

# Measurement of the N-jettiness variables in the production of Z boson events with the CMS detector and performance of its electromagnetic calorimeter

*Mesure des variables de N-jettiness dans la production  
d'événements de bosons Z avec le détecteur CMS et  
performance de son calorimètre électromagnétique*

## Thèse de doctorat de l'université Paris-Saclay et de l'université du Monténégro

Ecole doctorale n°576 : Particules, hadrons, énergie et noyau :  
instrumentation, imagerie, cosmos et simulation (PHENICS)  
Spécialité de doctorat : Physique des particules  
Graduate School : Physique, Référent : Faculté des sciences d'Orsay

Thèse préparée dans les unités de recherche **Département de Physique des  
Particules (Université Paris-Saclay, CEA) et Physics department  
(University of Montenegro)** sous la direction de **Federico FERRI**,  
ingénieur-chercheur et **Nataša RAICEVIĆ**, professeure, le co-encadrement de  
**Philippe GRAS**, ingénieur-chercheur

Thèse soutenue à Podgorica, Monténégro, le 07 Juin 2022, par

**Jelena MIJUŠKOVIĆ**

### Composition du jury

**Ivana PIĆURIĆ**

Professeure, University of Montenegro

**Ulla BLUMENSCHEN**

Maîtresse de Conférences, Queen Mary University  
of London

**Marco DELMASTRO**

Directeur de recherche, LAPP Annecy

**Simone ALIOLI**

Professeur Associé, University of Milano-Bicocca

**Lydia ICONOMIDOU-FAYARD**

Directrice de recherche, CNRS, IJCLab Université  
Paris-Saclay

**Federico FERRI**

Ingénieur-chercheur, CEA IRFU

Présidente

Rapporteuse & Examinatrice

Rapporteur & Examineur

Examineur

Examinatrice

Directeur de thèse



**Titre:** Mesure des variables de N-jettiness dans la production d'événements de bosons Z avec le détecteur CMS et performance de son calorimètre électromagnétique

**Mots clés:** Jets, Bosons Z, Grand collisionneur de hadrons, Modèle standard

**Résumé:** La mesure de sections efficaces différentielles de la production de bosons Z en association avec des jets dans les collisions proton-proton à l'énergie du centre de masse de 13 TeV est présentée dans cette thèse. Les données analysées ont été collectées par l'expérience CMS du LHC au cours de l'année 2018. Elles correspondent à une luminosité intégrée de  $59 \text{ fb}^{-1}$ . La section efficace est mesurée en fonction de variables de forme d'événement (event shape variable) utilisant les traces des particules, la zero-jettiness, l'un-jettiness, la somme des moments transversaux des particules, ainsi des variables de forme d'événement utilisant les jets :  $\tau_{\text{sum}}$  et  $\tau_{\text{max}}$ . Les variables de forme d'événement sont mesurées sur des événements avec une paire de muons produite par la désintégration d'un boson Z réel de masse invariante comprise entre 76 et 106 GeV. Elles sont également mesurées pour un boson Z virtuel pour différents intervalles de masse, entre 125 et 150 GeV, 150 et 350 GeV, 350 et 1500 GeV, et

dans quatre régions de moment transversal de bosons Z différentes. Les mesures sont comparées avec trois prédictions théoriques, avec des précisions QCD LO, NLO et NNLO, obtenues avec deux générateurs Monte Carlo, MADGRAPH5\_amc@NLO et GENEVA.

La seconde partie de la thèse est consacrée à l'étude des performances et de l'intercalibration du calorimètre électromagnétique (ECAL) pendant la période de prise de données Run 2 (2016, 2017 et 2018). Ce sous-détecteur est crucial pour la mesure des photons et des électrons et donc très important pour de nombreuses recherches à CMS. Grâce à un monitoring et un étalonnage constants, d'excellentes performances ont été obtenues. La résolution pour le Run 2 est de 1,7% dans la région à bas pseudorapidité. Il a également été démontré que les performances du Run 2 sont très proches de celles du Run 1 malgré le vieillissement du détecteur et une luminosité instantanée beaucoup plus élevée.

**Title:** Measurement of the N-jettiness variables in the production of Z boson events with the CMS detector and performance of its electromagnetic calorimeter

**Keywords:** Jets, Z boson, Large Hadron Collider, Standard Model

**Abstract:** This thesis presents the measurements of the differential cross section of the Z boson production in association with jets in proton-proton collisions at a center-of-mass energy of 13 TeV. The analyzed data have been collected by the CMS (Compact Muon Solenoid) experiment of LHC (Large Hadron Collider) during 2018 and corresponds to an integrated luminosity of  $59 \text{ fb}^{-1}$ . The cross section is measured as a function of track-based event shape variables: zero-jettiness, one-jettiness and sum of the transverse momentum of particles, and jet-based event shape variables ( $\tau_{\text{sum}}$  and  $\tau_{\text{max}}$ ).

The measurement of event shape variables is performed using events where pairs of muons are produced in the decay of a real Z boson with an invariant mass between 76 and 106 GeV. Track-based variables are also measured for off-shell Z bosons with an invariant mass between 125 and 150 GeV, 150 and 350 GeV, and 350 and 1500 GeV, as well as in four different Z

boson transverse momentum regions. The measurements have been compared with three types of theoretical predictions with LO, NLO, and NNLO QCD accuracies obtained with two Monte Carlo generators, MADGRAPH5\_amc@NLO and GENEVA.

Part of the thesis is devoted to the studies of the performance and intercalibration of the Electromagnetic calorimeter (ECAL) during the Run 2 data-taking period (2016, 2017, and 2018). This subdetector is crucial for the detection of photons and electrons; therefore, it is very important for many searches at CMS. Regular monitoring and calibration allowed an excellent performance to be achieved: the energy resolution was maintained within 1.7% in the central part of the detector. It is also shown that the performance from Run 2 is very close to the one from Run 1 despite ageing of the detector and much higher instantaneous luminosity.



# Contents

|                                                         |            |
|---------------------------------------------------------|------------|
| <b>List of Figures</b>                                  | <b>vii</b> |
| <b>List of Tables</b>                                   | <b>xii</b> |
| <b>Introduction</b>                                     | <b>1</b>   |
| <b>1 Theory overview</b>                                | <b>3</b>   |
| 1.1 Standard Model . . . . .                            | 3          |
| 1.1.1 The interactions in the Standard Model . . . . .  | 4          |
| 1.1.2 Beyond the Standard Model . . . . .               | 11         |
| 1.2 Proton-proton collisions . . . . .                  | 11         |
| 1.3 Drell-Yan process . . . . .                         | 13         |
| 1.4 Monte Carlo Simulation . . . . .                    | 15         |
| 1.4.1 MADGRAPH5_aMC@NLO . . . . .                       | 16         |
| 1.4.2 GENEVA . . . . .                                  | 17         |
| <b>2 Experimental setup</b>                             | <b>19</b>  |
| 2.1 Large Hadron Collider . . . . .                     | 19         |
| 2.1.1 The design of the Large Hadron Collider . . . . . | 20         |
| 2.1.2 Performance of the LHC . . . . .                  | 22         |
| 2.1.3 LHC operations . . . . .                          | 24         |
| 2.2 Compact Muon Solenoid . . . . .                     | 25         |
| 2.2.1 Coordinate system . . . . .                       | 26         |
| 2.2.2 Superconducting solenoid . . . . .                | 28         |
| 2.2.3 Tracker . . . . .                                 | 28         |
| 2.2.4 The electromagnetic calorimeter . . . . .         | 31         |
| 2.2.5 The hadronic calorimeter . . . . .                | 32         |
| 2.2.6 Muon system . . . . .                             | 34         |
| 2.2.7 Trigger . . . . .                                 | 35         |
| 2.3 Physics object reconstruction . . . . .             | 37         |
| 2.3.1 Track and vertex reconstruction . . . . .         | 37         |

|          |                                                            |           |
|----------|------------------------------------------------------------|-----------|
| 2.3.2    | Calorimeter clustering . . . . .                           | 38        |
| 2.3.3    | Muons . . . . .                                            | 39        |
| 2.3.4    | Particle flow algorithm description . . . . .              | 39        |
| 2.4      | CMS upgrade for HL-LHC . . . . .                           | 40        |
| <b>3</b> | <b>ECAL Calibration</b>                                    | <b>41</b> |
| 3.1      | Energy reconstruction . . . . .                            | 41        |
| 3.2      | Signal reconstruction . . . . .                            | 43        |
| 3.3      | Laser monitoring . . . . .                                 | 43        |
| 3.4      | ECAL intercalibration . . . . .                            | 46        |
| 3.5      | The $Z \rightarrow e^+e^-$ calibration method . . . . .    | 48        |
| 3.5.1    | IJazZ tool . . . . .                                       | 48        |
| 3.5.2    | $\eta$ scale calibration . . . . .                         | 49        |
| 3.5.3    | Intercalibration along $\phi$ . . . . .                    | 50        |
| 3.5.4    | Resolution and combination . . . . .                       | 52        |
| 3.5.5    | Simulation studies . . . . .                               | 56        |
| 3.6      | ECAL upgrade for High Luminosity LHC . . . . .             | 57        |
| <b>4</b> | <b>Measurements of track-based event shape observables</b> | <b>60</b> |
| 4.1      | Observable definition . . . . .                            | 61        |
| 4.2      | Data and simulation samples . . . . .                      | 62        |
| 4.3      | Selection . . . . .                                        | 63        |
| 4.3.1    | Event selection . . . . .                                  | 64        |
| 4.3.2    | Particle selection . . . . .                               | 66        |
| 4.4      | Simulation corrections . . . . .                           | 68        |
| 4.4.1    | Pileup . . . . .                                           | 68        |
| 4.4.2    | Momentum corrections . . . . .                             | 68        |
| 4.4.3    | Scale factors for ID and isolation . . . . .               | 69        |
| 4.4.4    | Trigger Scale factors . . . . .                            | 69        |
| 4.4.5    | Comparison of data and simulation . . . . .                | 72        |
| 4.5      | Unfolding method . . . . .                                 | 76        |
| 4.6      | Uncertainties . . . . .                                    | 78        |
| 4.6.1    | Theoretical predictions uncertainties . . . . .            | 81        |
| 4.7      | Results . . . . .                                          | 81        |
| <b>5</b> | <b>Measurement of jet-based event shape observables</b>    | <b>99</b> |
| 5.1      | Observable definition . . . . .                            | 99        |
| 5.2      | Jet reconstruction and selection . . . . .                 | 101       |
| 5.3      | Uncertainties . . . . .                                    | 105       |
| 5.4      | Results . . . . .                                          | 106       |

|                     |            |
|---------------------|------------|
| <b>Conclusion</b>   | <b>108</b> |
| <b>Synthèse</b>     | <b>110</b> |
| <b>Bibliography</b> | <b>115</b> |

# List of Figures

|      |                                                                                                                                                                                                             |    |
|------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 1.1  | Overview of the fundamental particles of the Standard model [11]. .                                                                                                                                         | 3  |
| 1.2  | Schematic view of the hadronization process. . . . .                                                                                                                                                        | 7  |
| 1.3  | Illustration of the Higgs potential [18] . . . . .                                                                                                                                                          | 10 |
| 1.4  | The scale dependence of the NNPDF3.1 set of PDFs [28]. . . . .                                                                                                                                              | 12 |
| 1.5  | Feynman diagram of the Drell-Yan process. . . . .                                                                                                                                                           | 13 |
| 1.6  | The Drell-Yan process at LO (a), at NLO with initial state radiation (b), at NLO with a gluon loop at the initial state (c), and at NLO with a quark-gluon initial state and with outgoing hadronization. . | 14 |
| 1.7  | Sketch of a proton-proton collision as simulated by a Monte Carlo event generator [31]. . . . .                                                                                                             | 15 |
| 2.1  | Aerial view of Geneva region, with the position of LHC, sketched in yellow [40]. . . . .                                                                                                                    | 20 |
| 2.2  | Schematic view of accelerator system at CERN [41]. . . . .                                                                                                                                                  | 21 |
| 2.3  | Schematic view of the LHC ring [4]. . . . .                                                                                                                                                                 | 22 |
| 2.4  | Integrated luminosity collected by CMS in 2018 [45]. . . . .                                                                                                                                                | 23 |
| 2.5  | Pileup distribution observed by CMS [45]. . . . .                                                                                                                                                           | 24 |
| 2.6  | The timeline of the LHC operations [46]. . . . .                                                                                                                                                            | 25 |
| 2.7  | Schematic view of the CMS detector [47]. . . . .                                                                                                                                                            | 26 |
| 2.8  | The longitudinal and transverse view of the CMS detector [48]. . . .                                                                                                                                        | 27 |
| 2.9  | The magnetic field produced by a superconducting magnet. The right part presents the lines of the field, while on the left the strength of the magnetic field is shown [50]. . . . .                        | 28 |
| 2.10 | Schematic view of CMS tracker layout [53]. . . . .                                                                                                                                                          | 29 |
| 2.11 | Comparison of the upgraded pixel detector with the original one [54].                                                                                                                                       | 30 |
| 2.12 | Schematic view of the longitudinal layout of ECAL detector [55]. . .                                                                                                                                        | 31 |
| 2.13 | Schematic view of the ECAL structure [9]. . . . .                                                                                                                                                           | 32 |
| 2.14 | Schematic view of the longitudinal layout of HCAL detector [9] . .                                                                                                                                          | 33 |
| 2.15 | Schematic view of the longitudinal layout of the muon detector system [58]. . . . .                                                                                                                         | 34 |
| 2.16 | A flowchart of the L1 trigger [60]. . . . .                                                                                                                                                                 | 36 |

|      |                                                                                                                                                                                                                                                                                                                                       |    |
|------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 2.17 | Slice of the CMS detector with the particles signature in the sub-detectors [62]. . . . .                                                                                                                                                                                                                                             | 38 |
| 3.1  | Example of fitted pulses for simulated events with 20 average pileup interactions and 25 ns bunch spacing, for a signal in the barrel (left) and endcap (right) [77]. . . . .                                                                                                                                                         | 44 |
| 3.2  | Relative response to laser light (440 nm in 2011 and 447 nm from 2012 onwards) injected in the ECAL crystals, measured by the ECAL laser monitoring system, averaged over all crystals in bins of pseudorapidity. The bottom plot shows the delivered instantaneous LHC luminosity [79]. . . . .                                      | 45 |
| 3.3  | The stability of the relative energy scale measured from the invariant mass distribution of $\pi^0 \rightarrow \gamma\gamma$ decays in the EB and plotted as a function of time, over a period of 3 hours during an LHC fill [80]. . . . .                                                                                            | 46 |
| 3.4  | The fit performed in the wide (left) and adapted (right) invariant mass window. . . . .                                                                                                                                                                                                                                               | 50 |
| 3.5  | The ratio of scale parameters between data and simulation, in gray for $R_9 > 0.94$ electrons and in green for $R_9 < 0.94$ electrons. . . . .                                                                                                                                                                                        | 51 |
| 3.6  | Map of the intercalibration constants for ECAL barrel region. . . . .                                                                                                                                                                                                                                                                 | 51 |
| 3.7  | Corrections applied to account for the effect of gaps between the supermodules. . . . .                                                                                                                                                                                                                                               | 52 |
| 3.8  | The final map of the intercalibration constants for barrel and endcaps for 2018. . . . .                                                                                                                                                                                                                                              | 53 |
| 3.9  | The overall precision of the different IC measurement methods as well as their combination for 2016 (top left), 2017 (top right), and 2018 (bottom). . . . .                                                                                                                                                                          | 54 |
| 3.10 | Energy resolution as a function of the pseudorapidity for the 2018 data set with the preliminary and refined calibration. . . . .                                                                                                                                                                                                     | 55 |
| 3.11 | Energy resolution with the refined calibration as a function of the pseudorapidity with 2016, 2017, and 2018 data periods for all electrons (left) and for low bremsstrahlung electrons (right). . . . .                                                                                                                              | 55 |
| 3.12 | Energy resolution with the refined calibration as a function of the pseudorapidity comparing Run 1 (2012) data-taking period with Run 2 (2016, 2017, and 2018) data-taking periods. The resolution is derived from all events (left) and from the events with the number of reconstructed vertices between 25 and 35 (right). . . . . | 56 |
| 3.13 | Contribution to the ECAL energy resolution from different effects on the (left) and cumulative energy resolution obtained by adding up the contribution from different effects. . . . .                                                                                                                                               | 57 |
| 3.14 | The ECAL readout system. . . . .                                                                                                                                                                                                                                                                                                      | 58 |
| 3.15 | Schematic view of the new ECAL barrel electronics. . . . .                                                                                                                                                                                                                                                                            | 59 |

|      |                                                                                                                                                        |    |
|------|--------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 3.16 | Energy resolution (left) and time resolution (right) obtained in the test beam campaign with the CATIA ASIC connected to a commercial ADC[87]. . . . . | 59 |
| 4.1  | The distribution of zero-jettiness at the generator level (blue) and reconstructed level (red). . . . .                                                | 66 |
| 4.2  | The distribution of zero-jettiness computed with charged, neutral, and all particles from the event . . . . .                                          | 67 |
| 4.3  | Vertex quality association flag for matched and non matched tracks. . . . .                                                                            | 68 |
| 4.4  | Fit results for one of the central pseudorapidity bins for probe muons that pass (left) and fail (right) selections. . . . .                           | 70 |
| 4.5  | Single and double muon trigger efficiencies in data (left) and simulation (right) shown in bins of pseudorapidity. . . . .                             | 71 |
| 4.6  | The scale factors obtained for 2018 data taking period shown in bins of pseudorapidity. . . . .                                                        | 72 |
| 4.7  | Data to simulation comparison of the invariant mass distribution with linear (left) and logarithmic (right) scale. . . . .                             | 73 |
| 4.8  | Data to simulation comparison of muon pair distributions: rapidity (left) and transverse momentum (right). . . . .                                     | 73 |
| 4.9  | Data to simulation comparison of reconstructed muon distributions: transverse momentum, pseudorapidity, and azimuthal angle. . . . .                   | 74 |
| 4.10 | Data to simulation comparison of charged particle distributions: transverse momentum and pseudorapidity. . . . .                                       | 75 |
| 4.11 | Comparison of transverse momentum and pseudorapidity of charged particles distributions for CP5 and CUTEPM1 tunes at generator level. . . . .          | 76 |
| 4.12 | Response matrices for zero-jettiness (left) and one-jettiness (right). . . . .                                                                         | 78 |
| 4.13 | The fit function used for reweighting one-jettiness . . . . .                                                                                          | 80 |
| 4.14 | Uncertainty breakdown for zero-jettiness and one-jettiness . . . . .                                                                                   | 82 |
| 4.15 | Data to simulation comparison and differential cross section as a function of zero-jettiness. . . . .                                                  | 83 |
| 4.16 | The distribution of zero-jettiness computed for the events without MPI (left) and with MPI (right). . . . .                                            | 83 |
| 4.17 | Data to simulation comparison and differential cross section as a function of zero-jettiness for $p_T^Z < 6$ GeV. . . . .                              | 84 |
| 4.18 | Data to simulation comparison and differential cross section as a function of zero-jettiness for $6 < p_T^Z < 12$ GeV. . . . .                         | 85 |
| 4.19 | Data to simulation comparison and differential cross section as a function of zero-jettiness for $12 < p_T^Z < 25$ GeV. . . . .                        | 85 |
| 4.20 | Data to simulation comparison and differential cross section as a function of zero-jettiness for $p_T^Z > 25$ GeV. . . . .                             | 86 |

|      |                                                                                                                                                                                 |    |
|------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 4.21 | Data to simulation comparison and differential cross section as a function of zero-jettiness for $125 < M_{\mu\mu} < 150$ GeV. . . . .                                          | 87 |
| 4.22 | Data to simulation comparison and differential cross section as a function of zero-jettiness for $150 < M_{\mu\mu} < 350$ GeV. . . . .                                          | 87 |
| 4.23 | Data to simulation comparison and differential cross section as a function of zero-jettiness for $350 < M_{\mu\mu} < 1500$ GeV. . . . .                                         | 88 |
| 4.24 | Data to simulation comparison and differential cross section as a function of one-jettiness. . . . .                                                                            | 89 |
| 4.25 | Data to simulation comparison and differential cross section as a function of one-jettiness for $6 < p_T^Z < 12$ GeV. . . . .                                                   | 89 |
| 4.26 | Data to simulation comparison and differential cross section as a function of one-jettiness for $12 < p_T^Z < 25$ GeV. . . . .                                                  | 90 |
| 4.27 | Data to simulation comparison and differential cross section as a function of one-jettiness for $25 < p_T^Z < 35$ GeV. . . . .                                                  | 90 |
| 4.28 | Data to simulation comparison and the differential cross section as a function of one-jettiness for $p_T^Z > 35$ GeV. . . . .                                                   | 91 |
| 4.29 | Data to simulation comparison and differential cross section as a function of one-jettiness for $125 < M_{\mu\mu} < 150$ GeV. . . . .                                           | 92 |
| 4.30 | Data to simulation comparison and differential cross section as a function of one-jettiness for $150 < M_{\mu\mu} < 350$ GeV. . . . .                                           | 92 |
| 4.31 | Data to simulation comparison and differential cross section as a function of one-jettiness for $350 < M_{\mu\mu} < 1500$ GeV. . . . .                                          | 93 |
| 4.32 | Data to simulation comparison and differential cross section as a function of the sum of the transverse momentum of charged particles. . . . .                                  | 94 |
| 4.33 | Data to simulation comparison and differential cross section as a function of the sum of the transverse momentum of charged particles for $p_T^Z < 6$ GeV. . . . .              | 94 |
| 4.34 | Data to simulation comparison and differential cross section as a function of the sum of the transverse momentum of charged particles for $6 < p_T^Z < 12$ GeV. . . . .         | 95 |
| 4.35 | Data to simulation comparison and differential cross section as a function of the sum of the transverse momentum of charged particles for $12 < p_T^Z < 25$ GeV. . . . .        | 95 |
| 4.36 | Data to simulation comparison and differential cross section as a function of the sum of the transverse momentum of charged particles for $p_T^Z > 25$ GeV. . . . .             | 96 |
| 4.37 | Data to simulation comparison and differential cross section as a function of the sum of the transverse momentum of charged particles for $125 < M_{\mu\mu} < 150$ GeV. . . . . | 97 |

|      |                                                                                                                                                                                  |     |
|------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 4.38 | Data to simulation comparison and differential cross section as a function of the sum of the transverse momentum of charged particles for $150 < M_{\mu\mu} < 350$ GeV. . . . .  | 97  |
| 4.39 | Data to simulation comparison and differential cross section as a function of the sum of the transverse momentum of charged particles for $350 < M_{\mu\mu} < 1500$ GeV. . . . . | 98  |
| 5.1  | Rapidity weighting functions for $\tau_B$ and $\tau_C$ [8]. . . . .                                                                                                              | 100 |
| 5.2  | Comparison of the leading jet transverse momentum distribution at generated level, reconstructed level with CHS and reconstructed level with the PUPPI algorithm. . . . .        | 104 |
| 5.3  | Data to simulation comparison of the leading jet transverse momentum and inclusive number of jets. . . . .                                                                       | 104 |
| 5.4  | Data to simulation comparison of the leading jet transverse momentum and inclusive number of jets. . . . .                                                                       | 105 |
| 5.5  | Uncertainty breakdown for $\tau_{\max}$ and $\tau_{\text{sum}}$ . . . . .                                                                                                        | 106 |
| 5.6  | Data to simulation comparison and differential cross section as a function of $\tau_{\max}$ . . . . .                                                                            | 107 |
| 5.7  | Data to simulation comparison and differential cross section as a function of $\tau_{\text{sum}}$ . . . . .                                                                      | 107 |

# List of Tables

|     |                                                              |     |
|-----|--------------------------------------------------------------|-----|
| 4.1 | List of data samples used in the analysis. . . . .           | 62  |
| 4.2 | List of simulated samples and their cross section. . . . .   | 64  |
| 4.3 | Medium ID criteria used for the muon selection [65]. . . . . | 65  |
| 4.4 | PV association flags description. . . . .                    | 67  |
| 5.1 | Criteria for identification of jets [108]. . . . .           | 103 |

# Introduction

The Standard Model theory, developed in the 1960s, describes the elementary particles and the interactions between them. Since its proposal, enormous work has been done to experimentally verify its predictions. The results obtained from many high energy physics experiments such as Tevatron at Fermilab (1983 to 2011) [1], the Large Electron Positron Collider at CERN (1989 to 2000) [2], the Hadron Electron Ring Accelerator at DESY (1992 to 2007) [3], and the Large Hadron Collider at CERN (from 2010) [4] have shown a good agreement of the measurements with the theoretical predictions. However, it is not a complete theory and there are still unanswered questions that theories beyond the Standard Model attempt to explain.

In 2012 the existence of the Higgs boson was confirmed at the LHC by the CMS [5] and ATLAS [6] experiments. It was the last missing piece of the Standard Model. The LHC experiments continued their work with studies of the properties of the Higgs boson, high-precision measurements of well-known processes, and searches for new physics beyond the Standard Model.

The processes of interest are identified by looking for a signal with specific numbers of leptons, photons, or jets. To discriminate the signal from the background events, it is often needed to use a veto on the hadronic activity in an event. This veto is typically based on jets but could also be implemented using track-based event shape variables such as N-jettiness [7] or jet-based event shape variables [8]. One of the advantages of using inclusive event shape variables is that the summation of logarithms to next-to-next-to-leading logarithmic order can be performed. From the experimental point of view, the event shape variables provide an efficient method to veto jets. Using the jet-based variables, the central jets can be vetoed while the phase space constraints are not strict. Before using these variables in analyses, it is essential to ensure that they are well described by the Monte Carlo simulations, which can be done by comparing the predictions with measurements.

The research presented in this thesis are measurements of track-based and jet-based event shape observables in events with one Z boson produced at the LHC in proton-proton collisions at a center-of-mass energy of 13 TeV.

The measurements are done using data collected in 2018 by the CMS detector

[9]. The measurements of track-based event shape variables are performed using events where pairs of muons are produced in the decay of a on-shell Z boson with an invariant mass between 76 and 106 GeV, and also for off-shell Z bosons with an invariant mass between 125 and 150 GeV, 150 and 350 GeV and 350 and 1500 GeV. In addition, these variables are measured in four different Z boson transverse momentum regions.

Part of this thesis is devoted to the intercalibraion of the Electromagnetic Calorimeter (ECAL) [10] of CMS. This subdetector is crucial for the detection of photons and electrons and its calibration is very important for many searches performed with CMS.

This thesis is organized as follows. The theoretical introduction to the Standard Model is described in Chapter 1. In Chapter 2, the experimental setup is presented. The intercalibration and performance of the ECAL are described in Chapter 3. The measurements of track-based event shape variables are presented in Chapter 4, and the measurements of jet-based event shape variables are presented in Chapter 5. A summary and conclusion follow.

# Chapter 1

## Theory overview

### 1.1 Standard Model

The Standard Model of particle physics (SM) is the theory developed in the 1960s that describes the fundamental particles and the interactions between them. This theory models the electromagnetic, weak, and strong interactions using the Quantum Field Theory (QFT) formulation. The gravitational force, which is negligible at the subatomic level, is not included in the SM.

According to the SM, all matter is made of particles with spin  $\frac{1}{2}$  called fermions, and the interactions between them are mediated by particles with integer spin called bosons (Figure 1.1).

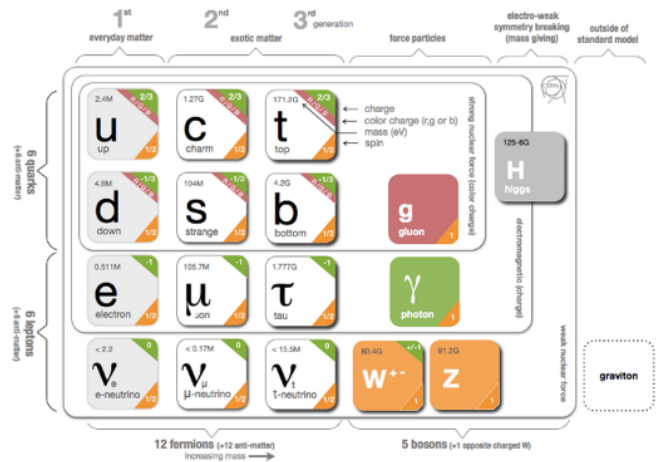


Figure 1.1: Overview of the fundamental particles of the Standard model [11].

Fermions can be classified into two groups: leptons and quarks. Each of the

twelve fermions has its corresponding antiparticle.

Leptons are particles that interact via electromagnetic and weak interactions. Their charge is integer or null. The charged leptons are electron (e), muon ( $\mu$ ), and tau ( $\tau$ ). The corresponding neutral leptons, which interact only through weak interaction, are the electron neutrino ( $\nu_e$ ), the muon neutrino ( $\nu_\mu$ ), and the tau neutrino ( $\nu_\tau$ ).

Quarks interact through electromagnetic, weak, and strong interactions. The electric charge of quarks is fractional and it is  $\frac{2}{3}$  for the up (u), charm (c), and top (t) quarks and  $-\frac{1}{3}$  for the down (d), strange (s) and bottom (b) quarks. Quarks do not exist as a free state but they form bound states called hadrons. Hadrons composed of three quarks (antiquarks) are called baryons (antibaryons), while the bound states formed by a quark-antiquark pair are called mesons. The charge of the hadrons is null or integer.

Fermions are classified into three generations. Those from the first generation, which are the lightest ones, make the ordinary matter. The second and third generation particles, except neutrinos, are unstable and are accessible at higher energies.

Bosons, with spin 1, are the mediators of the electromagnetic, weak, and strong interaction. The photon ( $\gamma$ ) is massless and is the mediator of electromagnetic interaction, while the  $W^\pm$  and Z bosons are massive and are the mediators of the weak interaction. The strong force is carried by massless gluons (g). In addition to the spin-1 bosons, there is a Higgs boson with the spin 0, mediator of the scalar Higgs field. It was discovered in 2012 by the CMS and ATLAS collaborations [5, 6].

### 1.1.1 The interactions in the Standard Model

The SM is a gauge theory based on the  $SU(3)_C \times SU(2)_L \times U(1)_Y$  symmetry, where the strong interaction is associated with the  $SU(3)_C$  symmetry and  $SU(2)_L \times U(1)_Y$  is the symmetry group for the electromagnetic and weak interaction.

#### Quantum electrodynamics

Quantum electrodynamics (QED) is the theory that describes electromagnetic interaction. To deliver the QED Lagrangian, we start with the Dirac Lagrangian [12] that describes the motion of the free fermion in each point of space-time  $x$ :

$$L^{Dirac} = \bar{\psi}(x)(i\gamma^\mu\partial_\mu - m)\psi(x), \quad (1.1)$$

where  $\gamma$  are Dirac matrices and  $\psi(x)$  the fermionic field. This Lagrangian must be invariant under the gauge transformation of the field  $\psi(x)$ :

$$\psi(x) \rightarrow \psi'(x) = e^{ie\alpha(x)}\psi(x), \quad (1.2)$$

where  $\alpha(x)$  is any function of  $x$  and  $e$  is the dimensionless coupling strength of electrodynamics. In order to achieve invariance, the derivative  $\partial_\mu$  must be replaced by the covariant derivative:

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + ieA_\mu, \quad (1.3)$$

where  $A_\mu$  is the gauge field that corresponds to the photon and has transformation property:

$$A_\mu \rightarrow A'_\mu = A_\mu - \partial_\mu \alpha(x). \quad (1.4)$$

To complete the QED Lagrangian, the kinematic term describing the propagation of photons needs to be included. This is done by introducing the field strength tensor  $F^{\mu\nu}$  defined as:

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu, \quad (1.5)$$

which leads to the photon propagation term:

$$L^{gauge} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \quad (1.6)$$

The final Lagrangian of the QED can be written as:

$$L_{QED} = \bar{\psi}(x)(i\gamma^\mu\partial_\mu - m)\psi(x) - e(\bar{\psi}(x)\gamma^\mu\psi(x))A_\mu - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \quad (1.7)$$

The first term of the Lagrangian corresponds to the free propagation of fermions, the second term represents the interactions of fermions and photons, and the last term takes into account the propagation of photons.

### Quantum chromodynamics

The strong interaction is described by the Quantum Chromodynamics (QCD) theory which is based on the gauge symmetry group SU(3). The particles that interact with the strong force (quarks and gluons) have an additional quantum number called color. The quarks can have one of the three colors: red, blue or green, while the antiquarks have anticolors. The eight gluons are carrying the combination of color and anticolor:  $r\bar{b}$ ,  $r\bar{g}$ ,  $b\bar{r}$ ,  $b\bar{g}$ ,  $g\bar{r}$ ,  $g\bar{b}$ ,  $\frac{1}{2}(r\bar{r} - g\bar{g})$  and  $\frac{1}{\sqrt{6}}(r\bar{r} + b\bar{b} - 2g\bar{g})$ .

Following the procedure for QED described in the previous section, the starting point for QCD is also the Dirac Lagrangian defined in Eq. 1.1. The interaction is introduced by requiring the Lagrangian to be invariant under the following gauge transformation:

$$\psi(x) \rightarrow \psi'(x) = U\psi(x) = (e^{ig_s \sum_a \theta_a(x) T_a})\psi(x), \quad (1.8)$$

where  $g_s$  is the dimensionless coupling strength of the strong force and  $T_a$  are the eight generators of SU(3) related with the Gell-Mann matrices as  $T_a = \lambda^a/2$ . To make the Lagrangian invariant under this transformation, similarly to QED, a covariant derivative is introduced:

$$D^\mu = \partial^\mu + ig_s T_a G_a^\mu, \quad (1.9)$$

where  $G_a^\mu$  are eight gauge fields ( $a = 1, \dots, 8$ ). The gauge transformation properties of the gauge fields are defined as:

$$G_b^\mu \rightarrow G_b'^\mu = G_b^\mu - \partial^\mu \theta_b - g_s \lambda_{abc} \theta_a G_c^\mu. \quad (1.10)$$

where  $\theta$  denoting eight functions of space-time coordinates. The field strength can be written as:

$$F_a^{\mu\nu} = (\partial^\mu G_a^\nu - \partial^\nu G_a^\mu) - g_s \lambda_{abc} G_b^\mu G_c^\nu. \quad (1.11)$$

Therefore, the QCD Lagrangian has the form:

$$\begin{aligned} L_{QCD} = & \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi - g_s(\bar{\psi}\gamma^\mu T_a \psi)G_a^\mu - \frac{1}{4}(\partial^\mu G_a^\nu - \partial^\nu G_a^\mu)(\partial_\mu G_{a\nu} - \partial_\nu G_{a\mu}) \\ & + \frac{1}{2}g_s \lambda_{abc} G_b^\mu G_c^\nu (\partial_\mu G_{a\nu} - \partial_\nu G_{a\mu}) - \frac{1}{4}g_s^2 \lambda_{abc} \lambda_{ars} G_b^\mu G_c^\nu G_{r\mu} G_{s\nu}. \end{aligned} \quad (1.12)$$

In this Lagrangian, the first term describes free quark propagation. The second term represents the quark-gluon interaction. The gluon propagation term corresponds to the third term. The last two terms are the triple coupling and the quartic coupling, which introduce the gluons self-coupling.

The fact that the quarks do not exist as free particles and can only be detected in bound states, can be explained by gluon self-coupling. If we consider two quarks pulled at some distance, the exchange of gluons and the interaction between the gluons themselves would squeeze the field and increase the force. Since the field becomes proportional to the distance between quarks, an infinite amount of energy would be needed to separate them at infinity. Therefore, the quarks cannot be detected as free particles, but only in bound color singlet states. This property of strong interaction is called color confinement. The consequence of color confinement is the process of hadronization. As shown in Figure 1.2, the quark-antiquark pairs are separated and the energy of the strong field between them increases as they are moving apart from each other. When this energy becomes high enough, a new quark-antiquark pair is produced. This process is repeated until the quarks have energy low enough to create hadrons. The produced hadrons are often the results of boosted interactions, which makes the particles to be collinear and form what is called a jet.

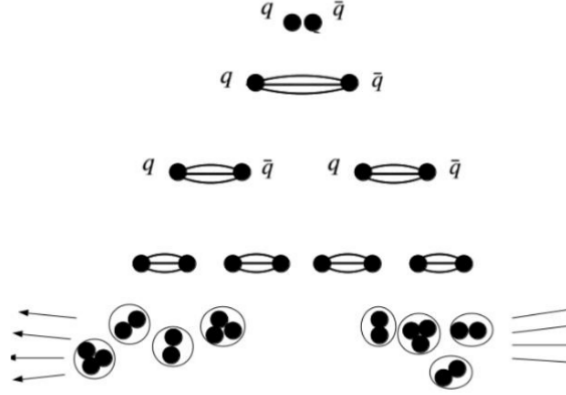


Figure 1.2: Schematic view of the hadronization process.

The running of the strong coupling constant  $\alpha_s$  can be obtained by the renormalization group equation (RGE):

$$Q^2 \frac{\partial \alpha}{\partial Q^2} = \beta(\alpha), \quad (1.13)$$

where  $Q$  is the energy scale and  $\beta$  function encodes the loop calculations. Solving the equation 1.13 at leading order, with renormalization scale  $\mu_R^2$  gives:

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu_R^2)}{1 + \alpha_s(\mu_R^2) \beta \ln \left( \frac{Q^2}{\mu_R^2} \right)} \quad (1.14)$$

At high energies, the coupling  $\alpha_s$  decreases, which means the interaction between quarks and gluons becomes weaker. This leads to the QCD fundamental property called "asymptotic freedom", where quarks and gluons can be considered as free, non-interacting particles and they are referred to as partons.

### Electroweak interaction

The electroweak theory, proposed by Glashow, Weinberg, and Salam [13, 14, 15], is the theory that unites the electromagnetic and weak interaction. The electroweak interaction follows a  $SU(2)_L \times U(1)_Y$  symmetry, which requires three fields for  $SU(2)_L$ :  $W^1$ ,  $W^2$ ,  $W^3$  and one field for  $U(1)_Y$ :  $B$ . The generators are the weak

isospin  $T$  and the weak hypercharge  $Y$ . Their relation to the electric charge is:

$$Q = T_3 + \frac{Y}{2}. \quad (1.15)$$

In the electroweak theory, it has to be taken into account that the right-handed and the left-handed projections of the fields do not behave the same. They are defined by the chirality operators  $P_{R,L}$ :

$$\psi_{R,L} = P_{R,L}\psi = \frac{1}{2}(1 \pm \gamma^5)\psi. \quad (1.16)$$

Fermions associated with the left-handed projection of the field form weak isospin doublets, while the right-handed fermions are weak isospin singlets.

$$\begin{aligned} \psi_L &= \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix} \\ \psi_R &= (\nu_{eR}), (e_R). \end{aligned} \quad (1.17)$$

As in the case of QED and QCD, a gauge transformation is introduced and the invariance under this transformation is required by introducing the covariant derivative:

$$D^\mu = \partial^\mu + igT_k W_k^\mu + i\frac{g'}{2}YB^\mu. \quad (1.18)$$

The interaction term in the Lagrangian has the form:

$$L_{EWK}^{interaction} = -\frac{g'}{2}(\bar{\psi}\gamma_\mu Y\psi)B^\mu - g(\bar{\psi}T^k Y\psi)W^{k\mu}. \quad (1.19)$$

The gauge bosons  $W^\pm$ ,  $Z$  and  $\gamma$  cannot be directly determined from the gauge fields. By introducing the rotation angle  $\theta_W$  (the Weinberg angle), the electroweak interaction Lagrangian can be expressed in terms of the gauge bosons:

$$\begin{aligned} B &= A\cos\theta_W - Z\sin\theta_W \\ W^3 &= A\cos\theta_W + Z\sin\theta_W \\ W^{1,2} &= \frac{W^- \pm W^+}{\sqrt{2}}. \end{aligned} \quad (1.20)$$

The rotation angle is defined by the coupling strengths:

$$e = g\sin\theta_W = g'\cos\theta_W. \quad (1.21)$$

Finally, the interaction term of Lagrangian can be expressed as:

$$\begin{aligned} L_{EWK}^{interaction} &= -\frac{g}{\sqrt{2}}(\bar{\nu}_{eL}\gamma^\mu W^- e_L + \bar{e}_L\gamma^\mu W^+ \nu_L) - e\bar{\psi}\gamma^\mu A Q\psi - \\ &\quad \frac{e}{2\sin\theta_W\cos\theta_W}\bar{\psi}\gamma^\mu Z[T^3 - 2Q\sin^2\theta_W - T_L^3\gamma^5]\psi. \end{aligned} \quad (1.22)$$

The first term in the equation corresponds to the weak interaction with the change of electric charge mediated by  $W^\pm$  bosons. They interact only with left-handed fermions. The second term represents the photon-fermion interaction, which was present in the QED Lagrangian (Eq. 1.7). The last term is the neutral weak interaction mediated by the Z boson.

