Z boson production in association with jets: measurement and phenomenology

Dissertation

zur Erlangung des Doktorgrades an der Fakultät
für Mathematik, Informatik und Naturwissenschaften
Fachbereich Physik
der Universität Hamburg

vorgelegt von
Mikel Mendizabal Morentin
aus
Lezama, Baskenland, Spanien

Hamburg
2023

Gutachter der Dissertation:	Dr. Hannes Jung
	Prof. Dr. Elisabetta Gallo
Zusammensetzung der Prüfungskommission:	Prof. Dr. Daniela Pfannkuche
	Prof. Dr. Peter Schleper
	Prof. Dr. Elisabetta Gallo
	Dr. Alexander Glazov
	Dr. Hannes Jung
Vorsitzender der Prüfungskommission:	Prof. Dr. Daniela Pfannkuche
Datum der Disputation:	09.08.2023
Vorsitzender des Fach-Promotionsausschusses Physik:	Prof. Dr. Günter H. W. Sigl
Leiter des Fachbereichs Physik:	Prof. Dr. Wolfgang J. Parak
Dekan der Fakultät MIN:	Prof. Dr.-Ing. Norbert Ritter

Eidesstattliche Versicherung / Declaration on oath

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt habe. Die eingereichte schriftliche Fassung entspricht der auf dem elektronischen Speichermedium. Die Dissertation wurde in der vorgelegten oder einer ähnlichen Form nicht schon einmal in einem früheren Promotionsverfahren angenommen oder als ungenügend beurteilt.

I hereby declare, on oath, that I have written the present dissertation by my own and have not used any resources and aids other than those acknowledged. The written version submitted corresponds to the one stored electronically. The dissertation presented in this form, has not already been accepted in an earlier doctoral procedure or assessed as unsatisfactory.

Hamburg, July 6, 2023

Mikel Mendizabal Morentin

Abstract

In this thesis, we study two Z boson production mechanisms: the Drell-Yan process and real electroweak corrections.

The transverse momentum spectrum of the Drell-Yan lepton pair is of general interest as it is sensitive to the hadron structure. Moreover, the small transverse momentum region , $q_T < m_{ll}$, is sensitive to the 3D structure of the hadron. To describe the $q_T < m_{ll}$ region, resummation to all orders is needed. However, the $q_T \ll m_{ll}$ region is non-perturbative and has to be modelled. This region is modelled via a Gaussian distribution of the intrinsic transverse momentum of the incoming partons. Nonetheless, this is insufficient, and very-soft gluon radiation is needed. We investigate the effect of very-soft gluon radiations in the $q_T \ll m_{ll}$ region within the Parton Branching framework, and how it relates to the Collins-Soper kernel of the TMD factorisation. The very-soft gluon radiation is removed through the soft-gluon resolution scale in the parton showers of Monte-Carlo event generators, such as Pythia8 and Herwig7. We demonstrate that this leads to a centre-of-mass energy dependent intrinsic transverse momentum, by tuning the intrinsic transverse momentum.

At high partonic centre-of-mass energies, electroweak corrections become large, and the Z boson might be produced as a real electroweak correction. In QCD processes, at high jet transverse momentum, virtual electroweak corrections become large, and at the TeV scale, their contribution to the cross-section is of the order of 10%. We measure the contribution of real electroweak corrections to QCD events as a function of the jet transverse momentum. We analyse proton-proton collisions at a centre-of-mass energy of 13 TeV at the LHC, recorded by the CMS experiment during 2016, corresponding to an integrated luminosity of 36.3 fb^{-1}. We measure the contribution of the real electroweak correction to a 1 TeV jet transverse momentum to be 1-2 %. Moreover, we investigate different observables sensitive to real electroweak corrections to test their modelling in different theoretical scenarios.

Zusammenfassung

In dieser Arbeit untersuchen wir zwei Produktionsmechanismen von Z-Bosonen: den Drell-Yan-Prozess und reelle elektroschwache Korrekturen.

Das transversale Impulsspektrum des Drell-Yan-Leptonenpaares ist von allgemeinem Interesse, da es Aufschluss über die Hadronenstruktur gibt. Transversalimpulsen bei $q_T < m_{ll}$ ist die 3D-Struktur des Hadrons zugänglich. Um den Bereich $q_T < m_{ll}$ zu beschreiben, ist eine Resummation über alle Ordnungen erforderlich. Der $q_T \ll m_{ll}$ Bereich ist jedoch nicht-perturbativ und muss modelliert werden. Dieser Bereich wird durch eine Gaußsche Normalverteilung des intrinsischen Transversalimpulses der ankommenden Partonen beschrieben. Nichtsdestotrotz ist dies nicht ausreichend und sehr weiche Gluonenstrahlung wird benötigt. Wir untersuchen die Auswirkung sehr weicher Gluonenstrahlung im $q_T \ll m_{ll}$-Bereich im Rahmen des Parton Branching Modells, und wie sie mit dem Collins-Soper-Kernel der TMD-Faktorisierung zusammenhängt. Die sehr weiche Gluonenstrahlung wird durch die Auflösungsskala in Partonenschauern der Monte-Carlo-Ereignisgeneratoren wie Pythia8 und Herwig7 vernachlaessigt. Wir zeigen, dass dies zu einem von der Schwerpunktsenergie abhängigen intrinsischen Transversalimpuls führt, indem wir den intrinsischen Transversalimpuls in Pythia8 und Herwig7 für unterschiedliche Drell-Yan-Messungen tunen.

Bei hohen partonischen Schwerpunktsenergien werden die elektroschwachen Korrekturen groß, und das Z-Boson kann als reelle elektroschwache Korrektur auftreten. Bei QCD-Prozessen werden die virtuellen elektroschwachen Korrekturen bei hohem Jet-Transversal -impuls groß, und, im TeV-Bereichliegt ihr Beitrag zum Wirkungsquerschnitt in der Größenordnung von 10 %. Wir messen den Beitrag der reellen elektroschwachen Korrekturen zu QCD-Ereignissen als Funktion des Jet-Transversalimpulses. Bei einer Schwerpunktsenergie von 13 TeV am LHC untersuchen wir Proton-Proton-Kollisionen, vom CMS-Experiment im Jahr 2016 aufgezeichnet, was einer integrierten Luminosität von 36.3 fb^{-1} entspricht. Den Beitrag der reellen elektroschwachen Korrektur zu einem 1 TeV Jet-Transversalimpuls bestimmer wir zu 1-2 %. Darüber hinaus untersuchen wir verschiedene Observablen, die auf reelle elektroschwache Korrekturen sensitiv sind, um deren Modellierung in verschiedenen theoretischen Szenarien zu testen.

Contents

CHAPTER 1

Introduction

The increase of access to higher education over the last century has benefited and shaped modern science[1]. High energy physics stands as one of the bigger benefactors of this socio economical situation, where international collaborations work on an attempt to understand the working principles of the universe. A clear example are the collaborations in the Large Hadron Collider at CERN: ALICE, ATLAS, CMS and LHCb, where as for Summer 2023 the CMS collaboration consists of almost 6000 participants spread across 59 countries in 257 institutes [1].

The Standard Model of particle physics (SM) is the outcome of such collaborations across nations and cultures, where theorists and experimentalists contribute to the completion of the SM. The SM is the theory describing the interactions between the building blocks of the universe (excluding the gravitational interaction). While the SM is one of the most successful theories, it is not able to describe all the experimental observations, such as the mass of the neutrinos. To explain such phenomena the SM has to be extended. The extensions to the SM are commonly known as beyond the Standard Model (BSM) theories. Many BSM theories have been proposed to describe the missing pieces of the SM, however, none of these extensions has been verified by experimental measurements. Thus, revisiting the SM with precision measurements and theoretical predictions is a key to find new physics.

In this thesis we revisit the electroweak sector of the SM, where the force carriers of the electroweak interactions are the Z and $W^{\pm}$ bosons. More precisely, we investigate the production of the Z boson. Its leptonic decay is one of the cleanest signatures to measure in a detector, making it the perfect playground to test our theories. The Z boson was theoretised as part of the electroweak theory by S. Glashow [2], A. Salam [3]

[1] It is important to note that access to science is still highly restricted.

and S. Weinberg [4], and later discovered in the Super Proton Synchrotron at CERN in 1983 [5]. In this work we investigate two production modes of the Z boson: the Drell-Yan production and the real electroweak correction.

The Drell-Yan process occurs in hadronic collisions, where a quark and anti-quark collide and create a Z boson. It was proposed as a mechanism to describe the production of lepton pairs in 1970 by S. D. Drell and T.-M. Yan [6] and it has been an object of study since then. The Drell-Yan process is of big interest as it is sensitive to the proton structure, moreover, at low transverse momentum, q_T, of the Drell-Yan lepton pair with mass m_{ll} it is sensitive to the 3D structure of incoming hadrons. To describe the $q_T < m_{ll}$ region resummation has to be introduced. However, the $q_T \ll m_{ll}$ region is non-perturbative and has to be modelled. We study the transverse momentum of the Z boson, with emphasis on the non-perturbative regime, within the Parton Branching framework and in the parton showers of Monte-Carlo event generators, such as PYTHIA8 and HERWIG7.

As the centre-of-mass energy of the hadronic collisions increases, electroweak corrections become large, and the Z boson might arise as product of a radiation, what can lead to an overestimate of the jet transverse momentum. We measure the contribution of the real electroweak corrections in high energetic Quantum Chromodynamic (QCD) processes in hadronic collisions. For the measurement we make use of the proton-proton collisions recorded by the CMS experiment at the LHC at a centre-of-mass energy of 13 TeV in 2016. Further, we measure observables sensitive to real electroweak radiation and we compare them to different theoretical calculations.

This thesis is organised as follows. In chapter 2 we present the theoretical premisses, where we give an overview of the SM and we present the theoretical tools to describe proton-proton collisions with a focus on the Drell-Yan process. In chapter 3 we introduce the main theoretical aspects of electroweak corrections and we investigate the modelisation of real electroweak radiations in parton showers. In chapter 4 we present the experimental set-up for the measurement of the electroweak corrections: the LHC apparatus and the CMS experiment. In chapter 5 we measure the contribution of real electroweak radiations in QCD processes. We finalise the thesis by drawing conclusions and presenting the results from a global perspective in chapter 6.

CHAPTER 2

Theoretical premises

In this chapter we present the theoretical premises to describe hadronic collisions. In section 2.1 we give a short introduction of the Standard Model of particle physics, the theory describing the interaction between the building blocks of the universe. In section 2.2 we explain the theoretical elements involved in describing hadronic collisions, with a focus on the Drell-Yan production.

2.1 The standard model of particle physics

The standard model of particle physics (SM) is the theory that describes the building blocks of visible matter and their interactions. Such interactions are: the electromagnetic, the weak and the strong interactions. Gravitational interactions are not included. We can divide the building blocks in different types of particles regarding their spin. Fermions are particles with a half integer spin and bosons particles with integer spin.

The SM fermions are particles with half integer spin. We can categorise the SM fermions depending on the interactions they undergo: leptons and quarks. Leptons do not interact through strong interactions. Moreover, within the leptons there are electromagnetically charged leptons and neutral leptons. Charged leptons are: electrons, muons and taus. Neutral leptons or neutrinos are the electron neutrino, muon neutrino and tau neutrino. Neutrinos only interact through weak interactions.

Quarks are fermions that interact through strong, weak and electromagnetic interactions. There are six types of quarks or flavours: up, down, strange, charm, bottom and top quarks. Moreover, all the SM fermions have a corresponding antiparticle.

The SM is a quantum field theory. It respects the laws of special relativity and quantum

mechanics and the symmetries of nature. Particles are embedded into quantum fields, with quantum numbers, such as, mass and spin. The SM fermions are embedded into spinor fields, bosons with spin 1 into vector fields and spin 0 particles into scalar fields.

The Lagrangian density characterises the kinematics of the fields. Besides, it has to respect two major types of symmetries: global and local symmetries. Global symmetries are physical symmetries, that involve physical consequences. For every conserved continuous global symmetry there is a conserved current. This is Noether's theorem [7], one of the most important foundations of modern physics based on symmetries, such as the SM.

Local symmetries, more known as gauge invariance, are not natural symmetries. The physics does not depend on a specific time and space. Local symmetries are artefacts that allows us to study nature locally, hence, we can describe the dynamics of the particles with a local Lagrangian, a Lagrangian that depends on space-time.

The SM is a theory based on symmetries, and it follows the group structure of

$$\mathrm{SU}(3)_c \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y, \tag{2.1}$$

here SU(N) stands for special unitary algebraic groups that act on N dimensions. Each group has $N^2 - 1$ objects that generated the group, called generators. The generators can be represented through $N \otimes N$ matrices, for SU(2) the three Pauli matrices and for SU(3) the eight Gell-Mann matrices. These are rotational and boost matrices. The $\mathrm{SU}(3)_c$ matrices, where the c stands for colour, act on the colour space and $\mathrm{SU}(2)_L$ on weak isospin space.

Now we are going to study the different interactions: electromagnetic, weak and strong interactions. First we present the electroweak interactions, which is the unification of the electromagnetic and weak interactions. The electroweak interactions correspond to the SU(2)⊗U(1) group. Then, the strong interactions which correspond to the SU(3) group.

From now on, we represent the spinor fields with the ψ Greek letter, the vector fields with a Latin capital letter A_μ and scalar fields with the Φ letter.

