



**STUDY OF $Z + b$ -JETS PRODUCTION AT LHC ENERGIES USING
CMS DETECTOR**

A THESIS

**Submitted to the
PANJAB UNIVERSITY, CHANDIGARH**

For the Award of

DOCTOR OF PHILOSOPHY

2022

in

FACULTY OF SCIENCE

**BY
MEENA**

GUIDE

Prof. SUMAN BALA BERI

Dr. SUNIL BANSAL

**DEPARTMENT OF PHYSICS
PANJAB UNIVERSITY
CHANDIGARH, INDIA**



November 2022

DECLARATION

I declare that the thesis entitled “**STUDY OF Z + b-JETS PRODUCTION AT LHC ENERGIES USING CMS DETECTOR**” has been prepared by me under the guidance of Prof. Suman Bala Beri, Department of Physics, Panjab University, Chandigarh and Dr. Sunil Bansal, UIET, Panjab University, Chandigarh. No part of this thesis has previously formed the basis for the award of any degree or fellowship.



Meena,
Department of Physics, Panjab University
Chandigarh.

DATE : 23-11-2022

CERTIFICATE

We clarify that Meena has prepared her thesis entitled **"STUDY OF Z + b-JETS PRODUCTION AT LHC ENERGIES USING CMS DETECTOR"** for the award of Ph.D. degree of the Panjab University, Chandigarh under our guidance. She has carried out the work at Department of Physics, Panjab University, Chandigarh.

S. Beri

Prof. Suman Bala Beri
Department of Physics,
Panjab University, Chandigarh.

SW

Dr. Sunil Bansal
UIET,
Panjab University, Chandigarh.

DATE : 23-11-2022

*Dedicated To My
Family
&
Friends*

Acknowledgements

I want to express my sincere gratitude to my supervisors, Prof. Suman Bala Beri and Dr. Sunil Bansal for giving me never-ending support in each possible way during this wonderful journey of PhD. Prof. Suman Bala Beri and Dr. Sunil Bansal's valuable guidance, patience, and suggestions helped me a lot to keep going toward this thesis completion. I owe my sincere gratitude to Prof. Suman Bala and Dr. Sunil Bansal for giving me the opportunity to work in the world's largest accelerator. I am always honoured to be their student.

Prof. Suman Bala Beri is a humble soul. She has been mentoring me in the best way possible and discussions with her on any aspect, were always enlightening and motivating. Her enthusiasm toward research was stimulating and her thorough understanding of high energy physics helped a lot during the course of this study.

Dr. Sunil Bansal's enthusiasm and passion towards research, motivated me to complete Ph.D under. He has not only guided me to complete the research work presented in this thesis, but also taught us other very basic skills such as good documentation or power point making skills.

I am thankful to Prof. Rajiv Kumar Puri, Chairman, Department of Physics, Panjab University, Chandigarh, for providing adequate facilities to work in the department. I am also thankful to the former Chairmans Prof. Navdeep Goyal and Prof. Devinder Mehta for their support. I would like to thank Prof. Devinder Mehta for his continuous support and innumerable physics discussions which have helped me to improve my knowledge of the subject.

This research work has been carried out based on the CMS experiment at the CERN LHC. I would like to thank the whole CMS collaboration, the scientists, as well as technical and administrative staff, at the LHC accelerator complex and the CMS experiment. Many thanks to INDIA-CMS Collaboration for giving valuable suggestions on my work during regular meetings.

I would also like to thank Prof. J.B. Singh, Prof. Manjit Kaur, Prof. Vipin

Bhatnagar, and Dr. Sushil Singh Chauhan for their continuous help and support

I appreciate all the efforts of the Engineers and CMS office staff to take care of all our needs. I am highly grateful to our CMS technical staff, Mr. Shiv, Mr. Navjot, and Mr. Amarjeet, for assisting in the GEM Laboratory. I would also like to thank Mr. Ajit, Mr. Gyan, and Mr. Subash for their assistance with administrative tasks.

I am extremely thankful to the technical, library, and purchase section staff at PU who contributed directly or indirectly throughout this research work and of course Harry Bhaiya's "Tea".

I would like to thank all my seniors and my fellow colleagues for making a joyful environment in the PU-EHEP lab. Special thanks to Dr. Krishna, Dr. Priyanka Pathania, Sandeep dudi, Dr. Rajat Gupta, Amandeep Kaur Viridi, Jyoti Babbar, Amandeep kaur, Anureet Kaur, Harjot Kaur, Sandeep Hundle, Gitanjali, Ayushi Singla, Sunil Sarohak, Prachi, Bhumika, and Latika.

Words cannot express my gratitude to my family, who have supported me throughout this journey. This thesis owes most to them. I am indebted to my mother, Mrs. Krishna (Santosh), for her love and unconditional support. I could never have reached this far without the efforts and support of my father, Mr. Baljeet. Thank you for having faith in me. Mr. Sunil Luhach, Sanjay Luhach, Harbir Dalal, and Jasvir Dalal, my brothers, have been tremendously supportive throughout this journey. I would like to thank my sister-in-law Mrs. Sunita Luhach, Sonia Luhach, and Manisha Dalal. I would not miss mentioning my nieces and nephew, Sahil, Samiti, Arman, Savita, and Jatin. I am obliged to my school teachers, Ajit kumar Arya, Jaswant sing, and Parveen for encouraging me to pursue my carrier in physics.

I duly acknowledge the financial support of University Grants Commission. This research would definitely not have been possible without the financial support of UGC.

Date: 23-11-2022


(Meena)

Abstract

The most successful theory in particle physics is the Standard Model (SM), which describes the fundamental particles and their interactions. The majority of SM predictions have been tested successfully with excellent precision. Despite its achievements, SM fails to resolve many key problems such as the inclusion of gravity, absence of dark matter, and unification of the gauge couplings at high energies, etc. Many scenarios beyond Standard Model (BSM) have been proposed to explain some of these aspects. To test the predictions of the SM and BSM scenarios several experiments have been designed and built. Among them, the Large Hadron Collider at the European Council for Nuclear Research is the largest and most powerful particle accelerator ever built. The proton beams are made to collide at four different locations along the LHC ring. At these positions, four giant detectors are placed to record every possible outcome of the collision. The Compact Muon Solenoid (CMS) is one of its general purpose detectors designed for SM measurements and BSM searches due to its high precision detection of particles from collision events. This thesis is based on two aspects: hardware work and physics analysis.

This present work includes the assembly and testing of the Gas Electron Multiplier based GE1/1 subdetectors, which are used for the upgrade of the first endcap muon station of the CMS detector during the Long Shutdown2 period (December 2018–April 2022). The GE1/1 subdetectors have added redundancy to the muon subsystem of the CMS detector in the forward regions (higher η) and improved the triggering and muon tracks reconstruction capabilities. The GE1/1 chambers have an excellent detection efficiency of more than 97% with time and position resolution of 10 ns and $260\text{ }\mu\text{m}$ to $340\text{ }\mu\text{m}$, respectively. The CMS GE1/1 upgrade project required the fabrication and testing of 144 GE1/1 chambers. India contributed to the CMS GE1/1 upgrade project, which included the assembly and testing of the GE1/1 chambers, which was carried out at Panjab University, Chandigarh, in collaboration with Delhi University, Delhi, and Bhabha Atomic Research Center, Mumbai. The assembly technique includes material preparation in the lab (pre-assembly) and final

detector assembly in the class-100 clean room. I participated in the assembly and quality control tests of eight chambers at Panjab University. All chambers passed quality control tests and have been installed in the CMS experiment.

The physics analysis presented in this thesis is based on the measurement of the fiducial cross section for the production of a Z boson in association with at least one bottom quark jet. This is the first measurement of these processes based on data collected by the CMS experiment at a center-of-mass energy of 13 TeV during the 2016–18 data-taking period. The dataset corresponds to an integrated luminosity of 137 fb^{-1} . The Z + b-jets is a crucial process at the LHC for precision testing of electroweak and perturbative quantum chromodynamics (pQCD) calculations as it involves an electroweak and strong vertex in hadronic collisions. The b quark mass is above QCD scales so their production is not affected by low energy hadronization effects and is directly described by pQCD more precisely than a light-flavored jet. Understanding and modeling of pQCD interactions are important because they can affect the accuracy levels when we proceed with theoretical predictions. In pQCD, the amplitude for Z + b-jets processes can be computed using different parton distribution functions which provide information about the b quark content of the proton and also helps us to scrutinize the theoretical calculations by comparing predictions with experimental data. This comparison can help to improve the modeling of the production of b quarks. Furthermore, the Z + b-jets process constitutes an irreducible background to many SM and BSM physics processes that occur at hadron colliders, such as the production of Higgs bosons production in the HZ mode, the new generation of heavy quarks (b' , t') decaying into Z + b-jets, supersymmetric Higgs bosons produced in association with b quarks, etc. This study may improve the purity and sensitivity of these relatively rare processes if one can measure and control the Z + b-jets process precisely enough, which can also help to constrain the theoretical modeling of data. Therefore, the study and understanding of the b-jets processes are very important to probe new physics and improve the precision of the measured properties of known processes and particles.

The measurements Z + b-jets processes are performed with transverse momenta

of the b-jets greater than 30 GeV and pseudorapidity $|\eta| < 2.4$. The integrated and differential cross sections are measured separately in the electron and muon decay channels and also in the combined channel. The differential cross section distributions are measured as functions of various kinematic observables which are sensitive to the b quark parton distribution function, pQCD calculations, initial-state-radiation, final-state-radiation, and multiple parton interactions. The detector-level distributions are corrected for the selection efficiencies or detector resolution effects and compared with different Monte Carlo event generators. It is observed that the measured fiducial cross sections and shapes of the kinematic observables are not completely described by all predictions.

Contents

List of Figures

List of Tables

1	Introduction	1
1.1	The Standard model	3
1.1.1	Fundamental particles	5
1.1.2	Fundamental forces	7
1.1.3	Electroweak theory and Higgs mechanism	10
1.1.4	Quantum chromodynamics	12
1.1.5	Limitations of the SM	15
1.2	Associate production of Z boson with b-jets at LHC	17
1.3	Physics motivation	20
1.4	Thesis organization	21
	Bibliography	24
2	Experimental Setup	31
2.1	The Large Hadron Collider	32
2.1.1	The CERN accelerator complex	34

2.1.2	LHC parameters and performance	39
2.1.3	Proton-Proton collisions at the LHC	43
2.2	The CMS detector	45
2.2.1	The CMS coordinate system and kinematic variables	48
2.2.2	Silicon tracker	49
2.2.3	Electromagnetic calorimeter	53
2.2.4	Hadron calorimeter	56
2.2.5	The solenoid magnet	60
2.2.6	Muon detector	60
2.3	The CMS event trigger and data acquisition	65
2.3.1	Level-1 trigger	66
2.3.2	High level trigger	67
	Bibliography	70
3	Generation, Simulation, and Reconstruction of the Events	75
3.1	CMSSW framework	77
3.2	Event generation	77
3.3	Monte Carlo event generators	86
3.3.1	PYTHIA	88
3.3.2	SHERPA	88
3.3.3	POWHEG	89
3.3.4	MADGRAPH5_aMC@NLO	89
3.4	Detector simulation	90
3.5	Physics object reconstruction	92
3.5.1	Beamspot reconstruction	94
3.5.2	Track reconstruction	94

3.5.3 Primary Vertex Reconstruction	95
3.5.4 Secondary vertex reconstruction	96
3.5.5 Electron reconstruction	99
3.5.6 Muon reconstruction	102
3.5.7 Lepton Isolation	104
3.5.8 Missing Transverse Energy	106
3.5.9 Jets Reconstruction	106
3.5.10 Jet corrections	110
3.5.11 b/c jet Tagging	112
Bibliography	116
4 Assembly and testing of Gas Electron Multiplier	127
4.1 Motivations for the GE1/1 muon detector upgrade	129
4.2 Summary of performance requirements for the GE1/1 chamber	131
4.3 GE1/1: construction, principle, and working	132
4.4 Assembly	139
4.4.1 Pre-assembly	139
4.4.1.1 Preparation of the readout board	139
4.4.1.2 Preparation of the drift board	141
4.4.1.3 Preparation of the internal and external frames:	144
4.4.2 Assembly in clean room	145
4.4.2.1 Assembly of the GEM stack	146
4.4.2.2 Closing of the chamber:	152
4.5 GE1/1 quality assurance/quality control test	157
4.5.1 QC2 - leakage current test	158
4.5.2 QC3 - gas leakage test	161

4.5.3	QC4 - HV test in CO ₂	164
4.6	Further QC testing, integration and validation with cosmic-ray muons	169
4.7	GE1/1 installation and early commissioning	175
	Bibliography	177
5	Measurement of the production cross section for Z + b-jets in proton-proton collisions at $\sqrt{s} = 13$ TeV	181
5.1	Introduction	181
5.2	Signal process	182
5.3	Background processes	183
5.4	Data and Monte Carlo samples	186
5.5	Observables	190
5.6	Events selection criteria	191
5.7	Background estimation	203
5.8	Unfolding and systematic uncertainties	206
5.8.1	Experimental uncertainties	210
5.8.2	Theoretical uncertainties	214
5.9	Results and discussion	217
	Bibliography	226
6	Summary	235
6.1	Assembly and testing of Gas Electron Multipliers	236
6.2	Measurement of the production cross section for Z + b-jets in proton-proton collisions at $\sqrt{s} = 13$ TeV	237
6.3	Outlook	239
	List of Publications	241

CONTENTS

Reprints

251

List of Figures

1.1	Spectra of particle discoveries [17].	3
1.2	Standard Model particles [22].	8
1.3	Different interactions among SM particles [24].	9
1.4	Pictorial view of quark confinement.	14
1.5	Diagram showing the asymptotic freedom phenomenon in which at high values of Q^2 , the coupling constant becomes very small [39]. . .	15
1.6	Comparison between theoretically given and observed cross sections by the CMS experiment for various SM processes [47].	18
1.7	Examples of Feynman diagrams for Z + 1 b-jet production in the 4FS (left) and 5FS (right).	19
1.8	Outline of the thesis.	22
2.1	Overview of the timeline for the HL-LHC upgrade [5].	33
2.2	Detailed diagram of the LHC injection series for protons and Pb nuclei. The proton and lead ion trajectories are indicated with red and blue arrows, respectively. The position of the four interaction points is also included.	36
2.3	LHC schematic and the four main interaction points [6].	37
2.4	Cumulative luminosity delivered to the CMS experiment during pp collisions at various $\sqrt{s}$ (left). Total integrated luminosity delivered by LHC and recorded by the CMS detector during Run2 data-taking period [16] (right).	40
2.5	Distribution of pileup and mean number of interactions per bunch crossing for the 2011–2012 and 2015–2018 for pp collisions recorded by the CMS experiment [16].	44
2.6	Schematic of a hard scattering pp collision [18].	45
2.7	A perspective view of CMS detector and its subdetectors [19].	46

2.8	Schematic view of a transverse slice of the CMS detector, illustrating the specific signatures of different types of detected particles.	47
2.9	Schematic diagram of the coordinate system used in the CMS experiment.	49
2.10	A slice through the CMS tracker detector in the r-z plane. The separate subsystem regions have been highlighted and are classified as the Tracker Inner Barrel and Disks (TIB and TIDs), the Tracker Outer Barrel (TOB), and the Tracker Endcaps (TECs). The Pixel subsystem is highlighted in red. Back-to-back stereo strip modules are highlighted in blue [20].	50
2.11	The layout of pixel detectors in the CMS tracker. The magenta wedges on the endcap disks are carbon-fiber blades, which hold plaquettes-rectangular arrangements of pixel sensors that come in five different sizes. The rectangles in the barrel layers are the barrel modules- 2×8 rectangular arrangements of pixel sensors, mounted on rectangular carbon-fibre blades.	51
2.12	The layout of the CMS Phase-1 pixel detector: Pixel barrel layers geometry previously and after the upgrade (yellow area) (left). Pixel longitudinal view comparing Phase-1 in the upper part with Phase-0 in the lower part (right).	52
2.13	The ECAL geometrical view and mechanical layout. For the barrel, crystals are organized into modules and super-modules, and the endcap has two half-disks, or dees. Almost the entire endcap surface is covered by the preshower detector.	54
2.14	Longitudinal view of an ECAL quadrant. The pseudorapidity coverage of the barrel, endcap, and preshower systems are indicated.	54
2.15	Longitudinal view of a CMS HCAL subdetectors.	57
2.16	Schematic view of CMS muon detector.	61
2.17	A CMS Muon spectrometer quadrant with the DT chambers (yellow), RPC (light blue), and CSC (green). The dashed box contains the positions of the new forward muon detectors for the HL-LHC phase. They are coloured red for the GEM stations (ME0, GE1/1, and GE2/1) and violet for the upgraded RPC stations (RE3/1 and RE4/1).	65
2.18	Flow diagram of the L1 trigger system of CMS.	67
2.19	CMS data flow from HLT to physics datasets.	69

3.1	CMS simulation workflows [4].	76
3.2	Parton distribution functions of the proton from NNPDF3.0 at NNLO as a function of the fraction of the proton momentum at two energy scales: (a) $Q^2 = 10^2$ GeV and (b) $Q^2 = 10^4$ GeV. Each curve repre- sents $x f(x, \mu^2)$. For display purposes, the gluon PDF is divided by 10 and the PDFs are weighted by the momentum fraction. Due to this weight, the graphs represent the momentum density associated to a parton type instead of its number density.	81
3.3	Double-counting: In the downward vertical direction Feynman dia- grams of LO processes with increasing multiplicity (displayed in red) which are used by ME generators are displayed. The horizontal direc- tion shows the Feynman diagrams of the corresponding LO processes with given multiplicity while the number of additional partons cre- ated by a PS (displayed in blue) is increased. A comparison between diagonal Feynman diagrams shows that they look alike which gives rise to the double-counting called overlap in the phase space [23]. . .	83
3.4	Overview of a pp collision as simulated by MC event generator. The three green lines coming from both the left and right of the diagram represent the valence quarks of the protons and the green ellipses represent PDFs of protons. The large dark red circle represents the hard scattering. The particles produced in hard scattering further de- cay, as indicated by the small red circles, and also emits other gluons and quarks (FSR), in red. Hard scatter surrounded by ISR in blue and FSR in red. Hadronization is represented by light green ellipses. Unstable hadrons further decay until stable final state particles are formed represented by dark green circles. The purple ellipse indicates an MPI. The remaining particles of the incoming protons make up the BBR, in cyan, also part of the UE. The yellow lines represent soft photon radiation [30].	87
3.5	Reconstruction algorithm for physics objects in CMS.	93
3.6	A pp interaction event recorded by the CMS experiment. It features two muons (indicated by red lines) detected in the muon chambers (indicated by red slabs) and two jets of particles (visualized by cones). The upper-right corner of the picture is a zoom-in view of the pp interaction vertex [61].	93
3.7	Seeding (left), building (middle), and fitting (right) [65].	96

3.8	Illustration of a heavy flavor jet with a SV from the decay of a b or c hadron resulting in charged particle tracks (including possibly a soft lepton) that are displaced with respect to the PV, and hence with a large IP value.	97
3.9	An electron (and its bremsstrahlung photons) passing through the tracker and creating clusters of energy in the electromagnetic calorimeter (ECAL).	100
3.10	Reconstruction of different types of muons.	104
3.11	Illustration of a jet to which bundles of partons, hadrons, or detector measurements are grouped together [76].	107
3.12	Parton-level jets clustered with cone size $R=1$ using anti- k_T jet algorithm [82].	110
3.13	Jet correction approach.	110
4.1	A quadrant of the muon system, shows DT chambers (yellow), RPC (light blue), and CSC (green). The new GEM chamber stations (ME0, GE1/1, and GE2/1) are indicated in red [4].	129
4.2	Comparison of L1 muon trigger rates as a function of trigger threshold with CSC only (blue curve) and with the combination of CSC and GEM (purple curve) trigger information [4].	130
4.3	Scanning Electron Microscope (SEM) picture of a GEM foil [3] (left) and schematic view of the electric field lines (white), electron flow (blue), and ion flow (purple) through a bi-conical GEM hole (middle), and sketch of a GE1/1 chamber (right) [13].	134
4.4	The GEM electronics readout system.	135
4.5	Exploded view of GE1/1 chamber [17].	137
4.6	First CMS muon endcap station where the inner ring is equipped with 18 long (pink) and 18 short (light blue) GE1/1 chambers.	138
4.7	List of components required for the readout board preparation: readout board (left), gas connectors (middle), and brass inserts (right).	140
4.8	Clamping of brass insert using clamping hand.	140
4.9	Making the thread in the gas inlet/outlet (left), applying glue to the gas connector (middle), and fixing the gas connector onto the board (right).	141

4.10 List of components required for the drift board preparation: drift board (top left), pull-outs (top middle), M3×6 screws (top right), nylon washers (bottom left), HV contact pins (bottom middle), and SMD components (bottom right).	142
4.11 Mounting the pull-outs on the drift board using torque screw-driver.	142
4.12 Positioning of the HV pins and SMD components on the drift board.	142
4.13 Soldering the HV pins (left), preparing the SMD pads (middle), and soldering the SMD components (right).	143
4.14 List of components required for frame preparation: internal frames, brass inserts, external frame, and O-ring (from left to right).	144
4.15 Removing the internal frames (left) and inserting brass inserts into the 3 mm internal spacer (right).	145
4.16 Inserting the VITON O-ring into the external frame.	145
4.17 List of components required for the assembly of the GEM stack: GEM foil (top left), prepared internal frame (top middle), M2×6 screw (top right), M2.5 nuts (bottom left), guiding pins (bottom middle), and M2×6 long screws (bottom right).	146
4.18 Cleaning the GEM foil with a anti-static roll (left) and mounting the HV clip to the foils (right).	147
4.19 Location of the HV pads.	148
4.20 Cutting off the spare HV pads.	148
4.21 Cleaning the assembly baseplate (left) and inserting the guiding pins (right).	149
4.22 Mounting the 3 mm internal frames.	149
4.23 Placing GEM1 (left) and pre-stretching points (right).	150
4.24 Mounting the 1 mm internal frames.	150
4.25 Placing GEM2 foil (left) and the metallic nuts into the internal frame (right).	151
4.26 Placing the Plexiglas cover (left), attaching the Plexiglas cover to the GEM stack (middle), and closing the internal frame (right).	152
4.27 Cleaning the GEM stack (left), cutting the excess Kapton foil (middle), and removing the excess Kapton foil (right).	154
4.28 Cleaning the drift board with the vacuum cleaner (left), removing the guiding pins (middle), and removing the Plexiglas cover from the GEM stack (right).	154

4.29 Description of the stretching procedure (left) and stretching of the foils (right).	155
4.30 Inserting the external frame between the GEM stack and assembly jig (left and middle) and mounting the readout board on the stack (right).	156
4.31 Sealing the chamber (left), placing Panasonic-to-Lemo adapter (middle), and testing the induction gap using an insulation meter (right).	157
4.32 Schematic overview of the quality assurance and QC procedures of the GE1/1 chambers [17].	159
4.33 QC3 gas leak test experimental setup.	162
4.34 Front (left) and back view (right) of QC3 gas leak test Stand.	162
4.35 The QC3 test results for the GE1/1 chambers assembled at PU.	164
4.36 QC4 HV test experimental setup.	167
4.37 QC4 overview of the chamber with the grounding plate and 50 Ω terminations (left) and QC4 close view of the sample point on the drift PCB (right).	167
4.38 The QC4 HV test result (left) and intrinsic noise rate as a function of monitored current (right) for 8 GE1/1 chambers assembled at PU.	169
4.39 GEM chamber with GEB (left) and GEM chamber with all electronics components (right).	170
4.40 Recorded CFG_THR_ARM the DAC register value for every channel of the VFATS (left) and partially unplugged VFAT on the readout board (right).	171
4.41 Noise rate as a function of the threshold for a normally behaving VFAT (left) and for misplaced/faulty VFAT (right).	172
4.42 S-curve showing one VFAT having 100 percent efficiency for each DAC value (left) and VFAT having dead channels (right).	173
4.43 Threshold of VFAT3 (left) and Noise (right).	174
4.44 Chamber with cooling plate (left), with the thermal screen (middle), and superchamber assembly (right).	174
4.45 Comparison of Equivalent Noise Charge (ENC) of chamber GE1/1-X-S-INDIA-0006 measured at QC7 and QC8, and in the CMS experimental cavern after the installation (a) per VFAT (b) per readout channel [26].	175

5.1	Examples of Feynman diagrams for $Z + 1$ b-jet (left) and $Z + 2$ b-jets (middle and right).	183
5.2	Examples of Feynman diagrams for $Z + c$ -jets (left) and $Z +$ jets (right).	183
5.3	Examples of Feynman diagrams for $t\bar{t}$ process.	184
5.4	Decay scheme of top quark.	184
5.5	Examples of Feynman diagrams for various diboson production processes.	185
5.6	Examples of Feynman diagrams for the single top quark production in LO s-channel (left), NLO t-channel (middle), and NLO tW (right) subprocess.	186
5.7	Examples of Feynman diagrams for the $W +$ jets production process.	186
5.8	Mean number of interactions per bunch crossing for the 2016 and 2018 data-taking periods [11].	187
5.9	Single electron trigger efficiency with respect to the offline reconstruction, as a function of the electron p_T , for different η regions using the 2018 dataset (left). The bottom panel shows the ratio of data over simulation events. The efficiency measurements combine the effects of the L1 and HLT triggers [38]. Single muon trigger efficiency as a function of muon p_T (right) [39]. Only statistical uncertainties are shown in both plots.	192
5.10	Distributions of the number of reconstructed primary vertices n_{PV} , after the application of pileup reweighting for the 2016 (left) and 2018 (right) years. The bottom panel shows the ratio of data over total MC simulated events. The band on ratio plots include only the statistical uncertainties.	193
5.11	Distributions as a function of ee transverse momentum, leading b-jet transverse momentum (top), leading b-jet absolute pseudorapidity, $\Delta\phi^{(Z, \text{b-jet})}$ between the Z and leading b-jet (middle), and $\Delta y^{(Z, \text{b-jet})}$, $\Delta R^{(Z, \text{b-jet})}$ between the Z and leading b-jet (bottom) for ee channel after $Z + \geq 1$ b-jet events selection for Run2 [46]. The bottom panel shows the ratio of data over total MC simulated events. The band on ratio plots includes only the statistical uncertainties.	201

5.12	Distributions as a function of $\mu\mu$ transverse momentum, leading b-jet transverse momentum (top), leading b-jet absolute pseudorapidity, $\Delta\phi^{(Z, \text{b-jet})}$ between the Z and leading b-jet (middle), and $\Delta y^{(Z, \text{b-jet})}$, $\Delta R^{(Z, \text{b-jet})}$ between the Z and leading b-jet (bottom) for $\mu\mu$ channel after $Z + \geq 1$ b-jet events selection for Run2 [46]. The bottom panel shows the ratio of data over total MC simulated events. The band on ratio plots includes only the statistical uncertainties.	202
5.13	Distributions as a function of leading b-jet p_T (left) and after applying scale factor (right) for $e\mu + \geq 1$ b-jet events for 2018 [46]. The bottom panel shows the ratio of data over total MC simulated events. The band on ratio plots includes only the statistical uncertainties.	204
5.14	Distributions as a function of leading jet p_T for $Z + \geq 1$ jet events for ee (left) and $\mu\mu$ (right) for the 2018 year [46]. The bottom panel shows the ratio of data over total MC simulated events. The band on ratio plots includes only the statistical uncertainties.	205
5.15	Distributions as a function of leading jet p_T for $Z + \geq 1$ jet events after applying light scale factor on $Z + \text{jet}$ MC for ee (left) and $\mu\mu$ (right) for the 2018 year [46]. The bottom panel shows the ratio of data over total MC simulated events. The band on ratio plots includes the statistical and experimental (Jet energy scale, jet energy resolution, p_T^{miss} , lepton selection, pileup reweighting, pileup jet identification, luminosity) uncertainties.	205
5.16	Distributions as a function of leading c-jet p_T for $Z + \geq 1$ c-jet events after applying scale factor from CR1 on $Z + \text{light-jet}$ MC for ee (left) and $\mu\mu$ (right) for the 2018 year [46]. The bottom panel shows the ratio of data over total MC simulated events. The band on ratio plots includes only the statistical uncertainties.	207
5.17	Distributions as a function of leading c-jet p_T for $Z + \geq 1$ c-jet events after applying scale factor on $Z + \text{light-jet}$ and $Z + \text{c-jet}$ MC from CR1 and CR3 region, respectively, for ee (left) and $\mu\mu$ (right) for the 2018 year [46]. The bottom panel shows the ratio of data over total MC simulated events. The band on ratio plots includes only the statistical uncertainties.	207

5.18	Normalized differential cross section distributions as a function of p_T^Z and $p_T^{\text{b-jet}}$ (only two of them are shown here) obtained by unfolding MG5_aMC (NLO) with the response matrix constructed with MG5_aMC (LO) and compared with generator-level prediction of MG5_aMC (NLO) in the combined channel for Run2. The band on ratio plots includes only the statistical uncertainties of the MCs [46].	210
5.19	Normalized differential cross section distributions as a function of p_T^Z and $p_T^{\text{b-jet}}$ (only two of them are shown here) obtained by unfolding the electron channel with the response matrix constructed for the muon channel for Run2. The band on ratio plots includes only the statistical uncertainties of the MCs [46].	211
5.20	Differential cross section (left) and the normalized differential cross section (right) distributions as a function of Z transverse momenta for $Z + \geq 1$ b-jet events. The uncertainties in the predictions are shown as colored bands in the bottom panel around the theoretical predictions including statistical, PDF, scale, and α_s uncertainties for the MG5_aMC (NLO, NNPDF 3.1, CP5), MG5_aMC (NLO, NNPDF 3.0, CUETP8M1) and the statistical and PDF uncertainties for the SHERPA predictions. The statistical, theoretical, and total uncertainties in data are indicated by the vertical bars, yellow, and hatched bands centered at 1, respectively.	220
5.21	Differential cross section (left) and the normalized differential cross section (right) distributions as a function of b-jet transverse momenta for $Z + \geq 1$ b-jet events. The uncertainties in the predictions are shown as colored bands in the bottom panel around the theoretical predictions including statistical, PDF, scale, and α_s uncertainties for the MG5_aMC (NLO, NNPDF 3.1, CP5), MG5_aMC (NLO, NNPDF 3.0, CUETP8M1) and the statistical and PDF uncertainties for the SHERPA predictions. The statistical, theoretical, and total uncertainties in data are indicated by the vertical bars, yellow, and hatched bands centered at 1, respectively.	221

5.22	Differential cross section (left) and the normalized differential cross section (right) distributions as functions of b-jet absolute pseudorapidity for $Z + \geq 1$ b-jet events. The uncertainties in the predictions are shown as colored bands in the bottom panel around the theoretical predictions including statistical, PDF, scale, and α_s uncertainties for the MG5_aMC (NLO, NNPDF 3.1, CP5), MG5_aMC (NLO, NNPDF 3.0, CUETP8M1) and the statistical and PDF uncertainties for the SHERPA predictions. The statistical, theoretical, and total uncertainties in data are indicated by the vertical bars, yellow, and hatched bands centered at 1, respectively.	221
5.23	Differential cross section (left) and the normalized differential cross section (right) distributions as functions of $\Delta\phi^{(Z, \text{b-jet})}$ between the Z boson and the leading b-jet for $Z + \geq 1$ b-jet events. The uncertainties in the predictions are shown as colored bands in the bottom panel around the theoretical predictions including statistical, PDF, scale, and α_s uncertainties for the MG5_aMC (NLO, NNPDF 3.1, CP5), MG5_aMC (NLO, NNPDF 3.0, CUETP8M1) and the statistical and PDF uncertainties for the SHERPA predictions. The statistical, theoretical, and total uncertainties in data are indicated by the vertical bars, yellow, and hatched bands centered at 1, respectively.	223
5.24	Differential cross section (left) and the (right) normalized differential cross section (right) distributions as functions of $\Delta y^{(Z, \text{b-jet})}$ between the Z boson and the leading b-jet for the $Z + \geq 1$ b-jet events. The uncertainties in the predictions are shown as colored bands in the bottom panel around the theoretical predictions including statistical, PDF, scale, and α_s uncertainties for the MG5_aMC (NLO, NNPDF 3.1, CP5), MG5_aMC (NLO, NNPDF 3.0, CUETP8M1) and the statistical and PDF uncertainties for the SHERPA predictions. The statistical, theoretical, and total uncertainties in data are indicated by the vertical bars, yellow, and hatched bands centered at 1, respectively.	224

5.25	Differential cross section (left) and the normalized differential cross section (right) distributions as functions of $\Delta R^{(Z, \text{b-jet})}$ between the Z boson and the leading b-jet for the $Z + \geq 1 \text{ b-jet}$ events. The uncertainties in the predictions are shown as colored bands in the bottom panel around the theoretical predictions including statistical, PDF, scale, and α_s uncertainties for the MG5_aMC (NLO, NNPDF 3.1, CP5), MG5_aMC (NLO, NNPDF 3.0, CUETP8M1) and the statistical and PDF uncertainties for the SHERPA predictions. The statistical, theoretical, and total uncertainties in data are indicated by the vertical bars, yellow, and hatched bands centered at 1, respectively.	224
5.26	Distributions of the cross section ratios, $\sigma(Z + \geq 2 \text{ b-jets})/\sigma(Z + \geq 1 \text{ b-jet})$, as functions of the (left) leading b-jet transverse momentum and (right) absolute pseudorapidity. The uncertainties in the predictions are shown as colored bands in the bottom panel around the theoretical predictions including statistical, PDF, scale, and α_s uncertainties for the MG5_aMC (NLO, NNPDF 3.1, CP5), MG5_aMC (NLO, NNPDF 3.0, CUETP8M1) and the statistical and PDF uncertainties for the SHERPA predictions. The statistical, theoretical, and total uncertainties in data are indicated by the vertical bars, yellow, and hatched bands centered at 1, respectively.	225

List of Tables

1.1	Properties of Quarks and Leptons [21].	7
1.2	Properties of fundamental interactions.	10
1.3	Branching fraction of Z Boson decay modes [21].	19
4.1	Main specifications and parameters for the design and operation of the GE1/1 chambers [17].	138
4.2	Results from QC4 HV test showing the value of measured resistance ($R_{\text{multimeter}}$) by multimeter, resistance (R_{IV}) by IV curve, and the maximum intrinsic noise rate (at the value of current $\sim 999 \mu\text{A}$) for all assembled chambers at PU.	168
5.1	HLT triggers path and their p_{T} thresholds used to collect events from a given sample for the 2016, 2017, and 2018 data-taking periods. . . .	192
5.2	Cut-based electron identification requirements for the tight working point in the barrel and in the endcaps.	195
5.3	Muon tight identification criteria.	197
5.4	Jet identification criteria in 2016 and 2017–2018 data-taking periods. .	198
5.5	Summary of event selection criteria for $Z + \geq 1$ b-jet in the dimuon and the dielectron channels.	199
5.6	Summary of event yields after $Z + \geq 1$ b-jet selection in the dimuon and the dielectron channel for 2016, 2017, 2018, and Run2.	200

5.7	Fiducial region definition at generator-level.	208
5.8	Summary of uncertainties (%) in the integrated cross sections for the dielectron, dimuon, and combined channels in the $Z + \geq 1$ b-jet. . . .	216
5.9	Summary of the uncertainties (%) in the differential (normalized differential) cross section distributions for the combined channel in the $Z + \geq 1$ b-jet events.	217
5.10	Measured and predicted cross sections (in pb) for the $Z + \geq 1$ b-jet and $Z + \geq 2$ b-jets final states. The cross section ratios between the $Z + \geq 2$ b-jets and $Z + \geq 1$ b-jet are shown in the last three rows for the dielectron, dimuon, and combined channels. The MG5_aMC (NLO) predictions include theoretical uncertainties (PDF, and renormalization and factorization scales).	219

Chapter 1

Introduction

For a human being, the most fundamental questions are regarding the beginning of the universe and its basic constituents. All matter in our universe is made up of a few fundamental constituents known as elementary particles. The study of elementary constituents of matter and the interactions between them is known as “particle physics” or “high energy physics”. The term “elementary” refers to particles that have no known structure, such as those that are tiny or can not be split down into smaller parts. According to the famous Einstein equation $E = mc^2$, the energy of the colliding particles can be transformed into the rest mass of new particles. To look more deeply into the structure of matter or to create new heavier particles, colliding particles must have tremendously high energy as well as a high intensity of collisions. To investigate elementary particles, we must move to a small scale (wavelength $\propto 1/E$) so that wavelengths do not overlap with those of the nearest particle, resulting in better resolution. Many elementary particles are extremely massive so a significant amount of energy (mc^2) is required to generate them. Therefore, accelerators are required to produce high energy. Fixed target and collider are the two types of high energy accelerators. In a fixed target setup, particles collide with a fixed target, whereas in a collider setup, collisions occur between two beams of high energy particles. Collider experiments bring together beams with equal but

opposing momentum, resulting in an initial momentum of zero, allowing all of the energy in both beams to be transformed into new particles. As a result, colliding beams have a significant advantage over fixed beams. Collider accelerators include the Tevatron [1] at Fermilab in the United States, the Large Electron Positron [2], and Large Hadron Collider (LHC) [3] at the European Council for Nuclear Physics (CERN), etc.

At the end of the 19th century, scientists indeed believed that the atom ($\sim 10^{-8}$ cm) was the fundamental constituent of all matter in nature. However, J. J. Thomson in 1897, discovered the first elementary particle, the “electron” (e^-) ($\sim 10^{-16}$ cm) in cathode rays experiments [4] by bending cathode rays in the presence of a magnetic field. Later in 1911, the famous Rutherford scattering experiment [5, 6] demonstrated that most of the mass of the atom was concentrated in the small, dense, positively charged “nucleus” ($\sim 10^{-12}$ cm). The discovery of the “neutron” (n) by J. Chadwick in 1932 [7] completed the atomic picture and allowed the explanation of matter formed in atoms by these fundamental particles (proton (p), e^- , and n). In 1933, C. Anderson experimentally observed the first anti-particle “positron” (e^+), in a cosmic ray experiment [8]. In 1934, Yukawa Heidki proposed the presence of “pion” (π) mesons as the transmitter of short-range nuclear forces between protons and neutrons. These force carriers were later observed in 1947 by Powell and his coworkers in the photographic-emulsions experiment at Bristol, who discovered that cosmic rays contain two middle-weight particles known as pion [9] and muon (μ^-) [10]. A more detailed description of the history of particle physics can be found in [11]. In the 20th century, the field of elementary particle physics was explored and successive discoveries of new particles were made.

A spectrum of particles discovered in different collider experiments, starting from the discovery of the electron in 1897, to the observation of the Higgs boson at the Large Hadron Collider (LHC) in 2012 is shown in Figure 1.1. All those discovered

particles could not be fundamental, so a simple and elegant theory was developed in the 1970s which successfully describes all the fundamental particles and forces via which they interact. This successful theoretical framework of the “Standard Model” (SM) [12–16] will be discussed in the next Section.

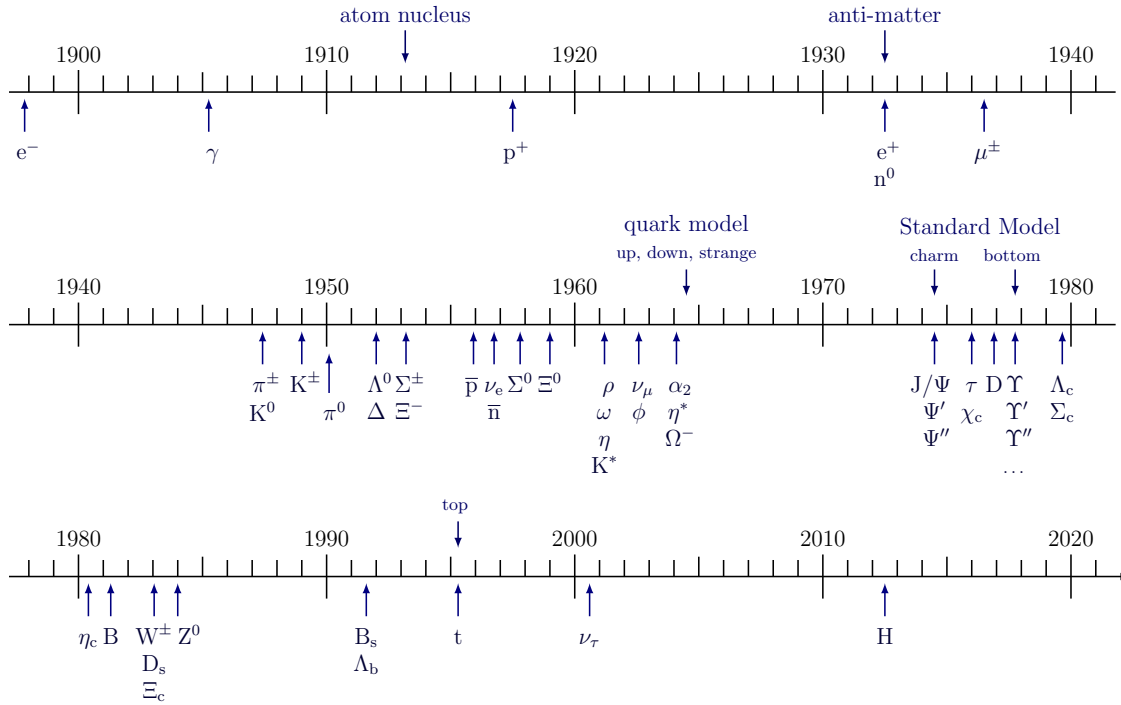


Figure 1.1: Spectra of particle discoveries [17].

1.1 The Standard model

The most widely accepted theory of elementary particle physics which explains most of the experimental results is the SM. The world of macroscopic objects in everyday life is governed by classical mechanics. The classical rules are modified by special relativity for objects that travel very fast (at speeds comparable to c). For objects that are very small (approximately comparable to the size of atoms), classical mechanics is superseded by quantum mechanics. For things that are both fast and small

(elementary particles), a theory is required to incorporate relativity and quantum principles: Quantum Field Theory (QFT). The SM of particle physics is a QFT that gives a description of the fundamental building blocks of matter and the interactions between them [18].

In the SM, all elementary particles correspond to excited states of quantum fields. The propagation and interactions of fields in a QFT are fully determined by a Lagrangian density, $\mathcal{L}(\psi_i, \partial_j \psi_i, x^j)$, that depends on the fields ψ_i , their derivatives $\partial_j \psi_i$, and the space-time coordinates, $x^j = (t, x, y, z)$. Lagrangian density is the dimension of the energy density to the fourth power. Its total space integral corresponds to the total kinetic energy of a system subtracted by the potential energy. Interactions between matter particles are introduced into the theory by imposing gauge symmetry conditions on the quantum fields. In QFT, elementary particles satisfy two main types of symmetries: space-time symmetries, acting directly on the coordinates x^j , and another one is internal symmetry, which corresponds to the transformation of a quantum field ψ without making their dependence on space or time. Symmetry can be local or global. A transformation is said to be global when it acts on the field in the exact same way at every point of space-time. According to Noether's theorem, each continuous global symmetry leads to the existence of a conservation law, and its conserved currents allow us to define a charge for the particles. The global symmetry can be made local, so that the factor $\alpha(\alpha = \alpha(x^j))$ can depend on the position in space-time, which leads to additional terms in the Lagrangian that break the invariance of the Lagrangian under this symmetry. This invariance can be restored by introducing additional gauge terms that couple to the conserved currents of the global symmetry. These gauge terms are vector fields, which correspond to the fundamental forces or interactions between the particles. Making a local symmetry from global symmetry is known as gauging the symmetry. The strength of the forces is driven by the gauge coupling constants.

The SM is a gauge theory of the electroweak and strong interactions with the

gauge group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ (here C, L, and Y stand for the color charge, left-handed chirality for weak isospin, and weak hypercharge $Y = B + S$, respectively, where B represents Baryon Number and S represents Strangeness). SM Lagrangian is required to remain invariant under the local gauge transformation of these gauge groups. The carriers of the forces are the generators of the symmetry group which are discussed in the next Subsection. The general form of SM Lagrangian can be written as:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_f + \mathcal{L}_g + \mathcal{L}_{Yukawa} + \mathcal{L}_\phi \quad (1.1)$$

here, $\mathcal{L}_f$ refers to Lagrangian describes the matter fermions and their interaction via the spin-1 mediators (gauge boson), $\mathcal{L}_g$ corresponds to gauge boson and how they interact with each other, $\mathcal{L}_{Yukawa}$ is the Yukawa interaction between the Higgs field and fermions, and $\mathcal{L}_\phi$ corresponds to boson coupling to the Higgs field and Higgs field potential, describing the Higgs boson self-coupling. The basic components of the SM are discussed in the following Subsections.

1.1.1 Fundamental particles

In the SM the fundamental particles are divided into two groups: fermions and bosons. The fermions constitute the building blocks of matter, while the bosons are the mediators of the interactions. The fermions must obey the Pauli Exclusion principle that is (i.e.) two fermions cannot exist in the same state. The fermions are classified into two types: leptons (l) and quarks (q) on basis of their interaction. Leptons interact exclusively through electromagnetic and weak interactions, while quarks interact through electromagnetic, weak, and strong interactions. In general, each particle has an anti-particle partner, often denoted by a bar ($\bar{y}$) with the opposite charge and quantum numbers.

- **Leptons:** The electron and its anti-particle, the positron are the stable charged leptons in free space. The muon and tau (τ^-), and their anti-particles

differ from the e^- and e^+ only in their masses and finite lifetimes. The neutral leptons are called neutrinos, denoted by the generic symbol ν that do not have any charge and are stable. All leptons (anti-leptons) except neutrinos carry an electric charge of -1 ($+1$).

- **Quarks:** Quarks have a non-integral electric charge and a strong color charge. The three colors are blue (B), green (G), and red (R) with the corresponding anti-colors being anti-blue (yellow or $\bar{B}$), anti-green (magenta or $\bar{G}$), and anti-red (cyan or $\bar{R}$). Quarks are grouped into two groups that differ only by one unit of electric charge. The quark type known as the flavor is denoted by the symbols: u for “up”, d for “down”, s for “strange”, c for “charm”, b for “bottom”, and t for “top”. The quark types are coarsely separated into light flavor quarks, which consist of u, d, and s quarks with masses in the MeV (considered $c=1$) region and heavy flavor quarks, which cover the remaining flavors and have masses in the GeV region. The heaviest quark is the top quark with a mass of 173.2 GeV [19,20]. All quarks have fractional charges i.e. $+2e/3$ for u, c, t and $-e/3$ for d, s, b.

In SM, all fermions are arranged in three generations of increasing mass, where each generation consists of two leptons (a charged lepton and a neutrino) and two flavors of quarks. The first generation fermions include e, “electron-neutrino” (ν_e) leptons and u, d quarks. Most of the visible matter in the universe is made up of first generation particles. The matter is made up of atoms which in turn consist of a nucleus (protons and neutrons) and electrons. The u and d quarks are the constituents of the proton (uud) and neutron (udd). The second generation includes μ , “muon-neutrino” (ν_μ) leptons and c, s quarks. The third generation consists of τ , “tau-neutrino” (ν_τ) leptons and t, b quarks. With the exception of neutrinos, second and third generation fermions are mostly unstable and decay via weak interactions to lower generations due to their short lifetime and can only be studied with

accelerators. The properties of the leptons and quarks are summarized in Table 1.1. The interactions between these elementary particles occur by mediator particles. The SM consists of 12 vector bosons (W^+ , W^- , Z , photon, and 8 gluons (g)) which are mediators of interactions, and a scalar boson is known as the “Higgs boson”. Figure 1.2 gives the pictorial view of the fundamental particles of the SM.

Table 1.1: Properties of Quarks and Leptons [21].

Generation		Flavor	Mass (MeV)	Charge	Spin	Interactions
I	Leptons	e	0.51	-1	1/2	EM, Weak
		ν_e	$< 2.2 \times 10^{-6}$	0	1/2	Weak
	Quarks	u	2.2	+2/3	1/2	EM, Weak, Strong
		d	4.7	-1/3	1/2	EM, Weak, Strong
II	Leptons	μ	105.66	-1	1/2	EM, Weak
		ν_μ	< 0.17	0	1/2	Weak
	Quarks	c	1.27×10^3	+2/3	1/2	EM, Weak, Strong
		s	96	-1/3	1/2	EM, Weak, Strong
III	Leptons	τ	1.78×10^3	-1	1/2	EM, Weak
		ν_τ	< 15.5	0	1/2	Weak
	Quarks	t	173.21×10^3	+2/3	1/2	EM, Weak, Strong
		b	4.18×10^3	-1/3	1/2	EM, Weak, Strong

1.1.2 Fundamental forces

All visible matter in the universe interacts via strong, electromagnetic, weak, and gravitational interactions. Interactions between SM particles are mediated by the exchange of force carriers i.e. integral spin gauge bosons. The exchange particles transfer momentum and energy between the two objects as well as change the quantum number. Due to their character, they are frequently called exchange particles. The SM describes three forces of nature, except gravitational force. Gravitational force, though prominent in the macroscopic lives of humans, is negligible at the minuscule level of particles. Each force is characterized by its strength and the range over which they are effective, which in turn depends upon the mass of its mediators. Figure 1.3 depicts how different particles interact via different forces. Table 1.2

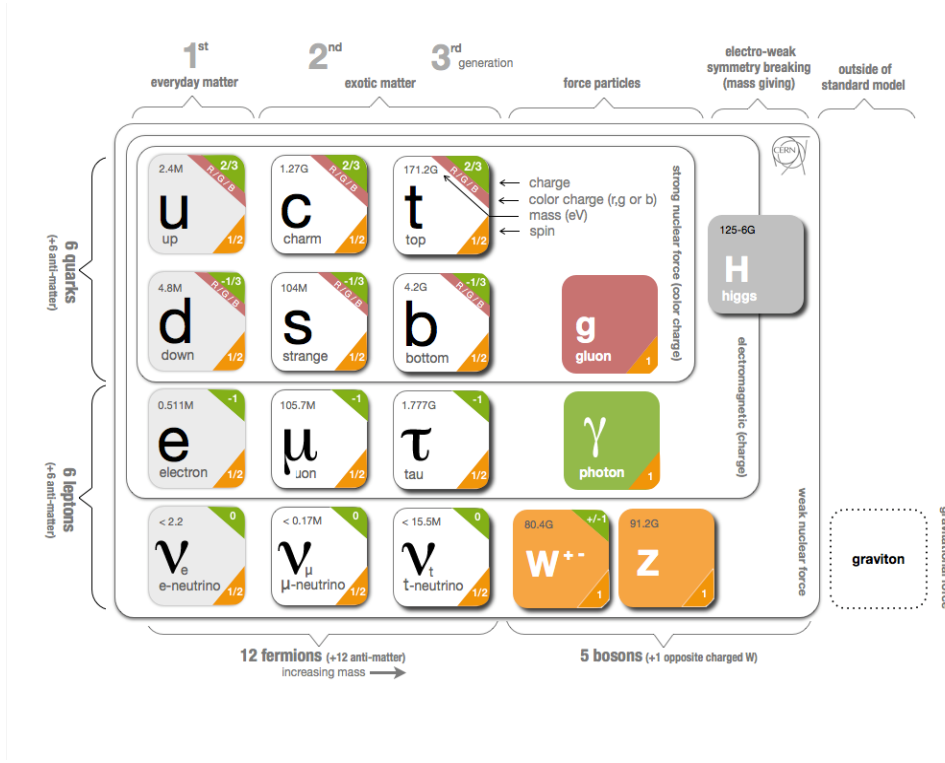


Figure 1.2: Standard Model particles [22].

shows the properties of four fundamental forces [23]. A brief description of these fundamental forces is given below.

- Strong interaction:** The strong interaction is experienced by particles having non-zero color charge i.e. quarks and gluons which are known as “Partons”. Leptons do not participate in strong interactions. The strong interaction is mediated by massless gluons. At the subatomic scale, this is the strongest force among all forces as it holds the protons and neutrons inside the nucleus despite the repulsive electromagnetic force. The range of strong force is limited to ≈ 1 fm. The strong interaction is described theoretically by Quantum Chromodynamic (QCD) [25, 26].
- Electromagnetic interaction:** This interaction is experienced by all particles having non-zero electric charge and is mediated by a massless and chargeless particle, known as a photon (γ). It is responsible for the bound states

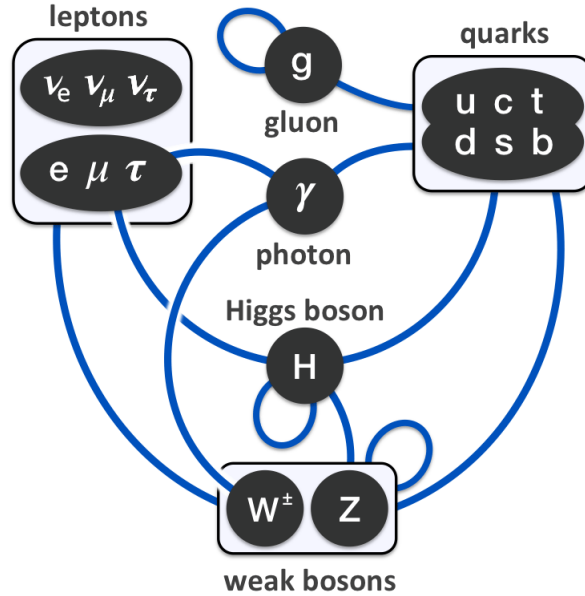


Figure 1.3: Different interactions among SM particles [24].

of electrons with nuclei, i.e. atoms and molecules, and for the intermolecular forces in liquids and solids. Electromagnetic interaction has an infinite range. The theory that describes electromagnetic interaction is known as **Q**uantum **E**lectrodynamics (QED) [27–29].

- **Weak interaction:** This interaction is experienced by leptons and quarks and is mediated by heavy vector bosons namely charged $W^\pm$ (called charged current interactions) and neutral Z^0 (called neutral current interaction). The weak force is a short-range ($\sim 10^{-18}$ m) force that is effective only at the subatomic scale. This force is responsible for the radioactive decay of unstable nuclei that occurs in nature. The charge responsible for this interaction is weak isospin. The neutral current interactions do not result in the exchange of leptons or quarks, whereas the charged current interactions include both. This is described by the **E**lectroweak theory [30, 31].

- **Gravitational interaction:** It is the weakest of all the fundamental interactions and acts between all particles. It is attractive in nature and is responsible for holding the planets inside the solar system. This force is supposed to be mediated by Graviton – massless particle of spin 2, that is not yet experimentally discovered. This interaction is almost 10^{38} times weaker than strong interaction and thus it is negligible at scales of elementary particles. Due to this fact, gravitational interaction is neglected in the SM. Gravitational force is described by the special theory of relativity. As of now, there is no successful theory of quantum gravity that can combine the theory of relativity with quantum mechanics.

Table 1.2: Properties of fundamental interactions.

Interactions	Mediators	Charge	Mass (GeV)	Range (m)	Coupling Constant (α)
Strong	8 Gluons	0	0	10^{-15}	1
Electromagnetic	Photon	0	0	∞	$\frac{1}{137}$
Weak	$W^{\pm}$	1	80.38 ± 0.012	10^{-18}	10^{-9}
	Z^0	0	91.188 ± 0.002		
Gravitational	Graviton	0	0	∞	10^{-38}

1.1.3 Electroweak theory and Higgs mechanism

QED is the theory that describes electromagnetic interactions. QED is characterized by the weak hypercharge gauge symmetry group $U(1)_Y$ that leaves the Lagrangian invariant under global symmetry and leads to the conservation of electric charge. Since there is only one generator and $U(1)_Y$ transformations are commutative, this means that the associated boson can not interact with each other i.e the photon

is electrically neutral. In QED, the coupling constant, known as the fine structure constant, is given as $\alpha = e^2/4\pi$. The coupling strength in QED is proportional to the electric charge of the particle and is independent of photon momentum transferred.

A theory of weak interactions was postulated by Enrico Fermi in 1933 as a zero-range contact interaction between four fermions but this theory only proved valid up to about 100 GeV. The weak interaction is mediated by W (W^+ and W^-) and Z bosons. The W boson couples only to particles of left-handed helicity and anti-particles of right-handed helicity. Left-handed helicity indicates that the spin of particle is orientated in the opposite direction of motion. The Z boson couples to both left- and right-handed fermions. The weak charged currents mediated by the W boson change the flavor of quark or lepton flavor and violate the parity. However, neutral weak currents mediated by the Z boson do not contribute to any flavor changing currents. The weak interaction is explained by $SU(2)_L$ which is the weak isospin (I) gauge group describing transformations of the left-handed isospin doublets. But, in order to make this theory a gauge invariant, there is a need to add another symmetry group for flavor invariance.

The coupling strength of the weak force becomes comparable to that of electromagnetism above the unification energy, on the order of 246 GeV, and they can merge into a single force. Sheldon Glashow, Abdus Salam, and Steven Weinberg unified the theory of electromagnetic and weak interactions. The photon contributes to neutral currents as the Z boson, hence giving way to unified electromagnetic and weak theories. So for flavor invariance, $U(1)_Y$ symmetry is added to this theory. The unified electroweak theory is represented by $SU(2)_L \otimes U(1)_Y$ symmetry group. The weak isospin and hypercharge are connected to the electromagnetic charge (Q) as $Y = 2(Q - I_3)$, where I_3 represents the third component of the weak isospin.

The electroweak theory involves the existence of four massless carriers, to mediate the unified electroweak interaction. But this is in direct conflict with what is observed in nature. If the mass terms for the particles were simply added to the

Lagrangian, the gauge symmetry would be broken. So, an alternative solution to this problem is the Electroweak symmetry breaking Higgs mechanism. This mechanism leaves the Lagrangian, but not the vacuum state, invariant under electroweak transformations, a principle called spontaneous symmetry breaking [32-34]. In the language of quantum mechanics, a symmetry breaking occurs when a field, or some components of a field, acquire a non-zero vacuum expectation value. In the Higgs mechanism an additional scalar field, the Higgs field, is introduced into the theory, which has a potential function of a form allowing degenerate vacuum solutions with a non-zero vacuum expectation value. It is said to be spontaneous because no external agent is responsible for it. Interaction with the Higgs field allows the W and Z bosons as well as other particles to acquire mass [35]. The Higgs boson is predicted to be quite massive and its experimental observation required the collision of particles at very high energy. On 4 July 2012, the two main experiments at the LHC, A Toroidal LHC ApparatuS (ATLAS) [36] and Compact Muon Solenoid (CMS) [37], independently reported the discovery of a new resonance with a mass of about 125 GeV, consistent with the Higgs boson.

1.1.4 Quantum chromodynamics

QCD is a quantum field theory and its constituents are a set of fermions, quarks, gauge bosons, and gluons. QCD is a gauge theory symmetric under the 3-dimensional special unitary group, $SU(3)_C$. QCD theory depicts the interaction of colored particles through the exchange of gluons having a certain momentum scale. Unlike in QED, the QCD coupling constant depends on the momentum exchange which makes it complicated. The conserved charge associated with this symmetry is the color charge. The $SU(3)_C$ group is non-abelian and the local invariance of this theory adds $8 (3^2 - 1 = 8)$ gauge fields. Gluons also couple to other gluons because each gluon carries a color charge. The colored objects are always confined to the states with zero color charge i.e. the states with non-zero color charge cannot exist

as free particles. In two quark states, color and its anti-color combine to make a colorless state called mesons. Three quark states known as baryons are also colorless but through a different combination: the three quarks are colored B, G, and R (or the corresponding anti-colors), which combine (like colored light) to form a state with no color. All mesons are unstable and the only known stable baryon is the proton.

The QCD coupling constant is represented by g_s and given as $g_s = \sqrt{4\pi\alpha_s}$, where α_s is the strength of interaction between quarks and gluons. The strong coupling constant depends on the energy scale of the interaction, referred to as running α_s

$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 2n_f)\ln(Q^2/\Lambda^2)} \quad (1.2)$$

here, Q is the momentum transferred in an interaction, Λ is the QCD scaling parameter, and n_f is the number of quark flavors (i.e 6). At $\Lambda_{QCD} \approx 200 \text{ MeV}$, the strong coupling constant becomes unity. For energy scales $Q \gg \Lambda$ ($\approx 1-2 \text{ GeV}$) the perturbative approximation is valid because $\alpha_s \ll 1$. QCD behaves differently at high and low energy scales, resulting in two important phenomena, which are discussed in detail below.

- **QCD confinement:** Hadrons are color singlet (i.e. color neutral) combinations of quarks, anti-quarks, and gluons. At low energies ($Q^2 \leq 1 \text{ GeV}$) or at a distance $\sim 1 \text{ fm}$ (typical size of hadrons), α_s is large and quarks inside the proton are confined in a bound state with other quarks. As a result, quarks cannot exist freely, which explains the stability of hadrons. If one of the quarks in a given hadron is pulled away from its neighbours that are in the bound state, the color force field increases between that quark and neighbour quark. More energy applied to separate the quark tends to add more energy to the color force. This energy is then utilized to form new quark-antiquark pairs while conserving the energy and other quantum numbers as the energy of the

color force field is converted into the mass of new quarks. Hence, quarks cannot exist independently. This phenomenon is known as “Quark Confinement” as shown in Figure 1.4. Confinement makes the QCD calculations of hadronic cross sections very complicated, due to the fact that bound states exist and perturbation series in the coupling constant cannot be applied.

- Asymptotic freedom:** At very high energies, where momentum exchange is large ($Q^2 \gg 1 \text{ GeV}$) or at a short distance regime, the coupling constant (α_s) decreases indicating quarks behaving as almost free or non-interacting. As $Q^2 \rightarrow \infty, \alpha_s \rightarrow 0$, indicating that at large Q^2 , interaction becomes weak and vanishes to zero asymptotically at high energies, justifying the term “Asymptotic Freedom” [38]. In this regime, perturbative QCD (pQCD) calculation can be used to make very accurate predictions. For example, a cross section of a given physics process is expressed by $\sigma = \sigma_0 + \sigma_1\alpha_s + \sigma_2\alpha_s^2 + \dots$, where σ_i are the cross sections at different perturbative orders. The leading order (LO) is the lowest order in the calculation, followed by the next-to-leading order (NLO), the next-to-next-to-leading order (NNLO), and so on up to a certain perturbative order. Figure 1.5 shows the running behaviour of the strong coupling constant as a function of the energy measured by different experiments. Q^2 is given by the sum of four vectors of the interacting partons.

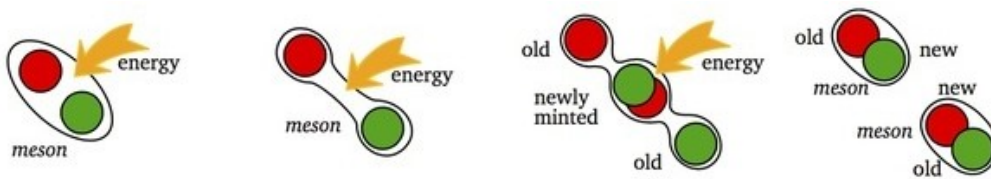


Figure 1.4: Pictorial view of quark confinement.

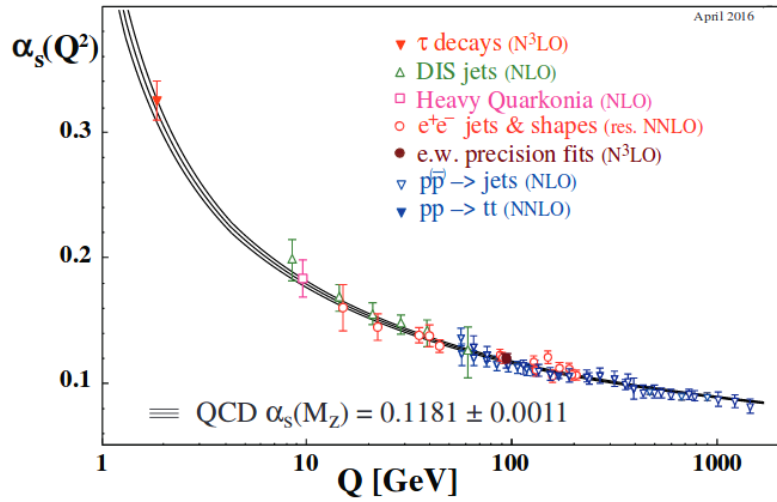


Figure 1.5: Diagram showing the asymptotic freedom phenomenon in which at high values of Q^2 , the coupling constant becomes very small [39].

1.1.5 Limitations of the SM

The SM has provided an elegant theoretical framework and most phenomena observed in particle physics experiments can be explained within this framework. The majority of SM predictions have been tested successfully to a high degree of precision by various experiments. However, it has several shortcomings and some of them are listed below.

- SM does not include Gravitational force and also, can not explain why gravitational force is much weaker than the EM or nuclear strong force.
- SM is not able to explain the large difference between the electroweak scale (10^2 GeV) and Planck's scale ($\approx 10^{19}$ GeV), called the Hierarchy or fine-tuning problem. The Higgs boson provides larger mass to scalar particles through large radiative correction [40]. But the Higgs boson mass is quite smaller than the Planck Scale, despite the huge quantum corrections. This implies that the bare mass of the Higgs boson should be tuned in such a way that it completely cancels out those quantum corrections.

- SM only explains visible matter (5%) but is not able to describe the existence of dark matter (26%) [41] and dark energy (69%) [42].
- At the time of the Big Bang around 13.8 billion years ago, it is believed that an equal amount of matter and anti-matter should have been created. However, physical objects in this universe around us are made up of matter and there exists no anti-matter which is the biggest challenge in physics. The SM fails to explain this large asymmetry between the matter and anti-matter.
- In the SM, neutrinos are considered massless and chargeless particles. But experimental observations show that neutrinos oscillate and transform into one another as they move [43, 44]. This feature is only possible if neutrinos have a non-zero mass.
- The SM does not explain why there are only three particle generations and why there is so much difference in masses of particles among different generations.

It is quite possible that the SM explains only a part of the bigger picture that might include new physics which is hidden either deep inside the subatomic world or in the darkness of this gigantic universe. Various beyond Standard Model (BSM) scenarios have been developed to resolve these key issues, which need to be experimentally verified [45].

A number of experiments have been designed and built to test the predictions of the SM and BSM scenarios. Among them, the LHC is the world's largest and most powerful particle accelerator ever built. Precision measurements of SM processes can provide a reference for the understanding of detector response as well as probes of new physics which are complementary to direct searches. One of the current challenges in particle physics is to improve the precision of the measured properties of known processes and particles in order to study any discrepancies between the measurements and SM predictions. The measurement of SM processes also forms the basis for testing and improving the Monte Carlo (MC) event generators that aim to

describe these processes in high energy hadron collisions at the LHC. The majority of SM predictions have been verified experimentally with tremendous precision, the most recent being the discovery of a new boson in 2012 at the LHC, which appears to be the Higgs boson as predicted by the SM. The Higgs boson discovery, along with the massive amount of data collected at the LHC, has provided physicists a sensitive tool to search for incompatibilities with the SM, which would result in new theoretical developments and could provide an answer to some of the questions stated above. Figure 1.6 compares cross sections that were predicted theoretically to those that were actually measured with the CMS experiment for several SM processes. Up to now, most measurements have shown consistency with the SM predictions within uncertainties, displaying the remarkable success of SM 46.

This thesis is based on the measurements of one of the important SM processes, the Z boson having an associated b-jets, using the CMS experiment at the LHC, as discussed in the next Section.

1.2 Associate production of Z boson with b-jets at LHC

The production of a Z boson in association with b-jets (a collimated spray of particles originating from a bottom quark) can occur from quark-gluon, quark-antiquark, and gluon-gluon interactions in pp collisions. Theoretically, there are two methods for calculating the amplitude of Z + b-jets processes, the 4 flavor scheme (4FS) and 5 flavor scheme (5FS) 48,49. In the 4FS, the b quarks are produced by hard initial-state gluon splitting, described by the QCD matrix element (ME); this process leads to at least one b-jet in the final state. Figure 1.7 (left) represents the Feynman diagram of the 4FS process. The 4FS ME treats the b quark as massive and thus it is not considered in the proton PDF evolution and perturbative calculation. In the 4FS, the proton PDFs only contain the four lightest quark flavors (u, d, s, and c)

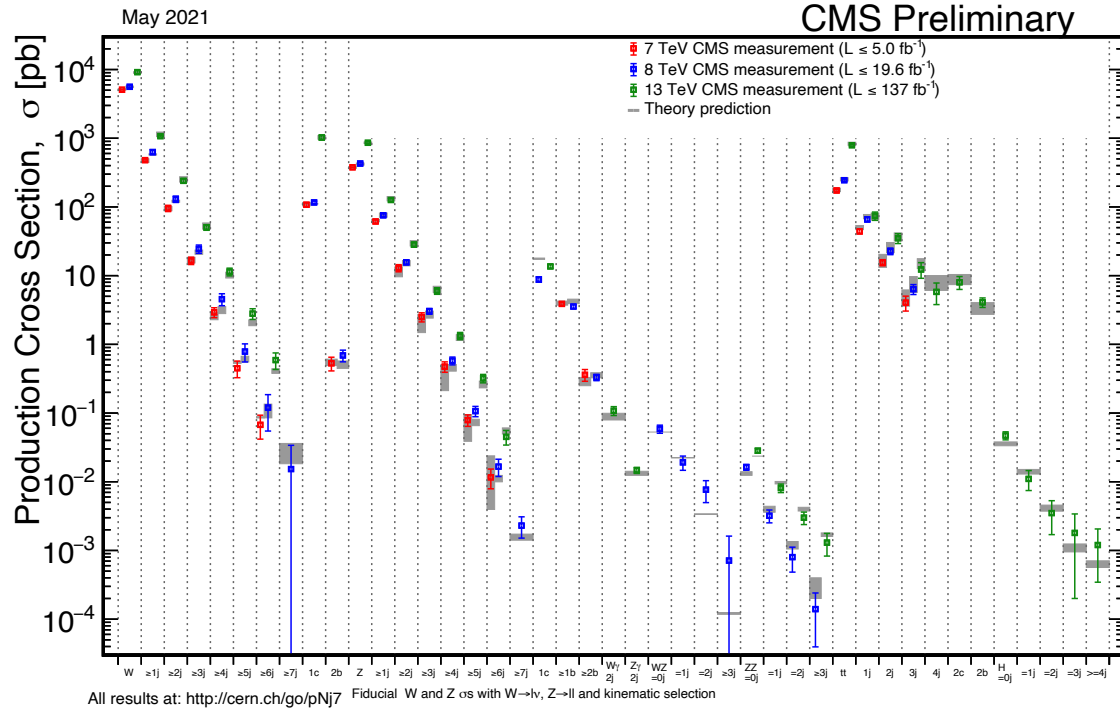


Figure 1.6: Comparison between theoretically given and observed cross sections by the CMS experiment for various SM processes [47].

and the b quarks are produced from an explicit gluon splitting $g \rightarrow b\bar{b}$ process. A collinear $g \rightarrow b\bar{b}$ splitting is assumed with one of the b quarks escaping down the beam line and the other quark participating in the hard scattering. The b quarks are treated as originating from the proton sea. In the 5FS, the gluon splitting contribution is included within the b quark PDF, and the b quark mass is set to zero in the ME calculation. Figure 1.7 (right) represents the Feynman diagram of the 5FS process.

The Z boson decays instantaneously due to the short mean lifetime (3×10^{-25} s) and can only be detected through its decay products. The decay of the Z boson can occur in several ways as listed in Table 1.3. The Z boson decays to strongly interacting particles approximately 70% of the time. It is experimentally challenging to detect the Z boson through the hadronic channels among the abundant QCD

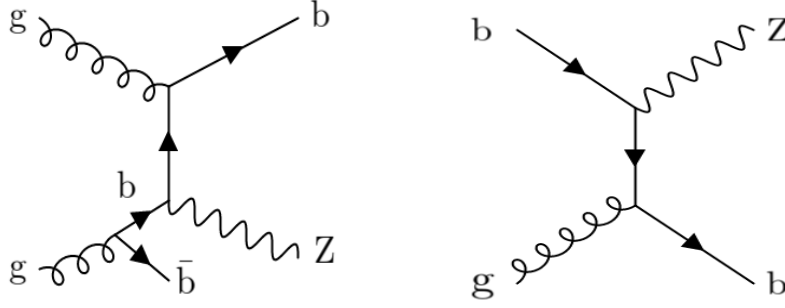


Figure 1.7: Examples of Feynman diagrams for Z + 1 b-jet production in the 4FS (left) and 5FS (right).

background. The Z boson decays to a charged lepton and anti-lepton pair (e^-e^+ , $\mu^-\mu^+$, and $\tau^-\tau^+$) approximately 10% of the time. The Z boson decays to $e^-e^+/\mu^-\mu^+$ providing a very clean signature at hadron colliders. Tau particles decay very quickly to other particles and so they are not directly detected.

For the present Z + b-jets analysis, the leptonic final state of Z ($\rightarrow e^+e^-/\mu^+\mu^-$) is chosen which makes the signal almost background free with high efficiency of leptons reconstruction.

Table 1.3: Branching fraction of Z Boson decay modes [21].

Mode	Fraction
e^-e^+	$3.363 \pm 0.004\%$
$\mu^+\mu^-$	$3.366 \pm 0.004\%$
$\tau^+\tau^-$	$3.370 \pm 0.008\%$
Invisible	$20.00 \pm 0.06\%$
Hadrons (total)	$69.91 \pm 0.06\%$

1.3 Physics motivation

The $Z + b$ -jets is a crucial process at the LHC for precision tests of electroweak and pQCD calculations as it involves an electroweak and strong vertex in hadronic collisions. Understanding and modeling of QCD interactions are important because they can affect the accuracy levels of theoretical predictions. In pQCD, the amplitude for $Z + b$ -jets processes can be computed using the 4FS and 5FS. Since the 5FS considers b quarks as the fifth flavor in the proton PDFs, so it provides information about the b quark content of the proton and also helps us to scrutinize the theoretical calculations by comparing predictions with experimental data. This comparison can help to improve the modeling of the b quarks production. These measurements provide valuable input for the tuning of perturbative and non-perturbative models and to constrain the model parameters of the advanced parton shower (PS) MCs programs.

Within the SM, the $Z + b$ -jets final state is the dominant background for the Higgs boson production in the HZ mode, in which the Higgs boson decays into a $b\bar{b}$ pair [50]. Many physics scenarios BSM predict final states involving b quarks and Z bosons: new generations of heavy quarks (b' , t') decaying into $Z + b$ -jets [51], supersymmetric Higgs bosons produced in association with b quarks [52], and extended SM scenarios involving additional $SU(2)$ doublets with enhanced Zb coupling [53]. This study may also provide useful information to describe an analogous process where a W boson is produced, which is more difficult to measure because of higher background contamination. The study of the $Z + b$ -jets processes can also help to estimate the theoretical systematic uncertainty coming from the PDF choice in the W mass measurement [54]. Furthermore, the final state of the Z polarization asymmetry measurement allows to study the Zb couplings combination, which also appears in the LEP measurement of the forward-backward asymmetry of b quarks [55]. Since the LEP result shows a 2.5 standard deviation discrepancy from the SM prediction in the electroweak fit, such measurement might contribute to investigating this long-

standing issue. If the $Z + b$ -jets process can be measured and controlled properly, it may improve the purity and sensitivity of these relatively rare processes, which can also help to constrain the theoretical modeling of data.

Measurements of heavy flavor production in association with a Z boson have attracted the interest of the Tevatron and LHC experimental communities in the last decades. Previous measurements have been performed on $Z + b$ -jets at center-of-mass energy ($\sqrt{s}$) = 1.96 TeV by the CDF [56] and D0 [57] Collaborations at the Fermilab Tevatron in $p\bar{p}$ collisions. The CMS and ATLAS Collaborations performed a measurement of the $Z + b$ -jets using pp collision at $\sqrt{s} = 7$ TeV, 8 TeV, and 13 TeV [58–61]. Previous measurements were compared with various MC predictions based on LO and NLO MEs interfaced with a PS simulation, which showed substantial mismodeling.

The main goal of this thesis in terms of physics analysis is to perform an integrated cross section of $Z + \geq 1$ b -jet processes, along with the differential cross section distributions as a function of various kinematic observables. These observables are sensitive to both the hard scattering process and associated soft QCD radiation, as well as to the PDFs. This is the first measurement of $Z + \geq 1$ b -jet processes at $\sqrt{s} = 13$ TeV based on pp collisions data collected by CMS experiment during the 2016–2018 LHC running periods. The data sample corresponds to an integrated luminosity of 137 fb^{-1} .

1.4 Thesis organization

The work presented in this thesis is divided into two parts: hardware work and physics analysis. The hardware work is based upon the assembly and testing of Gas Electron Multiplier (GEM) detectors, that have been installed in the CMS detector during Long Shutdown2, i.e., LS2 (2018–2022), in order to add redundancy to the existing muon system. The physics analysis involves the measurement of the cross

section for the production of a Z boson, decaying to e^+e^- or $\mu^+\mu^-$, in association with at least one b quark jet. The data sample corresponds to an integrated luminosity of 137 fb^{-1} , collected by the CMS experiment in pp collisions. The detailed outline of the thesis (see Figure 1.8) is as follows:

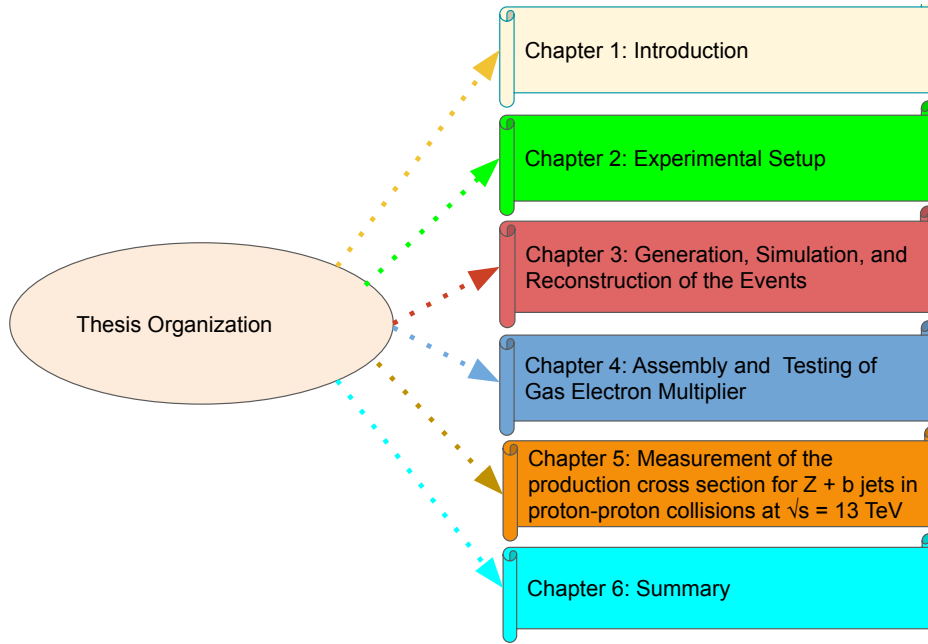


Figure 1.8: Outline of the thesis.

- The present **Chapter** provides a brief overview of how SM particles evolved with time. The SM theory, its successes, and shortcomings are discussed. A brief introduction and motivation for the measurements of the Z + b-jets processes are also given.
- **Chapter 2.3.2** gives an overview of the LHC machine and its parameters. Various experimental facilities at CERN are briefly discussed, along with the setup of the CMS detector used for the present work.
- **Chapter 3** This chapter aims to describe how information from the CMS subdetectors is combined to reconstruct physical objects and their algorithms.

It gives a brief overview of event generation, simulation, and different types of MC generators. Also, a detailed information of the reconstruction of the physics objects used in the present work such as primary vertex, secondary vertex, electrons, muons, and b-jets has been provided.

- **Chapter 4** gives a detailed description of the assembly and testing of the GEM chambers performed at Panjab University, Chandigarh and CERN, which is used for the upgrade of the muon system of the CMS detector.
- **Chapter 5** describes in detail the analysis performed to measure the production cross section of the $Z + b$ -jets process. This chapter includes a description of data, signal, and background MC samples, event selection, detector-level comparison, unfolding to correct for the detector effects, sources of systematic uncertainties, and results.
- **Chapter 6** provides a summary of the hardware and analysis work presented in this thesis.

The present work is done in collaboration with the CMS experiment at Geneva. But, the analysis and hardware work shown in this thesis is done by me.

Bibliography

- [1] F.T.Cole et al., “Design Report Tevatron 1 Project”, Batavia, IL (1984).
- [2] Bachy, G. and Hofmann, A. and Myers, S. and Picasso, E. and Plass, G., “The LEP collider: construction, project status and outlook”, *Particle Accelerators*, vol. 26 (1990) 19.
- [3] Lyndon Evans and Philip Bryant, “LHC Machine”, *Journal of Instrumentation*, vol. 3 (2008) S08001, [doi:10.1088/1748-0221/3/08/S08001](https://doi.org/10.1088/1748-0221/3/08/S08001).
- [4] J. J. Thomson, “XL. Cathode Rays - by J.J. Thomson”, *Philosophical Magazine*, vol. 44 (1897) 303, [doi:10.1080/09500830701306165](https://doi.org/10.1080/09500830701306165).
- [5] E. Rutherford, “The scattering of alpha and beta particles by matter and the structure of the atom”, *Philosophical Magazine and Journal of Science*, vol. 21 (1911) 669, [doi:10.1080/14786440508637080](https://doi.org/10.1080/14786440508637080).
- [6] E. Rutherford, “The structure of the atom”, *Nature*, vol. 92 (1913) 423, [doi:10.1038/092423a0](https://doi.org/10.1038/092423a0).
- [7] J. Chadwick, “Possible Existence of a Neutron”, *Nature*, vol. 129 (1932) 312, [doi:10.1038/129312a0](https://doi.org/10.1038/129312a0).
- [8] C. D. Anderson, “The positive electron”, *Phys. Rev.*, vol. 43 (1933) 491, [doi:10.1103/PhysRev.43.491](https://doi.org/10.1103/PhysRev.43.491).

-
- [9] C. M. G. Lattes, G. P. S. Occhialini, and C. F. Powell, “Observations on the Tracks of Slow Mesons in Photographic Emulsions”, *Nature*, vol. 160 (1947) 486, [doi:10.1038/160486a0](https://doi.org/10.1038/160486a0).
- [10] S. H. Neddermeyer and C. D. Anderson, “Note on the nature of cosmic-ray particles”, *Phys. Rev.*, vol. 51 (1935) 10, [doi:10.1103/PhysRev.51.884](https://doi.org/10.1103/PhysRev.51.884).
- [11] David J. Griffiths, “Introduction to elementary particles”, Physics textbook Wiley, 2008. <https://books.google.co.in/books?id=w9Dz56myXm8C> ISBN 9783527406012.
- [12] M. Herrero, “The Standard model”, *NATO Sci. Ser. C*, vol. 534 (1999) 1, [doi:10.1007/978-94-011-4689-0_1](https://doi.org/10.1007/978-94-011-4689-0_1).
- [13] S. Weinberg, “A Model of Leptons”, *Phys. Rev. Lett.*, vol. 19 (1967) 1264, [doi:10.1103/PhysRevLett.19.1264](https://doi.org/10.1103/PhysRevLett.19.1264).
- [14] D. J. Gross and F. Wilczek, “Ultraviolet Behavior of Non-Abelian Gauge Theories”, *Phys. Rev. Lett.*, vol. 30 (1973) 1346, [doi:10.1103/PhysRevLett.30.1343](https://doi.org/10.1103/PhysRevLett.30.1343).
- [15] H. D. Politzer, “Reliable Perturbative Results for Strong Interactions”, *Phys. Rev. Lett.*, vol. 30 (1973) 1346, [doi:10.1103/PhysRevLett.30.1346](https://doi.org/10.1103/PhysRevLett.30.1346).
- [16] F. Halzen and A. Martin, “Quarks and leptons: an introductory course in modern particle physics”, Wiley, 1984. <https://books.google.co.in/books?id=zwDvAAAAMAAJ>.
- [17] CMS Wiki Pages, <https://wiki.physik.uzh.ch/cms/latex:tikz:timescales>.
- [18] M. E. Peskin and D. V. Schroeder, “An Introduction to quantum field theory”, Addison-Wesley, Reading, USA, 1995. <http://www.slac.stanford.edu/mpe-skin/QFT.html>. ISBN 9780201503975, 0201503972.

-
- [19] CDF Collaboration, “Observation of top quark production in $p\bar{p}$ collisions with the Collider Detector at Fermilab”, *Phys. Rev. Lett.*, vol. 74, (1995) 2626, [doi:10.1103/PhysRevLett.74.2626](https://doi.org/10.1103/PhysRevLett.74.2626).
- [20] D0 Collaboration, “Search for high mass top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV”, *Phys. Rev. Lett.*, vol. 74, (1995) 2422, [doi:10.1103/physrevlett.74.2422](https://doi.org/10.1103/physrevlett.74.2422).
- [21] Particle Data Group, P. A. Zyla *et al.*, “Review of Particle Physics”, *Prog. Theor. Exp. Phys.*, vol. 2020 (2020) 083C01, [doi:10.1093/ptep/ptaa104](https://doi.org/10.1093/ptep/ptaa104).
- [22] http://www.yorku.ca/menary/HEP_resources.html.
- [23] C. Patrignani *et al.*, “Review of Particle Physics”, *Chinese Physics C*, vol. 40 (2016) 100001, [doi:10.1088/1674-1137/40/10/100001](https://doi.org/10.1088/1674-1137/40/10/100001).
- [24] https://en.wikipedia.org/wiki/Standard_Model.
- [25] D. Gross and F. Wilczek, “Asymptotically Free Gauge Theories – I”, *Phys. Rev. D*, vol. 8 (1973) 3633, [doi:10.1103/PhysRevD.8.3633](https://doi.org/10.1103/PhysRevD.8.3633).
- [26] H. D. Politzer, “Asymptotic Freedom: An Approach to strong Interactions”, *Phys. Rept.*, vol. 14 (1974) 129, [doi:10.1016/0370-1573\(74\)90014-3](https://doi.org/10.1016/0370-1573(74)90014-3).
- [27] Feynman, Richard “Chapter 1, QED: The Strange Theory of Light and Matter”, *Princeton University Press.*, p.6. ISBN 978-0691125756 (1985).
- [28] E.Fermi, “Quantum Theory of Radiation”, *Rev. Mod. Phys.*, vol. 4 (1932) 87, [doi:10.1103/RevModPhys.4.87](https://doi.org/10.1103/RevModPhys.4.87).
- [29] P.A.M.Dirac, “The Quantum Theory of the Emission and Absorption of Radiation”, in *Proceedings of the Royal Society of London A*, vol. 114 (1927) 243, [doi:10.1098/rspa.1927.0039](https://doi.org/10.1098/rspa.1927.0039).

-
- [30] S. L. Glashow, “Partial Symmetries of Weak Interactions”, *Nucl. Phys.*, vol. 22 (1961) 579, [doi:10.1016/0029-5582\(61\)90469-2](https://doi.org/10.1016/0029-5582(61)90469-2).
- [31] A. Salam and J. C. Ward, “Electromagnetic and weak interactions”, *Phys. Lett.*, vol. 13 (1964) 168, [doi:10.1016/0031-9163\(64\)90711-5](https://doi.org/10.1016/0031-9163(64)90711-5).
- [32] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons”, *Phys. Rev. Lett.*, vol. 13 (1964) 321, [doi:10.1103/PhysRevLett.13.321](https://doi.org/10.1103/PhysRevLett.13.321).
- [33] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons”, *Phys. Rev. Lett.*, vol. 13 (1964) 508, [doi:10.1103/PhysRevLett.13.508](https://doi.org/10.1103/PhysRevLett.13.508).
- [34] Suman B. Beri, “Concepts and Methods of High Energy and Nuclear Physics”, 2007. ISBN 81-88901-27-X.
- [35] ATLAS Collaboration and CMS Collaboration, “Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments”, *Phys. Rev. Lett.*, vol. 114 (2015) 191803, [doi:10.1103/PhysRevLett.114.191803](https://doi.org/10.1103/PhysRevLett.114.191803).
- [36] ATLAS Collaboration, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, *Phys. Lett. B*, vol. 716 (2012) 1, [doi:10.1016/j.physletb.2012.08.020](https://doi.org/10.1016/j.physletb.2012.08.020).
- [37] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, *Phys. Lett. B*, vol. 716 (2012) 30, [doi:10.1016/j.physletb.2012.08.021](https://doi.org/10.1016/j.physletb.2012.08.021).
- [38] H. D. Politzer, “Asymptotic Freedom: An Approach to Strong Interactions”, *Phys. Rept.* 14 (1974) 129, [doi:10.1016/0370-1573\(74\)90014-3](https://doi.org/10.1016/0370-1573(74)90014-3).
- [39] M. Tanabashi *et al.*, “Review of Particle Physics”, *Phys. Rev. D*, vol. 98 (2018) 030001, [doi:10.1103/PhysRevD.98.030001](https://doi.org/10.1103/PhysRevD.98.030001).

- [40] S. P. Martin, “A Supersymmetry primer”, *Advanced Series on Directions in High Energy Physics*, (1998) 1, [doi:10.1142/9789812839657_0001](https://doi.org/10.1142/9789812839657_0001).
- [41] “Planck 2015 Constraints on Dark Energy and Inflation”, <https://darkmatterdarkenergy.com/tag/planck-2015/>.
- [42] <https://science.nasa.gov/astrophysics/focus-areas/what-is-dark-energy>.
- [43] M. P. Decowski, “KamLAND’s precision neutrino oscillation measurements”, *Nucl. Phys. B*, vol. 908 (2016) 52, [doi:10.1016/j.nuclphysb.2016.04.014](https://doi.org/10.1016/j.nuclphysb.2016.04.014).
- [44] T2K Collaboration, “Precise Measurement of the Neutrino Mixing Parameter θ_{23} from Muon Neutrino Disappearance in an Off-Axis Beam”, *Phys. Rev. Lett.*, vol. 112 (2014) 181801, [doi:10.1103/PhysRevLett.112.181801](https://doi.org/10.1103/PhysRevLett.112.181801).
- [45] A. Pomarol, “Beyond the Standard Model”, in *High-energy Physics. Proceedings, 18th European School, ESHEP 2010, Raseborg, Finland, 20 Jun. - 3 Jul., 2010*, (2012) 115–151, [arXiv:10.48550/arXiv.1202.1391](https://arxiv.org/abs/10.48550/arXiv.1202.1391).
- [46] CMS Collaboration, “Highlights from the Compact Muon Solenoid CMS Experiment”, *Universe*, vol. 5 (2019) 28, [doi:10.3390/universe5010028](https://doi.org/10.3390/universe5010028).
- [47] <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined>.
- [48] J. M. Campbell, R. K. Ellis, F. Maltoni and S. Willenbrock, “Associated production of a Z Boson and a single heavy quark jet”, *Phys. Rev. D*, vol. 69 (2004) 074021, [doi:10.1103/PhysRevD.73.054007](https://doi.org/10.1103/PhysRevD.73.054007).
- [49] J. M. Campbell, R. K. Ellis, F. Maltoni and S. Willenbrock, “Production of a Z boson and two jets with one heavy-quark tag”, *Phys. Rev. D*, vol. 73 (2006) 054007, [arXiv:10.48550/arXiv.hep-ph/0510362](https://arxiv.org/abs/10.48550/arXiv.hep-ph/0510362).

-
- [50] S.Chatrchyan *et al.* [CMS], “Search for the Standard Model Higgs Boson Produced in Association with a W or a Z Boson and Decaying to Bottom Quarks”, *Phys. Rev. D*, vol. 89 (2014) 012003, [doi:10.1103/PhysRevD.89.012003](https://doi.org/10.1103/PhysRevD.89.012003).
- [51] B.Holdom, W.S.Hou, T.Hurth, M.L.Mangano, S.Sultansoy and G.Unel, “Four Statements about the Fourth Generation”, *PMC Phys. A*, vol. 3 (2009) 4, [doi:10.1186/1754-0410-3-4](https://doi.org/10.1186/1754-0410-3-4).
- [52] L.J.Hall, D.Pinner and J.T.Ruderman, “A Natural SUSY Higgs Near 125 GeV”, *JHEP*, vol. 04 (2012) 131, [doi:10.1007/JHEP04\(2012\)131](https://doi.org/10.1007/JHEP04(2012)131).
- [53] D.Choudhury, T.M.P.Tait and C.E.M.Wagner, “Beautiful mirrors and precision electroweak data”, *Phys. Rev. D*, vol. 65 (2002) 053002, [doi:10.1103/physrevd.65.053002](https://doi.org/10.1103/physrevd.65.053002).
- [54] G.Bozzi, J.Rojo and A.Vicini, “The Impact of PDF uncertainties on the measurement of the W boson mass at the Tevatron and the LHC”, *Phys. Rev. D*, vol. 83 (2011) 113008, [doi:10.1103/PhysRevD.83.113008](https://doi.org/10.1103/PhysRevD.83.113008).
- [55] M.Beccaria, N.Orlando, G.Panizzo, F.M.Renard and C.Verzegnassi, “The Relevance of polarized bZ production at LHC”, *Phys. Lett. B*, vol. 713 (2012) 457, [doi:10.1016/j.physletb.2012.06.026](https://doi.org/10.1016/j.physletb.2012.06.026).
- [56] CDF collaboration, “Measurement of cross sections for b jet production in events with a Z boson in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV”, *Phys. Rev. D*, vol. 79 (2009), [doi:10.1103/PhysRevD.79.052008](https://doi.org/10.1103/PhysRevD.79.052008).
- [57] D0 Collaboration, “A measurement of the ratio of inclusive cross sections of the ratio of inclusive cross sections $\sigma(p\bar{p} \rightarrow Z + \text{bjet})/\sigma(p\bar{p} \rightarrow Z + \text{jet})$ at $\sqrt{s} = 1.96$ TeV”, *Phys. Rev. D*, vol.83 (2011) 031105, [doi:10.1103/PhysRevD.83.031105](https://doi.org/10.1103/PhysRevD.83.031105).

-
- [58] CMS Collaboration, “Measurement of the production cross sections for a Z boson and one or more b jets in pp collisions at $\sqrt{s} = 7$ TeV”, *JHEP*, vol. 6 (2014) 120, [doi:10.1007/JHEP06\(2014\)120](https://doi.org/10.1007/JHEP06(2014)120).
- [59] ATLAS Collaboration, “Measurement of differential production cross-sections for a Z boson in association with b-jets in 7 TeV proton-proton collisions with the ATLAS detector”, *JHEP*, vol. 10 (2014) 141, [arXiv:10.48550/arXiv.1407.3643](https://arxiv.org/abs/10.48550/arXiv.1407.3643).
- [60] , CMS Collaboration, “Measurements of the associated production of a Z boson and b jets in pp collisions at $\sqrt{s} = 8$ TeV”, *Eur. Phys. J. C*, vol. 77 (2017) 751, [doi:10.1140/epjc/s10052-017-5140-y](https://doi.org/10.1140/epjc/s10052-017-5140-y).
- [61] ATLAS Collaboration, “Measurements of the production cross-section for a Z boson in association with b-jets in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *JHEP*, vol.7 (2020) 044, [doi:10.1007/JHEP07\(2020\)044](https://doi.org/10.1007/JHEP07(2020)044).

