

The measurement of  $W^\pm$  boson charge asymmetry at  $\sqrt{s} = 13$  TeV with the CMS detector using the 2015 data set and a global QCD analysis on this asymmetry.

## Dissertation

zur Erlangung des Doktorgrades  
des Department Physik  
der Universität Hamburg

vorgelegt von

VLADYSLAV DANILOV

Hamburg  
2020

---

Gutachter/in der Dissertation:	Prof. Dr. Elisabetta Gallo Dr. Katarzyna Wichman
Gutachter/in der Disputation:	Prof. Dr. Elisabetta Gallo Prof. Dr. Johannes Haller Dr. Hannes Jung Prof. Dr. Berndt Kniehl Dr. Katarzyna Wichman
Vorsitzender des Prüfungsausschusses:	Prof. Dr. Berndt Kniehl
Datum der Disputation:	29.05.2020
Vorsitzender des Promotionsausschusses:	Prof. Dr. Günter Hans Walter Sigl
Dekan des Fachbereichs Physik:	Prof. Dr. Wolfgang Hansen
Dekan der Fakultät MIN:	Prof. Dr. Heinrich Graener

---

# Abstract

This thesis presents physics results accomplished during the 2017-2019 years in the CMS experiment. Three main contributions are discussed: a measurement of diamond sensors for the BCM1F luminometer and monitoring of its performance during the whole Run 2; a measurement of differential cross-sections of  $W^\pm$  boson production in the muon channel as a function of pseudorapidity at  $\sqrt{s} = 13$  TeV using 2015 data set, and extraction of  $W^\pm$  boson charge asymmetry values; a global QCD analysis with the obtained asymmetry results.

During 2016-2017 in DESY-Zeuthen laboratories, twelve poly-crystalline Chemical Vapour Deposited (pCVD), and five single-crystalline Chemical Vapour Deposited (sCVD) diamonds were tested to select the most suitable sensors for the BCM1F upgrade, scheduled during a short technical stop at the end of 2017. In these studies, each sensor was tested for the leakage current durability, current over time stability, and charge collection distance constancy. During Run 2, the stability of diamond sensors performance was monitored and analyzed. These results allowed to broaden our knowledge about technical aspects of luminosity measurement with diamond sensors in severe conditions of radiation damage and were considered in the next design generation of luminometers.

The extraction of  $W^\pm$  boson charge asymmetry was performed starting from the analysis of the "Measurement of inclusive  $W^\pm$  and  $Z^0$  boson production cross-sections in pp collisions and luminosity calibration at  $\sqrt{s} = 13$  TeV" using a data sample collected with the CMS detector in 2015 with a corresponding integrated luminosity of  $2.2 \pm 0.05 \text{ fb}^{-1}$ . The asymmetry was calculated by measuring differential cross-sections of  $W^\pm$  boson production in the muon channel as a function of pseudorapidity. The obtained results of differential cross-sections and asymmetry values were compared with theoretical predictions at NNLO, produced using different PDF sets. The final results are presented with the full set of systematic uncertainties, showing good agreement with theoretical predictions, within the uncertainty range.

The final part is dedicated to two global QCD analyses made using the measured asymmetry. Studies were performed using a global QCD fit approach, implemented in the xFitter framework. In the first analysis, the measured asymmetry was tested on its sensitivity to proton structure using combined DIS results from the H1 and ZEUS experiments. The inclusion of the asymmetry values showed a good improvement of valence quark distributions in a range of  $10^{-3} \leq x \leq 10^{-1}$ . The second analysis was performed with DIS data and CMS results of previous measurements of  $W^\pm$  boson charge asymmetry and  $W^\pm$  boson + charm quark production. The impact of the  $W^\pm$  boson charge asymmetry, measured in this thesis, is presented with the full set of experimental, model, and parameterization systematic uncertainties. The final results showed an improvement in the distribution of the up valence quark distribution.

# Zusammenfassung

In dieser Arbeit werden Ergebnisse des CMS Experiments aus den Jahren 2017-2019 präsentiert. Es werden drei Beiträge vorgestellt: Messungen, die durch Diamant-Sensoren innerhalb des BCM1F Luminometers aufgezeichnet wurden und die Überwachung dieser Komponenten während des gesamten Run 2; die Bestimmung des differentiellen Wirkungsquerschnitts von  $W^\pm$  Bosonen im Myon-Zerfallskanal, als Funktion der Pseudorapidität, bei einer Schwerpunktsenergie von  $\sqrt{s} = 13$  TeV, sowie die Bestimmung der Ladungsasymmetrie-Werte von  $W^\pm$ ; die anschließende globale QCD Analyse, in der die zuvor gemessenen  $W^\pm$  Ergebnisse verwendet werden.

In den Jahren 2016 bis 2017 wurden in den DESY-Laboratorien in Zeuthen zwölf poly-crystalline Chemical Vapour Deposited (pCVD) und fünf single-crystalline Chemical Vapour Deposited (sCVD) Diamanten getestet, um die am besten passenden Sensoren für das BCM1F Upgrade zu finden, welches während eines kurzen Shut-Downs 2017 stattfinden sollte. In diesen Studien wurde jeder Sensor einzeln auf seine Leakage Current Durability, Current Over Time Stability und Charge Collection Distance Constancy getestet. Diese Ergebnisse liefern Erkenntnisse über die technischen Aspekte von Luminositätsmessungen mit Diamant-Sensoren unter extremen Bedingungen, wie hoher Strahlenbelastung, und werden einen Einfluss auf das Design zukünftiger Generationen von Luminometern haben.

Die Bestimmung der Ladungsasymmetrie von  $W^\pm$ -Bosonen wurde als Beitrag zur Analyse “Measurement of inclusive  $W^\pm$  and  $Z^0$  boson production cross-sections in pp collisions and luminosity calibration at  $\sqrt{s} = 13$  TeV” durchgeführt, welche von einer Gruppe am MIT geleitet wurde. Für die Messung wurden Daten verwendet, welche durch den CMS-Detektor im Jahr 2015 aufgezeichnet wurden und einer integrierten Luminosität von  $2.2 \pm 0.05 \text{ fb}^{-1}$  entsprechen. Die Asymmetrie wurde durch die Messung des differentiellen Wirkungsquerschnitts von  $W^\pm$ -Bosonen im Myon-Zerfallskanal, als eine Funktion der Pseudorapidität, berechnet. Die Ergebnisse dieser Messungen wurden mit theoretischen Vorhersagen in Next-to-Next-to-Leading Order (NNLO) verglichen, bei denen verschiedene Sets an Parton-Verteilungsfunktionen, eng. parton distribution functions (PDF) verwendet werden. Die Ergebnisse werden zusammen mit einem vollständigen Satz an systematischen Unsicherheiten präsentiert, und zeigen eine gute Übereinstimmung mit den theoretischen Vorhersagen innerhalb der Unsicherheiten.

Der letzte Teil der Arbeit ist den beiden globalen QCD Analysen gewidmet, in der die zuvor gemessene Asymmetrie verwendet werden. Für die Studien wurde ein globaler QCD-Fit Ansatz verwendet, wie er innerhalb des xFitter Frameworks implementiert ist. In der ersten Analyse wurde der Grad der Sensitivität der gemessenen Asymmetrie auf die Protonstruktur, durch einen Vergleich mit den kombinierten DIS-Ergebnissen der H1 und Zeus Experimente, getestet. Durch das Einbeziehen der Asymmetrie-Werte in die QCD Analyse kann eine Verbesserung bezüglich der Unsicherheiten in der Valenzquark Verteilung im Bereich  $10^{-3} \leq x \leq 10^{-1}$  beobachtet werden. In der zweiten Analyse wurden die DIS-Daten, sowie die Ergebnisse früherer Messungen der  $W^\pm$  Ladungsasymmetrie und  $W^\pm$ +charm Produktion verwendet. Der Einfluss der  $W^\pm$  Ladungsasymmetrie, welche in dieser Arbeit gemessen wurde, wird mit allen Unsicherheiten bezüglich des experimentellen Aufbaus, des gewählten Modells und der verwendeten Parametrisierung präsentiert. Die Ergebnisse zeigen, dass die Einbeziehung der Messung zu einer Verbesserung in der Up-Quark Verteilung führt.

# Contents

<b>Introduction</b>	<b>1</b>
<b>1 Theoretical overview</b>	<b>3</b>
1.1 Standard Model . . . . .	3
1.1.1 Strong interaction . . . . .	5
1.1.2 Electroweak interaction . . . . .	7
1.2 Proton structure . . . . .	10
1.2.1 Parton Distribution Functions . . . . .	10
1.2.2 Factorization theorem and PDF evolution equations . . . . .	14
1.2.3 PDF extraction using a global QCD fit . . . . .	16
1.3 $W^\pm$ boson production at LHC . . . . .	17
1.3.1 $W^\pm$ boson charge asymmetry . . . . .	19
<b>2 LHC and CMS experiment</b>	<b>23</b>
2.1 Large Hadron Collider . . . . .	23
2.2 Compact Muon Solenoid . . . . .	26
2.2.1 CMS coordinate system . . . . .	27
2.2.2 Physical processes behind detection techniques . . . . .	29
2.2.3 Tracker Detector . . . . .	29
2.2.4 Calorimeters . . . . .	30
2.2.5 Superconducting Magnet . . . . .	33
2.2.6 Muon Detectors . . . . .	34
2.2.7 Triggering and Data Acquisition . . . . .	36
2.2.8 Acceptance . . . . .	37
<b>3 Event reconstruction</b>	<b>39</b>
3.1 Particle Flow algorithm . . . . .	39
3.2 Vertex reconstruction . . . . .	40
3.3 Pile-up Per Particle Identification . . . . .	41
3.4 Reconstruction of muons . . . . .	42
3.4.1 Hit and segment reconstruction . . . . .	43
3.4.2 Muon track reconstruction . . . . .	43
3.4.3 Muon identification types . . . . .	44
3.4.4 Momentum determination . . . . .	45
3.4.5 Muon isolation . . . . .	45
3.4.6 Reconstruction, identification and isolation efficiencies . . . . .	46
3.5 Reconstruction of missing transverse energy . . . . .	48

<b>4</b>	<b>Luminosity measurement at CMS</b>	<b>51</b>
4.1	Luminosity measurement using BCM1F . . . . .	52
4.2	Fast Beam Condition Monitor Detector Upgrade during the Run 2 . . . . .	55
4.2.1	Diamond sensors measurements . . . . .	56
4.3	BCM1F status monitoring using VME ADC . . . . .	68
4.3.1	VME ADC data processing . . . . .	68
4.3.2	Baseline position . . . . .	69
4.3.3	Test pulse spectra characteristics . . . . .	70
4.3.4	Signal spectra characteristics . . . . .	71
4.4	BCM1F performance . . . . .	71
4.4.1	BCM1F monitoring upgrades . . . . .	71
4.4.2	BCM1F performance during the Run 2 . . . . .	74
4.4.3	Conclusions . . . . .	79
<b>5</b>	<b>Measurement of the <math>W^\pm</math> boson charge asymmetry in the muon channel using the 2015 data set</b>	<b>81</b>
5.1	Analysis strategy . . . . .	82
5.2	Data sample . . . . .	82
5.3	Muon selection . . . . .	83
5.4	Muon energy scale and resolution corrections . . . . .	84
5.4.1	Muon charge misidentification correction . . . . .	85
5.4.2	Estimation of muon efficiencies . . . . .	85
5.4.3	Muon tracking, identification and isolation efficiency . . . . .	87
5.4.4	Standalone muon efficiency . . . . .	88
5.4.5	Muon trigger efficiency . . . . .	88
5.4.6	Total muon efficiency . . . . .	89
5.5	Missing transverse energy . . . . .	90
5.6	$W^\pm$ boson signal extraction . . . . .	92
5.6.1	Electroweak background . . . . .	92
5.6.2	Background from quantum chromodynamics processes . . . . .	94
5.6.3	Simultaneous fit . . . . .	95
5.6.4	Fit results . . . . .	96
5.7	Acceptance . . . . .	99
5.8	Systematic uncertainties . . . . .	100
5.8.1	Muon efficiencies . . . . .	101
5.8.2	Muon momentum scale and resolution correction . . . . .	101
5.8.3	Missing transverse energy . . . . .	101
5.8.4	$W^\pm$ boson signal extraction fit . . . . .	102
5.8.5	Luminosity uncertainties . . . . .	102
5.9	$W^\pm$ boson charge asymmetry measurement . . . . .	103
5.9.1	Differential cross-sections . . . . .	103
5.9.2	Asymmetry . . . . .	104
5.9.3	Total uncertainties . . . . .	105
5.10	Results . . . . .	108

<b>6</b>	<b>QCD analysis of <math>W^\pm</math> boson charge asymmetry</b>	<b>111</b>
6.1	Previous results . . . . .	111
6.2	xFitter . . . . .	112
6.3	QCD analysis . . . . .	113
6.3.1	Estimation of PDF uncertainties . . . . .	116
6.3.2	QCD analysis with the new $W^\pm$ boson charge asymmetry and HERA data. . . . .	116
6.3.3	QCD analysis with the new $W^\pm$ boson charge asymmetry and other sensitive processes. . . . .	119
<b>7</b>	<b>Summary and conclusions</b>	<b>127</b>
	<b>Acknowledgements</b>	<b>131</b>
	<b>List of Figures</b>	<b>133</b>
	<b>List of Tables</b>	<b>143</b>
	<b>Bibliography</b>	<b>145</b>
<b>A</b>	<b>BCM1F sensor measurement results</b>	<b>157</b>
A.1	Poly-crystalline Chemical Vapor Deposited (pCVD) diamonds . . . . .	157
A.1.1	pCVD 07B10412 . . . . .	157
A.1.2	pCVD 07B10415 . . . . .	158
A.1.3	pCVD 07B10417 . . . . .	159
A.1.4	pCVD 07B10421 . . . . .	160
A.1.5	pCVD 07B10422 . . . . .	161
A.1.6	pCVD 07B10423 . . . . .	162
A.1.7	pCVD 07B10430 . . . . .	163
A.1.8	pCVD 07B10431 . . . . .	164
A.1.9	pCVD 07B10433 . . . . .	165
A.1.10	pCVD 07B10434 . . . . .	166
A.1.11	pCVD 07B10435 . . . . .	167
A.1.12	pCVD 07B10436 . . . . .	168
A.2	Single-crystalline Chemical Vapor Deposited (sCVD) diamonds . . . . .	169
A.2.1	sCVD PLT S55(2-1-2) . . . . .	169
A.2.2	sCVD PLT S115(2-1-6) . . . . .	170
A.2.3	sCVD PLT S119(2-1-3) . . . . .	171
A.2.4	sCVD Batch390 2499089-1 (3-3-4) . . . . .	172
A.2.5	sCVD 2713547-5 "Number 5" . . . . .	173
<b>B</b>	<b>Measurement of the <math>W^\pm</math> charge asymmetry</b>	<b>175</b>
<b>C</b>	<b>APPLgrid upgrade</b>	<b>183</b>
C.1	The APPLgrid project . . . . .	183
C.1.1	PDF representation on a grid . . . . .	184
C.1.2	Weights representation in the case of two incoming hadrons . . . . .	184
C.1.3	A Monte Carlo for femtobarn processes (MCFM) . . . . .	186
C.1.4	Technical implementation . . . . .	186
C.2	MCFM-interface upgrade . . . . .	188

C.2.1	Processes mapping . . . . .	188
C.2.2	Steering file implementation . . . . .	190



*In memory of my parents and grandparents.*



# Introduction

In the era of hadron-hadron collisions, the factorization theorem became a beneficial concept of cross-sections calculation. Due to the hard work of many theorists, a description of the non-perturbative nature of proton structure became available with the introduction of the Parton Distribution Functions (PDFs) concept. Various fields of precision measurements, as well as searches, contain PDF uncertainty as main contribution to the total uncertainty. For that reason, various groups continuously work on the improvement of PDF precision by studying effects sensitive to proton structure.

In proton-proton collisions at the LHC the CMS collaboration has studied various processes, sensitive to proton structure. One of such processes is the  $W^\pm$  boson charge asymmetry.  $W^\pm$  bosons can be produced in  $u + \bar{d} \rightarrow W^+$  and  $d + \bar{u} \rightarrow W^-$  processes. As it is known, a proton consists of two up and one down valence quarks. Thus, the increased production of  $W^+$  compared to  $W^-$  arises from the higher up valence quark density, compared to down valence quark.

In this thesis, the  $W^\pm$  boson charge asymmetry is measured in a data sample collected at  $\sqrt{s} = 13$  TeV with the CMS detector corresponding to an integrated luminosity of up to  $2.2 \pm 0.05 \text{ fb}^{-1}$ . The analysis is performed by selecting  $W^\pm$  bosons decaying in the muon channel,  $W^\pm \rightarrow \mu^\pm \nu_\mu$ . Since neutrino cannot be directly observed in the detector, a certain amount of undetected, missing transverse energy, carried by a neutrino is always present in each event. The muon can be reconstructed in the CMS detector with high efficiency, which allows to use it as an event selection criterion. The event selection procedure is organized in a way that allows to significantly reduce the number of selected events with the same signature but different origin. Nevertheless, background events cannot be fully avoided. For that reason, the extraction of the number of produced  $W^\pm$  bosons is performed using a fit to the reconstructed missing transverse energy. The  $W^\pm$  boson charge asymmetry is calculated differentially as a function of pseudorapidity. This requires extraction of  $W^\pm$  boson yields in bins of pseudorapidity. The final results are presented as differential distributions for  $W^+$  and  $W^-$  bosons, asymmetry values, and full sets of systematic uncertainties for each measurement.

The second part of the dissertation is dedicated to the QCD analysis, an estimation of PDF improvement using the obtained asymmetry values. The analysis is performed using a global QCD fit approach in the open-source QCD framework xFitter [1]. In this approach, the PDF for each parton is defined as parameterization as a function of Bjorken  $x$ . The exact parameterization scheme is determined using a parameterization scan procedure. To estimate the sensitivity of the obtained asymmetry values to the proton structure, in the first step, the QCD analysis is performed using combined DIS results from H1 [2] and ZEUS [3] experiments, measured at HERA during phase I and II. At this stage two fits are performed, the first one is obtained using only HERA I+II [4] data, the second fit is produced using HERA data with  $W^\pm$  boson charge asymmetry values. A comparison of relative uncertainties of both PDFs allowed to estimate the sensitivity of the measured asymmetry values to the proton structure. In the second stage the QCD analysis is done using HERA data and previous CMS measurements of

$W^\pm$  boson charge asymmetry at  $\sqrt{s} = 7$  TeV and  $\sqrt{s} = 8$  TeV. Also, due to the poor sensitivity of  $W^\pm$  boson charge asymmetry to the strange quark content of proton, previous CMS measurements of  $W^+$ charm quark at  $\sqrt{s} = 7$  TeV,  $\sqrt{s} = 13$  TeV are also used. The improvement of PDFs uncertainty after including new values of  $W^\pm$  boson charge asymmetry is estimated comparing PDF results obtained with and without new asymmetry values.

The thesis is organized as follows. The first chapter is dedicated to an introduction to the theoretical background of the Standard Model, proton structure, and  $W^\pm$  boson charge asymmetry. A short description of the LHC and CMS experiments are given in the second chapter. In particular, the CMS detector is described with a special emphasis on the detection of muons and missing energy. The interpretation of detector signals in terms of reconstructed physics objects is given in Chapter 3. The main focus of this chapter is dedicated to muon and missing energy reconstruction algorithms. Chapter 4 describes the author's contribution to luminosity measurements in CMS as a part of his responsibilities as a member of the CMS collaboration. The chapter describes diamond sensor measurements that were done in preparation for the BCM1F luminometer upgrade in 2017, and BCM1F sensor monitoring during the Run 2 period. The description of differential cross-sections measurement of  $W^\pm$  boson production as well as  $W^\pm$  boson charge asymmetry extraction, and systematic uncertainties evaluation is given in Chapter 5. The QCD analysis of extracted  $W^\pm$  boson charge asymmetry values is given in Chapter 6. The outlook and conclusions are given in Chapter 7.

## Chapter 1

# Theoretical overview

This chapter presents a brief overview of the Standard Model and proton structure functions, crucial for the understanding of this thesis.

The first part of this chapter is dedicated to the Standard Model, including general ideas behind the electroweak and strong interactions, as well as the main approach in theoretical cross-section calculation. The second part introduces a modern picture of the proton structure with a special emphasis on concepts and methods that are used in this thesis. The last part brings a short explanation of the theoretical background behind direct  $W^\pm$  boson production at the LHC, and why it is important for the proton structure determination.

### 1.1 Standard Model

The Standard Model introduces a classification of all known elementary particles and describes the fundamental forces of interaction. It evolved from a Quantum Field Theory (QFT), a theoretical framework that combines classical field theory, quantum mechanics, and special relativity. The framework is built around a concept of quantized fields, propagating in space-time, in which excited states are interpreted as observed elementary particles. Dynamic properties of fields are derived using Emmy Noether Theorem, a mathematical theorem that relates symmetry with a corresponding conservation law [5]. In this sense, conserved quantities and dynamic properties of fields can be explored using corresponding symmetries through the introduction of local gauge transformations. Depending on field properties, fields can interact with similar and different fields through the mediation of gauge bosons. So far, all observed interactions are summarized in four main types: electromagnetic, weak, strong, and gravitational. While the gravitational force remains unreachable on the elementary level, the standard model describes electromagnetic, weak, and strong interactions.

All experimentally observed particles are classified according to their properties. Particles with non-integer spin are defined to be fermions and obey Fermi-Dirac statistics [6] [7], integer spin particles, bosons, obey Bose-Einstein statistics [8]. Currently, 12 fundamental fermions are known; they give rise to all observed matter. Four gauge bosons, gluon ( $g$ ), photon ( $\gamma$ ),  $W^\pm$  and  $Z^0$ , are introduced using a field theory framework based on the  $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$  symmetry gauge group.

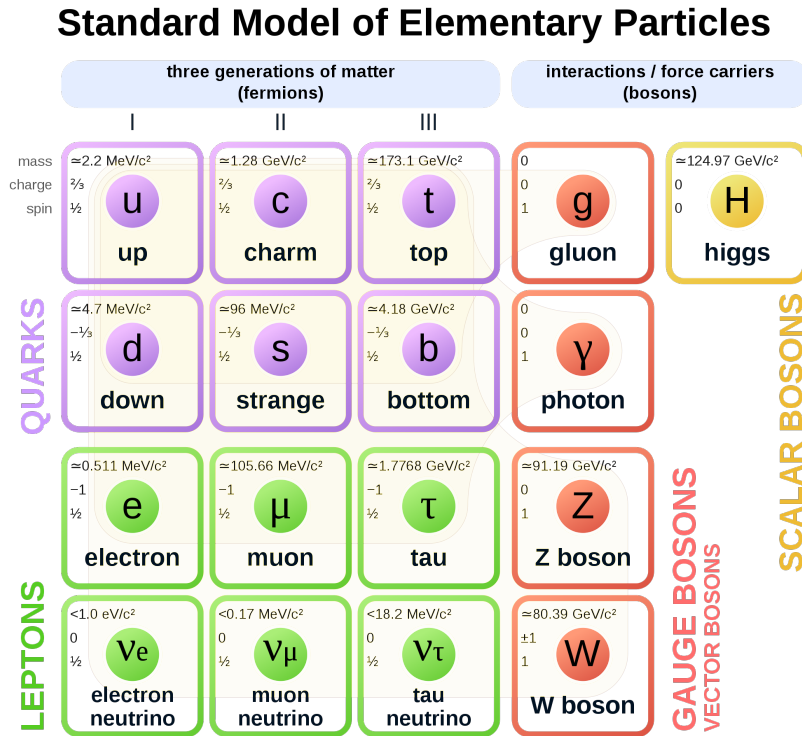


FIGURE 1.1: Elementary particles in the Standard model. The three left columns correspond to fermions. The right column is vector gauge bosons, and the last has the scalar boson of Higgs field. For each particle, mass, charge, and spin values are given in their left upper frame corner [9].

Figure 1.1 presents a table of the Standard Model particles. The upper two rows of fermions are dedicated to different types (flavors) of quarks, the only elementary fermions that can interact via strong, electromagnetic and weak forces. The branch of the Standard Model that describes strong interactions is Quantum Chromodynamics (QCD). Quarks are classified as up (up, charm, top) and down (down, strange, bottom) types; this classification appeared as an observation of flavor transformation through weak interaction. Quarks are also categorized as three generations, shown as columns in Fig. 1.1. Each generation has two different types of quarks; generations differ with their flavor quantum numbers and masses. Quarks are the only known particles carrying a fractional electric charge ( $+\frac{2}{3}$  for up type quarks, and  $-\frac{1}{3}$  for down type).

The lower two rows are dedicated to leptons, elementary particles that interact via the electromagnetic and weak forces. The first row contains electron, muon, and tau lepton; these particles have an electric charge of  $\pm 1.6 \times 10^{-19} \text{ C}$ . The second row represents neutrinos, particles without an electric charge, and relatively small masses, which interact only via the weak force. Like quarks, leptons are also categorized into generations, differing in lepton flavor numbers and masses. The theory that describes electromagnetic interactions is called Quantum Electrodynamics (QED). The last type of interaction described by the Standard Model is weak interaction; the corresponding theory is called the Theory of Weak Interaction.

The SM contains four types of interaction bosons: photon,  $\gamma$  (QED), gluon,  $g$  (QCD),  $W^\pm$  and  $Z^0$  bosons (weak interaction). The gluon can carry color charges, has

zero mass, and is being a self-interacting boson.  $W^\pm$ ,  $Z^0$ , and photon are combinations of electroweak gauge bosons. Unlike the rest of the bosons,  $W^\pm$  and  $Z^0$  have a very high mass, shown in the upper left corner. Each of the bosons has a spin of 1. More details on each of them will be given in a description of the corresponding type of interaction. The last but not least part of this table is a Higgs boson, a quantum excitation of the Higgs fields, responsible for the spontaneous symmetry breaking and mass generation [10], [11]. The observation was announced on 4th of July 2012 by CERN using data analysis results from two independent experiments ATLAS [12] and CMS [13].

The Standard Model still has unsolved questions, including the origin of dark matter, matter-antimatter asymmetry, dark energy interpretation, and others.

### 1.1.1 Strong interaction

#### Quantum Chromodynamics

QCD describes the interaction between hadrons, particles that interact via the strong force. The hadrons consist of quarks, depending on a number of quarks, two or three, hadrons are called mesons or baryons, respectively. Each quark carries a specific type of charge called "color-charge", schematically denoted as red (r), blue (b), and green (g). Hadronic wave-function states can be derived introducing a concept of isospin quantum number, with the SU(2), ud, and SU(3)<sup>1</sup>, uds, flavor symmetries. The color states are derived using a non-abelian gauge theory, with a color symmetry group SU(3)<sub>C</sub>, where C is denoted as "color".

The Lagrangian density of QCD is given by

$$\mathcal{L}_{\text{QCD}} = \sum_f \bar{\psi}_f^i (i\gamma_\mu D^\mu - m_f)_{ij} \psi_f^j - \frac{1}{4} F_a^{\mu\nu} F_{\mu\nu}^a, \quad (1.1)$$

here the f corresponds to a quark flavor,  $\psi$  is the field of a quark.  $\bar{\psi} = \psi^\dagger \gamma^0$  is the Dirac adjoint,  $\gamma_\mu$  represent the Dirac matrices,  $m_f$  is the quark mass, and  $D^\mu$  is the covariant derivative given by the

$$D_{ij}^\mu = \delta_{ij} \partial^\mu + ig(t^a)_{ij} A_a^\mu, \quad (1.2)$$

where  $t_{ij}^a$  are  $3 \times 3$  hermitian matrices, with elements  $(\lambda^a)_{ij}/2$  and  $\lambda^a$  are the Gell-Mann matrices. The gluon field strength tensor  $F_a^{\mu\nu}$  is given by

$$F_a^{\mu\nu} = \partial^\mu A_a^\nu - \partial^\nu A_a^\mu + gf^{abc} A_b^\mu A_c^\nu, \quad (1.3)$$

where  $A_a$  ( $a = 1..8$ ) are the gluon fields, the last term corresponds to the interaction of the gluons with themselves as they also carry a color-charge. One of the unique properties of QCD is the so-called color confinement, which means that colored quark cannot exist in a free state. Another QCD property is asymptotic freedom, a reduction of the interaction coupling  $\alpha_s$  at small distances, and high values of the interaction scale.

#### Perturbative QCD and renormalization

Calculations of some interaction processes in QFT originate from the Fermi's Golden Rule and perturbation expansion of the Transition Matrix Element. The QCD part, which is used in the perturbative regime, is called pQCD. This regime is based on

<sup>1</sup>Not exact symmetry due to difference in quark masses.

perturbation series calculation of considered process. At higher-order correction calculations, various mathematical obstacles lead to divergences in integration. Some of them are called infrared and collinear divergences and are common for QCD and QED. They appear from the multiple emission of particles with very low energy, or from considered very small angles between the particle and its radiation. For example, a simplified cross-section expression for the process where a soft photon emerges from an electron after some scattering process ( $e^-(p+k) \rightarrow \gamma(k) + e^-(p)$ ) is given by:

$$d\sigma \sim \left| \frac{\alpha}{(p+k)^2} \right|^2 \frac{d^3k}{E_k} \sim \frac{\alpha^2}{E_p^2(1-\cos\vartheta_{e\gamma})} \frac{dE_k}{E_k} d\Omega, \quad (1.4)$$

where  $E_p$ ,  $E_k$ , are the energies of the final electron and photon, the angle between them is given by  $\vartheta_{e\gamma}$ . From this expression, infrared singularity arises after integration of the  $dE_k/E_k$  spectrum down to zero energy. Collinear singularity appears when  $\vartheta_{e\gamma} \rightarrow 0$ . Calculation of a finite result requires the usage of the approach called "renormalization". Various techniques were proposed over the years, some of them are Pauli and Villars approach, the minimal subtraction scheme of t'Hooft and Veltman (MS), and the modified minimal subtraction scheme ( $\overline{\text{MS}}$ ). Usage of these techniques leads to dependence on the artificially included parameter, renormalization scale,  $\mu_r$ .

### Strong coupling

Considering processes at higher energies, a contribution from virtual pairs, caused by the uncertainty principle, cannot be neglected. This makes a coupling constant to be dependent on the scale at which observation is performed. This effect is known as a "running of the coupling", in QFT it is expressed through a beta-function,  $\beta(g)$ , defined by:

$$\beta(g) \equiv \mu \frac{\partial g}{\partial \mu} = \frac{\partial g}{\partial \ln(\mu)}, \quad (1.5)$$

here  $\mu$  corresponds to the energy scale of the given process. In QCD, the beta-function can be calculated as an expansion in  $\alpha_s$

$$\beta(\alpha_s) = -b_0\alpha_s^2 - b_1\alpha_s^3 - b_2\alpha_s^4 + O(\alpha_s^5), \quad (1.6)$$

which depends on the number of active flavors,  $n_f$ :

$$\begin{aligned} b_0 &= \frac{33 - 2n_f}{12\pi}, \\ b_1 &= \frac{153 - 19n_f}{24\pi^2}, \\ b_2 &= \frac{77139 - 15099n_f + 325n_f^2}{3445\pi^3}. \end{aligned} \quad (1.7)$$

The QCD beta-function takes negative values, which finds its reflection in asymptotic freedom, meaning the interaction intensity decreases as the process scale increases. The



equation for the running coupling at leading order can be written as:

$$\alpha(Q^2) = \frac{\alpha(\mu^2)}{1 + \alpha(\mu^2)b_0\ln(\frac{Q^2}{\mu^2})}. \quad (1.8)$$

Figure 1.2 shows the experimental evidence of the strong coupling dependence on the scale of  $Q$ . The world average value of the strong coupling in perturbative regime is  $\alpha_s = 0.1181 \pm 0.0011$  at  $Q^2 = M_{Z^0}^2$  [14].

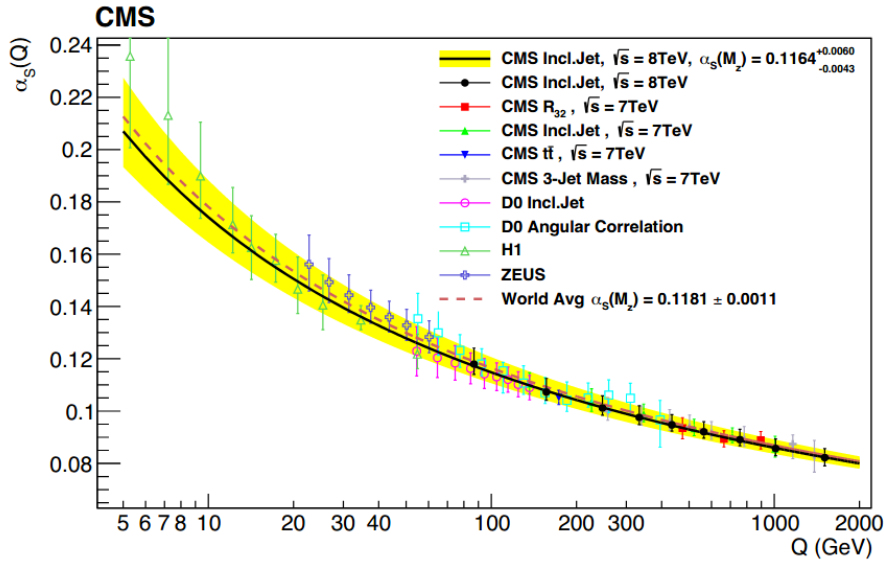


FIGURE 1.2: The running of  $\alpha_s(Q)$  as a function of the scale  $Q$ . The solid line and the uncertainty band are calculated by evolving the extracted  $\alpha_s(m_Z)$  values using the 2-loop 5-flavor renormalization group equations [15].

## 1.1.2 Electroweak interaction

### Quantum Electrodynamics

Quantum Electrodynamics (QED) is a branch of the SM that describes electromagnetic interactions on the quantum level, formulated in terms of relativistic quantum field theory. QED describes interactions between fermions with an electric charge through the mediation of a mass-less gauge boson, photon ( $\gamma$ ). Unlike gluons, photons do not carry a charge; hence they are not self-interacting. The beta function for the QED coupling gives an increase with increasing energy. In 2019, the Committee on Data for Science and Technology recommended the value of the  $\alpha_{EM}$  constant as [16]

$$\alpha_{EM} \equiv \frac{e^2}{4\pi\epsilon_0\hbar c} = 0.0072973525693(11). \quad (1.9)$$

Up to date, the most precise experimental measurement of  $\alpha$  is obtained measuring the magnetic moment of the electron using a one-electron quantum cyclotron apparatus [17].

$$\alpha^{-1} = 137.035999174(35). \quad (1.10)$$

The running of the coupling is also inherited to  $\alpha_{\text{EM}}$ , but unlike  $\alpha_s$ , at higher energies,  $Q^2 = m_W^2$ , the value of the coupling is higher [14]:

$$\alpha_{\text{EM}}^{-1} \sim 128. \quad (1.11)$$

The mathematical formulation of QED is built using the abelian gauge field with the symmetry group  $U(1)_{\text{em}}$ . The QED Lagrangian density is given by:

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(i\gamma_\mu D^\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \quad (1.12)$$

Here  $\psi$  is a bi-spinor field of a charged lepton,  $D_\mu$  is the QED gauge covariant derivative defined by:

$$D_\mu = \partial_\mu + ieA_\mu, \quad (1.13)$$

where  $e$  is a coupling constant,  $m$  is a mass of a charged lepton,  $A_\mu$  is the covariant four-potential of the electromagnetic field generated by the charged lepton,  $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$  is the electromagnetic field tensor.

### Weak interaction

All known fermions can interact via the weak force, exchanging three types of bosons,  $W^+$ ,  $W^-$ , and  $Z^0$ . Unlike for QED and QCD, these bosons have a very high mass,  $m_{W^\pm} = 80.379 \pm 0.012$  GeV,  $m_{Z^0} = 91.1876 \pm 0.0021$  GeV [14], which limits the distance of interaction to  $10^{-17} - 10^{-16}$  m.

In 1956, in the Wu experiment it was discovered that P-parity is violated by the weak interaction [18]. The explanation was proposed in 1957 in the "universal theory of weak interactions", or "V-A theory" by Feynman, Lee, and Yang introducing projection operators that can transform a four-component spinor into a two-component left or right-handed spinor:

$$P_R \equiv \frac{1 + \gamma^5}{2}, \quad P_L \equiv \frac{1 - \gamma^5}{2}, \quad (1.14)$$

where  $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ , and  $\gamma^i$  are the gamma matrices. This quantity is introduced as chirality, in which mathematical representation reflects whether the particle transforms in a left- or right-handed representation of the Poincare group. The weak symmetry is introduced through the weak isospin quantum number, usually marked as T or I. In quark flavor transformation the absolute value of the third component of the weak isospin,  $|T_3|$  is conserved. In this process, a quark can change its flavor only into one of the flavors of the opposite type. Figure 1.3 shows interaction relation between all six quarks, grouped into doublets of up and down with  $T_3 = +\frac{1}{2}$  and  $T_3 = -\frac{1}{2}$ , respectively. Up and down types of quarks and leptons are shown in Fig. 1.1.

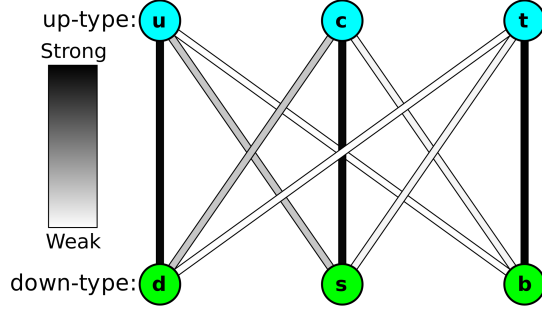


FIGURE 1.3: The strengths of the weak interactions between the six quarks. The "intensities" of the lines are determined by the elements of the CKM matrix [19], [20].

In particle-antiparticle relation, left-handed particles have  $T_3 = -\frac{1}{2}$ , which have corresponding right-handed anti-particles with  $T_3 = +\frac{1}{2}$ . However, right-handed fermions and left-handed anti-fermions do not interact via the weak force and hence do not form a doublet. From CPT symmetry [21], the weak interaction also violates charge conjugation and time-reversal symmetries. One of the first observations of indirect CP-violation was done in 1964 during neutral kaons transforming into their anti-particles and vice-versa when such a process did not occur with the same probability in both cases. In 1990 the NA31 experiment at CERN had the first evidence of the direct CP violation, which was proved in 1999 by the NA48 at CERN [22] and KTeV experiment at Fermilab. In 2001, 2011, 2013, and 2019 direct CP violation was observed by BaBar [23], Belle [24], and LHCb [25] [26] experiments in decays of B and D mesons.

### Unified electroweak theory

In 1968 Sheldon Glashow, Steven Weinberg, and Abdus Salam proposed a model (GWS model) that unifies electromagnetic and weak interactions above the scale of 246 GeV [27]. The model requires the corresponding Lagrangian to be invariant under transformations described by  $SU(2)_L \otimes U(1)_Y$  symmetry group, where  $U(1)_Y$  is an analog of  $U(1)_{em}$  group but with hypercharge,  $Y$ , defined as

$$Y = 2(Q - T_3), \quad (1.15)$$

where  $Q$  is electric charge, and  $T_3$  is the third component of weak isospin. The gauge fields are the weak hypercharge field  $B$ , and weak isospin fields  $W_1, W_2, W_3$ . The Lagrangian of the electroweak interaction is defined by

$$\mathcal{L}_{EW} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}W_{\mu\nu}W^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}, \quad (1.16)$$

where  $W_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu - g(W_\mu \times W_\nu)$ , and  $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$  are field strength tensors. The covariant derivative is given by

$$D_\mu = \partial_\mu - ig'\frac{1}{2}YB_\mu - igT^a W_\mu^a, \quad (1.17)$$

where  $g'$  is electromagnetic and  $g$  is weak coupling constant. The gauge fields in the Lagrangian satisfy the symmetry requirement being mass-less. In GWS model the gauge

fields  $W_3$  and  $B$  transform into  $\gamma$  and  $Z^0$  through a weak mixing angle, Weinberg angle,  $\vartheta_W$ :

$$\begin{pmatrix} \gamma \\ Z^0 \end{pmatrix} = \begin{pmatrix} \cos \vartheta_W & \sin \vartheta_W \\ -\sin \vartheta_W & \cos \vartheta_W \end{pmatrix} \begin{pmatrix} B \\ W_3 \end{pmatrix}, \quad (1.18)$$

while  $W^+$  and  $W^-$  are a superposition of  $W_1$  and  $W_2$

$$W^\pm = \frac{1}{\sqrt{2}} W_1 \mp W_2. \quad (1.19)$$

The interaction of the Higgs field with  $W^\pm$  and  $Z^0$  generates the mass of these particles, making weak interaction weak and different from the electromagnetic force.

## 1.2 Proton structure

One of the first experiments on probing the proton structure was made in 1956 by E.E. Chambers and R. Hofstadter [28]. In this experiment, the proton structure (charge distribution) was studied analyzing scattered high-energy electrons from protons in polyethylene at the energies in the laboratory system of 200, 300, 400, 500, 550 MeV. Results showed a strong deviation from a point-like proton model. Since that time, many various experiments have proved the non-triviality of proton structure.

Nowadays, proton is classified as the lightest baryon in the baryon octet with the mass  $m_p = 938.27$  MeV, a lower bound on its half-life is estimated to be  $>5.8 \times 10^{29}$  years [14]. In 2018 two independent groups measured a proton charge radius to be near  $r_p = 0.833 \pm 0.01$  fm [29] and  $r_p = 0.831 \pm 0.007 \pm 0.012$  fm [30]. It is known that proton consists of partons, three so-called "valence quarks" (uud), gluons and sea quarks (fluctuating  $g \leftrightarrow q\bar{q}$  states). It has a charge of +1, isospin  $I_3 = \frac{1}{2}$ , and parity  $P = +1$ . The spin of a proton is  $S = \frac{1}{2}$ , but its nature is still a topic for debates due to experimental results, which showed that the valence quark spin contribution into the total proton spin is at the level of 4% – 24% [31]. Despite various experimental results in this area, a theoretical model that would provide a prediction for proton structure dynamics is not reachable due to non-perturbative QCD phenomena and the absence of valid mathematical methods. So far, the Parton Distribution Functions (PDF) are estimated using experimental data and phenomenological assumptions.

### 1.2.1 Parton Distribution Functions

Significant progress in understanding of the proton structure was achieved after the beginning of "Deep Inelastic Scattering" (DIS) experiments. In these experiments, high energetic leptons ( $e^\pm$ ,  $\mu^-$ ,  $\nu$ ) were scattering on hadrons. Early experiments with the fixed target and lepton ( $e^-$ ,  $\mu^-$ ,  $\nu$ ) beams were performed in SLAC and CERN. Later,  $e^\pm p$  scattering was performed at the HERA accelerator at DESY. The experiments were able to reach the level of quark and gluon resolution, probing the hadron state. The lepton-proton scattering diagrams are shown in Fig. 1.4. The interaction is performed through a vector boson exchange ( $W^\pm$ ,  $Z^0$ , or  $\gamma$ ). Depending on the boson charge, processes are defined as Neutral Current (NC) and Charged Current (CC).

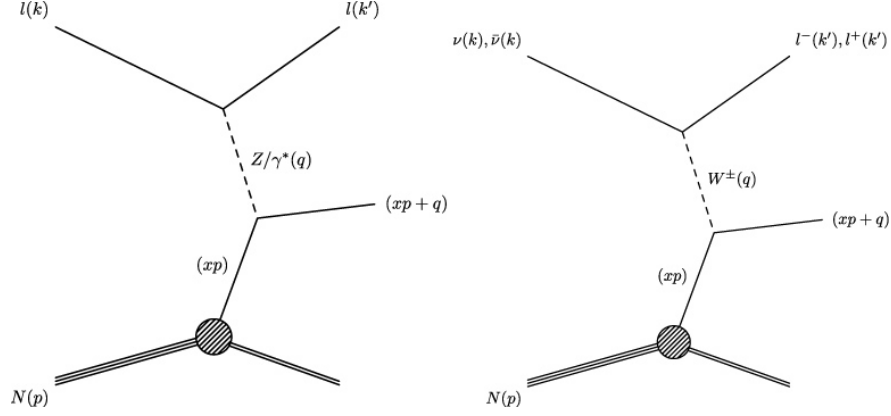


FIGURE 1.4: Diagrams represent deep inelastic scattering via exchange of  $Z^0/\gamma^*$  boson on the left, and  $W^\pm$  on the right. Lepton four-momentum before and after interaction is shown as  $k$  and  $k'$  respectively. The proton initial state momentum is shown as  $p$ .

The four-momentum of the boson is given by

$$q \equiv k' - k. \quad (1.20)$$

The center of mass energy squared is given by

$$s = (p + k)^2. \quad (1.21)$$

The squared invariant mass of the virtual exchanged boson is defined as

$$Q^2 \equiv -q^2. \quad (1.22)$$

The fraction of the momentum of the incoming proton taken by the struck quark is called Bjorken  $x$  variable

$$x = \frac{Q^2}{2p \cdot q}. \quad (1.23)$$

The measure of the energy amount transferred between hadron and lepton systems is called inelasticity and is given by

$$y = \frac{p \cdot q}{p \cdot k}. \quad (1.24)$$

In the deep inelastic scattering formalism, the structure functions,  $F_1(x, Q^2)$ ,  $F_2(x, Q^2)$ ,  $F_3(x, Q^2)$ , and  $F_L(x, Q^2)$  are defined as a measure of the partonic structure. The differential cross-sections for NC and CC, are defined

$$\frac{d^2\sigma_{\text{NC}}^\pm}{dx dQ^2} = \frac{2\pi\alpha_{\text{EM}}^2}{Q^4 x} [Y_+ \tilde{F}_2^\pm(x, Q^2) - y^2 \tilde{F}_L^\pm(x, Q^2) \mp Y_- \tilde{F}_3^\pm(x, Q^2)], \quad (1.25)$$

$$\frac{d^2\sigma_{\text{CC}}}{dx dQ^2} = \frac{M_W^4}{(Q^2 + M_W^2)^2} [Y_+ W_2^\pm(x, Q^2) - y^2 W_L^\pm(x, Q^2) \mp Y_- x W_3^\pm(x, Q^2)] \frac{G_F^2}{4\pi x}. \quad (1.26)$$

Here  $Y_\pm = 1 \pm (1-y)^2$ ,  $G_F$  is the Fermi constant. The generalized structure functions,  $\tilde{F}_{2,3}(x, Q^2)$ , may be written as linear combinations of the proton structure functions  $F_2$ ,  $F_{2,3}^{\gamma Z}$ , and  $F_{2,3}^Z$ .  $W_2^\pm(x, Q^2)$ ,  $xW_3^\pm(x, Q^2)$  and  $W_L^\pm(x, Q^2)$  are the structure functions

for CC  $e^\pm p$  scattering. In DIS experiments, structure functions were measured as a function of double differential cross-sections,  $d^2\sigma/dxdQ^2$ . Until the present day, data collected by H1 and ZEUS are considered as one of the most important contributors to parton structure determinations.

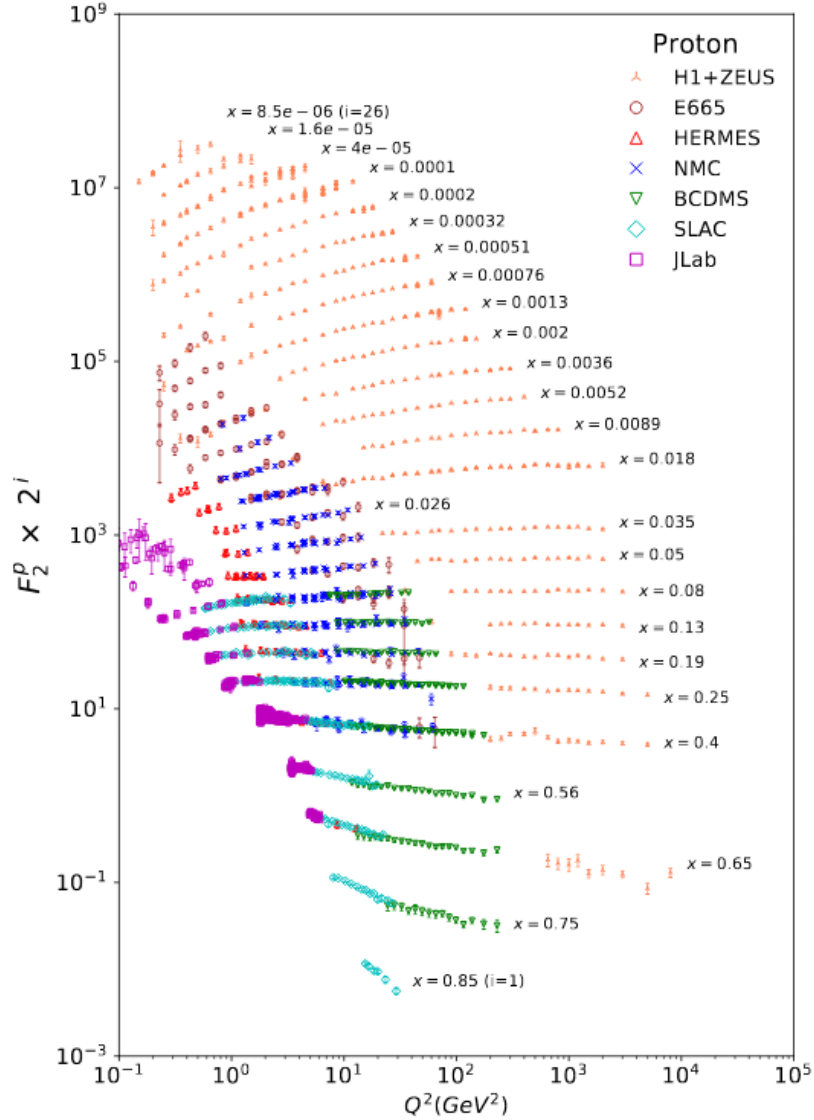


FIGURE 1.5: The proton structure function  $F_2^p$  measured in electromagnetic scattering of electrons and positrons on protons, and for electrons/positrons (SLAC, HERMES, JLAB) and muons (BCDMS, E665, NMC) on a fixed target. Statistical and systematic errors added in quadrature are shown. The H1+ZEUS combined values are obtained from the measured reduced cross-section and converted to  $F_2^p$  with a HERA-PDF NLO fit, for all measured points where the predicted ratio of  $F_2^p$  to reduced cross-section was within 10% of unity. The data are plotted as a function of  $Q^2$  in bins of fixed  $x$  [14].

Figure 1.5 shows experimental results on the structure function measurement of  $F_2^p$  collected in different experiments, including H1 and ZEUS.

With the start of the LHC experiments, many new measurements, sensitive to the proton structure, became available. The kinematic regions probed so far<sup>2</sup> by various experiments, including LHC, are shown in Fig. 1.6. In the early DIS experiment results it was found [32] that the structure function  $F_2(x, Q^2)$  does not depend<sup>3</sup> on the scale  $Q^2$ . The effect is called Bjorken scaling [33] meaning

$$\lim_{Q^2 \rightarrow +\infty} F_i(x, Q^2)|_{x=\text{const}} \rightarrow \tilde{F}_i(x).$$

To explain that phenomena, Feynman proposed the idea of partons, developed in a model called "Parton model" (or later "Quark-Parton Model" (QPM)) [34]. The model states that a hadron (in this case proton) consists of point-like non-interacting scattering centers, partons.

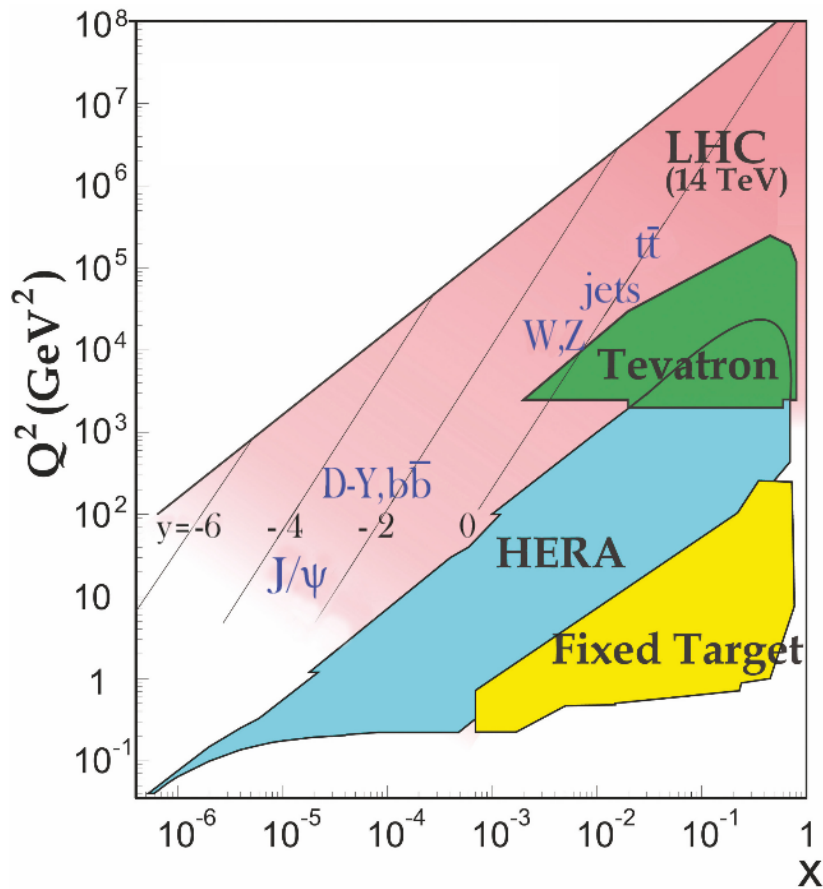


FIGURE 1.6: Kinematic domains in  $x$  and  $Q^2$  probed by fixed-target and collider experiments, where  $Q^2$  can refer either to the literal  $Q^2$  for deep inelastic scattering, or the hard scale of the process in hadron-hadron collisions, e.g. invariant mass or transverse momentum  $p_T^2$ . Some of the final states accessible at the LHC are indicated in the appropriate regions, where  $y$  is the rapidity. The incoming partons have  $x_{1,2} = (Q/14 \text{ TeV}) e^{\pm y}$  where  $Q$  is the hard scale of the process shown in blue in the figure. [14].

<sup>2</sup>Up to the fall of 2019.

<sup>3</sup>In the first order approximation.

In the QPM frame, the transverse structure function can be re-written as

$$F_2(x) = \sum_i e_i^2 x f_i(x), \quad (1.27)$$

where the sum over partons with a given charge  $e_i$  and function  $f_i(x)$  is unknown, however, the  $f_i(x)$  is interpreted as a probability that the  $i$ -th parton is carrying  $x$  fraction of the proton momentum, parton momentum density.

Various quark-parton model tests have shown that the proton consists not only of the valence quarks but also has sea-quarks, which consists of quark-antiquark pairs with no overall flavor. Taking this into account

$$F_1(x) = \sum_i e_i^2 [x q_i(x) + x \bar{q}_i(x)]. \quad (1.28)$$

Knowing the number of valence quarks allows to define some additional constraints, the so-called "sum rules"

$$\int_0^1 u_v(x) dx = 2, \quad \int_0^1 d_v(x) dx = 1, \quad (1.29)$$

where  $u_v$  and  $d_v$  are distributions for valence up and down quarks, respectively. A very important momentum sum rule defines the sum over all quark and antiquark types in the proton

$$\int_0^1 x \Sigma(x) dx = 1, \quad (1.30)$$

where  $\Sigma(x) = u(x) + \bar{u}(x) + d(x) + \bar{d}(x) + s(x) + \bar{s}(x) + c(x) + \bar{c}(x) + b(x) + \bar{b}(x)$ <sup>4</sup>.

Neutrino-nucleus and anti-neutrino-nucleus scattering experiments have shown [35] that only 50% of proton momentum is coming from quarks, the rest was contributed by particles that interact only via strong force - gluons. The gluon was discovered in 1979 with the PETRA collider at DESY [36]. To account for gluon contribution, the momentum sum rule is given by

$$\sum_i \int_0^1 x f_i(x) dx + \int_0^1 x g(x) dx = 1 \quad (1.31)$$

### 1.2.2 Factorization theorem and PDF evolution equations

Cross-section calculations of the hard interactions, calculable in pQCD, rely on the assumption that the time scale of the hard scatter is relatively much shorter than the final state hadronization [37]. The approach is formulated as a Factorization theorem, which allows factorizing the cross-section into the probability of the hard scattering and a probability density for finding a parton that carries a certain amount of hadron momentum [38]. In this scheme, the parton-parton interaction cross-section can be calculated in the pQCD regime using a general approach at LO, NLO, NNLO, etc. with the corresponding strong coupling defined at renormalization scale  $\mu_r$  and order  $p$ . The non-perturbative part of the process, the PDF, is separated from the pQCD, introducing a factorization scale,  $\mu_f$ , which is not necessarily equivalent to  $\mu_r$ . Infrared singularities are handled by renormalization at the factorization scale; collinear singularities are

<sup>4</sup>The top quark may be also included.



absorbed into the parton density at scale  $\mu_f$ . Figure 1.7 shows a schematic picture of the factorized process in the proton-proton collision.

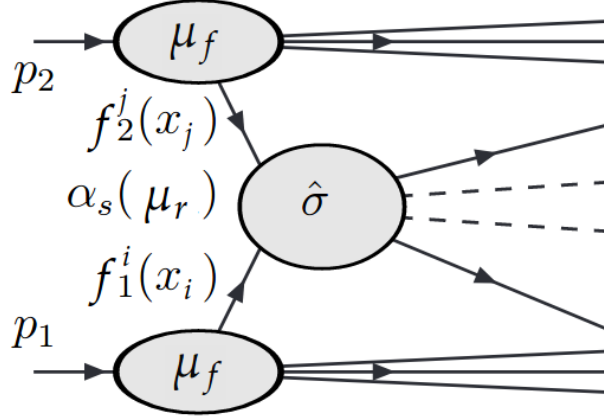


FIGURE 1.7: Representation of proton-proton collision in the factorization approach. Each proton carries momentum  $p$  with its non-perturbative PDF,  $f(x)$ . Strong coupling at renormalization scale is shown as  $\alpha_s(\mu_r)$ . Parton-parton interaction area is shown with  $\sigma$ .

Formula 1.32 reflects the factorization formulation. The sums are performed over partons  $i, j$  of PDFs  $f_1(x)$ ,  $f_2(x)$ , respectively. The order of  $\alpha_s$  is shown as  $p$ , the differential cross-section as a function of  $dx$  for each sub-process  $ij$  at order  $p$  is given by  $d\sigma_p^{ij}/dx$ .

$$\frac{d\sigma}{dx} \sim \sum_i \sum_j \sum_p \int_0^1 \int_0^1 f_{i/1}(x_i, \mu_f^2) \alpha_s^p(\mu_r^2) \frac{d\sigma_p^{ij}}{dx}(x_1, x_2, \mu_f^2, \mu_r^2; s) f_{j/2}(x_j, \mu_f^2) dx_i dx_j \quad (1.32)$$

This approach became a useful tool for the cross-section calculation of various processes. However, these calculations became dependent on the factorization scale. From a theoretical point, factorization theorem is a complicated approach that should be treated carefully, more details on that topic may be found in appropriate papers [38], [39].

Another theoretical approach in proton structure determination became available after numerous and continuous developments of Y. Dokshitzer, V. Gribov, L. Lipatov, G. Altarelli, and G. Parisi. In their papers [40] [41] they studied the PDF dependence on  $\mu_f$  using renormalization group equations. As a result, PDF evolution equations for quarks and gluons (DGLAP equations) were proposed. The equations describe how the PDF changes (evolves) through splitting functions  $P_{qq}$ ,  $P_{qg}$ ,  $P_{gq}$ , and  $P_{gg}$  as the factorization scale changes. DGLAP equations consist of gluon and quark parts and are given by

$$\frac{\partial q_i(x, \mu^2)}{\partial \ln \mu^2} \equiv \frac{\alpha_s}{2\pi} \int_x^1 \frac{dz}{z} \left[ q_i(z, \mu^2) P_{qq}(x/z) + g(z, \mu^2) P_{qg}(x/z) \right], \quad (1.33)$$

$$\frac{\partial g(x, \mu^2)}{\partial \ln \mu^2} \equiv \frac{\alpha_s}{2\pi} \int_x^1 \frac{dz}{z} \left[ \sum_i q_i(z, \mu^2) P_{gq}(x/z) + g(z, \mu^2) P_{gg}(x/z) \right]. \quad (1.34)$$

The splitting functions  $P_{ij}$  ( $i, j = q, g$ ) represent the probability that a parton  $i$  will

radiate a parton  $j$ , which carries  $z$  momentum fraction of parton  $i$ . The corresponding LO diagrams are shown in Fig. 1.8. The exact expressions for splitting functions may be found in already mentioned papers.

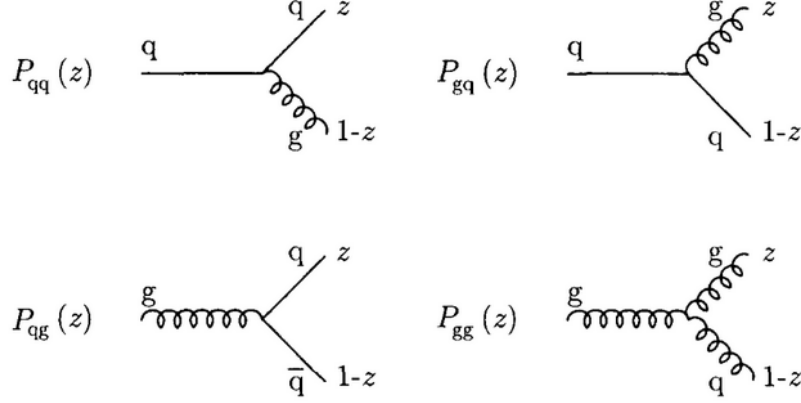


FIGURE 1.8: Feynman diagrams of  $P_{qq}$ ,  $P_{gq}$ ,  $P_{qg}$ , and  $P_{gg}$  splitting functions of DGLAP equations. [37].

### 1.2.3 PDF extraction using a global QCD fit

Perturbative QCD methods allow to evolve parton distributions with any scale  $Q^2$ ; however, the analytic form for  $x$  dependence remains unknown. For that reason, the global fit approach was proposed.

The technique starts with analytic<sup>5</sup> shapes for the parton distributions at a certain starting scale  $Q_0^2$ , at which  $\alpha_s$  is small enough to use pQCD approach ( $1.0 < Q_0^2 < 5.0 \text{ GeV}^2$ ). Valence quarks,  $u_v$ ,  $d_v$ , total sea quarks,  $S$ , gluon,  $g$  are usually chosen as a default scheme. Considering analytic function as  $x$ -dependent parameterized function, a typical choice for  $u_v$ ,  $d_v$ ,  $S$ , and  $g$  is

$$xu_v = A_u x^{\lambda_u} (1-x)^{\eta_u} P(x, u), \quad (1.35)$$

$$xd_v = A_d x^{\lambda_d} (1-x)^{\eta_d} P(x, d), \quad (1.36)$$

$$xS = A_S x^{-\lambda_S} (1-x)^{\eta_S} P(x, S), \quad (1.37)$$

$$xg = A_g x^{-\lambda_g} (1-x)^{\eta_g} P(x, g). \quad (1.38)$$

Here  $P(x, i)$  is a polynomial as a function of  $x$  or  $\sqrt{x}$ , or sum of Chebyshev polynomials. The normalization parameters  $A_i$  can be used for the sum rules normalization.

Using DGLAP equations, these functions are evolved to considered scales, where they are convoluted with coefficient functions. The obtained results are used to provide predictions for the structure functions or cross-section calculations, which are then fitted to the data sets, sensitive to proton structure. The fit is usually performed over data points distributed in a wide kinematic domain, as shown in Fig. 1.6. The  $\chi^2$ -method is used in order to validate the fit quality.

An example of fitted PDF distribution as a function of  $x$  is shown in Fig. 1.9. The plot shows results obtained from HERA data and CMS  $W^\pm$  production at  $\sqrt{s} = 7 \text{ TeV}$ ,

<sup>5</sup>Derived from non-perturbative methods or by parameterizing the  $x$  dependent function.

PDFs are given for  $xu_v$ ,  $xd_v$ ,  $x\Sigma = 2x(U + D)$ ,  $xg$  for  $Q^2 = 1.9 \text{ GeV}^2$  and  $Q^2 = m_W^2 \text{ GeV}^2$ .

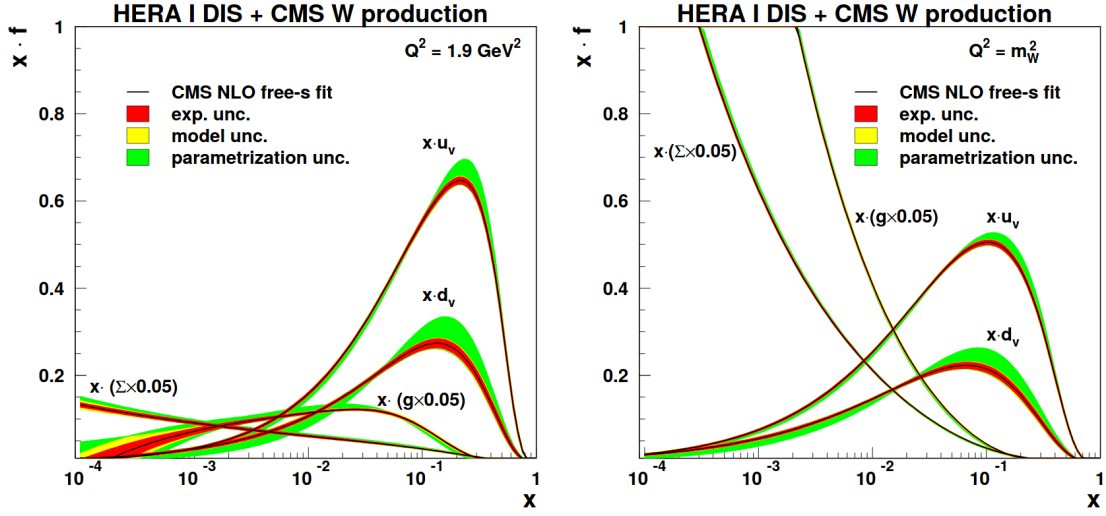
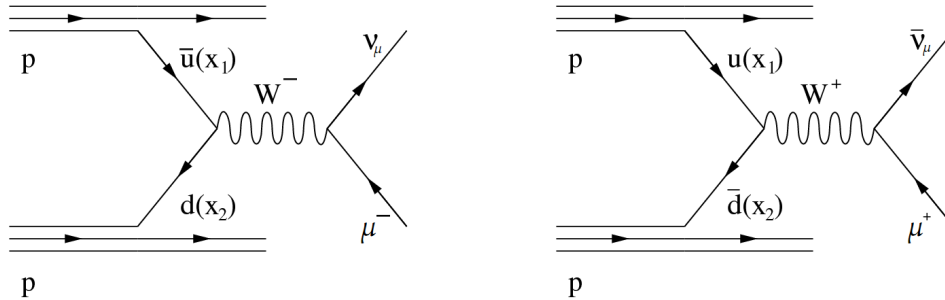


FIGURE 1.9: The parton distribution functions of valence quarks,  $xu_v$ ,  $xd_v$ , scaled sea quarks,  $x\Sigma$ , and gluons,  $xg$  at scale  $Q^2 = 1.9 \text{ GeV}^2$  (left plot) and  $Q^2 = m_W^2 \text{ GeV}^2$  (right plot) [42].

### 1.3 $W^\pm$ boson production at LHC

A direct production of  $W^\pm$  boson at LHC is possible through the Drell-Yan process of  $q\bar{q}$  annihilation. Leading order diagrams are shown in Fig. 1.10



(A)  $\bar{u}d \rightarrow W^- \rightarrow \mu^- \nu_\mu$ .

(B)  $\bar{d}u \rightarrow W^+ \rightarrow \mu^+ \bar{\nu}_\mu$ .

FIGURE 1.10: Feynman diagram for  $W^\pm$  boson production in  $pp$  collisions, decaying in the muon channel. Left figure show the diagram for  $W^-$  production,  $W^+$  is shown on the right.

Figure 1.11 shows the percentage of parton flavor decomposition of the  $W^\pm$  total cross-section in  $pp$  and  $p\bar{p}$  collisions as a function of center of mass energy. Dominant contribution is given by valence-sea-quark annihilation,  $u\bar{d}$  and  $\bar{u}d$ . The contribution from processes  $s\bar{c}$  and  $c\bar{s}$  slowly rises with the scale; the rest is suppressed by off-diagonal CKM matrix elements. The  $W^\pm$  decay modes are dominated by hadronic channel with

$67.41 \pm 0.27\%$  of all decays. The rest are lepton modes  $e^\pm\nu_e$ ,  $\mu^\pm\nu_\mu$ ,  $\tau^\pm\nu_\tau$  with fractions  $10.71 \pm 0.16\%$ ,  $10.63 \pm 0.15\%$ ,  $11.38 \pm 0.21\%$ , respectively [14]. From the experimental point of view, lepton channels are of particular interest due to available approaches to clean signal extraction. Figure 1.12 shows the total cross-section results as a function of center of mass energy obtained from  $p\bar{p}$  and  $pp$  collisions.

The total cross-section of  $W^\pm$  production in  $pp$  collisions can be calculated using factorization theorem 1.32. The  $W^\pm$  boson sub-process cross-section production is given by

$$\hat{\sigma}(q_i\bar{q}_j \rightarrow W) = \frac{\pi}{3}\sqrt{2}G_F M_W^2 |V|^2 \delta(\hat{s} - M_W^2), \quad (1.39)$$

where  $i \neq j$ ,  $V$  is a corresponding matrix element. Renormalization and factorization scales are chosen to be at the level of  $W^\pm$  boson mass,  $\mu_r = \mu_f = M_W$ .

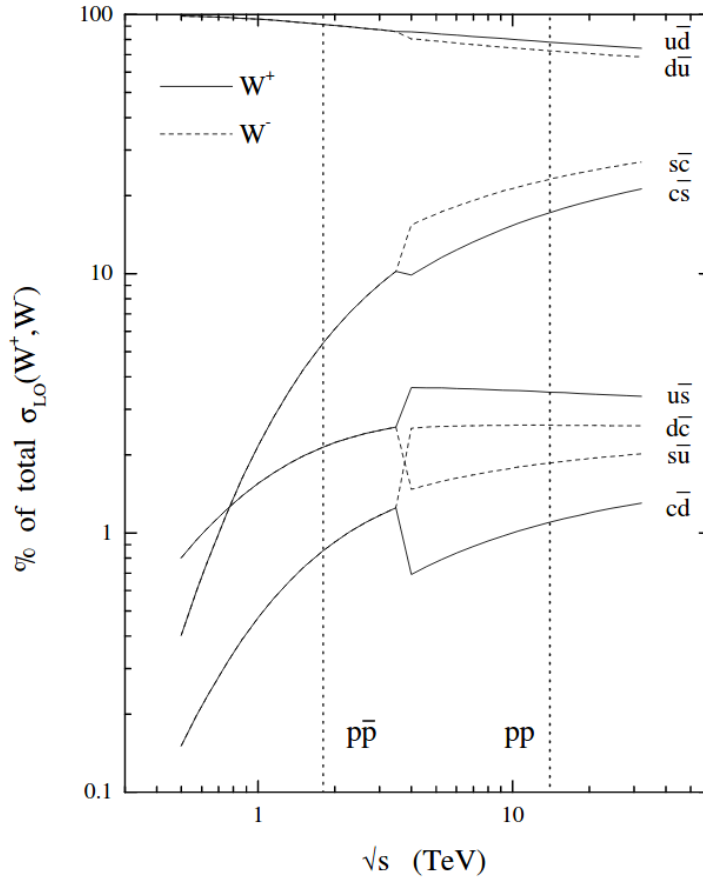


FIGURE 1.11: Individual contributions of flavor decomposition into the total  $W^+$  (solid line) and  $W^-$  (dashed line) cross-section at LO [43].

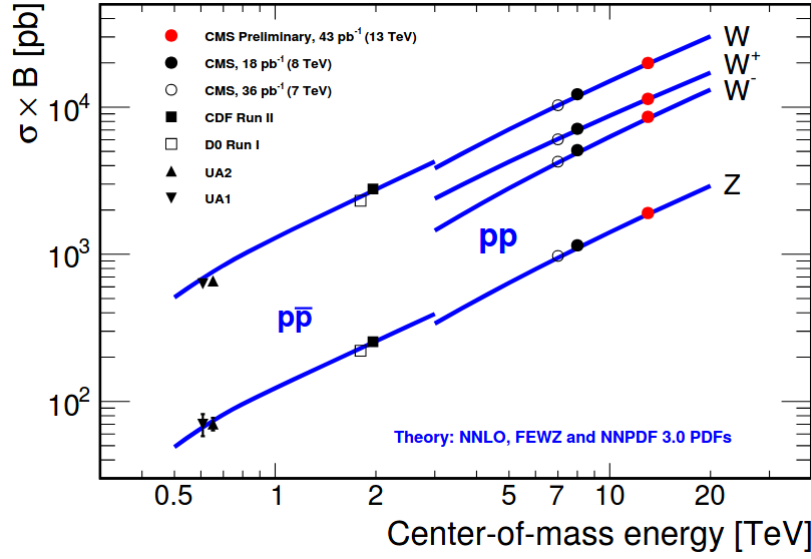


FIGURE 1.12: Measurements of the total  $W^+$ ,  $W^-$ ,  $W$ , and  $Z^0$  production cross sections times branching fractions as a function of center of mass energy for CMS and experiments at lower-energy colliders [44].

