

BOSTON UNIVERSITY
GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

**MEASUREMENT OF THE TOTAL AND DIFFERENTIAL
CROSS SECTION OF Z BOSON PRODUCTION IN
ASSOCIATION WITH JETS IN PROTON-PROTON
COLLISIONS AT $\sqrt{s}=13$ TEV**

by

DANIEL J. ARCARO

B.S., Boston University, 2013

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

2020

© 2020 by
DANIEL J. ARCARO
All rights reserved

Approved by

First Reader

Lawrence R. Sulak, PhD
Professor of Physics

Second Reader

John M. Butler, PhD
Professor of Physics

**MEASUREMENT OF THE TOTAL AND DIFFERENTIAL
CROSS SECTION OF Z BOSON PRODUCTION IN
ASSOCIATION WITH JETS IN PROTON-PROTON
COLLISIONS AT $\sqrt{s}= 13$ TEV**

DANIEL J. ARCARO

Boston University, Graduate School of Arts and Sciences, 2020

Major Professor: Lawrence R. Sulak, PhD
Professor of Physics

ABSTRACT

Z boson production in association with jets at the LHC is a major source of background to standard model searches and it provides a sensitive evaluation of the accuracy of perturbative QCD predictions. The production of a Z boson, decaying to two charged muons ($\mu^+\mu^-$), in association with jets in proton-proton collisions at a center-of-mass energy of 13 TeV is presented. The cross sections and their ratios are measured with data recorded by the CMS experiment at the LHC in 2016, corresponding to an integrated luminosity of 35.9 fb^{-1} . The cross sections are measured as a function of observables including the transverse momentum and rapidity of the Z boson, the jet transverse momentum and rapidity for the five highest momentum jets, and jet multiplicity. The measurements are compared with predictions from a multi-leg next-to-leading-order Monte Carlo generator and a next-to-next-to-leading order calculation with next-to-next-to-leading logarithm resummation and parton showering.

Contents

1	Introduction	1
2	The CMS Hadron Calorimeter	2
2.1	Hardware	2
2.1.1	The CMS Hadron Calorimeter	2
2.1.2	Back End Electronics	6
3	Z+jets Measurement	14
3.1	Analysis	14
3.2	Data and Simulated Samples	14
3.3	Event Reconstruction, Object Selection and Corrections	22
3.3.1	Z Boson Selections	27
3.3.2	Jet Selections	27
3.4	Background Estimation	29
3.5	Observables	42
3.6	Phenomenological Models and Theoretical Calculations	47
3.7	Unfolding Procedure	50
3.8	Systematical Uncertainties	55
3.8.1	Jet Energy Scale	58
3.8.2	Jet Energy Resolution	58
3.8.3	Muon Energy Scale and Resolution	58
3.8.4	Muon Efficiencies	58
3.8.5	Integrated Luminosity	59
3.8.6	Pile Up	59
3.8.7	Background Cross Section	59
3.8.8	Parton Distribution Function, α_s	60

3.8.9	Unfolding Statistics	60
3.8.10	Unfolding	60
3.8.11	Software	61
3.9	Results	62
3.10	Summary	92
4	Conclusions	94
	References	95
	Curriculum Vitae	98

List of Tables

2.1	AMC13 Trigger Throttling State	9
3.1	Single Muon Data Samples	15
3.2	Monte Carlo Signal Samples	16
3.3	DY npNLO Weights	16
3.4	Monte Carlo Resonant Samples	17
3.5	Monte Carlo Non-Resonant Samples	17
3.6	Loose Jet ID Parameters	29
3.7	Unfolding Numbers	65
3.8	Differential cross section in 1 st jet $ \eta $ ($N_{\text{jets}} \geq 1$) and break down of the systematic uncertainties for the combination of both decay channels. Uncertainties are in percent.	88
3.9	Differential cross section in 1 st jet p_T ($N_{\text{jets}} \geq 1$) and break down of the systematic uncertainties for the combination of both decay channels. Uncertainties are in percent.	88
3.10	Differential cross section in exclusive jet multiplicity and break down of the systematic uncertainties for the combination of both decay channels. Uncertainties are in percent.	89
3.11	Differential cross section in p_T^Z and break down of the systematic uncertainties for the combination of both decay channels. Uncertainties are in percent.	90
3.12	Differential cross section in $ y_Z $ ($N_{\text{jets}} \geq 1$) and break down of the systematic uncertainties for the combination of both decay channels. Uncertainties are in percent.	91

List of Figures

2-1	HCAL layout and segmentation in the Z - η plane.	3
2-2	AMC13 data format with detailed header information.	11
3-1	Jet multiplicity for the inclusive and binned samples. The binned samples are stacked. Normalized to total cross section of the samples (5931.9fb for inclusive, 5818.37fb for binned).	16
3-2	Example of Tag and Probe fit for passing and failing probes.	19
3-3	Number of vertices in data and MC. Minimum bias of 69.2mb is used for the data profile.	20
3-4	Number of vertices in data and MC with minimum bias variations. Minimum Bias of 69.2mb varied 4.6% up (left) and down (right).	20
3-5	Number of vertices in data and MC without re-weighting the MC.	21
3-6	Number of vertices in era B-F (Left) and G-H (Right). Minimum bias of 69.2mb is used to calculate the profile.	21
3-7	Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Lepton p_T (left), η (center), ϕ (right) with one jet inclusive.	32
3-8	Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Dimuon mass and Z candidate $ y $ with one jet inclusive.	33
3-9	Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Jet multiplicity inclusive and exclusive.	33

3·10	Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Z candidate p_T 0 and 1 jets inclusive.	34
3·11	Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). First, second, and third jet absolute rapidity.	35
3·12	Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). First, second, and third jet p_T	36
3·13	Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Fourth and fifth jet p_T	37
3·14	Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Invariant mass of leading and subleading jet.	37
3·15	Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Total hadronic p_T one, two, and three jets inclusive.	38
3·16	Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Difference in rapidity of the Z boson and leading jet with one jet (left) and two jets (right) inclusive.	39
3·17	Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Rapidity sum of the Z boson and leading jet with one jet (left) and two jets (right) inclusive.	39
3·18	Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Difference (left) and sum (right) of rapidity of the Z boson and subleading jet.	40

3·19	Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Difference (left) and sum (right) of rapidity of the leading and subleading jet.	40
3·20	Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Difference (left) and sum (right) of rapidity of the Z boson and subleading jet.	41
3·21	Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Difference in azimuthal angle between the Z boson and leading jet with one(left), two(right), and three(right) jet inclusive.	42
3·22	Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Difference in azimuthal angle between the leading and subleading jet with two(left) and three(right) jet inclusive.	43
3·23	Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Difference in azimuthal angle between the leading and third jet (left) and subleading and third jet (right) with three jet inclusive.	43
3·24	Muon p_T and η	45
3·25	Muon $\Delta\phi$ and ΔR	45
3·26	Z p_T and $ y $	46
3·27	Jet Multiplicity	47
3·28	Jet p_T and $ y $	48
3·29	Invariant mass of leading and subleading jet.	49
3·31	Leading jet p_T and $ y $ response matrices	52
3·32	Second jet p_T and $ y $ response matrices	53

3-33	Dijet mass response matrix	53
3-34	Z p_T response matrix	54
3-35	Jet multiplicity response matrix	54
3-36	Leading jet p_T and $ y $ differential cross section	66
3-37	Second jet p_T and $ y $ differential cross section	67
3-38	Third jet p_T and $ y $ differential cross section	68
3-39	Fourth jet p_T and $ y $ differential cross section	69
3-40	Fifth jet p_T and $ y $ differential cross section	70
3-41	Total hadronic p_T differential cross section	71
3-42	Dijet mass differential cross section	72
3-43	Jet multiplicity differential cross section	73
3-44	Z p_T and $ y $ differential cross section	74
3-45	Δy leading and sub-leading jet	75
3-46	Δy Z and leading jet ($\geq 1jet$)	76
3-47	Δy Z and leading jet ($\geq 2jet$)	77
3-48	Δy Z and sub-leading jet ($\geq 2jet$)	78
3-49	Δy Z and di-jet ($\geq 2jet$)	79
3-50	$\Delta\phi$ Z and leading jet	80
3-51	$\Delta\phi$ Z and sub-leading jet	81
3-52	$\Delta\phi$ Z and third jet	82
3-53	$\Delta\phi$ leading and sub-leading jet	83
3-54	$\Delta\phi$ leading and third jet	84
3-55	Double differential as a function of leading jet p_T and $ y $	85
3-56	Double differential as a function of leading jet $ y $ and Z $ y $	86
3-57	Double differential as a function of Z p_T and $ y $	87

List of Abbreviations

AMC	Advanced Mezzanine Card
CMS	Compact Muon Solenoid
JHEP	Journal of High Energy Physics
JINST	Journal of Instrumentation
uTCA	micro Telecommunications Architecture
QIE HPD SiPM		

Chapter 1

Introduction

Modeling the strong interaction, described by quantum chromodynamics (QCD), has been a long standing challenge in the field of particle physics. The challenge arises from the failure of perturbation theory at low energies where the coupling constant becomes large. To add to this the particles taking part in this interaction, the quarks, cannot be probed directly and are measured only through the resulting hadrons.

To probe the strong dynamics we take advantage of the worlds highest energy p-p collisions at the Large Hadron Collider where the strong force dominates. The p-p collisions are inherently noisy due to multiple scatterings so purely strong interactions are difficult to reconstruct accurately. Instead, we employ a tagging particle, the Z boson, to select events and use as a well understood recoil object to the outgoing hadrons. With this technique we can examine non-perturbative QCD as well as TeV scale dynamics in a low background environment.

Chapter 2

The CMS Hadron Calorimeter

2.1 Hardware

2.1.1 The CMS Hadron Calorimeter

Calorimeter Design

The Hadron Calorimeter (HCal) of the CMS detector is a Shashlik sampling calorimeter designed to measure the energy of outgoing hadrons. It uses heavy absorber layers to induce showers from incident hadrons and thin layers of active material to sample the energy of the shower. The shower from an incident hadron typically has two distinct components; the hadronic and electromagnetic components. The hadronic component consists of nuclear recoils in the absorber and has a radiation length of $\sim 15\text{cm}$. From these nuclear interaction π^0 particles are created and quickly decay to two photons that initiate the electromagnetic component of the shower. The photons undergo pair production when interacting with a nucleus and produce electron-positron pairs that then produce more photons from bremsstrahlung. This process cascades into a shower and has a characteristic length of $\sim \text{cm}$; much shorter than the hadronic component. Any charged particles such as $\pi^{+,-}$ or protons in the hadronic component or electrons in the EM component will excite electrons in the active material to the conduction band. The excited electrons relax through doping sites in the plastic and emit light in the blue spectrum. The light is extracted from the active material through a plastic wavelength shifting fiber wrapped once around the active layer. The wavelength is shifted to green to match the greatest conversion efficiency of the photo detectors. Clear fibers are spliced to the wavelength shifting fiber to

bring the light to the photo detector mounted on the outside of CMS.

The geometry of the detector shown in Figure 2.1 is optimized to capture a full shower along the radial direction out from the interaction point. The detector is broken up into many trapezoidal towers that each point towards the interaction point and the light from a single tower is integrated with approximately three photo-detectors. This geometry gives the energy of a shower deposited in one eta-phi area and approximately three radial sections. Each tower includes ~ 17 layers of 5cm of brass and 1cm of scintillator. One tower will produce ~ 20 photo-electrons from a minimum ionizing particle (MIP) passing completely through a tower and achieves the sensitivity design goal of 3σ certainty of a MIP traversing a tower.

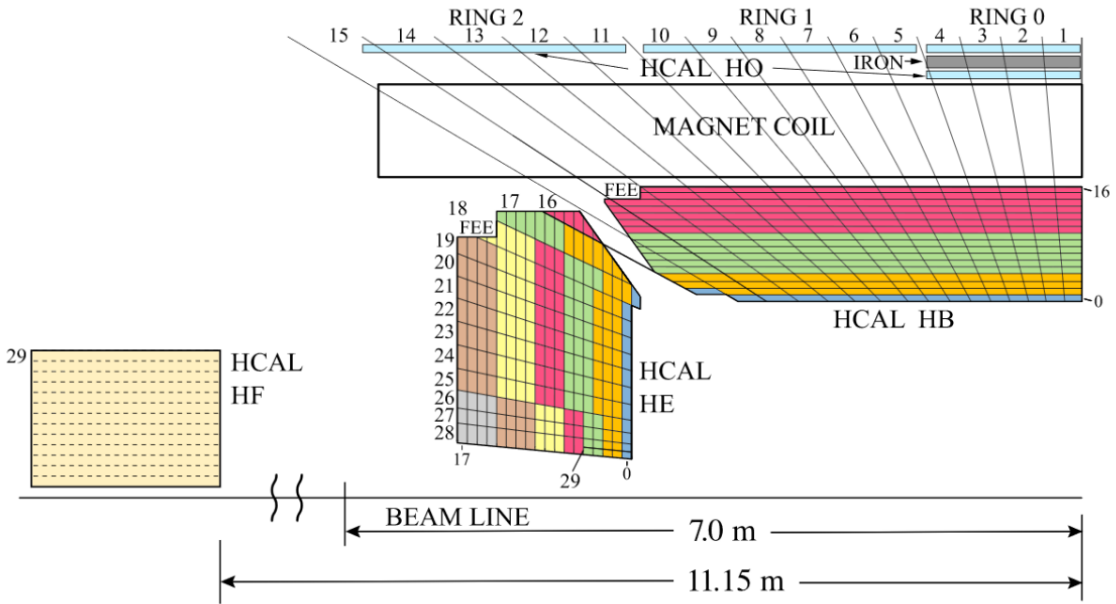


Figure 2.1: HCAL layout and segmentation in the Z - η plane.

The Hcal Barrel (HB) covers the range of $0 < |\eta| < 1.3$ and the full circle in ϕ . The barrel is broken up into 36 wedges in the ϕ direction and 17 in the η direction. The Hcal Endcap (HE) covers $1.3 < |\eta| < 3.0$ with 36 wedges in ϕ . HB and HE are joined such that each covers the projection to the interaction point with a fraction of

a tower. HO has two layers of scintillator with one placed outside the solenoid and another after the first steel return yoke. These two layers cover out to $|\eta| = 1.26$.

The Hcal forward calorimeters (HF) are located 11.2 m from the interaction point and extend the pseudorapidity coverage overlapping with the endcap from $|\eta| = 2.9$ down to $|\eta| = 5$. HF uses a steel absorber with embedded quartz fibers as the active media. The light is generated from Cherenkov radiation and not scintillation light.

Two forms of active calibration are built into the Hcal system: sourcing and laser. The sourcing is done through a series of tubes where a radioactive source (typically Cobalt-60) attached to a wire can be moved around Hcal. The laser system consists of another set of quartz fibers to inject laser light directly into the scintillating tiles. These two calibration methods will allow for monitoring of gain of the photo-detectors and light yield of the scintillator tiles.

Signal Transducers

Three technologies for converting light into electrical signals are used within Hcal for Run II. HB remained original with Hybrid Photo Diodes (HPD), HE was upgraded with Silicon Photo Multipliers (SiPM), and HF used Photomultiplier tubes (PMT).

The HPD uses the photoelectric effect to convert the incident light into electrons much like a PMT. The electrons are then accelerated with a strong field ($\sim 10^6 V/m$) in a vacuum into doped silicon. The resulting shower from electron drifts to the back of the doped silicon to an array of avalanche diodes that produce the final signal. Fibers from each towers are bundled together (18 fibers) and mounted to a single HPD so that an HPD output is the sum of a full tower. The gain of the HPD at ~ 7000 is 2000 and produces $\sim 0.3fC$ /photoelectron or $\sim 0.2GeV/fC$.

Silicon photomultipliers (SiPM) are used in HO (During LS1 HO was converted to full SiPM readout) for all of Run II and will be installed in HE during the EYETS. SiPMs have a gain of $\sim 10^6$ vs the HPD gain of $\sim 10^3$ and do not require high

voltage. The higher gain will improve MIP detection and compensate for the lower light yield from the scintillator from radiation damage. The QIE chips can relax the lowest sensitivity from 1fC to as much as 3fC and still retain a 10:1 signal ratio for a single photon incident on the array (1 pixel on). The gains of the HF PMTs and SiPMs are the same order of magnitude so the newer QIE10 chips can have the same integration range for both positive and negative inputs. Some of the signal can be used to measure the pulse length through a TDC on the QIE10. More sensitivity as well as a smaller form factor for the device is allowing for more segmentation in the towers by a factor of ~ 3 after EYETS.

Each SiPM consists of a pixel array of Avalanche photodiodes operating in Geiger mode. This means any carriers produced in the depletion region of the photodiode will create an avalanche and turn the pixel completely "on". Summing the current essentially counts the number of pixels turned on and is proportional to amount of light incident on the device. Since the number of pixels is finite there can be saturation for higher amounts of light incident on the device. This creates a non-linear scale for converting current to incident photons that must be taken into account when calculating energy deposits.

Geiger mode is achieved by bringing the reverse bias higher than the break down voltage of the silicon. One carrier created by a photon can cause a full avalanche and allow the device to detect single photons. The current generated by the break down will continue until the reverse bias voltage is brought below the breakdown voltage. This is usually done passively using a resistor in parallel with the diode. Once the current dissipates the voltage will return to above the breakdown voltage and the pixel will be reset. The full recovery time is 5-10ns which is fast enough to have a single pixel fire twice from a single scintillator signal. Since the SiPM consists of a large number of pixels many scintillator tiles can be paired with a single SiPM. Fibers

are bundled like for the HPD case and mounted to a single SiPM.

Since HF is out of the strong magnetic field it uses conventional photomultiplier tubes (PMT). Quartz fibers embedded in the steel produce blue Cherenkov light that is routed to PMTs with fibers. The Cherenkov radiation produces few photons compared to scintillation so the PMT gain is much higher at 15-75 fC/photoelectron.

Digitization and Readout

All of the transducer signals are digitized using an ASIC called the Charge (Q) Integrator Encoder (QIE). The main challenges for the QIE are to achieve accurate integration over each 25ns bunch crossing with no deadtime and maintain a dynamic range of $\sim 10^4$. The 25ns integration is achieved by using four integrators in each QIE chip that cycle every bunch crossing. Once 25ns has elapsed one of the integrators sends the amount of charge to the next stage and the second integrator takes over for the next 25ns. The cycle of the full chip to produce the final digital signal is done in four stages: integrate, compare, digitize, reset and takes 100ns.

The range of the QIE with a transducer gain of 15-75 fC/photoelectron is ~ 2.7 fC - 27pC. In HF this corresponds to an energy of a few GeV to 5TeV.

2.1.2 Back End Electronics

Introduction to FPGA's

A Field Programmable Gate Array (FPGA) is a large array of logic blocks ($\sim 10^6$) that can be re-wired through an electrical signal. The ability to re-wire enables an FPGA to change its output without changing the hardware. This provides a useful platform for projects that rely on frequent modification and bug fixes. Designing an ASIC to do the same job typically results in faster and more efficient computation, but future changes need to be planned for or would be impossible without building an entirely new ASIC.

The wiring of an FPGA is determined typically from proprietary software provided by the chip manufacture (called place-and-route software) using the logic code of the user. The logic code is written in a Hardware Descriptor Language (HDL) and a network of logic operations and interconnects called a netlist is created that replicates the HDL. The place-and-route then creates a file to load on the FPGA that will replicate the netlist using the specific FPGA.

Since the logic operations are evaluated on a clock cycle the timing of the system is predictable and operations can be performed in parallel. Modern FPGA have millions of logic block and operations can be done in parallel.

Micro Telecommunications Crate

The micro-Telecommunications (μTCA) crate is defined by a backplane that houses 12 Advanced Mezzanine Card (AMCs), two uTCA Carrier Hubs (MCH), and two power modules. For use in CMS the second MCH is replaced by a custom control card called the AMC13 for clock, timing distribution, and data acquisition. The 12 AMCs are subdetector specific and can provide 1.0Gb/s to 3.125 Gb/s of throughput, fast data processing, and slow control for downstream components. The stock MCH provides Gigabit Ethernet (GbE) for remote control of all cards in the crate. All AMCs are connected to the MCH slots through the backplane through bidirectional links to send data and control signals.

Hcal Trigger and Readout Card

The Hcal Trigger and Readout Card is an AMC designed for Hcal to calculate trigger information for the regional calorimeter trigger (RCT) and store the event fragments until the trigger signal is returned through the AMC13. The trigger primitives (TPs) are calculated by converting the ADC values from the front end to E_T and sum portions of the detector. The conversion of ADC to energy is done by a look-up table

(LUT) stored on the uHTR that contains the conversion for all ADC values for the specific region of Hcal. Each tower of Hcal is then summed before being sent to the RCT. In parallel to the TPs, the raw event fragments are stored in a first in first out (FIFO) pipeline of length $\sim 3.2\mu s$. If the event passes the RCT and then the global trigger an L1A is propagated through the AMC13 to the uHTR and the corresponding event fragment is sent to the AMC13. If no trigger is received the event fragment is overwritten after the $\sim 3.2\mu s$.

AMC13

Since its installation in 2015 the AMC13 has been in operation for many subsystems within the Compact Muon Solenoid detector at CERN. Along with other components of CMS the AMC13 has undergone several developments throughout its run to the hardware configuration and firmware. For the 2016-2018 run there are 15-20 AMC13 cards installed in the Hadronic Calorimeter (HCal) subsystem running continuously for 9 months. The card provides clock and timing commands to the HCal daq electronics, builds events from AMC modules, and propagates the status of the μTCA crate upstream. The card must handle the 40MHz signal rate set by the Large Hadron Collider and process events at 100KHz with event sizes ranging from 4-5kB depending on run conditions and Hcal region. Details of performance at various bandwidths will be reported and compared with the luminosity to predict future run conditions

For HCal each AMC13 is located in a micro controller hub slot of a micro telecommunications architecture crate. Hcal trigger and readout cards fill the 12 AMC slots of the crate and provide the event fragments to the AMC13. Additional uTCA crates with an AMC13 were needed for the 2017 run to provide clock and configuration commands to on-detector electronics but not event building. A general diagram of connectivity is shown in figure 1.

The first job of the AMC13 is to keep the back-end in sync by distributing the

LHC clock and other timing controls received by the TCDS (TCDS) subsystem. The clock and control signals are sent from TCDS encoded over a single data stream. The AMC13 separates the clock and control signal and distributes them over the backplane to the uHTRs. The clock frequency is measured and reported to the software for monitoring. The number of various control signals are saved in memory on the board along with a short history of the signals last received. The signal number counting is crucial for debugging where for example the number of L1As received by the AMC13 can be compared to the uHTR to see where an event was lost.

Table 2.1: AMC13 Trigger Throttling State

Value	State	Comment
0000	Disconnected	Hardware failure or broken cable
0001	Warning overflow	Imminent buffer overflow
0010	Out of sync	FED is no more synchronized with the TTC values
0100	Busy	Cannot accept triggers
1000	Ready	Ready to accept triggers
1100	Error	Any other state that prevents correct functioning
1111	Disconnected	Hardware failure or broken cable

The trigger throttling system (TTS) relays electronic state information to the trigger system in order to throttle the rate of L1As. The AMC13 has 6 state levels given by table 2.1. If the state of the AMC13 is in Overflow the queue of events for the event builder is 90% full and the rate of triggers is throttled 0 until the queue is cleared to 10%. Currently the uTCA system can reach a maximum of $\sim 250\text{kHz}$ rate of events with $\sim 4\text{kB}$ event size. Since the Level-1 trigger has a maximum rate of 100kHz the uTCA system rarely throttles L1As during collisions. Any throttling is a sign of an error with the uTCA electronics since they are not operating at full capability. After the EYETS HCAL will undergo extensive remapping and layers will be read out more finely in the radial direction. This means for more channels and hence larger event sizes for the uTCA system. During this time the AMC13 will be running closer to its design capabilities and the TTS system will be more vital to monitor.

In order to determine if an AMC13 performs adequately for data taking during collisions a stress test is conducted with each card. The stress test consists of checking basic functionality and running at high rates for an extended period of time. The basic functions include reading and writing to all registers, loading different FPGA firmware, and power cycling. The card is then ran at the maximum rate of $\sim 250\text{kHz}$ determined by the uHTRs and other crates in the system. The AMC13 fails if any errors are seen with the card in ~ 10 minutes of running and yields a failure rate of approximately 10%. The high rate in the stress test provides a fast way to determine the most capable and reliable cards. Since the failed cards passed the previous tests they can be used for non-critical systems such as test stands.

The uTCA crates contain 12 uHTR boards with each storing data in the pipeline from the Front End QIE cards. The uHTRs send an event from the pipeline once an L1A is received from the AMC13. The AMC13 combines the data from all 12 uHTRs into a single event fragment. This includes generating header and trailer information to identify the particular event (Event, bunch crossing, orbit numbers), to provide a full event size and information about the individual uHTR data, and to calculate a cyclic redundancy check on the data to catch data corruption errors later on. All of these functions are done on the Tier 1 board which houses the Kintex FPGA. The event is sent through the multi-mode fibers to a FEROL in the central data acquisition system. Currently the AMC13 has three multi-mode output modules on the front of the board and utilizes two for sending data out.

During the development of the AMC13 and when firmware updates are implemented the simplest validation is to check the output data format. This can quickly spot timing issues by looking at the event/orbit/bunch numbers, missing fragments, and data corruption. The simplified data format is shown in figure 2.2. The header of an AMC13 event contains most relevant event timing information such as the event,

orbit, and bunch number. The header also provides a way to navigate the AMC13 event with the total number of 64-bit numbers (or words). A parser was written to automate this process and provide the error checks that can be used to validate new firmware.

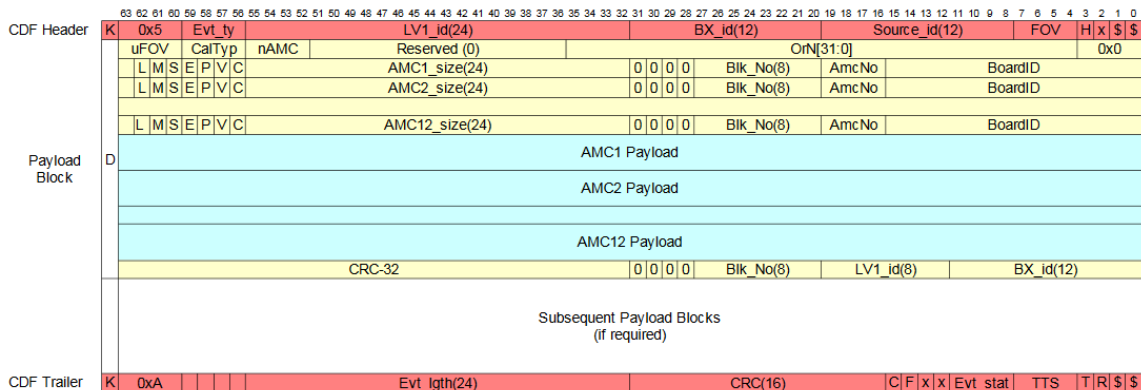


Figure 2-2: AMC13 data format with detailed header information.

During the beginning of run 2 in 2015 Hcal required a data throughput of less than 10Gbps for each uTCA crate. This meant the AMC13 could utilize one of its three fiber outputs to handle the data flow to cDAQ. The Hcal front ends were scheduled for upgrades after the 2016 run and all data rates were planned to increase. To handle this data rate the AMC13 would need to split the data flow through two fiber outputs. We decided to commission this functionality with the AMC13 during 2016 collisions. The data splitting functionality was built into the AMC13 firmware and could be enabled with register setting on the card. No software support for reading out the split data from the flash event buffer was developed before 2016.