Chapter 2

Experimental Setup

The European Council for Nuclear Research (CERN) was established in 1954 as a collaborative effort of twelve European countries. It is recognized as the most esteemed particle physics research laboratory worldwide ever constructed. The series of particle accelerators and other infrastructure at CERN are used to investigate the fundamental nature of matter. The incredible breakthroughs of CERN before the LHC [1] era, includes the discovery of weak neutral currents inside the Gargamelle bubble chamber in 1973, W and Z bosons by colliding protons and anti-protons in 1983, and Charge-Parity (CP) violation in 1964, etc. Nevertheless, the discovery of the Higgs boson with a mass of 125 GeV on 4th July 2012 is the most significant physics discovery at CERN by the LHC experiments. This resulted in the Nobel Prize in physics being awarded in 2013 to theorists Peter Higgs and Francois Englert, who predicted the presence of the Higgs boson theoretically [2-4].

The physics results presented in this thesis used data collected at $\sqrt{s} = 13$ TeV by the CMS detector at the LHC. This Chapter describes the experimental setup used for data collection.

2.1 The Large Hadron Collider

The LHC at CERN is the world’s most powerful circular hadron accelerator to date that accelerates the protons and heavy ions very close to the speed of light. The machine was designed to produce high energy pp collisions at $\sqrt{s} = 14 \text{ TeV}$ and peak instantaneous luminosity up to $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, resulting in a pp crossing rate of 40 MHz. The proton beams are accelerated in bunches and each bunch contains around 1.15×10^{11} protons. Each beam has 2808 bunches and each bunch spacing is 25 ns. The record-breaking energy achieved by LHC enables it to study the SM and BSM processes with high precision and search for novel physics signatures in the phase spaces that were inaccessible before.

The LHC is located around 50–175 m underground between the Geneva lake (Switzerland), Jura Mountains, and Cessy (France) in a tunnel spanning 27 km in circumference. The collider tunnel is made up of two adjacent parallel beam pipes that meet at four points, each containing a proton beam, which travels in opposite directions around the ring. It also accelerates lead ion ($^{208}\text{Pb}^{82+}$) beams, resulting in heavy-ion (HI) collisions. This thesis work is based on the data collected from pp collisions at the LHC.

The LHC injected its first beam on 10th September 2008, but was forced to stop due to a defective electrical connection between the superconducting magnets. Finally, a year later, on 20th November 2009, the first pp collision occurred at a $\sqrt{s}$ of 900 GeV. Then, on 30th March 2010, the energy of the proton beam was boosted to 3.5 TeV, resulting in a $\sqrt{s} = 7 \text{ TeV}$. The second period of data collection at $\sqrt{s} = 8 \text{ TeV}$ started from 5th April 2012 and ended on 6th February 2013. This collision time period (2010–2013) is known as **LHC “Run1”**. After a two-year long shutdown (LS1) for the upgrades, the LHC resumed its operations again in 2015 known as **LHC “Run2”**, with each proton beam of energy 6.5 TeV resulting in a

$\sqrt{s} = 13$ TeV. On 3rd December 2018, Run2 ended after delivering a large amount of data and the LHC entered its second long shutdown phase (LS2). After LS2, **LHC “Run3”** is planned for $\sqrt{s} = 14$ TeV between 2022–2024, with an expected integrated luminosity of 300 fb^{-1} . Further, during the long shutdown phase (LS3), the accelerator infrastructure will be modified to deliver unprecedented instantaneous luminosity from 5.0 to $7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ providing a total of 3000 to 4000 fb^{-1} data. The project is called HL-LHC, which stands for High-Luminosity LHC. The detailed schedule of LHC execution procedures is shown in Figure 2.1.

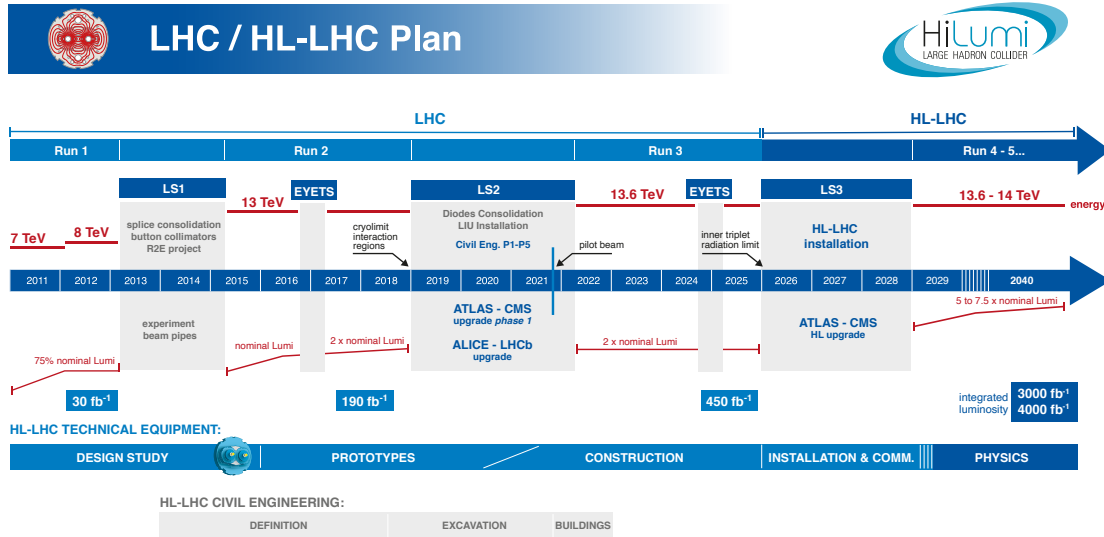


Figure 2.1: Overview of the timeline for the HL-LHC upgrade [5].

Since it is impossible to obtain a single proton traveling almost at the speed of light, so proton must be accelerated. Furthermore, achieving the speed of light with a single linear accelerator is not possible because it would require a long length apparatus. So, before being injected into the LHC machine, the protons pass through a series of accelerators, each of which increases the energy of the protons and transfers them to the next stage. On the other hand, cyclic accelerators known as synchrotrons

can use an applied magnetic field to keep the particles on their circular path and use an electric field to accelerate them. To keep the energetic proton beams moving throughout the LHC ring, 1232 dipole magnets with core fields ranging from 0.54 T (for 450 GeV proton energy) to 8.33 T (for 7 TeV proton energy) are utilized. Besides the dipole magnets, the LHC uses around 850 quadrupole magnets, out of which 392 are more powerful lattice quadrupoles. These quadrupoles are used to confine proton beams along the four LHC collision locations, hence enhancing the probability of bunch interactions. The size and position of the beam can also be corrected by using the octupole and sextuple magnets. The LHC magnets are composed of niobium-titanium (NbTi) and are cooled to 1.9°K in the cryogenic system using ~ 96 tonnes of superfluid helium to provide the required strong magnetic field at the LHC. The radio-frequency (RF) cavities are situated along the beam pipe and provide acceleration to the proton beams using alternating electric fields. Two independent RF systems are provided, one for each beam. Due to restrictions on the magnetic field corresponding to the ring circumference, a synchrotron can only be operated at a limited range of particle energies. The best strategy is to use several different synchrotrons of different sizes to gradually accelerate the particles. The LHC accelerator complex is discussed in the next Subsection.

2.1.1 The CERN accelerator complex

The proton acceleration process begins with a gas tank filled with H_2 molecules. The electric field is used to remove electrons from the hydrogen atoms and convert them into protons. These protons are then injected into a linear accelerator, LINAC2, that uses RF cavities to accelerate them to 50 MeV. Following that, the accelerated protons are sent to the Proton Synchrotron Booster (PSB), which is made of four superimposed synchrotron rings that assemble the protons into bunches and accelerate them to 1.4 GeV. The size of the bunch and trajectory of the particles are controlled by a series of superconducting magnets. Dipole magnets are used to bend the track

of the particles, whereas quadrupole magnets are utilized to focus the beam. Six proton bunches are transferred consecutively to the Proton Synchrotron (PS), where they are accelerated to 25 GeV and sorted into 72 bunches divided by 25 ns in time. The separated proton bunches from the PS are accelerated to 450 GeV in the Super Proton Synchrotron (SPS). The SPS splits a single beam into two counter-rotating beams, which are subsequently injected into the LHC beam pipes. The PSB, PS, and SPS accelerators use common electromagnets to keep the particles moving. Proton beams are accelerated for many hours inside the rings to achieve the desired energy before the collisions. The beam energy was 3.5 TeV and 4 TeV in Run1 and then in Run2, it was boosted to 6.5 TeV.

Besides the proton beams, the LHC has a proton-lead and lead-lead nuclei collision program at various center-of-mass energies. The goal of the heavy-ion program is to better understand the state of matter known as Quark-Gluon Plasma (QGP), which defined the early stages of the universe. It allows us to investigate QCD matter under extreme conditions of temperature, density, and parton momentum fraction. The Pb ions are initially accelerated using the LINAC3 linear accelerator, and the Low-Energy Ion Ring (LEIR) is used as an ion storage and cooling unit. The ions are then further accelerated by the PS and SPS before being injected into the LHC ring. The detail of the LHC injection is shown in Figure 2.2. There are four main sites at which the two counter-rotating beams are focused and made to intersect. At each of the four intersection points (IP), detectors are placed to record the collision events as illustrated in Figure 2.3. They are as follows:

- **A Toroidal LHC ApparatuS (ATLAS):** It is a general-purpose particle detector positioned at IP1, designed to study the SM processes as well as the search for BSM physics [7].
- **Compact Muon Solenoid (CMS):** It is a general-purpose particle detector placed at IP5, with the same layout and identical physics goals as an ATLAS

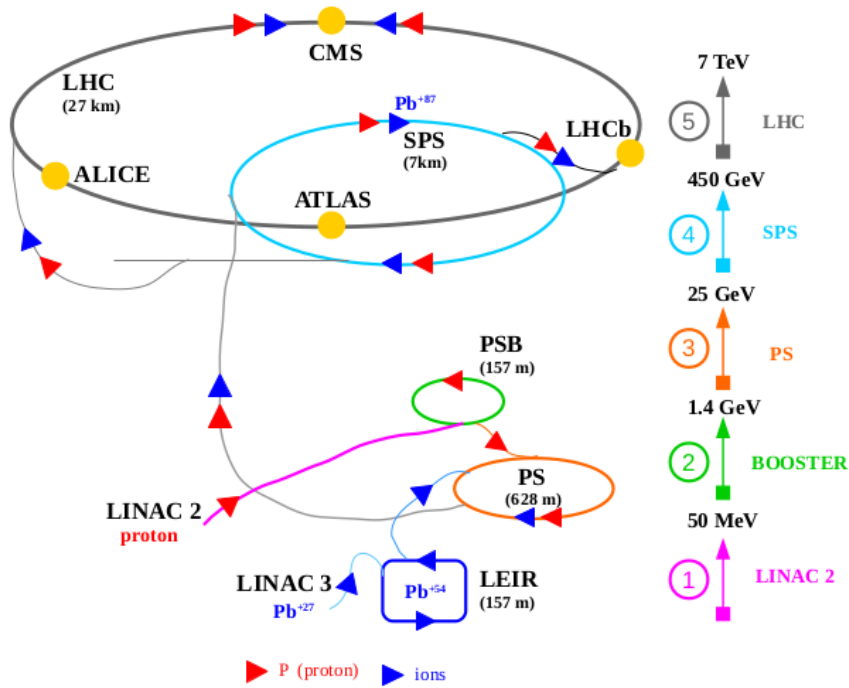


Figure 2.2: Detailed diagram of the LHC injection series for protons and Pb nuclei. The proton and lead ion trajectories are indicated with red and blue arrows, respectively. The position of the four interaction points is also included.

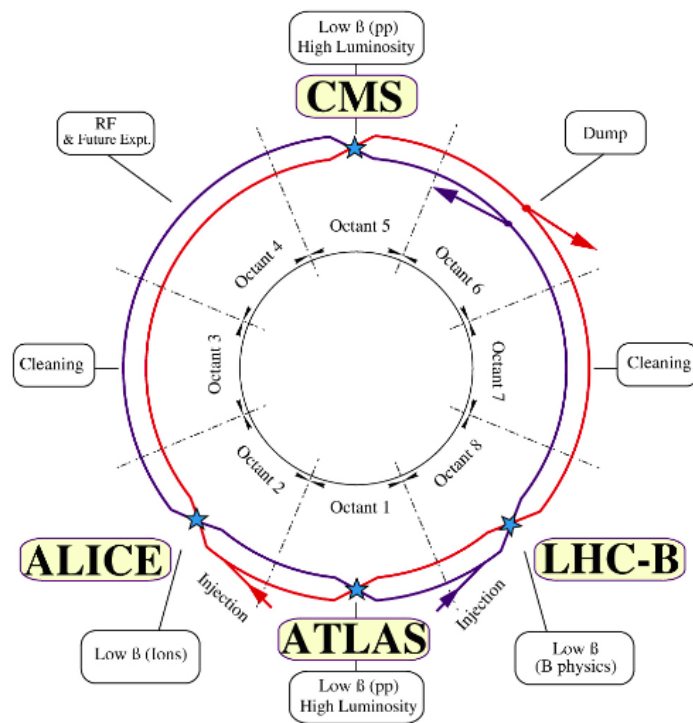


Figure 2.3: LHC schematic and the four main interaction points [6].

detector [8].

- **A Large Ion Collider Experiment (ALICE)**: This is a dedicated particle detector located at IP2, which is used to measure the characteristics of nuclear matter at high energy densities [9]. The ALICE detector allows the study of QGP [10] as well as other HI physics events such as proton-lead and lead-lead collisions.
- **Large Hadron Collider beauty (LHCb)**: It is a single-sided spectrometer installed at IP8 to explore the physics of the b quarks, the mystery of matter-antimatter asymmetry, and the measurement of the CP-violation parameters [11].

In addition to these, there are three smaller experiments, **Large Hadron Collider forward (LHCf)** [12], **TOTAL Elastic and diffractive cross section Measurement (TOTEM)** [13], and **Monopole and Exotics Detector At the LHC (MoEDAL)** [14] share the underground caverns of the ATLAS, CMS, and LHCb experiments, respectively. The primary goal of the LHCf experiment is to better understand the cosmic rays by using the particles produced in a forward direction inside the LHC. It is useful for calibrating particle detectors and hard interaction models employed in studies of cosmic radiation. It is the smallest experiment at the LHC, located near the ATLAS experiment region and very close to the beam pipe, consisting of two very forward imaging calorimeters. The TOTEM detector is designed for elastic low angle scattering of protons and is placed in the very forward region close to the CMS detector, which is inaccessible by other LHC experiments. It is made up of Roman Pots and two telescopes (a Cathode Strip Chamber and a Gas Electron Multiplier) for measuring the momentum of the charged particles. The MoEDAL experiment was first used in 2011 in order to look for magnetic monopoles or strongly ionizing **Stable Massive Particles (SMPs)** predicted by BSM theories [15].

The performance of the LHC depends on many parameters, which are discussed

in the next Subsections.

2.1.2 LHC parameters and performance

The LHC is a two-ring superconducting proton accelerator and collider designed to explore the TeV energy scale. The design of LHC is based on extremely basic concepts associated with the latest technologies. The following are the key components of the LHC:

- **Measurements of beam luminosity at the LHC:** Luminosity is one of the main parameters which is used to characterize the performance of the LHC. It measures the ability of a collider to produce the required number of interactions. When beams collide at high luminosities, there will be more particle interactions. All collider experiments aim to maximize their luminosities so that physicists can have more data to analyze. The expected event rate (dN/dt) produced in a particular process with a cross section (σ) is calculated as follows:

$$\frac{dN}{dt} = \sigma \mathcal{L} \quad (2.1)$$

where $\mathcal{L}$ is the instantaneous luminosity. The $\mathcal{L}$ of two colliding beams depends on several beam parameters:

$$\mathcal{L} = \frac{\gamma f k_B N_p^2}{4\pi \epsilon_n \beta^*} F \quad (2.2)$$

where f represents the revolution frequency at the LHC, γ is the relativistic Lorentz factor, k_B is the number of bunches, N_p is the number of protons/bunch, β^* is the betatron function defined at IP, ϵ_n is the normalized transverse beam emittance, and F represents the reduction factor which depends on crossing angle, length of the proton bunch, and the transverse beam

size.

The integrated luminosity (L) is obtained by integrating the instantaneous luminosity over a specific time interval. The integrated luminosity is a measurement of the size of the collected data. The integral is taken over the sensitive time, i.e., excluding any possible dead time (the time when the system is not able to record another event):

$$L = \int_{t_0}^{t_1} \mathcal{L} dt \quad (2.3)$$

Figure 2.4 (left) shows the cumulative luminosity delivered to the CMS detector during Run1 (2010–2012) and Run2 (2015–2018). The LHC delivered a total integrated luminosity of $\sim 163 \text{ fb}^{-1}$ during Run2 out of which CMS recorded $\sim 150 \text{ fb}^{-1}$, as seen in Figure 2.4 (right). This thesis work is based on the analysis of data collected by the CMS detector during 2016–2018 leading to total integrated luminosity of $\sim 137 \text{ fb}^{-1}$.

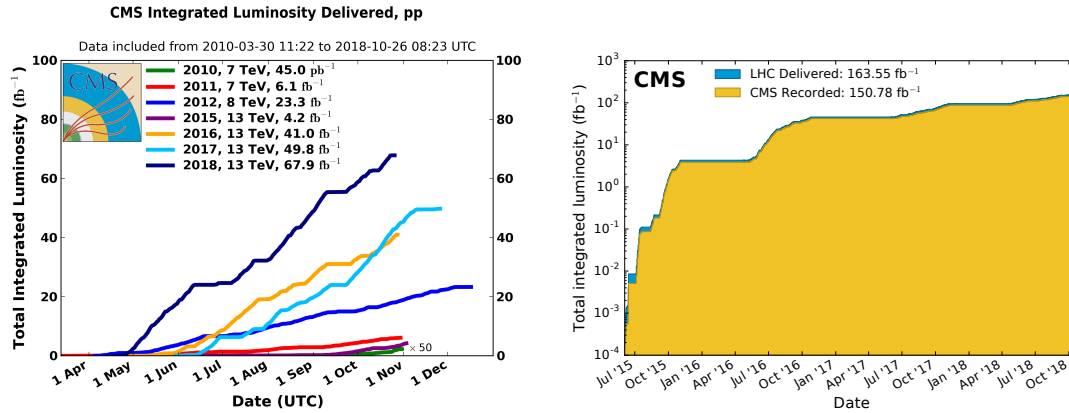


Figure 2.4: Cumulative luminosity delivered to the CMS experiment during pp collisions at various $\sqrt{s}$ (left). Total integrated luminosity delivered by LHC and recorded by the CMS detector during Run2 data-taking period [16] (right).

During the HL-LHC phase, several upgrades are planned for the accelerator complex, which should allow for an increase in the peak instantaneous lumi-

luminosity to at least 5 times the nominal value of Run3 [17]. The plan for the HL-LHC is to collect data of 3000–4000 fb⁻¹ by 2040, increasing the total recorded luminosity by at least 10 times of Run3.

- **Center-of-mass energy:** The center-of-mass energy ($\sqrt{s}$) of a system of particles, is the energy measured in the center-of-mass reference frame (E_{com}). Consider the collision of two particles with energy, momentum, and rest mass (E_1, p_1, m_1), (E_2, p_2, m_2) at an incident angle θ . The parameter “s” is called Mandelstam variable which is Lorentz invariant given as:

$$\begin{aligned}
 E_{\text{com}} = s &= (E_1 + E_2)^2 - (\vec{p}_1 \cdot c + \vec{p}_2 \cdot c)^2 \\
 s &= (E_1^2 + E_2^2 + 2E_1E_2) - (p_1^2c^2 + p_2^2c^2 + 2p_1c p_2c \cos\theta) \\
 s &= (E_1^2 - p_1^2c^2) + (E_2^2 - p_2^2c^2) + (2E_1E_2 - 2p_1c p_2c \cos\theta) \quad (2.4) \\
 s &= (m_1c^2)^2 + (m_2c^2)^2 + 2(E_1E_2 - p_1c p_2c \cos\theta) \\
 s &= m_1^2 + m_2^2 + 2(E_1E_2 - p_1p_2 \cos\theta)
 \end{aligned}$$

where $\vec{p}$ denotes momentum, E represents the energy of the two initial state particles, and $c = 1$ is considered as “natural unit” at high energy.

In collider experiments, when two identical particles of equal momentum collide head-on ($\theta = 180^\circ$), their center-of-mass energy is determined as follows:

$$\begin{aligned}
 s &\sim 4E_1E_2 \quad \text{if } E_i \gg m_i c^2, E_i \sim p_i c \\
 \sqrt{s} &= 2E_{\text{beam}} \quad \text{if } m_i = m, E_i = E_{\text{beam}}
 \end{aligned} \quad (2.5)$$

where E_{beam} and m_i represent the energy of the incoming beams and the mass of the interacting particles, respectively. The $\sqrt{s}$ is the amount of energy

available for new particle production as a result of the collision in the center-of-mass frame. So, in the case of pp collision at LHC, with 7 TeV per proton, the $\sqrt{s}$ is 14 TeV.

In the case of the fixed target, if m_1 is the projectile and m_2 is the target, then $p_2 = 0$ and $E_2 = m_2 c^2$, in the center-of-mass frame. Their center-of-mass energy will be as follows.

$$s \sim 2E_1 m_2 c^2 \quad \text{if } E_1 \gg m_1 c^2 \quad (2.6)$$

$$\sqrt{s} \sim \sqrt{2E_1 m_2 c^2}$$

So, for a fixed target, considering a proton moving with energy E_1 colliding with a fixed target formed from a proton at rest ($m_2 c^2 \sim 10^{-3} \text{ TeV}$), so to get $\sqrt{s} = 14 \text{ TeV}$, the amount of E_1 will be:

$$\sqrt{s} \sim \sqrt{2E_1 m_2 c^2}$$

$$14 = \sqrt{2E_1 10^{-3}}$$

$$E_1 \sim 10^5 \text{ TeV}$$

As a result, the collider is preferred over a fixed target.

- **Pileup:** At high luminosity, the probability of multiple pp collisions in the bunch crossing is increased. These multiple collisions will result in overlapping of particle interaction events in the detector. The number of additional collisions that result in low momentum transfer occurrences is referred to as pileup (PU). Depending on time, the pileup can be classified as in-time or out-of-time. In-time PU refers to additional pp collisions occurring in the same bunch crossing as the collision of interest. While out-time PU refers to

additional pp collisions occurring in bunch crossings just before and after the collision of interest. The PU collisions are substantially influenced by the beam squeezing via the reduction of the β^* parameter. The major improvements in the beam squeezing occurred during the end-of-year technical stops between 2015 and 2016, and in the late part of 2017. This resulted in an increase in the number of PU interactions during Run2. The pileup events or an average number of events per crossing are given by:

$$\langle N_p \rangle = \sigma_{\text{inel}} \cdot \mathcal{L} \cdot \tau_b \quad (2.7)$$

where σ_{inel} is the inelastic pp cross section and τ_b is the bunch spacing. For example, at the typical values of LHC parameter at nominal LHC collision energy $\sqrt{s} = 13 \text{ TeV}$ for 2017, the $\sigma_{\text{inel}} = 80.0 \text{ mb}$, and $\tau_b = 25 \text{ ns}$, the average number of PU events is ~ 38 . The LHC standard values for the minimum bias cross sections are taken from the theoretical prediction from Pythia. Figure 2.5 shows the average number of interactions per crossing for pp collisions in 2011–2012 and 2015–2018. The overall mean values and the minimum bias cross sections are also shown. These plots only show data that passed the “golden” certification (i.e., all CMS subdetectors were flagged to be ok for any kind of usage in physics analysis).

2.1.3 Proton-Proton collisions at the LHC

At the LHC, protons are accelerated close to the speed of light before colliding. Since each proton is made up of valence quarks (uud) surrounded by gluons and sea quarks which are created in the form of $q\bar{q}$ pairs through the vacuum fluctuations ($g \rightarrow q\bar{q}$). The quarks inside the proton behave as free particles in the LHC environment where pp collide at very high energy (large momentum transfer or short distance interaction). However, as quarks can emit gluons ($q \rightarrow qg$) and gluons themselves can split

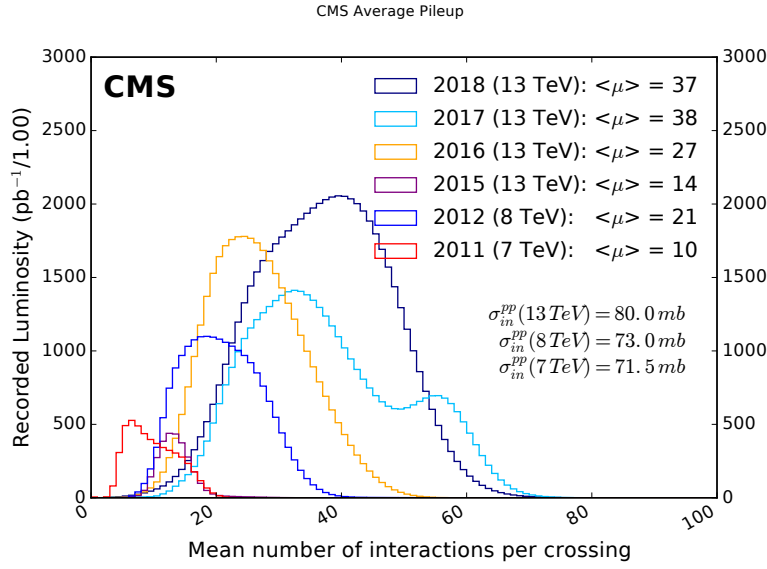


Figure 2.5: Distribution of pileup and mean number of interactions per bunch crossing for the 2011–2012 and 2015–2018 for pp collisions recorded by the CMS experiment [16].

into additional gluons ($g \rightarrow gg$) or gluons can split into quark-antiquark pairs ($g \rightarrow q\bar{q}$), and one of these quarks can collide. A parton-parton interaction is referred to as a hard interaction if it involves a large momentum transfer and produces very high energetic particles. Besides the hard interaction, there are many soft parton-level interactions that take place with low momentum transfer. As particles are accelerated, they can emit gluons or photons (via bremsstrahlung), and gluons further can emit gluons or produce $q\bar{q}$ pairs. The continuous emission of these partons is called a parton shower (PS). The PS produced by the incoming partons is called initial-state-radiation (ISR), while the emission associated with outgoing partons is called final-state-radiation (FSR). In addition to these processes, there are large numbers of partons that do not participate in the parton-parton interaction. These partons are of low transverse momentum and travel along the beam pipe with a small contribution to the particle activities in the forward detectors of the experiment. This contribution is referred to as beam-beam remnants (BBR). The parton densities become too large inside the proton at high energy collisions, several parton pairs may collide within one single pp collision. This class of events is known as

Multiple Parton Interactions (MPI). The MPI is broadly divided into soft and hard contributions. The soft contribution of MPI together with other processes such as ISR, FSR, and BBR is known as the Underlying event (UE). After the final state PS at energies of the order 1 GeV, the strong interaction coupling rises at a long distance and the process of hadronization happens i.e. recombination of partons into colorless hadrons. The collimated sprays of hadrons known as “hadronic jets” are observed in the surrounding of the collision point in the detector. Figure 2.6 depicts the different processes during pp collisions.

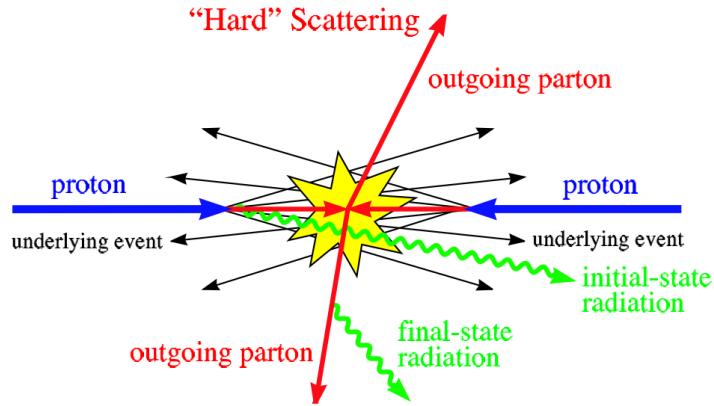


Figure 2.6: Schematic of a hard scattering pp collision [18].

2.2 The CMS detector

The CMS is one of the two general-purpose detectors at the LHC and offers an extensive variety of physics opportunities by pp and HI collisions to precisely measure SM processes and hunt for BSM physics. The CMS detector is situated in an underground hall about 100 m below the ground at the IP5 of the LHC in the Cessy village of France. The detector was given its name as a result of three distinguishing characteristics:

- It has compact geometry, a weight of 14,000 tonnes, a length of 28.7 m, and an overall diameter of 15 m.

- It is designed to achieve optimal muon identification and momentum resolution.
- It has a solenoid magnet that produces a 3.8 T magnetic field with significant bending strength for precise measurement of charged particle momenta. This is one of the biggest superconducting solenoids ever constructed.

The CMS has an onion-like structure, with each layer working as a subdetector to stop, monitor, or measure various types of particles that originate from pp or HI collisions, as shown in Figure 2.7. The central subdetector parts that are parallel to the beamline and close to the primary interaction point are known as the barrel regions, while subdetectors that are placed further from the interaction point and detect particles emitted at a low angle with respect to the beamline are known as the endcap regions.

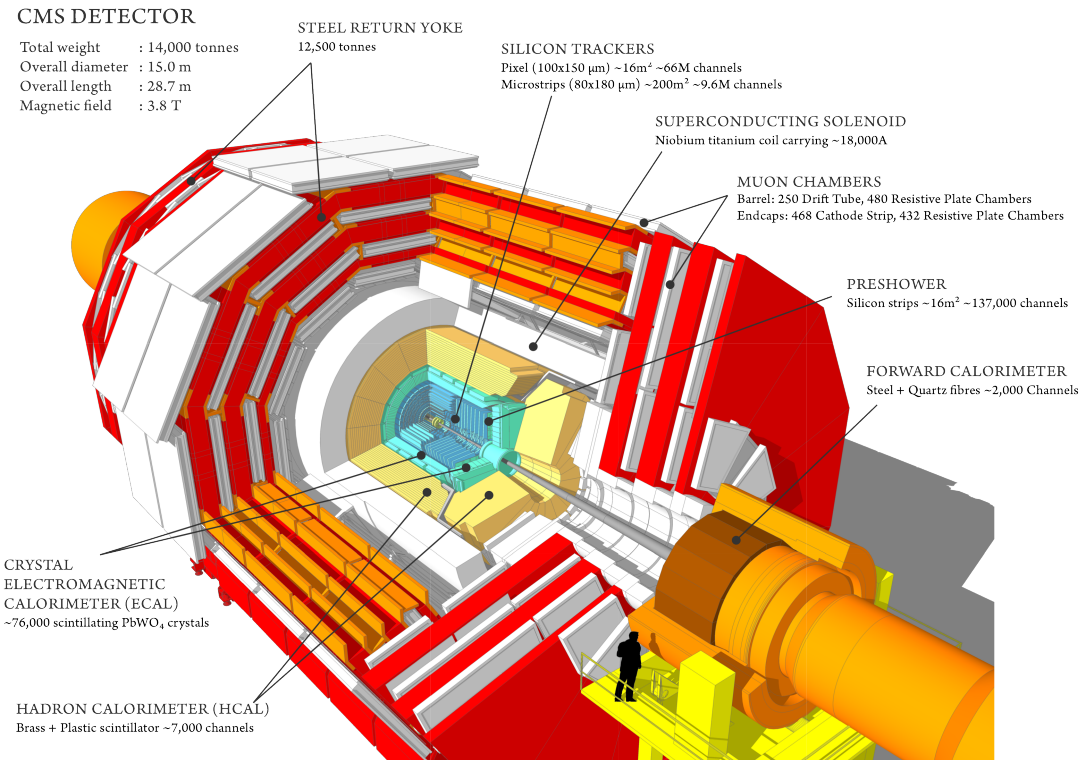


Figure 2.7: A perspective view of CMS detector and its subdetectors [19].

A 3.8 T magnetic field is produced by the superconducting solenoid in the center of the CMS detector. The tracker system finds the tracks of charged particles that are coming from collisions. Calorimeters (Electromagnetic and Hadron) are placed outside the tracker to stop the particles and measure their energy. The tracker and both calorimeters are placed inside the coil of the magnet making it very compact. The iron return yoke with the return field of the solenoid around 1.8 T is located within the muon system. The details of all subdetectors of the CMS are discussed in the next Sections. Figure 2.8 shows the kind of signatures left by different types of particles passing through the CMS detector.

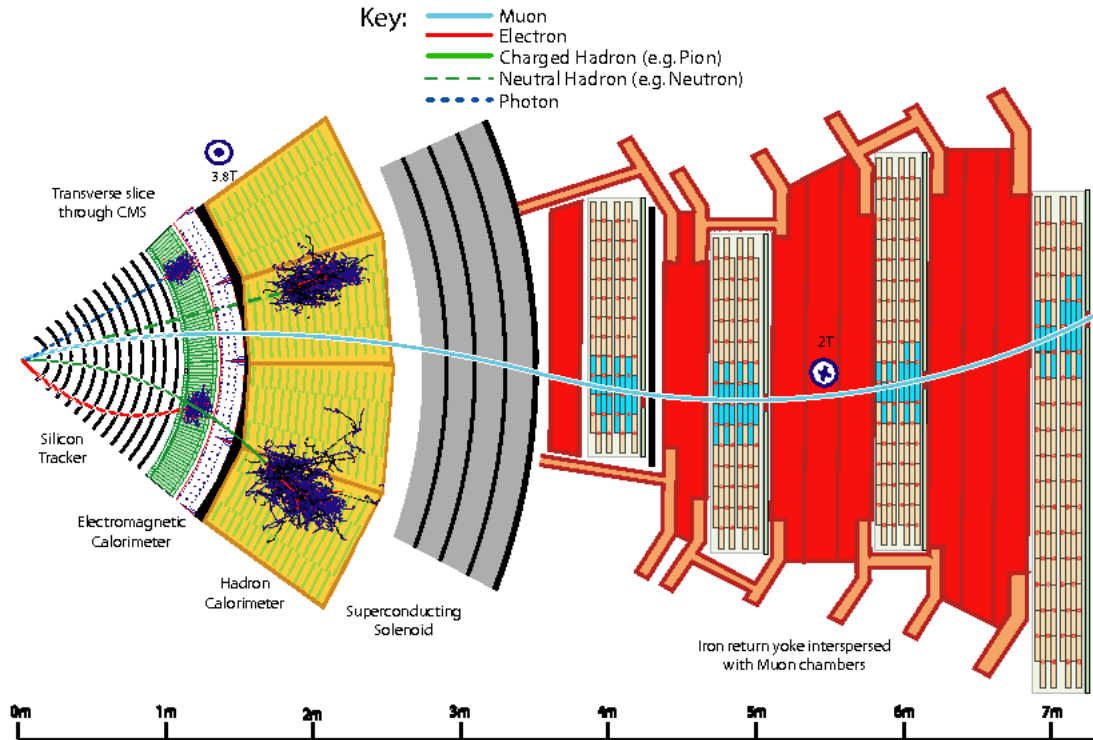


Figure 2.8: Schematic view of a transverse slice of the CMS detector, illustrating the specific signatures of different types of detected particles.

Understanding of the Cartesian coordinate system used in the CMS detector is important before providing a detailed description of the CMS detector and its subdetectors.

2.2.1 The CMS coordinate system and kinematic variables

For measuring physical quantities, the CMS detector uses a cartesian (x, y, z) or cylindrical (r, θ, ϕ) coordinate system with its origin at the nominal interaction point. The z -axis is collinear along the beam and the x -axis points radially toward the center of the LHC ring. Its y -axis is perpendicular to its x - z plane and oriented vertically upward. Figure 2.9 gives an illustration of the CMS coordinate system. The azimuthal angle, ϕ , is measured in the x - y plane with respect to the x -axis. The polar angle θ is defined with respect to the z -axis over the y - z plane such that θ is zero at the positive and π at the negative z -axis. Instead of θ , a more useful angular variable is the pseudorapidity η , which is also a good approximation of the rapidity to avoid energy dependence. Rapidity is defined as follows:

$$y = \frac{1}{2} \log \left(\frac{E + p_z}{E - p_z} \right) \quad (2.8)$$

where E and p_z are the energy and longitudinal component of the momentum of the particle, respectively.

A definition of pseudorapidity is:

$$\eta = -\ln \left(\tan \left(\frac{\theta}{2} \right) \right) \quad (2.9)$$

The central detectors are those detector elements that are located in a geometrical region defined by $|\eta| < 2.5$, whereas forward detectors are those that are located in $|\eta| > 2.5$ regions.

Another often used angular variable is ΔR , which quantifies the separation between two particles in the η - ϕ plane. It is defined as follows:

$$\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2} \quad (2.10)$$

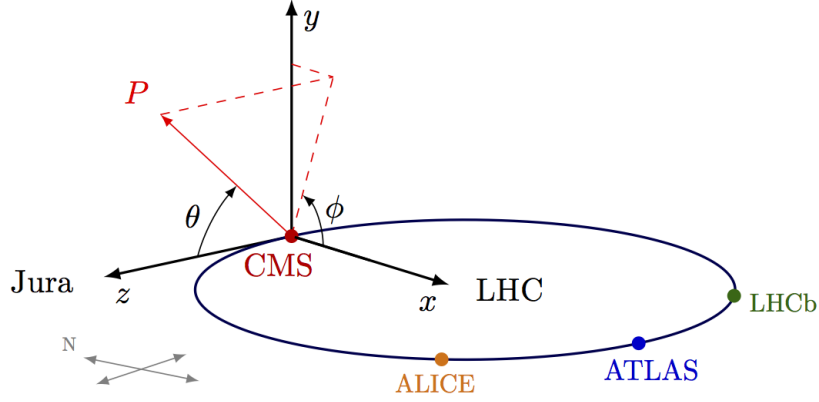


Figure 2.9: Schematic diagram of the coordinate system used in the CMS experiment.

where $\Delta\eta = \eta_1 - \eta_2$ and $\Delta\phi = \phi_1 - \phi_2$.

The initial momenta p_z of the colliding partons is unknown because of the complex partonic structure of hadrons in pp collisions. So, the transverse component of the momentum, denoted by the formula $\vec{p}_T = \vec{p}_x + \vec{p}_y$, is another useful variable. The transverse momentum is very helpful because the p_T of the incoming parton is approximately zero. Momentum conservation is then used to calculate the imbalance in the energy of the outgoing particles. In the next Subsections, each subdetector and its properties will be discussed briefly.

2.2.2 Silicon tracker

The tracking system receives a high amount of radiation because it is the closest subdetector to the IP and beamline of the CMS detector. The tracking system has been designed to sustain and operate in a high flux environment, with thousands of particles passing through it in every bunch crossing. The tracker also allows for precise measurement of secondary vertices resulting from the decay of heavy flavor hadrons, which are produced in several physics channels. Particles produced within $|\eta| < 2.5$ may be tracked by the entire volume of the tracker, which has a length of 5.8 m, a diameter of 2.5 m, and lies within a 3.8 T homogeneous magnetic field.

Figure 2.10 shows a schematic cross section of the CMS tracker. The tracker is divided into two subsystems: pixel and strip tracker.

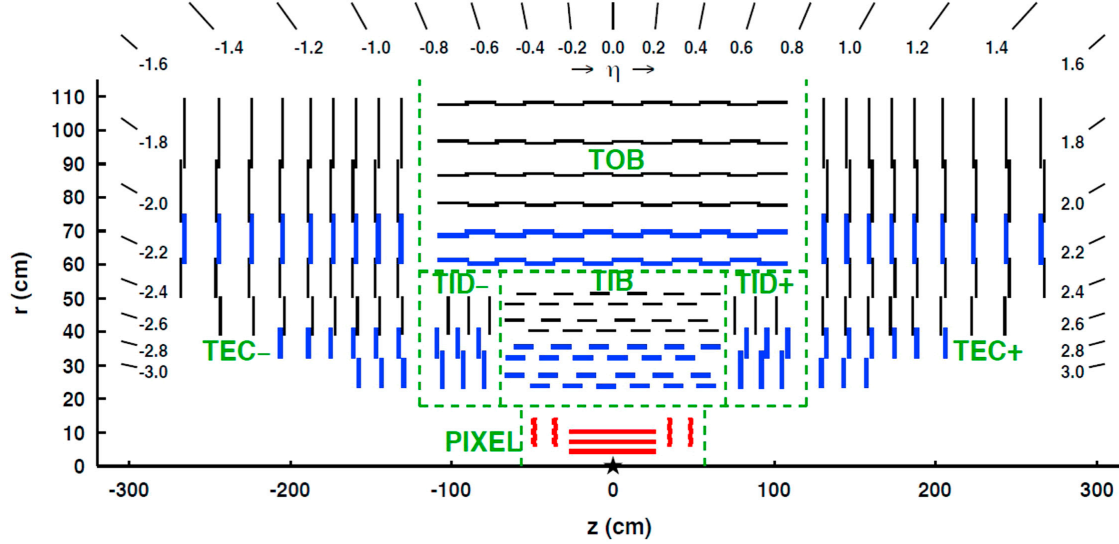


Figure 2.10: A slice through the CMS tracker detector in the r - z plane. The separate subsystem regions have been highlighted and are classified as the Tracker Inner Barrel and Disks (TIB and TIDs), the Tracker Outer Barrel (TOB), and the Tracker Endcaps (TECs). The Pixel subsystem is highlighted in red. Back-to-back stereo strip modules are highlighted in blue [20].

- **Pixel detector:** The pixel detectors are placed closest to the interaction vertex, where the particle flux is highest ($10^7/\text{s}$ at radius (r) ≈ 10 cm). So, in order to achieve the complete track reconstruction and precise vertex determination in a high particle flux environment, a small scale pixel geometry with an almost square pixel shape of $100 \times 150 \mu\text{m}^2$ in both the (r, ϕ) and z coordinates has been adopted. Figure 2.11 depicts the pixel detector arrangement, which includes barrel layers and endcap discs.

The pixel detectors were upgraded during the technical stop at the end of 2016 in order to retain their performance [21, 22]. In 2017, the innermost layer of a barrel pixel (BPix) detector moved closer to the IP, from 4.4 cm to 2.9 cm, and a new fourth outer layer was added at 16 cm. The two forward disks (FPix) are replaced with total 3 disks. So, the CMS Phase-1 pixel detector is made

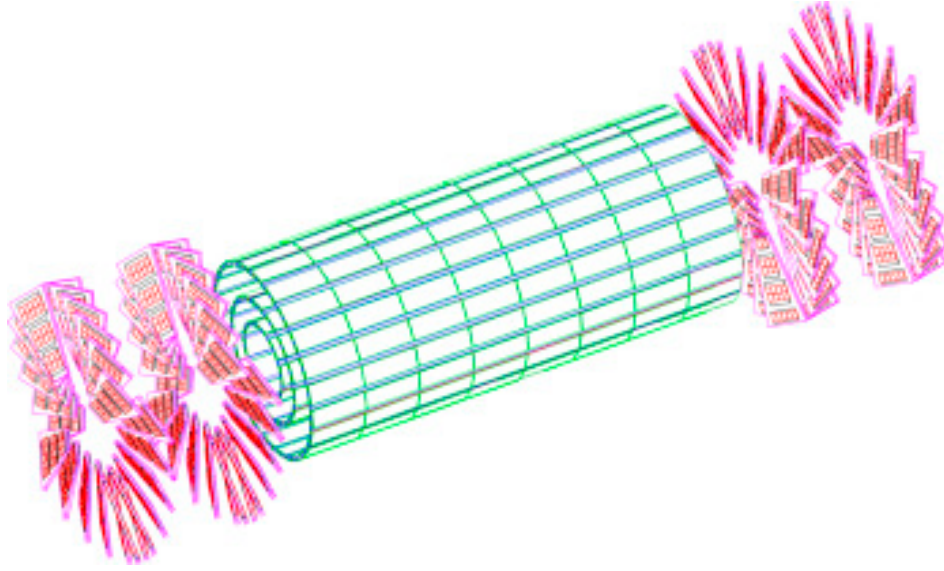


Figure 2.11: The layout of pixel detectors in the CMS tracker. The magenta wedges on the endcap disks are carbon-fiber blades, which hold plaquettes-rectangular arrangements of pixel sensors that come in five different sizes. The rectangles in the barrel layers are the barrel modules- 2×8 rectangular arrangements of pixel sensors, mounted on rectangular carbon-fibre blades.

up of four concentric barrel layers (L1-L4) at radii of 2.9 cm, 6.8 cm, 10.9 cm, and 16.0 cm, and three disks (D1-D3) on each end at distances of 29.1 cm, 39.6 cm, and 51.6 cm from the center of the detector. The Phase-1 design and its differences with respect to Phase-0 are illustrated in Figure 2.12. The CMS Phase-1 pixel detector has a total silicon area of 1.9 m^2 . The CMS Phase-1 pixel detector is built from 1856 segmented silicon sensor modules (1184 BPix and 672 FPix) and spans the pseudorapidity range of $|\eta| < 2.5$. Each module consists of a sensor with 160×416 pixels connected to 16 readout chips (ROCs). In total there are 124 million readout channels. The layout of the CMS Phase-1 Pixel detector has been optimized to have four-hit coverage and improved track reconstruction. The spatial resolution is approximately $13 \mu\text{m}$ for the r - ϕ measurement and about $20 \mu\text{m}$ for the z measurement.

- **Silicon strip detector:** The strip detector, like the pixel detector, has cylindrical elements in the barrel and disks in the endcap and is divided into four

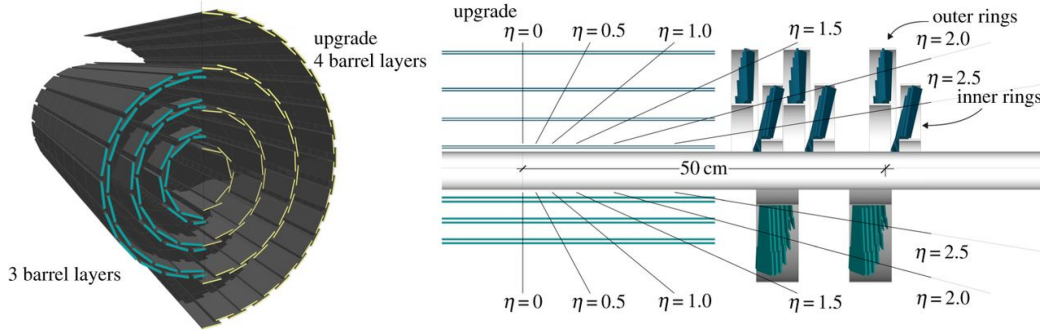


Figure 2.12: The layout of the CMS Phase-1 pixel detector: Pixel barrel layers geometry previously and after the upgrade (yellow area) (left). Pixel longitudinal view comparing Phase-1 in the upper part with Phase-0 in the lower part (right).

sections. The barrel tracker region is divided into two parts: four barrel layers (Tracker Inner Barrel, TIB) and six barrel layers (Tracker Outer Barrel, TOB). The endcaps are divided into the Tracker End Cap (TEC) with nine endcap disks and Tracker Inner Disks (TID) with three endcap layers. The silicon microstrip detectors are placed in the barrel at radii between 20 cm and 110 cm. In the TIB region ($20 \text{ cm} < r < 55 \text{ cm}$), the particle flux is low enough to allow the use of silicon microstrip detectors with a minimum cell size of $10 \text{ cm} \times 80 \mu\text{m}$. In the TOB region ($55 \text{ cm} < r \leq 110 \text{ cm}$), the particle flux has dropped significantly to allow the use of larger-pitch silicon microstrips with a maximum cell size of $25 \text{ cm} \times 180 \mu\text{m}$. Each TEC extends into the region $120 \text{ cm} < |z| < 280 \text{ cm}$ and radially from $r = 20 \text{ cm}$ to $r = 100 \text{ cm}$ to cover the range up to $|\eta| = 2.5$. Each TID is made up of 3 small disks that fill the gap between the TIB and TEC.

In the TIB (TOB), the tracker hit point resolution varies from 23 to 35 (35 to 53) μm in the r - ϕ plane and is 230 (530) μm in the z direction. The transverse momentum resolution is $\approx 1.5\%$ for particles with tracks of $p_T < 10 \text{ GeV}$ and $\approx 2.8\%$ for $p_T = 100 \text{ GeV}$ in the central region of the CMS experiment [20]. During Run1, they were operated at $+4^\circ\text{C}$ using a fluorocarbon cooling sys-

tem, while since 2017 (Run2), they have been operated at a lower temperature (-20° C) to mitigate long-term damages. Tracks reconstructed in pixel and strip detectors are used for vertex reconstruction, particle identification, charge identification, and momenta measurements of charged particles. The reconstruction of tracks and vertices is discussed in Chapter 3.

2.2.3 Electromagnetic calorimeter

The Electromagnetic Calorimeter (ECAL) is a hermetic, homogeneous calorimeter that is located within the solenoid and hadron calorimeter [23]. The ECAL purpose is to measure the energy of electromagnetically interacting particles (mainly photons and electrons) by slowing and finally stopping them. The incident particles interact with the material in the calorimeter, producing showers of secondary particles at lower energies. This process remains continues until the energy of the incident particle falls below the critical energy and the material can no longer be ionized or excited. The ECAL is made up of 61200 lead tungstate (PbWO_4) scintillating crystals mounted in the barrel part, closed by 7324 crystals in each of the 2 endcaps. The reason to choose PbWO_4 material is due to its high density (8.28 g/cm^3), small Molière length of 0.89 cm , short radiation length of 2.2 cm , a fast scintillation response (80% of the light is emitted within 25 ns), and radiation hardness (up to $10 \text{ milliradian (mrad)}$). The total thickness of ECAL is greater than 25 radiation lengths (X_0). The radiation length of a material is the mean length to reduce the energy of a particle by the factor $1/e$. Figures 2.13 and 2.14 represent the layout of ECAL. The ECAL is sub-divided into the following three regions:

- **ECAL barrel:** The ECAL barrel (EB) section has an inner radius of 129 cm and an outer radius is 177 cm . The EB covers the $|\eta| < 1.479$ and has a volume 8.14 m^3 . It consists of 61200 tapered (23 cm long) crystals with a $(2.2 \times 2.2) \text{ cm}^2$ front face ($25.8 X_0$) and a $(2.6 \times 2.6) \text{ cm}^2$ back face. The EB

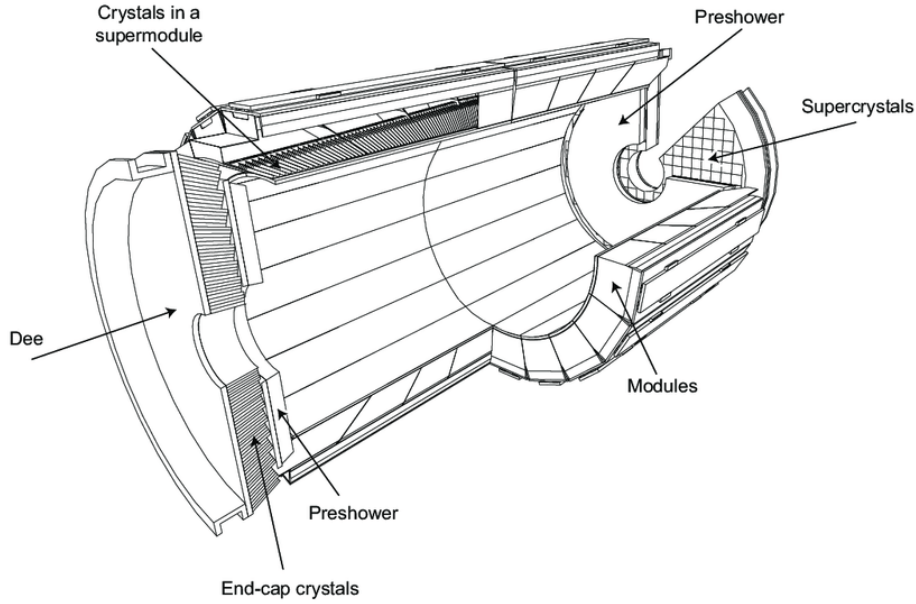


Figure 2.13: The ECAL geometrical view and mechanical layout. For the barrel, crystals are organized into modules and super-modules, and the endcap has two half-disks, or dees. Almost the entire endcap surface is covered by the preshower detector.

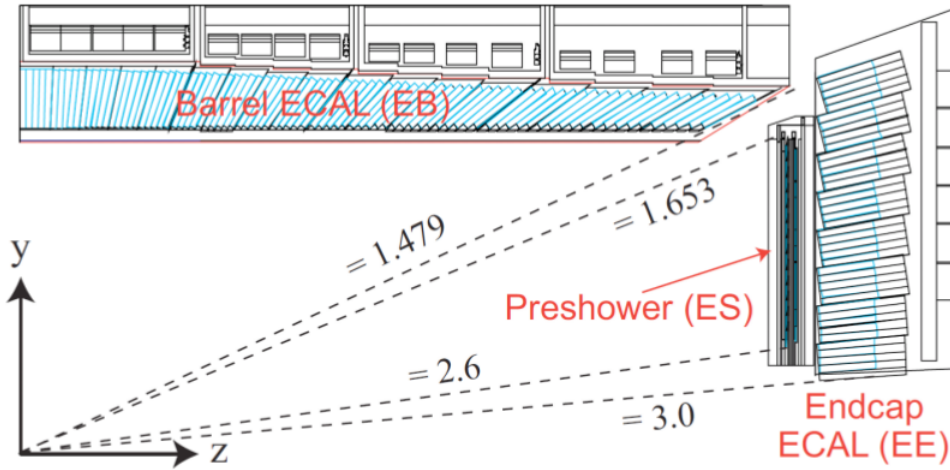


Figure 2.14: Longitudinal view of an ECAL quadrant. The pseudorapidity coverage of the barrel, endcap, and preshower systems are indicated.

crystals are arranged into modules, with four modules ($85 \times 20 = 1700$ crystals) packed into one supermodule. The granularity of the crystal in the η - ϕ plane is 0.0174×0.0174 . The Avalanche PhotoDiodes (APD) are mounted in pairs on the back face of each crystal and measure the scintillating light produced

in the EB. The total amount of scintillating light produced is proportional to the energy that electrons and photons deposit in the crystals.

- ECAL endcap:** The ECAL endcaps (EE) are located at a longitudinal distance of $|z|$ at 314 cm to 390 cm from the IP and cover a pseudorapidity range of $1.479 < |\eta| < 3.0$. Each EE ring is divided into two halves, known as “dees”, each consisting of 3662 PbWO_4 crystals. The 7324 crystals of an EE disk correspond to a thickness of $24.7 X_0$. In order to create supercrystals (SCs), the crystals measuring 22 cm in length are arranged in groups of 5×5 , and each dee has 156 SCs within it. The Vacuum PhotoTriodes (VPT) photomultiplier is mounted to the back face of each EE crystal and measures scintillating photons. The angles at which the crystals are oriented with respect to the z-axis ranging from 2 to 8 degrees.
- ECAL preshower:** The ECAL preshower (PS) detector is situated between EB and EE and has a thickness of 20 cm, an inner radius of 45.7 cm, and an outer radius of 123 cm. Its purpose is to discriminate between prompt photons and photons produced from the decay of neutral pions ($\pi^0 \rightarrow \gamma\gamma$) in the pseudorapidity range of $1.65 < |\eta| < 2.6$. Additionally, it is also used to improve electrons and photons position resolution in the ECAL, as well as to distinguish electrons from the background of minimum ionizing particles (MIP). It is composed of two layers of lead absorbers with a thickness of $2X_0$ and $1X_0$ interleaved with the two planes of silicon strip sensors. The lead planes are divided into two “dees”. The overall thickness of preshower silicon detector is 20 cm ($3X_0$). The silicon strips are used to measure the amount of energy deposited in the absorbers as well as the transverse profile of the shower.

ECAL resolution: Several factors influence the energy resolution of the ECAL, including fluctuation in lateral shower development, the energy released in

the preshower detector, crystal non-uniformities, calibration errors, and electrical noise in the photodetectors. The relative energy resolution of the ECAL [24] is parameterized as a function of measured energy E via:

$$\left(\frac{\sigma(E)}{E}\right)^2 = \left(\frac{0.028}{\sqrt{E}}\right)^2 + \left(\frac{0.12}{E}\right)^2 + (0.003) \quad (2.11)$$

The energy in this case is measured in GeV units, and the result is in good agreement with the design-goal performance of a completely calibrated calorimeter. The transition region ($1.4442 < |\eta| < 1.566$) between the EB and the EE results in smaller reconstruction efficiencies and is not considered in analyses requiring good electron and/or photon reconstruction.

2.2.4 Hadron calorimeter

The Hadron Calorimeter (HCAL) is a sampling calorimeter which completely surrounds the ECAL and covers the pseudorapidity range up to $|\eta| < 5.2$. The HCAL is not entirely immersed in the solenoid magnet coil due to space constraints between ECAL and solenoid [25]. The HCAL is made up of plastic scintillator tiles (fibers) and alternating layers of the brass absorber. An important requirement of HCAL is to minimize the non-Gaussian tails in the energy resolution and to provide good containment and hermeticity for the hadronic jet measurements. The HCAL can also be used to detect neutrinos indirectly using momentum conservation in the transverse plane, i.e., missing transverse energy (MET). Brass has been chosen as absorber material because it has a reasonably short interaction length (16.24 cm) and is non-magnetic. The interaction length (λ_I) of a material is the average distance travelled by a hadron before an inelastic nuclear interaction. The thickness of the HCAL varies between $10\text{--}15\lambda_I$ depending on η , and hence provides good hadronic shower containment. When hadrons interact strongly with the brass nuclei, a cascade of new

secondary particles known as the hadronic shower is produced. The total amount of light produced by a shower in the scintillator is collected by wavelength-shifting fibers (WLS) and transmitted to photodetectors for readout. The HCAL is divided into four parts into the different regions as shown in Figure 2.15: the HCAL barrel (HB) and HCAL outer (HO) in the central region and the HCAL endcap (HE) and HCAL forward (HF) in the endcap region.

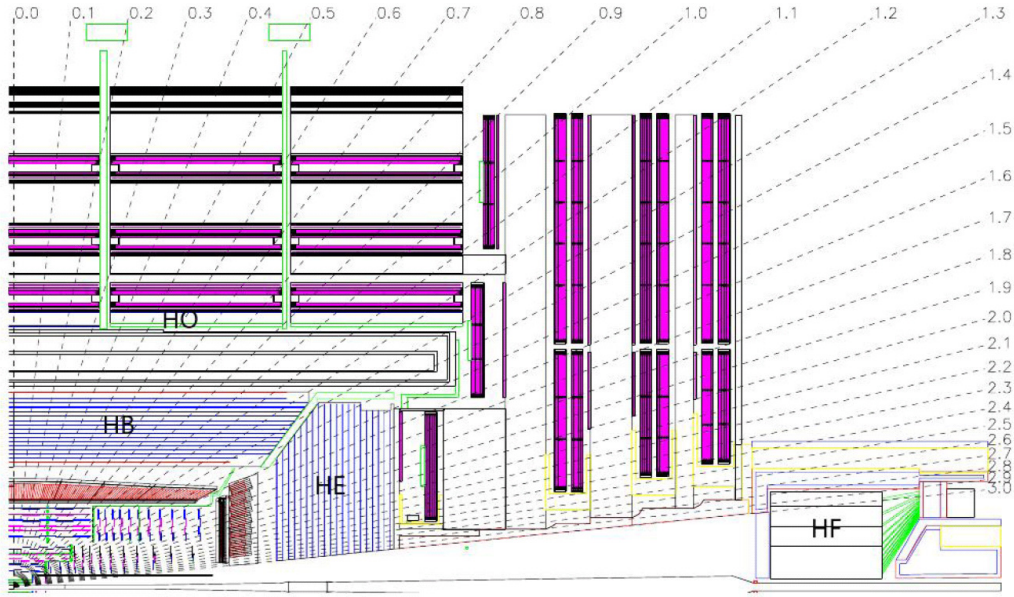


Figure 2.15: Longitudinal view of a CMS HCAL subdetectors.

- Hadron barrel region:** The HB is located at radii of 177.5 cm and 287.65 cm and covers $|\eta| < 1.39$. The HB is divided into two half barrels in the direction along the beam (z). It consists of 70 k scintillator tiles distributed in 2304 towers with a granularity of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$. Each tower in HB is composed of 15 layers of 5 cm thick absorbers brass (an alloy with 70% copper and 30% zinc) and 3.7 mm thick plastic scintillators. The additional innermost and outermost absorber plates are made of stainless steel to strengthen the mechanical structure. The reason for the absorber material to be non-magnetic is that it must not affect the magnetic field. The energy is readout with multi-

channel Hybrid Photo-Detectors (HPDs) that can operate in a high magnetic field and give an amplified response for a large range of particle energies. During LS2, HPDs in HB are replaced by Silicon Photomultipliers (SiPMs) resulting in increased photon detection efficiency and depth segmentation [26].

- Hadron endcap region:** Each HE covers $1.3 < |\eta| < 3.0$ partially overlapping HB. Here brass is used as the absorber material. The innermost surface of HE is located 400.65 cm from the interaction point and is radially divided into 14 rings containing 20 k scintillator tiles distributed in 2304 towers per endcap disk. The granularity begins from $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ at $|\eta| \leq 1.6$ up to $\Delta\eta \times \Delta\phi = 0.17 \times 0.17$ at $1.6 < |\eta| < 3.0$. Starting from Run1 till 2017 data-taking, HPDs have been used as photodetectors in HE. In 2018, the HPDs were replaced by SiPMs, increasing gain from 2×10^2 to 10^6 [26].
- Hadron outer region:** Due to space restrictions inside the magnet and to increase the number of interaction lengths in the barrel region, an additional HO calorimeter is placed on the outer surface of the solenoid. It is used to measure the energy of hadron showers that was left behind after the HB, and therefore HO serves as a “tail-catcher” after the magnet coil. The HO calorimeter covers $|\eta| < 1.26$, increasing the minimal interaction lengths in the barrel region to $11.8 \lambda_I$. The HO uses the steel return yoke and magnet material of CMS as an absorber material. In HO, the HPDs have been replaced by SiPM due to their sensitivity to not well known magnetic field in the return yoke. The scintillating light is collected by WLS and transported to SiPMs. By adding the HO into the HCAL, the hadronic shower energy leakage and accuracy of missing transverse energy measurements are significantly improved [27]. Panjab University in Chandigarh collaborated with the Tata Institute of Fundamental Research in Mumbai developed the HO subdetector. These institutes were responsible for the fabrication, processing, and characterization of the scintillation tiles and wavelength shifting fibers, which were

subsequently successfully installed at the CMS detector in 2004.

- **Hadron forward region:** The HF calorimeter surrounds the beam pipe covering the radial range from 12.5 cm to 157 cm at a distance $|z| = 11.2$ m from the IP, thus spanning the pseudorapidity $2.8 < |\eta| < 5.19$. The HF is placed in order to tolerate high amounts of radiation since the beam collisions cause a significant energy deposit in the forward region. Plastic scintillators are not suitable for such high radiation environments, hence they are made of 5 mm thick steel absorber plates interleaved with 0.6 mm thick layers of radiation resistant quartz fibers, resulting in a thickness of $10\lambda_I$. The incoming particles release energy in the form of Cerenkov light in the quartz fibers, which is then channeled by the fibers to photomultipliers. The fibers are of two different lengths, the longer ones running over the full depth (165 cm) of the detector and the shorter ones starting 143 cm from the front of the detector. This allows to distinguish between the electromagnetic and hadronic showers, as the electromagnetic showers mostly get terminated before reaching the short fibers.

HCAL resolution: Muons and neutrinos produced in the nuclear interactions are not recognized in the HCAL except for a minor muon MIP energy loss. The energy resolution of the HCAL in the HB and HF subdetector can be parameterized as a function of the energy given as:

$$\frac{\sigma(E)}{E} = \frac{0.9}{\sqrt{E}(\text{GeV})} \oplus 0.045 \quad (\text{HB}) \quad (2.12)$$

$$\frac{\sigma(E)}{E} = \frac{1.72}{\sqrt{E}(\text{GeV})} \oplus 0.09 \quad (\text{HF}) \quad (2.13)$$

The HCAL performs worse than the ECAL due to statistical fluctuations in sampling and energy loss through leakage.

2.2.5 The solenoid magnet

The central feature of the CMS detector is a superconducting solenoid magnet of length 12.5 m and 6 m internal radius that is designed to achieve a magnetic flux density of 4 T [28]. However, in order to increase the longevity of the solenoid, the CMS collaboration chooses to maintain the working magnetic field at 3.8 T. The produced magnetic field is approximately 100,000 times stronger than the magnetic field of the earth. The magnet occupies the $2.9 \text{ m} < r < 3.8 \text{ m}$, $|\eta| < 1.5$ region. The magnet is made up of four layers of aluminum-coated NbTi superconducting coil through which $\sim 18 \text{ kA}$ current flows. The entire solenoid is placed within a vacuum tank that has been cooled down to 4.7 K using a liquid helium cooling system. Stronger magnetic fields cause particle trajectories to be more curved, allowing for better momentum measurements of charged particles up to $|\eta| < 2.4$ and preventing low p_T particles ($p_T \leq 0.75 \text{ GeV}$) from reaching the surface of the ECAL, which helps to prevent PU. The magnetic field lines are confined to the volume of the detector through a steel return yoke with a diameter of 14 m and a length of 21.6 m. The yoke weights over 10,000 tonnes. The magnetic field intensity in the return yoke is 1.8 T and is directed in the opposite direction with respect to the magnetic field in the central region.

2.2.6 Muon detector

The muon system is designed outside the solenoid magnet and embedded in the iron return yoke for muon identification, precise momentum measurement, and trigger capabilities [29]. The muon system is a powerful component of CMS for the SM measurements and the discovery of new interesting physics processes over the strong background rate with high luminosity. Since muons have an average lifetime of $2.2 \mu\text{s}$ [30], their decay length is much larger than the size of the CMS detector. Like neutrinos, muon also escapes the tracker and layers of calorimeters without

interacting because muon is the minimum ionizing particle, but loose small energy fractions in the form of minimum ionization loss and gets trapped in the muon system. The muon system has a cylindrical barrel and endcap region as shown in Figure 2.16. Each muon subsystem was constructed using a distinct technology due to differences in particle flux, radiation levels, and residual magnetic field in the various regions covered by the muon detectors. In the barrel and endcaps regions, the neutron induced background, muon rate, and residual magnetic field are low and high, respectively.

Muon system consists of three types of gaseous detectors namely, Drift Tubes (DT) in the barrel part ($|\eta| < 1.2$), Cathode Strip Chamber (CSC) in the endcap ($0.9 < |\eta| < 2.4$), and Resistive Plate Chamber (RPC) in both barrel and endcap regions ($|\eta| < 1.9$). The 1400 muon chambers (250 DTs, 540 CSCs, and 610 RPCs) cover a total surface area of 25000 m^2 and contain about 1 million electronic channels. The details of three muon subsystems are given below.

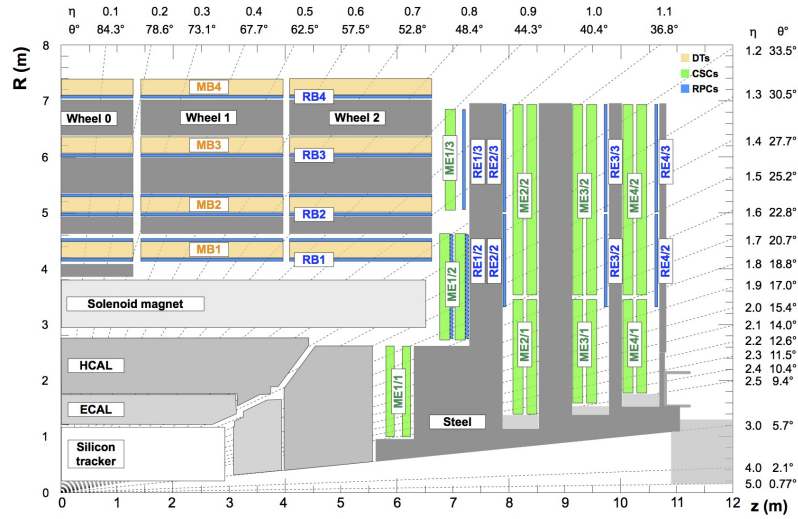


Figure 2.16: Schematic view of CMS muon detector.

- **DTs:** The DTs are organized in 4 layers (four stations named MB1, MB2, MB3, and MB4) that alternate with the segmented return yoke [31]. The segmentation along the beam direction follows the 5 wheels of the yoke (labeled

YB-2, YB-1, YB, YB+1, YB+2 where YB-2 (YB+2) represents the farthest wheel in $-z$ ($+z$). Each of the 5 wheels is divided into 12 sectors, with each covering a 30° azimuthal angle. Each sector comprises four chambers: one chamber inside the magnet return yoke, two chambers located in the gap between the magnet return yoke, and one chamber outside the magnet return yoke. The placement of chambers outside the solenoid magnet and surrounded by a return yoke allow the reduction of the local magnetic field inside the chambers. The chambers in different stations are placed so that a high- p_T muon produced near a sector boundary crosses at least 3 out of the 4 stations.

Each DT chamber consists of 12 layers of drift cells arranged in three groups of four and has a size of about $2\text{ m} \times 2.5\text{ m}$. Each DT cell is made up of a high voltage carrying wire in the center (as the anode) of a 4 cm diameter gas tube, surrounded by two parallel aluminium planes and I-shaped aluminium beams (as the cathodes). The working principle of DT is based on the ionization of charged particles when they pass through the gas, the electrons drift toward the wire and induce a fast signal on the wire. A gas mixture of argon (15%) and CO_2 (85%) is used which is good for quenching and provides a saturated drift velocity. A single cell has an efficiency of about 99.8% and a spatial resolution of $180\text{ }\mu\text{m}$.

- **CSCs:** The CSCs have four stations namely ME1, ME2, ME3, and ME4 [32]. These chambers have trapezoidal shapes and are arranged in a series of concentric rings centered on the beamline. The stations are separated by the iron disks of the flux return yoke, which are thick enough to isolate the electrons in the shower. The ME1 station has three rings of chambers (ME1/1, ME1/2, ME1/3) with increasing radius, while the other three stations are composed of two rings of chambers (ME n /1 and ME n /2). All CSCs except those in the third ring of the first endcap disk (ME1/3) are overlapped in ϕ to avoid gaps in the muon acceptance. There are 36 chambers in each ring of a muon station,

except for the innermost rings of the second, third, and fourth disks (ME2/1, ME3/1, and ME4/1) where there are 18 chambers.

The CSCs are multi-wire proportional counters consisting of 6 gas gaps, each gap having a plane of radial cathode strips and a plane of anode wires running almost perpendicularly to the strips. When muons pass through CSC, they ionize the gas atoms, and the electron drifts towards anode wires forming an avalanche of electrons. Positive ions move away from the wire and towards the copper cathode, also inducing a charge pulse in the strips. By measuring the signal on both wires and strips, two-dimensional position coordinates can be recorded for each muon passing the CSC chamber. The six gap structure of a CSC allows for the recording of third position coordinates as well. Hence, each CSC measures up to 6 space coordinates (r, ϕ, z). The gas mixture used for CSC chambers is 40% Ar, 50% CO₂, and 10% CF₄ with a spatial resolution of 80 μm . The CF₄ is included to prevent polymerization on the wires. The signal on the wires is fast and used in the Level-1 (L1) trigger. The CSCs have a fast response time and more radiation-resistant than the DTs. The spatial resolution provided by each chamber from the strips is typically about 200 μm (100 μm for ME1/1). The angular resolution in ϕ is of order 10 mrad. The timing resolution per layer is approximately 5 ns.

- **RPCs:** The RPCs mainly contribute to the muon trigger and offer a redundant system covering both the barrel (together with DTs) and endcap (along with CSCs) [33]. Five wheels are placed in the barrel region and eight disks (4 for the positive side and 4 for the negative side from the IP) are placed in the endcap region. There are 480 RPC chambers in the barrel region placed parallel to the wires of the DT chambers and 36 RPC chambers parallel to the CSC strips in the endcap. A resistive plate chamber is made up of two parallel plates of bakelite planes with good surface flatness and high bulk resistivity to prevent continuous breakdown. The separation between plates is

2 mm which acts as a gas gap. In CMS, the double gas gap geometry is used, which means that there are two gas gaps located on both sides of the readout anode. This choice of detector geometry allows us to measure double signal size per incident particle and also results in a reduction in voltage and hence in breakdown probability without any loss of signal quality. The bakelite is coated with graphite paint to form electrodes and readout is made by means of aluminum strips outside the bakelite plates, insulated from the electrodes. The RPC gas mixture contains 95.2% $\text{C}_2\text{H}_2\text{F}_4$ (tetrafluoroethane - Freon), 4.5% $\text{i-C}_4\text{H}_{10}$ (isobutane), and 0.3% SF_6 (sulphur hexafluoride). The RPCs work on the principle of production of electron-ion pairs due to the ionization of gas under a strong electric field when a charged particle passed through it. The electron-ion pair thus produced secondary ionization and hence led to an avalanche of electrons. The high bulk resistivity of the electrodes allows the electric charge to die off quickly in a minimal area around a point where the discharge occurs. The signal produced by discharge is collected by the metallic readout strips and is further read and transmitted using readout electronics. RPCs are fast gaseous detectors that provide a muon trigger system with a time resolution better than 3 ns, complementary to those given by the DT and CSC chambers. The transverse momentum resolution for muons in CMS with $20 < p_T < 100 \text{ GeV}$ ranges from 1.3–2.0% in the barrel region and less than 6% in the endcap region.

During the LHC shutdown period, the muon detector is upgraded in order to ensure the performance of the muon triggering and tracking systems during the HL-LHC [34] as shown in Figure 2.17. The CSCs upgrading will concentrate on electronics. The Cathode front-end boards (CFEB) replacement on the inner chambers is required to accommodate the increased L1 trigger latency and the rate at HL-LHC. Since DTs have a low radiation tolerance so improvements in electronics have increased trigger rate capability, performance, and maintainability. During

LS1, **RPCs** are installed in the (RE-4) RPC endcap station in the η region up to $|\eta|=1.8$. The key objective of the muon trigger is to give a high enough resolution to identify high p_T tracks. Also, it provides finer timing and resolution to the corresponding CSC station while maintaining a low p_T threshold. The present RPC system can be upgraded by installing a few improved RPC (iRPC) detectors in high- η regions. The GE1/1 triple Gas Electron Multiplier (GEM) detectors have been already installed in two forward regions at $1.6 < |\eta| < 2.4$ during LS2, and GE2/1 will be installed during LS3 [35]. Panjab University is one of the production sites for the production of GE1/1 detectors and detail will be discussed in Chapter 4

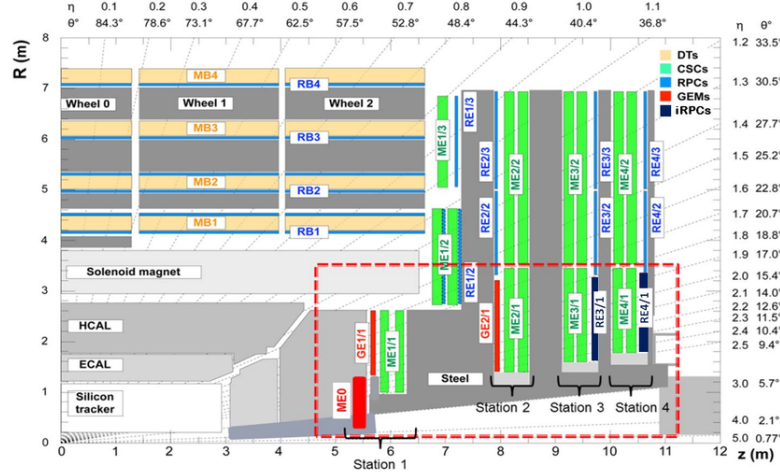


Figure 2.17: A CMS Muon spectrometer quadrant with the DT chambers (yellow), RPC (light blue), and CSC (green). The dashed box contains the positions of the new forward muon detectors for the HL-LHC phase. They are coloured red for the GEM stations (ME0, GE1/1, and GE2/1) and violet for the upgraded RPC stations (RE3/1 and RE4/1).

2.3 The CMS event trigger and data acquisition

The CMS is a multi-purpose detector designed for the precision measurement of leptons, photons, and jets, among other physics objects, in pp and HI collisions at the CERN LHC. The LHC is designed to collide protons at a $\sqrt{s}$ of 14 TeV and a

luminosity of $10^{34} \text{cm}^{-2} \text{s}^{-1}$ in every 25 ns (40 MHz) at each IP. At design luminosity, the pp interaction rate exceeds 1 GHz. Only a small fraction of these collisions contain events of interest to the CMS physics program, and these can be stored for later offline analysis. The role of the trigger system is to select interesting events for offline storage from the bulk of the inelastic collision events. The majority of the events, which are referred to as minimum bias events, are caused by a soft interaction between the protons. The CMS has created a trigger system and a Data Acquisition (DAQ) system to transport the data to Tier-0 (T0) and for further processing in order to remove unnecessary data and to save interesting information [36, 37]. To maintain high efficiency for the research of physics phenomena, the required rejection power of $\mathcal{O}10^5$ is too high to attain in a single processing step. For this reason, the full selection task is split into two steps. The first step (hardware based L1 Trigger) is designed to reduce the rate of events accepted for further processing to less than 100 kHz. The second step (software based High-Level Trigger or HLT) is designed to reduce this maximum L1 accept rate of 100 kHz to a final output rate of 1 kHz. A more detailed overview of the two systems is provided below.

2.3.1 Level-1 trigger

The L1 trigger is a fixed latency hardware system [38]. Within $4 \mu\text{s}$ of a collision, the system has to decide if an event should be tentatively accepted or rejected based on raw information from the calorimeter and muon detectors. A schematic of the L1 trigger is shown in Figure 2.18. The trigger primitives (TP) from the ECAL, HCAL, and muon detectors (DTs, CSCs, and RPCs) are processed in several steps before the combined event information is evaluated in the global trigger (GT) and a decision is made on whether or not to accept the event.

The GT is the final step of the CMS L1 trigger system and implements a menu of triggers, a set of selection requirements applied to the final list of objects (i.e.

electrons/photons, muons, jets, or τ leptons) required by the HLT algorithms to meet the physics data-taking objectives. This menu includes trigger criteria ranging from simple single object selections with E_T above a predefined threshold to selections requiring coincidences of several objects with topological conditions among them. A maximum of 128 separate selections can be implemented in a menu.

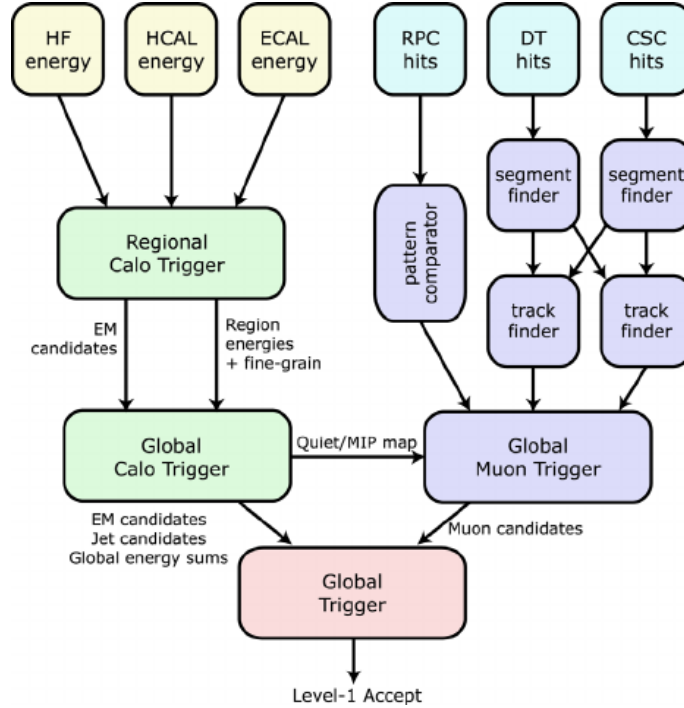


Figure 2.18: Flow diagram of the L1 trigger system of CMS.

2.3.2 High level trigger

The HLT employs cutting-edge software and a processor farm made up of a number of multi-core processors in order to lower the event rate from 100 kHz to 1 kHz with a processing delay time of around 100 ms [39]. The event selection method at the HLT is similar to that used in offline processing. For each event, objects such as electrons, muons, and jets are reconstructed and identification criteria are applied in order to select only those events which are of possible interest for data analysis. The HLT hardware consists of a single processor farm composed of commodity computers, the

event filter farm (EVF), which runs Scientific Linux. In the EVF, individual event fragments from the detector are assembled to form complete events and unpack this assembled raw data into detector specific data structures and perform the event reconstruction and trigger filtering. The combined set of L1 and HLT selection criteria for a given signature is called the trigger path. An event is selected if any of the trigger paths accept it. Upon completion, accepted events are then sent to another software process, called the storage manager, for archival storage. The event data is stored locally on disk and eventually transferred to the CMS T0 computing center for offline processing and permanent storage. The T0 center accepts RAW data, stores it on tape, and perform the prompt reconstruction on RAW data to produce reconstructed (RECO), Analysis Object Data (AOD), and miniAOD datasets. The T0 center acts as a buffer between CMS and other computing centers located across the world. Using graphs and histograms that are specifically made to spot issues, Data Quality Monitoring (DQM) verifies the accuracy of recorded data. Depending on this, datasets can then be classified as good or bad in the database and communicated to the subdetector experts. A list of datasets authorized to be used for physics analysis is generated at the conclusion of the DQM chain. From the T0 center, the data sets are replicated to a set of seven remote Tier-1 (T1) centers, where the second copy of RAW data is stored. The T1 centers also reconstruct the RAW data again (re-reconstruction) after all the calibrations. Finally, a large number of smaller Tier-2 (T2) centers hosted by various CMS member groups provide computing services and infrastructure for physics analysis users. Figure 2.19 shows the CMS data flow from the HLT to the datasets used for physics analysis.

In a brief, this Chapter provides an overview of the LHC machine as well as the CMS detector. The construction and operation of each CMS subdetector have been discussed briefly. In order to select the events that need to be kept for physics analysis, the CMS detector employs a two-level trigger. The various techniques that were used for the particle reconstructions and event flow descriptions will be

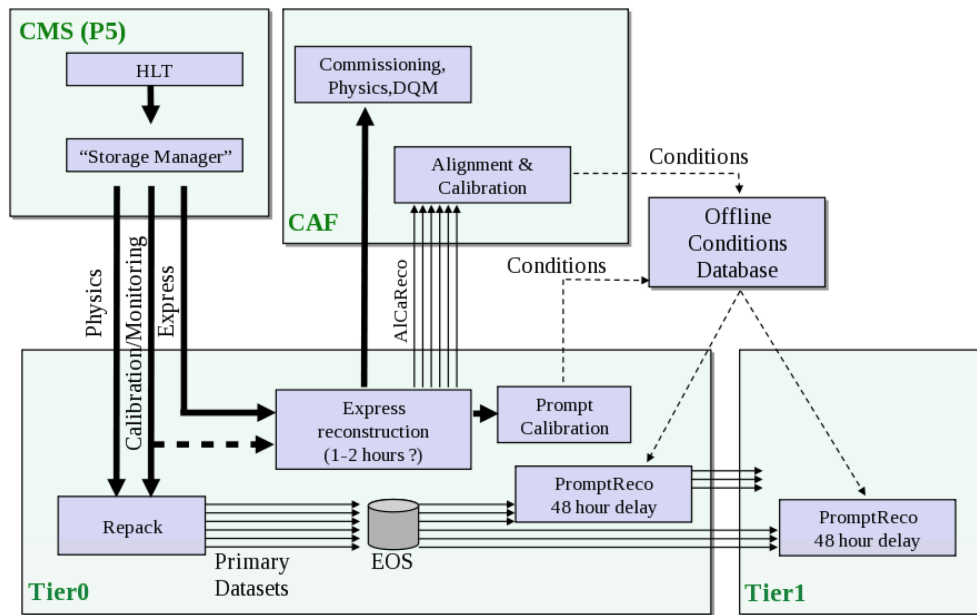


Figure 2.19: CMS data flow from HLT to physics datasets.

discussed in the next Chapter.

Bibliography

- [1] Lyndon Evans and Philip Bryant, “LHC Machine”, *JINST*, vol. 3 (2008) S08001, [doi:10.1088/1748-0221/3/08/s08001](https://doi.org/10.1088/1748-0221/3/08/s08001).
- [2] CMS Collaboration, “Evidence for the direct decay of the 125 GeV Higgs boson to fermions”, *Nature Phys*, vol. 10 (2014) 557, [doi:10.1038/nphys3005](https://doi.org/10.1038/nphys3005).
- [3] CMS Collaboration, “Observation of a new boson at amass of 125 GeV with the CMS experiment at the LHC”, *Phys. Lett. B*, vol. 716 (2012) 30, [doi:10.1016/j.physletb.2012.08.021](https://doi.org/10.1016/j.physletb.2012.08.021).
- [4] ATLAS Collaboration, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, *Phys. Lett. B*, vol. 716 (2012) 1, [doi:10.1016/j.physletb.2012.08.020](https://doi.org/10.1016/j.physletb.2012.08.020).
- [5] The HL-LHC Project, <https://hilumilhc.web.cern.ch/content/hl-lhc-project>.
- [6] <https://cds.cern.ch/record/841573/files/lhc-pho-1997-060.jpg>.
- [7] ATLAS Collaboration, G.Aad *et al.*, “The ATLAS Experiment at the CERN Large Hadron Collider”, *JINST*, vol. 3 (2008) S08003, [doi:10.1088/1748-0221/3/08/s08003](https://doi.org/10.1088/1748-0221/3/08/s08003).
- [8] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST*, vol. 3 (2008) S08004, [doi:10.1088/1748-0221/3/08/s08004](https://doi.org/10.1088/1748-0221/3/08/s08004).

-
- [9] ALICE Collaboration, “The ALICE experiment at the CERN LHC”, *JINST*, vol. 3 (2008) S08002, [doi:10.1088/1748-0221/3/08/s08002](https://doi.org/10.1088/1748-0221/3/08/s08002).
- [10] R. S. Bhalerao and R. V. Gavai, “Heavy Ions at LHC: A Quest for Quark-Gluon Plasma”, (2009) 105, [doi:10.1007/978-81-8489-295-6_8](https://doi.org/10.1007/978-81-8489-295-6_8).
- [11] LHCb Collaboration, “The LHCb Detector at the LHC”, *JINST*, vol. 3 (2008) S08005, [doi:10.1088/1748-0221/3/08/s08005](https://doi.org/10.1088/1748-0221/3/08/s08005).
- [12] LHCf Collaboration, “The LHCf detector at the CERN Large Hadron Collider,” *JINST*, vol. 3 (2008) S08006. [doi:10.1088/1748-0221/3/08/S08006](https://doi.org/10.1088/1748-0221/3/08/S08006).
- [13] TOTEM Collaboration, “The TOTEM experiment at the CERN Large Hadron Collider,” *JINST*, vol. 3 (2008) S08007. [doi:10.1088/1748-0221/3/08/S08007](https://doi.org/10.1088/1748-0221/3/08/S08007).
- [14] MoEDAL Collaboration, “Technical Design Report of the MoEDAL Experiment”, MoEDAL Technical Design Report, CERN-LHCC-2009-006, CERN, Geneva, 2009. <https://cds.cern.ch/record/1181486>.
- [15] J. L. Pinfold, “The MoEDAL experiment - a new light on LHC physics”, *EPJ Web Conf.*, vol. 95 (2015) 03030, [doi:10.1051/epjconf/20159503030](https://doi.org/10.1051/epjconf/20159503030).
- [16] “Public CMS Luminosity Information”, <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>.
- [17] CMS Collaboration, “Technical Proposal for the Phase-II Upgrade of the CMS Detector,” CMS Technical Design Report, CERN-LHCC-2015-010, CERN Geneva, (2015). <https://cds.cern.ch/record/2020886?ln=en>.
- [18] E. Halkiadakis. “Physics of the large and the small: Introduction to the lhc experiments” in *Proceedings for tasi 2009 summer school 2010*, [doi:10.48550/arXiv.1004.5564](https://doi.org/10.48550/arXiv.1004.5564).

-
- [19] CMS Collaboration, “Highlights from the compact muon solenoid (CMS) experiment”, *Universe*, vol. 5 (2019) 28, [doi:10.3390/universe5010028](https://doi.org/10.3390/universe5010028).
- [20] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker”, *JINST*, vol. 9 (2014), [doi:10.1088/1748-0221/9/10/p10009](https://doi.org/10.1088/1748-0221/9/10/p10009), [arXiv:1405.6569v2](https://arxiv.org/abs/1405.6569v2).
- [21] CMS Collaboration, “CMS Technical Design Report for the Pixel Detector Upgrade”, CMS Technical Design Report, CERN-LHCC-97-033, CERN, Geneva (2012). <https://cds.cern.ch/record/1481838?ln=en>.
- [22] CMS Collaboration, “Upgrade of the CMS tracker”, *JINST*, vol. 9 (2014) C03041, [doi:10.1088/1748-0221/9/03/C03041](https://doi.org/10.1088/1748-0221/9/03/C03041).
- [23] CMS Collaboration, “The CMS electromagnetic calorimeter project: Technical Design Report”, CMS Technical Design Report, CERN-LHCC-97-033, CERN, Geneva (1997). <https://cds.cern.ch/record/349375?ln=en>.
- [24] The CMS Collaboration, “Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp Collisions at $\sqrt{s} = 7$ TeV”, *JINST*, vol. 8 (2013) P09009, <http://dx.doi.org/10.1088/1748-0221/8/09/P09009>.
- [25] CMS Collaboration, “The CMS hadron calorimeter project: Technical Design Report”, CMS Technical Design Report, CERN-LHCC-97-031, CERN, Geneva (1997). <https://cds.cern.ch/record/357153?ln=en>.
- [26] CMS Collaboration, “Phase 1 Upgrade of the Hadron Calorimeter”, CMS Technical Design Report, vol. 273 (2012) 1002, [doi:10.1016/j.nuclphysbps.2015.09.157](https://doi.org/10.1016/j.nuclphysbps.2015.09.157).
- [27] S. Banerjee, “Performance of Hadron Calorimeter with and without HO”, CMS-NOTE-1999-063, CERN, Geneva, (1999). <http://cds.cern.ch/record/687178>.

- [28] CMS Collaboration, “The CMS Magnet Project: Technical Design Report”, CMS Technical Design Report, CERN, Geneva (1997). <https://cds.cern.ch/record/331056>.
- [29] CMS Collaboration, “The CMS muon project: Technical Design Report”, CMS Technical Design Report, CERN, Geneva (1997). <https://cds.cern.ch/record/343814>.
- [30] P. D. Group, “Review of Particle Physics,” *PTEP*, vol. 2020 (2020) 083C01.
- [31] CMS Collaboration, “Performance of the CMS drift tube chambers with cosmic rays”, *JINST*, vol. 5 (2010) T03015 [doi:10.1088/1748-0221/5/03/t03015](https://doi.org/10.1088/1748-0221/5/03/t03015).
- [32] CMS Collaboration, “Performance of the CMS cathode strip chambers with cosmic rays”, *JINST*, vol. 5 (2010) T03018 [doi:10.1088/1748-0221/5/03/t03018](https://doi.org/10.1088/1748-0221/5/03/t03018).
- [33] G. Pugliese, The CMS Muon Collaboration, “The RPC system for the CMS experiment”, *2006 IEEE NSS Conference Record N24-3*, (2007).
- [34] CMS Collaboration, “The Phase-2 Upgrade of the CMS Muon Detectors”, CMS Technical Design Report, CERN-LHCC-2017-012, CERN, Geneva (2017). <http://cds.cern.ch/record/2283189>.
- [35] CMS Collaboration, “CMS Technical Design Report for the Muon Endcap GEM Upgrade”, CMS Technical Design Report, CERN-LHCC-2015-012, CERN, Geneva (2015). <https://cds.cern.ch/record/2021453>.
- [36] Vardan Khachatryan *et al.*, “The CMS trigger system”, *JINST* 12 (2017) P01020 [doi:10.1088/1748-0221/12/01/p01020](https://doi.org/10.1088/1748-0221/12/01/p01020).

- [37] Andrea Perrotta, “Performance of the CMS High Level Trigger”, *Journal of Physics: Conference Series 664*, vol. 664 (2015) 082044, [doi:10.1088/1742-6596/664/8/082044](https://doi.org/10.1088/1742-6596/664/8/082044).
- [38] CMS Trigger and Data Acquisition Group, “The CMS high level trigger”, *Eur. Phys. J.C*, vol. 46 (2006) 605, [doi:10.1140/epjc/s2006-02495-8](https://doi.org/10.1140/epjc/s2006-02495-8).
- [39] CMS Collaboration, “CMS The TriDAS Project:Technical Design Report”. CMS Technical Design Report, CERN-LHCC-2000-038, CERN, Geneva (2000). <https://cds.cern.ch/record/706847>.

Chapter 3

Generation, Simulation, and Reconstruction of the Events

This Chapter describes the events generation, detector simulation, and reconstruction process of physics objects. At the LHC, extracting physics from the data will present several significant challenges. First, pp collisions at very high energies produce final states that involve a large number of particles. Second, even in the presence of a clear anomaly with respect to the SM prediction, its interpretation in terms of underlying phenomena or theory could be extremely difficult. Before one or a few candidate theories can be selected, accurate measurements of the corresponding parameters (like masses, couplings, spin, and charges) will be needed. Production rates and branching ratio measurements, for example, will provide constraints only if we are able to connect them to the fundamental parameters of a model through an accurate calculation, at least at NLO in perturbative QCD. In this context, there is no doubt that MC simulations play a key role at each stage of the exploration of the TeV scale, i.e., from the discovery and identification of BSM physics to the measurement of its properties. In High Energy Physics (HEP) experiments, a large number of variables are involved hence the MC methods make use of random numbers to solve difficult physical or mathematical systems [1-3].

Figure 3.1 shows a schematic view of the workflow of the simulation program. The complete chain of pp collision events is divided into the following steps:

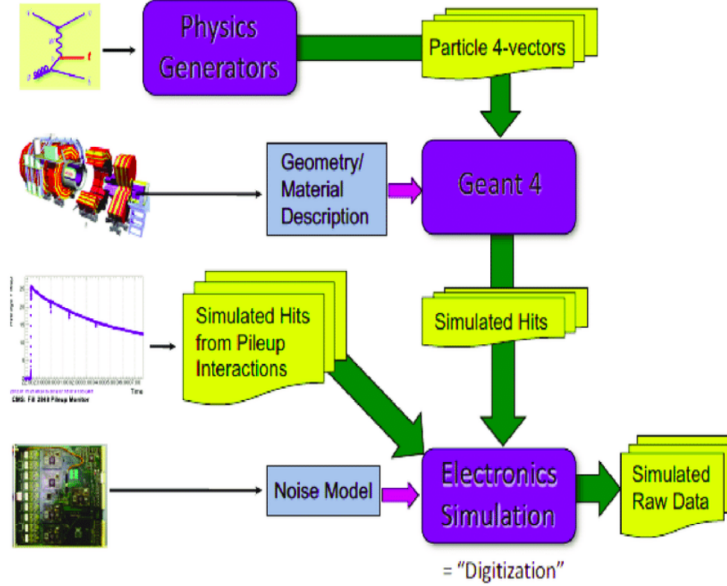


Figure 3.1: CMS simulation workflows [4].