### 1.3.1 $W^\pm$ boson charge asymmetry

In LHC collisions the transverse energy of the particles is relatively smaller than the longitudinal one. Due to this condition, it is convenient to use the rapidity,  $y$ , defined as:

$$y \equiv \frac{1}{2} \ln \left( \frac{E + p_z}{E - p_z} \right). \quad (1.40)$$

Here  $E$  is particle energy, and  $p_z$  is a  $z$ -th component of particle momentum along the beam direction. It has a finite range, approximately  $(\ln(m/2E), \ln(2E/m))$ . The rapidity difference is invariant in the case of Lorentz boosts along the  $Z$ -axis. For the wide range of processes in high energy collisions the masses can be ignored; in that case  $y$  reduces to a quantity  $\eta$ , pseudorapidity, which is defined as:

$$\eta \equiv -\ln \left( \tan \left( \frac{\vartheta}{2} \right) \right). \quad (1.41)$$

It covers a range  $(-\infty, \infty)$  that corresponds to  $(-1 < \cos(\vartheta) < 1)$ . Considering the direct  $W^\pm$  boson production at the LHC, the boson four-momentum can be expressed as a sum of quarks

$$p_W = p_q + p_{\bar{q}}, \quad (1.42)$$

where  $p_q$  and  $p_{\bar{q}}$  are quarks momentum. Quark four-momentum in the center of mass frame can be given by

$$p_q = \frac{\sqrt{s}}{2} (x_q, 0, 0, -x_q), \quad p_{\bar{q}} = \frac{\sqrt{s}}{2} (x_{\bar{q}}, 0, 0, -x_{\bar{q}}). \quad (1.43)$$

Using these equations together with Eq. 1.40, it can be shown how the boson rapidity depends on quarks momentum

$$y_W = \frac{1}{2} \ln \left( \frac{x_q}{x_{\bar{q}}} \right). \quad (1.44)$$

The scale of interaction is usually taken as the boson mass,  $q^2 = m_{W^\pm}^2$ , combining this relation with  $s = 4E^2$  it can be shown how the boson mass is connected with the momentum fraction of quarks

$$m_{W^\pm}^2 = x_q x_{\bar{q}} s. \quad (1.45)$$

The final expression for quark momentum fraction as a function of the boson mass, rapidity, and the proton center of mass energy can be derived combining Eq. 1.45 and Eq. 1.44

$$x_q = \frac{m_{W^\pm}}{\sqrt{s}} e^{y_{W^\pm}}, \quad x_{\bar{q}} = \frac{m_{W^\pm}}{\sqrt{s}} e^{-y_{W^\pm}}. \quad (1.46)$$

Taking into account individual contributions of flavor decomposition shown in Fig. 1.11, the attention is focused on  $u\bar{d}$  and  $\bar{u}d$  production mechanism.  $W^+$  bosons are predominantly produced in the direction of up quark, while  $W^-$  are produced in the direction of a down quark. The differential cross-section as a function of rapidity (or pseudorapidity) can be studied to explore the proton structure. It can be derived using the Cabibbo mixing approximation [45] together with an SU(3) symmetric sea approximation [46]

$$\frac{d\sigma}{dy} \sim \frac{2\pi G_F}{3\sqrt{2}} x_a x_b [u(x_a)\bar{d}(x_b) + \bar{d}(x_a)u(x_b)]. \quad (1.47)$$

Since proton has two up and one down quark, up quark has higher chances to be found in the proton than the down quark (see Fig. 1.9), that leads to quantitatively asymmetric production of  $W^+$  bosons over  $W^-$ . The asymmetry can be defined through the  $W^\pm$  boson production as a function rapidity by

$$A_{W^\pm}(y) \equiv \frac{d\sigma_{W^+}/dy - d\sigma_{W^-}/dy}{d\sigma_{W^+}/dy + d\sigma_{W^-}/dy}, \quad (1.48)$$

where  $d\sigma_{W^\pm}/dy$  is a differential cross-section value as a function of boson rapidity. This equation can be re-written in terms of the production mechanism

$$A_{W^\pm}(y) \approx \frac{u(x)\bar{d}(x) - d(x)\bar{u}(x)}{u(x)\bar{d}(x) + d(x)\bar{u}(x)}, \quad (1.49)$$

where  $u(x)$  and  $d(x)$  are PDFs of valence up and down quarks, while  $\bar{u}$  and  $\bar{d}$  are anti-up and anti-down PDFs of sea quarks. The sensitivity to valence quarks can only be obtained for a region of low  $x$  of sea quarks, when approximation  $\bar{d} \sim \bar{u} \sim \bar{q}$  is valid

$$A_{W^\pm}(y) \sim \frac{u(x) - d(x)}{u(x) + d(x)}. \quad (1.50)$$

$W^\pm$  boson charge asymmetry is usually studied using the lepton channel of  $W^\pm$  decay. Lepton asymmetry can be defined by

$$A_{l^\pm}(\eta) \equiv \frac{d\sigma/d\eta_{l^+} - d\sigma/d\eta_{l^-}}{d\sigma/d\eta_{l^+} + d\sigma/d\eta_{l^-}}. \quad (1.51)$$

Due to the vector-axial nature of the leptonic  $W^\pm$  decay, the lepton asymmetry has more complicated explanation. The laws of weak interaction impose limitations on  $W^\pm$  boson and fermion mechanism of coupling. The  $W^-$  boson can interact only with left-handed fermions or right-handed anti-fermions and vice-versa. This makes the dependence of  $W^\pm$  matrix element on the angle between lepton and the boson, sensitive to different boson polarization.

$$|M_-|^2 = g_W^2 m_W^2 \frac{1}{4} (1 + \cos \vartheta)^2, \quad (1.52)$$

$$|M_L|^2 = g_W^2 m_W^2 \frac{1}{2} \sin^2 \vartheta, \quad (1.53)$$

$$|M_+|^2 = g_W^2 m_W^2 \frac{1}{4} (1 - \cos \vartheta)^2. \quad (1.54)$$

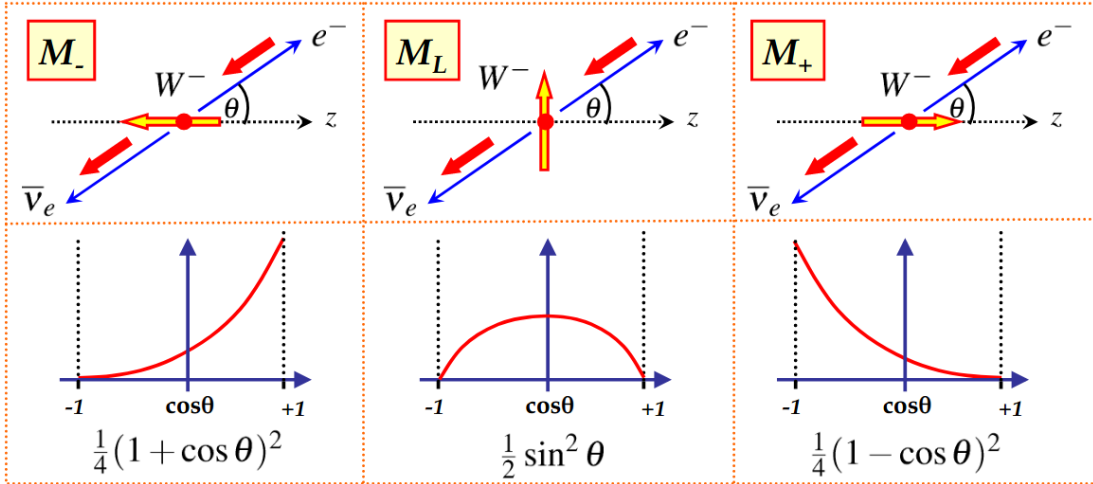


FIGURE 1.13:  $W^\pm$  boson in different helicity states decaying into electron channel in a rest frame. The first row corresponds to three different helicity states, left-handed (negative), longitudinal, and right-handed (positive). Blue arrows are momentum; red arrows are helicity states of daughter particles, while the yellow arrow shows the boson spin projection on the  $z$ -axis. The second row shows the matrix element dependence on the  $\theta$  angle for each helicity state [47].

Figure 1.13 in the first row illustrates different helicity states of  $W^-$  boson in the rest frame. The incoming quark is left-handed, while anti-quark is right-handed, due to angular momentum conservation, the helicity of the  $W^\pm$  must be left-handed, especially for high rapidity. In some cases, anti-quark might have a higher fraction of proton momentum than the valence quark, making the  $W^\pm$  boson polarization right-handed. As a result, angular momentum conservation law, V-A formalism, and matrix element dependence on helicity state cause the difference in  $d\sigma/d\eta$  distributions for  $W^+$  and  $W^-$ . When  $u\bar{d}$  annihilates into  $W^+$ , the dominant part of bosons has left-handed

polarization, which decay with the left-handed neutrino forward and right-handed anti-lepton backward. The  $d\bar{u}$  annihilate into  $W^-$ , and the dominant part of bosons have the left-handed polarization. They decay forward with the right-handed neutrino and backward with the left-handed lepton. Figure 1.14 shows comparison of  $W^\pm$  and  $\mu^\pm$  as a function of  $y_W$  or  $\eta_\mu$ .

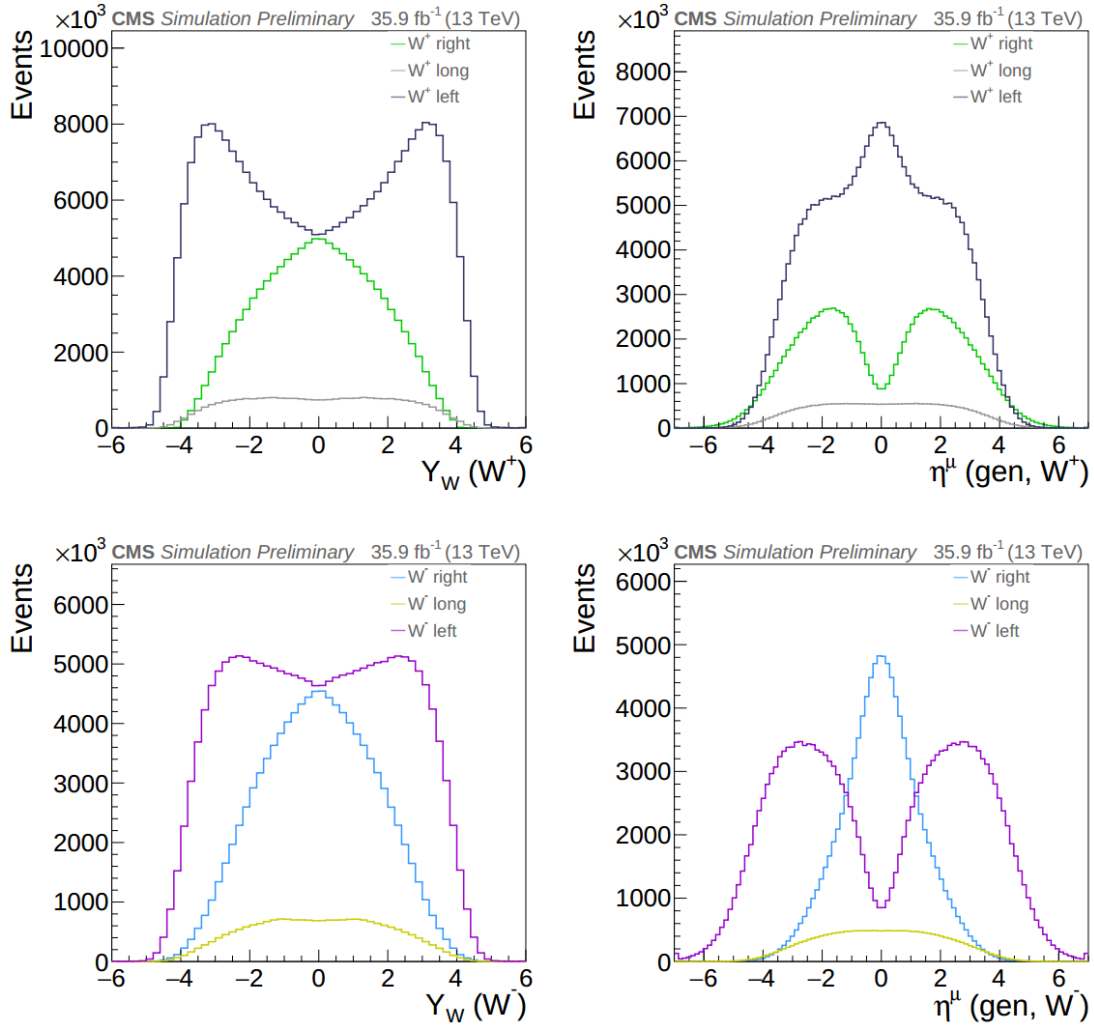


FIGURE 1.14: Two left plots show the comparison  $W^+$  and  $W^-$  events distribution as a function of the boson rapidity for different states of the boson helicity. The right plots shows the same distributions but for muon events as a function of muon pseudorapidity and  $W^\pm$  boson helicity state [48].



## Chapter 2

# LHC and CMS experiment

The leading player on the experimental high energy field today is the Large Hadron Collider (LHC), together with its four main experiments: A Large Ion Collider Experiment (ALICE), A Toroidal LHC Apparatus (ATLAS), Compact Muon Solenoid (CMS), and LHC-beauty (LHCb). Four years after the LHC launch, in July 2012, CMS and ATLAS, independently, reported on the observation of the boson that was predicted to be responsible for the mass generation mechanism. [12], [13]. The observation of the Higgs boson became a great success and a triumph of modern particle physics.

Studies and results presented in this thesis are achieved performing an analysis of data collected with the CMS detector at the LHC. The next section describes the LHC purpose and design. After that, a section is dedicated to the CMS detector and its main components.

## 2.1 Large Hadron Collider

The Large Hadron Collider, LHC, is a superconducting two-ring hadron accelerator and collider. The project was approved in 1994 by the Council of European Organization for Nuclear Research (CERN - Conseil Européen pour la Recherche Nucléaire). The LHC was built between 1998 and 2008 and installed in the 26.7 km LEP tunnel around the border between France and Switzerland. The LEP tunnel has eight arcs and eight straight sections that lie at a depth of 45 m to 170 m under the surface. Figure 2.1 shows the geographical location of the accelerator, and its smaller pre-accelerator ring (Super Proton Synchrotron).

The LHC purpose is to collide hadrons, in particular, protons and ions. In this thesis, the results are obtained, analyzing data from the proton-proton collisions, for that reason, the main focus is dedicated to the proton acceleration, shown in Fig. 2.2. At first, the protons, derived from the hydrogen gas ionization, are passed to the first stage of acceleration in LINAC 2. Linac 2 is a linear accelerator equipped with radio-frequency quadrupole with an output current of up to 180 mA; at this stage, the protons are accelerated to 50 MeV. After Linac 2, the protons proceed to a circular proton accelerator Proton Synchrotron Booster (PSB). The PSB has four stacked 25 meters radius rings with 16 periods of double-dipole magnets each. The protons formed in beams are accelerated up to 1.4 GeV. Afterward, they are transmitted to the Proton Synchrotron (PS) - a 100 meters radius synchrotron supplied with a hundred dipoles magnets.

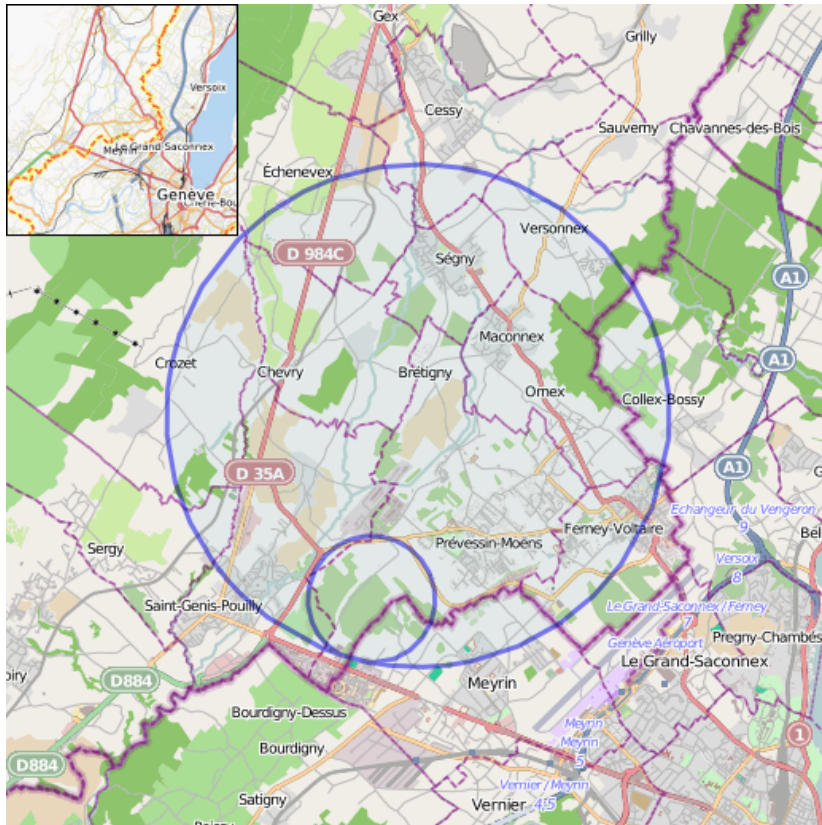


FIGURE 2.1: Geographic location of the LHC [49].

At this stage, protons gain 23.6 GeV and are ready to be directed to the Super Proton Synchrotron (SPS). The last stage of pre-acceleration is completed in a SPS circular tunnel of 6.9 kilometers at 450 GeV. Finally, the proton beams, formed in up to  $2,808^1$  bunches per ring, each filled with  $1.15 \times 10^{11}$  protons in a beam size of 3.5 micrometers<sup>2</sup>, are injected in LHC to be accelerated to 7, 8 or 13 TeV.

The main acceleration stage is done in the 27 km LHC ring. The whole cycle is maintained by a state of the art machinery and is described very briefly. The central acceleration technology is a superconducting NbTi Rutherford cables cooled to a superconducting temperature and operated at a magnetic field near 8.3 Tesla. In total near 150 tons of liquid helium is needed to keep the whole 27 km beam pipe at 1.9 Kelvin temperature. The dipole magnets responsible for a precise beam alignment and its stability are called "lattice magnets". They consist of 1232 main dipoles, weight 35 tons, and are 15 meters long each. The lattice magnets are also supplied with sextupole, octupole, and decapole magnets for a fine-tuning of the magnetic field. When the target energy is reached, the collisions are performed in the points of interactions using special magnets, where the detectors are located. Before the collision, entering the detector area, the bunches are squeezed closer together to increase the probability of interaction with incoming bunches from the opposite direction. This procedure is regulated by the inner triplet systems. Each main detector (ALICE, ATLAS, CMS, LHCb) has two inner

<sup>1</sup>other bunch filling schemes are available.

<sup>2</sup>since 2016 a new beam production technology Batch Compression Merging and Splitting (BCMS) was developed that reduced the beam size to 2.5 $\mu$ m.

triplet systems from both sides to squeeze the beams right before the interaction from 0.2 millimeters to 16 micrometers across.

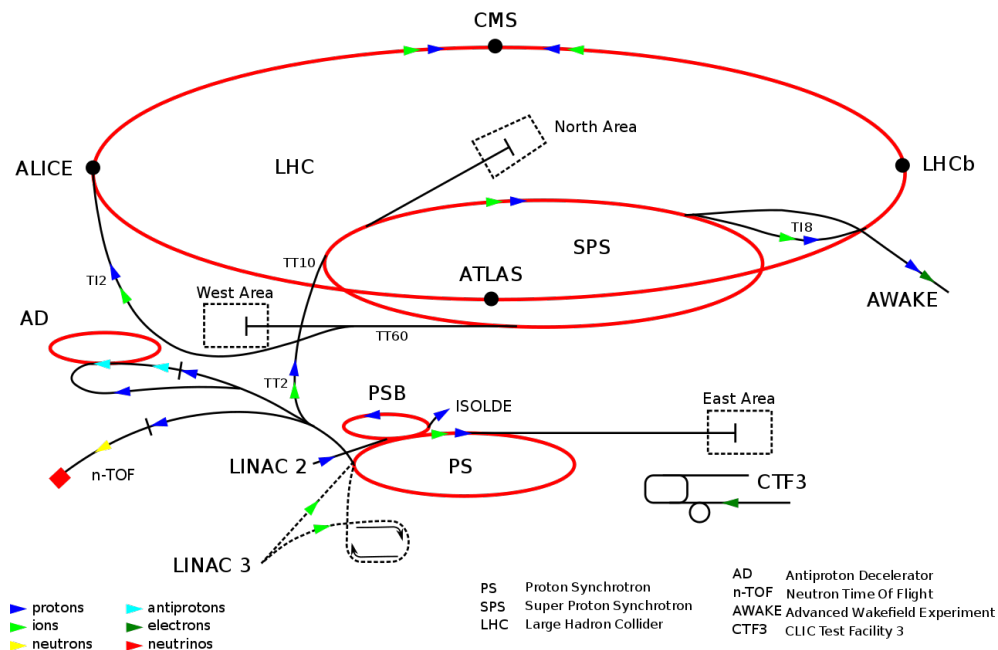


FIGURE 2.2: Acceleration scheme [50].

A tight packaging of bunches results in a small interval between the bunches collisions. Twenty-five nanosecond gap between the interactions forms a 40 MHz collision rate that creates a significant challenge to the detecting technologies and readout systems. After a collision, using dipole magnets, the beams are split back to be used in further collisions. At a certain point, the sufficient amount of protons diminishes to the threshold value, and the bunches are deflected from the LHC through the beam dump procedure.

The LHC goal is to provide a good quality interaction for the main experiments with the center of mass collision energy up to 14 TeV. Run 1 took place in 2009-2013; during that period, LHC maintained collisions at 7 and 8 TeV center of mass energy. The amount of data recorded during that period was sufficient to report the observation of the Higgs boson. A long shut down in 2013-2015 was necessary for accelerator and detectors to upgrade. The second operational run (Run 2) started in 2015 with the first-ever reached center of mass energy of 13 TeV and ended in late 2018. The results presented in this thesis are obtained after analyzing data collected with the CMS detector during the first period (2015) of Run 2. Since December 2018 and until 2021, the LHC and all detectors are on a long shutdown to replace few components. One of the goals of this period is to prepare for the High Luminosity LHC (HL-LHC), which is supposed to increase luminosity by a factor of 5-7. The next Run 3 is planned to start in 2021 and to last until 2024, while HL-LHC will start in 2027, in the present schedule.

A straightforward characteristic of the collisions is the number of collision events generated per second,  $R$ , given by the expression:

$$R = L \times \sigma_{\text{event}}, \quad (2.1)$$

where  $L$  is the machine luminosity and  $\sigma$  is the cross-section. The luminosity variable depends on the parameters of the beams and can be defined as:

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

where  $\epsilon_n$  is the normalized transverse beam emittance,  $\beta^*$  is the beta function at the collision point, and  $F$  corresponds to the geometric luminosity reduction factor due to the crossing angle at the interaction point.  $N_b$  is the number of particles in a bunch,  $\gamma_r$  the relativistic gamma factor,  $n_b$  is the number of bunches per beam,  $f_{\text{rev}}$  is the revolution frequency [51]. More details on luminosity measurements is given in chapter 4.

One of the main goals of the LHC is to deliver the designed luminosity  $L = 10^{34} \text{cm}^{-2} \text{s}^{-1}$  for the high luminosity experiments ATLAS and CMS. In September 2018, the LHC has reported in "The final days of Run 2" the successful delivery of designed luminosity for CMS and ATLAS [52]. The report was given a week before the shutdown, but already then it was clear that the machine achieved one of its main tasks. Figure 2.3 shows the integrated luminosity during Run 1 and Run 2.

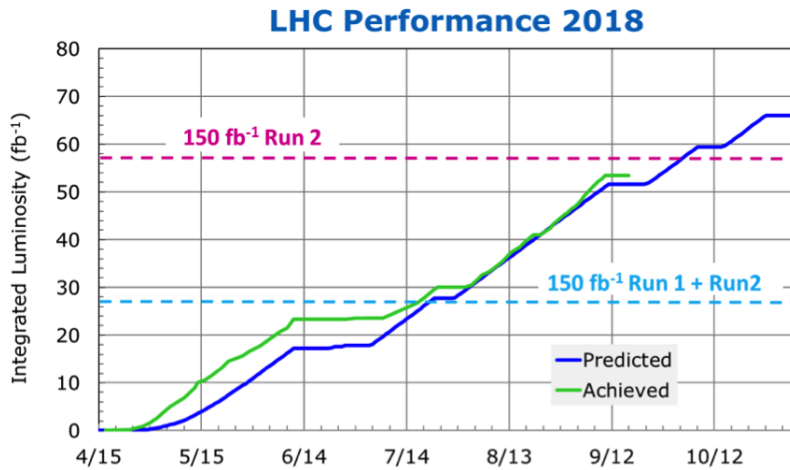


FIGURE 2.3: LHC report on the delivered luminosity during RUN 1 and RUN 2.

## 2.2 Compact Muon Solenoid

The CMS project approval in 1993 initiated a formation of the CMS collaboration that consists of over 4000 particle physicists, computer scientists, technicians, engineers, and students from near 200 universities and institutes from more than 40 countries. The objectives of the physics program are:

- Study for the excited state of the Higgs field (the Higgs Boson) discovered in 2012. The limit on the Higgs mass, derived from the Tevatron data analysis, predicted high branching ratio for the hadronic final state. The LHC has a high QCD background which reduces the quality of the results in hadronic channels. A good lepton detection system was a key point for this search.

- Search of supersymmetric particles like gluinos and squarks.
- New vector boson search (Z') involves identification of very high mass bosons that decay leptonically. Different models are expected to be probed using forward-backward asymmetry measurement that depends on good momentum resolution at high  $p_T$  and pseudorapidity coverage up to  $|\eta| < 2.4$ .
- Extra dimensions searches include different scenarios that could allow to determine the Hawking radiation temperature, number of extra dimensions, etc. Some of the models predict a graviton emission that escapes into extra dimension meaning a presence of missing transverse energy.
- The observation of standard model processes with higher precision level might lead to an indication of deviations from the predictions or reduce uncertainty level.
- Heavy-ion physics studies allow to shed a light on the mysterious and poorly understood quark-gluon matter.

The CMS detector is located in a cavern near Cessy in France, across the border from Geneva. It has a cylinder shape and weighs about 14,000 tonnes with sizes 21 meters long, and 15 m in diameter (Fig. 2.4). The detecting sub-systems are designed in a five-layer structure. The inner layer surrounds an interaction region located at the cylinder axis where the beam pipe is placed. This layer holds the CMS inner tracking detector. The second and third layers are electromagnetic and hadronic calorimeters. Layer four corresponds to a large solenoid magnet, and the muon system forms the fifth layer.

### 2.2.1 CMS coordinate system

The CMS coordinate system is a right-handed Cartesian coordinate system. The zero coordinates are placed in the center of the detector, which is considered as the nominal interaction point. The basis vector of the X-axis points towards the center of the LHC ring, the basis vector of the Y-axis is placed perpendicular to the LHC ring, and as the Z-axis basis vector the counterclockwise direction of the beams is used. The X-Y axes form a transverse plane where the measure of an angle is defined by the  $\varphi$  notation, and the radial component is denoted by  $r$ . The angle in Y-Z plane from the Z axis is defined as  $\Theta$ . The measure of distance between objects in the  $\eta - \varphi$  space is defined by the variable  $\Delta R$  as:

$$\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2}, \quad (2.3)$$

where  $\eta$  is pseudorapidity, defined in Eq.1.41,  $\Delta\eta = \eta_i - \eta_j$  and  $\Delta\varphi = \varphi_i - \varphi_j$  correspond to the difference between the  $\eta$  and  $\varphi$  coordinates of the objects. This variable is used in this thesis as a discriminant variable for the isolation of the muons (see, for example, section 5.3).

The particle transverse momentum,  $p_T$ , is defined as:

$$p_T \equiv \sqrt{p_x^2 + p_y^2}, \quad (2.4)$$

where  $p_x$  and  $p_y$  are the particle momentum components in X and Y planes.

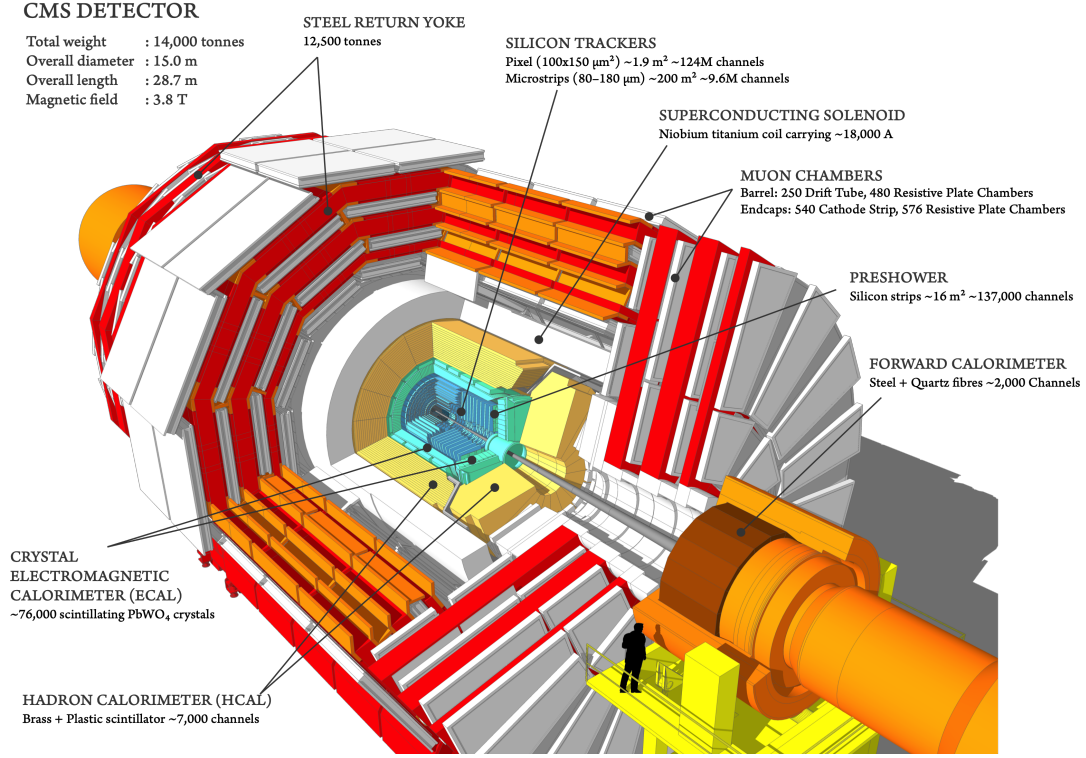


FIGURE 2.4: The figure shows the CMS detector, its main components, and characteristics. A black human figure on the right corner below is made of the real scale to the detector size [53].

In high energy collisions, the initial transverse components of the colliding protons can be neglected. This feature allows to define a transverse momentum conservation law – the vector sum of the transverse momenta of all particles in initial and final states are considered to be equal zero:

$$\sum_{i=1}^{\text{all}} \vec{p}_{T_i} = 0. \quad (2.5)$$

This feature finds its reflection in numerous techniques and is directly used in this thesis for the missing transverse energy estimation:

$$E_T^{\text{miss}} \equiv - \left| \sum_{i=1}^{\text{rec}} \vec{p}_{T_i} \right|, \quad (2.6)$$

where the vectorial sum is performed over all reconstructed particles in the event. The missing transverse energy is a lack of detected energy, which is carried away by the neutrinos, produced in weak decays.

The expression nominal interaction point (or nominal primary vertex) is often referred to as a hypothetical center of the detector on the beam-line where a collision is taking place. Practically, collisions have a pileup - multiple proton-proton collisions per bunch crossing, including hard scattering, elastic scattering, and diffractive processes. Depending on the origin of the interaction in terms of bunch crossings, the pileup events can be in-time - originate from the same bunch crossing as an event under consideration and out-of-time - pileup events from previous bunch crossing.

### 2.2.2 Physical processes behind detection techniques

The next sections present an overview of the CMS layers and its main components. Each of them has its own goals and unique, custom design and technology. However, physics processes and detection principles are conventional for each of them. All<sup>3</sup> particles created in pp collisions lose their energy moving through the detector materials. For charged particles, these processes are caused by the interaction of the charged particle electromagnetic field with the EM fields of electrons and nucleus in the matter. The energy loss may happen through atoms excitement, which returns to the ground state emitting a photon with characteristic energy. This process is called scintillation, and detectors that use this effect as a detecting principle are called scintillation detectors. Some detector uses bremsstrahlung as a detection principle. In bremsstrahlung, the decelerated charged particle loses its kinetic energy emitting photons. Deceleration mainly happens due to a charged particle (usually electron) deflection in the electric field of a nucleus. Another way<sup>4</sup> of energy loss is ionization. When the energy, transmitted to an electron, is high enough to overcome (or penetrate through) a potential barrier of an atom, a free electron and ion can appear. These processes are of a subatomic scale, and hence are being invariant from a phase state of a detecting matter. High energy photons interact with matter creating  $e^+e^-$  pairs. The general idea behind particle registration is to collect products of interaction as effective as possible. The produced electric charge signal is assembled with a readout electronics.

### 2.2.3 Tracker Detector

The Tracker Detector (TD) can be called the heart of the CMS since it is placed closest to the interaction point, where the charged particle flux has its highest value ( $\approx 10^7$  per second). One of the main objectives of this detector includes detection of decay vertices and particle path (track). According to the Lorentz force, a charged particle moving perpendicular to the magnetic field will have a curved trajectory. The information about the trajectory is used to calculate transverse momentum and the sign of the particle. Particles with a high curvature of a track have a small momentum and vice-versa. A high granularity of the TD allows track reconstruction with a high level of precision and efficiency.

The TD detector has a total length of 5.8 m and 2.5 m in diameter [54]. A detailed sketch of the TD is shown in Figure 2.5. Depending on the distance from the interaction point (hence - particle flux), the detector is divided into three parts.

The first part is a pixel detector. It has a radius of 10 cm and is made of pixel sensors with a size  $\approx 100 \times 150 \mu\text{m}^2$  each. During pixel sensor ionization, charged particle creates electron-hole pairs; then the electrons are collected with an electric current from each pixel. Pixel sensors are mounted on three barrel layers with 768 sensors in total and two endcap disks on each side of them. The radii of the barrel layers are of 4.4 cm, 7.3 cm, and 10.2 cm, the length of each is 53 cm [55]. Both endcap disks with inner, and outer radii 6 cm and 15 cm, respectively, are located at a distance  $|z| = 34.5$  cm and 46.5 cm, respectively, containing 672 pixel sensors in total. The spatial resolution is near  $10 \mu\text{m}$  in the  $r$ - $\phi$  plane and  $20 \mu\text{m}$  in the  $z$  plane. The pixel read-out system consists of almost 16,000 chips.

<sup>3</sup>Except hypothetical particles and neutrinos.

<sup>4</sup>Cherenkov radiation is not reviewed.

The second part of TD is an intermediate region with radius  $20 < r < 55$  cm. The Tracker Inner Barrel (TIB) consists of four layers of silicon and strip pitch sensors with total length  $|z| < 65$  cm [55]. The first two layers are made with stereo modules (shown in blue Fig. 2.5) allowing measurements in the  $r$ - $\phi$  and the  $r$ - $z$  planes. The Tracker Inner Disks (TID) are located at each side of the detector. They are arranged in three rings, two of which are supplied with stereo modules.

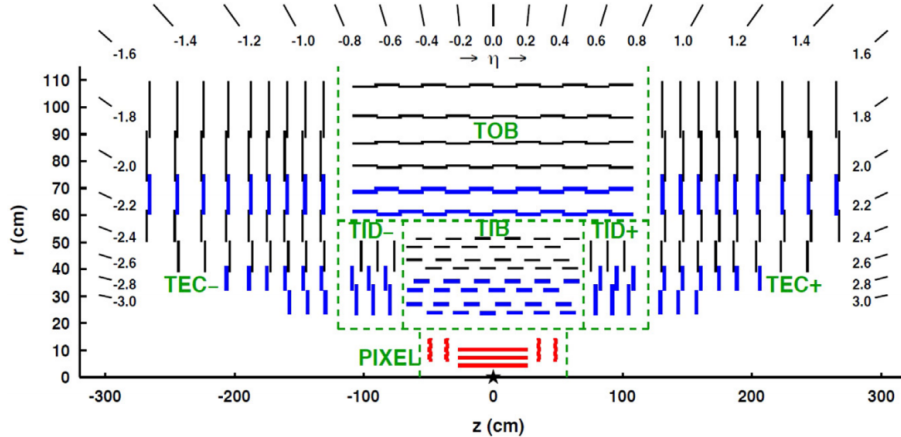


FIGURE 2.5: The CMS tracker slice view in the  $r$ - $z$  plane. Single-sided strip modules are marked with black thin lines, pixel modules are depicted in red, and strip stereo modules are shown as blue thick lines. [55]

The third, outermost region ( $r > 55$  cm) holds larger-pitch silicon microstrip sensors with maximum cell size of  $25 \text{ cm} \times 180 \mu\text{m}$  [54]. The Tracking Outer Barrel (TOB) is 220 cm ( $|z| < 110$  cm) in length and consists of six layers of silicon sensors. The first two layers are made in such a way to be able to measure in  $r$ - $\phi$  and  $r$ - $z$  coordinates, they are shown in Figure 2.5 in TOB section with blue lines. The Tracker End Cap (TEC) is placed on both sides of the TOB and TID. Each of them consists of 9 disks, that propagate from 120 cm to 280 cm in  $|z|$ . The thickness of sensors in the innermost rings is  $320 \mu\text{m}$ , for the rest of TEC the thickness is  $500 \mu\text{m}$ .

## 2.2.4 Calorimeters

Measurement of the energy of the particles produced in  $pp$  collisions is one of the most important goals of the CMS detector. Efficient reconstruction of the energy deposits, from various particles produced in the event can help to estimate the missing energy contribution, which is vital for searches of the new particles predicted by models of supersymmetry and other processes.

In order to have a high radiation resistance and a proper resolution, the CMS collaboration uses solid body sensors in the calorimeters. When a particle is moving through the sensor, it loses its energy creating a shower of particles (mainly photons and electrons). Additionally to a particle energy measurement, sensors can be used to identify the type of a particle through the shower shape recognition. The main characteristics of the shower depends on the particle type, nature of the interaction, and sensor material. The showers are often categorized by three parameters: Moliere radius, nuclear interaction length,  $\lambda_I$ , and radiation length,  $X_0$ . The Moliere radius characterizes the width of the shower and is defined as a radius of a hypothetical cylinder that contains near 90%



of the shower's energy. The nuclear interaction length,  $\lambda_I$ , describes the mean distance that hadron travels before an inelastic nuclear interaction occurs. The radiation length characterizes electromagnetic showers and represents an average distance during which the moving electron loses energy, by a factor of  $\frac{1}{e}$ .

For the purpose of high resolution and detector efficiency the calorimeter system is divided in two main sub-components. Depending on the type of the detecting particles, there are Electromagnetic Calorimeter (ECAL) and Hadron Calorimeter (HCAL). Both of them are placed between the Superconducting Magnet and the Tracker Detector forming two barrel and three endcap layers (a brief overview of the ECAL and HCAL systems are given in the next two paragraphs).

**Electromagnetic calorimeter.** The ECAL measures energies of particles that predominantly lose their energy through the electromagnetic interaction. In pp collisions at high energy, these particles are primarily photons, electrons, and positrons.

The ECAL is a hermetic, homogeneous detector made of scintillating crystals of lead tungstate ( $\text{PbWO}_4$ ). Homogeneous calorimeters use the material of the sensor for the deposited energy measurement. The detector consists of three main parts: Barrel ECAL (EB), Endcap ECAL (EE), and Pre shower (ES).

The EB is placed in the barrel region and covers a pseudorapidity interval of  $0 < |\eta| < 1.479$ , the inner radius of EB section is 129 cm [54]. The ECAL sketch in the Y-Z plane is given in the right of Fig. 2.6. The  $\text{PbWO}_4$  crystals have a length of 230 mm and a front face cross-section of near  $22 \times 22 \text{ mm}^2$ . The crystals are tilted at  $3^\circ$  with respect to the line from the nominal interaction point position. They are arranged in a way that allows each of them to cover  $\approx 1^{02}$  of a solid angle (approximately 0.0174 in  $\Delta\phi$  and  $\Delta\eta$ ).

The endcaps are placed at a distance of 314 cm away from the interaction point, covering the pseudorapidity range  $1.479 < |\eta| < 3.0$ . The crystals are arranged in an x-y grid and have a front cross-section of  $28.6 \times 28.6 \text{ mm}^2$  and a length near 220 mm.

The preshower part is placed in front of the endcaps and covers  $1.653 < |\eta| < 2.6$ . It consists of two planes of lead radiator/silicon strip sensors. The ES allows improving the identification of the neutral pions and the position resolution of the detected particles.

The  $\text{PbWO}_4$  crystals were chosen because of their well-suited properties. The crystals have a good Moliere radius (2.2 cm), short radiation length (0.89 cm), and a high density ( $8.3 \text{ g/cm}^3$ ). An example of such crystal is shown in the left image of Figure 2.6. The energy resolution is measured through the fitting procedure of a Gaussian model to the reconstructed energy distribution. The function is parametrized as a function of energy:

$$\frac{\sigma}{E} = \sqrt{\left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2}, \quad (2.7)$$

where S corresponds to a stochastic term, N is the noise and C is the constant term. The values of S, N, and C terms may vary depending on corrections applied and other conditions of the event selection. In CMS terms are estimated to be  $S \approx 2.8\%$ ,  $N \approx 12\%$ ,  $C \approx 0.3\%$  [54].

**Hadron Calorimeter.** The HCAL measures the energy of particles (hadronic jets) that predominantly interact via the strong interaction. The HCAL is also used for the missing transverse energy estimation. In order to do so, the HCAL is required to have a suitable containment, energy resolution, and hermeticity [58]. The HCAL is located

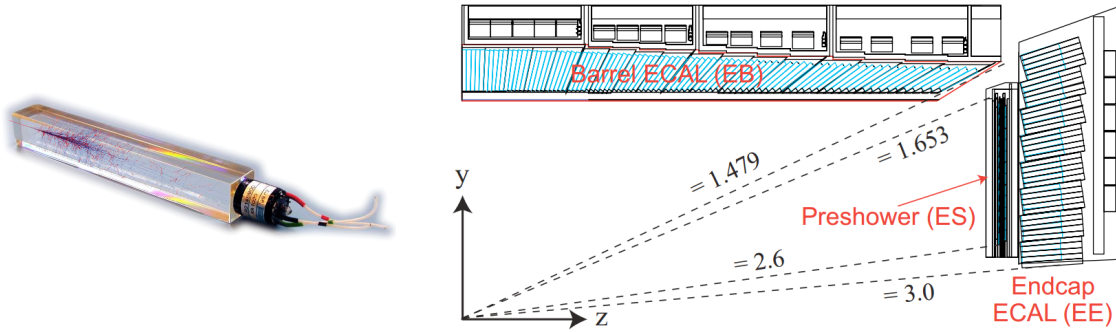


FIGURE 2.6: The left figure illustrates a  $\text{PbWO}_4$  crystal with a graphical representation of a shower inside [56]. The right figure represents the sketch of the one-fourth of the ECAL in the Y-Z plane. Tilted  $\text{PbWO}_4$  crystals are shown with the blue lines. The dashed lines show the pseudorapidity coverage of the EB, EE, and ES [57].

inside the magnetic coil between the ECAL and the magnet. Space limitations together with the magnet parameters and the design requirements define the detection principles of the HCAL and its material characteristics. To satisfy such demands, the detector is divided into active zones (detecting parts based on plastic scintillator technology) and the passive zones (the absorber material). Brass is used as an absorber material due to characteristics that perfectly fits the purpose - it is easy to produce, has a relatively short length of nuclear interaction, and satisfies necessary anti-ferromagnetic demands. As for the active part, the scintillator tiles readout with embedded wavelength-shifting (WLS) fibers are used [59].

The HCAL consists of four main parts: hadron barrel (HB), hadron outer (HO), hadron endcap (HE), and hadron forward (HF) calorimeter (Figure 2.7). The readout system consists of cells that are arranged into towers projecting to the nominal interaction point in  $\eta - \phi$  space. The size of the cells is different for different HCAL parts. In total the hadron calorimeter has 4176 towers.

The hadron barrel (HB) covers a pseudorapidity region of  $-1.4 < \eta < +1.4$  and consists of 2304 towers, with a size  $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$  each [59]. The thickness of the first scintillating layer after ECAL is 9 mm, while the rest of the detecting layers are 3.7 mm. The HB has 15 plates of absorbing material (brass) with a thickness of 5 cm each. Very high energy hadrons can not be absorbed in the brass and scintillators of the HCAL. Therefore, the HCAL is extended outside the coil to the hadron outer calorimeter (HO). The HO is located between the magnet and the muon system and divided into five sections. In Figure 2.7 these sections are shown as rings and numbered as -2,-1,0,1,2. The HCAL-HO consists of two layers of a scintillator that are placed before and after the first layer of the return yoke (absorber), which is shown as IRON in Figure 2.7.

The endcap hadron calorimeter (HE) is mounted from both sides of the HB and covers a pseudorapidity region of  $1.3 < |\eta| < 3.0$  [59]. The HE has a good hermiticity, moderate energy resolution, and a good transverse granularity, which matches the EM granularity. The total depth of the HE is ten absorption lengths.

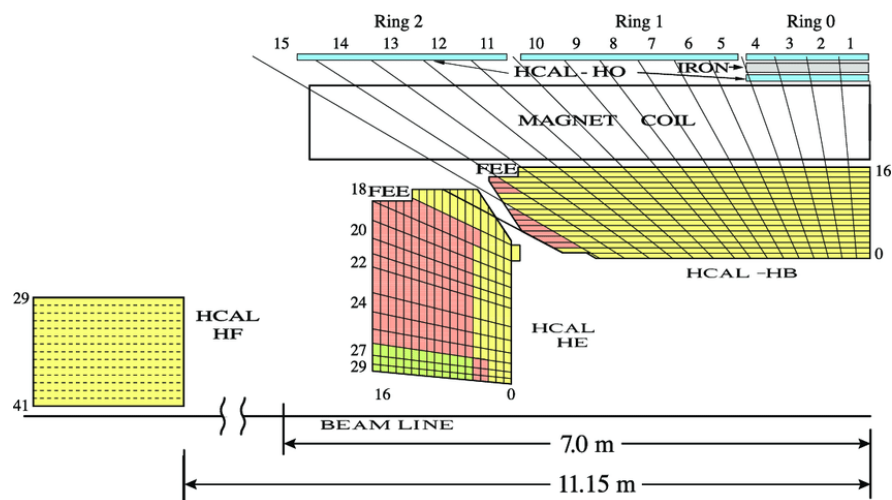


FIGURE 2.7: The figure represents the sketch of the quarter slice of the HCAL in the Y-Z plane. "FEE" corresponds to the locations of the Front End Electronics for HB and HE. The signals from the tower cells shown with the same color are added optically, providing a "longitudinal" segmentation of the HCAL. HB and HE are sliced in 29 azimuthal pieces with  $\Delta\varphi \approx 20$  degrees each [60].

The absorber (brass) is bolted in one solid structure with 19 gaps filled with scintillator trays. The signal from the scintillator is transferred through optical cables to decoder boxes in which they are optically mixed together, forming a single tower. The very forward hadron calorimeter (HF) is located at a distance  $\pm 11.1$  m from the nominal primary vertex and covers a pseudorapidity range of  $3.0 < |\eta| < 5.0$ . The high- $\eta$  region suffers from a severe particle flux and as a result, high radiation field, neutron fluxes, and material activation. For that reason, the detector is built using a quartz fiber calorimetry technique, which is based on Cherenkov radiation effect [61], [62], [63]. The HF is built as a block of copper absorber with embedded quartz fibers, mounted in parallel to the beam axis. Each hadron forward calorimeter is divided into 12 pieces in  $\eta$  and 36 pieces in  $\varphi$ . In total, each side has 432 towers, which corresponds to physical channels. Depending on its sensitivity to a particle type, the fibers are separated into two categories S and L. The S fibers are sensitive mainly to hadrons while the L fibers are sensitive to electrons, photons, and hadrons.

### 2.2.5 Superconducting Magnet

The current method of charged particle momentum determination relies on the Lorentz force, which requires the presence of a strong magnetic field. Therefore, a magnet, as a source of a magnetic field, is a vital part of the CMS detector. The configuration of the magnet was chosen to be a solenoid. In solenoid, it is easier to obtain a high magnetic field, uniformly distributed near the center. Comparing to the toroidal magnet, the magnetic field of a solenoid is parallel to the beams, in this case, the curvature of a muon track is in the transverse plane. Another advantage is that the momentum measurement in solenoids starts at  $r = 0$ , while for a toroidal magnet, it starts after absorber,  $r > 4$  m. The magnet size is relatively small for a magnetic field of 3.8 T, it is 13 m long with an inner diameter of 5.9 m [58]. The nominal current is

near 19.500 A with magnet inductance approximately 14 H, stored energy is near 2.6 GJ. Such a powerful field is achieved using superconducting technologies, which makes this magnet to be one of the most potent artificial magnets on Earth.

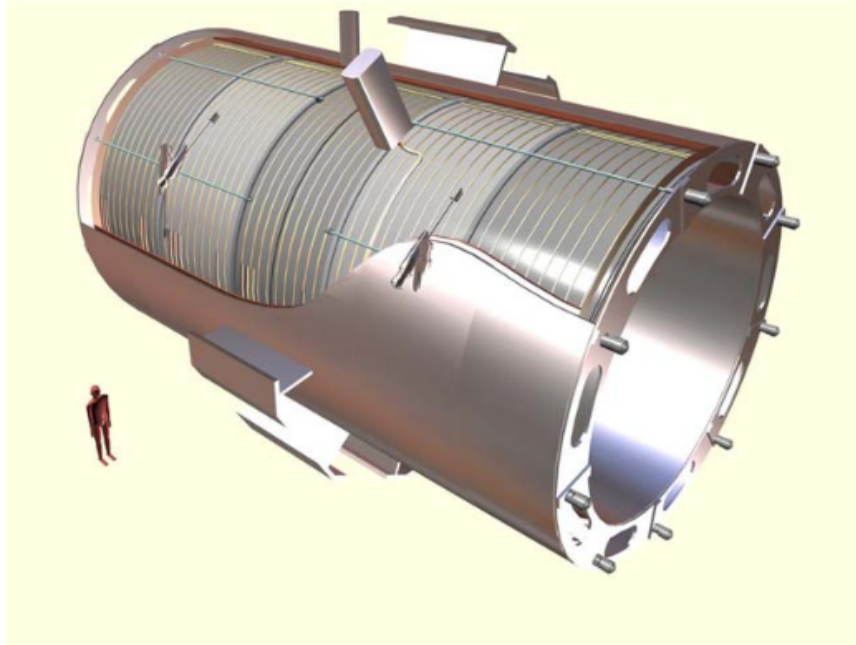


FIGURE 2.8: Representation of the 5 modules composing the cold mass inside the cryostat, with the supporting system (vertical, radial and longitudinal tie rods) [54].

The main parts of the magnet include a superconducting coil (cold mass) and magnet yoke (Barrel and End Cap). The CMS magnet is shown in Figure 2.8. The coil has four layers of winding made of a stabilized reinforced NbTi conductor. The operational temperature is 4.6 K obtained using a liquid helium cryostat. The magnetic flux produced by the coil is returned using a 1.5 m thick, saturated iron yoke, separated on the barrel and endcap parts. The barrel yoke is an 11 m twelve-sided cylinder that surrounds the coil and consists of 5 rings, 2.5 m each. Each of the rings has three iron layer-rings with the thickness of the inner part 295 mm, and 630 mm for the rest. The two inner endcap disks are 600 mm thick while the outer disks are only 250 mm thick, each endcap weighs 2300 tonnes [64].

### 2.2.6 Muon Detectors

The lepton identification and reconstruction is of a high priority for the CMS experiment. In particular, the muon detection is of special interest. The muon system has three main functions: identification of the muon, its momentum measurements, and triggering. The muon detecting principles are built around the gaseous particle detector technology. The detectors are located on the outermost layer of CMS and consist of three types of gas ionization chambers: Cathod Strip Chambers (CSC), Drift Tube Chambers (DTC), and Resistive Plate Chambers (RPC). The detailed quarter view of the muon system is shown in Figure 2.9.

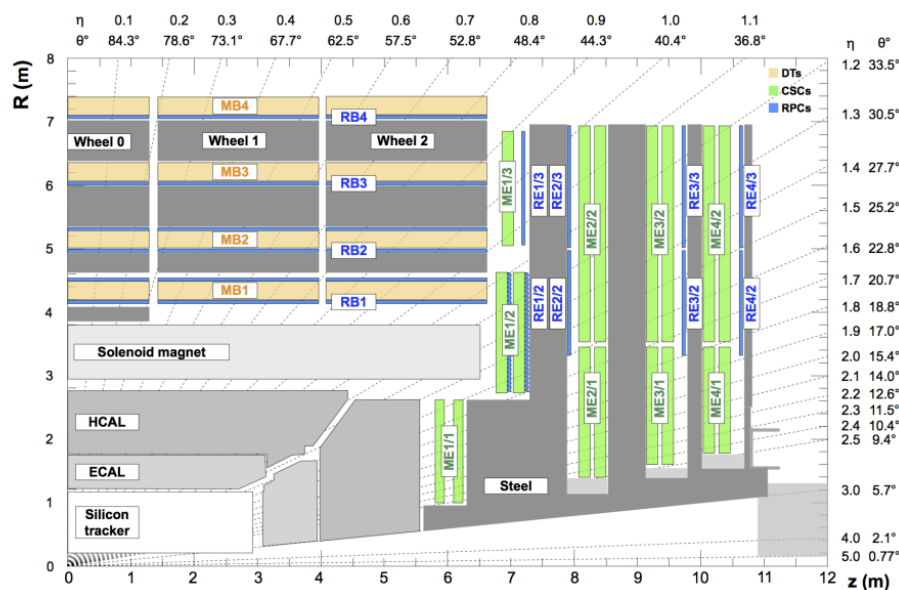


FIGURE 2.9: CMS muon system in the Y-Z plane. The rings of the barrel yoke 0, 1, 2 are shown as "Wheel", the endcap yoke is shown as "Steel". The Drift Tube stations (DTs) are shown with a beige color and labeled as MB ("Muon Barrel"). The color of Cathode Strip Chambers (CSCs) is lime-green; they are labeled as ME ("Muon Endcap"). Resistive Plate Chambers (RPCs) are mounted in both the barrel and endcaps of CMS, they are labeled RB and RE, respectively, and represented with a blue color. The dotted lines show the pseudorapidity coverage of the muon system with equidistant steps of 0.1.[65]

The DT chambers are located in a barrel region and cover the pseudorapidity range of  $|\eta| < 1.2$ . The tubes are installed between the layers of the return yoke and the magnet (except MB4 and RB4, Figure 2.9). The first three layers allow measurements in the  $r\phi$  plane, while the last one only in the  $z$ -direction. The elementary detecting part of the DT is the drift cell filled with the gas mixture 85%/15% of Ar/CO<sub>2</sub> [54]. The cell transverse size is  $42 \times 13 \text{ mm}^2$  with a high voltage (3600 V) gold-plated stainless steel anode wire in the center. The maximum drift time is about 400 ns.

The endcap region has higher background levels and non-uniform magnetic field. These conditions require usage of the system which has a fast-response and tolerance to the non-uniform magnetic field. The CSCs cover the pseudorapidity region of  $0.9 < |\eta| < 2.4$ . Each endcap consists of six layers of rings with radial cathode strip placement, which measure the muon position in the  $r-\phi$  plane. All chambers use a gas mixture of 10% CF<sub>4</sub>, 40% Ar, and 50% CO<sub>2</sub>. [65]

Besides the DTs and CSCs tracking detectors, the muon system has a special detector for triggering which is able to measure the beam crossing time at highest rates. The Resistive Plate Chambers (RPC) are mounted in the barrel and endcap regions, covering the pseudorapidity range of  $|\eta| < 1.6$ . The RPCs consists of double-gap chambers, each gap consists of two resistive plate (2 mm thick) and separated by a gas gap (2 mm thick). The first half of a gas mixture consists of 95.2% freon (C<sub>2</sub>H<sub>2</sub>F<sub>4</sub>), 4.5% isobutane (i-C<sub>4</sub>H<sub>10</sub>), and 0.3% sulfur hexafluoride (SF<sub>6</sub>). The second half is a water vapor. The operating voltage is about 9.6 kV [66].

### 2.2.7 Triggering and Data Acquisition

The CMS trigger system is the leading and final stage of data quality control. During the LHC data harvest<sup>5</sup>, pileup level is different for different acceleration parameters. In 2015 it was in the range of 10-20 collisions per bunch crossing, which resulted in a GHz interaction rate [67]. The main objective of the trigger systems is to select valid events that are of scientific interest. The CMS trigger system consists of the L1 (Level one) trigger system, which is compounded of custom-designed hardware, and HLT (High-Level Trigger) array of commercially available computers that run special selection programs, close to the offline reconstruction and selection.

The L1 trigger selects events that contains detector signals identified as a physics object (photon, leptons, jets,  $E_T^{\text{miss}}$  etc.). The Global Trigger, GT, collects and processes the information from the calorimeters and muon detectors to implement a menu of triggers. The trigger menu is a list of selection requirements applied to the final set of identified objects, demanded by the specifications of the HLT algorithms. In order to limit the output rate to 100 kHz, the trigger thresholds are adjusted to the instantaneous luminosity of LHC.

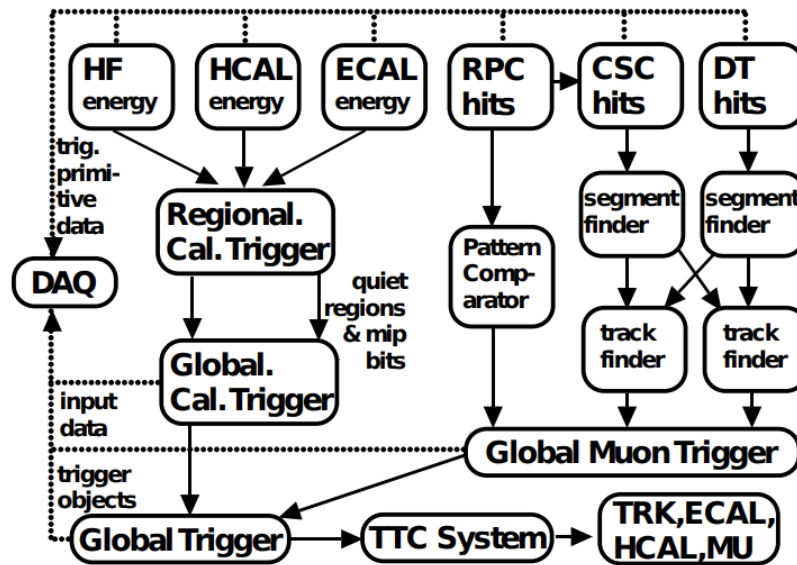


FIGURE 2.10: The scheme represents the CMS L1 trigger system. The data from the electromagnetic calorimeter (ECAL), hadronic calorimeters (HF and HCAL) are processed first to regional calorimeter trigger (RCT), then to global calorimeter trigger (GCT). Energy deposits (hits) from the muon detector (RPC, CSC, and DT) are processed via a system of a track- and segment-finders or via a pattern comparator, and sent to a global muon trigger (GMT). The final trigger decision is made using a global trigger (GT), which combines the information from the GCT and GMT. The decision is then sent to the tracker (TRK), ECAL, HCAL, or muon systems (MU) using the Trigger, Timing and Control system (TTC). Afterwards, the data from various subsystems is transferred to the acquisition system (DAQ) for offline storage.[68]

<sup>5</sup>CERN slang.

The general scheme of the L1 trigger stages is represented in Figure 2.10. The HLT hardware consists of a processor farm that consists of commodity computers. The trigger algorithms are implemented in software, that reduces an output stream to an average rate of 100 Hz for offline event storage. The HLT algorithms sort out the events, applying reconstruction and identification criteria to each object. The analog of the trigger menu in the HLT trigger is called HLT path. It has more sophisticated algorithms that simultaneously perform physics object reconstruction and selection of these objects. The selection criteria include various kinematic thresholds ( $p_T$ ,  $\eta$ ,  $E_T^{\text{miss}}$ ) as much as a specific detection condition (quality of a fit, number of matched stations, etc.). Depending on the signatures of a specific process of interest, different HLT path schemes may be used. To reduce the amount of data to be stored, trigger paths with high event rates may be prescaled by an individual factor  $X$ . This factor means that only one event out of  $X$  will be recorded.

### 2.2.8 Acceptance

The detector acceptance,  $\mathbb{A}$  - is a measure of the detector's performance, defined as the fraction of events that are detected, reconstructed, and passed selection requirements,  $N^{\text{pass}}$ , to a total number of all generated events,  $N^{\text{total}}$ , Formula 2.8. The acceptance depends on the detector phase space.

$$\mathbb{A} = \frac{N^{\text{pass}}}{N^{\text{total}}} \quad (2.8)$$

The detector phase space embodies characteristics of the detector (solid angle of the detector coverage, detecting volume, etc.), detector efficiency, reconstruction algorithms efficiencies, features of sub-detector alignment, and many others. The full phase space may be defined<sup>6</sup> as the full solid angle of events production within a corresponding volume. The detector acceptance in full phase space is estimated using theory predictions. The acceptance is calculated as the fraction of generated events in the detector phase space that pass the selection requirements,  $N_{\text{det}}^{\text{pass}}$ , over the number of generated events in the full phase space,  $N_{\text{full}}^{\text{gen}}$ , Formula 2.9.

$$\mathbb{A} = \frac{N_{\text{det}}^{\text{pass}}}{N_{\text{full}}^{\text{gen}}}. \quad (2.9)$$

CMS experiment has an extensive physics program so that the detector phase space may be different for different studies. In other words, depending on the properties of studied physics objects, the detector phase space can be customized, becoming a fiducial phase space. The fiducial phase space may be defined as a phase-space with a set of specific requirements of geometric, kinematic, and reconstruction nature. For example, the calculation of differential cross-section values in  $W^\pm$  asymmetry studies requires estimation of the fiducial acceptance in each pseudorapidity bin. Here, the additional constraints are the pseudorapidity cuts, which define the geometric boundaries of the detector phase space for each  $\eta$  bin. More details on acceptance definition for the  $W^\pm$  boson charge asymmetry studies is given in section 5.7.

---

<sup>6</sup>In the context of CMS experiment.





## Chapter 3

# Event reconstruction

The reconstruction of  $W^\pm$  boson candidate events includes recognition and interpretation of the detector signals in terms of physics objects. The products of  $W^\pm$  boson decay in lepton channel are highly energetic muon and muon neutrino, which is indirectly detected as missing energy,  $E_T^{\text{miss}}$ . The muon reconstruction procedure combines information from several detector systems and processes it via specific algorithms. The  $E_T^{\text{miss}}$  reconstruction procedure relies on sophisticated methods that combine the information of all reconstructed particles in the event.

The next chapter describes methods that were used to reconstruct the event candidates. First, an introduction to the algorithms that were used to identify physics objects is given. The next part briefly explains the vertex reconstruction algorithms. Afterward, the main steps in muon and missing transverse energy reconstruction is presented.

### 3.1 Particle Flow algorithm

The Particle Flow algorithm, (PF) [69] is a "link" between detector signals and physics objects used in the analysis. The main goal of PF is to reconstruct and identify all particles in an event. In this sense, particles are referred to as photons, electrons, muons, charged, and neutral hadrons. In particular, PF aims to determine a particle type of each PF candidate, the direction of its propagation, and its energy. To do that, it combines all information of the event from all CMS sub-detectors.

The general PF algorithm has three main stages. In the first stage, information from sub-detectors is analyzed to create elementary objects of the algorithm, so-called "elements". They typically include charged-particle tracks, muon tracks, and calorimeter clusters. The tracks are reconstructed during five iterations of iterative-tracking strategy [70]. Calorimeter clusters are identified using the "Clustering Algorithm" [71]. This algorithm is performed individually in ECAL barrel, ECAL endcap, HCAL barrel, HCAL endcap, first PS layer, and second PS layer. The algorithm reconstructs the direction and energy of stable neutral particles, separates neutral particles energy deposits from charged hadron deposits, measures the energy of charged hadrons, identifies and reconstructs electron energy deposits [69].

The second stage is a link algorithm. Reconstructed "elements" are linked to each other, forming "blocks". An essential part of this stage is to avoid double-counting. Depending on a variety of elements in the event, the linking schemes may be different.

In the last stage, the blocks are used to reconstruct and identify the particle type and its properties. In each block, the algorithm is executed in a particular order. First, the algorithm reconstructs the muon candidates. After reconstruction, a corresponding PF element is removed from the block. Next, the electron reconstruction is done; it also

collects the bremsstrahlung photons. In this step, energetic and isolated photons are identified as well. The rest of the elements in the block are then processed through the cross-identification routine to identify charged hadrons, neutral hadrons, and photons. At the same stage, secondary particles from nuclear interactions are reconstructed as well.

The schematic representation of PF algorithm performance is shown in Figure 3.1

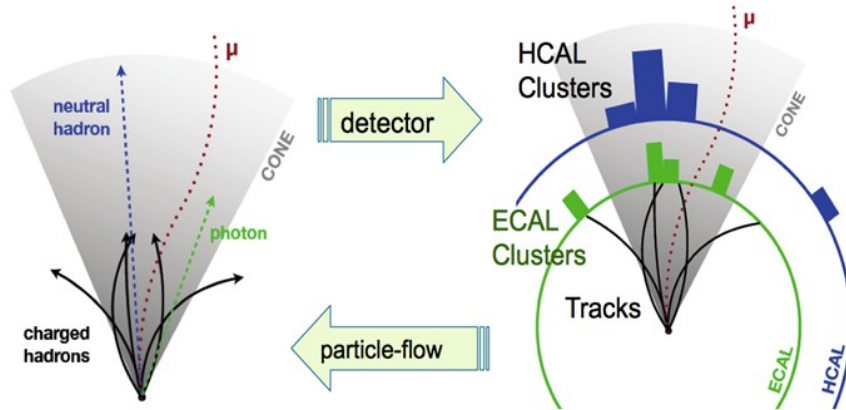


FIGURE 3.1: Schematic view of Particle Flow algorithm as a link between detector signals and reconstructed physics objects. The left side illustrates an event representation in terms of physics objects. The right part represents the CMS detector in the transverse plane; ECAL clusters represent an amount of energy detected from leptons and photons in the hypothetical event. The HCAL cluster shows amount of detected energy from hadrons. Tracks are shown with black lines. The muon that penetrates all layers is shown with the red dotted line [72].

## 3.2 Vertex reconstruction

Every event starts from an interaction point - a region in time-space, which can be identified only with a certain, limited, precision in which partons interact with each other. The pixel detector, in combination with the tracker detector and multiple algorithms, provides a vertex reconstruction with high efficiency. The restored location of the particle's interaction point is called the primary vertex<sup>1</sup>.

The algorithm that is used to reconstruct  $W^\pm$  boson vertices is called the Deterministic Annealing clustering algorithm (DA) [73]. The algorithm uses such methods as pattern recognition and classification, compression, and statistical regression. A vital feature of this algorithm is the ability to identify primary vertex in noisy conditions without previous knowledge of the number of vertices to be found. The first step of the algorithm is to calculate the apex points using PF reconstructed tracks. At this stage, DA uses the Apex Point clustering approach with the Minimum Two Values algorithm [74]. Afterward, the iterative procedure of DA is applied to the set of apex points. In each iteration, a created prototype of a vertex seed is associated with the apex points through a weight factor. In the next stage, a set of associated apex points are replaced by their corresponding tracks. After that, each prototype becomes a vertex seed. In the last step, the fit is performed on all vertex seeds.

<sup>1</sup>For the  $W^\pm$  boson production.

### 3.3 Pile-up Per Particle Identification

One of the advantages of the 2015 data set is that the pile-up level for that run period was relatively small, compared to other run periods. During the 2015 run period, the average pile-up value was equal to thirteen. The average values of pile-up for run II are shown in Figure 3.2.

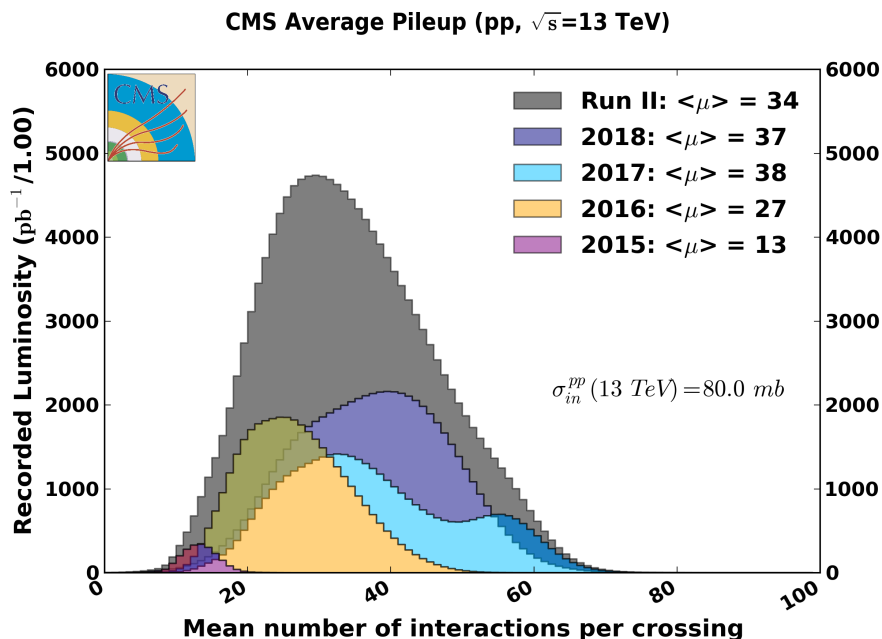


FIGURE 3.2: Interactions per crossing (pile-up) for 2015-2018 [75]

To solve the problem of high pile-up, a new method for pile-up mitigation was proposed – "pile-up per particle identification" (PUPPI) [76]. In this method, every particle in an event matches an  $\alpha$ -shape variable that combines event pile-up properties information, tracking information, and local shape values. The  $\alpha$ -shape variable for the  $i$ -th particle is defined as shown in Formula 3.1.

$$\alpha_i = \log \sum_{j \in \text{event}} \zeta_{ij} \times \Theta(R_{\min} \leq \Delta R_{ij} \leq R_0), \quad (3.1)$$

$$\zeta_{ij} = \frac{p_{Tj}}{\Delta R_{ij}}.$$

In this formula, the  $j$ -th particle belongs to a considered event. Here  $\Theta(R_{\min} \leq \Delta R_{ij} \leq R_0)$  is a shorthand notation of  $\Theta(\Delta R_{ij} - R_{\min}) \times \Theta(R_0 - \Delta R_{ij})$ , and  $\Theta$  is the Heaviside step function,  $\Delta R_{ij}$  is the distance between particles  $i$  and  $j$  in  $\eta - \phi$  space,  $p_{Tj}$  is the transverse momentum of particle  $j$ , and  $R_0$  is a cone defined in Formula 2.3. Particles in the cone around  $i$ -th particle are taken into calculation of  $\alpha_i$ . Any particle which is closer to  $i$ -th particle more than  $R_{\min}$  is discarded from the calculation. Default values are  $R_{\min} = 0.02$  and  $R_0 = 0.3$ .

Conditions on this variable allow to separate parton shower-like radiation from pile-up-like radiation. The event-level characterization of the charged pile-up distribution is obtained using the root mean square (RMS), and median values of the  $\alpha$  values. Each

particle is tagged with a weight by comparing its  $\alpha$  value to the median of the charged pile-up distribution. The weight means the probability of a particle to occur from the non-pile-up event. Afterward, the weight is used to re-scale the transverse momentum to discard particles with a very small  $p_T$ . The association of the particle to pile-up or leading vertex is obtained using  $p_T$  spectrum property – being the pile-up  $p_T$  spectrum falling much faster than the leading vertex  $p_T$  spectrum. After that, the set of pile-up corrected particles may be used for the missing transverse energy calculation or in the jet algorithm. The performance of PUPPI algorithm is shown in Figure 3.3. The  $E_T^{\text{miss}}$  estimation using PUPPI methods will be discussed in section 3.5.

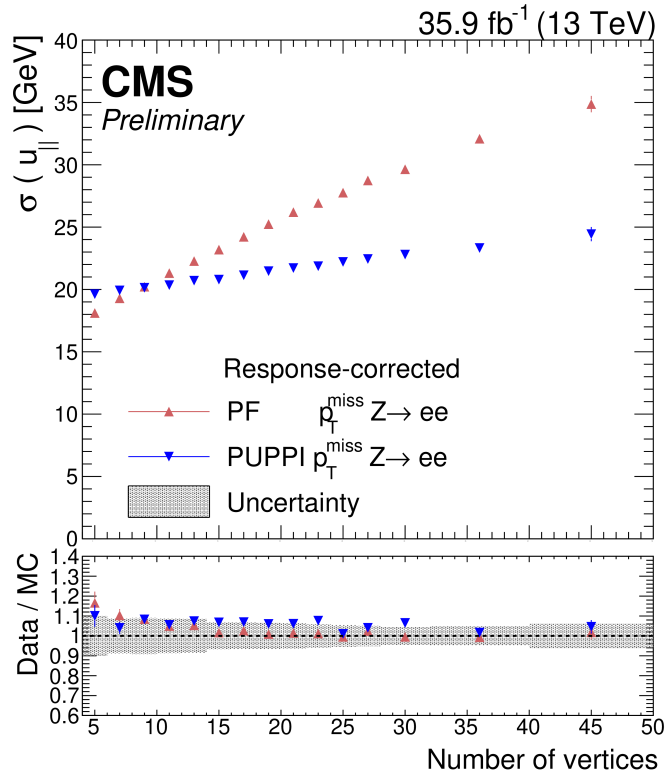


FIGURE 3.3: Resolution of missing transverse momenta of  $Z^0 \rightarrow e^+e^-$  as a function of reconstructed pp collision vertices. PF (red triangles) corresponds to events reconstructed only with the particle flow algorithm. PUPPI (blue triangles) illustrates PF events corrected with the PUPPI method. [77]

### 3.4 Reconstruction of muons

Muons are widely used in various analyses due to the clarity of signal events. The muon reconstruction consists of five main steps: muon hit and segment reconstruction, muon track reconstruction, muon identification, momentum determination, and muon isolation. The next sections briefly describe each of the steps.

### 3.4.1 Hit and segment reconstruction

Muon reconstruction starts with a collection of "hits" - electric signals produced on the wires and strips after muon gas ionization. So-called "local" algorithms collect information from the drift tubes (DT), residual plate chambers (RPC) and cathode strip chambers (CSC) [78].

In DT, the charge from the gas ionization is collected using anode wire under a voltage. The arrival time is obtained using a time-to-digital converter (TDC). This time,  $T_{\text{TDC}}$ , is corrected by the pedestal time,  $T_{\text{ped}}$ , and multiplied by the drift velocity ( $v$ ). The DT hit position calculation is calculated as

$$\text{Hit}_{\text{DT}} = (T_{\text{TDC}} - T_{\text{ped}}) \times v. \quad (3.2)$$

The CSC detector measures the hit position by combining the information from the anode wires and cathode strips. The radial location of the strips allows measuring the  $\varphi$  angle precisely. Traveling through the gap between two plates of RPC, the muon produces an electron shower, that is collected on an external strip readout plane as a charge. That charge produces a signal on a readout plate of an external strip allowing to identify muons with a precision of a few ns. The charge from ionization can be distributed among a few strips. For that reason, neighboring strips are clustered together [54]. Such configuration allows reconstructing hits more accurately.

### 3.4.2 Muon track reconstruction

In the standard CMS track reconstruction scheme, the tracks are reconstructed independently in the tracker system and muon system. After that, the information about reconstructed tracks proceeds for muon track reconstruction, which consists of three different types [79].

