

**Measurement of the  $Z(\rightarrow \mu\mu) + \text{Jets}$   
Cross Section at  $\sqrt{s} = 13 \text{ TeV}$  with  
the CMS Experiment and Studies of  
the Proton Structure**

Anna Carlotta Friedel

MASTERARBEIT

An der Fakultät für Physik des  
Karlsruher Institut für Technologie (KIT)  
Institut für Experimentelle Kernphysik (IEKP)

Abgabedatum: 02.05.2017

Referent: Prof. Dr. Günter Quast  
Korreferent: Prof. Dr. Thomas Müller



---

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Theoretical Foundations</b>	<b>5</b>
2.1	Electroweak Theory . . . . .	6
2.2	Quantum Chromodynamics . . . . .	8
2.2.1	Confinement and Asymptotic Freedom . . . . .	8
2.2.2	Parton Distribution Functions . . . . .	9
2.2.3	Cross Section Calculation with the Factorisation Theorem . . . . .	11
2.2.4	Monte Carlo Event Generation and Simulation . . . . .	12
<b>3</b>	<b>Experimental Setup</b>	<b>13</b>
3.1	The CERN and the Large Hadron Collider . . . . .	13
3.2	The CMS Detector . . . . .	14
3.2.1	Coordinates and Transverse Quantities . . . . .	15
3.2.2	Detector Components . . . . .	16
3.2.3	Event Triggering . . . . .	19
3.2.4	Muon Reconstruction . . . . .	20
3.3	Luminosity and Cross Section . . . . .	21
3.4	Software Tools used for Data Analysis . . . . .	21
3.4.1	Artus and Kappa Frameworks . . . . .	21
<b>4</b>	<b>Measurement of the <math>Z(\rightarrow \mu\mu) + \text{jets}</math> Cross Section</b>	<b>23</b>
4.1	Event Selection . . . . .	24
4.2	Muon Triggering and Reconstruction . . . . .	25
4.2.1	Verification of Efficiency Weights . . . . .	27
4.3	Background Estimation . . . . .	29
4.4	Detector Level Comparisons . . . . .	30
4.5	Unfolding . . . . .	34
4.6	Final State Radiation Studies . . . . .	37
4.7	Experimental Uncertainties . . . . .	40

<b>5</b>	<b>Theory Predictions for the <math>Z(\rightarrow \mu\mu) + \text{jets}</math> Cross Section</b>	<b>43</b>
5.1	Computation Workflow . . . . .	43
5.1.1	fastNLO . . . . .	44
5.2	Sensitivity of the Theory Predictions to the PDFs . . . . .	44
5.3	Theory Uncertainties . . . . .	45
<b>6</b>	<b>Comparison of Measurement and Theory Predictions and PDF Determination</b>	<b>49</b>
6.1	Comparison with NLO Predictions . . . . .	49
6.2	PDF Determination . . . . .	54
6.2.1	Estimation of Uncertainties on the PDFs . . . . .	55
6.2.2	Results of the PDF Fits . . . . .	56
<b>7</b>	<b>Conclusion</b>	<b>63</b>
<b>A</b>	<b>Appendix</b>	<b>65</b>
	<b>List of Figures</b>	<b>79</b>
	<b>List of Tables</b>	<b>81</b>
	<b>Bibliography</b>	<b>83</b>

# Introduction

The Standard Model of particle physics provides a consistent and well-proven description of most observed phenomena, although there are still unresolved questions. These range from the precise knowledge about the values of certain free parameters to the nature of gravity, which is not explained in the Standard Model.

In the search for answers, increasingly more powerful particle colliders have been built. The currently most advanced one is the Large Hadron Collider (LHC). To reach as high center-of-mass energies as the LHC does, heavy particles like protons were considered to be the most feasible objects to collide. The challenge of proton-proton collisions is their more complex event topology compared to, for example, electron-positron scattering. One reason for this is that protons themselves are composed of many constituents, referred to as partons.

The structure of the proton is described by the parton distribution functions (PDFs). The PDFs have to be determined experimentally, since they are not predicted by the corresponding theory, the quantum chromodynamics (QCD).

The uncertainty on the PDFs has to be taken into account for each theory prediction calculated for proton-proton collisions. Those predictions themselves will become more precise in the future since physicists develop more complex calculations, enabled by the increasing computing power. Hence, the uncertainties on the PDFs are expected to be even more relevant in the future.

On the other hand, the sensitivity of proton-proton collisions on the PDFs allows probing the current PDF sets and contributing to PDF constraints using data from the CMS experiment at LHC.

For this purpose, event classes with small statistical and systematic uncertainties on the measurement are suitable. One of these processes is the Z boson production in its leptonic decay channels, on which this work is focussed.

In this thesis,  $Z \rightarrow \mu^+ \mu^-$  cross section measurements are performed and PDF fits are presented.

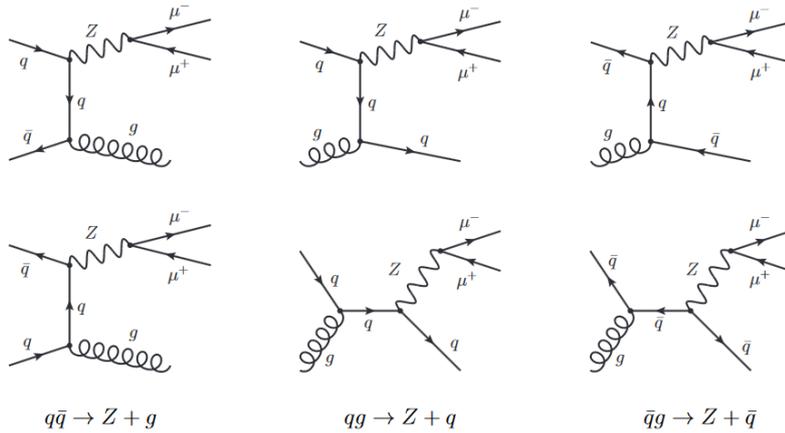
In Chapter 2, the theoretical foundations for the understanding of the studied process and the PDF determination are outlined. The CMS experiment and the software

framework used for this analysis are introduced in Chapter 3. In Chapter 4, the measurement of the double-differential  $Z(\rightarrow \mu\mu) + \text{jets}$  cross section is presented. The event selection, the background contributions and corrections for detector effects, including an unfolding procedure, are described in this context. In addition, the experimental uncertainties are estimated. The workflow used to obtain theory predictions for the  $Z(\rightarrow \mu\mu) + \text{jets}$  cross sections is presented in Chapter 5. In Chapter 6, measurement and theory predictions are combined to perform a fit of the PDFs.

## Theoretical Foundations

The Standard Model of particle physics describes all known elementary particles and their interactions except for gravity. These interactions include the *strong force* as well as the *electroweak force*. Each interaction is mediated by a *gauge boson* while everything that we call matter consists of *fermions*. The latter are divided into two different types. Fermions interacting via the strong force are called *quarks*, and the remaining fermions are termed *leptons*.

The process analysed in this thesis and shown in Figure 2.1 consists of two components. The Z boson production is an electroweak process since the Z boson is a gauge boson of the electroweak force. The gauge boson of the strong force is referred to as *gluon*. Hence, the quark or gluon radiation is described in the quantum field theory of the strong force called *quantum chromodynamics*. The momenta of quarks and gluons in the initial state of the process are determined by the, mainly strong, interactions inside the colliding protons and described by the *parton distribution functions* (PDFs). This chapter is dedicated to these aspects of the Standard Model.



**Figure 2.1:** Leading order Feynman diagrams for  $Z(\rightarrow \mu\mu) + \text{jets}$  events. Taken from [1]

## 2.1 Electroweak Theory

The electroweak theory was established by Glashow, Salam and Weinberg [2–4] as a unification of the theories of electromagnetism and the weak interaction. Mathematically, the electroweak symmetry is defined as an  $SU(2) \times U(1)$  gauge group.

The gauge bosons corresponding to the  $SU(2)$  subgroup are named  $W^1, W^2$  and  $W^3$ . The physical  $W^\pm$  bosons are the creation and annihilation operators:

$$W^\pm = \frac{1}{\sqrt{2}}(W^1 \mp iW^2) \quad (2.1)$$

The charge associated with this symmetry group is the *weak isospin*  $T$ . Fermions are split into a left-handed doublet (negative chirality,  $T = 1/2$ ) and a right-handed singlet (positive chirality,  $T = 0$ ). The  $W^\pm$  and  $W^3$  bosons form an isospin triplet with  $T = 1$ . Because of the isospin conservation, only couplings between left-handed fermions and  $W$  bosons are possible, which explains the parity violation observed in weak processes.

The charge corresponding to the  $U(1)$  subgroup is the *weak hypercharge*  $Y$  which is a combination of the weak isospin and the electrical charge  $Q$ :

$$Y = 2(Q - T_3) \quad (2.2)$$

The gauge boson is the  $B^0$ . The Lagrangian is given by:

$$\mathcal{L} = \bar{L}\not{D}L + \bar{R}\not{D}R - \frac{1}{4}W_a^{\mu\nu}W_{\mu\nu}^a - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} \quad (2.3)$$

Where  $L$  and  $R$  are the left- and right-handed fermion fields.  $B_{\mu\nu}$  and  $W_{\mu\nu}^a$  are the field strength tensors for the isospin and hypercharge fields.  $D$  is the covariant derivative of the electroweak theory. The Feynman slash notation was used in the equation.

This Lagrangian represents the electroweak symmetry without considering the spontaneous symmetry breaking that is described in the Brout-Englert-Higgs mechanism [5]. The symmetry breaking changes the Lagrangian and includes a kinetic part, that predicts a mass for the gauge bosons

$$\mathcal{L}_{\text{kin}} = \frac{g^2 v^2}{4} W_\mu^+ W^{\mu,-} + (B^0, W^3) \frac{v^2}{8} \begin{pmatrix} g'^2 & -gg' \\ -gg' & g^2 \end{pmatrix} \begin{pmatrix} B^0 \\ W^3 \end{pmatrix} \quad (2.4)$$

where  $g$  and  $g'$  represent the coupling constants for the isospin and hypercharge, respectively, and  $v$  is the vacuum expectation value of the Higgs boson.

The mass eigenstates and eigenvalues of the second term are the photon  $\gamma$  and the  $Z$  boson. They can be obtained by diagonalising the matrix:

$$\begin{pmatrix} B^0 \\ W^3 \end{pmatrix} = \begin{pmatrix} \cos(\theta_W) & -\sin(\theta_W) \\ \sin(\theta_W) & \cos(\theta_W) \end{pmatrix} \begin{pmatrix} \gamma \\ Z \end{pmatrix} \quad (2.5)$$

Here, the Weinberg angle was introduced:  $\cos(\theta_W) \equiv \frac{g}{\sqrt{g^2+g'^2}}$ .

This leads to the gauge boson masses:

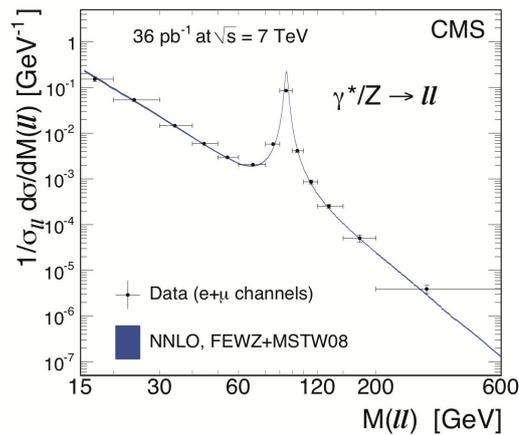
$$m_W = \frac{g^2 v^2}{4} \quad (2.6)$$

$$m_Z = \frac{g^2 v^2}{4 \cos^2(\theta_W)} \quad (2.7)$$

$$m_\gamma = 0 \quad (2.8)$$

The value of the Weinberg angle as well as W and Z masses are free parameters of the Standard Model and determined experimentally.

The annihilation of a quark antiquark pair and lepton pair production mediated by an interference of Z boson and photon is called *Drell-Yan* [6] process. The invariant mass spectrum of lepton pairs is characterised by a peak around the Z boson mass (see Figure 2.2). In the proximity of the Z peak, the contribution of the, highly virtual, photons  $\gamma^*$  to Drell-Yan processes is small and therefore often, also in this thesis, Z is written instead of  $Z/\gamma^*$  in the description of Drell-Yan processes.



**Figure 2.2:** Theory predictions and measurement of the lepton pair invariant mass spectrum observed in Drell-Yan processes. The horizontal error bar indicates the bin size. A peak at the Z mass is visible. Taken from [7].

## 2.2 Quantum Chromodynamics

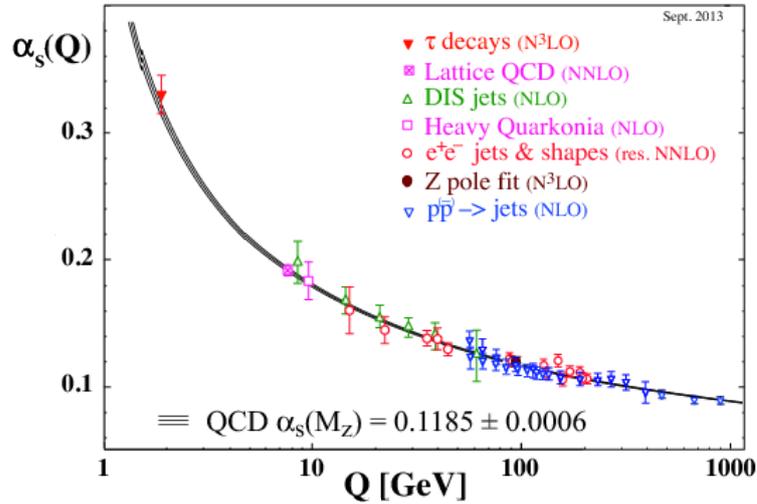
Like the electroweak theory, quantum chromodynamics (QCD) is a quantum field theory. Since there are three different types of charge in the QCD, referred to as *colours*, the underlying symmetry group is the  $SU(3)$ . The Lagrangian is given by:

$$\mathcal{L} = \bar{\psi}_i (i(\not{D})_{ij} - m \delta_{ij}) \psi_j - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu} \quad (2.9)$$

In this formula,  $D$  is the covariant derivative of the QCD,  $\psi_{i,j}$  are the quark fields and  $G_{\mu\nu}^a$  is the gluon field strength tensor. Therefore, the first part of the Lagrangian describes the quark-gluon interactions while the second term corresponds to the self-coupling of gluons, which is a fundamental property of the QCD.

### 2.2.1 Confinement and Asymptotic Freedom

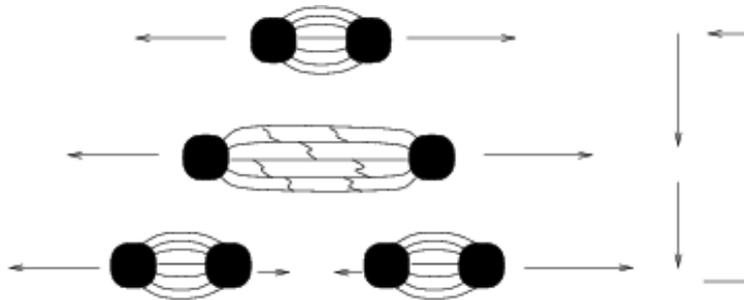
With increasing distance between colour charged particles, the self-interaction of the massless gluons leads to an increasing coupling constant  $\alpha_s$ . Hence, colour charged particles are always combined to colour neutral combinations at the femtometre distance scale. This property of the strong interaction, called *confinement*, is responsible for the binding of the quarks inside the proton. The running of  $\alpha_s$  as a function of the energy scale of the interaction  $Q$  is sketched in Figure 2.3.



**Figure 2.3:** Running of the strong coupling constant  $\alpha_s$  as a function of the energy scale  $Q$  of the corresponding interaction. The energy scale is reciprocal to the distance of the interacting particles. Taken from [8].

For collider experiments, the confinement of quarks and gluons implies, that it is impossible to detect these particles isolated. Instead, if strongly interacting particles are separated from each other, quark-antiquark pairs are produced from the vacuum. This process, accompanied by subsequent gluon emission, is called *parton shower* and shown in Figure 2.4. In the end, stable hadrons are formed in a process termed *hadronisation*. Hence, a quark or gluon in the final state of a process leads to a cone-shaped bunch of particles, which is called a *jet*.

In contrast, at the sub-femtometre scale, which corresponds to an energies above approximately 200 MeV, the strong coupling constant becomes small enough to be used in perturbative calculations. This short distance behaviour is called *asymptotic freedom* and the related theory tool is the *perturbative QCD* (pQCD).



**Figure 2.4:** Parton splitting that occurs if quarks or gluons are separated from each other. Taken from [9].

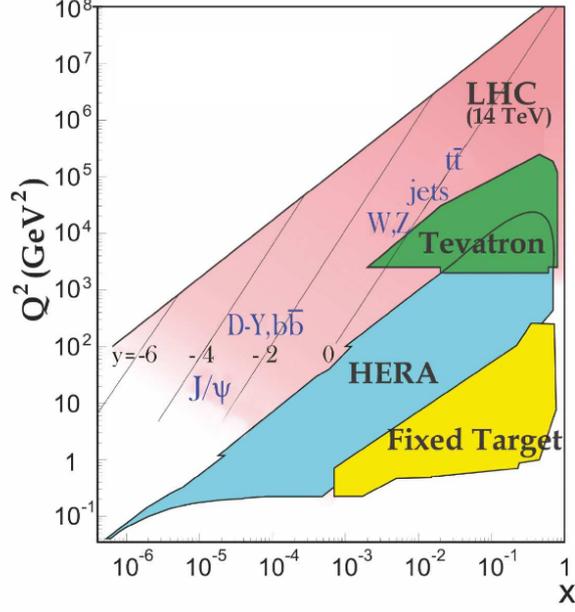
### 2.2.2 Parton Distribution Functions

Besides the valence quarks of the proton, which define its quantum numbers, a high number of gluons and quark-antiquark pairs, named *sea quarks*, exist inside the proton. They originate from the interactions between the valence quarks.

Therefore, in a proton-proton collision, those gluons and sea quarks, as well as the valence quarks, carry a fraction of the proton momentum and take part in interactions between the protons. All these proton constituents are summarised as *partons*.

For this reason, the knowledge of the probability density function of that momentum fraction  $x$  at a given energy scale  $Q$  for each parton is essential for the theoretical predictions of processes occurring at the LHC experiments. The probability density at a given energy scale is described by the parton distribution functions (PDFs).

Since the interactions inside the proton correspond to energies below the perturbative QCD, the PDFs have to be determined experimentally. For this purpose, fitting procedures are performed by various collaborations using different datasets. The most relevant experiments are shown in Figure 2.5. Thus, PDF fits were mainly



**Figure 2.5:** Experiments providing the data for PDF sets and their covered phase space in  $x$  and  $Q^2$ . The main contribution comes from DIS measurements at the HERA collider and fixed-target experiments. Taken from [10]

done with data from *deep inelastic scattering* (DIS) measurements at HERA [11] and from fixed-target experiments. A minor contribution comes from experiments at the Tevatron proton-antiproton collider [12]. Since a few years, the LHC experiments are working on further PDF constraints.

Examples for collaborations that develop PDF sets are CTEQ [13], NNPDF [14], MMHT [15], ABM [16] and HERAPDF [17]. An overview can be found in the LHAPDF library [18].

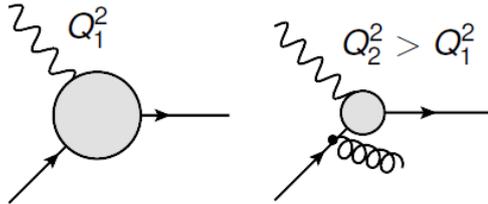
### DGLAP Equations

Given the PDFs at an energy scale  $Q_0^2$ , they can be calculated for another  $Q^2 > Q_0^2$  using the DGLAP [19–21] equations. In leading order, they can be expressed as

$$Q^2 \frac{\partial}{\partial Q^2} \begin{pmatrix} f_i(x, Q^2) \\ \bar{f}_i(x, Q^2) \\ g(x, Q^2) \end{pmatrix} = \frac{\alpha_s(Q^2)}{2\pi} \sum_j \int_x^1 \frac{d\xi}{\xi} \begin{pmatrix} P_{q_i q_j}(x/\xi) & 0 & P_{q_i g}(x/\xi) \\ 0 & P_{\bar{q}_i \bar{q}_j}(x/\xi) & P_{\bar{q}_i g}(x/\xi) \\ P_{g q_j}(x/\xi) & P_{g \bar{q}_j}(x/\xi) & P_{g g}(x/\xi) \end{pmatrix} \begin{pmatrix} f_j(\xi, Q^2) \\ \bar{f}_j(\xi, Q^2) \\ g(\xi, Q^2) \end{pmatrix} \quad (2.10)$$

where  $f_i(x, Q^2)$  and  $\bar{f}_i(x, Q^2)$  refer to the PDFs with (anti-)quark flavour  $i$  and  $g(x, Q^2)$  to the gluon PDF.

$P_{AB}(x/\xi)$  are named splitting functions and can be calculated perturbatively. They describe the probability of a quark or gluon with momentum fraction  $x$  to originate from another parton with higher momentum fraction  $\xi$ . The knowledge of those probabilities is necessary for the evolution to another energy scale because the  $Q$ -dependence of the PDFs is caused by the increasing vertex resolution at higher energies, as illustrated in Figure 2.6.



**Figure 2.6:** Schematic diagram of the vertex resolution at different energy scales.

### 2.2.3 Cross Section Calculation with the Factorisation Theorem

In this section, the cross section calculation for the production of an arbitrary final state  $X$  in proton-proton collisions is described. The factorisation theorem of QCD allows separating the computation into two parts: A short-range part  $\hat{\sigma}_{ij \rightarrow X}$ , corresponding to the hard process, that can be calculated within perturbative QCD and a part, that describes long-range effects with non-perturbative but universal functions. In the case of proton-proton collisions, these functions are the proton PDFs  $f_{i,j}$ .

$$\sigma_{pp \rightarrow X}(\mu_r, \mu_f) = \sum_{i,j} \sum_{n=0}^N \int dx_1 dx_2 f_i(x_1, \mu_f) f_j(x_2, \mu_f) \times \hat{\sigma}_{ij \rightarrow X}^n(x_1, x_2, \mu_r, \mu_f, \alpha_s(\mu_r)) \quad (2.11)$$

The sum runs over all parton flavours  $i$  and  $j$ , while  $N$  is the perturbative order.

The integral covers the momentum fractions  $x_{1,2}$  of the interacting partons while  $\mu_f$  and  $\mu_r$  are the *factorisation* and *renormalisation scales*. Emissions with  $p_T$  below  $\mu_f$  are included in the PDFs. The renormalisation scale belongs to the renormalisation theory [22] introduced to eliminate divergences in higher-order corrections. Usually,  $\mu_f$  and  $\mu_r$  are set to the energy scale of the hard process. For  $N \rightarrow \infty$ , the total cross section is independent of  $\mu_f$  and  $\mu_r$ . However, the perturbative series practically is truncated, leading to a dependence on the scale choice. Typical orders for cross section calculations are the leading order (LO), corresponding to  $N = 0$  and the next-to-leading order (NLO) with  $N = 1$ .

The integral is usually solved with Monte Carlo techniques. The hard process cross section  $\hat{\sigma}_{ij \rightarrow X}^n$  is closely linked to the *matrix element* [23] of the process.

### 2.2.4 Monte Carlo Event Generation and Simulation

Many programs are established for perturbative cross section calculations predictions. For the simulation of an event as it occurs at the LHC, generally referred to as *Monte Carlo event*, steps beyond the fixed-order computation are necessary. These steps are the simulation of parton showering and subsequent hadronisation of the colour charged partons (see Section 2.2.1) as well as the simulation of the *underlying event*. The underlying event summarises secondary interactions of the beam remnants and multi-parton interactions during the proton-proton collision.

The fixed-order theory predictions can be interfaced to programs performing the additional simulations, e.g. PYTHIA 8 [24]. However, it has to be taken care not to double count higher-order contributions from fixed-order predictions that are present already in the parton shower approach.

As a last step, after decaying unstable particles (decay length below 2 cm), the interactions of the remaining final state particles with the detector material can be simulated in detail using, for example, GEANT 4 [25].

## Experimental Setup

This chapter describes the CMS experiment at the Large Hadron Collider (LHC) as well as the software tools used for data analysis.

### 3.1 The CERN and the Large Hadron Collider

Today, the LHC is the largest and most powerful particle accelerator in the world. It is operated by the European Organization for Nuclear Research, better known as CERN. It derives its name from Conseil Européen pour la Recherche Nucléaire, which was the council commissioned to found the international research facility.

CERN currently has 22 member states and about 2400 employees. In addition, about 10000 scientists from about 600 universities and institutes all over the world are working at the CERN experiments<sup>1</sup> [26].

The LHC is in operation since 2009 and used for proton-proton collisions as well as for heavy ion physics. In the proton mode, the centre-of-mass-energy at the beginning was 7 TeV. In 2012, it was increased to 8 TeV. After an expansion phase from 2013 until 2015, the LHC was relaunched with 13 TeV centre-of-mass-energy.

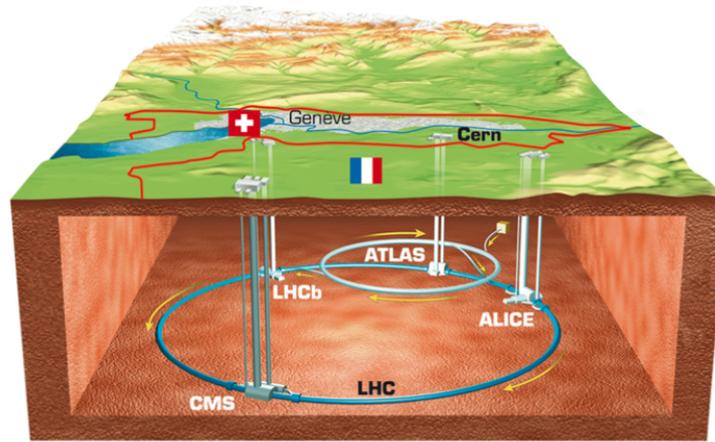
A chain of pre-accelerators is needed before the protons are injected into the LHC. The acceleration process starts in linear colliders and continues with the predecessors of the LHC called Proton Synchrotron (PS) and Super Proton Synchrotron (SPS). The LHC is built inside the tunnel that was once used for the Large Electron-Positron Collider (LEP). It has a circumference of 27 km and is located at the France-Switzerland border. An overview of the LHC is shown in Figure 3.1.

Since the LHC is a synchrotron accelerator, the protons (or lead ions) are circulating in well-defined bunches. They are accelerated by a radio frequency cavity with each revolution and kept on the desired orbit by about 10000 superconducting magnets [28].

The four major experiments located at the LHC are called ALICE [29], ATLAS [30], CMS [31] and LHCb [32]. ALICE is mainly a heavy ion experiment with the aim to

---

<sup>1</sup>state: 2013



**Figure 3.1:** The location of the LHC and the four main experiments. Also shown (around ATLAS): The SPS, which is the final pre-accelerator [27].

study the properties of the quark-gluon plasma, a state that is assumed to be similar to the very beginning of the universe.

LHCb is arranged to measure especially hadrons with bottom and charm quarks and examines CP violation as well as rare decay channels.

ATLAS and CMS are competing multi-purpose experiments designed to cover a wide range of particle physics analyses. Collaborations that independently operate separate experiments are necessary to be able to mutually support a possible discovery, as done in 2012, where the detection of a Higgs-like boson was announced by both collaborations separately.

Since the data used in this thesis are provided by the CMS collaboration, the CMS experiment is described in more detail in the next section.

## 3.2 The CMS Detector

The name CMS stands for Compact Muon Solenoid and outlines the features of the detector:

- **Compact:** CMS is smaller but heavier than the ATLAS detector.
- **Muon:** In the design of the CMS detector, a focus was laid on the muon detection, since it was expected to find the Higgs boson in its decay channels with muons in the final state.

- **Solenoid:** Most detector components are located inside a superconducting solenoid magnet. The field strength is up to 4 T - twice as much as reached inside the ATLAS detector.

In the following sections, the coordinate system and the detector components are explained briefly.

### 3.2.1 Coordinates and Transverse Quantities

The CMS collaboration uses a right-handed coordinate system, which originates at the nominal collision point inside the CMS detector. The x-axis points roughly towards the centre of the LHC, the y-axis points vertically upwards and the z-axis points west along one of the beams. For most purposes, a cylindrical coordinate system is used with the azimuthal angle  $\varphi$  in the x-y plane with  $\varphi = 0$  pointing along the x-axis and the polar angle  $\theta$  in the y-z plane with  $\theta = 0$  pointing along the z-axis.

Instead of the polar angle  $\theta$ , usually the pseudorapidity  $\eta$  and rapidity  $y$  are used:

$$\eta = -\ln \left[ \tan \left( \frac{\theta}{2} \right) \right] = \frac{1}{2} \cdot \ln \left( \frac{p + p_z}{p - p_z} \right) \quad (3.1)$$

$$y = \frac{1}{2} \cdot \ln \left( \frac{E + p_z}{E - p_z} \right) \quad (3.2)$$

Here,  $E$  is the energy of the particle and  $p_z$  the momentum along the z-axis.

In the limit of massless particles, i.e. in the highly relativistic limit,  $y$  and  $\eta$  become identical.

The choice of rapidity instead of the polar angle offers several benefits. At first, compared to the  $\theta$ , the particle production is flat as a function of  $y$ . In addition, differences in rapidity are Lorentz invariant under boosts along the z-axis.

