



Measurement of jet production in association with a Z boson at the LHC
and jet energy correction calibration at high level trigger in CMS

PEKING UNIVERSITY
UNIVERSITÉ LIBRE DE BRUXELLES

Fengwangdong Zhang

Thesis submitted for the degree of

Doctor of Philosophy

May, 2017

Doctoral Jury:

- Prof. Yong BAN (Co-supervisor) (Peking University)
- Prof. Laurent FAVART (Co-supervisor) (Université Libre de Bruxelles)
- Dr. Philippe GRAS (CEA Saclay)
- Prof. Yajun MAO (President) (Peking University)
- Prof. Pascal VANLAER (Secretary) (Université Libre de Bruxelles)
- Prof. Shouhua ZHU (Peking University)

Abstract

This PhD dissertation presents the measurement of the cross section of jet production in association with a Z boson in proton-proton collisions at the Large Hadron Collider in CERN, with a center-of-mass energy of 8 TeV in 2012 and of 13 TeV in 2015. The data used for this analysis were collected by the Compact Muon Solenoid (CMS) detector, with an integrated luminosity of 19.6 fb^{-1} in 2012 and of 2.25 fb^{-1} in 2015. The differential cross section is measured as a function of jet multiplicity, jet transverse momentum and rapidity, and the scalar sum of jet transverse momenta. The rapidity correlations between the Z boson and jets are also measured benefiting from the large statistics of data taken in 2012. All distributions of measured observables are obtained after correcting detector effects using unfolding approach, and the results of two leptonic decaying channels of Z boson are combined. Coming along with the systematic and statistical uncertainties, the measurement is compared to different theoretical predictions at different accuracy levels. The predictions are from MADGRAPH 5, SHERPA 2 (for 8 TeV analysis only), MADGRAPH_AMC@NLO, and fixed next-to-next-to-leading order (for 13 TeV analysis only). Thanks to the unprecedented high energy and the large statistics of data, precision measurement is accomplished in a physical phase space never reached before. This measurement provides precise systematics for different theoretical models. It also quantifies the improvement with higher order of perturbative quantum chromodynamics calculations on matrix elements relative to the leading order multi-leg approach. In particular to the rapidity correlation study, new matching schemes (FxFx and MEPS@NLO) for next-to-leading order matrix elements and parton shower show significant improvements with respect to the MLM matching scheme for leading order multi-leg matrix elements and parton shower. This measurement also gives precise background estimation for the measurements of many other processes in Standard Model like top quark production and gauge boson couplings, and for new physics searches such as Supersymmetry.

In this thesis, the jet energy correction and calibration for the high level trigger system of CMS are also depicted. From 2012 to 2015, the Large Hadron Collider was upgraded, not only with the center-of-mass energy of the beams enlarged, but also with the instantaneous luminosity increased. The time distance between two particle bunches in a beam is reduced. As a result, the reconstructed momenta of the jets produced in each bunch crossing are significantly contaminated by multiple interactions. A dedicated technical approach has been developed for correcting the reconstructed jet momenta. The corrections have been calibrated and configured for the data taking in 2015 and 2016.

Key words: Quantum Chromodynamics, Standard Model Measurement, Z Boson, Jet, Large Hadron Collider, Compact Muon Solenoid, High Level Trigger

Acknowledgments

This joint PhD is supported by Peking University and Université Libre de Bruxelles, and the research projects presented in this thesis were also funded by China Scholarship Council.

First of all, I would like to deliver my greatest gratitude to my two co-supervisors, Prof. Yong Ban and Prof. Laurent Favart. When I started my PhD in Peking University, Prof. Ban provided me several lectures about high energy experiments. During the daily study time, he not only answered plenty of my academic questions, but also continuously guided me to the correct direction of research. Once I encountered complicated administrative issues, he helped me to solve them promptly. He also cared about my daily life in the university, asking me whether I need assistance very often. Thanks to him, I got my first chance to visit CERN in Geneva during the summer of 2013, which was also the first time for me to go abroad, when I was just in the first year of my PhD. It opened a new horizon to me, and I was deeply impressed by the major accelerator and detector. During the three months, I gained precious experiences on the detector operation. I felt a new academic atmosphere and culture in Europe, and a beautiful environment. Thanks to his contact and the assistance from Laurent, I obtained the opportunity to begin my co-PhD in 2014. Laurent welcome me in a great pleasure when I came to Belgium at the first time. At the beginning, Laurent held private lectures about QCD theory for me weekly, and he tried his best to inspire my strong interests and enthusiasm. He really impressed me a lot with his strongly patient attitude and rigorous cogitation to answer my substantial questions. After I started my main data analysis, he discussed with me about the progresses and problems very often, and I also learned more interesting theories during our meetings. Thanks to him, I had significant progress in English, as well as French. He also supported me to attend several international conferences and schools in other European countries. Through the frequent missions and conference talks, I improved my ability on research, social communication, and independence. He also cared about my daily life in Belgium, and removed my several confusions about the different culture and social phenomena in Europe. He invited me several times to his home to have meals together with his warm family, so I was very touched and excited. Thanks to him, Brussels has become my second home.

I am also grateful to my Chinese co-supervisor Prof. Qiang Li, who worked with me together on my first data analysis for several years. He offered me a smooth beginning of physics data analysis work, and guided me the proper way of doing research during the working time. The second project presented in this thesis could not be launched if without his assistance. Meanwhile, I also thank the team leaders of my two cooperated institutes, Prof. Yajun Mao and Prof. Pascal Vanlaer, for their administrative supports.

My main data analysis work could not stand well if without the help from Dr. Alexandre Léonard, who spent plenty of time showing me many details of analysis programs hand by hand, and also enlightened me with the physical ideas behind them. We also had frequent discussions and figured out many issues together. Besides that, I also accomplished an online DAQ software development project in our laboratory with the help of him. Including him and my supervisors, a powerful research

team together with me were working on this data analysis work. I also would like to thank my another three colleagues Dr. Philippe Gras, Dr. Tomislav Seva, and Bugra Bilin, who helped me a lot with the analysis and had lively debugging and discussion meetings with me quite often. There were quite much fun with them together.

I am also very thankful to my colleague Dr. Anastasia Grebenyuk, who worked with me together on the same topic of physics analysis using the 13 TeV data. We enjoyed a striving and pleased team work, and shared colorful stories during the daily life.

I must thank two colleagues at CERN, Dr. Konstantinos Kousouris and Dr. Henning Kirschenmann. They did a lot of favors for me on the second project presented in this thesis. Through frequent email exchanges or discussions in site, I learned plenty of knowledge from them, making this project progressing efficiently.

Besides the people above, I want to thank my other colleagues including Dr. Hugues Brun, Dr. Cécile Caillol, Zhaoru Zhang, and Daneng Yang, for their kindly helps with my study or research before. However, there are still so many other colleagues and friends to be acknowledged, but I am afraid that this section would become extremely long. I thank all of them here.

Finally, I express my greatest love to my dear parents for their successive support on my study and research career. No matter where I am, they always care about my daily life and wish to know whether I am in good status. Once I got stuck by any issue, they would try their best to help me or guide me to solve it. Therefore this PhD also contains their abundant hardships.

Contents

1	Introduction	9
2	Theoretical Introduction	12
2.1	Standard Model	12
2.2	The formulation of SM	13
2.2.1	The Electroweak Interaction	14
2.2.2	The Quantum Chromodynamics	16
2.2.3	The Running of Coupling Constants	17
2.2.4	Perturbative QCD	18
2.2.5	Parton Distribution Functions	19
2.3	Application to Z + jets in proton-proton collision	20
2.3.1	Parton Shower and Hadronization	24
2.3.2	Multiple Parton Interaction	25
2.3.3	Jet Definition	26
2.4	Monte Carlo Simulations	28
2.4.1	PYTHIA	30
2.4.2	MADGRAPH 5	31
2.4.3	MADGRAPH_AMC@NLO	32
2.4.4	SHERPA	33
2.4.5	Matching between matrix elements and parton shower	33
3	The Experimental Setup	36
3.1	The Large Hadron Collider	36
3.1.1	Introduction	36
3.1.2	Beam luminosity	38
3.1.3	Injection chain	41
3.2	The CMS experiment	43
3.2.1	Introduction	43
3.2.2	The Superconducting solenoid magnet	46
3.2.3	The Tracker	46

3.2.4	The Calorimeters	48
3.2.5	The Muon Detector	52
3.2.6	The Trigger system	57
3.2.7	Data handling application	61
4	Jet production in association with Z boson	62
4.1	Motivation	62
4.2	Data and Simulation samples	64
4.3	Event reconstruction and selection	66
4.3.1	Trigger	67
4.3.2	Track and vertex reconstruction	70
4.3.3	Particle flow reconstruction	72
4.3.4	Muons	74
4.3.5	Electrons	77
4.3.6	Z Bosons	79
4.3.7	Jets	80
4.3.8	Pileup	86
4.4	Results at detector reconstructed level	91
4.4.1	Kinematic variable distributions at 8 TeV	91
4.4.2	Kinematic variable distributions at 13 TeV	113
4.4.3	Background subtraction	119
4.5	Measured observables	119
4.6	Data corrections for detector effects	120
4.6.1	Phase space selection	120
4.6.2	Introduction to Unfolding approach	121
4.6.3	Response matrix	123
4.6.4	Comparison between p_T order matching and angular matching	134
4.7	Uncertainties	138
4.8	Non-perturbative correction	148
4.9	Differential cross sections in a single channel	151
4.9.1	Comparison of single channels	151
4.9.2	Results of single channels at 8 TeV	151
4.9.3	Results of single channel at 13 TeV	160
4.10	Combined differential cross sections	164
4.10.1	Combination methodology	164
4.10.2	Combined results at 8 TeV	165
4.10.3	Combined results at 13 TeV	175
4.11	Recent Results on Z+Jets Measurements	177

5	Jet energy correction and calibration at high level trigger	188
5.1	Motivation	188
5.2	Track and vertex reconstruction at HLT	189
5.3	Jet reconstruction at HLT	190
5.4	Pileup at HLT	190
5.5	Jet energy correction	192
5.5.1	Pileup energy subtraction (L1 correction)	193
5.5.2	Relative (L2) and absolute (L3) corrections	195
5.5.3	Residual L2L3 corrections	197
5.5.4	Calibration workflow in simulations and results	198
6	Conclusion	210
A	Tables of combined results at 8 TeV	221
B	Results of single channel ($Z \rightarrow e^+e^-$) at 13 TeV	233

Chapter 1

Introduction

From ancient time to the current period, human beings have been keeping researches for our nature. The vast universe has supplied an infinite space for our investigation. Through the development of physical science, we have explored and discovered more and more about the nature. Now we know that the universe contains innumerable matters and energy. Only a small fraction of them is visible to us, which are made of elementary particles that we have discovered so far. The rest part includes dark matter and dark energy that dominate the ingredients of the universe, but we are not able to observe them directly. Even for the particles that we know, many aspects of their interaction mechanism are still unclear to us. For instance, among the four fundamental interactions known in the nature, the strong, electromagnetic, weak, and gravital forces, only the three first can be described in terms of elementary particles exchanges. It is still a question whether any new coupling mechanism along them exists. Therefore, we probably still know very little about the universe.

Our research approaches are mainly divided in three branches, including theoretical hypotheses, derived predictions based on the former, and experimental observations. These three elements are always accompanying each other to get themselves further developed. To be specific, the theoretical assumptions should be validated by experimental observations, which requires appropriate computations in consequence to provide predictions on the observables that can be probed or measured by experiments. At this point, technical equipments have to be built to realize the experimental measurement or search. The accuracy of theoretical predictions is also crucial for experimentalists to gain an intuitive view of the natural processes happening inside the experimental apparatus, and to understand better the fundamental mechanism. Moreover, if there is any new phenomenon happening in experiments, it gives an inspiration of new hypotheses and theoretical interpretations.

Since the second part of last century, the particle physics has been established and developed for many decades. It relies on a comprehensive theory called Standard Model (SM), an elaboration of the elementary particles (quarks, leptons, bosons) that have been discovered together with their fundamental interactions. This model has encoded results of particle physics experiments over seven decades and is still supported by substantial modern high energy experiments. Among them, thanks to the endeavor of many physicists and engineers, a major particle accelerator called Large Hadron

Collider (LHC) has been built at CERN with the unprecedented high energy that human beings have made. This machine accelerates protons or heavy ions to extremely high energy and makes them collide, which creates huge amount of particles for particle physics researches. There are two major experiments at the LHC, which are A Toroidal LHC Apparatus (ATLAS) and Compact Muon Solenoid (CMS). They discovered the existence of Higgs boson in 2012. Then LHC has reached one of its main goals.

Quantum Chromodynamics (QCD) is an important part of SM, providing a detailed explanation about strong interaction mechanism. Accurate theoretical computations and precise measurements for it are crucial for further development of particle physics. The LHC being a hadron collider, strong interactions take place in almost all processes studied in ATLAS and CMS. In particular the production of jets is abundant and would be a very limitative effect on the LHC discovery potential without a precise knowledge of jet production mechanism. In this dissertation, a precise measurement of jet production in association with a Z boson in proton-proton collisions at the LHC is performed and therefore provides stringent tests on QCD calculations. The process is detected by the CMS detector. Because the characteristics of Z bosons are well known, and Z bosons are easily detected, this process has a clean signature among many other processes. Thanks to the high energy of beams at the LHC and the substantial statistics of recorded data, this process leads to a large amount of data, which makes possible to reach extended phase spaces for precision measurements. For theoretical predictions, a higher precision means more complicated calculations. Based on the calculation, the predictions are achieved through Monte Carlo programs.

Furthermore, at the LHC, experimentalists wish to measure many other SM process precisely, such as top quark production and Higgs boson production. Meanwhile, physicists are also aimed to search for new physics beyond SM, dark matter for instance. Many of these measurements or searches suffer from a significant background of Z + jets process, which has to be distinguished and estimated accurately. As a result, the precise measurement of Z + jets process is also very important for the accuracy of their background subtraction from the data.

Although the existence of Higgs boson has been confirmed by the LHC, this major accelerator still has its unfinished tasks, including Higgs boson precision measurement, new physics searches, and more precision measurements of other elementary particles. Thus its maximal reached energy and the data statistics are still needed to be increased. In order to handle such huge amount of data, trigger system is mandatory for the detector, because the components of the detector are not able to record and process the produced data from the collisions in such a high frequency (about 40 MHz). Trigger system filters the data and selects the data at a proper frequency, so they must be working efficiently. In the trigger system, jets are reconstructed from the data. Due to the high frequency of beam collisions, there are substantial energy from extra interactions mixed with the energy of the jets produced in a specific physical process of interests. Moreover, the limited energy resolution of the detector gives rise to a smearing on the measured jet energy spectrum, which directly influences the efficiency of jet trigger system. At this point, the second topic of this thesis presents an approach

for correcting the jet energy at trigger level.

This dissertation is divided into six chapters in total. The first chapter is the present introduction. The second chapter provides a brief interpretation about the relevant theory elements and implementation of the Monte Carlo programs, in particular to their applications to the jet production in association with a Z boson process in proton-proton collisions. The third chapter gives a review of the LHC and the CMS experiment. The forth chapter explains the measurement of the jet production in association with a Z boson in proton-proton collisions with a center-of-mass energy of 8/13 TeV in detail. The fifth chapter presents the jet energy correction and calibration approach at high level trigger in the CMS detector. The last chapter summarizes and concludes. A list of references is attached to the main body of this thesis. Finally, an appendix lists the measured differential cross sections versus different observables of Z + jets process at 8 TeV with the systematic and statistical uncertainties in tables, and the next one provides the preliminary results of Z + jets process measurement in $Z \rightarrow e^+e^-$ channel at 13 TeV compared to different theoretical predictions.

Chapter 2

Theoretical Introduction

Particle physics has been developed for searching the fundamental constituents which constructed nature. These constituents are called elementary particles, and they are substantial in the universe. As known in common, everything in the world is made of atoms. They interact with each other and are bundled to form matters. An atom consists of a nucleus with one or several electrons moving fast around it. A nucleus is composed of protons and neutrons, except for a hydrogen nucleus made of only one proton. If one wonders about the substructure of a proton or a neutron, according to Heisenberg Uncertainty principle, one has to use much higher energy to probe inside it. So far, we know that these substructures are quarks, which are elementary particles. Since an electron does not have substructure that is proved by experiments, it is also an elementary particle. The comprehensive theory called Standard Model (SM) explains these concepts and reproduces many experimental observations with high accuracy. This chapter is inspired generally by [1, 2, 3], which provides a general introduction about this knowledge.

2.1 Standard Model

The Standard Model is a comprehensive theory of particle physics, describing all of the elementary particles known so far in our universe, and also explaining three basic types of interactions (strong, electromagnetic, and weak). It was developed since the latter part of the 20th century. After the indirect observation of quarks in the middle of 1970s, the SM formulation was close to being steady. Later, the top quark, the tau neutrino, as well as the Higgs boson were discovered, which consolidated the further confidence of the SM. However, it does not cover the forth fundamental interaction - gravity, which is interpreted by the general relativity.

The elementary particles are classified as fermions and bosons. The formers include six quarks and six leptons (together with their corresponding antiparticles), which have half-integer spin and therefore obey the Pauli exclusion principle. Quarks are grouped in three generations, carrying color charge such that they interact through strong force, and are always bounded to one another (color confinement) to form hadrons. Leptons have three generations as well. Among them, electrons,

muons, and taus carry electric charge, which make them interact through electromagnetic and weak forces. Neutrinos are electrically neutral in three flavors, interacting through only weak force. Quarks also have electric charge, and they participate in electromagnetic and weak interactions as well.

Bosons have integer spin and yield the Bose-Einstein statistics, including gauge bosons (spin 1) and scalar boson (spin 0). Gauge bosons perform as the force **carriers** between the fermions. To be specific, gluons are exchanged between quarks carrying color charge, mediating the strong interaction. Photons are the electromagnetic force carrier, and $W^\pm/Z$ bosons are for the weak interactions. Scalar boson (Higgs boson) is the quantum of Higgs field that couples with massive fundamental particles, leading them obtaining their masses. It also interacts with itself making it massive. It has zero spin, no charge, and is unstable. Thanks to the large hadron collider at CERN, it was discovered experimentally in 2012, half a century after being predicted.

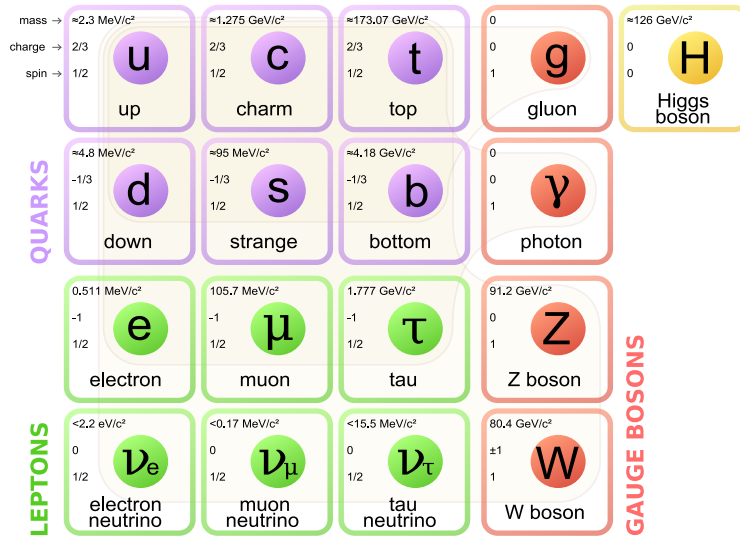


Figure 2.1: The schematic view of the elementary particles of the SM, including three generations of quarks and leptons, together with four gauge bosons as their force carriers, and the Higgs boson.

2.2 The formulation of SM

The SM is based on the quantum field theory, where particles are described as dynamical field in spacetime. The dynamics and kinematics of a field system are controlled by a gauge invariant Lagrangian, which is determined according to the particle content in this system and observes the potential symmetries. The SM follows a $SU(3) \otimes SU(2) \otimes U(1)$ gauge symmetry, where $SU(3)$ corresponds to the group of colour transformations for the interactions between gluons and quarks, which is interpreted by quantum chromodynamics (QCD) theory. The $SU(2) \otimes U(1)$ stands for the group of transformations for electromagnetic and weak (electroweak/EWK) interactions, and the electromagnetic force is described by quantum electrodynamics (QED) theory. Moreover, the

$SU(2) \otimes U(1)$ also reflects the symmetric characters of Brout-Englert-Higgs mechanism developed for the Higgs boson field.

In summary, all the fundamental particles of the SM are understood as quantum fields in space-time, which is called Lorentz space. They are labeled mathematically as:

- Fermi fields (matter particles): ψ
- Electroweak boson fields: W_1, W_2, W_3, B
- Gluon fields: G_a
- Higgs field: ϕ

These fields yield specific transformation regularity in the Lorentz space. The Fermi field behaves as a spinor under transformations, while the electroweak and gluon fields transform as a four-vector. The Higgs field behaves as a scalar.

2.2.1 The Electroweak Interaction

The electroweak theory is a unification of electromagnetism and weak mechanism, following the $SU(2)_L \otimes U(1)_Y$ gauge group. The subscript L means EWK coupling only to the fermions in left-handed state, and the subscript Y stands for the hypercharge of weak coupling. From $SU(2)_L$, there are three massless gauge boson states with three different weak isospin, corresponding to three components of $SU(2)$ gauge field, respectively, labeled as W_1, W_2, W_3 . From $U(1)_Y$, there is a single gauge boson state with weak hypercharge, represented by B . In order to unify the electromagnetic and weak interactions, the generator of $U(1)$ in QED is given by:

$$Q = \frac{1}{2}Y + T_3 \quad (2.1)$$

where the Q stands for the electric charge, related to the T_3 that is the third component of weak isospin, and the weak hypercharge Y . Due to the spontaneous symmetry breaking of the electroweak transformation group, the $W^\pm, Z$ bosons and photons are produced. They are connected by weak mixing angle θ_W :

$$\begin{pmatrix} \gamma \\ Z \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B \\ W_3 \end{pmatrix} \quad (2.2)$$

The $W^\pm$ is given by:

$$W^\pm = \frac{1}{\sqrt{2}}(W_1 \mp iW_2) \quad (2.3)$$

The masses of $W^\pm$ and Z bosons are induced by the spontaneous symmetry breaking according to the Brout-Englert-Higgs Mechanism, and they are related at tree level by the mixing angle:

$$M_W = M_Z \times \cos \theta_W \quad (2.4)$$

The Lagrangian of the electroweak interaction is gauge invariant, which is written as:

$$\mathcal{L}_{EWK} = i\bar{\psi}\gamma^\mu D_\mu\psi - \frac{1}{4}W_{\mu\nu}^i W^{i\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \quad (2.5)$$

for $i = 1, 2, 3$. The fermion field ψ is transformed as:

$$\psi \rightarrow \psi' = \exp\left(\frac{i}{2}g_W T^k \Lambda^k\right)\psi \exp\left(\frac{i}{2}g'_W Y \alpha\right) \quad (2.6)$$

where the g_W stands for the $SU(2)_L$ gauge coupling and g'_W represents the $U(1)_Y$ gauge coupling. The T^k is a tensor forming a representation of the $SU(2)_L$ weak isospin algebra and Y is a scalar holding a representation of the $U(1)_Y$ weak hypercharge algebra. The first exponential term means an $SU(2)_L$ rotation around the axis Λ^k in the space of weak isospin, while the second item gives rise to a phase transformation α in the space of weak hypercharge. The weak coupling constants are related to weak mixing angles and electric charge as well:

$$\sin^2 \theta_W = \frac{g_W'^2}{g_W^2 + g_W'^2} \quad (2.7)$$

$$e = g_W \sin \theta_W \quad (2.8)$$

The D_μ in the Lagrangian 2.5 represents a series of covariant derivatives, which make the Lagrangian gauge invariant:

$$(D_\mu)_{ij} = \delta_{ij}\partial_\mu + ig_W(T^k W_\mu^k)_{ij} + ig'_W Y \delta_{ij} B_\mu \quad (2.9)$$

These derivatives together with γ matrices in the first item of Lagrangian introduce the interaction between fermion field and gauge field. The $W_{\mu\nu}^i$ and $B_{\mu\nu}$ in the Lagrangian stand for the field strength tensors of $SU(2)_L$ gauge field W^i and $U(1)_Y$ gauge field B , respectively. They are expressed as:

$$W_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i - g_W \epsilon^{ijk} W_\mu^j W_\nu^k \quad (2.10)$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \quad (2.11)$$

The ϵ^{ijk} is an antisymmetric tensor with its components given by the structure constants of $SU(2)$, which satisfy the commutative rule as:

$$[T^i, T^j] = i\epsilon^{ijk} T^k; \epsilon^{123} = 1 \quad (2.12)$$

According to the definition of $W_{\mu\nu}^i$, the W^i gauge boson can have self interaction, which illustrates the non-Abelian character of $SU(2)$.

2.2.2 The Quantum Chromodynamics

This theory provides a comprehensive description of the interactions between quarks and gluons, which composes hadrons. Similar to the electric charge in QED, the colour is the **charge** responsible in QCD. There are three different types of colours, red, blue, and green. A quark can carry any colour, while an antiquark takes along anticolour (antired, antiblue, and antigreen). As the strong force carrier, a gluon typically carries combined colours to be exchanged between different quarks (antiquarks). Although the quarks and gluons have such colour creation and annihilation frequently in hadrons, the whole hadrons are colourless. A quark is not able to stand alone (free), but it is always bounded with another quarks or antiquarks to form hadrons, through forming gluon field. When a quark is going to be separated from other quarks, the energy in the gluon field is enough to create another quark pair. Then one of the quark pair will be combined with this quark to form a new bound state, and the other one will be bounded with the other quarks. This is called confinement that is one of the typical characters of QCD. The other feature is the asymptotic freedom, which indicates that quarks and gluons are seen as free particles when the reaction energy is very high.

QCD is also based on quantum field theory while coming with a gauge transformation yielding symmetry group $SU(3)$. The colour charge is taken to define a local symmetry. Because the strong interaction does not discriminate different flavors of quarks, there is an approximate flavor symmetry in QCD, which is broken by the differences of quark masses. The QCD Lagrangian is given by:

$$\mathcal{L}_{QCD} = \sum_{flavors} \bar{\psi}_{q'} (i\gamma^\mu D_\mu - m\delta_{qq'}) \psi_q - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu} \quad (2.13)$$

where ψ_q stands for the quark field with a specific flavor (q), and the Lagrangian sums on all of the quark flavors. Because QCD yields an $SU(3)$ gauge group, the transformations of quark field are given by:

$$\psi_q \rightarrow \psi'_q = \exp\left(\frac{i}{2} g_s T^a \mathcal{A}^a\right) \psi_q \quad (2.14)$$

The g_s is the $SU(3)$ gauge coupling strength, and T^a corresponds to a series of tensors forming the representation of the $SU(3)$ colour algebra (following the equation 2.12), with the index a summing over the eight colour degrees of freedom of gluon field. The $\mathcal{A}$ stands for the gluon field in the colour space.

The covariant derivatives D_μ are provided by:

$$(D_\mu)_{ij} = \partial_\mu \delta_{ij} + ig_s (T^a \mathcal{A}_\mu^a)_{ij} \quad (2.15)$$

where the indices i and j run over the triplet representation of the colour group. The last item of the Lagrangian represents the gluon field dynamics. The $G_{\mu\nu}^a$ is a tensor from gluon field strength, which is gauge invariant:

$$G_{\mu\nu}^a = \partial_\mu \mathcal{A}_\nu^a - \partial_\nu \mathcal{A}_\mu^a - g_s \epsilon^{abc} \mathcal{A}_\mu^b \mathcal{A}_\nu^c \quad (2.16)$$

where the ϵ are the structure constants of the $SU(3)$ colour group, which introduces the gluon self interaction showing the non-Abelian feature, contrasting to the electromagnetic interaction.

2.2.3 The Running of Coupling Constants

According to QED theory, the effective coupling constant of electromagnetic interaction is:

$$\alpha_{em} = \frac{e^2}{4\pi} \simeq \frac{1}{137} \quad (2.17)$$

This value was obtained from experiments. According to the QED theory, there are always spontaneous electric pair creation and annihilation happening in a vacuum system, which are called quantum fluctuation. If a free electron is inside such a local system, it induces those pairs to form electric dipoles surrounding itself, like producing a shielding cloud, and then the whole electromagnetic field pervades in the whole system. This effect is called vacuum polarization. Far away from the electron, one can observe the effective coupling constant as the value mentioned above. If getting closer to the electron through traversing the shielding cloud, one shows more and more sensitive to the bare electron charge. During this approaching process, the effective coupling constant becomes larger and larger, depending on the distance between the observer (photon) and the bare electron.

In terms of QCD, the coupling constant is defined by analogy with the fine structure constant of QED as:

$$\alpha_s = \frac{g_s^2}{4\pi} \quad (2.18)$$

Considering the non-Abelian character of the QCD interaction, there are not only pairs of quark and antiquark creation and annihilation, but also loops of gluons connected with the neighbour quarks/antiquarks in the vacuum. During the vacuum polarization, the former leads to colour charge dipoles to screen the colour field, while the latter strengthens the colour field. This two effects are in mixture and competition, but the latter dominates. The effective coupling strength increases when the distance is enlarged, which leads to quark confinement in hadrons. On the contrary, the quarks are approaching to liberty if the energy scale rises significantly, which is known as asymptotic freedom. At some point, the strong coupling constant is equal to unity, when the energy scale is defined as Λ_{QCD} , which is found to be about 200 MeV. If energy scale of an observer is beyond this value, the strong interactions can be computed using a perturbative QCD approach (see section 2.2.4), otherwise the non-perturbative effects dominate.

In general, the running of the coupling constant α is determined by the renormalization group equation:

$$Q^2 \frac{\partial \alpha}{\partial Q^2} = \beta(\alpha) \quad (2.19)$$

This β function encodes the dependence of the coupling constant on the energy scale, and its coefficients are extracted from the higher order (loop) corrections to the bare vertices of the theory. In QED, it is written as:

$$\beta_{QED}(\alpha_{em}) = \frac{1}{3\pi} \alpha_{em}^2 + \dots \quad (2.20)$$

The higher order terms dominate this function, which is related to the fact that the coupling parameter α_{em} increases when the energy scale becomes higher. According to the calculation in the first order, the coupling constant dependence is given by:

$$\alpha_{em}(Q^2) = \frac{\alpha_{em}(\mu_R^2)}{1 - \frac{\alpha_{em}(\mu_R^2)}{3\pi} \log \frac{Q^2}{\mu_R^2}} \quad (2.21)$$

where the μ_R^2 stands for an arbitrary energy scale chosen for QED renormalization. This formula describes the coupling constant evolution (running) between the renormalization scale and the physical energy scale.

In the case of QCD, the β function can be expanded in a perturbative way as:

$$\beta_{QCD}(\alpha_s) = -b\alpha_s^2(1 + b'\alpha_s + O(\alpha_s^2)) \quad (2.22)$$

where

$$b = \frac{33 - 2n_f}{12\pi}, b' = \frac{153 - 19n_f}{2\pi(33 - 2n_f)} \quad (2.23)$$

The n_f is the number of active flavors. As a matter of fact, the coupling constant of QCD can be dealt with perturbative expansion if the energy scale is beyond Λ_{QCD} . As the lower order terms dominate the expansion, concerning the first term, the coupling constant of QCD is written as:

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu_R^2)}{1 + \frac{33-2n_f}{12\pi} \alpha_s(\mu_R^2) \log \frac{Q^2}{\mu_R^2}} \quad (2.24)$$

The dependence of strong coupling constant on the energy scale is shown in figure 2.2, from which one can see the experimental validation for the running coupling constant of QCD.

2.2.4 Perturbative QCD

When the energy scale is beyond Λ_{QCD} , besides the strong coupling constant α_s , some observables can be treated in a perturbative way. The expression of an observable X is extended through perturbative series in powers of α_s :

$$X = \sum_{n=0}^N a_n \alpha_s^n = a_0 + a_1 \alpha_s + a_2 \alpha_s^2 + \dots \quad (2.25)$$

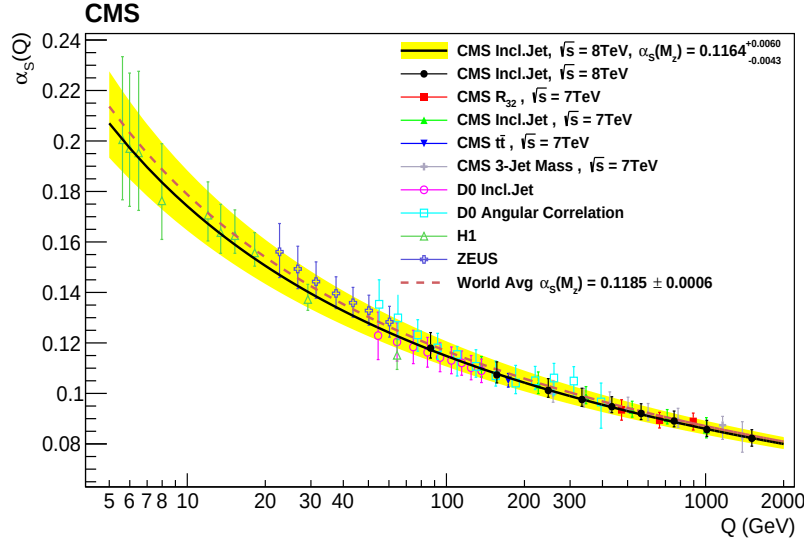


Figure 2.2: The running of α_s as a function of the scale Q is shown. The solid line and the uncertainty band are obtained by evolving the extracted α_s (from Z boson mass) values by using the 2-loop 5-flavour renormalization group equations. The dashed line represents the evolution of the world average value. Several experiments providing the measurements are presented, including CMS, D0, H1, ZEUS. The plot is taken from [4].

where a_n are the series of perturbative coefficients. If $\alpha_s \ll 1$, this expansion seems to converge quickly, so that the lower order terms already describe the accurate value of the observable sufficiently. On the other hand, there are still some divergences during the calculations, including ultraviolet or infrared divergences. The former enters the computation beyond leading order (LO) due to loop (virtual gluon) integral corrections (integration to infinite energy scale), which can be cancelled by either renormalization approach (see section 2.2.3) or regularization scenario. These two methodologies are independent from each other. The regularization is to redefine the quark charge, quark field, and gluon field in the Lagrangian to absorb the singularities. The latter results from the soft and collinear emissions of gluons, or from virtual gluon loops (integration from zero energy scale), which give rise to singularities at the boundaries of phase space. Since physical observables are always with finite values, they should be defined in an infrared safe way, and the short-distance effects should be separated from the divergent long-distance part, which could be absorbed into e.g. parton distribution functions (see section 2.2.5). Such a pattern is called collinear factorization.

2.2.5 Parton Distribution Functions

As explained in the last section, within the theoretical frame of collinear factorization, the parton distribution function (PDF) of e.g. a proton in QCD is defined as the probability density for finding a parton (quark or gluon) in the proton with a certain longitudinal momentum fraction x (Bjorken variable) at an energy scale Q^2 . Due to the inherent non-perturbative nature of partons that cannot be observed as free particles, parton distributions cannot be computed fully in perturbative QCD

(pQCD) approach. They have to be extracted from experimental measurements at specific energy scales Q_0^2 , and then evolved to another energy scales Q^2 . The extraction is typically realized by deep inelastic scattering measurements at the lepton-hadron collider (HERA) [5, 6] in DESY, and also by fixed target scattering experiments. The PDF of a hadron depends on the energy scale that an observer probes with, leading to various visible gluons and quarks inside the hadron. If the Q^2 is large, the fine structure inside the hadron is seen as substantial gluons and quarks carrying small fractions of the hadron momentum. On the other hand, a low Q^2 leads to a coarse resolution inside the hadron, and a few visible partons with large longitudinal momentum fractions of the hadron.

The PDF evolution is generally described by Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) matrix equation [7, 8, 9]:

$$Q^2 \frac{\partial}{\partial Q^2} \begin{pmatrix} q_i(x, Q^2) \\ g(x, Q^2) \end{pmatrix} = \frac{\alpha_s(Q^2)}{2\pi} \sum_{q_j, \bar{q}_j} \int_x^1 \frac{d\xi}{\xi} \times \begin{pmatrix} P_{q_i q_j}(\frac{x}{\xi}, \alpha_s(Q^2)) & P_{q_i g}(\frac{x}{\xi}, \alpha_s(Q^2)) \\ P_{g q_j}(\frac{x}{\xi}, \alpha_s(Q^2)) & P_{g g}(\frac{x}{\xi}, \alpha_s(Q^2)) \end{pmatrix} \begin{pmatrix} q_j(\xi, Q^2) \\ g(\xi, Q^2) \end{pmatrix} \quad (2.26)$$

where the $q_i(x, Q^2)$ and $g(x, Q^2)$ stand for the PDF corresponding to the quark or antiquark with flavor i and to the gluon, respectively. The $P_{ab}(\frac{x}{\xi}, \alpha_s)$ are series of splitting functions (forming the evolution kernel) that can be perturbatively expanded as a power series in α_s :

$$\begin{aligned} P_{q_i q_j}(z, \alpha_s) &= \delta_{ij} P_{qq}^{(0)}(z) + \frac{\alpha_s}{2\pi} P_{q_i q_j}^{(1)}(z) + \dots \\ P_{qg}(z, \alpha_s) &= P_{qg}^{(0)}(z) + \frac{\alpha_s}{2\pi} P_{qg}^{(1)}(z) + \dots \\ P_{gg}(z, \alpha_s) &= P_{gg}^{(0)}(z) + \frac{\alpha_s}{2\pi} P_{gg}^{(1)}(z) + \dots \end{aligned} \quad (2.27)$$

The leading order splitting functions $P_{ab}^{(0)}(z)$ indicates the probabilities of finding a parton of type a in a parton of type b with a fraction z of the longitudinal momentum of the parent parton and a transverse momentum squared much less than Q^2 [10]. This Q^2 is an arbitrary value called factorization scale providing the maximal radiation virtuality for PDF evolution, which corresponds to the boundary between the perturbative and non-perturbative regions. Figure 2.3 shows an example of PDF with different factorization scales.

2.3 Application to $Z + \text{jets}$ in proton-proton collision

As the main data analysis of this dissertation, the jet production in association with a Z boson process at the LHC invokes QCD factorization theorem into the Drell-Yan process [12]. Measurements can access the cross section of this process at particle (hadrons, leptons) level, while not able to directly probe the interactions among partons. This is due to the quark confinement effect. According to

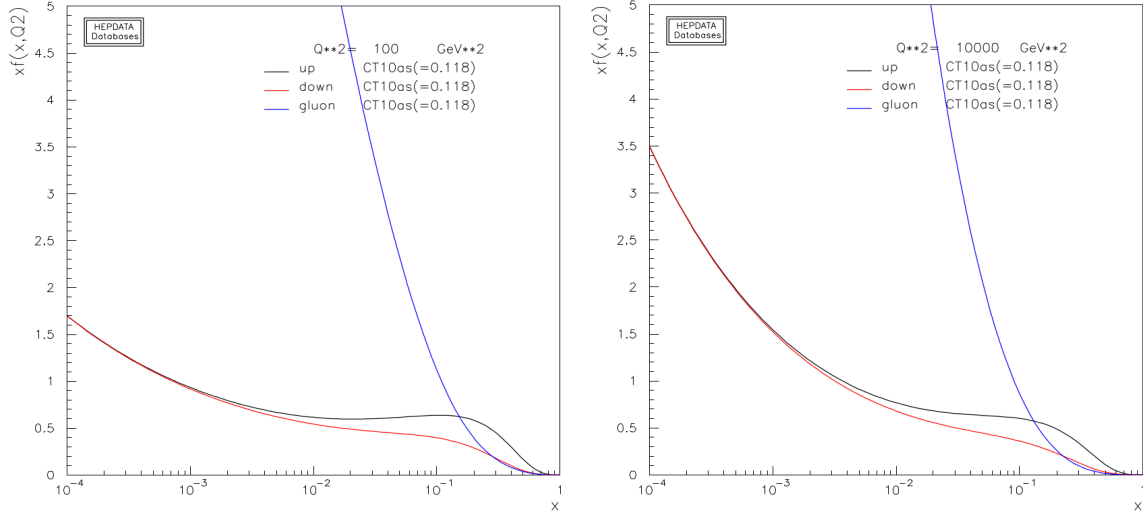


Figure 2.3: The PDF of gluon, up quark, and down quark at a scale of $Q^2 = 100 \text{ GeV}^2$ (left) and $Q^2 = 10000 \text{ GeV}^2$ (right). They are computed by CT10 [11] PDF with the nominal $\alpha_s = 0.118$.

the QCD factorization theorem, it is feasible to depict the process at parton level, which means computation of the cross section for parton-parton interactions. The relation between the cross section at particle level and the one at parton level is illustrated in figure 2.4. The two incident hadrons carry four-momenta of P_1 and P_2 , respectively. One parton in each of them participates in the interaction with a fraction of the hadron four-momentum as $p_1 = x_1 P_1$ and $p_2 = x_2 P_2$ respectively, where x_1 and x_2 are known as Bjorken variable (see section 2.2.5). The probability of a specific parton taking part in the interaction is interpreted by PDF at a designed factorization scale μ_F^2 . Then one can calculate the partonic cross section weighted by the PDF to obtain the cross section at particle (hadronic) level:

$$\sigma(P_1, P_2) = \sum_{ij} \int dx_1 dx_2 \left[f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) \times \sigma_{ij} \left(p_1, p_2, \alpha_s(\mu_R^2), \frac{Q^2}{\mu_R^2} \right) \right] \quad (2.28)$$

where i and j stands for each parton in the hadron. The partonic cross section depends on the renormalization scale μ_R^2 for the running coupling constant and the actual energy scale of the interaction.

The Drell-Yan process in proton-proton collisions can therefore be interpreted as lepton-antilepton pair production due to the annihilation of a quark from one proton and the corresponding antiquark from the other proton, by coupling to an intermediate Z boson or virtual photon. The center-of-mass energy of the annihilated quark-antiquark pair peaks around the Z boson mass of about 91.2 GeV. According to equation 2.28, the cross section of Drell-Yan process in proton-proton collisions is given by:

$$\sigma(pp \rightarrow l^+ l^-) = \sum_q \int dx_1 dx_2 [f_{q/p}(x_1, \mu_F^2) f_{\bar{q}/p}(x_2, \mu_F^2) \times \sigma(q\bar{q} \rightarrow l^+ l^-) + c.c.] \quad (2.29)$$

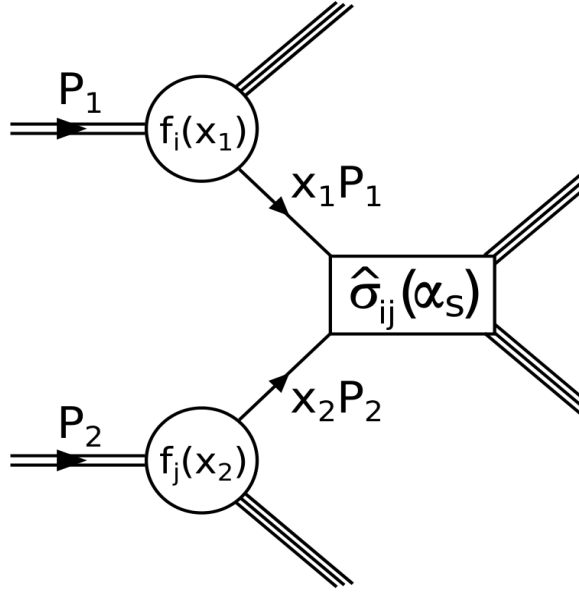


Figure 2.4: According to the QCD factorization theorem, the interaction between two hadrons is computed using parton model.

where the $f_{q/p}$ is the quark distribution function in a proton, similar for $f_{\bar{q}/p}$. $\sigma(q\bar{q} \rightarrow l^+l^-)$ gives the cross section of the lepton-antilepton pair production by a quark-antiquark pair. The *c.c* stands for the charge conjugation term. Due to the asymptotic freedom, the partonic cross section can be given by a perturbative expansion:

$$\sigma(q\bar{q} \rightarrow l^+l^-) = \sigma_0 + \alpha_s(\mu_R^2)\sigma_1 + \dots \quad (2.30)$$

Note that μ_R^2 and μ_F^2 dependence of the terms compensate each other for the integration if considering all the orders, which makes the total cross section invariant regardless of different choices of these factors. In practical way, the computation of infinite terms is not feasible, and special phase space selections are made for the specific physical study. The choice of μ_R^2 and μ_F^2 may not be fixed, which gives rise to different cross sections in theoretical calculation. As a result, there is also uncertainties introduced corresponding to the varied parameters. In terms of the Drell-Yan process, the typical choice of the factors is $\mu_F^2 = \mu_R^2 = M_{l^+l^-}^2$, which is the square of the invariant mass of the produced lepton pair.

Figure 2.5 shows some representative diagrams of the Drell-Yan process at leading order and higher order in α_s .

Based on the matrix element of Feynman diagram, one can compute both differential and integrated cross sections in a specific order. The details of calculation for the cross section of Drell-Yan process can be found for example in [13]. The total cross section at leading order is found to be:

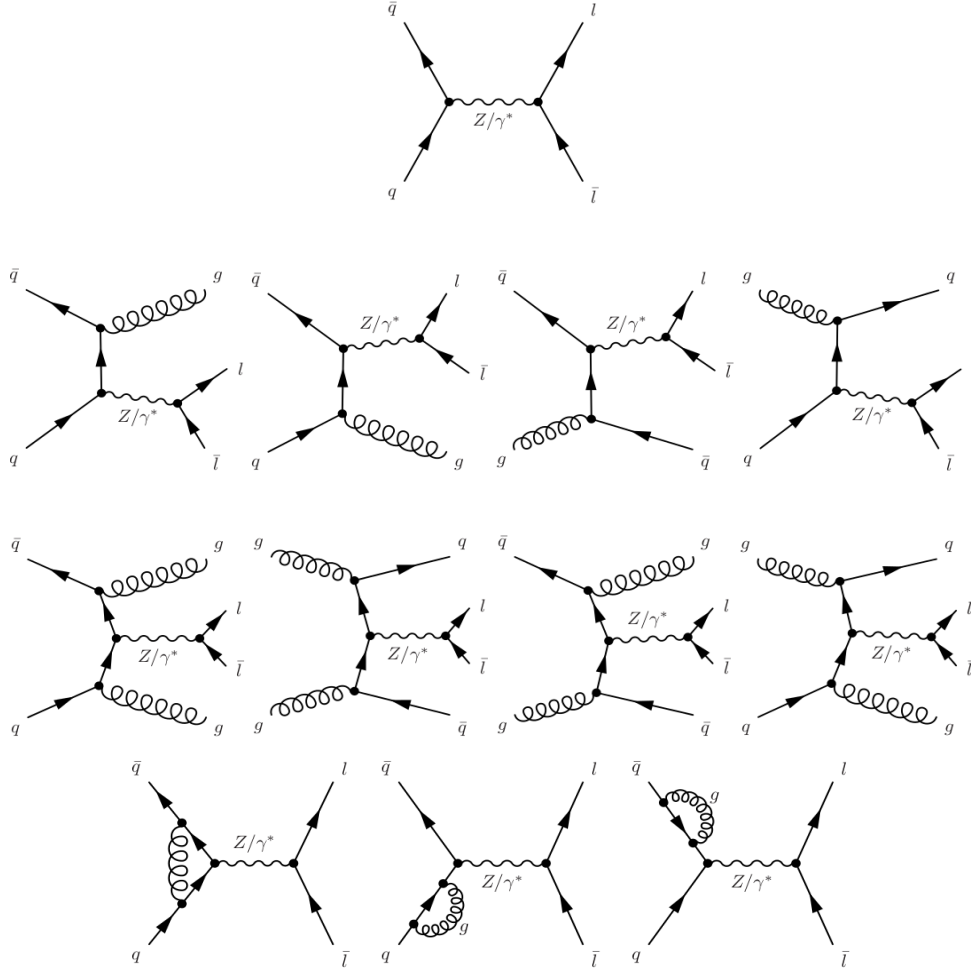


Figure 2.5: Feynman diagrams for the Drell-Yan process at different orders. The top diagram is the leading order matrix element of Drell-Yan process without strong coupling that has no parton production in the final state. The second row of diagrams shows the next-to-leading order of Drell-Yan process (first order in strong coupling) resulting in one parton produced in the final state. The third row represents the third order of Drell-Yan process (second order in strong coupling) leading to two partons in the final state. The last row provides the next-to-leading order of Drell-Yan process with virtual gluon loops in the initial state that will interfere with first order leading to α_s corrections. Figure is taken from [13].

$$\begin{aligned}
\sigma_0(q\bar{q} \rightarrow l^+l^-) &= \sigma_{\gamma^*} + \sigma_{interference} + \sigma_Z \\
&= \frac{4\pi\alpha^2}{3\hat{s}n_c} \left[Q_q^2 Q_l^2 + 2Q_q Q_l V_l V_q \times \frac{1}{\sin^2 2\theta_W} \times \frac{\hat{s}(\hat{s} - M_Z^2)}{(\hat{s} - M_Z^2)^2 + M_Z^2 \Gamma_Z^2} \right. \\
&\quad \left. + (A_l^2 + V_l^2)(A_q^2 + V_q^2) \times \frac{1}{\sin^4 2\theta_W} \frac{\hat{s}^2}{(\hat{s} - M_Z^2)^2 + M_Z^2 \Gamma_Z^2} \right] \quad (2.31)
\end{aligned}$$

where $V = T^3 - 2Q \sin^2 \theta_W$, $A = T^3$ corresponds to the weak mechanism for the quark and lepton,

and n_c is the number of quark colours. It takes into account the matrix element with exchanging virtual photon and Z boson, as well as the interference term. The $\hat{s} = x_1 x_2 s$ represents the center-of-mass energy of the two interacted partons coming from two incident hadrons with center-of-mass energy as s . The Q_q and Q_l are the electric charges of the quark and lepton, respectively. The Γ_Z is the total decay width of Z boson.

The matrix elements at next-to-leading order (NLO) start to involve strong coupling vertices. Gluons can appear in the initial state resulting in gluon-quark (gluon-antiquark) scattering to produce an additional parton in the final state in association with a Z boson. For even higher order cases like next-to-next-to-leading order (NNLO), gluon-gluon scattering can occur in the initial state to produce two partons in the final state with a Z boson. Moreover, virtual gluon loops also involve in the initial state. These higher order matrix elements in principle can be calculated in perturbative approach, which are considered as the pQCD corrections for the Drell-Yan process in leading order. Note that there are also interferences between each two diagrams with the same initial and final states.

It is certain that more higher order terms are taken into account, more accurate cross section of the process is obtained. For the higher order correction, gluon density increase significantly, which could dominate the contributions to the cross section compared with the densities of quark and antiquarks in the hadrons. A K-factor is introduced as the ratio of cross section with additional higher order terms to the cross sections with only lower order terms computed, such as (NLO+LO)/LO or (NNLO+NLO+LO)/(NLO+LO). Figure 2.6 shows the comparison of LO, NLO, and NNLO predictions on the differential cross section of inclusive Drell-Yan process (Z/γ^*+X) production in proton-proton collisions at center-of-mass energy of 14 TeV.

2.3.1 Parton Shower and Hadronization

Each parton presented in the hard interaction which is depicted by matrix elements, gives rise to successive gluon emissions. Quarks or antiquarks undergo gluon radiations. Unlike the QED radiation¹, gluons have self interaction feature, so they can couple to more additional gluons or quark-antiquark pairs, which leads to a shower effects of partons. This is called parton shower. It is dominated by collinear parton splitting and soft gluon emissions. The former can be similarly interpreted by the splitting function introduced by DGLAP parton evolution equation. After consecutive procedures of splitting and soft emissions, the parton shower stopped at the hadronization scale, which is in the magnitude order of 1 GeV. In order to determine the parton evolution, one could choose a kinematic variable to specify its value. The typical variables are transverse momenta or rapidities of partons, and the mass square of virtual partons. There are mainly two different parton shower approaches implemented in theoretical models, which are presented in section 2.4.1 and 2.4.4.

¹The leptons produced from the Z/γ^* radiate photons in the electromagnetic field. Muons have negligible QED radiations, while electrons suffer from significant radiations (bremsstrahlung effect). But this effect is marginal compared with QCD radiation in the final state.

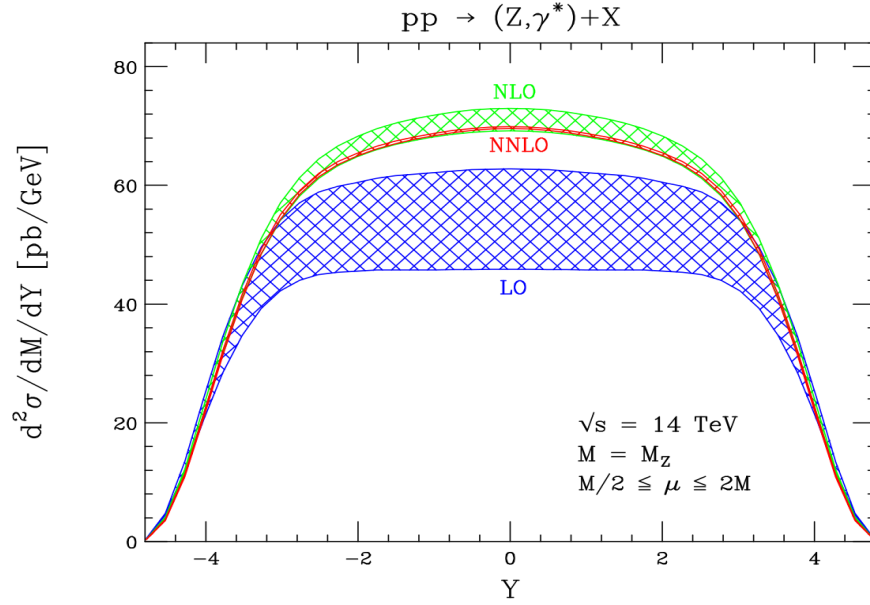


Figure 2.6: The rapidity distribution of an on-shell Z boson at the LHC. The LO, NLO, and NNLO results are included. The bands indicate the variation of the renormalization and factorization scales in the range of $\frac{M_Z}{2} \leq \mu_R(\mu_F) \leq 2M_Z$. Figure is fetched from [14].

Since it is impossible to calculate infinite terms of higher order corrections with matrix elements, parton shower supplies more partons to mimic the effects of the corrections. As the parton shower needs to be matched with matrix element computation, one needs to avoid double counting effects. The details of the matching approaches are presented in section 2.4.5.

When the energy scale decreases to the order of Λ_{QCD} (1 GeV), the partons formed during hard interaction and parton shower start to combine to form colorless hadrons. At this moment, the strong coupling strength becomes large enough that pQCD approach is not effective any more. One has to adopt phenomenological models (more details in section 2.4) to describe the subsequent process, which is called hadronization. In general, the gluons can split into quark-antiquark pairs that can be bounded with neighbouring pairs into different baryons or mesons, and can also become strings in the same hadrons. This is called string fragmentation model [15, 16], which is explained further in section 2.4.1.

There is a second scenario that is called cluster fragmentation model [17]. Partons close to each other enough are gathered into series of clusters, which form stable hadrons or heavy hadrons subsequently decaying into stable hadrons. The details of this model are interpreted in section 2.4.4. Figure 2.7 shows a schematic comparison of the two models.

2.3.2 Multiple Parton Interaction

At high energy of proton-proton collisions, there may be also additional parton-parton scattering effect in the same collision as the Drell-Yan process taking place. This kind of processes is called

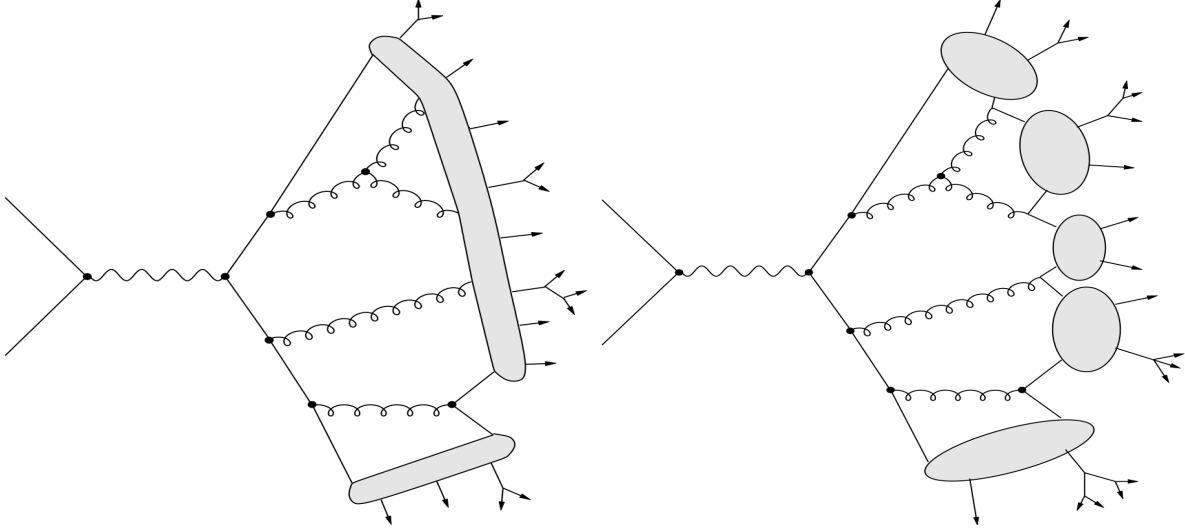


Figure 2.7: A schematic view of string fragmentation model (left) and cluster fragmentation model (right) for hadronization simulation. Figure is fetched from [17].

multiple parton interaction (MPI), resulting in the same final state as the Drell-Yan process with QCD interaction that have been mentioned before. Figure 2.8 shows an example of this kind of processes, with two additional gluons respectively from the two incident protons interacting to produce partons forming jets subsequently, which is called double parton scattering, compared with the Drell-Yan process with two partons produced in the hard interaction.

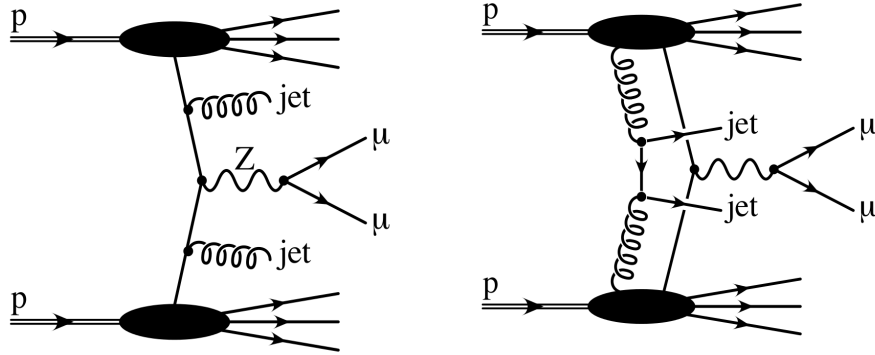


Figure 2.8: Comparison between Drell-Yan process with two partons produced (left) and double parton scattering process (right). Figure is fetched from [13].

2.3.3 Jet Definition

After the hadronization process, hadrons are formed and decay to more stable particles. They are mixed as particle streams. At this point, jets are defined as clusters of these particle streams, which connect the parton level interactions and the observed final states. Hence, an observed jet should be a signature of an explicit hadron cluster, leading to a question to the observer about the delineation

of jet clusters among plenty of particle streams. It is important to make the jet clusters from the hard interactions not influenced by soft emissions and collinear splittings, which means to be infrared and collinear safe.

There are several different algorithms developed for the jet cluster partition. The one for the analysis in this dissertation is called anti- k_T clustering algorithm [18, 19]. This algorithm involves the distance between each two objects (d_{ij}) in Lorentz invariant space, and the distance between each object and the beam (d_{iB}):

$$d_{ij} = \min(k_{t,i}^{-2}, k_{t,j}^{-2}) \times \frac{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}{R^2}, \quad (2.32)$$

$$d_{iB} = k_{t,i}^{-2} \quad (2.33)$$

where $k_{t,i}$, y_i , and ϕ_i are the transverse momentum, rapidity, and azimuthal angle of particle i , respectively. R is the algorithm parameter for users' choices. In the analysis of this thesis, we choose $R = 0.5$ at center-of-mass energy of 8 TeV and $R = 0.4$ at 13 TeV in proton-proton collisions.

After listing all the particles in the final state, this algorithm applies a sequential clustering, searching for the minimum of the distances. If i and j particles give the minimum d_{ij} , the particle i and j are combined into one particle ij , by summing their four-momenta [20], and then particle i and j are removed from the list of particles, while the new particle ij is added into the list. If d_{iB} is the minimum, the particle i is marked as a jet and subtracted from the particle list. The above implementation is repeated until each particle is subsumed into a jet. Note that the particles can be merged into a jet only when their distances are not larger than the parameter $2R$. This approach is specially suitable to account for soft or collinear radiated particles close to the hard particle (within $2R$), so they are clustered with the hard particle long before they are clustered along themselves. In this case, the jet is typically a conical one. However, if two hard particles have a distance $R < d_{ij} < 2R$, they will be distributed in two different jets, with the boundary clipped. Hence in this manner, the shape of jet is determined by the hard particles while being resilient for the soft particles. Figure 2.9 shows the jet cluster profiles with the anti- k_T algorithm. One can see that the regular cluster shape in this event sample.

In a more general content, the jet algorithm can be applied at three different levels: parton level, particle (hadron) level, and detector reconstruction level. In the first case, all the partons generated from matrix element and parton shower are considered into jet collections. The second case is stable particle level, which takes place after the hadronization and multiple partons scattering for all the produced partons. The last category is reconstruction level, in which case all the particles are proceeded with detector reconstruction, and then the jet algorithm is applied. In a word, the jets coming from these three cases are different but highly correlated.

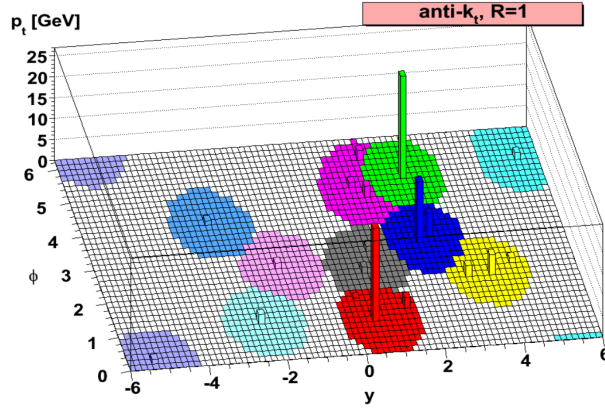


Figure 2.9: A sample event at parton level, including soft particles for anti- k_T jet algorithm, reflecting the shape of jets on the surface. This figure is fetched from [18].

2.4 Monte Carlo Simulations

In order to correct collected data for detector effects, one has to generate pseudo data based on Monte Carlo methodology (MC) in the frame of theoretical predictions. Namely, analytic and numerical computations are implemented to simulate realistic events by MC programs (generators). There are many MC generators developed and performed with different scenarios to simulate events. Such MCs also allow to compare the theoretical calculations with measurements. A general picture of an event during single proton-proton collision is shown in figure 2.10.

The common procedures of event generation for $Z + \text{jets}$ process in proton-proton collisions by MC generators are listed below:

- Computation for the differential cross section with two incident partons coming out of the collided protons in all the possible cases, i.e. $\sigma(q\bar{q} \rightarrow l^+l^-)$ in equation 2.29;
- The partonic cross sections are convoluted with suitable PDFs to get the hadronic colliding cross sections (see equation 2.29);
- With the calculated hadronic cross sections, the involved particles (including partons) are generated for each event within specific phase space;
- Modelling the initial state radiation (ISR) for the initial partons that may lead to complementary particles;
- Performing the decays of the particles with short lifetime (Z boson) to more stable particles;
- Modelling the final state radiation (FSR) for the particles (including parton shower) produced from the hard interaction;
- Simulating the multiple parton interaction effect in parallel with the hard parton interaction;

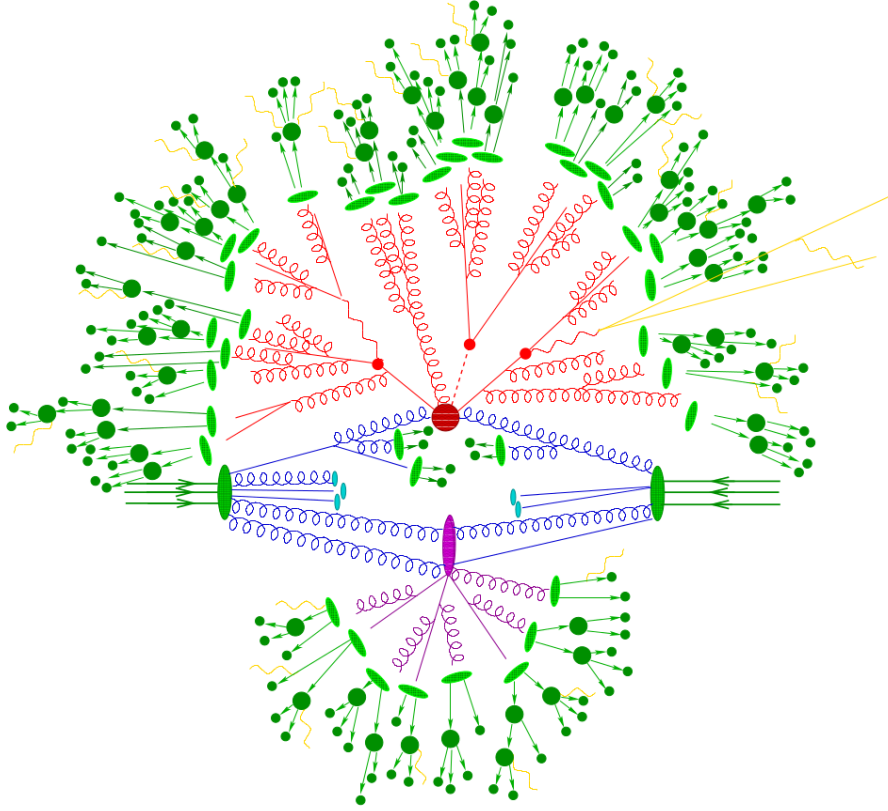


Figure 2.10: Sketch of a proton-proton collision as simulated by a complete Monte Carlo event generator. The red blob in the center represents the hard collision between two gluons respectively from two incident protons, producing quarks or antiquarks, gluons, and a boson. They are surrounded by a tree-like structure representing bremsstrahlung as simulated by parton showers. The boson subsequently decays to a pair of quark-antiquark. The purple blob indicates a secondary hard scattering event. Parton-to-hadron transitions are represented by light green blobs, dark green blobs indicate hadron decays, while yellow lines signal soft photon radiation. Figure is taken from [21].

- Concerning remnant collisions of partons for momentum balance and some physical quantity conservations;
- Simulating the hadronization for all the partons produced in the final state;
- Performing the unstable hadrons and leptons decaying before reaching the detector.

As mentioned before, there are several MC generators with different configurations. Some MC generators are only able to compute LO diagrams, and some of them cannot perform parton shower. The higher order with loop corrections are usually difficult to compute, and some generators can compute a limited number of processes with these loop corrections.

2.4.1 PYTHIA

The PYTHIA generator is a standard tool to generate high energy collision processes based on theoretical computations and phenomenological models. There are two main versions of programs, which are PYTHIA6 [22] and PYTHIA8 [23]. The former was constructed using Fortran language while the latter was rewritten in C++ code as a new release. The physical features of the two versions are the same. They implement matrix elements (ME) calculations at LO (no loop correction), initial state radiation (ISR), final state radiation (FSR), multiple parton interactions (MPI), and string fragmentation. They are interfaced with PDF libraries.

PYTHIA can generate hard QCD processes including $2 \rightarrow 1$ processes ($q\bar{q} \rightarrow Z/\gamma^*$) and $2 \rightarrow 2$ ($q\bar{q} \rightarrow Z/\gamma^*g$ and $qg \rightarrow Z/\gamma^*q$) ones. The latest version of PYTHIA8 can even produce events of some $2 \rightarrow 3$ processes. Due to lack of loop corrections, there are infrared divergences happening in the low transverse momentum (p_T) regions of the emitted partons. Because of the increased difficulty in higher order calculations, in order to take into account the significant multiple gluon emissions, one needs to implement a parton shower on top of the lower order of matrix elements to compensate for the gluon emission effects. The idea is to implement the parton evolution through computing the splitting functions of partons with varied virtuality energy scales. The evolution starts from the scale associated to the parton subject to parton shower down to the lower values until reaching a designed threshold.

In the case of FSR, the evolution forwards in physical time, with a single mother parton replaced by two daughter partons at each branching. Each daughter parton carries a fraction of the mother parton momentum, yielding the momentum conservation principle. The virtuality energy scale decreases through sequential gluons radiations or quark-antiquark pair productions, resulting in softer and softer radiations in the cascade. The virtuality scale can be also replaced by the transverse momentum of the splitted parton as the evolution variable.

In contrast to FSR, ISR is implemented through backwards in physical time, which means that the parton evolution begins from the hard interaction to the radiation position happening far before it. As a consequence, the emitted gluons becomes harder and harder when approaching to the hard scattering center. The virtuality energy scale decreases when complementing an additional sister parton to the parton in each branching to form the original mother parton. While, the Bjorken variable x increases through the backward evolution. The last scale of the ISR parton evolution is the factorization scale.

By use of the parton splitting functions, a differential branching probability per unit of evolution time for FSR and ISR is respectively given by:

$$\begin{aligned}\frac{d\mathcal{P}_{FSR}}{dp^2} &= \frac{1}{p^2} \int \frac{\alpha_s}{2\pi} P(z) dz \\ \frac{d\mathcal{P}_{ISR}}{dp^2} &= \frac{1}{p^2} \int \frac{\alpha_s}{2\pi} P(z) \frac{f(x/z, p^2)}{zf(x, p^2)} dz\end{aligned}\tag{2.34}$$

where z stands for the momentum fraction of a daughter parton radiated from a mother parton. The evolution variable p is given by:

$$\begin{aligned} p^2 &= (1 - z)Q_{ISR}^2 \\ p^2 &= z(1 - z)Q_{FSR}^2 \end{aligned} \tag{2.35}$$

where Q^2 represents the virtuality energy scale corresponding to each branching. After integration in a specific evolution time range $[t_0, t_1]$, the probability must be weighted by another factor to conserve the total branching probability. This factor is called Sudakov form factor, standing for the probability that no emission happens during this evolution time range.

When the virtuality energy scale is close to Λ_{QCD} , the outcoming partons from parton shower start to be bounded to form particles, which is the hadronization mechanism. In PYTHIA, Lund string fragmentation framework [15, 16] is applied, which is depicted in figure 2.7. The model postulates that partons within 1 fm are connected by gluon strings, the energies of which are proportional to their lengths. When the length between two partons gets larger, the energy of their connected string is high enough to create a new pair of partons, each of which is bounded with one of the original partons respectively to form a new string system. A series of strings keep breaking up until the created coloured quarks in their ends can form on-mass-shell hadrons, which can remain stably or decay to more stable particles. This phenomenological depiction is modelled according to observed hadronic spectroscopy and lattice QCD studies.

2.4.2 MADGRAPH 5

MADGRAPH 5 [24] is a MC tool that automatically generates matrix elements, such as decays and $2 \rightarrow n$ scattering processes. It can be interfaced with a second MC program like PYTHIA to implement parton shower and hadronization effects. This generator used to generate events at tree level only (without any loop correction)². At tree level, this generator can compute the matrix elements for any Lagrangian based model (renormalizable or effective). The matrix element evaluation is treated using helicity wavefunctions and amplitudes [25]. Without loop corrections, the matrix element calculation at tree level suffers from the soft and collinear divergences. In order to get rid of these singularities, one needs to specify the phase space to perform matrix elements away from the soft and collinear divergence regions. The missing phase space will be taken into account by parton shower.

Before generating the matrix elements, users need to specify a process with the initial and final state particles, then the program is executed based on the following iterative algorithm:

- Generate all of the probable interaction vertices and the intermediate particle currents according to the specific model (Lagrangian).

²The new version of MADGRAPH 5 has been incorporated with another generator AMC@NLO (see next section) which is available to provide NLO loop corrections for many processes.

- Take into account all the external legs (particles) including the particle and antiparticle exchange for the incident particles. A boolean variable called “from_group” flag, which is used to indicate whether a given particle results from a combination of particles in the previous step (true), or just copied from the previous step (false). For the moment, this flag is initialized to “true” for all the external particles.
- If there is one vertex combining N external particles, create the combination $[(1, 2, 3, \dots, N)]$ for them if at least two particles have the “from_group” equal to true.
- Create all appropriate mergings of particles (resulting from interactions) with at least one intermediate particle current with “from_group” equal to true.
- Set the “from_group” to true for those newly combined particles, and to false for the particles that are not merged in this iteration.
- Repeat the procedure from the third step for the reduced set of external particles if there are at least three external particles remained.
- Stop the algorithm when at most two external particles left.

Based on such a recursive approach, the program create subdiagrams from the diagrams through merging legs, with the flag avoiding double counting effect.

For the produced Feynman diagrams, the code for matrix element evaluation is implemented in terms of successive calls to a helicity amplitude function library, which is either HELAS [25] or ALOHA [26]. This library is able to compute the amplitudes of arbitrary diagrams at tree level. Then the generator can obtain the differential cross sections which are proportional to the amplitude squares.

2.4.3 MADGRAPH_AMC@NLO

The MADGRAPH_AMC@NLO program [27] is a further development on the top of MADGRAPH 5. It provides the computations of the matrix elements in both LO and NLO for an arbitrary process, and their matching to parton shower. The methodology of LO matrix elements is adopted from MADGRAPH 5 generator, and the NLO computation is done by AMC@NLO framework [28]. It also includes some additional capabilities such as the merging of event samples with different light parton multiplicities. At the LO, the theoretical model is constructed based on a given Lagrangian with its Feynman rules. However, in order to compute NLO corrections, additional specific terms are needed for the reduction of loop amplitudes such as Ultraviolet counterterms, which cannot be held by the automatic FEYNRULES program [29] for the moment. Fortunately, these terms are finite and small number of quantities such that they can be given by hand.

In the case of LO computation, one can always obtain a set of physical unweighted events from the short-distance cross sections, using the Les Houches event files [30]. While, the NLO cross sections

cannot be unweighted. The NLO computations suffer from soft and collinear divergences. In order to solve this issue, AMC@NLO has been implemented with special configurations, which is called FKS subtraction [31, 32, 33]. The weighted events contain the normal events that are within infrared safe region, and the counterevents corresponding to soft and collinear singularities.

2.4.4 SHERPA

SHERPA [34, 35] is also a MC event generator that provides complete hadronic final states in simulations of high energy particles collisions. It has two main sets of versions that are SHERPA 1 and SHERPA 2, the former dealing with LO matrix element computations (up to 8 partons in Born level diagrams), while the latter able to integrate NLO amplitudes for the implemented models. Moreover, SHERPA also implements parton shower simulation and the matching with its calculated matrix elements.

The implementation of parton shower is called dipole showering, which takes into account soft gluon coherence. In this approach, each quark or antiquark is combined with a colour partner (parton) to form a colour charge dipole, and each gluon is connected with two colour partners to form two colour charge dipoles respectively. Each dipole splits into two dipoles when radiating a gluon, which should be within the two cones of a designed opening angle around the directions of the two colour charges. The four-momenta of the produced dipoles are dependent on the gluon emission kinematics. Such kind of dipole splitting processes occur in sequential way until close to a minimum scale of dipole momentum. Therefore in this manner, the gluon emission is generated from the dipole radiation, instead of parton splitting (in the case of PYTHIA). This pattern adopted by SHERPA has been found to successfully reproduce the leading logarithmic terms of parton shower [17].

The hadronization simulation in SHERPA is the cluster fragmentation model as introduced before, which is different from the configuration in PYTHIA generator. The emitted gluons are forced to split into quark-antiquark pairs, and they are combined with their colour partners to form dipoles, which are merged into clusters that form hadrons (see figure 2.7).

2.4.5 Matching between matrix elements and parton shower

As mentioned before, the MC generators like PYTHIA, MADGRAPH 5 relying on matrix elements at tree level only suffer from soft and collinear divergences, which are not appropriate for describing the substructure of a jet. At this point, parton shower is needed for overcoming these divergences and mimicking higher order corrections. Moreover, the hadronization effects are also necessary to apply to partonic final state, with aim to compare with the observation. In a practical way, some MC generators have to be interfaced with another generator properly, such that their calculated matrix elements are matched with parton shower to depict a realistic physical process. For instance, the production of n partons in a matrix element computation could be replaced by an $n - 1$ partons production in a matrix element with complementary one parton from parton shower.

In the case of jet production with Drell-Yan process, there could be multiple partons produced in the final state, which leads to a large jet multiplicity production. Due to the limited order computation in MC generators, parton shower are implemented for each external leg starting from the virtuality scale of the parent partons. As shower programs only give a leading logarithmic (LL) approximation, to partly supplement the lack of hard emissions, the shower scale has to be increased to a large value. Furthermore, if the matrix element calculations are beyond leading order, their matching with parton shower becomes more complicated due to the subtraction of double counting effect.

For the interests of this thesis, we briefly introduce two matching schemes at tree level, CKKW and MLM matching schemes. In general, they define a jet measured in the final state and take into account the relevant cross sections for different numbers of jets within the maximal capability of jet multiplicity consideration. The parton shower evolution is constrained by the required jet multiplicity in the configurations, which means that it could not produce any extra jet if a higher jet multiplicity would be reached.

The CKKW (Catani-Krauss-Kuhn-Webber) matching prescription [36] was proposed in the e^-e^+ annihilations into hadronic final states [37]. In this scheme, it defines a jet through a k_T algorithm [38]. In the case of proton-proton collisions, two partons in the final state give their relative transverse momentum squared as:

$$k_{Tij}^2 = \min(p_{Ti}, p_{Tj})^2 \times \frac{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}{R^2} \quad (2.36)$$

where p_{Ti} , y_i , and ϕ_i represent the transverse momentum, rapidity, and azimuthal direction of parton i , respectively. R is a designed parameter of the jet algorithm, which is the chosen cone size of a jet opening angle. If k_{Tij} is larger than a designed value k_{T0} , the two partons belong to two different jets, otherwise they are incorporated into the same jet. This k_{T0} is called merging scale, which separate the matrix elements (ME) and parton shower (PS) in two regions, with ME contributing to the space above the merging scale while PS domains below the scale. The generated matrix elements and parton shower come along with reweighting factors related to different scales in running α_s and Sudakov form factors. Note that the starting scale for parton shower evolution is determined from the mother parton mass in ME, and the Sudakov weight of parton shower is attached to the one of ME. As a result, there is no hard jet produced during the jet evolution. This scheme is adopted by SHERPA generator.

The MLM (Michelangelo Mangano) matching scheme [39] begins with the event generation of parton-level configurations for all parton multiplicities up to an allowed number in the final state. Their transverse momenta p_T^{parton} and the angular distance in (η, ϕ) plane between each two partons ΔR_{ij} have minimal thresholds p_T^{min} and R_{min} , while their absolute rapidities $|y_{parton}|$ should be below a threshold y_{max} :

$$p_T^{parton} > p_T^{min}, |y_{parton}| < y_{max}, \Delta R_{ij} > R_{min} \quad (2.37)$$

Those thresholds are called generation parameters that are initialized as the same values for all the parton multiplicities generated. After the matrix element generation, the parton shower is implemented using PYTHIA for instance. The parton evolution begins at the scale that is determined by default PYTHIA configurations according to the kinematics and colour connections of the event, constrained by the hard interaction (matrix element) scale. After the evolution, a jet cone algorithm is applied to the partons produced in the perturbative phase of the shower. Similar to the event generation settings (expression 2.37), a set of matching parameters, including minimal transverse momentum, maximal rapidity, and a constraint on the cone size of jet opening angle, are required for the jets defined from those partons. The matching procedure starts from the hardest parton and searches for the closest jet in (η, ϕ) plane. If the angular distance between the parton and the jet center is smaller than 1.5 times of the jet cone size, they are matched. Then the program moves to next partons and looks for another adjacent jets to perform the subsequent matchings. If a matching fails, which means a parton cannot be matched with any jet, the event is rejected. As in CKKW method, each event comes with a Sudakov reweighting to prevent double counting. For the parton samples with jet multiplicity larger than parton multiplicity generated by matrix elements, the extra jets are originated from parton shower and are required to have lower transverse momenta than the matched jets.

In the case of NLO matrix elements, generated events are classified into two groups, which are called S events and H events respectively. The formers are produced with NLO matrix elements implemented with truncated parton shower, which is NLO-vetoed to avoid double counting. The latter are generated through LO matrix elements attached with parton shower similar to the way of MLM. If one wants a parton sample with parton multiplicity of n , the S events with n partons and the H events with $n - 1$ partons in the matrix elements are needed to be merged. Each H event provides an additional parton from shower. The merging scheme that adopts this approach in MADGRAPH_AMC@NLO is called FxFx [40], and MEPS@NLO [41] in SHERPA generator.

In summary, by comparing the predictions of different MC generators that adopt different approaches, one could figure out the potential problems or success of them. The prediction accuracy is influenced by the ME calculation, and the matched parton showering model as well. One could also perform a MC generator with different PDF sets that give rise to different cross sections. This can help physicists to understand the fundamental particle interactions more accurately.

Chapter 3

The Experimental Setup

This chapter provides a general introduction about the experimental facilities that are supporting the present physical analysis work. The Large Hadron Collider (LHC) is described firstly, including its geometric structure and several crucial physical characteristics. Then it explains the beam conditions and implementations in the unprecedented accelerator. The second part of this chapter presents the Compact Muon Solenoid (CMS) detector in details, which is complied to detect the particles produced in proton-proton beam collisions at LHC. Furthermore, the acquisition and operation methodology with respect to the data collected by the detector for physical analyses is introduced. The comprehensive explanations about LHC and CMS can be found in [42] and [43], respectively.

3.1 The Large Hadron Collider

3.1.1 Introduction

The LHC was built by the European Organization for Nuclear Research (CERN) between 1998 and 2008, which is the largest particle accelerator all over the world. It is located inside a circular tunnel with 27 km length and 3.8 m width, which was dug underground as deep as about 100 m penetrating the France-Swiss border between Jura mountains and Geneva lake. It was constructed for the large electron-positron collider (LEP), which was replaced by LHC in 2000. The purpose was to establish it as a proton-proton collider with the central-of-mass energy as 14 TeV, each proton beam with nominal energy of 7 TeV. It was also designed to accelerate heavy ions up to an energy of 5.02 TeV per nucleon pair.

The first physical run of LHC started in spring of 2010 and finished at the end of 2013 (RunI). Each proton beam carried the energy as 3.5 TeV at the beginning, and increased to 4 TeV in 2012. After a shut down for two years, the LHC RunII has been launched in 2015, each beam with a nominal energy of 6.5 TeV.

The main motivations of designing the LHC include precise measurements of the processes predicted by Standard Model Theory (SM) of particle physics, and searches for new physics. The

discovery of Higgs boson at LHC in 2012 was one of the main objectives of the LHC. Then the subsequent measurements of Higgs boson properties started to be conducted. Benefiting from the large statistics of the data, we can perform precision measurements in a physical phase space which was never reached before. Due to the unprecedented energy, we can search for the large family of particles predicted by Supersymmetric theory beyond SM, as well as dark matter. By use of the heavy ion (lead) collisions, one can study the quark-gluon plasma which is a state before hadrons are formed, which is expected to have existed in the early universe.

The LHC tunnel is made of two adjacent parallel beamlines which intersect at four points, where the two beams collide from opposite directions. The four points correspond to four detector experiments, including ATLAS, CMS, LHCb, and ALICE, as shown in figure 3.1.

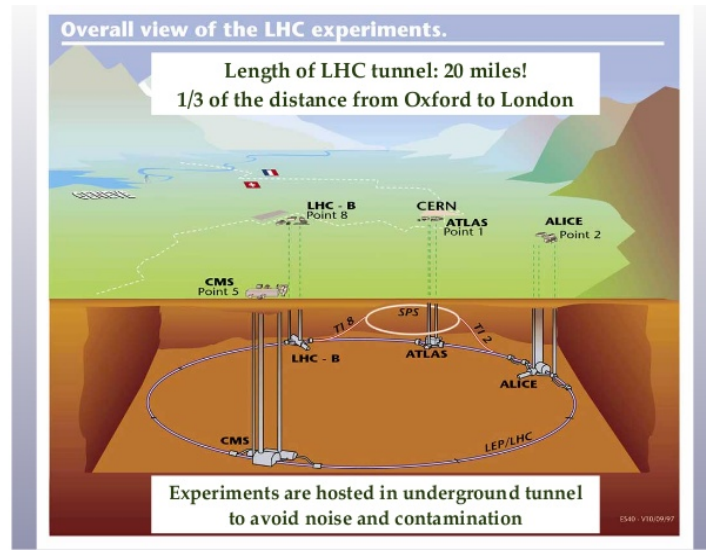


Figure 3.1: The locations of the main experiments of the LHC. The SPS is a pre-accelerator for the injected proton bunches into the LHC.

The whole collider ring consists of eight arcs and eight linear sections. The formers contain in total 1232 dipole magnets to realize the circular trajectories of the beams, and latters hold 392 quadrupole magnets and several accelerating cavities to collimate the beams and to accelerate each beam from 450 GeV to 6500 GeV. Especially, the quadrupole magnets become much stronger at the locations close to the four crossing points, which is aimed to maximize the probability of bunch interaction. There are also additional magnets of higher multipole orders to make up with small imperfections in the field geometry. The superconducting magnets are composed of copper-clad niobium-titanium. The peak magnetic field that they form can be above 8 Tesla, benefiting from the cryogenic system containing about 96 tonnes of superfluid helium 4, which keeps the operational temperature of 1.9 K (-271.3°C). Figure 3.2 and 3.3 show the dipole magnets of the LHC and the magnetic field profile of the dipole magnets, respectively.

The two beam pipes are maintained at ultrahigh vacuum, which is to reduce collisions between the protons and the gas molecules. Each proton beam consists of up to 2808 bunches at full intensity,

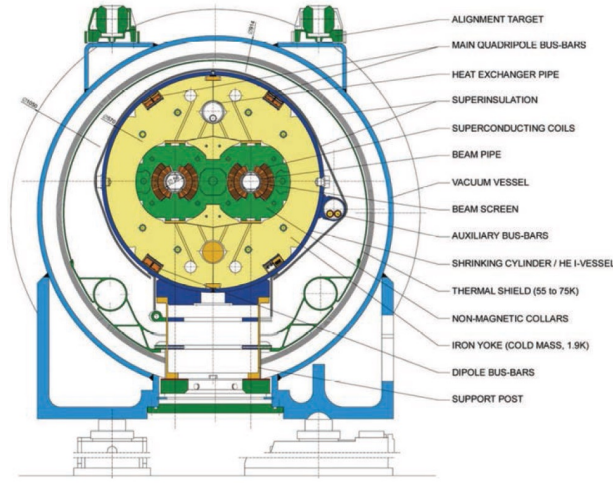


Figure 3.2: The dipole magnets of the LHC together with the outward facilities.

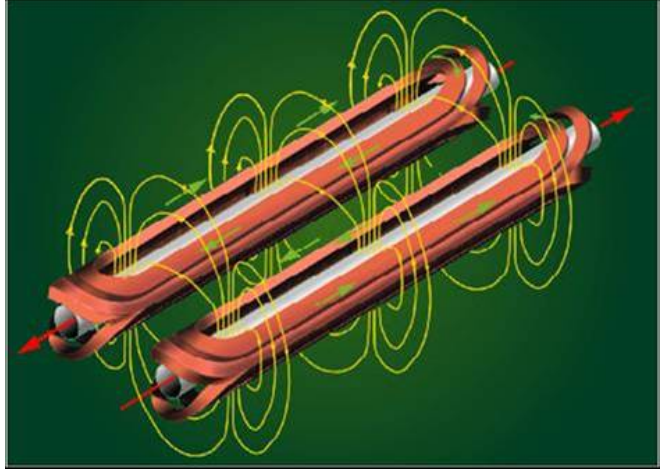
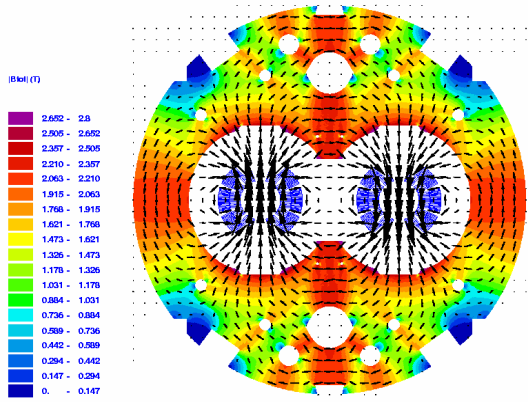


Figure 3.3: The magnetic field profile along the transverse cross section of the dipole magnets (left) and in three dimensional view (right).

and each bunch contains 115 billion protons. The bunch is typically with transverse dimensions of the order a mm, and it is squeezed to approximate 16 microns before a collision. The protons inside each bunch are ultra-relativistic with 0.999999991 times of the speed of light, therefore they spend less than 90 microseconds to travel along once the ring. In general, the distance between two bunches is 7.5 m, corresponding to 25 ns. If the nominal energy of each proton of the beam is 7 TeV, the total beam energy is approximately 362 MJ, which is equal to the energy of 77.4 kg TNT.

3.1.2 Beam luminosity

The rate of events of a given process is proportional to the cross section of this process. The luminosity is defined as the number of produced events (N_{event}) per cross section (σ_{event}). In general, it is written as:

$$L = \frac{N_{event}}{\sigma_{event}} \quad (3.1)$$

More precisely, it can be evaluated as instantaneous or integrated luminosities. The former is given by the number of events (dN) generated at a definite time (t) within a infinitesimal time interval (dt) with respect to the effective interaction cross section (σ):

$$L_{inst}(t) = \frac{1}{\sigma} \frac{dN}{dt} \quad (3.2)$$

The latter is the integral of the former during a certain time window:

$$L_{int} = \int_{t_0}^{t_1} L_{inst}(t) dt \quad (3.3)$$

The luminosity relies on the particle beam characteristics. For two beams with number of bunches as N_1 and N_2 , respectively, each bunch containing n_p protons, a beam crossing frequency is f , and transverse sizes of effective interaction of two beams $\sigma_x \sigma_y$, the instantaneous luminosity is expressed as:

$$L_{inst} = \frac{N_1 N_2 n_p f}{4\pi \sigma_x \sigma_y} \quad (3.4)$$

The $\sigma_x \sigma_y$ is equivalent to the combination of relativistic Lorentz boost factor γ , the normalized transverse beam emittance ϵ (a property for the average spread of particle coordinates in position-momentum phase space), and the beta function at the collision point β^* (the distance from the focus point at which the beam width is twice as wide as the focus point). Due to same numbers of bunches in the two beams (N_b), equation 3.4 can be rewritten as:

$$L_{inst} = \frac{N_b^2 n_p f \gamma}{4\pi \epsilon \beta^*} \quad (3.5)$$

The luminosity of the beam varies during each physical run at LHC, which decreases along with the degradation of the intensity and emittance of the beam. This phenomenon mainly results from the interactions between the two circulating beams. In addition, the intra scattering of each beam and the extra collisions with residual gas in the ring also consume the beams. The evolution of the instantaneous beam luminosity with respect to time can be approximated as:

$$L_{inst}(t) = \frac{L_{inst}(0)}{(1 + t/\tau)^2} \quad (3.6)$$

where $L_{inst}(0)$ corresponds to the initial instantaneous luminosity, and τ stands for the time when the luminosity approaches the $1/e$ of the initial value, and it is measured to be 14.9 hours as the average luminosity lifetime.

Figure 3.4 and 3.5 show the evolution of the maximal instantaneous luminosity and the integrated luminosity recorded with the four detector experiments for proton-proton collisions at LHC in 2012

and 2015, respectively.

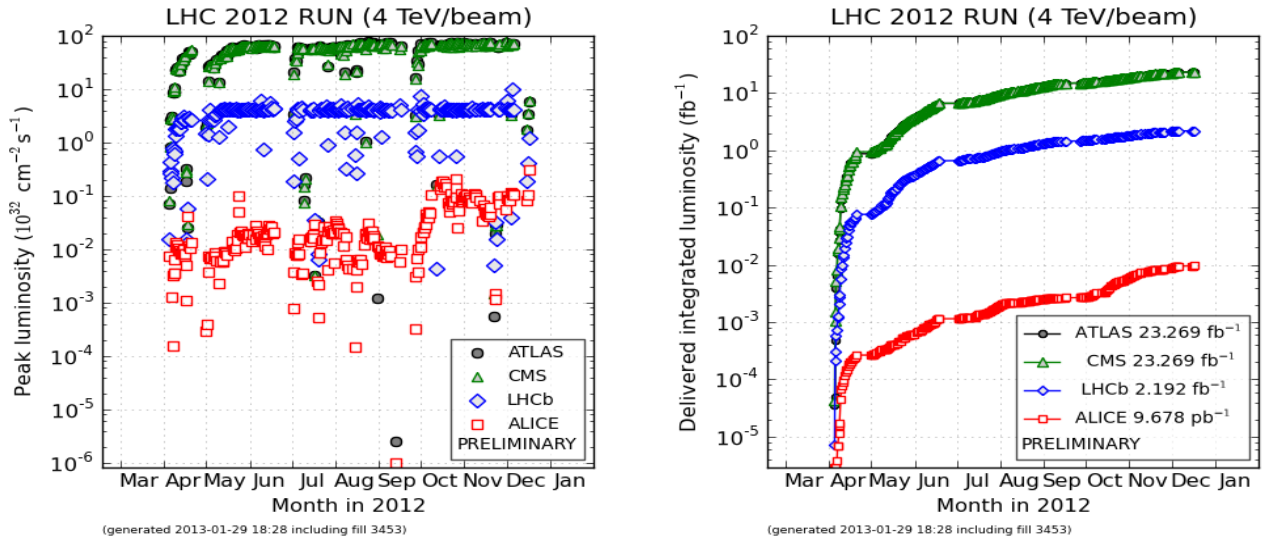


Figure 3.4: The maximal instantaneous (left) and integrated (right) luminosities recorded by LHC for proton-proton collisions at 8 TeV. The plots are taken from [44].

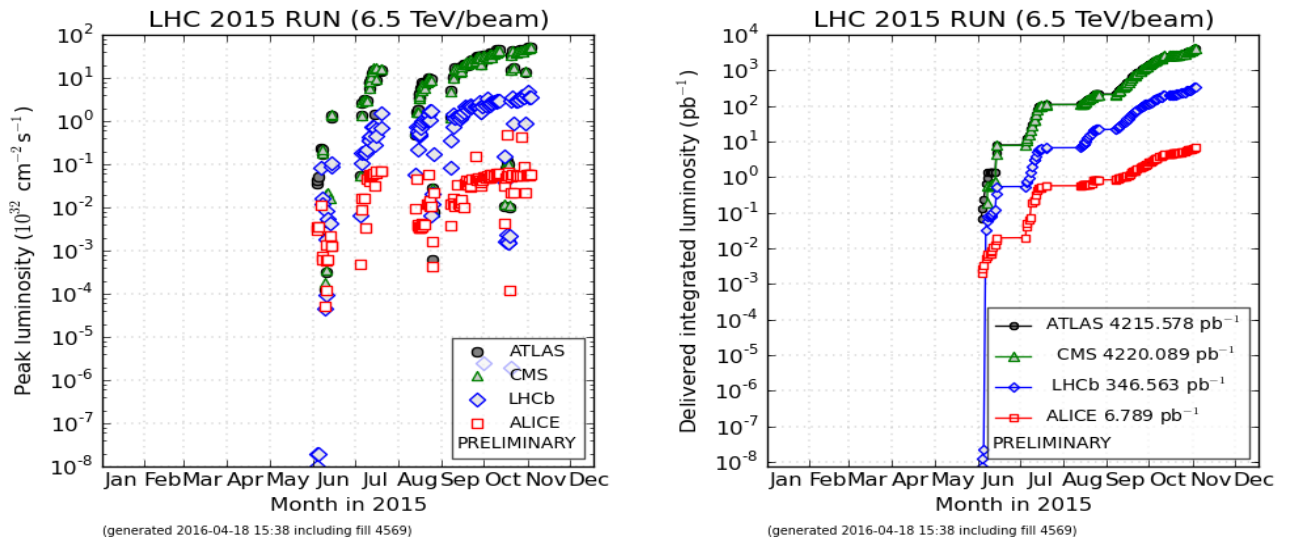


Figure 3.5: The maximal instantaneous (left) and integrated (right) luminosities recorded by LHC for proton-proton collisions at 13 TeV. The plots are taken from [44].

Figure 3.6 provides an schematic view of the LHC schedule. The collision energy is expected to reach 14 TeV with an integrated luminosity of 3000 fb^{-1} for HL-LHC.

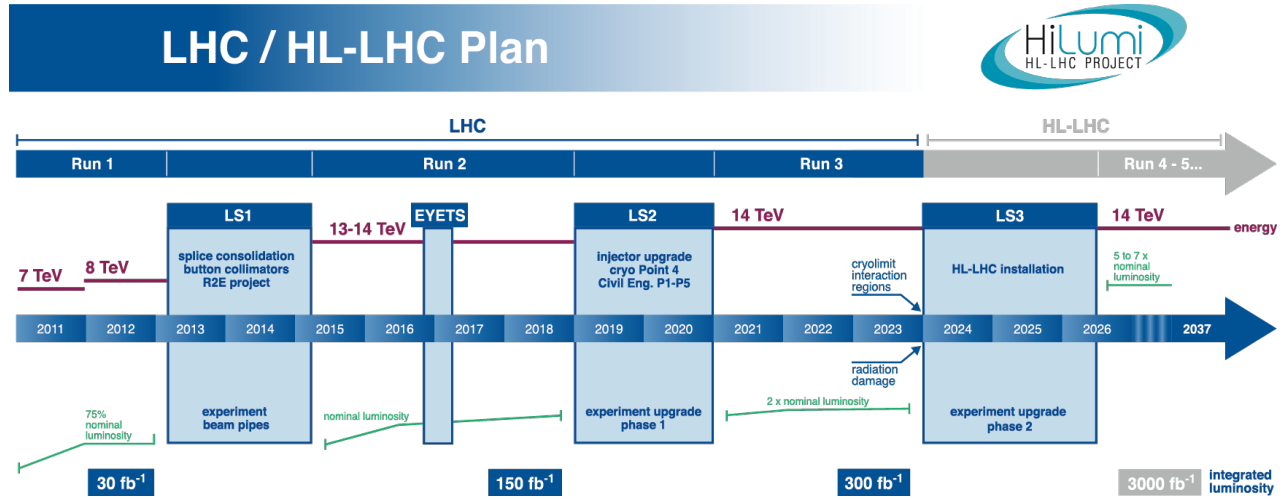


Figure 3.6: The global timeline view of LHC.

3.1.3 Injection chain

The LHC is a successful machine to accelerate the protons to super high energy. It needs the supports from the injection system, which is responsible to originate protons and to accelerate them to the given energy for LHC beam supply. The pre-accelerating system is composed of a succession of machines, each of which boosts the protons to higher energy and deliver them to the subsequent device. The graphic view of the system is shown in figure 3.7. We can see the sequential implementations. The injectors also provide beams to several other experiments.

- Proton source

The protons traveling and colliding in the LHC with unprecedented energy are obtained from a bottle of hydrogen, as shown in figure 3.8. The bottle is connected with a cylindrical device which is called duoplasmatron [45]. After some hydrogen gas is imported into this machine, there is an electric field used to strip the electrons from the hydrogen atoms to extract the protons. The average energy of these protons is about 100 KeV, and they are pulsed at each 100 μ s.

- Linear accelerator

The protons obtained from hydrogen gas are delivered to a linear accelerator of LINAC 2, which has length of 30 m. As shown in figure 3.8, it complies with multiple radiofrequency cavities to charge cylindrical conductors. When the protons pass through, the cylindrical conductors are alternately charged positive or negative, so that the protons are always accelerated between two adjacent conductors. There are also small quadrupole magnets to collimate the proton beams in time. After this step, the protons have a nominal energy of about 50 MeV.

- Proton Synchrotron Booster

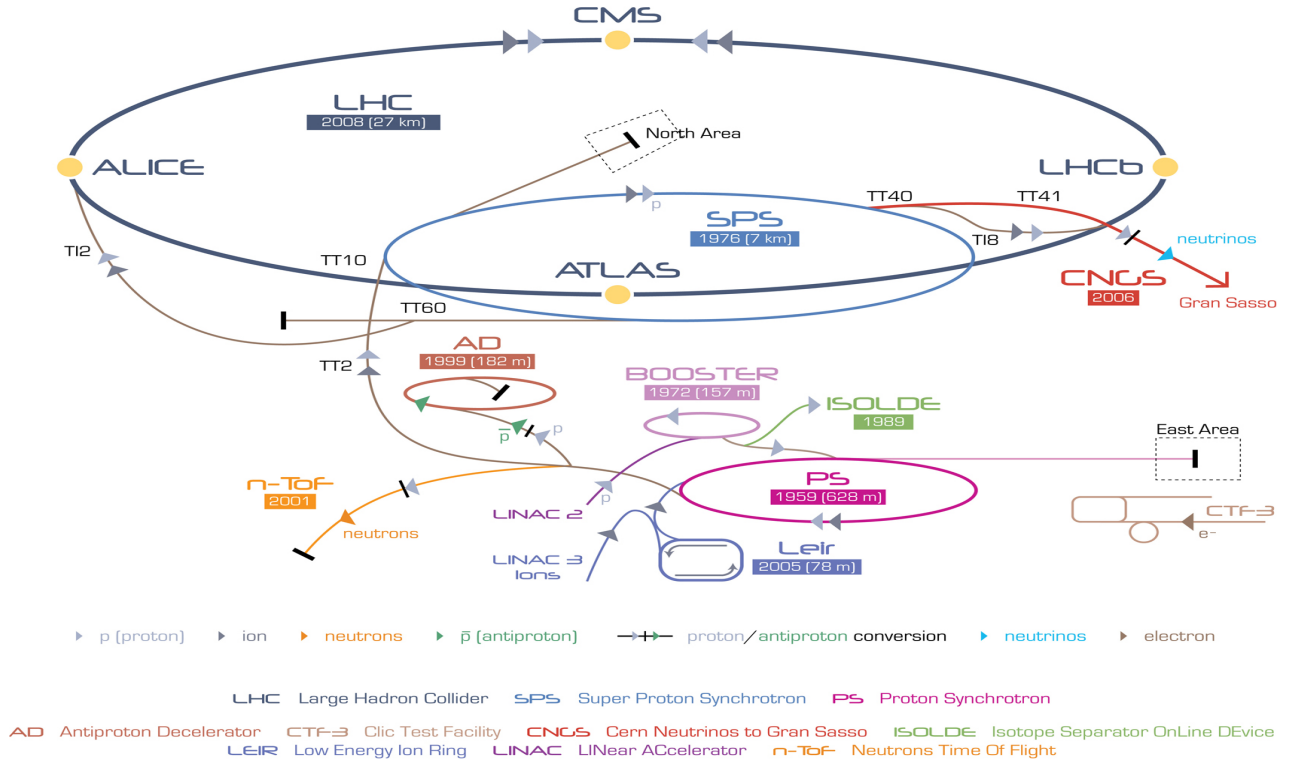


Figure 3.7: The accelerator complex of CERN. The injection system includes the LINAC 2 for originating proton beams and the LINAC 3 for producing ion beams, the BOOSTER and PS for primarily accelerating protons, and the SPS for furtherly increasing the beam energy to feed the LHC.

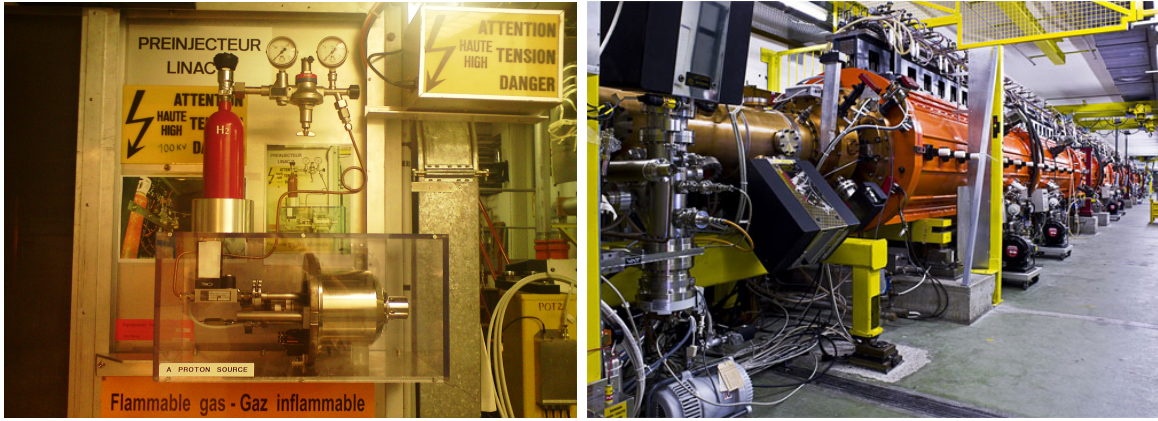


Figure 3.8: The bottle of hydrogen gas for manufacturing protons, connected with a duoplasmatron device [45] (left), the linear accelerator of LINAC 2 (right).