- **Event generation:** The generation of pp collision events is the preliminary step of the simulation process. So various MC event generators based on theoretical models are used to generate events from particle interaction to the final products of the collision. The output of the event generator is the MC truth information (vertices and particles four-vectors).
- **Detector simulation:** In the next step, generated events are required to pass through the virtual detector which is based on the GEANT4 (**GE**ometry **ANd** **T**racking) [5] toolkit. It includes full detector geometry simulation along with complete evolution of the event through space and time, along with the interaction of particles within the detector volume and the effect of the magnetic field. This step produces persistent hits in the sensitive detectors.
- **Digitization:** This step includes the simulation of readout electronics, including simulated hits from the PU integration detector noise model. This step

converts hits into digitizations (also known as “digis” or “raw data”) which correspond to the output of the CMS electronics.

- **Event reconstruction:** Reconstruction is the operation of constructing physics quantities for example electrons, muons, jets, etc. from the raw data collected in the experiment. The same reconstruction software chain is applied to both simulated and experimental data.

3.1 CMSSW framework

The CMS SoftWare (CMSSW) is a collection of software, that is built around an Event Data Model (EDM) framework, that processes event data for the simulation, calibration and alignment, and reconstruction modules [6]. The CMS EDM is centered around the concept of an Event. An Event is a C++ object container and configured via Python [7] for all raw and reconstructed data related to a particular collision. During processing, data is passed from one module to the next via the Event and is accessed only through the Event. Finally, all objects in the Event may be individually or collectively stored in ROOT files and accessed using the ROOT data analysis framework [8].

3.2 Event generation

Over the past decades, the description of multi-purpose MC event generators of the final states obtained in HEP experiments has been improved [9]. These programs can be run within the CMSSW framework using delicate interface libraries. Event generators divide the whole event into different components which are mentioned below.

- **Hard scattering processes:** The interesting physics process takes place through high momentum exchange between constituents of two colliding pro-

tons, which results in the production of very energetic particles called hard processes. The hard scattering component of the event consists of outgoing hard particles, along with hard partons from ISR and FSR that are simulated using a showering algorithm. The QCD factorization theorem [10] is used to calculate the cross section of a hadron collision. This theorem separated the hadronic collisions into high energy or short distance perturbative calculations of the hard scatter from low energy or long distance non-perturbative calculations. The hard scatter calculation is performed by evaluating the Feynman diagrams which represent the process of interest to yield the matrix element (ME). The ME describes the complex amplitude of the interaction. The hard scattering cross section is proportional to the squared sum of the MEs ($\sigma \propto |M^2|$). This part is known as the fixed-order (FO) or ME component. While the long distance effects that concern the proton structure are accounted by the PDFs [11, 12].

MEs can be calculated at different orders in QCD. The LO calculation takes only MEs at the lowest order and corresponding LO (tree-level or Born-level) Feynman diagrams. In the NLO and NNLO calculations, the hard process contains more Feynman diagrams or ME expands in the power of α_s to include high order corrections or extra radiation and final states with multiple partons. A cross section calculated at NLO accuracy is composed of three parts, the LO or Born-level part, and two corrections, the virtual and real emission. The real emission diagrams contain either Feynman diagrams with ISR or with FSR. In LO calculations, the absolute normalization is often badly described, due to missing large higher order corrections. One therefore often introduces a so called K-factor when comparing results from event generators with experimental data. This factor is normally a single factor multiplying to the LO cross section, typically obtained by taking the ratio of the total NLO or NNLO cross section to the LO one for the relevant process. The content of

the proton can be described in terms of PDFs, which specify the fraction of the proton longitudinal momentum (x) carried by a quasi-free constituent quark or gluon or parton, at a squared energy scale. Thus, the total cross section for any process in hadron collisions is evaluated by convoluting the partonic cross section with corresponding PDFs as given below:

$$\sigma_{pp \rightarrow X} = \int_0^1 dx_1 \int_0^1 dx_2 \sum_{i,j} f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) [\hat{\sigma}_{LO} + \alpha_s(\mu_R^2) \hat{\sigma}_{NLO} + \dots]_{ij \rightarrow X} \quad (3.1)$$

In this equation, the left hand side is the cross section of the production of the final state X in pp collisions and $\hat{\sigma}_{ij \rightarrow X}$ is the parton-level cross section for producing X when the partons i and j collide and f is the PDF. The μ_F is the factorization scale that separates the long distance and short distance phenomena. A parton with a transverse momentum less than μ_F is considered to be the part of the proton structure and it is absorbed in the PDF. Partons above this scale participate in the hard scattering with a cross section $\hat{\sigma}$. The choice of the scale μ_F depends on the process. The ME provides a good description of only wide angle and high energy particles because at low energy the α_s coupling diverges. Large logarithmic corrections from gluons emitted collinear (small angle) with the incoming quarks in the hard scatter are absorbed into the renormalized (μ_R) PDFs. So, μ_F and μ_R are introduced to remove divergences at low and high energies and are known as ultraviolet and infrared divergences, respectively. The dependence of the cross section on the factorization and renormalization scale can be described in pQCD via the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP) evolution equations [13–15]. These evolution equations allow us to determine PDFs at an energy scale of interest from the available measurements performed at a different energy scale. The PDF sets are determined through global fits on experimentally measured

cross sections values [16]. Usually, for the purpose of calculating the outcome of collisions, the values of the scales are taken to be $\mu_R^2 = \mu_F^2 = Q^2$.

As discussed in Chapter 1, two different approaches 4FS and 5FS can be followed to treat initial state b quarks. In the 4FS the b quark emission is simulated exactly at FO in perturbation theory, while in the 5FS its emission comes from the PS, so, in principle, the former should be more accurate. However, the gluon splitting in the former approach suffers from a collinear divergence that may spoil the convergence of the perturbation series. This divergence is on the other hand absorbed by the PDFs in the 5FS.

Figure 3.2 describes the probability density to find various proton constituents or partons, and the longitudinal momentum fraction (x) for the NNPDF3.0 PDF set at NNLO for two different energy scales: 10^2 GeV and 10^4 GeV [17]. At the lower energy scale, the b quark PDF is zero since the energy scale is below the value of the b quark mass. In contrast, at the higher energy scale, which is above the value of the b quark mass, the PDF of the b quark is non-zero. For about $x > 0.2$, the valence quark dominates with the contribution of u twice as large as of the d, while for about $x < 0.2$ the gluon PDF is the largest. As the energy scale increases, the valence quarks can radiate an increasing number of gluons, which can split into quarks. This distributes some of the proton momentum associated with the up and down quarks to the rest of the partons. Therefore, at high energy hadron colliders like the LHC, the processes initiated by a gluon-gluon interaction (gluon-fusion) dominate. Consequently, as the energy scale increases, the gluon, b, c, and s quarks PDFs increase at low x , while the up and down quark PDFs decrease at high x . Furthermore, since the b quark has a relatively heavy mass, its PDF is relatively small. This is because its PDF has not received contributions from the gluon splittings at the energy scales that are below the value of its mass, while the PDFs of the lighter quarks have received the contributions at those

energy scales. The width in each line represents the uncertainty in the PDF for that parton and it shows that at low Q^2 , the PDF is not known precisely for the small x .

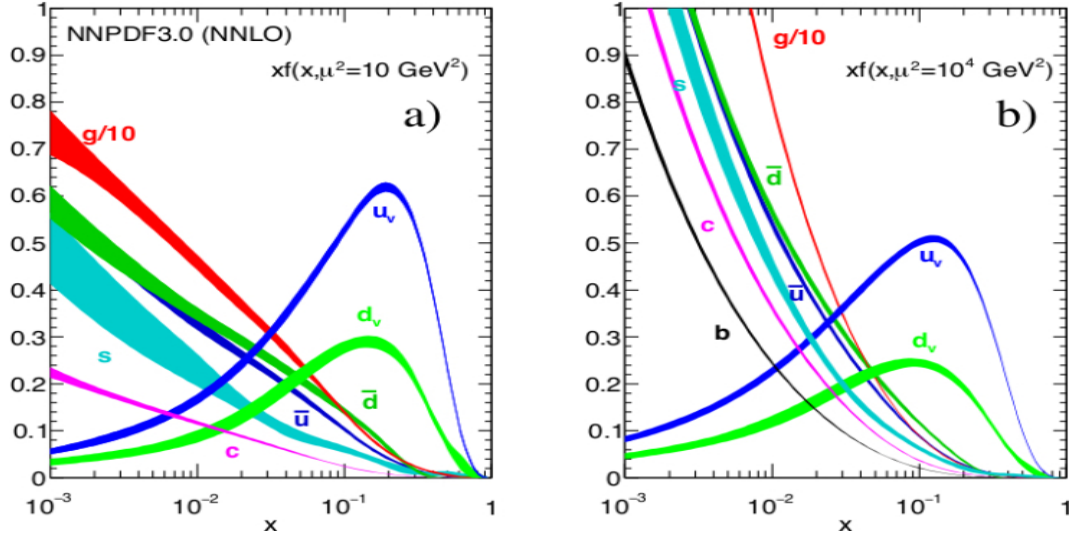


Figure 3.2: Parton distribution functions of the proton from NNPDF3.0 at NNLO as a function of the fraction of the proton momentum at two energy scales: (a) $Q^2 = 10^2 \text{ GeV}^2$ and (b) $Q^2 = 10^4 \text{ GeV}^2$. Each curve represents $xf(x, \mu^2)$. For display purposes, the gluon PDF is divided by 10 and the PDFs are weighted by the momentum fraction. Due to this weight, the graphs represent the momentum density associated to a parton type instead of its number density.

- Parton showering:** Due to the large momentum transfers during the hard process step, incoming and outgoing partons have high energies and are accelerated. Due to the acceleration, the strongly interacting partons carrying a color charge can emit QCD radiation by sequential emissions of quarks and gluons, including soft (small momentum) and collinear emissions. This leads to a PS [18, 19]. The extensively used algorithm for PS is the Sudakov form factor approach [20] which defines the probability for a parton that it does not split into other partons when evolving from a high scale to a smaller scale. This procedure is repeated until all partons in the event reach the ‘cut-off’ point of the hadronization scale. The PS produced by the incoming partons is called ISR, while the emission associated with outgoing partons is called

FSR. The PS approximation diverges for non-collinear parton splitting and to compensate this, a virtual loop corrections need to be included in the ME. As a consequence, the results from PS event generators are only reliable for the shape of distributions but not for the absolute normalization. The PS determines the flow of energy in an event without disturbing the total cross section.

- **Electroweak corrections:** Although strong interaction processes dominate at the LHC, but there are small and non-negligible contributions from electroweak processes. These electroweak corrections must be taken into account during the event generation.
- **Matching and merging:** The NLO and NNLO calculations have problems when trying to describe collinear and soft partons whereas the PSs are doing well in describing those. Clearly, a combination of both techniques would improve the description of any partonic state. PSs implement approximate higher order real emission corrections to arbitrary hard processes using the universal soft and collinear factorization properties of the hard cross sections. More sophisticated prescriptions are needed during combination due to the problems of double-counting hard radiation which is either generated by real emission or along with the PS and correctly covering phase space. For example, the Feynman diagram of an LO process with multiplicity $n + 1$ used by a ME generator looks exactly like the Feynman diagram of the same LO process with multiplicity n if one additional branching by a PS was taken into account as shown in Figure 3.3. This is because the PS contains all orders starting from the lowest order defined by the ME of the LO process with multiplicity zero. However, in NLO a further overlap arises due to including the additional contributions of the real emissions (ISR and FSR). One possibility for removing overlap is to replace the PS approximation at given orders in the strong coupling expansion by exact perturbative QCD results. This can be done in

two different ways [21, 22].

- **Matching:** The PS expression at FO is computed and subtracted from the higher order calculation to remove double-counting. The subtracted result is processed by the PS.
- **Merging:** A separate tree-level calculation is performed for each parton multiplicity of interest. Soft and collinear divergences of the hard MEs are regulated by resolution cuts. The PS is combined with all these calculations, and double-counting is removed by appropriate vetoes on shower branching.

In general, ME emissions are preferred to the PS ones, for their more accurate description of hard and angular separated radiations.

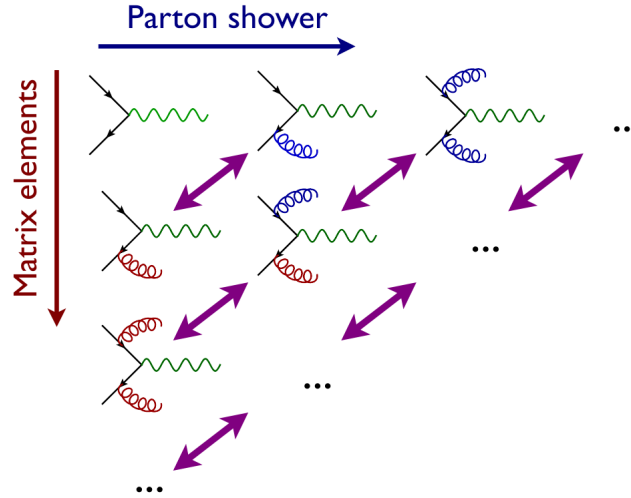


Figure 3.3: Double-counting: In the downward vertical direction Feynman diagrams of LO processes with increasing multiplicity (displayed in red) which are used by ME generators are displayed. The horizontal direction shows the Feynman diagrams of the corresponding LO processes with given multiplicity while the number of additional partons created by a PS (displayed in blue) is increased. A comparison between diagonal Feynman diagrams shows that they look alike which gives rise to the double-counting called overlap in the phase space [23].

- **Hadronization:** During the PS the involved partons lose energy by splitting and gluon emission. Once an energy scale of the order of 1 GeV (the exact

value depends on the hadronization model) is reached, the coupling between the individual color charged partons becomes so strong that they are bound into colorless hadrons which are detectable in the detector. This hadronization process cannot be described by the non-perturbative techniques currently available [24]. For this reason, hadronization models based on empirical data are needed for the description of this formation of hadrons. There are two most recent hadronization models: the “Lund string” and “cluster” models.

- **Lund string model:** This hadronization model [25] is built on the string analogy. In this model, the QCD potential between two quarks at low scales is parameterized by field lines seen to be compressed to a tube-like region by self interactions among soft gluons. The potential is given by $V = 4\pi/3r + kr$, where r is the distance between the two quarks, and k is the constant of the QCD potential ($k \approx 1 \text{ GeV/fb}$ from hadron spectroscopy). The first term represents the coulomb potential that has effects only at low distances, so it is only important for the internal organization inside hadrons rather than their formation. On stretching the string, the potential energy stored in the string will rise and kinetic energy decreases which break the string into two parts, creating additional color-singlet $q\bar{q}$ pairs, till the energy left to create another such pair, is less. The remaining partons can be identified with hadrons. The probability of breaking the string is given by the quantum mechanical probability of tunneling through a potential barrier. Hence, the fragmentation of heavy quark pairs is suppressed due to their higher masses. Charm and heavier quarks are not expected to be produced in those soft fragmentations. The fragmentation process is continues until all strings ending with quarks and antiquarks and all systems reach a cut-off scale. The string then transforms into a hadron or excited hadronic state based on its mass and flavor content.

- **Cluster hadronization model:** The cluster mode uses the property of pre-confinement [26] in QCD. This model starts with a non-perturbative splitting of gluons into $q\bar{q}$ pairs and quarks are clustered into colorless groups [27]. Clusters are made from the momentum space neighboring $q\bar{q}$ pairs and possess high invariant mass. If a cluster is too heavy it splits into two clusters, otherwise, these excited mesonic states decay directly into two hadrons according to the two hadron density of states. Even if the cluster model is less precise compared to the string one, it provides a complete description of events at high energy when combined with PS, by using a limited number of parameters.
- **The hadron decay stage:** During the hadronization, the excited and unstable hadrons, decay into further hadrons until only long-living or stable hadrons are left over. They are unstable enough to decay inside the detector but stable enough (like b hadron) to be detected before they decay. As a final step in the event simulation chain, the decay of these resonances into lighter hadrons is simulated.
- **Underlying events and Minimum Bias:** The UE is defined as the activity that is not associated with the particles originating from the hard scattering. The UE is softer but still produces new particles in the final state. The main contribution to the UE comes from color exchanges between the beam partons and is modeled in terms of BBR, MPI, and their associated ISR and FSR, and color reconnection (CR). The BBR includes hadrons from the fragmentation of spectator partons that do not exchange any appreciable transverse momentum in the collision. As a first approximation, it was assumed that these beam remnants keep on flying undisturbed. However, they can also split and emit gluons themselves and take part in the hadronization process together with their emitted daughter partons. The interactions among proton remnants involve both $1 \rightarrow 2$ and $2 \rightarrow 2$ processes, where the second is statistically

more probable. The production of these processes is divergent for $p_T \rightarrow 0$ and rapidly decreases for high- p_T : therefore in the generation mechanism, a lower threshold is introduced (~ 2 GeV). The MPI is additional parton-parton scatterings that occur within the same pp collision, and are softer in transverse momentum ($p_T \lesssim 3$ GeV) than the hard scattering. The majority of minimum bias collisions are soft, with a typical transverse momentum scale $p_T \lesssim 2$ GeV. Partons coming from hard scattering, MPI, and beam remnants can further interact by exchanging color between them. This color is usually represented by some color strings, which can overlap in space, and their interaction may lead to the merging of the two strings. This is known as color reconnection [28, 29]. The CR leads to the reduction of string length and hence a decrease in hadron multiplicity.

Figure 3.4 represent one of the pp collision events simulated by the MC event generator.

3.3 Monte Carlo event generators

The CMS collaboration is using a variety of event generators and different ways to simulate high energy processes. Some MC generators provide a possibility to generate a complete event while others provide only the hard interaction part and have to be supplemented by other MC generators to finalize the event generation. The PYTHIA [31], HERWIG [32–34], and SHERPA [35] generate ME up to the hadron formation and hadron decay also. While MADGRAPH5_aMC@NLO (denoted MG5_aMC) [36] and POWHEG [37] generate only ME and require PYTHIA or HERWIG to implement parton showering and hadronization steps. The MC event generators used in this thesis for the present work are described below.

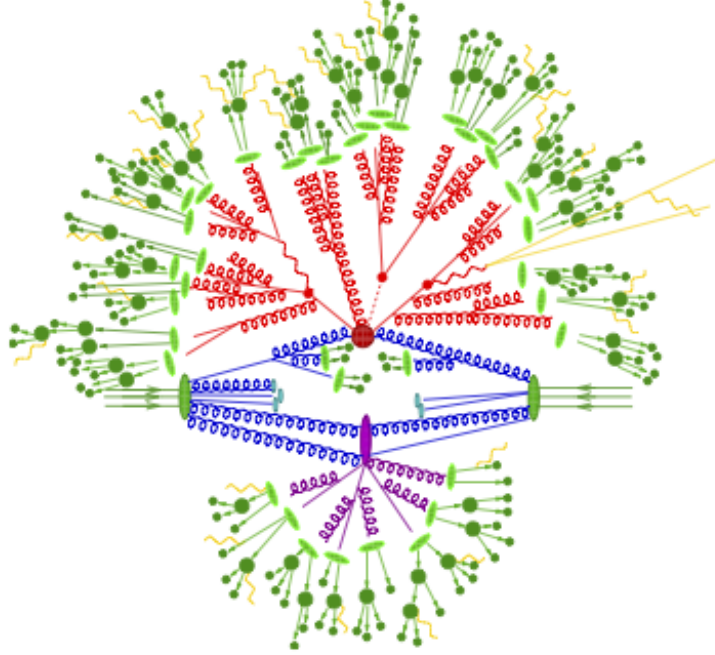


Figure 3.4: Overview of a pp collision as simulated by MC event generator. The three green lines coming from both the left and right of the diagram represent the valence quarks of the protons and the green ellipses represent PDFs of protons. The large dark red circle represents the hard scattering. The particles produced in hard scattering further decay, as indicated by the small red circles, and also emits other gluons and quarks (FSR), in red. Hard scatter surrounded by ISR in blue and FSR in red. Hadronization is represented by light green ellipses. Unstable hadrons further decay until stable final state particles are formed represented by dark green circles. The purple ellipse indicates an MPI. The remaining particles of the incoming protons make up the BBR, in cyan, also part of the UE. The yellow lines represent soft photon radiation [30].

3.3.1 PYTHIA

This general-purpose generator has two versions: PYTHIA6 [38] and PYTHIA8 [39] which are based on Fortran and C++, respectively. The MC samples used in this thesis are based on an advanced version of PYTHIA program i.e. PYTHIA8.2 [40]. PYTHIA is used to simulate mostly $2 \rightarrow 2$ processes ME at LO and has tools like p_T -ordered ISR, FSR, MPI, and many other processes. In PYTHIA8.2 event generator, the Lund string model is used for hadronization and dipole shower (it takes soft gluon coherence into account) for PS. In dipole shower, gluon emission is not generated by single parton splitting but by splitting of a dipole formed from a pair of partons into two dipoles. PYTHIA8 uses new MPI models [41], which are interleaved with parton showering. The CUETP8M1 [42] and CP5 [43] tunes, where CUET (read as “cute”) stands for “CMS UE”, P8M1 stands for “pythia8 Monash 1”, and CP stands for “CMS PYTHIA 8” are used for the 2016 and 2017–2018 data-taking periods, respectively. Typically, the values of strong coupling used for the simulation of the hard scattering are chosen to be consistent with the order of the PDFs set used. The CUETP8M1 tune includes the LO NNPDF 2.3 [44] PDFs set with the strong coupling $\alpha_s(m_Z)$ set at 0.137 for space- and time-like shower simulation. The CP5 tune uses the NNPDF3.1 PDF set at NNLO, with α_s values of 0.118, and running according to NLO evolution. But ISR emissions are ordered according to rapidity. The CMS collaboration widely uses PYTHIA for the hadronization and showering in the majority of the SM and SUSY physics.

3.3.2 SHERPA

SHERPA stands for the ‘Simulation of **H**igh **E**nergy Reactions of **P**Articles’ written in C++. SHERPA has two built-in ME generators AMEGIC++ [45] and Comix [46], which allow to generate LO and NLO MEs for a wide range of subprocesses and is not limited to $2 \rightarrow 2$ processes. While virtual MEs contributing to QCD NLO

corrections are provided by openloops [47]. It provides automated merging of LO and NLO ME results with its own PS model with the MC@NLO prescription. In this thesis, the SHERPA v2.2.4 [48] is used to generate $pp \rightarrow Z + n \text{ jets}$ events, with $n \leq 2$ at NLO and $n = 3, 4$ at LO. The merging cut parameter is set to $Q_{\text{cut}} = 20 \text{ GeV}$. The PS and MPI models in SHERPA are based upon PYTHIA model and a cluster model for the hadronization. In the ME calculation, the NNPDF3.0 PDF set at NLO is used with α_s values 0.118 at the Z boson mass.

3.3.3 POWHEG

As the name, POWHEG stands for **P**ositive **W**eight **H**ardest **E**mission Generator, it generates events only with positive weights at NLO precision. A package for the implementation of POWHEG for generic processes, known as the POWHEGBOX [49], has been created; to build the POWHEG implementation automatically. In the POWHEG method, the hardest radiation is generated first using the exact NLO MEs. Furthermore, because the algorithm does not depend on a particular Shower Monte Carlo (SMC) program, the POWHEG output can be easily interfaced with any modern SMC generator. It has been interfaced to HERWIG or PYTHIA that is either p_T -ordered, or allows the implementation of a p_T veto to perform parton and hadronization steps. The hardest emission always comes first in the case of shower generators, which are ordered according to p_T , but in this case, POWHEG simply replaces the hardest emission with its own NLO accurate emission. In this present work, the processes are generated at NLO by POWHEG v2.0 [37].

3.3.4 MADGRAPH5_aMC@NLO

MADGRAPH [50, 51] is a tool for automatically generating MEs for HEP processes, such as decays and $2 \rightarrow n$ scatterings. MG5_aMC (NLO) is a fusion of the MADGRAPH LO ME + PS and NLO + PS event generator aMC@NLO. In NLO the real emission MEs fill parts of the phase space which are also later filled by the PS, resulting

in an overlap called counterterms. The overlap is taken into account by adding or subtracting [52–54] these MC counterterms during the computation of the parton-level short distance cross sections and the PS can later be applied to the parton-level events. Due to the MC counterterms event samples generated by `amc@NLO` consist of a large number of events with positive weights, but also includes a small number of events with negative weights. The ratio of these two event classes depends on the specific process. The fraction of events with negative weight increases for processes with higher multiplicities. In the present thesis, the DY process simulated with two different predictions of `MG5_aMC`: `MG5_aMC` (LO) and `MG5_aMC` (NLO). A first sample is produced with `MG5_aMC` version 2.2.2 (2.4.2) for the 2016 (2017–2018) data-taking period using LO MEs for $pp \rightarrow Z + n \text{ jets}$ ($n \leq 4$), with the MLM matching scheme [55]. The NNPDF 3.0 LO and NNPDF 3.1 NNLO PDF sets are used for the 2016 and 2017–2018 data-taking periods, respectively. A second sample has exclusive jet multiplicity up to 2 at NLO precision by `MG5_aMC` version 2.3.2.2 for 2016 data and version 2.6.0 for the 2017–2018 data with the FFX [56] matching between the jets from ME calculations and PSs. The `MG5_aMC` generator is interfaced with `PYTHIA` v8.212 (v8.230) for parton showering, hadronization, and the underlying event simulation using tune CUETP8M1 (CP5) for the 2016 (2017–2018) data-taking period. Matching between the ME generators and the PS jets is done with the matching scales set at 19 and 30 GeV for the LO and NLO `MADGRAPH5` generators, respectively.

3.4 Detector simulation

High energy physics experiments depend critically on the accuracy of physics generators and detector simulations. Simulated events are used for detector design optimization, calibration, object identification, and physics analysis. The size of systematic uncertainties associated with particle discoveries, mass, or cross section

measurements are tightly associated with how accurately the simulations describe the actual performance of the detector in measuring leptons, photons, hadrons, etc. It is essential for the success of a HEP experiment to understand and tune the physics of the simulation tool to agree with measurements. The MC generators mentioned in the previous Section only model physics processes at parton or generator-level. Thus, the interaction of particles with the detector material has to be simulated separately. So, after the MC event generation, the final state particles pass through the CMS detector for the detailed simulation of the detector effect which is based on the GEANT4 toolkit. The GEANT4 10.0p02 [57, 58] is used in order to improve performance for large scale production for Run2. This includes the ability to construct a precise geometrical model of the detector itself and a model of the detector's magnetic field as well as the physical properties of the materials used. This package also handles the properties of the fundamental particles involved, how they behave in the modeled magnetic field, and how they interact (via bremsstrahlung, multiple scattering, etc.) with the detector material. This software is also capable to handle particle energy losses across each subdetector as they pass through them. The deposition of particle energy in each subdetector is simulated with GEANT4. These simulated hits are used as input to emulators that model the response or hits of the detector readout and trigger electronics (called GEN-SIM sample) and digitize this information by also taking into account noise and other factors. This step is known as digitization or raw data (called RAW-DIGI sample). After this, the full detector response is available (as in real operation), and the event can be reconstructed (called RECO-AOD sample). The physics analysis is generally performed by using the AOD/MINIAOD/NANOAOD samples which contain all the relevant information of reconstructed objects at the detector-level. The MINIAOD samples are used for the analysis work presented in this thesis. The CMS uses a tiered structure [59] for the storage and processing of the samples. MC samples are generally stored in Tier-2 centers and can be accessed for physics analysis. All the data and

MC simulated samples are analyzed by the CMSSW framework.

3.5 Physics object reconstruction

Reconstruction is the operation of constructing physics quantities from the raw data collected in the experiment. The reconstruction process can be divided into three steps, corresponding to local reconstruction within an individual subdetector module, global reconstruction within a whole subdetector, and combination of these reconstructed objects to produce higher-level objects, as shown in Figure [3.5](#). The reconstruction units provide local reconstruction in a subdetector module use as input real data from the DAQ system or simulated data. This data in either case is called “digis”. The output from the reconstruction units is reconstructed hits (“RecHits”) which are typically position measurements (from times or clusters of strips or pixels) in tracking type detectors (Muon and Tracker systems) and calorimetric clusters in Calorimeter systems. The RecHits are added as EDProducts to the event and used as the input for the global reconstruction. In the global reconstruction step information from the different modules of a single subdetector is combined to produce further objects which represent the best measurement from that subdetector. At this stage, no attempt is made to link the information from different subdetectors. For example, tracker RecHits are used to produce reconstructed charged particle tracks and muon RecHits are used to produce candidate muon tracks. After combining all the objects from each subdetector, the high-level objects like the electron, muon, b-jets, etc. are reconstructed by using the particle-flow (PF) event reconstruction algorithm [\[60\]](#). The analysis work presented in this thesis deals with the $Z + b$ -jets process, with the decay of the Z boson in the dimuon or dielectron final state. Figure [3.6](#) shows one of the events displayed in the CMS detector where two muons (indicated by red lines) were detected in the muon chambers (indicated by red slabs) and two jets of particles (visualized by cones). The

processes of reconstruction of all the objects used in this analysis are described in the next Subsections.

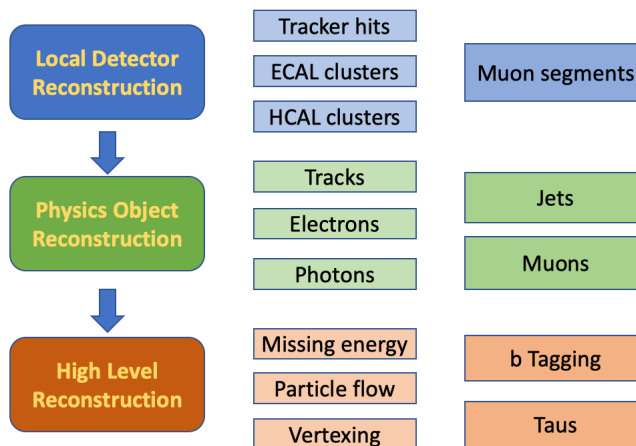


Figure 3.5: Reconstruction algorithm for physics objects in CMS.

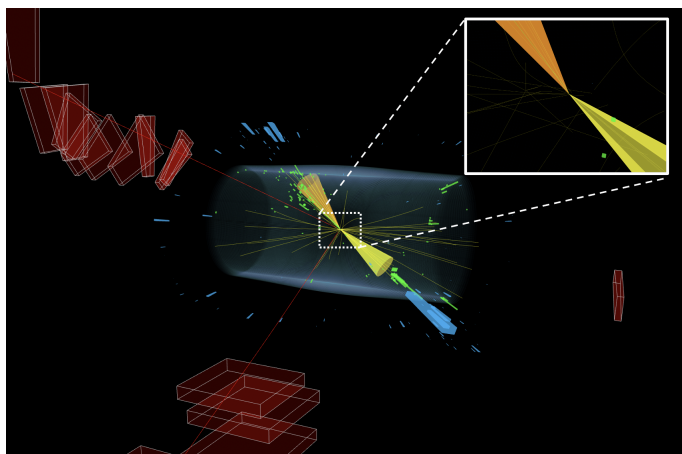


Figure 3.6: A pp interaction event recorded by the CMS experiment. It features two muons (indicated by red lines) detected in the muon chambers (indicated by red slabs) and two jets of particles (visualized by cones). The upper-right corner of the picture is a zoom-in view of the pp interaction vertex [61].

3.5.1 Beamspot reconstruction

The beamspot is the luminous region produced by the collisions of proton beams [62]. The position and spread of the beamspot need to be measured precisely because they are an important input for physics processes such as quark flavor tagging of jets, a quick check of the global alignment for HLT, and extraction of tracking impact parameter resolution. The algorithm used to calculate the transverse beam position is the so called d_0 - ϕ algorithm. For each track the transverse impact parameter d_0 —the shortest distance in the transverse plane between the track and the nominal beam axis—is determined. This distance is then signed according to the angular momentum convention. The polar coordinates of the beamspot are inferred from a simple least-squares fit with a sine function to the shape. This is a robust and fast χ^2 fit that just needs 1000 good tracks to reach micron precision. Using a Log-Likelihood fit, the beam width can be extracted using 20000 good tracks.

3.5.2 Track reconstruction

Event reconstruction starts with the reconstruction of charged particle tracks, which further contributes to the reconstruction of physics objects. First, local charge clusters are converted into hits using the digitized output of the tracking system. The local track reconstruction output is then buffered to the global track reconstruction, whose aim is to identify the hit combinations that match to possible trajectories of the charged particles present in the event. A combinatorial track finder [63] (CTF) based on Kalman Filtering (KF) [64] was used to reconstruct these tracks in four steps:

- **Seed generation:** Each iteration begins with a seed which is constructed using 2 or 3 hits in different layers of the pixel detector as shown in Figure 3.7 (left). A seed defines the initial estimate of the trajectory parameters and their uncertainties. Hits that are compatible with the interaction region above

a certain p_T threshold are considered charged particle track candidates. The pixel hits provide the best track seeding because of their three-dimensional position information and lower occupancy.

- **Track finding or pattern recognition:** Track finding is based on a KF. It extrapolates the seed trajectories along the expected flight path of a charged particle, searching for additional hits that can be assigned to the track candidate. The compatible hits are selected based on the χ^2 between the predicted and measured positions. This is done layer by layer and when the compatible hit is found, it gets attached to the track before going into the next layer. This procedure is continued until the last layer is reached as shown in Figure 3.7 (middle). At the end of each iteration, quality flags are set based on the compatibility of the hits with the extrapolated trajectory.
- **Track fitting:** As soon as the final track of the particle is found from the pattern recognition step, tracks are fitted using KF to get the track properties such as origin, transverse momentum, and its direction, as shown in Figure 3.7 (right).
- **Track selection:** The track finding procedure described above yields a significant fraction of fake tracks, where a fake track is defined as a reconstructed track that is not associated with a charged particle. Thus to avoid this, the track selection step is utilized which reduces the misreconstruction rate by implementing thresholds on various track parameters such as the number of tracker layers with hits, the number of tracker layers missing a hit where a hit was expected, quality of the track fit, impact parameters of the tracks, etc.

3.5.3 Primary Vertex Reconstruction

The reconstructed tracks are used as input to reconstruct all the true vertices in a collision, Primary Vertex (PV) along with the pileup vertices. The Deterministic

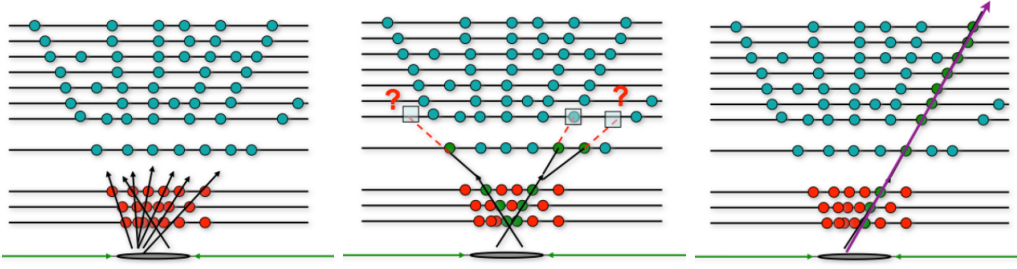


Figure 3.7: Seeding (left), building (middle), and fitting (right) [65].

Annealing (DA) clustering method [66] is employed to reconstruct offline PV which is done in two steps. The first step is to arrange the tracks in clusters according to the track points z coordinate near the beamline. Second, a three-dimensional vertex fit is performed for each of the tracks in the cluster. The fit decides whether to accept or reject the reconstructed vertex. If the vertex gets accepted, it should have four degrees of freedom (n_{dof}). This results in multiple primary vertices reconstruction and all of them get stored. Further, the PV must satisfy two conditions that are described below:

- 1) Its z position should be within 24 cm from beamspot.
- 2) Radial position within 2 cm from the beamspot.

After having a set of selected/feasible primary vertices, the one vertex which has the maximum sum of the square of the transverse momentum of the tracks associated with it, is chosen as the hard scattering vertex. The PV is then used to measure the location and associated uncertainty of pp interactions in each event.

3.5.4 Secondary vertex reconstruction

The $b(c)$ hadron travels ≈ 1 cm at an energy of ≈ 10 – 100 GeV in the laboratory frame before decaying to several particles, thus giving rise to displaced tracks from which a secondary vertex (SV) is reconstructed as illustrated in Figure 3.8. The displacement

of tracks with respect to the PV is characterized by their impact parameter (IP), which is defined as the distance between the PV and the tracks at their points of the closest approach. The vector pointing from the PV to the point of the closest approach is referred to as the IP vector. The IP value can be defined in three spatial dimensions (3D) or in the plane transverse to the beamline (2D). The longitudinal IP is defined along the beamline in one dimension. The IP is defined to be positive or negative, with a positive sign indicating that the track is produced “upstream”. Due to the long lifetime of the b/c hadrons, the IP from b/c jets is expected to be mainly positive, while for the light jets it is almost symmetric with respect to zero. An important feature of the IP is that it is Lorentz invariant. The IP significance is defined as the IP value divided by its uncertainty, $IP/\sigma(IP)$. The IP significance is used in order to take into account resolution effects.

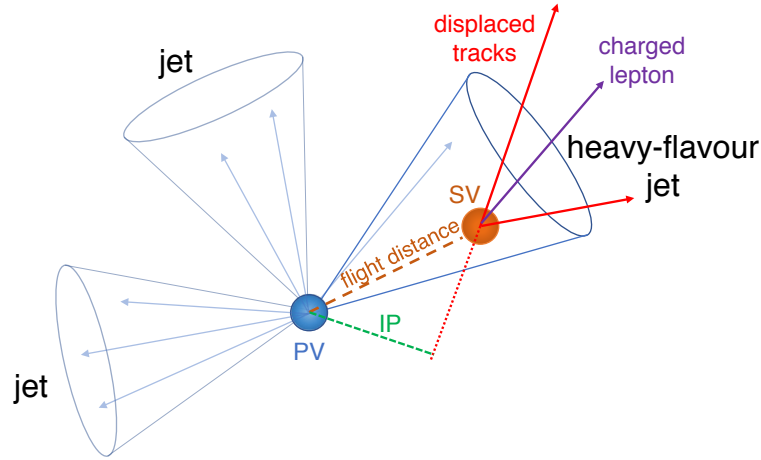


Figure 3.8: Illustration of a heavy flavor jet with a SV from the decay of a b or c hadron resulting in charged particle tracks (including possibly a soft lepton) that are displaced with respect to the PV, and hence with a large IP value.

The Inclusive Vertex Finding (IVF) algorithm is used for reconstructing secondary vertices [67]. The IVF uses as input all of the reconstructed tracks in the event with $p_T > 0.8 \text{ GeV}$ and a longitudinal IP $< 0.3 \text{ cm}$. The IVF procedure starts by identifying seed tracks with a 3D impact parameter value of at least $50 \mu\text{m}$ and a 2D impact parameter significance of at least 1.2. After identification of the seed

tracks, the IVF procedure includes the following steps:

- **Track clustering:** The compatibility between a seed track and any other track is evaluated using requirements on the distance at the point of closest approach of the two tracks and the angle between them. In addition, the distance between the seed track and any other track at their points of closest approach is required to be smaller than the distance between the track and the PV at their points of closest approach.
- **Secondary vertex fitting and cleaning:** In order to determine the position of the secondary vertices, the sets of clustered tracks are fitted with the adaptive vertex fitter algorithm. After the fit, secondary vertices with a 2D (3D) flight distance significance smaller than 2.5 (0.5) and 1.25 (0.25) are removed for the b and c tagging algorithm, respectively. In addition, if two secondary vertices share 70% or more of their tracks, or if the significance of the flight distance between the two secondary vertices is less than 2, one of the two secondary vertices is dropped from the secondary vertices collection.
- **Track arbitration:** At this stage, a track could be assigned to both the PV and SV. To resolve this ambiguity, a track is discarded from the SV if it is more compatible with the PV. This is the case if the angular distance between the track and SV flight direction $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} > 0.4$, and if the distance between the SV and the track is larger than the absolute IP value of the track.
- **Secondary vertex refitting and cleaning:** The SV position is refitted after track arbitration and if there are still two or more tracks associated with the SV. After refitting the SV positions, a second check for duplicate vertices is performed. This time, a SV is removed from the collection of secondary vertices when it shares at least 20% of its tracks with another SV and the

significance of the flight distance between the two secondary vertices is less than 10.

This results in multiple secondary vertices reconstruction and all of them get stored. Further, the secondary vertices are rejected when they share 80% or more of their tracks with the PV, and when the 2D flight distance significance is less than 2 (1.5) for secondary vertices used in b (c) tagging algorithms. The remaining secondary vertices are then associated with the jets by requiring the angular distance between the jet axis and the SV flight direction to satisfy $\Delta R < 0.3$.

3.5.5 Electron reconstruction

Electrons produce hits in the tracker layers and deposit almost all of their energy in the ECAL as shown in Figure 3.9. Electrons are reconstructed by associating a track reconstructed in the silicon detector with a cluster of energy in the ECAL [68]. The energy reconstruction algorithm starts with the formation of “topological” clusters by grouping cells with energy exceeding a predefined threshold to reject the noise (typically ~ 80 MeV (~ 300 MeV) for the EB (EE). A seed cluster is then defined as the one containing most of the energy deposited in any considered region, with a minimum transverse energy (E_T^{seed}) above 1 GeV. ECAL clusters within a certain geometric area (“window”) around the seed cluster are combined into superclusters (SC) to account for photon conversions and bremsstrahlung losses. This procedure is referred to as “superclustering” and the combining process uses two algorithms. The first is the “mustache” algorithm, which is particularly helpful to properly measure low energy deposits. It uses information only from the ECAL and preshower detector. The algorithm starts from a cluster above a given threshold, called seed cluster. Additional clusters are added if falling into a zone, whose shape is similar to a mustache in the transverse plane. The second superclustering algorithm is known as the “refined” algorithm. It utilizes tracking information to extrapolate bremsstrahlung tangents and conversion tracks to decide whether a cluster should

belong to a SC. It uses mustache SCs as a starting point but is also capable of creating its own SCs. The refined SCs are used for the determination of all ECAL-based quantities of electron objects.

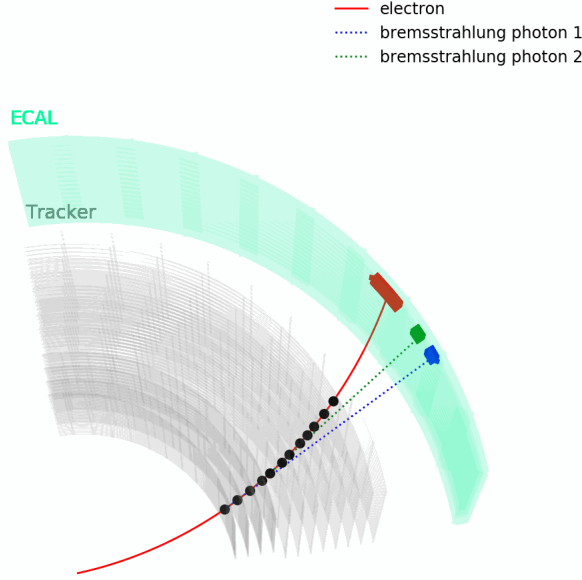


Figure 3.9: An electron (and its bremsstrahlung photons) passing through the tracker and creating clusters of energy in the electromagnetic calorimeter (ECAL).

The reconstruction of electron tracks begins with the identification of a hit pattern that might lie on an electron trajectory (“seeding”). The following two algorithms are used to generate trajectory seeds for the electrons:

- **ECAL-driven seeding:** This algorithm first selects mustache SCs with transverse energy $E_{\text{SC},T} > 4 \text{ GeV}$ and $H/E_{\text{SC}} < 0.15$, where E_{SC} and H are the SC energy and the sum of the energy deposits in the HCAL towers within a cone of $\Delta R = 0.15$ centered on the SC position. Each mustache SC is then compared in ϕ and z to a collection of track seeds that are formed by combining multiple hits in the inner tracker detector: triplets or doublets. The extrapolation towards the collision vertex neglects the effect of any photon emission. To deal with the highly non-Gaussian bremsstrahlung emission, a dedicated tracker

algorithm, namely Gaussian Sum Filter (GSF) [69] is deployed in the CMS detector. If the first two hits of a tracker seed are matched to the predicted trajectory for a SC under any charge hypothesis, it is selected for seeding a GSF track. The ECAL-driven approach performs better for high E_T isolated electrons with a larger than 95% seeding efficiency for $E_T > 10$ GeV for electrons from Z boson decay.

- **Tracker-driven:** This approach iterates over all generic tracks. The generic tracks are a collection of filtered tracks (not specific to electrons) with $p_T > 2$ GeV, reconstructed from hits in the pixel detector with KF. If any of these KF tracks is found to be compatible with an ECAL cluster, its track seed is used to seed a GSF track. The compatibility criteria are the logical OR of a cut-off based selection and a multivariate selection based on a boosted decision tree (BDT) [70], using track quality and track-cluster matching variables as inputs. Since it is computationally expensive to reconstruct all tracks in an event, tracker-driven seeding is performed only in the offline reconstruction and not in HLT. The tracker-driven approach is designed to recover efficiency for low p_T or non-isolated (electrons within jets) electrons with an efficiency more than $\sim 50\%$ for $p_T > 3$ GeV for electrons from b hadron decays. The tracker-driven approach also helps to recover efficiency in the ECAL regions with less precise energy measurements, such as in the barrel-endcap transition region and/or the gap between supermodules. The tracker-driven seeding has an efficiency of $\sim 50\%$ for electrons from Z decay with $p_T \sim 3$ GeV and drops to less than 5% for $p_T > 10$ GeV.

The GSF tracking algorithm is run on all ECAL- and tracker-driven seeds. The final collection of selected electron seeds (obtained by combining the ECAL-driven and tracker-driven seeds) is used to initiate the reconstruction of electron tracks using the PF reconstruction algorithm. The PF reconstruction algorithm links them to the PF-clusters in the ECAL. The electron seeds may not always originate

from isolated electrons coming from the PV of an event, but could be caused by background sources emanating from photon conversions into electron-positron pairs in the tracker material or misidentified jets. The conversion rejection variables are defined based on the patterns of hits registered in the tracker to reject secondary electrons produced from photon conversions. When a photon conversions take place inside the volume of the tracker, the first hit on electron tracks from the converted photons is often not located in the innermost layer of the tracker, and missing hits are therefore present in that region. For prompt electrons, whose trajectories start close to the beamline, no missing hits are expected in the inner layers. Calorimetric observables are used to separate genuine electrons (signal electrons or electrons from photon conversions) from misidentified electrons (jets with large electromagnetic components) based on the transverse shape of electromagnetic showers in the ECAL and exploiting the fact that electromagnetic showers are narrower than hadronic showers. The fractions of energy deposited in the HCAL (expected to be small, as electromagnetic showers are essentially fully contained in the ECAL) and the ECAL supercluster also provide the required discrimination against the misidentified electrons. Such variables are also used as an input to train BDTs and the final BDT discriminant is used to identify real electrons.

3.5.6 Muon reconstruction

Muons are identified with high efficiency by the muon system and precise measurement of their momentum involves the inner tracker [71]. Local reconstruction uses information from only a single muon chamber (RPC, CSC, or DT) to specify the passage of a muon through the chamber. The RPC chambers are single-layer chambers, while the CSC and DT chambers are multi-layer detectors where hits are reconstructed in each layer. From the reconstructed hits, straight line track “segments” are built within each CSC or DT chamber. The following muon collections are required for the offline muon reconstruction:

- **Standalone muon:** Muon tracks are built by fitting the DT, CSC, and RPC hits found along the muon trajectory in the muon system using a KF. Reconstruction starts with seeds made up of groups of DT or CSC segments.
- **Global muon:** Muon tracks are built “outside-in” by matching standalone muon tracks with tracker tracks. The matching is done by comparing the parameters of the two tracks propagated onto a common surface. A combined fit is performed with the KF using information from both the tracker track and standalone muon track. Global muon reconstruction improves the momentum resolution of high- p_T muons compared to the measurement obtained with a tracker only fit.
- **Tracker muon:** Muon tracks are built “inside-out” by matching tracker tracks to the muon system with at least one DT or CSC segment matching. Each tracker track with transverse momentum $p_T > 0.5 \text{ GeV}$ and a total momentum (p) greater than 2.5 GeV is extrapolated to the muon system taking into account the mean expected energy losses, the magnetic field strength, and the amount of coulomb scattering in the tracking material. The tracker muon reconstruction has the advantage of recovering low p_T muons, which usually fails in the global muon reconstruction sequence since they can not penetrate through the material until the last muon stations.

All such kind of muons and their reconstruction is demonstrated in Figure [3.10](#).

Tracker muons have high efficiency in regions of the CMS detector with less instrumentation (for routing of detector services) and for muons with low p_T . The tracker muons that are not global muons typically match only to segments in the innermost muon station, but no other stations. This increases the probability of muon misidentification since hadron shower remnants can reach this innermost muon station (punch-through). Global muon reconstruction, which uses standalone muon tracks, is designed to have high efficiency for muons penetrating through more than one muon station, which reduces the muon misidentification rate compared to tracker

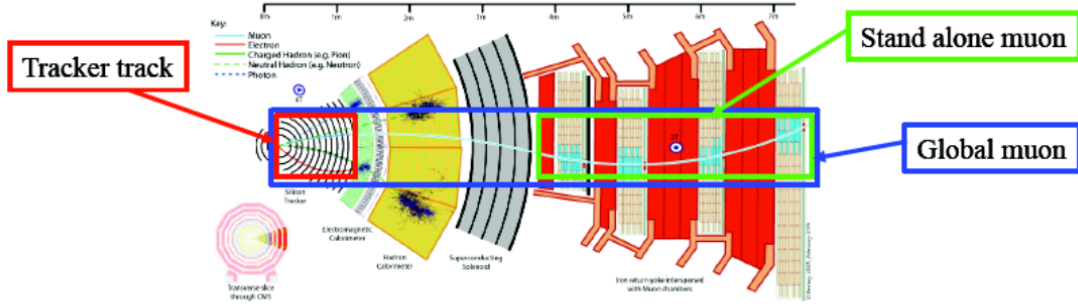


Figure 3.10: Reconstruction of different types of muons.

muons. By fully exploiting the information from both the inner tracker and muon system, the p_T measurement of global muons is also improved compared to tracker muons, especially for $p_T > 200$ GeV. Standalone muons have worse momentum resolution and a higher admixture of cosmic muons than global or tracker muons.

3.5.7 Lepton Isolation

In order to avoid double-counting of leptons both as individual particles and jet constituents, the isolated leptons are preventively excluded from the jet clustering sequence. It is necessary to discriminate between leptons produced in prompt decays and those produced by heavy flavor decays. Since non-prompt leptons are usually produced inside jets, the detection of other objects surrounding the lepton candidate can be used to discriminate them. Isolation variables are used for that purpose and built considering a cone in ΔR around the reconstructed lepton. The PF relative isolation variable is defined as:

$$I_{\text{rel}} = \frac{\sum p_{Ti}^{\text{charged}} + \max(0, \sum E_{Ti}^{\text{neutral}} + \sum E_{Ti}^{\gamma} - p_T^{\text{PU}})}{p_T^l}, \quad (3.2)$$

which is the sum of the p_T values for all PF candidates within a cone of radius $\Delta R = 0.3$ (0.4) centered on the electron (muon) track direction.

Here, the quantity $\sum p_{Ti}^{\text{charged}}$ represents the p_T sum of the charged hadrons associated with the PV and p_T^{PU} is the contribution from PU. The $\sum E_{Ti}^{\text{neutral}}$ and $\sum E_{Ti}^{\gamma}$ are the scalar sums of the transverse energies of neutral hadrons and photons, respectively. Finally, p_T^l stands for the lepton transverse momentum. The charged particle component of the isolation is very well measured due to the tracking system. However, the neutral component, which is composed of neutral hadrons and photons, requires extra corrections (subtraction of PU contribution). Since the contribution from PU cannot be directly subtracted, because no tracking information is available in the neutral component to discriminate between the different primary vertices. So, two different approaches are followed for PU subtraction: the $\Delta\beta$ corrections, and the effective area corrections. The $\Delta\beta$ corrections are applied to the neutral component of the isolation by computing the contribution from charged hadrons coming from PU in the cone. This quantity is scaled by a factor of 0.5, corresponding to the ratio between charged and neutral particle production rates at inelastic pp collisions, as observed in simulations. The effective area correction is similar to the jet areas method used to subtract the PU contribution from jets [72]. In this case, the PU correction is assumed to be $p_T^{\text{PU}} = \rho A_{\text{eff}}$, where ρ is the average density of an event, defined as the median of the energy density of particles within the area of any jet, reconstructed with a R parameter of 0.6. A_{eff} corresponds to the dependence of isolation as a function of the number of reconstructed vertices in the collision. For electrons, p_T^{PU} is evaluated using the “jet area” method. For muons, p_T^{PU} is taken to be half of the scalar p_T sum deposited in the isolation cone by charged particles not associated with the PV. For electron, the numerator part is denoted by (I_{combine}).

3.5.8 Missing Transverse Energy

Neutral weakly interacting particles, such as neutrinos, escape from typical collider detectors without producing any direct response in the detector elements. The presence of such particles must be inferred from the imbalance of total momentum. The vector momentum imbalance in the plane perpendicular to the beam direction is known as Missing Transverse Energy (MET or p_T^{miss}) [73]. The p_T^{miss} magnitude of the negative vector sum of the p_T of all the reconstructed particles in an event.

$$p_T^{\text{miss}} = \left| - \sum_i \vec{p}_T^i \right| \quad (3.3)$$

where the index i runs over all particle flow candidates. It is also an important variable in searches for new weakly interacting, long-lived particles. Many BSM scenarios, including supersymmetry, predict events with large p_T^{miss} . The reconstruction of p_T^{miss} is very sensitive to particle momentum mismeasurements, particle misidentification, detector malfunctions, particles impinging on poorly instrumented regions of the detector, and cosmic-ray particles which may result in artificial MET [74]. In the CMS detector, p_T^{miss} is calculated either using the energy depositions in calorimeters (known as CaloMET) or by using the momenta of all the reconstructed particles (PFMET). The physics analysis presented in this thesis uses the PFMET, which is more accurate than the CaloMET.

3.5.9 Jets Reconstruction

A jet can be defined as a collimated spray of stable particles arising from the fragmentation and hadronization of a parton after a collision as explained in Section 3.2. Jet reconstruction algorithms combine the calorimetry and tracking information to define jets. A basic illustration of a collision of two protons, the subsequent particle shower, and a reconstructed jet is shown in Figure 3.11. There are two main classes

of jet algorithms [75]. The first is the cone algorithms, of which the most important are the iterative cone with progressive removal (IC-PR), the iterative cone with the split-merge procedure (IC-SM), and the seedless infrared safe cone (SIScone). The second class is the sequential clustering algorithms which comprise of the k_T , anti- k_T , and the Cambridge/Aachen algorithms. Jets are classified into four types as discussed below:

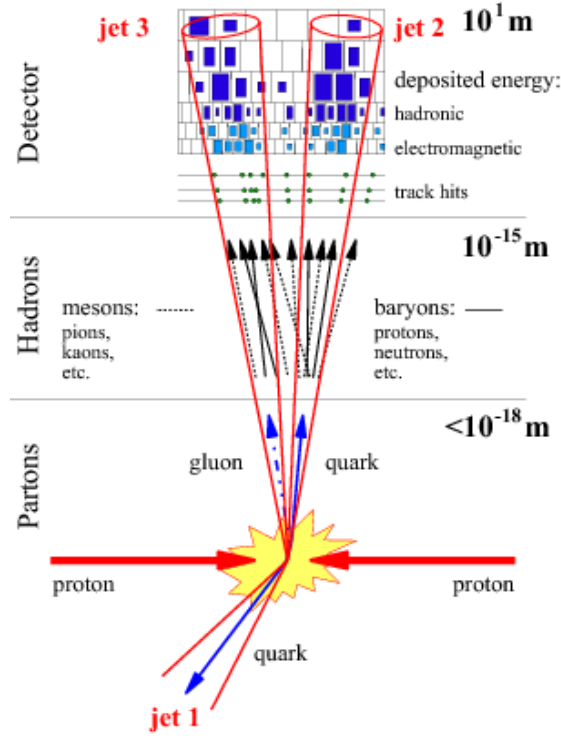


Figure 3.11: Illustration of a jet to which bundles of partons, hadrons, or detector measurements are grouped together [76].

- **Calorimeter Jets (CaloJets):** The CaloJets are reconstructed from energy deposits in ECAL crystal and calorimeter towers in the HCAL sub-detector [77]. CaloJets have a relatively poor resolution because of the compact size of HCAL.
- **Jet-plus-track (JPT) jets:** These jets are reconstructed using the information from the tracker and calorimeters [78] and rely on a high quality tracking

system. The JPT algorithm corrects both the energy and direction of the axis of the original calorimeter jet.

- **Particle Flow (PF) jets:** PF jets are reconstructed by clustering the four-momentum vectors of PF candidates. The PF algorithm combines the information from all relevant CMS subdetectors to identify and reconstruct all visible particles in the event, namely muons, electrons, photons, charged hadrons, and neutral hadrons. The PF jet momentum and spatial resolutions are greatly improved with respect to calorimeter jets, because the use of the tracking detectors and high granularity of ECAL allows the resolution and measurement of charged hadrons and photons inside a jet, which together constitute $\sim 85\%$ of the jet energy. In reconstructing the four-momentum of PF candidate, photons are assumed massless, and charged hadrons are assigned the charged pion mass.
- **MC particle jets or Truth jets:** The MC particle jets are reconstructed by clustering the four-momentum vectors of all stable ($c\tau > 1\text{ cm}$) particles generated in the simulation. In particular, there are two types of MC particle jets: those where the neutrinos are excluded from the clustering, and those where both the neutrinos and the muons are excluded. The former is used for the study of the PF and JPT jet response in the simulation, while the latter is used for the study of the Calo jet response (because muons are minimum ionizing particles and therefore do not contribute appreciably to the Calo jet reconstruction).

In the present work, jets are clustered from PF candidates using the infrared- and collinear-safe anti- k_T algorithm with a distance parameter of 0.4, as implemented in the FASTJET package [79]. Infrared and collinear safety is the property that if an event is modified by a collinear splitting or the addition of a soft emission, the set of hard jets that are found in the event should remain unchanged. The anti- k_T algorithm combines the PF objects to cluster into jets and calculates the distance

between beamline and detected particle (d_{iB}) as well as the distance between two particles i and j (i.e. pseudo-jets, particles), d_{ij} as shown below:

$$d_{iB} = k_{T,i}^{2p} \quad (3.4)$$

$$d_{ij} = \min(k_{T,i}^{2p}, k_{T,j}^{2p}) \frac{\Delta R_{ij}^2}{R^2} \quad (3.5)$$

where ΔR_{ij}^2 is the $(\eta-\phi)$ space distance between the two particles and R is the radius parameter that determines the final size of the jet and it is usually between 0.4–0.7. In the present study, a cone size of $\Delta R = 0.4$ is used, as this provides a good agreement between the energy of jet itself and external energy coming from overlap with other jets, PU interactions, ISR or FSR, etc. Variable “ p ” is added to govern the relative power of energy versus geometrical scales to differentiate the three algorithms i.e. if $p = 0$, it corresponds to Cambridge/Aachen algorithm [80], $p = 1$ and $p = -1$ corresponds to k_T [81] and anti- k_T [82] algorithms, respectively. The algorithm proceeds by finding the smallest of the two distances d_{ij} and d_{iB} . If d_{ij} is the smallest, combine i and j to have one particle (ij) using the summation of four-vectors after which i and j are removed from the list of particles. If d_{iB} is the minimum, i will be a part of final jet and it is removed from the list of particles. The distances are recalculated and the process is repeated until either all particles are part of a jet with the distance between the jet axis $R_{ij} > R$, which is inclusive clustering or until a desired amount of jets have been found, this is exclusive clustering. In anti- k_T the d_{ij} is dominated by high- p_T and hence this algorithm favors to cluster high- p_T particles first as shown in Figure 3.12. This algorithm is less susceptible to soft radiation i.e. UE and PU, and also infrared safety and collinear safe.

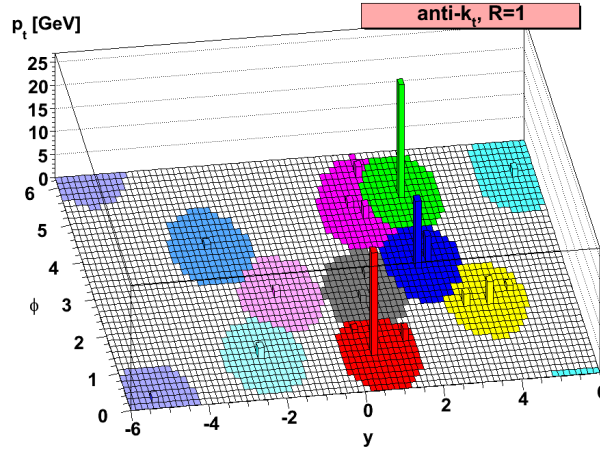


Figure 3.12: Parton-level jets clustered with cone size $R=1$ using anti- k_T jet algorithm [82].

3.5.10 Jet corrections

The energy of a PF jet might not optimal due to the non-linear response of the calorimeter to particles or due to PU contamination. Therefore, a set of jet corrections [83] are required for obtaining meaningful jet measurements that are consistent with the jets of partons emerging from a QCD hard collision. A factorized multi-level jet correction approach is chosen for the proper mapping of the measured jet energy with a high precision, as shown schematically in Figure 3.13. In this approach, each sub-correction is independent of the other and associated with the different detector and physics effects. Most of these corrections can be obtained from collision data. Initially, however, the corrections will be derived from MC tuned on test beam data. In the present work, the following corrections are used:

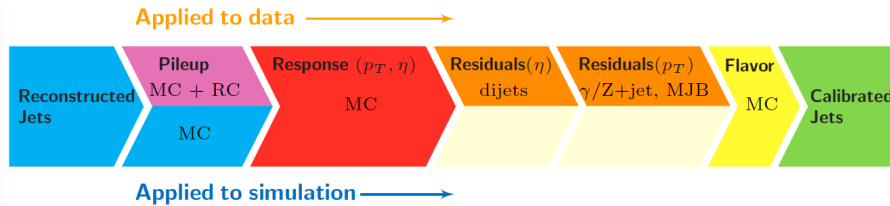


Figure 3.13: Jet correction approach.

- **Offset Level 1 (L1):** The L1 correction removes the contribution from electronic noise and pileup. The expected average contribution from pileup is estimated with ρ and the jet area [84] as inputs and is subtracted from the reconstructed jet.
- **Level 2 (L2) MC-truth correction:** The non-uniformities of the CMS Calorimeter and its non-linear response to the particles with different energies necessitate the development of corrections to the measured jet energies as a function of $|\eta|$. Relative- η corrections are applied in order to achieve a uniform jet energy response versus η . These corrections can be obtained using p_T balance in back-to-back dijet QCD events after MC based and L1 corrections. The L2 corrects the measured momentum to the true momentum, bringing the reconstructed jet in line with the generated jet.
- **Level 3 (L3) Absolute corrections:** The non-uniformities of the CMS Calorimeter and its non-linear response to the particles with different energies necessitate the development of corrections to the measured jet energies as functions of p_T , called absolute p_T correction. The Data/MC residual correction is derived to correct the response p_T dependency using $\gamma + \text{jets}$, $Z + \text{jet}$, and multijet events. The L3 corrects the measured momentum to match with the true momentum, bringing the reconstructed jet in line with the generated jet. One well measured object (for example, a jet near the center of the detector, a Z boson reconstructed from leptons) is balanced against a jet for which corrections are derived.
- **L2L3 Residual corrections:** These corrections are used to correct the remaining small difference between the true particle and measured jet energy. The Data/MC residual corrections are derived to correct for the response η and p_T dependency.

The JEC L1 and L2L3 MC-truth corrections are applied on MC, while L1, L2L3

MC-truth, and L2L3 Residuals are applied to data.

3.5.11 b/c jet Tagging

The heavy flavor b-jets (c-jets) are unique among jet flavors as they are experimentally identifiable with high efficiency and purity. The ability to identify jets with b (c) flavors is based on the properties of the b (c) hadrons and their decay. Firstly, because of the long lifetime of b (c) hadrons, of the order of 1.5 (1.1) ps, they will travel a significant distance in the transverse plane (≈ 1 cm at energy ≈ 10 –100 GeV in the laboratory frame) before decaying into several particles. Therefore, they will produce a new vertex called the secondary vertex which is displaced from the primary interaction point. The b and c quarks have a larger mass and harder fragmentation compared to the light quarks and massless gluons. As a result, the decay products of the heavy flavor hadron have, on average, a larger p_T relative to the jet axis than the other jet constituents. This causes b/c-jets to have a wider jet area and a larger charged particle multiplicity. In approximately 20% and 10% of the cases, a muon or electron is present in the decay chain of a heavy b ($b \rightarrow l$) and c ($b \rightarrow c \rightarrow l$) hadron.

Algorithms for heavy flavor jet identification use variables connected to the properties of heavy flavor hadrons present in jets resulting from the radiation and hadronization of b or c quarks. Several b/c tagging algorithms are used in CMS [85]. The b/c tagging algorithms mainly rely on the long life time, high mass, and large momentum fraction of b or c hadrons produced in b/c quark jets, as well as on the presence of soft leptons from semi-leptonic b decays.

In the present work, the deep combined secondary vertices (DEEPCSV) algorithm [67] based on deep machine learning [86] is used to identify the b/c-jets. In DEEPCSV algorithm, the significance of the SV displacement is combined with the invariant mass and impact parameter of the tracks associated with the SV in a

machine-learned multivariate analysis. The SV mass is corrected for the observed difference between its flight direction and its momentum, taking into account particles that were not reconstructed or failed to be associated with the SV. The corrected SV mass is defined as $\sqrt{(M_{\text{SV}}^2 + p^2 \sin^2 \theta)} + p \sin \theta$, where M_{SV} is the invariant mass of the tracks associated with the SV, p is the SV momentum obtained from the tracks associated with it, and θ the angle between the SV momentum and the vector pointing from the PV to the SV, which is referred to as the SV flight direction. It should be noted that the energy of a track is obtained using its momentum and assuming the $\pi^\pm$ mass [87]. Another example of a discriminating SV variable is its flight distance (significance), defined as the 2D or 3D distance between the primary and SV positions (divided by the uncertainty of the SV flight distance). At least six tracks per jet are required. When calculating the values of the track variables, the tracks are required to have an angular distance with respect to the jet axis of $\Delta R < 0.3$. Moreover, any combination of two tracks compatible with the mass of the K_S^0 meson in a window of 30 MeV is rejected. Jets that have neither a selected track nor a SV are assigned a default output discriminator value of -1 .

In the first step, various discriminating variables like SV, 2D flight distance significance, number of SV, corrected SV mass, a number of tracks from SV, etc. are combined in the algorithm. This step is called “training” of the algorithm. The training is performed on inclusive multijet events in three independent vertex categories:

- **RecoVertex:** The jet contains one or more secondary vertices.
- **PseudoVertex:** No SV is found in the jet but a set of at least two tracks with a 2D IP significance above two and a combined invariant mass at least 50 MeV away from the K_S^0 mass are found. Since there is no real SV reconstruction, no fit is performed, resulting in a reduced number of variables.
- **NoVertex:** Jets that have not been assigned to one of the previous two

categories. Only the information of the selected tracks is used.

For b-jets, the training of the deep neural network is performed using the KERAS [88] deep learning library, interfaced with the TENSOR FLOW [89] library that is used for low-level operations such as convolutions. The neural network uses four hidden layers that are fully connected, each with 100 nodes. The number of nodes in the hidden layer is different for the three different vertex categories and is set to twice the number of input variables. The discriminator values of the three vertex categories are combined with a likelihood ratio taking into account the fraction of jets of each flavor expected in multijet events. The distribution of all input variables is preprocessed to center the mean of each distribution around zero and to obtain a root-mean-square value of unity. In case a variable cannot be reconstructed, for example, because there are less than six selected tracks (or no SV), the variable values associated with the missing track or vertex are set to zero after the preprocessing. The output layer of the neural network contains five nodes corresponding to five jet flavor categories used in the training. These categories are defined according to whether the jet contains exactly one b hadron, at least two b hadrons, exactly one c hadron, and no b hadrons, at least two c hadrons and no b hadrons, or none of the aforementioned categories. Each of these categories is completely independent of the others.

For c-jets, the training of the algorithm was performed on simulated multijet events. Two trainings are performed: one for discriminating c-jets from light flavor jets (CvsL) and another one for discriminating c-jets from b-jets (CvsB). The training of the two discriminators was performed with the SCIKIT-LEARN package [90] using a gradient boosting classifier (GBC) as the implementation of the BDT.

The different algorithms provide as an output a value of discriminant for each jet. A set of working points are selected so that an average b/c tagging efficiency for a given misidentification rate is obtained. The three working points are defined as “loose” (L), “medium” (M), and “tight” (T), which correspond to thresholds on

the discriminator.

To summarize this Chapter, starting from the first Section, it explains the generation, simulation of events, and also the reconstruction of all physics objects used in the presented study. All kinds of processes emerging from pp collisions are discussed in detail.

Bibliography

- [1] M. Dobbs et al., “Les Houches Guidebook to Monte Carlo Generators for Hadron Collider Physics”, [doi:10.48550/arXiv.hep-ph/0403045](https://arxiv.org/abs/10.48550/arXiv.hep-ph/0403045).
- [2] C. P. Robert, “Monte Carlo Methods”, <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118445112.stat03876.pub2>.
- [3] J. Hammersley, “Monte Carlo methods for solving multivariable problems”, *Annals of the New York Academy of Sciences*, vol. 86 (2006) 844, [doi:10.1111/j.1749-6632.1960.tb42846.x](https://doi.org/10.1111/j.1749-6632.1960.tb42846.x).
- [4] Hildreth, M and Ivanchenko, V and Lange, D, “Upgrades for the CMS simulation”, *Journal of Physics: Conference Series*, vol. 898 (2017) p042040, [doi:10.1088/1742-6596/898/4/042040](https://doi.org/10.1088/1742-6596/898/4/042040).
- [5] S. Agostinelli et al., “GEANT4— simulation toolkit”, *Nucl. Instrum. Meth. A*, vol. 506 (2003) 250, [doi:10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8).
- [6] CMS Collaboration, “CMS Physics : Technical Design Report Volume 1: Detector Performance and Software”, CMS Technical Design Report, CERN-LHCC-2006-001, CERN, Geneva, (2006). <https://cds.cern.ch/record/922757?ln=en>.
- [7] M.Lutz, “Programming Python”, Sebastopol, CA:OReilly (1996).

-
- [8] R. Brun and F. Rademacher, “ROOT: An object oriented data analysis framework”, *Nucl. Instrum. Meth. A*, vol. 389 (1997) 81, [10.1016/S0168-9002\(97\)00048-X](#).
- [9] Buckley, Andy et al., “General-purpose event generators for LHC physics”, *Phys. Rept.*, vol. 504 (2011) 145, [doi:10.1016/j.physrep.2011.03.005](#).
- [10] J. C. Collins, D. E. Soper, and G. F. Sterman. “Factorization of Hard Processes in QCD”, In: *Adv. Ser. Direct. High Energy Phys.*, vol. 5 (1989) 094035, [doi:10.1142/9789814503266_0001](#).
- [11] M. Dittmar et al., “Parton Distributions”, (2009), [doi:10.48550/ARXIV.0901.2504](#), [arXiv:0901.2504](#).
- [12] I. Stewart, F. Tackmann, and W. Waalewijn, “Factorization at the LHC: From parton distribution functions to initial state jets”, *Phys. Rev. D*, vol. 81 (2010) 094035, [doi:10.1103/PhysRevD.81.094035](#).
- [13] V. N. Gribov and L. N. Lipatov, “Deep inelastic ep scattering in perturbation theory”, *Sov. J. Nucl. Phys.*, vol. 15 (1972) 438.
- [14] Y. L. Dokshitzer, “Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics”, *Sov. Phys. JETP*, vol. 46 (1977) 641.
- [15] G. Altarelli and G. Parisi, “Asymptotic freedom in Parton language”, *Nucl. Phys. B*, vol. 126 (1977) 298, [doi:10.1016/0550-3213\(77\)90384-4](#).
- [16] R. D. Ball et al. “Parton distributions for the LHC run II”, *JHEP*, vol. 04 (2015) 040, [doi:10.1007/JHEP04\(2015\)040](#).

-
- [17] J. Rojo et al., “The PDF4LHC report on PDFs and LHC data: results from Run I and preparation for Run II”, *Journal of Physics G: Nuclear and Particle Physics*, vol. 42 (2015) 103103, [doi:10.1088/0954-3899/42/10/103103](https://doi.org/10.1088/0954-3899/42/10/103103).
- [18] Z. Nagy and D. E. Soper, “What is a parton shower ?”, *Physical Review D*, vol. 98 (2018), [doi:10.1103/physrevd.98.014034](https://doi.org/10.1103/physrevd.98.014034).
- [19] S. Höche, “Introduction to parton-shower event generators”, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders*, (2015) 235, [doi:10.1142/9789814678766_0005](https://doi.org/10.1142/9789814678766_0005), [arXiv:1411.4085](https://arxiv.org/abs/1411.4085).
- [20] V. V. Sudakov, “Vertex parts at very high-energies in quantum electrodynamics”, *Sov. Phys. JETP*, vol. 3 (1956) 65.
- [21] S. Mrenna and P. Richardson, “Matching Matrix Elements and Parton Showers with HERWIG and PYTHIA”, *JHEP*, vol. 5 (2004) 040, [doi:10.1088/1126-6708/2004/05/040](https://doi.org/10.1088/1126-6708/2004/05/040), [arXiv:hep-ph/0312274](https://arxiv.org/abs/hep-ph/0312274).
- [22] J. Alwall et al., “Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions”, *Eur. Phys. J. C*, vol. 53 (2008) 473, [doi:10.1140/epjc/s10052-007-0490-5](https://doi.org/10.1140/epjc/s10052-007-0490-5), [arXiv:0706.2569](https://arxiv.org/abs/0706.2569).
- [23] F. Maltoni “Predictive Monte Carlo tools for the LHC. Lecture II”, in: CERN Academic Training Lectures – May 2012, Geneva, 2012.
- [24] B. Andersson, “THE LUND STRING MODEL”, in *7th European Symposium on Antiproton Interactions: From LEAR to the Collider and Beyond*, (1986) 447.
- [25] Andersson, B. and Gustafson, G. and Ingelman, G. and Sjöstrans, T., “Parton fragmentation and string Dynamics”, *Phys. Rept.*, vol. 97 (1983) 31, [doi:10.1016/0370-1573\(83\)90080-7](https://doi.org/10.1016/0370-1573(83)90080-7).

- [26] D. Amati and G. Veneziano, “Preconfinement as a property of perturbative QCD”, *Phys. Lett. B*, vol. 83 (1979) 87, [doi:10.1016/0370-2693\(79\)90896-7](https://doi.org/10.1016/0370-2693(79)90896-7).
- [27] B. R. Webber, “A QCD model for jet fragmentation including soft gluon interference”, *Nucl. Phys. B*, vol. 238 (1984) 492, [doi:10.1016/0550-3213\(84\)90333-X](https://doi.org/10.1016/0550-3213(84)90333-X).
- [28] G. Gustafson, U. Pettersson, and P. Zerwas, “Jet final states in WW pair production and colour screening in the QCD vacuum”, *Phys. Lett. B*, vol. 209 (1988) 90, [doi:10.1016/0370-2693\(88\)91836-9](https://doi.org/10.1016/0370-2693(88)91836-9).
- [29] T. Sjöstrand and V. A. Khoze, “On colour rearrangement in hadronic W^+W^- events”, *Z. Phys. C*, vol. 62 (1994) 281, [doi:10.1007/bf01560244](https://doi.org/10.1007/bf01560244), [arXiv:hep-ph/9310242](https://arxiv.org/abs/hep-ph/9310242).
- [30] S. Höche, “Introduction to parton-shower event generators”, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders*, (2015) 235, [doi:10.48550/arXiv.1411.4085](https://doi.org/10.48550/arXiv.1411.4085) [arXiv:1411.4085](https://arxiv.org/abs/1411.4085).
- [31] T. Sjöstrand, “The PYTHIA event generator: Past, present and future”, *Comput. Phys. Commun.*, vol. 246 (2020) 106910, [doi:10.1016/j.cpc.2019.106910](https://doi.org/10.1016/j.cpc.2019.106910).
- [32] M. Bähr et al., “Herwig++ physics and manual”, *Eur. Phys. J. C*, vol. 58 (2008) 639, [doi:10.1140/epjc/s10052-008-0798-9](https://doi.org/10.1140/epjc/s10052-008-0798-9), [arXiv:0803.0883](https://arxiv.org/abs/0803.0883).
- [33] J. Bellm et al., “Herwig 7.0/Herwig++ 3.0 release note”, *Eur. Phys. J. C*, vol. 76 (2016) 196, [doi:10.1140/epjc/s10052-016-4018-8](https://doi.org/10.1140/epjc/s10052-016-4018-8), [arXiv:1512.01178](https://arxiv.org/abs/1512.01178).
- [34] J. Bellm et al., “Herwig 7.1 Release Note”, (2017), [arXiv:1705.06919](https://arxiv.org/abs/1705.06919).
- [35] T. Gleisberg et al., “Event generation with SHERPA 1.1”, *JHEP*, vol. 02 (2009) 007, [doi:10.1088/1126-6708/2009/02/007](https://doi.org/10.1088/1126-6708/2009/02/007), [arXiv:0811.4622](https://arxiv.org/abs/0811.4622).

-
- [36] J. Alwall et al., “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”, *JHEP*, vol. 07 (2014) 079, [doi:10.1007/JHEP07\(2014\)079](https://doi.org/10.1007/JHEP07(2014)079).
- [37] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”, *JHEP*, vol. 11 (2007) 070, [doi:10.1088/1126-6708/2007/11/070](https://doi.org/10.1088/1126-6708/2007/11/070), [arXiv:0709.2092](https://arxiv.org/abs/0709.2092).
- [38] Sjöstrand, Torbjörn and Mrenna, Stephen and Skands, Peter, “PYTHIA 6.4 physics and manual”, *JHEP*, vol. 05 (2006) 026, [doi:10.1088/1126-6708/2006/05/026](https://doi.org/10.1088/1126-6708/2006/05/026).
- [39] Sjöstrand, Torbjörn and Mrenna, Stephen and Skands, Peter, “A brief introduction to PYTHIA 8.1”, *Comp. Phys. Comm.*, vol. 178 (2008) 852, [doi:10.1016/j.cpc.2008.01.036](https://doi.org/10.1016/j.cpc.2008.01.036).
- [40] T. Sjöstrand et al., “An introduction to PYTHIA 8.2”, *Comput. Phys. Commun.*, vol. 191 (2015) 159, [doi:10.1016/j.cpc.2015.01.024](https://doi.org/10.1016/j.cpc.2015.01.024).
- [41] T. Sjöstrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 physics and manual”, *JHEP*, vol. 05 (2006) 026, [doi:10.1088/1126-6708/2006/05/026](https://doi.org/10.1088/1126-6708/2006/05/026).
- [42] CMS Collaboration, “Event generator tunes obtained from underlying event and multiparton scattering measurements”, *Eur. Phys. J. C*, vol. 76 (2016) 155, [doi:10.1140/epjc/s10052-016-3988-x](https://doi.org/10.1140/epjc/s10052-016-3988-x), [arXiv:1512.00815](https://arxiv.org/abs/1512.00815).
- [43] CMS Collaboration, “Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements”, *Eur. Phys. J. C*, vol. 80 (2020) 4, [doi:10.1140/epjc/s10052-019-7499-4](https://doi.org/10.1140/epjc/s10052-019-7499-4), [arXiv:1903.12179](https://arxiv.org/abs/1903.12179).
- [44] NNPDF Collaboration, “Parton distributions with LHC data”, *Nucl. Phys. B*, vol. 867 (2013) 244, [doi:10.1016/j.nuclphysb.2012.10.003](https://doi.org/10.1016/j.nuclphysb.2012.10.003).

-
- [45] F. Krauss, R. Kuhn, and G. Soff, “AMEGIC++ 1.0: A Matrix element generator in C++”, *JHEP*, vol. 02 (2002) 044, [doi:10.1088/1126-6708/2002/02/044](https://doi.org/10.1088/1126-6708/2002/02/044), [arXiv:hep-ph/0109036](https://arxiv.org/abs/hep-ph/0109036).
- [46] T. Gleisberg and S. Höche, “Comix, a new matrix element generator”, *JHEP*, vol. 12 (2008) 039, [doi:10.1088/1126-6708/2008/12/039](https://doi.org/10.1088/1126-6708/2008/12/039).
- [47] F. Cascioli, P. Maierhofer, and S. Pozzorini, “Scattering Amplitudes with Open Loops”, *Phys. Rev. Lett.*, vol. 108 (2012) 111601, [doi:10.1103/PhysRevLett.108.111601](https://doi.org/10.1103/PhysRevLett.108.111601), [arXiv:1111.5206](https://arxiv.org/abs/1111.5206).
- [48] E. Bothmann et al., “Event generation with Sherpa 2.2”, *SciPost Phys.*, vol. 7 (2019) 034, [doi:10.21468/SciPostPhys.7.3.034](https://doi.org/10.21468/SciPostPhys.7.3.034), [arXiv:1905.09127](https://arxiv.org/abs/1905.09127).
- [49] S. Alioli, P. Nason, C. Oleari, and E. Re, “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”, *JHEP*, vol. 06 (2010) 043, [doi:10.1007/JHEP06\(2010\)043](https://doi.org/10.1007/JHEP06(2010)043), [arXiv:1002.2581](https://arxiv.org/abs/1002.2581).
- [50] J. Alwall et al., “MadGraph 5 : Going Beyond”, *JHEP*, vol. 06 (2011) 128, [doi:10.48550/arXiv.1106.0522](https://doi.org/10.48550/arXiv.1106.0522).
- [51] F. Maltoni and T. Stelzer, “MadEvent: automatic event generation with MadGraph”, *JHEP*, vol. 02 (2003) 027, [doi:10.1088/1126-6708/2003/02/027](https://doi.org/10.1088/1126-6708/2003/02/027), [arXiv:hep-ph/0208156](https://arxiv.org/abs/hep-ph/0208156).
- [52] R. Frederix, S. Frixione, F. Maltoni, and T. Stelzer, “Automation of next-to-leading order computations in QCD: The FKS subtraction”, *JHEP*, vol. 10 (2009) 003, [doi:10.1088/1126-6708/2009/10/003](https://doi.org/10.1088/1126-6708/2009/10/003), [arXiv:0908.4272](https://arxiv.org/abs/0908.4272).
- [53] S. Frixione, Z. Kunszt, and A. Signer, “Three jet cross sections to next-to-leading order”, *Nucl. Phys. B*, vol. 467 (1996) 399, [doi:10.1016/0550-3213\(96\)00110-1](https://doi.org/10.1016/0550-3213(96)00110-1), [arXiv:hep-ph/9512328](https://arxiv.org/abs/hep-ph/9512328).

- [54] S. Frixione, “A General approach to jet cross-sections in QCD”, *Nucl. Phys. B*, vol. 507 (1997) 295, [doi:10.1016/S0550-3213\(97\)00574-9](https://doi.org/10.1016/S0550-3213(97)00574-9), [arXiv:hep-ph/9706545](https://arxiv.org/abs/hep-ph/9706545).
- [55] M. L. Mangano, M. Moretti, and R. Pittau, “Multijet matrix elements and shower evolution in hadronic collisions: $Wb\bar{b} + n$ jets as a case study”, *Nucl. Phys. B*, vol. 632 (2002) 343, [doi:10.1016/S0550-3213\(02\)00249-3](https://doi.org/10.1016/S0550-3213(02)00249-3), [arXiv:hep-ph/0108069](https://arxiv.org/abs/hep-ph/0108069).
- [56] R. Frederix and S. Frixione, “Merging meets matching in MC@NLO”, *JHEP* vol. 12 (2012) 61, [doi:10.1007/jhep12\(2012\)061](https://doi.org/10.1007/jhep12(2012)061), [arXiv:1209.6215](https://arxiv.org/abs/1209.6215).
- [57] D J Lange and M Hildreth and V N Ivantchenko and I Osborne and, “Upgrades for the CMS simulation”, *Journal of Physics: Conference Series*, vol. 608 (2015) 012056, [doi:10.1088/1742-6596/608/1/012056](https://doi.org/10.1088/1742-6596/608/1/012056).
- [58] M Hildreth and V N Ivanchenko and D J Lange and M J Kortelainen, “CMS Full Simulation for Run-2”, *Journal of Physics: Conference Series*, vol. 664 (2015), [doi:10.1088/1742-6596/664/7/072022](https://doi.org/10.1088/1742-6596/664/7/072022).
- [59] Worldwide LHC Computing Grid Collaboration, “Memorandum of Understanding for Collaboration in the Deployment and Exploitation of the Worldwide LHC Computing Grid”. CERN-C-RRB-2005-01/Rev.1, Version 29 March 2011, <http://wlcg.web.cern.ch/collaboration/mou>.
- [60] CMS Collaboration, “Particle-flow reconstruction and global event description with the CMS detector”, *JINST*, vol. 12 (2017) P10003, [doi:10.1088/1748-0221/12/10/P10003](https://doi.org/10.1088/1748-0221/12/10/P10003).
- [61] CMS Collaboration, “The Z and B”, CMS Physics Analysis Summary, PAS-SMP-20-015, (2021). <https://cms.cern/news/z-and-b>.

-
- [62] T. Miao, N. Leioatts, H. Wenzel, and F. Yumiceva, “Beam Position Determination using Tracks”, CMS Technical Report, CMS-NOTE-2007-021, CERN, Geneva, (2007). <https://cds.cern.ch/record/1061285>.
- [63] P. Billoir and S. Qian, “Simultaneous pattern recognition and track fitting by the Kalman filtering method” *Nucl. Instrum. Methods Phys. Res. A*, vol. 294 (1990) 219, [doi:10.1016/0168-9002\(90\)91835-Y](https://doi.org/10.1016/0168-9002(90)91835-Y).
- [64] R. Fruhwirth, “Application of Kalman filtering to track and vertex fitting”, *Nucl. Instrum. Meth. A*, vol. 262 (1987) 444, [doi:10.1016/0168-9002\(87\)90887-4](https://doi.org/10.1016/0168-9002(87)90887-4).
- [65] CMS Collaboration, “Tracking at High Level Trigger in CMS”, *Nucl. Part. Phys. Proc.*, vol. 273 (2016) 2494, [doi:10.1016/j.nuclphysbps.2015.09.436](https://doi.org/10.1016/j.nuclphysbps.2015.09.436).
- [66] K. Rose, “Deterministic annealing for clustering, compression, classification, regression, and related optimization problems”, in *Proceedings of the IEEE*, vol. 86 (1998) 2210, [doi: 10.1109/5.726788](https://doi.org/10.1109/5.726788).
- [67] CMS Collaboration, “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV”, *JINST*, vol. 13 (2018) P05011, <https://doi.org/10.48550/arXiv.1712.07158>, [arXiv:1712.07158](https://arxiv.org/abs/1712.07158).
- [68] CMS Collaboration, “Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV”, *JINST*, vol. 10 (2015) P06005, [doi:10.48550/arXiv.1502.02701](https://doi.org/10.48550/arXiv.1502.02701).
- [69] Adam, Wolfgang and Fruhwirth, R and Strandlie, Are and Todor, T, “Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC”, CMS Technical Report, CMS-NOTE-2005-001, (2005). <https://cds.cern.ch/record/815410>.

- [70] Baffioni, S. and Charlot, C. and Ferri, F. and Futyan, D. and Meridiani, P. and Puljak, I. and Rovelli, C. and Salerno, R. and Sirois, Y., “Electron reconstruction in CMS”, *Eur. Phys. J. C*, vol. 49 (2007) 1099, [doi:10.1140/epjc/s10052-006-0175-5](https://doi.org/10.1140/epjc/s10052-006-0175-5).
- [71] CMS Collaboration, “Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JINST*, vol. 13 (2018) P06015, [doi:10.48550/arXiv.1804.04528](https://doi.org/10.48550/arXiv.1804.04528).
- [72] M. Cacciari and G. P. Salam, “Pileup subtraction using jet areas”, *Phys. Lett. B*, vol. 659 (2008) 119, [doi:10.48550/arXiv.0707.1378](https://doi.org/10.48550/arXiv.0707.1378).
- [73] Saranya Ghosh et al., “Performance of MET reconstruction and pileup mitigation techniques in CMS”, *Nucl. and Part. Phys. Proc.*, vol. 273 (2016) 2512, [doi:10.1016/j.nuclphysbps.2015.09.442](https://doi.org/10.1016/j.nuclphysbps.2015.09.442).
- [74] CMS Collaboration, “Performance of Missing Transverse Momentum Reconstruction Algorithms in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV with the CMS Detector”, *Eur. Phys. J. C* vol. 77 (2017) 241 [doi:10.1140/epjc/s10052-017-4780-2](https://doi.org/10.1140/epjc/s10052-017-4780-2).
- [75] G. P. Salam, “Towards Jetography”, *Eur. Phys. J.*, vol. 67 (2010) 637, [doi:10.48550/arXiv.0906.1833](https://doi.org/10.48550/arXiv.0906.1833), [arXiv:0906.1833](https://arxiv.org/abs/0906.1833).
- [76] Carli, Tancredi and Rabbertz, Klaus and Schumann, Steffen, “Studies of Quantum Chromodynamics at the LHC”, *The Large Hadron Collider: Harvest of Run 1*, (2015) 139, [doi:10.1007/978-3-319-15001-7_5](https://doi.org/10.1007/978-3-319-15001-7_5), [arXiv:1506.03239](https://arxiv.org/abs/1506.03239).
- [77] The CMS Collaboration, “Jet Performance in pp collisions at $\sqrt{s} = 7$ TeV”, *CMS-PAS-JME-10-03*, (2010), <https://cds.cern.ch/record/1279362>.

- [78] The CMS Collaboration, “Jet Plus Tracks Algorithm for Calorimeter Jet Energy Corrections in CMS”, CMS-PAS-JME-09-002, (2009), <https://cds.cern.ch/record/1190234>.
- [79] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet user manual”, *Eur. Phys. J. C*, vol. 72 (2012) 1896, [doi:10.48550/arXiv.1111.6097](https://doi.org/10.48550/arXiv.1111.6097).
- [80] The CMS Collaboration, “A Cambridge-Aachen (C-A) based Jet Algorithm forboosted top-jet tagging”, CMS Technical Report, CMS-PAS-JME-09-001, CERN, Geneva (2009). <https://cds.cern.ch/record/1194489>.
- [81] S. Catani et al., “Longitudinally invariant k_t clustering algorithms for hadron hadron collisions”, *Nucl. Phys. B*, vol. 406 (1993) 187, [doi:10.1016/0550-3213\(93\)90166-M](https://doi.org/10.1016/0550-3213(93)90166-M).
- [82] Matteo Cacciari, Gavin P. Salam, Gregory Soyez, “The anti- k_t jet clustering algorithm”, *JHEP*, vol. 4 (2008) 063, [doi:10.1088/1126-6708/2008/04/063](https://doi.org/10.1088/1126-6708/2008/04/063).
- [83] CMS Collaboration, “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV”, *JINST*, vol. 12 (2017) P02014, [doi:10.48550/arXiv.1607.03663](https://doi.org/10.48550/arXiv.1607.03663).
- [84] M. Cacciari, G. P. Salam, and G. Soyez, “The catchment area of jets”, *JHEP*, vol. 04 (2008) 005, [doi:10.1088/1126-6708/2008/04/005](https://doi.org/10.1088/1126-6708/2008/04/005), [arXiv:0802.1188](https://arxiv.org/abs/0802.1188).
- [85] CMS Collaboration, “Identification of b-quark jets with the CMS experiment”, *JINST*, vol. 8 (2013) P04013, [doi:10.1088/1748-0221/8/04/p04013](https://doi.org/10.1088/1748-0221/8/04/p04013).
- [86] D. Guest et al., “Jet flavor classification in high-energy physics with deep neural networks”, *Phys. Rev. D*, vol. 94 (2016) 112002, [doi:10.1103/PhysRevD.94.112002](https://doi.org/10.1103/PhysRevD.94.112002).

- [87] Particle Data Group Collaboration, “Review of Particle physics”, *Chin. Phys. C*, vol 40 (2016), 100001, [doi:10.1088/1674-1137/40/10/100001](https://doi.org/10.1088/1674-1137/40/10/100001).
- [88] F. Chollet, “KERAS”, 2015, <https://github.com/fchollet/keras>.
- [89] M. Abadi et al., “TensorFlow: Large-scale machine learning on heterogeneous systems”, (2015). Software available from tensorflow.org.
- [90] F. Pedregosa et al., “Scikit-learn: Machine learning in Python”, *J. Machine Learning Research*, vol. 12 (2011) 2825, [doi:10.48550/arXiv.1201.0490](https://doi.org/10.48550/arXiv.1201.0490).

Chapter 4

Assembly and testing of Gas Electron Multiplier

In this Chapter, the assembly and quality control (QC) testing of Gas Electron Multiplier (GEM) detectors in which I participated at the Panjab University, Chandigarh, India, and CERN sites will be discussed in detail.

As mentioned in Chapter [2.3.2](#), the CMS experiment will continue its current physics programs during the HL-LHC period, hence its sensitivity for electroweak and TeV scale searches must be maintained. Muons are crucial to research on a wide range of physics topics, from the “new physics” to the decay of the SM Higgs boson into ZZ or ZZ^* , where Z bosons further decay into 4 leptons, called the “golden channel” [\[1\]](#). The muon detectors play a central role in CMS, as stressed already by its name, Compact Muon Solenoid. The CMS muon detector was designed to be a highly hermetic and redundant system consisting of three types of gaseous ionization detectors: CSCs, DTs, and RPCs [\[2\]](#). The details of these detectors have been already discussed in Chapter [2.3.2](#). The DTs (CSCs) in the barrel (endcap) provide precision measurements and L1 triggering for acceptances up to $|\eta| < 1.2$ ($0.9 < |\eta| < 2.4$). RPCs provide redundant trigger and coarse position measurement in the barrel and endcap regions, but these are not used after $|\eta| > 1.9$ since they

cannot sustain the high background particle rates in this region. Furthermore, due to decrease in the magnetic field and higher background rates with increasing η , the contribution to the trigger rate after $|\eta| > 1.9$ is particularly large and difficult to control.

The forward region is particularly challenging for muon detection: on the one hand, relatively light particles to be studied or searched are produced predominantly with high pseudorapidities; on the other hand, the backgrounds in this region are the highest. Therefore, the forward region must be completed with missing muon detectors to enhance the redundancy and resolve the track reconstruction ambiguities. The first step in this direction is the installation of an additional set of muon detectors based on GEM technology [3] in the first muon endcap disks during LS2. That particular muon detector is known as GE1/1, where the letter G indicates the GEM, the letter E indicates this is an endcap muon station, the first “1” indicates that it is part of the first muon station encountered by particles from the interaction point, and the second “1” indicates that it is the first ring of muon detectors going outward in radius from the beam line. The fundamental element of a GE1/1 detector is the single chamber.