### Electroweak symmetry breaking

The mass terms are added the Lagrangian 1.22 through the Brout-Englert-Higgs mechanism [16, 17] that breaks spontaneously the gauge invariance symmetry. The following terms are added:

$$L_{Higgs} = (D^\mu \phi)^\dagger (D_\mu \phi) - V(\phi), \quad (1.23)$$

where  $\phi$  is a complex scalar field, while the  $V(\phi)$  is the potential of the field. In order to include the mass term, this field needs to have at least three degrees of freedom:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad (1.24)$$

The potential of the field, called Higgs potential, is defined as:

$$V(\phi) = \lambda(\phi^\dagger \phi)^2 + \mu^2 \phi^\dagger \phi. \quad (1.25)$$

If the  $\lambda$  term and  $\mu^2$  are positive, the potential is also positive with the single minimum at  $\phi = 0$ . In the case of  $\mu^2 < 0$  and  $\lambda > 0$ , the potential has an infinite set of minima at:

$$\phi^+ \phi = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = -\frac{\mu^2}{2\lambda} = \frac{1}{2}v^2. \quad (1.26)$$

In the second case, the spontaneous symmetry breaking can be achieved. The ground state is typically chosen as :

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (1.27)$$

The Lagrangian is invariant under the symmetry transformation but the potential around the minimum is not. The shape of potential is graphically illustrated in Figure 1.3.

Finally, using the field extension around the ground state, the Higgs Lagrangian

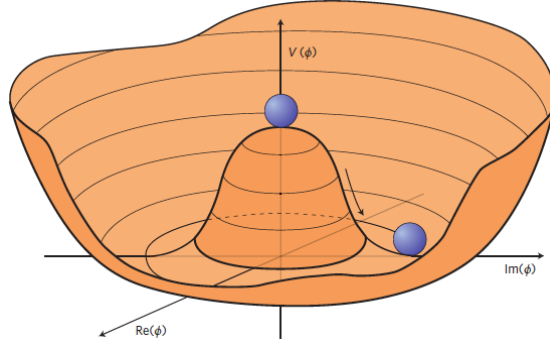


Figure 1.3: Illustration of the Higgs potential [18]

can be written as:

$$\begin{aligned}
L_{Higgs} = & \left\{ \frac{1}{2} \partial_\mu h \partial^\mu h - \frac{1}{2} 2v^2 \lambda h^2 \right\} + \left\{ \frac{1}{3!} 6v \lambda h^3 - \frac{1}{4!} 6 \lambda h^4 \right\} \\
& + \left\{ \frac{1}{2} \frac{v^2 g^2}{4} W_\mu^{-t} W^{-\mu} + \frac{1}{2} \frac{v^2 g}{4} W_\mu^{+t} W^{+\mu} \right\} \\
& + \left\{ \frac{1}{2} \frac{v^2 (g^2 + g'^2)}{4} \left( \frac{g W_\mu^3 - g' B^\mu}{\sqrt{g^2 + g'^2}} \right)^2 + 0 \left( \frac{g' W_\mu^3 + g B^\mu}{\sqrt{g^2 + g'^2}} \right)^2 \right\} \\
& + \left\{ \frac{1}{4} (2vh + h^2) \left[ g^2 W_\mu^- W^{+\mu} + \frac{1}{2} (g^2 + g'^2) \left( \frac{W_\mu^3 - g' B_\mu}{\sqrt{g^2 + g'^2}} \right)^2 \right] \right\}
\end{aligned} \tag{1.28}$$

The first line corresponds to the Higgs boson and its mass is given as:

$$m_H = v \sqrt{2\lambda} \tag{1.29}$$

The next two lines of the equation come from the kinetic term  $(D^\mu \phi)(D_\mu \phi)$  and they contain terms for masses of the gauge bosons:

$$\begin{aligned}
M_W &= \frac{1}{2} v g \\
M_Z &= \frac{1}{2} v \sqrt{g^2 + g'^2} \\
M_\gamma &= 0
\end{aligned} \tag{1.30}$$

The masses of the gauge bosons are measured experimentally with great precision, and their values are  $M_W = 80.379$  GeV [19] and  $M_Z = 91.1876$  GeV [20]. The experimentally measured mass of Higgs boson is  $M_H = 125$  GeV [5] [6].

### 1.1.2 Beyond the Standard Model

The Standard Model theory is so far consistent with all the experimental results; however, there are still open questions that are not addressed by this theory. Some of the shortcomings of the SM are:

- dark matter: Standard Model does not define a candidate that could explain the origin of the observed dark matter; therefore an extension of the SM is needed;
- asymmetry between matter and antimatter: the asymmetry in presence of matter and antimatter in the Universe is not explained by the SM; in order to provide an explanation, new theories need to be developed;
- neutrino masses: the observations of the neutrino mixing confirmed that the neutrinos are particles that have mass; according to the SSB mechanism, the neutrinos do not acquire mass, therefore it is needed to have the extension of SM which would explain this effect;
- gravity: the gravitational interaction is not included in SM; the effects due to the gravity are negligible at the accessible energy scale and it becomes relevant at the scale higher than the TeV scale; however, there are attempts to include gravity into the SM with a spin-2 mediator graviton.

## 1.2 Proton-proton collisions

Protons are baryons composed of two  $u$  quarks and one  $d$  quark. These quarks, called "valence" quarks, interact with each other within the proton and exchange gluons; gluons, in turn, also interact with each other and produce more gluons or quark-antiquark pairs called "sea" quarks. Quarks and gluons within protons are referred to as partons. According to the parton model [21], partons carry a fraction of the total proton momentum and are described by a parton distribution function (PDF), which gives the probability that parton has a fraction  $x$  of the total proton momentum  $P$  ( $p_i = xP$ ). The cross section of a proton-proton interaction cannot be computed easily, due to the complex structure of protons. Using the collinear factorisation [22], the cross section of the interactions of two protons with the final state  $X$  ( $pp \rightarrow X$ ), can be written as:

$$\sigma_{pp \rightarrow X} = \sum_{a,b} \int \int dx_a dx_b f_a(x_a, \mu_F) f_b(x_b, \mu_F) \sigma_{ab \rightarrow X}(x_a x_b, \mu_R, \mu_F), \quad (1.31)$$

where the sum runs over all the flavours of partons and  $\sigma_{ab \rightarrow X}(x_a, x_b, \mu_R, \mu_F)$  is the cross section at the partonic level. The partonic cross section depends on the

energy of partons ( $x_a, x_b$ ) and the scale  $\mu_F$  at which the factorization is performed. Since the calculations are performed with a perturbative expansion in  $\alpha_s^n$  that keeps only the first terms, the cross section depends on the renormalization scale  $\mu_R$  as well.

The PDFs are obtained mainly from the deep inelastic scattering (DIS) experiments such as lepton-hadron collider HERA [23] [23], hadron colliders, such as LHC, and the fixed-target experiments. The PDFs depend on the scale at which the hadron is probed, therefore it is important to determine the evolution of PDFs with the scale  $\mu_F$ . This evolution is described by the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [24, 25, 26]:

$$\mu_F \frac{\delta f_a(x, \mu_F)}{\delta \mu_F^2} = \sum \frac{\alpha_s(\mu_F^2)}{2\pi} \int_x^1 \frac{d\xi}{\xi} P_{a \rightarrow bc}(\frac{x}{\xi}, \mu_F) f_a(\xi, \mu_F), \quad (1.32)$$

where  $P_{a \rightarrow bc}$  is the Altarelli-Parisi splitting function, that gives the probability for a parton  $a$  to split into two partons  $bc$ . The resulting particle has a momentum fraction  $\xi$  of the quark momentum  $p_a$ . The produced parton  $c$  is absorbed by the proton sea quarks.

Because of the universality of the PDFs, it is possible to use PDFs extracted from well-known processes to obtain predictions corresponding to different scales or different final states. The modern PDF sets, which are available through the LHAPDF library [27], include data from several experiments and in several different final states. In Figure 1.4, the NNPDF3.1 PDF is shown.

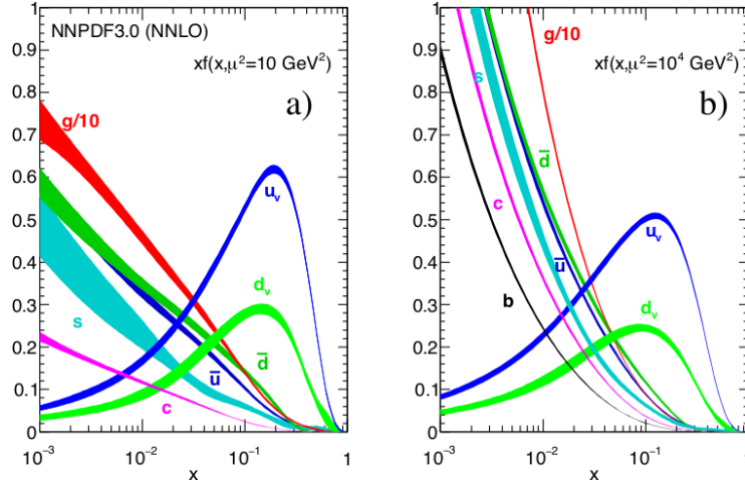


Figure 1.4: The scale dependence of the NNPDF3.1 set of PDFs [28].

### 1.3 Drell-Yan process

The process of creation of lepton pairs from hadron collisions was proposed by Sidney Drell and Tung-Mow Yan in 1970 [29] to test the parton model. This process, named Drell-Yan after them, consists of the annihilation of quark-antiquark pairs from hadrons with the creation of a Z boson or a virtual photon, which decays into a lepton-antilepton pair. The lowest-order Feynman diagram for this process is shown in Figure 1.5.

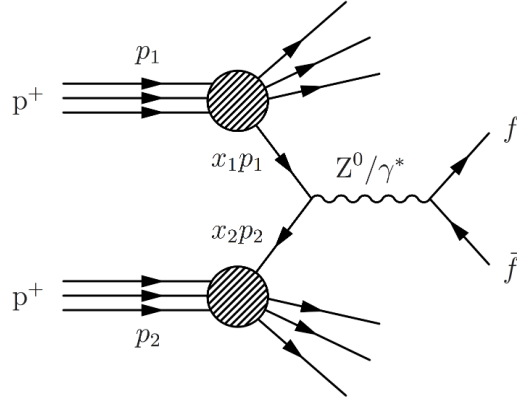


Figure 1.5: Feynman diagram of the Drell-Yan process.

The cross section of the Drell-Yan process, following the Eq 1.30, can be written as:

$$\sigma(pp \rightarrow l^+l^-) = \sum_q \int \int dx_q dx_{\bar{q}} f_q(x_q, \mu_F^2) f_{\bar{q}}(x_{\bar{q}}, \mu_F^2) \sigma(q\bar{q} \rightarrow l^+l^-), \quad (1.33)$$

where the  $\sigma(q\bar{q} \rightarrow l^+l^-)$  is the cross section of the lepton-antilepton production from a quark-antiquark pair. The renormalization and factorization scale for the Drell-Yan process is usually chosen to be equal to the invariant mass of leptons ( $\mu_F^2 = \mu_R^2 = M_Z^2$ ).

Following the perturbative QCD, the partonic cross section can be expanded in series with respect to the coupling constant  $\alpha_s$ :

$$\sigma(q\bar{q} \rightarrow l^+l^-) = \sigma_{LO} + \alpha_s \sigma_{NLO} + \dots \quad (1.34)$$

The partonic cross section can be calculated using the Matrix Element of the Feynman diagram. For the leading order (LO), on the Feynman diagram, it can be seen that there are no strong interactions. For higher orders, gluons can be exchanged between quarks. In Figure 1.6 examples of LO and next-to-leading order (NLO) diagrams are shown.

When performing calculations for higher order diagrams, loops of quarks and gluons must be included. With these additional terms, the logarithms of the form  $(\alpha_s \log Q^2/M^2)^n$ , where  $n$  is the number of quark (gluon) loops and  $M$  the renormalization point of  $\alpha_s$ , are introduced. The current best computations available are at NNLO.

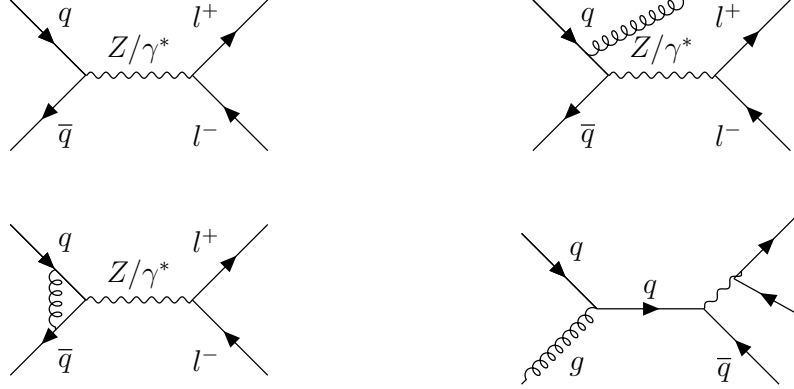


Figure 1.6: The Drell-Yan process at LO (a), at NLO with initial state radiation (b), at NLO with a gluon loop at the initial state (c), and at NLO with a quark-gluon initial state and with outgoing hadronization.

Besides the hard scattering, the process that involves large momentum transfer between the colliding particles, several different effects can occur in proton-proton collisions. The part of energy of the partons from the hard interaction can go to the radiation of gluons and photons. Radiated gluons and photons can create additional quarks and lepton pairs. Since gluons are particles that can self-interact, they can produce additional gluons or quark-antiquark. Such radiation created from partons is called parton shower. The radiation coming from the initial state particles is called initial state radiation (ISR); similarly, the radiation from final state particles is called final state radiation (FSR). Colored particles created in showers recombine to create hadrons, as explained in Section 1.1.1.

In high-energy proton-proton collisions, there are also additional soft interactions coming from the remaining partons of the protons participating in the interaction. These secondary interactions are called underlying event. The underlying event is a common name and denote for multiple parton interaction (MPI) and beam remnant (BBR) interaction.

## 1.4 Monte Carlo Simulation

The simulations that include physics processes and the response of the detector have a very important role in the measurements and discovery. In order to compare what is measured in the experiment with the theoretical prediction, it is necessary to simulate the proton-proton interactions, the interaction of the produced particles with the detector, and the response of the detector including its electronic.

To simulate processes from proton-proton collisions, which implies the computation of large integrals, the Monte Carlo (MC) technique is used [30]. To describe the typical high-energy event, event generators typically include the simulation of several physics effects that are schematically shown in Figure 1.7.

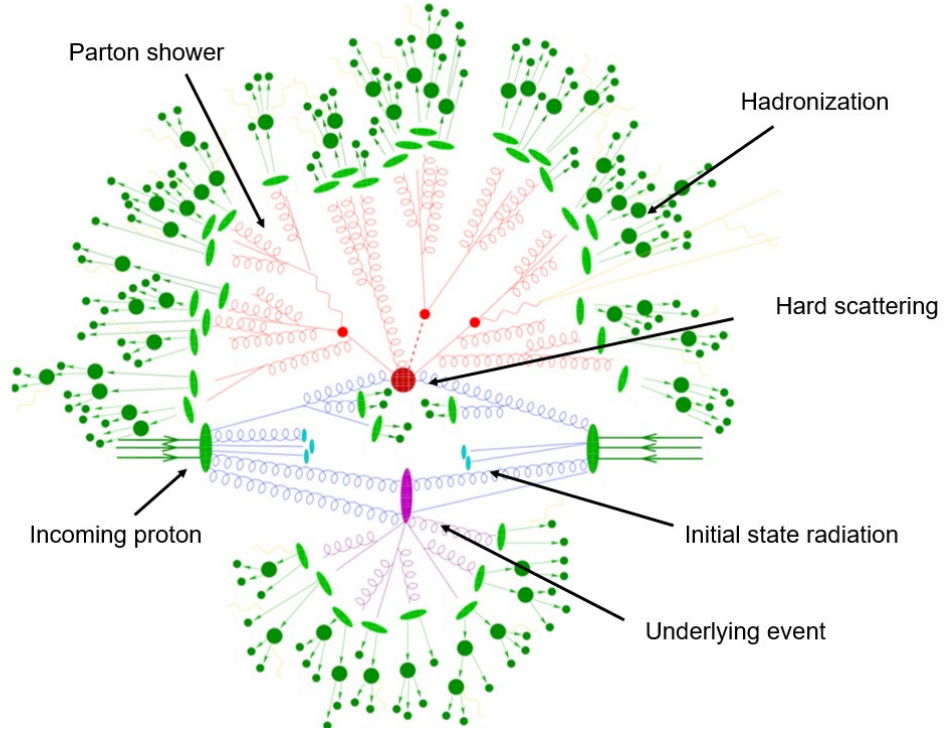


Figure 1.7: Sketch of a proton-proton collision as simulated by a Monte Carlo event generator [31].

The simulation of a proton-proton collision event is performed in the following steps:

- The evolution of an event in simulation starts with the two beam particles

that are colliding. The initial particles are generated according to the set of PDFs which provides information about the partonic substructure.

- The partons from beams start irradiating and the initial state shower is simulated.
- The incoming partons (one from each beam) enter the hard interaction and the outgoing particles are produced. In this step, according to the nature of the hard interaction, the main characteristics of the event are determined.
- In the hard process, short-lived resonances can be created and their decay is considered in this step.
- The outgoing particles undergo radiation and final state radiation is simulated.
- The simulation of underlying events.
- The process of hadronization is simulated.
- The decay of long-life particles, such as  $\tau$  leptons or B-hadrons.

The simulation of the interaction of the particles with the detector is done using the GEANT4 software [32]. The geometry of the CMS experiment is implemented in the software, as well as the information about the active and inactive volume. The signals created by the particles as they go through the detector are simulated and reconstructed using the same algorithm used for data.

The signal samples that are used for this analysis are simulated using MADGRAPH [33] and GENEVA [34, 35] MC generators. In both cases PYTHIA8 is used to simulate the initial and final state parton shower and hadronization [36].

### 1.4.1 MADGRAPH5\_aMC@NLO

MADGRAPH5\_aMC@NLO is a framework to compute cross sections and generate parton-level events that can be showered with a MC generator like PYTHIA8 or HERWIG. It computes automatically LO and NLO cross sections and provides the tools for the PS simulations. It generates automatically the Feynman diagram up to NLO and also computes automatically the loop contributions.

In order to describe a realistic physical process, it is important to combine ME calculations suitable to simulate separated hard partons processes with the PS algorithm that populate the soft and collinear region. Combining ME and PS prediction is not a trivial task and in the recent years a lot of effort has been dedicated for its development. One of the main difficulties is the separation of the

components of the event which belong to hard process from the ones developed during its evolution. Specific (n+1)-jet event can be obtained from the collinear radiation evolution of the appropriate (n+1)-parton final state or from an n-parton configuration where hard emission during its evolution leads to the extra jet. In order to avoid double counting where same event appears once for each path, the factorisation prescription called "matching scheme" or "merging scheme" is used. Matching scheme defines for each event which of the two paths should be followed, with providing the best possible approximation to a given kinematics.

In this thesis, two different predictions from MADGRAPH5\_aMC@NLO are used:

- MADGRAPH5\_aMC@NLO including ME computed at LO for up to 4 partons. The interface with PYTHIA8 is done using the  $k_T$  MLM scheme [37];
- MADGRAPH5\_aMC@NLO including ME at NLO for up to 2 partons. The interface with PYTHIA8 is done using FxFx merging scheme [36].

Both of these predictions are interfaced with PYTHIA8 for the parton showering. The effect of the underlying event in the simulation is modeled using PYTHIA8. PYTHIA8 has set of parameters to control the behavior of the event modeling, which are adjusted to better fit some of the aspects of the data. This set of parameter is referred to as a tune. For the mentioned samples, the CP5 tune of PYTHIA8 [38] is used.

### 1.4.2 GENEVA

GENEVA is an MC generator for Drell-Yan processes that matches analytic resummation to an NNLO fixed-order prediction. The NNLL' resummation of the global event shape variable N-jettiness ( $\tau_N$ ) is used [7]. N-jettiness is a variable designed as an N-jet resolution which quantifies how much given event looks like an event with N jets in the final state.

For computing the cross section, the phase space is divided using the variable  $\tau_N$ , and the zero, one and two jets spaces can be distinguished:

$$\begin{aligned}
 \Phi_{0\text{events}} &: \frac{d\sigma_0^{MC}}{d\Phi_0}(\tau_0^{cut}), \\
 \Phi_{1\text{events}} &: \frac{d\sigma_1^{MC}}{d\Phi_1}(\tau_0 > \tau_0^{cut}; \tau_1) \\
 \Phi_{2\text{events}} &: \frac{d\sigma_2^{MC}}{d\Phi_2}(\tau_0 > \tau_0^{cut}; \tau_1 > \tau_1^{cut}).
 \end{aligned} \tag{1.35}$$

The cross section for the 0-jet case is defined by the resolution cut  $\tau_0 < \tau_0^{cut}$ , similarly the 1-jet case is defined by  $\tau_0 > \tau_0^{cut}$  and  $\tau_1 < \tau_1^{cut}$ . The inclusive cross

section for 2-jets is defined with  $\tau_0 > \tau_0^{cut}$  and  $\tau_1 > \tau_1^{cut}$ . Therefore, the cross section for some observable  $X$  can be written as:

$$\sigma(X) = \int d\Phi_0 \frac{d\sigma_0}{d\Phi_0} M_X(\Phi_0) + \int d\Phi_1 \frac{d\sigma_1}{d\Phi_1} M_X(\Phi_1) + \int d\Phi_2 \frac{d\sigma_2}{d\Phi_2} M_X(\Phi_2), \quad (1.36)$$

where  $M_X(\Phi_N)$  is a measurement's function for computing  $X$  for the N-parton final state  $\Phi_N$ .

In this thesis, the GENEVA with  $\tau_0$  resummation is used. The PDF set used is PDF4LHC15 and  $\alpha_s(m_Z)$  is set to 0.118. The showering is done using a modified version of PYTHIA (version 8.235). The underlying event is modeled with the CUETP8M1 [39] tune of PYTHIA8.

# Chapter 2

## Experimental setup

The analysis presented in this thesis is done with data obtained from proton-proton collisions at a center-of-mass energy of 13 TeV. Protons are accelerated and collided in the LHC and the particles are detected with the CMS. CMS is a very complex detector that contains several different subdetectors systems. By combining the information from the subsystems, the complete picture of one collision is obtained: the produced particles are identified and their momentum at the interaction point is measured.

In this chapter, the accelerator system and the detector with its subsystems are presented. In Section 2.1 a brief description of LHC, its performance and future plans are presented. Section 2.2 is devoted to the CMS experiment. The reconstruction of particles inside the detector is described in Section 2.3.

### 2.1 Large Hadron Collider

The Large Hadron Collider [4] is a circular accelerator designed to collide protons or heavy ions. It is the largest and most powerful accelerator ever built. The circumference of the LHC is 27 km and the accelerator is located at the border between France and Switzerland, close to Geneva (Figure 2.1), at an underground depth between 45 and 175 m. The collider is placed in a circular tunnel built for the Large Proton Electron collider (LEP) which was operating until 2000 and had an essential role in studies of the Z and W bosons properties.

The LHC project was proposed in 1994 by the European Organization for Nuclear Research (CERN); the first results with this machine were obtained in 2010. One of the main goals of the LHC was the search for the Higgs boson, which was discovered by the CMS and ATLAS detectors in 2012 [5, 6]. After the discovery, the operations continued performing precision measurements to study the properties of the Higgs boson. In addition, at the energies reached by the



Figure 2.1: Aerial view of Geneva region, with the position of LHC, sketched in yellow [40].

collisions, it is possible to perform searches for new physics beyond by Standard Model.

### 2.1.1 The design of the Large Hadron Collider

The LHC is designed to study proton-proton collisions that can reach the center-of-mass energy of 14 TeV and heavy ion collisions at the center of mass energy of up to 2.76 TeV per nucleon. In order to achieve the design energy, before entering the LHC ring, beams of particles are accelerated in the sequence of accelerating machines shown in Figure (2.2).

The protons are obtained by hydrogen ionization. The first step in the LHC injection chain is the linear accelerator LINAC2 where the energy of 50 MeV is reached. Protons then enter the Proton Synchrotron Booster (PSB) and are accelerated to 1.4 GeV. In the next step, the Proton Synchrotron (PS) groups the protons into the bunches separated by 25 ns and accelerates them to 25 GeV. Each beam is divided into 2808 bunches where each bunch consists of about  $1.15 \cdot 10^{11}$  protons. After the PS, the protons are transferred to the Super Proton Synchrotron (SPS) and are accelerated to the energy of 450 GeV. With that energy, protons are injected into the LHC ring in two opposite directions.

Besides protons, heavy ions can also be accelerated through the chain of accelerators before coming to the LHC. They enter the linear accelerator (LINAC3),

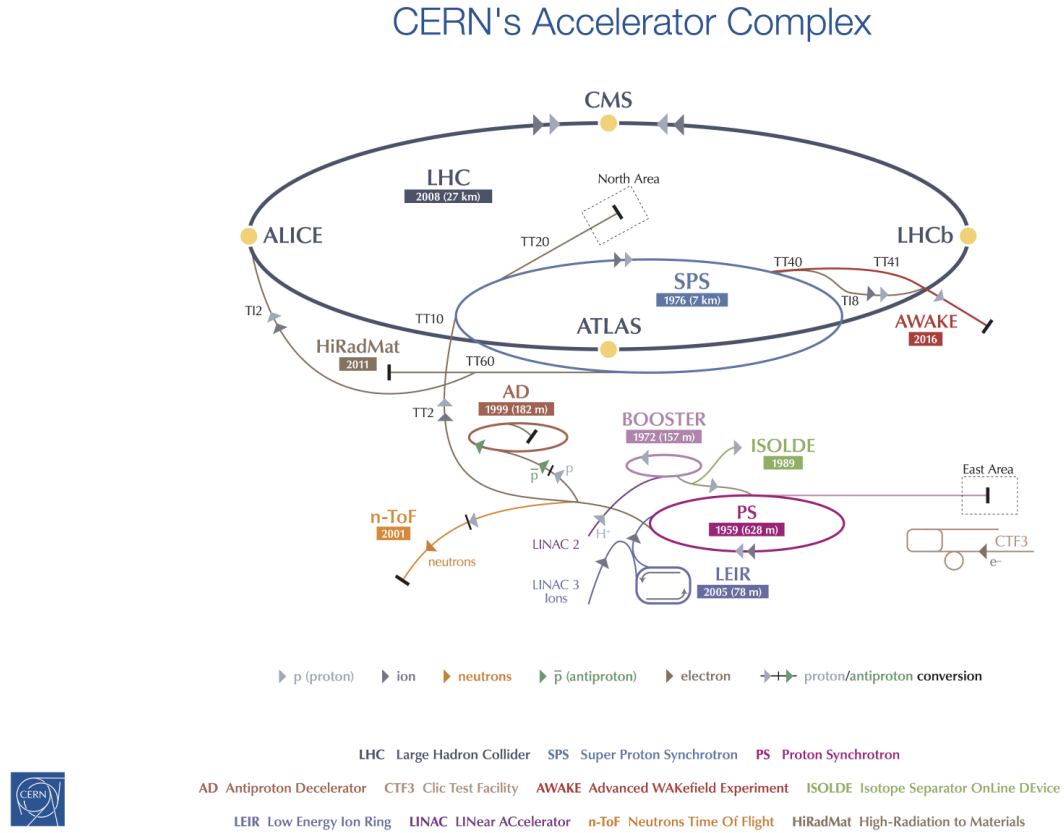


Figure 2.2: Schematic view of accelerator system at CERN [41].

Low Energy Ion Ring (LEIR) after which they are transferred to the PS, the SPS, and the LHC. Furthermore, the protons and ion beams from the PS and SPS can be sent to fixed-target experiments or to RD projects.

The LHC ring consists of eight arcs and eight straight sections. The region from the middle of one arc to the middle of the next arc is called octant (Figure 2.3). The beam crossings occur in four points along straight sections, where the main experiments are installed:

- CMS [9] and ATLAS (A Toroidal LHC AparatuS) [42] are two general purpose detectors that can study large spectrum of physics processes;
- LHCb (LHC beauty) [43] is an experiment designed for studying CP violation

and asymmetry between matter and antimatter by searching for rare decays of hadrons containing b quarks;

- ALICE (A Large Ion Collider Experiment) [44] studies heavy ion collisions through the production of quark-gluon plasma.

The other four straight sections are 3 and 7 where the system for the beam collimation is placed, point 4 which consists of two radio-frequency systems for the particle acceleration, point 6 where the beam dump extraction occurs using the combination of deflecting fast-pulsed magnets and vertically-deflecting double steel septum magnets.

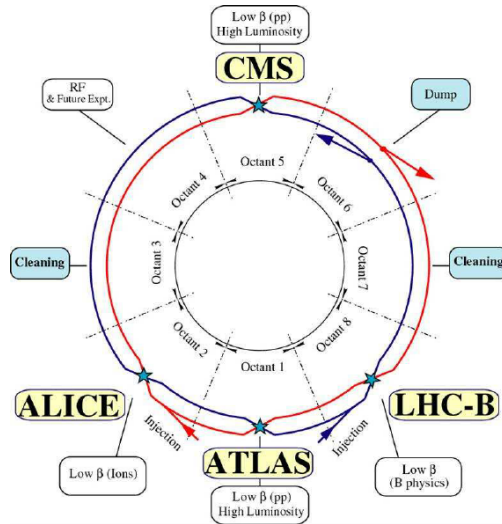


Figure 2.3: Schematic view of the LHC ring [4].

In the arcs of the LHC, superconducting magnets are placed in order to bend the trajectory of accelerated particles. The magnets are made of niobium-titanium (NbTi) and cooled down to 1.9 K using superfluid helium. The maximum magnetic field that can be reached is 8.3 T and this limits the achievable energy. In order to keep stable its trajectory, multipole magnets are placed to stabilize and focus the beam.

### 2.1.2 Performance of the LHC

The number of collisions per unit of time at colliders is defined by the relation:

$$N_{iter} = L\sigma_{inter}, \quad (2.1)$$

where  $\sigma_{inter}$  is the cross section of a given process and  $L$  is the luminosity of the machine. Luminosity is a very important parameter for colliders and, assuming that the two beams are round and with equal parameters, it can be written as:

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta^*} F, \quad (2.2)$$

where  $N_b$  and  $n_b$  are the number of protons in the bunch and number of bunches per beam respectively.  $f_{rev}$  is the revolution frequency of the LHC,  $\gamma_r$  is the relativistic  $\gamma$  factor,  $\epsilon_n$  is the normalized transverse beam emittance and  $\beta^*$  is the beta function at the collision point. The parameter  $F$  is the geometric factor that accounts for the luminosity reduction due to the crossing angle at the interaction point, and it can be written as:

$$F = \left( 1 + \left( \frac{\theta_c \sigma_z}{2\sigma^*} \right)^2 \right)^{-1/2} \quad (2.3)$$

$\theta_c$  is the crossing angle at the interaction point,  $\sigma_z$  is the RMS bunch length and  $\sigma^*$  is the transverse bunch RMS at the interaction point.

The integrated luminosity  $L$  for a specific time interval is defined as:

$$L_{int} = \int_0^t L dt. \quad (2.4)$$

In Figure 2.4, the integrated luminosity collected by the CMS experiment in 2018 is shown.

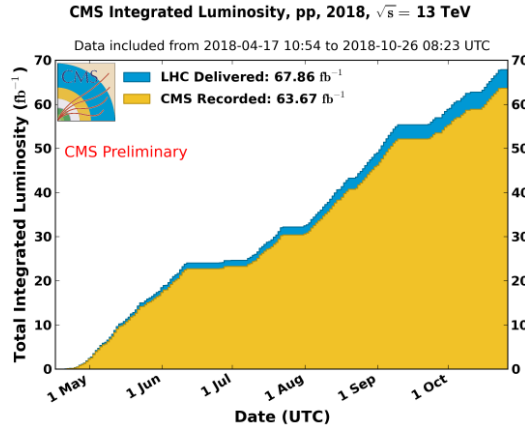


Figure 2.4: Integrated luminosity collected by CMS in 2018 [45].

At the nominal conditions at LHC, the number of bunch crossings per second is approximately 40 MHz with an average number of 22 proton-proton interaction

per bunch crossing. The average number of interactions for different data-taking years is shown in Figure 2.5 where it can be seen that the nominal values were exceeded starting from 2016 data-taking period

The effect of overlapping between the main interaction and interactions that are not coming from the hard scattering is called pileup. Since the proton bunches are separated by 25 ns and the response of the detector is not instantaneous, it is also possible to have overlapping interactions coming from different bunch crossings. This effect is known as out-of-time pileup.

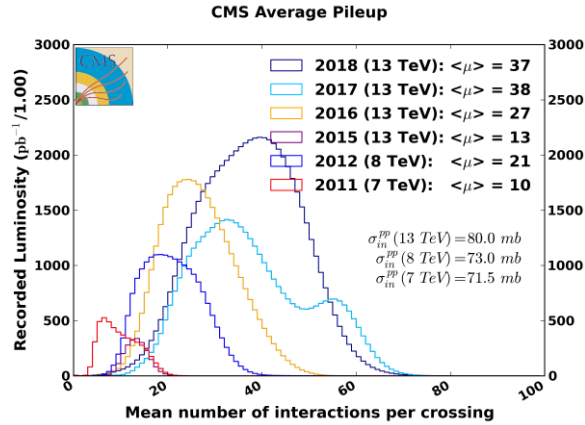


Figure 2.5: Pileup distribution observed by CMS [45].

### 2.1.3 LHC operations

The LHC program started with proton-proton collisions in 2009, while the first collisions at the center-of-mass energy of 7 TeV were recorded in 2010. Following the schedule of LHC operations (Figure 2.6), the first period of data-taking, referred to as the Run 1 period, lasted until 2012 included. During this period the energy was 7 TeV in 2010 and 2011, and 8 TeV in 2012. In the Run 2 data-taking period (2015-2018) the center-of-mass energy was increased to 13 TeV and the collected luminosity was 136 fb<sup>-1</sup>. The last stage of the so-called LHC Phase 1 will be Run 3, which will start in 2022 and last for 3 years. The total integrated luminosity that will be delivered during this period is planned to be about 300 fb<sup>-1</sup>. In between different data-taking periods, it was needed to prepare LHC and detectors for higher energy and luminosity (Long Shutdown 1 and Long Shutdown 2). Phase 1 will also be followed by a long shut-down that will allow the preparation of the detector and the machine for the High Luminosity LHC phase (HL-LHC). During this phase, which is planned to start in 2029, the integrated luminosity of 3500

$\text{fb}^{-1}$  will be reached. The center-of-mass energy will be increased from 13 TeV to 14 TeV. New data-taking conditions also include the increase of pileup by a factor of five compared to the one from Run 2. For the HL-LHC, the radiation level will be increased which is why it will be a highly challenging environment for the reconstruction of the events. In order to maintain the performance from Phase 1 and to deal with the harsh environment of HL-LHC, the detectors need to be upgraded.



Figure 2.6: The timeline of the LHC operations [46].

## 2.2 Compact Muon Solenoid

The Compact Muon Solenoid [9] is a multipurpose detector located in point 5 of the LHC ring, at Cessy in France, at around 100 m underground. It is a 22 m long and 15 m wide cylindrical detector that consists of several subdetector systems, each with a specific role in the detection of particles and measurement of their momentum at the interaction point. The central feature of CMS is a 12.5 m long superconducting solenoid magnet that provides a magnetic field of 3.8 T. The tracking system, electromagnetic, and hadronic calorimeters are placed within the solenoid. Outside the magnet, muon chambers embedded in the steel return yoke are placed. A schematic view of the CMS detector is shown in Figure 2.7.

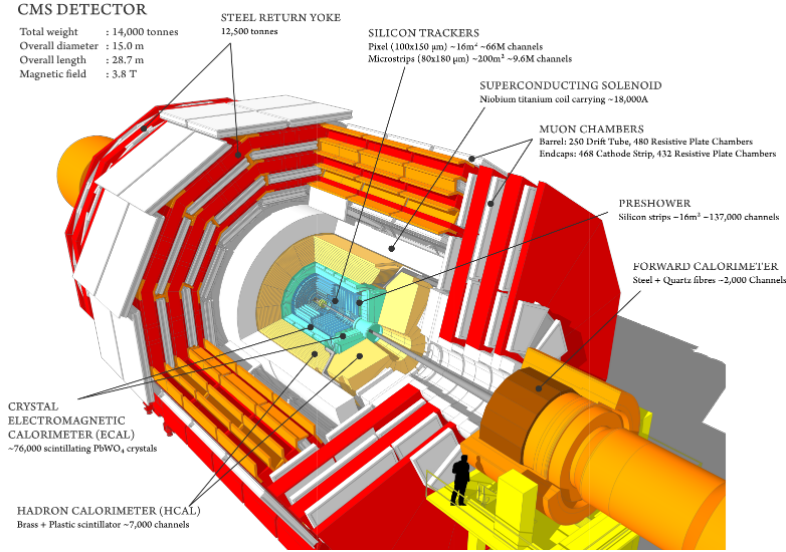


Figure 2.7: Schematic view of the CMS detector [47].

### 2.2.1 Coordinate system

In order to describe the interactions inside the detector, it is necessary to have a coordinate system. The CMS experiment uses a cylindrical coordinate system with origin at the interaction point. The  $y$ -axis is vertical and points upwards, while the  $x$ -axis is horizontal and points toward the center of the LHC ring. The  $z$ -axis point in the anticlockwise beam direction. The  $x$ - $y$  plane is called transverse plane; the projection of a vector in the transverse plane is determined by the azimuthal angle  $\phi$  with respect to the  $x$ -axis. The angle between the  $z$ -axis and the particles' momentum vector is the polar angle  $\theta$  (Figure 2.8).

According to these definitions, the momentum of a particle can be divided in two components: the longitudinal momentum ( $p_z$ ) and the transverse momentum ( $p_T$ ), defined as:

$$p_T = \sqrt{p_x^2 + p_y^2}. \quad (2.5)$$

The proton beams at the LHC carry almost no transverse momentum. Because of the momentum conservation, the vectorial sum of the transverse momenta of all the particles of an event is zero. However, the presence of neutrinos, which have very weak interaction with matter and escape the detection, leads to a momentum imbalance in the transverse plane. For this reason, the missing transverse

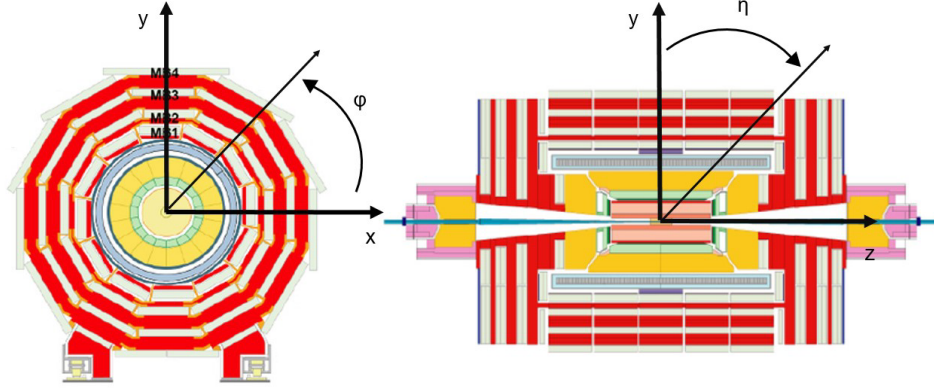


Figure 2.8: The longitudinal and transverse view of the CMS detector [48].

momentum is introduced. It is defined as the negative sum of the momenta of all reconstructed particles in an event:

$$\vec{p}_T^{miss} = - \sum_{i=1}^j \vec{p}_T^i \quad (2.6)$$

In the hadron colliders, the interaction take place in the centre-of-mass frame of the proton-proton system, which is not the centre-of-mass frame of the colliding partons. The colliding partons carry different longitudinal momentum fractions  $z$  and the rest frame of collision is longitudinally boosted. Therefore, it is more convenient to use variables that are not distorted by the center-of-mass boost. The angular distribution of particles is usually expressed using the variable called rapidity ( $y$ ):

$$y = \frac{1}{2} \ln \left( \frac{E + p_z}{E - p_z} \right). \quad (2.7)$$

In case of ultra-relativistic particles ( $p \gg m$ ) an approximation can be made and a variable named pseudorapidity ( $\eta$ ) can be used:

$$\eta = - \ln \left( \tan \frac{\theta}{2} \right). \quad (2.8)$$

The angular distance between particles is defined as:

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}. \quad (2.9)$$

The central part of the CMS detector is called “barrel”, while the forward rapidity regions are referred to as “endcaps”.

### 2.2.2 Superconducting solenoid

The superconducting magnet [49], which is the core of the CMS detector, is operating at temperature of 4.5 K and provides a magnetic field of 3.8 T within its volume. The magnet coil, with a diameter of 5.9 m, is made of a niobium-titanium (NbTi) conductor reinforced with an aluminum core. A steel return yoke guides a magnetic field of 1.8 T in the region outside of the magnet. It consists of five barrel layers and three disks for each of the endcaps. A map of the magnetic field is shown in Figure 2.9.

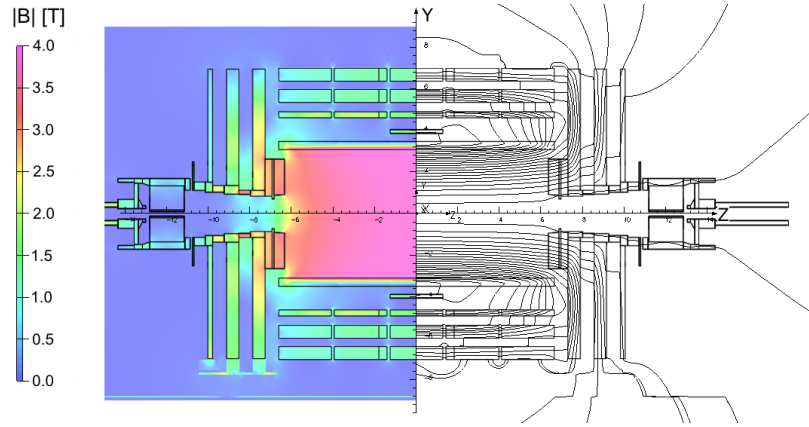


Figure 2.9: The magnetic field produced by a superconducting magnet. The right part presents the lines of the field, while on the left the strength of the magnetic field is shown [50].