Being a synchrotron, it is the first and smallest circular proton accelerator of the beam injection system. There are four superimposed rings with a radius of 25 m, which receive the protons from LINAC 2 to accelerate them up to 1.4 GeV. Figure 3.9 presents a photograph of the booster system. It uses an electric field to accelerate the protons in straight paths and a magnetic field in curved regions to bend the proton trajectory as well as focusing the beams.

- Proton Synchrotron

As shown in figure 3.9, the Proton Synchrotron has 277 conventional electromagnets, which include 100 dipoles to bend the beam trajectory along the storage ring. The total circumference is 628.3 m. After the protons come out of the Booster, they are sent to this synchrotron and accelerated up to 25 GeV. In addition to protons, the synchrotron can accelerate heavy ions as well. The proton bunches are splitted in the PS into smaller bunch distance of 25 ns.

- Super Proton Synchrotron

As the second largest machine in CERN accelerator complex, the Super Proton Synchrotron has 1317 conventional electromagnets with 744 dipoles for bending the beam trajectory. It acquires the protons from the Proton Synchrotron, and continuously accelerate them up to 450 GeV using its 7-km circumference. Similar to the Proton Synchrotron, it can also handle many different kinds of particles such as heavy nuclei, electrons, positrons, and antiprotons. Benefiting from this enormous machine, the W and Z bosons were discovered in last century. Figure 3.9 also shows the photograph of the Super Proton Synchrotron. After the proton beams are operated by it, they are directly transferred to the outmost LHC storage ring.

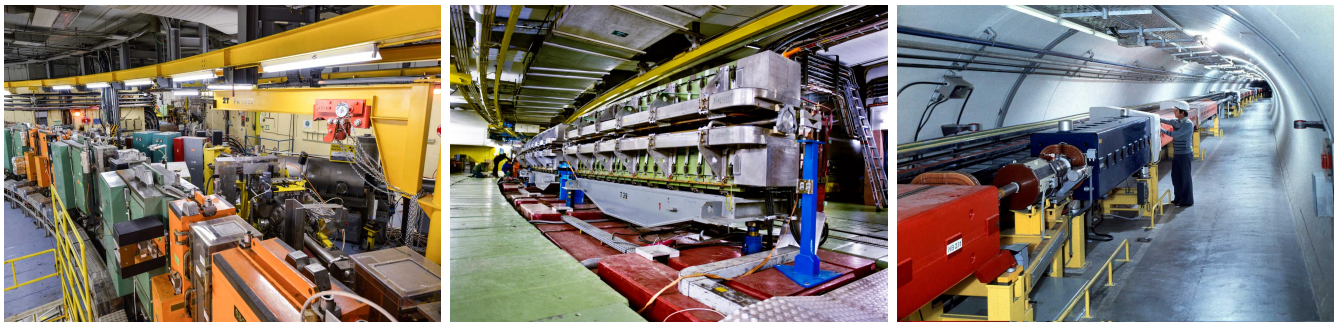


Figure 3.9: The Proton Synchrotron Booster (left), the Proton Synchrotron (middle), and the Super Proton Synchrotron (right).

3.2 The CMS experiment

3.2.1 Introduction

The Compact Muon Solenoid is a major detector for general purpose particle physics, which is built on the LHC. The goal of this experiment is to explore a wide scope of physics, including the search for Higgs boson, supersymmetry, extra dimensions, dark matter and so on. In 2012, the Higgs boson was discovered successfully in this experiment. This detector is one of the points where the proton-proton, proton-ion, and ion-ion beam collisions happen during the LHC run periods. It detects many

outcoming particles such as charged leptons, hadrons, and photons from the collision center and records the data for subsequent analyses.

The detector has a length of 28.7 m, and 15 m in diameter. Its total weight is 14000 tonnes, built in and around a huge solenoid magnet, made of a cylindrical coil of superconducting cable. The generated field is up to 3.8 Tesla, which is about 100000 times the magnetic field of earth. This field is confined by a steel yoke that composes the bulk of detector, resulting in the compact solenoid. Figure 3.10 shows the sectional view of the CMS detector.

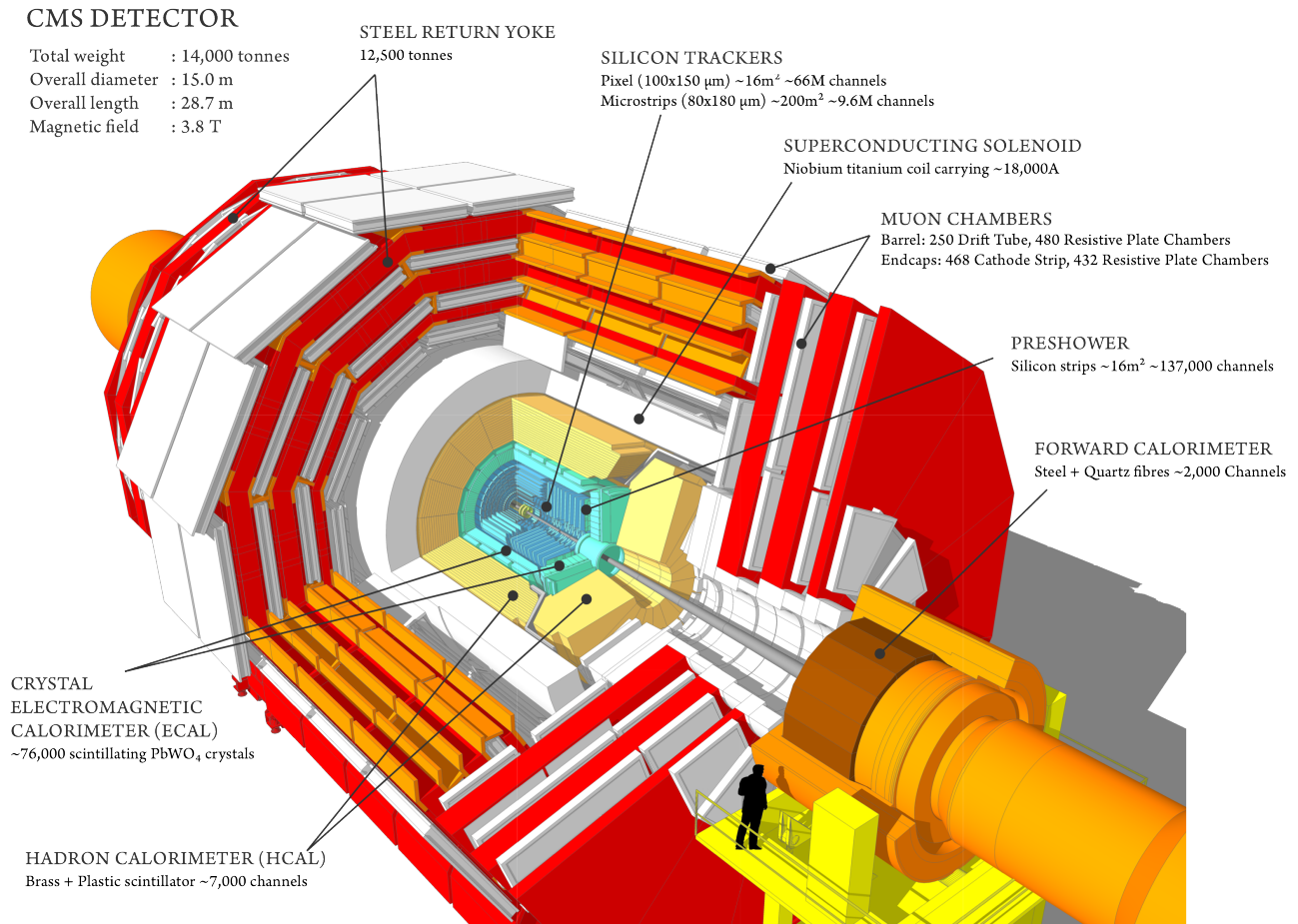


Figure 3.10: Sectional view of the CMS detector. The LHC beams travel in opposite directions along the central axis of the CMS cylinder colliding in the middle of the detector.

The requirements of the CMS detector design include:

- a high performance system to detect and measure muons;
- an excellent resolution calorimeter to detect and measure electrons and photons;
- a good quality central tracking system to provide accurate momentum measurements;
- a hermetic calorimeter to prevent hadrons from escaping.

Therefore, the huge magnet is an essential condition, which allows to precisely measure the momenta of charged particles such as muons. The higher momentum a charged particle carries, the less curvature of its bent trajectory in the magnetic field. Moreover, the large magnet also allows for a number of layers of muon detectors within the magnetic field, thus momentum can be measured inside and outside the coil. Figure 3.11 illustrates the schematic view of a slice of the CMS detector with different particles passing through the sub-detectors. The electrons are deflected by the magnetic field in the tracker, and deposit their energies in the electromagnetic calorimeter. The photons are not influenced by the magnetic field and are fully absorbed by the electromagnetic calorimeter. The trajectories of charged hadrons are bent in the tracker, and they deposit their energies through hadronic showers partially in the electromagnetic calorimeter and mainly in hadronic calorimeter. Neutral hadrons also cause hadronic showers to deposit their energies in the calorimeters, while their trajectories in the tracker are straight. The muons have quite few interactions and traverse the detector losing small fraction of energy, deflected by the strong magnetic field.

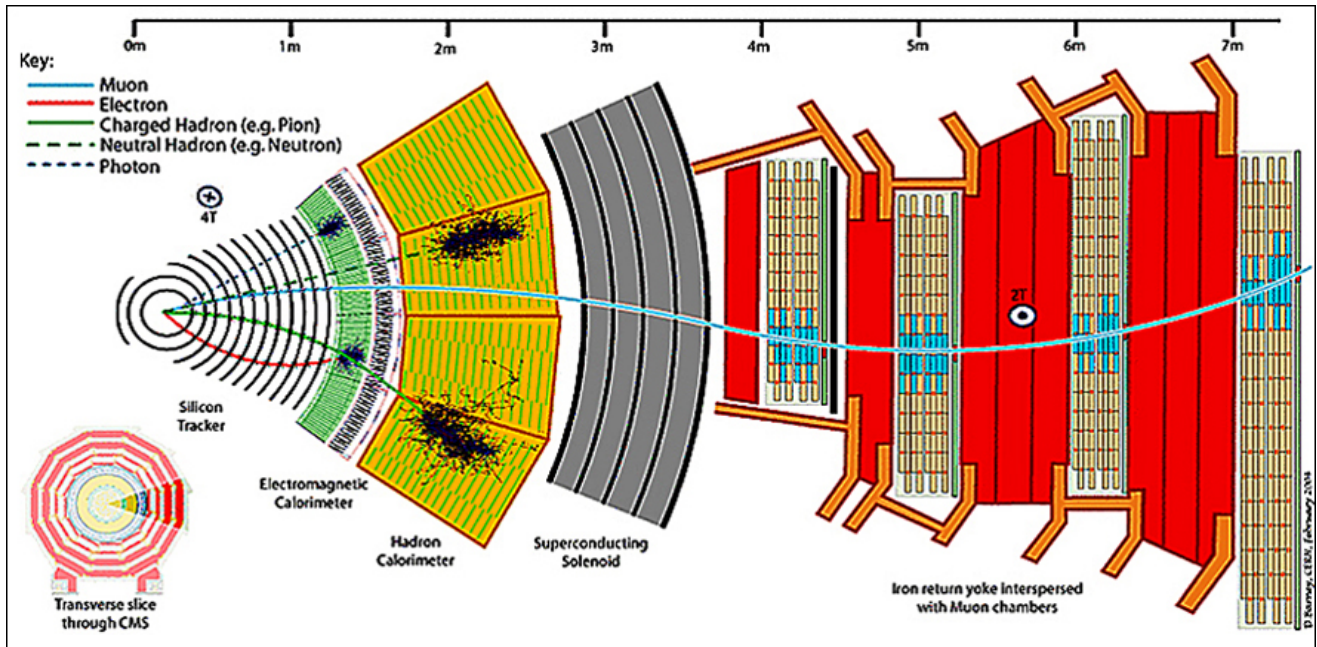


Figure 3.11: One transverse slice of the CMS detector with the particles passing through. The fragments from the left (inside) outward to right are interaction point, the silicon tracker, the electromagnetic calorimeter, the hadronic calorimeter, the superconducting solenoid magnet and the muon detectors.

At opposite ends of the detector, powerful magnets focus the two beams down to an extremely small diameter of about 16 microns at the collision point, which is aimed to increase the luminosity. The diameter of a proton is only about 2 fm, which means that the collision area is almost empty space. Therefore, the possibility of an effective collision is still small even with 115 billion protons from one beam crossing another one.

3.2.2 The Superconducting solenoid magnet

The superconducting magnet has a total length of 12.5 m, a diameter of 6 m, with a mass of 220 tonnes. It can provide a uniform magnetic field of 3.8 Tesla inside the coil, using a nominal electric current of 18500 amps with the stored energy of 2.6 GJ. The superconductivity is realized through chilling the coils, which are made of NbTi, to almost 0 K in temperature. Figure 3.12 provides an artistic view of the superconducting solenoid. It is called cryostat system. The yoke with a weight of 10000 tonnes supports the electric flux. It is composed of five wheels and two endcaps with three disks.

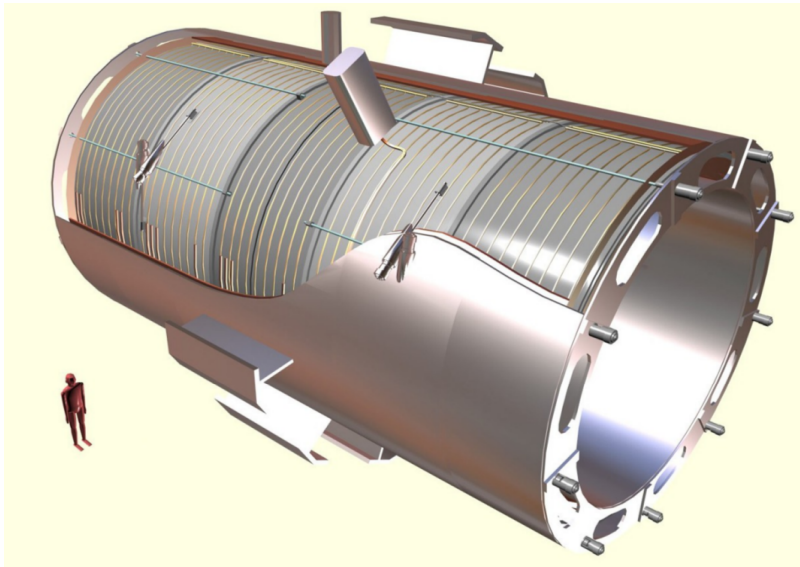


Figure 3.12: General artistic view of the superconducting magnet.

The strong magnetic field can bend the paths of charged particles emitted from the collisions. Their curved trajectories also depend on their charges and momenta.

3.2.3 The Tracker

The CMS tracking system has been designed to provide a precise measurement of the trajectories of charged particles emerging from the collisions, as well as the reconstruction of vertices. The CMS tracker occupies a cylindrical volume 5.8 m of length and 2.5 m of diameter, with its axis closely aligned to the LHC beam line, and it is immersed in the co-axial magnetic field of 3.8 Tesla. The schematic image of the CMS tracker is shown in figure 3.13. It comprises a large silicon strip tracker with a small silicon pixel tracker in it. In the barrel region, the pixel tracker comprises three co-axial layers of radii between 4.4 cm and 10.2 cm, and the strip tracker consists of ten co-axial layers extending outwards to a radius of 110 cm. The endcap regions (two sides of the barrel region) are also equipped of the subdetectors, including two disks of the pixel tracker, and three small plus nine large disks of the strip tracker. They extend the acceptance of the tracker up to a pseudorapidity

$$|\eta| < 2.5.$$

The pixel detector provides three dimensional position measurements of the hits from the interaction of charged particles with its sensors. The resolution of hit position is about $10 \mu\text{m}$ in the transverse coordinate and $20 - 40 \mu\text{m}$ in the longitudinal direction, and for the endcap wheels, the azimuthal plane is supplied by the position of sensor panels. There are in total 1440 modules containing 66 million pixels, and the covered area is 1 m^2 . Figure 3.14 shows the sketch of the pixel detector with its components.

The strip tracker contains four subsystems, including the tracker inner barrel (TIB) with disks (TID), tracker outer barrel (TOB), and tracker endcaps (TEC). The TIB and TID cover the transverse radius $r < 55 \text{ cm}$, and $|z| < 118 \text{ cm}$, with four layers and supplemented by three disks at each end. The resolution is $13 - 38 \mu\text{m}$. The TOB occupies $r > 55 \text{ cm}$, and $|z| < 118 \text{ cm}$. It is composed of six layers for the position measurements with a resolution about $18 - 47 \mu\text{m}$. The TEC lies on the region of $124 < |z| < 282 \text{ cm}$, each of which consists of nine disks. Each disk includes up to seven concentric rings of silicon strip modules, and the resolution is similar to that of TOB. In total, there are 15148 silicon modules in the strip tracker, which covers 198 m^2 area with 9.3 million strips. Figure 3.15 presents photographs of the silicon strip tracker and the TEC. More details can be found in [46].

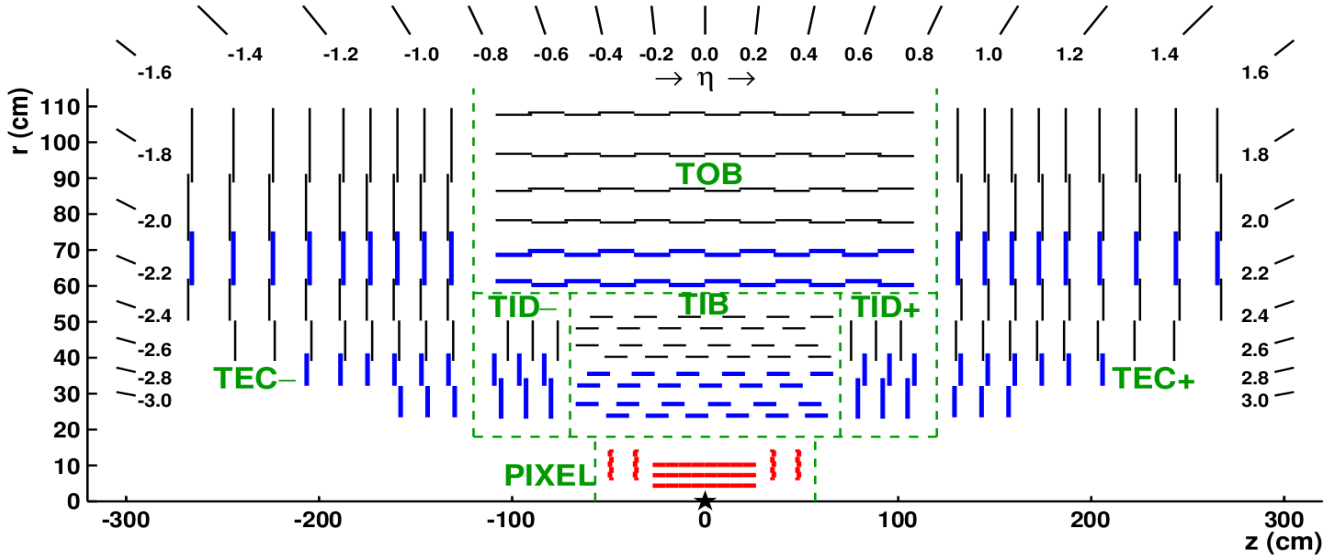


Figure 3.13: Schematic cross section through the CMS tracker in the r - z plane. The tracker is symmetric around the horizontal line $z = 0$. The star is the proton-proton collision center. The thin black lines show the two sensitive areas provided by strip tracker modules, while the thick blue lines stand for double modules permitting reconstruction of hit positions in three dimension. The red lines present the pixel modules that provide three dimensional hits. Within each layer, each modules are shifted slightly towards its neighbouring modules to have overlaps, in order to avoid gaps in the acceptance. The figure is taken from [46].

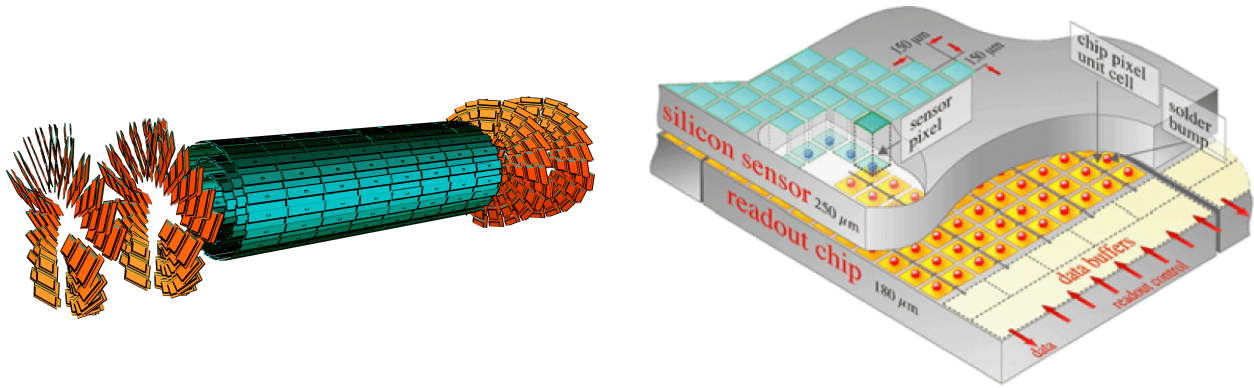


Figure 3.14: The schematic view of CMS silicon pixel detector (left) [47] and the pixel elements (right).

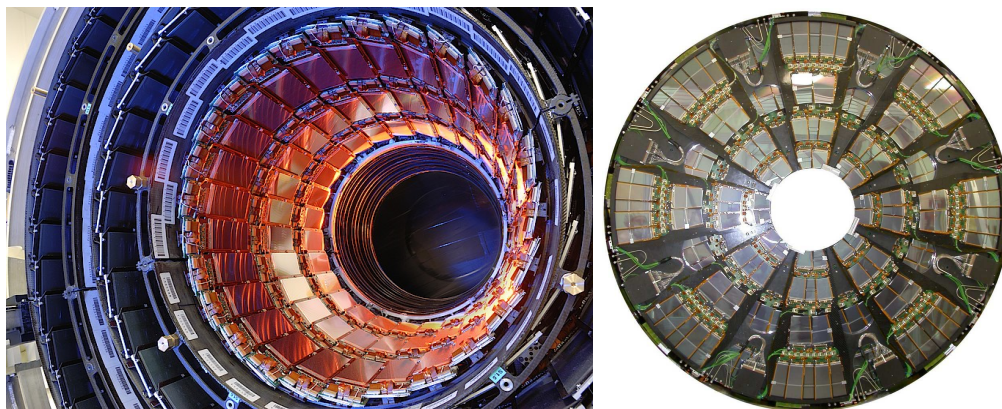


Figure 3.15: Photographs of silicon strip of tracker (left) and the TEC disk (right) [48].

3.2.4 The Calorimeters

The calorimeters are used to measure the energies of the particles, including the electrons, photons, and hadrons. Unlike the tracker system, the calorimeters can deal with both charged and neutral particles. They induce the particle showers inside the calorimeters. The total energy of the showers corresponds to the energy of the incoming particle. Muons lose only small fraction energy through ionization process when they traverse the calorimeters, while neutrinos penetrate the whole detector typically without interactions.

The calorimetry system is composed of electromagnetic calorimeter (ECAL) and hadronic calorimeter (HCAL). The former is designed to measure the energies of the particles participating in electromagnetic interactions (electrons, positrons and photons), and the latter is developed to collect the energies of the hadrons which are subject to strong interactions. Because the radiation lengths related to the electromagnetic interactions are much shorter ($X_0 = 0.89$ cm in ECAL crystals, with a total thickness of about $25X_0$) than the interaction lengths of hadrons ($\lambda_I = 16.2$ cm in HCAL components, with a total thickness of about $7-15\lambda_I$), placing the ECAL in front of the HCAL allows to obtain a good separation of electromagnetic showers from hadronic shower and optimized energy

resolutions.

- The electromagnetic calorimeter

The ECAL is made of 61200 lead tungstate (PbWO_4) crystals in the barrel region and 7324 crystals in the two endcaps. There is a preshower detector located in the front of each endcap. If one excludes the beam pipe, the ECAL is hermetic. Figure 3.16 shows the layout of CMS electromagnetic calorimeter. The barrel region of ECAL (EB) stands for the pseudorapidity range $|\eta| < 1.479$, and covers the whole azimuthal (ϕ) directions. The total crystal volume in the barrel region is 8.14 m^3 with a weight of 67.4 tonnes, and a front radius of 129 cm. The cross section of each crystal is about 0.0174×0.0174 in $\eta - \phi$ plane corresponding to $22 \times 22 \text{ mm}^2$ at the front panel and $26 \times 26 \text{ mm}^2$ at the rear area. Each crystal has a length of 230 mm. They are placed with their axes forming a tiny angle (3°) related to the vector from the nominal interaction vertex in both η and ϕ directions, in order to avoid particles traveling along the gap between two crystals.

The endcaps (EE) of ECAL ranges in the pseudorapidity as $1.479 < |\eta| < 3.0$. Their envelope are at a longitudinal distance of 315.4 cm from the collision point. Once the magnetic field is fully switched on, the endcap is expected to move towards to the nominal interaction position by 1.6 cm. Each endcap contains 7324 crystals that occupy totally 2.9 m^3 with a weight of 2.4 tonnes. There are two halves (dees), each of which holds 3662 crystals. They are arranged in 5×5 matrix units to form many supercrystals, and are directed to a position that is 130 cm beyond the nominal interaction point, rotated with small angles (about $2 - 8^\circ$) to prevent the emitted particles from passing along any crack between two crystals. Each crystal has a front cross section of $28.62 \times 28.62 \text{ mm}^2$, and a rear cross section of $30 \times 30 \text{ mm}^2$, and a length of 22 cm.

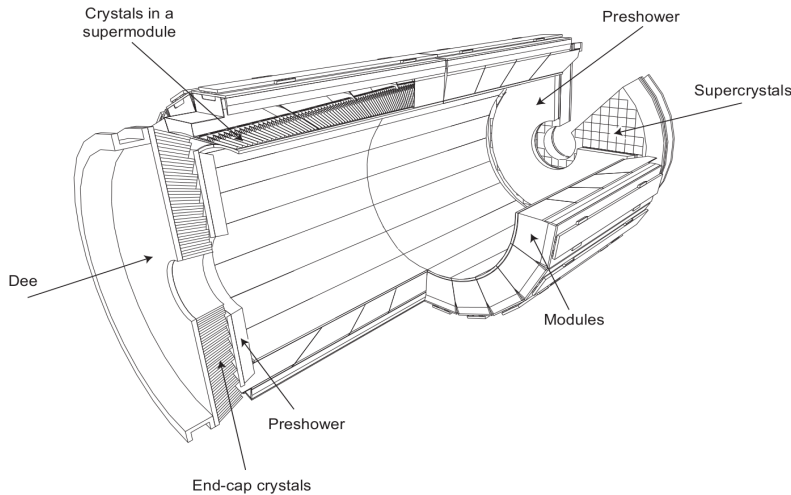


Figure 3.16: Layout of the CMS electromagnetic calorimeter showing the arrangement of crystal modules, supermodules and endcaps, with the preshower in front.

The compact calorimeter benefits from its fine granularity, which results from the PbWO_4 crystals with high density (8.28 g/cm^3), short radiation length (0.89 cm), and small Molière radius (2.2 cm).

The scintillation decay time of the crystals are very close to the bunch spacing at LHC, emitting 80% of light in 25 ns. In order to avoid cracks between each two crystals, the lead tungstate crystals are polished like a truncated pyramid, which are connected with photodetectors. These detectors have fast response and radiation tolerance in the strong magnetic field. They are made of avalanche photodiodes (APD) in the barrel and vacuum phototriodes (VPT) in the endcaps. Each APD has an active face of $5 \times 5 \text{ mm}^2$, with a signal amplification of 50 times at room temperature. The responsible area of a VPT is 280 mm^2 , with a nominal gain of 10.2 times and an efficiency of 90% under zero magnetic field. Figure 3.17 shows photographs of an APD and VPT.

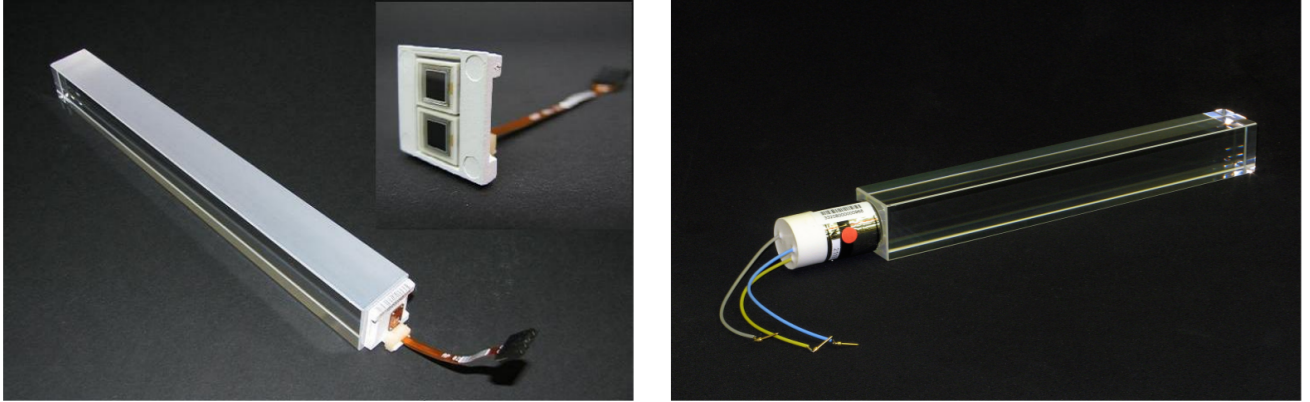


Figure 3.17: PbWO_4 crystals with photodetectors. Left: A barrel crystal with the upper area depolished and the APD capsule. In the insert, a capsule with the two APDs. Right: An endcap crystal and VPT. The figures are taken from [43].

The preshower is designed to provide the separation of photon and π^0 hadron. in the endcaps ($1.653 < |\eta| < 2.6$), and also to improve the spatial resolution for electrons and photons. It is a sampling calorimeter with two layers yielding a thickness of 20 cm in total. The incoming electrons or photons induce the consequent electromagnetic showers through the lead radiators, and then deposit their energies in the subsequent silicon strip sensors. The lead panels of the preshower is divided into two dees which are respectively placed in each side of beam pipes. Each silicon sensor has an active plane of $61 \times 61 \text{ mm}^2$ with 32 strips, and having a nominal thickness of 320 microns.

The energy resolution of ECAL is parametrized by quadratic sum of three terms:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{A}{\sqrt{E}}\right)^2 + \left(\frac{B}{E}\right)^2 + C^2 \quad (3.7)$$

In this formula, A stands for the stochastic term, which corresponds to fluctuations in the lateral shower constraint, photostatistics, and the energy deposition fluctuations in the preshower absorber in the silicon detector. The second term B is the noise term that is dominated by the noises from electronic, digitization, and pileup. The last term C is a constant term, taking into account several miscellaneous effects such as non-uniformity of the longitudinal light collection, intercalibration errors, and the leakage of the energy from the rear face of the crystal. According to the results from beam

tests [43], the resolution was measured to be:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{12\%}{E}\right)^2 + (0.30\%)^2 \quad (3.8)$$

where E is the incoming particle energy in GeV.

- The hadronic calorimeter

The hadronic calorimeter has been developed to measure the energies of hadronic jets, and indirectly to neutrinos and exotic particles that lead to large missing transverse energy. It is located outward from ECAL, covering the spatial region of $|\eta| < 5.2$. As is illustrated in figure 3.18, it consists of four fragments including hadronic barrel (HB), endcap (HE), outer (HO) and forward (HF) calorimeters.

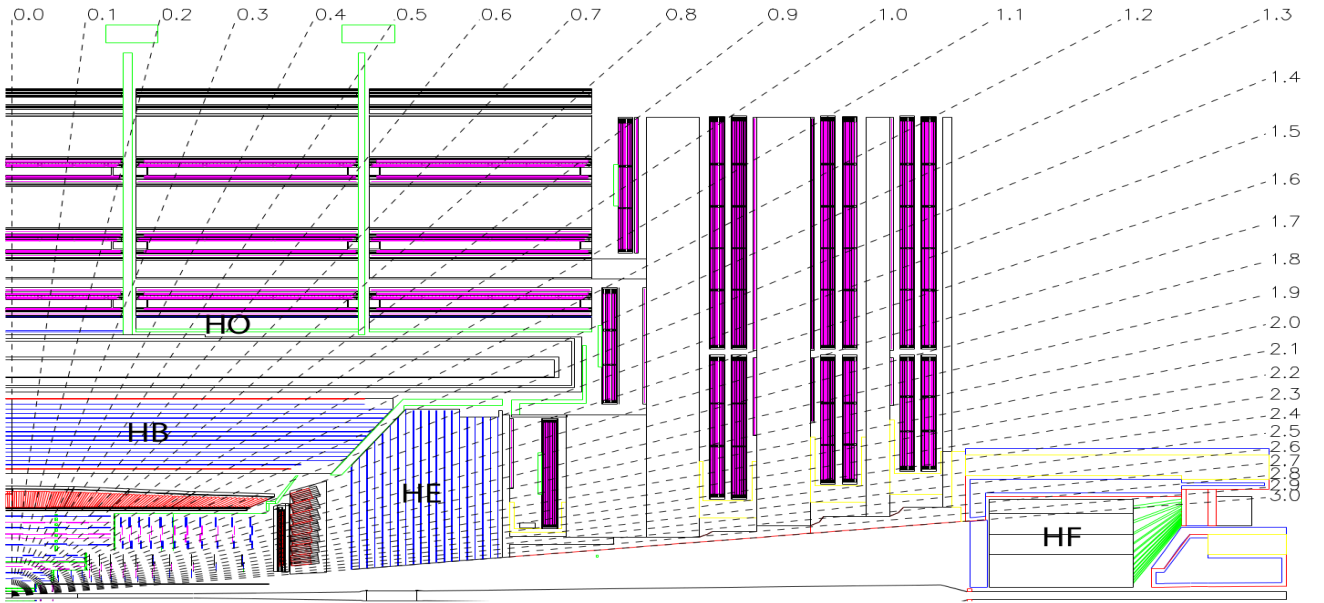


Figure 3.18: Longitudinal view of the CMS detector showing the locations of the hadron barrel (HB), endcap (EE), outer (HO) and forward (HF) calorimeters. The figure is taken from [43].

The HB is a sampling calorimeter within the pseudorapidity of $|\eta| < 1.3$, placed between the outer edge of the ECAL (1.77 m from collision center) and the inner frame of the magnet coil (2.95 m from interaction point), with a length of 9 m and radius of $178 < r < 288$ cm. It contains two half-cylinders, one inserted behind the barrel cryostat of the superconducting solenoid, while the other one with segments subsequently installed from the rails in the middle panel. These two sections are composed of 36 azimuthal wedges that are at the end of the flat absorber planes. The planes are aligned to be parallel with respect to the beam direction, and they are arranged in a crossed way to reduce effects of cracks in geometry. There are 14 brass plates and 2 stainless steel in different η directions to form the absorber, leading to a granularity of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$. Among them, the two steels are used to guarantee the frame solidity. Supported by the metal absorber, 16 sectors of 3.7-mm-thick plastic scintillators are inserted together with wavelength shifting fibres (WLS).

The HE spreads along the pseudorapidity range of $1.3 < |\eta| < 3$ with a fraction of 13.2% of the solid angle. It has the similar geometric manufacture. The absorber consists of 18 layers including 16 brass towers (79 mm thickness) and 2 stainless steels (10 cm thickness). The scintillators are planted into the gaps (9 mm) of the towers, and the light from them is collected by the WLS fibres. There are totally 36 sectors (wedges), each of which occupies an azimuthal angle of $\Delta\phi = 20^\circ$. In this partial calorimeter, the granularity is $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ for $|\eta| < 1.6$, and $\Delta\eta \times \Delta\phi = 0.17 \times 0.17$ for $|\eta| > 1.6$.

The HO is extended in the pseudorapidity range of $|\eta| < 1.3$, which complements the coverage of HB and HE for the hadronic showers. It contains five rings with a width of 2.54 m. Each ring is divided into 12 sectors in the azimuthal directions, and they are separated by stainless steels with 75 mm thickness holding the layers of scintillators. Each layer is placed at the radial distance of 4.07 m with a thickness of 40 mm, and the central ring has an additional layer at the radial point of 3.82 m. These layers are supported by aluminium honeycomb for the solidity. The granularity of HO is as the same as that of HB. Each sector is composed of six identical slices with plastic covers and scintillator tiles.

The HF is deployed in the spatial ranges of $3 < |\eta| < 5$, which was designed to detect the particle fluxes emitted to the forward directions. It is installed at 11.2 m away from the nominal collision point, and the radius of its cylindrical steel is 1.3 m. There are totally 18 wedges, each covers the azimuthal angle of 20° , composing the machinery in each side of the beam interaction position. The granularity is formed as $\Delta\eta \times \Delta\phi = 0.175 \times 0.175$ towers. Because the radiation dose is very high at this place, the detector body is encased by a hermetic shielding, which is made of steel (40 cm thickness), concrete (40 cm thickness) and polyethylene (5 cm). Instead of scintillators which are easily damaged by the high radiation, quartz fibres are used for the signal response. They are inserted in grooved plates of steel absorber, and emit Cherenkov photons when charged particles traverse them. Then the generated light is conveyed to photomultipliers.

Based on the compact machinery, the energy resolutions of the sub-detectors of HCAL were parametrized and measured to be [49, 50]:

$$HB + HE + HO : \left(\frac{\sigma}{E}\right)^2 = \left(\frac{1.2}{\sqrt{E}}\right)^2 + (0.095)^2 \quad (3.9)$$

$$HF : \left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8}{\sqrt{E}}\right)^2 + (0.11)^2 \quad (3.10)$$

where the E stands for the incoming particle energy of beams in GeV.

3.2.5 The Muon Detector

Muons have a mass 200 times heavier than electrons, and they are charged leptons. They can easily penetrate the iron calorimeters without interaction. Hence, dedicated sub-detector is developed and installed at the outmost space of the CMS detector, in order to fetch the signal when muons pass

through. This is the muon detector system, which contains totally 1400 muon chambers. These chambers are classified according to the different gaseous detection technologies, including 250 drift tubes (DT), 540 cathode strip chambers (CSC), and 610 resistive plate chambers. These three categories of components are arranged in large quantities, and they have different specialities, which make the whole muon detector robust and subject to low background level.

The detection system can measure the momenta of muons, and provide triggers. It also takes advantages of the high magnetic field provided by the solenoid with its flux-return yoke to realize an excellent muon momentum resolution and trigger capability. Figure 3.19 shows the positions of different components of muon detector around the tracker and calorimeters.

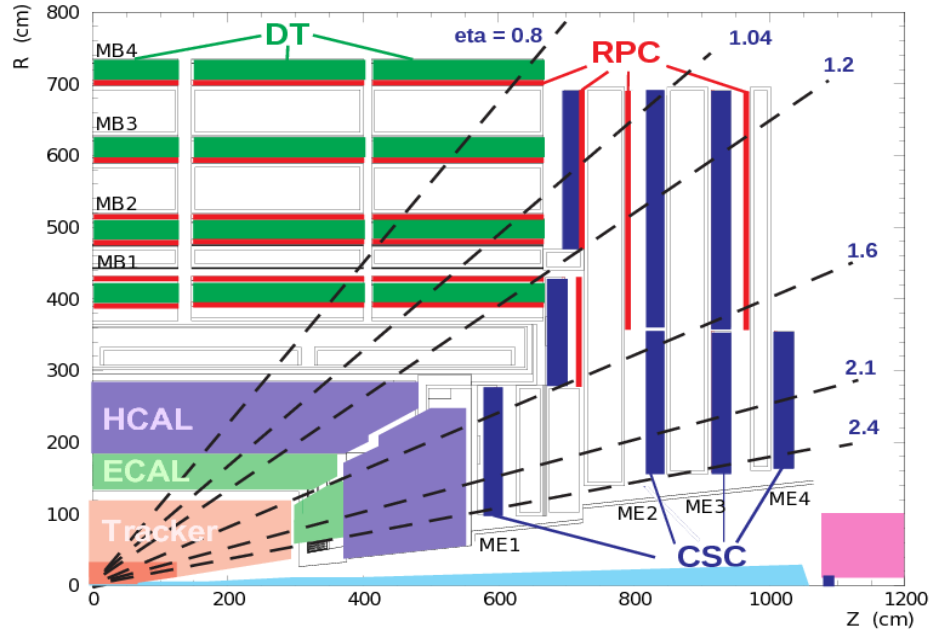


Figure 3.19: Layout of one quadrant of CMS with different sub-detectors highlighted, including the tracker, ECAL, HCAL, DTs, CSCs, and RPCs. This plot is taken from [51].

- Drift Tubes

The drift chambers are installed in the pseudorapidity range of $|\eta| < 1.2$ which is the barrel region of CMS detector. The background induced by neutron and the muon rate are low in this region. The magnetic field is uniform here and almost constrained in the iron yoke. The chambers are equipped with standard rectangular drift units, and they make up four concentric cylindrical stations around the beam line. Three of the stations with 180 drift chambers are used to measure the muon coordinate in the $r - \phi$ curved plane, and the last one with 70 drift chambers is dedicated to measure the z coordinate (beam line direction). Figure 3.20 shows the transverse cross section of the drift tubes structure.

Each drift tube unit, represented in figure 3.21, is filled with a gas mixture of 85% argon and 15% carbon dioxide. There is an anode wire and five electrode strips in the middle of the cell, which

forms a dense electric field. When a muon pass through the gas, it will release electrons from some gaseous atoms through ionization. Then these electrons are attracted and accelerated towards to the anode, and they induce more ionization processes in the gas. This results in an avalanche electric current in the wires. Benefiting from such many cells, the drift chamber gains an average resolution of $100 \mu\text{m}$ in $r - \phi$ plane.

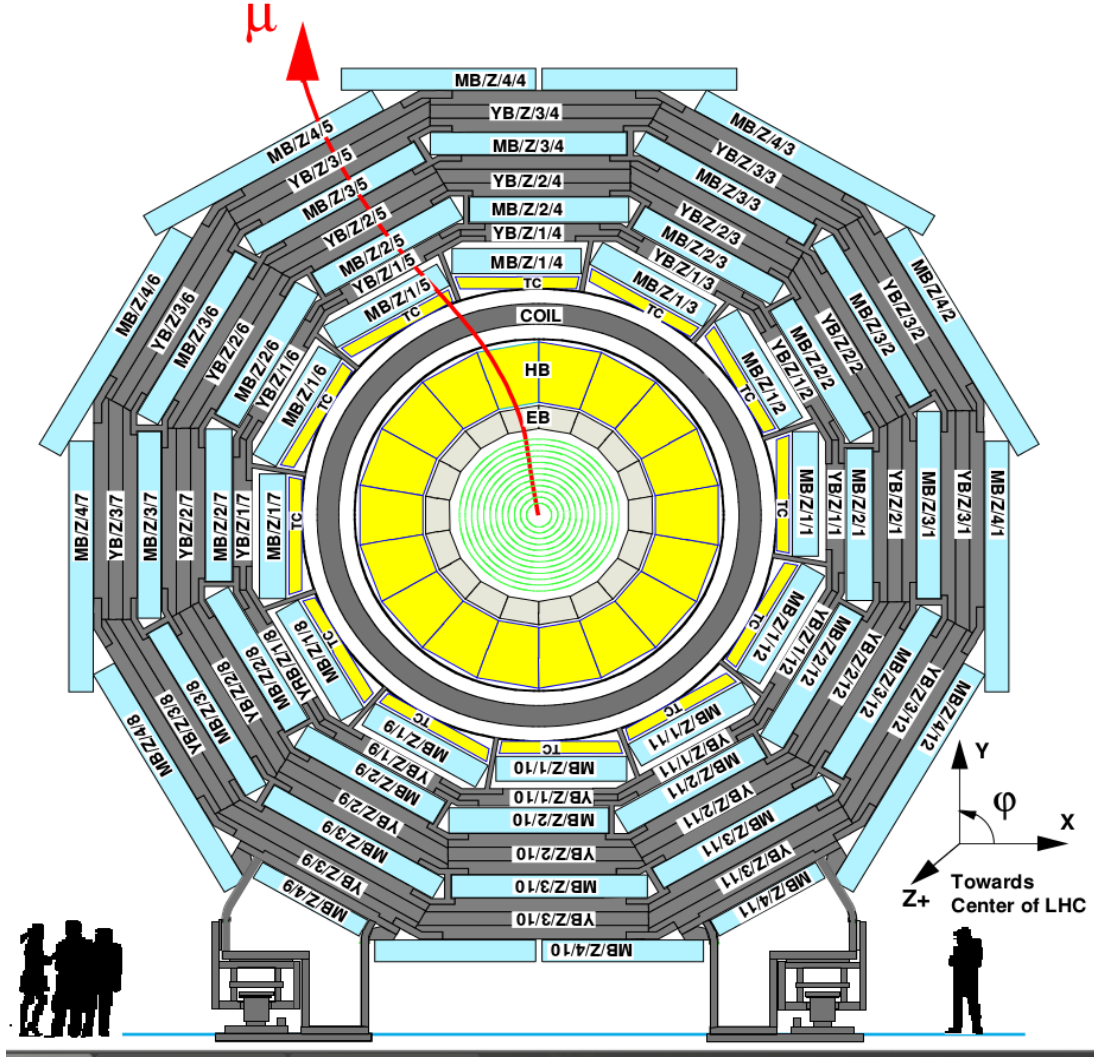


Figure 3.20: Layout of the CMS barrel muon drift tube chambers in one of the 5 wheels. In sectors 4 (top) and 10 (bottom) the MB4 chambers are cut in half to simplify the mechanical assembly and the global chamber layout. The figure is taken from [43].

- Cathode strip chamber

In the endcap region of CMS detector, the muon rates and neutron-induced background levels are high, and the magnetic field is uneven and strong. The cathode strip chambers are used in the disks here, which cover the pseudorapidity range of $0.9 < |\eta| < 2.4$. In each side of endcaps, there are

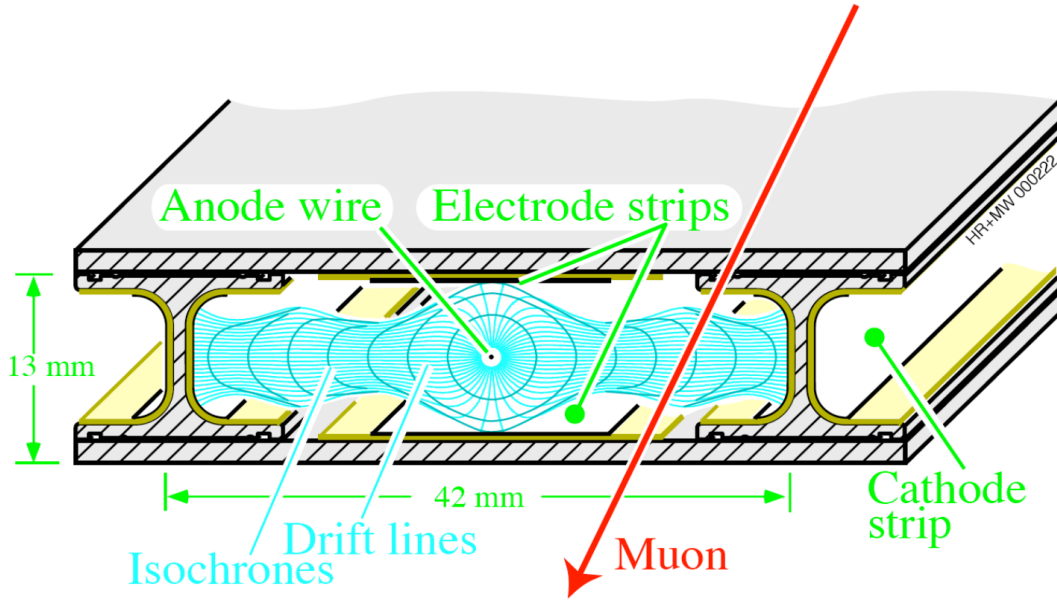


Figure 3.21: A unit of DT together with drift lines and isochrones when a muon pass through. The plates at the top and bottom of the segment are at ground potential. The voltages applied to the electrodes are +3600V for wires, +1800V for strips, and -1200V for cathodes. The figure is taken from [43].

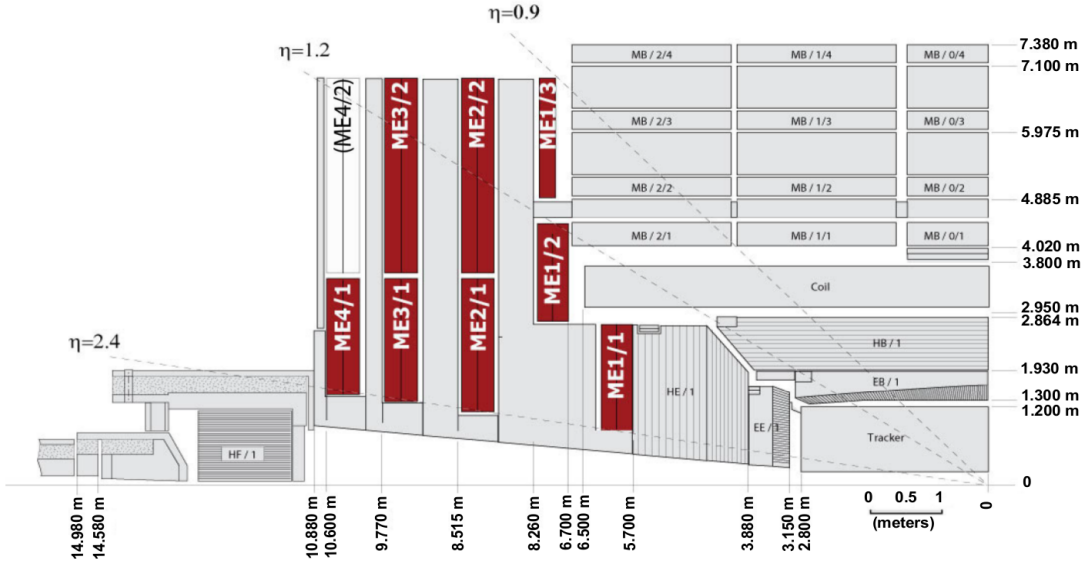


Figure 3.22: Partial view of the CMS detector. Cathode strip chambers of the Endcap Muon system are highlighted (ME_X/Y). The figure is taken from [43].

four sectors of CSCs, which are placed inside the flux planes perpendicular to the beam directions (as shown in figure 3.22). The largest chambers have a length of 3.4 m and a width of 1.5 m.

Each chamber contains 7 trapezoidal cathode plates intersected with 6 anode panels, forming 6 gaps filled with gas mixture of 40%Ar + 50%CO₂ + 10%CF₄. There are dense arrays of wires arranged on the anode panels, which are supplied with positive high voltage. The cathode plates are

made of copper strips, which are connected to the negative charged source. The anode wires and cathode strips are crossed, so that they can measure two coordinates for muons. Ionized electrons are accelerated in the electric field to the anode wires, resulting in avalanche currents. While the positive ions move away from the anode towards the cathode strips, and induce electric pulses there at right angles to the wire directions. Figure 3.23 shows a schematic view of the partial layout of the CSC and the principle of the CSC operation.

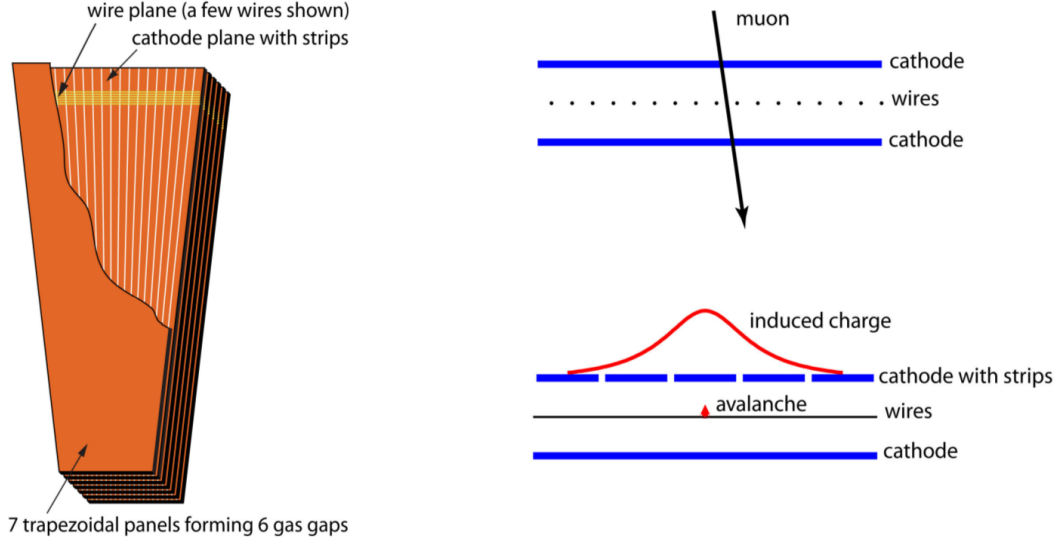


Figure 3.23: Layout of a CSC with 7 cathode strip plates and 6 anode panels forming gas volumes between them (left); An avalanche happens in a single chamber gap when a muon passes through (right). The figures are taken from [43].

Benefiting from this special structure, the CSCs have short response time, fine segmentation, and radiation resistance. They can implement a precise measurement in the $r - \phi$ curved plane and fast triggers, including robust identification for muons and efficient matching of their hits in the tracker system. Together with DTs, they can reach a momentum of resolution as 15% in barrel and 25% in endcaps.

- Resistive plate chamber

Although the DTs and CSCs have fast response time and fine position resolution, there are still large uncertainties in the background rates and the time of bunch crossings. Namely, the time resolution is comparable with the bunch spacing. In order to improve the time resolution for accurate measurements, which means to associate unambiguously the hits with the bunch crossing, the resistive plate chamber is used as complementary machinery. There are 6 layers of RPCs in the barrel region, planted parallel with the DT sectors. In the endcap region, three plates of RPCs are placed together with the CSC stations, perpendicular to the beam line. Figure 3.24 shows the structure of resistive plate chamber layers, and figure 3.25 provides a photograph of RPC plates in endcap region.

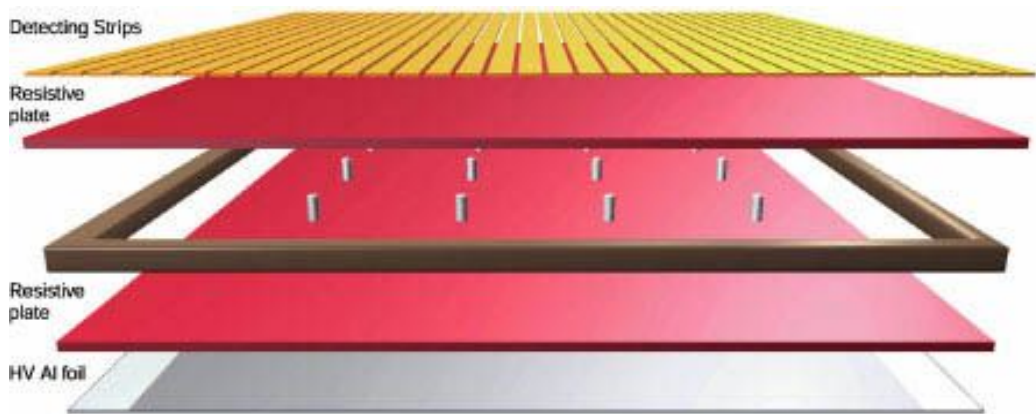


Figure 3.24: The schematic view of resistive plate chamber layers.

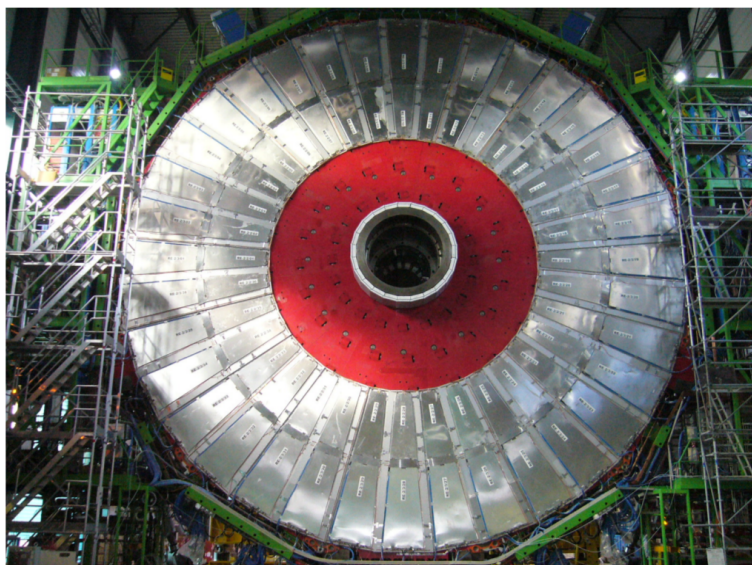


Figure 3.25: RPC station RE2 on the back side of the YE-1 yoke.

A chamber is composed of two parallel planes with a positive charged anode and a negative charged cathode, respectively. They are made of plastic materials and separated by two gaseous gaps. Similar to the principle of CSCs, the chambers also work in avalanche mode when muons come through them. The emitted electrons from ionizations are collected by external metallic strips after a certain time delay (one nanosecond), which is much shorter than the beam bunch spacing. This achieves a prompt measurement of the muon momentum. Compared with DTs and CSCs, RPCs have better time resolution while coarser spatial resolution. Thanks to these three components, it is feasible to have muon triggers at high frequency.

3.2.6 The Trigger system

During the RunI period, the instantaneous luminosity at LHC reached $7.7 \times 10^{33} \text{cm}^{-2}\text{s}^{-1}$, and the proton-proton bunch crossing frequency was 20 MHz, corresponding to 50 ns between individual

bunch collisions. While it comes to RunII time, the luminosity has been raised to $1.4 \times 10^{34} \text{cm}^{-2}\text{s}^{-1}$, and the nominal bunch cross frequency is 40 MHz (25 ns bunch spacing). The total collision rate is high, with in average 25 collisions (8 TeV) and 40 collisions (expected value at 13 TeV if with the same integrated luminosity as 8 TeV) with activity in the detector per bunch crossing. Among them, only a small fraction of collisions contains physical events of interests for the LHC priority program. In technical aspect, the detector storage could not handle the full massive data with such a high frequency. The offline reconstruction for physical analysis is available for the data frequency of several hundreds Hz. Therefore, a trigger system is necessary for filtering the event according to the physical programs and technical requirements.

The CMS detector utilizes two trigger levels. The first level is called L1 trigger, which is implemented in the hardware system. It gives a primary selection of events through the information from calorimeters and muon spectrometers, and complies with a global decision. At this step, the event frequency decreases to 100 kHz. The second level is High Level Trigger (HLT), which is based on streamlined software with computer clusters containing 16000 processors. It reduces the event output rate to 1 KHz by means of its programmable menus. With the aim to satisfy the requirement of offline data storage, the HLT can implement a further refining for the event purity, which results in an average rate of 400 Hz. Figure 3.26 shows the architecture of CMS trigger system.

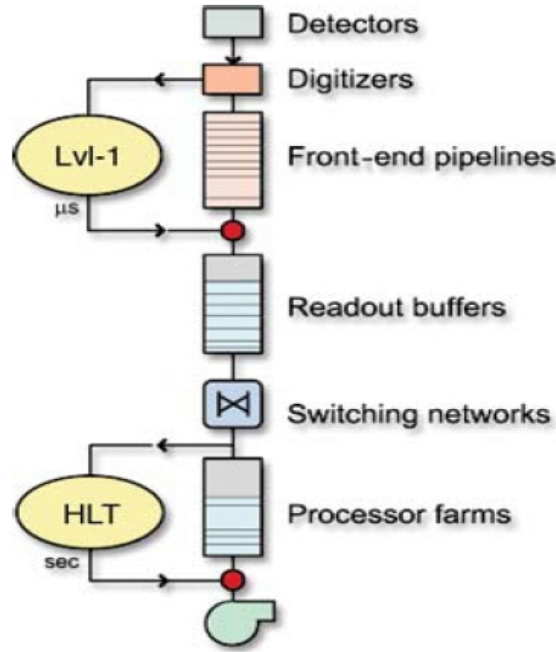


Figure 3.26: The simplified structure of CMS trigger system.

- CMS L1 trigger

The L1 trigger procedure consists of three trigger primitives, including the filtering decisions from calorimeters (both ECAL and HCAL), muon detectors, and the combinatorial information from the

previous two instruments. The data acquisition system is based on electronics that deliver the data (passed trigger filter) to subsequent trigger system. The schematic of L1 trigger is displayed in figure 3.27.

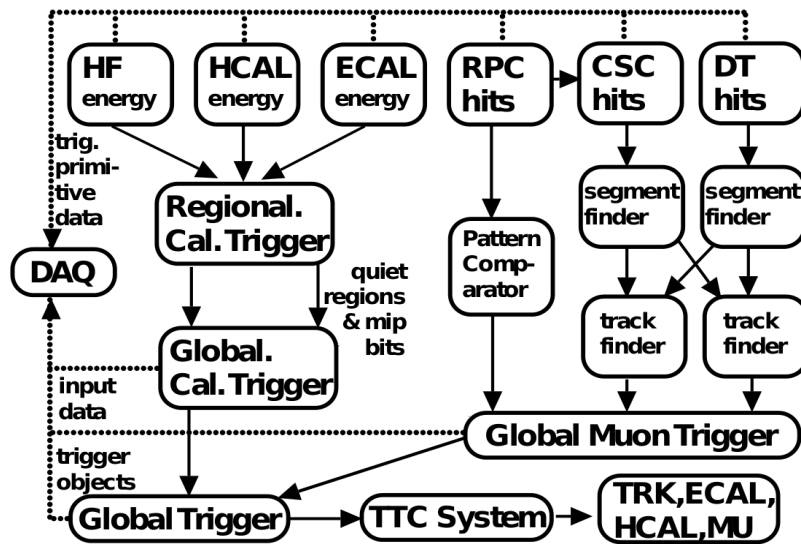


Figure 3.27: The architecture of CMS L1 trigger system.

The calorimeter trigger comprises a regional trigger and a global trigger. The former is responsible to collect the transverse momenta and processing quality flags from ECAL, HCAL and HF towers, and to provide a trigger in parallel, sending electron/photon candidates together with regional transverse energy sums measured in the crystals. The energy clusters are reconstructed using high level readout circuits called trigger primitive generators [52], which are transmitted to the regional calorimeter trigger. The global calorimeter trigger subsequently deals with the further sort of these candidates according to their transverse energy and forwards the leading four objects to the L1 global trigger. These triggers are based on dedicated electronics that uses programmable chips, so-called field programmable gate arrays (FPGA). The trigger objects supplied by the regional calorimeter trigger and computed by the global calorimeter trigger are [53]:

- 4 isolated and 4 non-isolated electrons/photons of highest energy
- 4 central, 4 forward jets, and 4 tau jets of highest energy
- scalar sums of transverse energy of all the calorimeter deposits
- missing jet transverse energy
- total missing transverse energy

The muon trigger benefits from the compact muon detector, which identifies the muon candidates using the regional track finders, and measures their momenta according to their trajectory curvatures

impacted by the magnetic field. The trigger performs a quality assignment for all the muon candidates to discard the ones that are poorly reconstructed by the muon track finders. Prior to the track finders, the local trigger track segments are formed in drift tubes and cathode strip chambers, through calculating the spatial coordinates and transverse momenta of muon candidates. The local trigger supplied by resistive plate chamber is complementary with its excellent time resolution, and it is based on the spatial and temporal coincidence of hits in different layers. The quality values of muon candidates, such as transverse momenta and angular coordinates, are assigned for their recognition. In order to synchronize and match those regional trigger information, a global muon trigger is compiled for further assignment and combination for the regional muon candidates. It picks all the muon candidates according to their reconstruction quality and sorts them by their transverse momenta, and sends the first four candidates to the global L1 trigger.

The global L1 trigger implements a trigger menu containing a maximum of 128 separated selections that are applied to the list of objects. Those selections specify different physical thresholds for special objects like muons, electrons/photons, jets, missing transverse energy and so on. Depending on the actual output rate, those selections may be prescaled for specific time tolerance [53]. In a word, this trigger is designed for the subsequent high level trigger (see next section) algorithms to process in a good condition.

- The High Level Trigger

The HLT handles data using HLT paths, which are sets of algorithmic processing procedures in specific orders. They are executed to reconstruct the objects and make selections on these objects. Each path performs as a sequence of filtering steps of increasing complexity, reconstruction refinement, and physics sophistication. Multiple HLT paths are typically configured as modular HLT menus, which executes the HLT paths in sequential way, as exemplified in figure 3.28. For each path, if an event fails to pass a filter, the rest of the path for this event will be skipped, and this path will not accept this event. As long as the event passes at least one single trigger path, the final trigger decision for this event will be positive.

However, in some cases, the input event rate is too high for some HLT paths to process efficiently. For instance, for a jet trigger path that requires low threshold of jet transverse momentum, more events will satisfy its criteria than for the path with higher threshold. In order to overcome this issue, these HLT paths are configured with “prescale”, which selects only a fraction of input events in real time. These approaches explained above accomplishes to lower the event rate from L1 trigger output. In 2012, the reconstruction and filtering for each event consumed up to 175 ms. More details on event reconstruction at HLT are introduced in 5.2.

Once event reconstruction and selection are completed at HLT, the accepted events are transferred to storage manager, which is based on software program as well. These event data are temporarily stored in the local disks, and are then delivered to CMS computing center for offline processing and permanent storage. In addition, some output events are also used for online data monitoring and calibration.

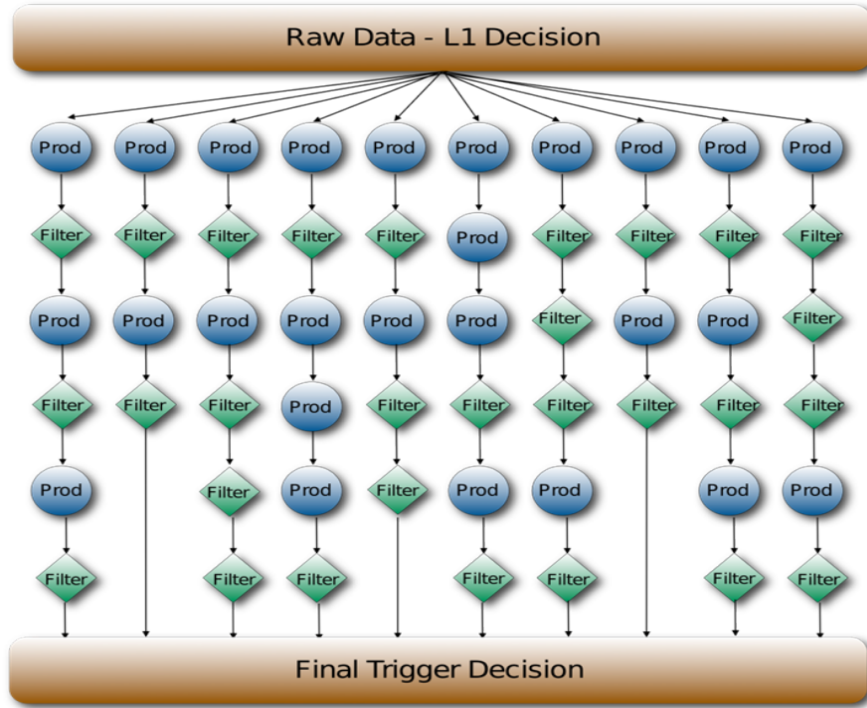


Figure 3.28: Schematic representation of a HLT menu with the HLT paths inside. The final trigger decision is the logical OR of the decisions of the single paths. The figure is taken from [54].

3.2.7 Data handling application

In order to perform the HLT and physical analysis on the recorded data in the storage, the CMS collaboration has developed a common software framework called CMSSW. It is modular using C++ as core language, and configures the plugged C++ modules using Python language. The data in the storage are configured into Event Data Model (EDM) format [55], which holds the information event by event. Each event is a C++ class container, preserving all the raw and reconstructed data corresponding to a particular collision, such as the specific information of trigger paths and event filtering, tracks, reconstructed objects, identified objects and so on. They are stored based on ROOT framework [56] for data analysis.

Besides the real data collected by the detector, we can also use Monte Carlo methodology to generate pseudo data, which are used for data calibration and comparison. In order to simulate the detector performances, we usually use the GEANT4 software [57], which is a common toolkit to simulate particles passing through matters. By taking into account the detector conditions including geometrical coverage, electromagnetic field distributions, and effective components, we can reconstruct the processes of the generated particles from the expected beam collision point encountering the detector system, and implement the detector effects such as energy shifts, energy smearing, efficiency loss and so on. All the pseudo events are simulated and then processed in the similar way as the real data. They are also stored in the ROOT files, containing the additional information of MC generated particle level, which is not in real data files.

Chapter 4

Jet production in association with Z boson

This chapter presents the measurement of the differential cross section of Z boson production associated with jets, in which case Z boson decays into muon or electron pairs, in proton-proton collisions at the LHC, respectively at a central energy of 8 TeV and 13 TeV. It includes analysis introduction, data and simulation samples, event selection, results at reconstructed level, corrections for detector effect, uncertainties, and results.

Regarding the case of 8 TeV, both $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ decay channels are measured and their cross sections are combined. The details of this analysis are introduced in [58] that has been published. The differential cross section is expressed as a function of jet multiplicity, the transverse momenta (p_T) of the Z boson and jets, the absolute rapidity (y) of the Z boson and jets, and the scalar sum (H_T) of jets transverse momenta. Moreover, the longitudinal angular correlation between Z boson and jets are also measured. The measured cross section is compared with three theoretical predictions, based on MADGRAPH, SHERPA 2, and MADGRAPH_AMC@NLO.

In the case of 13 TeV, only $Z \rightarrow \mu^+\mu^-$ decay channel is presented in details ¹, and is combined with the $Z \rightarrow e^+e^-$ decay channel at the cross section level. The details of whole analysis are interpreted in [59, 60]. The differential cross section is presented as a function of jet multiplicity, the transverse momenta of the Z boson and jets, the absolute rapidity of the Z boson and jets, and the scalar sum of jets transverse momenta. The observed results are compared with the theoretical predictions based on MADGRAPH_AMC@NLO, MADGRAPH, and fixed next-to-next-leading order computation.

4.1 Motivation

The hadronic jet production in Drell-Yan process has played a significant role in the fundamental understanding of Quantum Chromodynamics (QCD) since the 1980s. Due to its sizable cross section and subject to sizable QCD radiative corrections, the Z boson production in association with jets becomes a standard candle to provide stringent tests of perturbative QCD, and allow us to validate

¹The preliminary results of $Z \rightarrow e^+e^-$ at 13 TeV are shown in the appendix B. The details are not presented in this thesis, but will be included into Sandeep Kaur's PhD thesis (in preparation). They can be also found in [59].

and improve our understanding and prediction of jet productions. The Z boson, which can decay to two oppositely charged leptons (like $\mu^+\mu^-$ or e^+e^-) has a high reconstruction efficiency, and the background contamination in the peak region of Z boson mass is largely suppressed, resulting in a clean signature and almost background free samples. In addition, this process constitutes a significant background to several standard model processes, such as single top, $t\bar{t}$, vector boson fusion and scattering, scalar boson production and so on. The precision measurement of Z + jets process provides a better control of background rejection for those measurements, and reduces the related systematics. Furthermore, an improved understanding of jet production dynamics also influences the precision that can be reached for Higgs boson measurement, and limits on many searches beyond standard model like supersymmetry and exotic particle searches. Hence, the precision measurement of this process is central at the LHC.

In practice, we measure the differential cross section of this process as a function of several observables. The jet multiplicity observable is an important test on perturbative QCD. At high energy, many processes of interests contain high jet multiplicity that has to be predicted accurately, because high jet multiplicity is related to either higher order of radiation correction or parton shower. In order to perform that, several recent MCs combine ME calculation for many partons in the final state (multi-leg) matched with PS. The jet multiplicity measurement gives a fine overall control of this approach and limits of its validity. The modelling of the transverse momenta p_T of jets are influenced by the merging scheme between matrix element and parton shower. Moreover, the measurements of hard jet p_T are also sensitive to the parton distribution function (PDF) in a region where they are not very well known (large x). The H_T of jets is a suitable observable for supersymmetry (SUSY) searches with multiple jets production and missing energy, which may come from supersymmetric particles. The dominant background is Z + jets process in some cases, where the Z boson decays in neutrinos ($\nu\bar{\nu}$), resulting in the same final state as these searches. The neutrino pairs cannot be directly reconstructed by the detector, and they cause significant missing transverse energy. Therefore, the measured H_T (scalar sum of all the reconstructed jet p_T) together with the p_T of each jet are needed for estimating the energy and angle of the neutrinos.

The multi-leg approach, with ME computed at LO has been very successful to describe multi-jet events but has also its limitations. The lack of higher order ME could be visible (e.g. for angular correlations). This has stimulated our present analysis of the longitudinal angular correlation between Z boson and jets, namely, to measure the rapidity correlation. The rapidity of a particle is defined as $y = (1/2)\ln[(E + p_z)/(E - p_z)]$, where E represents the energy and p_z is the momentum component along the direction of the counterclockwise circulating proton beam on the LHC. In particular, the rapidity sum and rapidity difference between the Z boson and the leading jet, which has the largest transverse momentum among all the jets, reflect the longitudinal correlation between Z boson and jets. Based on the Z boson rapidity y_Z and leading jet rapidity y_{j1} , they are defined as:

$$y_{sum(Z,j_1)} = \frac{1}{2} |y_Z + y_{j1}| , \quad y_{diff(Z,j_1)} = \frac{1}{2} |y_Z - y_{j1}| , \quad (4.1)$$

The quantity $y_{sum(Z,j_1)}$ is related to the boost from the laboratory frame to the central mass frame of the Z boson and leading jet, and it mainly depends on the parton distribution functions. The $y_{diff(Z,j_1)}$ is the relative rapidity between Z boson and leading jet, which reflects the partonic differential cross section. These two variables are approximately uncorrelated, and correspond to a rotation in the phase space of y_Z and y_{j_1} . On the other hand, the y_Z and y_{j_1} are highly correlated in the laboratory frame, for a fixed invariant mass of the lepton pair from Z boson decay. The reason is that Z + jets production in most of the cases involves a valence quark carrying relatively high momentum interacting with a gluon or anti-quark which has a low momentum [61]. At leading order level, in order to reach the energy scale of Z boson mass, the initial quark comes from valence quarks ($x \simeq 0.15$) in a proton, while the initial antiquark is from sea quarks (small x). In the case of next-to-leading order, the initial quark could come from valence quarks or splitted gluons. The antiquark could be from sea quarks or splitted gluons. Among all the contributions, the quark-gluon and antiquark-gluon scattering processes are dominant compared to the quark-antiquark interaction with gluon radiations, because of high gluon density in the incident protons (small x). The relative contributions depend on the center-of-mass energy of the colliding protons.

4.2 Data and Simulation samples

The data used in the present analysis were collected by the CMS collaboration at CERN in 2012 and in 2015, with the recorded integrated luminosity of $19.6 fb^{-1}$ for 8 TeV and $2.5 fb^{-1}$ for 13 TeV. The corresponding data samples are listed in tables 4.1 and 4.2. They are restricted to the luminosity sections that are in the following JSON files. The analyses of 8 TeV and 13 TeV have been implemented by use of the CMSSW_5_3_X and CMSSW_7_6_X software releases respectively.

Besides the data sample, two lists of Monte Carlo samples have been used in these two analyses respectively, for subtracting background contamination, and detector level correction. The lists of those samples are in tables 4.3 and 4.4 for 8 TeV and 13 TeV, respectively. For the comparison to data, each Monte Carlo sample is normalized to the integrated luminosity of data samples.

For 8 TeV, the signal is modelled with five samples of Drell-Yan events, restricted to oppositely charged leptonic Z boson decay with an invariant mass above 50 GeV. The first simulated signal sample in table 4.3 contains the events with two charged leptons ($\mu^+\mu^-$, e^+e^- , $\tau^+\tau^-$) and with between zero and four partons in the final state, therefore it is named as inclusive signal sample. Due to the statistical limit of this sample, four additional exclusive samples which have been generated with Z + n partons, where $n = 1, 2, 3, 4$ respectively, are merged to the first one. Each of the four samples has been normalized to the predictions of the inclusive sample. The Z boson and partons are produced with the LO matrix element generator MADGRAPH [24] using the parton distribution function CTEQ6L1 [62]. The parton shower and hadronization effects are simulated by PYTHIA [22] with the Z2 tune [63]. The integrated cross section of the signal process is normalized to match the next-to-next-leading order (NNLO) prediction for inclusive Z boson production obtained

Sample Name	Events	Luminosity(fb^{-1})
/DoubleMu/Run2012A-22Jan2013-v1	5636274	0.88
/DoubleMuParked/Run2012B-22Jan2013-v1	29308627	4.40
/DoubleMuParked/Run2012C-22Jan2013-v1	36820243	6.95
/DoubleMuParked/Run2012D-22Jan2013-v1	38006513	7.35
/DoubleElectron/Run2012A-22Jan2013-v1	12964286	0.88
/DoubleElectron/Run2012B-22Jan2013-v1	23571931	4.40
/DoubleElectron/Run2012C-22Jan2013-v1	33843769	6.99
/DoubleElectron/Run2012D-22Jan2013-v1	34526899	7.34
/MuEG/Run2012A-22Jan2013-v1	1939498	0.88
/MuEG/Run2012B-22Jan2013-v1	13502523	4.41
/MuEG/Run2012C-22Jan2013-v1	20027780	7.03
/MuEG/Run2012D-22Jan2013-v1	20830686	7.34
JSON		
Cert_190456-208686_8TeV_22Jan2013ReReco_Collisions12_JSON.txt		

Table 4.1: Datasets and the corresponding number of recorded events with the integrated luminosity, and JSON file used for 8 TeV. The bunch spacing is 50 ns.

Sample Name	Events	Luminosity(pb^{-1})
/DoubleMuon/Run2015D-16Dec2015-v1	39702224	2250.29
JSON		
Cert_246908-260627_13TeV_PromptReco_Collisions15_25ns_JSON.txt		

Table 4.2: Datasets and the corresponding number of recorded events with the integrated luminosity, and JSON file used for 13 TeV. The bunch spacing is 25 ns.

with FEWZ [64], which is 3531.8 pb with NLO electroweak corrections for $Z \rightarrow l^+l^-$ summed on the three flavors and with the invariant mass of lepton pair larger than 50 GeV.

Background contamination is dominated by $t\bar{t}$ production with a decay chain ending in at least two oppositely charged leptons and jets (TTJets), as well as double electroweak boson production (WW, ZZ, WZ). They are modelled with MADGRAPH for matrix element and PYTHIA for parton shower. The cross section of $t\bar{t}$ is normalized to the next-to-next-to-leading order and logarithmic accuracy (NNLO + NNLL) computation from [65]. The double boson background samples are normalized to the next-to-next-leading order (NNLO) accuracy predicted by MCFM [66] generator using the CTEQ PDF set. Since the background coming from the Drell-Yan process with Z boson decaying into $\tau\bar{\tau}$ has been included in the signal sample, it needs to be taken into account in this analysis. Except the sample of W+jets (WJets), all the background samples are implemented with tau lepton production and decay using TAUOLA library [67], including single top productions (T_s, T_t, T_tW, Tbar_s, Tbar_t, Tbar_tW).

Regarding 13 TeV, the sample is generated by MADGRAPH_AMC@NLO using the FxFx merging scheme [40, 27], which is an inclusive Drell-Yan $Z + n$ partons ($n = 0,1,2,3$) sample. The NNPDF3.0 [68] is employed. The matrix element (ME) is computed with $Z + 0,1,2$ partons at next-

leading order (NLO) level, i.e. with leading order (LO) accuracy for $Z + 3$ partons. The parton shower and hadronization effects are performed with PYTHIA8 [23] using the CUETP8M1 tune [69].

Background Monte Carlo samples of $t\bar{t}$ and single top processes are produced using POWHEG [70, 71, 72] or MADGRAPH_AMC@NLO interfaced with PYTHIA8. The MC samples containing double vector boson electroweak production are generated either with MADGRAPH_AMC@NLO (for WZ production) or POWHEG (for WW production), both interfaced with PYTHIA8, or generated with PYTHIA 8 alone (for ZZ production). The sample corresponding to W boson production in association with jets is simulated using the same generator and merging scheme with the same parton multiplicity and accuracy in ME as the signal sample.

Sample Name	Events	$\sigma(\text{pb}) \times \text{BR}$
Signal		
/DYJetsToLL_M-50_TZ2Star_8TeV-mad-tarball	30459503	3531.8
/DY1JetsToLL_M-50_TZ2Star_8TeV-mad-tarball	24045248	666.30
/DY2JetsToLL_M-50_TZ2Star_8TeV-mad-tarball	21852156	214.97
/DY3JetsToLL_M-50_TZ2Star_8TeV-mad-tarball	11015445	60.69
/DY4JetsToLL_M-50_TZ2Star_8TeV-mad-tarball	6402827	27.36
Background		
/WJetsToLNu_TZ2Star_8TeV-mad-tarball	76102995	36864
/TTJets_MassiveBinDECAY_TZ2star_8TeV-mad-tauola	6923652	245
/T_s-channel_TZ2star_8TeV-powheg-tauola	259961	3.79
/T_t-channel_TZ2star_8TeV-powheg-tauola	3758227	56.4
/T_tW-channel_TZ2star_8TeV-powheg-tauola	497658	11.1
/Tbar_s-channel_TZ2star_8TeV-powheg-tauola	139974	1.76
/Tbar_t-channel_TZ2star_8TeV-powheg-tauola	1903681	30.7
/Tbar_tW-channel_TZ2star_8TeV-powheg-tauola	493460	11.1
/ZZJetsTo2L2Q_TZ2star_8TeV-mad-tauola	1936727	17.654×0.14
/ZZJetsTo2L2Nu_TZ2star_8TeV-mad-tauola	944911	17.654×0.04
/ZZJetsTo4L_TZ2star_8TeV-mad-tauola	4807893	17.654×0.01
/WZJetsTo2L2Q_TZ2star_8TeV-mad-tauola	3215990	33.21×0.07
/WZJetsTo3LNu_TZ2_8TeV-mad-tauola	1995334	33.21×0.03
/WWJetsTo2L2Nu_TZ2star_8TeV-mad-tauola	1933235	54.838×0.11

Table 4.3: Monte Carlo samples employed in the analysis for 8 TeV, followed by their number of events and cross section with branching ratios. In terms of the names, mad stands for madgraph, and TZ2 is for TuneZ2.

4.3 Event reconstruction and selection

This section explains how events are reconstructed and selected in this analysis. It will firstly list the trigger filter used to save the data. Secondly, one will explain the reconstruction approach and requirement on lepton candidates, including the constraint of Z boson reconstruction. Thirdly,

Sample Name	Events	$\sigma(\text{pb})$
Signal		
/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	28500769	6025.2
Background		
/TT_TuneCUETP8M1_13TeV-powheg-pythia8	87995547	831.76
/ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	1e+06	35.6
/ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	749400	35.6
/ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	998400	10.32
/ST_t-channel_5f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	2995200	216.99
/WJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	24156124	60290
/WWTo2L2Nu_13TeV-powheg	1979988	12.21
/WZJets_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	1937499	4.4
/ZZ_TuneCUETP8M1_13TeV-pythia8	985600	15.4

Table 4.4: Monte Carlo samples employed in the analysis for 13 TeV, followed by their number of event and cross section. For the name, 4f stands for 4 flavor quarks in the initial partons which does not include heavy flavor quarks (bottom/top), and 5f represents the similar scenario taking into account bottom quarks.

the reconstruction methodology and selection of reconstructed jets will be introduced. Finally, the treatment on pileup contamination will be discussed.

4.3.1 Trigger

The data samples in these analyses were collected with specific triggers, which are called Double Muon and Double Electron. These triggers are unrescaled².

The reconstructions of muon and electron candidates at trigger level are similar to the offline reconstructions, which are introduced in section 4.3.4 and section 4.3.5. For 8 TeV analysis, the following trigger is used to select the events that contain at least two muons:

HLT_Mu17_Mu8_v*

L1 trigger seed: L1_DoubleMu_11_4 or L1_DoubleMu_12_5

which requires the first muon candidate carrying a p_T larger than 17 GeV while a second one with a p_T larger than 8 GeV at high level trigger (HLT) level. The following trigger is applied for extracting the events including at least two electrons:

HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*

L1 trigger seed: L1_DoubleEG_15_10

²During a single run of LHC, the instantaneous luminosity may be quite high. Some triggers can be fired at too high frequency to process the sequence of incoming events in time. Therefore, according to the timing feature, prescale could need to be applied to these triggers to decrease the rate. For instance, if the prescale value is 100, the trigger will be kept randomly by only one out of one hundred incoming triggers.

which requires the first electron candidate with a p_T higher than 17 GeV together with a calorimeter identification (CaloIdT), a very loose calorimeter isolation (CaloIsoVL), a very loose tracker identification and isolation (TrkIdVL). Meanwhile, it picks the second electron passing the same sequences except for the lower p_T threshold of 8 GeV.

In 13 TeV analysis, the instantaneous luminosity being higher, the triggers had to be more restrictive to keep a reasonable trigger rate. This document only presents the $Z \rightarrow \mu^+ \mu^-$ channel. The unscaled triggers used for selecting two muons are:

```
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ
L1 trigger seed: L1_DoubleMu_11_4 or L1_DoubleMu_12_5
```

As similar as the triggers used in 8 TeV, this two triggers implement the requirements including the p_T thresholds (17 GeV and 8 GeV) of two muon candidates respectively and the tracker isolations (very very loose). Moreover, DZ stands for the distance along the beam axis between the extrapolation of the two muons' tracks. The trigger paths implement an upper threshold of 0.2 cm on DZ for background rejection. The events that pass at least one of this two trigger paths will be selected. Note that the only difference between the two trigger paths is the selection method for the second muon candidate, i.e. "Mu8" and "TkMu8". The former starts the probe in the muon detector. Once a muon candidate in the muon detector is found, the trigger system extrapolates the trajectory from muon detector to the tracker system to search for a matched track. While, the latter begins in the tracker by reconstructing a track, and extrapolates it to the muon detector to see if there is any matched track. If found, this muon candidate is selected and the search is stopped. Otherwise, it repeats the same procedures for next reconstructed track in the tracker system. In comparison, the former takes much less time than the latter.

Regarding the trigger selection and also other steps of selection for leptons, there is non-zero probability to miss a lepton which would have been poorly reconstructed or would have ended up into some dead part of the detector. In order to correct for such effects, the different efficiencies must be computed and implemented in the simulation. For measuring the efficiency of these double muon triggers, CMS collaboration deployed tag and probe method [73]. One firstly computes the efficiency of a single muon trigger (reference trigger) using the tag and probe method, and then evaluates the efficiency of the double muon trigger on di-muon events triggered by the reference trigger, which is done by fitting the invariant mass of the two muons around Z boson mass peak region. The final efficiency is the multiplication of both. The technical procedure for electron triggers is similar.

Figure 4.1 and 4.2 show the efficiencies of the muon trigger paths at 8 TeV and 13 TeV, respectively, by use of tag and probe approach. The efficiencies are measured for both data and MC and the ratios of data to MC. According to the results, one can see that the trigger efficiency on data and

MC samples are different. This is due to the variation of detector experimental condition. The MC always overestimates the efficiency. Hence, an additional correction factor, which is the ratio of trigger efficiencies of data and MC, needs to be applied to MC yields for the consistence. Typically, this factor varies as a function of η , and is close to 1. On the other hand, for double electron triggers, data and simulation are compatible.

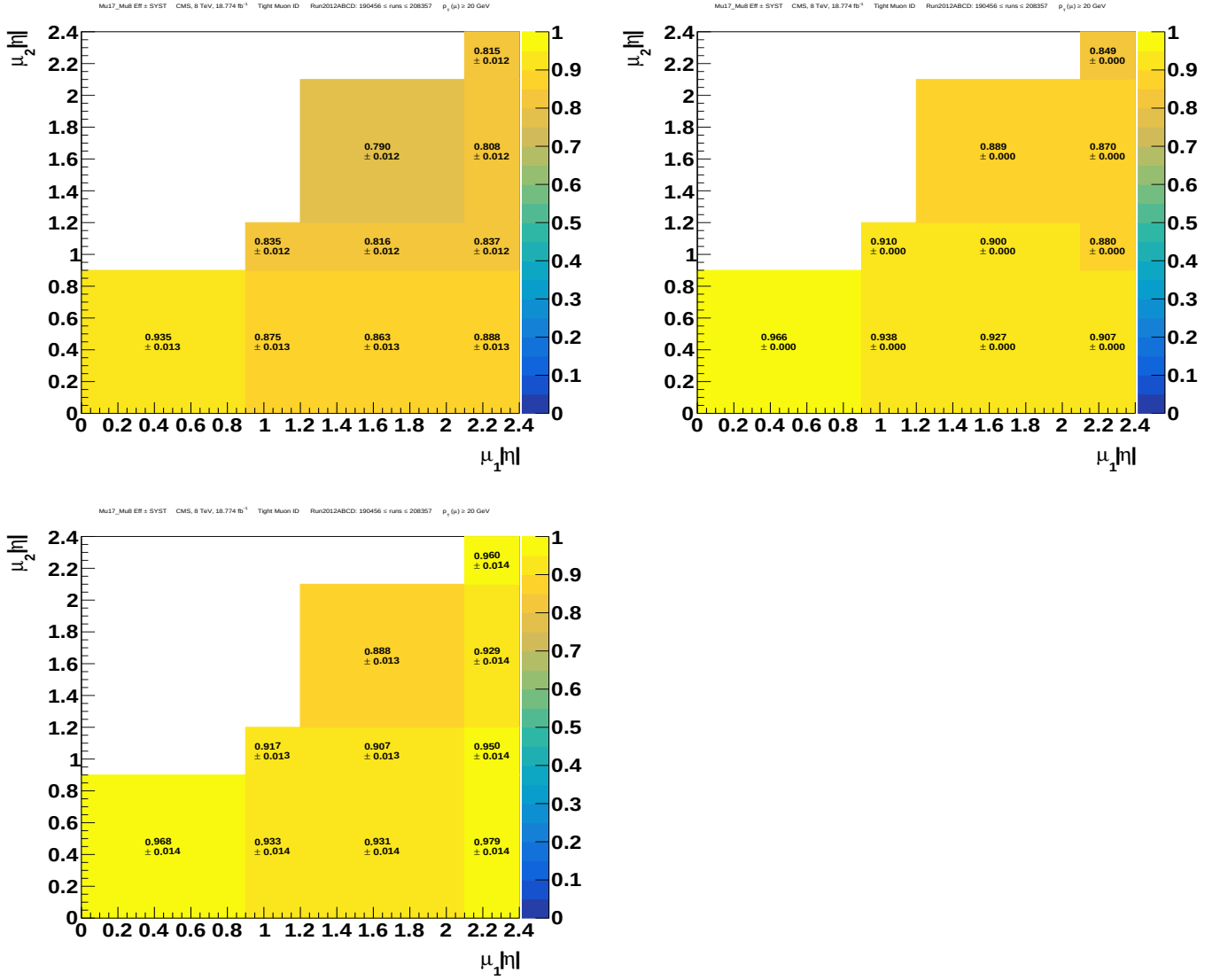


Figure 4.1: The double muon trigger efficiencies measured with tag and probe approach for the data collected in 2012 (top left), and simulated DY + jets sample (top right), and the ratio of data to MC (bottom), as a function of the pseudorapidity of the leading and subleading muon candidates, with $p_T > 20$ GeV. Note that the efficiency distributions are assumed to be symmetric. The results are taken from [74].

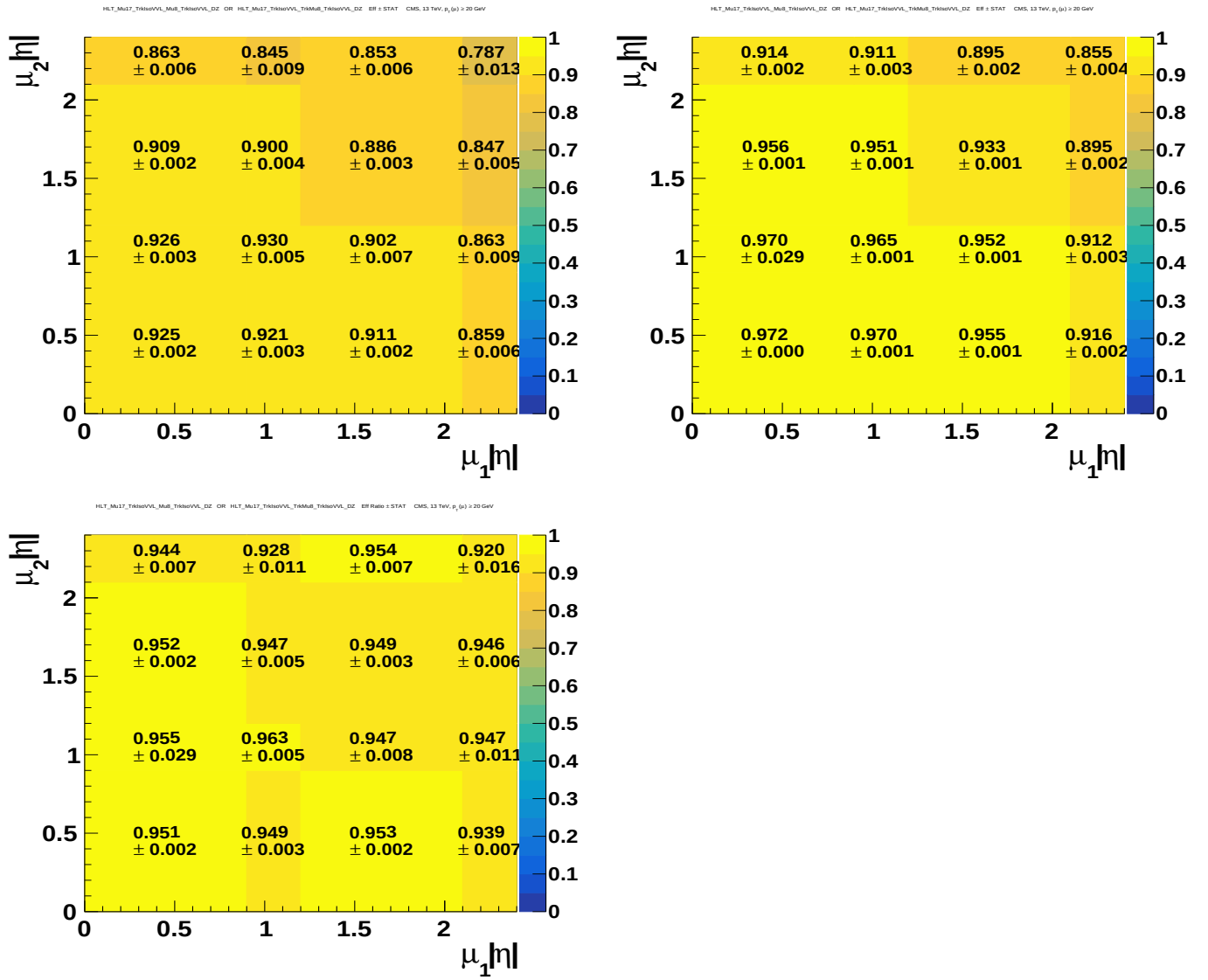


Figure 4.2: The double muon trigger efficiencies measured with tag and probe approach for the data collected in 2015 (top left), and simulated DY + jets sample (top right), and the ratio of data to MC (bottom), as a function of the pseudorapidity of the leading and subleading muon candidates, with $p_T > 20$ GeV. The results are taken from [74].

4.3.2 Track and vertex reconstruction

Before selecting the particle candidates such as muons, electrons, or jets, one needs to obtain the reconstructed objects using the detector response. First, track and vertex reconstruction is performed for the signal information.

The proton-proton collisions happen in the center of the detector, and the emitted charged particles can traverse the tracker system (section 3.2.3). These particles can directly hit the pixel and strip modules through their propagations and the hits are formed into clusters, whose positions are measured with uncertainties in the local coordinate system of each sensor. Since the tracker system

is in the quasi-uniform magnetic field, the charged particles pass helical paths due to the Lorentz force. In order to formulate a charged particle trajectory accurately, five parameters are used for the definition. The reconstruction system implements a seed generation, which provides the initial estimate of the trajectory parameters and their uncertainties for the initial track candidates using a few (2-3) hits. This initialization is based on the assumption that the particle is produced close to the beam spot.

Beginning with the trajectory seed generation, the track modules implement a Kalman filter approach [75] to find the accurate tracks. They rebuild the track candidates through taking into account the additional hits due to multiple scattering from successive detector layers, with the parameters of tracks updated. The order is from innermost hits outwards. For each adding step, a χ^2 test is executed to check the compatibility between the hits and the extrapolated trajectory. It is used to reject outliers during the track fitting. For instance, if one or more pixel hits are inconsistent with the trend expected for the reconstructed track angle, they are classified as invalid hits and excluded. Concerning the silicon modules in the layers, if trajectory intercepts the module surface more than three Gaussian standard deviations (convention) outside the module boundary, the module hits are ignored. On the other hand, the valid hits from each of the modules groups are collected to form a cluster respectively. At the same time, the uncertainties of the trajectory parameters are also propagated. The clusters will be associated to a track. This approach above is proceeded in an iterative pattern.

Considering the time consuming and the quality of the track reconstruction, there are some constraints on the fitted trajectory candidates, such as minimum number of hits, maximum number of lost hits, minimum track p_T . Meanwhile, if a track fitting fails, the track will be discarded.

However, this standard Kalman filter is not appropriate for electrons, which lose energy preliminarily through bremsstrahlung in top of ionization, and the distribution of the large energy loss is non-Gaussian. For instance, approximately one third of electrons radiate more than 70% of their initial energy before depositing on the electromagnetic calorimeter surrounding the tracker. In order to overcome this drawback, a different technique is designed for electron candidates (see section 4.3.5).

When a set of tracks are reconstructed and selected, they are grouped if several tracks appear to originate from the same primary interaction position. At this point, the primary vertices can be inferred preliminarily. By use of all the associated tracks together with the corresponding pixel hits, the track system implements fitting for the position of each vertex accurately. Once all the vertices are found, they are sorted according to the sum of the squared tracks transverse momenta. Among them, the vertex with the maximum summed value is treated as the hardest collision.

Figure 4.3 shows an event with 50 reconstructed vertices inside the CMS tracker system in 2012. According to this example, we can have an intuitive idea of the challenges that come with increasing the luminosity.

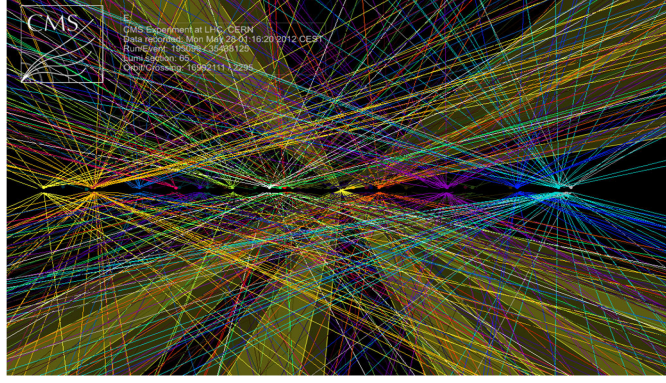


Figure 4.3: A side view of the center of CMS detector immediately after two beams pass through each other. One estimates that 50 independent collisions happen almost simultaneously.

4.3.3 Particle flow reconstruction

After the tracks and interaction vertices are reconstructed, one needs to combine them with calorimeter and muon detector signals to distinguish and classify the particles corresponding to them. At this point, the particle flow reconstruction scenario is designed by the CMS collaboration. It is used to reconstruct and identify all the stable particles from the proton-proton collision events, including muons, electrons, photons, charged hadrons, and neutral hadrons. Their directions, types, energies are determined optimally through a thorough combination of all CMS sub-detectors. The particle flow event reconstruction should have a high efficiency, while the misreconstruction rate should be low, even in the environment of high density. Hence, the advanced track and cluster algorithm is developed for that.

As introduced in section 4.3.2, the tracker is a cornerstone of event reconstruction process. The tracker should have high efficiency close to 100%. Moreover, it must be able to suppress misreconstruction rate of tracking, which has a randomly distributed momentum, and it may lead to large energy excesses. In order to achieve this aim, the iterative pattern is implemented. It has been found that the track originated close to the beam axis has an efficiency of about 99.5% for muons and more than 90% for charged hadrons [76]. All in all, the charged particles have a few percent of fake tracking rate.

In general, the track works efficiently for the charged particles, but it cannot identify the stable neutral particles including photons and neutral hadrons. Therefore, additional technique is needed to measure the position and energy of these particles, and separate them from the charged particles. It can happen that the track parameters of charged particles are not measured accurately at even high p_T tracks. Furthermore, the photons emitted by the electrons due to bremsstrahlung process are also missed by tracking procedure. Hence, another procedure is invoked, called calorimeter clustering.

The developed algorithm is composed of three steps. The first one is to identify the "cluster seeds" as local calorimeter cell energy maximum above a provided threshold. The second step, one grows the "topological clusters" from the seeds by aggregating the cells with at least one side together with

a cell already in the cluster, if the detected energy is beyond a given value. These threshold values stand for the two standard deviations above the electronic noise in the ECAL (80 MeV in the barrel and up to 300 MeV in the endcaps), and up to 800 MeV in the HCAL. Then during the last step, the "particle flow clusters" are supplied as seeds by the topological clusters. Since the calorimeter has high granularity, the energy of each cell among the particle flow clusters are shared. According to the cell-cluster distance, an iterative determination of the cluster energy and position is implemented.

Since some single particle can give rise to several particle flow elements (clusters) in different sub-detectors, these elements need to be connected with each other for the full reconstruction. Thus, a link algorithm is dedicated to this requirement. Benefiting from the high granularity of the detectors, the algorithm can aggregate the elements through quantifying each distance in the (η, ϕ) plane between the extrapolated track position and the cluster position. The specific procedures are divided by three parts. First, the track is extrapolated from the last detected hit in the tracker to the two layers of pre-shower system. Second, the extrapolation will be proceeded to the ECAL at the depth of typical electron shower profile. The last step is extending the track to the HCAL according to the expected hadron interaction length. The track can be linked to any cluster, as long as the extrapolation position in the corresponding calorimeter is within the edge of the cluster. Since there are gaps between the electronic cells in the calorimeter and cracks between each calorimeter modules, the envelope of each cluster can be extended to the size of a cell in each direction. This is also aimed to take into account the uncertainty of the shower maximum position and multiple scattering effect of the charged particles with low energy.

This algorithm can deal with the energy collection for the photons from bremsstrahlung effect of electrons, according to the tangents to the tracks from the intersection locations between the track and each layer, until reaching the ECAL. If the tangent points are within the boundaries of a cluster, this cluster will become a link to the track as a bremsstrahlung photon candidate.

It is also similar for a link between two calorimeter clusters. The link is constructed if the cluster position in the more granular calorimeter is within the cluster envelope in the less granular calorimeter. Furthermore, a link between the track of a charged particle in the tracker system and a muon track in the muon system can be established, via a global fit between the two tracks, once a suitable χ^2 is returned.

After the links are implemented, the blocks of clusters are formed. Then the particle flow algorithm will proceed with the identification and reconstruction of sets of particles. The "particle flow muon" is the first to be identified if the combined momentum is compatible with the value determined from the tracker. The next one is electron reconstruction and identification. By taking into account the short tracks and bremsstrahlung effect of electrons, the identified electron tracks are fitted with a Gaussian Sum Filter [77] (see section 4.3.5) to trace their trajectories. During each step, the corresponding track and cluster information is removed from the block once the particle flow candidate is identified.

Afterwards, one needs to apply special criteria for the remained tracks, and also calibrate the calorimetric clusters left [76]. If the tracks are compatible with the calorimetric energy, the momenta

of charged hadrons are extracted from the clusters by use of the combinatorial information from the tracker and deposited calorimeters. Finally, the rest clusters give rise to particle flow photons and neutral hadrons, according to the calorimetric cluster calibration.

4.3.4 Muons

The proton-proton collisions produce a lot of muons which can typically traverse the whole CMS detector. The muons have long mean lifetime of $2.2 \mu\text{s}$, and they lose small energy when traveling through the matters in the detector with the energy in the magnitude of GeV. The mass of a muon is 105 MeV, which is 200 times larger than the mass of an electron. This feature makes the muons have low radiation loss in the detector. During the period of passing through the detector, muons only leave few energy for the ionization effect around their trajectories.

Muons are reconstructed and identified using the combinatorial information from the tracker and muon chamber system. There are two methods for the reconstruction. The first one is implementing a matching on the tracks from the tracker and muon chamber respectively, by comparing the track parameters and combining the hits in the two systems. The global track is actually fitted using Kalman Filter approach. The other pattern is searching a match between any muon segment in the muon chamber and the tracks extrapolated from the tracker. The latter takes into account the elements including magnetic field, average expected energy loss, and multiple Coulomb scattering in the detector. Benefiting from the high efficiencies of the reconstruction in the tracker and muon system, approximate 99% of muons generated in the proton-proton collisions within the geometrical acceptance of the muon detector are reconstructed with the two approaches. More details can be found in [78]. As mentioned in section 4.3.1, the second method costs much more time than the first approach. It has to be noted that in high level trigger, the reconstruction stops once a muon candidate is reconstructed successfully.