2.1.1 Electroweak interactions

The theory of electroweak interactions arises from the unification of quantum electrodynamics and weak interactions at high energies. Quantum electrodynamics explains the interactions of electromagnetically charged particles at low energies. The theory of weak interactions was developed to explain flavour changing processes, such as β decays. In this section we will go through quantum electrodynamics and the theory of weak interactions. Then, shortly introduce the unification of the electroweak theory and its spontaneous symmetry breaking.

Quantum electrodynamics

Quantum electrodynamics (QED) is the theory that describes the interaction of electromagnetically charged particles, quarks and charged leptons, in other words the interaction between matter and light.

QED is a theory that is invariant under U(1) transformations, hence, it has a conserved current, the electromagnetic charge. We start from Diracs Lagrangian, which represents a free spinor field:

$$\mathscr{L}_{Dirac} = \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi \tag{2.2}$$

here ψ is a doublet of spinors which represents the different possible states of the spin of the 1/2 spin particles:

$$\begin{aligned} &\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \\ &\psi_R = (0, 1/2), \psi_L = (1/2, 0), \end{aligned} \tag{2.3}$$

γ are the Pauli matrices, m is the mass of the fermion field and $\bar{\psi} = \psi^{\dagger}\gamma^0$. The Dirac Lagrangian is invariant under U(1) transformations of the spinor fields

$$\psi \to e^{i\alpha}\psi, \tag{2.4}$$

where α is a real number. However, it is not invariant under a U(1) local transformation:

$$\psi \to e^{i\alpha(x)}\psi, \tag{2.5}$$

where we promote the real number α in equation 2.4 into a function depending on space-time. To ensure local invariance under a U(1) local transformation the interaction between a fermion and vector field is introduced through a covariant derivative D_{μ}:

$$D_{\mu} = \partial_{\mu} - ieA_{\mu} \tag{2.6}$$

here A_{μ} represents the vectorial field of a spin 1 particle and e the coupling strength. Then, we need to add a kinematic term for the A_{μ} field to the Lagrangian in equation 2.3. Moreover, a mass term for the A_{μ} field is not allowed, since it breaks the gauge invariance or local invariance. All these ingredients give the QED Lagrangian:

$$\mathscr{L}_{QED} = \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(D + m^2)\psi \tag{2.7}$$

where $F_{\mu\nu} = (\partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu})$ is the kinematic term of the vector field of the gauge boson A_{μ}. This Lagrangian is invariant under simultaneous transformations under U(1) of the A_{μ} and ψ fields. Now we can interpret the gauge boson A_{μ} as the photon, the force carrier of electromagnetic interactions.

Weak interactions

The theory of weak interactions was developed by E. Fermi [8] in 1933 to describe processes such as β decays:

$$\mathrm{n} \rightarrow \mathrm{p} + e^- + \bar{\nu}^e, \tag{2.8}$$

where a neutron (n) decays to a proton (p) and a electron (e^-) and electron neutrino (ν). These interactions happen at short distances and are weaker than QED interactions, hence the name. Fermi proposed weak interactions as a contact interaction, four fermions interact in a point, a direct four fermion interaction. In these type of interactions there is no force mediator involved. The Lagrangian for such interactions is given by:

$$\mathscr{L} = -G_F(\bar{\psi}_p\gamma^\mu\psi_n)(\bar{\psi}_e\gamma_\mu\psi_\nu) \tag{2.9}$$

where G_F is the coupling constant, known as the Fermi constant. Fermi's theory was successful describing processes at low energies. However, at larger energies the cross-sections calculated with such Lagrangian led to infinities. Thus, the introduction of gauge bosons was necessary. Moreover, to explain the weakness of the interactions these bosons had to be massive.

Unification of the electroweak theory

The failure of Fermi's theory to describe weak processes at large energies, led to the idea of unifying electromagnetic and weak interactions into an electroweak theory. It was postulated that the electromagnetic and weak interactions are the same at large energies. The theory of the electroweak interactions was introduced by S. Glashow [2], A. Salam [3] and S. Weinberg [4].

The theory of electroweak interactions is a SU(2)$\otimes$U(1) gauge theory. The three generators of the SU(2) group are the τ^a, $a = 1, 2, 3$, which are represented by the Pauli matrices. Thus, we will have three gauge fields related to the SU(2) group and a gauge field from the U(1) group. The three gauge fields of the SU(2) group are A^1_μ, A^2_μ and A^3_μ. The gauge field of the U(1) group is B_μ.

Electroweak interactions are symmetric under SU(2)$\otimes$U(1) transformations. Since they are invariant under two group transformations, we will have two conserved currents. The weak isospin for the SU(2) group and the hypercharge for the U(1) group. Then, the total conserved charge is the sum of the two conserved currents:

$$Q = \frac{Y}{2} + T_3 \tag{2.10}$$

where Y is the hypercharge and T_3 is the isospin charge. From here, we can identify Q as the electromagnetic charge of the particle.

From experimental observations we know that only fermions with left chiralities interact trough weak interactions. Then, left handed fermions transform as a doublet under SU(2), while the right-handed fermions transform as singlets. The gauge field of the U(1) group will interact with both left and right handed chiralities. We represent the weak doublet for the leptons as:

$$L = (\nu_l, l_L,)^T, \tag{2.11}$$

where l stands for the charged lepton and ν_l the corresponding neutrino. We write the right-handed lepton as l_R. Note that for quarks the doublet is formed with the quark generations, e.g.: $L = (u_L, d_L)$, where d_L is a mixed state of down type quarks trough the Cabibbo–Kobayashi–Maskawa matrix [9, 10]. The interaction of the gauge fields with the fermions is introduced by the covariant derivative

$$D_\mu = \partial_\mu - ig\frac{\tau^a}{2}A^a_\mu - ig'YB_\mu, \tag{2.12}$$

with g the coupling constant between A^a_μ and the fermion field and g' the coupling constant between B_μ and the fermion field. In this way the Lagrangian for the fermion fields is invariant under local $\mathrm{SU}(2)_L$ and $\mathrm{U}(1)_Y$ transformations. The kinematic term in the Lagrangian for the gauge fields will be

$$\begin{aligned} &-\frac{1}{4}\left(F^a_{\mu\nu}F^{\mu\nu}_a + B_{\mu\nu}B^{\mu\nu}\right), \\ &B_{\mu\nu} = (\partial_\nu B_\mu - \partial_\mu B_\nu), \\ &F^a_{\mu\nu} = (\partial_\nu A^a_\mu - \partial_\mu A^a_\nu + ig\tau^a f_{abc}A^b_\mu A^c_\nu), \end{aligned} \tag{2.13}$$

where f_{abc} is the structure constant of the SU(2) group. The self-interaction of the A^a_μ fields appears from the non-Abelian structure of the SU(2) group, where the generators do not commute. It is important to note that the A^a_μ and B_μ fields are not directly linked to the $W^\pm, Z$ and γ bosons, but from a combination of the fields

$$\begin{aligned} W^\pm_\mu &= \frac{1}{\sqrt{2}}(A^1_\mu \mp iA^2_\mu) \\ Z_\mu &= -B_\mu \sin\theta_W + A^3_\mu \cos\theta_W \\ A_\mu &= B_\mu \cos\theta_W + A^3_\mu \sin\theta_W \end{aligned} \tag{2.14}$$

where θ_W is the Weinberg angle. Then, we can write the Lagrangian for electroweak interactions as:

$$\mathscr{L}_{EW} = i\bar{L}\gamma^\mu D_\mu L + i\bar{\psi}_R\gamma^\mu D_\mu \psi_R - \frac{1}{4}A^a_{\mu\nu}A^{\mu\nu}_a - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \tag{2.15}$$

However, it does not allow mass terms for the gauge bosons. These terms are not invariant under gauge transformations. From experimental observations, the $W^\pm$ and Z boson are massive, but for a symmetric theory the gauge bosons are massless.

Hence, the SU(2)⊗U(1) symmetry is broken spontaneously. The Lagrangian is symmetric under global and local transformations, however, there is a broken internal symmetry. The Goldstone theorem states that for each spontaneously broken symmetry there is a massless boson [11], which are commonly known as Goldstone bosons.

F. Englert and R. Brout [12] and P. Higgs [13] explained a way how the massless vector bosons of a spontaneously broken symmetry can acquire mass. Nowadays, this phenomena is known as the Higgs mechanism, with a massive scalar boson.

The Higgs mechanism for SU(2)xU(1)

The Higgs mechanism explains how a gauge boson becomes massive when there is an spontaneously broken symmetry. Later, Weinberg and Salam used the Higgs mechanism to explain the masses of three of the gauge bosons of the electroweak interactions.

We suppose the existence of two complex scalar fields that transform SU(2) as a doublet, $\Phi = (\Phi^+, \Phi^0)$, that extends through space-time. We can write the Lagrangian for such field as

$$\begin{aligned} \mathscr{L} &= D_\mu \Phi^\dagger D^\mu \Phi - V(\Phi) \\ V(\Phi) &= -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \end{aligned} \tag{2.16}$$

where $V(\phi)$ is the Higgs potential, μ is the value of the vacuum and λ is the quartic Higgs coupling. For $\mu^2 > 0$, $\Phi = 0$ is the global minimum. However, for $\mu^2 < 0$, $\Phi = 0$ becomes a local maximum, which is unstable, and the new global minimum is $\Phi = \sqrt{2\mu^2/\lambda} e^{i\theta}$. Now one can decompose the Φ field in, radial $\sigma(x)$ and angular fields $\pi(x)$ around the new minimum, in the so called linear sigma model

$$\Phi(x) = e^{\frac{i}{v}\pi^a(x)\tau^a} \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(v + \sigma(x)) \end{pmatrix} \tag{2.17}$$

We introduce the parametrisation of the Φ field around the minimum in equation 2.16, where we get that the angular field, $\pi(x)$, is a Goldstone boson and the radial field, $\sigma(x)$, is massive. The mass of the $\sigma(x)$ field is given by

$$m_\sigma = \mu\sqrt{2} \tag{2.18}$$

To remove the massless Goldstone bosons from the Lagrangian we fix the exponential parametrisation to zero:

$$\Phi(x) = e^{\frac{i}{v}\pi^a(x)\tau^a} \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(v + \sigma(x)) \end{pmatrix} \rightarrow \Phi(x) = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(v + \sigma(x)) \end{pmatrix}. \tag{2.19}$$

We denote the radial field, $\sigma(x) = H(x)$, which we refer as the Higgs field. The gauge invariant Lagrangian for the Higgs field is

$$\begin{aligned} \mathscr{L}_{Higgs} =& (D_\mu \Phi^\dagger)(D_\mu \Phi) - \lambda \left(\Phi^\dagger \Phi - \frac{v^2}{2} \right)^2 \\ & D_\mu = \partial_\mu - igA^a_\mu \frac{\tau^a}{2} - ig' Y B_\mu \text{ and } a = 1, 2, 3 \end{aligned} \tag{2.20}$$

Where A^a_μ are the gauge fields corresponding to SU(2) and B_μ to U(1). Thus, the gauge fields of SU(2) become massive when the internal symmetry of SU(2)⊗U(1) is broken. The broken symmetry makes the vacuum to contain a new minimum, where a massive field arises, the Higgs field. The interaction between the Higgs field and the electroweak bosons generates the mass.

To extract the mass terms for the physical $W^\pm_\mu$ and the Z_μ fields we introduce equations 2.14 and 2.19 in the Lagrangian above. The mass terms for the bosons are given by the following relations

$$m_{W^\pm} = \frac{1}{2} g v \text{ and } m_Z = \frac{1}{2} v \sqrt{g^2 + g'^2}. \tag{2.21}$$

2.1.2 Quantum chromodynamics: SU(3)$_c$

In this subsection we present the theory of quantum chromodynamics (QCD), which describes the strong interactions. QCD is a SU(3) gauge theory. The conserved current of the SU(3) symmetry is the colour charge. Then, more precisely, QCD describes the interaction between particles with colour charge, the interaction between quarks. These transform as triplets under SU(3) transformations. The SU(3) group has eight generators, the eight Gell-Mann matrices. Thus, the strong interaction has eight massless gauge fields, G^a_μ, where a runs from one to eight. The gluons are the gauge bosons of the strong interactions.

The Lagrangian for the strong interactions, for the quark q and anti-quark $\bar{q}$ fields is given by

$$-\mathscr{L}_{QCD} = -\frac{1}{4} F^a_{\mu\nu} F^{\mu\nu}_a + \bar{q}(i\gamma^\mu D_\mu - m)q \tag{2.22}$$

$$F^a_{\mu\nu} = \partial_\nu G_\mu - \partial_\mu G_\nu + g_s f^{abc} A^b_\mu A^b_\nu \tag{2.23}$$

$$D_\mu = \partial_\mu - ig_s T^a G^a_\mu, \tag{2.24}$$

where we introduce the covariant derivative to keep an invariant Lagrangian under local SU(3) transformations. Besides, as a non-Abelian theory, the generators, T^a of the SU(3) group do not commute, the self interaction between the gauge fields is allowed. The Lagrangian in equation 2.22 is not the full QCD Lagrangian. We should introduce a gauge-fixing term, a term that ensures the vector field A^a_μ transforms as a vector under SU(3) transformations.

2.2 Hadronic collisions

Hadronic collisions at high energies are excellent tools to test QCD theory. At high energies, in the proton-proton collision a partonic collision happens: the constituents of one of the protons collide with the constituents of the other proton. As shown in section 2.3 the partons, the constituents of the proton, can carry different fractions of the momentum of the proton. Moreover, they can have different flavours. The collision of protons offers many different partonic reactions at different energy scales. This complicated environment is a perfect scenario to test the SM. Besides, it opens possibilities to processes that are not within the SM.