The design of the whole CMS experiment is based on the solenoid. In order to have the least possible amount of material in front of the tracker system and of the calorimeters, these subdetectors are placed within the solenoid. Given this space constraint, the calorimeters are designed to have a high density so that the electromagnetic and hadronic showers can be exhausted within their volume.

### 2.2.3 Tracker

The CMS subdetector closest to the beam is the tracking system. It is composed of a pixel detector and a strip tracker and has a radius of about 1.2 m and a length of 5.6 m [51, 52].

The role of the tracking system is to measure the trajectory of particles, with the precise determination of their momenta and vertex position. It needs to distinguish the primary vertex that corresponds to the hard interaction from the additional interactions coming from the pileup. The displaced vertices, coming from decays of heavy particles, such as  $\tau$  leptons and B hadrons, should also be identified.

Because of the large luminosity and proximity to the interaction point, this part of the detector system is exposed to large radiation. This drives the choice of the materials of the tracking system. In addition, the structure of this system needs to be optimized so that the amount of material in front of the calorimeters is minimized. In order to fulfill these requirements, both pixel and strip detectors are made of silicon and have different granularity for different positions in the detector system. The longitudinal scheme of the tracking system is shown in Figure 2.10.

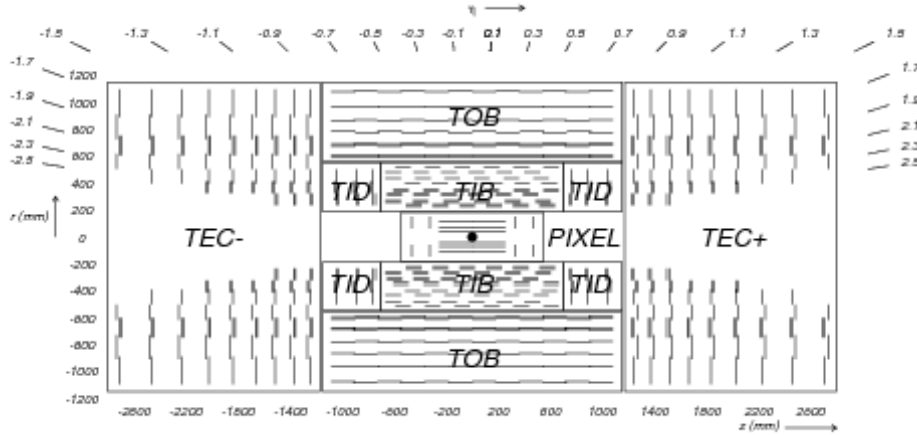


Figure 2.10: Schematic view of CMS tracker layout [53].

### The pixel detector

The pixel detectors of size  $100 \times 150 \mu\text{m}^2$  are placed closest to the interaction point ( $r \leq 10$  cm), and cover a pseudorapidity region of  $|\eta| < 2.5$ . In order to be able to cope with the increased luminosity, the pixel detector has been upgraded during the extended LHC technical stop in 2016/2017 [54]. The upgraded pixel detector has four layers in the barrel region instead of three, at radii of 2.9, 6.8, 10.9, and 16 cm, and three disks on each of the endcaps at 29.1, 39.6, and 51.6 cm from the center of the detector (Figure 2.11). The barrel region consists of 1185 segmented silicon sensor modules (BPIX), while in the forward disks there

are 672 modules (FPIX). Each of the modules has a sensor with  $160 \times 416$  pixels connected to the readout chips (ROCs).

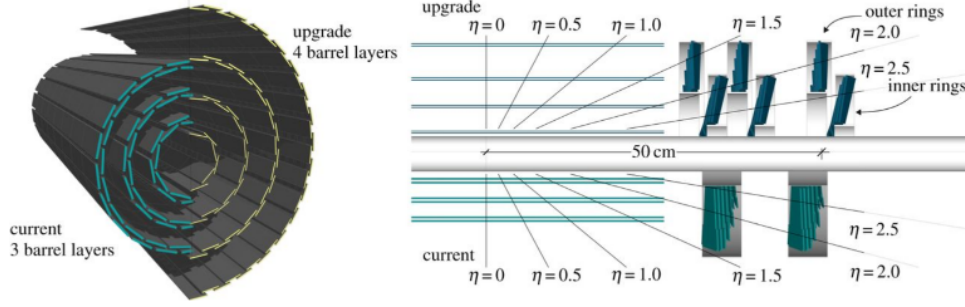


Figure 2.11: Comparison of the upgraded pixel detector with the original one [54].

### The strip detector

The pixel detectors are surrounded by a strip detector system divided into 10 different regions. In the barrel region, with  $|\eta| < 2.6$  the Tracker Inner Barrel (TIB) and Tracker Outer Barrel (TOB) are placed. The TIB covers the range  $20 \text{ cm} < r < 55 \text{ cm}$  and is made of four layers of silicon sensors. Such sensors are  $320 \text{ }\mu\text{m}$  thick and the inter-strip distance goes from 81 to  $118 \text{ }\mu\text{m}$ . The TOB, covering the region  $55 \text{ cm} < r < 120 \text{ cm}$ , consists of 6 layers of  $550 \text{ }\mu\text{m}$  thick silicon sensors. Their inter-strip distance varies from 120 to  $180 \text{ }\mu\text{m}$ . The sensors in the TIB and TOB regions have a rectangular shape with different size according to their position. In the inner region, it is  $6 \times 12 \text{ cm}^2$ , while in the outer region size it is  $10 \times 9 \text{ cm}^2$ . The endcap region of the strip detector consists of the Tracker Inner Disk (TID) and the Tracker Endcaps (TEC). In the TID, sensors with thickness of  $320 \text{ }\mu\text{m}$  are used and are divided into 3 disks. The TEC is made of nine disks where the thickness of the sensors depends on the distance from the center of the detector (from  $320 \text{ }\mu\text{m}$  to  $500 \text{ }\mu\text{m}$ ). The sensors in the endcap regions are wedge shaped.

The performance of the tracker allows precise measurements of the charged particle momentum. Momentum resolution, for muons with the high transverse momentum (100 GeV) in the central region ( $|\eta| < 1.5$ ), is about 2% and the impact parameter resolution achieved by the inner tracker is about  $15 \text{ }\mu\text{m}$ .

### 2.2.4 The electromagnetic calorimeter

The ECAL [10] is a hermetic, homogeneous, and high-granularity calorimeter dedicated to the measurement of the energy of electrons and photons. It is made of lead tungstate ( $\text{PbWO}_4$ ) crystals. The chosen material is convenient because of its fast light emission and resilience to irradiation. Due to its short scintillation time, about 80% of light can be collected in 25 ns. In addition, because of the properties such as high density ( $8.8 \text{ g/cm}^3$ ), small Molière radius (2.2 cm), and short radiation length (0.89 cm) of the  $\text{PbWO}_4$ , it was possible to construct a very compact calorimeter with high granularity. The total number of 75848 trapezoidal crystals is divided into a barrel area (61200) and two endcaps (7324 each). The longitudinal view of the ECAL is shown in Figure 2.12. The barrel region of the detector covers a rapidity range up to  $|\eta| < 1.48$ . The length of the crystals in the barrel region is 23 cm which corresponds to 25.8 radiation lengths ( $X_0$ ), and the surface that they cover is  $22 \times 22 \text{ mm}^2$ . The endcaps cover the pseudorapidity range  $1.48 < |\eta| < 3.0$ . The crystals in this region are 22 cm long ( $24.7 X_0$ ), with the front face area of  $28.6 \times 28.6 \text{ mm}^2$ .

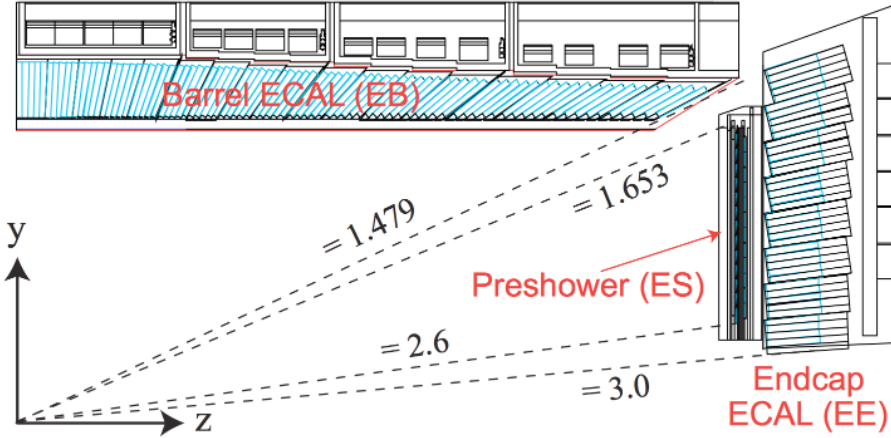


Figure 2.12: Schematic view of the longitudinal layout of ECAL detector [55].

The crystals in the barrel area are forming modules (400 or 500 crystals in one module), which are then grouped into supermodules. Each supermodule covers half of the barrel length along  $z$  and has a width of 20 along  $\phi$ . In the endcap regions, crystals are arranged into units of  $5 \times 5$  crystals called supercrystals, assembled in two semi-circular "Dees" per endcap. The longitudinal view of ECAL is shown in Figure 2.13.

In front of each endcap, a preshower detector (ES) is placed. They are covering the pseudorapidity range of  $1.65 < |\eta| < 2.6$ . The preshower detector consists of

two layers of lead plates with the radiation length of  $2X_0$  and  $1X_0$  and two layers of silicon strip detectors. The ES helps to distinguish the photons coming from the pion decay from single photons and to separate electrons from minimum ionizing particles.

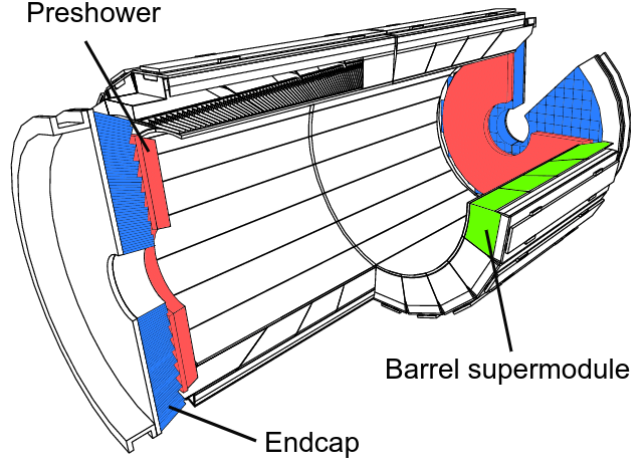


Figure 2.13: Schematic view of the ECAL structure [9].

The scintillation light emitted by the crystals is amplified and collected by fast, radiation-tolerant photodetectors. Because of the different magnetic fields and radiation flow in the different regions of the detector, two types of photodetectors are used: crystals are read out by avalanche photodiodes (APD) in the barrel region and by vacuum phototriodes (VPT) in the endcaps. Since crystal response and the APD gain are temperature dependent, the temperature is stabilized at  $18^\circ\text{C}$  using the cooling system. The target for the cooling is set by the  $(\text{PbWO}_4)$  light-yield dependence on temperature which is  $-2\%/^\circ\text{C}$ .

### 2.2.5 The hadronic calorimeter

The hadronic calorimeter [56], installed between ECAL and the solenoid, is the subdetector measuring the energy of charged and neutral hadrons by estimating the energy and position of hadronic jets. Moreover, the HCAL contributes to the missing energy determination. To perform these measurements with satisfactory precision, the HCAL needs to be hermetically closed, which means that it must cover the largest possible solid angle. The pseudorapidity range covered by the HCAL detector is up to  $|\eta| = 5.2$ . Since the ECAL covers the range up to  $\eta = 3$ , in the forward region of the HCAL the electromagnetic energy is also measured.

The HCAL is composed of alternating layers of absorber and scintillator materials and it is divided into four different regions (Figure 2.14). The Barrel Hadronic Calorimeter (HB) covers the pseudorapidity range of  $|\eta| < 1.4$ , while the endcap hadronic calorimeters (HE) cover the range up to  $|\eta| = 3$ . In these regions, the absorber material used is brass (10% Cu and 30% Zn), which has a short interaction length. The thickness of the absorber in the HB is 5.8 and 10.6 hadronic interaction lengths for  $\eta = 0$  and  $\eta = 1.3$  respectively, while in the endcaps the thickness is 10.6 hadronic interaction lengths. The scintillation material is plastic. Outside the solenoid, the outer hadronic calorimeter is placed, which covers the range of  $|\eta| < 1.4$ . The HO is used to improve the confinement of hadronic showers: it increases the thickness of the material so that they are fully absorbed before the muon system. In order to have a high hermeticity of the HCAL, the forward hadronic calorimeter (HF) is placed at a distance of 11 m from the interaction point, to cover the range up to  $|\eta| = 5.2$ . The HF is Cherenkov-based calorimeter composed of steel absorbers and radiation hard quartz fibers.

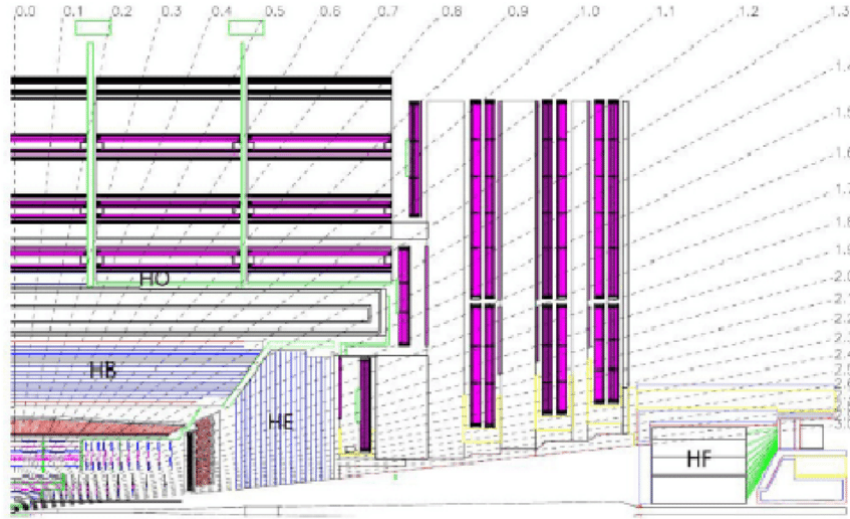


Figure 2.14: Schematic view of the longitudinal layout of HCAL detector [9]

The photodetectors used for the readout of the HB and HE subdetectors are hybrid photodiodes. The HO calorimeter uses silicon photomultipliers (SiPM), while in the forward region the Cherenkov light is read out by photomultiplier tubes. The energy resolution of HCAL + ECAL is measured in the test beam with high-energy pions and it is parametrized as:

$$\frac{\Delta E}{E} = \frac{84.7\%}{\sqrt{E}} \oplus 7.4\%. \quad (2.10)$$

The first term in the formula represents the stochastic effects, while the constant term corresponds to the effects independent of the energy, such as imperfect calorimeter calibration.

### 2.2.6 Muon system

Since the muons are particles that can penetrate several meters of material without being absorbed, the subdetector systems of CMS described above can not stop them. The muons are detected in the tracker system, they lose a small amount of their energy in the calorimeters and are finally detected in the muon system placed at the outermost part of the experiment [57]. The muon system, shown in Figure 2.15, is made of three different types of detectors: Drift Tubes (DT) in the range  $|\eta| < 1.2$ , Cathode Strip Chambers (CSC) for  $0.9 < |\eta| < 2.4$  and Resistive Plate Chambers (RPC) that cover the pseudorapidity range up to  $|\eta| = 1.6$ .

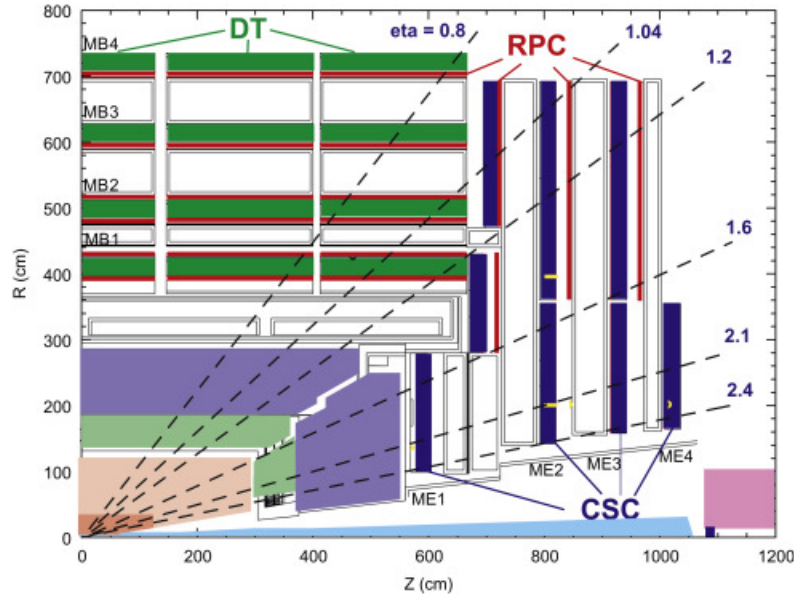


Figure 2.15: Schematic view of the longitudinal layout of the muon detector system [58].

Drift tube chambers are placed in the central part of the detector and they are organized in four stations intertwined with the return yoke, which provides a uniform magnetic field. The four stations are placed at a distance of 4.0, 4.9, 5.9, and 7 m from the interaction point and consist of 250 drift tubes in total. The basic constituents of the DT system are drift cells filled with a mix of Ar (85%) and CO<sub>2</sub> (15%). The dimensions of drift cells are  $13 \times 42 \text{ mm}^2$ . Each of the

cells consists of an aluminum I-shaped cathode and a stainless steel anode. After passing through a cell, a muon ionizes the mixture of gas; thus, the position of the muon can be determined by measuring the drift time of the electrons. The spatial resolution of DT in the  $(r, \phi)$  plane ranges from  $80 \mu\text{m}$  to  $120 \mu\text{m}$ , and in  $z$  direction from  $200 \mu\text{m}$  to  $300 \mu\text{m}$ .

In the endcap region, where the magnetic fields are less uniform and the flux of particles is large, the CSC detectors are used. The cathode chambers have a trapezoidal shape, with dimensions that vary with their position in the detector (the largest ones are  $3.4 \times 1.5 \text{ m}^2$ ). In the four layers (ME1-ME4) 468 CSCs are deployed. The cathode chambers consist of seven cathode panels and six anode wire planes. The space in between layers is filled with a mix of Ar (40%),  $\text{CO}_2$  (50%), and  $\text{CF}_4$  (10%). The cathode strips are placed in the radial direction, which allows measurements in the  $r - \phi$  plane to be performed. Anodes wires, instead, are almost perpendicular to the strips so that measurements in the  $\eta$  planes can be made. When passing through the CSC, the muon creates an avalanche, which induces a charge on the cathode strips. The spatial resolution of this detector is ranging from  $30 \mu\text{m}$  to  $150 \mu\text{m}$ .

The RPC detectors, which cover both barrel and endcaps regions, are composed of 4 bakelite layers that form 2 mm wide gas gaps. These gaps are filled with a mixture of  $\text{C}_2\text{H}_2\text{F}_4$  (95.2%),  $\text{C}_4\text{H}_{10}$  (4.5%) and  $\text{SF}_6$  (0.3%). The barrel part has 480 RPCs, while in each of the endcaps 288 chambers are distributed across four layers. The space resolution obtained in RPC detectors is not as good as in other detectors from the muon system, but the time resolution is excellent and better than 3 ns.

### 2.2.7 Trigger

At a center-of-mass energy of 13 TeV and with an instantaneous luminosity of  $10^{34} \text{cm}^{-2} \text{s}^{-1}$ , the LHC produces close to 40 million collisions each second. Because of the technical limits in data processing and storage, it is not possible to store information from all events. Since not all of the events are of interest for physics studies, a trigger system [59] was developed to keep interesting events and discards the others. The CMS trigger works in two stages: the Level-1 (L1) trigger, which reduces the event rate to 100 kHz, and the High-Level trigger (HLT), which reduces the event rate to 1 kHz.

The L1 trigger operates at the hardware level and involves the calorimeters and the muon system. It performs a fast readout and selects events that have ionization signals in the muon system or energy clusters in calorimeters. The information from calorimeters and muon detectors are processed in two separate flows, as shown on the flowchart of the L1 trigger in Figure 2.16. Muon candidates are built in the muon trigger flow, while the calorimeter trigger builds photons, electrons, jets, tau

lepton candidates, as well as complex quantities (such as the missing transverse momentum and jets energy sums). The information is then combined and passed to the global trigger, where the final decision is made. The final selection is based on a menu that selects the event based on specific criteria involving conditions on the objects, and passes the event to HLT paths for more sophisticated selections. The time used to make a decision about rejecting or selecting events is less than 4  $\mu\text{s}$ .

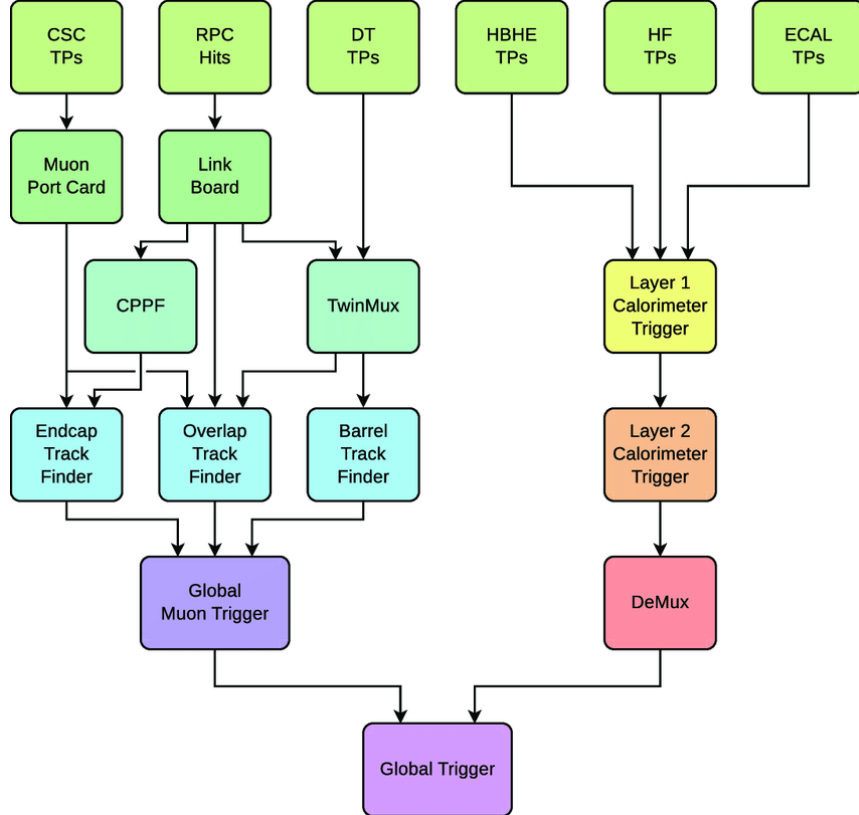


Figure 2.16: A flowchart of the L1 trigger [60].

For the events that pass the L1 Trigger, the full readout of the CMS detector is performed and the HLT, which is implemented at the software level, is used to further suppress the event rate. The HLT object reconstruction is based on the software used for the standard event reconstruction, but the configuration is optimized in order to have a faster reconstruction. To deal with the information coming from the L1, HLT runs on a computer farm with 13000 CPUs. The HLT objects are using L1 objects as a seed. A given sequence of requirements performed

by the HLT reconstruction algorithm is called an "HLT path" and it is used for selecting events of interest for a specific physics analysis.

For both L1 and HLT, it can happen that the trigger rate is too high. In that case, only a fraction of the events is accepted and the specific trigger is said to be "prescaled".

## 2.3 Physics object reconstruction

The particles created in the proton-proton collision go through the detector leaving a specific trace in the subdetector systems of CMS. The interactions of different types of particles in the subdetector systems are shown in Figure 2.17: a muon produces hits in the tracker and muon system; an electron produces hits in the tracker and an electromagnetic shower in the ECAL; charged hadrons produce hits in the tracker and energy deposits in the HCAL; photons and neutral hadrons, finally, do not produce hits in the tracker, but only electromagnetic showers and energy deposits in the ECAL and HCAL, respectively.

The reconstruction of particles in the CMS experiment is based on the Particle Flow (PF) algorithm [61]. This algorithm collects information from the subdetector systems and combines them to infer the nature of the particles in the event and reconstruct them, and to build higher level objects and quantities.

### 2.3.1 Track and vertex reconstruction

The reconstruction of the trajectory of the charged particles in the tracker system uses iterative methods to achieve high tracking efficiency [63]. The first step in the track reconstruction is the initial seed generation, where the initial track candidate and trajectory parameters are determined. At this step, the information about the position of the beam spot and interaction position is needed. This information is obtained using a very fast track and vertex reconstruction algorithm that uses only hits from the pixel detector. Once the seeds are found, the hits from all tracker layers are gathered along the trajectory in the track finding (pattern recognition) step. Compatible hits are found using a Kalman filter [64]. A final fit is then performed which provides the best possible estimate of the track parameters: origin, transverse momentum, and direction. Quality flags are set based on specific criteria for the tracks such as the number of hits or the quality of the fit. If a track fails to meet the required quality criteria, it is discarded.

The interaction point (vertex) is determined using reconstructed tracks which are compatible with the beam spot [63]. After selecting compatible tracks, a deterministic annealing algorithm is used to cluster the tracks that originate from the same vertex. Those tracks are then fitted using the adaptive vertex fitter to

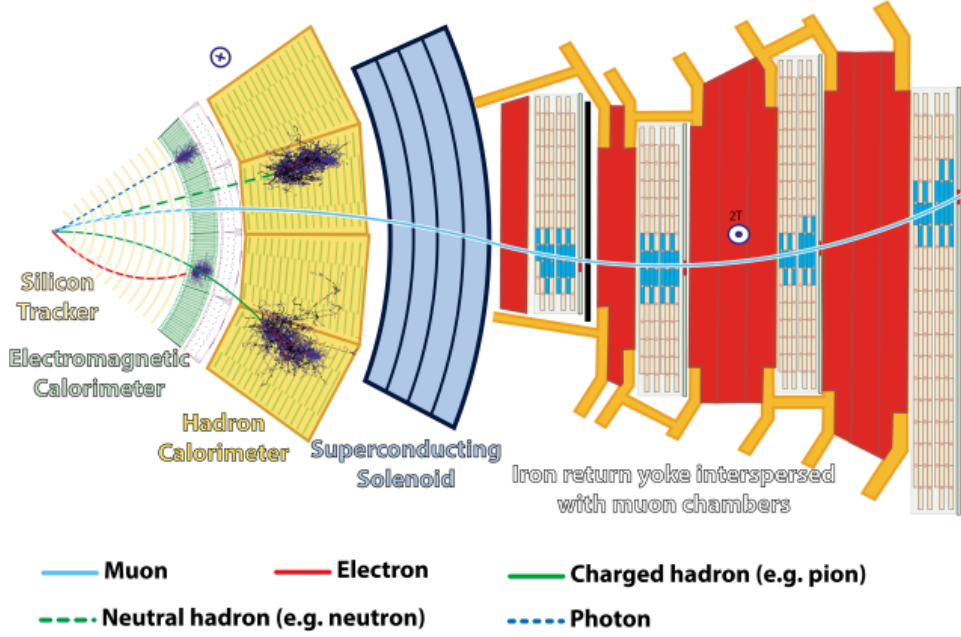


Figure 2.17: Slice of the CMS detector with the particles signature in the sub-detectors [62].

compute the best estimate of the vertex parameters. The vertex with the highest probability to originate from the hard interaction is called the primary vertex (PV) and it is determined as the one with the maximum sum of the transverse momentum of associated tracks.

### 2.3.2 Calorimeter clustering

The energy and the position of electrons, photons, and charged and neutral hadrons is measured in calorimeters, where they deposit their energy and form electromagnetic or hadronic shower. A clustering algorithm is used for the particle reconstruction in the calorimeters. The main goal of this algorithm is to reconstruct the energy and direction of stable neutral particles such as photons and neutral hadrons, to reconstruct and identify electrons and accompanying bremsstrahlung photons, and to improve measurement of charged hadrons for which the track parameters were not determined accurately. The clustering algorithm starts with determining a "seed", which is the calorimeter cell with energy larger than a specific threshold and larger than the energy deposited in surrounding cells. After that, topological clusters are grown starting from the seed and grouping iteratively cells that have at least one corner in common with those already in the

cluster. Their energy also has to be above a fixed threshold. Finally, an iterative expectation-maximization algorithm is used to reconstruct the clusters within a topological cluster.

### 2.3.3 Muons

The reconstruction of muons is done using information from the tracker and from the muon system [65]. Depending on the subdetector system where they are detected, muons can be:

- stand-alone muons - reconstructed using the information from muon chambers only
- tracker muons - reconstructed in the tracker, where the track is compatible with at least one track segment in one of the muon detectors;
- global muons - the track from the inner tracker is compatible with the one in the muon chambers.

About 99% of the muons, that are produced in the geometrical acceptance of the muon system, are reconstructed as a global muon or a tracker muon or as both. Global muon and tracker muon that have the same tracker track are merged together.

The momentum resolution for the muons with momenta up to 100 GeV is 1% in the barrel region of the detector and 3% in the endcaps.

### 2.3.4 Particle flow algorithm description

A particle created in a collision can leave traces in a few of the subdetector systems. It can be detected as a track in the tracker system, create calorimeter clusters, and (or) a track in the muon system. In order to avoid double counting of particles, the specific link algorithm is used to build the physics objects. The quality of the link is estimated by defining the distance between elements in the event. If the link satisfies the criteria, the flow of reconstruction starts.

The algorithm starts with identifying muons. The compatible muon track and calorimeter deposit are assigned to this particle and removed for further consideration. After that, the reconstruction of electrons is performed. If an electron passes identification criteria, its track and ECAL cluster are removed from the algorithm. Then the charged hadrons are determined by matching the clusters created in the ECAL and the HCAL, which are compatible with the momentum measured in the tracker system. If there are no matched signals from the tracker, the neutral hadron or photon is created. Finally, the photons are reconstructed from the remaining clusters created in the ECAL.

## 2.4 CMS upgrade for HL-LHC

In order to maintain a good performance and to deal with the harsh environment of the HL-LHC described in Section 2.1.3, the CMS detector needs to be upgraded. The Phase 2 upgrade for CMS includes:

- The replacement of the entire tracker with a new detector with higher granularity and better resistance to radiation. In the inner tracker, the new pixels will be smaller and the sensors will be thinner, which will improve the tracks separation and the impact parameter resolution. Since there will be additional pixel disks (up to 10), the pseudorapidity coverage of the tracker will be extended to  $|\eta| = 4$ . The outer tracker will be lighter, with shorter silicon sensor strips than the current ones. Furthermore, with the new design of the module, the L1 trigger information will be available at the tracker level [54].
- The replacement of the calorimeter endcaps with a silicon-based High Granularity Calorimeter (HGCAL). The electromagnetic part will have tungsten absorbers, while in the hadronic part the absorber will be made of lead [66].
- New detectors in the muon system. In the region,  $1.5 \leq |\eta| \leq 2.4$  with the four chambers of the CSC, new four stations will be added equipped with Gas Electron Multiplier (GEM) chambers and Resistive Plate Chambers (RPC). This will improve the momentum resolution of the standalone muon trigger, the matching with the tracks for global muons; it will also provide a better timing resolution. The acceptance of the muon system will be increased by additional GEM detectors in the space behind the new endcaps [67].
- The new electronics for the L1 trigger system. The latency of the L1 trigger will be increased from the current  $3.4 \mu\text{s}$  to  $12.5 \mu\text{s}$ , which will provide enough time for track reconstruction and matching tracks to the information from the calorimeters or the muon systems. In addition, the L1 trigger rate will increase from 100 kHz to 750 kHz [68]. Because of the higher event rate, the subdetectors electronics will also need to be upgraded. The details about the ECAL readout electronics upgrade are in Section 3.6.

# Chapter 3

## ECAL Calibration

The CMS electromagnetic calorimeter measures the energy of photons and electrons with high resolution, that played important role in the discovery of the Higgs boson in the  $H \rightarrow \gamma\gamma$  channel during the LHC Run 1 [69]. The performance of the ECAL is also crucial for many analyses involving physics beyond the SM such as high-mass resonances or detection of final states with energetic electromagnetic particles [70], and for SM precision measurements [71].

The energy and signal reconstruction in the ECAL are described in Sections 3.1 and 3.2. Since the response of the ECAL varies with time due to the crystal transparency loss induced by irradiation, constant calibration of the detector was performed during Run 1 and Run 2 in order to keep the excellent performance of this detector. The laser system used to monitor the crystal transparency is described in Section 3.3.

Part of my PhD work was devoted to the intercalibration of ECAL channels, which is described in Section 3.4. Using  $Z \rightarrow e^+e^-$  method (Section 3.5), I delivered the intercalibration constants for all three years of the Run 2. In addition, I worked on resolution studies for Run 2 and on comparisons of the performance of ECAL during Run 1 and Run 2. On behalf of the collaboration I presented the results for calibration and performance during LHC Run 2 at The APS Division of Particles and Fields (DPF) Meeting, and plans for electronics readout upgrade at International Workshop on Radiation Imaging Detectors (iWoRiD) [72]. The upgrade plans are described in Section 3.6.

### 3.1 Energy reconstruction

At the CMS experiment, electrons are reconstructed by combining the measurement from the tracker and the ECAL [73]; the reconstruction of photons, in absence of conversion in the tracker material, is performed only with the ECAL [74].

Electrons and photons produce a shower of secondary particles in the  $\text{PbWO}_4$  crystals. On average, 94% of the total energy of the electron or photon passing through the calorimeter is deposited in a  $3 \times 3$  crystal matrix centered on the hit crystal, and 97% in a  $5 \times 5$  crystal matrix. Since electrons and photons can undergo respectively bremsstrahlung and photon conversion in the tracker material and because of the presence of a strong magnetic field, the electromagnetic shower is spread over more crystals. To take into account the spread of energy, the reconstruction is done using a clustering algorithm. This algorithm starts with grouping together crystals with an energy greater than a specific threshold ( $\approx 80$  MeV for EB and  $\approx 300$  MeV for EE) in one cluster. The cluster which contains most of the energy deposited in a specific region is called the seed cluster. Because of the showering of electrons and photons when transversing the tracker, clusters are grouped into the superclusters (SC) belonging to the original electron or photon. A SC is formed by grouping together clusters from the geometric area around the seed. The superclustering step combines two different algorithms to form SCs: the "mustache" algorithm, which uses information from the preshower detector and is used to measure low-energy deposits; and the "refined" algorithm which uses information from the tracker and is used for determining the electron and photon object quantities [75]. The reconstruction of the energy of photons and electrons is rather complete (around 95%) even for the electrons or photons that undergo bremsstrahlung or conversion in the material in front of the ECAL.

The energy in a supercluster can be expressed as:

$$E_{e/\gamma} = F_{e/\gamma} \left[ G \cdot \sum_i C_i \cdot S_i(t) \cdot A_i + E_{ES} \right]. \quad (3.1)$$

The sum runs over all the crystals that belong to the SC and the terms represent:

- $F_{e/\gamma}$  - supercluster energy correction that accounts for several effects such as biases in the energy reconstruction due to the geometry of the detector, leakage of the electromagnetic shower, and the clustering of energy emitted by bremsstrahlung or photon conversions. The small difference in shower development of electrons and photons is also taken into account;
- $G$  - the conversion factor between ADC counts and energy, prior to any radiation damage. Two different values are derived for barrel and for endcaps from  $Z \rightarrow e^+e^-$  events ( $\approx 40$  MeV/ADC in EB and  $\approx 60$  MeV/ADC in EE);
- $C_i$  - the intercalibration term which equalizes relative differences in the crystal response;

- $S_i(t)$  - the time-dependent transparency correction that takes into account the changes in response of crystals;
- $A_i$  - the reconstructed amplitude in ADC counts;
- $E_{ES}$  - the energy deposited in the preshower, set to zero in the region not covered by this detector;

## 3.2 Signal reconstruction

The scintillation light emitted by  $\text{PbWO}_4$  crystals is measured by photodetectors and read out as an analog signal by the front-end electronics. The electrical signal from the photodetectors is pre-amplified, shaped, and processed by a multi-gain pre-amplifier. The output is sampled at 40MHz and digitized by a 12-bit ADC. Ten consecutive samples are stored in the recorded events. For estimating the signal amplitude, during the LHC Run 2, a new algorithm called "multi-fit" was developed [76]. The method used during Run 1 was not suitable because of the increased pileup during Run 2. The multi-fit algorithm estimates the in-time signal amplitude and up to 9 out-of-time amplitudes by the minimization of a  $\chi^2$ , given by:

$$\chi^2 = \sum_{i=1}^{10} \frac{\sum_{j=1}^M (A_j p_{ij} - S_i)^2}{\sigma_{S_i}^2}, \quad (3.2)$$

where  $A_j$  are the amplitudes of up to ten sampled signals. For each bunch crossing  $j$ , the so-called pulse templates  $p_{ij}$  have the same shape of the in-time signal and differ by a 25 ns shift in time. The total electronic noise  $S_i$  and its associated covariance matrix  $\sigma_{S_i}$  are measured from dedicated pedestal runs. The least-square method is used to perform the  $\chi^2$  minimization. The fitted amplitudes are all constrained to be positive. Examples of a fit for signals in the barrel and in the endcap are shown in Figure 3.1, for an average pileup of 20 and for 25 ns bunch spacing. The red distributions represent the in-time pulses while the other light colors represent the out-of-time pulses with positive amplitude. The distribution in dark blue corresponds to the sum of all fitted contributions.

## 3.3 Laser monitoring

The main source of ECAL response degradation is the damage to crystals and to photodetectors due to the high radiation level produced at LHC collisions. The optical transmission within the crystals is affected by the color centers produced by the ionizing electromagnetic radiation, which leads to the reduction of transparency of the crystal. The color centers partially anneal with thermal energy,

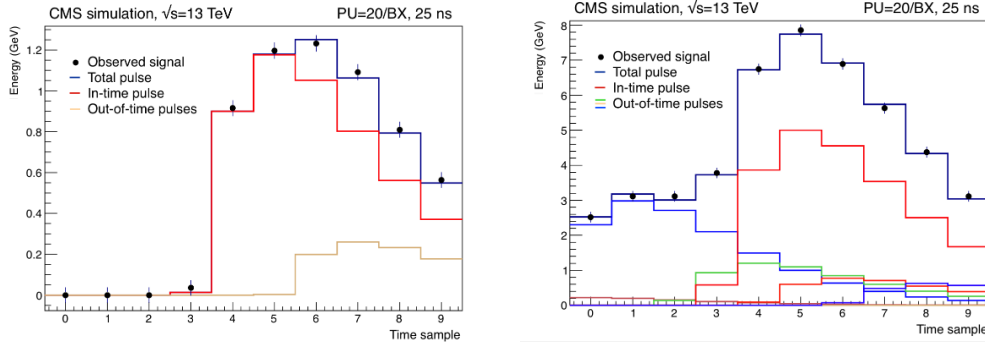


Figure 3.1: Example of fitted pulses for simulated events with 20 average pileup interactions and 25 ns bunch spacing, for a signal in the barrel (left) and endcap (right) [77].

which allows the light output to be partially recovered in absence of LHC collisions, e.g. in the time between fills and during technical stops.

The crystal transparency is monitored during data-taking with dedicated laser system, which provides a measurement of the crystal and photodetector response every 40 minutes [78]. A blue laser (447 nm) is used to measure and correct for changes in crystal and photodetector response. The laser light is directed to crystals using a system of optical fibers and splitters. At the last splitting stage, a bundle of 200 fibers, called "harness", directs the light into 200 crystals and two PN diodes. The transparency variation is measured through the ratio between the amplitude measured by the photodetectors reading the crystals and the one measured with the PN diodes. The evolution of the relative crystal response to laser light for Run 1 (2011-2012) and Run 2 (2015-2018) is shown in Figure 3.2. By construction, the first point is set to one. The response variation is shown for different pseudorapidity regions. The loss of transparency is larger at the very forward region because of the higher absorbed dose of radiation in that part of the detector. The observed response variation is up to 10% in the barrel region, and 50% at  $\eta = 2.5$  which is the limit of the tracker acceptance, while the changes go up to 98% in the region closest to the beam pipe. The recovery of the transparency during the periods without collision is also visible.