AMC13 Software

The AMC13 firmware run settings are determined by values in a built in flash memory. The memory can be read and written to by the CERN IPBus protocol and the AMC13 software is a high level wrapper of the IPBus software. IPBus is a hardware control

protocol wrapped by IP for reliable remote control of the system.

To communicate with the AMC13 two C++ classes are used called AMC13Simple and AMC13. The class AMC13Simple provides base commands that can read from and write to registers on the flash memory. The FPGAs read the memory from their respective Tiers and carry out commands based on the values. For example, if a 1 is written to the first bit of memory on the Tier 1 board the card will enter Run mode where it will build events when it receives a Level 1 accept. These write commands are the building blocks of all of the controls to the AMC13.

The AMC13 class inherits all of the AMC13Simple commands and uses them to form user friendly and complex functionality. This allows for applications to the class and utilize the AMC13 board without having to know memory registers or the proper order of read and writes.

The last element of the supported software is the AMC13 Tool. The tool is an interactive session that allows for AMC13 class commands to be called directly on the command line. This type of communication is not used for collision operations but instead is a key method for testing and debugging. The tool can generate arbitrary events with varying trigger rate, active AMCs, AMC event sizes, trigger rules and more. Around 1k events can be saved in the AMC13 event buffer and saved for processing using the event parsing software. During collisions the command line tool can still be used to monitor specific cards that may be exhibiting issues.

There are two components to the AMC13 online monitoring: Event builder and Event buffer. The event builder component refers to comparing key elements of the data as they are being built by the event builder. The three most important checks are the event, orbit, and bunch crossing numbers, but many more checks are available and can be made available through firmware. The AMC13 is able to look into the payloads received from the uHTRs and compare the values to its own internal counters. If the

values do not match a register specific to the value being checked is incremented. This is done for every event accepted by the level one trigger and can be reliably measured down to one mismatch in millions of events. During collisions operations with 100kHz L1A rate we expect the number of false positives to be less than 1 in 5×10^{10} events (or 12 hours of collisions for one fill). Since the mismatch counter is so reliable it is an immediate sign of a problem with the detector.

The second component of online monitoring, referred to here as the event buffer, has not yet been implemented to its full potential. Currently it is used during local runs as a convenient tool to record data since the usual DAQ subsystem is not used for local runs. For future runs the event buffer monitoring will offer online data quality monitoring of full Hcal events even during collisions. Hcal currently monitors only the trigger primitives online and the full event offline, so the event buffer would fill the gap online where we are blind and potentially save data. The event builder on the AMC13 works by writing the event from a crate to on board memory with a size limit of 1k events. The data can be read off at any point during running without effecting the crate or AMC13. Once the events are read data is overwritten using a simple fifo and the rate at which events are read is limited only to the read rate at 100Hz. The events can then undergo the standard set of data quality check done for the data offline. Even with low statistics from the limited read rate

To bridge the gap between an error in the event builder system and Hcal operations is the alarmer. The program reads a table that is updated with all of the error register values on the uTCA cards and takes action corresponding to the errors. The action usually consists of changing the status of HCal from "Running" to "Running Degraded" as well as sending out emails and text alerts with information about the errors. That status can immediately be seen in the control room and signals the shifter to call the HCal on-call to get instructions on what actions to take.

Chapter 3

Z+jets Measurement

3.1 Analysis

3.2 Data and Simulated Samples

The data sample analyzed corresponds to an integrated luminosity of 35.9 fb^{-1} , collected in 13 TeV proton-proton collisions with the CMS detector during the 2016 data taking period including eras B-H. Candidate events are selected online using single-lepton triggers listed in 3.1, which require at least one isolated electron (muon) with $p_T > 25(24) \text{ GeV}$ and $|\eta^l| < 2.4$. The total trigger efficiency for events within the acceptance of this analysis is greater than 90%. The analysis uses CMSSW version CMSSW_9_4_10 with the miniAOD v3 format. The Global tag and JSON file are listed in below.

Data Global Tags:

- Era B-H: 94X_dataRun2_v10

MC Global Tag:

- 94X_mcRun2_asymptotic_v3

JSON File:

- Cert_271036-284044_13TeV_23Sep2016ReReco_Collisions16_JSON.txt

Simulated events for both signal and background are produced using various Monte Carlo (MC) event generators, with the CMS detector response modeled using the

Table 3.1: Single Muon Data Samples

Sample
/SingleMuon/Run2016B-17Jul2018_ver2-v1/MINIAOD
/SingleMuon/Run2016C-17Jul2018-v1/MINIAOD
/SingleMuon/Run2016D-17Jul2018-v1/MINIAOD
/SingleMuon/Run2016E-17Jul2018-v1/MINIAOD
/SingleMuon/Run2016F-17Jul2018-v1/MINIAOD
/SingleMuon/Run2016G-17Jul2018-v1/MINIAOD
/SingleMuon/Run2016H-17Jul2018-v1/MINIAOD

GEANT4 [Allison et al. (2006)] program. These events are then reconstructed using the same algorithms as used to reconstruct collision data and are normalized to the integrated luminosity of the data sample using their respective cross sections. For the simulation of the signal, we use a sample generated with MADGRAPH5_AMC@NLO (shortened as MG5_AMC) [Alwall et al. (2008)] using the FxFx merging scheme [Frederix and Frixione (2012)]. Parton shower and hadronization are simulated with PYTHIA8 [Sjöstrand et al. (2008)] using the CUETP8M1 tune [CMS Collaboration (2015)]. The matrix elements include $Z + 0,1,2$ jets at NLO, giving a LO accuracy for $Z + 3$ jets.

To gain more statistics for the signal MC a binned sample in the number of outgoing partons (npNLO) at matrix element level is used. All signal samples and their corresponding event number, effective event number, and cross section are shown in 3.2. The number of effective events is the sum of event weights before any offline selections. The effective number of events is smaller than the number of events because the event weight can be negative. This is a result of the the method used by Madgraph to compute NLO accurate cross sections. To combine the inclusive and binned samples a weight as a function of npNLO is derived from the formula:

$$weight(npNLO) = \frac{(Int.Lumi) * \sigma_{npNLO}}{NEffectiveEvents} \quad (3.1)$$

The weights as a function of npNLO is given in table 3.3. The weighted inclusive

and binned sample npNLO distributions match to within 1%. After the final event selections the inclusive and binned jet multiplicity have a trend in the ratio as seen in figure 3-1. This difference is caused by different madgraph versions being used in the production of the samples: version (2.2 and 2.3). The difference is covered by statistical uncertainty alone without considering systematics.

Table 3.2: Monte Carlo Signal Samples

MC	Events	Eff. Events	XSec (pb)
DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	122055388	81781064	5931.9
DYToLL_0J_13TeV-amcatnloFXFX-pythia8	93832853	76690000	4620.52
DYToLL_1J_13TeV-amcatnloFXFX-pythia8	91500283	41572416	859.59
DYToLL_2J_13TeV-amcatnloFXFX-pythia8	90299356	26282782	338.26

Table 3.3: DY npNLO Weights

npNLO	Weight
0	1.15
1	0.57
2	0.38

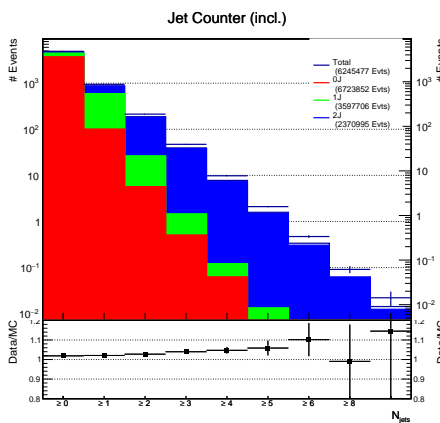


Figure 3-1: Jet multiplicity for the inclusive and binned samples. The binned samples are stacked. Normalized to total cross section of the samples (5931.9fb for inclusive, 5818.37fb for binned).

The production of $Z(\rightarrow \ell^+\ell^-) + \text{jets}$ can be mimicked by various background sources: ($t\bar{t}$) events, single top, dibosons (WW, WZ, ZZ), tri-bosons (ZZZ, WWZ, WZZ), and W bosons produced in association with jets, as well as Z + jets events in which

the Z boson decays into $Z \rightarrow \tau^+ \tau^-$. Background processes are split into two components: the resonant and nonresonant background. The irreducible resonant background comes from events with a real Z boson in the final state (WZ, ZZ, tribosons, etc.) and it is estimated using MC samples. The nonresonant background comes from events which do not have a Z boson in the final state ($t\bar{t}$) and is estimated using events with both an electron and muon. The non-resonant calculation is discussed further in sec. 3.4. $Z \rightarrow \tau^+ \tau^-$ events are considered background and are estimated using the MG5_AMC signal sample.

Background samples corresponding to diboson electroweak production [Campbell et al. (2011)] are generated at NLO with POWHEG [Nason (2004); Frixione et al. (2007a); Alioli et al. (2010); Frixione et al. (2007b)] interfaced to PYTHIA 8 or MADGRAPH5_AMC@NLO interfaced to PYTHIA 8. The background sample corresponding from tribosons are generated at NLO using MADGRAPH5_AMC@NLO interfaced with PYTHIA 8. Multi-boson samples with their corresponding number of events and cross section are shown in 3.4.

As a cross-check for the non-resonant control region technique we include MC samples that estimate the same event. The list of samples is shown in 3.5. The closure tests are shown and discussed more in sec. 3.4.

Table 3.4: Monte Carlo Resonant Samples

MC	Events	XSec (pb)
ZZTo2L2Nu_13TeV_powheg_pythia8		
ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8		
ZZTo4L_13TeV_powheg_pythia8		
WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8		
WZTo3LNu_TuneCUETP8M1_13TeV_powheg_pythia8		
WWZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8		
WZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8		
ZZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8		

Table 3.5: Monte Carlo Non-Resonant Samples

MC	Events	XSec (pb)
ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	1000000	3.35
ST_t-channel_antitop_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	38811017	136.0
ST_t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	67240808	80.95
ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	6952830	35.6
ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	6933094	35.6
TT_TuneCUETP8M2T4_13TeV-powheg-pythia8	77081156	831.7
WJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8		61526
WWTo2L2Nu_13TeV-powheg		12.21

Rochester Correction

To correct misalignment of the CMS detector in both data and MC the Rochester corrections are applied. This provides corrections as a function of muon charge, p_T , and η .

Muon Scale Factors

To calculate the cross section of Z boson production we must approximate the efficiency of the detector and make up for missed Z decays. Within the fiducial region of CMS there are many ways to miss a muon including inactive material, measurement noise, and limits on fitting algorithms. Many of these efficiencies could be increased to improve acceptance but would come at the cost of mistaking other objects as muons. The efficiencies of each step in the measurement process can be approximated by the tag and probe method. Tag and probe uses a high quality muon (Table 1) as a "tag" and picks other muons, called "probes", in the same event that result in an invariant mass of $\pm 20\text{GeV}$ from the Z mass.

This sample of probes is then put through any selection criteria being measured including reconstruction, identification, isolation, and triggers. The ratio of probes that pass the selection to the total number of probes is estimated to be the efficiency. This method of measuring the efficiency assumes muons only from Z boson decays so a background subtraction is applied before the final count of passing and total probes is measured. The total passing and failing mass spectra are fit using a convolution of a function that looks like the signal (typically bright-wigner) and one that looks like the background (e.g. exponential) as shown in Figure 3.2. The area of the signal function alone is then the number of muons from Z decays. The efficiencies are calculated as a function of muon p_T and η using the same binning scheme as the main analysis. At the reconstruction level the MC is weighted so the efficiency

matches data by use of scale factors. The scale factors are simply the ratio of the data and MC efficiency for each pt, eta bin. Once the data is unfolded it is divided by the efficiency to produce a cross section measurement.

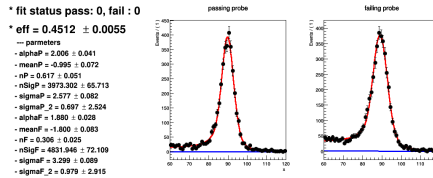


Figure 3-2: Example of Tag and Probe fit for passing and failing probes.

- Muon Medium ID SF: p_T , η bins; provided by Muon POG
- Muon Isolation SF: p_T , η bins; provided by Muon POG
- Single Muon Trigger SF: To be calculated

Pileup

Pileup scale factors are applied to the MC events to account for the difference in pileup profiles between data and MC. This is done with the data pileup estimated from Brilcalc with eras B-H and the MC pileup profile used to generate the events. A value of 69.2mb is used for the minimum bias cross section in the calculation. As a check on the pile up reweighting figure 3-3 shows the comparison of the number of vertices in data and mc. The minimum bias is then varied with its 4.6% uncertainty to obtain a different data pileup profile and compared again to mc in figure 3-4.

The number of vertices without pile re-weighting is shown in figure 3-5. To look at the effect of highly ionizing particles (HIP) in the tracker Figure 3-6 shown the number of vertices in eras B-F and G-H.

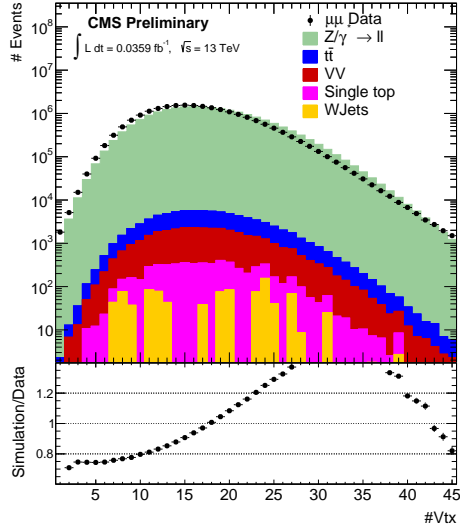


Figure 3-3: Number of vertices in data and MC. Minimum bias of 69.2mb is used for the data profile.

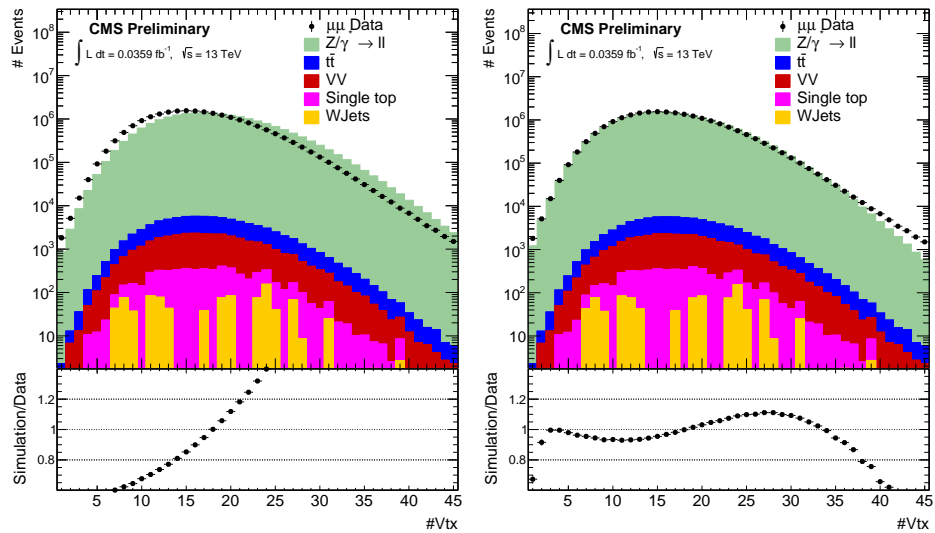


Figure 3-4: Number of vertices in data and MC with minimum bias variations. Minimum Bias of 69.2mb varied 4.6% up (left) and down (right).

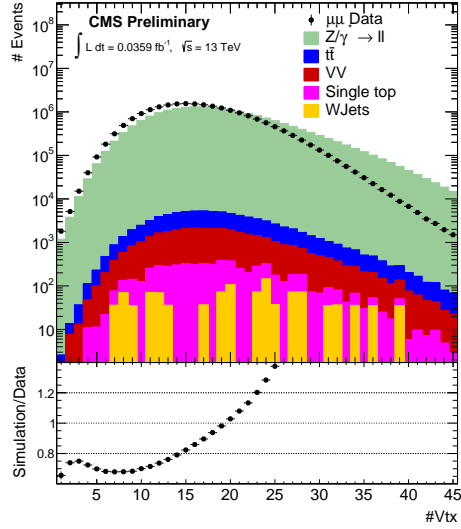


Figure 3-5: Number of vertices in data and MC without re-weighting the MC.

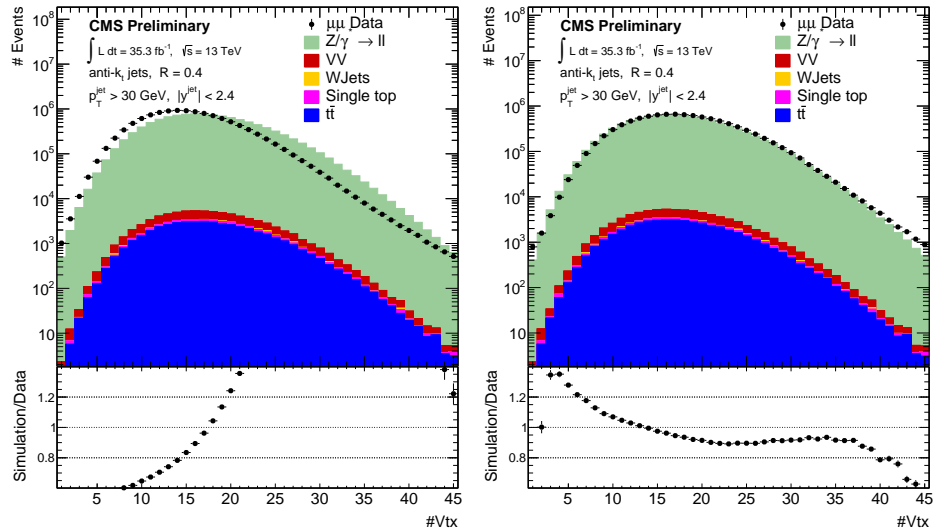


Figure 3-6: Number of vertices in era B-F (Left) and G-H (Right). Minimum bias of 69.2mb is used to calculate the profile.

L1 Prefire Weights

Due to a timing drift of the ECAL with respect to the calorimeter trigger primitives (TP), high eta TPs could be assigned to the previous bunch crossing BX-1(negative 1) instead of BX0. A TP corresponding to the BX0 would then not be able to generate a level 1 accept (L1A) because of the CMS trigger rules. The BX-1 TP event will most likely be a minimum bias event and get rejected at the high level trigger. These missing events are not simulated in the MC and the MC must be corrected using weights.

The event weights are calculated using the p_T and η of all photons and jets in an event and generate a probability for the event to be lost due to prefiring. The weight is applied to each MC event in the analysis. For the full analysis selection the average of the prefiring weights is 0.98.

3.3 Event Reconstruction, Object Selection and Corrections

The global event reconstruction (also called particle-flow event reconstruction [Sirunyan et al. (2017)]) aims to reconstruct and identify each individual particle in an event, with an optimized combination of all sub-detector information. In this process, the identification of the particle type (photon, electron, muon, charged hadron, neutral hadron) plays an important role in the determination of the particle direction and energy.

The reconstructed vertex with the largest value of summed physics-object p_T^2 is taken to be the primary pp interaction vertex. The physics objects are the jets, clustered using the jet finding algorithm [Cacciari et al. (2008, 2012)] with the tracks assigned to the vertex as inputs, and the associated missing transverse momentum, taken as the negative vector sum of the p_T of those jets.

Electron candidates are reconstructed by combining the information from the

ECAL and from the silicon tracker. The energy of electrons is determined from a combination of the electron momentum at the primary interaction vertex as determined by the tracker, the energy of the corresponding ECAL cluster, and the energy sum of all bremsstrahlung photons spatially compatible with originating from the electron track. The Super Cluster (SC) reconstruction efficiency for $E_T^{SC} > 5$ GeV is close to 100%. The momentum resolution for electrons with $p_T \approx 45$ GeV from $Z(\rightarrow e^+e^-)$ decays ranges from 1.7% to 4.5%. It is generally better in the barrel region than in the endcaps, and depends on the bremsstrahlung energy emitted by the electron, as it traverses the material in front of the ECAL [Khachatryan et al. (2015)]. To reduce the electron misidentification rate, electron candidates are subjected to additional identification criteria, which are based on the distribution of the electromagnetic shower in the ECAL, a matching of the trajectory of an electron track with the cluster in the ECAL, and its consistency with originating from the selected primary vertex.

Muon candidates are reconstructed with a global fit using both the inner tracking system and the muon spectrometer [Chatrchyan et al. (2012)]. The momentum of the muons is obtained from the curvature of the corresponding track. Muons are selected as Z decay product candidates from the particle flow (PF) objects. The efficiency of the muon reconstruction algorithm is higher than 0.95% for muons with $p_T > 20$ GeV. The relative p_T resolution at 100 GeV is 2% in the barrel and 6% in the endcap and increases to 10% in the endcap at 1 TeV.

Jets are formed from the particles reconstructed by the PF algorithm, using the FAST-JET software package and the anti- k_T jet clustering algorithm [Cacciari et al. (2008)] with a distance parameter of 0.4. The jet four-momentum is obtained according to the E-scheme (vector sum of the four-momenta of the constituents). The technique of charged-hadron subtraction (CHS) [Collaboration (2014)] is used to re-

duce the pileup contribution by removing charged particles that originate from pileup vertices. The jet four-momentum is corrected for the difference observed in the simulation between jets built from PF candidates and generator-level particles. The jet mass and direction are kept constant for the corrections. An offset correction is applied to jet energies to take into account the contribution from additional proton-proton interactions within the same or nearby bunch crossings. Further jet energy corrections are applied for differences between data and simulation in the pileup in zero-bias events and in the p_T balance in dijet, $Z + \text{jet}$, and $\gamma + \text{jet}$ events. To maximize the reconstruction efficiency while reducing the instrumental background and contamination from pileup jets, tight identification quality criteria are applied on jets, based on the energy fraction carried by charged and neutral hadrons, as well as charged leptons and photons. A minimum threshold of 30 GeV on the p_T of jets is required to ensure that they are well measured and to reduce the pileup contamination. Jets are required to have $|\eta| < 2.4$ and to be separated from all selected lepton candidates by at least $R \leq 0.4$.

To compare the measured distributions with the theoretical predictions, various experimental corrections are applied after subtracting the total expected background from the observed number of events in each bin. A correction for the detector resolution effects is implemented using an unfolding technique (see details in section 3.7). The event acceptance and selection efficiency are estimated using simulation and are used to correct the data. To correct for differences in efficiencies between data and simulation for lepton reconstruction, identification, isolation, and trigger, efficiency corrections ("scale factors (SFs)") are determined from the data using the tag-and-probe method (T&P) [Chatrchyan et al. (2011)].

We select events with one isolated electron (muon) with transverse momentum of at least 25 (24) GeV. After offline reconstruction, two leptons are required with

the first having $p_T > 30 \text{ GeV}$ and the second having $p_T > 20 \text{ GeV}$. We require that the two electrons (muons) with highest transverse momenta form a pair of oppositely charged leptons with an invariant mass in the range $91 \pm 20 \text{ GeV}$. Electron candidates are required to be reconstructed within $|\eta| < 2.4$, excluding the barrel-to-endcap ($1.444 < |\eta| < 1.566$) transition regions of the ECAL. Electrons and muons are considered isolated based on the scalar p_T sum of the nearby PF candidates with a distance $R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.4$. For both electrons and muons medium identification criteria is applied. To correct misalignment of the CMS detector in both data and MC for the muon channel, corrections [Bodek et al. (2012)] are applied.

Reco Muons

Global and tracker muons are selected as Z decay product candidates from the particle flow (PF) collection. For global muons tracks are constructed after fitting hits in the DTs, CSCs, and RPCs of the muon spectrometer to form standalone muons. For each standalone muon a search for a matching track in the tracker is done by a fitting algorithm similar to the algorithm for standalone muons. The efficiency of the global muon algorithm is higher than 0.95% for muons with $p_T > 20 \text{ GeV}$. To achieve high efficiency at low p_T tracker muons are used as well. Tracker muon reconstruction begins with all tracker tracks with $p_T > 0.5 \text{ GeV}$ or $p > 2 \text{ GeV}$. The tracks are then extrapolated to the muon system taking into account the interaction with the detector medium. If a hit in a DT or CSC is within 3cm or 4σ of the muon track (in the local coordinate system of the muon chamber) the track is considered a tracker muon. No fit is done for the combined tracker track and muon system hits so the tracker muon p_T is calculated from the tracker track curvature. The efficiency at low muon p_T is greater than global muons because of the low p_T seeds and the requirement of only one hit in the muon system. Further quality cuts must be applied

Tracking algorithm: Kalman Filtering [CMS Note CMS-NOTE-2006-041, 2006].

A Preliminary charged track is found as a seed. All layers of the tracker are then combined to get a full track. Fitting is then done to determine the curvature and location. Multiple iterations are done with varying requirements on the number of pixel or strip hits and the p_T of the particle. Each iteration removes any tracks that satisfy the criteria and allows for the next iteration to lower thresholds while keeping the computation time manageable. The p_T acceptance can be as low as 200MeV.

Electron tracking: similar iterative method used, all tracks from the iterative method with larger than 2GeV are electron candidates. Different fit methods are used (Gaussian-sum filter) to handle the bremsstrahlung radiation when the electron moves through the tracker. Also a BDT is used that uses GSF + KF fits and the chi-squared of the fits and distance between ECAL deposit and track direction.

Standalone muons: hits in DT or CDC used to get track seeds. The seeds go to the full track measurement using the full muon spectrometer. Tracker muons: $p_T > 0.5\text{GeV}$, $p > 2.5\text{GeV}$ and at least one hit in the muon system matches the track extrapolation. A match is made if the muon system hit is less than 3cm away from the track or less than 4 sigma from the track.

Global muons: A standalone muon that has a matching track.

The analysis cuts are as follow:

- Muon ID:
 - Medium ID
- Isolation:
 - Relative isolation < 0.15 (Tight)
 - PF based, $\Delta\beta$ -corrected, combined (charged and neutral hadrons, photons), normalized to muon p_T

- $P_t > 25\text{GeV}$
- $\eta < 2.4$

Gen Muons

Generator level muons with status 1 are used and dressed with status 1 photons within a cone size of $R = 0.1$ to account for FSR. The selections are the same as the reconstruction level muons after the dressing is done.

The particle level objects are defined to be particles with a lifetime of $>1\text{cm}$ (excluding neutrinos) and identified using the same algorithms as for the data. In addition, leptons are stable particle from Z decays, dressed by adding the momenta of all photons within $\Delta R < 0.1$ from their directions. The momenta of the leading leptons are summed to obtain the particle level Z momentum. The particle level objects are required to pass the same kinematic selections as at detector level.

3.3.1 Z Boson Selections

Only the two highest pt muons that pass the lepton selection criteria are considered for the Z boson candidate. The muons must be oppositely charged and have an invariant mass $71\text{GeV} < m_{\mu\mu} < 111\text{GeV}$. Selections are the same for both reco and gen level.