In figure 2.1 we depict a simplistic diagram of a proton-proton collision to depict the main process in a proton-proton collision. Note that the colors do not represent color charge, but the different processes. We can disentangle the proton-proton collision into different stages. First, we have two incoming protons (blue lines and blobs), this stage is the initial state. The constituents of the protons collide, which we refer as the hard-process (violet blob). The quarks and gluons produced in the collision evolve from the scale of the hard-process to the hadronic energy scale forming hadrons (orange ellipses). Some of the newly formed hadrons will decay into more stable hadrons. While the partons evolve, both in the inital state and after the hard-process, they emit QCD radiation (red). The remnants of the protons (green) will also hadronise. When all the leptons and hadrons have reached a stable state we have reached the final state of the proton-proton collision. It is important to note that quarks and gluons in the event are connected through colour, as the colour charge has to be conserved in QCD interactions.

In the experiments we are only able to measure the leptons and hadrons left in the final state. The measurement of a given process is given in terms of cross-sections. The cross-section is the probability for a given process to occur, in this case, in a proton-proton collision.

In proton-proton collisions we can find many different energy scales. Starting from the proton, with a mass around 1 GeV, to the hard-process, which at 13 TeV centre of mass energies can reach up to the TeV scale, and back to non-perturbative scales in the hadronisation processes. Factorising the proton-proton collision is a good approach to deal with the different energy scales. In this section we introduce the idea of the factorisation of the cross-section and the collinear factorisation theorem.

2.2.1 Factorisation of the cross-section

With the factorisation theorem we separate the different processes at different energy scales in the proton-proton collisions into different parts. For simplicity we study the factorisation theorem for the cross-section of the Drell-Yan (DY) processes. In the DY

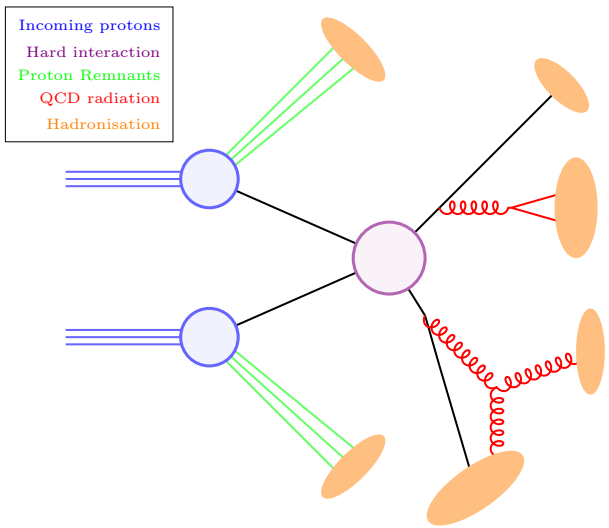

Figure 2.1: Diagram of a proton-proton collision.

processes a quark from a proton and an anti-quark from the other proton collide producing a photon or a Z boson where the boson can decay to a lepton anti-lepton pair. By only considering the lepton pair we do not take hadronisation into account. In the DY process we find two different energy scales: the protons and the quark anti-quark collision. We separate the cross-section of the DY processes into three functions, two functions for the two incoming protons, and a function for the quark-anti quark collision.

The PDFs describe the probability to find the incoming quark (or anti-quark) with a fraction of momentum of the proton, x, inside the proton. The quark anti-quark collision is described by a hard-scattering function, which is computed in perturbation theory. With the factorisation theorem we also introduce an energy scale, the factorisation scale. We embed radiation with energies below the factorisation scale into the PDFs and the processes with energies above the factorisation scale into the hard-scattering function.

In the following we describe the collinear factorisation theorem, where we give an overview of the hard-scattering function and the factorisation scale. In section 2.3 we present the PDFs and PDF TMDs.

Collinear factorisation theorem

The cross-section for the DY processes for protons with momentum P_1 and P_2 is factorised in the collinear factorisation theorem into the PDFs of the two incoming hadrons and the scattering of the incoming quarks:

$$\sigma_h = \sum_{i,j} \int_0^1 dx_i \int_0^1 dx_j f_{i/A}(x_i, \mu_\mathrm{F}) f_{j/B}(x_j, \mu_\mathrm{F}) \hat{\sigma}_{ij}(x_i, x_j, Q, \mu_\mathrm{F}, \mu_\mathrm{R}), \tag{2.25}$$

where i and j run over all the parton flavours, A and B are the incoming protons, x_i and x_j are the momentum fraction carried by the parton from their respective proton, $\hat{\sigma}_{ij}$ is the partonic cross-section, Q is the energy scale of the partonic interaction, μ_{F} the factorisation scale and μ_{R} the renormalisation scale. The partonic cross-section, defined as

$$\begin{aligned} \frac{d\hat{\sigma}_{i,j}}{dt} &= \frac{1}{16\pi\hat{s}^2}\sum |M(\hat{s},\hat{t},\hat{u})|^2 \text{ with} \\ \hat{s} &= (p_1+p_2)^2 = (p_3+p_4)^2 \\ \hat{t} &= (p_1-p_3)^2 = (p_4-p_2)^2 \\ \hat{u} &= (p_1-p_4)^2 = (p_3-p_2)^2, \end{aligned} \tag{2.26}$$

is a $2 \to 2$ process where two incoming particles interact generating two outgoing particles, with $\hat{s}$, $\hat{t}$ and $\hat{u}$ the Mandelstam variables, where p_1 and p_2 are the four momenta of the incoming particles and p_3 and p_4 the four momenta of the outgoing particles. M is the hard-scattering function and defines the transition probability of the incoming particles into the outgoing particles, e.g.: in the DY process, the probability of the two incoming quarks to transform into a lepton anti-lepton pair. M is computed using perturbation theory. In general, in the calculation there are ultraviolet and infrared divergences, where ultraviolet divergences are divergences in the calculation caused by large momenta. These divergences are absorbed by the renormalisation procedure, which introduces a new scale, the renormalisation scale μ_{R}. Infrared divergences are divergences in the calculation caused by small momenta and are absorbed into the PDFs. Then, μ_{F} works as cut-off scale to avoid infrared divergences.

It is important to note that μ_{F} and μ_{R} are not physical but arbitrary energy scales. The cross-section must not depend on any of the two scales

$$\frac{d\sigma_h}{d\mu_{\mathrm{F}}} = \frac{d\sigma_h}{d\mu_{\mathrm{R}}} = 0 \tag{2.27}$$

The calculation of the matrix element is performed through a perturbative expansion and the expansion is truncated at some order. The truncation leads to a dependence on μ_{F} and μ_{R}. This dependence is large when we truncate the expansion at low orders. By introducing more orders into the calculation the dependence on the two scales is reduced. However, the computation of higher orders is a difficult and time consuming task. The variation of the μ_{F} and μ_{R} in the calculations can be used to estimate the contribution of higher order terms of the expansion that are not included by the truncation. The collinear factorisation is a powerful tool to compute the cross-section of different processes. However, it fails to describe distributions where two different energy scales appear, more precisely, when one of the energy scales is much smaller than the other. In section 2.4 we study this effect in the transverse momentum of the Z boson and introduce the concept of resummation.

2.3 The proton structure

The structure of hadrons plays a big role in hadronic collisions, moreover, real electroweak corrections change the structure of the hadrons at high energies. For this, we present the basic ideas of the hadron structure only taking into account its QCD aspects. In chapter 3 we will investigate the effects of electroweak corrections to the hadron structure.

First, we need to understand the nature of hadrons. These are particles composed by quarks and gluons. The confinement of quarks into hadrons arises from the behaviour of the coupling strength of the strong interactions (α_s):

$$\alpha_s(\mu) = \frac{1}{\beta_0 \ln(\mu^2/\Lambda^2)} \tag{2.28}$$

where μ is the energy scale, Λ is the QCD energy scale of the order of hundred MeVs and β_0 is a constant. The coupling strength diverges for $\mu \to 0$ producing the confinement of quarks into hadrons by exchange of gluons. At large energy scales the coupling strength becomes small and quarks and gluons can be treated as free entities, this is known as the asymptotic freedom. Hence, at low energies quarks and gluons cannot be isolated, they form color neutral structures called hadrons. Hadrons are divided into two categories depending on the number of their valence quarks. If a hadron is composed by three (two) valence quarks the hadron is a baryon (meson). Valence quarks are quarks and anti-quarks that define the quantum numbers of the hadron, such as charge and isospin. The mass of a hadron is a result of strong interactions, not of its interaction with the Higgs boson.

Besides, for large values of μ, the coupling strength is small enough to apply a perturbative expansion of the QCD interactions, for small values of μ, at energy scales of hadrons ($\mu \sim$ 1 GeV) the coupling strength is too large to apply perturbation theory and perturbative QCD calculations cannot be made. Hence, the structure of hadrons has to be extracted from measurements.

One of the first models for the proton structure was proposed by R. Feynmann in 1969 [14], the parton model. In this model the constituents of the proton are massless point-like particles called partons. Later, partons were identified as quarks and gluons. This led to improve the parton model using QCD theory. From now on we will call the constituents of hadrons (quarks and gluons) the partons. The study of the composition of hadrons and their structure is done through parton distribution functions (PDFs). In section 2.3.1 we introduce the collinear structure of the hadrons trough PDFs and their evolution. In section 2.3.2 the transverse momentum dependent parton distribution functions (TMDs) are introduced.

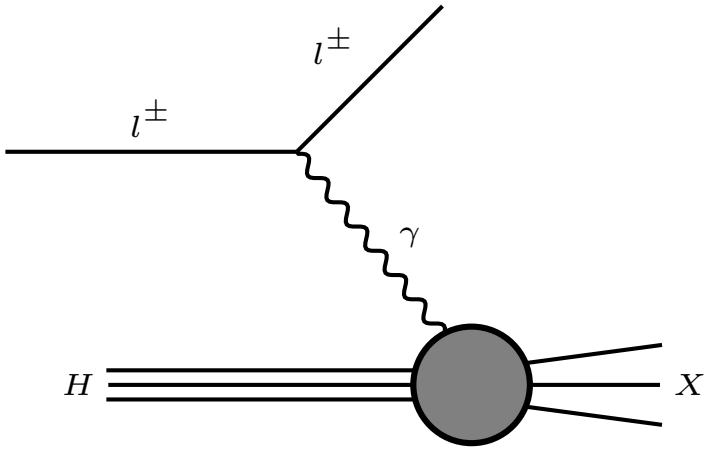

Figure 2.2: Deep inelastic scattering between a charged lepton ($l^{\pm}$) and a hadron (H) through a photon (γ), where the hadron breaks into remnants (X).

2.3.1 Parton distribution functions

With collinear parton distribution functions (PDFs) the collinear structure of hadrons is studied. They give the probability of a parton of a given flavour carrying a fraction of the total momentum of the hadron, x, at a given energy scale μ. We can derive the theoretical definition of the PDF from deep inelastic scattering (DIS). In DIS a high energetic charged lepton collides with a constituent of the hadron, breaking the hadron. The interaction is a QED interaction, the charged lepton and the parton interact via a photon (see figure 2.2). The DIS process can be factorised into an electromagnetic current, which involves the incoming lepton and the photon, and a hadronic current, which corresponds to the hadron. The definition of the PDF is then derived from the hadronic current:

$$f_{b/h}(x) \propto \int \frac{dk_T}{(2\pi)^4} \quad P \;\; k \tag{2.29}$$

here $f_{b/h}$ is the bare PDF for the parton of flavour b inside the hadron h, k_T is the transverse component of the momentum of the parton and the diagram is the hadron current in DIS. As bare quantity $f_{b/h}$ contains ultra violet divergences and it is renormalised to remove the divergences. The renormalisation procedure introduces a dependence on an arbitrary energy scale, μ, in the PDF:

$$f_{b/h}(x) \xrightarrow{Renormalisation} f_{b/h}(x, \mu) \tag{2.30}$$

As μ is an arbitrary scale $f_{b/h}(x, \mu)$ should not depend on the choice of μ. We can derive the renormalisation group equations (RGEs) that explain the dependence of the PDF on

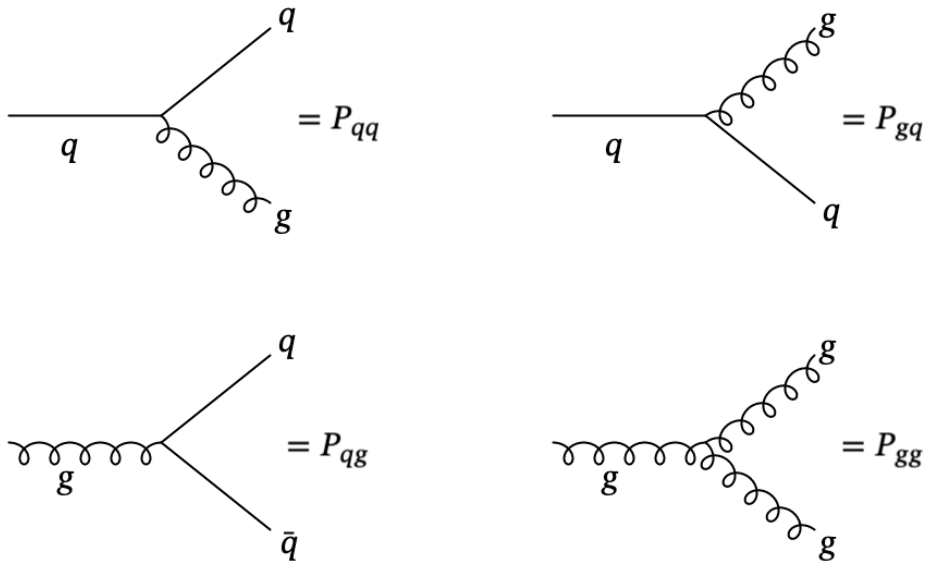

Figure 2.3: Feynmann diagrams at leading order for splittings.

the energy scale:

$$\frac{\partial f_{b,h}(x,\mu)}{\partial \ln \mu} = \sum_a \int_x^1 \frac{dz}{z} P_{ba}(z, \alpha_s(\mu)) f_{a,h}(x/z, \mu). \tag{2.31}$$

The RGEs for the PDFs are known as the DGLAP evolution equations [15–17] and P_{ba}, the evolution kernels, are known as the splitting functions.