After the trigger filters, the muon candidates are selected from the particle flow (PF) collection [76]. There is requirement about the matching between trigger objects and PF objects, which is made in $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ with $\Delta R \leq 0.3$, where $\Delta\eta$ is the pseudorapidity difference between the muon candidates in trigger and PF levels respectively, and $\Delta\phi$ stands for the difference in azimuthal angle of the two particle candidates. This is aimed for applying trigger scale factors on simulation correctly due to the difference between data and simulation on trigger level, and also for correcting the trigger efficiency. Furthermore, the muon candidates must pass the identification criteria in table 4.5.

In order to suppress the muons not originating from Z boson decay, an isolation variable is defined:

$$R_{iso} = [\sum^{charged} p_T + \max(0, \sum^{neutral} p_T + \sum^{EM} p_T - 0.5 \sum^{PU} p_T)] / p_T^\mu \quad (4.2)$$

where the sums run over the PF candidates in a cone of radius $\Delta R = 0.4$ around the direction of the muon candidate track, divided by muon p_T (p_T^μ). The PF candidates include: charged hadrons

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

Table 4.5: The list of criteria applied for the tight identification of muons, the definition of each parameter is explained in [79] for 8 TeV and [80] for 13 TeV.

produced from the primary vertex of the event, neutral hadrons, photons (EM), and charged hadrons not originating from the primary vertex due to pileup events (PU, see section 4.3.8). Since neutral pileup particles deposit collectively on average half as much energy as charged pileup particles, the contamination in the isolation cone from neutral coming from pileup interactions is estimated as $0.5 \sum^{PU} p_T$, which is subtracted in R_{iso} . More details are described in [81]. The required cut is $R_{iso} \leq 0.2$ for 8 TeV and $R_{iso} \leq 0.25$ for 13 TeV. Furthermore, the analyses have kinematically constrained on muons with criteria 4.3. Note that the p_T cut is higher than the one of trigger path (17 GeV). This is to select the muon candidates that pass the trigger when the trigger efficiency is very high (nearly 100%).

$$p_T^\mu > 20 \text{ GeV}, |\eta_\mu| < 2.4 \quad (4.3)$$

In a word, these selection criteria are aimed to extract the muons with high purity, quality and efficiency. However, the efficiencies of the selections between data taken by detector and MC simulation are different. Hence, additional corrections are implemented, which are muon p_T and η dependent factors. They will be deployed to the simulated events to take into account the residual disagreement between data and MC. These correction factors are close to unity [74]. Figure 4.4 and figure 4.5 show the measured efficiencies of muon identification, reconstruction, and isolation, which compares the different performances between data and simulation at 8 TeV and 13 TeV respectively. Besides those, the misalignment of the tracker of detector can also result in the difference between data and MC in reconstructed muon momentum ³.

There is a global efficiency loss after the kinematic selections, because of the finite resolution and non-linear response of the detector. This effect can be corrected using unfolding methodology, which will be introduced later.

³The dedicated correction is called rochester momentum correction [82]. It is derived as a function of muon charge and angular observables, by means of Z boson peak which has been already well understood. Regarding this analysis, this effect has influenced the statistics by within 1%.

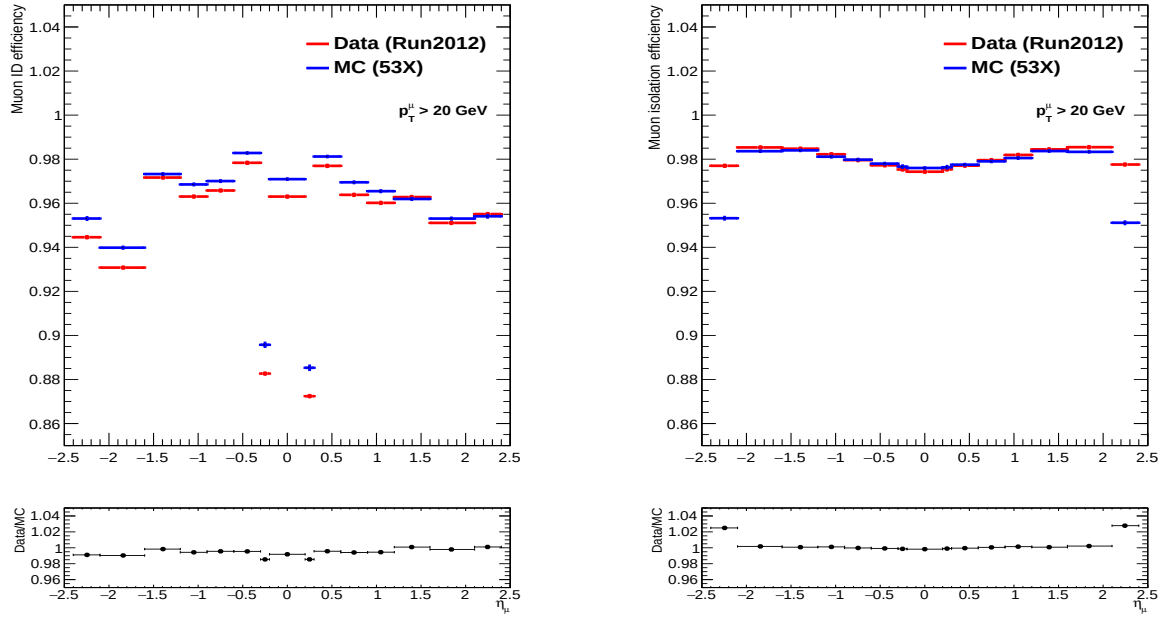


Figure 4.4: The measured tight muon identification and reconstruction efficiency (left) and the isolation efficiency (right), as a function of the pseudorapidities of the leading muon candidate, with $p_T > 20$ GeV. The data were collected in 2012, compared with simulation. The results are taken from [74].

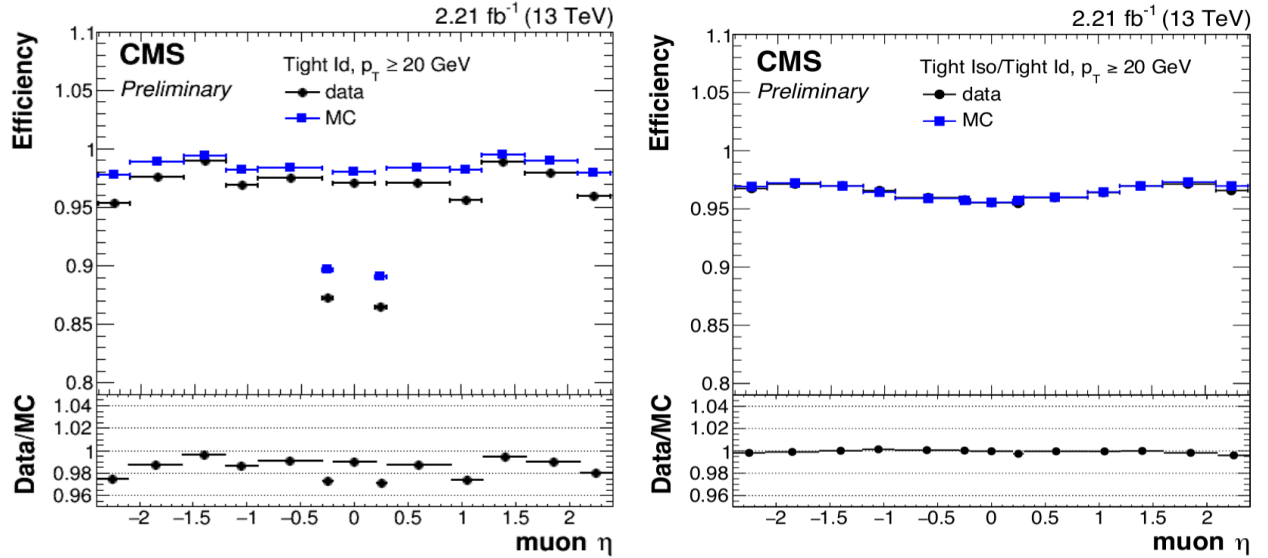


Figure 4.5: The measured tight muon identification efficiency (left) and the isolation efficiency (right), as a function of the pseudorapidities of the leading muon candidate, with $p_T > 20$ GeV. The data were collected in 2015, compared with simulation. The results are fetched from [83].

4.3.5 Electrons

Electrons are identified and reconstructed from the track they leave in the tracker matching with an electromagnetic clusters in the ECAL.

As was mentioned in section 4.3.2, electrons lose energy through bremsstrahlung effect in top of ionization, so they suffer from larger energy loss than other charged particles. Approximate 35% of electrons from the interaction center radiate more than 70% of their initial energy before arriving at ECAL. The profile of such a kind of energy loss is not a Gaussian distribution. As a consequence, the standard Kalman Filter approach is not suitable for the electron track reconstruction. Therefore, a dedicated method is developed for solving this issue. This method is to reconstruct electrons through associating tracks in the silicon detector with the clusters of energy deposited in ECAL from the radiated photons which spread along the electrons trajectory. Due to the strong CMS magnetic field, the electron trajectories are bent along the azimuthal (ϕ) coordinate, while not in longitudinal (η) coordinate. At this point, the photons from bremsstrahlung, which have the similar η value and within some ranges of ϕ , are merged into the electron candidate clusters.

In order to measure the initial energy of the electron accurately, two clustering algorithms are developed to take into account the spread energy of electrons over the multiple crystals of the ECAL. One is called “hybrid algorithm” that is applied in the barrel region. It starts at a seed crystal that contains most of the energy deposited in the whole barrel, and the transverse energy should be above a specific threshold (1 GeV). Then arrays of 5×1 crystals in the (η, ϕ) plane are considered around the seed crystal, in a range of 17 crystals along the azimuthal directions if their energy are beyond a minimal threshold (0.1 GeV). Then the contiguous arrays are collected into clusters. If each cluster has an energy exceeding a threshold (0.35 GeV), it is grouped into a final global cluster, which is called supercluster (SC). The other algorithm is called “multi- 5×5 ” method that is used in endcap region of ECAL. Similar to the first algorithm, it looks for the seed crystals which receive the maximal energy locally above a threshold (0.18 GeV), and then collects the energy in clusters of 5×5 crystals around these seeds starting from the one with largest transverse energy. If the total transverse energy of these clusters is larger than a threshold (1 GeV), they are incorporated into a supercluster, which is within the range of (0.07, 0.3 rad) in the (η, ϕ) direction around each seed crystal. Note that the contribution of each cluster inside the supercluster is weighted by its energy. The positions of these clusters are then extrapolated to the plane of preshower. At this point, the preshower energy, within the range of extra (0.15 rad) with respect to the spread range of these clusters along the ϕ direction, is merged into the supercluster energy.

Electron track reconstruction starts from seeding, which is to find and select the two or three hits in the tracker for track initialization. The procedure is similar to the standard track reconstruction (see section 4.3.2). Because of the bremsstrahlung process, electrons’ tracks usually cannot be fitted smoothly or with few associated hits. A Gaussian Sum Filter (GSF) approach [77] is implemented after the track seeding procedure. The GSF methodology treats the fractional energy loss distribution as the superposition of different Gaussian functions, each of which comes with a weight value that

stands for its associated probability. That means a reconstructed track is a mixture of several trajectory components, and the parameters of each component are with Gaussian uncertainties. During the processing time, the less probable components are thrown away or merged with the similar components, and the weight of each component is updated according to the compatibility with the hits. It has been calibrated such that this approach improves the resolution of reconstructed electrons significantly compared to the traditional Kalman Filter method [84].

By associating reconstructed tracks and the superclusters energies, one can obtain the reconstructed electron candidates (“gsfElectron”). Actually, most of the bremsstrahlung photons are recovered by means of GSF scenario. However, there are still some converted photons missed, for which a dedicated MVA algorithm [85] is deployed. It usually happens to the supercluster with energy leakage or loss, and additional energy from pileup (see section 4.3.8), there is an MVA regression technique [86] used for the energy correction. At trigger level, the electron reconstruction is stopped once an electron candidate is reconstructed successfully, which is aimed to reduce time consume.

Figure 4.6 shows electron reconstruction efficiency as a function of the clustering energy deposited in ECAL, comparing data and simulation with the GSF approach. We can see the high reconstruction efficiency and excellent agreement between data and simulation.

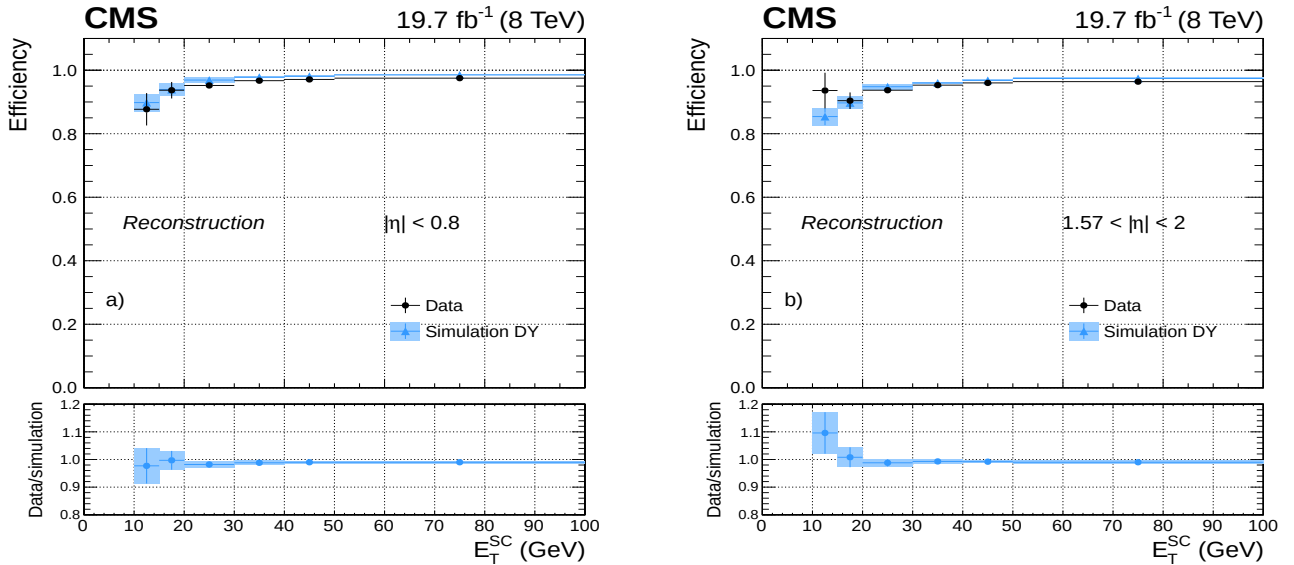


Figure 4.6: Electron reconstruction efficiency measured in dielectron events, as a function of super-cluster transverse energy in ECAL, by comparing data and simulation. The uncertainties shown in the plots correspond to the quadratic sum of the statistical and systematic contributions. The results are taken from [87].

After the electron candidates are selected from the “gsfElectron” collection [88], they are required to fulfill a set of medium identification criteria in table 4.6. Note that the electrons coming from converted photons are also removed through track fitting together with missing hits requirement.

We apply also an isolation requirement for all the electron candidates in the medium identification conditions, based on the PF algorithm:

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

Table 4.6: The list of criteria applied for the medium identification of electrons. The definition of each variable is explained detailedly in [88].

$$PFIsoRhoCorr = [\sum^{charged} p_T + \max(0, \sum^{neutral} p_T + \sum^{EM} p_T - \rho EA)]/p_T^e \leq 0.15 \quad (4.4)$$

where the sums run over all the PF candidates within a cone size that is chosen as $\Delta R \leq 0.3$. The item ρEA stands for pileup effect correction, where the ρ represents the level of diffuse noise related to the number of additional transverse momentum per unit area, and the EA is the effective areas of jets. In this way, a possible double counting of energy assigned to particle candidates is avoided. Therefore, other objects that can be misidentified as electrons are rejected efficiently, as well as pileup subtraction from reconstructed electron candidates.

Based on the constraints above, additional kinematic restriction is implemented to improve the signal purity using the detector region, where the reconstruction quality is high. They are listed in expression 4.5. Note that the η is the pseudorapidity of the supercluster associated to electron candidates. In general, these selections lead to increase of efficiency on both data and simulation. As similar to the case of muons, the efficiencies between data and MC using the above selection criteria are different. Therefore, correction factors for compensating this difference are estimated by means of tag and probe method.

$$p_T^e > 20 GeV, |\eta_e| < 2.4 \quad (4.5)$$

Figure 4.7 shows the efficiency as a function of electron p_T , based on the medium identification criteria. The efficiency is lower than the one for muon identification. By comparing the performances of data and simulation, we see good agreement between them ⁴.

4.3.6 Z Bosons

Based on the selected charged lepton pair with the highest transverse momenta, the Z boson candidates can be reconstructed. By summing over the four-momenta of both leptons, one can obtain

⁴We remind that the electron pair channel is not detailed in the present document for the 13 TeV case, which is referred to [59].

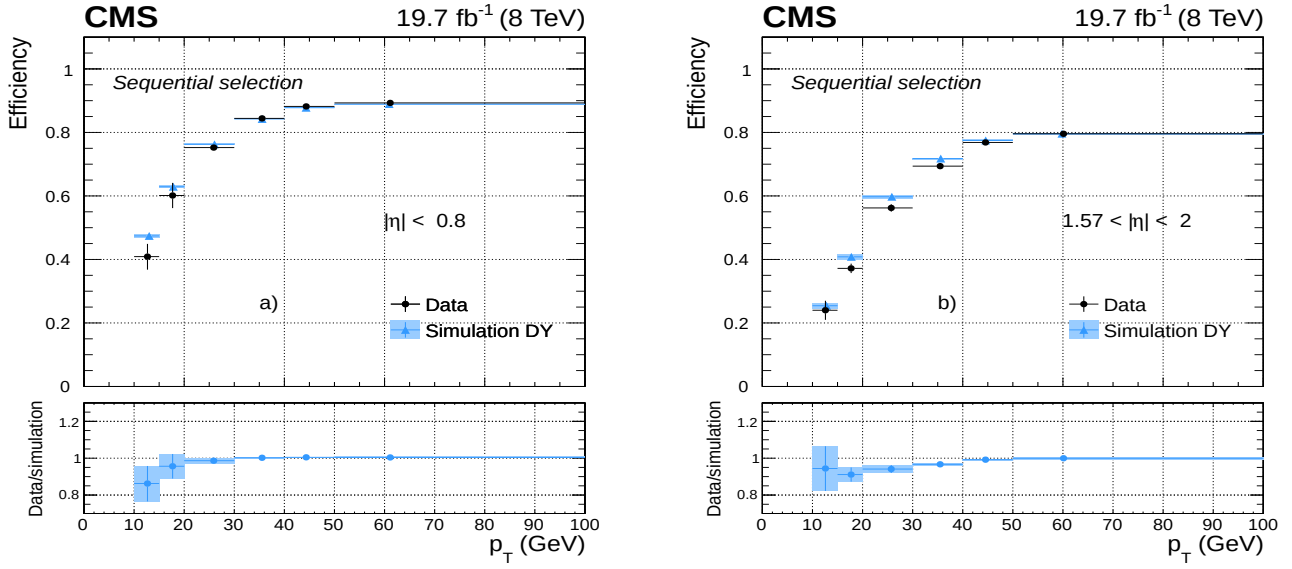


Figure 4.7: Electron identification efficiency measured in dielectron events, as a function of electron p_T , by comparing data and simulation, for the medium selection criteria. The uncertainties shown in the plots correspond to the quadratic sum of the statistical and systematic contributions. The results are taken from [87].

the four-momentum of the Z boson. The two leptons must have opposite charges, and the Z boson invariant mass should be within the window from 71 GeV to 111 GeV.

4.3.7 Jets

Jets are defined as collimated streams made of stable particles, including hadrons (mainly) and photons, produced from fragmentation and hadronization of partons (quarks and gluons) from proton-proton collisions. The characteristics of a jet is given by the constituted hadrons inside it. The hadrons contains not only charged particles but also neutral ones. The jet energy fractions typically carried by charged particles, neutral hadrons, and photons are 65%, 10%, and 25%, respectively. Most of the particles in jets penetrate ECAL with small energy loss, and deposit their energy in HCAL. In addition, the charged hadrons generate tracks in the CMS tracker system. Therefore, jets are reconstructed using the energy deposition in both ECAL and HCAL, and the associated tracks (see section 4.3.2). Figure 4.8 shows a simulated jet profile in the CMS detector.

- Jet reconstruction

The jets are defined using an anti- k_T clustering algorithm (see section 2.3.3). By use of the hit HCAL cells and the geometrically corresponding ECAL crystals, calorimeter towers are formed reflecting the energy deposition of particles (like hadrons, photons). A four-momentum is associated with each tower deposit above a designed threshold. Moreover, the tracker provides the tracks of charged particles. Through the combinatorial information of tracks and calorimeter towers, the par-

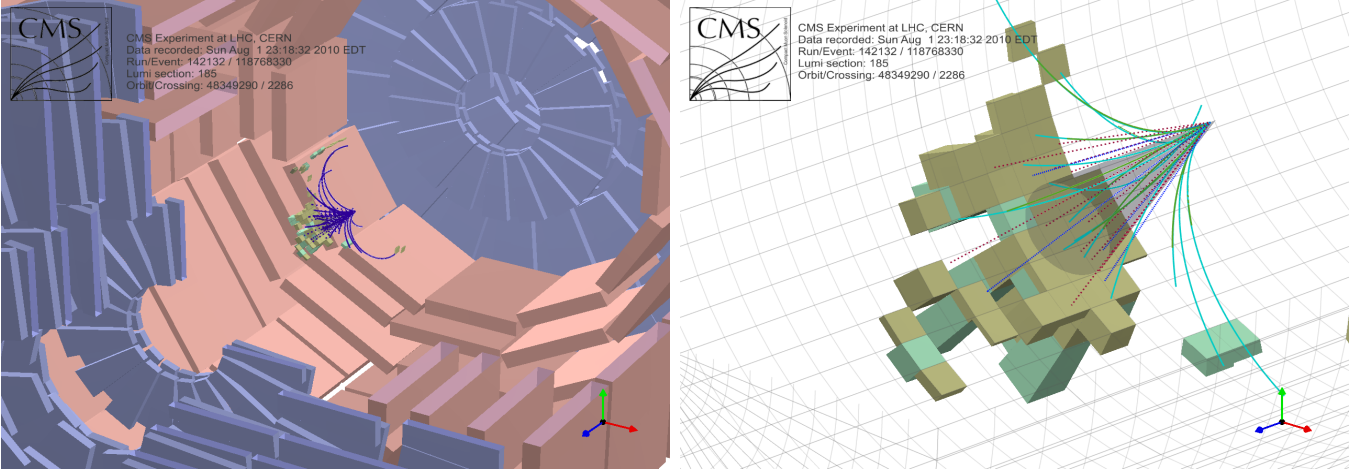


Figure 4.8: Left: A produced jet in the CMS detector. Right: A completed jet from a primary vertex, recorded by CMS detector, with its tracks (the green lines), energy deposit in ECAL (yellowish squares), photons (dotted purplish lines), neutral hadrons (dotted blue lines), energy deposit in HCAL (teal rectangles), charged hadrons (bright blue lines).

particle flow candidates are reconstructed. The jets are reconstructed by clustering the four-momentum vectors of the particle flow candidates.

- Jet energy calibration

After jets are reconstructed using the one algorithm, we need to calibrate their energy scale before selecting them for offline analyses [89]. The jet energy correction is parametrized with sequential levels, which correspond to the scaling of the four-momentum of the jet as a function of different observables, such as number of pileup, jet p_T , η , flavour. The jet energy resolution is generally worse than the resolutions of muons, electrons, and photons reconstructed by the CMS detector. The jet energy resolution can smear the spectra measurement precision or resonance decay searches. It is also influenced by the jet energy scale. Hence, it is important to calibrate the jet energy after applying jet energy correction.

The first correction step (L1) is to subtract the energy that are not associated with the hard interaction. The contributions to this kind of energy are from electronic noise, and multiple proton-proton collisions in the same or adjacent bunch crossings (pileup, see section 4.3.8). Both of the sources lead to increasing the jet energy. In addition, the pileup energy varies in different instantaneous luminosity delivered by CMS detector, which is also taken into account in this correction procedure. This is called offset energy correction. It is parametrized as a function of jet energy density, area, number of pileup, jet p_T and η , which is determined from simulation. It can happen that data differ from simulation with respect to the offset energy due to dead materials not well considered in simulation. These differences also depend on the varying pileup amount. Then additional offset scale factor for data/simulation is determined from zero-bias data and neutrino gun simulation (+ noise and pileup simulation), with the random cone method (see section 5.5.1). The zero-bias data sample is collected

using a random trigger during a beam crossing. Because these events are not triggered by any specific energy deposit, the main sources of the energy deposits in these events are detector noise and pileup. Figure 4.9 shows the pileup offset correction factor with respect to jet p_T and η .

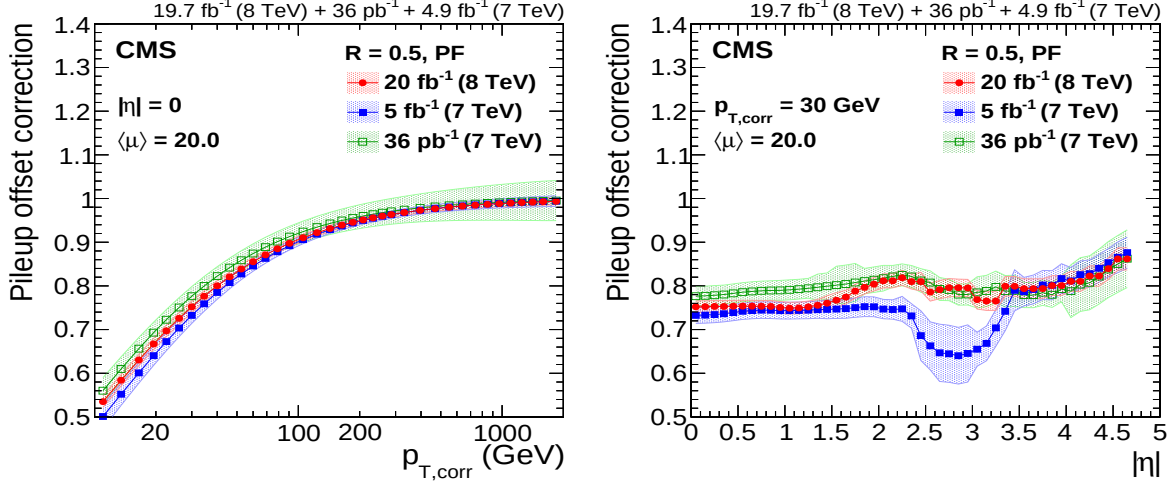


Figure 4.9: Pileup offset correction including data/MC scale factors, as a function of corrected jet p_T (left) and absolute η (right), with systematic uncertainty band, for the average conditions of 20 pileup ($\langle \mu \rangle$). The results are compared between 7 TeV and 8 TeV, with respect to the particle flow jet cluster using the anti- k_T ($R = 0.5$) algorithm. This figure is fetched from [90].

The second level (L2/relative) is to correct the non-uniform response of the jets in the η direction of the CMS detector, so that the response distribution becomes flat in the whole η scope. This correction parametrization can be determined using MC truth, through matching each generated jet with the closest reconstructed (RAW) jet, according to their angular distance (ΔR), and then calculating the ratio of their p_T as the correction factor for both data and simulation. Figure 4.10 shows the L2 correction factor and the response after correction as a function of jet $|\eta|$, using simulations.

The third step is called absolute correction (L3), which is dedicated to deal with the jet p_T variations due to the non-linear response effect of the calorimeter. The scaling factors are dependent on the reconstructed jet p_T . At an average level, we retrieve the true jet momenta through applying the factors to the reconstructed jet momenta. On purpose, the average response is set to unity in the whole p_T region. Similar to the L2 correction, the L3 correction can be determined using MC truth. Figure 4.11 shows the L3 correction factor profiles and the response after the correction applied, using QCD enriched sample for simulations.

These three levels of corrections are applied to both data and simulation. After that, the jet response performance of simulation has excellent results. However, the results of data are still with a few discrepancies. In order to solve this issue, an additional correction is implemented for data. It is named as L2L3 residual correction, which is composed of scaling factors dependent on jet η and p_T . For data, there are only reconstructed jets. One can deal with the events with two body processes, where the two objects can have good p_T balance. We select a tag object, which can be a jet in the central region ($|\eta| < 1.3$), or an efficiently reconstructed object like a Z boson or a photon,

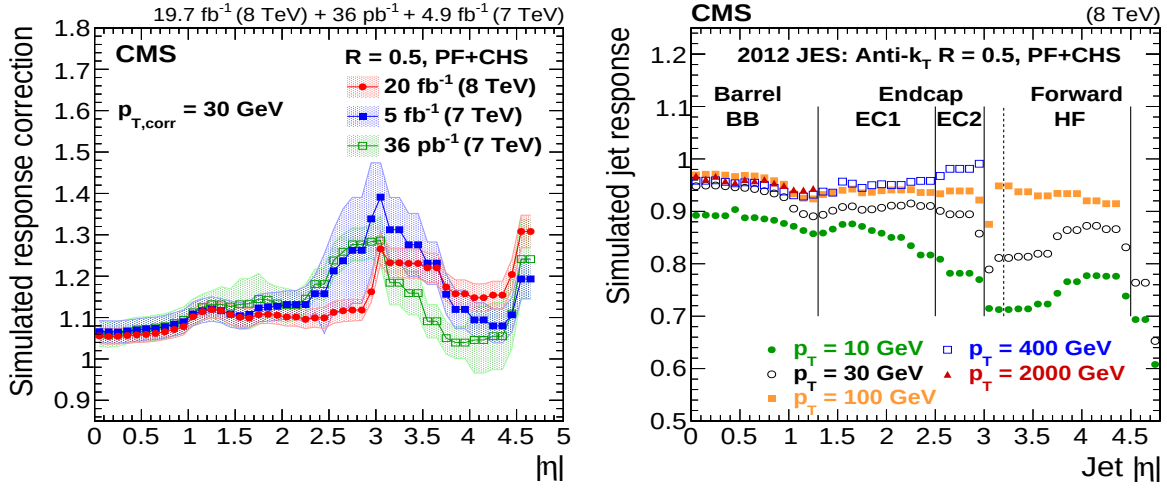


Figure 4.10: Simulated relative correction (left) factor in p_T and response with L2 correction applied (right), as a function of jet $|\eta|$, with systematic uncertainty band. This figure is fetched from [90].

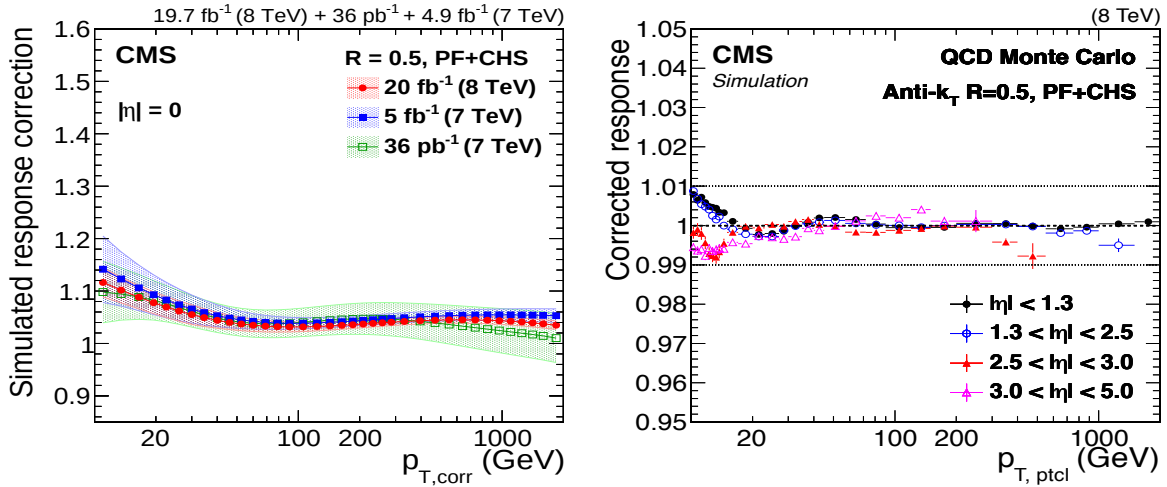


Figure 4.11: Simulated absolute correction (left) in p_T and response with L3 correction applied (right), as a function of jet p_T , with systematic uncertainty band. This figure is fetched from [90].

and then search for a probe jet in the unrestricted η regions for the response (RAW p_T divided by tag p_T) estimation. This method is called p_T balance approach. After this calibration, the data and simulation have compatible performances. Figure 4.12 shows the residual correction distributions with respect to jet $|\eta|$ and p_T .

After the full jet energy corrections are applied [91], the resolution can be obtained from the width of the response (p_T^{reco}/p_T^{gen}) distributions using fit methodology. The profiles of responses are not strictly symmetric, due to inactive ECAL areas, high p_T particles not leaving their full energy in HCAL, or particle level imbalance with neutrinos produced. These effects give rise to longer tail at low response scope compared to the other side. In approximate ways, jet energy responses are usually modeled using Gaussian functions or Crystal Ball functions, for instance. The full tables for the jet energy calibration are listed in [92].

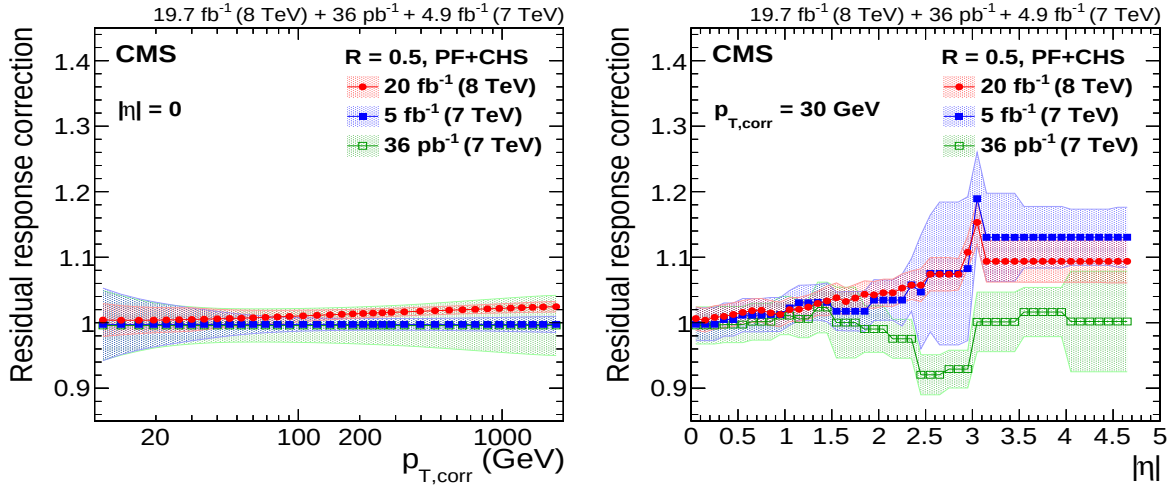


Figure 4.12: L2L3 Residual correction as a function of jet p_T (left) and $|\eta|$ (right), with systematic uncertainty band. This figure is fetched from [90].

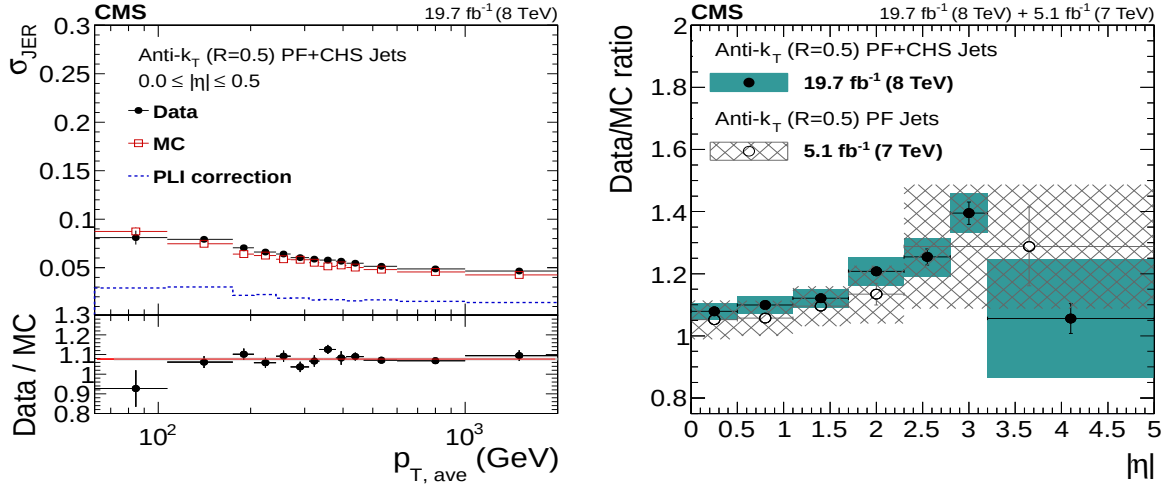


Figure 4.13: Jet momentum resolution as a function of jet p_T (left) and $|\eta|$ (right), comparing data and simulation. The left plot contains the distribution with particle level imbalance (PLI) effect corrected. This figure is fetched from [90].

Figure 4.13 illustrates the jet transverse momentum resolution as a function of jet p_T and $|\eta|$, comparing data and MC. The calibration is performed using events with two jets in the final state. One can notice that simulations give rise to better resolution than measurements in general, which is a common phenomenon for the jet momentum resolution calibration.

In order to improve the agreement, one needs to smear the simulation with the following formula to adjust the jet momenta:

$$p_T^{smear} = \max(0, p_T^{gen} + C(\eta) \times (p_T^{reco} - p_T^{gen})), E^{smear} = E^{reco} \times (p_T^{smear} / p_T^{reco}) \quad (4.6)$$

where $p_T^{reco}(E^{reco})$ and $p_T^{gen}(E^{smear})$ are the reconstructed and generated transverse momentum(energy) of a jet. The $C(\eta)$ is a factor that takes into account the smearing extent, and it comes with systematic uncertainties. They are shown in table 4.7 for 8 TeV and in table 4.8 for 13 TeV.

$ \eta $	$C(\eta)$	$C^{up}(\eta)$	$C^{down}(\eta)$
0 - 0.5	1.079	1.105	1.053
0.5 - 1.1	1.099	1.127	1.071
1.1 - 1.7	1.121	1.150	1.092
1.7 - 2.3	1.208	1.254	1.162
2.3 - 2.8	1.254	1.316	1.192
2.8 - 3.2	1.395	1.458	1.332
3.2 - 5.0	1.056	1.247	0.865

Table 4.7: Smearing factors for jet energy resolution correction at 8 TeV. This table is taken from [93].

$ \eta $	$C(\eta)$	$\sigma^{C(\eta)}$
0 - 0.8	1.061	0.023
0.8 - 1.3	1.088	0.029
1.3 - 1.9	1.106	0.030
1.9 - 2.5	1.126	0.094
2.5 - 3.0	1.343	0.123
3.0 - 3.2	1.303	0.111
3.2 - 5.0	1.320	0.286

Table 4.8: Smearing factors for jet energy resolution correction at 13 TeV. This table is taken from [93].

- Jet selection

Since we need the jets production in association with a Z boson, the jets coming from the same primary vertices as Z bosons are selected. Because particle flow particles are parts of jets, they can come from the hard interaction or from the pileup (see section 4.3.8). In order to subtract the pileup contamination, an approach called Charged Hadron Subtraction (CHS) is applied, which retains all the charged particles using the reconstructed primary vertices that fulfill the selection criteria below:

- The z component of primary vertex: $z < 24$ cm
- The radius of x-y plane of primary vertex: $\rho < 2$ cm
- The number of degrees of freedom for the vertex fit (number of hits used to reconstruct the tracks belonging to the vertex): $ndof > 4$

The charged particle tracks associated to pileup vertices are removed in this procedure. This approach accounts for about half of the offset energy from pileup in the region that the tracker covers, and the rest offset energy is subtracted by use of jet area method. After the jets are reconstructed, they are required to have a distance ($\Delta R > 0.5$ for 8 TeV and $\Delta R > 0.4$ for 13 TeV) with each of the two leptons from Z boson decay, otherwise are removed from jet collection. This is to avoid well reconstructed muons passing the jet selection criteria. Moreover, in order to reject misidentification and noise contamination, more criteria on the reconstructed jets are implemented as shown in table 4.9. In parallel, there is an unique variable that represents jet identification based on both vertex and jet shape information used, and a loose cut is applied on it for further pileup suppression [94]. The details are interpreted in section 4.3.8.

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

Table 4.9: List of criteria on jets for good identification. The “EM” represents the electromagnetic energy deposit in ECAL.

Moreover, additional criteria 4.7⁵ are required for jets for reducing pileup contamination and large uncertainty on the energies of jets. With the aim to guarantee a good quality of the tracking information, forward jets are removed.

$$p_T^{jet} > 30 GeV, |y_{jet}| < 2.4 \quad (4.7)$$

4.3.8 Pileup

Because the running of LHC has high intensities, multiple proton-proton collisions happen with large probability per bunch interaction. As a consequence, we need to separate the different interactions and to identify and reconstruct the event of interest, among the substantial background with these extra collisions. Such additional interactions are called pileup (or in-time pileup), and their rates is defined as the number of additional interactions per bunch crossing. The pileup interactions often produce jets with low p_T and induce additional tracks in the tracker system and energy deposition in the calorimeters. Since the signal development in the detector is longer than 25 ns, the pileup produced in one collision is possible to be recorded in the previous or following reconstructed event. This kind of pileup is named as out-of-time (OOT) pileup. The bunch spacing at 8 TeV is 50 ns, the OOT pileup has much smaller effect than the in-time pileup (IT pileup). On the other hand,

⁵The mass of a jet is assumed to be zero, which leads the jet η to be almost equal to the value of jet rapidity.

the bunch spacing at 13 TeV decreases to 25 ns, which together with the increased instantaneous luminosity result in the significant OOT pileup.

The amount of pileup is estimated according to the number of reconstructed primary vertices (see section 4.3.2) or the offset energy density [20]. Figure 4.14 shows the relations between them, using the data and simulations at 8 TeV. One can see that the amount of pileup depends almost linearly in the number of primary vertices or the offset energy density. Furthermore, in order to model the correlation between IT pileup and OOT pileup, we can generate the number of interactions for each bunch crossing using a Poisson distribution with the same mean value of pileup [95].

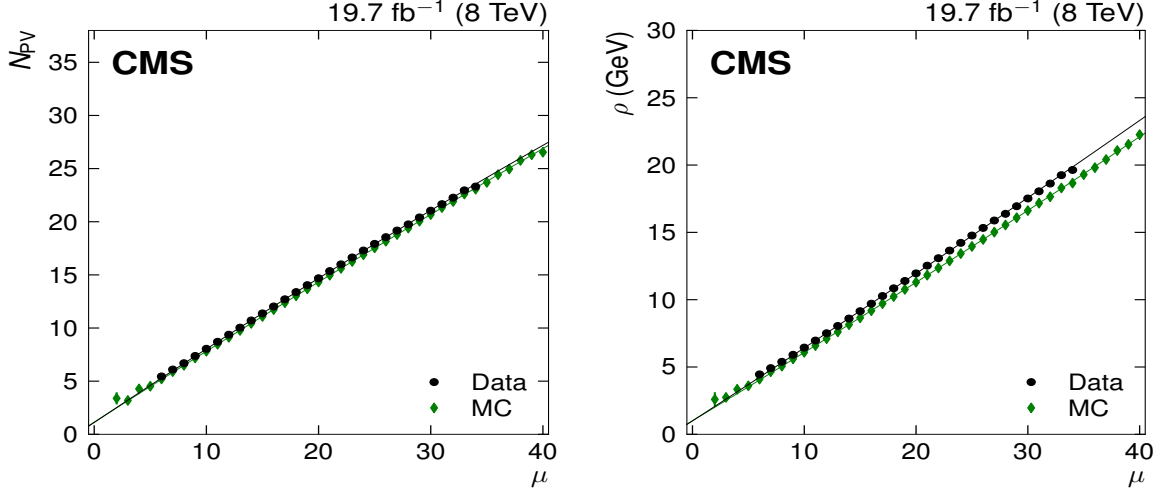


Figure 4.14: Mean number of primary vertices per event (N_{PV} left), mean offset energy density (ρ right), versus the average number of pileup per bunch crossing. The data are compared to the simulation using PYTHIA6. This figure is fetched from [90].

During the jet reconstruction, some jets originate from pileup collisions, so they are called pileup jets. The shape distributions of pileup jets and jets in association with Z boson from hard interactions are very different. Figure 4.15 illustrates the rate distributions of jets in association with Z boson (real jets) and pileup jet (PU jets) versus the number of primary vertices. Although the data and MC are not in good agreement, we can see the shape difference between real jet and pileup jet profiles.

According to the relevant analysis [94], charged particle flow objects contribute approximately half of pileup, and the other half contains neutral particles. Due to the coverage of the tracker volume ($|\eta| < 2.5$), the charged particles (mainly hadrons) contained in pileup jets can be efficiently distinguished and removed according to the reconstructed vertices. This is to reduce half of the pileup contamination in particle flow jets. Figure 4.16 shows the pileup offset energy per pileup interaction versus the jet η , comparing data and MC at 8 TeV. We can see good agreement between them in the central region, where the charged hadrons dominate the pileup energy.

Pileup jets are identified using the "Pileup Jet Id Algorithm" [94], which distinguishes pileup jets using vertex observables or shape related variables. The former has much higher efficiency than the latter, since it relies on the tracker system, which results in the efficiency of vertex reconstruction

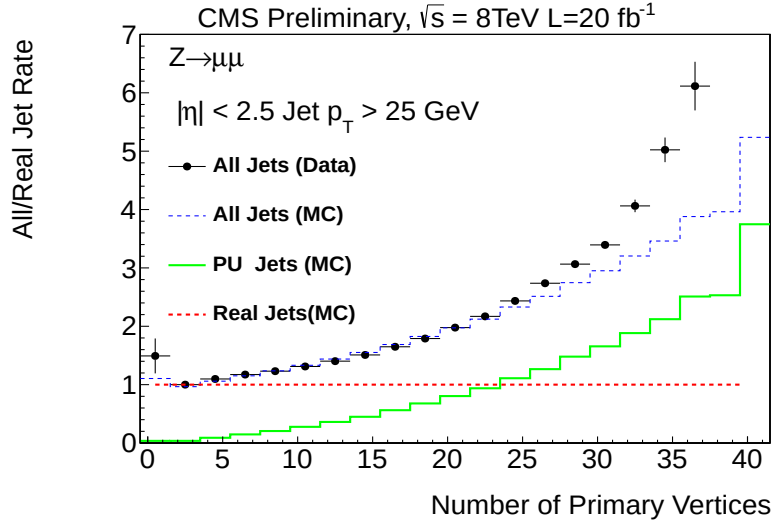


Figure 4.15: Rate of PU jets with $p_T > 25$ GeV relative to the expected rate of real jets as a function of the number of reconstructed primary vertices, comparing data and simulations. This figure is fetched from [94].

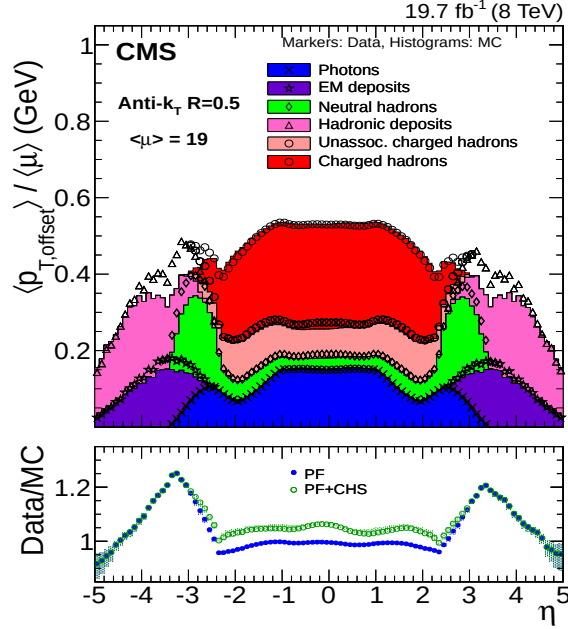


Figure 4.16: Offset energy normalized by the average number of pileup interactions ($\langle \mu \rangle$), comparing data and simulations. Different types of particle flow objects are overlapped. The fraction labeled "charged hadrons" is removed by CHS. This figure is fetched from [90].

close to 70%. While, it is only available within the region $|\eta| < 2.5$, and the latter has to be used in the scope without tracker volume, by handling the diffuseness of jet showers. Since we select the jets with $|\eta| < 2.4$, we always have the pileup jets well subtracted using the reconstructed primary vertices. With the aim to optimize the jet identification, more complicated approach is developed

based on MVA technique [94]. Figure 4.17 shows the pileup jet identification efficiency using the advanced MVA approach, with respect to jet η , p_T , and number of reconstructed vertices. Concerning the jet selection criteria in this analysis, we actually have the efficiency close to unity.

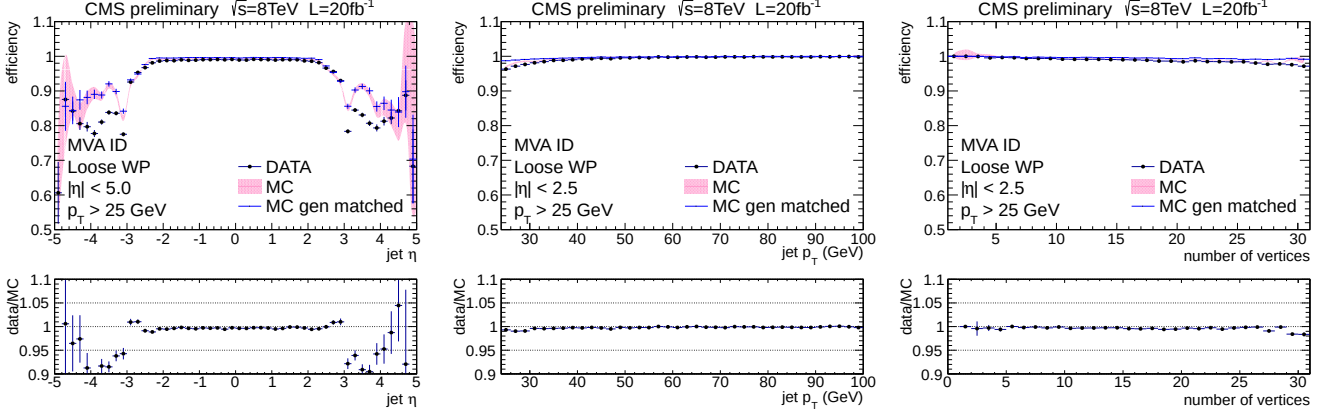


Figure 4.17: Pileup jet identification efficiency using $Z \rightarrow \mu^+\mu^- + \text{jets}$ events, as a function of jet η (left), p_T (middle), and number of vertices (right). This figure is fetched from [94].

The pileup conditions depends on proton-proton inelastic cross section and on the actual run conditions of data taking, which include the instantaneous luminosity. In order to roughly cover the pileup conditions expected for each data-taking period, many MC productions are implemented including the number of pileup interactions. Due to the unpredicted run conditions and different analysis scenarios, data and MC in the pileup distributions do not match a priori. Although the deterministic primary vertex reconstruction is shown to be efficient and behaves well up to high levels of pileup, there is a non-linear correlation between amount of pileup and number of primary vertices. Actually, the final number of reconstructed primary vertices is still sensitive to the details of the reconstruction, and gives rise to a difference in the underlying event between data and MC. In addition, the number of reconstructed vertices can be potentially influenced by the offline selection criteria and even by the trigger.

In order to improve the agreement between data and simulation, there are two alternative approaches that reweigh the MC to data. One is to simply count the number of reconstructed primary vertices, and make correspond the average number of additional vertices approximately linearly to the number of pileup interactions which concerns the total number of events. The other method is to use the measured instantaneous luminosity for a single proton-proton collision for each bunch pair, and multiply by the total inelastic cross section (69.4 mb for 8 TeV) using emulation, and then obtain the average number of interaction for this bunch crossing. The second method is called minimum bias cross section method [90].

In this Z+jets at 8/13 TeV analysis, we adopt the second method and implement double check with the first approach. The results are shown in figure 4.18. The performance at 13 TeV has fluctuations, which is related to technical issues on vertex reconstruction inefficiency in some periods. This discrepancy is covered by the corresponding systematic uncertainty.

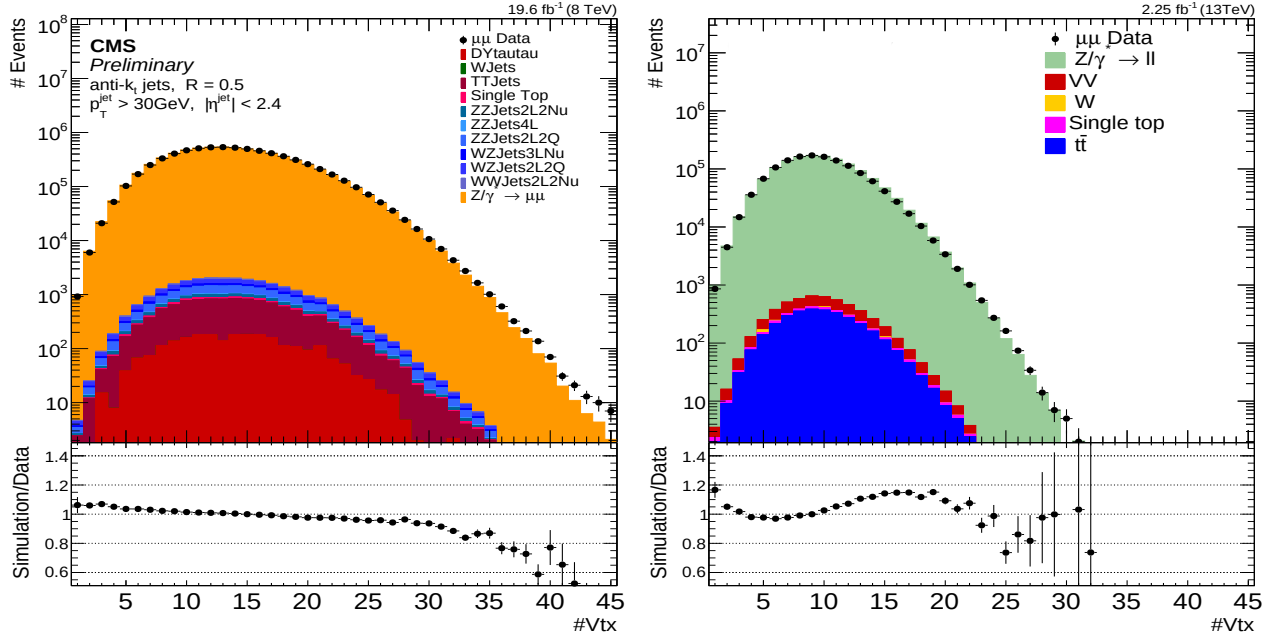


Figure 4.18: The data and simulation comparison at reconstruction level as a function of number of primary vertices after pileup re-weighting using minimum bias cross section method, at 8 TeV (left) and 13 TeV (right).

Figure 4.19 provides an event display of Z + jet process in the CMS detector during a data taking period.

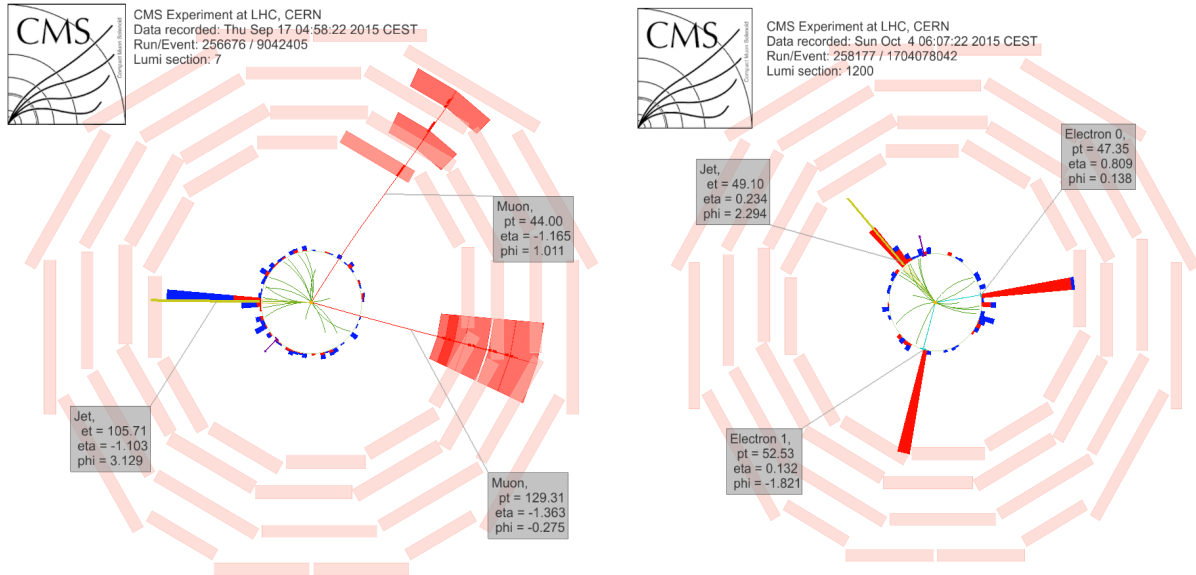


Figure 4.19: A reconstructed event in CMS with a jet associated with a pair of muons (left) / electrons (right). The number of primary vertices are seven (left) and four (right).

4.4 Results at detector reconstructed level

Based on the event selection, it is interesting to see the remained events of both data and simulation at the detector level. From that, one can have a preliminary impression about the statistics of data and the relative background contaminations, and know about the fraction of signal process in data, but also approximately judge whether the simulations are estimated precisely within statistical fluctuation.

For the comparison, the simulations include the signal process and all the significant background processes, each of which needs to be normalized to the integrated luminosity of data. Then they are superimposed to a histogram distribution and overlapped with the histogram from data. The histograms show the number of events as a function of an observable. The error bars in the comparison histogram represent only statistical uncertainties. Regarding the ratio of MC to data, the error bars are the propagation of statistical uncertainties of both of them. The comparison including systematic uncertainties will be presented for the measurements of cross sections later.

In this section, the results at 8 TeV and 13 TeV respectively will be shown. The former contains both $\mu^+\mu^-$ and e^+e^- channels, while the latter only involves $\mu^+\mu^-$ channel. Note that they are implemented up to different jet multiplicities due to the difference in statistics.

The presented profiles include all the distributions that will lead to differential cross sections, and give the cross check for the data quality. Note that only statistical uncertainty involves, so there is no firm conclusion about the agreement level.

4.4.1 Kinematic variable distributions at 8 TeV

Figure 4.20 and 4.21 are showing the lepton transverse momentum and pseudorapidity respectively. One can see generally good agreement between data and overlapped simulations. The peak at 45 GeV is due to the peak of Z boson mass. There is a drop around $|\eta| = 1.5$ region in e^+e^- channel that is because of the transition area between barrel and endcap parts of ECAL. Hence it is removed by our selection.

The invariant mass of lepton pair is shown in figure 4.22, with the constraint from 71 GeV to 111 GeV. There is a shape discrepancy not covered by the statistical uncertainty in both channels, which may be due to the tracker misalignment and residual magnetic field mismodelling on muons, and energy smearing effect on electrons. The approach to this issue is to apply energy correction and smear to the simulated DYJets sample to compensate such an effect. It has been checked that for $Z \rightarrow \mu^+\mu^-$ channel this issue affects the other distributions by less than 1%⁶. Figures 4.23 and 4.24 show the azimuthal distance and the separation, respectively, of the two leptons. Good agreements are seen in most of the ranges. Figures 4.25 and 4.26 present the Z boson p_T and rapidity respectively, without requirement on the jet multiplicity. From the former, one can see a small p_T dependent discrepancy between data and simulation. The jumps of the spectrum at 210 GeV and 320

⁶Such a correction is not applied in the present analysis.

GeV are due to the changed binning. The latter has a discrepancy in the endcap region, especially for the e^+e^- decay channel, which may result from that energy smearing for leptons are not implemented for simulations. Additionally, there is a requirement of lepton $|\eta| < 2.4$, and there is a small difference between rapidity and pseudorapidity due to the mass of Z bosons, which results in the rise towards $|y| = 2.4$. Figure 4.27 and 4.28 show the Z boson p_T and rapidity respectively by requiring the jet multiplicity at least to be one. One can see the shift of the peak of Z boson p_T due to the balance with the jet which is applied with p_T cut. Good agreements between data and simulation have been seen.

In order to see the precise statistics that the data can reach, it is instructive to see the jet multiplicity distributions. Figure 4.29 and 4.30 show the exclusive jet multiplicity up to seven jets, and inclusive jet multiplicity respectively. The latter can gain a bit more statistics than the former in each bin. Certainly, the events with lower jet multiplicity are with higher cross section than those with larger jet multiplicity. One can see excellent agreement in detector level up to four jets, which are generated with matrix element using MADGRAPH and matched with parton shower. The discrepancy starting from the five jets case can not only result from the missing tree level computation while only relying on parton shower, but also the mismodelling of $t\bar{t}$ process, which contributes significantly in large jet multiplicity region. The number of events as a function of exclusive jet multiplicity in $\mu^+\mu^-$ and e^+e^- channels are summarized in table 4.10 and 4.11 respectively.

From figure 4.31 to figure 4.35, the distributions of transverse momentum of each jet with different requirements of jet multiplicity have been shown. Note that the starting point of p_T distribution is 20 GeV, which is below the p_T cut of jet i.e. 30 GeV as for the distributions of other observables. This is because one needs to take into account the event migration effect from the underflow bins to the region of p_T above 30 GeV, when implementing the detector effect correction for jet p_T later. Namely, this "buffer" region (20-30 GeV) holds the events that probably migrate to the region above threshold of 30 GeV. According to the control plots, the MC yields have underestimated the events by 10-20%⁷ generally in this region. There is a rising slope at high p_T region of the first jet, and the fluctuations can be seen in the same regions of other jets.

The distributions of absolute rapidity of each jet with different options of jet multiplicity are presented from figure 4.36 to figure 4.40. In general, good agreement has been seen except for the undershooting of MC with 10-20% at the endcap region. There is also apparent fluctuation with higher jet multiplicity.

Shifted peak can be observed in scalar sum of jets transverse momenta. From figure 4.41 to figure 4.45, note that the lowest H_T values are changing with respect to different jet multiplicity requirements. Similarly, the MC yields underestimate the data at low H_T region, where one will see later the systematic uncertainty is large.

In contrast to the overall agreement between data and MC in the reconstruction level, large discrepancy has been seen in figure 4.46 to figure 4.53, which are the event yields as functions of rapidity

⁷It has been inspected that this issue is not originated from a potential pileup misestimation by checking the yields of reconstructed vertices.

correlation observables. All of them have significant discrepancy between data and simulation. Regarding the average sum in rapidity, there is about 25% disagreement for Z boson and the first jet, while about 15% for Z boson and the second jet. The discrepancy of rapidity sum rises to about 30% for Z boson and the di-jets. The rapidity sum has about 20% for the two leading jets.

With respect to the average difference in rapidity, the shape change between data and simulation is approximately 50% for Z boson and the first jet, compared to 40% for Z boson and the second jet. Both of them contribute to the discrepancy for Z boson and di-jets. In parallel, the shape difference is about 30% for the two leading jets. These obvious discrepancies seen in the detector level will propagate to the comparison between measurement and prediction on the differential cross section as a function of rapidity correlation observables. The relevant results will be shown later.

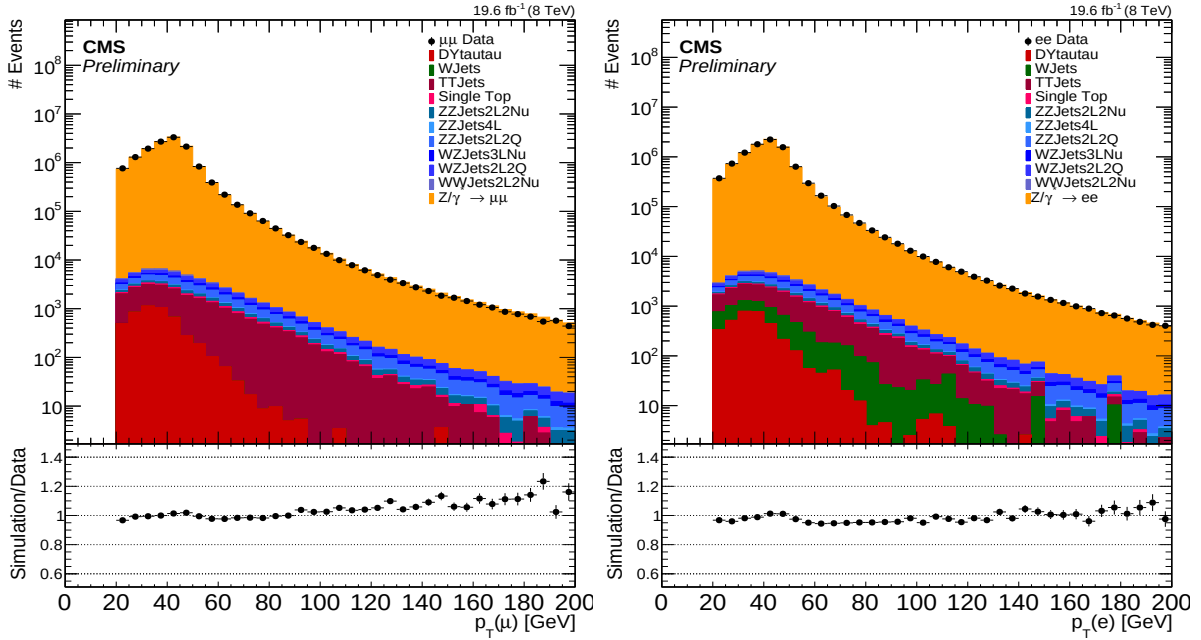


Figure 4.20: The data and simulation comparison of lepton p_T with $N_{jets} \geq 0$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

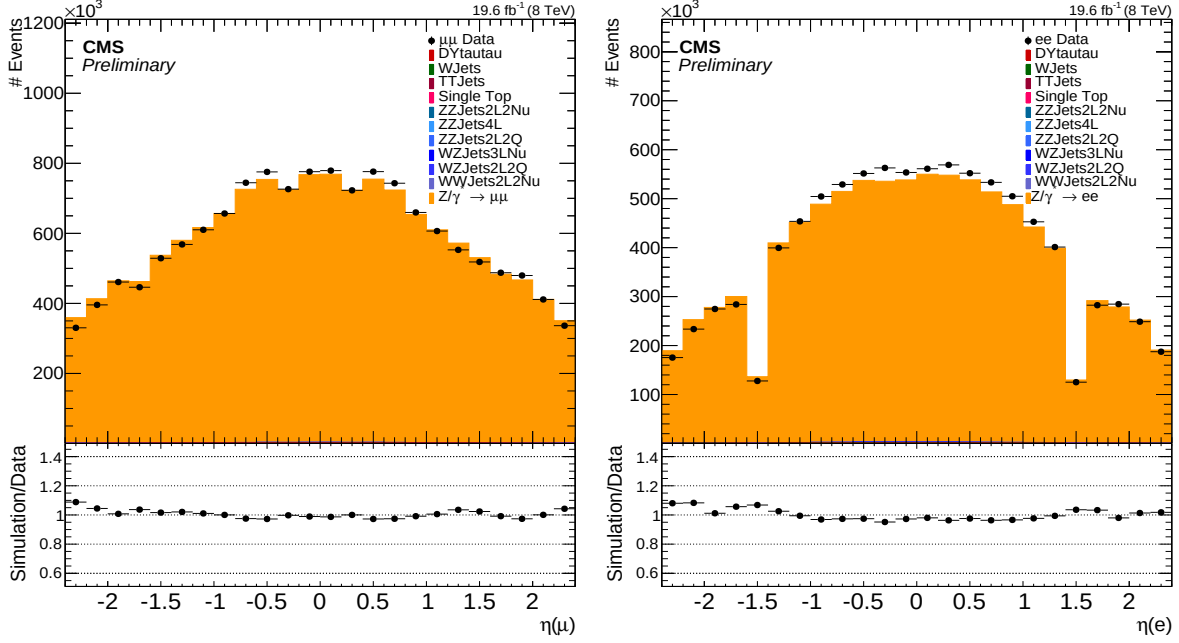


Figure 4.21: The data and simulation comparison of lepton η with $N_{jets} \geq 0$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

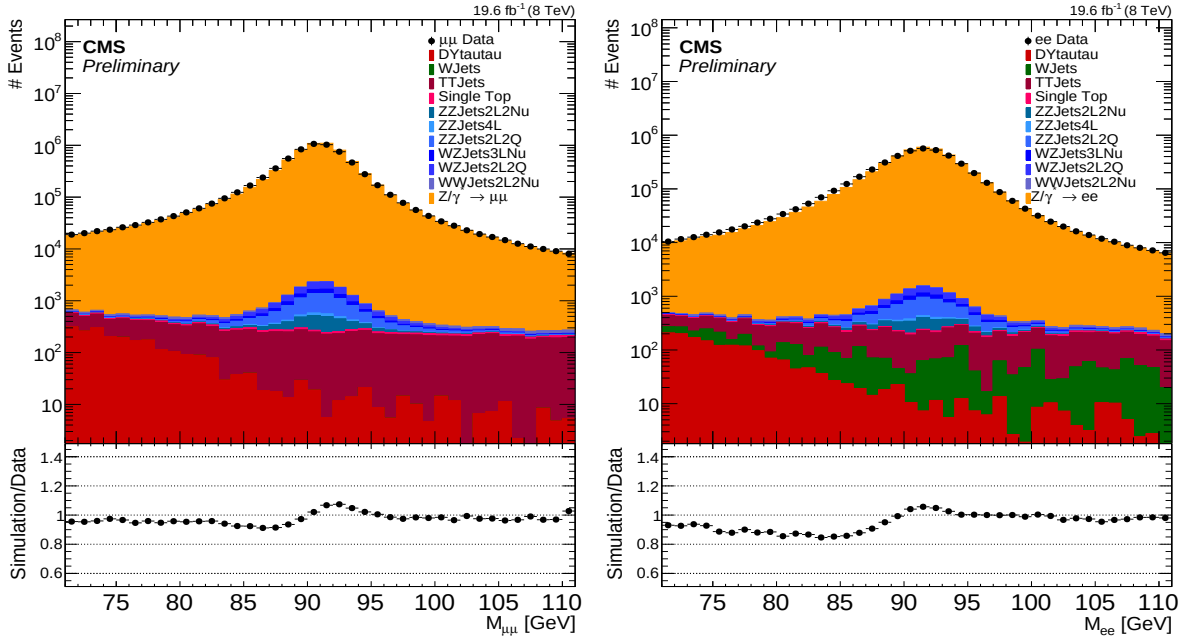


Figure 4.22: The data and simulation comparison of invariant mass of lepton pair with $N_{jets} \geq 0$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

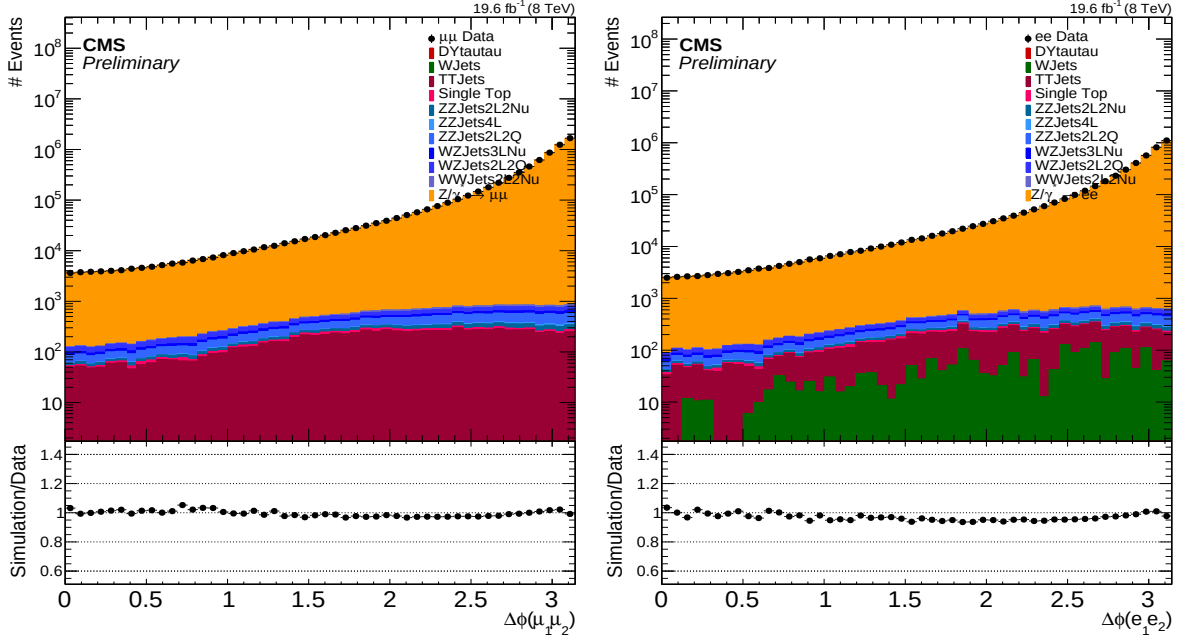


Figure 4.23: The data and simulation comparison of azimuthal distance between two leptons with $N_{jets} \geq 0$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

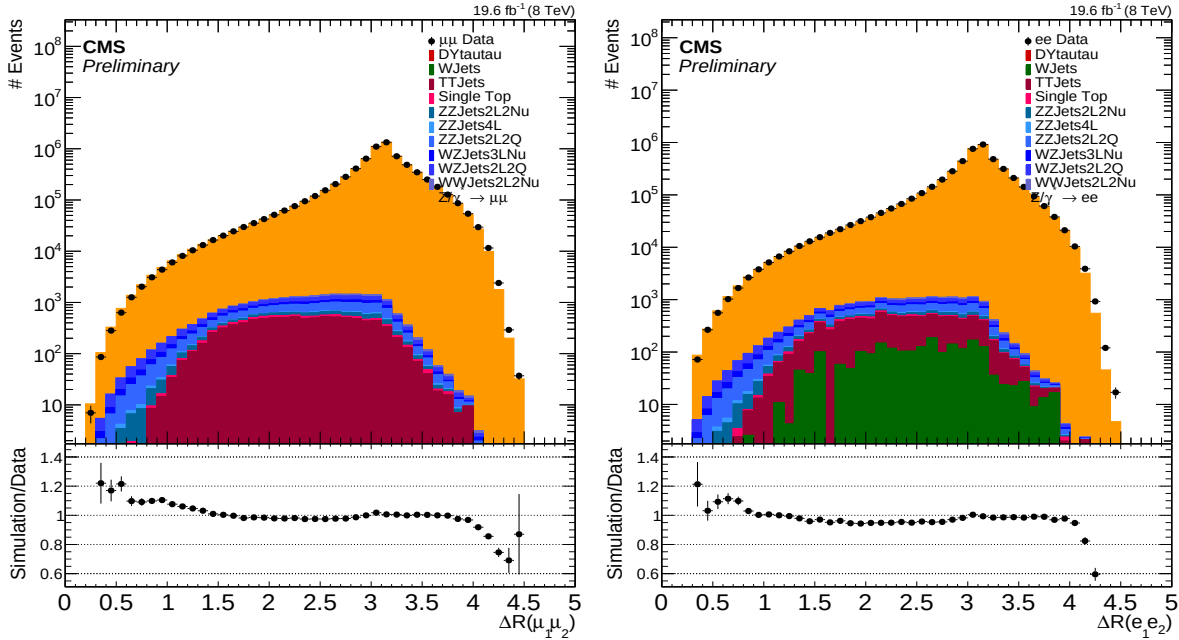


Figure 4.24: The data and simulation comparison of separation of two leptons with $N_{jets} \geq 0$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

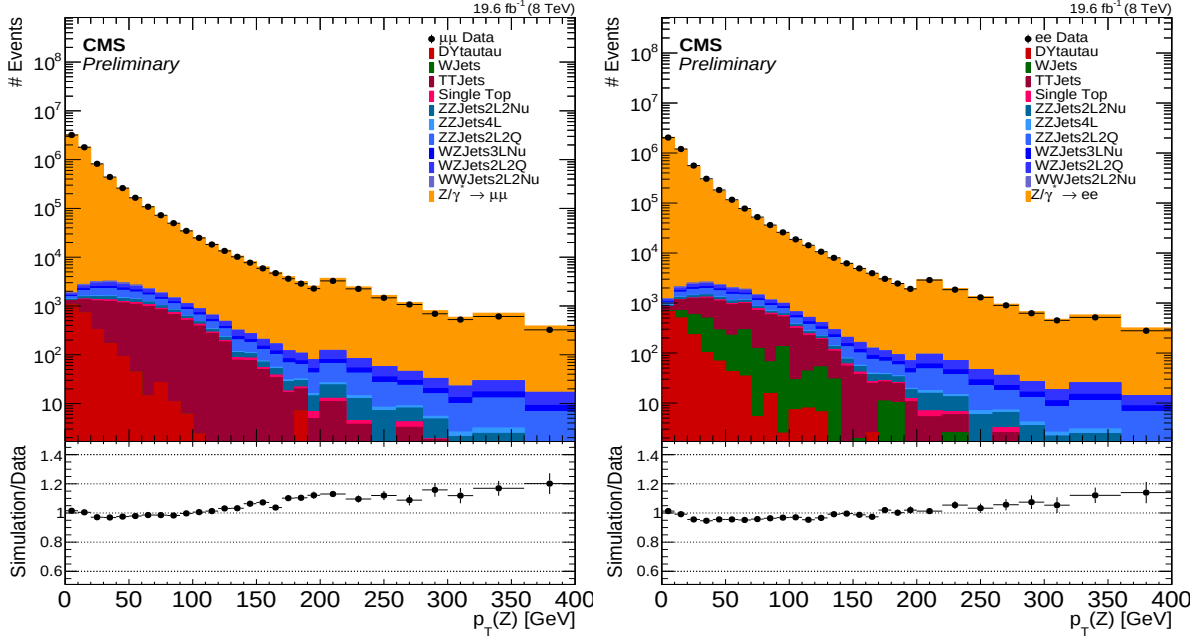


Figure 4.25: The data and simulation comparison of Z boson p_T with $N_{jets} \geq 0$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

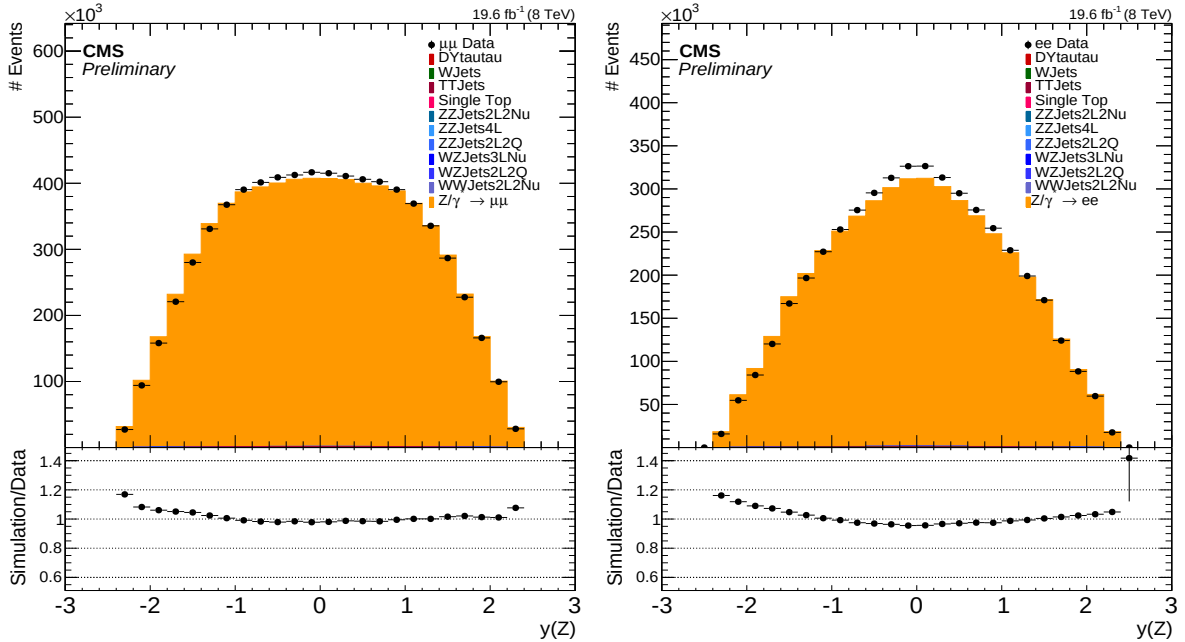


Figure 4.26: The data and simulation comparison of Z boson rapidity with $N_{jets} \geq 0$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

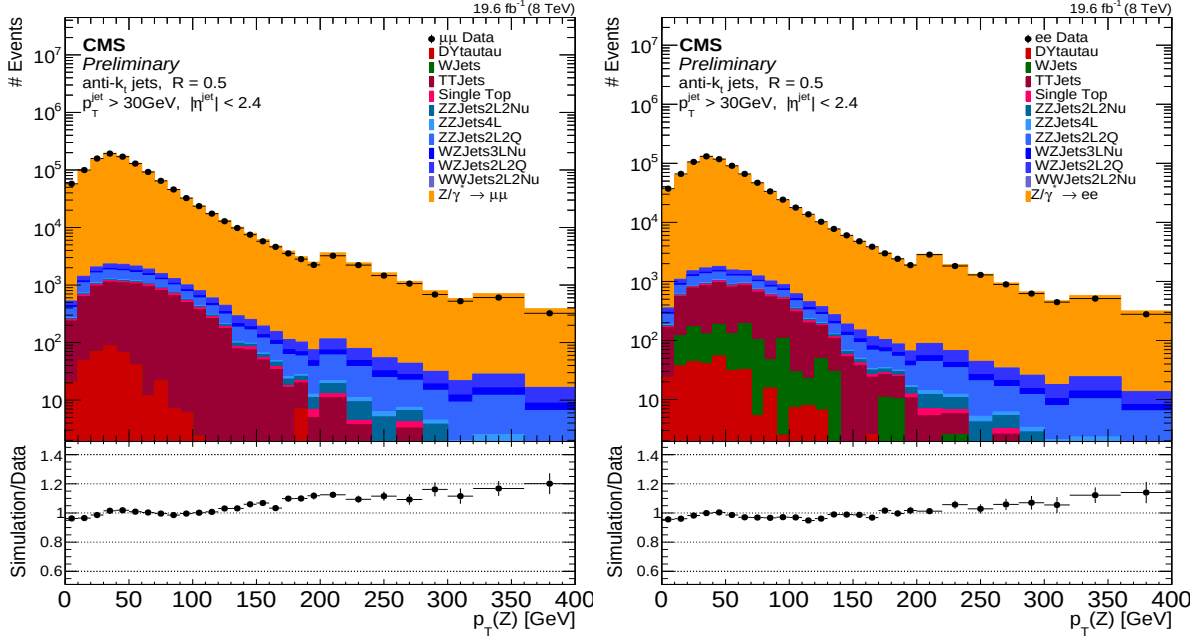


Figure 4.27: The data and simulation comparison of Z boson p_T with $N_{jets} \geq 1$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

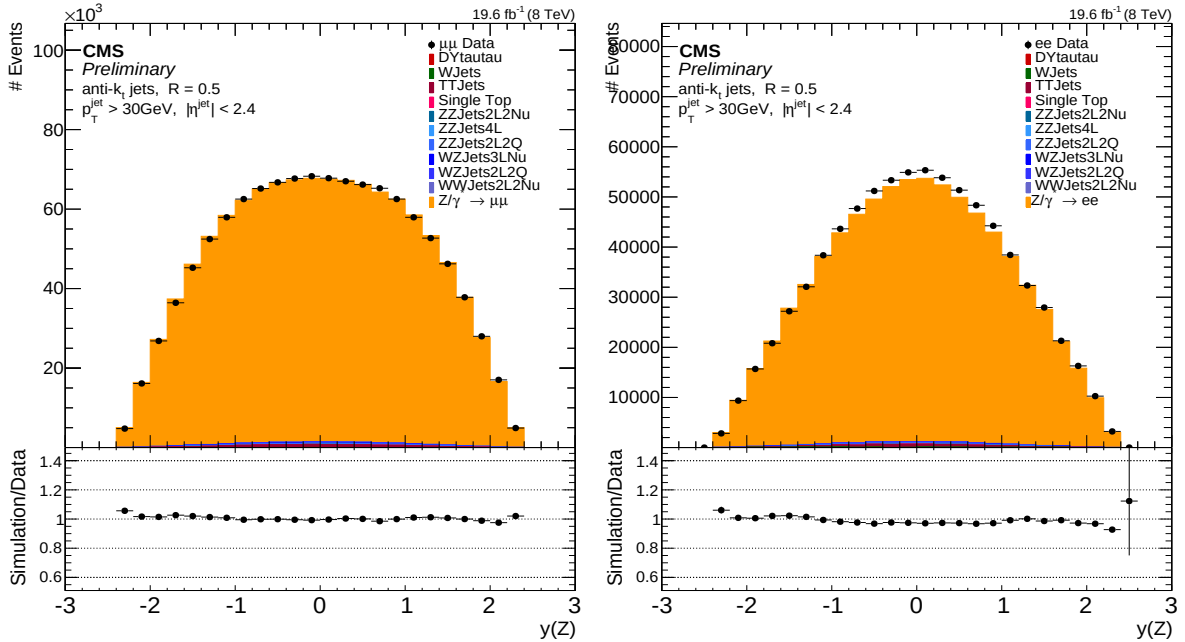


Figure 4.28: The data and simulation comparison of Z boson rapidity with $N_{jets} \geq 1$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

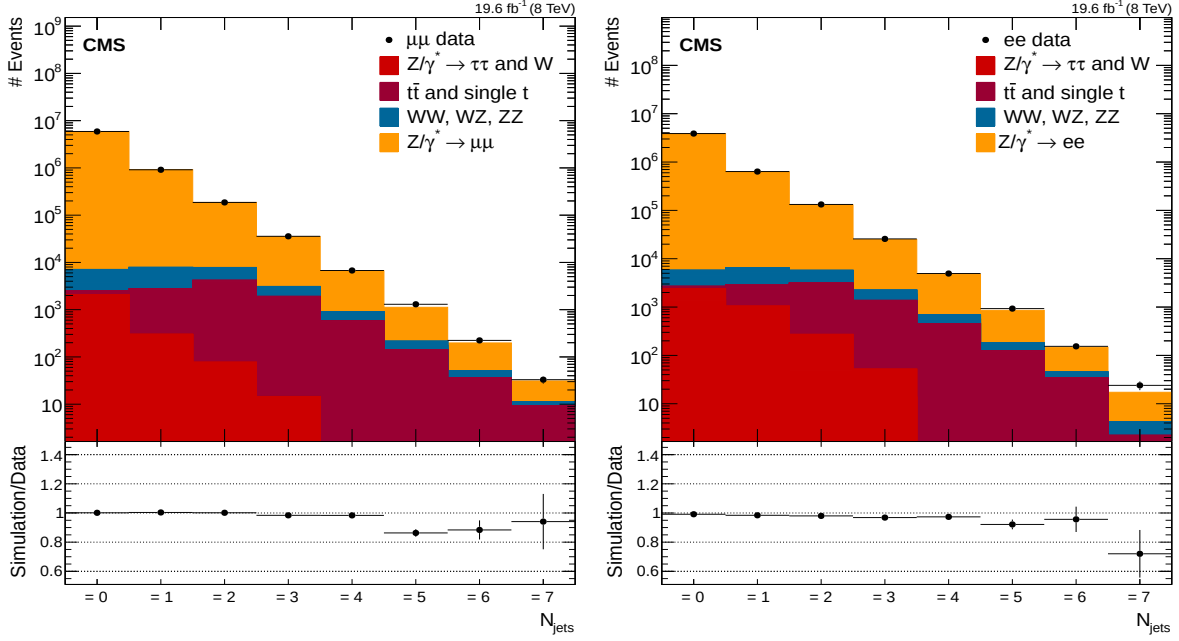


Figure 4.29: The data and simulation comparison of exclusive jet multiplicity for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

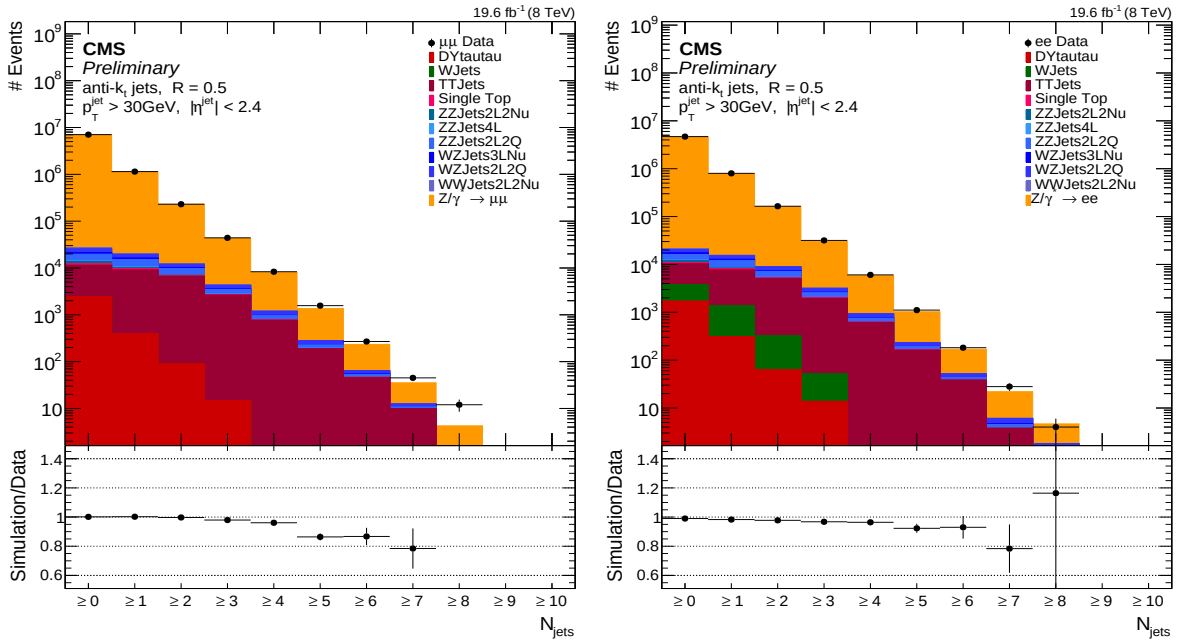


Figure 4.30: The data and simulation comparison of inclusive jet multiplicity for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

	$N_{jets} = 0$	$N_{jets} = 1$	$N_{jets} = 2$	$N_{jets} = 3$	$N_{jets} = 4$	$N_{jets} = 5$	$N_{jets} = 6$	$N_{jets} = 7$
WJets	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DYTauTau	2075.6	300.8	77.0	14.6	0.0	0.0	0.0	0.0
ZZJets4L	130.7	150.4	107.6	29.4	7.0	1.5	0.3	0.0
Single Top	127.2	410.4	196.8	60.1	9.7	0.9	1.3	0.0
ZZJets2L2Nu	1185.5	332.1	89.7	19.9	3.8	0.6	0.2	0.0
ZZJets2L2Q	719.6	2088.2	1803.9	511.2	130.7	26.4	5.7	0.7
WWJets2L2Nu	1112.0	362.0	107.3	22.6	4.5	0.4	0.0	0.0
WZJets2L2Q	712.2	1387.1	1193.0	522.8	167.4	39.2	8.2	1.3
WZJets3LNU	727.2	774.1	315.8	81.8	18.9	3.2	0.8	0.1
TTJets	312.8	2071.4	3933.1	1840.2	570.9	142.7	35.0	9.2
DYJets	5903239.3	908971.9	178530.9	32043.0	5727.1	908.3	146.5	19.8
TOTAL(MC)	5910342.1	916848.4	186354.9	35145.6	6640.0	1123.1	198.0	31.0
Data	5902750	913503	186087	35718	6751	1301	224	33
Ratio	1.0013	1.0037	1.0014	0.9840	0.9836	0.8633	0.8838	0.9409

Table 4.10: Number of events in data and MC samples for exclusive jet multiplicity in muon channel after full selection.

	$N_{jets} = 0$	$N_{jets} = 1$	$N_{jets} = 2$	$N_{jets} = 3$	$N_{jets} = 4$	$N_{jets} = 5$	$N_{jets} = 6$	$N_{jets} = 7$
WJets	1016.5	801.9	222.7	38.4	0.0	0.0	0.0	0.0
DYTauTau	1392.2	247.9	49.6	13.8	0.0	0.0	0.0	0.0
ZZJets4L	86.4	110.1	81.2	22.5	5.5	1.1	0.2	0.0
Single Top	67.7	301.3	131.0	38.0	9.5	3.2	0.0	0.0
ZZJets2L2Nu	810.3	239.0	65.0	13.8	2.9	0.5	0.0	0.0
ZZJets2L2Q	486.0	1488.7	1284.8	371.9	93.1	19.9	3.9	0.7
WWJets2L2Nu	745.8	246.5	77.0	16.5	3.2	0.9	0.3	0.1
WZJets2L2Q	480.4	977.0	854.2	381.1	123.3	30.0	6.5	1.2
WZJets3LNU	490.6	567.6	237.5	62.3	13.7	2.9	0.4	0.1
TTJets	222.7	1531.3	2772.8	1286.7	447.0	122.2	34.3	2.2
DYJets	3843135.3	619067.1	124236.0	22649.6	4106.4	675.3	101.8	13.0
TOTAL(MC)	3848934.0	625578.3	130011.9	24894.4	4804.8	856.1	147.3	17.3
Data	3882357	635598	132659	25704	4936	929	154	24
Ratio	0.9914	0.9842	0.9800	0.9685	0.9734	0.9215	0.9565	0.7202

Table 4.11: Number of events in data and MC samples for exclusive jet multiplicity in electron channel after full selection.

