01	0.55$\pm$0.01	0.54$\pm$0.01	0.53$\pm$0.02

Acceptance

The acceptance $\mathcal{A}$ of the detector is the fraction of events theoretically able to be detected within it. We used Madgraph signal MC with detector simulation to estimate the acceptance for the fiducial and kinematic cuts. The acceptance numerator counts the number of events satisfying the following requirements at the detector reconstructed level: $|\eta| < 1.442$ or $1.566 < |\eta| < 2.5$, and transverse energies above 20 GeV, and Dielectron invariant mass cut: 70 GeV $<$ Mee $<$ 110 GeV. The acceptance denominator counts the number of events with generated Zee particles that are within the di-electron invariant mass window. The acceptance for inclusive Zee is estimated to be: 0.42

Absolute cross-section calculation

After applying all the corrections, based on the integrated luminosity ($4.890 fb^{-1}$), the inclusive Z/$\gamma*$ production cross-section times branching fraction into electrons is calculated. Table 7.3 represents the measurements of the Z/$\gamma * (\rightarrow e^+e^-)$ production cross-section

for different exclusive jet multiplicities and total inclusive Z/$\gamma*$ production cross-section times branching fraction into electrons is found to be in agreement with CMS results from 2010 data [137].

Table 7.3: The values of $\sigma \times B.R$ calculated from the data. The quoted uncertainties are statistical, with contribution from efficiency errors included.

j	S_j	$\langle\epsilon_{ee}\rangle(j)$	$\sigma \times B.R$ $(\mathcal{A}_{Full})(pb)$	$\sigma \times B.R$ $(\mathcal{A}_{Det})$
$= 0$	906824±1387	0.55±0.01	802.7±14.6	337.2±6.15
$= 1$	118928±582	0.55±0.01	105.3±1.98	44.2±0.83
$= 2$	24059±264	0.54±0.01	21.69±0.46	9.11±0.19
$= 3$	4606±100	0.53±0.005	4.231±0.10	1.77±0.04
$= 4$	885±42	0.53±0.005	0.813±0.04	0.34±0.01
Total			934.81±14.5	392.6±6.1

7.2.2 Differential Cross-sections

The inclusive jet differential cross sections as a function of a given variable α are defined as:

$$\frac{d\sigma}{d\alpha} = \frac{1}{\mathcal{L}} \times \frac{N_{jets}^{corr}}{\Delta\alpha}, \alpha \equiv p_T^{jet}, Y^{jet} \tag{7.2}$$

where N_{jets}^{corr} denotes the background subtracted event yield in a given bin in α, $\Delta\alpha$ is the size of the bin and $\mathcal{L}$ is the luminosity. The jet kinematic distributions obtained after unfolding step are now used for obtaining the differential cross sections as function of p_T ($\frac{d\sigma}{dp_T^{n^{th}Jet}}$) and Y ($\frac{d\sigma}{dY^{n^{th}Jet}}$) of n^{th} jet. For this, the jet p_T and Y distributions are bin by bin divided by the integrated luminosity and also the corresponding bin width. Measurements are performed in well defined kinematic regions for the leptons and jets in the final state. Acceptance corrections extrapolating the measurements to the full phase space are not applied in presented results. The measurements are also compared to the predictions from LO plus parton shower Monte Carlo generators.

The results so obtained for the inclusive jet differential cross section in Z+Njets production for Njets$\geq$ 1, 2, 3 and 4, respectively, as a function of p_T of leading jet with $p_T > 30$ GeV and $|\eta|$ <2.4; are given in Tables 7.4, 7.5, 7.6 and 7.7. Figures 7.4, 7.5, 7.6

Table 7.4: The measured $d\sigma/dp_T^{1^{st}Jet}$×B.R($Z \to e^+e^-$) for Z+≥1 jet production with total systematic and statistical uncertainties

p_T^{Jet} (GeV)	Bin center (GeV)	$\frac{d\sigma}{dp_T^{Jet}} \pm (stat.) \pm (syst)$ (fb/(GeV/c))
30 - 40	35	$1945.76 \pm 39.868^{+106.818}_{-102.798}$
40 - 55	47.5	$1006.76 \pm 21.3007^{+43.4852}_{-47.8857}$
55 - 75	65	$495.739 \pm 11.2652^{+22.79}_{-19.6845}$
75 - 105	90	$208.768 \pm 5.0007^{+8.57968}_{-9.29314}$
105 - 150	127.5	$71.4061 \pm 1.90747^{+3.07369}_{-2.88938}$
150 - 210	180	$19.4257 \pm 0.715191^{+0.920646}_{-0.925027}$
210 - 315	262.5	$4.56391 \pm 0.177674^{+0.207159}_{-0.22737}$
315 - 500	407.5	$0.549841 \pm 0.0406062^{+0.0215273}_{-0.019757}$

and 7.7 depict the inclusive jet differential cross section in Z+Njets production for Njets≥ 1, 2, 3 and 4, respectively, as a function of p_T of leading jet with $p_T > 30$ GeV and $|\eta| < 2.4$.

The measurements are compared to the Madgraph matrix element LO generator shown in each cross section distribution and are showing a good agreement between data and theory within their uncertainties. Every source of uncertainty (Section 7.1) is treated as independent from each other. The differential jet cross sections as a function of Y for Njets≥1, 2, 3 and 4 are represented in Figures 7.8, 7.9, 7.10 and 7.11. The good agreement between the data and the theoretical predictions is observed in this case also. Results have been given in Tables 7.8, 7.9, 7.10 and 7.11.

Table 7.5: The measured $d\sigma/dp_T^{2ndJet}$×B.R($Z \rightarrow e^+e^-$) for Z+≥2 jet production with total systematic and statistical uncertainties

p_T^{Jet} (GeV)	Bin center (GeV)	$\frac{d\sigma}{dp_T^{Jet}} \pm (stat.) \pm (syst)$ (fb/(GeV/c))
30 - 40	35	$517.518 \pm 10.8765^{+38.5068}_{-35.545}$
40 - 55	47.5	$206.1 \pm 5.18085^{+12.5995}_{-12.0027}$
55 - 75	65	$84.2334 \pm 2.32764^{+4.63664}_{-4.11697}$
75 - 105	90	$29.7532 \pm 1.12976^{+1.22198}_{-1.38705}$
105 - 150	127.5	$8.69581 \pm 0.413691^{+0.430346}_{-0.402443}$
150 - 210	180	$2.99285 \pm 0.21378^{+0.122888}_{-0.105022}$
210 - 315	262.5	$0.496426 \pm 0.0549734^{+0.01936}_{-0.030072}$
315 - 450	382.5	$0.0906596 \pm 0.0204194^{+0.0133732}_{-0.00550597}$

Table 7.6: The measured $d\sigma/dp_T^{3rdJet}$×B.R($Z \rightarrow e^+e^-$) for Z+≥3 jet production with total systematic and statistical uncertainties

p_T^{Jet} (GeV)	Bin center (GeV)	$\frac{d\sigma}{dp_T^{Jet}} \pm (stat.) \pm (syst)$ (fb/(GeV/c))
30 - 40	35	$133.004 \pm 3.30783^{+11.7085}_{-12.0082}$
40 - 55	47.5	$43.8953 \pm 1.65653^{+2.36192}_{-2.66324}$
55 - 75	65	$14.9332 \pm 0.834067^{+0.958082}_{-0.651697}$
75 - 105	90	$3.77943 \pm 0.331109^{+0.196708}_{-0.254594}$
105 - 150	127.5	$0.833325 \pm 0.134723^{+0.0407207}_{-0.0985635}$
150 - 300	225	$0.0681704 \pm 0.0166145^{+0.00665458}_{-0.00239628}$

Table 7.7: The measured $d\sigma/dp_T^{4^{th}Jet}$×B.R($Z \to e^+e^-$) for Z+≥4 jet production with total systematic and statistical uncertainties

p_T^{Jet} (GeV)	Bin center (GeV)	$\frac{d\sigma}{dp_T^{Jet}} \pm (stat.) \pm (syst)$ (fb/(GeV/c))
30 - 40	35	$29.2173 \pm 1.12263^{+3.04426}_{-2.968}$
40 - 55	47.5	$8.70114 \pm 0.838472^{+0.948246}_{-0.836236}$
55 - 75	65	$2.66646 \pm 0.320072^{+0.0688438}_{-0.154029}$
75 - 105	90	$0.307675 \pm 0.0778895^{+0.0470819}_{-0.0143773}$
105 - 200	152.5	$0.0429226 \pm 0.0143807^{+0.0012442}_{-0.00122329}$

Table 7.8: The measured $d\sigma/dY^{1^{st}Jet}$×B.R($Z \to e^+e^-$) for Z+≥1 jet production with total systematic and statistical uncertainties

$\lvert Y^{Jet}\rvert$	Bin center	$\frac{d\sigma}{d\lvert Y^{Jet}\rvert} \pm (stat.) \pm (syst)$ (fb)
0 - 0.3	0.15	$31881 \pm 670.711^{+1428.19}_{-1401.25}$
0.3 - 0.6	0.45	$31577.4 \pm 676.843^{+1399.1}_{-1354.94}$
0.6 - 0.9	0.75	$28983.6 \pm 613.388^{+1281.07}_{-1241.22}$
0.9 - 1.2	1.05	$26275 \pm 571.265^{+1243.97}_{-1204.72}$
1.2 - 1.5	1.35	$22711.4 \pm 486.578^{+1161.77}_{-1131.16}$
1.5 - 1.8	1.65	$19034.6 \pm 419.697^{+1055.1}_{-981.787}$
1.8 - 2.1	1.95	$15445.4 \pm 332.513^{+908.029}_{-841.007}$
2.1 - 2.4	2.25	$11135.3 \pm 254.313^{+706.64}_{-770.506}$