3.3.2 Jet Selections

Reco Jets

Jets are reconstructed using the anti-kT clustering algorithm[] with a cone size of $R = 0.4$. Any jets within $\Delta R = 0.4$ of the two leading leptons are discarded from the jet collection. This eliminate jets in the collection that are clustered from the

energy deposits of the lepton in the calorimeter. Cuts concerning the jet composition are applied in order to avoid mis-identification of pile-up and to increase noise rejection. This list of cuts (see table 3.6) constitutes the Loose identification criterion provided by the JetMET POG. To reduce pile-up jets an MVA discriminator is used with working points from the JetMET POG. The working points are MVA cuts as a function of jet η and the loose working point is used for this analysis. Reconstructed jets in MC have their transverse momenta smeared in order to more closely match what is observed in data. Smearing is done according to the formula given by the JetMET group. If a reco level jet is found to have a gen level jet match the reco level p_T is pushed away from the gen jet value by a factor given by:

$$c_{JER} = 1 + (s_{JER} - 1) \frac{p_T^{reco} - p_T^{gen}}{p_T^{reco}} \quad (3.2)$$

Where s_{JER} is the data-MC jet resolution scale factor dependent on the jet η and p_T , p_T^{reco} is the jet transverse momentum at reco level, p_T^{gen} is the jet transverse momentum at gen level. The final smeared p_T is the given by:

$$p_T^{smeared} = p_T * c_{JER} \quad (3.3)$$

The criteria for a reco-gen match is as follows:

$$\Delta R < \frac{R_{cone}}{2} \quad |p_T^{reco} - p_T^{gen}| < 3\sigma_{JER}p_T \quad (3.4)$$

Where ΔR is the angular separation between the reco and gen jet, R_{cone} is the cone size for the jets clustering algorithm: $R = 0.4$. p_T^{reco} and p_T^{gen} are defined previously. $\sigma_{JER}p_T$ is the relative p_T resolution in MC.

If a match is not found for the reco jet a Gaussian smearing is done which pushes the jet p_T in a random direction an amount dependent on the resolution.

$$c_{JER} = 1 + Gauss(0, \sigma_{JER}) \sqrt{\max(s_{JER}^2 - 1, 0)} \quad (3.5)$$

where all values have been defined previously. Lastly if a match is found and the scale factor s_{JER} is less than 1.0 no smearing is done. A systematic error for this technique is discussed in the systematics section.

A cut of 30 GeV on jet p_T is applied to reduce the pileup contamination as well as large uncertainty on the energy measurement. To ensure a good quality of the tracking information, jets with $|\eta| < 2.4$ are removed from the collection. Finally, the jet collection is ordered by decreasing p_T values.

Table 3.6: Loose Jet ID Parameters

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

Gen Jets

Generator level jets are clustered from stable status 1 particles using the same anti-kT algorithm used for jet reconstruction. Prior to clustering neutrinos are removed and photons from final state radiation in a cone of $R=0.4$ are added. The kinematic cuts applied are the same as the reco level selections. The jets are then ordered by p_T .

3.4 Background Estimation

Background events are split into two categories: resonant and non-resonant. The resonant background, which consists mainly of multi-boson events with at least one Z boson in the final state, are estimated using MC samples. The non-resonant events contain two leptons primarily from W decays, such as $t\bar{t}$, and are estimated from a

data-driven method. The $Z \rightarrow \tau^+ \tau^-$ is considered a background and is estimated from the MG5_AMC signal MC sample.

The data driven method for the non-resonant background uses a control region containing events with one electron and muon ($e^\pm \mu^\mp$) passing all other signal region criteria. The control region is then used to estimate the non-resonant background in the signal region by applying a conversion factor to account for cross section and lepton efficiency differences. Assuming lepton flavor symmetry, the cross section for a $e^\pm \mu^\mp$ final state and a $e^+ e^-$ or $\mu^+ \mu^-$ final state differs only by a factor of 2. The difference in efficiency between muons and electrons is estimated using the total yields of the two channels. Resonant signal and background are estimated in the control region by the same signal region MC and subtracted to avoid double counting.

Data and simulation of jet multiplicity and kinematics of the Z boson and leading jet are shown in 3.26 - 3.28. The number of background events is small compared to the signal and it amounts to approximately 1% for ≥ 0 jets and increases to 10% at 5 or more jets. For transverse momentum variables the background increases from 1% below 100 GeV to 10% in the high-pt tails.

In order to obtain cross section values for Z production background events are estimated using either MC samples or data-driven methods and subtracted prior to unfolding. The resonant background involving a Z boson in the final state are estimated using MC samples listed in 3.4. These include multi-boson production where at least one of the bosons in a Z . The non-resonant background coming dominantly from two leptonic W decays are estimated from an $e^\pm \mu^{mp}$ sample. First, an $e^\pm \mu^{mp}$ control region is created from the SingleMuon data set by inverting the flavor of one of the final state muons. For the DY and resonant MC this also means applying electron SFs for the electron ID and reconstruction and a single muon trigger SF. The DY and resonant MC are then subtracted from the data in the $e^\pm \mu^{mp}$ control

region. The remaining data distribution is multiplied by a constant dependent on the difference in cross section between $e^\pm\mu^{mp}$ and mu/mu and the difference in efficiency of an electron and muon. Since the decay rate of the W is the same for all lepton flavors the cross section for $e^\pm\mu^{mp}$ is double the cross section of $\mu\mu$. The difference in efficiencies is estimated from the total signal region yields:

$$k_{e\mu} = \frac{\sqrt{N_{\mu\mu}}}{\sqrt{N_{ee}}} = \frac{\epsilon_\mu}{\epsilon_e} \quad (3.6)$$

Combining the two we can estimate the number of non-resonant events in $\mu\mu$ from $e^\pm\mu^{mp}$:

$$N_{\mu\mu} = \frac{1}{2} N_{e\mu} k_{e\mu} \quad (3.7)$$

As a closure test for this background estimation method we use MC samples including $t\bar{t}$, WW, single top, and W+Jets. In figures 3·7-3·23 the MC samples are compared for the jet and Z kinematics and differ by 5%-10%. The MC in the jet multiplicity plot (figure 3·9) deviates from data especially at higher multiplicities. In previous analyses scale factors were applied to correct for this difference.

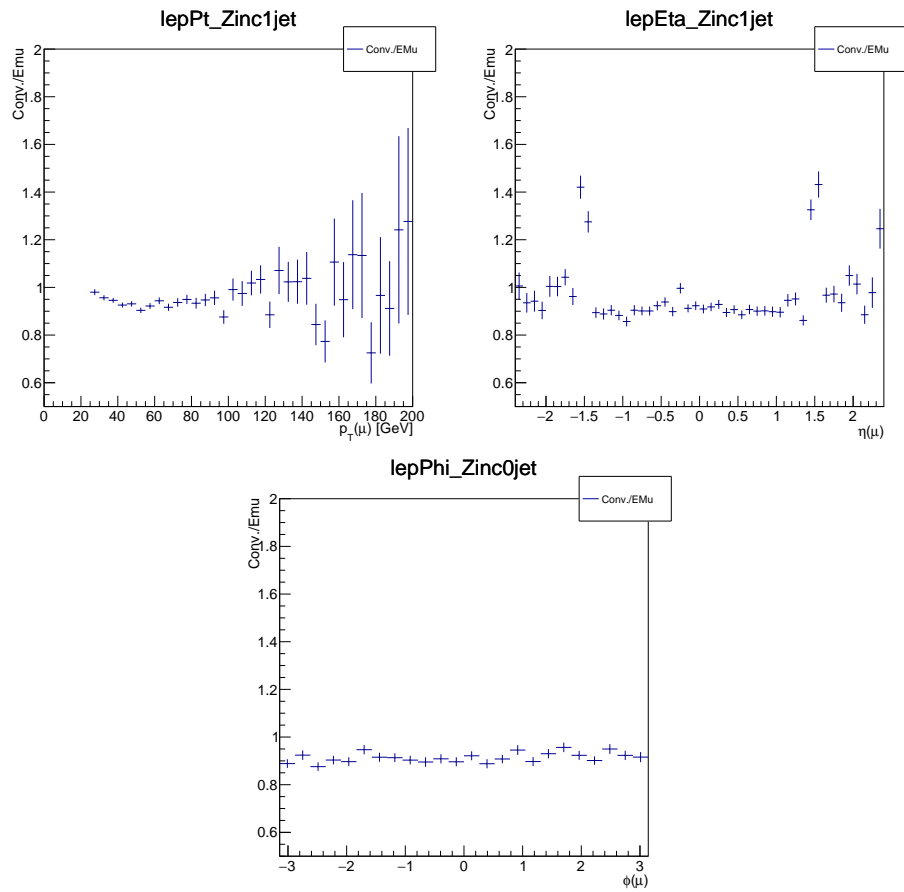


Figure 3·7: Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Lepton p_T (left), η (center), ϕ (right) with one jet inclusive.

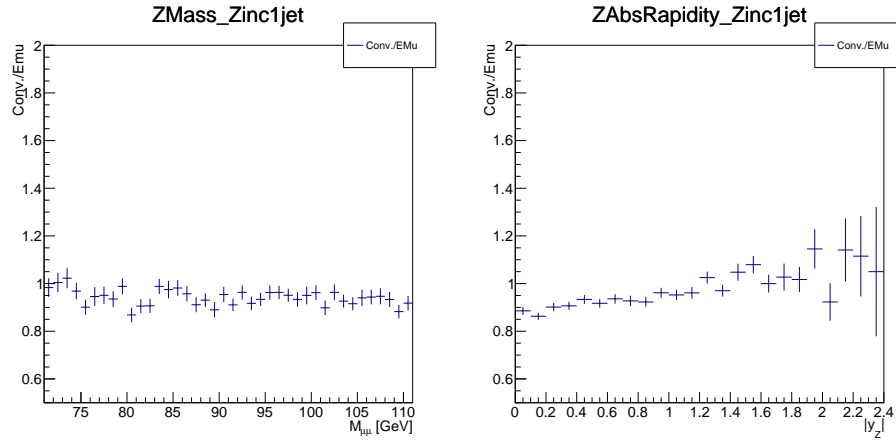


Figure 3.8: Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Dimuon mass and Z candidate $|y|$ with one jet inclusive.

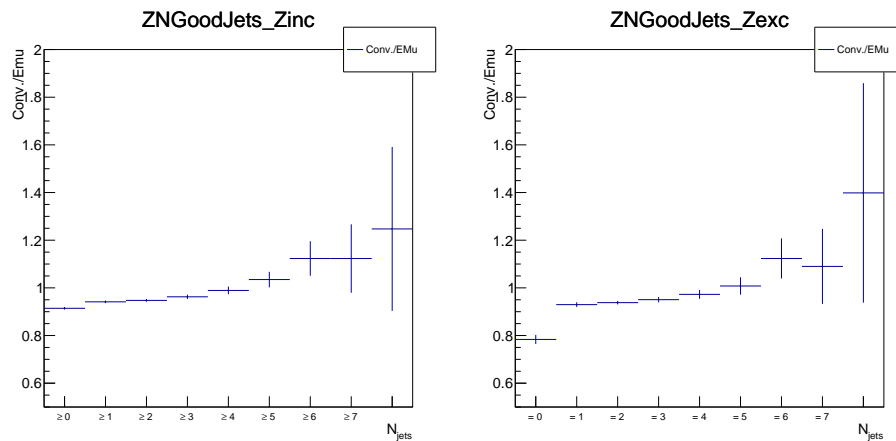


Figure 3.9: Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Jet multiplicity inclusive and exclusive.

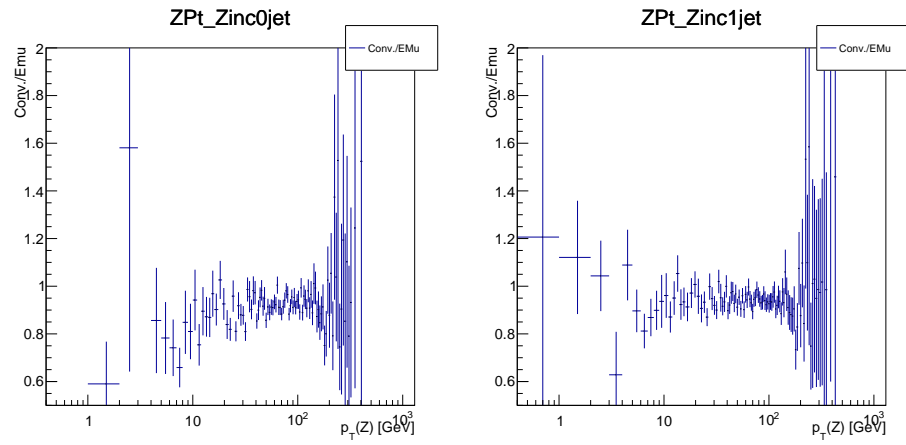


Figure 3-10: Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Z candidate p_T 0 and 1 jets inclusive.

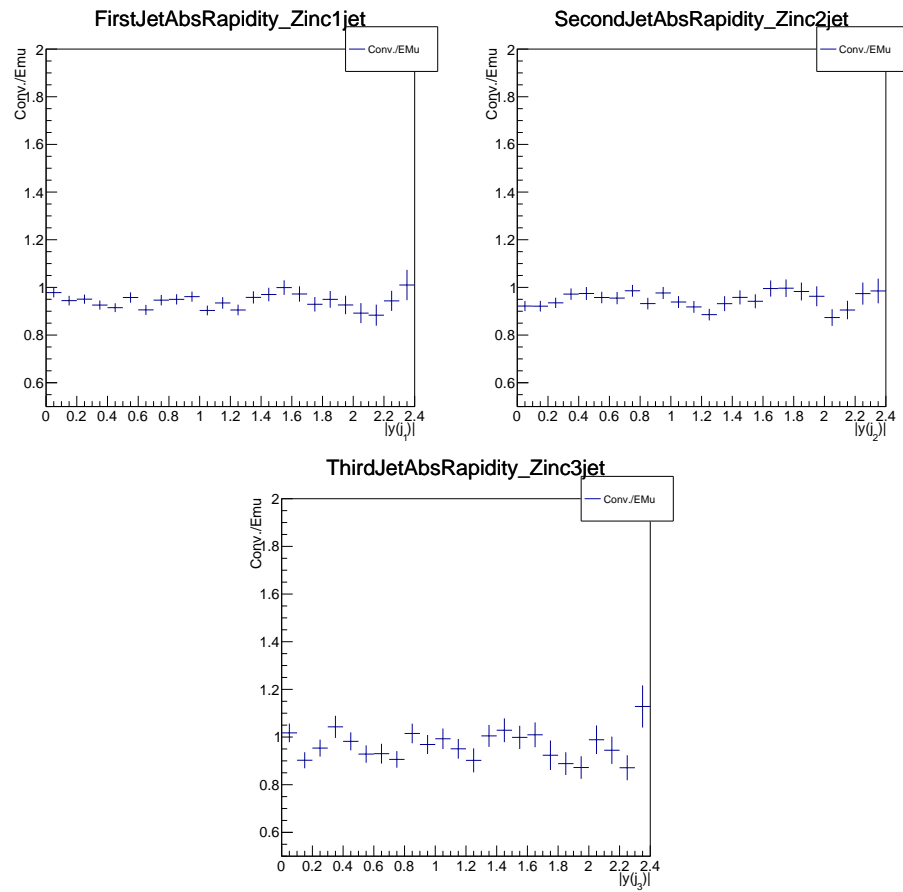


Figure 3.11: Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). First, second, and third jet absolute rapidity.

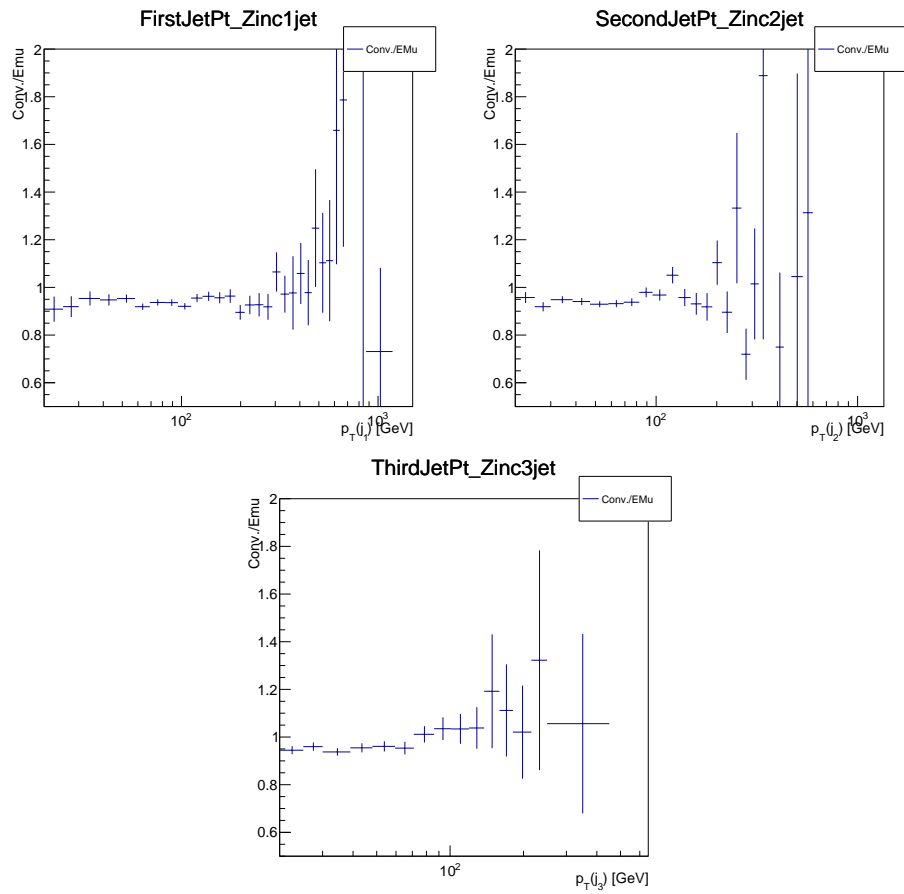


Figure 3.12: Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). First, second, and third jet p_T .

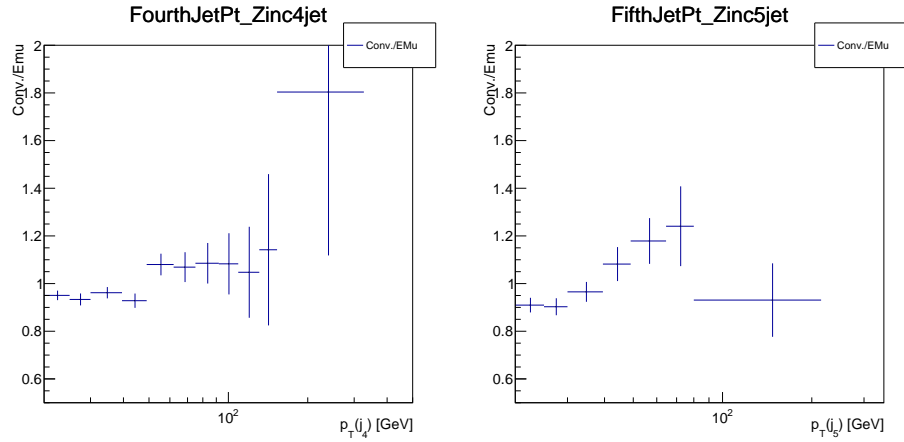


Figure 3-13: Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Fourth and fifth jet p_T .

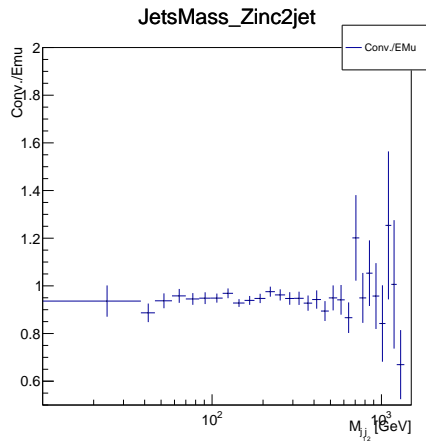


Figure 3-14: Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Invariant mass of leading and subleading jet.

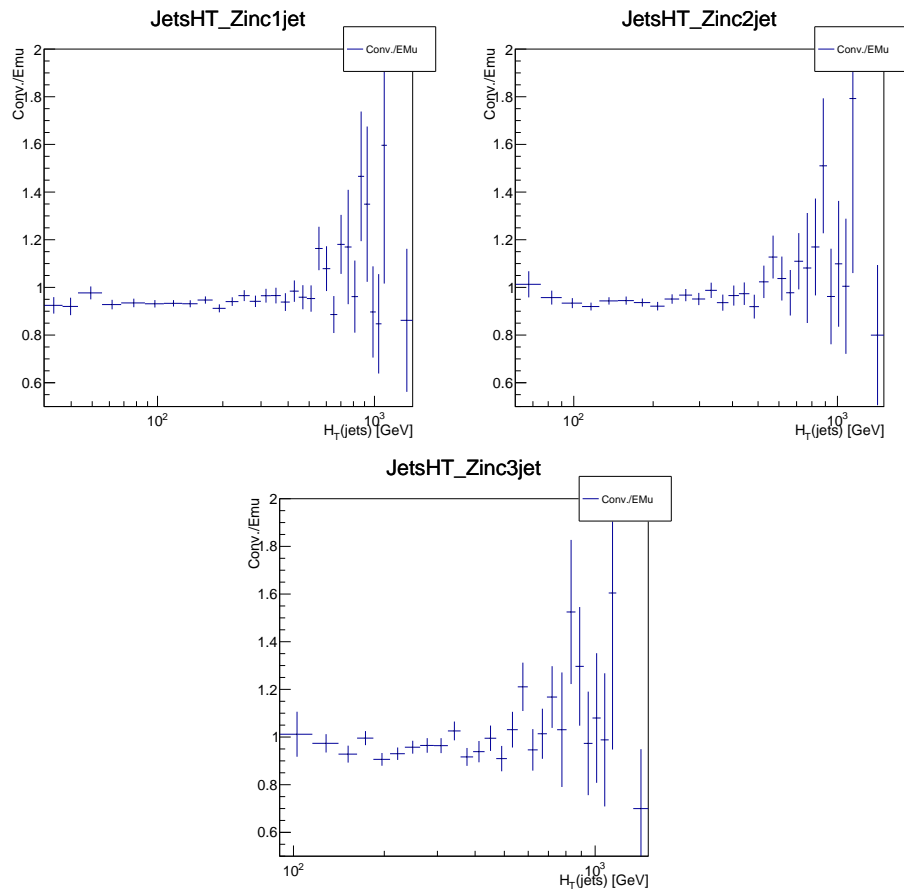


Figure 3-15: Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Total hadronic p_T one, two, and three jets inclusive.

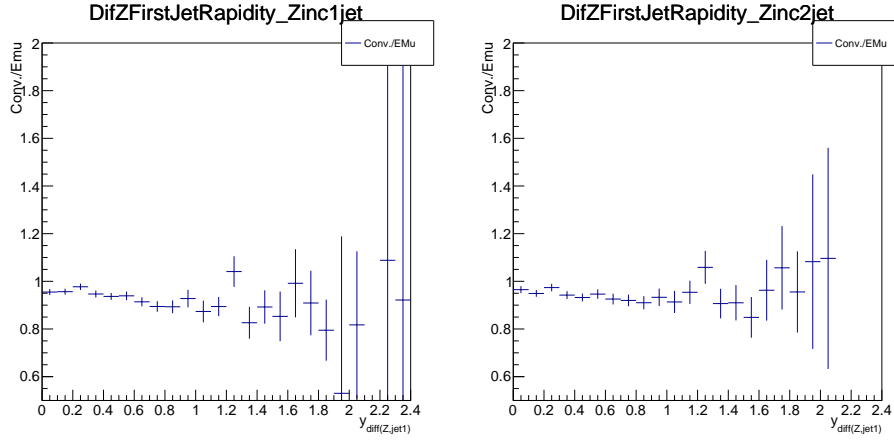


Figure 3-16: Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Difference in rapidity of the Z boson and leading jet with one jet (left) and two jets (right) inclusive.

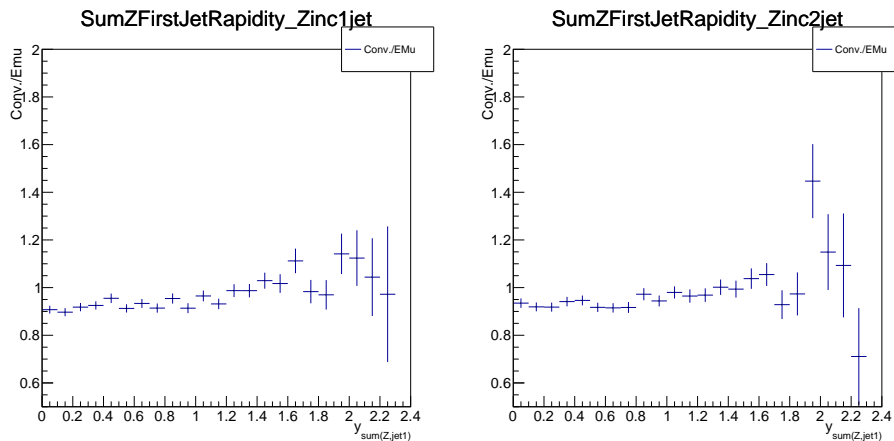


Figure 3-17: Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Rapidity sum of the Z boson and leading jet with one jet (left) and two jets (right) inclusive.

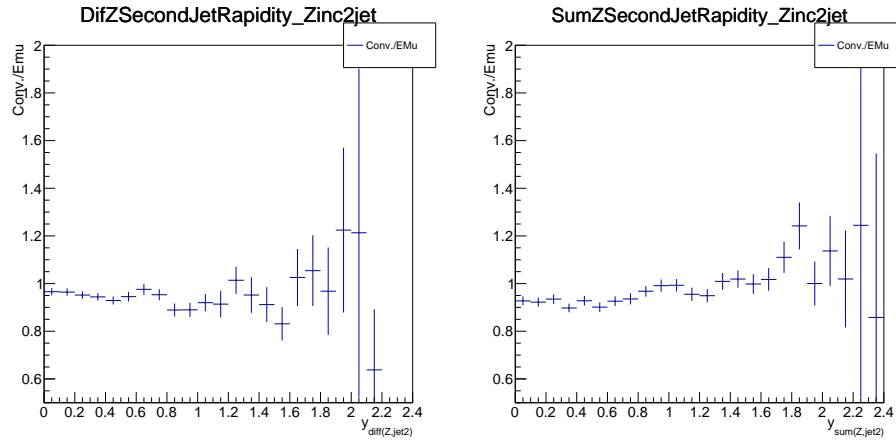


Figure 3-18: Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Difference (left) and sum (right) of rapidity of the Z boson and subleading jet.

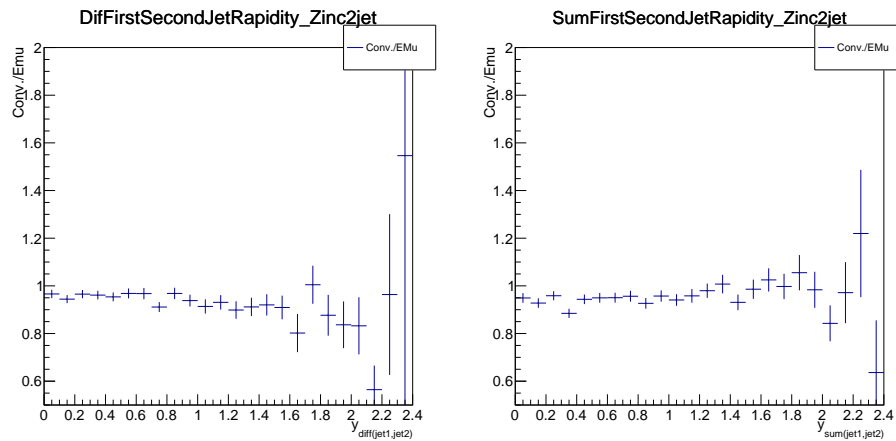


Figure 3-19: Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Difference (left) and sum (right) of rapidity of the leading and subleading jet.