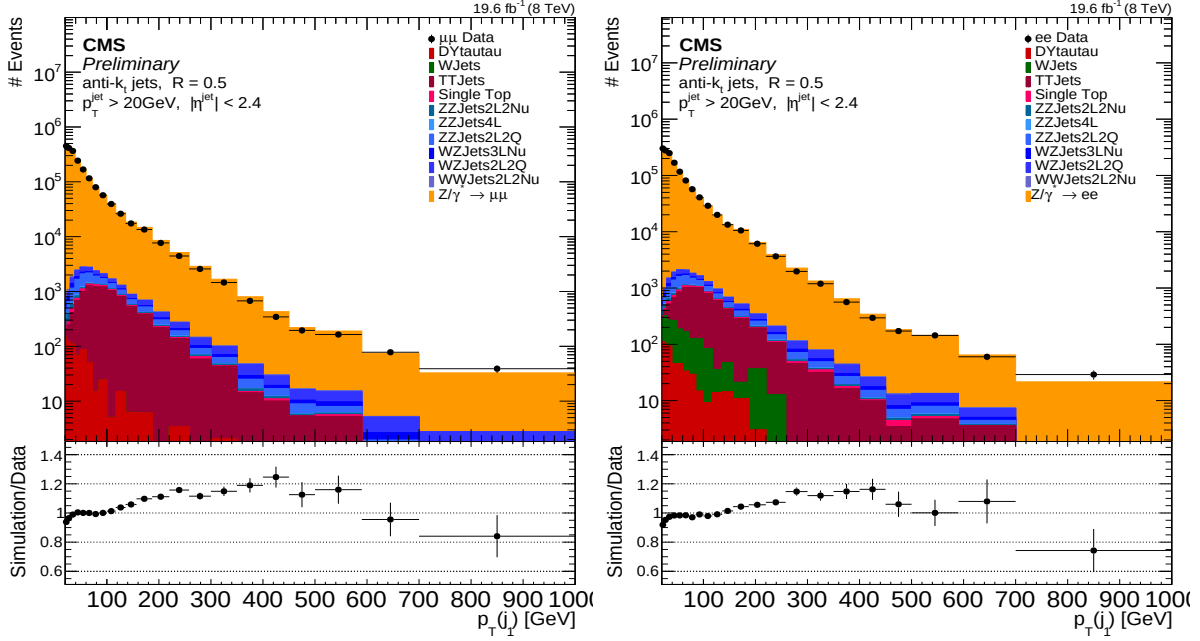


Figure 4.31: The data and simulation comparison of first jet p_T with $N_{jets} \geq 1$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

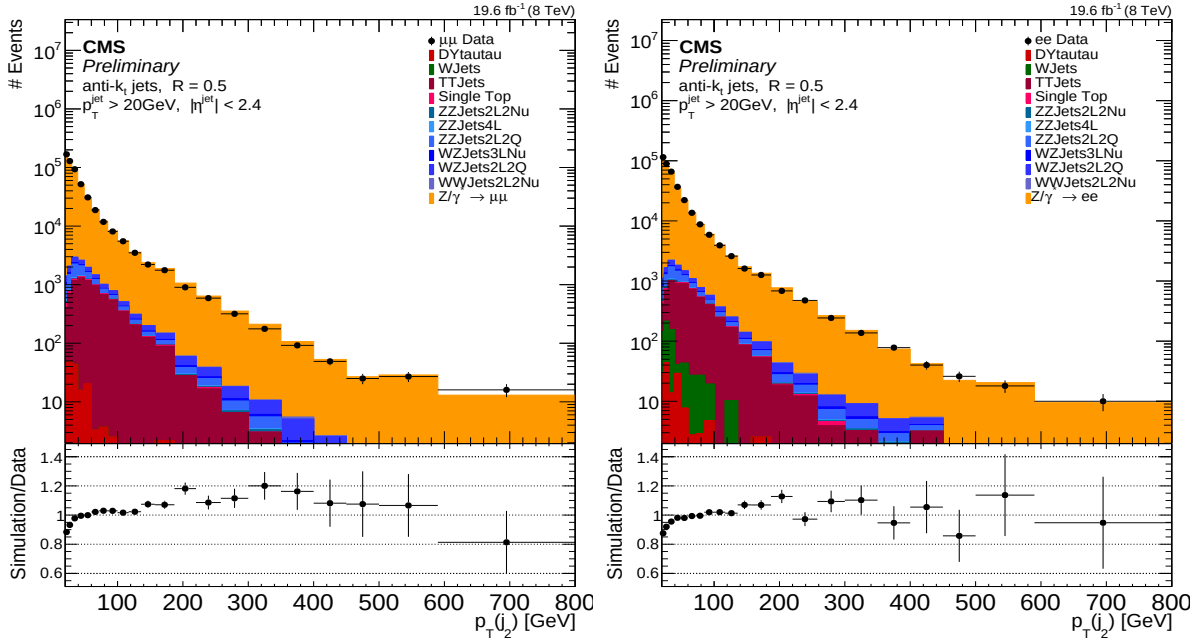


Figure 4.32: The data and simulation comparison of second jet p_T with $N_{jets} \geq 2$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

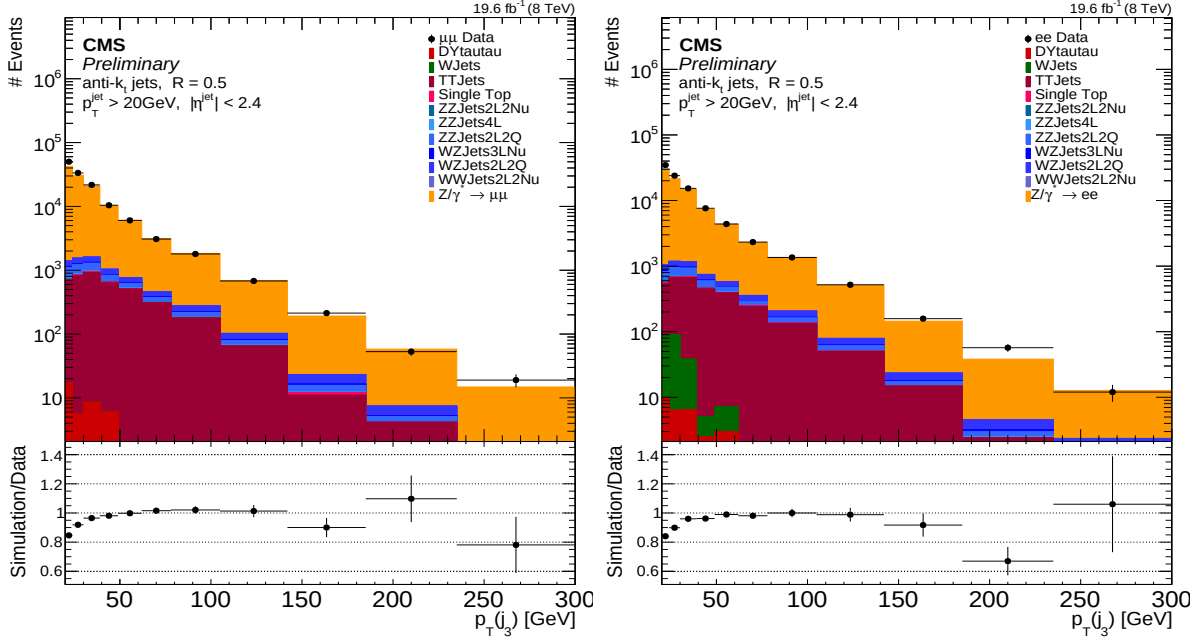


Figure 4.33: The data and simulation comparison of third jet p_T with $N_{jets} \geq 3$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

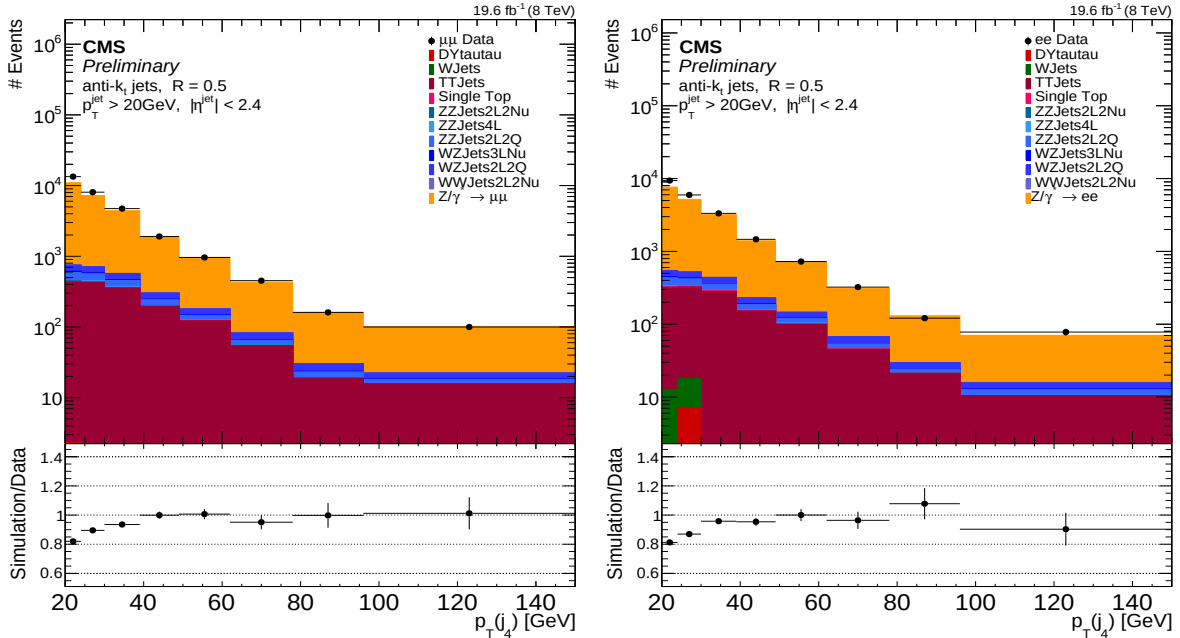


Figure 4.34: The data and simulation comparison of fourth jet p_T with $N_{jets} \geq 4$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

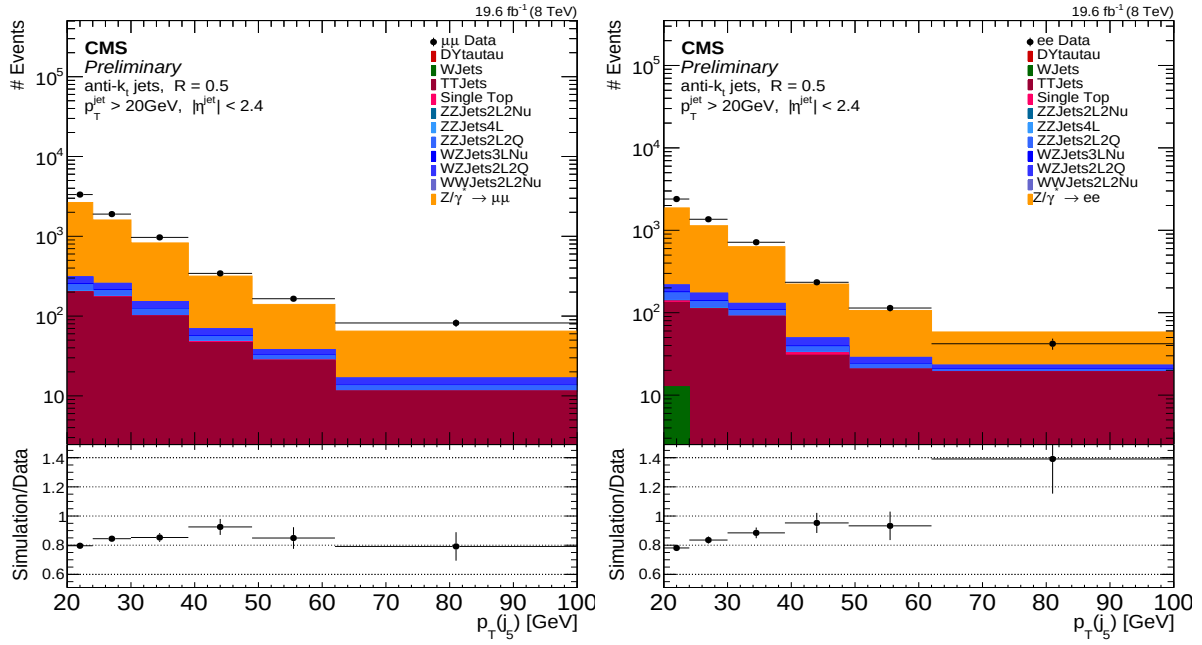


Figure 4.35: The data and simulation comparison of fifth jet p_T with $N_{jets} \geq 5$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

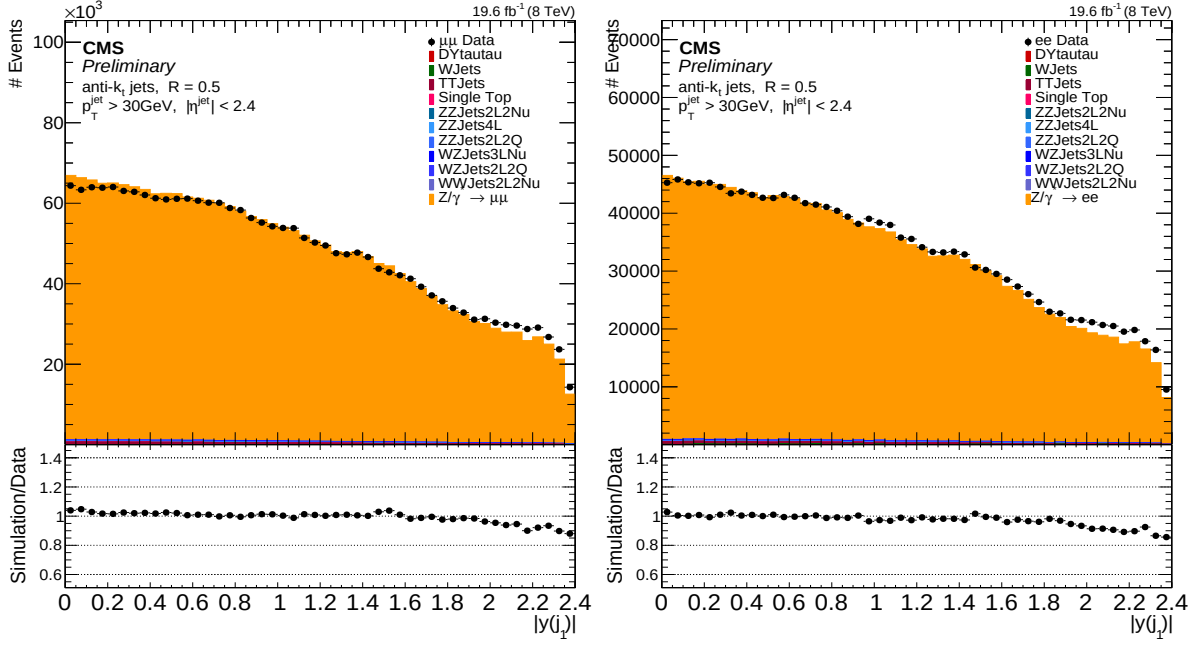


Figure 4.36: The data and simulation comparison of first jet absolute rapidity with $N_{jets} \geq 1$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

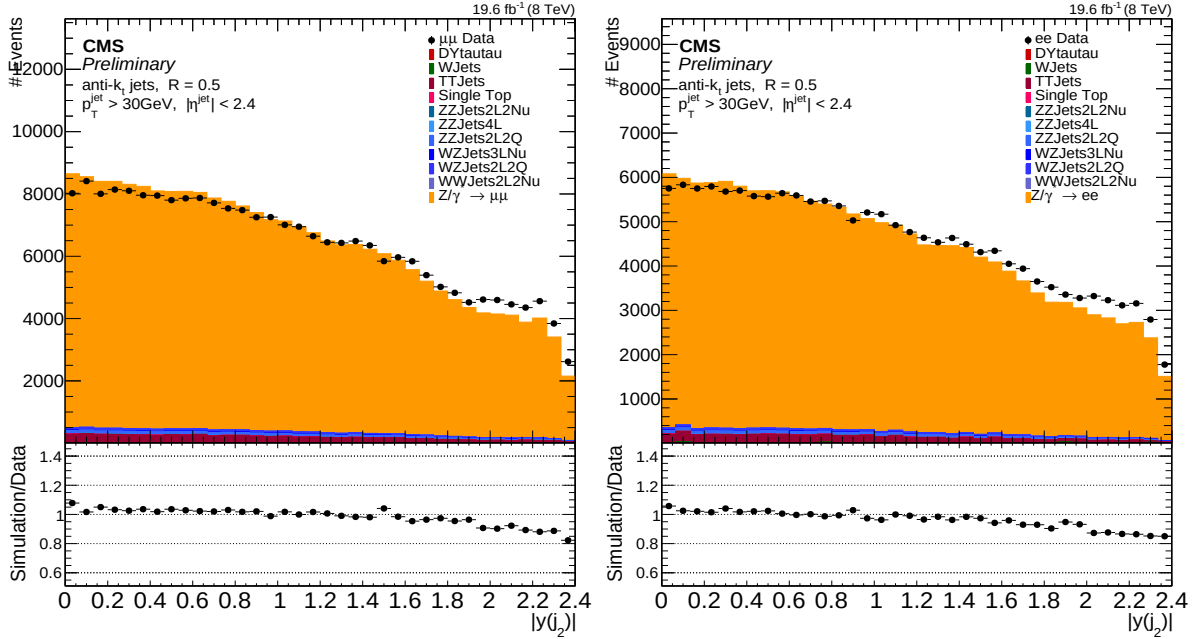


Figure 4.37: The data and simulation comparison of second jet absolute rapidity with $N_{jets} \geq 2$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

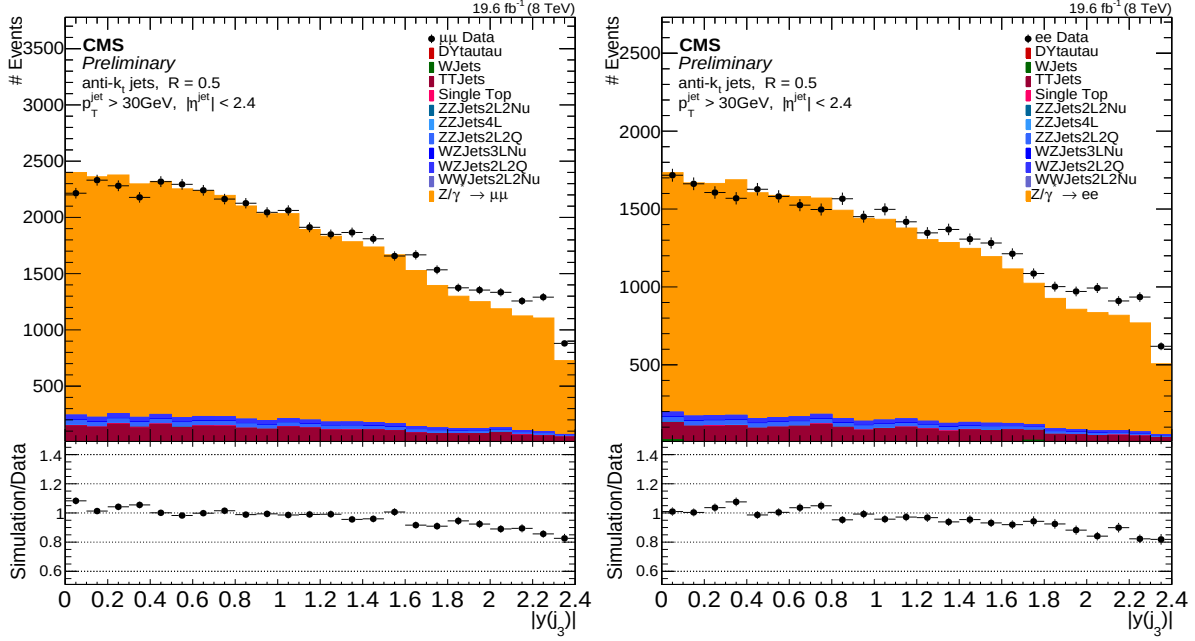


Figure 4.38: The data and simulation comparison of third jet absolute rapidity with $N_{jets} \geq 3$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

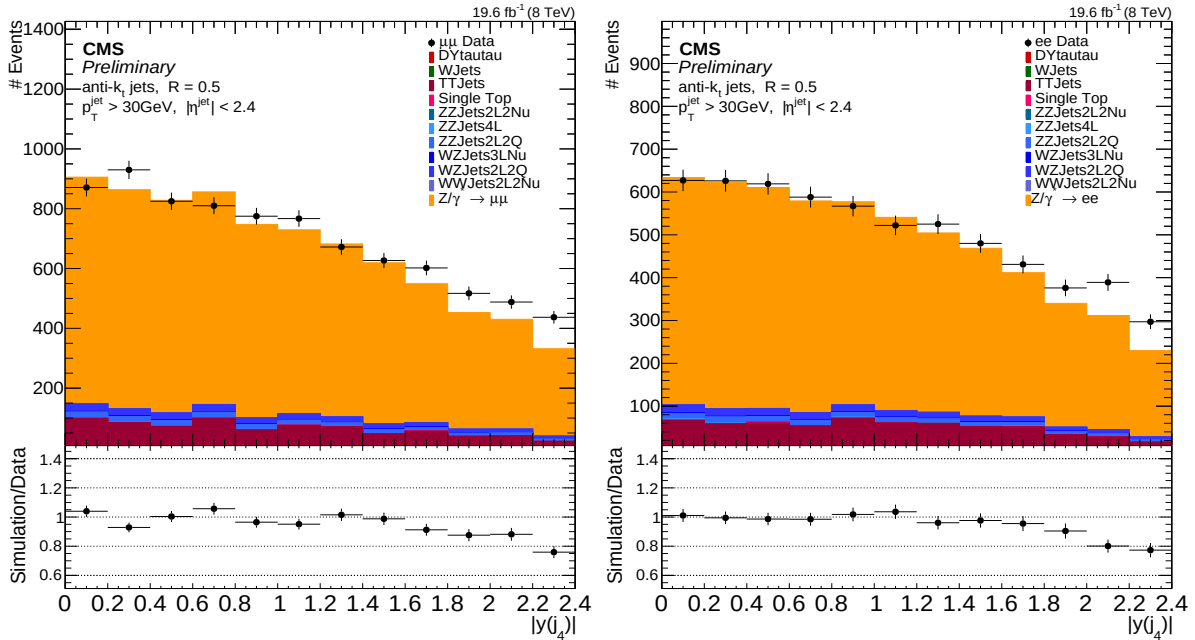


Figure 4.39: The data and simulation comparison of fourth jet absolute rapidity with $N_{jets} \geq 4$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

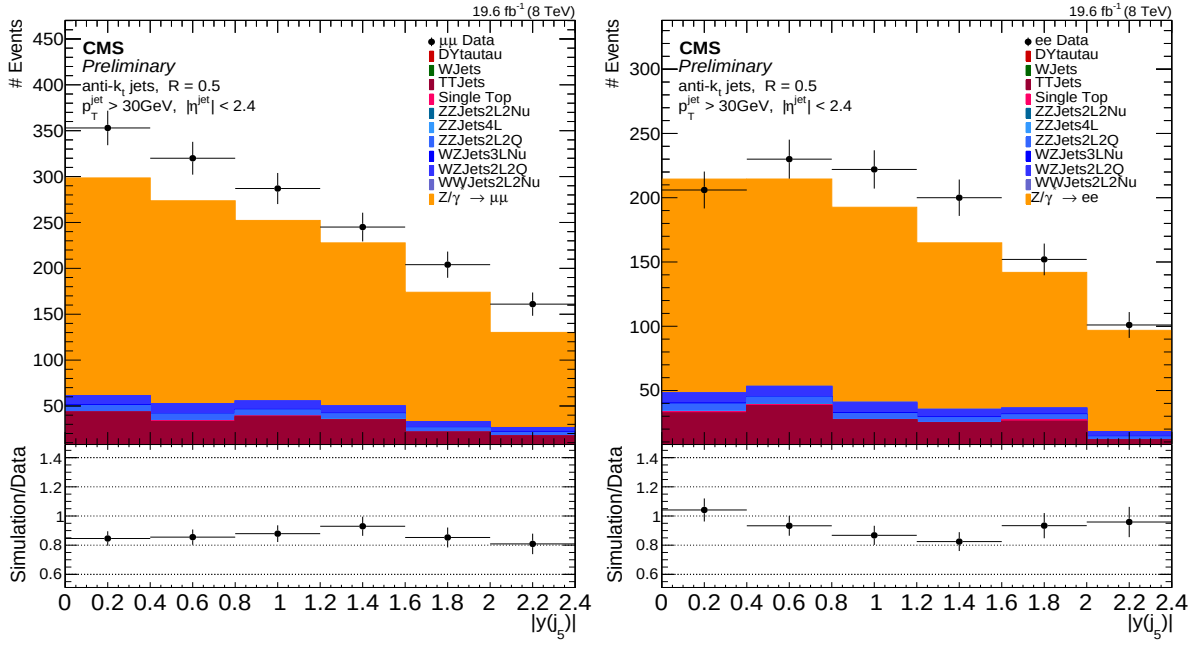


Figure 4.40: The data and simulation comparison of fifth jet absolute rapidity with $N_{jets} \geq 5$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

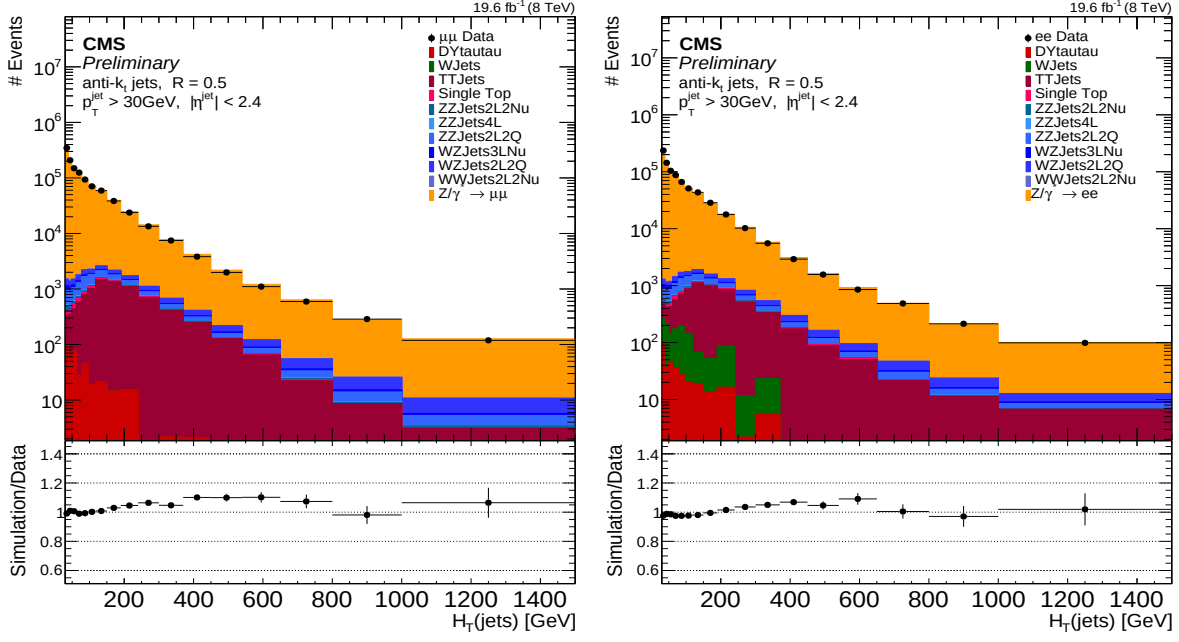


Figure 4.41: The data and simulation comparison of scalar sum of jets p_T with $N_{jets} \geq 1$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

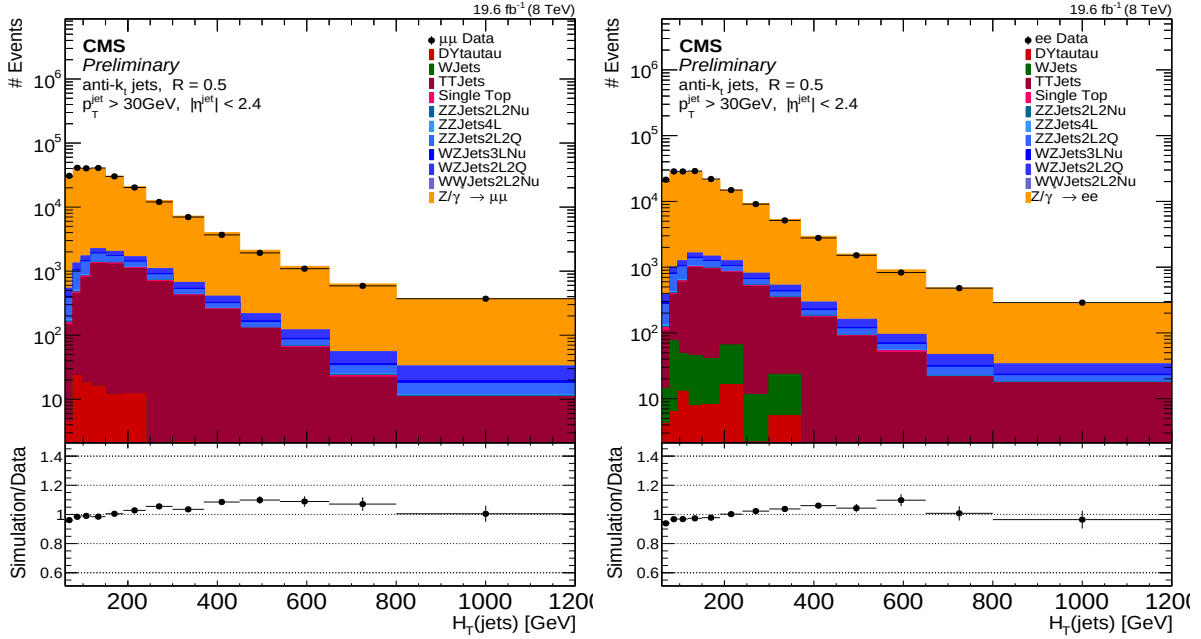


Figure 4.42: The data and simulation comparison of scalar sum of jets p_T with $N_{jets} \geq 2$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

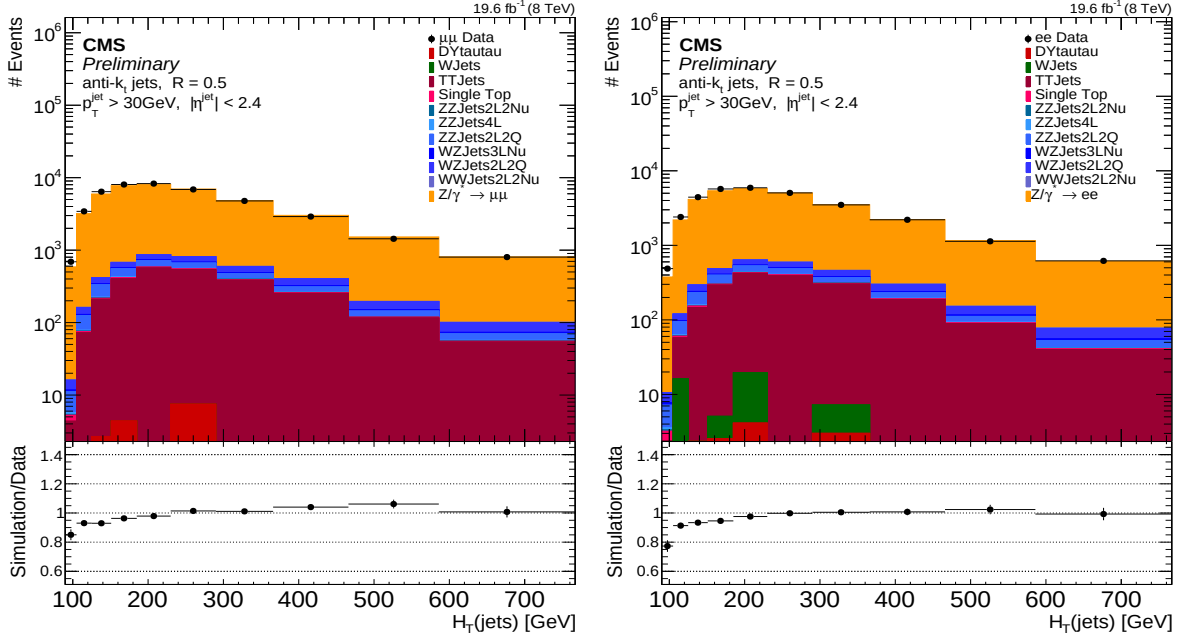


Figure 4.43: The data and simulation comparison of scalar sum of jets p_T with $N_{jets} \geq 3$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

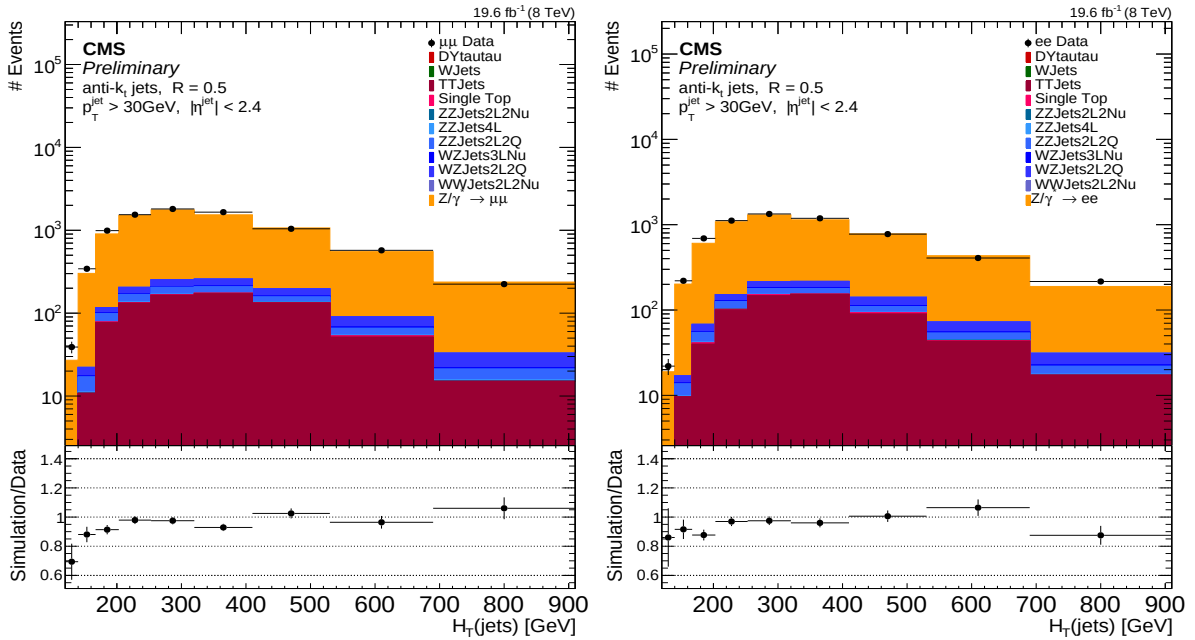


Figure 4.44: The data and simulation comparison of scalar sum of jets p_T with $N_{jets} \geq 4$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

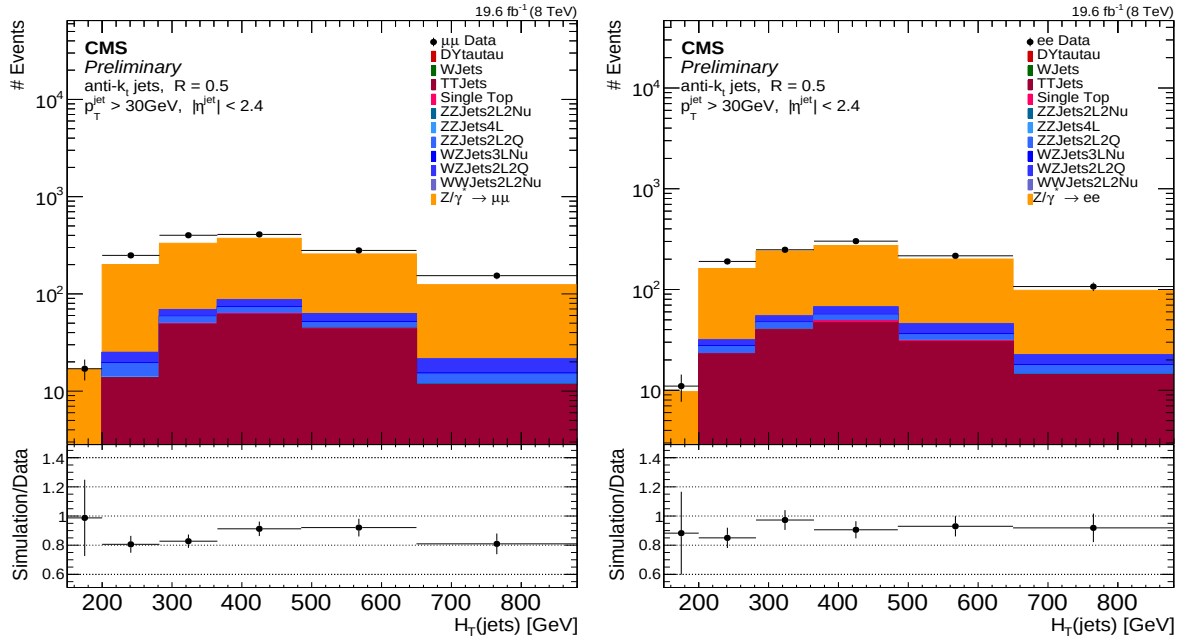


Figure 4.45: The data and simulation comparison of scalar sum of jets p_T with $N_{jets} \geq 5$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

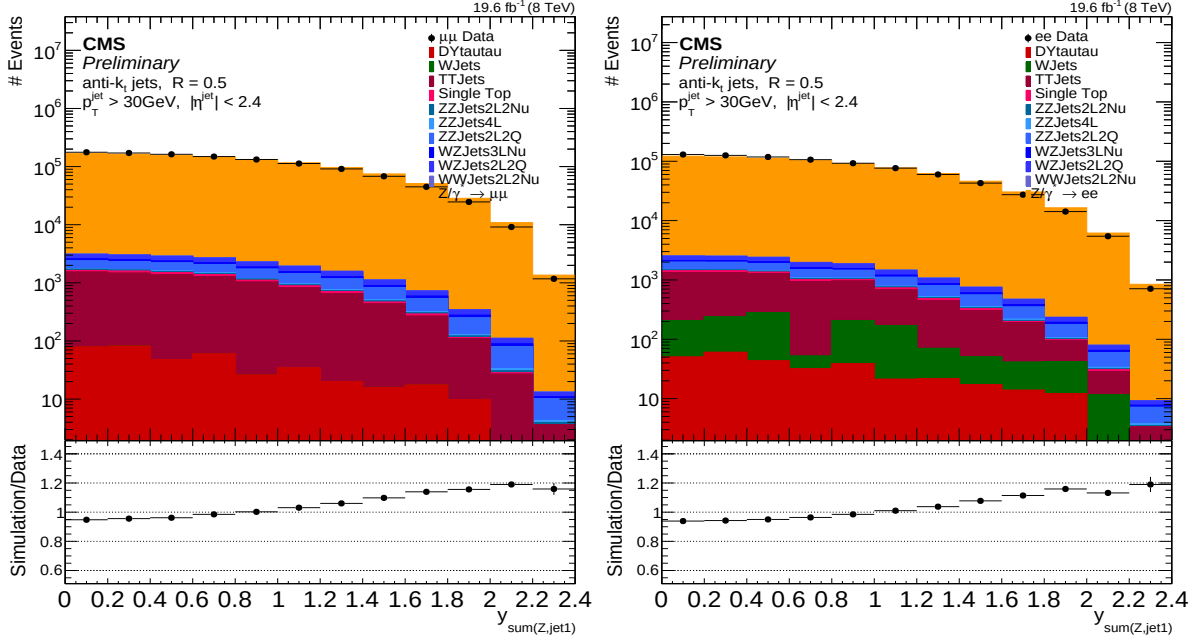


Figure 4.46: The data and simulation comparison in rapidity sum of Z boson and the first jet with $N_{jets} \geq 1$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

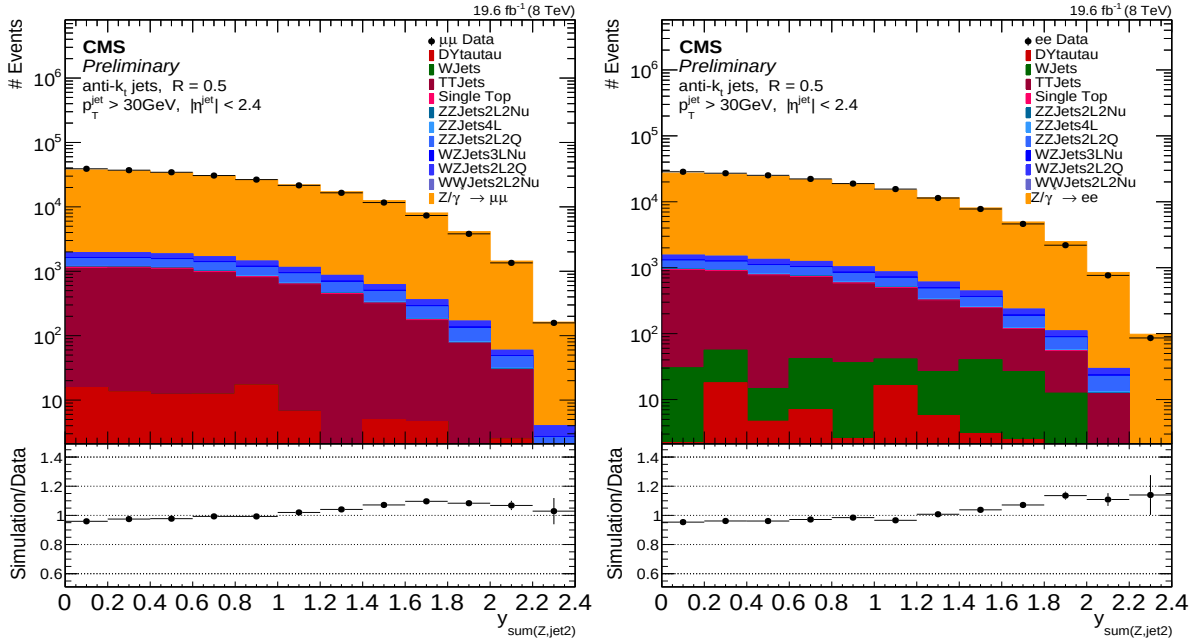


Figure 4.47: The data and simulation comparison in rapidity sum of Z boson and the second jet with $N_{jets} \geq 2$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

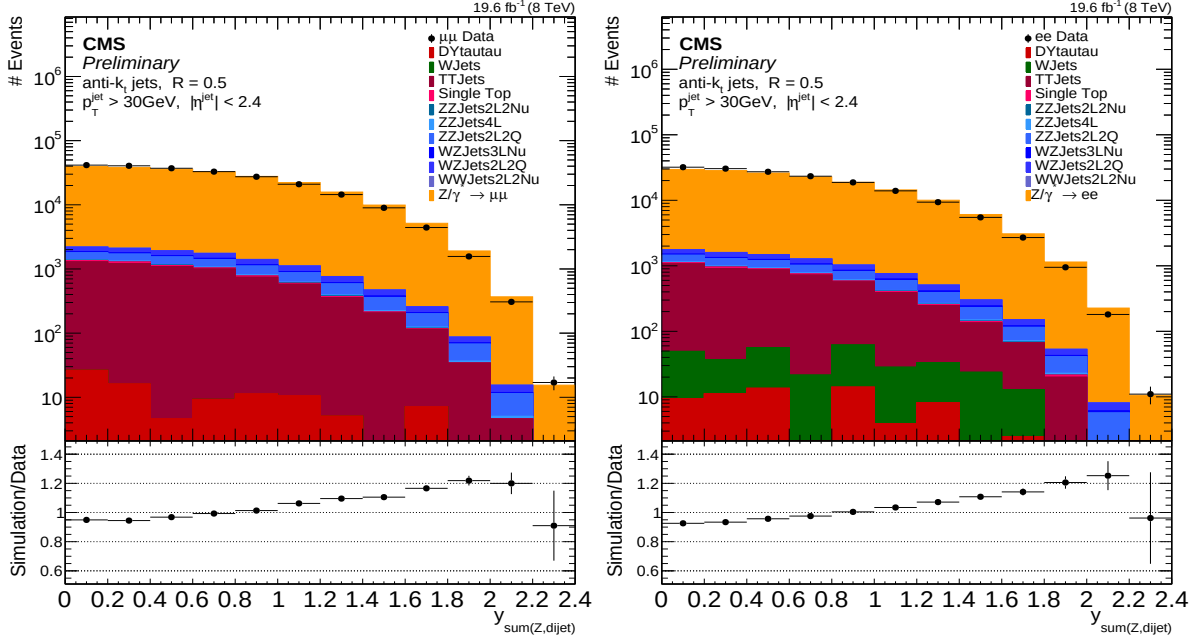


Figure 4.48: The data and simulation comparison in rapidity sum of Z boson and the first two jets with $N_{jets} \geq 2$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

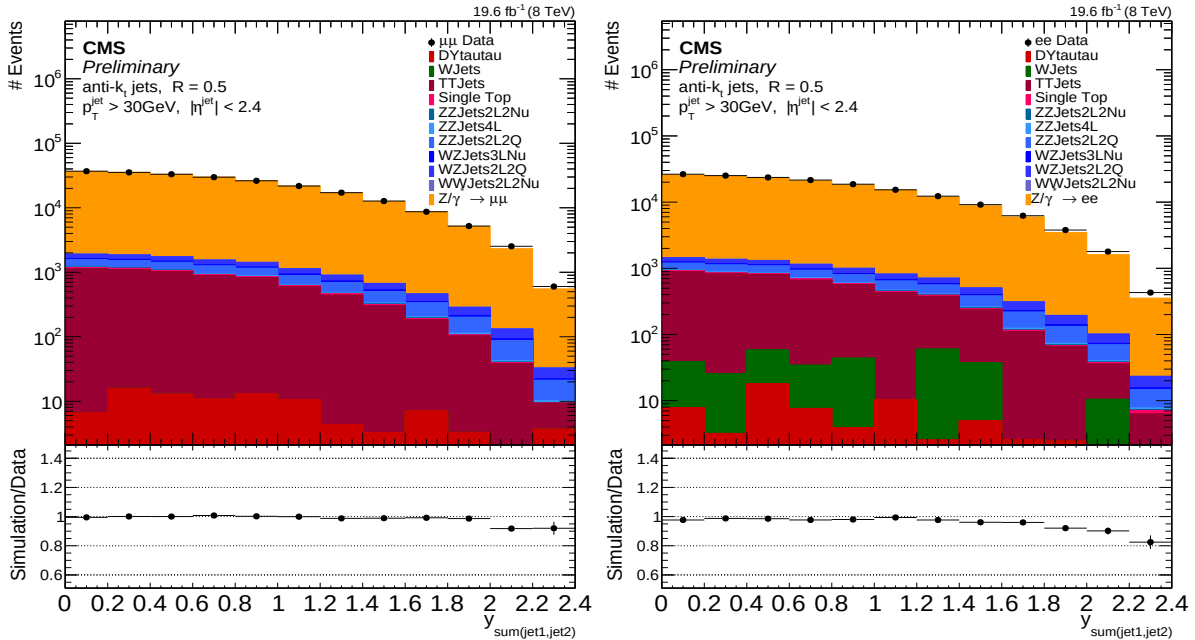


Figure 4.49: The data and simulation comparison in rapidity sum of the first jet and the second jet with $N_{jets} \geq 2$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

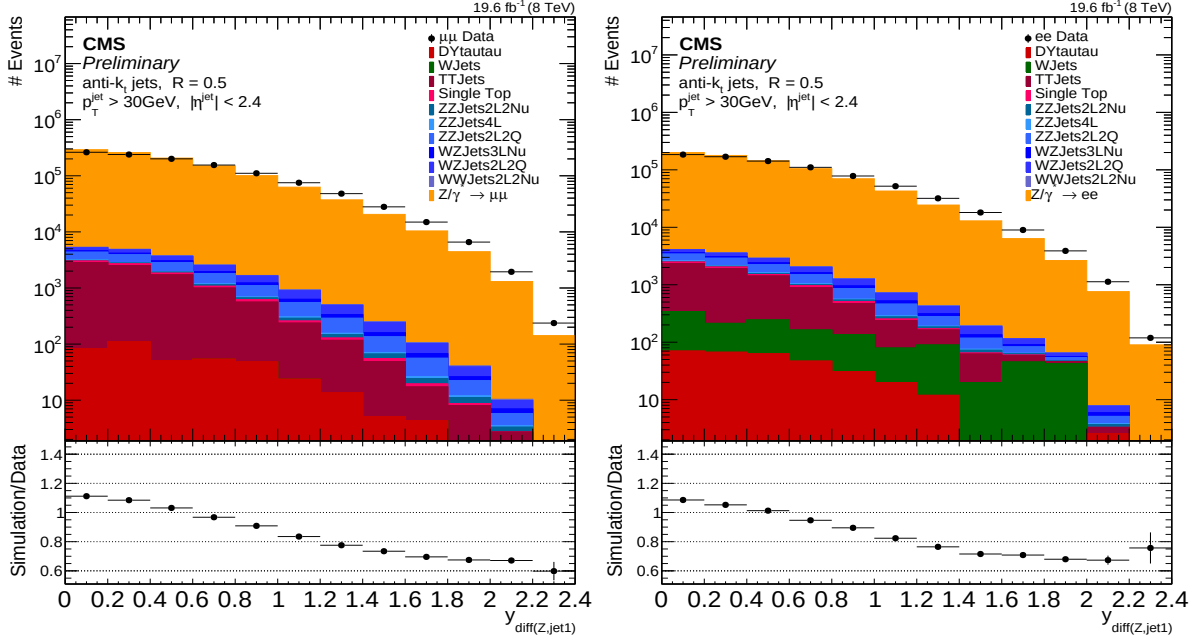


Figure 4.50: The data and simulation comparison in rapidity difference of Z boson and the first jet with $N_{jets} \geq 1$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

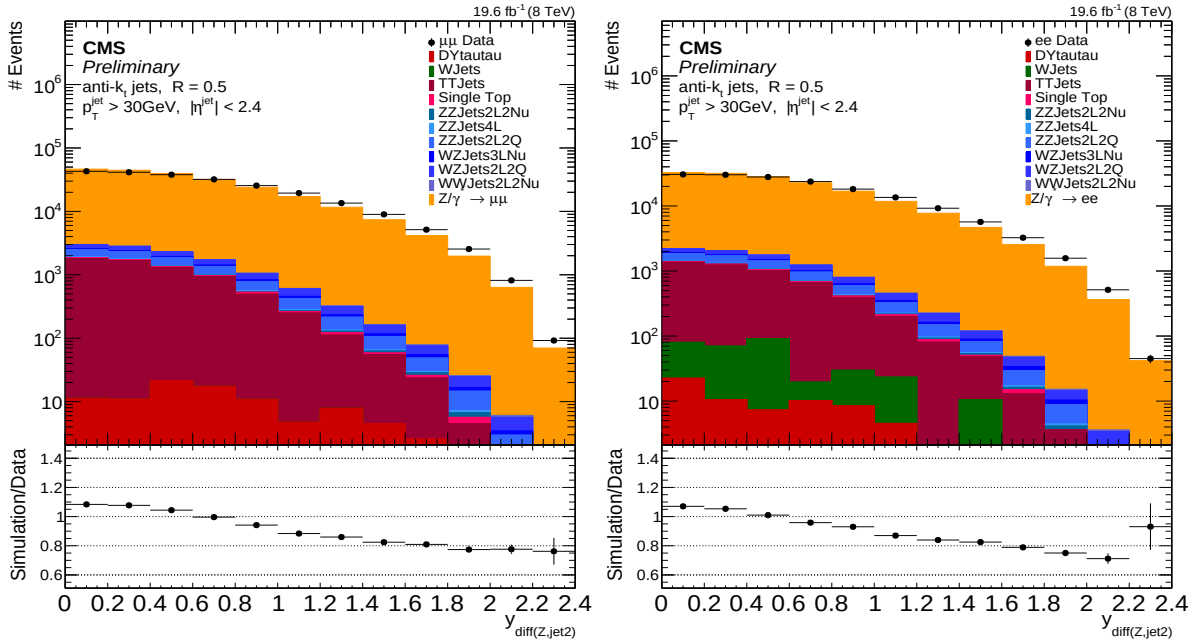


Figure 4.51: The data and simulation comparison in rapidity difference of Z boson and the second jet with $N_{jets} \geq 2$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

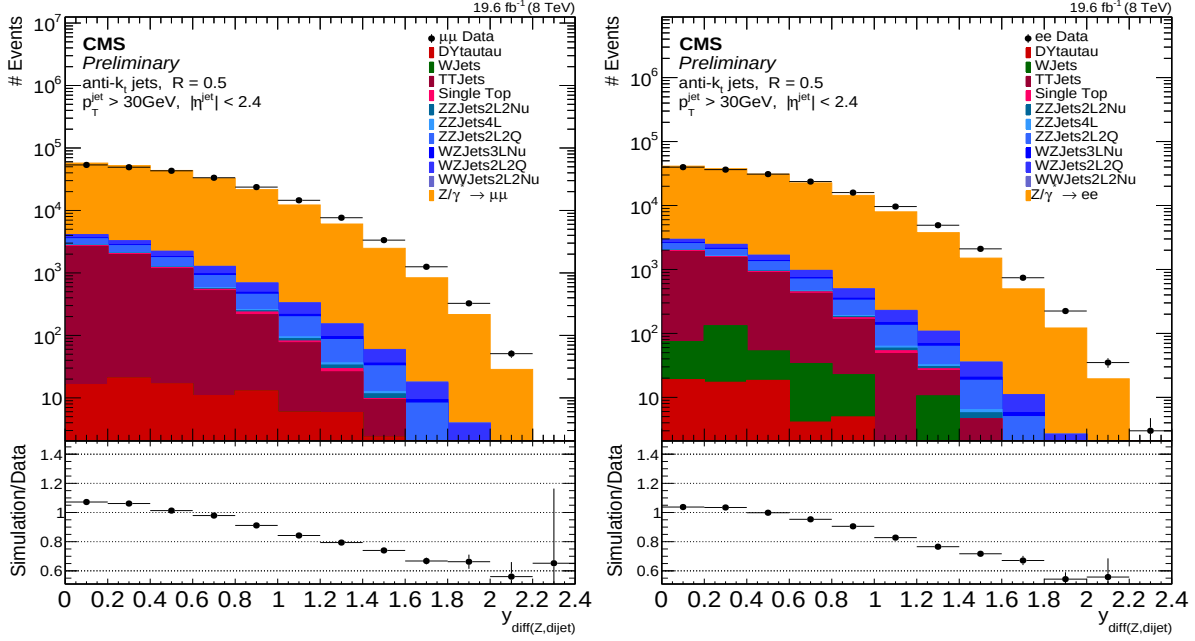


Figure 4.52: The data and simulation comparison in rapidity difference of Z boson and the first two jets with $N_{jets} \geq 2$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

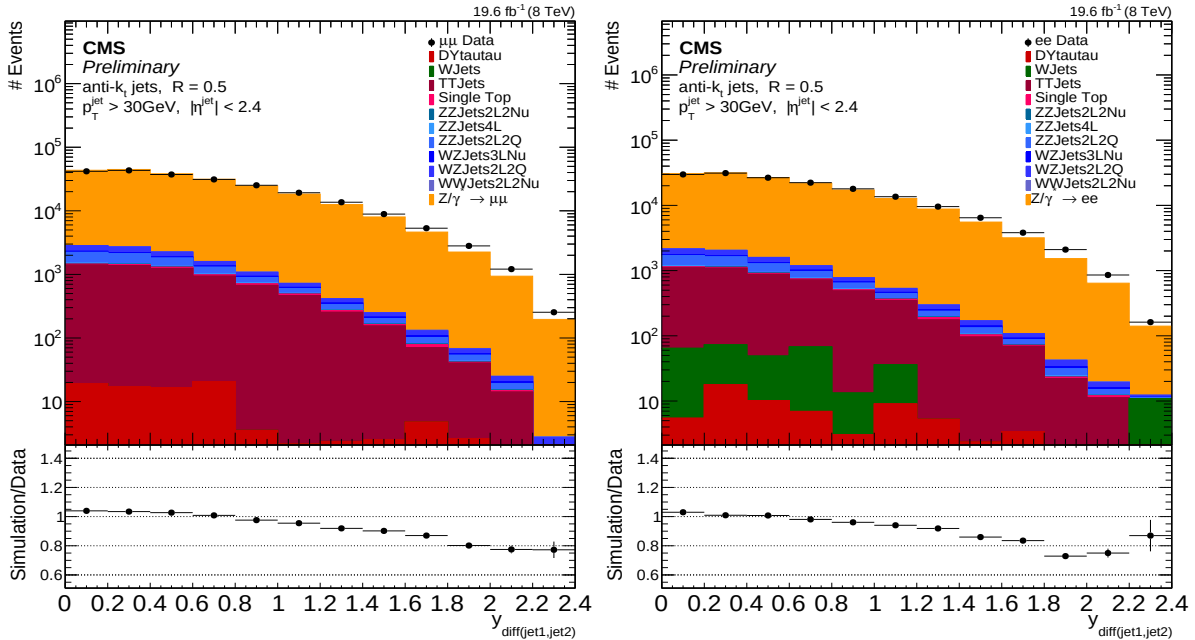


Figure 4.53: The data and simulation comparison in rapidity difference of the first jet and the second jet with $N_{jets} \geq 2$ for $\mu^+\mu^-$ (left) and e^+e^- (right) channels at reconstruction level.

4.4.2 Kinematic variable distributions at 13 TeV

The data at 13 TeV analyzed here are the first data even collected at such a high energy. In this section we compare the data and the superposition of MC yields in reconstruction level.

Figure 4.54 shows the muon p_T and η distributions at detector level, from which one can see a good agreement between data and MC yields. The simulation at large muon p_T region has large statistical fluctuation. Note that some of the backgrounds from simulation are summed up. To be specific, the "VV" is di-boson process where the "V" can be a W boson or Z boson. The "single top" includes top (anti-top) quark production in association with a W boson, and also the production in both t-channel and s-channel. The "W" means a W boson with a possible production of associated jets. These symbols are employed in all the following control plots.

Figure 4.55 shows the azimuthal distance and separation between two muons, and figure 4.56 presents the invariant mass of di-muon and rapidity of Z boson, without a requirement of jet multiplicity. Since the rochester correction (see section 4.3.4) has been applied to the reconstructed muons before kinematic selection, one can see good agreement between data and simulation at the whole selected region of Z mass window. The distribution of rapidity of Z boson is symmetric in general, and it is interrupted at $|y_{\mu\mu}| = 2.4$ due to the kinematic cut.

The event yields of exclusive jet multiplicity and inclusive jet multiplicity are shown in figure 4.57. As mentioned for 8 TeV analysis before, it is interesting to see how much precision the data can reach in detector level with the current statistics. The MC signal sample is generated using higher order calculation for low jet multiplicity production, which is expected to perform better than the sample with leading order computation. This two plots provide a primary view of the comparison between data and simulation with different production condition of jets. We can see an up shift about 10% with the requirement of $N_{jets} \geq 1, 2$, and 3. On the other hand, the ratio drops by 20-25% at high jet multiplicity region. The statistical uncertainty becomes large as well. The statistics corresponding to exclusive jet multiplicity in all the samples are summarized in table 4.12.

From figure 4.58 to figure 4.60, the plots reveal the p_T and absolute rapidity distributions of the first jet, the second jet, and the third jet respectively. One can see a global overshooting of 10% with simulation for all these variables within the statistical uncertainty. The fluctuation becomes more apparent for the variables of third jet.

The scalar sums of jets p_T , with the requirements of jet multiplicity as 1, 2, and 3 respectively, are shown in figure 4.61. Note that the low H_T limits of the three cases are different. One can see also a overall discrepancy of 10% within statistical uncertainty between data and simulation. There is a drop in the low H_T region for the case of $N_{jets} \geq 3$, which has also been seen at 8 TeV.

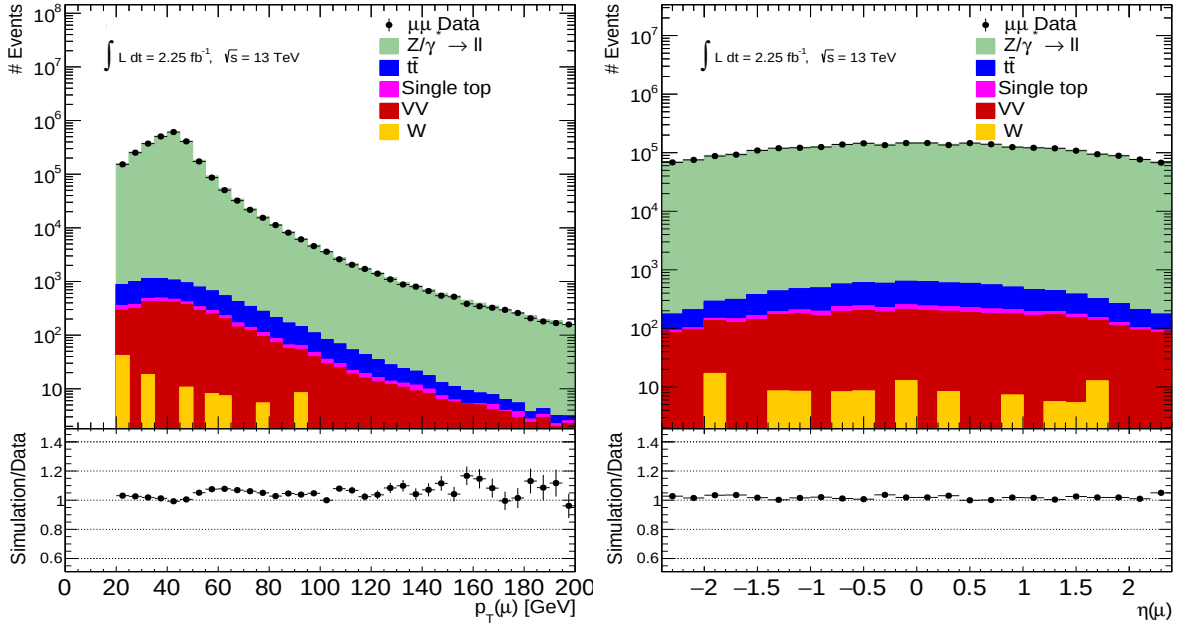


Figure 4.54: The data and simulation comparison of lepton p_T (left) and η (right) with $N_{jets} \geq 0$ for $\mu^+\mu^-$ channel at reconstruction level.

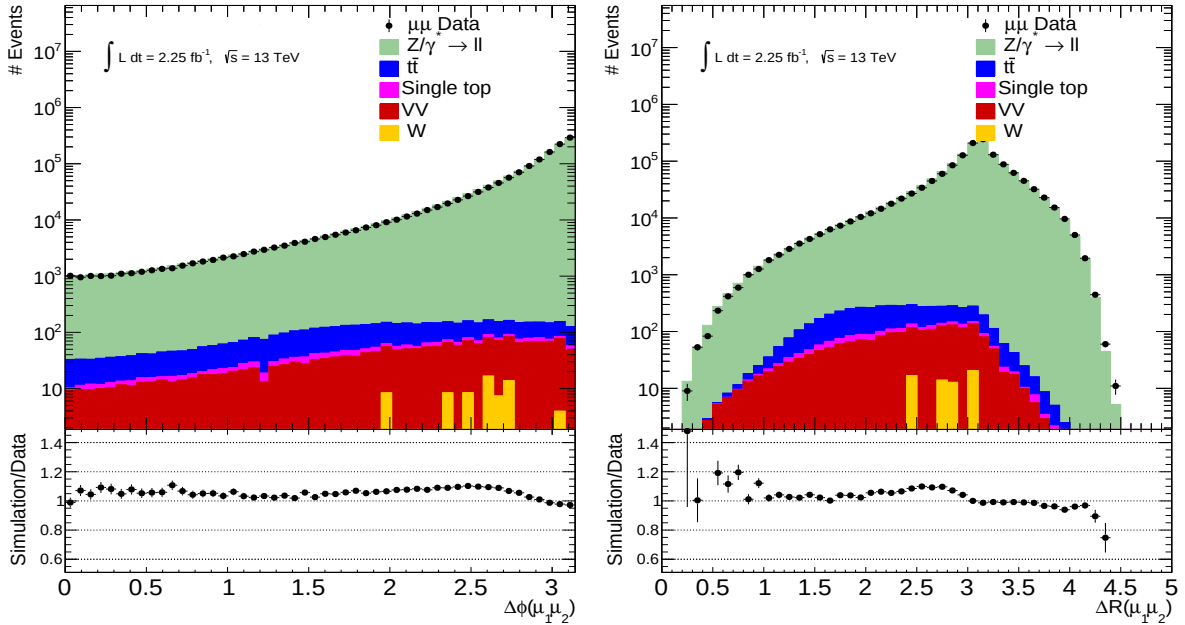


Figure 4.55: The data and simulation comparison of azimuthal distance (left) and separation (right) between two leptons with $N_{jets} \geq 0$ for $\mu^+\mu^-$ channel at reconstruction level.

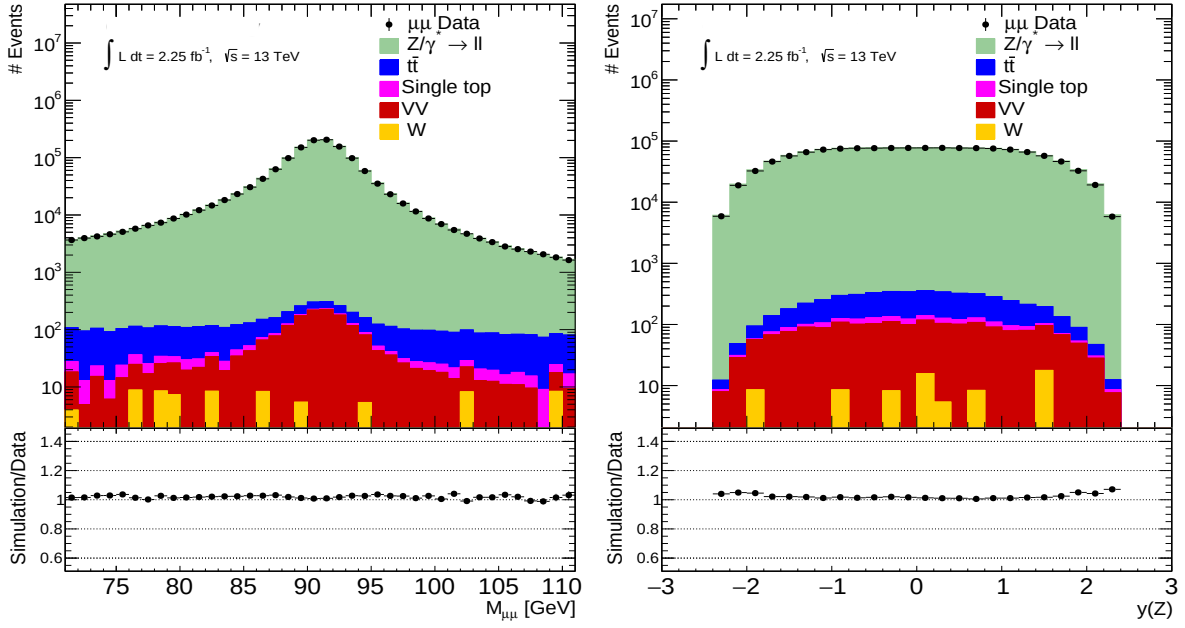


Figure 4.56: The data and simulation comparison of Z boson mass (left) and rapidity (right) with $N_{jets} \geq 0$ for $\mu^+\mu^-$ channel at reconstruction level.

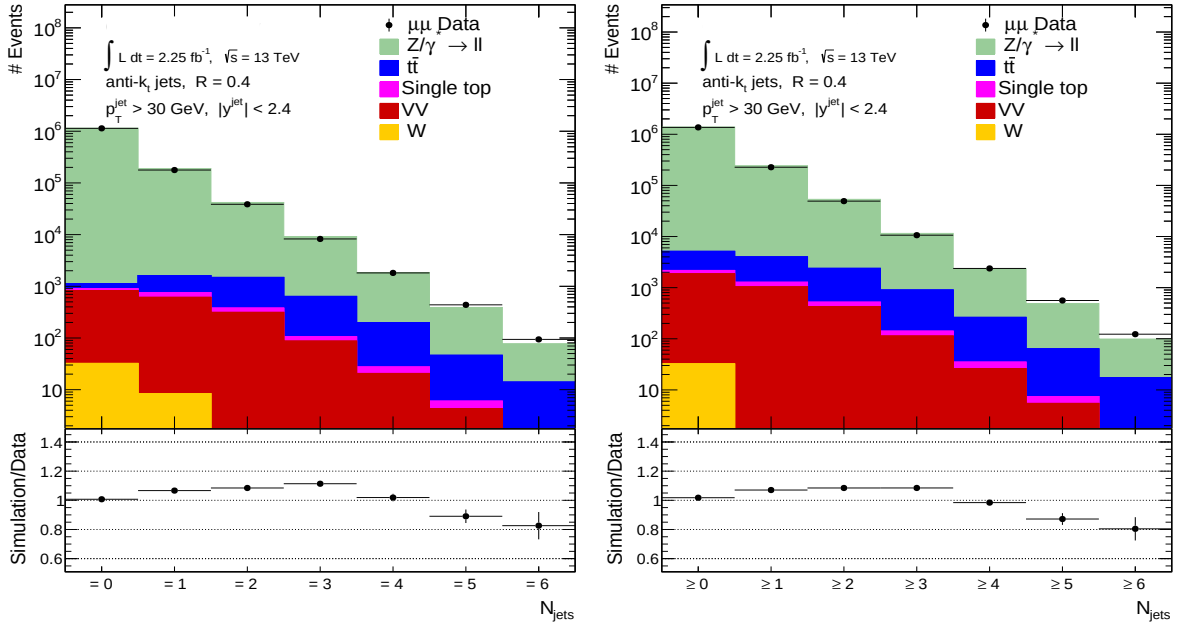


Figure 4.57: The data and simulation comparison of exclusive jet multiplicity (left) and inclusive jet multiplicity (right) for $\mu^+\mu^-$ channel at reconstruction level.

	$N_{jets} = 0$	$N_{jets} = 1$	$N_{jets} = 2$	$N_{jets} = 3$	$N_{jets} = 4$	$N_{jets} = 5$	$N_{jets} = 6$
TTbar	228	865	1109	532	171	40	13
Single top	71	143	66	19	7	2	0
ZZ	293	293	174	46	11	2	0
WW	248	82	21	5	1	0	0
Wjets	32	8	0	0	0	0	0
WZ	253	228	127	35	9	2	0
DYJets	1144940	187790	40411	8546	1652	344	64
TOTAL(MC)	1146065	189409	41908	9183	1851	390	77
Data	1137820	177590	38634	8247	1816	438	94
Ratio	1.01	1.07	1.08	1.11	1.02	0.89	0.82

Table 4.12: Number of events in data and MC samples for exclusive jet multiplicity in muon channel after full selection at 13 TeV.

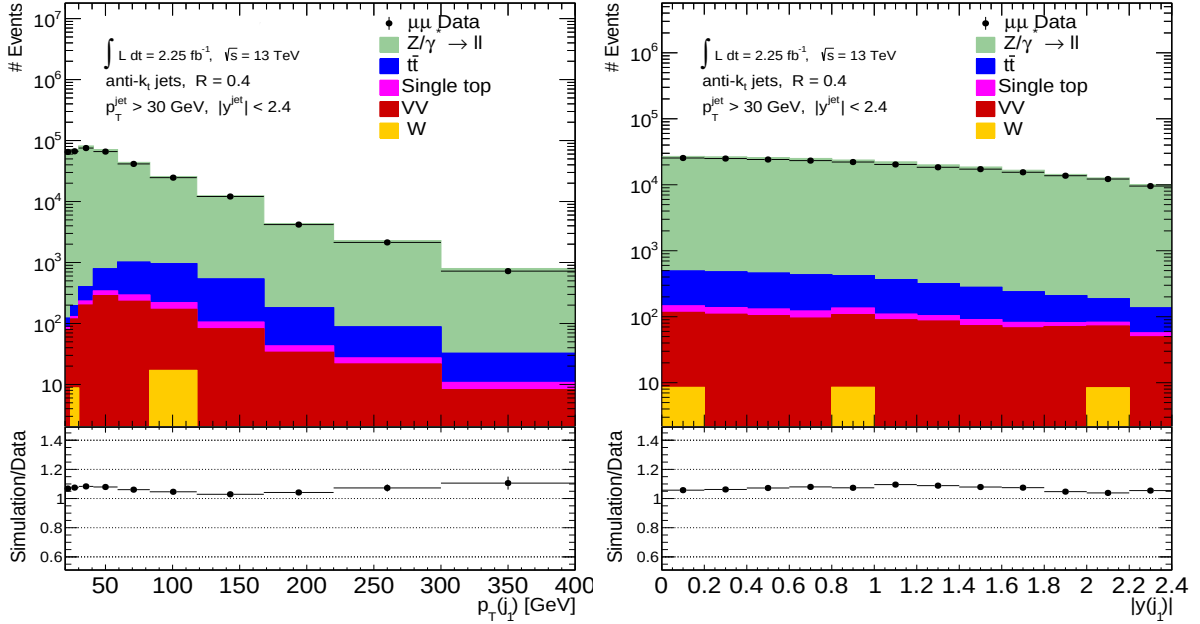


Figure 4.58: The data and simulation comparison of the first jet p_T (left) and absolute rapidity (right) with $N_{jets} \geq 1$ for $\mu^+\mu^-$ channel at reconstruction level.

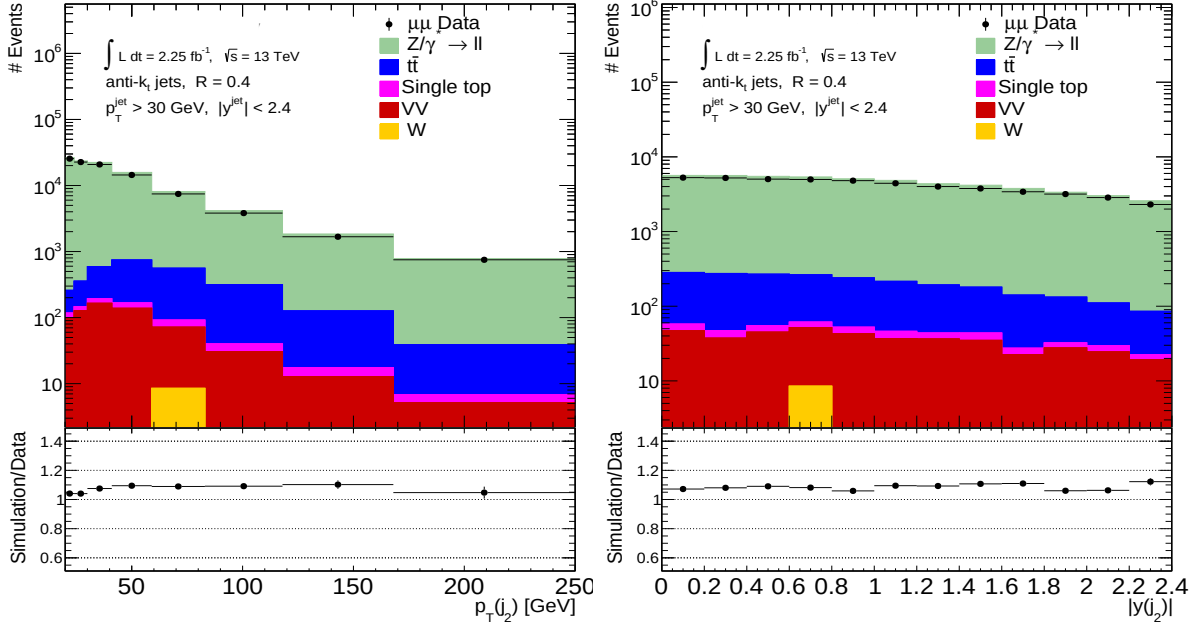


Figure 4.59: The data and simulation comparison of the second jet p_T (left) and absolute rapidity (right) with $N_{jets} \geq 2$ for $\mu^+\mu^-$ channel at reconstruction level.

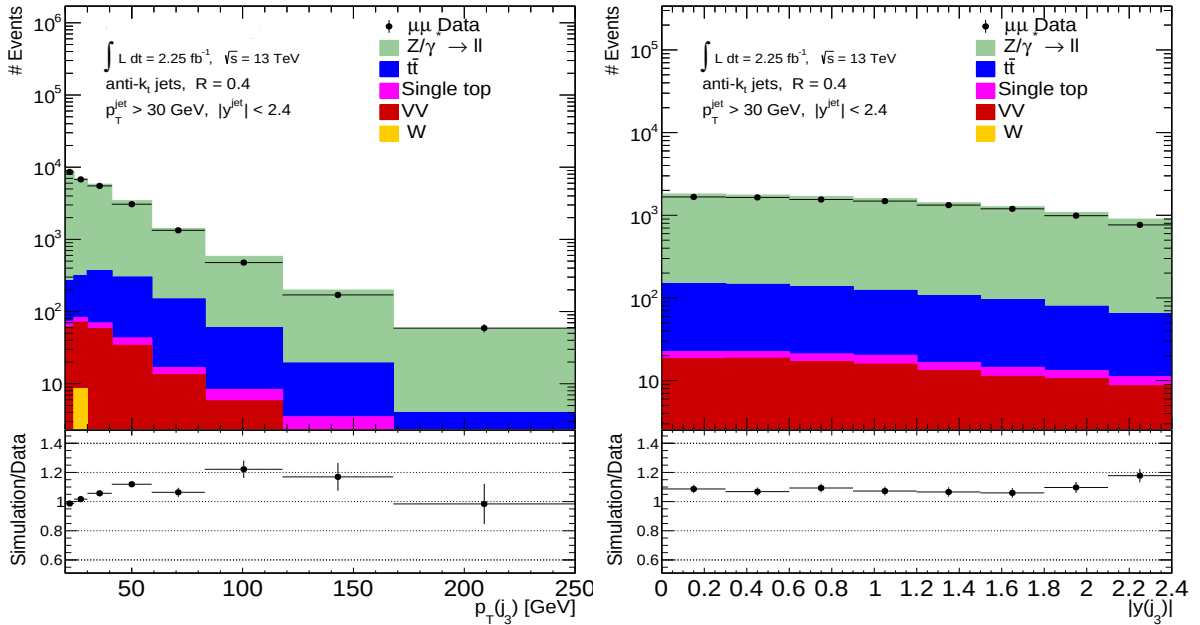


Figure 4.60: The data and simulation comparison of the third jet p_T (left) and absolute rapidity (right) with $N_{jets} \geq 3$ for $\mu^+\mu^-$ channel at reconstruction level.

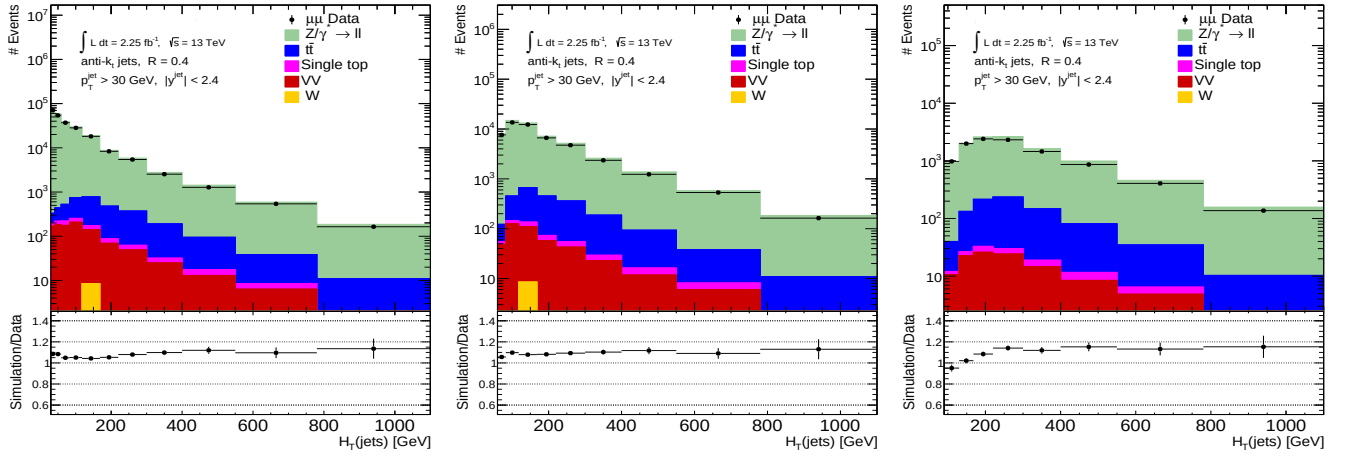


Figure 4.61: The data and simulation comparison of the scalar sum of jets p_T with $N_{jets} \geq 1$ (left), $N_{jets} \geq 2$ (middle), and $N_{jets} \geq 3$ (right) for $\mu^+\mu^-$ channel at reconstruction level.

4.4.3 Background subtraction

In order to measure the cross section of the signal process with data, we first need to subtract the background contamination from data at reconstruction level. Hence, the comparison between data and MC yields is important, and it provides with the fraction of background with the same final state signature as the signal process in the data, bin by bin. The subtraction of background in this level relies on the cross section computation of all the background processes and the quality of simulation. Therefore, the extracted signal element from data suffers from the theoretical uncertainties and systematic uncertainties of all the backgrounds. Certainly, if a background contamination is tiny, then the uncertainties originated from this background will be negligible for signal extraction.

It has to be noted that the taus, which are the third type of charged leptons, can also come from Z boson decay. They are unstable particles with a lifetime of approximate 2.9×10^{-13} s and a mass of 1776.82 MeV. They can hardly arrive at any fragment of the CMS detector since they decay immediately after producing in the collision center, so they are reconstructed using their decay products. A tau lepton decay involves a virtual W boson and a tau neutrino. The W boson decays subsequently to lighter flavor leptons (muons or electrons) together with neutrinos, or to hadrons. Among them, the neutrinos are typically recognized as significant missing energy, which cannot be reconstructed directly with the detector system. Anyway, the contribution of $\tau^+\tau^-$ leptonic decay channel is classified as a part of backgrounds in this analysis. By use of the lepton identification in the Drell-Yan MC sample, this contribution is subtracted.

For 13 TeV analysis, we extract a control sample dominated by $t\bar{t}$ process using, instead of double electrons or double muons, an electron-muon sample in data and simulation, and rescale the simulated jet multiplicity distribution according to data distribution. The scaling factor becomes significant for four jets and more. This reweighted distribution of $t\bar{t}$ background is subtracted from data.

4.5 Measured observables

The differential cross section is measured as a function of the kinematic observables as below for 8 TeV or 13 TeV:

- exclusive and inclusive jet multiplicities (N_{jets});
- transverse momentum of Z boson for $N_{jets} \geq 1$ (only for 8 TeV);
- absolute rapidity of Z boson for $N_{jets} \geq 1$ (only for 8 TeV);
- transverse momentum of the n^{th} jet for $N_{jets} \geq n$, where $n = 1-5$ for 8 TeV and 1-3 for 13 TeV;
- absolute rapidity of the n^{th} jet for $N_{jets} \geq n$, where $n = 1-5$ for 8 TeV and 1-3 for 13 TeV;
- scalar sum of jet transverse momenta for $N_{jets} \geq n$, where $n = 1-5$ for 8 TeV and 1-3 for 13 TeV;

Based on section 4.1, it has to be noted that the rapidity distributions depend on how many jet multiplicities are taken into account. If one requires at least two jets, then the second leading jet, which has second largest momentum among all the jets, may compensate the unbalance between Z boson and the leading jet. At 8 TeV, we have also measured the differential cross section as a function of the following correlation variables:

- $y_{sum(Z,j_n)}$, where $n = 1, 2$;
- $y_{sum(Z,dijet)}$, where dijet is the vector sum of the two leading jets;
- $y_{sum(j_1,j_2)}$;
- $y_{diff(Z,j_n)}$, where $n = 1, 2$;
- $y_{diff(Z,dijet)}$, where dijet is the vector sum of the two leading jets;
- $y_{diff(j_1,j_2)}$.

4.6 Data corrections for detector effects

The measurement relies on the particle reconstruction techniques, while the detector has a limited precision. For instance, the reconstructed transverse momentum of a charged lepton or a jet usually is not equal to the true transverse momentum, which is due to the finite resolution of the reconstruction system. Additionally, the angular values of a particle, such as rapidity and azimuth, can also be biased. As a result, the global observable spectrum at produced level is seen as smeared at reconstruction level. Although one may apply the energy correction for the data and smear the energy for simulation for the reconstructed objects, there is still a global efficiency loss.

In order to correct this, an unfolding method is applied. This solution is based on MC technique, and needs the information of both generated values from the MC generator and reconstructed values after the detector simulation. The reconstructed events by definition have passed the kinematic selection. Generated events are selected in a specific phase space. To limit uncertainties due to detector effects, the cross sections are measured in a phase space limited by experimental cuts, like the minimal jet transverse momenta.

4.6.1 Phase space selection

Among the generated events, first we need to pick up all the muons and electrons. Note that taus and neutrinos are excluded from our analysis. They can be produced with two categories, which are leptons before (status 3) and after (status 1) quantum electrodynamics (QED) FSR. In this analysis, the MADGRAPH generator generates both types of leptons as it is interfaced to PYTHIA, while the SHERPA generator only has the leptons after QED FSR. To compare the theoretical predictions

using different generators, the leptons with status 1 are selected into the phase space. On the other hand, the leptons before QED FSR are needed to reconstruct accurately the generated Z bosons. In order to do that, we implement an approximated approach, through scanning the photons in a cone around the leptons after FSR. We compensate the four-momentum of each lepton with the four-momentum of photons that are close to the lepton. The formula is: $p_l^{cor} = p_l + \sum_{\gamma} p_{\gamma}$, where p_l or p_l^{cor} is the four-momentum of a lepton before or after this correction, respectively, and p_{γ} is the four-momentum of a photon that has to be within the range of $\Delta R = 0.1$ around the lepton. After that, we say that the leptons are dressed with hard photon radiation⁸. From these dressed leptons, we select the two hardest ones with opposite charges, and require the same p_T and η cuts as for selecting the reconstructed leptons. Furthermore, their invariant mass must lie in a Z boson mass window as mentioned before.

With respect to the jet selection in phase space, each generated jet must have an isolation of at least $\Delta R = 0.5$ with respect to the selected generated leptons, and satisfy the same p_T and η requirements as reconstructed jets. Table 4.13 summarizes the phase space that we select.

Variable	criterion
$p_T(l)$	$> 20 \text{ GeV}$
$ \eta(l) $	< 2.4
$M(l^+l^-)$	$(71 - 111) \text{ GeV}$
$p_T(j)$	$> 30 \text{ GeV}$
$ y(j) $	< 2.4
$\Delta R(j, l)$	> 0.5

Table 4.13: List of criteria on the phase space selection. The "l" and "j" stand for a lepton (electron or muon) and a jet, respectively.

4.6.2 Introduction to Unfolding approach

The unfolding procedure is MC dependent, because it needs the values of an observable in both reconstructed level and generated level, and implements a matching between the values. Generated (true) distributions of an observable are predicted by a specific MC generator, which computes the matrix element and QCD/QED radiative corrections. Afterwards, the reconstructed value will be dealt with a MC simulation, which takes into account the detector material geometry and response. Therefore it allows for estimating the finite acceptance and resolutions. To be specific, MC tools can generate plenty of events and proceed with simulations. After the event selection, one can obtain two histograms at reconstructed level and true level respectively for an observable. The matching

⁸It has been checked that these dressed leptons are not fully equivalent to the real generated leptons before FSR. For example, a hard photon not from FSR is possibly overlapping a lepton, and then it will be taken into account the correction. On the other hand, a photon from FSR can also have large isolation with the mother lepton, while it is ignored. Moreover, this residual effect also depends on the cone size (ΔR) that one chooses. It has been checked the this effect affects the jet related observables with 2% in $\mu^+\mu^-$ channel and 2% - 4% in e^+e^- channel [13].

approach is to make a two-dimensional matrix, which is called response matrix, where the element in i^{th} row and j^{th} column corresponds to the probability of the observable in the i^{th} bin of the true level histogram while in the j^{th} bin of the reconstructed level histogram. Certainly, there is also an possibility that the observable in the i^{th} of true level cannot match with that in any bin of reconstructed level, which is defined as missing events. On the contrary, i.e. the observable in the j^{th} of reconstructed level does not match with that in any bin of true level, it is named as fake events. Then one can correct the observable distribution in data using the inverted response matrix, after the subtraction of fake events and before adding the missing events. This leads to the data distribution at particle level.

According to the above principle, the most difficult operation will be the inversion of the response matrix. There is a RooUnfold package [96] supplying a common framework to calculate and deploy different unfolding methods. A direct method is named bin-by-bin correction, which simply computes the inefficiency factor from true level to reconstructed level in each bin, and then corrects the data by multiplying these factors bin by bin. The correlations between different bins are ignored in this method. Hence, the result is influenced by the migrations between contiguous bins only. Since the correlation between different bins indeed are not negligible in this analysis, in particular for jet variables, this method is not precise enough and largely model dependent.

The second method is the singular value decomposition (SVD), which has been introduced in [97]. It implements a decomposition procedure for inverting the response matrix. In the least square senses, one needs to optimize the regularization parameters to make smooth unfolded distributions satisfying the initial linear algorithm. The regularization solution has suppressed the spurious and quickly oscillating components, remaining statistically significant elements. However, as will be mentioned later, we will use the number of events in the response matrix, and the SVD requires the initial MC statistics are at least one or two orders of magnitude higher than the statistics of the measured data, which is not the case in this analysis. Hence, although the SVD method can solve the issue of bin migrations for bin-by-bin scenario and consider the instability corresponding to the inversion of response matrix, it is still not the best choice. Furthermore, there is no reason why the inverse matrix should exist.

The third method, which is applied in this analysis, is called the Bayes approach [98]. It does not require the inversion of response matrix, and it takes into account the correlations and migrations between adjacent bins. Based on Bayes theorem, the unfolding formula can be written as:

$$n(T_i) = \frac{1}{\epsilon_i} \sum_{j=1}^{N_R} n(R_j) P(T_i|R_j) \quad (4.8)$$

where $n(T_i)$ is the number of events in the i^{th} bin of true distribution; ϵ_i is the inefficiency factor⁹ for this bin; $n(R_j)$ is the number of events in the j^{th} bin of the reconstructed distribution; N_R is the number of bins of the reconstructed distribution. The $P(T_i|R_j)$ is called unfolding matrix, which

⁹The $n(T_i)$ will be set to zero if ϵ_i is zero, because the experiment is not sensitive to this case of T_i .

describes cell-to-cell migration of the distributions. It is defined as:

$$P(T_i|R_j) = \frac{P(R_j|T_i)P_0(T_i)}{\sum_{l=1}^{N_T} P(R_j|T_l)P_0(T_l)} \quad (4.9)$$

where $P(R_j|T_i)$ stands for the conditional probability that a event seen in the j^{th} bin of reconstructed distribution was generated in the i^{th} bin of true distribution; $P_0(T_i)$ represents the initial probability of the causes of true distribution. It is noted that $\sum_{i=1}^{N_T} P_0(T_i) = 1$ for the normalization requirement. As mentioned before, the backgrounds estimated with simulations have been subtracted from data distribution before unfolding. Moreover, by use of MC signal sample, we have estimated the events passing kinematic selection in reconstructed level, while failing the phase space requirement in generated level. These events are named as fake events, which are re-scaled bin-by-bin according to the event difference between MC signal and data minus backgrounds, and these events are also removed from data bin-by-bin for each observable.

In equation 4.8, the unfolding matrix $P(T_i|R_j)$ can be computed using the MC signal sample, since it contains both $n(T_i)$ and $n(R_j)$. After initializing the P_0 , the unfolding program implements an iterative procedure to compute the optimal values in the unfolding matrix. The initialization starts from the true MC distribution. In principle, the number of iterations should be enough once the program reaches a convergence, and the result will not be changed with more additional iterations. However, it has to be noticed that too many iterations can lead to large discrepancies. Hence, it is necessary to have a maximum number of iterations. In this analysis, during each time of iteration, we apply the raw response matrix to the unfolded distribution, which is called folding pattern, and we compare the folded distribution with the initial data distribution using χ^2 approach. If the χ^2 divided by number of freedom is less than $1/\sqrt{2}$, the two distributions are compatible within their statistical uncertainties. By use of this kind of monitor, we use 4 iterations for all of the observables at 8 TeV and for most of the ones at 13 TeV, except 1 iteration for absolute jet rapidity at 13 TeV due to the large statistical fluctuations of Drell-Yan MC sample.

4.6.3 Response matrix

In order to prepare the unfolding response matrix, one needs to use a suitable MC sample which contains both generated events and reconstructed events. We use the MC signal sample produced by MADGRAPH (LO) at 8 TeV, while using the sample generated by MADGRAPH_AMC@NLO at 13 TeV. The first question is how to match each reconstructed object spectrum with the related generated object spectrum, such as the p_T profiles of a Z boson or jets. At reconstructed level, we select two leptons which are sorted by decreasing p_T , and associated jets which are ordered by decreasing p_T as well. Meanwhile, we use the generated objects in the similar way. We match the leading reconstructed jet with the leading generated jet event by event, and the same story happens in the following jets. This pattern is realized according to the p_T order of objects, but it can happen with the angular correlation between the reconstructed object and the generated one. For instance,

the second generated jet is reconstructed as the leading jet due to the finite momentum resolution of the detector. Although this reconstructed jet is closer (in angle) to the first generated jet than to the second generated jet. Therefore it would be matched with the wrong generated jet in the p_T ordering scenario. This issue affects the response matrices of jet observables for unfolding.

Besides the p_T order matching for the object distributions, we study a second choice called angular matching approach, which can compensate the drawback above. Within the phase space constraint and kinematic selection, if there is a generated jet, the program will seek for a jet at reconstruction level that has the smallest angular distance with this generated jet, by looping among all of the reconstructed jets. There is an upper threshold of the isolation cone, which is chosen to be $\Delta R = 0.25$. A generated jet which cannot be matched with any reconstructed jet inside the cone will be treated as a missed jet. On the other hand, if a reconstructed jet has no matched jet "partner" in generator level, it is classified as a fake jet. This two cases are induced by the detector effect as explained before. Through implementing the angular match procedure, we obtain a new response matrix. On the other hand, this approach also has disadvantages, such that one may have to resort either the generated jets or the reconstructed jets in p_T . For example, both the reconstructed jets and generated jets were ordered by their p_T . If the leading generated jet is matched with the second leading reconstructed jet, this reconstructed jet will be re-recorded as the leading reconstructed jet, and vice versa. However, there are only reconstructed jets recorded in the data, and they are sorted by their p_T . For this reason, before MC model computation of unfolding procedure, one has to reorder the generated jets for the input response matrices. After unfolding data, the leading jet at reconstruction level may not be the leading jet at produced level any more, and the same for other jets. It causes the difficulty to make one to one correspondance between leading jet prediction and leading jet for unfolding. In other word, leading jet prediction becomes dependent on the unfolding procedure.

Therefore, it is worth to compare the results with these two patterns. First, the comparisons of response matrices corresponding to several kinematic observables of jets are shown from figure 4.62 to figure 4.69, by presenting the performances of 8 TeV in $Z \rightarrow \mu^+\mu^-$ decay channel. They are built using MADGRAPH (LO) sample.

The number in each matrix element is the probability (in percentage) of an event located in the region after phase space requirement in generator level and kinematic selection in reconstruction level, normalized on the sum of events for a given bin of generated level. In general, one can see that the diagonal elements are dominant in both sides, which means that most of the generated events within the phase space are well reconstructed. To be specific, the kinematic distributions with angular matching method have the events more concentrated around diagonal line for the rapidity distributions of jets, which means the reconstructed observable corresponds to the generated one much better, compared to the distributions with p_T order matching. It results in exactly the same response matrix in jet multiplicity with the two patterns. For the distributions of jet p_T and H_T , the angular matching scheme does not have apparent improvement in the response matrices, relative to

the p_T ordering matching scheme. These response matrices are filled with the weighted events in the analysis chain. The binning choice affects the response matrix values due to the bin migration effect. In order to allow a proper unfolding, we choose an optimal bin width with each diagonal bin purity of above 60%.

Regarding the variables like p_T and H_T of jets, the filled histograms including the response matrices have overflow bins, if the histograms do not cover the full ranges of the observable spectra¹⁰. The unfolding method takes into account the migration effect between the normal bins and those overflow bins. It has to be noted that the lower threshold of jets p_T has been adjusted to 20 GeV when we fill the histograms of jets p_T , while we keep them as 30 GeV as the condition for all the other observable histograms. The reason is that we treat the region between 20 GeV and 30 GeV as “underflow” region, where the possible migration effect can be corrected by the unfolding procedure.

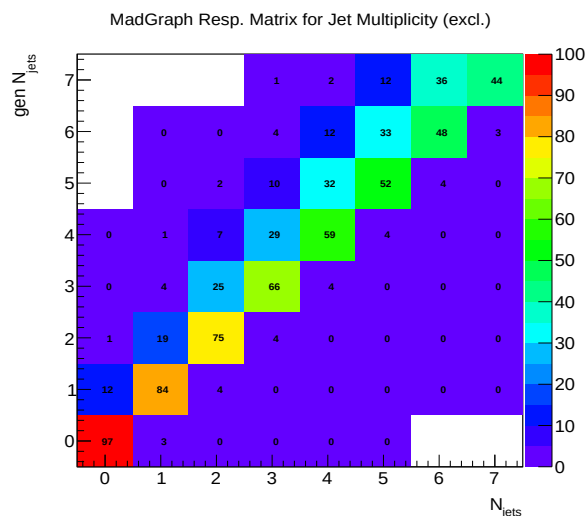


Figure 4.62: The response matrix of exclusive jet multiplicity, constructed from MADGRAPH (LO) at 8 TeV, with p_T order matching or angular matching (same).

Regarding 13 TeV analysis, the same response matrices with p_T ordering matching and angular matching are compared too. These response matrices are built with the reconstructed and generated distributions using MADGRAPH_AMC@NLO generator. Note that the bin width at 13 TeV is larger than that at 8 TeV for the kinematic observables and with less jet multiplicity at 13 TeV. This is because of less statistics at 13 TeV analysis. Figure 4.70 to 4.77 show the comparisons, from which we can also see the angular matching makes the response matrices of jet rapidity more diagonal.

¹⁰Since there are p_T lower cut and absolute rapidity upper threshold for jets, there is no underflow bin or overflow bin for jet absolute rapidity distribution, and no underflow bin for jet p_T and H_T distributions. There is an overflow bin for the distribution of jet multiplicity, but it is limited by statistics.

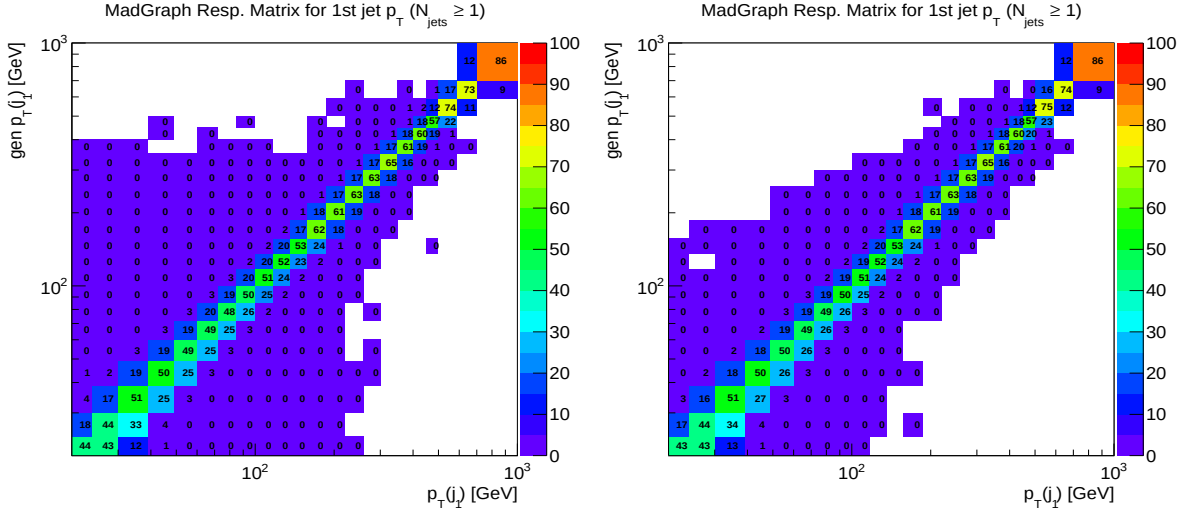


Figure 4.63: The response matrix of the first jet p_T for $N_{jets} \geq 1$, constructed from MADGRAPH (LO) at 8 TeV, with p_T order matching (left) and angular matching (right).

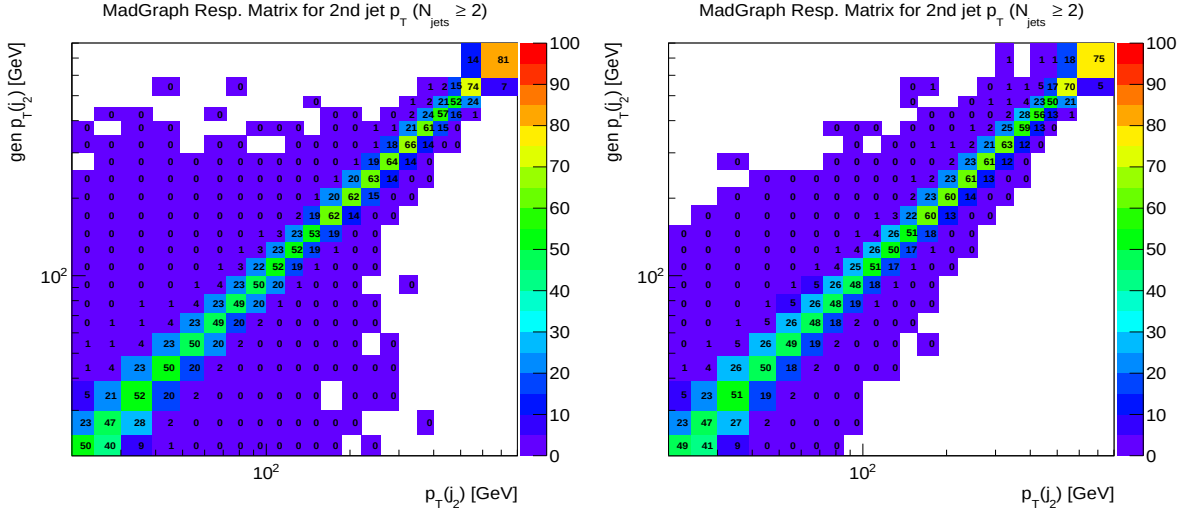


Figure 4.64: The response matrix of the second jet p_T for $N_{jets} \geq 2$, constructed from MADGRAPH (LO) at 8 TeV, with p_T order matching (left) and angular matching (right).

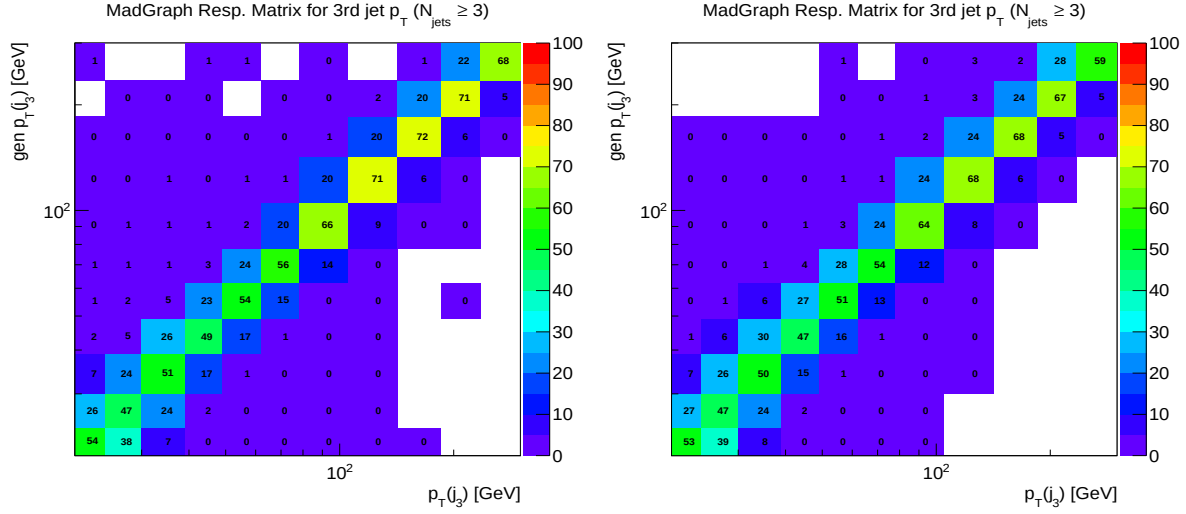


Figure 4.65: The response matrix of the third jet p_T for $N_{jets} \geq 3$, constructed from MADGRAPH (LO) at 8 TeV, with p_T order matching (left) and angular matching (right).

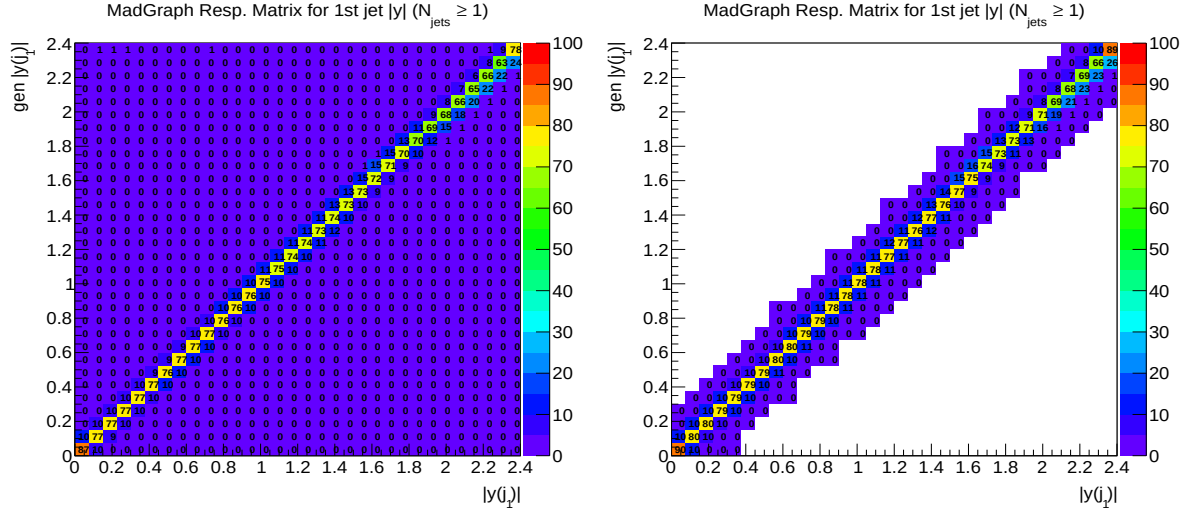


Figure 4.66: The response matrix of the first jet rapidity for $N_{jets} \geq 1$, constructed from MADGRAPH (LO) at 8 TeV, with p_T order matching (left) and angular matching (right).

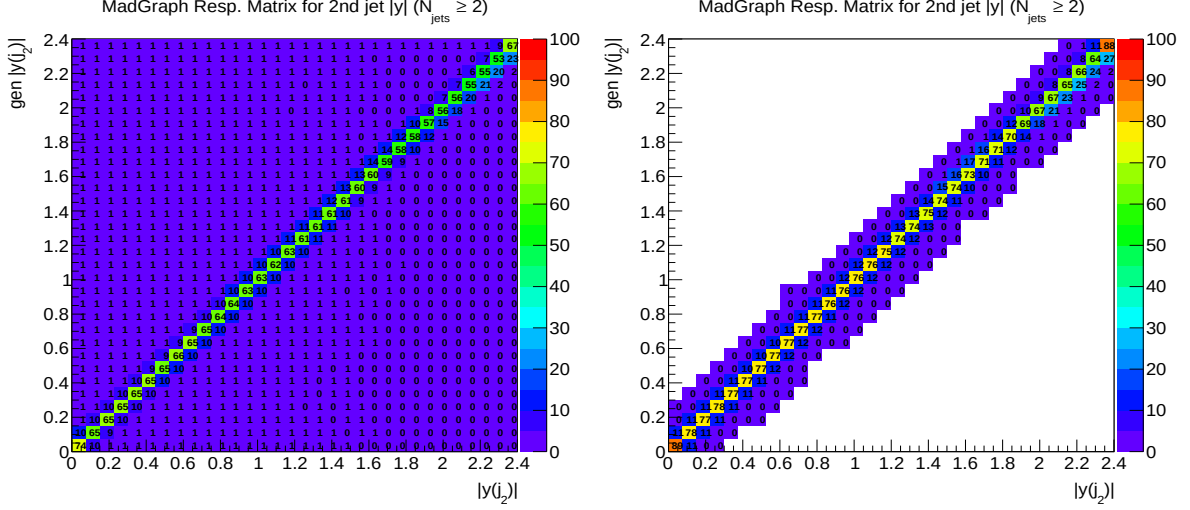


Figure 4.67: The response matrix of the second jet rapidity for $N_{jets} \geq 2$, constructed from MADGRAPH (LO) at 8 TeV, with p_T order matching (left) and angular matching (right).

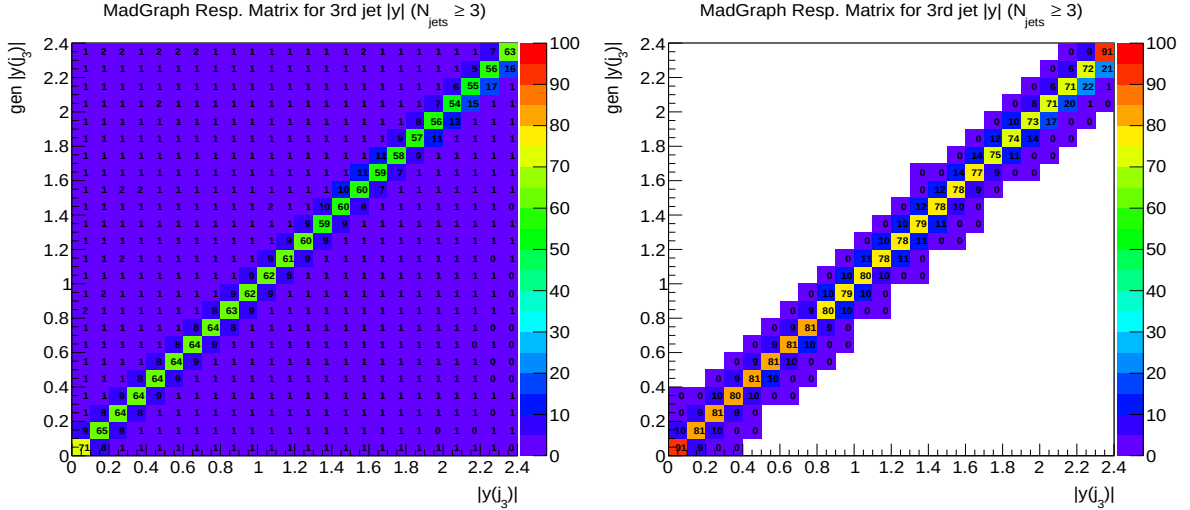


Figure 4.68: The response matrix of the third jet rapidity for $N_{jets} \geq 3$, constructed from MADGRAPH (LO) at 8 TeV, with p_T order matching (left) and angular matching (right).

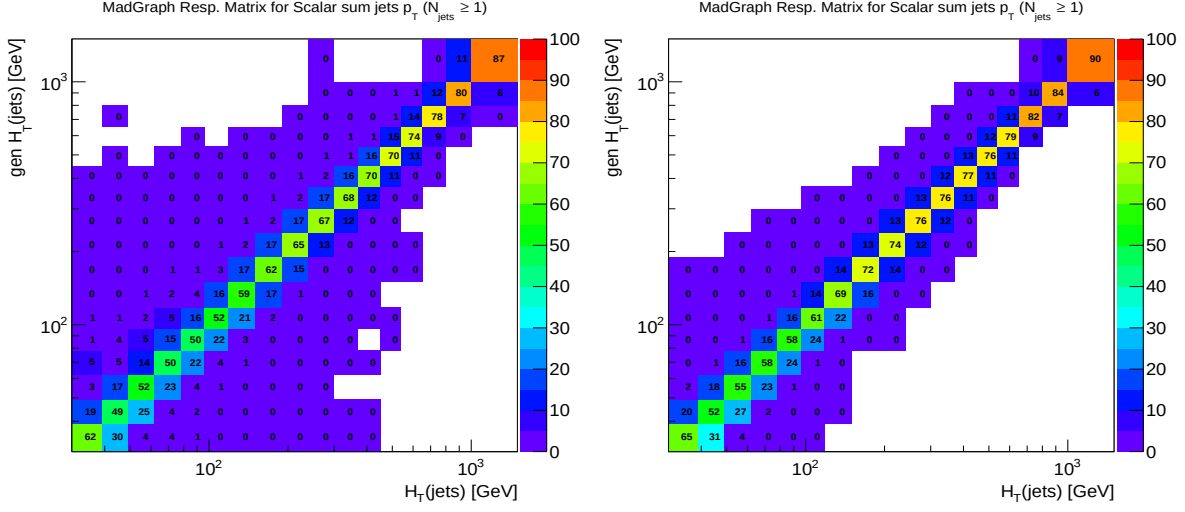


Figure 4.69: The response matrix of the scalar sum of jets p_T for $N_{jets} \geq 1$, constructed from MADGRAPH (LO) at 8 TeV, with p_T order matching (left) and angular matching (right).

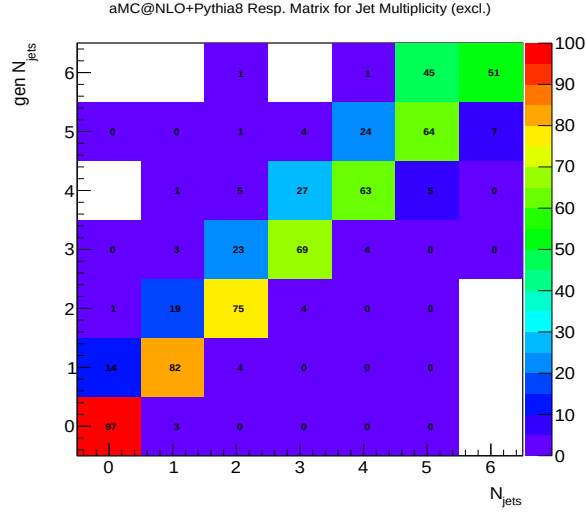


Figure 4.70: The response matrix of exclusive jet multiplicity, constructed from MADGRAPH_AMC@NLO at 13 TeV, with p_T order matching or angular matching (same).