- Standalone-muon tracks are created using the information about the position of the hits in the muon detector. In particular, it combines the data from DT, RPC, and CSC. In this reconstruction scheme, the Kalman-filter technique is widely used [80]. The reconstruction starts with reconstructed seeds in DT groups and CSC segments.
- Tracker-muon tracks are reconstructed using a chain of slightly different tracking algorithms in an iterative approach. The tracks are built "inside-out" by extrapolating tracker tracks with  $p_{\text{T}} > 0.5$  GeV and total momentum  $p > 2.5$  GeV to the muon system to match tracks reconstructed in DT or CSC parts. The tracker track is determined as a tracker-muon track if at least one muon segment matches the extrapolated track and if the absolute value of the difference between positions of the tracks in the x coordinate is smaller than 3 cm.
- Global muon tracks, unlike the tracker muon tracks, are built in the opposite direction - "outside-in". Using Kalman filter and information about the standalone muon track and tracker track, the algorithm matches the tracks through the fit. At large transverse momenta,  $p_{\text{T}} > 200$  GeV, the resolution can be improved with the information from the global muon fit. The hadron shower remnants can punch through the innermost muon station (punch-through effect) [55]. Such events may be a source of misidentification for those muons, which match only one segment

in the innermost muon station. However, requiring more than one muon station hit during the reconstruction can reduce the impact of this effect.

### 3.4.3 Muon identification types

The muon identification criteria may be different depending on the desired balance between efficiency and purity of the muon selection. The identification types may vary depending on the  $\chi^2$  of the track fit, level of matching between standalone track and tracker track, number of hits per track, etc. In total, different identification schemes may be categorized into five main types:

- Loose muon identification (ID) is a set of requirements that reconstruct a muon using the PF algorithm. Such muon is required to be identified as a global or a tracker muon, originate from a primary vertex, and heavy flavor decays. That ID type has a low rate of misidentification of charged hadrons as muons.
- Medium muon ID is used for prompt muons and muons originating from heavy-flavor decay. It is required to be a loose muon that uses more than 80% of hits in the inner tracker track during the fit. The muon tracker segment compatibility should be more than 0.451 [78]. The goodness of the global fit is required to be less than three,  $\chi^2/\text{dof} < 3$ . The matching between the standalone muon and the tracker muon must be  $\chi^2/\text{dof} < 12$ .
- Tight muon ID is a loose muon that uses hits from at least six layers of the inner tracker, including one pixel hit. The muon is required to be reconstructed as a tracker and global muon. The goodness of the global muon fit is required to be  $\chi^2/\text{dof} < 10$  and include at least one hit in the muon system. The longitudinal impact parameter of the primary vertex must be  $|dz| < 0.5$  cm and a transverse impact parameter  $|d_{XY}| < 0.2$  cm. Such conditions suppress muons that appear in decays on a flight and reduces misidentification of hadronic punch-through events. This muon ID was chosen for the muon selection in the analysis described in chapter 5.
- Soft muon ID is mostly used for quarkonia analysis and in B-physics analysis. The reconstruction of the muon tracker track uses hits from at least six tracker layers, and at least one pixel hit. Also, the track is required to satisfy a high purity flag [81]. The primary vertex requirements are  $|d_{XY}| < 0.3$  cm and  $|dz| < 20$  cm.
- High momentum muon ID defines a muon as a reconstructed global muon and tracker muon with  $p_T > 200$  GeV. Most of the requirements are the same as for the tight muon. However, when high  $p_T$  muons move through the steel flux return yoke, they radiate large electromagnetic showers. The showers are the source of additional hits in the muon chambers. To prevent inefficiencies during the fit, the requirement on the goodness of the global muon fit  $\chi^2/\text{dof}$  is removed [65]. The relative  $p_T$  uncertainty is required to be  $\sigma(p_T)/p_T < 30\%$ .

### 3.4.4 Momentum determination

The  $p_T$  value of each muon is obtained through four central fitting schemes described below:

- **Inner-Track fit** collects the information from the inner tracker. This fit has proved to be highly effective for muons with  $p_T < 200$  GeV. The  $p_T$  values of muons in this analysis (Chapter 5) were calculated using this fit.
- **Tracker-Plus-First-Muon-Station fit** performs a refit of the global muon track hits using information from the inner tracker track and the innermost muon station. The innermost muon station has the best information about the muon  $p_T$  in the muon station.
- **Picky fit** may be used in the events where the muon chambers contain showers. The algorithm combines information from the global muon track and chambers with showers. Using the condition on the  $\chi^2$  values the refit selects the hits from the shower which match to the extrapolated global muon track trajectory.
- **Dynamic-Truncation fit** is used when the muon energy losses significantly bend the muon trajectory. In this case, the algorithm is performed in an iterative way. The tracker tracks are extrapolated to the innermost station. The closest hits to the propagated trajectory are then used in the fit. If the fit converges, the next layer of the muon station is added, and the procedure repeats until it is possible. The algorithm stops if it finds no suitable hits in two consecutive muon stations.

To determine the muon momentum, the Tune-P algorithm is used [78]. The algorithm processes result from each of the fitting schemes. To choose the  $p_T$  value, it analyzes the goodness of the fit and momentum resolution.

### 3.4.5 Muon isolation

The muon isolation procedure is targeted to distinguish between prompt muons and those from the weak decays in jets. The isolation is a procedure of summing the energy in the geometrical cone  $\Delta R$ . Two main strategies are available - isolation based on the summation of reconstructed tracks in the cone and PF isolation, where the procedure is performed using the PF reconstructed objects - charged hadrons and neutral particles.

In the PF isolation scheme, the sum of the transverse momenta of the charged hadrons, the neutral hadrons, and photons originating from the primary vertex within the cone is used as the discrimination variable (PFIso). The cone also contains contributions from other vertices due to the high pile-up. In order to reduce this dependence,  $\Delta\beta$  correction is used - half of the summed transverse momenta of the charged particles from other vertices is subtracted from the cone. The factor of 0.5 is estimated from the simulation of inelastic proton-proton collisions [65]. It corresponds to the ratio of neutral particles to charged hadron production. To reflect the level of the muon cone contamination with respect to the muon  $p_T$  the PFIso value, after correction, is divided by the muon  $p_T$ .

### 3.4.6 Reconstruction, identification and isolation efficiencies

The muon efficiency is studied with the Tag-and-Probe method [82] using 2015 data and MC sample. The total efficiency value is calculated by factorizing its components

$$\varepsilon_{\mu} = \varepsilon_{\text{track}} \times \varepsilon_{\text{reco+ID}} \times \varepsilon_{\text{iso}} \times \varepsilon_{\text{trig}}, \quad (3.3)$$

where  $\varepsilon_{\text{track}}$  corresponds to the track reconstruction efficiency,  $\varepsilon_{\text{reco+ID}}$  is the combined efficiency of the ID criteria and reconstruction in the muon detector. Efficiency of the muon isolation is shown as  $\varepsilon_{\text{iso}}$ , and  $\varepsilon_{\text{trig}}$  is the trigger efficiency.

The  $\varepsilon_{\text{track}}$  for the isolated muon with  $1 \text{ GeV} < p_{\text{T}} < 100 \text{ GeV}$  is  $> 99\%$ . Muons pierce through the whole tracker detector, and unlike electrons, muons have a negligible amount of bremsstrahlung radiation. The dominant contribution to the altering of the muon trajectory is Coulomb scattering [70]. These qualities assure a high reconstruction efficiency. The results are shown in Figure 3.4.

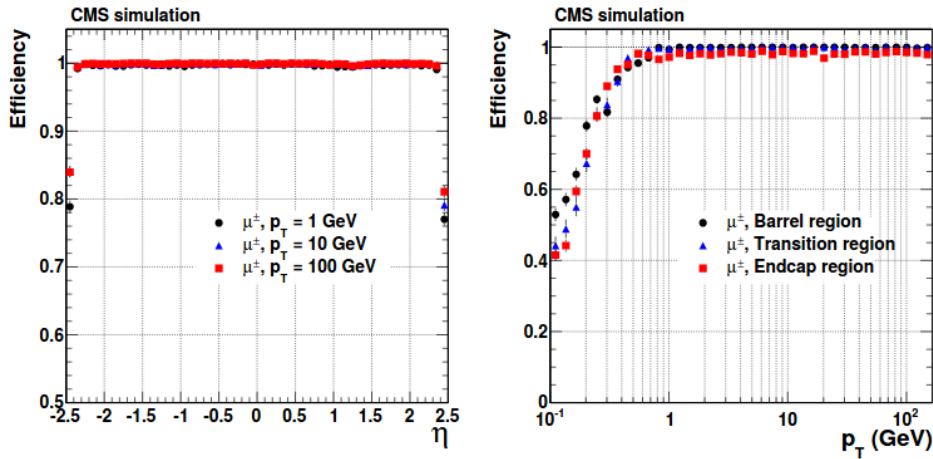


FIGURE 3.4:  $\varepsilon_{\text{track}}$  for isolated muons shown as a function of pseudorapidity (left) with  $p_{\text{T}} = 1, 10, 100 \text{ GeV}$  and transverse momentum (right) for barrel ( $0.0 < |\eta| < 0.9$ ), transition ( $0.9 < |\eta| < 1.4$ ), and endcap ( $1.4 < |\eta| < 2.5$ ) regions [70].

The  $\varepsilon_{\text{reco+ID}}$  efficiency estimation is defined for tight ID and loose ID. The loose ID efficiency is around 99% across the whole pseudorapidity range, the MC and data agree up to 1%. The tight ID efficiency varies depending on  $\eta$  region from 95% to 99%. Data and MC agree up to 1% – 3%. Results, as a function of  $\eta$ , are shown in Figure 3.5.

In some cases (decays in flight, "punch-through" effect, and random matching) hadron may be misidentified as a muon. The probability of a hadron to be misidentified as a muon is obtained using data samples of kaons and pions produced in resonance decays. According to the results, the kaons can be misidentified as a loose muon in data and simulation with probability 0.5%, as a tight muon 0.3%. For the pions, results are 0.2% and 0.1%, respectively [78].



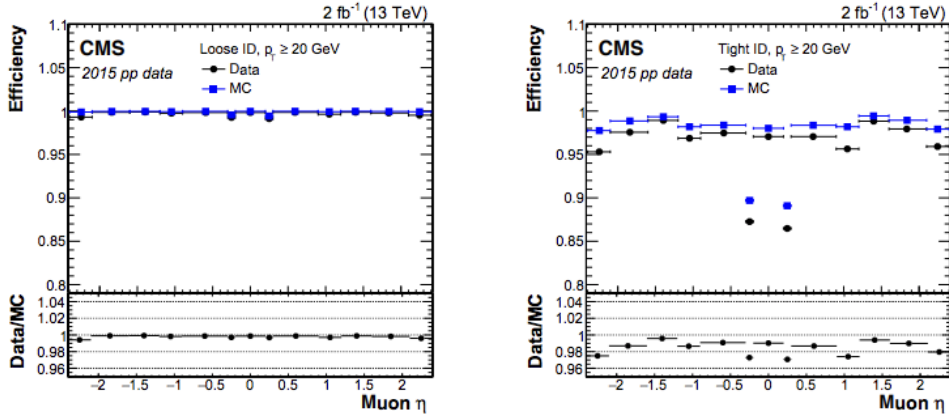


FIGURE 3.5: Tag-and-probe  $\epsilon_{\text{reco+ID}}$  efficiency for loose ID (left) and tight ID (right) as a function of  $\eta$ . Detector regions with a smaller amount of instrumentation cause the dips in efficiency. The statistical uncertainties are tiny [65].

The muon isolation efficiency,  $\epsilon_{\text{iso}}$ , is studied relatively to a probe that passes the tight ID criteria. The estimated agreement between data and MC is around 0.5%. The  $\epsilon_{\text{iso}}$  efficiency results for tight ID as a function of  $\eta$  and  $p_{\text{T}}$  is shown in Figure 3.6. The studies on the probability of incorrect labeling of muons within jets were performed using simulated QCD events enriched in muon decays. According to them, the probability of a muon with  $p_{\text{T}} > 20$  GeV that passes tight muon ID to satisfy tight isolation conditions is near 5% for the barrel, and 15% for the endcap.

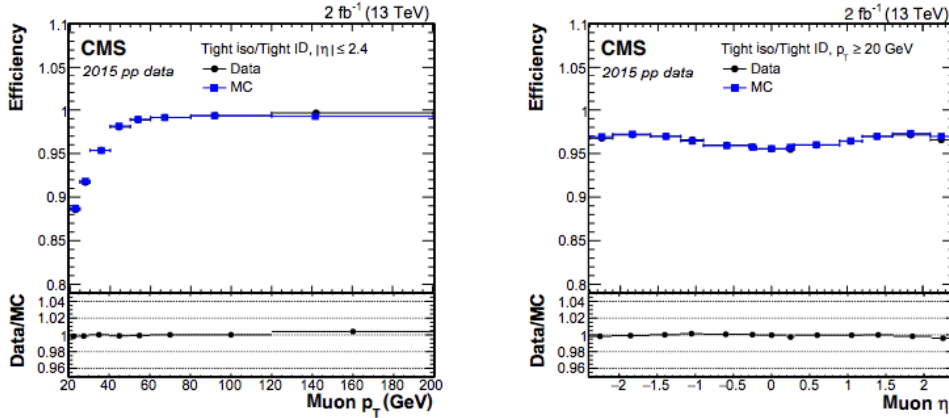


FIGURE 3.6: Tag-and-probe  $\epsilon_{\text{iso}}$  efficiency for tight ID as a function of  $p_{\text{T}}$  (left) and  $\eta$  (right). The statistical uncertainties are smaller than the symbols [65].

The trigger efficiency is estimated using the Tag-and-Probe technique. The tag is matched geometrically to the HLT trigger that selected the event. This is done to avoid a possible bias of the  $\epsilon_{\text{trig}}$  measurement. Backgrounds are reduced by requiring the tag to satisfy PF isolation criteria and tight ID. Figure 3.7 represents the  $\epsilon_{\text{trig}}$  efficiency as a function of reconstructed muon  $p_{\text{T}}$  and  $\eta$ .

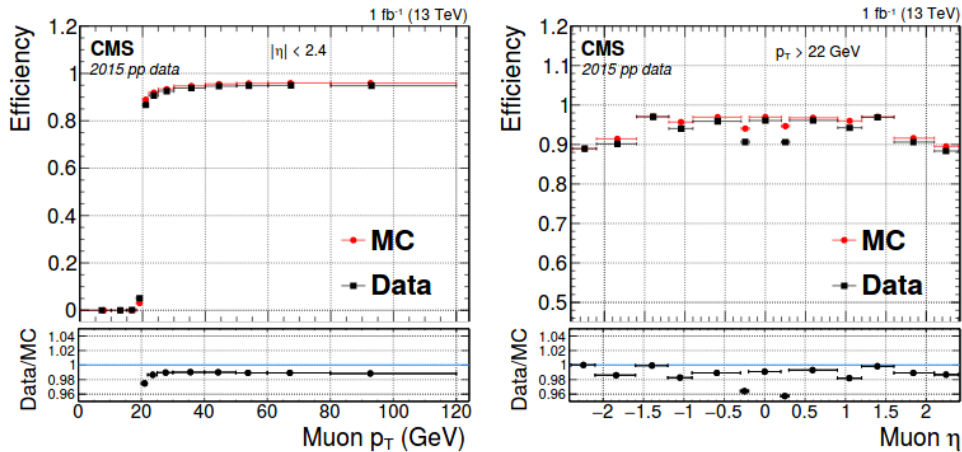


FIGURE 3.7: The  $\epsilon_{\text{trig}}$  efficiency of an isolated muon is presented. The left plot shows  $\epsilon_{\text{trig}}$  efficiency as a function of muon  $p_{\text{T}}$  while the right corresponds to a function of muon  $\eta$ . The statistical uncertainties are smaller than the markers [65].

The dips in the distribution of the muon efficiency as a function of pseudorapidity are caused by geometrical features of the detector. Table 3.1 represents the trigger efficiency of L1 and HLT track reconstruction.

Step	$\eta$ region	Data eff. [%]	Scale factor
L1 w.r.t. offline	$0.0 <  \eta  < 0.9$	$96.8 \pm 0.02$	$0.9914 \pm 0.0005$
	$0.9 <  \eta  < 2.4$	$94.38 \pm 0.02$	$0.9947 \pm 0.0005$
HLT w.r.t L1	$0.0 <  \eta  < 0.9$	$99.67 \pm 0.02$	$0.9967 \pm 0.0005$
	$0.9 <  \eta  < 2.4$	$99.46 \pm 0.02$	$0.9957 \pm 0.0005$
Online isolation w.r.t HLT	$0.0 <  \eta  < 0.9$	$97.95 \pm 0.02$	$0.9906 \pm 0.0005$
	$0.9 <  \eta  < 2.4$	$98.28 \pm 0.02$	$0.9931 \pm 0.0005$

TABLE 3.1: The isolated single-muon trigger efficiency using 2015 data. The first row represents the Level-1 efficiency ( $p_{\text{T}}$  threshold 16 GeV) with respect to (w.r.t.) offline muons. The second two rows show the HLT efficiency ( $p_{\text{T}} < 20$  GeV) w.r.t. to offline muons geometrically linked to L1 candidates. The last rows correspond to online isolation efficiency w.r.t. offline muons firing HLT. The last column shows the scale factor between MC and data. Represented uncertainties are statistical [65].

### 3.5 Reconstruction of missing transverse energy

The missing transverse energy (MET),  $E_{\text{T}}^{\text{miss}}$ , is an essential part of the LHC event reconstruction. Many searches of physics beyond the standard model are focused on the missing transverse energy reconstruction and its interpretation in terms of new theories. Another considerable application is an estimation of the number of produced  $W^{\pm}$  bosons, decaying in the lepton channel. In this channel, the  $W^{\pm}$  boson decay to a pair of lepton and lepton neutrino due to the lepton number conservation law. Direct neutrino detection is not possible due to its low cross-section of interaction; for this reason, the amount of energy carried away by the neutrino is not detected. According to the momentum conservation in the plane transverse to the beams, the total transverse

momentum of all particles produced in the event is equal to zero. Precise measurement of all available particles in the event allows to construct the transverse momentum distribution and spot the imbalance that corresponds to undetected particles. The  $\vec{E}_T^{\text{miss}}$  is computed as a negative of the vectorial sum of particles momenta, and is defined by

$$\vec{E}_T^{\text{miss}} = - \sum_{i=1}^{\text{all}} \vec{p}_{T_i}, \quad (3.4)$$

where  $\vec{p}_{T_i}$  corresponds to the transverse momenta of all reconstructed particles in the event. The modulus of  $\vec{E}_T^{\text{miss}}$  value is referred to as  $E_T^{\text{miss}}$  and is used in the analysis (Chapter 5) as the main variable for the signal yields extraction. CMS has four main algorithms to estimate  $E_T^{\text{miss}}$ : Particle Flow MET [69], PUPPI  $E_T^{\text{miss}}$  [76], Calo MET method [83], and Track-Corrected  $E_T^{\text{miss}}$  (TC) algorithm [84].

The Particle Flow MET uses  $p_T$  values of all particles reconstructed with the PF algorithm [69]. The values of MET are obtained using formula 3.4. The corrections that are applied to the MET values are briefly described later in this section.

PUPPI  $E_T^{\text{miss}}$  calculation applies PUPPI corrections to the PF objects [76]. The definition of local shape variable  $\alpha$  is redefined for PF objects. The  $\alpha$  variable of  $i$ -th PF object is defined as

$$\alpha_i = \log \sum_{j \neq i, \Delta R_{ij} < 0.4} \left( \frac{p_{T_j}}{\Delta R_{ij}} \right)^2 \begin{cases} \text{for } |\eta_i| < 2.5, j \text{ are charged PF candidates from PV,} \\ \text{for } |\eta_i| > 2.5, j \text{ are all kinds of PF candidates.} \end{cases} \quad (3.5)$$

Here  $j$  represents charged PF candidates that originate from the primary vertex (PV) in a cone of radius  $R$  in  $\eta - \phi$  space around PF candidate  $i$ ,  $\Delta R_{ij}$  is a distance between  $i$  and  $j$  in  $\eta - \phi$  space. If a charged PF candidate is not associated with the PV, but the distance between the track and PV is less than 0.3 cm,  $d_z < 0.3$  cm, the candidate is used in the calculation [85]. The  $\chi^2$  approximation is used to determine the likelihood that the PF candidate is originating from the pile-up:

$$\chi_i^2 = \frac{(\alpha_i - \bar{\alpha}_{\text{PU}})^2}{\text{RMS}_{\text{PU}}^2}. \quad (3.6)$$

Here  $\bar{\alpha}_{\text{PU}}$  is a median value of the  $\alpha_i$  distribution for pile-up particles of the event. The root mean square of the  $\alpha_i$  distribution is shown as  $\text{RMS}_{\text{PU}}$ . The  $\chi^2$  values are used to calculate the weight  $w$ , which means a probability of the PF candidate to originate from the PV. The expression for the weight is given as:

$$w_i = F_{\chi^2, \text{NDF}=1}(\chi_i^2). \quad (3.7)$$

The  $F_{\chi^2, \text{NDF}=1}$  is the cumulative distribution function that approximates the  $\chi^2$  distribution with one degree of freedom of all PF candidates in the event [85]. After the weight is defined for every PF candidate, the  $E_T^{\text{miss}}$  can be calculated using formula 3.4.

Calo MET method is based on calorimeter tower geometry and calorimeter energies. The MET is calculated using the same formula 3.4, but the information about the transverse energy is taken from the energy deposits in calorimetry towers [71].

The TC  $E_T^{\text{miss}}$  algorithm is based on the Calo MET method and includes reconstructed tracks from the inner tracker [55]. For the tracks that are not identified as

electron or muon, the simulation of energy deposit from charged pion is used [84].

Nonlinearities in detector response, inefficiencies of the tracker, and minimum energy thresholds in the calorimeters can lead to overestimated or underestimated values of the  $E_T^{\text{miss}}$ . For that reason, CMS uses a few types of MET corrections:

- Type-0 correction aims to consider the impact of the degraded MET reconstruction, which appears from the pile-up interaction. The MET reconstruction degrades as the pile-up increases. Type-0 correction applies Charged Hadron Subtraction (CHS) scheme to the MET. The scheme removes contributions from charged hadrons that originate from the pile-up events [83]. The correction is developed for the PF MET and cannot be used for the Calo MET.
- Type-I correction corresponds to the propagation of jet energy correction (JEC [86]) to MET. In few words, this correction replaces the four-vectors of jets with the JEC four-vectors.
- Type-II correction corrects the  $\vec{p}_T$  of unrescaled particles by uniformly scaling it with a constant factor. The correction was originally made for the Calo MET and is not recommended for PF MET.
- The XY-Shift correction aims to reduce the MET  $\varphi$  modulation. The experimental distribution of MET shows that it has a non-uniform distribution in  $\varphi$ . At the same time, due to the axial symmetry of events production, MET values are expected to have uniform distribution across the  $\varphi$  coordinate. The observed  $\varphi$  modulation can be explained by the presence of the anisotropic detector response, inactive calorimeter cells, the detector misalignment, and the displacement of the beam spot. The modulation can be reduced by shifting the origin of the coordinate in the transverse plane.

## Chapter 4

# Luminosity measurement at CMS

One of the primary goals of high energy physics experiments is to measure the cross-section value of specific physics processes. The measurements depend on a luminosity value - a measure that reflects how effectively the collider can produce collision rates. Uncertainty from a luminosity measurement is present in every cross-section or limit result obtained from data analysis. For that reason, the precise luminosity measurement is one of the crucial parts of the experiment.

The luminosity,  $L$ , at the LHC can be defined as:

$$L = \frac{N_1 \cdot N_2 \cdot f \cdot n_b}{A_{\text{eff}}}. \quad (4.1)$$

Here  $N_1$  and  $N_2$  are numbers of particles in bunches,  $n_b$  is number of colliding bunches,  $f$  is the revolution frequency [51] of the collider and  $A_{\text{eff}}$  corresponds to the area of the luminous region. Values of  $N_1$ ,  $N_2$ ,  $n_b$ , and  $f$  can be obtained from the LHC measurements. The luminous region is defined, assuming Gaussian distributions of the particle density in colliding bunches, by the formula:

$$A_{\text{eff}} = 4\pi \cdot \sigma_x \cdot \sigma_y, \quad (4.2)$$

where  $\sigma_x$  and  $\sigma_y$  are the widths of the Gaussian distributions of the lateral particle density in the bunches. This feature finds its reflection in the acceleration scheme: as it was already mentioned (see Chapter 2.1), each detector has two inner magnet systems from both sides to squeeze the beams from 0.2 mm to 16  $\mu\text{m}$  right before the interaction.

One of the luminosity measurement in the CMS experiment is performed by a subgroup of the Beam Radiation Instrumentation and Luminosity project (BRIL).

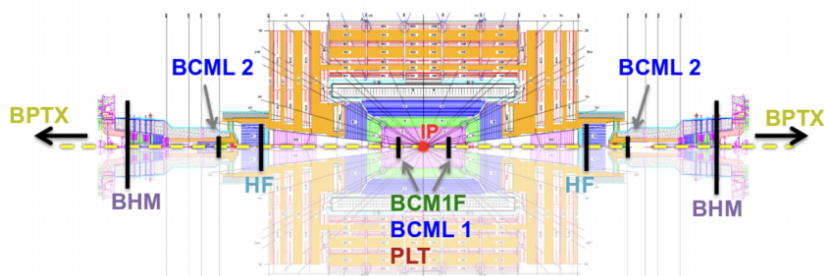


FIGURE 4.1: The picture shows the CMS detector in the Y-Z plane. Names of the BRIL subsystems are shown in different colors. The black lines indicate positions of the subsystems. The BPTX is at a distance of 175 m from the IP and is not shown on the scheme [87].

BRIL consists of subgroups whose responsibilities also include monitoring beam conditions and machinery induced background. These subgroups also design and develop other systems, such as The Beam Condition Monitor “Leakage” (BCML1 and BCML2), Beam Halo Monitor (BHM), and The Beam Pick-up Timing for the experiment (BPTX). Currently, the luminosity measurements are done using four independent systems: the Hadron Forward calorimeter (HF) [59], silicon pixel detector [88], Pixel Luminosity Telescope (PLT) [89], and Fast Beam Condition Monitor (BCM1F) [90]. Luminometers and other BRIL subsystems are presented in Figure 4.1.

## 4.1 Luminosity measurement using BCM1F

The BCM1F is a beam halo monitor and an online, bunch-by-bunch luminometer with diamond and silicon sensors that have a time resolution near one ns. It was developed from the requirement to have a very fast, sensitive to the beam conditions monitoring system, to prevent the inner pixel detector from damage due to rapid beam losses [90]. The BCM1F is installed inside the pixel volume at a distance of 6.94 cm from the beam axis in r-plane, and 1.8 m away from the interaction point in z-axis. The detector consists of four half-rings called C-shapes [91], each of which is equipped with six sensors, as shown in Fig. 4.2. Two C-shapes form a ring around the beam pipe, perpendicularly to the direction of the beams. Depending on their location relative to the interaction point and to the center of the LHC ring, the C-shapes are named "Z+", "Z-", and "NEAR", "FAR", respectively.

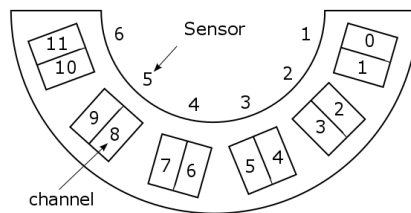


FIGURE 4.2: C-shape equipped with six sensors. Each sensor has two channels [91].

The C-shapes are located at a distance of six nanoseconds before the nominal interaction point, registering incoming and outgoing particles during 12.5 ns. The detector has a good enough time resolution and fast front-end electronics to separate machine induced background (MIB) rates from luminosity measurements. Figure 4.3 shows a detecting cycle of one bunch crossing.

The principle of BCM1F beam condition and luminosity measurement is based on the online monitoring of the hit rates produced from incoming MIB and particles after a collision. An example of measured rates for one bunch crossing as a function of time, collected during 2.5 minutes, is shown in Fig. 4.4. Each bunch crossing has 4 bins, with a time interval of near 6 ns each. The plot shows the ability to discriminate between incoming and outgoing particles. The first peak corresponds to the beam halo from the incoming beam, while the second peak is the sum of halo particles and collision products.

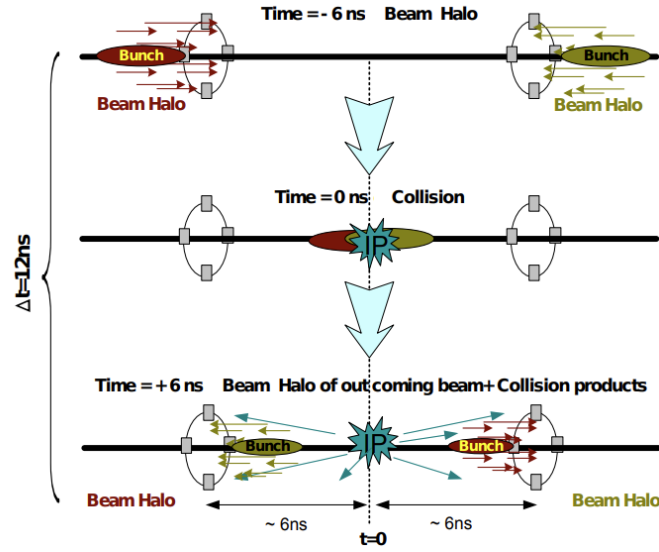


FIGURE 4.3: The scheme represents arrival time before and after the bunch crossing. The thick black line is the beam-line, circles with gray parallelograms are the BCM1F rings with schematic sensors. The first row shows proton bunches, and their halos passing the BCM1F right before the collision, which corresponds to a timing -6 ns. The second row shows the collision at the interaction point (IP). The third row is dedicated to the outgoing bunches after the collision with some collision products passing the BCM1F sensors, the timing of this stage corresponds to 6 ns after the collision. Overall timing of each collision is 12 ns [87].

Each histogram from each sensor channel is integrated over 4096 LHC turns (a luminosity nibble), or  $4096 \times 89 \mu\text{s}$ . The luminosity is measured using the mean value of the hit rates,  $\mu$ , recorded in the second peak. In this approach luminosity,  $L$ , can be expressed as a function of  $\mu$  with the formula:

$$L = \frac{n_b f \mu}{\sigma_{\text{visible}}}. \quad (4.3)$$

Here,  $n_b$  is a number of bunches,  $f$  is a revolution frequency,  $\sigma_{\text{visible}}$  is the visible cross-section in the detector. The mean value of the hit rates is obtained from the "Zero counting" method. In this method, hit rates distribution is assumed to follow the Poisson distribution. The probability of observing  $n$  hit rates is then given by the following formula:

$$P(n) = \frac{\mu^n e^{-\mu}}{n!}, \quad (4.4)$$

where  $\mu$  is the mean value. The probability to have zero hits,  $P(0)$ , is given by:

$$P(0) = e^{-\mu} \quad (4.5)$$

$$\mu = -\ln[P(0)] = -\ln[1 - P(> 0)] \quad (4.6)$$

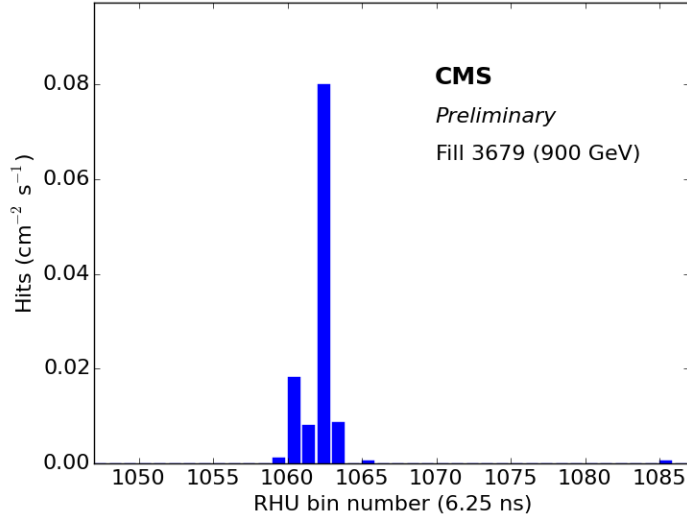


FIGURE 4.4: Hit rates as a function of time for one bunch crossing collected during 2.5 minutes [92].

The visible cross-section in the detector,  $\sigma_{\text{visible}}$ , can be obtained combining equations 4.1, 4.2, and 4.3:

$$\sigma_{\text{visible}} = \mu \frac{2\pi \Sigma_x \Sigma_y}{N_1 N_2}, \quad (4.7)$$

where  $\Sigma_x$  and  $\Sigma_y$  are measured beam width in Van der Meer scans (VdM)

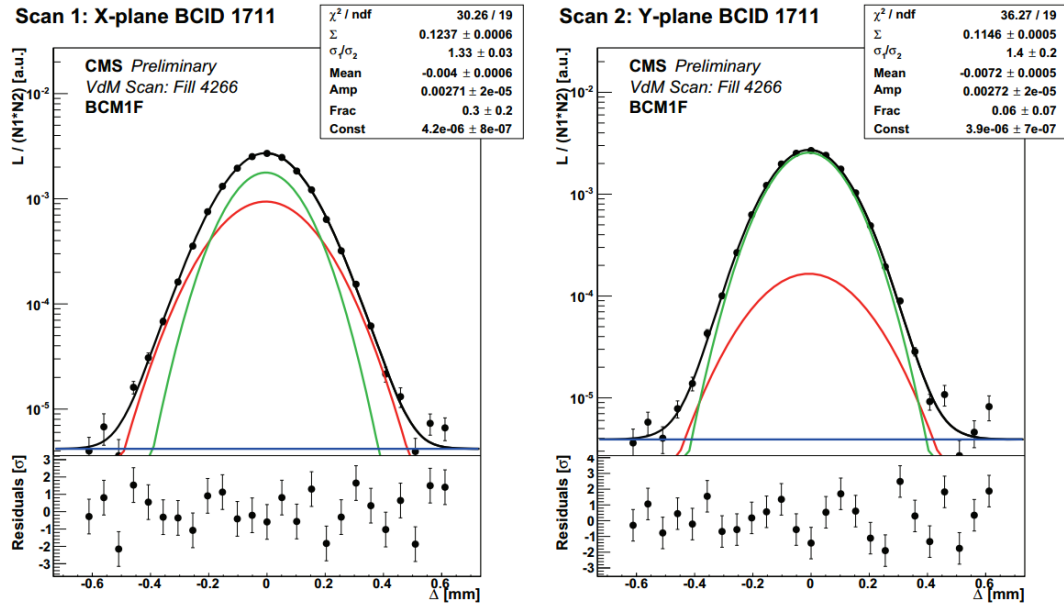


FIGURE 4.5: Van der Meer scan as performed during the fill 4266. The left plot corresponds to X-plane, while the right is Y-plane. The plots show the measured values for each displacement as a function of separation distance. Black dots correspond to measured rate values at each step. Red and green lines correspond to a Gaussian fit, constant term is shown with a blue line. The resulting fit is shown with a black line [93].



[94], which are equivalent to  $\sigma_x$  and  $\sigma_y$ . The VdM scan is performed in specific LHC runs, where the beams are varied from their centered positions with small steps in x and y directions. During such variations, hit rates are measured as a function of the beam position. The more beam is varied from its position, the fewer hit rates are observed. Figure 4.5 shows an example of the Van der Meer scan. The average number of proton interactions,  $\mu$ , is plotted as a function of the beam separation distance [95].

## 4.2 Fast Beam Condition Monitor Detector Upgrade during the Run 2

During the short shut down in 2017, BCM1F was upgraded with 14 diamond and 10 silicon (Si) sensors of near  $4.5 \times 4.5 \text{ mm}^2$  area and  $500 \text{ }\mu\text{m}$  thickness each. Two types of diamond sensors are used: Poly-crystalline Chemical Vapor Deposited (pCVD) diamond and Single-crystalline Chemical Vapor Deposited (sCVD) diamond. The pCVD diamonds are grown on wafers with a single Si crystal, while for more expensive and high-quality sCVD diamonds production, a high-temperature-high-pressure (HTHP) diamond substrate with a single-crystal structure is used [96]. The main difference between these two approaches is that the sCVD is grown as a single crystal, while the pCVD diamond consists of many small crystals. Figure 4.6 shows an illustrative example of a pCVD diamond surface image obtained using a scanning electron microscope.

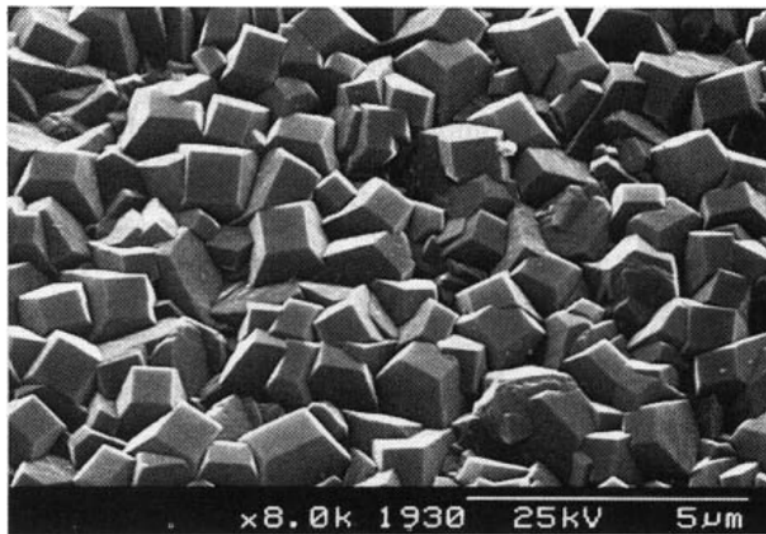
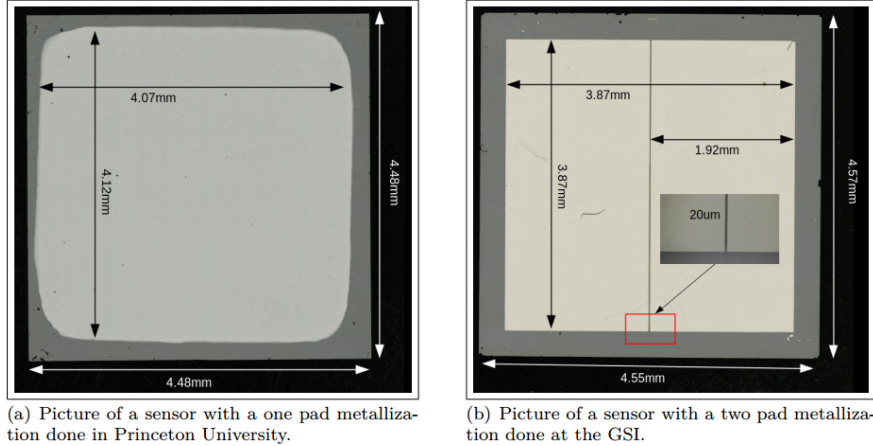


FIGURE 4.6: The graph of an electron scan of typical poly-crystalline diamond films grown on Si by CVD, using 0.5% methane in hydrogen.[97]

Each sensor, from both sides, is covered with the metallization pads using the photolithography technique [98]; the ionization charge is collected by applying a high voltage on the pads. The metallization is performed in two different laboratories and with different metal alloys. One is done at Princeton University with an alloy 50:50 of atoms of tungsten and titanium (W/Ti). The second alloy is applied at the GSI Helmholtz Center for Heavy Ion Research using a 50 nm thick layer of chromium (Cr) and a 150 nm thick layer of gold (Au). One side of the sensor has a metallization with rounded corners with the size near  $4 \times 4 \text{ mm}^2$ . The second side of sensors has a two pad metallization with sizes approximately  $3.87 \times 1.92 \text{ mm}^2$  each. Two pad metallization allows

having two channels per sensor. The distance between the metallizations on the two pad metallization side is approximately  $20\ \mu\text{m}$ . Optical inspection of a metallized sensor is shown in Figure 4.7.



(a) Picture of a sensor with a one pad metallization done in Princeton University.

(b) Picture of a sensor with a two pad metallization done at the GSI.

FIGURE 4.7: Geometrical characteristics of a sensor and its one pad metallization (Picture 4.7 a) and two pad metallization (Picture 4.7 b) [99].

The ionization signal, produced in a sensor, is transmitted in two channels through a right and left pads to a custom fast front-end application-specific integrated circuit (ASIC) [100]. In ASIC, signals are amplified and passed through the left and right outputs to Analog-Opto-Hybrids (AOHs), in which they are converted into optical signals [101]. The AOH settings are configured using digital-opto-hybrid (DOH) [102]. Figure 4.8 shows a schematic view of the signal chain from a sensor to AOHs. Each C-shape is served by six ASICs (per channel) and four AOHs. Optical signals are transferred to the counting room using a ribbon fiber with 12 channels. Afterward, to convert optical signals into electric ones, signals are fed into four optical receiver boards (one per C-Shape), with 12 channels each. BCM1F has two different approaches in signal treatment at the back-end stage - using Versa Module Europa Bus (VME) [103] and Micro Telecommunications Computing Architecture ( $\mu\text{TCA}$ ) [104]. The optical receiver produces a positive copy of each channel for  $\mu\text{TCA}$  and a negative copy to feed the VME [93].

During Run 2,  $\mu\text{TCA}$  was used for testing and development for future replacement of the ADC in the next Run.

The VME back-end consists of read-out Analog-Digital-Converters (ADC) [105] and deadtime-free Real-time Histogramming Units (RHUs) with fast discriminators [87]. The system performance and sensors monitoring are tracked using ADC data, while MIB and luminosity measurements are calculated with the RHU raw rates. Details on VME ADC data studies and sensor performance during Run 2 is given in section 4.3.

#### 4.2.1 Diamond sensors measurements

The sCVD and pCVD sensors have a list of advantages that make them a perfect tool for MIB and luminosity measurements. Diamonds have a high level of radiation hardness with an excellent time resolution (about nanoseconds). The operation regime does not require the usage of the cooling system; this feature makes diamonds especially

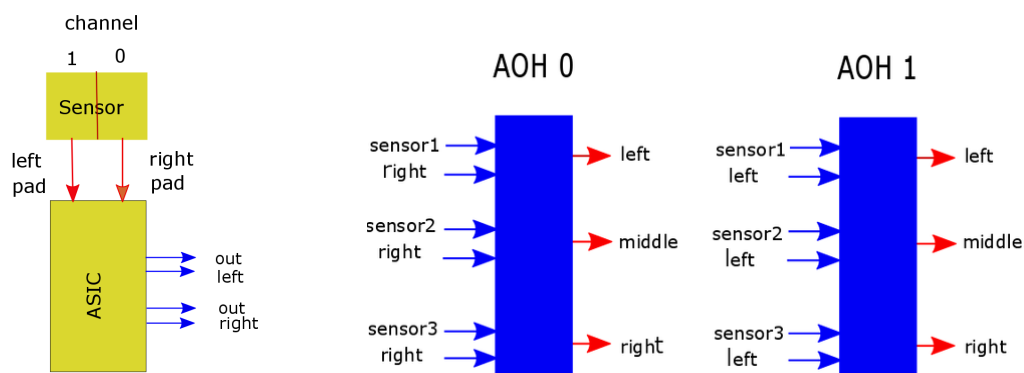


FIGURE 4.8: Schematic representation of signal transferring from sensor channels to AOH. The left plot shows a sensor with two channels and two pads that are connected to the ASIC. Blue arrows correspond to ASIC output, which is connected to AOHs. The right plot shows two AOHs with incoming signals marked as blue arrows. Each AOH processes only right or left ASIC output signal from three different sensors. Red arrows correspond to outgoing optical fibers marked as left, middle, and right.

attractive due to space limitations near the beam pipe. The diamond sensors show a satisfying efficiency with a relatively small size ( $5 \times 5 \text{ mm}^2$ ).

During the Run 1, diamond sensors have proven their reliability and high-performance [106], they were chosen to be used again during the Run 2. In 2016-2017, 20 diamonds (sCVD and pCVD) were tested in laboratories of DESY-Zeuthen and DESY Hamburg to select fourteen as nominal sensors for four main C-shapes.

This section describes techniques that were used to probe diamond performance, selected sensor measurement results, and main conclusions from these studies. Presented measurements, unless mentioned, including various sensors measurements (including BCM1F sensors monitoring during the whole Run 2 and multiple improvements in the BCM1F data analysis, see Section 4.3), were performed and obtained by the author as a part of his CMS member responsibilities. The general information on diamond energy loss is presented in the first paragraph. The next three paragraphs represent performed tests: leakage current estimation, signal stability tests, and charge collection efficiency measurement.

### Energy loss of diamond solid-state sensor.

Basic concepts of physical processes that stand behind various detecting techniques were briefly introduced in section 2.2.2. The energy loss of charged ionizing particle,  $\frac{dE}{dx}$  that is moving through a diamond is a sum of two main mechanisms, ionization,  $\frac{dE}{dx}|_{\text{ioniz}}$ , and bremsstrahlung,  $\frac{dE}{dx}|_{\text{brems}}$ :

$$\frac{dE}{dx} = \frac{dE}{dx}|_{\text{ioniz}} + \frac{dE}{dx}|_{\text{brems}}. \quad (4.8)$$

The mean ionization energy loss for relativistic charged heavy particle is described by the Bethe-Bloch formula:

$$-\left\langle \frac{dE}{dx} \right\rangle = K Z^2 \frac{1}{A \beta} \left[ \frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 T_{\max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right], \quad (4.9)$$

$$K = 4\pi N_A r_e^2 m_e c^2,$$

$$T_{\max} = \frac{2m_e c^2 \beta^2 \gamma^2}{1 + 2\gamma \frac{m_e}{m_0} + \left(\frac{m_e}{m_0}\right)^2}.$$

Here  $N_A$  is Avogadro number,  $r_e$  is the classical electron radius,  $m_e$  mass of electron,  $Z$  is atomic absorber number,  $A$  is the atomic mass of absorber,  $\beta = v/c$ , and  $\gamma$  is Lorentz  $\gamma$ -factor.  $T_{\max}$  is the maximum kinetic energy that can be transferred to a free electron in a single collision,  $m_0$  is the rest mass of incoming particle,  $I$  is the mean excitation energy, and  $\delta(\beta\gamma)$  is the density effect correction to the ionization energy loss [107]. In the bremsstrahlung effect, the energy loss of charged particles is proportional to their energy. For electron energy above O(100 MeV), the contribution of bremsstrahlung processes dominates over ionization. The bremsstrahlung intensity is inverse proportional to the squared mass of the particle; for that reason, it's stronger for light particles, like electrons. The energy loss of electrons through bremsstrahlung is described by

$$\left. \frac{dE}{dx} \right|_{\text{brems}} = 4\alpha N_A \frac{Z^2}{A} r_e^2 E \ln \left( \frac{183}{Z^{1/3}} \right). \quad (4.10)$$

Here  $\alpha$  is the QED fine structure constant,  $N_A$  is the Avogadro number,  $Z$  is the atomic absorber number,  $A$  is the atomic mass of absorber,  $r_e$  is the classical electron radius, and  $E$  is the energy of the electron.

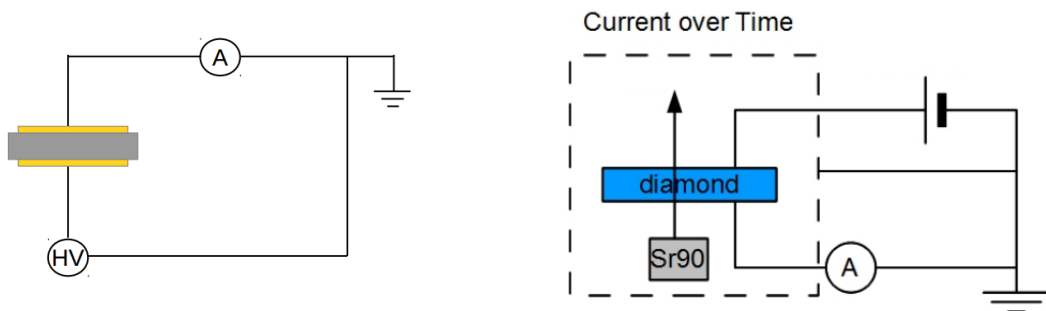
The shape of the energy loss of MIPs is normally distributed around its mean value forming a Gauss distribution; however, a contribution from  $\delta$ -electrons<sup>1</sup> extends the distribution towards higher energy losses. Due to this effect, the signal is often described using Landau distribution or with a convolution of Landau and Gauss distributions.

### Leakage current estimation

The leakage current typically appears in a non-ionized sensor under an applied voltage due to the presence of free charges and is considered as a background. Diamond is an insulator, and pure samples may have a minimal leakage current; however, crystal defects (especially in pCVD diamonds) and admixtures may lead to deformed energy bands, causing increasing leakage current. An estimation of leakage current, further referred to as IV measurement, is obtained measuring the electric current,  $I$ , that appears in the sensor, with the applied voltage,  $V$ . Selected sensors are required to have a leakage current of at most 1 pA for up to 1 kV applied bias voltage; for that reason, measurements are done with a ramping voltage from 0 to 1000 V. To accomplish such measurement, a simple scheme, shown in Figure 4.9a, is used. In this scheme, the sensor is mounted in a custom plastic frame that has individual cables with metal strips, which are soldered to tiny metal bonds; the bonds are brazed to metallization pads on each side of a sensor, as shown schematically in Figure 4.10. A special bond is placed to connect two parts of the two pad metallization side; this is done to apply a uniform electric

<sup>1</sup>Electrons, products of ionization, which has a high enough energy to become a source of ionization.

field between the total area of metallization pads. Depending on the sensor polarity, the



(A) Electric scheme of IV measurement [99]. (B) Electric scheme of CT measurement [108].

FIGURE 4.9: Electric schemes for IV and CT measurements. Generally, the schemes for the IV and CT are the same, except that the CT has a source of  $\beta^-$  decay  $^{90}\text{Sr}$ . HV is a power supply; the gray and blue rectangles are a diamond sensor, metallization pads are shown with a yellow color. A is ampere meter, and earth ground is shown as a dashed triangle.

leakage current can be different. The positive applied voltage defines positive polarity and vice versa. The polarity can be determined by applying the voltage in the range  $-1000$  to  $+1000$  V. During such measurements, the voltage was ramped in few schemes, with 50 V steps and 120 sec delay between steps, 20 V step with 40 sec delay, and others. Some selected results are shown in Figure 4.11.

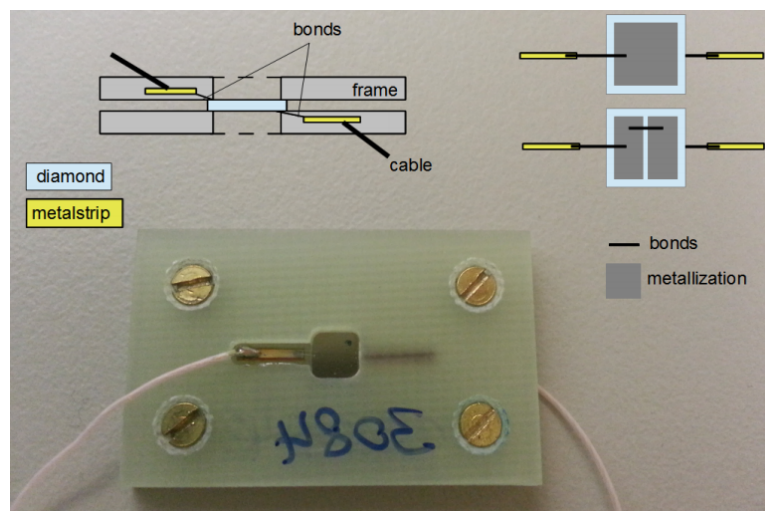
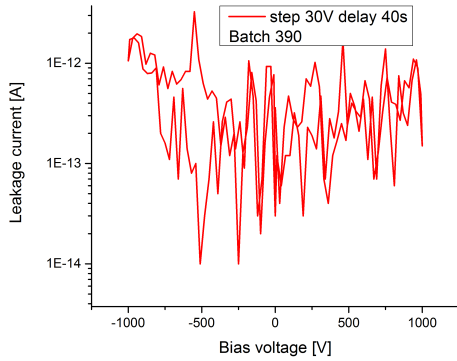
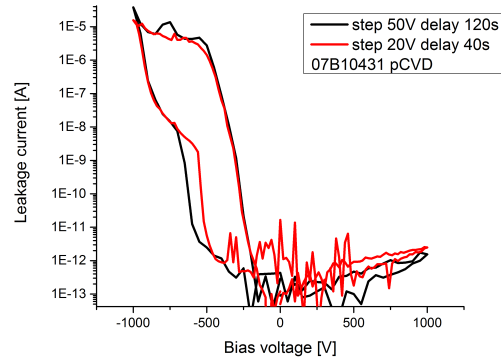


FIGURE 4.10: A diamond sensor in a plastic frame that is used in IV and CT measurements. Schematic pictures show the binding of a diamond sensor with the metal strips [108].

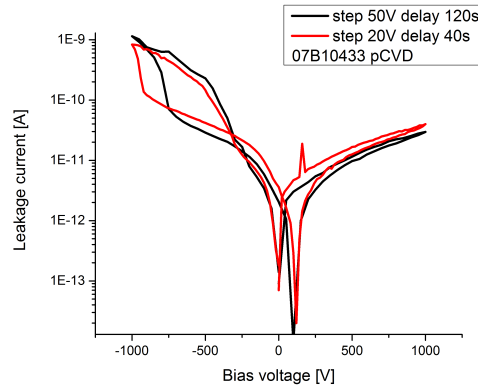
Three different cases are presented: an excellent sensor with very low leakage current, a sensor that can be used only with a specific polarity, and a sensor with unsatisfactory leakage current. The first sensor among presented is sCVD "Batch 390", shown in Fig. 4.11a. The measurement is done in 30 V steps with 40 sec delay. The current varies from  $8 \cdot 10^{-11}$  -  $10^{-14}$  A, such stability makes "Batch 390" one of the best sensors



(A) Good sensor with low leakage current in both polarities.



(B) High leakage current in negative polarity and low in positive.



(c) Bad sensor with high leakage current.

FIGURE 4.11: Selected results of IV measurement for sCVD sensor "Batch 390" (plot 4.11a), and pCVD sensors 07B10431 (plot 4.11b), 07B10433 (plot 4.11c). The voltage is applied in order 0 - +1000 - 0 - -1000 - 0 V.

tested. Figure 4.11b shows the result for pCVD sensor 07B10431. It was measured in two different regimes, 50 V steps with 120 sec delay (black line), and 20 V steps with 40 sec delay (red line). The sensor shows high leakage current in negative polarity regime, while the positive polarity remains stable, in the range  $10^{-11} - 10^{-12}$  A. Sensors with similar behavior are marked as satisfactory with a remark on polarity. Figure 4.11c represents the result for pCVD sensor 07B10433 measured in the same regimes as previous. That sensor has a high leakage current in both polarities. Sensors with such IV characteristics are considered as low-efficient and are removed from the selection.

### Current over time measurement

Another essential characteristic of a diamond sensor is the stability of a sensor current in time. The measurement is done using the same electric scheme as for IV, but with a source of ionization, see Fig. 4.9b. As the radiation source  $^{90}\text{Sr}$  is used. The experimental setup is the same as for the IV measurement, including the sensor frame, see in Fig. 4.10. Measurement of current over time, further referred to as CT measurement, is studied by tracking electric current values at different applied voltage

with different polarity<sup>2</sup> during a more extended amount of time (typically 3-4 hours). The main criterion of a suitable sensor is the stability of the current over time at all applied voltages. Some selected results of the CT measurements are shown in Fig. 4.12. Figure 4.12a shows CT result of pCVD sensor 07B10436. Studies were performed at 200 V, 500 V, 750 V, and 1000 V in negative polarity. The observation time is one hour

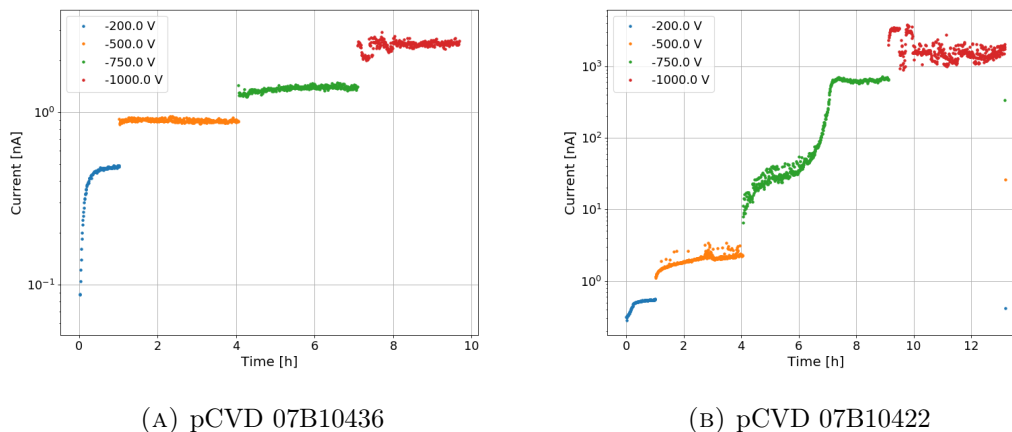


FIGURE 4.12: CT measurements of 07B10436 (plot A) and 07B10422 (plot B) pCVD sensors.

for 200 V and three hours for each of the rest applied voltages. In all measurements, the electric current value was recorded approximately every 20-60 seconds. The sensor shows apparent stability at all applied voltages<sup>3</sup> and is an example of an excellent CT measurement. In contrast, Fig. 4.12b represents an unstable behavior of pCVD sensor 07B10422. Measurements for this sensor were performed during 1, 3, 5, and 4 hours for -200, -500, -750, and -1000 V, respectively.

The sCVD diamonds showed different CT results than pCVD. Figure 4.13 shows the CT measurements of sCVD "Batch 390" (Fig. 4.13a) and "Number 5" (Fig. 4.13b). Measurements are performed for one hour at 200 V and three hours at 500 V, 750 V, and 1000 V, each in negative polarity. The current values are very stable and are almost the same at any applied voltage. This is different from pCVD, where the current values are different at different bias voltages. On average, sCVD sensors have a lower electric current than pCVDs.

### Measurement of charge collection distance

The Charge Collection Distance (CCD) is a sufficient distance in a sensor from which a potential charge carrier can be registered. This characteristic is closely related to a charge collection efficiency (CCE):

$$\text{CCD} = \text{CCE} \times d, \quad (4.11)$$

<sup>2</sup>It depends on the sensor behavior obtained from IV measurements.

<sup>3</sup>Fast current increase at -200 V is interpreted as a sensor's "warm-up" and, in this case, is not considered as instability.

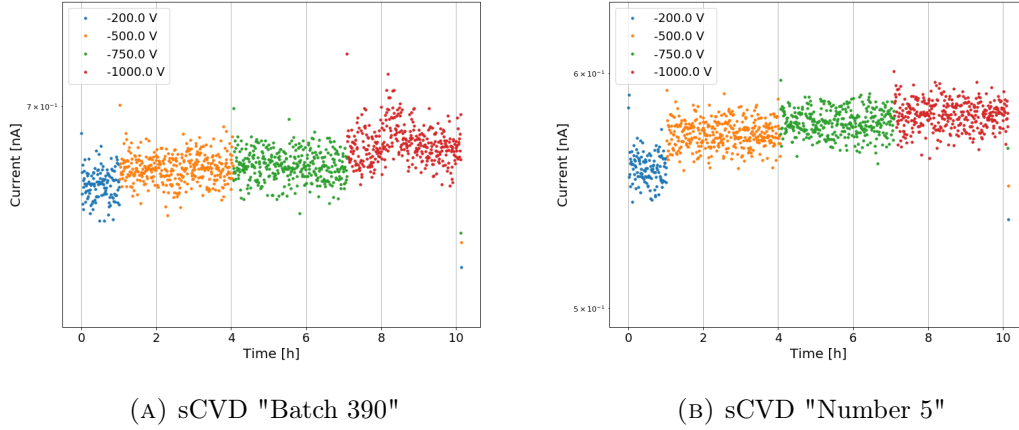


FIGURE 4.13: CT measurements of "Batch 390" (Fig. A) and "Number 5" (Fig. B) sCVD sensors.

where  $d$  is the diamond thickness. The CCE is defined as the ratio of the collected charge,  $Q_{\text{col}}$ , over induced charge,  $Q_{\text{ind}}$ :

$$\text{CCE} = \frac{Q_{\text{col}}}{Q_{\text{ind}}}. \quad (4.12)$$

The collected charge is the charge that reaches the surface of the sensor after ionization and is being registered using read-out electronics. The induced charge is an average amount of created electron-hole pairs during ionization, in other words - a charge produced from ionization, potentially available for collection. Due to various effects in a diamond sensor, a certain amount of produced electron-pairs is never registered. The discrepancy between collected charge and potentially available charge reflects the CCE of a diamond sensor. The induced charge is defined as:

$$Q_{\text{ind}} = \rho_{\text{MIP}} \cdot d, \quad (4.13)$$

where  $d$  is the sensor thickness,  $\rho_{\text{MIP}} = 36 (\mu\text{m})^{-1}$  is the average amount of created electron-hole pairs in diamond per  $\mu\text{m}$  [109]. In this study, CCD was measured as a function of time and as a function of different bias voltages. Measurements were performed using a "beta-setup box". In this setup, the sensor is placed on a high voltage plate on top of the collimated source of  $\beta^-$  decay,  $^{90}\text{Sr}$ . As signal readout, a thin iron needle is placed on top of the sensor (see Fig. 4.14). Two scintillators are placed above the sensor; their assignment is to provide a signal for a gate trigger (more details see below).  $^{90}\text{Sr}$  radiates electrons that pass through a diamond sensor creating electron-hole pairs. The applied high voltage creates an external field that allows collecting the signal with the needle. The signal is amplified with a pre-amplifier, which has a test capacitance, that allows a charge calibration. Hence, the signal response,  $Q_{\text{calib}}$ , is measured as a function of a voltage step,  $\Delta V$  of a test capacitance input,  $C_{\text{T}}$ :

$$Q_{\text{calib}} = C_{\text{T}} \cdot \Delta V. \quad (4.14)$$

Afterward, the signal goes to a shaper with 100 ns shaping time [110].





FIGURE 4.14: A diamond sensor is in the opened "beta-setup" box before the measurement. The needle is on top of the diamond sensor, which is mounted in the center of the high voltage golden plate. The black hole in an iron pyramid is a dry-air source (5% oxygen + 95% dry nitrogen). During the measurement, this scheme is covered with a black wooden box. Inside on top, the box has two mounted layers of scintillation detectors.

At the next stage, the signal is fed to Fan-In Fan-Out (CAEN Model N625 [111]), where it duplicates; one of the signals goes as an input in ADC while the other is being delayed for 400 ns to be used in ADC as a pedestal<sup>4</sup>, see red and blue signal chains in Fig. 4.15.

The ADC reads input values only after activation of the ADC gate - a time interval during which the sensor signal is integrated and processed by ADC. The gate activation scheme is arranged as follows. Each time when an electron from  $^{90}\text{Sr}$  ionizes the studied sensor, it is registered with two scintillators, mounted on top of the box, right above the collimator. Transformed electric signals from scintillators are then fed into Low-Threshold Discriminator (LTD), see magenta signal chains in Fig. 4.15. At this stage, LTD defines if the signal is high enough ( $>0.4$  MeV) to be considered as a non-background signal. Afterward, the Coincidence Unit creates a logic AND-signal, which triggers the Dual Timer (DT). DT allows creating a rectangular signal, the gate, with a customizing width and a delay time. For this particular study, the rectangular signal is adjusted to measure the signal input in ADC, without the pedestal interval.

The results of such measurement for pCVD sensor 07B10417 at 500 V are shown in Figure 4.16. The presented results were obtained after 10000 ADC counts. The upper plot shows a clear Gaussian distribution of ADC counts that corresponds to the pedestal channel. The lower plot shows the distribution of signal ADC count with a Landau distribution shape. The signal size can be obtained from histograms using the Most Probably Value (MPV) of the Landau distribution. The pedestal value, Ped, is derived from the mean value of the Gaussian distribution.

<sup>4</sup>Pedestal is used to estimate the noise of the whole readout chain.

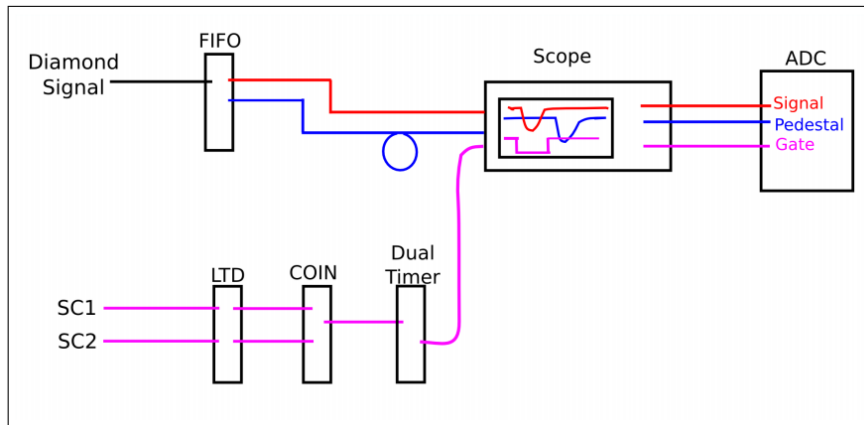


FIGURE 4.15: The ionization signal is registered with the needle and amplified with the pre-amplifier forming a Diamond Signal. FIFO is a Fan-In Fan-Out, Scope is an oscilloscope. SC1 and SC2 - two scintillators, LTD - Low-Threshold Discriminator, COIN - Coincidence Unit. The red line is a signal input in ADC; the blue line is a duplicated signal with a delay (blue circle between the FIFO and the Scope). The magenta line is a gate, formed from two scintillator signals. The Scope shows the opened gate right in the time interval of the signal, excluding the pedestal time interval [99].

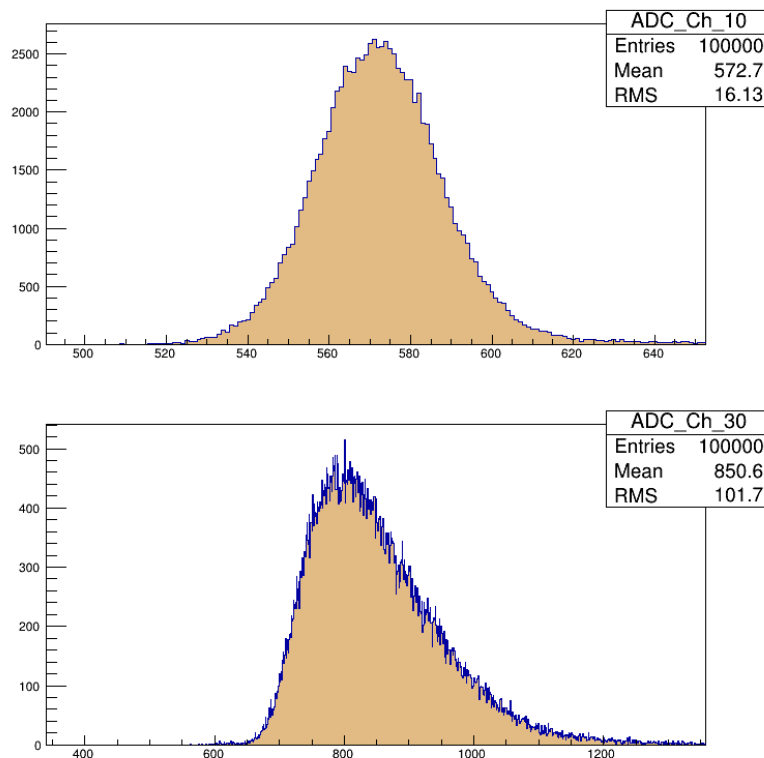


FIGURE 4.16: Two histograms of 10000 ADC counts of pCVD sensor 07B10417 measured at 500 V. The pedestal ADC counts with the Gaussian distribution shape are shown in the upper histogram. The lower histogram shows signal ADC counts with the Landau distribution shape.

The signal charge value,  $Q_m$  can be calculated by subtracting pedestal value from MPV, times calibration factor, Calib:

$$Q_m = (\text{MPV} - \text{Ped}) \cdot \text{Calib}.$$

The calibration, Calib, is defined as the ratio between the charge and the ADC channel. Knowing the capacitance and a charge, Calib can be defined through a fit of ADC channels as a function of charge fed in. Figure 4.17 shows an example of one of the calibration fits, obtained on December 27th, 2016, in the Zeuthen laboratory. The value of the slope of the line,  $k$  is the Calib value. The calibration depends on the performance of electronics, which can have a deviation in time; for that reason, the calibration is recalculated before each set of measurements.

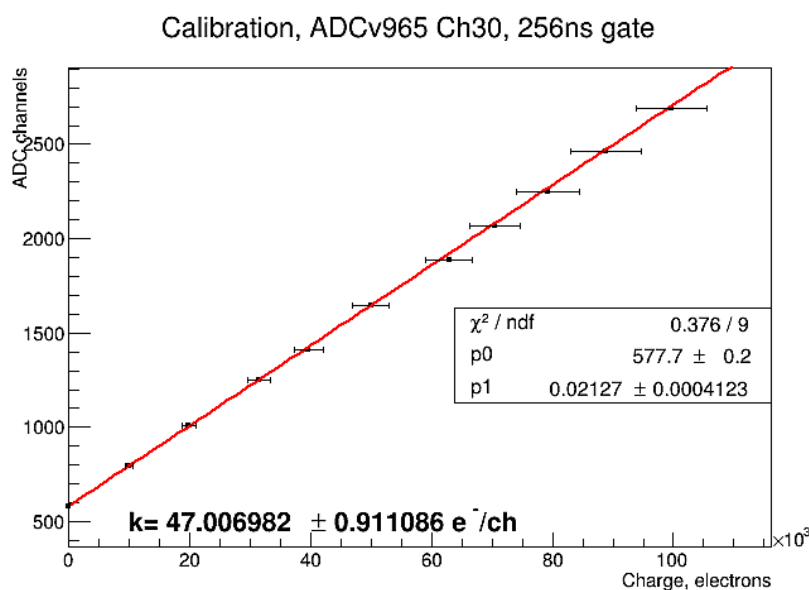


FIGURE 4.17: Calibration line of ADC channels as a function of charge. The Calib value obtained from this fit is  $k=47$ .

The generation of  $\delta$ -electrons can increase the expected number of produced electron-hole pairs. The track length can be a bit higher than the thickness of the sensor. To take these features into account, a certain correction, obtained from a simulation, must be applied. The correction,  $\text{sr}_{\text{corr}}$ , derived from simulations, for  $320 \mu\text{m}$  thick diamond sensor is approximately  $\text{sr}_{\text{corr}} = 1/0.95$  [112]. The charge collection distance, CCD, can be calculated by:

$$\text{CCD} = \text{CCE} \cdot d = \frac{(\text{MPV} - \text{Ped}) \cdot \text{Calib}}{\rho_{\text{MIP}} \cdot \text{sr}_{\text{corr}}}. \quad (4.15)$$

To investigate CCD and its stability in time at different bias voltages, measurements were performed for each pCVD and sCVD sensor at 500 V, 750 V, and 1000 V.

CCE values were measured, counting 10000 ADC counts in a twenty minutes time step during six hours at each bias voltage. Figure 4.18 represents example of such measurement, obtained for two pCVD sensors, 07B10415 and 07B10434 (Figures 4.18a and 4.18b, respectively). Figure 4.18a shows good CCD values at 500 V, 750 V, and 1000 V, while Fig. 4.18b shows lower CCD values. Measurements of sCVD sensors showed different characteristics compared to pCVD sensors. In general, sCVD has higher CCD

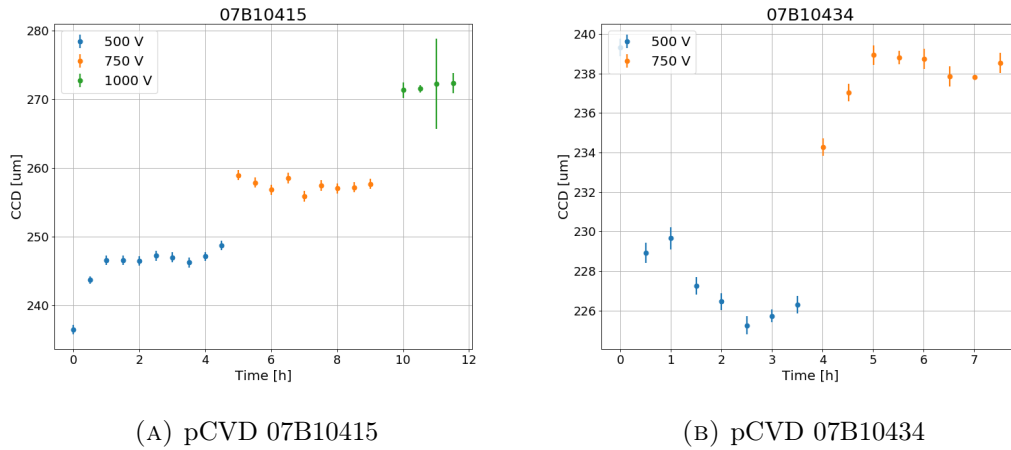


FIGURE 4.18: CCD measurements as a function of time of 07B10415 (plot 4.18a) and 07B10434 (plot 4.18b) pCVD sensors.

and is much more stable for different bias voltages. Results for two sCVD sensors are shown in Fig. 4.19.

Another kind of studies is the CCD stability as a function of bias voltage. In this measurement, CCD values are obtained measuring 10000 ADC counts every 50 V or 100 V starting from 25 V and up to 800 V or 1000 V (depending on sensor type). Figure 4.20 shows the results of such measurement for pCVD and sCVD sensors. The pCVD sensor 07B10415, shown in Fig. 4.20a, is measured starting from 200 V with a step of 100 V up to 800 V. CCD values logarithmically rise from approximately 195 nm at 200 V to 253 nm at 800 V. The sCVD sensor "Batch 390" is presented in Fig. 4.20b. It is measured starting from 50 V with a step of 25 V up to 150 V, and then with a step of 50 V up to 800 V.

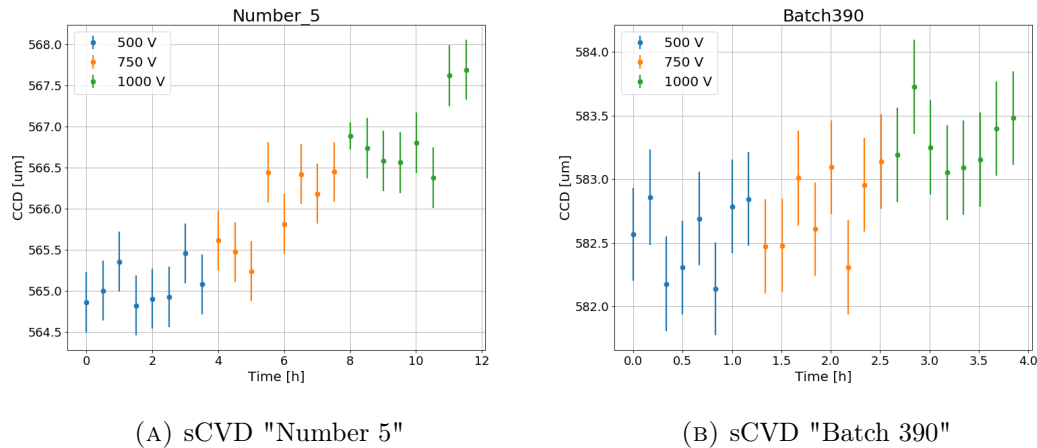


FIGURE 4.19: CCD measurements as a function of time of "Number 5" (plot 4.19a) and "Batch 390" (plot 4.19b) sCVD sensors. Both sensors show high charge collection distance and stability over time for all bias voltages.