Many research institutes and industrial companies from Belgium, China, Bulgaria, Colombia, Egypt, Finland, France, Germany, Hungary, India, Italy, Korea, Pakistan, Qatar, and the United States are involved in several CMS GEM projects. India contributed to the CMS GEM upgrading project for the assembly and testing of the GE1/1 chambers which was carried out at Panjab University (PU), Chandigarh, in collaboration with Delhi University, Delhi, and Bhabha Atomic Research Center, Mumbai. The production of GE1/1 chambers was started at CERN in April 2017 and at other production sites in Autumn 2017. Total 144 GE1/1 chambers are required to be fabricated and tested for the GE1/1 upgrade project.

4.1 Motivations for the GE1/1 muon detector upgrade

A quadrant cross section of the CMS with the muon detector station where GE1/1 is installed in $1.6 < |\eta| < 2.2$ region during LS2 is shown in Figure 4.1. Because of the earlier plan of forward RPCs at this location, there is a compact space available for a new muon detector/chamber. The GEM technology is an excellent choice for this region, due to its thin profile and ability to operate at particle fluxes much higher than those expected in the forward area under HL-LHC conditions. The maximum hit rate at the first muon station is expected to be around 5 kHz/cm^2 during HL-LHC, whereas GE1/1 can sustain 10 kHz/cm^2 for 20 years of HL-LHC operation.

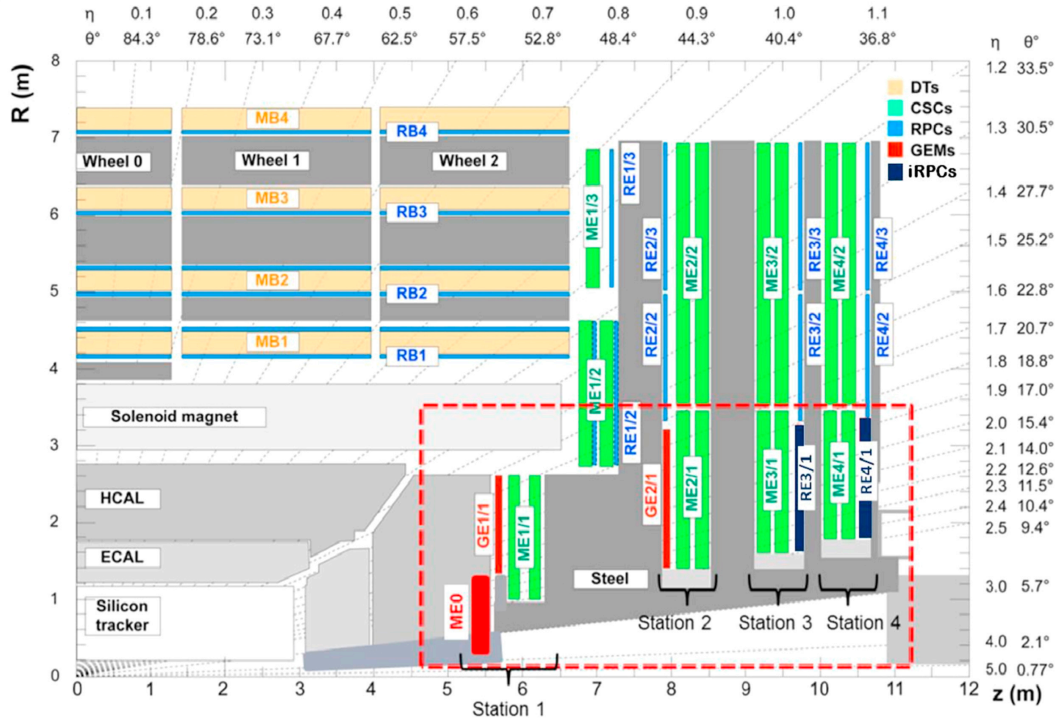


Figure 4.1: A quadrant of the muon system, shows DT chambers (yellow), RPC (light blue), and CSC (green). The new GEM chamber stations (ME0, GE1/1, and GE2/1) are indicated in red [4].

The installation of the GE1/1 chamber increase the path length traversed by

muons within the first muon station and muons bend more there than at succeeding stations. As a result, the information of the bending angle between the GE1/1 and CSC and increased path length significantly improved the muon momentum resolution and reduced the trigger rate of L1 standalone muons [5]. Figure 4.2 shows the single muon trigger rate curves before and after the GE1/1 upgrade for the region $1.6 < |\eta| < 2.2$ at a luminosity of $2 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ and at a constant efficiency of 94% for the first endcap muon station (MS1/1). With the upgrade, the L1 muon trigger thresholds can be maintained at low p_T values, hence the efficiency for capturing interesting physics processes featuring soft leptons will become high. Maintaining low muon trigger p_T thresholds is important for a broad spectrum of physics studies ranging from the Higgs sector measurements to new physics searches [6–11].

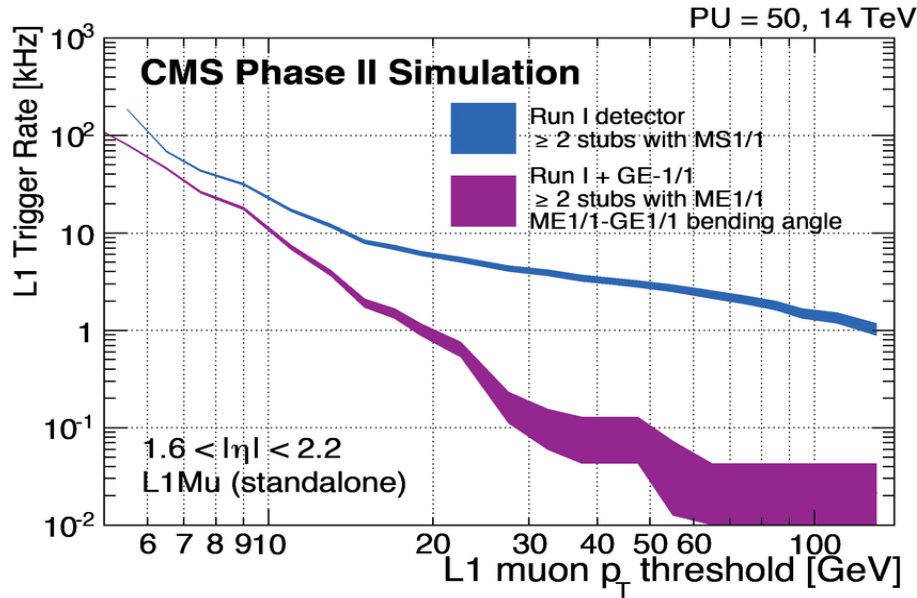


Figure 4.2: Comparison of L1 muon trigger rates as a function of trigger threshold with CSC only (blue curve) and with the combination of CSC and GEM (purple curve) trigger information [4].

The greatest benefit of the early installation of the GE1/1 muon station is that it improves the L1 muon trigger during LHC running before the installation of a new silicon tracker and its associated track trigger [12] in LS3. The L1 muon trigger will be used with the new silicon tracker and track trigger in LS3 to make

a “combined muon trigger”. The next Section will describe the summary of the performance criteria for the GE1/1 chamber.

4.2 Summary of performance requirements for the GE1/1 chamber

The GE1/1 chamber needs to perform certain requirements in terms of detection. These requirements are:

- **Maximum geometry acceptance within a particular CMS coverage area:** This is because maximum acceptance allows maximum physics yield.
- **Atleast 10 kHz/cm² rate capability:** For HL-LHC ($5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$), the maximum expected hit rate with GE1/1 acceptance is roughly 5 kHz/cm², therefore multiplying by a safety factor of two, requires a hit rate capability of 10 kHz/cm².
- **The detection efficiency of a single chamber for minimum ionizing particles is equal to or greater than 97%:** When the signals from the two chambers are merged as a logical OR, the efficiency of a superchamber is greater than 99.9%.
- **Angular resolution of 300 μrad or better on difference $\Delta\phi = \phi_{\text{GE1/1}} - \phi_{\text{ME1/1}}$:** This allows the trigger to differentiate between high and low p_T muons by combining the angular position of the muon detected in GE1/1 and ME1/1. At the outer radius ($r=2.6 \text{ m}$) of GE1/1 chambers, the azimuthal precision of 300 μrad corresponds to a 0.8 mm resolution in the azimuthal $\hat{\phi}$ direction.
- **Timing resolution of 10 ns or better for a single chamber:** This timing is sufficient to reliably match GE1/1 hits to ME1/1 stubs when the LHC is

running with a 25 ns bunch crossing time.

- **Gain uniformity of 15% or better inside a single chamber or between chambers:** This gain uniformity ensures that there is no bias in trigger or reconstruction due to geometry.
- **No gain loss due to aging after 200 mC/cm² integrated charge:** It is expected that the total charge to be integrated for a GE1/1 chamber at the highest $|\eta|$ over 20 years of operation will be about 100 mC/cm². The stated requirement of 200 mC/cm² includes an additional safety factor of two.

Furthermore, certain technological limits and criteria must be considered when building the GE1/1 chambers, as follows:

- The chamber must be able to operate only with gases that have a low global warming impact.
- The material budget must be small enough that multiple scattering within the GE1/1 does not interfere with the muon track measurement in the GE1/1-CSC trigger.
- The chambers must be designed in such a way that a superchamber is less than 10 cm thick and can easily fit into the available slot in the muon endcap nose.

4.3 GE1/1: construction, principle, and working

The GE1/1 detectors are based on the principle that radiation (particle) detection occurs when it passes through a gas mixture that is enclosed between two electrodes with an electric field applied between them. The GE1/1 chamber is made up of a gas volume that holds a stack of three large area trapezoidal-shaped GEM foils between

a drift electrode and a readout board. A standard gas mixture of Ar/CO₂ 70:30 is used for operating a GE1/1 chamber. Since working under a purely noble gas medium can cause avalanche creation beyond the limits which leads to sparks and cause permanent damage to the chamber, hence a smaller amount of a quenching gas in the medium is required. The carbon-dioxide (CO₂) works as a quencher in the gas mixture because electrons are attracted and attached to its molecules due to its electronegativity. The foils are made of Kapton, which is 50 μm thick and has a dielectric constant of 3.5; it is coated on both sides with 5 μm conductive layers of copper. Kapton is a Polyamide film that is an excellent insulator and remains stable over a wide range of temperatures. The GEM holes are truncated double cones with larger (outer) diameters of roughly 70 μm and smaller (inner) diameters around of 50 μm , spaced in a hexagonal pattern with a pitch of 140 μm as shown in Figure 4.3 (left). The voltage applied between the two copper-clad surfaces of a foil produces a strong electric field in the GEM hole that amplifies the signal generated during the primary and secondary ionization processes as shown in Figure 4.3 (middle). The GE1/1 chamber has a drift region of 3 mm between the drift cathode (at negative potential) and the first GEM foil (GEM1), transfer regions of 1 mm and 2 mm between the second and third GEM foils (GEM2 and GEM3), and a 1 mm induction region. These regions are created by placing spacers at the edges of the adjacent GEM foils. The top surface of the GEM foil facing toward the drift board is segmented into sectors, whereas the bottom surface of the GEM foil facing toward the readout board is a single continuous conductor. A single high voltage (HV) divider, consisting of a chain of resistors, is used to supply appropriate voltages to the drift plane and three GEM foils. The drift region is where the primary electrons are generated and they move towards the first GEM foil due to the field present in this region. This drift field depends on the potential difference between the cathode and the top metal of the first GEM foil. The electric field generated inside the holes is ~ 100 times stronger than the field present in the

drift region, where each hole works as an individual proportional counter. Once electrons begin to experience the very intense electric field in the holes, they gain enough kinetic energy to produce secondary ionization in the gas. The first transfer gap was chosen to minimize the charge released after the first GEM foil. Electrons were further created and accelerated towards the second GEM foil due to the transfer field present between the bottom of the first GEM foil and the top of the second GEM foil. This process is repeated for the succeeding foil with asymmetric potential across the foils and gaps. Finally, all of the charges generated by the avalanche are deposited on the anode plate by the induction field, which is present between the last GEM foil and the ground anode. The induction gap was intended to provide an induction field of up to 5 kV/cm without significantly raising the potential of the entire structure. The GE1/1 chamber has advantage of operating at low voltage to avoid any electrical breakdown problem. Although it provides a high total charge gain factor (up to 10^5) since the gains of the individual foils (≥ 20) multiply to produce the total gain. A schematic view of this operation principle is shown in Figure 4.3 (right), which also defines the drift region, two transfer regions, and induction region inside the GE1/1 chamber.

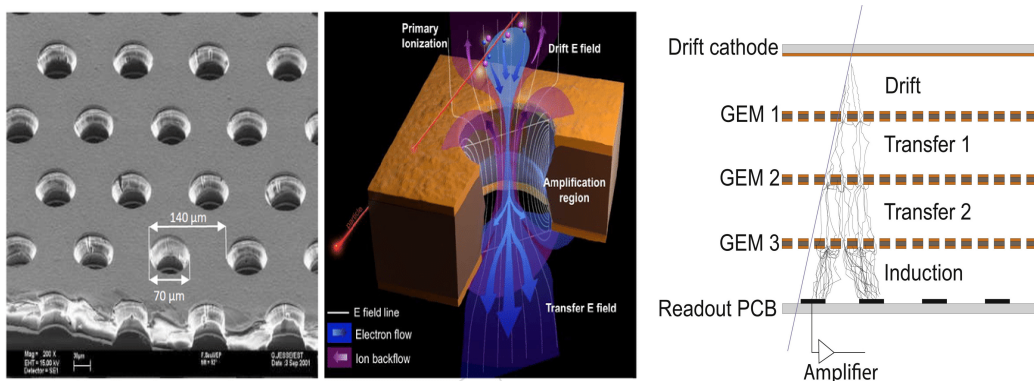


Figure 4.3: Scanning Electron Microscope (SEM) picture of a GEM foil [3] (left) and schematic view of the electric field lines (white), electron flow (blue), and ion flow (purple) through a bi-conical GEM hole (middle), and sketch of a GE1/1 chamber (right) [13].

The electronics component of readout processes GEM signals, which are separated into two main sections, namely On-Detector (frontend) and Off-Detector

(backend) [14], as shown in Figure 4.4. The readout board is segmented into three ϕ portions along its longitudinal axis and eight η sections or sectors perpendicular to its longitudinal axis resulting in 24 individual chamber segments. Each sector has an area of about 100 cm^2 . Each segment is then further subdivided into 128 copper strips with a trapezoidal shape, which is read by the Very Forward Atlas and Totem (VFAT3) application-specific integrated circuits (ASICs) [15]. The VFAT3 has 128 charge-sensitive analog channels, which can be programmed to respond to positive or negative signals. The VFATs have a charge-sensitive preamplifier, shaper, and constant fraction discriminator for each channel. The avalanche charge induced on the strips is converted into binary information by the VFAT. Once the digital information about the hits has been elaborated, it takes two paths: the full-granularity data path and the half-granularity trigger path. The digital signals are sent from the VFATs to the OptoHybrid (OH) board through the GEM Electronics Board (GEB).

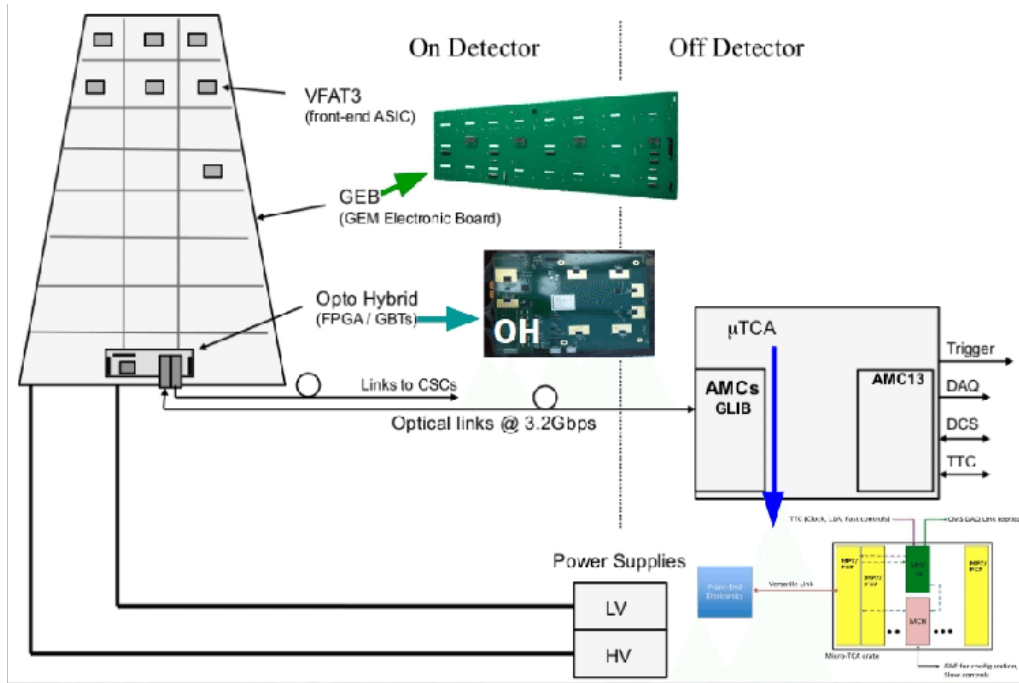


Figure 4.4: The GEM electronics readout system.

An OH board contains the GigaBit Transceiver (GBT) chip set, a Field Pro-

grammable Gate Array (FPGA), as well as optical receivers and transmitters to provide the link to the Off-Detector region. After a final signal processing in the OH, the information is sent to the GEM backend and to the CSC trigger motherboard (TMB) via optical links. There are two optical paths to the OH. The first is bidirectional and runs between the OH and micro-Telecommunications Computing Architecture (μ -TCA) [16] crates are located in the counting room at P5. The return path is used for the VFAT3 tracking and triggering data packets as well as to return slow control data. The second path is unidirectional and takes the VFAT3 fixed latency trigger data from the GEM system to the Cathode Strip Chamber (CSC) system.

The operating voltages of the different GE1/1 electronics components are not uniform; they vary from 1 V to 2.55 V. The FEST (Radiation tolerant DC-DC converter) mounted locally on the GEB, provided the required operating voltage to the pieces of electronics transforming the 8 V given at the input with efficiencies above 70%. A total of nine FEASTs are installed on the chambers, among them four FEASTs providing 1.3 V to a domain of six VFATs and the other five FEASTs providing the power to the OH. Each FEAST is thermally coupled with the cooling plate with a non-conductive thermal paste, which carries away the heat produced in the voltage conversion.

Figure 4.5 shows the exploded view of the GE1/1 chamber, including thermal screens, covers, readout boards, OH, and GEM foils. In the GE1/1 muon system, a pair of such GE1/1 chambers is combined to form a “superchamber” that provides two measurement planes in the muon endcap that complement the existing ME1/1 chambers and maximizes the detection efficiency. Each superchamber covers a $\approx 10^\circ$, so 72 superchambers or 36 (18 long and 18 short) in each endcap are required as shown in Figure 4.6. The basic parameters and specifications for the construction of the GE1/1 chambers and their operation in CMS are summarized in Table 4.1.

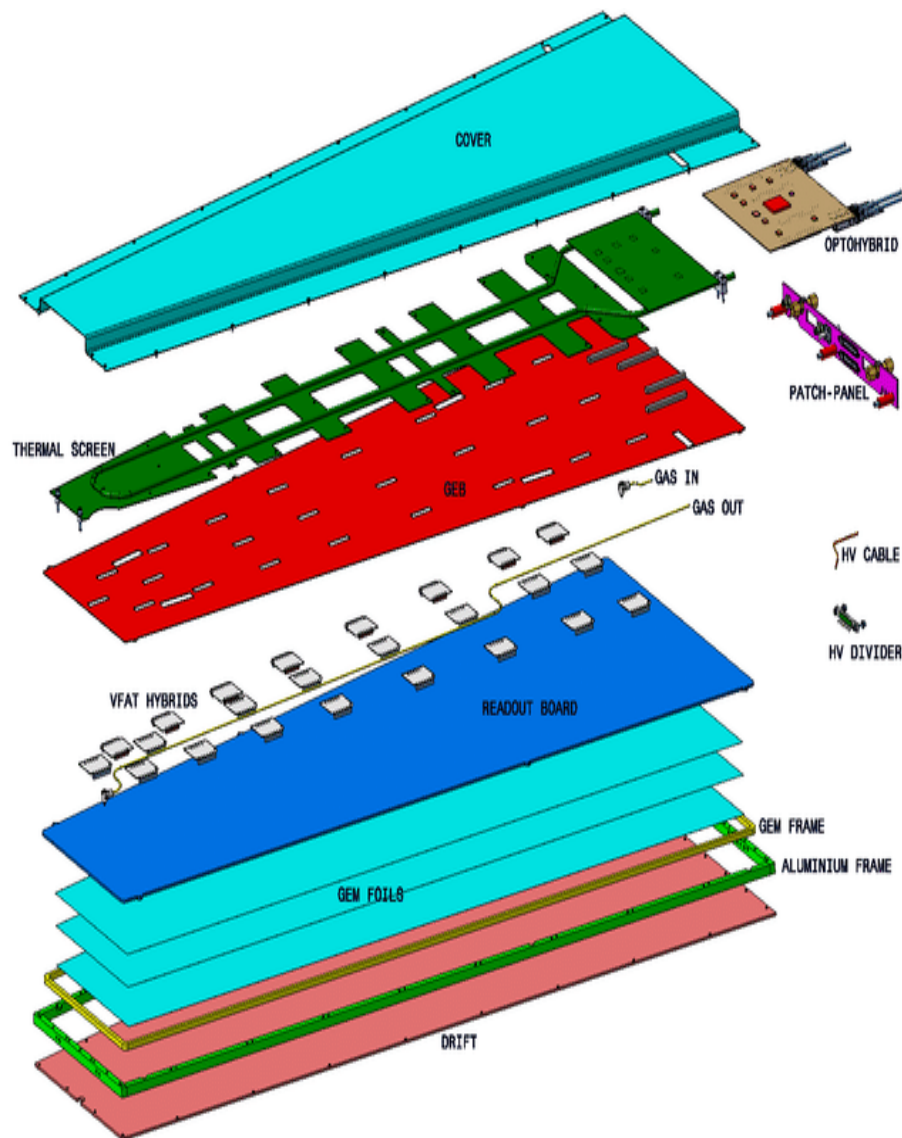


Figure 4.5: Exploded view of GE1/1 chamber [17].

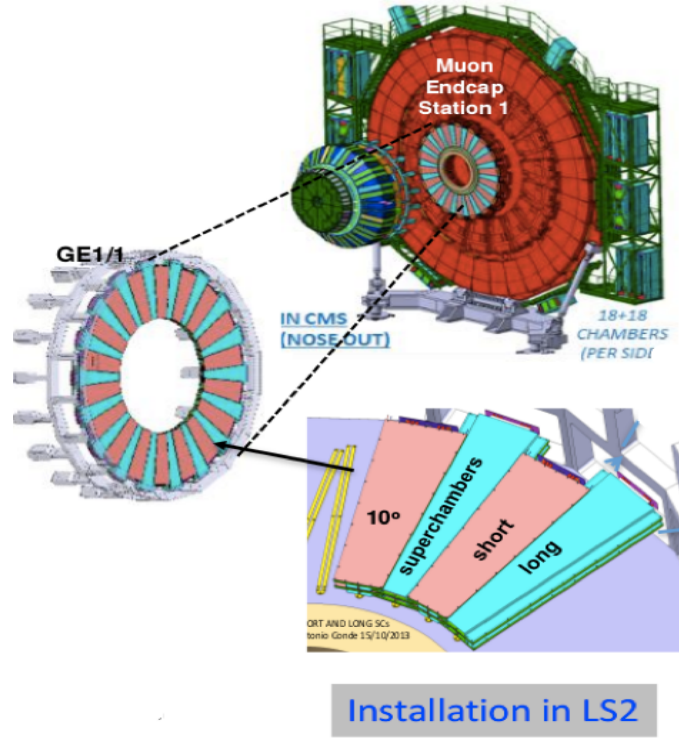


Figure 4.6: First CMS muon endcap station where the inner ring is equipped with 18 long (pink) and 18 short (light blue) GE1/1 chambers.

Table 4.1: Main specifications and parameters for the design and operation of the GE1/1 chambers [17].

Specification / Parameter	GE1/1
Detector technology	Gaseous detector; micro-pattern gas detector (MPGD)
Charge amplification element	GEM foil (triple, cascaded, tensioned at ≈ 5 N/cm)
Number of chambers in overall system	144 (72 in each endcap)
Chamber shape (active readout area)	Trapezoidal; opening angle 10.15°
Active area overlap in adjacent chambers	2.6 mrad (corresponds to 5.7 readout strip pitches)
Short chamber dimensions (active vol.)	L: 106.1 cm (center line), W: (23.1 - 42.0) cm, D: 0.7 cm
Long chamber dimensions (active vol.)	L: 120.9 cm (center line), W: (23.1 - 44.6) cm, D: 0.7 cm
Total chamber thickness	D: 3.5 cm
Active readout area	0.345 m^2 (short ch.); 0.409 m^2 (long ch.)
Active chamber volume	2.6 liters (short ch.); 3 liters (long ch.)
Radial distance from beam line	130.2 cm (at inner edge of active readout area)
Geometric acceptance in η	1.61 - 2.18 (short ch.); 1.55 - 2.18 (long ch.)
Signal readout structure	Truly radial readout strips
Readout strip dimensions	$230 \mu\text{rad}$ angular strip width; $463 \mu\text{rad}$ angular pitch
Number of η -segments in readout	8
Number of readout strips per η -segment	384
Number of readout strips per chamber	3,072
Counting gas mixtures	Ar/CO ₂ 70:30 or Ar/CO ₂ /CF ₄ 45:15:40
Nominal operational gas flow	1 chamber volume per hour
Number of gas inlets	1
Number of gas outlets	1
Nominal HV applied to drift electrode	3200 V (Ar/CO ₂); 4000 V (Ar/CO ₂ /CF ₄)
Nominal operational gas gain	$1-2 \times 10^4$
Demonstrated rate capability	100 MHz/cm ²

Eight out of 144 GE1/1 chambers were assembled at PU, and all chambers passed the QC test at both PU and CERN before being installed in the CMS chamber during LS2. The assembly technique and QC tests performed at the PU and CERN sites, in which I participated, will be discussed in detail in the next Section.

4.4 Assembly

The GE1/1 entire assembly is divided into two parts: pre-assembly outside the clean room and assembly inside the clean room. More information and technical details about the CMS GE1/1 chamber assembly procedure can be found in Ref. [17–19].

4.4.1 Pre-assembly

The first day is dedicated to pre-assembly, which includes preparing the readout board, drift board, internal frames, and external frame of the chamber. During the pre-assembly of GE1/1 chambers, the following steps were taken:

4.4.1.1 Preparation of the readout board

The GE1/1 readout board was prepared by mounting the brass inserts in the lateral flanges, threading of the gas holes, and gluing the gas connectors. All components required for the preparation of the board are shown in Figure 4.7.

- First, the readout board was cleaned using a vacuum cleaner, then it was placed on a table with the strips facing up (never touch the readout strips with bare fingers). It is recommended to use a soft material or a sheet of foam in between to avoid scratching on the Panasonic connectors.
- Using the clamping hand, the brass inserts are clamped in the dedicated hole in the readout board flanges (threaded side of insert sitting inside the hole),

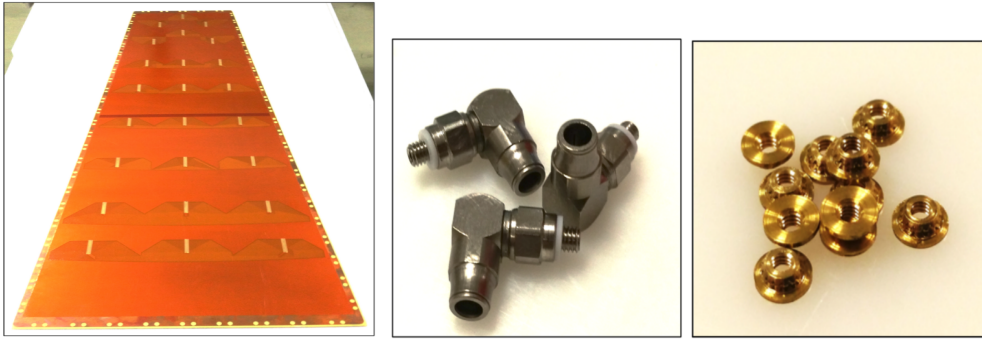


Figure 4.7: List of components required for the readout board preparation: readout board (left), gas connectors (middle), and brass inserts (right).

until they are getting fixed in it as shown in Figure 4.8. Finally, the brass inserts are checked for appropriate insertion by running fingertips over the insert heads.

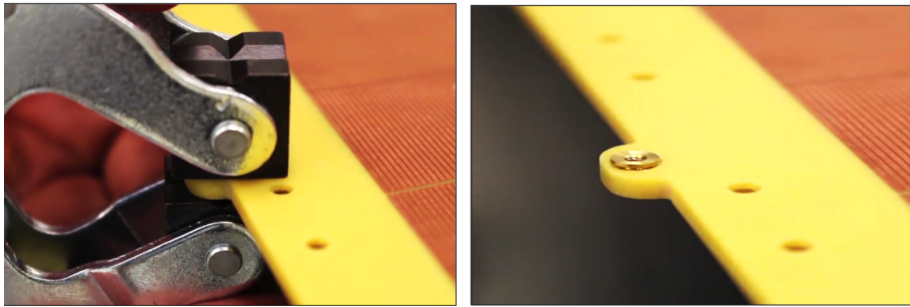


Figure 4.8: Clamping of brass insert using clamping hand.

- On the other side of the readout board, a 3 mm tap was inserted vertically in the hole dedicated to the gas connector to create the thread pattern in the material by slowly screwing by hand as shown in Figure 4.9 (left).
- Dust was removed using a vacuum cleaner and clean tissue paper with ethanol was used to clean both sides of the readout board around the hole.
- On a metallic surface which was cleaned by ethanol, the two parts of epoxy glue were mixed until they became opaque and uniform. The Teflon washer was then removed from the gas plugs and the metallic stick was used to apply a thin ring of glue between the body and the threaded part of the connector as shown in Figure 4.9 (middle).

- The gas connector was gently screwed onto the readout board until the base of the connector body touched the readout board. The glue should form a neat and smooth ring all around the base of the connector as shown in Figure 4.9 (right) and the readout board was then cleaned.

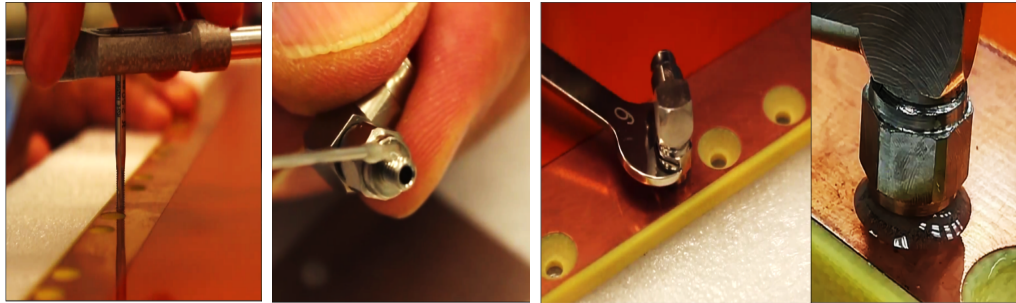


Figure 4.9: Making the thread in the gas inlet/outlet (left), applying glue to the gas connector (middle), and fixing the gas connector onto the board (right).

4.4.1.2 Preparation of the drift board

The GE1/1 drift board was prepared by cleaning the board, mounting the pull-outs, soldering the HV pins, and Surface Mount Devices (SMD) components (one 10 M Ω resistor, one 100 k Ω resistor, and one 330 pF capacitor) at the dedicated position. All components required for the preparation of the drift boards are shown in Figure 4.10.

- The active area of the drift board was facing up (never touch the active area of the drift with bare fingers), and the pull-outs were placed with screws and nylon washers in the dedicated holes. Figure 4.11 shows how the screws were tightened using a torque screw-driver at a maximum torque of 1.2 Nm.
- Previous steps were repeated all around the board, except for the large base of the trapezoid, where the HV components were initially soldered.
- First, the HV circuit was cleaned with a clean piece of tissue and ethanol. Figure 4.12 shows the HV and SMD components at their respective places.

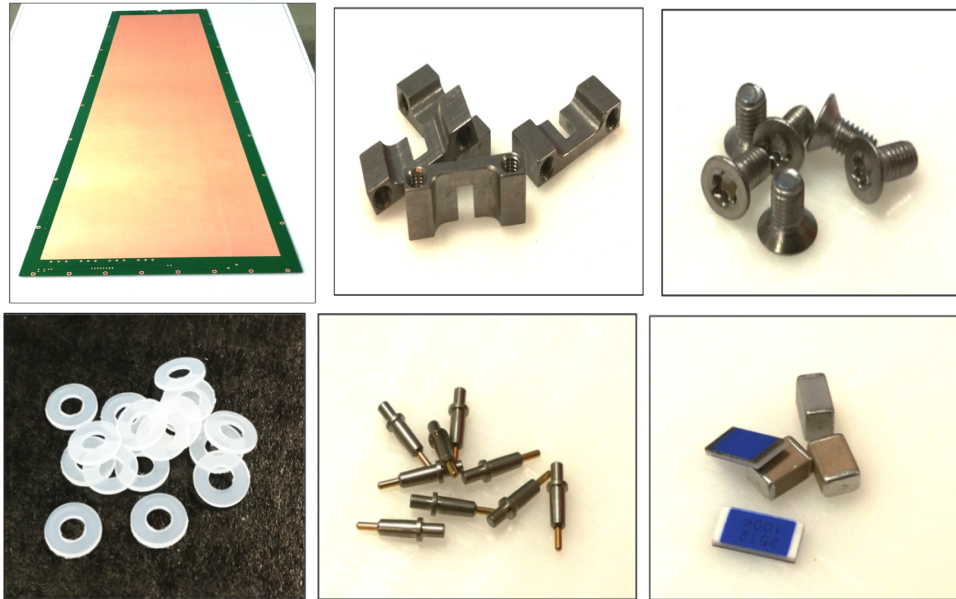


Figure 4.10: List of components required for the drift board preparation: drift board (top left), pull-outs (top middle), M3×6 screws (top right), nylon washers (bottom left), HV contact pins (bottom middle), and SMD components (bottom right).

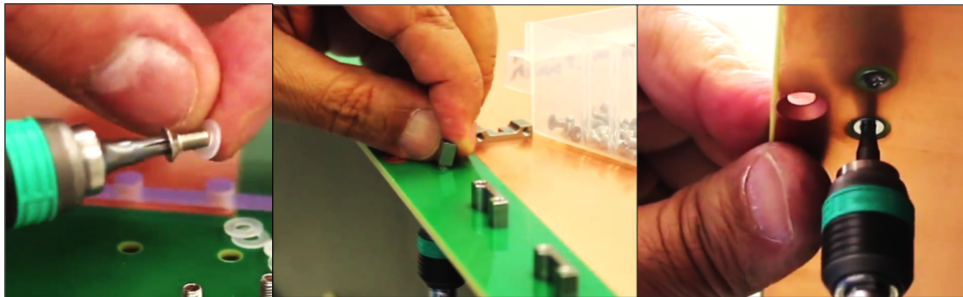


Figure 4.11: Mounting the pull-outs on the drift board using torque screw-driver.

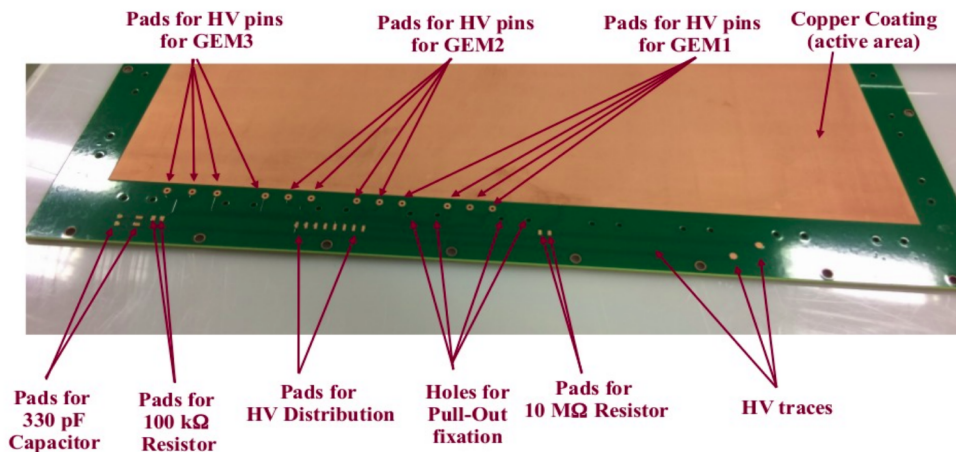


Figure 4.12: Positioning of the HV pins and SMD components on the drift board.

- Then, as shown in Figure 4.13 (left), the 6.1 mm pins in GEM1 pads, 8 mm pins in GEM2 pads, and 9.28 mm in the GEM3 pads were placed vertically aligned within the holes. The soldering tin was placed all around the base of the pins while heating the pads with the iron. The temperature was kept below 350° to avoid melting of the glue that holds the copper pads to the Printed Circuit Board (PCB).
- The SMD component pads were cleaned with tissue and ethanol. A little amount of soldering tin was used to “wet” the pads before attaching the SMD components as shown in Figure 4.13 (middle). After that, the SMD components were placed on the pad and held in position with soldering tweezers until the tin solidified completely as shown in Figure 4.13 (right).

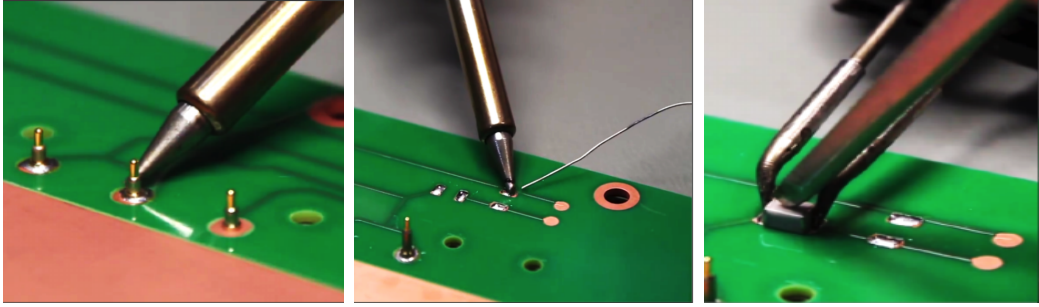


Figure 4.13: Soldering the HV pins (left), preparing the SMD pads (middle), and soldering the SMD components (right).

- The drift board was cleaned in the tub by placing it vertically with the large base facing down. The tub was inclined to ensure that the acetone flowed before drying and left impurities on the edge of the PCB. Acetone was poured gently over the HV circuit while brushing all soldering points. This step was repeated several times to ensure that all contaminants were removed. The rest of the drift board was cleaned with a clean piece of tissue and ethanol, paying special attention to the regions around the pull-outs.
- The remaining pull-outs were positioned in front of the HV circuit in the same manner as previously indicated.

4.4.1.3 Preparation of the internal and external frames:

The GE1/1 frames were prepared by inserting brass inserts into the internal frame (10 pieces) and mounting an O-ring on the outer gas frame. All components required for the preparation of frames are shown in Figure 4.14.

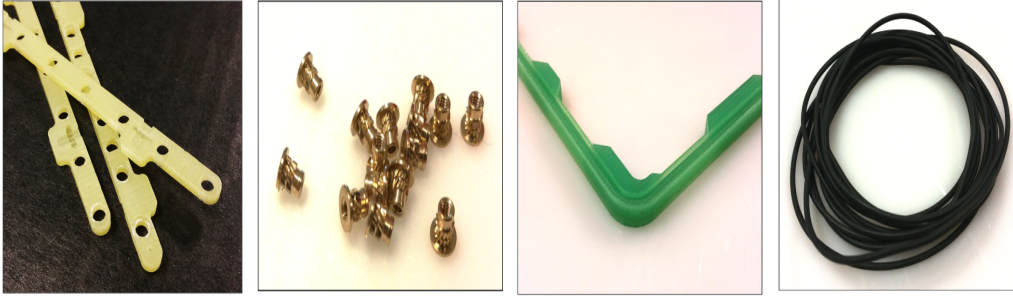


Figure 4.14: List of components required for frame preparation: internal frames, brass inserts, external frame, and O-ring (from left to right).

- As shown in Figure 4.15 (left), the internal frames were first taken out from their support by holding them in place with one hand while gently removing the tape, as frames are very thin and can easily break.
- The frame baseplate (FR4) was attached to the table, and the internal frame piece was placed on top with the hole of the frame aligned with the hole of the baseplate. The brass piece was put into the hole and gently hammered with the mallet until it was fully inserted as shown in Figure 4.15 (right). The same procedure was followed for the entire and all of the remaining frames. Also, carefully checked with fingers that there were no loose brass inserts that could fall out during the assembly.
- The O-rings were first cleaned in an ultrasonic bath for a few minutes before being inserted in the dedicated groove of the external frame. The O-ring was gently pulled into the trapezoidal shape of the groove by pulling simultaneously at each of its four corners. The O-ring was then gently pressed uniformly with the finger to insert it into the trenches as shown in Figure 4.16. If not, some

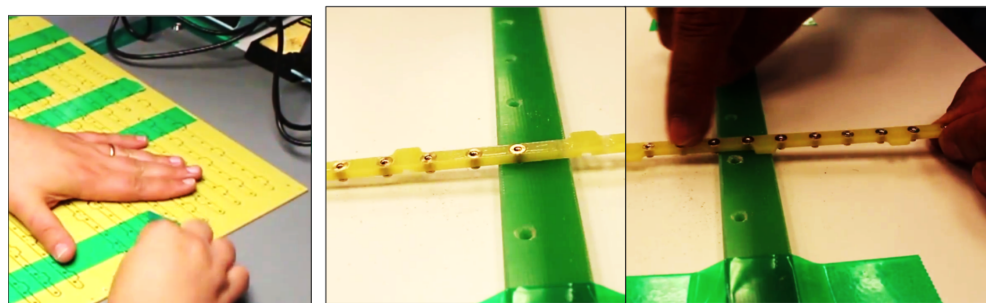


Figure 4.15: Removing the internal frames (left) and inserting brass inserts into the 3 mm internal spacer (right).

parts of the O-ring may be too compact, which can later deform the PCB when closing the chamber, or create leak points. Also, checked by eye and with finger that the O-ring was properly fitted and slightly exceeding the frame's housing.

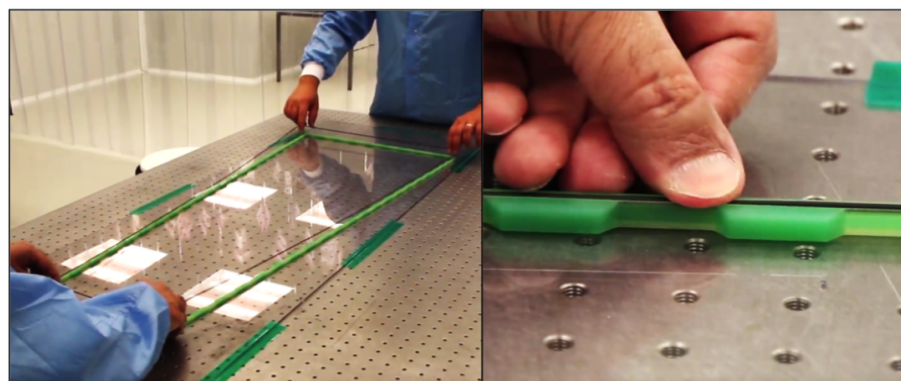


Figure 4.16: Inserting the VITON O-ring into the external frame.

4.4.2 Assembly in clean room

The major assembly of the chamber was completed on the second day, which included the assembly of the GEM stack and the closing of the chamber. Because the GEM technology is strongly dependent on the quality of the GEM foils. Any kind of contamination can damage a GE1/1 chamber as particles passing through the micron-sized holes in the GEM foils can cause electrical shorts. The GE1/1 chambers at PU were also assembled in a class-100 clean room with air purity values of no more than 100 particles of $\geq 0.5 \mu\text{m}$ size per cubic meter of the inside air.

Furthermore, no more than four persons may work at the same time and everyone must wear aprons, dust free gloves, facial masks, shoes, headcovers, etc.

4.4.2.1 Assembly of the GEM stack

The assembly of the GE1/1 stack includes the preparation of the components like internal frames, screws, short or long metallic nuts, washers, guiding pins, the cleaning and testing of the GEM foils, and the mounting of the stack. All components required for the assembly of the GEM stack are shown in Figure 4.17.



Figure 4.17: List of components required for the assembly of the GEM stack: GEM foil (top left), prepared internal frame (top middle), M2×6 screw (top right), M2.5 nuts (bottom left), guiding pins (bottom middle), and M2×6 long screws (bottom right).

- **Preparation of the GEM foils:**

- After pre-assembly, the GEM foils have to be extracted one at a time by at least two trained operators while maintaining good tension on the foils to avoid the formation of folds. The operators should not wear gloves but carefully wash their hands with ethanol.

- Then, every GEM foil (together with its frames) was fixed to an FR4 frame, aluminum, or any other rigid that had been positioned vertically against a wall.
- All the possible dust was removed from both sides of the foils using the anti-static roll on the foil, including the active and excess Kapton areas on the sides as shown in Figure 4.18 (left).
- To test the GEM foils, the HV clips were placed on the top and bottom of the foil as shown in Figure 4.18 (right).

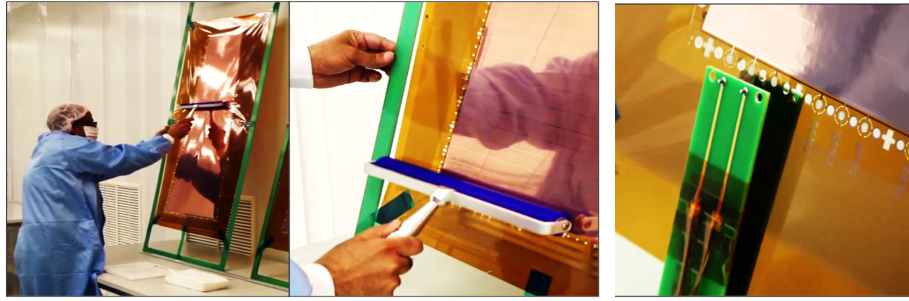


Figure 4.18: Cleaning the GEM foil with a anti-static roll (left) and mounting the HV clip to the foils (right).

- With relative humidity, of less than 40%, the multi Mega-ohmmeter (Megger)/insulation meter was adjusted to 550 V and this voltage was applied for several minutes. After a few seconds, the impedance of the foils should exceed $20\text{ G}\Omega$, and the number of sparks should be counted every 30s for the next 10 minutes. The spark rate should be less than 2 Hz during the last two or three minutes of test. After the test, the GEM foil was discharged. If the quantity of sparks exceeds this limit, the foil must be cleaned with the anti-static roll once again. If the problem persists after several cleaning attempts, the foil must be sent back to the production workshop and cleaned with Deionized water.
- The uncut areas of the pre-cut circular spare HV pads were then cut with a very sharp blade to define the position of the GEM1, GEM2, and

GEM3 foil in the chamber, as shown in Figure 4.19. The GEM2 and GEM3 pads were removed for the GEM1 foil position, the GEM1 and GEM3 pads were removed for the GEM2 foil position, and the GEM2 and GEM3 pads were removed for the GEM3 foil position. A horizontal cut must be followed by a vertical cut as shown in Figure 4.20.

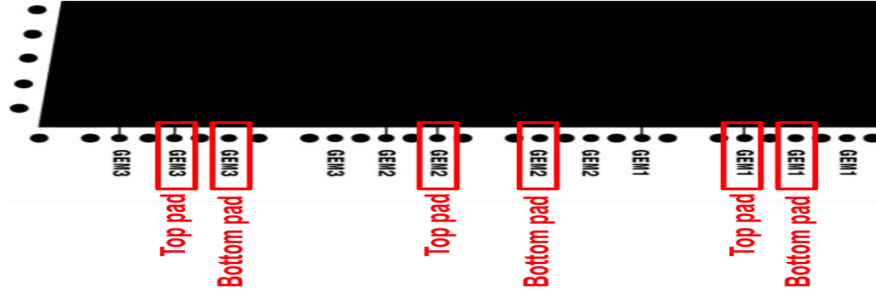


Figure 4.19: Location of the HV pads.

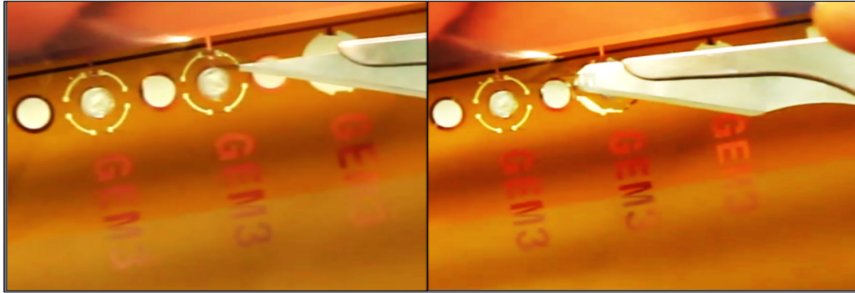


Figure 4.20: Cutting off the spare HV pads.

- **Mounting of the GEM stack:**

- The Plexiglas baseplate was placed and fixed to the assembly table using tape and the entire surface of the baseplate was cleaned with fresh tissue paper and ethanol. The guiding pins were inserted in all of the single holes and one of the grouped holes all around the trapezoid as shown in Figure 4.21.
- The 3 mm internal frames were cleaned from both sides using the anti-static roll, being careful to ensure that no brass inserts fell out of the

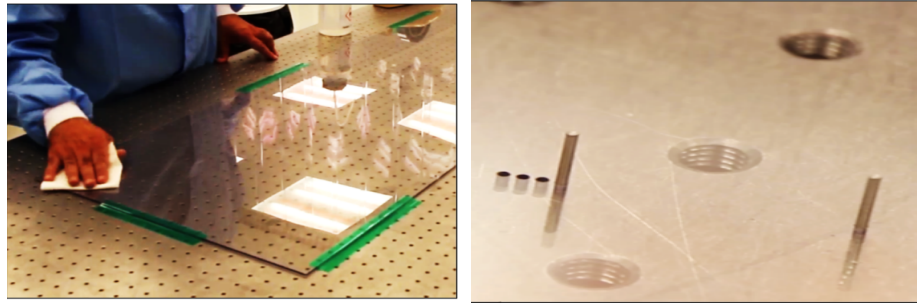


Figure 4.21: Cleaning the assembly baseplate (left) and inserting the guiding pins (right).

frame. The drift gap will be formed by this set of frames with brass inserts.

- The internal frames were placed on the baseplate all the way around the trapezoid using the guiding pins. The direction and location of each frame piece were determined by matching the geometry of the grouped holes on the frame and baseplate. The flanges must always face the inside of the trapezoid as shown in Figure [4.22](#).

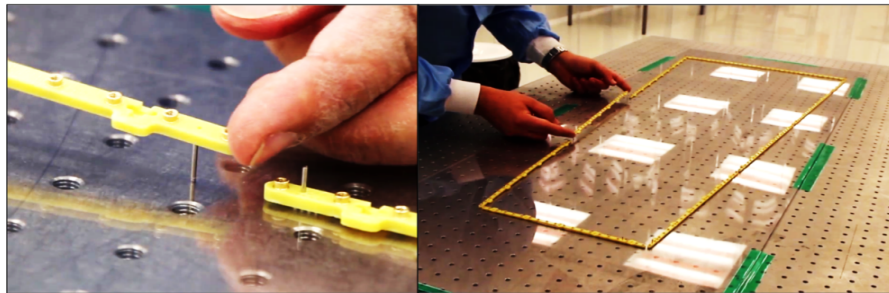


Figure 4.22: Mounting the 3 mm internal frames.

- The GEM1 foil was placed on top of the 3 mm frames using the guiding pins for proper alignment. The foil should be well aligned to prevent scratching of the active area by the guiding pin. A sharp blade was used to cut the tape after inserting the foil in order to detach it from the frame.
- The foil was stretched by two people working in opposite directions of the table. The tape on the Kapton area was connected during pre-stretching, on both the large and small bases of the trapezoid, and then gently pulled

towards the operator before being fixed the tape on the table, as illustrated in Figure 4.23 (left). Excessive stretching strength will result in the deformation of the GEM foil. Figure 4.23 (right) shows the position of this pre-stretching and the order in which it must be performed.

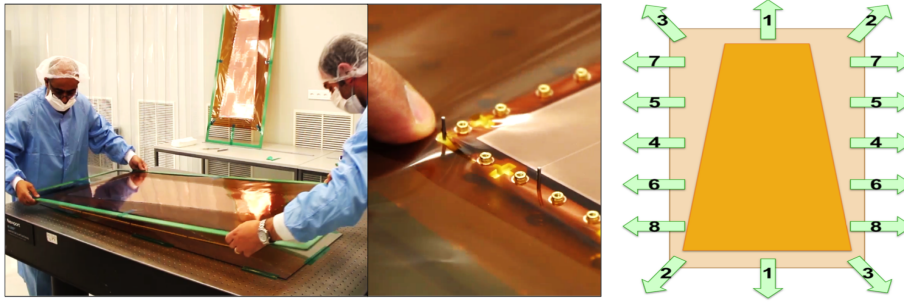


Figure 4.23: Placing GEM1 (left) and pre-stretching points (right).

- Pins of the Megger connected to the pads corresponding to the top and bottom of GEM1 foil. A voltage of 550 V is applied between the two electrodes for a few minutes, and the resistance must reach at least 20 G Ω after a few seconds. After the test, the GEM foil was discharged.
- The set of 1 mm thick internal frames was cleaned which will form the first transfer gap. These frames have cross-shaped holes to accommodate the metallic nuts for the final stretching. The frames were positioned and orientated properly on the stack by matching the shape of the grouped holes as shown in Figure 4.24.

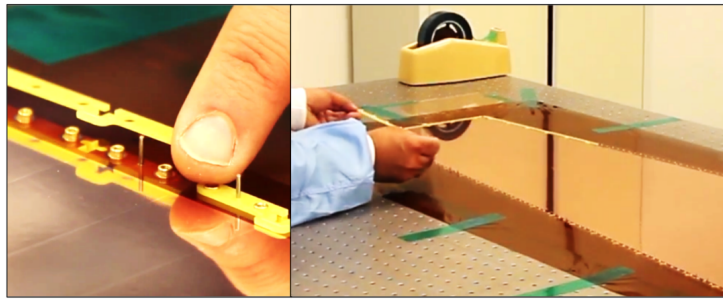


Figure 4.24: Mounting the 1 mm internal frames.

- Before mounting the frame on the large base of the trapezoid, the GEM1

foil was tested with the insulation meter at 550 V again. After the test, the GEM foil was discharged.

- The second GEM foil was then tested with an insulation meter before being placed into the stack with the guiding pins for alignment, as shown in Figure 4.25 (left). The same procedures were used to detach it from the frame, and the foil was stretched in every direction using green tape.
- The set of 2 mm thick frames was cleaned and placed on the stack to create the second transfer gap, and then the M2.5 metallic nuts were inserted into the dedicated housing of internal frames as shown in Figure 4.25 (right).

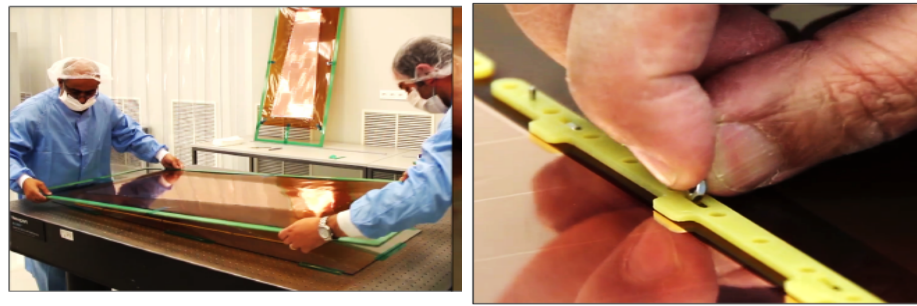


Figure 4.25: Placing GEM2 foil (left) and the metallic nuts into the internal frame (right).

- The impedance of the second foil was verified using 550 V before putting the final component on the large base of the trapezoid. The GEM foil was then discharged, and the final piece of internal frame was installed, along with the remaining nuts.
- The third GEM foil was tested and placed on top of the stack, and the tape that held the foil and its frame together was removed.
- The last set of 1 mm internal frames was cleaned and placed on the stack to produce the induction gap, and the foil was tested by applying 550 V on the pads.

- **Protecting the GEM stack:**

- The Plexiglas was cleaned using a anti-static roll and ethanol before being mounted with guiding pins on top of the GEM stack, as shown in Figure 4.26 (left).
- The long M2 screws and washers are used to attach the Plexiglas cover to the stack. Screws were placed into the dedicated holes to catch the brass inserts that were already fixed in the 3 mm internal frames. These screws were placed on the borders of each piece of the internal frame and in the corners of the trapezoid as shown in Figure 4.26 (middle).
- M2×6 mm screws are placed in the dedicated holes of the Plexiglas cover to hold the layers of internal frames together. These screws are tightened with screw-driver without putting too much strain on the thin frame sections as shown in Figure 4.26 (right).

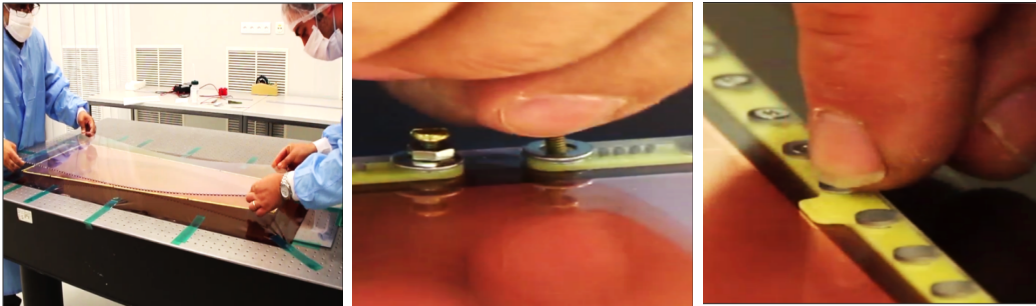


Figure 4.26: Placing the Plexiglas cover (left), attaching the Plexiglas cover to the GEM stack (middle), and closing the internal frame (right).

4.4.2.2 Closing of the chamber:

The closing of the GE1/1 chamber included the final preparation of the GEM stack, stretching of the foils, and electrical test after closure. The assembled GEM stack, prepared drift board and readout board, stretching screws M2.5, M3×6, and nylon washers are required for chamber closing.

- **Finalizing the GEM stack:**

- The GEM stack was cleaned with the vacuum cleaner, focusing on the edges of the internal frame and holes in the Plexiglas cover as shown in Figure 4.27 (left).
- The excess Kapton foil was then cut away with a sharp knife and the Plexiglas cover served as a stencil as shown in Figure 4.27 (middle).
- The extra Kapton was removed while being careful not to detach the baseplate from the assembly table as shown in Figure 4.27 (right). The remaining piece of Kapton that exceeds the frame stack should be no longer than a fraction of a millimeter in length. The vacuum cleaner was used to remove any shavings that may have remained after cutting.
- The drift board was then cleaned with the vacuum cleaner, with a focus on the pull-out and the electrical circuit area as shown in Figure 4.28 (left), and the active area was cleaned with the anti-static roll. The drift board was placed next to the GEM stack (without touching the active area).
- The stack was then separated from the Plexiglas baseplate by gently lifting it upward without touching the bottom of GEM3 with fingers or the guiding pins that may remain attached to the baseplate. The stack was then placed on the drift board in the area enclosed by the pull-outs.
- All of the remaining guiding pins in the frame stack were removed with tweezers, as shown in Figure 4.28 (middle). The drift board clamped to the assembly table with the aluminum bars and chambers should be orientated towards the inner side of the trapezoid. The aluminium bars were then screwed to the assembling jig. The amount of force required to ensure that the system maintains the flatness of the board while performing the stretching of the foils.
- The long M2 screws that hold the Plexiglas cover and GEM stack together were removed as shown in Figure 4.28 (right). It was replaced with the

final M2×6 mm screws. The Plexiglas cover was left in place without screws to protect the foil. The internal frames were free to move during the stretching of the foils.

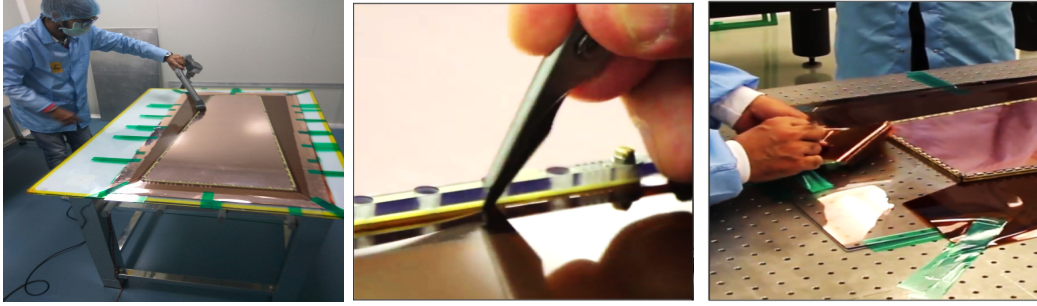


Figure 4.27: Cleaning the GEM stack (left), cutting the excess Kapton foil (middle), and removing the excess Kapton foil (right).

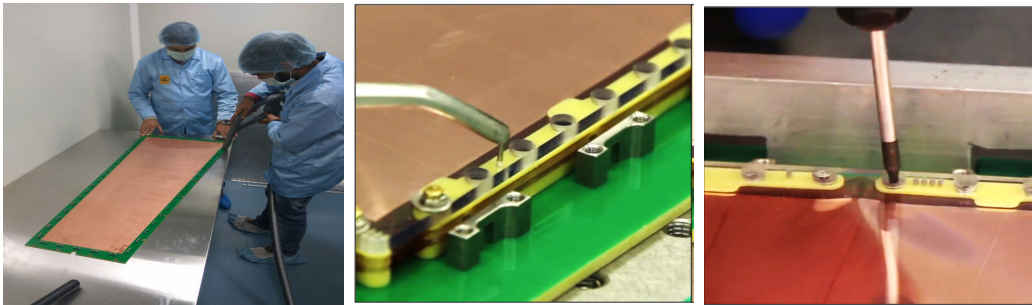


Figure 4.28: Cleaning the drift board with the vacuum cleaner (left), removing the guiding pins (middle), and removing the Plexiglas cover from the GEM stack (right).

- **Stretching of the GEM foils:**

- The stretching of the GEM foil was performed exactly in the same order as shown in Figure 4.29 (left) to avoid any formation of the waves on the foils, especially between the different parts of internal frames.
- The M2.5 screws were placed in the lateral holes of the aluminum bars, starting with a large and small base of the trapezoid to the lateral sides as shown in Figure 4.29 (right). Stretching should be done with the two persons facing in opposite directions.
- After all of the screws were tightened, the Plexiglas cover was removed

and the top surface of the GEM stack was cleaned with the anti-static roll, being careful not to damage the foils.

- The three GEM foils were then tested one after the other at 550 V using the dedicated pads on the drift board. After each test, the foil was discharged. The impedance of GEM foils must reach $20\text{ G}\Omega$ after a few seconds. However, the impedance of the gaps must reach greater than $100\text{ G}\Omega$ immediately after applying the voltage.
- After the completion of the test, the electronic screw-driver was used to complete the stretching of the foils. The nominal strength value on the side of the trapezoid is 9 cNm , although it can be slightly greater in the corners, reaching up to 15 cNm .

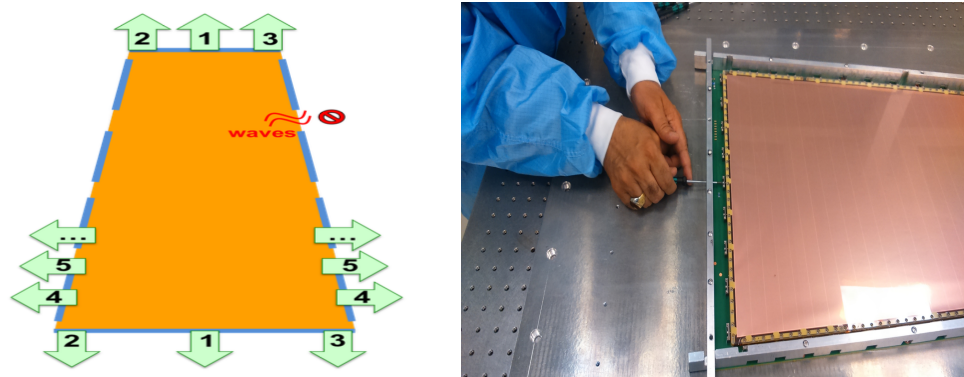


Figure 4.29: Description of the stretching procedure (left) and stretching of the foils (right).

- **Closing and electrical test:**

- The external frame was placed between the GEM stack and aluminium bars as shown in Figure 4.30 (left and middle). The vacuum cleaner was used on the side of the bars to pick up any dust that may have been formed during the stretching and to properly vacuum the gap between the stack and the bars. The anti-static roll was used once more to gently wipe the top surface of the foil.

- The vacuum cleaner was used to clean the active area of the readout board, taking care not to scratch the copper traces.
- The readout board was placed on top of the GEM stack without contacting the strips with fingers as shown in Figure 4.30 (right).

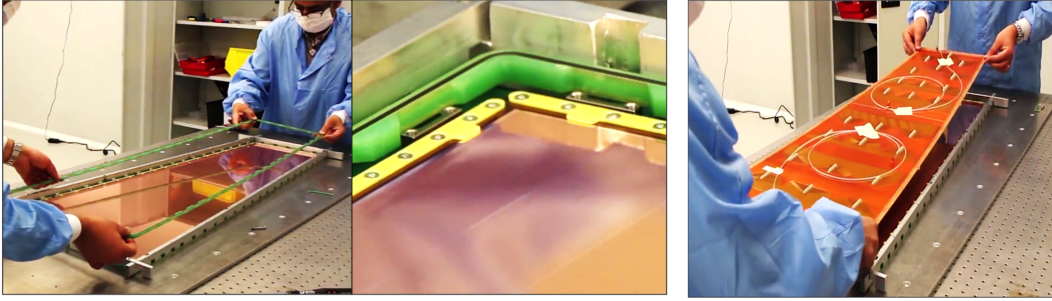


Figure 4.30: Inserting the external frame between the GEM stack and assembly jig (left and middle) and mounting the readout board on the stack (right).

- The Polyamide washers were placed into the conical holes and the M3 screws were inserted as shown in Figure 4.31 (left). The screws were tightened at 1.2 Nm using the manual torque screw-driver, as indicated in Section 4.4.1.2. When all of the screws were in place, the aluminum bars were removed to allow access to the chamber.
- The GEM foils and gaps are tested at 550 V once again and after each test, the foils were discharged.
- The induction gap can also be checked in this setup. The Panasonic-to-Lemo adaptor was attached to one of the readout sectors as shown in Figure 4.31 (middle).
- The 550 V was applied between the GEM3 bottom and the signal pad of the Panasonic-to-Lemo converter while the chamber was held vertically. The location of the GEM3 bottom pad is shown in Figure 4.31 (right). The impedance should immediately reach 100 G Ω .
- If the impedance of the induction gap is lower than 100 G Ω , or several discharges experienced at 550 V, there might be a short circuit of a weak

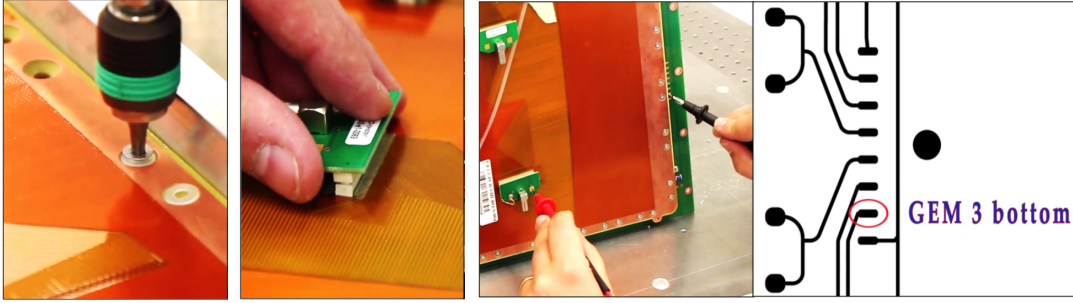


Figure 4.31: Sealing the chamber (left), placing Panasonic-to-Lemo adapter (middle), and testing the induction gap using an insulation meter (right).

point caused by the bending of the readout PCB toward the third GEM foil. In this case, the problem can be localized by removing all the 0Ω terminations from the Panasonic-to-LEMO adapters. Then negative and positive pins of Megger were connected to the HV pad corresponding to the GEM3 bottom and the signal pad of the Panasonic-to-LEMO adapter, respectively on the readout sector $i\eta=1$, $i\phi=1$. The impedance was then measured after one minute and the number of sparks was counted for 1-2 minutes. This procedure was repeated for all 24 readout sectors.

- The same procedure was used for all of the readout sectors of the chambers.

The GE1/1 chamber is now ready for the next round of quality checks.

4.5 GE1/1 quality assurance/quality control test

It is important that the chamber must pass all Quality Assurance (QA) and QC tests in order to ensure fully efficient. The standard QA and QC procedures have been established by the GEM Collaboration at CERN to ensure that the chamber does not have any mechanical or electrical issues that could impact its operation after installation in the Muon endcap in the CMS experiment. These procedures are the same at all GEM assembly and QC sites worldwide. Out of the few tests, some are carried out at CERN while others are carried out at the production sites.

Several steps of the overall process for various QA and QC procedures are shown in Figure 4.32. There are mainly three phases:

- Purchasing, manufacturing, and QC testing of assembly components at the CERN production site;
- Single GE1/1 chamber assembly and QC test at all production sites;
- Superchambers were built and tested at the CERN manufacturing site using chambers delivered from other production sites before being installed in the CMS muon system.

The following Subsections detail the QC2-QC4 tests performed during and after the assembling of the GE1/1 chambers at the PU production site:

- GEM-foils: QC2 fast - leakage current test;
- GE1/1 chamber: QC3 - gas leak test in CO₂;
- GE1/1 chamber: QC4 - HV test in CO₂;

4.5.1 QC2 - leakage current test

QC2 - leakage current test determines the quality of a GEM foil by measuring the maximum leakage current flowing on the surface of the GEM holes. As a result, the QC2 test must be performed before, during, and after the assembly of the chamber. When a voltage is applied across the GEM electrodes, current begins to flow from the top to the bottom of the foil due to the surface conductivity of the Polyamide (Kapton). This current is referred to as the leakage current because it always flows across the Polyamide surface and cannot be eliminated. The quality of GEM foil and cleanliness both have a significant impact on its conductivity. Any dust, contamination or defect could act as a bridge between the top and bottom

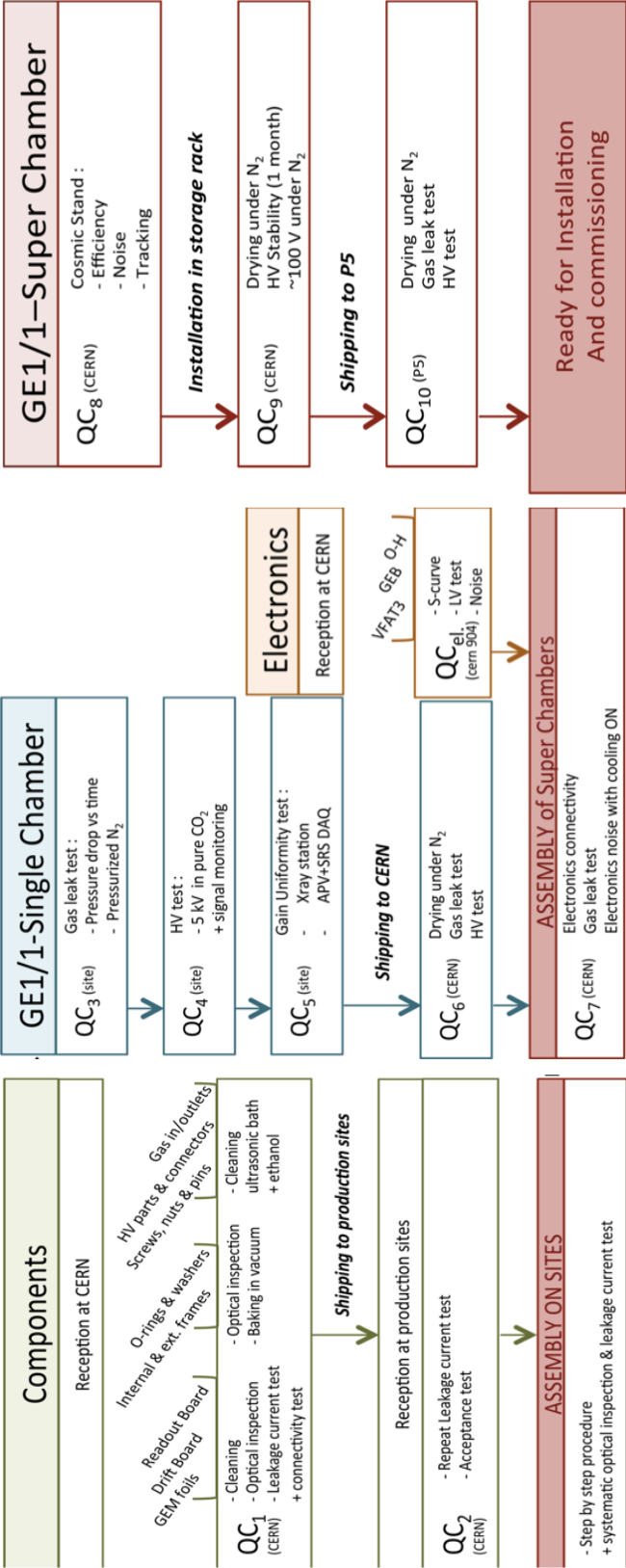


Figure 4.32: Schematic overview of the quality assurance and QC procedures of the GE1/1 chambers [17].

layers of foil and hence leads to an increase in leakage current through the Polyamide surface. To ensure the perfect functionality of the GEM foils, the QC2 procedure is divided into five steps:

- GEM foil preparation and cleaning: All activities involving GEM foil must be performed in a clean room, as previously mentioned in Subsection [4.4.2.1](#).
- QC2 fast (GEM acceptance test I): As previously mentioned in Subsection [4.4.2.1](#).
- QC2 long (GEM acceptance test II): Before the shipping of the foils to the production sites, the HV long-term stability of the GEM foils was measured in a dry environment. It consists of monitoring the leakage current and possible sparks when the GEM foil is subjected to HV, typically up to 600 V during a period of 30 minutes to 1 hour. More detail can be found in Ref. [\[20–24\]](#).
- QC2 fast post-assembly in the clean room: As previously discussed in Subsection [4.4.2.2](#).
- QC2 fast in dry gas: The QC2 fast test in dry gas must be performed just before mounting the HV circuit on the chamber. The GE1/1 chamber was flushed with pure CO₂ for at least 5 hours at a gas flow rate of 2.5 L/h. A detailed overview of the gas system is discussed in the next Sections [4.5.2](#) and [4.5.3](#). The CO₂ gas is useful to dry out the GEM foils, prevent electron amplification, and therefore avoid the possible propagation of discharges. After flushing the chamber, the fast test was performed in the same manner as the QC2 fast test in a clean room (previous test). After one minute of no repetitive discharges, the impedance of the GEM foils and the gap must be greater than 20 GΩ and 100 GΩ, respectively. A few discharges may occur after a few seconds due to dust deposition on the GEM foils. When the dust particles have been blown away, the GEM should be completely quiet.

All eight chambers are called “GE1/1_X_S_INDIA_00YY” (where ‘YY’ stand for ‘02’, ‘07’, ‘08’, ‘09’, ‘10’, ‘15’, ‘16’, ‘17’) assembled at PU site passed the QC2 - leakage current test and are then taken to next QC test, called QC3 - gas leak test.

4.5.2 QC3 - gas leakage test

The QC3 - gas leak test aims to determine the gas leak rate of a GE1/1 chamber by monitoring the drop of the internal over-pressure as a function of time. There is a chance that gas will leak from the drift board and readout board since they are secured with hundreds of screws. This might happen because some screws were left loose during assembly or were tightened so forcefully that they broke inside. Gas leaks have an impact on the functioning of the chamber in addition to producing gas waste, which pollutes the environment. A schematic representation of the QC3 gas circuit is shown in Figure 4.33 (top). The gas input of the chamber is connected to a CO₂ gas bottle equipped with a two-stage pressure regulator with the output pressure being set to a maximum of 50 millibars (mBar). The input flow-meter, together with its manual valve, allows for controlling the gas flow rate at the entrance of the chamber while the output flow-meter helps to monitor the output flow and therefore identify large leaks. The input and output valves are used to isolate the chamber from the rest of the gas system during the gas leak test. The calibration of the gas system is done before connecting the chamber, as indicated in Figure 4.33 (bottom). A pressure indicator is connected between the input and output valves in order to monitor the over-pressure inside the system only or the chamber + system. Additional atmospheric and ambient temperature sensors are located nearby the test stand. All sensors are then connected to a micro-controller (e.g. ARDUINO Mega 2560), which is responsible for the data acquisition (DAQ) and slow control. A typical QC3 system test stand prepared at the PU site is shown in Figure 4.34.

The QC3 gas leak test is divided into two steps: system calibration and chamber + system leak measurement.

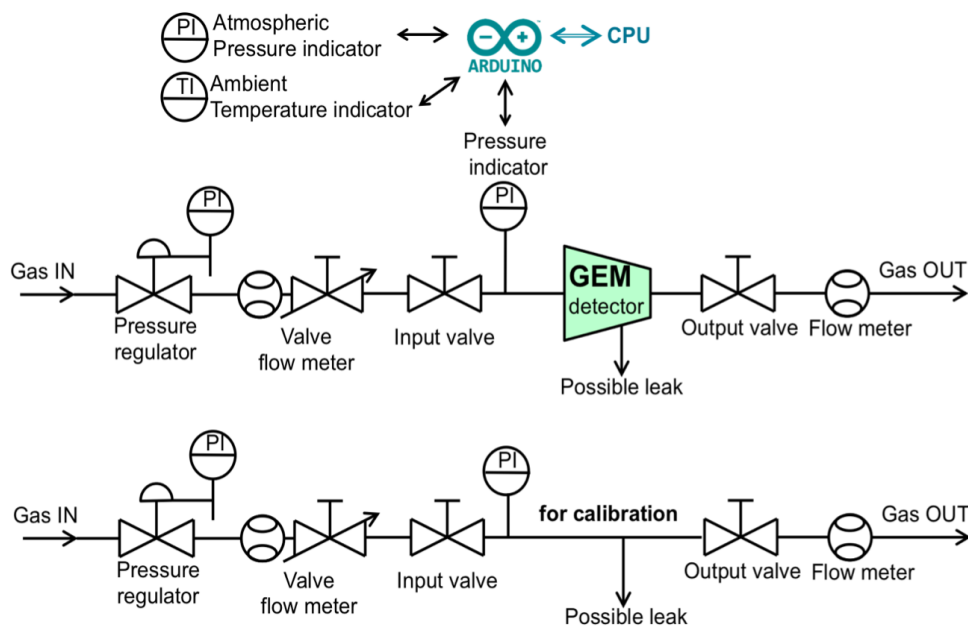


Figure 4.33: QC3 gas leak test experimental setup.

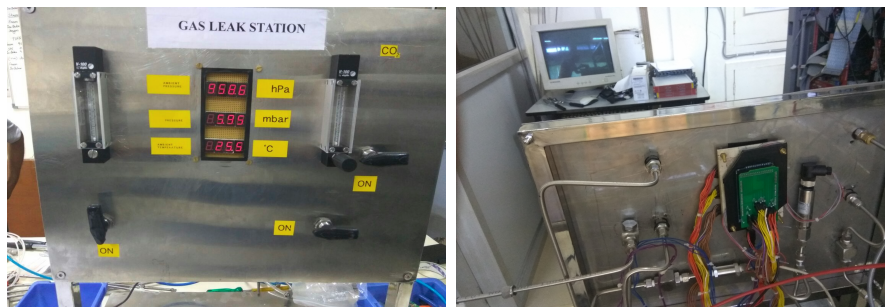


Figure 4.34: Front (left) and back view (right) of QC3 gas leak test Stand.

- **System calibration:** The calibration of the gas system was performed by connecting the input and output lines of the gas system. The input and output valves of the gas system are then opened. The input gas flow rate of CO₂ was set to 5 L/h by adjusting the manual valve of the input flow-meter. The output flow rate was instantly visible on the output flow-meter. The output valve of the gas system was then closed and the pressure inside the system began to rise. The over-pressure in the system was set to about 25 hectopascal (hPa), and the input valve was closed as soon as this pressure was reached. The DAQ system monitored the drop in internal over-pressure as a function of time as well as other environmental factors (room temperature, relative humidity, and atmospheric pressure). The pressure drop in the gas system should not exceed 1 hPa per hour. When the system passed this test then the chamber + system gas leak is measured.
- **Chamber + system leak measurement:** The gas inlet and outlet of the chamber were connected to the gas system. The input and output valves of the gas system are then opened. The manual valve of the input flow-meter was adjusted to allow CO₂ to pass through the chamber. The input gas flow rate of CO₂ was set at 5 L/h and the chamber was flushed for at least 30 minutes. The output flow rate was immediately visible on the flow-meter. The output valve of the gas system was then closed and the pressure inside the chamber began to rise. The over-pressure in the chamber was set to about 25 hPa, and the input valve was closed to isolate the chamber as soon as this pressure was reached. The valve on the input flow-meter was then closed. Because the GE1/1 chambers cannot sustain internal over-pressures greater than 40–50 hPa, before closing the output valve, confirmed that the gas system is under control. The DAQ system records the over-pressure inside the chamber as a function of time as well as other environmental factors. The GE1/1 chamber is validated if the pressure drop in the chamber + gas system does not exceed

7 hPa per hour.

The internal gas pressure is parametrized by $P(t)=P_0\exp(-t/\tau)$, where P_0 is the initial internal over-pressure and parameter τ specifies how fast the over-pressure inside the chamber decreases as a function of time. The QC3 acceptance limit is $\tau \geq 3.04$ hours, which corresponds to a maximum acceptable chamber + gas system leak rate of around 7 hPa/hr. All the chambers assembled at the PU production site successfully passed the QC3 gas leak test because gas leak reduction rate is less than 7 hPa/hr as shown in Figure 4.35.

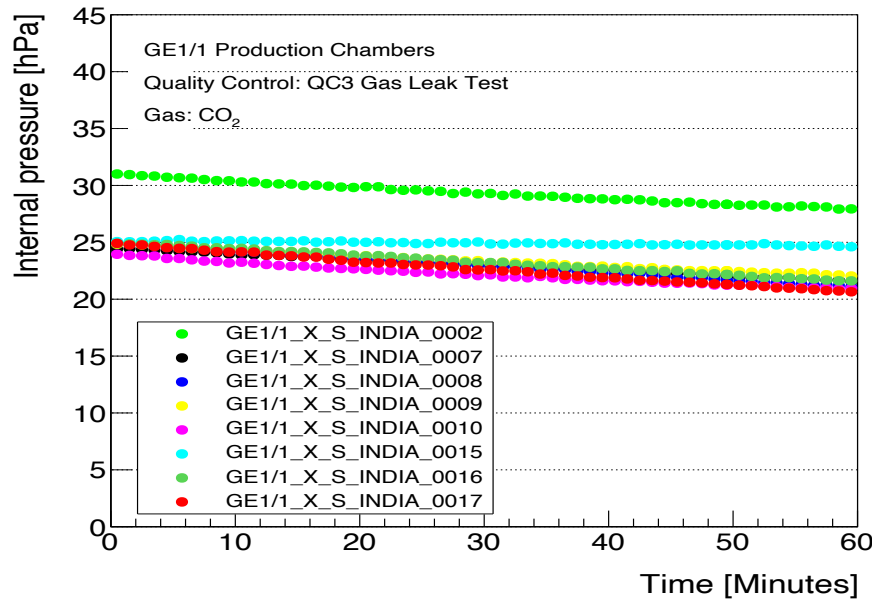


Figure 4.35: The QC3 test results for the GE1/1 chambers assembled at PU.

4.5.3 QC4 - HV test in CO₂

After the QC3 test, the QC4 test is performed which is a high voltage test to determine the voltage versus current curve of the chamber and to identify possible malfunctions, defects in the HV circuit, and intrinsic noise rate in the GEM chamber. A programmable HV power supply (CAEN SY5527) provides the negative HV

(V_{SET}) which allows to control the current limit (I_{SET}), the steps to ramp up and down the voltage, the maximum voltage, and the trip time. This power supply must be able to deliver a current up to 1 mA with a monitoring resolution less than $1\text{ }\mu\text{A}$ in order to identify possible leaks, sparks, or unusual current fluctuations.

The QC4 data acquisition setup for spurious signal measurements is shown in Figure 4.36. A HV low pass filter was placed between the power supply and chamber to reduce the electronics noise (by rejecting the high frequency noise from the power supply) so that a stable voltage can be provided to the electrodes. The signal generated at the anode is processed using charged sensitive pre-amplifier connected to the bottom electrode of the third GEM foil through a decoupling RC circuit already existing on the drift board of the chamber. Further, amplified and recorded signal obtained from readout strips was sent to the timing filter amplifier + shaper unit. The output is then fed into the discriminator which converts it into a rectangular digital pulse, which is then fed into a dual timer, which provides a veto signal to prevent the fake counts. The generated digital pulse is then transferred to a scalar and counter unit for rate measurement. The following steps were taken during the QC4 test:

- The chamber was connected to the QC3 gas system and flushed with pure CO_2 gas for at least 5 hours at a gas flow rate of 5 L/h. A ground shielding was placed on top of the readout board to insulate the chamber from background noise. The Panasonic-to-LEMO adapters with 50Ω terminations were connected to all the readout sectors. The ground pad of the PCBs was connected to the grounding plate with copper tape as shown in Figure 4.37 (left).
- Then the input wire of the pre-amplifier is soldered to the dedicated pad on the drift PCB as shown in Figure 4.37 (right).
- The HV voltage cable was connected to the Safe High Voltage (SHV) connector on the chamber and an HV low-pass filter was connected between the

SHV connector of the chamber and power supply. In this configuration, the maximum noise level should be less than 100 mV. The power supply was set up as follows: Ramp up = 50 V/s; Ramp down = 250 V/s; Max voltage = 5050 V; Trip time = 0 s, and Trip action = KILL.

- The amplifier was set to Coarse gain = 4, Fine gain = 4.5, Integrate = 500 ns, and Differential = 500 ns.
- Copper ribbons are also utilized to properly ground the chamber, pre-amplifier, and HV filter.
- The discriminator threshold and scalar clock were both set at -140 mV and 60 s, respectively.
- A digital multimeter was used to measure the total resistance of the HV circuit (the divider plus all HV filters). The total resistance was used to calculate the expected current in the divider for every HV point.
- The current limit I_{set} of the power supply was set according to the expected target value plus $5 \mu\text{A}$, and the corresponding high voltage V_{set} was switched on. After the operating current and voltage get stable, the actual voltage V_{mon} and current I_{mon} , and intrinsic noise rate of the chamber were recorded. The intrinsic noise rate is defined as the rate of signals that do not arise due to the ionization of the gas. The total number of counts was measured using the scalar unit for the intervals of 60 s. During this step, the oscilloscope was also examined to identify possible spurious signals, unusual or repetitive signals in the chamber.
- The previous steps were repeated until the chamber voltage reached 4900 V. The chamber voltage was ramped up to 3000 V in 200 V steps and up to 4900 V in 100 V steps. However, in order to bring the chamber to a new HV point, first the current limit I_{SET} is changed according to the value given in the template file, adding $5 \mu\text{A}$. The voltage V_{set} was then changed and confirmed.

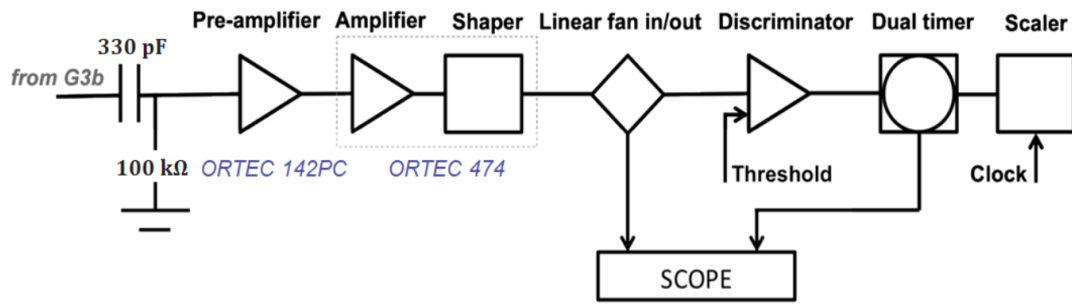


Figure 4.36: QC4 HV test experimental setup.

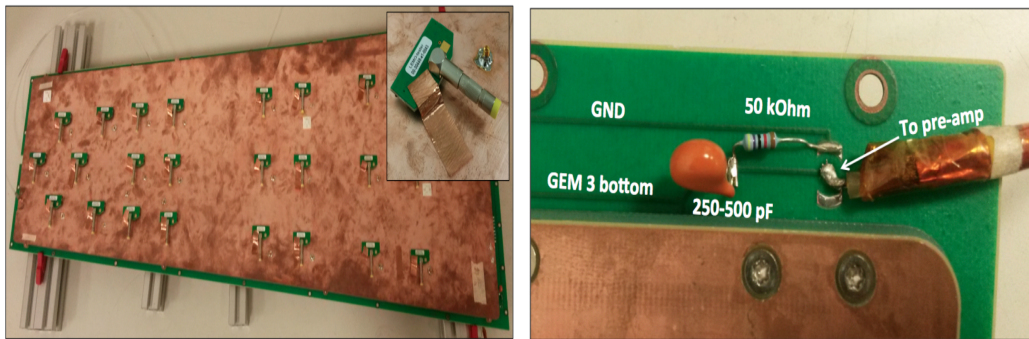


Figure 4.37: QC4 overview of the chamber with the grounding plate and 50 Ω terminations (left) and QC4 close view of the sample point on the drift PCB (right).

The I-V curve is plotted, which must be linear, and is fitted by a straight line to determine a resistance (“fitted resistance”). The reciprocal of its slope gives the resistance (R_{IV}) of the HV distribution system. This calculated resistance R_{IV} is then compared with the value of measured resistance by a digital multimeter ($R_{\text{multimeter}}$). The chamber will pass the test if the deviation in the resistance between the $R_{\text{multimeter}}$ and R_{IV} is $\leq 2\%$, because a deviation greater than 2% indicates a loss of linearity. The maximum intrinsic noise rate should be ≤ 100 Hz at the value of current $\sim 999 \mu\text{A}$, which is less than 10^{-2} Hz/cm^2 , which is a negligible value in comparison to the 4.5 kHz/cm^2 background rate expected in CMS in standard operating conditions. Table 4.2 shows the value of measured resistance ($R_{\text{multimeter}}$) by multimeter, resistance (R_{IV}) by IV curve, and maximum intrinsic noise rate (at the value of current $\sim 999 \mu\text{A}$) for all the chambers assembled at PU. Figure 4.38 (left) shows the monitored voltage (V_{mon}) as a function of the monitored current (I_{mon}) and Figure 4.38 (right) shows the intrinsic noise rate as a function of the monitored current for different GE1/1 production chambers. All of the chambers successfully passed the QC4 HV test because the deviation in resistance is $< 2\%$ and the maximum intrinsic noise rate is 11.88 Hz which is very less than 100 Hz. Some data points of the eight chambers overlap, because the I-V relation of each chamber is similar.

Table 4.2: Results from QC4 HV test showing the value of measured resistance ($R_{\text{multimeter}}$) by multimeter, resistance (R_{IV}) by IV curve, and the maximum intrinsic noise rate (at the value of current $\sim 999 \mu\text{A}$) for all assembled chambers at PU.

Chamber ID	Resistance by multimeter (R_{IV})(M Ω)	Resistance by I-V curve ($R_{\text{multimeter}}$) (M Ω)	Intrinsic noise rate(HZ)
GE1/1_X_S.INDIA_0002	4.67	4.69	11.88
GE1/1_X_S.INDIA_0007	5.00	5.01	3.50
GE1/1_X_S.INDIA_0008	5.00	5.02	6.57
GE1/1_X_S.INDIA_0009	5.00	5.01	5.62
GE1/1_X_S.INDIA_0010	4.97	5.02	5.72
GE1/1_X_S.INDIA_0015	4.96	5.01	11.30
GE1/1_X_S.INDIA_0016	4.97	5.02	8.58
GE1/1_X_S.INDIA_0017	5.00	5.02	8.25

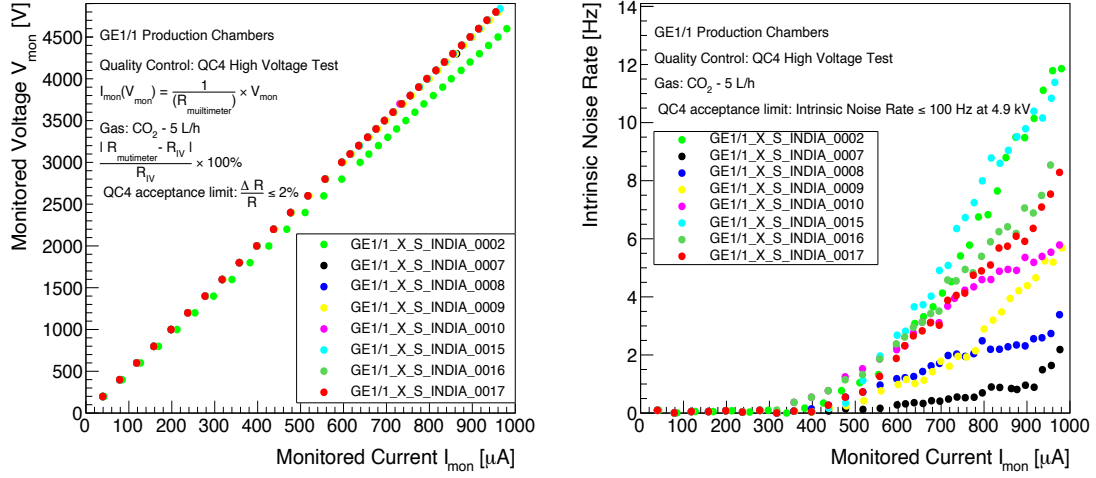


Figure 4.38: The QC4 HV test result (left) and intrinsic noise rate as a function of monitored current (right) for 8 GE1/1 chambers assembled at PU.