To compute the evolution kernels with a perturbative expansion over α_S we have to replace the hadron by an on-shell parton. This is possible as the evolution kernels do not depend on the hadron or the hadron state. Then, by replacing the hadron by an on-shell parton we can calculate the kernels with a perturbative expansion from QCD vertices, given that μ is in the perturbative region. In figure 2.3 we present the leading order diagrams of the QCD vertices corresponding to the different evolution kernels and in equation 2.32 the evolution kernels at leading order.

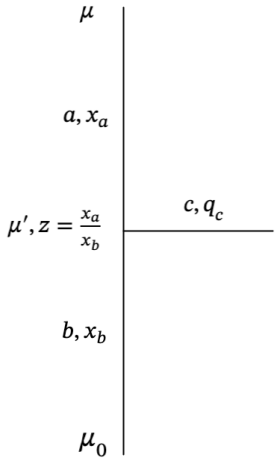

Figure 2.4: Evolution of a parton from a starting energy scale μ_0 to a scale μ, that undergoes a splitting at scale μ'. The parton b splits into a parton a and c at the energy scale μ'. x_a and x_b represent the fraction of momentum of the proton carried by the parton a and b consecutively.

$$
\begin{aligned}
P_{qq}(z) &= C_F \left[\frac{1+z^2}{(1-z)_+} + \frac{3}{2}\delta(1-z) \right] \\
P_{gq}(z) &= C_F \frac{1+(1-z)^2}{z} \\
P_{qg}(z) &= T_R \left[z^2 + (1-z)^2 \right] \\
P_{gg} &= 2C_A \left[\frac{z}{(1-z)_+} + \frac{1-z}{z} + z(1-z) \right] + \frac{11C_A - 4n_f T_R}{6}\delta(1-z)
\end{aligned}
\tag{2.32}
$$

where, C_A, C_B and T_R are the color factors contributing to the strong coupling and n_f is the number of flavours. The different splitting functions diverge for different values of z. P_{gg} and P_{qq} give the probability to radiate a gluon by a gluon and a quark respectively. These two functions diverge for $z \to 1$. For values of $z \sim 1$ the momentum transfer is large, hence, the momentum of the emitted gluon $(1-z)$ is small. We refer to these emissions as soft gluon radiations. P_{gg} and P_{gq} diverge for $z \to 0$. Now we can interpret the DGLAP evolution as a parton splitting process, where the evolving quark of a given flavour reduces its momentum by QCD bremsstrahlung, emission of gluons, or where the decay of a gluon to a quark anti-quark pair of a given flavour contributes to the PDF evolution. In figure 2.4 we show the schematic process of a splitting, where an incoming parton b splits into a parton a by emitting a parton c at scale μ' when evolving from μ_0 to μ. The momentum transfer from quark b to a is defined by z.

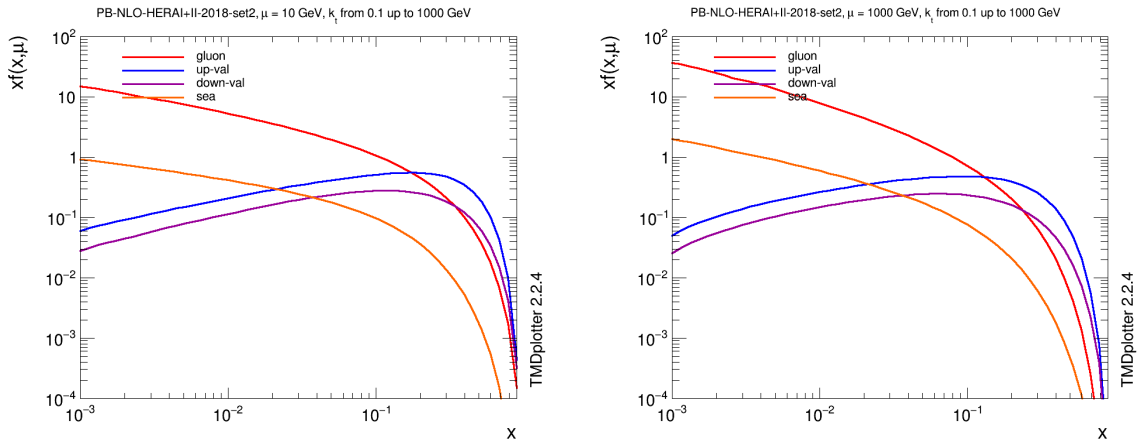

Figure 2.5: The momentum weighted collinear parton distribution function, $xf(x,\mu)$, extracted using HERA data by the Parton Branching collaboration for the valence quarks, sea quarks and gluons at different energy scales, μ, [18]. Plots made with TMDplotter2.2.4 [19].

While we can calculate the evolution of the PDF using perturbative physics, the structure of the hadron can not be determined from first principles, but it must be modelled. The structure of the hadron is modelled by parametrising the content of the different parton flavours in the hadron. A general pararmetrisation of a given parton flavour q has the following form:

$$xp(x) = Ax^B(1-x)^C + \cdots , \tag{2.33}$$

where x is the fraction of momenta carried by the parton, $p(x)$ is the PDFs for the parton p, and A, B, C are the parameters to fit. These parametrisations are fitted to measurements sensitive to the content of the hadrons and the PDF is determined at a given energy scale, μ_0. We refer to $f_{b/h}(x,\mu_0)$ as the starting distribution. The starting distribution can then be evolved to any other energy scale using the DGLAP evolution equation.

In figure 2.5 we present the Parton Branching collinear PDFs extracted using HERA data for different flavours at different energy scales: $\mu = 100$ GeV on the left panel and $\mu = 1000$ GeV. For both energy scales at large x, $x \sim 1/3$, the valence quarks carry most of the momentum, two up quarks and a down quark. For lower x values the gluon distribution increases. The gluons carry a large fraction of the momentum of the proton. Sea quarks, the quarks generated from gluon splitting into quark anti quark pairs, contribute also at low x. Parton radiation plays a large role in the structure of the hadron. The contribution of radiation increases with increasing energy scales.

Using collinear PDFs the hadron structure in one dimension can be studied. We can study the three dimensional structure of the hadron by not integrating over the transverse

momentum in equation 2.29. These three dimensional PDFs are known as unintegrated PDFs or transverse momentum dependent PDFs (TMD PDFs).

2.3.2 Transverse momentum dependent parton distributions

Unintegrated PDFs or transverse momentum PDFs (TMDs) study the structure of the hadrons in three dimensions. The TMD is defined by not integrating the transverse component of the momentum of the partons in equation 2.29. To parametrise the TMDs we can do a simultaneous parametrisation of the transverse and collinear structure. This approach is done in the GBW type of TMDs, which focuses on saturation effects.

In Parton Branching (PB) TMDs and TMDs based on the TMD factorisation, the collinear and transverse structure of the hadron are factorised. First, the collinear parametrisation of the hadron is performed at a given energy, in other words a collinear PDF is derived. On top of the collinear starting distribution the transverse momentum components are parametrised and fitted. In the following we focus on TMD factorisation and PB TMDs. We investigate the non-perturbative components and the evolution of the two approaches.

TMD factorisation TMDs

In the TMD factorisation the TMDs are defined as [20]

$$\begin{aligned}\tilde{f}_{j,h}(x, b_T, Q^2, \mu) = &\exp\left[-g_{j/A}(x_A, b_T; b_{max}) - g_D(b_T; b_{max}) \ln\frac{Q}{Q_0}\right] \\ &\times \exp\left[D(b_*, \mu_{b_*}) \ln\frac{Q}{\mu_{b_*}} + \int_{\mu_{b_*}}^{\mu} \frac{d\mu'}{\mu'}\left(\ln\frac{Q}{\mu'}\gamma_d + \gamma_j\right)\right] \\ &\times \sum_j F_{j/h}(x, \mu)\end{aligned} \tag{2.34}$$

where x is the collinear fraction of momentum of the hadron h carried by the parton j, b_T is the transverse position variable (the transverse momentum in position space), Q^2 is the rapidity factorisation scale, which arises from rapidity divergences and it is normally written as ζ, and μ, the factorisation scale of ultra-violet divergences. In section 2.4 more details on the TMD factorisation are given. The function F corresponds to the collinear distribution of the proton structure mapped to b space. The dependence on the Q and μ factorisation scales give rise to the evolution equations. The evolution equation in rapidity is given by

$$\begin{aligned}&\frac{\partial \tilde{f}_{j,h}(x, b_T, \zeta, \mu)}{\partial \ln\sqrt{\zeta}} = D(b_T, \mu) \text{ with } \frac{dD(b_T, \mu)}{d\ln\mu} = -\gamma_d \\ &D(b_T, \mu) = -\frac{\alpha_S}{\pi} C_F\left[\ln\frac{b_T^2\mu^2}{4} + \gamma_E\right] + \mathcal{O}(\alpha_S^2)\end{aligned} \tag{2.35}$$

The function D is commonly known as the Collins-Soper (CS) kernel. It is independent of the parton flavour, the hadron and x. Moreover, it is process independent and it is the same for fragmentation functions and TMDs. Fragmentation functions will be later introduced in section 2.6. For larger values of b_T γ_d and hence D become non-perturbative and can not be calculated. To avoid the non-perturbative region b_T is saturated to a value of b_* and μ_{b_*} is the energy scale for b_*. For $b_T > b_*$, the non-perturbative part of the evolution in rapidity is modelled in the g_D function. In addition to g_D, the intrinsic transverse momentum of the partons is non-perturbative and it is defined by the $g_{j/f}(x, b_T)$ function. Hence, to determine a TMD in the CSS formalism we need to extract the two non-perturbative functions g_D and $g_{j/f}$. The RG for the TMD is given by

$$\frac{\partial \tilde{f}_{j,h}(x, b_T, \zeta, \mu)}{\partial \ln \mu^2} = \gamma_j - \frac{1}{2}\gamma_d \ln \frac{\zeta}{\mu} \tag{2.36}$$

where γ_j is a perturbative coefficient dependent on the parton flavour. The values of γ_d and γ_j can be found in [20].

In summary, in the TMD factorisation TMDs we have two evolution equations: the CS kernel, which governs the rapidity evolution, and the RG, the evolution in μ. To determine the TMD the non-perturbative functions g_D, the non-perturbative function of the CS kernel, and $g_{j/A}$, the intrinsic transverse momentum, need to be fitted to data.

PB TMDs

In the simplest case in the PB method the starting TMDs can be defined as

$$\tilde{A}_{0,b/h}(x, k_T, \mu_0) = x f_{0,b/h}(x, \mu_0) \frac{1}{q_s^2} e^{-k_T^2/q_s^2}, \tag{2.37}$$

generated by introducing a transverse momentum k_T with a Gaussian distribution to the collinear starting distribution, $f_{0,b/h}$, where q_s is the width of the Gaussian distribution. Hence, to determine the TMD one can extract q_s from measurements. In [21] we have tuned the q_s parameter with measurements of the transverse momentum of the Drell-Yan pair at different centre-of-mass energies and Drell-Yan masses, were we have found $q_s = 1$ GeV.

The evolution of the PB TMDs is derived from the DGLAP evolution equations [22],

$$\begin{aligned} \tilde{A}(x, k_T, \mu^2) =& \Delta_a(\mu^2)\tilde{A}(x, k_T, \mu_0^2) + \sum_b \int_{\mu_0^2}^{\mu^2} \frac{d^2\mu'^2}{\pi\mu'^2} \frac{\Delta_a(\mu^2)}{\Delta(\mu'^2)} \\ & \times \int_x^{z_M} dz P_{ab}^{(R)}(\alpha_S(\mu'), z) \tilde{A}(x/z, \vec{k}_T - \vec{q}_T, \mu'^2) \end{aligned} \tag{2.38}$$

where Δ_a is the Sudakov form factor, which is interpreted as the probability of no emission between two scales, $\tilde{A}$ is the momentum weighted TMD, z_M is the maximum momentum transfer from parton b to a, $P_{ab}^{(R)}$ is the real part of the splitting function, μ' is the scale of the radiation and q_T is the transverse momentum of the radiated parton. The Sudakov form factor has the following form

$$\Delta_a(\mu_0, \mu) = \exp\left(-\int_{\mu_0}^{\mu} \frac{d\mu'^2}{\mu'^2} \int_0^{z_M} dz z P_{ab}^{(R)}(\alpha_S, z)\right). \tag{2.39}$$

We require that in the collinear limit the PB TMD evolution recovers the DGLAP evolution, as the PB evolution is derived from DGLAP. For that, we need to impose a $z_M \to 1$. Another important aspect of the evolution is the ordering of the emissions in the evolution. In the PB evolution angular ordering is used, which orders the emissions by the radiation angle. The first emission has the smallest angle, in the consequent radiations the angle increases. The angular ordering relates the p_T of the radiation with the radiation scale μ' as

$$p_T = (1 - z)\mu'. \tag{2.40}$$

The angular ordering arises from the condition to have a coherent branching [23].