The transparency is measured continuously, in parallel with the collision data-taking. The time-dependent corrections are derived on the fly from the measured response to the laser light. The correction for the  $i$ -th crystal at time  $t$  is given by:

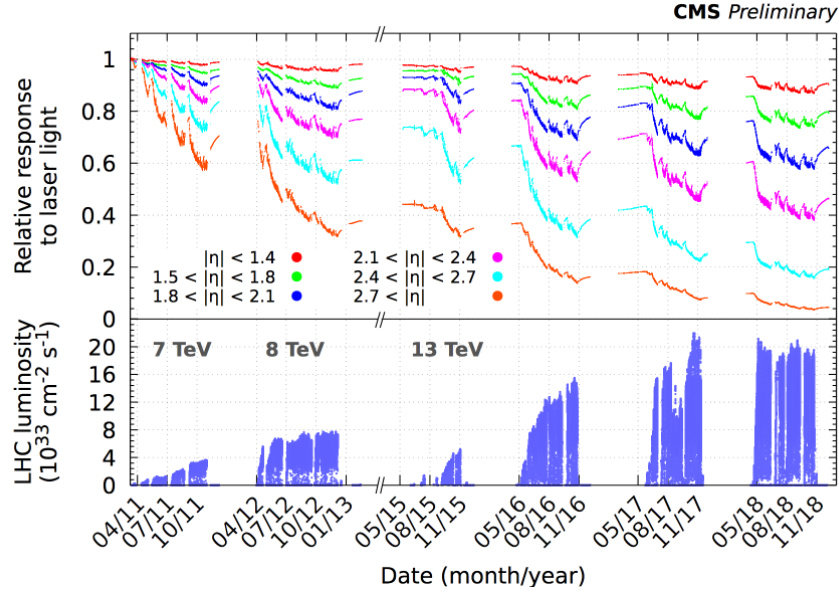


Figure 3.2: Relative response to laser light (440 nm in 2011 and 447 nm from 2012 onwards) injected in the ECAL crystals, measured by the ECAL laser monitoring system, averaged over all crystals in bins of pseudorapidity. The bottom plot shows the delivered instantaneous LHC luminosity [79].

$$LC_i(t) = \left[ \frac{R_i(0)}{R_i(t)} \right]^\alpha. \quad (3.3)$$

The response  $R_i(0)$  is the response to the laser light at the beginning of each year of data-taking and the parameter  $\alpha$  takes into account the difference in the optical paths between laser and scintillation light. The laser corrections are used in the event reconstruction which occurs 48h after a run is ended. The stability of the energy scale is monitored using the diphoton invariant mass in  $\pi^0 \rightarrow \gamma\gamma$  decays and by comparing the energy measured in the ECAL to the track momentum measured in the tracker ( $E/p$ ) in  $W \rightarrow e\nu$  events. The stability plot obtained with the  $\pi^0$  method using the 2017 dataset is shown in Figure 3.3, where the ratio of the diphoton mass and  $\pi^0$  mass is shown before and after applying the laser corrections.

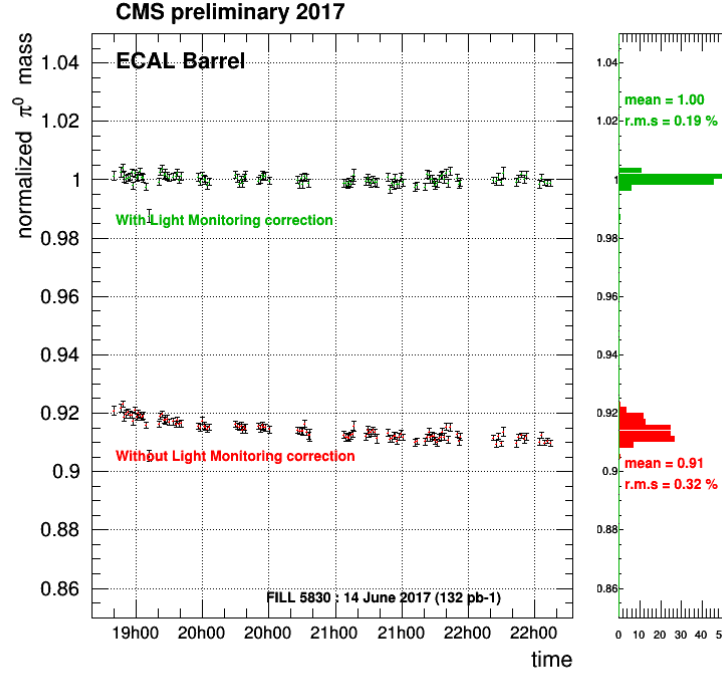


Figure 3.3: The stability of the relative energy scale measured from the invariant mass distribution of  $\pi^0 \rightarrow \gamma\gamma$  decays in the EB and plotted as a function of time, over a period of 3 hours during an LHC fill [80].

### 3.4 ECAL intercalibration

The electromagnetic calorimeter is designed to provide high energy resolution for photons and electrons. The resolution can be expressed by the formula:

$$\frac{\sigma(E)}{E} = \frac{S}{\sqrt{E}} \oplus \frac{N}{E} \oplus C, \quad (3.4)$$

where  $S$  is a stochastic term that includes statistical effects,  $N$  corresponds to the electronic noise contribution, and the  $C$  parameter is a constant term related to the calibration of the calorimeter. In order to measure the performance of the ECAL in ideal conditions, a test beam was performed before its installation in CMS [81]. During the test beam, there is no magnetic field and no material in front of the ECAL and the electron beams with energy ranging from 20 to 250 GeV are used. The parameters obtained are :  $S = 2.8\%$ ,  $N = 12\%$ , and  $C =$

0.3%. For high-energy photons, the constant term is expected to dominate the energy resolution. The design performance of the ECAL corresponds to an energy resolution of 1% for photons from a  $H \rightarrow \gamma\gamma$  decay, which is why the constant term  $C$  must be kept below 0.5%. The main contribution to the constant term comes from the non-uniformity of the light collection, instabilities in the ECAL operations, and the accuracy of the intercalibration constant.

The intercalibration is performed continuously. Several independent methods have been developed to compute the IC constants, and the results are combined to provide the final number for each crystal. These methods are:

- the  $\phi$ -symmetry method - based on the assumption that the total deposited transverse energy ( $\sum E_T$ ) is the same for all the crystals at the same pseudorapidity ( $\eta$  ring). The average energy is equalized in channels in the same  $\eta$  region; the intercalibration in  $\phi$  is performed by comparing the  $\sum E_T$  deposited in one crystal with the total transverse energy collected by the crystals in the  $\eta$ -ring. Although this method can profit from a large amount of data, it provides lower precision compared to other methods because of the presence of  $\phi$  asymmetric material in front of the ECAL. Therefore, the  $\phi$ -symmetry method is typically used for validation or to monitor the relative time drift.
- the  $\pi^0$  method - it uses  $\pi^0 \rightarrow \gamma\gamma$  decays to calibrate the ECAL response. The  $\pi^0$  invariant mass distribution is fitted with a Gaussian function for the signal and with a fourth order polynomial for the background. The intercalibration constants are updated iteratively in order to correct the fitted mass value in each channel. Since this method is based on low energy photons, its precision is limited by the electronic noise and by pileup events, with non-negligible background contributions. Because of these limitations, the  $\pi^0$  method cannot be used in the region  $|\eta| > 2.5$ .
- E/p method - compares the energy measured in the calorimeter (E) with the momentum measured in the tracker (p) for electrons in  $Z \rightarrow e^+e^-$  and  $W \rightarrow e\nu$  events. This method provides the highest precision in the barrel but, because of the statistical uncertainties, the precision is worse in the endcaps.
- $Z \rightarrow e^+e^-$  method - it uses the electrons from  $Z \rightarrow e^+e^-$  decays. In the endcaps, this method is the most precise one and is also used in the region not covered by the tracker  $|\eta| \in [2.5, 3]$ . The  $Z \rightarrow e^+e^-$  data set is small compared to other processes, which is why the statistical uncertainty is higher than in the E/p method. This method will be described in detail in the next section.

### 3.5 The $Z \rightarrow e^+e^-$ calibration method

The properties of the  $Z$  boson were measured by the LEP experiments with very good precision [82]. The measurements of its mass had a relative uncertainty of  $2 \cdot 10^{-5}$ :

$$\begin{aligned} M_Z &= 91.1876 \pm 0.0021 \text{ GeV} \\ \Gamma_Z &= 2.4952 \pm 0.0023 \text{ GeV}. \end{aligned} \tag{3.5}$$

Because of the precise knowledge of its properties, the  $Z$  boson is a good candidate for measuring the performance of the detector. The  $Z \rightarrow e^+e^-$  method for calibration of the ECAL was developed for the Run 2 data-taking period by the CEA-Saclay group.

#### 3.5.1 IJazZ tool

IJazZ is a tool that uses the  $Z$  resonance for measuring the energy scale and resolution of the ECAL by probing the phase space with a very fine granularity. In order to use the available data in the best possible way, this tool is based on the maximization of an unbinned likelihood [83]. The likelihood compares the lineshape of the reconstructed invariant mass with the expected one. The resolution and the energy scale, which are vectors of free parameters in different regions of the detector, are determined by the fit. The likelihood of the invariant mass is obtained from a simulation embedding a realistic description of the background contributions and of detector effects. Because of the complexity to model the energy resolution and a large number of free parameters, a simplified description of the invariant mass lineshape is used with the assumption that:

- the underlying invariant mass distribution from  $Z$  boson is well modeled by a non-relativistic Breit-Wigner distribution:

$$\rho_Z(M) = \frac{1}{\pi} \frac{\Gamma_Z}{2} \frac{1}{M - M_Z + \frac{\Gamma_Z^2}{4}}; \tag{3.6}$$

- the energy response function of the ECAL is described by a Gaussian distribution and the non Gaussian tails are neglected;
- the background contamination in the di-electron sample is negligible.

The probability distribution of the invariant mass is computed using an approximate form of the Voigtain function [84], i.e. Breit-Wigner convolved with a Gaussian distribution. The invariant dilepton mass is calculated from the kinematics of the two leptons:

$$m_{ee} = \sqrt{2E_1 E_2 (1 - \cos \theta)}, \quad (3.7)$$

where  $E_i$  is the energy of the two leptons and  $\theta$  is their angular separation. The energy scale ( $\mu_i$ ) and resolution ( $\sigma_i$ ) of  $i$ -th electron, which are two parameters of the Gaussian smearing, are defined as the:

$$\mu = \sqrt{\mu_1 \cdot \mu_2} \quad (3.8)$$

$$\sigma = 0.5 \sqrt{\sigma_1^2 + \sigma_2^2}. \quad (3.9)$$

To obtain a better fit and avoid the effect of asymmetric response tails due to the energy loss in the material in front of ECAL or empty spaces between the supermodules, the mass window is truncated. The mass window is obtained by defining the approximate full width at half maximum [85] as:

$$\frac{F_{tot}}{2\sqrt{2\ln 2}} = \sigma_Z \times \left( 1 - c_0 c_1 + \sqrt{\left( \frac{\Gamma_Z/2}{\sigma_Z \sqrt{2\ln 2}} \right)^2 + 2c_1 \frac{\Gamma_Z/2}{\sigma_Z \sqrt{2\ln 2}} + c_0^2 c_1^2} \right), \quad (3.10)$$

where  $\sigma_Z$  is the resolution and the constants are  $c_0 = 2.0056$  and  $c_1 = 1.0593$ . The fitting window is then defined as:

$$\mu_Z - 0.9 \frac{F_{tot}}{2\sqrt{2\ln 2}} < m_{ee} < \mu_Z + 0.9 \frac{F_{tot}}{2\sqrt{2\ln 2}}, \quad (3.11)$$

where  $\mu_Z$  is the peak position of the invariant mass histogram. An example of the fit is shown in Figure 3.4.

The  $Z \rightarrow e^+ e^-$  method is used for three different purposes:

- absolute calibration and energy equalisation along  $\eta$  ( $\eta$  scale );
- intercalibration measurements along  $\phi$ ;
- resolution estimation and IC combination.

### 3.5.2 $\eta$ scale calibration

In the  $\eta$  scale calibration procedure, the energy response is equalized for each ECAL  $\eta$  ring. In the barrel region of the ECAL, there are 170  $\eta$  rings, while in the endcap region there are 39  $\eta$  rings.

For defining the  $\eta$  scale the electrons that are less affected by bremsstrahlung and less dependent on the upstream material included in the simulation are used. The selection of these electrons is based on the R9 variable which is defined as:

$$R_9 = \frac{E_{3 \times 3}}{E_{SC}}, \quad (3.12)$$

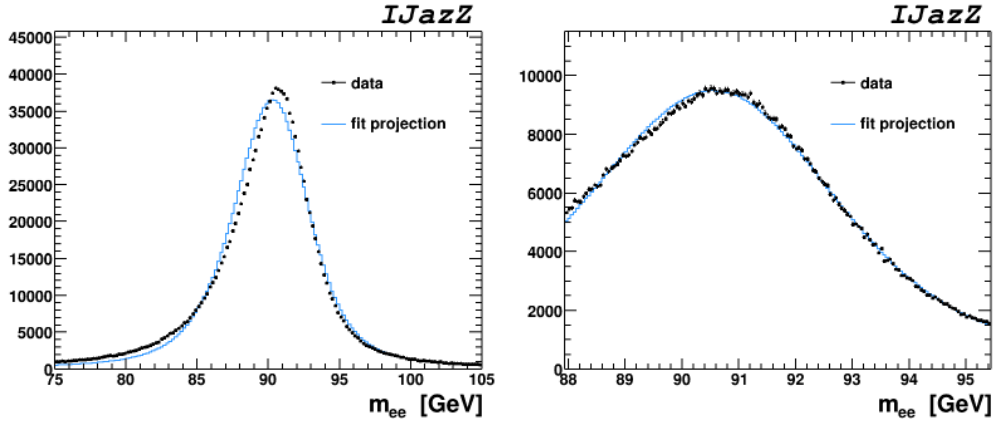


Figure 3.4: The fit performed in the wide (left) and adapted (right) invariant mass window.

where  $E_{3\times3}$  is the energy deposited in a  $3\times3$  crystal matrix around the seed, and  $E_{SC}$  is the supercluster energy. Only the electrons with a high value of R9 (typically  $> 0.94$ ) are used for computing the  $\eta$  scale.

The fit is performed with one free parameter per  $\eta$ -ring (the ring energy scale) and 20 free parameters for the energy resolution (2 bins in R9 and 10 bins in  $|\eta|$ ). In the fit, the energy of the electron is rescaled according to the ring energy scale of the seed crystal in the supercluster. Since the tracker covers the endcaps only up to  $|\eta| \leq 2.5$  ( $|i\eta| \leq 117$ ), the  $\eta$ -scale calibration uses the reconstructed electrons in that region; in the very forward region the calibration is done with pairs of one electron and one supercluster, where the electron is in the tracker coverage.

The fit consists of three steps: the calibration of the pairs EB-EB, the calibration of EB-EE and EE-EE with EB parameters fixed, and SC with  $|\eta| > 2.5$  with electron (SC pairs) with the ring energy scale for  $|\eta| < 2.5$  fixed. The fitting procedure is done for both data and simulation and the  $\eta$  scale is defined as the ratio of the ring energy scale measured in data and in simulation. The resulting  $\eta$  scale is used as a scale factor applied to the calibration constants for data. In Figure 3.5 the ratio of scale parameters in 2018 data and simulation is shown. The energy scales are computed after all the corrections for time-dependent effects are applied. Results are shown for both low bremsstrahlung (Golden) and high bremsstrahlung (Brems) electrons.

### 3.5.3 Intercalibration along $\phi$

The  $Z \rightarrow e^+e^-$  method is one of the methods used for equalizing the crystal energy response along  $\phi$ . Before the derivation of intercalibration constants, the

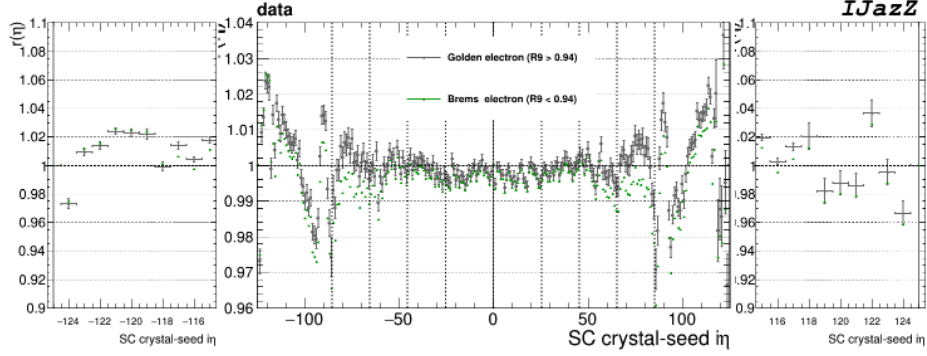


Figure 3.5: The ratio of scale parameters between data and simulation, in gray for  $R_9 > 0.94$  electrons and in green for  $R_9 < 0.94$  electrons.

absolute  $\eta$ -scale is applied such that all  $\eta$  rings have the same average energy response. The fit is performed with a free energy scale parameter for each crystal and 20 resolution parameters as described previously. To reduce the number of free parameters within a single fit, the fit is performed over bands of 10  $\eta$ -rings overlapping one with another. To ensure that the ICs are measured with each SC fully contained in the window, only the scale parameters of the 5 rings in the center of the window are used as calibration constants. This procedure is repeated in different  $\eta$  regions to scan the whole EB. The intercalibration map obtained with this method for 2018 is shown in Figure 3.6. The white squares in the map correspond to non-responding ("dead") crystals and to dead trigger towers.

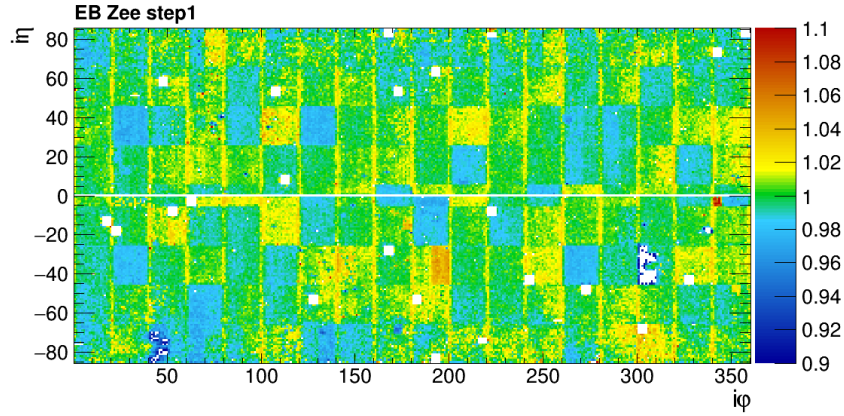


Figure 3.6: Map of the intercalibration constants for ECAL barrel region.

This method for delivering IC constants is sensitive to gaps between ECAL su-

permodules (every 20 crystals around  $\phi$ ) where the energy loss is not fully recovered by regression [75]. Therefore it is needed to perform additional corrections to account for this effect. This correction is derived by performing a fit on the whole EB. Since the effect should be the same for all modules, data in all supermodules are folded together; the EB positive and negative regions are folded together as well. The results of this fit are shown in Figure 3.7. After using the results of

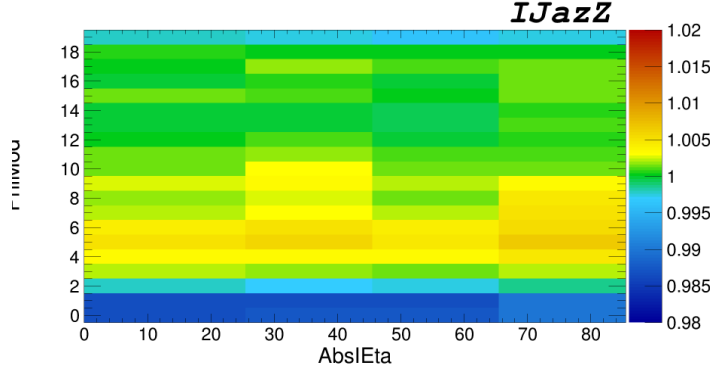


Figure 3.7: Corrections applied to account for the effect of gaps between the supermodules.

the fit as a multiplicative factor to the ICs, the intercalibration map is corrected and shown in Figure 3.8. For the calibration of endcaps, the fit is performed over the full  $EE^+$  and  $EE^-$ , and no additional corrections are needed. The calibration with this method is performed using 90% of the available  $Z \rightarrow e^+e^-$  events, the remaining 10% are used for validation.

### 3.5.4 Resolution and combination

The final IC constants are obtained as a combination of those derived from the methods described in Section 3.4. Each of these methods uses different data sets and has different statistical and systematical uncertainty. To obtain the final intercalibration constants, a combination of the overall precision of each method with systematic uncertainties included is performed. The uncertainties for the  $\pi^0$  and E/p methods are obtained by estimation of the impact of the calibration on the lineshape of the Z invariant mass distribution for the  $Z \rightarrow e^+e^-$  events.

The  $Z \rightarrow e^+e^-$  method is chosen as the reference method because it is the least sensitive for effects such as tracker momentum calibration, pileup, and upstream

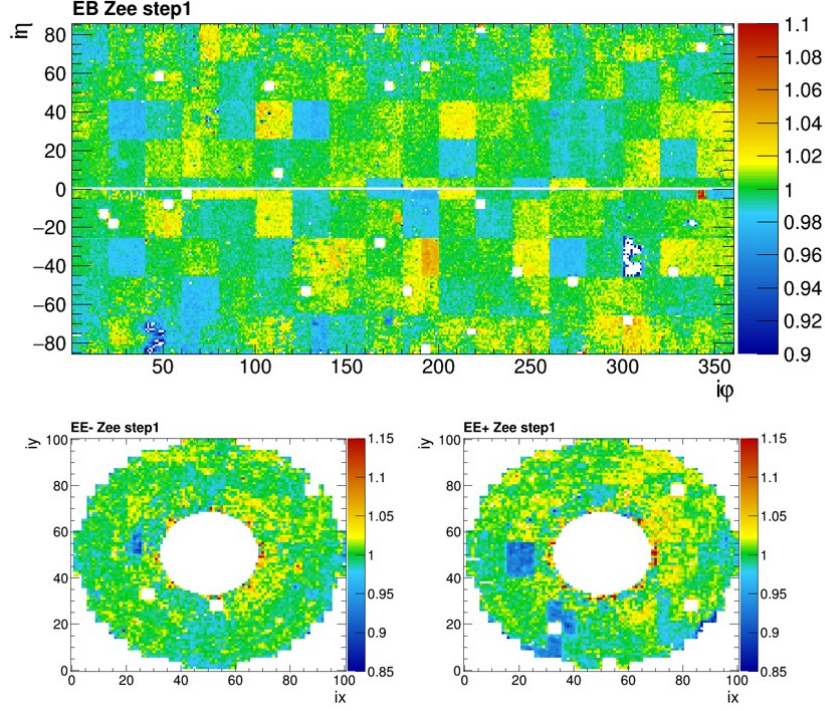


Figure 3.8: The final map of the intercalibration constants for barrel and endcaps for 2018.

material. Therefore, the IC constants for a specific method  $C$  are computed as:

$$\sigma_C = \sqrt{\sigma_{ref}^2 + \rho^{-2} \left[ \left( \frac{\sigma_E}{E} \right)_C^2 - \left( \frac{\sigma_E}{E} \right)_{ref}^2 \right]}, \quad (3.13)$$

where  $\rho$  is a parameter that corresponds to the IC precision on the lepton energy resolution and  $\sigma_E/E$  is the relative energy resolution per electron.

The combination of the constants obtained with the different methods is done by attributing to each of them a weight based on the relative IC precision:

$$w_C = \frac{\frac{1}{(\sigma_C)^2}}{\sum_k \frac{1}{(\sigma_k)^2}}. \quad (3.14)$$

The index  $k$  runs over the calibration methods:  $Z \rightarrow e^+ e^-$ ,  $E/p$ ,  $\pi^0$ .

The precision of the combined ICs, assuming that all measurements are independent, is obtained as:

$$\sigma_{comb} = \sqrt{\frac{1}{\sum_k \frac{1}{(\sigma_k)^2}}}. \quad (3.15)$$

The IC precisions obtained with different methods for all years from the Run 2 are shown in Figure 3.9. In the barrel region the  $\pi^0$  method is slightly worse than  $Z \rightarrow e^+e^-$  and the E/p method. In the endcaps, the most precise method is  $Z \rightarrow e^+e^-$ . In the very forward region ( $|\eta| > 2.5$ ), which is outside of the tracker acceptance, only the  $Z \rightarrow e^+e^-$  method is used.

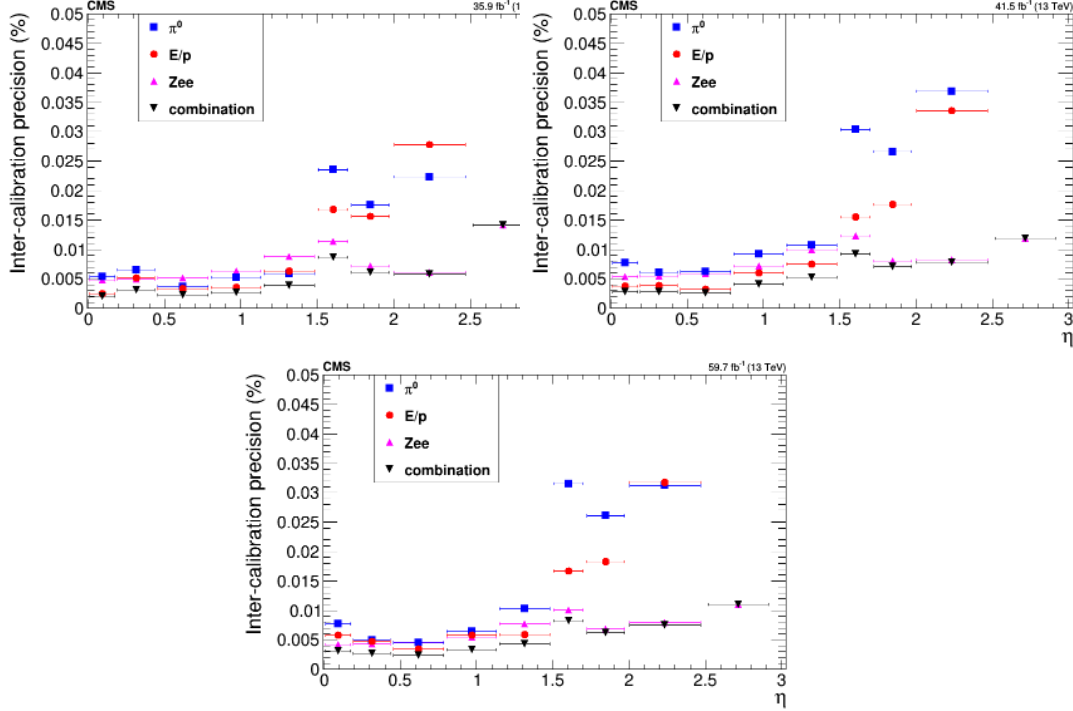


Figure 3.9: The overall precision of the different IC measurement methods as well as their combination for 2016 (top left), 2017 (top right), and 2018 (bottom).

For each year of the Run 2 of the LHC, a refined calibration of the ECAL was performed. The improvement after this calibration can be seen in Figure 3.10, where the electron energy resolution for 2018 is compared using the preliminary calibration performed at the end of each year and to the refined calibration. The resolution is shown as a function of the pseudorapidity, and  $|\eta| = 0, 0.45, 0.8$ , and  $1.15$  corresponds to inter-module boundaries in the barrel.

The resolution through the Run 2 is shown in Figure 3.11 for all electrons and for low bremsstrahlung electrons only. It can be seen that the performance is comparable during the three years. In the central region of the ECAL, the resolution of electrons from Z boson decays is at the level of 1.7%; at large pseudorapidity it is at the level of 3%.

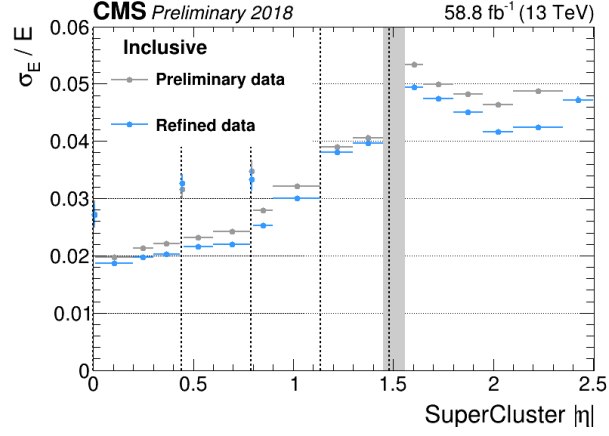


Figure 3.10: Energy resolution as a function of the pseudorapidity for the 2018 data set with the preliminary and refined calibration.

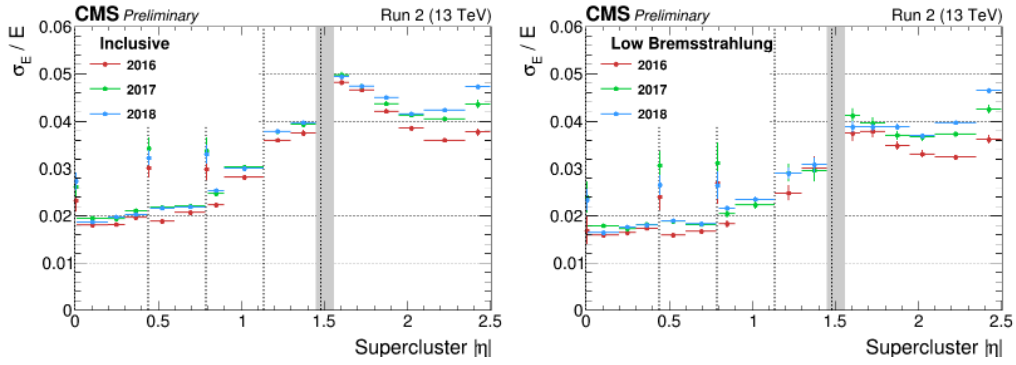


Figure 3.11: Energy resolution with the refined calibration as a function of the pseudorapidity with 2016, 2017, and 2018 data periods for all electrons (left) and for low bremsstrahlung electrons (right).

In order to see how the performance of the ECAL changed through time, going from 8 TeV in Run 1 to 13 TeV in Run 2, the resolution is compared for 2016, 2017 and 2018 (Run 2) with 2012 (Run 1) data-taking period. In Figure 3.12 (left) the energy resolution in the barrel region of the ECAL is shown and it can be seen that the resolution during Run 2 differs from the one obtained during Run 1. One of the effects that contribute to this resolution degradation is the higher number of interactions per bunch crossing during Run 2 (Figure 2.5). To exclude the pileup effect, the resolution is also derived for events in a narrow range of number of reconstructed vertices (between 25 and 35) for all years. In addition, these events are reweighted to match the pileup distribution from the 2012 data

set. It can be seen in Figure 3.12 (right) that good resolution is maintained and, despite the much larger instantaneous luminosity and ageing of the detector, the performance during Run 2 is very close to the one from Run 1.

Even though the performance is almost entirely recovered after excluding pileup effects, there are still some differences between the Run 1 and Run 2 resolution. Studies on the further breakdown of contributions to the energy resolution will be discussed in the next section.

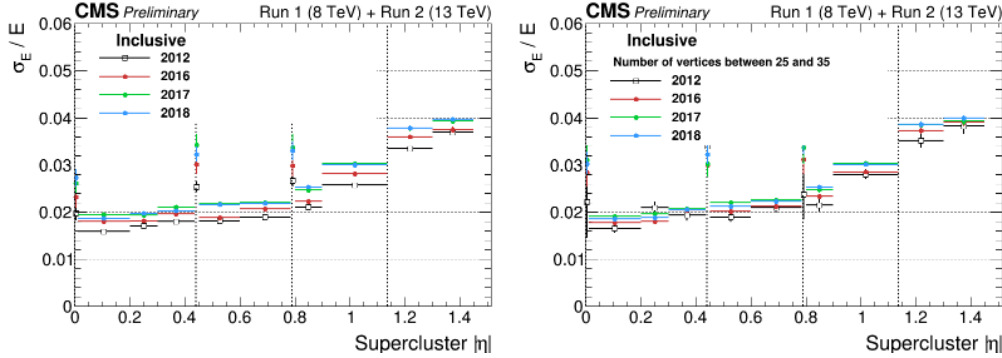


Figure 3.12: Energy resolution with the refined calibration as a function of the pseudorapidity comparing Run 1 (2012) data-taking period with Run 2 (2016, 2017, and 2018) data-taking periods. The resolution is derived from all events (left) and from the events with the number of reconstructed vertices between 25 and 35 (right).

### 3.5.5 Simulation studies

The energy resolution has contributions from many different effects such as pileup, noise, and accuracy of the calibration. In order to study these effects, dedicated simulated samples of  $Z \rightarrow e^+e^-$  with specific effects are produced:

- simulated sample with calibration;
- simulated sample with calibration, energy threshold in PF cluster reconstruction, and realistic noise;
- simulated sample with calibration, energy threshold in PF cluster reconstruction, realistic noise, and pileup.

These studies are done assuming 2018 data-taking conditions.

Using these samples it is possible to estimate the individual contributions of the different effects to the resolution, as shown in Figure 3.13. The cumulative

ECAL energy resolution, obtained by adding up the different contributions to the resolution is also shown. On these plots, the simulation is assumed to be an ideal ECAL simulation without electronic noise, without pileup, and assuming a perfect ECAL calibration (Geant4 simulation).

It can be seen that the effects that are not modeled in the simulation have a significant contribution to the energy resolution. These effects indicate the dependence of the performance on ageing or luminosity. The pileup and noise have almost equal contributions to the resolution, while the impact of intercalibration is negligible.

This study indicates that, in order to maintain a performance during Run 3, the noise has to be mitigated.

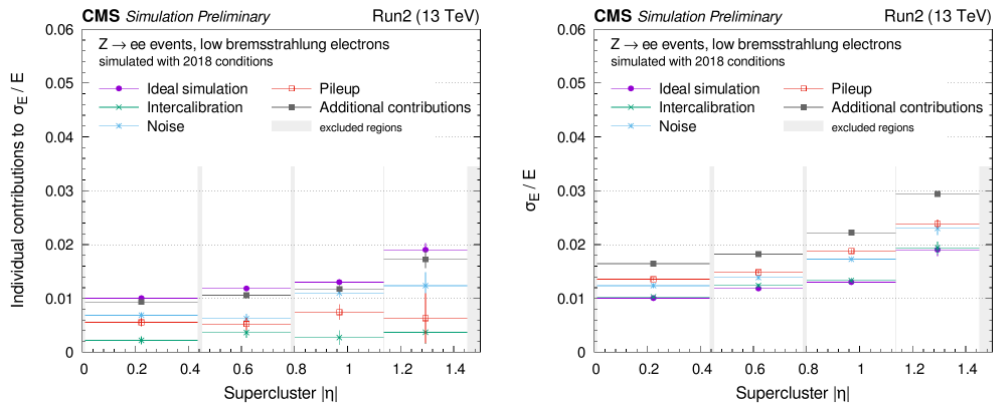


Figure 3.13: Contribution to the ECAL energy resolution from different effects on the (left) and cumulative energy resolution obtained by adding up the contribution from different effects.

### 3.6 ECAL upgrade for High Luminosity LHC

In order to maintain performance from Phase 1 and to cope with the harsh environment of the HL-LHC, the ECAL needs to be upgraded [86].

The plan for the upgrade is a full replacement of the endcaps and upgrade of the barrel readout electronics.

The current readout system of the ECAL is shown in Figure 3.14. On each crystal of the supermodule, there are two APDs that are connected to a Very Front End (VFE) card through a Kapton cable. VFE card consists of five readout channels with analog-to-digital converters (ADC) and multi-gain pre-amplifiers (MPGA). Three outputs with gains  $\times 1$ ,  $\times 6$ , and  $\times 12$  are provided for each channel by the MPGA, and the conversion of outputs is done using the 12-bit, 40 MS/s ADC

chip. After the VFE, the signal is sent to the Front End (FE) card that contains an optical transceiver and the trigger generator circuit based on an ASIC. Thus, the output is sent to the DAQ and trigger system using the optical transceiver FENIX.

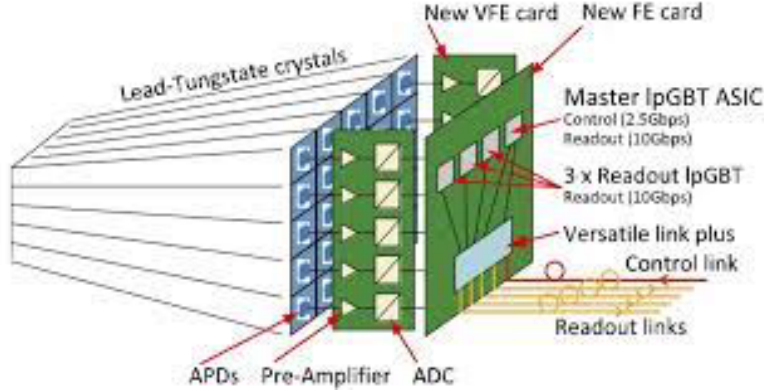


Figure 3.14: The ECAL readout system.

In order to deal with the trigger latency of  $12.5 \mu s$  (instead of the current  $4.5 \mu s$ ) and trigger rate of 750 kHz (instead of the current 100kHz), the ECAL barrel electronics need to be modified and designed to maintain good performance. The schematic view of the updated electronic for ECAL is shown in Figure 3.15. In the upgraded electronics, the MPGAs from the VFE card will be replaced with the Trans Impedance Amplifier (TIA) called CATIA [87]. This will improve the separation of the electromagnetic signals and the signals coming from the ionization in APDs. Instead of the multi-channel APDs, new readout electronics will have LiTE-DTU ASIC (Lisbon-Torino ECAL Data Transmission Unit) [88]. This ASIC samples the signal at 160MS/s with 12-bit resolution. The upgrade will also include moving the trigger primitive generation from the on-detector to the off-detector system. The upgraded FE card will use Low Power Gigabit Transceivers (lpGBT) optical transceivers [89] and Versatile Link plus [90] for data transmission. The off-detector system will be based on the Barrel Calorimeter Processor (BCP) card [91] and it will use FPGAs for the read-out of the detector and to generate trigger primitive.

The prototypes for the upgraded ECAL readout have been produced and tested while further tests are underway. The results of the tests of the CATIA prototype, performed in test beam campaigns at the H4/H2 beamline of the CERN SPS [87], are shown in Figure 3.16. The measured resolution in the test beam matches the one obtained in the beam test for legacy electronics [92]. The timing resolution is measured to be better than 30 ps for the electrons with an energy above 50 GeV,

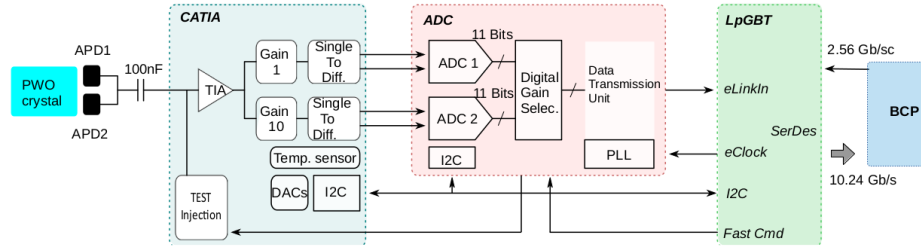


Figure 3.15: Schematic view of the new ECAL barrel electronics.

which complies with the specification.

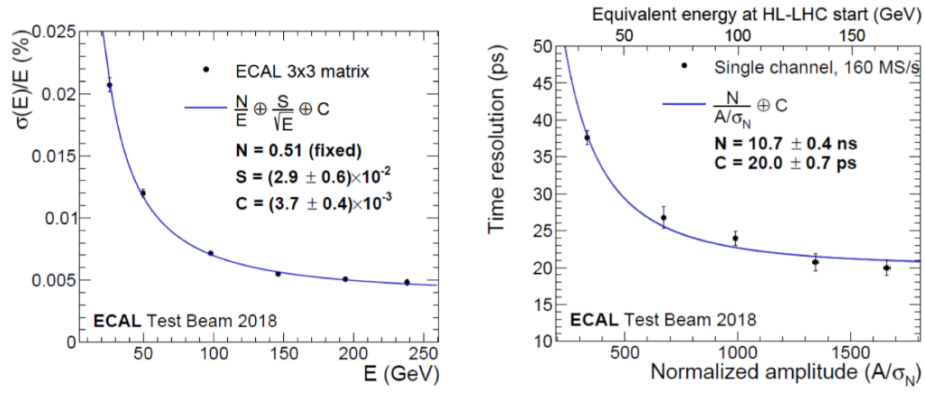


Figure 3.16: Energy resolution (left) and time resolution (right) obtained in the test beam campaign with the CATIA ASIC connected to a commercial ADC[87].

## Chapter 4

# Measurements of track-based event shape observables

The production of the Z boson in association with jets is essential for the modeling and understanding of QCD interactions. The comparison between the measurements and simulations improves the understanding of the prediction accuracy. It is also useful for guiding the improvement of calculations and MC generator techniques. The production of a Z boson in association with jets is also a major background to many processes within the Standard Model and beyond the SM. Therefore, measuring this process with the highest possible precision has wide application.

Several measurements and searches at the LHC require vetoing background events based on their hadronic activity. In particular, in the case of a signal produced through vector boson fusion, the low hadronic activity in the central region is typically exploited to select signal events. The veto implies corrections from higher-order QCD contributions, whose uncertainties are estimated by scale variations.