4.6 Further QC testing, integration and validation with cosmic-ray muons

- QC5 - Gain calibration test:** After QC4, the next test is QC5 which is performed in two steps: measuring the effective gain as a function of the voltage applied to the divider and response uniformity of the chamber. The measurement of the effective gain consists of comparing the primary current induced in the drift gap by the X-ray source with the output current after amplification. The effective gain is always measured on the central readout sector $(i\eta, i\phi) = (4, 2)$ as a function of the current through the divider using an X-ray tube with an Ag target and beam energy maxima at 22.5 keV and 25 keV. More details can be found in Ref. [20–24]. After completing the QC5 test, the chambers were shifted to CERN for further QC tests and superchamber assembly. At CERN, I participated in QC7 test and superchamber assembly.
- QC6 - HV test:** After chambers are returned to CERN, the first test performed is QC6, which involves measuring the I-V curve before and after the HV connector is replaced with the standard cable installed at CMS.

- **QC7 - Electronics connectivity, gas leak, and electronic noise with cooling test:** All connections of electronic components such as VFAT, FEST, OH, and so on of the GEM chamber are evaluated during the QC7 test. Figure 4.39 (left) shows the GEM with GEB, and Figure 4.39 (right) shows the GEM after all of the electronic components have been placed on GEB.

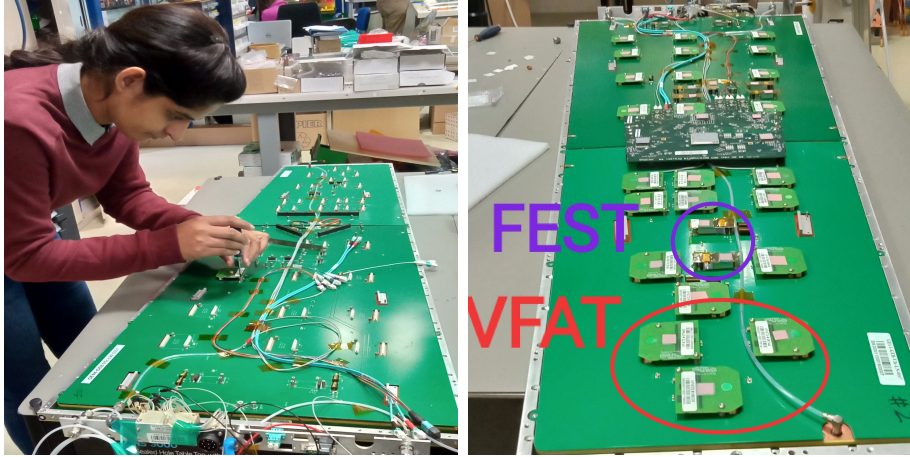


Figure 4.39: GEM chamber with GEB (left) and GEM chamber with all electronics components (right).

To analyze the chamber characteristics, the following tests are performed:

- **Connectivity and DAC threshold scan test:** During the connection test, all electronic components are examined. Each of the VFAT DACs (Digital to Analog Converter) is calibrated when the proper communication between the frontend and backend electronics has been established. The VFAT has a programmable register called a “DAC” that controls charge amplification, charge shaping, bias current, and other functions. The DAC register setting corresponds to the nominal value of the ADC (Analog to Digital Converter) are provided by the VFAT3 designers. The appropriate DAC value may change due to the radiation, and a calibration of the global VFAT DACs is performed for each VFAT. Figure 4.40 (left) shows the calibrated global DAC register values (CFG_THR_ARM) for each channel of the VFATS.

As there is no charge injection, the connection between the VFATs and the chamber readout board is checked during the threshold scan test by allowing 100 Hz of noise. The presence of noise hits at a fixed threshold is checked 100 times for each of the global threshold `CFG_THR_ARM` registers scanned values. The absence of noise hits in the channels indicates a disconnection of VFATs from the readout strips. This is because when the capacitance at the channel inputs is reduced, there is less intrinsic noise. Figure 4.40 (right) shows an example of a partially unplugged VFAT, where channels 55–127 behave differently than those between 0 and 54. The solution is as simple as re-plugging the VFAT.

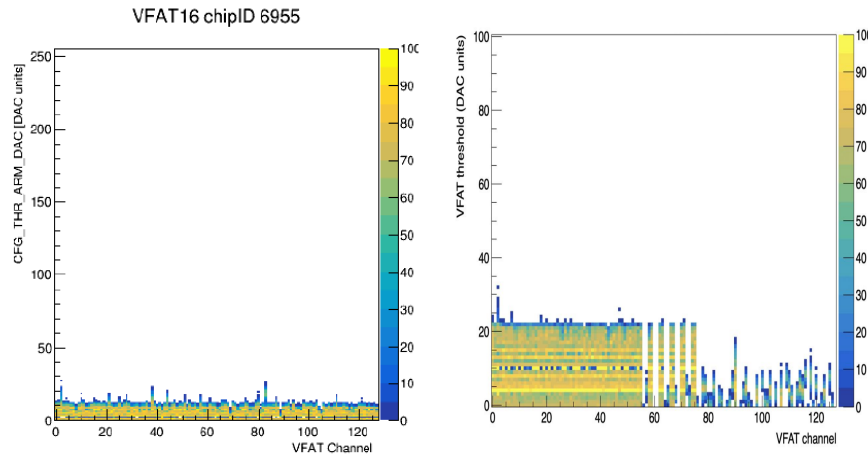


Figure 4.40: Recorded `CFG_THR_ARM` the DAC register value for every channel of the VFATS (left) and partially unplugged VFAT on the readout board (right).

- **S-bit rate scan:** This test confirms that no issues exist in the trigger path by comparing the noise rate recorded by the VFAT3 trigger as a function of the global thresholds by scanning the `CFG_THR_ARM` register values from 0 to 255. The `CFG_THR_ARM` DAC register value corresponding to 100 Hz noise is selected independently for each VFAT. To pass this test, the S-bit scan curve must be smoothly falling as shown in Figure 4.41 (left), and the threshold must be less than 50 for each of the 24 VFAT. The steeply falling part of the curve represents the reduction

in noise rate when the threshold is increased. As shown in Figure 4.41 (right) deviation from a smoothly falling curve indicates faulty channels or poorly connected VFAT. Poorly connected VFATs cause interrupted (“broken”) S-bit lines between the VFAT and the OH, which may be fixed by unplugging and replugging the electronic components and cleaning all connections.

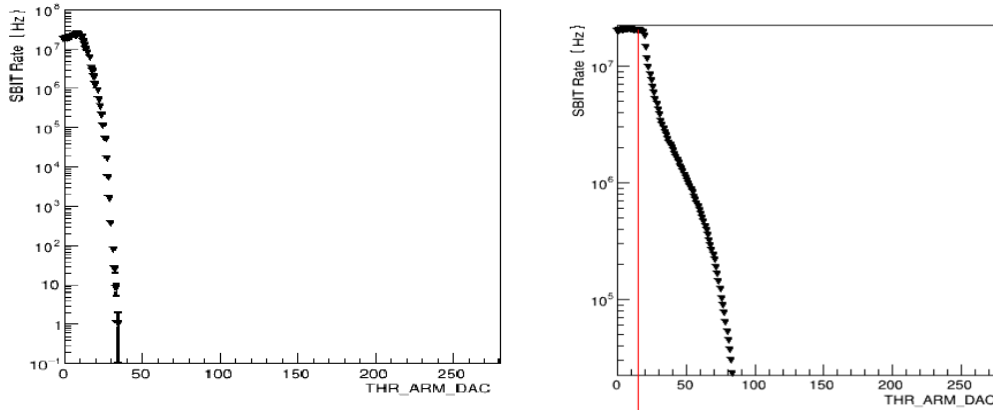


Figure 4.41: Noise rate as a function of the threshold for a normally behaving VFAT (left) and for misplaced/faulty VFAT (right).

- **S-curve scan:** In this test, a calibration pulse is sent to each channel and whether or not a hit is registered is recorded. The S-curve represents the over all response of a channel to the amount of injected charge. As the amount of the injected charge increased, the response of the channel also increased, and eventually, the efficiency will reach 100 percent. The shape of the graph is similar to S, therefore it is called S-curve as shown in Figure 4.42 (left). Sometimes few channels in some VFATs have zero efficiency for each DAC value, which has an effect on chamber efficiency as shown in Figure 4.42 (right). This test determines which channels are dead, and if there are less than three, they should be masked. If more than three channels per row are dead, this is unacceptable, and it must be resolved before rerunning the test.

The strip capacitance, which is inversely proportional to their area, de-

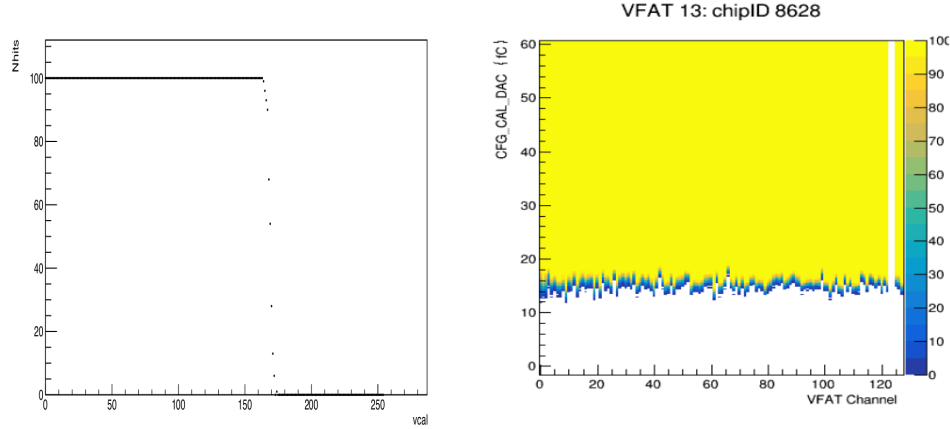


Figure 4.42: S-curve showing one VFAT having 100 percent efficiency for each DAC value (left) and VFAT having dead channels (right).

terminates the intrinsic noise. Because the noise level is higher for larger readout partitions, the frontend thresholds must be higher to achieve the same noise rate. This may be verified by plotting the mean (threshold) and sigma for the S-curves that correspond to a noise hit at 100 Hz. Figure 4.43 shows the mean and sigma for all VFATs of a single chamber. In Figure 4.43, the colorful box contains 50% of the channels, the dashed lines have 100% of the channels, and the mean (median) is represented by the circle (line) inside the colored box. Noise increases from the narrowest to the widest eta sector, with additional noise for the VFATs closest to the OH. To pass this test, the threshold and noise of each VFAT must be less than 10 fC and 1 fC, respectively.

- **Superchamber assembly:** The above mentioned tests were repeated once the cooling plate and thermal screen were installed, as illustrated in Figure 4.44 (left and middle). After passing the all electronics test, the superchamber is formed by coupling two GEM chambers, as shown in Figure 4.44 (right).
- **QC8 - cosmic ray test:** This test is performed to validate the performance of the superchamber and its electronics. The major purpose of this test is to measure the efficiency and time resolution of the chamber.

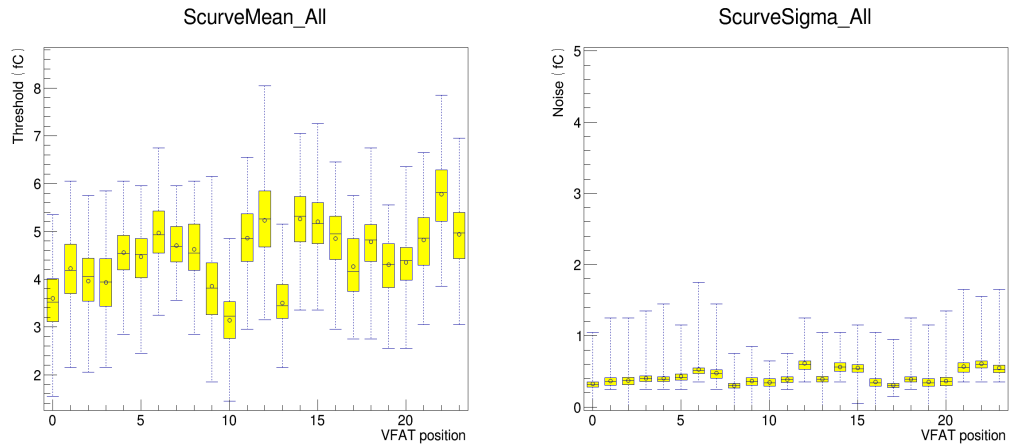


Figure 4.43: Threshold of VFAT3 (left) and Noise (right).



Figure 4.44: Chamber with cooling plate (left), with the thermal screen (middle), and supercavity assembly (right).

After passing all QC tests, the superchamber is installed in the Muon endcap.

4.7 GE1/1 installation and early commissioning

The first GE1/1 endcap was successfully installed in September and October 2019, with the second endcap installed in June 2020. Before installing the first endcap, a pilot exercise was carried out in July 2019 by installing two superchambers in order to validate the installation technique and obtain early feedback on the electronics response in the CMS environment. The improved grounding of the CMS experiment helped in improving the response of the electronic frontend, and lowering the Equivalent Noise Charge (ENC) per VFAT chip and per single channel. Figure 4.45 (left) shows the comparison of the ENC per VFAT derived from S-curves (sigma of the S-curves fit) taken in QC7, QC8, and at the early commissioning in the CMS experiment at P5. The boxes represent 50% of the channels and the bars 100% of them. Figure 4.45 (right) shows the distribution of the ENC derived from S-curves (sigma of the fit of the S-curves) taken in QC7, QC8, and P5. All the 3072 (28×124) channels of the chamber are contributing to the distributions [25].

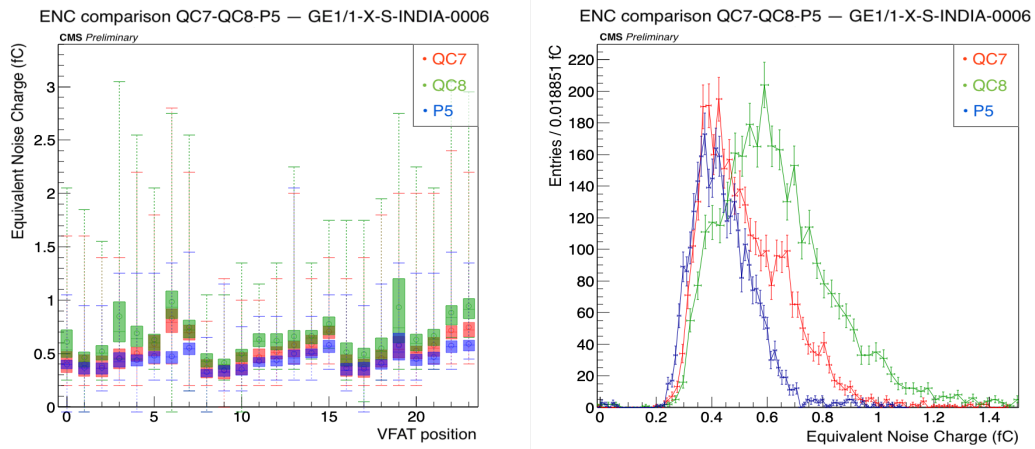


Figure 4.45: Comparison of Equivalent Noise Charge (ENC) of chamber GE1/1-X-S-INDIA-0006 measured at QC7 and QC8, and in the CMS experimental cavern after the installation (a) per VFAT (b) per readout channel [26].

This Chapter describes the assembly and QC tests (QC2, QC3, and QC4) of eight GEM chambers performed at the PU cite as well as further QC tests (QC5, QC6, QC7, and QC8) of other chambers performed at the CERN cite. All chambers have been successfully installed in the endcap of the CMS muon spectrometer. The physics analysis performed at 13 TeV is discussed in the next Chapter.

Bibliography

- [1] CMS Collaboration, “Combined results of searches for the standard model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV”, *Phys. Lett. B*, vol. 710 (2012) 26, [arXiv:1202.1488](#).
- [2] CMS Collaboration, “The CMS muon project: Technical Design Report”. Technical Design Report CMS. CERN, Geneva, (1997). <https://cds.cern.ch/record/343814>.
- [3] F. Sauli, “GEM: A new concept for electron amplification in gas detectors”, *Nucl. Instrum. Meth. A*, vol. 386 (1997) 531, [doi:10.1016/S0168-9002\(96\)01172-2](#).
- [4] CMS GEM Collaboration, “Physics motivations and expected performance of the CMS muon system upgrade with triple-GEM detectors”, *PoS EPS-HEP2015*, (2015) 225, [doi:10.22323/1.234.0225](#).
- [5] CMS Collaboration, “CMS Technical Design Report for the Level-1 Trigger Upgrade”, Technical Report CERN-LHCC-2013-011. CMS-TDR-12, (2013). <https://cds.cern.ch/record/1556311>.
- [6] N. Arkani-Hamed and S. Dimopoulos, “Supersymmetric unification without low energy supersymmetry and signatures for fine-tuning at the LHC”, *JHEP*, vol. 0506 (2005) 073, [doi:10.1088/1126-6708/2005/06/073](#).

-
- [7] G. Giudice and A. Romanino, “Split supersymmetry”, *Nucl. Phys. B*, vol. 699 (2004) 65, [doi:10.1016/j.nuclphysb.2004.11.048](https://doi.org/10.1016/j.nuclphysb.2004.11.048).
- [8] G. Giudice, M. Luty, H. Murayama, and R. Rattazzi, “Gaugino mass without singlets”, *JHEP*, vol. 9812 (1998) 027, [doi:10.1088/1126-6708/1998/12/027](https://doi.org/10.1088/1126-6708/1998/12/027).
- [9] L. Randall and R. Sundrum, “Out of this world supersymmetry breaking”, *Nucl. Phys. B*, vol. 557 (1999) 79, [doi:10.1016/S0550-3213\(99\)00359-4](https://doi.org/10.1016/S0550-3213(99)00359-4).
- [10] C.-Y. Chen and S. Dawson, “Exploring two Higgs doublet models through Higgs production”, *Phys. Rev. D*, vol. 87 (2010) 055016, [doi:10.1103/PhysRevD.87.055016](https://doi.org/10.1103/PhysRevD.87.055016).
- [11] J. Z. Neil Turok, “Electroweak baryogenesis in the two doublet model”, *Nucl. Phys. B*, vol. 358 (1991) 471, [doi:10.1016/0550-3213\(91\)90356-3](https://doi.org/10.1016/0550-3213(91)90356-3).
- [12] F. Palla, “Tracking Triggers for the High Luminosity LHC”, *PoS Vertex2013*, (2013) 034.
- [13] M. Brice, D. Abbaneo, and F. Sauli, “CMS Technology Transfer: Homeland security application of CMS GEMs”, CMS-PHO-MUON-2017-001, CERN, Geneva, (2017). <https://cds.cern.ch/record/2257210>.
- [14] E.R. Starling, “Status of the Readout Electronics for the Triple-GEM Detectors of the CMS GE1/1 System and Performance of the Slice Test in the 2017-18 LHC Run”, *PoS(TWEPP2018)132*, (2018).
- [15] P. Aspell et al., “VFAT3: A Trigger and Tracking Front-end ASIC for the Binary Readout of Gaseous and Silicon Sensors”, in *IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC)*, (2018).
- [16] PICMG, “MicroTCA Overview”, <https://www.picmg.org/openstandards/microtca>.

-
- [17] A. Colaleo, A. Safonov, A. Sharma, and M. Tytgat, “CMS Technical Design Report for the Muon Endcap GEM Upgrade”, CMS Technical Report CERN-LHCC-2015-012, CERN, Geneva, (2015). <https://cds.cern.ch/record/2021453>.
- [18] CMS Collaboration, “The Phase-2 Upgrade of the CMS Muon Detectors”, Technical Report CERN-LHCC-2017-012. CMS-TDR-016, CERN, Geneva, 2017. <https://cds.cern.ch/record/2283189>.
- [19] CMS Muon Collaboration, “Layout and Assembly Technique of the GEM Chambers for the Upgrade of the CMS First Muon Endcap Station”, *Nucl. Instrum. Meth. A*, vol. 918 (2018) 67, [doi:10.1016/j.nima.2018.11.061](https://doi.org/10.1016/j.nima.2018.11.061), [arXiv:1812.00411](https://arxiv.org/abs/1812.00411).
- [20] CMS Collaboration, “Production and quality control of the new chambers with GEM technology in the CMS muon system”, Conference Report CMS-CR-2018-081, CERN, Geneva, (2018). <https://cds.cern.ch/record/2629842>.
- [21] CMS Collaboration, “Upgrade of the CMS muon spectrometer in the forward region with the GEM technology”, Conference Report CMS-CR-2018-325, CERN, Geneva, (2018). <https://cds.cern.ch/record/2647110>.
- [22] A. Shah et al., “Layout and Assembly Technique of the GEM Chambers for the Upgrade of the CMS First Muon Endcap Station”, Detector Note, CMS-DN-2015-020, CERN, Geneva, (2015). <https://cds.cern.ch/record/2645530>.
- [23] CMS Muon Collaboration, “Performance of prototype GE1/1 chambers for the CMS muon spectrometer upgrade”, *Nucl. Instrum. Meth. A*, vol. 972 (2020) 164104, [doi:10.1016/j.nima.2020.164104](https://doi.org/10.1016/j.nima.2020.164104), [arXiv:1903.02186](https://arxiv.org/abs/1903.02186).
- [24] CMS Muon Collaboration, “GEM detectors for the Upgrade of the CMS Muon Forward system”, *J. Phys. : Conf. Ser.*, vol. 1390 (2019) 012116, [doi:10.1088/1742-6596/1390/1/012116](https://doi.org/10.1088/1742-6596/1390/1/012116).

- [25] “Proceedings, 15th Topical Seminar on Innovative Particle and Radiation Detectors (IPRD19): Siena, Italy, October 14-17, 2019”, vol. 15 (2020).
- [26] CMS Collaboration, “GEM DPG Public Page”.
<https://twiki.cern.ch/twiki/bin/view/CMSPublic/GEMDPGPublic>, (2021).

Chapter 5

Measurement of the production cross section for $Z + b$ -jets in proton-proton collisions at $\sqrt{s} = 13$ TeV

5.1 Introduction

As discussed in Chapter [1](#), the measurement of $Z + b$ -jets processes can be used to test pQCD predictions by comparing the measured cross sections with theoretical predictions. The $Z + b$ -jets measurements help to extract PDFs for the b quarks [\[1\]](#) and constraint the PDF uncertainty in the measurement of the W boson mass [\[2\]](#). These processes are the dominant background for Higgs boson production in association with a Z boson (ZH , $H \rightarrow b\bar{b}$) in the standard model. These processes are also a source of background in many BSM scenarios, such as the production of supersymmetric Higgs bosons in association with b quark, and new generations of heavy quarks decaying to a Z boson and a b quark, etc. Therefore, the measurement of $Z + b$ -jets processes is one of the important objectives of collider experiments.

The CDF [3] and D0 [4] Collaborations have measured the cross section of $Z + b$ -jets processes using proton-antiproton collisions data at $\sqrt{s} = 1.96$ TeV. At the CERN LHC, the measurements of the $Z + b$ -jets processes have been performed at 7 TeV by the CMS [5] and ATLAS [6] Collaborations, at 8 TeV by CMS [7], and at 13 TeV by ATLAS [8], using the pp collisions data.

In this thesis, we present the first measurement of $Z + b$ -jets cross section at 13 TeV using data collected by the CMS experiment. The integrated cross sections together with differential distributions of the $Z + \geq 1$ b-jet processes are measured and compared with various MC predictions.

5.2 Signal process

The Z boson accompanied by b -jets can originate from quark-gluon, quark-antiquark, and gluon-gluon interactions in pp collisions. Examples of Feynman diagrams of the signal process with quark-gluon, quark-antiquark, and gluon-gluon interactions for $Z + 1$ b-jet and $Z + 2$ b-jets final states are shown in Figure 5.1. The decay of the Z boson can occur in several ways. It prefers to decay hadronically, with almost 70% of the decay being to strong interacting particles. Invisible decays to neutrinos account for 20% of the decays and the remaining are leptonic decays to electrons, muons, and taus in almost equal amounts. It is experimentally very difficult to detect the hadronic channels among the abundant QCD background but leptonic decays offer a clean signature at hadron colliders. For the present analysis, the signal event comprises two oppositely charged leptons ($e^+e^-/\mu^+\mu^-$) originating from the decay of the Z boson and at least one b quark jet.

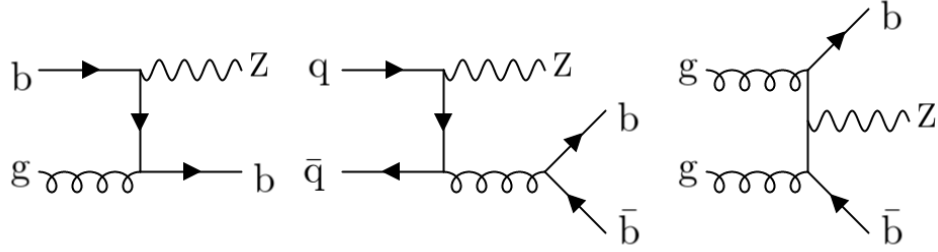


Figure 5.1: Examples of Feynman diagrams for $Z + 1$ b-jet (left) and $Z + 2$ b-jets (middle and right).

5.3 Background processes

Background events are those that can mimic the characteristics of the signal but are not real signal events. In the present analysis, the final state consists of two leptons and b-jets. There are many processes that can produce similar final state as discussed below.

- **Z + jets processes:** The $Z + \text{jets}$ processes include signal events as well as background events from $Z + \text{c-jet}$ and $Z + \text{light-jet}$ as shown in Figure 5.2 in which a jet is initiated by a charm quark (c-jet) and a light quark or gluon (light-jet), respectively is misidentified as a b-jet.

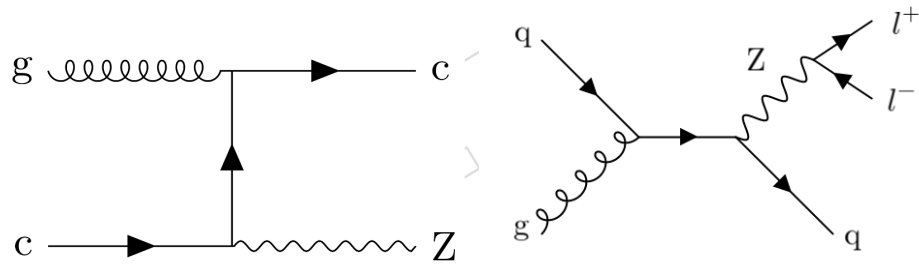


Figure 5.2: Examples of Feynman diagrams for $Z + \text{c-jets}$ (left) and $Z + \text{jets}$ (right).

- **$t\bar{t}$ process:** The top quark-antiquark pairs are produced via quark annihilation ($q\bar{q} \rightarrow t\bar{t}$) and gluon fusion ($gg \rightarrow t\bar{t}$) processes, as shown in Figure 5.3. The top quark can decay into a W boson and a b quark jets, with a branching

fraction of more than 95% [9, 10], as shown in Figure 5.4. These W bosons can further decay either leptonically or hadronically into jets. So, $pp \rightarrow t\bar{t} \rightarrow W^+W^-b\bar{b} \rightarrow l^+l^-b\bar{b} + E_T^{\text{miss}}$ process is the dominant non-(Z+jets) background source for the Z+b-jets process.

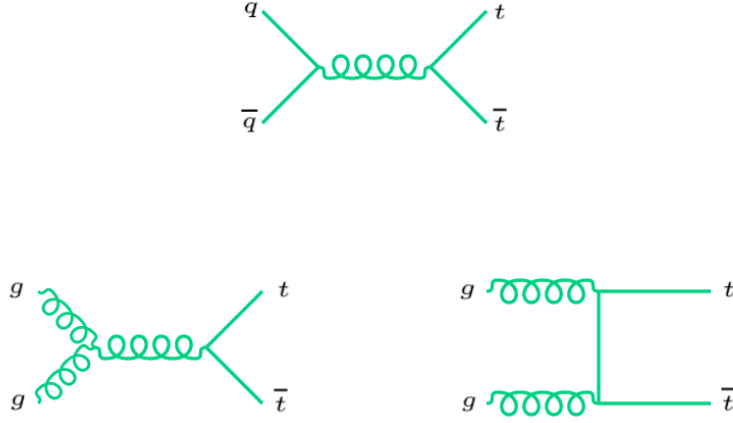


Figure 5.3: Examples of Feynman diagrams for $t\bar{t}$ process.

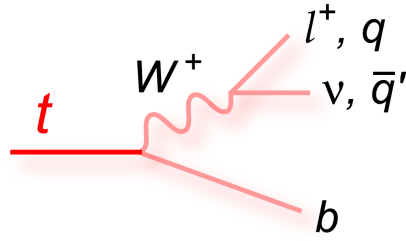


Figure 5.4: Decay scheme of top quark.

- **Diboson processes:** The major diboson processes which can mimic the signal are ZZ, WZ, and WW processes. Tree-level diagrams illustrating the associated production of bosons are shown in Figure 5.5. Among these diboson processes, the following decay modes can lead to two leptons in the final state: $Z \rightarrow ll$, $W \rightarrow l\nu$. Diboson events may mimic the Z+b-jets final state when missing the reconstruction of one or more leptons or when one W or Z boson

decays hadronically into $q\bar{q}$ pairs (in particular, a Z boson may decay into a $b\bar{b}$).

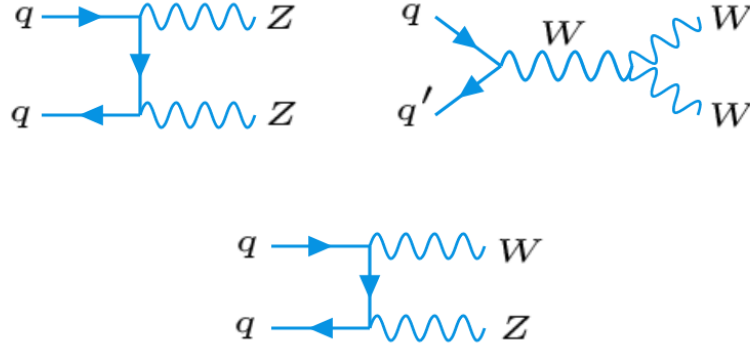


Figure 5.5: Examples of Feynman diagrams for various diboson production processes.

- Single top processes:** Another source of contamination for the $Z + b$ -jets final state comes up from the single top production in the s-channel, t-channel, and tW-channel as shown in Figure 5.6. The s-channel results in the production of a time-like W boson which then decays into a top and a b quark, while the t-channel involves the exchange of a space-like W boson. The tW-channel mode leads to the production of a top quark in association with a W boson. In all cases, the final state contains at least one W boson and at least one b-jet, and can thus reproduce the signal process. In the first two cases, the single top is produced with a b-jet, leading to a final state with an isolated lepton and at least two b-jets, and thus they mimic the $Z + b$ -jets final state when a secondary lepton is misidentified. In the signal process, single top background contributions are significantly low as compared to $t\bar{t}$ and $Z + c$ -jet processes.
- W + jets process:** The W + jets events as shown in Figure 5.7 with $W \rightarrow \mu\nu$ having a branching fraction of about 11% can mimic the signal when one of the associated jets (mainly the leading one) is misidentified as an additional

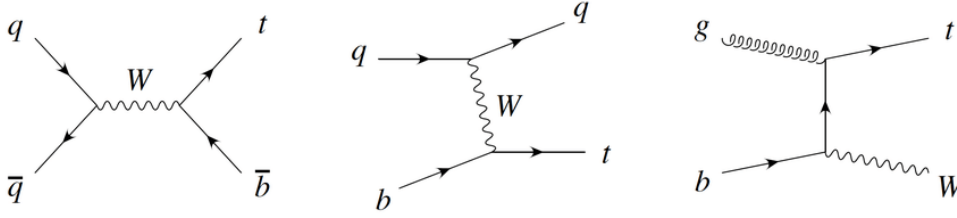


Figure 5.6: Examples of Feynman diagrams for the single top quark production in LO s-channel (left), NLO t-channel (middle), and NLO tW (right) subprocess.

lepton and another jet as a b-jet. The $W + \text{jets}$ contribution to the signal process is negligible.

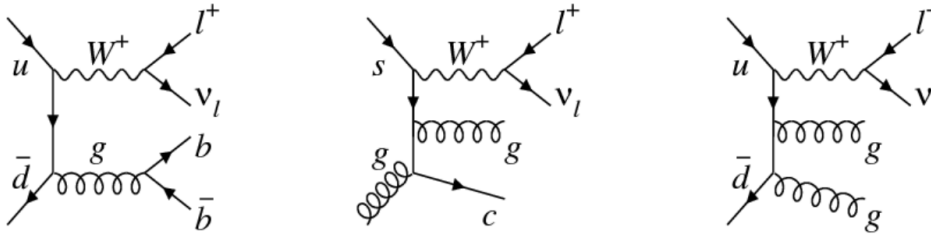


Figure 5.7: Examples of Feynman diagrams for the $W + \text{jets}$ production process.

5.4 Data and Monte Carlo samples

The present analysis is performed using the pp collision data and various MC samples as discussed below.

- **Data samples:** This is the first measurement of the $Z + \text{b-jets}$ process performed using a data sample of pp collisions at a $\sqrt{s} = 13 \text{ TeV}$ collected by the CMS experiment during the 2016 (35.9 fb^{-1}), 2017 (41.5 fb^{-1}), and 2018 (59.7 fb^{-1}) data-taking periods with an integrated luminosity of 137 fb^{-1} . The mean numbers of interactions per bunch crossing are 23, 33, and 32 for the 2016, 2017, and 2018 data-taking periods, respectively as shown in Figure [5.8](#).
- **Monte Carlo samples:** Several MC event generators with different versions, tunes, and PDFs are used to simulate signal and background processes.

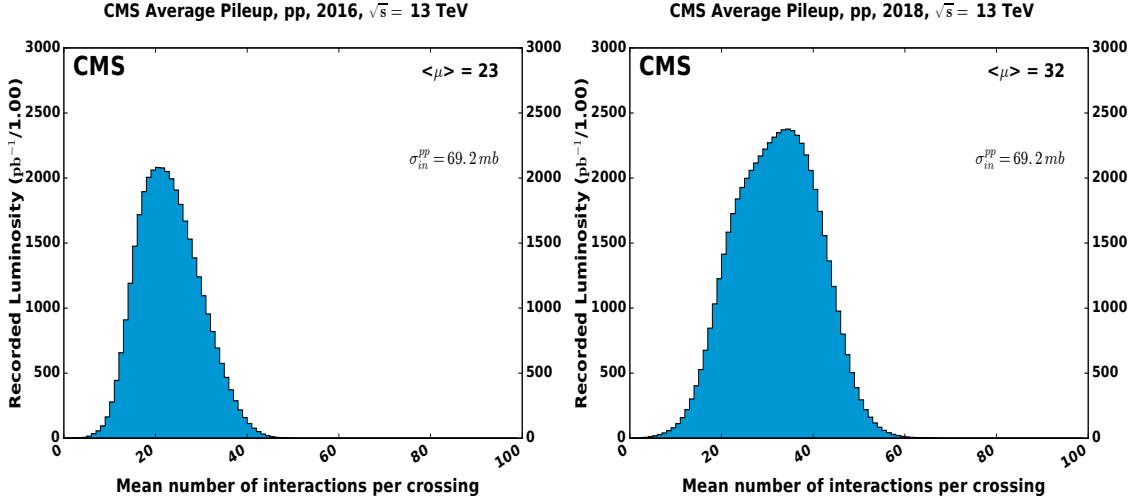


Figure 5.8: Mean number of interactions per bunch crossing for the 2016 and 2018 data-taking periods [11].

These MC samples are employed to optimize event selection, validate data-MC performance, estimate the background, and correct the measurement for the detector effects.

- **Signal samples:** The Z with an exclusive jet multiplicity up to 2 is simulated at NLO precision by MADGRAPH5_aMC@NLO (MG5_aMC) [12] version 2.3.2.2 for 2016 data and version 2.6.0 for the 2017–2018 data with the FXFX [13] matching between the jets from ME calculations and PSs. A second sample is produced with MG5_aMC version 2.2.2 (2.4.2) for the 2016 (2017–2018) data-taking period using LO MEs for $pp \rightarrow Z + n\text{jets}$ ($n \leq 4$), with the MLM matching scheme [14].

The MG5_aMC generator is interfaced with PYTHIA v8.212 (v8.230) [15] for PS, hadronization, and the UE simulation using tune CUETP8M1 [16] (CP5 [17]) for the 2016 (2017–2018) data-taking period. Matching between the ME generators and the PS jets is done with the matching scales set at 19 and 30 GeV for the MG5_aMC LO and NLO, respectively. The NNPDF 3.0 LO (NLO) and NNPDF 3.1 NNLO (NNLO) PDF sets [18] are used for MG5_aMC LO (NLO) for the 2016 and 2017–2018 data-taking

periods, respectively. In the ME calculation, the value of the NNPDF 3.0 (NNPDF 3.1) α_s at the Z boson mass is set to 0.130 (0.118) for MG5_aMC at LO and is set to 0.118 (0.118) for MG5_aMC at NLO for the 2016 (2017–2018) data-taking period.

A third inclusive sample has been produced with SHERPA v2.2.4 [19] to generate $pp \rightarrow Z + n \text{ jets}$ events, with $n \leq 2$ at NLO and $n = 3, 4$ at LO. The merging with the SHERPA PS is done with a matching scale of 20 GeV. The NNPDF 3.0 NLO PDF and a dedicated set of tuned PS parameters developed by the SHERPA authors are used.

All three signal samples use the 5FS with b quark density allowed in the initial state via the b and c quark PDF, with the b and c quarks typically considered massless in the ME calculation.

These $Z + \text{jets}$ samples are inclusive in the jet flavors (b, c, and light). In the simulation, the classification of reconstructed $Z + \text{jet}$ events into $Z + \text{b-jet}$, $Z + \text{c-jet}$, and $Z + \text{light-jet}$ categories are based on the flavors of reconstructed jets. They are classified as b- or c-jets if they are matched to the generated b or c hadrons, respectively. The matching is done using the ghost association procedure in which the modulus of the hadron four-momentum is set to a small number, retaining only the directional information [20]. In the case when both c and b hadrons are matched, the jet is considered as a b-jet. Based on reconstructed jets with defined flavors, events are classified as $Z + \text{b-jet}$ events if they contain at least one b-jet. Of the remaining events, those that contain at least one c hadron are considered as $Z + \text{c-jet}$ events, and those that contain neither b nor c hadrons are classified as $Z + \text{light-jet}$ events. Hence, $Z + \text{jets}$ processes include both signal events ($Z + \text{b-jet}$) as well as background events ($Z + \text{c-jet}$ and $Z + \text{light-jet}$).

- **Background samples:** The $t\bar{t}$ ($W + \text{jets}$) process is generated at NLO

(LO) by POWHEG v2.0 [21–23] (MG5_aMC) and MG5_aMC (MG5_aMC) for the 2016 and 2017–2018 data-taking periods, respectively. The $t\bar{t}$ ($W + \text{jets}$) process generators are interfaced with PYTHIA for parton showering and hadronization using tune CUETP8M2T4 [24] (CUETP8M1) and CP5 (CP5) for the 2016 and 2017–2018 data-taking periods, respectively. The single top quark t-channel and tW-channel processes are simulated at NLO with POWHEG v2.0 and POWHEG v1.0, respectively, whereas the s-channel process is simulated with MG5_aMC at NLO, both interfaced with PYTHIA.

For $t\bar{t}$ ($W + \text{jets}$) process PYTHIA v8.212 (v8.212), v8.226 (v8.230), and v8.230 (v8.230) are used in 2016, 2017, and 2018, respectively. The PYTHIA v8.212, v8.212, and v8.205 are used for single top quark t-channel, tW-channel, and s-channel processes, respectively, in 2016, while PYTHIA v8.226 and v8.230 are used in 2017 and 2018, respectively, for all single top quark processes. The diboson background processes are simulated at LO using PYTHIA v8.205, v8.226, and v8.230 for 2016, 2017, and 2018, respectively.

All $Z + \text{jets}$ processes are scaled to NNLO accuracy in QCD based on the $Z + \text{jets}$ cross sections calculated with FEWZ v3.1 [25]. The $W + \text{jets}$, single top quark, and diboson predictions are scaled to NNLO accuracy [26,27]. The prediction for $t\bar{t}$ production is normalized to NNLO calculations in pQCD including resummation of NNL logarithmic soft-gluon terms with TOP++ 2.0 [28,34].

The MC samples are processed with the full CMS detector simulation based on GEANT4 [35]. To compare data with MC simulation, the event yields of MC simulated samples are normalized by multiplying a weight factor (w). This weight factor is calculated from the production cross section of the MC samples (σ), integrated luminosity (L) of collision data, and the total number of MC simulated events (N)

for the corresponding process and is defined as follows:

$$w = \frac{\sigma \cdot L}{N}. \quad (5.1)$$

5.5 Observables

In this analysis, the differential cross section for $Z + \geq 1$ b-jet events is measured as a function of various kinematic observables: the transverse momentum of the dilepton (p_T^Z); the highest p_T (leading) b-jet ($p_T^{\text{b-jet}}$); the absolute pseudorapidity of the b-jet ($|\eta^{\text{b-jet}}|$); the azimuthal angle difference between the Z boson and the b-jet ($\Delta\phi^{(Z, \text{b-jet})}$); the rapidity difference between the Z boson and the b-jet ($\Delta y^{(Z, \text{b-jet})}$); and the angular correlation variable, which is defined as the separation in the pseudorapidity and azimuthal angle between the Z boson and the b-jet [$\Delta R^{(Z, \text{b-jet})} = \sqrt{(\Delta\eta^{(Z, \text{b-jet})})^2 + (\Delta\phi^{(Z, \text{b-jet})})^2}$]. Here, the normalized differential cross section is also measured which can better constraint the theoretical predictions because of the cancellation of the correlated systematic uncertainties. The p_T^Z , $p_T^{\text{b-jet}}$, and $|\eta^{\text{b-jet}}|$ observables are directly sensitive to the b quark PDF and details of the pQCD calculations. Searches for physics processes BSM in lorentz-boosted topology events depend on precise knowledge of the Z boson p_T distribution. The distribution of $\Delta\phi^{(Z, \text{b-jet})}$ is also sensitive to the presence of boosted particles decaying into a Z boson and b quarks. The angular separation is also sensitive to ISR, FSR, and MPIs where extra jet(s) affect the corresponding variables [9].

To understand the dynamics of $Z + \geq 1$ b-jet and $Z + \geq 2$ b-jets, the ratio of their differential cross section is also measured as a function of p_T and $|\eta|$ of the leading b-jet. These measurements allow a further reduction in various experimental uncertainties such as those related to the jet energy scale measurement. The $Z + \geq 2$ b-jets analysis has been performed by other collaborative group and details can

be found in Ref. [36].

5.6 Events selection criteria

The measurement has a Z boson ($\rightarrow ee/\mu\mu$) and b-jets in the final state. The measurement with this final state requires the number of conditions on kinematic and identification properties of the final state objects as discussed below.

- Trigger:** The collision events required to pass single lepton triggers as listed in Table 5.1. At the trigger level, the events are required to have at least one electron (muon) candidate with a minimum p_T of 27, 32, and 32 GeV (24, 27, and 24 GeV) during 2016, 2017, and 2018 data-taking periods, respectively. The single electron and muon trigger efficiencies are measured by the CMS collaboration using a data-driven method based on the tag-and-probe (T&P) technique [37] which exploits $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ events. Figure 5.9 shows the efficiency of the electron trigger [38] for the 2018 and of the muon trigger [39] for the three years. The single electron and muon trigger efficiency is about 80% and exceeds 90% over the full η range, respectively, at the efficiency plateau. The differences in trigger efficiencies between data and simulation are corrected by applying the corresponding scale factors to MC simulation. The scale factors are calculated from the ratio of the efficiencies estimated on data and MC simulation.
- Vertex selection:** Each event is required to contain at least one well reconstructed PV in order to select the good collision events. So, events are removed which are not coming from pp collisions, such as those from beam-gas interactions or beamscraping in the transport system near the interaction point. The reconstructed primary vertices are required to have $|z|$ position within 24 cm around beamspot. Vertex is needed to have a radial position within 2 cm of the beamspot axis and a vertex fit exceeding four degrees of freedom. Out

Table 5.1: HLT triggers path and their p_T thresholds used to collect events from a given sample for the 2016, 2017, and 2018 data-taking periods.

Year	channel	Trigger name	p_T threshold
2016	ee	HLT_Ele27_WPTight_Gsf_v*	> 27 GeV
	$\mu\mu$	HLT_IsoMu24_v* or HLT_IsoTkMu24_v*	> 24 GeV
2017	ee	HLT_Ele32_WPTight_Gsf_L1DoubleEG_v* and hltEGL1SingleEGOrFilter	> 32 GeV
	$\mu\mu$	HLT_IsoMu27_v*	> 27 GeV
2018	ee	HLT_Ele32_WPTight_Gsf_v*	> 32 GeV
	$\mu\mu$	HLT_IsoMu24_v*	> 24 GeV

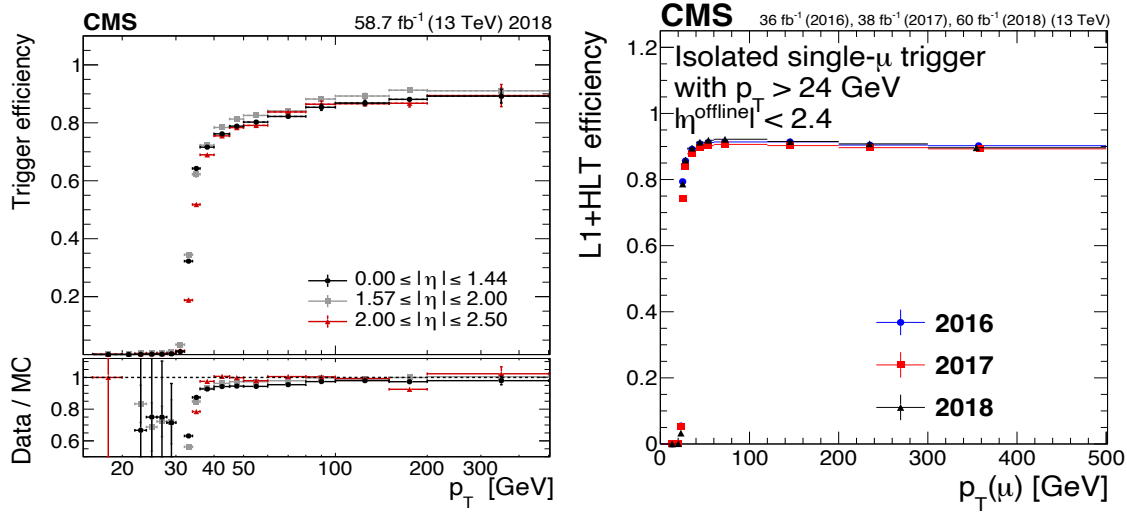


Figure 5.9: Single electron trigger efficiency with respect to the offline reconstruction, as a function of the electron p_T , for different η regions using the 2018 dataset (left). The bottom panel shows the ratio of data over simulation events. The efficiency measurements combine the effects of the L1 and HLT triggers [38]. Single muon trigger efficiency as a function of muon p_T (right) [39]. Only statistical uncertainties are shown in both plots.

of the selected vertices, the largest value of summed physics-object p_T^2 is the primary pp interaction vertex corresponding to the hard scattering process. In this analysis, events are required to have at least one vertex satisfying above mentioned selections.

The simulated events are reconstructed using the same algorithms as used for the data and include additional interactions per bunch crossings, referred to as PU. Simulated events are reweighted with a weight factor (w_{PU}) so that the

PU distribution reproduces the one observed in the data. This weight factor is estimated to obtain “n” true interactions by taking the ratio of probabilities in data ($P_{\text{data}}(n)$) and MC ($P_{\text{MC}}(n)$), and is given as:

$$w_{\text{PU}}(n) = \frac{P_{\text{data}}(n)}{P_{\text{MC}}(n)}. \quad (5.2)$$

Figure 5.10 shows the vertex multiplicity distributions in which data is compared with MC simulation, after pileup reweighting, for the 2016 and 2018 years. There is a discrepancy in data and MC simulation in the vertex multiplicity even after reweighting. The effect of this discrepancy is evaluated as a part of the systematic, which will be discussed later in the Section 5.8.1.

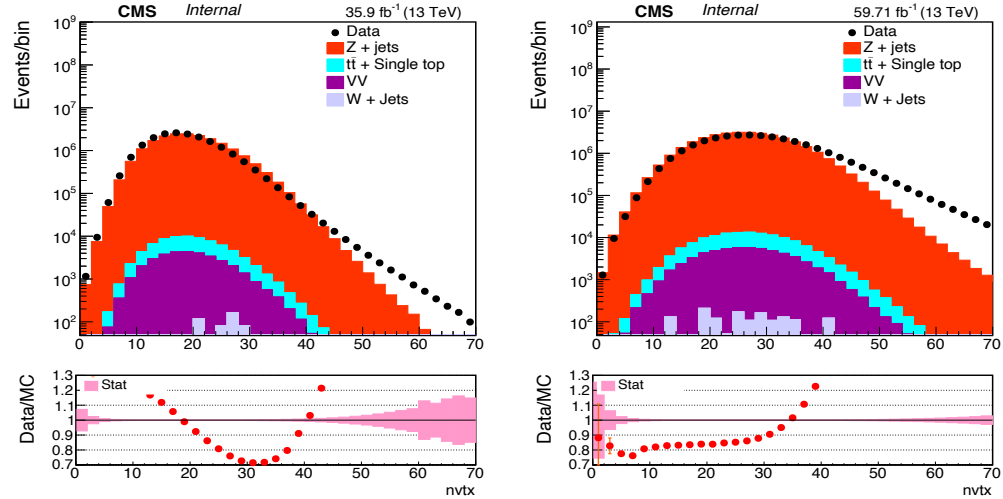


Figure 5.10: Distributions of the number of reconstructed primary vertices n_{PV} , after the application of pileup reweighting for the 2016 (left) and 2018 (right) years. The bottom panel shows the ratio of data over total MC simulated events. The band on ratio plots include only the statistical uncertainties.

- **Electron selection:** The electron candidates in the event are required to originate from the selected PV. For prompt electrons (produced at the PV), background sources can originate from photon conversions, hadrons misidentified as electrons (e.g. jets with large electromagnetic components), and secondary electrons from semileptonic decays of b or c quarks. In order to maximize the

identification efficiency and minimize the fake rate, the electrons are identified by a set of selection of tight working point as listed in Table 5.2.

The tracker-related variables $|\Delta\eta_{\text{in}}^{\text{seed}}|$, $|\Delta\phi_{\text{in}}|$, and $|1/E - 1/p|$ are used for the identification of electrons. Purely calorimetric observables $\sigma_{\text{in}\eta}$ and H/E ratio are used to separate genuine electrons (prompt electrons or electrons from photon conversions) from hadrons misidentified as electrons. The $\sigma_{\text{in}\eta}$ variable describes the lateral extension of the shower along η . The H/E variable is defined as the ratio of the energy deposited in the HCAL in a cone of radius $\Delta R = 0.15$ around the SC direction to the energy of the electron candidate.

The isolation variable (defined in Equation 3.2 of Chapter 3) is built considering a cone in $\Delta R = 0.3$ around the reconstructed electron to reject electrons that are embedded in jets. For electrons, the relative isolation variable I_{rel} is E_{T} dependent; $I_{\text{rel}} < 0.021$ (0.040) for 25 GeV and $I_{\text{rel}} < 0.015$ (0.029) for 35 GeV in the barrel (endcap) region [40], respectively. Because the larger the energy of the incoming electrons, the larger the amount of energy spread around its direction in the various subdetectors. To reject secondary electrons produced in conversions of photons in the tracker material, the number of missing hits in the innermost layers of the tracker should not exceed one. Also, longitudinal ($|d_z|$) and transverse ($|d_{xy}|$) impact parameters of the electron with respect to the beam line are required to be < 0.10 cm and < 0.05 cm for the barrel ($|\eta| < 1.48$) and $|d_z| < 0.20$ cm and $|d_{xy}| < 0.10$ cm for endcap ($1.48 < |\eta| < 3.0$) regions, respectively.

For the tight working point, the efficiency for genuine electrons is around 70%. There are small residual differences in the trigger, identification, and isolation efficiencies between data and simulation, which are measured using T&P methods and taken into account by applying scale factors as a function of their p_{T} and $|\eta|$ to the simulated events. Due to the ECAL-Tracker momentum

combination when applying the scale and smearing, the energy scale shifts causing a discontinuity. To fix this post processing energy correction is also applied.

For this analysis, after identification and all correction, leading and subleading electrons are required to have $p_T > 35$ GeV and $p_T > 25$ GeV, respectively, and supercluster $|\eta| < 2.4$ with EB-EE gap removal ($|\eta| < 1.4442$ or $|\eta| > 1.556$). The lowest unscaled single electrons have a p_T threshold of 32 GeV, and chosen electrons are well inside the plateau of the trigger efficiency curve.

Table 5.2: Cut-based electron identification requirements for the tight working point in the barrel and in the endcaps.

Variable	Barrel (tight WP)	Endcaps (tight WP)
$\sigma_{i\eta i\eta}$	<0.010	<0.035
$ \Delta\eta_{\text{in}}^{\text{seed}} $	<0.0025	<0.005
$ \Delta\phi_{\text{in}} $	<0.022 rad	<0.024 rad
H/E	$<0.026 + 1.15 \text{ GeV}/E_{\text{SC}}$	$<0.019 + 2.06 \text{ GeV}/E_{\text{SC}}$
	$+0.032\rho/E_{\text{SC}}$	$+0.183\rho/E_{\text{SC}}$
I_{combined}/E_T	$<0.029 + 0.51 \text{ GeV}/E_T$	$<0.0445 + 0.963 \text{ GeV}/E_T$
$ 1/E - 1/p $	$<0.16 \text{ GeV}^{-1}$	$<0.0197 \text{ GeV}^{-1}$
Number of missing hits	≤ 1	≤ 1
Pass conversion veto	Yes	Yes

- Muons selection:** Muon candidates are reconstructed using the PF algorithm. The efficiency to reconstruct and identify muons is greater than 96%. Matching muons to tracks measured in the silicon tracker results in a relative transverse momentum resolution of 1% in the barrel and 3% in the endcaps, for muons with p_T up to 100 GeV. The muons are identified by a set of selection of tight working point as described in Table 5.3 for Run2. A tight muon requires to be a global muon, the tracker muon must have segment matching in at least two of the muon stations to suppress punch-through and accidental track-to-segment matches. A tracker track must have hits from at least six layers of the inner tracker including at least one pixel hit to suppress muons from decay in flight. To suppress hadronic punch-through and muons from

decays in flights of π and K mesons, an additional requirement on the χ^2/ndof for the global muon track fit to be less than 10 is added and includes at least one hit from the muon system. A tight muon must be compatible with the PV to suppress contamination due to cosmic muons, the threshold on the muon tracker track transverse impact parameter ($|d_{xy}|$) and longitudinal distance ($|d_z|$) is required to be less than 2 mm and 5 mm, respectively. In order to reduce the misidentification rate, I_{rel} is required to be less than 0.15 in a cone in $\Delta R = 0.4$ around the reconstructed muon and muon isolation efficiency is $> 95\%$.

The muon transverse momentum scale and resolution corrections [41] are also applied in both data and MC to remove biases of the muon momentum from detector misalignments or magnetic field uncertainties. This provides corrections as a function of muon charge, p_T , and η . The small residual differences in the muon trigger, identification, and isolation efficiencies between data and simulation are measured using T&P methods and are taken into account by applying scale factors as a function of their p_T and $|\eta|$ to simulated events.

After identification and all corrections, leading and subleading muons are required to have $p_T > 35 \text{ GeV}$ and $p_T > 25 \text{ GeV}$, respectively, within $|\eta| < 2.4$. The p_T thresholds are chosen for consistency with the electron channel and selected muons are well inside the plateau of the muon trigger efficiency curve.

- **Missing transverse momentum:** The missing transverse momentum vector (MET) $\vec{p}_T^{\text{miss}}$ is computed as the negative vector sum of the transverse momenta of all the PF candidates in an event and its magnitude is denoted as p_T^{miss} [42]. The p_T^{miss} is modified to include corrections to the energy scale of the reconstructed jets in the event. In this analysis, MET is used to reduce contributions from $t\bar{t}$ by requiring $p_T^{\text{miss}} < 50 \text{ GeV}$.
- **Z boson selection:** To select a Z candidate, two same flavored leptons with

Table 5.3: Muon tight identification criteria.

Selection	Cut
Global Muon	yes
PF muon	yes
Track fit χ^2/ndof	< 10
Muon chamber hits	≥ 1
Muon station segments	≥ 2
$ d_{xy} $	< 2
$ d_z $	< 5
Pixel hits	> 1
Tracker layers hits	> 5

opposite charges are combined. Here, leading and subleading leptons must have $p_T > 35 \text{ GeV}$ and $p_T > 25 \text{ GeV}$, respectively. Further, the dilepton invariant mass is required to have in the range of $[71, 111] \text{ GeV}$. With the above discussed selection criteria, the total 77 million and 34 million Z bosons candidate events are found for the dimuon and dielectron final states, respectively.

- Jets selection:** Jets are clustered from PF candidates using the infrared- and collinear-safe anti- k_T algorithm with a distance parameter of 0.4, as implemented in the FASTJET package [43]. The jet momentum is the vector sum of all particle momenta in the jet and is found from simulation to be within 5–10% of the true momentum over the entire p_T spectrum and detector acceptance. The jets are corrected for electronic noise and pileup, non-uniformities of the CMS Calorimeter, and its non-linear response to the particles with different energies and data/MC residual corrections as already discussed in Chapter 3. The JER amounts typically to 16% at 30 GeV and 8% at 100 GeV. After these corrections, jets are required to pass tight PF identification criteria, as listed in Table 5.4, to reduce the contamination from misreconstruction and instrumental noise. The jet identification efficiency is observed to be greater than 98–99% for the entire η and p_T range. In order to reduce the contamination from pileup jets, jets are required to be matched to the same PV as

the leptons. Further, to remove the jet coming from the PU, the jet MVA pileup ID discriminant is required to be greater than 0.61 for the jet with $p_T < 50$ GeV [44]. For jets with $p_T > 50$ GeV, the pileup jet ID discriminator is not applied, since here the pileup event contribution is negligible [45]. The scale factors related to pileup jet MVA ID are also applied in order to account for possible differences between data and MC.

The jets are selected with $p_T > 30$ GeV, $|\eta| < 2.4$, and the overlap between the jets and a lepton from the Z boson decay (isolated lepton) is removed by requiring a minimum distance of $\Delta R > 0.4$ between them.

Table 5.4: Jet identification criteria in 2016 and 2017–2018 data-taking periods.

Selection	Cut (2016)	Cut (2017–2018)
Neutral Hadron Fraction	< 0.90	< 0.90
Neutral EM Fraction	< 0.90	< 0.90
Number of Constituents	> 1	> 1
Charged Hadron Fraction	> 0	> 0
Charged Multiplicity	> 0	> 0
Charged EM Fraction	< 0.99	—

- b/c-jets selection:** After jets selection, a neural network algorithm (DEEPCSV) is used to discriminate b-jets from c- and light-jets. In the DEEPCSV algorithm, the significance of the secondary vertex displacement is combined with track-based information in a machine-learned multivariate analysis to increase the efficiency compared with cut-based algorithms. In this analysis, tight working points are used for the identification of the b-jet which corresponds to the cut of 0.8953, 0.8001, and 0.7527 on the discriminator values for the 2016, 2017, and 2018 data-taking periods, respectively. Similarly, DEEPCSV tight working points of c tagger are used to discriminate c-jets from b- and light-jets, which corresponds to the cut of $C_{vsL} > 0.59$ & $C_{vsB} > 0.05$ for 2016, $C_{vsL} > 0.8$ & $C_{vsB} > 0.1$ for 2017, and $C_{vsL} > 0.66$ & $C_{vsB} > 0.1$ for 2018.

The tagging efficiency of Z + b-jet (Z + c-jet) events is estimated as the fraction of events that pass the b (c) hadron and b tagging (c tagging) requirement over the number of events that pass just the b (c) hadron requirement. In simulations, the b-jets tagging efficiencies with the tight operating point are 50–60% and misidentification probabilities are about 2–5% and 0.1% for a c- and light-jet, respectively. The c-jets tagging efficiencies with the tight operating point are 25% and misidentification probabilities for a b- and light-jet are $\approx 24\%$ and $< 1.5\%$, respectively. Data from $t\bar{t}$, W + jets, and multijet events that are enriched with a given flavor of interest are used to measure the b-, c-, and light-jets tagging efficiencies. These are compared with the efficiencies calculated from the simulation and small differences between data and simulation are corrected by applying scale factors to the simulation.

Table 5.5 summarizes the selection criteria used in this analysis. Events yields with

Table 5.5: Summary of event selection criteria for Z + ≥ 1 b-jet in the dimuon and the dielectron channels.

Objects	Selection
Muon	$p_{T_\mu} > 35$ (leading) and > 25 (subleading) GeV, $ \eta_\mu < 2.4$, tight ID, $I_{\text{rel}} < 0.15$
Electron	$p_{T_e} > 35$ (leading) and > 25 (subleading) GeV, $ \eta_e < 2.4$, $ \eta_{\text{SC}} < 1.4442$ or $ \eta_{\text{SC}} > 1.556$, cut-based tight WP ID, I_{rel} Barrel $\rightarrow d_{xy}$ (d_z) 0.05 (0.10), Endcap $\rightarrow d_{xy}$ (d_z) 0.10 (0.20)
Z boson	Two oppositely charged leptons and each leading (subleading) lepton $p_T > 35$ (25) GeV, $71 < m_{ll} < 111$ GeV
jets	$p_{T_{\text{jets}}} > 30$ GeV, $ \eta_{\text{jets}} < 2.4$, PU MVA ID > 0.61 ($p_T < 50$ GeV)
b-tagged jets	2016: deep CSV tight WP (> 0.8953) 2017: deep CSV tight WP (> 0.8001) 2018: deep CSV tight WP (> 0.7527)
p_T^{miss}	$p_T^{\text{miss}} < 50$ GeV

a Z boson and an associated b-jet with the above mentioned selection criteria are reported in Table 5.6. The MC simulation overestimates the data up to 10–15% for

the 2016 and 2018 data-taking periods. Data and MC validation are similar in the electron and the muon channels but there is some difference across the years which can be due to differences in tagging conditions. Another reason can be the PDF and tune used in the MC simulation which is the same in 2017 and 2018 but different from 2016. The selected events with a Z boson and at least one b-jet have $\approx 18\%$ background contamination, with the major contributions arising from the $t\bar{t}$, Z + c-jet, and Z + light-jet processes. We can not completely rely on modeling of these events in MC simulated samples, hence, to validate MC simulation, data-driven methods are used to extract the background contribution in the signal region. The control regions that are enriched with particular background processes of interest are constructed and will be discussed later in the background estimation Section 5.7.

Table 5.6: Summary of event yields after Z + ≥ 1 b-jet selection in the dimuon and the dielectron channel for 2016, 2017, 2018, and Run2.

Process	2016		2017		2018		Run2	
	ee	$\mu\mu$	ee	$\mu\mu$	ee	$\mu\mu$	ee	$\mu\mu$
Data	49481	107030	57578	118892	87377	189049	194436	414972
Z + b-jets	47214	105082	48594	99639	77670	163498	173478	368221
Z + light-jets	2114	4220	1888	3368	5774	10280	13700	29254
Z + c-jets	3907	9240	3438	7136	6354	12877	9777	17868
$t\bar{t}$	3387	7528	4554	9421	7018	14953	14959	31903
VV	396	906	501	981	775	1651	1672	3537
Single top	253	530	330	678	518	1112	1101	2320
Total simulation	127430	57309	59306	121263	98112	204371	214727	453142
Data/Total simulation	0.84	0.86	0.97	0.98	0.89	0.92	0.90	0.91

It is interesting to look at the comparison between data and MC for each observable to check the detector simulation's reliability. The data needs to be described as good as possible by simulations. Figures 5.11 and 5.12 show the data and MC simulation comparison of p_T^Z , $p_T^{b\text{-jet}}$, $|\eta^{b\text{-jet}}|$, $\Delta\phi^{(Z, b\text{-jet})}$, $\Delta y^{(Z, b\text{-jet})}$, and $\Delta R^{(Z, b\text{-jet})}$ after Z + ≥ 1 b-jet events selection for the electron and muon channels, respectively. The MC simulation overestimates the data up to 10–15% for all observables. The detailed data-MC validation studies at the detector-level for each year and channel can be found in our analysis note Ref. [46].

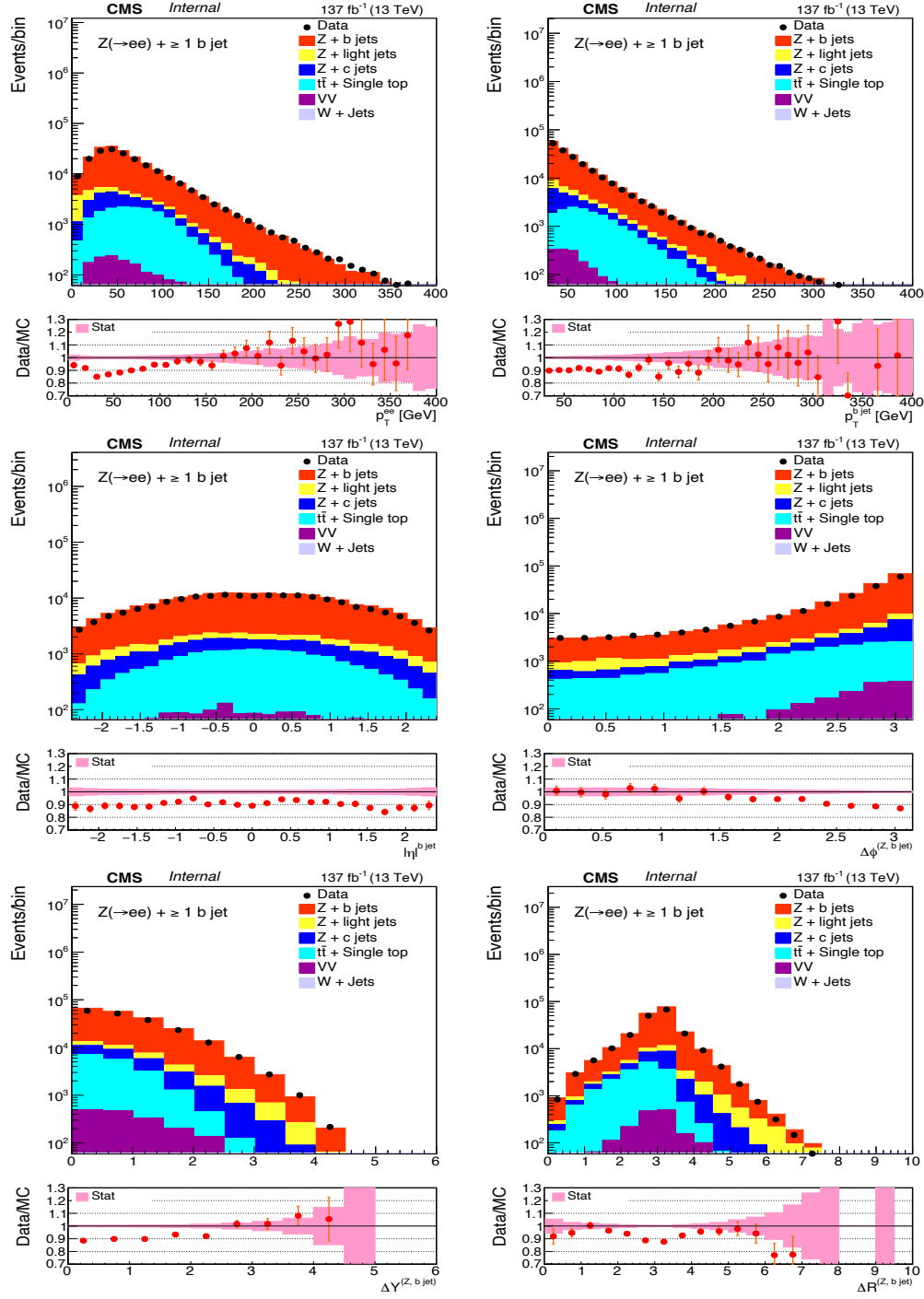


Figure 5.11: Distributions as a function of ee transverse momentum, leading b-jet transverse momentum (top), leading b-jet absolute pseudorapidity, $\Delta\phi^{(Z, b\text{-jet})}$ between the Z and leading b-jet (middle), and $\Delta y^{(Z, b\text{-jet})}$, $\Delta R^{(Z, b\text{-jet})}$ between the Z and leading b-jet (bottom) for ee channel after $Z + \geq 1$ b-jet events selection for Run2 [46]. The bottom panel shows the ratio of data over total MC simulated events. The band on ratio plots includes only the statistical uncertainties.

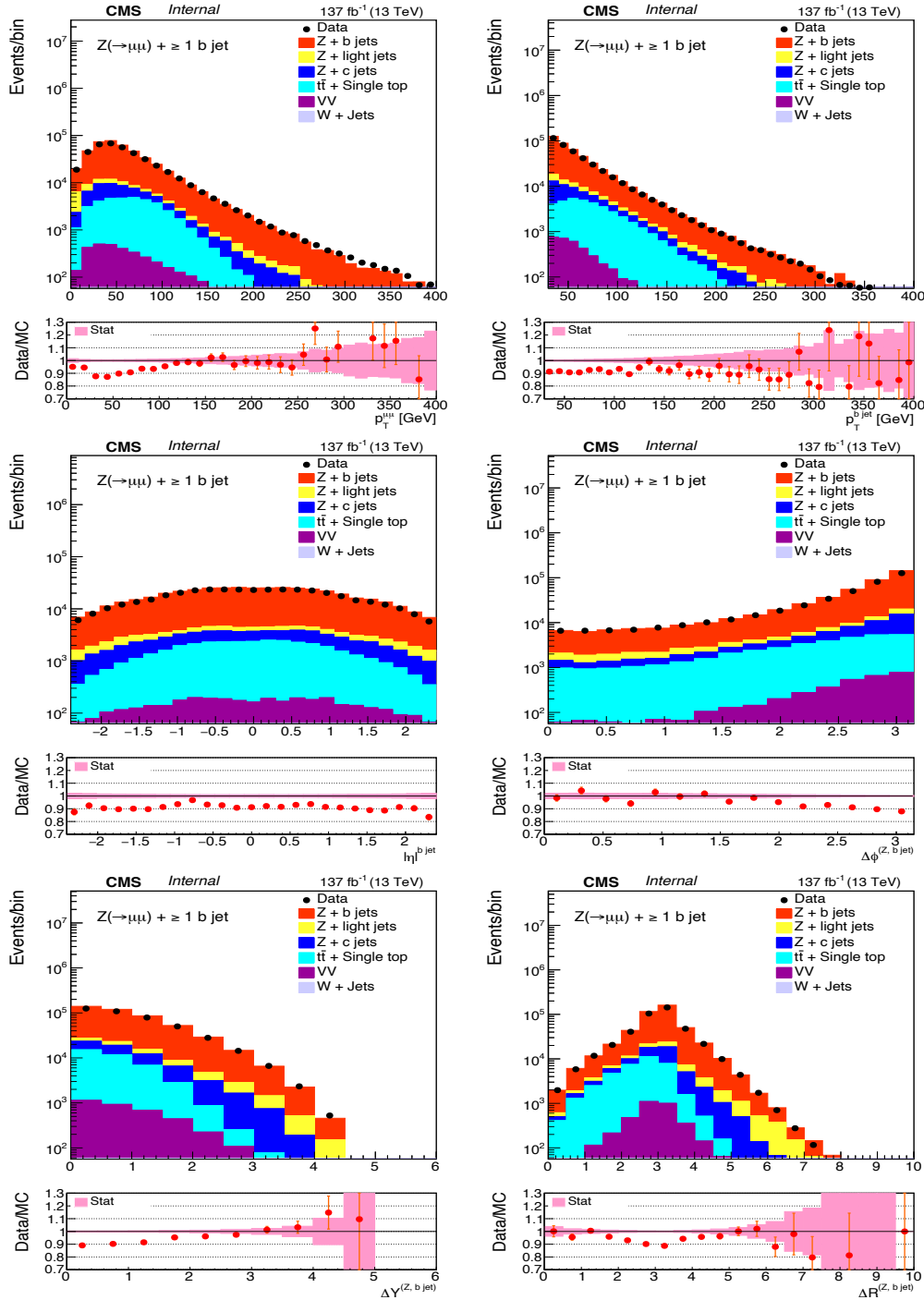


Figure 5.12: Distributions as a function of $\mu\mu$ transverse momentum, leading b-jet transverse momentum (top), leading b-jet absolute pseudorapidity, $\Delta\phi^{(Z, b\text{-jet})}$ between the Z and leading b-jet (middle), and $\Delta y^{(Z, b\text{-jet})}$, $\Delta R^{(Z, b\text{-jet})}$ between the Z and leading b-jet (bottom) for $\mu\mu$ channel after $Z + \geq 1$ b-jet events selection for Run2 [46]. The bottom panel shows the ratio of data over total MC simulated events. The band on ratio plots includes only the statistical uncertainties.

5.7 Background estimation

As discussed earlier, the selected events in the signal region have $\approx 18\%$ background contamination. The dominant background arises from the $t\bar{t}$ ($\approx 7\%$), $Z + c\text{-jet}$ ($\approx 7\%$), and $Z + \text{light-jet}$ ($\approx 4\%$) processes. To validate the simulation, control regions that are enriched with processes of interest are constructed.

- Control region for $t\bar{t}$ ($e\mu + \geq 1 \text{ b-jet}$):** The $t\bar{t}$ control region is constructed by requiring an oppositely charged electron and muon pair with at least one b-jet ($e\mu + \geq 1 \text{ b-jet}$) with the same selection criteria as mentioned in Table 5.5. A sample containing about 92% of $t\bar{t}$ events is obtained for each year of data-taking. The data is described by MC simulation within 10% for 2016 and within 20% for 2017 and 2018 years. Figure 5.13 (left) shows the data and MC simulation comparison as a function leading b-jet p_T for the 2018 year. The non- $t\bar{t}$ contributions to this region are $Z + \text{jets}$, single top, $W + \text{jets}$, and diboson events. As the control region $t\bar{t}$ is very much clean, a scale factor $(\text{data} - (\text{non-}t\bar{t}) \text{ background}) / (t\bar{t} \text{ MC})$, is extracted from this region. It is 0.98 (0.87) for the 2016 (2017–2018) data-taking period and is applied in the signal region to the $t\bar{t}$ simulation in the ee and $\mu\mu$ channel. Figure 5.13 (right) shows the data and MC simulation comparison as a function leading b-jet p_T after applying SF on $t\bar{t}$ MC simulation for $e\mu + \geq 1 \text{ b-jet}$ control region for the 2018 year. There is a significant improvement between data and MC after applying SF.
- Control region for $Z + \text{light-jet}$ ($Z + \geq 1 \text{ jet}$):** The $Z + \text{light-jet}$ control region is studied with two different selection criteria. The first $Z + \text{light-jet}$ control region (CR1) is constructed by requiring events to have inclusive jets without any condition on the tagger discriminator. As a cross-check, another $Z + \text{light-jet}$ control region (CR2) is defined by requiring events to pass anti-b and anti-c tagging tight operating points of b and c quark discriminators to

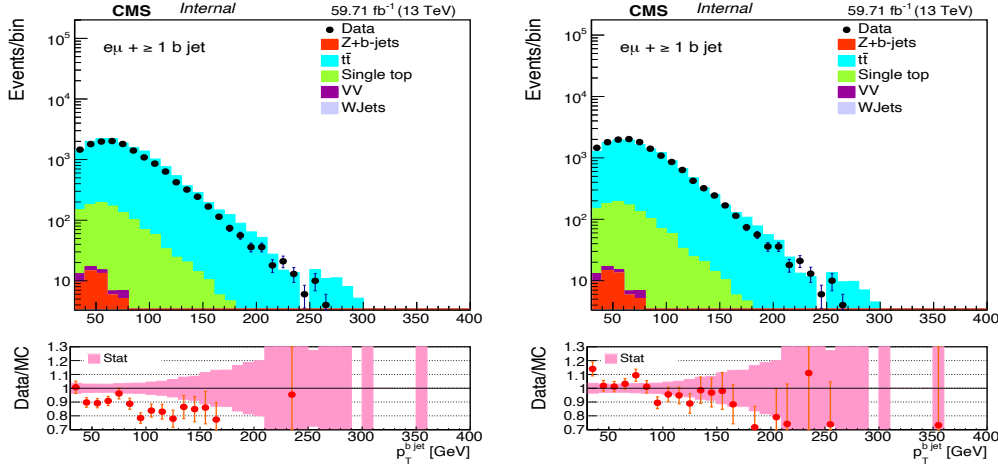


Figure 5.13: Distributions as a function of leading b-jet p_T (left) and after applying scale factor (right) for $e\mu + \geq 1$ b-jet events for 2018 [46]. The bottom panel shows the ratio of data over total MC simulated events. The band on ratio plots includes only the statistical uncertainties.

reject Z + b-jet and Z + c-jet events. A sample containing about 83–85% and 89–90% of Z + light-jet events is obtained in CR1 and CR2, respectively, for each year of data-taking. The data is described by MC simulation within 10% for the 2016 and 2017 data-taking periods whereas, for 2018 deviation goes up to 20% at higher jet p_T . Figure 5.14 shows the data and MC simulation comparison of leading jet p_T in CR1 for the dielectron and the dimuon channels for the 2018 year.

The non-(Z + light-jet) backgrounds are Z + c-jet, Z + b-jet, diboson, $t\bar{t}$, single top, and W + jets events. A scale factor, $(\text{data} - \text{non-(Z + light-jet) background}) / (\text{Z + light-jet MC})$, is extracted and applied in the signal region (Z + ≥ 1 b-jet) to the Z + light-jet MC. The scale factors in CR1 and CR2 are consistent within 1–2%. Therefore, the scale factors 0.94, 0.98, and 0.90 are taken from CR1 and applied in the signal region (Z + ≥ 1 b-jet) for the 2016, 2017, and 2018 data-taking periods, respectively. The comparison of data and MC simulation after applying the scale factor to Z + light-jet MC are shown in Figure 5.15 for the ee and $\mu\mu$ channels for the 2018 year. The measurements are described by MC predictions within uncertainties (statistical and

experimental) discussed in Subsection 5.8.1.

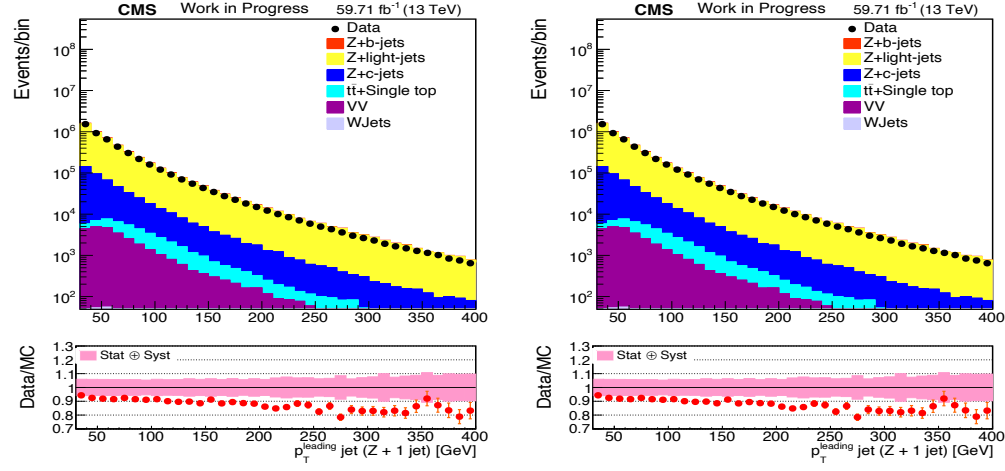


Figure 5.14: Distributions as a function of leading jet p_T for $Z + \geq 1$ jet events for ee (left) and $\mu\mu$ (right) for the 2018 year [46]. The bottom panel shows the ratio of data over total MC simulated events. The band on ratio plots includes only the statistical uncertainties.

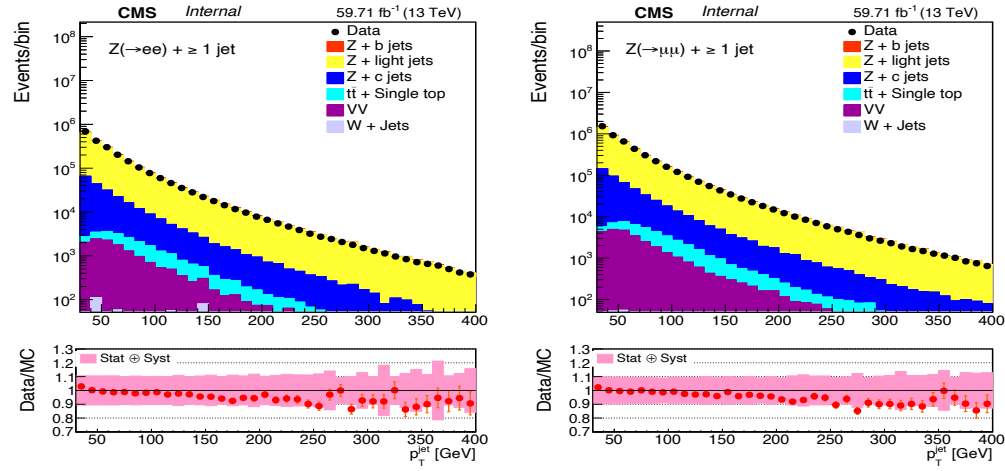


Figure 5.15: Distributions as a function of leading jet p_T for $Z + \geq 1$ jet events after applying light scale factor on $Z + \text{jet}$ MC for ee (left) and $\mu\mu$ (right) for the 2018 year [46]. The bottom panel shows the ratio of data over total MC simulated events. The band on ratio plots includes the statistical and experimental (Jet energy scale, jet energy resolution, p_T^{miss} , lepton selection, pileup reweighting, pileup jet identification, luminosity) uncertainties.

- **Control region for $Z + \text{c-jet}$ ($Z + \geq 1 \text{ c-jet}$):** The $Z + \text{c-jet}$ control region is studied with two different tagger conditions. The one $Z + \text{c-jet}$ control region (CR3) is constructed by requiring c-jet events with the tight operating points of the c tag discriminator. As a cross-check, another $Z + \text{c-jet}$ control region

(CR4) is defined by requiring events to pass tight operating points of the c quark discriminator and inverted b quark discriminator (to reject $Z + b$ -jet). The scale factors extracted from $Z + \text{light-jet}$ CR1 as discussed above are applied to $Z + \text{light-jet}$ MC in $Z + c$ -jet CR3 and CR4. A sample containing about 40–53% and 44–62% of $Z + c$ -jet events is obtained in CR3 and CR4, respectively, for each year of data-taking and is consistent with previous $Z + c$ -jet analysis [47]. The observed agreement between data and MC simulation varies from 10–20% across the years but is consistent across the channels as expected. Figure 5.16 shows validation of leading c -jet p_T for $Z + c$ -jet events in CR3 in the electron and the muon channel.

The non- $Z + c$ -jet contributions to this region are $Z + b$ -jet, $Z + \text{light-jet}$, $t\bar{t}$, single top, $W + \text{jets}$, and diboson events. A scale factor, $(\text{data} - \text{non-}(Z + c\text{-jet}) \text{ background}) / (Z + c\text{-jet MC})$, is extracted and applied in the signal region ($Z + \geq 1 b\text{-jet}$) to the $Z + c$ -jet MC. The scale factors in CR3 and CR4 are consistent within 1–2%. Therefore, the scale factors 0.75, 1.08, and 0.99 are taken from CR3 and applied in the signal region ($Z + \geq 1 b\text{-jet}$) for 2016, 2017, and 2018 data-taking periods, respectively. Figure 5.17 shows validation of leading c -jet p_T for $Z + \geq 1 c\text{-jet}$ control region after applying scale factors extracted from $Z + \text{light-jet}$ CR1 to $Z + \text{light-jet}$ MC and from $Z + c\text{-jet}$ CR3 to $Z + c\text{-jet}$ MC, in the electron and the muon channels. The agreement between data and simulation is improved as compared to when scale factors are not applied and it is within 10%.

5.8 Unfolding and systematic uncertainties

The measured observables, presented in this analysis, differ from the corresponding generator-level values due to the detector effects such as detector resolution and acceptance. So, detector-level distributions have to be corrected for these detector

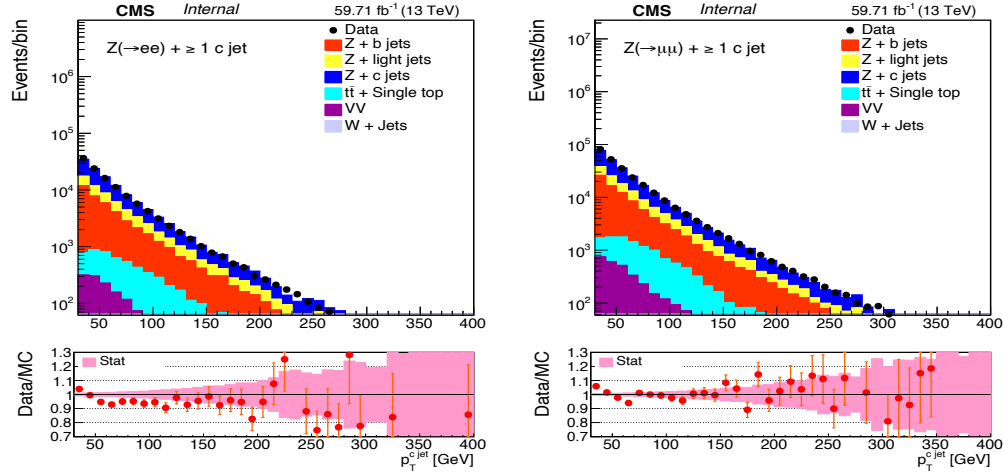


Figure 5.16: Distributions as a function of leading c-jet p_T for $Z + \geq 1$ c-jet events after applying scale factor from CR1 on $Z + \text{light-jet}$ MC for ee (left) and $\mu\mu$ (right) for the 2018 year [46]. The bottom panel shows the ratio of data over total MC simulated events. The band on ratio plots includes only the statistical uncertainties.

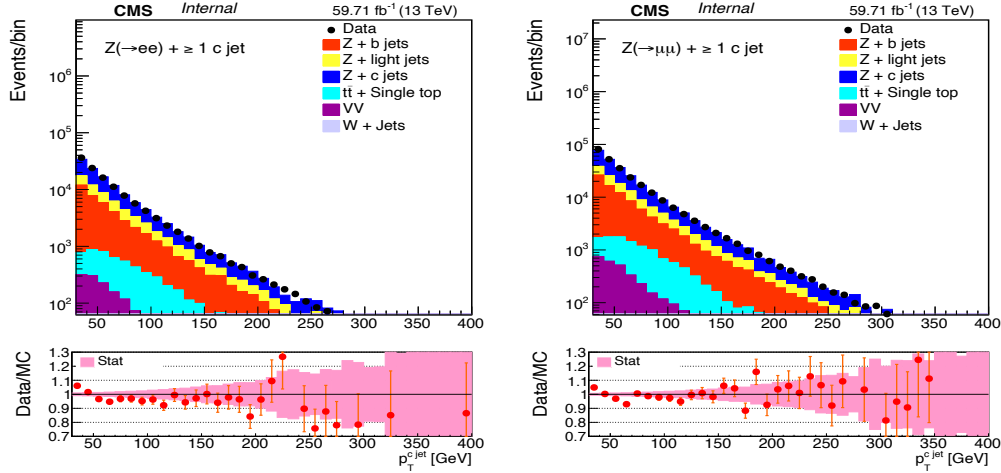


Figure 5.17: Distributions as a function of leading c-jet p_T for $Z + \geq 1$ c-jet events after applying scale factor on $Z + \text{light-jet}$ and $Z + \text{c-jet}$ MC for CR1 and CR3 region, respectively, for ee (left) and $\mu\mu$ (right) for the 2018 year [46]. The bottom panel shows the ratio of data over total MC simulated events. The band on ratio plots includes only the statistical uncertainties.

effects in order to compare the measured observables to the theoretical predictions. The reconstructed distributions are unfolded to generator-level quantities using the TUNFOLD package (v17.5) [48], which is based on the least squares fit. The generated $Z + \text{b-jets}$ events in the simulation are classified according to generator-level information. The generator-level jets are formed using stable particles ($c\tau > 1 \text{ cm}$), except neutrinos, using the same anti- k_T algorithm as for the reconstructed jets.

Any overlap between generator-level jets and a pair of leptons from the Z boson decays is removed by requiring $\Delta R < 0.4$ between them. If an event consists of at least one generator-level jet containing a b hadron, the event is classified as a $Z + \geq 1$ b-jet event. The generator-level leptons are “dressed” by adding the momenta of all photons within $\Delta R \leq 0.1$ around the lepton direction to account for the final-state radiation effects. The distributions are unfolded to the generator-level in the fiducial region defined in Table 5.7.

Table 5.7: Fiducial region definition at generator-level.

Object	Selection
Dressed leptons	p_T (leading) > 35 GeV, p_T (subleading) > 25 GeV, $ \eta < 2.4$
Z boson	$71 < m_{ll} < 111$ GeV
Generator-level b-jet	b hadron jet, $p_T > 30$ GeV, $ \eta < 2.4$

The reconstructed distributions are unfolded with a response matrix (two-dimensional histogram) that describes the migration probability between the generator- and reconstruction-level quantities. The response matrix is constructed by spatially matching of the reconstruction-level leptons and b-jet objects to the corresponding generator-level objects within $\Delta R < 0.3$. The selected $Z + \geq 1$ b-jet sample at the reconstruction-level can have background events that do not match to generator-level objects. This contribution also includes the events where selected b-jets are from the pileup. For this analysis, the total contribution of pileup b-jets is about 2% in the dimuon and dielectron for each year [46]. The background events are channel dependent and they are subtracted from reconstructed events for each year and channel separately. The background subtracted reconstruction-level distributions are unfolded. The unfolded distributions are then corrected for those events that have generator-level objects in the fiducial volume, which do not match with reconstructed objects. Finally, the unfolded distributions are scaled with the integrated luminosity to obtain the differential cross section. The data distributions are unfolded with a response matrix constructed using the $Z +$ b-jet events in the $Z +$ jets sample that is simulated with the MG5_aMC (NLO) generator with CP5 tune.

The choice of a binning is important in order to limit migrations from one bin to others and to ensure that the correction applied to the measured quantities at reconstructed-level, inside a particular bin, is based on events coming mainly from the same bin at the generator-level. So in this analysis, the binning of different variables are chosen on the basis of resolution, purity, and stability. The low purity and stability values indicate strong migrations between bins, meaning that the bin size reaches the resolution limit of the observable. The binwidth is kept more than the resolution of the respective variable and the stability and purity in each bin are more than 50% [46]. The number of bins for reconstructed-level distributions are double than the generator-level distributions bins.

The effect of regularization, a procedure to control large fluctuations of unfolded distributions due to statistical uncertainties of observed quantities [49], is negligible for the selected observables. Therefore, all studies are performed without regularization.

The unfolding procedure is also validated with statistically independent tests in which distributions with MG5_aMC (NLO) are unfolded with MG5_aMC (LO) and compared with generator-level MG5_aMC (NLO) prediction. In Figures 5.18 the results of this closure test are shown only for the combined channel for Run2. There is a good agreement between the unfolded and the generator-level predictions.

The resolution of variables to be measured in the present analysis is not expected to differ for the electron and muon channels. The feasibility of this procedure is validated by unfolding reconstruction-level MC distributions for the electron channel with the response matrix constructed for the muon channel. The background and acceptance distributions are taken from the electron channel only. The resulting unfolded distributions are compared with the unfolded results when the electron channel is used for the construction of the response matrix as shown in Figures 5.19. A good agreement is observed between the two unfolded results, which establishes that the response matrices for the two channels are consistent with each other.

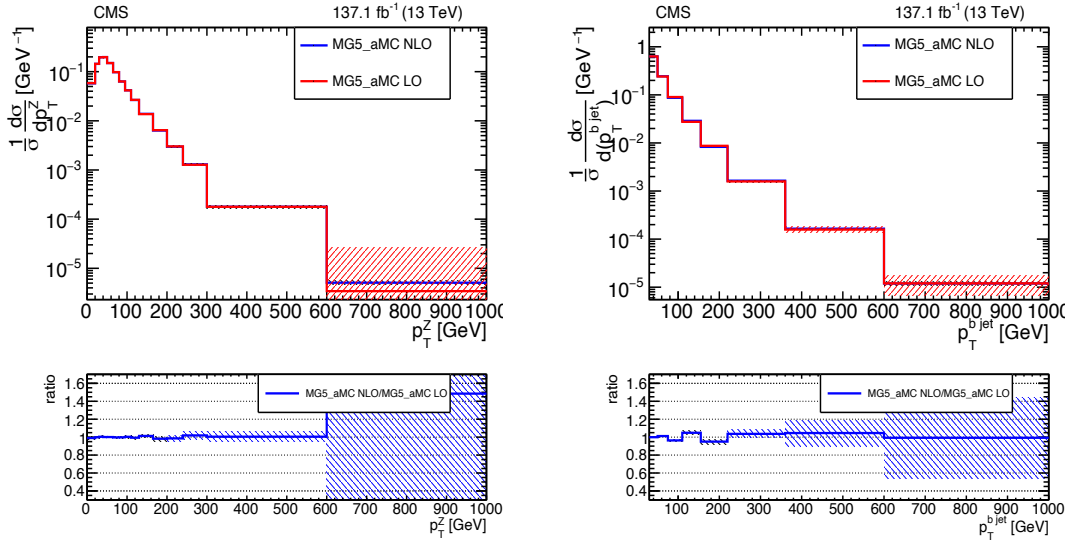


Figure 5.18: Normalized differential cross section distributions as a function of p_T^Z and $p_T^{b\text{-jet}}$ (only two of them are shown here) obtained by unfolding MG5_aMC (NLO) with the response matrix constructed with MG5_aMC (LO) and compared with generator-level prediction of MG5_aMC (NLO) in the combined channel for Run2. The band on ratio plots includes only the statistical uncertainties of the MCs [46].

Therefore, the electron and muon data are added and unfolded with the same or common response matrix.

Closure tests, reported above, are also performed for dimuon, dielectron, and combined channel for each year and can be found in Ref. [46]. All of the closure tests are on the expected lines for the combined unfolding.

5.8.1 Experimental uncertainties

There are a number of experimental sources which can affect the measurements. These are discussed below.

- **Integrated luminosity:** The integrated luminosities of the 2016–2018 data-taking periods have uncertainties in the 1.2–2.5% range. These uncertainties are partially correlated and correspond to a total uncertainty of 1.6% for Run2 [50–52]. The luminosity uncertainty affects the integrated cross section but not the normalized distributions.