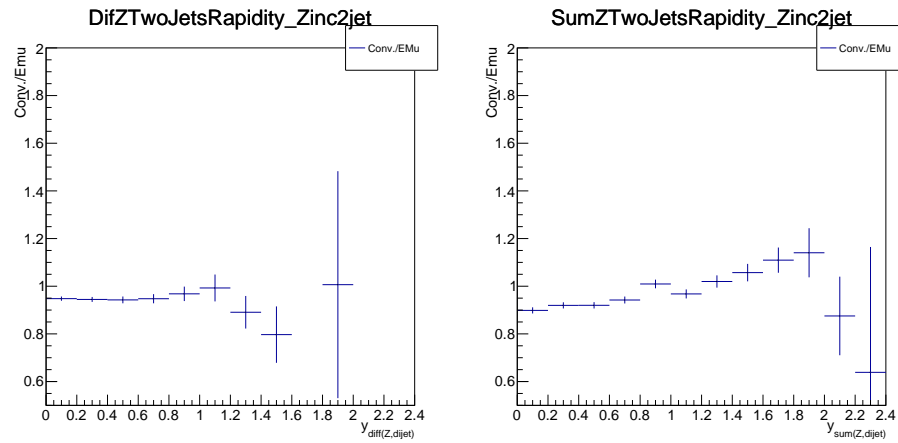


Figure 3-20: Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Difference (left) and sum (right) of rapidity of the Z boson and subleading jet.

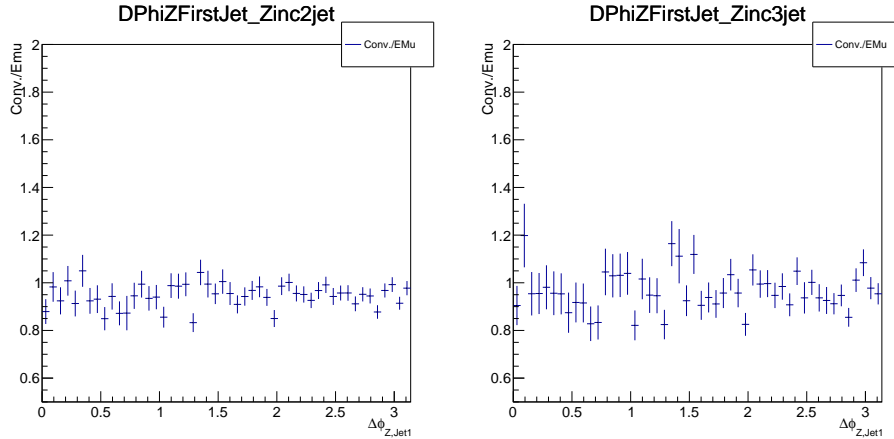


Figure 3-21: Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Difference in azimuthal angle between the Z boson and leading jet with one(left), two(right), and three(right) jet inclusive.

3.5 Observables

In this paper, the cross sections are presented as functions of several kinematic and angular observables to characterize the production mechanisms of $Z(\rightarrow \ell^+ \ell^-) + \text{jets}$ events.

The differential cross section has been measured as functions of the exclusive and inclusive jet multiplicities for a total number of up to eight jets in the final state, of the jet kinematic variables including jet p_T , the jet rapidity (y) and the scalar sum of the jet transverse momenta (H_T) for $N_{jets} \geq 1, 2, 3, 4, 5$.

The jet multiplicity distributions provide a general view about the agreement between measurement and different predictions. The comparisons can test different MC generators, and also estimate their accuracy in different configuration as number of jets.

The measurement of the distribution of $p_T(Z)$ for events with at least one jet is vital to understand the balance in the transverse momentum between the jets and the

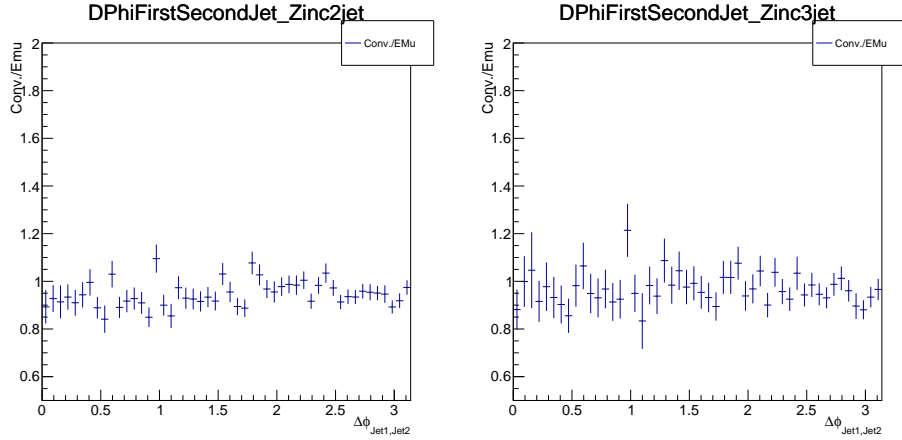


Figure 3-22: Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Difference in azimuthal angle between the leading and subleading jet with two(left) and three(right) jet inclusive.

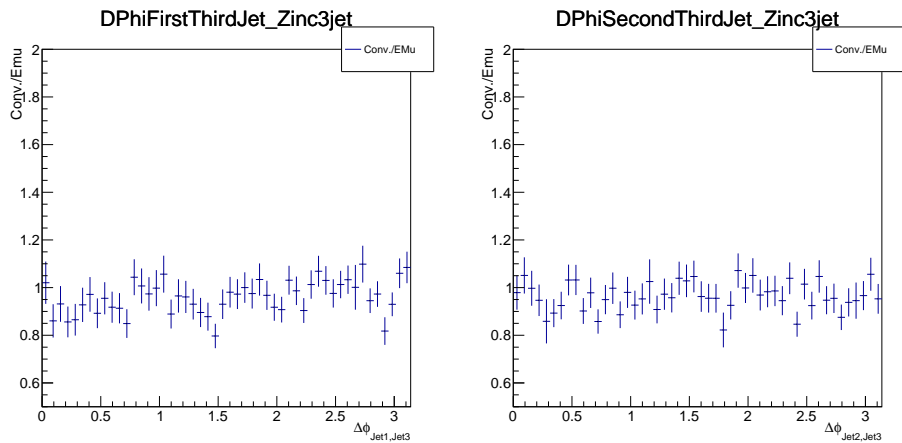


Figure 3-23: Comparison of two background estimation methods: MC samples only (Conv.) and data-driven method (EMu). Difference in azimuthal angle between the leading and third jet (left) and subleading and third jet (right) with three jet inclusive.

Z boson, and can be used for the comparison amongst theoretical predictions which achieve multiple soft-gluon emissions with different ways.

The rapidity (y) of Z boson ($y(Z)$) is related to the momentum fraction (x) carried by the parton in the forward-going (backward-going) proton. Therefore, the y distribution directly reflects the PDFs of the interacting partons. At the LHC, the $y(Z)$ distribution is expected to be symmetric around zero, therefore the appropriate measurement is the distribution of Z bosons as a function of the $|y|$.

The jet kinematic variables and H_T are sensitive to the effects of higher order corrections and these variables make it possible to specify the level of agreement between data and theory.

In terms of angular correlations between jets, cross sections are measured as a function of the difference in $\Delta y(j_i, j_k)$, and of the difference in azimuthal angle $\Delta\phi(j_i, j_k)$, between the i^{th} and k^{th} jets from the p_T -ordered list of jets in the event. From the point of angular correlations between Z boson and jets, cross sections are measured as a function of the difference in $\Delta y(Z, j_k)$, and of the difference in azimuthal angle $\Delta\phi(Z, j_k)$, between the i^{th} and k^{th} jets from the p_T -ordered list of jets in the event. The azimuthal angle separation ($\Delta\phi$) between the final state Z boson and jet is sensitive to the soft gluon radiation. The advantage of studying the ϕ distribution is that it only depends on the directions of the final state Z boson and jet.

Lastly, double differential cross sections are measured as functions of leading jet p_T and y , leading jet and $y(Z)$, $p_T(Z)$ and y . The measured cross sections are corrected for detector effects and compared with theoretical predictions to LO and NLO matched with parton showering as implemented in MC generators.

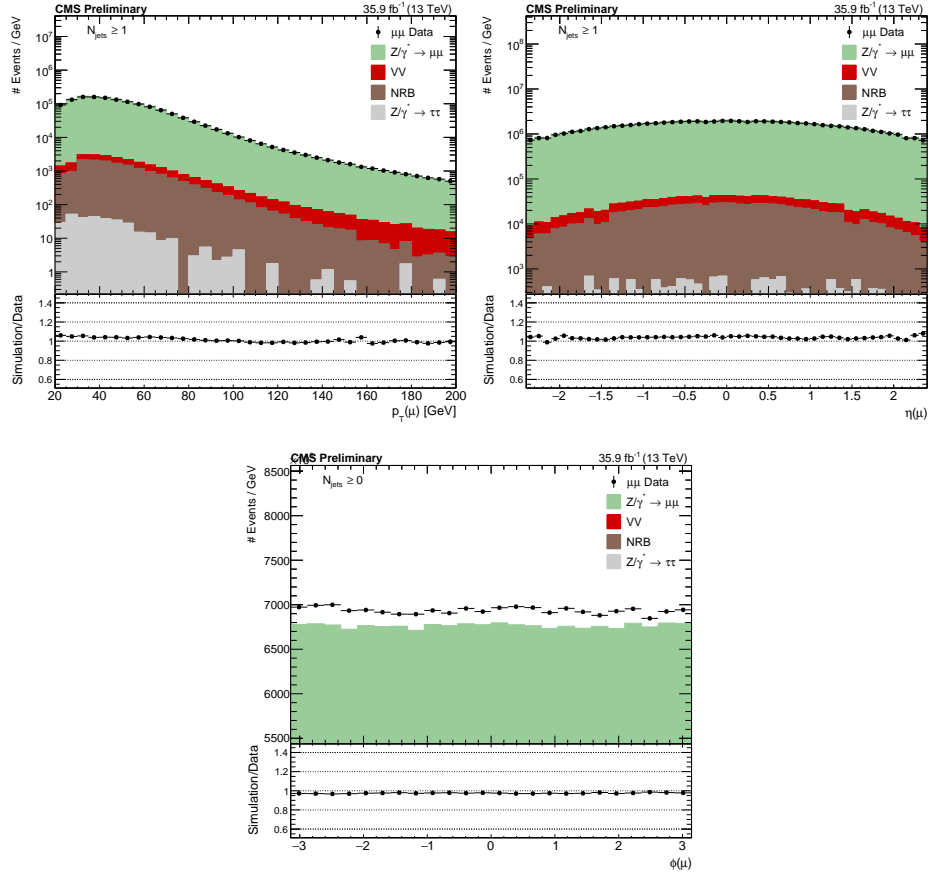


Figure 3-24: Lepton p_T (left), η (center), ϕ (right) with one jet inclusive.

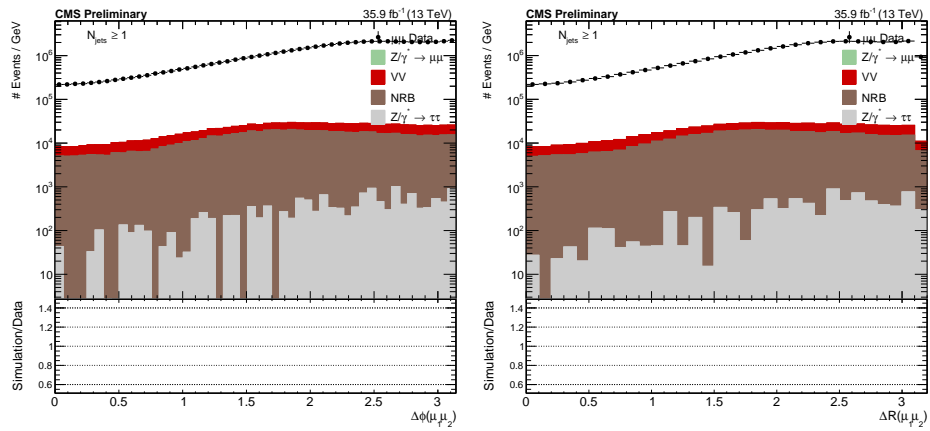


Figure 3-25: Difference in ϕ and R of the two Z decay leptons with one jet inclusive.

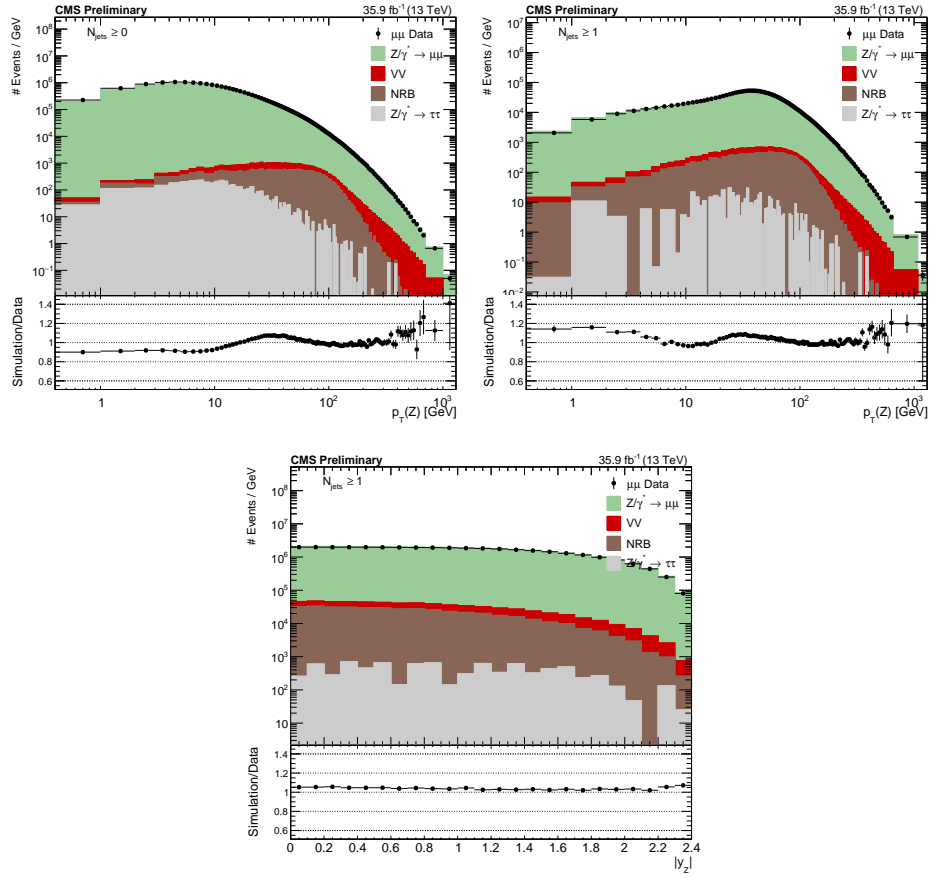


Figure 3-26: Z candidate p_T (upper) and $|y|$ (lower) with at least zero jets (left) and at least one jet (right). The background is estimated from both simulation and data driven methods describe in section ???. The ratio shows the combined statistical uncertainty of the data and total simulation.

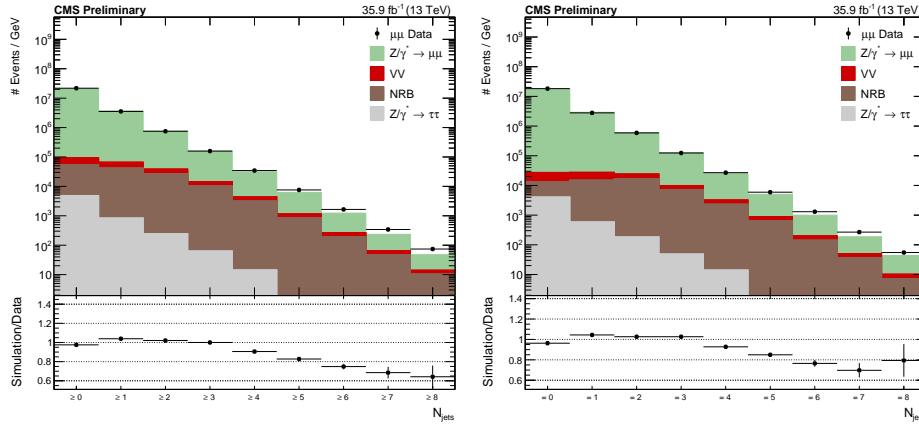


Figure 3.27: Inclusive (left) and exclusive (right) jet multiplicity. The background is estimated from both simulation and data driven methods describe in section ?? . The ratio shows the combined statistical uncertainty of the data and total simulation.

3.6 Phenomenological Models and Theoretical Calculations

We compare the measured $Z + \text{jets}$ differential cross sections to three calculations: MG5_AMC at next-to leading order (NLO), MG5_AMC at leading order (LO), and the GENEVA MC program. The two MG5_AMC calculations (version 2.2.2) [Alwall et al. (2014)] are interfaced with PYTHIA 8 (version 8.212) [Sjöstrand et al. (2015)]. For the LO MG5_AMC, the generator calculates LO MEs for five processes: $pp \rightarrow Z + N_{\text{jets}}$ with $N = 0 \dots 4$. The NNPDF 3.0 LO PDF [Ball et al. (2015)] is used and $\alpha_S(m_Z)$ is set to 0.130. The NLO MG5_AMC prediction includes NLO ME calculations for $pp \rightarrow Z + N_{\text{jets}}$ with N up to 2. The NNPDF 3.0 NLO PDF is used and $\alpha_S(m_Z)$ is set to 0.118. Both predictions use PYTHIA 8 to model the initial+final state parton showers (PS) and hadronization with the CUETP8M1 [Khachatryan et al. (2016)] tune that includes the NNPDF 2.3 [Ball et al. (2013)] LO PDF and $\alpha_S(m_Z) = 0.130$. ME and PS matching is done using the k_T -MLM [Alwall et al. (2008, 2009)] scheme with the matching scale set at 19 GeV for the LO MG5_AMC and the FxFx [Frederix and Frixione (2012)] scheme with the matching scale set to

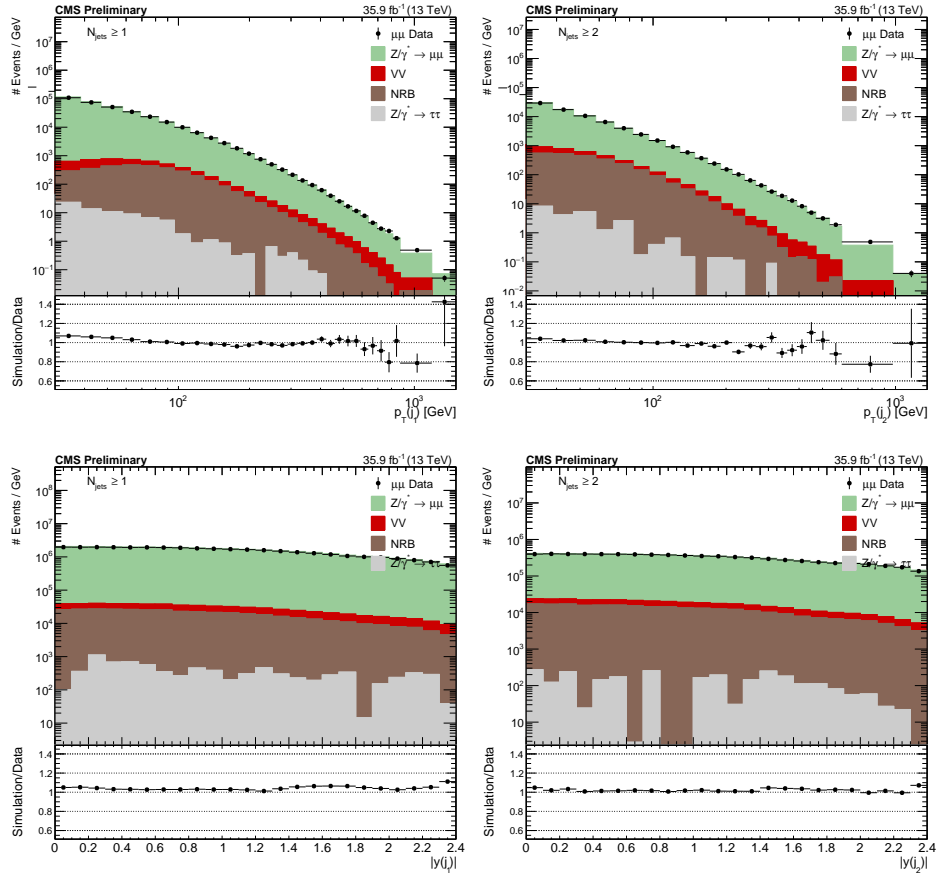


Figure 3-28: Jet p_T (upper) and $|y|$ (lower) for the leading jet (left) and sub-leading jet (right). The background is estimated from both simulation and data driven methods describe in section ???. The ratio shows the combined statistical uncertainty of the data and total simulation.

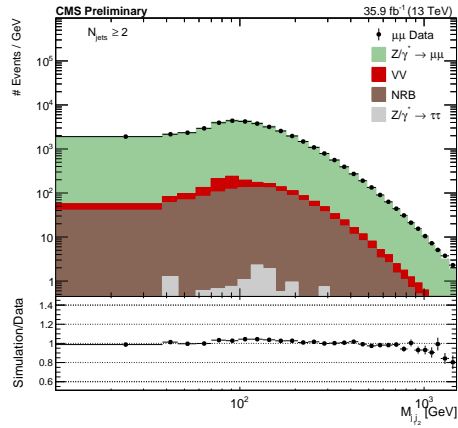


Figure 3-29: Invariant mass of leading and subleading jet.

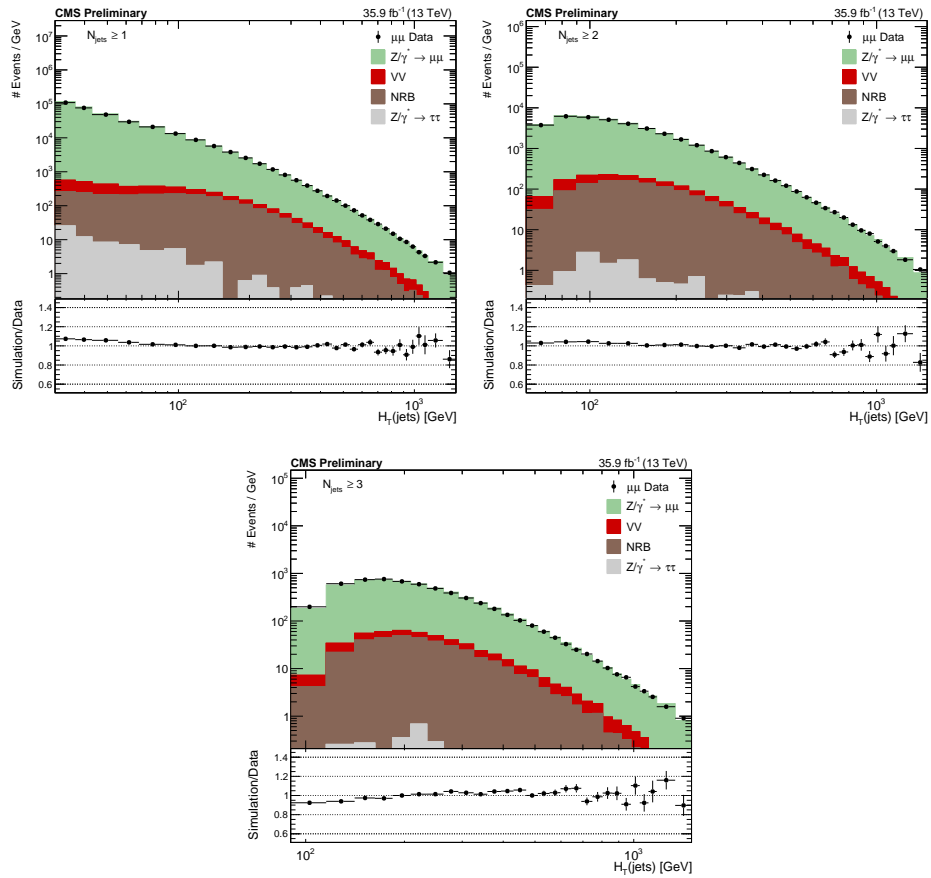


Figure 3-30: Total hadronic p_T one, two, and three jets inclusive.

30 GeV for the NLO MG5_AMC.

In this analysis uncertainties in the ME calculation are estimated for the NLO MG5_AMC calculation as recommended by the authors of the respective generators. Fixed-order cross section calculations depend on the renormalization (μ_R) and factorization (μ_F) scales. The uncertainty coming from missing terms in the fixed-order calculation is estimated by varying the μ_R and μ_F scales by factors of 0.5 and 2. Uncertainties in PDF and α_S values are also estimated in the case of the FxFx-merged sample. The PDF uncertainty is estimated using the set of 100 replicas of the NNPDF 3.0 NLO PDF, and the uncertainty in the α_S value used in the ME calculation is estimated by varying it by ± 0.001 . These two uncertainties are added in quadrature to the ME calculation uncertainties. All these uncertainties are obtained using the reweighting method [Frederix et al. (2012)] implemented in these generators.

3.7 Unfolding Procedure

Unfolding is done in order to reverse the effects of the detector and estimate the particle or generator (GEN) level distributions in data. With a particle level distribution any theory or experimental data can be readily compared to the distributions without a need for detector simulation. Unfolding requires understanding the finite resolutions of the detector where an observable is smeared about its true value by a Gaussian of width equal to the resolution. The measured observables are also taken from a Poisson distribution with the true smeared or reconstructed (RECO) value as the mean. When considering binned distribution the unfolding problem amounts to solving the matrix equation $Rx = y - b$ where R is the response matrix, x is the unknown particle level distribution, y is the measured reconstructed distribution, and b is the reconstructed background. The MADGRAPH5_AMC@NLO MC sample is used

to extract the nominal detector transformation.

The response matrix R is created by a pair of observables where one value is at GEN level and the other at RECO level. Pairs are created by ordering the particles in p_T and selecting both with the same index. The ordering can be done by ΔR matches and can lower the mismatch rate for a GEN and RECO pair. Since the mismatch rate is at the percent level the response matrix and the unfolding is insensitive to how the matching is done and the p_T ordering is used for simplicity. If a GEN (RECO) observable does not have a matching RECO (GEN) observable it is put into the corresponding GEN (RECO) underflow bin. The GEN underflow bins are used in the TUnfold package [Schmitt (2012)] to calculate the efficiency since the underflow events represent events that were not found after reconstruction. The RECO underflow bins are events that do not have a GEN match and are considered fakes. The matrix is normalized over a single GEN bin including the underflow bin. The normalization transforms the response matrix into probabilities for a GEN value in bin i to have a RECO match in bin j . The RECO background b and fakes are subtracted from the RECO distribution y bin by bin with the background approximated by both MC sample and data driven methods described in section 3.4.

The unfolding is done using the TUnfold package which utilizes a χ^2 minimization using a Tikhonov [Tikhonov (1963)] regularization term:

$$\chi^2 = \chi_0^2 + \chi_{reg}^2 = (y - Ax)^T V_{yy}^{-1} (y - Ax) + \tau^2 x^T L^{-1} L x \quad (3.8)$$

where x and y have been previously defined, V is the RECO covariance matrix, τ is the regularization strength, and L defines the regularization conditions. The regularization term is needed to dampen Poisson fluctuation which tend to be amplified during the unfolding process. In this analysis the regularization term does not have a bias vector and is done in curvature mode where the matrix L is the discrete second

derivative. To approximate the amount of regularization needed for a specific variable the condition of the response matrix can be calculated. The condition is the ratio of the largest and smallest single values of the response matrix. If the condition is low (>10) then no regularization is most likely needed and if it is large ($>1E5$) strong regularization is needed. In this analysis variable associated with angles such as η and $|y|$ have low condition numbers and are unfolded without regularization. Momentum variables have condition numbers in between the extremes defined above and need some regularization. Condition values are shown on the response matrices for select variable in figures 3-31 - 3-35 and all condition numbers are shown in table 3.7. The strength of the regularization τ is chosen using the L-Curve method [Calvetti et al. (2004)]. This plots the minimized values of χ_0^2 vs χ_{reg}^2 with different value of τ . The location of highest curvature is chosen as the optimal amount of regularization.