Hadronic activity vetoes are typically based on jets, yielding complicated phase space restrictions, and reliance on leading-log (LL) parton-shower simulations. In order to reduce uncertainties on cross sections measurements, a new method based on using event shape variables such as N-jettiness has been proposed [7]. The factorization formula allows the summation of the logarithms to NNLL order to be performed.

In this chapter, the measurement of the differential cross section for the production of a Z boson as a function of event shape variables and the sum of the transverse momentum of charged particles is presented. The study is performed in the Z boson decay channel into two muons. The event shape variables are computed using charged particles only in order to constrain the contributions coming from pileup particles. The measurements are performed in four transverse Z mo-

momentum ( $p_T^Z$ ) regions and in three dilepton mass ( $M_{\mu\mu}$ ) regions.

In the first part of this chapter (Section 4.1), the measured variables are defined. In Section 4.2, data samples and simulated samples are presented. The selection criteria for the events and for the particles used for the N-jettiness calculation are described in Section 4.3. The corrections applied to simulations are summarized in Section 4.4. The techniques for the deconvolution of detector effects are presented in Section 4.5. The uncertainty sources that affect the measurements are described in Section 4.6. Finally, in Section 4.7 the measurements for the differential cross section are presented.

## 4.1 Observable definition

The inclusive event shape variable N-jettiness defined in [7] gives a global view of the event. It tests the compatibility of events with a topology with N jets and it can be used to discriminate signal from background with higher jet multiplicity.

N-jettiness ( $\tau_N$ ) is defined as:

$$\tau_N = \frac{2}{Q^2} \sum_k \min \{q_a \cdot p_k, q_b \cdot p_k, q_1 \cdot p_k, \dots, q_N \cdot p_k\}. \quad (4.1)$$

where  $p_k$  is the four-momentum of particle  $k$ ,  $q_a$  and  $q_b$  are the four-momenta of the beams,  $q_1, \dots, q_N$  represents the four-momenta of N jets in the event, and the scale  $Q^2$  is the typical scale of the hard scattering process. The sum runs over all final state particles except the signal leptons or photons.

The four product  $q_a \cdot p_k$  or  $q_j \cdot p_k$  represents the distance between the final state particle with the momentum  $p_k$  and the beam or jet  $j$ . The distance measure can be adapted for a particular case, while the general properties of the variable remain the same. The choice of the distance measure defines the shape of the area that is assigned to the jet. The closest distance between  $p_k$  and the beam or jet axis is the minimum of the Eq. 4.1. This estimated sum of minima has a small contribution from soft particles and from energetic particles that are close to jets or beams, while energetic particles that are far from jet and beam axes give a large contribution. For events with at least N jets, in the limit  $\tau_N \rightarrow 0$ , the event contains N narrow jets. In the case where the N-jettiness is greater than 0, the jets are wider and there is radiation between beams and jets or the number of jets is higher than N.

In this analysis, the zero-jettiness  $\tau_0$  (also called "beam-thrust") and one-jettiness  $\tau_1$  are computed using a geometrical measure in the reference frame where the Z boson rapidity is zero:

$$\tau_0 = \sum_k p_{Tk} e^{-|y_k - Y|} \quad (4.2)$$

$$\tau_1 = \sum_k \min \{p_{Tk}e^{Y-\eta_k}, p_{Tk}e^{-Y+\eta_k}, p_{Tk}(2 \cosh \Delta\eta_{J1,k} - 2 \cos \Delta\phi_{J1,k})\}, \quad (4.3)$$

where  $Y$  is the rapidity of the Z boson, and  $\Delta\eta_{J1,k}$  and  $\Delta\phi_{J1,k}$  are the pseudorapidity and azimuthal angles between the particle  $k$  and the jet  $J$ . In addition to these variables, the sum of the transverse momenta  $\sum p_k$  of final state particles excluding the signal components is measured.

Since the N-jettiness assigns particles to one of the jets or beams axes by finding the minimum distance, it can be used as an exclusive jet algorithm. Therefore, in this analysis, the axes of the jets are computed by minimizing the one-jettiness. The minimization is performed using the XCone algorithm [93]. The algorithm starts with the seed axes that are obtained with the anti- $k_T$  clustering algorithm or by looping over all the possible axes combinations to find the minimum (if anti- $k_T$  jets are not defined). These axes are iteratively moved to find a minimum of the one-jettiness. Through this step, the particles are assigned to one of the jets or beams regions. Using the information from the jet constituents, the jet axes are recomputed and updated. The assigning and updating axes steps are repeated until the axes change their orientation by more than  $10^{-4}$ .

## 4.2 Data and simulation samples

The analysis is based on data collected by the CMS experiment from 28th of April to 3rd of December 2018 in proton-proton collisions with the center-of-mass energy of 13 TeV. This data set corresponds to an integrated luminosity of  $59.4 \text{ fb}^{-1}$  [94].

Collected data are divided into different primary data sets according to the trigger selection: in this analysis, the DoubleMuon data set is used. During 2018 there were four run periods named A, B, C, and D with different beam and detector conditions (Table 4.1). The samples with the latest data processing and the simulation with the most updated conditions (such as calibrations, energy scale corrections, etc) for all the subsystems and object reconstruction algorithms are used.

| Data sample      | L[fb <sup>-1</sup> ] |
|------------------|----------------------|
| DoubleMuon Run A | 13.704               |
| DoubleMuon Run B | 7.061                |
| DoubleMuon Run C | 6.895                |
| DoubleMuon Run D | 31.742               |

Table 4.1: List of data samples used in the analysis.

The samples for the simulation of the signal and background processes are listed

in Table 4.2. The signal samples are generated with the MADGRAPH5\_aMC@NLO generator [33] interfaced with PYTHIA8 [], as explained in Section 1.4. In order to have a larger number of simulated signal events, several samples are merged: three samples with a different number of outgoing partons (npNLO) and a sample with an invariant mass of leptons greater than 50 GeV. The showering and hadronization are performed with PYTHIA8 using the CP5 Tune [38]. The matrix element (ME) is computed to NLO for up to two partons. The PDF set used is NNPDF 3.1 [28] and the strong coupling  $\alpha_s$  is set to 0.118. The cross section for these samples is computed by the generator.

The dominant background is the production of top quark pairs that decay into leptons. This process, as well as the single top production in the t channel and single top production in association with a W boson, are simulated at NLO in  $\alpha_s$  using the POWHEGBOX [95]. The single top quark production in the s channel is simulated using MADGRAPH5\_aMC@NLO with the  $\alpha_s$  at NLO. The production of the Z boson in association with an additional electroweak boson Z or W is simulated at LO with PYTHIA8. The showering and hadronization for signal and background simulations are performed with PYTHIA8 CP5 Tune.

In order to compare the signal and background simulations with the experimental data, the simulated events are normalized according to the observed luminosity of the data samples. The weight factor used for normalization is computed as:

$$\text{weight} = \frac{\sigma L}{N_{\text{processed}}} \quad (4.4)$$

where  $\sigma$  is the cross section of each process as listed in Table 4.2,  $L$  is the integrated luminosity of the data sample and  $N_{\text{processed}}$  is the number of generated events. The cross sections are calculated at NNLO + NNLL using TOP++ version 2.0 [96] for the tt process and using HATHOR version 2.1 at NLO in  $\alpha_s$  for single top production. The cross section for ZZ and ZW processes is calculated using MCFM 6.6 and the tW cross section is taken from the [97].

### 4.3 Selection

In this analysis, events with two opposite charged muons are studied. In order to reject the events coming from the background processes listed in Table 4.2, a set of selection criteria is applied. In addition, for the computation of the N-jettiness variables, a selection is applied to particles to exclude those coming from pileup events.

| Process                              | $\sigma$ [pb] |
|--------------------------------------|---------------|
| Signal                               |               |
| Z+jets ( $m_{ll} > 50$ GeV)          | 5931.9        |
| Z+jets (npNLO = 0)                   | 4620.52       |
| Z+jets (npNLO = 1)                   | 859.59        |
| Z+jets (npNLO = 2)                   | 338.26        |
| Background                           |               |
| $t\bar{t}$                           | 831.7         |
| $tW$                                 | 35.6          |
| $\bar{t}W$                           | 35.6          |
| WJetsToLNu                           | 61526.7       |
| $t \rightarrow lX$ (s channel)       | 10.32         |
| $t \rightarrow lX$ (t channel)       | 136.02        |
| $\bar{t} \rightarrow lX$ (t channel) | 80.5          |
| $WW \rightarrow 2L2Nu$               | 12.21         |
| WZ                                   | 23.5          |
| ZZ                                   | 15.4          |

Table 4.2: List of simulated samples and their cross section.

### 4.3.1 Event selection

Events are selected using an unprescaled trigger with the requirements for the transverse momenta of the leading<sup>1</sup> muons to be greater than 17 GeV and 8 GeV. There is a loose track isolation requirement and the invariant mass of the leptons is required to be above 3.8 GeV. In addition, the longitudinal distance between the two muon tracks and the vertex of the hard interaction has to be lower than 0.2 cm.

The measurements are performed for the muon pairs where the leading muon has  $p_T > 25$  GeV and the subleading has  $p_T > 20$  GeV. The transverse momentum is required to be above the threshold of the trigger, in order to reduce a possible bias coming from the trigger efficiency. Both muons are selected within the pseudorapidity  $|\eta| < 2.4$ .

The muons are identified using the PF algorithm described in Section 2.3. The selection criteria for the identification used in this analysis correspond to the so-called *Medium ID* muons [65]. A *Medium ID* muon is either a tracker or a global muon that leaves a signal in 80% of the inner tracker layers it transverses. For the *Medium ID* muons, the muon segment compatibility, that evaluates the number

---

<sup>1</sup>The two leading muons in an event are the two muon candidates with the highest  $p_T$

of matched segments in all stations and the closeness of the matching in position and direction, has to be greater than 0.451 for the tracker muons and 0.303 for the global muons. In addition, if the muon is global, the global fit needs to have a  $\chi^2$  per degree of freedom less than 3, and a  $\chi^2$  for the position match between the tracker and the standalone muon less than 12. The *Medium ID* muon also includes a selection based on the kink-finding algorithm. This algorithm splits the track at several places and for each split compares the two tracks. If the  $\chi^2$  is large, the two tracks are not compatible with a single track. The  $\chi^2$  is required to be lower than 20. The overall reconstruction efficiency for the muons from W and Z events is 99.5%.

The details of the Medium ID identification requirement are summarized in Table 4.3.

| Variable                                          | Selection |
|---------------------------------------------------|-----------|
| Global Muon                                       | Yes       |
| Particle-Flow muon                                | Yes       |
| $\chi^2/\text{ndof}$ of the global muon track fit | $< 3$     |
| Tracker-Standalone position match                 | $< 12$    |
| Kink-finder                                       | $< 20$    |
| Segment compatibility                             | $> 0.45$  |

Table 4.3: Medium ID criteria used for the muon selection [65].

In order to distinguish the muons originating from the Z boson from the ones produced in jets, an isolation criterion is applied. The relative muon isolation is determined as:

$$I_{rel}^\mu = \frac{1}{p_T^\mu} \left[ \sum p_T^{h_{PV}^\pm} + \max \left( 0, \sum E_T^{h^0} + \sum E_T^\gamma - \frac{1}{2} \sum p_T^{h_{PU}^\pm} \right) \right]_{\Delta R < 0.4}, \quad (4.5)$$

where the sum runs over the PF candidates in a cone of radius  $R = 0.4$  around the direction of the muon candidate track. The energy contributions considered are the ones originating from the primary vertex ( $h_{PV}^\pm$ ), neutral hadrons ( $h^0$ ), and photons ( $\gamma$ ). Since the neutral particles deposit on average half as much energy as charged particles, the contribution of the neutral hadrons from pileup is estimated as 1/2 of charged particles coming from pileup  $h_{PU}^\pm$ . For this analysis, the relative muon isolation is required to be lower than 0.15, with the achieved efficiency of 95%.

### 4.3.2 Particle selection

The N-jettiness variables as defined in Section 4.1 are computed using all final state particles but the final state leptons. The particles correspond to the PF candidates, that were reconstructed as described in Section 2.3.

In Figure 4.1 the zero-jettiness distribution built from generator quantities and from reconstruction quantities is shown. At the reconstruction level, there are many particles coming from pileup that are contributing to the sum and therefore shifting the peak of the distribution toward higher values.

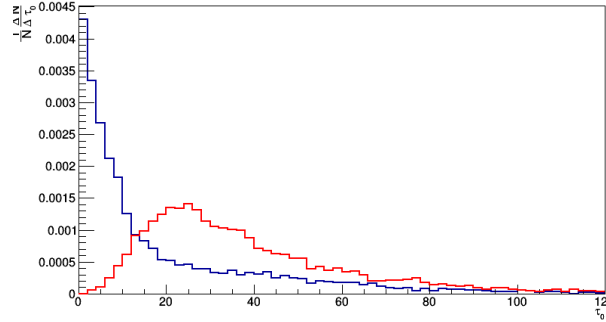


Figure 4.1: The distribution of zero-jettiness at the generator level (blue) and reconstructed level (red).

Many particles measured in the final state originate from pileup and specific selections need to be applied to exclude them from the calculation of the  $\tau_N$ .

Since the neutral particles are detected only in the calorimeter it is hard to distinguish those coming from the primary vertex from those coming from the pileup. Therefore, in this analysis, only the charged particles are considered in calculations of the  $\tau_N$ .

In the following, the charged particles corresponds to the PF tracks that are seeded with two hits in consecutive layers in the pixel detector. They are required to have at least eight hits in total and at most one missing hit along way, and to originate within a cylinder of a few mm radius centered around the beam axis.

To estimate the effect of choosing only charged particles, a comparison of the distributions of the  $\tau_0$  variable are computed with only charged particles, only neutral particles, and of  $\tau_0$  computed with all the particles is performed at generator level (Figure 4.2). The number of particles in all three distributions is the same and particles are selected randomly from the event.

Charged particles are required to have a  $p_T > 1$  GeV and  $|\eta| < 2.4$ .

To suppress charged particles from pileup, a further selection is made according to their association with the primary vertex. The flags that show how tight the

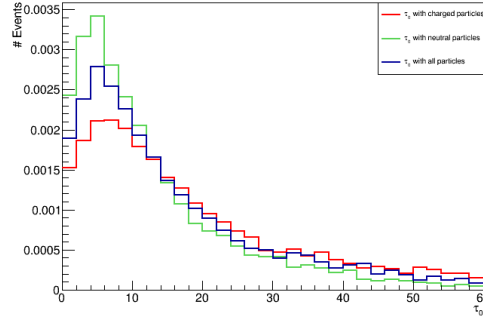


Figure 4.2: The distribution of zero-jettiness computed with charged, neutral, and all particles from the event

association with the primary vertex (PV) is, are summarized in Table 4.4.

| PV association | meaning                                                                                       |
|----------------|-----------------------------------------------------------------------------------------------|
| PVUsedInFit    | a the track that is used in the PV fit                                                        |
| PVTight        | a the track that is not used in the fit of any of the other PVs and is closest in z to the PV |
| PVLoose        | a the track that is closest in z to a PV other than the PV                                    |
| NoPV           | a the track that is used in the fit of another PV                                             |

Table 4.4: PV association flags description.

In order to estimate the contribution of charged particles coming from pileup and the ones coming from the main PV, for different PV association selections, dedicated studies are performed. Using simulated events from Drell-Yan process, the charged particles are identified as coming from the PV by searching for the generated particles that can be associated with them. If there is a generator level particle that can be matched with the reconstructed charged particle we consider that particle as coming from the main PV. The matching requires a distance between the generator-level particle and the reconstructed track lower than 0.005 and a relative difference between their transverse momenta lower than 3%. To be identified as a particle that comes from pileup, the reconstructed track is required not to be matched with any of the particles at the generator level, which in this case means to have a  $\Delta R$  distance greater than 0.3 or to have a relative  $p_T$  difference greater than 20%. In Figure 4.3 it can be seen that most of the particles that originate from the main PV are used in the primary vertex fit or are closest in z to the main primary vertex. For the N-jettiness computation, only the particles that satisfy one of these two PV association qualities are considered. In addition, the longitudinal distance of the charged particle from the PV is required to be lower than 0.3 cm.

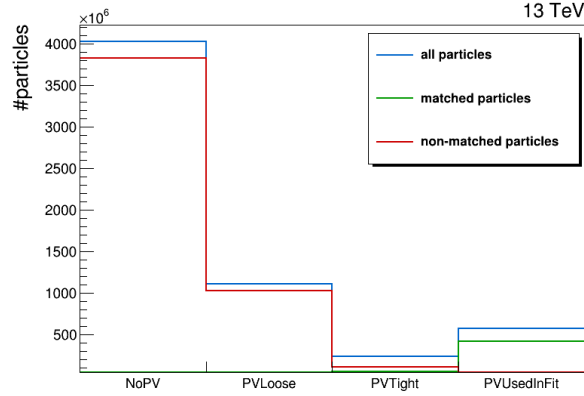


Figure 4.3: Vertex quality association flag for matched and non matched tracks.

With these selections applied, the particles coming from pileup are rejected while the efficiency of reconstructing particles is kept.

## 4.4 Simulation corrections

In order to improve the agreement of the simulation with data, several corrections need to be applied to simulated events. These corrections are applied as an event weight and include several effects such as pileup, misalignment of the reconstruction efficiency, identification and isolation efficiency, and trigger efficiency.

### 4.4.1 Pileup

The average number of pileup events depends on the beam conditions and varies during the data-taking. The pileup profile in data is estimated by using the instantaneous luminosity and a total proton-proton cross section of 69.2 mb [98]. The pileup profile in simulated samples is not the same as the one in the data since it cannot be known in advance and be included in the simulation. Therefore, the simulation needs to be corrected in order to match the pileup distribution from the data. The correction is performed by delivering weights as a function of the number of vertices per event, which are then applied to simulated events.

### 4.4.2 Momentum corrections

The measurements of muon momentum are biased due to the detector misalignment, reconstruction algorithm, or uncertainties in the knowledge of the magnetic

field. The so-called Rochester method [99] is used to correct the biases. It consists of two steps. Firstly, the mean inverse transverse muon momentum  $\langle 1/p_{T,\mu} \rangle$  of muons from the Z boson decays is required to be the same as the one derived from a perfectly aligned sample. The corrections are derived as a function of charge, pseudorapidity, and azimuthal angle and they are derived for both data and simulation. Secondly, the average invariant mass  $\langle M_{\mu\mu} \rangle$  is used to tune the corrections and remove the bias coming from the mismodeling of the detector efficiency in the simulation.

#### 4.4.3 Scale factors for ID and isolation

Differences between data and simulation can be introduced by the selection criteria and imperfect modeling of objects used for these selections. The efficiency of the isolation and identification selections, described in Section 4.3, are computed for data and simulation using the Tag-and-Probe method [100] as a function of the muon transverse momentum and pseudorapidity. The correction factors, called scale factors (SF) are computed as:

$$SF = \frac{\epsilon_{data}(p_T, \eta)}{\epsilon_{sim}(p_T, \eta)}. \quad (4.6)$$

and applied as weight by event.

#### 4.4.4 Trigger Scale factors

To account for the efficiency of the trigger used in the analysis, the corresponding scale factors need to be computed and applied to the simulation. For the double muon trigger used in this analysis, I calculated the scale factors using the "reference trigger" method that is suitable for complex trigger efficiency computation. In this method, the efficiency is computed by estimating its efficiency for the events selected by a reference trigger first. After computing the reference trigger efficiency, the double muon trigger efficiency can be then calculated for all events.

The first step in this method is choosing the reference trigger which should be the one that has high efficiency on the events that pass complex trigger. For the trigger used in this analysis, a good choice for a reference trigger is a single muon trigger with a  $p_T$  threshold of 17 GeV. For the dimuon events, the reference trigger efficiency is computed as:

$$\epsilon_{ref} = 1 - (1 - \epsilon_{ref}^{\mu_1}) \cdot (1 - \epsilon_{ref}^{\mu_2}). \quad (4.7)$$

The efficiency of the complex trigger is computed for events that are selected with the reference trigger ( $\epsilon_{DMu|ref}$ ). Finally, the efficiency of the double muon trigger efficiency is computed regardless of the reference trigger:

$$\epsilon_{DMu} = \epsilon_{DMu|ref} \cdot \epsilon_{ref}. \quad (4.8)$$

In simulation samples, the information of generated and reconstructed number of events is available, and it is exploited to compute the efficiency. In data samples, the efficiencies in data are calculated using the Tag-and-Probe method. The efficiencies are computed for events that pass the selection defined in Section 4.3. In order to avoid the case where both muons are matched to the same trigger object, it is required that the  $\Delta R$  between the two muons is greater than 0.3. The efficiency measurements are performed as a function of the absolute pseudorapidity of the muons.

In the Tag-and-Probe method, the signal and background are modeled using analytical functions: a Gaussian for the signal and an exponential function for the background. A fit is performed to model the invariant mass of muons. The example of a fit for probe muons that pass the selection and for the ones that fail the selection is shown in Figure 4.4.

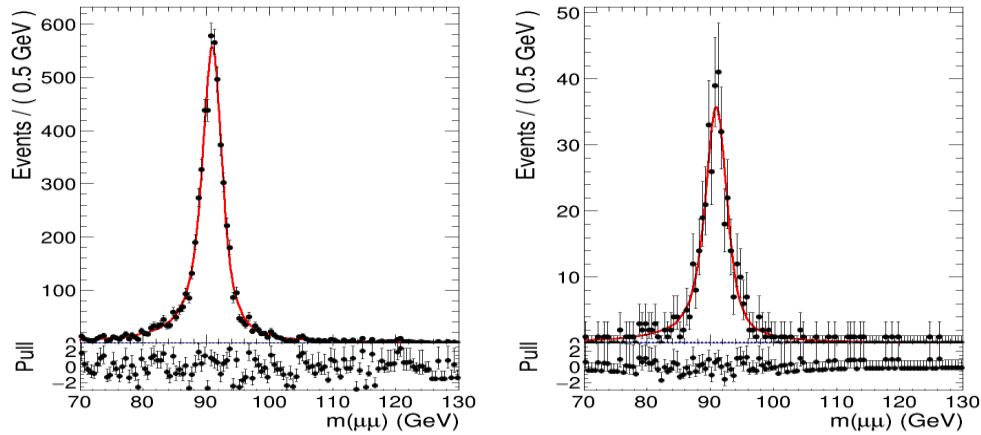


Figure 4.4: Fit results for one of the central pseudorapidity bins for probe muons that pass (left) and fail (right) selections.

The efficiencies computed in data and in simulation are shown in Figure 4.5. The uncertainties shown in these plots are including both statistical and systematic uncertainty. The systematic uncertainties include the uncertainties related to the modification of the histograms that are fitted, such as variation in the binning and range of the invariant mass histograms fitted, and uncertainties related to the fitting procedure, more precisely the fitting function used for the signal fit. These uncertainties are summed in quadrature. The final scale factors are shown in Figure 4.6:

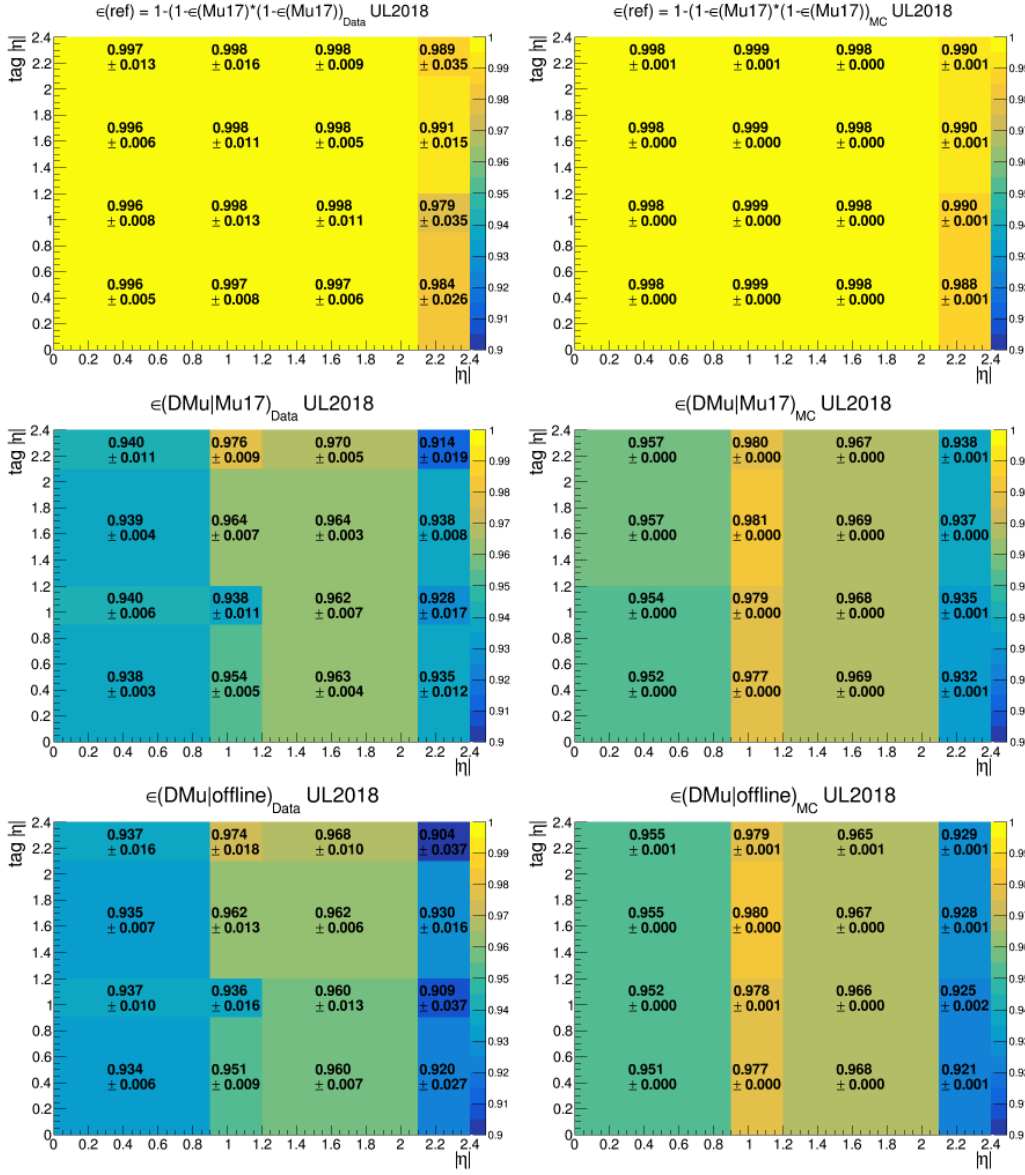


Figure 4.5: Single and double muon trigger efficiencies in data (left) and simulation (right) shown in bins of pseudorapidity.

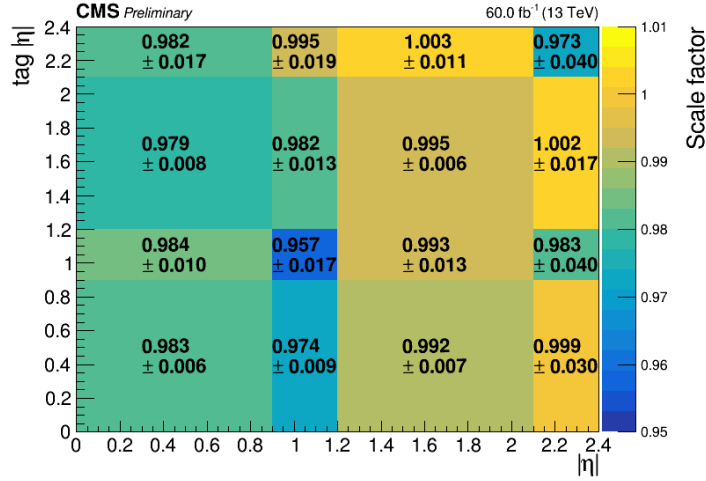


Figure 4.6: The scale factors obtained for 2018 data taking period shown in bins of pseudorapidity.

#### 4.4.5 Comparison of data and simulation

In order to estimate the contribution of the background, the distributions of variables at the reconstruction level in data and simulations (samples from Table 4.1 and Table 4.2) are compared. The comparison of data and simulation also allows the effectiveness of the corrections applied to simulated samples to be checked.

##### Z boson variables

The Z boson candidates are reconstructed using the two opposite-charge muons in the event with the highest transverse momenta. The distributions are shown for events that pass the selections defined in Section 4.3. For the invariant mass distribution shown in Figure 4.7, the invariant mass bounds are removed. It can be seen that after applying the Rochester corrections, a satisfactory description of the invariant mass is obtained. In order to see the contributions of each background considered, the distribution of the invariant mass is also shown in a logarithmic scale. It can be seen that in the whole range, the background contributions are rather small.

The distributions of the Z boson rapidity and transverse momentum are shown in Figure 4.8. These distributions are shown for events with real Z boson with the invariant mass in the range from 76 to 106 GeV. The rapidity distribution shows good agreement between data and simulation, while in the low region of  $p_T^Z$  (less than 30 GeV), differences are observed.

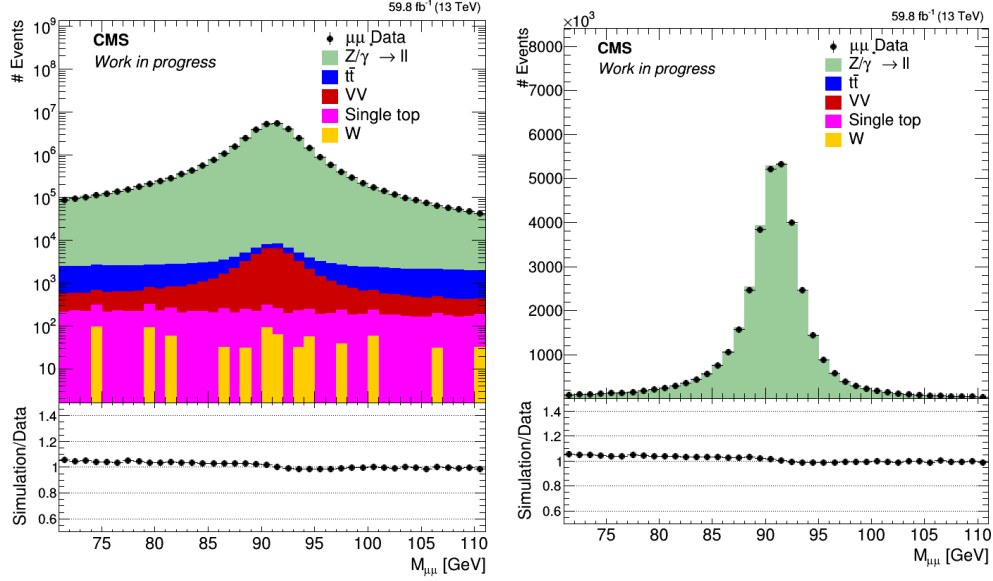


Figure 4.7: Data to simulation comparison of the invariant mass distribution with linear (left) and logarithmic (right) scale.

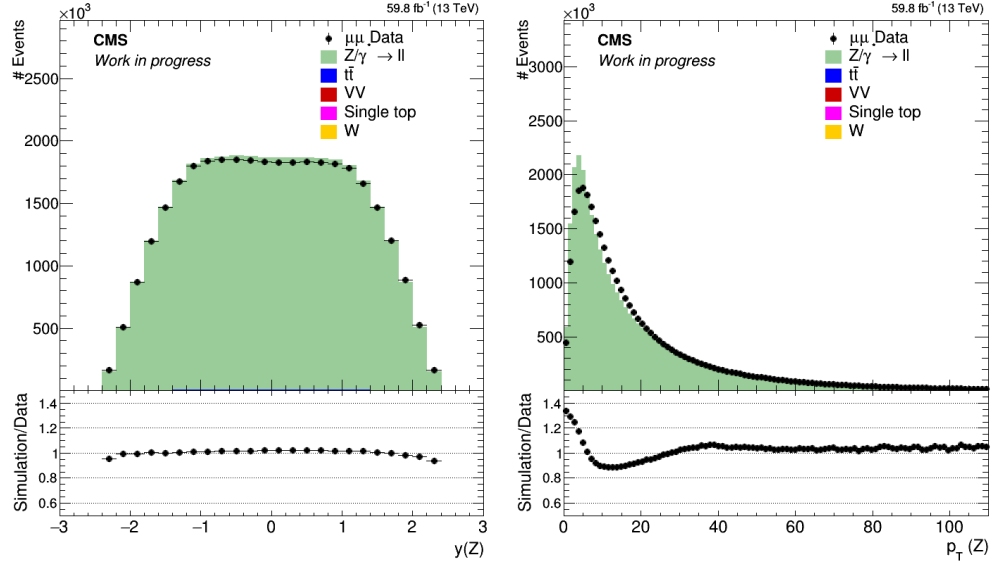


Figure 4.8: Data to simulation comparison of muon pair distributions: rapidity (left) and transverse momentum (right).

### Muon variables

The data and simulation comparison of the muon transverse momenta, pseudorapidity, and azimuthal angle are shown in Figure 4.9. The distributions are shown for the two muons in the selected events; therefore, histograms are filled twice for each event. After applying all necessary correction factors, the distributions show a good agreement between data and simulation.

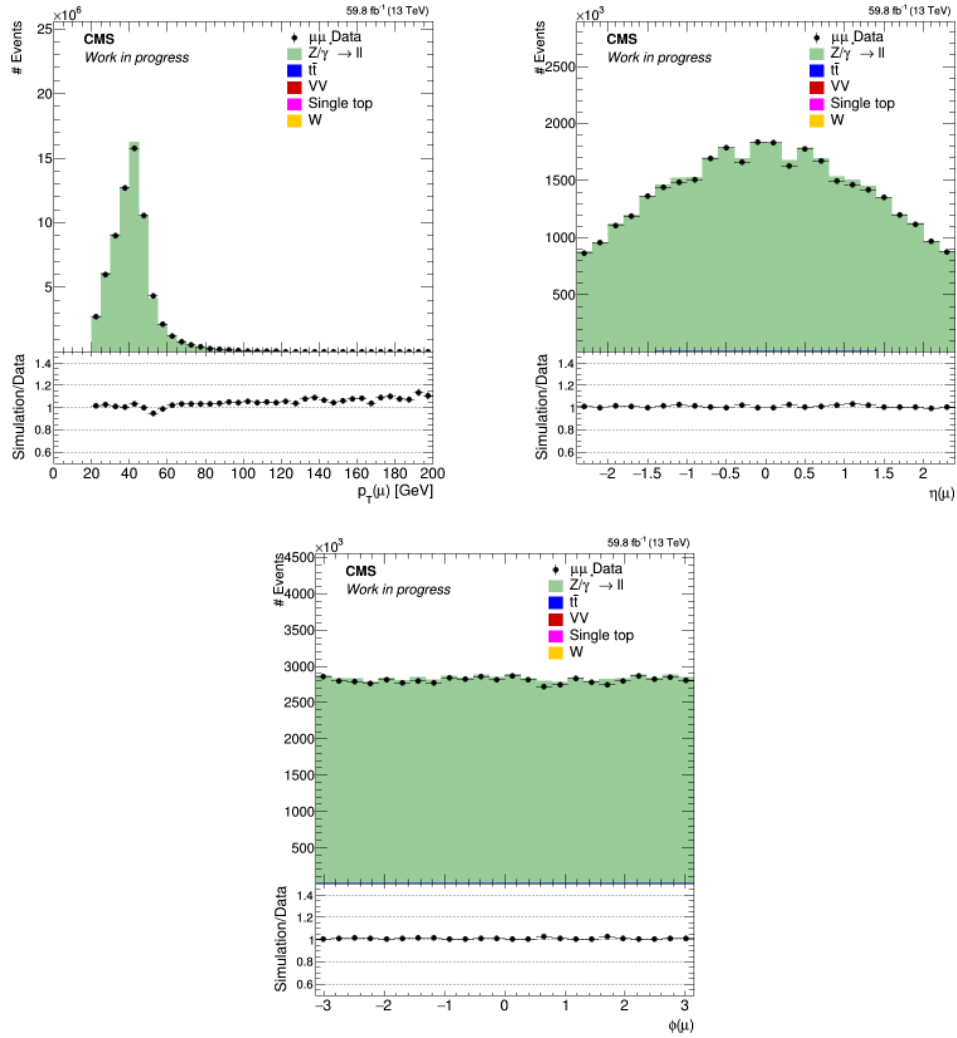


Figure 4.9: Data to simulation comparison of reconstructed muon distributions: transverse momentum, pseudorapidity, and azimuthal angle.

### Charged particle variables

The transverse momentum and pseudorapidity distributions of charged particles that contribute to the calculation of the N-jettiness are shown in Figure 4.10. In the high pseudorapidity region, it can be seen that there is a discrepancy between data and simulation. Since this observable is sensitive to the non-perturbation effects, additional studies are done using simulations obtained with a different PYTHIA8 tune. The transverse momentum and pseudorapidity distributions are compared at the generator level for two different tunes, CP5 and CUETP8M1. Figure 4.11, shows a difference in the high pseudorapidity region for these two tunes and according to this comparison, the CUETP8M1 tune will provide a better agreement with data.

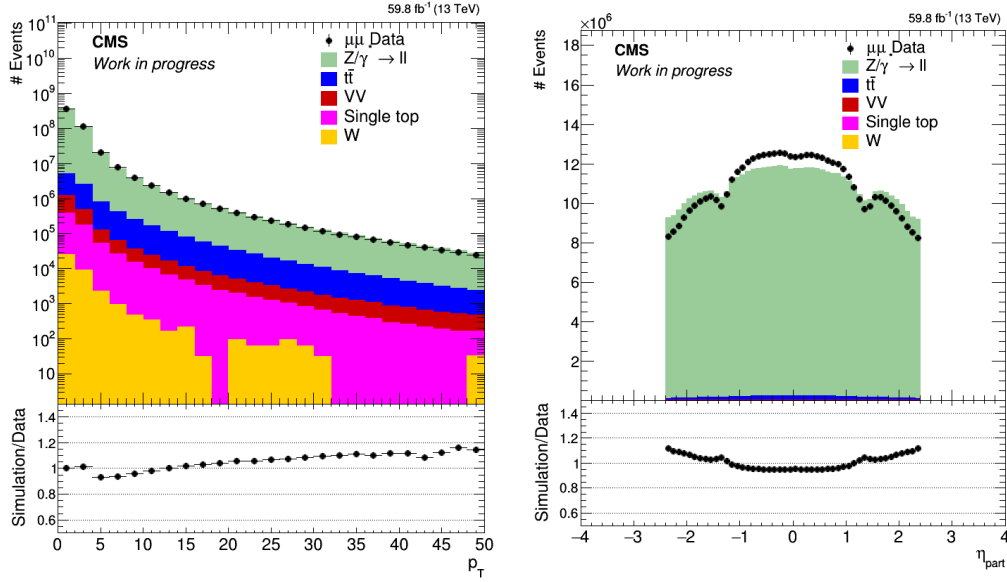


Figure 4.10: Data to simulation comparison of charged particle distributions: transverse momentum and pseudorapidity.

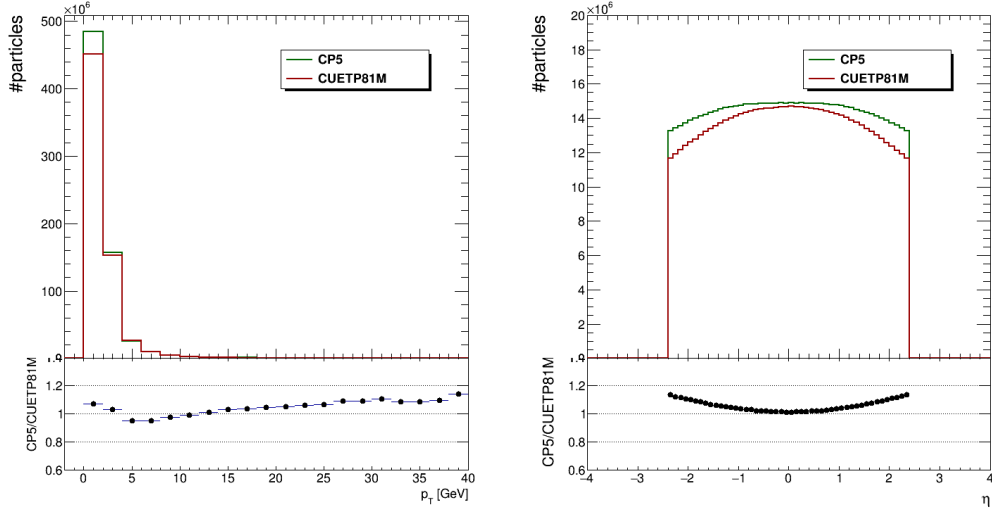


Figure 4.11: Comparison of transverse momentum and pseudorapidity of charged particles distributions for CP5 and CUETP81M tunes at generator level.

## 4.5 Unfolding method

In order to compare the obtained results with theoretical predictions or with results from other experiments, the measurements have to be corrected for the effects coming from the detector such as object reconstruction, efficiency inside the acceptance of the detector and the misidentification of the objects of interest. In addition, since the reconstruction system has a finite resolution, the measured value of the observable usually does not correspond to the true one.