The last three measurements are done at 800, 900, and 1000 V. The sensor shows a stable performance with a quick rise in a region of 25 - 100 V.

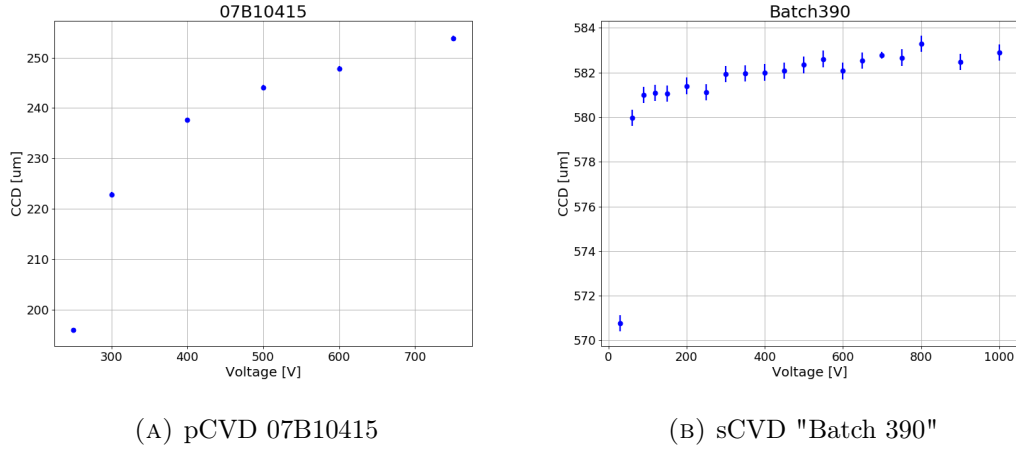


FIGURE 4.20: CCD measurements as a function of applied bias voltage of pCVD sensor 07B10415 (plot 4.20a) and sCVD sensor "Batch 390" (plot 4.20b). Both sensors show a stable logarithmic rise.

### Summary on diamond sensor measurements

The presented measurements allowed to chose the most suitable pCVD and sCVD sensors for the BCM1F upgrade in 2017. More details on sensor measurement studies is given in Appendix A. The location of the selected sensors per C-shape is given in Table 4.1. The main highlights from BCM1F sensors and electronics monitoring during the Run 2 and conclusions on diamond sensors' performance during that period are given in the next section.

"+Z FAR:	"-Z FAR:
1) pCVD "31"	1) pCVD "23"
2) Si "22"	2) Si "8"
3) pCVD "12"	3) pCVD "19"
4) Si "15"	4) Si "20"
5) Si "I-5"	5) Si "3"
6) sCVD "PLT S115"	6) sCVD "S119"
"+Z NEAR:	"-Z NEAR:
1) pCVD "30"	1) pCVD "17"
2) Si "7"	2) Si "19"
3) pCVD "36"	3) pCVD "15"
4) Si "3"	4) Si "17"
5) pCVD "21"	5) pCVD "35"
6) sCVD "5"	6) sCVD "390"

TABLE 4.1: Position of sensors on C-shapes. The numbers in quotes correspond to the personal name of the sensors.

### 4.3 BCM1F status monitoring using VME ADC

BCM1F performance control exists to give a clear picture of how reliable the detector results are. The detector efficiency is expected to decrease due to sensors and electronics degradation from the radiation damage. Also, uncertainties may arise from electronics productivity due to the changes in CMS magnetic field, temperature fluctuation, etc. While the RHU unit is called to calculate online luminosity, the goal of VME ADC readout is to provide the data for the monitor sensor performance and electronics stability. Sensor efficiency is studied in each channel, analyzing the amplitude spectra of particle hits. Such measurements allow to track spectra degradation in time and, when it's possible, to take actions to improve sensor performance. Electronics behavior is observed separately from the sensors' response, using a Test Pulse (TP).

#### 4.3.1 VME ADC data processing

The BCM1F status monitoring starts from the VME ADC raw data processing, which is a record of ADC counts as a function of time. The timeline ends when a bunch makes a full circle in LHC, approximately in  $90 \cdot 10^{-6}$  seconds. Such data, recorded per one ring, is called "orbit". Figure 4.21 shows an example of recorded orbit. The first signal belongs to a TP, and the rest are bunches of the fill. Bunches are gathered in groups called "trains". The so-called "pilot" bunch train follows the test pulse and consists of 12 bunches. The rest are twelve trains with 70 bunches in each.

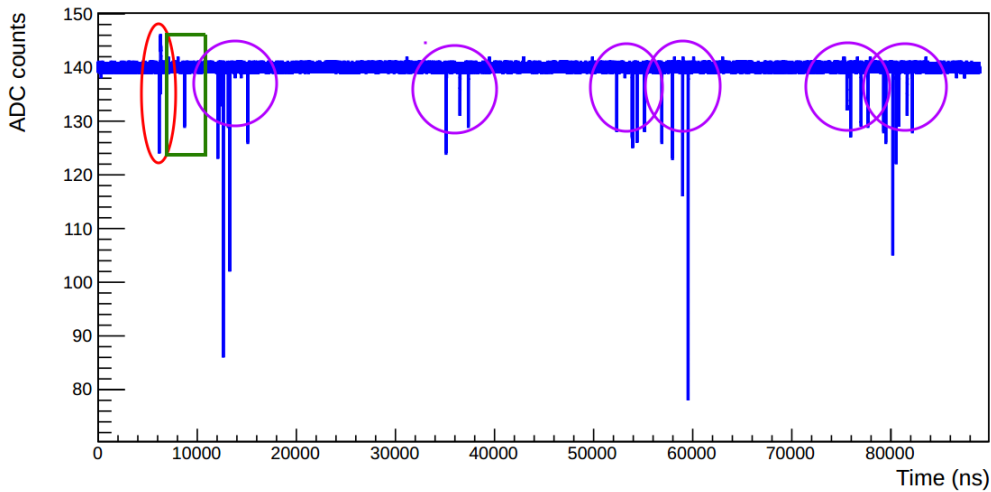


FIGURE 4.21: Recorded orbit from the LHC Fill 4364 with 852 colliding bunches [87]. The baseline is set on 140 ADC counts, signals are shown as dips in a baseline position. Red ellipse shows a TP, green rectangle corresponds to the "pilot" train, and each magenta circle is a train of bunches.

The arrival time is obtained by applying a 3 ADC counts threshold to the orbit. An example of acquired arrival time distribution is shown in Figure 4.22, where the TP signal is cut out, the "pilot" train is first, and the rest are twelve bunches grouped in two.

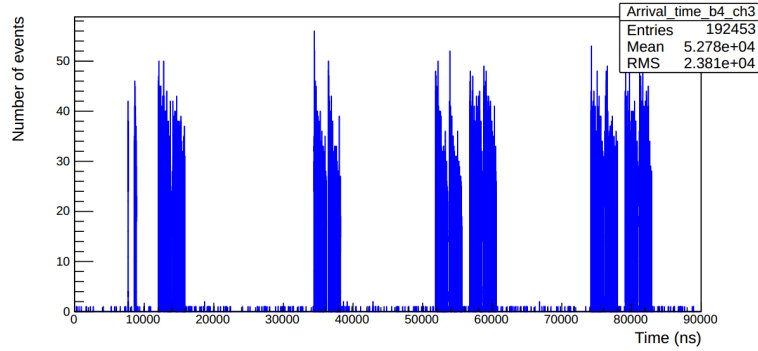
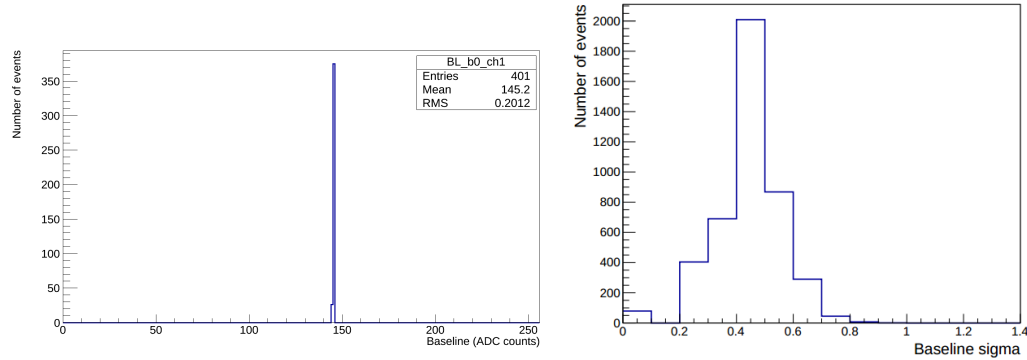


FIGURE 4.22: Arrival time distribution [87].

### 4.3.2 Baseline position

Baseline monitoring helps to control the readout chain stability. The baseline calibration is done for each channel individually; actual values are determined from the data samples taken during 40 ns in the abort gap [54]. An example of the baseline distribution is shown in Figure 4.23a.



(A) An example of the baseline position distribution for channel one from a random fill.

(B) An example of the baseline standard deviation distribution for channel one from a random fill.

FIGURE 4.23: Baseline distributions for channel 1 taken from one of the first fills. Plot 4.23a corresponds to the baseline position distribution, while plot 4.23b shows distribution of its standard deviation.

The definition of the standard deviation of the baseline distribution is defined as [93]:

$$BL_{\text{sigma}} = \sum_{i=1}^N \sqrt{\frac{(BL_{\text{average}} - BL_i)^2}{N}}.$$

Here  $BL_{\text{average}}$  is the averaged baseline value,  $BL_i$  the  $i$ -th baseline value,  $N$  is a number of samples. The baseline standard deviation plays a vital role in orbit data analysis because the threshold value is estimated taking into account the baseline standard deviation. An example of baseline standard deviation is shown in Figure 4.23b.

### 4.3.3 Test pulse spectra characteristics

The TP is fed in a special ASIC input with a fixed amplitude and arrival time. Usual average TP amplitude is around 70 mV and time window less than 20 ns. Fixed arrival time allows for tracking fluctuations in signal timing, while deviations from the constant amplitude show signal distortion. During the analysis of the orbit data, the TP is cropped out using known arrival time. Depending on the distribution of the TP arrival time, the cut-out window could be up to 100 ns. Figure 4.24 shows an example of the zoomed TP signal. Arrival time is defined as a time that corresponds to a half of the signal.

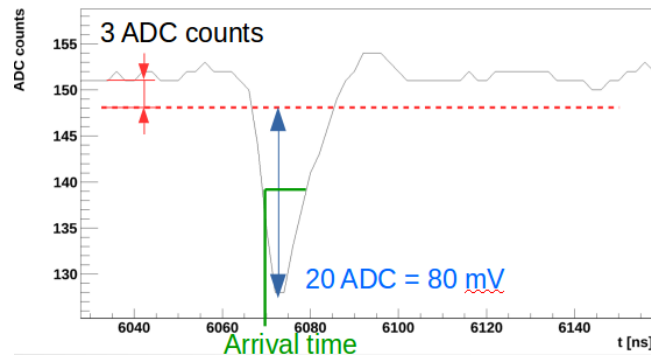
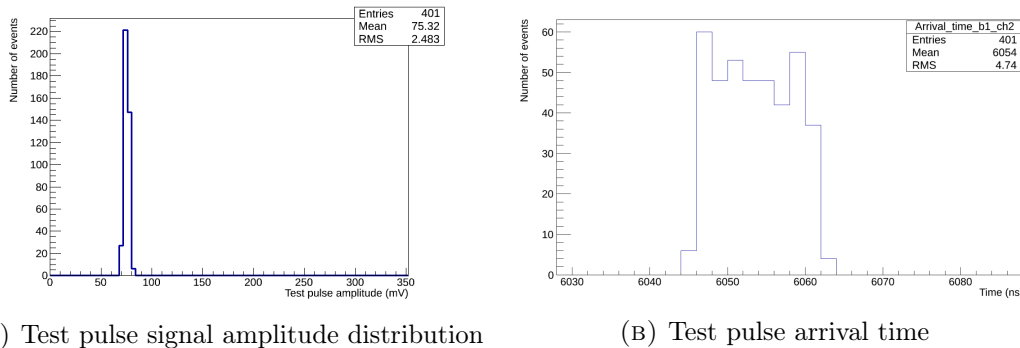


FIGURE 4.24: The zoomed test pulse signal is a dip in the baseline, defined between 6070 and 6080 ns. The red dashed line corresponds to the baseline threshold, which is 3 ADC counts. The blue line is a signal value, which corresponds to 20 ADC or 80 mV. The green line shows which part of the signal is taken as the arrival time.

An example of the TP signal amplitude distribution is shown in Figure 4.25. Figure 4.25 A presents the position of the TP collected from 401 orbits. Two hundred twenty orbits have a TP with the amplitude near 80 mV, 150 with 84 mV, 25 with 76 mV, and 10 with approximately 88 mV. Figure 4.25 B shows how the test pulse is distributed in time. Ideally, TP is expected to arrive at one time for in all orbits, but due to various reasons, its arrival time shifts. More details on usage and tracking of such distributions will be given further in this chapter.



(A) Test pulse signal amplitude distribution

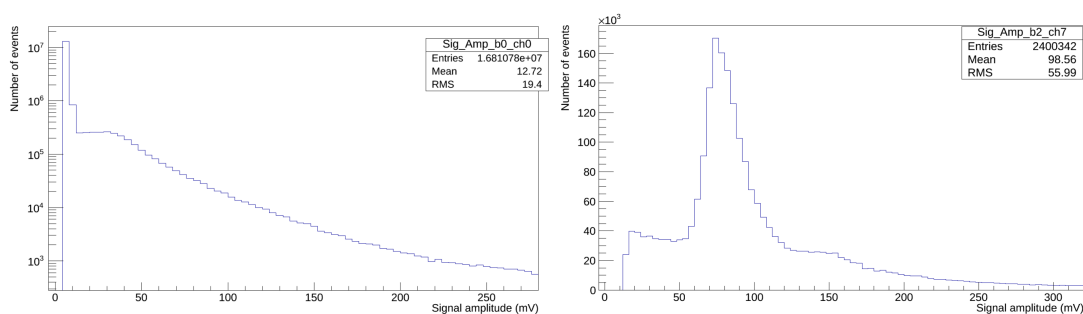
(B) Test pulse arrival time

FIGURE 4.25: Test pulse distributions for channel 2 taken from one of the first fills, obtained after processing 401 orbits. Plot 4.25a shows distribution of TP signal amplitude. Plot 4.25b shows distribution of TP arrival time.



### 4.3.4 Signal spectra characteristics

Signal recording starts if it is below the threshold value longer than six ns. The time count starts from the beginning of the signal and is recorded until it reaches the threshold value. A signal amplitude distribution is built using collected info from many orbits. An example of such spectrum is shown in Figure 4.26. The distribution in plot 4.26a belongs to the pCVD sensor "07B10430", the first two bins correspond to a noise, a so-called "pedestal". An important part is to separate it from the signal; for that reason, the signal amplitude should be increased. Hence, a bias voltage applied to the sensor is increased in a way to keep CCE near 100%. After that, the pedestal is removed, applying a slightly higher threshold, which can be different for different sensors, especially after extensive radiation damage. An example of a clean peak of sCVD "S115" is shown in plot 4.26b.



(A) Signal amplitude spectra of pCVD sensor 07B10430 with a threshold of 1 ADC. (B) Signal amplitude spectra of sCVD sensor S115 with a threshold of 3 ADC.

FIGURE 4.26: Signal amplitude spectra taken from Fill 5722 of pCVD sensor 4.26a in log scale, and sCVD sensor 4.26b in linear.

## 4.4 BCM1F performance

The BCM1F status was monitored during the whole Run 2 of LHC. Each time when the LHC beam status "Stable beams" is declared, a special program, starts a timer. After 10 minutes, all 48 ADC channels are enabled to record the orbit data into a file with a ROOT extension (a root file). Recording performs up to 8000 orbits, resulting in three data files. During each fill, this data acquisition cycle repeats every three hours.

Recorded root files are then analyzed by a special algorithm, briefly described in sections 4.3.3 and 4.3.4. The algorithm creates plots with distributions shown in Fig. 4.23, 4.25, and 4.26. At this stage, various improvements were implemented to increase the BCM1F performance monitoring.

### 4.4.1 BCM1F monitoring upgrades

**Reference histogram implementation.** Sensor and electronics radiation damage finds its reflection in shape changes of the distributions of the signal amplitude, baseline, baseline standard deviation, and a test pulse. Observing the shape changes became possible after the implementation of a reference histogram, taken from a previous fill where considered characteristics can be observed very clear. The reference histogram was added into the algorithm to be plotted together with every newly analyzed signal

amplitude distribution. During Run 2, the reference histogram was changed a few times, mainly due to a change of the bias voltage at which sensors are operated. Reference fills were chosen to be 5730, 5840, and 6620.

Figure 4.27 shows an example of such implementation. Figure 4.27 A corresponds to the signal amplitude distribution of channel 0, for pCVD 07B10430. The blue histogram corresponds to actual fill 6570 in 2018, with a maximum bin at 18 mV, while the red histogram is a reference from fill 5840 in 2017, and the gray histogram is a reference from fill 6620 in 2018. It's clearly seen that the peak shifts toward lower values compared to 2017.

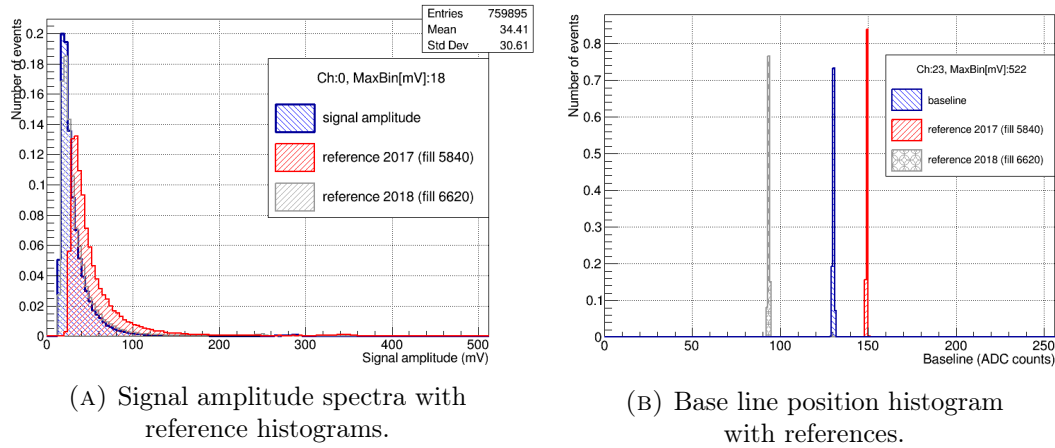


FIGURE 4.27: Signal amplitude spectra with reference histograms (plot A), base line position histogram with references (plot B).

The reference histogram was also implemented for a test pulse spectra and a baseline position distribution. Figure 4.27 B shows the baseline position for channel 23. Here the reference histogram from 2017 is at 150; reference from 2018 is at 90, and the actual baseline position is at 130 ADC counts.

**Implementation of actual high voltage.** New sensors have started their operation at an applied bias voltage of 500 V for pCVD and 750 V for sCVD. Due to radiation damage, individual sensors could have a sudden so-called "erratic current" - an increase of a signal current, which does not originate from the ionization processes. In this case, the voltage on that sensor switches off, and when the bias voltage is applied again, it is lowered to a reasonable value. For that reason, it's important to track at which bias voltage the signal amplitude and other characteristics were measured. The HV values of each sensor were recorded from the Data Interchange Protocol (DIP) [113]. Before the algorithm starts orbit data analysis, the received HV values from the DIP are stored to be recorded on a legend of each plot for each sensor. An example is shown in Fig. 4.28, the reference signal amplitude was collected at 500 V, while the actual distribution is collected at 450 V.

**Most probable value extraction.** The most probable values (MPV) of the signal amplitude distribution and TP are used to track sensor and readout-electronic performance, as it will be shown in the next paragraph. Special software, Web monitor, collects these values from each fill to record a behavior chronology of each sensor.

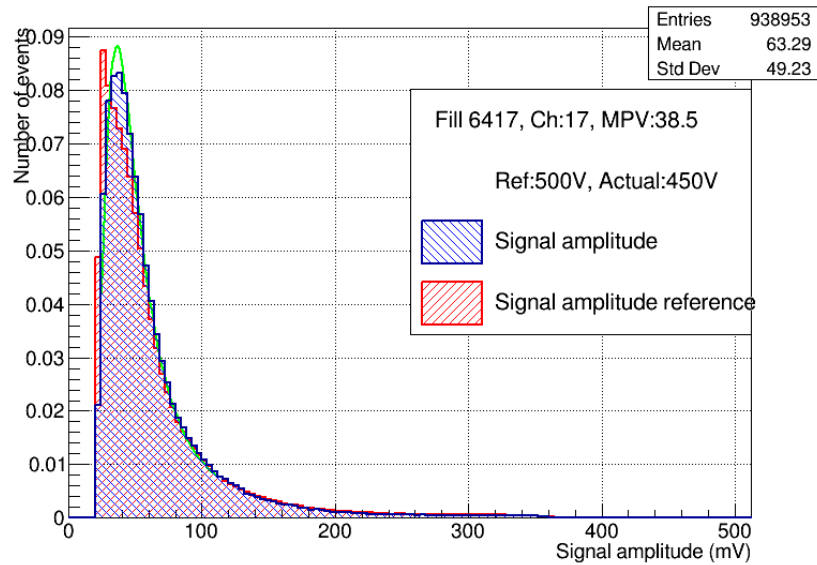


FIGURE 4.28: Signal amplitude distribution from channel 17 (pCVD 07B10512), collected in Fill 6417.

An MPV value of the signal amplitude distribution is derived from fitting Landau distribution to actual data distribution. An example can be seen in Fig. 4.28, where the Landau fit is shown with a green line, and MPV is equal to 38.5.

An MPV of the TP is derived from the fit using the Gaussian function.

**Web monitoring tool.** A visualization of the BCM1F online monitoring is implemented using a dedicated server: <http://srv-s2d16-22-01.cms/webmonitor/bcm1f-adc>, a so-called "Web monitor" [114]. An example test pulse spectrum provided by the Web monitor is shown in Fig. 4.29.

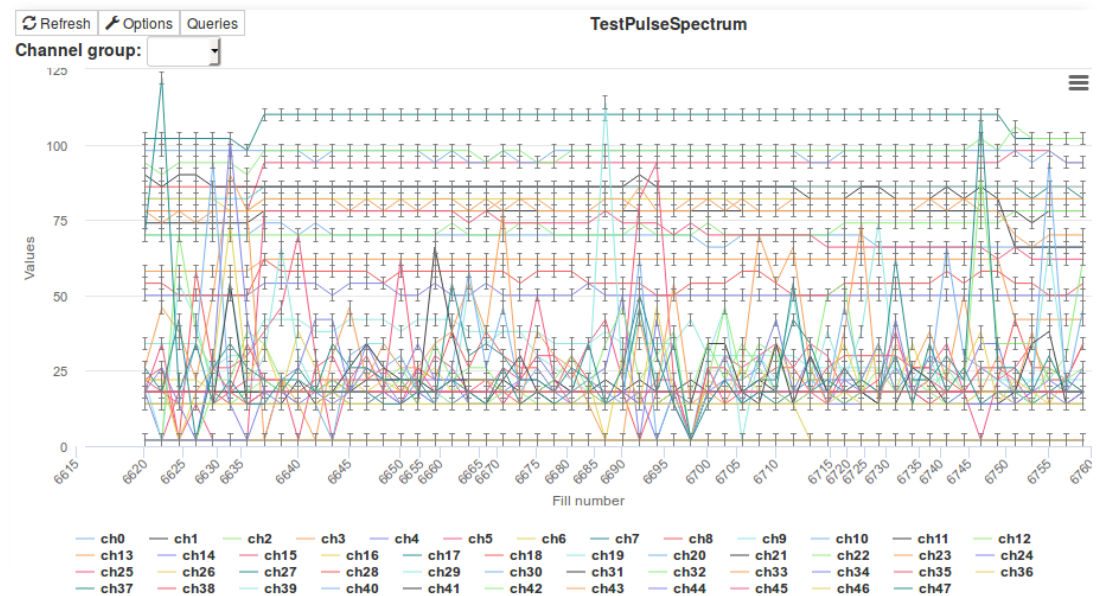


FIGURE 4.29: Test pulse spectrum in Web monitor tool.

In this plot, the MPV values of a test pulse are given as a function of a fill number. Different colored lines correspond to different BCM1F channels. The plot shows results for fill numbers from 6620 to 6760. Additional options include channel grouping, channel selection, etc. Similar distributions are available for the signal amplitude, baseline position and baseline standard deviation. The signal amplitude distribution like in Fig. 4.28, together with the TP spectrum and baseline position plots, are also available via the Web monitor.

#### 4.4.2 BCM1F performance during the Run 2

BCM1F was in operation since its re-installation during a technical shut down in 2017 and until fill 7334 on 23rd of October in 2018, when the last data was recorded. Most of the time, it was used only as a beam-induced background monitor, and only occasionally as the main luminometer. The main problem with the luminosity measurement is that diamond sensors were not able to handle such high rates produced by the LHC, especially for the 25 ns operation period with dense filling schemes. With time, the high rates have damaged pCVD and sCVD sensors in a way which created very weak signal or erratic behavior. The main aspects of sensors, test pulse, and baseline degradation are given below.

##### Sensor degradation

The sCVD and pCVD sensors, due to differences in the crystal structure, had a different response to the radiation damage. Results for each of them are given separately.

**sCVD sensors.** The sCVD sensors, already during the testing stage, showed excellent characteristics and promising results. Mostly due to a minimal amount of defects in their crystal structure, comparing to the pCVD sensors. However, during the run period, they showed a significant efficiency loss, proportional to an ionization rate. In other words, the sCVD sensors can not be used for high intensities measurement. Figure 4.30 shows signal amplitude spectra for the sCVD "Number 5", one of the most outstanding sensors according to the test measurements performed in DESY-Zeuthen.

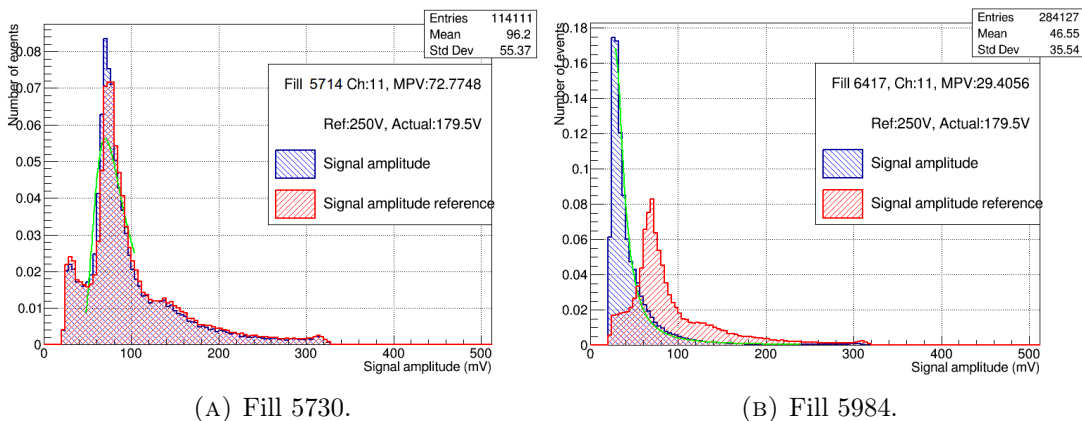


FIGURE 4.30: Signal amplitude spectra of sCVD "Number 5", channel 11. Plot A shows distribution obtained from fill 5730, 30.05.2017. Plot B is taken from fill 5984, 24.07.2017.

Figure 4.30 A was taken at the beginning of the operation period (fill 5730). The signal has a very clear peak, and the second small peak corresponds to events with a double ionization. The third peak, in the end, is a result of the baseline position limitation. Figure 4.30 B was taken during the fill 5984, two months after Fig. 4.30 A. The signal is completely shifted into the pedestal, and can not be used. Normalized results on the time behavior as a function of a fill number, for the sCVD sensors "Number 5" and "Batch390", are shown in Fig. 4.31. The fill number range is 5730 to 6400. Figure 4.31 A shows the sCVD sensor, "Number 5". It can be seen how the MPV of signal amplitude in both channels has lost  $\approx 60\%$  in the fill range from 5840 to 5980. The test pulse slowly decreases due to the radiation damage of ASIC (more details will be given in the next paragraph). Figure 4.31 B shows the sCVD sensor "Batch 390" which, in general, repeats behavior of "Number 5", losing  $\approx 70\%$ . The rest of the sCVD sensors ("S115" and "S119") exhibit the same results.

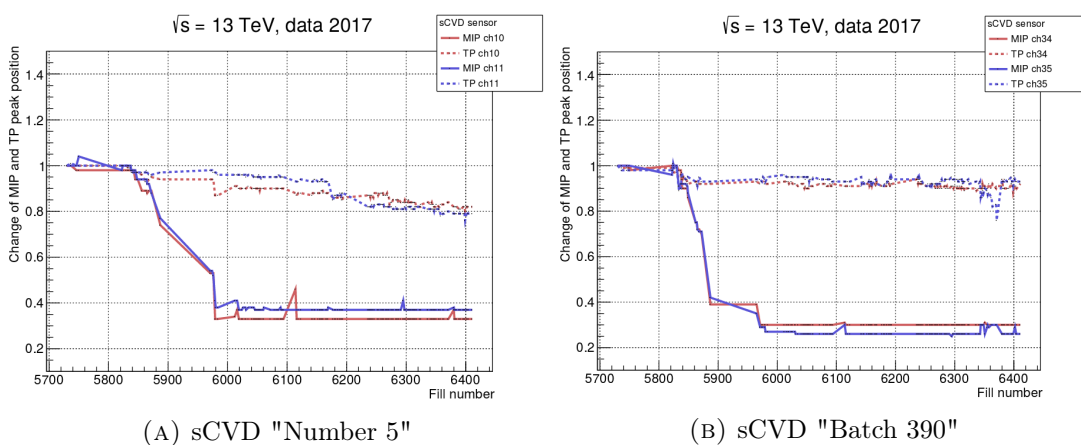


FIGURE 4.31: Signal amplitude and TP MPV values as a function of a fill number for sCVD "Number 5" (plot A) and sCVD "Batch390" (plot B). Each plot shows results from both channels, in different colors. The solid line corresponds to MPV values taken from signal amplitude distribution, while dashed lines correspond to MPV taken from test pulse spectra.

**pCVD sensors.** As was already mentioned, comparing to the sCVDs, the pCVD sensors have more defects in the diamond lattice, which has its advantages and disadvantages. During the operational period, the pCVD sensors also showed efficiency loss due to the hits rate, but not as high as the sCVD. The efficiency loss originates from a simultaneous emergence of two effects, sensor polarization, and "erratic dark currents".

The polarization effect appears due to a large number of locations in sensor volume, which restricts the motion of charge carriers. Such locations are called traps, and they exist as a result of the diamond lattice defects. They also can appear due to radiation damage. When a pCVD is ionized, some amounts of electrons and holes are produced. Since the sensor is under a bias voltage, the charge carriers start moving to electrodes, near which they can get stuck in a trap. After some time, the amount of trapped electrons and holes is high enough to create their own space-charge distribution that reduces applied bias voltage. Polarization effects can be reduced, maintaining a high bias voltage.

The so-called "erratic dark currents" (or "erratic currents"), which are a sudden increase of a signal without appropriate ionization, are not clearly understood yet. It is known that silicon sensors suffer from this effect, and sensor cooling is used to reduce this effect. The pCVD sensors may also experience it, but less than the Si, and magnetic field presence, oriented perpendicular to a sensor plane, increases current-over-time stability of the pCVD. The erratic currents emerge after some time only above a certain applied voltage. Results of CT measurements at -750 V, shown in Fig. 4.12 B, can be interpreted as the erratic behavior.

With radiation damage, a pCVD sensor loses its current-over-time stability; it means that with a high voltage, the erratic currents start to appear more often. The erratic behavior can be improved by lowering a bias voltage, and at some point, contributions from polarization and erratic currents begin to overlap. From one side, the sensor polarization increases, and the bias voltage increase is the solution, from another - the higher voltage leads to erratic currents. As a result, the collection of smaller, well-understood signals is better than not understood and chaotic signals. For that reason, a life cycle of a pCVD sensor consists of a succession of the erratic signals and the bias voltage decrease. During the running cycle, the pCVD sensors lasted longer than the sCVD sensors but were almost never used as a main luminometer.

An example of the pCVD sensor 07B10436 behavior is shown in Fig. 4.32. Figure 4.32 A shows the signal amplitude distribution of channel 5, taken from fill 5737. Here the actual peak is measured at 500V; it has a very clear form and is well fitted, almost the same as the reference histogram. Figure 4.32 B is taken from fill 7090; it shows the same channel of the same sensor at an applied bias voltage of 79.5 V. The peak has shifted into the pedestal, such shape is not valid for usage.

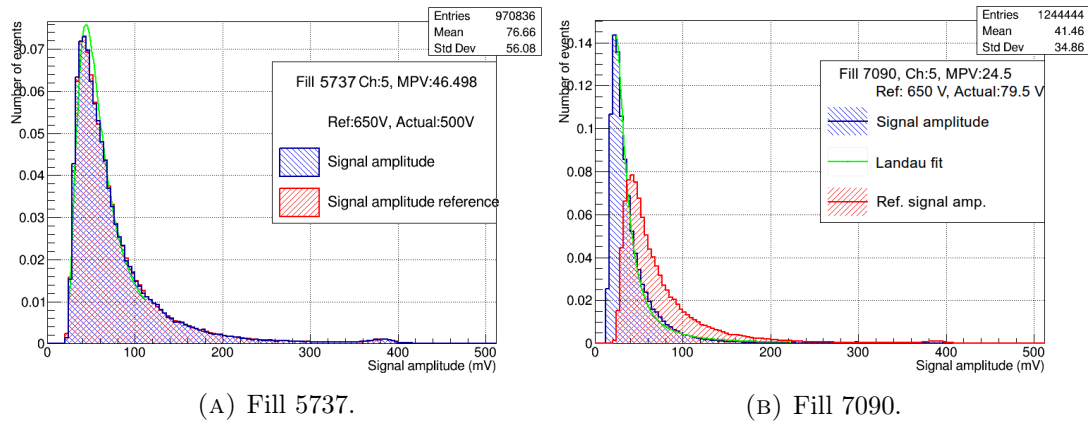


FIGURE 4.32: Signal amplitude spectra of pCVD sensor 07B10436, channel 5. Plot A shows distribution obtained from fill 5737, 03.06.2017. Plot B is taken from fill 7090, 26.08.2018.

During the running operation, various inefficiencies were corrected and improved. One of such examples is a problem with AOH bias, which appeared during fill 5849, on 20-21.06.2017, on channel 24. Figure 4.33 shows how signal amplitude distribution has changed for 18 hours. The plot in the middle shows a distribution taken at 15:40:02 UTC; the shape is stable and coincides with its reference histogram. The plot on the right is the same distribution but taken at 18:05:01 UTC. The shape is distorted; signal MPV moves to higher values, baseline edge peak rises. The left plot is taken at 09:05:02

UTC 21.06.2017; the shape is very distorted, the MPV moved to higher values, the baseline edge peak is big. The problem was cured by resetting (increase ) of the AOH bias.

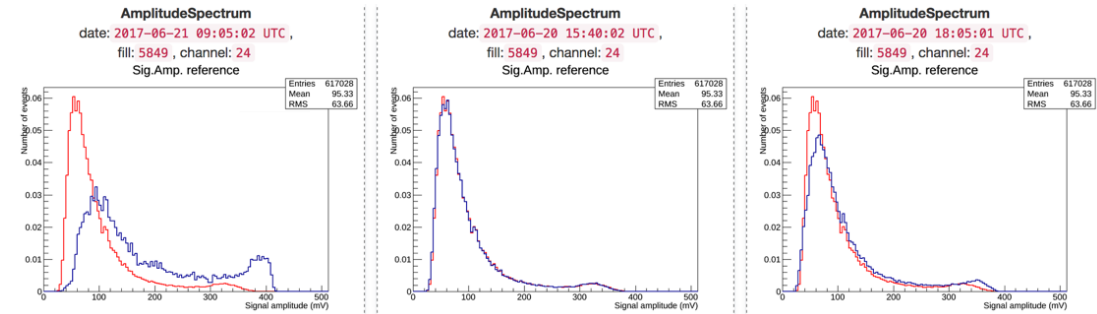


FIGURE 4.33: Signal amplitude spectrum in Web monitor tool.

### TP degradation

During the running period, the test pulse of each channel showed a decrease of its signal and various shifts in the time window. This is mainly caused by degraded electronics, in particular ASICs, which are installed on C-shapes close to the sensors. Due to its location, the electronics suffers from the same radiation damage as the sensors. Figure 4.34 shows an example of a change of the TP taken from fill 6417. The normalized TP signal amplitude value has lowered from  $\approx 100$  mV, shown with reference histogram in red, to  $\approx 60$  mV, shown in blue.

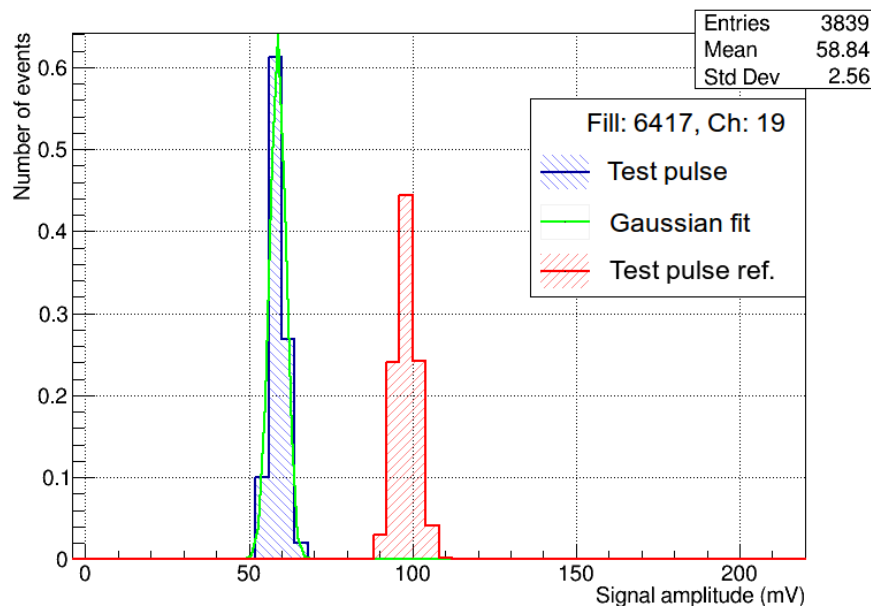
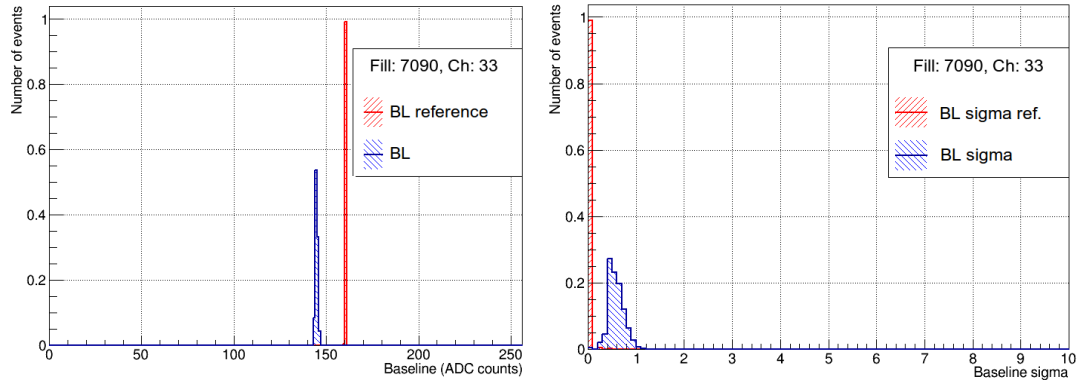


FIGURE 4.34: Test pulse spectrum degradation.

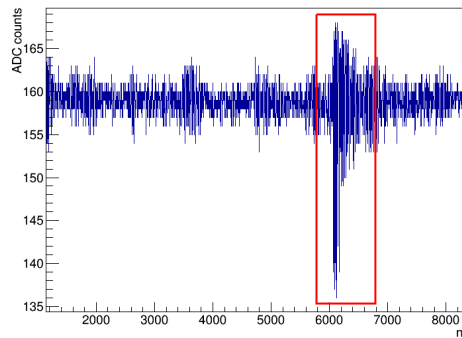
## Baseline degradation

The baseline positions define a signal amplitude registration range. To register as many diverse signals as possible, the baseline position is kept as high as possible. However, the electric chain is very complicated, and with time it may start having unwanted currents that lower the bias voltage or create a distorted baseline. The distortions make a baseline deviation wider, which creates additional noise in the signal processing. In some instances, the baseline can be leveled up using an AOH and DOH amplification setup, but it also has its amplification limit. The noise can be reduced by leveling the threshold value. Examples of the baseline position decrease, and a noisy baseline signal are shown in Fig. 4.35. Figure 4.35 A shows how the baseline changes its position with respect to the reference histogram. The actual baseline signal is wider than the reference histogram. Figure 4.35 B illustrates the baseline standard deviation of Figure 36 A. The actual histogram has a one sigma deviation while the reference one has zero. Figure 4.35 C shows an example of how a noisy baseline looks at the orbit plot (near a test pulse).



(A) Baseline position shift.

(B) Baseline standard deviation spread.



(C) Noisy baseline near the TP.

FIGURE 4.35: Baseline position distortions. Plot A shows the baseline position shift, and plot B is a baseline standard deviation spread. Plot C is a zoomed region of an orbit record around TP (shown in red rectangle).

This plot shows a noisy baseline.



### 4.4.3 Conclusions

The studied diamond sensors have shown a great performance for the severe conditions in which they were operated. The diamond sensors are a very reliable detecting material; however, only when the ionization rate is lower than during Run 2. During the running period, each C-shape was equipped with one silicon sensor. Despite the fact that those silicon sensors didn't have a proper cooling system, they were mounted in a cold place. They showed a good performance on the same level as the sCVD sensors (at their best) and lasted longer. This is one of the reasons why, for the next upgrade, silicon sensors with a cooling system will be used.



## Chapter 5

# Measurement of the $W^\pm$ boson charge asymmetry in the muon channel using the 2015 data set

This chapter describes the  $W^\pm$  boson charge asymmetry measurement using  $W^\pm$  bosons decaying to a muon,  $\mu^\pm$ , and a muon neutrino,  $\nu_\mu$ . The studied decay is characterized by an isolated, prompt, energetic muon and a neutrino that carries a significant amount of undetectable or missing transverse energy,  $E_T^{\text{miss}}$ . The asymmetry,  $\mathcal{A}(\eta)$ , is extracted using the differential  $W^\pm$  boson cross-section as a function of pseudorapidity,  $\frac{d\sigma^\pm}{d\eta}$ . The pseudorapidity,  $\eta$ , is chosen as observable due to its sensitivity to the proton momentum fraction. The definition of differential  $W^\pm$  boson cross-section is shown in Formula 5.1.

$$\frac{d\sigma^\pm}{d\eta} = \frac{1}{2\Delta\eta} \frac{N(\eta)^\pm}{\varepsilon(\eta)^\pm \varepsilon(\eta)_{\text{FSR}}^\pm \mathbb{A}(\eta)^\pm \mathcal{L}_{\text{int}}}, \quad (5.1)$$

where the factor of two appears due to pseudorapidity representation in absolute values,  $\Delta\eta$  is the  $\eta$  bin width, muon efficiency corrections as a function of the pseudorapidity are marked as  $\varepsilon(\eta)^\pm$ ,  $\varepsilon(\eta)_{\text{FSR}}^\pm$  corresponds to a ratio of number of events within the  $\eta$  and  $p_T$  acceptance before and after the final state radiation,  $\mathcal{L}_{\text{int}}$  is the integrated luminosity. The detector acceptance, estimated for  $W^+$  or  $W^-$  boson production in each  $\eta$  bin, is shown as  $\mathbb{A}(\eta)^\pm$ . Finally, the number of estimated  $W^+$  or  $W^-$  bosons in each pseudorapidity bin is represented as  $N(\eta)^\pm$ .

The  $W^\pm$  boson charge asymmetry is defined in Formula 5.2.

$$\mathcal{A}(\eta) = \frac{\frac{d\sigma}{d\eta}(W^+ \rightarrow \mu^+ \nu_\mu) - \frac{d\sigma}{d\eta}(W^- \rightarrow \mu^- \bar{\nu}_\mu)}{\frac{d\sigma}{d\eta}(W^+ \rightarrow \mu^+ \nu_\mu) + \frac{d\sigma}{d\eta}(W^- \rightarrow \mu^- \bar{\nu}_\mu)}. \quad (5.2)$$

The  $W^\pm$  boson lifetime is  $10^{-25}$  seconds, such a short-lived particle cannot be directly observed using existing technologies. However, good theoretical understanding and state of the art detecting techniques allow to reconstruct the events up to the  $W^\pm$  boson production and investigate its properties through observation of its decay products.

The next sections present the concrete steps of  $W^\pm$  boson charge asymmetry measurement. The analysis strategy is explained in section 5.1. The data set and the main triggers are introduced in section 5.2. The event selection strategy and the corresponding corrections with efficiency calculations are described in section 5.3. Muon energy scale and resolution corrections are given in section 5.4. Missing energy estimation and

recoil corrections are shown in section 5.5. Background adjustment and signal yields extraction are represented in section 5.6. Acceptance and systematic uncertainties are explained in sections 5.7 and 5.9. The  $W^\pm$  boson charge asymmetry estimation is described in section 5.8. The final results and the summary for this chapter are presented in section 5.10.

## 5.1 Analysis strategy

The  $W^\pm$  boson charge asymmetry analysis is based on the unpublished analysis "Measurement of inclusive  $W^\pm$  and  $Z^0$  boson cross-section in pp collisions at  $\sqrt{s} = 13$  TeV using 2015 data set recorded with CMS detector" by MIT group (in particular A. Apyan, K. Bierwagen, S. Brandt, M. D'Alfonso, M. Klute, A. Marini, X. Niu, J. Salfeld and S. Tkaczyk) [115]. The main goal of this chapter is to extract the differential cross-sections and asymmetry adapting the existing inclusive analysis. The selection of  $W^\pm$  boson candidate events is performed using its muon decays. The muon candidate selection assures a high probability that the physics object under the study is a muon. This muon is supposed to originate from the primary vertex and to be prompt, energetic, isolated, and should satisfy the fiducial volume requirements, including the pseudorapidity sorting for the asymmetry extraction. Additional requirements must be applied to reject on flight decaying muons, punch-through effects, possible accidental matching, and cosmic muons. To reduce correlations in  $\eta$  bins, efficiencies and corrections are required to be calculated in the chosen pseudorapidity binning.

The selected muon candidates can originate from the signal processes and background processes. The main background processes for the  $W^\pm$  boson production are QCD multi-jet events, Drell-Yan,  $W^\pm \rightarrow \tau^\pm \nu_\tau$ ,  $Z^0 \rightarrow \tau^+ \tau^-$ , top-antitop quark, and di-boson production. The number of  $W^\pm$  bosons is estimated using the missing transverse energy,  $E_T^{\text{miss}}$ . Signal yields are obtained through a simultaneous fit of a sum of MC signal templates and various background predictions to the data distribution of  $E_T^{\text{miss}}$ .

The differential cross-sections and the charge asymmetry are calculated using the signal yields from the fit, together with their statistical and systematic uncertainties.

## 5.2 Data sample

The year 2015 was the first period of Run 2 of the LHC and the year when the collisions have reached a center of mass energy of 13 TeV. The data set for this analysis was collected using the CMS detector during the 25 ns running period of the LHC with an average of nearly 13 interactions per bunch crossing (both for in-time and out-of-time pileup). The run is taken from the re-reco reconstruction level. The  $W^\pm$  boson event selection is based on a single muon trigger. The events are selected using the lowest  $p_T$  unprescaled triggers. The integrated luminosity of the corresponding data set is in the range  $\mathcal{L}_{\text{int}} = 2.2 \pm 0.1 \text{ fb}^{-1}$

The data set names and trigger path are listed in Table 5.1.

Dataset name	/SingleMuon/Run2015C(D)-PromptReco-v1/
HLT Path	HLT Iso Mu20
L1 Seed	L1 SingleMu16

TABLE 5.1: Summary of data set names and trigger paths

### 5.3 Muon selection

Each selected muon candidate is required to satisfy the following selection cuts:

- The nominal vertex is required to be in a range of 2 cm in the transverse plane and 24 cm in the z-axis from the center of the detector. The fit that is applied to reconstruct the vertex is required to have at least 4 degrees of freedom [73].
- The impact from the cosmic rays contamination is reduced through a requirement for the muon candidate to pass impact parameter cuts. The muon candidate is rejected if the distance measured at the point of closest approach is higher than  $|d_{0,pv}| > 0.02$  cm in the transverse plane and  $|d_{z,pv}| > 0.5$  cm in longitudinal.
- The muon is required to have at least one pixel hit in the tracking detector. This condition reduces the number of muons originating from decays on flight (see Section 3.4.3).
- The Global Muon algorithm is used to reconstruct the muon originating from a  $W^+$  or  $W^-$  boson decay. The muon is required to be identified as a tracker muon and as a global muon simultaneously. This condition prevents from selecting on-flight-decaying muons, punch-through events, accidentally wrong-matching of a global muon, and noisy segments of tracker muons. In order to have a good  $p_T$  resolution tracker layers must have at least five hits. The global muon track fit must contain at least one good muon hit [55].
- Muon presence must be detected in at least two stations of the muon detectors to exclude punch-through effect and accidental matching.
- The muon should be identified by the Particle-Flow Algorithm (PFA) with a muon ID. The quality of the global fit should be less than ten,  $\chi^2/\text{ndof} < 10$ .
- The absolute value of the pseudorapidity of the muon is supposed to be less than 2.4,  $|\eta| < 2.4$ . The selection is performed in eleven  $\eta$  bins. The binning is shown in Table 5.2.
- The muon transverse momentum value is determined using the Inner-Track fit. The muon is required to have a transverse momentum  $p_T > 25$  GeV.
- The muon candidate is required to be isolated in the cone  $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.4$ . To suppress amount of passing multi-jet QCD events, the contribution from other processes within the cone is required to be less than 13% of the muon  $p_T$ ,  $\text{PFIso} < 13\%$ .
- The event candidates are vetoed if an additional muon, a so-called loose muon, passes muon veto requirements. The loose muon is supposed to be in the fiducial pseudorapidity region  $|\eta| < 2.4$  and is required to have  $p_T > 10$  GeV. It must be identified by the PFA and is required to be recognized by the Global Muon or Tracker Muon algorithms. The background contamination in the loose muon cone must be less than 20% of its  $p_T$  ( $\text{PFIso}/p_T < 0.20$ ). This criteria rejects  $Z^0 \rightarrow \mu^+\mu^-$  events.

The muon selection requirements are summarized in Table 5.3.

Bin number	$ \eta $ bin range
1	0.0 - 0.2
2	0.2 - 0.4
3	0.4 - 0.6
4	0.6 - 0.8
5	0.8 - 1.0
6	1.0 - 1.2
7	1.2 - 1.4
8	1.4 - 1.6
9	1.6 - 1.85
10	1.85 - 2.1
11	2.1 - 2.4

TABLE 5.2: Pseudorapidity binning

Observable	Muon candidate	Loose muon
$p_T$	$> 25$ GeV	$> 10$ GeV
$ \eta $	$< 2.4$	$< 2.4$
Id	Global muon $\wedge$ Tracker muon	Global muon $\vee$ Tracker muon
PF muon	True	True
PFIso/ $p_T$	$< 0.13$	$< 0.20$
$\chi^2/\text{ndof}$	$< 10$	-
# Valid muon hits	$> 0$	-
# Matched stations	$> 1$	-
# Tracker layers	$> 5$	-
# Valid pixel hits	$> 0$	-
$ d_{0,pv} $	$> 0.02$	-
$ d_{z,pv} $	$> 0.5$	-

TABLE 5.3: Muon selection requirements

## 5.4 Muon energy scale and resolution corrections

The precision of the experimental muon momentum determination is biased by the detector alignment, reconstruction algorithms, and the uncertainties in the magnetic field. The precision bias of MC events originates from the event reconstruction stage, inaccurate inputs for the detector alignment, and uncertainties in the magnetic field modeling. Even though the bias of measured and generated muon momentum may have similar sources, quantitatively, they might have different values. For that reason, to be able to compare data and reconstructed MC events, the bias must be removed from both. Section 5.5 will show that the  $E_T^{\text{miss}}$  is sensitive to the muon  $p_T$  value. This connection is making the momentum determination of particular importance.

The group from the University of Rochester introduced a two-step data-driven method for the misalignment extraction and the muon energy scale correction. This method uses the POWHEG [116] generated sample of  $Z^0/\gamma^* \rightarrow \mu^+\mu^-$  events, which, after corrections, is used as perfectly aligned. To calibrate the generated distributions to the published CMS data [117], the muon transverse momentum and rapidity distributions are weighted by a correction factor. After that, the sample is required to be

modified to have perfectly aligned detector by smearing generated muon momentum with an experimental resolution function that depends on pseudorapidity.

In the first step, to obtain initial momentum scale corrections, data and simulation are re-scaled in a way that the average value of  $1/p_T^\mu$  exactly matches the perfectly aligned sample. In the second step, the reconstructed  $Z^0$  masses from data and MC are required to be the same as for the perfectly aligned sample. This correction removes dependencies on the assumptions that were made during the production of  $Z^0/\gamma^* \rightarrow \mu^+\mu^-$  events and in the detector efficiencies modeling. The corrections are calculated in terms of  $\eta-\phi$  bins, separately for positive and negative charges of the muon candidates. More details on the exact methodology may be found in the appropriate source [118].

The Rochester corrections are implemented as a package of a few C++ classes to be applied directly during the event's selection. For this analysis, the "2015 Rochester Correction" package is used.

#### 5.4.1 Muon charge misidentification correction

Charge identification is essential for the  $W^\pm$  boson charge asymmetry measurement. The muon charge identification is performed through its track curvature reconstruction, hence the primary source of possible charge misidentification is related to the curvature measurement. Studies with cosmic rays have shown that the CMS tracker detector has a high charge identification efficiency. In the  $p_T$  range of this study, the rate of misidentification is negligible,  $< 10^{-5}$ , which gives insignificant impact on the asymmetry measurement,  $< 0.01\%$  [119].

#### 5.4.2 Estimation of muon efficiencies

A variety of physical processes like pileup, the precision level of the reconstruction algorithms, and many other issues do not allow to have a perfect detection system. This is taken into account by introducing the detector efficiency variable, defined as the ratio of the number of detected objects over the total amount of produced objects. In practice, it is more convenient to use the efficiency of physics object detection and reconstruction.

The muon efficiency is estimated in data and simulation using the Tag and Probe Method [82]. The idea of Tag and Probe is simple – find a good source of  $l^+l^-$  simultaneous pair production and require one of the leptons (the tag) to pass harsh identification and isolation requirements. Afterwards, find a way to check if the second lepton (the probe) has passed the studied conditions. The efficiency is calculated using formula 5.3.

$$\varepsilon = \frac{N_{\text{pass}}}{N_{\text{pass}} + N_{\text{fail}}}, \quad (5.3)$$

where  $N_{\text{pass}}$  and  $N_{\text{fail}}$  are the numbers of passing and failing probes, respectively. This method usually is performed using  $Z^0$  bosons decaying in muon channel ( $Z^0 \rightarrow \mu^+\mu^-$ ), which are characterized by pure, high  $p_T$  muon pair production. Having a well-reconstructed one muon from the pair and taking into consideration a constraint on the  $Z^0$  boson mass ( $60 \text{ GeV} < M_{Z^0} < 120 \text{ GeV}$ ), it is possible to determine if the second high  $p_T$  muon was reconstructed using unbiased conditions. During the check, two categories are considered: muons that passed or failed the probe. The usage of two independent categories makes their statistical uncertainty uncorrelated and simplifies

the uncertainty calculation. The background contribution is estimated performing a fit on the dilepton signal invariant mass spectrum.

The difference in  $Z^0$  and  $W^\pm$  boson production kinematics is taken into account by representing kinematic efficiencies in two dimensional bins of  $p_T$  (Table 5.4) and  $\eta$  (Table 5.2). The charge-dependent difference is taken into account by separate measurement for positively and negatively charged leptons.

$p_T$ bins [GeV]	25	30	35	40	45	50	55	13000
------------------	----	----	----	----	----	----	----	-------

TABLE 5.4: Transverse momentum binning

The signal of the passing and failing probe distributions can be extracted in each bin by using fitting techniques with an appropriate signal model. Two different signal models are considered:

1. A two-parameter convolution of a Gaussian spectrum with the reconstructed MC mass spectrum. The Gaussian distribution is used to account for energy scale shifts and resolution degradation in data relative to the simulation.
2. A four-parameter convolution of a Breit-Wigner line shape with the Crystal Ball function (BW\*CB).

For the nominal fit the first model is used while the second model is used as an alternative model for the estimation of systematic uncertainties (more details in section "Systematic uncertainties" 5.8.1). Muons, which fail the required conditions, might produce a distorted signal shape. Distortions are caused mainly by an excessive amount of final-state radiation (FSR) or bremsstrahlung.

The background contributions are extracted using various fitting functions:

1. A one-parameter exponential model is used when the background is significantly smaller than the signal.
2. A one-parameter power law model is an alternative to the exponential model.
3. A three parameter quadratic model is a polynomial model.

Background for the pass and fail samples is modeled mostly using the exponential model. Systematic uncertainties are estimated using the Power law model (see section 5.8.1).

The measured efficiencies are used to determine the scaling corrections on simulation to account for difference between data and simulation (Formula 5.4).

$$\varepsilon_{W^\pm, \text{data}}(\eta, p_T) = \varepsilon_{W^\pm, \text{MC}}(\eta, p_T) \frac{\varepsilon_{\text{T\&P, data}}(\eta, p_T)}{\varepsilon_{\text{T\&P, MC}}(\eta, p_T)}. \quad (5.4)$$

Here  $\varepsilon_{W^\pm, \text{data}}(\eta, p_T)$  corresponds to the reconstruction efficiency of muons estimated from  $W^\pm$  boson decays using data events, same values but from MC events are shown as  $\varepsilon_{W^\pm, \text{MC}}(\eta, p_T)$ . Efficiencies measured using Tag and Probe technique in data and MC samples are  $\varepsilon_{\text{T\&P, data}}(\eta, p_T)$  and  $\varepsilon_{\text{T\&P, MC}}(\eta, p_T)$ , respectively.

The full muon reconstruction relies on multiple detection systems and reconstruction algorithms. Each of them has its efficiency that can be estimated using the Tag and Probe technique. Different efficiencies might be correlated with each other, and these



corrections should be taken into account appropriately. In order to rectify the problem, the passing and selection criteria from different reconstruction steps are combined. The selection criteria applied are always the passing criteria of the previous efficiency step. For example, the efficiency of the trigger is measured with respect to a well-identified muon.

### 5.4.3 Muon tracking, identification and isolation efficiency

The combined efficiency,  $\epsilon_{\text{tracking+ID+ISO}}$ , is calculated in this analysis using the Tag and Probe method in corresponding bins of  $p_T$  and  $\eta$ , separately for  $\mu^-$  and  $\mu^+$ . Tracker track reconstruction,  $\epsilon_{\text{tracking}}$ , has a high efficiency for muons with low  $p_T$  and in regions without CMS detector services.

Muon identification techniques combine information from different levels of muon reconstruction into a set of variables. The balance between purity and efficiency is obtained through the identification algorithm that provides different types of identified muons. The highest efficiency is provided by the tight muon ID type,  $\epsilon_{\text{ID}}$ . It combines all available information to select high-quality muons suppressing on flight decay muons and punch-through effects.

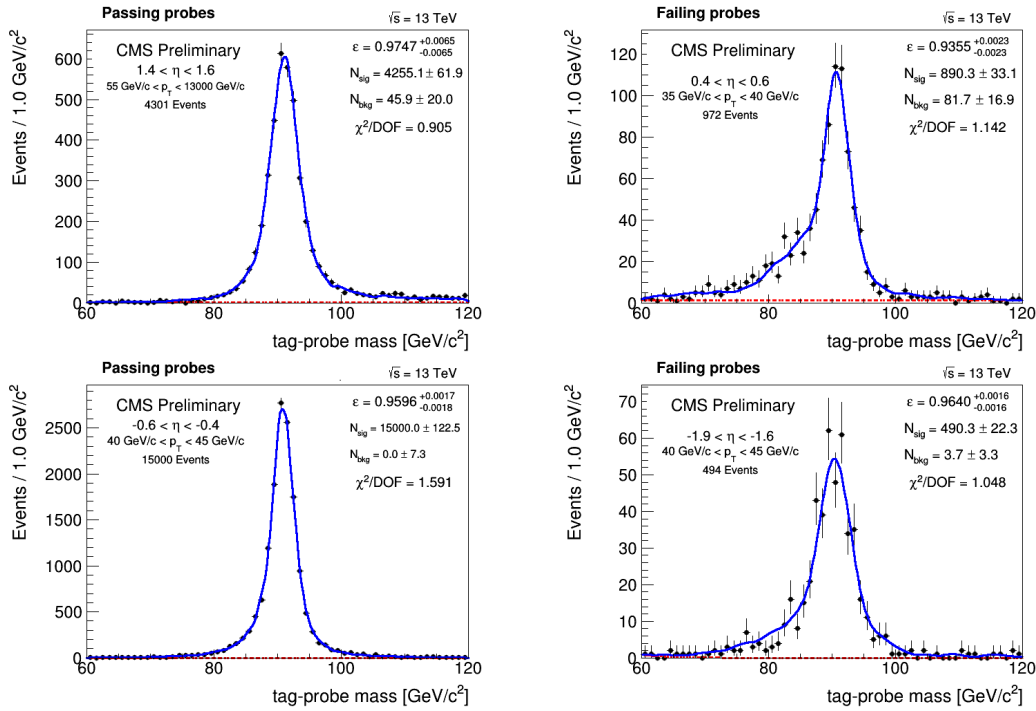


FIGURE 5.1: Muon tracking, identification and isolation efficiency in different  $\eta$  and  $p_T$  bins. Passing and failing probes are represented in the left and right columns, respectively. Results for  $\mu^-$  and  $\mu^+$  are shown in upper and lower rows, respectively.

The muon isolation is an important step to distinguish a prompt, energetic muon from weak decay within a jet. Isolation is performed using PF identified charged hadrons and neutral particles.

Figure 5.1 illustrates passing and failing probes of  $\mu^-$  and  $\mu^+$  in different  $p_T$  and  $\eta$  bins. The probe for muon tracking, identification, and isolation efficiency is defined to

be a stand-alone muon. The background contribution in passing and failing probes is modeled using the exponential model.

#### 5.4.4 Standalone muon efficiency

The efficiency calculation is performed using the Tag and Probe method in corresponding  $p_T$  and  $\eta$  bins for positive and negative muons separately. The probe of the standalone muon efficiency is defined to be a tracker track. Background estimation is performed using a polynomial function in passing and failing probes. Figure 5.2 shows examples for passing and failing probes in various  $p_T$  and  $\eta$  bins for positive and negative muons.

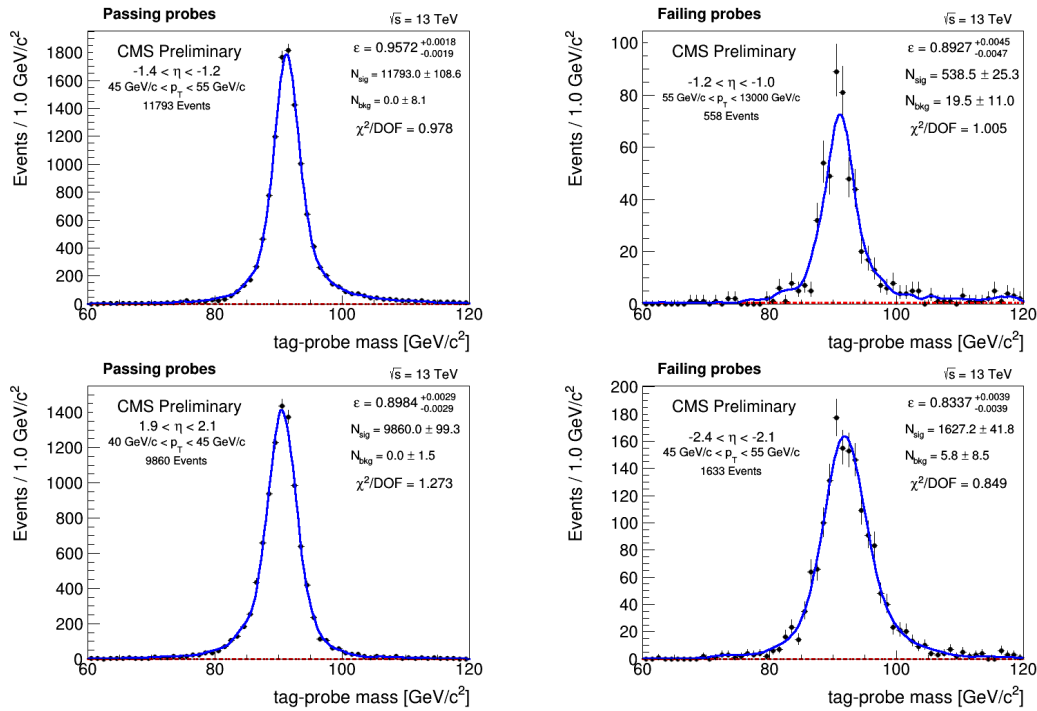


FIGURE 5.2: Muon standalone efficiency in different  $\eta$  and  $p_T$  bins. Passing and failing probes are represented in the left and right columns, respectively. Results for  $\mu^-$  and  $\mu^+$  are shown in upper and lower rows, respectively.

#### 5.4.5 Muon trigger efficiency

The trigger efficiency is estimated using well-identified muon that passes the selection criteria described in section 5.3 (table 5.3). The background in the passing category is small, so the signal yields are extracted by counting. For the failing category, the exponential model is used as a background model. Figure 5.3 illustrates example distributions of passing and failing probes for positive and negative muons in different  $p_T$  and  $\eta$  regions.

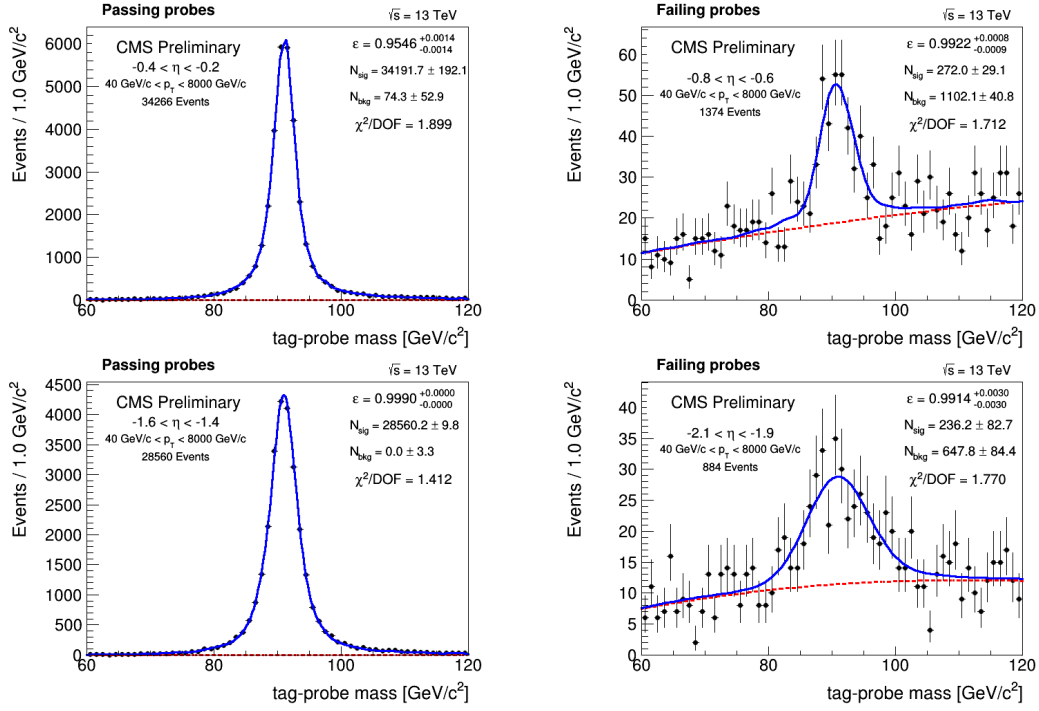


FIGURE 5.3: Muon trigger efficiency in different  $\eta$  and  $p_T$  bins. Passing and failing probes are represented in the left and right columns, respectively. Results for  $\mu^-$  and  $\mu^+$  are shown in upper and lower rows, respectively.

#### 5.4.6 Total muon efficiency

The total muon efficiency is factorized as follows:

$$\epsilon_{\text{total}} = \epsilon_{\text{tracking+ID+ISO}} \times \epsilon_{\text{STA}} \times \epsilon_{\text{trigger}}, \quad (5.5)$$

where  $\epsilon_{\text{tracking+ID+ISO}}$  is the efficiency that a standalone muon matches to a global muon that passes the identification and isolation criteria,  $\epsilon_{\text{STA}}$  is the efficiency that a tracker track from a muon matches to a global muon, and  $\epsilon_{\text{trigger}}$  is the efficiency that a fully identified and isolated muon passes the trigger (HLT and Level-1) requirements.

The total efficiency derived from Tag and Probe method using  $Z^0 \rightarrow \mu^+\mu^-$  from data and simulation is shown in Figure 5.4, the quantitative results are shown in Table 5.5.

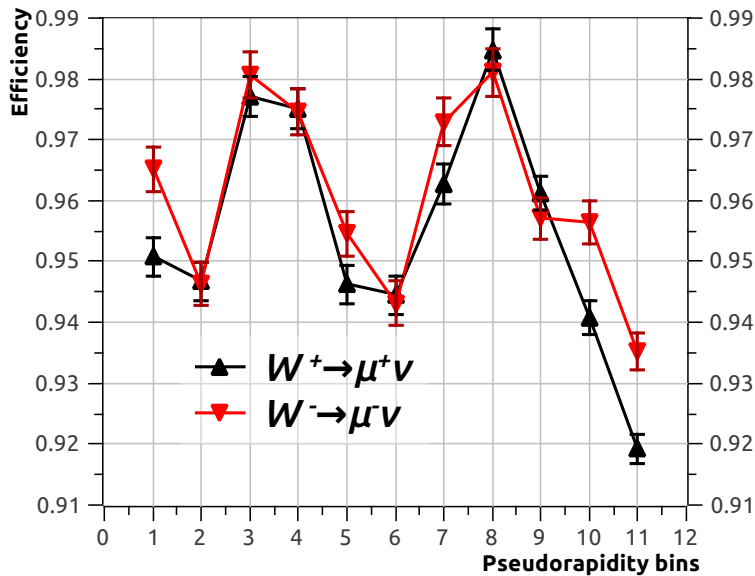


FIGURE 5.4: Muon efficiency as a function of pseudorapidity.

Bin number	$\mu^+$ efficiency	$\pm$ uncert.	$\mu^-$ efficiency	$\pm$ uncert.
1	0.9507	0.0032	0.9652	0.0036
2	0.9467	0.0031	0.9463	0.0034
3	0.9771	0.0033	0.9805	0.0037
4	0.9751	0.0033	0.9746	0.0037
5	0.9462	0.0031	0.9545	0.0035
6	0.9444	0.0032	0.9431	0.0035
7	0.9626	0.0033	0.9728	0.0038
8	0.9847	0.0033	0.9811	0.0039
9	0.9612	0.0028	0.9571	0.0033
10	0.9408	0.0028	0.9563	0.0035
11	0.9193	0.0024	0.9351	0.0030

TABLE 5.5: Muon efficiency with uncertainty as a function of pseudorapidity

## 5.5 Missing transverse energy

Missing transverse energy,  $E_T^{\text{miss}}$  or MET, is one of the main signatures of  $W^\pm$  boson production in the lepton channel. Previous studies of MIT group (see [115]) have shown that for this particular analysis it is more suitable to use  $E_T^{\text{miss}}$  estimated using PUPPI algorithm. Unlike the PF MET, the PUPPI  $E_T^{\text{miss}}$  values are free of  $\phi$  modulation effect. Mainly this is because PUPPI produces an accurate event interpretation without the need for further pileup correction.

The number of  $W^\pm$  bosons is estimated from the MC signal ( $W^\pm \rightarrow \mu^\pm \nu_\mu$ ) template after the fit, so that an accurate prediction of  $E_T^{\text{miss}}$  is essential for this analysis. The MC  $E_T^{\text{miss}}$  corrections depend on the vector-boson recoil modeling techniques. An exact calculation of the recoil processes involves numerous detector effects simulation, but

due to various reasons, not all of them can be modeled with a necessary precision level. Discrepancies can appear from an incomplete description of underlying events or/and insufficient quality of the calorimeter response modeling. For this reason, an additional correction must be applied.

### Recoil correction

The MET is measured in every selected event using momentum conservation in the transverse plane. In this plane, momentum vectors can be categorized into three parts: transverse energy of a muon,  $\vec{E}_T^1$ , missing transverse energy (another designation is  $\vec{E}_T$ ), and recoil vector,  $\vec{u}_T$ , figure 5.5a. The recoil is determined in  $W^\pm \rightarrow \mu^\pm \nu_\mu$  process as the negative vectorial sum of the transverse energy vectors of all reconstructed particles not related to a daughter muon:

$$\vec{u}_T = - \left( \sum_{i=1}^N \vec{E}_{T_i} - \vec{E}_{T_\mu} \right). \quad (5.6)$$

Here N corresponds to a number of all PF reconstructed particles in the event,  $\vec{E}_{T_\mu}$  is the transverse energy of a daughter muon.