By combining a $z_M \sim 1$ with an angular ordering the PB evolution favours soft radiation. PB-NLO-HERAI+II-2018-set1 (PBSet1) and PB-NLO-HERAI+II-2018-set2 (PBSet2) are TMDs derived using the PB method introduced above [18]. These two TMD sets are extracted using DIS measurements, more precisely precision measurements at HERA. These measurements are sensitive to the content of the hadrons at small fraction of momentum. The evolution of the TMDs is done at next-to-leading order. The starting scale of the TMDs is $\mu_0 = 1.3$ GeV with a width of the Gaussian distribution $q_s = 0.5$ GeV (see equation 2.37). The only difference between these two TMD sets is the treatment of the strong coupling α_S. PBSet1 uses the energy scale of the splitting, μ', to calculate α_S, while PBSet2 uses the transverse momentum of the splitting, p_T, where both are related by the angular ordering condition. The angular ordering condition $p_T < \mu'$, which translates into $\alpha_S(\mu') < \alpha_S(p_T)$, increases the probability of emission in the PBSet2 evolution. It is important to point out that in PBSet2 for values of $p_T < 1$ GeV α_S is frozen to $\alpha_S(1\text{ GeV})$, in order to not spoil the perturbative calculation due to the large value of α_S.

To understand the effects of increasing the radiation probability through the treatment of α_S we compare the up valence quark distribution of the two TMD sets: PBSet1 and PBSet2. Studying the up valence quark distribution gives a simpler understanding of radiations, as the number of up-valence quarks in the proton is constant and only gluon radiation affects their distribution. In figure 2.6 we show the collinear distribution (left panel) and the transverse momentum distribution (right panel). We choose an upper limit of the evolution of $\mu = 100$ GeV in order to allow more radiation. The collinear

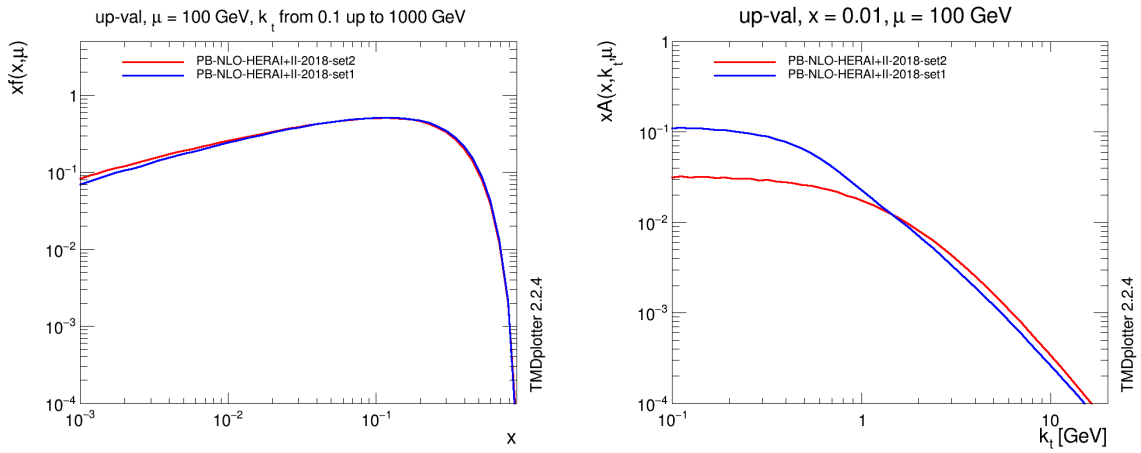

Figure 2.6: Up valence quark three dimensional contribution to the proton at $\mu = 100$ GeV for two different TMDs: PB-NLO-HERAI+II-2018-set2 and PB-NLO-HERAI+II-2018-set1 [18]. The difference between the two PDFs is the treatment of α_S, $\alpha_S(p_T)$ in set2, where p_T is the transverse momentum of the emitted parton, and $\alpha_S(\mu')$ in set1, where μ' is the scale of the branching. Left panel: momentum weighted collinear distribution. Right Panel: transverse momentum distribution for $x = 0.01$. Plots made with TMDplotter2.2.4 [19].

distribution is the same for most of the momentum fraction, x, range. However, there is a noticeable difference at low x. There is a larger probability to find an up-valence quark at low x in the PBSet2 PDF due to an increased probability of radiation. This is enhanced in the transverse momentum distribution. In the PBSet1 distribution the Gaussian distribution of the intrinsic transverse momentum dominates at low transverse momentum, while in PBSet2 the intrinsic transverse momentum is smeared by the extra radiations. In summary, the treatment of α_S does not affect much the evolution when looking in the collinear distributions, but it changes drastically the transverse momentum distribution. This difference will become noticeable and important when looking at proton-proton collisions.

2.4 Resummation

The QCD radiation of the two incoming particles modifies their transverse momentum. In the DY process the lepton pair is produced such that the pair recoils against the QCD radiation:

$$\vec{q}_T = -\sum_{i=0}^{n} \vec{p}_{T,i} \tag{2.41}$$

where q_T is the transverse momentum of the lepton pair and $p_{T,i}$ is the transverse momentum of i-th emission. Most of the DY lepton pairs are produced with low transverse momentum, where the main production channel is $q\bar{q} \to l^+l^- + g$, where the gluon has been radiated from one of the incoming quarks. This process is a leading order correction to the tree level production diagram. We follow the calculation in [24]. The cross-section of the DY process with a gluon radiation reads as

$$\frac{d\sigma}{dq_T^2} \propto (d\sigma)_{Born} \times \alpha_s \frac{1}{q_T^2} \ln \frac{m_{ll}^2}{q_T^2} \tag{2.42}$$

with $(d\sigma)_{Born}$ is the Born level cross-section for the DY process, α_S is the strong coupling and m_{ll} is the mass of the lepton pair. The cross section diverges for $q_T \to 0$ and for $q_T \ll m_{ll}$ the logarithm and α_S become large. The divergence behaviour arises from the lack of higher order corrections. We must sum all possible gluon radiations to the cross-section to all orders in order to remove the divergency. We include n gluon radiations in the DY process, $q\bar{q} \to l^+l^- + ng$. The computation of such matrix element is not possible and thus, we need to find an approximation. One of these approximations is the collinear limit, were we suppose that the radiation angle of these radiations is small. The cross-section for n radiations in the collinear limit is

$$\frac{d\sigma}{dq_T^2}^{(n)} \propto \left(d\sigma)_{Born} \times \prod_{i=1}^{n} \int \frac{dp_{T,i}^2}{p_{T,i}^2} \ln \frac{m_{ll}^2}{p_{T,i}^2} \delta \left(\sum_{i=1}^{n} \vec{p}_{T,i} - \vec{q}_T \right) \right. . \tag{2.43}$$

To proceed we need to find another approximation such that we can write the contribution of n radiations in a finite series. For that, we can impose a strong ordering condition of the radiations, where $p_{T,n} \gg p_{T,n-1} \gg \cdots \gg p_{T,1}$. In this way the δ function in the integral only applies in the last radiation. With the strong ordering condition we reach a geometric series such that

$$\prod_{i=1}^{n} \int \frac{dp_{T,i}^2}{p_{T,i}^2} \ln \frac{m_{ll}^2}{p_{T,i}^2} \delta \left(\vec{p}_{T,n} - \vec{q}_T \right) = \exp \left(\frac{1}{2} \ln^2 \frac{m_{ll}^2}{p_t^2} \right) \tag{2.44}$$

which resemble the well known Sudakov form factors. Now, we can define a finite cross-section by resumming gluon radiations to all orders to the DY process

$$\frac{d\sigma_{Res}}{dq_T} = \frac{d\sigma}{dq_T} \times e^S, \tag{2.45}$$

where σ_{Res} is the resummed and finite cross-section and S englobes any type series product of the different approximations. There is a small caveat for $q_T \ll m_Z$, the logarithm and α_S in equation 2.42 become too large and non-perturbative and thus this region has to be modelled. A good first approximation is modelling the intrinsic transverse momentum of the incoming partons with a Gaussian distribution (see section 2.5).

In summary, we can divide the Z boson transverse momentum spectrum in three main regions: the non-perturbative region ($q_T \ll m_Z$), the resummation region ($q_T < m_Z$) and the hard region ($q_T \geq m_Z$). The last step for a correct treatment of collinear and soft radiation is to match the resummation region with the hard region to avoid double counting of emissions that might happen.

It is important to note that resummation is necessary when two different scales appear in a calculation and one of them is much smaller in comparison to the other. Hence, many different resummation procedures can be found, such as, small x resummation, where the momentum fraction, x, is way smaller than the hard energy scale. However, in the following we only focus in the resummation for the transverse momentum of the DY lepton pair.

The resummation of the soft and collinear radiation is done in different ways, with transverse momentum factorisation theorems, with TMDs or with parton showers. In the following sections we will give an overview of a couple of transverse momentum factorisation theorems and a short introduction to parton showers. We will focus on the treatment of the $q_T \ll m_Z$ region for the different cases.

Transverse momentum factorisation theorems

We can find different factorisation theorems that treat transverse momentum resummation. We will briefly present two of these factorisations: the Collins-Soper-Stermann (CSS), the TMD and the high energy factorisation theorems.

The CSS factorisation was introduced in 1984 [25] to describe the transverse momentum of the electroweak bosons produced in proton-proton collisions. More precisely, the region where the transverse momentum of the electroweak boson is smaller than its mass, the resummation region. The cross-section for the electroweak boson production in proton-proton collisions given by the CSS factorisation theorem is

$$\begin{aligned}\frac{d\sigma}{dQ^2 dy dQ_{\mathrm{T}}^2} \sim \frac{4\pi^2\alpha^2}{9Q^2 s}\{(2\pi)^{-2} \int d^2b^{iQ_{\mathrm{T}}\cdot b} \\ \sum_j e_j^2 \tilde{W}_j(b_*; \frac{C_2^2 Q^2}{\mu^2} x_A, x_B)_{pert} \times \exp\left[S(b_*; Q, C_1, C_2)\right] \\ \times \exp\left[-\ln\big(Q^2/Q_0^2\big) g_D(b) - g_{j/A}(x_A, b) - g_{j/B}(x_B, b)\right] \\ + Y(Q_T; Q, x_A, x_B)\},\end{aligned} \tag{2.46}$$

where we sum over all the parton flavours i, j. The calculation is done is b space, the Fourier transform of the transverse momentum. Large (small) values of b correspond to small (large) values of transverse momentum. In the second line of equation 2.46 we find the perturbative functions of the factorisation theorem. The $\tilde{W}(b_*)_{pert}$ function

corresponds to the convolution of the collinear PDFs with the the hard-scattering function. The exponential with the S function takes care of the resummation with:

$$S(b_T, Q, Q_0, \mu_0) = -\int_{\mu_0}^{\mu_Q} \frac{d\mu'}{\mu'} \left(A(\alpha_S) \ln \frac{Q^2}{\mu'} + B(\alpha_S) \right), \tag{2.47}$$

where A and B are perturbative functions. It is important to note that when b is large, small transverse momentum, we are in a non-perturbative regime of the calculation, in the $q_T \ll m_Z$, which for the DY corresponds to $q_T < 1,2$ GeV ($b > 0.5,1$ GeV^{-1}). The other b regions can be calculated perturbatively. To avoid the large b region and calculate $\tilde{W}(b_*)_{pert}$ perturbatively, the b_* prescription is introduced

$$b_* = \frac{b}{\sqrt{1 + b^2/b_{max}^2}}, \tag{2.48}$$

in this way b_{max} is used as a cut-off to avoid the non-perturbative region. For $b < b_{max}$ $b_* \sim b$ and for $b > b_{max}$ we have $b_* \sim b_{max}$. Then, b_* is always in the perturbative regime and we can compute $\tilde{W}(b_*)_{pert}$ perturbatively. In the third line of equation 2.46 we have the non-perturbative functions of the CSS factorsation theorem. The g_j is the intrinsic transverse momentum of the incoming parton. The g_D function is the non-perturbative function describing the $q_T \ll m_{ll}$ region. This function, together with the intrinsic transverse momentum of the partons, has to be extracted from data. To extract g_D we need to measure the production of the electroweak boson at different energy scales. The Y function gives important corrections for $Q \sim Q_{\mathrm{T}}$ and negligible corrections at Q_{T}.

The CSS factorisation theorem for the electroweak boson production introduces the resummation together with the ideas of TMDs. In [26] it is discussed that the CSS factorisation theorem comes from a more general factorisation theorem, the TMD factorisation theorem:

$$\frac{d\sigma}{dQ^2 dy dq_T^2} \sim \frac{4\pi^2\alpha^2}{9Q^2 s} \sum_j H_{j\bar{j}}^{DY}(Q, \mu, \alpha_S) \int db_T^2 e^{iq_T \cdot b_T} f_{j/A}(x_A, b_T; \zeta_A, \mu) f(x_B, b_T; \zeta_B, \mu), \tag{2.49}$$

where f are the TMDs presented in section 2.3.2 and H^{DY} is the hard-scattering function for the partonic interaction for the DY process. In [20] a comparison of the CSS factorisation theorem with the newer TMD factorisation theorem is done. From the TMD factorisation theorem the TMD evolution presented in section 2.3.1 is derived.