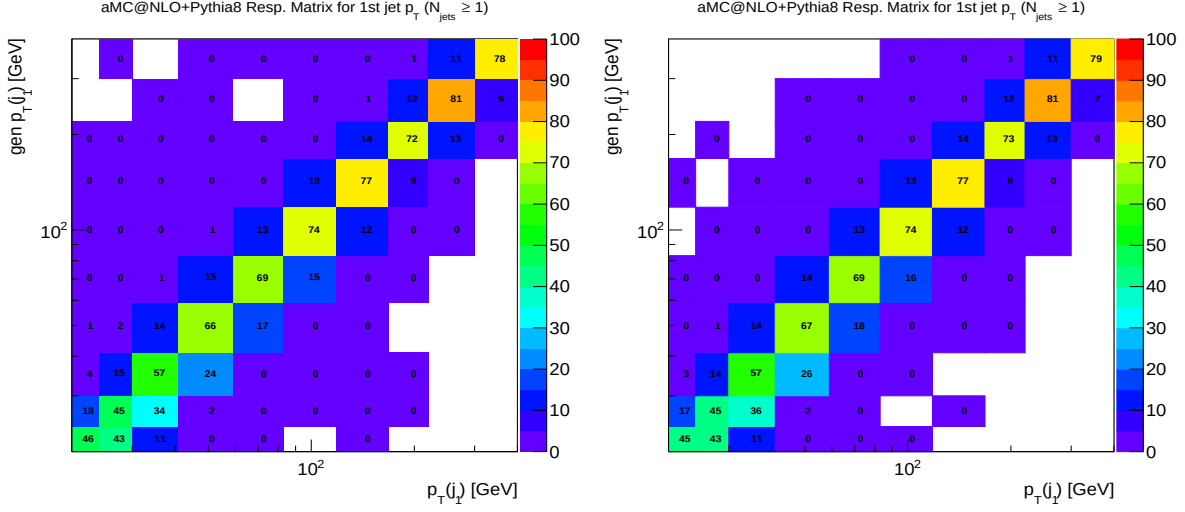


Figure 4.71: The response matrix of the first jet p_T for $N_{jets} \geq 1$, constructed from MADGRAPH_AMC@NLO at 13 TeV, with p_T order matching (left) and angular matching (right).

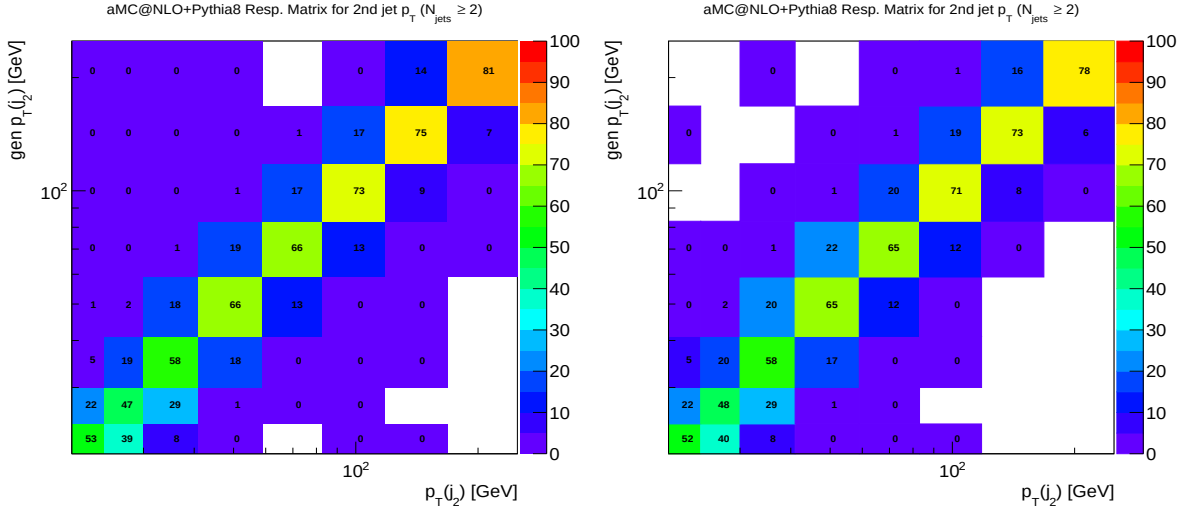


Figure 4.72: The response matrix of the second jet p_T for $N_{jets} \geq 2$, constructed from MADGRAPH_AMC@NLO at 13 TeV, with p_T order matching (left) and angular matching (right).

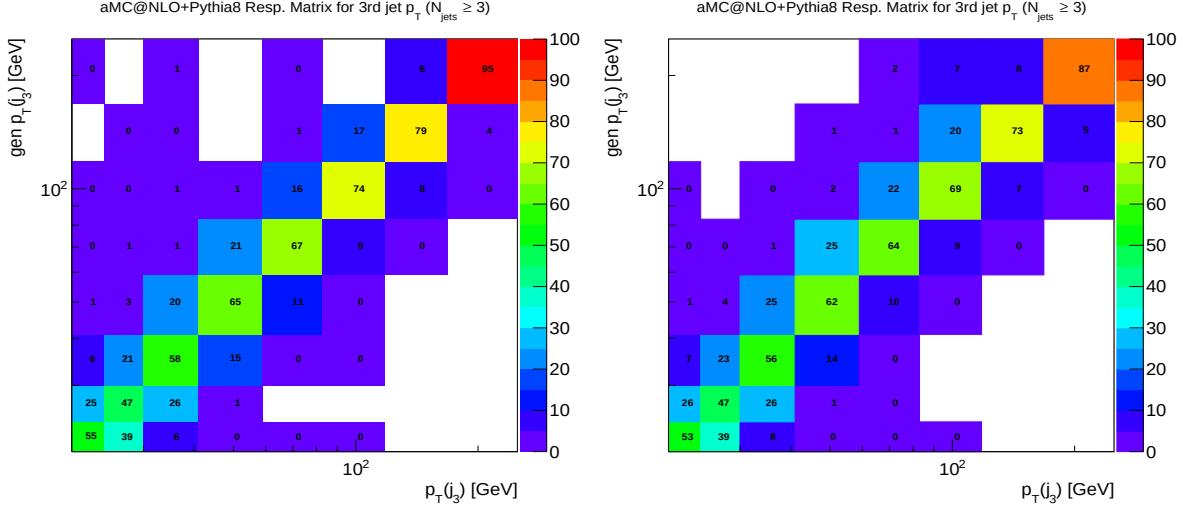


Figure 4.73: The response matrix of the third jet p_T for $N_{jets} \geq 3$, constructed from MADGRAPH_AMC@NLO at 13 TeV, with p_T order matching (left) and angular matching (right).

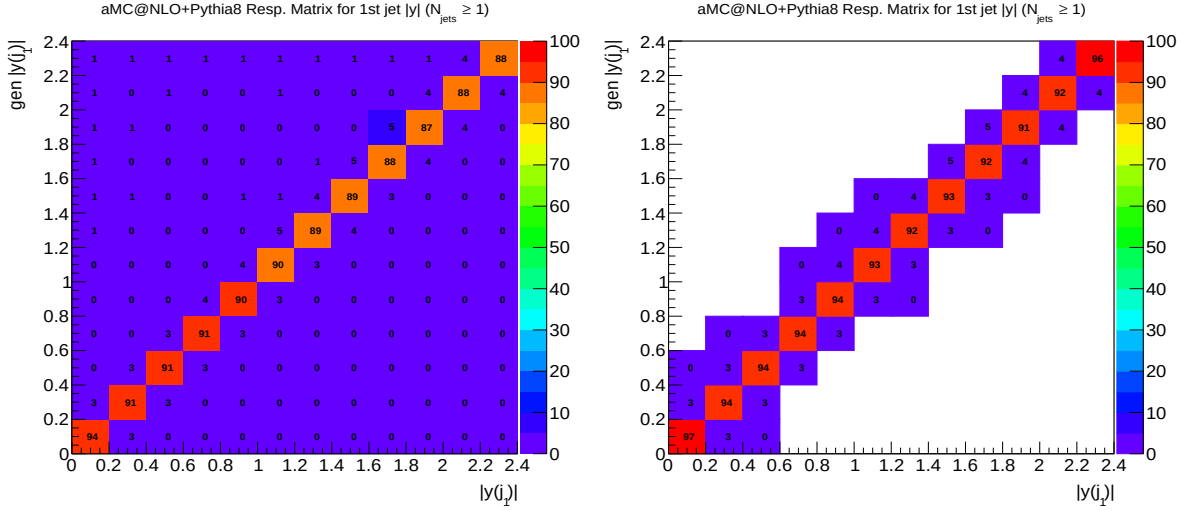


Figure 4.74: The response matrix of the first jet rapidity for $N_{jets} \geq 1$, constructed from MADGRAPH_AMC@NLO at 13 TeV, with p_T order matching (left) and angular matching (right).

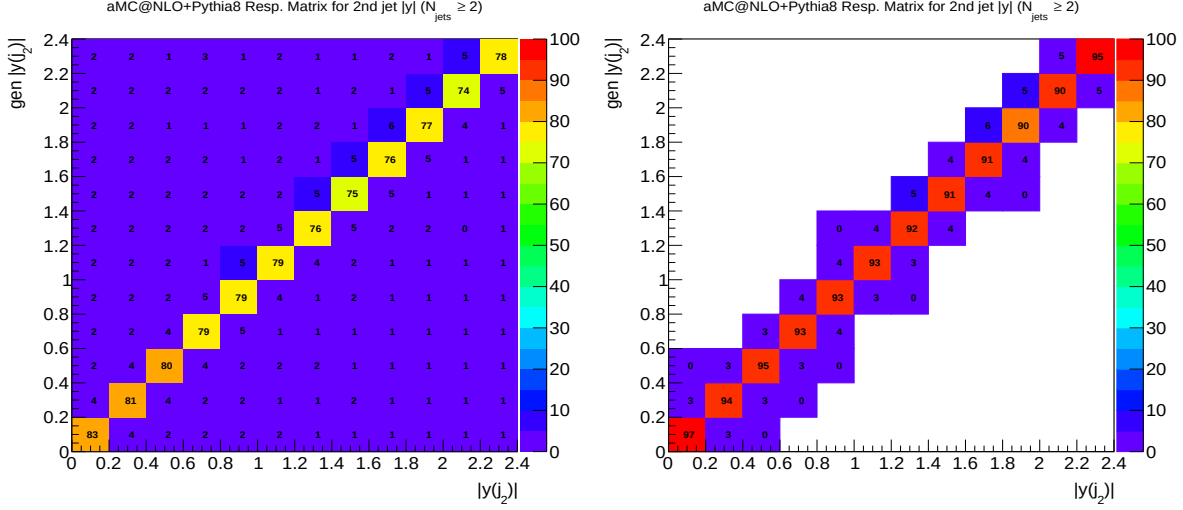


Figure 4.75: The response matrix of the second jet rapidity for $N_{jets} \geq 2$, constructed from MADGRAPH_AMC@NLO at 13 TeV, with p_T order matching (left) and angular matching (right).

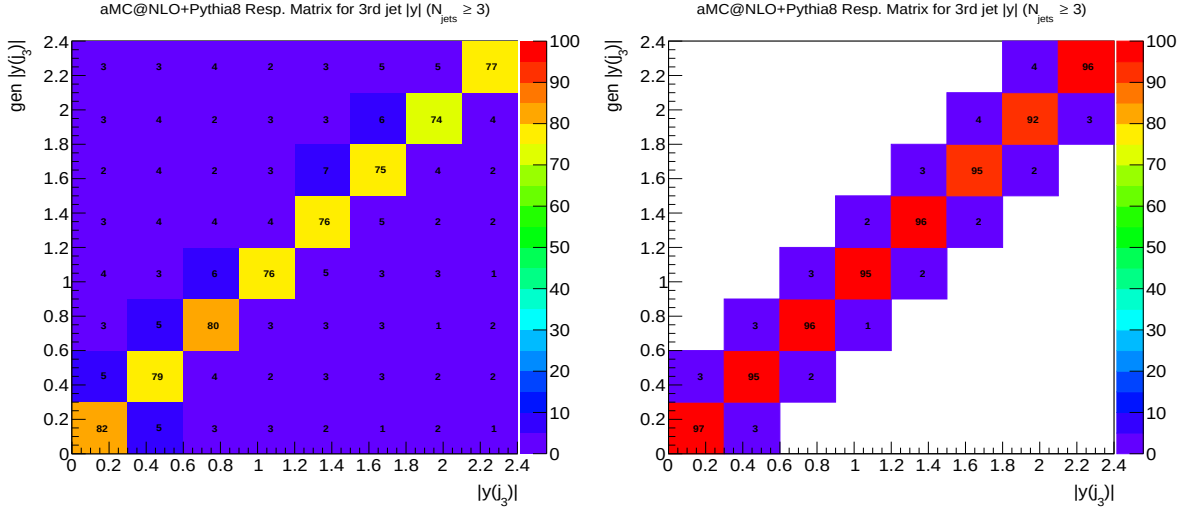


Figure 4.76: The response matrix of the third jet rapidity for $N_{jets} \geq 3$, constructed from MADGRAPH_AMC@NLO at 13 TeV, with p_T order matching (left) and angular matching (right).

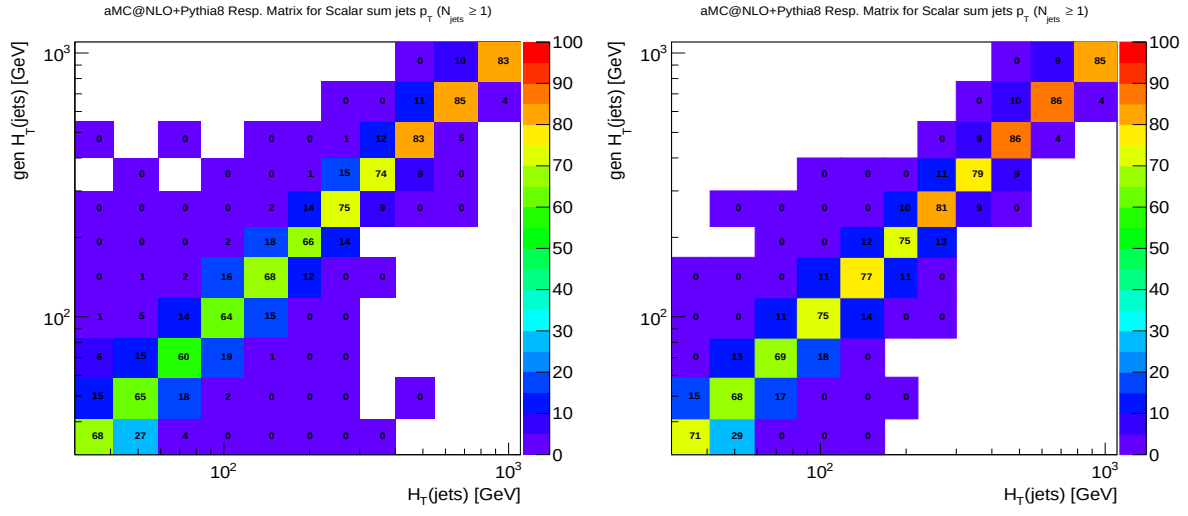


Figure 4.77: The response matrix of the scalar sum of jets p_T for $N_{jets} \geq 1$, constructed from MADGRAPH_AMC@NLO at 13 TeV, with p_T order matching (left) and angular matching (right).

4.6.4 Comparison between p_T order matching and angular matching

Since we have seen the response matrices of jet rapidity are improved by the angular matching approach, it is interesting to quantify how much the two matching scenarios affect the measured differential cross sections after unfolding data due to the difference in the response matrices. The comparisons are illustrated from figure 4.78 to figure 4.81 at 8 TeV. By definition there is no difference in the jet multiplicity distribution. The difference between the two matching schemes is seen in the absolute rapidity distributions of the second, third, forth, and fifth jet. There is also a bit difference at low p_T region for those jets. In general, the difference is a few percent. As mentioned before, the angular matching scheme leads to the confusion for corresponding jets between unfolded data and theoretical predictions. Hence, we choose the p_T order matching scheme in this analysis.

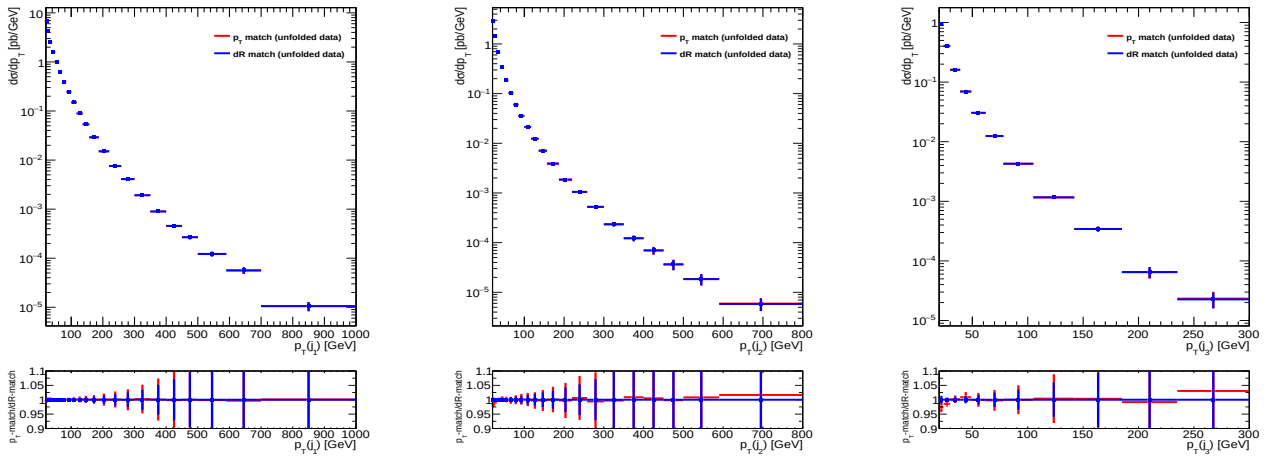


Figure 4.78: The measured differential cross section as a function of the leading jet (left), the second jet (middle), and the third jet (right) p_T in $\mu^+\mu^-$ decay channel at 8 TeV, comparing p_T order matching and angular matching scenarios. The error bars stand for statistical uncertainties.

The comparisons at 13 TeV are illustrated from figure 4.82 to 4.84. The response matrices are constructed by means of the events generated using MADGRAPH_AMC@NLO. There is no apparent discrepancy in all the distributions using the two matching approaches. Therefore we also choose the p_T ordering matching scheme for this analysis.

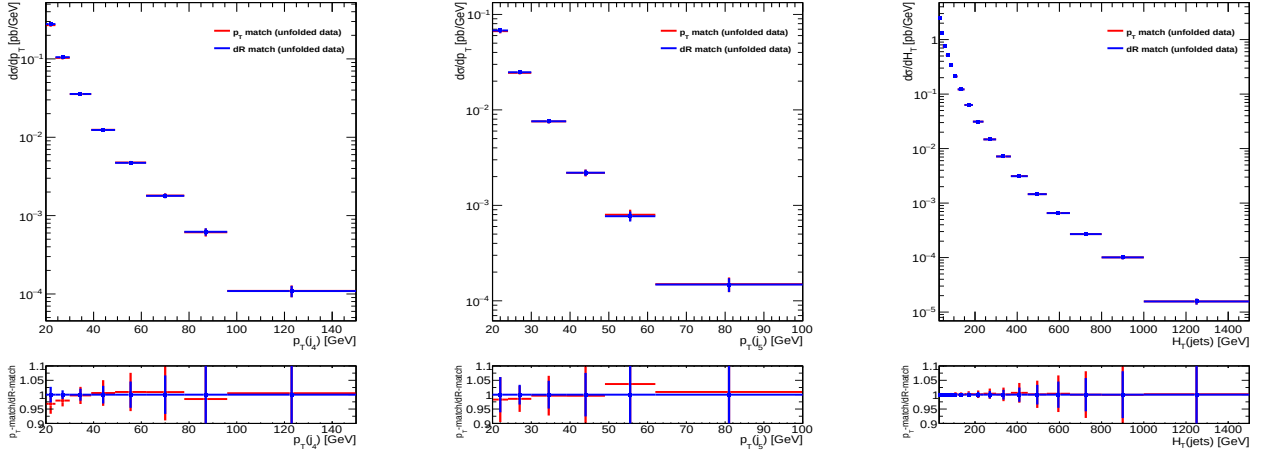


Figure 4.79: The measured differential cross section as a function of the fourth jet p_T (left), the fifth jet p_T (middle), and the scalar sum of jets p_T with $N_{jets} \geq 1$ (right) in $\mu^+\mu^-$ decay channel at 8 TeV, comparing p_T order matching and angular matching scenarios. The error bars stand for statistical uncertainties.

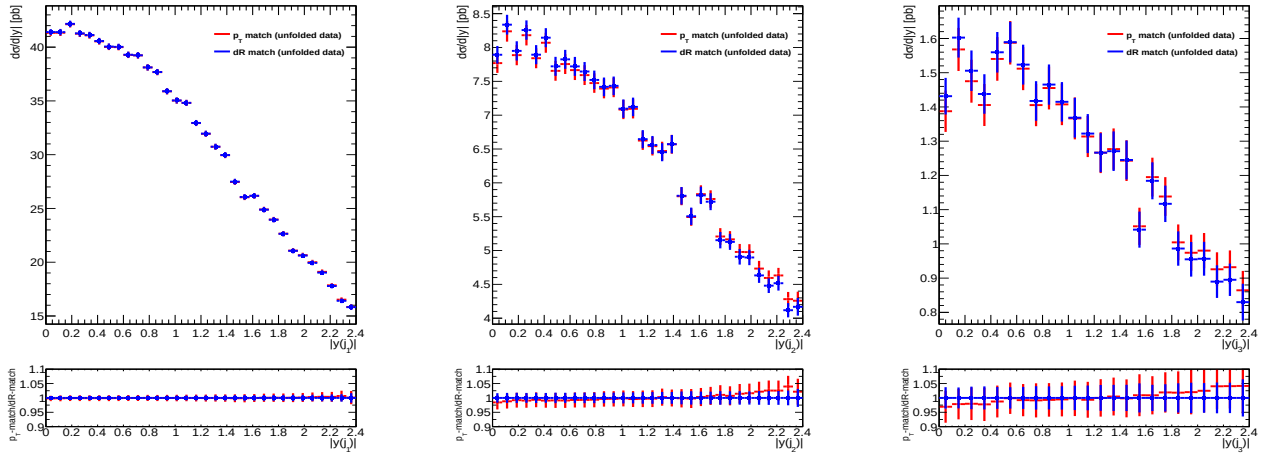


Figure 4.80: The measured differential cross section as a function of the leading jet (left), the second jet (middle), and the third jet (right) absolute rapidity in $\mu^+\mu^-$ decay channel at 8 TeV, comparing p_T order matching and angular matching scenarios. The error bars stand for statistical uncertainties.

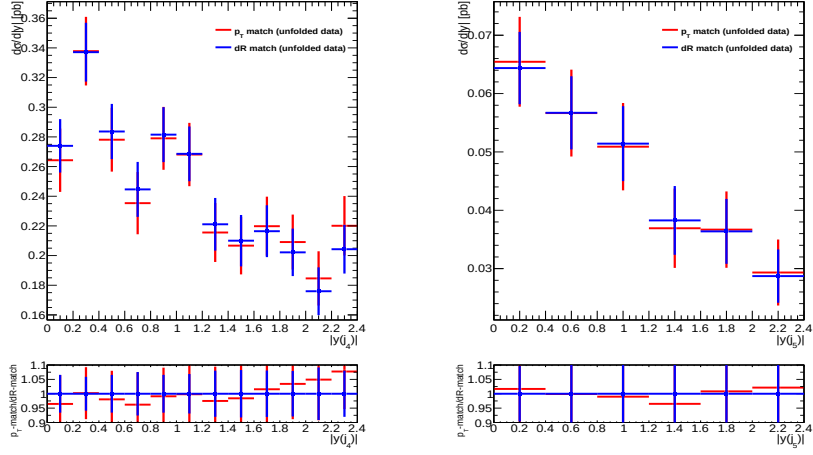


Figure 4.81: The measured differential cross section as a function of the forth jet (left) and the fifth jet (right) rapidity in $\mu^+\mu^-$ decay channel at 8 TeV, comparing p_T order matching and angular matching scenarios. The error bars stand for statistical uncertainties.

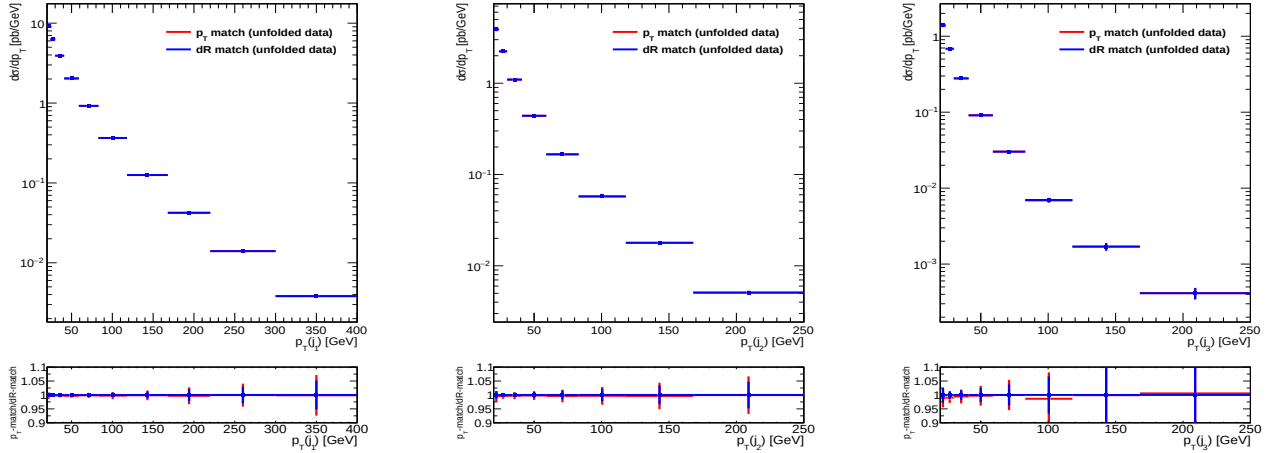


Figure 4.82: The measured differential cross section as a function of the leading jet (left), the second jet (middle), and the third jet (right) p_T in $\mu^+\mu^-$ decay channel at 13 TeV, comparing p_T order matching and angular matching scenarios. The error bars stand for statistical uncertainties.

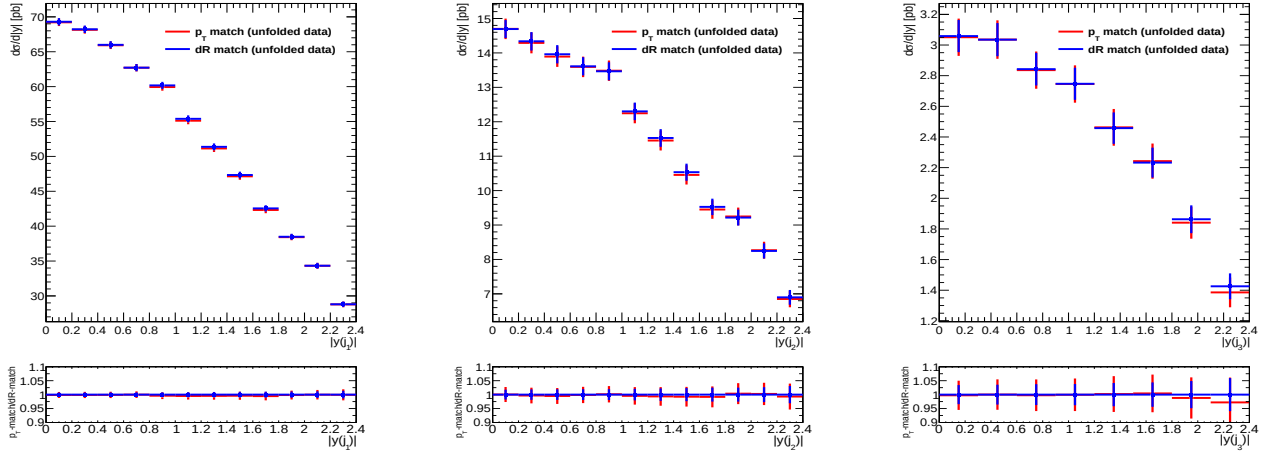


Figure 4.83: The measured differential cross section as a function of the leading jet (left), the second jet (middle), and the third jet (right) absolute rapidity in $\mu^+\mu^-$ decay channel at 13 TeV, comparing p_T order matching and angular matching scenarios. The error bars stand for statistical uncertainties.

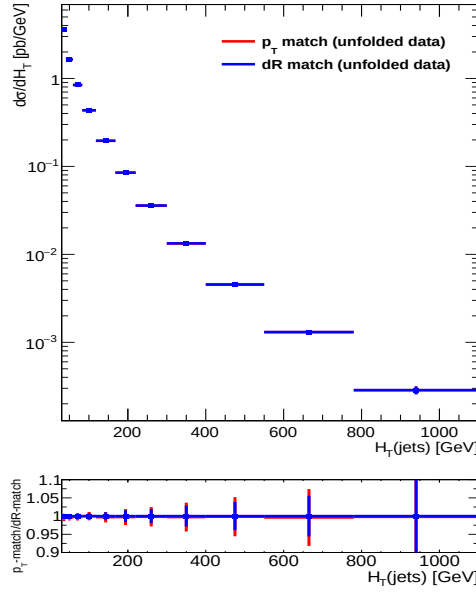


Figure 4.84: The measured differential cross section as a function of the scalar sum of jets p_T in $\mu^+\mu^-$ decay channel at 13 TeV, comparing p_T order matching and angular matching scenarios. The error bars stand for statistical uncertainties.

4.7 Uncertainties

The uncertainties in this analysis are induced by the systematic errors on measurement, the finite statistics of the data and MC samples, and the uncertainty in the estimation of theoretical models. The first two sources lead to the systematic uncertainty and statistical uncertainty respectively, and they will be propagated analytically to the measured differential cross sections through the covariance matrices computation using the unfolding approach. Meanwhile, the statistical uncertainty of signal MC samples and the theoretical uncertainty will be also propagated to the predicted differential cross sections. It needs to be noted that the theoretical uncertainty computation is not available for every theoretical prediction in this analysis, due to some technical constraints of MC productions.

Considered systematic uncertainties sources and propagations are:

- Jet energy scale (JES)

By rescaling the reconstructed jet p_T distribution up and down in data, one can compute the difference between each of them and the nominal spectrum as the uncertainty. The rescaling factors (computed by CMS collaboration) for up and down variations are assigned as jet p_T and η dependent numbers, and they vary from 1% up to 5% within the kinematic selection constraint. The difference is propagated into the differential cross section computation. More details can be found in [91] and [99].

- Jet energy resolution (JER)

As mentioned before, the jet resolution estimated using MC simulation is a bit higher than the one for data, so the correction has to be taken into account for the MC case. The solution is to smear the jet momentum, which has been introduced in equation 4.6. This effect is propagated by smearing the response matrix of the signal MC sample. Actually, the uncertainty is small (within 1%). More details about this correction have been interpreted in [89] and [90].

- Pileup (PU)

The pileup modelling uncertainty is estimated through varying the minimum bias cross section by $\pm 5\%$. The effect is within 1% except for the large jet multiplicity region.

- Luminosity (Lumi)

The integrated luminosities are measured at 8 TeV [100] and at 13 TeV [101] respectively. For each center-of-mass energy, since all the MC distributions are normalized to the integrated luminosity, the uncertainty from the luminosity should be propagated to all of them. In this analysis, the distributions of all the simulated processes are varied up and down by 2.6% at 8 TeV or by 2.7% at 13 TeV, which will be propagated to the uncertainties of the measured differential cross sections. It has to be noted that the propagated luminosity uncertainty is dependent on the cross sections of simulated background processes.

- Background cross section (Bgnd)

According to the control plots which are shown before, the $t\bar{t}$ and diboson processes dominate the background contributions. Since they are produced by simulation and normalized to the luminosity of data, the subtraction procedure for data can involve in the uncertainties, one of which results from the cross section predicted by theoretical modelling, while the other from the integrated luminosity. The uncertainty resulting from luminosity has been taken into account before. For the uncertainty of theoretical modelling, we implement a variation on the distribution of $t\bar{t}$ by 10% [102] (at 8 TeV analysis) and independent variations on diboson processes (WW/WZ/ZZ) respectively by 3%. For 13 TeV analysis, through selecting a μ and an electron for data and simulation samples, we obtain the distribution of $t\bar{t}$ process, and reweigh the simulated distribution according to the data. The uncertainty on the reweighting factor is treated as the systematic uncertainty coming from $t\bar{t}$ modelling.

- Lepton efficiency correction (Eff)

Considering the corrections applied for lepton trigger, reconstruction, identification, and isolation, we conservatively implement a global variation for each nominal distribution of the signal process that is predicted by MC. The variation factor is $\pm 2.5\%$ for $\mu^+\mu^-$ channel and $\pm 2.6\%$ for e^+e^- channel. The propagated uncertainty is smaller than 5%. More explanations can be found in [78] and [87].

- Lepton energy scale (LES)

Similar to the method for estimating JES uncertainty, we rescale the simulated signal distribution up and down by 0.2% [78] for $\mu^+\mu^-$ channel and 0.6% (EB) / 1.5% (EE) ¹¹ [87] for e^+e^- channel. After propagation, this uncertainty becomes quite negligible ($< 0.5\%$).

- Lepton energy resolution (LER)

This uncertainty is assessed by smearing the momenta of leptons in emulated signal sample, by means of the equation similar to 4.6, where all the variables correspond to the leptons. Note that the smearing factor C now is a random Gaussian distributed value, of which the nominal value is 1 with a width of 0.6% for $\mu^+\mu^-$ channel and 6% for e^+e^- channel. This uncertainty is very small ($< 0.5\%$) after propagation to the differential cross sections.

- Unfolding (Unf)

In the analysis at 8 TeV, we use the sample produced by MADGRAPH (LO) to estimate the unfolding matrices. Since the unfolding matrices are dependent on the chosen MC sample, different choices of MC models may lead to different results. Alternatively, we use another sample produced by SHERPA 1.4, which is also a MC generator in LO accuracy while with a different model of parton shower and hadronization. Through comparing the unfolding matrices between the two samples, we treat the difference between them as the systematic uncertainty related to

¹¹EB is for barrel region while EE is for endcap region.

unfolding (Unf sys). Regarding 13 TeV, the sample simulated by MADGRAPH_AMC@NLO is used for unfolding procedure, and there is no another suitable sample for unfolding matrices comparison. Hence we use another method to estimate it, through weighting the simulation to be consistent with data in each distribution and constructing a re-weighted response matrix to unfold the data distribution. The reweighting is done using a fine binning allowing to be sensitive to different shapes within bins. The difference between the unfolded results obtained using the weighted response matrix and the nominal results is taken as the systematic uncertainty associated with the unfolding response matrix. Moreover, the statistical uncertainties of MADGRAPH at 8 TeV and of MADGRAPH_AMC@NLO at 13 TeV are respectively treated as the unfolding statistical uncertainties (Unf stat).

These systematic uncertainties above are summed in quadrature through the assumption that each uncertainty source is independent from others, yielding to the total systematic uncertainty. The covariance matrices associated to each systematic source are calculated by referring to [103], which is applied in the combination of the two decaying channels (see section 4.10.1). In each matrix, the diagonal elements are the squares of the systematic uncertainties listed above:

$$\Delta\sigma_{ii} = \Delta\sigma_i^2, \quad (4.10)$$

where $\Delta\sigma_i = |\sigma_i^{up} - \sigma_i^{down}|/2$, and σ_i corresponds to the effect on cross section of each systematic. The off diagonal members are calculated based on the hypothesis of 100% correlation between each systematic source, also considering the sign of the source variation difference:

$$\Delta\sigma_{ij} = \text{sign}(\sigma_i^{up} - \sigma_i^{down}) * \text{sign}(\sigma_j^{up} - \sigma_j^{down}) * \Delta\sigma_i * \Delta\sigma_j, \quad (4.11)$$

Different systematic sources actually have different effect on the measured differential cross section, as illustrated from figure 4.85 to 4.90 at 8 TeV. One can see that the dominant uncertainty is from jet energy scale variation. It increases significantly when the jet multiplicity is larger, while it more or less maintains flatness for the leading jet p_T . The luminosity uncertainty and muon efficiency uncertainties are almost flat. The uncertainties resulting from the background modelling and pileup variations are enlarged at high jet multiplicity region. The other uncertainties are almost negligible.

The systematic effects at 13 TeV are shown from figure 4.91 to 4.92. One can see the similar performances to the case of 8 TeV. The uncertainty from jet energy scale variation increases when the jet multiplicity becomes higher, while it keeps flat for the distribution of leading jet p_T . The systematic uncertainty due to the background modelling is only apparent in large jet multiplicity region. The other systematic uncertainties are negligible, so they are not shown in the plots.

Regarding the MADGRAPH_AMC@NLO generator, the theoretical uncertainty and PDF uncertainty need to be computed as well. The generated sample contains eight weight values for each event, corresponding to eight combinations of factorization (μ_F) and renormalization (μ_R) scales. According to the convention, each scale is varied as half or twice of the nominal (default) value [104].

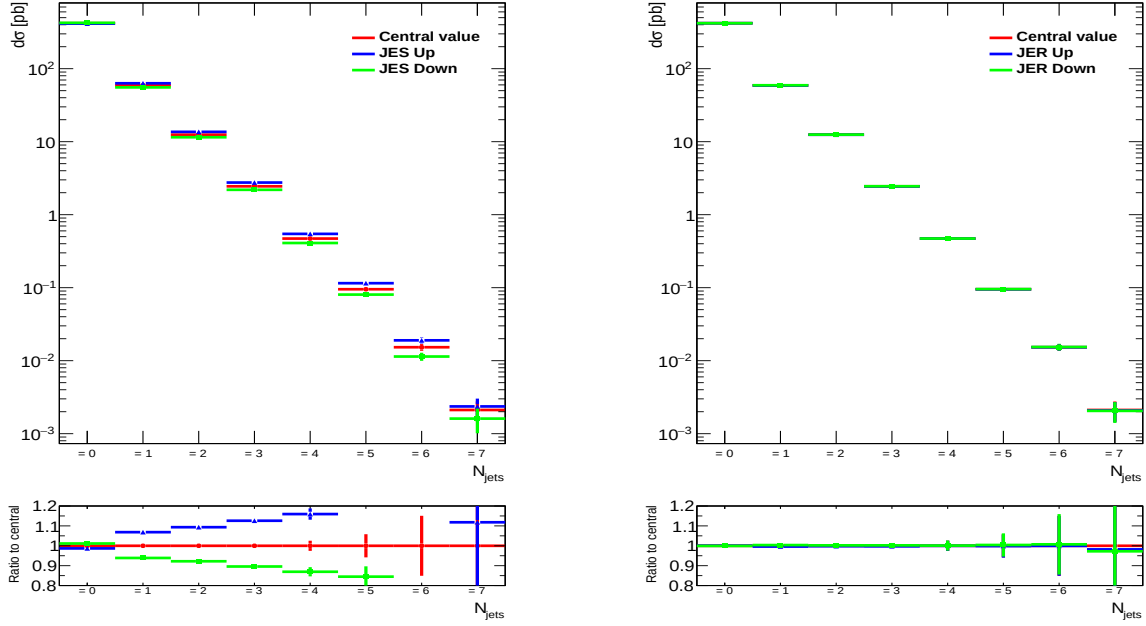


Figure 4.85: The systematic effects on the measured differential cross section as a function of exclusive jet multiplicity in $\mu^+\mu^-$ decay channel at 8 TeV, due to the variations of jet energy scale (left) and jet energy resolution (right).

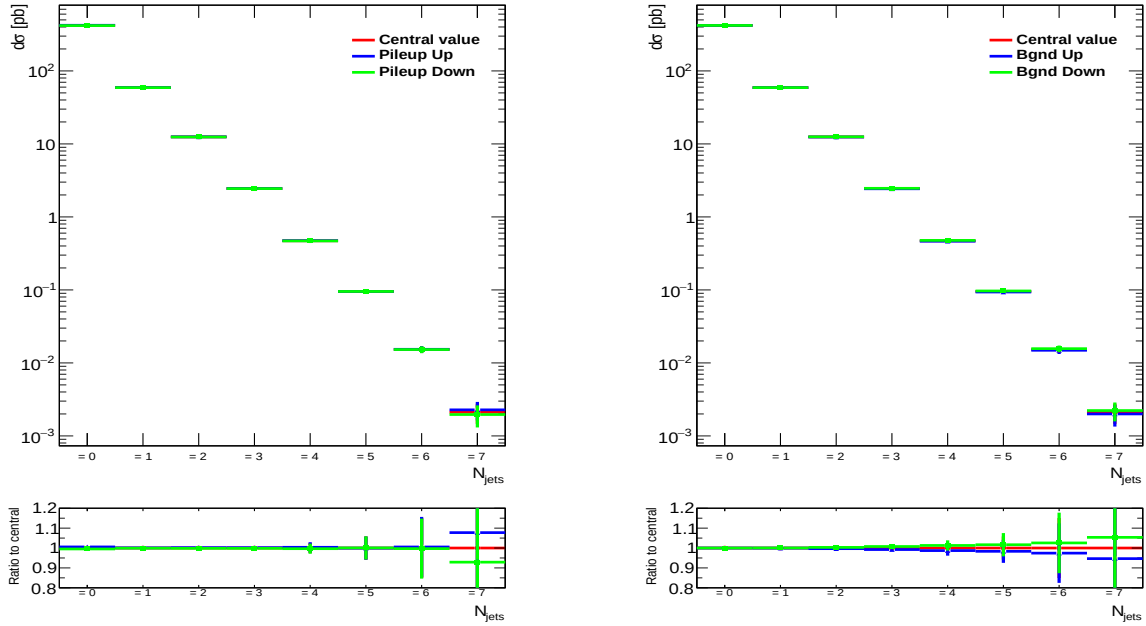


Figure 4.86: The systematic effects on the measured differential cross section as a function of exclusive jet multiplicity in $\mu^+\mu^-$ decay channel at 8 TeV, due to the variations of pileup (left) and background modelling (right).

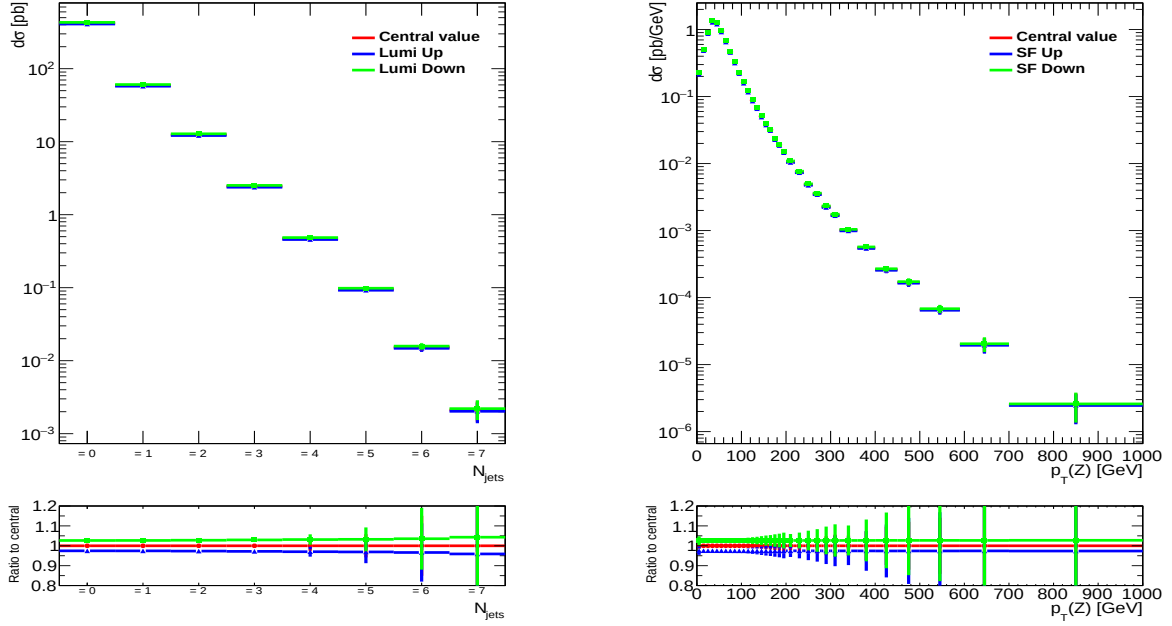


Figure 4.87: The systematic effects on the measured differential cross section as a function of exclusive jet multiplicity (left) and Z boson p_T (right) in $\mu^+\mu^-$ decay channel at 8 TeV, due to the variations of luminosity (left) and lepton efficiency correction (right).

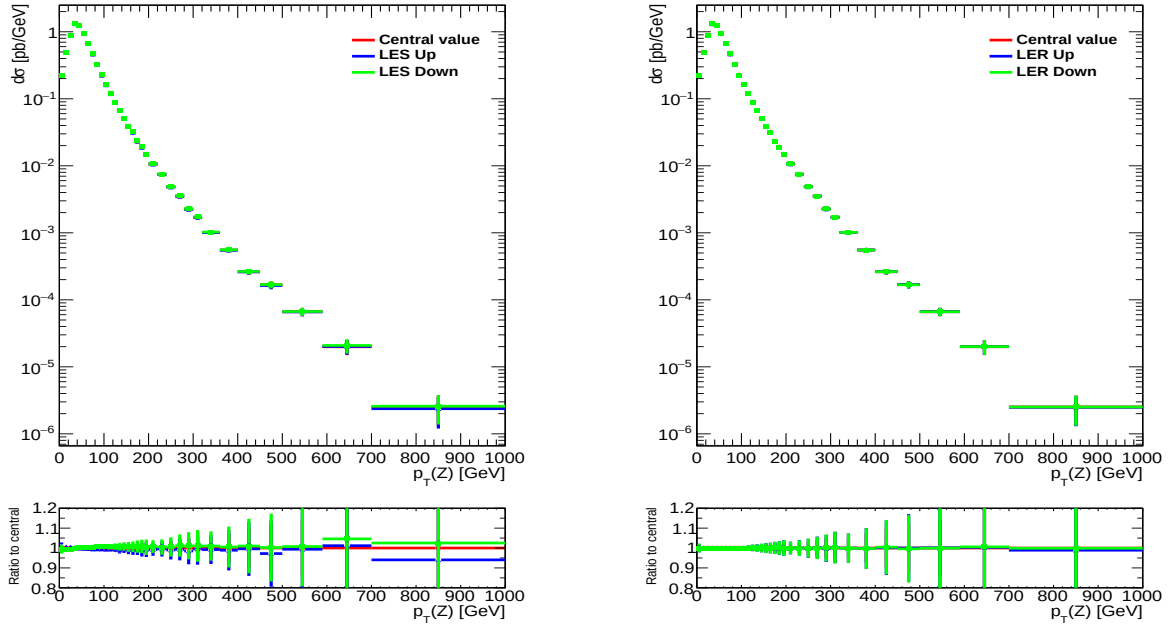


Figure 4.88: The systematic effects on the measured differential cross section as a function of Z boson p_T in $\mu^+\mu^-$ decay channel at 8 TeV, due to the variations of lepton energy scale (left) and lepton energy resolution (right).

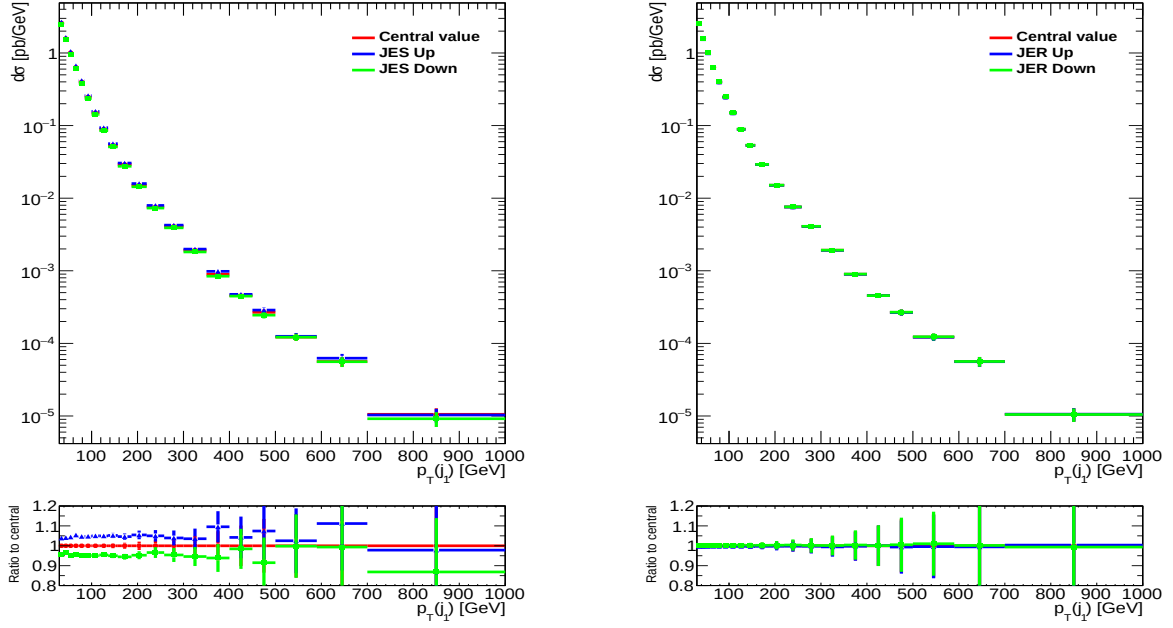


Figure 4.89: The systematic effects on the measured differential cross section as a function of leading jet p_T in $\mu^+\mu^-$ decay channel at 8 TeV, due to the variations of jet energy scale (left) and jet energy resolution (right).

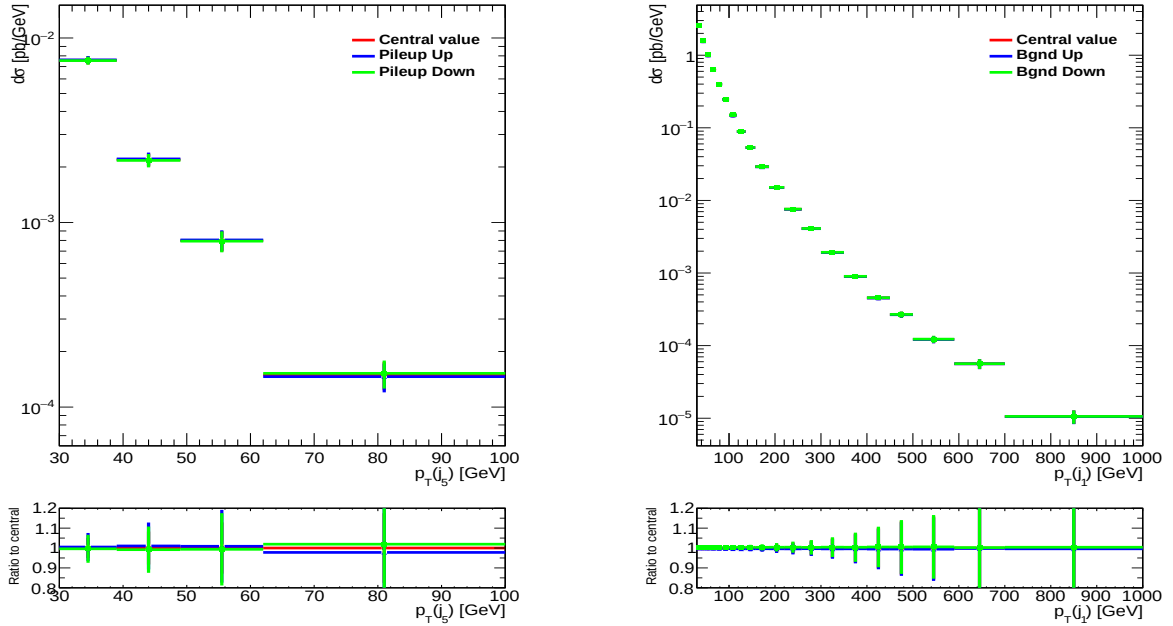


Figure 4.90: The systematic effects on the measured differential cross section as a function of the fifth jet p_T (left) and the leading jet p_T (right) in $\mu^+\mu^-$ decay channel at 8 TeV, due to the variations of pileup (left) and background modelling (right).

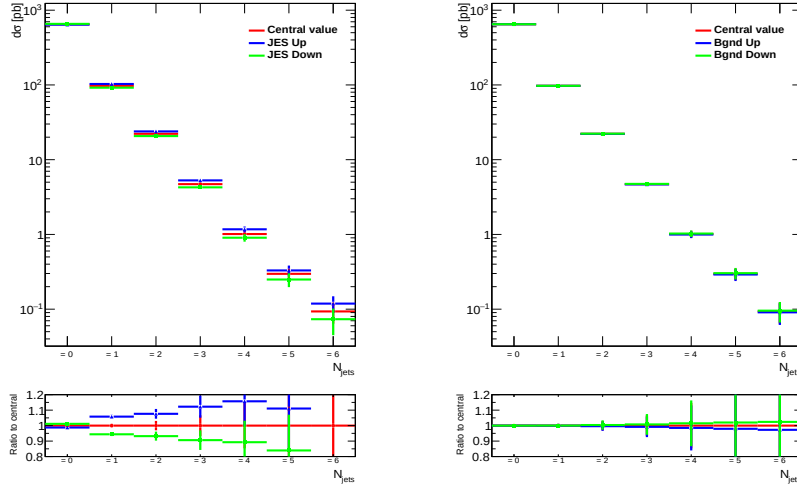


Figure 4.91: The systematic effects on the measured differential cross section as a function of exclusive jet multiplicity in $\mu^+\mu^-$ decay channel at 13 TeV, due to the variations of jet energy scale (left), and background modelling (right).

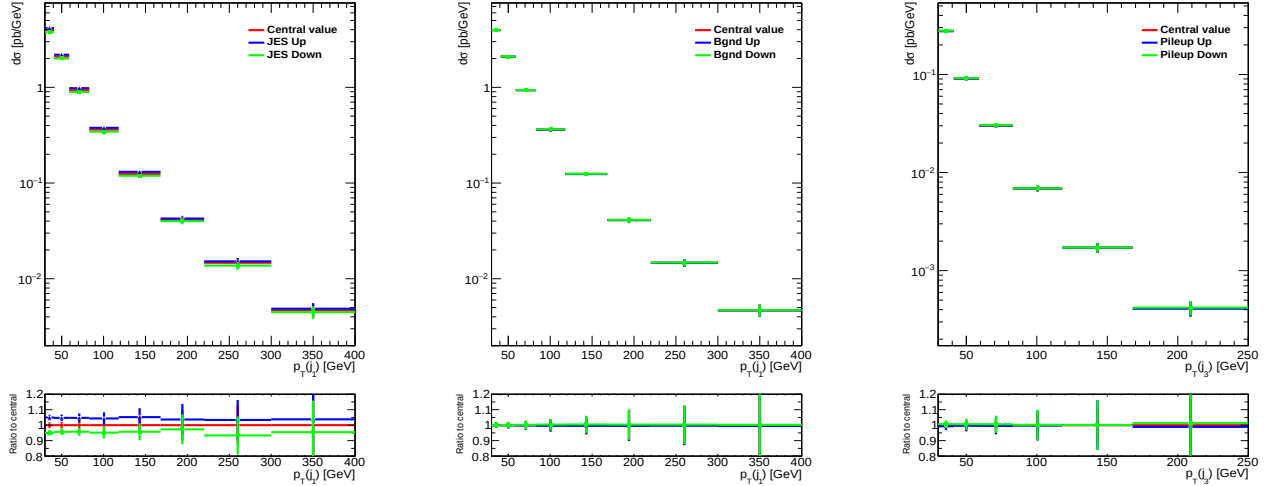


Figure 4.92: The systematic effects on the measured differential cross section as a function of leading jet p_T (left and middle) and third jet p_T (right) in $\mu^+\mu^-$ decay channel at 13 TeV, due to the variations of jet energy scale (left), background modelling (middle), and pileup (right).

Since it overestimates the uncertainty to have the extreme case with one scale varying to the half while the other to twice, there are six combinations considered. We calculate the nominal differential cross sections that are predicted by this generator using the default weight value of each event and the weight values corresponding to the six scale combinations, respectively. Then we take the envelope of the variation, which is with the maximum up/down variation with respect to the default distribution, to compute the theoretical uncertainty. The calculated theoretical uncertainties at 8 TeV and 13 TeV are illustrated in figure 4.93 and 4.94, respectively. One can see that the uncertainties are within

15%.

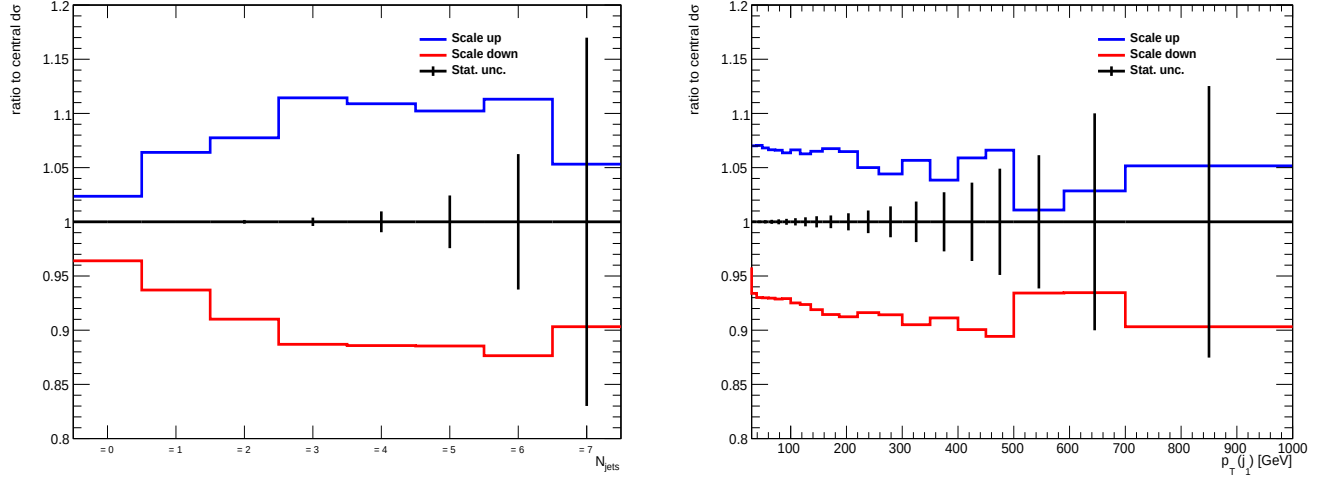


Figure 4.93: The theoretical (scale) uncertainty on the predicted differential cross section as a function of jet multiplicity (left) and leading jet p_T (right) in $\mu^+\mu^-$ decay channel at 8 TeV, due to the variations of factorization and renormalization. The predictions are provided by MADGRAPH_AMC@NLO generator.

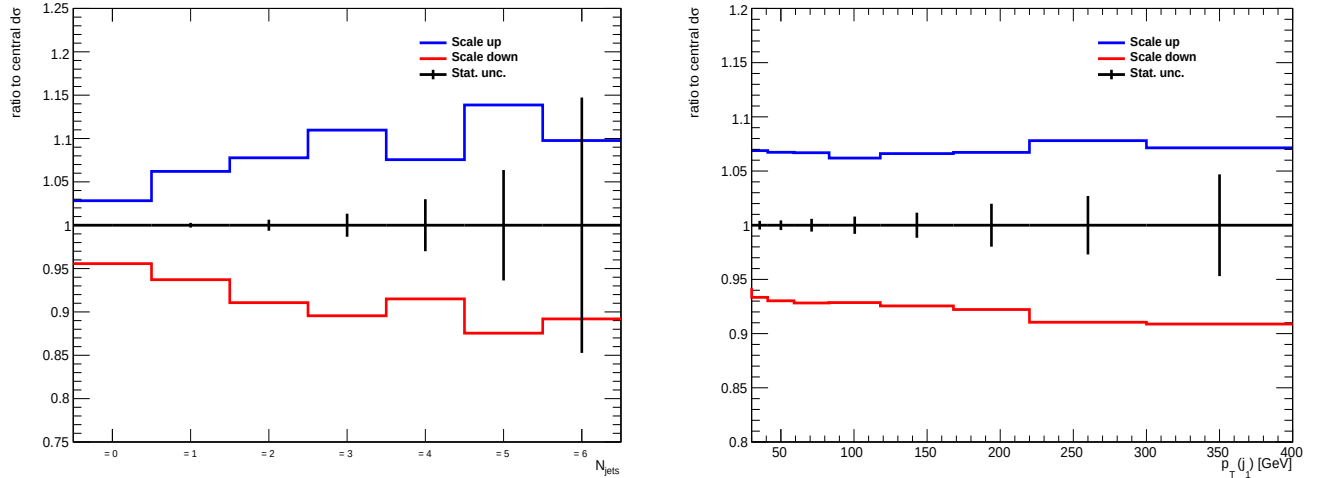


Figure 4.94: The theoretical (scale) uncertainty on the predicted differential cross section as a function of jet multiplicity (left) and leading jet p_T (right) in $\mu^+\mu^-$ decay channel at 13 TeV, due to the variations of factorization and renormalization. The predictions are provided by MADGRAPH_AMC@NLO generator.

The parton distribution function (PDF) sets of the MADGRAPH_AMC@NLO sample is NNPDF 3.0 [105], which is estimated at next-to-leading-order level. The value of α_s is chosen to be 0.118. Each event in this sample comes with several different weights, which are corresponding to the default prediction for parton distribution function and varied PDF settings (replicas) respectively. Typically,

there are 100 different weights that are floating around the nominal weight, which are marked as w_k ($k = 0, 1, 2, \dots, 100$), where w_0 stands for the nominal weight. As a result, one can obtain 100 different cross sections using the various replicas, and the PDF uncertainty is computed using the RMS [106]:

$$\text{RMS} = \sqrt{\frac{1}{N_{\text{rep}} - 1} \times \sum_{k=1}^{N_{\text{rep}}} (F[q(k)] - \langle F[q] \rangle)^2}, \quad (4.12)$$

where the N_{rep} represents the number of PDF replicas (100), and $F[q(k)]$ stands for the cross section with each replica weight, and $\langle F[q] \rangle$ is the average value of the cross sections using the 100 replica weights. The RMS is taken as the absolute PDF uncertainty, and the relative PDF uncertainty will be RMS divided by the cross section computed using nominal weight. Figure 4.95 and 4.96 show the uncertainties on the predicted differential cross sections because of the PDF variations at 8 TeV and 13 TeV, respectively, from which one can see that the uncertainty is no more than 10%. The uncertainty also contains the contribution of α_s uncertainty, which is estimated through varying the α_s value with ± 0.001 and calculating their difference in measured cross section. It is combined with the two other uncertainties by quadratic sum.

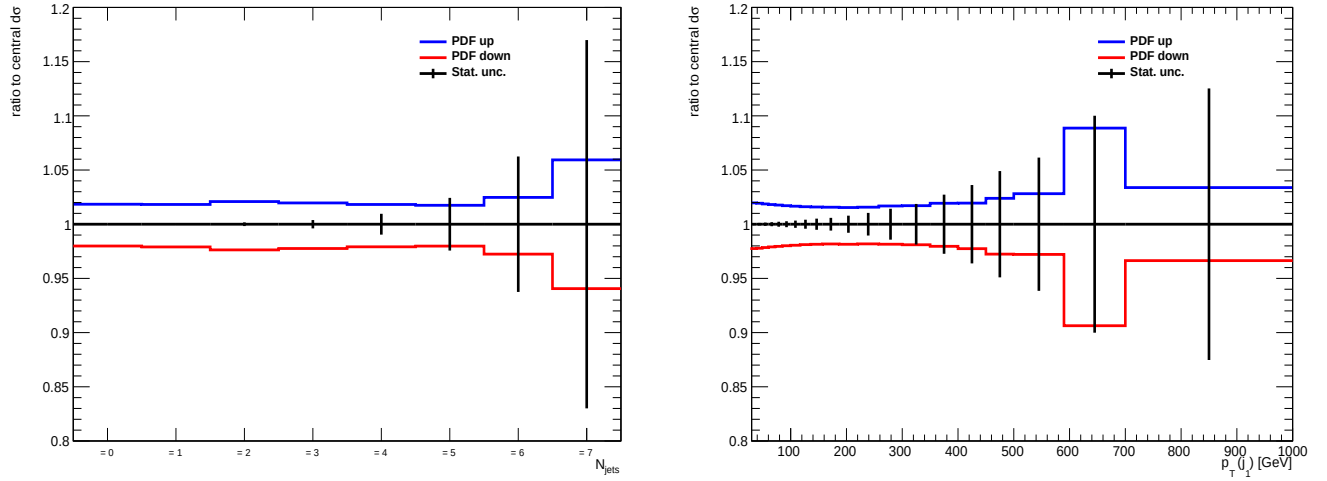


Figure 4.95: The uncertainty of the predicted differential cross section as a function of jet multiplicity (left) and leading jet p_T (right) in $\mu^+\mu^-$ decay channel at 8 TeV, due to the variations of PDF. The predictions are provided by MADGRAPH_AMC@NLO generator.

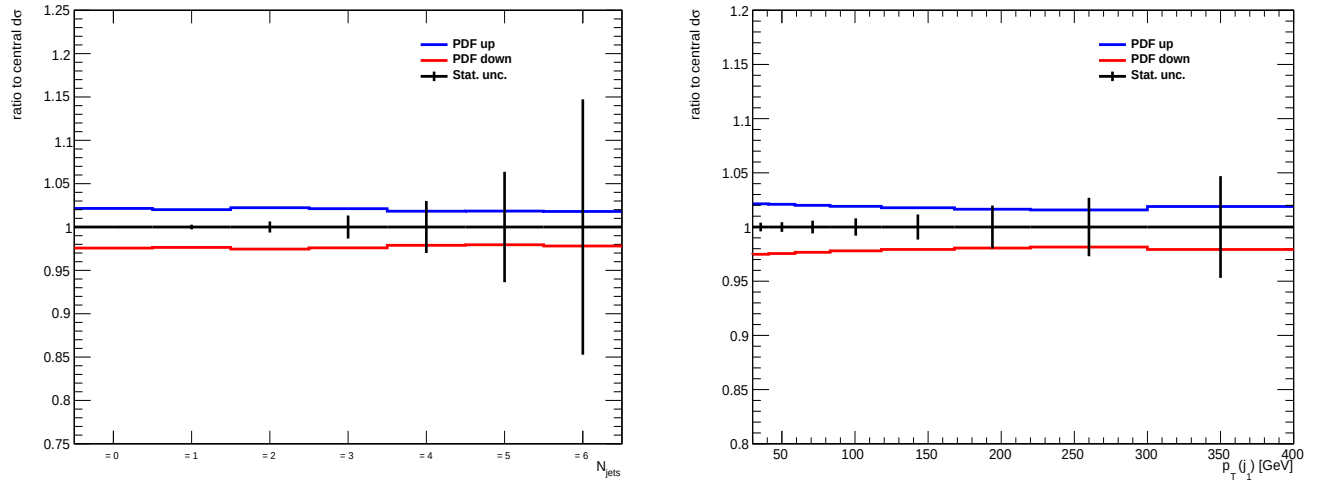


Figure 4.96: The uncertainty of the predicted differential cross section as a function of jet multiplicity (left) and leading jet p_T (right) in $\mu^+\mu^-$ decay channel at 13 TeV, due to the variations of PDF. The predictions are provided by MADGRAPH_AMC@NLO generator.

4.8 Non-perturbative correction

Regarding the 13 TeV analysis, we have the theoretical predictions for Z boson plus one jet at next-to-next-leading-order (NNLO) level [104, 107], concerning the differential cross section as a function of p_T and $|y|$ of the leading jet, and H_T . These computations are at parton level restricted to perturbative QCD calculation, while not considering the hadronization process and multiple parton scattering (MPI) effect, which contains non-perturbative QCD region. On the other hand, our measurement is based on the data obtained from the detector, where the hadronization and MPI processes are involved. At this point, we quantify the non-perturbative correction as differential factors to apply on the predicted differential cross sections. The calculation for the correction is realized using Rivet [108] framework, which allows for the validation and tuning of MC event generators.

We choose MADGRAPH_AMC@NLO as the MC generator for the estimation. It is interfaced with PYTHIA8 [23], which can implement parton shower, hadronization, and MPI process. Through switching on/off the hadronization and MPI processes with PYTHIA8 during the simulation, one can obtain the events in parton level and hadron level, respectively. Then Rivet can compute the differential cross sections. It has to be noted that Rivet makes use of the FastJet anti- k_T algorithm [20, 18] to form jets from the outgoing particles. The phase space for the cross section of Z boson plus jets process are as same as the ones introduced in section 4.3.

Figure 4.97 and figure 4.98 show the predicted differential cross sections with or without the non-perturbative correction, from which one can see that the differences are within 7%.

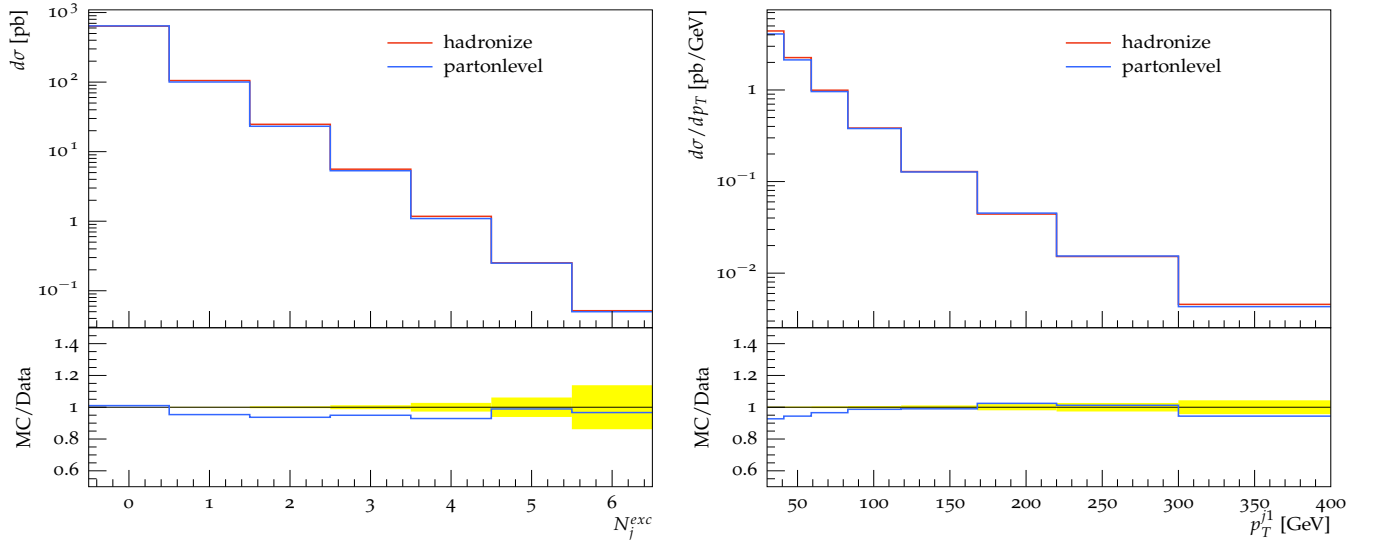


Figure 4.97: Comparison of differential cross section as a function of exclusive jet multiplicity (left) and leading jet p_T (right) in $\mu^+\mu^-$ decay channel at 13 TeV, with and without hadronization and MPI effects, respectively. Note that the "Data" and "MC" in the bottom panel stands for the former and latter, respectively. There is no real data involved. The predictions are provided by MADGRAPH_AMC@NLO generator interfaced with PYTHIA8 in the frame of Rivet.

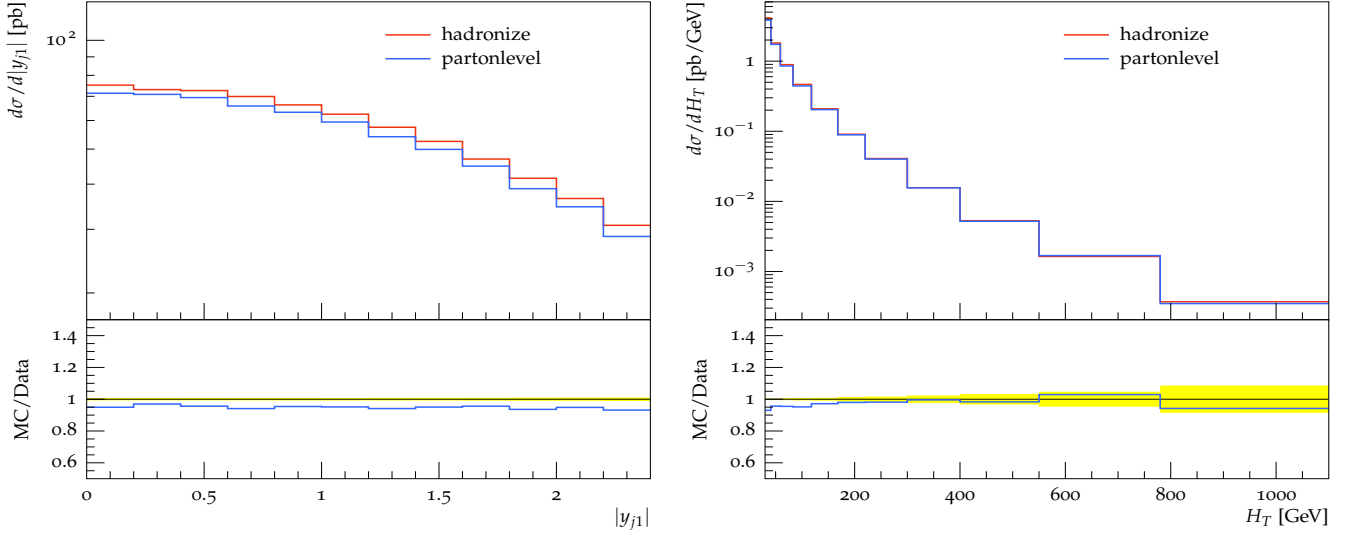


Figure 4.98: Comparison of differential cross section as a function of leading jet absolute rapidity (left) and jets H_T (right) in $\mu^+\mu^-$ decay channel at 13 TeV, with and without hadronization and MPI effects, respectively. Note that the "Data" and "MC" in the bottom panel stands for the former and latter, respectively. There is no real data involved. The predictions are provided by MADGRAPH_AMC@NLO generator interfaced with PYTHIA8 in the frame of Rivet.

Since this study is done by Rivet, it is important to check the consistency between Rivet and MADGRAPH_AMC@NLO generator. We compare the differential cross sections predicted by Rivet and MADGRAPH_AMC@NLO generator, respectively. As shown in 4.99 and 4.100, we can see that they are indeed compatible within their statistical uncertainties.

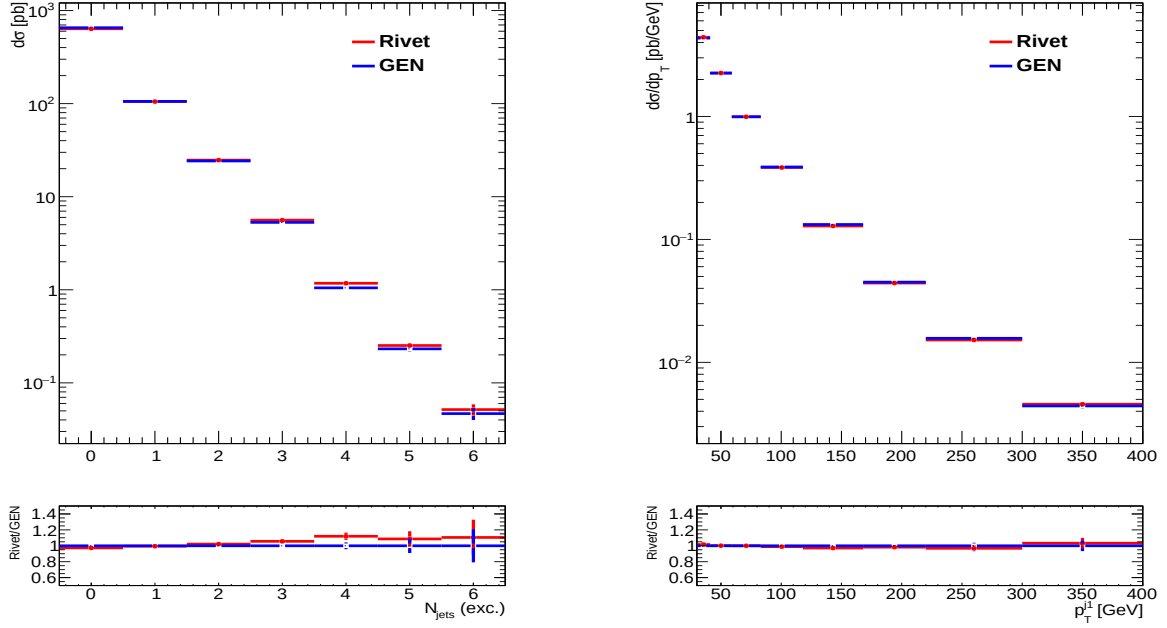


Figure 4.99: The comparison of differential cross section as a function of exclusive jet multiplicity (left) and leading jet p_T (right) in $\mu^+\mu^-$ decay channel at 13 TeV, with Rivet and MADGRAPH_AMC@NLO (GEN), respectively.

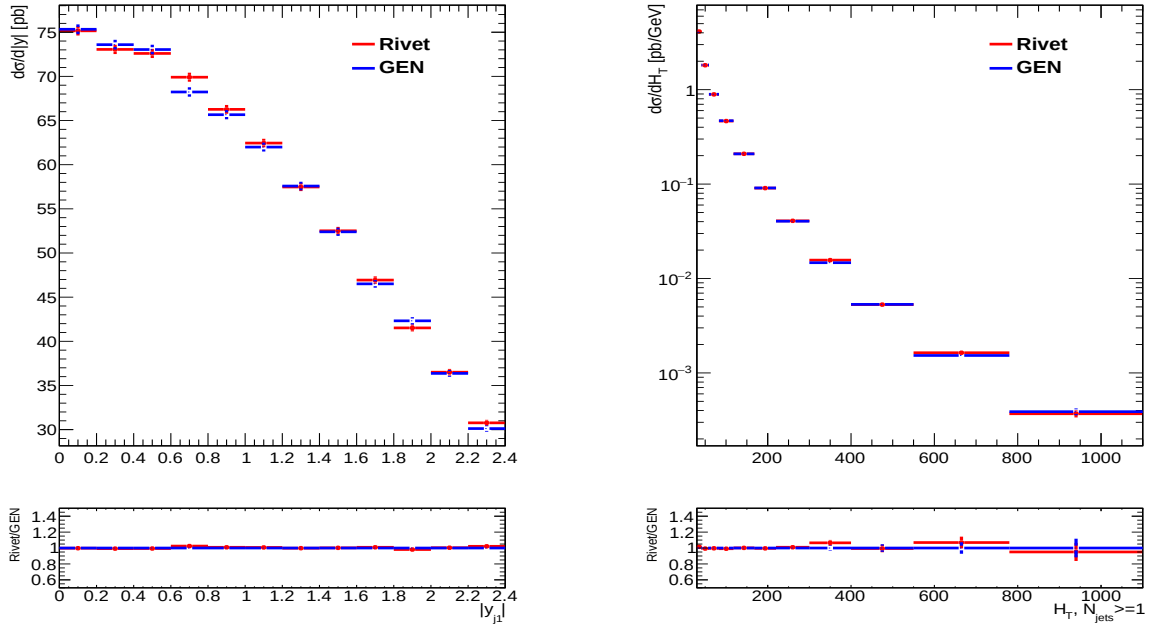


Figure 4.100: The comparison of differential cross section as a function of leading jet absolute rapidity (left) and jets H_T (right) in $\mu^+\mu^-$ decay channel at 13 TeV, with Rivet and MADGRAPH_AMC@NLO (GEN), respectively.

4.9 Differential cross sections in a single channel

After all the background contaminations are subtracted from the data, and unfolding procedure is implemented for the data, one can obtain the differential cross sections as functions of observables in $\mu^+\mu^-$ and e^+e^- decay channels, respectively, through dividing the number of events by the integrated luminosity and bin width for each bin.

4.9.1 Comparison of single channels

The measurements in two leptonic decay channels should be consistent, and they are combined to reduce the statistical uncertainty. Before that, the compatibility between the results of two decay channels should be checked. Therefore, the comparison of the two decay channels at 8 TeV in the differential cross sections and their ratios, together with their statistical uncertainties are illustrated from figures 4.101 to 4.102.

Regarding the differential cross section as a function of jet multiplicity, the two decay channels are compatible. The statistical fluctuation becomes large at high jet multiplicity region. In terms of the distribution of leading jet p_T , there is a discrepancy between the two channels which increases up to about 15% at 240 GeV within the statistical uncertainty, while overcome by systematic uncertainty. The statistical uncertainty covers the difference at higher p_T region. For the distributions of leading jet absolute rapidity and scalar sum of jets p_T by requiring at least one jet, there is no apparent discrepancy between the two decay channels.

Figure 4.103 and 4.104 show the single channel comparisons at 13 TeV. There is an apparent discrepancy at large jet multiplicity region ($N_{jets} \geq 5$) within statistical uncertainty. In terms of the leading jet p_T distribution, there is a discrepancy about 8% at around 200 GeV. Regarding the scalar sum of jets p_T distribution, an inconsistency arises to 15% from 300 GeV to 400 GeV. These discrepancies are compensated by systematic uncertainties. For the other distributions, no discrepancy is seen concerning the statistical uncertainties.

4.9.2 Results of single channels at 8 TeV

The measured differential cross sections, with full systematic and statistical uncertainties, are presented in this section. To compare the measurement, three different theoretical predictions are involved:

- MADGRAPH: The matrix element computed at tree level (LO) up to four partons in the final state, using the CTEQ6L1 [109] PDF and interfaced with PYTHIA6 Tune Z2* for parton shower and hadronization. The inclusive cross section is normalized to the NNLO cross section predicted by FEWZ (see section 4.2). The statistical uncertainty of this generator prediction is considered in the comparison.

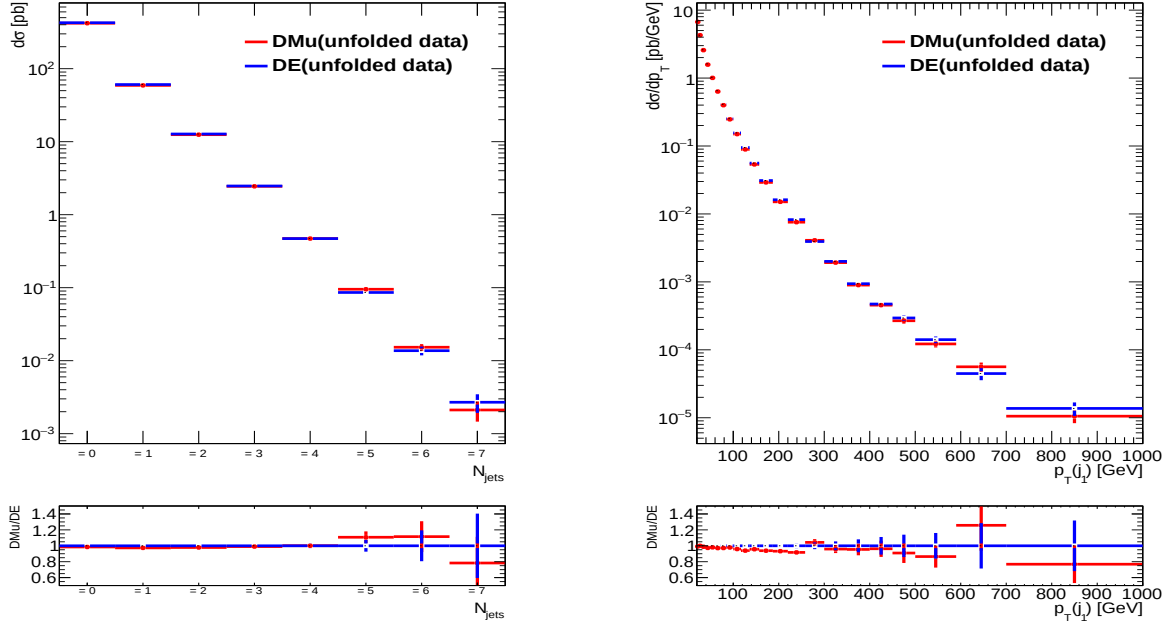


Figure 4.101: The comparison in differential cross section as a function of jet multiplicity (left) and leading jet p_T (right) of $\mu^+\mu^-$ (DMu) and e^+e^- (DE) decay channels at 8 TeV, concerning statistical uncertainties.

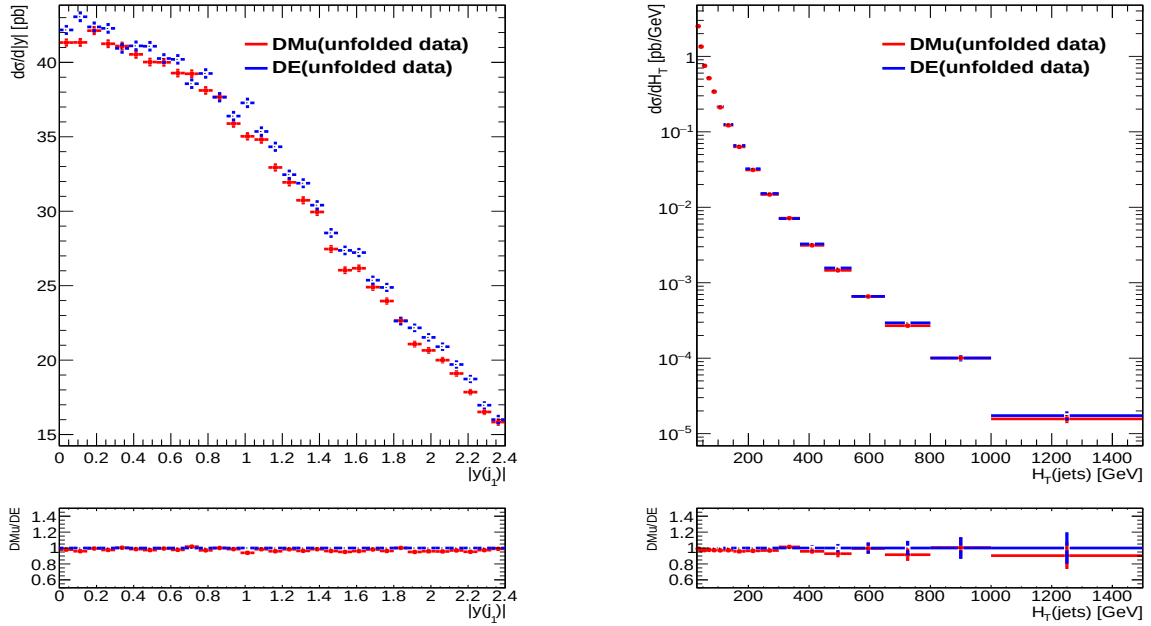


Figure 4.102: The comparison in differential cross section as a function of leading jet absolute rapidity (left) and jets H_T (right) of $\mu^+\mu^-$ (DMu) and e^+e^- (DE) decay channels at 8 TeV, concerning statistical uncertainties.

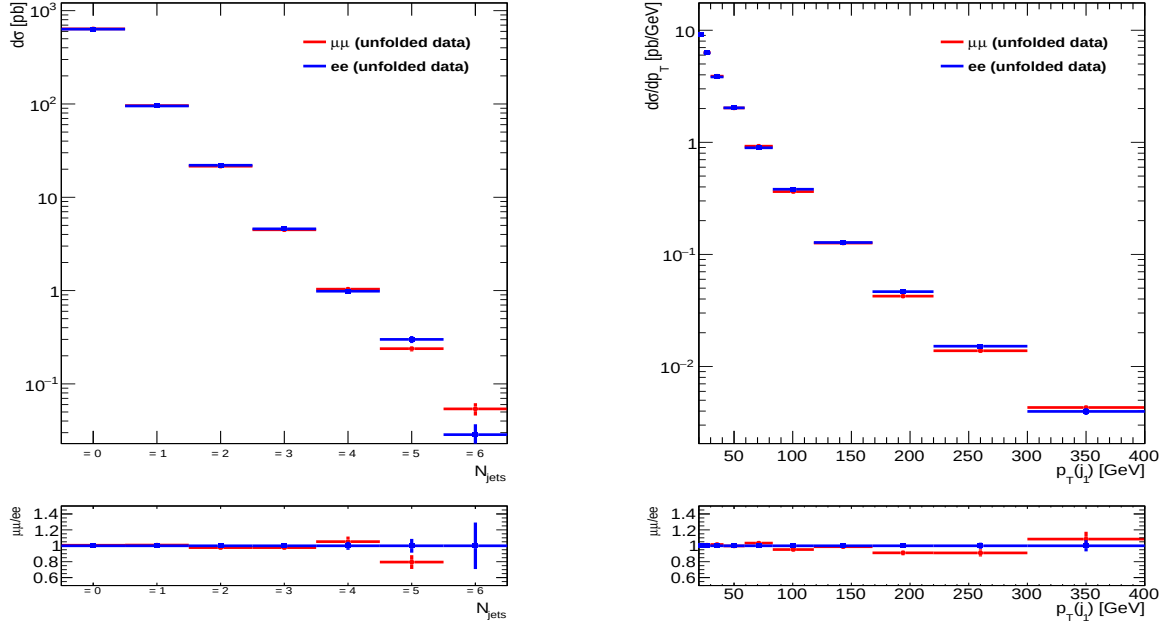


Figure 4.103: The comparison in differential cross section as a function of jet multiplicity (left) and leading jet p_T (right) of $\mu^+\mu^-$ and e^+e^- decay channels at 13 TeV, concerning statistical uncertainties.

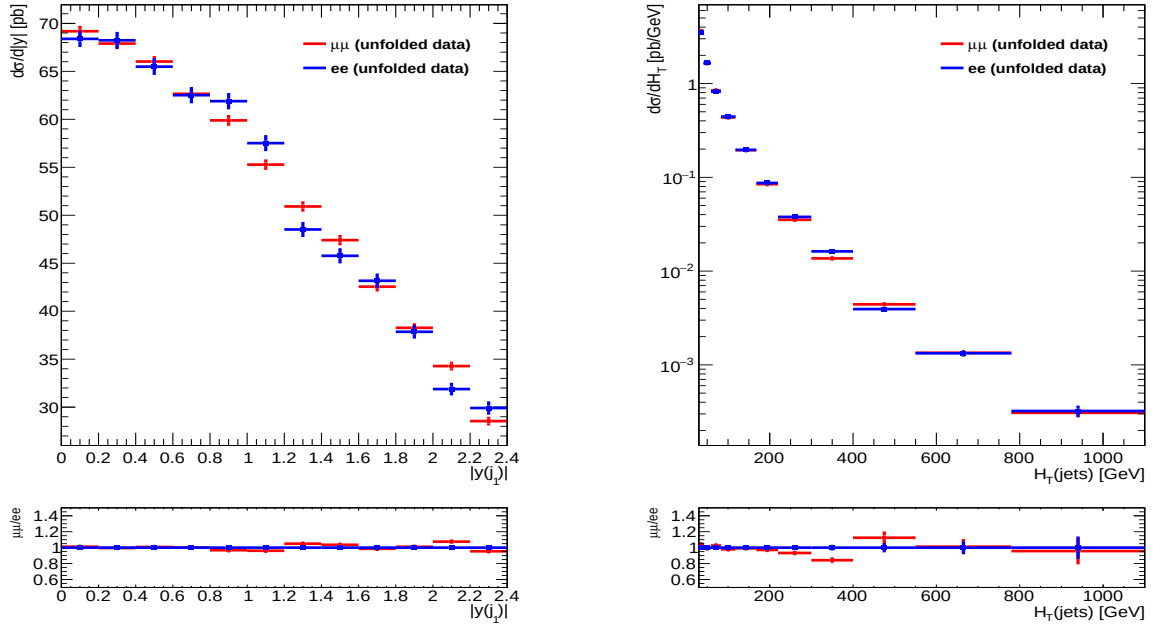


Figure 4.104: The comparison in differential cross section as a function of leading jet absolute rapidity (left) and jets H_T (right) of $\mu^+\mu^-$ and e^+e^- decay channels at 13 TeV, concerning statistical uncertainties.

- SHERPA2: The matrix element computed at NLO accuracy with 0, 1, 2 partons in the final state, and with 3, 4 partons at tree level. The parton shower and hadronization are implemented

by its own model. It includes the configuration of BlackHat [110, 111] to take into account one loop corrections. The merging of matrix element and parton shower is done with the MEPS@NLO method [41], with the merging scale of 20 GeV. It uses CT10 [11] PDF sets. The statistical uncertainty of this prediction is shown in the comparison.

- **MADGRAPH_AMC@NLO:** The matrix element calculated in NLO accuracy with 0, 1, 2 partons in the final state, and with 3 partons in tree level, using NNPDF [68]. It is interfaced with PYTHIA8 for parton shower, hadronization, and multiple parton scattering. The FxFx [40, 27] merging scheme is used with a merging scale parameter set to 30 GeV. The strong coupling constant for the matrix element computation is set to $\alpha_s = 0.118$. While for the showering and underlying events, the used value of CUETP8M1 tune for PYTHIA8 is 0.130. The statistical, theoretical (scale), and PDF uncertainties are shown in the comparison. These uncertainties are added in quadrature.

Figure 4.105 to figure 4.113 show several comparisons between measurements and predictions of the differential cross sections. In terms of the distribution of jet multiplicity, the three theoretical predictions are generally compatible with the measurement in both decay channels. The MADGRAPH_AMC@NLO underestimates the cross sections at large jet multiplicity region. Regarding the distributions of Z boson p_T and leading jet p_T , there is a common discrepancy between measurement and MADGRAPH prediction up 420 GeV approximately, while this disagreement disappears for MADGRAPH_AMC@NLO prediction with higher order calculation. The SHERPA prediction is apparently lack in statistics. For the leading jet absolute rapidity, scalar sum of jets p_T (at least one jet), and the second jet p_T and rapidity, all the MC predictions agree with data well.

It has to be noted that the differential cross sections as functions of rapidity sum/difference of Z boson and leading jet has been normalized to the total cross sections. Since we are interested in the shape comparison to validate the same measurement (see section 4.5) at 7 TeV [61] (also see section 4.11) with larger energy scale and extended observable region. With the full statistical uncertainties taken into account, we find dramatic discrepancy between MADGRAPH predictions and measurements. On the other hand, there is no disagreement generally with MADGRAPH_AMC@NLO and SHERPA2 predictions. This problem observed with 7 TeV data has been studied extensively by MADGRAPH 5 and MADGRAPH_AMC@NLO authors, adapting and improving the ME and parton shower matching (MLM approach) for MADGRAPH 5 to a new merging scheme (FxFx), which is developed for merging multiple parton multiplicity at NLO. These results confirm the success of this new approach.

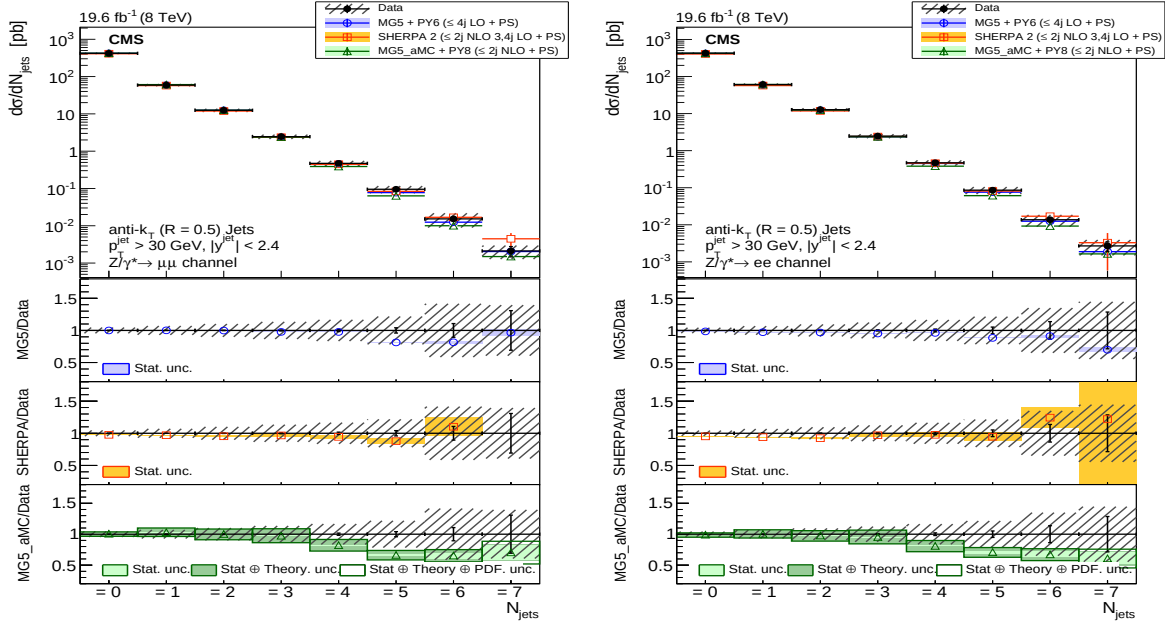


Figure 4.105: The differential cross section as a function of exclusive jet multiplicity in $\mu^+\mu^-$ (left) and e^+e^- (right) decay channels at 8 TeV, showing the full uncertainties. The measurement is compared with MADGRAPH (MG5), SHERPA2, and MADGRAPH_AMC@NLO (MG5_aMC).

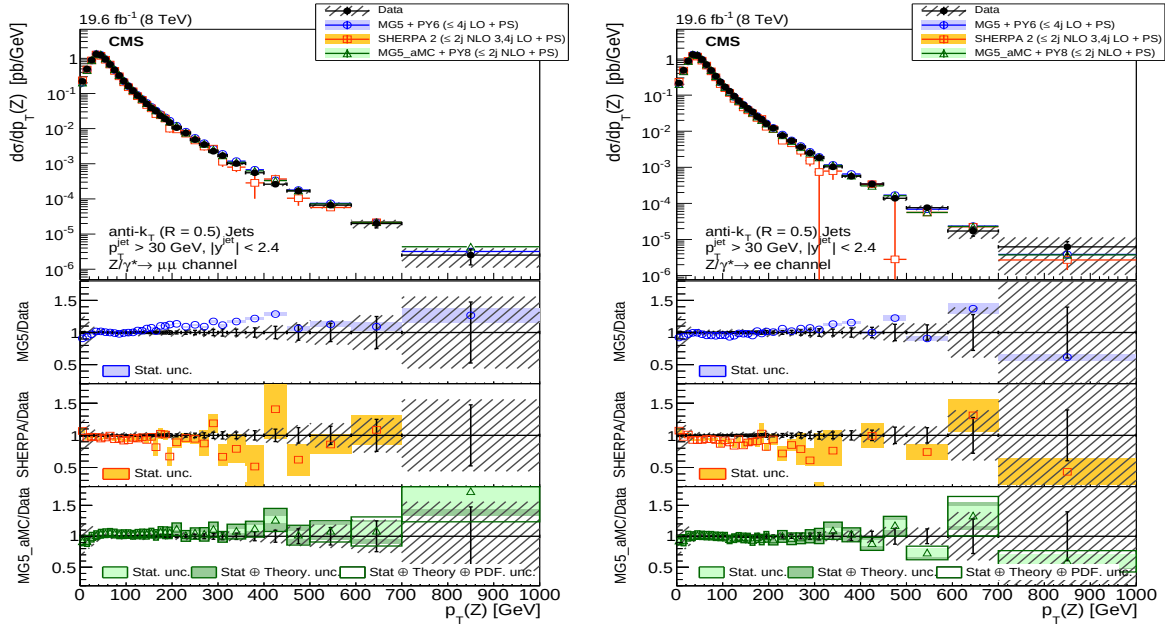


Figure 4.106: The differential cross section as a function of Z boson p_T in $\mu^+\mu^-$ (left) and e^+e^- (right) decay channels at 8 TeV, showing the full uncertainties. The measurement is compared with MADGRAPH (MG5), SHERPA2, and MADGRAPH_AMC@NLO (MG5_aMC).

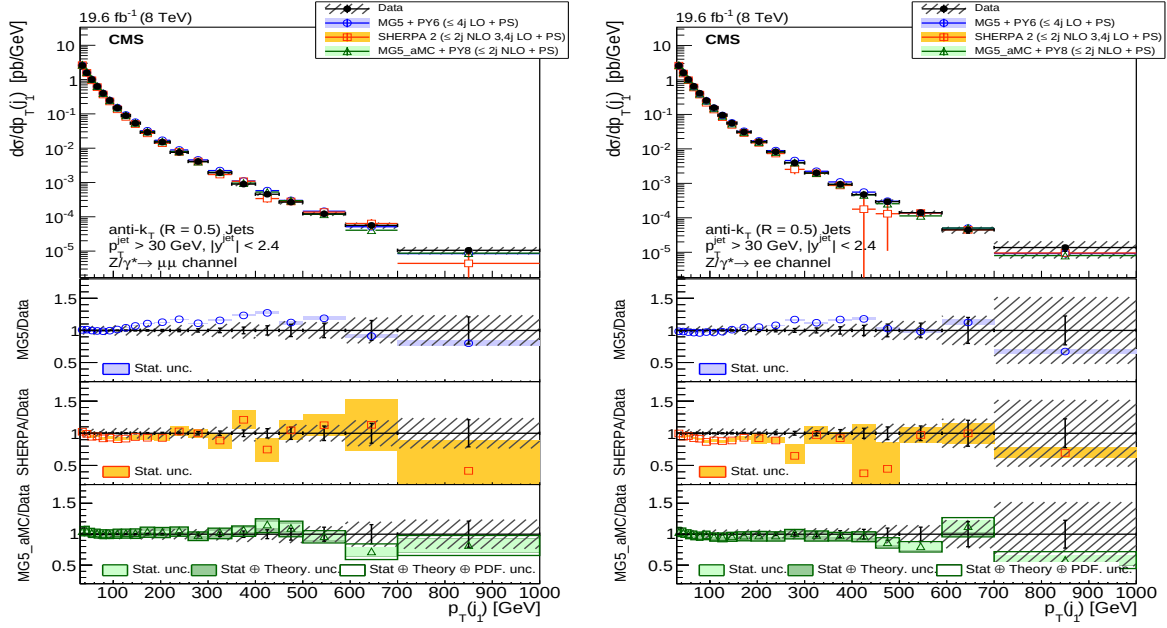


Figure 4.107: The differential cross section as a function of leading jet p_T in $\mu^+\mu^-$ (left) and e^+e^- (right) decay channels at 8 TeV, showing the full uncertainties. The measurement is compared with MADGRAPH (MG5), SHERPA2, and MADGRAPH_AMC@NLO (MG5_aMC).

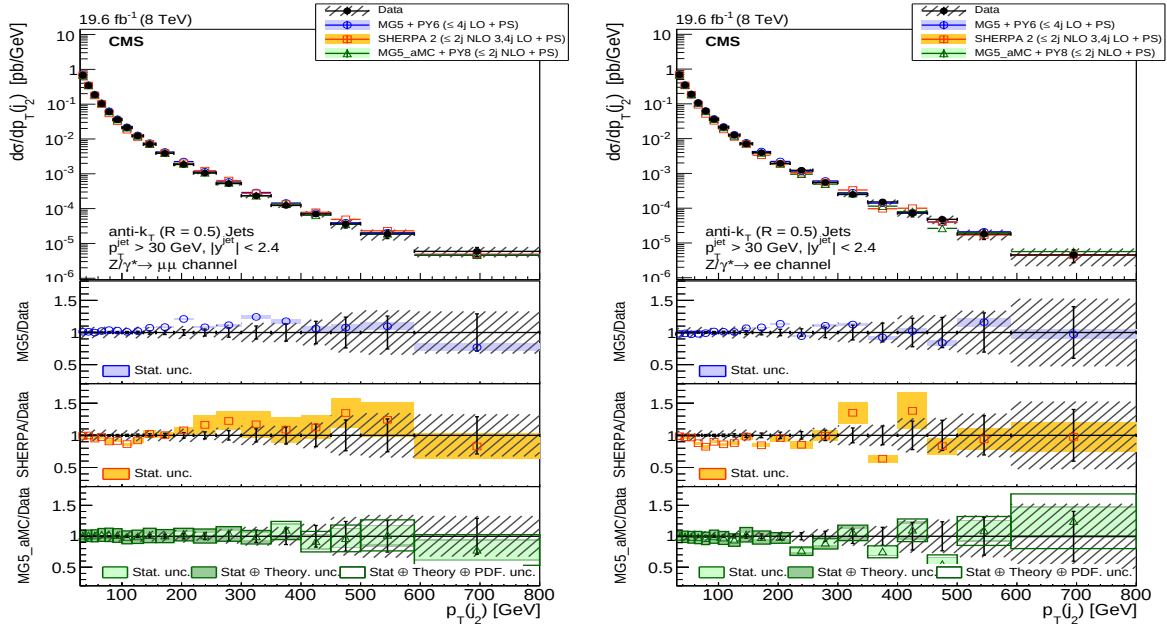


Figure 4.108: The differential cross section as a function of the second jet p_T in $\mu^+\mu^-$ (left) and e^+e^- (right) decay channels at 8 TeV, showing the full uncertainties. The measurement is compared with MADGRAPH (MG5), SHERPA2, and MADGRAPH_AMC@NLO (MG5_aMC).