The initial state momentum along the beam axis is unknown in hadron collider experiments since the momentum fraction  $x$  of the interacting partons is not known (see Section 2.2.2). For this reason, in many cases, only the quantities perpendicular to the beam axis are analysed. These observables are denoted with an index "T" for transverse. For example,  $p_T$  represents the transverse momentum.

The  $p_T$  of the colliding partons is often approximated to be zero since it is small compared to the energy of the collision. With this assumption, a resulting  $p_T$  after a collision, denoted as  $E_T^{\text{miss}}$ , shows either one or more undetected particles (neutrinos) or a mismeasurement in the  $p_T$  of at least one detected object.

#### The $\phi_\eta^*$ Observable

In addition to transverse momentum and rapidity, a variable called  $\phi_\eta^*$  is studied in this thesis. It is defined for event topologies with two leptons that are balanced by a jet.

The definition of  $\phi_\eta^*$  consists of two parts. The first part depends on the azimuthal opening angle  $\Delta\phi$  between the leptons. The topology is illustrated in Figure 3.2.

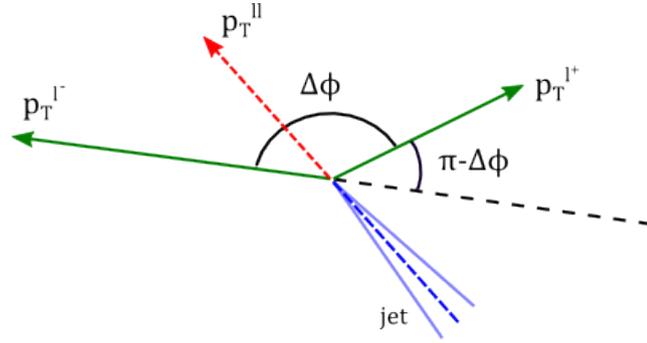
The second part is related to a Lorentz boost along the beam direction into the system where both leptons are back-to-back. The pseudorapidity difference of the positively and negatively charged leptons  $\eta^{l^\pm}$  defines the scattering angle of the leptons with respect to the beam direction in this frame  $\theta^*$  [33]:

$$\cos(\theta^*) = \tanh\left(\frac{\eta^{l^-} - \eta^{l^+}}{2}\right) \quad (3.3)$$

Finally,  $\phi_\eta^*$  is defined as:

$$\phi_\eta^* = \tan\left(\frac{\pi - \Delta\phi}{2}\right) \cdot \sin(\theta^*) \quad (3.4)$$

It depends only on angular parameters and can therefore be measured more accurately than momentum dependent observables. The  $\phi_\eta^*$  observable is a measure for the deviation from the back-to-back topology of the leptons. Hence, it is correlated to the ratio of the lepton pair's transverse momentum to its mass  $p_T^{ll}/m^{ll}$  [34]. If only a small mass range is analysed, as done in this thesis,  $\phi_\eta^*$  is expected to show a behaviour similar to  $p_T^{ll}$ .

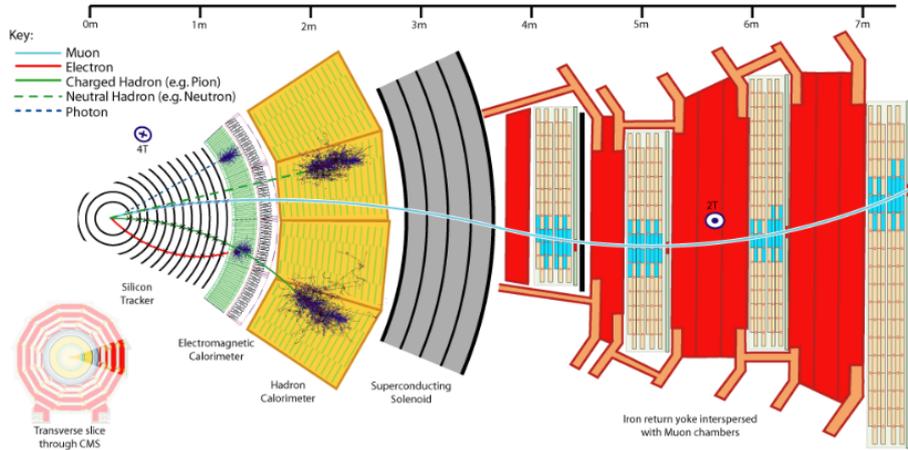


**Figure 3.2:** Illustration of the event topology and the definition of the azimuthal opening angle  $\Delta\phi$ . Two oppositely charged leptons with transverse momentum  $p_T^{l^\pm}$  are balanced by a jet. The vectorial addition of  $p_T^{l^\pm}$  yield the lepton pairs transverse momentum  $p_T^{ll}$ .

### 3.2.2 Detector Components

The CMS detector has a cylindrical structure containing several layers of subdetectors. This includes a tracker, electromagnetic and hadronic calorimeters as well as a muon

detection system. The determination of different particle types using combined information from the subdetectors is shown in Figure 3.3.



**Figure 3.3:** A slice through the CMS detector shows the identification of different particle types from the combination of signals in the subdetectors. Taken from [35]

The CMS detector consists of a central *barrel* region and *endcaps* on both sides as shown in Figure 3.4.

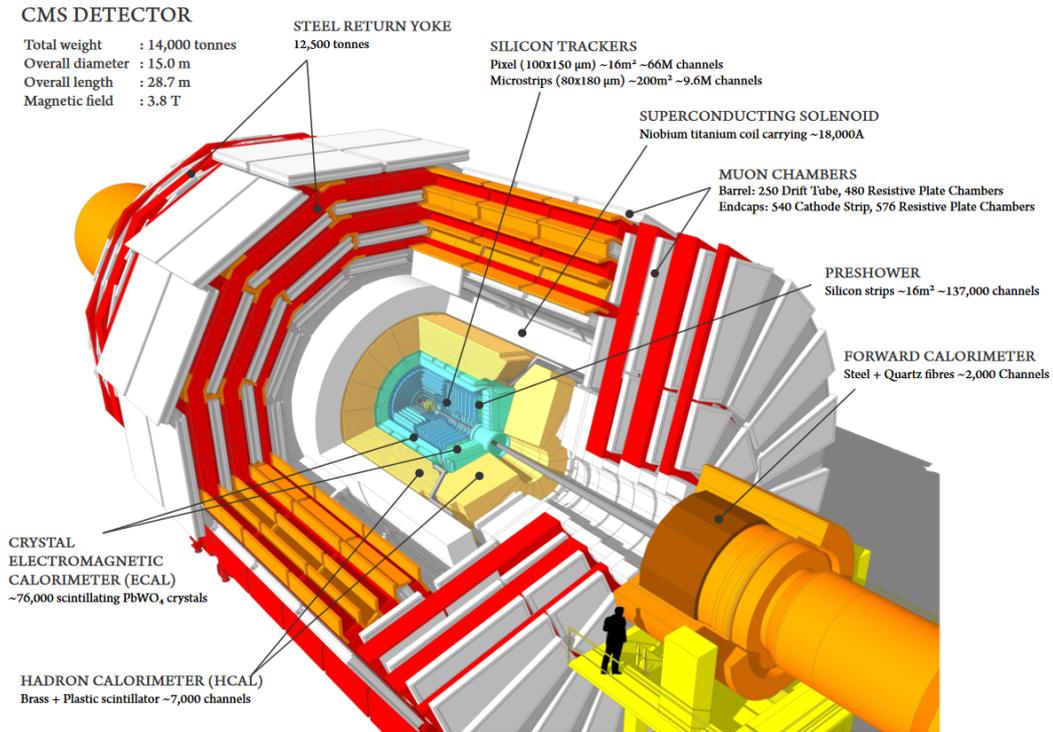
### Tracker

The innermost component is the tracker. Its purpose is to measure the trajectories of charged particles. Due to the magnetic field within the CMS detector and the resulting Lorentz force, the particle tracks are bent. Therefore, the tracker is able to measure the particle's momentum and to differentiate between positively and negatively charged particles. In addition, the track can be extrapolated to its origin, and it can be concluded, whether a particle originates from a hard interaction.

The tracker consists of several layers of silicon diodes that are reverse biased. A charged particle passing the diode generates electron-hole pairs that cause a pulse of electrical current.

In the barrel region, the first three layers are structured in a pixel design to reach the highest possible resolution. In the endcap, two layers of pixel detectors are installed. The following layers consist of silicon-strips.

### 3 Experimental Setup



**Figure 3.4:** The CMS detector consists of a barrel region complemented with endcaps on both sides. It contains several components, arranged in layers. The tracker is located nearest to the interaction point, followed by electromagnetic and hadronic calorimeters, the superconducting magnet and the muon chambers. Taken from [36].

#### Calorimeters

The second and third components of the detector are calorimeters.

The aim of calorimetry is to induce a particle shower when a particle enters, and finally to absorb the incoming particle's energy completely. The deposited energy is measured in an active material such as a scintillator. The active material can be the same as the absorber material, or two different materials can be ordered in alternating layers. The energy of the incoming particle can be calculated from e.g. the amount of light emitted by the active material.

The inner calorimeter of the CMS detector is optimised for electrons, positrons and photons and therefore called electromagnetic calorimeter (ECAL). It consists of lead-tungstate crystals, which is a dense, scintillating material. The high density corresponds to a short radiation length, which is  $X_0 = 0.89$  cm. The total length of

the ECAL is  $25.8 X_0$ .

Since the nuclear interaction length  $\lambda_I$  exceeds the radiation length of lead-tungstate by far, hadrons deposit only a small part of their energy in the ECAL and afterwards enter the hadronic calorimeter (HCAL). The CMS HCAL consists of alternating layers of brass and a plastic scintillator.

The resolution of calorimeters  $\frac{\sigma_E}{E}$  can be parametrised as follows:

$$\frac{\sigma_E}{E} = \sqrt{\left(\frac{A}{\sqrt{E}}\right)^2 + A^2 + \left(\frac{C}{E}\right)^2} \quad (3.5)$$

The first term corresponds to the statistical fluctuations of the particle shower that propagate into the energy measurement. The second term takes the possibility into account that the particle energy is not deposited completely. The third term describes the noise contribution.

### Muon Chambers

Because of their high mass, muons are much less affected by bremsstrahlung than electrons and are not absorbed in the calorimeters. Therefore, an additional tracking system, especially for muon detection, is installed outside the solenoid. Three different types of gaseous detectors are used. Drift Tubes (DT) are applied in the barrel region. They consist of positively charged wires inside a gas tank. Muons passing these chambers ionise the gas. The knocked out electrons drift towards the wires causing an electrical current. In addition, the distance from the wire is calculated from the time until the electrons reach the wire.

In the endcaps, where a large amount of background from the beam remnants exists, and the magnetic field is inhomogeneous, Cathode Strip Chambers (CSC) are used instead. Basically, negatively charged copper strips are added perpendicular to the wires. This adjustment increases the drift speed of the electrons. Another signal is caused by the ions that are drifting towards the strips. Hence, the position calculation is less dependent on the magnetic field.

In addition to both systems, Resistive Plate Chambers (RPC) are utilised to provide a faster response, which is needed for the triggering that will be discussed in the following section. They consist of two parallel and oppositely charged plates, providing a powerful and uniform electric field. Behind the anode, detecting strips are installed to determine the muon position roughly.

The muon chambers cover a rapidity range of  $|\eta| < 2.4$ .

### 3.2.3 Event Triggering

At the 2016 run period, bunch crossings are taking place every 25 ns, which leads to several billions of proton-proton collisions per second. It is not feasible to read out

and save all data from these events. Therefore, a mechanism is necessary to select events, which are potentially interesting for physics analyses. The CMS collaboration uses a multilevel trigger system for this purpose.

The first component is the hardware-based Level 1 (L1) trigger that receives only reduced information from the calorimeters and the muon system. It evaluates every single collision and preselects events according to simple signatures e.g. one or more high  $p_T$  objects. The L1 trigger reduces the number of valid events by about four orders of magnitude to a maximum of 100 kHz.

The preselected events are pipelined to the High Level Trigger (HLT) for further evaluation. The software-based HLT reconstructs candidates for all kinds of particles, now also including the tracker information and applies further selection criteria to them. Depending on the reconstructed particle candidates, the event is matched to predefined HLT paths. For example, an event with two muon candidates is matched to double muon HLT paths as well as to single muon HLT paths, if kinematic criteria required by these paths are fulfilled.

In summary, less than 800 events each second are passing the HLT algorithm and are stored in the computing centre at CERN.

#### 3.2.4 Muon Reconstruction

There can be various detector signals that indicate a certain particle. The analyst has some flexibility to determine the requirements for the signal to be identified as that particle. Since this thesis is based on muon measurements, the reconstruction of muons is explained exemplarily in this chapter.

The first part of the muon reconstruction refers to the quality of the signal. Three different working points called *IDs* are defined by the CMS collaboration: loose, medium and tight. The loose working point rejects signals in the muon chambers without matched hits in the inner tracker. The medium working point adds additional criteria on the quality of the hits inside the tracker or the muon chambers while the tight selection requires, besides other criteria, a global fit including hits in the muon chambers and the tracker.

While the loose ID is highly efficient, the tight ID is designed to contain almost no objects classified as muons wrongly.

Besides this, the *isolation* of the muons can be an additional requirement. Muons can occur inside of jets as a result of hadron or tau decays. In that case, the momentum of hadrons and photons measured by tracker and calorimeters is elevated in a cone around the muon. A loose and a tight isolation working point are established with different thresholds regarding the ratio between muon  $p_T$  and  $p_T$  of the particles around it.

### 3.3 Luminosity and Cross Section

The *cross section*  $\sigma$  is a measure of the probability that a certain process occurs. It is related to the number of events  $N$  by the integrated *luminosity*  $\mathcal{L}_{\text{int}}$ :

$$\sigma = \frac{N}{\mathcal{L}_{\text{int}}} \quad (3.6)$$

The luminosity depends on several parameters of the LHC proton beams and is obtained by the CMS collaboration with so-called *Van der Meer scans* [37]. The idea of this method is to use runs with different beam separations to calibrate the absolute luminosity scale. The luminosity in a certain run period in relation to the absolute scale is determined from the particle flux measured in the silicon pixel detector. The uncertainty of the luminosity measurement is estimated to be 2.5% [38].

### 3.4 Software Tools used for Data Analysis

In 2016, the CMS experiment stored several petabytes of event information in computing centres all over the world summarised as the Worldwide LHC Computing Grid (WLCG). It is structured hierarchically around the computing centre at CERN.

A lot of tools needed for data analysis are provided centrally in the CMS software framework (CMSSW) [39]. The software is based on the Event Data Model (EDM) concept where events are handled completely independent from each other. Hence, the workflow is highly parallelisable.

The CMSSW tools are included in several analysis frameworks developed by the working groups. In this thesis, frameworks developed at the KIT called KAPPA [40] and ARTUS [41] were used.

#### 3.4.1 Artus and Kappa Frameworks

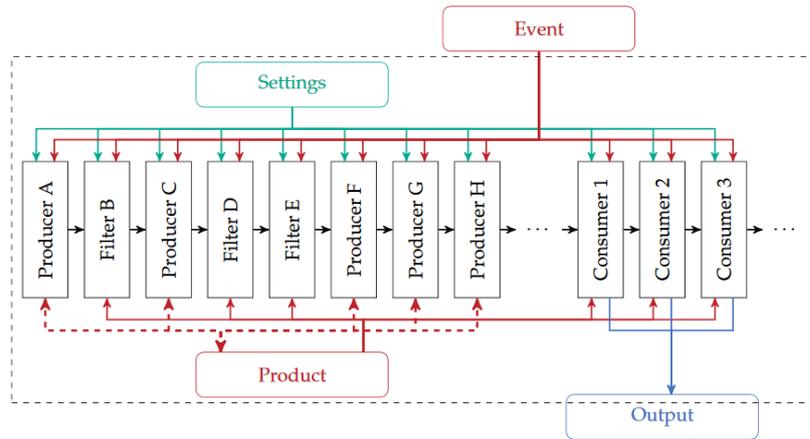
The first step of the analysis is called *skimming* and done with (KAPPA). During the skimming step, the objects required for further analysis are written out in a KAPPA specific data format. This step also can be used for a first event selection, e.g. cuts on several objects can be performed, and HLT paths (see Section 3.2.3) can be chosen. In addition, jet algorithms are applied in the skimming. These algorithms cluster particles, assumed to originate from the same hadron, to a jet object (see Section 2.2). The Kappa data format is optimised to reduce disk space. KAPPA can handle different CMSSW data formats. Hence, another advantage of the Kappa output is its independence from the input format and the CMSSW version used for skimming.

The Kappa output is further processed with the ARTUS framework. ARTUS has a modular structure, where modules are classified as producer, filter or consumer.

### 3 Experimental Setup

---

Producers calculate additional quantities, define objects and apply corrections. Filters reject events and consumers are responsible for writing out the desired observables to a ROOT [42] file. ARTUS has an expansion particular for Z+jet studies called EXCALIBUR [43]. Because of the code separation, ARTUS can be used efficiently for many kinds of analyses and easily modified by adding additional modules. ARTUS is used by several working groups including the group performing the official Z+jet energy calibration for CMS.



**Figure 3.5:** Event pipelining in ARTUS. The event information from the KAPPA file is stored in the event class. Configuration is done via a global setting file. Several producers, filters and consumers can access additional variables stored in the product class. The consumers write the results to the output file. Taken from [41].

---

## Measurement of the $Z(\rightarrow \mu\mu) + \text{jets}$ Cross Section

This chapter is dedicated to the measurement of the cross section of Drell-Yan processes.

These processes are characterised by a large cross section and a clear signature, which leads to a low background fraction for invariant lepton pair masses around the  $Z$  boson mass (see Section 2.1). In addition, muons are the most accurately measured particles in the CMS experiment. Therefore, exclusively the muon decay channel of the  $Z$  boson is studied in this analysis.

The vast majority of Drell-Yan events occurs in association with one or more jets, which balance the transverse momentum  $p_T$  of the  $Z$  boson. This process is denoted as  $Z(\rightarrow \mu\mu) + \text{jets}$  in this thesis.

The measurement of jets is challenging and therefore, measured jet properties suffer from large uncertainties. For this reason, no selections with respect to the occurring jets were performed in this thesis, and no jet observables were studied.

The  $Z(\rightarrow \mu\mu) + \text{jets}$  cross section is measured double-differentially as a function of  $p_T$  and absolute rapidity  $|y|$  of the dimuon system identified as  $Z$  decay. In addition, a second cross section measurement is performed as a function of  $\phi_\eta^*$  and dimuon  $|y|$ . The  $\phi_\eta^*$  observable is explained in Section 3.2.1. These measurements are compared to study the advantages and disadvantages of both observables for the purpose of  $Z(\rightarrow \mu\mu) + \text{jets}$  cross section measurements and PDF studies.

The first part of the data collected by the CMS detector in the year 2016 from proton-proton collisions at a centre-of-mass energy of 13 TeV were analysed for this purpose. These data correspond to an integrated luminosity of  $12.5 \text{ fb}^{-1}$ .

The CMS collaboration provides simulated events for all relevant processes with a detailed simulation of the CMS detector included in addition to the data. These simulations will be referred to as *Monte Carlo samples*. The event status before the detector simulation is named *generator level*. The properties of particles after the detector simulation is applied, are referred to as *detector level* properties.

More information on the used data and Monte Carlo samples is given in Table A.1

in the appendix.

## 4.1 Event Selection

The detector signature of a  $Z(\rightarrow \mu\mu) + \text{jets}$  event is characterised by two isolated muons, i.e. muons that are not embedded in a jet, as explained in Section 3.2.4. For the  $Z$  candidate reconstruction, each pair of oppositely charged muons is considered. If more than two muons occur in an event and pass the following selection criteria, the combination with dimuon invariant mass closest to the  $Z$  boson mass is selected. All muon properties discussed in this thesis are related to the muons chosen for  $Z$  reconstruction. In this thesis, the index  $Z$  in a formula symbol refers to the dimuon system reconstructed as  $Z$  decay.

The phase space region considered in the measurement is selected to achieve accurate results in experimental terms as well as in theoretical terms.

Trigger efficiencies are characterised by a turn-on curve (will be discussed in more detail in Section 4.2). A lower limit is required for the transverse momentum of the muons  $p_{\text{T}}^{\mu}$  to achieve a high trigger efficiency:

$$p_{\text{T}}^{\mu} > 27 \text{ GeV} \quad (4.1)$$

Muon reconstruction is most accurate in the centre of the detector, where the detector is well-calibrated. Therefore, events are selected based on muon pseudorapidity  $|\eta^{\mu}|$ :

$$|\eta^{\mu}| < 2.3 \quad (4.2)$$

The  $Z$  boson peak in the dimuon invariant mass spectrum was utilised to reject the vast majority of background. The invariant mass of the dimuon system  $m^Z$  has to be close to the  $Z$  boson mass [44]:

$$|m^Z - 91.1876 \text{ GeV}| < 10 \text{ GeV} \quad (4.3)$$

For an additional background rejection, a threshold for the maximum number of muons  $n^{\mu}$  is introduced:

$$n^{\mu} < 4 \quad (4.4)$$

For measurements as a function of  $\phi_{\eta}^*$ , only events with  $p_{\text{T}}^Z > 5 \text{ GeV}$  are selected.

The binnings chosen for the double-differential cross section measurements are shown in Table 4.1.

Low  $p_{\text{T}}^Z$  and low  $\phi_{\eta}^*$  regimes are not included in the analysis since theory predictions including soft gluon resummation would be necessary to achieve reliable theory predictions in this phase space [45].

A comparison between different Monte Carlo generators and data as a function of  $|y^Z|$  at 13 TeV is presented in [46]. In the high rapidity region, significant discrepancies

between the Monte Carlo generators are visible, indicating insufficient constraints on the theory. The phase space region  $|y^Z| > 2.0$  is therefore not included in the double-differential measurements for the purpose of PDF extractions.

Observable	Binning
$p_T^Z / \text{GeV}$	[40, 45, 50, 55, 60, 70, 75, 80, 90, 110, 130, 150, 170, 200, 250, 400]
$\phi_\eta^*$	[0.40, 0.45, 0.50, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 1.0, 1.25, 1.5, 2, 3, 4, 6, 12, 25]
$ y^Z $	[0, 0.4, 0.8, 1.2, 1.6, 2.0]

**Table 4.1:** Binning for the double-differential cross section measurements performed either as a function of  $p_T^Z$  and  $|y^Z|$  or as a function of  $\phi_\eta^*$  and  $|y^Z|$ .

## 4.2 Muon Triggering and Reconstruction

Not each  $Z(\rightarrow \mu\mu) + \text{jets}$  event occurring in a proton-proton collision is identified. This leads to a lower observed event rate than the real one. The efficiency of triggering and reconstructing muons is measured double-differentially as a function of  $p_T^\mu$  and  $|\eta^\mu|$  with the tag and probe method [47, 48] to quantify this effect.

The first step of the event selection is the triggering described in Section 3.2.3. In this analysis, HLT paths are chosen, which require at least one isolated muon. The efficiency of this trigger as a function of  $p_T^\mu$  can be seen in Figure 4.1 on the left side. This distribution is typical for trigger efficiencies. A sharp turn-on curve proceeds to a flat plateau area for higher transverse momenta. It is sufficient if one muon triggers the event. Therefore, the total trigger efficiency  $\epsilon_{\text{trig}}$  for events with two muons is

$$\epsilon_{\text{trig}} = 1 - \left[ (1 - \epsilon_{\text{trig}}^{\mu_1}) \cdot (1 - \epsilon_{\text{trig}}^{\mu_2}) \right] \quad (4.5)$$

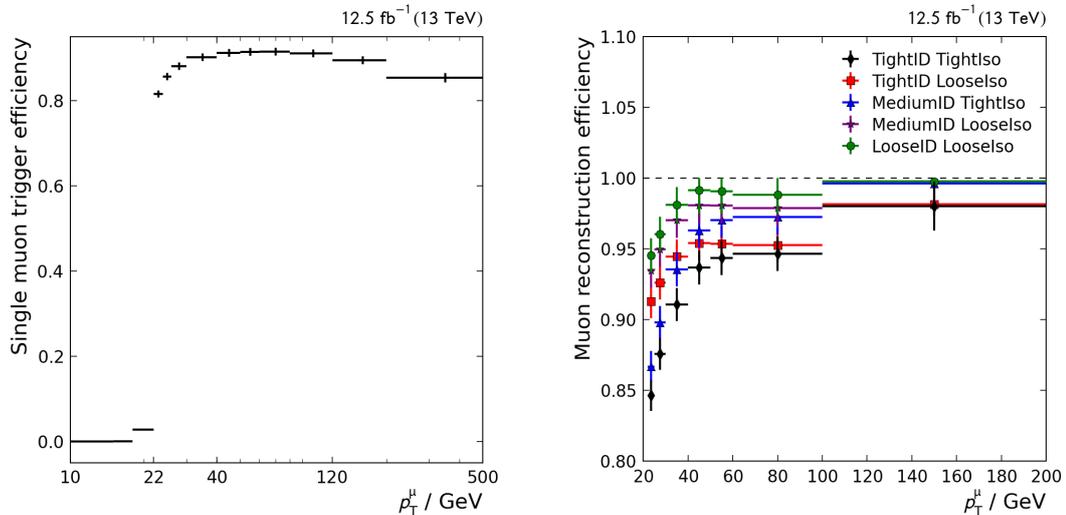
where  $\epsilon_{\text{trig}}^{\mu_i}$  refers to the trigger efficiency of the first or second muon selected for  $Z$  reconstruction, ordered by  $p_T$ .

In Section 3.2.4, different ID and isolation working points of a muon were discussed. The efficiency of the muon reconstruction depends on the working point applied in the analysis.

The efficiencies for all combinations of ID and isolation working points as a function of  $p_T^\mu$  are shown on the right side of Figure 4.1. The isolation strongly depends on the transverse momentum. Therefore the working points with tight isolation are much less efficient in the low  $p_T$  region than those with loose isolation, while at high  $p_T$ , the efficiency primarily depends on the chosen ID.

The efficiencies as functions of  $|\eta^\mu|$  are visualised in Figure A.1.

In addition to ID and isolation, an efficiency on the tracking has to be taken into account.



**Figure 4.1:** Efficiency of the used HLT paths (left) and the reconstruction efficiency for the provided combinations of muon ID and isolation (right) as a function of  $p_T^\mu$ . The trigger shows a turn-on curve with deflection point at 22 GeV. Horizontal error bars represent the bin sizes.

The tracking inefficiency is caused by highly ionising particles (HIPs) saturating the amplifier in the silicon strips, leading to the temporary blindness of the affected strips. This issue is referred to as HIP effect [49].

Since both muons have to be identified, the reconstruction efficiency  $\epsilon_{\text{reco}}$  is calculated as follows:

$$\epsilon_{\text{reco}} = \epsilon_{\text{ID}}^{\mu_1} \cdot \epsilon_{\text{ID}}^{\mu_2} \cdot \epsilon_{\text{Iso}}^{\mu_1} \cdot \epsilon_{\text{Iso}}^{\mu_2} \cdot \epsilon_{\text{Trk}}^{\mu_1} \cdot \epsilon_{\text{Trk}}^{\mu_2} \quad (4.6)$$

Here,  $\epsilon_{\text{ID}}$  refers to the ID efficiency,  $\epsilon_{\text{Iso}}$  to the isolation efficiency and  $\epsilon_{\text{Trk}}$  to the tracking efficiency. All particular efficiencies depend on  $p_T$  and  $|\eta|$  of the related muon.

A weight  $w_{\text{eff}}$  was applied to each event to compensate for the efficiency:

$$w_{\text{eff}} = \frac{1}{\epsilon_{\text{reco}} \cdot \epsilon_{\text{trig}}} \quad (4.7)$$

To decide, which ID and isolation working points to use, a *same sign test* was performed:

Since the charge of objects misidentified as muons is randomly distributed, muons with the same charge were taken for the Z reconstruction algorithm to estimate the number of events including misidentified muons. As shown in Figure 4.2 (left), the loose ID leads to a much higher number of events with a same sign muon pair than the other working points. Therefore, this working point is inappropriate for a precision

study. The performance of the medium ID and tight ID are similar. As expected, the tight isolation point leads to fewer events with a same sign muon pair as the loose isolation, but the difference is much smaller than that between loose and tight ID. In this analysis, the loose isolation was chosen to benefit from the high efficiency of this working point.

For the analysed datasets, there were issues with the medium ID due to the HIP effect mentioned above. For this reason, the definition of the medium ID was changed during the 2016 data taking. The tight ID was used in this analysis to ensure consistency with the data recorded later in 2016, which might be included in further analyses.

Another task of the muon reconstruction workflow is to correct for the bias on the muon  $p_T$  caused by tracker misalignments. The approach is based on shifted dimuon mass spectra near the peak at the Z mass depending on charge, pseudorapidity  $\eta$  and azimuth angle  $\phi$  of the muon [50, 51].