To account for these effects, a procedure called unfolding is performed. It is based on simulation which can provide information about the measured and original values of the specific observables. The true value, which is the result of the simulation procedure described in Section 1.4 will be referred to as generated observable, while the measured value after the detector simulation will be referred to as reconstructed observable. The reconstructed events have to pass the selections described in Section 4.3, implying that also the corresponding generated events are selected in specific phase space.

### Phase space

The phase space at the generator level is chosen to match the selections at the reconstruction level. To take into account the effects of final state QED radiation of muons on measured observables, so-called "dressed" muons are used. At the

generator level, the photons that are inside of the cone  $\Delta R(l, \gamma) < 0.1$  around the muon are merged with that muon. The merging is done such that the four-momenta of photons are added to the closest muon:

$$p_\mu^{\text{dressed}} = p_\mu^{\text{bare}} + \sum_\gamma p_\gamma, \quad (4.9)$$

where the term bare (dressed) corresponds to the lepton before (after) correction. Two dressed leptons with opposite charges and the largest transverse momentum are selected and required to have  $p_T$  larger than 25 GeV (first muon) and 20 GeV (second muon). The leptons are required to be within pseudorapidity region  $|\eta| < 2.4$  and the invariant mass of leptons is required to be within a certain interval.

Charged particles that are used for track-based variables calculation are required to have  $p_T$  larger than 1 GeV and to be within pseudorapidity region  $|\eta| < 2.4$ .

### Unfolding procedure

The relation between the reconstructed and generated observable can be defined as:

$$\bar{y}_i = \sum_{j=1}^m A_{i,j} \bar{x}_j + b_i, \quad 1 \leq i \leq n, \quad (4.10)$$

where the sum runs over the  $m$  bins of the generated distribution while the number of bins at the reconstruction level is  $n$ .  $A_{i,j}$  represents the matrix of probabilities that describes the bin-to-bin migration at the reconstructed level. The average expected value is  $\bar{y}_i$ . The background contribution for the specific bin is  $b_i$ .

The unfolding procedure is done using the TUnfold package [101] where the estimation of the true value is done using the least square method with Tikhonov regularization [102]. By determining the stationary point in the Langrangian, the statistical fluctuation of the expected value is amplified which causes fluctuation in  $x$ . Typically, to address these fluctuations regularization can be used. In the measurements presented in this thesis, regularization is not used. Therefore, the unfolding procedure is performed by minimizing the following expression:

$$\chi^2 = (y - Ax)^T V_{yy}^{-1} (y - Ax). \quad (4.11)$$

The response matrices  $A$ , are obtained from the simulated signal sample. The example of the response matrices of the measured variables is shown in Figure 4.12. The vertical axis on the histogram corresponds to the generated variable, while on the horizontal axis the reconstruction one is shown. On the histogram numbers

represent the probability that a value at the generator level is reconstructed as the specific value at the reconstruction level. The events that are used for determining the response matrix are the ones that pass the selection required for both reconstruction and generator levels.

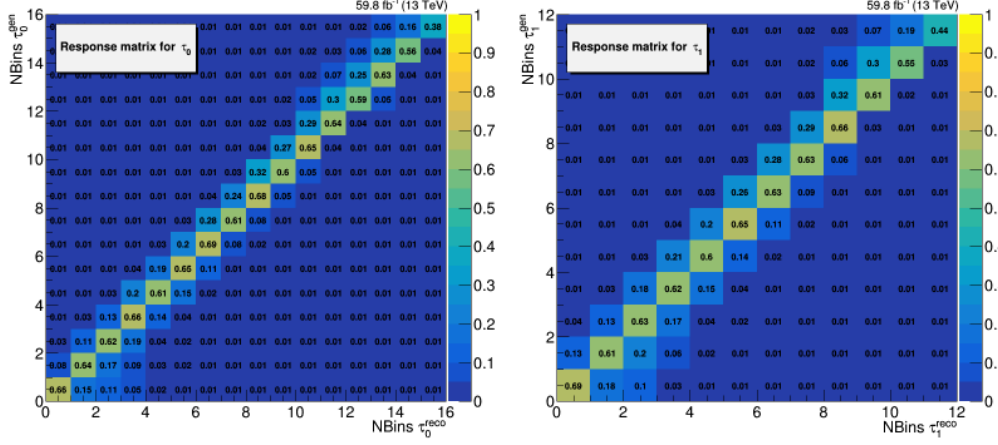


Figure 4.12: Response matrices for zero-jettiness (left) and one-jettiness (right).

## 4.6 Uncertainties

Two types of uncertainties are considered in the analysis: statistical uncertainties and systematic uncertainties. The statistical uncertainties are coming from limitations in the amount of data and from response matrices (amount of simulated events) which are estimated using the unfolding library. Systematic uncertainties include uncertainties from several different sources. Most of the systematical uncertainties are estimated by varying the corresponding parameter up and down by one standard deviation. The measurement is then performed with the modified parameter and the average of the difference with respect to the central value is taken as its standard deviation. Systematical uncertainties are added in quadrature assuming that each uncertainty is independent.

### Pileup

The uncertainty associated to the pileup reweighting applied to simulated events is obtained by varying the minimum bias cross section by 4.7% up and down. The unfolding procedure is then performed for both resulting pileup profiles.

### **Luminosity**

The uncertainty assigned to the luminosity is 2.5% [94]. It is applied as a global scale factor to extract the cross section and for the normalization of the simulated samples used to estimate the background.

### **Background**

The uncertainty in the estimation of the background that is subtracted from data is determined by varying the cross section by the largest uncertainty in the cross sections used for normalization. This uncertainty corresponds to the  $t\bar{t}$  process, it is 6% [96] and it takes into account PDF and scale uncertainties. The variation of the cross section is done before the background subtraction.

### **Lepton Energy Scale**

The uncertainty in the muon energy scale is estimated from the uncertainty in the Rochester corrections. The uncertainty includes the statistical component and that relative to the fitting method used (fitting function, Z mass window). These uncertainties are added in quadrature and used for varying the Rochester corrections up and down.

### **Lepton Energy Resolution**

The uncertainty in the lepton energy reconstruction is estimated in the signal samples by smearing the lepton by 0.6% with respect to the corresponding lepton at the generated level.

### **Lepton reconstruction and trigger efficiency**

The uncertainty in the reconstruction and trigger efficiency is estimated by varying the corresponding scale factors up and down by one standard deviation. For the identification and isolation scale factors, the considered uncertainties are of statistical nature. Trigger scale factors include both statistical and systematic uncertainties as explained in Section 4.4.

### **Track $p_T$**

The uncertainty in the track  $p_T$  is estimated for each track of each event. It is propagated analytically to the measured track-based variables assuming uncorrelated Gaussian uncertainties. The uncertainty in the reconstruction-level data distribution is derived from this event-by-event uncertainty with a toy Monte-Carlo using 100 replicas. The unfolding matrix is used to derive the uncertainty on the

measured differential cross sections, following the same method as the one used to propagate the data statistical uncertainties.

### Track efficiency

The uncertainty in the track reconstruction efficiency is estimated by varying the measured variable by 2.3% in the low- $p_T$  region (less than 20 GeV) and 1% in the higher- $p_T$  region. The 2.3% and 1% uncertainties are estimated from charged pions [103] and muons [104].

### Unfolding model

The uncertainty associated to the unfolding model is estimated by reweighting the signal simulation to match the data and using it as an alternative model for the unfolding. The weights are obtained from the ratio histograms of background-subtracted data to signal. The ratios are fitted with a polynomial to smooth out statistical fluctuations. An example of one of the fits is shown in Figure 4.13 for the  $\tau_1$  variable. The results of the fits are used to obtain weights for each event. After reweighting, a new response matrix is generated and used for the unfolding. The difference with respect to the central value is taken as the uncertainty.

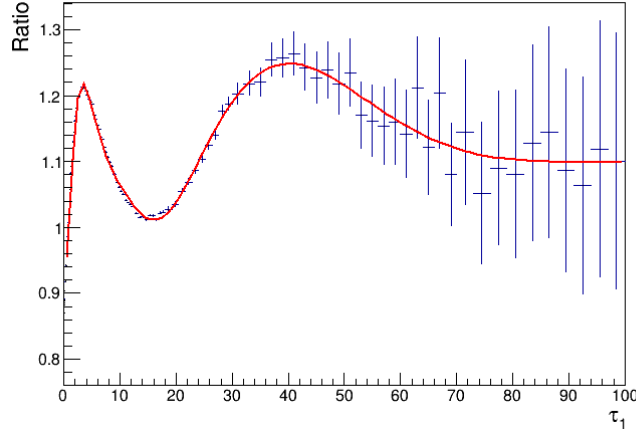


Figure 4.13: The fit function used for reweighting one-jettiness

### 4.6.1 Theoretical predictions uncertainties

The measured cross sections are compared with three different theoretical descriptions MADGRAPH5\_aMC@NLO (MG5\_aMC) at NLO, MG5\_aMC at LO, and GENEVA described in Section 1.4. The statistical uncertainties originating from the size of the simulated sample are estimated for each of these predictions. In addition, for MG5\_aMC NLO and GENEVA, the uncertainties in ME calculations are estimated.

For the MG5\_aMC NLO sample, the uncertainty originating from missing terms in the fixed-order calculation is obtained by varying the renormalization ( $\mu_R$ ) and factorization scales ( $\mu_F$ ) by factors 0.5, 1, and 2. The envelope of the variation is considered as uncertainty, with the excluded cases where the scales are varied in the opposite direction. The uncertainty in the extraction of the PDF set is estimated using the 100 replicas of NNPDF 3.0 NLO, where the standard deviation is taken as uncertainty. The uncertainty in  $\alpha_S$  is estimated by variation of the scale by its uncertainty (0.001) up and down.

The theoretical variations in the GENEVA sample are estimated using the 7-fold variation of the renormalization and factorization scale and using the variation of the scales used for resummation [105]. For each event, the calculation of its cross section with a different set of profile scales is performed and the corresponding weights are produced. The profile variations include variation up and down of the scales and variation of the transition points by  $\pm 0.05$ . The maximum absolute deviation from the central value among these six profiles is considered the resummation uncertainty. The total perturbative uncertainty is obtained by adding the fixed order uncertainty.

## 4.7 Results

The cross sections are obtained after background subtraction and unfolding procedures are computed and are compared to the theoretical prediction described in Section 1.4.

The uncertainty breakdown plots for zero-jettiness and one-jettiness are shown in Figure 4.14, for the results obtained in the Z peak region,  $76 < M_{\mu\mu} < 106$  GeV. It can be seen that the dominant source of uncertainty is the unfolding model. With the alternative model for unfolding, the uncertainty goes to 15% for the lowest values of the track-based variables. The second largest contribution to the total uncertainty comes from the track efficiency uncertainty.

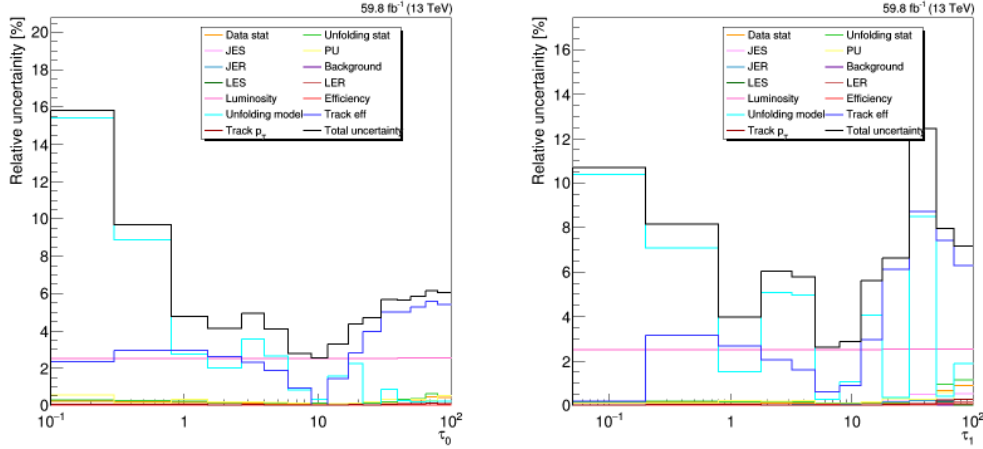


Figure 4.14: Uncertainty breakdown for zero-jettiness and one-jettiness

Because of the large uncertainties that come from the unfolding model itself, along with the unfolded results, the distributions at the reconstructed level are shown. For the detector level distributions, the uncertainties include only the statistical uncertainty coming from data. The background samples are included in the detector level distribution but have small contributions in the region of the Z peak, therefore they are not visible in the plots shown with a linear scale.

The measured cross sections as a function of zero-jettiness and the distribution at the detector level are shown in Figure 4.15. At low zero-jettiness, the NNLL GENEVA gives a better description than NLO and LO MG5\_aMC. For the high zero-jettiness region, all three predictions show a fair agreement with the measurement. The difference in the peak region between data and simulation can be caused by the different modeling of multiple parton interaction in simulation. The sensitivity of zero-jettiness on MPI is shown in Figure 4.16. Using the samples generated with GENEVA and showered with PYTHIA8, the zero-jettiness is computed in the Z peak region with and without MPI modeling included. It can be seen that the MPI affects the shape of the distribution, especially in the peak region. For MG5\_aMC and GENEVA, different PYTHIA8 tunes were used, CP5 and CUETP8M1 respectively.

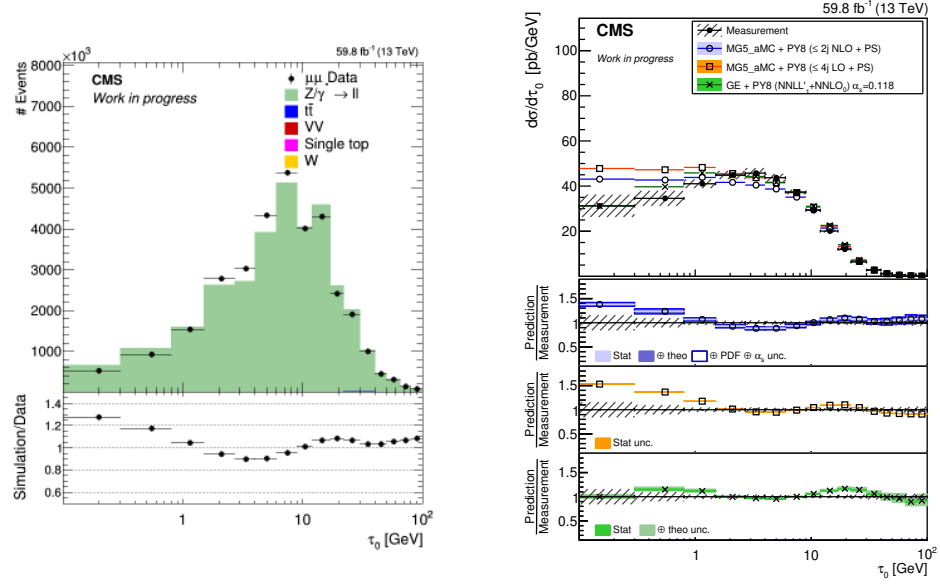


Figure 4.15: Data to simulation comparison and differential cross section as a function of zero-jettiness.

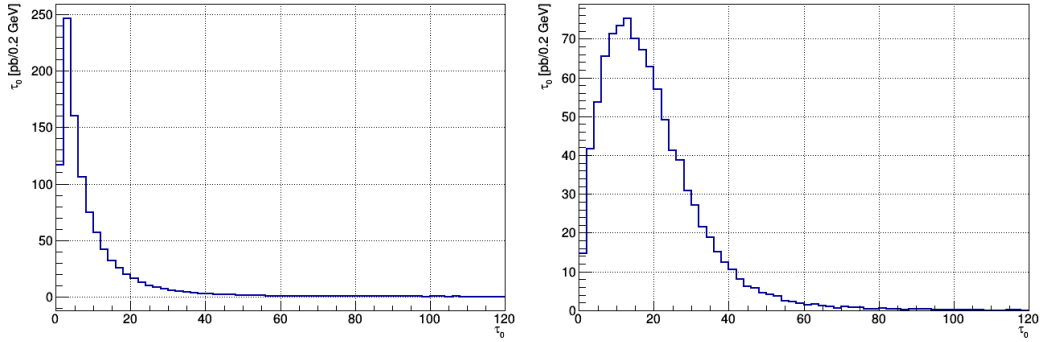


Figure 4.16: The distribution of zero-jettiness computed for the events without MPI (left) and with MPI (right).

The measured cross section as a function of zero-jettiness and the detector level comparison of zero-jettiness distribution in different  $p_T^Z$  bins are presented in Figure 4.17 - Figure 4.20. The four regions observed are :  $p_T^Z < 6$  GeV,  $6 < p_T^Z < 12$  GeV,  $12 < p_T^Z < 25$  GeV and  $p_T^Z > 25$  GeV.

In the lowest  $p_T^Z$  region, it is expected to have low jet activity from the primary collision and to be particularly sensitive to UE characteristics. In this region, none of the generators describe successfully the data. NLO and LO MG5\_aMC overestimate data while GENEVA underestimates the measurements. In the regions with  $p_T^Z$  6-12 GeV and 12-25 GeV, a better description of data by predictions is observed. In the region with  $p_T^Z > 25$  GeV, where it is expected to have at least one jet with high transverse momentum, there is still a significant difference in the MC generators with respect to the measurement. In this region, NLO MG5\_aMC describes the measurement better than the other two generators. While LO MG5\_aMC underestimates data, the GENEVA prediction overestimates the data in the whole range.

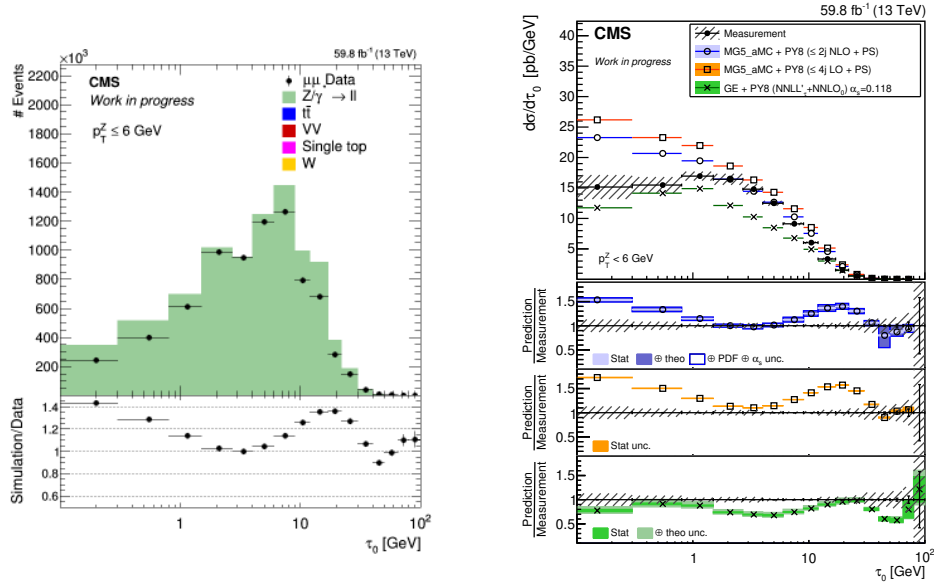


Figure 4.17: Data to simulation comparison and differential cross section as a function of zero-jettiness for  $p_T^Z < 6$  GeV.

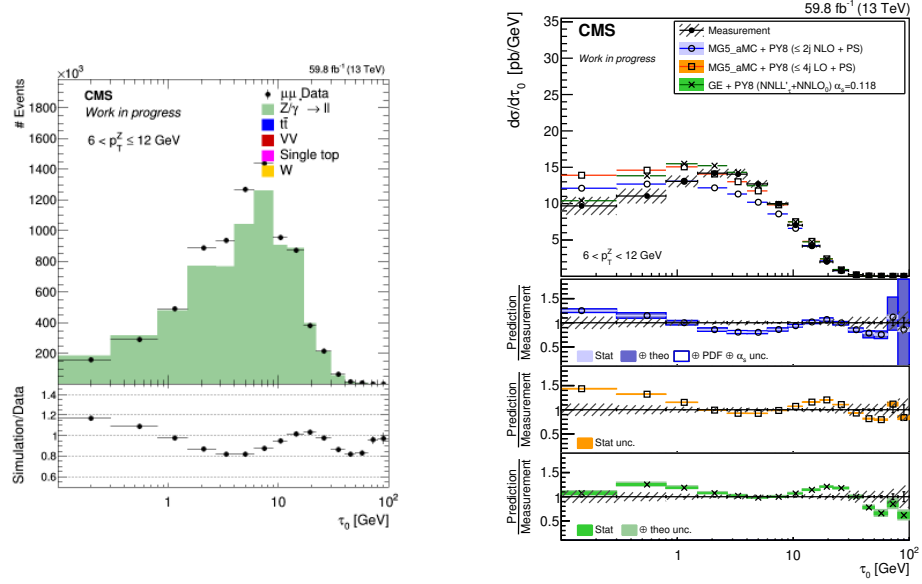


Figure 4.18: Data to simulation comparison and differential cross section as a function of zero-jettiness for  $6 < p_T^Z < 12 \text{ GeV}$ .

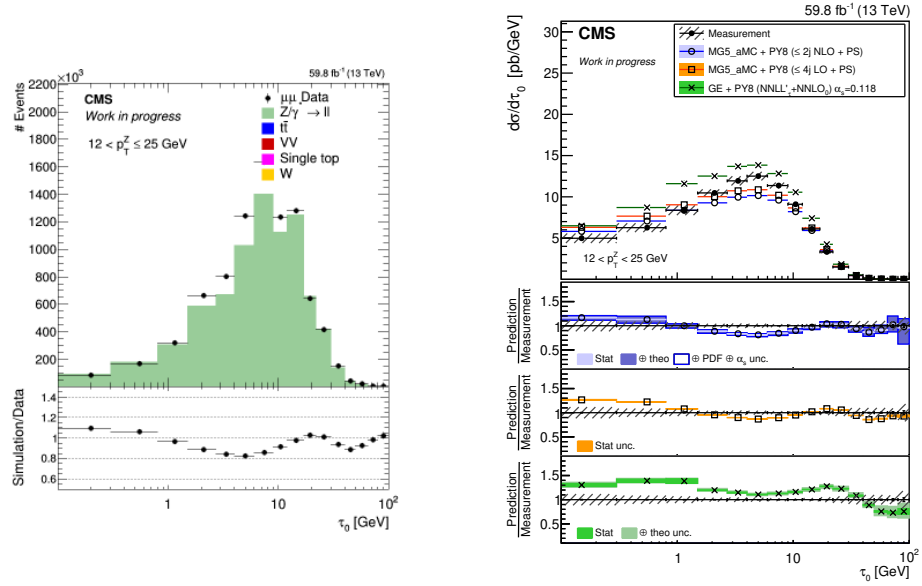


Figure 4.19: Data to simulation comparison and differential cross section as a function of zero-jettiness for  $12 < p_T^Z < 25 \text{ GeV}$ .

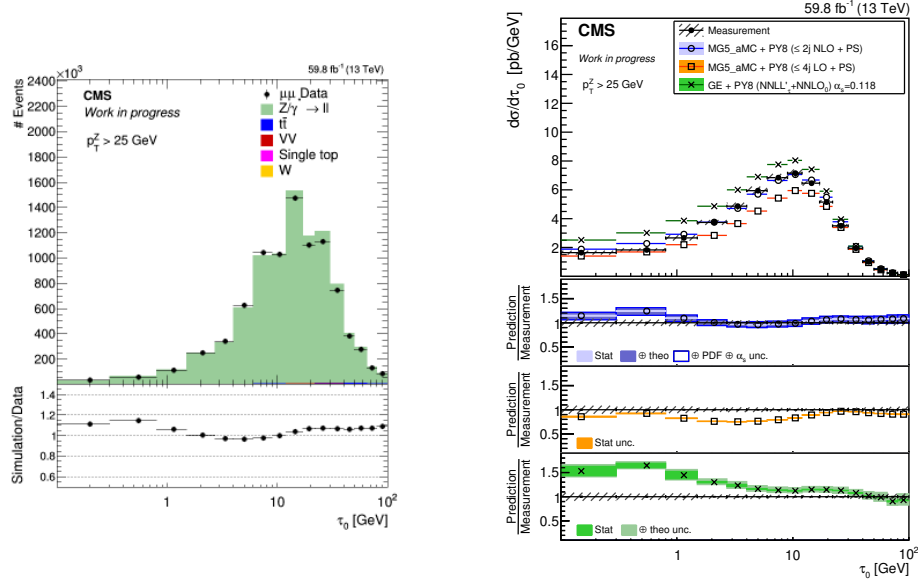


Figure 4.20: Data to simulation comparison and differential cross section as a function of zero-jettiness for  $p_T^Z > 25$  GeV.

The zero-jettiness, one-jettiness and the sum of the transverse momentum of particles are measured in the three additional invariant mass bins:  $125 < M_{\mu\mu} < 150$  GeV,  $150 < M_{\mu\mu} < 350$  GeV and  $350 < M_{\mu\mu} < 1500$  GeV.

The measured cross section as a function of zero-jettiness and the detector level comparison of zero-jettiness distribution in different  $M_{\mu\mu}$  bins are presented in Figure 4.21 - Figure 4.23.

From the detector level distributions, it can be seen that as going higher in the invariant mass regions, the contribution of the background becomes significant, in particular, from the  $t\bar{t}$  process.

In the invariant mass bin 125-150 GeV, the GENEVA prediction describes the data better than the other predictions. In the low zero-jettiness region, NLO and LO MG5\_aMC overestimate the data. As going towards a higher invariant mass bin, the description of data by MG5\_aMC improves.

Going to the higher mass bins, the dominant uncertainty source originates from the statistical limitations of data and simulation.

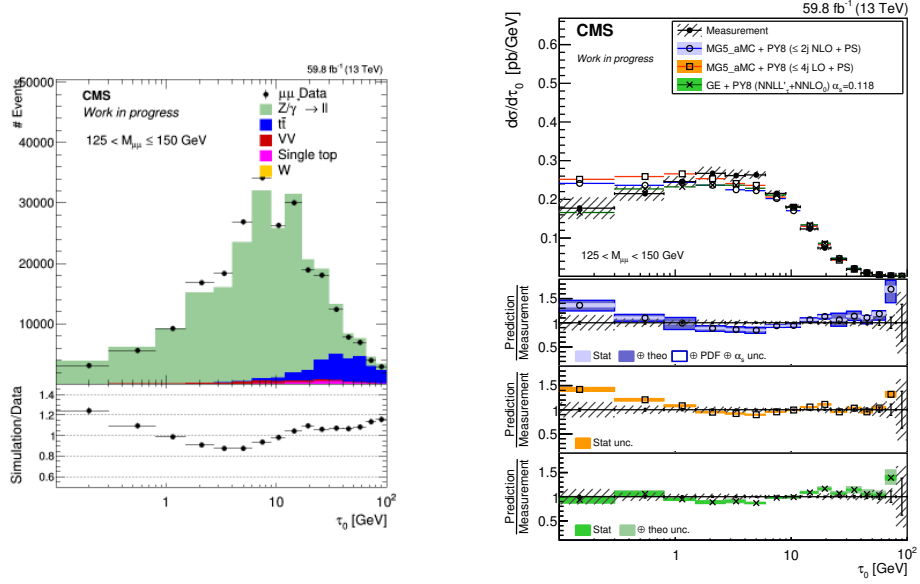


Figure 4.21: Data to simulation comparison and differential cross section as a function of zero-jettiness for  $125 < M_{\mu\mu} < 150$  GeV.

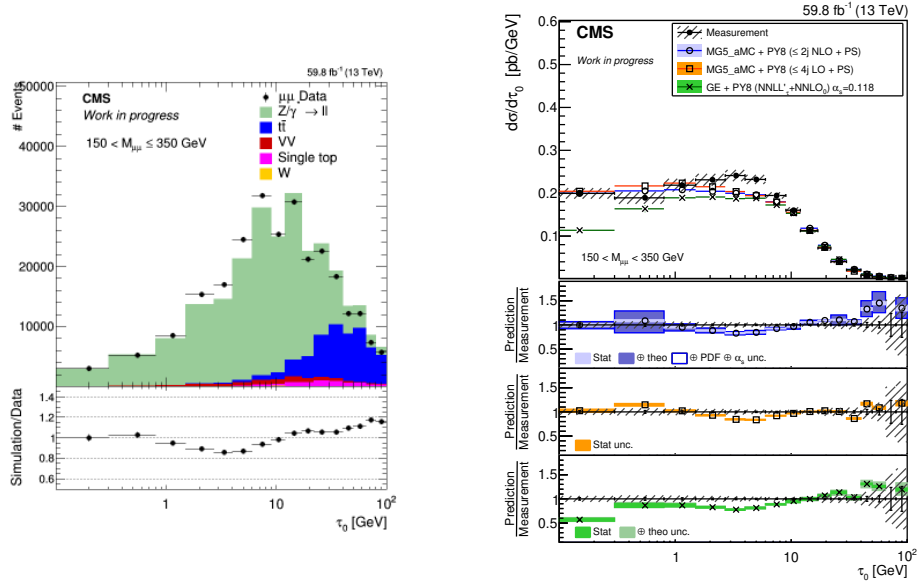


Figure 4.22: Data to simulation comparison and differential cross section as a function of zero-jettiness for  $150 < M_{\mu\mu} < 350$  GeV.

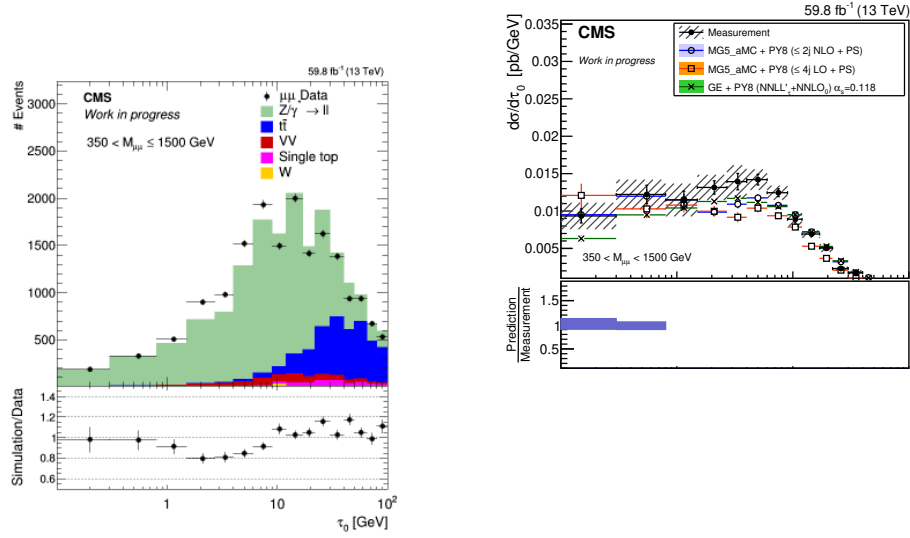


Figure 4.23: Data to simulation comparison and differential cross section as a function of zero-jettiness for  $350 < M_{\mu\mu} < 1500$  GeV.

The measured cross sections as a function of the one-jettiness and the distribution at the detector level are shown in Figure 4.24.

Similarly to the zero-jettiness, the low zero-jettiness region is better described by GENEVA than NLO and LO MG5\_aMC. For the higher one-jettiness, all three predictions show fair agreement with the measurement.

The measured cross section as a function of one-jettiness and the detector level comparison of the one-jettiness distribution in different  $p_T^Z$  bins are presented in Figure 4.25 - Figure 4.28. Unlike zero-jettiness, the one-jettiness is not measured in the bin where  $p_T^Z$  is lower than 6 GeV. In this region a high energy jet is emitted with a rather low probability and the distribution would be misleading.

The four regions where one-jettiness is observed are:  $6 < p_T^Z < 12$  GeV,  $12 < p_T^Z < 25$  GeV,  $25 < p_T^Z < 35$  GeV and  $p_T^Z > 35$  GeV.

In the lowest  $p_T^Z$  region, none of the predictions describes well the data. In the low one-jettiness region, all the predictions overestimate the data. For the higher  $p_T^Z$ , data is better described in the high one-jettiness region compared with the  $p_T^Z$  6-12 GeV. In the region  $p_T^Z > 35$  GeV, the LO MG5\_aMC describes rather well the data, while NLO MG5\_aMC and GENEVA show disagreements, especially in the low one-jettiness region.

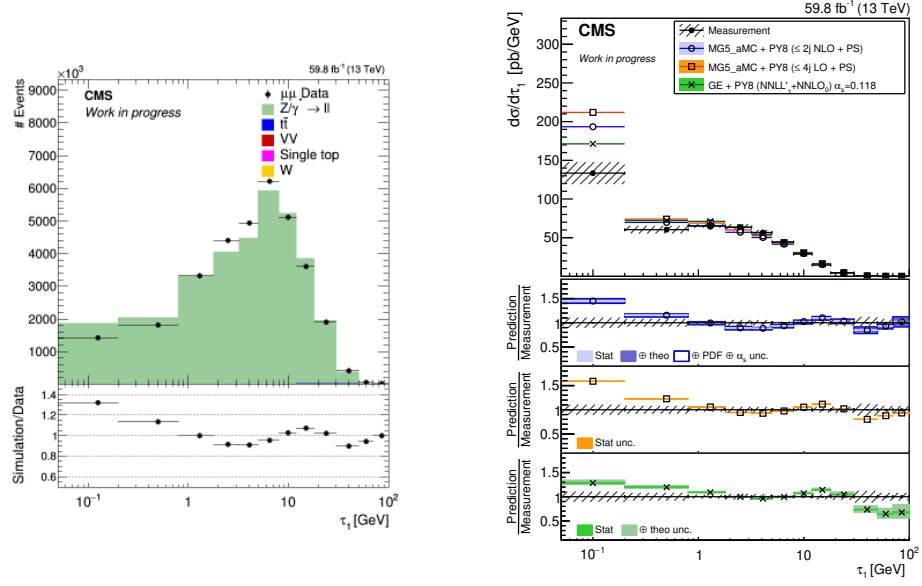


Figure 4.24: Data to simulation comparison and differential cross section as a function of one-jettiness.

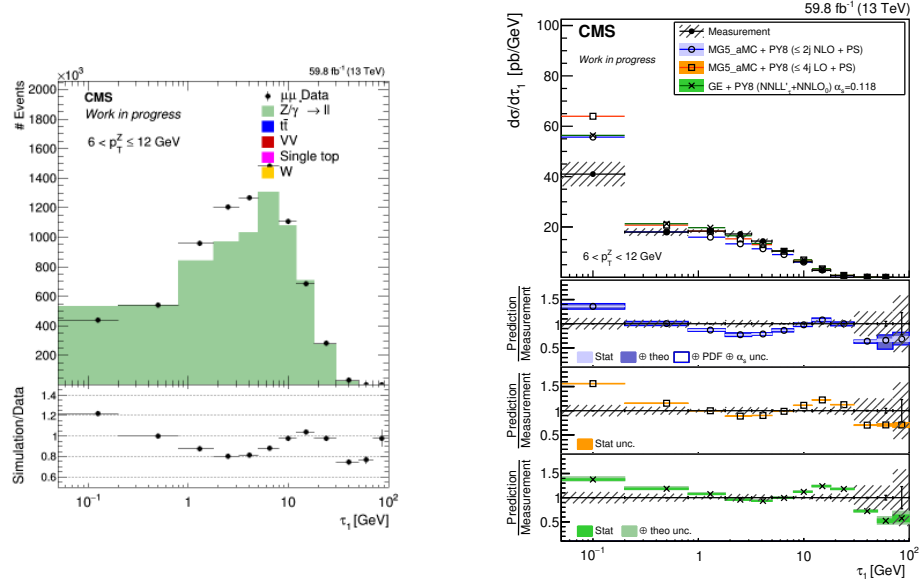


Figure 4.25: Data to simulation comparison and differential cross section as a function of one-jettiness for  $6 < p_T^Z < 12$  GeV.

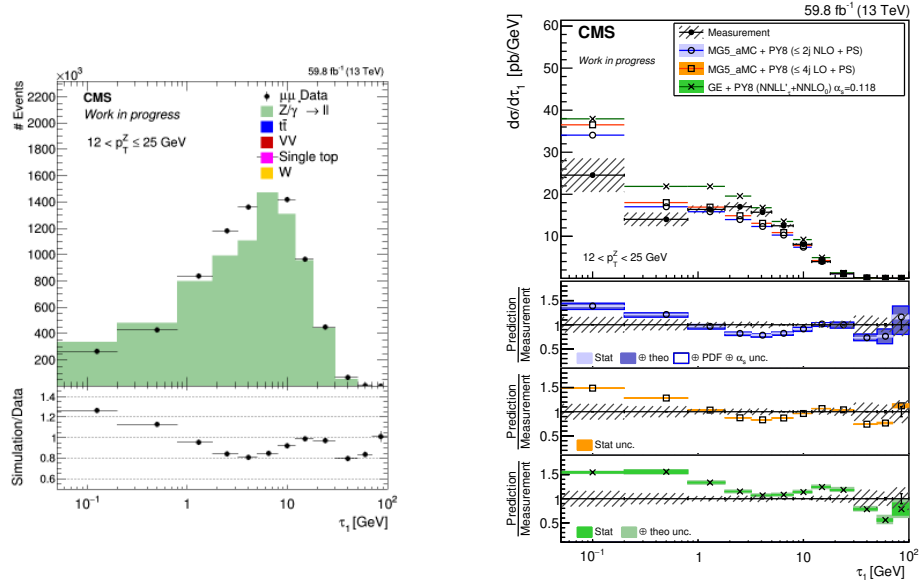


Figure 4.26: Data to simulation comparison and differential cross section as a function of one-jettiness for  $12 < p_T^Z < 25 \text{ GeV}$ .

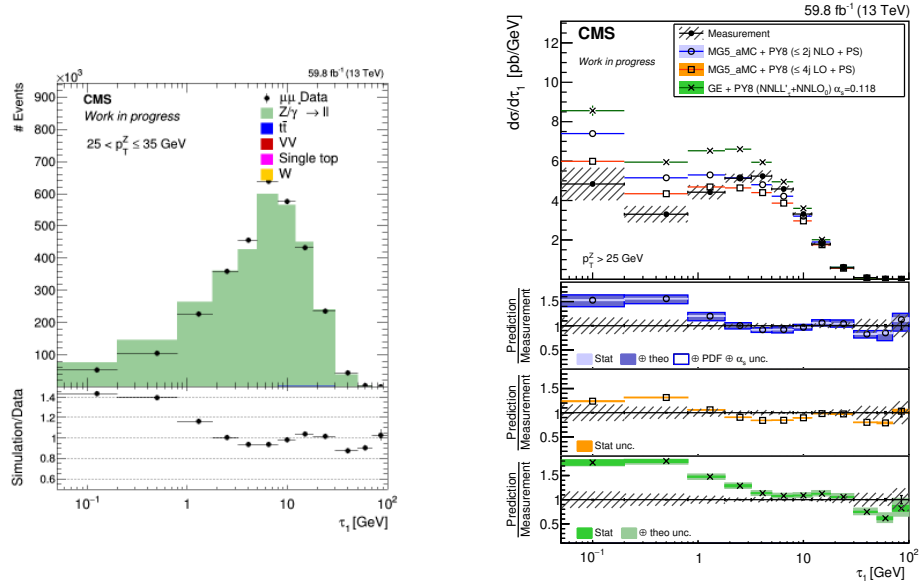


Figure 4.27: Data to simulation comparison and differential cross section as a function of one-jettiness for  $25 < p_T^Z < 35 \text{ GeV}$ .

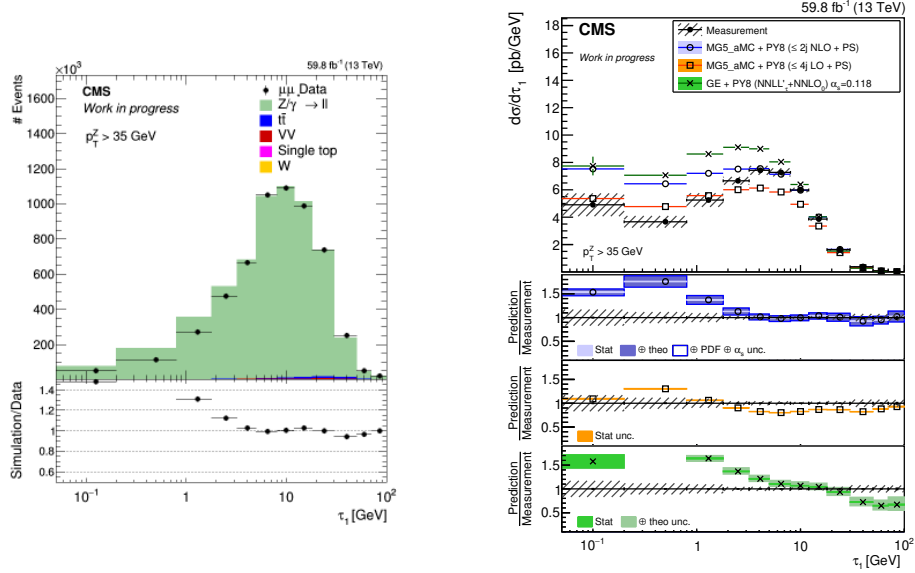


Figure 4.28: Data to simulation comparison and the differential cross section as a function of one-jettiness for  $p_T^Z > 35$  GeV.

The measured cross section as a function of the one-jettiness and the detector level comparison of one-jettiness distribution in different  $M_{\mu\mu}$  bins are presented in Figure 4.29 - Figure 4.31.