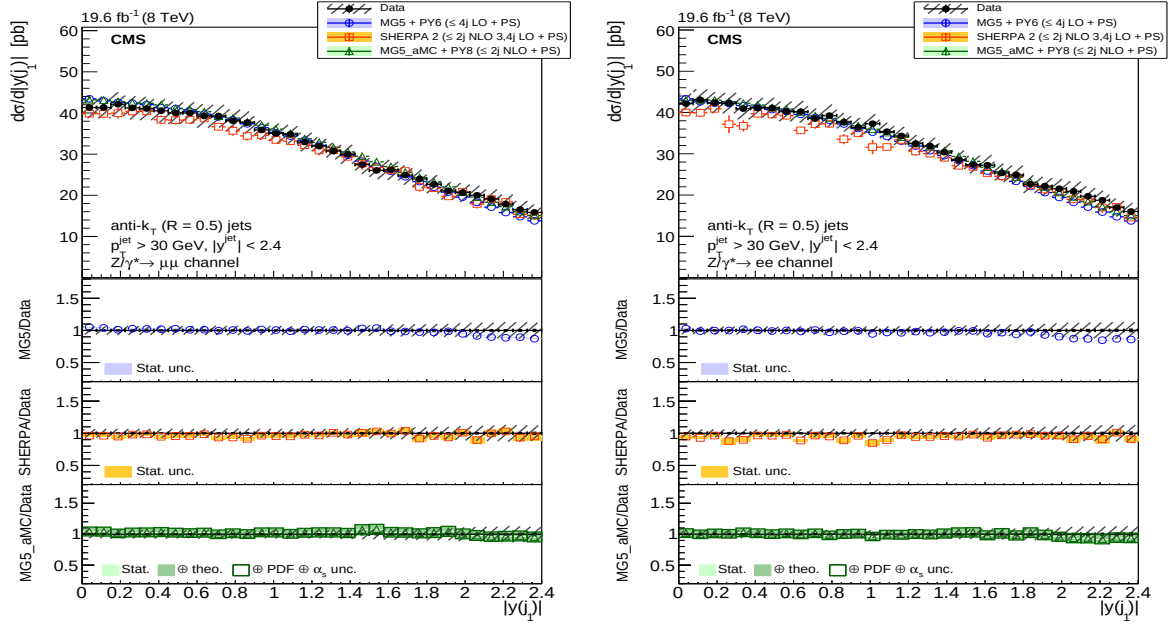


Figure 4.109: The differential cross section as a function of leading jet absolute rapidity in $\mu^+\mu^-$ (left) and e^+e^- (right) decay channels at 8 TeV, showing the full uncertainties. The measurement is compared with MADGRAPH (MG5), SHERPA2, and MADGRAPH_AMC@NLO (MG5_aMC).

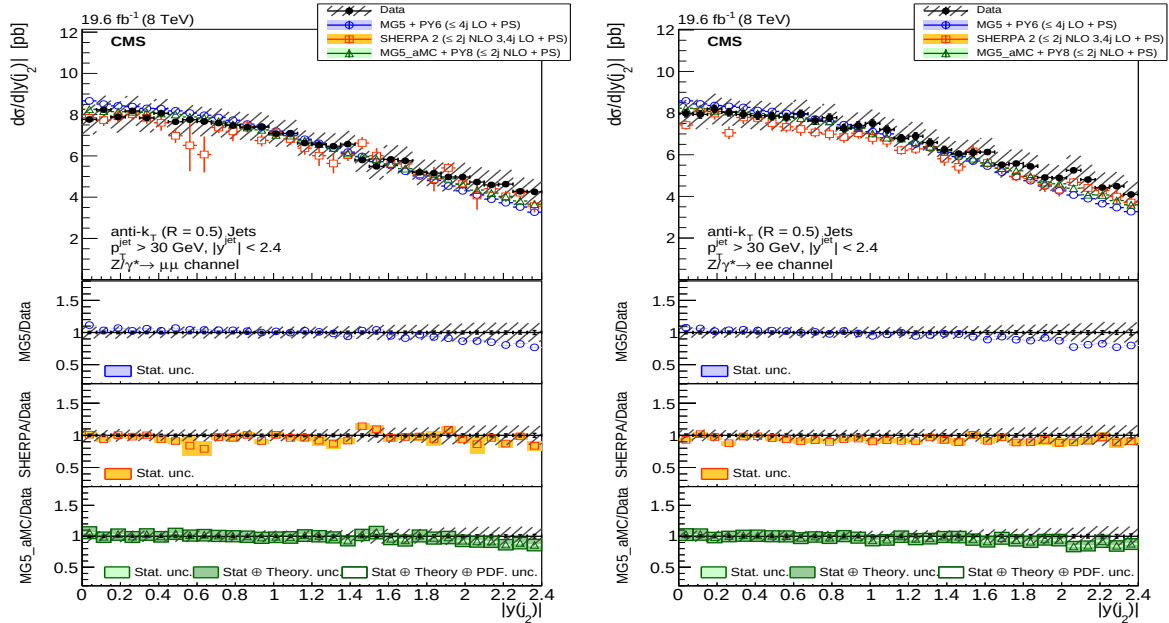


Figure 4.110: The differential cross section as a function of the second jet absolute rapidity in $\mu^+\mu^-$ (left) and e^+e^- (right) decay channels at 8 TeV, showing the full uncertainties. The measurement is compared with MADGRAPH (MG5), SHERPA2, and MADGRAPH_AMC@NLO (MG5_aMC).

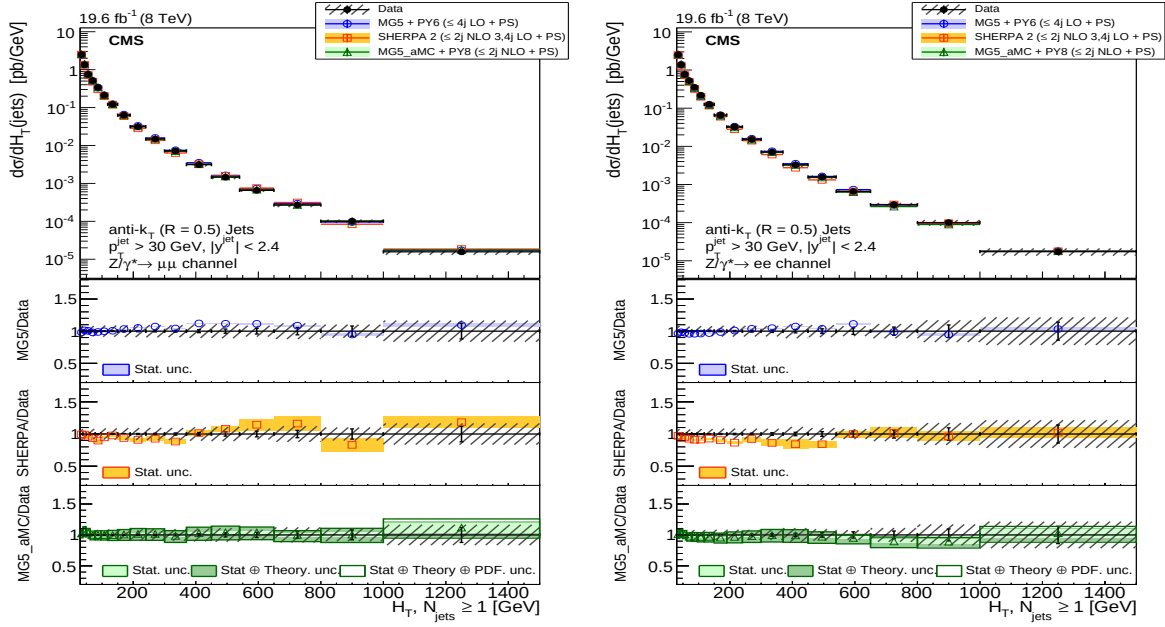


Figure 4.111: The differential cross section as a function of scale sum of jets p_T by requiring at least one jet in $\mu^+\mu^-$ (left) and e^+e^- (right) decay channels at 8 TeV, showing the full uncertainties. The measurement is compared with MADGRAPH (MG5), SHERPA2, and MADGRAPH_AMC@NLO (MG5_aMC).

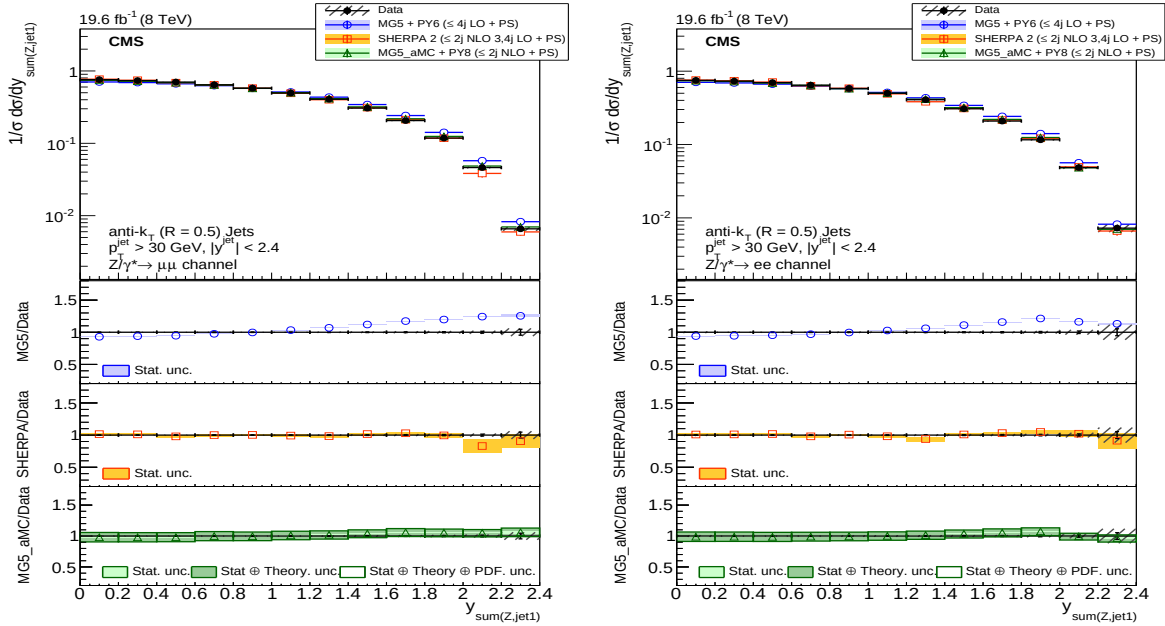


Figure 4.112: The normalized differential cross section as a function of rapidity sum of Z boson and leading jet in $\mu^+\mu^-$ (left) and e^+e^- (right) decay channels at 8 TeV, showing the full uncertainties. The measurement is compared with MADGRAPH (MG5), SHERPA2, and MADGRAPH_AMC@NLO (MG5_aMC).

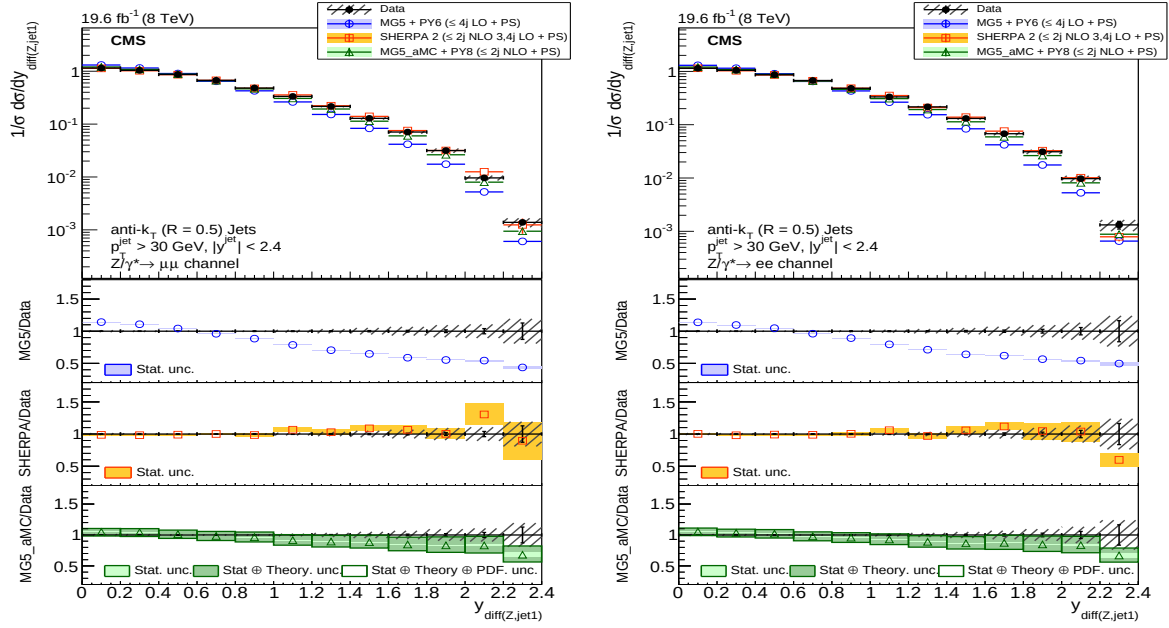


Figure 4.113: The normalized differential cross section as a function of rapidity difference of Z boson and leading jet in $\mu^+\mu^-$ (left) and e^+e^- (right) decay channels at 8 TeV, showing the full uncertainties. The measurement is compared with MADGRAPH (MG5), SHERPA2, and MADGRAPH_AMC@NLO (MG5_aMC).

4.9.3 Results of single channel at 13 TeV

As mentioned at the beginning of this chapter, only $\mu^+\mu^-$ decay channel is presented in details at 13 TeV, and the preliminary results of e^+e^- decay channel are shown in appendix B. The measured differential cross sections are as functions of observables for a jet multiplicity less than four. The full systematic uncertainties considered at 13 TeV include the ones from jet energy scale and resolution, luminosity, background modelling, and unfolding. The theoretical predictions for comparison are:

- **MADGRAPH_AMC@NLO:** The matrix element computed in NLO accuracy with 0, 1, 2 partons in the final state, and with 3 partons in tree level, using NNPDF set [68]. It is interfaced with PYTHIA8 using CUETP8M1 tune for parton shower and hadronization. The FxFx [40] merging scheme is used with a merging scale parameter set to 30 GeV, and the strong coupling constant is set by default to 0.118 at the Z boson mass scale. For the showering and underlying events, the value of CUETP8M1 tune $\alpha_s = 0.130$ is used. The statistical uncertainty is taken into account in the analysis.
- **MADGRAPH:** The matrix element computed at tree level (LO) up to four partons in the final state, using the CTEQ6L1 [109] PDF and interfaced with PYTHIA8 Tune CUETP8M1 for parton shower and hadronization. The default strong coupling constant for CUETP8M1 tune is 0.130 at the Z boson mass scale. The matrix element calculation is matched with the parton showering using the k_T -MLM [39, 112] scheme with the merging scale set at 19 GeV. The inclusive cross section is normalized to the NNLO cross section predicted by FEWZ (see section 4.2), which is 2008^{+88}_{-83} pb. The statistical uncertainty of this generator prediction is shown in the comparison.
- **NNLO Z + 1 jet:** This prediction is referred to [104, 107]. It is computed at parton level, at NNLO level with one parton in the final state, using N-jettiness subtraction scheme. In order to make the comparison with measurement, non-perturbative corrections are applied on it (see section 4.8). Both theoretical (scale) and statistical uncertainties are considered. The latter is negligible (within 1%) compared with the former.

The uncertainties corresponding to the renormalization and factorization scales, and PDFs are considered for the MADGRAPH_AMC@NLO prediction. These uncertainties are added in quadrature to be propagated to the differential cross sections.

The results in $\mu^+\mu^-$ decay channel at 13 TeV are illustrated from figure 4.114 to figure 4.118. In general, the MADGRAPH_AMC@NLO and NNLO predictions describe the measurement very well within the systematic uncertainty and theoretical uncertainty. The prediction at LO accuracy agrees in particular with data for the jet multiplicity and the third jet rapidity distributions.

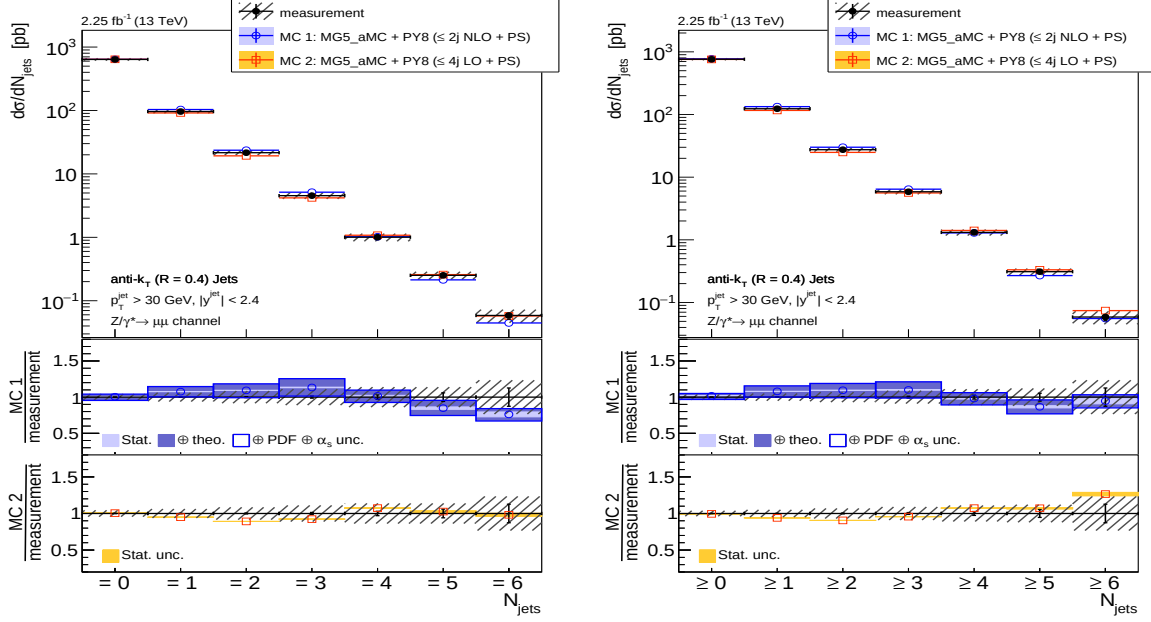


Figure 4.114: The differential cross section as a function of exclusive (left) and inclusive (right) jet multiplicity in $\mu^+\mu^-$ decay channel at 13 TeV, showing the full uncertainties. The measurement is compared with MADGRAPH_AMC@NLO (NLO) and MADGRAPH (LO) predictions.

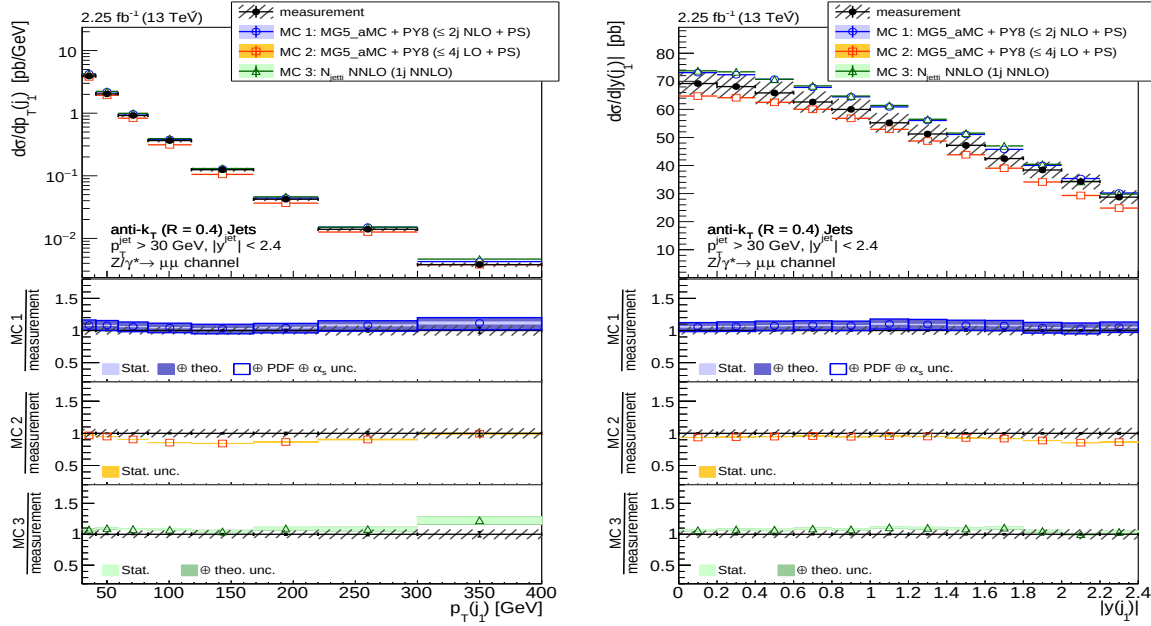


Figure 4.115: The differential cross section as a function of leading jet p_T (left) and absolute rapidity (right) in $\mu^+\mu^-$ decay channel at 13 TeV, showing the full uncertainties. The measurement is compared with MADGRAPH_AMC@NLO (NLO), MADGRAPH (LO), and NNLO predictions.

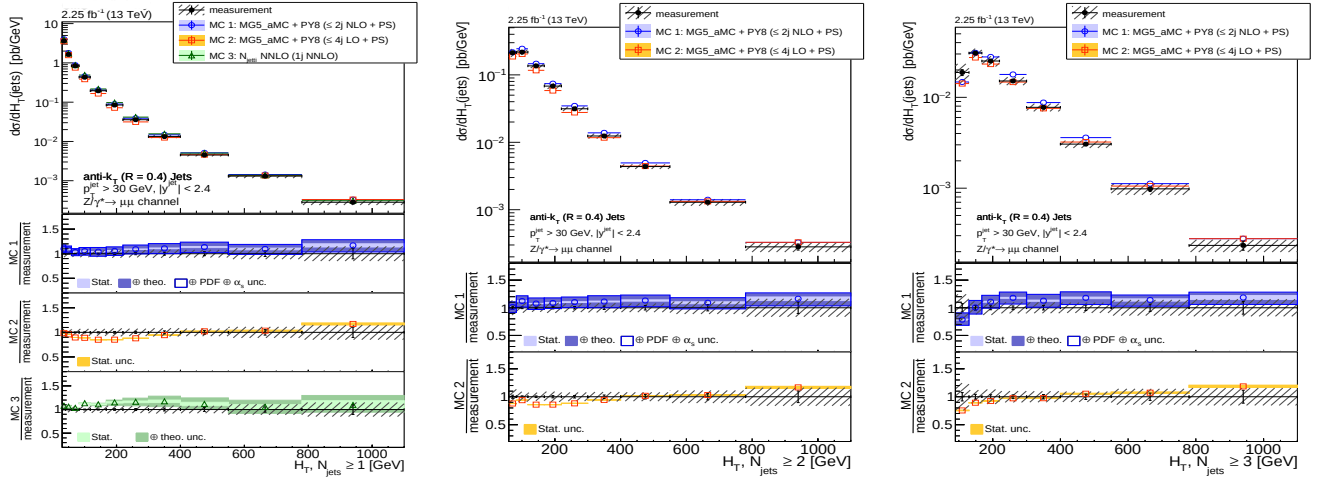


Figure 4.118: The differential cross section as a function of scalar sum of jets p_T for $N_{\text{jets}} \geq 1$ (left), $N_{\text{jets}} \geq 2$ (middle), and $N_{\text{jets}} \geq 3$ (right) in $\mu^+\mu^-$ decay channel at 13 TeV, showing the full uncertainties. The measurement is compared with MADGRAPH_AMC@NLO (NLO) and MADGRAPH (LO) predictions. The left plot also shows the NNLO prediction.

4.10 Combined differential cross sections

Once we obtain the differential cross sections with the corresponding uncertainties in single leptonic flavor channels, we can combine them for estimating the differential cross sections in l^+l^- decay channel. This is called combined measurement. Because the probabilities of a Z boson decaying to $\mu^+\mu^-$ and e^+e^- are same, and the coupling between a Z boson and the two leptons is leptonic flavor independent, the observations of the differential cross sections in the two decay channels should be the same. In order to gain more precision, we combine the measurement in both decay channels.

However, the uncertainties in this two channels are not independent. For instance, the associated jets are reconstructed in the same way, so the jet energy scale and resolution of the two channels are correlated. The same story happens to the luminosity and pileup. With respect to the simulation, the events of two channels are generated using the same MC generators for the observable distributions (not only signal sample but also background samples), and the same unfolding method is applied for them. The uncertainties related to each source above between the two channels are correlated. At this point, a dedicated approach is developed for the combination.

4.10.1 Combination methodology

The differential cross sections are combined using a weighted mean approach of their uncertainties. The general formula is written as:

$$\sigma_\alpha^{comb.} = \sum_{i=1}^{2n} \lambda_{\alpha i} \sigma_i = \sum_{i=1}^n \lambda_{\alpha i} \sigma_i^{ee} + \sum_{i=n+1}^{2n} \lambda_{\alpha i} \sigma_{i-n}^{\mu\mu} \quad (4.13)$$

where $\sigma_\alpha^{comb.}$ represents the measured differential cross section related to an observable in the α th bin with two channels combined. α stands for the bin number and ranges from 1 to n . The superscripts ee and $\mu\mu$ correspond to the decay channels. The series of σ_i are vectors with respect to the measured differential cross section in each bin. The $\lambda_{\alpha i}$ are $n \times 2n$ matrices containing the coefficients (weights) for both channels. For the case of our analysis, this formula is simplified as:

$$\sigma_\alpha^{comb.} = \lambda_{\alpha\alpha} \sigma_\alpha^{ee} + \lambda_{\alpha\alpha+n} \sigma_\alpha^{\mu\mu} = \frac{(\Delta\sigma_\alpha^{ee})^{-2}}{(\Delta\sigma_\alpha^{ee})^{-2} + (\Delta\sigma_\alpha^{\mu\mu})^{-2}} \sigma_\alpha^{ee} + \frac{(\Delta\sigma_\alpha^{\mu\mu})^{-2}}{(\Delta\sigma_\alpha^{ee})^{-2} + (\Delta\sigma_\alpha^{\mu\mu})^{-2}} \sigma_\alpha^{\mu\mu} \quad (4.14)$$

where $\Delta\sigma_\alpha^{ee}$ and $\Delta\sigma_\alpha^{\mu\mu}$ are the total uncertainties in the α th bin related to the measurements for the two channels, respectively. Note that the correlation between adjacent bins for a single decay channel is already concerned in the unfolding approach (see section 4.6.2). As is explained later, the correlations are taken into account in the uncertainty estimation through covariance matrices propagation.

In order to estimate the uncertainties correctly after combination, one needs to compute the propagated covariance matrices using linear estimates [113]. By means of the coefficients $\lambda_{\alpha i}$ in

formula 4.13, the covariance matrix of the combined differential cross section is given by:

$$\Delta\sigma_{\alpha\beta}^{comb.} = \sum_{i=1}^{2n} \sum_{j=1}^{2n} \lambda_{\alpha i} M_{ij} \lambda_{\beta j}, \quad (4.15)$$

where M_{ij} is a $2n \times 2n$ covariance matrix of the measurement in both leptonic decay channels. Concerning all of the systematic and statistical uncertainties in this analysis, the matrix is defined as below [13]:

$$\begin{aligned} M &= \sum_{s \in \text{unc.sources}} M^s \\ &= \sum_{s \in \{JES, JER, PU, Bgnd, Lumi\}} \left(\begin{array}{c|c} (\Delta\sigma^{ee})^s & \pm \sqrt{(\Delta\sigma_{\alpha}^{ee})^s \times (\Delta\sigma_{\beta}^{\mu\mu})^s} \\ \hline \pm \sqrt{(\Delta\sigma_{\alpha}^{ee})^s \times (\Delta\sigma_{\beta}^{\mu\mu})^s} & (\Delta\sigma^{\mu\mu})^s \end{array} \right) \\ &+ \sum_{s \in \{Stat, Eff, LES, LER, Unf\}} \left(\begin{array}{c|c} (\Delta\sigma^{ee})^s & 0 \\ \hline 0 & (\Delta\sigma^{\mu\mu})^s \end{array} \right) \end{aligned} \quad (4.16)$$

where the $(\Delta\sigma^{ll})^s$ are the covariance matrices of single channel for the uncertainty source s , which is described in section 4.7. Note that the "Stat" represents the sources corresponding to the statistical uncertainties of both data and MC. The sign of the off diagonal block is the same as the sign of the diagonal element of $(\Delta\sigma^{ll})^s$. The uncertainties resulting from JES, JER, PU, Bgnd, and Lumi of the two decay channels are correlated, so the corresponding covariance matrices are crossed in the off diagonal blocks. On the other hand, the statistical uncertainties and the lepton efficiency/energy/resolution uncertainties of the two channels are independent, therefore their covariance matrices only appear in the diagonal blocks. The systematic uncertainties coming from unfolding modelling of the two channels are not fully correlated, so we adopt a conservative estimation assuming that they are independent. In this way we obtain the combined covariance matrices related to each kind of uncertainty, as well as the total uncertainty.

4.10.2 Combined results at 8 TeV

The 8 TeV combined differential cross sections are presented in this section and compared to theoretical predictions.

- Jet multiplicity

The differential cross section as a function of exclusive and inclusive jet multiplicities is measured concerning from zero up to seven jets, benefiting from the large statistics of data at 8 TeV. The comparison of measurement and predictions are shown in figure 4.119. One can see excellent agreement with MADGRAPH tree level calculation and SHERPA2 prediction, both

take into account a maximum of four partons in the final state for the matrix element computation. In terms of MADGRAPH_AMC@NLO, which includes maximum three final state partons into the matrix element, it is consistent with data up to three jets. The discrepancy at larger jet multiplicity regions originating from the parton shower are within one standard deviation. The values of the differential cross sections together with the systematic and statistical uncertainties are summarized in table 4.14.

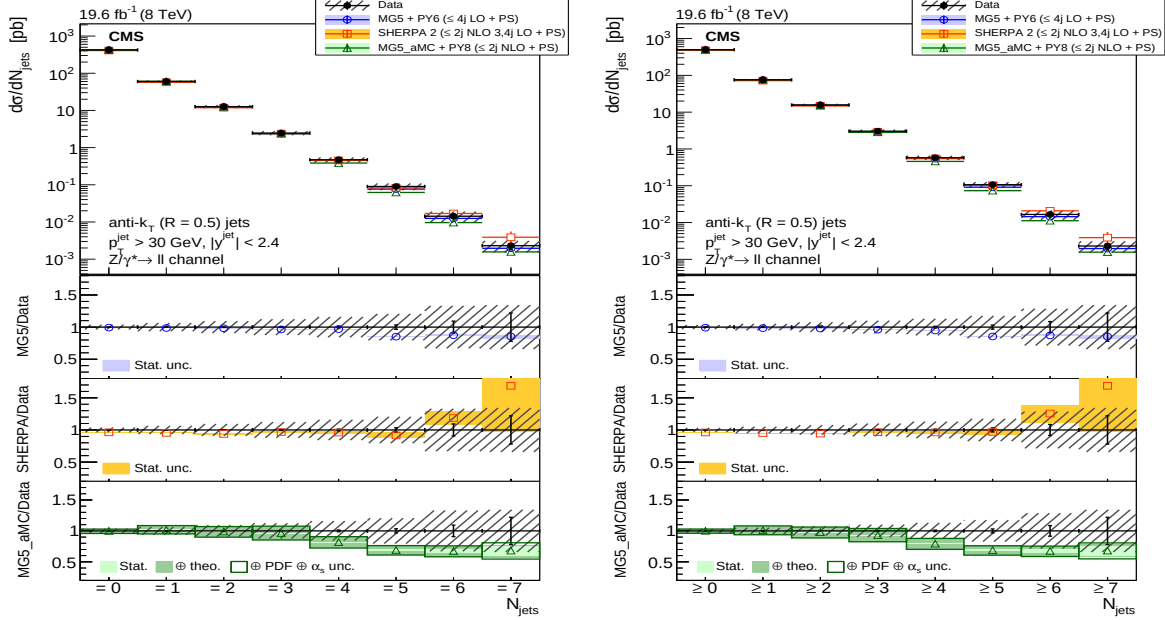


Figure 4.119: The measured differential cross section as a function of exclusive (left) and inclusive (right) jet multiplicity in l^+l^- decay channel at 8 TeV, concerning the full uncertainties. The measurement is compared to MADGRAPH (MG5), SHERPA2 and MADGRAPH_AMC@NLO (MG5_aMC) predictions.

- Z boson and jet transverse momenta

The differential cross section as a function of Z boson, jet p_T for the first, second, third, fourth, and fifth jet are displayed in figures 4.120, 4.121, and 4.122. The cross sections fall rapidly with increasing p_T . Among them, the cross section for the leading jet is measured for p_T values between 30 GeV and 1 TeV and decreases by more than five orders of magnitude over this range. The cross section for the fifth jet is measured for p_T values between 30 and 100 GeV and decreases even faster, mainly because of the phase space covered.

Regarding the Z boson and leading jet transverse momenta, the agreement between MADGRAPH LO prediction and measurement keeps good until about 150 GeV. There are discrepancies observed approximately from 150 GeV to 450 GeV. A similar excess has been seen in the CMS experiment at 7 TeV [114] with the same MC generators. It appears as well at the ATLAS experiment [115], which used ALPGEN [116] interfaced to HERWIG [117] for the predictions

N_{jets}	$\frac{d\sigma}{dN_{jets}}$ [pb]	Tot unc [%]	Stat [%]	JES [%]	JER [%]	Bgnd [%]	PU [%]	Unf stat [%]	Unf sys [%]	Lumi [%]	Eff [%]
=0	423	3.7	0.034	1.2	0.06	0.002	0.71	0.05	1.2	2.6	1.8
=1	59.9	6.3	0.11	5.3	0.23	0.042	0.26	0.075	1.4	2.6	1.8
=2	12.6	9.2	0.25	8.4	0.22	0.33	0.35	0.12	1.7	2.7	1.9
=3	2.46	12	0.6	11	0.22	0.76	0.42	0.22	2.7	2.9	2.0
=4	0.471	16	1.4	15	0.16	1.3	0.57	0.43	3.5	3.1	2.1
=5	0.0901	20	3.4	19	0.28	1.9	0.75	1.0	4.6	3.2	2.3
=6	0.0143	33	9.3	28	0.72	3.3	1.9	2.4	5.5	3.7	2.6
=7	0.00230	34	22	23	0.61	4.3	5.6	6.3	6.4	3.9	2.8

Table 4.14: Cross section results obtained from the combination of the muon and electron channels as a function of the exclusive jet multiplicity and details of the systematic uncertainties. The column denoted “Tot unc” contains the total uncertainty; the column denoted “Stat” contains the statistical uncertainty; the remaining columns contain the systematic uncertainties.

(see section 4.11). This discrepancy disappears with NLO prediction. The SHERPA2 prediction shows about 10% discrepancy at low p_T region, and it is in statistic shortage at higher p_T region. The second jet p_T distribution has the similar performance. For the third jet p_T distribution, both MADGRAPH and MADGRAPH_AMC@NLO describe the data very well, while the shape of SHERPA2 prediction is a bit away. The three predictions have the similar performance in the distributions of forth and fifth jet p_T , while some discrepancies are observed for SHERPA2 and MADGRAPH_AMC@NLO predictions.

- Z boson and jet absolute rapidities

The differential cross section as a function of absolute rapidity for Z boson and first to fifth jet, respectively, are shown from figure 4.123 to 4.125. Overall, MADGRAPH and MADGRAPH_AMC@NLO have the similar shape in all the distributions. SHERPA2 has generally good agreement with data, while it suffers from statistical limitation with fine binning.

The rapidity correlation distributions of Z boson, first and second jets, are illustrated from figure 4.126 to 4.130. As has been seen in the detector reconstruction level (see section 4.4.1), there are prominent discrepancies in the shape distributions of $y_{sum}(Z, j_1)$ (figure 4.130) and $y_{diff}(Z, j_1)$ (figure 4.126), with respect to the MADGRAPH LO predictions. This phenomenon first also observed at 7 TeV [61] (also see section 4.11), and it is now confirmed again with extended rapidity range and benefiting from increased statistical precision at 8 TeV. From the same figures, we can see that the discrepancies are reduced significantly when applying a specific threshold on Z boson p_T . The observation above supports the attributions of the discrepancy to the matching procedure between the matrix element and parton shower, which has been addressed in [61]. On the other hand, the other two NLO predictions have no hard disagreement in the same distributions, thanks to the new matching schemes between matrix

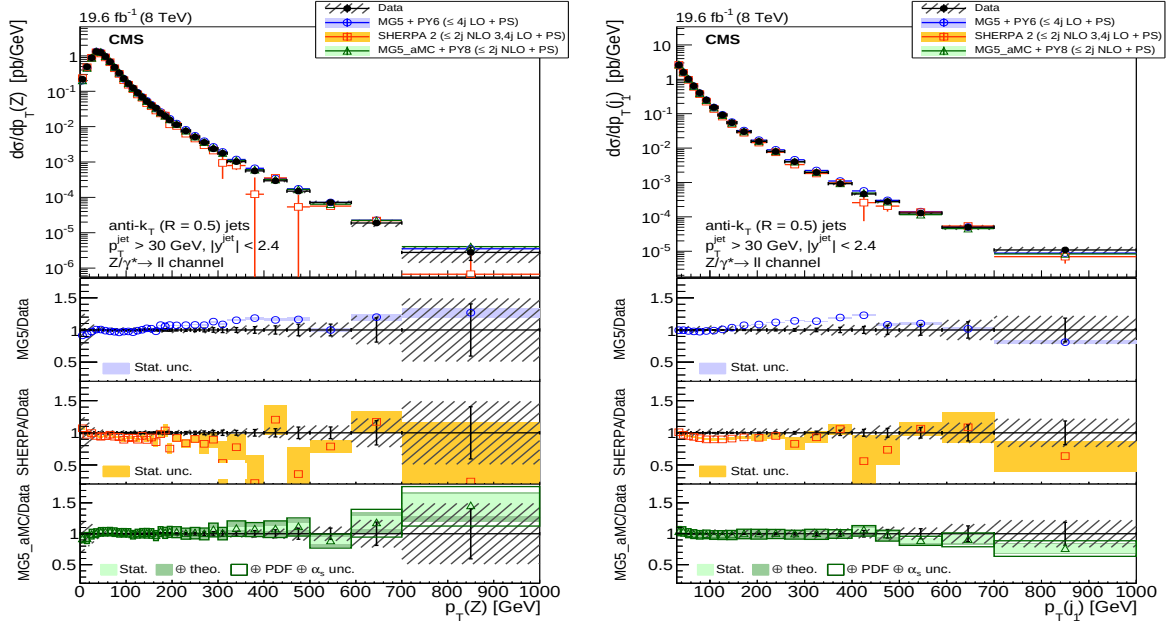


Figure 4.120: The measured differential cross section as a function of Z boson p_T (left) and leading jet p_T (right) in l^+l^- decay channel at 8 TeV, concerning the full uncertainties. The measurement is compared to MADGRAPH (MG5), SHERPA2 and MADGRAPH_AMC@NLO (MG5_aMC) predictions.

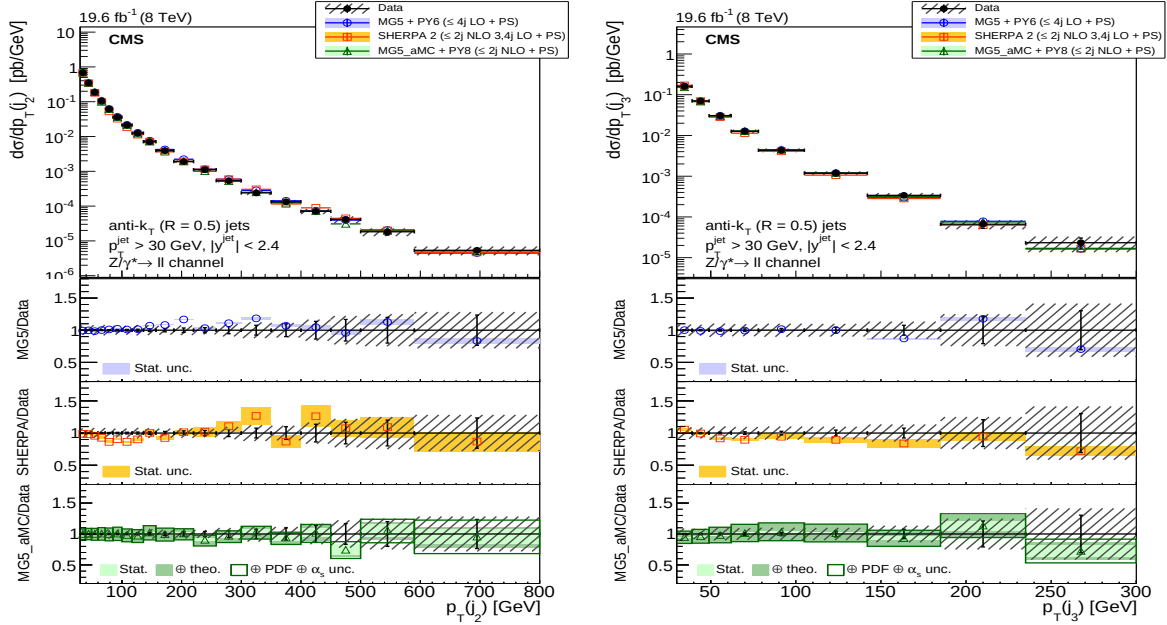


Figure 4.121: The measured differential cross section as a function of second jet p_T (left) and third jet p_T (right) in l^+l^- decay channel at 8 TeV, concerning the full uncertainties. The measurement is compared to MADGRAPH (MG5), SHERPA2 and MADGRAPH_AMC@NLO (MG5_aMC) predictions.

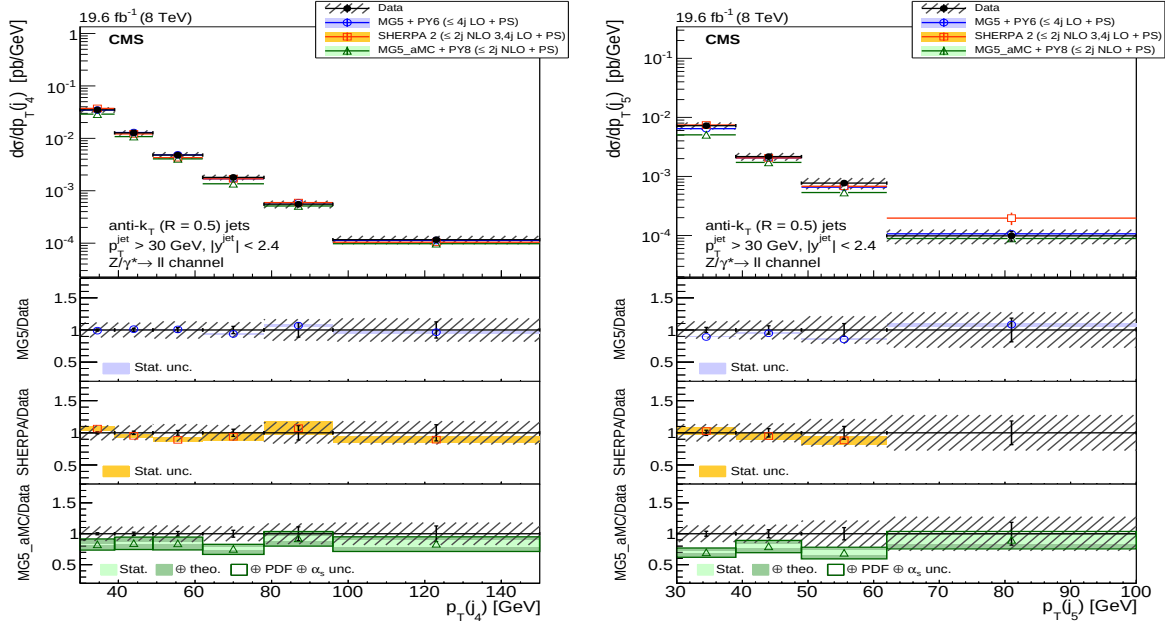


Figure 4.122: The measured differential cross section as a function of forth jet p_T (left) and fifth jet p_T (right) in l^+l^- decay channel at 8 TeV, concerning the full uncertainties. The measurement is compared to MADGRAPH (MG5), SHERPA2 and MADGRAPH_AMC@NLO (MG5_aMC) predictions.

elements and parton shower. Note that the SHERPA2 is short of statistics for the distributions with Z boson p_T threshold applied.

The presence of additional jets in the event should reduce the dependence on the matching between matrix element and parton shower for the first jet, because this jet will have larger p_T on the average. According to figure 4.127 and 4.129, we can see that the shape discrepancies of MADGRAPH in the distribution of $y_{sum}(Z, j_1)$ and $y_{diff}(Z, j_1)$ are decreased, by considering at least two jets. Moreover, the MADGRAPH prediction also fails to describe the shape distribution of $y_{sum}(Z, j_2)$ and $y_{diff}(Z, j_2)$, with the same level of difference as $y_{sum}(Z, j_1)$ and $y_{diff}(Z, j_1)$, respectively. However, if considering the dijet events, the discrepancy of MADGRAPH becomes much larger. The rapidity correlation between the two leading jets regardless of the Z boson, as shown in figure 4.128, all of the predictions describe the $y_{sum}(j_1, j_2)$ well, while MADGRAPH prediction disagrees with the $y_{diff}(j_1, j_2)$ distribution. On the other hand, the NLO predictions agree with the measurement in general.

- Scalar sum of jet transverse momenta

The differential cross section as a function of the jets H_T , for inclusive one up to five jets production are displayed in figure 4.131 and 4.132. All of the predictions agree with the measurement for the cases of $N_{jets} \geq 1$ and $N_{jets} \geq 2$. Nevertheless, for the distributions with $N_{jets} \geq 3$ and $N_{jets} \geq 4$, all of the predictions are away from the measurement at the low

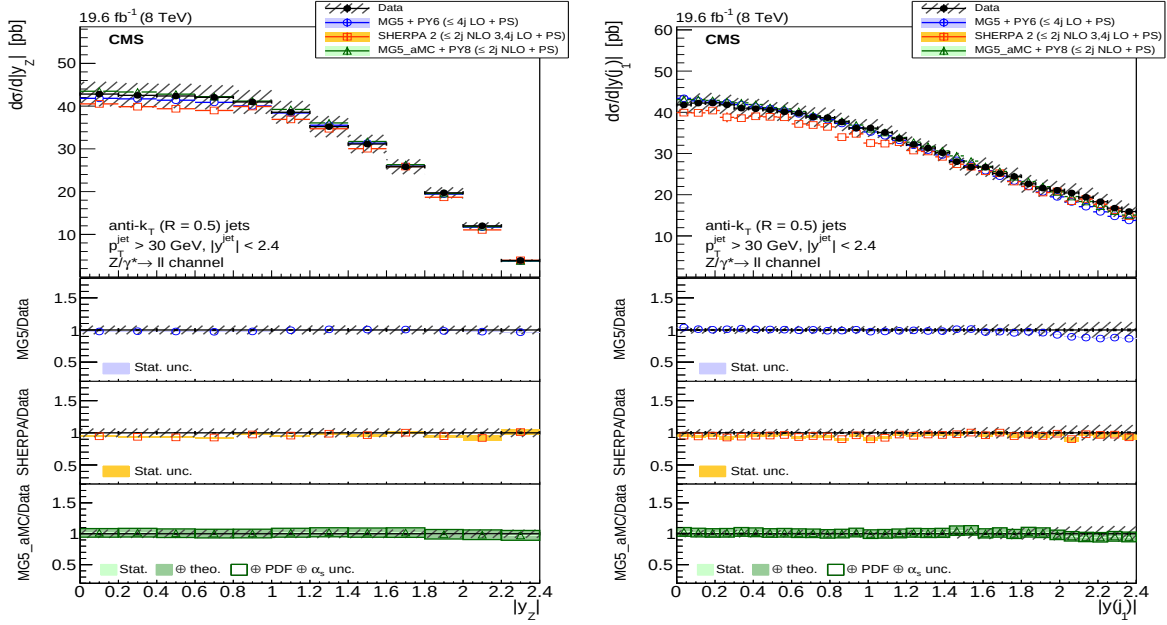


Figure 4.123: The measured differential cross section as a function of Z boson absolute rapidity (left) and leading jet absolute rapidity (right) in l^+l^- decay channel at 8 TeV, concerning the full uncertainties. The measurement is compared to MADGRAPH (MG5), SHERPA2 and MADGRAPH_AMC@NLO (MG5_aMC) predictions.

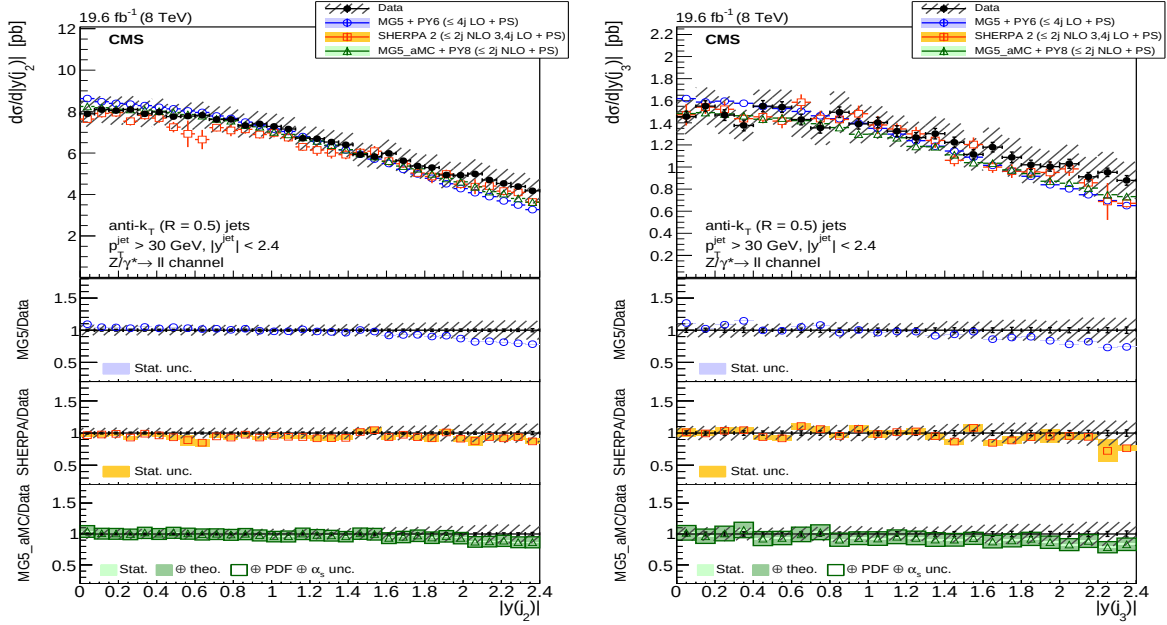


Figure 4.124: The measured differential cross section as a function of second jet absolute rapidity (left) and third jet absolute rapidity (right) in l^+l^- decay channel at 8 TeV, concerning the full uncertainties. The measurement is compared to MADGRAPH (MG5), SHERPA2 and MADGRAPH_AMC@NLO (MG5_aMC) predictions.

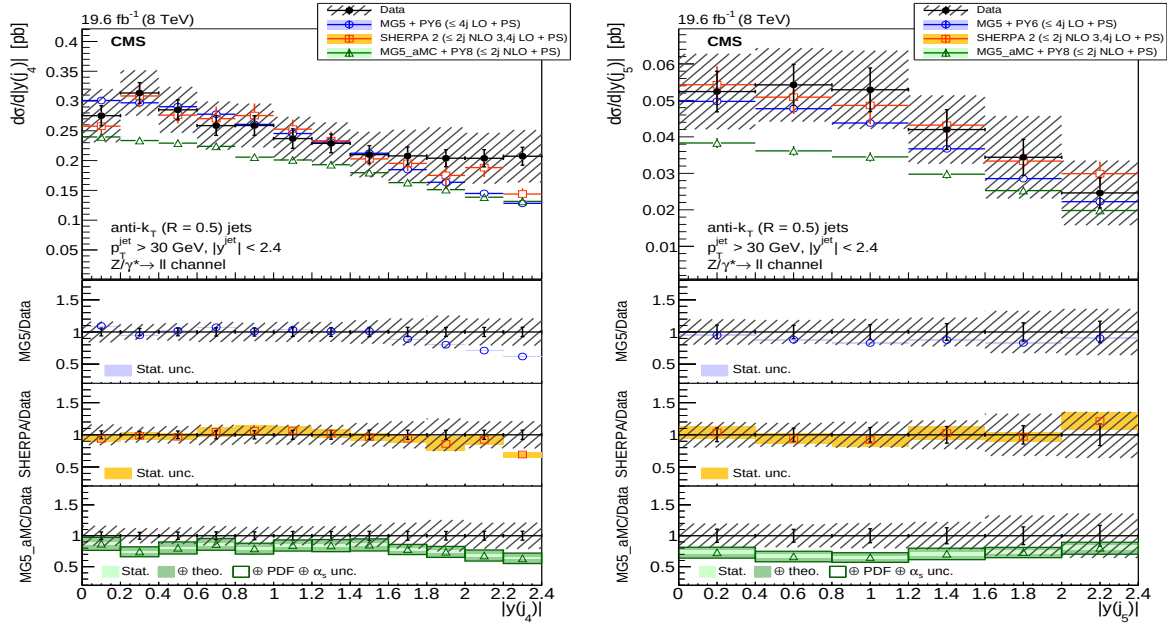


Figure 4.125: The measured differential cross section as a function of fourth jet absolute rapidity (left) and fifth jet absolute rapidity (right) in l^+l^- decay channel at 8 TeV, concerning the full uncertainties. The measurement is compared to MADGRAPH (MG5), SHERPA2 and MADGRAPH_AMC@NLO (MG5_aMC) predictions.

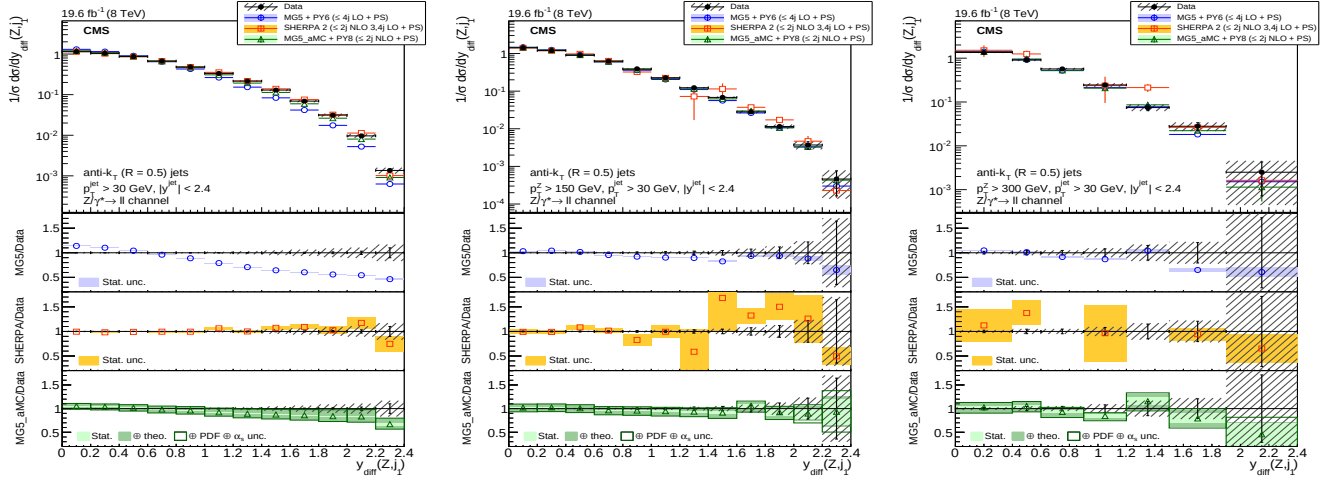


Figure 4.126: The normalized differential cross section as a function of rapidity difference between Z boson and leading jet by requiring at least one jet in l^+l^- decay channel at 8 TeV, with $p_T^Z > 0$ GeV (left), $p_T^Z > 150$ GeV (middle), and $p_T^Z > 300$ GeV (right), concerning the full uncertainties. The measurement is compared to MADGRAPH (MG5), SHERPA2 and MADGRAPH_AMC@NLO (MG5_aMC) predictions.

H_T region, where the p_T threshold (30 GeV) is applied for each jet. Regarding the distribution with $N_{jets} \geq 5$, SHERPA2 seems to have the best description.

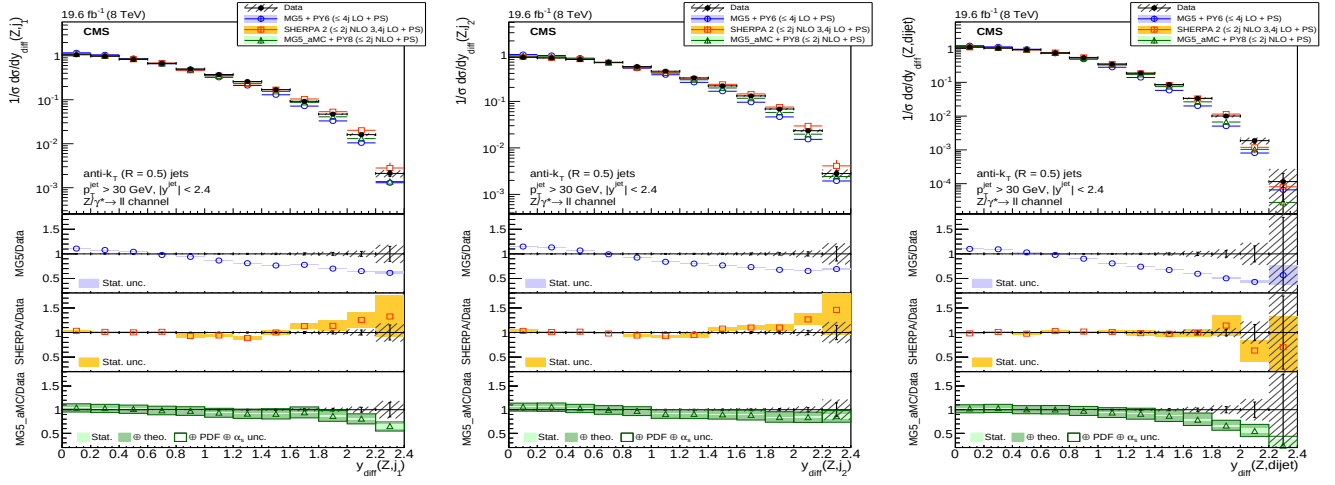


Figure 4.127: The normalized differential cross section as a function of rapidity difference between Z boson and leading jet (left), between Z boson and second jet (middle), between Z boson and di-jet (right) by requiring at least two jets in l^+l^- decay channel at 8 TeV, concerning the full uncertainties. The measurement is compared to MADGRAPH (MG5), SHERPA2 and MADGRAPH_AMC@NLO (MG5_aMC) predictions.

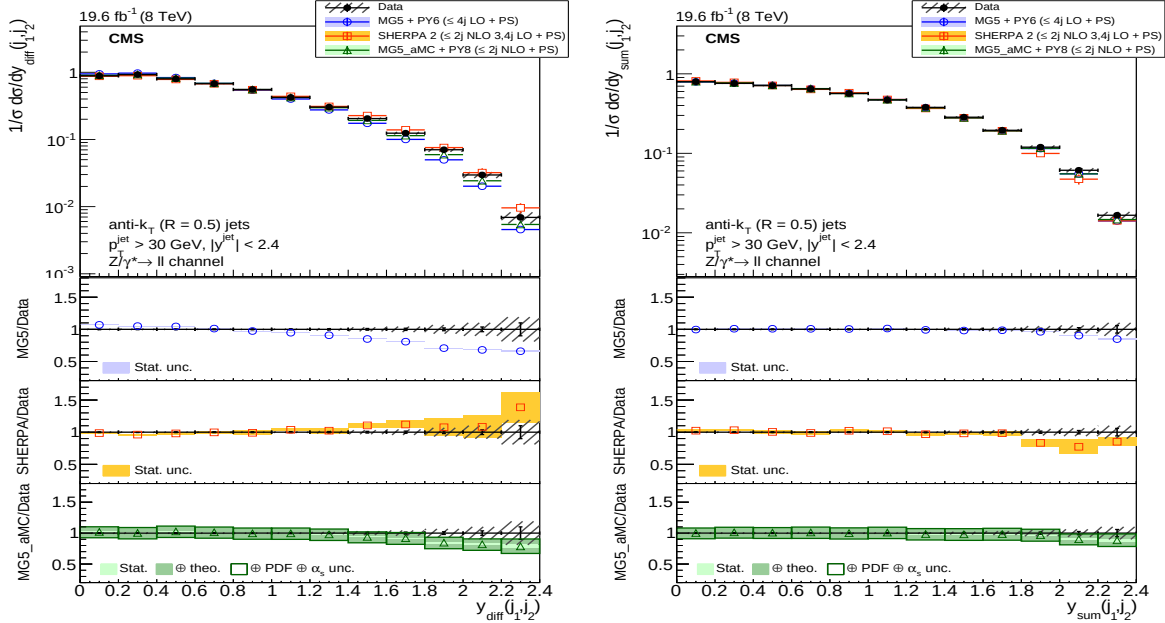


Figure 4.128: The normalized differential cross section as a function of rapidity difference (left) and rapidity sum (right), between the leading jet and second jet by requiring at least two jets in l^+l^- decay channel at 8 TeV, concerning the full uncertainties. The measurement is compared to MADGRAPH (MG5), SHERPA2 and MADGRAPH_AMC@NLO (MG5_aMC) predictions.

The measured differential cross section values with systematic and statistical uncertainties are summarized in the appendix A of this thesis.

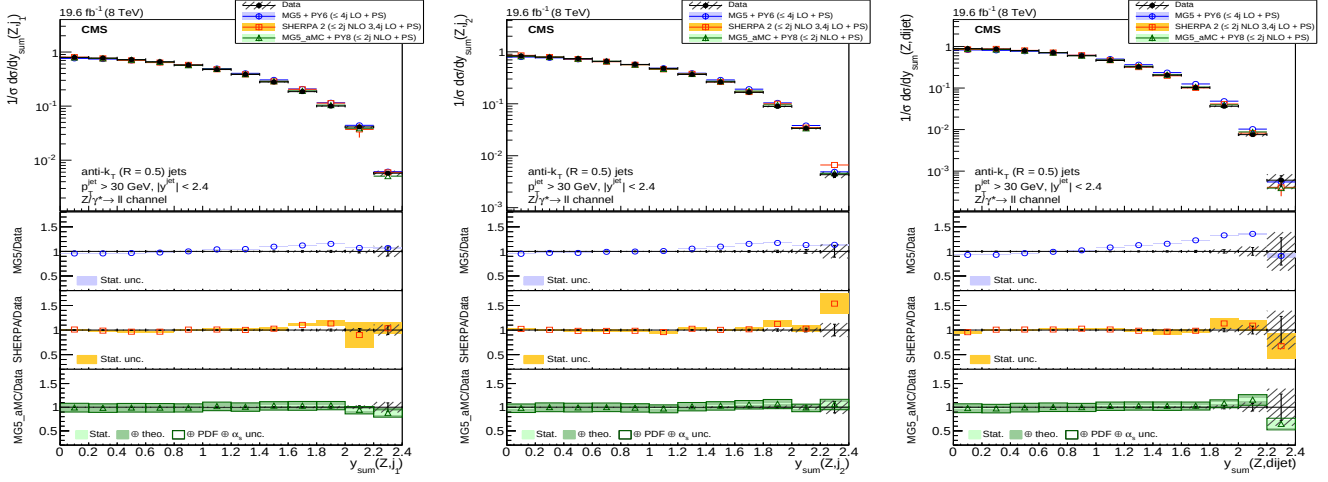


Figure 4.129: The normalized differential cross section as a function of rapidity sum between Z boson and leading jet (left), between Z boson and second jet (middle), between Z boson and di-jet (right) by requiring at least two jets in l^+l^- decay channel at 8 TeV, concerning the full uncertainties. The measurement is compared to MADGRAPH (MG5), SHERPA2 and MADGRAPH_AMC@NLO (MG5_aMC) predictions.

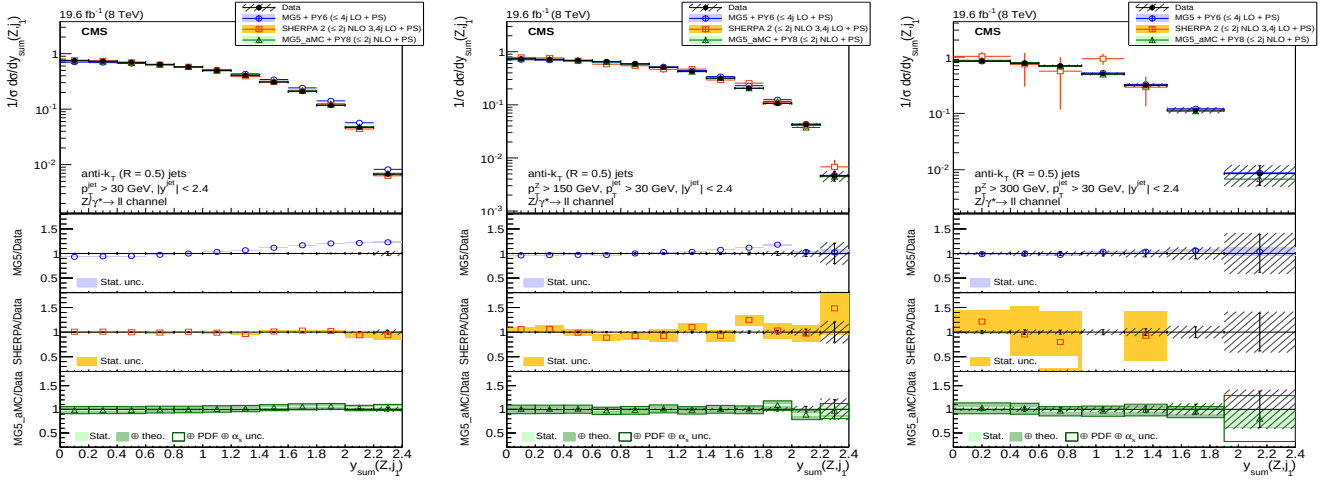


Figure 4.130: The normalized differential cross section as a function of rapidity sum between Z boson and leading jet by requiring at least one jet in l^+l^- decay channel at 8 TeV, with $p_T^Z > 0$ GeV (left), $p_T^Z > 150$ GeV (middle), and $p_T^Z > 300$ GeV (right), concerning the full uncertainties. The measurement is compared to MADGRAPH (MG5), SHERPA2 and MADGRAPH_AMC@NLO (MG5_aMC) predictions.

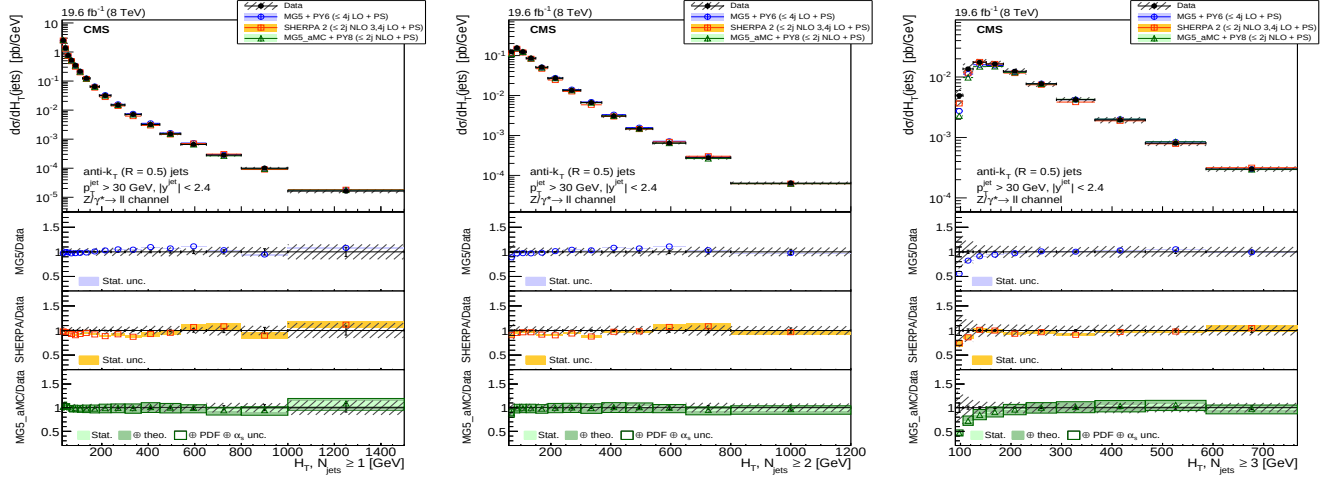


Figure 4.131: The measured differential cross section as a function of scalar sum of jets p_T by requiring at least one (left), two (middle) and three (right) jets in l^+l^- decay channel at 8 TeV, concerning the full uncertainties. The measurement is compared to MADGRAPH (MG5), SHERPA2 and MADGRAPH_AMC@NLO (MG5_aMC) predictions.

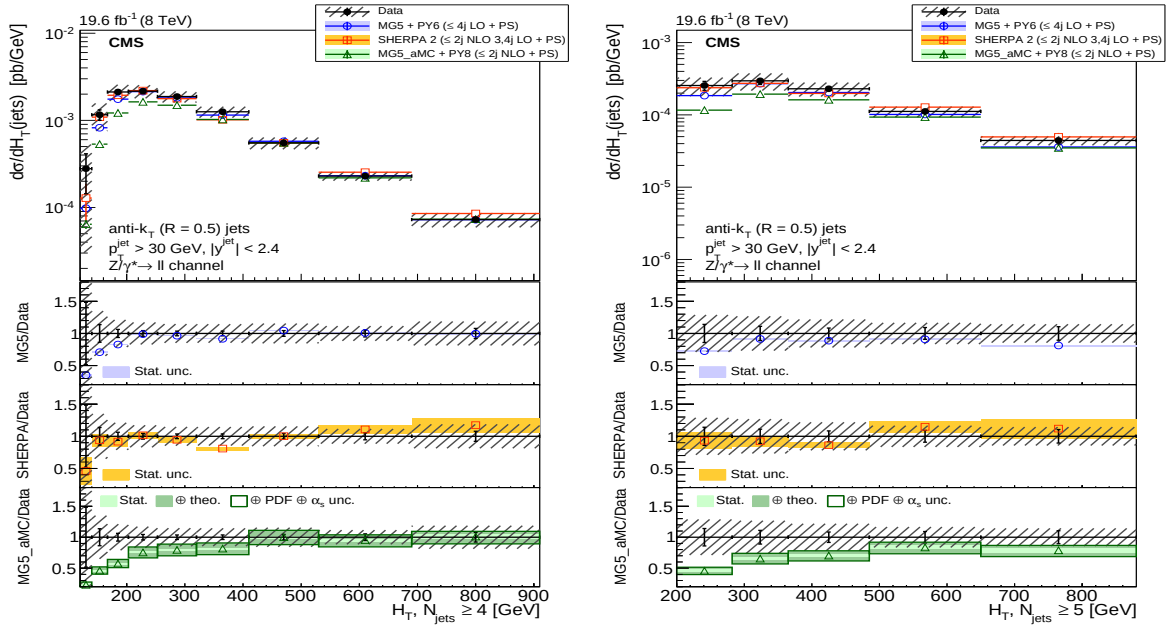


Figure 4.132: The measured differential cross section as a function of scalar sum of jets p_T by requiring at least four (left) and five (right) jets in l^+l^- decay channel at 8 TeV, concerning the full uncertainties. The measurement is compared to MADGRAPH (MG5), SHERPA2 and MADGRAPH_AMC@NLO (MG5_aMC) predictions.

4.10.3 Combined results at 13 TeV

The measured differential cross sections in two single leptonic decay channels at 13 TeV are combined in the same way as that of 8 TeV analysis. Together with the theoretical predictions, the combined differential cross sections with the full uncertainties ¹² at 13 TeV are shown as follows:

- Jet multiplicity

Figure 4.133 shows the differential cross section as a function of jet multiplicity in exclusive and inclusive cases. Within uncertainties, both MADGRAPH and MADGRAPH_AMC@NLO are compatible with measurement, except for the highest jet multiplicity region ($N_{jets} \geq 6$ for MADGRAPH). This is because of a statistical fluctuation.

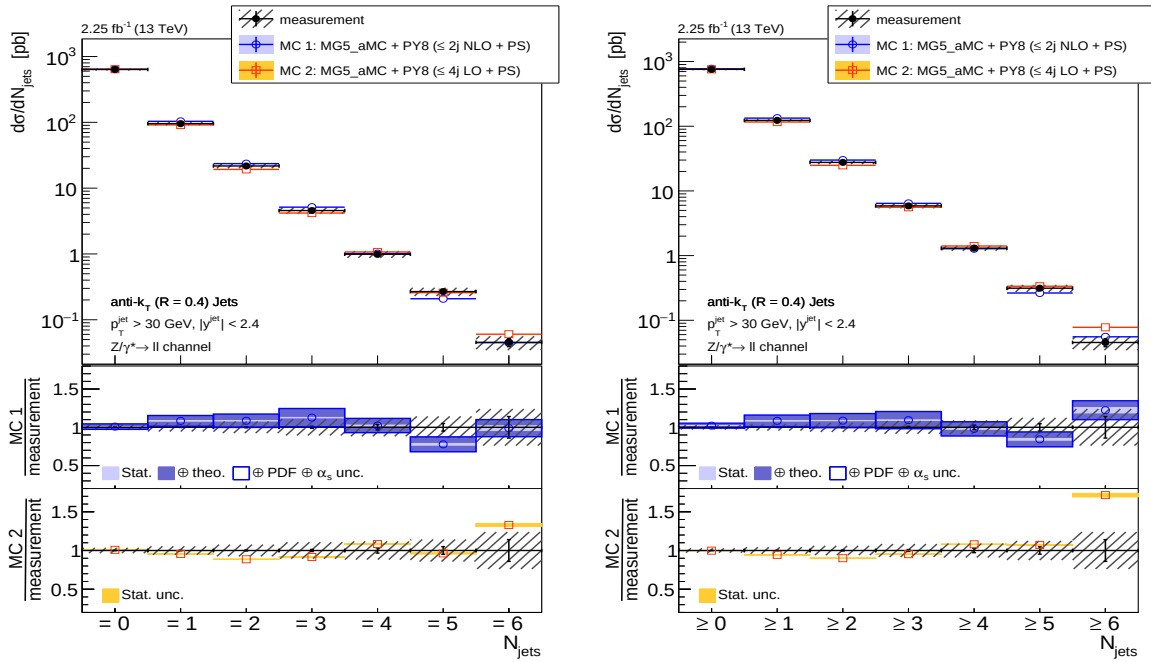


Figure 4.133: The measured differential cross section as a function of exclusive (left) and inclusive (right) jet multiplicity in l^+l^- decay channel at 13 TeV, concerning the full uncertainties. The measurement is compared to MADGRAPH, and MADGRAPH_AMC@NLO predictions.

- Jet transverse momentum

The differential cross sections as a function of p_T of the first, second, third jet, are shown in figure 4.134. The shape distributions of MADGRAPH and MADGRAPH_AMC@NLO are similar. Considering the full uncertainties, the MADGRAPH_AMC@NLO prediction agrees with data for all the distributions. The MADGRAPH prediction has a discrepancy up to 15% from 60 GeV to 300 GeV for the distribution of leading jet p_T , and gives another undershooting

¹²This is a reminder that the e^+e^- channel results are shown in the appendix. The details of e^+e^- channel are given in [59].

up to 10% ranging from 50 GeV to 120 GeV for the second jet p_T distribution. It overestimates the third jet p_T distribution up to 20% from 120 GeV to 250 GeV. The NNLO prediction have good consistency with the measurement for the leading jet p_T distribution.

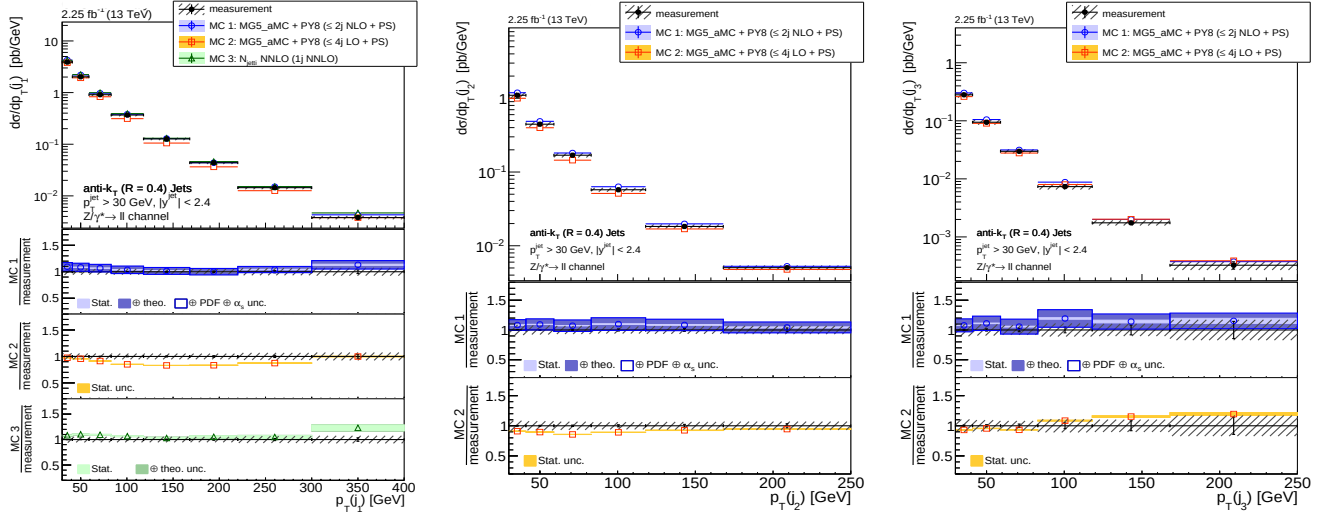


Figure 4.134: The measured differential cross section as a function of the leading jet p_T (left), second jet p_T (middle), and third jet p_T (right) in l^+l^- decay channel at 13 TeV, considering the full uncertainties. The measurement is compared to MADGRAPH, and MADGRAPH_AMC@NLO predictions. For the leading jet p_T distribution, a prediction of Z+1jet (N_{jetti} NNLO) with non-perturbative correction applied is compared to the measurement too.

- Jet absolute rapidity

The differential cross section as a function of absolute rapidity of the first, second, third jet, are illustrated in figure 4.135. In general, the MADGRAPH_AMC@NLO prediction describes the data very well within the full uncertainties. The MADGRAPH prediction reproduces the shape of data, while has small fluctuation. It undershoots the distribution of leading jet absolute rapidity up to about 15% ranging from 2 to 2.4. Almost at the same region, it underestimates the second jet rapidity distribution up to 20%, which is also similar to the third jet rapidity distribution. The NNLO prediction agrees well with the measurement for the leading jet rapidity distribution.

- Scalar sum of jet transverse momenta

The differential cross section as a function of jets H_T by considering $N_{jets} \geq 1$, $N_{jets} \geq 2$, and $N_{jets} \geq 3$, are illustrated in figure 4.136. The MADGRAPH_AMC@NLO and NNLO predictions have good agreement with measurements in general. The MADGRAPH prediction has an undershooting of about 15% ranging from 100 GeV to 400 GeV for the jets H_T ($N_{jets} \geq 1, 2$) distributions, but it describes the third distribution well within the full uncertainties.

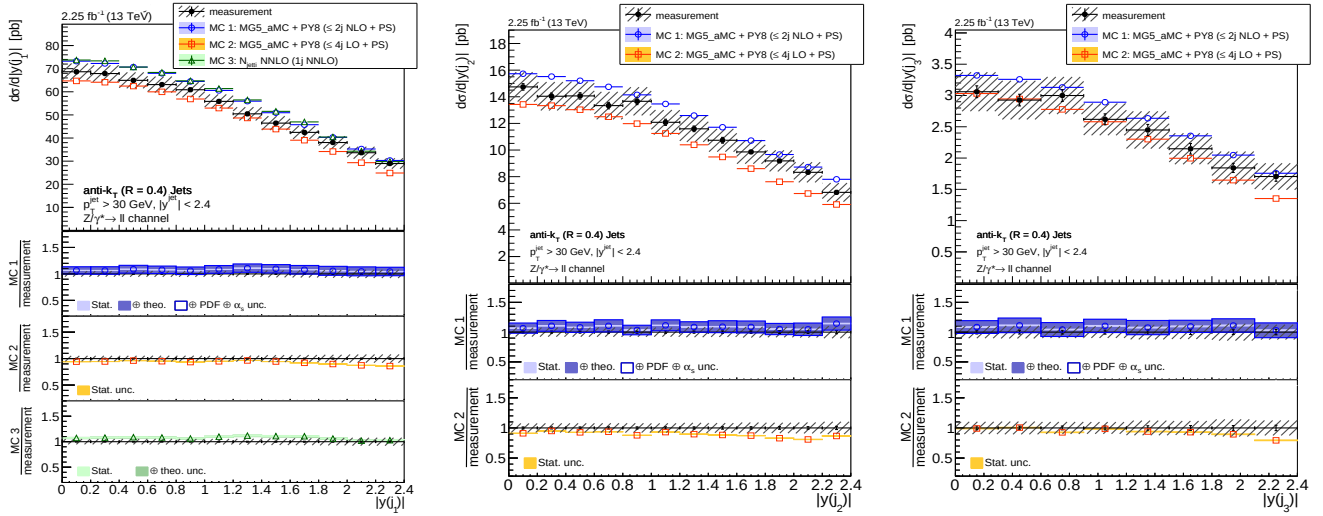


Figure 4.135: The measured differential cross section as a function of absolute rapidity of the first jet (left), second jet (middle), and third jet (right) in l^+l^- decay channel at 13 TeV, concerning the full uncertainties. The measurement is compared to MADGRAPH, and MADGRAPH_AMC@NLO predictions. For the leading jet absolute rapidity distribution, a prediction of Z+1jet (N_{jetti} NNLO) with non-perturbative correction applied is compared to the measurement too.

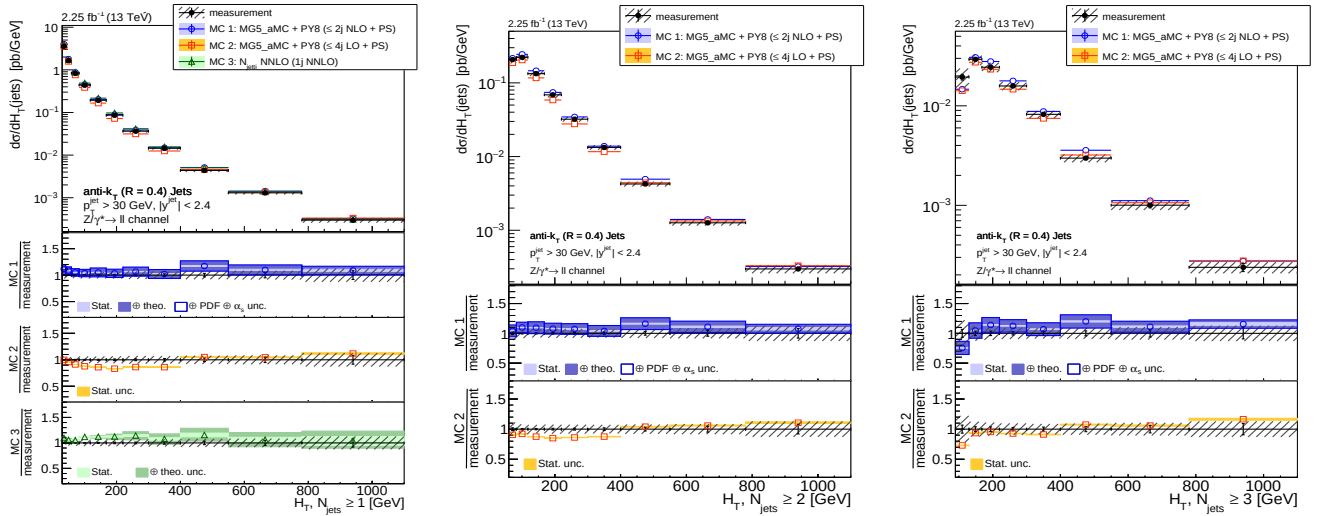


Figure 4.136: The measured differential cross section as a function of jets H_T , by requiring $N_{jets} \geq 1$ (left), $N_{jets} \geq 2$ (middle), and $N_{jets} \geq 3$ (right) in l^+l^- decay channel at 13 TeV, concerning the full uncertainties. The measurement is compared to MADGRAPH, and MADGRAPH_AMC@NLO predictions. For the distribution of H_T ($N_{jets} \geq 1$), a prediction of Z+1jet (N_{jetti} NNLO) with non-perturbative correction applied is compared to the measurement too.

4.11 Recent Results on Z+Jets Measurements

The Z boson was discovered by the UA1 and UA2 experiments at CERN in 1983, benefiting from the Super Proton Synchrotron accelerator used in proton-antiproton mode. From that moment,

measurements of Z boson plus jets started to be studied. During the early time of these measurements, the limited statistics of data and the low center-of-mass energy did not allow for high precision. After several decades, new accelerators were built, such as the Tevatron and LHC, which provide energy high enough and abundant statistics for studying the Z boson plus jets process in detail. The Tevatron is located at the Fermi National Accelerator Laboratory in US, used to provide proton-antiproton collisions with a center-of-mass energy of 1.96 TeV between 2001 and 2011, collecting the data with an integrated luminosity of 10 fb^{-1} . The LHC has provided proton-proton collisions at a center-of-mass energy of 7 TeV in 2011, of 8 TeV in 2012, and of 13 TeV in 2015. The integrated luminosity was about 5 fb^{-1} , 20 fb^{-1} , 4 fb^{-1} respectively.

Due to the differences of incident particles for collisions and center-of-mass energies, the two accelerators lead to several different consequences. For instance, the Tevatron provides higher probability of quark-antiquark annihilations to form gluons producing hard jets, while the LHC gives more chance to quark-quark scattering to originate quarks resulting in hard jets. As a result, they give rise to different ratios of quark and gluon productions in the final state. Due to the higher collision energy, the LHC provides a larger amount of heavy flavour quarks in the initial/final state than the Tevatron, and leads to a higher transverse momentum of jets. These differences result in different PDF constraints from the data.

The Z boson plus jets process has been precisely measured by two major experiments at the LHC, including ATLAS and CMS. It was also measured by the CDF and D0 experiments at the Tevatron. Because the differences in the experimental setup configurations, in order to gain the optimal precision, the analyses based on these experiments correspond to different phase spaces. Moreover, jets are defined using anti- k_T algorithm in CMS and ATLAS experiments. The CDF and D0 experiments applied Midpoint cone algorithm and seed-based cone algorithm to define jets, respectively.

This section provides a brief summary of the results of Z + jets measurements using:

- D0 data collected with center-of-mass energy of 1.96 TeV with an integrated luminosity of 1 fb^{-1} [118]
- CDF data collected with center-of-mass energy of 1.96 TeV with an integrated luminosity of 9.6 fb^{-1} [119]
- ATLAS data collected with center-of-mass energy of 7 TeV with an integrated luminosity of 4.6 fb^{-1} [115]
- CMS data collected with center-of-mass energy of 7 TeV with an integrated luminosity of 4.9 fb^{-1} [114]
- ATLAS data collected with center-of-mass energy of 13 TeV with an integrated luminosity of 3.16 fb^{-1} [120]

Those experiments make their measurements in a bit different phase spaces. For example, the p_T low threshold of the leptons from Z boson decaying in CDF and D0 analyses is 25 GeV, while being 20 GeV in CMS and ATLAS analyses. The absolute rapidity of the leptons is required to be less than 1 in CDF analysis, and less than 2.5 in D0 analysis, while 2.4 in CMS and ATLAS cases. In terms of jet selection, the p_T are required to be higher than 30 GeV in CDF, CMS and ATLAS analyses, while higher than 20 GeV in D0 case. The absolute rapidity of jets is within 2.1 (CDF), 2.4 (CMS), 2.5 (D0 and ATLAS). In the case of ATLAS 7 TeV measurement, the maximal absolute rapidity of jets is extended to 4.4.

Figure 4.137 shows the inclusive jet multiplicity distributions measured by CDF, CMS and ATLAS experiments, and the comparisons to several theoretical predictions. Benefiting from large statistics of data, those measurements can reach the multiplicity up to 4, 6 and 7 respectively. Different generator predictions show good agreements with the CDF and CMS measurements. In the case of ATLAS measurement at 7 TeV, the ALPGEN, SHERPA, and BLACKHAT predictions are consistent with measurements, while the MC@NLO has large offsets to the data for large jet multiplicities because its parton shower underestimates the observed rate of additional jet emissions. At 13 TeV, the predictions are in agreement with the observations, except for the SHERPA, ALPGEN, MG5_aMC+PY8 FxFx predictions in high jet multiplicity region, where additional jets are produced by the parton shower.

Figures 4.138 and 4.139 show the leading jet p_T distributions for the four experiments. Good agreements are seen between the theoretical predictions and the measurement by CDF. Regarding the D0 case, MCFM NLO shows excellent agreement while its LO prediction underestimates the distribution by 20%. PYTHIA QW gives a falling distribution with respect to the measurement, and the HERWIG + JIMMY shows the similar trend at high p_T region. ALPGEN + PYTHIA undershoots the observed distribution by about 20%, while SHERPA overshoots the distribution at high p_T region. In general, the theoretical (scale) uncertainties of these predictions are ranging from 10% to about 50%. In the case of CMS analysis, good agreements are seen with SHERPA2 β 2 and POWHEG + PYTHIA6, while an overestimation is seen with MADGRAPH + PYTHIA6 ranging from 150 GeV to 350 GeV, which is also confirmed in figure 4.120. Regarding the ATLAS analysis at 7 TeV, a similar excess with ALPGEN prediction ranging from 200 GeV to 450 GeV is shown, and the MC@NLO has about one order magnitude discrepancy for large p_T resulting from a quite soft spectrum in prediction. For the 13 TeV analysis from ATLAS, the MG5_aMC+PY8 CKKWL overestimates the distribution at high p_T scope, and all the other predictions are in good agreements.

The second jet p_T distributions are illustrated in figures 4.140 and 4.141. One can see excellent agreements between the theoretical predictions and the measurement by CDF. In terms of the D0 analysis, there is a factor seen as 1.1 for MCFM NLO prediction and about 1.4 for MCFM LO prediction. Both PYTHIA and HERWIG predictions have more steeply falling spectra than the observed distribution. ALPGEN + PYTHIA again undershoots the measurement, while SHERPA overestimates the observed distribution. The theoretical uncertainty ranges from approximately 30% to 50%. Regarding the CMS analysis, SHERPA tends to underestimates the high p_T region, while

both POWHEG + PYTHIA6 and MADGRAPH + PYTHIA6 overshoot the distribution by about 20-30%. In the case of ATLAS analysis, SHERPA underestimates the distribution for about 15% in the low p_T scope, and the other predictions are in good agreements.

The third jet p_T distributions are shown in figure 4.142. There is about 30% discrepancy between BLACKHAT + SHERPA prediction and the measurement by CDF. In the case of D0 analysis, the MCFM LO overestimates the distribution for about 60% with large theoretical uncertainty. Both PYTHIA and HERWIG + JIMMY show a more steeply dropping distribution versus the data distribution. The ALPGEN + PYTHIA predicts a lower cross section while the shape is consistent. The SHERPA agrees with the observed distribution. In terms of the analysis from CMS, POWHEG and MADGRAPH show good agreements with data. SHERPA2 β 2 seems to underestimate the distribution while in consistence within the large theoretical uncertainties. For the ATLAS analysis, all the predictions agree with the measurement.

Figure 4.143 shows the forth jet p_T distributions measured by CMS and ATLAS experiments at 7 TeV, respectively. In general, good agreements between measurements and all the predictions are formed. POWHEG prediction has about 30% higher prediction in the range from 70 GeV to 110 GeV. ALPGEN drops by about 20% with respect to the ATLAS measurement.

The rapidity correlation between Z boson and jet for Z + jet process at 7 TeV measured by CMS collaboration is illustrated in figure 4.144. The leptons (muons or electrons) from Z boson decaying are required to have p_T larger than 20 GeV, and their absolute pseudorapidity should be smaller than 2.1. The jet p_T is required to be above 30 GeV, and its absolute rapidity range is selected to be within 2.4 to assure good tracker coverage. The measurement is compared to several theoretical predictions in different PDF sets. Note that the differential cross section is normalized to unity by the integral cross section. For the distribution of y_{sum} , both SHERPA and MCFM predictions describe the observation well, while the MADGRAPH prediction is much less consistent with the measurement. In case of the y_{dif} distribution, the MCFM prediction agrees with the data within the range of $y_{dif} < 1.0$, and the SHERPA describes the measurement very well. The most interesting phenomenon is that the MADGRAPH prediction fails to reproduce the shape of data. This is confirmed again at 8 TeV analysis (see figure 4.126).

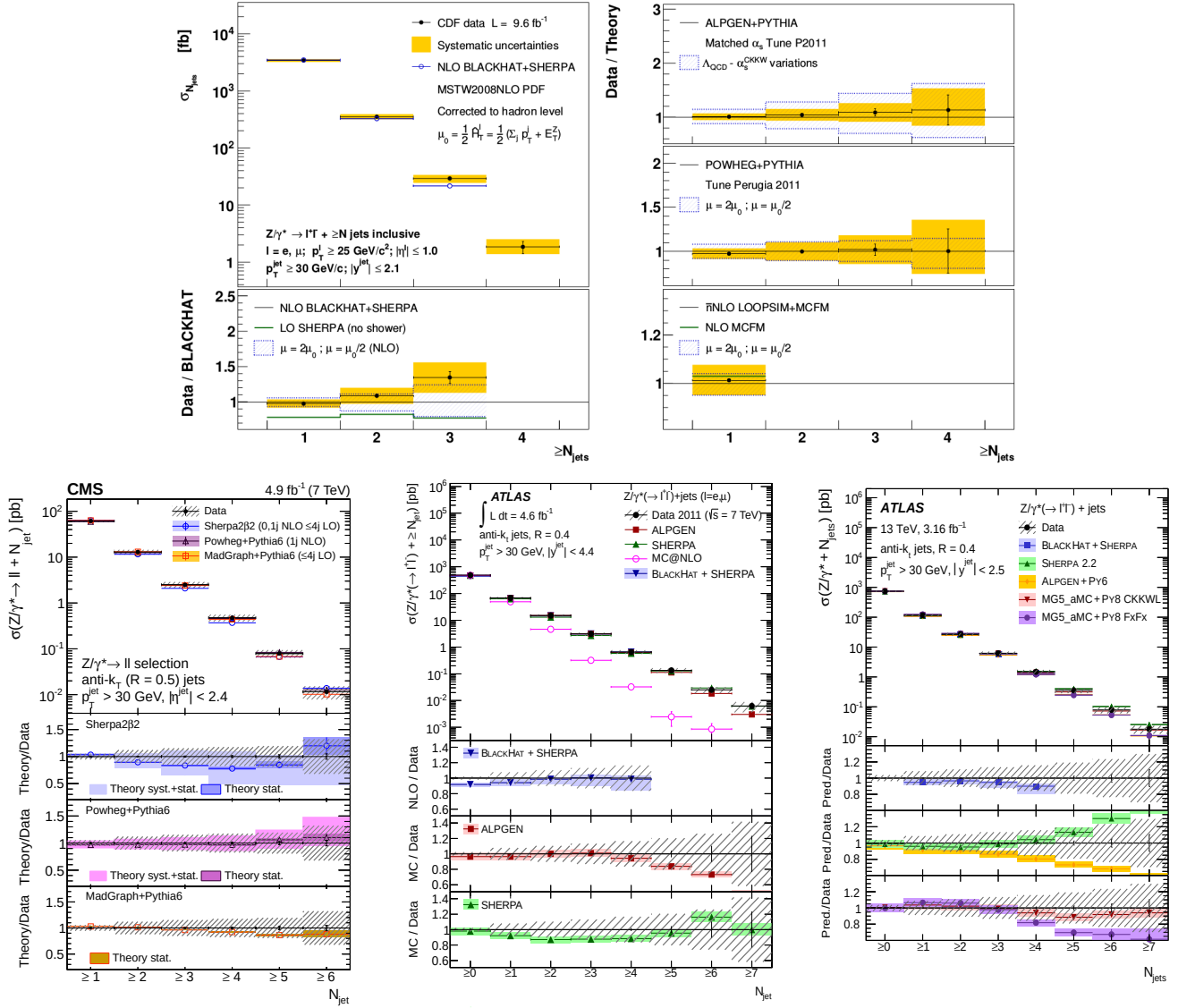


Figure 4.137: The differential cross section as a function of inclusive jet multiplicity for $Z + \text{jets}$ process measured by CDF experiment with $\sqrt{s} = 1.96$ TeV (top), by CMS experiment with $\sqrt{s} = 7$ TeV (bottom left), by ATLAS experiment with $\sqrt{s} = 7$ TeV (bottom middle) and with $\sqrt{s} = 13$ TeV (bottom right). The measurements are compared to different theoretical predictions. Figures are taken from [119, 114, 115, 120].

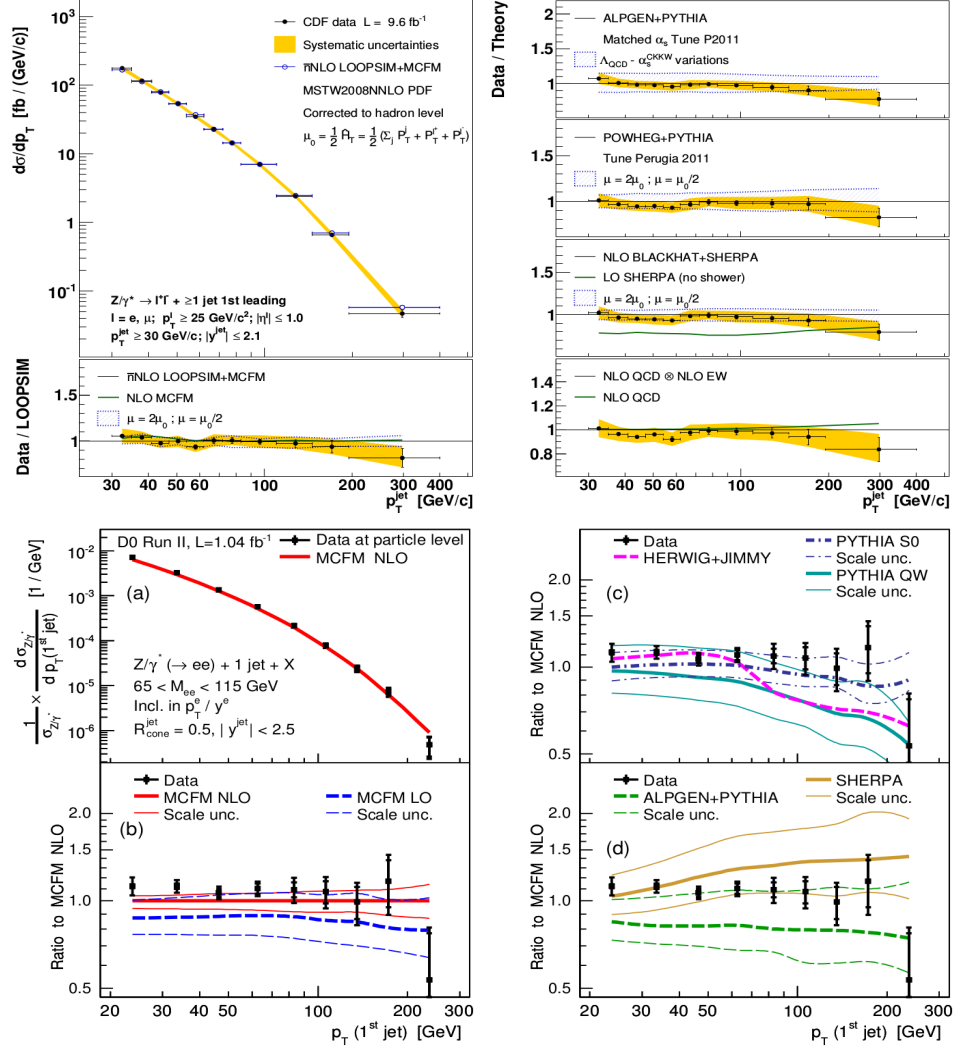


Figure 4.138: The differential cross section as a function of leading jet p_T for $Z + \text{jets}$ process measured by CDF experiment (top) and by D0 experiment (bottom) with $\sqrt{s} = 1.96$ TeV. The latter is normalized to the total cross section. The measurements are compared to different theoretical predictions. Figures are taken from [119, 118].

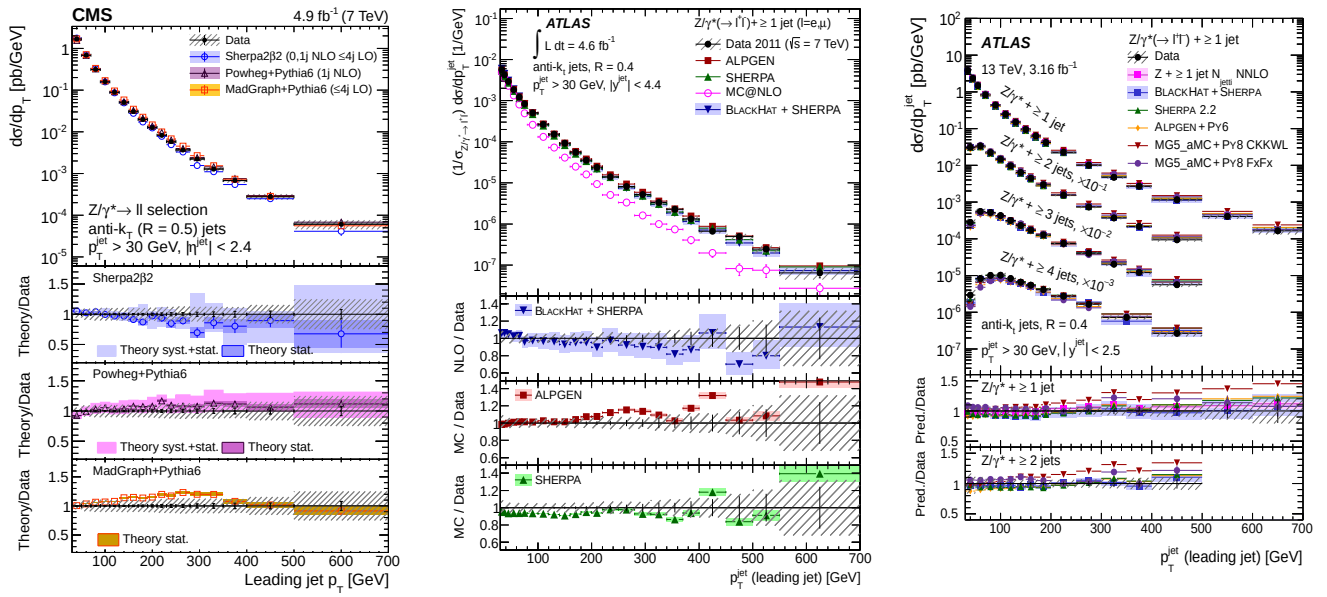


Figure 4.139: The differential cross section as a function of leading jet p_T for $Z + \text{jets}$ process measured by CMS experiment with $\sqrt{s} = 7$ TeV (left), by ATLAS experiment with $\sqrt{s} = 7$ TeV (middle) and with $\sqrt{s} = 13$ TeV (right). The measurements are compared to different theoretical predictions. Figures are taken from [114, 115, 120].

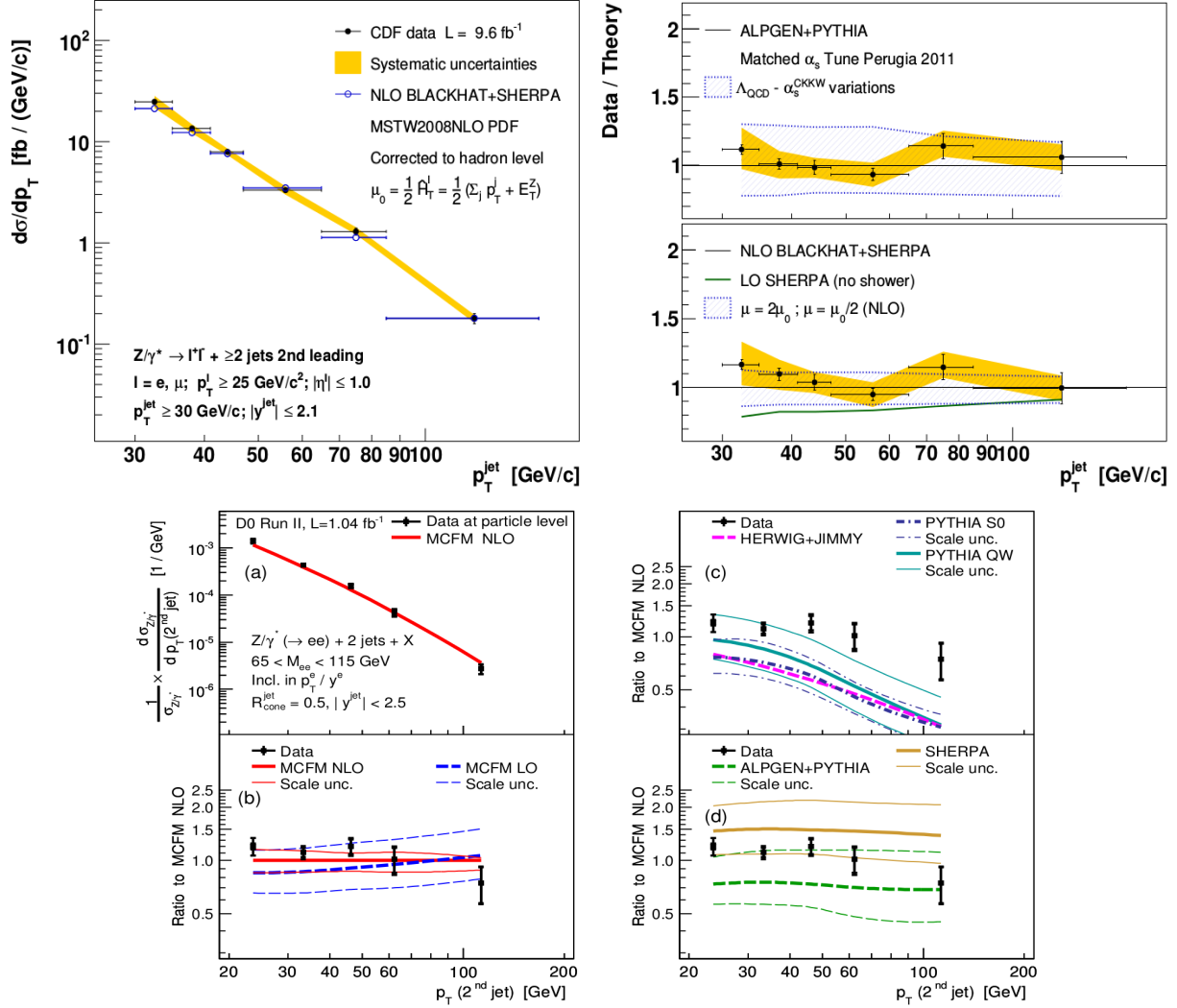


Figure 4.140: The differential cross section as a function of the second jet p_T for $Z + \text{jets}$ process measured by CDF experiment (top) and by D0 experiment (bottom) with $\sqrt{s} = 1.96 \text{ TeV}$. The latter is normalized to the total cross section. The measurements are compared to different theoretical predictions. Figures are taken from [119, 118].