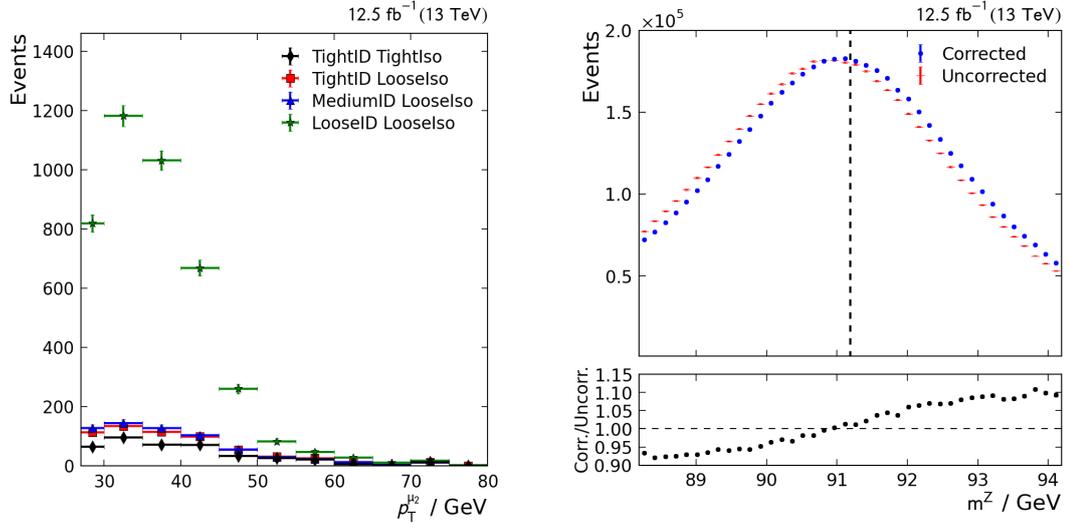
On the right side of Figure 4.2, the dimuon mass is visualised with and without momentum corrections applied. A shift of the peak towards the Z mass (91.1876 GeV [44]) is visible for the corrected data.

### 4.2.1 Verification of Efficiency Weights

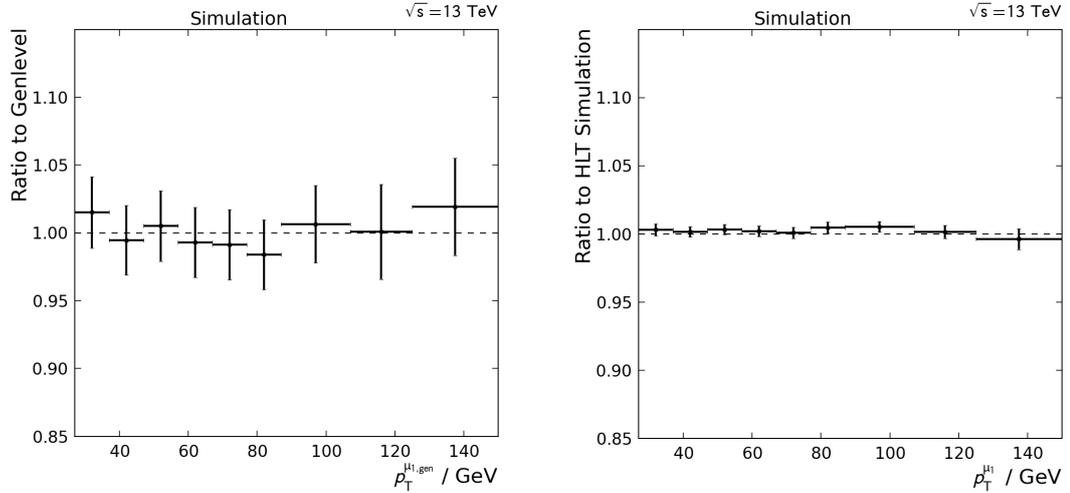
To validate the weights that were applied to correct for muon trigger and reconstruction efficiency, the signal Monte Carlo sample was used to compare reweighted distributions to reference distributions, which are independent of the particular efficiency.

For the verification of the trigger efficiency weight, the  $p_T$  distribution of the leading muon was analysed without a simulation of the CMS trigger system and with the trigger simulation and reweighting (according to Eq. 4.5). To verify the reconstruction efficiency, the event selection specified in Section 4.1 was applied at generator level. The leading muon  $p_T$  distribution obtained with this selection was compared to the distribution, where the same generator level restrictions were applied and in addition, two reconstructed muons without further restrictions were required. The latter distribution was reweighted according to Eq. 4.6. As shown in Figure 4.3, the differences between the distributions in each case are of the order of 1% and within the uncertainties on the particular weights.

## 4 Measurement of the $Z(\rightarrow \mu\mu) + jets$ Cross Section



**Figure 4.2:** Left: Subleading muon  $p_T$ , considering only events, where a Z boson could be reconstructed from muons with the same sign (same sign test) with different reconstruction working points and related weights applied. Right: Dimuon mass spectrum with and without momentum corrections applied. The Z boson mass (91.1876 GeV [44]) is illustrated to show that the corrections shift the peak towards it.



**Figure 4.3:** Validation of the muon reconstruction (left) and the event triggering (right). For the validation of the trigger weights, Monte Carlo samples with and without trigger simulation were compared. To verify the reconstruction efficiency, the generator level distribution without any detector level restrictions was compared with a selection where two reconstructed muons without further restrictions were required in addition. Weights to compensate for trigger and reconstruction efficiency were applied. The error bars represent the uncertainty on the applied weights.

### 4.3 Background Estimation

Besides  $Z(\rightarrow \mu\mu) + \text{jets}$  events, many other processes can lead to the observed final state in the detector. To estimate their contribution, the most important background processes are simulated:

- $t\bar{t}$ : Since top quarks always decay into a W boson, the muons originate from both W bosons decaying into a muon and a muon neutrino. A high number of jets and  $E_T^{\text{miss}}$  are characteristic of this process.
- WZ: With the Z decaying into a muon pair, the final state of this process is not distinguishable from a  $Z(\rightarrow \mu\mu) + \text{jets}$  event. Since the muons originate from a Z decay, WZ events are expected to follow the same dependency on the Z mass.
- ZZ: Like WZ events, two muons in the final state are observed if one Z decays into a muon pair.
- WW and tW: Similar to  $t\bar{t}$ , two muons are observed if both W bosons decay into a muon and a muon neutrino. Therefore, missing transverse energy ( $E_T^{\text{miss}}$ ) is typical for these events.
- W+jets and QCD processes: These events do not lead to a final state with two muons, but a jet can be misidentified as a muon. These processes are not simulated, but taken into account with the same sign test explained in Section 4.2.
- $Z \rightarrow \tau\tau$ : If both taus decay into a muon, a final state with two muons is observed. Since neutrinos occur at the tau decay, these events can be identified by their  $E_T^{\text{miss}}$ . In addition, the invariant mass spectrum of the dimuons is expected to be shifted towards lower masses since the neutrinos are not taken into account for the Z boson reconstruction.

The contributions of the different background processes can be seen in Figure 4.4 on the left. Their cross sections and the number of observed events are reported in Table 4.2.

For the ZZ and WZ samples, the expected dependence on the dimuon mass can be observed. Likewise, the shift to lower dimuon masses is visible for  $Z \rightarrow \tau\tau$  events, leading to an effective suppression of these events due to the  $m^Z$  restriction.

The largest background contribution comes from  $t\bar{t}$  due to the large cross section of this process compared to diboson production. It would be possible to reject large parts of  $t\bar{t}$  events as well as other background processes with a threshold on  $E_T^{\text{miss}}$ . However, the detector calibration for 2016 data is still in progress and mismeasurements in jet  $p_T$  can lead to  $E_T^{\text{miss}}$ , which would reject  $Z(\rightarrow \mu\mu) + \text{jets}$  events wrongly.

## 4 Measurement of the $Z(\rightarrow \mu\mu) + \text{jets}$ Cross Section

Process	Cross Section / pb	Cross Section Dimuon /pb	Events
$Z \rightarrow ll$	5760 [52]	1920	68588757
$t\bar{t}$	830 [53]	9.7	7590
WZ	42.3 [54]	1.5	6034
ZZ	15.4 [55]	1.1	4250
WW	119 [56]	1.4	947
tW	70 [57]	0.8	714
W+jets/QCD data	-	-	609
$Z \rightarrow \tau\tau$	1920 [52]	58	460

**Table 4.2:** Overview of background processes and their cross sections for  $\sqrt{s} = 13$  TeV. The signal  $Z \rightarrow ll$  Monte Carlo sample includes  $Z \rightarrow \tau\tau$  events. The total cross section is reported as well as the cross section in the relevant decay channels. No indirect decays via taus are included in the calculation of the cross section that leads to the same final state as  $Z(\rightarrow \mu\mu) + \text{jets}$ . These events are expected to be filtered by the event selection according to  $m^Z$ . Branching ratios are taken from [44]. The number of expected events is normalised to a luminosity of  $12.5 \text{ fb}^{-1}$ .

The background-to-signal ratio does not exceed 3% in a single phase space region, as shown in Figure 4.4 (right) for the dimuon mass distribution and in Figure A.2 as functions of  $p_T^Z$  and  $\phi_\eta^*$ . Hence, further event selections to decrease the background fraction are not expedient, and no restriction with respect to  $E_T^{\text{miss}}$  is applied.

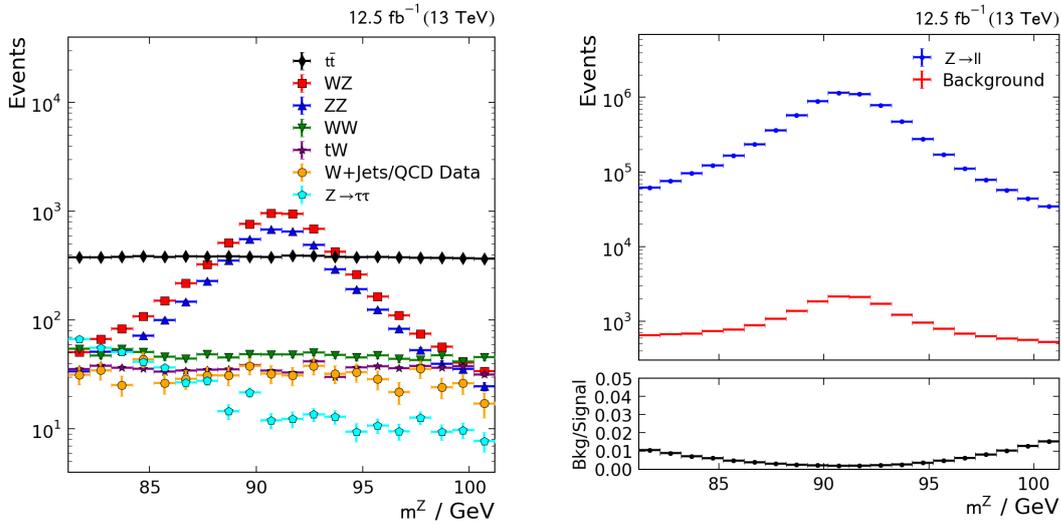
The relative background contribution is lowest at the Z mass and increases further away from the peak. Hence the dimuon mass limits, as expected, decrease the background-to-signal ratio.

The data based cross check for processes, where muons with uncorrelated charge occur, such as W+Jets and QCD processes with jets falsely identified as muons, yield a minor contribution compared to the other background processes.

Unless otherwise stated, the background contributions were subtracted from the data. An exception is the  $Z \rightarrow \tau\tau$  contribution, which is included in the signal  $Z \rightarrow ll + \text{jets}$  Monte Carlo sample and therefore not subtracted but considered in the unfolding procedure that will be explained in Section 4.5.

### 4.4 Detector Level Comparisons

Comparisons between measurement and detector simulation are an essential method to validate the detector simulation as well as the used software configuration. These comparisons were made using signal and background Monte Carlo samples. The  $Z \rightarrow \tau\tau$  contribution was removed from the signal Monte Carlo sample to obtain a



**Figure 4.4:** Contribution of the different background processes (left) and estimated background to signal fraction (right). Both as a function of the dimuon mass.

simulation of  $Z(\rightarrow \mu\mu) + \text{jets}$  events.

For illustration purposes, the background was not subtracted from the data. Instead, the background contribution is shown separately.

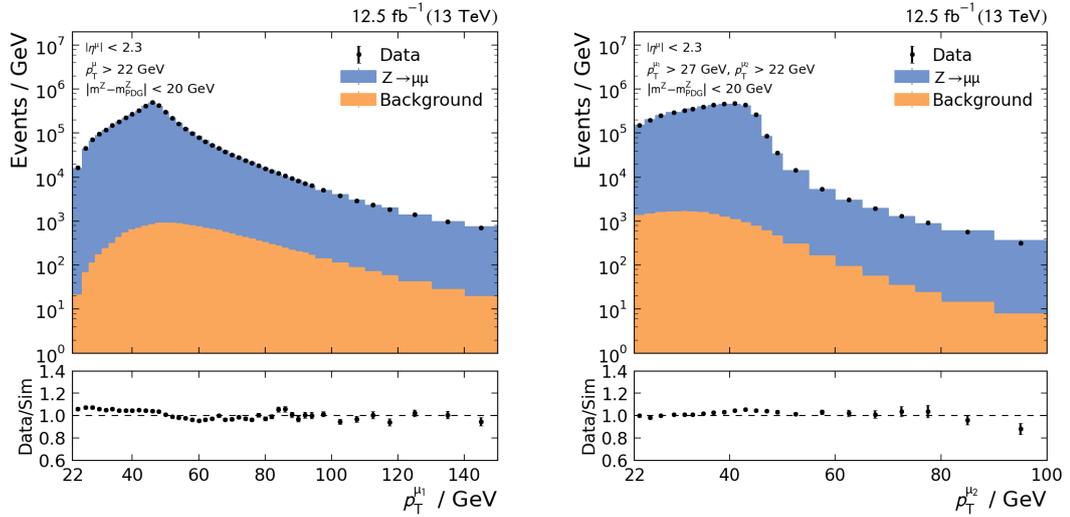
In this section, an extended phase space region was studied to analyse the impact of the event selection introduced in Section 4.1. The restriction on the muon pseudorapidity was left at  $|\eta^\mu| < 2.3$ . The muon  $p_T$  thresholds were lowered to 22 GeV, and the observed invariant dimuon mass range was increased to 20 GeV around the Z mass.

Afterwards, the event selection was tightened successively after studying the particular observable.

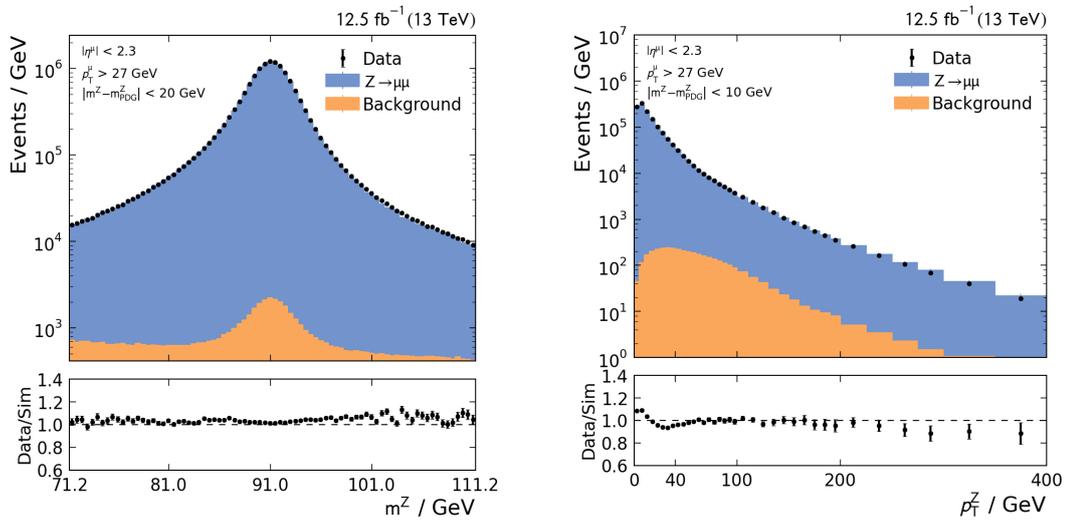
Figure 4.5 shows the transverse momenta of the muons. Differences between data and simulation are within 5%. For the leading muon, a shift towards higher  $p_T$  is observed in the simulation. This phenomenon is also observed in other analyses, e.g. in [46]. The differences are largest below 27 GeV, where the trigger efficiency is low. Therefore, this phase space region was excluded in the further analysis. For the second muon, no sizeable deviations between data and Monte Carlo are observed.

Figure 4.6 shows invariant mass and transverse momentum of the dimuon system. The position of the peak in the invariant mass distribution agrees well between data and simulation. In the mass range below the peak, the background-to-signal ratio increases for lower masses, caused by the increasing contribution of  $Z \rightarrow \tau\tau$  events and the decreasing number of signal events. In the high mass region, an excess in data is visible. These are the reasons for the restrictive dimuon mass requirement

#### 4 Measurement of the $Z(\rightarrow \mu\mu) + \text{jets}$ Cross Section



**Figure 4.5:** Comparison of detector level data and simulation for transverse momentum of the leading (left) and subleading (right) muon. The leading muon distribution is shown with a lower threshold of 22 GeV on both muons, and a dimuon mass range of 20 GeV around the Z mass. In the second muon distribution, the threshold on the leading muon was increased to 27 GeV.



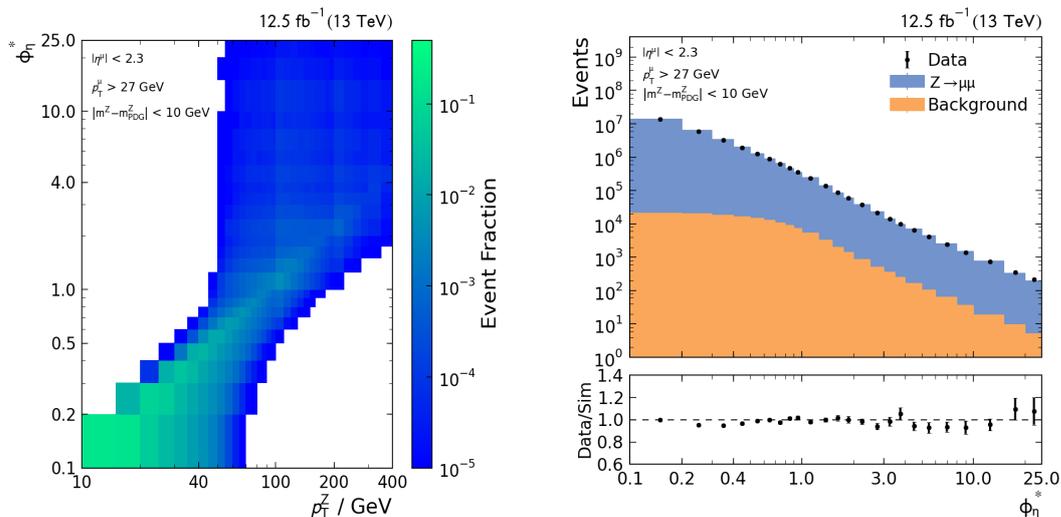
**Figure 4.6:** Comparison between data and detector level simulation for invariant mass (left) and transverse momentum (right) of the dimuon system. The mass distribution is shown in an extended range.

(see Section 4.1) chosen for the cross section measurement.

Concerning the dimuon momentum distribution, deviations of up to 10% are observed between data and simulation in the low  $p_{\text{T}}^Z$  region. The observed behaviour can be caused by the missing resummation calculations in the theory predictions used for the simulation (see Section 4.1). For  $p_{\text{T}}^Z > 250$  GeV, an increasing excess in the simulation is visible.

Both effects are also described in other analyses [46, 58, 59].

The relationship between  $\phi_{\eta}^*$  and the dimuon transverse momentum can be seen in Figure 4.7 on the left. Both observables are clearly correlated, as assumed in



**Figure 4.7:** Left: Distribution of the data as a function of  $p_{\text{T}}^Z$  and  $\phi_{\eta}^*$ . A clear correlation between both observables is visible. Right: Detector level comparison of data and simulation as a function of  $\phi_{\eta}^*$ . All event selection criteria described in 4.1 are applied.

Section 3.2.1. The comparison of the  $\phi_{\eta}^*$  distribution between data and simulation (right side of Figure 4.7) shows a similar behaviour as the  $p_{\text{T}}^Z$  distribution. This analogy includes the good agreement between data and simulation in the medium  $\phi_{\eta}^*$  region and discrepancies at low and high  $\phi_{\eta}^*$ , although they are less prominent than in the  $p_{\text{T}}^Z$  distribution.

The angular distributions of the dimuon system are shown in A.3. Concerning  $|y^Z|$ , discrepancies between data and simulation are observed in the challenging outer detector region that is not included in the double-differential cross section measurements, as outlined in Section 4.1. The number of observed events is distributed uniformly in  $\phi^Z$ .

## 4.5 Unfolding

The measured observables differ from their true values due to the detector response. These deviations concern the shape of the measured distributions as well as the total cross section in the observed phase space. For example, the peak in the dimuon mass spectrum is smeared at detector level, which leads to a higher event loss due to the mass restriction.

To compare observations with theory predictions, without applying a detector simulation to each of the latter, the measurement has to be corrected for these effects. This is done by *unfolding*.

Mathematically, the detector response can be described as

$$g(x) = \int R(x|y) \cdot f(y) dy \quad (4.8)$$

where  $g(x)$  is the measured distribution,  $f(y)$  the true distribution and  $R(x|y)$  the response function. These functions can be determined using Monte Carlo samples with a simulation of the CMS detector. Since the analysed distributions are binned in this thesis, Eq. 4.8 turns into a discrete problem, and the response function becomes the response matrix. Hence, from a purely mathematical point of view, the response matrix has to be inverted to calculate  $f(y)$  for the data.

Although the matrix inversion method is a statistically correct way to solve the unfolding problem, it turned out that, in many cases, it is not working as desired. The reason is that this method is not able to distinguish statistical fluctuations from the structure of the detector response. As a result, the anti-correlation between the bins is often overestimated, and the solution shows an oscillating behaviour.

Another simple method is the bin-by-bin correction. Thereby, the response matrix is replaced by a correction factor according to the difference between true and observed number of events in each bin. However, this method is also inappropriate in many cases, as it ignores correlations between bins completely.

Therefore, several unfolding algorithms are established in high energy physics, which rely on a kind of regularisation of the response matrix. In this analysis, the iterative D'Agostini method described in [60] is used. It is based on Bayes' Theorem:

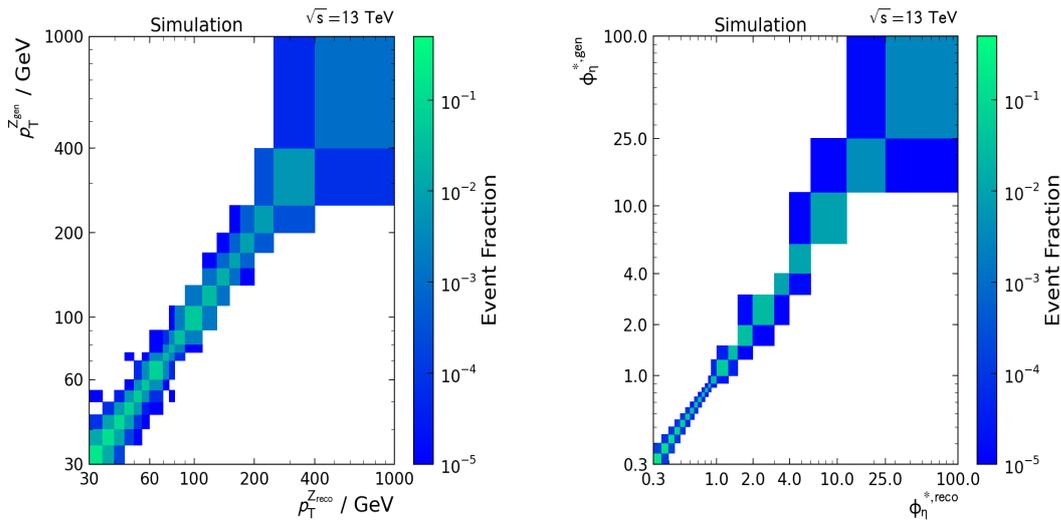
$$R(y|x) = \frac{R(x|y)p(y)}{\sum_y R(x|y)p(y)} \quad (4.9)$$

As included in the ROOUNFOLD [61] package, which was used in this analysis, at the beginning of the unfolding procedure, the generator level distribution is taken as prior  $p(y)$ . For each further iteration, the result of the former iteration is taken as new prior. The level of regularisation depends on the number of iterations. For infinite iterations, this method converges to matrix inversion [62].

The event selection explained in Section 4.1 was applied at generator level (representing the true distribution) and detector level (representing the measured distribution).

The signal Monte Carlo sample includes also  $Z \rightarrow \tau\tau$  events. They were identified and removed at generator level. Hence, the subtraction of those events is part of the unfolding, as mentioned in Section 4.3.

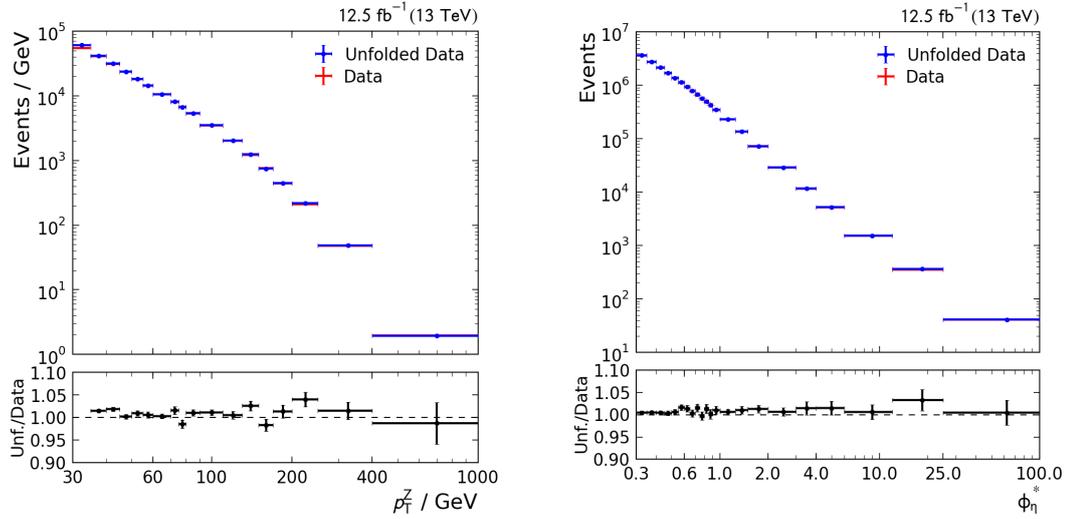
In Figure 4.9, the normalised response matrices are shown for dimuon transverse momentum and  $\phi_\eta^*$ . For both distributions, but especially for  $\phi_\eta^*$ , the vast majority



**Figure 4.8:** Normalised response matrices for  $p_T^Z$  (left) and  $\phi_\eta^*$  (right). The detector level distributions are shown on the x-axis and the generator level distributions on the y-axis.

of entries are located in the diagonal elements. A low migration between adjacent bins is visible and only very few events migrate to bins further away from their true values. Thus, it can be concluded that the resolution of the muon reconstruction in the CMS detector is high compared to the bin sizes. The asymmetric shape of the  $\phi_\eta^*$  related response matrix is caused by the limited number of events in the Monte Carlo sample. The off-diagonal elements are only filled with one or two individual events in the high  $\phi_\eta^*$  region. Thus, as predicted in Section 3.2.1, the influence of migration effects between the bins can be decreased by replacing  $p_T^Z$  with  $\phi_\eta^*$  in the analysis.

The comparison between the unfolded data and the data before unfolding is shown in Figure 4.9. The differences are at the level of only a few percent. In the first bin of the  $p_T^Z$  distribution, a somewhat higher impact of the unfolding is visible, which is caused by the neglect of migration effects from outside the unfolded phase space. Extended  $p_T^Z$  and  $\phi_\eta^*$  ranges were unfolded to avoid this effect in phase space regions included in the cross section measurement.



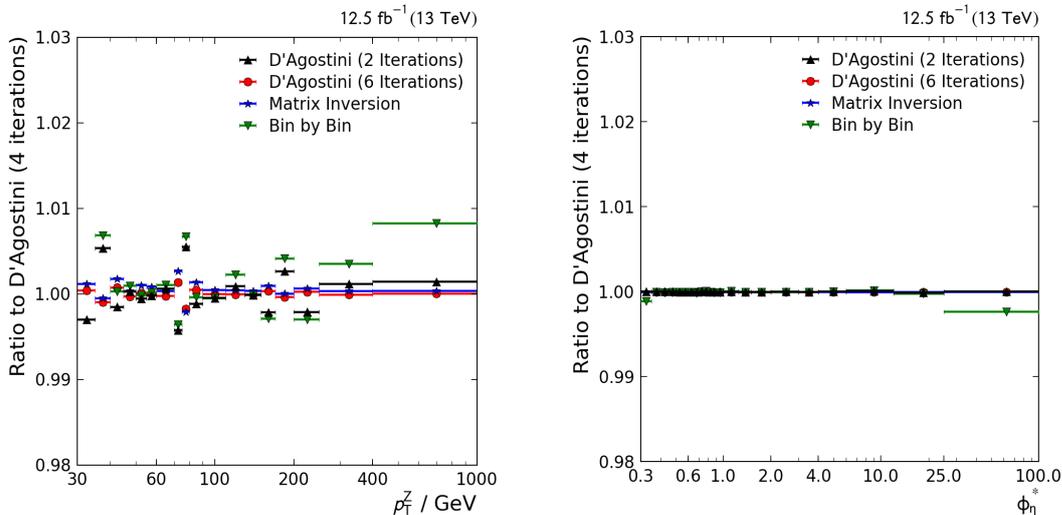
**Figure 4.9:** Comparison between the unfolded data and the background subtracted data before unfolding. The unfolding was performed with respect to the shown observable,  $p_T^Z$  (left) and  $\phi_n^*$  (right), respectively.