The recoil corrections are derived from the  $Z^0 \rightarrow \mu^+ \mu^-$  boson recoil in data and specified for every event in simulated samples of  $Z^0 \rightarrow \mu^+ \mu^-$  and  $W^\pm \rightarrow \mu^\pm \nu_\mu$ . In this method, the recoil vector is decomposed into two components - a perpendicular ( $u_\perp$ ) and a parallel ( $u_\parallel$ ) as a projection of the recoil vector on the boson  $q_T$ . Figure 5.5b

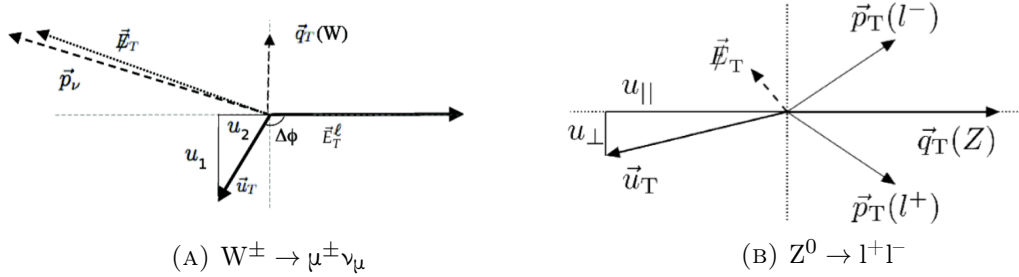


FIGURE 5.5: Schematic representation of  $W^\pm$  and  $Z^0$  boson event decomposition in transverse plane.

shows the illustration of  $Z^0 \rightarrow l^+ l^-$  event kinematics in a transverse plane. The  $\vec{u}_T$ ,  $u_\perp$  and  $u_\parallel$  corresponds to the recoil vector and its components,  $\vec{q}_T(Z^0)$ ,  $\vec{p}_T(\mu^-)$  and  $\vec{p}_T(\mu^+)$  are transverse momentum of  $Z^0$  boson with its muons, and  $\vec{E}_T$  is missing energy. Every distribution of the recoil components is fitted independently in bins of boson  $p_T$  and  $\eta$  with a triple Gaussian, whose width and means vary with the boson  $p_T$  (Figure 5.6). The  $p_T$  is binned in 56 bins in range from 0.0 to 300.0 GeV, the pseudorapidity bins are shown in Tab. 5.2.

The obtained width and mean values are fitted by polynomials  $f_\perp(p_T^{Z^0})$ ,  $f_\parallel(p_T^{Z^0})$  and  $\sigma_\perp(p_T^{Z^0})$ ,  $\sigma_\parallel(p_T^{Z^0})$ , producing resolution and response curves respectively. Figure 5.7 shows the fit of the recoil response curve as a function of boson  $p_T$ . To apply obtained recoil corrections to generated  $W^\pm$  boson events, the parallel and perpendicular planes of  $W^\pm$  boson recoil must be determined. The propagation direction of the generated  $W^\pm$  boson is calculated using its  $p_T$  and  $\phi$  values. In each simulated event, the recoil

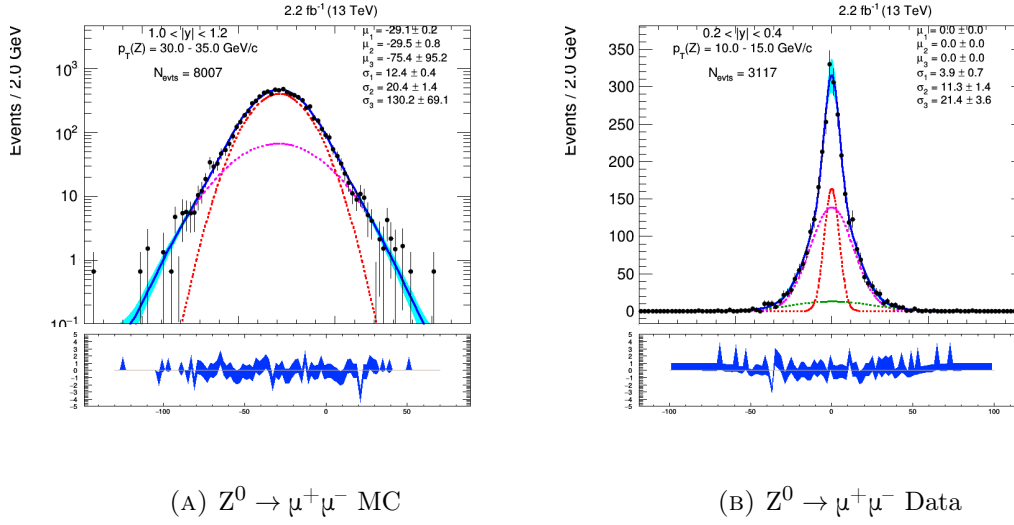


FIGURE 5.6: The figures represent an example of the fit of parallel (left figure) and perpendicular (right) recoil components in a random rapidity and boson  $p_T$  bins. The legend contains the mean and width values of each Gaussian function.

value corresponding to transverse momentum of  $W^\pm$  is corrected for the difference in between simulated  $Z^0 \rightarrow \mu^-\mu^+$  and  $Z^0 \rightarrow \mu^-\mu^+$  from the data. The recoil corrections are also applied to the modeled background events. The corrections are calculated in bins of  $W^\pm$  boson absolute rapidity for a few reasons:

- The difference between the MC and data for  $W^\pm$  and  $Z^0$  models can be resolved in a first approximation through  $W^\pm/Z^0$  boson rapidity parameterization;
- The recoil against the boson after a hard scattering can cover different parts of the detector.

The recoil correction for the background from the top quark production is considered through an additional modified fit. The fit is performed with a composed model of the triple Gaussian plus the top template from MC. The MET value is estimated after adding back the energy of the muon from  $W^\pm$  decay.

## 5.6 $W^\pm$ boson signal extraction

Some events in pp collisions have a different physics origin, but the same signature of the process as  $W^\pm \rightarrow \mu^\pm \nu_\mu$  production: missing energy and a prompt, energetic muon. Such events are considered as background events. The measure of the presence of each component may be identified using theoretical predictions of the signal and background events in a fitting procedure. This section describes the background modeling used to extract the signal yields and the simultaneous fit procedure.

### 5.6.1 Electroweak background

Electroweak background (EWK BKG) is a background that consists mainly of the electroweak decay products that can mimic the signal muon. The main EWK background processes are:

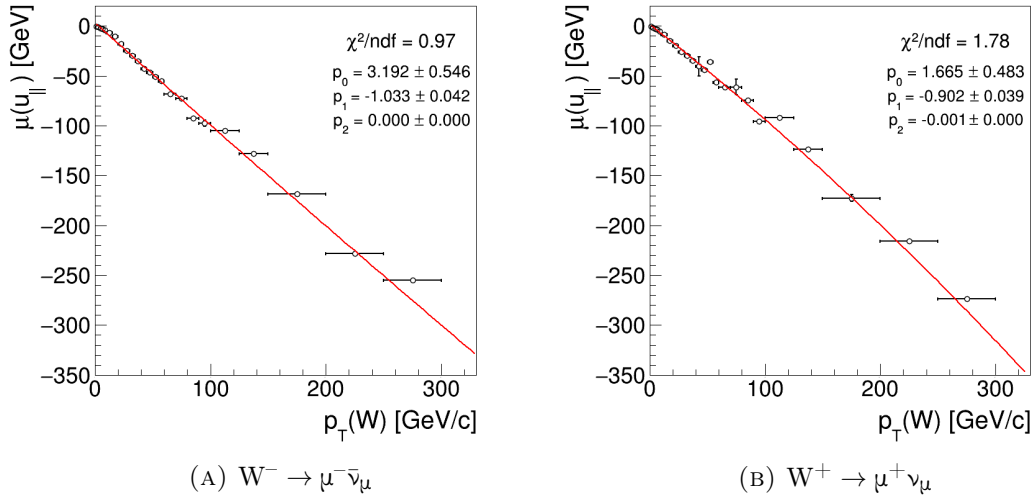


FIGURE 5.7: The response curve of the recoil scale as a function of boson  $p_T$ . The left figure represents the fit of the mean values of the first Gaussian for the parallel recoil component in  $W^- \rightarrow \mu^- \bar{\nu}_\mu$  in the first  $\eta$  bin. The right figure represents the fit of the mean values of the second Gaussian for the parallel recoil component in  $W^+ \rightarrow \mu^+ \nu_\mu$  in the tenth  $\eta$  bin.

- Drell-Yan process ( $q^f q^{\bar{f}} \rightarrow \mu^+ \mu^-$ ) - flavor-antiflavor quark annihilation to a pair of muons. Events from such processes become a background when one of the pair leaves the fiducial volume of the detector remaining undetected.
- Events from  $Z^0 \rightarrow \tau^+ \tau^-$  and  $W^\pm \rightarrow \tau^\pm \nu_\tau$  processes when a tau lepton decays to a muon. Such muons usually have low momentum and are strongly suppressed by the  $p_T$  requirement.
- The events of diboson ( $W^\pm W^\pm$ ,  $W^\pm Z^0$  and  $Z^0 Z^0$ ) production are relatively rare, and the contribution from the sum of the three is small.
- Top-antitop quark annihilation is not related to electroweak background. 99.9% of  $t\bar{t}$  pair decays to the bottom quark and  $W^\pm$  boson, 10.5% of which decay into muons [14]. The contribution from  $t\bar{t}$  events for the matter of convenience was added to the EWK BKG.

The  $E_T^{\text{miss}}$  distributions of all electroweak and  $t\bar{t}$  contributions are modeled with Monte Carlo methods using theoretical predictions. Drell-Yan,  $Z^0 \rightarrow \tau^+ \tau^-$  and  $W^\pm \rightarrow \tau^\pm \nu_\tau$  samples were produced using aMC@NLO [120] interfaced to PYTHIA 8 [121] for parton shower evolution. The  $t\bar{t}$  sample was produced using aMC@NLO using NNPDF3.0 NLO PDFs [122], interfaced with PYTHIA 8, and using the CUETP8M1 tune for a parton shower evolution. Diboson ( $W^\pm Z^0$ ,  $Z^0 Z^0$ ) events were modeled using PYTHIA 8 with NNPDF2.3 LO PDFs [123] and the CUETP8M1 tune. The  $W^\pm W^\pm$  production was modeled with POWHEG using NNPDF3.0 NLO PDFs, interfaced to PYTHIA 8 and the CUETP8M1 tune. Table 5.6 summarizes the MC files which are used for the EWK background modeling. A quantitative contribution from each source will be presented in section 5.6.4.

Set name	$\sigma$ [pb]
WJetsToLNu TuneCUETP8M1 13TeV-amcatnloFXFX-pythia8	59100
DYJetsToLL M-50 TuneCUETP8M1 13TeV-amcatnloFXFX-pythia8	5835
TTJets TuneCUETP8M1 13TeV-amcatnloFXFX-pythia8	831.76
WW To2L2Nu 13TeV-powheg	1.3845
WZ TuneCUETP8M1 13TeV-pythia8	58.8
ZZ TuneCUETP8M1 13TeV-pythia8	15.4

TABLE 5.6: Monte Carlo sets for the electroweak background

### 5.6.2 Background from quantum chromodynamics processes

Unlike the other background contributions, the QCD background (QCD BKG) arises from multijets misidentification. The complexity of a multijet structure does not allow to distinguish with a high level of confidence a real origin of the event. The main sources of events with a fake isolated muon are meson decay and semi-leptonic heavy flavor decay. Usage of the Tight muon identification type and other requirements prevents the selection of decay on flight events and significantly suppresses the contribution from the meson decay events. Remaining semi-leptonic heavy flavor events are studied using data-driven techniques.

The methodology of the method relies on two data-driven background samples: signal region and control region. Both of them represent the background selected from the events with a different level of misidentification rate. The QCD background estimation is based on the assumption that the shape of its distribution remains the same in signal and control regions. The signal region is related to the control region through a particular factor that can be defined from the simultaneous fit. In this analysis control region is defined by alternative particle selection using variables that define a level of misidentification rate. The right candidate for such a variable is a condition on the muon cone isolation (PFIso). The access to events with a different level of muon cone isolation is obtained using the data recorded with a non-isolated muon trigger. Due to a vast amount of such events, the event record was prescaled with a specific value, X, meaning one event record out of X events. The prescale value is sensitive to various setups and is not a constant. In the control region, the events with a high muon cone contamination are selected using  $PFIso/p_T > 30\%$  condition. That condition assures to have misidentified events as the main contribution in the  $E_T^{\text{miss}}$  control region distribution.

The shape of the QCD  $E_T^{\text{miss}}$  distribution is modeled using modified Rayleigh distribution assuming that each MET component is uncorrelated, normally distributed with equal variance, given by

$$f_{\text{QCD}}(E_T^{\text{miss}}) = E_T^{\text{miss}} \exp\left(-\frac{E_T^{\text{miss}^2}}{aE_T^{\text{miss}^2} + bE_T^{\text{miss}} + c}\right) * (aE_T^{\text{miss}^2} + bE_T^{\text{miss}} + c), \quad (5.7)$$

where a, b, c are fitted parameters and MET is in the range  $0 < E_T^{\text{miss}} < 150$  GeV. The fit is performed simultaneously in signal and control regions for the events with positive and negative muons. Figure 5.8 shows distribution of  $E_T^{\text{miss}}$  from events in the control region. Plots are shown for  $W^+ \rightarrow \mu^+ \nu_\mu$  and  $W^- \rightarrow \mu^- \bar{\nu}_\mu$  channels in two pseudorapidity bins,  $0.2 < |\eta| < 0.4$  and  $1.2 < |\eta| < 1.4$ . As seen from the plots, an overwhelming majority of control region events originate from the QCD background.



It is considered that the QCD background is the same for the events from  $W^+$  and  $W^-$  decays. This assumption is varied for the systematic uncertainty. The values of "a" parameter in the fit are considered to be the same for signal and control regions; this condition is varied for the systematic uncertainty. More details on systematic uncertainties from the QCD background estimation procedure are discussed in section 5.8.4. Quantitative results on the QCD background estimation are briefly summarized in section 5.6.4.

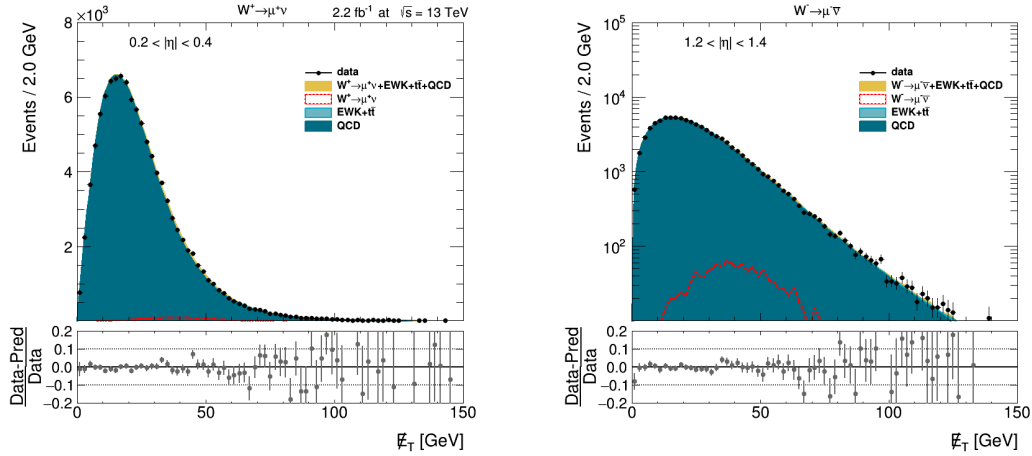


FIGURE 5.8:  $E_T^{\text{miss}}$  fit in the control region defined by inverting isolation criteria. The fit is performed separately for  $W^+$  and  $W^-$  samples in different pseudorapidity regions.

### 5.6.3 Simultaneous fit

The  $W^\pm \rightarrow \mu^\pm \nu_\mu$  MC signal template is modeled with aMC@NLO using NNPDF3.0 NLO PDFs, interfaced with PYTHIA 8 and the CUETP8M1 tune. The signal shape is modeled with a Gaussian distribution as an external constraint during the fit. Technical implementation of the fit is performed using libraries of the RooFit toolkit [124]. The probability density functions (p.d.f.) and its integration interval for the signal shape and EWK BKG are obtained separately for  $W^- \rightarrow \mu^- \bar{\nu}_\mu$  and  $W^+ \rightarrow \mu^+ \nu_\mu$  channels in each  $\eta$  bin. The integration interval for the signal is derived from its distribution and defined as a range from zero to its total value. The QCD interval is derived from the data and defined as a range from zero to the total number with a starting value of 30% of the data distribution. The EWK BKG integration interval is defined as a signal interval multiplied by a constant ratio ( $\text{ewk}_r$ ) of a total EWK BKG over total signal distribution. The constant ratio is defined separately for each channel, in signal and control regions ( $\text{ewk}_{r(S)}^\pm$  and  $\text{ewk}_{r(C)}^\pm$ ). The integration interval definitions are considered to be the same for  $W^- \rightarrow \mu^- \bar{\nu}_\mu$  and  $W^+ \rightarrow \mu^+ \nu_\mu$  channels in signal and control regions in each eta bin. The data  $E_T^{\text{miss}}$  distribution consists of the signal, EWK BKG and QCD BKG contributors (see MET distribution in control or signal regions, Fig. 5.8 and Fig. 5.9, respectively). The sum probability density function (p.d.f.) per channel per region per  $\eta$  bin consists of signal p.d.f., EWK BKG p.d.f., and QCD BKG fit. The final fit contains the phase space of signal and control regions, due to the data-driven technique, and a common QCD phase space for  $W^+$  and  $W^-$

due to a common QCD background assumption. The unbinned maximum likelihood fit is performed with an extended likelihood term. As a minimization tool, the MINUIT package [125] in the most precise mode (strategy 2) is used. The  $W^\pm \rightarrow \mu^\pm \nu_\mu$  signal yields, the QCD background yields, and the parameters of the Rayleigh distribution (a, b, c) are considered to be free during the fit.

#### 5.6.4 Fit results

Selective results of simultaneous fit for different  $\eta$  regions are presented in Figure 5.9. The full set of plots for each  $\eta$  region, for each channel in signal and control regions is given in Appendix B. The extracted yields of  $W^+ \rightarrow \mu^+ \nu_\mu$ , EWK BKG, and QCD from the signal region are summarized in the table 5.7. Results corresponding to the control region of  $W^+ \rightarrow \mu^+ \nu_\mu$  channel are summarized in the table 5.8. Tables 5.9 and 5.10 show results obtained in  $W^- \rightarrow \mu^- \bar{\nu}_\mu$  channel in signal and control regions, respectively.

$ \eta $ bin	Signal region						
	Total sel.	$W^+ \rightarrow \mu^+ \nu_\mu$	$\pm$ err	QCD	$\pm$ err	EWK	$\pm$ err
0.0-0.2	908163	774951	1794	82240	1684	50641	117
0.2-0.4	848544	721812	1351	76362	1147	50045	93
0.4-0.6	951901	803483	1807	89727	1710	57956	130
0.6-0.8	950977	793812	1058	95478	682	61940	82
0.8-1.0	877905	723295	1515	95523	1380	59392	124
1.0-1.2	910356	746482	1406	98038	1238	65896	124
1.2-1.4	1004326	807860	1621	118811	1506	77332	155
1.4-1.6	1023373	820048	1740	120057	1665	83079	176
1.6-1.85	1153777	921030	1610	132763	1472	99819	174
1.85-2.1	1170676	933515	1611	131799	1479	104784	180
2.1-2.4	1207398	978976	1625	110690	1490	117687	195
0.0-2.4	11007396	9122783	5616	1043100	5282	839350	516

TABLE 5.7: The table represents the results of the fitting procedure for an absolute value of each pseudorapidity region, including the total  $|\eta|$  range. "Total sel." corresponds to the total number of events selected from the data. The extracted  $W^+ \rightarrow \mu^+ \nu_\mu$  yields and background contribution in the signal region are shown with its uncertainties.

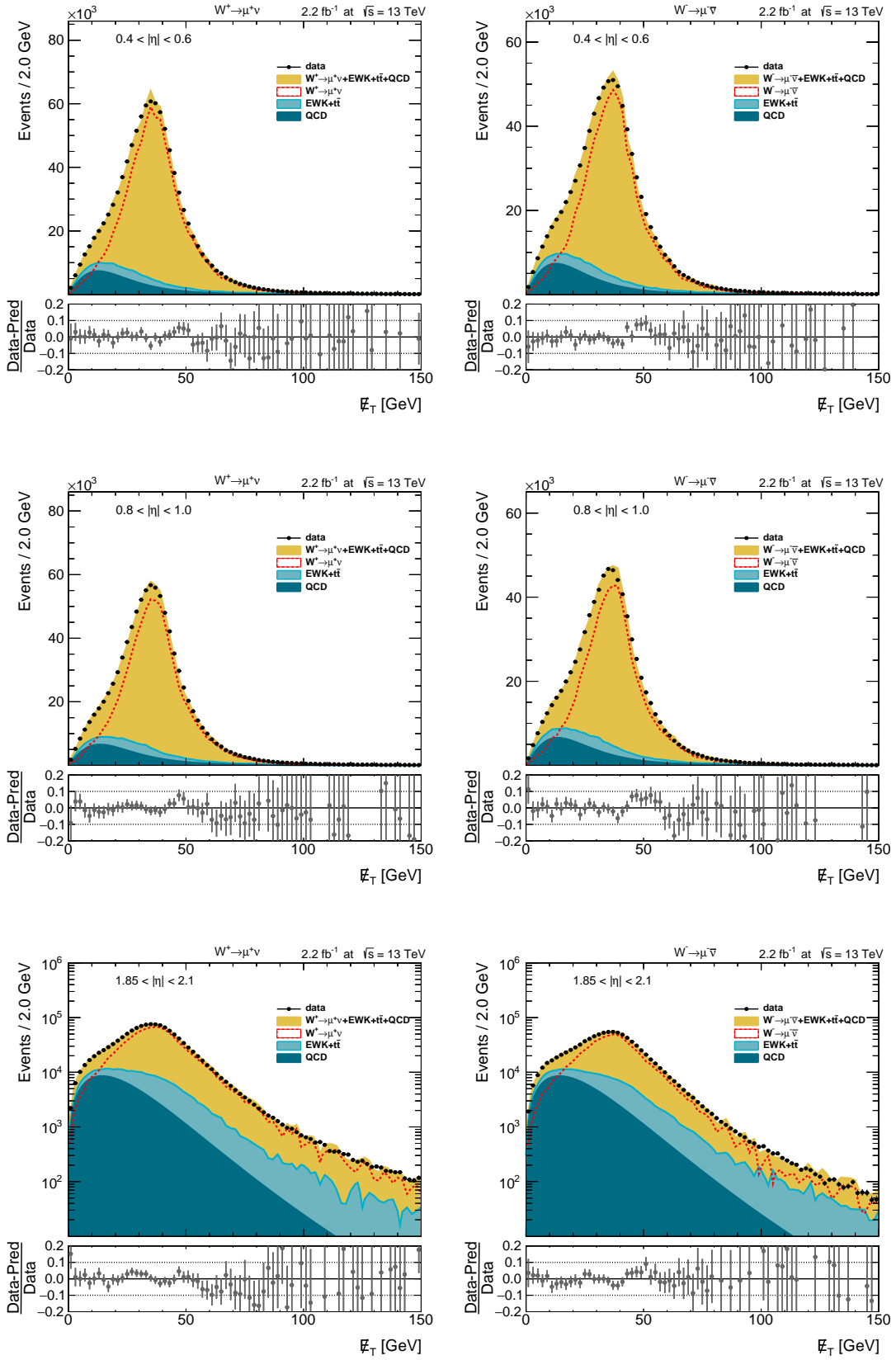


FIGURE 5.9: Simultaneous fit of the  $E_T^{\text{miss}}$  distribution in signal region in the  $0.4 < |\eta| < 0.6$  pseudorapidity range (upper row),  $0.8 < |\eta| < 1.0$  (middle row) and  $1.85 < |\eta| < 2.1$  (lower row). The fit is performed separately for  $W^+$  (left column) and  $W^-$  (right column).

$ \eta $ bin	Control region						
	Total sel.	$W^+ \rightarrow \mu^+ \nu_\mu$	$\pm$ err	QCD	$\pm$ err	EWK	$\pm$ err
0.0-0.2	115913	2280	224	112349	344	798	78
0.2-0.4	104095	1860	204	101520	326	699	77
0.4-0.6	114610	2085	224	111338	348	770	83
0.6-0.8	107018	2062	206	104338	318	745	74
0.8-1.0	92169	1623	184	89970	294	577	65
1.0-1.2	87345	1887	179	84515	283	660	62
1.2-1.4	85750	1553	171	83775	272	499	55
1.4-1.6	79107	1618	170	76902	267	506	53
1.6-1.85	78210	1565	167	75743	254	479	51
1.85-2.1	65627	1527	163	63654	237	395	42
2.1-2.4	52280	1500	168	50532	230	359	40
0.0-2.4	982124	21636	1408	952310	1908	7154	465

TABLE 5.8: The table represents the results of the fitting procedure for an absolute value of each pseudorapidity region, including the total  $|\eta|$  range. "Total sel." corresponds to the total number of events selected from the data. The extracted  $W^+ \rightarrow \mu^+ \nu_\mu$  yields and background contribution in the control region are shown with its uncertainties.

$ \eta $ bin	Signal region						
	Total sel.	$W^- \rightarrow \mu^- \bar{\nu}_\mu$	$\pm$ err	QCD	$\pm$ err	EWK	$\pm$ err
0.0-0.2	785783	655427	1738	82240	1684	48441	128
0.2-0.4	730670	608222	1298	76362	1147	46406	99
0.4-0.6	813480	671092	1755	89727	1710	53390	139
0.6-0.8	801520	650326	987	95478	682	55457	84
0.8-1.0	736393	586365	1454	95523	1380	54169	134
1.0-1.2	740903	583685	1334	98038	1238	59116	135
1.2-1.4	806372	619859	1543	118811	1506	68021	169
1.4-1.6	804511	611094	1652	120057	1665	73490	198
1.6-1.85	882305	662417	1502	132763	1472	87286	198
1.85-2.1	863917	641325	1485	131799	1479	91367	211
2.1-2.4	869639	655214	1486	110690	1490	103777	235
0.0-2.4	8835493	7041192	5363	1043100	5282	753364	573

TABLE 5.9: The table represents the results of the fitting procedure for an absolute value of each pseudorapidity region, including the total  $|\eta|$  range. "Total sel." corresponds to the total number of events selected from the data. The extracted  $W^- \rightarrow \mu^- \bar{\nu}_\mu$  yields and background contribution in the signal region are shown with its uncertainties.

$\eta$   bin	Control region						
	Total sel.	$W^- \rightarrow \mu^- \bar{\nu}_\mu$	$\pm$ err	QCD	$\pm$ err	EWK	$\pm$ err
0.0-0.2	113920	1420	193	112349	344	633	86
0.2-0.4	103657	1499	183	101520	326	653	80
0.4-0.6	113088	1489	193	111338	348	676	88
0.6-0.8	106639	1480	180	104338	318	672	81
0.8-1.0	92090	1454	158	89970	294	641	70
1.0-1.2	86005	1230	149	84515	283	529	64
1.2-1.4	85526	1149	137	83775	272	517	61
1.4-1.6	78121	991	135	76902	267	409	55
1.6-1.85	76755	1015	129	75743	254	415	52
1.85-2.1	64853	876	114	63654	237	369	48
2.1-2.4	52162	1096	125	50532	230	418	47
0.0-2.4	972816	15005	1283	952310	1908	6508	556

TABLE 5.10: The table represents results for an absolute value of each pseudorapidity region, including the total  $|\eta|$  range. The extracted  $W^- \rightarrow \mu^- \bar{\nu}_\mu$  yields and background contribution (in particular, the contribution of QCD background yields) in the control region are shown with its uncertainties.

## 5.7 Acceptance

In this analysis, the fiducial acceptance is calculated using the same selection requirements as described in section 5.3. The fiducial acceptance,  $\mathbb{A}(\eta)$  in  $\eta$  pseudorapidity bin is defined as:

$$\mathbb{A}(\eta) = \frac{N_{\text{fid}}^{\text{pass}}(\eta)}{N_{\text{fid}}^{\text{gen}}(\eta)}. \quad (5.8)$$

Here  $N_{\text{fid}}^{\text{pass}}(\eta)$  is the number of MC events in corresponding  $\eta$  bin that pass the selection requirements in a corresponding fiducial phase space,  $N_{\text{fid}}^{\text{gen}}(\eta)$  is the number of generated events in the same bin. The event selection is performed using the Monte Carlo sample, which is used as a signal template in the simultaneous fit procedure.

Calculated acceptance value in each pseudorapidity bin for  $W^+ \rightarrow \mu^+ \nu_\mu$  and  $W^- \rightarrow \mu^- \bar{\nu}_\mu$  events are given in the Tab. 5.11.

$ \eta $ bin	$W^+ \rightarrow \mu^+ \nu_\mu$	$\pm \text{err}$	$W^- \rightarrow \mu^- \bar{\nu}_\mu$	$\pm \text{err}$
0.0-0.2	0.852	0.003	0.861	0.003
0.2-0.4	0.799	0.003	0.804	0.003
0.4-0.6	0.889	0.003	0.894	0.003
0.6-0.8	0.883	0.003	0.885	0.003
0.8-1.0	0.806	0.003	0.815	0.003
1.0-1.2	0.828	0.003	0.824	0.003
1.2-1.4	0.893	0.003	0.898	0.003
1.4-1.6	0.896	0.003	0.899	0.003
1.6-1.85	0.805	0.003	0.815	0.003
1.85-2.1	0.828	0.003	0.841	0.003
2.1-2.4	0.723	0.002	0.748	0.003

TABLE 5.11: Table represents fiducial acceptance values in different pseudorapidity regions for  $W^+ \rightarrow \mu^+ \nu_\mu$  and  $W^- \rightarrow \mu^- \bar{\nu}_\mu$  channels.

Figure 5.10 shows acceptance as a function of the absolute values of pseudorapidity. The results have a periodic structure that reflects the design of the CMS detector. The gaps in the muon coverage originate from the sub-detectors design, space for the detector services, and alignment challenges. Sensitivity to this feature is reported in various technical documentation [68], [65].

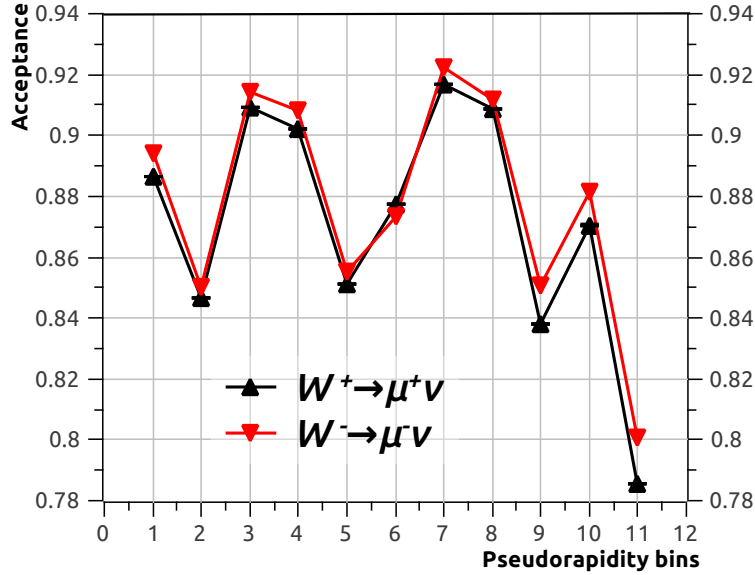


FIGURE 5.10: Acceptance values as the function of absolute pseudorapidity for  $W^+ \rightarrow \mu^+ \nu_\mu$  and  $W^- \rightarrow \mu^- \bar{\nu}_\mu$  channels. The systematic uncertainties are smaller than the size of the markers used in the plot.

## 5.8 Systematic uncertainties

This section describes the systematic uncertainties from every step of the analysis and approaches to treat them.

### 5.8.1 Muon efficiencies

Lepton efficiency estimation is a complicated procedure that uses various assumptions in its methodology. The main sources of systematic uncertainties originate from the signal and background shape modeling, variations in uncertainty range, and alternative binning.

**Breit-Wigner \* Crystal Ball signal model.** For the nominal signal model, a convolution of the Gaussian spectrum with reconstructed MC mass spectrum is used. As an alternative model the four-parameter convolution of Breit-Wigner [126] line shape with the Crystal Ball function [127] is used. The Breit-Wigner shape is fixed to the PDG values of the  $Z^0$  width and mass. The Crystal Ball function has free parameters intended to account for energy scale shifts, mass resolution, and low mass tails from FSR and bremsstrahlung. The deviation from the nominal differential cross-section values across the pseudorapidity bins varies between 1 – 1.5% for  $W^+$  and  $W^-$  channels.

**Power law background model.** The power-law model is used as an alternative background model due to its wide usage for background estimation. It does a good description of protracted energy tails after energy loss events. The power-law distribution is defined as:

$$f(x) = A \cdot x^{-r}, \quad (5.9)$$

where  $A$  is a constant,  $-r$  is a power-law index. The systematic uncertainty from using this model is below 1%.

**Variation in estimated uncertainty range.** Another considered systematic uncertainty is a variation of lepton efficiencies in uncertainty range of its  $\eta - \varphi$  bins. The correlation among the  $\eta$  bins is reduced due to a representation of  $\eta$  part of 2D bins in asymmetry pseudorapidity binning (Table 5.2). The uncertainty from the variation lies in the range of 0.6 – 1.1% for both channels.

### 5.8.2 Muon momentum scale and resolution correction

The systematic uncertainty estimation of the muon energy scale and resolution is based on the smearing of the correction using one sigma of the statistical error in  $\eta$  and  $\varphi$  bins following Gaussian distribution. The actual smearing is implemented in the Rochester C++ class and can be calculated through a generation of a toy run with a random seed number. Systematic uncertainty is calculated running a hundred of toy runs with different seeds, and taking an RMS value of all toy runs. The uncertainty value in each channel varies from bin to bin in a range of 0.5 – 1.2%.

### 5.8.3 Missing transverse energy

The primary sources of systematic uncertainties for the  $E_T^{\text{miss}}$  calculation come from an alternative recoil binning, the pileup variation, and usage of alternative fitting functions.

**Inclusive recoil.** To produce the systematic uncertainty from the recoil fit, an alternative, inclusive ( $0.0 < \eta < 2.4$ ) binning of rapidity is performed. The results varies in a range from 1.0–3.5% in  $W^- \rightarrow \mu^- \bar{\nu}_\mu$  and 0.9–3.5% in  $W^+ \rightarrow \mu^+ \nu_\mu$  channels.

**Pileup variation.** Additional systematic uncertainty is coming from the variations in the pileup uncertainty range. The systematic uncertainty is derived using recoil corrections calculated using an alternative simulation, which was produced with varied pileup in its uncertainty range. Obtained results are varying in range 0.6 – 2.9% in  $W^+ \rightarrow \mu^+ \nu_\mu$  channel and 0.8 – 3.4% in  $W^- \rightarrow \mu^- \bar{\nu}_\mu$  channel.

**Alternative fitting function of the recoil component.** The systematic uncertainty coming from the fitting function choice is estimated using an alternative approach. The fit is performed using the non-parametric function in the RooFit toolkit. In this approach, the p.d.f. of the recoil components is represented as a superposition of Gaussians with the same surface but different dynamic width, which varies depending on event density [124]. The obtained uncertainty is varying in the range of 0.7%–1.04% and 0.7% – 1.2% for  $W^+ \rightarrow \mu^+ \nu_\mu$  and  $W^- \rightarrow \mu^- \bar{\nu}_\mu$  channels, respectively. For the asymmetry, the uncertainty is in the range of 0.1% – 0.4%.

#### 5.8.4 $W^\pm$ boson signal extraction fit

**Separate "a" parameter for the QCD shape functions.** As an alternative approach in QCD background estimation, the fit is performed without binding the "a" parameter between signal and control regions. The new fit is performed for each  $\eta$  bin in each channel. The systematic uncertainty varies from bin to bin in a range 0.5–3.0% in  $W^+ \rightarrow \mu^+ \nu_\mu$  channel and 0.6 – 3.6% in  $W^- \rightarrow \mu^- \bar{\nu}_\mu$ .

**Separate QCD shape functions.** Additional fits are performed with a separate QCD shape functions for  $W^+ \rightarrow \mu^+ \nu_\mu$  and  $W^- \rightarrow \mu^- \bar{\nu}_\mu$ . The systematic uncertainty in  $W^+ \rightarrow \mu^+ \nu_\mu$  and  $W^- \rightarrow \mu^- \bar{\nu}_\mu$  channels varies for different  $\eta$  bins in a range 0.5 – 1.7% and 0.5 – 1.5% respectively.

#### 5.8.5 Luminosity uncertainties

The total uncertainty of luminosity is 2.3% [128]; it is fully correlated between different channels and pseudorapidity bins. The source of each uncertainty component is summarized in the table 5.12.



	Systematic	Correction (%)	Uncertainty (%)
Integration	Stability	-	1
	type 1	7-9	0.6
	type 2	0-4	0.7
	CMS dead time	-	0.5
	Dynamic Inefficiency	-	0.4
Normalization	XY-Correlations	1.1	1.5
	Beam current calibration	-	0.3
	Ghosts and satellites	-	0.2
	Length scale	-0.5	0.5
	Orbit Drift	-	0.4
	Beam-beam deflection	1.8	0.4
	Dynamic- $\beta$	-	0.5
	Total		2.3

TABLE 5.12: The table represents components of the CMS luminosity uncertainty measurement during the 2015 data-taking period at  $\sqrt{s} = 13$  TeV in proton-proton collisions [128].

## 5.9 $W^\pm$ boson charge asymmetry measurement

### 5.9.1 Differential cross-sections

The differential cross-sections are calculated using the formula:

$$\frac{d\sigma^\pm}{d\eta} = \frac{1}{2\Delta\eta} \frac{N(\eta)^\pm}{\varepsilon(\eta)^\pm \varepsilon(\eta)_{\text{FSR}}^\pm \mathbb{A}(\eta)^\pm \mathcal{L}_{\text{int}}}. \quad (5.10)$$

Here the factor of two appears due to a representation of pseudorapidity in absolute values,  $\Delta\eta$  is  $\eta$  bin width, efficiency values are marked as  $\varepsilon(\eta)^\pm$  and integrated luminosity is shown as  $\mathcal{L}_{\text{int}}$ .  $\varepsilon(\eta)_{\text{FSR}}^\pm$  is a factor, defined as a ratio of number of events within the  $\eta$  and  $p_T$  acceptance before and after the final state radiation. The detector acceptance estimated for  $W^+$  or  $W^-$  in each  $\eta$  bin is marked as  $\mathbb{A}(\eta)^\pm$ . Finally, the number of estimated  $W^+$  or  $W^-$  bosons in each  $\eta$  bin is represented as  $N(\eta)^\pm$ . The final result of differential cross-sections as a function of pseudorapidity are shown in Figure 5.11. The results are compared to theoretical predictions, calculated using MCFM-8.3 [129] at NNLO using CT14NNLO [130], MSTW8nnlo [131], and MMHTnnlo [132] PDF sets with  $\alpha(s) = 0.118$ . The results are in agreement with theory predictions, CT14NNLO shows the best compatibility while the others disagree in some pseudorapidity bins. One bin ( $1.85 < \eta < 2.1$ ) of  $\frac{d\sigma}{d\eta}(W^- \rightarrow \mu^- \bar{\nu}_\mu)$  is in a small disagreement with predictions, the level of disagreement is near 1%.

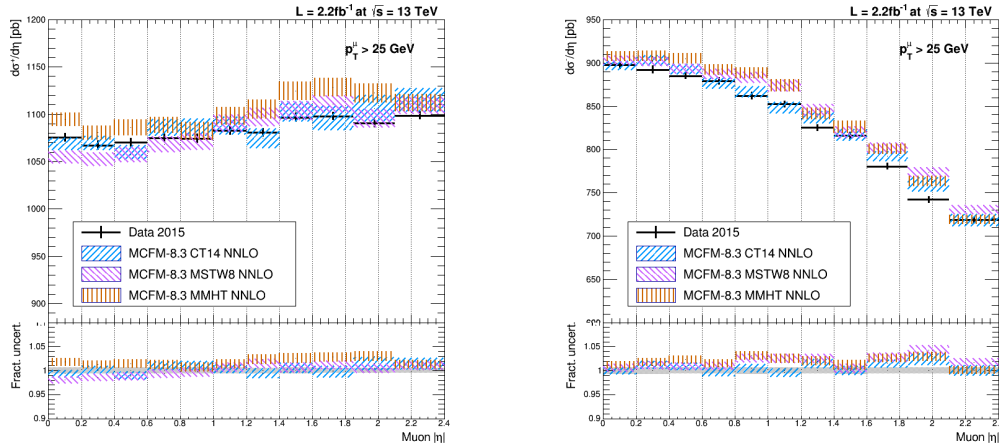


FIGURE 5.11: Differential cross-section values as a function of pseudorapidity represented for  $W^+ \rightarrow \mu^+ \nu_\mu$  and  $W^- \rightarrow \mu^- \bar{\nu}_\mu$  channels. The results are compared to predictions at NNLO calculated in MCFM-8.3 using CT14 NNLO, MSTW8 NNLO, and MMHT NNLO PDF sets.

### 5.9.2 Asymmetry

The  $W^\pm$  boson charge asymmetry is calculated using extracted  $W^\pm$  boson cross-sections using Formula 5.11.

$$\mathcal{A}(\eta) = \frac{\frac{d\sigma}{d\eta}(W^+ \rightarrow \mu^+ \nu_\mu) - \frac{d\sigma}{d\eta}(W^- \rightarrow \mu^- \bar{\nu}_\mu)}{\frac{d\sigma}{d\eta}(W^+ \rightarrow \mu^+ \nu_\mu) + \frac{d\sigma}{d\eta}(W^- \rightarrow \mu^- \bar{\nu}_\mu)}. \quad (5.11)$$

Figure 5.12 represents the final result of the  $W^\pm$  boson charge asymmetry. The asymmetry values are compared to theory predictions calculated from differential cross-sections discussed in the previous paragraph. The overall result is in good agreement with theoretical predictions, as with the differential cross-sections, the asymmetry results are in the best agreement with the predictions calculated using the CT14NNLO PDF set. The overall uncertainty of the asymmetry is at the level of 2% – 5%, depending on  $\eta$  region. The obtained result will be used in the QCD analysis to improve the parton density functions of distributions sensitive to this process.

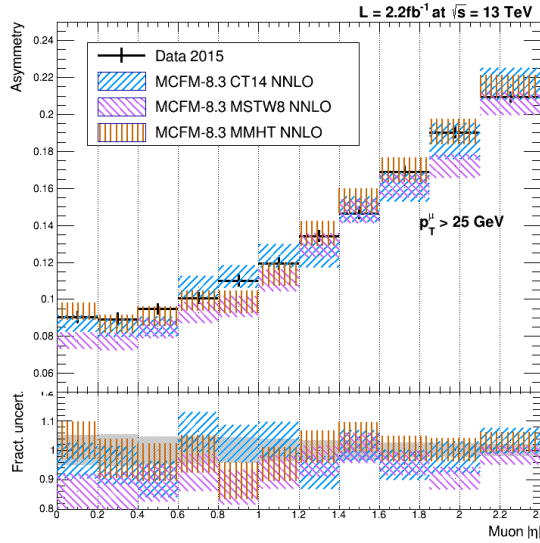


FIGURE 5.12:  $W^\pm$  boson charge asymmetry distribution in the full pseudorapidity range as a function of pseudorapidity. The results are compared to predictions at NNLO calculated in MCFM-8.3 using different PDF sets.

### 5.9.3 Total uncertainties

The final results on relative uncertainties of differential cross-sections of  $W^+$  and  $W^-$  production in muon channel as a function of pseudorapidity are presented in Table 5.13. The table includes statistical and systematic uncertainties. Relative uncertainties were calculated as the highest deviation from the nominal result, together with its error. In the case of asymmetric errors, the error was symmetrized by the highest deviation. Figure 5.13 shows the contribution of each systematic source into the relative uncertainty of differential cross-sections for each channel in each pseudorapidity bin. Similar results, calculated for the  $W^\pm$  boson charge asymmetry are presented in Tab. 5.14.

$\eta$ region	Relative systematic uncertainties in $W^+ \rightarrow \mu^+ \nu_\mu$ channel [%]										
	Stat.	2 pdf	Br. Wig. $\times$ CB	Incl. rec.	Err up	Err down	"a"	pileup up	pileup down	Recoil	Rochester
0.0–0.2	4.78	1.22	1.37	3.02	0.7	0.7	0.91	2.73	2.85	0.94	1.04
0.2–0.4	4.77	0.98	1.01	2.44	0.68	0.79	0.9	1.84	1.97	0.8	0.61
0.4–0.6	4.53	0.8	1.27	2.55	0.66	0.66	1.18	2.18	2.02	1.0	0.87
0.6–0.8	4.65	0.73	1.25	1.61	0.63	0.6	3.12	1.2	1.25	1.04	0.73
0.8–1.0	4.64	0.85	0.87	1.38	0.73	0.68	1.29	1.04	1.28	1.02	0.75
1.0–1.2	4.76	1.68	0.91	0.93	0.66	1.33	0.7	1.0	0.94	0.86	0.59
1.2–1.4	4.63	0.63	0.7	1.17	0.65	0.64	0.65	0.69	0.62	0.71	0.82
1.4–1.6	4.45	0.72	0.9	1.32	1.12	0.92	1.68	1.36	0.84	0.73	0.6
1.6–1.85	4.77	1.0	0.74	1.45	0.58	0.58	0.93	0.72	0.71	0.69	0.62
1.85–2.1	4.36	1.13	0.96	1.41	0.59	0.59	1.19	0.77	0.72	0.84	1.1
2.1–2.4	4.31	0.89	0.5	1.67	0.58	0.57	1.16	0.65	0.72	0.7	0.5
$\eta$ region	Relative systematic uncertainties in $W^- \rightarrow \mu^- \bar{\nu}_\mu$ channel [%]										
	Stat.	2 pdf	Br. Wig. $\times$ CB	Incl. rec.	Err up	Err down	"a"	pileup up	pileup down	Recoil	Rochester
0.0–0.2	4.41	1.32	1.59	3.49	0.78	0.78	1.02	3.13	3.33	1.06	0.72
0.2–0.4	4.42	0.86	1.07	2.81	0.76	0.84	1.01	2.1	2.25	0.9	0.65
0.4–0.6	4.17	0.83	1.46	2.98	0.74	0.75	1.37	2.54	2.36	1.14	0.96
0.6–0.8	4.25	0.81	1.42	1.89	0.7	0.67	3.72	1.45	1.39	1.2	0.87
0.8–1.0	4.30	0.88	0.91	1.63	0.84	0.78	1.53	1.24	1.52	1.19	0.86
1.0–1.2	4.17	1.59	1.06	1.08	0.76	1.62	0.81	1.22	1.17	1.02	0.69
1.2–1.4	3.94	0.75	0.83	1.41	0.76	0.76	0.76	0.85	0.71	0.85	1.11
1.4–1.6	4.24	0.77	1.13	1.62	1.28	0.99	2.15	1.7	1.05	0.89	0.73
1.6–1.85	3.72	1.01	0.98	1.88	0.7	0.7	1.17	0.92	0.91	0.84	0.78
1.85–2.1	3.64	1.28	1.24	1.89	0.74	0.73	1.59	1.02	0.89	1.08	0.8
2.1–2.4	3.27	1.07	0.79	2.31	0.73	0.73	1.58	0.82	0.96	0.92	0.64

TABLE 5.13: Relative systematic uncertainties of differential cross-sections in  $W^+ \rightarrow \mu^+ \nu_\mu$  (upper table) and  $W^- \rightarrow \mu^- \bar{\nu}_\mu$  (lower table) channels.

$\eta$ region	Relative systematic uncertainties of the $W^\pm$ boson charge asymmetry [%]										
	Stat.	2 pdf	Br.Wig. $\times$ CB	Incl. rec.	Err up	Err down	"a"	pileup up	pileup down	Recoil	Rochester
0.0 – 0.2	0.003	0.001	0.885	2.161	0.044	0.044	0.279	1.985	2.408	0.277	2.066
0.2 – 0.4	0.003	0.001	0.050	1.748	0.052	0.056	0.227	1.126	1.246	0.233	0.100
0.4 – 0.6	0.003	0.001	0.690	1.989	0.030	0.040	0.540	1.601	1.492	0.409	0.123
0.6 – 0.8	0.003	0.001	0.577	1.081	0.044	0.027	2.697	0.944	0.387	0.432	0.422
0.8 – 1.0	0.003	0.001	0.174	0.780	0.067	0.110	0.667	0.560	0.715	0.388	0.156
1.0 – 1.2	0.003	0.001	0.245	0.245	0.052	0.845	0.086	0.558	0.508	0.304	0.096
1.2 – 1.4	0.003	0.002	0.101	0.534	0.045	0.046	0.030	0.220	0.145	0.142	0.708
1.4 – 1.6	0.003	0.002	0.377	0.637	0.180	0.144	1.127	0.562	0.321	0.153	0.067
1.6 – 1.85	0.003	0.002	0.354	0.934	0.020	0.021	0.386	0.274	0.260	0.100	0.162
1.85 – 2.1	0.003	0.002	0.365	0.883	0.037	0.039	0.673	0.310	0.095	0.287	1.054
2.1 – 2.4	0.003	0.002	0.144	1.154	0.031	0.034	0.625	0.071	0.239	0.182	0.002

TABLE 5.14: Relative systematic uncertainties of the  $W^\pm$  boson charge asymmetry.

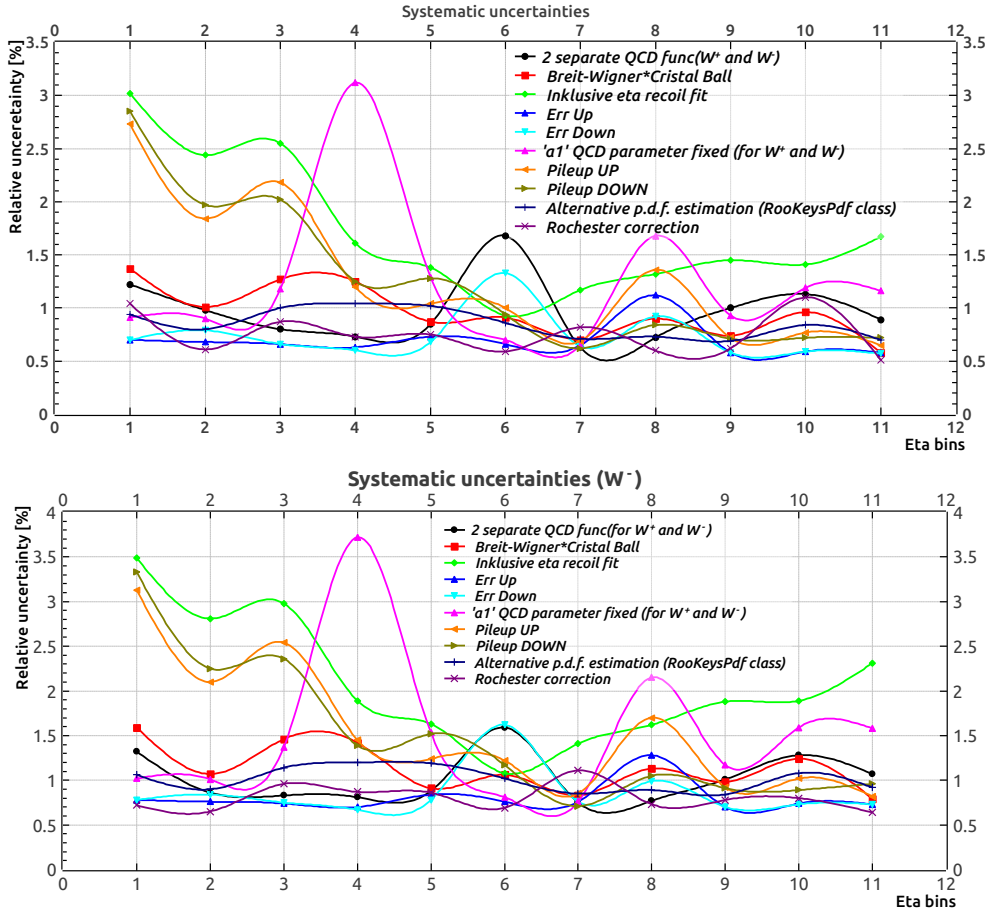


FIGURE 5.13: Systematic uncertainty figures represents the contribution of each source of systematic as a deviation from the cross-section results. Results are shown for  $W^+ \rightarrow \mu^+ \nu_\mu$  and  $W^- \rightarrow \mu^- \bar{\nu}_\mu$  channels on the left and right figures respectively. Measured values are presented with statistical and systematic uncertainties.

## 5.10 Results

The differential cross-sections of  $\frac{d\sigma}{d\eta}(W^+ \rightarrow \mu^+ \nu_\mu)$  and  $\frac{d\sigma}{d\eta}(W^- \rightarrow \mu^- \bar{\nu}_\mu)$  boson production, in muon decay channel were measured using existing, non-published analysis "Measurement of inclusive  $W^\pm$  and  $Z^0$  boson cross-section in pp collisions at  $\sqrt{s} = 13$  TeV using 2015 data set recorded with CMS detector" of MIT group (in particular A. Apyan, K. Bierwagen, S. Brandt, M. D'Alfonso, M. Klute, A. Marini, X. Niu, J. Salfeld and S. Tkaczyk) [115]. The analysis was adopted for the purpose of  $W^\pm$  boson charge asymmetry extraction. In particular, various stages of the unpublished analysis were changed in order to measure differential cross-sections and to reduce correlations between pseudorapidity bins. Eleven sources of systematic uncertainties are considered. Obtained results are compared to theory predictions calculated using MCFM-8.3 with different PDF sets. Final results of measurement of  $W^+$  and  $W^-$  boson differential cross-section as a function of pseudorapidity are presented in tables 5.15 and 5.16, respectively. Tables show measured values with its uncertainty range, compared to theory predictions calculated using CT14 NNLO, MSTW8 NNLO, and MMHT NNLO PDF sets. The final results of the asymmetry measurement are shown in Tab. 5.17.

$\eta$ region	$\frac{d\sigma}{d\eta}(W^+ \rightarrow \mu^+ \nu_\mu)$			
	Data	CT14 NNLO	MSTW8 NNLO	MMHT NNLO
0.0 – 0.2	$1075.60 \pm 4.78 \pm 5.65$	$1069.0 \pm 6.79$	$1055.0 \pm 6.78$	$1094.5 \pm 7.35$
0.2 – 0.4	$1067.17 \pm 4.77 \pm 4.27$	$1070.0 \pm 7.54$	$1052.5 \pm 7.15$	$1080.5 \pm 7.48$
0.4 – 0.6	$1070.22 \pm 4.53 \pm 4.66$	$1060.0 \pm 7.97$	$1057.5 \pm 8.1$	$1086.0 \pm 8.32$
0.6 – 0.8	$1075.10 \pm 4.65 \pm 4.46$	$1086.5 \pm 8.8$	$1068.5 \pm 8.6$	$1088.5 \pm 8.68$
0.8 – 1.0	$1074.50 \pm 4.76 \pm 3.23$	$1086.5 \pm 9.31$	$1071.5 \pm 8.85$	$1082.5 \pm 9.61$
1.0 – 1.2	$1083.00 \pm 4.63 \pm 3.20$	$1088.5 \pm 9.44$	$1090.5 \pm 9.28$	$1098.0 \pm 9.71$
1.2 – 1.4	$1080.84 \pm 4.45 \pm 2.37$	$1075.0 \pm 10.4$	$1097.0 \pm 9.79$	$1105.5 \pm 10.24$
1.4 – 1.6	$1096.63 \pm 4.77 \pm 3.42$	$1103.0 \pm 10.48$	$1104.0 \pm 10.01$	$1124.5 \pm 10.03$
1.6 – 1.85	$1097.92 \pm 4.36 \pm 2.67$	$1096.12 \pm 12.54$	$1110.9 \pm 9.28$	$1128.5 \pm 10.03$
1.85 – 2.1	$1090.44 \pm 4.31 \pm 3.06$	$1106.24 \pm 14.49$	$1095.8 \pm 9.57$	$1122.3 \pm 10.26$
2.1 – 2.4	$1098.71 \pm 4.01 \pm 2.77$	$1115.5 \pm 12.48$	$1109.6 \pm 8.99$	$1112.4 \pm 9.16$

TABLE 5.15: The table represents differential cross-section values as a function of pseudorapidity of  $W^+$  boson production in  $W^+ \rightarrow \mu^+ \nu_\mu$  channel. The results are compared to theory predictions calculated using MCFM-8.3 with CT14 NNLO, MSTW8 NNLO, and MMHT NNLO PDF sets. Measured values are presented with statistical and systematic uncertainties, respectively.

$\eta$ region	$\frac{d\sigma}{d\eta}(W^- \rightarrow \mu^- \bar{\nu}_\mu)$			
	Data	CT14 NNLO	MSTW8 NNLO	MMHT NNLO
0.0 – 0.2	$897.35 \pm 4.41 \pm 6.43$	$897.5 \pm 5.42$	$902.65 \pm 5.38$	$907.65 \pm 5.55$
0.2 – 0.4	$892.21 \pm 4.42 \pm 4.79$	$902.9 \pm 5.89$	$901.9 \pm 5.75$	$907.95 \pm 5.85$
0.4 – 0.6	$884.97 \pm 4.17 \pm 5.39$	$893.05 \pm 5.87$	$893.2 \pm 5.91$	$905.1 \pm 6.08$
0.6 – 0.8	$878.71 \pm 4.25 \pm 5.23$	$876.65 \pm 6.58$	$887.0 \pm 6.06$	$892.05 \pm 6.23$
0.8 – 1.0	$861.67 \pm 4.30 \pm 3.74$	$866.65 \pm 6.39$	$883.55 \pm 6.43$	$888.35 \pm 6.72$
1.0 – 1.2	$852.26 \pm 4.17 \pm 3.62$	$848.75 \pm 6.68$	$873.9 \pm 6.77$	$874.05 \pm 6.87$
1.2 – 1.4	$825.20 \pm 3.94 \pm 2.87$	$837.45 \pm 7.14$	$845.95 \pm 6.85$	$841.35 \pm 6.91$
1.4 – 1.6	$816.24 \pm 4.24 \pm 4.18$	$816.95 \pm 6.65$	$819.85 \pm 6.71$	$825.95 \pm 7.04$
1.6 – 1.85	$780.55 \pm 3.72 \pm 3.30$	$792.36 \pm 5.45$	$802.12 \pm 6.19$	$800.52 \pm 6.13$
1.85 – 2.1	$741.95 \pm 3.64 \pm 3.75$	$759.88 \pm 7.98$	$774.04 \pm 5.76$	$763.2 \pm 5.98$
2.1 – 2.4	$718.27 \pm 3.27 \pm 3.69$	$718.53 \pm 7.19$	$730.63 \pm 5.21$	$718.86 \pm 5.24$

TABLE 5.16: The table represents differential cross-section values as a function of pseudorapidity of  $W^-$  boson production in  $W^- \rightarrow \mu^- \bar{\nu}_\mu$  channel. The results are compared to theory predictions calculated using MCFM-8.3 with discussed PDF sets. Measured values are presented with statistical and systematic uncertainties, respectively.

$\eta$ region	Asymmetry			
	Data	CT14 NNLO	MSTW8 NNLO	MMHT NNLO
0.0 – 0.2	$0.0903 \pm 0.0033 \pm 0.0086$	$0.0872 \pm 0.0045$	$0.0778 \pm 0.0045$	$0.0933 \pm 0.0047$
0.2 – 0.4	$0.0892 \pm 0.0033 \pm 0.0065$	$0.0846 \pm 0.0049$	$0.0770 \pm 0.0047$	$0.0867 \pm 0.0048$
0.4 – 0.6	$0.0947 \pm 0.0031 \pm 0.0076$	$0.0854 \pm 0.0052$	$0.0842 \pm 0.0052$	$0.0908 \pm 0.0053$
0.6 – 0.8	$0.1005 \pm 0.0032 \pm 0.0079$	$0.1068 \pm 0.0057$	$0.0928 \pm 0.0055$	$0.0991 \pm 0.0055$
0.8 – 1.0	$0.1099 \pm 0.0033 \pm 0.0065$	$0.1125 \pm 0.0060$	$0.0961 \pm 0.0057$	$0.0985 \pm 0.0061$
1.0 – 1.2	$0.1192 \pm 0.0032 \pm 0.007$	$0.1237 \pm 0.0062$	$0.1102 \pm 0.0060$	$0.1135 \pm 0.0062$
1.2 – 1.4	$0.1341 \pm 0.0031 \pm 0.0065$	$0.1242 \pm 0.0069$	$0.1292 \pm 0.0064$	$0.1356 \pm 0.0067$
1.4 – 1.6	$0.1465 \pm 0.0034 \pm 0.0096$	$0.1489 \pm 0.0069$	$0.1476 \pm 0.0066$	$0.1530 \pm 0.0066$
1.6 – 1.85	$0.1689 \pm 0.0031 \pm 0.0092$	$0.1608 \pm 0.0080$	$0.1614 \pm 0.0062$	$0.1700 \pm 0.0066$
1.85 – 2.1	$0.1901 \pm 0.0032 \pm 0.0118$	$0.1856 \pm 0.0098$	$0.1720 \pm 0.0065$	$0.1904 \pm 0.0069$
2.1 – 2.4	$0.2093 \pm 0.0030 \pm 0.0127$	$0.2164 \pm 0.0088$	$0.2059 \pm 0.0063$	$0.2149 \pm 0.0064$

TABLE 5.17: The table represents the asymmetry values as a function of pseudorapidity. The results are compared to theory predictions calculated using MCFM-8.3 with discussed PDF sets. Measured values are presented with statistical and systematic uncertainties, respectively.



## Chapter 6

# QCD analysis of $W^\pm$ boson charge asymmetry

This chapter describes the QCD analysis of the  $W^\pm$  boson charge asymmetry measurement extracted in the previous section. The QCD analysis is performed at NLO using a global QCD fit approach, implemented in the xFitter framework [1].

### 6.1 Previous results

The importance of  $W^\pm$  boson charge asymmetry as a source of information on the parton distributions for the proton was first explored by E.L. Berger, F. Halzen, C.S. Kim and S. Willenbrock in [133] and [134]. In  $p\bar{p}$  collisions at Tevatron, the main channels of the direct  $W^\pm$  boson production are  $u\bar{d} \rightarrow W^+$  and  $d\bar{u} \rightarrow W^-$ , where antiquarks are more likely to originate from valence quarks of anti-proton. D0 and CDF experiments have reported on successful measurement of the  $W^\pm$  boson charge asymmetry using  $9.7 \text{ fb}^{-1}$  and  $1.0 \text{ fb}^{-1}$  data sets, respectively, [135], [136]. In the LHC era, the asymmetry studies of CMS collaboration have made an important contribution in a proton structure determination. Unlike the  $p\bar{p}$  collisions, the antiquarks in the direct  $W^\pm$  boson production originate only from sea-quarks. This makes the  $W^\pm$  boson charge asymmetry, extracted from  $pp$  collisions, more sensitive to the sea antiquarks distributions than similar effect obtained in  $p\bar{p}$  collisions. In particular, during the Run 1, asymmetries were extracted at  $\sqrt{s} = 7 \text{ TeV}$  and  $\sqrt{s} = 8 \text{ TeV}$  resulting in two main articles [137] and [138], respectively. Figure 6.1 shows the parton density functions of valence up and down quarks, estimated using combined DIS data from HERA with the CMS result of  $W^\pm$  boson charge asymmetry measurement at  $\sqrt{s} = 8 \text{ TeV}$  using a data sample with an integrated luminosity of  $18.8 \text{ fb}^{-1}$ . Distributions are presented with total uncertainties. To compare obtained results with previous, additional fit with only HERA data was performed. In this comparison, the impact of the new data can be seen in the reduction of the total uncertainty. Improvement in the precision of valence quark distributions is demonstrated in the range of  $10^{-3} < x < 10^{-1}$ . This result, together with a similar measurement performed at  $\sqrt{s} = 7 \text{ TeV}$ , motivates to use of these asymmetries in proton structure studies and encourages future measurements of this effect.

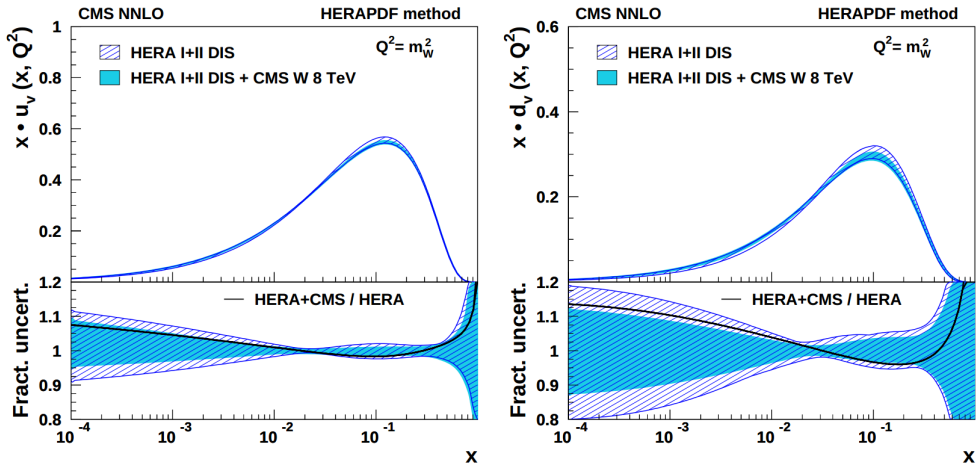


FIGURE 6.1: Distributions of  $x \cdot d_v$  valence (right) and  $x \cdot u_v$  valence (left) quarks as functions of  $x$  at the scale  $Q^2 = m_W^2$ ,  $\text{GeV}^2$ . The results of the fit to HERA data only (hatched band), and to the HERA data and muon asymmetry measurements (light shaded band) are compared. In the bottom panels, the distributions are normalized to 1 for a direct comparison of the uncertainties. The change of the PDFs with respect to the HERA-only fit is represented by a solid line. [138].

## 6.2 xFitter

xFitter (former HERAFitter) [1] is an open-source QCD framework that contains various theoretical and analytical methods to include experimental results in PDF fit. The framework can be categorized into four main constituents, as presented in Fig. 6.2.

xFitter holds a collection of various experimental results, sensitive to different components of proton structure. DIS measurements performed in neutral current (NC) and charged current (CC) channels at HERA are well known due to its direct sensitivity to quark distributions, and indirect sensitivity to gluons. The HERA experimental results are broadly used in different PDF groups; the main part of the data covers the low and middle  $x$  region, as shown in Fig. 1.6. Constraints on heavy quark contribution as well as gluons at high  $x$  are obtained from experimental results of various Drell-Yann processes, top quark production, jets production, collected from  $p\bar{p}$ , and  $pp$  collisions at Tevatron and LHC.

The theoretical part of the xFitter provides various parameterization schemes of PDF. One of the most conventional schemes is Standard Polynomials that can be used for all distributions (valence and sea quarks, gluons). It uses a polynomial as a function of  $x$ , and for each given parton  $j$  is given by

$$xf_j(x) = A_j x^{B_j} (1-x)^{C_j} P_j(x), \quad (6.1)$$

where polynomial  $P_j(x)$  can be different. In this thesis, HERAPDF-style parameterization is used, the polynomial is defined with  $P_j(x) = (1 + \epsilon_j \sqrt{x} + D_j x + E_j x^2)$ .

PDF evolution from the starting scale  $Q_0^2$  to a considered scale can be performed with a few different QCD evolution schemes. xFitter provides DGLAP formalism [40] [41], embodied in QCDNUM program [139], and CCFM evolution scheme [140], implemented

in uPDFevolv [141]. Theoretical predictions of cross-sections can be used from various tools, including MCFM [142].

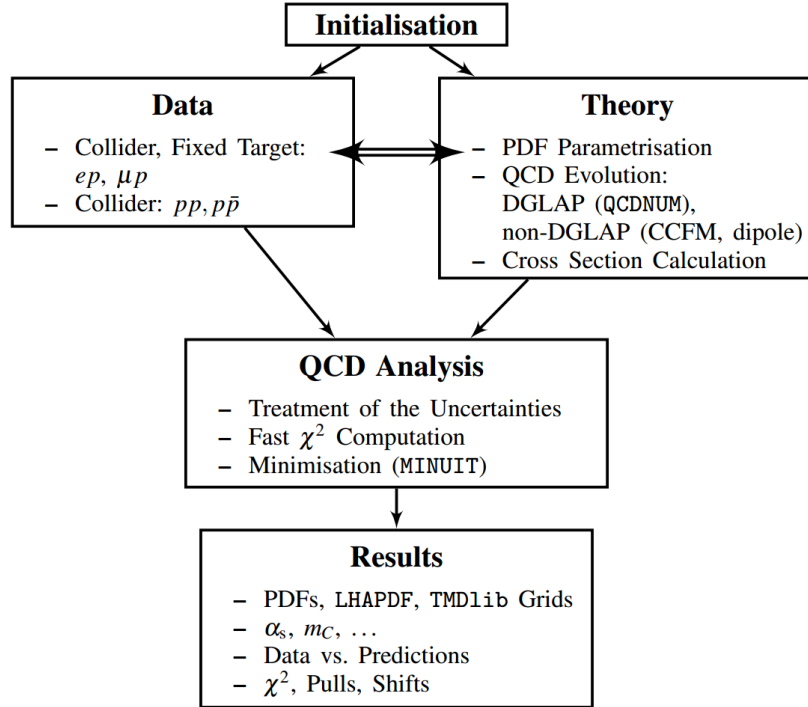


FIGURE 6.2: Schematic representation of the xFitter constituents [1].

Parameters of PDFs are estimated through the  $\chi^2$  function minimization using the MINUIT program [125]. The xFitter provides a list of options for the account of correlated statistical and systematic uncertainties in various forms of the  $\chi^2$  representation. The list includes the Full Covariance Matrix approach, use of Nuisance Parameters, and Mixed Form. In particular, the treatment of nuisance parameters can be performed using Offset or Hessian methods. More details on these methods can be found in [1].

Theoretical uncertainties from renormalization and factorization scale variations, including strong coupling variation and predictions with different PDF sets, are time/resource consuming. For that reason, fast techniques like k-factor and fast grid (APPLgrid [143] and fastNLO [144]) are used. In particular, a certain contribution to the APPLgrid modernization was performed by the author in order to prepare it for the NNLO upgrade. More details on that study may be found in Appendix C.

### 6.3 QCD analysis

Two separate QCD analyses are presented. The main difference is the data sets used in these analyses and some of the settings. In both analyses the parton distribution functions are defined for valence quarks,  $xu_v(x)$ ,  $xd_v(x)$ , anti-up,  $x\bar{u}(x)$ , and anti-down,  $x\bar{d}(x)$ , quarks, anti-strange quark<sup>1</sup>,  $x\bar{s}(x)$ , and gluon,  $xg(x)$  using standard polynomials parameterization scheme (Eq.6.1).

The impact of  $W^\pm$  boson charge asymmetry measurement at  $\sqrt{s} = 13$  TeV on the proton structure determination is performed in combination with other data sets,

<sup>1</sup>Parameterization is done only in one analysis.

sensitive to the proton structure. The list of selected data sets includes DIS processes of neutral and charged current scattering, collected at HERA [4]. This data represents combined results from H1 [2] and ZEUS [3] experiments, obtained in  $e^\pm p$  collisions at proton beam energies of 920, 820, 575, 460 GeV and an electron/positron beam energy of 27.5 GeV. This data is a core of any QCD analysis due to its high sensitivity to valence and sea quark distributions in a wide kinematic range. The neutral current (NC) cross-sections are sensitive to valence and sea quark distributions in a range of  $6 \cdot 10^{-7} \leq x \leq 0.65$ . The charged current (CC) cross-sections provide access to probe mainly valence quark distributions in a range of  $1.3 \cdot 10^{-2} \leq x \leq 0.4$ .

Previous results of  $W^\pm$  boson charge asymmetry measurements with the CMS detector at  $\sqrt{s} = 7$  TeV [145] and  $\sqrt{s} = 8$  TeV [138] are also used. In these measurements, muons were selected with the transverse momentum higher than 25 GeV,  $p_T > 25$  GeV, in eleven bins of absolute values of pseudorapidity using the same  $\eta$  binning that was used in the  $W^\pm$  boson charge asymmetry measurement, described in this thesis. In addition to the valence quarks sensitivity,  $W^\pm$  asymmetry is sensitive to the strange quark parton density function through suppressed processes  $u_\nu + \bar{s} \rightarrow W^+$  and  $\bar{u} + s \rightarrow W^-$ . Therefore, an additional result, sensitive to proton strange quark content is important to include. The  $W^\pm +$  charm quark processes provide information on strange–anti-strange proton content in a range of  $10^{-3} \leq x \leq 10^{-1}$ . The contribution of the sea quarks into the proton structure is usually expressed through a strangeness suppression factor, defined as

$$k_s(Q^2) = \frac{\int_0^1 [\bar{s}(Q^2, x) + s(Q^2, x)] dx}{\int_0^1 [\bar{u}(Q^2, x) + \bar{d}(Q^2, x)] dx}. \quad (6.2)$$

The strangeness suppression  $r_s$  is defined as an integrand of eq. 6.2. The list of selected data sets includes measurements of associated  $W^\pm +$  charm production with the CMS detector at  $\sqrt{s} = 7$  TeV [146] and  $\sqrt{s} = 13$  TeV [147]. Each measurement was performed in five bins of absolute values of pseudorapidity. Results of extracted  $W^\pm$  boson charge asymmetry at  $\sqrt{s} = 7$  TeV,  $\sqrt{s} = 8$  TeV, and measurements of differential cross-sections of associated  $W^\pm +$  charm production at  $\sqrt{s} = 7$  TeV,  $\sqrt{s} = 13$  TeV, used in this analysis are shown in Fig. 6.3.

The nuisance parameter representation is used in the fit for the treatment of systematic uncertainties from data. The  $\chi^2$  is defined as

$$\chi^2(\mathbf{m}, \mathbf{s}) = \sum_i \frac{\left[ m^i - \sum_j \gamma_j^i m^i s_j - \mu^i \right]^2}{\delta_{i,\text{stat}}^2 \mu^i m^i + \delta_{i,\text{uncor}}^2 (m^i)^2} + \sum_j s_j^2 + \sum_i \ln \frac{\delta_{i,\text{stat}}^2 \mu^i m^i + (\delta_{i,\text{uncor}} m^i)^2}{(\delta_{i,\text{stat}}^2 + \delta_{i,\text{uncor}}^2) (\mu^i)^2}. \quad (6.3)$$

Here  $\mu_i$  is the measured value at the point  $i$ ,  $\sum_j \gamma_j^i$  are the relative correlated systematic uncertainties,  $\delta_{i,\text{stat}}$  and  $\delta_{i,\text{uncor}}$  are relative statistical and uncorrelated systematic uncertainties. Vector  $\mathbf{m}$  corresponds to a vector of results of the averaging for the cross-sections [148]. Summation over  $i$  is performed overall points on the  $(x, Q^2)$  grid. Summation over  $j$  runs overall correlated systematic uncertainties. Vector  $\mathbf{s}$  represents correlated shifts of the cross-sections in units of sigma of the corresponding correlated systematic uncertainties. Logarithmic term is used to minimize biases [149].