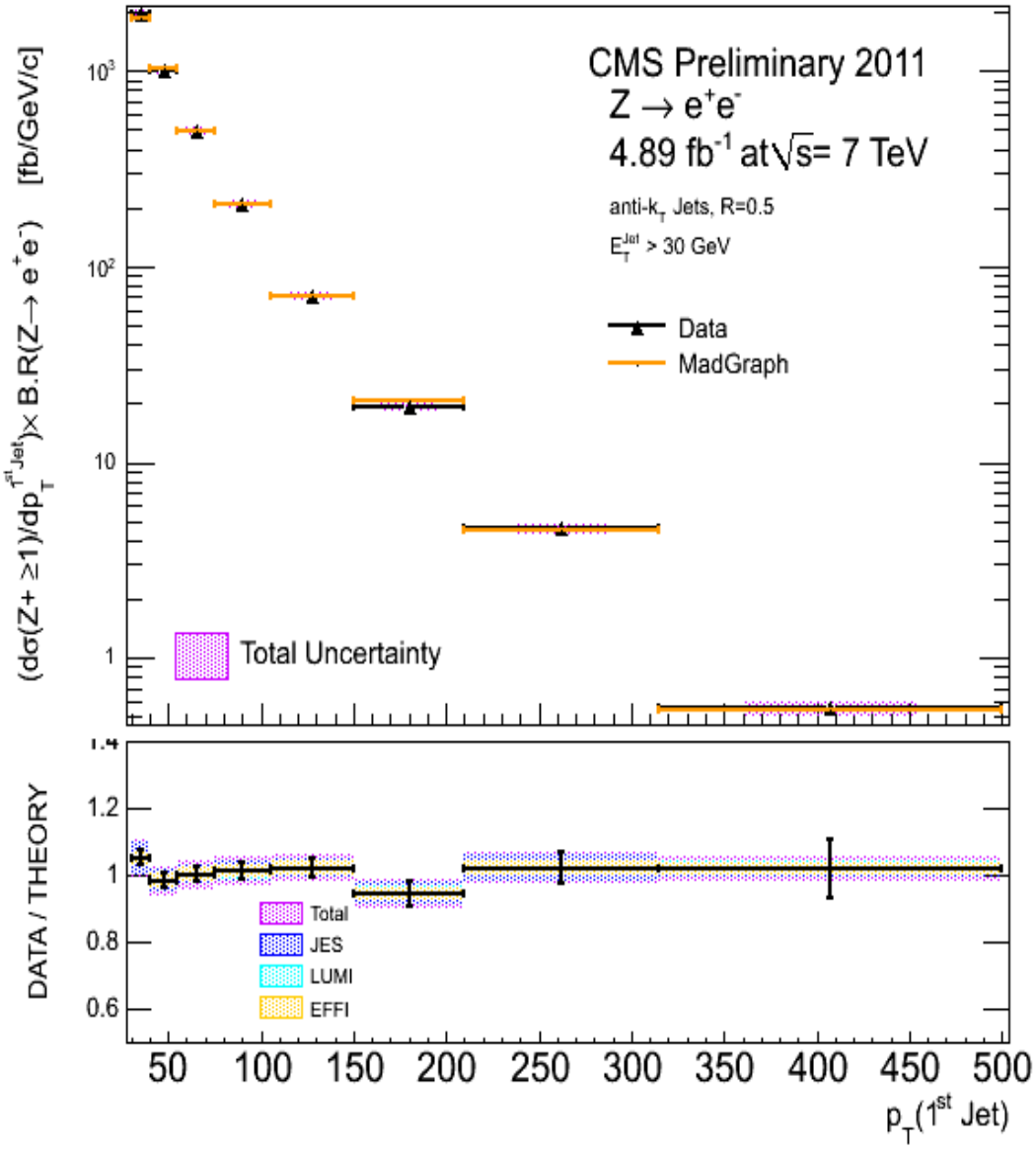


Figure 7.4: The measured $d\sigma/dp_T^{1^{st}Jet}$ ×B.R($Z \to e^+e^-$) for Z+≥1 jet production (black triangles) compared to Madgraph predictions (yellow lines). The violet band is the total systematics uncertainty that has been evaluated by summing in quadrature all the components shown in the legend (bottom). The bottom plot shows a data/theory comparison.

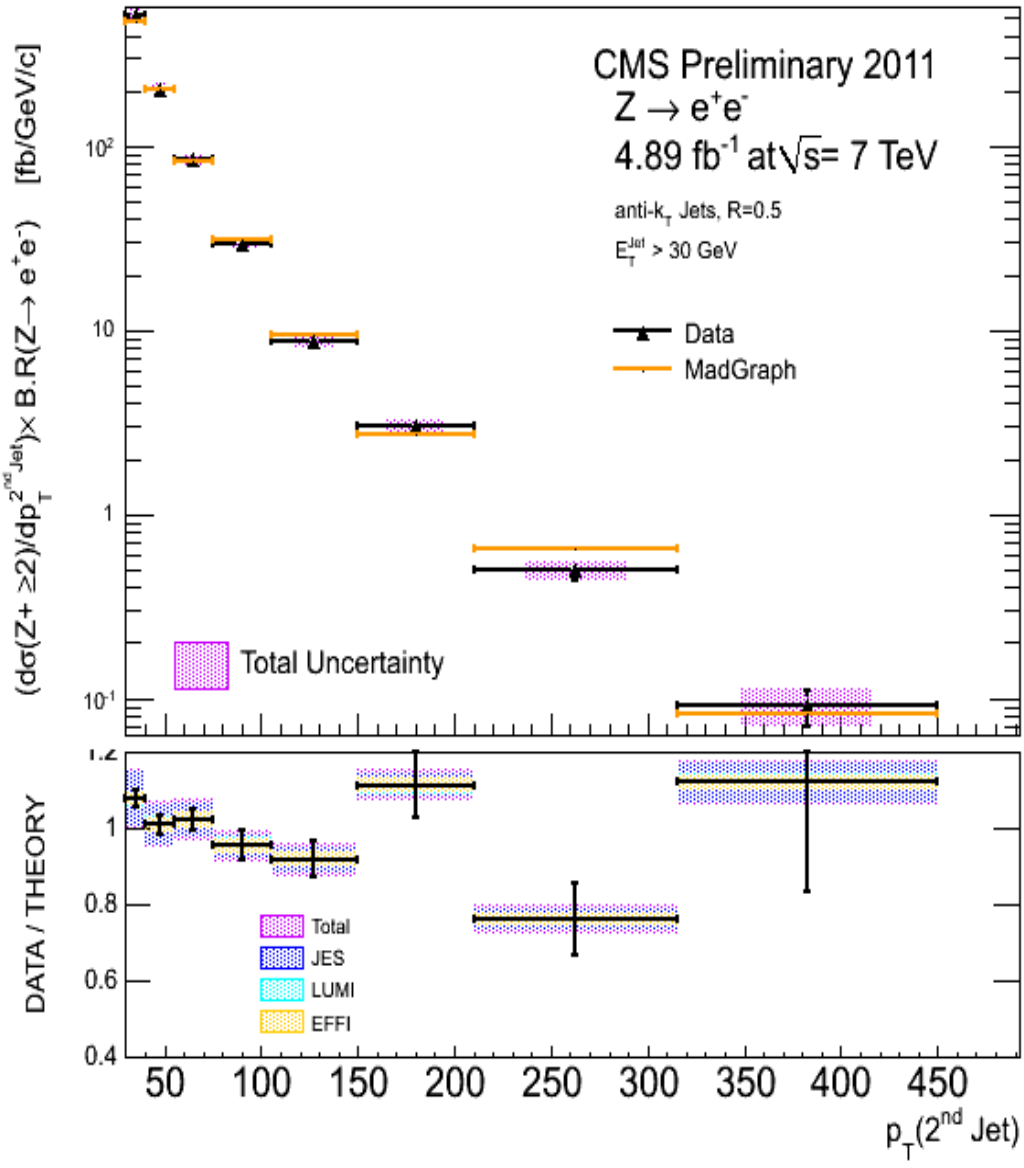


Figure 7.5: The $d\sigma/dp_T^{2^{nd}Jet}$×B.R($Z \rightarrow e^+e^-$) for Z+≥2 jet production. The total systematics has been evaluated by summing in quadrature all the components shown in the legend. The bottom plot shows a comparison with the Madgraph predictions.