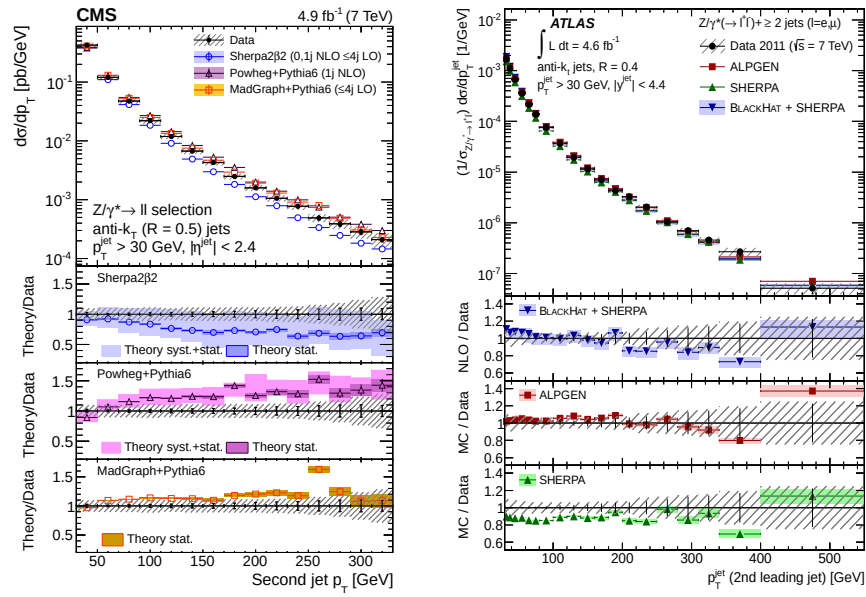


Figure 4.141: The differential cross section as a function of the second jet p_T for $Z + \text{jets}$ process measured by CMS experiment with $\sqrt{s} = 7$ TeV (left), by ATLAS experiment with $\sqrt{s} = 7$ TeV (right). The measurements are compared to different theoretical predictions. Figures are fetched from [114, 115].

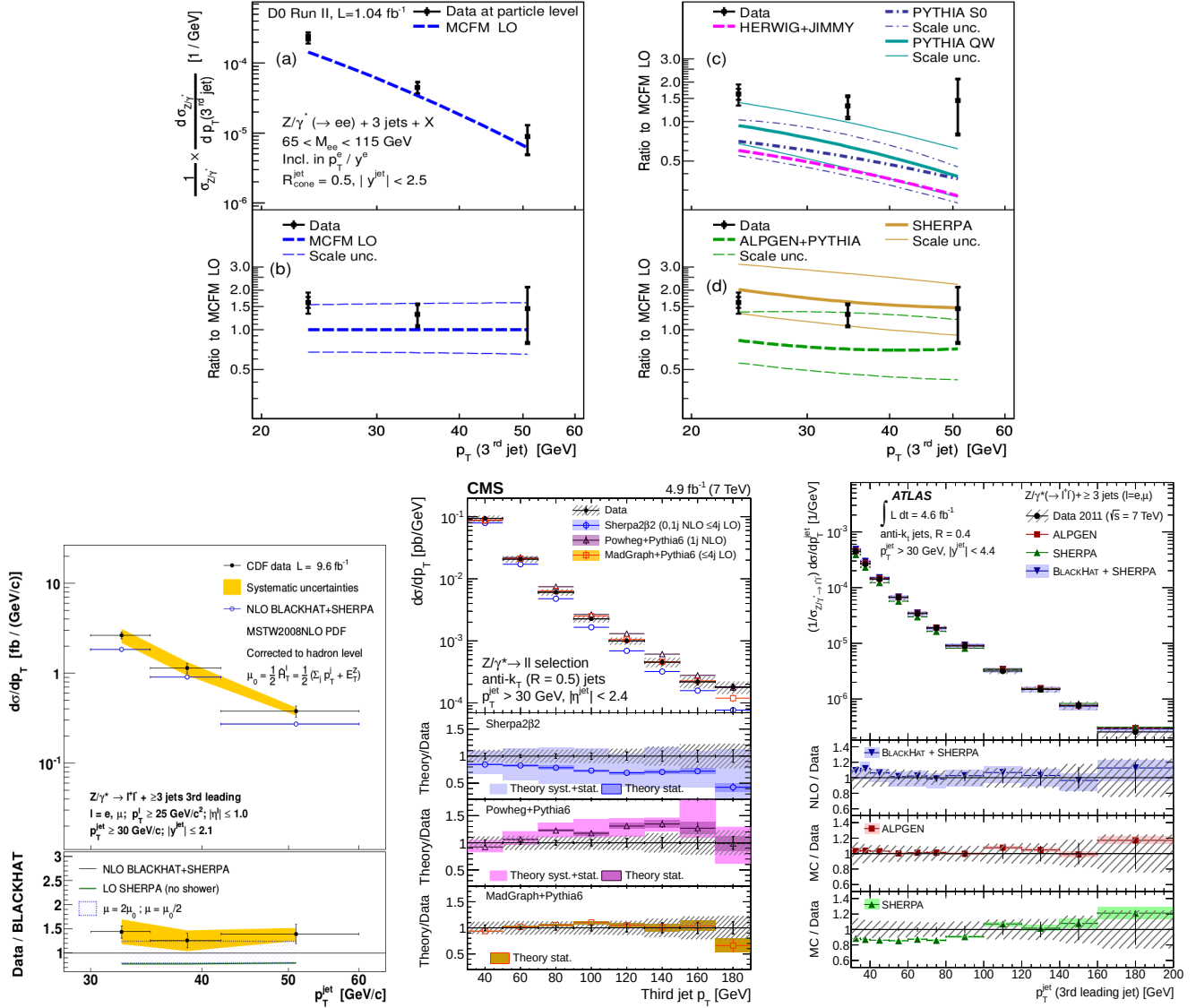


Figure 4.142: The differential cross section as a function of the third jet p_T for $Z + \text{jets}$ process measured by D0 experiment with $\sqrt{s} = 1.96 \text{ TeV}$ (top), by CDF experiment with $\sqrt{s} = 1.96 \text{ TeV}$ (bottom left), by CMS experiment with $\sqrt{s} = 7 \text{ TeV}$ (bottom middle), by ATLAS experiment with $\sqrt{s} = 7 \text{ TeV}$ (bottom right). The measurements are compared to different theoretical predictions. Figures are taken from [119, 118, 114, 115].

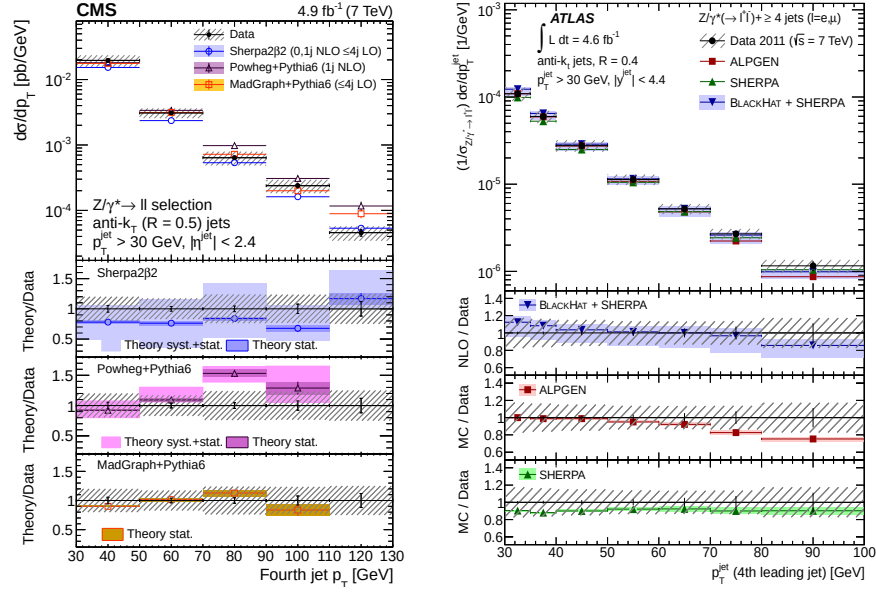


Figure 4.143: The differential cross section as a function of the forth jet p_T for Z + jets process measured by CMS experiment with $\sqrt{s} = 7$ TeV (left), by ATLAS experiment with $\sqrt{s} = 7$ TeV (right). The measurements are compared to different theoretical predictions. Figures are fetched from [114, 115].

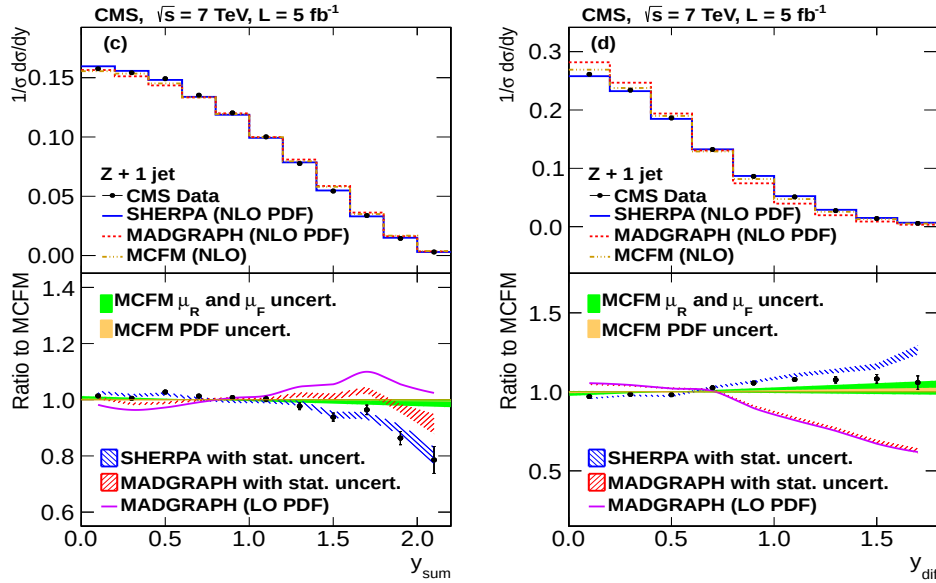


Figure 4.144: The differential cross section as a function of absolute rapidity sum (left) and difference (right) of Z boson and jet for Z + 1 jet process. It is normalized to unity. The measurement is compared to different theoretical predictions with NLO or LO accuracy. Figures are taken from [61].

Chapter 5

Jet energy correction and calibration at high level trigger

This chapter presents the development of the jet energy correction (JEC) for an upgrade of the high level trigger (HLT) of the CMS detector for LHC RunII (run at 13 TeV). The jet energy corrections are determined and implemented to HLT jet objects for full calibration. After being successfully calibrated, they are configured in the database for the data-taking of the detector and MC sample productions. This work is presented starting from its motivation, followed by the description of physical assignment and technical implementation. Finally, some relevant results are presented. The calibrated configurations were applied for the data-taking in 2015 and 2016.

5.1 Motivation

At high level trigger, the jet reconstruction is influenced by the detector effects, such as inactive regions in the tracker system, the non-linear response to energy depositions, and the finite energy resolutions of calorimeters. What's more, the pileup interactions contribute to bias its energy. As a consequence, the reconstructed momentum of a jet at HLT differs from the real produced momentum of the jet at particle level. This difference varies in different spatial regions of the detector, and depends on the instantaneous luminosity of each bunch crossing, which results in varying pileup interactions. The spectrum of reconstructed jet momenta can be seen as a dispersion with respect to the distribution of produced jet momenta. As a consequence, the jet energy correction is needed to subtract the pileup energy contamination, to compensate the detector inefficiencies in different angular directions, and to adjust the reconstructed jet momentum to the produced jet momentum. It is important for HLT to possess the corrected jet momenta, such that the triggers with specific jet momentum thresholds can efficiently filter the jets carrying the momenta below the thresholds. Then the trigger may have a proper rate. For a physical point of view, efficient jet triggers are important for the precision measurements of QCD and the searches of new physics with jet productions in the final state.

5.2 Track and vertex reconstruction at HLT

The HLT takes advantage of the same software as the offline track reconstruction (section 4.3.2). Due to the timing constraint at HLT, the reconstruction configurations are simplified related to offline case. For the track reconstruction, the HLT uses the hits in both pixel detectors and strip detectors. It may only take into account the pixel hits to save time, depending on the trigger type. Then during the iterative tracking step, these hits are used for track seeding, when the HLT may implement only one iteration with high p_T requirement for seeding to reduce the time. For some triggers, track reconstruction is performed in a specific spatial region for a timing requirement. During the iterative track reconstruction, the interaction vertex is reconstructed using a simple gap-clustering algorithm [53] for the associated tracks, which is also called pixel vertex. Once a track seed is compatible with the position of reconstructed pixel vertex along the beam direction, and satisfies the specific tracking conditions (like minimal p_T threshold, angular ranges), this seed is kept for the next step. Meanwhile, in order to reduce time expense, the iterative track building program stops.

By means of such patterns above, the vertex reconstruction at HLT is much faster (time reduced by a factor of 4 - 6) than the offline standard vertex reconstruction. Figure 5.1 shows the amount of reconstructed HLT pixel vertices versus the expected number of primary interactions, comparing the two different pileup conditions. We can see the good linearity of the HLT pixel reconstruction. The figure also provides the comparison of the tracking efficiency for HLT (online) and offline track reconstruction. Due to the inefficiency of the pixel detector for hits, the HLT loses some isolated tracks.

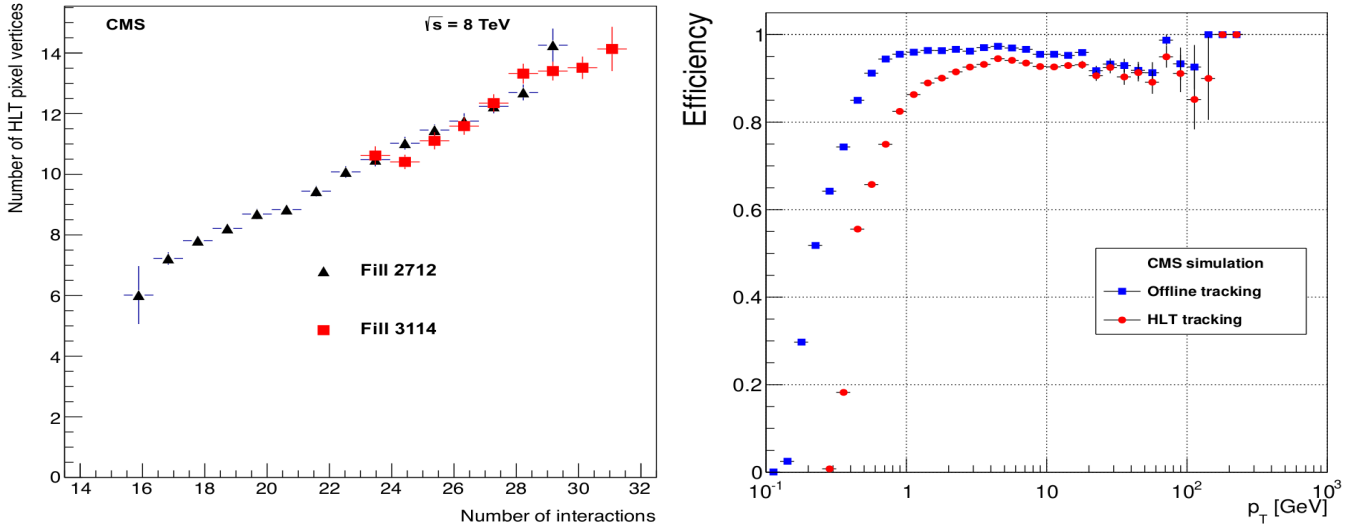


Figure 5.1: Left: Number of vertices as a function of the number of proton-proton interactions measured by the forward calorimeter with the fills taken in two different pileup conditions at 8 TeV. Right: Tracking efficiency versus the reconstructed track momentum from simulated $t\bar{t}$ events at 8 TeV, comparing the online and offline methods. The figures are taken from [53].

5.3 Jet reconstruction at HLT

At HLT, the jets are reconstructed using the anti- k_T algorithm for the offline jet reconstruction (section 4.3.7). The parameter R is chosen as 0.5 at 8 TeV while as 0.4 at 13 TeV. This change is because the smaller value reduces the pileup contamination. The jet candidates for the input of this algorithm can be from either calorimeter towers (Calojet) or particle flow objects (PFjet). The former is reconstructed according to only the energy deposits in the calorimeters, while the latter implements a particle flow algorithm similar to the offline reconstruction (see section 4.3.3). Because the track reconstruction scenario at HLT is different from offline track reconstruction, and some tracks are lost at HLT, the PFjets at HLT differ from those in offline case. The Calojets cost much less time to be reconstructed than the PFjets, so the PFjets can be selected from Calojets with their combinatorial algorithms. While, before this procedure is complied, one needs to calibrate the Calojets and PFjets, and they should be matched.

The efficiency of jet trigger at HLT is important, which affects the quantity of collected data. The typical jet trigger paths are specified with thresholds of reconstructed jet transverse momenta to select jet candidates. In principle, a specific threshold for a jet trigger path results in a turn-on step from zero to unity for the efficiency versus the reconstructed jet transverse momentum at HLT, which means that only and all of the jets with transverse momenta beyond the threshold can pass the trigger. In practice, people concern more about jet trigger efficiency versus the offline reconstructed jet transverse momenta. Because the offline jet reconstruction may be based on the data selected by jet triggers, the reconstructed jets at HLT and at offline level are correlated, but have different behaviors due to the different reconstruction scenarios. As a consequence, the trigger efficiency as a function of offline reconstructed jet p_T gives rise to a slanting turn-on curve, which reflects the energy resolution between HLT jets and offline jets. Figure 5.2 shows examples of jet trigger efficiencies versus the offline jet p_T with different trigger paths (different HLT jet p_T thresholds). The more steep turn-on curve to the plateau reflects the more precision of threshold application of trigger path, which has a high efficiency for the jets with p_T above the threshold while a large rejection for the jets with p_T under the threshold.

5.4 Pileup at HLT

As for offline jet reconstruction, the online jet reconstruction also suffers from pileup jet contamination. The inclusive cross section of proton-proton collision increases from 69.4 mb at 8 TeV to 71.3 mb at 13 TeV [121]. At 8 TeV, the average number of pileup interaction was 20, with the bunch spacing of 50 ns. Because of the increased instantaneous luminosity and reduced bunch spacing time (25 ns) compared to 8 TeV, the mean amount of pileup interaction at 13 TeV increased to 30-40. By consequence, the overlapping energy deposited in the detector from the in-time pileup and out-of-time pileup at 13 TeV were increased significantly. This means a large contribution of artificial jets from multiple proton-proton interactions, and these jets increased the rates of jet triggers, which

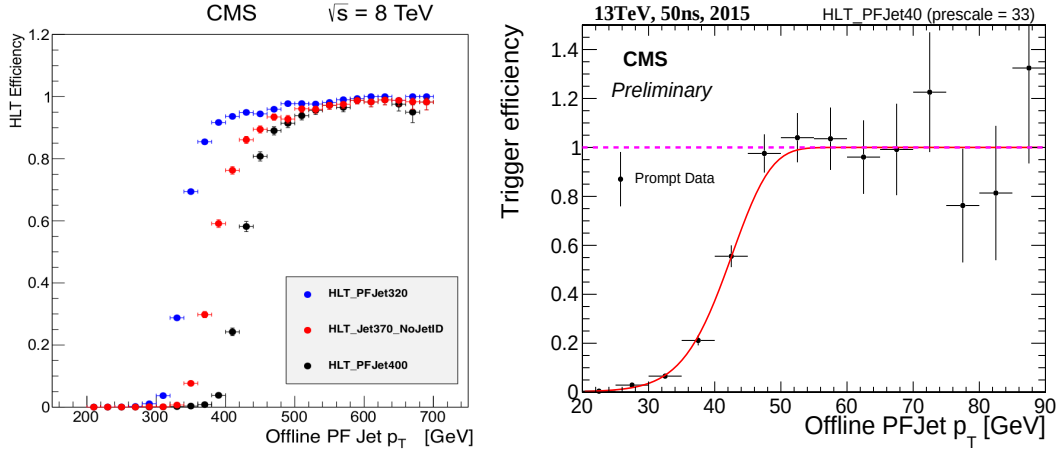


Figure 5.2: The HLT efficiency as a function of p_T for Calojet and PFjet trigger at 8 TeV (left), and of PFjet trigger at 13 TeV with 50 ns bunching spacing (right). In the plots, the numbers in the trigger paths stand for the p_T thresholds of HLT jets. The left figure is taken from [53]. Note that the right plot shows the efficiency above unity with prescale corrected in some bins, which is due to statistical fluctuations.

changed the trigger efficiency and required a higher threshold. Therefore, we aimed to distinguish these energy contaminations and subtract them from reconstructed jet cluster energy. Then we could concentrate on the jets produced in physical processes of interests, such as $Z + \text{jets}$ process.

Figure 5.3 shows the ratio of the amount of pileup jets to the total number of reconstructed jets using three different reconstruction scenarios. The first one is the default hadron calorimeter reconstruction at 8 TeV, which uses the combinatorial information of tracks and calorimetric energy deposition to reconstruct jets. In order to extract pileup contributions, a template fit approach is implemented. Based on the foretime run calibration, different template pulses corresponding to different bunch crossings that formed in the grouped crystals of ECAL are recorded in the database¹. For a reconstructed energy profile with a specific bunch crossing, ten different sampling pulses are selected accordingly and are combined to fit this profile [122]. The similar approach is applied in HCAL for the pileup jets. According to the fitted results, one can extract the in-time pileup energy, while it does nothing for the out-of-time pileup contribution. During RunII with 25 ns bunch spacing, the OOT pileup contribution is significant, hence the second and third patterns have been developed. Thanks to the RunII period with 50 ns bunch spacing, we could perform the first method to estimate the in-time pileup contribution for the case of 25 ns bunch spacing. Considering the time resolution of HCAL, the new patterns select only three sampling pulses in HCAL to perform the fitting. Then one could extract the OOT pileup contribution together with the in-time pileup [123]. On the other aspect, the second pattern reconstructs all the track candidates, which is suitable for offline reconstruction while not allowed at HLT due to consuming plenty of time. At this point, the third pattern is an adaptation of the second one, while implementing with deterministic fit approach [124] for track reconstruction to improve the timing feature for HLT. These three scenarios are proceeded

¹The number of interactions per bunch crossing has a probability density yielding a Poisson distribution.

in the simulations of QCD jet production at 13 TeV. One can see that the second and third scenarios effectively reduced the pileup contamination in the jet clusters.

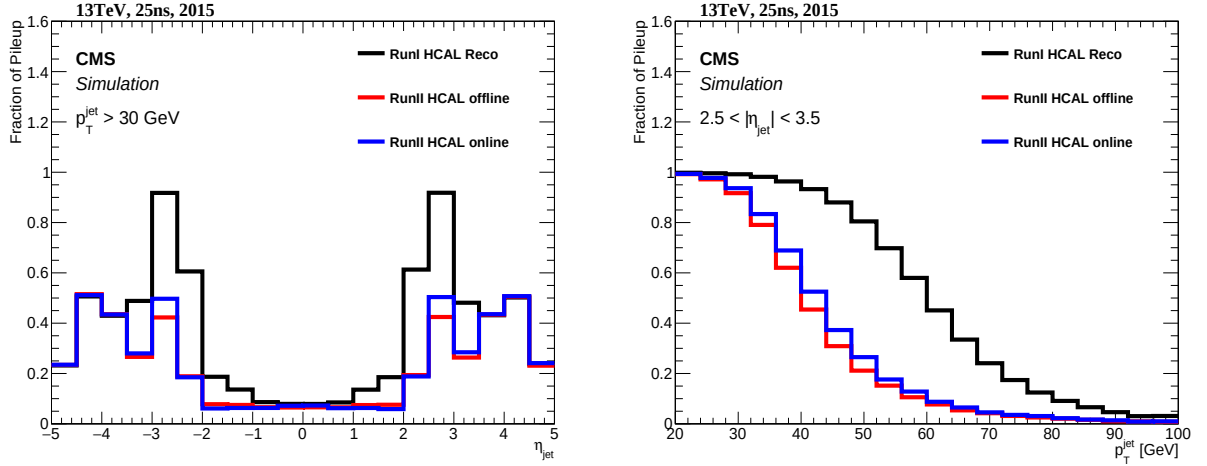


Figure 5.3: The rate of pileup jets in all of the reconstructed jets using simulation at 13 TeV with bunch spacing of 25 ns, as a function of reconstructed jet η (left) and of jet p_T (right), comparing three different reconstruction scenarios.

The pileup offset is defined as the difference between the vectorial sum of p_T of all the jets and the p_T of jets coming from hard scattering. It is an observable for monitoring the quality of pileup identification of reconstruction pattern. Figure 5.4 shows the offset p_T divided by number of pileup for 25 ns and 50 ns of time separation between each bunch crossing at HLT and offline reconstruction. The similarity of the distributions for 25 ns and 50 ns bunch spacings is achieved by use of the third reconstruction scenario to mitigate the effect of “OOT pileup”. With the track and energy deposit of particles, we can distinguish different categories of objects. The offline case has abundant time to reconstruct robust vertices. In comparison, the HLT runs at much higher frequency to treat the incoming data, so it only reconstructs pixel vertices. Relatively, the pileup offset of offline method is more reasonable than that of online method, so we need to adjust HLT reconstruction to improve the confidence level. Especially, the tracking at HLT is optimized to take into account only the primary vertex of the hard interaction², mimicking the behavior of removing charged hadrons from pileup vertices in the offline reconstruction.

5.5 Jet energy correction

Due to the pileup energy deposition within the cone size of reconstructed jet clusters, the observed jet momenta are generally larger than the produced jet momenta. Furthermore, the jet reconstruction is also influenced by several detector effects, such as the less active regions, non-linear response to

²At the same bunch crossing, the vertex with the largest momentum is typically from hard (physical) interaction, and the others are from pileup.

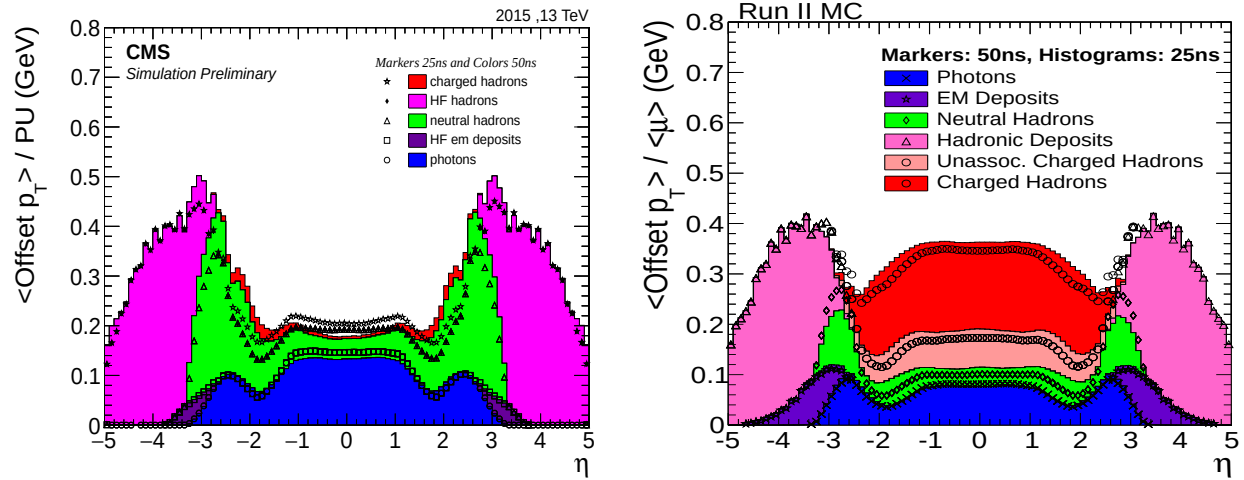


Figure 5.4: Average offset per pileup event (“PU” or “ $\langle \mu \rangle$ ”) in the HLT reconstruction (left) and offline reconstruction (right) on different categories of jets, with the comparison of the 25 ns and 50 ns of separation time between each bunch crossing. Those kinds of jets are respectively: charged hadrons, neutral hadrons, forward hadrons (“HF hadrons” (left) and “Hadronic Deposits” (right)), forward electrons (“HF em deposits” (left) and “EM Deposits” (right)), photons. The “Unassoc.Charged Hadrons” in the right plot represents the charged hadrons without track. More explanations can be found in [90].

the jets with low energy, detector noise, and limited resolution of the detector. These effects result in the reconstructed jet momentum spectrum is dispersed with respect to the true jet momentum spectrum. Therefore, the jet energy correction program is divided by three steps to correct each of them in order, which is similar to the offline case (section 4.3.7). The corrections in different levels are determined as scaling factors for the reconstructed jet transverse momenta, by means of MC simulations.

5.5.1 Pileup energy subtraction (L1 correction)

There are two different approaches for the L1 correction. The first one is called random cone method. This method uses a neutrino particle gun sample, which is generated for pure pileup simulation. Since neutrinos cannot be directly reconstructed, they are not included in the reconstructed pileup jets. The IT pileup and OOT pileup are fully distinguished in the simulation. As is shown in figure 5.5, the simulated pileup ingredients include charged hadrons, neutral hadrons, and photons. The $\eta - \phi$ plane is divided in fine η slices. For each slice, one constructs a circle of radius R centered in the middle of the slice and at random ϕ angle event by event. In such a circle, the total four-momentum of the particle candidates is computed for each event. Through considering the average over all the events across the ϕ ranges, one can obtain the total transverse momenta of the particles in the circle as a function of η and the number of IT pileup.

As mentioned in section 4.3.8, the amount of pileup can be estimated through calculating the

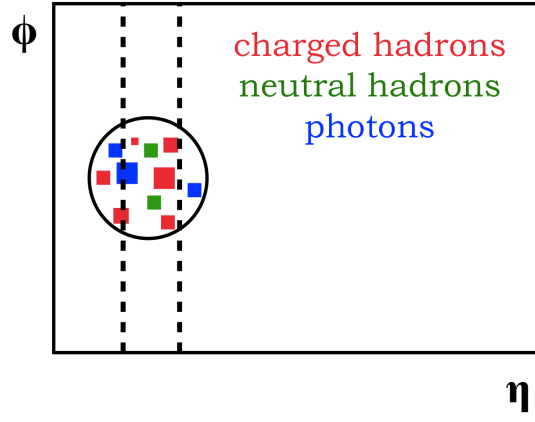


Figure 5.5: The profile of random cone method for pileup energy correction.

diffuse offset energy density ρ , which is defined as the median of jet transverse momenta (p_T) divided by the their area ($A \approx \pi R^2$), where this R is the cone size of a jet cluster. This density is insensitive to the jets produced in a hard QCD interaction, which reduces the contribution from the hard jets. By taking advantage of this variable, we can parametrize the correction as below:

$$C(\eta) = \frac{1 - \beta(\eta) * \rho * A}{p_T} \quad (5.1)$$

where p_T stands for the total reconstructed jet transverse momentum before correction, and the β is a jet η dependent parameter that corresponds to the fitted slope of offset p_T/A versus ρ , as illustrated in figure 5.6 for instance. The extrapolation to $\rho = 0$ is driven by noise, residual effects, and non-linearities [125].

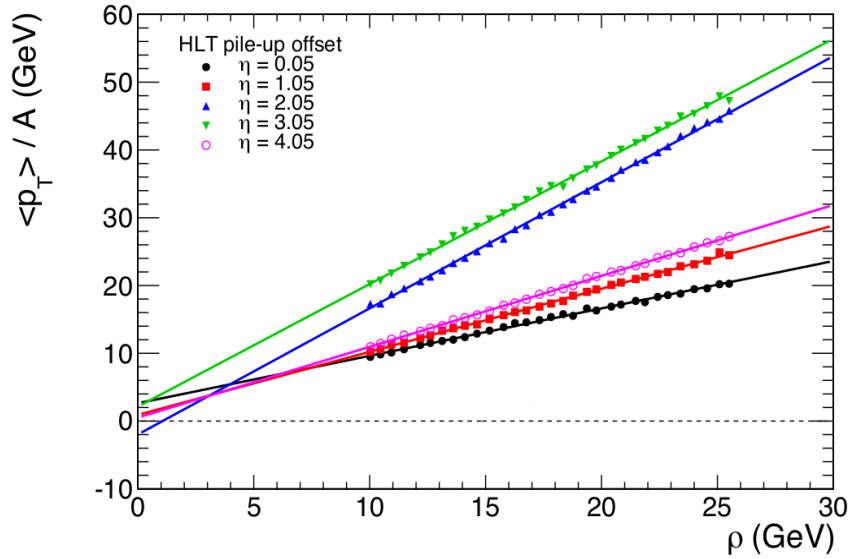


Figure 5.6: The pileup offset versus jet density in different η ranges in simulations (25 ns bunch spacing).

The second pattern for L1 correction is called hybrid jet area approach. It does not only take into account the jet area and offset density for the η dependent corrections, but also adds a logarithmic reconstructed jet p_T (uncorrected) dependent term for the parametrization:

$$C(\eta) = \frac{1 - [\rho_0(\eta) + \rho * \beta(\eta) * (1 + \gamma(\eta) * \log p_T)] * A}{p_T} \quad (5.2)$$

where the $\rho_0(\eta)$ contains the ingredient of multiple parton scattering density [89]. The $\beta(\eta)$ corrects the non-uniformity of IT and OOT pileup offsets. The $\gamma(\eta)$ is the coefficient of the logarithmic p_T dependence. These three parameters are obtained from shaping of the offset correction versus jet η , and they can be determined using simulations. In our case, we use two MC samples with enriched QCD productions, and they are generated by the same MC generator with the same events. The difference between them is that one sample is simulated with pileup, the profile of which is usually flat with respect to the expected nominal number of pileup interactions. The average number of pileup bins are mapped to the average value of ρ for the subsequent parametrization. Note that the OOT pileup is also included in the simulation. The other sample is simulated without pileup, which is used for calibrating the pileup energy subtraction. Then we loop on the corresponding events of the two samples at the same time. Once we find a reconstructed jet in the pileup sample is within the distance ($\Delta R < 0.2$ at 13 TeV) to a reconstructed jet in the no-pileup sample, we match them assuming that they are from the same jet at particle level³. According to the differences of p_T of the matched couples, the parametrizations are conducted for the corrections.

In order to evaluate the L1 correction, we can perform the jet response as a function of true number of pileup. In MC truth, the jet response is defined as the ratio of reconstructed jet to the original jet at particle level. They should be matched according to their distance ($\Delta R < 0.2$). This option ensures a high matching efficiency and provides a unique match for the anti- k_T defined jets. After the L1 correction is applied, the jet response should be insensitive to the number of pileup, which means the profile is a flat distribution. At HLT, the flatness of this distribution is specially important for the stable jet trigger rate in various pileup contributions.

Through comparing the two L1 correction methods, we have found that the second one is much better, especially at low jet p_T region. Figure 5.7 illustrates a representative comparison of the jet response using the same MC sample, applied with the corrections determined by the two methods respectively.

5.5.2 Relative (L2) and absolute (L3) corrections

Because the components in different longitudinal directions along the CMS detector have different reconstruction efficiencies and energy/time resolutions, the jet response distribution with respect to the whole longitudinal scope is not uniform. In order to correct this detector effect, the L2 correction

³It has been checked that the matching efficiency, for the reconstructed jets in no pileup simulated samples and the reconstructed jets in samples with average pileup interactions of 40, is above 90% when requiring jet $p_T > 30$ GeV.

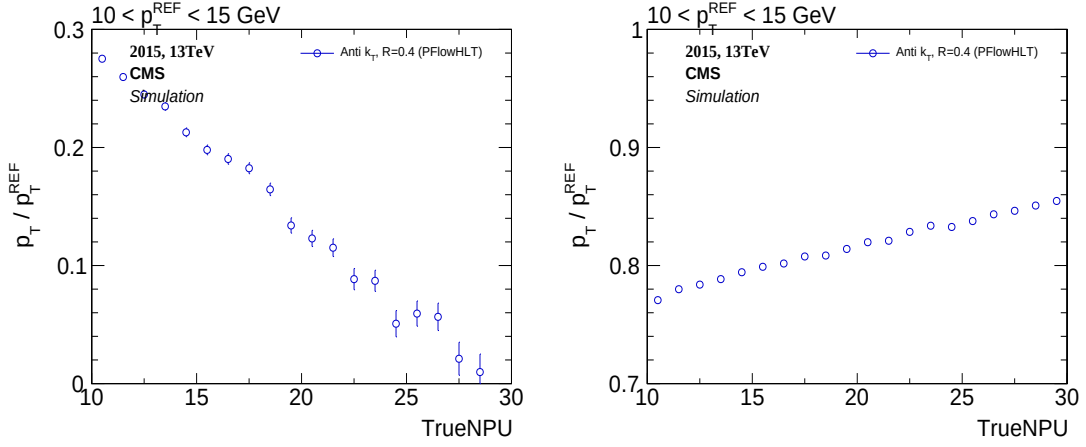


Figure 5.7: Jet response as a function of true number of pileup interactions, applied with the corrections determined by random cone method (left) and hybrid jet area method (right), using simulations at 13 TeV. The PFjets are used with the anti- k_T algorithm (distance parameter $R = 0.4$). The “ p_T^{REF} ” is p_T of the original jets in particle level that are matched with reconstructed jets at HLT.

is designed to retrieve this uniformity. The correction is factorized as a function of pseudorapidity (η).

The detector system has non-linear response to the jets with different transverse momenta, especially when jets carry p_T below a few tens of GeV. In addition, the limited resolution of the detector leads the reconstructed jet p_T spectra smeared. At this point, the L3 correction has to be applied, which brings the jet response back to unity in different jet p_T regions. It can also helps to improve the jet p_T resolution. Therefore, this correction is parametrized to reconstruct the jet p_T .

This two levels of corrections are determined using simulations, and are typically derived from the jets with L1 correction already applied. Hence, these two corrections are correlated with the L1 corrections. In a practical way, the L2 and L3 corrections are usually parametrized together. By use of MC samples, we can extract all the generated jets at particle level, and match each of them with a reconstructed jet if their distance $\Delta R < 0.2$. According to the p_T difference between each matched couple, we can calculate the correction factor as a function of reconstructed jet p_T in different η bins. This is based on a fitting approach, which provides smooth curves to describe the distributions of the differential correction factors. We mainly have Calojets and PFjets at HLT, and their response performances are different, so we have two basic formulas to parametrize the correction for them respectively:

- Calojets:

$$C'(\eta) = \alpha_0 + \frac{\alpha_1}{(\log_{10} p_T)^{\alpha_2} + \alpha_3} \quad (5.3)$$

- PFjets:

$$C'(\eta) = \alpha_0 + \frac{\alpha_1}{(\log_{10} p_T)^2 + \alpha_2} + \alpha_3 * \exp(-\alpha_4 * (\log_{10} p_T - \alpha_5)^2) \quad (5.4)$$

where the α_i ($i = 0, 1, 2, 3$ for Calojets while $i = 0, 1, \dots, 5$ for PFjets) are various parameters for the fitting in each η bin. Note that the p_T now stands for the reconstructed jet p_T after pileup energy subtraction. Figure 5.8 exemplifies the fitted L2 and L3 correction distributions with respect to reconstructed jet p_T in three different η ranges, each of which belongs to the barrel, endcap, and HF regions, respectively. The ranges of those curves depend on the effective p_T scopes of the reconstructed jets in each η bin. In general, the correction factors of PFjets are much smaller than the ones of Calojets at low p_T region, while they approach to each other at high p_T values. Because particle flow reconstruction relies heavily on tracking for the jets with low p_T , the sensitivity to the detector calibration is significantly reduced related to only calorimetry reconstruction. On the other hand, the jets with high p_T usually traverse the tracker system with too straight tracks for a reliable p_T measurement, when the tracking efficiency decreases in the dense jet kernel due to the limited energy resolution. Meanwhile, the deeper hadronic showers happen in the calorimeters, which enlarges the sensitivity of PFjets to the detector calibration.

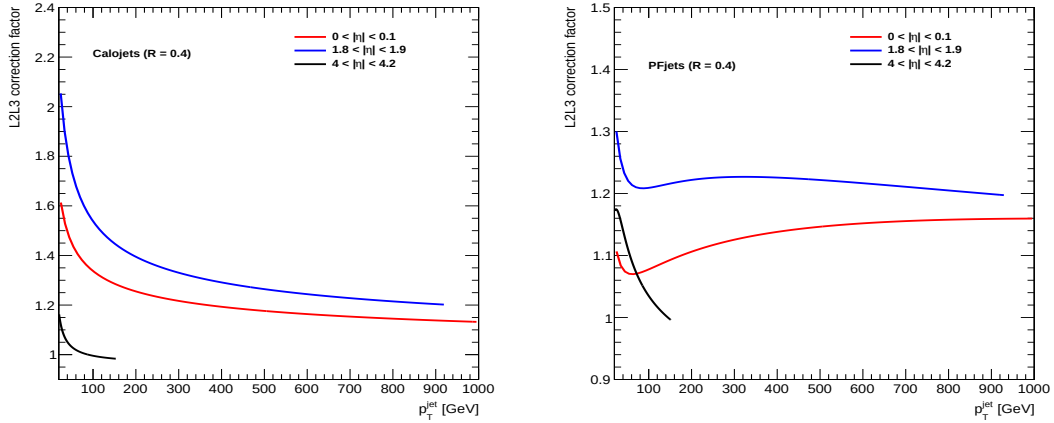


Figure 5.8: Combined L2 and L3 correction factor as a function of reconstructed jet p_T in different η regions, using fitting approach. The Calojets (left) and PFjets (right) at HLT are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

5.5.3 Residual L2L3 corrections

With the full jet energy corrections applied, the data and simulations may still have some discrepancies, which can result from overestimated jet energy resolution by simulations, and also from inactive segments of the detector system that are not performed effectively in simulations. In order to compensate this drawback at HLT, the residual correction is designed, i.e. performing the calibration with data. Similar to the offline residual correction introduced in section 4.3.7, the correction is factorized as a function of jet p_T in different η regions. It can be determined from the jet response

difference of data and MC using the p_T balance approach. Through considering the events with a reconstructed Z boson in association with a jet, which are typically recoiling back to back, we treat the Z boson p_T as the denominator (reference) of jet response to proceed ⁴. Figures 5.9 to 5.11 display the jet response comparisons between data and MC (Drell-Yan) simulations for PFjets in different η ranges using p_T balance method, before and after the previous three levels of corrections applied. The selected Z boson are with p_T larger than 20 GeV. In general, the residual correction factors (data/MC) are close to unity.

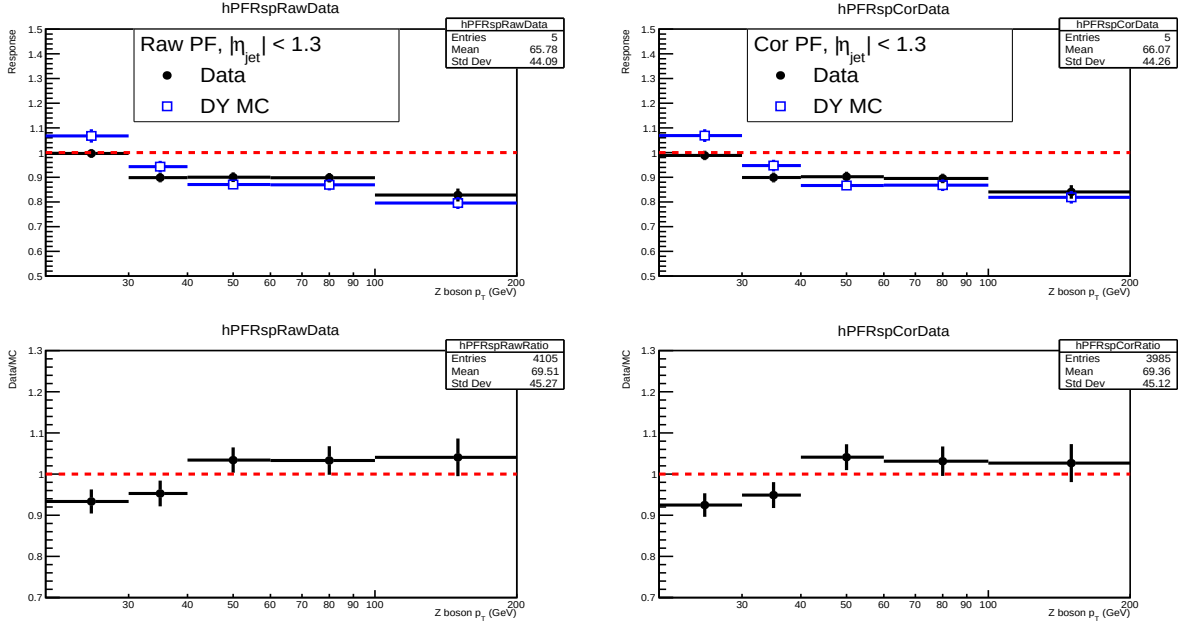


Figure 5.9: HLT PFjet response as a function of reconstructed Z boson p_T using p_T balance approach, comparing data and simulations before (left) and after (right) L1/L2/L3 corrections applied for $|\eta_{jet}| < 1.3$. The PFjets are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

5.5.4 Calibration workflow in simulations and results

The determination and calibration for jet energy correction at HLT are realized using the JetResponseAnalyzer [126, 127], which is a modular program based on CMSSW software (see section 3.2.7). As mentioned before, the calibration for the jet response is conducted by means of MC simulations. The special MC samples used for determining and calibrating the jet energy corrections (JEC) prepared for the data taking in 2016 are listed in table 5.1.

Based on the large statistics of the sample with QCD production, we can calculate the jet response for each matched jet couple event by event. Because the reconstruction system has finite resolution for the jet momenta, the profile of the jet response yields approximately a Gaussian distribution or

⁴Once more additional jets appear in the events, they likely take off some momenta, which may break the p_T balance between the Z boson and the leading jet.

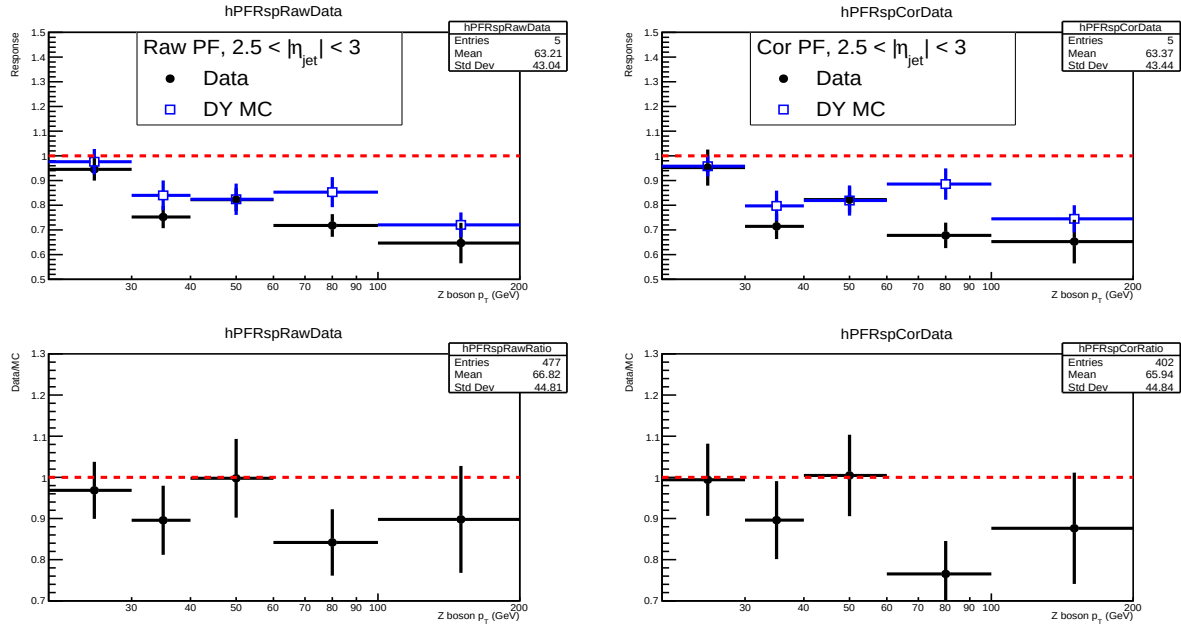


Figure 5.10: HLT PFjet response as a function of reconstructed Z boson p_T using p_T balance approach, comparing data and simulations before (left) and after (right) L1/L2/L3 corrections applied for $2.5 < |\eta_{jet}| < 3$. The PFjets are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

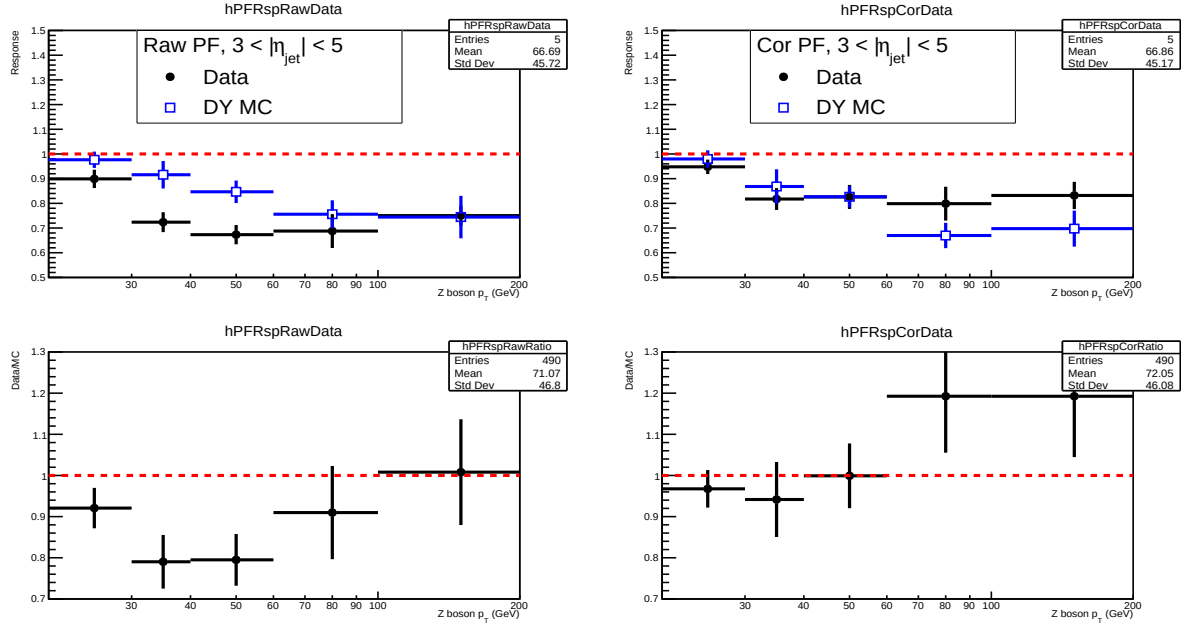


Figure 5.11: HLT PFjet response as a function of reconstructed Z boson p_T using p_T balance approach, comparing data and simulations before (left) and after (right) L1/L2/L3 corrections applied for $3 < |\eta_{jet}| < 5$. The PFjets are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

non-Gaussian shape (also see section 4.3.7). Figure 5.12 provides two examples of jet response profiles in symmetric and asymmetric distributions in two different spatial regions using simulations, after full

L1 correction determination (Random Cone Method)	Events
(1) /SingleNeutrino/RunIISpring15DR74-HFscaleFlat10to30Asympt25ns_MCRUN2	9847776
(2) /SingleNeutrino/RunIIFall15DR76-25nsFlat0to50ZsecalNzshcalRaw_76X_mcRun2_asymptotic_2016EcalTune_30fb	19803954
L2L3 correction determination and full JEC calibration	
/QCD_Pt-15to3000_TuneCUETP8M1_Flat_13TeV_pythia8/RunIIFall15DR76-25nsFlat0to50ZsecalNzshcalRaw_76X_mcRun2_asymptotic_2016EcalTune_30fb	9870818

Table 5.1: MC sample datasets and the corresponding number of events are used for determining and calibrating the jet energy corrections, which were applied for the data taking at 13 TeV in 2016. The bunch spacing was 25 ns. The first sample was simulated with flat distribution of pileup amount from 10 to 30, while the last two samples were produced later with the extended ranges of pileup amount from 0 to 50 according to the various expected pileup conditions.

jet energy corrections are applied. In our analysis, we fit each distribution using Gaussian functions only around the peak position, in order to estimate the average response value of each distribution corresponding to the maximal value of the fitted profile. In addition, the root-mean-square (RMS) value of each fitted curve is treated as the jet p_T resolution in this region.

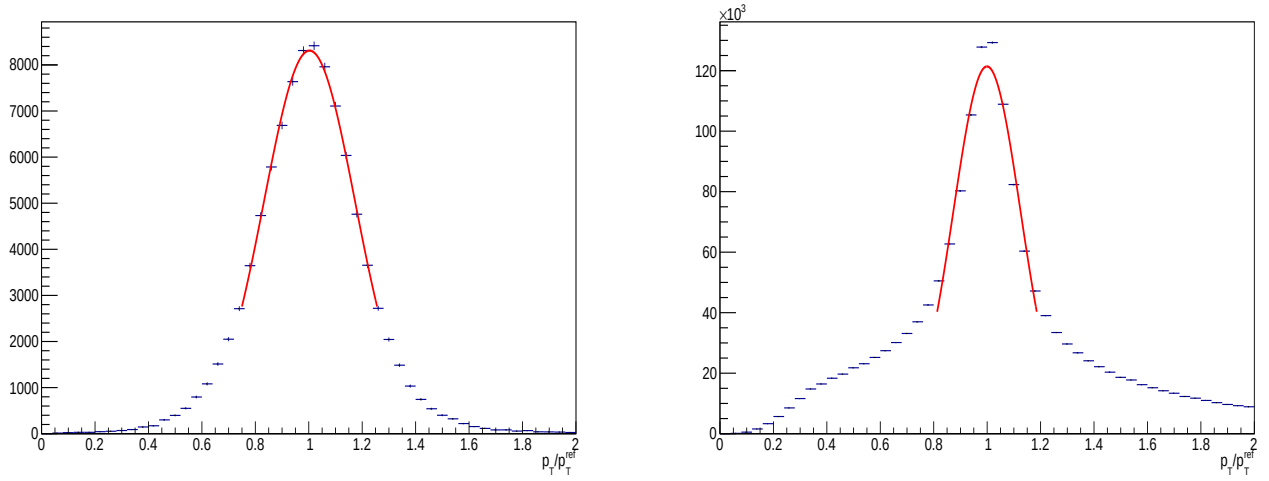


Figure 5.12: The profiles of PFjet response in Gaussian (left) or non-Gaussian (right) distributions with full jet energy corrections applied. The blue points are the real distributions, and the red lines stand for the fitted curves with Gaussian approximation around the maximal values. The p_T and p_T^{ref} are the reconstructed jet p_T and the generated jet p_T in particle level, respectively. The PFjets are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

Since there are two sets of MC samples for determining the L1 correction, we have two groups of calibration results. The jet response calibration results (closure) are illustrated from figure 5.13 to figure 5.19. In general, the PFjets have much better responses than Calojets, because they are implemented with more effective pileup subtraction and have better consistence between the differential correction factor distributions and the fitted curves of L2L3 corrections, compared with Calojets. The two categories of jets have similar response distributions at high jet p_T region, which are close to unity. Through comparing the response of the same kind of reconstructed jets applied with

the L1 corrections determined from the two different samples respectively, one can see that sample (1) is more suitable for Calojets with less fluctuation on response distribution after corrections, while sample (2) is more reliable for PFjets.

The jet transverse momentum resolutions with the jet energy correction applied are shown from figure 5.20 to figure 5.24. Apparently, the resolution of PFjets (about 2%) is better than that of Calojets (about 2-3%), which is because that PFjets are reconstructed with both tracks and calorimetry, while Calojets are reconstructed with only calorimetry. The resolutions of Calojets and PFjets at high p_T regions are approaching to each other due to the decreased tracking efficiency for PFjets. The barrel region shows the best jet resolution, while the resolutions in endcap and HF regions give some fluctuations due to the inactive gaps and the absence of efficient tracking.

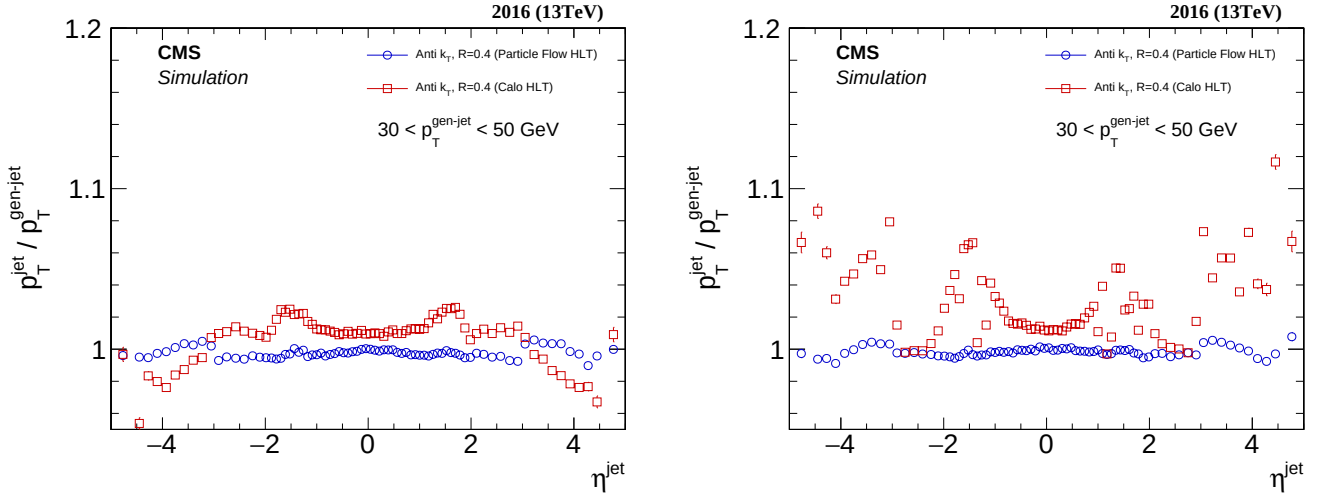


Figure 5.13: The compared responses between Calojets and PFjets at HLT as functions of reconstructed jet η in the region of generated jet p_T from 30 GeV to 50 GeV, with full jet energy corrections applied. The L1 correction is derived from sample (1) (left) and from sample (2) (right) listed in table 5.1. The jets are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

As is seen in section 5.5.1, the hybrid jet area method improved significantly the offset energy subtraction related to the random cone scenario ⁵, especially at low jet transverse momenta scope. Hence, this method is adopted for the L1 correction at HLT in 2017. At this point, the improved jet energy corrections are needed to be determined and calibrated at simulation level. The whole procedures are similar to the previous ones for 2016 data taking, except for the step of determining L1 correction. The MC samples for this round are listed in table 5.2. They are generated from the same MC generator at particle level, while being simulated with different pileup conditions. The first sample is simulated with pileup varying from 0 to 70 yielding a flat distribution, and the second one does not include pileup interaction. Moreover, the first one is also used to estimate L2L3 corrections and to implement the calibration of jet response and energy resolution in simulations.

⁵The random cone method can be also used to calibrate the offset p_T versus η distribution, similar to figure 5.4, in order to validate the simulated pileup amount.

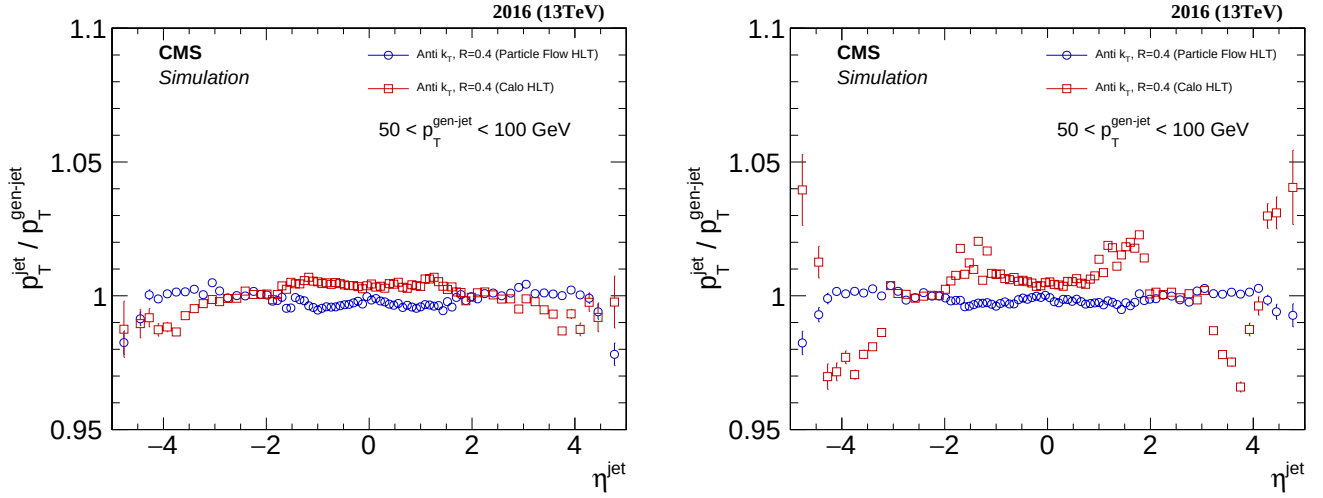


Figure 5.14: The compared responses between Calojets and PFjets at HLT as functions of reconstructed jet η in the region of generated jet p_T from 50 GeV to 100 GeV, with full jet energy corrections applied. The L1 correction is derived from sample (1) (left) and from sample (2) (right) listed in table 5.1. The jets are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

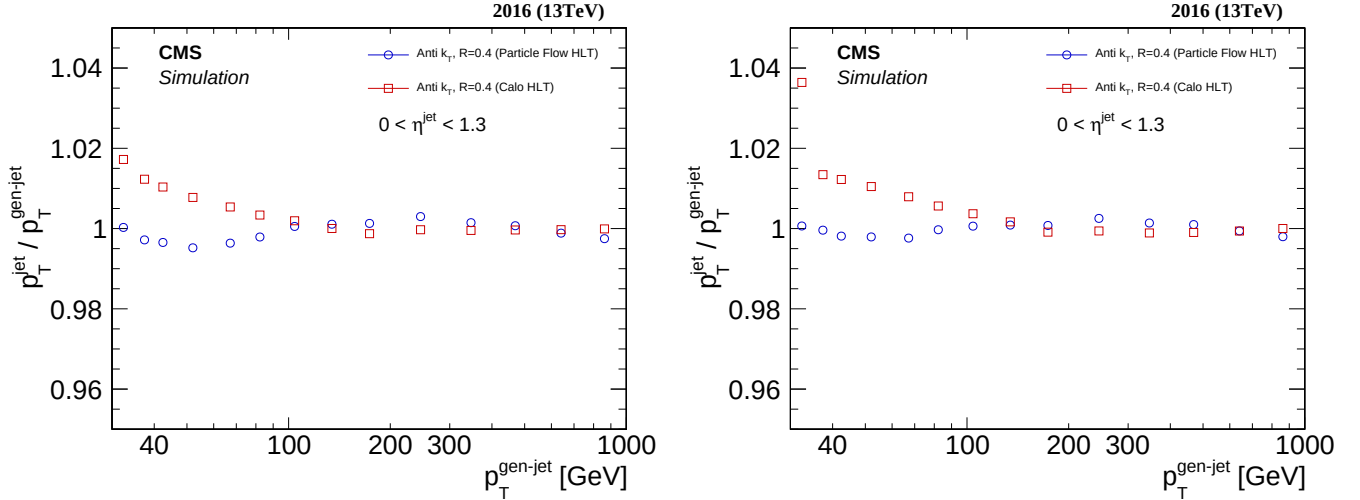


Figure 5.15: The compared responses between Calojets and PFjets at HLT as functions of generated jet p_T in the barrel region, with full jet energy corrections applied. The L1 correction is derived from sample (1) (left) and from sample (2) (right) listed in table 5.1. The jets are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

In terms of the calibration this time, we extend the range of generated jet p_T down to 10 GeV for both Calojets and PFjets, in order to take into account the correction at low reconstructed p_T region. Figure 5.25 shows the jet response as a function of reconstructed jet η in different jet p_T regions. The distributions become flatter and flatter when the jet p_T increase. The responses of Calojets have the discrepancies of approximate 15% dropping to about 1%, comparing with the fluctuations of PFjets responses reduced from 10% to 1% as well. The inactive gap between endcap and HF ($|\eta| = 3$) gives

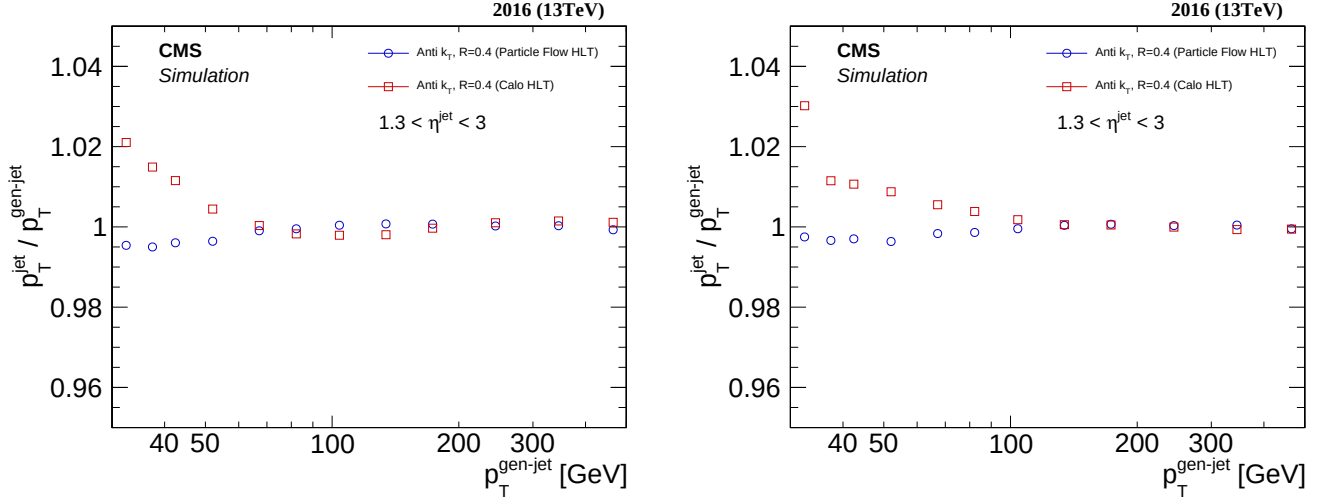


Figure 5.16: The compared responses between Calojets and PFjets at HLT as functions of generated jet p_T in the endcap region, with full jet energy corrections applied. The L1 correction is derived from sample (1) (left) and from sample (2) (right) listed in table 5.1. The jets are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

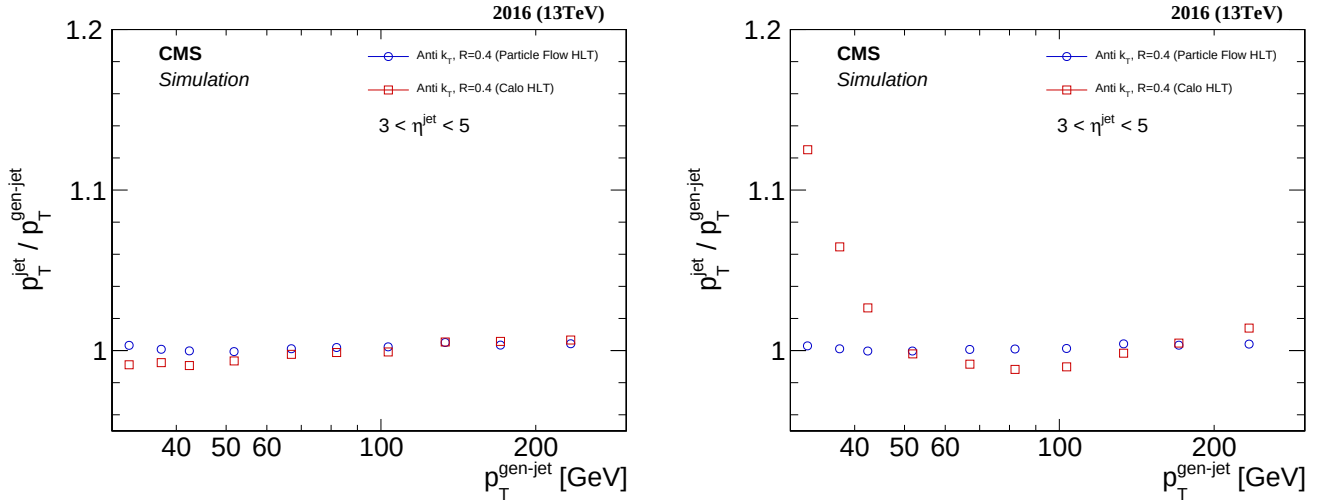


Figure 5.17: The compared responses between Calojets and PFjets at HLT as functions of generated jet p_T in the HF region, with full jet energy corrections applied. The L1 correction is derived from sample (1) (left) and from sample (2) (right) listed in table 5.1. The jets are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

rise to a bump in response.

Figure 5.26 presents the jet response as a function of generated jet p_T in different longitudinal angular directions. The responses show about 3-6% discrepancies below 20 GeV for PFjets, while the Calojets have 10-20% difference in the same region.

The reason why the corrected jet transverse momenta versus jet p_T and η in some regions still have some discrepancies is that the raw jet responses have large fluctuations in those regions, which bias

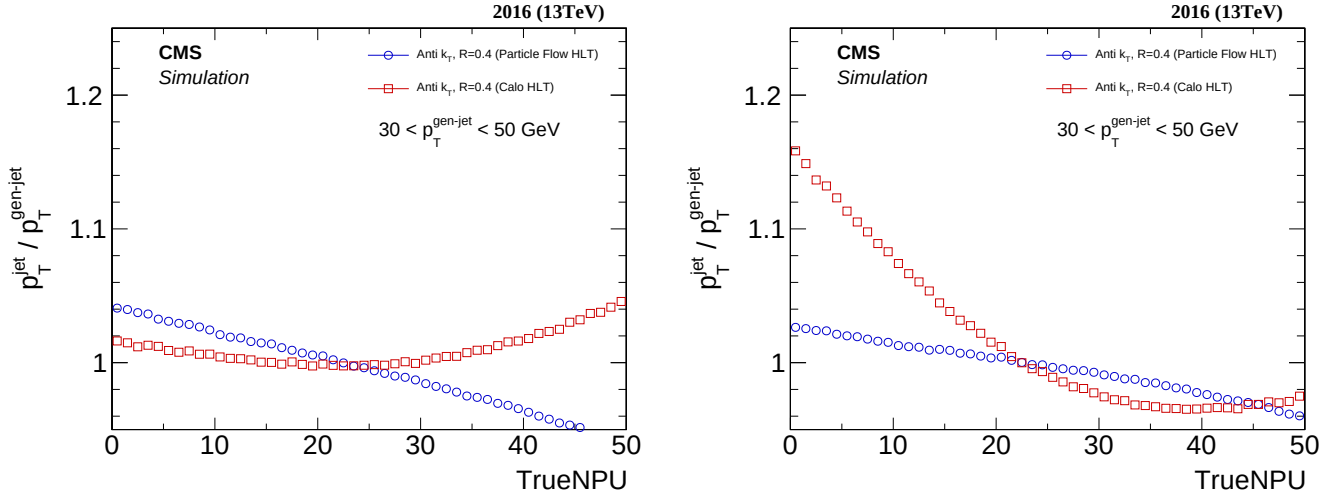


Figure 5.18: The compared responses between Calojets and PFjets at HLT as functions of true number of pileup in different generated jet p_T regions, with full jet energy corrections applied. The L1 correction is derived from sample (1) (left) and from sample (2) (right) listed in table 5.1. The jets are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

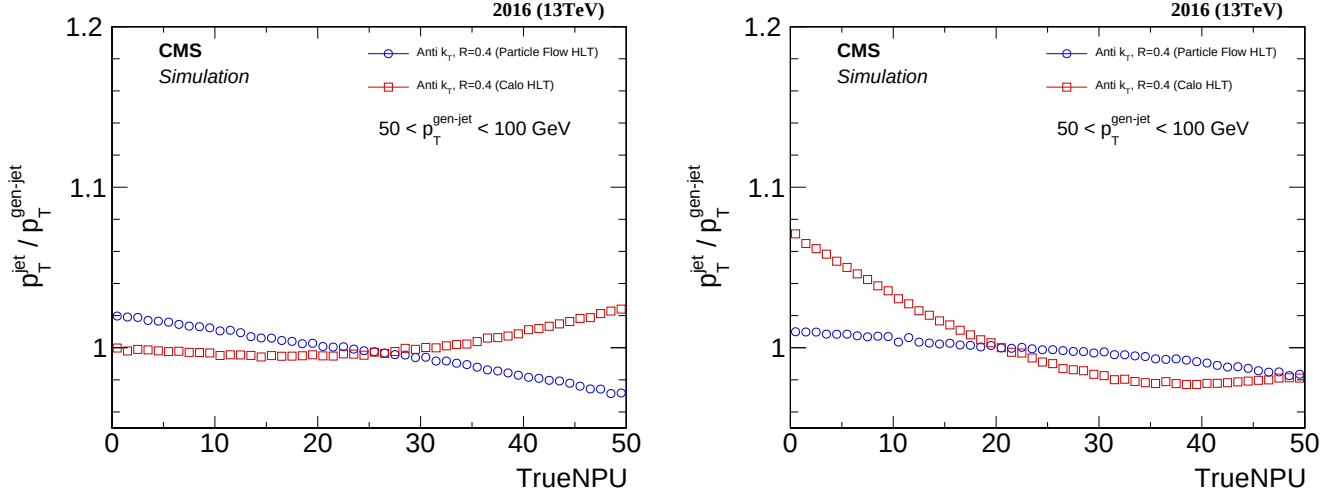


Figure 5.19: The compared responses between Calojets and PFjets at HLT as functions of true number of pileup in different generated jet p_T regions, with full jet energy corrections applied. The L1 correction is derived from sample (1) (left) and from sample (2) (right) listed in table 5.1. The jets are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

the parametrizations of the estimated L2L3 corrections. It is affected by the pileup energy subtraction and the statistics in simulations. Therefore, it is a challenge to optimize the parametrization scenario of L2L3 corrections, which is still under investigation.

Figure 5.27 shows the jet response as a function of true number of pileup in different jet p_T ranges. The flatness of PFjets response (varying within 5%) is much better than that of Calojets response (varying from 5% to 15%). Since the tracking technique is applied in the reconstruction of PFjets at

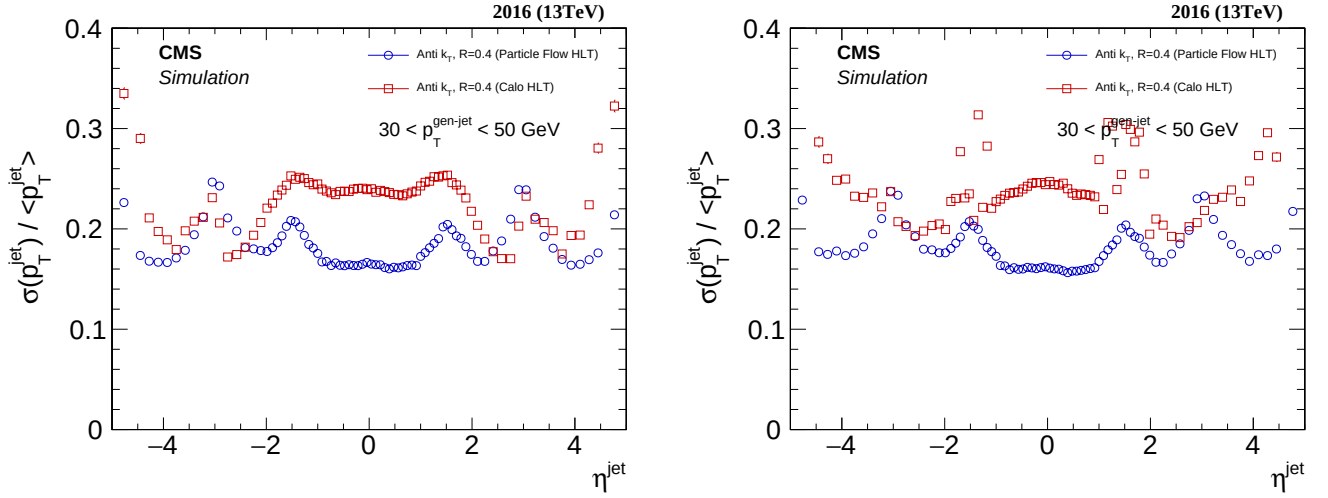


Figure 5.20: The compared p_T resolutions between Calojets and PFjets at HLT as functions of reconstructed jet η in the region of generated jet p_T from 30 GeV to 50 GeV, with full jet energy corrections applied. The L1 correction is derived from sample (1) (left) and from sample (2) (right) listed in table 5.1. The jets are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

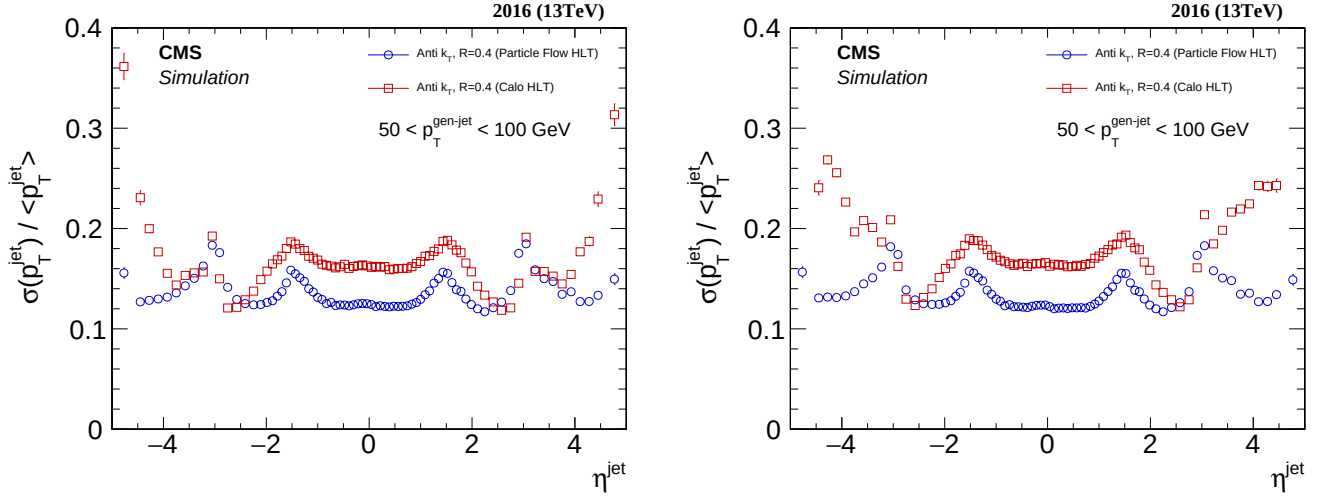


Figure 5.21: The compared p_T resolutions between Calojets and PFjets at HLT as functions of reconstructed jet η in the region of generated jet p_T from 50 GeV to 100 GeV, with full jet energy corrections applied. The L1 correction is derived from sample (1) (left) and from sample (2) (right) listed in table 5.1. The jets are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

HLT, the pileup energy subtraction is much more efficient compared with the case of Calojets. When the jet p_T get larger, the PFjet responses approach to the Calojet responses due to the reduction of pileup contamination and decreased tracking efficiency.

Figure 5.28 displays the jet p_T resolution versus jet η in different jet p_T regions. The PFjets p_T resolutions vary approximately from 30% to 15% when the jet p_T increase, compared with the

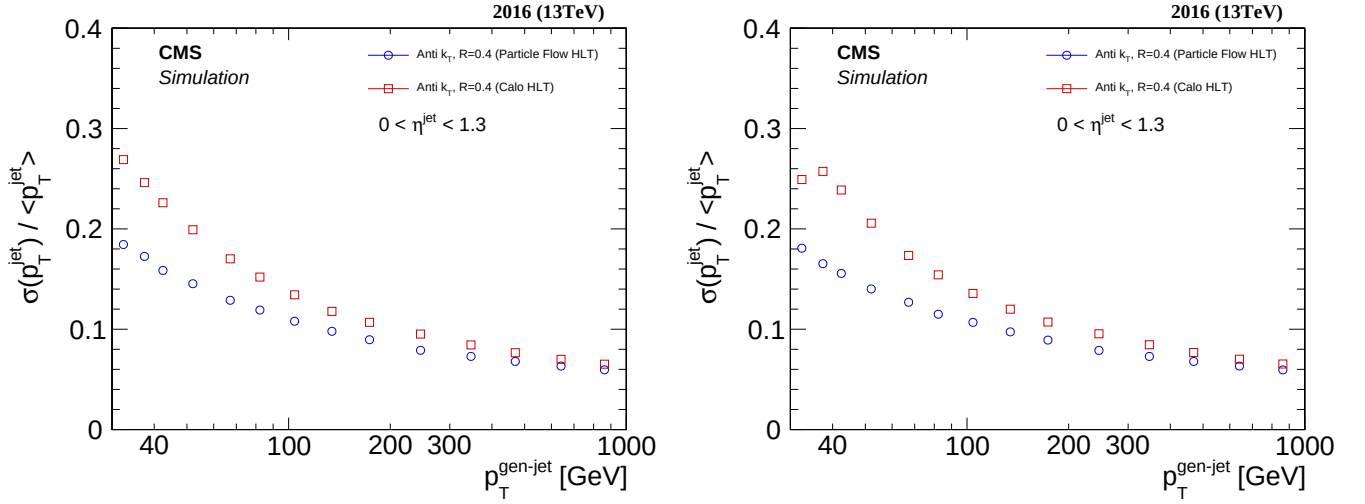


Figure 5.22: The compared p_T resolutions between Calojets and PFjets at HLT as functions of generated jet p_T in the barrel region, with full jet energy corrections applied. The L1 correction is derived from sample (1) (left) and from sample (2) (right) listed in table 5.1. The jets are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

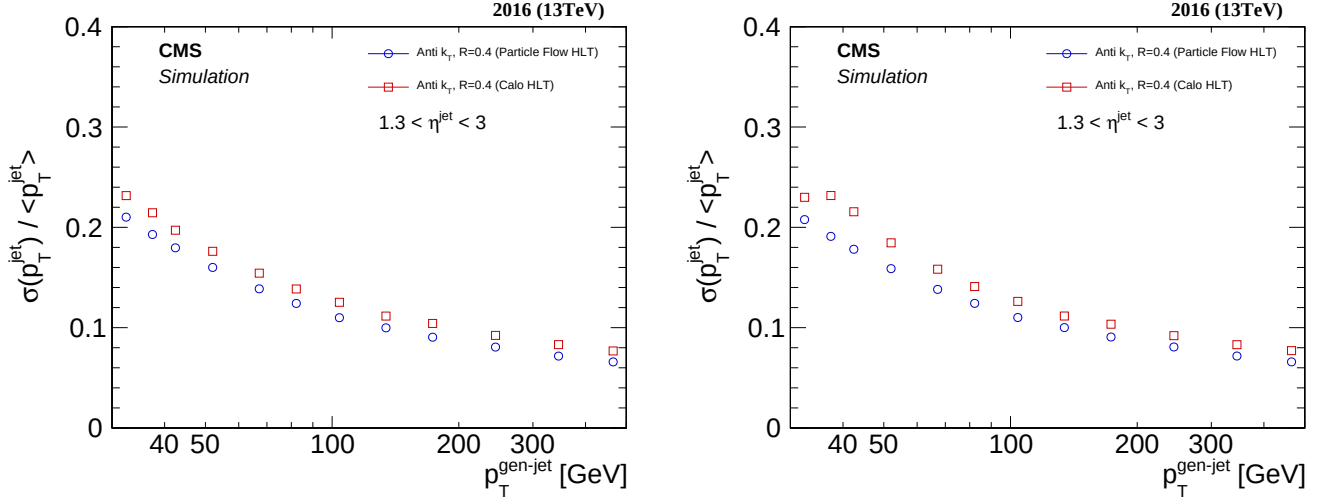


Figure 5.23: The compared p_T resolutions between Calojets and PFjets at HLT as functions of generated jet p_T in the endcap region, with full jet energy corrections applied. The L1 correction is derived from sample (1) (left) and from sample (2) (right) listed in table 5.1. The jets are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

p_T resolutions of Calojets changing from 50% to 20%. The off-uniformity results from different reconstruction efficiencies and subtracted pileup quantities along different longitudinal directions.

The jet p_T resolution with respect to generated jet p_T in barrel, endcap, and HF regions, respectively. In general, one can see again the PFjets perform better than Calojets. The significantly increased relative resolution at low p_T region is mainly due to the pileup energy subtraction, which suppresses the average p_T of the reconstructed jet clusters.

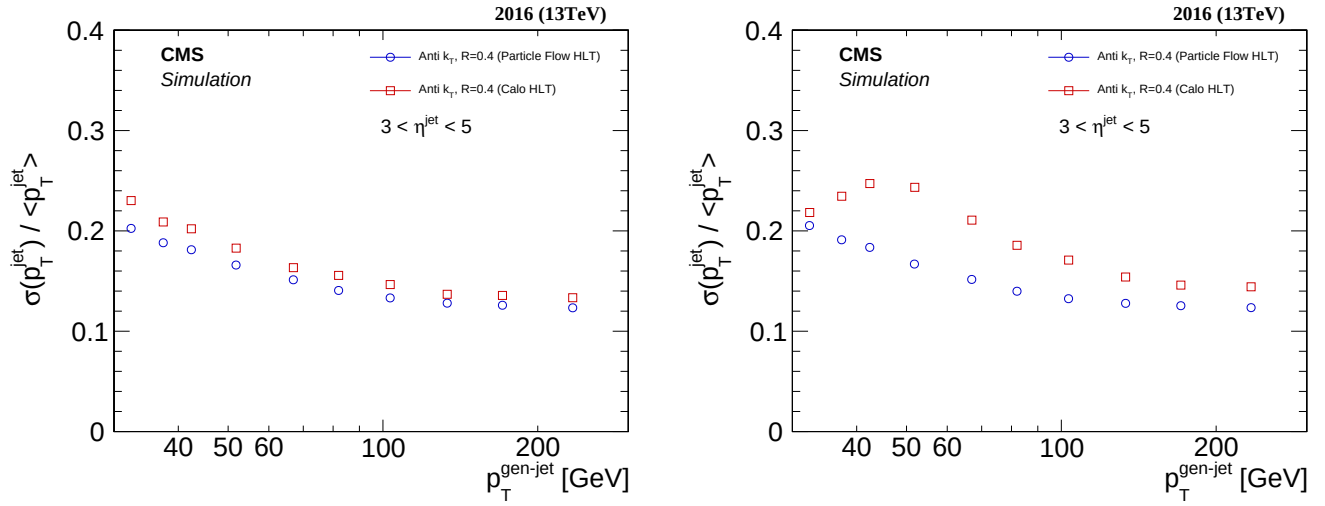


Figure 5.24: The compared p_T resolutions between Calojets and PFjets at HLT as functions of generated jet p_T in the HF region, with full jet energy corrections applied. The L1 correction is derived from sample (1) (left) and from sample (2) (right) listed in table 5.1. The jets are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

L1 correction determination (Hybrid Jet Area Method)	Events
/QCD_Pt-15to3000_TuneCUETP8M1_Flat_13TeV_pythia8/RunIISummer16DR80-PUFlat0to70HcalNZSRAWAODSIM_80X_mcRun2_asymptotic_2016_TracheIV/	9923654
/QCD_Pt-15to3000_TuneCUETP8M1_Flat_13TeV_pythia8/RunIISummer16DR80-NoPUHcalNZSRAWAODSIM_80X_mcRun2_asymptotic_2016_TracheIV/	9942392
L2L3 correction determination and full JEC calibration	
/QCD_Pt-15to3000_TuneCUETP8M1_Flat_13TeV_pythia8/RunIISummer16DR80-PUFlat0to70HcalNZSRAWAODSIM_80X_mcRun2_asymptotic_2016_TracheIV/	9923654

Table 5.2: MC sample datasets and the corresponding number of events are used for determining and calibrating the jet energy corrections, which will be applied for the data taking at 13 TeV in 2017. The bunch spacing will be 25 ns. The first sample is simulated with flat distribution of pileup amount from 0 to 70, while the second sample is produced without pileup at the same time.

Finally, it has to be noted that the technique of jet energy correction at HLT is still developing, and it is a long-term project.

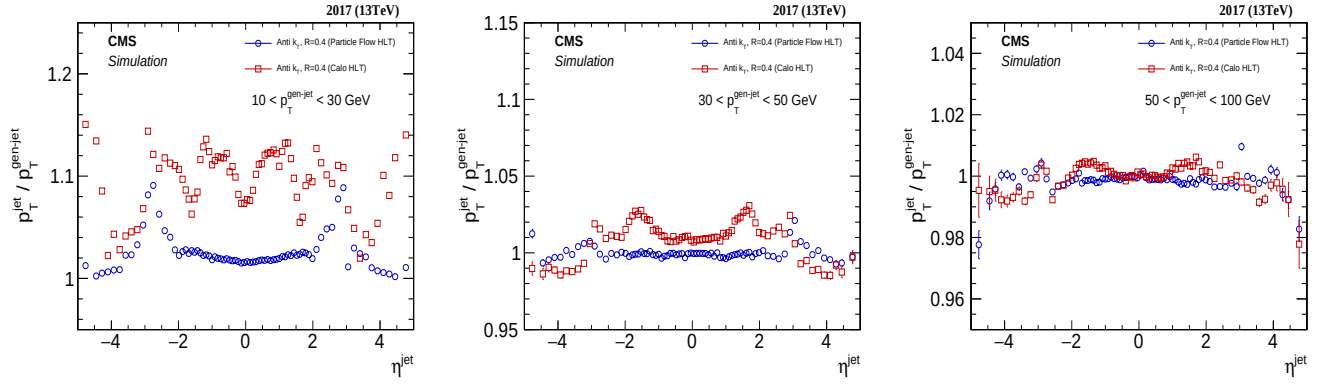


Figure 5.25: The compared jet responses between Calojets and PFjets at HLT as functions of reconstructed jet η in different generated jet p_T ranges, with full jet energy corrections applied. The jets are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

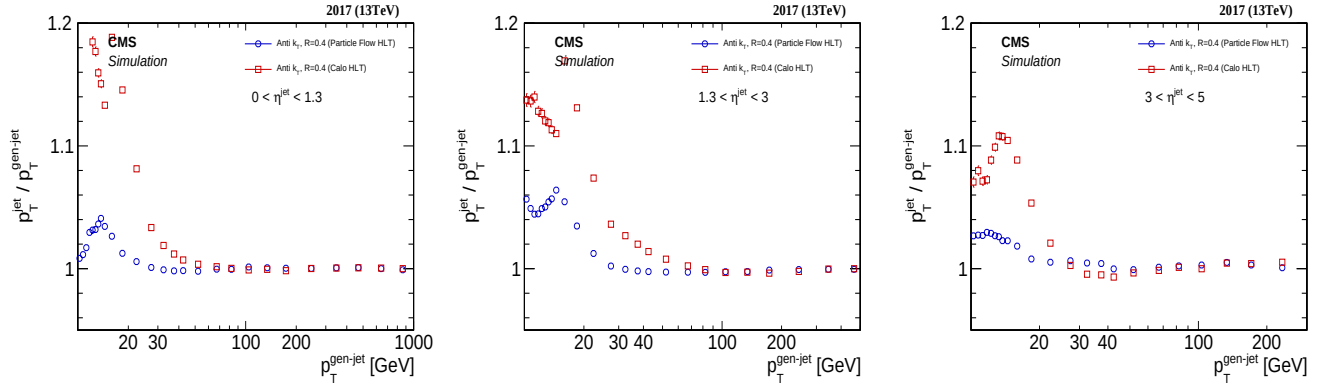


Figure 5.26: The compared jet responses between Calojets and PFjets at HLT as functions of generated jet p_T in different reconstructed jet η ranges, with full jet energy corrections applied. The jets are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

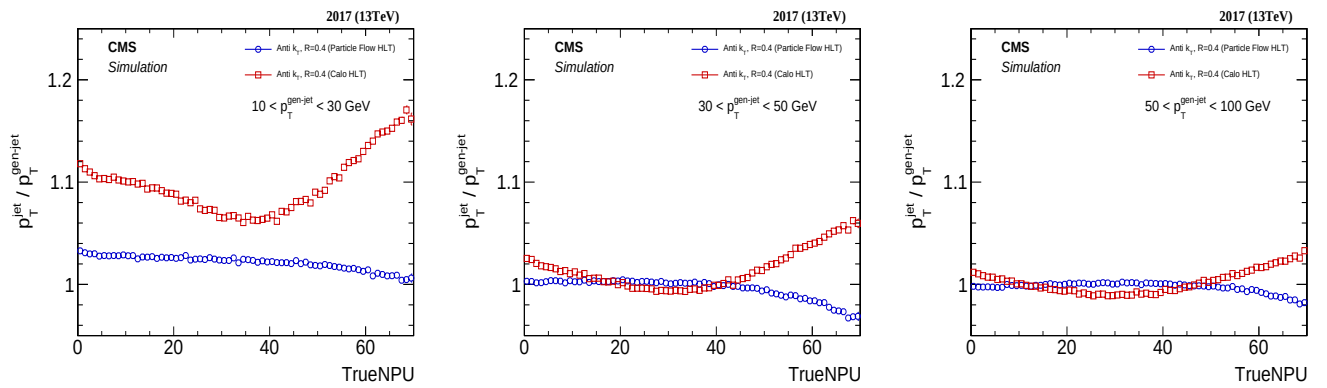


Figure 5.27: The compared jet responses between Calojets and PFjets at HLT as functions of true number of pileup in different generated jet p_T ranges, with full jet energy corrections applied. The jets are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

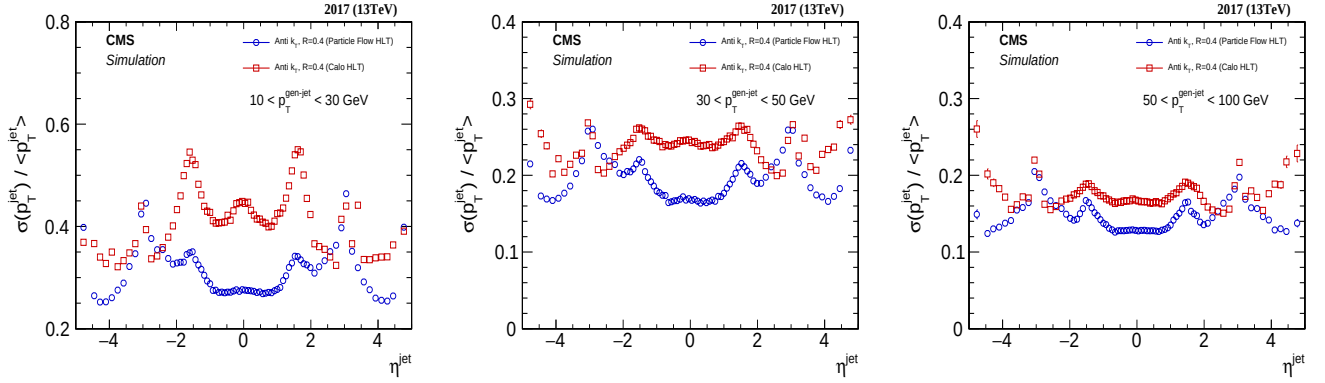


Figure 5.28: The compared jet p_T resolutions between Calojets and PFjets at HLT as functions of reconstructed jet η in different generated jet p_T ranges, with full jet energy corrections applied. The jets are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

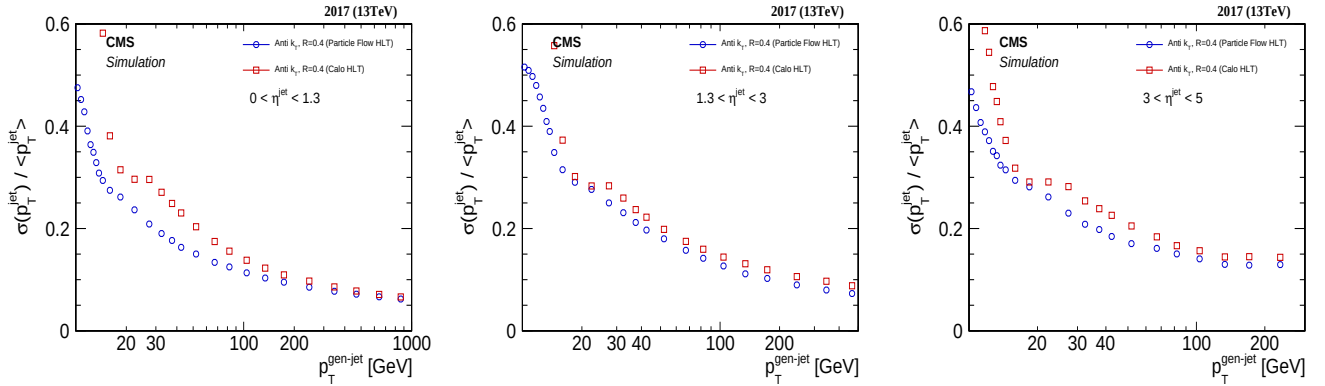


Figure 5.29: The compared jet p_T resolutions between Calojets and PFjets at HLT as functions of generated jet p_T in different reconstructed jet η ranges, with full jet energy corrections applied. The jets are reconstructed with the anti- k_T algorithm (distance parameter $R = 0.4$).

Chapter 6

Conclusion

The Standard Model was established based on quantum field mechanism in the second half of 20th century, and got improved in accuracy during many decades. Thanks to the continuous efforts of physicists on high energy experiments, including precision measurements and extensive searches, the visible elementary particles in the nature together with their interaction mechanisms are completely depicted in the theory. The Higgs boson has been discovered at the Large Hadron Collider in 2012, which has complemented the solidity of the model.

The Large Hadron Collider occupies the highest energy that human beings has made so far in particle interactions, and it offers us an unprecedented sight of the microscopic world. Using this major accelerator, we have realized precision measurements of many SM predicted processes in a phase space never reached before. This machine also makes possible to probe new physics beyond the SM. Benefiting from high integrated luminosity of the collided beams, the LHC provides huge amount of statistics for our physical analyses. As one of the general purpose detectors at the LHC, the Compact Muon Solenoid has confirmed the existence of the Higgs boson.

The study on jet production at the LHC is important for our accurate understanding of perturbative QCD theory based on the SM. Due to the large amount of data collected by the CMS experiment, we can reach significant statistics with large jet multiplicity production, which requires high accuracy level of theoretical prediction. In order to gain a clean selection for the jet study, we require the produced jets to be associated with a Z boson, which is easily detectable and efficiently reconstructed. This process has high cross section and almost free of background contamination, which is an ideal laboratory for jet production study. This thesis presents the measurement of the differential cross section of this process in several observables, including jet multiplicity, jet transverse momentum, rapidity, and scalar sum of jet transverse momenta. Several theoretical predictions from different Monte Carlo generators are available for their comparisons with the measurement, so that these predictions could be confronted to the observations. At this point, the measurement of jet multiplicity distribution provides an effective constraint on the validity of theoretical predictions, which perform ME calculations at different accuracy level matched with parton shower. Jet transverse momentum is interesting to be measured, because its modelling is sensitive to the matching scheme between ME

and parton shower, and to PDF as well. The scalar sum of jet transverse momenta gives a hint to supersymmetry searches with multiple hadronic jets production and large missing transverse energy. This measurement also provides an accurate estimation of the background from $Z + \text{jets}$ process for SUSY searches.

The rapidity correlation between Z boson and jets is also measured, because it makes the limitation of multi-leg leading order ME approach visible, which helps us to understand QCD processes much more accurately. Thanks to the large enough statistics of data, we manage to measure the correlation observable distributions precisely. There are two uncorrelated variables, $y_{sum}(Z, jet)$ and $y_{diff}(Z, jet)$. The former is related to a boost from the laboratory frame to the center-of-mass frame of Z boson and jet, which relies on the PDFs. The latter is the relative rapidity between Z boson and jet, which reflects the partonic differential cross section.

The measurements have been performed using two sets of data collected by the CMS detector with proton-proton collisions at the LHC in 2012 and 2015, respectively. They correspond to an integrated luminosity of 19.6 fb^{-1} and 2.25 fb^{-1} , with a center-of-mass energy of 8 TeV and 13 TeV, respectively. The first compared theoretical prediction for 8 TeV analysis is from MADGRAPH 5 with LO multi-leg ME calculation up to four final state partons interfaced with PYTHIA 6 for parton shower and hadronization. The second one for 8 TeV is from SHERPA 2 with NLO multi-leg ME computation plus one virtual loop correction up to two partons in the final state, implementing parton shower and hadronization with its own model. The third one for 8 TeV is from MADGRAPH_AMC@NLO with NLO multi-leg approach up to two partons in the final state, and interfaced with PYTHIA8 for parton shower and hadronization. In the case of 13 TeV analysis, the predictions of MADGRAPH 5 and MADGRAPH_AMC@NLO are available, and they are both interfaced with PYTHIA8. Moreover, a prediction of NNLO ME computation based on N-jettiness scheme is provided for the leading jet kinematic observables.

In conclusion, this analysis work has accomplished the precision measurement of $Z + \text{jets}$ process, and provides the precise systematics of jet production for different theoretical models (including the implementations of PDF, ME calculation, parton shower, hadronization). The improvement of higher order (NLO/NNLO) of perturbative QCD ME computations relative to the LO multi-leg ME approach has been quantified and is significant for jet transverse momentum distributions. A discrepancy arises at large jet multiplicity when the predictions are only relying on parton shower. An important disagreement is observed in the rapidity correlation between Z boson and jets, while comparing the data to the MADGRAPH 5 LO multi-leg approach with MLM matching scheme in the distributions of $y_{sum}(Z, j_1)$ and $y_{diff}(Z, j_1)$. On the other hand, the NLO predictions with new matching schemes (FxFx and MEPS@NLO) reach good agreement. This observation confirms the fact that although the LO multi-leg approach includes higher order diagrams with a gluon propagator yielding more forward rapidities, the corresponding parton shower matching in MLM approach is failing. The present measurement provides a precise validation of the improvements of the new matching schemes for NLO predictions. These conclusions based on 8 TeV measurement are (preliminarily)

confirmed by 13 TeV measurement with lower statistical significance.

All the results of 8 TeV analysis have been incorporated in a major CMS publication for RunI. For the 13 TeV analysis with a larger center-of-mass energy, the results in $Z \rightarrow \mu^+\mu^-$ channel have been approved by the collaboration and shown in conferences. The combined results of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ channels are foreseen to be summarized as a physical article for publication in future. Since there were already more statistics obtained during 2016, and the integrated luminosity will continue increasing significantly through the data taking in future, the measurement of this process at 13 TeV will gain more precision and allow more differential studies. The statistical gain will allow to study regions of lower cross section, like the large Z boson and jet p_T and large Z boson mass. In large p_T region, electroweak corrections are expected to become significant; in large Z boson mass domain, the effect of the increased scale on many soft gluon radiations has to be investigated.

In 8 TeV analysis, I have contributed to the investigation and application of unfolding approach, and estimated the theoretical and PDF uncertainties, and the combination of two leptonic decay channels. I am the main investigator of the rapidity correlation between Z boson and jets. In 13 TeV measurement, I am the main analyzer for the di-muon channel, and have contributed to the channel combination. In 2015, I gave the physical pre-approval talk for this analysis in the collaboration.