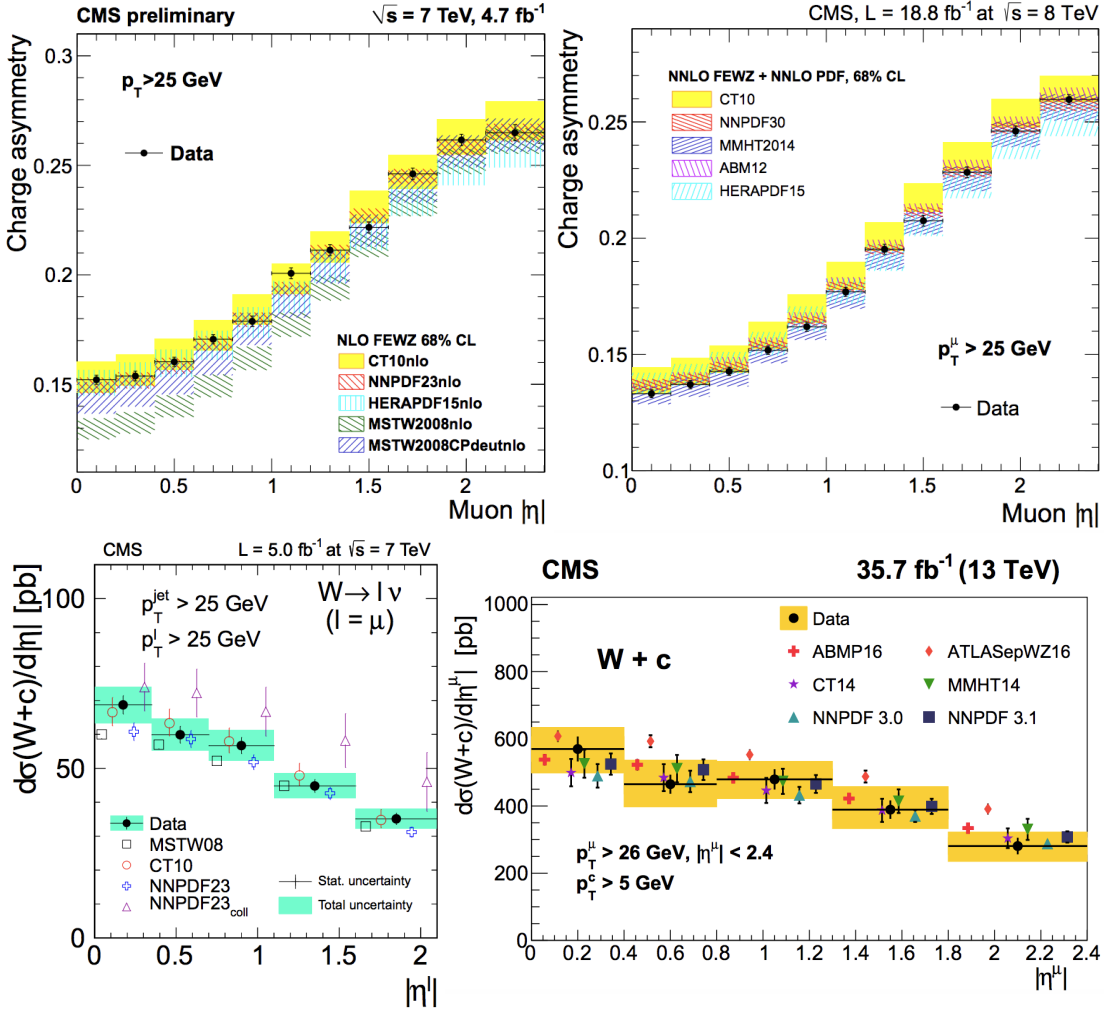


FIGURE 6.3: The  $W^\pm$  boson charge asymmetry values, extracted at  $\sqrt{s} = 7$  TeV (left) [145] and  $\sqrt{s} = 8$  TeV (right) [150], are shown in the first row. Asymmetries are shown as comparison with theory predictions calculated at NLO and NNLO, using different PDF sets. The differential cross-sections of the  $W^\pm + c$  production as a function of absolute values of pseudorapidity, measured at  $\sqrt{s} = 7$  TeV (left) [146], and  $\sqrt{s} = 13$  TeV (right) [147], are shown in the second row. The measurements are compared with theory predictions calculated at NLO, using different PDF sets.

The QCD evolution of parton densities is performed using DGLAP formalism [40, 41], implemented in QCDNUM package of version 07-01-13 [139]. Starting scale for HERA data, for pQCD methods to be applicable, is chosen to be  $Q_{\min}^2 \geq 3.5 \text{ GeV}^2$ . The starting scale of the fit is  $Q_0^2 = 1.9 \text{ GeV}^2$ . The treatment of heavy-quark contributions is done using the Thorne-Roberts general mass variable flavor number scheme [151, 152]. Heavy quark masses are set to be  $m_{\text{charm}} = 1.5 \text{ GeV}$  and  $m_{\text{bottom}} = 4.5 \text{ GeV}$ , the strong coupling is set to  $\alpha_s = 0.118$ , renormalization and factorization scales are of the scale of interaction,  $\mu_f = \mu_r = Q$ .

### 6.3.1 Estimation of PDF uncertainties

The PDF uncertainties are evaluated in accordance with the HERAPDF 2.0 strategy [4]. Three types of uncertainties are considered: experimental, model, and parameterization uncertainties. Propagation of experimental uncertainties to PDF, with the tolerance criterion of  $\Delta\chi^2 = 1$ , which corresponds to 68% confidence level, is performed using the Hessian method [153].

Model uncertainty is evaluated from fit results, performed with a different charm and bottom quark masses. Variation of charm quark mass was performed in two separate fits with  $m_{\text{charm}} = 1.45$  GeV and  $m_{\text{charm}} = 1.55$  GeV. Variation of the bottom quark mass was obtained in a similar way with  $m_{\text{bottom}} = 4.25$  GeV and  $m_{\text{bottom}} = 4.75$  GeV. In addition  $Q_{\text{min}}^2$  is set to 2.5 GeV and 5 GeV. All model uncertainties are added in quadrature.

Parameterization uncertainty in each analysis is calculated as an envelope of a set of fits obtained with the inclusion of an additional D and E parameters. An additional systematic includes a variation of the starting scale,  $Q_0^2 = 1.6$  GeV<sup>2</sup>, and  $Q_0^2 = 2.2$  GeV<sup>2</sup>.

In the first QCD analysis, model uncertainty of strange quark content originates from varying values of strange fraction,  $f_s$ . Details are given in the next subsection. The second QCD analysis uses the parameterization of strange quarks. Details of parameterization and systematic uncertainty evaluation are given in 6.3.3. The total uncertainty for each of the QCD analyses is calculated, adding in quadrature all three uncertainty types.

### 6.3.2 QCD analysis with the new $W^\pm$ boson charge asymmetry and HERA data.

The first analysis is done to estimate the sensitivity of the new  $W^\pm$  boson charge asymmetry result to the proton structure. For that reason, the global fit is performed using only the HERA data. The sensitivity can be shown comparing total relative uncertainties of two fits, with and without the  $W^\pm$  boson charge asymmetry obtained in this thesis. Parameters of the standard polynomials and hence their form, are defined using a so-called "parameterization scan". In this procedure, suitable parameters are defined in an iterative approach, through the  $\chi^2$  monitoring. The first fit starts only with ten parameters, with all D, E, and  $A_g$  being set to zero. Each next fit is modified with one new parameter until no more further  $\chi^2$  improvement is observed. After this procedure is done, the resulting fit with thirteen parameters is given by:

$$xg(x) = A_g x^{B_g} (1-x)^{C_g}, \quad (6.4)$$

$$xu_v(x) = A_{u_v} x^{B_{u_v}} (1-x)^{C_{u_v}} (1 + D_{u_v} x + E_{u_v} x^2), \quad (6.5)$$

$$xd_v(x) = A_{d_v} x^{B_{d_v}} (1-x)^{C_{d_v}}, \quad (6.6)$$

$$x\bar{u}(x) = A_{\bar{u}} x^{B_{\bar{u}}} (1-x)^{C_{\bar{u}}} (1 - D_{\bar{u}} x), \quad (6.7)$$

$$x\bar{d}(x) = A_{\bar{d}} x^{B_{\bar{d}}} (1-x)^{C_{\bar{d}}}. \quad (6.8)$$

Here the normalization parameters  $A_{u_v}$ ,  $A_{d_v}$ , and  $A_g$  are determined from the QCD sum rules. The B parameter describes PDF behavior at small x values with additional constraints  $A_{\bar{u}} = A_{\bar{d}}$ ,  $B_{\bar{u}} = B_{\bar{d}}$ . The C parameter corresponds to a shape of a distribution at  $x \rightarrow 1$ . The asymmetry itself does not provide strong constraints on the strange

quarks distribution. For that reason, the strange quark contribution is introduced with  $x$ -independent  $f_s$  value, defined as  $x\bar{s} = f_s x\bar{D}$  at  $\mu_{f_0}^2$ . In this analysis  $f_s = 0.31$ . As a model uncertainty, two additional fits with  $f_s = 0.22$  and  $f_s = 0.4$  are performed.

The quality of the global fit is estimated using a  $\chi^2$  method. Corresponding fit results is shown in Tab. 6.1. The table compares fitted results with and without new measurement of  $W^\pm$  boson charge asymmetry. The partial  $\chi^2$  per number of data points, is shown in the second and third columns for each data set. The final  $\chi^2/N_{\text{dof}}$ , where  $N_{\text{dof}}$  is a number of degrees of freedom, is shown in the last row of the table. The new measurement of  $W^\pm$  boson charge asymmetry is well fitted together with the HERA data. The compatibility of theoretical predictions for the asymmetry with measured values is shown in Fig. 6.4. The  $W^\pm$  boson charge asymmetry measurement at  $\sqrt{s} = 13$  TeV is compared to theory predictions at NLO, using MCFM linked to the APPLgrid. Black dots represent measured values; the yellow band corresponds to the total uncertainty of the measurement. The solid red line corresponds to the fitted theoretical predictions. The red dashed line corresponds to the final result of the most consistent fit obtained using the nuisance parameters methodology of treatment of correlated systematic uncertainties. This method allows to collectively move experimental data points according to their correlated systematic uncertainties. To display the result of such fits and avoid shifting measured data points, the theory curve is shifted instead. Parton density functions are shown in Fig. 6.5. The impact of the new data on the proton structure determined with the HERA data is shown in Fig. 6.6. Distributions of relative total uncertainty of  $\delta x_{u_v}/x_{u_v}$  and  $\delta x_{d_v}/x_{d_v}$  are shown at the scale of 10 GeV<sup>2</sup> (upper row) and  $m_{W^\pm}^2$  (second row). The new result can improve constraints on the valence quarks, as expected.

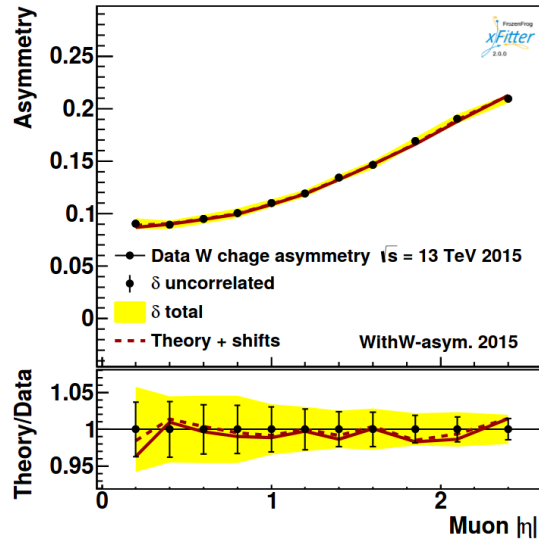


FIGURE 6.4: The comparison of theoretical prediction to the  $W^\pm$  boson charge asymmetry measurement.

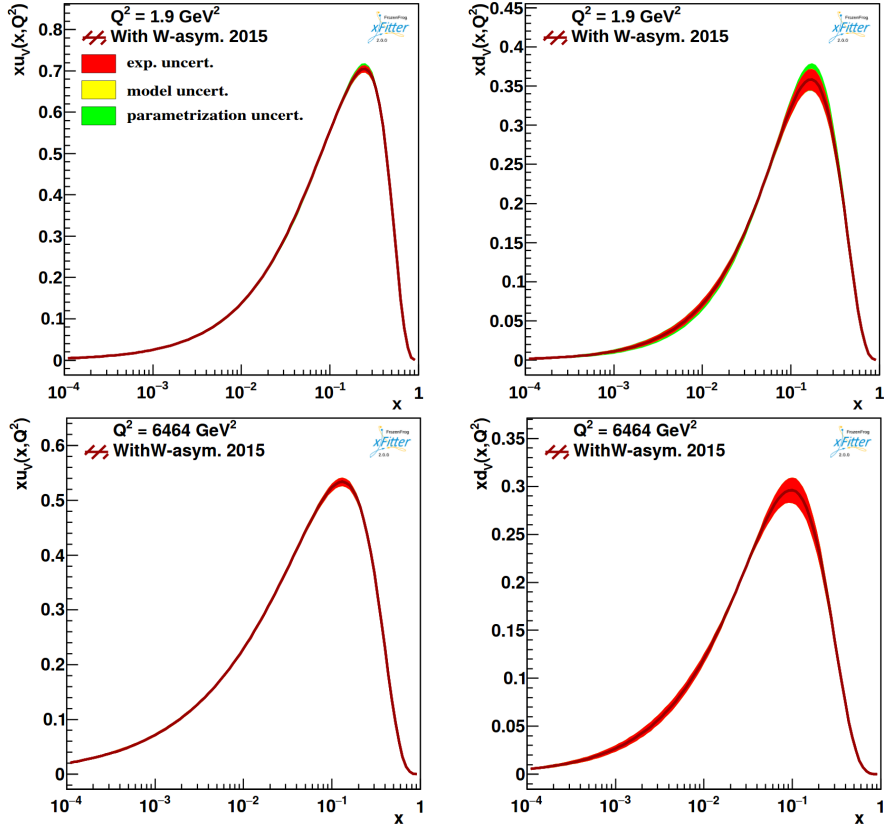


FIGURE 6.5: Obtained parton density functions for  $x \cdot u_v$  and  $x \cdot d_v$ , using HERA data and new  $W^\pm$  boson charge asymmetry results. The first row represents results obtained at the starting scale  $Q^2 = 1.9 \text{ GeV}^2$ , while the second row shows results at the scale  $Q^2 = m_{W^\pm}^2$ . The left plots correspond to  $x \cdot u_v$  distributions, while right are  $x \cdot d_v$ .

Dataset	With W-asym. 2015	No W-asym. 2015
HERA 1+2 CC ( $e^+p$ ) $E_p = 920 \text{ GeV}$	41/39	38/39
HERA 1+2 CC ( $e^-p$ ) $E_p = 920 \text{ GeV}$	53/42	53/42
HERA 1+2 NC ( $e^-p$ ) $E_p = 920 \text{ GeV}$	218/159	218/159
HERA 1+2 NC ( $e^+p$ ) $E_p = 820 \text{ GeV}$	69/70	69/70
HERA 1+2 NC ( $e^+p$ ) $E_p = 920 \text{ GeV}$	440/337	435/337
HERA 1+2 NC ( $e^+p$ ) $E_p = 460 \text{ GeV}$	217/204	217/204
HERA 1+2 NC ( $e^+p$ ) $E_p = 575 \text{ GeV}$	220/254	220/254
CMS $W^\pm$ muon asymmetry 13 TeV (2015)	2.3/11	-
Correlated $\chi^2$	86	87
Log penalty $\chi^2$	+5.5	+8.0
Total $\chi^2/N_{\text{dof}}$	1352/1143	1346/1132

TABLE 6.1: The first column corresponds to the data set name, the second is the  $\chi^2$  per number of data points for the analysis with the asymmetry measured in this thesis, the third column shows results of the fit without the asymmetry.



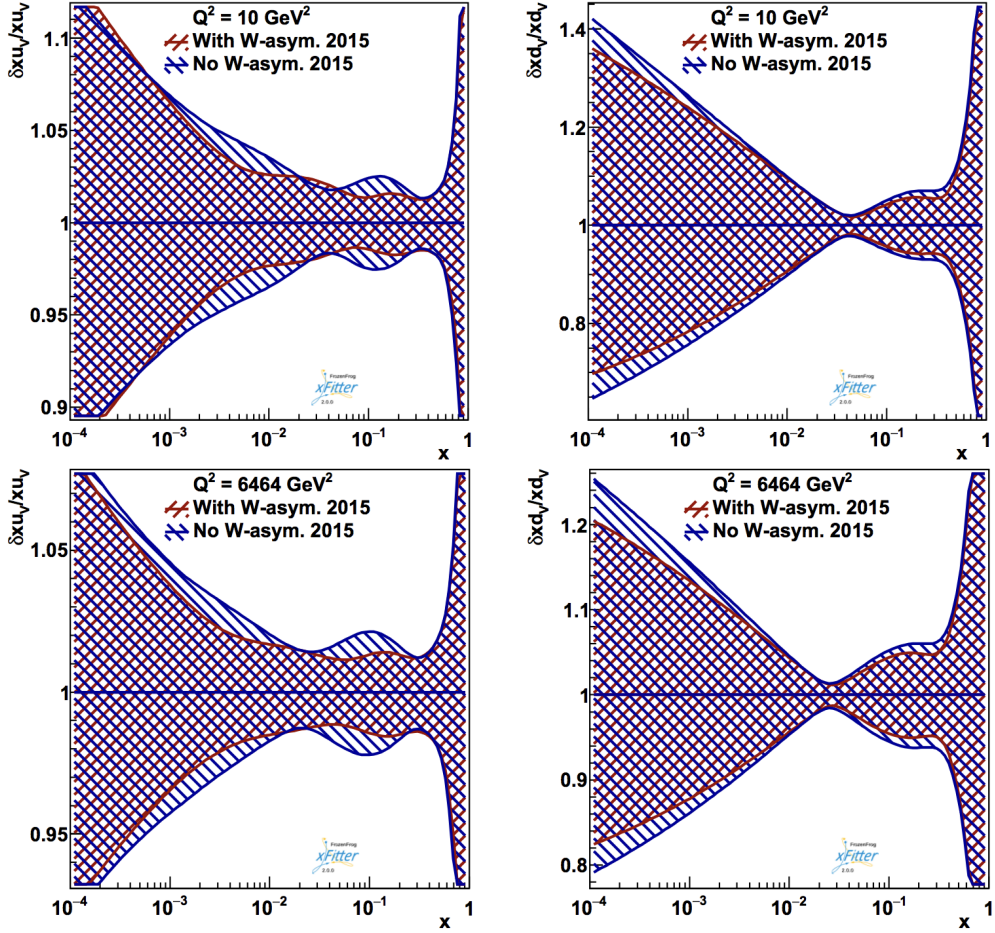


FIGURE 6.6: Distributions of relative uncertainty  $\delta x_{q_v}/x_{q_v}$  of up (left) and down (right) quarks at the scale of  $10 \text{ GeV}^2$ . The second row presents the same distributions at the scale of  $6464 \text{ GeV}^2$ .

### 6.3.3 QCD analysis with the new $W^\pm$ boson charge asymmetry and other sensitive processes.

The second analysis is performed to estimate the impact of the new asymmetry in a context of already existing, similar measurements. In this case, the QCD fit is performed using already mentioned HERA data and the new asymmetry result, with previous results of  $W^\pm$  boson charge asymmetry measurements at  $\sqrt{s} = 7 \text{ TeV}$  [145],  $\sqrt{s} = 8 \text{ TeV}$  [138], and  $W$ +charm measurements at  $\sqrt{s} = 7 \text{ TeV}$  [146] and  $\sqrt{s} = 13 \text{ TeV}$  [147], obtained by CMS. The parameterization scheme is similar to the first QCD analysis, except for some parameters. Since the  $W$ +charm data is used, the strange quark distribution can now be partially fitted. For that reason, additional parameterization for strange quark distribution is introduced, instead of a single  $f_s$  value. The parameterization is

given by

$$xg(x) = A_g x^{B_g} (1-x)^{C_g} - A'_g x^{B'_g} (1-x)^{C'_g}, \quad (6.9)$$

$$xu_v(x) = A_{u_v} x^{B_{u_v}} (1-x)^{C_{u_v}} (1 + E_{u_v} x^2), \quad (6.10)$$

$$xd_v(x) = A_{d_v} x^{B_{d_v}} (1-x)^{C_{d_v}}, \quad (6.11)$$

$$x\bar{u}(x) = A_{\bar{u}} x^{B_{\bar{u}}} (1-x)^{C_{\bar{u}}} (1 - D_{\bar{u}} x), \quad (6.12)$$

$$x\bar{d}(x) = A_{\bar{d}} x^{B_{\bar{d}}} (1-x)^{C_{\bar{d}}}, \quad (6.13)$$

$$x\bar{s}(x) = A_{\bar{s}} x^{B_{\bar{s}}} (1-x)^{C_{\bar{s}}}, \quad (6.14)$$

where the normalization parameters  $A_{u_v}$ ,  $A_{d_v}$ , and  $A_g$  are derived from the QCD sum rules. Additional constraints are given by  $A_{\bar{u}} = A_{\bar{d}}$ ,  $B_{\bar{u}} = B_{\bar{d}}$ . The anti-strange parameter  $B_{\bar{s}}$  is set to be equal to  $B_{\bar{d}}$ , strange–anti-strange quarks are treated as equal,  $x_s = x_{\bar{s}}$ , with  $A_{\bar{s}}$  and  $C_{\bar{s}}$  being as free parameters, resulting in a fifteen parameters fit. For the gluon distribution, an additional term of the form  $A'_g x^{B'_g} (1-x)^{C'_g}$  is subtracted;  $C'_g$  is fixed to  $C'_g = 25$ . The negative gluon term is discussed, for example, in [4].

Partial  $\chi^2$  per number of data points,  $\chi^2/N_{\text{DP}}$ , for each data set is shown in the second column of Table 6.2. The  $\chi^2/N_{\text{dof}}$  of the final fit is shown in the last row. The  $W^\pm$  boson charge asymmetry, extracted in this thesis, shows good compatibility with overall fit resulting in a partial  $\chi^2/N_{\text{DP}} = 6.0/11$ . The total  $\chi^2$  per number of degrees of freedom is shown in the last row,  $\chi^2/N_{\text{dof}} = 1391/1172$ , resulting in a final value close to 1.

The comparison of theory predictions with measurements<sup>2</sup> that are used in the fit are shown in Fig. 6.7. The  $W^\pm$  boson charge asymmetry measurements at  $\sqrt{s} = 7$  TeV and  $\sqrt{s} = 8$  TeV, and  $W^\pm + \text{charm}$  at  $\sqrt{s} = 7$  TeV and  $\sqrt{s} = 13$  TeV are compared to theory predictions at NLO, using MCFM linked to the APPLgrid. Black dots corresponds to measured values; the yellow band represent the total uncertainty of the measurement. A solid red line corresponds to the starting values of theoretical predictions, while the red dashed line corresponds to the shifted predictions.

Additional variation of low- $x$  sea quarks is performed releasing constraints on  $A_{\bar{q}}$  and  $B_{\bar{q}}$  parameters,  $A_{\bar{u}} \neq A_{\bar{d}}$ ,  $B_{\bar{u}} \neq B_{\bar{d}}$ . The final result of parton distribution functions at starting scale  $Q^2 = 1.9 \text{ GeV}^2$  for  $x \cdot u_v$ ,  $x \cdot d_v$ ,  $x \cdot \bar{u}$ ,  $x \cdot \bar{d}$ ,  $x \cdot s$ ,  $x \cdot (s + \bar{s})/(\bar{u} + \bar{d})$ ,  $x \cdot g$  and  $x \cdot \Sigma$  are shown in Fig. 6.8. Experimental uncertainties are shown with the red band, while model uncertainties are presented with the yellow band. The green band shows uncertainties from parameterization variation. For most of the distributions, experimental uncertainties are dominant.

<sup>2</sup>Here the comparison is shown only for CMS results.

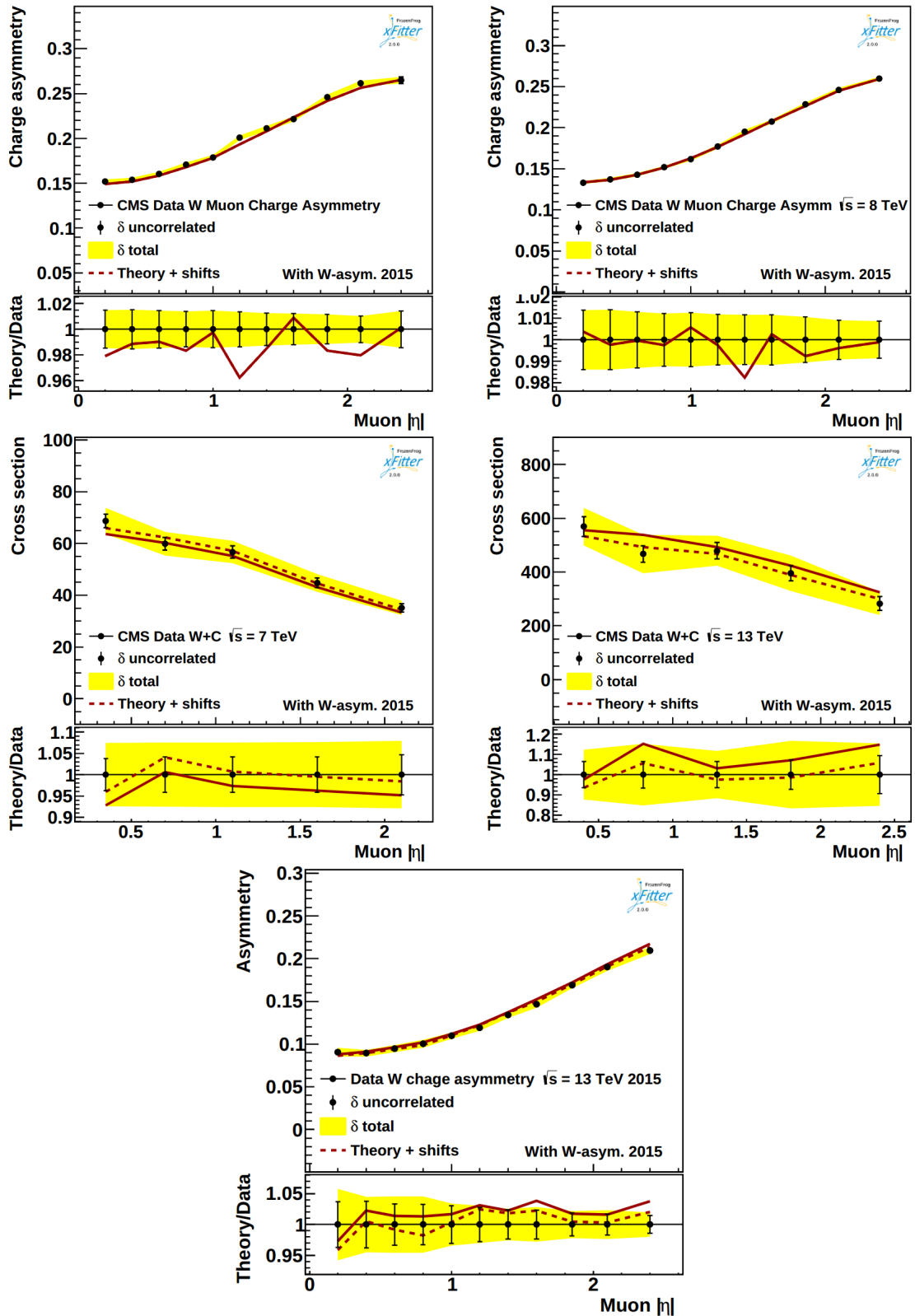


FIGURE 6.7: Comparison of theory predictions to data used in the fit. The first row shows results for the  $W^\pm$  boson charge asymmetry measured at  $\sqrt{s} = 7$  TeV (left) and  $\sqrt{s} = 8$  TeV (right). The second row shows comparison for  $W^\pm$ +charm process, measured at  $\sqrt{s} = 7$  TeV (left) and  $\sqrt{s} = 13$  TeV (right).  $W^\pm$  boson charge asymmetry measured in this thesis is shown in the last row.

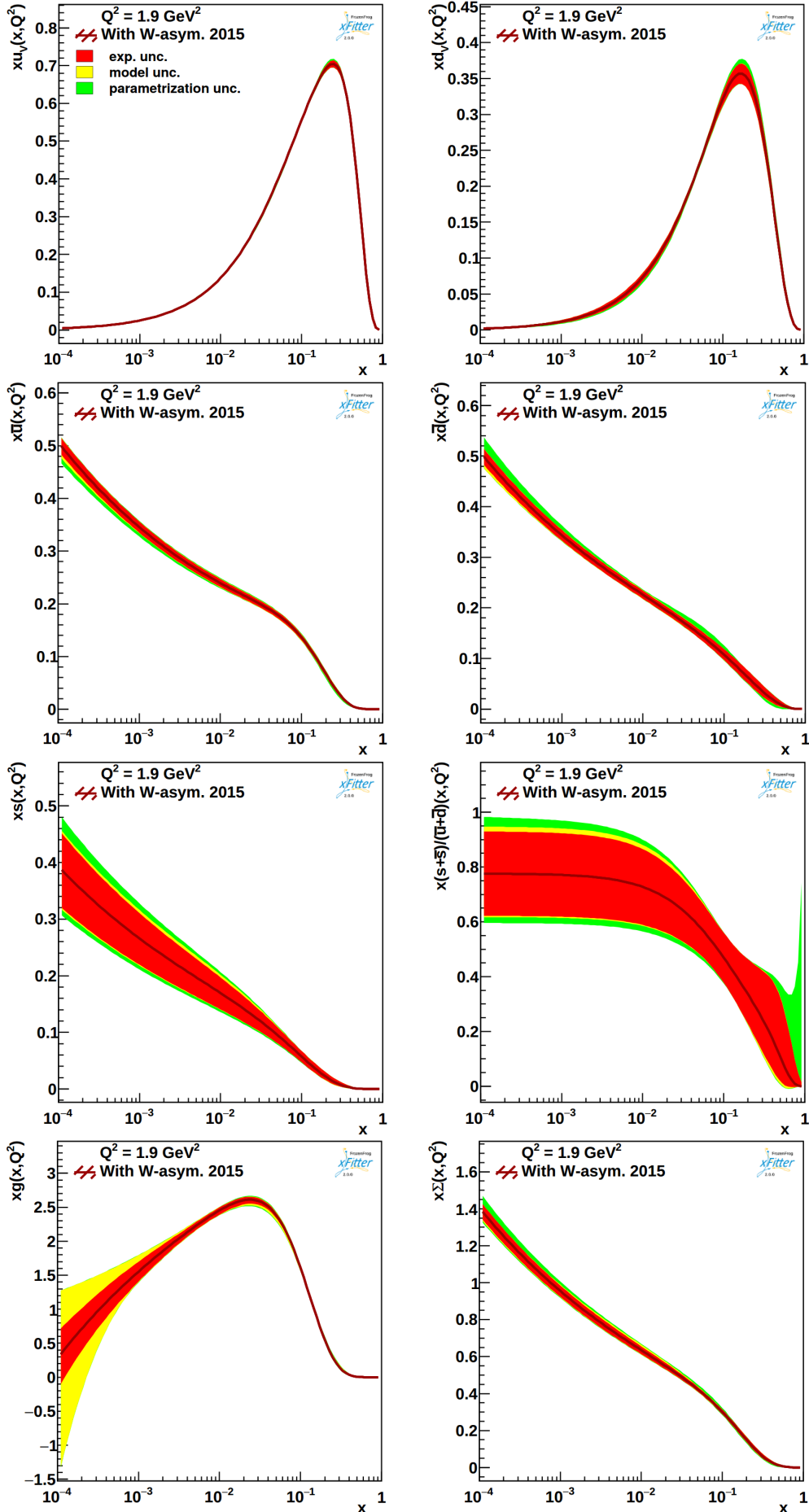


FIGURE 6.8: Obtained parton density functions for  $x \cdot u_v$  and  $x \cdot d_v$  (first row),  $x \cdot \bar{u}$  and  $x \cdot \bar{d}$  (second),  $x \cdot s$  and  $x \cdot (s + \bar{s})/(\bar{u} + \bar{d})$  (third),  $x \cdot g$  and  $x \cdot \Sigma$  (fourth) at starting scale  $Q^2 = 1.9 \text{ GeV}^2$ .

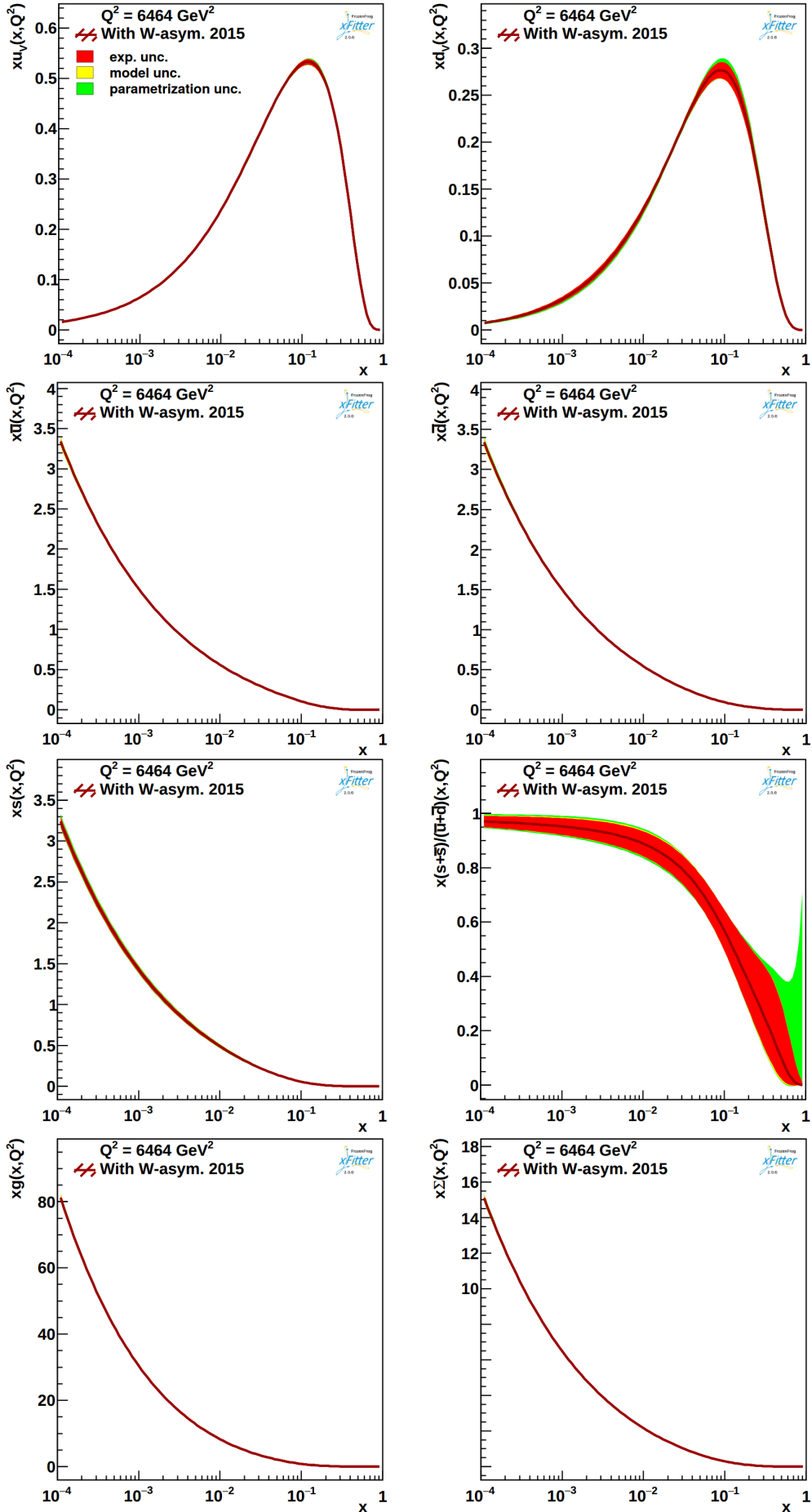


FIGURE 6.9: Obtained parton density functions for  $x \cdot u_v$  and  $x \cdot d_v$  (first row),  $x \cdot \bar{u}$  and  $x \cdot \bar{d}$  (second),  $x \cdot s$  and  $x \cdot (s + \bar{s})/(\bar{u} + \bar{d})$  (third),  $x \cdot g$  and  $x \cdot \Sigma$  (fourth) at starting scale  $Q^2 = m_{W^\pm}^2 \text{ GeV}^2$ .

All mentioned PDFs have also sizable parameterization uncertainty, except for the gluon,  $x \cdot g$ , which at low  $x$  shows high sensitivity to model variations. Similar results at the scale of the  $M_{W^\pm}^2$  are shown in Fig. 6.9. The uncertainties are much smaller except for  $x \cdot (s + \bar{s}) / (\bar{u} + \bar{d})$ , that shows high experimental and parameterization uncertainties.

Estimation of the impact of the new result on the proton structure is done performing already mentioned fit with and without the measured  $W^\pm$  boson charge asymmetry values. The total  $\chi^2/N_{\text{dof}}$  of both fits are shown in the Tab. 6.2. The results show similar total chi-square values for both cases, resulting in  $\chi^2/N_{\text{dof}} \approx 1.19$ .

Data set	With W-asym. 2015	No W-asym. 2015
HERA 1+2 CC ( $e^+p$ ) $E_p = 920$ GeV	41/39	41/39
HERA 1+2 CC ( $e^-p$ ) $E_p = 920$ GeV	60/42	59/42
HERA 1+2 NC ( $e^-p$ ) $E_p = 920$ GeV	221/159	219/159
HERA 1+2 NC ( $e^+p$ ) $E_p = 820$ GeV	69/70	69/70
HERA 1+2 NC ( $e^+p$ ) $E_p = 920$ GeV	443/377	443/377
HERA 1+2 NC ( $e^+p$ ) $E_p = 460$ GeV	218/204	217/204
HERA 1+2 NC ( $e^+p$ ) $E_p = 575$ GeV	220/254	220/254
CMS $W^\pm$ + charm 7 TeV	2.2/5	2.2/5
CMS $W^\pm$ + charm 13 TeV	2.3/5	2.3/5
CMS $W^\pm$ muon asymmetry 7 TeV	15/11	14/11
CMS $W^\pm$ muon asymmetry 8 TeV	4.2/11	3.6/11
CMS $W^\pm$ muon asymmetry 13 TeV (2015)	6.0/11	-
Correlated $\chi^2$	86	85
Log penalty $\chi^2$	+3.7	+6.3
Total $\chi^2/N_{\text{dof}}$	1391/1172	1382/1161

TABLE 6.2: The quality of two fits is presented. The first column corresponds to the data set name, the second is the  $\chi^2$  per number of data points for the analysis with the asymmetry measured in this thesis, the third column shows results of the fit without the asymmetry. The last row shows the total  $\chi^2/N_{\text{dof}}$  of both fits.

Figure 6.10 illustrates the relative total uncertainty for valence quarks, and strange quark at the starting scale  $Q^2 = 1.9$  GeV<sup>2</sup>. The new data reduces uncertainty of the valence up quark for  $x$  in a range of  $0.02 \leq x \leq 10^{-4}$ . The distribution for a down quark shows that the new data slightly reduces uncertainty in the region  $10^{-3} \leq x \leq 10^{-4}$ . An improvement is also seen in the distribution for strange quarks and  $x(s + \bar{s}) / (\bar{u} + \bar{d})$ . Similar results at the scale of  $W^\pm$  boson mass,  $m_{W^\pm}^2$  are shown in Fig. 6.11. At this scale, the only distribution that has a reduction of the relative uncertainty is the valence up quark.

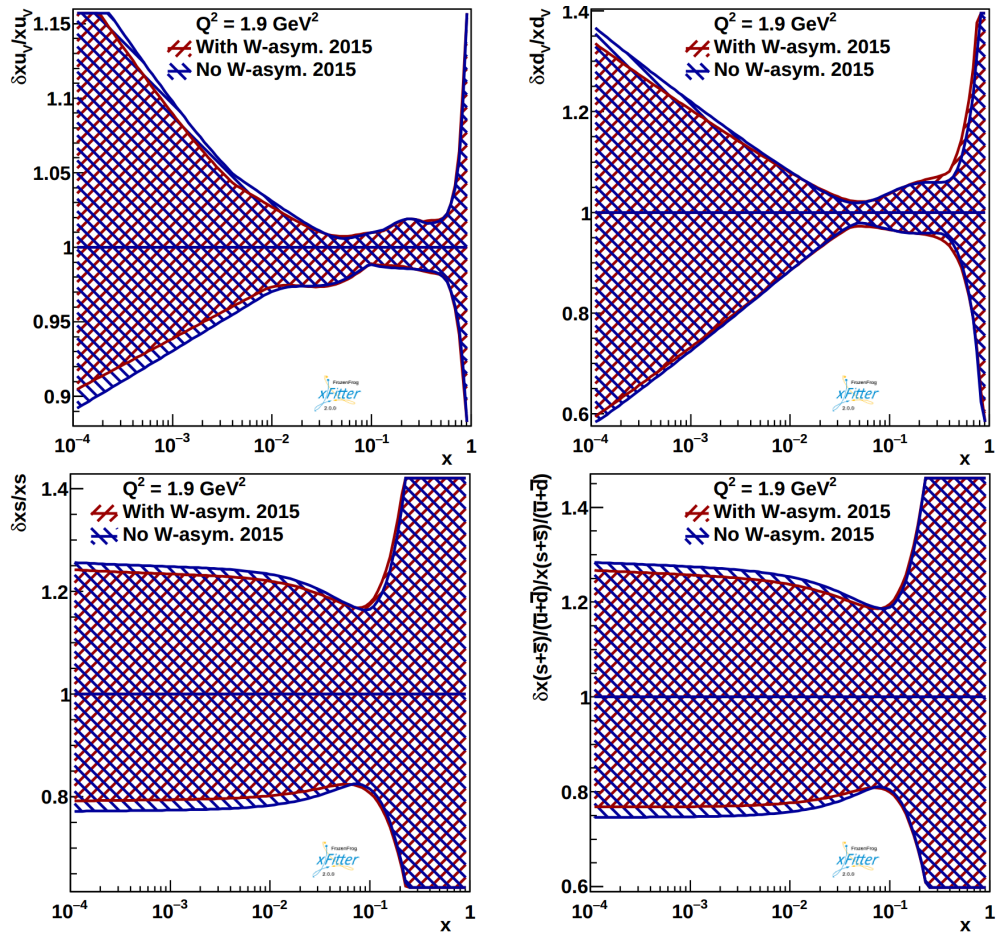


FIGURE 6.10: Distributions of relative uncertainty  $\delta x_{q_v}/x_{q_v}$  of up (1st row, left plot) and down (1st row, right plot) quarks. The second row presents distributions relative uncertainty  $\delta x_s/x_s$  for strange quark (left plot) and  $\delta x(s+\bar{s})/(\bar{u}+\bar{d})/x(s+\bar{s})/(\bar{u}+\bar{d})$  (right plot). All plots show results at the scale of  $1.9 \text{ GeV}^2$ .

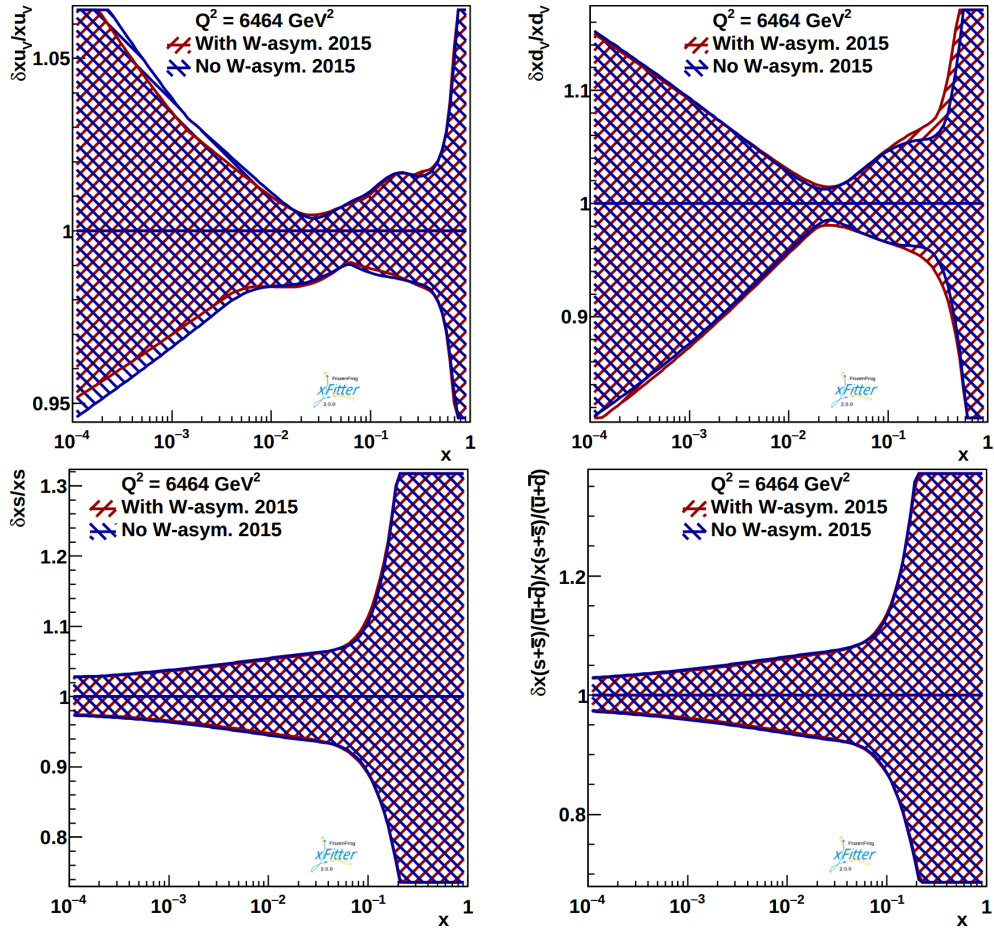


FIGURE 6.11: Distributions of relative uncertainty  $\delta x_{q_v}/x_{q_v}$  of up (1st row, left plot) and down (1st row, right plot) quarks. The second row presents distributions for gluon (left plot) and sea quarks (right plot). Anti-up,  $\delta x_{\bar{u}}/x_{\bar{u}}$  and anti-down,  $\delta x_{\bar{d}}/x_{\bar{d}}$ , quarks are shown on the left and right plots, respectively. All plots show results with and without 2015  $W^\pm$  boson charge asymmetry at the scale of  $m_{W^\pm}^2$ .



## Chapter 7

# Summary and conclusions

A measurement of  $W^\pm$  boson charge asymmetry in the data sample collected at  $\sqrt{s} = 13$  TeV with the CMS detector corresponding to an integrated luminosity of up to  $2.2 \pm 0.05 \text{ fb}^{-1}$  and the QCD analysis with measured values are presented.

Differential cross-sections are measured using  $W^\pm$  bosons events decaying in the muon channel. The event selection is performed selecting muons that satisfy kinematic and fiducial requirements. The estimation of  $W^\pm$  boson yields is obtained from the missing transverse energy distributions using an unbinned maximum-likelihood fit. Calculation of systematic uncertainties of differential cross-sections showed high sensitivity to inclusive  $W^\pm$  boson recoil correction and pileup. The obtained values are compared to theoretical predictions calculated at NNLO using MCFM-8.3 package with CT14 NNLO, MSTW8 NNLO, MMHT NNLO PDF sets. A comparison of differential cross-sections showed good agreement with prediction calculated using the CT14 NNLO PDF set. Asymmetry values are in good agreement with predictions produced using CT14 NNLO and MMHT NNLO PDF sets. The sensitivity of  $W^\pm$  boson charge asymmetry to valence and sea content of proton in the range of  $10^{-3} \leq x \leq 10^{-1}$  was tested by performing a QCD analysis to probe the improvement on the PDF determination. The analysis is done using the open-source QCD framework xFitter. Quantitative estimation of the measured asymmetry sensitivity to proton structure is obtained using HERA I+II data. Experimental uncertainties associated with the input data sets are propagated to PDF uncertainties with the Hessian method, with the tolerance criterion of  $\Delta\chi^2 = 1$ . The results show that the measured asymmetry values can improve PDF uncertainties for valence quarks at different scales. After this result is obtained, the measured asymmetry is used in combination with previous CMS measurements of  $W^\pm$  boson charge asymmetry at  $\sqrt{s} = 7$  TeV and  $\sqrt{s} = 8$  TeV, measurements of  $W$ +charm quark production at  $\sqrt{s} = 7$  TeV,  $\sqrt{s} = 13$  TeV, and HERA I+II data. The inclusion of similar data sets into the global fit significantly constrains the phase space of valence quark distributions.



# Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt zu haben.

Die eingereichte schriftliche Fassung entspricht der auf dem elektronischen Speichermedium.

Die Dissertation wurde in der vorgelegten oder einer ähnlichen Form nicht schon einmal in einem früheren Promotionsverfahren angenommen oder als ungenügend beurteilt.

Hamburg den 29.05.2020

\_\_\_\_\_

Vladyslav Danilov



## *Acknowledgements*

First of all, I would like to thank my supervisors Elisabetta Gallo and Katarzyna Wichmann, for giving me the opportunity to do this wonderful research and gain a lot of experience. I am very grateful for all that support and experience that they gave me and their help in my aspiration to be better. I also would like to thank all CMS DESY group members who were always open to questions, discussions, and chats. Much helpful advice I received from Katerina Lipka, Svenja Pflitsch, Akshansh Singh, Pavel Starovoitov, Oleksi Turkot, and many others. Also, I'm very thankful to Stephanie Brandt, who was always ready to answer my questions and share her opinion in various discussions. I was glad to participate in the APPLgrid upgrade with Pavel Starovoitov. During that time, I have learned a lot. Many thanks to Moritz Guethoff, Olena Karacheban, and Anastasia Velyka for being great teachers in my technical responsibilities.

I would like to thank my aunt, Valentyna Harnaga, and cousin, Yaroslav Tyshenko, for all support they gave me. Finally, I would like to thank my girlfriend, Valeria, who made my life better than it was before.

I thank you all.



# List of Figures

- 1.1 Elementary particles in the Standard model. The three left columns correspond to fermions. The right column is vector gauge bosons, and the last has the scalar boson of Higgs field. For each particle, mass, charge, and spin values are given in their left upper frame corner [9]. . . . . 4
- 1.2 The running of  $\alpha_s(Q)$  as a function of the scale  $Q$ . The solid line and the uncertainty band are calculated by evolving the extracted  $\alpha_s(m_Z)$  values using the 2-loop 5-flavor renormalization group equations [15]. . . . . 7
- 1.3 The strengths of the weak interactions between the six quarks. The "intensities" of the lines are determined by the elements of the CKM matrix [19], [20]. . . . . 9
- 1.4 Diagrams represent deep inelastic scattering via exchange of  $Z^0/\gamma^*$  boson on the left, and  $W^\pm$  on the right. Lepton four-momentum before and after interaction is shown as  $k$  and  $k'$  respectively. The proton initial state momentum is shown as  $p$ . . . . . 11
- 1.5 The proton structure function  $F_p^2$  measured in electromagnetic scattering of electrons and positrons on protons, and for electrons/positrons (SLAC, HERMES, JLAB) and muons (BCDMS, E665, NMC) on a fixed target. Statistical and systematic errors added in quadrature are shown. The H1+ZEUS combined values are obtained from the measured reduced cross-section and converted to  $F_p^2$  with a HERA-PDF NLO fit, for all measured points where the predicted ratio of  $F_p^2$  to reduced cross-section was within 10% of unity. The data are plotted as a function of  $Q^2$  in bins of fixed  $x$  [14]. . . . . 12
- 1.6 Kinematic domains in  $x$  and  $Q^2$  probed by fixed-target and collider experiments, where  $Q^2$  can refer either to the literal  $Q^2$  for deep inelastic scattering, or the hard scale of the process in hadron-hadron collisions, e.g. invariant mass or transverse momentum  $p_T^2$ . Some of the final states accessible at the LHC are indicated in the appropriate regions, where  $y$  is the rapidity. The incoming partons have  $x_{1,2}=(Q/14 \text{ TeV}) e^{\pm y}$  where  $Q$  is the hard scale of the process shown in blue in the figure. [14]. . . . 13
- 1.7 Representation of proton-proton collision in the factorization approach. Each proton carries momentum  $p$  with its non-perturbative PDF,  $f(x)$ . Strong coupling at renormalization scale is shown as  $\alpha_s(\mu_r)$ . Parton-parton interaction area is shown with  $\sigma$ . . . . . 15
- 1.8 Feynman diagrams of  $P_{qq}$ ,  $P_{gq}$ ,  $P_{qg}$ , and  $P_{gg}$  splitting functions of DGLAP equations. [37]. . . . . 16
- 1.9 The parton distribution functions of valence quarks,  $xu_v$ ,  $xd_v$ , scaled sea quarks,  $x\Sigma$ , and gluons,  $xg$  at scale  $Q^2 = 1.9 \text{ GeV}^2$  (left plot) and  $Q^2 = m_W^2 \text{ GeV}^2$  (right plot) [42]. . . . . 17

1.10	Feynman diagram for $W^\pm$ boson production in pp collisions, decaying in the muon channel. Left figure show the diagram for $W^-$ production, $W^+$ is shown on the right. . . . .	17
1.11	Individual contributions of flavor decomposition into the total $W^+$ (solid line) and $W^-$ (dashed line) cross-section at LO [43]. . . . .	18
1.12	Measurements of the total $W^+$ , $W^-$ , $W$ , and $Z^0$ production cross sections times branching fractions as a function of center of mass energy for CMS and experiments at lower-energy colliders [44]. . . . .	19
1.13	$W^\pm$ boson in different helicity states decaying into electron channel in a rest frame. The first row corresponds to three different helicity states, left-handed (negative), longitudinal, and right-handed (positive). Blue arrows are momentum; red arrows are helicity states of daughter particles, while the yellow arrow shows the boson spin projection on the z-axis. The second row shows the matrix element dependence on the theta angle for each helicity state [47]. . . . .	21
1.14	Two left plots show the comparison $W^+$ and $W^-$ events distribution as a function of the boson rapidity for different states of the boson helicity. The right plots shows the same distributions but for muon events as a function of muon pseudorapidity and $W^\pm$ boson helicity state [48]. . . . .	22
2.1	Geographic location of the LHC [49]. . . . .	24
2.2	Acceleration scheme [50]. . . . .	25
2.3	LHC report on the delivered luminosity during RUN 1 and RUN 2. . . . .	26
2.4	The figure shows the CMS detector, its main components, and characteristics. A black human figure on the right corner below is made of the real scale to the detector size [53]. . . . .	28
2.5	The CMS tracker slice view in the r-z plane. Single-sided strip modules are marked with black thin lines, pixel modules are depicted in red, and strip stereo modules are shown as blue thick lines. [55] . . . . .	30
2.6	The left figure illustrates a $PbWO_4$ crystal with a graphical representation of a shower inside [56]. The right figure represents the sketch of the one-fourth of the ECAL in the Y-Z plane. Tilted $PbWO_4$ crystals are shown with the blue lines. The dashed lines show the pseudorapidity coverage of the EB, EE, and ES [57]. . . . .	32
2.7	The figure represents the sketch of the quarter slice of the HCAL in the Y-Z plane. "FEE" corresponds to the locations of the Front End Electronics for HB and HE. The signals from the tower cells shown with the same color are added optically, providing a "longitudinal" segmentation of the HCAL. HB and HE are sliced in 29 azimuthal pieces with $\Delta\phi \approx 20$ degrees each [60]. . . . .	33
2.8	Representation of the 5 modules composing the cold mass inside the cryostat, with the supporting system (vertical, radial and longitudinal tie rods) [54]. . . . .	34



2.9	CMS muon system in the Y-Z plane. The rings of the barrel yoke 0, 1, 2 are shown as "Wheel", the endcap yoke is shown as "Steel". The Drift Tube stations (DTs) are shown with a beige color and labeled as MB ("Muon Barrel"). The color of Cathode Strip Chambers (CSCs) is lime-green; they are labeled as ME ("Muon Endcap"). Resistive Plate Chambers (RPCs) are mounted in both the barrel and endcaps of CMS, they are labeled RB and RE, respectively, and represented with a blue color. The dotted lines show the pseudorapidity coverage of the muon system with equidistant steps of 0.1.[65] . . . . .	35
2.10	The scheme represents the CMS L1 trigger system. The data from the electromagnetic calorimeter (ECAL), hadronic calorimeters (HF and HCAL) are processed first to regional calorimeter trigger (RCT), then to global calorimeter trigger (GCT). Energy deposits (hits) from the muon detector (RPC, CSC, and DT) are processed via a system of a track- and segment-finders or via a pattern comparator, and sent to a global muon trigger (GMT). The final trigger decision is made using a global trigger (GT), which combines the information from the GCT and GMT. The decision is then sent to the tracker (TRK), ECAL, HCAL, or muon systems (MU) using the Trigger, Timing and Control system (TTC). Afterwards, the data from various subsystems is transferred to the acquisition system (DAQ) for offline storage.[68] . . . . .	36
3.1	Schematic view of Particle Flow algorithm as a link between detector signals and reconstructed physics objects. The left side illustrates an event representation in terms of physics objects. The right part represents the CMS detector in the transverse plane; ECAL clusters represent an amount of energy detected from leptons and photons in the hypothetical event. The HCAL cluster shows amount of detected energy from hadrons. Tracks are shown with black lines. The muon that penetrates all layers is shown with the red dotted line [72]. . . . .	40
3.2	Interactions per crossing (pile-up) for 2015-2018 [75] . . . . .	41
3.3	Resolution of missing transverse momenta of $Z^0 \rightarrow e^+e^-$ as a function of reconstructed pp collision vertices. PF (red triangles) corresponds to events reconstructed only with the particle flow algorithm. PUPPI (blue triangles) illustrates PF events corrected with the PUPPI method. [77] . . . . .	42
3.4	$\epsilon_{\text{track}}$ for isolated muons shown as a function of pseudorapidity (left) with $p_T = 1, 10, 100$ GeV and transverse momentum (right) for barrel ( $0.0 <  \eta  < 0.9$ ), transition ( $0.9 <  \eta  < 1.4$ ), and endcap ( $1.4 <  \eta  < 2.5$ ) regions [70]. . . . .	46
3.5	Tag-and-probe $\epsilon_{\text{reco+ID}}$ efficiency for loose ID (left) and tight ID (right) as a function of $\eta$ . Detector regions with a smaller amount of instrumentation cause the dips in efficiency. The statistical uncertainties are tiny [65]. . . . .	47
3.6	Tag-and-probe $\epsilon_{\text{iso}}$ efficiency for tight ID as a function of $p_T$ (left) and $\eta$ (right). The statistical uncertainties are smaller than the symbols [65]. . . . .	47
3.7	The $\epsilon_{\text{trig}}$ efficiency of an isolated muon is presented. The left plot shows $\epsilon_{\text{trig}}$ efficiency as a function of muon $p_T$ while the left corresponds to a function of muon $\eta$ . The statistical uncertainties are smaller than the markers [65]. . . . .	48

4.1	The picture shows the CMS detector in the Y-Z plane. Names of the BRIL subsystems are shown in different colors. The black lines indicate positions of the subsystems. The BPTX is at a distance of 175 m from the IP and is not shown on the scheme [87]. . . . .	51
4.2	C-shape equipped with six sensors. Each sensor has two channels [91]. . . . .	52
4.3	The scheme represents arrival time before and after the bunch crossing. The thick black line is the beam-line, circles with gray parallelograms are the BCM1F rings with schematic sensors. The first row shows proton bunches, and their halos passing the BCM1F right before the collision, which corresponds to a timing -6 ns. The second row shows the collision at the interaction point (IP). The third row is dedicated to the outgoing bunches after the collision with some collision products passing the BCM1F sensors, the timing of this stage corresponds to 6 ns after the collision. Overall timing of each collision is 12 ns [87]. . . . .	53
4.4	Hit rates as a function of time for one bunch crossing collected during 2.5 minutes [92]. . . . .	54
4.5	Van der Meer scan as performed during the fill 4266. The left plot corresponds to X-plane, while the right is Y-plane. The plots show the measured values for each displacement as a function of separation distance. Black dots correspond to measured rate values at each step. Red and green lines correspond to a Gaussian fit, constant term is shown with a blue line. The resulting fit is shown with a black line [93]. . . . .	54
4.6	The graph of an electron scan of typical poly-crystalline diamond films grown on Si by CVD, using 0.5% methane in hydrogen.[97] . . . . .	55
4.7	Geometrical characteristics of a sensor and its one pad metallization (Picture 4.7 a) and two pad metallization (Picture 4.7 b) [99]. . . . .	56
4.8	Schematic representation of signal transferring from sensor channels to AOH. The left plot shows a sensor with two channels and two pads that are connected to the ASIC. Blue arrows correspond to ASIC output, which is connected to AOHs. The right plot shows two AOHs with incoming signals marked as blue arrows. Each AOH processes only right or left ASIC output signal from three different sensors. Red arrows correspond to outgoing optical fibers marked as left, middle, and right. . . . .	57
4.9	Electric schemes for IV and CT measurements. Generally, the schemes for the IV and CT are the same, except that the CT has a source of $\beta^-$ decay $^{90}\text{Sr}$ . HV is a power supply; the gray and blue rectangles are a diamond sensor, metallization pads are shown with a yellow color. A is ampere meter, and earth ground is shown as a dashed triangle. . . . .	59
4.10	A diamond sensor in a plastic frame that is used in IV and CT measurements. Schematic pictures show the binding of a diamond sensor with the metal strips [108]. . . . .	59
4.11	Selected results of IV measurement for sCVD sensor "Batch 390" (plot 4.11a), and pCVD sensors 07B10431 (plot 4.11b), 07B10433 (plot 4.11c). The voltage is applied in order 0 - +1000 - 0 - -1000 - 0 V. . . . .	60
4.12	CT measurements of 07B10436 (plot A) and 07B10422 (plot B) pCVD sensors. . . . .	61
4.13	CT measurements of "Batch 390" (Fig. A) and "Number 5" (Fig. B) sCVD sensors. . . . .	62

4.14	A diamond sensor is in the opened "beta-setup" box before the measurement. The needle is on top of the diamond sensor, which is mounted in the center of the high voltage golden plate. The black hole in an iron pyramid is a dry-air source (5% oxygen + 95% dry nitrogen). During the measurement, this scheme is covered with a black wooden box. Inside on top, the box has two mounted layers of scintillation detectors. . . . .	63
4.15	The ionization signal is registered with the needle and amplified with the pre-amplifier forming a Diamond Signal. FIFO is a Fan-In Fan-Out, Scope is an oscilloscope. SC1 and SC2 - two scintillators, LTD - Low-Threshold Discriminator, COIN - Coincidence Unit. The red line is a signal input in ADC; the blue line is a duplicated signal with a delay (blue circle between the FIFO and the Scope). The magenta line is a gate, formed from two scintillator signals. The Scope shows the opened gate right in the time interval of the signal, excluding the pedestal time interval [99]. . . . .	64
4.16	Two histograms of 10000 ADC counts of pCVD sensor 07B10417 measured at 500 V. The pedestal ADC counts with the Gaussian distribution shape are shown in the upper histogram. The lower histogram shows signal ADC counts with the Landau distribution shape. . . . .	64
4.17	Calibration line of ADC channels as a function of charge. The Calib value obtained from this fit is $k=47$ . . . . .	65
4.18	CCD measurements as a function of time of 07B10415 (plot 4.18a) and 07B10434 (plot 4.18b) pCVD sensors. . . . .	66
4.19	CCD measurements as a function of time of "Number 5" (plot 4.19a) and "Batch 390" (plot 4.19b) sCVD sensors. Both sensors show high charge collection distance and stability over time for all bias voltages. . . . .	66
4.20	CCD measurements as a function of applied bias voltage of pCVD sensor 07B10415 (plot 4.20a) and sCVD sensor "Batch 390" (plot 4.20b). Both sensors show a stable logarithmic rise. . . . .	67
4.21	Recorded orbit from the LHC Fill 4364 with 852 colliding bunches [87]. The baseline is set on 140 ADC counts, signals are shown as dips in a baseline position. Red ellipse shows a TP, green rectangle corresponds to the "pilot" train, and each magenta circle is a train of bunches. . . .	68
4.22	Arrival time distribution [87]. . . . .	69
4.23	Baseline distributions for channel 1 taken from one of the first fills. Plot 4.23a corresponds to the baseline position distribution, while plot 4.23b shows distribution of its standard deviation. . . . .	69
4.24	The zoomed test pulse signal is a dip in the baseline, defined between 6070 and 6080 ns. The red dashed line corresponds to the baseline threshold, which is 3 ADC counts. The blue line is a signal value, which corresponds to 20 ADC or 80 mV. The green line shows which part of the signal is taken as the arrival time. . . . .	70
4.25	Test pulse distributions for channel 2 taken from one of the first fills, obtained after processing 401 orbits. Plot 4.25a shows distribution of TP signal amplitude. Plot 4.25b shows distribution of TP arrival time. . . .	70
4.26	Signal amplitude spectra taken from Fill 5722 of pCVD sensor 4.26a in log scale, and sCVD sensor 4.26b in linear. . . . .	71
4.27	Signal amplitude spectra with reference histograms (plot A), base line position histogram with references (plot B). . . . .	72

4.28	Signal amplitude distribution from channel 17 (pCVD 07B10512), collected in Fill 6417. . . . .	73
4.29	Test pulse spectrum in Web monitor tool. . . . .	73
4.30	Signal amplitude spectra of sCVD "Number 5", channel 11. Plot A shows distribution obtained from fill 5730, 30.05.2017. Plot B is taken from fill 5984, 24.07.2017. . . . .	74
4.31	Signal amplitude and TP MPV values as a function of a fill number for sCVD "Number 5" (plot A) and sCVD "Batch390" (plot B). Each plot shows results from both channels, in different colors. The solid line corresponds to MPV values taken from signal amplitude distribution, while dashed lines correspond to MPV taken from test pulse spectra. . . . .	75
4.32	Signal amplitude spectra of pCVD sensor 07B10436, channel 5. Plot A shows distribution obtained from fill 5737, 03.06.2017. Plot B is taken from fill 7090, 26.08.2018. . . . .	76
4.33	Signal amplitude spectrum in Web monitor tool. . . . .	77
4.34	Test pulse spectrum degradation. . . . .	77
4.35	Baseline position distortions. Plot A shows the baseline position shift, and plot B is a baseline standard deviation spread. Plot C is a zoomed region of an orbit record around TP (shown in red rectangle). This plot shows a noisy baseline. . . . .	78
5.1	Muon tracking, identification and isolation efficiency in different $\eta$ and $p_T$ bins. Passing and failing probes are represented in the left and right columns, respectively. Results for $\mu^-$ and $\mu^+$ are shown in upper and lower rows, respectively. . . . .	87
5.2	Muon standalone efficiency in different $\eta$ and $p_T$ bins. Passing and failing probes are represented in the left and right columns, respectively. Results for $\mu^-$ and $\mu^+$ are shown in upper and lower rows, respectively. . . . .	88
5.3	Muon trigger efficiency in different $\eta$ and $p_T$ bins. Passing and failing probes are represented in the left and right columns, respectively. Results for $\mu^-$ and $\mu^+$ are shown in upper and lower rows, respectively. . . . .	89
5.4	Muon efficiency as a function of pseudorapidity. . . . .	90
5.5	Schematic representation of $W^\pm$ and $Z^0$ boson event decomposition in transverse plane. . . . .	91
5.6	The figures represent an example of the fit of parallel (left figure) and perpendicular (right) recoil components in a random rapidity and boson $p_T$ bins. The legend contains the mean and width values of each Gaussian function. . . . .	92
5.7	The response curve of the recoil scale as a function of boson $p_T$ . The left figure represents the fit of the mean values of the first Gaussian for the parallel recoil component in $W^- \rightarrow \mu^- \bar{\nu}_\mu$ in the first $\eta$ bin. The right figure represents the fit of the mean values of the second Gaussian for the parallel recoil component in $W^+ \rightarrow \mu^+ \nu_\mu$ in the tenth $\eta$ bin. . . . .	93
5.8	$E_T^{\text{miss}}$ fit in the control region defined by inverting isolation criteria. The fit is performed separately for $W^+$ and $W^-$ samples in different pseudorapidity regions. . . . .	95

5.9	Simultaneous fit of the $E_T^{\text{miss}}$ distribution in signal region in the $0.4 <  \eta  < 0.6$ pseudorapidity range (upper row), $0.8 <  \eta  < 1.0$ (middle row) and $1.85 <  \eta  < 2.1$ (lower row). The fit is performed separately for $W^+$ (left column) and $W^-$ (right column). . . . .	97
5.10	Acceptance values as the function of absolute pseudorapidity for $W^+ \rightarrow \mu^+ \nu_\mu$ and $W^- \rightarrow \mu^- \bar{\nu}_\mu$ channels. The systematic uncertainties are smaller than the size of the markers used in the plot. . . . .	100
5.11	Differential cross-section values as a function of pseudorapidity represented for $W^+ \rightarrow \mu^+ \nu_\mu$ and $W^- \rightarrow \mu^- \bar{\nu}_\mu$ channels. The results are compared to predictions at NNLO calculated in MCFM-8.3 using CT14 NNLO, MSTW8 NNLO, and MMHT NNLO PDF sets. . . . .	104
5.12	$W^\pm$ boson charge asymmetry distribution in the full pseudorapidity range as a function of pseudorapidity. The results are compared to predictions at NNLO calculated in MCFM-8.3 using different PDF sets. . . . .	105
5.13	Systematic uncertainty figures represents the contribution of each source of systematic as a deviation from the cross-section results. Results are shown for $W^+ \rightarrow \mu^+ \nu_\mu$ and $W^- \rightarrow \mu^- \bar{\nu}_\mu$ channels on the left and right figures respectively. Measured values are presented with statistical and systematic uncertainties. . . . .	108
6.1	Distributions of $x \cdot d_v$ valence (right) and $x \cdot u_v$ valence (left) quarks as functions of $x$ at the scale $Q^2 = m_W^2 \text{ GeV}^2$ . The results of the fit to HERA data only (hatched band), and to the HERA data and muon asymmetry measurements (light shaded band) are compared. In the bottom panels, the distributions are normalized to 1 for a direct comparison of the uncertainties. The change of the PDFs with respect to the HERA-only fit is represented by a solid line. [138]. . . . .	112
6.2	Schematic representation of the xFitter constituents [1]. . . . .	113
6.3	The $W^\pm$ boson charge asymmetry values, extracted at $\sqrt{s} = 7 \text{ TeV}$ (left) [145] and $\sqrt{s} = 8 \text{ TeV}$ (right) [150], are shown in the first row. Asymmetries are shown as comparison with theory predictions calculated at NLO and NNLO, using different PDF sets. The differential cross-sections of the $W^\pm + c$ production as a function of absolute values of pseudorapidity, measured at $\sqrt{s} = 7 \text{ TeV}$ (left) [146], and $\sqrt{s} = 13 \text{ TeV}$ (right) [147], are shown in the second row. The measurements are compared with theory predictions calculated at NLO, using different PDF sets. . . . .	115
6.4	The comparison of theoretical prediction to the $W^\pm$ boson charge asymmetry measurement. . . . .	117
6.5	Obtained parton density functions for $x \cdot u_v$ and $x \cdot d_v$ , using HERA data and new $W^\pm$ boson charge asymmetry results. The first row represents results obtained at the starting scale $Q^2 = 1.9 \text{ GeV}^2$ , while the second row shows results at the scale $Q^2 = m_{W^\pm}^2$ . The left plots correspond to $x \cdot u_v$ distributions, while right are $x \cdot d_v$ . . . . .	118
6.6	Distributions of relative uncertainty $\delta x_{q_v}/x_{q_v}$ of up (left) and down (right) quarks at the scale of $10 \text{ GeV}^2$ . The second row presents the same distributions at the scale of $6464 \text{ GeV}^2$ . . . . .	119

6.7	Comparison of theory predictions to data used in the fit. The first row shows results for the $W^\pm$ boson charge asymmetry measured at $\sqrt{s} = 7$ TeV (left) and $\sqrt{s} = 8$ TeV (right). The second row shows comparison for $W^\pm$ +charm process, measured at $\sqrt{s} = 7$ TeV (left) and $\sqrt{s} = 13$ TeV (right). $W^\pm$ boson charge asymmetry measured in this thesis is shown in the last row. . . . .	121
6.8	Obtained parton density functions for $x \cdot u_v$ and $x \cdot d_v$ (first row), $x \cdot \bar{u}$ and $x \cdot \bar{d}$ (second), $x \cdot s$ and $x \cdot (s + \bar{s})/(\bar{u} + \bar{d})$ (third), $x \cdot g$ and $x \cdot \Sigma$ (fourth) at starting scale $Q^2 = 1.9 \text{ GeV}^2$ . . . . .	122
6.9	Obtained parton density functions for $x \cdot u_v$ and $x \cdot d_v$ (first row), $x \cdot \bar{u}$ and $x \cdot \bar{d}$ (second), $x \cdot s$ and $x \cdot (s + \bar{s})/(\bar{u} + \bar{d})$ (third), $x \cdot g$ and $x \cdot \Sigma$ (fourth) at starting scale $Q^2 = m_{W^\pm}^2 \text{ GeV}^2$ . . . . .	123
6.10	Distributions of relative uncertainty $\delta x_{q_v}/x_{q_v}$ of up (1st row, left plot) and down (1st row, right plot) quarks. The second row presents distributions relative uncertainty $\delta x_s/x_s$ for strange quark (left plot) and $\delta x(s + \bar{s})/(\bar{u} + \bar{d})/x(s + \bar{s})/(\bar{u} + \bar{d})$ (right plot). All plots show results at the scale of $1.9 \text{ GeV}^2$ . . . . .	125
6.11	Distributions of relative uncertainty $\delta x_{q_v}/x_{q_v}$ of up (1st row, left plot) and down (1st row, right plot) quarks. The second row presents distributions for gluon (left plot) and sea quarks (right plot). Anti-up, $\delta x_{\bar{u}}/x_{\bar{u}}$ and anti-down, $\delta x_{\bar{d}}/x_{\bar{d}}$ , quarks are shown on the left and right plots, respectively. All plots show results with and without 2015 $W^\pm$ boson charge asymmetry at the scale of $m_{W^\pm}^2$ . . . . .	126
A.1	CT at 200V 400V (negative attenuated) . . . . .	157
A.2	CT at 500V 750V 1000V (positive attenuated) . . . . .	157
A.3	CCD at 500V . . . . .	157
A.4	CCD at 500V 750V 1000V . . . . .	158
A.5	CCD vs HV . . . . .	158
A.6	CT at 200V 400V 600V (negative attenuated) . . . . .	158
A.7	CT at 500V 750V (negative attenuated) . . . . .	158
A.8	CT at 1000V (negative attenuated) . . . . .	158
A.9	CT at 200V 400V 600V (negative attenuated) . . . . .	159
A.10	CT at 500V 750V 1000V(negative attenuated) . . . . .	159
A.11	CCD at 500V 750V 1000V . . . . .	159
A.12	CT at 200V 500V 750V 1000V (negative attenuated) . . . . .	160
A.13	CCD at 500V 750V . . . . .	160
A.14	CT at 200V 500V 750V 1000V (negative attenuated) . . . . .	161
A.15	CT at 500V 750V 1000V(positive attenuated) . . . . .	161
A.16	CCD at 500V 750V 1000V . . . . .	161
A.17	CT at 200V 400V 600V 750V 1000V (negative attenuated) . . . . .	162
A.18	CCD at 500V 750V . . . . .	162
A.19	CT at 200V 500V 750V 1000V (negative attenuated) . . . . .	163
A.20	IV . . . . .	163
A.21	CCD at 500V 750V 1000V . . . . .	163
A.22	CCD 500V 4h check . . . . .	163
A.23	CT at 200V 500V 750V 1000V . . . . .	164
A.24	IV . . . . .	164
A.25	CCD at 500V . . . . .	164

A.26	CT at 200V 500V 750V 1000V	165
A.27	IV	165
A.28	CCD at 500V	165
A.29	CT at 200V 500V 750V 1000V	166
A.30	IV	166
A.31	CCD at 500V 750V	166
A.32	CT at 200V 500V 750V 1000V	167
A.33	IV	167
A.34	CCD at 500V 750V 1000V	167
A.35	CT at 200V 500V 750V 1000V	168
A.36	IV	168
A.37	CCD at 500V 750V 1000V	168
A.38	CCD vs HV Hysteresis	168
A.39	CT at 200V 500V 750V 1000V	169
A.40	IV	169
A.41	CCD at 500V 750V	169
A.42	CCD vs HV	169
A.43	CT at 200V 500V 750V 1000V	170
A.44	IV	170
A.45	CCD at 500V 750V 1000V	170
A.46	remeasured HV at 500V	170
A.47	CCD vs HV	170
A.48	CT at 200V 500V 750V 1000V	171
A.49	IV	171
A.50	CCD at 500V 750V 1000V	171
A.51	CCD vs HV	171
A.52	CT at 200V 500V 750V 1000V	172
A.53	IV	172
A.54	CCD at 500V 750V 1000V	172
A.55	CCD vs HV	172
A.56	CT at 200V 500V 750V 1000V	173
A.57	IV	173
A.58	CCD at 500V 750V 1000V	173
A.59	CCD vs HV	173
B.1	Results of simultaneous fit of the $E_T^{\text{miss}}$ distribution in signal (left) and control (right) regions for $W^+ \rightarrow \mu^+ \nu_\mu$ . The upper row corresponds to $0.0 <  \eta  < 0.2$ region and the last row to $0.2 <  \eta  < 0.4$ .	175
B.2	Results of simultaneous fit of the $E_T^{\text{miss}}$ distribution in signal (left) and control (right) regions for $W^+ \rightarrow \mu^+ \nu_\mu$ . The upper row corresponds to $0.4 <  \eta  < 0.6$ region, middle row to $0.6 <  \eta  < 0.8$ , and the last row to $0.8 <  \eta  < 1.0$ .	176
B.3	Results of simultaneous fit of the $E_T^{\text{miss}}$ distribution in signal (left) and control (right) regions for $W^+ \rightarrow \mu^+ \nu_\mu$ . The upper row corresponds to $1.0 <  \eta  < 1.2$ region, middle row to $1.2 <  \eta  < 1.4$ , and the last row to $1.4 <  \eta  < 1.6$ .	177

B.4	Results of simultaneous fit of the $E_T^{\text{miss}}$ distribution in signal (left) and control (right) regions for $W^+ \rightarrow \mu^+ \nu_\mu$ . The upper row corresponds to $1.6 <  \eta  < 1.85$ region, middle row to $1.85 <  \eta  < 2.1$ , and the last row to $2.1 <  \eta  < 2.4$ . . . . .	178
B.5	Results of simultaneous fit of the $E_T^{\text{miss}}$ distribution in signal (left) and control (right) regions for $W^- \rightarrow \mu^- \bar{\nu}_\mu$ . The upper row corresponds to $0.0 <  \eta  < 0.2$ region, middle row to $0.2 <  \eta  < 0.4$ , and the last row to $0.4 <  \eta  < 0.6$ . . . . .	179
B.6	Results of simultaneous fit of the $E_T^{\text{miss}}$ distribution in signal (left) and control (right) regions for $W^- \rightarrow \mu^- \bar{\nu}_\mu$ . The upper row corresponds to $0.6 <  \eta  < 0.8$ region, middle row to $0.8 <  \eta  < 1.0$ , and the last row to $1.0 <  \eta  < 1.2$ . . . . .	180
B.7	Results of simultaneous fit of the $E_T^{\text{miss}}$ distribution in signal (left) and control (right) regions for $W^- \rightarrow \mu^- \bar{\nu}_\mu$ . The upper row corresponds to $1.2 <  \eta  < 1.4$ region, middle row to $1.4 <  \eta  < 1.6$ , and the last row to $1.6 <  \eta  < 1.85$ . . . . .	181
B.8	Results of simultaneous fit of the $E_T^{\text{miss}}$ distribution in signal (left) and control (right) regions for $W^- \rightarrow \mu^- \bar{\nu}_\mu$ . The upper row corresponds to $1.85 <  \eta  < 2.1$ region, the last row to $2.1 <  \eta  < 2.4$ . . . . .	182



# List of Tables

3.1	The isolated single-muon trigger efficiency using 2015 data. The first row represents the Level-1 efficiency ( $p_T$ threshold 16 GeV) with respect to (w.r.t.) offline muons. The second two rows show the HLT efficiency ( $p_T < 20$ GeV) w.r.t. to offline muons geometrically linked to L1 candidates. The last rows correspond to online isolation efficiency w.r.t. offline muons firing HLT. The last column shows the scale factor between MC and data. Represented uncertainties are statistical [65]. . . . .	48
4.1	Positions of sensors on C-shapes . . . . .	67
5.1	Summary of data set names and trigger paths . . . . .	82
5.2	Pseudorapidity binning . . . . .	84
5.3	Muon selection requirements . . . . .	84
5.4	Transverse momentum binning . . . . .	86
5.5	Muon efficiency . . . . .	90
5.6	Electroweak background Monte Carlo sets . . . . .	94
5.7	Extracted $W^+ \rightarrow \mu^+ \nu_\mu$ yields in signal region . . . . .	96
5.8	Extracted $W^+ \rightarrow \mu^+ \nu_\mu$ yields in control region . . . . .	98
5.9	Extracted $W^- \rightarrow \mu^- \bar{\nu}_\mu$ yields in signal region . . . . .	98
5.10	Extracted $W^- \rightarrow \mu^- \bar{\nu}_\mu$ yields in control region . . . . .	99
5.11	Acceptance . . . . .	100
5.12	Luminosity uncertainty table . . . . .	103
5.13	Table of differential cross-section systematic uncertainties . . . . .	106
5.14	Table of the asymmetry systematic uncertainties . . . . .	107
5.15	Differential cross-sections of $W^+$ boson production . . . . .	109
5.16	Differential cross-sections of $W^-$ boson production . . . . .	109
5.17	Measured asymmetry results . . . . .	110
6.1	Comparison of the fit quality of the first QCD analysis . . . . .	118
6.2	Comparison of the fit quality . . . . .	124

- C.1 Table of MCFM-8.0 processes available in the APPLgrid. The MCFM number of the process is given in the first column. Process channel, as well as numbering of outgoing 4-vectors of particles ( $p_3 - p_8$ ), is shown in the second column. A maximum available computational perturbation order is given in the third column. [rad.in.dk] - processes includes only the corrections in the semileptonic decay of the top quark. (uncorr) - no spin correlations in the decay of the top quarks. [rad.in.top.dk] - include only the radiative corrections in the decay of the top quark without including the radiative corrections in the hadronic decay of the W-boson. [rad.in.W.dk] - includes only the radiative corrections in the hadronic decay of the W-boson coming from the anti-top(top). NLO+F signifies that the calculation can be performed at NLO including the effects of experimental isolation and photon fragmentation. . . . . 189

# Bibliography

- [1] V. Bertone et al. *xFitter 2.0.0: An Open Source QCD Fit Framework*. 2017. arXiv: [1709.01151](https://arxiv.org/abs/1709.01151) [hep-ph].
- [2] *The H1 experiment main page*. URL: <https://h1.desy.de/>.
- [3] *ZEUS experiment main page*. URL: <https://www-zeus.desy.de/>.
- [4] H1 and ZEUS Collaborations. “Combination of Measurements of Inclusive Deep Inelastic  $e^\pm p$  Scattering Cross Sections and QCD Analysis of HERA Data”. In: (2015). arXiv: [1506.06042](https://arxiv.org/abs/1506.06042) [hep-ex].
- [5] Emmy Noether. “Invariante Variationsprobleme”. In: *Machrichten von der Koniglichen Gesellschaft der Wissenschaften* (July 1918), pp. 17, 235.
- [6] Alberto Zannoni. *On the Quantization of the Monoatomic Ideal Gas*. 1999. arXiv: [cond-mat/9912229](https://arxiv.org/abs/cond-mat/9912229) [cond-mat.stat-mech].
- [7] P. A. M. Dirac. “On the Theory of Quantum Mechanics”. In: *Proceedings of the Royal Society of London Series A* 112.762 (Oct. 1926), pp. 661–677. DOI: [10.1098/rspa.1926.0133](https://doi.org/10.1098/rspa.1926.0133).
- [8] Bose. “Plancks Gesetz und Lichtquantenhypothese”. In: *Zeitschrift fur Physik* 26.1 (Dec. 1924), pp. 178–181. DOI: [10.1007/BF01327326](https://doi.org/10.1007/BF01327326).
- [9] Cush MissMJ. *Standard Model of Elementary Particles*. 2019. URL: [https://commons.wikimedia.org/wiki/File:Standard\\_Model\\_of\\_Elementary\\_Particles.svg](https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg).
- [10] Peter W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. In: *Phys. Rev. Lett.* 13 (16 Oct. 1964), pp. 508–509. DOI: [10.1103/PhysRevLett.13.508](https://doi.org/10.1103/PhysRevLett.13.508). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.508>.
- [11] F. Englert and R. Brout. “Broken Symmetry and the Mass of Gauge Vector Mesons”. In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323. DOI: [10.1103/PhysRevLett.13.321](https://doi.org/10.1103/PhysRevLett.13.321). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.321>.
- [12] G. Aad et al. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Physics Letters B* 716.1 (Sept. 2012), pp. 1–29. ISSN: 0370-2693. DOI: [10.1016/j.physletb.2012.08.020](https://doi.org/10.1016/j.physletb.2012.08.020). URL: <http://dx.doi.org/10.1016/j.physletb.2012.08.020>.
- [13] S. Chatrchyan et al. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: *Physics Letters B* 716.1 (Sept. 2012), pp. 30–61. ISSN: 0370-2693. DOI: [10.1016/j.physletb.2012.08.021](https://doi.org/10.1016/j.physletb.2012.08.021). URL: <http://dx.doi.org/10.1016/j.physletb.2012.08.021>.