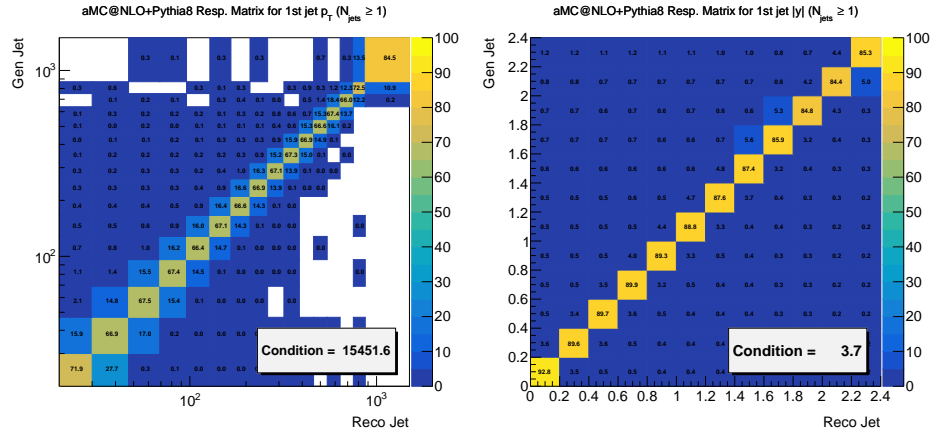


Figure 3-31: The response matrices of the leading jet $|y|$ (left) and p_T (right) with at least one jet. The reco and gen values are taken from the NLO MG5_AMC.

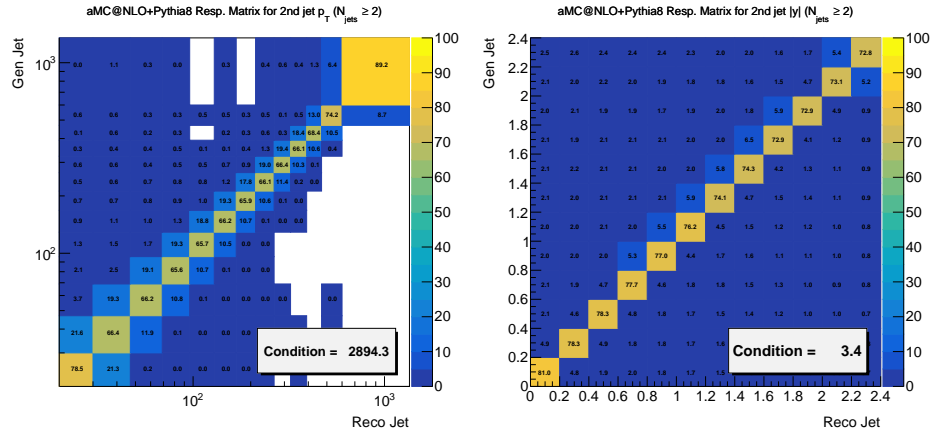


Figure 3-32: The response matrices of the second jet $|y|$ (left) and p_T (right) with at least two jets. The reco and gen values are taken from the NLO MG5_AMC.

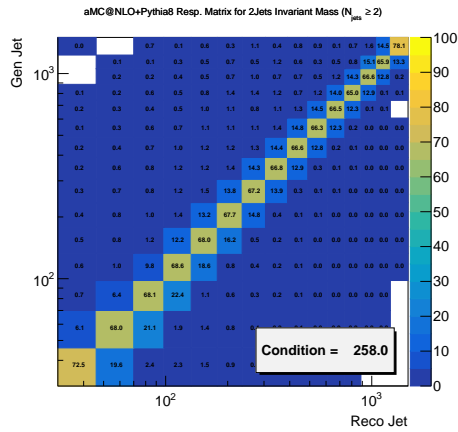


Figure 3-33: The response matrix of the dijet mass with at least two jets. The reco and gen values are taken from the NLO MG5_AMC.

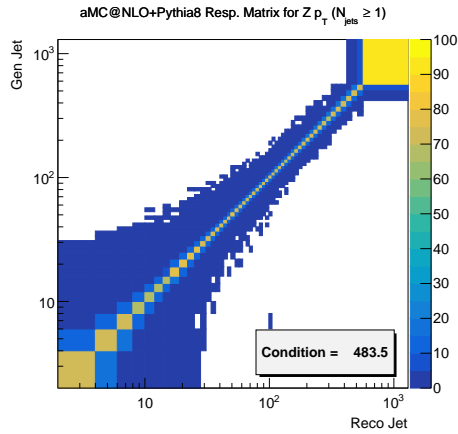


Figure 3-34: The response matrix of the $Z p_T$ with at least one jet. The reco and gen values are taken from the NLO MG5_AMC.

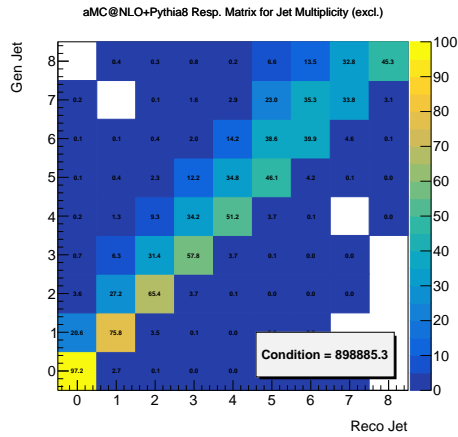


Figure 3-35: The response matrix of the jet multiplicity exclusive. The reco and gen values are taken from the NLO MG5_AMC.

The statistical uncertainties from both the data distribution y and response matrix R are calculated analytically in the TUnfold package. The systematic uncertainties are calculated from an envelope of two unfolded distributions after varying either the response matrix or data distribution. More details of the systematic errors are given in section 3.8.

Unfolding is validated using the bottom line test and comparing the folded-unfolded data distributions to the RECO distributions as a closure test. For the bottom line test χ^2 is calculated for the unfolded and RECO space to check if $\chi_{Unf}^2 \leq \chi_{Reco}^2$. A table of the χ^2 and degrees of freedom are shown in table 3.7. The number of degrees of freedom is approximately the number of bins in the respective space. The number of RECO bins is always two times the number of GEN bins as mentioned in the binning strategy above.

The binning for all variables is based on the detector resolution and statistics. A measure of detector resolution is migration defined as the fraction of events of a gen bin that are reconstructed to a different reco bin. In this analysis at least 68% of the gen events must be reconstructed in the same reco bin. For better performance with unfolding the number of reco bins is larger than the number of gen bins. For all variables the number of reco bins is double with each gen bin split evenly in half. All bins are also adjusted so that at least 100 signal events in each bin of the data distribution.

3.8 Systematical Uncertainties

The sources of experimental uncertainties are divided into the following categories: Jet Energy Scale (JES) and Jet Energy Resolution (JER), Lepton Efficiencies (identification, isolation, and track reconstruction), Lepton Energy Scale (LES) and Resolution (LER), Trigger efficiency, Luminosity, Pileup, Background and Unfolding uncertain-

ties. The listed uncertainties are assumed to be independent such that each can be computed individually and added in quadrature to obtain a total uncertainty. To compute the systematic uncertainty from each source, the analysis is repeated using the source values increased and decreased by 1σ from the central value. This results in bin-by-bin uncertainty contributions from each source in the unfolded distributions.

The JES uncertainty originates mainly from the uncertainty on the single particle response and it is the dominant source of systematic uncertainty. It affects both the reconstruction of the transverse energy of the selected jets and also the reconstructed kinematic variables measured with the calorimeter. In this analysis jet energy corrections (JEC) were applied to take into account inefficiencies, non-linearities and finite resolutions in energy and position of the reconstructed jets. The effect of the JES uncertainty is studied by scaling up and down the reconstructed jet energy by p_T and η -dependent scale factors. A similar procedure is followed for the JER. The uncertainties due to the JES and the JER vary in the range 1-11% as a function of jet multiplicity.

Scale factors for lepton efficiencies are applied on an object-by-object basis so that the simulation samples reflect the inefficiencies observed in data. The lepton identification, isolation, track reconstruction and trigger efficiencies in simulation are corrected with scaling factors derived with a T&P method and applied as a function of lepton p_T and η . To estimate the uncertainties, the total yield is recomputed with the scaling factors varied up and down by the fit uncertainties. The uncertainties associated with lepton efficiency in the electron channel is 1% while in the muon channel 0.5%.

Another small source of the lepton uncertainties is LES and LER and is $\sim 1\%$ for both channels.

A normalization uncertainty is assigned to the imperfect knowledge of the in-

tegrated luminosity. This is applied as an overall normalization uncertainty on all processes stemming from MC simulation and takes a value of 2.5% [CMS (2017)].

To match the pileup conditions in data and in MC simulation, pileup reweighing is applied for the simulated samples. The reweighing factors depend on the minimum-bias cross section. We vary the nominal minimum-bias cross section of 69.2 mb up and down by its uncertainty of 4.6% when reconstructing the response matrices, and take the difference in the unfolded data as the uncertainty.

The uncertainty on the unfolding procedure is due to both the statistical uncertainty in the response matrix coming from the finite size of the MC sample used to compute it and to the possible event generator dependence of the response matrix itself. Because of the finite binning a different distribution will lead to a different response matrix. This uncertainty is estimated by weighting the MC to agree with the data in each distribution and building a new response matrix. The weighting is done using a finer binning than for the measurement. The difference between the nominal results and the results unfolded using the alternative response matrix is taken as the systematic uncertainty. An additional uncertainty comes from the finite size of the MC sample used to build the response matrix. This source of uncertainty is called unfolding statistics ("unf stat") and is included in the systematic uncertainty of the measurement as well. Statistical fluctuations in the response matrix are propagated analytically in the TUnfold package.

Lastly, the background samples are varied by their corresponding cross section uncertainty before being subtracted from data prior to unfolding. The systematic uncertainties used for the combination of the electron and muon channels are summarized in Tables 3.8 - 3.12.

3.8.1 Jet Energy Scale

The energy of each jet in the event is increased or decreased by a percent determined from tables in the JEC database: <https://github.com/cms-jet/JECDatabase>. All independent sources are treated as correlated and combined into uncertainty. The table is binned in p_T and η . The jet energy variation is performed before any p_T requirements in the analysis.

3.8.2 Jet Energy Resolution

The energy of the MC jets are smeared to improve the matching to the data jet resolution as described in the jet selections section. The values used for smearing are varied up and down using the uncertainties obtained from the JER database: JER database at <https://github.com/cms-jet/JRDatabase>. The resolution uncertainty table is binned in p_T , η , and ρ . The jet resolution variation is performed before any p_T requirements in the analysis.

3.8.3 Muon Energy Scale and Resolution

The muon energy scale and resolution systematics are estimated from variations in the Rochester correction described in 3.2. The variations included in the package are statistical and parameters used in the fitting method such as the fitting function and Z mass window. For each event the RMS is calculated for the different statistical values and the maximum deviation is taken as the systematic for all other variations. All systematics are added in quadrature and the Rochester correction is pushed up or down by the total systematic about the central value.

3.8.4 Muon Efficiencies

Efficiency uncertainties are calculated from variations done with the Tag and Probe method. The correlated sources include variations in the fit procedure, MC sample,

and tag selection. A Breit-wigner is used as the alternate fit function to the nominal crystal ball function. The leading order Madgraph MC sample is used as the alternate sample to the nominal NLO Madgraph sample. For the alternate tag selection the p_T cut is increased to 35GeV from the nominal of 30GeV. All of these sources are treated as correlated and varied together.

For the statistical uncertainty each bin is treated as an independent uncertainty source and does not have any correlated with other bins. The analysis must be repeated then for each bin of the scale factor varied up and down by the statistical uncertainty. To reduce computation time the analysis is run only once and copies of the response matrices are produced corresponding to each scale factor bin variation. This is possible because the scale factor only changes the overall event weight and does not change the kinematics of any object. The data is then unfolded using all response matrix variations and the width is taken as the uncertainty. All variations are added in quadrature for each bin to obtain the total uncorrelated efficiency uncertainty.

3.8.5 Integrated Luminosity

Lumi: official value 2.5%

3.8.6 Pile Up

Pileup: Minimum bias cross section, used to obtain the data pile up profile, varied $\pm 4.6\%$ from 69.2mb when calculating pileup scale factors.

3.8.7 Background Cross Section

The background samples are varied by their corresponding cross section uncertainty before being subtracted from data prior to unfolding. All background sources are varied at the same time and in the same direction.

3.8.8 Parton Distribution Function, alpha s

Weights are provided in the MC signal sample corresponding to variations in the Parton Distribution Function (PDF) by changing (fill in). Each weight is used to fill the generator level histograms and creates an envelop for each histogram corresponding to the PDF variations. The standard deviation in each bin is calculated from:

$$\sigma^2 = \frac{\Sigma(var_0 - var_i)^2}{NPDF - 1} \quad (3.9)$$

where var_0 corresponds to the bin value with no PDF variations, var_i is the bin value of the i th variation and $NPDF$ is the total number of variations. The strong coupling constant α_s uncertainty is obtained in a similar way as the PDF where the variation is calculated from an envelope of variations.

3.8.9 Unfolding Statistics

Statistical fluctuations in the response matrix are propagated analytically in the TUnfold package.

3.8.10 Unfolding

The technique for estimating the unfolding uncertainty comes from the 2015 Z +jets analysis ?. The procedure mainly estimates the uncertainty from the shape within bins. The ratio of the background subtracted data and signal MC with 5 times finer binning is fit with Chebyshev polynomials. The function is then used to reweight the signal MC reco level when repeating the analysis to create a varied response matrix. All unfolding steps are then repeated and the difference with the central value is taken to be the uncertainty.

3.8.11 Software

The analysis software has multiple stages in order to deal with the high level physics object from CMS events and reduce them to variable distributions. The process of creating the high level physics objects from raw detector data is done centrally and the reconstruction `c++` code and object classes are defined in the CMS software (CMSSW). The raw detector events are reconstructed with algorithms such as Particle Flow to pick out particles from an event and save information such as a particle four vector. Two layers of skim are then done to reduce the data set size and remove any information not necessary for physics analyses. The first skimmed data set is called Analysis Object Data (AOD) and the second is miniAOD. The miniAOD data set is used as the starting point of this analysis.

The first step of the analysis software called Shears framework is to reduce the miniAOD data set further and remove the class structure. This is done in two steps called Baobab and Bonsai. Baobab is an EDAnalyzer and interacts directly with miniAOD. The skimming the Shears is done by saving only lepton, jet, and trigger information from each event. The miniAOD is organized in classes such as “muon” where the objects have member variables such as the four vector. This makes reading and writing events to file difficult and adds unnecessary complexity to the analysis. Shears takes the information in “muon” and saves it as vectors where the index is the i th muon in the event. This format is typically called flat ntuples. This means an event will contain a vector for the muons p_T and the first element will be the p_T of the first muon in the event. Care has to be taken so that the vectors of the same physics object all have the same size and the indices match. Preliminary cuts on the objects saved can be done during the Baobab stage to further reduce the size, but all events must be counted to preserve the luminosity estimate. Commonly a soft momentum cut is used for both the leptons and jets around 10GeV and events must have at least

one lepton to be saved. In the Bonsai stage of Shears we skim based on the final state particles we are looking for in the analysis. For this Z analysis we require at least two leptons and cut lepton p_T at 15GeV.

The ntuples are less than 5TB for the full $36fb^{-1}$ and can be stored on a local cluster or private space on a Tier 2/3 machine. The final step of producing variable distributions uses the ntuples as the input events. When running over events in the ntuple all vectors are filled with information from one event at a time. The MC weights and any corrections are done on the fly and respect the order in which the corrections are derived such as applying the muon ID scale factors before applying the Rochester corrections.

3.9 Results

The measurements from the electron and muon channels are consistent within the statistical and systematic uncertainties, and hence they are combined. To combine the two channels, a hybrid method based on weighted mean and the Best Linear Unbiased Estimates (BLUE) method [Lyons et al. (1988); Valassi (2003)] is used to calculate the cross section values. This method requires the construction of a covariance matrix (including statistical and systematic uncertainties) with all correlations determined externally.

The 2016 data statistically allows us to determine the differential cross sections of jet multiplicities up to eight jets and to study the cross sections as a function of several kinematic observables up to five jets. The combined single-differential cross sections are shown in figures 3-36-3-54, while double-differential cross sections are given in figures 3-55-3-57. All results are compared with theoretical prediction from MG5_AMC at LO and MG5_AMC at NLO and compared to the GENEVA MC program for results with at least one or two jets.

The jet transverse momenta and rapidities up to five leading jets can be seen in figures 3·36-3·40. For both quantities data distributions are well reproduced by the simulations. The MG5_AMC at LO, MG5_AMC at NLO, describe the data well in general. The GENEVA prediction shows good agreement for the measured p_T and y of the first jet, while it undershoots the data at low p_T for the second jet.

In addition, the inclusive jet differential cross sections as a function of H_T for events with at least one, two, three jets respectively are presented in figure 3·41. Both MG5_AMC at LO and MG5_AMC at NLO are compatible with the measurement. The contribution at higher values of H_T is slightly overestimated, but the discrepancy is compatible with the theoretical and experimental uncertainties. The slopes of the distributions for the first two jet multiplicities predicted by GENEVA samples do not fully describe the data.

In figure 3·43 the measured cross sections as a function of the exclusive jet multiplicity, for a total number of up to 8 jets in the final state, are shown. The trend of the jet multiplicity represents the expectation of the pQCD prediction for an exponential decay with the number of jets. The agreement is very satisfactory for the exclusive distributions for all the theoretical estimations, within the uncertainties and going up to the maximum number of final state partons included in the ME, namely 4 in the MC generators used here. The GENEVA predictions do not model the jet multiplicity for events with greater than 2 jets.

The $p_T(Z)$ distribution in figure 3·44 is described well above the jet cut of 30 GeV where the kinematics are dominated by jets modeled at NLO accuracy. Below the jet cut of 30 GeV non-perturbative QCD effects become dominant and the predictions show significant deviations from data. In many regions the total uncertainty in data is smaller than the theoretical uncertainty and greatly reduces the predictive power of the sample.

Overall the MG5_AMC at NLO predictions describe the data within theoretical uncertainties over a wide variety of kinematics. In regions of NLO accuracy, such as the first and second jet p_T and y , the agreement is within 10% of data up to the TeV scale.

Table 3.7: Unfolding Numbers

Variable	Condition	χ_{reco}^2	χ_{unf}^2	dof
ZNGoodJets_Zexc	898885.30	14935.47	11828.81	9
FirstJetPt_Zinc1jet	15451.57	537053.62	14540.67	16
SecondJetPt_Zinc2jet	2894.30	154113.95	1959.30	13
ThirdJetPt_Zinc3jet	629.64	34252.99	225.41	8
FourthJetPt_Zinc4jet	339.46	5810.82	489.48	6
FifthJetPt_Zinc5jet	49.86	736.75	349.78	4
FirstJetAbsRapidity_Zinc1jet	3.66	124313.38	2747.19	12
SecondJetAbsRapidity_Zinc2jet	3.41	34030.45	644.80	12
ThirdJetAbsRapidity_Zinc3jet	3.44	7930.90	338.65	12
FourthJetAbsRapidity_Zinc4jet	4.08	632.03	94.84	12
FifthJetAbsRapidity_Zinc5jet	3.85	73.18	136.64	12
JetsHT_Zinc1jet	3420.74	283062.72	12113.22	16
JetsHT_Zinc2jet	578.42	54492.54	1893.86	15
JetsHT_Zinc3jet	107.43	9221.13	215.51	14
JetsHT_Zinc4jet	18.83	581.50	319.55	12
JetsHT_Zinc5jet	6.43	49.74	177.80	11
JetsMass_Zinc2jet	258.05	39265.41	1868.91	15
ZPt_Zinc0jet	8522.58	75441.58	70094.66	55
ZPt_Zinc1jet	483.48	208237.88	9300.63	54
ZPt_Zinc2jet	92.32	38695.40	1737.56	54
ZAbsRapidity_Zinc1jet	10.80	120353.26	1732.65	12
SumZFirstJetRapidity_Zinc1jet	114.36	121250.61	2499.35	12
DifZFirstJetRapidity_Zinc1jet	1411.84	120186.84	3619.02	12
SumZFirstJetRapidity_Zinc2jet	142.76	33426.92	1194.56	12
DifZFirstJetRapidity_Zinc2jet	873.23	33422.14	769.24	12
SumZSecondJetRapidity_Zinc2jet	214.57	33478.54	963.12	12
DifZSecondJetRapidity_Zinc2jet	479.07	33442.64	859.86	12
SumZTwoJetsRapidity_Zinc2jet	155.73	33445.96	657.31	6
DifZTwoJetsRapidity_Zinc2jet	2325.62	33404.99	563.19	6
SumFirstSecondJetRapidity_Zinc2jet	72.41	33799.86	2880.53	12
DifFirstSecondJetRapidity_Zinc2jet	238.31	33460.06	1174.13	12
DPhiZFirstJet_Zinc1jet	185.33	132910.98	8507.17	25
DPhiZFirstJet_Zinc2jet	37.45	33627.87	2072.65	25
DPhiZFirstJet_Zinc3jet	17.11	7774.54	586.37	25
DPhiZSecondJet_Zinc2jet	2.92	33739.31	1634.70	25
DPhiZSecondJet_Zinc3jet	2.82	7709.30	402.89	25
DPhiZThirdJet_Zinc3jet	4.52	7931.24	705.20	25
DPhiFirstSecondJet_Zinc2jet	7.73	33541.43	1932.58	25
DPhiFirstSecondJet_Zinc3jet	9.67	7829.29	537.04	25
DPhiFirstThirdJet_Zinc3jet	4.33	7820.15	415.85	25
DPhiSecondThirdJet_Zinc3jet	2.96	8440.03	923.12	25

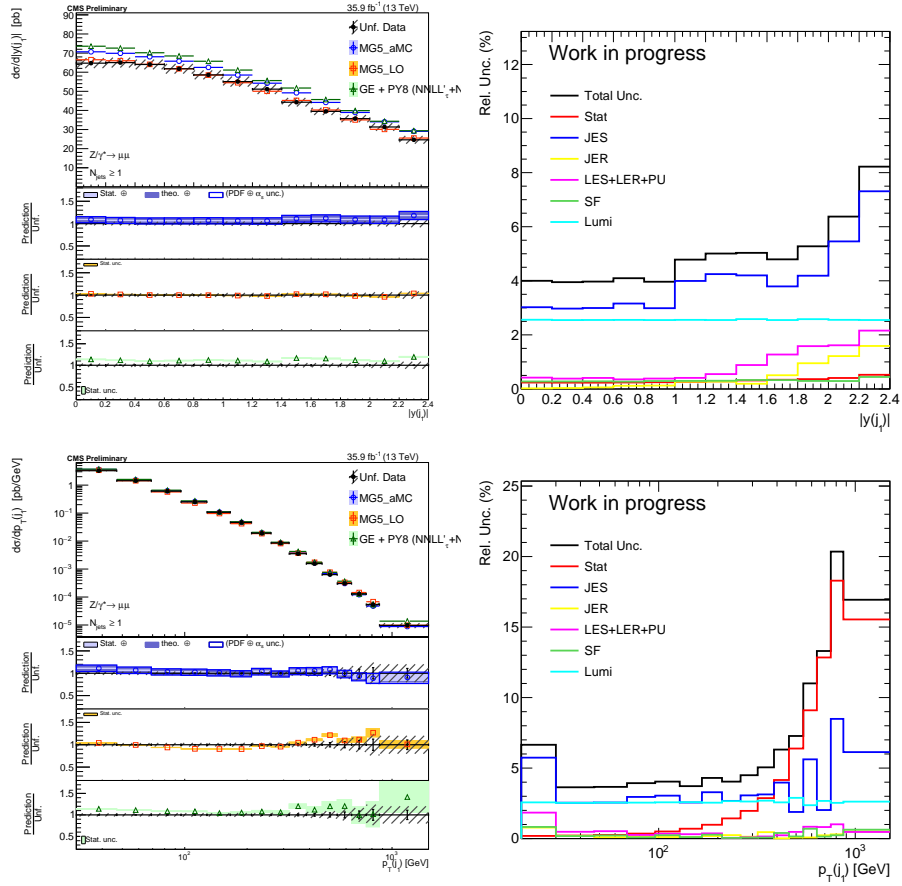


Figure 3-36: The measured differential cross section as a function of leading jet $|y|$ (left) and p_T (right) with at least one jet for the combined channel. For data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_AMC, LO MG5_AMC, and GENEVA. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_AMC and statistical only for the GENEVA and LO MG5_AMC predictions.

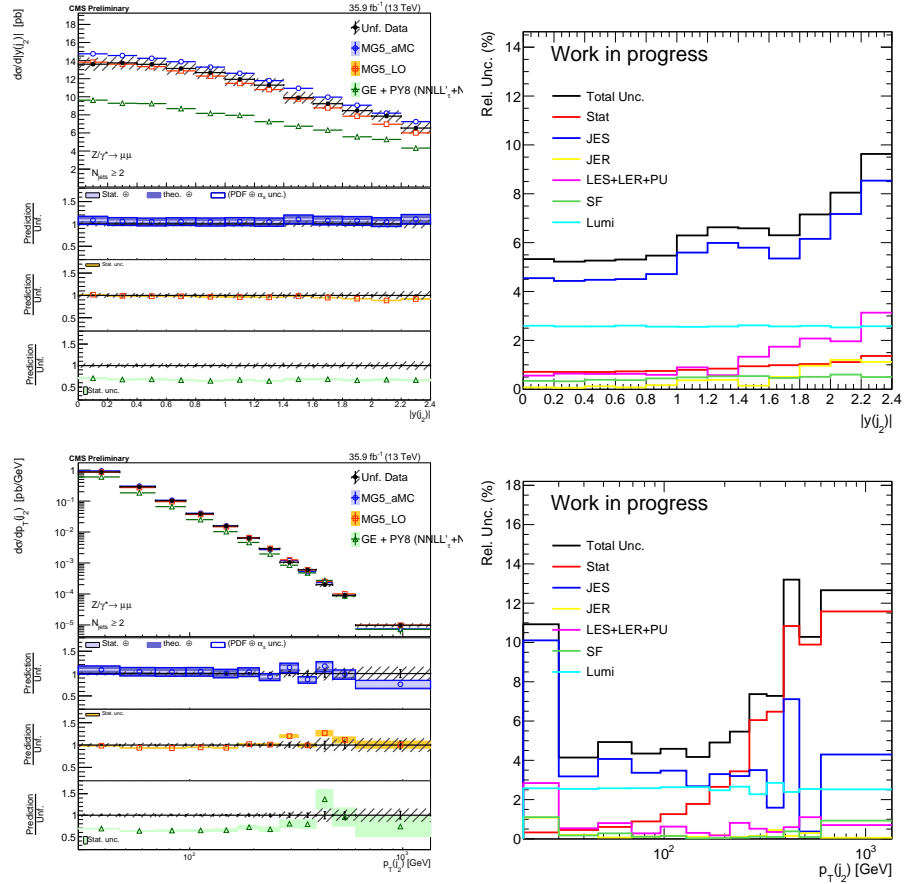


Figure 3-37: The measured differential cross section as a function of second jet $|y|$ (left) and p_T (right) with at least two jets for the combined channel. For data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_AMC, LO MG5_AMC, and GENEVA. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_AMC and statistical only for the GENEVA and LO MG5_AMC predictions.