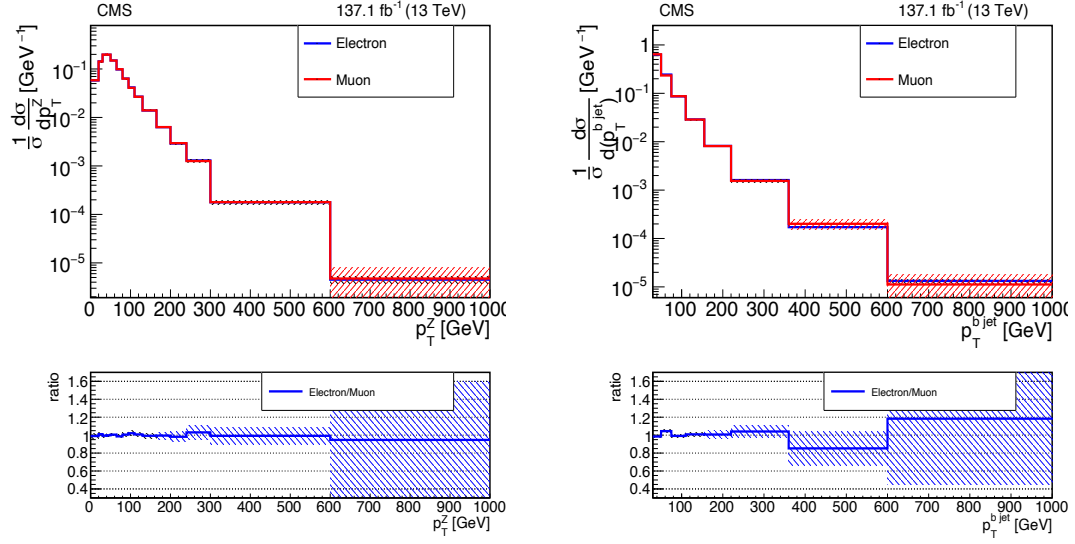


Figure 5.19: Normalized differential cross section distributions as a function of p_T^Z and $p_T^{b\text{-jet}}$ (only two of them are shown here) obtained by unfolding the electron channel with the response matrix constructed for the muon channel for Run2. The band on ratio plots includes only the statistical uncertainties of the MCs [46].

- Jet energy scale and resolution:** The effect of JES and JER is evaluated by varying up and down the p_T values of jets with the corresponding uncertainty factors. The JES and JER uncertainties are also propagated to p_T^{miss} . The JES and JER scale factors are correlated with the JES and JER components of p_T^{miss} , so scale factors are varied up and down simultaneously. The JES introduces a systematic uncertainty of $\approx 3.0\%$ on the integrated cross section whereas the effect of JER is less than 1%. The JES has varying effects on the differential and normalized differential cross section distributions for the different observables, with maximum effects of 9% and 7%, respectively, in the low event count bins.
- b tagging/mistagging:** The uncertainty due to b tagging and mistagging scale factors is evaluated by varying them up and down within their uncertainties. The scale factor for tagging b-jets is correlated with the c-jets scale factor, so for tagging b and mistagging c-jets, scale factors are varied up and down simultaneously, while the scale factor for mistagging light-jets is varied

independently. The combined b tagging uncertainty is then calculated as the quadratic sum of these variations. The uncertainties in the b tagging and mistagging scale factors affect the cross section measurements by 3%. The differential and normalized differential cross section distributions are affected by up to 8% and 5%, respectively.

- Unclustered energy of p_T^{miss} :** The effect of the uncertainty in the remaining component of p_T^{miss} , that is unclustered energy, is estimated by varying its contributions within the uncertainties, which are treated as correlated across years and channels. This uncertainty source affects the integrated cross section by up to 2.8%. The differential and normalized differential cross section distributions of different observables are affected by up to 3.8% and 1.1%, respectively.
- Lepton selection:** Leptons are corrected for the trigger, reconstruction, and identification efficiencies. The effects of these corrections are estimated by varying the corresponding scale factors within their uncertainties. The integrated cross section has a systematic uncertainty around 1.5% and $< 1\%$ for electron and muon selection, respectively. The differential and normalized differential cross section distributions of different observables for electron (muon) channels are affected by up to 2.8% (1%) and 1.3% (1%), respectively.
- Pileup reweighting:** The PU distribution in the simulated samples is reweighted to match that of the data. The corresponding uncertainty is estimated by varying the total inelastic cross section by $\pm 4.6\%$ [53]. The uncertainty in the PU simulation affects the integrated cross section by 1.7%. For the differential and normalized differential cross section distributions the effects are less than 2.7% and 1.3%, respectively.
- Background estimation:** The Z + c-jet and Z + light-jet contributions are estimated using the simulated sample and validated in control regions as dis-

cussed in Section 5.7. The difference between data and MC simulation is extracted from each control region and is applied as a scale factor for the $Z + c\text{-jet}$ and $Z + \text{light-jet}$ MC predictions in the signal region, respectively. These scale factors have associated systematic uncertainties from experimental sources (JES, b tagging, PU, lepton selection), although these are already considered while evaluating systematic uncertainties in the signal region. For all of the control regions, the data and MC simulation values are within 10%. Hence, an additional systematic uncertainty of 10% is assigned to the background modeling. The effect of this uncertainty is about 2.2% on the integrated cross section, whereas the effect of uncertainties related to $Z + c\text{-jet}$ and $Z + \text{light-jet}$ backgrounds is about 1.5%. The corresponding effects on the differential and normalized differential cross section distributions are up to 4.8% and 3.3%, respectively.

- **Model dependence:** The unfolding model uncertainty is estimated by reweighting the signal MC (MG5_aMC at NLO) with a scale factor that is calculated by fitting the ratio of background-subtracted data to signal simulation of normalized distributions and using it as an alternative model for the unfolding. It affects the differential and normalized differential cross section distributions of different observables up to 1.5% and 1.2%, respectively.
- **Pileup jet identification:** The PU jet identification criteria have different performances in the data and simulated events. This difference is corrected by applying a scale factor. The effect of uncertainties associated with the scale factor is less than 0.7% in the integrated cross section. The differential and normalized differential cross section distributions of different observables are affected up to 0.5%.
- **L1 prefiring:** During the 2016–2017 data-taking period there was a timing issue in the $2.0 < |\eta| < 3.0$ region of the ECAL where L1 trigger primitives were mistakenly associated with the previous bunch crossing. This led to a

self-vetoing of events with significant energy in the $2.0 < |\eta| < 3.0$ region and a loss of trigger efficiency. Correction factors are determined for the efficiency loss from data and applied to the simulation of the dielectron and dimuon channels. The uncertainty due to the prefiring scale factors is evaluated by varying them up and down within their uncertainties. For the differential and normalized differential cross section distributions as well as on the integrated cross section the effects are less than 0.5%.

In the cross section ratio measurement $\sigma(Z + \geq 2 \text{ b-jets})/\sigma(Z + \geq 1 \text{ b-jet})$, the luminosity uncertainty is effectively cancelled. All other experimental uncertainties except the background estimation are evaluated as follows. For an uncertainty of interest, the $Z + \geq 2 \text{ b-jets}$ and $Z + \geq 1 \text{ b-jet}$ cross sections are obtained for up and down systematic variations described above and the corresponding ratios are obtained. The uncertainty is derived by comparing these ratios with the nominal ones. The background estimations in the $Z + \geq 2 \text{ b-jets}$ and $Z + \geq 1 \text{ b-jet}$ events are based on independent methods and control regions. Therefore, the background uncertainties are uncorrelated and derived from quadrature sums of those from the $Z + \geq 2 \text{ b-jets}$ and $Z + \geq 1 \text{ b-jet}$ measurements.

5.8.2 Theoretical uncertainties

The uncertainties coming from the theory calculations and MC generators are estimated as follows.

- **Renormalization and factorization scales:** The scale uncertainties are estimated using a set of weights corresponding to variations of renormalization (μ_R) and factorization (μ_F) scales by factors of 0.5 and 2.0. The unfolded distributions are obtained for all combinations, except for $\mu_F/\mu_R = 0.25$ or 4.0, and their envelope is quoted as the uncertainty. The integrated cross sections have scale uncertainties of 2.6%, 2.9%, and 2.1% in the dielectron, dimuon, and combined channels, respectively. The differential and normalized differential

cross section distributions of different observables are affected up to 10% and 9.4%, respectively. Similarly, uncertainties at generator-level are evaluated for MG5_aMC (NLO) with CP5 tune and the effect on the integrated cross section at the generator-level is around 6.6%. The normalized and differential distributions are effected by up to 5–10%.

- **PDF:** The uncertainty in the parton distribution functions is estimated using different Hessian eigenvectors of the NNPDF 3.1 PDF sets. Applying the master formula of the Hessian PDF error calculation [54], this uncertainty is estimated on a bin-by-bin basis for the differential distributions. The integrated cross sections have PDF uncertainties around 0.4%. The effect of this uncertainty on the differential and normalized distributions is up to 4.2% and 3.3%, respectively. The PDF uncertainties are treated as correlated across the years and channels. Similarly, the effect of PDF (NNPDF 3.1) uncertainty in the generator-level distributions is obtained on a bin-by-bin basis of generator-level distributions by applying PDF weights. The effect on the integrated cross section at the generator-level is around 1.6% and up to 1% for differential and normalized distributions, respectively. The uncertainties in the PDFs are also estimated for SHERPA at generator-level using the 102 replicas of the NNPDF 3.0 PDF set. The integrated cross sections have PDF uncertainties around 1.4%.
- **Strong coupling (α_s):** The uncertainties due to α_s are estimated by using two additional PDF members corresponding to upper (0.120) and lower (0.116) values with respect to the central (0.118) α_s value. The integrated cross sections have α_s uncertainties around 0.3%. The effect on the differential and normalized distributions are up to 1.3% and 1.1%, respectively. The effect on the integrated cross section at the generator-level is 1.2% and up to 4.5% and 3.5% for differential and normalized distributions, respectively.

All theoretical uncertainties in the cross section ratios are the quadrature sum of those from the $Z + \geq 2$ b-jets and $Z + \geq 1$ b-jet measurements.

Table 5.8 lists the uncertainties in the $Z + \geq 1$ b-jet integrated cross sections. All systematic uncertainties are treated as 100% correlated across the years, however, JER is uncorrelated across the year but its effect is small, so this is also considered as correlated. In summary, the $Z + \geq 1$ b-jet integrated cross sections have total experimental systematic uncertainties of 5.9% and 6.1%, respectively, in the dimuon and combined channels. For the electron channel, the uncertainty is a bit higher, up to 7.6%, owing to the electron selection uncertainty. For the measured integrated cross section the statistical uncertainty is less than 1%, thus measurement is systematically limited. The total experimental systematic uncertainty for the integrated cross section of $\sigma(Z + \geq 2 \text{ b-jets})/\sigma(Z + \geq 1 \text{ b-jet})$ is 7%.

Table 5.9 summarize the range of uncertainty in the differential and normalized differential cross section distributions for the $Z + \geq 1$ b-jet combined channel.

Table 5.8: Summary of uncertainties (%) in the integrated cross sections for the di-electron, dimuon, and combined channels in the $Z + \geq 1$ b-jet.

Uncertainty (%)	ee	$\mu\mu$	ll
Statistical	1.0	0.7	0.6
JES, JER	2.7	3.0	2.9
b tagging/mistagging	3.0	2.9	2.9
Unclustered energy of p_T^{miss}	2.8	2.8	2.8
Background estimation	2.2	2.0	2.1
Pileup reweighting	2.4	1.7	1.9
Electron selection	4.6	—	1.5
Luminosity	1.6	1.6	1.6
Muon selection	—	0.6	0.4
Pileup jet identification	0.3	0.3	0.3
L1 prefiring	0.3	0.2	0.2
μ_R and μ_F scales	2.6	2.9	2.1
PDF	0.4	0.3	0.3
α_s	0.3	0.2	0.2
Total experimental	7.6	5.9	6.1
Total theoretical	2.6	2.9	2.1

Table 5.9: Summary of the uncertainties (%) in the differential (normalized differential) cross section distributions for the combined channel in the $Z + \geq 1$ b-jet events.

Observable/Uncertainty (%)	p_T^Z	$p_T^{\text{b-jet}}$	$ \eta^{\text{b-jet}} $	$\Delta\phi^{(Z, \text{b-jet})}$	$\Delta y^{(Z, \text{b-jet})}$	$\Delta R^{(Z, \text{b-jet})}$
Statistical	1.0–5.2 (1.0–5.2)	1.0–12 (1.0–12)	0.7–1.4 (0.7–1.4)	0.7–2.2 (1.0–2.1)	0.7–4.7 (0.7–4.7)	1.0–5.3 (0.8–3.3)
JES, JER	1.0–9.0 (1.0–5.8)	2.0–7.2 (1.0–6.7)	2.0–7.0 (0.5–4.5)	2.0–4.8 (0.5–2.1)	2.0–6.4 (0.5–4.2)	1.6–6.1 (0.5–4.2)
b tagging/mistagging	2.3–6.0 (0.5–3.9)	2.0–8.3 (0.5–5.6)	2.6–4.3 (0.5–1.7)	2.5–4.9 (0.5–2.9)	2.7–5.5 (0.5–3.7)	2.6–7.2 (0.5–1.7)
Unclustered energy of p_T^{miss}	2.2–3.6 (0.5–0.7)	2.2–3.4 (0.5–0.7)	0.5–1.8 (0.5–1.1)	2.5–3.8 (0.5–1.0)	2.7–3.0 (0.5)	2.3–3.8 (0.5–1.0)
Background estimation	1.6–3.6 (0.5–1.9)	1.8–3.0 (0.6–1.7)	1.9–3.2 (0.5–1.4)	1.5–3.9 (0.5–2.0)	2.0–3.4 (0.5–2.4)	2.6–4.8 (0.5–3.3)
Pileup reweighting	1.0–2.2 (0.5–0.9)	0.7–2.2 (0.5–1.3)	1.8–2.4 (0.5)	1.8–2.6 (0.5–0.7)	1.8–2.2 (0.5)	1.5–2.5 (0.5–0.8)
Model dependency	0.5–1.4 (0.5–1.4)	0.5–1.2 (0.5–0.8)	0.5 (0.5)	0.5–1.5 (0.5–1.2)	0.5–0.6 (0.5)	0.5–1.1 (0.5–0.9)
Electron selection	1.3–2.8 (0.5–1.3)	1.2–1.9 (0.5)	1.4–1.8 (0.5)	1.3–1.9 (0.5)	1.5–2.4 (0.5–0.9)	1.2–2.2 (0.5–0.6)
Muon selection	0.5–1.0 (0.5–1.0)	0.5–1.0 (0.5)	0.5 (0.5)	0.5 (0.5)	0.5 (0.5)	0.5–0.6 (0.5)
Pileup jet identification	0.5–0.6 (0.5)	0.5 (0.5)	0.5–0.9 (0.5)	0.5 (0.5)	0.5–0.6 (0.5)	0.5 (0.5)
L1 prefiring	0.5 (0.5)	0.5 (0.5)	0.5–0.7 (0.5)	0.5 (0.5)	0.5–0.9 (0.5)	0.5 (0.5)
μ_R and μ_F scales	1.0–7.2 (1.0–7.0)	1.0–10.0 (0.5–9.4)	0.7–3.3 (0.5–2.0)	1.5–2.7 (1.0–2.1)	1.0–2.8 (0.5–2.1)	1.5–4.0 (0.5–3.7)
PDF	0.5–2.9 (0.5–2.0)	0.5–2.1 (0.5–1.6)	1.5–2.8 (0.5–0.7)	0.5–2.1 (0.5–1.8)	0.5–1.6 (0.5–1.6)	0.5–4.2 (0.5–3.3)
α_s	0.5–1.0 (0.5–1.0)	0.5–0.7 (0.5–0.7)	0.5 (0.5)	0.5–1.1 (0.5–0.9)	0.5 (0.5)	0.5–1.3 (0.5–1.1)

5.9 Results and discussion

The production cross sections, within fiducial region, are measured to be 6.52 ± 0.04 (stat) ± 0.40 (syst) ± 0.14 (theo) pb and 0.65 ± 0.03 (stat) ± 0.07 (syst) ± 0.02 (theo) pb for the $Z + \geq 1$ b-jet and $Z + \geq 2$ b-jets processes, respectively. The measurements are consistent for the dielectron, dimuon, and combined channels. The measured cross sections are compared with predictions from MG5_aMC obtained with two settings corresponding to the 2016 and 2017–2018 data-taking periods. The 2017–2018 settings use more up-to-date MG5_aMC versions (2.6.0 and 2.4.2 for NLO and LO, respectively), PDFs (NNPDF 3.1), and underlying event tune (CP5), and are labeled with NNPDF 3.1, CP5 in the following. The 2016 data-

taking period uses earlier versions of the MG5_aMC (2.3.2.2 and 2.2.2 for NLO and LO, respectively), PDFs (NNPDF 3.0), and underlying event tune (CUETP8M1), and are labeled with NNPDF 3.0, CUETP8M1. In the $Z + \geq 1$ b-jet final state, the measured integrated cross section values are well described by MG5_aMC LO for both settings, whereas MG5_aMC (NLO, NNPDF 3.1, CP5) and MG5_aMC (NLO, NNPDF 3.0, CUETP8M1) overestimate data by $\approx 10\%$ and $\approx 18\%$, respectively. The SHERPA simulation overestimates the measured cross section values by $\approx 24\%$. In the $Z + \geq 2$ b-jets final state, the measured integrated cross section is in good agreement with MG5_aMC (LO). The MG5_aMC (NLO, NNPDF 3.1, CP5), MG5_aMC (NLO, NNPDF 3.0, CUETP8M1), and SHERPA predictions overestimate the measured cross section values by 29%, 38%, and 21%, respectively. The measured value of the cross section ratio of the $Z + \geq 2$ b-jets and $Z + \geq 1$ b-jet for the combined channel is 0.100 ± 0.005 (stat) ± 0.007 (syst) ± 0.003 (theo) and, as expected, has smaller uncertainties. This measurement is in agreement with the MG5_aMC (LO, NNPDF 3.0, CUETP8M1) and SHERPA values given the large uncertainties associated with the predictions.

The measured differential cross section distributions and the corresponding ones normalized by the integrated fiducial cross sections compared with different MC predictions in the combined channel for the selected $Z + \geq 1$ b-jet events are shown in Figures [5.20](#)–[5.25](#). In general, the differential cross section distributions predicted by SHERPA tend to overestimate data by 20–30%, depending on the kinematic variables.

The differential and normalized differential cross section distributions as a function of the p_T^Z in the selected $Z + \geq 1$ b-jet event are shown in Figure [5.20](#). In the low p_T region (up to 100 GeV), where soft radiative effects play a role, all the predicted shapes exhibit different trends from those in the data. The shapes of p_T^Z distributions are described best by MG5_aMC (LO, NNPDF 3.1, CP5), while MG5_aMC (NLO) and SHERPA predictions vary up to 30% depending on p_T^Z . The differential cross section distribution of p_T^Z agrees within 10% with MG5_aMC (LO).

Table 5.10: Measured and predicted cross sections (in pb) for the $Z + \geq 1$ b-jet and $Z + \geq 2$ b-jets final states. The cross section ratios between the $Z + \geq 2$ b-jets and $Z + \geq 1$ b-jet are shown in the last three rows for the dielectron, dimuon, and combined channels. The MG5_aMC (NLO) predictions include theoretical uncertainties (PDF, and renormalization and factorization scales).

	Channel	Measured	MG5_aMC		MG5_aMC		MG5_aMC		SHERPA
			LO		NLO		NLO		
			NNPDF 3.0	NNPDF 3.1	NNPDF 3.0	NNPDF 3.1	NNPDF 3.0	NNPDF 3.1	
			CUETP8M1	CP5	CUETP8M1	CP5	CUETP8M1	CP5	
Z + ≥ 1 b-jet	ee	6.45 ± 0.06 (stat) ± 0.49 (syst) ± 0.17 (theo)	6.25	6.33	7.86 ± 0.52 (theo)	7.05 ± 0.48 (theo)	7.05 ± 0.48 (theo)	8.05	
	$\mu\mu$	6.55 ± 0.05 (stat) ± 0.39 (syst) ± 0.19 (theo)	6.26	6.34	7.86 ± 0.51 (theo)	7.02 ± 0.47 (theo)	7.02 ± 0.47 (theo)	7.98	
	ll	6.52 ± 0.04 (stat) ± 0.40 (syst) ± 0.14 (theo)	6.25	6.34	7.86 ± 0.51 (theo)	7.03 ± 0.47 (theo)	7.03 ± 0.47 (theo)	8.02	
Z + ≥ 2 b-jets	ee	0.66 ± 0.05 (stat) ± 0.07 (syst) ± 0.02 (theo)	0.62	0.72	0.89 ± 0.08 (theo)	0.77 ± 0.07 (theo)	0.77 ± 0.07 (theo)	0.84	
	$\mu\mu$	0.65 ± 0.04 (stat) ± 0.06 (syst) ± 0.02 (theo)	0.64	0.71	0.91 ± 0.09 (theo)	0.77 ± 0.07 (theo)	0.77 ± 0.07 (theo)	0.84	
	ll	0.65 ± 0.03 (stat) ± 0.07 (syst) ± 0.02 (theo)	0.63	0.71	0.90 ± 0.09 (theo)	0.77 ± 0.07 (theo)	0.77 ± 0.07 (theo)	0.84	
Ratio	ee	0.102 ± 0.008 (stat) ± 0.008 (syst) ± 0.004 (theo)	0.100	0.113	0.113 ± 0.016 (theo)	0.110 ± 0.013 (theo)	0.110 ± 0.013 (theo)	0.104	
	$\mu\mu$	0.100 ± 0.006 (stat) ± 0.006 (syst) ± 0.004 (theo)	0.103	0.112	0.116 ± 0.016 (theo)	0.110 ± 0.013 (theo)	0.110 ± 0.013 (theo)	0.105	
	ll	0.100 ± 0.005 (stat) ± 0.007 (syst) ± 0.003 (theo)	0.102	0.112	0.114 ± 0.016 (theo)	0.110 ± 0.013 (theo)	0.110 ± 0.013 (theo)	0.105	

The MG5_aMC (NLO) predictions tend to deviate up to 20% for very low and very high p_T^Z values, while SHERPA tends to deviate up to 35% for low values of p_T^Z and agrees within 10% for high p_T^Z values.

The $p_T^{\text{b-jet}}$ and $|\eta^{\text{b-jet}}|$ differential and normalized differential cross section distributions are shown in Figures 5.21 and 5.22, respectively. The shapes of distributions are well described by all simulations, except for the MG5_aMC (LO, NNPDF 3.1, CP5) $p_T^{\text{b-jet}}$ spectrum, which deviates up to 25% in the higher p_T region. The differential cross section distribution of $p_T^{\text{b-jet}}$ agrees within 10% with MG5_aMC (LO) and MG5_aMC (NLO), except for MG5_aMC (LO, NNPDF 3.1, CP5), which deviates up to 25% in the higher p_T region, while SHERPA tends to deviate by 10–30%. The $|\eta^{\text{b-jet}}|$ differential cross section distribution is well described by MG5_aMC (LO), while MG5_aMC (NLO, NNPDF 3.0, CUETP8M1) and SHERPA predictions describe it within 25% and 30%, respectively.

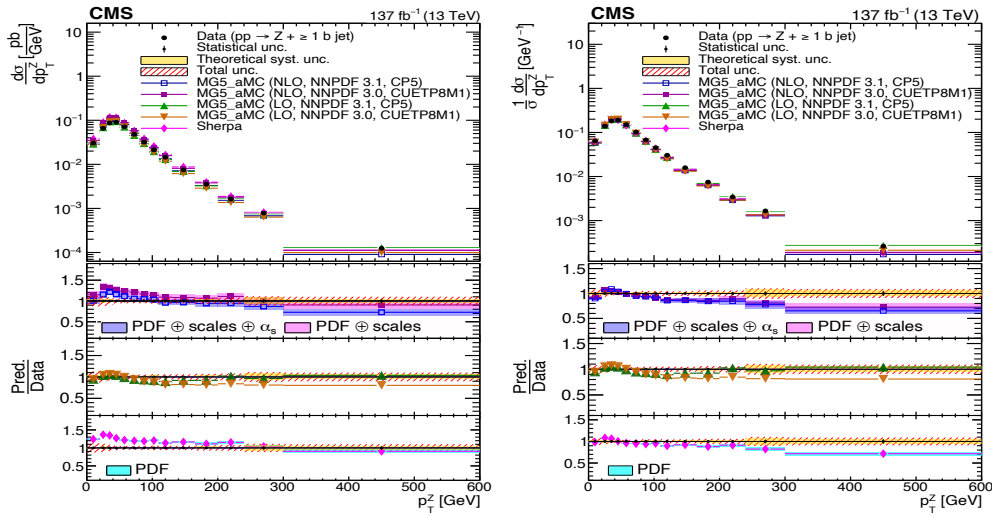


Figure 5.20: Differential cross section (left) and the normalized differential cross section (right) distributions as a function of Z transverse momenta for $Z + \geq 1$ b-jet events. The uncertainties in the predictions are shown as colored bands in the bottom panel around the theoretical predictions including statistical, PDF, scale, and α_s uncertainties for the MG5_aMC (NLO, NNPDF 3.1, CP5), MG5_aMC (NLO, NNPDF 3.0, CUETP8M1) and the statistical and PDF uncertainties for the SHERPA predictions. The statistical, theoretical, and total uncertainties in data are indicated by the vertical bars, yellow, and hatched bands centered at 1, respectively.

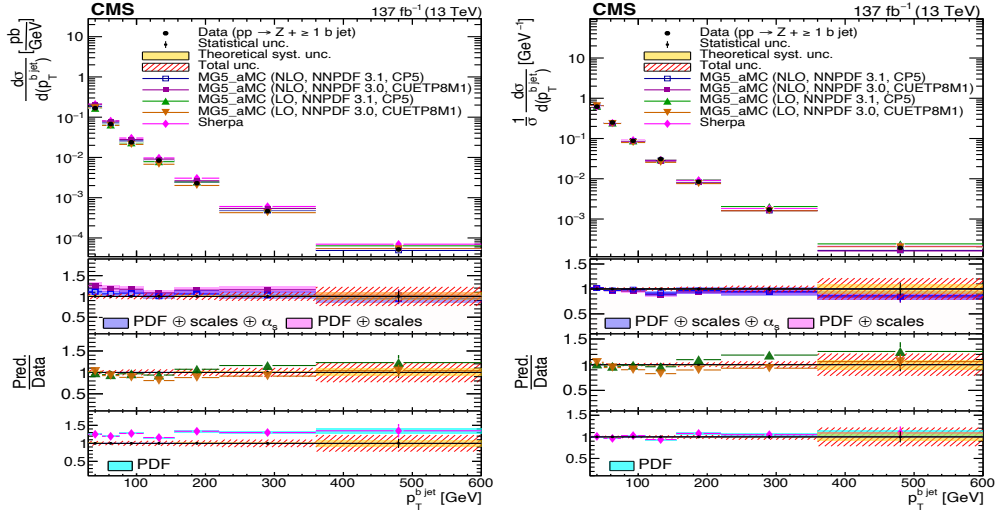


Figure 5.21: Differential cross section (left) and the normalized differential cross section (right) distributions as a function of b-jet transverse momenta for $Z + \geq 1$ b-jet events. The uncertainties in the predictions are shown as colored bands in the bottom panel around the theoretical predictions including statistical, PDF, scale, and α_s uncertainties for the MG5_aMC (NLO, NNPDF 3.1, CP5), MG5_aMC (NLO, NNPDF 3.0, CUETP8M1) and the statistical and PDF uncertainties for the SHERPA predictions. The statistical, theoretical, and total uncertainties in data are indicated by the vertical bars, yellow, and hatched bands centered at 1, respectively.

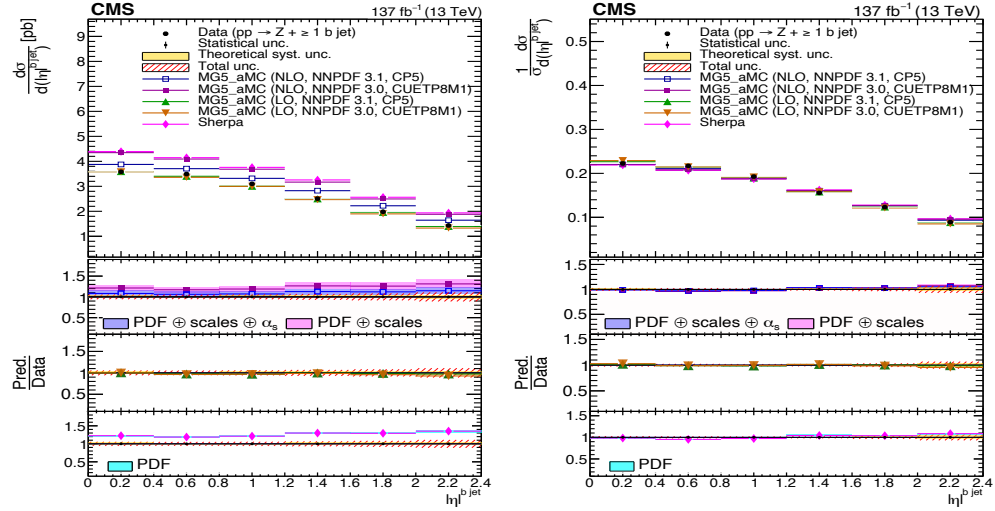


Figure 5.22: Differential cross section (left) and the normalized differential cross section (right) distributions as functions of b-jet absolute pseudorapidity for $Z + \geq 1$ b-jet events. The uncertainties in the predictions are shown as colored bands in the bottom panel around the theoretical predictions including statistical, PDF, scale, and α_s uncertainties for the MG5_aMC (NLO, NNPDF 3.1, CP5), MG5_aMC (NLO, NNPDF 3.0, CUETP8M1) and the statistical and PDF uncertainties for the SHERPA predictions. The statistical, theoretical, and total uncertainties in data are indicated by the vertical bars, yellow, and hatched bands centered at 1, respectively.

The differential and normalized differential cross section distributions of $\Delta\phi^{(Z, \text{b-jet})}$, $\Delta y^{(Z, \text{b-jet})}$, and $\Delta R^{(Z, \text{b-jet})}$ are shown in Figures 5.23–5.25. The shapes of these distributions are best described by the SHERPA predictions. The $\Delta\phi^{(Z, \text{b-jet})}$ differential cross section distribution is well described by MG5_aMC (LO), while MG5_aMC (NLO, NNPDF 3.0, CUETP8M1) and SHERPA predictions describe within 25%. The MG5_aMC (NLO) and MG5_aMC (LO) predictions both provide a similar agreement for the normalized distribution, although the agreement is somewhat better in the higher $\Delta\phi^{(Z, \text{b-jet})}$ region as compared to the lower $\Delta\phi^{(Z, \text{b-jet})}$. The lower region of $\Delta\phi^{(Z, \text{b-jet})}$, which is most sensitive to additional QCD radiation as well as soft corrections, is poorly modeled; these effects are not fully captured in the non-perturbative corrections applied to that prediction. It also appears from the normalized distribution that higher order calculations of MG5_aMC do not improve the description of the distribution but different MPI models in MG5_aMC and SHERPA have different agreements. So, this variable can be used for tuning the MPI models.

The shape of $\Delta y^{(Z, \text{b-jet})}$ and $\Delta R^{(Z, \text{b-jet})}$ distributions are nicely described by the SHERPA predictions, whereas MG5_aMC (LO) and MG5_aMC (NLO) deviate in the higher $\Delta y^{(Z, \text{b-jet})}$ and $\Delta R^{(Z, \text{b-jet})}$ regions. The LO calculation by MG5_aMC underestimates the distributions in the higher region by 20–30%, which is improved significantly with the NLO prediction by MG5_aMC for normalized as well as differential cross section distributions. The MG5_aMC (LO) and MG5_aMC (NLO) are simulated with the same MPI model of PYTHIA 8 but have different level of agreement with the $\Delta y^{(Z, \text{b-jet})}$ and $\Delta R^{(Z, \text{b-jet})}$ distributions which can be attributed to the effect of higher order corrections on the simulation of $\Delta y^{(Z, \text{b-jet})}$ and $\Delta R^{(Z, \text{b-jet})}$ observables. The SHERPA predictions tend to deviate up to 20–30% for both observables. In summary, SHERPA prediction provides a good description of different kinematic observables except for the higher p_T^Z region for the normalized differential

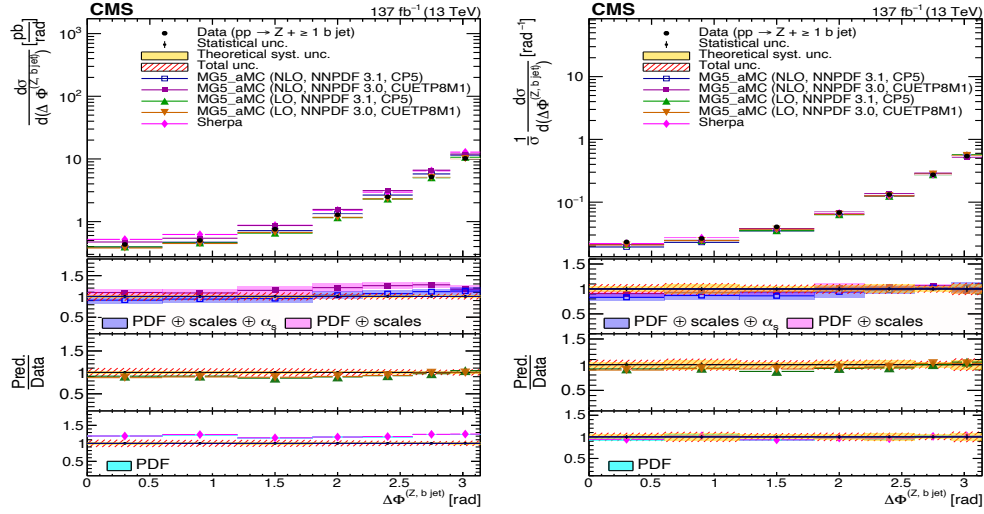


Figure 5.23: Differential cross section (left) and the normalized differential cross section (right) distributions as functions of $\Delta\phi^{(Z, b\text{-jet})}$ between the Z boson and the leading b-jet for $Z + \geq 1$ b-jet events. The uncertainties in the predictions are shown as colored bands in the bottom panel around the theoretical predictions including statistical, PDF, scale, and α_s uncertainties for the MG5_aMC (NLO, NNPDF 3.1, CP5), MG5_aMC (NLO, NNPDF 3.0, CUETP8M1) and the statistical and PDF uncertainties for the SHERPA predictions. The statistical, theoretical, and total uncertainties in data are indicated by the vertical bars, yellow, and hatched bands centered at 1, respectively.

distributions. The MG5_aMC LO and NLO simulations provide varying levels of agreement for the observables.

The $\sigma(Z + \geq 2 \text{ b-jets})/\sigma(Z + \geq 1 \text{ b-jet})$ cross section ratio distributions as functions of the leading b-jet transverse momentum and absolute pseudorapidity are shown in Figure 5.26. The ratio gradually increases (from 0.05 to 0.25) with the leading b-jet p_T (ranging from 30 to 200 GeV), but is nearly independent of the pseudorapidity of the leading b-jet. The increase in the ratio is due to a kinematic effect; at larger p_T of the leading jet, the kinematic acceptance for the subleading jet would increase. Predictions from MG5_aMC (LO), MG5_aMC (NLO), and SHERPA describe the measured ratios within uncertainties.

To summarize, a measurement of fiducial cross sections of the $Z + \geq 1$ b-jet process, along with the differential and normalized differential cross section distributions of different kinematic observables is presented. This measurement is performed using pp collisions data at $\sqrt{s} = 13 \text{ TeV}$ collected by the CMS experiment at the

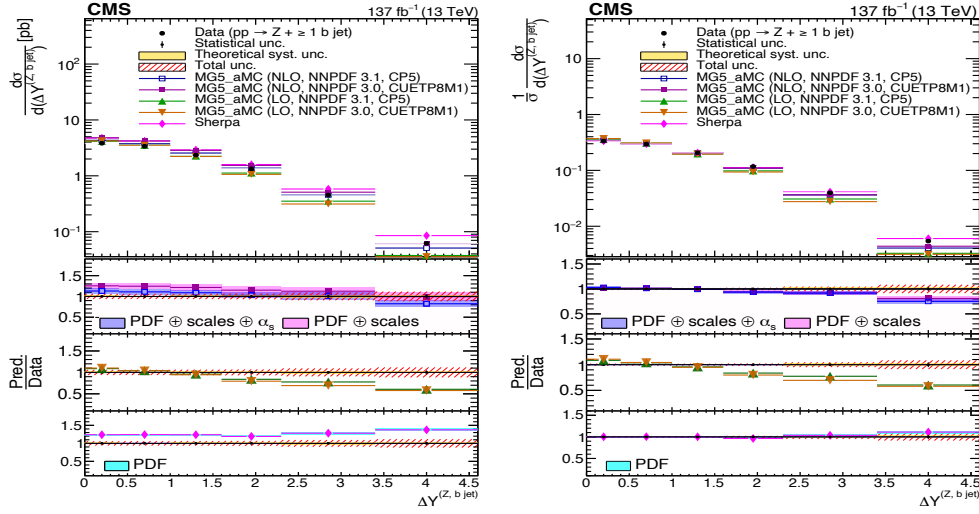


Figure 5.24: Differential cross section (left) and the (right) normalized differential cross section (right) distributions as functions of $\Delta Y^{(Z, \text{b-jet})}$ between the Z boson and the leading b-jet for the $Z + \geq 1 \text{ b-jet}$ events. The uncertainties in the predictions are shown as colored bands in the bottom panel around the theoretical predictions including statistical, PDF, scale, and α_s uncertainties for the MG5_aMC (NLO, NNPDF 3.1, CP5), MG5_aMC (NLO, NNPDF 3.0, CUETP8M1) and the statistical and PDF uncertainties for the SHERPA predictions. The statistical, theoretical, and total uncertainties in data are indicated by the vertical bars, yellow, and hatched bands centered at 1, respectively.

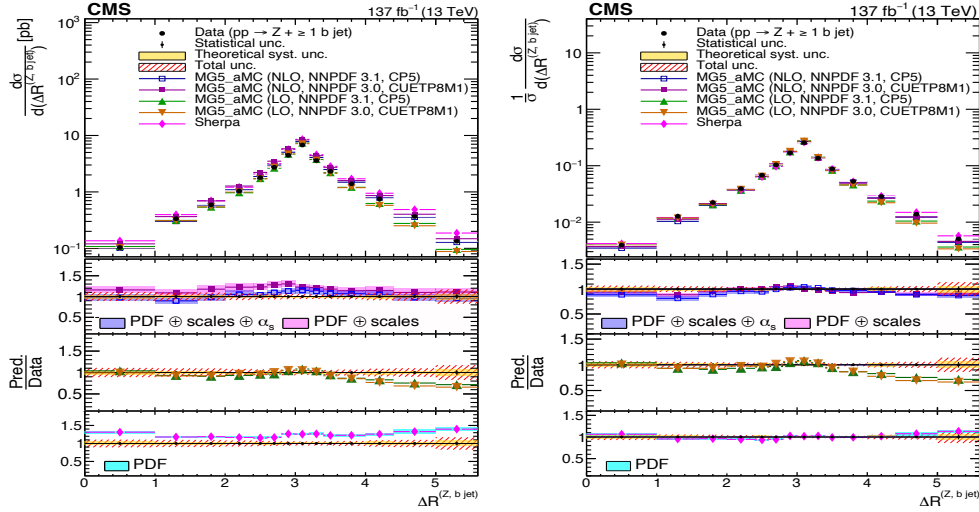


Figure 5.25: Differential cross section (left) and the normalized differential cross section (right) distributions as functions of $\Delta R^{(Z, \text{b-jet})}$ between the Z boson and the leading b-jet for the $Z + \geq 1 \text{ b-jet}$ events. The uncertainties in the predictions are shown as colored bands in the bottom panel around the theoretical predictions including statistical, PDF, scale, and α_s uncertainties for the MG5_aMC (NLO, NNPDF 3.1, CP5), MG5_aMC (NLO, NNPDF 3.0, CUETP8M1) and the statistical and PDF uncertainties for the SHERPA predictions. The statistical, theoretical, and total uncertainties in data are indicated by the vertical bars, yellow, and hatched bands centered at 1, respectively.

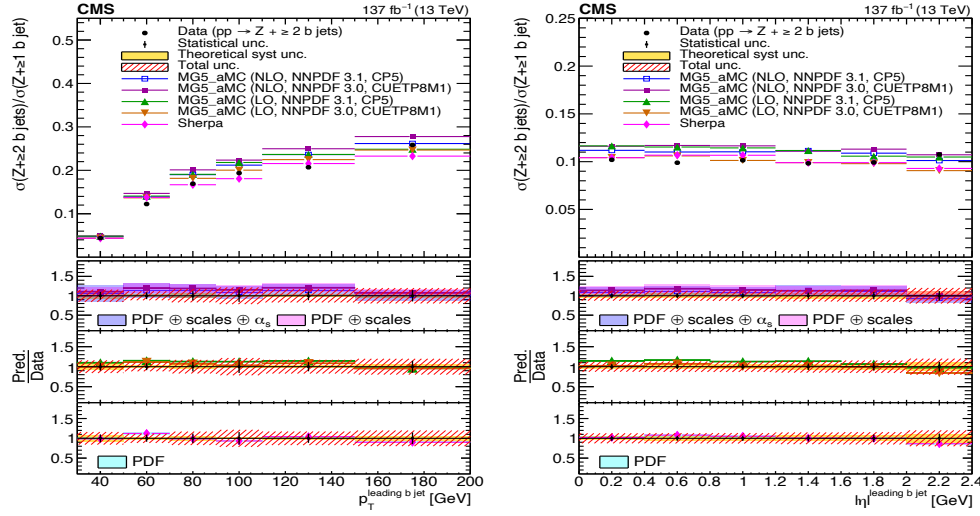


Figure 5.26: Distributions of the cross section ratios, $\sigma(Z + \geq 2 \text{ b-jets})/\sigma(Z + \geq 1 \text{ b-jet})$, as functions of the (left) leading b-jet transverse momentum and (right) absolute pseudorapidity. The uncertainties in the predictions are shown as colored bands in the bottom panel around the theoretical predictions including statistical, PDF, scale, and α_s uncertainties for the MG5_aMC (NLO, NNPDF 3.1, CP5), MG5_aMC (NLO, NNPDF 3.0, CUETP8M1) and the statistical and PDF uncertainties for the SHERPA predictions. The statistical, theoretical, and total uncertainties in data are indicated by the vertical bars, yellow, and hatched bands centered at 1, respectively.

CERN LHC. This is the first measurement of these processes based on data collected during the 2016–2018 LHC running period. The measured fiducial cross sections are better described by the MG5_aMC (LO) simulation but overestimated by MG5_aMC (NLO) and SHERPA predictions. The measured value of the cross section ratio of the $Z + \geq 2 \text{ b-jets}$ to $Z + \geq 1 \text{ b-jet}$ is well described by the MG5_aMC (LO, NNPDF 3.0, CUETP8M1) and SHERPA calculations but overestimated by MG5_aMC (NLO) prediction. Present measurements can be used as an input for the further optimization of the simulation parameters and b-quark PDF. This work has been published [55].

Bibliography

- [1] Brodsky, S. J. and Kusina, A. and Lyonnet, F. and Schienbein, I. and Spiesberger, H. and Vogt, R., “A review of the intrinsic heavy quark content of the nucleon”, *Adv. High Energy Phys.*, vol. 2015 (2015) 231547, [doi:10.1155/2015/231547](https://doi.org/10.1155/2015/231547) [arXiv:1504.06287](https://arxiv.org/abs/1504.06287).
- [2] Bozzi, G. and Rojo, J. and Vicini, A., “The impact of the parton distribution function uncertainties on the measurement of the W boson mass at the Tevatron and the LHC,” *Phys. Rev. D* , vol. 83 (2011) 113008, [doi:10.1103/PhysRevD.83.113008](https://doi.org/10.1103/PhysRevD.83.113008).
- [3] CDF collaboration, “Measurement of cross sections for b jet production in events with a Z boson in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV”, *Phys. Rev. D*, vol. 79 (2009) 052008, <https://doi.org/10.1103/PhysRevD.79.052008>.
- [4] D0 Collaboration, “A measurement of the ratio of inclusive cross sections $\sigma(p\bar{p} \rightarrow Z + \text{b-quark jet})/\sigma(p\bar{p} \rightarrow Z + \text{jet})$ at $\sqrt{s} = 1.96$ TeV”, *Phys. Rev. D*, vol. 83 (2011) 031105, <https://doi.org/10.1103/PhysRevD.83.031105>.
- [5] CMS Collaboration, “Measurement of the production cross sections for a Z boson and one or more b jets in pp collisions at $\sqrt{s} = 7$ TeV”, *JHEP*, vol. 06 (2014) 120, [https://doi.org/10.1007/JHEP06\(2014\)120](https://doi.org/10.1007/JHEP06(2014)120), [arXiv:1402.1521](https://arxiv.org/abs/1402.1521).
- [6] ATLAS Collaboration, “Measurement of differential production cross-sections for a Z boson in association with b-jets in 7 TeV proton-

- proton collisions with the ATLAS detector”, *JHEP*, vol. 10 (2014) 141, [doi:10.1007/JHEP10\(2014\)141](https://doi.org/10.1007/JHEP10(2014)141), [arXiv:1407.3643](https://arxiv.org/abs/1407.3643).
- [7] CMS Collaboration, “Measurements of the associated production of a Z boson and b jets in pp collisions at $\sqrt{s} = 8$ TeV”, *Eur. Phys. J. C*, vol. 77 (2017) 751, [doi:10.1140/epjc/s10052-017-5140-y](https://doi.org/10.1140/epjc/s10052-017-5140-y).
- [8] ATLAS Collaboration, “Measurements of the production cross-section for a Z boson in association with b-jets in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *JHEP*, vol. 07 (2020) 044, [doi:10.1007/JHEP07\(2020\)044](https://doi.org/10.1007/JHEP07(2020)044).
- [9] Particle Data Group, P. A. Zyla et al., “Review of particle physics”, *Prog. Theor. Exp. Phys.*, vol. 2020 (2020) 083C01, [doi:10.1093/ptep/ptaa104](https://doi.org/10.1093/ptep/ptaa104).
- [10] ATLAS Collaboration, “Measurements of the top quark branching ratios into channels with leptons and quarks with the ATLAS detector”, *Phys. Rev. D*, vol. 92 (2015) 072005, [doi:10.1103/PhysRevD.92.072005](https://doi.org/10.1103/PhysRevD.92.072005), [arXiv:1506.05074](https://arxiv.org/abs/1506.05074).
- [11] <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>
- [12] J. Alwall et al., “MadGraph 5 : Going Beyond”, *JHEP*, vol. 06 (2011) 128, [doi:10.1007/JHEP06\(2011\)128](https://doi.org/10.1007/JHEP06(2011)128), [arXiv:1106.0522](https://arxiv.org/abs/1106.0522).
- [13] R. Frederix and S. Frixione, “Merging meets matching in MC@NLO”, *JHEP* vol. 12 (2012) 061, [doi:10.1007/jhep12\(2012\)061](https://doi.org/10.1007/jhep12(2012)061), [arXiv:1209.6215](https://arxiv.org/abs/1209.6215).
- [14] M. L. Mangano, M. Moretti, and R. Pittau, “Multijet matrix elements and shower evolution in hadronic collisions: $Wb\bar{b} + n$ jets as a case study”, *Nucl. Phys. B*, vol. 632 (2002) 343, [doi:10.1016/S0550-3213\(02\)00249-3](https://doi.org/10.1016/S0550-3213(02)00249-3), [arXiv:hep-ph/0108069](https://arxiv.org/abs/hep-ph/0108069).

-
- [15] T. Sjöstrand et al., “An introduction to PYTHIA 8.2”, *Comput. Phys. Commun.*, vol. 191 (2015) 159, [doi:10.1016/j.cpc.2015.01.024](https://doi.org/10.1016/j.cpc.2015.01.024).
- [16] CMS Collaboration, “Event generator tunes obtained from underlying event and multiparton scattering measurements”, *Eur. Phys. J. C*, vol. 76 (2016) 155, [doi:10.1140/epjc/s10052-016-3988-x](https://doi.org/10.1140/epjc/s10052-016-3988-x), [arXiv:1512.00815](https://arxiv.org/abs/1512.00815).
- [17] CMS Collaboration, “Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements”, *Eur. Phys. J. C*, vol. 80 (2020) 4, [doi:10.1140/epjc/s10052-019-7499-4](https://doi.org/10.1140/epjc/s10052-019-7499-4), [arXiv:1903.12179](https://arxiv.org/abs/1903.12179).
- [18] R. D. Ball et al. “Parton distributions for the LHC run II”, *JHEP*, vol. 04 (2015) 040, [doi:10.1007/JHEP04\(2015\)040](https://doi.org/10.1007/JHEP04(2015)040).
- [19] E. Bothmann et al., “Event generation with Sherpa 2.2”, *SciPost Phys.*, vol. 7 (2019) 034, [doi:10.21468/SciPostPhys.7.3.034](https://doi.org/10.21468/SciPostPhys.7.3.034), [arXiv:1905.09127](https://arxiv.org/abs/1905.09127).
- [20] CMS Collaboration, “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV”, *JINST*, vol. 13 (2018) P05011, [doi:10.1088/1748-0221/13/05/P05011](https://doi.org/10.1088/1748-0221/13/05/P05011), [arXiv:1712.07158](https://arxiv.org/abs/1712.07158).
- [21] Nason, Paolo, “A new method for combining NLO QCD with shower Monte Carlo algorithms”, *JHEP*, vol. 11 (2004) 040, [doi:10.1088/1126-6708/2004/11/040](https://doi.org/10.1088/1126-6708/2004/11/040), [arXiv:hep-ph/0409146](https://arxiv.org/abs/hep-ph/0409146).
- [22] Frixione, Stefano and Nason, Paolo and Oleari, Carlo, “Matching NLO QCD computations with parton shower simulations: the POWHEG method”, *JHEP*, vol. 11 (2007) 070, [doi:10.1088/1126-6708/2007/11/070](https://doi.org/10.1088/1126-6708/2007/11/070), [arXiv:0709.2092](https://arxiv.org/abs/0709.2092).
- [23] Alioli, Simone and Nason, Paolo and Oleari, Carlo and Re, Emanuele, “A general framework for implementing NLO calculations in shower Monte

- Carlo programs: the POWHEG BOX”, *JHEP*, vol. 06 (2010) 043, [10.1007/JHEP06\(2010\)043](https://arxiv.org/abs/10.1007/JHEP06(2010)043), [arXiv:1002.2581](https://arxiv.org/abs/1002.2581).
- [24] CMS Collaboration, “Investigations of the impact of the parton shower tuning in PYTHIA 8 in the modelling of $t\bar{t}$ at $\sqrt{s} = 8$ and 13 TeV”, CMS Physics Analysis Summary, CMS-PAS-TOP-16-021, CERN, Geneva, (2016). <https://cds.cern.ch/record/2235192>.
- [25] Y. Li and F. Petriello, “Combining QCD and electroweak corrections to dilepton production in FEWZ”, *Phys. Rev. D* vol. 86 (2012) 094034, [doi:10.1103/PhysRevD.86.094034](https://arxiv.org/abs/doi:10.1103/PhysRevD.86.094034), [arXiv:1208.5967](https://arxiv.org/abs/1208.5967).
- [26] M. Grazzini, S. Kallweit, and M. Wiesemann, “Fully differential NNLO computations with MATRIX”, *Euro. Phys. J. C*, vol. 78 (2018) 537, [doi:10.1140/epjc/s10052-018-5771-7](https://arxiv.org/abs/doi:10.1140/epjc/s10052-018-5771-7), [arXiv:1711.06631](https://arxiv.org/abs/1711.06631).
- [27] N. Kidonakis, “Top quark production”, in *Proceedings, Helmholtz International Summer School on Physics of Heavy Quarks and Hadrons (HQ 2013): JINR, Dubna, Russia, July 15-28, 2013*, (2014) 139. [doi:10.3204/DESY-PROC-2013-03/Kidonakis](https://arxiv.org/abs/doi:10.3204/DESY-PROC-2013-03/Kidonakis), [arXiv:1311.0283](https://arxiv.org/abs/1311.0283).
- [28] M. Czakon, P. Fiedler, and A. Mitov, “Total top-quark pair-production cross section at hadron colliders through $O(\alpha_s^4)$ ”, *Phys. Rev. Lett.*, vol. 110 (2013) 252004, [doi:10.1103/PhysRevLett.110.252004](https://arxiv.org/abs/doi:10.1103/PhysRevLett.110.252004), [arXiv:1303.6254](https://arxiv.org/abs/1303.6254).
- [29] P. Bärnreuther, M. Czakon, and A. Mitov, “Percent level precision physics at the Tevatron: first genuine NNLO QCD corrections to $q\bar{q} \rightarrow t\bar{t} + X$ ”, *Phys. Rev. Lett.*, vol. 109 (2012) 132001, [doi:10.1103/PhysRevLett.109.132001](https://arxiv.org/abs/doi:10.1103/PhysRevLett.109.132001), [arXiv:1204.5201](https://arxiv.org/abs/1204.5201).

-
- [30] M. Beneke, P. Falgari, S. Klein, and C. Schwinn, “Hadronic top-quark pair production with NNLL threshold resummation”, *Nucl. Phys. B*, vol. 855 (2012) 695, [doi:10.1016/j.nuclphysb.2011.10.021](https://doi.org/10.1016/j.nuclphysb.2011.10.021), [arXiv:1109.1536](https://arxiv.org/abs/1109.1536).
- [31] M. Cacciari et al., “Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation”, *Phys. Lett. B*, vol. 710 (2012) 612, [doi:10.1016/j.physletb.2012.03.013](https://doi.org/10.1016/j.physletb.2012.03.013), [arXiv:1111.5869](https://arxiv.org/abs/1111.5869).
- [32] M. Czakon and A. Mitov, “NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels”, *JHEP*, vol. 12 (2012) 054, [doi:10.1007/JHEP12\(2012\)054](https://doi.org/10.1007/JHEP12(2012)054), [arXiv:1207.0236](https://arxiv.org/abs/1207.0236).
- [33] M. Czakon and A. Mitov, “NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction”, *JHEP*, vol. 01 (2013) 080, [doi:10.1007/JHEP01\(2013\)080](https://doi.org/10.1007/JHEP01(2013)080), [arXiv:1210.6832](https://arxiv.org/abs/1210.6832).
- [34] M. Czakon and A. Mitov, “Top++: A program for the calculation of the top-pair cross-section at hadron colliders”, *Comput. Phys. Commun.*, vol. 185 (2014) 2930, [doi:10.1016/j.cpc.2014.06.021](https://doi.org/10.1016/j.cpc.2014.06.021), [arXiv:1112.5675](https://arxiv.org/abs/1112.5675).
- [35] S. Agostinelli et al., “GEANT4—simulation toolkit”, *Nucl. Instrum. Meth. A*, vol. 506 (2003) 250, [doi:10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8).
- [36] CMS Collaboration, “Measurement of the Z boson production in association with at least two b jets in pp collisions at $\sqrt{s}=13$ TeV”, Analysis Note, CMS AN-20-187, CERN, Geneva, (2020).
- [37] CMS Collaboration, “Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s}=7$ TeV”, *JHEP*, vol. 01 (2011) 080, [arXiv:1012.2466](https://arxiv.org/abs/1012.2466).
- [38] CMS Collaboration, “Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC” *JINST*, vol. 16 (2021) P05014, [doi:10.1088/1748-0221/16/05/P05014](https://doi.org/10.1088/1748-0221/16/05/P05014), [arXiv:2012.06888](https://arxiv.org/abs/2012.06888).

- [39] CMS Collaboration, "Performance of the CMS muon trigger system in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JINST*, vol. 16 (2021) P07001, [doi:10.1088/1748-0221/16/07/P07001](https://doi.org/10.1088/1748-0221/16/07/P07001), [arXiv:2102.04790](https://arxiv.org/abs/2102.04790).
- [40] CMS Collaboration, "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC", *JINST*, vol. 16 (2021) P05014, [doi:10.1088/1748-0221/16/05/p05014](https://doi.org/10.1088/1748-0221/16/05/p05014), [arXiv:2012.06888](https://arxiv.org/abs/2012.06888).
- [41] Bodek, Arie and van Dyne, Aaron and Han, J. and Sakumoto, W. and Strel'nikov, A, "Extracting Muon Momentum Scale Corrections for Hadron Collider Experiments", *Eur. Phys. J. C*, vol.72 (2012) 2194, [doi:10.1140/epjc/s10052-012-2194-8](https://doi.org/10.1140/epjc/s10052-012-2194-8).
- [42] Saranya Ghosh et al., "Performance of MET reconstruction and pileup mitigation techniques in CMS", *Nucl. and Part. Phys. Proc.*, vol. 273 (2016) 2512, [doi:10.1016/j.nuclphysbps.2015.09.442](https://doi.org/10.1016/j.nuclphysbps.2015.09.442).
- [43] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet user manual", *Eur. Phys. J. C*, vol. 72 (2012) 1896, [doi:10.1140/epjc/s10052-012-1896-2](https://doi.org/10.1140/epjc/s10052-012-1896-2), [arXiv:1111.6097](https://arxiv.org/abs/1111.6097).
- [44] CMS Collaboration, "Pileup mitigation at CMS in 13 TeV data", *JINST*, vol. 15 (2020) P09018, [doi:10.1088/1748-0221/15/09/p09018](https://doi.org/10.1088/1748-0221/15/09/p09018), [arXiv:2003.00503](https://arxiv.org/abs/2003.00503).
- [45] CMS Collaboration, "Pileup jet identification", CMS Physics Analysis Summary, CMS-PAS-JME-13-005, CERN, Geneva, (2013). <https://cds.cern.ch/record/1581583>.
- [46] M. Meena, S. Bansal, S. B. Beri, and R. Gupta, "Measurement of $Z + b$ jets cross section at $\sqrt{s} = 13$ TeV with full sample of Run 2", Analysis Note, CMS AN-19-179, CERN, Geneva, (2019).

- [47] CMS collaboration, “Measurement of differential cross sections for Z bosons produced in association with charm jets in pp collisions at $\sqrt{s} = 13$ TeV”, *JHEP* 04 (2021) 109, [doi:10.1007/JHEP04\(2021\)109](https://doi.org/10.1007/JHEP04(2021)109).
- [48] S. Schmitt, “TUnfold: an algorithm for correcting migration effects in high energy physics”, *JINST*, vol. 7 (2012) T10003, [doi:10.1088/1748-0221/7/10/t10003](https://doi.org/10.1088/1748-0221/7/10/t10003), [arXiv:1205.6201](https://arxiv.org/abs/1205.6201).
- [49] A. N. Tikhonov, “Solution of incorrectly formulated problems and the regularization method”, *Soviet Math. Dokl.*, vol. 4 (1963) 1035.
- [50] CMS Collaboration, “Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS”, *Eur. Phys. J. C*, vol. 81 (2021) 800, [doi:10.1140/epjc/s10052-021-09538-2](https://doi.org/10.1140/epjc/s10052-021-09538-2), [arXiv:2104.01927](https://arxiv.org/abs/2104.01927).
- [51] CMS Collaboration, “CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV”, CMS Physics Analysis Summary, CMS-PAS-LUM-17-004, CERN, Geneva, (2018). <http://cds.cern.ch/record/2621960>.
- [52] CMS Collaboration, “CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV”, CMS Physics Analysis Summary, CMS-PAS-LUM-18-002, CERN, Geneva, (2019). <https://cds.cern.ch/record/2676164>.
- [53] CMS Collaboration, “Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV”, *JHEP*, vol. 07 (2018) 161, [doi:10.1007/JHEP07\(2018\)161](https://doi.org/10.1007/JHEP07(2018)161), [arXiv:1802.02613](https://arxiv.org/abs/1802.02613).
- [54] D. Bourilkov, R. C. Group, and M. R. Whalley, “LHAPDF: PDF use from the Tevatron to the LHC”, in *TeV4LHC Workshop - 4th meeting Batavia, Illinois, October 20-22, 2005* (2006), [arXiv:hep-ph/0605240](https://arxiv.org/abs/hep-ph/0605240).

- [55] CMS Collaboration, “Measurement of the production cross section for $Z + b$ jets in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *Phys. Rev. D*, vol. 105 (2022) 092014, [doi:10.1103/physrevd.105.092014](https://doi.org/10.1103/physrevd.105.092014), [arXiv:2112.09659](https://arxiv.org/abs/2112.09659).

Chapter 6

Summary

Several experiments have been designed and built to test the predictions of the Standard Model (SM) and beyond SM (BSM) scenarios. Among them, the Large Hadron Collider (LHC) at the European Council for Nuclear Research (CERN) is the largest and most powerful particle accelerator ever built. The proton beams are made to collide at four different locations along the LHC ring. At these positions, four giant detectors are placed to record every possible outcome of the collision. The Compact Muon Solenoid (CMS) is one of its general purpose detectors for SM measurements and BSM searches due to its high precision detection of particles from collision events. The work presented in this thesis is based on the contribution to CMS detector upgrade and physics measurement. This thesis mainly reports two tasks:

- Assembly and quality control testing of Gas Electron Multiplier (GEM) subdetectors for the upgrade of the first endcap muon station of the CMS detector.
- The cross sections measurement of the $Z+b$ -jets processes using proton-proton (pp) collisions data at $\sqrt{s} = 13$ TeV.

These tasks are briefly summarized in the following Sections.

6.1 Assembly and testing of Gas Electron Multipliers

To study a wide range of physics topics in SM and BSM scenarios muons play a crucial role as they act as the cleanest signal. The CMS muon detector system earlier had three different types of gaseous subdetectors: cathode strip chambers (CSC), drift tubes (DT), and resistive plate chambers (RPC) for collecting data up to Run2. These detectors used different detection technologies to improve the timing and spatial resolution of the system. Precision measurements and Level 1 (L1) triggering are provided by DT (CSC) in the barrel (endcap) region. Additionally, RPC provide redundant trigger and coarse position measurement in both barrel and endcap regions, but were not implemented beyond $|\eta| > 1.9$ because they cannot sustain the high background particle rates in this region. As a result, the forward region ($1.6 < |\eta| < 2.2$) is completed with new muon subdetectors known as GE1/1 based on GEM technology to further improve muon trigger and spatial resolution. The GEM technology is an excellent choice for this region, due to its thin profile and ability to operate at particle fluxes much higher than those expected in the forward region under HL-LHC conditions. The GE1/1 chambers have a detection efficiency of more than 97% with time and position resolution of 10 ns and $260\ \mu\text{m}$ to $340\ \mu\text{m}$, respectively.

Total 144 GE1/1 chambers were required to upgrade the first endcap muon station of the CMS detector during the Long Shutdown (LS2). Eight GE1/1 chambers were assembled and tested at Panjab University in Chandigarh. The assembly technique included the preparation of different parts of chambers called pre-assembly and then the final assembly of these parts in the class-100 clean room. During pre-assembly the readout board, drift board, internal frames, and external frame of the chamber are prepared. The final assembly of the chamber was completed in the clean room, which included the assembly of the GEM foil stack and closing of the

chamber.

After the assembly of eight chambers, various quality control (QC) tests up to QC4 are performed at the Panjab University site and QC5 at Delhi University to ensure that the chamber does not have any mechanical or electrical issues that could impact its operation. The chambers are shipped to CERN for further superchamber assembly and QC testing (upto QC8) after passing all tests up to QC5. The first QC test, known as QC2, was carried out to check the quality of a GEM foil before, during, and after chamber assembly. Then, QC3 and QC4 tests are performed to check for gas leaks and possible malfunctions, as well as defects in the HV circuit, and intrinsic noise rate, respectively. Then, QC5, QC6, QC7, and QC8 tests are performed to measure effective gas gain, voltage vs. current characteristics, all electronic component connections, and the efficiency and time resolution the chamber.

All chambers passed the QC tests performed at the Panjab University site and later at the CERN site. These chambers have already been installed during LS2 in the endcap of the CMS muon spectrometer and are part of the data taking.

6.2 Measurement of the production cross section for $Z + b$ -jets in proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$

In this thesis, the measurement of the fiducial cross section for the production of a Z boson in association with at least one bottom quark jet is performed. This is the first measurement of these processes based on data collected by the CMS experiment during the 2016–2018 LHC running period. The $Z + b$ -jets is a crucial process at the LHC for precision testing of electroweak and perturbative quantum chromodynamics prediction (pQCD) calculations by comparing the measured cross sections to theoretical predictions. The $Z + b$ -jets measurements help to extract

PDFs for the b quarks and constraint the PDF uncertainty in the measurement of the W boson mass. These processes are the dominant irreducible backgrounds to various SM and BSM processes.

The measurement is performed using pp collision data corresponding to an integrated luminosity of 137 fb^{-1} at a $\sqrt{s}$ of 13 TeV. The integrated cross section as well as the differential and normalized differential cross section are measured for the electron, muon, and combined channels. The differential cross section for $Z + \geq 1$ b-jet event is measured as a function of various kinematic observables: the transverse momentum of the dilepton; the highest p_T (leading) b-jet ($p_T^{\text{b-jet}}$); the absolute pseudorapidity of the b-jet ($|\eta^{\text{b-jet}}|$); the azimuthal angle difference between the Z boson and the b-jet; the rapidity difference between the Z boson and the b-jet; and the angular correlation variable. These observables are directly sensitive to the b quark PDF, details of the pQCD calculations, initial-state-radiation (ISR), final-state-radiation (FSR), and multiple parton interactions (MPIs). The normalized differential cross section is measured which can better constraint the theoretical predictions because of the cancellation of the correlated systematic uncertainties. To understand the dynamics of $Z + \geq 1$ b-jet and $Z + \geq 2$ b-jets, the ratio of their differential cross section is measured as a function of $p_T^{\text{b-jet}}$ and $|\eta^{\text{b-jet}}|$, which allows a further reduction in various experimental uncertainties such as those related to the jet energy scale measurement. The $Z + \geq 2$ b-jets analysis has been performed by other collaborating group.

The fiducial phase space regions are defined by selecting Z boson with at least one b jet with $p_T > 30 \text{ GeV}$ and $|\eta| < 2.4$. The production cross sections within the fiducial region, are measured to be $6.52 \pm 0.04 \text{ (stat)} \pm 0.40 \text{ (syst)} \pm 0.14 \text{ (theo) pb}$ and $0.65 \pm 0.03 \text{ (stat)} \pm 0.07 \text{ (syst)} \pm 0.02 \text{ (theo) pb}$ for the $Z + \geq 1$ b-jet and $Z + \geq 2$ b-jets processes, respectively. In the $Z + \geq 1$ b-jet and $Z + \geq 2$ b-jets final states, the measured integrated cross section values are well described by both MG5_aMC (leading order (LO), NNPDF 3.0, CUETP8M1) and MG5_aMC (next-to-

LO (NLO), NNPDF 3.1, CP5) simulations. Whereas MG5_aMC (NLO, NNPDF 3.1, CP5), MG5_aMC (NLO, NNPDF 3.0, CUETP8M1), and SHERPA overestimate data by $\approx 10\%$ (29%), $\approx 18\%$ (38%) and $\approx 24\%$ (21%), respectively, for the $Z + \geq 1$ b-jet ($Z + \geq 2$ b-jets) final states. The measured value of the cross section ratio of the $Z + \geq 2$ b-jets and $Z + \geq 1$ b-jet for the combined channel is 0.100 ± 0.005 (stat) ± 0.007 (syst) ± 0.003 (theo), which is well described by the MG5_aMC (LO, NNPDF 3.0, CUETP8M1) and SHERPA calculations but overestimated by MG5_aMC (NLO) prediction. The shapes of various kinematic observables are well described by the SHERPA prediction but are not completely described by MG5_aMC (LO) and MG5_aMC (NLO) predictions. Since all predictions are normalized to the inclusive $Z + \text{jets}$ next-to-NLO cross section, differences between MG5_aMC (NLO) and MG5_aMC (LO) results could be attributable to variations in shapes of observables and settings (parton distribution functions, Monte Carlo (MC) tunes, matching schemes) used in those simulations. This measurements can be used as an input for the further optimization of the simulation parameters and b quark PDF.

6.3 Outlook

During LS2, the GE/1 chambers have been installed in the first endcap muon station of the CMS detector. It has been observed that this upgrade maintained the L1 muon trigger thresholds at low p_T values, which increased the efficiency for capturing interesting physics processes involving soft leptons. Maintaining low p_T muon trigger thresholds is important for a wide range of physics studies ranging from the Higgs sector measurements to new physics searches. Furthermore, during the LS3, the accelerator infrastructure will be updated to deliver unprecedented instantaneous luminosity ranging from 5.0 to $7.5 \times 10^{34} \text{cm}^{-2} \text{s}^{-1}$ providing an integrating luminosity of 3000 to 4000 fb^{-1} . During the LS3 period, the muon detector will be upgraded to improve and maintain the performance of the muon triggering and

tracking systems during the HL-LHC. There will be an installation of an additional set of GEM subdetectors known as GE2/1 and ME0 in the muon system during LS3. These additional GEM subdetectors will be used with the new silicon tracker and track trigger installed during LS3 to make a “combined muon trigger”. As a result, these additional GEM subdetectors will add redundancy to the forward region and further improve the L1 muon trigger.

In this thesis, the measurement of the cross section for the production of a Z boson in association with at least one bottom quark jet with pp collision data at $\sqrt{s} = 13 \text{ TeV}$ is reported. The measurement of Z + b-jets processes is important for all collider experiments. The study of Z + b-jets processes provides an excellent environment to check numerous pQCD predictions and helps to improve the understanding of the current theoretical prediction of b-jets. In the present analysis, the datasets were compared with various MC predictions based on leading order and next-to-leading order matrix elements with the most recent different parton distribution functions, MC tunes, matching schemes, etc. The current MC simulation is found to be mismodelled. Present measurements can be used as input to further optimization of the simulation parameters and b quark PDF to describe data well.

Publications/Conferences

Publications and CMS Analysis Notes with Direct Contribution

1. CMS Collaboration “Measurement of the production cross section for $Z + b$ jets in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *Phys. Rev. D*, vol. 105 (2022) 092014, [doi:10.1103/physrevd.105.092014](https://doi.org/10.1103/physrevd.105.092014), [arXiv:2112.09659](https://arxiv.org/abs/2112.09659).
2. CMS Collaboration “Benchmarking LHC background particle simulation with the CMS triple-GEM detector”, *JINST*, vol. 16 (2021) P12026, [doi:10.1088/1748-0221/16/12/P12026](https://doi.org/10.1088/1748-0221/16/12/P12026).
3. CMS Collaboration “Measurements of Z bosons plus jets using variables sensitive to double parton scattering in pp collisions at 13 TeV” *JHEP*, vol 10 (2021) 176, [doi:10.1007/JHEP10\(2021\)176](https://doi.org/10.1007/JHEP10(2021)176).
4. CMS Collaboration “Interstrip capacitances of the readout board used in large triple-GEM detectors for the CMS Muon Upgrade”, *JINST*, vol. 15 (2020) P12019, [doi:10.1088/1748-0221/15/12/p12019](https://doi.org/10.1088/1748-0221/15/12/p12019).
5. CMS Collaboration “Performance of the CMS Level-1 Trigger in proton-proton collisions at 13 TeV”, *JINST*, vol. 15 (2020) P10017, [doi:10.1088/1748-0221/15/10/p10017](https://doi.org/10.1088/1748-0221/15/10/p10017).
6. CMS Collaboration, “Measurement of $Z + b$ jets cross section in proton-proton collisions at $\sqrt{s} = 13$ TeV”, CMS PAS-20-015 (2021), <https://cds.cern.ch/record/2776981>.
7. CMS Collaboration, “Measurement of distributions sensitive to double parton scattering using Z bosons produced in association with jets at 13 TeV”, CMS

PAS-SMP-20-009 (2021), <https://cds.cern.ch/record/2755115>.

8. CMS Collaboration, “Level-jets and energy sums trigger performance with full 2017 dataset”, CMS-DP-2018/004 (2018), <https://cds.cern.ch/record/230554>.

* **CMS AN - CMS Analysis Note**

* **CMS DN - CMS Detector Note**

* **CMS DP - CMS Detector Performance Note**

Published conference proceedings

1. M. Meena, S. Bansal, S. B. Beri on behalf of the CMS Collaboration, “Study for vector bosons production in association with heavy-flavor jets in proton-proton collisions”, *Springer Proc. Phys.*, vol. 277 (2022) 161, https://doi.org/10.1007/978-981-19-2354-8_29.
2. M. Meena on behalf of the ATLAS and CMS Collaborations, “Recent jet and jet substructure measurements at the LHC, and ML based tagging”, *SciPost. Phys. Proc.*, (2022) 022, [doi:10.21468/SciPostPhysProc.10.022](https://doi.org/10.21468/SciPostPhysProc.10.022).
3. M. Meena on behalf of the ALICE, ATLAS, CMS and LHCb Collaborations, “Experimental measurement of Heavy Flavors and jets”, *PoS LHCP*, 2021 (2021) 087 [doi:10.22323/1.397.0087](https://doi.org/10.22323/1.397.0087).
4. M. Meena, J. Babbar, S. Bansal, S. B. Beri, V. Bhatnagar, S. Chauhan, R. Gupta, A. Kaur, H. Kaur, S. Kumar, P. Kumari, J.B. Singh and A.K. Viridi “Fabrication and Characterization of Gaseous Detector for the identification of High Energy Particles”, *IOP Conference Series: Materials Science and Engineering*, vol. 1033 (2021) 012055, [doi:1088/1757-899x/1033/1/012055](https://doi.org/10.1088/1757-899x/1033/1/012055).
5. M. Meena, A. K. Viridi, P. Kumari, R. Gupta, J. S. Shahi, S. Bansal, V. Bhatnagar, J. B. Singh and Manisha, “Quality Control Testing of GEM Detector”, *Proceedings of the DAE Symp. on Nucl. Phys.*, vol. 62 (2017), http://inis.iaea.org/search/search.aspx?orig_q=RN:49014821.

Conference proceedings under review

1. M. Meena on behalf of the CMS Collaboration, “Production of vector bosons in association with jets in CMS”, *Proceeding under review in Zenodo*, <https://cds.cern.ch/record/2798327>, CMS-CR-2022-053 (2022).

Papers presented in Conferences, Workshops, Symposiums

1. M. Meena on behalf of the CMS Collaboration, “Azimuthal correlation measurements from Z+jets and dijets”, in *REF22: Resummation, Evolution, Factorization 2022*, 31st October-4th November 2022, University of Montenegro, Montenegro.
2. M. Meena on behalf of the CMS Collaboration, “Vector boson associated with jets in CMS”, in 30th *International Symposium on Lepton Photon Interactions at High Energies*, was held on 10th-14th January, 2022 at the University of Manchester.
3. M. Meena on behalf of the CMS Collaboration, “Recent measurement of heavy flavors jets with CMS detector” in *Posters@LHCC* on 18th November, 2021.
4. M. Meena on behalf of the ATLAS and CMS Collaborations, “Recent jet and jet substructure measurements at the LHC, and ML based tagging”, in 50th *International Symposium on Multiparticle Dynamics (ISMD2021)*, held on 12th-16th July, 2021.
5. M.Meena on behalf of the ALICE, ATLAS, CMS, and LHCb Collab-

- orations, “Experimental measurement of Heavy Flavors and jets”, in *The Ninth Annual Conference on Large Hadron Collider Physics - LHCP2021*, held on 7th-12th June, 2021.
6. M. Meena on behalf of the CMS Collaboration, “Vector boson plus heavy-flavor jets measurements at CMS”, in *Pheno2021: Phenomenology 2021 Symposium*, University of Pittsburgh, Pennsylvania, held on 24th-26th May, 2021.
 7. M. Meena for the CMS Collaboration, “Study for vector bosons production in association with heavy-flavor jets” in 24th *DAE-BRNS High Energy Physics Symposium 2020*, Bhubaneswar, India held on 14th-20th December, 2020.
 8. M. Meena on behalf of the CMS Collaboration, “Study Of Z+ Heavy Flavor Jets In Proton Proton Collision” in 14th *Chandigarh Science Congress (CHASCON)*, held at Panjab University, Chandigarh, India on 17th-19th December, 2020.
 9. M. Meena, “Fabrication and Characterization of Gaseous Detector for the identification of High Energy Particles” in *International Conference on Integrated Interdisciplinary Innovations in Engineering* held at UIET, Panjab University, Chandigarh, India held on 28th-30th August, 2020.
 10. M. Meena on behalf of the CMS Collaboration, “Study for Z boson production in association with b-jets in proton-proton collision at 13 TeV” in 23rd *DAE-BRNS High Energy Physics Symposium* held at IIT Madras, India on 10th-14th December, 2018.
-

Other workshops and schools attended:

1. 9th International workshop on Multiple Partonic Interactions at the LHC held 11th-15th December, 2017, Shimla, India.
2. XI SERC School Experimental High Energy Physics from 7th-27th November, 2017 NISER at Odisha, India.

Publications as co-author with CMS Collaboration

1. CMS Collaboration “Search for a W’ boson decaying to a vector-like quark and a top or bottom quark in the all-jets final state at $\sqrt{s} = 13$ TeV”, *JHEP*, vol. 9 (2022) 88, [doi:10.1007/JHEP09\(2022\)088](https://doi.org/10.1007/JHEP09(2022)088).
2. CMS Collaboration, “Measurement of the Drell-Yan forward-backward asymmetry at high dilepton masses in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JHEP*, vol. 08 (2022) 063, [doi:10.1007/JHEP08\(2022\)063](https://doi.org/10.1007/JHEP08(2022)063).
3. CMS Collaboration, “Nuclear modification of Υ states in pPb collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *Phys. Lett. B*, vol. 835 (2022) 137397, [doi:10.1016/j.physletb.2022.137397](https://doi.org/10.1016/j.physletb.2022.137397).
4. CMS Collaboration, “Search for Higgs boson pair production in the four b quark final state in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *Phys. Rev. Lett.*, vol. 129 (2022) 081802, [doi:10.1103/PhysRevLett.129.081802](https://doi.org/10.1103/PhysRevLett.129.081802).
5. CMS Collaboration, “Inclusive nonresonant multilepton probes of new phenomena at $\sqrt{s} = 13$ TeV”, *Phys. Rev. D*, vol. 105 (2022) 112007, [arXiv:2202.08676](https://arxiv.org/abs/2202.08676).
6. CMS Collaboration, “Measurement of the Higgs boson width and evidence of its off-shell contributions to ZZ production”, *Nature Phys.*, vol. 18 (2022) 1329, [arXiv:2202.06923](https://arxiv.org/abs/2202.06923).
7. CMS Collaboration, “Search for new physics in the lepton plus missing transverse momentum final state in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JHEP*, vol. 07 (2022) 067, [arXiv:2202.06075](https://arxiv.org/abs/2202.06075).

8. CMS Collaboration, “Search for invisible decays of the Higgs boson produced via vector boson fusion in proton-proton collisions at $\sqrt{s}=13$ TeV”, *Phys. Rev. D*, vol. 105 (2022) 092007, [arXiv:2201.11585](#).
9. CMS Collaboration, “Search for resonances decaying to three W bosons in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *Phys. Rev. Lett.*, vol. 129 (2022) 021802, [arXiv:2201.08476](#).
10. CMS Collaboration, “Identification of hadronic tau lepton decays using a deep neural network”, *JINST*, vol. 17 (2022) P07023, [arXiv:2201.08458](#).
11. CMS Collaboration, “Search for charged-lepton flavor violation in top quark production and decay in pp collisions at $\sqrt{s} = 13$ TeV ”, *JHEP*, vol. 06 (2022) 082, [arXiv:2201.07859](#).
12. CMS Collaboration, “Precision measurement of the W boson decay branching fractions in proton-proton collisions at $\sqrt{s} = 13$ TeV ”, *Phys. Rev. D*, vol. 105 (2022) 072008, [arXiv:2201.07861](#).
13. CMS Collaboration, “Measurement of the inclusive and differential $t\bar{t}\gamma$ cross sections in the dilepton channel and effective field theory interpretation in proton-proton collisions at $\sqrt{s} = 13$ TeV ”, *JHEP*, vol. 05 (2022) 091, [arXiv:2201.07301](#).
14. CMS Collaboration, “Search for long-lived heavy neutral leptons with displaced vertices in proton-proton collisions at $\sqrt{s}=13$ TeV”, *JHEP*, vol. 07 (2022) 081, [arXiv:2201.05578](#).
15. CMS Collaboration, “Search for higgsinos decaying to two Higgs bosons and missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV ”, *JHEP*, vol. 05 (2022) 014, [arXiv:2201.04206](#).

16. CMS Collaboration, “Observation of the B_c^+ Meson in Pb-Pb and pp Collisions at $\sqrt{s_{NN}}=5.02$ TeV and Measurement of its Nuclear Modification Factor ”, *Phys. Rev. Lett.*, vol. 128 (2022) 252301, [arXiv:2201.02659](#).
17. CMS Collaboration, “Search for single production of a vector-like T quark decaying to a top quark and a Z boson in the final state with jets and missing transverse momentum at $\sqrt{s} = 13$ TeV ”, *JHEP*, vol. 05 (2022) 093, [arXiv:2201.02227](#).
18. CMS Collaboration, “Search for long-lived particles decaying into muon pairs in proton-proton collisions at $\sqrt{s} = 13$ TeV collected with a dedicated high-rate data stream ”, *JHEP*, vol. 04 (2022) 062, [arXiv:2112.13769](#).
19. CMS Collaboration, “Search for resonances decaying to three W bosons in the hadronic final state in proton-proton collisions at $\sqrt{s}=13$ TeV”, *Phys. Rev. D*, vol. 106 (2022) 012002, [arXiv:2112.13090](#).
20. CMS Collaboration, “Probing Charm Quark Dynamics via Multi-particle Correlations in Pb-Pb Collisions at $\sqrt{s_{NN}}=5.02$ TeV”, *Phys. Rev. Lett.*, vol. 129 (2022) 022001, [arXiv:2112.12236](#).

Complete list of 269 published papers as a co-author in CMS: <https://inspirehep.net/literature?sort=mostrecent&size=25&page=1&q=meena%20meena&subject=Experiment-HEP&collaboration=CMS>

*Selected
Reprints*

Measurement of the production cross section for $Z + b$ jets in proton-proton collisions at $\sqrt{s} = 13$ TeV

A. Tumasyan *et al.*^{*}
(CMS Collaboration)



(Received 17 December 2021; accepted 19 April 2022; published 31 May 2022)

The measurement of the cross section for the production of a Z boson, decaying to dielectrons or dimuons, in association with at least one bottom quark jet is performed with proton-proton collision data at $\sqrt{s} = 13$ TeV. The data sample corresponds to an integrated luminosity of 137 fb^{-1} , collected by the CMS experiment at the LHC during 2016–2018. The integrated cross sections for $Z + \geq 1 b$ jet and $Z + \geq 2 b$ jets are reported for the electron, muon, and combined channels. The fiducial cross sections in the combined channel are $6.52 \pm 0.04(\text{stat}) \pm 0.40(\text{syst}) \pm 0.14(\text{theo}) \text{ pb}$ for $Z + \geq 1 b$ jet and $0.65 \pm 0.03(\text{stat}) \pm 0.07(\text{syst}) \pm 0.02(\text{theo}) \text{ pb}$ for $Z + \geq 2 b$ jets. The differential cross section distributions are measured as functions of various kinematic observables that are useful for precision tests of perturbative quantum chromodynamics predictions. The ratios of integrated and differential cross sections for $Z + \geq 2 b$ jets and $Z + \geq 1 b$ jet processes are also determined. The value of the integrated cross section ratio measured in the combined channel is $0.100 \pm 0.005(\text{stat}) \pm 0.007(\text{syst}) \pm 0.003(\text{theo})$. All measurements are compared with predictions from various event generators.

DOI: [10.1103/PhysRevD.105.092014](https://doi.org/10.1103/PhysRevD.105.092014)

I. INTRODUCTION

A Z boson accompanied by bottom (b) jets can originate from quark-gluon, quark-antiquark, and gluon-gluon interactions in proton-proton (pp) collisions. These processes can be used to test perturbative quantum chromodynamics (pQCD) predictions by comparing the measured cross sections with theoretical predictions. They also provide information on the b quark parton distribution functions (PDFs) [1]. The measurement of $Z + b$ jets helps to estimate the uncertainty coming from the PDF choice in the W boson mass measurement [2]. These processes are the dominant background for Higgs boson production in association with a Z boson (ZH , $H \rightarrow b\bar{b}$) in the standard model. They are also a source of background in many scenarios beyond the standard model such as the production of supersymmetric Higgs bosons in association with a b quark, and new generations of heavy quarks decaying to a Z boson and a b quark. Examples of Feynman diagrams of the signal process with quark-gluon, quark-antiquark, and gluon-gluon interactions for $Z + 1 b$ jet and $Z + 2 b$ jets final states are shown in Fig. 1. Two different approaches are available for calculating the $Z + b$ jets production, the

five-flavor and four-flavor schemes, and, as shown in Ref. [3], they both yield consistent results.

Previous measurements have been performed on $Z + b$ jets at $\sqrt{s} = 1.96$ TeV by the CDF [4] and D0 [5] Collaborations at the Fermilab Tevatron in proton-antiproton collisions. At the CERN LHC similar measurements have been performed at 7 TeV by the CMS [6] and ATLAS [7] Collaborations, at 8 TeV by CMS [8], and at 13 TeV by ATLAS [9], all in pp collisions.

In this paper, the integrated cross sections together with differential distributions of the $Z(\rightarrow \ell\ell) + \geq 1 b$ jet and $Z(\rightarrow \ell\ell) + \geq 2 b$ jets are reported (where $\ell\ell = ee$ or $\mu\mu$). The differential distributions normalized to the corresponding fiducial cross sections are presented as well. In the measurements of normalized distributions, several systematic uncertainties cancel, which allows a more precise comparison with theory. The differential and normalized differential cross section distributions for $Z + \geq 1 b$ jet production are measured with respect to six kinematic observables: the transverse momentum of the dilepton (p_T^Z); the highest p_T (leading) b jet ($p_T^{b \text{ jet}}$); the absolute pseudorapidity of the b jet ($|\eta^{b \text{ jet}}|$); the azimuthal angle difference between the Z boson and the b jet ($\Delta\phi^{(Z, b \text{ jet})}$); the rapidity difference between the Z boson and the b jet ($\Delta y^{(Z, b \text{ jet})}$); and the angular correlation variable, which is defined as the separation in the pseudorapidity and azimuthal distance between the Z boson and the b jet [$\Delta R^{(Z, b \text{ jet})} = \sqrt{(\Delta\eta^{(Z, b \text{ jet})})^2 + (\Delta\phi^{(Z, b \text{ jet})})^2}$].

The differential and normalized differential cross sections for the $Z + \geq 2 b$ jets production are measured as

^{*}Full author list given at the end of the article.

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

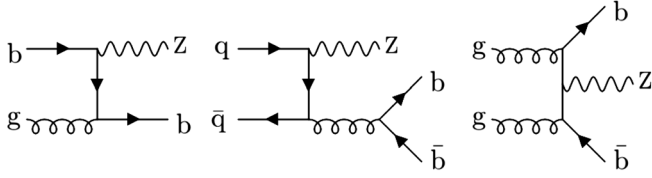


FIG. 1. Examples of Feynman diagrams for $Z + 1 b$ jet (left) and $Z + 2 b$ jets (middle and right).

functions of nine kinematic observables describing the final state: the transverse momenta of the Z boson and the two highest p_T (leading and subleading) b jets; the absolute pseudorapidity of the highest p_T b jet; the angular correlations between the Z boson and the b jets; the invariant mass of the two b jets (m_{bb}); the invariant mass of the Z boson and the two b jets (m_{Zbb}); the angular separation between two b jets (ΔR_{bb}); the minimum separation between the Z boson and the two b jets ($\Delta R_{Zbb}^{\min} = \min(\Delta R_{Z,b_1}, \Delta R_{Z,b_2})$, where b_1 and b_2 are the two leading b jets); and the asymmetry of the $Z + \geq 2 b$ jets system ($A_{Zbb} = [\Delta R_{Zbb}^{\max} - \Delta R_{Zbb}^{\min}] / [\Delta R_{Zbb}^{\max} + \Delta R_{Zbb}^{\min}]$, where $\Delta R_{Zbb}^{\max} = \max(\Delta R_{Z,b_1}, \Delta R_{Z,b_2})$ is the maximum separation between the Z and the two b jets).

In addition, the integrated cross section ratios of $Z + \geq 2 b$ jets to $Z + \geq 1 b$ jet are measured. The differential cross section ratios are measured as functions of p_T and $|\eta|$ of the leading b jet. These measurements allow a further reduction in various experimental uncertainties such as those related to the jet energy scale measurement.

This study at higher $\sqrt{s}$ allows the exploration of the more extreme regions of phase space. The measurement of $p_T^{b \text{ jet}}$ and p_T^Z are sensitive to the b quark PDF and the details of the pQCD calculations. The angular separation is sensitive to initial- and final-state radiation and multiparton interactions where extra jet(s) affect the corresponding variables [10]. The invariant mass distributions are sensitive to two- b -jet resonances. Measurements are compared with the predictions from various Monte Carlo (MC) event generators to test the higher-order pQCD calculations and simulations of parton showering, hadronization, and multiparton interactions.

Tabulated results are provided in the HEPData record for this analysis [11].

The outline of this paper is as follows. Section II gives a brief description of the CMS detector. Section III describes the data and simulated samples used for the validation and unfolding studies, and Sec. IV describes the event selection criteria. The unfolding procedure and systematic uncertainties are discussed in Sec. V, and the results are presented in Sec. VI. Finally, the paper is summarized in Sec. VII.

II. THE CMS DETECTOR

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a

magnetic field of 3.8 T. Within the solenoid volume there is a silicon pixel and strip tracker covering a pseudorapidity region $|\eta| < 2.5$ (3) in 2016 (2017–2018) data-taking period [12], a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter. Forward calorimeters, made of steel and quartz fibers, extend the coverage to $|\eta| < 5.0$ provided by the barrel and end cap detectors. Muons are detected in gas-ionization chambers embedded in the steel flux-return yoke outside the solenoid. Events of interest are selected using a two-tiered trigger system [13]. The first level (L1), composed of custom hardware processors, uses information from the calorimeters and muon detectors to select events at a rate of around 100 kHz within a latency of less than 4 μ s [14]. The second level, known as the high-level trigger, consists of a farm of processors running a version of the full event reconstruction software optimized for fast processing and reduces the event rate to around 1 kHz before data storage.

A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, is reported in Ref. [15].

III. DATA AND SIMULATED SAMPLES

This analysis is performed using a data sample of pp collisions at a center-of-mass energy of 13 TeV collected by the CMS experiment during the 2016 (35.9 fb $^{-1}$), 2017 (41.5 fb $^{-1}$), and 2018 (59.7 fb $^{-1}$) data-taking periods with an integrated luminosity of 137 fb $^{-1}$.

Several Monte Carlo event generators are used to simulate signal and background processes. The Drell–Yan (DY) process with exclusive jet multiplicity up to 2 is simulated at next-to-leading order (NLO) precision by MadGraph5_aMC@NLO (denoted MG5_aMC) [16] version 2.3.2.2 for 2016 data and version 2.6.0 for the 2017–2018 data with the FxFX [17] matching between the jets from matrix element calculations and parton showers. The NNPDF 3.0 NLO and NNPDF 3.1 next-to-NLO (NNLO) PDF sets [18] are used for the 2016 and 2017–2018 data-taking periods, respectively.

A second sample is produced with MG5_aMC version 2.2.2 (2.4.2) for the 2016 (2017–2018) data-taking period using leading order (LO) matrix elements for $pp \rightarrow Z + n$ jets ($n \leq 4$), with the MLM matching scheme [19]. The NNPDF 3.0 LO and NNPDF 3.1 NNLO PDF sets are used for the 2016 and 2017–2018 data-taking periods, respectively. The MG5_aMC generator is interfaced with PYTHIA v8.212 (v8.230) [20] for parton showering, hadronization, and the underlying event simulation using tune CUETP8M1 [21] (CP5 [22]) for the 2016 (2017–2018) data-taking period. Matching between the matrix element generators and the parton shower jets is done with the matching scales set at 19 and 30 GeV for the LO and NLO MG5_aMC generators, respectively.

A third inclusive sample has been produced with SHERPA v2.2.4 [23] to generate $pp \rightarrow Z + n\text{jets}$ events, with $n \leq 2$ at NLO and $n = 3, 4$ at LO. The merging with the SHERPA parton shower is done via the MEPS@NLO prescription [24–26] with a matching scale of 20 GeV. The NNPDF 3.0 NLO PDF and a dedicated set of tuned parton shower parameters developed by the SHERPA authors are used. In the matrix element calculation, the value of the NNPDF 3.0 strong coupling (α_s) at the Z boson mass is set to 0.130 (0.118) for MG5_aMC at LO and is set to 0.118 (0.118) for MG5_aMC at NLO for the 2016 (2017–2018) data-taking period. All three signal samples use the five-flavor number scheme with b quark density allowed in the initial state via a b and charm (c) quark PDF, with the b and c quarks typically being massless.

The DY process includes signal events with a Z boson and a jet initiated by a b quark (b jet) as well as background events from $Z + c$ jet and $Z + \text{light jet}$ in which a jet is initiated by a charm quark (c jet) and a light quark or gluon (light jet), respectively. All DY processes are scaled to NNLO accuracy in QCD based on the $Z + \text{jets}$ cross sections calculated with FEWZ v3.1 [27]. Other background contributions come from top quark-antiquark ($t\bar{t}$) and single top quark (s - and t -channel, and tW) production as well as from diboson (WW , WZ , and ZZ) processes. The $t\bar{t}$ process is generated at NLO by POWHEG v2.0 [28–30] in 2016 and MG5_aMC in 2017–2018 data-taking periods and interfaced with PYTHIA v8.212 (CUETP8M2T4 [31]), v8.226 (CP5), and v8.230 (CP5) for 2016, 2017, and 2018, respectively, for parton showering and hadronization. The single top quark t -channel and tW processes are simulated at NLO with POWHEG v2.0 and POWHEG v1.0, respectively, whereas the s -channel process is simulated with MG5_aMC at NLO, both interfaced with PYTHIA. The PYTHIA v8.212, v8.212, and v8.205 are used for single top quark t -channel, tW , and s -channel processes, respectively, in 2016, while PYTHIA v8.226 and v8.230 are used in 2017 and 2018, respectively, for all single top quark process. The diboson background processes are simulated at LO using PYTHIA v8.205, v8.226, and v8.230 for 2016, 2017, and 2018, respectively.

All simulation samples, except those based on SHERPA, use PYTHIA with the CUETP8M1 or CUETP8M2T4 or CP5 tune. The CUETP8M1 tune includes the LO NNPDF 2.3 [32] PDFs set with the strong coupling $\alpha_s(m_Z)$ set to 0.137 for space- and time-like shower simulation. The CUETP8M2T4 tune is based on the CUETP8M1 tune, which includes the NNPDF30_lo_as_0130 PDF set, but uses a lower value of $\alpha_s = 0.111$ for the initial-state radiation component of the parton shower. The CP5 tune use the NNPDF3.1 PDF set at NNLO, with α_s values of 0.118, and running according to NLO evolution. The diboson and single top quark predictions are scaled to NNLO accuracy [33,34]. The prediction for $t\bar{t}$ production is normalized to NNLO calculations in pQCD including

resummation of next-to-next-to-leading logarithmic soft-gluon terms with TOP++ 2.0 [35–41].