The increased detector resolution in the  $\phi_n^*$  distribution leads to a more regular unfolding correction compared to  $p_T^Z$ .

For the double-differential measurement, the unfolding was performed in the particular  $|y^Z|$  regions. This approach neglects fluctuations between the  $|y^Z|$  bins. However, these migration effects were studied and found to be negligible. The corresponding plots for the double-differential measurements can be found in the Figures A.4 and A.5.

As an additional study, the unfolding was performed with the matrix inversion and the bin-by-bin method. Furthermore, the number of iterations used in the iterative D’Agostini approach was varied. The results are shown in Figure 4.10. The differences between the unfolding methods are less than 1%. The highest deviations to the D’Agostini method with four iterations are observed for the bin-by-bin unfolding as a function of  $p_T^Z$  while the unfolding results are robust against an increase in the number of iterations up to matrix inversion.

For unfolding with respect to  $\phi_n^*$ , an excellent agreement between all studied methods can be observed.



**Figure 4.10:** Comparison of different unfolding methods for  $p_T^Z$  (left) and  $\phi_n^*$  (right). Analysed were matrix inversion and bin-by-bin unfolding. Also, the number of iterations used for the D'Agostini method was varied. Shown is the ratio of the data, unfolded with the particular methods to the data unfolded with the D'Agostini method using four iterations. No uncertainty is shown, since the uncertainty that arises from the choice of the unfolding algorithm was not studied.

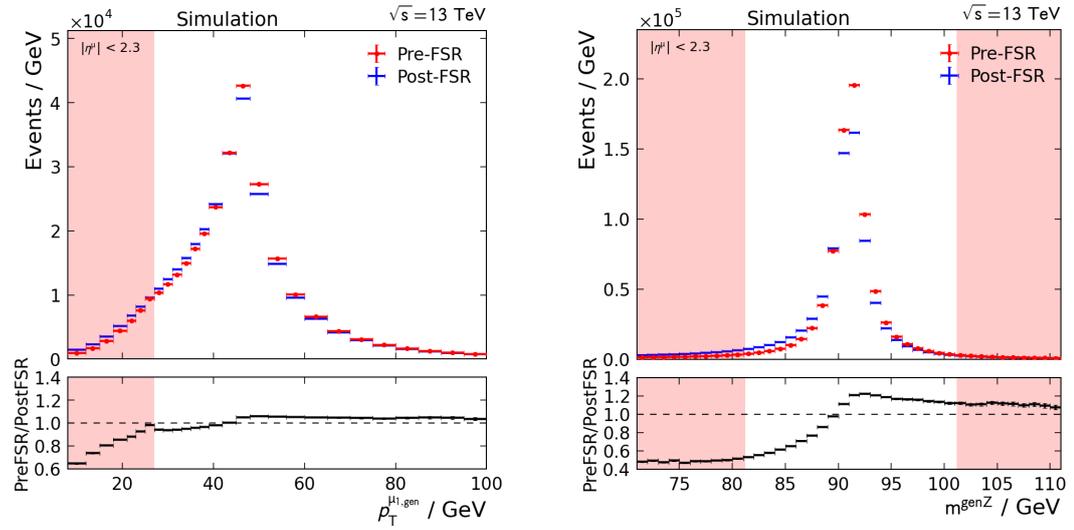
## 4.6 Final State Radiation Studies

Until now, the Z boson was reconstructed from muons as they appear in the final state of the event. It has to be considered that a muon can emit photons and therefore changes its properties after its emergence in the Z boson decay. The impact of this *final state radiation* (FSR) was studied using the generator level information of the Monte Carlo. All generated particles and their decay products are listed in the event information so that the muons occurring in the final state of the event generation can be retraced to the original muons.

On the left side of Figure 4.11, the leading muon  $p_T$  is shown exemplarily before and after FSR. The muons lose a noticeable amount of their  $p_T$  due to FSR.

Figure 4.11 (right) shows the dimuon mass distribution of pre-FSR and post-FSR muons. The spectrum is shifted and smeared due to FSR. Especially the tail below the Z mass is enriched, because of migration from the peak. Both effects lead to a lower amount of events passing the selection after FSR is considered. To correct the data for the discrepancy, the FSR effect was studied as a function of the dimuon transverse momentum,  $\phi_n^*$  and absolute dimuon rapidity, as shown in Figure 4.12.

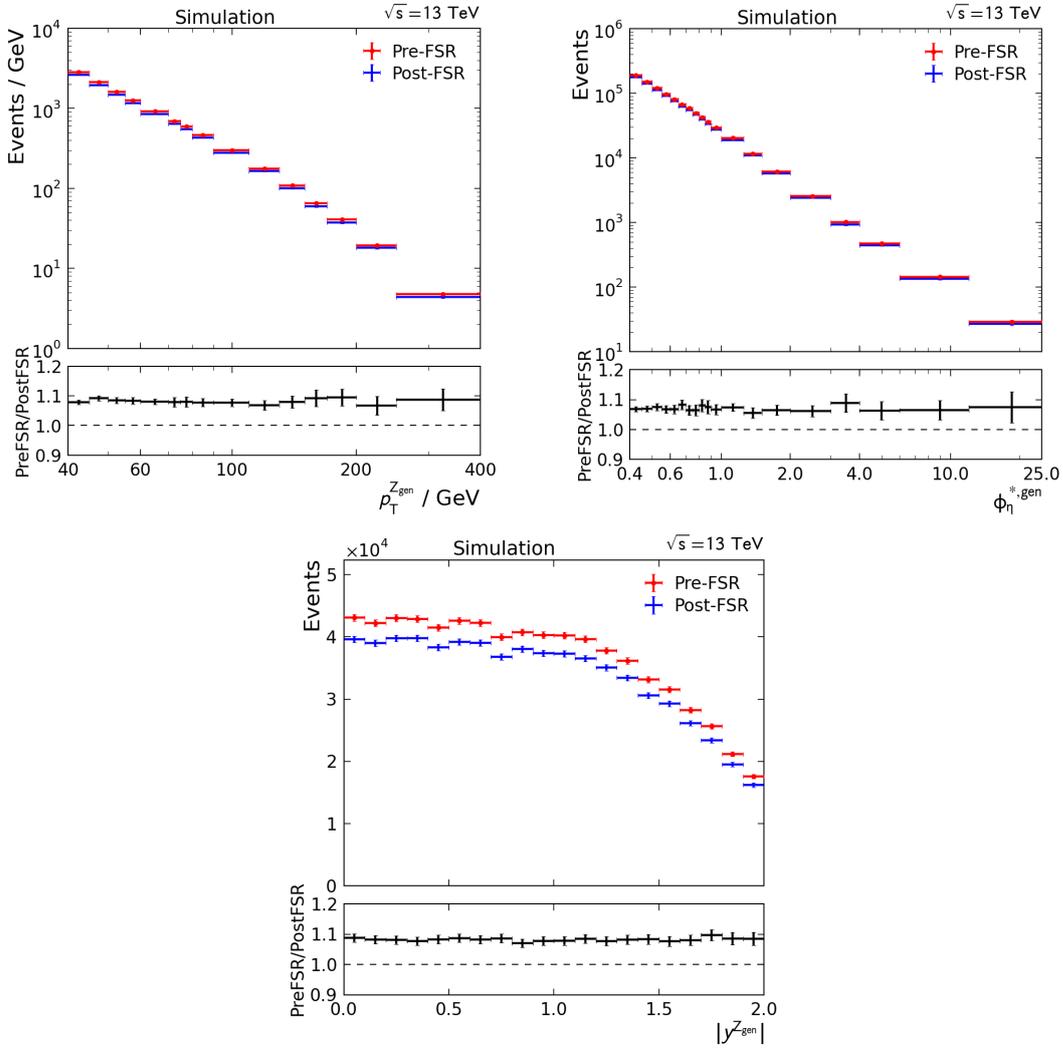
The ratio between the pre-FSR distribution and post-FSR distribution is nearly



**Figure 4.11:** Comparison of the generator level leading muon  $p_T$  and the dimuon mass spectrum before and after FSR. Only  $|\eta^\mu| < 2.3$  is required. The red area is not considered in the cross section measurement. The muons are ordered by  $p_T$  before FSR. The second muon  $p_T$  and the momentum of the reordered post FSR muons are shown in Figure A.6.

independent of  $|y^Z|$ . In the dimuon  $p_T$  and  $\phi_\eta^*$  distribution, due to the steeper spectrum, uncertainties are larger. Regardless, the ratio is in agreement with a constant factor in the observed phase space. Therefore, a constant was fitted to the rapidity ratio to determine the FSR correction weight  $w_{\text{FSR}}$ . Fits to the other distributions lead to similar results. To estimate an uncertainty on this weight, further studies with different Monte Carlo generators are expedient. However, this could not be achieved within this thesis. Hence, a preliminary ad hoc estimation was done:

$$w_{\text{FSR}} = 8\% \pm 2\% \quad (4.10)$$



**Figure 4.12:** Comparison of the generator level dimuon  $p_T$  (top left),  $\phi_n^*$  (top right) and  $|y^Z|$  (bottom) distributions before and after FSR. The event selection as described in Section 4.1 is applied on generator level. The ratio is in agreement with a constant correction factor in the observed phase space regions.

## 4.7 Experimental Uncertainties

The approaches to estimate the different uncertainties are based on the  $Z(\rightarrow ee)+\text{jets}$  analysis at  $\sqrt{s}=8$  TeV performed by Dominik Haitz [59].

- Trigger and reconstruction efficiencies: The uncertainties on the trigger and reconstruction efficiency are provided as functions of muon  $p_T$  and pseudorapidity. To propagate them to the observables of interest, the event weights are recalculated varying the efficiencies one standard deviation upwards and downwards. The larger deviation is taken as uncertainty.
- Background: Since the background contribution is low (see Section 4.4), the uncertainty on each Monte Carlo sample is not studied in detail. Instead, a conservative ad hoc estimation is done: The uncertainty is estimated to be 50% of the total background distribution. The reason for the high uncertainty is that the background Monte Carlo samples are mainly leading order simulations normalised to higher order cross sections. Hence, the background distribution in a certain phase space region is not simulated accurately.
- Luminosity: The uncertainty on the integrated luminosity is 2.5% for all data recorded in 2016 (see Section 3.3).
- Statistical and unfolding: The statistical uncertainties on data and Monte Carlo are given by the square root of the number of entries in each bin. To estimate the influence of unfolding,  $10^7$  pseudo-experiments are generated, where measured distributions and response matrix are varied simultaneously according to their statistical uncertainties. The mean and standard deviation are calculated from the results of these pseudo-experiments.
- Final state radiation: The uncertainty on the FSR modelling was estimated to be 2% as discussed in Section 4.6.

The uncertainty on the muon energy scale was not studied in the course of this thesis. It is expected to be relevant, if at all, only in the high  $p_T^Z$  phase space region [46].

Figure 4.13 shows the values obtained for the different uncertainty sources as functions of  $p_T^Z$  and  $\phi_\eta^*$ . The distribution is similar for both observables. One notable difference is the uncertainty on the reconstruction efficiency, which shows a flatter behaviour for  $\phi_\eta^*$  than for dimuon  $p_T$ .

The uncertainties in the particular dimuon rapidity regions are shown in the Figures A.7 and A.8.

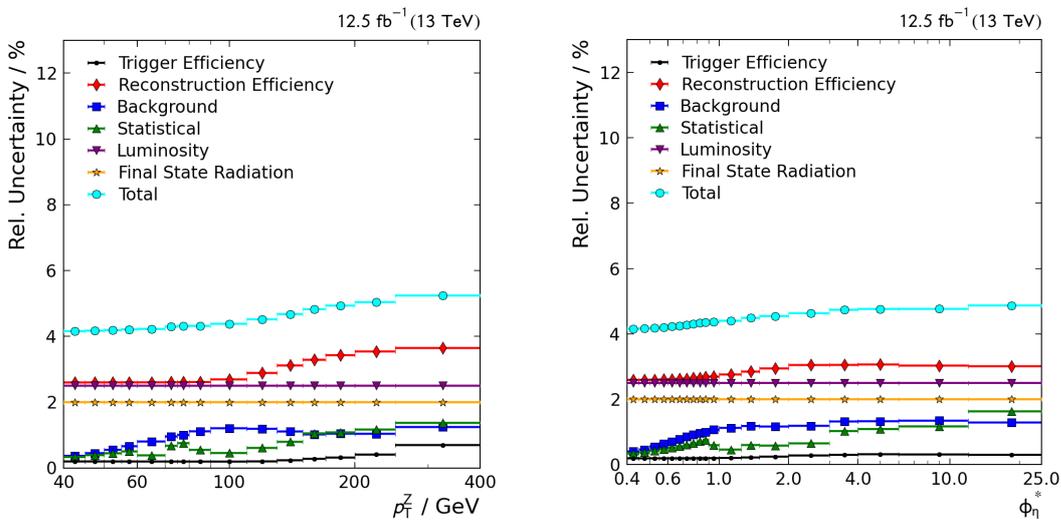
The sources are assumed not to be correlated to each other, and therefore, the total uncertainty is obtained by adding them quadratically. Besides the statistical uncertainty, the FSR uncertainty was assumed to be uncorrelated between the different

bins. The reason for this assumption is that the uncertainty on the FSR effect is a preliminary estimation and the suggestion of a constant factor (see Section 4.6) still has to be validated. This approach possibly overestimates the statistic component of the FSR uncertainty. On the other hand, the other uncertainties were treated as completely correlated between the bins, and the statistical part on the efficiency uncertainties was not estimated separately. This part is small in most phase space regions, although in the high rapidity region, it may have an impact on the total uncorrelated uncertainty.

The main sources of uncertainty are the reconstruction efficiency, luminosity and FSR. The accuracy of the reconstruction efficiency is expected to increase with further calibration of the detector for 2016 data, especially the statistical part of the uncertainty will be negligible for the final efficiency calculation. Hence, the assumption of a fully correlated uncertainty will be reasonable using the complete 2016 dataset [63].

The luminosity measurement is already published for the analysed dataset. Hence, this uncertainty is the final result. Concerning FSR, a more precise estimation of the correction factor can be achieved by further studies.

The total uncertainty amounts to about 4% to 5% for the single differential measurement and increases for high  $p_T^Z$  and  $\phi_\eta^*$ . In the outermost rapidity bin, the uncertainty reaches up to 9% in the highest  $\phi_\eta^*$  bin analysed.



**Figure 4.13:** Estimation of experimental uncertainties as functions of  $p_T^Z$  (left) and  $\phi_\eta^*$  (right). The total uncertainty is obtained by adding the values obtained for the individual sources quadratically.



---

## Theory Predictions for the $Z(\rightarrow \mu\mu) + \text{jets}$ Cross Section

In this chapter, the workflow to obtain theory predictions for the  $Z(\rightarrow \mu\mu) + \text{jets}$  cross sections using perturbative QCD (see Section 2.2.3) is presented. In addition, the sensitivity of the cross sections to the PDFs is studied and theory uncertainties are estimated.

### 5.1 Computation Workflow

SHERPA [64] was used to calculate fixed-order theory predictions for the  $Z(\rightarrow \mu\mu) + \text{jets}$  cross sections at next-to-leading order using the same phase space selections as applied to the measurements (see Chapter 4.1). The virtual corrections were calculated with BLACKHAT [65].

The transverse energy of the dimuon system was chosen as the factorisation and renormalisation scale of each event:

$$\mu_r = \mu_f = \sqrt{m_Z^2 + p_{T,Z}^2} \quad (5.1)$$

The generated events were analysed with RIVET [66]. This analysis includes, for instance, the Z boson reconstruction from the muons and the calculation of the double-differential cross sections using the binning specified in Section 4.1.

FASTNLO [67, 68] was used to evaluate the events with different PDF sets via the RIVET interface MCGRID [69]. The FASTNLO interpolation method is summarised in the next section. This approach allows recalculating a cross section prediction with another PDF within seconds from a look-up table and avoids redoing the event generation procedure.

### 5.1.1 fastNLO

The basic idea of FASTNLO is to decouple the PDFs from the cross section  $\sigma_{pp \rightarrow X}$  calculation done with Eq. 2.11. To maintain simplicity, the equation is rewritten in a slightly different manner with  $\mu_r = \mu_f = \mu$ :

$$\sigma_{pp \rightarrow X}(\mu) = \sum_{k,n} \alpha_s^n(\mu) \int F_k(x_1, x_2, \mu) \cdot c_k^n(x_1, x_2, \mu) dx_1 dx_2 \quad (5.2)$$

The PDFs of both interacting partons are combined to the function  $F_k$ , where  $k$  runs over all combinations of parton flavours. The hard process cross section is replaced by perturbative coefficients  $c_k^n$  for each combination of parton flavours and each order  $n$  in perturbation theory.

Eigenfunctions  $e^{l,m}$  and  $b^o$  are introduced around fixed interpolation values  $x_1^l, x_2^m$  and  $\mu^o$ , respectively:

$$F_k(x_1, x_2, \mu) \approx \sum_{l,m,o} F_k(x_1^l, x_2^m, \mu^o) \cdot e^l(x_1) e^m(x_2) b^o(\mu) \quad (5.3)$$

The interpolation values form a finite lattice.

The Monte Carlo integration can now be performed independently of the PDFs:

$$\begin{aligned} & \int F_k(x_1, x_2, \mu) \cdot c_k^n(x_1, x_2, \mu) dx_1 dx_2 \\ &= \sum_{l,m,o} \int \left[ c_k^n(x_1, x_2, \mu) \cdot e^l(x_1) e^m(x_2) b^o(\mu) dx_1 dx_2 \right] F_k(x_1^l, x_2^m, \mu^o) \\ &= \sum_{l,m,o} \hat{\sigma}_{k,l,m,o}^n \cdot F_k(x_1^l, x_2^m, \mu^o) \end{aligned} \quad (5.4)$$

The perturbative coefficients  $\hat{\sigma}_{k,l,m,o}^n$  are calculated once and stored in a table.

Any further calculations with other PDF sets only require to calculate the sum:

$$\sigma_{pp \rightarrow X}(\mu) \approx \sum_{n,k,l,m,o} \hat{\sigma}_{k,l,m,o}^n(\mu) \alpha_s^n(\mu) F_k(x_1^l, x_2^m, \mu^o) \quad (5.5)$$

## 5.2 Sensitivity of the Theory Predictions to the PDFs

The aim of this analysis is to combine theory predictions with measured cross sections to improve the precision of the PDFs. For this purpose, the sensitivity of the studied process on the PDFs is a basic prerequisite.

The sensitivity can be studied through the correlation between the particular PDF flavours and the cross section predictions. The method to calculate the correlation coefficients from NNPDF replicas is described in [70].

Figure 5.1 shows the correlation between the  $Z(\rightarrow \mu\mu) + \text{jets}$  cross section and the PDFs exemplarily as a function of dimuon  $p_T$  and the momentum fraction  $x$  of the parton. The correlation as a function of  $|y^Z|$  and  $\phi_\eta^*$  is shown in the Figures A.9 and A.10.

The deep red and deep blue areas determine the phase space regions, where the PDFs and the cross section predictions are highly correlated. These are the regions where the theory predictions for the cross section are expected to be most sensitive to the PDFs. The correlation to the gluon is the highest among all flavours, in particular in the region  $10^{-3} < x < 3 \cdot 10^{-2}$ . This is evident, taking into account that gluons are involved in about two-thirds of  $Z(\rightarrow \mu\mu) + \text{jets}$  events [59].

The high sensitivity on the gluon PDF is promising for PDF fitting, considering that the uncertainties on the gluon PDF are currently larger than those on the quark PDFs. The reason is that the current PDF sets are mainly determined from electron-proton collisions, referred to as deep inelastic scattering (DIS) data (see Chapter 2.2.2). These data are not directly sensitive to the gluon PDF. Hence, the gluon PDF is mainly constrained through scaling violations [71].

A high correlation between the cross section predictions and the sea quark PDF is observed for  $6 \cdot 10^{-3} < x < 2 \cdot 10^{-1}$ . The correlation to the valence quark PDFs, in general, is small. Somewhat higher correlations are only visible at high  $p_T^Z$  and high  $x$ . This can be explained with the rarity of gluons and sea quarks that carry a high momentum fraction of the proton. Hence, an increasing fraction of valence quarks involved in high  $p_T$  interactions can be assumed.

## 5.3 Theory Uncertainties

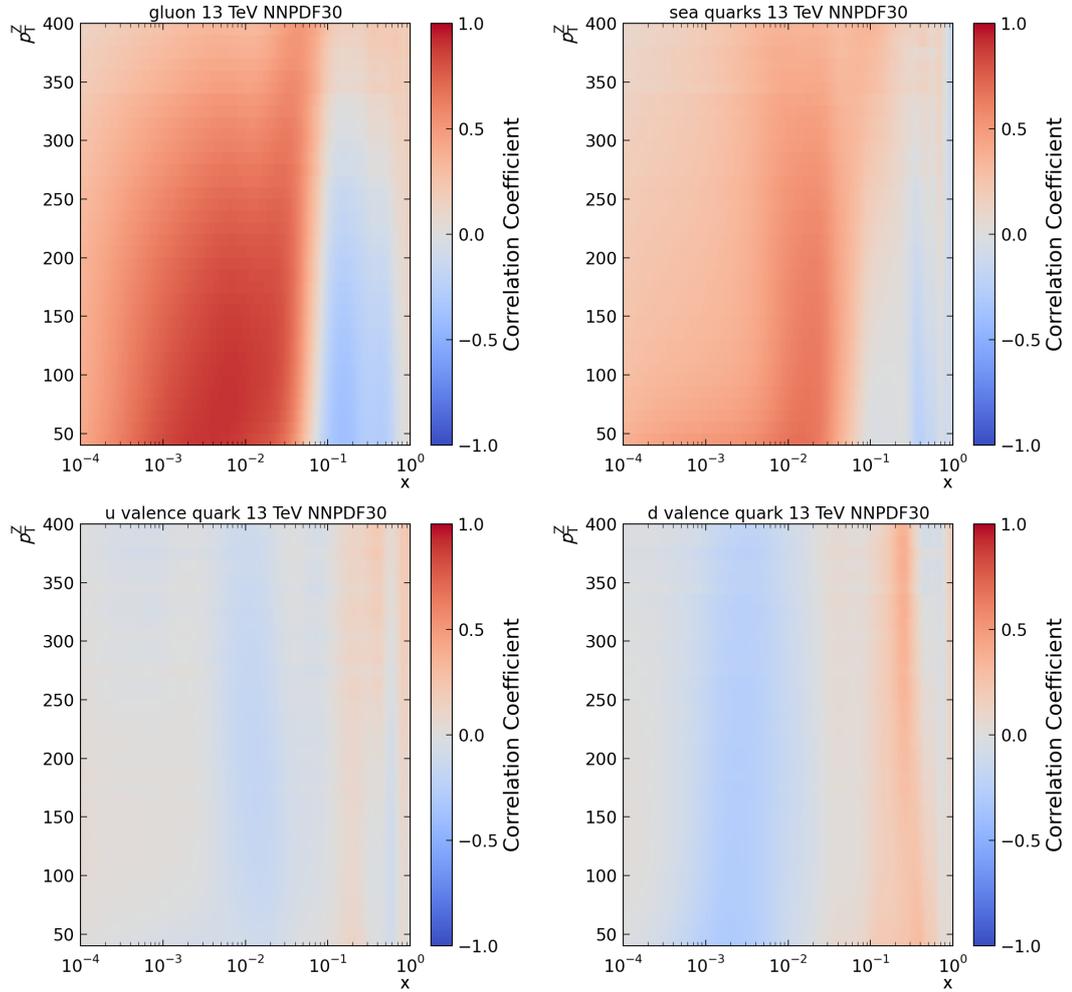
For the theory predictions, the main sources of uncertainty are the PDFs and the choice of the renormalisation and factorisation scales. The latter arises from the truncation of the perturbation series as explained in Section 2.2.3. In addition, a statistical uncertainty on the numeric integration methods of 0.5% is assumed.

In this analysis, a common approach was used to estimate the scale uncertainty. The cross section was recalculated with several variations of the renormalisation and factorisation scale  $C_{\mu_r, f}$  determined in Eq. 5.1:

$$(C_{\mu_r}, C_{\mu_f}) = (1/2, 1/2), (1/2, 1), (1, 1/2), (1, 2), (2, 1), (2, 2) \quad (5.6)$$

The scale uncertainty is asymmetric. Variations that are leading to lower and those leading to higher cross section predictions are treated separately. In both cases, the largest deviation from the central value defines the scale uncertainty.

The variations were calculated with the HOPPET [72] package interfaced to FASTNLO. The different PDF sets prescribe different methods to derive the PDF

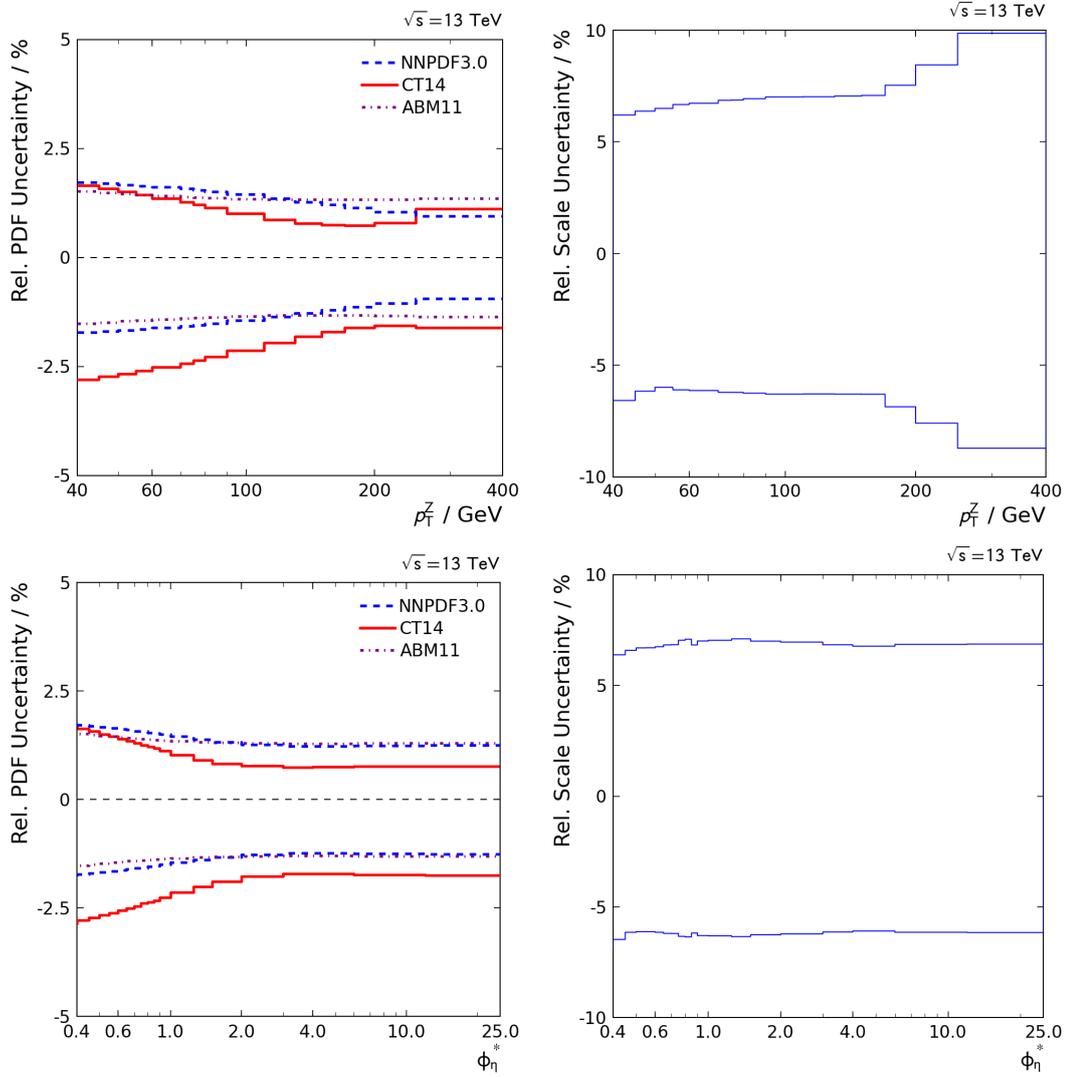


**Figure 5.1:** Correlation between PDF and  $Z(\rightarrow \mu\mu) + \text{jets}$  cross section predictions as a function of  $x$  and dimuon transverse momentum, calculated with NNPFD 3.0. The correlation is shown in particular for the gluon (upper left), the sea quarks (upper right) and up (lower left) and down (lower right) valence quarks.

uncertainty. It was consistently calculated for each PDF set at a 68% confidence level using FASTNLO.

Figure 5.2 shows the relative scale and PDF uncertainties as functions of dimuon  $p_T$  and  $\phi_\eta^*$ . The PDF uncertainty amounts to less than 3% for all studied PDF sets, shows a similar behaviour for  $p_T^Z$  and  $\phi_\eta^*$ , and decreases for higher values of those observables.

The scale uncertainty amounts to more than 6% of the central value in every bin and is the largest uncertainty in this analysis. It increases at high  $p_T^Z$  but remains constant for  $\phi_\eta^*$  in the observed range. Thus, the experimentally accessible phase space range above  $\phi_\eta^* = 0.4$  is continuously well described by the NLO theory predictions, which indicates an advantage over  $p_T^Z$  for the PDF determination.