- [14] M. Tanabashi et al. “Review of Particle Physics”. In: *Phys. Rev. D* 98 (3 Aug. 2018), p. 030001. DOI: [10.1103/PhysRevD.98.030001](https://doi.org/10.1103/PhysRevD.98.030001). URL: <https://link.aps.org/doi/10.1103/PhysRevD.98.030001>.
- [15] V. Khachatryan et al. “Measurement and QCD analysis of double-differential inclusive jet cross sections in pp collisions at  $s=8\sqrt{s}=8$  TeV and cross section ratios to 2.76 and 7 TeV”. In: *Journal of High Energy Physics* 2017.3 (Mar. 2017). ISSN: 1029-8479. DOI: [10.1007/jhep03\(2017\)156](https://doi.org/10.1007/jhep03(2017)156). URL: [http://dx.doi.org/10.1007/JHEP03\(2017\)156](http://dx.doi.org/10.1007/JHEP03(2017)156).
- [16] *Fine structure constant*. 2019. URL: <https://physics.nist.gov/cgi-bin/cuu/Value?alph>.
- [17] Rym Bouchendira et al. “New Determination of the Fine Structure Constant and Test of the Quantum Electrodynamics”. In: *Physical Review Letters* 106.8 (Feb. 2011). ISSN: 1079-7114. DOI: [10.1103/physrevlett.106.080801](https://doi.org/10.1103/physrevlett.106.080801). URL: <http://dx.doi.org/10.1103/PhysRevLett.106.080801>.
- [18] C. S. Wu et al. “Experimental Test of Parity Conservation in Beta Decay”. In: *Physical Review* 105.4 (Feb. 1957), pp. 1413–1415. DOI: [10.1103/PhysRev.105.1413](https://doi.org/10.1103/PhysRev.105.1413).
- [19] Nicola Cabibbo. “Unitary Symmetry and Leptonic Decays”. In: 10.12 (June 1963), pp. 531–533. DOI: [10.1103/PhysRevLett.10.531](https://doi.org/10.1103/PhysRevLett.10.531).
- [20] KJBurns TimothyRias. *Quark weak interactions*. 2009. URL: [https://commons.wikimedia.org/wiki/File:Quark\\_weak\\_interactions.svg](https://commons.wikimedia.org/wiki/File:Quark_weak_interactions.svg).
- [21] Julian Schwinger. “The Theory of Quantized Fields. I”. In: *Phys. Rev.* 82 (6 June 1951), pp. 914–927. DOI: [10.1103/PhysRev.82.914](https://doi.org/10.1103/PhysRev.82.914). URL: <https://link.aps.org/doi/10.1103/PhysRev.82.914>.
- [22] Heinrich Wahl. “First observation and measurement of direct CP violation by the NA31 and NA48 experiments at CERN”. In: *Subnucl. Ser.* 37 (2001), pp. 378–386. DOI: [10.1142/9789812811585\\_0013](https://doi.org/10.1142/9789812811585_0013).
- [23] B. Aubert et al. “Measurement of CP-Violating Asymmetries in  $B^0$  Decays to CP Eigenstates”. In: *Physical Review Letters* 86.12 (Mar. 2001), pp. 2515–2522. ISSN: 1079-7114. DOI: [10.1103/physrevlett.86.2515](https://doi.org/10.1103/physrevlett.86.2515). URL: <http://dx.doi.org/10.1103/PhysRevLett.86.2515>.
- [24] K. Abe et al. “Observation of Large CP Violation in the Neutral  $B$  Meson System”. In: *Physical Review Letters* 87.9 (Aug. 2001). ISSN: 1079-7114. DOI: [10.1103/physrevlett.87.091802](https://doi.org/10.1103/physrevlett.87.091802). URL: <http://dx.doi.org/10.1103/PhysRevLett.87.091802>.
- [25] R. Aaij et al. “First Observation of CP Violation in the Decays of  $B_s^0$  Mesons”. In: *Physical Review Letters* 110.22 (May 2013). ISSN: 1079-7114. DOI: [10.1103/physrevlett.110.221601](https://doi.org/10.1103/physrevlett.110.221601). URL: <http://dx.doi.org/10.1103/PhysRevLett.110.221601>.
- [26] R. Aaij et al. “Observation of CP Violation in Charm Decays”. In: *Phys. Rev. Lett.* 122 (21 May 2019), p. 211803. DOI: [10.1103/PhysRevLett.122.211803](https://doi.org/10.1103/PhysRevLett.122.211803). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.122.211803>.

- [27] Steven Weinberg. “A Model of Leptons”. In: *Phys. Rev. Lett.* 19 (21 Nov. 1967), pp. 1264–1266. DOI: [10.1103/PhysRevLett.19.1264](https://doi.org/10.1103/PhysRevLett.19.1264). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.19.1264>.
- [28] E. E. Chambers and R. Hofstadter. “THE STRUCTURE OF THE PROTON”. In: (Apr. 1956). DOI: [10.2172/4345570](https://doi.org/10.2172/4345570).
- [29] N. Bezginov et al. “A measurement of the atomic hydrogen Lamb shift and the proton charge radius”. In: *Science* 365.6457 (Sept. 2019), pp. 1007–1012. DOI: [10.1126/science.aau7807](https://doi.org/10.1126/science.aau7807).
- [30] W. Xiong et al. “A small proton charge radius from an electron–proton scattering experiment”. In: *Nature* 575.7781 (2019), pp. 147–150. DOI: [10.1038/s41586-019-1721-2](https://doi.org/10.1038/s41586-019-1721-2).
- [31] J. Ashman et al. “A measurement of the spin asymmetry and determination of the structure function  $g_1$  in deep inelastic muon-proton scattering”. In: *Physics Letters B* 206.2 (1988), pp. 364–370. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(88\)91523-7](https://doi.org/10.1016/0370-2693(88)91523-7). URL: <http://www.sciencedirect.com/science/article/pii/0370269388915237>.
- [32] J I Friedman and H W Kendall. “Deep Inelastic Electron Scattering”. In: *Annual Review of Nuclear Science* 22.1 (1972), pp. 203–254. DOI: [10.1146/annurev.ns.22.120172.001223](https://doi.org/10.1146/annurev.ns.22.120172.001223). eprint: <https://doi.org/10.1146/annurev.ns.22.120172.001223>. URL: <https://doi.org/10.1146/annurev.ns.22.120172.001223>.
- [33] J. D. Bjorken. “Asymptotic Sum Rules at Infinite Momentum”. In: *Phys. Rev.* 179 (5 Mar. 1969), pp. 1547–1553. DOI: [10.1103/PhysRev.179.1547](https://doi.org/10.1103/PhysRev.179.1547). URL: <https://link.aps.org/doi/10.1103/PhysRev.179.1547>.
- [34] F Close. “The quark parton model”. In: *Reports on Progress in Physics* 42.8 (Aug. 1979), pp. 1285–1335. DOI: [10.1088/0034-4885/42/8/001](https://doi.org/10.1088/0034-4885/42/8/001). URL: <https://doi.org/10.1088%2F0034-4885%2F42%2F8%2F001>.
- [35] T. Eichten et al. “Measurement of the neutrino-nucleon and antineutrino-nucleon total cross sections”. In: *Physics Letters B* 46.2 (1973), pp. 274–280. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(73\)90702-8](https://doi.org/10.1016/0370-2693(73)90702-8). URL: <http://www.sciencedirect.com/science/article/pii/0370269373907028>.
- [36] T. Eichten et al. “Measurement of the neutrino-nucleon and antineutrino-nucleon total cross sections”. In: *Physics Letters B* 46.2 (1973), pp. 274–280. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(73\)90702-8](https://doi.org/10.1016/0370-2693(73)90702-8). URL: <http://www.sciencedirect.com/science/article/pii/0370269373907028>.
- [37] Robin Devenish and Amanda Cooper-Sarkar. *Deep Inelastic Scattering*. Oxford Scholarship Online, 2003.
- [38] John C. Collins, Davison E. Soper, and George Sterman. “Factorization for short distance hadron-hadron scattering”. In: *Nuclear Physics B* 261 (1985), pp. 104–142. ISSN: 0550-3213. DOI: [https://doi.org/10.1016/0550-3213\(85\)90565-6](https://doi.org/10.1016/0550-3213(85)90565-6). URL: <http://www.sciencedirect.com/science/article/pii/0550321385905656>.

- [39] John C. Collins, Davison E. Soper, and George Sterman. “Soft gluons and factorization”. In: *Nuclear Physics B* 308.4 (1988), pp. 833–856. ISSN: 0550-3213. DOI: [https://doi.org/10.1016/0550-3213\(88\)90130-7](https://doi.org/10.1016/0550-3213(88)90130-7). URL: <http://www.sciencedirect.com/science/article/pii/0550321388901307>.
- [40] V. N. Gribov and L. N. Lipatov. “Deep inelastic e p scattering in perturbation theory”. In: *Sov. J. Nucl. Phys.* 15 (1972). [*Yad. Fiz.*15,781(1972)], pp. 438–450.
- [41] G. Altarelli and G. Parisi. “Asymptotic freedom in parton language”. In: *Nuclear Physics B* 126.2 (1977), pp. 298–318. ISSN: 0550-3213. DOI: [https://doi.org/10.1016/0550-3213\(77\)90384-4](https://doi.org/10.1016/0550-3213(77)90384-4). URL: <http://www.sciencedirect.com/science/article/pii/0550321377903844>.
- [42] CMS Collaboration. *Measurement of the lepton charge asymmetry in inclusive W production in pp collisions at  $\sqrt{s} = 7\text{TeV}$* . 2012. URL: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSMP12021>.
- [43] A.D. Martin et al. “Parton distributions and the LHC: W and Z production”. In: *The European Physical Journal C* 14.1 (May 2000), pp. 133–145. ISSN: 1434-6052. DOI: [10.1007/s100520000324](https://doi.org/10.1007/s100520000324). URL: <http://dx.doi.org/10.1007/s100520000324>.
- [44] *Measurement of inclusive W and Z boson production cross sections in pp collisions at  $\sqrt{s}=13\text{ TeV}$* . Tech. rep. CMS-PAS-SMP-15-004. Geneva: CERN, 2015. URL: <https://cds.cern.ch/record/2093537>.
- [45] V. Barger and R. J. N. Phillips. *Collider Physics 1993*. 1993. arXiv: [hep-ph/9309250](https://arxiv.org/abs/hep-ph/9309250) [[hep-ph](https://arxiv.org/abs/hep-ph/9309250)].
- [46] Vernon D. Barger and R. J. N. Phillips. *COLLIDER PHYSICS*. 1987. ISBN: 9780201149456.
- [47] Prof. Mark Thomson. *Lecture: The Weak Interaction*. 2011.
- [48] CMS Marco Cipriani. “Measurement of the helicity of the W boson with the CMS experiment”. In: *Inspirehep* (Nov. 2019). URL: <http://inspirehep.net/record/1763934>.
- [49] Wikipedia. *Large Hadron Collider*. 2005. URL: [https://en.wikipedia.org/wiki/Large\\_Hadron\\_Collider](https://en.wikipedia.org/wiki/Large_Hadron_Collider).
- [50] Dukwon Forthommel. *File:Cern-accelerator-complex.svg*. 2016. URL: <https://commons.wikimedia.org/wiki/File:Cern-accelerator-complex.svg>.
- [51] Lyndon Evans and Philip Bryant. “LHC Machine”. In: *JINST* 3 (2008), S08001. DOI: [10.1088/1748-0221/3/08/S08001](https://doi.org/10.1088/1748-0221/3/08/S08001).
- [52] LHC. “The final days of Run 2”. In: (Sept. 2018). URL: <https://home.cern/news/news/accelerators/lhc-report-final-days-run-2>.
- [53] Thomas McCauley Tai Sakuma. “Detector and Event Visualization with SketchUp at the CMS Experiment”. In: (2014). URL: <https://iopscience.iop.org/article/10.1088/1742-6596/513/2/022032>.
- [54] CMS Collaboration. “The CMS experiment at the CERN LHC”. In: *Journal of Instrumentation* 362.14 (Aug. 2008). URL: <https://iopscience.iop.org/article/10.1088/1748-0221/3/08/S08004/meta>.

- [55] CMS Collaboration. “CMS tracker performance and readiness for LHC Run II”. In: *ELSEVIER* (Sept. 2015). URL: <http://dx.doi.org/10.1016/j.nima.2015.09.046>.
- [56] Marzena Lapka. *Materials for CMS ECAL POSTER (CMS POSTER 7.4)*. 2013. URL: <https://cms-docdb.cern.ch/cgi-bin/PublicDocDB/ShowDocument?docid=12030>.
- [57] A Benaglia. “The CMS ECAL performance with examples”. In: *Journal of Instrumentation* 9.02 (Feb. 2014), pp. C02008–C02008. DOI: 10.1088/1748-0221/9/02/c02008. URL: <https://doi.org/10.1088%5C%2F1748-0221%5C%2F9%5C%2F02%5C%2F02008>.
- [58] CMS Collaboration. “CMS Physics. Technical Design Report. Detector Performance and Software”. In: I (Feb. 2006). URL: <http://cdsweb.cern.ch/record/922757/files/lhcc-2006-001.pdf>.
- [59] CMS Collaboration. “HCAL Technical Design Report”. In: (1997). URL: [https://cds.cern.ch/record/357153/files/CMS\\_HCAL\\_TDR.pdf](https://cds.cern.ch/record/357153/files/CMS_HCAL_TDR.pdf).
- [60] Jorge Molina et al. “Performance of CMS hadron calorimeter timing”. In: *Journal of Instrumentation* 5 (Mar. 2010), T03013. DOI: 10.1088/1748-0221/5/03/T03012.
- [61] J.P. Merlo and K. Cankocak. “Radiation-hardness studies of high OH- content quartz fibers irradiated with 24 GeV protons”. In: *CMS CR 2006/005* (2006).
- [62] CMS Collaboration. “Radiation-Hard Quartz Cerenkov Calorimeters”. In: *CMS Conference Report* (Oct. 2006), p. 8. URL: [https://cds.cern.ch/record/1000399/files/CR2006\\_075.pdf](https://cds.cern.ch/record/1000399/files/CR2006_075.pdf).
- [63] Y. Onel A. Penzo. “The CMS-HF quartz fiber calorimeters”. In: *Journal of Physics: Conference Series* 160 (2009) 012014 (2009). URL: [http://inspirehep.net/record/822895/files/jpconf9\\_160\\_012014.pdf](http://inspirehep.net/record/822895/files/jpconf9_160_012014.pdf).
- [64] CMS Collaboration. “The Magnet Project Technical Design Report”. In: (1997). URL: <http://cds.cern.ch/record/331056/?ln=ru>.
- [65] CMS Collaboration. “Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at a centre-of-mass energy of  $\sqrt{s} = 13$  TeV”. In: *arXiv.org* 12 (Apr. 2018), pp. 5–10, 16–19, 28–30. URL: <https://arxiv.org/abs/1804.04528>.
- [66] CMS Collaboration. “The CMS Muon system”. In: *CMS Conference Report* (Apr. 2005). URL: [http://cds.cern.ch/record/927394/files/CR2006\\_006.pdf?version=1](http://cds.cern.ch/record/927394/files/CR2006_006.pdf?version=1).
- [67] CMS Collaboration. *CMS Luminosity - Public Results*. URL: [https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#2015\\_proton\\_proton\\_collisions](https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#2015_proton_proton_collisions).
- [68] The CMS Collaboration. “The CMS trigger system”. In: *ArXiv* (Jan. 2017).
- [69] CMS Collaboration. “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and  $E_T^{\text{miss}}$ ”. In: *CMS PAS PFT-09-001* (Apr. 2009).
- [70] CMS Collaboration. “Description and performance of track and primary-vertex reconstruction with the CMS tracker”. In: *arxiv* (Oct. 2014), pp. 5–9, 13–20, 27, 46–54. URL: <https://arxiv.org/pdf/1405.6569.pdf>.

- [71] CMS Collaboration. “Performance of Jet Algorithms in CMS”. In: (Aug. 2009). URL: <http://cds.cern.ch/record/1198227/files/?ln=ru>.
- [72] Cesare Calabria on behalf of CMS Collaboration. *Tau trigger, reconstruction and identification at CMS*. URL: <http://www2.hepl.phys.nagoya-u.ac.jp/indico/getFile.py/access?contribId=48&sessionId=26&resId=0&materialId=slides&confId=0>.
- [73] N. Estre E. Chabanat. “Deterministic annealing for vertex finding at CMS”. In: *Inspire HEP* (Sept. 2008), pp. 287–290. URL: <http://inspirehep.net/record/707291?ln=en>.
- [74] W. Waltenberger R. Fruhwirth et al. “New vertex reconstruction algorithms for CMS”. In: *Computing in High Energy and Nuclear Physics, La Jolla, California* (Mar. 2003), pp. 288–289. URL: <https://www.ip2i.in2p3.fr/cms/cmstraces/Notes/TULT013.PDF>.
- [75] CMS Collaboration. *Public CMS Luminosity Information*. URL: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>.
- [76] Philip Harris Daniele Bertolini et al. “Pileup Per Particle Identification”. In: *arXiv* 1407.6013 (Sept. 2014). URL: <https://arxiv.org/abs/1407.6013>.
- [77] CMS Collaboration. *How CMS weeds out particles that pile up*. URL: <https://cms.cern/news/how-cms-weeds-out-particles-pile>.
- [78] CMS Collaboration. “Performance of CMS muon reconstruction in pp collision events at  $\sqrt{s} = 7$  TeV”. In: *arXiv.org* (Mar. 2013), pp. 5–7. URL: <https://arxiv.org/pdf/1206.4071.pdf>.
- [79] CMS Collaboration. “Performance of CMS muon reconstruction in cosmic-ray events”. In: *arxiv* (Jan. 2010). URL: <https://arxiv.org/abs/0911.4994>.
- [80] R. Frühwirth. “Application of Kalman filtering to track and vertex fitting”. In: *Science Direct* Volume 262 (Dec. 1987). URL: <https://www.sciencedirect.com/science/article/pii/0168900287908874?via%3Dihub>.
- [81] CMS Collaboration. “Performance of CMS muon reconstruction in pp collision events at  $\sqrt{s} = 7$  TeV”. In: *arxiv* (Mar. 2013). URL: <https://arxiv.org/pdf/1206.4071.pdf>.
- [82] J. Berryhill N. Adam et al. “Generic Tag and Probe Tool for Measuring Efficiency at CMS with Early Data”. In: (Sept. 2009).
- [83] CMS Collaboration. *MET Analysis*. URL: [https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookMetAnalysis#7\\_7\\_6\\_MET\\_Corrections](https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookMetAnalysis#7_7_6_MET_Corrections).
- [84] CMS Collaboration. “Performance of Track-Corrected Missing Transverse Energy in CMS”. In: *CMS-PAS-JME* 09-010 (Dec. 2009). URL: <http://cdsweb.cern.ch/record/1228297>.
- [85] CMS Collaboration. “Performance of missing transverse momentum reconstruction in proton-proton collisions at  $\sqrt{s} = 13$  TeV using the CMS detector”. In: (Mar. 2019). URL: <https://arxiv.org/abs/1903.06078>.
- [86] CMS Collaboration. “Plans for Jet Energy Corrections at CMS”. In: *CMS-PAS-JME* 07-002 (July 2008). URL: <http://cds.cern.ch/record/1194485?ln=en>.



- [87] Olena Karacheban. “Luminosity Measurement at the Compact Muon Solenoid Experiment of the LHC”. PhD thesis. Cham: Brandenburg Tech. U., 2017. DOI: [10.3204/PUBDB-2017-11439](https://doi.org/10.3204/PUBDB-2017-11439), [10.1007/978-3-319-93139-5](https://doi.org/10.1007/978-3-319-93139-5).
- [88] CMS Collaboration. “CMS luminosity measurement for the 2015 data-taking period”. In: *CERN CDS* (Feb. 2017).
- [89] Andreas Kornmayer. *The CMS Pixel Luminosity Telescope*. Tech. rep. CMS-CR-2015-121. Geneva: CERN, June 2015. DOI: [10.1016/j.nima.2015.09.104](https://doi.org/10.1016/j.nima.2015.09.104). URL: <https://cds.cern.ch/record/2039978>.
- [90] E. Castro A. Bell et al. “Fast Beam Conditions Monitor BCM1F for the CMS Experiment”. In: (Dec. 2009). URL: <https://arxiv.org/pdf/0911.2480.pdf>.
- [91] CMS Collaboration. *THE BCM1F DOCUMENTATION*. URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BcmDocumentation>.
- [92] CMS Collaboration. *BCM1F performance plots 2015*. URL: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/BCM1FforEPS2015>.
- [93] Olena Karacheban. *Luminosity Measurement at the Compact Muon Solenoid Experiment of the LHC*. Springer International Publishing, 2018. URL: <https://link.springer.com/book/10.1007%2F978-3-319-93139-5#toc>.
- [94] S van der Meer. *Calibration of the effective beam height in the ISR*. Tech. rep. CERN-ISR-PO-68-31. ISR-PO-68-31. Geneva: CERN, 1968. URL: <http://cds.cern.ch/record/296752>.
- [95] CMS Collaboration. “Fast Beam Condition Monitor for CMS: performance and upgrade”. In: *Nuclear Instruments and Methods in Physics Research Section Accelerators Spectrometers Detectors and Associated Equipment* 765 (May 2014). DOI: [10.1016/j.nima.2014.05.008](https://doi.org/10.1016/j.nima.2014.05.008).
- [96] Moritz Guthoff. “Radiation Damage to the diamond-based Beam Condition Monitor of the CMS Detector at the LHC”. Doctoral thesis. Karlsruher Insituts fur Technologie, 2014.
- [97] M. Ashfold, C.A Rego, and N. Everitt. “Thin Film Diamond by Chemical Vapor Deposition Methods”. In: *Chemical Society Reviews - CHEM SOC REV* 23 (Feb. 1994). DOI: [10.1039/cs9942300021](https://doi.org/10.1039/cs9942300021).
- [98] Robert Visinka Carmen Simons. “Microstructuring of diamonds with laser lithography”. 2nd ADAMAS Workshop @GSI. 2013. URL: [http://www-adamas.gsi.de/ADAMAS02/talks/simons\\_ADAMAS02.pdf](http://www-adamas.gsi.de/ADAMAS02/talks/simons_ADAMAS02.pdf).
- [99] Maria Hempel. “Development of a novel diamond based detector for machine induced background and luminosity measurements”. Master thesis. BTU Cottbus - Senftenberg, 2017.
- [100] D. Przyborowski. “Development of Front-End Electronics for Beam-Condition Monitor at CMS.” Contribution at the 1st ADAMAS Workshop, GSI, Darmstadt, 2012. URL: <http://www-adamas.gsi.de/ADAMAS01/talks/przyborowski.pdf>.
- [101] Bheesette Srinidhi. “Characterisation of the BCM1F Analog Opto Hybrids for the CMS experiment. BCM1F-Fast Beam Condition Monitoring ”. In: (Aug. 2013). URL: <https://cds.cern.ch/record/1596774>.

- [102] *Optical Links for CMS webpage*. URL: <http://cms-tk-opto.web.cern.ch/cms-tk-opto/>.
- [103] Richard Grigonis. *Versa Module Europa (VME) Bus*. URL: <https://www.globalspec.com/reference/41898/203279/chapter-v-versa-module-europa-vme-bus-vpim>.
- [104] PICMG. *MicroTCA Overview*. URL: <https://www.picmg.org/openstandards/microtca/>.
- [105] <http://www.onmyphd.com>. *Analog-Digital Converters*. URL: <http://www.onmyphd.com/?p=analog.digital.converter>.
- [106] R S Schmidt et al. “Performance of the Fast Beam Conditions Monitor BCM1F of CMS in the first running periods of LHC”. In: *Journal of Instrumentation* 6.01 (Jan. 2011), pp. C01004–C01004. ISSN: 1748-0221. DOI: 10.1088/1748-0221/6/01/c01004. URL: <http://dx.doi.org/10.1088/1748-0221/6/01/C01004>.
- [107] PDG Collaboration. “Review of Particle Physics”. In: *Phys. Rev. D* 86 (1 July 2012), p. 010001. DOI: 10.1103/PhysRevD.86.010001. URL: <https://link.aps.org/doi/10.1103/PhysRevD.86.010001>.
- [108] Martin Stegler. “Characterisation of non-irradiated and irradiated diamond sensors before and after application at LHC”. Master thesis. Brandenburg University of Technology, 2015.
- [109] C. Canali et al. “Electrical properties and performances of natural diamond nuclear radiation detectors”. In: *Nuclear Instruments and Methods* 160.1 (1979), pp. 73–77. ISSN: 0029-554X. DOI: [https://doi.org/10.1016/0029-554X\(79\)90167-8](https://doi.org/10.1016/0029-554X(79)90167-8). URL: <http://www.sciencedirect.com/science/article/pii/0029554X79901678>.
- [110] Circuit Cellar. *Pulse-Shaping Basics*. URL: <https://circuitcellar.com/cc-blog/pulse-shaping-basics/>.
- [111] CAEN. *DIP Description*. 2004. URL: <https://edms.cern.ch/file/457113/2/DIPDescription.doc>.
- [112] Alexander Sinclair Howard. *Diamond detectors for particle physics*. University of London, 1998.
- [113] W. Salter et al. *CAEN N625*. URL: <https://www.caen.it/products/n625/>.
- [114] William Badgett et al. “Web Based Monitoring in the CMS Experiment at CERN”. In: (2014). arXiv: 1409.1133 [physics.ins-det].
- [115] CMS Collaboration. “Measurement of inclusive W and Z boson cross section in pp collisions at  $\sqrt{s} = 13$  TeV”. In: *CMS AN - unpublished* 2016/250 (2016). URL: [http://cms.cern.ch/iCMS/jsp/db\\_notes/noteInfo.jsp?cmsnoteid=CMS%20AN-2016/250](http://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%20AN-2016/250).
- [116] C. Oleari. “The POWHEG BOX”. In: *Nuclear Physics B - Proceedings Supplements* 205-206 (Aug. 2010), pp. 36–41. ISSN: 0920-5632. DOI: 10.1016/j.nuclphysbps.2010.08.016. URL: <http://dx.doi.org/10.1016/j.nuclphysbps.2010.08.016>.

- [117] CMS Collaboration. “Measurement of the Rapidity and Transverse Momentum Distributions of Z Bosons in pp Collisions at  $\sqrt{s}=7$  TeV”. In: *arXiv:1110.4973* (22 Oct 2011). URL: <https://arxiv.org/abs/1110.4973>.
- [118] A. van Dyne A. Bodek et al. “Extracting Muon Momentum Scale Corrections”. In: *ArXiv:1208.3710v3* (Sept. 2012), p. 6.
- [119] CMS Collaboration. “Measurement of the charge ratio of atmospheric muons with the CMS detector”. In: *ArXiv* (May 2010). URL: <https://arxiv.org/pdf/1005.5332.pdf>.
- [120] Roberto Pittau. *Status of MadLoop/aMC@NLO*. 2012. arXiv: 1202.5781 [hep-ph].
- [121] Torbjörn Sjöstrand. *PYTHIA 8 Status Report*. 2008. arXiv: 0809.0303 [hep-ph].
- [122] Richard D. Ball et al. “Parton distributions for the LHC run II”. In: *Journal of High Energy Physics* 2015.4 (Apr. 2015). ISSN: 1029-8479. DOI: 10.1007/jhep04(2015)040. URL: [http://dx.doi.org/10.1007/JHEP04\(2015\)040](http://dx.doi.org/10.1007/JHEP04(2015)040).
- [123] Richard D. Ball et al. “Parton distributions with LHC data”. In: *Nuclear Physics B* 867.2 (Feb. 2013), pp. 244–289. ISSN: 0550-3213. DOI: 10.1016/j.nuclphysb.2012.10.003. URL: <http://dx.doi.org/10.1016/j.nuclphysb.2012.10.003>.
- [124] Wouter Verkerke and David Kirkby. *The RooFit toolkit for data modeling*. 2003. arXiv: physics/0306116 [physics.data-an].
- [125] *MINUIT homepage*. URL: <https://seal.web.cern.ch/seal/snapshot/work-packages/mathlibs/minuit/>.
- [126] G. Breit and E. Wigner. “Capture of Slow Neutrons”. In: *Phys. Rev.* 49 (7 Apr. 1936), pp. 106–109. DOI: 10.1103/PhysRev.49.519. URL: <https://link.aps.org/doi/10.1103/PhysRev.49.519>.
- [127] T. Skwarnicki. *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*. 1986. DESY F31-86-02: *PhDThesis* (physics.data-an).
- [128] *CMS luminosity measurement for the 2015 data-taking period*. Tech. rep. CMS-PAS-LUM-15-001. Geneva: CERN, 2017. URL: <http://cds.cern.ch/record/2138682>.
- [129] Radja Boughezal et al. “Color-singlet production at NNLO in MCFM”. In: *The European Physical Journal C* 77.1 (Dec. 2016). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4558-y. URL: <http://dx.doi.org/10.1140/epjc/s10052-016-4558-y>.
- [130] Sayipjamal Dulat et al. “New parton distribution functions from a global analysis of quantum chromodynamics”. In: *Physical Review D* 93.3 (Feb. 2016). ISSN: 2470-0029. DOI: 10.1103/physrevd.93.033006. URL: <http://dx.doi.org/10.1103/PhysRevD.93.033006>.
- [131] A. D. Martin et al. “Extended parameterisations for MSTW PDFs and their effect on lepton charge asymmetry from W decays”. In: *The European Physical Journal C* 73.2 (Feb. 2013). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-013-2318-9. URL: <http://dx.doi.org/10.1140/epjc/s10052-013-2318-9>.

- [132] L. A. Harland-Lang et al. “Parton distributions in the LHC era: MMHT 2014 PDFs”. In: *The European Physical Journal C* 75.5 (May 2015). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-015-3397-6](https://doi.org/10.1140/epjc/s10052-015-3397-6). URL: <http://dx.doi.org/10.1140/epjc/s10052-015-3397-6>.
- [133] E.L. Berger et al. In: *Madison preprint* University of Wisconsin.MAD/PH/400 (1988).
- [134] E. L. Berger et al. “Weak-boson production at Fermilab Tevatron energies”. In: *Phys. Rev. D* 40 (1 July 1989), pp. 83–91. DOI: [10.1103/PhysRevD.40.83](https://doi.org/10.1103/PhysRevD.40.83). URL: <https://link.aps.org/doi/10.1103/PhysRevD.40.83>.
- [135] J. L. Holzbauer. *Measurements of W charge asymmetry*. 2015. arXiv: [1510.01716](https://arxiv.org/abs/1510.01716) [hep-ex].
- [136] T. Aaltonen et al. “Direct Measurement of the W Production Charge Asymmetry in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.96$  TeV”. In: *Physical Review Letters* 102.18 (May 2009). ISSN: 1079-7114. DOI: [10.1103/physrevlett.102.181801](https://doi.org/10.1103/physrevlett.102.181801). URL: <http://dx.doi.org/10.1103/PhysRevLett.102.181801>.
- [137] S. Chatrchyan et al. “Measurement of the muon charge asymmetry in inclusive  $pp \rightarrow W + X$  production at  $\sqrt{s} = 7$  TeV and an improved determination of light parton distribution functions”. In: *Physical Review D* 90.3 (Aug. 2014). ISSN: 1550-2368. DOI: [10.1103/physrevd.90.032004](https://doi.org/10.1103/physrevd.90.032004). URL: <http://dx.doi.org/10.1103/PhysRevD.90.032004>.
- [138] V. Khachatryan et al. “Measurement of the differential cross section and charge asymmetry for inclusive  $pp \rightarrow W^\pm + X$   $\sqrt{s} = 8$  TeV”. In: *The European Physical Journal C* 76.8 (Aug. 2016). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-016-4293-4](https://doi.org/10.1140/epjc/s10052-016-4293-4). URL: <http://dx.doi.org/10.1140/epjc/s10052-016-4293-4>.
- [139] M. Botje. “QCDNUM: Fast QCD evolution and convolution”. In: *Computer Physics Communications* 182.2 (Feb. 2011), pp. 490–532. ISSN: 0010-4655. DOI: [10.1016/j.cpc.2010.10.020](https://doi.org/10.1016/j.cpc.2010.10.020). URL: <http://dx.doi.org/10.1016/j.cpc.2010.10.020>.
- [140] Emil Avsar and Edmond Iancu. “CCFM evolution with unitarity corrections”. In: *Nuclear Physics A* 829.1-2 (Oct. 2009), pp. 31–75. ISSN: 0375-9474. DOI: [10.1016/j.nuclphysa.2009.08.001](https://doi.org/10.1016/j.nuclphysa.2009.08.001). URL: <http://dx.doi.org/10.1016/j.nuclphysa.2009.08.001>.
- [141] F. Hautmann, H. Jung, and S. Taheri Monfared. “The CCFM uPDF evolution uPDFevol Version 1.0.00”. In: *The European Physical Journal C* 74.10 (Oct. 2014). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-014-3082-1](https://doi.org/10.1140/epjc/s10052-014-3082-1). URL: <http://dx.doi.org/10.1140/epjc/s10052-014-3082-1>.
- [142] John Campbell et al. *MCFM - Monte Carlo for FeMtobarn processes*. 2001. URL: <https://mcfm.fnal.gov/>.
- [143] Tancredi Carli et al. “A posteriori inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders: the APPLGRID project”. In: *The European Physical Journal C* 66.3-4 (Feb. 2010), pp. 503–524. ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-010-1255-0](https://doi.org/10.1140/epjc/s10052-010-1255-0). URL: <http://dx.doi.org/10.1140/epjc/s10052-010-1255-0>.

- [144] Michal Czakon, David Heymes, and Alexander Mitov. *fastNLO tables for NNLO top-quark pair differential distributions*. 2017. arXiv: 1704.08551 [hep-ph].
- [145] CMS Collaboration. “Measurement of the lepton charge asymmetry in inclusive W production in pp collisions at  $\sqrt{s} = 7\text{TeV}$ ”. In: *ArXive* (2011). URL: <https://arxiv.org/abs/1103.3470>.
- [146] S. Chatrchyan et al. “Measurement of associated W + charm production in pp collisions at  $\sqrt{s}=7\text{ TeV}$ ”. In: *Journal of High Energy Physics* 2014.2 (Feb. 2014). ISSN: 1029-8479. DOI: 10.1007/jhep02(2014)013. URL: [http://dx.doi.org/10.1007/JHEP02\(2014\)013](http://dx.doi.org/10.1007/JHEP02(2014)013).
- [147] A. M. Sirunyan et al. “Measurement of associated production of a W boson and a charm quark in proton–proton collisions at  $\sqrt{s} = 13\text{ TeV}$ ”. In: *The European Physical Journal C* 79.3 (Mar. 2019). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-6752-1. URL: <http://dx.doi.org/10.1140/epjc/s10052-019-6752-1>.
- [148] F. D. Aaron et al. “Measurement of the inclusive ep scattering cross section at low Q<sup>2</sup> and x at HERA”. In: *The European Physical Journal C* 63.4 (Sept. 2009), pp. 625–678. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-009-1128-6. URL: <http://dx.doi.org/10.1140/epjc/s10052-009-1128-6>.
- [149] H1 Collaboration. *Inclusive Deep Inelastic Scattering at High Q<sup>2</sup> with Longitudinally Polarised Lepton Beams at HERA*. 2012. arXiv: 1206.7007 [hep-ex].
- [150] CMS Collaboration. “Measurement of the differential cross section and charge asymmetry for inclusive pp to W + X production at  $\sqrt{s} = 8\text{TeV}$ ”. In: *ArXive* (2016). URL: <https://arxiv.org/abs/1603.01803>.
- [151] A. D. Martin et al. “Parton distributions for the LHC”. In: *The European Physical Journal C* 63.2 (July 2009), pp. 189–285. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-009-1072-5. URL: <http://dx.doi.org/10.1140/epjc/s10052-009-1072-5>.
- [152] R. S. Thorne. “Effect of changes of variable flavor number scheme on parton distribution functions and predicted cross sections”. In: *Physical Review D* 86.7 (Oct. 2012). ISSN: 1550-2368. DOI: 10.1103/physrevd.86.074017. URL: <http://dx.doi.org/10.1103/PhysRevD.86.074017>.
- [153] M Botje. “Error estimates on parton density distributions”. In: *Journal of Physics G: Nuclear and Particle Physics* 28.5 (Apr. 2002), pp. 779–789. ISSN: 0954-3899. DOI: 10.1088/0954-3899/28/5/305. URL: <http://dx.doi.org/10.1088/0954-3899/28/5/305>.
- [154] Mark Sutton et al. *The APPLgrid project*. 2010. URL: <https://applgrid.hepforge.org/index.htm>.
- [155] Zoltan Nagy. *NLOJET++*. 2011. URL: <http://www.desy.de/~znagy/Site/NLOJet++.html>.
- [156] Andy Buckley et al. *Les Houches Approach*. 2019. URL: <https://lhpdf.hepforge.org/>.
- [157] Andy Buckley et al. “LHAPDF6: parton density access in the LHC precision era”. In: *The European Physical Journal C* 75.3 (Mar. 2015). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-015-3318-8. URL: <http://dx.doi.org/10.1140/epjc/s10052-015-3318-8>.

- [158] Particle Data Group. *MONTE CARLO PARTICLE NUMBERING SCHEME*. URL: <http://pdg.lbl.gov/2007/reviews/montecarlohpp.pdf>.
- [159] John M. Campbell, R. Keith Ellis, and Ciaran Williams. *MCFM - Monte Carlo for FeMtobarn processes*. 2020. URL: <https://mcfm.fnal.gov/>.
- [160] J. M. Campbell and R. K. Ellis. “Update on vector boson pair production at hadron colliders”. In: *Physical Review D* 60.11 (Nov. 1999). ISSN: 1089-4918. DOI: [10.1103/physrevd.60.113006](https://doi.org/10.1103/physrevd.60.113006). URL: <http://dx.doi.org/10.1103/PhysRevD.60.113006>.
- [161] John M. Campbell, R. Keith Ellis, and Ciaran Williams. “Vector boson pair production at the LHC”. In: *Journal of High Energy Physics* 2011.7 (July 2011). ISSN: 1029-8479. DOI: [10.1007/jhep07\(2011\)018](https://doi.org/10.1007/jhep07(2011)018). URL: [http://dx.doi.org/10.1007/JHEP07\(2011\)018](http://dx.doi.org/10.1007/JHEP07(2011)018).
- [162] Iain W. Stewart, Frank J. Tackmann, and Wouter J. Waalewijn. “NJetteness: An Inclusive Event Shape to Veto Jets”. In: *Physical Review Letters* 105.9 (Aug. 2010). ISSN: 1079-7114. DOI: [10.1103/physrevlett.105.092002](https://doi.org/10.1103/physrevlett.105.092002). URL: <http://dx.doi.org/10.1103/PhysRevLett.105.092002>.
- [163] G.P. Salam and J. Rojo. “A Higher Order Perturbative Parton Evolution Toolkit (HOPPET)”. In: *Computer Physics Communications* 180.1 (Jan. 2009), pp. 120–156. ISSN: 0010-4655. DOI: [10.1016/j.cpc.2008.08.010](https://doi.org/10.1016/j.cpc.2008.08.010). URL: <http://dx.doi.org/10.1016/j.cpc.2008.08.010>.

## Appendix A

# BCM1F sensor measurement results

The Appendix presents detailed results of pCVD and sCVD sensor measurements, performed by the author at DESY-Zeuthen laboratories in 2016-2017. In particular, results are presented for twelve pCVD and five sCVD sensors. Each sensor is shown with its current over voltage (IV) measurement, current over time (CT) measurement, and charge collection distance (CCD) estimation.

## A.1 Poly-crystalline Chemical Vapor Deposited (pCVD) diamonds

### A.1.1 pCVD 07B10412

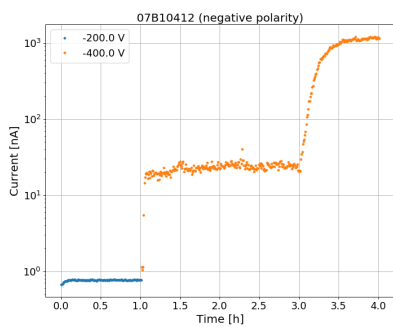


FIGURE A.1:  
CT at 200V  
400V (negative  
attenuated)

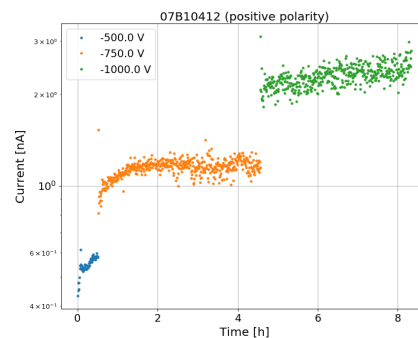


FIGURE A.2: CT  
at 500V 750V  
1000V (positive  
attenuated)

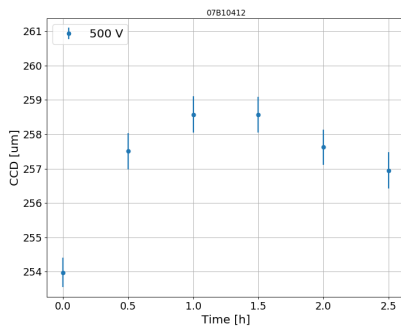


FIGURE A.3:  
CCD at 500V

### A.1.2 pCVD 07B10415

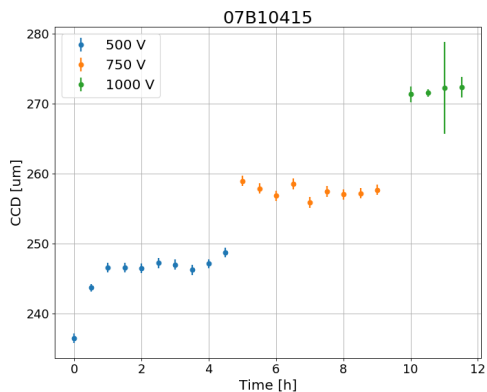


FIGURE A.4: CCD at 500V 750V 1000V

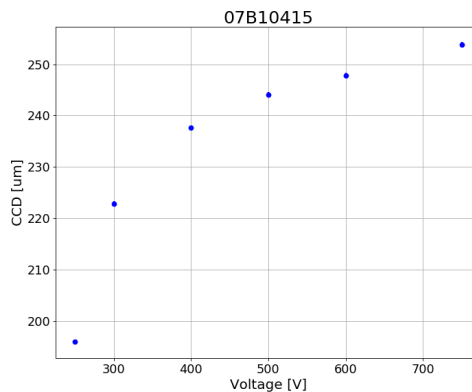


FIGURE A.5: CCD vs HV

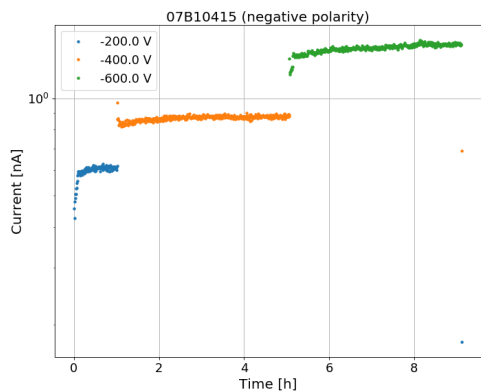


FIGURE A.6: CT at 200V 400V 600V (negative attenuated)

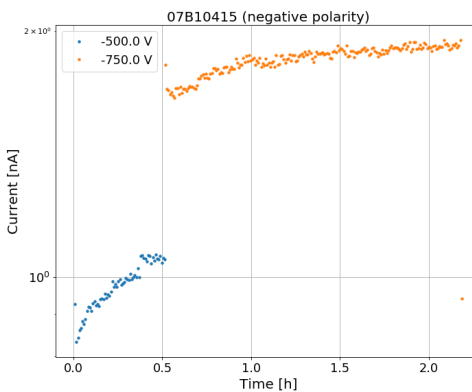


FIGURE A.7: CT at 500V 750V (negative attenuated)

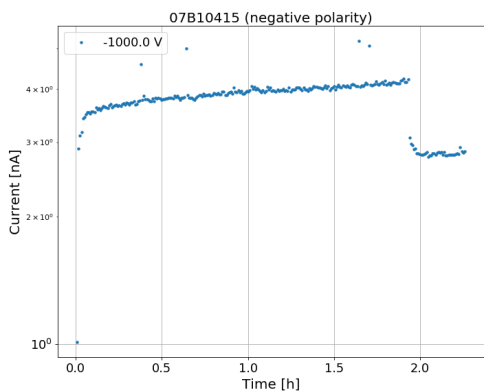


FIGURE A.8: CT at 1000V (negative attenuated)



A.1.3 pCVD 07B10417

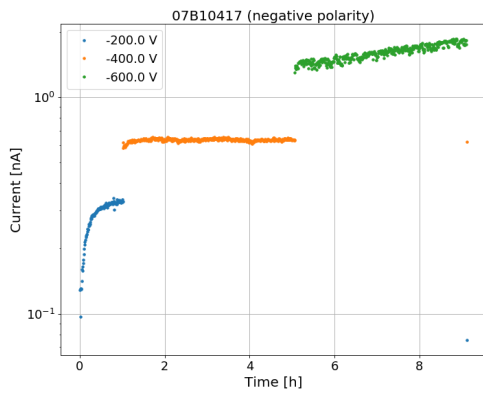


FIGURE A.9: CT at 200V 400V 600V (negative attenuated)

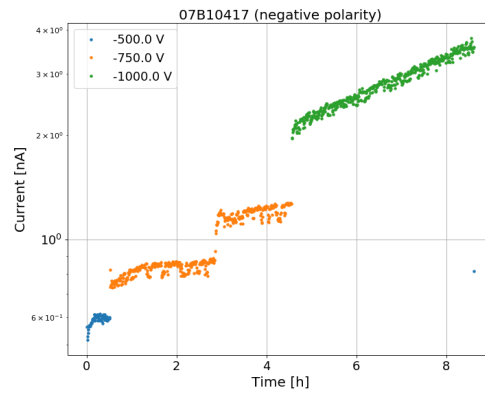


FIGURE A.10: CT at 500V 750V 1000V(negative attenuated)

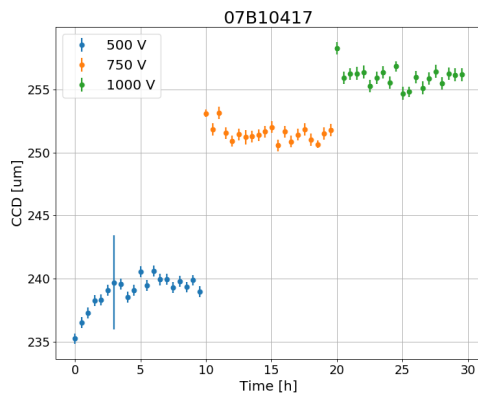


FIGURE A.11: CCD at 500V 750V 1000V

## A.1.4 pCVD 07B10421

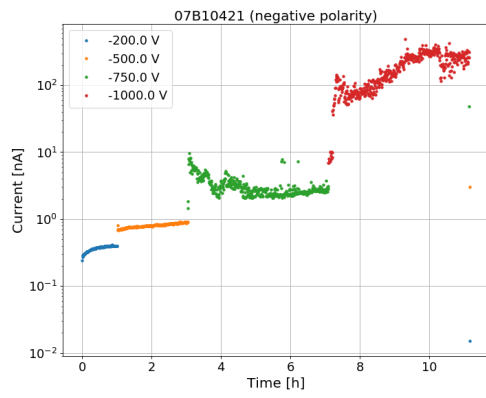


FIGURE A.12: CT at  
200V 500V 750V 1000V  
(negative attenuated)  
Medium

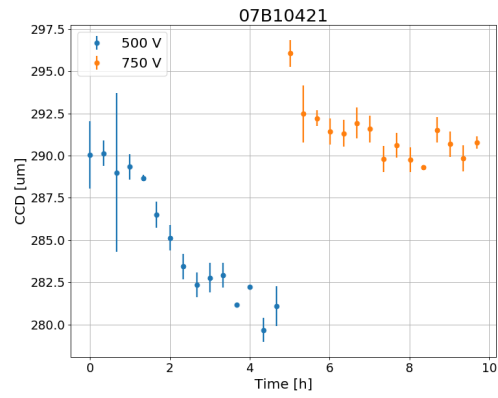


FIGURE A.13: CCD at  
500V 750V

A.1.5 pCVD 07B10422

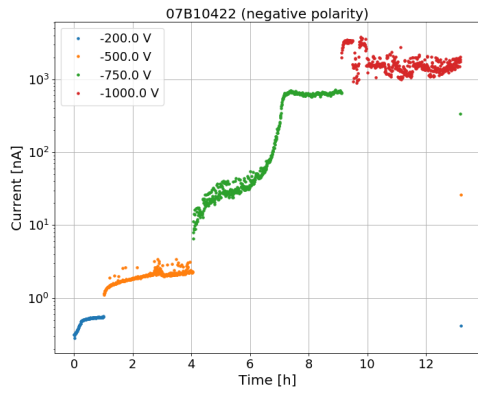


FIGURE A.14: CT at 200V 500V 750V 1000V (negative attenuated)

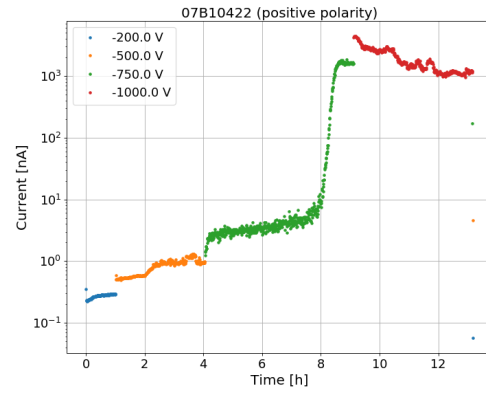


FIGURE A.15: CT at 500V 750V 1000V (positive attenuated)

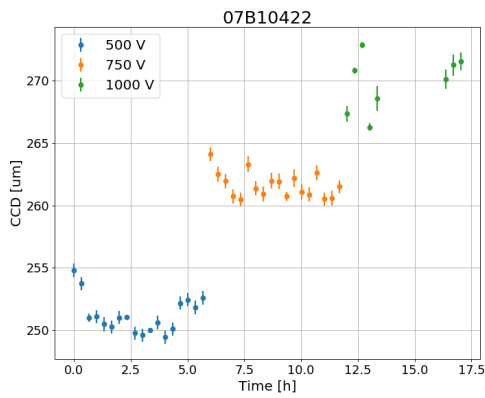


FIGURE A.16: CCD at 500V 750V 1000V

## A.1.6 pCVD 07B10423

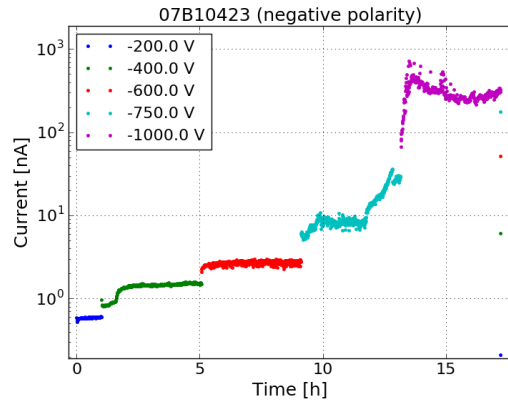


FIGURE A.17: CT at  
200V 400V 600V 750V  
1000V (negative attenu-  
ated)  
Medium

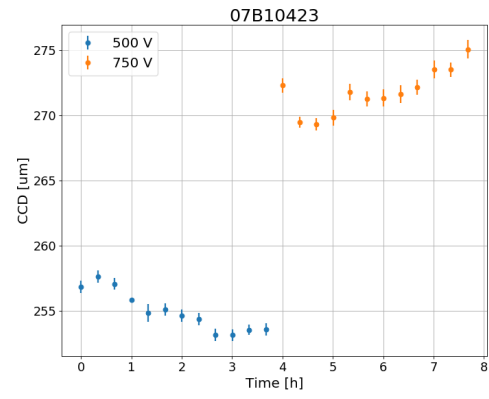


FIGURE A.18: CCD at  
500V 750V

A.1.7 pCVD 07B10430

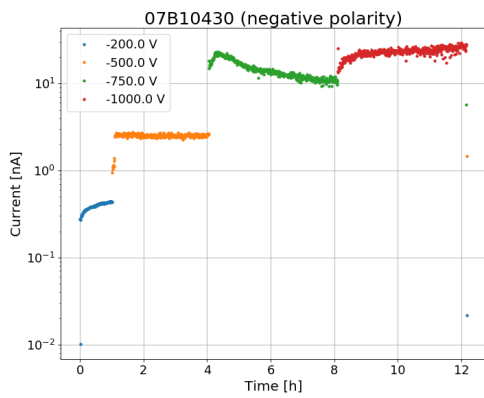


FIGURE A.19: CT at 200V 500V 750V 1000V (negative attenuated)

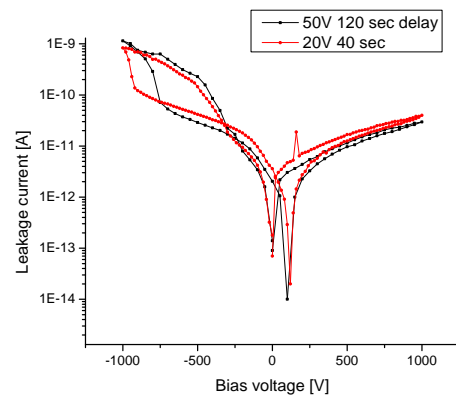


FIGURE A.20: IV

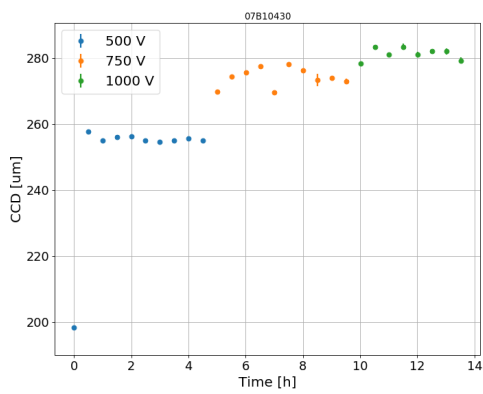


FIGURE A.21: CCD at 500V 750V 1000V

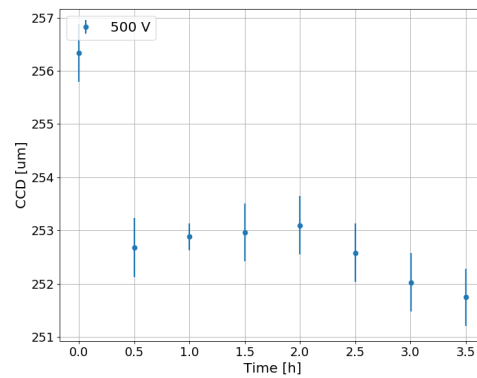


FIGURE A.22: CCD 500V 4h check

Good, but we had to remeasure CCD many times!