The MC samples are processed with the full CMS detector simulation based on GEANT4 [42]. The simulated events are reconstructed using the same algorithms used for the data, and include additional interactions per bunch crossings, referred to as pileup (PU). Simulated events are weighted so that the PU distribution reproduces the one observed in data, which has an average of about 23 (32) interactions per bunch crossing in 2016 (2017–2018).

IV. EVENT SELECTION

The final state in this analysis contains a Z boson ($Z \rightarrow ee/\mu\mu$) and at least one jet coming from a b quark. Therefore, the measurement requires the reconstruction of electrons, muons, and b jets. Events are reconstructed using a particle-flow (PF) algorithm [43], which identifies each individual particle candidate using information from various subdetectors. Energy deposits are measured in the calorimeters, and charged particles are identified in the central tracking and muon systems.

The collision events are required to pass single lepton triggers. They must have at least one electron (muon) candidate with a minimum p_T of 27, 32, and 32 GeV (24, 27, and 24 GeV) during the 2016, 2017, and 2018 data-taking periods, respectively.

The candidate vertex with the largest value of summed physics-object p_T^2 is the primary pp interaction vertex. The physics objects are the leptons and jets, clustered using the anti- k_T algorithm [44] with the tracks assigned to candidate vertices as inputs, and the associated missing transverse momentum ($\vec{p}_T^{\text{miss}}$), the negative vector p_T sum of all tracks. The leptons in the analysis are required to originate from the primary vertex.

Electrons are identified as a primary charged particle track and potentially many ECAL energy clusters. These clusters correspond to the electron track extrapolation to the ECAL and to possible bremsstrahlung photons emitted by electrons along the way through the tracker material [45]. The electron momentum is estimated by combining the energy measurement in the ECAL with the momentum measurement in the tracker. The longitudinal (transverse) impact parameters with respect to the beam line for the barrel ($|\eta| < 1.48$) and end cap ($1.48 < |\eta| < 3.0$) regions are required to be less than 0.10 (0.05) cm and 0.20 (0.10) cm, respectively. The momentum resolution for electrons with $p_T \approx 45$ GeV ranges from 1.7 to 4.5% depending on η . It is generally better in the barrel region than in the end caps, and also depends on the bremsstrahlung energy emitted by the electron as it traverses the material in front of the ECAL. The dielectron mass resolution for $Z \rightarrow ee$ decays when both electrons are in the ECAL barrel is 1.9%, and degrades to 2.9% when both electrons are in the end caps.

Muons are measured in the pseudorapidity range $|\eta| < 2.4$. The single muon trigger efficiency exceeds 90% over the full η range, and the efficiency to reconstruct and identify muons is greater than 96% for $p_T > 25$ GeV. Efficiencies are measured using data-driven tag-and-prove methods. Matching muons to tracks measured in the silicon tracker results in a transverse momentum resolution for muons with p_T up to 100 GeV of 1% in the barrel and 3% in the end caps [46]. The corrections for the muon transverse momentum scale and resolution [47] are also applied to remove biases of the muon momentum from detector misalignments or magnetic field uncertainties.

To reduce the misidentification rate, leptons are required to be isolated. The relative isolation variable is defined as:

$$I_{\text{rel}} = \frac{\sum p_{Ti}^{\text{charged}} + \max(0, \sum E_{Ti}^{\text{neutral}} + \sum E_{Ti}^{\gamma} - p_T^{\text{PU}})}{p_T^{\ell}}, \quad (1)$$

which is the sum of the p_T values for all PF candidates within a cone of radius $\Delta R = 0.3$ (0.4) centered on the electron (muon) track direction. Here, $\sum E_{Ti}^{\text{neutral}}$ and $\sum E_{Ti}^{\gamma}$ are the scalar sums of the transverse energies of neutral hadrons and photons, respectively. The quantity $\sum p_{Ti}^{\text{charged}}$ represents the p_T sum of the charged hadrons associated with the primary vertex and p_T^{PU} is the contribution from PU. For electrons, p_T^{PU} is evaluated using the “jet area” method described in [48]. For muons, p_T^{PU} is taken to be half of the scalar p_T sum deposited in the isolation cone by charged particles not associated with the primary vertex. The factor of one half corresponds approximately to the ratio of neutral to charged hadrons produced in the hadronization of PU interactions. Finally, p_T^{ℓ} stands for the lepton (electron or muon) transverse momentum. For muons I_{rel} is required to be less than 0.15. For electrons I_{rel} is p_T dependent; for 25 GeV $I_{\text{rel}} < 0.021$ (0.040) and for 35 GeV $I_{\text{rel}} < 0.015$ (0.029) in the barrel (end cap) region [49].

The small residual differences in the lepton (electron and muon) trigger, identification, and isolation efficiencies between data and simulation are measured using a “tag-and-probe” method [50] and are included by applying scale factors to simulated events.

Jets are clustered from PF candidates using the infrared- and collinear-safe anti- k_T algorithm with a distance parameter of 0.4, as implemented in the *FastJet* package [51]. The jet momentum is the vector sum of all particle momenta in the jet, and is found from simulation to be within 5%–10% of the true momentum over the entire p_T spectrum and detector acceptance. Pileup interactions could result in additional tracks and calorimetric energy depositions, possibly increasing the measured jet momentum. To mitigate this effect, tracks identified as originating from PU vertices are discarded and an offset correction [48] is applied to correct for remaining contributions.

The reconstructed jet energy scale (JES) is corrected using a factorized model to compensate for the nonlinear and nonuniform response in the calorimeters. Corrections are derived from simulation to bring, on average, the measured response of jets to that of particle-level jets. In situ measurements of the momentum balance in dijet, multijet, photon + jet, and leptonically decaying Z + jet events are used to account for any residual differences between the JES in data and simulation [52]. The jet energy resolution (JER) in the simulation is smeared to reproduce what is observed in data. The JER amounts typically to 16% at 30 GeV and 8% at 100 GeV. Additional selection criteria are applied to remove jets potentially dominated by anomalous contributions from various subdetector components or reconstruction failures [53]. Jets identified as likely to originate from PU are also removed by using a pileup jet ID discriminator for jets with $30 < p_T \leq 50$ GeV [54]. For jets with $p_T > 50$ GeV, the pileup jet ID discriminator is not applied, since here the pileup event contribution is negligible [55].

The missing transverse momentum vector $\vec{p}_T^{\text{miss}}$ is computed as the negative vector sum of the transverse momenta of all the PF candidates in an event, and its magnitude is denoted as p_T^{miss} [56]. The $\vec{p}_T^{\text{miss}}$ is modified to include corrections to the energy scale of the reconstructed jets in the event.

A neural network algorithm (DeepCSV) [57] is used to discriminate b jets from c and light jets. In the DeepCSV algorithm, the significance of the secondary vertex displacement is combined with track-based information in a machine-learned multivariate analysis to increase the efficiency compared with cut-based algorithms.

The Z boson candidate is reconstructed from two same flavor leptons with opposite charge having invariant mass ($m_{\ell\ell}$) between 71–111 GeV. Here, leading and subleading leptons must have $p_T > 35$ GeV and $p_T > 25$ GeV, respectively. The Z boson event must contain at least one b jet with $p_T > 30$ GeV, $|\eta| < 2.4$, and have $p_T^{\text{miss}} < 50$ GeV. The b jets are selected with the tight operating point of the b tag discriminator. The overlap between the b jet and a lepton from the Z boson decay is removed by requiring a minimum distance of $\Delta R > 0.4$ between them.

In simulation, the classification of reconstructed Z + jet events into Z + b jet, Z + c jet, and Z + light jet categories is based on the flavors of reconstructed jets with $p_T > 30$ GeV and $|\eta| < 2.4$. They are classified as b or c jets if they are matched to the generated b or c hadrons. The matching is done using the ghost association procedure in which the modulus of the hadron four-momentum is set to a small number, retaining only the directional information [57]. In the case when both c and b hadrons are matched, the jet is considered as a b jet. Based on reconstructed jets with defined flavors, events are classified as Z + b jet events if they contain at least one b jet. Of the remaining events, those

that contain at least one c hadron are considered as $Z + c$ jet events and those that contain neither b nor c hadrons are classified as $Z + \text{light jet}$ events. In this analysis the tagging efficiency of $Z + b$ jet ($Z + c$ jet) events is estimated as the fraction of events that pass the b (c) hadron and b tagging (c tagging) requirement over the number of events that pass just the b (c) hadron requirement. In simulations for the tight operating point, the tagging efficiencies for b (c) quarks are 50%–60% (25%) and misidentification probabilities are about 0.1% (<1.5%) and 2%–5% ($\approx 24\%$) for a light (light) and c (b) jet.

Data from $t\bar{t}$, $W + \text{jets}$, and multijet events that are enriched with a given flavor of interest are used to measure the b , c , and light jet tagging efficiencies. These are compared with the efficiencies calculated from simulation. Small differences between data and simulation are corrected through scale factors applied to simulation.

The selected events with a Z boson and at least one b jet have $\approx 18\%$ background contamination, with the major contributions arising from the $Z + c$ jet ($\approx 7\%$), $t\bar{t}$ ($\approx 7\%$), and $Z + \text{light jet}$ ($\approx 4\%$) processes. In the $Z + \geq 2 b$ jets signal region a significant source of background originates from $t\bar{t}$ production, accounting for $\approx 30\%$ of the events in the signal region and $\approx 85\%$ of the background. To validate the simulation, control regions that are enriched with processes of interests are constructed.

For the $Z + \geq 1 b$ jet measurement, the $t\bar{t}$ control region is constructed by requiring an oppositely charged electron and muon pair with at least one b jet ($e\mu + \geq 1 b$ jet) with the same selection criteria discussed above. A sample containing about 92% of $t\bar{t}$ events is obtained for each year of data taking. The non- $t\bar{t}$ contributions to this region are DY, single top, $W + \text{jets}$, and diboson events. A scale factor, $(\text{data} - \text{non-}t\bar{t} \text{ background}) / (t\bar{t} \text{ MC})$, is extracted from this region. It is 0.98 for 2016 and 0.87 for 2017–2018 data-taking periods and is applied in the signal region to the $t\bar{t}$ simulation in the ee and $\mu\mu$ channel.

In the $Z + \geq 2 b$ jets signal region, the $t\bar{t}$ background is estimated by a data-driven method using an $e\mu + \geq 2 b$ jets sample where the purity of the $t\bar{t}$ events is $\approx 96\%$. A normalization factor is derived by comparing the numbers of $t\bar{t}$ events seen in the sidebands of the dilepton invariant mass distributions from $e\mu + \geq 2 b$ jets and signal ($Z + \geq 2 b$ jets) regions. This factor is applied to the number of $t\bar{t}(\rightarrow e\mu) + \geq 2 b$ jets events within the Z boson mass window to obtain the $t\bar{t}$ background contribution in the signal region.

The $Z + \text{light jet}$ control region (CR1) is constructed by requiring events to have inclusive jets without any condition on the tagger discriminator. A sample containing about 83%–85% of $Z + \text{light jet}$ events is obtained for each year of data taking. The non- $(Z + \text{light jet})$ backgrounds are $Z + c$ jet, $Z + b$ jet, diboson, $t\bar{t}$, single top, and $W + \text{jets}$ events. A scale factor, $(\text{data} - \text{non-}(Z + \text{light jet}) \text{ background}) / (Z + \text{light jet MC})$, is extracted from

this region and applied in the signal region to the $Z + \text{light jet MC}$. The scale factors 0.94, 0.98, and 0.90 are used for the 2016, 2017, and 2018 data-taking periods, respectively. As a cross-check, another $Z + \text{light jet}$ control region (CR2) is defined by requiring events to pass anti- b and anti- c tagging tight operating points of b and c quark discriminators. The scale factors in CR1 and CR2 are consistent within 1%–2%.

The $Z + c$ jet control region (CR3) is constructed by requiring c jet events with the tight operating points of the c tag discriminator. A sample containing about 41%–55% of $Z + c$ jet events is obtained for each year of data taking. Scale factors, $(\text{data} - \text{non-}(Z + c \text{ jet}) \text{ background}) / (Z + c \text{ jet MC})$, 0.75, 1.08, and 0.99 for 2016, 2017, and 2018 data-taking periods, are extracted from this region and are applied to the $Z + c$ jet MC sample in the signal region. The non- $Z + c$ jet contributions to this region are $Z + b$ jet, $Z + \text{light jet}$, $t\bar{t}$, single top, $W + \text{jets}$, and diboson events. As a cross-check, another $Z + c$ jet control region (CR4) is defined by requiring events to pass tight operating points of the c quark discriminator and inverted b quark discriminator. The scale factors in CR3 and CR4 are consistent within 1%–2%.

The collective contribution of single top, $W + \text{jets}$, and diboson processes is small (<2%). Thus, data-driven estimates are not considered for these processes and their contributions are determined from simulation.

V. UNFOLDING AND SYSTEMATIC UNCERTAINTIES

The reconstructed distributions are unfolded to generator-level quantities using the TUnfold package (v17.5) [58], which is based on a least squares fit, to correct for the detector resolution and the selection efficiencies. The simulated events are classified according to generator-level information. The generator-level jets are formed from stable particles ($c\tau > 1 \text{ cm}$), except neutrinos, using the same anti- k_T algorithm as for the reconstructed jets. Any overlap between generator-level jets and a pair of leptons from the Z boson decays is removed by requiring a minimum distance of 0.4 between them. If an event consists of at least one (two) generator-level jet(s) containing a b hadron, the event is classified as a $Z + \geq 1 b$ jet ($Z + \geq 2 b$ jets) event. The generator-level leptons are “dressed” by adding the momenta of all photons within $\Delta R \leq 0.1$ around the lepton direction to account for the final-state radiation effects. The distributions are unfolded to the generator-level in the fiducial region defined in Table I.

The reconstructed distributions are unfolded with a response matrix that describes the migration probability between the particle- and reconstruction-level quantities. The response matrix is constructed by spatially matching the reconstruction-level leptons and b jet objects to the corresponding generator-level objects within $\Delta R < 0.3$. The selected $Z + \geq 1 b$ jet and $Z + \geq 2 b$ jets samples at

TABLE I. Fiducial region definition at generator-level.

Object	Selection
Dressed leptons	$p_T(\text{leading}) > 35 \text{ GeV}$, $p_T(\text{subleading}) > 25 \text{ GeV}$, $ \eta < 2.4$
Z boson	$71 < m_{\ell\ell} < 111 \text{ GeV}$
Generator-level b jet	b hadron jet, $p_T > 30 \text{ GeV}$, $ \eta < 2.4$

reconstruction-level can have background events that do not match to generator-level objects. The background events are channel dependent and they are subtracted from reconstructed events for each year and channel separately. The background subtracted reconstruction-level distributions are unfolded. The unfolded distributions are corrected for those events that have generator-level objects in the fiducial volume, which do not match with reconstructed objects. Events generated with MG5_aMC (NLO) and CP5 tune are used to construct the response matrix. The MG5_aMC (NLO) with CUETP8M1 tune, MG5_aMC (LO) with CP5 and CUETP8M1 tunes, and SHERPA are used for comparison with unfolded data.

The effect of regularization, a procedure to control large fluctuations of unfolded distributions due to statistical uncertainties of observed quantities [59], is negligible for the selected observables. Therefore, all studies are performed without regularization. The unfolding procedure is validated with statistically independent tests in which MG5_aMC NLO distributions are unfolded with LO MG5_aMC and compared with generator-level NLO MG5_aMC predictions. There is good agreement between the unfolded and the generator-level predictions. The resolution of variables to be measured in the present analysis is not expected to differ for the electron and muon channels. This procedure is validated by unfolding generator-level MC distributions for the electron channel with the response matrix constructed for the muon channel. The background and acceptance distributions are taken from the electron channel only. The resulting unfolded distributions are compared with the unfolded results when the electron channel is used for the construction of the response matrix. A good agreement is observed between the two unfolded results, which establishes that the response matrices for the two channels are consistent with each other. Therefore, the electron and muon data are added and unfolded with the combined response matrix.

The data distributions for the $Z+ \geq 1 b \text{ jet}$ ($Z+ \geq 2 b \text{ jets}$) analysis are unfolded with a response matrix constructed using the $Z + b \text{ jet}$ events in the DY sample that is simulated with the MG5_aMC (NLO) generator.

A. Experimental uncertainties

There are a number of sources of uncertainty that have a significant impact on the cross section measurements. These include

- (i) *Integrated luminosity*: The integrated luminosities of the 2016–2018 data-taking periods have uncertainties in the 1.2–2.5% range. These uncertainties are partially correlated and correspond to a total uncertainty of 1.6% for the Run 2 (2016–2018) [60–62]. The luminosity uncertainty affects the integrated cross section but not the normalized distributions.
- (ii) *Jet energy scale and resolution*: The effect of JES and JER is evaluated by varying up and down the p_T values of jets with the corresponding uncertainty factors. The JES and JER uncertainties are also propagated to p_T^{miss} . The JES and JER scale factors are correlated with the JES and JER components of p_T^{miss} , so scale factors are varied up and down simultaneously. For the $Z+ \geq 1 b \text{ jet}$ ($Z+ \geq 2 b \text{ jets}$) analysis, the JES introduces a systematic uncertainty of $\approx 3\%$ ($\approx 6\%$) on the integrated cross section whereas the effect of JER is less than 1%. The JES has varying effects on the differential and normalized differential cross section distributions for the different observables of $Z+ \geq 1 b \text{ jet}$ ($Z+ \geq 2 b \text{ jets}$) events, with maximum effects of 9% and 7% (24% and 18%), respectively, in the low event count bins.
- (iii) *b tagging/mistagging*: The uncertainty due to b tagging and mistagging scale factors is evaluated by varying them up and down within their uncertainties. The scale factor for tagging b jets is correlated with the c jets scale factor, so for tagging b and mistagging c jets, scale factors are varied up and down simultaneously, while the scale factor for mistagging light jets is varied independently. The combined b tagging uncertainty is then calculated as the sum in quadrature of these variations. For the $Z+ \geq 1 b \text{ jet}$ ($Z+ \geq 2 b \text{ jets}$) analysis the uncertainties in the b tagging and mistagging scale factors affect the cross section measurements by 3.0% (5.8%). The differential and normalized differential cross section distributions are affected by 2%–8% and 0.5%–5.0%, respectively, for selected observables of $Z+ \geq 1 b \text{ jet}$ events. These uncertainties are less than 14.5 and 9.5% in the $Z+ \geq 2 b \text{ jets}$ differential and normalized differential cross section measurements, respectively.
- (iv) *Unclustered energy of p_T^{miss}* : The effect of the uncertainty in the remaining component of p_T^{miss} , that is unclustered energy, is estimated by varying its contributions within the uncertainties, which are treated as correlated across years and channels. For $Z+ \geq 1 b \text{ jet}$ ($Z+ \geq 2 b \text{ jets}$), this uncertainty source affects the integrated cross section by up to 2.8% (3.6%). The differential and normalized differential cross section distributions of different observables of $Z+ \geq 1 b \text{ jet}$ ($Z+ \geq 2 b \text{ jets}$) events are affected by up to 3.8% (8.4%) and 1.1% (7.1%), respectively.

- (v) *Lepton selection*: Electrons and muons are corrected for trigger, reconstruction, and identification efficiencies and the effects of these corrections are estimated by varying the corresponding scale factors within their uncertainties. For the $Z + \geq 1 b$ jet ($Z + \geq 2 b$ jets) measurement, the uncertainty in the electron selection efficiency results in an uncertainty in the integrated cross section of about 4.6% (4.3%) for the electron channel, but for the combined channel the effect is about 1.5% (1.4%). Normalized distributions are not biased due to electron selection efficiency because associated uncertainties cancel in the ratio. The integrated cross section has a systematic uncertainty due to muon selection of less than 1%.
 - (vi) *Pileup reweighting*: The PU distribution in the simulated samples is reweighted to match that of the data. The corresponding uncertainty is estimated by varying the total inelastic cross section by $\pm 4.6\%$ [63]. For $Z + \geq 1 b$ jet ($Z + \geq 2 b$ jets) the uncertainty in the PU simulation affects the integrated cross section by 1.7%–2.4% (2.1%–2.9%). For the $Z + \geq 1 b$ jet ($Z + \geq 2 b$ jets) differential and normalized differential cross section distributions the effects are less than 2.7% (6.8%) and 1.3% (4.3%), respectively.
 - (vii) *Background estimation*: In the $Z + \geq 1 b$ jet cross section measurements the $Z + c$ jet and $Z + \text{light jet}$ contributions come from simulation and are validated in control regions as discussed in Sec. IV. The difference between data and MC simulation is extracted from each control region and is applied as a scale factor for the $Z + c$ jet and $Z + \text{light jet}$ MC predictions in the signal region, respectively. These scale factors can have associated systematic uncertainties from experimental sources (JES, b tagging, PU, lepton selection), although these are already considered while evaluating systematic uncertainties in the signal region. For all of the control regions the data and MC simulation values are within 10%. Hence, an additional systematic uncertainty of 10% is assigned to the background modeling. For $Z + \geq 1 b$ jet, the effect of these uncertainties is $\approx 2.2\%$ on the integrated cross section, whereas the effect of uncertainties related to $Z + c$ jet and $Z + \text{light jet}$ backgrounds is about 1.5%. The corresponding effects on the differential and normalized differential cross section distributions range from 1.5% to 4.8% and 0.5% to 3.3%, respectively. In the $Z + \geq 2 b$ jets measurements the background uncertainty originates from the $t\bar{t}$, Drell–Yan, diboson, and single top estimations. The $t\bar{t}$ uncertainty is the dominant source, and is derived by comparing the normalization coefficients obtained using fits in the dilepton invariant mass and missing transverse momentum distributions. The background uncertainty in the $Z + \geq 2 b$ jets integrated cross section is 2.4%. In the $Z + \geq 2 b$ jets differential and normalized differential cross sections this uncertainty is less than 4.9% and 2.7%, respectively.
 - (viii) *Model dependence*: The unfolding model uncertainty is estimated by reweighting the signal MC (MG5_aMC at NLO) with a scale factor that is calculated by fitting the ratio of background-subtracted data to signal simulation of normalized distributions and using it as an alternative model for the unfolding. It affects the differential and normalized differential cross section distributions of different observables in the $Z + \geq 1 b$ jet analysis up to 1.5% and 1.2%, respectively.
 - (ix) *Pileup jet identification*: The PU jet identification criteria have different performance in the data and simulated events. This difference is corrected as a scale factor. For $Z + \geq 1 b$ jet and $Z + \geq 2 b$ jets measurements, the effect of uncertainties associated with the scale factor is less than 0.7% in the integrated cross section. The differential and normalized differential cross section distributions of different observables of $Z + \geq 1 b$ jet ($Z + \geq 2 b$ jets) events are affected up to 0.5% (1.2%).
 - (x) *L1 prefiring*: During the 2016–2017 data-taking period there was a timing issue in the $2.0 < |\eta| < 3.0$ region of the ECAL where L1 trigger primitives were mistakenly associated with the previous bunch crossing. This led to a self-vetoing of events with significant energy in the $2.0 < |\eta| < 3.0$ region and a loss of trigger efficiency. Correction factors are determined from data for the efficiency loss and applied to the simulation of the dielectron and dimuon channels. The uncertainty due to the prefiring scale factors is evaluated by varying them up and down within their uncertainties. For the $Z + \geq 1 b$ jet and $Z + \geq 2 b$ jets differential and normalized differential cross section distributions as well as on the integrated cross section the effects are less than 0.5%.
- In the cross section ratio measurement, the luminosity uncertainty is effectively canceled. All other experimental uncertainties except the background estimation are evaluated as follows. For an uncertainty of interest, the $Z + \geq 2 b$ jets and $Z + \geq 1 b$ jet cross sections are obtained for up and down systematic variations described above and the corresponding ratios are obtained. The uncertainty is derived by comparing these ratios with the nominal ones. The background estimations in the $Z + \geq 2 b$ jets and $Z + \geq 1 b$ jet events are based on independent methods and control regions. Therefore, the background uncertainties are uncorrelated and derived from quadrature sums of those from the $Z + \geq 2 b$ jets and $Z + \geq 1 b$ jet measurements.

B. Theoretical uncertainties

The uncertainties coming from the theory calculations and MC generators are estimated as follows.

- (i) *Renormalization and factorization scales*: The scale uncertainties are estimated using a set of weights provided by the generator that corresponds to variations of renormalization (μ_R) and factorization (μ_F) scales by factors of 0.5 and 2. The unfolded distributions are obtained for all combinations, except for $\mu_F/\mu_R = 0.25$ or 4, and their envelope is quoted as the uncertainty. The integrated cross sections for $Z+ \geq 1 b\text{jet}$ ($Z+ \geq 2 b\text{jets}$) have scale uncertainties of 2.6% (2.5%), 2.9% (2.3%), and 2.1% (2.5%) in the dielectron, dimuon, and combined channels, respectively. The uncertainties in the $Z+ \geq 1 b\text{jet}$ ($Z+ \geq 2 b\text{jets}$) differential distributions vary from 1% to 10% (0.5% to 4.9%) and for the normalized differential distributions they vary from 0.5% to 9.4% (0.5% to 5.0%). Similarly, uncertainties are evaluated for MG5_aMC (NLO) with CP5 tune generator-level distributions and the effect on the $Z+ \geq 1 b\text{jet}$ ($Z+ \geq 2 b\text{jets}$) integrated cross section at the generator-level is around 6.6% (9.0%). The renormalization and factorization scale uncertainties in the cross section ratios are the quadrature sum of those from the $Z+ \geq 2 b\text{jets}$ and $Z+ \geq 1 b\text{jet}$ measurements.
- (ii) *PDF*: The uncertainty in the parton distribution functions is estimated using different Hessian eigenvectors of the NNPDF 3.1 PDF sets. Applying the master formula of the Hessian PDF error calculation [64], this uncertainty is estimated on a bin-by-bin basis for the differential distributions. The integrated cross sections for the $Z+ \geq 1 b\text{jet}$ ($Z+ \geq 2 b\text{jets}$) process have PDF uncertainties around 0.3%–0.4% (0.3%) in the dielectron, dimuon, and combined channels. For the $Z+ \geq 1 b\text{jet}$ ($Z+ \geq 2 b\text{jets}$) analysis, the uncertainties in the differential distributions for the dielectron, dimuon, and combined channels are 0.5% to 4.2% (0.5%–4.0%), whereas the uncertainties in the normalized differential cross sections for these channels are in the range 0.5%–3.3% (0.5%–3.2%). The PDF uncertainties are treated as correlated across the years and channels. Similarly, the effect of PDF (NNPDF 3.1) uncertainty in the generator-level distributions is obtained on a bin-by-bin basis of generator-level distributions with PDF weights applied. The effect on the integrated cross sections of the $Z+ \geq 1 b\text{jet}$ and $Z+ \geq 2 b\text{jets}$ at the generator-level is around 1.0%. The $Z+ \geq 2 b\text{jets}$ and $Z+ \geq 1 b\text{jet}$ PDF uncertainties are summed in quadrature for the cross section ratios. The uncertainties in the PDFs are also estimated for SHERPA at generator-level using the 102 replicas of the NNPDF 3.0 PDF set.

The integrated cross sections for the $Z+ \geq 1 b\text{jet}$ ($Z+ \geq 2 b\text{jets}$) process have PDF uncertainties around 1.4% and 2.0%.

- (iii) *Strong coupling (α_s)*: The uncertainties due to α_s are estimated by using two additional PDF members corresponding to upper (0.120) and lower (0.116) values with respect to the central (0.118) α_s value. The integrated cross sections for the $Z+ \geq 1 b\text{jet}$ ($Z+ \geq 2 b\text{jets}$) process have α_s uncertainties around 0.2%–0.3% (0.1%) in the dielectron, dimuon, and combined channels. For the $Z+ \geq 1 b\text{jet}$ ($Z+ \geq 2 b\text{jets}$) analysis, the uncertainty in the differential distributions for the dielectron, dimuon, and combined channels is 0.5%–1.3% (0.5%–1.6%), and the uncertainties in the normalized differential cross sections for these channels are in the range 0.5 to 1.1% (0.5%–1.3%). The effect on $Z+ \geq 1 b\text{jet}$ ($Z+ \geq 2 b\text{jets}$) integrated cross section at the generator-level is 1.2% (0.1%). The $Z+ \geq 2 b\text{jets}$ and $Z+ \geq 1 b\text{jet}$ α_s uncertainties are summed in quadrature for the cross section ratios.

Table II lists the uncertainties in the $Z+ \geq 1 b\text{jet}$ and $Z+ \geq 2 b\text{jets}$ integrated cross sections. Tables III and IV summarize the range of uncertainty in the differential and normalized differential cross section distributions for the $Z+ \geq 1 b\text{jet}$ combined channel, respectively. For the $Z+ \geq 2 b\text{jets}$ measurements, Tables V and VI show the uncertainty in the differential and normalized differential cross sections, respectively. All systematic uncertainties are treated as 100% correlated across the years. In summary, the $Z+ \geq 1 b\text{jet}$ ($Z+ \geq 2 b\text{jets}$) integrated cross sections have total experimental systematic uncertainties of 5.9%

TABLE II. Summary of the uncertainties (in percent) in the integrated cross sections for the dielectron, dimuon, and combined channels in the $Z+ \geq 1 b\text{jet}$ and $Z+ \geq 2 b\text{jets}$ events.

Uncertainty (%)	$Z+ \geq 1 b\text{jet}$			$Z+ \geq 2 b\text{jets}$		
	ee	$\mu\mu$	$\ell\ell$	ee	$\mu\mu$	$\ell\ell$
Statistical	1.0	0.7	0.6	7.7	5.9	4.6
JES, JER	2.7	3.0	2.9	6.9	5.4	5.8
b tagging/mistagging	3.0	2.9	2.9	5.4	6.0	5.8
Unclustered energy of p_T^{miss}	2.8	2.8	2.8	3.5	3.7	3.6
Background estimation	2.2	2.0	2.1	2.3	2.4	2.4
Pileup reweighting	2.4	1.7	1.9	2.9	2.1	2.4
Electron selection	4.6	...	1.5	4.3	...	1.4
Luminosity	1.6	1.6	1.6	1.6	1.6	1.6
Muon selection	...	0.6	0.4	...	1.0	0.7
Pileup jet identification	0.3	0.3	0.3	0.6	0.7	0.7
L1 prefiltering	0.3	0.2	0.2	0.3	0.2	0.3
μ_R and μ_F scales	2.6	2.9	2.1	2.5	2.3	2.5
PDF	0.4	0.3	0.3	0.3	0.3	0.3
α_s	0.3	0.2	0.2	0.1	0.1	0.1
Total experimental	7.6	5.9	6.1	11.4	9.7	9.9
Total theoretical	2.6	2.9	2.1	2.5	2.3	2.5

TABLE III. Summary of the uncertainties (in percent) in the differential cross section distributions for the combined channel in the $Z + \geq 1 b$ jet events.

Observable/Uncertainty (%)	p_T^Z	$p_T^{b \text{ jet}}$	$ \eta^{b \text{ jet}} $	$\Delta\phi^{(Z, b \text{ jet})}$	$\Delta y^{(Z, b \text{ jet})}$	$\Delta R^{(Z, b \text{ jet})}$
Statistical	1.0–5.2	1.0–12	0.7–1.4	0.7–2.2	0.7–4.7	1.0–5.3
JES, JER	1.0–9.0	2.0–7.2	2.0–7.0	2.0–4.8	2.0–6.4	1.6–6.1
b tagging/mistagging	2.3–6.0	2.0–8.3	2.6–4.3	2.5–4.9	2.7–5.5	2.6–7.2
Unclustered energy of p_T^{miss}	2.2–3.6	2.2–3.4	0.5–1.8	2.5–3.8	2.7–3.0	2.3–3.8
Background estimation	1.6–3.6	1.8–3.0	1.9–3.2	1.5–3.9	2.0–3.4	2.6–4.8
Pileup reweighting	1.0–2.2	0.7–2.2	1.8–2.4	1.8–2.6	1.8–2.2	1.5–2.5
Model dependency	0.5–1.4	0.5–1.2	0.5	0.5–1.5	0.5–0.6	0.5–1.1
Electron selection	1.3–2.8	1.2–1.9	1.4–1.8	1.3–1.9	1.5–2.4	1.2–2.2
Muon selection	0.5–1.0	0.5–1.0	0.5	0.5	0.5	0.5–0.6
Pileup jet identification	0.5–0.6	0.5	0.5–0.9	0.5	0.5–0.6	0.5
L1 prefire	0.5	0.5	0.5–0.7	0.5	0.5–0.9	0.5
μ_R and μ_F scales	1.0–7.2	1.0–10.0	0.7–3.3	1.5–2.7	1.0–2.8	1.5–4.0
PDF	0.5–2.9	0.5–2.1	1.5–2.8	0.5–2.1	0.5–1.6	0.5–4.2
α_S	0.5–1.0	0.5–0.7	0.5	0.5–1.1	0.5	0.5–1.3

TABLE IV. Summary of the uncertainties (in percent) in the normalized differential distributions for the combined channel in the $Z + \geq 1 b$ jet events.

Observable/Uncertainty (%)	p_T^Z	$p_T^{b \text{ jet}}$	$ \eta^{b \text{ jet}} $	$\Delta\phi^{(Z, b \text{ jet})}$	$\Delta y^{(Z, b \text{ jet})}$	$\Delta R^{(Z, b \text{ jet})}$
Statistical	1.0–5.2	1.0–12	0.7–1.4	1.0–2.1	0.7–4.7	0.8–3.3
JES, JER	1.0–5.8	1.0–6.7	0.5–4.5	0.5–2.1	0.5–4.2	0.5–4.2
b tagging/mistagging	0.5–3.9	0.5–5.6	0.5–1.7	0.5–2.9	0.5–3.7	0.5–1.7
Unclustered energy of p_T^{miss}	0.5–0.7	0.5–0.7	0.5–1.1	0.5–1.0	0.5	0.5–1.0
Background estimation	0.5–1.9	0.6–1.7	0.5–1.4	0.5–2.0	0.5–2.4	0.5–3.3
Pileup reweighting	0.5–0.9	0.5–1.3	0.5	0.5–0.7	0.5	0.5–0.8
Model dependency	0.5–1.4	0.5–0.8	0.5	0.5–1.2	0.5	0.5–0.9
Electron selection	0.5–1.3	0.5	0.5	0.5	0.5–0.9	0.5–0.6
Muon selection	0.5–1.0	0.5	0.5	0.5	0.5	0.5
Pileup jet identification	0.5	0.5	0.5	0.5	0.5	0.5
L1 prefire	0.5	0.5	0.5	0.5	0.5	0.5
μ_R and μ_F scales	1.0–7.0	0.5–9.4	0.5–2.0	1.0–2.1	0.5–2.1	0.5–3.7
PDF	0.5–2.0	0.5–1.6	0.5–0.7	0.5–1.8	0.5–1.6	0.5–3.3
α_S	0.5–1.0	0.5–0.7	0.5	0.5–0.9	0.5	0.5–1.1

TABLE V. Summary of the uncertainties (in percent) in the differential cross section distributions for the combined channel in the $Z + \geq 2 b$ jets events. The symbols, $b1$ and $b2$, stand for leading and subleading b jets, respectively.

Observable/Uncertainty (%)	p_T^{b1}	p_T^{b2}	$ \eta ^{b1}$	p_T^Z	ΔR_{bb}	$\Delta R_{Zbb}^{\text{min}}$	A_{Zbb}	m_{Zbb}	m_{bb}
Statistical	8.6–15.5	7.0–37.6	6.2–11.9	6.2–25.3	5.6–29.2	5.3–43.5	4.5–25.0	8.1–22.1	11.3–25.9
JES, JER	6.0–12.8	5.1–14.3	6.8–8.2	5.6–14.8	5.9–14.9	4.5–23.8	4.3–13.2	4.8–13.7	4.2–17.5
b tagging/mistagging	4.4–7.7	4.3–9.2	4.3–6.2	4.1–13.7	3.5–6.4	3.8–14.5	5.3–7.3	4.2–7.2	4.5–6.6
Unclustered energy of p_T^{miss}	1.3–6.8	2.1–8.0	2.7–4.2	3.1–6.8	1.1–5.0	0.7–4.8	2.1–4.8	2.1–5.1	1.3–8.4
Background simulation	0.6–3.7	0.8–4.9	1.0–2.9	0.6–3.8	0.5–3.0	0.5–3.8	1.9–3.1	0.6–4.1	1.1–3.5
Pileup reweighting	1.4–4.1	1.6–6.8	1.7–2.7	1.6–3.3	1.9–3.1	0.9–3.9	1.7–4.3	1.7–3.2	1.5–3.5
Electron selection	0.9–1.7	1.1–2.0	1.2–1.5	0.9–1.7	0.7–1.5	0.5–1.8	1.1–1.6	1.0–1.8	0.9–1.6
Muon selection	0.5–0.8	0.5–0.9	0.7–3.3	0.4–0.9	0.5–0.8	0.5–0.8	0.6–0.8	0.5–0.9	0.6–0.8
Pileup jet identification	0.6–1.2	0.5–1.0	0.8–0.9	0.7–0.9	0.8–1.1	0.8–0.9	0.7–0.8	0.7–1.1	0.6–1.0
μ_R and μ_F scales	0.5–3.2	0.7–4.6	1.8–4.7	0.7–4.5	0.6–4.1	0.7–4.7	1.3–4.9	1.7–3.8	1.1–4.8
PDF	0.5–0.8	0.5–1.4	0.5–0.7	0.5–1.1	0.5–2.5	0.5–4.0	0.5–0.9	0.5–0.8	0.5–0.9
α_S	0.5	0.5	0.5	0.5	0.5–1.6	0.5–0.9	0.5	0.5	0.5

TABLE VI. Summary of the uncertainties (in percent) in the normalized differential distributions for the combined channel in the $Z+ \geq 2 b$ jets events. The symbols, $b1$ and $b2$, stand for leading and subleading b jets, respectively.

Observable/Uncertainty (%)	p_T^{b1}	p_T^{b2}	$ \eta ^{b1}$	p_T^Z	ΔR_{bb}	$\Delta R_{Zbb}^{\min}$	A_{Zbb}	m_{Zbb}	m_{bb}
Statistical	9.7–16.2	8.3–37.9	7.7–12.8	7.7–25.7	7.2–29.6	7.0–43.7	6.4–25.4	9.3–22.6	12.2–26.3
JES, JER	3.2–7.7	1.6–9.4	0.8–1.3	0.9–7.9	0.8–7.9	1.0–17.5	0.9–6.1	2.2–8.1	3.5–12.5
b tagging/mistagging	0.5–2.2	0.7–3.8	0.6–1.7	0.5–8.4	0.5–2.4	0.5–9.5	0.5–1.6	0.5–1.7	0.5–1.3
Unclustered energy of p_T^{miss}	0.5–3.1	0.8–7.1	0.5–0.9	0.5–3.0	0.5–2.5	0.5–6.1	0.5–3.9	0.7–3.1	0.5–4.6
Background simulation	0.5–1.8	0.5–2.7	0.5–1.4	0.5–1.9	0.5–2.1	0.5–2.3	0.5–0.7	0.5–1.8	0.5–1.3
Pileup reweighting	0.5–1.8	0.5–4.3	0.5–0.8	0.5–2.7	0.5–0.8	0.5–1.7	0.5–1.7	0.5–1.1	0.5–1.3
Electron selection	0.5	0.5–0.7	0.5	0.5	0.5–0.6	0.5–1.0	0.5	0.5	0.5
Muon selection	0.5	0.5	0.9–2.3	0.5	0.5	0.5	0.5	0.5	0.5
Pileup jet identification	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
μ_R and μ_F scales	0.5–2.9	0.6–4.6	1.4–3.9	1.4–3.8	1.3–5.0	1.1–4.8	1.7–4.6	2.0–4.7	1.3–4.9
PDF	0.5–0.6	0.7–1.3	0.5–0.6	0.5–1.0	0.5–2.1	2.6–3.2	0.5–0.9	0.5–0.8	0.5–0.8
α_S	0.5	0.5	0.5	0.5	0.5–1.3	0.5–0.6	0.5	0.5	0.5

and 6.1% (9.7% and 9.9%), respectively, in the dimuon and combined channels. For the electron channel the uncertainty is a bit higher, up to 7.6% ($Z+ \geq 1 b$ jet) and 11.4% ($Z+ \geq 2 b$ jets), owing to the electron selection uncertainty. For the $Z+ \geq 1 b$ jet ($Z+ \geq 2 b$ jets) measured integrated cross section the statistical uncertainty is less than 1% (5%).

VI. RESULTS

The measured $Z+ \geq 1 b$ jet and $Z+ \geq 2 b$ jets integrated cross sections in the fiducial region described in Table I, are $6.52 \pm 0.04(\text{stat}) \pm 0.40(\text{syst}) \pm 0.14(\text{theo})$ pb and $0.65 \pm 0.03(\text{stat}) \pm 0.07(\text{syst}) \pm 0.02(\text{theo})$ pb in the combined channel, respectively. The measurements are consistent for the dielectron, dimuon, and combined channels as reported in Table VII, therefore the differential and normalized differential distributions are shown only for the combined channel. The measured cross sections are compared with predictions from MG5_aMC obtained with two settings

corresponding to 2016 and 2017–2018 data-taking periods. The 2017–2018 settings use more up-to-date MG5_aMC versions (2.6.0 and 2.4.2 for NLO and LO, respectively), PDFs (NNPDF 3.1), and underlying event tune (CP5), and are labeled with NNPDF 3.1, CP5 in the following. The 2016 data-taking period uses earlier versions of the MG5_aMC (2.3.2.2 and 2.2.2 for NLO and LO, respectively), PDFs (NNPDF 3.0), and underlying event tune (CUETP8M1), and are labeled with NNPDF 3.0, CUETP8M1. In the $Z+ \geq 1 b$ jet final state, the measured integrated cross section values are well described by MG5_aMC LO for both settings, whereas MG5_aMC (NLO, NNPDF 3.1, CP5) and MG5_aMC (NLO, NNPDF 3.0, CUETP8M1) overestimate data by ≈ 10 and $\approx 18\%$, respectively. The SHERPA simulation overestimates the measured cross section values by $\approx 24\%$. In the $Z+ \geq 2 b$ jets final state, the measured integrated cross section is in good agreement with MG5_aMC (LO). The MG5_aMC (NLO, NNPDF 3.1, CP5), MG5_aMC (NLO, NNPDF 3.0, CUETP8M1), and SHERPA predictions overestimate the

TABLE VII. Measured and predicted cross sections (in pb) for the $Z+ \geq 1 b$ jet and $Z+ \geq 2 b$ jets final states. The cross section ratios between the $Z+ \geq 2 b$ jets and $Z+ \geq 1 b$ jet are shown in the last three rows for the dielectron, dimuon, and combined channels. In the measured results the first, second, and third uncertainties correspond to the statistical, systematic, and theoretical sources, respectively. The MG5_aMC (NLO) predictions include theoretical uncertainties (PDF, and renormalization and factorization scales).

			MG5_aMC LO NNPDF 3.0	MG5_aMC LO NNPDF 3.1	MG5_aMC NLO NNPDF 3.0	MG5_aMC NLO NNPDF 3.1	
Channel	Measured		CUETP8M1	CP5	CUETP8M1	CP5	SHERPA
$Z+ \geq 1\ b\text{ jet}$	ee	$6.45 \pm 0.06 \pm 0.49 \pm 0.17$	6.25	6.33	7.86 ± 0.52	7.05 ± 0.48	8.05
	$\mu\mu$	$6.55 \pm 0.05 \pm 0.39 \pm 0.19$	6.26	6.34	7.86 ± 0.51	7.02 ± 0.47	7.98
	$\ell\ell$	$6.52 \pm 0.04 \pm 0.40 \pm 0.14$	6.25	6.34	7.86 ± 0.51	7.03 ± 0.47	8.02
$Z+ \geq 2\ b\text{ jets}$	ee	$0.66 \pm 0.05 \pm 0.07 \pm 0.02$	0.62	0.72	0.89 ± 0.08	0.77 ± 0.07	0.84
	$\mu\mu$	$0.65 \pm 0.04 \pm 0.06 \pm 0.02$	0.64	0.71	0.91 ± 0.09	0.77 ± 0.07	0.84
	$\ell\ell$	$0.65 \pm 0.03 \pm 0.07 \pm 0.02$	0.63	0.71	0.90 ± 0.09	0.77 ± 0.07	0.84
Ratio	ee	$0.102 \pm 0.008 \pm 0.008 \pm 0.004$	0.100	0.113	0.113 ± 0.016	0.110 ± 0.013	0.104
	$\mu\mu$	$0.100 \pm 0.006 \pm 0.006 \pm 0.004$	0.103	0.112	0.116 ± 0.016	0.110 ± 0.013	0.105
	$\ell\ell$	$0.100 \pm 0.005 \pm 0.007 \pm 0.003$	0.102	0.112	0.114 ± 0.016	0.110 ± 0.013	0.105

measured cross section values by 29, 38, and 21%, respectively. The measured value of the cross section ratio of the $Z + \geq 2 b$ jets and $Z + \geq 1 b$ jet for the combined channel is $0.100 \pm 0.005(\text{stat}) \pm 0.007(\text{syst}) \pm 0.003(\text{theo})$ and, as expected, has smaller uncertainties. This measurement is in agreement with the MG5_aMC (LO, NNPDF 3.0, CUETP8M1) and SHERPA values given the large uncertainties associated with the predictions. Table VII summarizes the measured and predicted cross sections for the $Z + \geq 1 b$ jet and $Z + \geq 2 b$ jets processes.

The measured differential cross section distributions and the corresponding ones normalized by the integrated fiducial cross sections compared with different MC predictions in the combined channel are shown in Figs. 2–16. In general, the differential cross section distributions predicted by SHERPA tend to overestimate data by 20%–30%, depending on the kinematic variables.

The differential and normalized differential cross section distributions as functions of the p_T^Z in the selected $Z + \geq 1 b$ jet events are shown in Fig. 2. The shapes of these distributions are described best by MG5_aMC (LO, NNPDF 3.1, CP5), while MG5_aMC (NLO) and SHERPA predictions vary up to 30% depending on p_T^Z . The $p_T^{b\text{jet}}$ and $|\eta^{b\text{jet}}|$ differential and normalized differential cross section

distributions are shown in Figs. 3 and 4, respectively. The shapes of distributions are well described by all simulations, except for the MG5_aMC (LO, NNPDF 3.1, CP5) $p_T^{b\text{jet}}$ spectrum, which deviates up to 25% in the higher p_T region. The differential and normalized differential cross section distributions of $\Delta\phi^{(Z,b\text{jet})}$, $\Delta y^{(Z,b\text{jet})}$, and $\Delta R^{(Z,b\text{jet})}$ are shown in Figs. 5–7. The shapes of these distributions are best described by the SHERPA predictions. The MG5_aMC (LO) predictions show the largest deviation from data in the high $\Delta y^{(Z,b\text{jet})}$ and $\Delta R^{(Z,b\text{jet})}$ regions, which is significantly improved with the MG5_aMC (NLO) prediction. In summary, SHERPA simulations provide a good description of different kinematic observables except for the higher p_T^Z region for the normalized differential distributions. The MG5_aMC LO and NLO simulations provide varying levels of agreement for the observables.

For the $Z + \geq 2 b$ jets final states, differential and the normalized differential cross sections as functions of the leading b jet p_T and $|\eta|$, subleading b jet p_T , and Z boson transverse momentum are shown in Figs. 8–11. The shapes of distributions in data are well-described by predictions from MG5_aMC (LO and NLO; with both PDFs) and SHERPA, except for a couple of bins in the high- p_T and high- $|\eta|$ regions. In the differential cross sections MG5_aMC

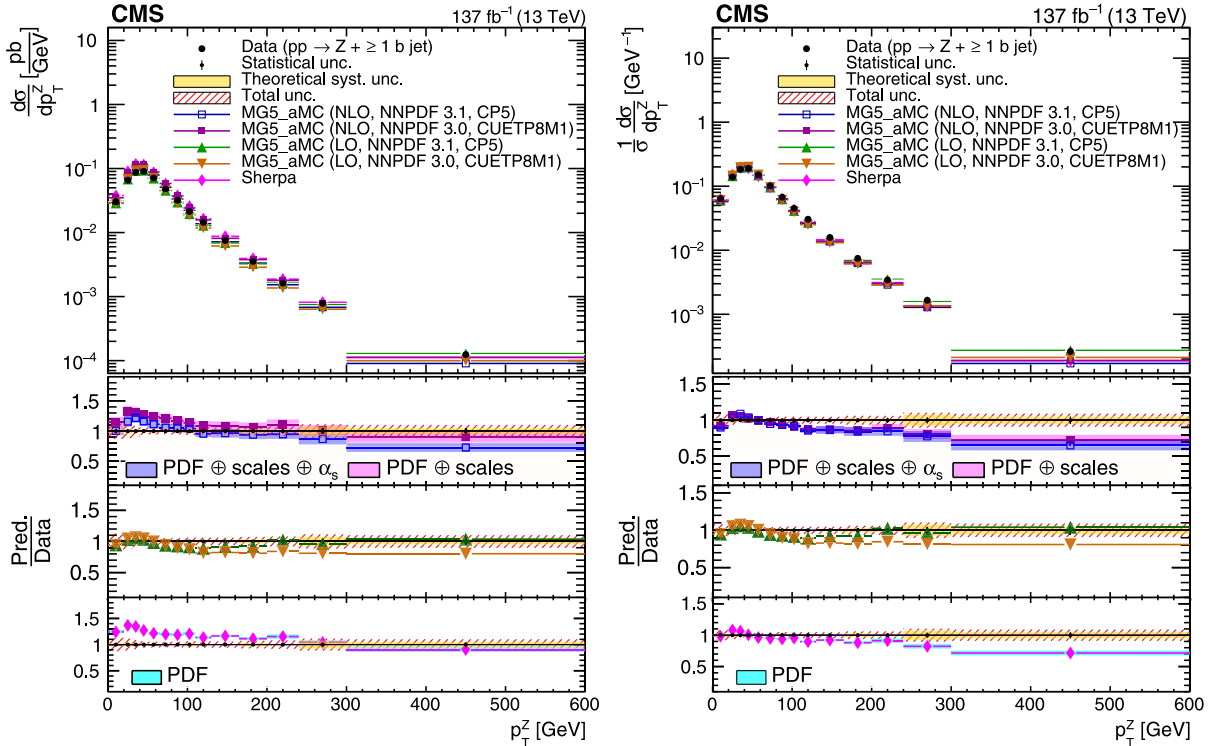


FIG. 2. (left) Differential cross section and the (right) normalized differential cross section distributions as a function of Z transverse momenta for $Z + \geq 1 b$ jet events. The uncertainties in the predictions are shown as colored bands in the bottom panel around the theoretical predictions including statistical, PDF, scale, and α_s uncertainties for the MG5_aMC (NLO, NNPDF 3.1, CP5), MG5_aMC (NLO, NNPDF 3.0, CUETP8M1) and the statistical and PDF uncertainties for the SHERPA predictions. The statistical, theoretical, and total uncertainties in data are indicated by the vertical bars, yellow, and hatched bands centered at 1, respectively. The same description applies to all of the remaining figures.

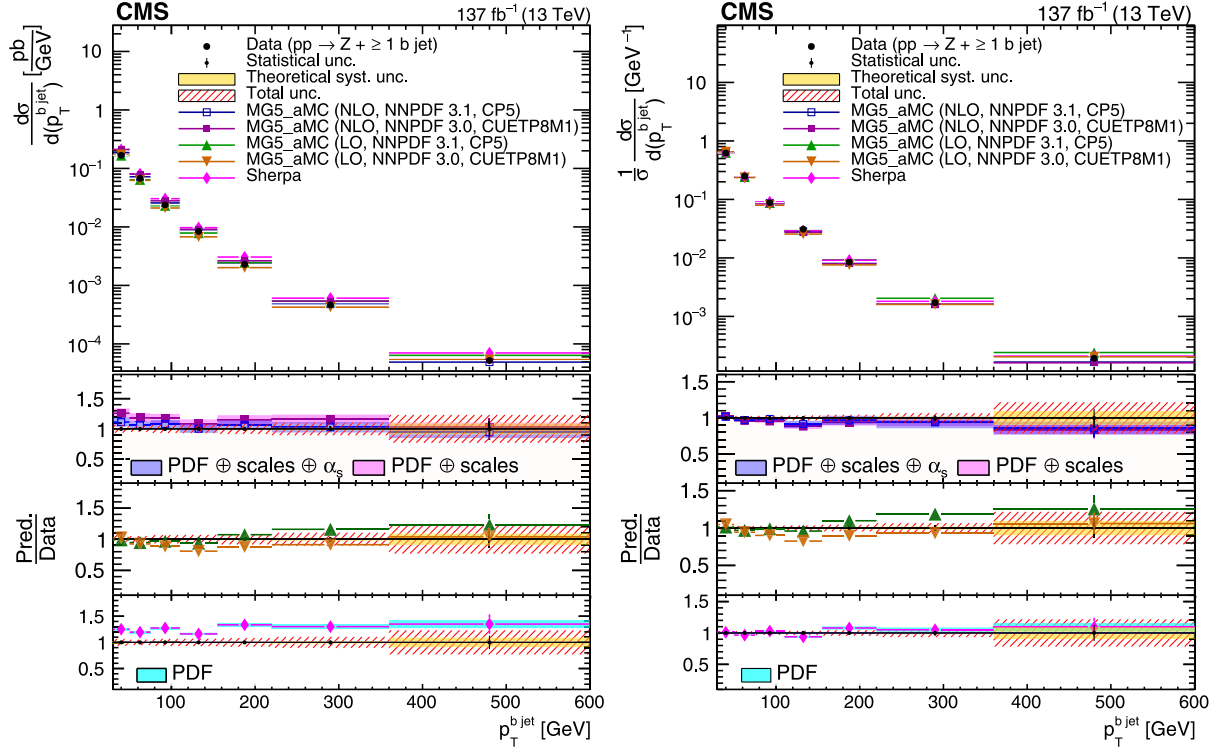


FIG. 3. (left) Differential cross section and the (right) normalized differential cross section distributions as a function of b jet transverse momenta for $Z + \geq 1 \text{ b jet}$ events.

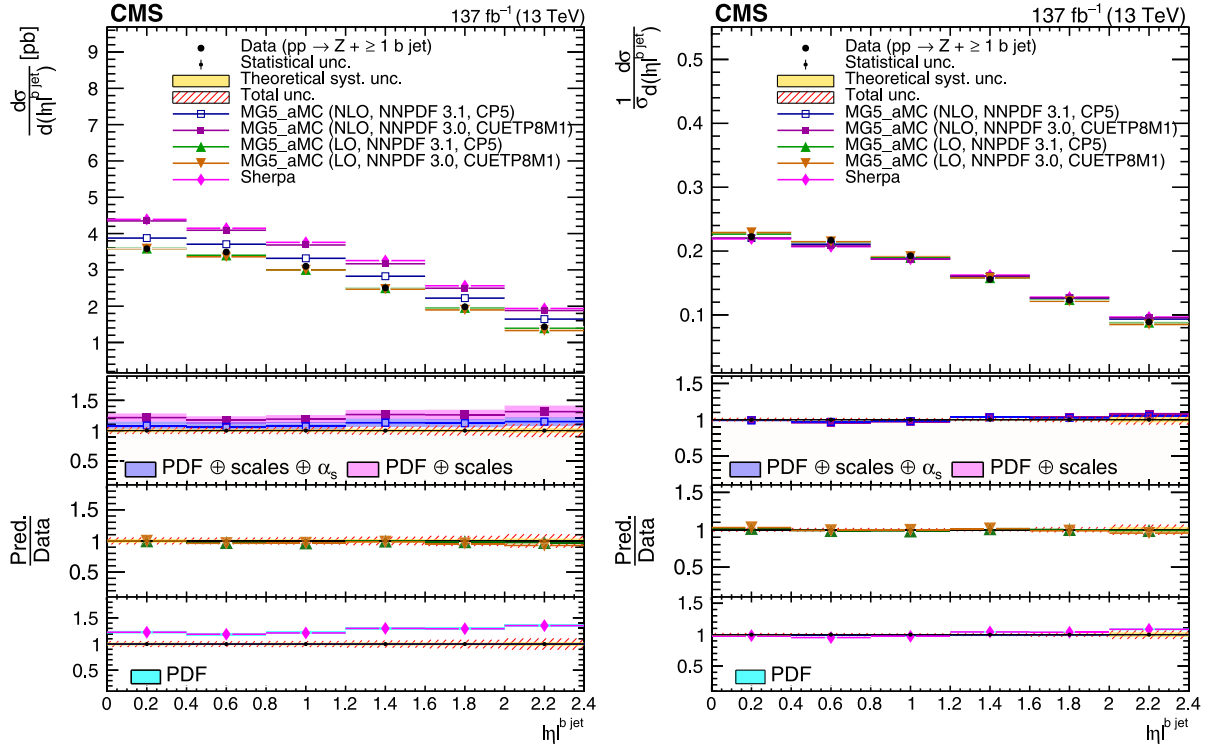


FIG. 4. (left) Differential cross section and the (right) normalized differential cross section distributions as functions of b jet absolute pseudorapidity for $Z + \geq 1 \text{ b jet}$ events.

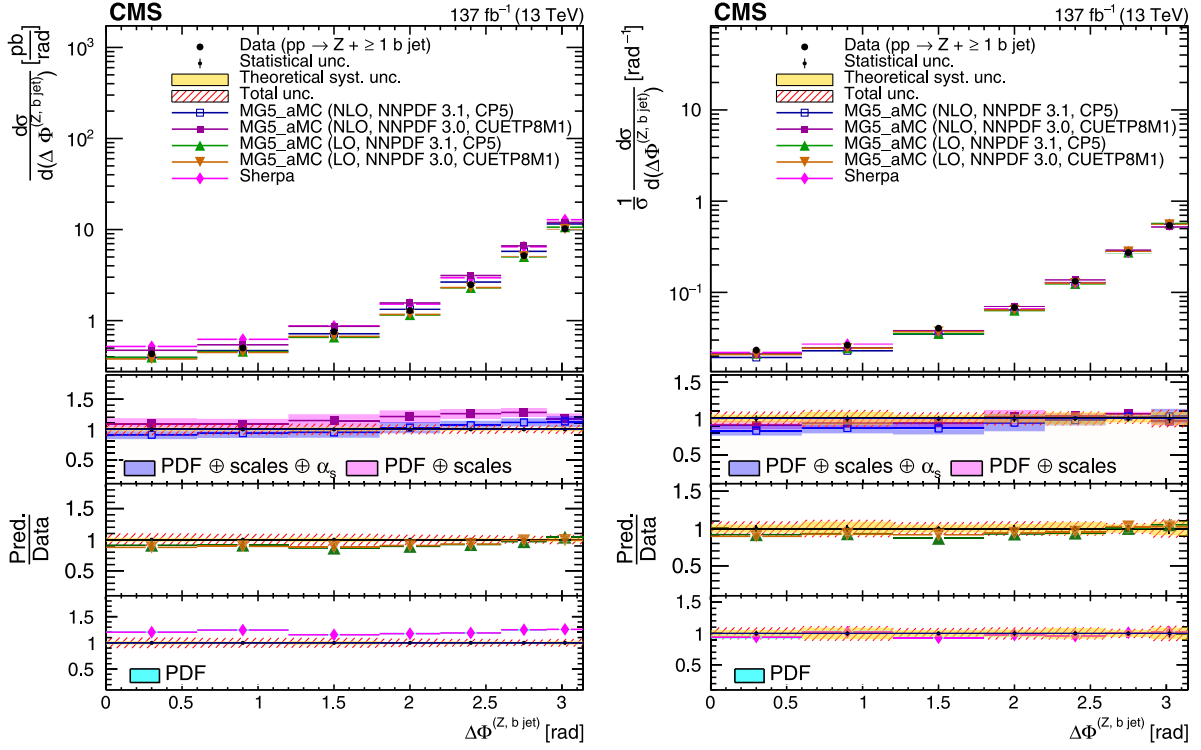


FIG. 5. (left) Differential cross section and the (right) normalized differential cross section distributions as functions of $\Delta\phi^{(Z, b \text{ jet})}$ between the Z boson and the leading b jet for $Z + \geq 1$ b jet events.

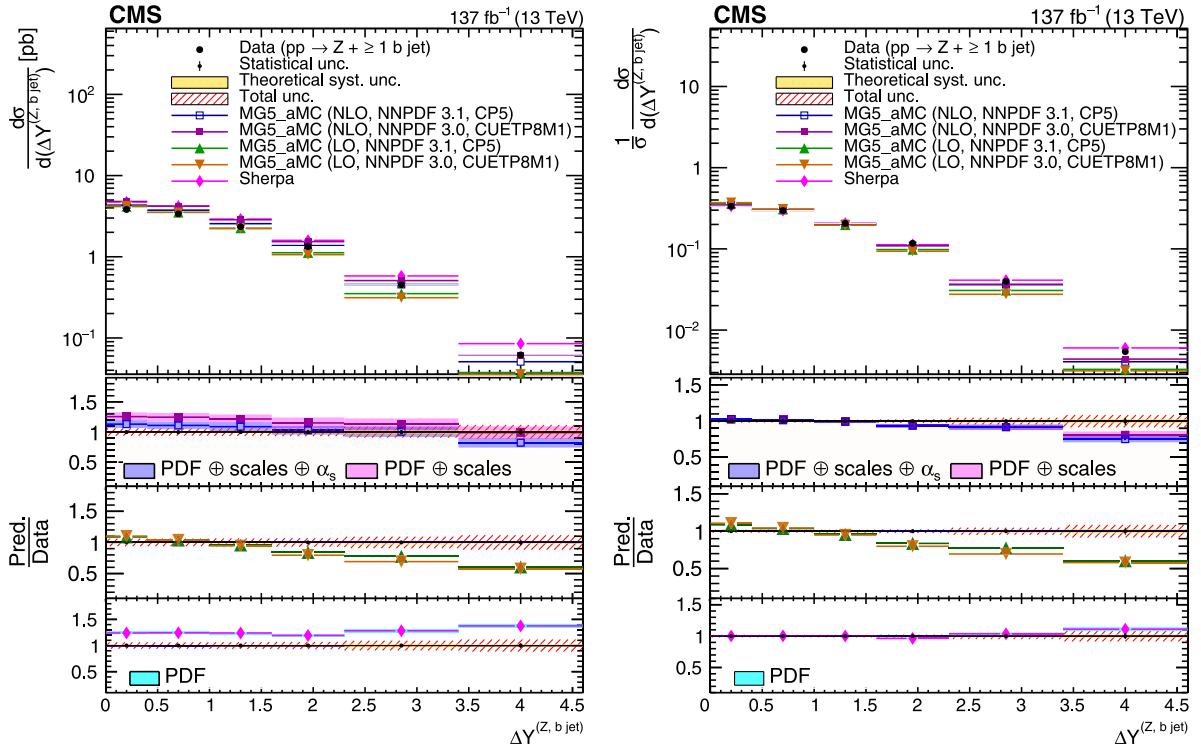


FIG. 6. (left) Differential cross section and the (right) normalized differential cross section distributions as functions of $\Delta Y^{(Z, b \text{ jet})}$ between the Z boson and the leading b jet for the $Z + \geq 1$ b jet events.

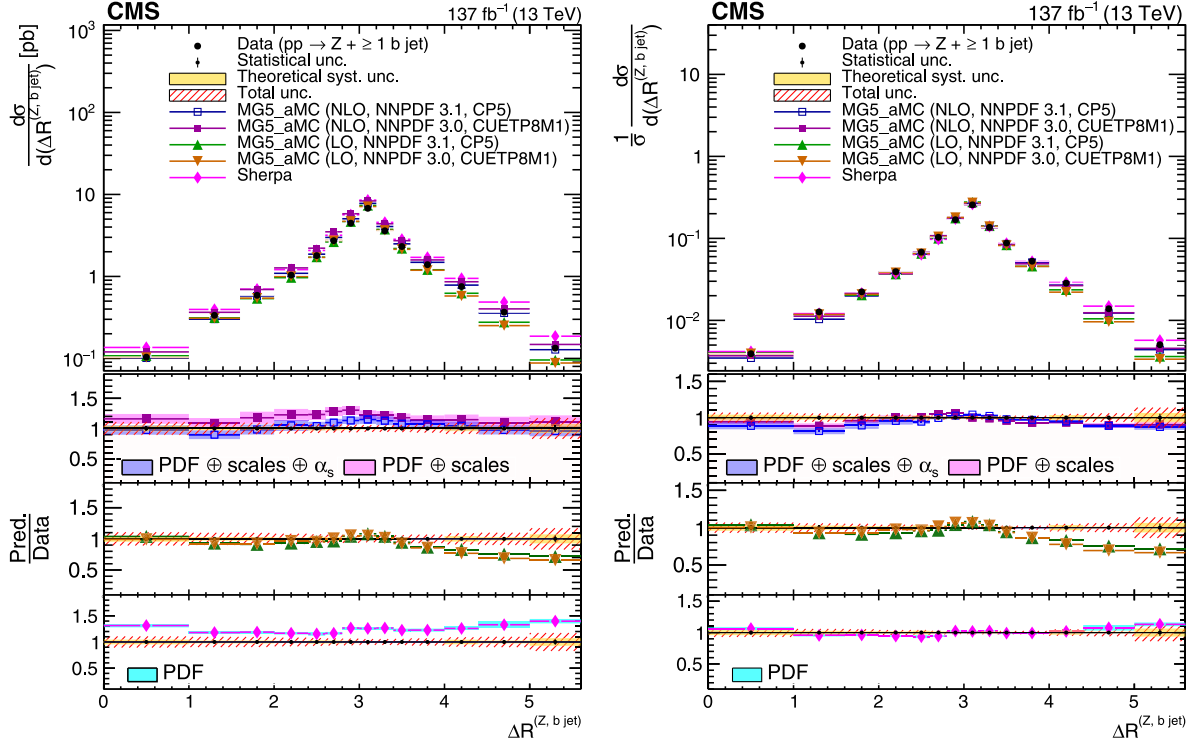


FIG. 7. (left) Differential cross section and the (right) normalized differential cross section distributions as functions of $\Delta R^{(Z, \text{b jet})}$ between the Z boson and the leading b jet for the $Z + \geq 1 \text{ b jet}$ events.

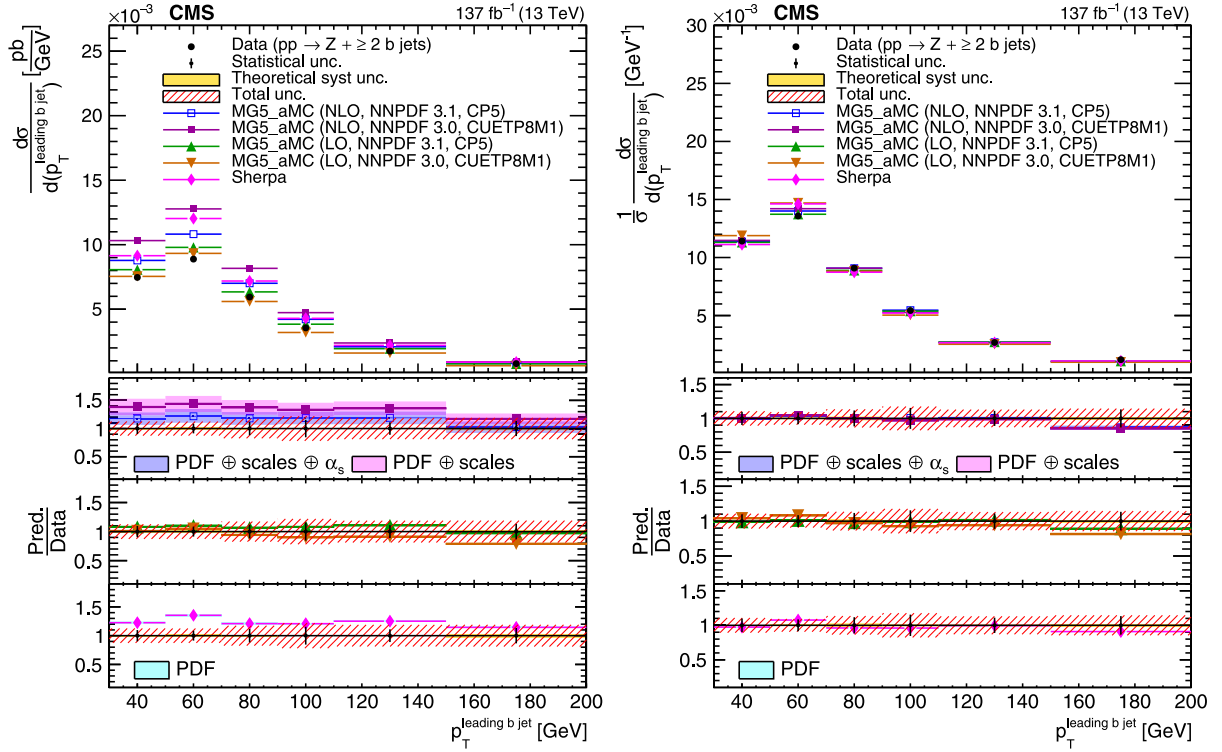


FIG. 8. (left) Differential cross section and the (right) normalized differential cross section distributions as functions of the leading b jet transverse momentum for the $Z + \geq 2 \text{ b jets}$ events.

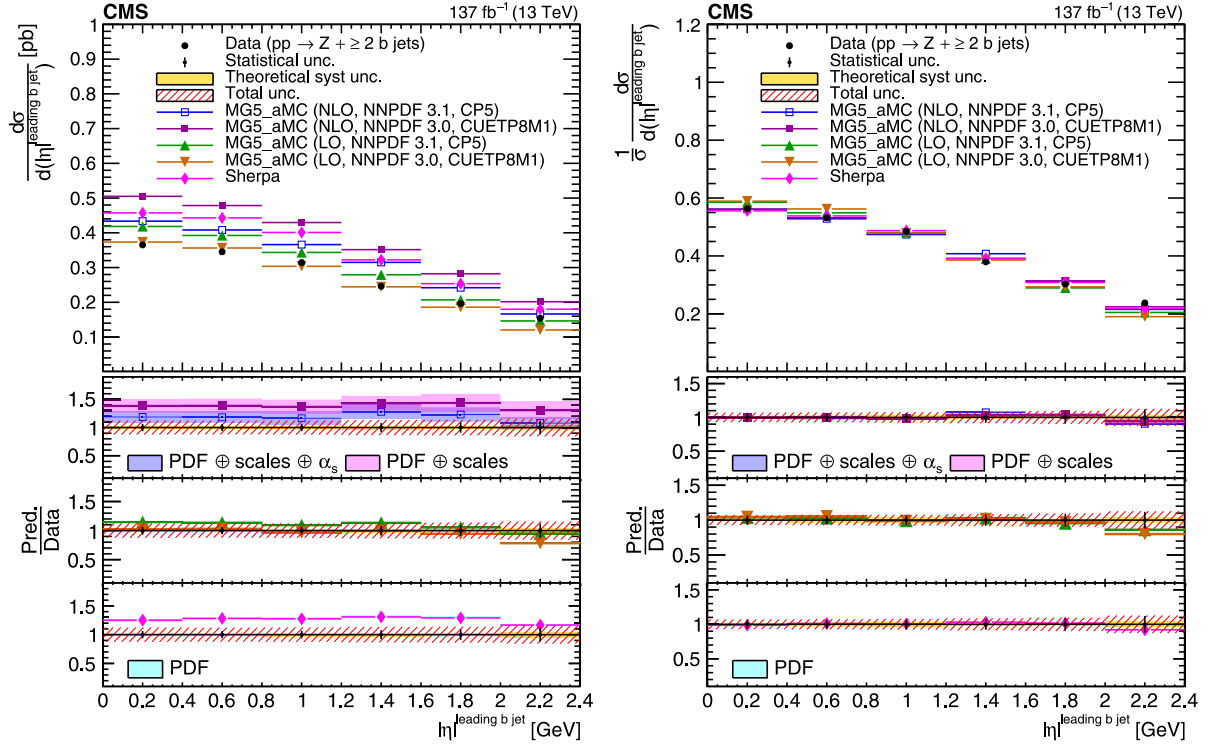


FIG. 9. (left) Differential cross section and the (right) normalized differential cross section distributions as functions of the leading b jet absolute pseudorapidity for the $Z + \geq 2$ b jets events.

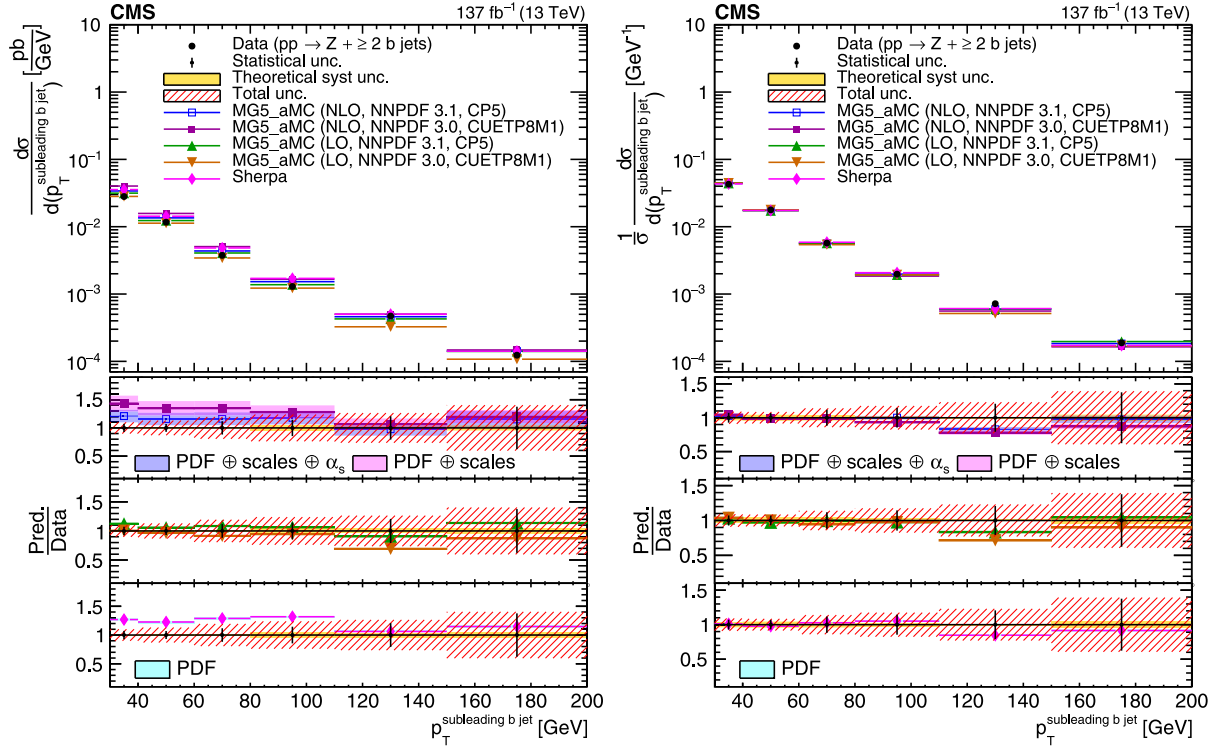


FIG. 10. (left) Differential cross section and the (right) normalized differential cross section distributions as functions of the subleading b jet transverse momentum for the $Z + \geq 2$ b jets events.

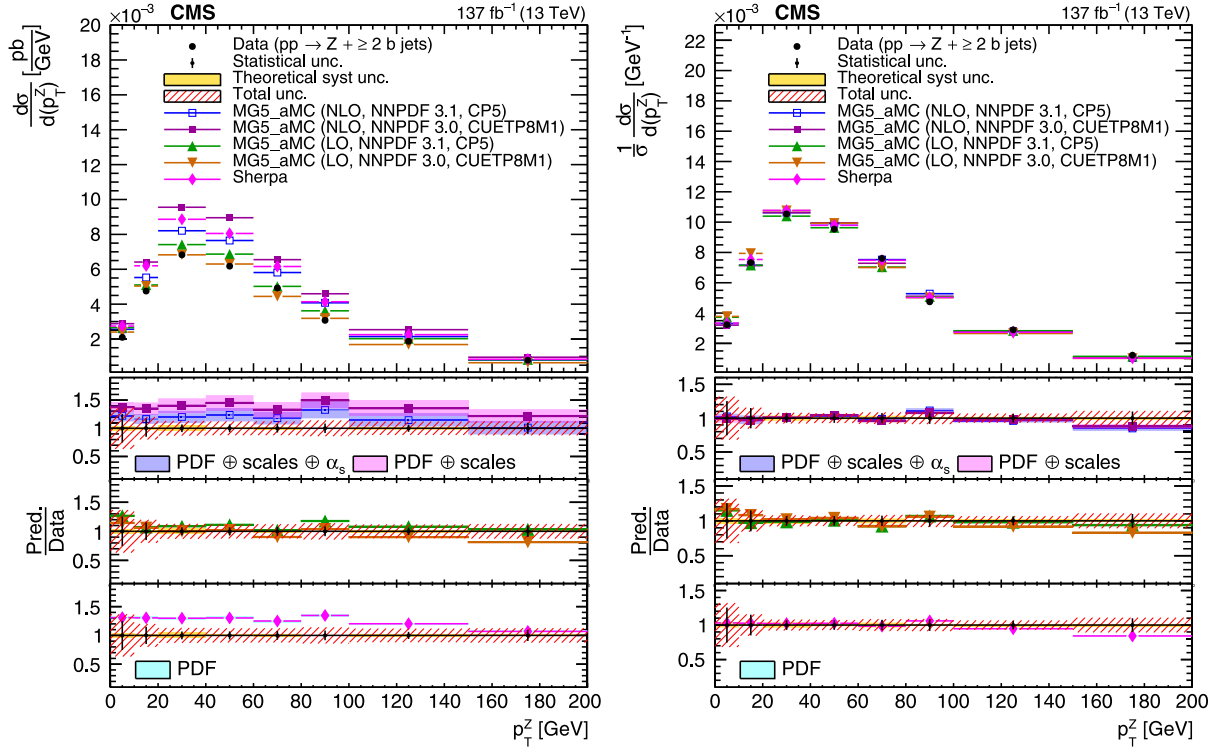


FIG. 11. (left) Differential cross section and the (right) normalized differential cross section distributions as functions of the Z boson transverse momentum for the $Z + \geq 2 \text{ b jets}$ events.

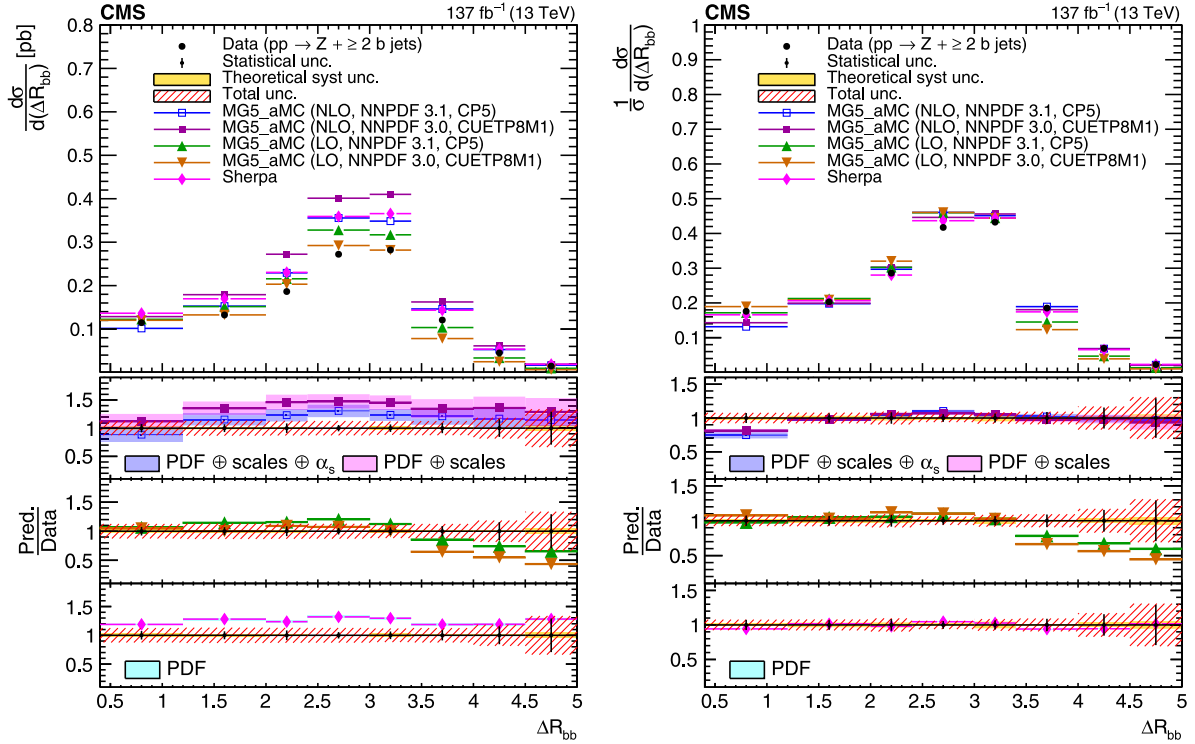


FIG. 12. (left) Differential cross section and the (right) normalized differential cross section distributions as functions of the angular separation between two b jets, ΔR_{bb} for the $Z + \geq 2 \text{ b jets}$ events.

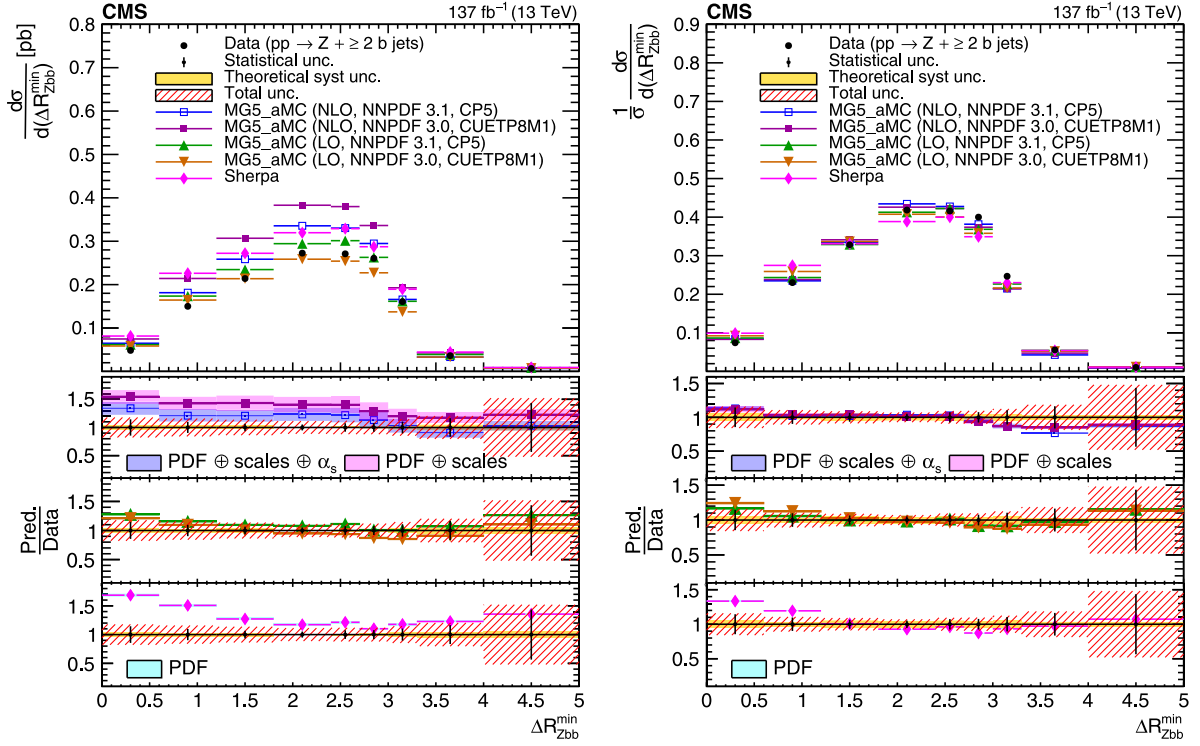


FIG. 13. (left) Differential cross section and the (right) normalized differential cross section distributions as functions of the minimum angular separation between the Z boson and two b jets, $\Delta R_{Zbb}^{\min}$ for the $Z + \geq 2 \text{ b jets}$ events.

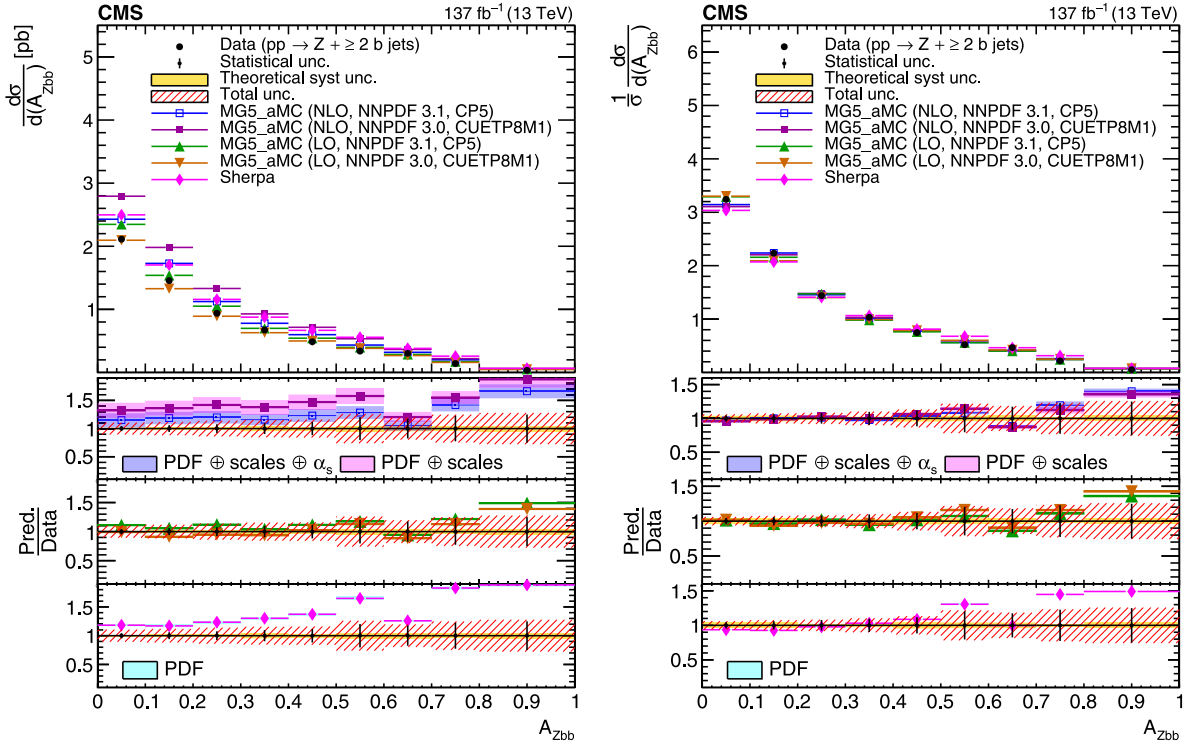


FIG. 14. (left) Differential cross section and the (right) normalized differential cross section distributions as functions of the asymmetry of the $Z + \geq 2 \text{ b jets}$ system, A_{Zbb} .

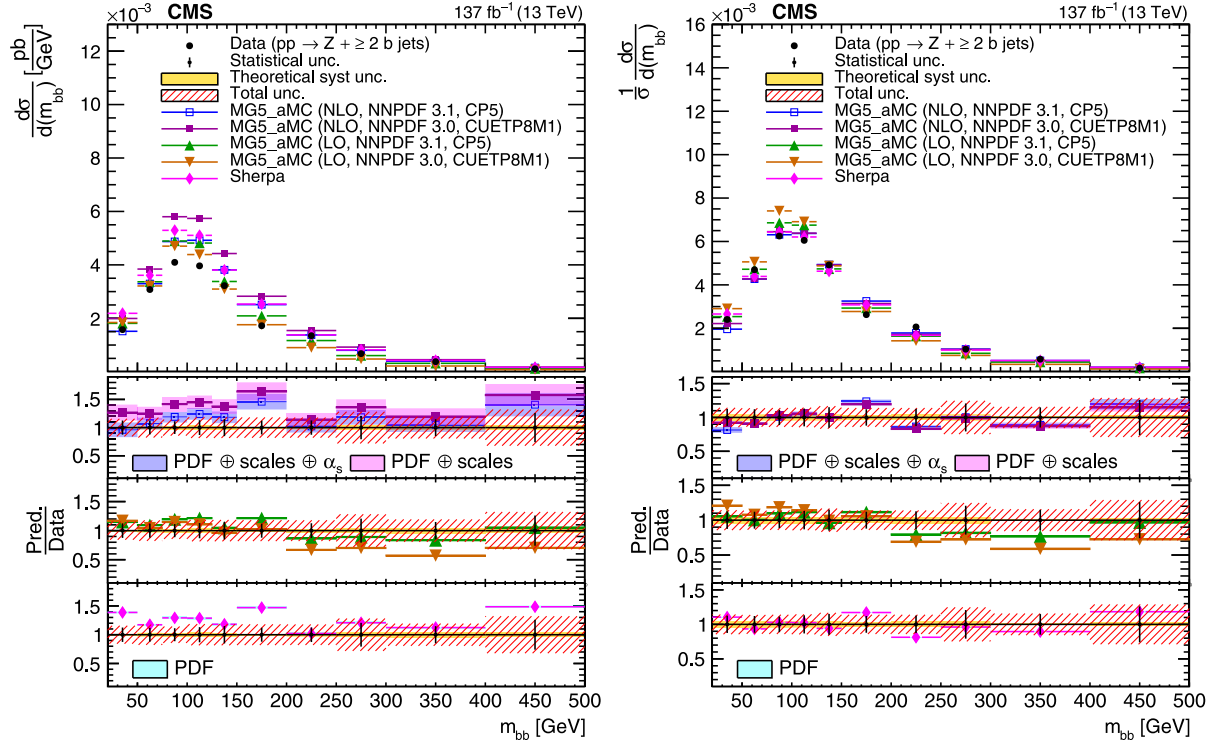


FIG. 15. (left) Differential cross section and the (right) normalized differential cross section distributions as functions of the invariant mass of the two b jets.

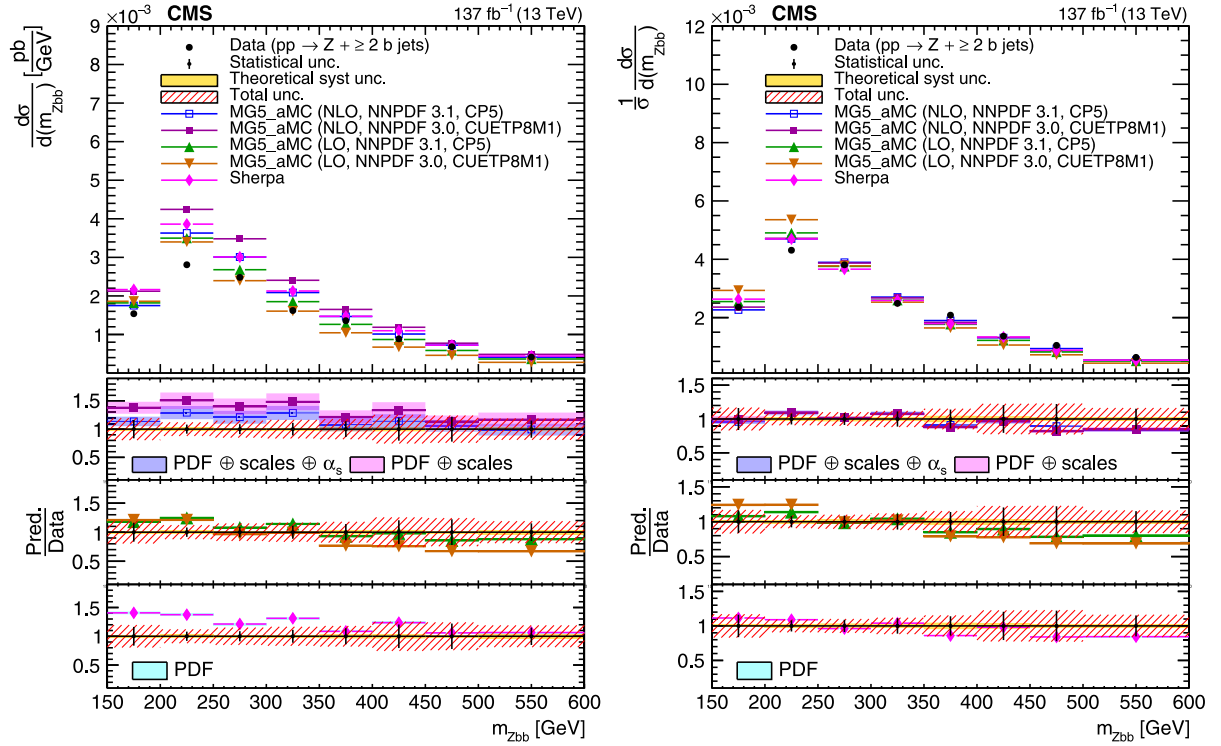


FIG. 16. (left) Differential cross section and the (right) normalized differential cross section distributions as functions of the invariant mass of the Z boson and two b jets.

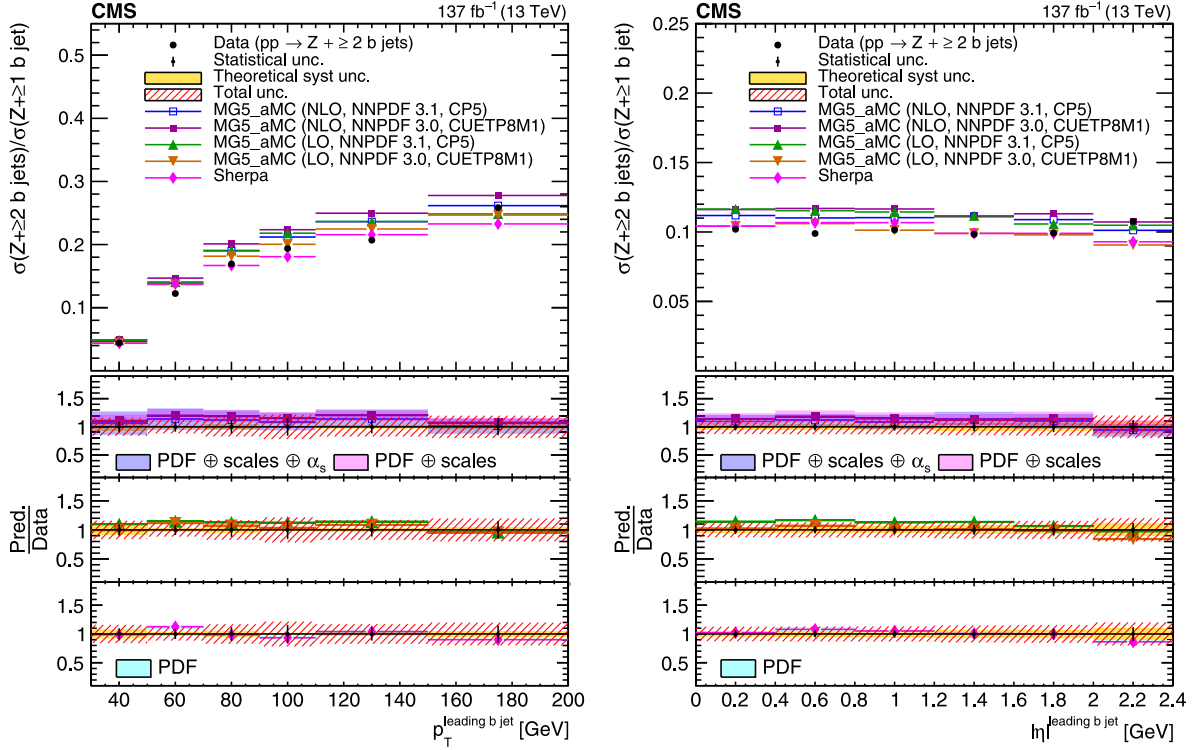


FIG. 17. Distributions of the cross section ratios, $\sigma(Z+ \geq 2 \text{ b jets})/\sigma(Z+ \geq 1 \text{ b jet})$, as functions of the (left) leading b jet transverse momentum and (right) absolute pseudorapidity.