The large uncertainties, especially at the high one-jettiness originate from the statistical limitations of data and simulation.

In the mass bin  $M_{\mu\mu}$  in range 125-150 GeV, GENEVA shows a fair agreement with data. NLO and LO MG5\_aMC overestimate the data in the low one-jettiness region. In the higher invariant mass bins, all the predictions perform rather well.

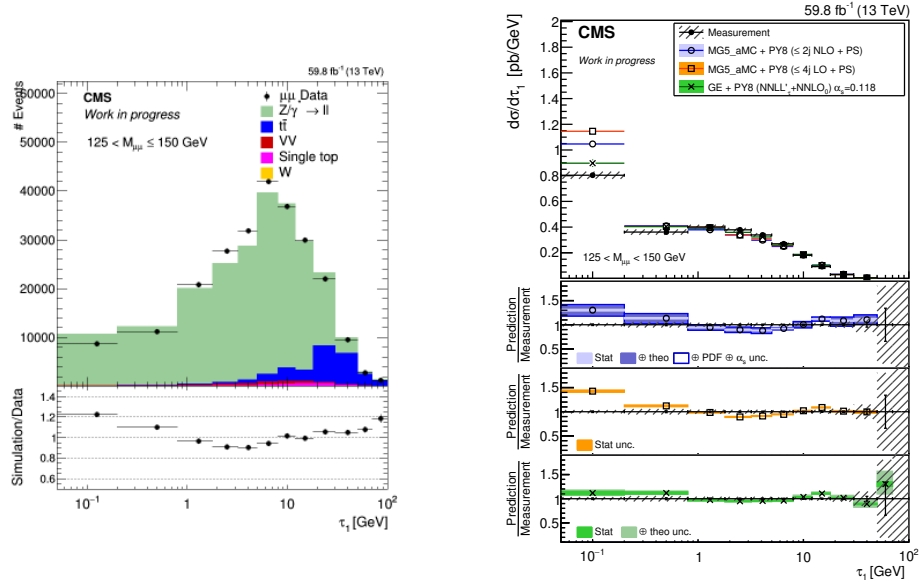


Figure 4.29: Data to simulation comparison and differential cross section as a function of one-jettiness for  $125 < M_{\mu\mu} < 150$  GeV.

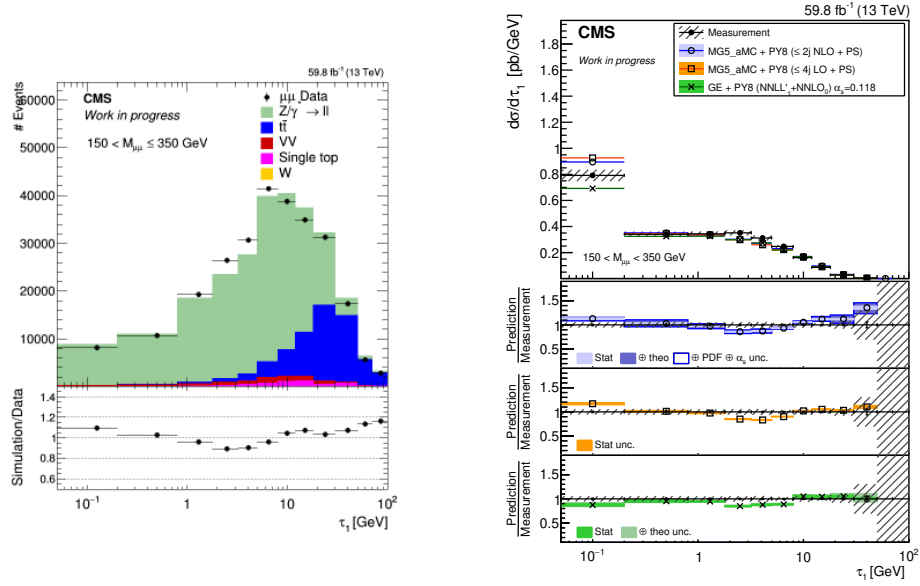


Figure 4.30: Data to simulation comparison and differential cross section as a function of one-jettiness for  $150 < M_{\mu\mu} < 350$  GeV.

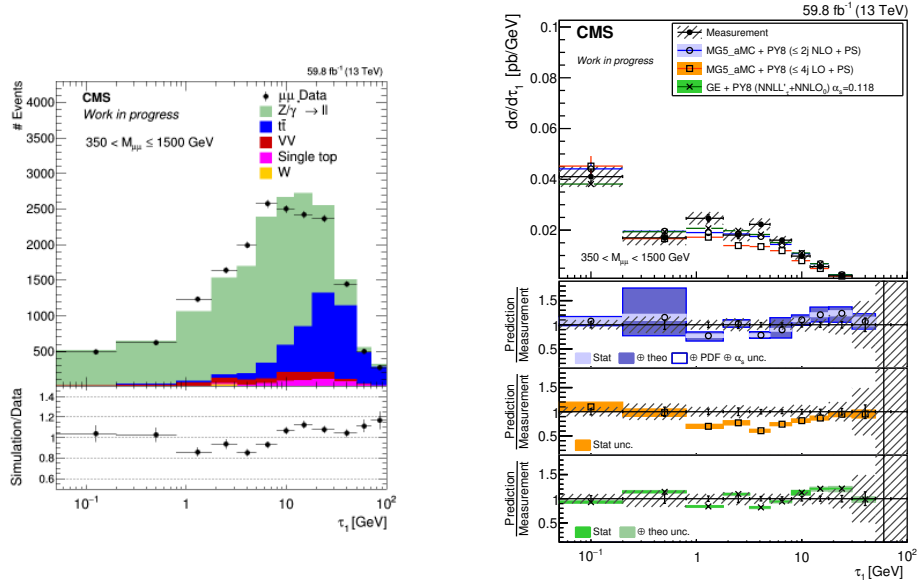


Figure 4.31: Data to simulation comparison and differential cross section as a function of one-jettiness for  $350 < M_{\mu\mu} < 1500$  GeV.

The measured cross sections as a function of the sum of the transverse momentum of the charged particles and the distribution at the detector level are shown in Figure 4.32.

In the low  $p_{T,sum}$  region, none of the predictions describes well the data, while in the higher  $p_{T,sum}$  region there is a good agreement with data observed for all the predictions.

The measurement of the sum of the transverse momentum of particles is also performed in four different  $p_T^Z$  bins:  $p_T^Z < 6$  GeV,  $6 < p_T^Z < 12$  GeV,  $12 < p_T^Z < 25$  GeV and  $p_T^Z > 25$  GeV. The measured cross section as a function of  $p_{T,sum}$  and the detector level comparison of the  $p_{T,sum}$  distribution in different  $p_T^Z$  bins are presented in Figure 4.33 - Figure 4.36.

In the region with the lowest  $p_T^Z$ , none of the predictions describe well the data. As going towards the higher  $p_T^Z$  regions, where it is expected to have higher jet activity, data are better described by all predictions. In the region where  $p_T^Z$  is higher than 25 GeV, MG5\_aMC is describing data better than other predictions.

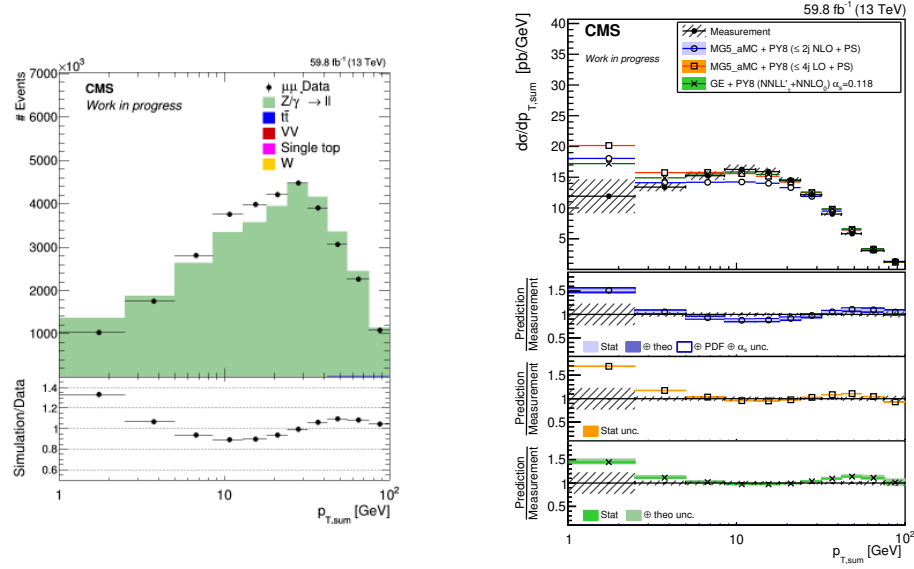


Figure 4.32: Data to simulation comparison and differential cross section as a function of the sum of the transverse momentum of charged particles.

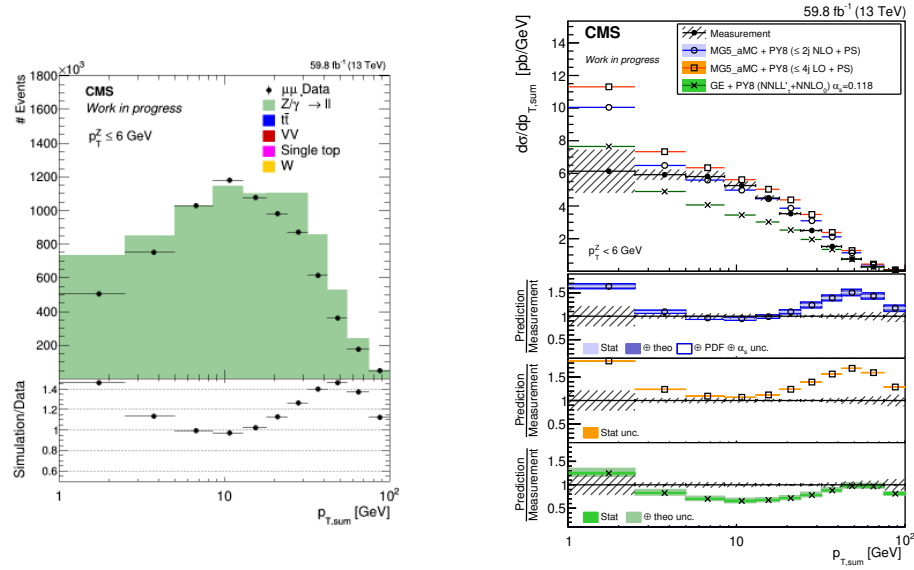


Figure 4.33: Data to simulation comparison and differential cross section as a function of the sum of the transverse momentum of charged particles for  $p_T^Z < 6$  GeV.

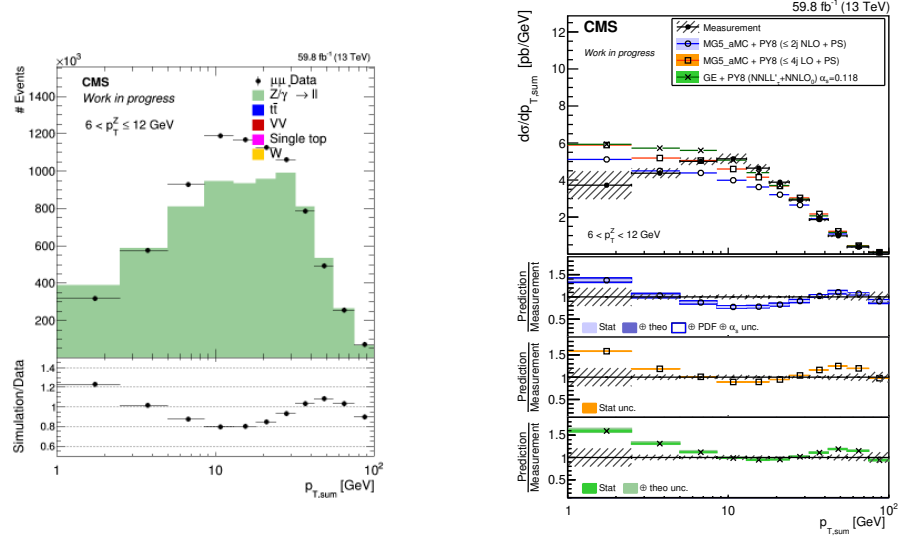


Figure 4.34: Data to simulation comparison and differential cross section as a function of the sum of the transverse momentum of charged particles  $6 < p_T^Z < 12 \text{ GeV}$ .

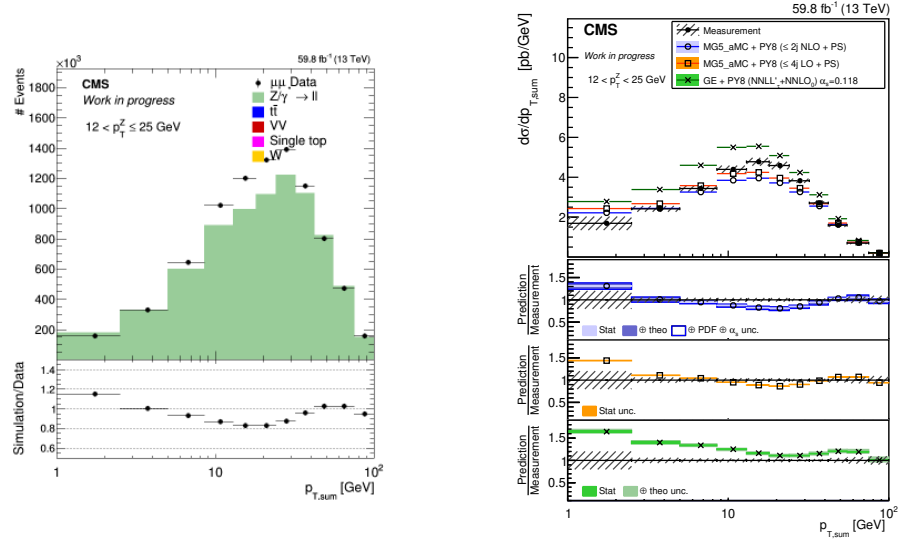


Figure 4.35: Data to simulation comparison and differential cross section as a function of the sum of the transverse momentum of charged particles for  $12 < p_T^Z < 25 \text{ GeV}$ .

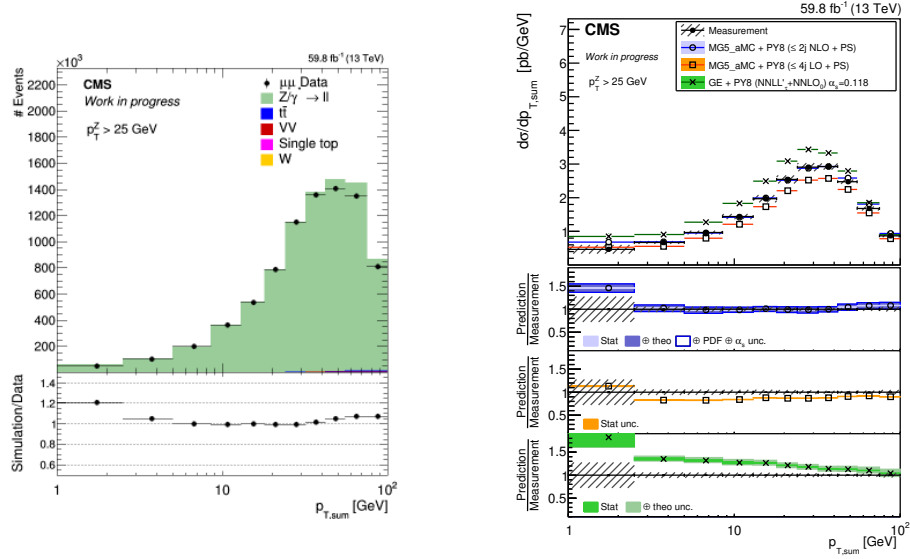


Figure 4.36: Data to simulation comparison and differential cross section as a function of the sum of the transverse momentum of charged particles for  $p_T^Z > 25$  GeV.

The measured cross section as a function of  $p_{T,sum}$  and the detector level comparison of the  $p_{T,sum}$  distribution in different  $M_{\mu\mu}$  bins are presented in Figure 4.37 - Figure 4.39. The large uncertainties, especially at the low  $p_{T,sum}$  region have large contribution from data and MC statistical uncertainties.

In the region where the invariant mass is in the range 125-150 GeV, all the predictions describe well data within the uncertainties. In the higher mass bins, the data is underestimated by the LO MG5\_aMC and GENEVA.

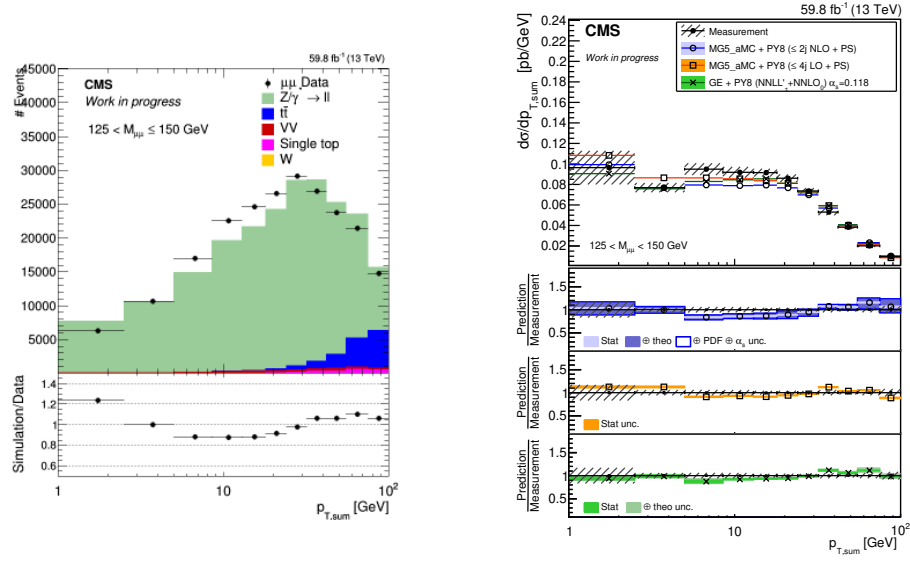


Figure 4.37: Data to simulation comparison and differential cross section as a function of the sum of the transverse momentum of charged particles for  $125 < M_{\mu\mu} < 150$  GeV.

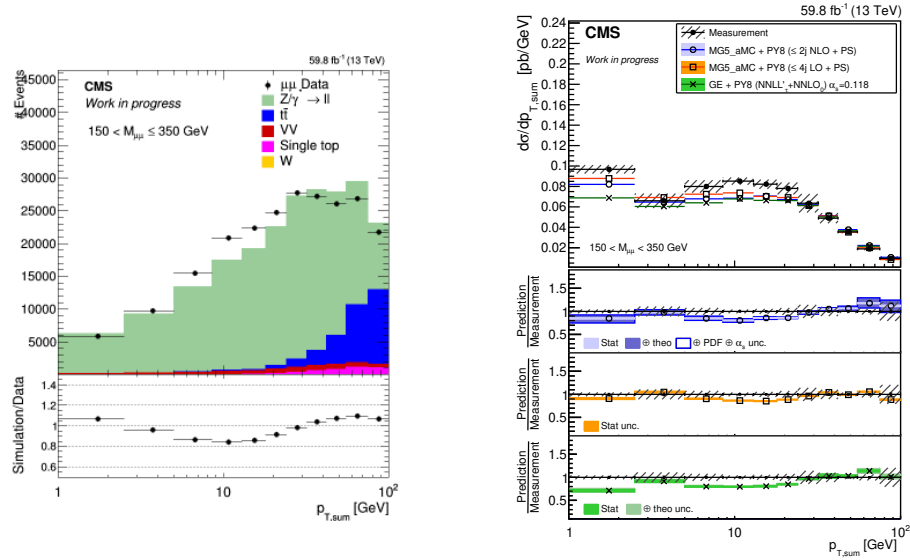


Figure 4.38: Data to simulation comparison and differential cross section as a function of the sum of the transverse momentum of charged particles for  $150 < M_{\mu\mu} < 350$  GeV.

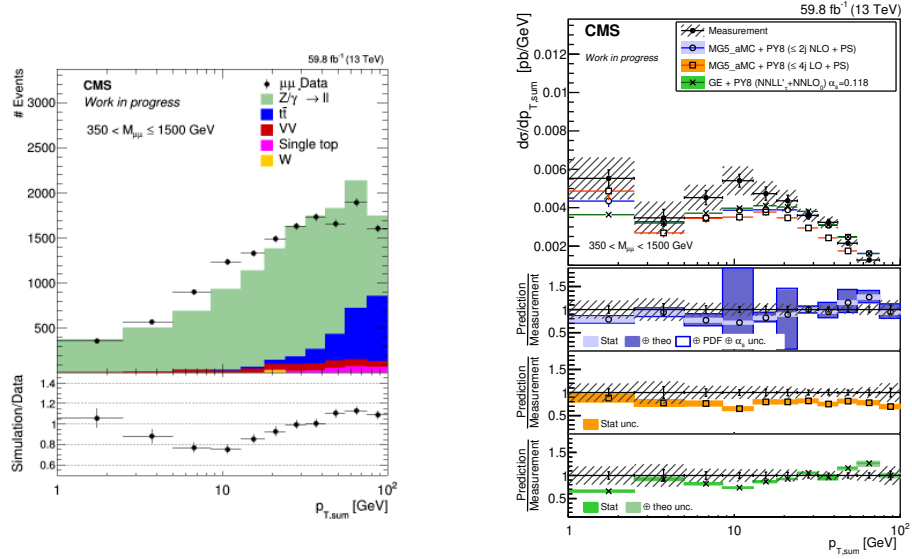


Figure 4.39: Data to simulation comparison and differential cross section as a function of the sum of the transverse momentum of charged particles for  $350 < M_{\mu\mu} < 1500$  GeV.

# Chapter 5

## Measurement of jet-based event shape observables

In this chapter, the measurements of jet-based event shape observables in the production of the Z boson in association with jets are presented. These observables are measured in events with an on-shell Z boson, where the invariant mass of the decay leptons is required to be 76 to 106 GeV. The definition of the jet-based observables is presented Section 5.1. The data samples and simulation are the same as defined in Section 4.2. In addition, the corrections for jets and the criteria for selecting events with a Z boson and one or more jets are presented in Section 5.2. To obtain the result, the TUnfold method is used. Additional uncertainties related to jets are defined in Section 5.3. The final results are presented in Section 5.4.

### 5.1 Observable definition

The vetoes on the hadronic activity are used for classifying events with the minimum number of jets (inclusive) or with an exact number of jets (exclusive), to suppress background, or to increase sensitivity for a particular process of interest. To select the events based on their jet multiplicity, the usual requirement is to have a specific number of jets with their transverse momentum above some thresholds. Besides the transverse momentum, additional variables can be used for selection as defined in [8]. These variables depend on the jet transverse momentum  $p_T^j$  and a weighting function depending on the jet rapidity  $f(y_i)$ :

$$\tau_j = p_T^j f(y_i). \tag{5.1}$$

Two different weighting function are considered:

$$\begin{aligned}\tau_{Bj} : f(y_j) &= e^{-|y_j - Y|} \\ \tau_{Cj} : f(y_j) &= \frac{1}{2 \cosh(y_j - Y)},\end{aligned}\tag{5.2}$$

where  $Y$  is the rapidity of the dimuon system.

The weighting functions are decreasing functions of the absolute jet rapidity. Therefore, small values of  $\tau_j$  are obtained either for small values of the transverse momentum or for large values of absolute rapidity. In Figure 5.1 the weighting functions are shown. It can be seen that for values of  $|y_j|$  greater than 1.5 the  $\tau_B$  and  $\tau_C$  are equivalent, while for rapidities around 0,  $\tau_C$  is smoother than  $\tau_B$ . Using  $\tau_B$  and  $\tau_C$  for the event selection allows to have a constraint that is tight at central rapidity and gets looser for the forward rapidity. This type of selection is convenient for processes that have jets well separated in rapidity, such as the vector boson fusion.

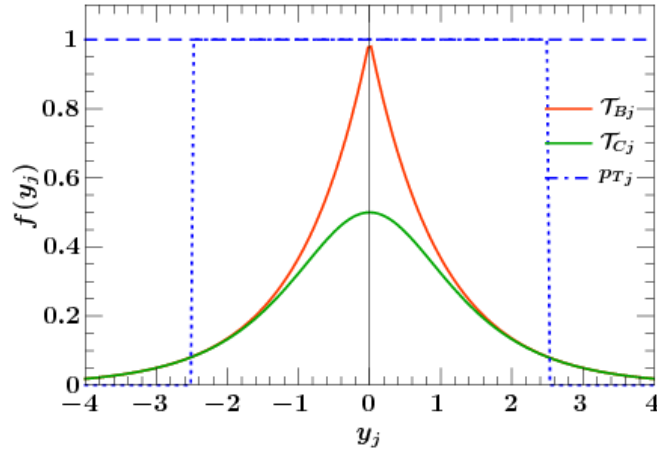


Figure 5.1: Rapidity weighting functions for  $\tau_B$  and  $\tau_C$  [8].

Compared to the  $\tau_B$ , the  $\tau_C$  variable has an experimental advantage because it can be measured to smaller values which is why this chapter focuses on it.

The differential cross section is measured for two variables derived from Eq. 5.2:

$$\begin{aligned}\tau_{max} &= \max_j \tau_{Cj} = \max \left( m_{Tj} \frac{1}{2 \cosh(y_j - Y)} \right) \\ \tau_{sum} &= \sum_j \tau_{Cj} = \sum_J \frac{m_{Tj}}{2 \cosh(y_j - Y)},\end{aligned}\tag{5.3}$$

where  $m_{Tj} = \sqrt{p_{Tj}^2 + m_j^2}$ , for a jet with mass  $m_j$ .

The variable  $\tau_{\max}$  corresponds to the highest  $\tau_{Cj}$  in the event, while  $\tau_{\text{sum}}$ , which is analog to the zero-jettiness, is the scalar sum of the  $\tau_{Cj}$  in the event.

## 5.2 Jet reconstruction and selection

The jet reconstruction algorithms are used for combining the information from the calorimetry and the tracking system in order to build jets. In this thesis, jets are reconstructed with the anti- $k_T$  jet clustering algorithm [106] applied on the particle flow candidates.

The anti- $k_T$  groups particles based on the momentum space. It starts with determining the distance between entities  $i$  and  $j$  ( $d_{ij}$ ) and the distance between the beam and entity  $i$  ( $d_{iB}$ ):

$$\begin{aligned} d_{ij} &= \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2} \\ d_{iB} &= k_{ti}^{2p}, \end{aligned} \tag{5.4}$$

where  $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$  and  $k_{ti}$ ,  $y_i$  and  $\phi_i$  are transverse momentum, rapidity and azimuthal angle of the entity  $i$ .  $R$  is the radius parameter that determines the size of the jet and  $p$  is the parameter that sets the relative power of the energy with respect to the geometrical scales. In this analysis, the  $R$  is set to 0.4 and  $p$  is set to 1. The jet clustering algorithm then determines the minimum of the two distances. If  $d_{ij}$  is the minimum, the entries  $i$  and  $j$  are combined into one entry by summing their four-momentum. If  $d_{iB}$  is the minimum,  $i$  is considered to be a final jet. The clustering process is repeated over all entries.

The jet reconstruction is affected by the particles coming from pileup. In order to mitigate the pileup events the techniques such as Charged Hadron Subtraction (CHS) [61] or Pileup Per Particle Identification (PUPPI) [107] can be used.

The CHS algorithm uses the information from the tracking system to identify particles that are originating from pileup vertices. These charged particles are then removed from the event, and the remaining charged particles and all neutral particles are used for jet reconstruction.

The PUPPI algorithm considers the particle level candidates and assigns them weight in the range from 0 to 1, where the value 1 is assigned to the particles that come from the primary vertex while the particles originating from pileup have a weight 0. For charged particles, the weight assignment is based on the information from the tracking system. Charged particles involved in the fit of the primary vertex are assigned a weight close to 1, while charged particles associated with

pileup vertices are assigned a weight close to 0. The weights for neutral particles are assigned based on the discriminating variable  $\alpha$ :

$$\alpha_i = \log \sum_{j \neq i, \Delta R_{ij} < \Delta R} \left( \frac{p_{T,j}}{\Delta R_{ij}} \right)^2, \quad (5.5)$$

where the sum runs over the particles in a cone with radius 0.4 of particle  $i$ ,  $p_{T,j}$  is the transverse momentum of particle  $j$  and  $\Delta R_{ij}$  is the distance between the particles  $i$  and  $j$  in the  $\eta - \phi$  frame. In the region with  $|\eta| < 2.5$ ,  $j$  are charged particles from the primary vertex while in the region  $|\eta| > 2.5$   $j$  are all the reconstructed particles. The  $\alpha_i$  value of neutral particles is compared with the median ( $\bar{\alpha}_{PU}$ ) and the RMS ( $\alpha_{PU}^{RMS}$ ) of the  $\alpha$  distribution obtained for charged particles from pileup, using a signed  $\chi^2$  approximation:

$$\text{signed}\chi^2 = \frac{(\alpha_i - \bar{\alpha}_{PU})|\alpha_i - \bar{\alpha}_{PU}|}{(\alpha_{PU}^{RMS})^2}. \quad (5.6)$$

The neutral particles with large signed  $\chi^2$  is large are most likely originating from the primary vertex. The final weights for neutral particles are computed with a cumulative distribution function of the signed  $\chi^2$  with one degree of freedom:

$$w_i = F_{\chi^2, NDF=1}(\text{signed}\chi^2). \quad (5.7)$$

The resulting weights are used for rescaling the four-momentum of the particles in the jet to reduce the pileup contribution.

To account for inefficiencies, nonlinearities, and for the finite resolution in energy and position of the reconstructed jets, jet energy corrections (JEC) are derived from simulations and applied to both CHS and PUPPI jets. In addition, since the jet energy resolution is different in data and simulation, a smearing of the simulated jet energies is performed. The four-momentum of a reconstructed jet is rescaled with the factor:

$$C_{\text{JER}} = 1 + (S_{\text{JER}} - 1) \frac{p_T - p_T^{\text{gen}}}{p_T}, \quad (5.8)$$

where  $p_T$  is the transverse momentum of the reconstructed jet,  $p_T^{\text{gen}}$  is the transverse momentum of the corresponding jet at generator level and  $S_{\text{JER}}$  is the data-to-simulation core resolution scale factor. If the corresponding generator level jets are not found, a stochastic smearing is performed. In that case, the correction factor is computed as:

$$C_{\text{JER}} = 1 + N(0, \sigma_{\text{JER}}) \sqrt{\max(s_{\text{JER}}^2 - 1, 0)}, \quad (5.9)$$

where  $\sigma_{\text{JER}}$  is the relative  $p_T$  resolution in simulation and  $N(0, \sigma_{\text{JER}})$  is the random number from the normal distribution with a zero mean and variance  $\sigma^2$ .

In order to decide which jet algorithm to use, studies on the performance of the CHS and PUPPI methods are performed. In the following, jets are required to be within the pseudorapidity range of  $|\eta| < 2.5$  and events considered are the ones with at least one jet with transverse momentum above 20 GeV. In order to avoid misidentification and to reject the background, the identification selection summarized in Table 5.1 is used [?].

| Variable                    | Selection |
|-----------------------------|-----------|
| Charged Hadron Function     | $> 0$     |
| Charged Hadron Multiplicity | $> 0$     |
| Charged EM Fraction         | $< 0.80$  |
| Neutral Hadron Fraction     | $< 0.90$  |
| Neutral EM Fraction         | $< 0.90$  |
| Muon Fraction               | $< 0.80$  |

Table 5.1: Criteria for identification of jets [108].

In addition, the jets are required to be separated from the muons by selecting only the jets with  $\Delta R(\text{jet}, \text{muon}) > 0.4$ .

In Figure 5.2, the distribution of the leading jet transverse momentum is shown at generator level and at reconstructed level for both CHS and PUPPI jets. For  $p_T$  above 30 GeV, both the CHS and PUPPI reconstruction show good agreement with the generator level jets  $p_T$ . For lower jet transverse momentum (from 20 to 30 GeV), a better description is given by PUPPI jets.

To see the performance of the pileup identification, the ratio of pileup jets to genuine jets is studied. Jets are classified according to the difference between the azimuthal angle  $\phi$  of the leading jet and the Z boson. Pileup jets are required to have  $\Delta\phi(\text{Z}, \text{jet}) < 1.5$ , while non pileup jets have  $\Delta\phi(\text{Z}, \text{jet}) > 2.5$ . In Figure 5.3 it can be seen that CHS jets have a strong dependence on the number of vertices in the detector, especially for the events where it is required to have a leading jet with transverse momentum higher than 10 GeV and 20 GeV. PUPPI shows a stable behavior even for a leading jet  $p_T$  of 20 GeV.

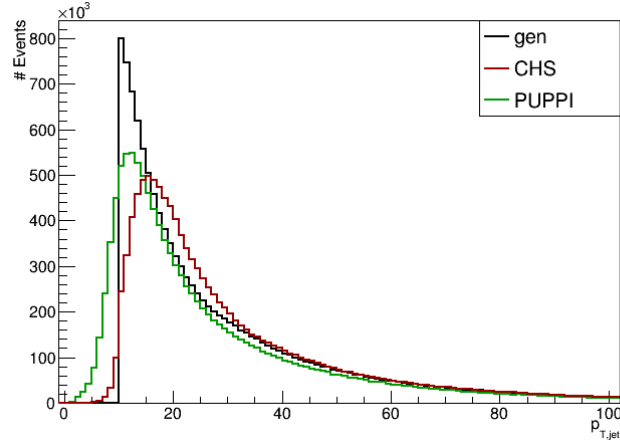


Figure 5.2: Comparison of the leading jet transverse momentum distribution at generated level, reconstructed level with CHS and reconstructed level with the PUPPI algorithm.

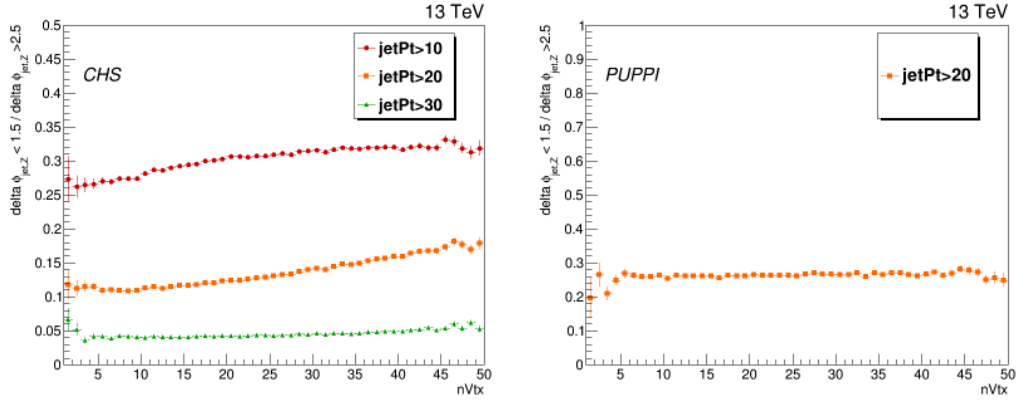


Figure 5.3: Data to simulation comparison of the leading jet transverse momentum and inclusive number of jets.

Since for measuring jet-based event shape variables, it is important to go as low as possible in  $p_T$ , PUPPI jets are used and events are selected such that there is at least one jet with transverse momentum above 20 GeV.

After the selections and corrections are applied, the comparison of the transverse momentum of the leading jet (without restrictions on the leading jet  $p_T$ ) from the event and of the number of reconstructed jets with  $p_T > 20$  GeV in data

and simulation is shown in Figure 5.4. It can be seen that there is a 10% difference between data and simulation in the distributions of the jet transverse momentum.

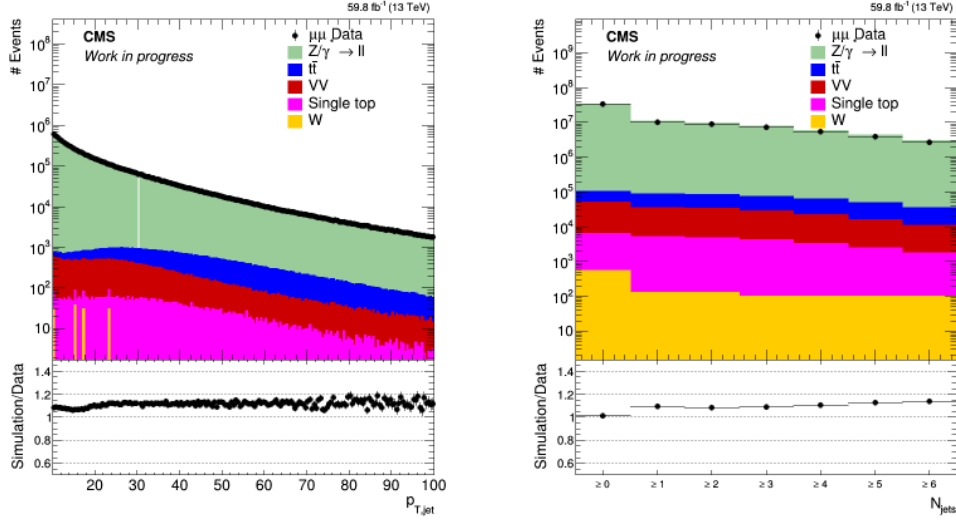


Figure 5.4: Data to simulation comparison of the leading jet transverse momentum and inclusive number of jets.

## 5.3 Uncertainties

In addition to those defined in Section 4.6, uncertainties are assigned to account for the jet energy scale and the jet energy resolution.

### Jet energy scale

The uncertainty in the jet energy scale is estimated by scaling the jet momentum in data by uncertainties that are common for all CMS analyses. These factors are  $p_T$  and  $\eta$  dependent.

### Jet energy resolution

The uncertainties in the resolution of jets are determined by varying the smearing factor used in simulations by their uncertainty.

## 5.4 Results

The jet-based variables are measured in events with the production of a real boson in an invariant mass range of  $76 < m_{ll} < 106$  GeV. The main source of uncertainty comes from the jet energy correction and the jet energy scale. In the low  $\tau_{\max}$  and low  $\tau_{\text{sum}}$  regions, these uncertainties go up to %. The uncertainties breakdown plots are shown in Figure 5.5.

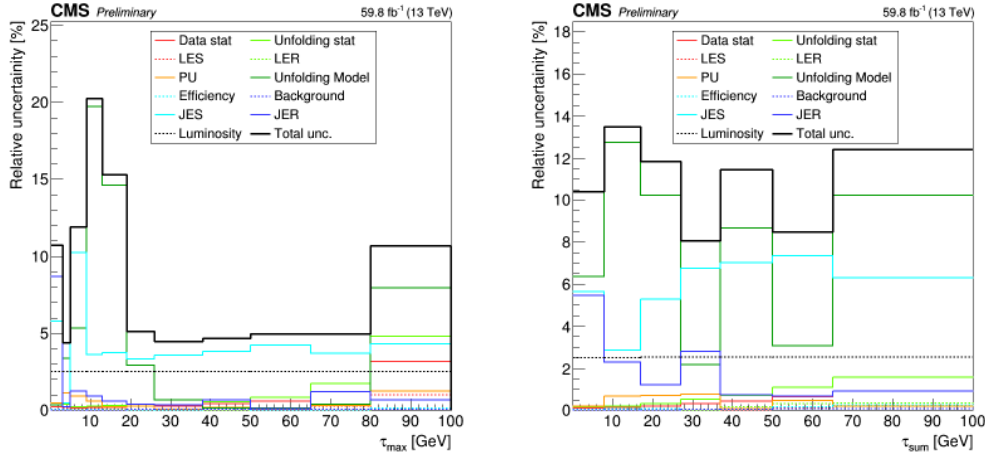


Figure 5.5: Uncertainty breakdown for  $\tau_{\max}$  and  $\tau_{\text{sum}}$ .

The measured cross section as a function of  $\tau_{\max}$  and the detector level comparison of the  $\tau_{\max}$  distribution are shown in Figure 5.6. At low  $\tau_{\max}$ , the best generator for describing the measurement is NLO MG5\_aMC. In the peak region of the distribution, GENEVA overestimates the data while at the high  $\tau_{\max}$  NLO MG5\_aMC and GENEVA show good agreement and LO MG5\_aMC underestimates data. Since it shows good agreement with predictions,  $\tau_{\max}$  can be used as a veto for jets.

The measured cross section as a function of  $\tau_{\text{sum}}$  and the detector level comparison of the  $\tau_{\text{sum}}$  distribution are shown in Figure 5.7. The low  $\tau_{\text{sum}}$  region is best described by the LO MG5\_aMC. Both NLO MG5\_aMC and GENEVA overestimate the data. At high  $\tau_{\text{sum}}$ , NLO MG5\_aMC gives the best description of data, while LO MG5\_aMC and GENEVA predict fewer events in this region.

The differences with the GENEVA prediction were observed in the previous measurements performed [109]. It was observed that the GENEVA prediction has a smaller accuracy for the higher jet multiplicities, where one or more jets arise from the parton shower.