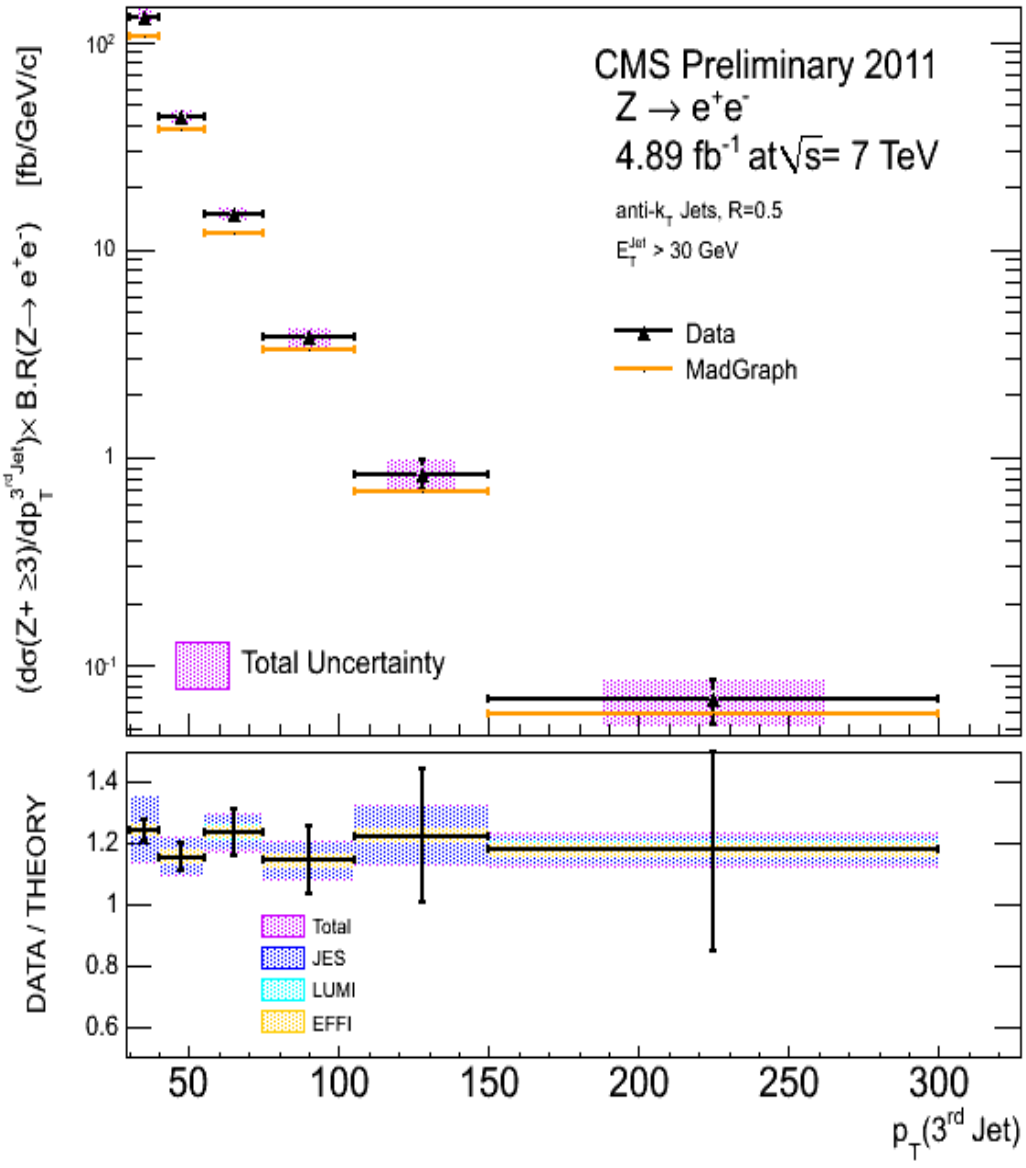


Figure 7.6: The $d\sigma/dp_T^{3rdJet}$×B.R($Z \rightarrow e^+e^-$) for Z+≥3 jet production. The total systematics has been evaluated by summing in quadrature all the components shown in the legend. The bottom plot shows a comparison with the Madgraph predictions.

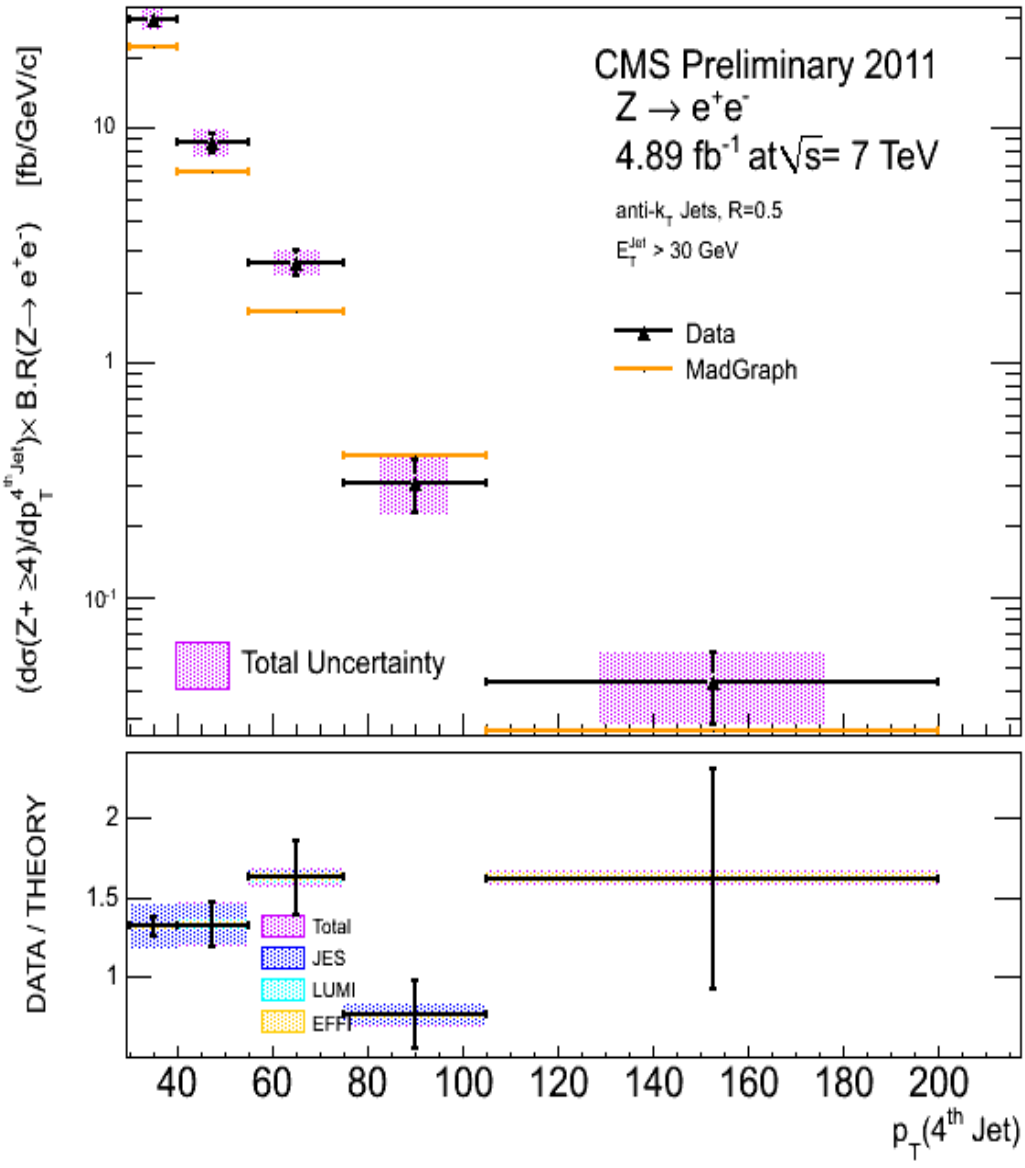


Figure 7.7: The $d\sigma/dp_T^{4^{th}Jet}$ ×B.R($Z \rightarrow e^+e^-$) for Z+≥4 jet production. The total systematics has been evaluated by summing in quadrature all the components shown in the legend. The bottom plot shows a comparison with the Madgraph predictions.

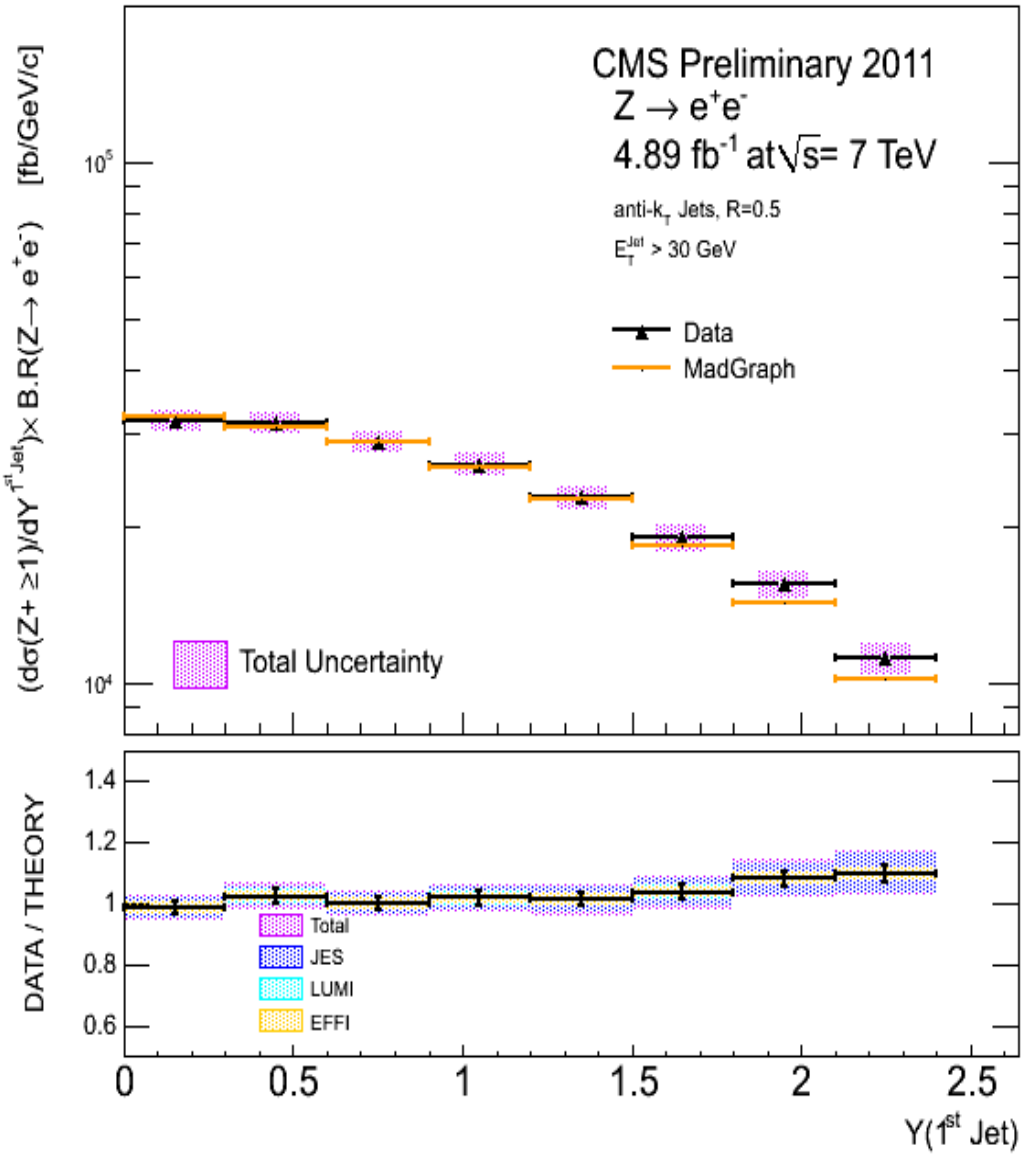


Figure 7.8: The measured $d\sigma/dY^{1^{st}Jet}$×B.R($Z \to e^+e^-$) for Z+≥1 jet production (black triangles) compared to Madgraph predictions (yellow lines). The violet band is the total systematics uncertainty that has been evaluated by summing in quadrature all the components shown in the legend (bottom). The bottom plot shows a data/theory comparison.