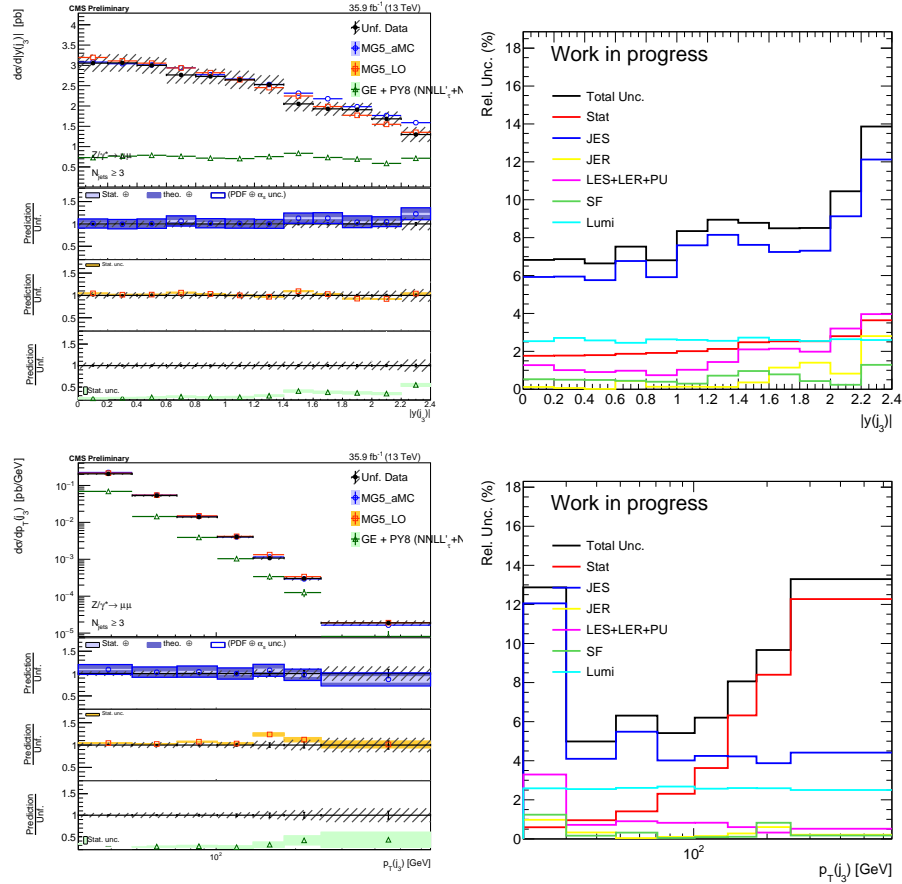


Figure 3-38: The measured differential cross section as a function of third jet $|y|$ (left) and p_T (right) with at least three jets for the combined channel. For data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_aMC and LO MG5_aMC. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_aMC and statistical only for the LO MG5_aMC predictions.

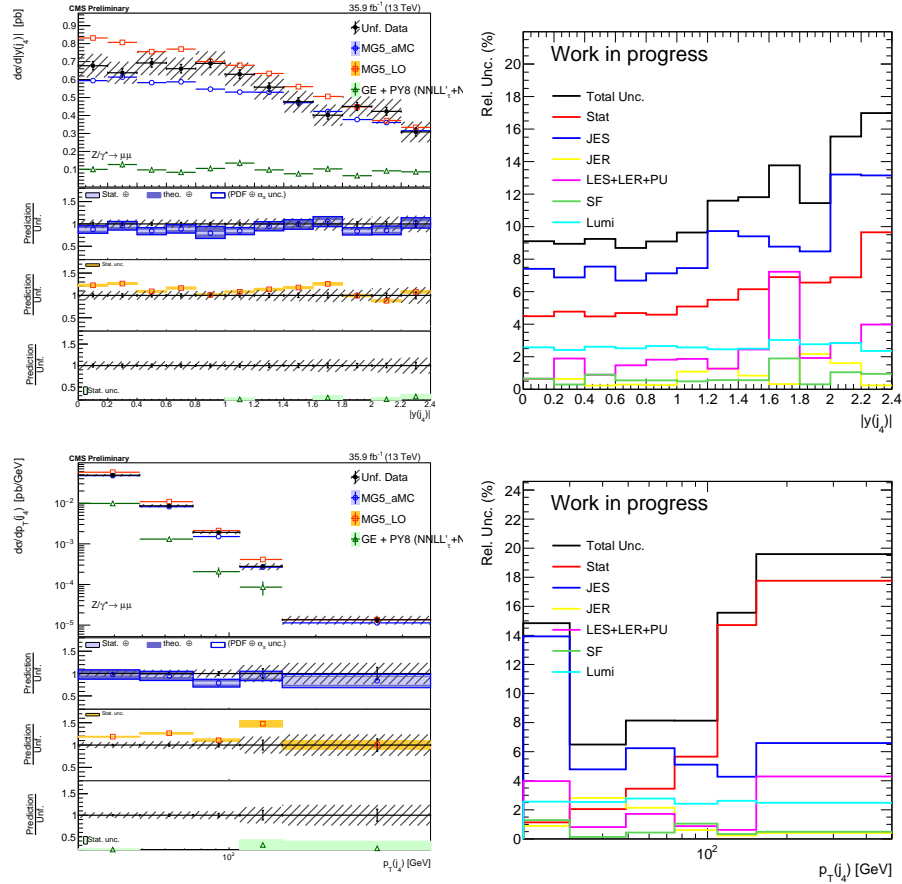


Figure 3-39: The measured differential cross section as a function of fourth jet $|y|$ (left) and p_T (right) with at least four jets for the combined channel. For data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_aMC and LO MG5_aMC. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_aMC and statistical only for the LO MG5_aMC predictions.

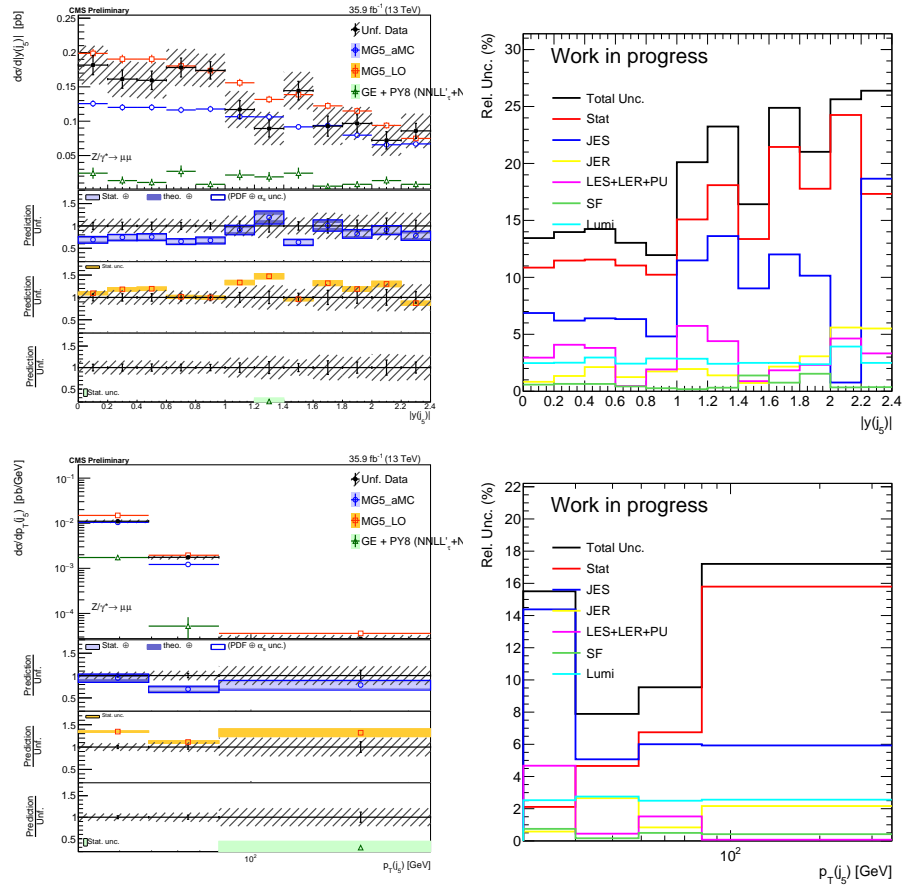


Figure 3-40: The measured differential cross section as a function of fifth jet $|y|$ (left) and p_T (right) with at least five jets for the combined channel. For data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_aMC and LO MG5_LO. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_aMC and statistical only for the LO MG5_LO predictions.

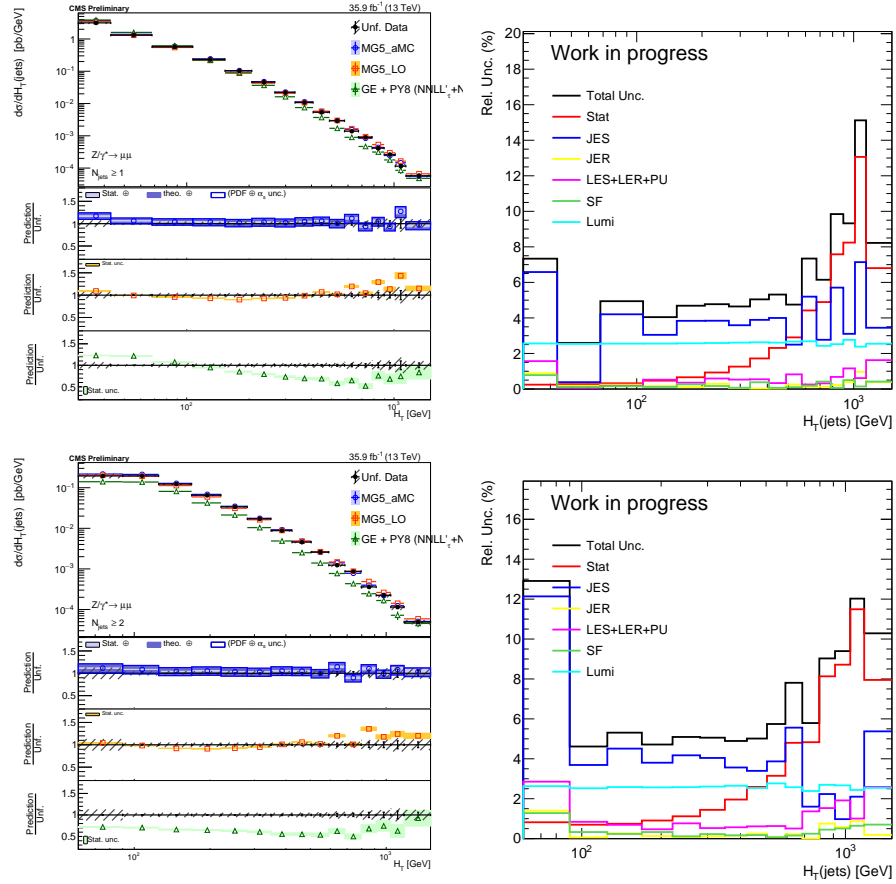


Figure 3-41: The measured differential cross section as a function of total hadronic p_T with at least one(left), two(middle), and three(right) jets. For data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_AMC and LO MG5_AMC. The measurements with at least one and two jets is also compared to GENEVA. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_AMC and statistical only for the GENEVA and LO MG5_AMC predictions.

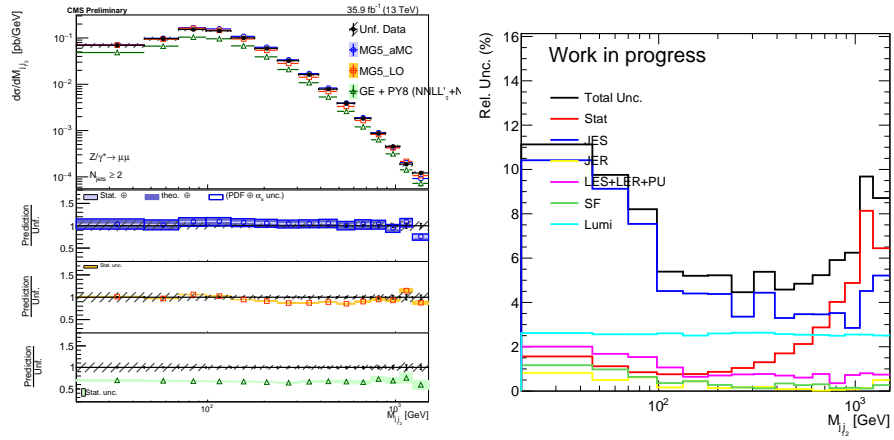


Figure 3.42: The measured differential cross section as a function of dijet mass with at least two jets. For data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_AMC, LO MG5_AMC, and GENEVA. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_AMC and statistical only for the GENEVA and LO MG5_AMC predictions.

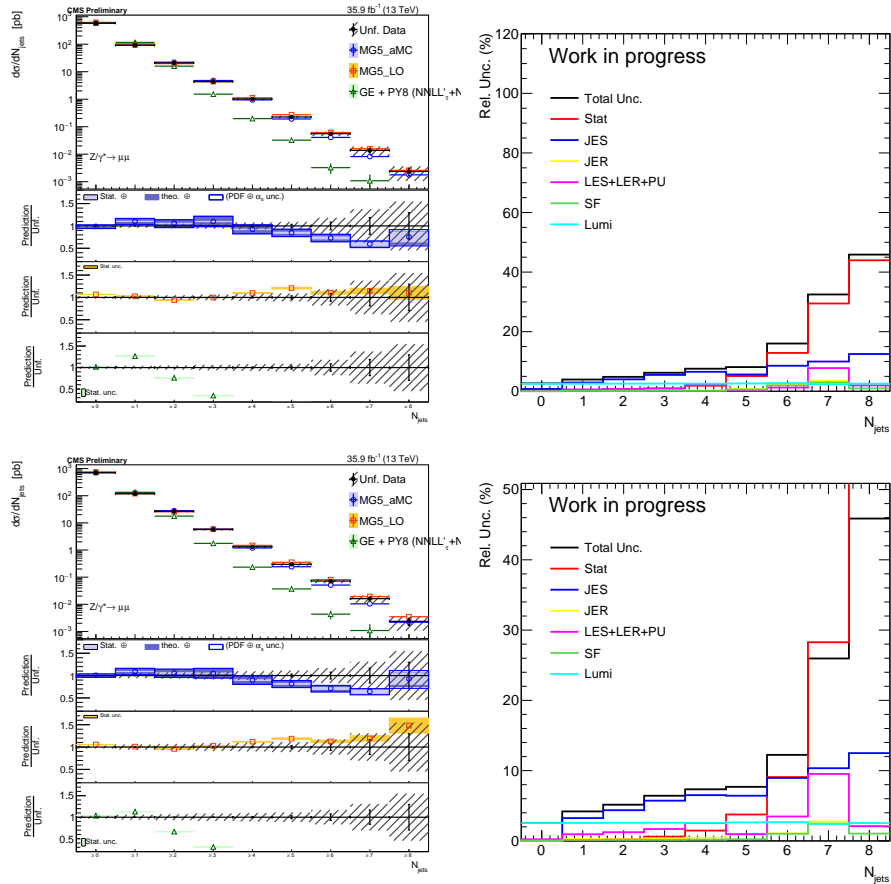


Figure 3.43: The measured differential cross section as a function of jet multiplicity inclusive(left) and exclusive(right). For data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_AMC, LO MG5_AMC, and GENEVA. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_AMC and statistical only for the GENEVA and LO MG5_AMC predictions.

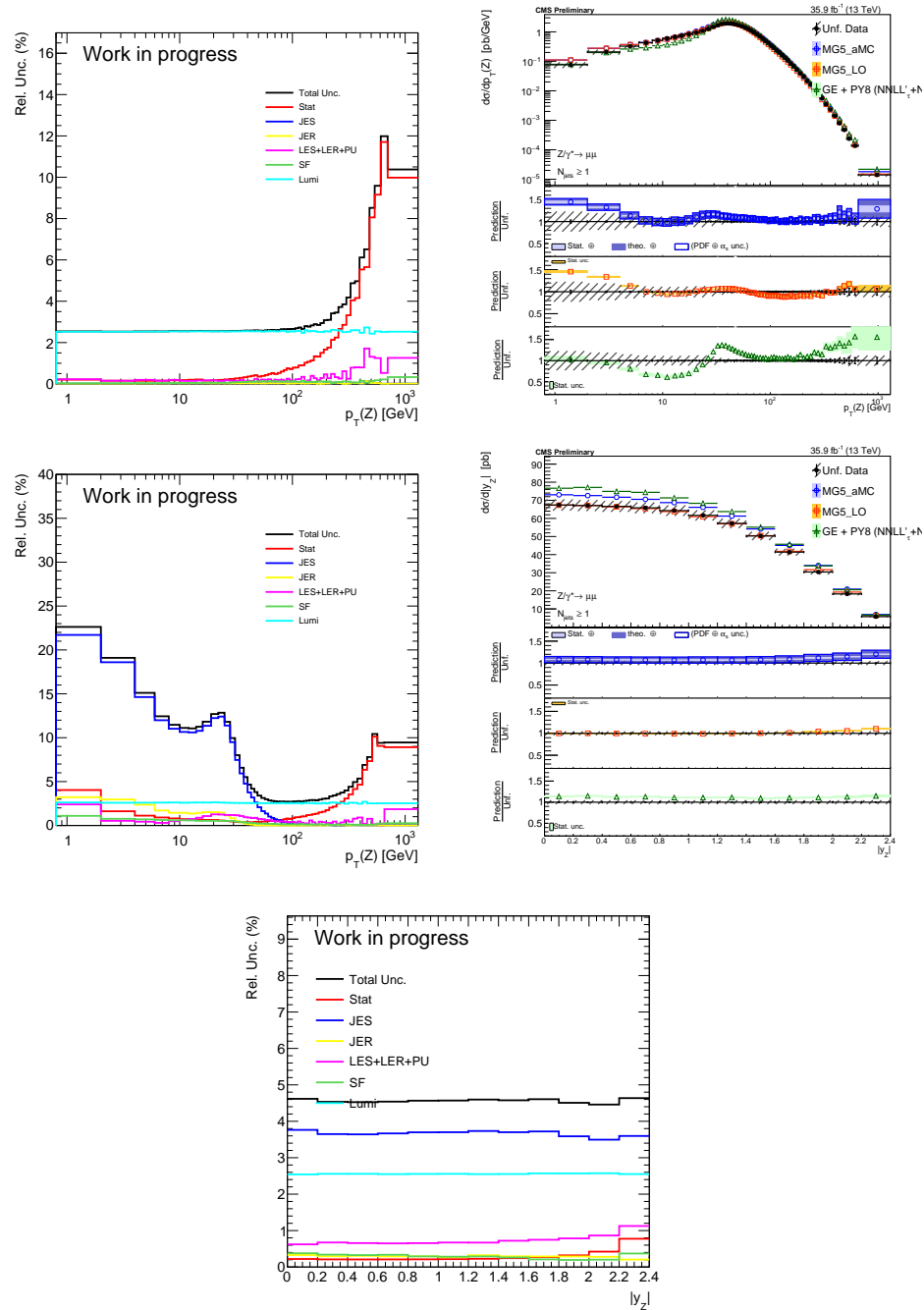


Figure 3-44: The measured differential cross section as a function of Z candidate p_T (left) and Z absolute rapidity (right) with at least one jet. For data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_AMC, LO MG5_AMC, and GENEVA. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_AMC and statistical only for the GENEVA and LO MG5_AMC predictions.

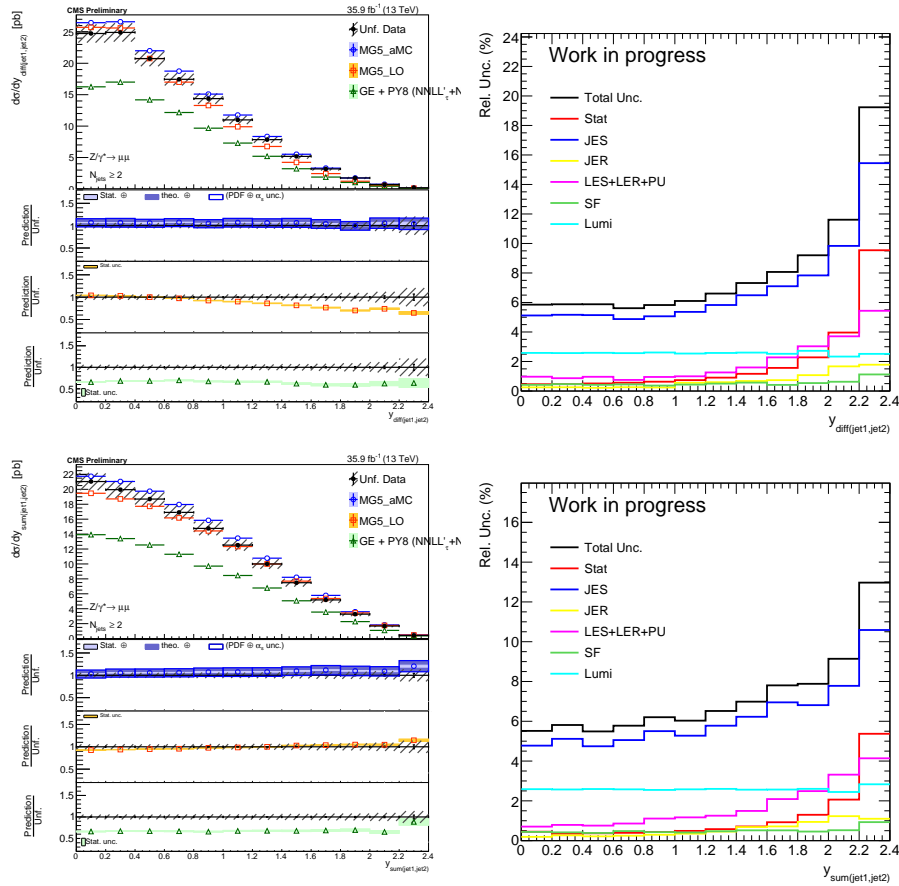


Figure 3-45: The measured differential cross section as a function of the leading and subleading jet rapidity difference(left) and sum(right) with at least two jets. For data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_AMC, LO MG5_AMC, and GENEVA. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_AMC and statistical only for the GENEVA and LO MG5_AMC predictions.

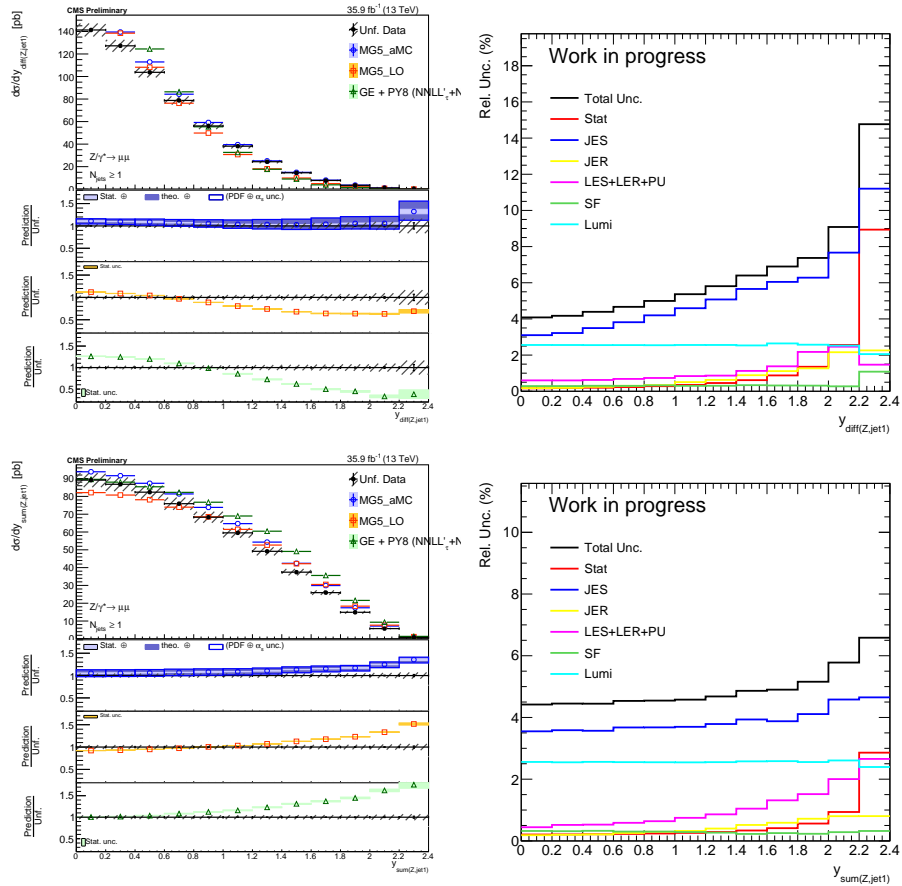


Figure 3-46: The measured differential cross section as a function of the Z boson and leading jet rapidity difference(left) and sum(right) with at least one jet. For data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_AMC, LO MG5_AMC, and GENEVA. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_AMC and statistical only for the GENEVA and LO MG5_AMC predictions.

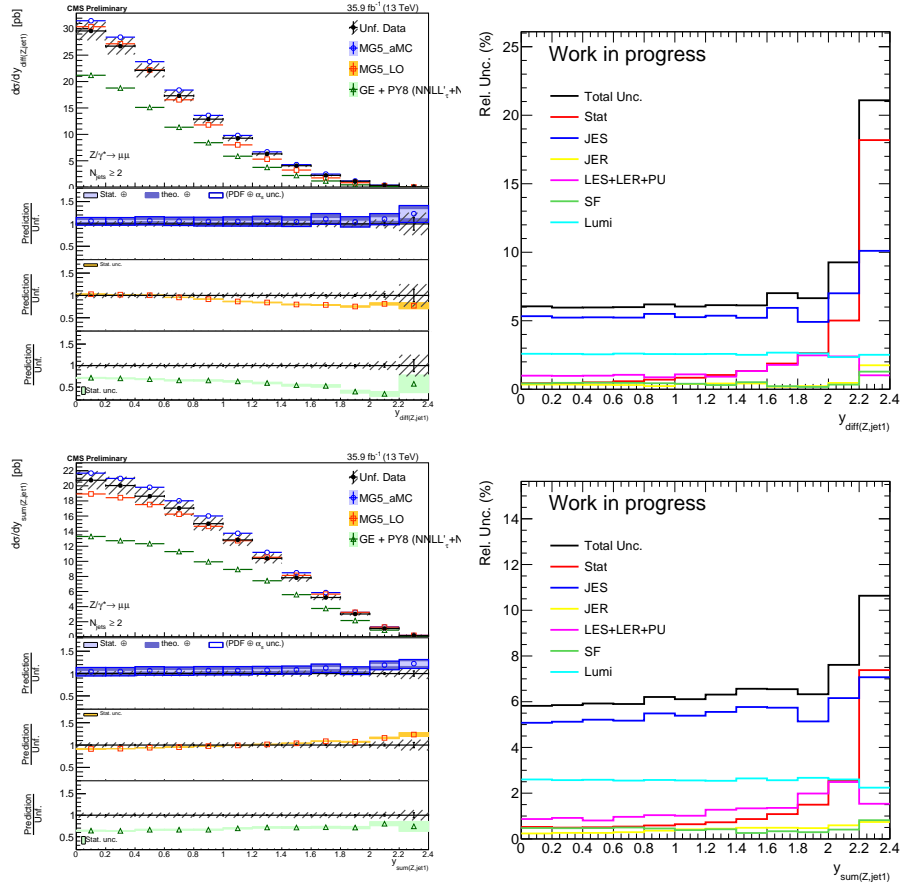


Figure 3-47: The measured differential cross section as a function of the Z boson and leading jet rapidity difference(left) and sum(right) with at least two jets. For data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_AMC, LO MG5_AMC, and GENEVA. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_AMC and statistical only for the GENEVA and LO MG5_AMC predictions.