Another general transverse momentum dependent factorisation theorem is the high-energy factorisation theorem [27]. The high energy factorisation of the cross-section for a hadronic collision is similar to the collinear factorisation in equation 2.25:

$$\begin{aligned} \sigma = \sum_{i,j} \int_0^1 dx_i \int_0^1 dx_j \int dk_{T,i} \int dk_{T,j} \; & f_{i/A}(x_i, k_{T,i}, \mu_F) f_{i/B}(x_j, k_{T,j}, \mu_F) \\ & \times \hat{\sigma}_{ij}(x_i, x_j, k_{T,i}, k_{T,j}, Q, \mu_F, \mu_R) \end{aligned} \tag{2.50}$$

where we sum over all the parton flavours, f are the TMD functions for the incoming hadrons, $k_{T,i}$ is the transverse momentum of the incoming parton and $\hat{\sigma}$ is the hard-scattering function which depends on the transverse momentum. The high-energy factorisation is only valid for small values of x. In the small-x region the t-channel diagrams dominate and the u-channel diagrams and the interference terms are negligible. However, at large x, where the u-channel and interference terms are not negligible, the high energy factorisation may not to hold.

The Parton Branching method

Transverse momentum factorisation theorems, such as, CSS and TMD (high energy) factorisation theorems, are only valid for low values of transverse momentum (momentum fractions) and might not hold at higher transverse momentums (momentum fractions). Moreover, as analytical calculations of CSS and TMD factorisation theorems are focused on inclusive final states. Calculating exclusive final states such as DY+0,1,2 partons within these frameworks becomes cumbersome.

The Parton Branching (PB) method includes TMD physics together with high transverse momentum physics in a fully exclusive way. To achieve a good description of exclusive final states predictions are performed on an event basis (see section 2.7). We can represent the PB method in a naive way as:

$$\sigma = \sum_{ij} \int dx_i dx_j \hat{\sigma}_{ij}(x_i, x_i, \mu) \int dk_{T,i} A_i(x_i, k_{T,i}, \mu) \int dk_{T,j} A_j(x_j, k_{T,j}, \mu) \tag{2.51}$$

where we sum over all the partonic flavours i and j, k_T is the transverse momentum of the incoming partons, μ is the scale of the partonic interaction, $\hat{\sigma}_{ij}$ is the partonic cross-section and A is the TMD. The transverse momentum of the incoming particles is extracted from the TMDs, which is included into the final state such that the integrated cross-section is unchanged.

Parton Showers

Parton showers resum the cross-section on an event basis. In section 2.7 we give an overview on the procedure of event generation. Parton showers resum the gluon emissions by evolving the incoming partons iteratively. The evolution of the incoming partons is performed as a backward evolution, since the final state of the process is already determined, and a backward evolution is more efficient.

Conventionally parton showers are based in the DGLAP evolution. The Sudakov form factor corresponding to the DGLAP evolution equations is

$$\Delta_{ab}(\mu_0, \mu) = \exp\left(\int_{\mu_0}^{\mu} \frac{d\mu'}{\mu} \int_0^{1-\epsilon} dz P_{ab}(z)\right). \tag{2.52}$$

The Sudakov form factor can be interpreted as the probability of a parton to evolve from a scale μ_0 to μ without radiating. By rewriting the DGLAP evolution equation with the Sudakov form factor,

$$f_a(x,\mu) = f_a(x,\mu_0)\Delta(\mu) + \sum_b \int_{\mu_0}^{\mu} \frac{d\mu'}{\mu'} \frac{\Delta(\mu)}{\Delta(\mu')} \int \frac{dz}{z} P_{ab}(z) f_b(x/z,\mu'), \tag{2.53}$$

we can solve the evolution equation iteratively by applying Monte-Carlo techniques. The first step to evolve a parton from μ_0 to μ iteratively is to generate z, the momentum transfer in the emission, according to the splitting function, P_{ab}. This will allow us to compute the energy scale of the splitting, μ', with the Sudakov form factor in equation 2.52. If μ' does not fall in between μ_0 and μ there will be no branching in the evolution. For a $\mu' \in (\mu_0, \mu)$ a branching is generated.

In a backward evolution we evolve from a high energy scale μ down to μ_0. For that, we need to add a constraint to the Sudakov form factor in equation 2.52:

$$\Delta_a(\mu)^{ISR} = \exp\left(\int_{\mu_0^\mu} \frac{d\mu}{\mu} \int_x^{z_{\max}} dz P_{ab}(z) \frac{f_b(x/z,\mu)}{f_a(x/z,\mu)} \right). \tag{2.54}$$

In this way we ensure that the parton previous to splitting was present in the hadron. The sum of the transverse momentum of the radiations together with the intrinsic transverse momentum of the incoming partons is then added to the final state particles. The non-perturbative intrinsic transverse momentum is generated with a Gaussian distribution in the same fashion as in PB.

2.5 Non-perturbative transverse momentum of the Drell-Yan lepton pair

Previously we have shown different ways to resum the soft and collinear radiation of the incoming partons. The correct treatment of these radiations is crucial for a good description of the transverse momentum of the DY lepton pair. However, for the $q_T \ll m_{ll}$ region perturbative resummation is no longer valid. In transverse momentum factorisation theorems this region is modelled and fitted from data. In the TMD factorisation theorem two functions are defined for that, $g_{h/A}$, the intrinsic transverse momentum and g_D the non-perturbative evolution of the TMD.

In iterative DGLAP evolutions, such as, the PB TMD evolution and parton showers, the non-perturbative treatment of the $q_T \ll m_{ll}$ is not explicit, as there is no specific function such the g_D function of the CSS and TMD factorisation theorems. In the DY process the lepton pair recoils against the QCD radiations of the incoming partons and its transverse

momentum is the sum off the transverse momentum of all the incoming partons. Hence, we find different scenarios for producing very soft lepton pairs where $q_T \ll m_Z$ with iterative evolution equations.

1. Events where incoming partons did not undergo much radiation. The transverse momentum of the incoming partons is given by the modelling of their intrinsic transverse momentum. Thus, the transverse momentum of the lepton pair is a product of the intrinsic transverse momentum of the incoming partons.
2. Events where the momentum of the incoming partons compensate each other, such that, the transverse momentum of the produced lepton pair is very small.
3. Events where the incoming partons have radiated very soft radiation, which smears the modelling of the intrinsic transverse momentum.

The first scenario is the largest contribution to the $q_T \ll m_{ll}$ region. Usually, the intrinsic transverse momentum is modelled with a Gaussian distribution. Let us follow the calculation in [23], where the transverse momentum distribution of the DY lepton pair is approximated to the product of the intrinsic transverse momentum of the incoming partons:

$$\frac{1}{\sigma}\frac{d\sigma}{dq_T^2} = \int dk_{T,1}dk_{T,2}\delta^{(2)}(\vec{k}_{T,1} + \vec{k}_{T,2} - \vec{q}_T)h(k_{T,1})h(k_{T,2}) \quad \text{with } h(k_T) = \frac{b}{\pi}\exp\left(-bk_T^2\right) \tag{2.55}$$

where b is the width of the Gaussian distribution and $k_{T,i}$ is the intrinsic transverse momentum of the incoming parton. The transverse momentum distribution of the DY lepton pair is the product of the two Gaussian distributions:

$$\frac{1}{\sigma}\frac{d\sigma}{dq_T^2} = \frac{b}{2\pi}\exp\frac{-bq_T^2}{2} \tag{2.56}$$

In figure 2.7 we compare the Gaussian approximation in equation 2.56 to measured DY lepton pair transverse momentum distributions by the CFS collaboration [28], in the same spirit as in [23]. From these results we observe that the modelling of the intrinsic transverse momentum of the incoming partons is a good approximation to describe the low transverse momentum region. In addition, the width of the Gaussian distribution is important for a good description of the data. For the two different centre of mass energies different values of the width of the Gaussian are needed. For $p_{lab} = 200$ GeV a value of $\langle k_T \rangle = 660$ MeV is needed, and a value of $\langle k_T \rangle = 760$ MeV for $p_{lab} = 400$ GeV. As the centre of mass increases we need larger values of intrinsic transverse momentum. This is the effect of very soft radiation. As we increase the energy there is more room for radiation and the very soft radiation starts to play a key role in describing the $q_T \ll m_{ll}$ region.

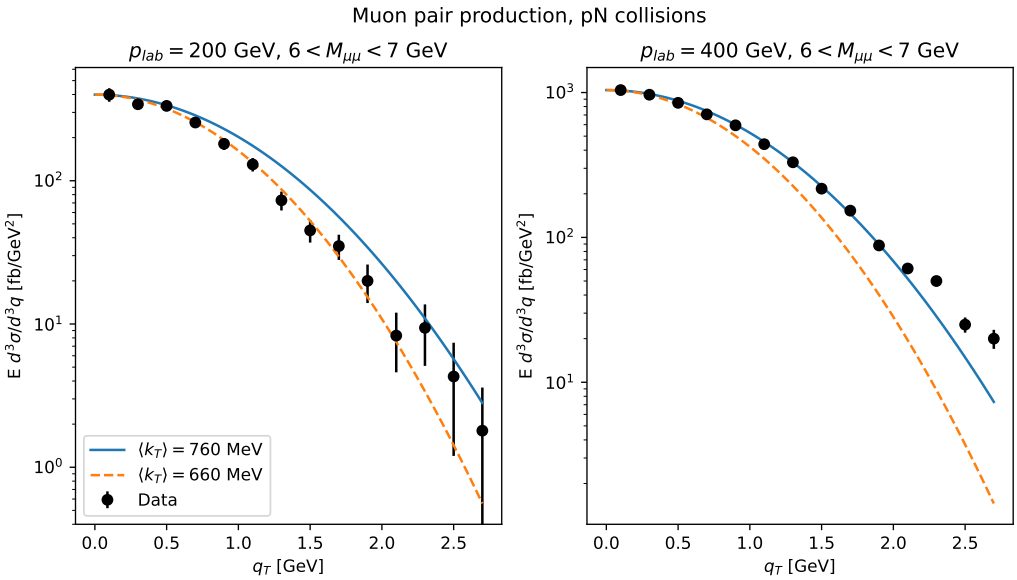

Figure 2.7: Cross-section of the Drell-Yan lepton pair measured by the CFS collaboration [28]. Lines represent the predictions of the cross-section with a Gaussian distribution of the intrinsic transverse momentum of the incoming partons.

By including very soft radiation in the evolution the Gaussian distribution is smeared, allowing the intrinsic transverse momentum distribution to stay constant. Thus, the third scenario presented above becomes important. We need very soft radiation to have predictive power over the $q_T \ll m_{ll}$ region at high energies and physical values of the intrinsic transverse momentum that do not exceed the size of the proton.

Let us consider the PB-TMD evolution, an angular ordered evolution, where the energy scale of the radiation, μ', and the transverse momentum of the emission are related as $p_T = (1-z)\mu'$. The softest radiations are hence produced when the momentum transfer is large, $z \sim 1$. The DGLAP splitting functions can be written as:

$$P_{ab}(\alpha_S, z) = \delta_{ab} d_a(\alpha_S)\delta(1-z) + \delta_{ab} k_a(\alpha_S)\frac{1}{(1-z)_+} + R_{ab}(\alpha_S, z), \tag{2.57}$$

where the plus prescription for any test function, $f(z)$, is defined as

$$\int_0^1 \frac{1}{(1-z)_+} f(z) dz = \int_0^1 \frac{1}{1-z}\left(f(z) - f(1)\right) dz, \tag{2.58}$$

which characterises the divergent behaviour for $z \to 1$, and d_a, k_a and R_{ab} are perturbative coefficients. We define the real part of the splitting function with

$$P^R_{ab}(\alpha_S, z) = \delta_{ab} k_a(\alpha_S)\frac{1}{1-z} + R_{ab}(z, \alpha_S), \tag{2.59}$$

which at leading order correspond to equations 2.32, and the virtual contribution with

$$P_a^V(\alpha_S, z) = k_a(\alpha_S)\frac{1}{1-z} - d_a(\alpha_S)\delta(1-z). \tag{2.60}$$

We make use of the momentum sum rule of the DGLAP splitting functions,

$$\sum_b \int_0^1 dz z P_{ab} = 0 \Leftrightarrow \sum_b \int_0^1 z P_{ba}^R = \int_0^1 dz P_a^V, \tag{2.61}$$

and we rewrite the PB Sudakov form factor in equation 2.39 in terms of virtual splitting functions

$$\Delta_a(\mu_0, \mu) = \exp\left(-\int_{\mu_0}^{\mu} \frac{d\mu'^2}{\mu'^2}\left\{\int_0^{z_M} k_a(\alpha_S)\frac{dz}{1-z} - d_a(\alpha_S)\right\}\right). \tag{2.62}$$