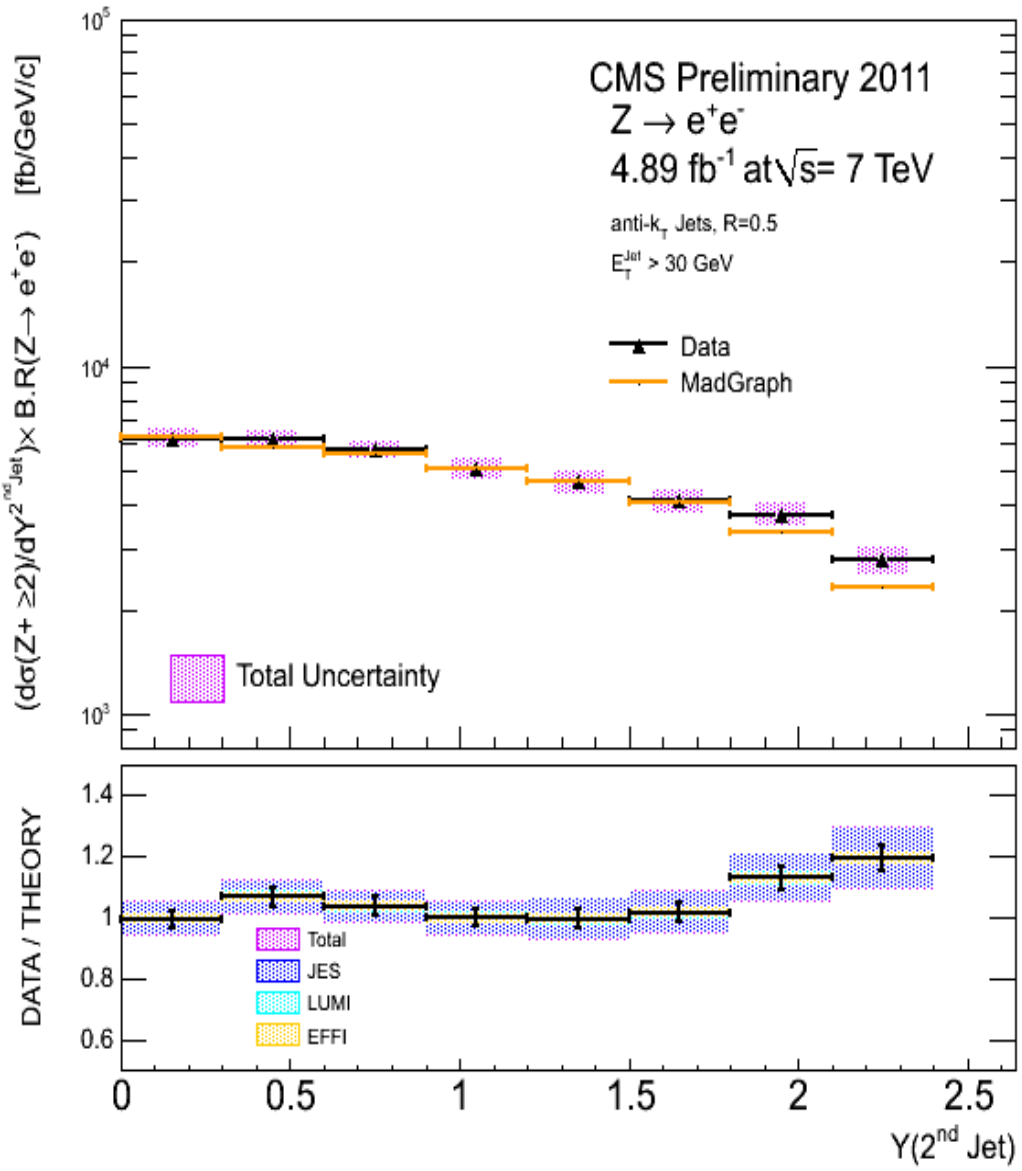


Figure 7.9: The measured $d\sigma/dY^{2^{ndJet}}$×B.R($Z \to e^+e^-$) for Z+≥2 jet production (black triangles) compared to Madgraph predictions (yellow lines). The violet band is the total systematics uncertainty that has been evaluated by summing in quadrature all the components shown in the legend (bottom). The bottom plot shows a data/theory comparison.

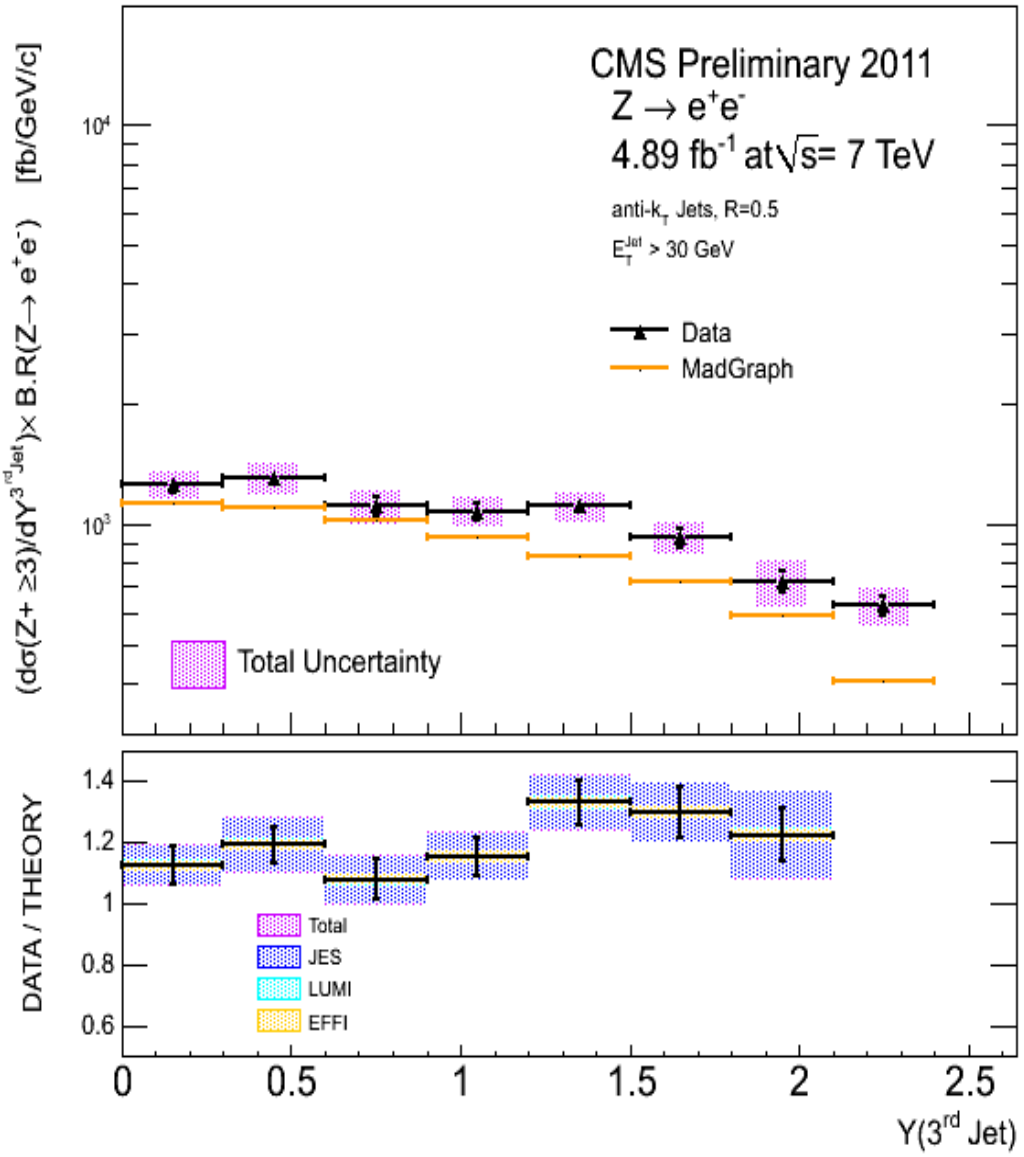


Figure 7.10: The measured $d\sigma/dY^{3^{rd}Jet}$ ×B.R($Z \to e^+e^-$) for Z+≥3 jet production (black triangles) compared to Madgraph predictions (yellow lines). The violet band is the total systematics uncertainty that has been evaluated by summing in quadrature all the components shown in the legend (bottom). The bottom plot shows a data/theory comparison.