## A.1.8 pCVD 07B10431

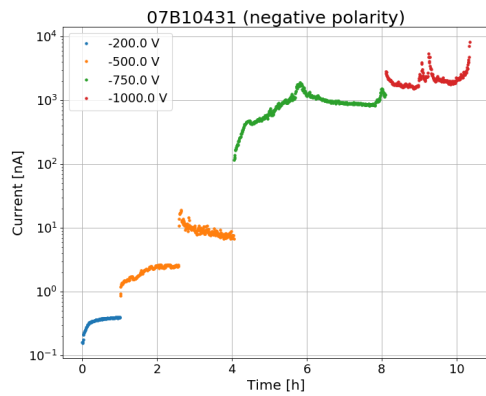


FIGURE A.23: CT at 200V 500V 750V 1000V

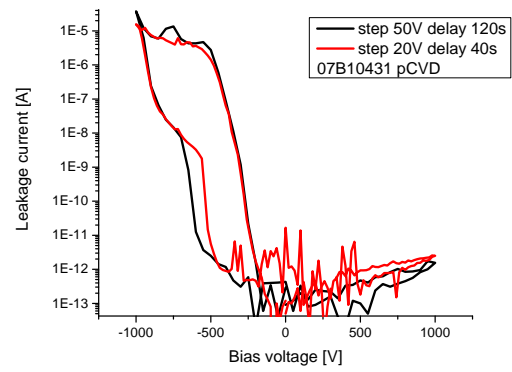


FIGURE A.24: IV

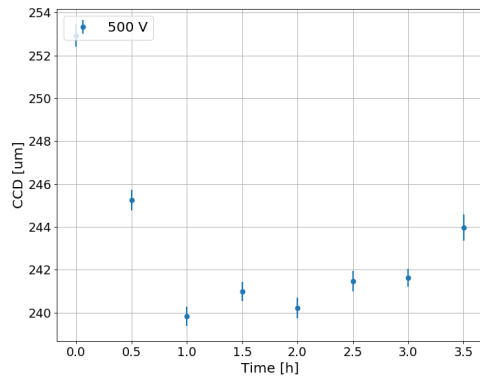


FIGURE A.25: CCD at 500V

A.1.9 pCVD 07B10433

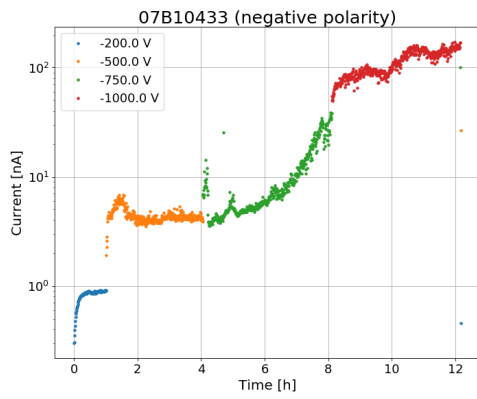


FIGURE A.26: CT at 200V 500V 750V 1000V

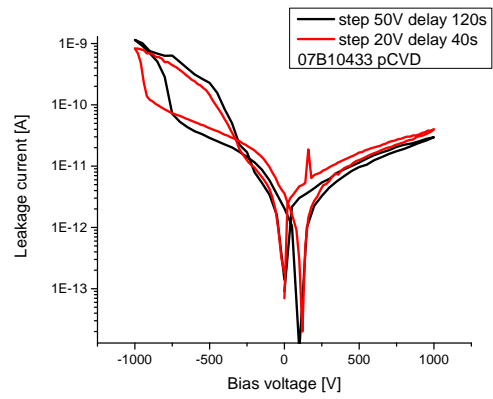


FIGURE A.27: IV

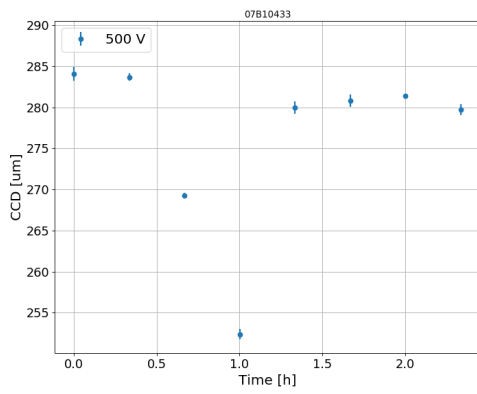


FIGURE A.28: CCD at 500V

## A.1.10 pCVD 07B10434

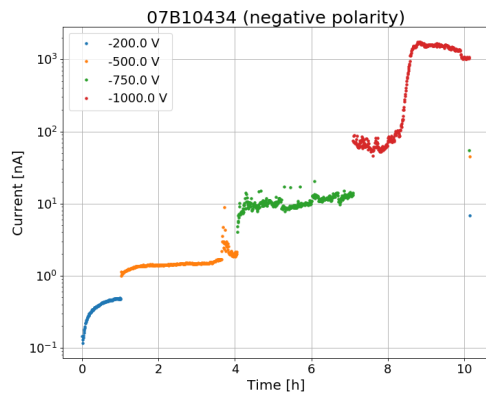
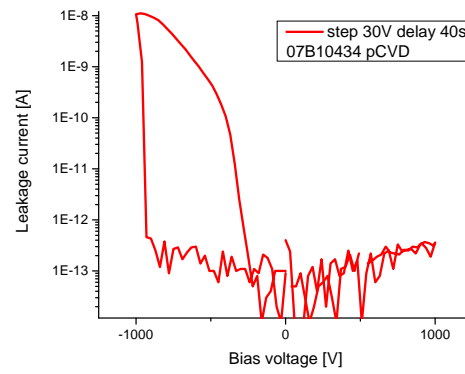
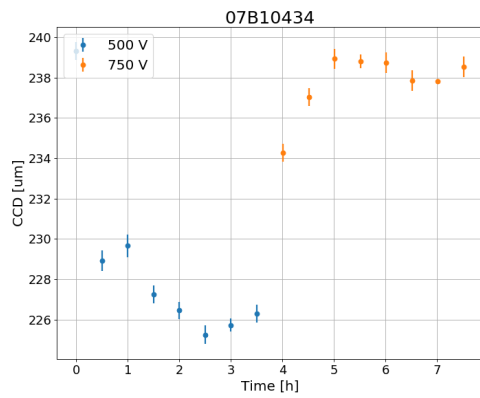
FIGURE A.29: CT at  
200V 500V 750V 1000V

FIGURE A.30: IV

FIGURE A.31: CCD at  
500V 750V



A.1.11 pCVD 07B10435

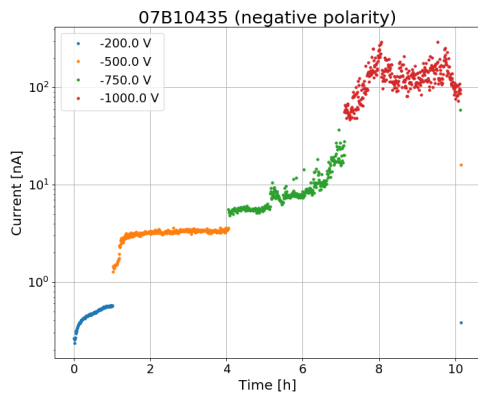


FIGURE A.32: CT at 200V 500V 750V 1000V

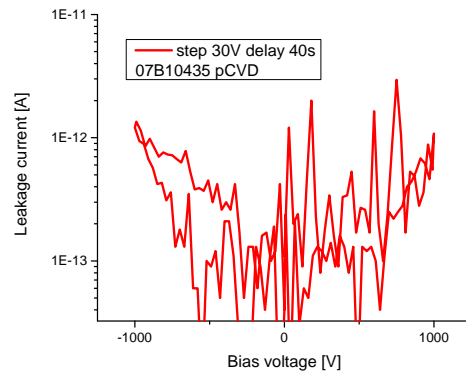


FIGURE A.33: IV

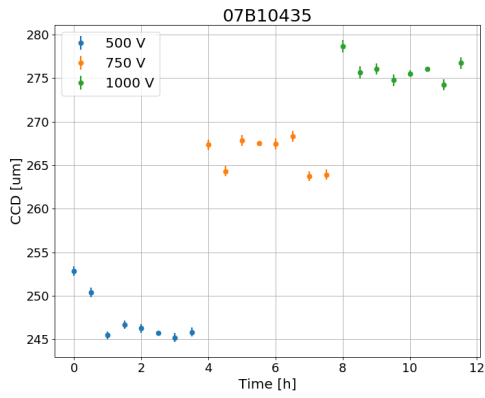


FIGURE A.34: CCD at 500V 750V 1000V

A.1.12 pCVD 07B10436

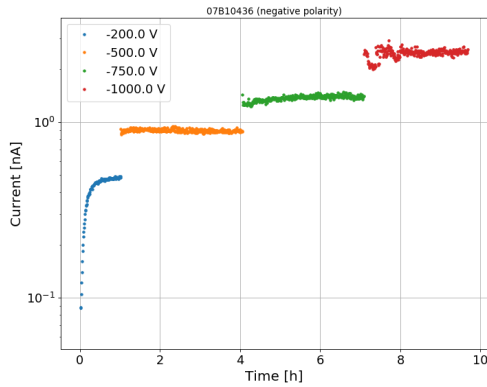


FIGURE A.35: CT at 200V 500V 750V 1000V

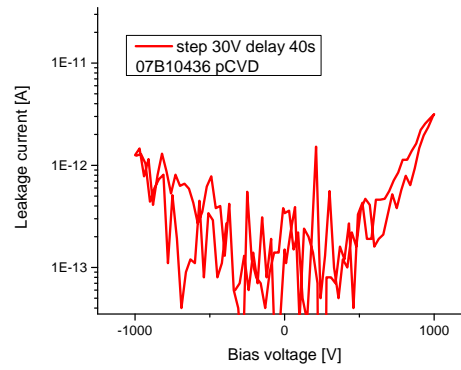


FIGURE A.36: IV

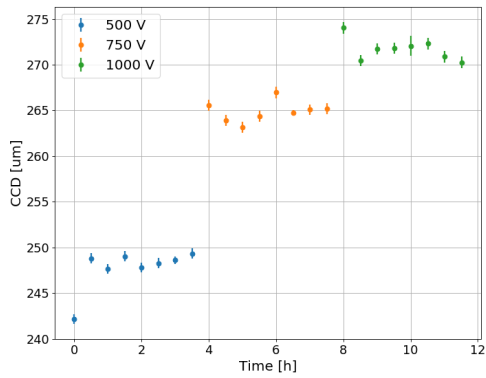


FIGURE A.37: CCD at 500V 750V 1000V

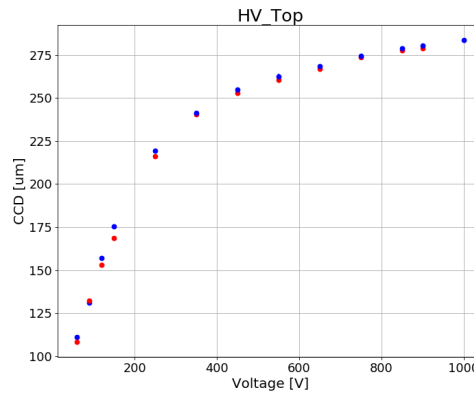


FIGURE A.38: CCD vs HV Hysteresis

## A.2 Single-crystalline Chemical Vapor Deposited (sCVD) diamonds

### A.2.1 sCVD PLT S55(2-1-2)

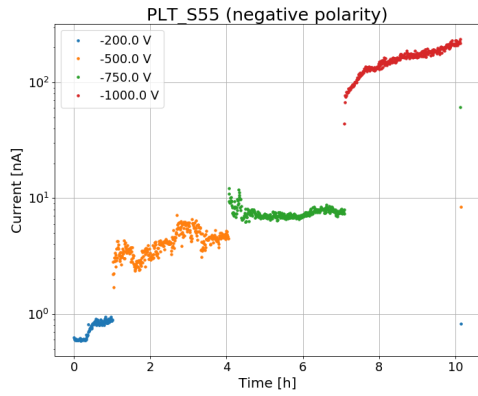


FIGURE A.39: CT at 200V 500V 750V 1000V

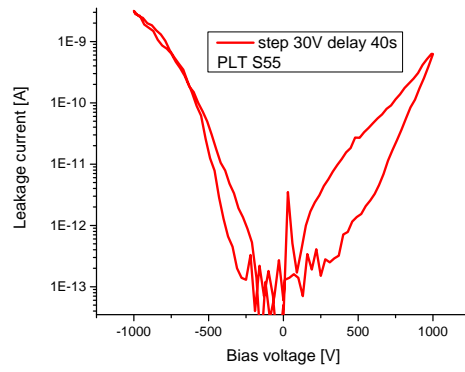


FIGURE A.40: IV

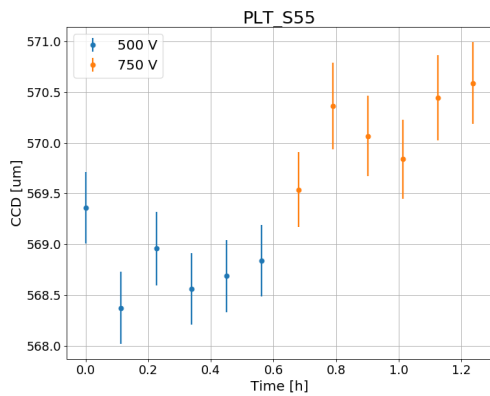


FIGURE A.41: CCD at 500V 750V

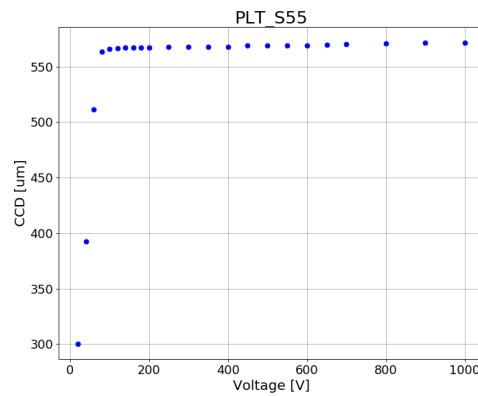


FIGURE A.42: CCD vs HV

### A.2.2 sCVD PLT S115(2-1-6)

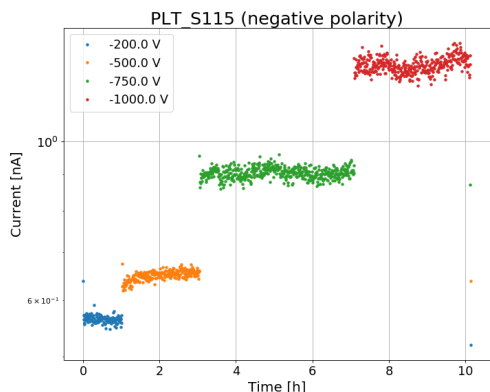


FIGURE A.43: CT at 200V 500V 750V 1000V

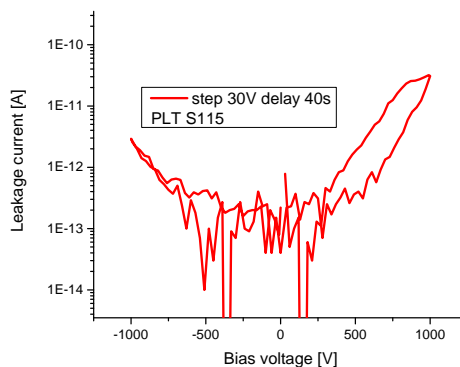


FIGURE A.44: IV

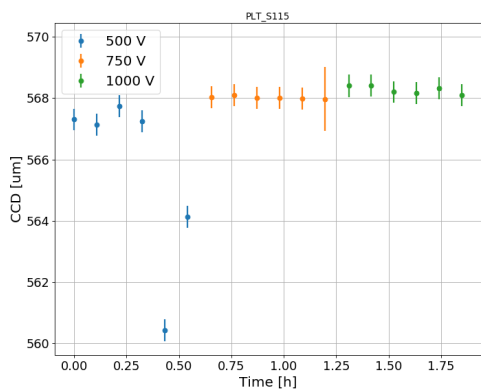


FIGURE A.45: CCD at 500V 750V 1000V

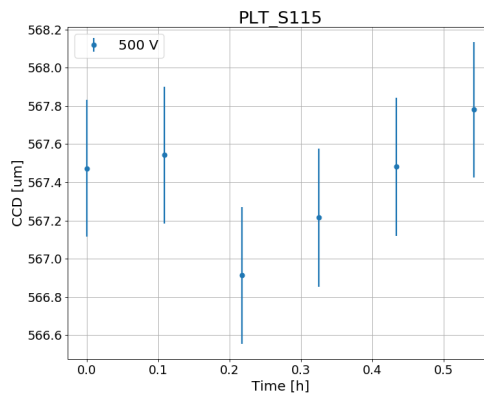


FIGURE A.46: remeasured HV at 500V

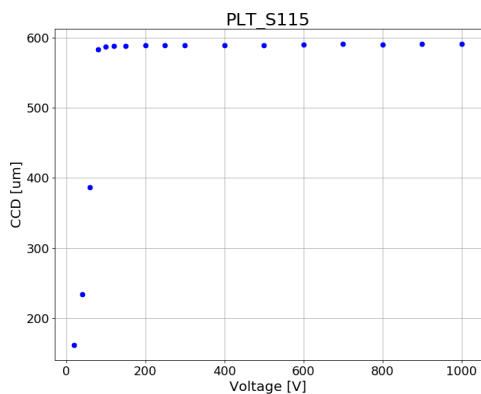


FIGURE A.47: CCD vs HV

A.2.3 sCVD PLT S119(2-1-3)

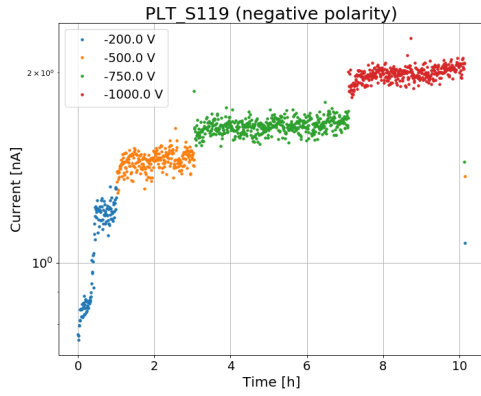


FIGURE A.48: CT at 200V 500V 750V 1000V

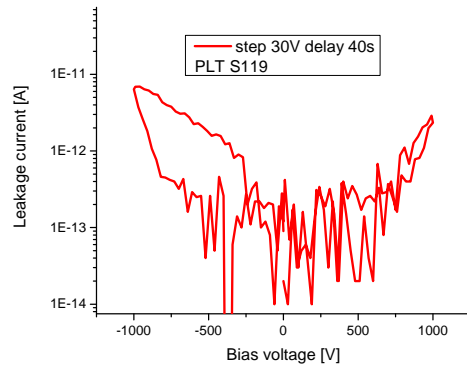


FIGURE A.49: IV

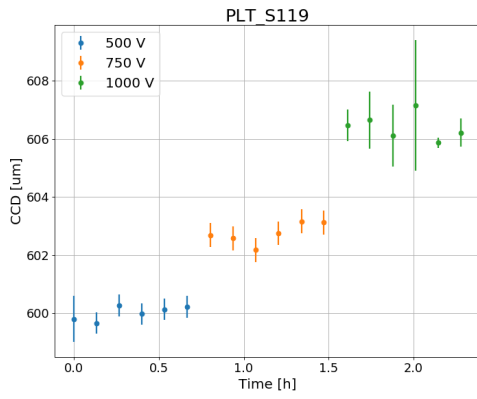


FIGURE A.50: CCD at 500V 750V 1000V

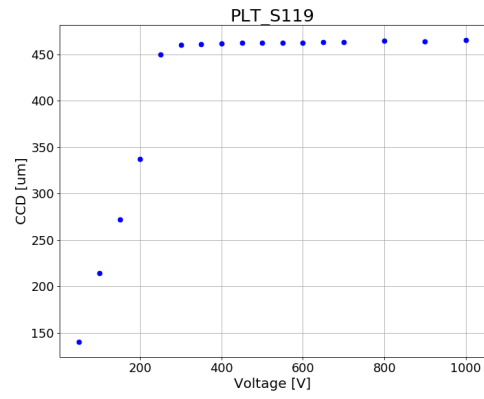


FIGURE A.51: CCD vs HV

### A.2.4 sCVD Batch390 2499089-1 (3-3-4)

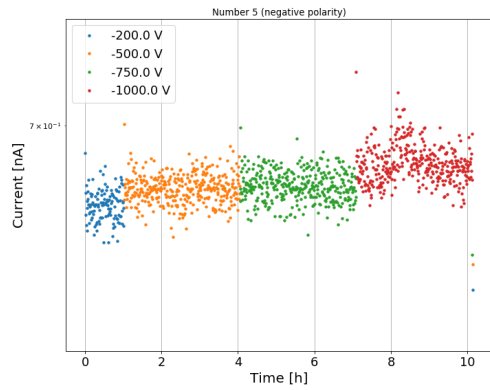


FIGURE A.52: CT at 200V 500V 750V 1000V

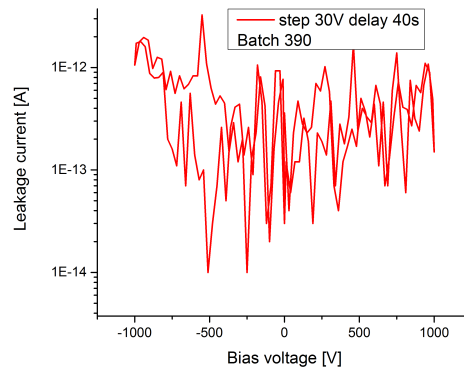


FIGURE A.53: IV

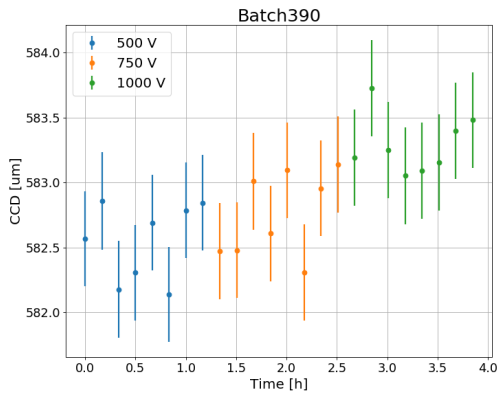


FIGURE A.54: CCD at 500V 750V 1000V

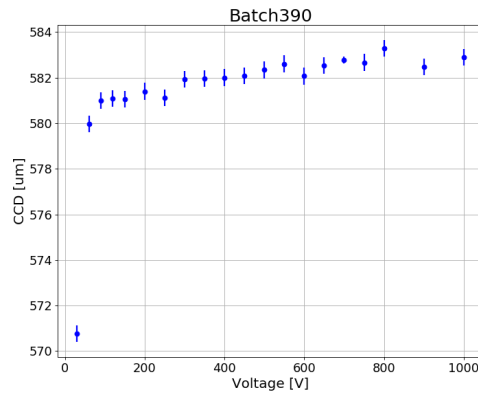


FIGURE A.55: CCD vs HV

A.2.5 sCVD 2713547-5 "Number 5"

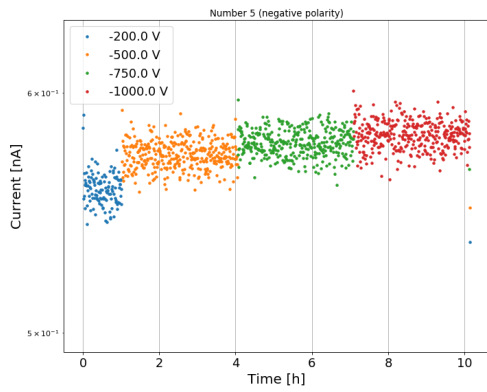


FIGURE A.56: CT at 200V 500V 750V 1000V

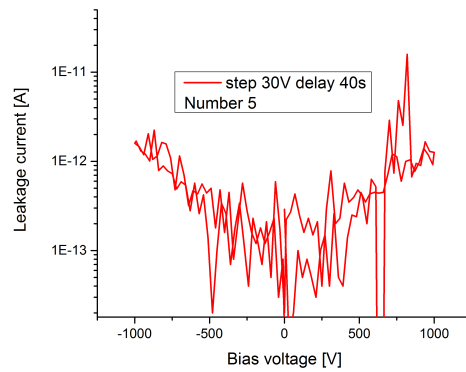


FIGURE A.57: IV

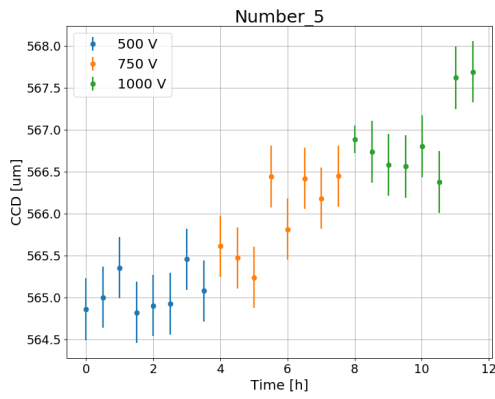


FIGURE A.58: CCD at 500V 750V 1000V

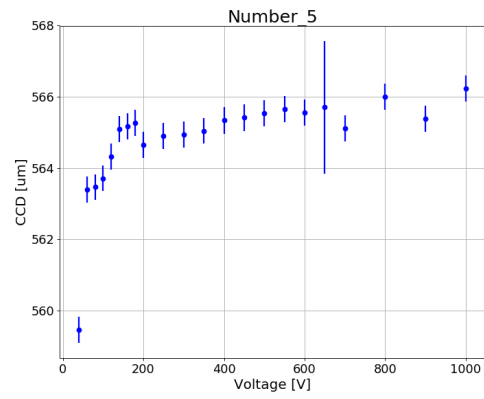


FIGURE A.59: CCD vs HV





## Appendix B

# Measurement of the $W^\pm$ charge asymmetry

This appendix presents the full set of results from  $E_T^{\text{miss}}$  fits of  $W^\pm$  boson yields extraction per pseudorapidity region. The plots are presented in logarithmic scale. Results for each pseudorapidity bin are shown from signal (left plot) and control (right plot) regions. Figures B.1-B.4 are dedicated to the  $W^+ \rightarrow \mu^+ \nu_\mu$  channel, while Fig. B.5-B.8 to the  $W^- \rightarrow \mu^- \bar{\nu}_\mu$  channel.

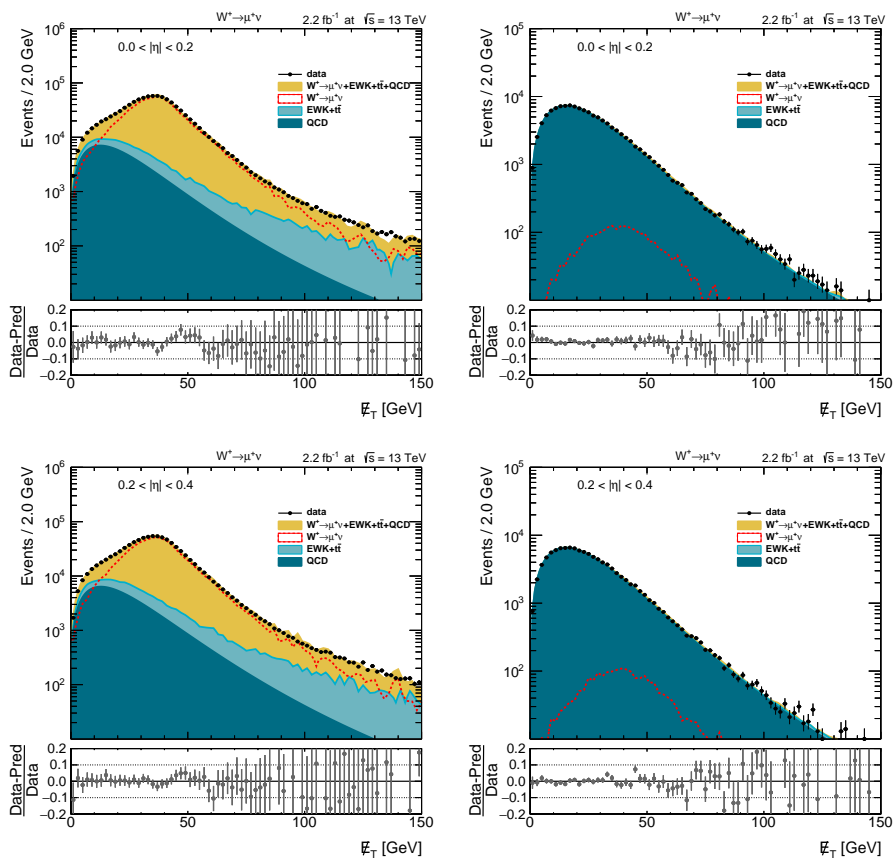


FIGURE B.1: Results of simultaneous fit of the  $E_T^{\text{miss}}$  distribution in signal (left) and control (right) regions for  $W^+ \rightarrow \mu^+ \nu_\mu$ . The upper row corresponds to  $0.0 < |\eta| < 0.2$  region and the last row to  $0.2 < |\eta| < 0.4$ .

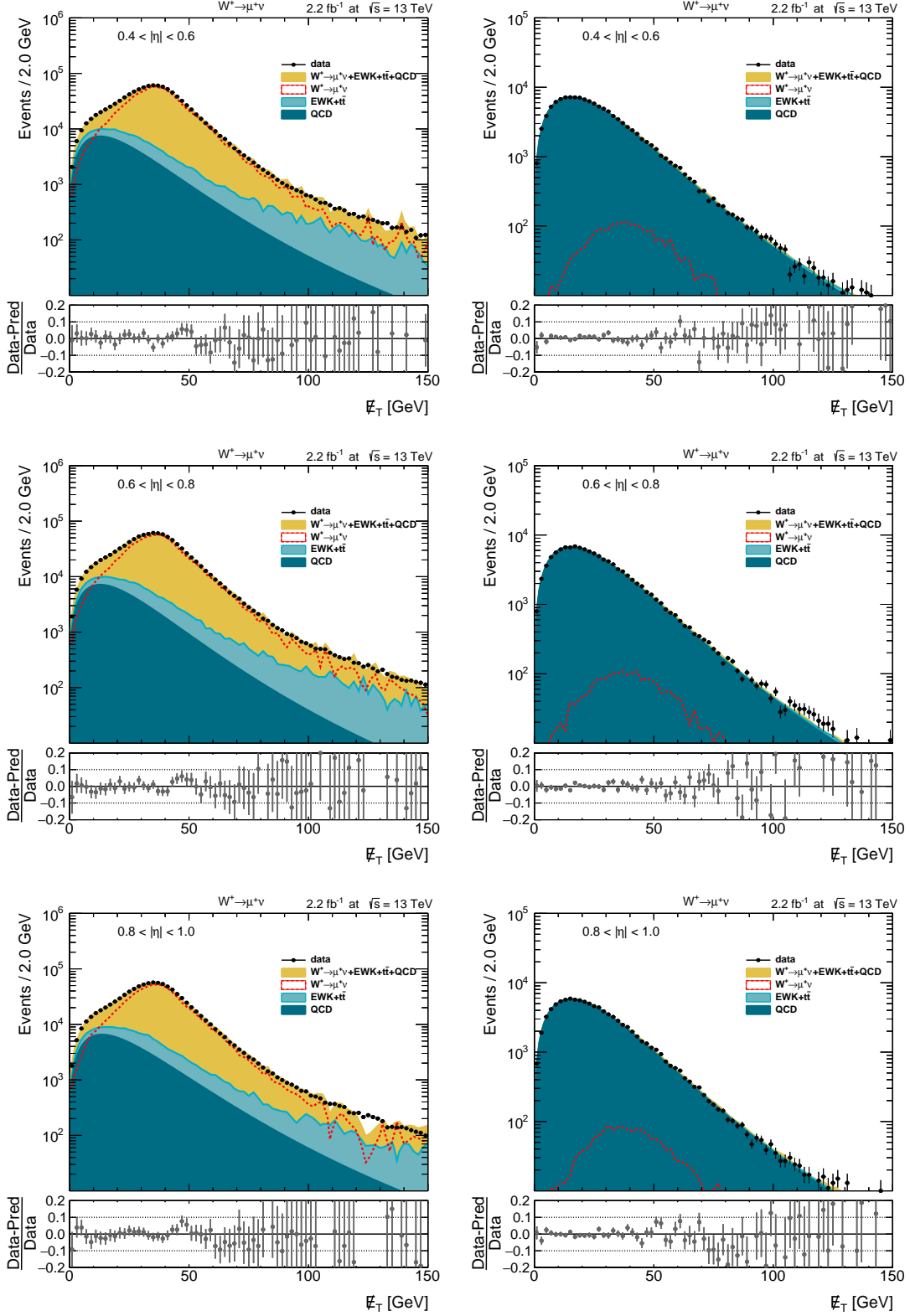


FIGURE B.2: Results of simultaneous fit of the  $E_T^{\text{miss}}$  distribution in signal (left) and control (right) regions for  $W^+ \rightarrow \mu^+ \nu_\mu$ . The upper row corresponds to  $0.4 < |\eta| < 0.6$  region, middle row to  $0.6 < |\eta| < 0.8$ , and the last row to  $0.8 < |\eta| < 1.0$ .

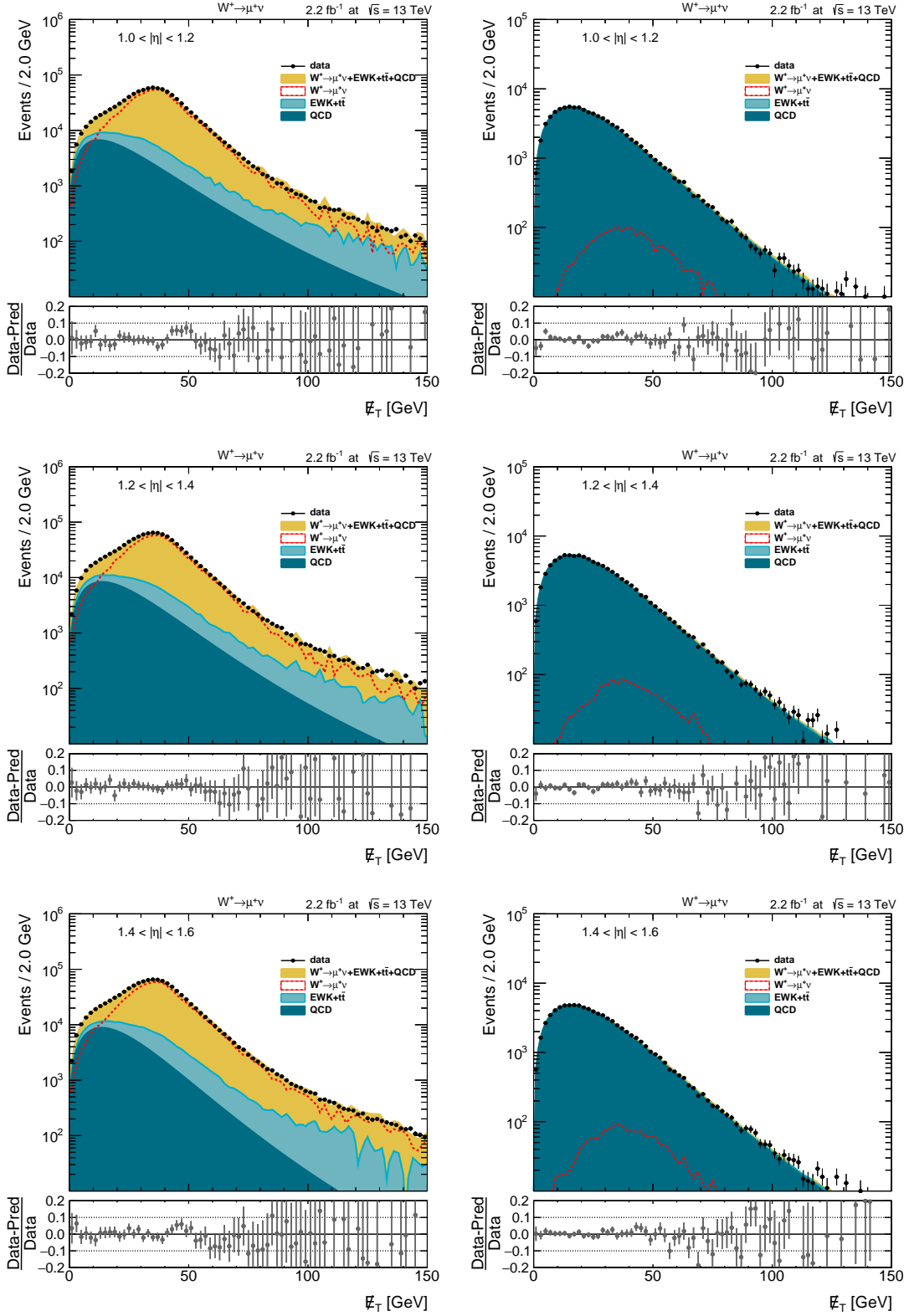


FIGURE B.3: Results of simultaneous fit of the  $E_T^{\text{miss}}$  distribution in signal (left) and control (right) regions for  $W^+ \rightarrow \mu^+ \nu_\mu$ . The upper row corresponds to  $1.0 < |\eta| < 1.2$  region, middle row to  $1.2 < |\eta| < 1.4$ , and the last row to  $1.4 < |\eta| < 1.6$ .

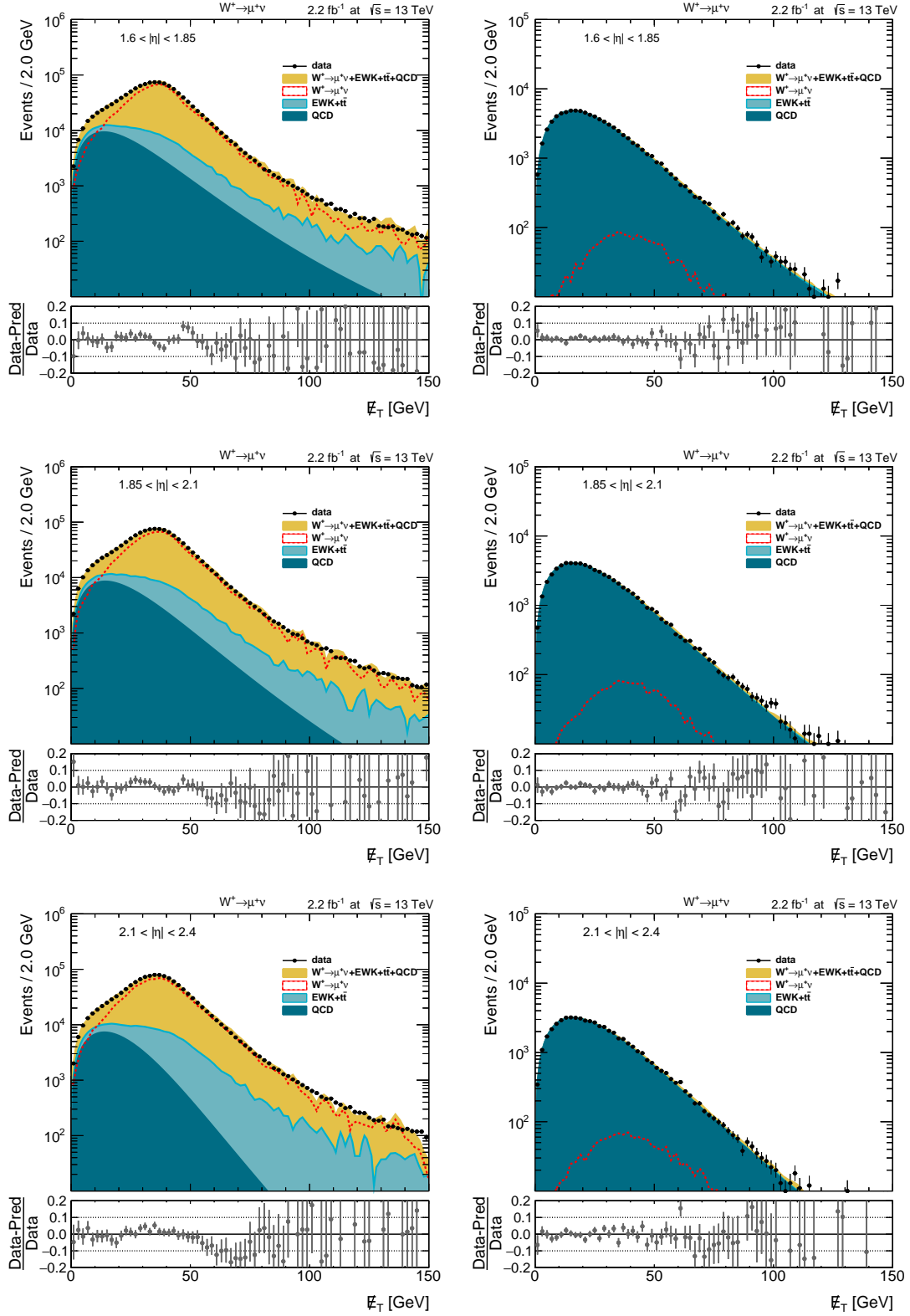


FIGURE B.4: Results of simultaneous fit of the  $E_T^{\text{miss}}$  distribution in signal (left) and control (right) regions for  $W^+ \rightarrow \mu^+ \nu_\mu$ . The upper row corresponds to  $1.6 < |\eta| < 1.85$  region, middle row to  $1.85 < |\eta| < 2.1$ , and the last row to  $2.1 < |\eta| < 2.4$ .

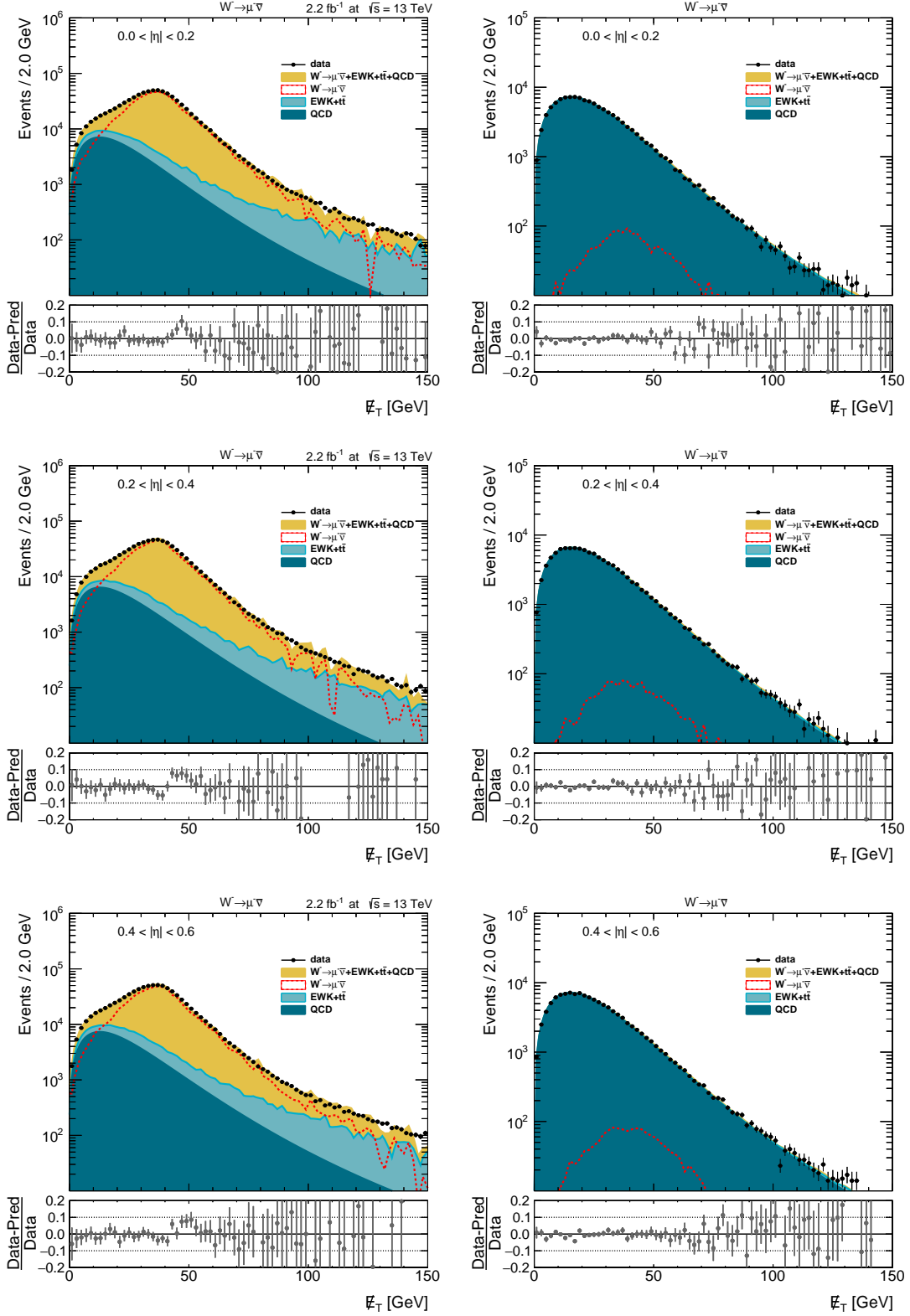


FIGURE B.5: Results of simultaneous fit of the  $E_T^{\text{miss}}$  distribution in signal (left) and control (right) regions for  $W^- \rightarrow \mu^- \bar{\nu}_\mu$ . The upper row corresponds to  $0.0 < |\eta| < 0.2$  region, middle row to  $0.2 < |\eta| < 0.4$ , and the last row to  $0.4 < |\eta| < 0.6$ .

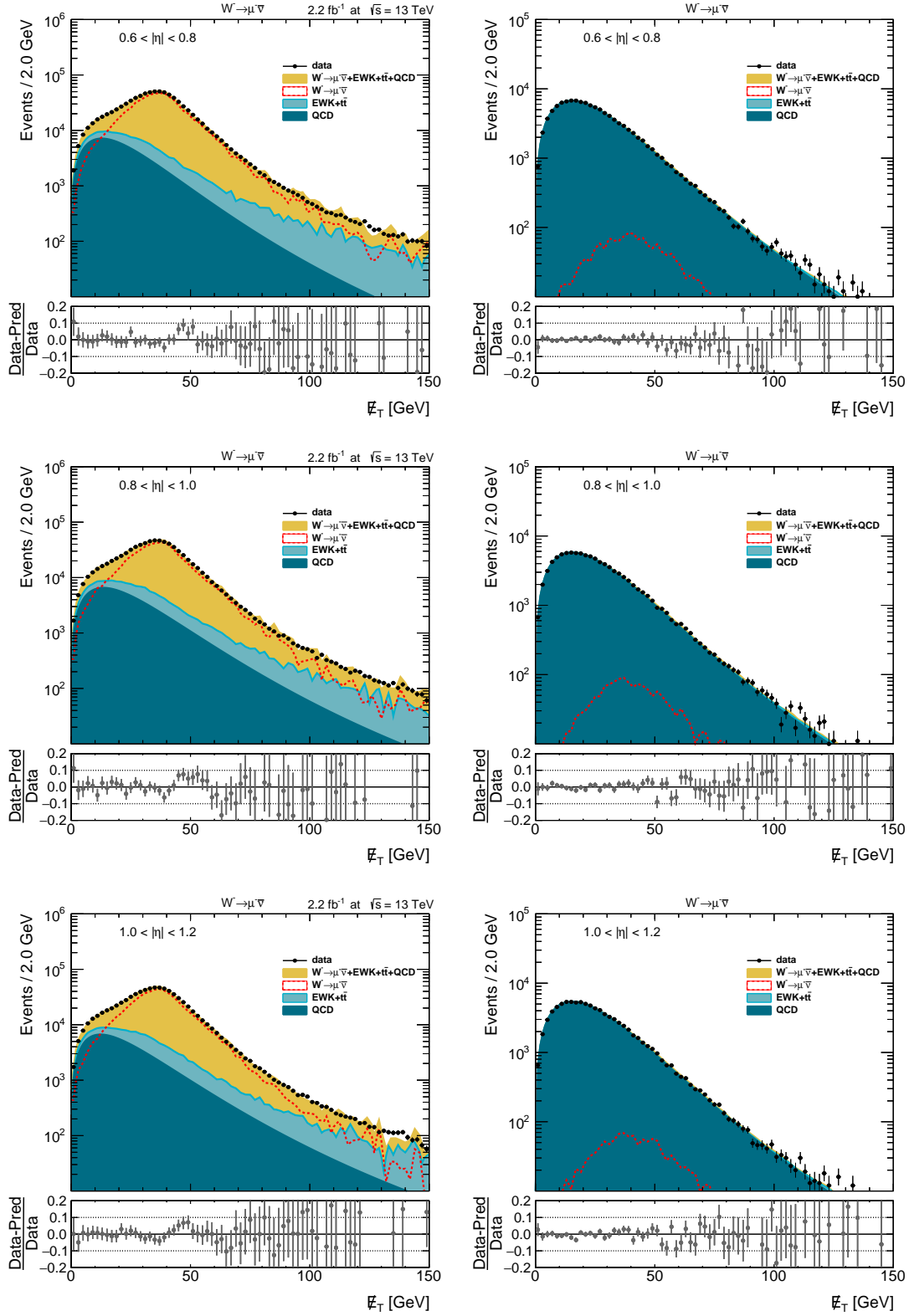


FIGURE B.6: Results of simultaneous fit of the  $E_T^{\text{miss}}$  distribution in signal (left) and control (right) regions for  $W^- \rightarrow \mu^- \bar{\nu}_\mu$ . The upper row corresponds to  $0.6 < |\eta| < 0.8$  region, middle row to  $0.8 < |\eta| < 1.0$ , and the last row to  $1.0 < |\eta| < 1.2$ .

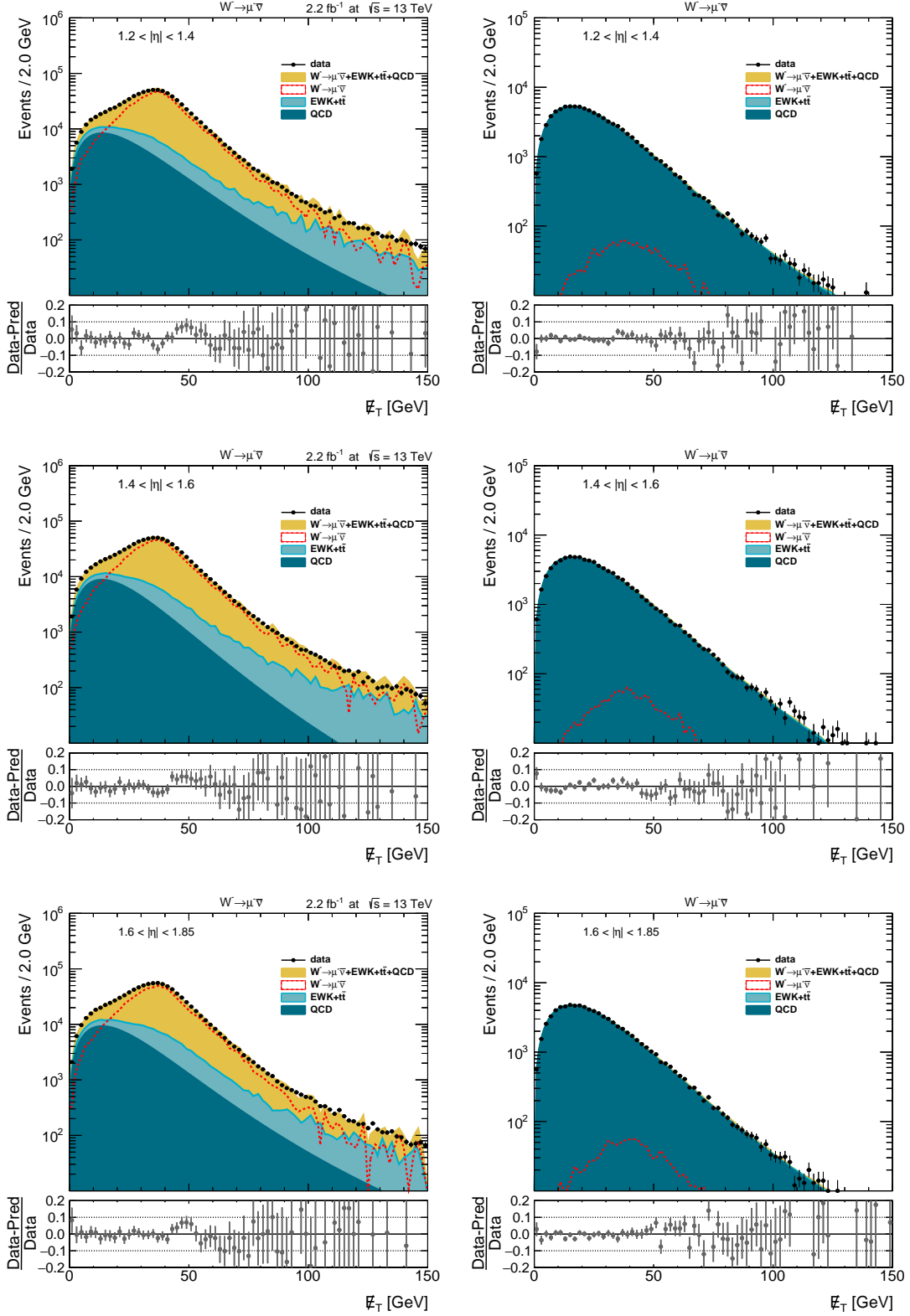


FIGURE B.7: Results of simultaneous fit of the  $E_T^{\text{miss}}$  distribution in signal (left) and control (right) regions for  $W^- \rightarrow \mu^- \bar{\nu}_\mu$ . The upper row corresponds to  $1.2 < |\eta| < 1.4$  region, middle row to  $1.4 < |\eta| < 1.6$ , and the last row to  $1.6 < |\eta| < 1.85$ .

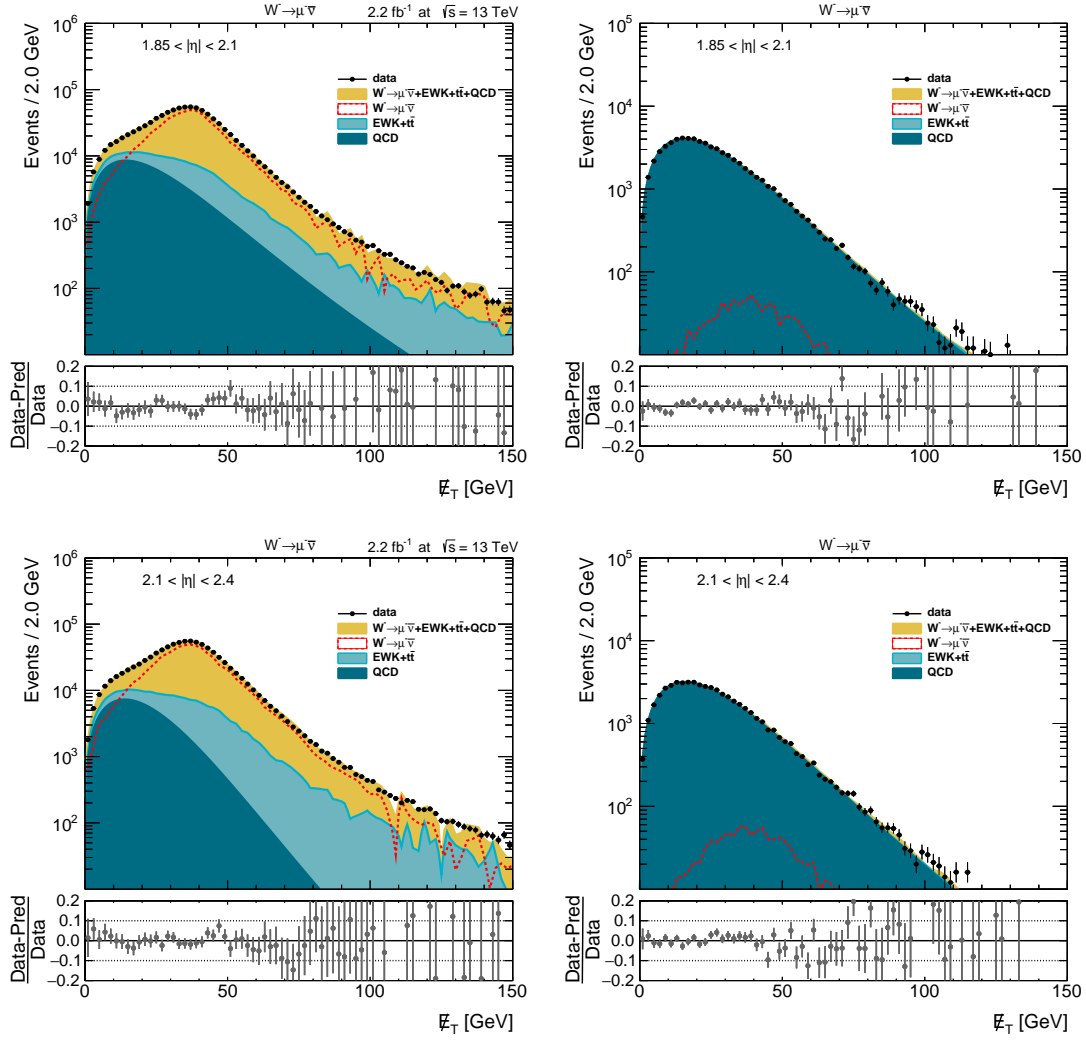


FIGURE B.8: Results of simultaneous fit of the  $E_T^{\text{miss}}$  distribution in signal (left) and control (right) regions for  $W^- \rightarrow \mu^- \bar{\nu}_\mu$ . The upper row corresponds to  $1.85 < |\eta| < 2.1$  region, the last row to  $2.1 < |\eta| < 2.4$ .



## Appendix C

# APPLgrid upgrade

An intense development of theoretical approaches, numerical methods, and computational hardware allow calculations of many different physical processes up to NNLO. Computation of processes from proton collisions (like deep inelastic scattering or proton-proton collisions) depends on errors from PDF uncertainties. The calculation of final state observables using a particular PDF set, and estimation of theoretical uncertainties from a variation of the strong coupling,  $\alpha_s$ , renormalization,  $\mu_r$ , and factorization,  $\mu_f$ , scales becomes very time-consuming. Different approaches to PDF determination raise a necessity to compare cross sections obtained using various PDF sets, which requires recalculation of the whole process using different PDF sets. On average, depending on a process, the CPU time to produce a precise result can take from days to weeks. For that reason, less resource-demanding methods are being developed.

The APPLgrid project [154] emerged as one of proposed solutions and proved to be a reliable algorithm allowing to handle such problems. APPLgrid is library that allow to collect cross-section weights from Monte Carlo generators<sup>1</sup> though additional software interfaces. The weights are stored on a 3D look-up table (in momentum transfer,  $Q$ , and momentum fraction,  $x$ ), which can be afterwards combined with an arbitrary PDF, stored in an APPLgrid format. In this approach, PDF values are stored in a two-dimensional grid of points. PDF-independent cross-section weights representation allows the implementation of renormalization and factorization scale variations. Using the APPLgrid methods, a total cross section is calculated convoluting the PDF values with weights from the look-up table.

The APPLgrid package is widely used not only as a package interfaced into MC tools but also as a part of complex computational frameworks like xFitter [1]. In that context, APPLgrid became an important part of modern numerical tools of particle physics.

This chapter is dedicated to the APPLgrid modification for its upgrade to NNLO. The first part covers a brief introduction to the APPLgrid methodology, while the second is focused on implemented improvements.

### C.1 The APPLgrid project

The APPLgrid project appeared as a result of the collaboration between Tancredi Carli, Dan Clements, Amanda Cooper-Sarkar, Claire Gwenlan, Gavin P. Salam, Frank Siegert, Pavel Starovoitov, and Mark Sutton [143]. Detailed instructions on installation and execution, as well as various produced grids together with convolution and calculation codes, can be found on its official web-page <https://mcfm.fnal.gov>.

---

<sup>1</sup>MCFM [142] or NLOJET++ [155].

### C.1.1 PDF representation on a grid

In the APPLgrid approach, parton density function values for different  $x$  and  $Q^2$  are stored on a two-dimensional grid of points with an  $n$ -th order of interpolation between them. The parton momentum fraction  $x$  and the factorization scale  $Q^2$  are re-written in more convenient terms,  $y(x)$  and  $\tau(Q^2)$  to provide a uniform distribution of points on the grid:

$$y(x) = \ln \frac{1}{x} + a(1-x), \quad (C.1)$$

$$\tau(Q^2) = \ln \left( \ln \frac{Q^2}{\Lambda^2} \right). \quad (C.2)$$

The  $\Lambda$  parameter may be chosen as of the order of  $\Lambda_{\text{QCD}}$ , but not mandatory identical, the parameter  $a$  regulates a uniform distribution of points increasing density of points in the large  $x$  region. Using these notations, the PDF<sup>2</sup>,  $f(x, Q^2)$ , can be re-written in two-dimensional grid ( $i_y \delta y, i_\tau \delta \tau$ ), where  $i_y$  and  $i_\tau$  correspond to PDF values on the grid, and  $\delta y$  with  $\delta \tau$  are the grid spacing, obtained by interpolation:

$$f(x, Q^2) = \sum_{i=0}^n \sum_{j=0}^{n'} f_{h+i, l+j} I_i^{(n)} \left( \frac{y(x)}{\delta y} - h(x) \right) I_j^{(n')} \left( \frac{\tau(Q^2)}{\delta \tau} - l(Q^2) \right). \quad (C.3)$$

Here  $n$  and  $n'$  are the interpolation order,  $I_i^{(n)}(u)$  is interpolation function equal to 1 for case  $u = i$  and for the rest is given by:

$$I_i^{(n)}(u) = \frac{(-1)^{n-i}}{i!(n-i)!} \frac{u(u-1)\dots(u-n)}{u-i}. \quad (C.4)$$

Definitions of  $h(x)$  and  $l(Q^2)$  are given by:

$$h(x) = \text{int} \left( \frac{y(x)}{\delta y} - \frac{n-1}{2} \right), \quad (C.5)$$

$$l(Q^2) = \text{int} \left( \frac{\tau(Q^2)}{\delta \tau} - \frac{n'-1}{2} \right). \quad (C.6)$$

### C.1.2 Weights representation in the case of two incoming hadrons

The APPLgrid method of storing cross-section weights in a look-up table allows the convolution of a chosen PDF set with a given weight, including  $\mu_r$ ,  $\mu_f$ , and  $\alpha_s$  scale variation. Let's consider an NLO MC software that produces  $N$  events  $m = 1 \dots N$  from two incoming hadrons, with a weight  $w_m^b$  of a particular  $b$ -th sub-process, at a scale  $Q^2$ , as a  $Q_m^2$ , and a certain momentum fraction from two protons  $x_1$  and  $x_2$ , as  $x_{1m}$ , and  $x_{2m}$ . Defining  $p_m$  as the number of powers of  $\alpha_s$  in event  $m$ , one would expect the final result of the MC integration for one sub-process to be:

$$W = \sum_{m=1}^N \sum_{b=1}^{n_{\text{sub}}} w_m^{(b)} \left( \frac{\alpha_s(Q_m^2)}{2\pi} \right)^{p_m} F^{(b)}(x_{1m}, x_{2m}, Q^2), \quad (C.7)$$

<sup>2</sup>Les Houches Approach is used as the main PDF format [156], [157].

where  $n_{\text{sub}}$  is a number of sub-processes contributing to the considered event,  $F^{(b)}(x_{1m}, x_{2m}, Q^2)$  are PDF combinations of sub-processes defined later. In the APPLgrid approach, a weight grid  $W_{i_{y1}, i_{y2}, i_{\tau}}^{(p)(b)}$  updates with each event by a portion of the grid with:

$$W_{h+i, l+j}^{(p)(b)} \rightarrow W_{h+i, l+j}^{(p)(b)} + w_m^{(b)} I_i^{(n)(b)} \left( \frac{y(x_m)}{\delta y} - h(x) \right) I_j^{(n')(b)} \left( \frac{\tau(Q_m^2)}{\delta \tau} - l(Q_m^2) \right), \quad (\text{C.8})$$

$i = 0 \dots n, j = 0 \dots n'$ .

A general expression for  $W$ , for any PDF and any  $\alpha_s$  can be expressed by:

$$W = \sum_P \sum_{b=0}^{n_{\text{sub}}} \sum_{i_{y1}} \sum_{i_{y2}} \sum_{i_{\tau}} W_{i_{y1}, i_{y2}, i_{\tau}}^{(p)(b)} \left( \frac{\alpha_s(Q^{2(i_{\tau})})}{2\pi} \right)^P F^{(b)}(x_1^{(i_{y1})}, x_2^{(i_{y2})}, Q^{2(i_{\tau})}). \quad (\text{C.9})$$

Depending on considered processes, combinations of the incoming parton densities can be different. Here, the case of  $W^+$  sub-processes is presented, details on some other cases, as well as more details on the APPLgrid project can be found in appropriate paper [143]. Sub-processes of the  $W$ -boson production are defined from the assumptions of mass-less quarks, using CKM matrix elements. At NLO, six initial states are given by

$$\begin{aligned} \bar{D}U : F^{(0)}(x_1, x_2, Q^2) &= S_{12}(x_1, x_2), \\ U\bar{D} : F^{(1)}(x_1, x_2, Q^2) &= S_{21}(x_1, x_2), \\ \bar{D}g : F^{(2)}(x_1, x_2, Q^2) &= \bar{D}_1(x_1)G_2(x_2), \\ Ug : F^{(3)}(x_1, x_2, Q^2) &= U_1(x_1)G_2(x_2), \\ g\bar{D} : F^{(4)}(x_1, x_2, Q^2) &= G_1(x_1)\bar{D}_2(x_2), \\ gU : F^{(5)}(x_1, x_2, Q^2) &= G_1(x_1)U_2(x_2). \end{aligned} \quad (\text{C.10})$$

Generalized PDFs are defined as:

$$\begin{aligned} G_H(x) &= f_{0/H}(x, Q^2), \\ U_H(x) &= f_{2/H}(x, Q^2)(V_{ud}^2 + V_{us}^2) + f_{4/H}(x, Q^2)(V_{cd}^2 + V_{cs}^2), \\ \bar{D}_H(x) &= f_{-1/H}(x, Q^2)(V_{ud}^2 + V_{cd}^2) + f_{-3/H}(x, Q^2)(V_{us}^2 + V_{cs}^2), \\ S_{12}(x_1, x_2) &= f_{-3/H_1}(x_1, Q^2)f_{2/H_2}(x_2, Q^2)V_{us}^2 + \\ &\quad f_{-3/H_1}(x_1, Q^2)f_{4/H_2}(x_2, Q^2)V_{cs}^2 + \\ &\quad f_{-1/H_1}(x_1, Q^2)f_{2/H_2}(x_2, Q^2)V_{ud}^2 + \\ &\quad f_{-1/H_1}(x_1, Q^2)f_{4/H_2}(x_2, Q^2)V_{cd}^2, \\ S_{21}(x_1, x_2) &= f_{2/H_1}(x_1, Q^2)f_{-3/H_2}(x_2, Q^2)V_{us}^2 + \\ &\quad f_{4/H_1}(x_1, Q^2)f_{-3/H_2}(x_2, Q^2)V_{cs}^2 + \\ &\quad f_{2/H_1}(x_1, Q^2)f_{-1/H_2}(x_2, Q^2)V_{ud}^2 + \\ &\quad f_{4/H_1}(x_1, Q^2)f_{-1/H_2}(x_2, Q^2)V_{cd}^2. \end{aligned}$$

Here  $V_{ij}$  are the CKM matrix elements. The PDF of a particular parton with corresponding PDG Id [158] of incoming  $k$ -th hadron (proton) is given as  $f_{Id/H_k}$ . In that sense, numbers 1...5 correspond to quarks from up to bottom, and anti-quarks with the opposite sign, zero is dedicated to gluon, the contribution from the top quark is neglected. Details on scale dependence as well as re-weighting to a different center of mass can be found in the paper [143].

### C.1.3 A Monte Carlo for femtobarn processes (MCFM)

MCFM-6.8 [159] is a parton-level Monte Carlo program, that produces predictions up to next-to-leading-order for various processes at hadron colliders [160], [161]. The program has a wide range of various setups as input parameters. A user can define different modes of MCFM, producing an output of a specific format. MCFM allows setting heavy quark masses, as well as PDF set, jet definition, event cuts, settings for photon processes, and anomalous couplings of the  $W^\pm$  and  $Z^0$  bosons. The program provides calculation of renormalization and factorization scales on an event-by-event basis, with different possible schemes of scales treatment. Besides that, it has many different setups dedicated to particular processes. More details can be found on the official web-page [154].

Since version 8.0, MCFM can calculate NNLO predictions for color-singlet processes [129]. The NNLO corrections are calculated using the non-local N-jettiness subtraction approach [162]. This upgrade, together with APPLgrid compatibility, makes the latest versions of MCFM (8.3-present) even more attractive in terms of a variety of its usage. Representation of NNLO weights on the APPLgrid grid can find its application on various areas of particle physics, from a great computation time reduction in theory uncertainties calculation, to NNLO weights usage in xFitter [1] PDF fits. For that reason, the APPLgrid upgrade to NNLO weights processing is a crucial task.

### C.1.4 Technical implementation

Studies described in the second half of this chapter are performed for the APPLgrid part, which is linked to the MCFM Monte Carlo program. For that reason, an example of APPLgrid technical implementation is given in the frame of MCFM.

An important element of the APPLgrid is a so-called "mcfm-bridge", in this description, version 0.0.35 is used. It is a package of C++ codes that contains a library, libmcfmbridge.a, with instructions for weights extraction from MCFM and settings for a grid constructor. The bridge also includes a utility called "mcfmbridge-config"; this utility provides options required to link APPLgrid with MCFM. A user access to specify all available parameters for grid production is provided by a C++ file called "mcfm-bridge-0.0.35/src/mcfm\_interface.cxx". The setting list includes:

- A number of grids and the binning information for the grid constructor:
  - lower and upper limits of  $x$  and  $Q^2$ ;
  - number of  $x$  and  $Q^2$  bins;
  - order of interpolation for  $x$  and  $Q^2$ ;

Default values for each process are also available.

- Derived observables from 4-vectors of available final state particles, provided by MCFM, in particular:

- number of observables;
  - binning information (flexibility in binning definition);
  - kinematic cuts.
- A type of PDF decomposition and some other technical details.

When the grid parameters are set and ready, the `mcfm-bridge-0.0.35` package, as well as `MCFM-6.8` has to be recompiled for changes to be applied<sup>3</sup>.

The software implementation of a grid structure is fulfilled as a multidimensional array (for  $x_1$ ,  $x_2$ ,  $Q^2$ ) of the `TH3D`-class of the ROOT analysis framework. Each such three-dimensional grid is stored for each sub-process per  $\alpha_s$  order (see eq. C.9). The weight grid  $W_{h+i,l+j}^{(p)(b)}$  is filled for each cross-section bin during the `MCFM-6.8` run. To reduce a memory usage and to increase a grid readability time, the `MCFM-6.8` is required to run twice. The first run is needed for `APPLgrid` to estimate and allocate a proper amount of required weight space, while the second is to fill the actual grid(s).

The master class for a given cross-section is called `appl::grid` and contains instances of an internal class for each order of  $\alpha_s$ . The internal class `appl::igrd` contains all necessary tools to calculate the cross-section for one particular bin at a specific order. The class also contains *x-to-y*, *y-to-x*, and  $Q^2$ -to- $\tau$ ,  $\tau$ -to- $Q^2$  transformations, and others. A produced grid(s) is saved to a ROOT file with all necessary information required for the cross-section calculation:

- The center of mass energy at which the weight grid has been produced,
- The choice of coordinate transformation functions, with the default definition given by eq. C.1 and C.2,
- The order of interpolation given in eq. C.4,
- The number of grid points to be used in each dimension of  $x_1$ ,  $x_2$ , and  $Q^2$ ,
- The definition of the sub-processes in a 13 x 13 matrix,
- The CKM matrix elements and others.

Calculation of a considered cross section is possible using a convolution procedure, built in a dedicated sub-class of the master class. Members of this sub-class fulfill a convolution procedure over a given  $x_1$ ,  $x_2$ , and  $Q^2$  for each sub-process. The final cross-section value in each bin is then computed by the master class collecting and summing all cross-section values from the sub-classes for each order. To perform a convolution, as an input information, user has to specify a PDF set and  $\alpha_s$ . Example of such a procedure may look as shown in Listing C.1.

---

LISTING C.1: APPLgrid commands to perform convolution.

---

```

1
2 // read the grid
3 appl::grid grid_etal("grid-incljets06-etal.root");
4
5 // perform the convolution
6 std::vector<double> xsec_etal = grid_etal.vconvolute(evolpepdf_ , alphasqcd_);

```

---

<sup>3</sup>This issue is successfully removed by the author, discussed in detail in the second section of this chapter.

Where "grid-incljets06-eta1.root" is a grid file, "vconvolute" is a convolution function, "evolvepdf\_" is a PDF set under consideration, and "alphasqcd\_" is a strong coupling parameter [154]. DGLAP splitting functions are calculated using the HOPPET program [163].

## C.2 MCFM-interface upgrade

The mcfm-bridge upgrade is a necessary part of the APPLgrid modification. Technically, the APPLgrid modification to MCFM NNLO weights processing will require various studies, accuracy checks, etc. natural for any development process. The mcfm-bridge upgrade, discussed in this section, includes modifications that will allow saving time and simplifying the development process. The upgrade was performed in cooperation with Pavel Starovoitov, one of the authors of APPLgrid and xFitter.

### C.2.1 Processes mapping

The first step is dedicated to an improvement of APPLgrid grid default setups for various processes. The list of processes with default grid setups is shown in Table C.1. Each of the processes has recommended grid setup that consists of following parameters:

- Name of the processes,
- Type of a PDF decomposition,
- String label in a grid name ,
- Lower and upper limits of  $Q^2$ ,
- Number of the bins in  $Q^2$  range,
- Order of interpolation for the  $Q^2$  dimension,
- Lowest  $\alpha_s$  order,
- Lower and upper limits of x,
- Number of the bins in x range,
- Order of interpolation for the x dimension,
- Number of loops.

To re-arrange the default grid settings, the table is taken out from mcfm\_interface.cxx into separate files. The parameter list for each process is implemented into a special structure "info", described in a header file "mcfm-bridge-0.0.35/src/mcfm\_procmap.h". The file contains a standard C++ associative container object "map", which returns process parameters through a corresponding key (MCFM process number, see the first column in Tab. C.1). Access to the process parameter can be provided directly through a structure member name. The C++ structure syntax together with the definition of "info" is given in the Listing C.2.

nproc	$f(p_1) + f(p_2) \rightarrow \dots$	Order
1	$W^+(\rightarrow \nu(p_3) + e^+(p_4))$	NNLO
6	$W^-(\rightarrow e^-(p_3) + \bar{\nu}(p_4))$	NNLO
11	$W^+(\rightarrow \nu(p_3) + e^+(p_4)) + f(p_5)$	NLO
13	$W^+(\rightarrow \nu(p_3) + e^+(p_4)) + \bar{c}(p_5)$	NLO
16	$W^-(\rightarrow e^-(p_3) + \bar{\nu}(p_4)) + f(p_5)$	NLO
18	$W^-(\rightarrow e^-(p_3) + \bar{\nu}(p_4)) + c(p_5)$	NLO
31	$Z(\rightarrow e^-(p_3) + e^+(p_4))$	NNLO
41	$Z(\rightarrow e^-(p_3) + e^+(p_4)) + f(p_5)$	NLO
42	$Z_0(\rightarrow 3 \times (\nu(p_3) + \bar{\nu}(p_4))) + f(p_5)$	NLO
43	$Z(\rightarrow b(p_3) + \bar{b}(p_4)) + f(p_5)$	NLO
141	$t(\rightarrow \nu(p_3) + e^+(p_4) + b(p_5)) + \bar{t}(\rightarrow b(p_6) + e^-(p_7) + \bar{\nu}(p_8))$	NLO
142	$t(\rightarrow \nu(p_3) + e^+(p_4) + b(p_5)) + \bar{t}(\rightarrow b(p_6) + e^-(p_7) + \bar{\nu}(p_8))$ [rad.in.dk]	NLO
144	$t(\rightarrow \nu(p_3) + e^+(p_4) + b(p_5)) + \bar{t}(\rightarrow b(p_6) + e^-(p_7) + \bar{\nu}(p_8))$ (uncorr)	NLO
145	$t(\rightarrow \nu(p_3) + e^+(p_4) + b(p_5)) + \bar{t}(\rightarrow b(p_6) + e^-(p_7) + \bar{\nu}(p_8))$ [rad.in.dk],uncorr	NLO
146	$t(\rightarrow \nu(p_3) + e^+(p_4) + b(p_5)) + \bar{t}(\rightarrow b(p_6) + q(p_7) + \bar{q}(p_8))$	NLO
147	$t(\rightarrow \nu(p_3) + e^+(p_4) + b(p_5)) + \bar{t}(\rightarrow b(p_6) + q(p_7) + \bar{q}(p_8))$ [rad.in.top.dk]	NLO
148	$t(\rightarrow \nu(p_3) + e^+(p_4) + b(p_5)) + \bar{t}(\rightarrow b(p_6) + q(p_7) + \bar{q}(p_8))$ [rad.in.W.dk]	NLO
149	$t(\rightarrow q(p_3) + \bar{q}(p_4) + b(p_5)) + \bar{t}(\rightarrow b(p_6) + e^-(p_7) + \bar{\nu}(p_8))$	NLO
150	$t(\rightarrow q(p_3) + \bar{q}(p_4) + b(p_5)) + \bar{t}(\rightarrow b(p_6) + e^-(p_7) + \bar{\nu}(p_8))$ [rad.in.top.dk]	NLO
151	$t(\rightarrow q(p_3) + \bar{q}(p_4) + b(p_5)) + \bar{t}(\rightarrow b(p_6) + e^-(p_7) + \bar{\nu}(p_8))$ [rad.in.W.dk]	NLO
157	$t\bar{t}$ [for total Xsect]	NLO
158	$b\bar{b}$ [for total Xsect]	NLO
159	$c\bar{c}$ [for total Xsect]	NLO
280	$\gamma(p_3) + f(p_4)$	NLO+F
282	$f(p_1) + f(p_2) \rightarrow \gamma(p_3) + f(p_4) + f(p_5)$	LO
283	$f(p_1) + f(p_2) \rightarrow \gamma(p_3) + b(p_4)$	LO
284	$f(p_1) + f(p_2) \rightarrow \gamma(p_3) + c(p_4)$	LO
285	$f(p_1) + f(p_2) \rightarrow \gamma(p_3) + \gamma(p_4)$	NNLO
286	$f(p_1) + f(p_2) \rightarrow \gamma(p_3) + \gamma(p_4) + f(p_5)$	LO

TABLE C.1: Table of MCFM-8.0 processes available in the APPLgrid. The MCFM number of the process is given in the first column. Process channel, as well as numbering of outgoing 4-vectors of particles ( $p_3 - p_8$ ), is shown in the second column. A maximum available computational perturbation order is given in the third column. [rad.in.dk] - processes includes only the corrections in the semileptonic decay of the top quark. (uncorr) - no spin correlations in the decay of the top quarks. [rad.in.top.dk] - include only the radiative corrections in the decay of the top quark without including the radiative corrections in the hadronic decay of the W-boson. [rad.in.W.dk] - includes only the radiative corrections in the hadronic decay of the W-boson coming from the anti-top(top). NLO+F signifies that the calculation can be performed at NLO including the effects of experimental isolation and photon fragmentation.

LISTING C.2: Structure of a grid settings.

---

```

1 // Structure definition in mcfm_procmmap.h
2 struct info{
3     std::string chan;
4     std::string pdf_fun;
5     std::string glab;
6     double q2low;
7     double q2up;
8     int nq2bins;
9     unsigned int qOrder;
10    unsigned int LowestOrder;
11    int nxbins;
12    unsigned int xOrder;
13    double xlow;
14    double xup;
15    unsigned int nloops;
16 };

```

---

## C.2.2 Steering file implementation

An important part of the upgrade was an introduction of the steering file "mcfm-bridge-0.0.35/src/grid\_setup.DAT". The main idea behind this implementation is to simplify user access to the grid settings, avoiding changes in the source file. To do so, a special class "grid\_input" is implemented in the corresponding header and source files in "mcmf-bridge-0.0.35/src". The class allows reading the input steering file grid\_setup.DAT and store grid settings into a class object. Every time the class object is created, the steering file is read and stored in the memory. Each steering file parameter has a personal class member which returns its value.

The grid\_setup.DAT structure is shown in the Listing C.3. The file consists of three parts: grid settings, kinematic cuts, and custom binning. The first section allows specifying already mentioned grid settings.

In the second part, in the first line, a user can choose a number of particles to perform cuts to. The second line represents a format of input specified for kinematic cuts. The "p\_i" is a number of the 4-vector of the final state particle (see second column in Table C.1), each such line is dedicated to a particular 4-vector; "pt\_cut" is a transverse momentum cut (in GeV), and the last two are pseudorapidity cuts. The example given below represents setup for process number 141<sup>4</sup> ( $t\bar{t}$  is decaying in dilepton channel), two 4-vectors of particles 4 ( $e^+(p_4)$ ) and 7 ( $e^-(p_7)$ ) with  $p_T$  cuts of 25 GeV, and  $-\eta$ ,  $+\eta$  cuts of -2.4, 2.4 respectively.

The third section is dedicated to a custom binning of an observable distribution. As in section two, the first line represents a format of custom binning input. The "Obs\_name" is a string name of the given observable, "p\_i" is a number of the 4-vector of the final state particle, which is used for the distribution; "abs" is an absolute value option, "NBins" specifies the number of bins, and the rest are bin edges.

---

<sup>4</sup>The process number is not present in the steering file; it is automatically transmitted from MCFM to the bridge.



LISTING C.3: Example of the steering file "grid\_setup.DAT"

---

```

1
2 default 'Lowest x limit '
3 default 'Highest x limit '
4 40      'x number of bins '
5 6       'Interpolation order of x value '
6 1       'Amount of loops '
7
8 default 'pdf_fun '
9
10 default 'glab '
11 1.0     'Lowest Q^2 limit '
12 1.6e7   'Highest Q^2 limit '
13 15      'Q^2 number of bins '
14 3       'Interpolation order of Q value '
15 default 'lowest order '
16
17 2       'Amount of grids '
18
19
20 ''Kinematic cuts ''
21 2       'Amount of particles with kinematic cuts '
22 '[p_i pt_cut -eta_cut eta_cut]'
23 [4 25 -2.4 2.4]
24 [7 25 -2.4 2.4]
25
26
27 ''Custom binning ''
28 '[Obs_name p_i abs(1 or 0) NBins bin1 bin2 bin3 ...]'
29 [Eta 4 0 10 -2.4 -1.8 -1.5 -0.8 -0.4 0.0 0.4 0.8 1.5 1.8 2.4]
30 [pt 7 1 20 0 2 4 6 8 10 14 18 22 26 30 35 40 50 60 90 120 150 200 300]

```

---

The last two lines show settings of pseudorapidity and transverse momentum distributions for  $e^+(p_4)$  and  $e^-(p_7)$  respectively. The number of requested grids is specified by the variable "Amount of grids", located between the first and the second sections; the number of histograms, defined in the third section, should coincide with this value. The dynamic memory allocation for an arbitrary number of grids is performed using a vector of pointers to an APPLgrid grid; this is an important implementation since it helps to avoid recompilation procedure each time a user wants changes to be applied.

The `grid_input` class object, together with the class members, are used in functions of `mcfm_interface.cxx` code. The interface code consists of three main functions, `book_grid()`, `fill_grid()`, `write_grid()`, and few auxiliary functions (`getObservable()`, `cuts()` and others).

The `book_grid()` function was rewritten in a way that allows to use the information from the steering file, accessing it through the class members. The processes map,

described in section C.2.1, is combined with the steering file through additionally implemented flag "default"<sup>5</sup>. In this case, parameters tagged as "default" are taken from the processes map.

Function `fill_grid()` selects valid weights and fills them in previously created grid using functions `getObservable()` and `cuts()`. Particles, selected for kinematic selection, as well as observables, are determined in `getObservable()` function using the `customKinematics(p, c)` and `customGrid(k, t)` class members. Kinematic selection is performed in function `cuts()`; the function is modified to use `customKinematics(p,c)` and `nKinPar()` members to provide information from the "Kinematic cuts" section of the `grid_setup.DAT`. Selected events are then filled in a grid using `fill_grid()` function.

Additionally, to mentioned improvements, various checks and minor optimizations were also performed.

---

<sup>5</sup>The steering file example (List.C.3) contains variables replaced with the word "default".