**Figure 5.2:** PDF (left) and scale (right) uncertainty as a function of  $p_T^Z$  (top) and  $\phi_\eta^*$  (bottom). The PDF uncertainty is shown for the NNPDF 3.0, CT14 and ABM11 PDF sets.

---

## Comparison of Measurement and Theory Predictions and PDF Determination

In the following, the double-differential cross section measurements performed in Chapter 4 are compared to the NLO theory predictions obtained with the workflow described in Chapter 5. Both are used in a combined fit with HERA I+II data to study the impact of the  $Z(\rightarrow \mu\mu) + \text{jets}$  cross section measurement on the PDF determination.

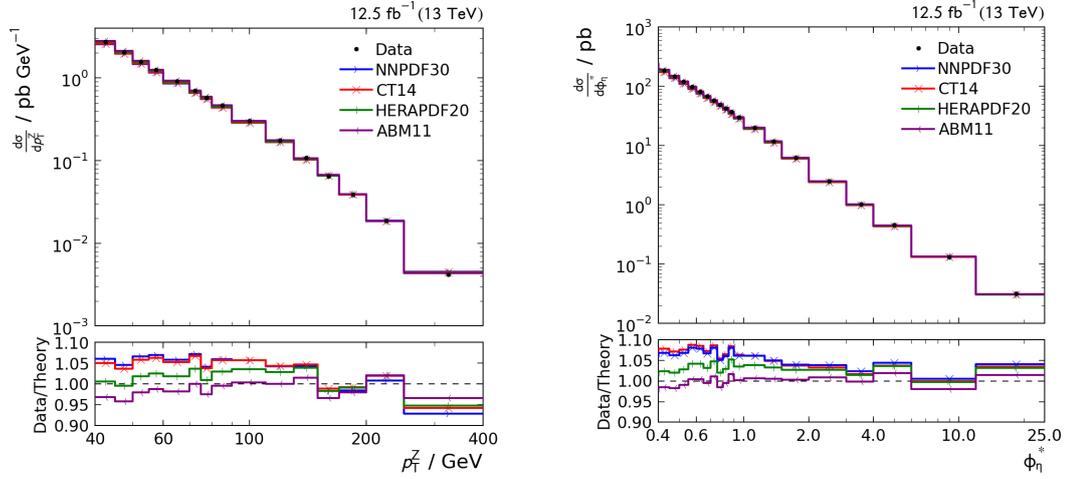
### 6.1 Comparison with NLO Predictions

As a first comparison between the measurement and the theory predictions, the single-differential cross sections with respect to  $p_T^Z$  and  $\phi_\eta^*$  were examined. The results are shown in Figure 6.1. The FASTNLO tables are evaluated with the NNPDF 3.0, CT14, HERAPDF 2.0 and ABM11 PDF sets.

A noticeable discrepancy among the different theory predictions is visible in both distributions, especially in the low  $p_T^Z$  and low  $\phi_\eta^*$  phase space regions. This indicates a high sensitivity of these predictions to the PDFs.

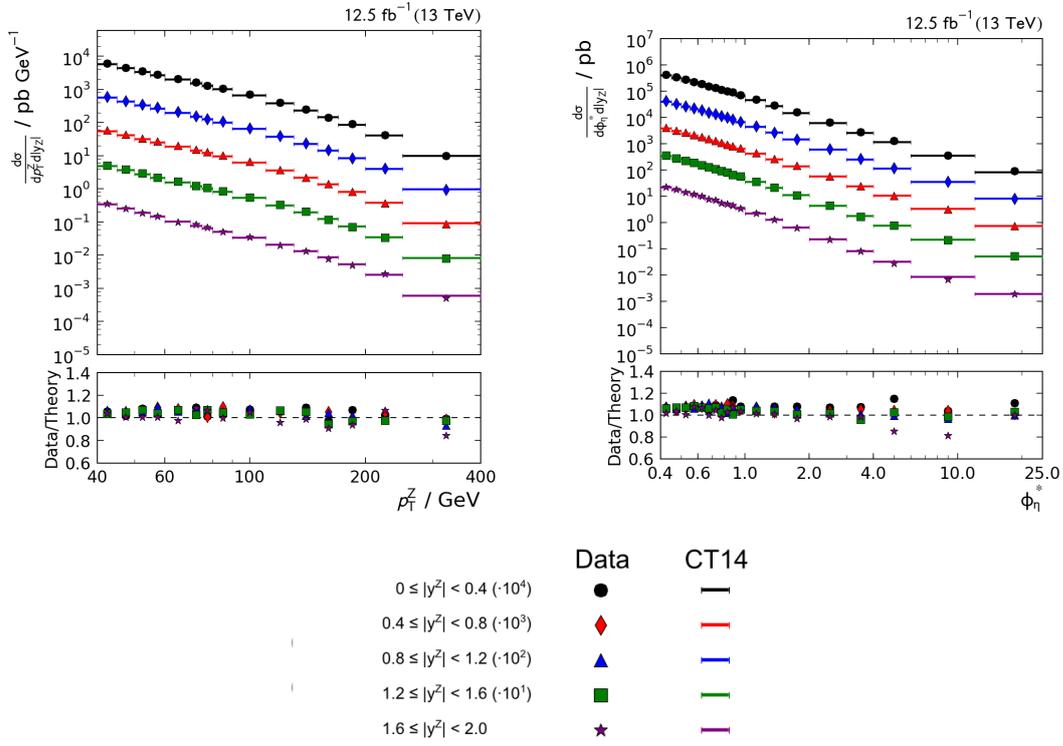
The difference between data and theory predictions obtained with the HERAPDF 2.0 and ABM11 PDF sets are at the level of 2-4% percent in almost all phase space regions. Theory predictions with CT14 and NNPDF 3.0 systematically underestimate the data by 4-7%, except for the high  $p_T^Z$  region, where the predicted cross section exceeds the measured one by about 5% for all PDF sets. As discussed in Section 4.4, this behaviour is known and probably explained by missing higher-order corrections, including electroweak effects [73], in the theory predictions.

The measured double-differential cross sections compared to the theory predictions obtained with the CT14 PDF set are shown in Figure 6.2 as an overview. The ratio of the data to theory predictions with different PDF sets in the particular  $|y^Z|$  bins is visualised in the Figures 6.3 and 6.4. The double-differential distributions are consistent with the observations made for the single-differential measurements. At



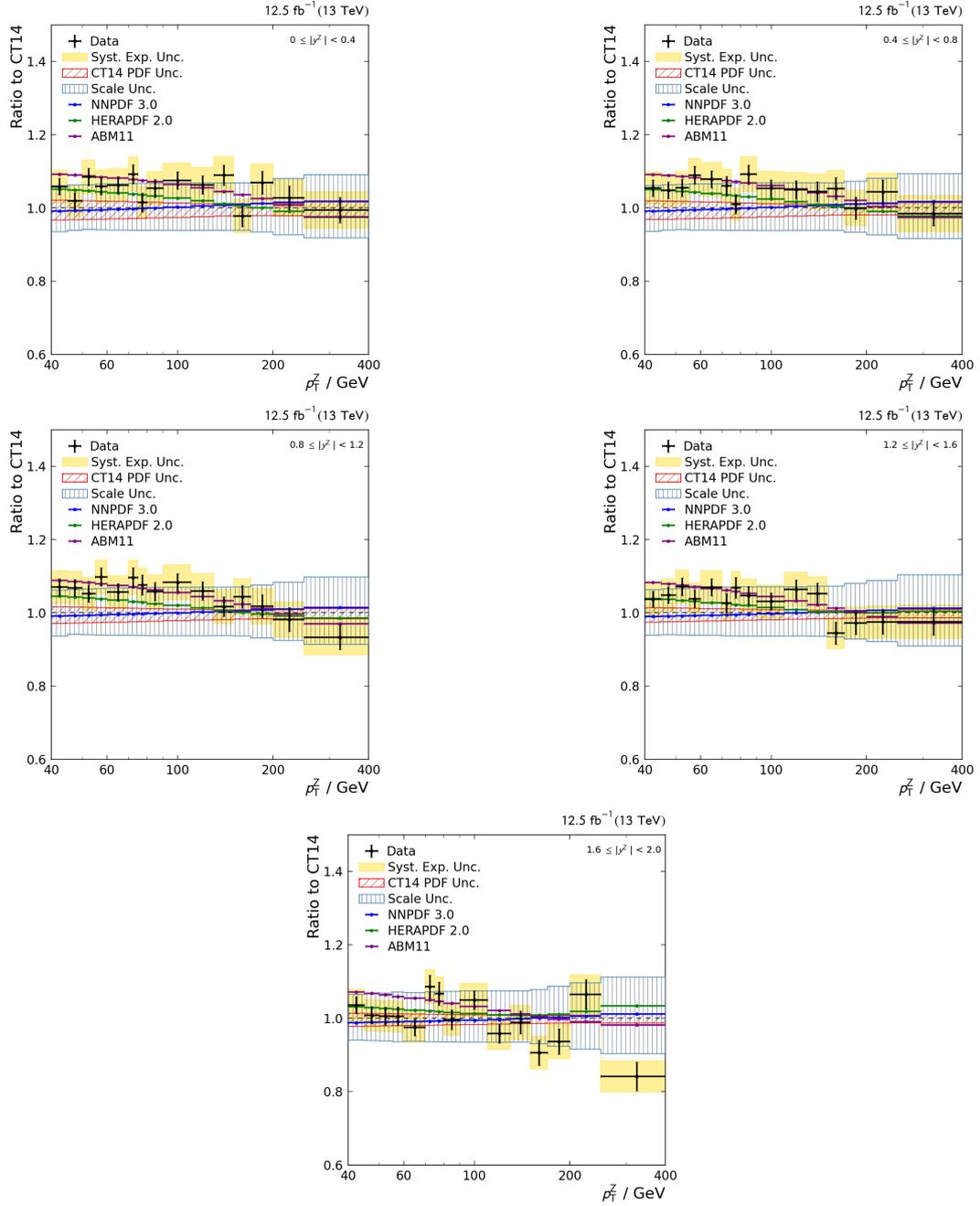
**Figure 6.1:** Single-differential cross section as a function of  $p_T^Z$  (left) and  $\phi_n^*$  (right). Theory predictions at NLO are evaluated with different PDF sets.

outermost rapidities, the observed cross section is somewhat lower than predicted. Several explanations can be identified: First, the calibration of the outer detector region is challenging, and the detector calibration in this region is not as accurate as in the centre of the detector, especially since latest data were used in this analysis, and studies of detector behaviour are still in progress. Second, as mentioned in Section 4.1, the high rapidity region is described inaccurately by the theory predictions. Although the region most affected by this was not considered in the measurement, an impact is still possible for  $|y^Z| < 2.0$ .



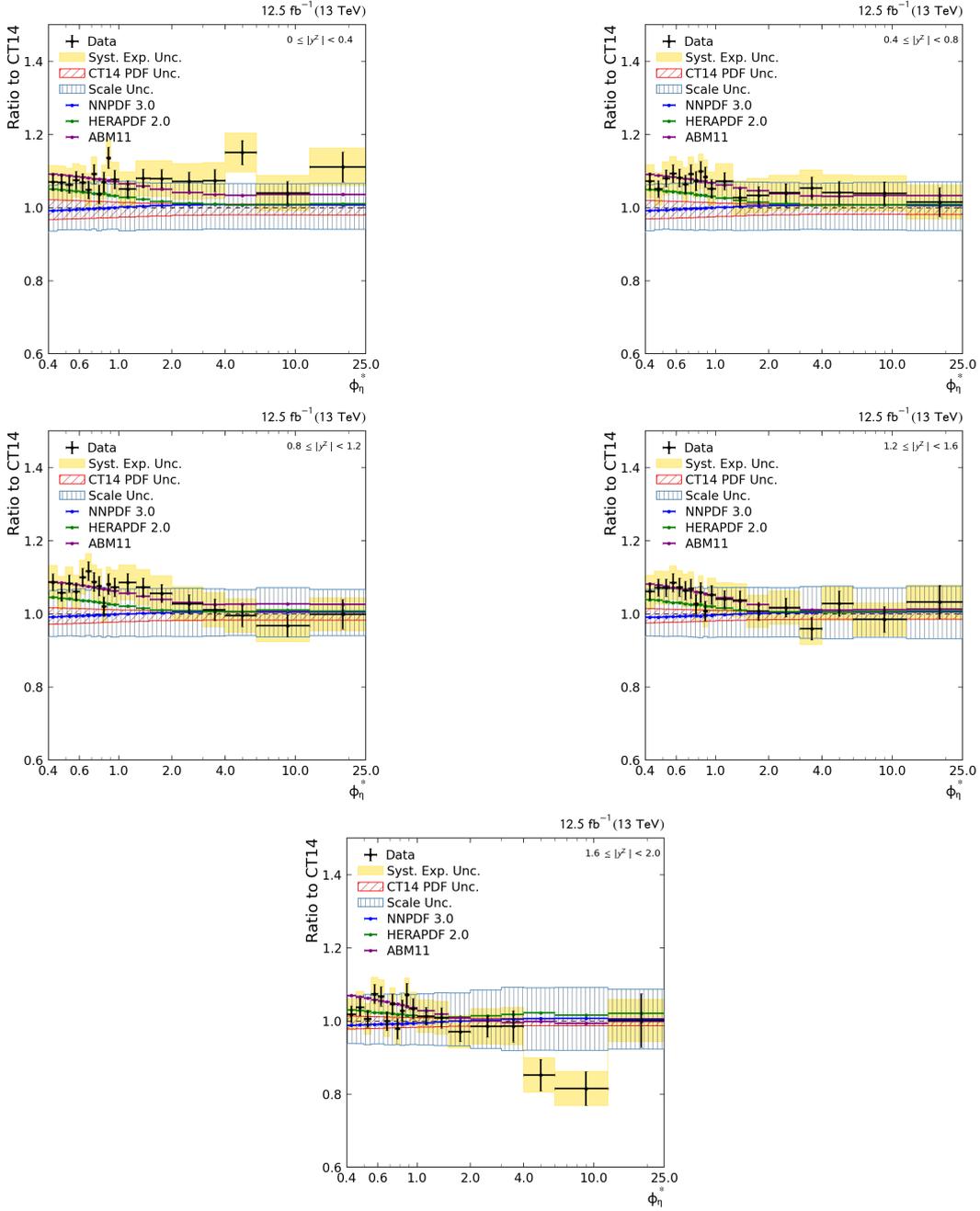
**Figure 6.2:** Double-differential cross section as a function of  $p_T^Z$  (left) and  $\phi_n^*$  (right) in bins of  $|y^Z|$ . Both measurements are compared to theory predictions at NLO evaluated with the CT14 PDF set.

## 6 Comparison of Measurement and Theory Predictions and PDF Determination



**Figure 6.3:** Ratio to theory predictions as a function of  $p_T^Z$  in bins of  $|y^Z|$  obtained with the CT14 PDF set for data and NNPDF 3.0, HERAPDF 2.0 and ABM11 PDF sets. PDF and scale uncertainties are shown separately as hatched areas. The error bars show the statistical and FSR uncertainties added in quadrature. The total systematic uncertainty (FSR excluded) on the measurement is shown as yellow squares.

## 6.1 Comparison with NLO Predictions



**Figure 6.4:** Ratio to theory predictions as a function of  $\phi_{\eta}^*$  in bins of  $|y^Z|$  obtained with the CT14 PDF set for data and NNPDF 3.0, HERAPDF 2.0 and ABM11 PDF sets. PDF and scale uncertainties are shown separately as hatched areas. The error bars show the statistical and FSR uncertainties added in quadrature. The total systematic uncertainty on the measurement (FSR excluded) is shown as yellow squares.

## 6.2 PDF Determination

The xFITTER [74] framework was used to determine the proton PDFs from the comparison between data and theory predictions. In addition to the  $Z(\rightarrow \mu\mu) + \text{jets}$  cross section measurements, DIS data from the HERA I+II experiments were included into the fits. The DIS cross section predictions were calculated with QCDNUM [75] at NLO.

The PDF fitting was performed with a configuration similar to the one used for the determination of the HERAPDF 2.0 PDF set [76] as well as for studies with CMS Z+jet data at  $\sqrt{s} = 8 \text{ TeV}$  [59].

The parametrisation of the PDFs is chosen as follows:

$$xg(x) = A_g x^{B_g} (1-x)^{C_g} - A'_g x^{B'_g} (1-x)^{C'_g} \quad (6.1)$$

$$xu_v(x) = A_{u_v} x^{B_{u_v}} (1-x)^{C_{u_v}} (1 + E_{u_v} x^2) \quad (6.2)$$

$$xd_v(x) = A_{d_v} x^{B_{d_v}} (1-x)^{C_{d_v}} \quad (6.3)$$

$$x\bar{U}(x) = A_{\bar{U}} x^{B_{\bar{U}}} (1-x)^{C_{\bar{U}}} (1 + D_{\bar{U}} x) \quad (6.4)$$

$$x\bar{D}(x) = A_{\bar{D}} x^{B_{\bar{D}}} (1-x)^{C_{\bar{D}}} \quad (6.5)$$

Here,  $g$  refers to the gluon,  $u_v$  and  $d_v$  to the valence quarks, and  $\bar{U}$ ,  $\bar{D}$  to the up- and down-type antiquarks. Relations between the parameters guarantee the physical validity of the PDFs, and hence reduce the number of free parameters:

- The strange quark PDF is assumed to be proportional to the antidown quark PDF with a strangeness fraction of  $f_s = 0.4$ .
- At low  $x$ , the behaviour of up- and down-type antiquarks is assumed to be equal and therefore  $B_{\bar{U}} = B_{\bar{D}}$  is enforced.
- For similar reasons, the normalisations of the sea quark PDFs are assumed to satisfy the relation  $A_{\bar{U}} = (1 - f_s)A_{\bar{D}}$ .
- The normalisation factors  $A$  for the gluon and valence quark PDFs are constrained by the quark flavour and momentum sum rules

$$\int dx (u_v(x) - \bar{U}(x)) = 2, \quad (6.6)$$

$$\int dx (d_v(x) - \bar{D}(x)) = 1, \text{ and} \quad (6.7)$$

$$\sum_i \int dx x f_i(x) = 1 \quad (6.8)$$

where  $i$  runs over all parton flavours. The sum rules follow from the assumption that the proton consists of one down and two up valence quarks.

- $C'_g$  is set to 25. This ensures that the negative gluon term only contributes at small  $x$ .

The parametrisation defines the PDFs at the starting scale of  $Q^2 = 1.9 \text{ GeV}^2$ . The DGLAP equations (see Section 2.2.2), implemented in QCDNUM, are used to evolve the PDF to the scale of the measurement.

The MINUIT [77] tool is used by xFITTER to perform the fit. As starting values for the parameters, the result of [76] is chosen. The fit is done minimising a least-squares  $\chi^2$  goodness of fit estimator

$$\chi^2 = \sum_{i,j} \left( D_i - T_i - \sum_k \beta_{ik} r_k \right) C_{ij}^{-1} \left( D_j - T_j - \sum_k \beta_{jk} r_k \right) + \sum_k r_k^2 + \sum_i \ln \frac{\Delta_{i,\text{stat}}^2 D_i T_i + \Delta_{i,\text{uncor}}^2 T_i^2}{(\Delta_{i,\text{stat}}^2 + \Delta_{i,\text{uncor}}^2) D_i^2} \quad (6.9)$$

with data points  $D_i$  and the corresponding theory predictions  $T_i$  as described in [70]. The correlated systematic uncertainties  $\beta_i$  are treated using nuisance parameters  $r_k$ . The logarithmic term is an additional correction that takes the transition of the Gaussian distribution to the  $\chi^2$  distribution into account. Thereby,  $\Delta_{i,\text{stat}}$  refers to the statistical uncertainty, while  $\Delta_{i,\text{uncor}}$  summarises other uncertainties that are uncorrelated between data points.

### 6.2.1 Estimation of Uncertainties on the PDFs

The uncertainty estimation follows the procedure developed by the HERAPDF collaboration [76].

There are three contributions to the total uncertainty on the PDFs.

#### Experimental Uncertainty

The uncertainties of the data used to constrain the PDFs are propagated to the PDFs using the Hessian method [78]. In general, the correlated PDF parameters are transformed into an eigenvector base. Then, the  $n$  eigenvalues are separately varied upwards and downwards until  $\chi^2 - \chi_{\text{min}}^2 = 1$ . This approach leads to  $2n$  different PDFs. The experimental PDF uncertainty is finally derived by adding the differences between the central fit and the PDF achieved with each varied eigenvalue in quadrature.

### Model Uncertainty

The theory predictions depend on several physical parameters with considerable uncertainties. In particular, the charm and bottom quark masses  $m^c$  and  $m^b$ , the strangeness fraction  $f_s$  and the value of the strong coupling constant at the Z boson mass  $\alpha_s(m^Z)$  were separately varied, in order to estimate the impact of those parameters on the PDF fit. In addition, an uncertainty arising from the lower limit on the momentum transfer in DIS data  $Q_{\min}^2$  is estimated.

Like the experimental uncertainty, the model uncertainty is obtained by adding the differences between the central fit and the each fit with a varied parameter in quadrature.

### Parametrisation Uncertainty

The parametrisation of the PDFs is not specified by theoretical considerations. Hence, an uncertainty arises from the choice of the parametrisation. It is taken into account by expanding the parametric form of the PDFs, leading to

$$xg(x) = A_g x^{B_g} (1-x)^{C_g} (1 + D_g x + E_g x^2) - A'_g x^{B'_g} (1-x)^{C'_g} \quad (6.10)$$

for the gluon and

$$xf(x) = A_f x^{B_f} (1-x)^{C_f} (1 + D_f x + E_f x^2) \quad (6.11)$$

for the quarks.

Additionally, the starting scale for the PDF evolution  $Q_0^2$  (see Section 2.2.2) is varied. The envelope obtained by combining the differences between all PDFs achieved this way and the central fit is chosen as the parametrisation uncertainty.

The total PDF uncertainty is obtained by adding the three contributions in quadrature.

All varied parameters, their central values and upwards and downwards variation values are listed in Table A.2.

## 6.2.2 Results of the PDF Fits

The PDF fits were performed using the double-differential cross section measurements discussed in Section 6.1. The  $\chi^2$  values obtained with and without the CMS Z( $\rightarrow \mu\mu$ ) + jets measurements included are given in Table 6.1.

For  $1.6 < |y^Z| < 2.0$ , a  $\chi^2/n_{\text{Data}}$  larger than two was observed in both measurements, indicating systematic discrepancies between data and theory in this area, which are not considered in the uncertainty estimation. Possible explanations for this

Dataset	$n_{\text{Data}}$	HERA $\chi^2/n_{\text{Data}}$	+CMS( $p_{\text{T}}^Z$ ) $\chi^2/n_{\text{Data}}$	+CMS( $\phi_{\eta}^*$ ) $\chi^2/n_{\text{Data}}$
HERAI+II CC $e^-$	42	1.18	1.16	1.15
HERAI+II CC $e^+$	39	1.09	1.24	1.27
HERAI+II NC $e^-$	159	1.39	1.39	1.39
HERAI+II NC $e^+$ $E_p = 460$	204	1.03	1.04	1.03
HERAI+II NC $e^+$ $E_p = 575$	254	0.83	0.85	0.84
HERAI+II NC $e^+$ $E_p = 820$	70	0.93	0.96	0.96
HERAI+II NC $e^+$ $E_p = 920$	377	1.10	1.13	1.12
CMS Z+jet $ y^Z  < 0.4$	15 19	-	1.44	1.76
CMS Z+jet $0.4 \leq  y^Z  < 0.8$	15 19	-	0.79	0.36
CMS Z+jet $0.8 \leq  y^Z  < 1.2$	15 19	-	0.87	1.13
CMS Z+jet $1.2 \leq  y^Z  < 1.6$	15 19	-	0.96	0.59
CMS Z+jet $1.6 \leq  y^Z  < 2.0$	15 19	-	3.20	2.08

Dataset	$n_{\text{dof}}$	$\chi^2/n_{\text{dof}}$	$\chi^2/n_{\text{dof}}$	$\chi^2/n_{\text{dof}}$
HERA I+II	1131	1.17	-	-
HERA I+II & CMS Z+jet ( $p_{\text{T}}^Z$ )	1206	-	1.20	-
HERA I+II & CMS Z+jet ( $\phi_{\eta}^*$ )	1226	-	-	1.19

**Table 6.1:**  $\chi^2$  between data and theory in the central fit for HERA I+II datasets only and in combined fits with CMS  $Z(\rightarrow \mu\mu) + \text{jets}$  data using the double-differential cross sections as a function of  $p_{\text{T}}^Z$  and  $\phi_{\eta}^*$ , respectively, both in bins of  $|y^Z|$ . The number of data points differs between the  $p_{\text{T}}^Z$  and  $\phi_{\eta}^*$  related analyses. In the case of two values in the second column, the first one corresponds to the fourth column and the second to the fifth column.

mismatch are discussed in Section 6.1. Especially the assumption of a fully correlated reconstruction uncertainty could be a reason for the high  $\chi^2/n_{\text{Data}}$  in this phase space region, as discussed in Section 4.7. For  $|y^Z| < 1.6$ , the  $\chi^2/n_{\text{Data}}$  is in the order of one, which indicates the compatibility between HERA I+II and CMS  $Z(\rightarrow \mu\mu) + \text{jets}$  data within their uncertainties. Including the measurement as a function of  $\phi_{\eta}^*$  leads to a lower  $\chi^2/n_{\text{dof}}$  than the fit with the  $p_{\text{T}}^Z$  related measurement.

The resulting PDFs for the HERA I+II-only fit and the combined fit using the CMS measurement as a function of  $p_{\text{T}}^Z$  are shown in the Figures 6.5 and 6.6. The sea quark PDF is defined as  $2x \cdot (\bar{U} + \bar{D})$  [76]. This definition implies the assumption that the quark and antiquark distributions are equal for the sea quarks.

The different uncertainty contributions explained in Section 6.2.1 are shown separately. In addition, a comparison of the total PDF uncertainties with and without the CMS data can be seen in Figure 6.7.

The inclusion of CMS  $Z(\rightarrow \mu\mu) + \text{jets}$  data leads to a clear reduction of the parametrisation uncertainty for all flavours. The most significant effect can be observed for the gluon PDF at medium  $x$ , where the measurement is expected to be most sensitive to the PDFs (see Section 5.2). Furthermore, the hump on the parametrisation uncertainty for the sea quark PDF around  $x = 10^{-2}$  is no longer present.

The clear reduction of the parametrisation uncertainty of the valence quarks (especially the u quark) is surprising, since the correlations between the  $Z(\rightarrow \mu\mu) + \text{jets}$  cross section predictions and the PDFs are small for valence quarks (see 5.2). The reason could be the correlation between the parameters of all PDF flavours. Hence, constraints on the gluon and sea quark PDF indirectly lead to a lower parametrisation dependency of the other PDFs.

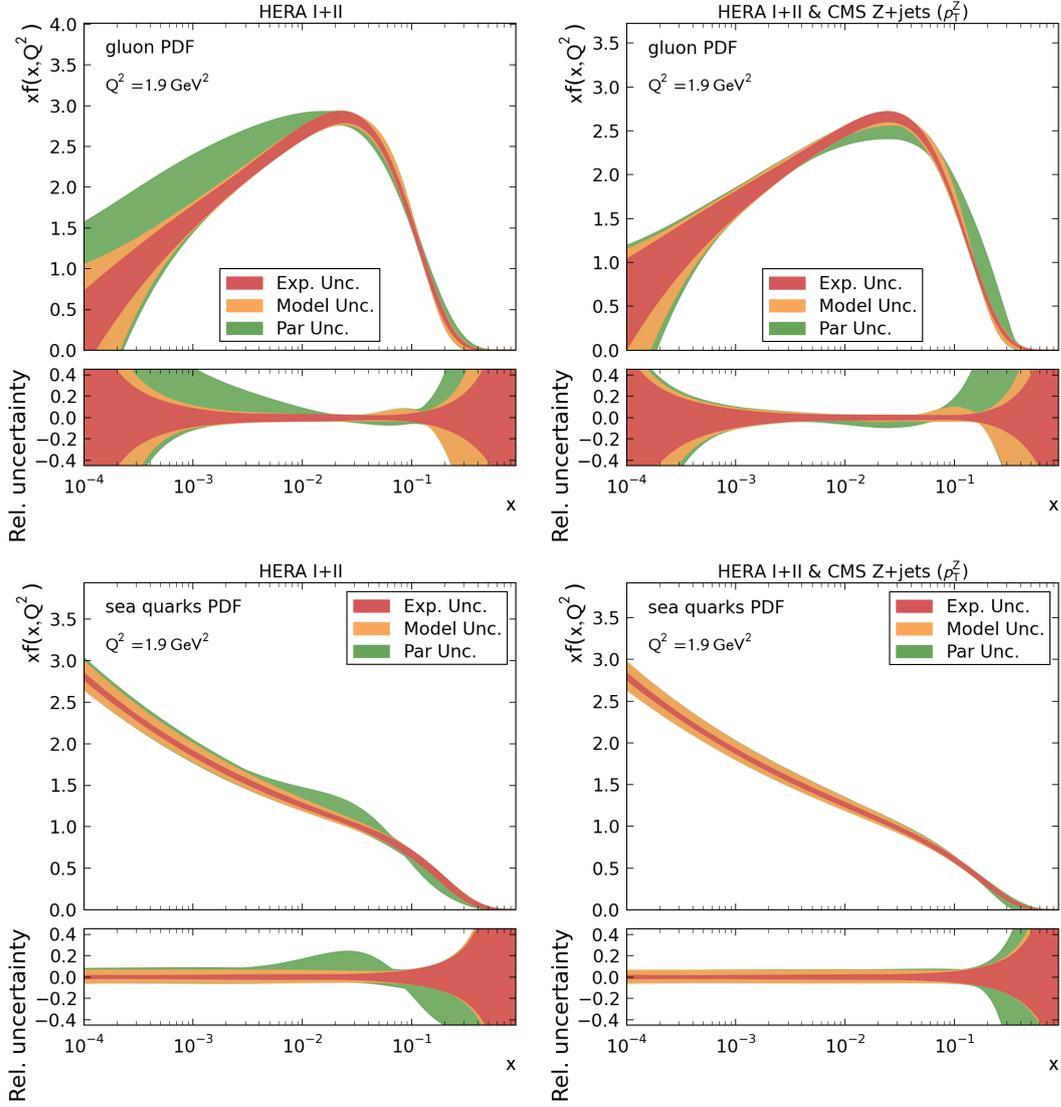
At high  $x$ , the parametrisation uncertainty on the gluon and d valence quark PDF increases for the fit with CMS data included. This area is constrained only by extrapolation. Hence, differences between HERA and CMS data can even increase the uncertainty in this phase space region.