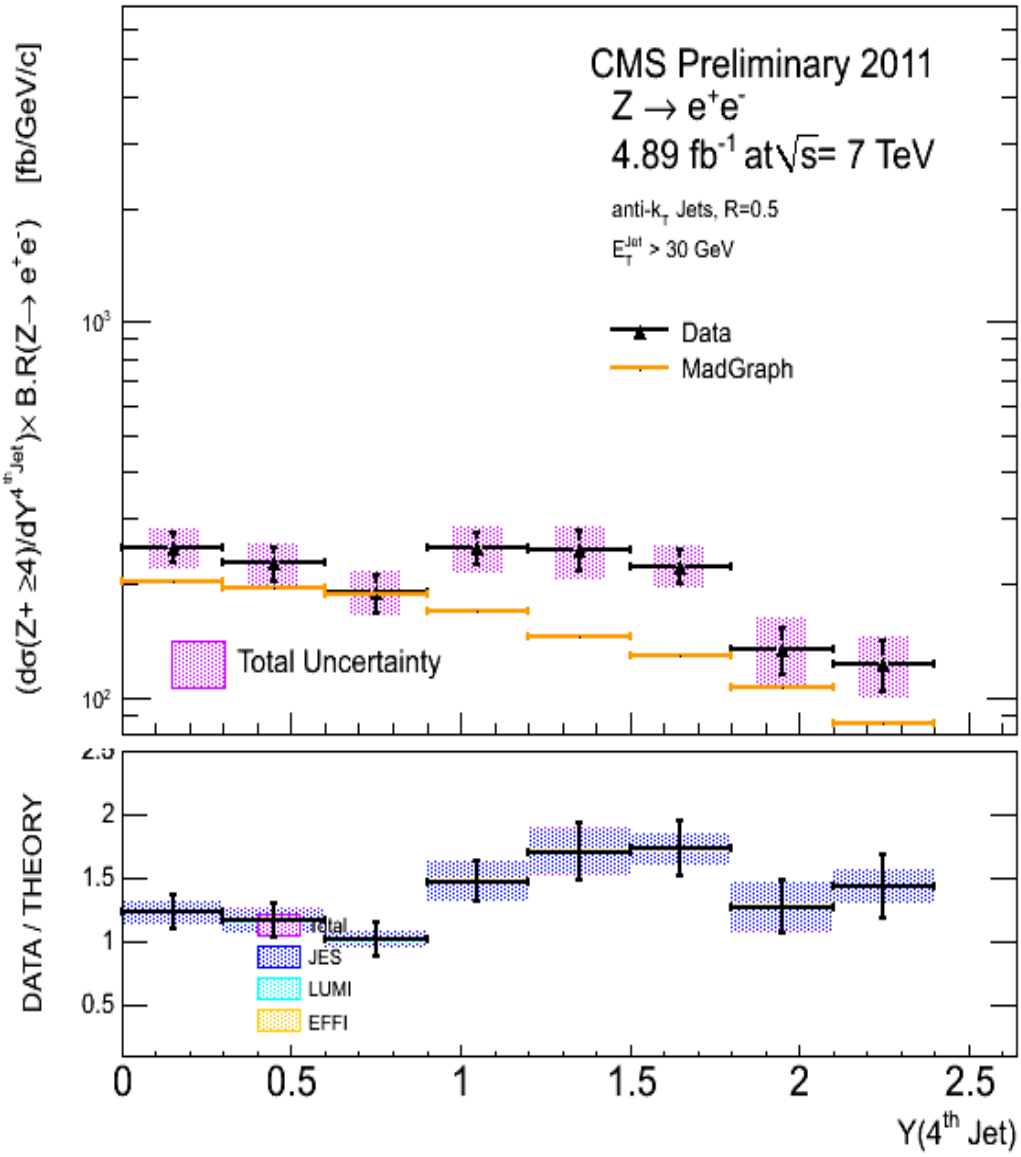


Figure 7.11: The measured $d\sigma/dY^{4^{th}Jet}$ ×B.R($Z \to e^+e^-$) for Z+≥4 jet production (black triangles) compared to Madgraph predictions (yellow lines). The violet band is the total systematics uncertainty that has been evaluated by summing in quadrature all the components shown in the legend (bottom). The bottom plot shows a data/theory comparison.

Table 7.9: The measured $d\sigma/dY^{2^{ndJet}}$×B.R($Z \rightarrow e^+e^-$) for Z+≥2 jet production with total systematic and statistical uncertainties

$\lvert Y^{Jet}\rvert$	Bin center	$\frac{d\sigma}{d\lvert Y^{Jet}\rvert} \pm (stat.) \pm (syst)$ (fb)
0 - 0.3	0.15	$6207.88 \pm 163.155^{+359.665}_{-356.169}$
0.3 - 0.6	0.45	$6204.31 \pm 171.213^{+351.647}_{-327.663}$
0.6 - 0.9	0.75	$5794.51 \pm 158.769^{+305.291}_{-311.79}$
0.9 - 1.2	1.05	$5093.4 \pm 138.099^{+322.57}_{-274.756}$
1.2 - 1.5	1.35	$4648.04 \pm 139.348^{+364.092}_{-293.766}$
1.5 - 1.8	1.65	$4102.19 \pm 125.481^{+296.471}_{-291.917}$
1.8 - 2.1	1.95	$3755.18 \pm 123.599^{+290.979}_{-252.823}$
2.1 - 2.4	2.25	$2764.81 \pm 89.8539^{+229.75}_{-238.583}$

Table 7.10: The measured $d\sigma/dY^{3^{rdJet}}$×B.R($Z \rightarrow e^+e^-$) for Z+≥3 jet production with total systematic and statistical uncertainties

$\lvert Y^{Jet}\rvert$	Bin center	$\frac{d\sigma}{d\lvert Y^{Jet}\rvert} \pm (stat.) \pm (syst)$ (fb)
0 - 0.3	0.15	$1262.89 \pm 62.297^{+68.6835}_{-88.5208}$
0.3 - 0.6	0.45	$1309.18 \pm 54.1063^{+115.621}_{-88.2146}$
0.6 - 0.9	0.75	$1111.55 \pm 60.5028^{+89.116}_{-79.3835}$
0.9 - 1.2	1.05	$1078 \pm 54.4976^{+84.1188}_{-64.0618}$
1.2 - 1.5	1.35	$1111.36 \pm 50.5928^{+68.1508}_{-85.25}$
1.5 - 1.8	1.65	$929.143 \pm 54.3431^{+68.2077}_{-73.803}$
1.8 - 2.1	1.95	$722.368 \pm 45.1351^{+85.8151}_{-84.2437}$
2.1 - 2.4	2.25	$627.474 \pm 38.109^{+48.9048}_{-70.486}$

Table 7.11: The measured $d\sigma/dY^{4^{thJet}}$×B.R($Z \to e^+e^-$) for Z+≥4 jet production with total systematic and statistical uncertainties

$\lvert Y^{Jet}\rvert$	Bin center	$\frac{d\sigma}{d\lvert Y^{Jet}\rvert} \pm (stat.) \pm (syst)$ (fb)
0 - 0.3	0.15	$251.065 \pm 23.1106^{+16.3504}_{-21.6388}$
0.3 - 0.6	0.45	$227.255 \pm 23.5016^{+20.9466}_{-17.5586}$
0.6 - 0.9	0.75	$190.427 \pm 22.1871^{+12.1708}_{-14.1026}$
0.9 - 1.2	1.05	$250.262 \pm 23.4852^{+28.3652}_{-26.1389}$
1.2 - 1.5	1.35	$246.51 \pm 29.5171^{+20.7578}_{-32.6287}$
1.5 - 1.8	1.65	$223.539 \pm 23.8963^{+22.4735}_{-9.59334}$
1.8 - 2.1	1.95	$134.178 \pm 19.6644^{+23.2217}_{-18.0814}$
2.1 - 2.4	2.25	$122.21 \pm 18.7221^{+8.0449}_{-15.188}$

Chapter 8

Summary and Conclusions

The main result of this thesis is a measurement of the differential production cross section of one, two, three and four jets in association with the Z boson as a function of p_T and Y of leading jets. The end-states considered for this study consisted of ones where the Z boson decay leptonically into an electron positron pair. The measurements are based on an integrated luminosity of 4.89 fb^{-1} collected by the CMS detector at the CERN, LHC collider during Run2011A and Run2011B scenario of high energy proton proton collisions at $\sqrt{s} = 7\ TeV$.

The physics motivation for the study of Z+Jet(s) production can hardly be overestimated. This measurement not only serves as a stringent test of the perturbative QuantumChromodynamis (pQCD) but is also sensitive to the production of new particles which decay into Z+Jet final states. Also due to relatively large production cross section and clean signal in leptonic decay modes, it is possible to perform measurements with a high degree of precision. Such measurements can then serve as important benchmarks for testing and tuning of the various Monte Carlo event generators. Finally, but equally important is the fact that these processes often constitute dominant background for several new physics searches.