The Sudakov form factor contains a resolution scale, z_M, which does not allow emissions with $z > z_M$. Thus, to allow very soft radiations we set $z_M \sim 1$. From now on we refer to $z_M \sim 1$ as z_{fix}. However, we know that these soft emissions enter the non-perturbative regime of the transverse momentum of the Z boson. We introduce a minimum transverse momentum, q_0, for the radiated particles in the same spirit as b_{max} in the CSS factorisation. Radiations with $q_T > q_0$ are then considered perturbative and radiations with $q_T < q_0$ non-perturbative. We can introduce this condition with a dynamical resolution scale motivated with angular ordering $z_{dyn} = 1 - q_0/\mu'$. This will divide the Sudakov form factor in equation 2.62 into a perturbative (Δ_a^P) and a non-perturbative (Δ_a^{NP}) Sudakov form factors:

$$\begin{aligned}\Delta_a(\mu_0, \mu^2) =& \exp\left(-\int_{\mu_0}^{\mu} \frac{d\mu'^2}{\mu'^2}\left\{\int_0^{z_{dyn}} k_a(\alpha_S)\frac{dz}{1-z} - d_a(\alpha_S)\right\}\right) \\ &\times \exp\left(-\int_{\mu_0}^{\mu} \frac{d\mu'^2}{\mu'^2}\int_{z_{dyn}}^{z_{fix}} k_a(\alpha_S)\frac{dz}{1-z}\right) \\ =& \Delta_a^{(P)}(\mu_0, \mu^2)\Delta_a^{NP}(\mu_0, \mu^2).\end{aligned} \tag{2.63}$$

The effect of very soft radiation is encoded in non-perturbative Sudakov form factor:

$$\Delta_a^{NP}(\mu_0, \mu^2) = \exp\left(-\int_{\mu_0}^{\mu} \frac{d\mu'^2}{\mu'^2}\int_{1-\frac{q_0}{\mu'}}^{1-\epsilon} k_a(\alpha_S)\frac{dz}{1-z}\right), \tag{2.64}$$

where we have exchanged the integration limits $z_{fix} = 1-\epsilon$ with $\epsilon \ll 1$ and $z_{dyn} = 1-q_0/\mu'$. The perturbative coefficient k_a cannot be computed as we are in the non-perturbative region. Let us define h_k as the non-perturbative function of k_a, which does not depend on α_S. Integrating over z and μ' we get that the non-perturbative Sudakov form factor is

$$\Delta_a^{NP}(\mu_0, \mu) = \exp\left(-\frac{h_k}{2}\ln\left(\frac{\mu^2}{\mu_0^2}\right)\ln\left(\frac{q_0^2}{\epsilon^2\mu_0\mu}\right)\right), \tag{2.65}$$

which takes care of the very soft radiation. From here, we can observe that the non-perturbative Sudakov form factor derived from the PB evolution resembles the non-perturbative Sudakov form form factor of the CSS factorisation theorem, the g_D function (see equation 2.46). We can extract the corresponding g_D function of the CSS factorisation in the PB evolution:

$$\Delta_a^{NP}(\mu_0,\mu) = \exp\left(g_D^{PB}\ln\left(\frac{\mu^2}{\mu_0^2}\right)\right) : g_D^{PB} = \frac{h_k}{2}\ln\left(\frac{q_0^2}{\epsilon^2\mu_0\mu}\right), \tag{2.66}$$

which was introduced in the original CSS factorisation to describe the $q_T \ll m_{ll}$ region. Moreover, the g_D^{PB} function can be calculated analytically provided $\alpha_{\rm S}$ is in the perturbative region.

In conclusion, we can understand the non-perturbative Sudakov form factor of the TMD and CSS factorisation theorems from the PB evolution as the radiation of very soft gluons. However, g_D^{PB} can not be directly compared with g_D with g_D^{PB} as we will show later when comparing the TMD factorisation evolution and the PB evolution.

In figure 2.8 we show the gluon PDFs obtained with the Parton Branching method using the starting distribution of the PBSet2 BermudezMartinez:2018fsv for an evolution with a non-perturbative Sudakov form factor (PB-NLO-HERAI+II-2018-set2), and without the non-perturbative Sudakov form factor (PB-NLO-set2=qs0.5-Q0ord1.0). We observe that the collinear distribution is affected in the whole momentum fraction, x, range, where at small x there is a 40% difference between the two distributions. Thus, we can conclude that the non-perturbative Sudakov form factor has a large effect in the collinear distribution.

In the following we will discuss the evolution of different parton showers, Pythia8 and Herwig7, with a focus on their treatment of the non-perturbative evolution. Later we will compare the evolution of the PB evolution to the one of TMD factorisation.

Parton showers

The ordering of the evolution plays a large role in allowing very soft radiation. In p_T ordered evolutions the energy scale of the radiation, μ', is set as the transverse momentum of the radiation, $p_T = \mu'$. For an initial scale with $\mu_0 \sim 1-2$ GeV the softest radiation is $p_T^{min} = 1-2$ GeV and emissions below p_T^{min} are not considered. This will cause the non-perturbative region of the Z boson transverse momentum not to be treated (remember that for the Z boson the $q_T < 1,2$ GeVs is considered as non-perturbative). An angular ordered shower will allow softer radiation, as the transverse momentum of the emission is defined as $p_T = (1-z)\mu'$, therefore with angular ordering $p_T^{min} \sim 0$ GeV can be reached if we allow large momentum transfers in the radiation, $z \to 1$. Evolutions that do not allow large momentum transfers or that are p_T ordered will not have a predictive power for the

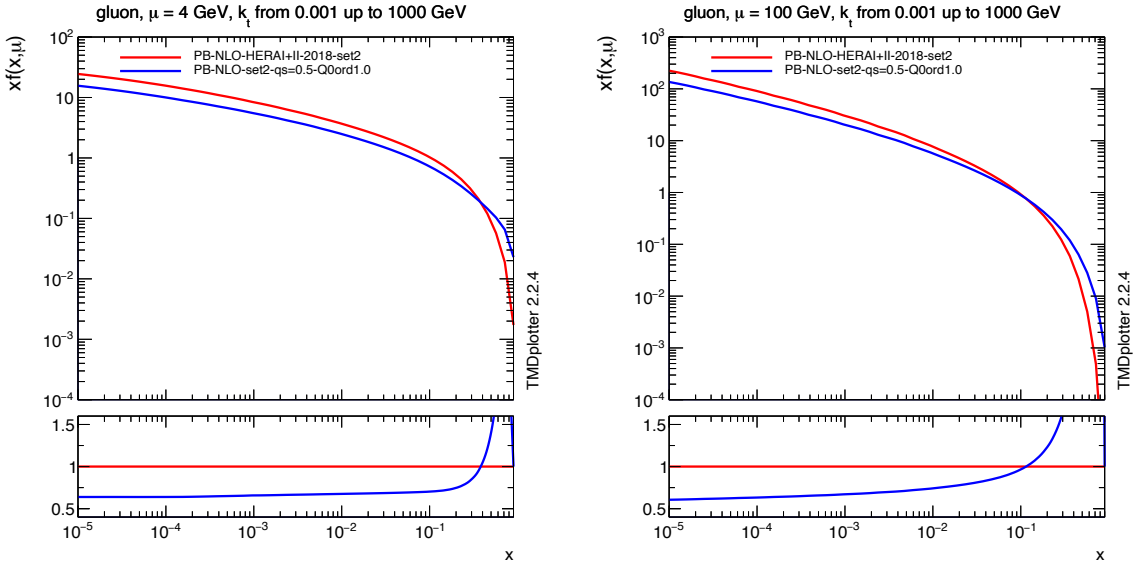

Figure 2.8: Collinear gluon distribution functions at two different energy scales, $\mu =$ 4 GeV (left panel) and $\mu = 100$ GeV (right panel), for two TMDs based on the PBSet2 starting distribution [18]: PB-NLO-HERAI+II-2018-set2 includes soft radiation in the $z \to 1$ region, while PB-NLO-set2=qs0.5-Q0ord1.0 which does not include emissions with $p_T < 1$ GeV. With these two distributions we observe the effect of the removing very soft radiation from the evolution. Plots made with TMDplotter2.2.4 [19].

$q_T \ll m_Z$ region, as the very soft radiation will not smear the Gaussian distribution of the intrinsic transverse momentum and thus, we will need larger values for the width of the Gaussian distributions.

To seize the effect of not considering soft radiation we tune the intrinsic transverse momentum distribution to different DY measurements with different parton showers, Herwig7 and Pythia8. With these two parton showers we cover two scenarios mentioned above, a p_T ordered shower (Pythia8) and angular ordered shower (Herwig7). Moreover, both parton showers introduce a dynamical resolution scale, z_M, which does not allow momentum transfers in radiations above z_M. Pythia8 and Herwig7 model the intrinsic transverse momentum with a Gaussian distribution, where the width of the distribution can be tuned. For a good description of the DY transverse momentum spectrum, more precisely the region of $q_T \ll m_{DY}$, the intrinsic transverse momentum distribution has to accommodate the untreated soft radiations. These radiations increase with increasing centre of mass energies, hence, the intrinsic transverse momentum has to also increase.

To tune the intrinsic transverse momentum parameters Pythia8 and Herwig7 we use the Professor2 software [29]. We generate the DY predictions at next-to-leading order accuracy in QCD with MC@NLO with the NNPDF3.1 PDF set at next-to-next-to leading order in QCD, with $\alpha_S(m_Z) = 0.118$. We supplement the events with the Pythia8 or Herwig7 Monte-Carlo event generator with a full simulation of the underlying event. For Pythia8 we use the next-to-leading order CP3, CP4 and CP5 tunes of the underlying event [30] and for Herwig7 the next-to-leading order CH2 tune of the underlying event [31]. For each different tune of the underlying event we tune the intrinsic transverse momentum parameter. In [32] we show that the variation of the intrinsic transverse momentum parameters does not affect the observables sensitive to the underlying event. We tune the intrinsic transverse momentum parameter to different DY measurements at different centre-of-mass energies and different DY lepton pair masses. In appendix A we present the DY measurements used for the tuning of the intrinsic transverse momentum, together with their corresponding tuned value and a small explanation on the considered uncertainties.

In figure 2.9 the tuned values of the intrinsic transverse momentum for a good description of the $q_T \ll m_{DY}$ region for both parton showers are shown. Both parton showers show a linear dependence on $\ln\sqrt{s}$ with a similar slope. Besides, the p_T ordered shower in Pythia8 needs larger intrinsic transverse momentum compared to Herwig7, the angular ordered shower. Besides, we extract the value of the intrinsic transverse momentum for the Pythia8 CP5 tune of the underlying event for a correct description of the DY transverse momentum spectrum at a centre-of-mass energy of 13 TeV:

$$2.67 \pm 0.02(\text{stat.}) \pm 0.03(\text{range}) \pm 0.04(\text{int.}),$$

where stat. stands for the statistical precision, range for DY transverse momentum region

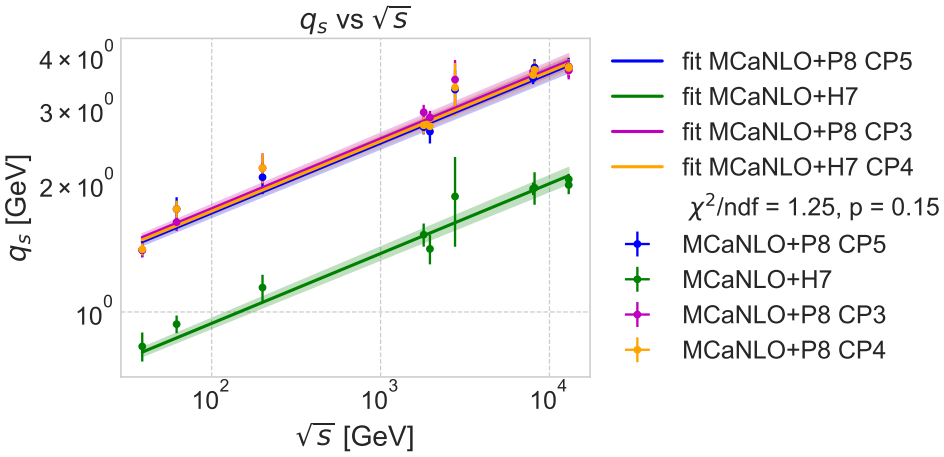

Figure 2.9: Tune of the intrinsic transverse momentum distribution for the PYTHIA8 and HERWIG7 parton showers for different centre of mass energies and different Drell-Yan lepton pair masses.

considered for the tune and int. for the interpolation uncertainty.

In comparison, the PB method shows no sensitivity to the width of the Gaussian distribution at different centre of mass energies [33]. Moreover, it shows an incredible agreement with the different measurements of the DY transverse momentum distribution. Recently a tune of the width of the Gaussian distribution of the intrinsic transverse momentum was performed in [21]. In this study, the width of the Gaussian distribution showed no dependence on the mass of the DY lepton pair or in the centre-of-mass energy.

The PB and TMD factorisation evolution

In this section we compare the evolution of the TMDs in the TMD factorisation theorem with PB evolution as a cross-check of the treatment of the resummation region of the transverse momentum distribution of the DY process. In the TMD factorisation theorem, the TMD evolution is governed by the following Sudakov form factor:

$$\Delta^{TMD} = \exp\left[(D(b_*,\mu)\ln\frac{Q}{\mu_{b_*}} + g_D(b_T,b_{max})\ln\frac{Q}{Q_0}\right] \times \exp\left[\int_{\mu_{b_*}}^{\mu}\frac{d\mu'}{\mu'}\gamma_d\ln\frac{Q}{\mu'} + \gamma_j\right]. \tag{2.67}$$

The first exponent represents the rapidity evolution, the CS kernel, with g_D the non-perturbative rapidity evolution and D the perturbative rapidity evolution. The second exponent represents the evolution in μ.