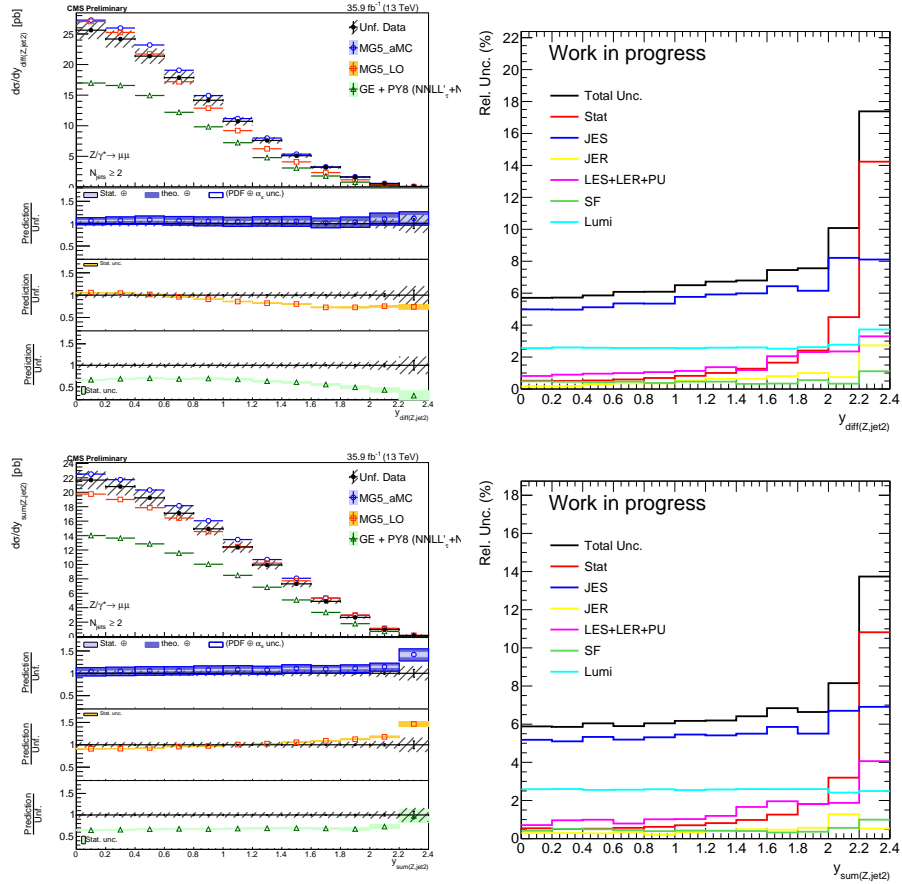


Figure 3-48: The measured differential cross section as a function of the Z boson and subleading jet rapidity difference(left) and sum(right) with at least two jets. For data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_AMC, LO MG5_AMC, and GENEVA. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_AMC and statistical only for the GENEVA and LO MG5_AMC predictions.

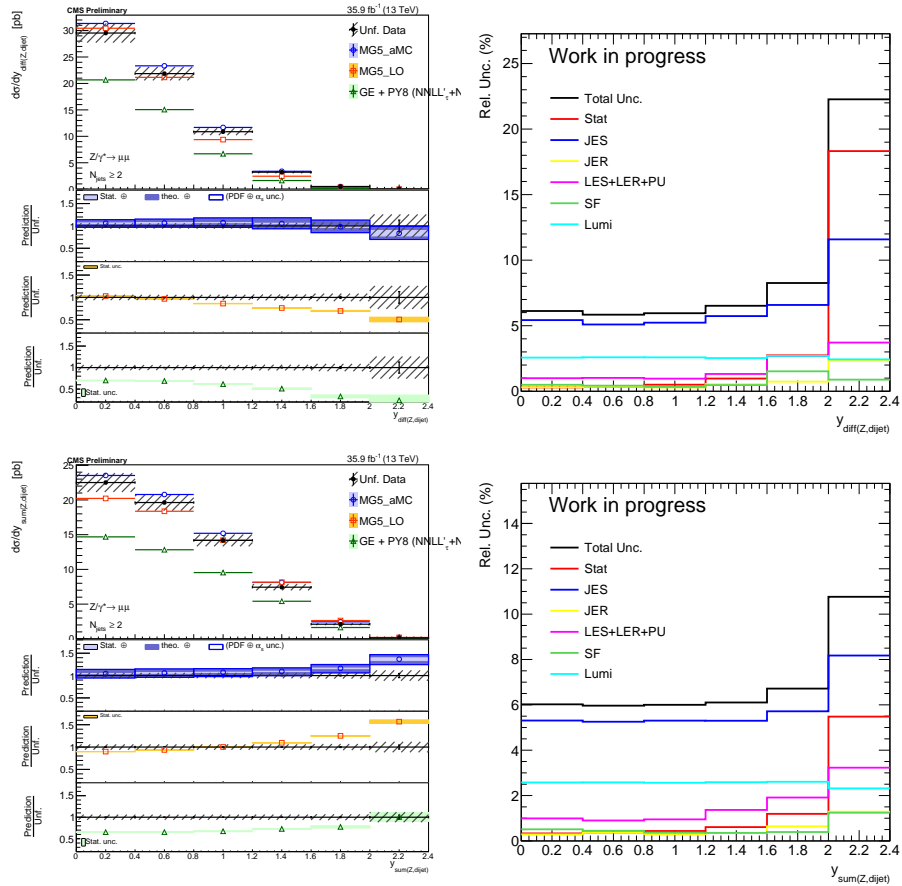


Figure 3-49: The measured differential cross section as a function of the Z boson and dijet rapidity difference(left) and sum(right) with two jets inclusive. For data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_AMC, LO MG5_AMC, and GENEVA. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_AMC and statistical only for the GENEVA and LO MG5_AMC predictions.

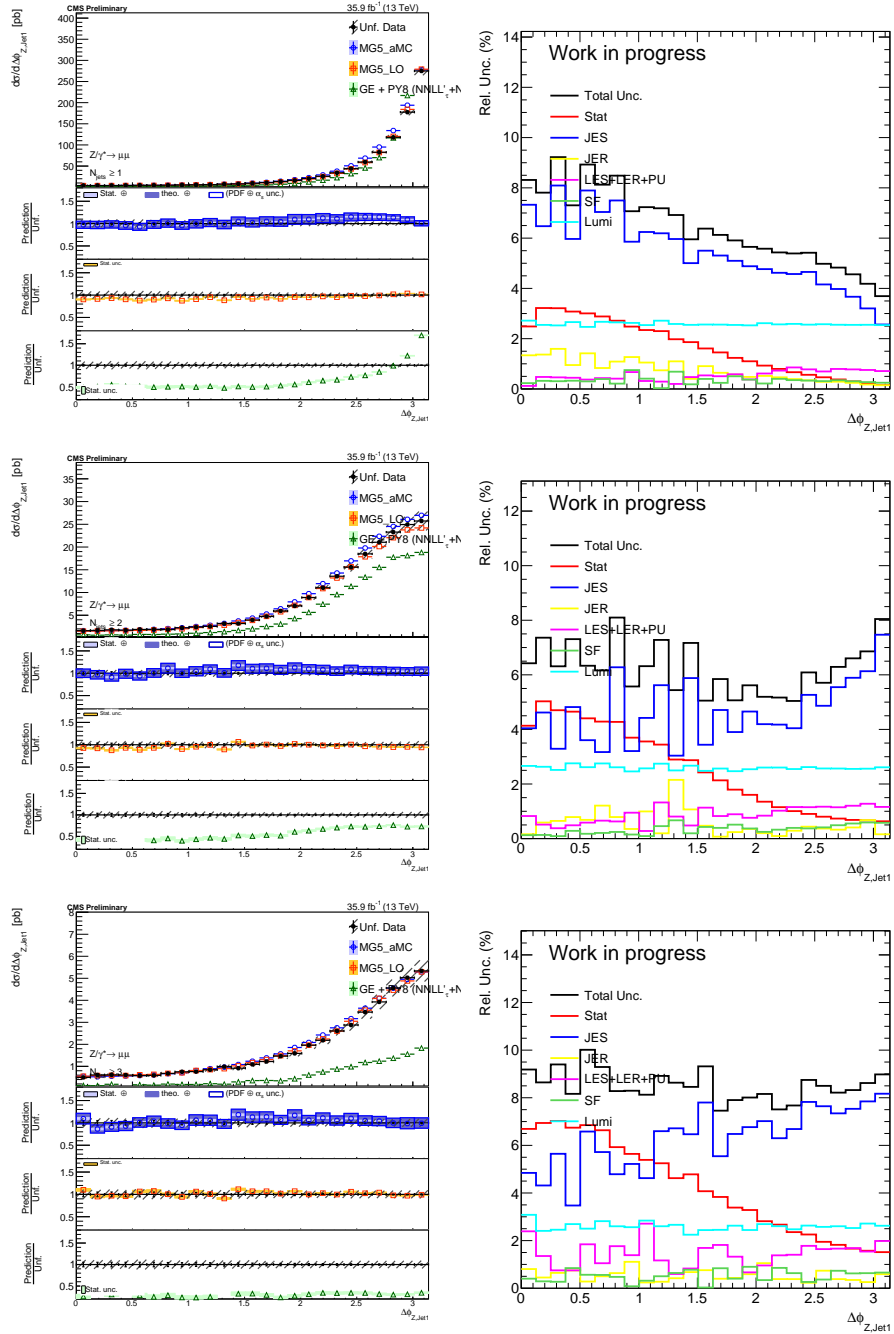


Figure 3-50: The measured differential cross section as a function of the Z boson and leading jet azimuthal difference with at least one(left), two(middle), and three(right) jets. For PU data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_AMC and LO MG5_AMC. The measurements with at least one and two jets is also compared to GENEVA. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_AMC and statistical only for the GENEVA and LO MG5_AMC predictions.

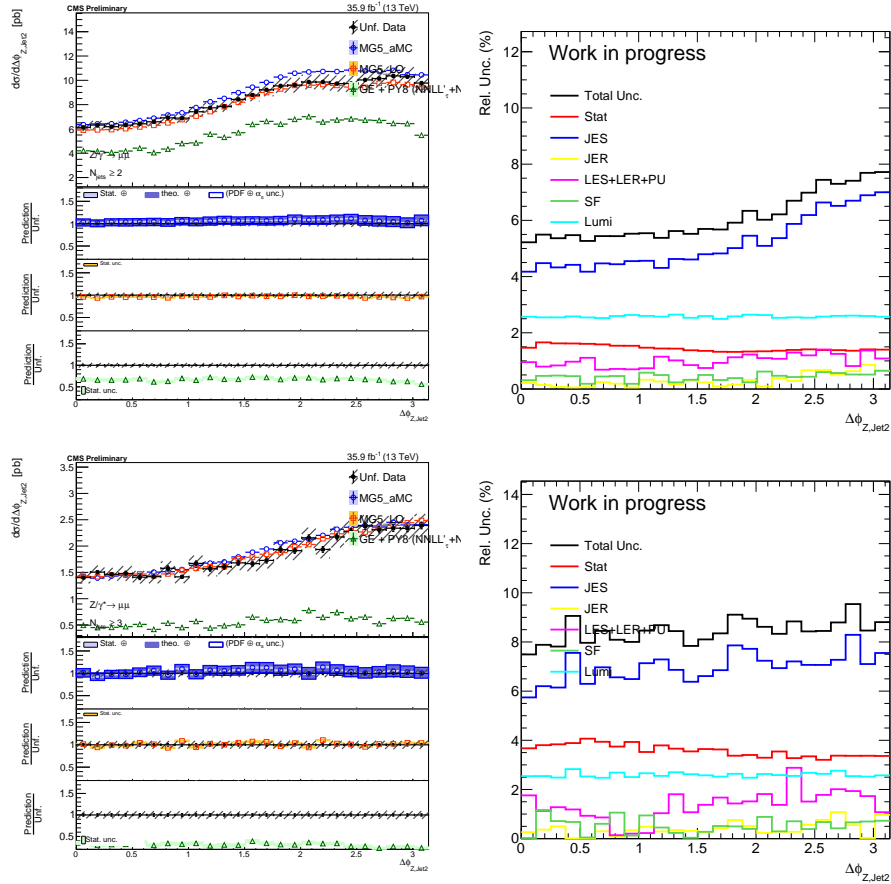


Figure 3-51: The measured differential cross section as a function of the Z boson and subleading jet azimuthal difference with at least two(left) and three(right) jets. For data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_AMC and LO MG5_AMC. The measurement with at least two jets is also compared to GENEVA. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_AMC and statistical only for the GENEVA and LO MG5_AMC predictions.

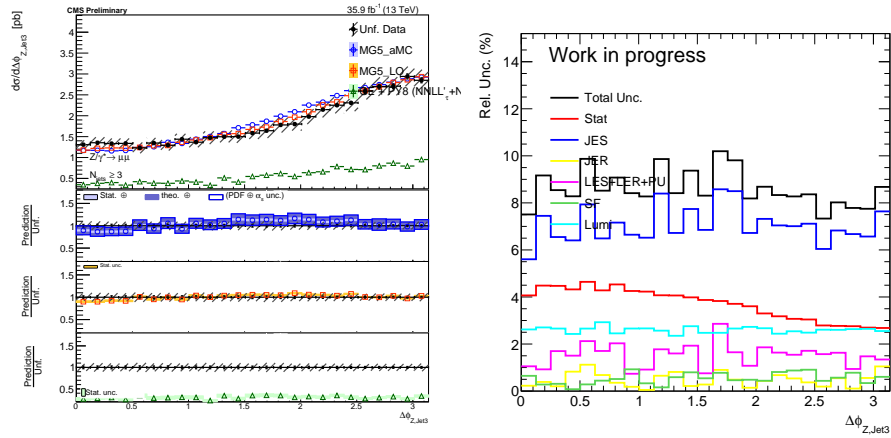


Figure 3.52: The measured differential cross section as a function of the Z boson and third jet azimuthal difference with at least three jets. For data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_AMC and LO MG5_AMC. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_AMC and statistical only for the LO MG5_AMC predictions.

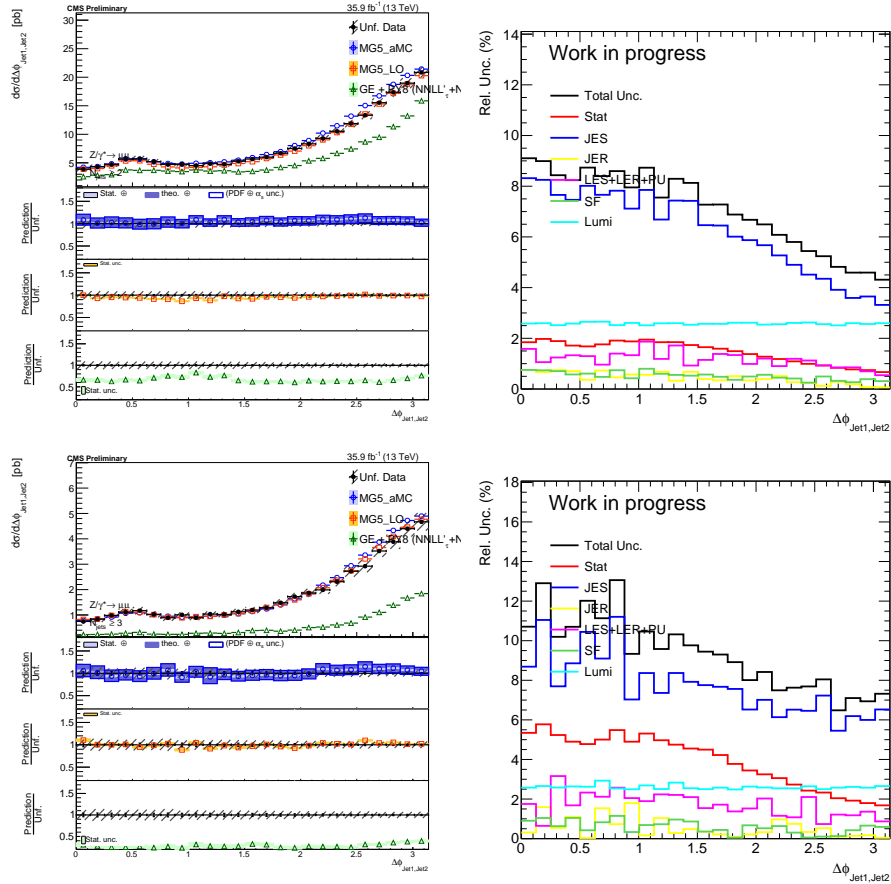


Figure 3-53: The measured differential cross section as a function of the leading and subleading jet azimuthal difference with at least two(left) and three(right) jets. For data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_AMC and LO MG5_AMC. The measurement with at least two jets is also compared to GENEVA. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_AMC and statistical only for the GENEVA and LO MG5_AMC predictions.

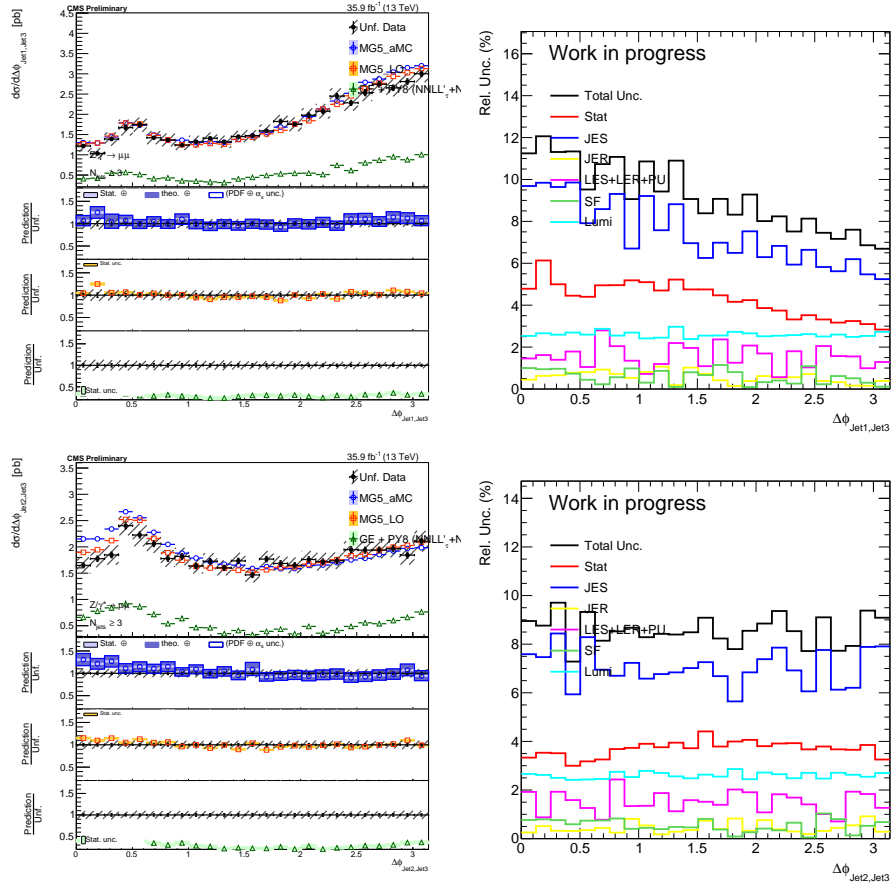


Figure 3-54: The measured differential cross section as a function of the leading and third jet azimuthal difference (left) and subleading and third jet azimuthal difference (right) with at least three jets. For data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_AMC and LO MG5_AMC. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_AMC and statistical only for the LO MG5_AMC predictions.

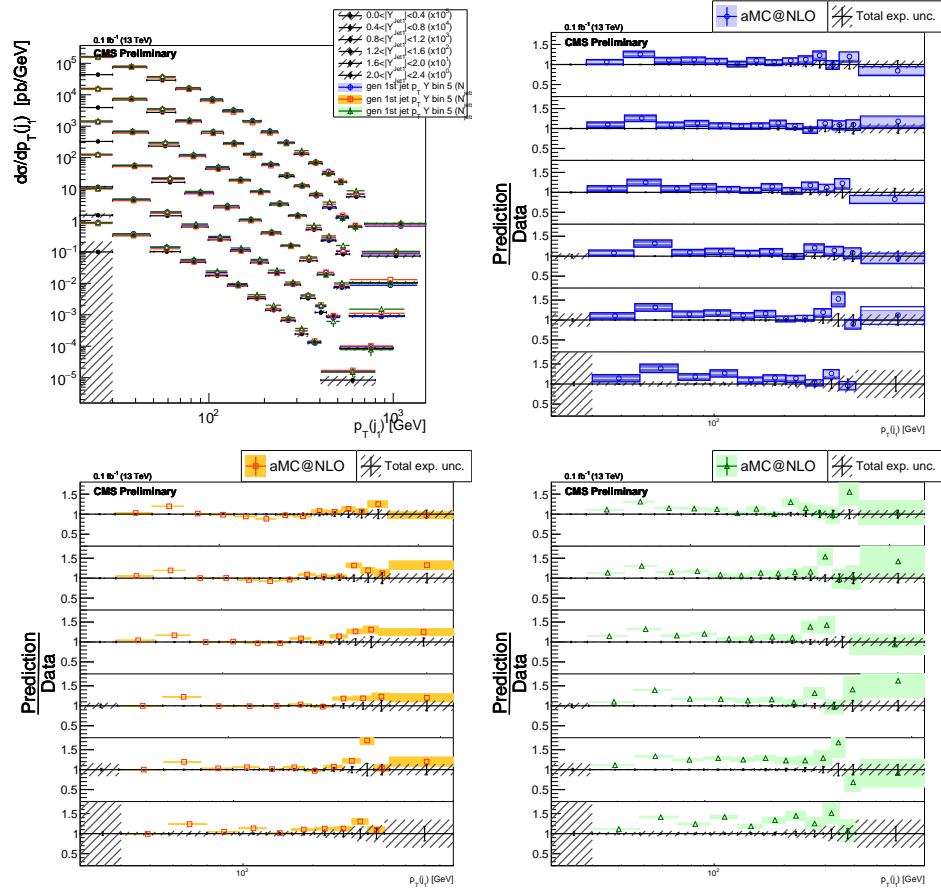


Figure 3-55: Double differential cross section as a function of leading jet p_T and rapidity with at least one jet (upper left). For data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_AMC, LO MG5_AMC, and GENEVA and the ratios are shown in the upper right, lower left, and lower right plots, respectively. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_AMC and statistical only for the GENEVA and LO MG5_AMC predictions.

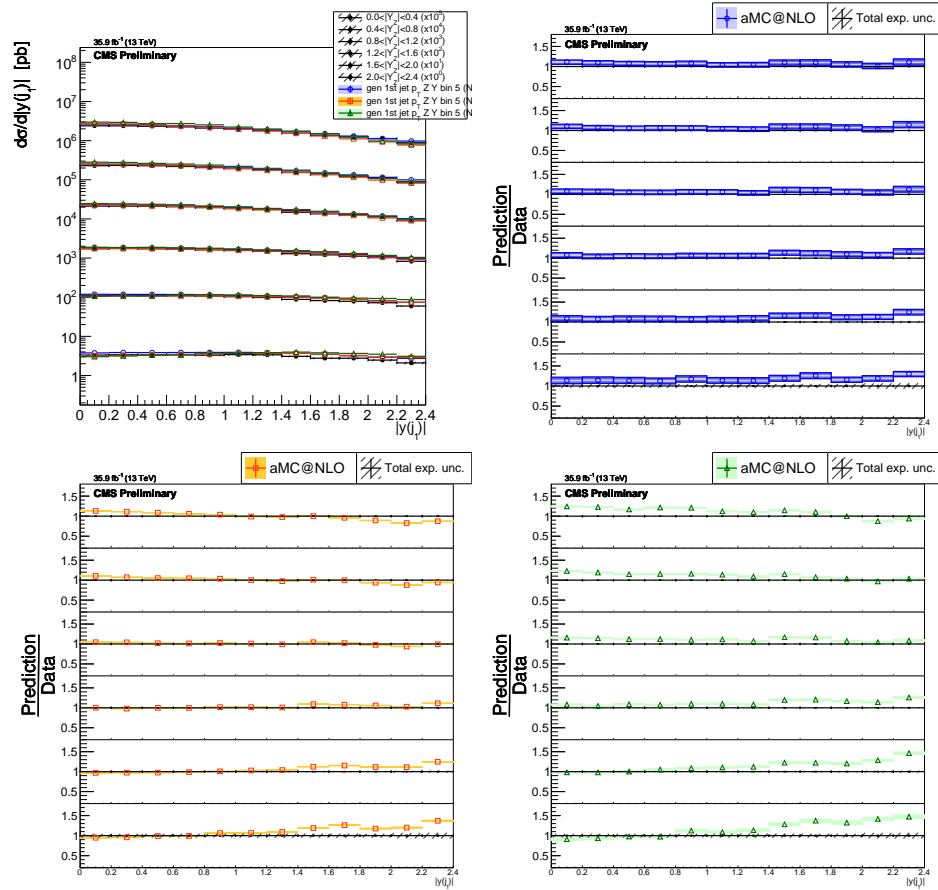


Figure 3-56: Double differential cross section as a function of leading jet and Z boson rapidity with at least one jet. For data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_AMC, LO MG5_AMC, and GENEVA and the ratios are shown in the upper right, lower left, and lower right plots, respectively. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_AMC and statistical only for the GENEVA and LO MG5_AMC predictions.

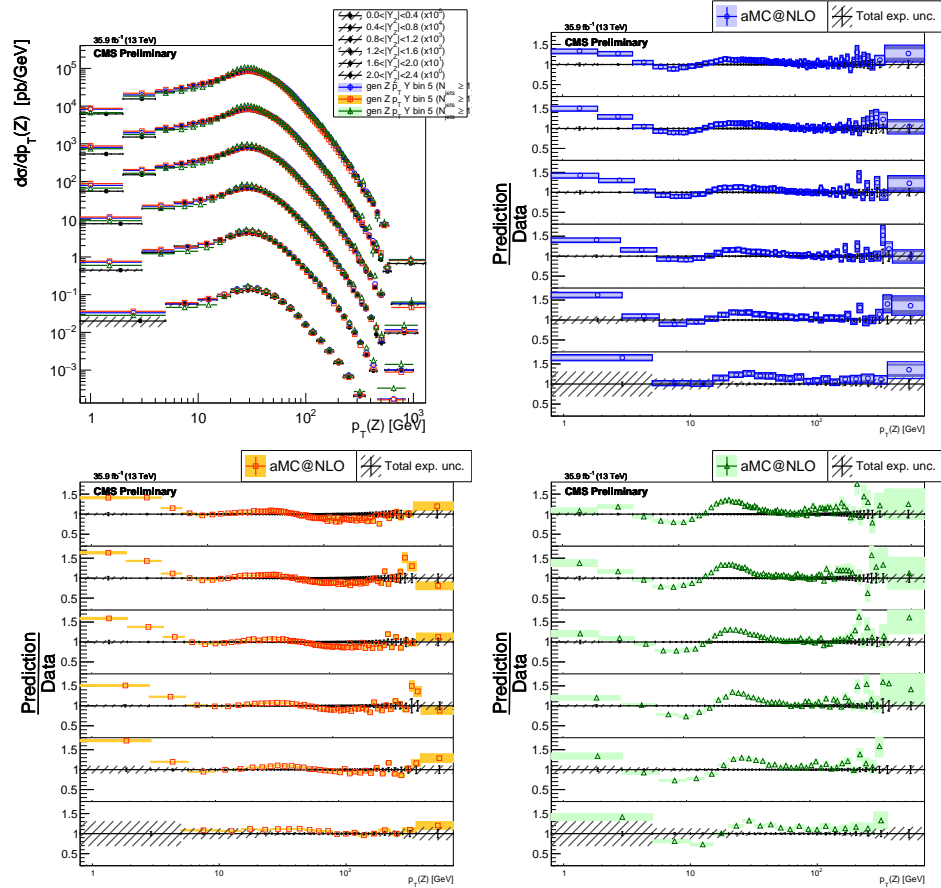


Figure 3.57: Double differential cross section as a function of Z boson p_T and rapidity with at least one jet. For data the black bars show the statistical uncertainty and the hashed area shows the total uncertainty. The measurement is compared to NLO MG5_AMC, LO MG5_AMC, and GENEVA and the ratios are shown in the upper right, lower left, and lower right plots, respectively. The uncertainty for predictions is shown only in the ratio plots with statistical, PDF, and scale uncertainties for the NLO MG5_AMC and statistical only for the GENEVA and LO MG5_AMC predictions.