Current study begins by selecting the events having at-least two well reconstructed electrons. An electron is reconstructed from its energy deposits inside the electro-

magnetic calorimeter and the tracks it leaves inside the tracker. Naturally, we are limited by the geometric acceptance of the ECAL, which is defined by $|\eta|$ <2.4 with an exclusion of the uninstrumented gaps in the $1.442 < |\eta| < 1.556$ region. To ensure that we are using well reconstructed electrons, we further apply several criteria including a minimum threshold on the transverse momentum p_T of the electrons and some identification requirements on shower shape topology. The identification consists of constraints based on spatial matching between the ECAL supercluster and the track in silicon detector within the η and ϕ coordinates, on the supercluster energy distribution in η direction and on the energy leakage into the HCAL detector. Also to reduce the contamination from converted photons, a minimal track transverse impact parameter significance is required. If there are no associated track hits found in the first tracker layers or if a conversion partner candidate is found, the electrons are rejected from the selection. Finally we ensure that we are not using the electrons from within the jets by requiring them to be isolated. This is based on the three isolation variables which compute the energy deposits of stable particles of any type (e.g. photon, neutral or charged hadron, etc). The momentum around the electron candidate is summed within a cone of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.4$ where $\Delta\eta$ and $\Delta\phi$ are the differences in the pseudo-rapidity and azimuthal angle between the electron and energy deposit, lepton itself is excluded from the summation. We then define the relative combined isolation as $\Sigma p_T(charged\ hadrons)+\Sigma p_T(neutral\ hadrons)+\Sigma p_T(\ photons)/p_T$, where first sum runs over the transverse momenta of all particle flow charged hadrons and second is over the transverse momenta of all neutral hadrons and third over all the photons within the cone. The Z bosons are reconstructed from the two selected leading electrons, and the events are kept if the invariant mass of the pair lied between $70 \leq M_{ee} \leq 110$. The selected events are then classified into the bins of jet multiplicity, by counting the jets in the event. The final results refer to hadron level jets with p_T of jet $> 30\ GeV/c^2$ and $|\eta| < 2.4$ and reconstructed using anti-kT clustering algorithm with a size parameter R=0.5, from the particle collection created with the particle flow algorithm. Jet energy

corrections (JEC) are applied to account for the jet energy response as a function of η and p_T. The pileup and underlying events, if not taken care of can potentially affect the multiplicity and also p_T, Y spectrum of the jets, although a relatively high jet p_T threshold of 30 GeV/c^2 greatly reduces the probability of a jet from pileup to enter the counting. Among the selected jets, the contribution from pileup is subtracted by removing all the charged particles who can not be associated with the leading primary vertex. Finally since the electrons can appear inside the jet list too, we avoid them by choosing only those jets, which are $\Delta R \geq 0.3$ away from ($Z \rightarrow$ ee) electrons in the $\eta - \phi$ space.

The results are not corrected for the detector acceptance, but rather quoted them within the acceptance, as defined by the lepton and jet fiducial and kinematic selections given above and this is helpful in providing model-independent results. The efficiencies for on-line event selection, offline lepton reconstruction, and identification are evaluated from data using Z→ee events. For this, we used a sample which is triggered using two electrons, one of which is identified using very loose criteria. And during offline analysis, in every event we require the invariant mass of di-electron system to fall in the range 60-120 GeV. Then one of the electrons is required to pass full selection criteria and have a match with the tighter leg of the trigger. This is called a tag. The other electron candidate is selected with criteria which depend on efficiency being measured, and this is called a probe. This sample is then divided into two exclusive sub-samples depending on whether the probe lepton passes or fails the selection criteria currently under investigation. Due to the presence of events from backgrounds, the signal yields have to be obtained via a fit to invariant mass distribution of the di-electron system. The measured efficiency is then deduced from relative yields of signal in sub-samples with passing or failing probes. This procedure is performed separately in different bins of probe pseudo-rapidity (η) to obtain results as function of probe η. Finally the data to Monte Carlo scale factors are deduced by dividing efficiencies in data to the ones obtained from MC using exactly same procedure. We use scale factors instead of raw efficiencies, in order to benefit from

possible cancellation of systematic uncertainties associated with the procedure. In order to estimate the behaviour of the jets at the particle-level jet, we apply an unfolding procedure that removes the effects of jet energy resolution and reconstruction efficiency. This study uses the Iterative Bayesian Method. One of the dominant sources of systematic uncertainties in the V+jets measurements is the determination of the jet energy, which affects the jet counting. We analysed the Jet energy scale, Jet energy resolution, pileup , luminosity, selection efficiencies and unfolding as major sources of systematics.

Finally we present absolute and differential cross-section results in various bins of jet multiplicity obtained after background removal and unfolding procedure. The $Z/\gamma* \rightarrow e^+e^-$ inclusive cross section is measured to provide a basic cross check for some of the techniques used in the final measurement of the $Z/\gamma* \rightarrow e^+e^- + \geq$ n jet cross sections. We present the Z+nJets, for n=1 to 4, cross-section measurements as a function of jet p_T and Y. The results obtained in the present study are given in Tables 7.4 to 7.11. The data measurements are compared to LO pQCD predictions, which are computed using the Madgraph generator with CTEQ6L1 PDFs, with the renormalization and factorization scales set to $\mu^2 = (M_Z)^2 + (p_T^Z)^2$, and using a anti-kT algorithm with R = 0.5 to reconstruct jets at the parton level. It is observed that the measured cross sections are well described by Madgraph predictions normalized to NNLO cross-sections calculated from FEWZ (Fully Exclusive W and Z Production). Current analysis constitutes a baseline for many CMS studies, where Z/$\gamma*$+Jets final states will be used to understand Standard Model processes as well as to look for physics beyond the Standard Model.

In addition to this analysis, online data monitoring shifts and service work related to Trigger Performance and EGamma object selection efficiencies had been undertaken during my visits to CERN.

Bibliography

[1] C. Itzykson, J-B. Zuber, "Quantum Field Theory", McGraw-Hill, 3rd edition (1987)

[2] F. Halzen, A. D. Martin, "Quarks and Leptons: An Introductory Course in Modern Particle Physics", Wiley (1982)

[3] B. R. Martin, G. Shaw, "Particle Physics", Wiley (1992)

[4] D. H. Perkins, "Introduction to High Energy Physics", Addison-Wesley (1987)

[5] S. B. Beri., "Concepts and Methods of High Energy and Nuclear Physics", Elite Publishing House (2007)

[6] S. Eidelman et al. (Particle Data Group), Phys. Lett. B 592 (2004)

[7] P. A. M. Dirac, "The Quantum Theory of the Electron", Proc. Royal. Soc., A117,610 (1928)

[8] I. J. R. Aitchison, A. J. G. Hey, "Gauge Theories in Particle Physics", IOP Publishing (1989)

[9] D. Bailin, A. Love, "Introduction to Gauge Field Theory", IOP Publishing,Revised Edition (1993)

[10] T-P. Cheng, L-F. Li, "Gauge Theory of Elementary Particle Physics", Oxford University Press (1984)

[11] Feynman, Richard "Chapter 1. QED: The Strange Theory of Light and Matter". Princeton University Press. p. 6. ISBN 978-0691125756 (1985)

[12] E. Fermi "Quantum Theory of Radiation". Reviews of Modern Physics 4: 87132. Bibcode 1932RvMP....4...87F. doi:10.1103/RevModPhys.4.87 (1932)

[13] P.A.M. Dirac "The Quantum Theory of the Emission and Absorption of Radiation". Proceedings of the Royal Society of London A 114 (767): 243265. Bibcode 1927RSPSA.114..243D. doi:10.1098/rspa.1927.0039. (1927)

[14] S. L. Glasow, "Partial Symmetries of Weak Interactions", Nucl. Phys. 22 (1961)

[15] A. Salam, J. C. Ward, "Electromagnetic and Weak Interactions", Phys. Lett. 13 (1964)

[16] A. Salam, "Proceedings of the Eigth Nobel Symposium", ed. N. Svartholm, Amkvist and Wiksell, Stockholm, 367-377 (1968)

[17] S. Weinberg, "A Model of Leptons", Phys. Rev. Lett. 19, 1264-1266, (1967)

[18] K. Nakamura et al. "Review of Particles Physics: The CKM Quark-Mixing Matrix". Journal of Physics G 37 (75021):150, (2010)

[19] P. W. Higgs, "Brocken Symmetries, Massless Particles and Gauge Fields", Phys. Lett. 12, 132-133, (1964)

[20] P. W. Higgs, "Brocken Symmetries and the Masses of Gauge Bosons", Phys.Rev. Lett. 13, 508-509, (1964)

[21] R.K. Ellis, W.J. Stirling and B.R. Webber, "QCD and Collider Physics", Cambridge University Press, (1996)

[22] D. J. Gross, F. Wilczek, "Asymptotically Free Gauge Theories", Phys. Rev.D8, 3633-3652, (1973)

[23] H. D. Politzer, "Asymptotic Freedom: An Approach to Strong Interactions",Phys. Rept. 14, 129-180, (1974)

[24] W.M. Yao et al. "Searches for Higgs Bosons -Review of Particle Physics". Journal of Physics G 33: 1. arXiv:astro-ph/0601168 (2006)

[25] T. Aaltonen (CDF and D Collaborations) "Combination of Tevatron searches for the standard model Higgs boson in the W+W decay mode". Physical Review Letters 104 (6). arXiv:1001.4162 (2010)

[26] The CDF and D0 Collaborations, "Combined CDF and D0 Search for Standard Model Higgs Boson Production with up to 10.0 fb^{-1} of Data". arXiv:1203.3774 (16 Mar 2012)

[27] ATLAS collaboration, "Combined search for the Standard Model Higgs boson using up to 4.9 fb^{-1} of pp collision data at sqrt(s) = 7 TeV with the ATLAS detector at the LHC", Physics Letters B, Volume 710, Issue 1, (29 March 2012)

[28] CMS collaboration, "Combined results of searches for the standard model Higgs boson in pp collisions at sqrts=7 TeV",Phys. Lett. B 710 p26-48,(2012)

[29] The CMS Collaboration, *"Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC"*, CMS-PAS-HIG-12-028

[30] The ATLAS Collaboration, *"Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC"*, CERN-PH-EP-2012-218.