The experimental uncertainty also decreases slightly in most areas, due to the additional data points.

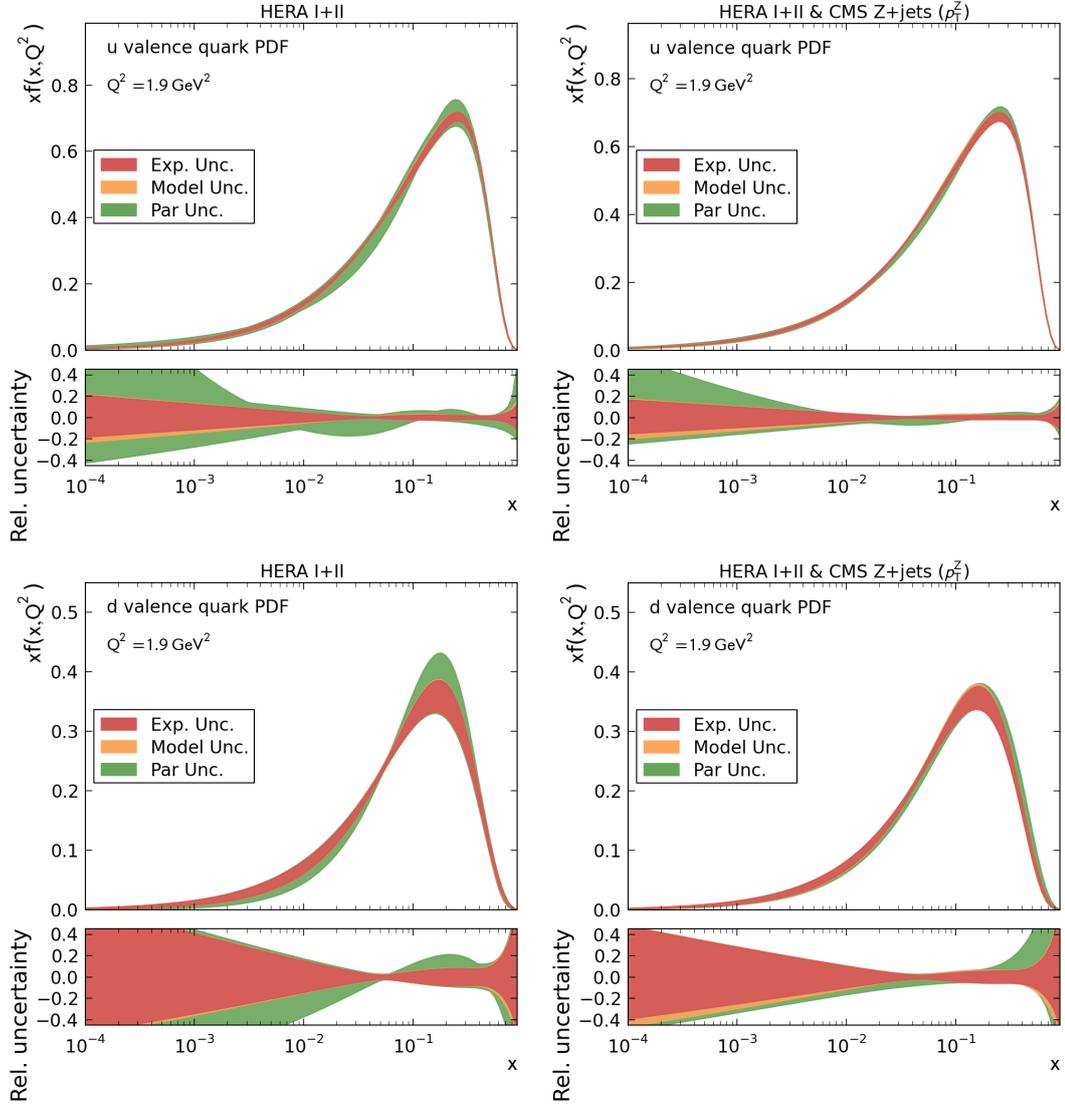
Changes concerning the central value of the fit can be observed primarily for the gluon PDF. Here, the peak at medium  $x$  decreases in the combined fit, while higher values of the gluon contribution are visible for low and high  $x$  values. This trend is also observed in other CMS studies, e.g. the inclusive jet measurement at 8 TeV [79].

In Figure 6.8, the PDFs obtained with the  $p_T^Z$  and  $\phi_\eta^*$  related measurements are compared to each other. The obtained PDFs are similar, apart from small differences concerning the gluon PDF. The results show that  $\phi_\eta^*$  is a reasonable choice as observable for the purpose of Z+jet studies and measurements of the proton structure. The lower  $\chi^2/n_{\text{dof}}$  along with the visual impressions obtained in Section 6.1 shows the better theoretical description of the measurement as a function of  $\phi_\eta^*$  compared to the  $p_T^Z$  related measurement. Regardless, the uncertainties on the PDFs are almost equal for both measurements. In summary, consistent results are obtained for  $p_T^Z$  and  $\phi_\eta^*$ . The contributions of the different PDF uncertainties obtained for the  $\phi_\eta^*$  related measurement are shown in Figure A.11.

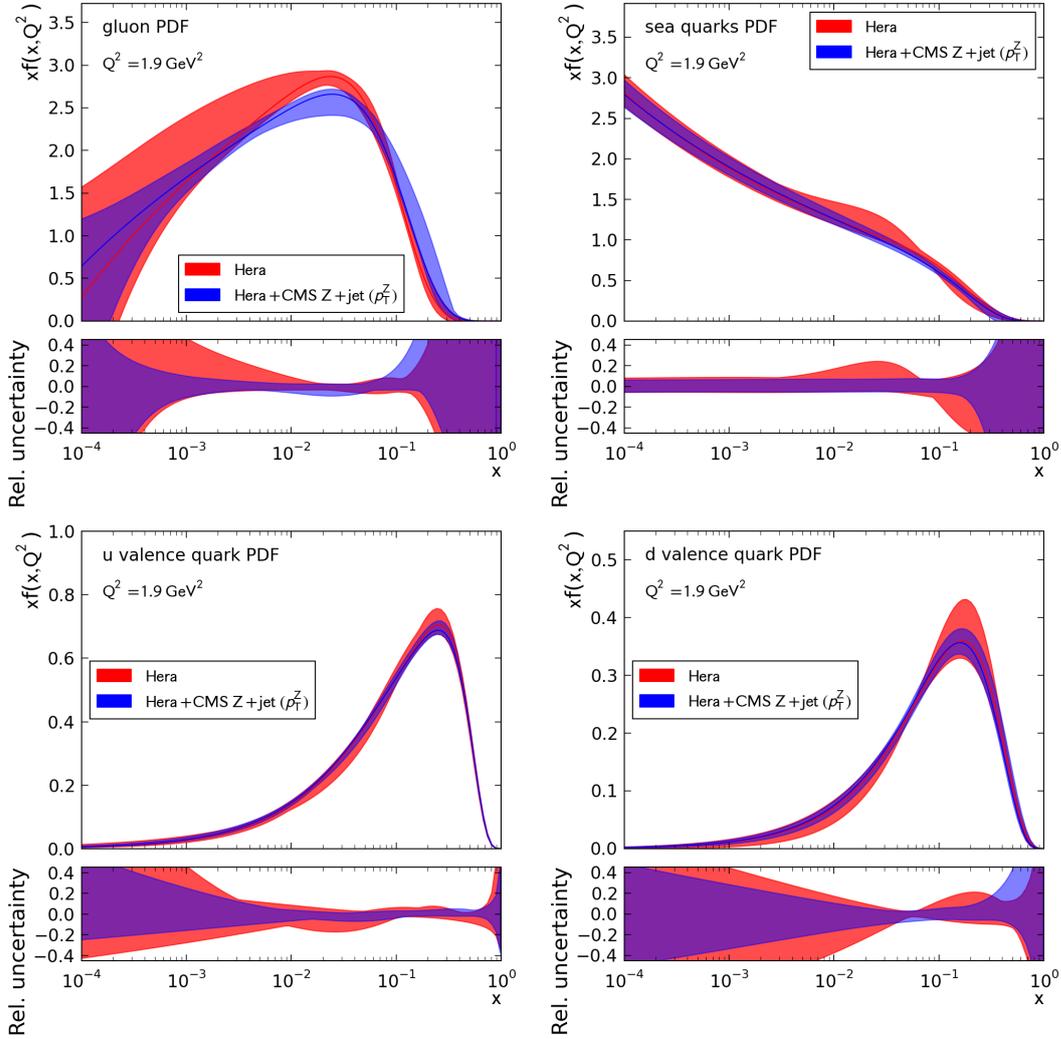
The parameter values obtained in the central fits of both measurements are given in Table A.3.



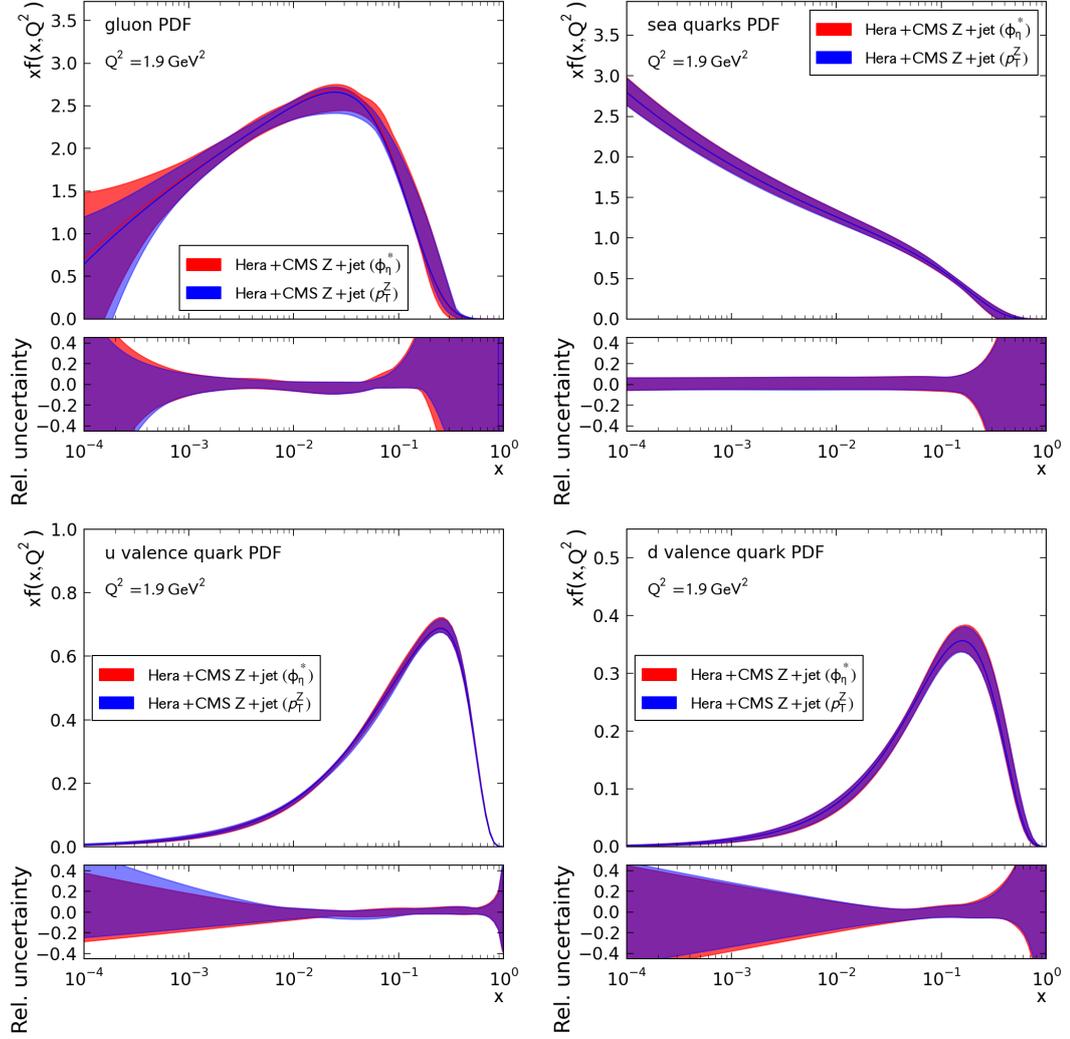
**Figure 6.5:** Gluon (top) and sea quark (bottom) PDFs obtained with HERA I+II data (left) and HERA I+II combined with the  $Z(\rightarrow \mu\mu) + \text{jets}$  double-differential cross section measurement as a function of  $p_T^Z$  and  $|y^Z|$  (right). The PDFs were evaluated at the starting scale of  $Q^2 = 1.9 \text{ GeV}^2$ .



**Figure 6.6:** Valence quark PDFs for the up (top) and down (bottom) quarks obtained with HERA I+II data only (left) and combined with the  $Z(\rightarrow \mu\mu) + \text{jets}$  double-differential cross section measurement as a function of  $p_T^Z$  in bins of  $|y^Z|$  (right). The PDFs were evaluated at the starting scale of  $Q^2 = 1.9 \text{ GeV}^2$ .



**Figure 6.7:** Comparison of the PDFs obtained with HERA I+II data only and HERA I+II and CMS data combined, using the double-differential cross section measurement as a function of  $p_T^Z$ . Shown are the gluon (top left), sea quark (top right), u valence quark (bottom left) and d valence quark (bottom right) PDFs evaluated at  $Q^2 = 1.9 \text{ GeV}^2$ .



**Figure 6.8:** Comparison of the PDFs obtained with HERA I+II data combined with the double-differential  $Z(\rightarrow \mu\mu) + \text{jets}$  cross section measurements as functions of  $p_{\text{T}}^{\text{Z}}$  and  $\phi_{\eta}^*$ , respectively. Both are measured in bins of  $|y^{\text{Z}}|$ . Shown are the gluon (top left), sea quark (top right), u valence quark (bottom left) and d valence quark (bottom right) PDFs evaluated at  $Q^2 = 1.9 \text{ GeV}^2$ .

---

## Conclusion

In this thesis, a subset of the data recorded by the CMS detector in 2016 at  $\sqrt{s} = 13$  TeV has been analysed. These data correspond to an integrated luminosity of  $12.5 \text{ fb}^{-1}$ , offering a statistical precision comparable to the entire data recorded at  $\sqrt{s} = 8$  TeV. This allows the probing of QCD predictions at an energy scale unprecedented in collider physics.

For this purpose, the  $Z(\rightarrow \mu\mu) + \text{jets}$  cross section was measured double-differentially as a function of dimuon transverse momentum  $p_{\text{T}}^Z$  and absolute rapidity  $|y^Z|$ . In addition, an observable correlated to  $p_{\text{T}}^Z$  named  $\phi_{\eta}^*$  was studied. The cross section was measured as a function of  $\phi_{\eta}^*$  and  $|y^Z|$  in order to compare  $\phi_{\eta}^*$  and  $p_{\text{T}}^Z$  with regard to proton structure studies.

The event yields were corrected for trigger and reconstruction efficiencies and the background processes were analysed and subtracted from the data. Afterwards, the data were unfolded. In this context, it was demonstrated that migration effects between the bins were much smaller for  $\phi_{\eta}^*$  than for  $p_{\text{T}}^Z$ . Thus, the dependence on the detector simulation and the unfolding algorithm can be minimised in  $Z + \text{jet}$  studies by using  $\phi_{\eta}^*$  instead of the more common  $p_{\text{T}}^Z$ .

Theory predictions were calculated with SHERPA and BLACKHAT at NLO, evaluated with different PDF sets using FASTNLO and compared to the data. It was shown that the studied  $Z(\rightarrow \mu\mu) + \text{jets}$  cross section predictions are sensitive to the PDFs, leading to variations of up to 10% among the predictions obtained for different PDF sets. The data are well described by the theory predictions in almost the entire phase space analysed.

To study the influence of the measurement on the PDF determination, PDF fits were performed using both double-differential cross section measurements separately, each combined with data from the HERA I+II experiments. For this purpose, the xFITTER framework was used. Both combined fits led to similar PDFs with a slightly better goodness-of-fit result for the  $\phi_{\eta}^*$  related measurement. Compared to fits with only the HERA data, a reduction of the uncertainties on all PDF flavours and a shift in the central value of the gluon PDF was obtained. In summary, the results of the

fits show that the inclusion of  $Z$ +jet cross section measurements at  $\sqrt{s} = 13$  TeV can improve the PDF sets.

A few opportunities for further improvements of the analysis were identified: It was revealed in the analysis that the muon final state radiation (FSR) has a major influence on the measurement, due to the tight event selection. In this thesis, a preliminary correction was applied. Additional studies on this topic, for instance developing an algorithm to add the momentum of nearby photons to the muon momentum, could enhance the procedure and may lead to a lower uncertainty on the FSR effect.

The reconstruction efficiency introduces the largest experimental uncertainty for the used dataset. A more precise efficiency calculation was recently released for the full dataset recorded in 2016. An analysis on the complete dataset thus could benefit from this additional precision.

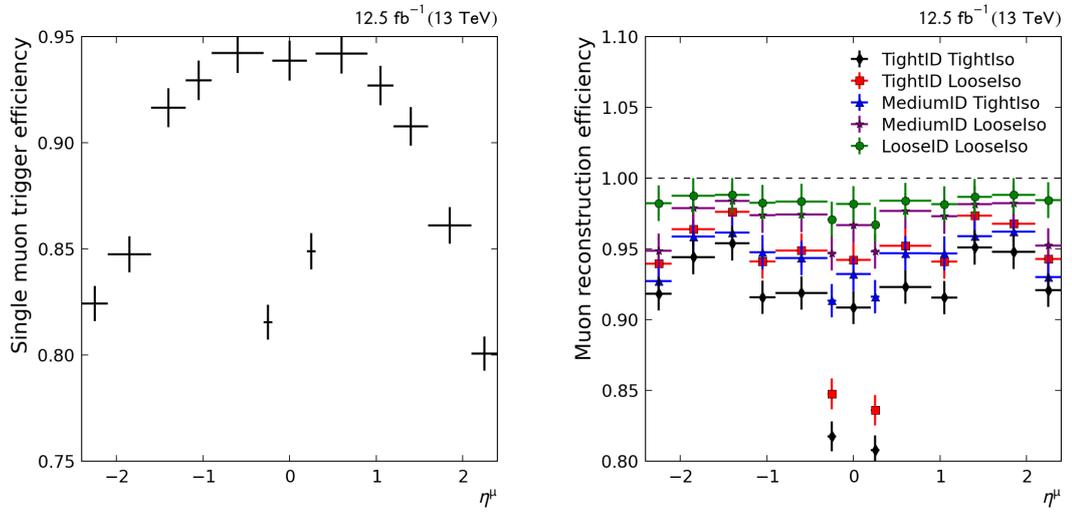
Among theoretical systematics, the scale uncertainty is dominant at NLO. Theory predictions at NNLO are needed to achieve a higher accuracy. They will be available in the near future and may improve the PDF fits.

## Appendix A

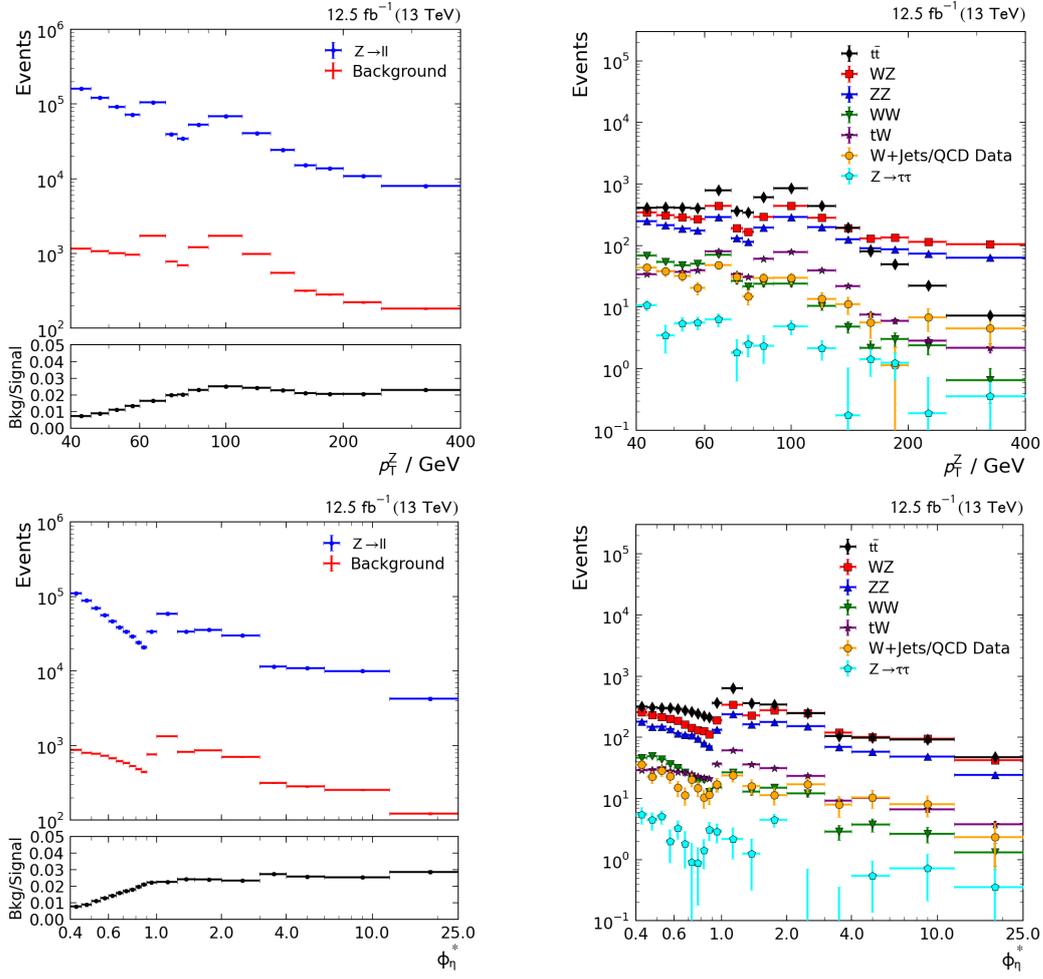
# Appendix

Dataset	Events
<b>Data:</b>	
SingleMuon/Run2016B-PromptReco-v2/MINIAOD	158188719
SingleMuon/Run2016C-PromptReco-v2/MINIAOD	68492270
SingleMuon/Run2016D-PromptReco-v2/MINIAOD	98175265
<b>Signal Monte Carlo sample:</b>	
DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/[Spring16]	28611654
<b>Background Monte Carlo samples:</b>	
TTJets_DiLept_TuneCUETP8M1_13TeV-madgraphMLM_pythia8/[Summer16]	24350202
WZ_TuneCUETP8M1_13TeVpythia8/[Summer16]	2995828
ZZ_TuneCUETP8M1_13TeV-pythia8/[Summer16]	998034
WW_TuneCUETP8M1_13TeV-pythia8/[Summer16]	6987124
ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1/[Summer16]	6952830
ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1/[Summer16]	6933094
DYJetsToTauTau_ForcedMuEleDecay_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8_ext1/[Summer16]	25649546
[Spring16] RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	
[Summer16] RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TracheIV_v6_ext1-v1/MINIAODSIM	
<b>HLT Path: HLT_IsoMu22 or HLT_IsoTkMu22</b>	

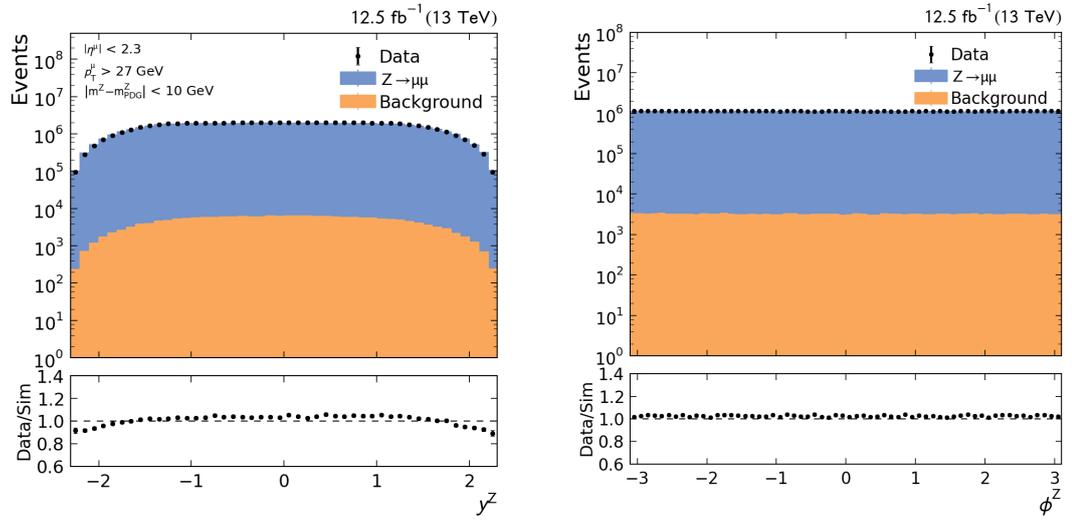
**Table A.1:** CMS internal names of the datasets and HLT paths used in this analysis.



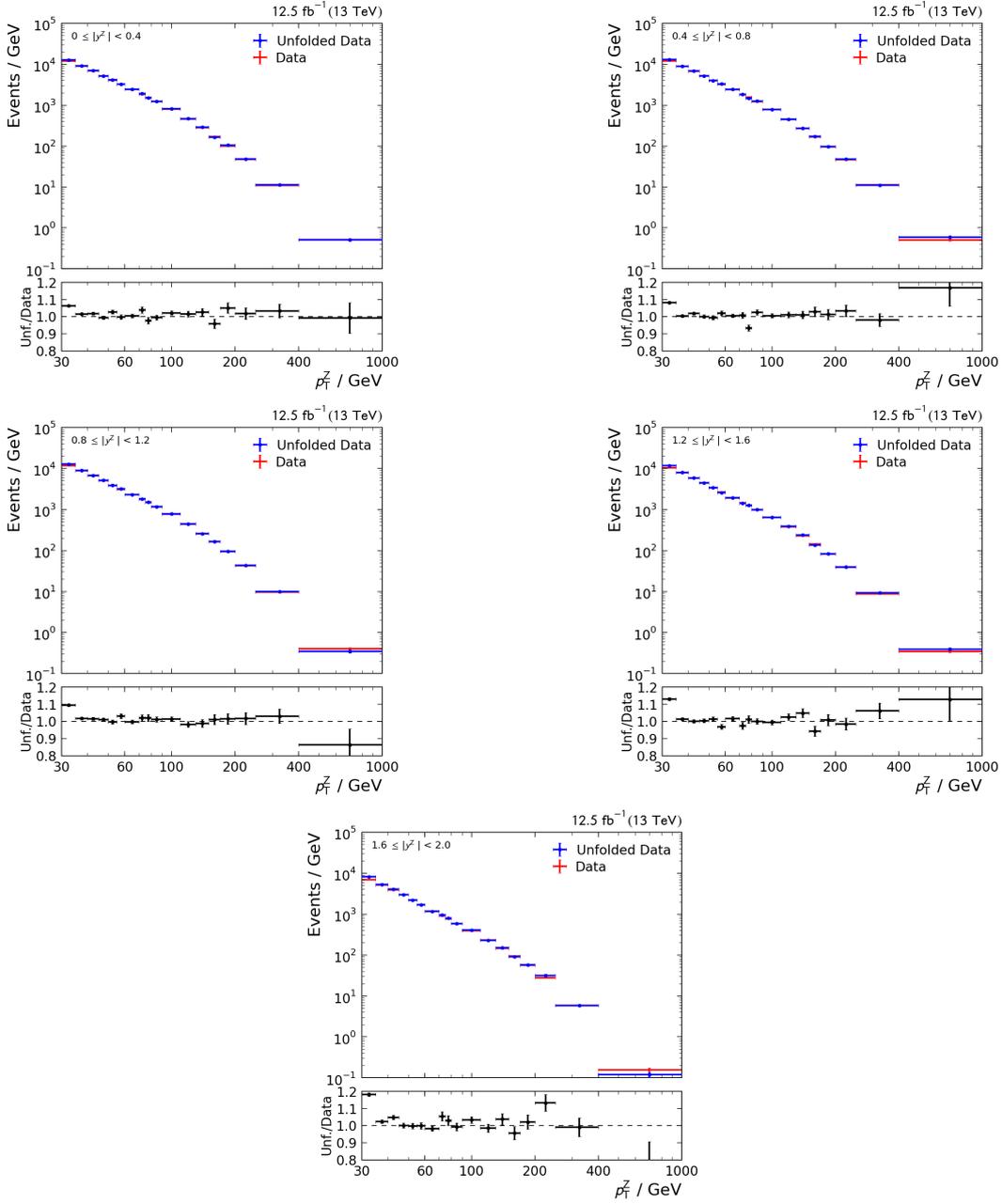
**Figure A.1:** Efficiency of the used HLT paths (left) and the reconstruction efficiency for the provided combinations of muon ID and isolation (right) as a function of  $|\eta^\mu|$ . The lower efficiency in the small bin around the centre of the detector is caused by the detector geometry.