Table 3.8: Differential cross section in 1st jet $|\eta|$ ($N_{\text{jets}} \geq 1$) and break down of the systematic uncertainties for the combination of both decay channels. Uncertainties are in percent.

$ y(j_1) $	$\frac{d\sigma}{d y(j_1) }$ [pb]	Tot	stat	JES	JER	Eff	Lumi	XSec	PU	LES+LER	Unf stat
0 – 0.2	64.1	4.1	0.20	3.0	0.018	0.90	2.6	0.031	0.58	0.42	0.11
0.2 – 0.4	64.4	4.0	0.19	2.9	0.032	0.90	2.5	0.026	0.55	0.40	0.11
0.4 – 0.6	63.4	4.1	0.19	3.0	0.053	0.91	2.6	0.028	0.60	0.44	0.11
0.6 – 0.8	61.2	4.2	0.19	3.1	0.099	0.94	2.6	0.030	0.55	0.41	0.11
0.8 – 1	58.0	4.1	0.20	2.9	0.12	0.93	2.5	0.025	0.54	0.39	0.12
1 – 1.2	54.7	4.9	0.21	4.0	0.25	0.98	2.6	0.027	0.62	0.45	0.13
1.2 – 1.4	50.7	5.2	0.23	4.2	0.28	1.0	2.6	0.034	0.79	0.58	0.14
1.4 – 1.6	43.9	5.2	0.25	4.2	0.16	1.1	2.6	0.039	1.1	0.77	0.15
1.6 – 1.8	39.1	5.0	0.27	3.7	0.51	0.99	2.5	0.024	1.5	1.1	0.17
1.8 – 2	35.3	5.6	0.29	4.2	0.95	1.0	2.6	0.045	1.8	1.3	0.18
2 – 2.2	31.0	6.7	0.33	5.5	1.2	1.1	2.6	0.035	1.8	1.3	0.20
2.2 – 2.4	24.4	8.5	0.43	7.3	1.6	1.3	2.5	0.023	2.3	1.6	0.27

Table 3.9: Differential cross section in 1st jet p_T ($N_{\text{jets}} \geq 1$) and break down of the systematic uncertainties for the combination of both decay channels. Uncertainties are in percent.

$p_T(j_1)$ [GeV]	$\frac{d\sigma}{dp_T(j_1)}$ [$\frac{\text{pb}}{\text{GeV}}$]	Tot	stat	JES	JER	Eff	Lumi	XSec	PU	LES+LER	Unf stat
30 – 47	3.25	3.7	0.18	2.5	0.18	0.098	2.5	0.017	0.62	0.45	0.12
47 – 69	1.42	3.8	0.21	2.5	0.10	0.87	2.5	0.027	0.71	0.51	0.15
69 – 96	0.605	4.0	0.29	2.9	0.14	0.60	2.5	0.027	0.37	0.27	0.20
96 – 128	0.251	4.1	0.40	3.0	0.13	0.77	2.6	0.038	0.56	0.42	0.27
128 – 166	0.106	3.8	0.55	2.6	0.075	0.80	2.6	0.044	0.49	0.36	0.38
166 – 210	0.0463	4.3	0.78	3.2	0.13	0.95	2.5	0.024	0.46	0.33	0.53
210 – 261	0.0194	4.1	1.1	2.6	0.21	0.84	2.6	0.082	0.31	0.29	0.74
261 – 319	0.00832	4.6	1.6	3.2	0.24	0.80	2.6	0.068	0.32	0.28	1.1
319 – 386	0.00357	4.7	2.3	2.5	0.37	0.80	2.7	0.099	0.38	0.33	1.5
386 – 460	0.00157	5.8	3.3	3.2	0.15	0.95	2.6	0.034	0.27	0.23	2.2
460 – 544	0.000652	6.8	4.9	1.3	0.23	0.90	2.9	0.054	0.54	0.42	3.3
544 – 638	0.000298	11.	7.1	7.1	0.32	0.54	2.4	0.059	0.92	0.81	4.8
638 – 751	0.000128	13.	10.	2.5	0.53	0.73	2.7	0.075	1.9	1.8	7.1
751 – 870	6.08e-05	19.	14.	6.6	0.23	0.96	2.5	0.024	1.3	1.3	10.
870 – 1500	7.73e-06	20.	15.	6.5	0.46	0.49	2.6	0.077	1.6	1.3	11.

Table 3.10: Differential cross section in exclusive jet multiplicity and break down of the systematic uncertainties for the combination of both decay channels. Uncertainties are in percent.

N_{jets}	$\frac{d\sigma}{dN_{\text{jets}}}$ [pb]	Tot	stat	JES	JER	Eff	Lumi	XSec	PU	LES+LER	Unf stat
= 0	580.	4.0	0.041	0.82	0.047	0.69	3.7	0.032	0.67	0.70	0.020
= 1	90.9	7.7	0.14	6.0	0.48	1.1	4.2	0.048	1.5	1.5	0.074
= 2	20.7	9.2	0.32	7.7	0.44	1.3	4.2	0.061	1.8	1.8	0.17
= 3	4.35	12.	0.81	10.	0.56	1.6	4.4	0.073	2.4	2.4	0.44
= 4	1.03	13.	1.9	11.	0.68	2.1	4.3	0.076	2.7	2.7	1.1
= 5	0.223	16.	4.7	13.	1.1	2.2	4.7	0.14	2.7	2.8	2.7
= 6	0.0586	22.	10.	16.	0.53	0.99	4.3	0.12	4.7	5.2	6.1
= 7	0.0128	41.	28.	20.	2.3	3.4	4.0	0.046	6.9	8.7	17.
= 8	0.00251	61.	50.	15.	2.2	1.1	2.5	0.019	1.1	2.2	30.

Table 3.11: Differential cross section in p_T^Z and break down of the systematic uncertainties for the combination of both decay channels. Uncertainties are in percent.

$p_T(Z)$ [GeV]	$\frac{d\sigma}{dp_T(Z)}$ [$\frac{\text{pb}}{\text{GeV}}$]	Tot	stat	JES	JER	Eff	Lumi	XSec	PU	LES+LER	Unf stat
0.8 - 2	0.0725	26.	3.7	25.	4.0	4.1	2.6	0.066	2.3	1.9	1.6
2 - 4	0.189	22.	1.5	21.	3.0	4.4	2.6	0.055	1.7	1.4	0.72
4 - 6	0.307	16.	1.0	16.	2.3	3.4	2.6	0.054	0.71	0.53	0.51
6 - 8	0.405	14.	0.82	13.	1.9	2.6	2.6	0.060	0.29	0.23	0.45
8 - 10	0.490	12.	0.73	12.	1.6	2.2	2.6	0.068	0.31	0.30	0.41
10 - 12	0.574	12.	0.65	11.	1.6	2.0	2.6	0.049	0.65	0.49	0.38
12 - 14	0.651	12.	0.61	11.	1.4	1.9	2.6	0.057	1.0	0.75	0.36
14 - 16	0.728	12.	0.60	11.	1.4	1.7	2.6	0.066	1.2	0.89	0.36
16 - 19	0.815	12.	0.58	12.	1.5	1.8	2.6	0.060	1.1	0.81	0.35
19 - 22	0.932	13.	0.57	12.	1.5	1.9	2.6	0.076	1.4	1.0	0.34
22 - 25	1.07	13.	0.52	12.	1.4	2.0	2.6	0.060	1.5	1.1	0.30
25 - 28	1.25	12.	0.46	11.	1.2	1.9	2.6	0.043	1.5	1.1	0.27
28 - 31	1.46	10.	0.40	9.4	0.80	1.6	2.6	0.037	1.5	1.1	0.24
31 - 34	1.66	8.2	0.36	7.5	0.49	1.3	2.6	0.035	1.5	1.0	0.22
34 - 37	1.82	6.8	0.33	6.0	0.36	1.1	2.6	0.030	1.3	0.95	0.21
37 - 40	1.91	5.7	0.32	4.7	0.19	0.92	2.6	0.035	1.2	0.89	0.20
40 - 43	1.91	4.9	0.32	3.8	0.12	0.87	2.6	0.031	1.1	0.77	0.20
43 - 46	1.87	4.4	0.32	3.2	0.097	0.79	2.5	0.021	1.1	0.76	0.21
46 - 49	1.79	4.0	0.36	2.6	0.047	0.69	2.5	0.026	1.1	0.77	0.24
49 - 53	1.67	3.6	0.34	2.1	0.035	0.66	2.6	0.031	0.96	0.68	0.22
53 - 57	1.52	3.3	0.38	1.6	0.027	0.63	2.5	0.015	0.84	0.60	0.25
57 - 61	1.37	3.1	0.40	1.3	0.022	0.62	2.5	0.022	0.78	0.55	0.27
61 - 65	1.23	3.0	0.43	1.0	0.016	0.56	2.5	0.016	0.79	0.57	0.29
65 - 69	1.07	2.9	0.48	0.86	0.017	0.52	2.5	0.023	0.68	0.49	0.32
69 - 73	0.955	2.9	0.52	0.72	0.015	0.46	2.5	0.022	0.62	0.44	0.35
73 - 78	0.811	2.8	0.47	0.58	0.015	0.55	2.5	0.018	0.65	0.48	0.32
78 - 83	0.691	2.8	0.53	0.46	0.027	0.47	2.6	0.040	0.63	0.46	0.35
83 - 88	0.580	2.8	0.58	0.40	0.019	0.51	2.5	0.026	0.61	0.44	0.39
88 - 93	0.488	2.8	0.64	0.33	0.016	0.50	2.5	0.0073	0.50	0.38	0.43
93 - 99	0.411	2.8	0.60	0.25	0.0093	0.53	2.6	0.033	0.51	0.37	0.40
99 - 105	0.342	2.8	0.66	0.25	0.0022	0.48	2.5	0.015	0.58	0.42	0.44
105 - 111	0.286	2.8	0.74	0.23	0.028	0.56	2.5	0.027	0.36	0.27	0.49
111 - 118	0.236	2.8	0.72	0.15	0.023	0.55	2.6	0.018	0.54	0.39	0.48
118 - 125	0.198	2.8	0.79	0.16	0.015	0.64	2.5	0.0044	0.43	0.32	0.53
125 - 133	0.162	2.8	0.79	0.11	0.018	0.60	2.6	0.014	0.24	0.19	0.53
133 - 141	0.132	2.9	0.90	0.10	0.022	0.65	2.6	0.028	0.40	0.34	0.60
141 - 150	0.109	2.9	0.90	0.096	0.016	0.71	2.6	0.028	0.32	0.35	0.60
150 - 160	0.0843	3.0	0.95	0.054	0.011	0.72	2.6	0.037	0.48	0.34	0.63
160 - 171	0.0673	2.9	1.0	0.083	0.017	0.70	2.5	0.0039	0.30	0.29	0.65
171 - 183	0.0510	3.0	1.1	0.071	0.018	0.76	2.5	0.0020	0.50	0.40	0.71
183 - 197	0.0394	3.0	1.1	0.065	0.012	0.72	2.6	0.035	0.37	0.29	0.73
197 - 212	0.0295	3.0	1.3	0.053	0.012	0.79	2.5	0.0043	0.25	0.21	0.82
212 - 228	0.0217	3.2	1.4	0.039	0.040	0.73	2.6	0.056	0.28	0.25	0.92
228 - 246	0.0161	3.3	1.5	0.024	0.066	0.82	2.6	0.063	0.34	0.40	0.98
246 - 266	0.0112	3.4	1.7	0.014	0.028	0.81	2.6	0.029	0.091	0.13	1.1
266 - 289	0.00793	3.5	1.8	0.014	0.025	0.65	2.6	0.024	0.32	0.36	1.2
289 - 314	0.00552	3.7	2.1	0.029	0.016	0.82	2.5	0.010	0.48	0.42	1.3
314 - 344	0.00355	3.9	2.4	0.038	0.036	0.60	2.6	0.023	0.37	0.31	1.5
344 - 377	0.00226	4.3	2.8	0.044	0.032	0.81	2.6	0.048	0.27	0.36	1.7
377 - 418	0.00139	4.5	3.1	0.021	0.024	0.85	2.5	0.0073	0.59	0.50	1.9
418 - 460	0.000820	5.5	4.0	0.074	0.025	0.86	2.7	0.051	0.20	0.15	2.4
460 - 511	0.000452	6.3	4.8	0.031	0.014	0.92	2.5	0.014	0.96	0.72	2.8
511 - 567	0.000250	7.7	6.2	0.012	0.0059	0.15	2.5	0.0061	0.80	0.79	3.7
567 - 655	0.000135	8.7	6.9	0.016	0.012	0.59	2.5	0.012	0.65	0.59	4.5
655 - 1300	1.40e-05	11.	8.9	0.046	0.013	0.26	2.5	0.013	1.6	1.9	4.4

Table 3.12: Differential cross section in $|y_Z|$ ($N_{\text{jets}} \geq 1$) and break down of the systematic uncertainties for the combination of both decay channels. Uncertainties are in percent.

$ y_Z $	$\frac{d\sigma}{d y_Z }$ [pb]	Tot	stat	JES	JER	Eff	Lumi	XSec	PU	LES+LER	Unf stat
0 – 0.2	66.8	4.7	0.18	3.7	0.32	0.75	2.5	0.021	0.82	0.60	0.099
0.2 – 0.4	66.5	4.6	0.18	3.6	0.29	0.76	2.6	0.030	0.81	0.58	0.098
0.4 – 0.6	65.9	4.6	0.18	3.6	0.29	0.81	2.6	0.030	0.79	0.56	0.10
0.6 – 0.8	65.0	4.6	0.18	3.6	0.28	0.90	2.5	0.026	0.84	0.60	0.10
0.8 – 1	63.5	4.7	0.18	3.6	0.29	1.0	2.6	0.029	0.87	0.62	0.11
1 – 1.2	60.8	4.7	0.19	3.7	0.27	1.1	2.6	0.032	0.87	0.62	0.11
1.2 – 1.4	56.7	4.8	0.20	3.8	0.30	1.2	2.6	0.028	0.88	0.63	0.11
1.4 – 1.6	50.0	4.8	0.22	3.8	0.31	1.0	2.6	0.030	0.91	0.64	0.13
1.6 – 1.8	41.1	4.9	0.24	3.7	0.28	1.4	2.6	0.042	0.97	0.69	0.14
1.8 – 2	30.2	4.8	0.28	3.7	0.29	1.2	2.6	0.040	1.0	0.72	0.16
2 – 2.2	18.3	4.8	0.36	3.6	0.29	0.99	2.6	0.042	1.1	0.79	0.20
2.2 – 2.4	5.72	5.0	0.68	3.6	0.21	1.1	2.6	0.029	1.4	1.0	0.38

3.10 Summary

The production of a Z boson decaying into two charged leptons in association with jets are measured in LHC proton-proton collisions at center-of-mass energy of 13 TeV with the CMS experiment, using data sets corresponding to an integrated luminosity of 35.9 fb^{-1} . Differential cross sections are measured for a Z boson decaying to electrons or muons with $p_T > 25 \text{ GeV}$ and $|\eta| < 2.4$ with at least one jet with $p_T > 30 \text{ GeV}$ and $|\eta| < 2.4$.

The cross section has been measured as functions of the exclusive and inclusive jet multiplicities up to 8, of the transverse momentum of the Z boson, jet kinematic variables including jet transverse momenta, the scalar sum of jet transverse momenta, and the jet rapidity for inclusive jet multiplicities up to five jets.

The results, corrected for all detector effects by means of regularized unfolding, have been compared with three different calculations. Two predictions are particle-level simulations where one uses multileg NLO predictions using the FxFx merging scheme and the other uses multileg LO predictions and the MLM matching scheme. The third calculation is the GENEVA MC program, where an NNLO calculation for Drell-Yan production is combined with higher-order resummation.

High precision has been achieved in the CMS measurements of cross sections and their ratios using latest experimental methods and larger data sets. Indeed, the larger data sets allow to extend the kinematic range of cross section measurements to higher values of p_T and mass, as well as open the possibility to investigate rare processes not yet observed. The measurements presented in this letter provide a detailed description of the topological structure of $Z(\rightarrow \ell^+ \ell^-) + \text{jets}$ production that is complementary to existing measurements of rates and associated jet multiplicities.

In summary the measured cross sections are generally described well by the predictions within the experimental and theoretical uncertainties. The predictions describe

the jet multiplicity within the uncertainties, with increasing deviations observed for jet multiplicities beyond three. A general agreement is observed for the distribution of the considered jet quantities. However, some deviations from data are seen.

The results suggest using multi-parton NLO calculation for the estimation of Z+jets contribution at the LHC in measurements and searches, and its associated uncertainty.

Chapter 4

Conclusions

With the strong performance of the Hadron Calorimeter and the DAQ, lead by the AMC13, the most precise measurements of the Z in association with jets at 13 TeV has been presented. The differential cross section are measured as a function of jet p_T , $|\eta|$, Z p_T , $|\eta|$, and numerous others. In order to relate these measurements to the strong interaction and progress the understanding of the theory many models are compared to the cross sections. Madgraph LO, Madgraph NLO, powheg, fixed order are compared. Overall the Madgraph NLO has proved to be a powerful model in describing phenomena at the LHC.

References

- (2017). CMS Luminosity Measurements for the 2016 Data Taking Period. Technical Report CMS-PAS-LUM-17-001, CERN, Geneva.
- Alioli, S., Nason, P., Oleari, C., and Re, E. (2010). A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043.
- Allison, J. et al. (2006). GEANT4 developments and applications. *IEEE Trans. Nucl. Sci.*, 53:270.
- Allwall, J., de Visscher, S., and Maltoni, F. (2009). QCD radiation in the production of heavy colored particles at the LHC. *JHEP*, 02:017.
- Allwall, J. et al. (2008). Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions. *Eur. Phys. J.*, C53:473–500.
- Allwall, J., Frederix, R., Frixione, S., Hirschi, V., Maltoni, F., Mattelaer, O., Shao, H.-S., Stelzer, T., Torielli, P., and Zaro, M. (2014). The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079.
- Ball, R. D., Bertone, V., Carrazza, S., Deans, C. S., Del Debbio, L., Forte, S., Guffanti, A., Hartland, N. P., Latorre, J. I., Rojo, J., and Ubiali, M. (2013). Parton distributions with LHC data. *Nucl. Phys. B*, 867:244.
- Ball, R. D. et al. (2015). Parton distributions for the LHC Run II. *JHEP*, 04:040.
- Bodek, A., van Dyne, A., Han, J. Y., Sakumoto, W., and Strelnikov, A. (2012). Extracting Muon Momentum Scale Corrections for Hadron Collider Experiments. *Eur. Phys. J.*, C72:2194.
- Cacciari, M., Salam, G. P., and Soyez, G. (2008). The anti- k_t jet clustering algorithm. *JHEP*, 04:063.
- Cacciari, M., Salam, G. P., and Soyez, G. (2008). The anti- k_t jet clustering algorithm. *Journal of High Energy Physics*, 2008(4):063.
- Cacciari, M., Salam, G. P., and Soyez, G. (2012). FastJet user manual. *Eur. Phys. J. C*, 72:1896.
- Calvetti, D., Reichel, L., and Shuibi, A. (2004). L-Curve and Curvature Bounds for Tikhonov Regularization. *Numerical Algorithms*, 35:301–314.

- Campbell, J. M., Ellis, R. K., and Williams, C. (2011). Vector boson pair production at the LHC. *JHEP*, 07:018.
- Chatrchyan, S. et al. (2011). Measurement of the Inclusive W and Z Production Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV. *JHEP*, 10:132.
- Chatrchyan, S. et al. (2012). Performance of CMS Muon Reconstruction in pp Collision Events at $\sqrt{s} = 7$ TeV. *JINST*, 7:P10002.
- CMS Collaboration (2015). Event generator tunes obtained from underlying event and multiparton scattering measurements. *arXiv e-prints*, page arXiv:1512.00815.
- Collaboration, C. (2014). Pileup Removal Algorithms.
- Frederix, R. and Frixione, S. (2012). Merging meets matching in MC@NLO. *Journal of High Energy Physics*, 2012:61.
- Frederix, R. and Frixione, S. (2012). Merging meets matching in MC@NLO. *JHEP*, 12:061.
- Frederix, R., Frixione, S., Hirschi, V., Maltoni, F., Pittau, R., and Torrielli, P. (2012). Four-lepton production at hadron colliders: aMC@NLO predictions with theoretical uncertainties. *JHEP*, 02:099.
- Frixione, S., Nason, P., and Oleari, C. (2007a). Matching NLO QCD computations with parton shower simulations: the POWHEG method. *JHEP*, 11:070.
- Frixione, S., Nason, P., and Ridolfi, G. (2007b). A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction. *JHEP*, 09:126.
- Khachatryan, V. et al. (2015). Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV. *JINST*, 10(06):P06005.
- Khachatryan, V. et al. (2016). Event generator tunes obtained from underlying event and multiparton scattering measurements. *Eur. Phys. J.*, C76(3):155.
- Lyons, L., Gibaut, D., and Clifford, P. (1988). How to Combine Correlated Estimates of a Single Physical Quantity. *Nucl. Instrum. Meth.*, A270:110.
- Nason, P. (2004). A new method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040.
- Schmitt, S. (2012). TUnfold, an algorithm for correcting migration effects in high energy physics. *Journal of Instrumentation*, 7(10):T10003–T10003.
- Sirunyan, A. M. et al. (2017). Particle-flow reconstruction and global event description with the cms detector. *JINST*, 12:P10003.

- Sjöstrand, T., Ask, S., Christiansen, J. R., Corke, R., Desai, N., Ilten, P., Mrenna, S., Prestel, S., Rasmussen, C. O., and Skands, P. Z. (2015). An introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159.
- Sjöstrand, T., Mrenna, S., and Skands, P. (2008). A brief introduction to PYTHIA 8.1. *Computer Physics Communications*, 178(11):852–867.
- Tikhonov, I. V. (1963). Solution of incorrectly formulated problems and the regularization method. *Soviet Mathematics*, 4(1):1023–1038.
- Valassi, A. (2003). Combining correlated measurements of several different physical quantities. *Nucl. Instrum. Meth.*, A500:391–405.

CURRICULUM VITAE

Daniel J. Arcaro

Boston University
Physics Department
590 Commonwealth Ave.
Boston, MA 02215

Email: djarcaro@bu.edu
Phone: 508-479-8049

EDUCATION

Physics Graduate Program, Boston University, 2013-Present
B.A. Physics, Boston University, 2013

EMPLOYMENT

Boston University - Physics Graduate Program

2013-Present Research Assistant, Boston University CMS Group

2013 Teaching Fellow, General Physics

2013-2014 Teaching Fellow, Advanced Laboratory

Boston University - College of Arts and Sciences

2011-2013 Research Assistant, Multi-functional Materials Spectroscopy Laboratory

PUBLICATIONS AND PRESENTATIONS

2018 “Electroweak and QCD aspects in V +jets” Pheno2018: Phenomenology 2018 Symposium, Pittsburgh, PA

2019 “Precision Measurement of Z +jets” DPF Grad Slam, APS April 2019, Denver, CO

2019 “Electroweak and QCD aspects in V +jets” APS April 2019, Denver, CO

TEACHING EXPERIENCE

Boston University

2014 *General Physics*

Teaching Assistant, BU PY211 Calculus-based general physics for scientists and engineers. Teaching three hour lab sections per week. Guiding student in successful completion of experiments. Teaching both fundamental concepts and advising hands-on troubleshooting skills. Topics include Newtonian mechanics, basic fluid mechanics, and thermal physics.

2013, 2014 *Advanced Laboratory*

Teaching Assistant, BU PY581 Advanced Laboratory. Teaching one eight-hour lab section per week. Assisting students on classical experiments in atomic and nuclear physics. Improving existing apparatuses, formulating new experiments, and guiding students' research projects. Experiments include magnetic resonance, nuclear-decay studies, Zeeman effect, holography, black-body radiation, cosmic ray detection, muon lifetime measurement, and X-ray diffraction.

RESEARCH AND LABORATORY EXPERIENCE

Boston University Physics Department

High Energy Physics Group, CMS Experiment

Developed software for monitoring the data integrity of the AMC13 data acquisition card used in uTCA crates for CMS. This software parsed output data to track errors in both AMC13 formats and specific sub-detectors including the Hadron Calorimeter and Muon detectors.

Worked on site at CERN as a 24/7 on call expert to diagnose and fix the Hadron Calorimeter data acquisition errors during data taking for CMS. Tested and integrated AMC13 firmware and software updates into the Hadron Calorimeter data acquisition system.

Boston University Physics Department

Prof. Lawrence Sulak Advanced Laboratory

Measured the lifetime of a muon using a water Cherenkov chamber equipped with photomultiplier tubes and fast electronics.

Resolved the angular distribution of cosmic muons and measured thermal neutron scattering off Helium-3 using a state of the art time projection chamber.

Gained experience with high-temperature superconductors and optical techniques including optical pumping and Faraday rotation measurements.

Boston University Physics Department

Prof. Richard Averitt Multi-functional Materials Spectroscopy Laboratory

Developed a high field Terahertz spectroscopy apparatus to investigate photo-induced phase transitions in correlated electron material such as Vanadium Oxide.

Collaborated with colleagues from the condensed matter theory group at Boston University to investigate phase transition in 5CB Liquid Crystals. The time dynamics of the first and second order transitions was measured and compared to computer simulated Ising model magnetic transitions.

PROCEDURAL AND EQUIPMENT EXPERIENCE

Extensive experience involving Terahertz spectroscopy techniques, class IV lasers, optical equipment, and electronics.

Proficient with analog, digital circuitry, and fast electronics.

Advanced programming skills in C++, C, MATLAB, Python, and LabView

Experience with the VHDL language and Xilinx FPGAs

AWARDS RECEIVED

2011 Undergraduate Research Opportunities Program at Boston University
Summer Funding recipient

2011-2013 Dean's List at Boston University

2013 Graduated Cum Laude at Boston University

2016 Department of Physics Teaching Fellow of the Year Award

REFERENCES

Dr. Lawrence Sulak
Professor, Boston University
Phone: 617-353-9454, Email: sulak@physics.bu.edu

Dr. Richard Averitt
Professor, UC San Diego
Email: raveritt@ucsd.edu