[31] H. P. Nilles, "Supersymmetry, Supergravity And Particle Physics" Phys.Rept. 110, (1984)

[32] N. Arkani-Hamed, A. G. Cohen and H. Georgi, "Electroweak symmetry breaking

from dimensional deconstruction", Phys. Lett. B 513, 232 [arXiv:hep-ph/0105239] (2011)

[33] K. R. Dienes, E. Dudas and T. Gherghetta, "Extra spacetime dimensions and unification," Phys. Lett. B 436, 55 [arXiv:hep-ph/9803466] (1998)

[34] L. Randall and R. Sundrum, "A large mass hierarchy from a small extra dimension," Phys. Rev. Lett. 83, 3370 [arXiv:hep-ph/9905221] (1999)

[35] Y. Fukuda et al. [Super-Kamiokande Collaboration], "Evidence for oscillation of atmospheric neutrinos," Phys. Rev. Lett. 81, 1562 [arXiv:hep-ex/9807003] (1998)

[36] S. Fukuda et al. [Super-Kamiokande Collaboration], "Constraints on neutrino oscillations using 1258 days of Super-Kamiokande solar neutrino data," Phys. Rev. Lett. 86, 5656 [arXiv:hep-ex/0103033] (2001)

[37] S. D. Drell and T.-M. Yan. "Massive Lepton-Pair Production in Hadron-Hadron Collisions at High Energies", Physical Review Letters, 25:316320, (August 1970)

[38] K Nakamura and Particle Data Group. "Review of particle physics". Journal of Physics G: Nuclear and Particle Physics, 37(7A):075021, (2010)

[39] A. Abulencia et al. [CDF Collaboration], "Combination of CDF Top Pair Production Cross-Section Measurements", CDF Public Note 7794 (2005)

[40] A. Abulencia et al. [CDF Collaboration], "Precision top quark mass measurement in the lepton + jets topology in p anti-p collisions at 1.96-TeV" Phys. Rev. Lett. 96, 022004 [arXiv:hep-ex/0510049] (2006)

[41] J. Campbell and R.K. Ellis, "Next-to-Leading Order Corrections to W + 2 jets and Z + 2 jets Production at Hadron Colliders", Phys. Rev. D 65, 113007 [arXiv:hep-ph/0202176] (2002)

[42] C. F. Berger, Z. Bern et al., "Vector Boson + Jets with BlackHat and Sherpa", Nucl.Phys.Proc.Suppl.205-206:92-97 (2010)

[43] S.L. Glashow, Nucl.Phys.22,579 (1961)

[44] S. Weinberg, "A Model of Leptons",Phys.Rev.Lett. 19,1264-1266 (1967)

[45] A.Salam in Elementary Particle Physics (Nobel Symp. N.8), Ed. N.Svartholm, Almquist and Wiksells, Stockholm, p.367, (1968)

[46] http://public.web.cern.ch/public/en/About/History-en.html

[47] L. Evans and P. Bryant (editors), "LHC Machine", JINST, vol. 3, no. S08001, (2008)

[48] LEP Design Group, "Design Study of a 22 to 130 GeV electron-positron Colliding Beam Machine (LEP)", CERN/ISR-LEP/79-33, (1979)

[49] The CMS Collaboration, "The CMS Experiment at the CERN LHC", JINST, vol. 3, no. S08004, (2008)

[50] The ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider", JINST, vol. 3, no. S08003, (2008)

[51] The LHCb Collaboration, "The LHCb Detector at the LHC", JINST, vol. 3, no. S08005, (2008)

[52] The ALICE Collaboration, "The ALICE Experiment at the CERN LHC", JINST, vol. 3, no. S08002, (2008)

[53] http://press.web.cern.ch/press/PressReleases/Releases2008/PR14.08E.html

[54] R. R. Wilson, "The Tevatron", FERMILAB-TM-0763, (1978)

[55] D0 Collaboration, "Run Ilb Upgrade Technical Design Report", FERMILAB-Publication-02/327-E, 482p (Dec 2002)

[56] CMS Collaboration, Search for supersymmetry in pp collisions at s=7 TeV in events with a single lepton, jets, and missing transverse momentum, J. High Energy Phys. 08 p156, arXiv:1107.1870. (2011)

[57] CMS Collaboration, Inclusive search for squarks and gluinos in pp collisions at s= 7 TeV, Phys. Rev. D 85 p012004, arXiv:1107.1279. (2012)

[58] CMS Collaboration, Search for Supersymmetry in Events with b Jets and Missing Transverse Momentum at the LHC, J. High Energy Phys. 07 p113, arXiv:1106.3272. (2011)

[59] CMS Collaboration, Search for New Physics with Jets and Missing Transverse Momentum in pp collisions at s = 7 TeV, J. High Energy Phys. 08 p155, arXiv:1106.4503. (2011)

[60] CMS Collaboration, Search for Supersymmetry in pp Collisions at s= 7 TeV in Events with Two Photons and Missing Transverse Energy, Phys. Rev. Lett. 106 p211802, arXiv:1103.0953. (2011)

[61] CMS Collaboration, Observation of Long-Range, Near-Side Angular Correlations in Proton-Proton Collisions at the LHC, J. High Energy Phys. 09 p091, arXiv:1009.4122. (2010)

[62] CMS Collaboration, Measurements of Inclusive W and Z Cross Sections in pp Collisions at s= 7 TeV, J. High Energy Phys. 01 p080, arXiv:1012.2466. (2011)

[63] CMS Collaboration, Upsilon production cross section in pp collisions at s= 7 TeV, Phys. Rev. D 83 p112004, arXiv:1012.5545. (2011)

[64] CMS Collaboration, Measurement of the B+ Production Cross Section in pp Collisions at s= 7 TeV, Phys. Rev. Lett. 106 p112001, arXiv:1101.0131. (2011)

[65] CMS Collaboration, Measurement of the tt production cross section and the top quark mass in the dilepton channel in pp collisions at s=7 TeV, J. High Energy Phys. 07 p049, arXiv:1105.5661. (2011)

[66] The CMS Collaboration, "The CMS Tracker System Project: Technical Design Report", CERN-LHCC-98-006, http://cdsweb.cern.ch/record/368412, (1998).

[67] P. Lecoq et al., Lead Tungstate (PBWO4) Scintillators for LHC EM Calorimetry, Nucl. Instrum. Meth., vol. A365, no. 291, (1995).

[68] A. A. Annenkov et al., Lead Tungstate (PBW04) Scintillation Material, Nucl. Instrum. Meth., vol. A490, no. 30, (2002).

[69] Riccardo Paramatti for the CMS Collaboration, Calibration of the CMS Electromagnetic Calorimeter at LHC startup, CMS CR 2010/139, (2011).

[70] P. Adzic et al., Energy Resolution of the Barrel of the CMS Electromagnetic Calorimeter, JINST, vol. 2, no. P04004, (2007).

[71] The CMS Collaboration, The Electromagnetic Calorimeter Project: Technical Design Report, CERN-LHCC-97-033, http://cdsweb.cern.ch/record/349375, (1997)

[72] A. Heister et al., Measurement of Jets with the CMS detector at the LHC, CMS-NOTE-2006-036, http://cdsweb.cern.ch/record/933705, (2006)

[73] P. Cushman et al., Measurement of Missing Transverse Energy with the CMS Detector at the LHC, Eur. Phys. J., vol. C46, no. 45, CMS NOTE 2006/035, http://cdsweb.cern.ch/record/933706, (2006)

[74] P. Cushman et al., Custom HPD Readout for the CMS HCAL, Nucl. Instrum. Meth., vol. A442, no. 289, (2000)

[75] The CMS Collaboration, The Hadron Calorimeter Project: Technical Design Report, CERN-LHCC-97-031, http://cdsweb.cern.ch/record/357153, (1997)

[76] V. I. Kryshkin and A. I. Ronzhin, An Optical Fiber Readout for Scintillator Calorimeters, Nucl. Instrum. Meth., vol. A247, no. 583, (1986)

[77] M. G. Albrow et al., A Uranium Scintillator Calorimeter with Plastic-Fibre Readout, Nucl. Instrum. Meth., vol. A256, no. 23, (1987)

[78] N. Akchurin and R. Wigmans, Quartz Fibers as Active Elements in Detectors for Particle Physics, Rev. Sci. Instrum., vol. 74, no. 2955, (2002)

[79] The CMS Collaboration, The CMS Muon Project, Technical Design Report, CERN-LHCC-97-032, http://cdsweb.cern.ch/record/343814, (1997)

[80] G. Charpak and F. Sauli, High-Accuracy, Two-Dimensional Read-Out in Multiwire Proportional Chambers, Nucl. Instrum. Meth., vol. 113, no. 381, (1973)

[81] R. Santonico and R. Cardarelli, Development of Resistive Plate Counters, Nucl. Instrum. Meth., vol. 187, no. 377, (1981).

[82] R. Cardarelli et al., Performance of a Resistive Plate Chamber Operating with pure CF3 Br, Nucl. Instrum. Meth., vol. A333, no. 399, (1993)

[83] The CMS Collaboration, The TriDAS Project, Technical Design Report. Volume 1: The Level-1 Trigger, CERN-LHCC-2000-038, http://cdsweb.cern.ch/record/706847, (2000)

[84] The CMS Collaboration, The TriDAS Project, Technical Design Report. Volume 2: Data Acquisition and High-Level Trigger, CERN-LHCC-2002-026, http://cdsweb.cern.ch/record/578006, (2000).

[85] https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookCMSSWFramework

[86] CMS Collaboration, "Particle Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET", CMS PAS PFT-09-001 (2009).

[87] CMS Collaboration, "Particle Floow reconstruction of 0.9 TeV and 2.36 TeV collision events in CMS", CMS PAS PFT-10-001 (2010).

[88] CMS Collaboration, "Missing Transverse Energy Performance in Minimum-Bias and Jet Events from Proton-Proton Collisions at $\sqrt{s} = 7$ TeV", CMS Physics Analysis Summary, CMS-PAS-JME-10-004 (2010).

[89] CMS Collaboration, "CMS MET Performance in Events Containing Electroweak Bosons from pp Collisions at $\sqrt{s} = 7$ TeV", CMS Physics Analysis Summary CMS-PAS-JME-10-005 (2010).