(NLO, NNPDF 3.1, CP5), MG5_aMC (NLO, NNPDF 3.0, CUETP8M1), and SHERPA overestimate data by 20–30%, 30–50%, and 20–30%, respectively, whereas the MG5_aMC (LO) predictions are in good agreement with data. The angular correlations and asymmetry distributions of the $Z+ \geq 2 \text{ b jets}$ system are presented in Figs. 12–14. Although MG5_aMC (NLO) predictions are consistent with data in the normalized differential cross section distributions, except for the ΔR_{bb} , A_{Zbb} regions of 0.5–1.5 and 0.8–1, respectively, they are generally higher than the measured differential cross sections by 20%–30% and 30%–50% for MG5_aMC (NLO, NNPDF 3.1, CP5) and MG5_aMC (NLO, NNPDF 3.0, CUETP8M1), respectively. The MG5_aMC (LO) predictions show large deviations from data in the ΔR_{bb} distributions above 3.4, as can be seen in Fig. 12. In the $\Delta R_{Zbb}^{\min}$ distribution, Fig. 13, the MG5_aMC (LO) predictions are consistent with data, although they are higher than data in the last bin of the A_{Zbb} distribution (0.8–1) in Fig. 14, similar to the MG5_aMC (NLO, NNPDF 3.1, CP5) case. The MG5_aMC (NLO, NNPDF 3.0, CUETP8M1) simulation overestimates data in the differential $\Delta R_{Zbb}^{\min}$ distribution. The SHERPA predictions describe well the shape of the measured ΔR_{bb} differential cross sections but they significantly overestimate data in the low (<1.2) and high (>0.5) regions of the $\Delta R_{Zbb}^{\min}$ and A_{Zbb} distributions, respectively. Figure 15 shows invariant mass for two b jets events, and Fig. 16 shows the invariant mass for the $Z+ \geq 2 \text{ b jets}$ events. The shapes of these distributions are

described better by the MG5_aMC (NLO) and SHERPA predictions than the MG5_aMC (LO) ones. As for the rates, the data are described best by the MG5_aMC (NLO, NNPDF 3.1, CP5) predictions.

The $\sigma(Z+ \geq 2 \text{ b jets})/\sigma(Z+ \geq 1 \text{ b jet})$ cross section ratio distributions as functions of the leading b jet transverse momentum and absolute pseudorapidity are shown in Fig. 17. The ratio gradually increases (from 0.05 to 0.25) with the leading b jet p_T (ranging from 30 to 200 GeV), but is nearly independent of the pseudorapidity of the leading b jet. The increase in the ratio is due to a kinematic effect; at larger p_T of the leading jet, the kinematic acceptance for the subleading jet would increase. Predictions from MG5_aMC (LO), MG5_aMC (NLO), and SHERPA describe the measured ratios within uncertainties.

VII. SUMMARY

A measurement of fiducial cross sections of the $Z+ \geq 1 \text{ b jet}$ and $Z+ \geq 2 \text{ b jets}$ processes, along with the differential and normalized differential cross section distributions of different kinematic observables, is performed using proton-proton collisions data at $\sqrt{s} = 13 \text{ TeV}$ collected by the CMS experiment at the CERN LHC. This is the first measurement of these processes based on data collected during the 2016–2018 LHC running period. The fiducial cross sections are measured to be $6.52 \pm 0.04(\text{stat}) \pm 0.40(\text{syst}) \pm 0.14(\text{theo}) \text{ pb}$ for the $Z+ \geq 1 \text{ b jet}$ and

$0.65 \pm 0.03(\text{stat}) \pm 0.07(\text{syst}) \pm 0.02(\text{theo})$ pb for the $Z + \geq 2b$ jets, which are better described by the MG5_aMC leading order (LO) simulation but overestimated by MG5_aMC next-to-LO (NLO) and SHERPA predictions. Since all predictions are normalized to the inclusive $Z + \text{jets}$ next-to-NLO cross section, differences between MG5_aMC (NLO) and MG5_aMC (LO) results could be attributable to variations in shapes of observables and settings (parton distribution functions, Monte Carlo tunes, matching schemes) used in those simulations. The SHERPA simulation overestimates the measured integrated cross section; however, it provides a good description of the shapes of various kinematic observables. The MG5_aMC (LO) and MG5_aMC (NLO) generators interfaced with PYTHIA describe the fiducial cross section better but do not completely describe the shapes of the kinematic observables. Present measurements can be used as an input for the further optimization of the simulation parameters. The measured value of the cross section ratio of the $Z + \geq 2b$ jets to $Z + \geq 1b$ jet is $0.100 \pm 0.005(\text{stat}) \pm 0.007(\text{syst}) \pm 0.003(\text{theo})$, which is well described by the MG5_aMC (LO, NNPDF 3.0, CUETP8M1) and SHERPA calculations but overestimated by MG5_aMC (NLO) prediction.

ACKNOWLEDGMENTS

We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC and thank the technical and administrative staffs at CERN and at other CMS institutes for their contributions to the success of the CMS effort. In addition, we gratefully acknowledge the computing centers and personnel of the Worldwide LHC Computing Grid and other centers for delivering so effectively the computing infrastructure essential to our analyses. Finally, we acknowledge the enduring support for the construction and operation of the LHC, the CMS detector, and the supporting computing infrastructure provided by the following funding agencies: BMBWF and FWF (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, FAPERGS, and FAPESP (Brazil); MES and BNSF (Bulgaria); CERN; CAS, MoST, and NSFC (China); MINCIENCIAS (Colombia); MSES and CSF (Croatia); RIF (Cyprus); SENESCYT (Ecuador); MoER, ERC PUT and ERDF (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Germany); GSRI (Greece); NKFI (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); MSIP and NRF (Republic of Korea); MES (Latvia); LAS (Lithuania); MOE and UM (Malaysia); BUAP, CINVESTAV, CONACYT, LNS, SEP, and UASLP-FAI (Mexico); MOS (Montenegro); MBIE (New Zealand); PAEC (Pakistan); MSHE and NSC (Poland); FCT (Portugal); JINR (Dubna); MON, RosAtom, RAS, RFBR, and NRC KI (Russia); MESTD (Serbia); MCIN/AEI and PCTI (Spain); MOSTR

(Sri Lanka); Swiss Funding Agencies (Switzerland); MST (Taipei); ThEPCenter, IPST, STAR, and NSTDA (Thailand); TUBITAK and TAEK (Turkey); NASU (Ukraine); STFC (United Kingdom); DOE and NSF (USA). Individuals have received support from the Marie-Curie program and the European Research Council and Horizon 2020 Grant, contracts No. 675440, No. 724704, No. 752730, No. 758316, No. 765710, 824093, No. 884104, and COST Action No. CA16108 (European Union); the Leventis Foundation; the Alfred P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Federal Science Policy Office; the Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture (FRIA-Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie (IWT-Belgium); the F.R.S.-FNRS and FWO (Belgium) under the “Excellence of Science—EOS”—be.h Project No. 30820817; the Beijing Municipal Science & Technology Commission, No. Z191100007219010; the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic; the Deutsche Forschungsgemeinschaft (DFG), under Germany's Excellence Strategy—EXC 2121 “Quantum Universe”—390833306, and under Project No. 400140256—GRK2497; the Lendület (“Momentum”) Program and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences, the New National Excellence Program ÚNKP, the NKFI research grants No. 123842, No. 123959, No. 124845, No. 124850, No. 125105, No. 128713, No. 128786, and No. 129058 (Hungary); the Council of Science and Industrial Research, India; the Latvian Council of Science; the Ministry of Science and Higher Education and the National Science Center, contracts Opus 2014/15/B/ST2/03998 and 2015/19/B/ST2/02861 (Poland); the Fundação para a Ciência e a Tecnologia, Grant No. CEECIND/01334/2018 (Portugal); the National Priorities Research Program by Qatar National Research Fund; the Ministry of Science and Higher Education, Projects No. 14.W03.31.0026 and No. FSWW-2020-0008, and the Russian Foundation for Basic Research, Project No. 19-42-703014 (Russia); MCIN/AEI/10.13039/501100011033, ERDF “a way of making Europe”, and the Programa Estatal de Fomento de la Investigación Científica y Técnica de Excelencia María de Maeztu, grant No. MDM-2017-0765 and Programa Severo Ochoa del Principado de Asturias (Spain); the Stavros Niarchos Foundation (Greece); the Rachadapisek Sompot Fund for Postdoctoral Fellowship, Chulalongkorn University and the Chulalongkorn Academic into Its 2nd Century Project Advancement Project (Thailand); the Kavli Foundation; the Nvidia Corporation; the SuperMicro Corporation; the Welch Foundation, Contract NO. C-1845; and the Weston Havens Foundation (USA).

- [1] S. J. Brodsky, A. Kusina, F. Lyonnet, I. Schienbein, H. Spiesberger, and R. Vogt, A review of the intrinsic heavy quark content of the nucleon, *Adv. High Energy Phys.* **2015**, 231547 (2015).
- [2] G. Bozzi, J. Rojo, and A. Vicini, Impact of the parton distribution function uncertainties on the measurement of the W boson mass at the Tevatron and the LHC, *Phys. Rev. D* **83**, 113008 (2011).
- [3] F. Maltoni, G. Ridolfi, and M. Ubiali, b-initiated processes at the LHC: A reappraisal, *J. High Energy Phys.* **07** (2012) 022.
- [4] T. Aaltonen *et al.* (CDF Collaboration), Measurement of cross sections for b jet production in events with a Z boson in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV, *Phys. Rev. D* **79**, 052008 (2009).
- [5] V. M. Abazov *et al.* (D0 Collaboration), A measurement of the ratio of inclusive cross sections of the ratio of inclusive cross sections $\sigma(p\bar{p} \rightarrow Z + b\text{jet})/\sigma(p\bar{p} \rightarrow Z + \text{jet})$ at $\sqrt{s} = 1.96$ TeV, *Phys. Rev. D* **83**, 031105 (2011).
- [6] CMS Collaboration, Measurement of the production cross sections for a Z boson and one or more b jets in pp collisions at $\sqrt{s} = 7$ TeV, *J. High Energy Phys.* **06** (2014) 120.
- [7] ATLAS Collaboration, Measurement of differential production cross-sections for a Z boson in association with b-jets in 7 TeV proton-proton collisions with the ATLAS detector, *J. High Energy Phys.* **10** (2014) 141.
- [8] CMS Collaboration, Measurements of the associated production of a Z boson and b jets in pp collisions at $\sqrt{s} = 8$ TeV, *Eur. Phys. J. C* **77**, 751 (2017).
- [9] ATLAS Collaboration, Measurements of the production cross-section for a Z boson in association with b-jets in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, *J. High Energy Phys.* **07** (2020) 044.
- [10] P. A. Zyla *et al.* (Particle Data Group), Review of particle physics, *Prog. Theor. Exp. Phys.* **2020**, 083C01 (2020).
- [11] HEPData record for this analysis (2021).
- [12] W. Adam, T. Bergauer, D. Blöchl, and M. D. others (CMS Tracker Group), The CMS phase-1 pixel detector upgrade, *J. Instrum.* **16**, P02027 (2021).
- [13] CMS Collaboration, The CMS trigger system, *J. Instrum.* **12**, P01020 (2017).
- [14] CMS Collaboration, Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV, *J. Instrum.* **15**, P10017 (2020).
- [15] CMS Collaboration, The CMS experiment at the CERN LHC, *J. Instrum.* **3**, S08004 (2008).
- [16] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations, *J. High Energy Phys.* **07** (2014) 079.
- [17] R. Frederix and S. Frixione, Merging meets matching in MC@NLO, *J. High Energy Phys.* **12** (2012) 061.
- [18] R. D. Ball *et al.* (NNPDF Collaboration), Parton distributions for the LHC Run II, *J. High Energy Phys.* **04** (2015) 040.
- [19] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, MadGraph5: Going beyond, *J. High Energy Phys.* **06** (2011) 128.
- [20] T. Sjöstrand, S. A. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, An introduction to PYTHIA8.2, *Comput. Phys. Commun.* **191**, 159 (2015).
- [21] CMS Collaboration, Event generator tunes obtained from underlying event and multiparton scattering measurements, *Eur. Phys. J. C* **76**, 155 (2016).
- [22] CMS Collaboration, Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements, *Eur. Phys. J. C* **80**, 4 (2020).
- [23] E. Bothmann, G. Singh Chahal, S. Höche, J. Krause, F. Krauss, S. Kuttimalai, S. Liebschner, D. Napoletano, M. Schönherr, H. Schulz, S. Schumann, and F. Siegert, Event generation with SHERPA2.2, *SciPost Phys.* **7**, 034 (2019).
- [24] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, QCD matrix elements + parton showers: The NLO case, *J. High Energy Phys.* **04** (2013) 027.
- [25] T. Gehrmann, S. Hoche, F. Krauss, M. Schönherr, and F. Siegert, NLO QCD matrix elements + parton showers in $e^+e^- \rightarrow \text{hadrons}$, *J. High Energy Phys.* **01** (2013) 144.
- [26] S. Höche, F. Krauss, M. Schönherr, J. M. Thompson, S. Pozzorini, and K. C. Zapp, Triple vector boson production through Higgs-strahlung with NLO multijet merging, *Phys. Rev. D* **89**, 093015 (2014).
- [27] Y. Li and F. Petriello, Combining QCD and electroweak corrections to dilepton production in FEWZ, *Phys. Rev. D* **86**, 094034 (2012).
- [28] P. Nason, A new method for combining NLO QCD with shower Monte Carlo algorithms, *J. High Energy Phys.* **11** (2004) 040.
- [29] S. Frixione, P. Nason, and C. Oleari, Matching NLO QCD computations with parton shower simulations: The POWHEG method, *J. High Energy Phys.* **11** (2007) 070.
- [30] S. Alioli, P. Nason, C. Oleari, and E. Re, A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX, *J. High Energy Phys.* **06** (2010) 043.
- [31] CMS Collaboration, Investigations of the impact of the parton shower tuning in PYTHIA 8 in the modelling of $t\bar{t}$ at $\sqrt{s} = 8$ and 13 TeV, CMS Physics Analysis Summary, Report No. CMS-PAS-TOP-16-021, 2016, <https://cds.cern.ch/record/2235192>.
- [32] R. D. Ball *et al.* (NNPDF Collaboration), Parton distributions with LHC data, *Nucl. Phys.* **B867**, 244 (2013).
- [33] M. Grazzini, S. Kallweit, and M. Wiesemann, Fully differential NNLO computations with MATRIX, *Eur. Phys. J. C* **78**, 537 (2018).
- [34] N. Kidonakis, Top quark production, in *Proceedings of the Helmholtz International Summer School on Physics of Heavy Quarks and Hadrons (HQ 2013): JINR, Dubna, Russia, 2013* (DESY, Hamburg, Germany, 2014), p. 139, 10.3204/DESY-PROC-2013-03/Kidonakis.
- [35] M. Czakon, P. Fiedler, and A. Mitov, Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_s^4)$, *Phys. Rev. Lett.* **110**, 252004 (2013).
- [36] P. Bärnreuther, M. Czakon, and A. Mitov, Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \rightarrow t\bar{t} + X$, *Phys. Rev. Lett.* **109**, 132001 (2012).
- [37] M. Beneke, P. Falgari, S. Klein, and C. Schwinn, Hadronic top-quark pair production with NNLL threshold resummation, *Nucl. Phys.* **B855**, 695 (2012).

- [38] M. Cacciari, M. Czakon, M. Mangano, A. Mitov, and P. Nason, Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation, *Phys. Lett. B* **710**, 612 (2012).
- [39] M. Czakon and A. Mitov, NNLO corrections to top-pair production at hadron colliders: The all-fermionic scattering channels, *J. High Energy Phys.* **12** (2012) 054.
- [40] M. Czakon and A. Mitov, NNLO corrections to top pair production at hadron colliders: The quark-gluon reaction, *J. High Energy Phys.* **01** (2013) 080.
- [41] M. Czakon and A. Mitov, Top++: A program for the calculation of the top-pair cross-section at hadron colliders, *Comput. Phys. Commun.* **185**, 2930 (2014).
- [42] S. Agostinelli *et al.* (GEANT4 Collaboration), GEANT4—a simulation toolkit, *Nucl. Instrum. Methods Phys. Res., Sect. A* **506**, 250 (2003).
- [43] CMS Collaboration, Particle-flow reconstruction and global event description with the CMS detector, *J. Instrum.* **12**, P10003 (2017).
- [44] M. Cacciari, G. P. Salam, and G. Soyez, The anti- k_T jet clustering algorithm, *J. High Energy Phys.* **04** (2008) 063.
- [45] CMS Collaboration, Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV, *J. Instrum.* **10**, P06005 (2015).
- [46] CMS Collaboration, Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV, *J. Instrum.* **13**, P06015 (2018).
- [47] A. Bodek, A. van Dyne, J. Y. Han, W. Sakumoto, and A. Strelnikov, Extracting muon momentum scale corrections for hadron collider experiments, *Eur. Phys. J. C* **72**, 2194 (2012).
- [48] M. Cacciari and G. P. Salam, Pileup subtraction using jet areas, *Phys. Lett. B* **659**, 119 (2008).
- [49] CMS Collaboration, Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC, *J. Instrum.* **16**, P05014 (2021).
- [50] CMS Collaboration, Measurement of the inclusive W and Z production cross sections in pp collisions at $\sqrt{s} = 7$ TeV with the CMS experiment, *J. High Energy Phys.* **10** (2011) 132.
- [51] M. Cacciari, G. P. Salam, and G. Soyez, FastJet user manual, *Eur. Phys. J. C* **72**, 1896 (2012).
- [52] CMS Collaboration, Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV, *J. Instrum.* **12**, P02014 (2017).
- [53] CMS Collaboration, Jet algorithms performance in 13 TeV data, CMS Physics Analysis Summary, Report No. CMS-PAS-JME-16-003, 2017, <https://cds.cern.ch/record/2256875>.
- [54] CMS Collaboration, Pileup mitigation at CMS in 13 TeV data, *J. Instrum.* **15**, P09018 (2020).
- [55] CMS Collaboration, Pileup jet identification, CMS Physics Analysis Summary, Report No. CMS-PAS-JME-13-005, 2013, <https://cds.cern.ch/record/1581583>.
- [56] CMS Collaboration, Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector, *J. Instrum.* **14**, P07004 (2019).
- [57] CMS Collaboration, Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV, *J. Instrum.* **13**, P05011 (2018).
- [58] S. Schmitt, TUnfold: An algorithm for correcting migration effects in high energy physics, *J. Instrum.* **7**, T10003 (2012).
- [59] A. N. Tikhonov, Solution of incorrectly formulated problems and the regularization method, *Sov. Math. Dokl.* **4**, 1035 (1963).
- [60] CMS Collaboration, Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS, *Eur. Phys. J. C* **81**, 800 (2021).
- [61] CMS Collaboration, CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV, CMS Physics Analysis Summary, Report No. CMS-PAS-LUM-17-004, 2018, <https://cds.cern.ch/record/2621960>.
- [62] CMS Collaboration, CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV, CMS Physics Analysis Summary, Report No. CMS-PAS-LUM-18-002, 2019, <https://cds.cern.ch/record/2676164>.
- [63] CMS Collaboration, Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV, *J. High Energy Phys.* **07** (2018) 161.
- [64] D. Bourilkov, R. C. Group, and M. R. Whalley, LHAPDF: PDF use from the Tevatron to the LHC, in *TeV4LHC Workshop—4th meeting Batavia, Illinois, 2005* (Fermilab, 2006).

A. Tumasyan,¹ W. Adam,² J. W. Andrejkovic,² T. Bergauer,² S. Chatterjee,² K. Damanakis,² M. Dragicevic,² A. Escalante Del Valle,² R. Frühwirth,^{2,b} M. Jeitler,^{2,b} N. Krammer,² L. Lechner,² D. Liko,² I. Mikulec,² P. Paulitsch,² F. M. Pitters,² J. Schieck,^{2,b} R. Schöffbeck,² D. Schwarz,² S. Templ,² W. Waltenberger,² C.-E. Wulz,^{2,b} V. Chekhovsky,³ A. Litomin,³ V. Makarenko,³ M. R. Darwish,^{4,c} E. A. De Wolf,⁴ T. Janssen,⁴ T. Kello,^{4,d} A. Lelek,⁴ H. Rejeb Sfar,⁴ P. Van Mechelen,⁴ S. Van Putte,⁴ N. Van Remortel,⁴ F. Blekman,⁵ E. S. Bols,⁵ J. D'Hondt,⁵ M. Delcourt,⁵ H. El Faham,⁵ S. Lowette,⁵ S. Moortgat,⁵ A. Morton,⁵ D. Müller,⁵ A. R. Sahasransu,⁵ S. Tavernier,⁵ W. Van Doninck,⁵ D. Beghin,⁶ B. Bilin,⁶ B. Clerbaux,⁶ G. De Lentdecker,⁶ L. Favart,⁶ A. Grebenyuk,⁶ A. K. Kalsi,⁶ K. Lee,⁶ M. Mahdavihorrani,⁶ I. Makarenko,⁶ L. Moureaux,⁶ L. Pétré,⁶ A. Popov,⁶ N. Postiau,⁶ E. Starling,⁶ L. Thomas,⁶ M. Vanden Bemden,⁶ C. Vander Velde,⁶ P. Vanlaer,⁶ T. Cornelis,⁷ D. Dobur,⁷ J. Knolle,⁷ L. Lambrecht,⁷ G. Mestdach,⁷ M. Niedziela,⁷ C. Roskas,⁷ A. Samalan,⁷ K. Skovpen,⁷ M. Tytgat,⁷ B. Vermassen,⁷ L. Wezenbeek,⁷ A. Benecke,⁸ A. Bethani,⁸ G. Bruno,⁸ F. Bury,⁸ C. Caputo,⁸ P. David,⁸ C. Delaere,⁸ I. S. Donertas,⁸ A. Giammanco,⁸ K. Jaffel,⁸ Sa. Jain,⁸ V. Lemaître,⁸ K. Mondal,⁸

L. Wiens,⁴⁴ C. Wissing,⁴⁴ S. Wuchterl,⁴⁴ R. Aggleton,⁴⁵ S. Albrecht,⁴⁵ S. Bein,⁴⁵ L. Benato,⁴⁵ P. Connor,⁴⁵ K. De Leo,⁴⁵ M. Eich,⁴⁵ F. Feindt,⁴⁵ A. Fröhlich,⁴⁵ C. Garbers,⁴⁵ E. Garutti,⁴⁵ P. Gunnellini,⁴⁵ M. Hajheidari,⁴⁵ J. Haller,⁴⁵ A. Hinzmann,⁴⁵ G. Kasieczka,⁴⁵ R. Klanner,⁴⁵ T. Kramer,⁴⁵ V. Kutzner,⁴⁵ J. Lange,⁴⁵ T. Lange,⁴⁵ A. Lobanov,⁴⁵ A. Malara,⁴⁵ A. Nigamova,⁴⁵ K. J. Pena Rodriguez,⁴⁵ O. Rieger,⁴⁵ P. Schleper,⁴⁵ M. Schröder,⁴⁵ J. Schwandt,⁴⁵ J. Sonneveld,⁴⁵ H. Stadie,⁴⁵ G. Steinbrück,⁴⁵ A. Tews,⁴⁵ I. Zoi,⁴⁵ J. Bechtel,⁴⁶ S. Brommer,⁴⁶ M. Burkart,⁴⁶ E. Butz,⁴⁶ R. Caspart,⁴⁶ T. Chwalek,⁴⁶ W. De Boer,^{46,a} A. Dierlamm,⁴⁶ A. Droll,⁴⁶ K. El Morabit,⁴⁶ N. Faltermann,⁴⁶ M. Giffels,⁴⁶ J. o. Gosewisch,⁴⁶ A. Gottmann,⁴⁶ F. Hartmann,^{46,u} C. Heidecker,⁴⁶ U. Husemann,⁴⁶ P. Keicher,⁴⁶ R. Koppenhöfer,⁴⁶ S. Maier,⁴⁶ M. Metzler,⁴⁶ S. Mitra,⁴⁶ Th. Müller,⁴⁶ M. Neukum,⁴⁶ A. Nürnberg,⁴⁶ G. Quast,⁴⁶ K. Rabbertz,⁴⁶ J. Rauser,⁴⁶ D. Savoie,⁴⁶ M. Schnepf,⁴⁶ D. Seith,⁴⁶ I. Shvetsov,⁴⁶ H. J. Simonis,⁴⁶ R. Ulrich,⁴⁶ J. Van Der Linden,⁴⁶ R. F. Von Cube,⁴⁶ M. Wassmer,⁴⁶ M. Weber,⁴⁶ S. Wieland,⁴⁶ R. Wolf,⁴⁶ S. Wozniowski,⁴⁶ S. Wunsch,⁴⁶ G. Anagnostou,⁴⁷ G. Daskalakis,⁴⁷ T. Gerasis,⁴⁷ A. Kyriakis,⁴⁷ D. Loukas,⁴⁷ A. Stakia,⁴⁷ M. Diamantopoulou,⁴⁸ D. Karasavvas,⁴⁸ G. Karathanasis,⁴⁸ P. Kontaxakis,⁴⁸ C. K. Koraka,⁴⁸ A. Manousakis-Katsikakis,⁴⁸ A. Panagiotou,⁴⁸ I. Papavergou,⁴⁸ N. Saoulidou,⁴⁸ K. Theofilatos,⁴⁸ E. Tziaferi,⁴⁸ K. Vellidis,⁴⁸ E. Vourliotis,⁴⁸ G. Bakas,⁴⁹ K. Kousouris,⁴⁹ I. Papakrivopoulos,⁴⁹ G. Tsipolitis,⁴⁹ A. Zacharopoulou,⁴⁹ K. Adamidis,⁵⁰ I. Bestintzanos,⁵⁰ I. Evangelou,⁵⁰ C. Foudas,⁵⁰ P. Gianneios,⁵⁰ P. Katsoulis,⁵⁰ P. Kokkas,⁵⁰ N. Manthos,⁵⁰ I. Papadopoulos,⁵⁰ J. Strologas,⁵⁰ M. Csanad,⁵¹ K. Farkas,⁵¹ M. M. A. Gadallah,^{51,aa} S. Lökös,^{51,bb} P. Major,⁵¹ K. Mandal,⁵¹ A. Mehta,⁵¹ G. Pasztor,⁵¹ A. J. Rádl,⁵¹ O. Surányi,⁵¹ G. I. Veres,⁵¹ M. Bartók,^{52,cc} G. Bencze,⁵² C. Hajdu,⁵² D. Horvath,^{52,dd} F. Sikler,⁵² V. Veszpremi,⁵² S. Czellar,⁵³ D. Fasanella,⁵³ J. Karancsi,^{53,cc} J. Molnar,⁵³ Z. Szillasi,⁵³ D. Teyssier,⁵³ P. Raics,⁵⁴ Z. L. Trocsanyi,^{54,ee} B. Ujvari,⁵⁴ T. Csorgo,^{55,ff} F. Nemes,^{55,ff} T. Novak,⁵⁵ S. Choudhury,⁵⁶ J. R. Komaragiri,⁵⁶ D. Kumar,⁵⁶ L. Panwar,⁵⁶ P. C. Tiwari,⁵⁶ S. Bahinipati,^{57,gg} C. Kar,⁵⁷ P. Mal,⁵⁷ T. Mishra,⁵⁷ V. K. Muraleedharan Nair Bindhu,^{57,hh} A. Nayak,^{57,hh} P. Saha,⁵⁷ N. Sur,⁵⁷ S. K. Swain,⁵⁷ D. Vats,^{57,hh} S. Bansal,⁵⁸ S. B. Beri,⁵⁸ V. Bhatnagar,⁵⁸ G. Chaudhary,⁵⁸ S. Chauhan,⁵⁸ N. Dhirga,^{58,ii} R. Gupta,⁵⁸ A. Kaur,⁵⁸ M. Kaur,⁵⁸ S. Kaur,⁵⁸ P. Kumari,⁵⁸ M. Meena,⁵⁸ K. Sandeep,⁵⁸ J. B. Singh,⁵⁸ A. K. Viridi,⁵⁸ A. Ahmed,⁵⁹ A. Bhardwaj,⁵⁹ B. C. Choudhary,⁵⁹ M. Gola,⁵⁹ S. Keshri,⁵⁹ A. Kumar,⁵⁹ M. Naimuddin,⁵⁹ P. Priyanka,⁵⁹ K. Ranjan,⁵⁹ A. Shah,⁵⁹ M. Bharti,^{60,jj} R. Bhattacharya,⁶⁰ S. Bhattacharya,⁶⁰ D. Bhowmik,⁶⁰ S. Dutta,⁶⁰ S. Dutta,⁶⁰ B. Gomber,^{60,kk} M. Maity,^{60,ll} P. Palit,⁶⁰ P. K. Rout,⁶⁰ G. Saha,⁶⁰ B. Sahu,⁶⁰ S. Sarkar,⁶⁰ M. Sharan,⁶⁰ B. Singh,^{60,jj} S. Thakur,^{60,jj} P. K. Behera,⁶¹ S. C. Behera,⁶¹ P. Kalbhor,⁶¹ A. Muhammad,⁶¹ R. Pradhan,⁶¹ P. R. Pujahari,⁶¹ A. Sharma,⁶¹ A. K. Sikdar,⁶¹ D. Dutta,⁶² V. Jha,⁶² V. Kumar,⁶² D. K. Mishra,⁶² K. Naskar,^{62,mm} P. K. Netrakanti,⁶² L. M. Pant,⁶² P. Shukla,⁶² T. Aziz,⁶³ S. Dugad,⁶³ M. Kumar,⁶³ S. Banerjee,⁶⁴ R. Chudasama,⁶⁴ M. Guchait,⁶⁴ S. Karmakar,⁶⁴ S. Kumar,⁶⁴ G. Majumder,⁶⁴ K. Mazumdar,⁶⁴ S. Mukherjee,⁶⁴ K. Alpna,⁶⁵ S. Dube,⁶⁵ B. Kansal,⁶⁵ A. Laha,⁶⁵ S. Pandey,⁶⁵ A. Rane,⁶⁵ A. Rastogi,⁶⁵ S. Sharma,⁶⁵ H. Bakhshiansohi,^{66,nn} E. Khazaie,⁶⁶ M. Zeinali,^{66,oo} S. Chenarani,^{67,pp} S. M. Etesami,⁶⁷ M. Khakzad,⁶⁷ M. Mohammadi Najafabadi,⁶⁷ M. Grunewald,⁶⁸ M. Abbrescia,^{69a,69b} R. Aly,^{69a,69b,qq} C. Aruta,^{69a,69b} A. Colaleo,^{69a} D. Creanza,^{69a,69c} N. De Filippis,^{69a,69c} M. De Palma,^{69a,69b} A. Di Florio,^{69a,69b} A. Di Pilato,^{69a,69b} W. Elmetenawee,^{69a,69b} L. Fiore,^{69a} A. Gelmi,^{69a,69b} M. Gul,^{69a} G. Iaselli,^{69a,69c} M. Ince,^{69a,69b} S. Lezki,^{69a,69b} G. Maggi,^{69a,69c} M. Maggi,^{69a} I. Margjeka,^{69a,69b} V. Mastrapasqua,^{69a,69b} S. My,^{69a,69b} S. Nuzzo,^{69a,69b} A. Pellecchia,^{69a,69b} A. Pompili,^{69a,69b} G. Pugliese,^{69a,69c} D. Ramos,^{69a} A. Ranieri,^{69a} G. Selvaggi,^{69a,69b} L. Silvestris,^{69a} F. M. Simone,^{69a,69b} Ü. Sözbilir,^{69a} R. Venditti,^{69a} P. Verwilligen,^{69a} G. Abbiendi,^{70a} C. Battilana,^{70a,70b} D. Bonacorsi,^{70a,70b} L. Borroni,^{70a} L. Brigliadori,^{70a} R. Campanini,^{70a,70b} P. Capiluppi,^{70a,70b} A. Castro,^{70a,70b} F. R. Cavallo,^{70a} M. Cuffiani,^{70a,70b} G. M. Dallavalle,^{70a} T. Diotallevi,^{70a,70b} F. Fabbri,^{70a} A. Fanfani,^{70a,70b} P. Giacomelli,^{70a} L. Gionni,^{70a,70b} C. Grandi,^{70a} L. Guiducci,^{70a,70b} S. Lo Meo,^{70a,rr} L. Lunerti,^{70a,70b} S. Marcellini,^{70a} G. Masetti,^{70a} F. L. Navarria,^{70a,70b} A. Perrotta,^{70a} F. Primavera,^{70a,70b} A. M. Rossi,^{70a,70b} T. Rovelli,^{70a,70b} G. P. Siroli,^{70a,70b} S. Albergo,^{71a,71b,ss} S. Costa,^{71a,71b,ss} A. Di Mattia,^{71a} R. Potenza,^{71a,71b} A. Tricomi,^{71a,71b,ss} C. Tuve,^{71a,71b} G. Barbagli,^{72a} A. Cassese,^{72a} R. Ceccarelli,^{72a,72b} V. Ciulli,^{72a,72b} C. Civinini,^{72a} R. D'Alessandro,^{72a,72b} E. Focardi,^{72a,72b} G. Latino,^{72a,72b} P. Lenzi,^{72a,72b} M. Lizzo,^{72a,72b} M. Meschini,^{72a} S. Paoletti,^{72a} R. Seidita,^{72a,72b} G. Sguazzoni,^{72a} L. Viliani,^{72a} L. Benussi,⁷³ S. Bianco,⁷³ D. Piccolo,⁷³ M. Bozzo,^{74a,74b} F. Ferro,^{74a} R. Mulargia,^{74a,74b} E. Robutti,^{74a} S. Tosi,^{74a,74b} A. Benaglia,^{75a} G. Boldrini,^{75a} F. Brivio,^{75a,75b} F. Cretorelli,^{75a,75b} F. De Guio,^{75a,75b} M. E. Dinardo,^{75a,75b} P. Dini,^{75a} S. Gennai,^{75a} A. Ghezzi,^{75a,75b} P. Govoni,^{75a,75b} L. Guzzi,^{75a,75b} M. T. Lucchini,^{75a,75b} M. Malberti,^{75a} S. Malvezzi,^{75a} A. Massironi,^{75a} D. Menasce,^{75a} L. Moroni,^{75a} M. Paganoni,^{75a,75b} D. Pedrini,^{75a} B. S. Pinolini,^{75a} S. Ragazzi,^{75a,75b} N. Redaelli,^{75a} T. Tabarelli de Fatis,^{75a,75b} D. Valsecchi,^{75a,75b,u} D. Zuolo,^{75a,75b} S. Buontempo,^{76a} F. Carnevali,^{76a,76b} N. Cavallo,^{76a,76c} A. De Iorio,^{76a,76b} F. Fabozzi,^{76a,76c} A. O. M. Iorio,^{76a,76b} L. Lista,^{76a,76b,tt} S. Meola,^{76a,76d,uu} P. Paolucci,^{76a,u} B. Rossi,^{76a} C. Sciacca,^{76a,76b} P. Azzi,^{77a} N. Bacchetta,^{77a} D. Bisello,^{77a,77b} P. Bortignon,^{77a} A. Bragagnolo,^{77a,77b} R. Carlin,^{77a,77b} P. Checchia,^{77a}

P. Sheldon,¹⁹⁰ S. Tuo,¹⁹⁰ J. Velkovska,¹⁹⁰ M. W. Arenton,¹⁹¹ B. Cox,¹⁹¹ G. Cummings,¹⁹¹ J. Hakala,¹⁹¹ R. Hirosky,¹⁹¹
M. Joyce,¹⁹¹ A. Ledovskoy,¹⁹¹ A. Li,¹⁹¹ C. Neu,¹⁹¹ C. E. Perez Lara,¹⁹¹ B. Tannenwald,¹⁹¹ S. White,¹⁹¹ E. Wolfe,¹⁹¹
N. Poudyal,¹⁹² K. Black,¹⁹³ T. Bose,¹⁹³ C. Caillol,¹⁹³ S. Dasu,¹⁹³ I. De Bruyn,¹⁹³ P. Everaerts,¹⁹³ F. Fienga,¹⁹³ C. Galloni,¹⁹³
H. He,¹⁹³ M. Herndon,¹⁹³ A. Hervé,¹⁹³ U. Hussain,¹⁹³ A. Lanaro,¹⁹³ A. Loeliger,¹⁹³ R. Loveless,¹⁹³
J. Madhusudanan Sreekala,¹⁹³ A. Mallampalli,¹⁹³ A. Mohammadi,¹⁹³ D. Pinna,¹⁹³ A. Savin,¹⁹³ V. Shang,¹⁹³ V. Sharma,¹⁹³
W. H. Smith,¹⁹³ D. Teague,¹⁹³ S. Trembath-Reichert,¹⁹³ and W. Vetens¹⁹³

(CMS Collaboration)

- ¹*Yerevan Physics Institute, Yerevan, Armenia*
- ²*Institut für Hochenergiephysik, Vienna, Austria*
- ³*Institute for Nuclear Problems, Minsk, Belarus*
- ⁴*Universiteit Antwerpen, Antwerpen, Belgium*
- ⁵*Vrije Universiteit Brussel, Brussel, Belgium*
- ⁶*Université Libre de Bruxelles, Bruxelles, Belgium*
- ⁷*Ghent University, Ghent, Belgium*
- ⁸*Université Catholique de Louvain, Louvain-la-Neuve, Belgium*
- ⁹*Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil*
- ¹⁰*Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil*
- ¹¹*Universidade Estadual Paulista, Universidade Federal do ABC, São Paulo, Brazil*
- ¹²*Institute for Nuclear Research and Nuclear Energy, Bulgarian Academy of Sciences, Sofia, Bulgaria*
- ¹³*University of Sofia, Sofia, Bulgaria*
- ¹⁴*Beihang University, Beijing, China*
- ¹⁵*Department of Physics, Tsinghua University, Beijing, China*
- ¹⁶*Institute of High Energy Physics, Beijing, China*
- ¹⁷*State Key Laboratory of Nuclear Physics and Technology, Peking University, Beijing, China*
- ¹⁸*Sun Yat-Sen University, Guangzhou, China*
- ¹⁹*Institute of Modern Physics and Key Laboratory of Nuclear Physics and Ion-beam Application (MOE) - Fudan University, Shanghai, China*
- ²⁰*Zhejiang University, Hangzhou, China, Zhejiang, China*
- ²¹*Universidad de Los Andes, Bogota, Colombia*
- ²²*Universidad de Antioquia, Medellin, Colombia*
- ²³*University of Split, Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, Split, Croatia*
- ²⁴*University of Split, Faculty of Science, Split, Croatia*
- ²⁵*Institute Rudjer Boskovic, Zagreb, Croatia*
- ²⁶*University of Cyprus, Nicosia, Cyprus*
- ²⁷*Charles University, Prague, Czech Republic*
- ²⁸*Escuela Politécnica Nacional, Quito, Ecuador*
- ²⁹*Universidad San Francisco de Quito, Quito, Ecuador*
- ³⁰*Academy of Scientific Research and Technology of the Arab Republic of Egypt, Egyptian Network of High Energy Physics, Cairo, Egypt*
- ³¹*Center for High Energy Physics (CHEP-FU), Fayoum University, El-Fayoum, Egypt*
- ³²*National Institute of Chemical Physics and Biophysics, Tallinn, Estonia*
- ³³*Department of Physics, University of Helsinki, Helsinki, Finland*
- ³⁴*Helsinki Institute of Physics, Helsinki, Finland*
- ³⁵*Lappeenranta University of Technology, Lappeenranta, Finland*
- ³⁶*IRFU, CEA, Université Paris-Saclay, Gif-sur-Yvette, France*
- ³⁷*Laboratoire Leprince-Ringuet, CNRS/IN2P3, Ecole Polytechnique, Institut Polytechnique de Paris, Palaiseau, France*
- ³⁸*Université de Strasbourg, CNRS, IPHC UMR 7178, Strasbourg, France*
- ³⁹*Institut de Physique des 2 Infinis de Lyon (IP2I), Villeurbanne, France*
- ⁴⁰*Georgian Technical University, Tbilisi, Georgia*
- ⁴¹*RWTH Aachen University, I. Physikalisches Institut, Aachen, Germany*
- ⁴²*RWTH Aachen University, III. Physikalisches Institut A, Aachen, Germany*
- ⁴³*RWTH Aachen University, III. Physikalisches Institut B, Aachen, Germany*
- ⁴⁴*Deutsches Elektronen-Synchrotron, Hamburg, Germany*
- ⁴⁵*University of Hamburg, Hamburg, Germany*

- ⁴⁶Karlsruher Institut fuer Technologie, Karlsruhe, Germany
- ⁴⁷Institute of Nuclear and Particle Physics (INPP), NCSR Demokritos, Aghia Paraskevi, Greece
- ⁴⁸National and Kapodistrian University of Athens, Athens, Greece
- ⁴⁹National Technical University of Athens, Athens, Greece
- ⁵⁰University of Ioánnina, Ioánnina, Greece
- ⁵¹MTA-ELTE Lendület CMS Particle and Nuclear Physics Group, Eötvös Loránd University, Budapest, Hungary
- ⁵²Wigner Research Centre for Physics, Budapest, Hungary
- ⁵³Institute of Nuclear Research ATOMKI, Debrecen, Hungary
- ⁵⁴Institute of Physics, University of Debrecen, Debrecen, Hungary
- ⁵⁵Karoly Robert Campus, MATE Institute of Technology, Gyongyos, Hungary
- ⁵⁶Indian Institute of Science (IISc), Bangalore, India
- ⁵⁷National Institute of Science Education and Research, HBNI, Bhubaneswar, India
- ⁵⁸Panjab University, Chandigarh, India
- ⁵⁹University of Delhi, Delhi, India
- ⁶⁰Saha Institute of Nuclear Physics, HBNI, Kolkata, India
- ⁶¹Indian Institute of Technology Madras, Madras, India
- ⁶²Bhabha Atomic Research Centre, Mumbai, India
- ⁶³Tata Institute of Fundamental Research-A, Mumbai, India
- ⁶⁴Tata Institute of Fundamental Research-B, Mumbai, India
- ⁶⁵Indian Institute of Science Education and Research (IISER), Pune, India
- ⁶⁶Isfahan University of Technology, Isfahan, Iran
- ⁶⁷Institute for Research in Fundamental Sciences (IPM), Tehran, Iran
- ⁶⁸University College Dublin, Dublin, Ireland
- ^{69a}INFN Sezione di Bari, Bari, Italy
- ^{69b}Università di Bari, Bari, Italy
- ^{69c}Politecnico di Bari, Bari, Italy
- ^{70a}INFN Sezione di Bologna, Bologna, Italy
- ^{70b}Università di Bologna, Bologna, Italy
- ^{71a}INFN Sezione di Catania, Catania, Italy
- ^{71b}Università di Catania, Catania, Italy
- ^{72a}INFN Sezione di Firenze, Firenze, Italy
- ^{72b}Università di Firenze, Firenze, Italy
- ⁷³INFN Laboratori Nazionali di Frascati, Frascati, Italy
- ^{74a}INFN Sezione di Genova, Genova, Italy
- ^{74b}Università di Genova, Genova, Italy
- ^{75a}INFN Sezione di Milano-Bicocca, Milano, Italy
- ^{75b}Università di Milano-Bicocca, Milano, Italy
- ^{76a}INFN Sezione di Napoli, Napoli, Italy
- ^{76b}Università di Napoli 'Federico II', Napoli, Italy
- ^{76c}Università della Basilicata, Potenza, Italy
- ^{76d}Università G. Marconi, Roma, Italy
- ^{77a}INFN Sezione di Padova, Padova, Italy
- ^{77b}Università di Padova, Padova, Italy
- ^{77c}Università di Trento, Trento, Italy
- ^{78a}INFN Sezione di Pavia, Pavia, Italy
- ^{78b}Università di Pavia, Pavia, Italy
- ^{79a}INFN Sezione di Perugia, Perugia, Italy
- ^{79b}Università di Perugia, Perugia, Italy
- ^{80a}INFN Sezione di Pisa, Pisa, Italy
- ^{80b}Università di Pisa, Pisa, Italy
- ^{80c}Scuola Normale Superiore di Pisa, Pisa, Italy
- ^{80d}Università di Siena, Siena, Italy
- ^{81a}INFN Sezione di Roma, Rome, Italy
- ^{81b}Sapienza Università di Roma, Rome, Italy
- ^{82a}INFN Sezione di Torino, Torino, Italy
- ^{82b}Università di Torino, Torino, Italy
- ^{82c}Università del Piemonte Orientale, Novara, Italy
- ^{83a}INFN Sezione di Trieste, Trieste, Italy
- ^{83b}Università di Trieste, Trieste, Italy

Study of Z boson plus jets events using variables sensitive to double-parton scattering in pp collisions at 13 TeV



The CMS collaboration

E-mail: cms-publication-committee-chair@cern.ch

ABSTRACT: Double-parton scattering is investigated using events with a Z boson and jets. The Z boson is reconstructed using only the dimuon channel. The measurements are performed with proton-proton collision data recorded by the CMS experiment at the LHC at $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of 35.9 fb^{-1} collected in the year 2016. Differential cross sections of $Z + \geq 1$ jet and $Z + \geq 2$ jets are measured with transverse momentum of the jets above 20 GeV and pseudorapidity $|\eta| < 2.4$. Several distributions with sensitivity to double-parton scattering effects are measured as functions of the angle and the transverse momentum imbalance between the Z boson and the jets. The measured distributions are compared with predictions from several event generators with different hadronization models and different parameter settings for multiparton interactions. The measured distributions show a dependence on the hadronization and multiparton interaction simulation parameters, and are important input for future improvements of the simulations.

KEYWORDS: Hadron-Hadron scattering (experiments), Jets

ARXIV EPRINT: [2105.14511](https://arxiv.org/abs/2105.14511)

Contents

1	Introduction	1
2	The CMS detector	3
3	Simulated event samples	3
4	Event selection	5
5	Observables and the unfolding	7
6	Systematic uncertainties	9
7	Results	10
8	Summary	14
	The CMS collaboration	23

1 Introduction

Proton-proton (pp) collisions at high energies result in many events with a Z boson produced in association with jets at large transverse momentum p_T . The Z+jets process is an important background for many standard model (SM) measurements and searches for physics beyond the SM. Measurements of differential cross sections for Z boson production in association with jets can be used to test models of initial-state radiation, and multiparton interactions (MPI). Monte Carlo (MC) models are used to simulate collision processes and play a crucial role in understanding background for the new physics searches. The description of MPI is an important part of these simulations. However, MPI can not be completely described by perturbative quantum chromodynamics (QCD), so this requires a phenomenological description involving parameters that must be tuned with the help of data.

The ATLAS, CMS, and LHCb experiments have reported various measurements of Z+jets processes at centre-of-mass energies of 7 TeV [1–6], 8 TeV [7–9], and 13 TeV [10, 11]. Z boson production in association with jets is a process with a clean experimental signature and is well understood theoretically. In this paper, measurements of Z+jets events are performed to explore observables sensitive to the presence of MPI. A possible contribution could come from the simultaneous occurrence of two parton-parton interactions and is usually, when the involved scale is large, called double-parton scattering (DPS). This analysis explores new signatures sensitive to DPS in events with a Z boson and one or more jets,

The CMS collaboration

Yerevan Physics Institute, Yerevan, Armenia

A. Tumasyan

Institut für Hochenergiephysik, Wien, Austria

W. Adam, J.W. Andrejkovic, T. Bergauer, S. Chatterjee, M. Dragicevic, A. Escalante Del Valle, R. Frühwirth¹, M. Jeitler¹, N. Krammer, L. Lechner, D. Liko, I. Mikulec, P. Paulitsch, F.M. Pitters, J. Schieck¹, R. Schöffbeck, M. Spanring, S. Templ, W. Waltenberger, C.-E. Wulz¹

Institute for Nuclear Problems, Minsk, Belarus

V. Chekhovsky, A. Litomin, V. Makarenko

Universiteit Antwerpen, Antwerpen, Belgium

M.R. Darwish², E.A. De Wolf, X. Janssen, T. Kello³, A. Lelek, H. Rejeb Sfar, P. Van Mechelen, S. Van Putte, N. Van Remortel

Vrije Universiteit Brussel, Brussel, Belgium

F. Blekman, E.S. Bols, J. D'Hondt, J. De Clercq, M. Delcourt, H. El Faham, S. Lowette, S. Moortgat, A. Morton, D. Müller, A.R. Sahasransu, S. Tavernier, W. Van Doninck, P. Van Mulders

Université Libre de Bruxelles, Bruxelles, Belgium

D. Beghin, B. Bilin, B. Clerbaux, G. De Lentdecker, L. Favart, A. Grebenyuk, A.K. Kalsi, K. Lee, M. Mahdavihorrani, I. Makarenko, L. Moureaux, L. Pétré, A. Popov, N. Postiau, E. Starling, L. Thomas, M. Vanden Bemden, C. Vander Velde, P. Vanlaer, D. Vannerom, L. Wezenbeek

Ghent University, Ghent, Belgium

T. Cornelis, D. Dobur, J. Knolle, L. Lambrecht, G. Mestdach, M. Niedziela, C. Roskas, A. Samalan, K. Skovpen, T.T. Tran, M. Tytgat, W. Verbeke, B. Vermassen, M. Vit

Université Catholique de Louvain, Louvain-la-Neuve, Belgium

A. Bethani, G. Bruno, F. Bury, C. Caputo, P. David, C. Delaere, I.S. Donertas, A. Giannamano, K. Jaffel, V. Lemaitre, K. Mondal, J. Prisciandaro, A. Taliencio, M. Teklishyn, P. Vischia, S. Wertz, S. Wuyckens

Centro Brasileiro de Pesquisas Fisicas, Rio de Janeiro, Brazil

G.A. Alves, C. Hensel, A. Moraes

Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil

W.L. Aldá Júnior, M. Alves Gallo Pereira, M. Barroso Ferreira Filho, H. BRANDAO MALBOUISSON, W. Carvalho, J. Chinellato⁴, E.M. Da Costa, G.G. Da Silveira⁵, D. De Jesus Damiao, S. Fonseca De Souza, D. Matos Figueiredo, C. Mora Herrera, K. Mota Amarilo, L. Mundim, H. Nogima, P. Rebello Teles, A. Santoro, S.M. Silva Do Amaral, A. Sznajder, M. Thiel, F. Torres Da Silva De Araujo, A. Vilela Pereira

National Institute of Science Education and Research, HBNI, Bhubaneswar, India

S. Bahinipati³¹, D. Dash, C. Kar, P. Mal, T. Mishra, V.K. Muraleedharan Nair Bindhu³², A. Nayak³², P. Saha, N. Sur, S.K. Swain, D. Vats³²

Panjab University, Chandigarh, India

S. Bansal, S.B. Beri, V. Bhatnagar, G. Chaudhary, S. Chauhan, N. Dhingra³³, R. Gupta, A. Kaur, M. Kaur, S. Kaur, P. Kumari, M. Meena, K. Sandeep, J.B. Singh, A.K. Virdi

University of Delhi, Delhi, India

A. Ahmed, A. Bhardwaj, B.C. Choudhary, R.B. Garg, M. Gola, S. Keshri, A. Kumar, M. Naimuddin, P. Priyanka, K. Ranjan, A. Shah

Saha Institute of Nuclear Physics, HBNI, Kolkata, India

M. Bharti³⁴, R. Bhattacharya, S. Bhattacharya, D. Bhowmik, S. Dutta, S. Dutta, B. Gomber³⁵, M. Maity³⁶, S. Nandan, P. Palit, P.K. Rout, G. Saha, B. Sahu, S. Sarkar, M. Sharan, B. Singh³⁴, S. Thakur³⁴

Indian Institute of Technology Madras, Madras, India

P.K. Behera, S.C. Behera, P. Kalbhor, A. Muhammad, R. Pradhan, P.R. Pujahari, A. Sharma, A.K. Sikdar

Bhabha Atomic Research Centre, Mumbai, India

D. Dutta, V. Jha, V. Kumar, D.K. Mishra, K. Naskar³⁷, P.K. Netrakanti, L.M. Pant, P. Shukla

Tata Institute of Fundamental Research-A, Mumbai, India

T. Aziz, S. Dugad, M. Kumar, U. Sarkar

Tata Institute of Fundamental Research-B, Mumbai, India

S. Banerjee, S. Bhattacharya, R. Chudasama, M. Guchait, S. Karmakar, S. Kumar, G. Majumder, K. Mazumdar, S. Mukherjee

Indian Institute of Science Education and Research (IISER), Pune, India

K. Alpana, S. Dube, B. Kansal, S. Pandey, A. Rane, A. Rastogi, S. Sharma

Department of Physics, Isfahan University of Technology, Isfahan, Iran

H. Bakhshiansohi³⁸, M. Zeinali³⁹

Institute for Research in Fundamental Sciences (IPM), Tehran, Iran

S. Chenarani⁴⁰, S.M. Etesami, M. Khakzad, M. Mohammadi Najafabadi

University College Dublin, Dublin, Ireland

M. Grunewald

INFN Sezione di Bari^a, Università di Bari^b, Politecnico di Bari^c, Bari, Italy

M. Abbrescia^{a,b}, R. Aly^{a,b,41}, C. Aruta^{a,b}, A. Colaleo^a, D. Creanza^{a,c}, N. De Filippis^{a,c}, M. De Palma^{a,b}, A. Di Florio^{a,b}, A. Di Pilato^{a,b}, W. Elmetenawee^{a,b}, L. Fiore^a, A. Gelmi^{a,b}, M. Gul^a, G. Iaselli^{a,c}, M. Ince^{a,b}, S. Lezki^{a,b}, G. Maggi^{a,c}, M. Maggi^a, I. Margjeka^{a,b},

Experimental measurement of Heavy Flavors and jets

Meena Meena^{a,b}

^a*Panjab University,
Chandigarh, India*

^b*On behalf of the ALICE, ATLAS, CMS, and LHCb Collaborations*

E-mail: meena.meena@cern.ch

Recent results of the heavy flavors and jet production in proton-proton collisions using the data collected by the ALICE, ATLAS, CMS, and LHCb detector at the Large Hadron Collider are presented. These measurements are crucial for precise tests of electroweak and next-to-leading-order pQCD calculations, improving the understanding of the proton structure. The measured results are compared with theoretical predictions from several Monte Carlo event generators using different parton shower simulations which provide valuable inputs for tuning of Monte Carlo event generators.

*The Ninth Annual Conference on Large Hadron Collider Physics - LHCP2021
7-12 June 2021
Online*

1. Introduction

In this document, an overview of recent results on heavy flavors (HF) and jet productions and their comparison with various theoretical predictions are presented in the proton-proton collision with the ALICE [1], ATLAS [2], CMS [3], and LHCb [4] experiments.

2. Measurement of the HF dijet and Z + HF jets cross section at the center-of-mass energy ($\sqrt{s}$) of 13 TeV

The first differential and integrated cross section measurements of inclusive $b\bar{b}$ - and $c\bar{c}$ -dijet [5] production are performed with the LHCb detector. The phase space regions are defined by requiring the two jets originating from the two b or c quarks with transverse momentum (p_T) greater than 20 GeV, pseudorapidity (η) in the range $2.2 < \eta < 4.2$, and with a difference in the azimuthal angle between the two jets greater than 1.5. The integrated $b\bar{b}$ - and $c\bar{c}$ -dijet cross sections are measured to be 53.0 ± 0.1 (stat) ± 9.5 (stat+syst) ± 2.1 (lumi) nb, and 73 ± 0.1 (stat) ± 16.1 (stat+syst) ± 2.9 (lumi) nb and the ratio between $c\bar{c}$ - and $b\bar{b}$ -dijet cross section is 1.37 ± 0.01 (stat) ± 0.27 (stat+syst). The differential cross section as a function of leading jet p_T is shown in Figure 1 (left).

The differential and integrated cross section measurements of $Z(\rightarrow ll) + \geq 1$ c jet [6], where $ll = e^+e^-$ or $\mu^-\mu^+$ and cross section ratios of Z+b jet and Z+c jet w.r.t Z + jet (R (b/j), R (c/j)) and Z+c jet w.r.t Z+b jet (R(c/b)) [7] are performed with the CMS detector. The fiducial phase space regions are defined by selecting Z boson with at least one inclusive or b jet or c jet selected with $p_T > 30$ GeV and $|\eta| < 2.4$. The total fiducial measured cross section of Z+c jet to be 405.4 ± 5.6 (stat) ± 24.3 (exp) ± 3.7 (theo) pb. The measured integrated cross section ratio values are R (c/j) = 0.102 ± 0.002 (stat) ± 0.009 (syst), R (b/j) = 0.0633 ± 0.0004 (stat) ± 0.0015 (syst), and R (c/b) = 1.62 ± 0.03 (stat) ± 0.15 (syst). The predictions for Z+b and Z+c jet production can be derived in either a 4-flavor number scheme (4FNS) or a 5-flavor number scheme (5FNS). In 5FNS, b quark density is allowed in the initial state via a b quark parton distribution functions (PDFs) of proton but not in 4FNS. The measured integrated cross section of Z+c jet, R (c/j), and R (b/j), $b\bar{b}$ - and $c\bar{c}$ -dijet are overestimated by MG5_aMC next-to-leading order (NLO) 5FNS, except R (c/b). The Z+c jet, R (c/j), and R (b/j) are well described by the MG5_aMC LO 5FNS and $b\bar{b}$ - and $c\bar{c}$ -dijet by PYTHIA8 LO but Z+c jet overestimated by the SHERPA NLO 5FNS. The R (c/j) and R (c/b) are underestimated and R (b/j) is overestimated by the MCFM predictions at NLO and LO.

A measurement of the inclusive and differential production cross sections of a $Z(\rightarrow ll) + \geq 1$ b jet and $Z(\rightarrow ll) + \geq 2$ b jets [8] are performed with the ATLAS detector. The measured inclusive cross sections for $Z + \geq 1$ b jet and $Z + \geq 2$ b jets for jet $p_T > 20$ GeV and rapidity $|y| < 2.5$ as shown in Figure 1 (middle and right), are 10.90 ± 0.03 (stat) ± 1.08 (syst) ± 0.25 (lumi) pb and 1.32 ± 0.01 (stat) ± 0.21 (syst) ± 0.04 (lumi) pb, respectively. The 5FNS simulations adequately predict the inclusive cross sections for both $Z + \geq 1$ b jet & $Z + \geq 2$ b jets. All 4FNS MC predictions are systematically lower than data for $Z + \geq 1$ b jet but agree well for $Z + \geq 2$ b jets except Alpgen+Py6 LO 4FNS which underestimates the data. The use of the NNPDF3.0lo PDF set in Alpgen predictions gives better agreement with data because of a higher acceptance in the fiducial region. The SHERPA fusing 4FNS+5FNS at NLO simulation agrees with Sherpa 5FNS NLO, showing that the effects of merging are minor.

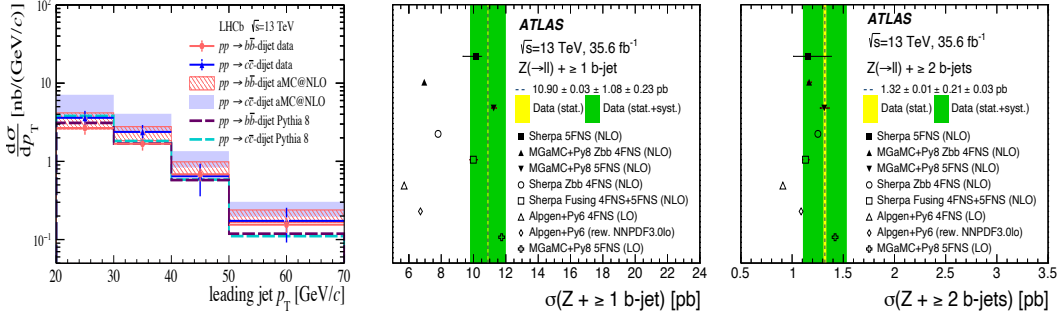


Figure 1: Differential $b\bar{b}$ - and $c\bar{c}$ -dijet cross section as a function of leading jet p_T and error bars represent the total uncertainties [5] (left). The integrated cross sections for $Z + \geq 1$ b jet (middle) and $Z + \geq 2$ b jets (right) [8].

3. Study of J/ψ and $\psi(2s)$ meson production inside jets

The first study of prompt $J/\psi \rightarrow \mu^+\mu^-$ mesons contained in jets produced in the gluon fragmentation dominated central rapidity region ($|\eta(\text{jet})| < 1$) with $p_T(\text{jet}) > 25$ GeV, is performed based on the fragmenting jet function (FJF) approach with the CMS detector at $\sqrt{s} = 8$ TeV [9]. The FJF analysis uses the non-relativistic quantum chromodynamics (NRQCD) approach to compute the cross section for the formation of a J/ψ meson from the $c\bar{c}$ system. The shape of measured differential cross section for J/ψ meson production as a jet constituent (z) is compared to the FJF prediction, using three different long-distance matrix element (LDME) parameter sets the Bodwin, Chung, Kim, and Lee (BCKL), Butenschoen, and Kniehl (BK), and Chao. The data agree best with FJF BCKL calculation that use an LDME parameter set in which prompt J/ψ mesons are predicted to be unpolarized for all three z ranges (0.40–0.45, 0.50–0.55, and 0.60–0.65). In contrast, the FJF BK and Chao predictions disagree with the data for all three z ranges. The FJF BCKL not only describes the production of high- p_T J/ψ mesons as constituents of jets but also predicts small J/ψ meson polarization.

The first measurement $J/\psi \rightarrow \mu^+\mu^-$ mesons produced in jets promptly and in b-hadron decay is studied in the forward region ($2.5 < \eta(\text{jet}) < 4.0$) with $p_T(\text{jet}) > 20$ GeV with the LHCb detector at $\sqrt{s} = 13$ TeV [10]. Fraction of J/ψ mesons produced in b-hadron decays is in the range 20–60%, depending on $z(J/\psi)$ and jet p_T . The observed normalized cross section distribution for J/ψ mesons produced in b-hadron decays is consistent with the PYTHIA8 prediction; however, the prompt- J/ψ results do not agree with the LO NRQCD-based predictions as implemented in PYTHIA8 as shown in Figure 2 (left). At small $z(J/\psi)$, PYTHIA8 predicts that most of the jet arises from a parton-parton scatter other than the one that produced the J/ψ meson.

The measurements of the prompt and non-prompt double differential cross sections of J/ψ and $\psi(2s)$ mesons with jet p_T between 60 and 360 GeV for J/ψ and 60 and 140 GeV for $\psi(2s)$ and $|\eta| < 2$ are performed using data recorded by the ATLAS detector at $\sqrt{s} = 13$ TeV between 2015 and 2018 [11]. Furthermore, the measurements of the non-prompt fractions of $\psi(2s)$ and J/ψ , and prompt and non-prompt production ratios of $\psi(2s)$ to J/ψ are presented. In J/ψ and $\psi(2s)$ both cases, p_T range goes well beyond the values reached so far, which may help discriminate various

theoretical models. The results show similar p_T -dependence for prompt and non-prompt double differential cross sections, with non-prompt fractions close to constant $\sim 70\%$ for both J/ψ and $\psi(2s)$. The prediction with the FONLL model is consistent with data for non-prompt production at the low end of the p_T range but exceeds the experimental values at large p_T as shown in Figure 2 (middle).

4. Groomed jet substructure measurements of charm jets tagged with D^0 mesons

This is the first measurement of groomed jet substructure observables for jets containing a fully reconstructed D^0 meson at $\sqrt{s} = 13$ TeV performed with the charged jet in $15 \leq p_T < 30$ GeV range with the ALICE detector. The Soft Drop grooming procedure is used to unwind the reclustered jets, starting from the final step until a fully reconstructed D^0 meson candidate is found or all splittings have been exhausted. The momentum fraction of groomed two-prong jet substructure (z_g), the aperture angle between these prongs (R_g), and the number of splittings satisfying the grooming condition (n_{SD}) are the jet substructure observables. In particular, n_{SD} shifted to a smaller value for the D^0 -tagged jets shows significant sensitivity to differences in the D^0 -tagged and inclusive-jets samples as shown in Figure 2 (right). This indicates that the fragmentation of the c quark is hard throughout the showering procedure. The PYTHIA8 simulations with Monash tune calculation describing the data qualitatively well for all measured observables.

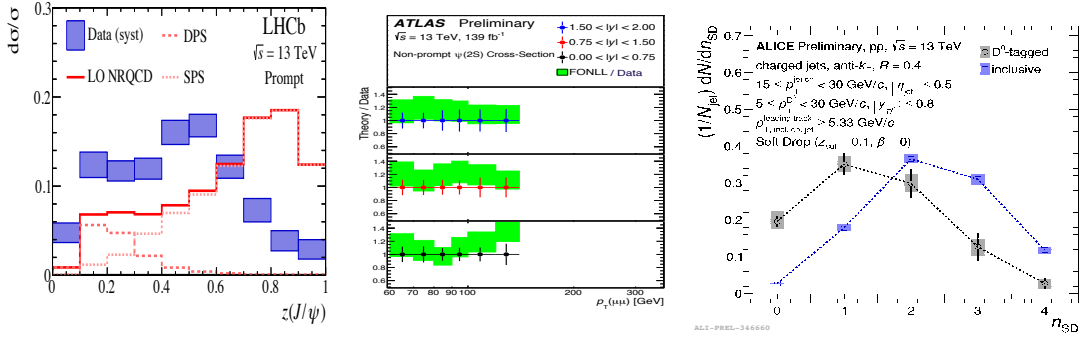


Figure 2: Normalized $z(J/\psi)$ distributions for J/ψ mesons produced in b-hadron decay [10] (left). The ratios of the FONLL prediction to the measured differential cross sections are shown for non-prompt $\psi(2s)$ mesons [11] (middle). The measured n_{SD} distributions, corrected to particle level, of prompt D^0 -tagged and inclusive jets are shown (right) [12] (right).

5. Summary

The ALICE, ATLAS, CMS, and LHCb experiments have a rich program on the measurements related to heavy flavors and jet. Here, we presented an overview of many recent measurements which are sensitive to different theoretical approaches based on perturbative and non-perturbative methods and b and c quark PDFs in different kinematic regimes of jets production.

References

- [1] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, [JINST 3 \(2008\) S08002](#).
- [2] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, [JINST 3 \(2008\) S08003](#)
- [3] CMS Collaboration, *The CMS Experiment at the CERN LHC*, [JINST 3 \(2008\) S08004](#)
- [4] LHCb Collaboration, *The LHCb Detector at the LHC*, [JINST 3 \(2008\) S08005](#)
- [5] LHCb Collaboration, *Measurement of differential $b\bar{b}$ - and $c\bar{c}$ -dijet cross-sections in the forward region of pp collisions at $\sqrt{s} = 13$ TeV*, [JHEP 02 \(2021\) 023](#)
- [6] CMS Collaboration, *Measurement of differential cross sections for Z bosons produced in association with charm jets in pp collisions at $\sqrt{s} = 13$ TeV*, [JHEP 04 \(2021\)](#)
- [7] CMS Collaboration, *Measurement of the associated production of a Z boson with charm or bottom quark jets in proton-proton collisions at $\sqrt{s} = 13$ TeV*, [Phys. Rev. D 102 \(2020\)](#)
- [8] ATLAS Collaboration, *Measurements of the production cross-section for a Z boson in association with b -jets in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, [JHEP 07 \(2020\) 044](#)
- [9] CMS Collaboration, *Study of J/ψ meson production inside jets in pp collisions at $\sqrt{s} = 8$ TeV*, [Phys. Lett. B 804 \(2020\) 135409](#)
- [10] LHCb Collaboration, *Study of J/ψ production in jets*, [Phys. Rev. Lett. 118 \(2017\) 19](#)
- [11] ATLAS Collaboration, *Measurement of the production cross-section of J/ψ and $\psi(2S)$ mesons at high transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, [ATLAS-CONF-2019-047](#)
- [12] ALICE Collaboration, *Groomed jet substructure measurements of charm jets tagged with D^0 mesons in pp collisions at $\sqrt{s} = 13$ TeV*, [ALICE-PUBLIC-2020-002](#)

Fabrication and Characterization of Gaseous Detector for the identification of High Energy Particles.

J.Babbar¹, S.Bansal², S.B.Beri¹, V.Bhatnagar¹, S.Chauhan¹,
R.Gupta¹, A.Kaur¹, H.kaur¹, S.Kumar¹, P.Kumari¹, M.Meena¹,
J.B.Singh¹, A.K.Virdi¹

¹ Department of Physics, Panjab University, Chandigarh, India

² UIET, Panjab University Chandigarh, India

E-mail: meenaluhach91@gmail.com

Abstract. There is a vast range of gases which get ionized, produce electron-ion pairs, on the passage of high energy charged particles. Such gases are extensively used in experiments such as LHC, Belle-II, RHIC, DAFNE, etc which produce high energy particles. Panjab University established a detector assembly and characterization laboratory dedicated to gaseous detectors such as Resistive Plate Chamber (RPC) and Gas Electron Multiplier (GEM) for the LHC experiment. Here, we present the recent work on the fabrication and characterizations of the GEM detector at Panjab University.

1. Introduction

Gaseous detectors generally have two electrodes separated by the gas medium in which primary charged particles interact and produced electron-ion pairs. The GEM detector is one of the most important Micro-Pattern Gaseous Detector (MPGD) first introduced in 1997 by F. Sauli [1] at the European Organization for Nuclear Research (CERN). It records the ionization produced by charged particles in the gaseous medium consists of three GEM foils embedded between the drift board (acting as the cathode) and the readout board (acting as the anode) [2]. The GEM detector abundantly used in High-Energy Physics (HEP) experiments like CMS [3], LHCb [4], TOTEM [5], COMPASS [6] at Large Hadron Collider (LHC), Belle II [7], etc, and in medical diagnostics [8] as well as in astronomy due to their high rate detection [9] capability up to 100 MHz/cm² with efficiency above 97%, an excellent spatial resolution of $\sim 340 \mu m$, an excellent time resolutions [10] of ~ 8 ns. In most of the experiments to achieve high gain a high voltage supposed to apply across the detector which leads to an electrical breakdown problem. This problem can be avoided by using multi-layer GEM foils placed at a relative distance of a few mm in a suitable gas mixture. In this case, amplification is shared over different gem foils [11], with a relatively low voltage on each GEM foil. Panjab University has been collaborated with CMS [12] and build high-energy physics laboratory to assemble and test [13] triple-GEM detectors (referred to GE1/1 detector) that are installed in the 1st ring in 1st muon station in both end-caps of the CMS detector. The GEM foils and other components of production chambers are delivered from CERN.



2. Description and working principle the GEM detector

The GEM detector works in a gas medium in which gas ionization takes place due to the acceleration of electrons under the influence of the electric field. Since among noble gas Argon (Ar) is having high specific ionization and is less expensive. It can cause avalanche creation beyond the limits which leads to sparks and cause permanent damage to the detector. There is a need for a small amount of quenching gas which can be used to reduce the spark probability and to stop fake avalanche due to the emission of UV photons. The Carbon-dioxide (CO₂) has the properties of electro-negativity, it attracts electrons and attached to its molecules. A lot of studies have been performed on the selection of gas mixture to be used for this detector [14]. The gas mixtures of Ar and CO₂ are used in proportions of 70:30. Basically, a triple GEM gaseous detector is made by a drift cathode, three GEM foils, and the Printed Circuit Board (PCB) anode (or readout board). Readout board has two gas connectors located at two diagonal corners, one for gas inlet and another one for gas outlet. GEM foils are made of Kapton of 50 μm each coated with a copper layer of 5 μm on each side. The Kapton is a good insulator which remains stable in a very wide range of temperature. Each foil has equally spaced small holes perforated by chemical etching in a hexagonal pattern with 50 μm internal and 70 μm outer diameter and a pitch of 140 μm as shown in Figure 1 (Left) where the signal gets amplified as shown in Figure 1 (right). The cms triple GEM production chamber has a 3/1/2/1 mm (drift/transfer1/transfer2/induction) gas gap configuration as shown in Figure 2 which is enclosed between two electrodes with an electric field applied between them. The primary electrons produced ionization in the gap between the cathode and the first GEM foil (drift gap). In transfer region electron-ion pairs are attracted by electric field applied on GEM foil where they get multiplied in the hole. After crossing the last GEM foil they drift to the anode in the induction gap where the signal is registered by the readout strips.

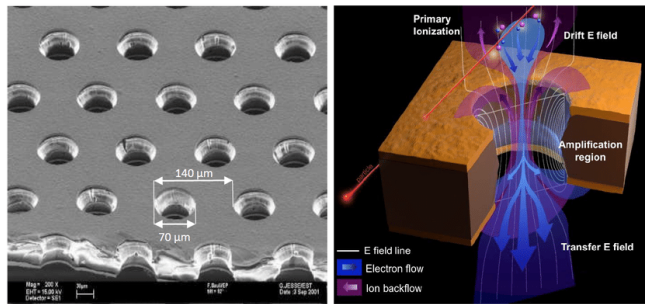


Figure 1. Left: Scanning Electron Microscope (SEM) picture of a GEM foil

Right : schematic view of the electric field lines (white), electron flow (blue), and ion flow (purple) through a bi-conical GEM hole

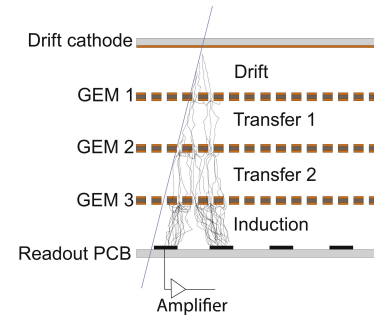


Figure 2. Sketch of a Triple-GEM chamber. Free electrons coming from ionization processes undergo three stages of amplifications.

3. Assembly and quality control testing

After receiving a kit of the different parts of the production chamber from CERN, they inspected very carefully for any possible damages that happened during transport. After that next step is the assembly of the different parts of the GE1/1 detector. Many quality control(QC) test have been performed during and after the assembly such as current leakage in foil, gas leak rate of the detector, noise, effective gas gain, etc.

3.1. Pre-assembly and assembly

3.1.1. Pre-assembly : The drift board and readout are cleaned using a vacuum cleaner. All the screws and pull-outs are cleaned with an ultrasonic bath to remove any dust. Then the drift board is prepared by placing high voltage pins as shown in Figure 3 (Left), mounting resistor and capacitor as shown in Figure 3 (Middle), and pull-outs as shown in Figure 3 (Right). Spacers are prepared by fixing the brass-inserts using pressure and are used for creating gap between the foils. Readout board screwed gently with the gas connector on the readout board as shown in Figure 4 (Left).

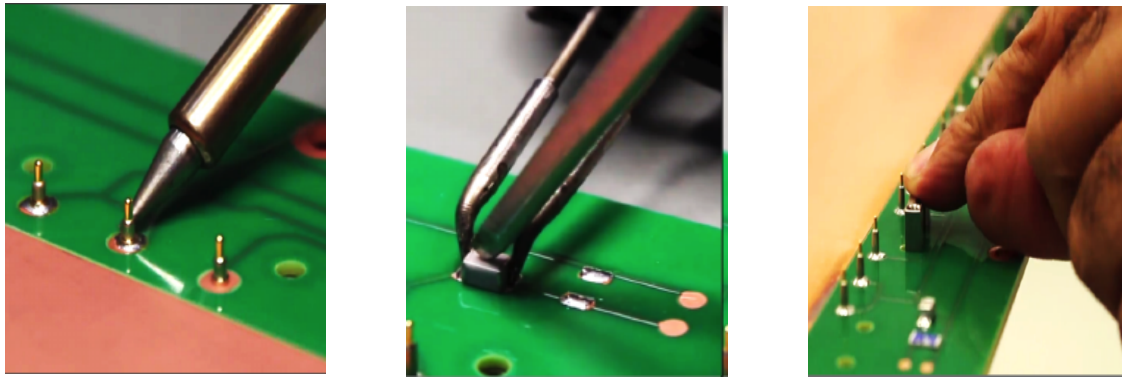


Figure 3. Left: Soldering of the HV pins, Middle : Soldering of the SMD components, Right: Mounting of the pull-outs

3.1.2. Assembly: A class 100 clean room is built at PU site for the assembly. All foils are stretched in the frames and cleaned gently by using an anti-static sticky roller over the foils on both sides in order to remove the dust particles as shown in Figure 4 (middle). Then foils are tested with the MegaOhm/insulation meter by applying 550V voltage to foil by putting a negative pin on the top and positive pin on bottom high voltage (HV) pads of the foils for several minutes. The foil is accepted if the resistance of foil reaches 10 G Ω after few seconds with relative humidity less than 40%. During assembly first cleaned plexiglass is placed then 3 mm clean spacer is inserted on the guiding pins. After that GEM1 foil is placed then 1 mm spacer as shown in Figure 4 (Right). Then in a similar way, GEM2 foil, 2 mm spacer, GEM foil3, 1 mm spacer are placed to form a stack called GEM foil stack as shown in Figure 5 (Left). Plexiglass is placed on the top of the GEM foil stack using guiding pins and then uses screws to the protect GEM stack. Then the excess amounts of Kapton foils are removed using a sharp blade and detached stack from bottom plexiglass and placed it on the drift board. The chamber moved to the assembly jig and clamp the drift board to the assembly table using aluminum bars as shown in Figure 5 (Middle). After placing all the screws in place, the plexiglass and aluminum bar is removed. After that, the GEM foils and all gaps are tested at 550V and the foils are discharged after the test. To test the induction gap, the Panasonic-to-Lemo adapters are connected to one of the readout sectors and 550 V applied between GEM3 bottom and signal pad of the Panasonic-to-Lemo adapter, keeping the chamber in a vertical position as shown in Figure 5 (Right). The impedance should immediately reach 100 G Ω . All QCs results reported in this paper are obtained from a GEM detector having the ID “GE1/1.X.S.INDIA-00YY”(where ‘YY’ stand for ‘02’, ‘07’, ‘08’, ‘09’, ‘10’, ‘15’, ‘16’, ‘17’).



Figure 4. Left: Fixing the gas connector onto the readout board, Middle : Cleaning of the GEM foil with the anti-static roller, Right : Placing the internal frame after testing.

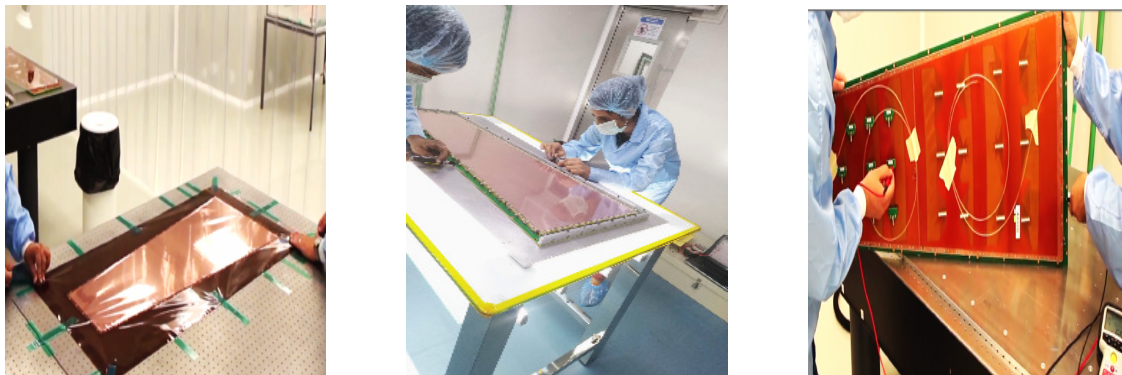


Figure 5. Left : Assembled GEM stack, Middle : Finalizing the stretching of the stack, Right : Assemble detector and testing the induction gap

3.2. Quality control testing

3.2.1. Post-assembly QC2 test: To check again GEM foils after assembly if they are clean or not by measuring the leakage current of the GEM foils and gas gap. The leakage current is that current which flows from the top surface of the foil to the bottom due to the surface conductivity. Due to deposition of some contaminants on the surface of foil in the form of moisture or dust particles will increase this leakage. The detector is accepted if the impedance of each foil and gap is above 10 G Ω and 100 G Ω respectively and the number of sparks are zero after every minute over a period of 10 minutes. Similarly, an impedance also measured over all 24 Panasonic connectors and it should be greater than 100 G Ω . If impedance does not fulfill these criteria then there are might be some short circuit and in that case, the detectors have to be reopened and fixed the issue in the clean-room. All the 8 detectors assembled at the Panjab University site qualified the QC2-Leakage current test.

3.2.2. QC3 (Gas leak) test: If the first test goes fine, then we proceed for the QC3 test which is Gas leak test to identify the gas leak rate of the detector by monitoring the drop of the internal over-pressure as a function of the time. A typical QC3 system has 2 flow-meters, a micro-controller based Arduino chip having 3 sensors for atmospheric pressure, internal gas pressure, and temperature measurements stand prepared at the Panjab University is shown in Figure 6 (Left). The detector is connected in between the inlet and outlet of the calibrated gas system

and filled with CO_2 at the rate of 5 L/h. The chambers cannot sustain internal over-pressure higher than 40-50 hPa. All chambers, except our first production chamber, were tested with an initial over-pressure of 25 hPa, and its internal pressure is monitored for 1 hour. The internal gas pressure is parametrized by $P(t)=P_0\exp(-t/\tau)$, where P_0 is the initial internal over-pressure [15]. The QC3 acceptance limit is $\tau \geq 3.04$ hour, corresponding to a maximum acceptable gas leak rate of about 7 hPa/hr. All the detectors successfully passed the QC3 gas leak test because gas leak reduction rate is less than 7 hPa/hr as shown in Figure 6 (Right).

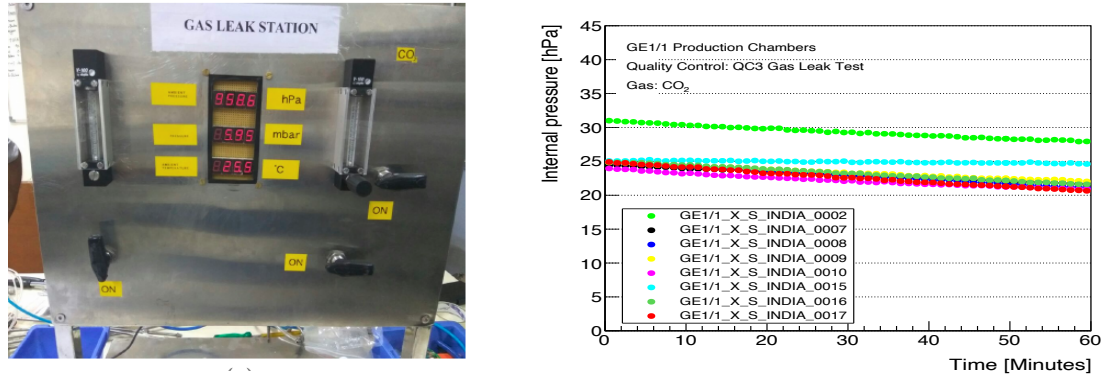


Figure 6. Left : Front view of QC3 Gas Leak Test Stand, Right : Results from the QC3 Gas Leak Test on the GE1/1 detectors assembled at Panjab University.

3.2.3. QC4 (High Voltage) test: After the QC3 test, the QC4 test is performed which is a high voltage test to determine the voltage vs. current curve of the detector and to identify possible malfunctions, defects in the HV circuit, and intrinsic noise rate in GEM detector. The detector is shielded from the background noise with a ground shielding placed on the top of the readout board. The GE1/1 detector under this test is first flushed with pure CO_2 gas at least for 5 hours at the rate of 5 L/h. The high voltage (VSET) is provided by programmable HV power supply (CAEN SY5527) which allows the user to control the current limit (ISET), the steps to ramp up and down the voltage, the maximum voltage, and the trip time. A high voltage low pass filter was placed between the power supply and detector to reduce the electronics noise (by rejecting the high frequency noise from the power supply) so that a stable voltage can be provided to the electrodes [16]. The signal generated at the anode is processed using charged sensitive pre-amplifier connected to the bottom electrode of the third GEM foil through a decoupling RC circuit. Further to amplify and record the signal obtained from readout strips signal sent to timing filter amplifier (TFA). The output of TFA is fed into the discriminator to convert it into a rectangular digital pulse which further passes through a dual timer which generates a veto signal to avoid the fake counts. Finally, the resulting digital pulse sent to a scaler and counter unit for the rate measurement.

The discriminator threshold and the scaler clock are set to -140 mV and 60 s respectively. The total resistance ($R_{\text{multimeter}}$) of the HV circuit (divider plus HV filters) is measured using a multimeter and is used to calculate the current in the divider for every high voltage points. The current limit ISET of the power supply is set according to the expected target value plus 5 μA , and the corresponding high voltage VSET is switched on. After the operating current and voltage get stable, the values of actual voltage V_{mon} , the current I_{mon} , and the intrinsic noise rate of the detector are recorded. The intrinsic noise rate is defined as the rate of signals not arising due to ionization of the gas. The number of counts are measured using the scaler unit

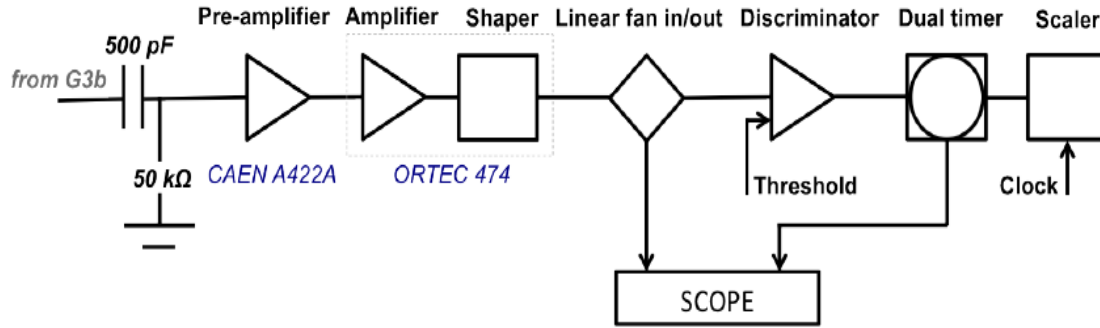


Figure 7. QC4 HV test experimental setup.

for interval of 60 s. These steps are repeated until the voltage reaches 4900 V on the detector. The detector is ramped up to 3000 V in the step of 200 V and up to 4900 V in the step of 100 V. The I-V curve is plotted, which must be a linear, and is fitted by a straight line to determine a resistance (“fitted resistance”). The reciprocal of its slope gives the resistance (R_{IV}) of the HV distribution system. This calculated resistance R_{IV} is then compared with the value of measured resistance $R_{multimeter}$. The detector is said to pass the test if the deviation in the resistance between $R_{multimeter}$ and R_{IV} is $\leq 2\%$ because the deviation greater than 2% indicates a loss of linearity. The maximum intrinsic noise rate should be ≤ 100 Hz at the value of current $\sim 999 \mu A$ and this rate is below 10^{-2} Hz/cm², which is a negligible value with respect to the 4.5 kHz/cm² background rate expected in CMS in standard operating conditions. Table 1 shows the value of measured resistance ($R_{multimeter}$) by multimeter, resistance (R_{IV}) by IV curve, and the maximum intrinsic noise rate (at the value of current $\sim 999 \mu A$) for all the detectors assembled at Panjab University. Figure 8 (Left) shows the monitored voltage (V_{mon}) as a function of the monitored current (I_{mon}) and Figure 8 (Right) shows intrinsic noise rate as a function of monitored current for different GE1/1 production chambers assembled at Panjab University site. All the detectors successfully passed the QC4 HV test as deviation in resistance is $< 2\%$ and maximum intrinsic noise rate is 11.88 Hz which is very less than 100 Hz. Some data points of 8 chambers are overlapped, as the I-V relation of each chamber is similar to each other.

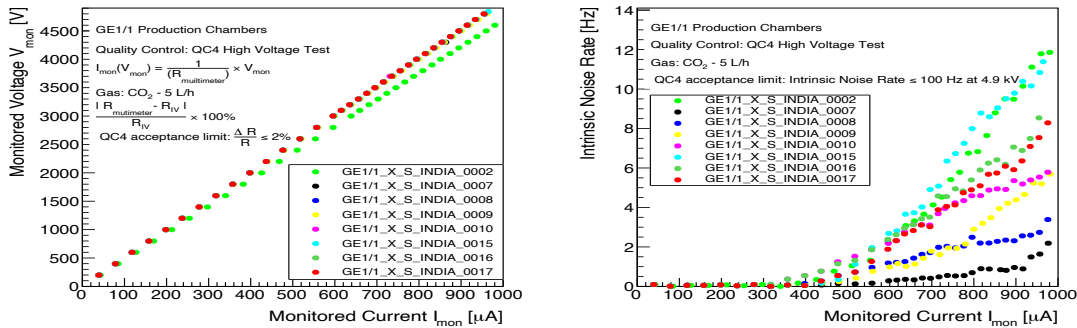


Figure 8. Left : Results from QC4 HV test, Right : Intrinsic Noise Rate as a function of monitored current for different assembled GE1/1 production chambers.

Table 1. Results from QC4 HV test showing the value of measured resistance ($R_{\text{multimeter}}$) by multimeter, resistance (R_{IV}) by IV curve, and the maximum intrinsic noise rate (at the value of current $\sim 999 \mu A$) for all the detectors assembled at Panjab University.

Detector ID	Resistance by Multimeter (R_{IV})(M Ω)	Resistance by I-V curve ($R_{\text{multimeter}}$) (M Ω)	Intrinsic Noise Rate(HZ)
GE1/1_X_S.INDIA.0002	4.67	4.69	11.88
GE1/1_X_S.INDIA.0007	5.00	5.01	3.50
GE1/1_X_S.INDIA.0008	5.00	5.02	6.57
GE1/1_X_S.INDIA.0009	5.00	5.01	5.62
GE1/1_X_S.INDIA.0010	4.97	5.02	5.72
GE1/1_X_S.INDIA.0015	4.96	5.01	11.30
GE1/1_X_S.INDIA.0016	4.97	5.02	8.58
GE1/1_X_S.INDIA.0017	5.00	5.02	8.25

4. Conclusion

All the QCs test have been performed at Panjab University and the results are consistent as recommended by the CMS GEM collaboration. In total, 8 production chambers have been assembled and tested at PU site and shipped to CERN, Geneva and successfully installed in the CMS detector. In future, Panjab university group will be receiving more kits of GEM detector for assembly and testing.

Acknowledgments

We would like to acknowledge the funding agency Council of Scientific and Industrial Research (CSIR), Department of Science and Technology (DST), and University Grants Commission (UGC) for providing the financial support to accomplish this work. We would also like to acknowledge the CMS GEM collaboration to provide the GEM detector, helps us to gain expertise in QC testing as well as to calibrate QCs set-ups.

References

- [1] F. Sauli 2016 The gas electron multiplier (GEM): operating principles and applications *Nucl. Instrum. Meth. A* **805** p 2.
- [2] Mohit Gola, Shivali Malhotra, Aashaq Shah, Asar Ahmed, Ashok Kumar and Md. Naimuddin 2020 Performance of the triple GEM detector built using commercially manufactured GEM foils in India *Nucl. Instrum. Meth. A* 951 162967
- [3] CMS Collaboration 2015 *CMS Technical Design Report for the Muon Endcap GEM Upgrade* CMS-TDR-013
- [4] A. Cardini, B. Bencivenni and P. De Simone 2012 *The operational experience of the triple-GEM detectors of the LHCb muon system: summary of 2 years of data taking* IEEE Nucl. Sci. Symp. Med. Imag. Conf. Rec. pp 759–762
- [5] G. Anelli et al. 2008 The TOTEM experiment at the CERN Large Hadron Collider *JINST* **3** S08007
- [6] COMPASS collaboration 2007 The COMPASS experiment at CERN *Nucl. Instrum. Meth. A* **577** pp 455–518
- [7] Z. Dolezal and S. Uno 2010 *Belle II Technical Design Report*
- [8] M. Bucciantonio and F. Sauli 2015 Proton computed tomography *Mod. Phys. Lett. A* **30** 1540024
- [9] Bachmann S., Kappler S., Ketzer B., Muller T., Ropelewski L., Sauli F. and Schulte E. 2002 *Nucl. Instrum. Meth. A* **478** pp 104–108
- [10] R. Alon et al. 2008 Time resolution of a Thick Gas Electron Multiplier (THGEM) - based detector *Journal of Instrumentation* **3** p 11001
- [11] F. Sauli 1997 GEM: a new concept for electron amplification in gas detectors *Nucl. Instrum. Meth. A* **386** p 531

- [12] CMS Muon Group Performance of prototype GE1/1 chambers for the CMS muon spectrometer upgrade (*Preprint* 1903.02186)
- [13] Manisha, A. Kaur, M. Meena, P. Kumari, R. Gupta, J. S. Shahi, S. Bansal, V. Bhatnagar and J. B. Singh 2017 *Proc. of the DAE Symp. on Nuclear Physics Quality control testing of GEM detector* vol 62
- [14] A. Buzulutskov et al. 1999 GEM operation in pure noble gases and the avalanche confinement *Nucl. Instrum. Meth. A* **433** p 471
- [15] Rosamaria Venditti 2020 *Production, Quality Control and Performance of GE1/1 Detectors for the CMS Upgrade* J. Phys.: Conf. Ser. **1498** 012055
- [16] A. Ahmed, M. Gola, A. Kumar and Md. Naimuddin. 2020 Development and qualification of triple-GEM detector built with large size single mask foils produced in India *JINST* **15**



ICIIE-2020

International Conference on Integrated Interdisciplinary Innovations in Engineering
(August 28-30, 2020)



Organized under twinning arrangement by

University Institute of Engineering and Technology, Panjab University, Chandigarh
and
Government College of Engineering and Technology, Jammu

Supported by :



TE@IP

This is to certified that Ms. Meena of Panjab University, Chandigarh has presented a paper entitled “Fabrication and characterization of gaseous detector for the identification of high energy particles” online in the ICIIE-2020 held at Panjab University, Chandigarh.

Minto Rattan

(Dr. Minto Rattan)
ICIIE-2020 Coordinator

Savita Gupta

(Prof. Savita Gupta)
Director, UIET Chandigarh

Certificate

Meena Meena



Prof. Bedangadas Mohanty
(Chair)

has successfully participated and presented a talk titled

Study for vector bosons production in association with
heavy-flavor jets in proton-proton collisions

XXIV DAE-BRNS Symposium on High Energy Physics

held **ONLINE** by the

National Institute of Science Education and Research

Jatni 752050 Odisha India

during

December 14-18, 2020



Prof. Sanjay Swain
(Co-Chair)

Sponsors & Organizers



IISER
BERHAMPUR