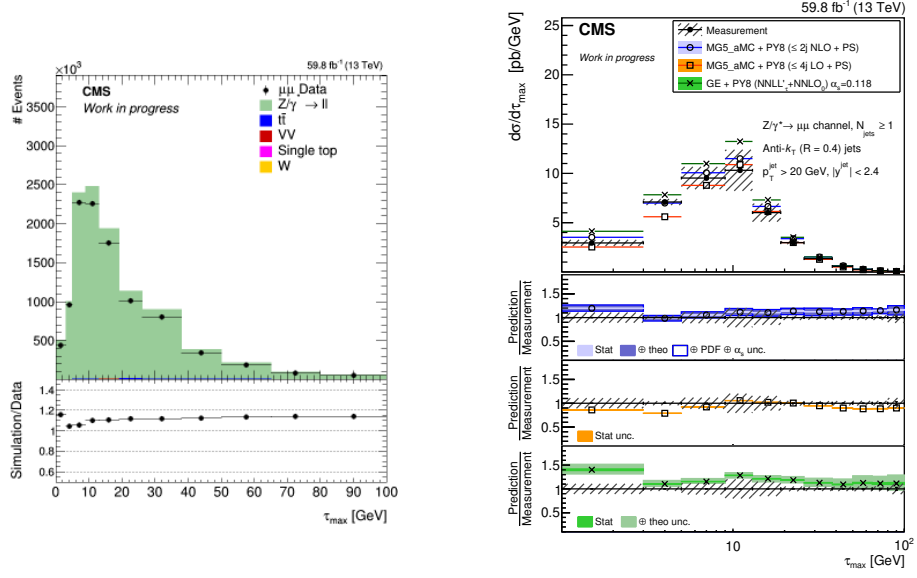


Figure 5.6: Data to simulation comparison and differential cross section as a function of  $\tau_{max}$ .

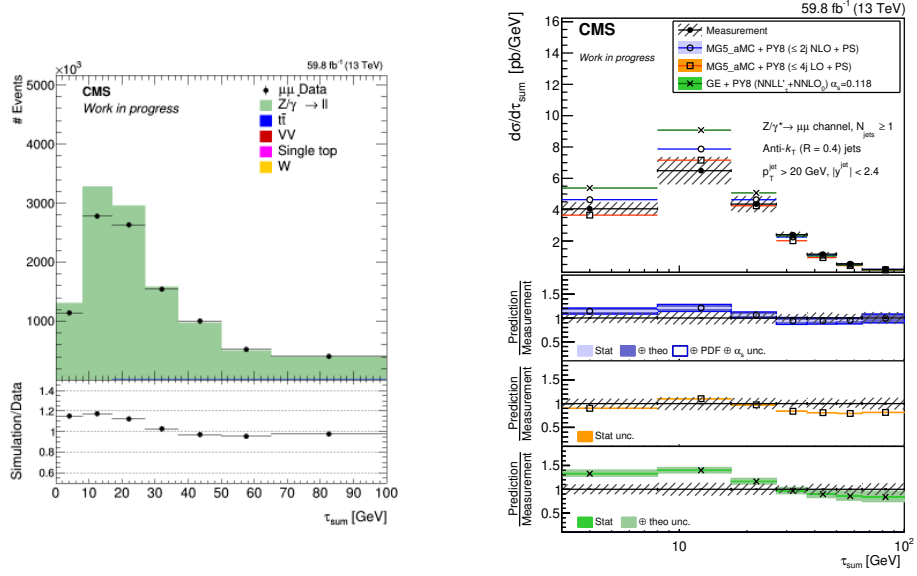


Figure 5.7: Data to simulation comparison and differential cross section as a function of  $\tau_{sum}$ .

# Conclusion

I have presented in this thesis the work I have performed during my PhD thesis. It has been carried out within the CMS Collaboration. The main focus has been on the measurement of the N-jettiness variables in the production of a Z boson in association with jets presented in Chapter 4 and the measurement of jet-based event shape variables presented in Chapter 5. A part of the thesis is devoted to intercalibration and performance studies of the CMS electromagnetic calorimeter, which is described in Chapter 3.

The excellent resolution of the CMS electromagnetic calorimeter plays an important role in many physics analyses performed at CMS. In particular, its precise measurement of electron and photon energy had an essential contribution to the discovery of the Higgs boson through the  $H \rightarrow \gamma\gamma$  channel during the LHC Run 1 period. The Run 2 data-taking period, with the increased pileup and radiation level, created a challenging environment for ECAL. In order to optimize the performance, a calibration of the relative response of the ECAL channels and corrections for the response variation in time are performed. My work included delivering the intercalibration constants using the  $Z \rightarrow e^+e^-$  method and studying the performance of ECAL. The intercalibration constants were derived for all three years of the Run 2 data-taking period. The constant monitoring and calibration resulted in excellent performance of ECAL during Run 2. The energy resolution for electrons from Z boson decays is at the level of 1.7% in the low pseudorapidity region. It was also shown that the performance with Run 2 data is very close to the one from Run 1, despite the ageing of the detector and much higher instantaneous luminosity provided by the LHC.

The second part of my work was dedicated to the measurement of the differential cross section of Z boson production in association with jets in proton-proton collisions at 13 TeV. The data recorded by the CMS detector during 2018, corresponding to an integrated luminosity of 59.4 fb<sup>1</sup>, has been analyzed. The measurements of the production process of the Z boson in association with jets are crucial for understanding and modeling the QCD interactions. In addition, this process is an important background for many Standard Model processes and for predicted processes beyond the SM. Therefore, a precise knowledge of the Z

boson production in association with jets has great importance for exploiting the potential of the LHC experiments.

Differential cross sections have been measured as a function of track-based event shape variables and as a function of jet-based event shape variables. The measurements have been compared with three types of theoretical predictions with LO, NLO, and NNLO QCD accuracies obtained with two generators, MADGRAPH5\_aMC@NLO and GENEVA.

The track-based variables that have been measured are the zero-jettiness, one-jettiness, and the sum of the transverse momentum of charged particles. These variables can be used as a veto for hard radiation or jets and to define a theoretically well-controlled exclusive N-jet cross section. Track-based variables are very sensitive to the underlying events and soft radiation; therefore studies of these variables give valuable input for event generator developments. The measurements are performed in the events with pairs of muons produced in the decay of on-shell Z bosons with an invariant mass between 76 and 106 GeV, and for off-shell Z bosons with an invariant mass between 125 and 150 GeV, 150 and 350 GeV and 350 and 1500 GeV. Track-based variables are also measured in four different Z boson transverse momentum regions. The measurements done for the on-shell Z boson show that the low zero-jettiness region in the inclusive case is best described by the GENEVA prediction. In the higher Z boson transverse momentum region, where it is expected to have one or more jets accompanying the Z boson, among the predictions MADGRAPH5\_aMC@NLO performs best. In the higher invariant mass regions, all predictions show a fair agreement with data. Measurements of these variables show a good potential for studies of the underlying events. By studying track-based variables for invariant masses above the Z peak, a regime similar to the one of the Higgs boson has been explored.

Jet-based variables that are measured are  $\tau_{\text{max}}$  and  $\tau_{\text{sum}}$ . These variables are defined using the jet transverse momentum weighted by a rapidity dependent function. Jet-based variables introduce a possibility to apply a tight veto on central jets while at forward rapidities the veto constrain gets looser. The  $\tau_{\text{max}}$  variable showed a good agreement with the predictions, especially with MADGRAPH5\_aMC@NLO. This variable can be used as a jet veto.

The measurements performed show good potential for testing resummation. Using N-jettiness variables as a jet veto in order to have better control on theory uncertainty, would require understanding better the predictions including resummation of zero-jettiness such as that obtained with GENEVA.

# Synthèse

Dans cette thèse est présentée la mesure de sections efficaces différentielles de la production d'un boson Z associé à des jets dans une collision proton-proton à une énergie de centre de masse de 13 TeV. Les données analysées ont été collectées sur l'expérience CMS (Compact Muon Solenoid) auprès du LHC (Large Hadron Collider) durant l'année 2018. Elles correspondent à une luminosité intégrée de  $59 \text{ fb}^{-1}$ . La section efficace est mesurée en fonction de variables de forme d'événement définies à partir de la cinématique des particules: zéro-jettiness, un-jettiness et somme des moments transversaux des particules, ainsi que des variables de forme d'événement s'appuyant sur les jets ( $\tau_{\text{sum}}$  et  $\tau_{\text{max}}$ ).

Le LHC est un accélérateur circulaire conçu pour faire entrer en collision des protons ou des ions lourds. Il s'agit du plus grand et plus puissant accélérateur jamais construit. Sa circonférence est de 27 km. Il est situé à la frontière entre la France et la Suisse, près de Genève, sous-terre, à une profondeur comprise entre 45 et 175 m. Ce collisionneur a été installé dans le tunnel circulaire initialement construit pour le Grand collisionneur de protons et d'électrons (LEP) qui a fonctionné jusqu'en 2000 et a joué un rôle essentiel dans l'étude des propriétés des bosons Z et W. Le projet LHC a été proposé en 1994 par l'Organisation européenne pour la recherche nucléaire (CERN). Les principales expériences sont installées aux quatre points de collision du LHC : ATLAS (Toroidal LHC ApparatuS), CMS, LHCb (LHC beauty) et ALICE (A Large Ion Collider Experiment). L'un des principaux objectifs du LHC était la recherche du boson de Higgs, découvert par les expériences CMS et ATLAS en 2012. Depuis cette découverte, des mesures de précision pour étudier les propriétés du boson de Higgs sont menées sur ce même accélérateur. En outre, grâce aux énergies atteintes par les collisions, sont effectuées des recherches de nouvelle physique au-delà du modèle standard.

L'expérience CMS repose sur un détecteur polyvalent situé à Cessy en France, à environ 100 m sous terre. C'est un détecteur cylindrique de 22 m de long et 15 m de large qui se compose de plusieurs systèmes de sous-détecteurs, chacun ayant un rôle spécifique dans la détection des particules et la mesure de leur quantité de mouvement au point d'interaction. L'élément central de CMS est un solénoïde supraconducteur de 12.5 m de long qui fournit un champ magnétique de 3.8 T.

Le trajectographe, les calorimètres électromagnétique et hadronique sont placés à l'intérieur du solénoïde. Des chambres à muons sont placées dans la culasse de retour en acier de l'aimant.

L'étude de la production d'un boson Z en association avec des jets est essentielle pour la modélisation et la compréhension des interactions QCD. La comparaison entre les mesures et les simulations permet une meilleure compréhension de la précision des prédictions. Elle est également utile pour permettre d'améliorer les calculs de sections efficaces et les techniques de générateur Monte-Carlo d'événements. La production d'un boson Z en association avec des jets est également un fond majeur pour de nombreux processus au sein du modèle standard et au-delà. Par conséquent, la mesure de ce processus avec la plus grande précision possible offre une large application. Plusieurs mesures et recherches effectuées au LHC nécessitent le rejet d'événements de fond en fonction de leur activité hadronique. En particulier, dans le cas d'un signal produit par la fusion de bosons vecteurs, la faible activité hadronique dans la région centrale est généralement exploitée pour sélectionner les événements du signal. Le veto implique des corrections provenant de contributions QCD d'ordres supérieurs, dont les incertitudes sont estimées en variant les échelles d'énergie. Les vetos de l'activité hadronique reposent typiquement sur les jets, ce qui entraîne des restrictions compliquées de l'espace de phase, et une dépendance au terme obtenu par génération de gerbes de partons avec une précision à l'ordre des logarithmes de  $\alpha_s$  dominant. Afin de réduire les incertitudes sur les mesures des sections efficaces, une nouvelle méthode basée sur l'utilisation des variables de forme des événements a été proposée. La formule de factorisation permet d'effectuer la sommation des logarithmes à l'ordre NNLL (next-to-next-to-leading log ou log sous-sous-dominant).

Les mesures ont été comparées à trois types de prédictions théoriques avec des précisions QCD d'ordre dominant (LO), d'ordre sous-dominant (NLO) et sous-sous-dominant (NNLO) obtenues avec deux générateurs, MADGRAPH5\_aMC@NLO et GENEVA. Le premier

MADGRAPH5\_aMC@NLO est un échantillon combiné (*merged samples*) FxFx qui inclut zéro, un et deux jets à la précision NLO, tandis que le deuxième, obtenu avec MADGRAPH5\_aMC@NLO et PYTHIA8 au niveau Born utilise le schéma de correspondance (*matching*)  $k_T$ -MLM et inclut jusqu'à quatre jets dans les éléments de matrice. La troisième prédiction, obtenue avec GENEVA, est effectuée à l'ordre NNLO avec une résommation d'ordre supérieur de la zéro-jettiness (zéro-jettiness originale incluant toutes les particules) à une précision logarithmique, dénotée NNLL', proche de l'ordre NNLL.

Les variables de forme de l'événement utilisant les traces des particules qui ont été mesurées sont la zéro-jettiness, la un-jettiness et la somme des moments transversaux des particules chargées. Ces variables peuvent être utilisées comme

veto d'activité hadronique, ou de jets, et pour définir une section efficace exclusive en nombre de jets bien contrôlée sur le plan théorique. Les variables utilisant les traces sont très sensibles aux événements sous-jacents et au rayonnement doux ; par conséquent, l'étude de ces variables fournit des informations précieuses pour le développement de générateurs d'événements. Les mesures sont effectuées dans les événements avec des paires de muons produites dans la désintégration de bosons  $Z$  sur la couche de masse avec une masse invariante comprise entre 76 et 106 GeV, et pour les bosons  $Z$  hors couche de masse avec une masse invariante entre 125 et 150 GeV, 150 et 350 GeV et 350 et 1500 GeV. Les variables reposant sur les traces sont également mesurées des valeurs du moment du boson  $Z$  compris dans quatre intervalles différents. Les mesures effectuées pour le boson  $Z$  sur la couche de masse montrent que la région de faible jettiness zero dans le cas inclusif est mieux décrite par la prédiction GENEVA. Lorsque le moment transversal du boson  $Z$  est restreint à une valeur inférieure à 6 GeV ou comprise entre 6 et 12 GeV, le niveau d'accord est similaire pour les prédictions avec et sans resommation. Dans l'intervalle de moment transversal du boson  $Z$  inférieur, on s'attend à une faible activité du jet provenant de l'interaction principale et à être particulièrement sensible aux interactions sous-jacentes de l'événement. Dans l'intervalle du moment transversal du boson  $Z$  le plus élevé, où l'on s'attend à ce qu'un ou plusieurs jets accompagnent le boson  $Z$ , MADGRAPH5\_aMC@NLO est le plus performant parmi les prédictions. Dans les régions de masse invariante supérieure, toutes les prédictions montrent un bon accord avec les données.

Pour la mesure de la un-jettiness autour du pic  $Z$ , la forme des trois prédictions est similaire et les différences portent principalement sur les valeurs de section efficace totale qui sont calculées à différents ordres. Dans la région où le moment transversal du boson  $Z$  est supérieur à 25 GeV, la prédiction qui inclut le nombre maximal de jets dans les éléments de matrice est la plus proche de la mesure. La prédiction GENEVA fournit une bonne description de la mesure pour toutes les régions de masse au-dessus du pic de masse.

Pour un espace de phase inclusif des valeurs du moment transversal du boson  $Z$ , la somme du moment transversal est assez bien décrite par les prédictions GENEVA et MADGRAPH5\_aMC@NLO indépendamment de la masse du dilepton, mais avec une distribution plus plate autour de 10 GeV pour MADGRAPH5\_aMC@NLO  $FxFx$  pour toutes les intervalles en masse et pour  $FxFx$  et  $k_T$  MLM pour les régions hors couche de masse. En outre, les prédictions MADGRAPH5\_aMC@NLO fournissent une bonne description lorsque l'espace de phase est restreint au moment transversal du boson  $Z$  supérieur à 25 GeV avec le boson dans le pic de masse.

Les mesures de ces variables présentent un bon potentiel pour l'étude des événements sous-jacents. En étudiant les variables reposant sur les traces pour les masses invariantes au-dessus du pic  $Z$ , un régime similaire à celui du boson de Higgs a été

exploré.

Les variables construites à partir des jets qui sont mesurées sont  $\tau_{\max}$  et  $\tau_{\text{sum}}$ . Elle sont définies en utilisant les moments transversaux des jets pondérés par une fonction dépendant de la rapidité. Les variables reposant sur les jets introduisent la possibilité d'appliquer un veto strict sur les jets centraux, tandis qu'à des rapidités plus élevées, la contrainte du veto se relâche. Tous les échantillons générés prédisent un pic plus proéminent de la distribution de  $\tau_{\text{sum}}$ , avec un effet plus prononcé pour GENEVA. La variable  $\tau_{\max}$  a montré un bon accord avec les prédictions, en particulier avec MADGRAPH5\_aMC@NLO. Cette variable peut être utilisée comme un veto de jet.

Une partie de la thèse est consacrée aux études des performances et de l'étalonnage entre canaux, appelé interétalonnage, du calorimètre électromagnétique (ECAL) pendant la période de prise de données Run 2 (2016, 2017 et 2018).

ECAL est un calorimètre hermétique, homogène et à haute granularité, dédié à la mesure de l'énergie des électrons et des photons. Il est constitué de cristaux de tungstate de plomb ( $\text{PbWO}_4$ ). Le matériau choisi est pratique en raison de son émission rapide de lumière et de sa résistance à l'irradiation. En raison de son temps de scintillation court, environ 80% de la lumière peut être collectée en 25 ns. En outre, grâce aux propriétés telles que la densité élevée (8,8 g/cm<sup>3</sup>), le petit rayon de Molière (2.2 cm) et la courte longueur de rayonnement (0.89 cm) du  $\text{PbWO}_4$ , il a été possible de construire un calorimètre très compact à granularité élevée. ECAL est composée d'une région centrale, en forme de tonneau, qui comprend 61200 cristaux et couvre la région de pseudo-rapacité  $|\eta| < 1.48$ , et de deux bouchons avec 14648 cristaux couvrant la plage  $1.48 < |\eta| < 3$ .

La haute résolution d'ECAL joue un rôle important dans de nombreuses analyses de physique réalisées à CMS. En particulier, sa mesure précise de l'énergie des électrons et des photons a été essentielle à la découverte du boson de Higgs par le canal  $H \rightarrow \gamma\gamma$  pendant la période du Run 1 du LHC. La période de prise de données du Run2, avec l'augmentation de l'empilement d'événements et du niveau de radiation, a créé un environnement difficile pour le détecteur. Afin d'optimiser les performances, un étalonnage de la réponse relative de chaque canal et des corrections pour la variation de la réponse dans le temps sont effectués. J'ai fourni à la collaboration les constantes d'interétalonnage en utilisant la méthode  $Z \rightarrow e^+e^-$  et j'ai étudié les performances d'ECAL. Les constantes d'interétalonnage ont été dérivées pour les trois années de la période de prise de données du Run 2. La surveillance et l'étalonnage constants ont permis d'obtenir d'excellentes performances de l'ECAL au cours du Run 2. La résolution en énergie des électrons issus de la désintégration du boson Z est de l'ordre de 1.7% dans la région de faible pseudo-rapacité. Les études que j'ai effectuée ont montré que les performances

avec les données du Run 2 sont très proches de celles du Run 1, malgré le vieillissement du détecteur et une luminosité instantanée fournie par le LHC beaucoup plus élevée.

# Bibliography

- [1] TeVI Group, “Design Report Tevatron 1 project,” 1984. FERMILAB-DESIGN-1984-01.
- [2] CERN, “LEP design report: Vol.2. The LEP Main Ring,” 1984. CERN-LEP-84-01.
- [3] G. E. Wolf, “HERA: Physics, Machine and Experiments,” *NATO Science Series B*, vol. 164, pp. 375–449, 1987. doi:10.1007/978-1-4684-5401-7\_6.
- [4] L. Evans and P. Bryant, “LHC machine,” *Journal of instrumentation*, vol. 3, no. 08, p. S08001, 2008. doi:10.1088/1748-0221/3/08/S08001.
- [5] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC,” *Physics Letters B*, vol. 716, pp. 30–61, 2012. doi:10.1016/j.physletb.2012.08.021.
- [6] ATLAS Collaboration, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC,” *Physics Letters B*, vol. 716, pp. 1–29, 2012. doi:10.1016/j.physletb.2012.08.020.
- [7] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, “N-jettiness: an inclusive event shape to veto jets,” *Physical review letters*, vol. 105, no. 9, p. 092002, 2010. doi:10.1103/PhysRevLett.105.092002.
- [8] S. Gangal, M. Stahlhofen, and F. J. Tackmann, “Rapidity-dependent jet vetoes,” *Physical Review D*, vol. 91, no. 5, 2015. doi:10.1103/physrevd.91.054023.
- [9] CMS Collaboration, “The CMS experiment at the CERN LHC,” *Journal of instrumentation*, vol. 3, p. S08004, 2008. doi:10.1088/1748-0221/3/08/S08004.

- [10] CMS Collaboration, “Overview of the CMS electromagnetic calorimeter,” *Nuclear Physics B*, vol. 78, pp. 186–191, 1999. doi:10.1016/S0920-5632(99)00544-7.
- [11] D. Galbraith, “Standard Model of the Standard Model.” Last accessed October 2021 <http://www.who.int/mediacentre/factsheets/fs282/fr>.
- [12] P. A. Dirac, “Theory of electrons and positrons,” *Nobel Lecture*, vol. 12, pp. 320–325, 1933.
- [13] S. L. Glashow, “Partial-symmetries of weak interactions,” *Nuclear physics*, vol. 22, no. 4, pp. 579–588, 1961. doi:10.1016/0029-5582(61)90469-2.
- [14] S. Weinberg, “A model of leptons,” *Physical review letters*, vol. 19, no. 21, p. 1264, 1967.
- [15] A. Salam, “Weak and Electromagnetic Interactions,” *Conf. Proc. C*, vol. 680519, pp. 367–377, 1968. doi:10.1142/9789812795915\_0034.
- [16] P. W. Higgs, “Broken symmetries and the masses of gauge bosons,” *Physical Review Letters*, vol. 13, no. 16, p. 508, 1964. doi:10.1103/PhysRevLett.13.508.
- [17] F. Englert and R. Brout, “Broken symmetry and the mass of gauge vector mesons,” *Physical Review Letters*, vol. 13, no. 9, p. 321, 1964. doi:10.1103/PhysRevLett.13.321.
- [18] J. Ellis, “Higgs Physics,” *2013 European School of High-Energy Physics*, pp. 117–168, 2015. doi:10.5170/CERN-2015-004.117.
- [19] Particle Data Group, “W boson mass.” <https://pdg.lbl.gov/2020/listings/rpp2020-list-w-boson.pdf>.
- [20] Particle Data Group, “Z boson mass.” <https://pdg.lbl.gov/2020/listings/rpp2020-list-z-boson.pdf>.
- [21] R. P. Feynman, “Very high-energy collisions of hadrons,” *Physical Review Letters*, vol. 23, pp. 1415–1417, 1969. doi:10.1103/PhysRevLett.23.1415.
- [22] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and collider physics*. Cambridge University Press, 2011. doi:10.1017/CBO9780511628788.
- [23] F. D. Aaron *et al.*, “Inclusive Deep Inelastic Scattering at High  $Q^2$  with Longitudinally Polarised Lepton Beams at HERA,” *Journal of High Energy Physics*, vol. 09, p. 061, 2012. doi:10.1007/JHEP09(2012)061.

- [24] Y. L. Dokshitzer, “Calculation of the structure functions for deep inelastic scattering and  $e^+e^-$  annihilation by perturbation theory in quantum chromodynamics,” *Zh. Eksp. Teor. Fiz.*, vol. 73, p. 1216, 1977.
- [25] V. N. Gribov and L. N. Lipatov, “Deep inelastic e p scattering in perturbation theory,” *Soviet Journal of Nuclear Physics*, vol. 15, pp. 438–450, 1972.
- [26] G. Altarelli and G. Parisi, “Asymptotic freedom in parton language,” *Nuclear Physics B*, vol. 126, no. 2, pp. 298–318, 1977. doi:10.1016/0550-3213(77)90384-4.
- [27] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr, and G. Watt, “LHAPDF6: parton density access in the LHC precision era,” *The European Physical Journal C*, vol. 75, no. 3, pp. 1–20, 2015. doi:10.1140/epjc/s10052-015-3318-8.
- [28] R. D. Ball *et al.*, “Parton distributions from high-precision collider data,” *The European Physical Journal C*, vol. 77, no. 10, pp. 1–75, 2017. doi:10.1140/epjc/s10052-017-5199-5.
- [29] S. D. Drell and T.-M. Yan, “Massive lepton-pair production in hadron-hadron collisions at high energies,” *Physical Review Letters*, vol. 25, no. 5, p. 316, 1970. doi:10.1103/PhysRevLett.25.316.
- [30] N. Metropolis and S. Ulam, “The Monte Carlo method,” *Journal of the American statistical association*, vol. 44, no. 247, pp. 335–341, 1949.
- [31] T. Gleisberg, S. Hoeche, F. Krauss, A. Schälicke, S. Schumann, and J. Winter, “SHERPA 1.  $\alpha$ , a proof-of-concept version,” *Journal of High Energy Physics*, vol. 2004, no. 02, p. 056, 2004. doi:10.1088/1126-6708/2004/02/056.
- [32] S. Agostinelli *et al.*, “GEANT4—a simulation toolkit,” *Nuclear instruments and methods in physics research section A*, vol. 506, no. 3, pp. 250–303, 2003. doi:10.1016/S0168-9002(03)01368-8.
- [33] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. 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,” *Journal of High Energy Physics*, vol. 2014, no. 7, pp. 1–157, 2014.
- [34] C. W. Bauer, F. J. Tackmann, and J. Thaler, “GenEvA (I): A new framework for event generation,” *Journal of High Energy Physics*, vol. 2008, no. 12, p. 010, 2008. doi:10.1088/1126-6708/2008/12/010.

- [35] C. W. Bauer, F. J. Tackmann, and J. Thaler, “GenEvA (II): A phase space generator from a reweighted parton shower,” *Journal of High Energy Physics*, vol. 2008, no. 12, p. 011, 2008. doi:10.1088/1126-6708/2008/12/011.
- [36] R. Frederix and S. Frixione, “Merging meets matching in MC@NLO,” *Journal of High Energy Physics*, vol. 2012, no. 12, p. 061, 2012. doi:10.1007/JHEP12(2012)061.
- [37] J. Alwall *et al.*, “Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions,” *The European Physical Journal C*, vol. 53, no. 3, p. 473–500, 2007. doi:10.1140/epjc/s10052-007-0490-5.
- [38] CMS Collaboration, “Extraction and validation of a new set of CMS PYTHIA 8 tunes from underlying-event measurements,” *The European Physical Journal C*, vol. 80, no. 1, p. 4, 2020. doi:10.1140/epjc/s10052-019-7499-4.
- [39] CMS Collaboration, “Event generator tunes obtained from underlying event and multiparton scattering measurements,” *The European Physical Journal C*, vol. 76, no. 3, pp. 1–52, 2016. doi:10.1140/epjc/s10052-016-3988-x.
- [40] R. E. Allen, “The Higgs Bridge,” *Physica Scripta*, vol. 89, no. 1, p. 018001, 2013. doi:10.1088/0031-8949/89/01/018001.
- [41] E. Mobs, “The CERN accelerator complex-August 2018,” 2018. <http://cds.cern.ch/record/2636343>.
- [42] ATLAS Collaboration, “The ATLAS experiment at the CERN Large Hadron Collider,” *Journal of instrumentation*, vol. 3, p. S08003, 2008. doi:10.1088/1748-0221/3/08/S08003.
- [43] LHCb Collaboration, “The LHCb detector at the LHC,” *Journal of instrumentation*, vol. 3, no. 08, p. S08005, 2008. doi:10.1088/1748-0221/3/08/S08005.
- [44] ALICE Collaboration, “The ALICE experiment at the CERN LHC. A Large Ion Collider Experiment,” *Journal of instrumentation*, vol. 3, p. S08002. doi:10.1088/1748-0221/3/08/S08002.
- [45] CMS Collaboration, “Public luminosity results.” <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>.
- [46] HiLumi Collaboration, “The HL-LHC project.” <https://hilumilhc.web.cern.ch/content/hl-lhc-project>.

- [47] T. Sakuma and T. McCauley, “Detector and Event Visualization with SketchUp at the CMS Experiment,” *Journal of Physics: Conference Series*, vol. 513, p. 022032, 2014. doi:10.1088/1742-6596/513/2/022032.
- [48] CMS Collaboration, “Performance of the CMS drift-tube local trigger with cosmic rays,” *Journal of Instrumentation*, vol. 5, p. T03003, 2010. doi:10.1088/1748-0221/5/03/T03003.
- [49] CMS Collaboration, “The CMS magnet project: Technical Design Report,” 1997. CERN-LHCC-97-10.
- [50] CMS Collaboration, “Precise mapping of the magnetic field in the CMS barrel yoke using cosmic rays,” *Journal of Instrumentation*, vol. 5, no. 03, p. T03021, 2010.
- [51] CMS Collaboration, “The CMS tracker system project: Technical Design Report,” 1997. CERN-LHCC-98-006.
- [52] CMS Collaboration, “The CMS tracker: addendum to the Technical Design Report,” 2000. CERN-LHCC-2000-016.
- [53] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker,” *Journal of Instrumentation*, vol. 9, no. 10, p. P10009, 2014.
- [54] CMS Collaboration, “CMS technical design report for the pixel detector upgrade,” 2012. CERN-LHCC-2012-016, doi:10.2172/1151650.
- [55] A. Benaglia, “The CMS ECAL performance with examples,” *Journal of Instrumentation*, vol. 9, p. C02008, 2014. doi:10.1088/1748-0221/9/02/C02008.
- [56] CMS Collaboration, “Studies of the response of the prototype CMS hadron calorimeter, including magnetic field effects, to pion, electron, and muon beams,” *Nuclear Instruments and Methods in Physics Research Section A*, vol. 457, no. 1-2, pp. 75–100, 2001. doi:10.1016/S0168-9002(00)00711-7.
- [57] CMS Collaboration, “The CMS muon project: Technical Design Report,” 1997. CERN-LHCC-97-032.
- [58] CMS Collaboration, “Performance of CMS muon reconstruction in  $pp$  collision events at  $\sqrt{s} = 7$  TeV,” *Journal of Instrumentation*, vol. 7, p. P10002, 2012. doi:10.1088/1748-0221/7/10/P10002.
- [59] CMS Collaboration, “The CMS trigger system,” *Journal of Instrumentation*, vol. 12, no. 01, p. P01020, 2017. doi:10.1088/1748-0221/12/01/P01020.

- [60] CMS Collaboration, “Performance of the CMS Level-1 trigger in proton-proton collisions at  $\sqrt{s} = 13$  TeV,” *Journal of Instrumentation*, vol. 15, no. 10, p. P10017, 2020. doi:10.1088/1748-0221/15/10/P10017.
- [61] CMS Collaboration, “Particle-flow reconstruction and global event description with the CMS detector,” *Journal of Instrumentation*, vol. 12, no. 10, p. P10003, 2017.
- [62] D. Barney, “CMS Detector Slice.” CMS Collection, 2016.
- [63] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker,” *Journal of Instrumentation*, vol. 9, no. 10, pp. P10009–P10009, 2014. doi:10.1088/1748-0221/9/10/p10009.
- [64] R. Fruhwirth, “Application of Kalman filtering to track and vertex fitting,” *Nuclear Instruments and Methods in Physics Research Section A*, vol. 262, pp. 444–450, 1987. doi:10.1016/0168-9002(87)90887-4.
- [65] CMS Collaboration, “Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at  $\sqrt{s}=13$  TeV,” *Journal of Instrumentation*, vol. 13, no. 06, p. P06015–P06015, 2018. doi:10.1088/1748-0221/13/06/P06015.
- [66] CMS Collaboration, “The Phase-2 Upgrade of the CMS Endcap Calorimeter,” 2017. CERN-LHCC-2017-023.
- [67] CMS Collaboration, “The Phase-2 Upgrade of the CMS Muon Detectors,” 2017. CERN-LHCC-2017-012.
- [68] CMS Collaboration, “The Phase-2 Upgrade of the CMS L1 Trigger Interim Technical Design Report,” 2017. CERN-LHCC-2017-013.
- [69] CMS Collaboration, “Observation of the diphoton decay of the Higgs Boson and measurement of its properties,” *The European Physical Journal C*, vol. 74, no. 10, p. 3076, 2014. doi:10.1140/epjc/s10052-014-3076-z.
- [70] CMS Collaboration, “Search for narrow resonances in dilepton mass spectra in proton-proton collisions at  $\sqrt{s} = 13$  TeV and combination with 8 TeV data,” *Physics Letters B*. doi:10.1393/ncc/i2018-18012-1.
- [71] CMS Collaboration, “Measurement of the ZZ production cross section and  $Z \rightarrow l^+l^-l'^+l'^-$  branching fraction in pp collisions at  $\sqrt{s} = 13$  TeV,” *Physics Letters B*, vol. 763, pp. 280–303, 2016. doi:10.1016/j.physletb.2016.10.054.

- [72] J. Mijuskovic, “The CMS electromagnetic calorimeter upgrade: high-rate readout with precise time and energy resolution,” *Journal of Instrumentation*, vol. 17, no. 01, p. C01004, 2022. doi:10.1088/1748-0221/17/01/c01004.
- [73] CMS Collaboration, “Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV,” *Journal of Instrumentation*, vol. 10, no. 06, p. P06005, 2015. doi:10.1088/1748-0221/10/06/P06005.
- [74] CMS Collaboration, “Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV,” *Journal of Instrumentation*, vol. 10, no. 08, pp. P08010–P08010, 2015. doi:10.1088/1748-0221/10/08/p08010.
- [75] CMS Collaboration, “Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC,” *Journal of Instrumentation*, vol. 16, no. 05, p. P05014, 2021. doi:10.1088/1748-0221/16/05/P05014.
- [76] CMS Collaboration, “Reconstruction of signal amplitudes in the CMS electromagnetic calorimeter in the presence of overlapping proton-proton interactions,” *Journal of Instrumentation*, vol. 15, no. 10, p. P10002, 2020. doi:10.1088/1748-0221/15/10/P10002.
- [77] E. Di Marco, “CMS electromagnetic calorimeter calibration and timing performance during LHC Run I and future prospects,” 2014. CMS-CR-2014-410.
- [78] CMS Collaboration, “Laser monitoring system for the CMS lead tungstate crystal calorimeter,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 594, no. 2, pp. 292–320, 2008. doi:10.1016/j.nima.2008.01.104.
- [79] A. Zghiche, “Performance of the CMS electromagnetic calorimeter during the LHC Run II,” *Journal of Physics: Conference Series*, vol. 1162, p. 012001, 2019. doi:10.1088/1742-6596/1162/1/012001.
- [80] CMS Collaboration, “Public ECAL DPG results.” <https://twiki.cern.ch/twiki/bin/view/CMSPublic/EcalDPGResultsCMSDPS2017023>.
- [81] CMS Collaboration, “Energy resolution of the barrel of the CMS electromagnetic calorimeter,” *Journal of Instrumentation*, vol. 2, no. 04, p. P04004, 2007. doi:10.1088/1748-0221/2/04/P04004.
- [82] Particle Data Group, “Review of Particle Physics,” *Chin. Phys. C*, vol. 40, no. 10, p. 100001, 2016. doi:10.1088/1674-1137/40/10/100001.

- [83] F. Couderc, *Quest for the Higgs boson(s) from D0 to CMS experiments*. Habilitation à diriger des recherches, Sorbonne Université, 2018.
- [84] T. Czosnyka and A. Trzcińska, “Unified analytical approximation of Gaussian and Voigtian lineshapes,” *Nuclear Instruments and Methods in Physics Research Section A*, vol. 431, no. 3, pp. 548–550, 1999. doi:10.1016/S0168-9002(99)00287-9.
- [85] J. J. Olivero and R. L. Longbothum, “Empirical fits to the Voigt line width: A brief review,” *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 17, pp. 233–236, 1977. doi:10.1016/0022-4073(77)90161-3.
- [86] CMS Collaboration, “The Phase-2 Upgrade of the CMS Barrel Calorimeters,” 2017. CERN-LHCC-2017-011.
- [87] O. Gevin, F. Guilloux, P. Baron, and M. Dejardin, “CATIA: APD readout ASIC for CMS phase 2 ECAL electronics upgrade,” *Proceeding of science*, vol. TWEPP2019, p. 001, 2020. doi:10.22323/1.370.0001.
- [88] S. Cometti and G. Mazza, “LiTE-DTU: online data selection, compression, and transmission ASIC for the upgraded front-end of the CMS Electromagnetic Calorimeter,” *8th International Conference on Modern Circuits and Systems Technologies*, pp. 1–4, 2019. doi:10.1109/MOCAST.2019.8741539.
- [89] J. M. Mendez, S. Baron, S. Kulis, and J. Fonseca, “New LpGBT-FPGA IP: Simulation model and first implementation,” *Proceeding of science*, vol. TWEPP2018, p. 059, 2019. doi:10.22323/1.343.0059.
- [90] C. Soós, S. Détraz, L. Olanterä, C. Sigaud, J. Troska, F. Vasey, and M. Zeiler, “Versatile Link PLUS transceiver development,” *Journal of Instrumentation*, vol. 12, no. 03, p. C03068, 2017. doi:10.1088/1748-0221/12/03/C03068.
- [91] CMS Collaboration, “The Barrel Calorimeter Processor demonstrator board for the Phase II Upgrade of the CMS ECAL Barrel,” 2018. <https://cds.cern.ch/record/2644903>.
- [92] CMS Collaboration, “Technical proposal for the Phase-II Upgrade of the CMS Detector,” 2015. CERN-LHCC-2015-010.
- [93] I. W. Stewart, F. J. Tackmann, J. Thaler, C. K. Vermilion, and T. F. Wilkerson, “Xcone: N-jettiness as an exclusive cone jet algorithm,” *Journal of High Energy Physics*, vol. 2015, no. 11, 2015. doi:10.1007/jhep11(2015)072.
- [94] CMS Collaboration, “CMS luminosity measurement for the 2018 data-taking period at  $\sqrt{s}=13$  TeV,” 2019. CMS-PAS-LUM-18-002.

- [95] S. Alioli, P. Nason, C. Oleari, and E. Re, “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX,” *Journal of High Energy Physics*, vol. 2010, no. 6, p. 043, 2010. doi:10.1007/jhep06(2010)043.
- [96] M. Czakon and A. Mitov, “Top++: A program for the calculation of the top-pair cross-section at hadron colliders,” *Computer Physics Communications*, vol. 185, no. 11, p. 2930–2938, 2014. doi:10.1016/j.cpc.2014.06.021.
- [97] A. Giammanco, “Single top quark production at the LHC,” *Reviews in Physics*, vol. 1, pp. 1–12, 2016. doi:10.1016/j.revip.2015.12.001.
- [98] CMS Collaboration, “Measurement of the inelastic proton-proton cross section at  $\sqrt{s} = 13$  TeV,” *Journal of High Energy Physics*, vol. 2018, no. 7, 2018. doi:10.1007/jhep07(2018)161.
- [99] A. Bodek, A. Van Dyne, J. Y. Han, W. Sakumoto, and A. Strelnikov, “Extracting muon momentum scale corrections for hadron collider experiments,” *The European Physical Journal C*, vol. 72, no. 10, pp. 1–7, 2012. doi:10.1140/epjc/s10052-012-2194-8.
- [100] CMS Collaboration, “Performance of CMS muon reconstruction in pp collision events at  $\sqrt{s} = 7$  TeV,” *Journal of Instrumentation*, vol. 7, no. 10, p. P10002–P10002, 2012. doi:10.1088/1748-0221/7/10/p10002.
- [101] S. Schmitt, “TUnfold, an algorithm for correcting migration effects in high energy physics,” *Journal of Instrumentation*, vol. 7, no. 10, p. T10003–T10003, 2012. doi:10.1088/1748-0221/7/10/t10003.
- [102] A. I. Tihonov, “Solution of incorrectly formulated problems and the regularization method,” *Soviet Math.*, vol. 4, pp. 1035–1038, 1963.
- [103] CMS Collaboration, “Tracking POG results for pion efficiency with the  $D^*$  meson using data from 2016 and 2017,” 2018. CMS-DP-2018-050.
- [104] CMS Collaboration, “CMS Tracking POG performance plots for 2018 dataset.” <https://twiki.cern.ch/twiki/bin/view/CMSPublic/TrackingPOGResults2018>.
- [105] S. Alioli, C. W. Bauer, C. Berggren, F. J. Tackmann, and J. R. Walsh, “Drell-Yan production at NNLL+NNLO matched to parton showers,” *Physical Review D*, vol. 92, no. 9, 2015. doi:10.1103/physrevd.92.094020.

- [106] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- $k_T$  jet clustering algorithm,” *Journal of High Energy Physics*, vol. 2008, no. 04, p. 063–063, 2008. doi:10.1088/1126-6708/2008/04/063.
- [107] D. Bertolini, P. Harris, M. Low, and N. Tran, “Pileup per particle identification,” *Journal of High Energy Physics*, vol. 2014, no. 10, pp. 1–22, 2014. doi:10.1007/JHEP10(2014)059.
- [108] CMS Collaboration, “Pileup mitigation at CMS in 13 TeV data,” *Journal of Instrumentation*, vol. 15, no. 09, pp. P09018–P09018, 2020. doi:10.1088/1748-0221/15/09/p09018.
- [109] CMS Collaboration, “Measurement of differential cross sections for Z bosons produced in association with charm jets in pp collisions at  $\sqrt{s} = 13$  TeV,” *Journal of High Energy Physics*, vol. 2021, no. 4, 2021. doi:10.1007/jhep04(2021)109.