The instantaneous luminosity during RunII is much higher in 2016 than the one was in RunI period. What is more, the bunch spacing during CMS RunII is reduced from 50 ns to 25 ns. These upgrades give rise to a challenge to the jet reconstruction at high level trigger system in the CMS detector, because there are much more extra interactions (pileup) contributing with their energies to the reconstructed jet energies. These effects influence the efficiency of jet trigger system and the trigger rate as well. Therefore, a jet energy correction methodology has been developed. The second part of this thesis has presented the jet energy correction approach and the calibration procedures.

There are three levels of corrections for reconstructed jet transverse momenta, which are pileup energy correction, relative (η) correction, and absolute (p_T) correction. The first level is to subtract the pileup energy contamination from the reconstructed jet p_T . The second step is to make the distribution of jet response versus jet pseudorapidity uniform (overcome η dependence). The third one is to correct the jet response as a function of jet p_T to unity, which means to revise the jet p_T to the value at produced level. The corrections are determined by use of simulated samples, and parametrized as differential scaling factors with respect to pileup amount, jet η and p_T at the three levels, respectively. Through the calibration step, parametrized residual corrections for data are performed according to their differences with MC simulations. As my contribution in this project, the performed procedures described in this thesis have been adopted by the CMS collaboration for RunII, and the determined optimal corrections were configured and used for the data taking in 2015 and 2016. New corrections will be also applied for the data taking in future.

In summary, this thesis establishes a connection between RunI and RunII for the precision measurement of $Z + \text{jets}$ process, focusing on the jet production study. At this point, jet performance at HLT for RunII upgrade is also important for the subsequent data analyses involving jet production.

Bibliography

- [1] Donald H. Perkins. *Introduction to High Energy Physics*. Cambridge University Press, 4 edition, 2000.
- [2] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory*. Frontiers in Physics. Avalon Publishing, 1995.
- [3] R.K. Ellis, W.J. Stirling, and B.R. Webber. *QCD and Collider Physics*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 2003.
- [4] Vardan Khachatryan et al. Measurement and QCD analysis of double-differential inclusive jet cross-sections in pp collisions at $\sqrt{s} = 8$ TeV and ratios to 2.76 and 7 TeV. *JHEP*, 03:156, 2017.
- [5] F. D. Aaron et al. Inclusive Deep Inelastic Scattering at High Q^2 with Longitudinally Polarised Lepton Beams at HERA. *JHEP*, 09:061, 2012.
- [6] H. Abramowicz et al. Measurement of high- Q^2 neutral current deep inelastic e^+p scattering cross sections with a longitudinally polarised positron beam at HERA. *Phys. Rev.*, D87(5):052014, 2013.
- [7] V. N. Gribov and L. N. Lipatov. Deep inelastic e p scattering in perturbation theory. *Sov. J. Nucl. Phys.*, 15:438–450, 1972. [*Yad. Fiz.*15,781(1972)].
- [8] Yuri L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics. *Sov. Phys. JETP*, 46:641–653, 1977. [*Zh. Eksp. Teor. Fiz.*73,1216(1977)].
- [9] Guido Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl. Phys.*, B126:298–318, 1977.
- [10] Mikhail A. Shifman, A. I. Vainshtein, and Valentin I. Zakharov. QCD and Resonance Physics. Theoretical Foundations. *Nucl. Phys.*, B147:385–447, 1979.
- [11] Marco Guzzi, Pavel Nadolsky, Edmond Berger, Hung-Liang Lai, Fredrick Olness, and C. P. Yuan. CT10 parton distributions and other developments in the global QCD analysis. *arXiv:1101.0561*, 2011.
- [12] S. D. Drell and Tung-Mow Yan. Massive Lepton Pair Production in Hadron-Hadron Collisions at High-Energies. *Phys. Rev. Lett.*, 25:316–320, 1970. [Erratum: *Phys. Rev. Lett.*25,902(1970)].
- [13] Alexandre Léonard. *Measurement of Z boson production in association with jets at the LHC and study of a DAQ system for the Triple-GEM detector in view of the CMS upgrade*. PhD thesis, Brussels U., 2015-04-16.

- [14] Charalampos Anastasiou, Lance J. Dixon, Kirill Melnikov, and Frank Petriello. High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO. *Phys. Rev.*, D69:094008, 2004.
- [15] Bo Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand. Parton Fragmentation and String Dynamics. *Phys. Rept.*, 97:31–145, 1983.
- [16] Torbjorn Sjostrand. The Lund Monte Carlo for Jet Fragmentation and e^+e^- Physics: Jetset Version 6.2. *Comput. Phys. Commun.*, 39:347–407, 1986.
- [17] B. Webber. Parton shower Monte Carlo event generators. *Scholarpedia*, 6(12):10662, 2011. revision #128236.
- [18] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Anti- $k(t)$ jet clustering algorithm. *JHEP*, 04:063, 2008.
- [19] Ryan Atkin. Review of jet reconstruction algorithms. *Journal of Physics: Conference Series*, 645(1):012008, 2015.
- [20] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet User Manual. *Eur. Phys. J.*, C72:1896, 2012.
- [21] Stefan Hoche. Introduction to parton-shower event generators. In *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014) Boulder, Colorado, June 2-27, 2014*, 2014.
- [22] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 05:026, 2006.
- [23] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852–867, 2008.
- [24] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. MadGraph 5 : Going Beyond. *JHEP*, 06:128, 2011.
- [25] H. Murayama, I. Watanabe, and Kaoru Hagiwara. HELAS: HELicity amplitude subroutines for Feynman diagram evaluations. *KEK-91-11*, 1992.
- [26] Priscila de Aquino, William Link, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. ALOHA: Automatic Libraries Of Helicity Amplitudes for Feynman Diagram Computations. *Comput. Phys. Commun.*, 183:2254–2263, 2012.
- [27] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.
- [28] Stefano Frixione and Bryan R. Webber. Matching NLO QCD computations and parton shower simulations. *JHEP*, 06:029, 2002.
- [29] Neil D. Christensen and Claude Duhr. FeynRules - Feynman rules made easy. *Comput. Phys. Commun.*, 180:1614–1641, 2009.

- [30] Johan Alwall et al. A Standard format for Les Houches event files. *Comput. Phys. Commun.*, 176:300–304, 2007.
- [31] Rikkert Frederix, Stefano Frixione, Fabio Maltoni, and Tim Stelzer. Automation of next-to-leading order computations in QCD: The FKS subtraction. *JHEP*, 10:003, 2009.
- [32] S. Frixione, Z. Kunszt, and A. Signer. Three jet cross-sections to next-to-leading order. *Nucl. Phys.*, B467:399–442, 1996.
- [33] S. Frixione. A General approach to jet cross-sections in QCD. *Nucl. Phys.*, B507:295–314, 1997.
- [34] T. Gleisberg, Stefan. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter. Event generation with SHERPA 1.1. *JHEP*, 02:007, 2009.
- [35] F. Cascioli, S. Höche, F. Krauss, P. Maierhöfer, S. Pozzorini, and F. Siegert. Automatic one-loop calculations with Sherpa+OpenLoops. *J. Phys. Conf. Ser.*, 523:012058, 2014.
- [36] F. Krauss. Matrix elements and parton showers in hadronic interactions. *JHEP*, 08:015, 2002.
- [37] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber. QCD matrix elements + parton showers. *JHEP*, 11:063, 2001.
- [38] S. Catani, Yuri L. Dokshitzer, M. H. Seymour, and B. R. Webber. Longitudinally invariant K_t clustering algorithms for hadron hadron collisions. *Nucl. Phys.*, B406:187–224, 1993.
- [39] Johan Alwall et al. Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions. *Eur. Phys. J.*, C53:473–500, 2008.
- [40] Rikkert Frederix and Stefano Frixione. Merging meets matching in MC@NLO. *JHEP*, 12:061, 2012.
- [41] Stefan Hoeche, Frank Krauss, Marek Schonherr, and Frank Siegert. QCD matrix elements + parton showers: The NLO case. *JHEP*, 04:027, 2013.
- [42] Lyndon Evans and Philip Bryant. LHC Machine. *Journal of Instrumentation*, 3(08):S08001, 2008.
- [43] S. Chatrchyan et al. The CMS Experiment at the CERN LHC. *JINST*, 3:S08004, 2008.
- [44] LHC Programme Coordination web pages. <http://lpc.web.cern.ch>.
- [45] The CERN Hadron Ion Sources. <https://linac2.web.cern.ch/linac2/sources/source.htm>.
- [46] Serguei Chatrchyan et al. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *JINST*, 9(10):P10009, 2014.
- [47] S Chatrchyan et al. Commissioning and Performance of the CMS Pixel Tracker with Cosmic Ray Muons. *JINST*, 5:T03007, 2010.

- [48] D. Sprenger, M. Weber, R. Adolphi, R. Brauer, L. Feld, K. Klein, A. Ostapchuk, S. Schael, and B. Wittmer. Validation of Kalman Filter alignment algorithm with cosmic-ray data using a CMS silicon strip tracker endcap. *JINST*, 5:P06007, 2010.
- [49] G. Bayatian et al. Design, performance and calibration of the CMS forward calorimeter wedges. *Eur. Phys. J.*, C53:139–166, 2008.
- [50] G. Baiatian et al. Energy Response and Longitudinal Shower Profiles Measured in CMS HCAL and Comparison With Geant4. *CERN-CMS-NOTE-2006-143*, 2007.
- [51] Min Suk Kim. CMS reconstruction improvement for the muon tracking by the RPC chambers. *PoS*, RPC2012:045, 2012. [JINST8,T03001(2013)].
- [52] CMS Collaboration. *Data Acquisition and High-Level Trigger*. <http://cmsdoc.cern.ch/cms/TDR/DAQ/daq.html>.
- [53] Vardan Khachatryan et al. The CMS trigger system. *JINST*, 12(01):P01020, 2017.
- [54] Andrea Perrotta. Performance of the CMS High Level Trigger. *Journal of Physics: Conference Series*, 664(8):082044, 2015.
- [55] CMS Public. *CMSSW Application Framework*. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookCMSSWFramework>.
- [56] CERN. *ROOT User’s Guide*. <https://root.cern.ch/guides/users-guide>.
- [57] CERN. *Geant4*. <https://geant4.web.cern.ch/geant4/>.
- [58] CMS Collaboration. Measurement of the differential production cross section of Z bosons in association with jets in pp collisions at $\sqrt{s} = 8$ TeV. *JHEP*, 2017(4):22, 2017.
- [59] A. Grebenyuk et al. Differential cross section measurements for the production of a Z boson in association with jets in proton-proton collisions at $\sqrt{s} = 13$ TeV. CMS Analysis Note AN-2015-219 (CERN), 2015.
- [60] CMS Collaboration. Measurement of the differential cross section of Z boson production in association with jets in proton-proton collisions at $\sqrt{s} = 13$ TeV. *CMS-PAS-SMP-15-010*, 2015.
- [61] Serguei Chatrchyan et al. Rapidity distributions in exclusive Z + jet and γ + jet events in pp collisions at $\sqrt{s} = 7$ TeV. *Phys. Rev.*, D88(11):112009, 2013.
- [62] Pavel Nadolsky, Jun Gao, Marco Guzzi, Joey Huston, Hung-Liang Lai, Zhao Li, Jon Pumplin, Dan Stump, and C. P. Yuan. Progress in CTEQ-TEA PDF Analysis. In *Proceedings, 20th International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS 2012): Bonn, Germany, March 26-30, 2012*, pages 417–420, 2012. [417(2012)].
- [63] Vardan Khachatryan et al. Charged particle multiplicities in pp interactions at $\sqrt{s} = 0.9, 2.36$, and 7 TeV. *JHEP*, 01:079, 2011.
- [64] Ryan Gavin, Ye Li, Frank Petriello, and 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.

- [65] Nikolaos Kidonakis. Differential and total cross sections for top pair and single top production. In *Proceedings, 20th International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS 2012): Bonn, Germany, March 26-30, 2012*, pages 831–834, 2012. [,831(2012)].
- [66] John M. Campbell and R. K. Ellis. MCFM for the Tevatron and the LHC. *Nucl. Phys. Proc. Suppl.*, 205-206:10–15, 2010.
- [67] Z. Was. TAUOLA the library for tau lepton decay, and KKMC / KORALB / KORALZ /... status report. *Nucl. Phys. Proc. Suppl.*, 98:96–102, 2001. [,96(2000)].
- [68] Richard D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015.
- [69] Vardan Khachatryan et al. Event generator tunes obtained from underlying event and multi-parton scattering measurements. *Eur. Phys. J.*, C76(3):155, 2016.
- [70] Paolo Nason. A New method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004.
- [71] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 11:070, 2007.
- [72] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010.
- [73] CMS Public. Tag and Probe. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/TagAndProbe>.
- [74] CMS Internal. Muon HLT efficiency. <https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs>.
- [75] P. Billoir and S. Qian. Simultaneous pattern recognition and track fitting by the kalman filtering method. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 294(1):219 – 228, 1990.
- [76] CMS Collaboration. Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. *CMS-PAS-PFT-09-001*, 2009.
- [77] W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov. Reconstruction of Electron Tracks With the Gaussian-Sum Filter. *CERN-CMS-RN-2003-001*, 2003.
- [78] Serguei Chatrchyan et al. Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV. *JINST*, 7:P10002, 2012.
- [79] CMS Public. *Baseline muon selections for Run-I*. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonId>.
- [80] CMS Internal. *Baseline muon selections for Run-II*. <https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2>.
- [81] CMS Collaboration. Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV. *CMS-PAS-PFT-10-002*, 2010.

- [82] J. Han and A. Bodek. *Rochester Muon Momentum Correction*. https://www-cdf.fnal.gov/~jyhan/cms_momscl/cms_rochcor_manual.html.
- [83] CMS Internal. Muon ID, Isolation and Trigger Efficiencies with 13 TeV, 25 ns Data. https://indico.cern.ch/event/502421/contributions/2012616/attachments/1234575/1811863/Muon_IDISO_Efficiencies.pdf.
- [84] Wolfgang Adam, R. Frühwirth, Are Strandlie, and T. Todor. Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC. *CERN-CMS-NOTE-2005-001*, 2005.
- [85] Andreas Hocker et al. TMVA - Toolkit for Multivariate Data Analysis. *PoS, ACAT:040*, 2007.
- [86] Jerome H. Friedman. Greedy function approximation: A gradient boosting machine. *Ann. Statist.*, 29(5):1189–1232, 10 2001.
- [87] Vardan Khachatryan et al. Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV. *JINST*, 10(06):P06005, 2015.
- [88] Vardan Khachatryan et al. Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV. *JINST*, 10(06):P06005, 2015.
- [89] Serguei Chatrchyan et al. Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS. *JINST*, 6:P11002, 2011.
- [90] Vardan Khachatryan et al. Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV. *JINST*, 12(02):P02014, 2017.
- [91] CMS Internal. *Recommended Jet Energy Corrections and Uncertainties For Data and MC*. <https://twiki.cern.ch/twiki/bin/view/CMS/JECDataMC>.
- [92] CMS Internal. *Jet Energy Corrections: reference sample for jet energy corrections and jet resolution*. <https://twiki.cern.ch/twiki/bin/view/CMS/JERCReference>.
- [93] CMS Internal. *Jet Energy Resolution*. <https://twiki.cern.ch/twiki/bin/view/CMS/JetResolution>.
- [94] CMS Collaboration. Pileup Jet Identification. *CMS-PAS-JME-13-005*, 2013.
- [95] CMS Public. *Pile-Up Events (PileUpProducer or MixingModule)*. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideFastSimPileUp>.
- [96] Tim Adye. Unfolding algorithms and tests using RooUnfold. In *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN, Geneva, Switzerland 17-20 January 2011*, pages 313–318, Geneva, 2011. CERN, CERN.
- [97] Andreas Hocker and Vakhtang Kartvelishvili. SVD approach to data unfolding. *Nucl. Instrum. Meth.*, A372:469–481, 1996.
- [98] G. D’Agostini. A Multidimensional unfolding method based on Bayes’ theorem. *Nucl. Instrum. Meth.*, A362:487–498, 1995.

- [99] CMS Public. *Jet Energy Corrections: Official Software Tools for applying JEC Corrections and Uncertainties*. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookJetEnergyCorrections>.
- [100] CMS Collaboration. CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update. *CMS-PAS-LUM-13-001*, 2013.
- [101] CMS Collaboration. CMS Luminosity Measurement for the 2015 Data Taking Period. *CMS-PAS-LUM-15-001*, 2016.
- [102] Michał Czakon, Paul Fiedler, and Alexander Mitov. Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_s^4)$. *Phys. Rev. Lett.*, 110:252004, 2013.
- [103] Richard Nisius. *a software package to combine correlated estimates of physics observables within ROOT using the Best Linear Unbiased Estimate method*. <https://blue.hepforge.org/Bluemanual.pdf>.
- [104] Radja Boughezal, Xiaohui Liu, and Frank Petriello. Phenomenology of the Z-boson plus jet process at NNLO. *Phys. Rev.*, D94(7):074015, 2016.
- [105] Richard D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015.
- [106] Juan Rojo et al. The PDF4LHC report on PDFs and LHC data: Results from Run I and preparation for Run II. *J. Phys.*, G42:103103, 2015.
- [107] Radja Boughezal, John M. Campbell, R. Keith Ellis, Christfried Focke, Walter T. Giele, Xiaohui Liu, and Frank Petriello. Z-boson production in association with a jet at next-to-next-to-leading order in perturbative QCD. *Phys. Rev. Lett.*, 116(15):152001, 2016.
- [108] Andy Buckley, Jonathan Butterworth, Leif Lonnblad, David Grellscheid, Hendrik Hoeth, James Monk, Holger Schulz, and Frank Siegert. Rivet user manual. *Comput. Phys. Commun.*, 184:2803–2819, 2013.
- [109] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, Pavel M. Nadolsky, and W. K. Tung. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 07:012, 2002.
- [110] C. F. Berger, Z. Bern, Lance J. Dixon, F. Febres Cordero, D. Forde, H. Ita, D. A. Kosower, and D. Maitre. One-Loop Calculations with BlackHat. *Nucl. Phys. Proc. Suppl.*, 183:313–319, 2008.
- [111] C. F. Berger, Z. Bern, Lance J. Dixon, F. Febres Cordero, D. Forde, T. Gleisberg, H. Ita, D. A. Kosower, and D. Maitre. Vector Boson + Jets with BlackHat and Sherpa. *Nucl. Phys. Proc. Suppl.*, 205-206:92–97, 2010.
- [112] Johan Alwall, Simon de Visscher, and Fabio Maltoni. QCD radiation in the production of heavy colored particles at the LHC. *JHEP*, 02:017, 2009.
- [113] A. Valassi. Combining correlated measurements of several different physical quantities. *Nucl. Instrum. Meth.*, A500:391–405, 2003.

- [114] Vardan Khachatryan et al. Measurements of jet multiplicity and differential production cross sections of Z +jets events in proton-proton collisions at $\sqrt{s} = 7$ TeV. *Phys. Rev.*, D91(5):052008, 2015.
- [115] Georges Aad et al. Measurement of the production cross section of jets in association with a Z boson in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *JHEP*, 07:032, 2013.
- [116] Michelangelo L. Mangano, Mauro Moretti, Fulvio Piccinini, Roberto Pittau, and Antonio D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 07:001, 2003.
- [117] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber. HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). *JHEP*, 01:010, 2001.
- [118] V. M. Abazov et al. Measurements of differential cross sections of $Z/\gamma^* + \text{jets} + X$ events in proton anti-proton collisions at $s^{**}(1/2) = 1.96$ -TeV. *Phys. Lett.*, B678:45–54, 2009.
- [119] Timo Antero Aaltonen et al. Measurement of differential production cross section for Z/γ^* bosons in association with jets in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev.*, D91(1):012002, 2015.
- [120] The ATLAS collaboration. Measurement of the Production Cross Sections of a Z Boson in Association with Jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector. *ATLAS-CONF-2015-041*, 2015.
- [121] CMS Collaboration. Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV. *CMS-PAS-FSQ-15-005*, 2016.
- [122] Andrea Massironi. Precision electromagnetic calorimetry at the energy frontier: CMS ECAL at LHC Run 2. In *Proceedings, Meeting of the APS Division of Particles and Fields (DPF 2015): Ann Arbor, Michigan, USA, 4-8 Aug 2015*, 2015.
- [123] Fengwangdong Zhang. Performance of the CMS Jets and Missing Transverse Energy trigger for the upgraded LHC. *PoS*, EPS-HEP2015:290, 2015.
- [124] E. Chabanat and N. Estre. Deterministic annealing for vertex finding at CMS. In *Computing in high energy physics and nuclear physics. Proceedings, Conference, CHEP'04, Interlaken, Switzerland, September 27-October 1, 2004*, pages 287–290, 2005.
- [125] CMS Internal. *Calorimeter reconstruction and pileup*. https://indico.cern.ch/event/362962/contributions/859630/attachments/721855/990894/JME_TSG_8Jan15.pdf.
- [126] CMS Public. *JetResponseAnalyzer*. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideJetResponseAnalyzer>.
- [127] Alexx Perloff. JetMETAnalysis/JetResponseAnalyzer software package: A how-to guide for checking out, compiling, and running the package. CMS Analysis Note AN-13-131 (CERN), 2013.

Appendix A

Tables of combined results at 8 TeV

This chapter provides a summary of the differential cross sections as a function of the measured observables in Z+jets analysis at 8 TeV, together with the systematic and statistical uncertainties.

N_{jets}	$\frac{d\sigma}{dN_{jets}}$ [pb]	Tot [%]	stat [%]	Unf stat [%]	JES [%]	JER [%]	PU [%]	Bgnd [%]	Lumi [%]	LES [%]	LER [%]	Unf sys [%]	Eff [%]
= 0	423.	3.7	0.034	0.050	1.2	0.060	0.71	0.0024	2.6	0.065	0.024	1.2	1.8
= 1	59.9	6.3	0.11	0.075	5.3	0.23	0.26	0.042	2.6	0.085	0.041	1.4	1.8
= 2	12.6	9.2	0.25	0.12	8.4	0.22	0.35	0.33	2.7	0.082	0.039	1.7	1.9
= 3	2.46	12.	0.60	0.22	11.	0.22	0.42	0.76	2.9	0.10	0.034	2.7	2.0
= 4	0.471	16.	1.4	0.43	15.	0.16	0.57	1.3	3.1	0.083	0.024	3.5	2.1
= 5	0.0901	20.	3.4	1.0	19.	0.28	0.75	1.9	3.2	0.076	0.072	4.6	2.3
= 6	0.0143	33.	9.3	2.4	28.	0.72	1.9	3.3	3.7	0.17	0.27	5.5	2.6
= 7	0.00230	34.	22.	6.3	23.	0.61	5.6	4.3	3.9	0.013	0.10	6.4	2.8

Table A.1: Differential cross section in Exclusive jet multiplicity and break down of the systematic uncertainties for the combination of both decay channels.

N_{jets}	$\frac{d\sigma}{dN_{jets}}$ [pb]	Tot [%]	stat [%]	Unf stat [%]	JES [%]	JER [%]	PU [%]	Bgnd [%]	Lumi [%]	LES [%]	LER [%]	Unf sys [%]	Eff [%]
≥ 0	499.	3.1	0.032	0.044	1.3	0.058	0.60	0.011	2.2	0.056	0.021	1.0	1.6
≥ 1	75.5	5.0	0.097	0.063	4.4	0.19	0.22	0.069	2.1	0.069	0.033	1.2	1.5
≥ 2	15.7	7.4	0.23	0.10	7.0	0.18	0.29	0.29	2.3	0.068	0.032	1.4	1.6
≥ 3	3.04	9.9	0.55	0.19	9.5	0.18	0.35	0.65	2.4	0.082	0.028	2.2	1.7
≥ 4	0.577	13.	1.3	0.39	12.	0.14	0.49	1.1	2.6	0.069	0.023	2.9	1.8
≥ 5	0.107	17.	3.2	0.93	17.	0.25	0.69	1.7	2.8	0.068	0.071	4.2	1.9
≥ 6	0.0166	29.	8.6	2.2	25.	0.63	1.8	2.9	3.2	0.14	0.24	5.5	2.3
≥ 7	0.00230	34.	22.	6.3	23.	0.61	5.6	4.3	3.9	0.013	0.10	6.4	2.8

Table A.2: Differential cross section in Inclusive jet multiplicity and break down of the systematic uncertainties for the combination of both decay channels.

p_T^Z [GeV]	$\frac{d\sigma}{dp_T^Z}$ [$\frac{pb}{GeV}$]	Tot [%]	stat [%]	Unf stat [%]	JES [%]	JER [%]	PU [%]	Bgnd [%]	Lumi [%]	LES [%]	LER [%]	Unf sys [%]	Eff [%]
0 – 10	0.219	31.	1.2	0.57	30.	1.9	5.1	0.029	2.6	1.6	0.18	1.2	1.8
10 – 20	0.493	20.	0.60	0.24	20.	1.9	1.4	0.064	2.6	1.1	0.048	1.3	1.8
20 – 30	0.885	16.	0.39	0.16	15.	1.2	0.74	0.064	2.6	1.3	0.063	2.1	1.8
30 – 40	1.32	8.4	0.26	0.13	7.5	0.42	0.67	0.066	2.6	0.45	0.062	1.9	1.8
40 – 50	1.25	5.0	0.24	0.14	3.5	0.081	0.71	0.084	2.6	0.40	0.020	0.98	1.8
50 – 60	0.957	3.9	0.26	0.15	1.8	0.056	0.68	0.10	2.6	0.39	0.056	1.0	1.8
60 – 70	0.682	3.6	0.31	0.17	1.0	0.039	0.62	0.14	2.7	0.61	0.021	0.92	1.8
70 – 80	0.476	3.7	0.37	0.20	0.61	0.021	0.53	0.16	2.7	0.82	0.046	1.4	1.9
80 – 90	0.330	3.8	0.45	0.24	0.38	0.018	0.54	0.19	2.7	0.89	0.032	1.5	1.9
90 – 100	0.230	3.8	0.54	0.27	0.25	0.0042	0.47	0.22	2.7	0.95	0.0079	1.5	1.9
100 – 110	0.165	4.0	0.63	0.32	0.17	0.0055	0.45	0.22	2.7	1.1	0.030	1.7	1.9
110 – 120	0.124	3.9	0.73	0.37	0.15	0.0068	0.47	0.20	2.7	1.0	0.048	1.6	1.9
120 – 130	0.0910	4.4	0.86	0.42	0.11	0.0050	0.35	0.21	2.7	1.1	0.048	2.5	1.9
130 – 140	0.0688	4.2	1.0	0.49	0.084	0.0037	0.39	0.16	2.7	1.3	0.050	1.9	1.9
140 – 150	0.0522	4.1	1.1	0.52	0.10	0.0038	0.33	0.17	2.7	1.0	0.024	1.7	1.9
150 – 160	0.0407	4.1	1.3	0.61	0.050	0.0025	0.32	0.16	2.7	1.1	0.057	1.7	1.9
160 – 170	0.0334	4.5	1.4	0.68	0.073	0.0043	0.25	0.14	2.7	1.3	0.12	2.2	1.9
170 – 180	0.0242	4.1	1.6	0.73	0.068	0.0040	0.32	0.13	2.7	1.3	0.070	1.1	1.9
180 – 190	0.0198	4.4	1.8	0.78	0.041	0.012	0.25	0.14	2.7	1.1	0.064	1.7	1.9
190 – 200	0.0154	5.0	2.0	0.89	0.079	0.0086	0.35	0.13	2.7	1.1	0.086	2.8	1.9
200 – 220	0.0114	4.4	1.6	0.69	0.053	0.0060	0.14	0.12	2.7	1.3	0.028	1.8	1.9
220 – 240	0.00756	4.8	2.0	0.85	0.034	0.0014	0.24	0.13	2.7	1.8	0.036	2.1	1.9
240 – 260	0.00510	4.5	2.4	1.0	0.062	0.0034	0.36	0.11	2.7	1.3	0.047	1.0	1.9
260 – 280	0.00360	5.0	2.9	1.2	0.078	0.0012	0.29	0.15	2.7	1.1	0.054	1.8	1.9
280 – 300	0.00236	5.8	3.5	1.4	0.021	0.0045	0.28	0.15	2.7	2.1	0.078	2.2	1.9
300 – 320	0.00177	5.9	4.0	1.6	0.11	0.00027	0.45	0.12	2.7	1.4	0.080	1.6	1.9
320 – 360	0.00102	5.9	3.7	1.4	0.099	0.00069	0.40	0.15	2.7	2.0	0.17	1.9	1.9
360 – 400	0.000560	6.8	5.1	1.9	0.0058	0.00016	0.25	0.16	2.7	1.2	0.18	2.1	1.9
400 – 450	0.000294	8.1	6.4	2.4	0.015	0.00076	0.41	0.17	2.8	1.9	0.19	2.1	1.9
450 – 500	0.000150	11.	8.9	3.3	0.033	0.0016	0.57	0.19	2.8	2.5	0.31	3.6	2.0
500 – 590	7.19e-05	11.	9.1	3.6	0.034	0.0016	0.37	0.18	2.8	2.1	0.21	4.5	2.0
590 – 700	1.88e-05	22.	19.	6.6	0.039	0.0018	0.23	0.18	2.8	3.4	0.25	8.5	2.0
700 – 1000	2.78e-06	49.	41.	14.	0.0017	8.1e-05	0.97	0.22	2.8	3.7	0.41	23.	2.3

Table A.3: Differential cross section in p_T^Z and break down of the systematic uncertainties for the combination of both decay channels.

$p_T(j_1)$ [GeV]	$\frac{d\sigma}{dp_T(j_1)} \frac{pb}{GeV}$	Tot [%]	stat [%]	Unf stat [%]	JES [%]	JER [%]	PU [%]	Bgnd [%]	Lumi [%]	LES [%]	LER [%]	Unf sys [%]	Eff [%]
30 – 39	2.59	5.4	0.18	0.11	4.0	0.37	0.52	0.033	2.6	0.079	0.036	1.6	1.8
39 – 49	1.60	5.1	0.19	0.11	3.6	0.37	0.54	0.061	2.6	0.076	0.039	1.5	1.8
49 – 60	1.02	5.9	0.22	0.13	4.6	0.31	0.49	0.10	2.7	0.095	0.046	1.8	1.9
60 – 72	0.643	6.3	0.26	0.14	4.9	0.25	0.53	0.17	2.7	0.10	0.047	2.0	1.9
72 – 85	0.405	5.9	0.31	0.17	4.4	0.30	0.49	0.23	2.7	0.099	0.050	2.1	1.9
85 – 100	0.250	6.3	0.38	0.19	4.6	0.27	0.50	0.29	2.7	0.10	0.049	2.7	1.9
100 – 117	0.153	6.8	0.45	0.22	5.1	0.30	0.50	0.34	2.7	0.090	0.039	3.0	1.9
117 – 136	0.0920	6.6	0.55	0.26	4.8	0.28	0.43	0.38	2.7	0.079	0.042	3.0	1.9
136 – 157	0.0550	6.3	0.71	0.33	4.7	0.27	0.55	0.37	2.7	0.045	0.042	2.4	2.0
157 – 187	0.0301	6.9	0.83	0.36	5.0	0.23	0.37	0.34	2.7	0.044	0.035	3.2	1.9
187 – 220	0.0157	7.0	1.2	0.47	4.9	0.27	0.41	0.38	2.8	0.042	0.032	3.4	1.9
220 – 258	0.00795	6.2	1.5	0.60	4.0	0.29	0.35	0.40	2.8	0.067	0.0099	2.9	2.0
258 – 300	0.00400	6.8	2.0	0.76	5.0	0.19	0.28	0.35	2.8	0.084	0.017	2.2	1.9
300 – 350	0.00196	8.0	2.6	0.97	4.9	0.36	0.45	0.43	2.8	0.072	0.011	4.5	1.9
350 – 400	0.000924	8.6	4.1	1.4	5.2	0.26	0.71	0.47	2.8	0.083	0.029	3.9	2.1
400 – 450	0.000462	7.5	5.3	1.8	3.5	0.13	0.26	0.54	2.9	0.15	0.022	0.37	2.0
450 – 500	0.000279	11.	6.9	2.4	6.4	0.39	0.20	0.44	2.8	0.15	0.024	4.7	2.0
500 – 590	0.000129	10.	8.2	2.7	3.4	0.53	0.42	0.58	2.9	0.11	0.042	1.6	2.0
590 – 700	4.99e-05	16.	13.	4.4	5.1	0.52	0.55	0.62	2.9	0.23	0.092	5.3	2.0
700 – 1000	1.09e-05	22.	18.	6.7	6.1	0.51	0.39	0.36	2.8	0.21	0.0023	6.6	2.4

Table A.4: Differential cross section in 1st jet p_T ($N_{jets} \geq 1$) and break down of the systematic uncertainties for the combination of both decay channels.

$p_T(j_2)$ [GeV]	$\frac{d\sigma}{dp_T(j_2)} \frac{pb}{GeV}$	Tot [%]	stat [%]	Unf stat [%]	JES [%]	JER [%]	PU [%]	Bgnd [%]	Lumi [%]	LES [%]	LER [%]	Unf sys [%]	Eff [%]
30 – 39	0.686	7.1	0.35	0.15	5.9	0.35	0.41	0.25	2.7	0.092	0.043	2.1	1.9
39 – 49	0.346	6.7	0.43	0.17	5.4	0.38	0.59	0.39	2.8	0.097	0.042	1.8	1.9
49 – 60	0.188	7.8	0.54	0.21	6.6	0.27	0.47	0.55	2.8	0.086	0.036	2.1	2.0
60 – 72	0.105	8.1	0.67	0.24	6.7	0.15	0.44	0.68	2.8	0.066	0.036	2.7	2.0
72 – 85	0.0614	7.5	0.86	0.30	5.8	0.13	0.55	0.77	2.9	0.077	0.034	3.1	2.0
85 – 100	0.0360	6.9	1.1	0.37	5.1	0.11	0.67	0.85	2.9	0.034	0.039	2.4	2.0
100 – 117	0.0214	7.0	1.3	0.44	5.5	0.14	0.55	0.81	2.9	0.10	0.032	2.1	2.0
117 – 136	0.0125	8.1	1.6	0.52	6.2	0.16	0.63	0.76	2.9	0.17	0.028	3.4	2.0
136 – 157	0.00709	7.6	2.0	0.65	5.6	0.15	0.44	0.71	2.9	0.071	0.032	3.0	2.0
157 – 187	0.00392	6.8	2.5	0.80	4.6	0.16	0.47	0.56	2.8	0.016	0.032	2.5	2.0
187 – 220	0.00191	7.4	3.5	1.1	4.5	0.047	0.54	0.39	2.8	0.044	0.024	3.0	2.0
220 – 258	0.00111	6.4	4.0	1.2	3.4	0.10	0.26	0.39	2.8	0.070	0.031	0.31	2.0
258 – 300	0.000536	9.7	5.5	1.7	6.5	0.20	0.38	0.30	2.7	0.12	0.098	2.6	1.9
300 – 350	0.000240	12.	7.6	2.3	7.8	0.61	0.35	0.34	2.8	0.086	0.11	3.8	1.9
350 – 400	0.000133	11.	9.9	3.0	1.7	0.66	0.48	0.33	2.8	0.022	0.034	2.6	2.0
400 – 450	7.07e-05	17.	14.	4.2	4.8	0.68	1.0	0.53	2.8	0.040	0.063	5.5	2.0
450 – 500	4.14e-05	22.	17.	5.4	11.	0.74	2.0	0.13	2.7	0.17	0.21	6.8	1.9
500 – 590	1.83e-05	25.	20.	6.4	8.4	0.64	2.3	0.11	2.7	0.25	0.28	9.6	1.9
590 – 800	5.37e-06	28.	24.	7.9	2.3	0.50	1.2	0.13	2.7	0.045	0.0072	13.	2.0

Table A.5: Differential cross section in 2nd jet p_T ($N_{jets} \geq 2$) and break down of the systematic uncertainties for the combination of both decay channels.

$p_T(j_3)$ [GeV]	$\frac{d\sigma}{dp_T(j_3)}$ $\left[\frac{pb}{GeV}\right]$	Tot [%]	stat [%]	Unf stat [%]	JES [%]	JER [%]	PU [%]	Bgnd [%]	Lumi [%]	LES [%]	LER [%]	Unf sys [%]	Eff [%]
30 – 39	0.161	8.8	0.75	0.25	7.6	0.29	0.76	0.69	2.9	0.099	0.035	2.6	2.0
39 – 49	0.0704	8.4	1.0	0.33	7.2	0.50	0.62	0.93	2.9	0.062	0.028	2.1	2.1
49 – 62	0.0307	10.	1.4	0.40	9.2	0.22	0.52	1.2	3.0	0.074	0.036	2.0	2.1
62 – 78	0.0127	9.3	2.0	0.57	7.6	0.20	0.52	1.5	3.1	0.14	0.048	2.7	2.2
78 – 105	0.00433	10.	2.7	0.75	8.2	0.17	0.56	1.4	3.1	0.20	0.026	3.2	2.1
105 – 142	0.00119	9.8	4.6	1.3	7.4	0.15	0.87	1.3	3.1	0.15	0.053	0.99	2.1
142 – 185	0.000339	13.	7.6	2.2	7.9	0.28	1.1	1.3	3.0	0.052	0.024	6.0	2.1
185 – 235	6.63e-05	25.	21.	5.6	7.5	0.84	2.7	1.3	3.1	3.4	0.11	7.4	2.8
235 – 300	2.33e-05	41.	30.	9.2	16.	1.2	2.0	0.67	3.0	2.0	0.0033	21.	2.8

Table A.6: Differential cross section in 3rd jet p_T ($N_{jets} \geq 3$) and break down of the systematic uncertainties for the combination of both decay channels.

$p_T(j_4)$ [GeV]	$\frac{d\sigma}{dp_T(j_4)}$ $\left[\frac{pb}{GeV}\right]$	Tot [%]	stat [%]	Unf stat [%]	JES [%]	JER [%]	PU [%]	Bgnd [%]	Lumi [%]	LES [%]	LER [%]	Unf sys [%]	Eff [%]
30 – 39	0.0350	12.	1.6	0.47	10.	0.35	0.82	1.3	3.1	0.11	0.017	3.8	2.1
39 – 49	0.0127	11.	2.6	0.74	9.9	0.41	0.64	1.7	3.2	0.027	0.041	2.5	2.2
49 – 62	0.00480	13.	3.7	0.97	12.	0.23	0.67	2.0	3.3	0.038	0.029	2.9	2.3
62 – 78	0.00180	12.	5.6	1.4	9.4	0.30	0.44	1.9	3.2	0.086	0.027	2.3	2.3
78 – 96	0.000554	17.	11.	2.8	10.	0.58	1.0	2.3	3.4	0.14	0.15	4.5	2.5
96 – 150	0.000117	18.	13.	3.1	11.	0.46	0.34	2.2	3.4	0.35	0.052	4.5	2.4

Table A.7: Differential cross section in 4th jet p_T ($N_{jets} \geq 4$) and break down of the systematic uncertainties for the combination of both decay channels.

$p_T(j_5)$ [GeV]	$\frac{d\sigma}{dp_T(j_5)}$ $\left[\frac{pb}{GeV}\right]$	Tot [%]	stat [%]	Unf stat [%]	JES [%]	JER [%]	PU [%]	Bgnd [%]	Lumi [%]	LES [%]	LER [%]	Unf sys [%]	Eff [%]
30 – 39	0.00723	13.	4.0	1.1	11.	0.70	0.43	1.9	3.3	0.077	0.076	4.1	2.3
39 – 49	0.00217	15.	6.5	1.7	12.	0.24	1.3	2.2	3.4	0.13	0.059	3.2	2.3
49 – 62	0.000771	21.	9.9	2.6	17.	0.41	1.2	2.9	3.5	0.24	0.071	5.2	2.4
62 – 100	9.85e-05	28.	18.	4.3	19.	0.24	2.6	4.7	4.0	0.63	0.015	1.3	2.8

Table A.8: Differential cross section in 5th jet p_T ($N_{jets} \geq 5$) and break down of the systematic uncertainties for the combination of both decay channels.

$ y(Z) $	$\frac{d\sigma}{d y(Z) }$ [pb]	Tot [%]	stat [%]	Unf stat [%]	JES [%]	JER [%]	PU [%]	Bgnd [%]	Lumi [%]	LES [%]	LER [%]	Unf sys [%]	Eff [%]
0 – 0.2	42.8	6.7	0.27	0.14	5.7	0.28	0.22	0.14	2.7	0.079	0.045	1.1	1.8
0.2 – 0.4	42.5	6.7	0.27	0.14	5.7	0.29	0.19	0.14	2.7	0.065	0.032	1.4	1.8
0.4 – 0.6	42.4	6.7	0.28	0.14	5.7	0.27	0.27	0.14	2.7	0.050	0.036	1.4	1.8
0.6 – 0.8	42.1	6.7	0.29	0.15	5.6	0.26	0.24	0.12	2.7	0.038	0.059	1.7	1.8
0.8 – 1	41.0	6.8	0.30	0.15	5.6	0.28	0.35	0.12	2.7	0.032	0.047	1.7	1.9
1 – 1.2	38.6	6.9	0.32	0.16	5.7	0.26	0.45	0.10	2.6	0.037	0.056	1.8	1.9
1.2 – 1.4	35.2	7.0	0.36	0.18	5.8	0.28	0.56	0.087	2.6	0.10	0.055	2.1	1.9
1.4 – 1.6	31.1	6.9	0.39	0.19	5.7	0.29	0.50	0.072	2.6	0.21	0.034	1.9	1.9
1.6 – 1.8	25.8	6.7	0.44	0.22	5.6	0.26	0.50	0.061	2.6	0.44	0.024	1.8	1.9
1.8 – 2	19.7	6.5	0.49	0.25	5.3	0.27	0.56	0.046	2.6	0.33	0.0026	1.8	1.9
2 – 2.2	12.0	6.7	0.61	0.32	5.5	0.25	0.68	0.033	2.6	0.16	0.0034	1.9	1.8
2.2 – 2.4	3.89	6.6	1.1	0.58	5.3	0.29	1.0	0.026	2.6	0.023	0.0084	1.8	1.8

Table A.9: Differential cross section in Z boson Absolute Rapidity ($N_{jets} \geq 1$) and break down of the systematic uncertainties for the combination of both decay channels.

$ y(j_1) $	$\frac{d\sigma}{d y(j_1) }$ [pb]	Tot [%]	stat [%]	Unf stat [%]	JES [%]	JER [%]	PU [%]	Bgnd [%]	Lumi [%]	LES [%]	LER [%]	Unf sys [%]	Eff [%]
0 – 0.075	41.8	5.4	0.53	0.27	4.1	0.18	0.31	0.13	2.7	0.084	0.039	1.5	1.9
0.075 – 0.15	42.2	5.5	0.56	0.29	4.0	0.14	0.32	0.14	2.7	0.078	0.048	1.9	1.8
0.15 – 0.225	42.3	5.7	0.58	0.29	4.2	0.14	0.42	0.13	2.7	0.087	0.038	2.0	1.9
0.225 – 0.3	41.8	5.7	0.58	0.29	4.3	0.13	0.26	0.13	2.7	0.10	0.037	1.7	1.9
0.3 – 0.375	41.0	5.9	0.58	0.29	4.5	0.18	0.47	0.13	2.7	0.099	0.048	2.0	1.9
0.375 – 0.45	40.9	5.8	0.59	0.30	4.4	0.15	0.43	0.13	2.7	0.075	0.041	1.7	1.9
0.45 – 0.525	40.6	5.8	0.58	0.30	4.3	0.17	0.38	0.12	2.7	0.097	0.052	1.9	1.9
0.525 – 0.6	40.1	5.7	0.59	0.30	4.4	0.24	0.36	0.12	2.7	0.10	0.028	1.4	1.9
0.6 – 0.675	39.8	6.2	0.59	0.30	4.9	0.21	0.38	0.13	2.7	0.086	0.046	2.0	1.9
0.675 – 0.75	38.8	6.0	0.61	0.31	4.8	0.16	0.43	0.12	2.7	0.10	0.051	1.5	1.9
0.75 – 0.825	38.7	6.3	0.59	0.30	4.7	0.19	0.50	0.12	2.7	0.090	0.037	2.5	1.8
0.825 – 0.9	37.7	5.9	0.63	0.32	4.5	0.15	0.26	0.11	2.7	0.087	0.052	1.6	1.9
0.9 – 0.975	36.1	6.0	0.63	0.32	4.8	0.23	0.48	0.12	2.7	0.080	0.034	1.3	1.8
0.975 – 1.05	36.2	5.9	0.62	0.32	4.3	0.22	0.41	0.12	2.7	0.071	0.039	2.1	1.8
1.05 – 1.125	35.1	5.8	0.64	0.33	4.5	0.22	0.29	0.099	2.6	0.051	0.031	1.6	1.8
1.125 – 1.2	33.7	5.9	0.66	0.34	4.4	0.15	0.20	0.11	2.6	0.089	0.035	2.1	1.8
1.2 – 1.275	32.2	5.9	0.68	0.34	4.4	0.25	0.37	0.11	2.7	0.087	0.040	2.1	1.8
1.275 – 1.35	31.3	5.7	0.69	0.35	4.3	0.24	0.23	0.10	2.6	0.084	0.017	1.9	1.8
1.35 – 1.425	30.2	6.7	0.73	0.36	5.5	0.29	0.28	0.10	2.6	0.096	0.029	2.0	1.9
1.425 – 1.5	28.0	7.6	0.76	0.37	6.6	0.29	0.24	0.088	2.6	0.11	0.040	1.7	1.8
1.5 – 1.575	26.7	7.5	0.78	0.38	6.5	0.33	0.18	0.097	2.6	0.081	0.020	1.7	1.8
1.575 – 1.65	26.7	8.7	0.76	0.38	7.7	0.16	0.25	0.089	2.6	0.055	0.035	2.1	1.8
1.65 – 1.725	25.1	8.6	0.81	0.40	7.7	0.57	0.24	0.082	2.6	0.088	0.051	1.9	1.8
1.725 – 1.8	24.4	8.7	0.81	0.40	7.7	0.62	0.59	0.089	2.6	0.065	0.032	2.3	1.8
1.8 – 1.875	22.6	9.3	0.85	0.42	8.5	0.57	0.70	0.089	2.6	0.095	0.039	1.6	1.8
1.875 – 1.95	21.6	9.6	0.85	0.42	8.7	0.65	0.72	0.075	2.6	0.056	0.032	2.1	1.8
1.95 – 2.025	21.1	11.	0.86	0.44	9.7	0.60	0.51	0.072	2.6	0.096	0.048	2.3	1.8
2.025 – 2.1	20.4	11.	0.87	0.45	11.	0.40	0.53	0.072	2.6	0.063	0.050	2.1	1.8
2.1 – 2.175	19.4	12.	0.91	0.47	11.	0.58	0.43	0.070	2.6	0.063	0.046	2.2	1.8
2.175 – 2.25	18.3	12.	0.96	0.49	12.	0.55	0.34	0.059	2.6	0.077	0.012	1.5	1.8
2.25 – 2.325	16.7	13.	1.0	0.51	12.	0.71	0.14	0.067	2.6	0.071	0.038	1.3	1.8
2.325 – 2.4	15.9	12.	1.3	0.61	11.	0.60	0.42	0.050	2.6	0.042	0.049	1.8	1.8

Table A.10: Differential cross section in $|y(j_1)|$ ($N_{jets} \geq 1$) and break down of the systematic uncertainties for the combination of both decay channels.

$ y(j_2) $	$\frac{d\sigma}{d y(j_2) }$ [pb]	Tot [%]	stat [%]	Unf stat [%]	JES [%]	JER [%]	PU [%]	Bgnd [%]	Lumi [%]	LES [%]	LER [%]	Unf sys [%]	Eff [%]
0 – 0.075	7.89	7.9	1.5	0.50	6.6	0.19	0.15	0.49	2.8	0.070	0.036	2.2	2.0
0.075 – 0.15	8.10	7.9	1.5	0.48	6.7	0.18	0.48	0.53	2.8	0.073	0.024	2.0	1.9
0.15 – 0.225	8.05	8.4	1.5	0.50	6.9	0.27	0.27	0.47	2.8	0.078	0.038	3.0	1.9
0.225 – 0.3	8.10	7.9	1.5	0.52	6.8	0.25	0.32	0.47	2.8	0.052	0.033	1.4	1.9
0.3 – 0.375	7.89	8.0	1.5	0.51	6.9	0.24	0.41	0.42	2.8	0.059	0.027	1.6	1.9
0.375 – 0.45	7.98	8.0	1.4	0.50	6.7	0.23	0.21	0.46	2.8	0.059	0.029	2.3	1.9
0.45 – 0.525	7.76	7.8	1.5	0.51	6.5	0.26	0.32	0.48	2.8	0.15	0.040	2.1	1.9
0.525 – 0.6	7.78	8.4	1.5	0.51	7.3	0.15	0.39	0.46	2.8	0.089	0.041	1.7	1.9
0.6 – 0.675	7.82	8.3	1.5	0.52	7.1	0.12	0.20	0.46	2.8	0.086	0.058	2.0	1.9
0.675 – 0.75	7.62	8.2	1.5	0.52	6.9	0.32	0.20	0.43	2.8	0.061	0.036	2.3	1.9
0.75 – 0.825	7.63	8.1	1.5	0.52	6.9	0.24	0.27	0.45	2.8	0.10	0.042	1.8	1.9
0.825 – 0.9	7.33	7.8	1.5	0.53	6.6	0.25	0.47	0.38	2.7	0.072	0.040	2.0	1.9
0.9 – 0.975	7.41	8.3	1.6	0.56	7.0	0.16	0.44	0.44	2.8	0.13	0.047	2.0	1.9
0.975 – 1.05	7.29	8.3	1.5	0.54	7.1	0.20	0.32	0.41	2.8	0.071	0.039	2.2	1.9
1.05 – 1.125	7.16	7.9	1.6	0.57	6.6	0.13	0.20	0.38	2.8	0.072	0.031	2.2	1.9
1.125 – 1.2	6.71	8.4	1.6	0.56	7.3	0.18	0.48	0.39	2.8	0.11	0.056	1.5	1.9
1.2 – 1.275	6.69	8.6	1.6	0.55	7.3	0.12	0.13	0.35	2.7	0.12	0.031	2.6	1.9
1.275 – 1.35	6.53	8.7	1.7	0.58	7.2	0.27	0.28	0.33	2.7	0.12	0.049	3.0	1.9
1.35 – 1.425	6.40	9.3	1.7	0.60	8.2	0.14	0.23	0.35	2.7	0.043	0.038	2.0	1.9
1.425 – 1.5	5.92	9.5	1.7	0.61	8.6	0.23	0.096	0.31	2.7	0.12	0.025	1.3	1.9
1.5 – 1.575	5.83	10.	1.8	0.64	8.9	0.33	0.68	0.36	2.7	0.094	0.034	2.4	1.9
1.575 – 1.65	5.98	11.	1.7	0.64	10.	0.22	0.37	0.31	2.7	0.069	0.021	0.79	1.9
1.65 – 1.725	5.63	11.	1.8	0.67	10.	0.19	0.34	0.31	2.7	0.096	0.048	1.5	1.9
1.725 – 1.8	5.38	11.	1.8	0.66	10.	0.29	0.63	0.26	2.7	0.090	0.022	2.1	1.9
1.8 – 1.875	5.30	12.	1.8	0.66	11.	0.21	0.82	0.26	2.7	0.068	0.041	1.8	1.9
1.875 – 1.95	4.94	13.	1.9	0.70	12.	0.24	0.91	0.28	2.7	0.084	0.049	3.2	1.9
1.95 – 2.025	4.93	13.	1.9	0.72	12.	0.20	1.1	0.21	2.7	0.025	0.050	2.2	1.9
2.025 – 2.1	4.99	14.	1.8	0.69	13.	0.31	0.69	0.19	2.7	0.045	0.039	1.5	1.9
2.1 – 2.175	4.70	15.	1.9	0.71	14.	0.30	0.71	0.20	2.7	0.077	0.0069	1.1	1.9
2.175 – 2.25	4.53	15.	1.9	0.72	15.	0.34	0.33	0.21	2.7	0.069	0.033	1.5	1.9
2.25 – 2.325	4.38	15.	1.9	0.74	14.	0.25	0.41	0.14	2.7	0.13	0.058	2.4	1.9
2.325 – 2.4	4.18	15.	2.5	0.92	14.	0.077	0.79	0.12	2.7	0.071	0.023	2.3	1.8

Table A.11: Differential cross section in $|y(j_2)|$ ($N_{jets} \geq 2$) and break down of the systematic uncertainties for the combination of both decay channels.

$ y(j_3) $	$\frac{d\sigma}{d y(j_3) }$ [pb]	Tot [%]	stat [%]	Unf stat [%]	JES [%]	JER [%]	PU [%]	Bgnd [%]	Lumi [%]	LES [%]	LER [%]	Unf sys [%]	Eff [%]
0 – 0.1	1.46	11.	3.3	0.85	9.4	0.20	0.60	0.95	3.0	0.14	0.027	1.9	2.0
0.1 – 0.2	1.55	11.	3.2	0.87	9.5	0.17	0.40	0.77	2.9	0.13	0.051	2.0	2.0
0.2 – 0.3	1.47	12.	3.3	0.91	10.	0.22	0.44	0.98	2.9	0.089	0.035	2.0	2.1
0.3 – 0.4	1.38	10.	3.5	0.95	8.6	0.13	0.33	0.90	2.9	0.061	0.029	1.7	2.1
0.4 – 0.5	1.55	9.6	3.3	0.99	7.9	0.17	0.34	0.83	2.9	0.066	0.023	2.2	2.0
0.5 – 0.6	1.54	10.	3.1	0.90	8.7	0.40	0.30	0.76	2.9	0.15	0.014	2.4	2.0
0.6 – 0.7	1.43	12.	3.4	0.96	10.	0.29	0.34	0.90	2.9	0.22	0.039	1.7	2.0
0.7 – 0.8	1.35	11.	3.5	0.95	9.9	0.27	0.17	1.1	3.0	0.14	0.024	1.6	2.1
0.8 – 0.9	1.50	13.	3.2	0.93	11.	0.082	0.069	0.76	2.9	0.048	0.054	3.3	2.0
0.9 – 1	1.39	13.	3.4	0.99	11.	0.22	0.18	0.69	2.9	0.083	0.032	1.7	2.0
1 – 1.1	1.40	10.	3.4	0.98	9.0	0.29	0.42	0.81	2.9	0.15	0.018	1.8	2.0
1.1 – 1.2	1.32	11.	3.7	1.0	8.8	0.21	0.66	0.94	2.9	0.13	0.037	2.7	2.1
1.2 – 1.3	1.27	11.	3.6	1.0	9.8	0.24	0.42	0.81	2.9	0.077	0.034	2.0	2.0
1.3 – 1.4	1.30	12.	3.6	1.0	10.	0.27	0.30	0.70	2.9	0.15	0.011	3.3	2.0
1.4 – 1.5	1.22	14.	3.8	1.1	12.	0.23	0.44	0.82	2.9	0.19	0.030	3.0	2.0
1.5 – 1.6	1.11	13.	3.8	1.1	12.	0.45	0.33	0.80	2.9	0.075	0.057	1.3	2.0
1.6 – 1.7	1.18	15.	3.7	1.0	14.	0.12	0.36	0.69	2.8	0.057	0.019	3.1	2.0
1.7 – 1.8	1.09	16.	4.4	1.2	15.	0.22	1.2	0.71	2.9	0.051	0.039	1.9	2.1
1.8 – 1.9	1.02	16.	4.1	1.1	15.	0.35	1.1	0.62	2.8	0.12	0.060	1.8	2.0
1.9 – 2	1.00	14.	4.1	1.2	13.	0.25	0.79	0.65	2.8	0.21	0.039	3.8	2.0
2 – 2.1	1.03	17.	3.9	1.2	15.	0.23	1.2	0.61	2.8	0.25	0.031	3.7	2.0
2.1 – 2.2	0.912	20.	4.2	1.3	19.	0.13	0.76	0.55	2.8	0.19	0.011	2.2	2.0
2.2 – 2.3	0.952	19.	4.0	1.3	18.	0.28	0.32	0.42	2.8	0.074	0.039	1.6	1.9
2.3 – 2.4	0.879	19.	5.0	1.5	17.	0.45	1.1	0.61	2.8	0.16	0.044	5.3	2.0

Table A.12: Differential cross section in $|y(j_3)|$ ($N_{jets} \geq 3$) and break down of the systematic uncertainties for the combination of both decay channels.

$ y(j_4) $	$\frac{d\sigma}{d y(j_4) }$ [pb]	Tot [%]	stat [%]	Unf stat [%]	JES [%]	JER [%]	PU [%]	Bgnd [%]	Lumi [%]	LES [%]	LER [%]	Unf sys [%]	Eff [%]
0 – 0.2	0.275	16.	6.1	1.4	14.	0.38	1.4	1.7	3.2	0.18	0.066	1.2	2.2
0.2 – 0.4	0.313	12.	5.6	1.3	9.7	0.23	0.31	1.2	3.0	0.18	0.032	2.5	2.1
0.4 – 0.6	0.285	14.	5.9	1.4	11.	0.26	0.39	1.2	3.0	0.15	0.036	3.6	2.1
0.6 – 0.8	0.259	16.	6.4	1.4	13.	0.27	0.51	1.6	3.2	0.030	0.030	2.7	2.2
0.8 – 1	0.259	14.	6.4	1.6	12.	0.23	0.96	1.5	3.1	0.027	0.018	2.5	2.2
1 – 1.2	0.237	15.	6.9	1.6	13.	0.24	0.087	1.6	3.2	0.24	0.045	2.4	2.2
1.2 – 1.4	0.229	15.	6.9	1.6	12.	0.58	0.29	1.6	3.1	0.11	0.020	4.8	2.2
1.4 – 1.6	0.210	18.	7.2	1.7	16.	0.44	0.67	1.1	3.0	0.11	0.012	3.2	2.1
1.6 – 1.8	0.208	21.	7.2	1.8	19.	0.30	1.1	1.5	3.1	0.26	0.015	0.31	2.2
1.8 – 2	0.204	25.	7.0	1.9	24.	0.36	0.88	0.91	2.9	0.22	0.017	0.95	2.1
2 – 2.2	0.204	21.	7.1	1.9	19.	0.10	1.0	0.83	2.9	0.32	0.019	1.4	2.0
2.2 – 2.4	0.207	21.	7.3	2.1	19.	0.019	0.62	0.28	2.7	0.031	0.047	4.9	1.9

Table A.13: Differential cross section in $|y(j_4)|$ ($N_{jets} \geq 4$) and break down of the systematic uncertainties for the combination of both decay channels.

$ y(j_5) $	$\frac{d\sigma}{d y(j_5) }$ [pb]	Tot [%]	stat [%]	Unf stat [%]	JES [%]	JER [%]	PU [%]	Bgnd [%]	Lumi [%]	LES [%]	LER [%]	Unf sys [%]	Eff [%]
0 – 0.4	0.0524	20.	11.	2.4	15.	1.1	1.0	2.0	3.3	0.12	0.0073	5.3	2.3
0.4 – 0.8	0.0543	19.	10.	2.4	14.	0.55	0.43	2.0	3.3	0.17	0.060	3.0	2.3
0.8 – 1.2	0.0529	20.	11.	2.6	15.	0.40	1.5	1.8	3.2	0.37	0.0067	5.4	2.2
1.2 – 1.6	0.0420	22.	13.	2.9	17.	0.30	0.89	2.0	3.3	0.31	0.032	4.8	2.3
1.6 – 2	0.0344	33.	14.	3.1	29.	0.59	2.2	1.9	3.2	0.68	0.19	4.4	2.2
2 – 2.4	0.0246	36.	17.	3.8	30.	0.99	0.90	1.1	3.0	1.2	0.029	7.9	2.1

Table A.14: Differential cross section in $|y(j_5)|$ ($N_{jets} \geq 5$) and break down of the systematic uncertainties for the combination of both decay channels.

$y_{diff}(Z, j_1)$	$\frac{1}{\sigma} \frac{d\sigma}{dy_{diff}(Z, j_1)}$	Tot	stat	Unf stat	JES	JER	PU	Bgnd	Lumi	LES	LER	Unf sys	Eff
		[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]
0 – 0.2	1.15	1.5	0.20	0.099	1.4	0.054	0.11	0.039	0.0099	0.0053	0.00075	0.21	0.0069
0.2 – 0.4	1.04	1.2	0.22	0.11	1.2	0.040	0.055	0.031	0.0084	0.0047	0.0011	0.23	0.0059
0.4 – 0.6	0.866	0.83	0.24	0.12	0.74	0.034	0.041	0.024	0.0060	0.016	0.0027	0.26	0.0042
0.6 – 0.8	0.673	0.83	0.28	0.15	0.71	0.050	0.059	0.021	0.0060	0.0099	0.0033	0.28	0.0042
0.8 – 1	0.482	1.1	0.34	0.18	0.98	0.037	0.075	0.038	0.0087	0.0099	0.0067	0.27	0.0061
1 – 1.2	0.333	2.1	0.42	0.23	2.0	0.062	0.070	0.063	0.018	0.015	0.0091	0.31	0.012
1.2 – 1.4	0.216	3.2	0.56	0.31	3.1	0.092	0.11	0.074	0.022	0.035	0.022	0.44	0.015
1.4 – 1.6	0.130	4.9	0.76	0.44	4.7	0.10	0.25	0.086	0.029	0.058	0.021	0.46	0.020
1.6 – 1.8	0.0692	6.5	1.1	0.65	6.1	0.21	0.37	0.096	0.029	0.067	0.055	1.7	0.021
1.8 – 2	0.0313	9.0	1.8	1.1	8.5	0.33	0.64	0.091	0.024	0.11	0.0073	1.6	0.020
2 – 2.2	0.00965	11.	3.3	2.1	9.5	0.21	1.6	0.12	0.051	0.15	0.063	2.3	0.035
2.2 – 2.4	0.00136	17.	10.	6.5	11.	1.1	1.6	0.15	0.061	0.17	0.13	4.1	0.042

Table A.15: Normalized differential cross section in $y_{diff}(Z, j_1)$ ($N_{jets} \geq 1$) and break down of the systematic uncertainties for the combination of both decay channels.

$y_{diff}(Z, j_1)$	$\frac{1}{\sigma} \frac{d\sigma}{dy_{diff}(Z, j_1)}$	Tot	stat	Unf stat	JES	JER	PU	Bgnd	Lumi	LES	LER	Unf sys	Eff
		[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]
0 – 0.2	1.06	0.85	0.54	0.18	0.38	0.038	0.029	0.22	0.062	0.0038	0.0016	0.45	0.043
0.2 – 0.4	0.976	0.72	0.58	0.19	0.28	0.036	0.057	0.17	0.049	0.0090	0.0039	0.18	0.036
0.4 – 0.6	0.830	0.74	0.62	0.21	0.23	0.027	0.096	0.12	0.034	0.021	0.0059	0.18	0.024
0.6 – 0.8	0.674	0.85	0.69	0.25	0.21	0.035	0.10	0.11	0.034	0.018	0.0082	0.32	0.024
0.8 – 1	0.502	0.96	0.80	0.30	0.12	0.017	0.058	0.18	0.047	0.019	0.0079	0.37	0.033
1 – 1.2	0.375	1.2	0.95	0.37	0.11	0.089	0.12	0.28	0.083	0.0087	0.010	0.45	0.061
1.2 – 1.4	0.260	1.6	1.1	0.48	0.69	0.077	0.13	0.33	0.11	0.019	0.0075	0.58	0.074
1.4 – 1.6	0.170	2.1	1.4	0.63	0.86	0.081	0.15	0.35	0.12	0.031	0.0077	0.91	0.083
1.6 – 1.8	0.0923	3.3	2.0	0.86	2.1	0.12	0.17	0.41	0.14	0.049	0.017	1.2	0.10
1.8 – 2	0.0471	4.4	2.9	1.4	3.0	0.18	0.16	0.39	0.13	0.12	0.018	0.51	0.095
2 – 2.2	0.0161	6.5	5.1	2.5	2.3	0.41	0.18	0.43	0.16	0.18	0.019	2.2	0.11
2.2 – 2.4	0.00211	19.	15.	8.0	4.9	0.19	2.2	0.48	0.18	0.24	0.25	4.5	0.13

Table A.16: Normalized differential cross section in $y_{diff}(Z, j_1)$ ($N_{jets} \geq 2$) and break down of the systematic uncertainties for the combination of both decay channels.

$y_{diff}(Z, j_2)$	$\frac{1}{\sigma} \frac{d\sigma}{dy_{diff}(Z, j_2)}$	Tot	stat	Unf stat	JES	JER	PU	Bgnd	Lumi	LES	LER	Unf sys	Eff
		[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]
0 – 0.2	0.885	1.4	0.62	0.20	1.1	0.047	0.075	0.25	0.080	0.020	0.00083	0.42	0.056
0.2 – 0.4	0.858	1.4	0.64	0.21	1.2	0.092	0.093	0.21	0.069	0.014	0.00074	0.30	0.048
0.4 – 0.6	0.796	1.1	0.66	0.22	0.71	0.073	0.061	0.11	0.037	0.013	0.0035	0.40	0.026
0.6 – 0.8	0.692	1.0	0.70	0.25	0.65	0.065	0.060	0.12	0.040	0.013	0.0041	0.22	0.028
0.8 – 1	0.560	1.0	0.79	0.29	0.46	0.036	0.068	0.12	0.038	0.021	0.0070	0.32	0.027
1 – 1.2	0.446	1.7	0.89	0.35	1.2	0.058	0.048	0.26	0.083	0.018	0.0083	0.67	0.060
1.2 – 1.4	0.320	2.0	1.1	0.43	1.5	0.050	0.14	0.35	0.11	0.026	0.015	0.70	0.082
1.4 – 1.6	0.216	2.9	1.3	0.55	2.4	0.099	0.32	0.38	0.13	0.032	0.026	0.50	0.091
1.6 – 1.8	0.131	3.6	1.9	0.84	2.8	0.31	0.35	0.42	0.15	0.041	0.049	0.88	0.10
1.8 – 2	0.0686	4.5	2.4	1.1	3.6	0.13	0.32	0.46	0.16	0.18	0.040	0.34	0.11
2 – 2.2	0.0234	7.0	4.3	2.1	4.9	0.34	0.82	0.47	0.17	0.082	0.072	0.61	0.12
2.2 – 2.4	0.00281	22.	15.	5.9	15.	1.1	1.2	0.51	0.19	0.16	0.40	2.9	0.14

Table A.17: Normalized differential cross section in $y_{diff}(Z, j_2)$ ($N_{jets} \geq 2$) and break down of the systematic uncertainties for the combination of both decay channels.

$y_{diff}(Z, dijet)$	$\frac{1}{\sigma} \frac{d\sigma}{dy_{diff}(Z, dijet)}$	Tot	stat	Unf stat	JES	JER	PU	Bgnd	Lumi	LES	LER	Unf sys	Eff
		[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]
0 – 0.2	1.10	0.68	0.51	0.16	0.18	0.041	0.12	0.31	0.089	0.042	0.0083	0.14	0.063
0.2 – 0.4	1.03	0.66	0.56	0.18	0.18	0.075	0.037	0.17	0.054	0.0024	0.0098	0.086	0.038
0.4 – 0.6	0.926	0.77	0.59	0.21	0.28	0.049	0.098	0.17	0.049	0.0092	0.0094	0.27	0.034
0.6 – 0.8	0.740	0.83	0.66	0.24	0.21	0.023	0.058	0.17	0.040	0.022	0.0042	0.33	0.028
0.8 – 1	0.536	1.1	0.77	0.30	0.52	0.043	0.10	0.27	0.081	0.029	0.0035	0.27	0.056
1 – 1.2	0.346	1.4	0.97	0.41	0.73	0.079	0.24	0.33	0.10	0.050	0.0093	0.42	0.070
1.2 – 1.4	0.187	2.1	1.4	0.63	0.88	0.12	0.35	0.36	0.11	0.028	0.0099	0.91	0.086
1.4 – 1.6	0.0857	3.5	2.1	1.0	2.2	0.15	0.029	0.37	0.12	0.043	0.020	1.2	0.089
1.6 – 1.8	0.0335	6.0	3.4	2.0	4.4	0.32	0.075	0.40	0.13	0.052	0.042	1.2	0.094
1.8 – 2	0.0101	9.1	6.6	3.7	4.4	0.85	1.0	0.42	0.14	0.12	0.12	2.1	0.10
2 – 2.2	0.00188	23.	16.	11.	4.2	3.8	2.1	0.41	0.14	0.25	0.20	9.3	0.10
2.2 – 2.4	0.000115	1.4e+02	74.	34.	50.	13.	7.5	0.46	0.17	0.099	3.0	95.	0.12

Table A.18: Normalized differential cross section in $y_{diff}(Z, dijet)$ ($N_{jets} \geq 2$) and break down of the systematic uncertainties for the combination of both decay channels.

$y_{diff}(j_1, j_2)$	$\frac{1}{\sigma} \frac{d\sigma}{dy_{diff}(j_1, j_2)}$	Tot	stat	Unf stat	JES	JER	PU	Bgnd	Lumi	LES	LER	Unf sys	Eff
		[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]
0 – 0.2	0.884	1.1	0.58	0.19	0.76	0.053	0.069	0.14	0.066	0.020	0.0042	0.43	0.046
0.2 – 0.4	0.925	1.1	0.58	0.19	0.75	0.0029	0.041	0.084	0.040	0.013	0.0023	0.61	0.029
0.4 – 0.6	0.798	0.95	0.62	0.21	0.57	0.022	0.038	0.070	0.026	0.015	0.0035	0.38	0.018
0.6 – 0.8	0.676	1.0	0.66	0.23	0.52	0.030	0.081	0.068	0.031	0.013	0.0024	0.47	0.022
0.8 – 1	0.555	0.91	0.73	0.26	0.43	0.016	0.13	0.045	0.037	0.019	0.0039	0.14	0.026
1 – 1.2	0.423	1.1	0.85	0.31	0.51	0.042	0.091	0.11	0.056	0.0088	0.0030	0.35	0.039
1.2 – 1.4	0.303	1.9	1.0	0.38	1.3	0.046	0.085	0.20	0.083	0.0089	0.0029	0.81	0.058
1.4 – 1.6	0.205	3.1	1.3	0.52	2.7	0.093	0.15	0.24	0.098	0.021	0.0026	0.70	0.071
1.6 – 1.8	0.124	4.3	1.7	0.69	3.7	0.17	0.071	0.25	0.10	0.022	0.0035	0.99	0.072
1.8 – 2	0.0703	6.8	2.2	0.89	6.3	0.15	0.25	0.29	0.11	0.060	0.031	0.65	0.079
2 – 2.2	0.0296	10.	3.7	1.7	8.4	0.27	0.21	0.31	0.15	0.23	0.020	3.9	0.10
2.2 – 2.4	0.00691	19.	10.	4.3	14.	0.28	0.69	0.40	0.12	0.15	0.11	7.0	0.097

Table A.19: Normalized differential cross section in $y_{diff}(j_1, j_2)$ ($N_{jets} \geq 2$) and break down of the systematic uncertainties for the combination of both decay channels.

$y_{sum}(j_1, j_2)$	$\frac{1}{\sigma} \frac{d\sigma}{dy_{sum}(j_1, j_2)}$	Tot	stat	Unf stat	JES	JER	PU	Bgnd	Lumi	LES	LER	Unf sys	Eff
		[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]
0 – 0.2	0.796	1.1	0.58	0.20	0.87	0.047	0.050	0.026	0.0018	0.011	0.0013	0.31	0.0013
0.2 – 0.4	0.758	1.0	0.62	0.22	0.77	0.034	0.034	0.028	0.0012	0.0081	0.0038	0.20	0.00084
0.4 – 0.6	0.711	1.1	0.64	0.22	0.83	0.010	0.016	0.030	0.0012	0.013	0.0063	0.35	0.0011
0.6 – 0.8	0.645	0.98	0.68	0.24	0.46	0.027	0.055	0.012	0.0026	0.019	0.0029	0.49	0.0018
0.8 – 1	0.566	0.96	0.73	0.25	0.29	0.032	0.064	0.010	0.0046	0.013	0.0031	0.49	0.0039
1 – 1.2	0.469	1.0	0.80	0.28	0.37	0.024	0.11	0.025	0.0040	0.011	0.0048	0.43	0.0029
1.2 – 1.4	0.379	1.2	0.92	0.32	0.50	0.048	0.048	0.020	0.0020	0.014	0.0029	0.48	0.0017
1.4 – 1.6	0.285	2.2	1.1	0.39	1.9	0.065	0.12	0.053	0.0071	0.017	0.0098	0.33	0.0067
1.6 – 1.8	0.195	3.1	1.3	0.48	2.8	0.095	0.17	0.11	0.0047	0.012	0.0038	0.38	0.0033
1.8 – 2	0.120	5.0	1.8	0.61	4.7	0.14	0.20	0.10	0.014	0.013	0.0033	0.58	0.0097
2 – 2.2	0.0612	6.1	2.6	0.94	5.4	0.23	0.35	0.13	0.013	0.021	0.025	0.49	0.0090
2.2 – 2.4	0.0166	10.	5.9	2.2	7.8	0.59	0.36	0.21	0.013	0.045	0.0082	1.0	0.0093

Table A.20: Normalized differential cross section in $y_{sum}(j_1, j_2)$ ($N_{jets} \geq 2$) and break down of the systematic uncertainties for the combination of both decay channels.

$y_{sum}(Z, j_1)$	$\frac{1}{\sigma} \frac{d\sigma}{dy_{sum}(Z, j_1)}$	Tot	stat	Unf stat	JES	JER	PU	Bgnd	Lumi	LES	LER	Unf sys	Eff
		[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]
0 – 0.2	0.756	0.64	0.27	0.14	0.41	0.049	0.12	0.0078	0.00086	0.011	0.0035	0.35	0.00062
0.2 – 0.4	0.734	0.45	0.26	0.14	0.24	0.065	0.13	0.0081	0.0014	0.0089	0.0041	0.19	0.0010
0.4 – 0.6	0.701	0.40	0.27	0.14	0.17	0.051	0.098	0.0089	0.0018	0.011	0.0049	0.16	0.0013
0.6 – 0.8	0.643	0.51	0.28	0.14	0.20	0.034	0.051	0.0073	0.0025	0.010	0.0053	0.34	0.0017
0.8 – 1	0.578	0.49	0.33	0.16	0.13	0.027	0.029	0.0077	0.0017	0.0069	0.0064	0.30	0.0014
1 – 1.2	0.496	0.55	0.35	0.17	0.20	0.038	0.018	0.0048	0.0016	0.0099	0.0035	0.34	0.0011
1.2 – 1.4	0.407	0.78	0.37	0.18	0.32	0.065	0.13	0.0090	0.00094	0.0063	0.0047	0.57	0.00081
1.4 – 1.6	0.307	1.0	0.44	0.21	0.61	0.13	0.21	0.017	0.0029	0.010	0.0038	0.62	0.0022
1.6 – 1.8	0.207	1.3	0.56	0.26	1.0	0.13	0.39	0.024	0.0051	0.064	0.025	0.40	0.0037
1.8 – 2	0.117	1.7	0.76	0.35	1.3	0.23	0.52	0.047	0.011	0.20	0.023	0.47	0.0079
2 – 2.2	0.0469	2.2	1.2	0.56	1.2	0.15	0.59	0.061	0.017	0.16	0.0061	1.2	0.012
2.2 – 2.4	0.00670	5.4	3.9	1.7	1.5	0.077	0.81	0.082	0.025	0.068	0.036	2.9	0.019

Table A.21: Normalized differential cross section in $y_{sum}(Z, j_1)$ ($N_{jets} \geq 1$) and break down of the systematic uncertainties for the combination of both decay channels.

$y_{sum}(Z, j_1)$	$\frac{1}{\sigma} \frac{d\sigma}{dy_{sum}(Z, j_1)}$	Tot	stat	Unf stat	JES	JER	PU	Bgnd	Lumi	LES	LER	Unf sys	Eff
		[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]
0 – 0.2	0.797	0.86	0.61	0.22	0.55	0.046	0.020	0.042	0.0096	0.030	0.0067	0.13	0.0072
0.2 – 0.4	0.774	0.71	0.63	0.23	0.16	0.076	0.077	0.028	0.0052	0.019	0.0062	0.11	0.0036
0.4 – 0.6	0.724	0.75	0.66	0.23	0.19	0.088	0.079	0.042	0.012	0.023	0.0068	0.15	0.0083
0.6 – 0.8	0.660	0.86	0.73	0.26	0.16	0.065	0.088	0.034	0.0081	0.011	0.0092	0.32	0.0068
0.8 – 1	0.574	0.87	0.75	0.26	0.29	0.046	0.12	0.036	0.0075	0.0097	0.015	0.11	0.0052
1 – 1.2	0.475	1.0	0.84	0.29	0.24	0.054	0.049	0.034	0.0091	0.017	0.0018	0.47	0.0077
1.2 – 1.4	0.385	1.1	0.95	0.32	0.17	0.061	0.10	0.050	0.010	0.028	0.0030	0.42	0.0086
1.4 – 1.6	0.279	1.3	1.2	0.37	0.20	0.10	0.11	0.080	0.019	0.010	0.0021	0.53	0.013
1.6 – 1.8	0.185	2.0	1.5	0.49	0.70	0.25	0.36	0.12	0.030	0.058	0.010	0.80	0.024
1.8 – 2	0.100	2.6	2.0	0.64	1.4	0.27	0.52	0.23	0.062	0.15	0.029	0.36	0.045
2 – 2.2	0.0412	3.7	3.1	1.1	0.93	0.32	0.37	0.30	0.093	0.21	0.016	1.2	0.065
2.2 – 2.4	0.00574	11.	9.2	2.9	2.4	0.45	1.1	0.41	0.13	0.40	0.058	5.1	0.096

Table A.22: Normalized differential cross section in $y_{sum}(Z, j_1)$ ($N_{jets} \geq 2$) and break down of the systematic uncertainties for the combination of both decay channels.

$y_{sum}(Z, j_2)$	$\frac{1}{\sigma} \frac{d\sigma}{dy_{sum}(Z, j_2)}$	Tot	stat	Unf stat	JES	JER	PU	Bgnd	Lumi	LES	LER	Unf sys	Eff
		[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]
0 – 0.2	0.836	0.94	0.62	0.23	0.61	0.068	0.12	0.023	0.0072	0.024	0.0033	0.24	0.0064
0.2 – 0.4	0.791	0.98	0.65	0.24	0.66	0.028	0.097	0.023	0.0056	0.033	0.0047	0.20	0.0043
0.4 – 0.6	0.740	0.81	0.68	0.24	0.35	0.064	0.054	0.019	0.0045	0.012	0.0052	0.12	0.0031
0.6 – 0.8	0.657	0.81	0.73	0.26	0.17	0.024	0.076	0.022	0.0070	0.026	0.0099	0.13	0.0057
0.8 – 1	0.572	0.88	0.79	0.28	0.13	0.025	0.057	0.015	0.0043	0.021	0.0088	0.20	0.0038
1 – 1.2	0.481	0.97	0.88	0.30	0.17	0.060	0.057	0.011	0.0029	0.0089	0.0092	0.22	0.0020
1.2 – 1.4	0.367	1.2	1.0	0.35	0.24	0.044	0.12	0.016	0.0040	0.022	0.0096	0.47	0.0036
1.4 – 1.6	0.262	2.0	1.3	0.41	1.4	0.044	0.31	0.017	0.0081	0.038	0.0092	0.068	0.0060
1.6 – 1.8	0.166	2.2	1.6	0.52	1.3	0.12	0.25	0.096	0.013	0.12	0.010	0.13	0.0090
1.8 – 2	0.0889	3.4	2.3	0.72	2.1	0.20	0.36	0.14	0.029	0.30	0.012	1.0	0.022
2 – 2.2	0.0338	4.7	3.8	1.3	1.5	0.47	0.79	0.17	0.051	0.36	0.031	1.8	0.036
2.2 – 2.4	0.00432	14.	12.	4.6	0.60	0.64	2.5	0.49	0.14	0.13	0.020	5.9	0.10

Table A.23: Normalized differential cross section in $y_{sum}(Z, j_2)$ ($N_{jets} \geq 2$) and break down of the systematic uncertainties for the combination of both decay channels.

$y_{sum}(Z, dijet)$	$\frac{1}{\sigma} \frac{d\sigma}{dy_{sum}(Z, dijet)}$	Tot	stat	Unf stat	JES	JER	PU	Bgnd	Lumi	LES	LER	Unf sys	Eff
		[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]
0 – 0.2	0.889	0.80	0.60	0.22	0.081	0.043	0.11	0.033	0.0022	0.058	0.0074	0.45	0.0021
0.2 – 0.4	0.868	0.72	0.59	0.22	0.074	0.049	0.045	0.0070	0.0023	0.022	0.0044	0.33	0.0017
0.4 – 0.6	0.798	0.72	0.62	0.22	0.12	0.042	0.088	0.0089	0.0040	0.010	0.0060	0.24	0.0028
0.6 – 0.8	0.709	0.77	0.67	0.24	0.12	0.034	0.081	0.013	0.0024	0.0054	0.0052	0.26	0.0022
0.8 – 1	0.600	0.83	0.74	0.25	0.10	0.070	0.056	0.018	0.0033	0.018	0.0017	0.24	0.0031
1 – 1.2	0.461	1.1	0.86	0.28	0.10	0.10	0.094	0.011	0.0029	0.020	0.0019	0.60	0.0025
1.2 – 1.4	0.324	1.4	1.1	0.33	0.41	0.14	0.22	0.041	0.0022	0.046	0.0090	0.73	0.0015
1.4 – 1.6	0.204	1.9	1.4	0.43	0.50	0.045	0.25	0.062	0.0013	0.032	0.029	1.1	0.0011
1.6 – 1.8	0.102	2.7	2.0	0.59	1.1	0.21	0.58	0.045	0.011	0.20	0.038	1.1	0.0095
1.8 – 2	0.0364	4.3	3.4	0.96	1.6	0.41	0.29	0.059	0.011	0.30	0.059	1.6	0.0083
2 – 2.2	0.00757	10.	7.9	2.1	3.0	1.2	1.5	0.20	0.032	0.75	0.093	4.6	0.023
2.2 – 2.4	0.000605	39.	28.	9.8	9.9	0.69	3.9	0.39	0.097	0.074	0.049	24.	0.074

Table A.24: Normalized differential cross section in $y_{sum}(Z, dijet)$ ($N_{jets} \geq 2$) and break down of the systematic uncertainties for the combination of both decay channels.

H_T	$\frac{d\sigma}{dH_T}$	Tot	stat	Unf stat	JES	JER	PU	Bgnd	Lumi	LES	LER	Unf sys	Eff
[GeV]	$\left[\frac{pb}{GeV}\right]$	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]
30 – 39	2.52	8.8	0.30	0.13	8.2	0.26	0.21	0.063	2.6	0.070	0.025	0.65	1.8
39 – 49	1.37	5.2	0.21	0.12	3.6	0.27	0.48	0.075	2.6	0.076	0.037	1.8	1.8
49 – 62	0.762	4.9	0.26	0.15	3.0	0.32	0.56	0.062	2.6	0.095	0.052	1.9	1.9
62 – 78	0.520	6.8	0.28	0.16	5.5	0.28	0.47	0.084	2.6	0.10	0.055	2.3	1.9
78 – 96	0.346	7.4	0.31	0.16	6.1	0.28	0.45	0.15	2.7	0.099	0.041	2.5	1.9
96 – 118	0.215	7.4	0.36	0.18	6.1	0.24	0.44	0.23	2.7	0.098	0.046	2.6	1.9
118 – 150	0.124	7.4	0.42	0.19	6.1	0.27	0.41	0.34	2.7	0.086	0.042	2.5	1.9
150 – 190	0.0647	7.9	0.53	0.22	6.4	0.24	0.45	0.45	2.8	0.067	0.036	3.1	1.9
190 – 240	0.0319	7.9	0.70	0.27	6.3	0.27	0.50	0.60	2.8	0.051	0.035	3.2	2.0
240 – 300	0.0150	8.2	0.95	0.34	7.0	0.15	0.49	0.67	2.8	0.087	0.024	2.2	2.0
300 – 370	0.00712	7.6	1.3	0.44	5.5	0.25	0.45	0.76	2.9	0.078	0.035	3.5	2.0
370 – 450	0.00319	9.5	1.8	0.55	7.6	0.26	0.53	0.86	2.9	0.028	0.016	3.9	2.0
450 – 540	0.00152	9.1	2.5	0.76	7.4	0.15	0.52	0.84	2.9	0.13	0.054	2.6	2.0
540 – 650	0.000659	6.6	3.4	1.0	3.6	0.064	0.36	0.84	2.9	0.15	0.064	2.0	2.1
650 – 800	0.000282	10.	4.3	1.3	7.9	0.065	0.38	0.60	2.9	0.054	0.0069	2.7	2.0
800 – 1000	0.000101	9.5	6.4	1.9	5.0	0.12	0.29	0.60	2.9	0.12	0.014	2.8	2.2
1000 – 1500	1.62e-05	15.	9.3	2.7	11.	0.30	0.92	0.68	2.9	0.044	0.028	3.0	2.1

Table A.25: Differential cross section in H_T ($N_{jets} \geq 1$) and break down of the systematic uncertainties for the combination of both decay channels.

H_T	$\frac{d\sigma}{dH_T}$	Tot	stat	Unf stat	JES	JER	PU	Bgnd	Lumi	LES	LER	Unf sys	Eff
[GeV]	$\left[\frac{pb}{GeV}\right]$	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]
60 – 78	0.125	18.	1.3	0.40	17.	0.12	0.37	0.35	2.5	0.067	0.045	0.53	1.8
78 – 96	0.155	9.6	0.64	0.24	8.9	0.13	0.34	0.26	2.7	0.085	0.029	1.1	1.9
96 – 118	0.124	8.3	0.59	0.22	7.4	0.29	0.44	0.29	2.7	0.10	0.043	1.6	1.9
118 – 150	0.0862	8.1	0.57	0.21	7.0	0.26	0.42	0.44	2.8	0.096	0.041	2.0	1.9
150 – 190	0.0512	8.6	0.65	0.23	7.3	0.23	0.50	0.55	2.8	0.084	0.039	2.8	2.0
190 – 240	0.0271	8.4	0.79	0.27	6.9	0.27	0.52	0.70	2.8	0.056	0.037	3.0	2.0
240 – 300	0.0135	8.6	1.0	0.34	7.4	0.11	0.47	0.74	2.9	0.090	0.027	2.2	2.0
300 – 370	0.00668	7.7	1.4	0.43	5.6	0.24	0.45	0.80	2.9	0.095	0.039	3.4	2.0
370 – 450	0.00306	9.6	1.8	0.55	7.8	0.25	0.53	0.89	2.9	0.029	0.016	3.8	2.0
450 – 540	0.00147	9.1	2.6	0.76	7.4	0.15	0.53	0.86	2.9	0.13	0.057	2.4	2.1
540 – 650	0.000644	6.7	3.5	1.0	3.7	0.046	0.42	0.85	2.9	0.15	0.066	2.2	2.1
650 – 800	0.000278	10.	4.4	1.3	7.9	0.069	0.37	0.61	2.9	0.052	0.0089	2.8	2.0
800 – 1200	6.41e-05	8.8	5.3	1.6	5.4	0.13	0.45	0.62	2.9	0.044	0.024	2.2	2.2

Table A.26: Differential cross section in H_T ($N_{jets} \geq 2$) and break down of the systematic uncertainties for the combination of both decay channels.

H_T [GeV]	$\frac{d\sigma}{dH_T}$ [$\frac{pb}{GeV}$]	Tot [%]	stat [%]	Unf stat [%]	JES [%]	JER [%]	PU [%]	Bgnd [%]	Lumi [%]	LES [%]	LER [%]	Unf sys [%]	Eff [%]
90 – 105	0.00492	34.	9.9	2.7	32.	0.74	2.0	1.5	2.1	0.062	0.039	1.4	1.6
105 – 125	0.0136	23.	3.3	0.96	22.	0.75	0.29	0.47	2.5	0.074	0.030	0.84	1.8
125 – 151	0.0175	13.	2.1	0.62	13.	0.53	0.32	0.48	2.7	0.13	0.058	0.86	2.0
151 – 185	0.0165	12.	1.6	0.48	12.	0.40	0.48	0.66	2.8	0.14	0.024	1.4	2.0
185 – 230	0.0124	12.	1.5	0.42	11.	0.25	0.59	0.96	2.9	0.039	0.032	2.5	2.0
230 – 290	0.00770	11.	1.6	0.45	9.5	0.26	0.54	1.1	3.0	0.13	0.028	2.1	2.1
290 – 366	0.00422	9.6	1.9	0.53	8.0	0.23	0.56	1.1	3.0	0.11	0.041	3.0	2.1
366 – 466	0.00194	12.	2.3	0.61	9.8	0.16	0.51	1.2	3.0	0.054	0.032	4.3	2.1
466 – 586	0.000801	9.7	3.2	0.86	7.9	0.20	0.49	1.1	3.0	0.089	0.038	2.2	2.1
586 – 767	0.000301	10.	3.9	1.1	7.8	0.25	0.71	0.94	3.0	0.054	0.0072	2.8	2.1

Table A.27: Differential cross section in H_T ($N_{jets} \geq 3$) and break down of the systematic uncertainties for the combination of both decay channels.

H_T [GeV]	$\frac{d\sigma}{dH_T}$ [$\frac{pb}{GeV}$]	Tot [%]	stat [%]	Unf stat [%]	JES [%]	JER [%]	PU [%]	Bgnd [%]	Lumi [%]	LES [%]	LER [%]	Unf sys [%]	Eff [%]
120 – 140	0.000279	89.	48.	16.	72.	1.4	8.6	2.0	1.9	0.60	0.27	12.	1.3
140 – 167	0.00116	35.	14.	3.2	32.	0.24	1.7	0.74	2.3	0.33	0.072	2.7	1.6
167 – 203	0.00212	21.	6.3	1.6	20.	0.71	0.81	0.90	2.8	0.061	0.027	1.7	2.0
203 – 253	0.00216	17.	4.4	1.0	16.	0.51	0.46	1.2	3.0	0.095	0.014	2.1	2.1
253 – 320	0.00188	15.	3.6	0.90	13.	0.098	0.52	1.3	3.1	0.037	0.021	2.2	2.1
320 – 410	0.00125	14.	3.7	0.91	13.	0.35	0.82	1.7	3.2	0.022	0.037	2.2	2.2
410 – 530	0.000551	15.	4.5	1.1	14.	0.24	0.61	1.9	3.2	0.045	0.042	3.0	2.2
530 – 690	0.000229	11.	5.4	1.3	8.6	0.28	0.59	1.4	3.1	0.12	0.025	3.0	2.2
690 – 910	7.27e-05	19.	7.8	2.0	16.	0.27	0.38	1.1	3.0	0.25	0.044	4.1	2.1

Table A.28: Differential cross section in H_T ($N_{jets} \geq 4$) and break down of the systematic uncertainties for the combination of both decay channels.

H_T [GeV]	$\frac{d\sigma}{dH_T}$ [$\frac{pb}{GeV}$]	Tot [%]	stat [%]	Unf stat [%]	JES [%]	JER [%]	PU [%]	Bgnd [%]	Lumi [%]	LES [%]	LER [%]	Unf sys [%]	Eff [%]
150 – 200	2.55e-06	5.3e+02	2.9e+02	47.	4.3e+02	13.	27.	4.3	0.97	1.2	0.21	42.	0.91
200 – 282	0.000255	29.	14.	3.1	24.	0.68	1.0	1.4	2.6	0.041	0.034	7.0	1.9
282 – 365	0.000296	27.	11.	2.3	23.	0.33	1.5	1.7	3.3	0.50	0.037	3.5	2.3
365 – 485	0.000230	22.	8.4	1.9	19.	0.21	0.74	2.4	3.4	0.34	0.025	1.2	2.4
485 – 650	0.000112	16.	9.2	2.2	12.	0.20	1.2	2.2	3.3	0.51	0.055	3.3	2.3
650 – 880	4.43e-05	14.	11.	2.8	8.2	0.28	0.22	1.3	3.1	0.25	0.021	1.1	2.2

Table A.29: Differential cross section in H_T ($N_{jets} \geq 5$) and break down of the systematic uncertainties for the combination of both decay channels.

Appendix B

Results of single channel ($Z \rightarrow e^+e^-$) at 13 TeV

This chapter presents the measured differential cross section as a function of jet observables with the full uncertainties in $Z \rightarrow e^+e^-$ channel for $Z + \text{jets}$ process, compared to the theoretical predictions of MADGRAPH_AMC@NLO and MADGRAPH (LO). It provides some supplementary results for Chapter 4. More details can be found in [59].

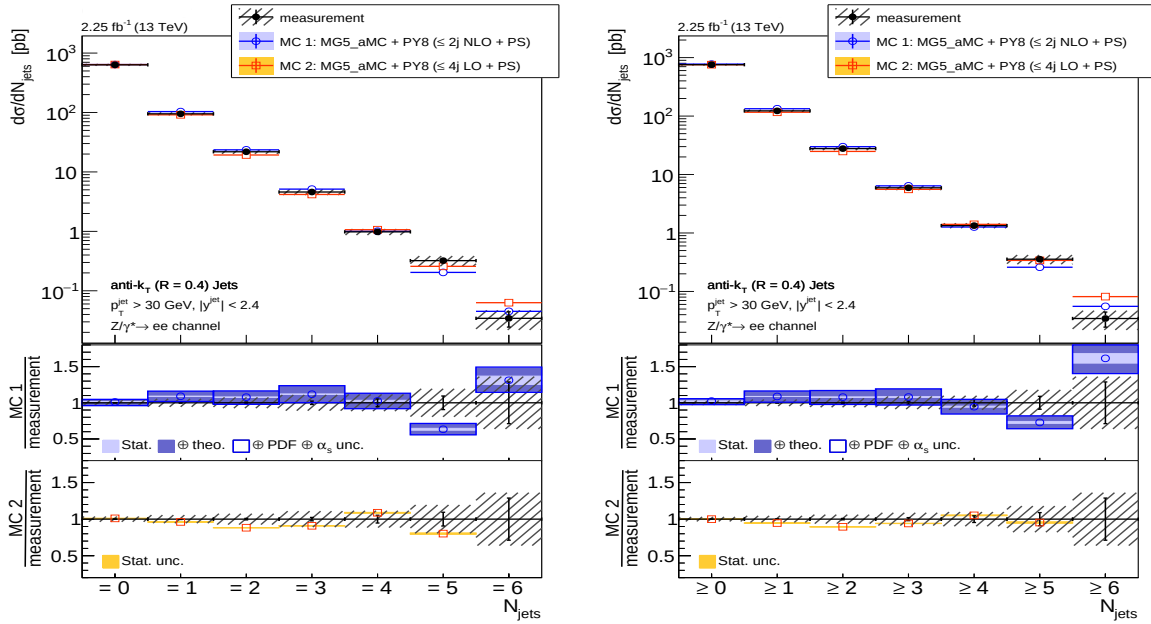


Figure B.1: The differential cross section as a function of exclusive (left) and inclusive (right) jet multiplicity in e^+e^- decay channel at 13 TeV, showing the full uncertainties. The measurement is compared with MADGRAPH_AMC@NLO (NLO) and MADGRAPH (LO) predictions.

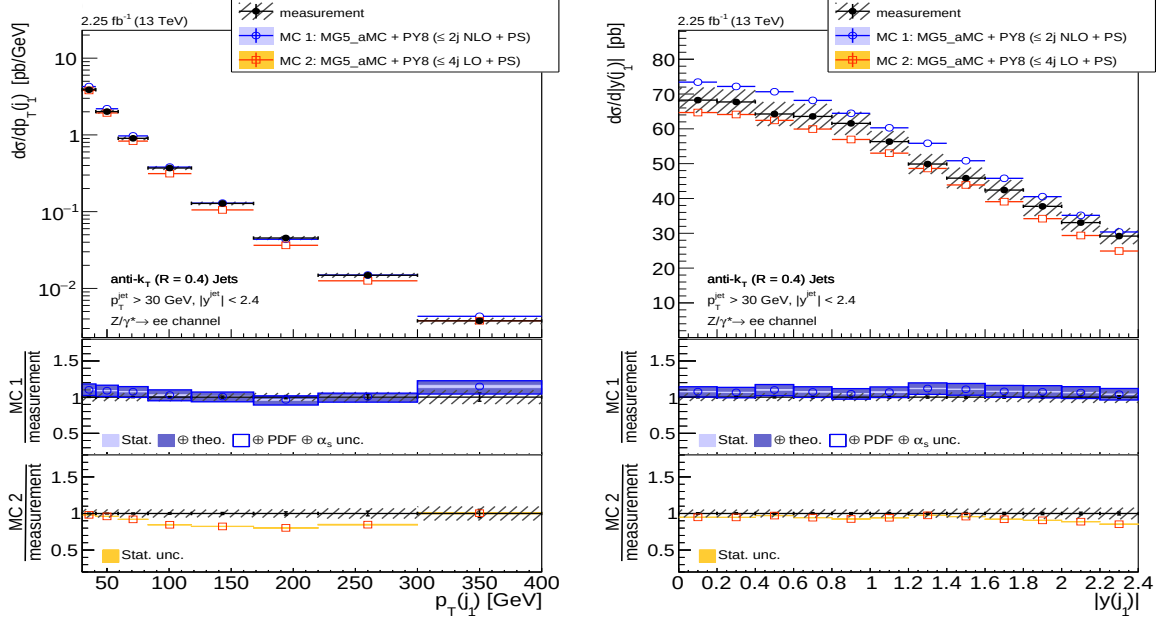


Figure B.2: The differential cross section as a function of leading jet p_T (left) and absolute rapidity (right) in e^+e^- decay channel at 13 TeV, showing the full uncertainties. The measurement is compared with MADGRAPH_AMC@NLO (NLO) and MADGRAPH (LO) predictions.

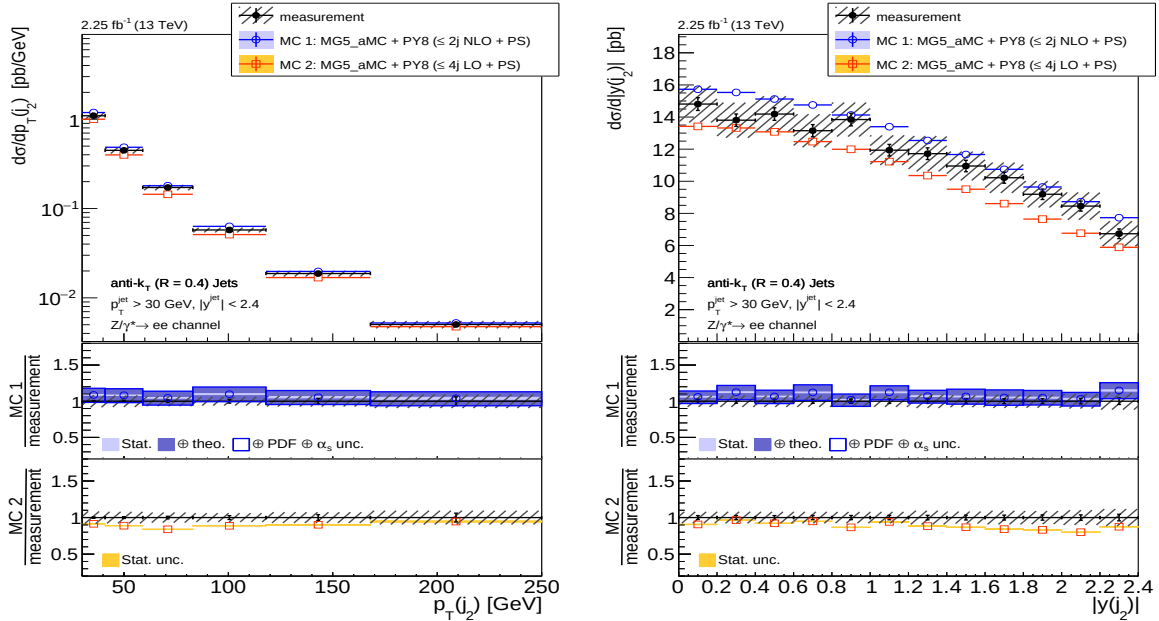


Figure B.3: The differential cross section as a function of the second jet p_T (left) and absolute rapidity (right) in e^+e^- decay channel at 13 TeV, showing the full uncertainties. The measurement is compared with MADGRAPH_AMC@NLO (NLO) and MADGRAPH (LO) predictions.

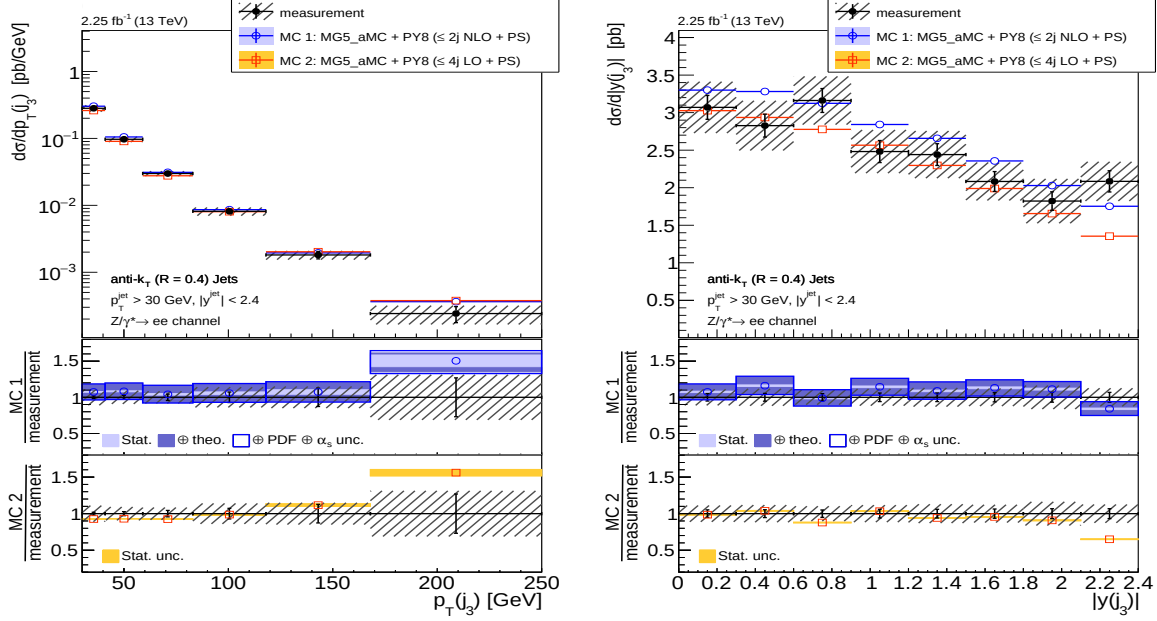


Figure B.4: The differential cross section as a function of the third jet p_T (left) and absolute rapidity (right) in e^+e^- decay channel at 13 TeV, showing the full uncertainties. The measurement is compared with MADGRAPH_AMC@NLO (NLO) and MADGRAPH (LO) predictions.

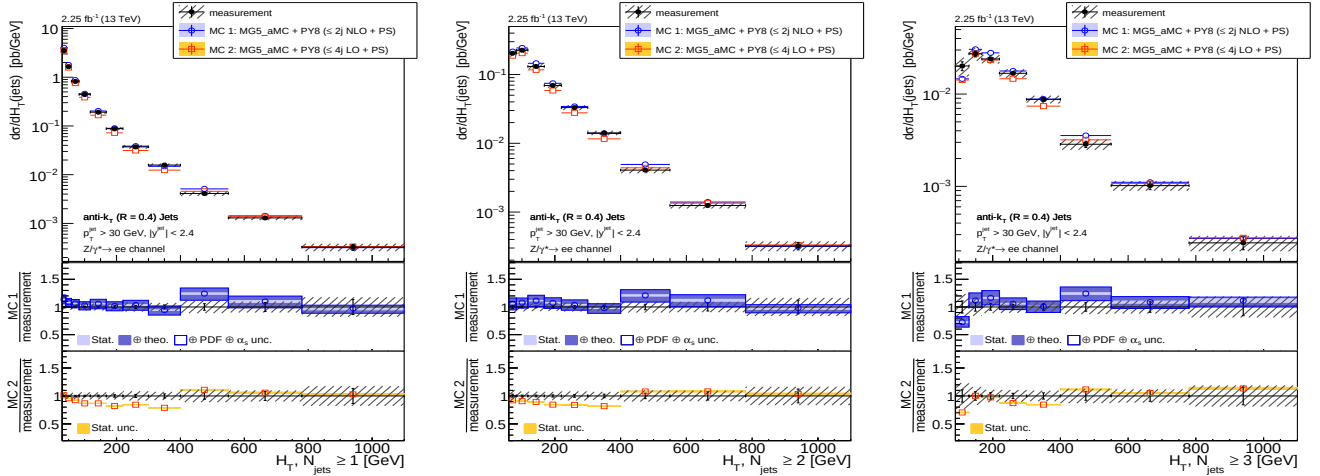


Figure B.5: The differential cross section as a function of scalar sum of jets p_T for $N_{jets} \geq 1$ (left), $N_{jets} \geq 2$ (middle), and $N_{jets} \geq 3$ (right) in e^+e^- decay channel at 13 TeV, showing the full uncertainties. The measurement is compared with MADGRAPH_AMC@NLO (NLO) and MADGRAPH (LO) predictions.