**Figure A.2:** Estimated background to signal fraction (left) and contribution of the different background processes (right). Both as functions of  $p_T^Z$  (top) and  $\phi_\eta^*$  (bottom).

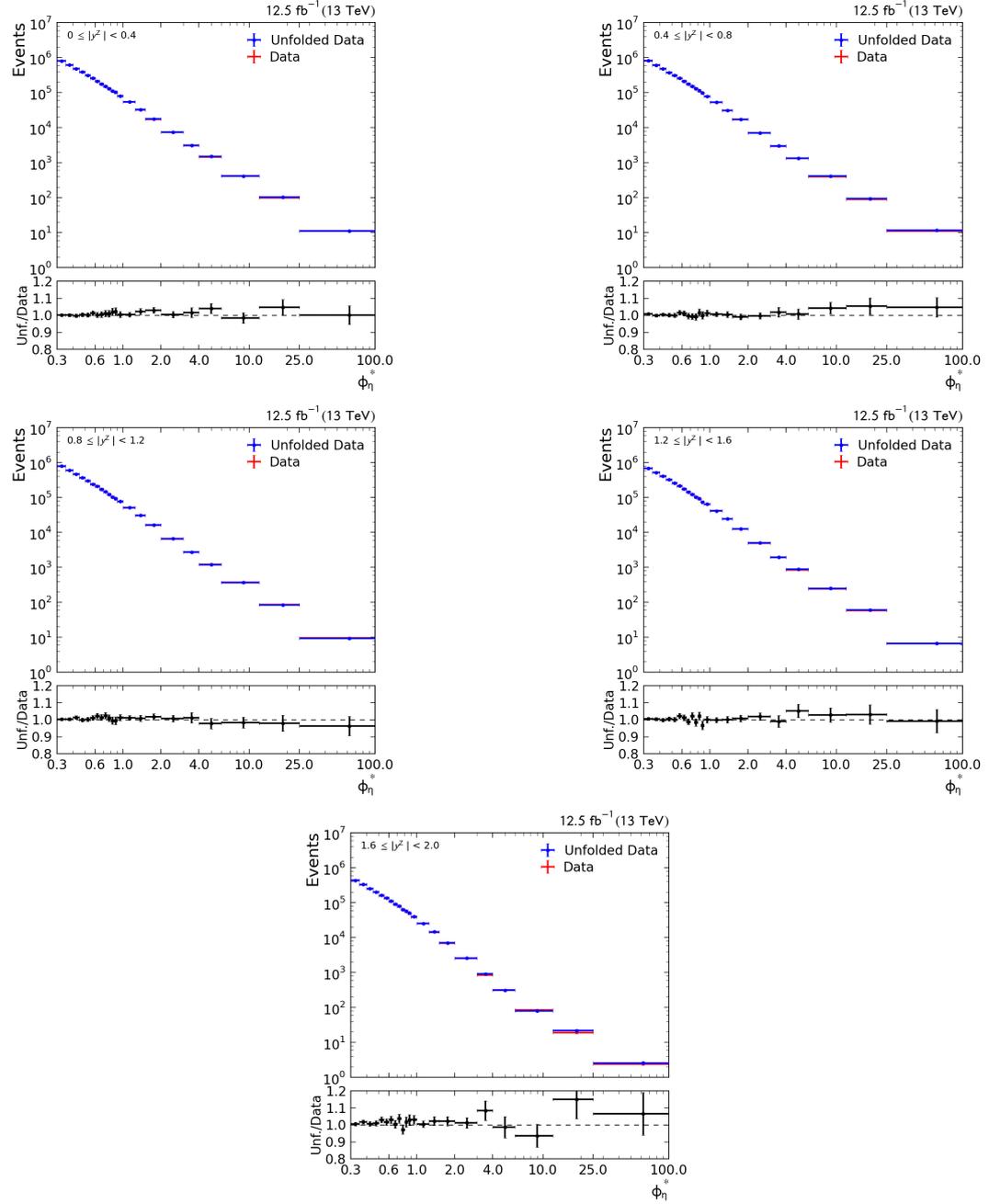


**Figure A.3:** Comparison of detector level data and simulation as a function of rapidity (left) and azimuthal angle (right) of the dimuon system.

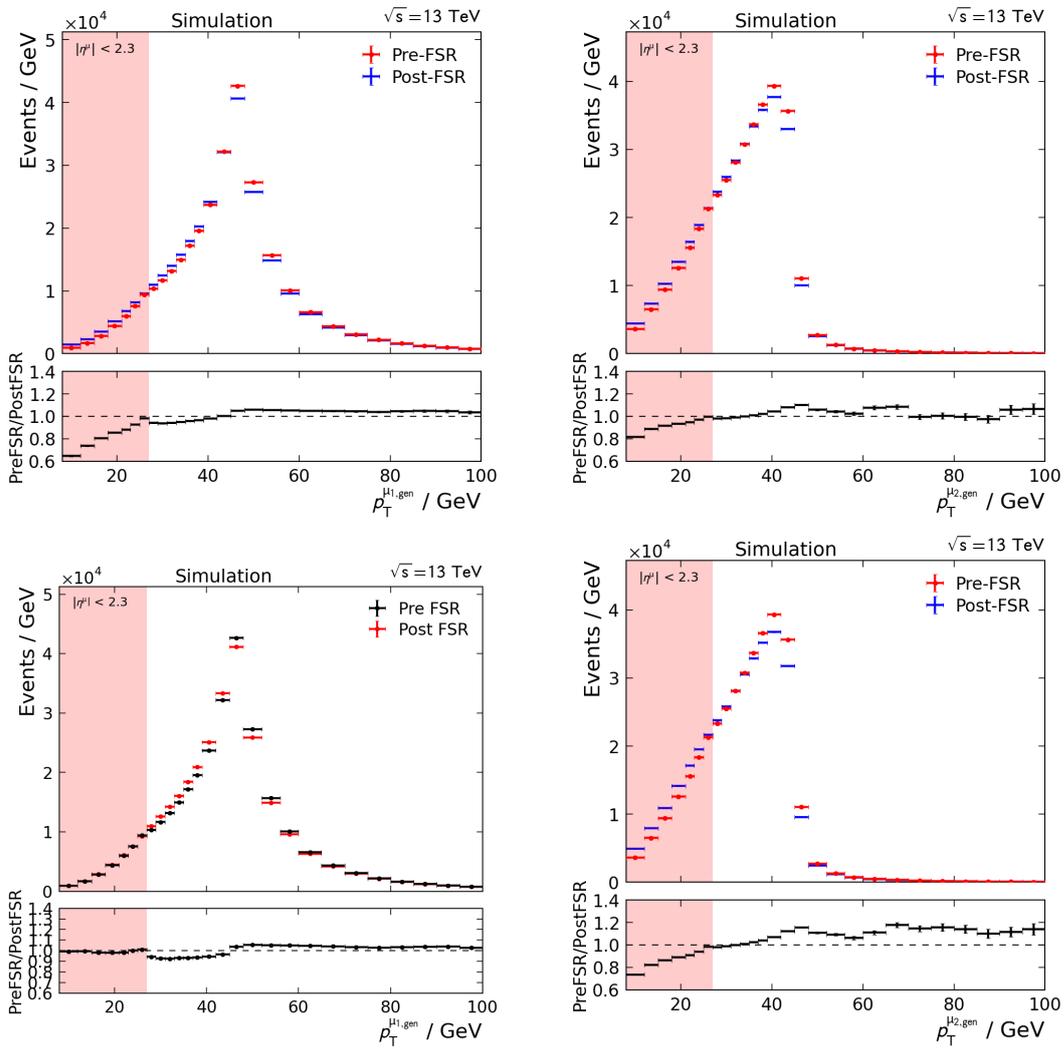


**Figure A.4:** Comparison between the unfolded data and the background subtracted data as a function of  $p_T^Z$  in the particular  $|y^Z|$  bins. The unfolding is performed with respect to  $p_T^Z$  in the particular  $|y^Z|$  region. The phase space in the final cross section measurement is restricted to  $40 < p_T^Z < 400$ .

## A Appendix

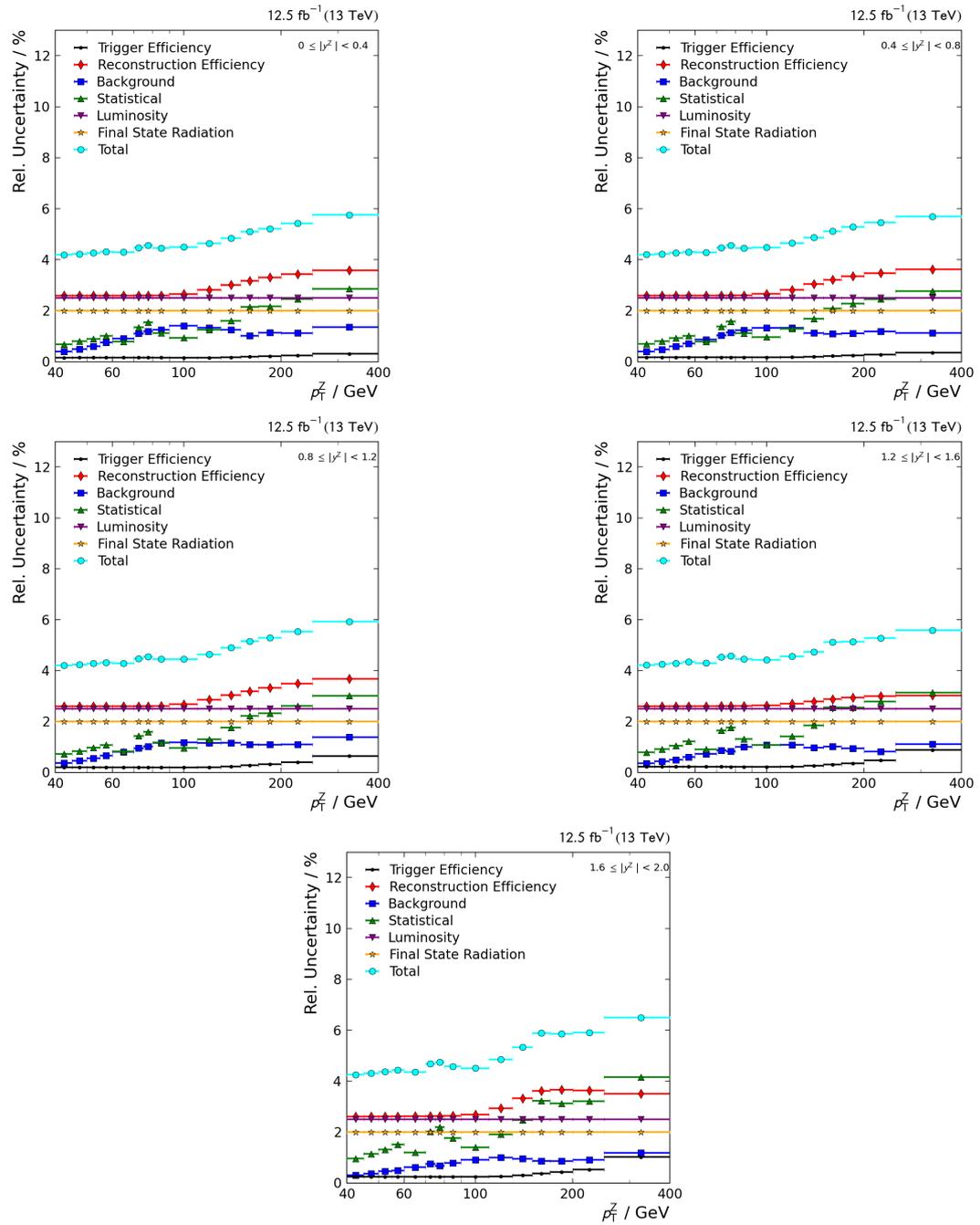


**Figure A.5:** Comparison between the unfolded data and the background subtracted data before unfolding as a function of  $\phi_n^*$  in the particular  $|y^Z|$  bins. The unfolding is performed with respect to  $\phi_n^*$  in the particular  $|y^Z|$  region. The phase space in the final cross section measurement is restricted to  $0.4 < \phi_n^* < 25$ .

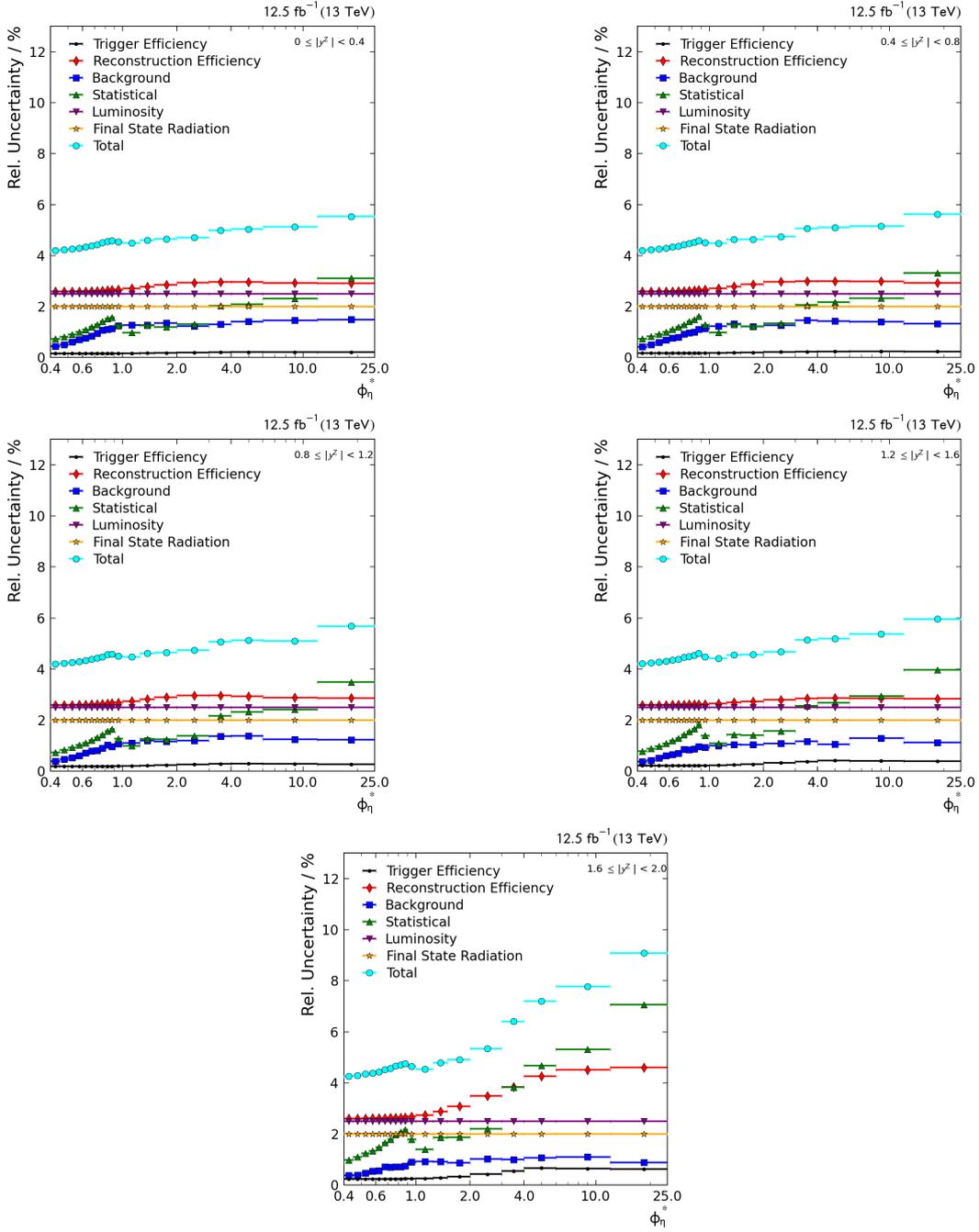


**Figure A.6:** Muon  $p_T$  comparison at generator level before and after final state radiation (FSR) for the leading and subleading muon. In the first row, the muon order is determined before FSR in both distribution. The second row shows possibly reordered muons after FSR. The red area is excluded in the cross section measurements.

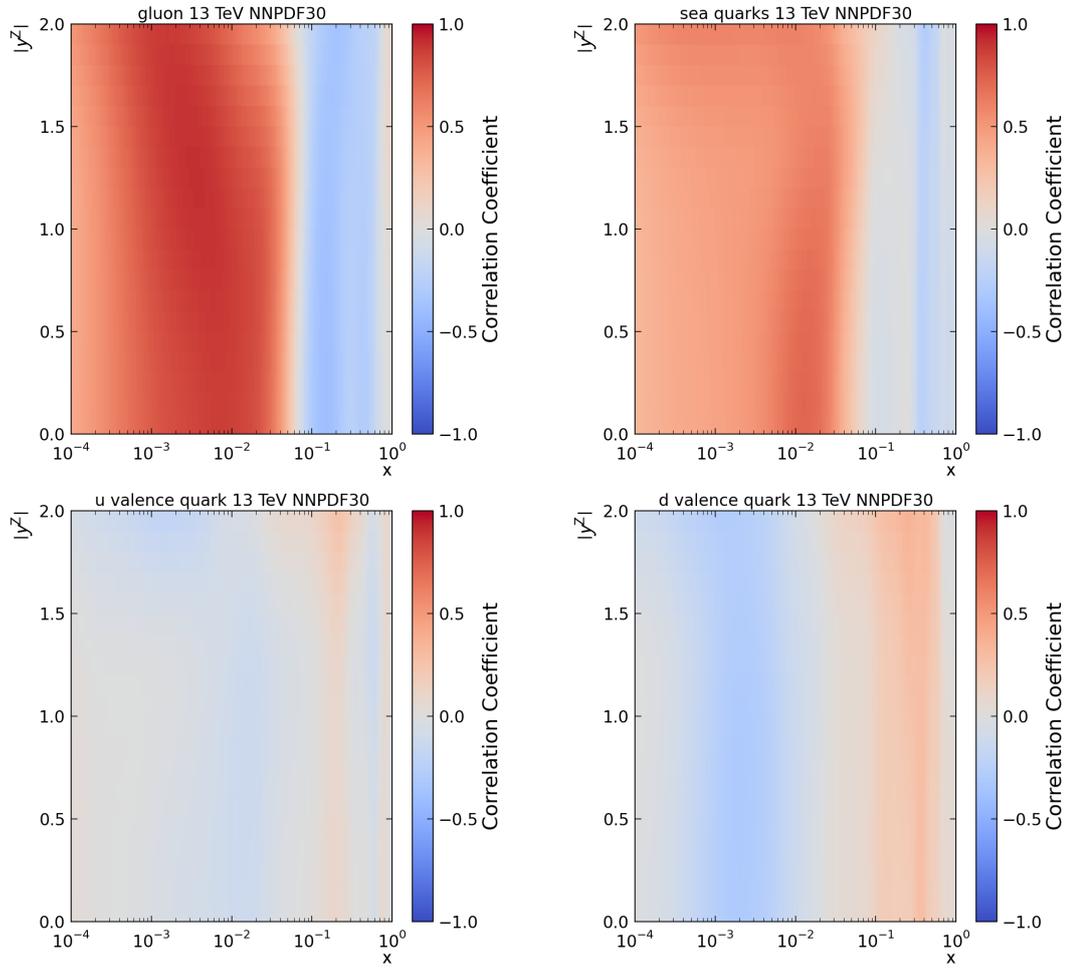
## A Appendix



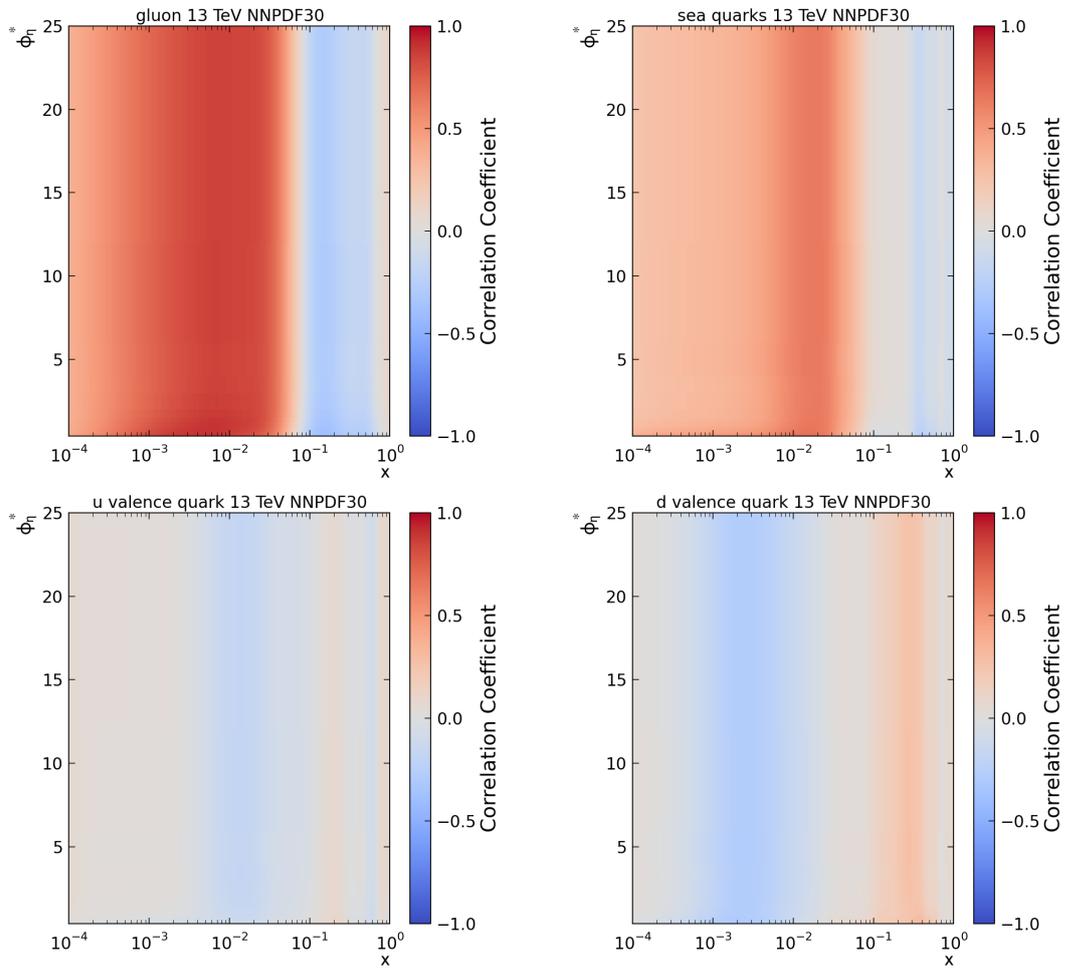
**Figure A.7:** Uncertainties on the measured cross section as a function of dimuon  $p_T$  in the five bins of the absolute dimuon rapidity.



**Figure A.8:** Uncertainties on the measured cross section as a function of  $\phi_n^*$  in the five bins of the absolute dimuon rapidity.



**Figure A.9:** Correlation coefficients between PDFs for different flavours and  $Z(\rightarrow \mu\mu) + \text{jets}$  cross section predictions as a function of  $x$  and dimuon rapidity calculated with NNPDF 3.0.



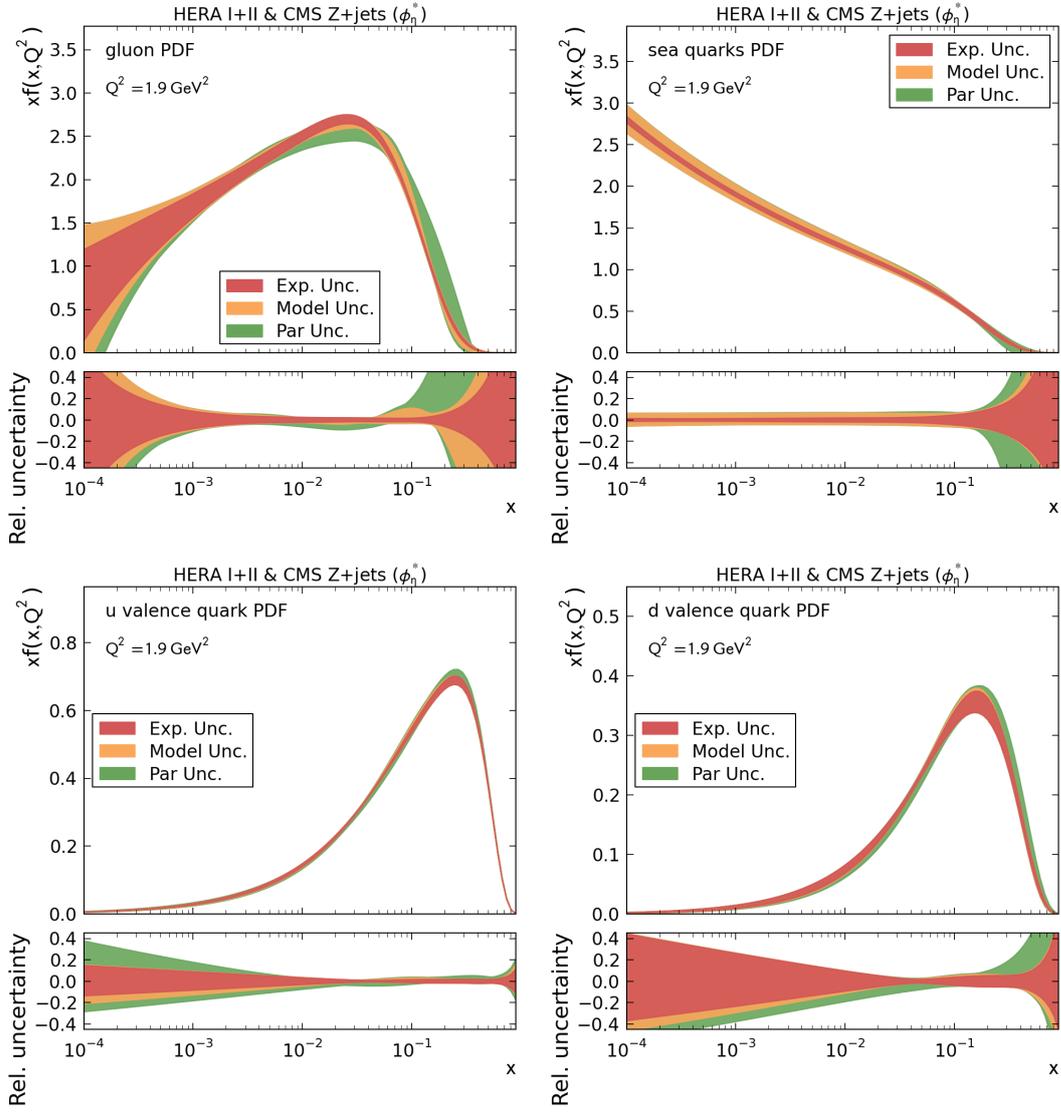
**Figure A.10:** Correlation coefficients between PDFs for different flavours and  $Z(\rightarrow \mu\mu) + \text{jets}$  cross section predictions as a function of  $x$  and  $\phi_n^*$  calculated with NNPDF 3.0.

Parameter	Central Value	Down Variation	Up Variation
$m^c / \text{GeV}$	1.47	1.41	1.53
$m^b / \text{GeV}$	4.5	4.25	4.75
$f_s$	0.4	0.3	0.5
$\alpha_s(m^Z)$	0.118	0.117	0.119
$Q_{min}^2 / \text{GeV}^2$	3.5	2.5	7.5
$Q_0^2 / \text{GeV}^2$	1.9	1.6	2.2

**Table A.2:** Parameters varied to obtain PDF uncertainties.

	$A$	$B$	$C$	$D$	$E$	$A'$	$B'$
$g$		-0.03(-0.05)	9.66(10.09)			1.76(2.18)	-0.15(-0.12)
$u_v$		0.72(0.72)	4.73(4.71)		12.24(11.71)		
$d_v$		0.82(0.82)	4.37(4.82)				
$\bar{U}$			6.92(6.80)	10.12(10.12)			
$\bar{D}$	0.19(0.19)	-0.17(-0.17)	6.53(6.69)				

**Table A.3:** PDF parameters obtained for the combined fit of HERA I+II and CMS Z+jet data. The values correspond to the double-differential measurement as a function of  $p_T^Z$  ( $\phi_\eta^*$ ), both in bins of  $|y^Z|$ .



**Figure A.11:** PDFs for the gluon (top left), sea quarks (top right), up valence quark (bottom left) and down valence quark (bottom right) obtained with HERA I+II data combined with the  $Z(\rightarrow \mu\mu) + \text{jets}$  cross section measurement as a function of  $\phi_\eta^*$  in bins of  $|y^Z|$ . The PDFs were evaluated at the starting scale of  $Q^2 = 1.9 \text{ GeV}^2$ .