[90] CMS Collaboration, "MET Performance in pp Collisions at sqrts=7 TeV", CMS-PAS-JME- 10-09, CERN-PH-EP-2011-051 (2010)

[91] CMS Collaboration, "Tracking and Vertexing Results from First Collisions" CMS PAS TRK-10-001 (2010).

[92] E. James et al., "Muon Identification in CMS," CMS Note 010 (2006).

[93] CMS Collaboration, "Reconstruction of Electron TracksWith the Gaussian-Sum Filter in the CMS tracker at LHC," CMS Note RN 2003-001 (2003).

[94] S. Baffioni, et al., "Electron Identification in CMS", CMS AN-2009/178.

[95] "Updated Measurements of the Inclusive W and Z Cross Sections at 7 TeV ", CMS AN-2010/264

[96] "Measurement of the W and Z inclusive production cross sections at sqrts=7 TeV with the CMS experiment at the LHC", CMS Collaboration, CMS-PAS-EWK-10-005, CDS Record:1337017

[97] W. Andrews et al., "Study of Photon Conversion Rejection", CMS AN-2009/159.

[98] M. Cacciari and G.P. Salam and G. Soyez, "The anti-kt jet clustering algorithm," JHEP 2008, no. 04, 063.(2008)

[99] S. Catani, Y.L. Dokshitzer, M.H. Seymour and B.R. Webber, "Longitudinally invariant K(t) clustering algorithms for hadron hadron collisions", Nucl. Phys. B 406 p187 (1993)

[100] S.D. Ellis and D.E. Soper, "Successive combination jet algorithm for hadron collisions", Phys. Rev. D 48 p3160 [hep-ph/9305266] (1993)

[101] Y.L. Dokshitzer, G.D. Leder, S. Moretti and B.R. Webber, "Better jet clustering algorithms", JHEP 08 p001 [hep-ph/9707323], (1997)

[102] Matteo Cacciari. FastJet: A Code for fast k(t) clustering, and more. (2006).

[103] CMS Collaboration, "Plans for Jet Energy Corrections at CMS", CMS-PASJME-07-002, (2007).

[104] Niki Saoulidou, "Particle Flow Jet Identification Criteria", CMS AN-2010/003, (2010)

[105] Baer, H. et al., "Les Houches Guidebook to Monte Carlo Generators for Hadron Collider Physics". hep-ph/0403045, (2004).

[106] Sjoestrand, T. et al.: PYTHIA 6.4 Physics and Manual. hep-ph/0308153, (2006).

[107] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand. Parton fragmentation and string dynamics. Physics Reports, 97(2-3):31145, 1983.

[108] Rick Field. "Early LHC Underlying Event Data - Findings and Surprises". (2010)

[109] Mangano, M. "Higher-order QCD processes via LO matrix-element MC's." (2004)

[110] Mangano, M. "Merging multijet matrix elements and shower evolution in hadronic collisions" (2001)

[111] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, Tim Stelzer. "MadGraph 5: Going Beyond", JHEP 1106 p128, arXiv:1106.0522 [hep-ph].(2011)

[112] Mrenna, S.: Mrenna Matched Datasets. http://cepa.fnal.gov/personal/mrenna/Matched Dataset Description.html.

[113] Mrenna, S. and Richardson, P.: Matching matrix elements and parton showers with HERWIG and PYTHIA. JHEP05(2004)040, (2004).

[114] Catani, S. et al.: QCD matix elements + parton showers. JHEP11(2001)063, (2001).

[115] J. Alwall et al., "A standard format for Les Houches event files", Comput. Phys. Commun. 303 176 p300304, arXiv:hep-ph/0609017. (2007)

[116] Ryan Gavin, Ye Li, Frank Petriello, Seth Quackenbush, "FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order", Comput.Phys.Commun.182:2388-2403, (2011)

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

[118] S. Agostinelli et al. "G4a simulation toolkit". Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 506(3):250303, (2003)

[119] R. Brun et al., "ROOT - An Object Oriented Data Analysis Framework," Nucl. Instrum. Meth. A 389 (1997).

[120] CMS collaboration, "Absolute Calibration of the Luminosity Measurement at CMS: Winter 2012 Update", CMS-PAS-SMP-12-008, (2012)

[121] https://twiki.cern.ch/twiki/bin/viewauth/CMS/StandardModelCrossSections

[122] Ryan Gavin, Ye Li, Frank Petriello, Seth Quackenbush, "FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order", arXiv:1011.3540v1 [hep-ph], (2010).

[123] J. Pumplin et al., "New generation of parton distributions with uncertainties from global QCD analysis", JHEP 07 p012 (2002)

[124] CMS Collaboration, "Tracking and Primary Vertex Results in First 7 TeV Collisions," CMS PAS TRK-10-005 (2010).

[125] W. Erdmann, "Offline Primary Vertex Reconstruction with Deterministic Annealing Clustering", CMS IN -2011/014, (2011)

[126] N. Adam et al., "Generic Tag and Probe Tool for Measuring Efficiency at CMS with Early Data", CMS AN 2009/111, (2009).

[127] F. Beaudette et. al., "Electron Reconstruction within the Particle Flow Algorithm", CMS-2010/034 (2010).

[128] F.A. Berends, W. Giele, H. Kuijf, R. Kleiss and W. Stirling, "Multi-jet production in W, Z events at $p\bar{p}$ colliders", Phys. Lett. B 224 p237 (1989)

[129] CMS collaboration, "Jet production rates in association with W and Z bosons in pp collisions at sqrts = 7 TeV", JHEP-102P-1011 (2011)

[130] RooUnfold: ROOT Unfolding Package and documentation. http://hepunx.rl.ac.uk/ adye/software/unfold/RooUnfold.html.

[131] G. D'Agostini. "A Multidimensional unfolding method based on Bayes theorem." Nucl. Instrum. Meth., A362:487498, (1995)

[132] Andreas Hoecker and Vakhtang Kartvelishvili. "SVD approach to data unfolding". Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 372(3):469481, (1996)

[133] https://twiki.cern.ch/twiki/bin/view/CMS/JECUncertaintySources

[134] CMS collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", arXiv:1107.4277v1 [physics.ins-det] (21 Jul 2011)

[135] https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetResolution

[136] https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupSystematicErrors

[137] Measurement of the W and Z inclusive production cross sections at $\sqrt{(}s)$=7 TeV with the CMS experiment at the LHC, DOI: 10.1007/JHEP10(2011)132, (2011)

Publications/Conferences

Selected Publications / CMS Public Notes

- EWK-11-017 – Study of the di-jet mass spectrum in W+2-jets events
- JME-10-009 – MET Performance in pp Collisions at $\sqrt(s) = 7$ TeV
- EWK-10-012 – Rates of Jets Produced in Association with W and Z Bosons
- EWK-10-005 – Measurement of the W and Z inclusive production cross sections at $\sqrt(s) = 7$ TeV with the CMS experiment at the LHC
- EWK-10-002 – Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt(s) = 7$ TeV

Published CMS papers: 192

Selected CMS Analysis Notes

- CMS AN-2012/135 – Measurement of Z/γ*+jet+X Angular Distributions in pp Collisions at $\sqrt(s) = 7$ TeV in the Electron Channel
- CMS AN-2012/033 – Measurement of $d\sigma(Z+n)/dp_T^{n^{th}Jet}$ and $d\sigma(Z+n)/dY^{n^{th}Jet}$ in Proton-Proton Collsions at $\sqrt{s} = 7\ TeV$
- CMS AN-2011/266 – Study of the dijet invariant mass distribution in W($\rightarrow l\nu$) + jets events in pp collisions at $\sqrt{s} = 7$ TeV
- CMS AN-2011/017 – Efficiency correction and unfolding for V+jets events
- CMS AN-2010/413 – Measurement of the Associated Production of Vector Bosons and Jets in proton-proton collisions at $\sqrt(s) = 7$ TeV

- CMS AN-2010/408 – Study of Z/γ^* +Jet(s) production at sqrt(s)=7TeV and Commissioning of Particle Flow MET

- CMS AN-2010/377 – SUSY Searches in the Z+jets+MET final state in 7 TeV pp collisions with a jet-Z balance method

- CMS AN-2010/395 – Measurement of the Inclusive W and Z Cross Section in pp Collisions at 7 TeV: Update with full 2010 statistics

- CMS AN-2010/264 – Updated Measurements of the Inclusive W and Z Cross Sections at 7 TeV

- CMS AN-2010/250 – Searching SUSY in Z+Jet(s)+MET Final State Using Particle Flow

- CMS AN-2010/246 – Trigger efficiency correction factors for inclusive W/Z cross section measurement

- CMS AN-2010/136 – Study of Production of Vector Bosons and Jets at 7 TeV for ICHEP2010 Conference

- CMS AN-2010/116 – Measurements of Inclusive W and Z cross-sections in pp collisions at $\sqrt(s) = 7$ TeV

- CMS AN-2010/036 – Z/γ^*+Jets in Muon Mode using PF Reconstruction

- CMS DP-2010/029 – Trigger Study Group: HLT performance

- CMS AN-2009/170 – A Study of W/Z+jets Events Using Electron Modes

- CMS AN-2008/093 – A Study of W/Z+jets Events Using Electron Modes

Conferences/Presentations

- The MCnet-LPCC school on Event Generators for LHC, 23-27 July 2012, CERN
- Lepton-Photon 2011, Mumbai, 22-27 August 2011
- DAE-BRNS symposium of HEP, JaiPur, India, Dec 13-18, 2010.
- 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research: Jaipur, India, February 22-27, 2010.
- Workshop on Physics at the LHC, Tata Institute of Fundamental Research, Mumbai, 21-27 October 2009.
- DAE-BRNS symposium of HEP, Banaras, India, Dec 14-18, 2008.
- Workshop on Physics at the LHC, Tata Institute of Fundamental Research, Mumbai, 2008.