---

## List of Figures

2.1	Z( $\rightarrow \mu\mu$ ) + jets Feynman diagrams . . . . .	5
2.2	Lepton pair invariant mass spectrum . . . . .	7
2.3	Running of $\alpha_s$ . . . . .	8
2.4	Parton splitting . . . . .	9
2.5	Experiments contributing to PDFs . . . . .	10
2.6	Vertex resolution at different energy scales . . . . .	11
3.1	Location of the LHC . . . . .	14
3.2	Event topology for $\phi_\eta^*$ definition . . . . .	16
3.3	Slice through the CMS detector . . . . .	17
3.4	CMS detector components . . . . .	18
3.5	Event pipelining in ARTUS . . . . .	22
4.1	Trigger and reconstruction efficiency ( $p_T^\mu$ ) . . . . .	26
4.2	Same sign test and muon $p_T$ corrections . . . . .	28
4.3	Validation of efficiency weights . . . . .	28
4.4	Background contributions ( $m^Z$ ) . . . . .	31
4.5	Detector level comparison ( $p_T^\mu$ ) . . . . .	32
4.6	Detector level comparison ( $m^Z$ and $p_T^Z$ ) . . . . .	32
4.7	Correlation between $p_T^Z$ and $\phi_\eta^*$ and $\phi_\eta^*$ detector level comparison . . . . .	33
4.8	Response matrices ( $p_T^Z$ and $\phi_\eta^*$ ) . . . . .	35
4.9	Comparison data before/after unfolding (single differential) . . . . .	36
4.10	Comparison unfolding methods . . . . .	37
4.11	FSR effect ( $p_T^\mu$ and $m^Z$ ) . . . . .	38
4.12	FSR effect ( $p_T^Z$ , $\phi_\eta^*$ , $ y^Z $ ) . . . . .	39
4.13	Experimental uncertainties (single differential) . . . . .	41
5.1	Correlation between PDF and cross section predictions ( $p_T^Z$ ) . . . . .	46
5.2	PDF and scale uncertainty . . . . .	48
6.1	Single-differential cross section . . . . .	50

6.2	Double differential cross section (overview)	51
6.3	Double differential cross section in bins of $ y^Z $ ( $p_T^Z$ )	52
6.4	Double differential cross section in bins of $ y^Z $ ( $\phi_\eta^*$ )	53
6.5	Gluon and sea quark PDFs	59
6.6	Valence quark PDFs	60
6.7	PDF comparison HERA/HERA+CMS	61
6.8	PDF comparison $p_T^Z$ and $\phi_\eta^*$ measurements	62
A.1	Trigger and reconstruction efficiency ( $ \eta^\mu $ )	66
A.2	Bkg contribution ( $p_T^Z$ and $\phi_\eta^*$ )	67
A.3	Detector level comparison ( $y^Z$ and $\phi^Z$ )	68
A.4	Comparison data before/after unfolding ( $p_T^Z$ in bins of $ y^Z $ )	69
A.5	Comparison data before/after unfolding ( $\phi_\eta^*$ in bins of $ y^Z $ )	70
A.6	FSR ( $p_T^\mu$ reorder effect)	71
A.7	Experimental uncertainties ( $p_T^Z$ in bins of $ y^Z $ )	72
A.8	Experimental uncertainties ( $p_T^Z$ in bins of $ y^Z $ )	73
A.9	Correlation between PDF and cross section predictions ( $ y^Z $ )	74
A.10	Correlation between PDF and cross section predictions ( $\phi_\eta^*$ )	75
A.11	PDFs with $\phi_\eta^*$ measurement	77

---

## List of Tables

4.1	Binning for cross section measurement . . . . .	25
4.2	Background processes . . . . .	30
6.1	$\chi^2$ of PDF fits . . . . .	57
A.1	Datasets . . . . .	65
A.2	Parameters varied for PDF fits . . . . .	76
A.3	PDF parameters obtained in the fits . . . . .	76



---

## Bibliography

- [1] Joram Berger, “Analysis of  $Z(\rightarrow \mu\mu) + \text{jets}$  Events at the LHC”, Diplomarbeit, Karlsruher Institut für Technologie (KIT), 2010, URL: <http://www-ekp.physik.uni-karlsruhe.de/~thesis/data/iekp-ka2010-23.pdf>.
- [2] S. L. Glashow, “Partial Symmetries of Weak Interactions”, *Nucl. Phys.* **22** (1961) 579–588, DOI: [10.1016/0029-5582\(61\)90469-2](https://doi.org/10.1016/0029-5582(61)90469-2).
- [3] Abdus Salam, “Weak and Electromagnetic Interactions”, *Conf. Proc.* **C680519** (1968) 367–377.
- [4] Steven Weinberg, “A Model of Leptons”, *Phys. Rev. Lett.* **19** (1967) 1264–1266, DOI: [10.1103/PhysRevLett.19.1264](https://doi.org/10.1103/PhysRevLett.19.1264).
- [5] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons”, *Phys. Rev. Lett.* **13** (1964) 321–323, DOI: [10.1103/PhysRevLett.13.321](https://doi.org/10.1103/PhysRevLett.13.321).
- [6] Sidney D. Drell and Tung-Mow Yan, “Massive Lepton-Pair Production in Hadron-Hadron Collisions at High Energies”, *Phys. Rev. Lett.* **25** (1970) 902–902, DOI: [10.1103/PhysRevLett.25.902.2](https://doi.org/10.1103/PhysRevLett.25.902.2).
- [7] Serguei Chatrchyan et al., “Measurement of the Drell-Yan Cross Section in pp Collisions at  $\sqrt{s} = 7$  TeV”, *J. High Energy Phys.* **10** (2011) 007. 37 p, URL: <http://cds.cern.ch/record/1372196>.
- [8] G. Dissertori, “The Determination of the Strong Coupling Constant”, *Adv. Ser. Direct. High Energy Phys.* **26** (2016) 113–128, DOI: [10.1142/9789814733519\\_0006](https://doi.org/10.1142/9789814733519_0006), [[arXiv:1506.05407](https://arxiv.org/abs/1506.05407)].
- [9] Ulrich Kerzel, *B-Quark Fragmentation*, 2002, URL: <http://www-ekp.physik.uni-karlsruhe.de/~kerzel/fragmentation.html>.
- [10] K. A. Olive et al., “Review of Particle Physics”, *Chin. Phys.* **C38** (2014) p. 090001, DOI: [10.1088/1674-1137/38/9/090001](https://doi.org/10.1088/1674-1137/38/9/090001).

- [11] H. Abramowicz et al., “Combination of measurements of inclusive deep inelastic  $e^\pm p$  scattering cross sections and QCD analysis of HERA data”, *Eur. Phys. J. C* **75** (2015) p. 580, DOI: [10.1140/epjc/s10052-015-3710-4](https://doi.org/10.1140/epjc/s10052-015-3710-4), [[arXiv:1506.06042](https://arxiv.org/abs/1506.06042)].
- [12] S. Camarda et al., “QCD analysis of  $W$ - and  $Z$ -boson production at Tevatron”, *Eur. Phys. J. C* **75** (2015) p. 458, DOI: [10.1140/epjc/s10052-015-3655-7](https://doi.org/10.1140/epjc/s10052-015-3655-7), [[arXiv:1503.05221](https://arxiv.org/abs/1503.05221)].
- [13] Sayipjamal Dulat et al., “New parton distribution functions from a global analysis of quantum chromodynamics”, *Phys. Rev. D* **93** (2016) p. 033006, DOI: [10.1103/PhysRevD.93.033006](https://doi.org/10.1103/PhysRevD.93.033006), [[arXiv:1506.07443](https://arxiv.org/abs/1506.07443)].
- [14] Richard D. Ball et al., “Parton distributions for the LHC Run II”, *JHEP* **04** (2015) p. 040, DOI: [10.1007/JHEP04\(2015\)040](https://doi.org/10.1007/JHEP04(2015)040), [[arXiv:1410.8849](https://arxiv.org/abs/1410.8849)].
- [15] L. A. Harland-Lang et al., “Parton distributions in the LHC era: MMHT 2014 PDFs”, *Eur. Phys. J. C* **75** (2015) p. 204, DOI: [10.1140/epjc/s10052-015-3397-6](https://doi.org/10.1140/epjc/s10052-015-3397-6), [[arXiv:1412.3989](https://arxiv.org/abs/1412.3989)].
- [16] S. Alekhin, J. Blümlein and S. Moch, “The ABM parton distributions tuned to LHC data”, *Phys. Rev. D* **89** (2014) p. 054028, DOI: [10.1103/PhysRevD.89.054028](https://doi.org/10.1103/PhysRevD.89.054028), [[arXiv:1310.3059](https://arxiv.org/abs/1310.3059)].
- [17] H. Abramowicz et al., “Combination of measurements of inclusive deep inelastic  $e^\pm p$  scattering cross sections and QCD analysis of HERA data”, *Eur. Phys. J. C* **75** (2015) p. 580, DOI: [10.1140/epjc/s10052-015-3710-4](https://doi.org/10.1140/epjc/s10052-015-3710-4), [[arXiv:1506.06042](https://arxiv.org/abs/1506.06042)].
- [18] Andy Buckley et al., “LHAPDF6: parton density access in the LHC precision era”, *Eur. Phys. J. C* **75** (2015) p. 132, DOI: [10.1140/epjc/s10052-015-3318-8](https://doi.org/10.1140/epjc/s10052-015-3318-8), [[arXiv:1412.7420](https://arxiv.org/abs/1412.7420)].
- [19] Yuri L. Dokshitzer, “Calculation of the Structure Functions for Deep Inelastic Scattering and  $e^+e^-$  Annihilation by Perturbation Theory in Quantum Chromodynamics.”, *Sov. Phys. JETP* **46** (1977) 641–653.
- [20] V. N. Gribov and L. N. Lipatov, “Deep inelastic  $e p$  scattering in perturbation theory”, *Sov. J. Nucl. Phys.* **15** (1972) 438–450.
- [21] Guido Altarelli and G. Parisi, “Asymptotic Freedom in Parton Language”, *Nucl. Phys.* **B126** (1977) 298–318, DOI: [10.1016/0550-3213\(77\)90384-4](https://doi.org/10.1016/0550-3213(77)90384-4).
- [22] Peter Skands, “Introduction to QCD”, *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Searching for New Physics at Small and Large Scales (TASI 2012): Boulder, Colorado, June 4-29, 2012*, 2013, 341–420, DOI: [10.1142/9789814525220\\_0008](https://doi.org/10.1142/9789814525220_0008), [[arXiv:1207.2389](https://arxiv.org/abs/1207.2389)].

- 
- [23] Steven Weinberg, “Feynman Rules for Any Spin. 2. Massless Particles”, *Phys. Rev.* **134** (1964) B882–B896, DOI: [10.1103/PhysRev.134.B882](https://doi.org/10.1103/PhysRev.134.B882).
- [24] Torbjorn Sjostrand, Stephen Mrenna and Peter Z. Skands, “A Brief Introduction to PYTHIA 8.1”, *Comput. Phys. Commun.* **178** (2008) 852–867, DOI: [10.1016/j.cpc.2008.01.036](https://doi.org/10.1016/j.cpc.2008.01.036), [[arXiv:0710.3820](https://arxiv.org/abs/0710.3820)].
- [25] S. Agostinelli et al., “Geant4—a simulation toolkit”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003) 250–303, DOI: [http://doi.org/10.1016/S0168-9002\(03\)01368-8](http://doi.org/10.1016/S0168-9002(03)01368-8).
- [26] CERN, *A global endeavour*, 2013, URL: <http://public-archive.web.cern.ch/public-archive/en/About/Global-en.html>.
- [27] Alexandre Léonard, “Measurement of Z boson production in association with jets at the LHC and study of a DAQ system for the Triple-GEM detector in view of the CMS upgrade”, PhD thesis, Université Libre de Bruxelles, 2015, URL: <https://cds.cern.ch/record/2065693/files/Thesis-2015-Leonard.pdf>.
- [28] Stephen Myers, “The Large Hadron Collider 2008-2013”, *Int. J. Mod. Phys. A* **28** (2013) 1330035. 65 p, DOI: [10.1142/S0217751X13300354](https://doi.org/10.1142/S0217751X13300354).
- [29] K. Aamodt et al., “The ALICE experiment at the CERN LHC”, *JINST* **3** (2008) S08002, DOI: [10.1088/1748-0221/3/08/S08002](https://doi.org/10.1088/1748-0221/3/08/S08002).
- [30] G. Aad et al., “The ATLAS Experiment at the CERN Large Hadron Collider”, *JINST* **3** (2008) S08003, DOI: [10.1088/1748-0221/3/08/S08003](https://doi.org/10.1088/1748-0221/3/08/S08003).
- [31] S. Chatrchyan et al., “The CMS Experiment at the CERN LHC”, *JINST* **3** (2008) S08004, DOI: [10.1088/1748-0221/3/08/S08004](https://doi.org/10.1088/1748-0221/3/08/S08004).
- [32] A. Augusto Alves Jr. et al., “The LHCb Detector at the LHC”, *JINST* **3** (2008) S08005, DOI: [10.1088/1748-0221/3/08/S08005](https://doi.org/10.1088/1748-0221/3/08/S08005).
- [33] Victor Mukhamedovich Abazov et al., “Measurement of the  $\phi_\eta^*$  distribution of muon pairs with masses between 30 and 500 GeV in  $10.4 \text{ fb}^{-1}$  of  $p\bar{p}$  collisions”, *Phys. Rev.* **D91** (2015) p. 072002, DOI: [10.1103/PhysRevD.91.072002](https://doi.org/10.1103/PhysRevD.91.072002), [[arXiv:1410.8052](https://arxiv.org/abs/1410.8052)].
- [34] The ATLAS Collaboration, “Measurement of angular correlations in Drell–Yan lepton pairs to probe  $Z/\gamma^*$  boson transverse momentum at  $\sqrt{s} = 7 \text{ TeV}$  with the ATLAS detector”, *Phys. Lett. B* (2013), [[arXiv:1211.6899](https://arxiv.org/abs/1211.6899)].
- [35] David Barney, *CMS Detector Slice*, URL: <http://cds.cern.ch/record/2120661>.
- [36] CMS Collaboration, *3D SketchUp images of the CMS detector*, URL: [https://cms-docdb.cern.ch/cgi-bin/PublicDocDB/RetrieveFile?docid=11514&filename=cms\\_160312\\_02.pdf&version=3](https://cms-docdb.cern.ch/cgi-bin/PublicDocDB/RetrieveFile?docid=11514&filename=cms_160312_02.pdf&version=3).

- [37] Vladislav Balagura, “Notes on van der Meer Scan for Absolute Luminosity Measurement”, *Nucl. Instrum. Meth.* **A654** (2011) 634–638, DOI: [10.1016/j.nima.2011.06.007](https://doi.org/10.1016/j.nima.2011.06.007), [[arXiv:1103.1129](https://arxiv.org/abs/1103.1129)].
- [38] CMS Collaboration, “CMS Luminosity Measurements for the 2016 Data Taking Period” (2017).
- [39] G L Bayatyan et al., *CMS computing: Technical Design Report*, Technical Design Report CMS, Submitted on 31 May 2005, Geneva: CERN, 2005, URL: <https://cds.cern.ch/record/838359>.
- [40] *Kappa*, URL: <https://github.com/KappaAnalysis/Kappa>.
- [41] Joram Berger et al., “ARTUS - A Framework for Event-based Data Analysis in High Energy Physics” (2015), [[arXiv:1511.00852](https://arxiv.org/abs/1511.00852)].
- [42] Rene Brun and Fons Rademakers, “ROOT — An object oriented data analysis framework”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **389** (1997) 81–86, DOI: [http://dx.doi.org/10.1016/S0168-9002\(97\)00048-X](https://doi.org/10.1016/S0168-9002(97)00048-X).
- [43] *Excalibur*, URL: <https://github.com/artus-analysis/Excalibur>.
- [44] C. Patrignani and Particle Data Group, “Review of Particle Physics”, *Chinese Physics C* **40** (2016) p. 100001, URL: <http://stacks.iop.org/1674-1137/40/i=10/a=100001>.
- [45] Anna Kulesza, George F. Sterman and Werner Vogelsang, “Joint resummation in electroweak boson production”, *Phys. Rev.* **D66** (2002) p. 014011, DOI: [10.1103/PhysRevD.66.014011](https://doi.org/10.1103/PhysRevD.66.014011), [[arXiv:hep-ph/0202251](https://arxiv.org/abs/hep-ph/0202251)].
- [46] *Measurements of inclusive and differential Z boson production cross sections in pp collisions at  $\sqrt{s} = 13$  TeV*, tech. rep., Geneva: CERN, 2016, URL: <https://cds.cern.ch/record/2140105>.
- [47] Serguei Chatrchyan et al., “The performance of the CMS muon detector in proton-proton collisions at  $\sqrt{s} = 7$  TeV at the LHC”, *JINST* **8** (2013) P11002, DOI: [10.1088/1748-0221/8/11/P11002](https://doi.org/10.1088/1748-0221/8/11/P11002), [[arXiv:1306.6905](https://arxiv.org/abs/1306.6905)].
- [48] G.Perrin, *Update on Muon Scale- Factors on 2016 data*, 2016, URL: [https://indico.cern.ch/event/546811/contributions/2217942/attachments/1299081/1938370/TnP\\_2016data\\_topup.pdf](https://indico.cern.ch/event/546811/contributions/2217942/attachments/1299081/1938370/TnP_2016data_topup.pdf).
- [49] *Loss of Hit Efficiency due to HIP*, 2016, URL: <https://twiki.cern.ch/twiki/CMS/SiStripHitEffLoss>.
- [50] The CMS Collaboration, *Misalignment and Muon Scale Corrections Extracted from the 2011A  $Z \rightarrow \mu\mu$  Sample*, tech. rep., Geneva, 2012.
- [51] *Rochester Correction*, 2016, URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/RochcorMuon>.

- 
- [52] *Measurements of inclusive and differential Z boson production cross sections in pp collisions at  $\sqrt{s} = 13$  TeV*, tech. rep., Geneva: CERN, 2016, URL: <https://cds.cern.ch/record/2140105>.
- [53] Albert M Sirunyan et al., “Measurement of the  $t\bar{t}$  production cross section using events with one lepton and at least one jet in pp collisions at  $\sqrt{s}=13$  TeV” (2017), [[arXiv:1701.06228](https://arxiv.org/abs/1701.06228)].
- [54] Vardan Khachatryan et al., “Measurement of the WZ production cross section in pp collisions at  $\sqrt{s} = 13$  TeV”, *Phys. Lett.* **B766** (2017) 268–290, DOI: [10.1016/j.physletb.2017.01.011](https://doi.org/10.1016/j.physletb.2017.01.011), [[arXiv:1607.06943](https://arxiv.org/abs/1607.06943)].
- [55] Georges Aad et al., “Measurement of the ZZ Production Cross Section in pp Collisions at  $\sqrt{s} = 13$  TeV with the ATLAS Detector”, *Phys. Rev. Lett.* **116** (2016) p. 101801, DOI: [10.1103/PhysRevLett.116.101801](https://doi.org/10.1103/PhysRevLett.116.101801), [[arXiv:1512.05314](https://arxiv.org/abs/1512.05314)].
- [56] T. Gehrmann et al., “ $W^+W^-$  Production at Hadron Colliders in Next to Next to Leading Order QCD”, *Phys. Rev. Lett.* **113** (2014) p. 212001, DOI: [10.1103/PhysRevLett.113.212001](https://doi.org/10.1103/PhysRevLett.113.212001), [[arXiv:1408.5243](https://arxiv.org/abs/1408.5243)].
- [57] Nikolaos Kidonakis, “Theoretical results for electroweak-boson and single-top production”, *PoS DIS2015* (2015) p. 170, [[arXiv:1506.04072](https://arxiv.org/abs/1506.04072)].
- [58] The ATLAS Collaboration, “Measurement of the Z boson transverse momentum distribution in pp collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector”, *JHEP* **09** (2014) p. 145, DOI: [10.1007/JHEP09\(2014\)14](https://doi.org/10.1007/JHEP09(2014)14), [[arXiv:1406.3660](https://arxiv.org/abs/1406.3660)].
- [59] Dominik Haitz, “Precision Studies of Proton Structure and Jet Energy Scale with the CMS Detector at the LHC”, PhD thesis, Karlsruher Institut für Technologie (KIT), 2016, URL: <https://ekp-invenio.physik.uni-karlsruhe.de/record/48801>.
- [60] Arnulf Quadt Katharina Bierwagen Ulla Blumenschein, *Bayesian Unfolding*, tech. rep., Göttingen: Georg-August-University Göttingen, URL: <https://indico.cern.ch/event/107747/contributions/32648/attachments/24323/35007/bierwagen.pdf>.
- [61] Kerstin Tackmann et al. Tim Adye Richard Claridge, *RooUnfold: ROOT Unfolding Framework*, 2016, URL: <http://hepunix.rl.ac.uk/~adye/software/unfold/RooUnfold.html>.
- [62] G. Cowan, “A survey of unfolding methods for particle physics”, *Conf. Proc.* **C0203181** (2002) 248–257.
- [63] G.Perrin, *Muon ID Scale-Factor on 2016 ReReco*, 2016, URL: [https://indico.cern.ch/event/595070/contributions/2405095/attachments/1388822/2114847/MC\\_12\\_12\\_2016.pdf](https://indico.cern.ch/event/595070/contributions/2405095/attachments/1388822/2114847/MC_12_12_2016.pdf).

- [64] T. Gleisberg et al., “Event generation with SHERPA 1.1”, *JHEP* **02** (2009) p. 007, DOI: [10.1088/1126-6708/2009/02/007](https://doi.org/10.1088/1126-6708/2009/02/007), [[arXiv:0811.4622](https://arxiv.org/abs/0811.4622)].
- [65] C. F. Berger et al., “An Automated Implementation of On-Shell Methods for One-Loop Amplitudes”, *Phys. Rev.* **D78** (2008) p. 036003, DOI: [10.1103/PhysRevD.78.036003](https://doi.org/10.1103/PhysRevD.78.036003), [[arXiv:0803.4180](https://arxiv.org/abs/0803.4180)].
- [66] Andy Buckley et al., “Rivet user manual”, *Comput. Phys. Commun.* **184** (2013) 2803–2819, DOI: [10.1016/j.cpc.2013.05.021](https://doi.org/10.1016/j.cpc.2013.05.021), [[arXiv:1003.0694](https://arxiv.org/abs/1003.0694)].
- [67] T. Kluge, K. Rabbertz and M. Wobisch, “FastNLO: Fast pQCD calculations for PDF fits”, *Deep inelastic scattering. Proceedings, 14th International Workshop, DIS 2006, Tsukuba, Japan, April 20-24, 2006*, 2006, 483–486, [[arXiv:hep-ph/0609285](https://arxiv.org/abs/hep-ph/0609285)].
- [68] Daniel Britzger et al., “New features in version 2 of the fastNLO project”, *Proceedings, 20th International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS 2012): Bonn, Germany, March 26-30, 2012*, 2012, 217–221, DOI: [10.3204/DESY-PROC-2012-02/165](https://doi.org/10.3204/DESY-PROC-2012-02/165), [[arXiv:1208.3641](https://arxiv.org/abs/1208.3641)].
- [69] Luigi Del Debbio, Nathan P. Hartland and Steffen Schumann, “MCgrid: projecting cross section calculations on grids”, *Comput. Phys. Commun.* **185** (2014) 2115–2126, DOI: [10.1016/j.cpc.2014.03.023](https://doi.org/10.1016/j.cpc.2014.03.023), [[arXiv:1312.4460](https://arxiv.org/abs/1312.4460)].
- [70] Georg Sieber, “Measurement of Triple-Differential Dijet Cross Sections with the CMS Detector at 8 TeV and PDF Constraints”, PhD thesis, Karlsruhe Institut für Technologie (KIT), 2016, URL: <https://ekp-invenio.physik.uni-karlsruhe.de/record/48871>.
- [71] Juan Rojo, “Constraints on parton distributions and the strong coupling from LHC jet data”, *Int. J. Mod. Phys.* **A30** (2015) p. 1546005, DOI: [10.1142/S0217751X15460057](https://doi.org/10.1142/S0217751X15460057), [[arXiv:1410.7728](https://arxiv.org/abs/1410.7728)].
- [72] Gavin P. Salam and Juan Rojo, “A Higher Order Perturbative Parton Evolution Toolkit (HOPPET)”, *Comput. Phys. Commun.* **180** (2009) 120–156, DOI: [10.1016/j.cpc.2008.08.010](https://doi.org/10.1016/j.cpc.2008.08.010), [[arXiv:0804.3755](https://arxiv.org/abs/0804.3755)].
- [73] W. Hollik et al., “Electroweak corrections to Z-boson hadroproduction at finite transverse momentum”, *Nucl. Phys. B* **900** (2015) 576–602, DOI: [10.1016/j.nuclphysb.2015.09.018](https://doi.org/10.1016/j.nuclphysb.2015.09.018), [[arXiv:1504.07574](https://arxiv.org/abs/1504.07574)].
- [74] S. Alekhin et al., “HERAFitter”, *Eur. Phys. J.* **C75** (2015) p. 304, DOI: [10.1140/epjc/s10052-015-3480-z](https://doi.org/10.1140/epjc/s10052-015-3480-z), [[arXiv:1410.4412](https://arxiv.org/abs/1410.4412)].
- [75] M. Botje, “QCDNUM: Fast {QCD} evolution and convolution”, *Computer Physics Communications* **182** (2011) 490–532, DOI: [http://dx.doi.org/10.1016/j.cpc.2010.10.020](https://doi.org/http://dx.doi.org/10.1016/j.cpc.2010.10.020).

- [76] H1 Zeus Collaborations, “Combination of Measurements of Inclusive Deep Inelastic  $e^\pm p$  Scattering Cross Sections and QCD Analysis of HERA Data”, *Eur. Phys. J.* **C75** (2015) p. 580, DOI: [10.1140/epjc/s10052-015-3710-4](https://doi.org/10.1140/epjc/s10052-015-3710-4), [[arXiv:1506.06042](https://arxiv.org/abs/1506.06042)].
- [77] F. James and M. Roos, “Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations”, *Comput. Phys. Commun.* **10** (1975) 343–367, DOI: [10.1016/0010-4655\(75\)90039-9](https://doi.org/10.1016/0010-4655(75)90039-9).
- [78] J. Pumplin et al., “Uncertainties of predictions from parton distribution functions. 2. The Hessian method”, *Phys. Rev.* **D65** (2001) p. 014013, DOI: [10.1103/PhysRevD.65.014013](https://doi.org/10.1103/PhysRevD.65.014013), [[arXiv:hep-ph/0101032](https://arxiv.org/abs/hep-ph/0101032)].
- [79] Vardan Khachatryan et al., “Measurement and QCD analysis of double-differential inclusive jet cross-sections in pp collisions at  $\sqrt{s} = 8$  TeV and ratios to 2.76 and 7 TeV”, *JHEP* **03** (2017) p. 156, DOI: [10.1007/JHEP03\(2017\)156](https://doi.org/10.1007/JHEP03(2017)156), [[arXiv:1609.05331](https://arxiv.org/abs/1609.05331)].



---

## Danksagung

Ganz besonders möchte ich mich bei Prof. Dr. Günter Quast für die Möglichkeit bedanken mein Studium mit einer Arbeit zu diesem interessanten Thema abzuschließen. Vor allem: Vielen Dank für die umfassende Betreuung und die vielen hilfreichen Ratschläge!

Ebenso möchte ich mich bei Prof. Dr. Müller für die Übernahme des Korreferats bedanken.

Weiterhin möchte ich mich bei Dr. Klaus Rabbertz, Dr. Stefan Wayand, Dr. Dominik Haitz, Dr. Georg Sieber, Thomas Berger und Christoph Heidecker bedanken, die mich in die Software eingeführt haben, immer für meine Fragen zur Verfügung standen, intensiv mit mir über Details meiner Arbeit diskutiert haben und mir damit von Anfang an sehr beim Erstellen dieser Arbeit geholfen haben.

Außerdem möchte ich allen Kollegen am Institut für experimentelle Teilchenphysik für die tolle Arbeitsatmosphäre und die gegenseitige Unterstützung danken. Das gilt insbesondere für diejenigen Kollegen, die meine Arbeit korrekturgelesen haben. Zudem ist auch die Unterstützung des Staates Baden-Württemberg durch das bwHPC-Konzept und der Deutsche Forschungsgemeinschaft (DFG) durch den Forschungszuschuss INST 39/963-1 FUGG herauszustellen, die die Bereitstellung wichtiger Computerressourcen für diese Arbeit ermöglicht haben.

Zudem möchte ich mich bei meinem Freund und meinen Eltern bedanken, die immer für mich da waren, mich bei Rückschlägen aufgebaut haben und mich insbesondere in der anstrengenden Phase beim Niederschreiben dieser Arbeit ertragen haben.



## Erklärung der selbständigen Anfertigung der Masterarbeit

Hiermit erkläre ich, dass ich die Masterarbeit mit dem Titel

»*Measurement of the  $Z(\rightarrow \mu\mu) + \text{Jets}$  Cross Section at  $\sqrt{s} = 13 \text{ TeV}$  with the CMS  
Experiment and Studies of the Proton Structure*«

selbständig und unter ausschließlicher Verwendung der angegebenen Hilfsmittel  
angefertigt habe.

---

Anna Carlotta Friedel  
Karlsruhe, den 02.05.2017