In the PB method the evolution is defined by the Sudakov form factor, which in equa-

tion 2.63 we have divided in a perturbative and non-perturbative form factor:

$$\Delta^{PB} = \exp\left[-\int \frac{d\mu}{\mu}\int_0^{1-q_0/\mu'} dz k_a(\alpha_S)\frac{1}{1-z} - d_a(\alpha_S)\right] \times \exp\left[g_D^{PB}\ln\frac{Q}{Q_0}\right] \tag{2.68}$$

By comparing the TMD factorisation and the PB Sudakov form factors we find direct relations for

$$\begin{aligned} \gamma_d &\to k_a, \\ \gamma_j &\to d_a, \\ g_D^{TMD} &\to g_D^{PB}. \end{aligned} \tag{2.69}$$

In [34] we show that the perturbative coefficients k_a and d_a agree with γ_d and γ_j correspondingly up to next-to leading logarithmic accuracy. At next-to-next-to leading logarithmic accuracy they do not longer agree. The disagreement is solved by the redefinition of α_S with the Catani-Marchesini-Webber (CMW) method [35]. However, this explanation surpasses the scope of this thesis. For the CS kernel, D, there is not such a relation, as there is no specific rapidity evolution in the PB method, only for the non-perturbative side of the evolution. Nonetheless, the PB evolution is angularly ordered, which might give rise to a rapidity evolution. To prove this statement one should differentiate the PB evolution in equation 2.38 in z, the momentum transfer, and relate z with rapidity using the angular ordering relation, $p_T = (1-z)\mu$. This becomes complicated as we need to factor out the Sudakov form factors in each emission in the evolution.

In [36] the extraction of the CS kernel from ratios of transverse momentum distributions of DY lepton pair at different DY lepton masses and centre of mass energies was proposed:

$$D(b,\mu) = \frac{\ln\frac{\Sigma_1}{\Sigma_2} - \ln Z(Q_1,Q_2) - 2\Delta_R(Q_1,Q_2;\mu_0)}{4\ln Q_1/Q_2} - 1 \tag{2.70}$$

where Q_1 and Q_2 are different DY lepton pair masses, Σ is the cross-section in b space for the corresponding Q and the Z and Δ functions are perturbative functions. With this method we are able to extract the CS kernel for the PB evolution from the cross-section ratios. In this work the rapidity evolution kernel of the PBSet2 TMD was extracted and compared to the rapidity evolution kernel of the TMD factorisation TMDs and to lattice QCD calculations of the g_D function. First, let us recall the main aspects of the PBSet2 TMD evolution. The emission in the evolution are angularly ordered and q_T is used as input for calculating α_S. In PBSet2 α_S is frozen for emissions with $p_T < 1$ GeV. We can interpret the cut in p_T of the emissions as $q_0 = 1$ GeV, the minimum transverse momentum to resolve non-perturbative and perturbative emissions. In figure 2.10 we show the results of [36], where the evolution kernel of PBSet2 is Cascade. We divide b, the distance parameter, in two regions, the $b < 1$ GeV^{-1} region, where the evolution kernel

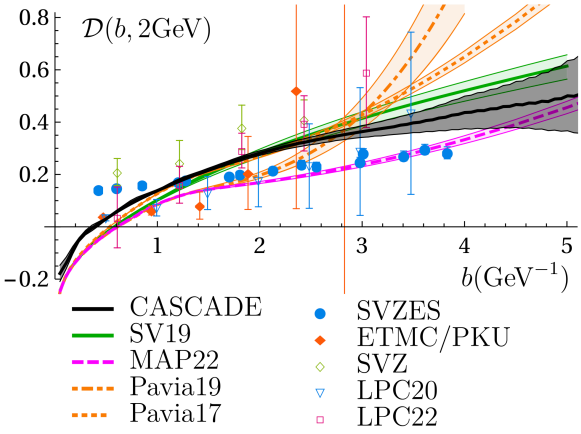

Figure 2.10: The TMD evolution kernel, D, for different TMD extractions. Cascade, correspond to the evolution kernel of PBSet2. Other solid lines correspond to TMDs extracted from fits: SV19 [37], MAP22 [38], Pavia19 [39] and Pavia17 [40]. Dots represent lattice calculations of the non-perturbative evoltion kernel: SVZES [41], ETMC/PKU [42], SVZ [43], LPC20 [42], LPC22 [44]. Figure taken from [36].

is perturbative and the $b > 1$ GeV^{-1}, where the evolution kernel is non-perturbative. The evolution kernel of PBSet2 shows a similar trend as the CSS evolutions in the perturbative region. In the non-perturbative region the parametrisation of the CSS evolutions start to differ. Lattice QCD calculations show a linear dependence on b, a similar trend to the PBSet2 trend.

To disentangle the different contributions to the non-perturbative function g_D in the PB evolution we compare the CS kernel of the PBSet2 TMD to the ones of PBSet1 and toy evolutions of the PBSet2 TMD with $z_m = z_{dyn} = 1 - q_0/\mu'$ with different values of q_0. In other words the toy models will give us information of not including soft radiations bellow q_0. The PBSet1 TMD was extracted together with PBSet2 in [18]. In PBSet1 μ' is used as input for α_S. Hence, in the evolution of PBSet1 soft radiations will not be as enhanced as in PBSet2. In figure 2.11 we show the evolution kernel, D, for the different TMDs. We observe by enhancing soft radiation the non-perturbative side of the evolution, $b > 1$ GeV^{-1}, becomes larger. In PBSet1 the evolution saturates at a value of $D \sim 0.1$, thus, showing that the treatment of α_S plays an important role in the rapidity evolution. The toy model with $q_0 = 1$ GeV saturates at $D \sim 0.2$ showing no b dependence in the large b region, and hence, not agreeing with the lattice QCD calculations shown in figure 2.10. The evolution kernel of the toy model with $q_0 = 0.5$ GeV behaves similar to PBSet2, however, the evolution kernel shows a stronger linear dependence compared to PBSet2. Besides, the calculation becomes unstable for the toy model with $q_0 = 0.5$ GeV as $\alpha_S(q_T = 0.5 \text{ GeV})$ is in the non-perturbative region.

From these results we can draw the following conclusion from the PB evolution:

- The PB evolution is able to reproduce the rapidity evolution of the TMD factorisation. The treatment of the very soft radiation has an impact on the non-perturbative side of the rapidity evolution, which in the TMD factorisation is modelled with the g_D function. Thus, we can understand that the g_D function in the TMD factorisation is composed by very soft radiations.
- The impact of the treatment of α_S in the rapidity evolution increases with b. The scenario with $\alpha_S(p_T)$ (PBSet2) has a larger rapidity evolution compared with the scenario with $\alpha_S(\mu')$ (PBSet1) due to the increased probability of emission with $\alpha_S(p_T)$.
- The integration range of the momentum transfer is important to get a b dependence of the rapidity evolution in the non-perturbative region. This also shows that the freezing of α_S for very soft raditions is a good approximation to describe the non-perturbative region of the rapidity evolution.

More importantly, phenomenological studies in [21,33], show that the measured transverse momentum distribution of the DY lepton pair is well described by resummed predictions with the PBSet2 evolution at different centre of mass energies and DY lepton pair masses. While predictions with PBSet1 and toy models with $q_0 > 1$ GeV need different values for the width of the Gaussian distribution modelling the intrinsic transverse momentum of the incoming particles to describe the $q_T \ll m_{ll}$ region.

2.6 Hadronisation

In the case that in the hard-process quarks or gluons are created, the hadronisation process becomes important. The outgoing quarks and gluons evolve from the hard-process energy scale to the hadronic energy scale. At this energy scale the coupling strength of the strong interaction is large enough to start the formation of hadrons. The hadronisation process is then a non-perturbative process.

Hadronisation can be treated through fragmentation functions or a modelling of the hadronisation process. Fragmentation functions (FF) give us the probability of a parton of a given flavour to hadronise into a certain hadron. FFs are similar to PDFs, they contain non-perturbative physics, which need to be extracted from data. FFs are supposed to be universal objects, the fragmentation of a parton into a given hadron should be independent of the process.

FFs can be used to calculate specific processes. In order to study all possible case scenarios we should extract the FFs for every quark or gluon decaying into a specific hadron. For a DY process where the Z boson decays into quark anti-quark pair and these hadronise

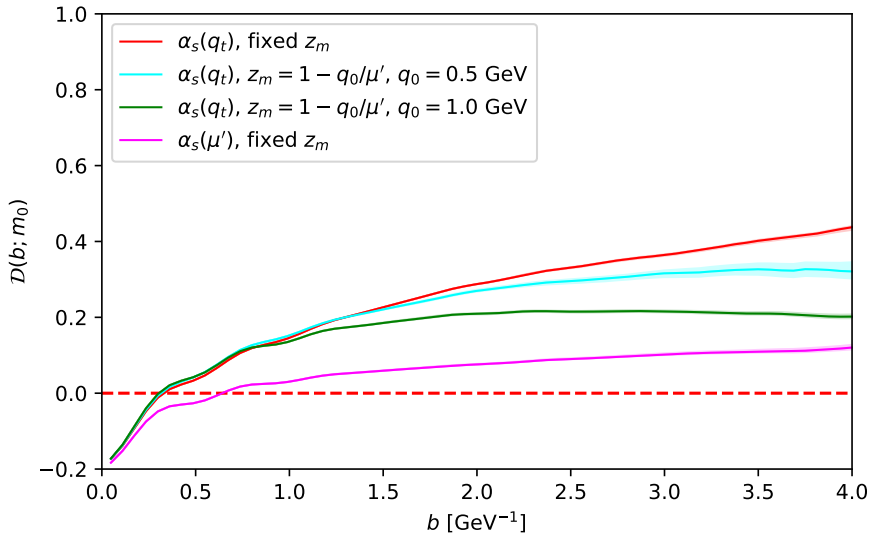

Figure 2.11: Rapidity evolution, CS Kernel, for different Parton Branching TMDs: PB-Set2 ($\alpha_{\mathrm{S}q_t}$, fixed z_m), PBSet1 ($\alpha_{\mathrm{S}}(\mu')$, fixed z_m) and toy models of the PBSet2 TMD where the $1 - q_0/\mu' < z < 1$ region is removed from the evolution, for different values of q_0.

into specific hadrons C and D the factorisation of the cross-section could be defined as

$$\sigma_h = \sum_{i,j,k,l} \int_0^1 dx_i \int_0^1 dx_j f_{i/A}(x_i, \mu_{\mathrm{F}}) f_{j/B}(x_j, \mu_{\mathrm{F}}) \hat{\sigma}_{ij\to kl}(x_i, x_j, Q, \mu_{\mathrm{F}}, \mu_{\mathrm{R}}) \times \int_0^1 dz_k \int_0^1 dz_l D_{k\to C}(z_k) D_{l\to D}(z_l), \tag{2.71}$$

where we sum over all parton flavours in the initial state i, j and final state k, l and $D_{k\to C}$ represents the FF of the outgoing parton k to hadronise into hadron C. If we are not interested in a specific type of hadron in the final state, we exchange the FF by a jet function. A jet is a collimated spray of hadrons arising from the QCD bremsstrahlung of the outgoing parton as it evolves from the hard energy scale to the hadronisation scale. Later, these shower of particles hadronises. The jet function describes the evolution of the outgoing particle from the hard energy scale to the hadronisation scale in a jet of radius R, $J(R, \mu)$. The evolution of the outgoing particles will depend on the factorisation of the cross-section. With a TMD factorisation of the cross-section, the evolution of the outgoing and incoming particle is the same.

The hadronisation is also modelled and introduced in parton shower Monte Carlo event generators. Two main models of the hadronisation are used:

- **The Lund string model** approximates the strong interactions at non-perturbative energy scales with a linear potential, which is represented by a string. After the hard-scattering the outgoing particles are evolved with a forward parton shower to the an hadronisation scale, which is set around 1-2 GeVs. Once all particles in the event are evolved to the hadronisation scale, the quark and anti-quarks are connected through a string forming quark anti-quark pairs. When the energy of the pair is larger than the one from the string, the string breaks. At the breaking position a new quark anti-quark pair is generated from the vacuum and the newly produced quark (anti-quark) will paired with the "old" anti-quark (quark).
- **The cluster model**, as the name suggests this model of hadronisation is based on clustering quarks and anti-quarks at the hadronisation scale. In the first step, gluons at the hadronisation scale are splitted non-perturbatively into quark anti-quark pairs. Then, the surrounding quarks and anti-quarks are combined forming color neutral objects, hadrons.

2.7 Event generators

Previously we have presented the different theoretical aspects involved in proton-proton collisions. This knowledge then can be applied to simulate proton-proton collisions. The generation of proton-proton collisions can be divided in two steps: the generation and the showering.

In the event generation, we need to define the process under study e.g. the DY process: two incoming protons collide generating a pair of leptons. The events are generated according to the collinear factorisation theorem (see equation 2.25). First, the transition amplitudes are calculated, the probability of a quark anti-quark to transform to a lepton pair for the DY case. Then, we need to compute the partonic-cross section. For that, we need the momentum of the two incoming particles. The momentum and flavour of the incoming partons are sampled using the PDFs. The integration of the cross-section will give us the probability for a specific configuration. Then, the events will be generated according to the proability distribution.

The generated events in the generation step are not physical, these need to be showered with a parton shower (see section 2.4). The parton showers divide the showering process into an initial and final state showers. Interference terms are neglected between initial and final state are neglected. The initial state shower evolves the initial state partons from the partonic scattering scale to the hadron scale with a backward evolution. For the case where gluons or quarks are generated in the final state, these outgoing partons are evolved from the partonic interaction scale to the hadronisation scale. The hadronisation scale is an arbitrary parameter around 1-2 GeVs.

In the case quarks and gluons are considered in the final state the hadronisation process is included. In section 2.6 we have presented the two main modelling of hadronisation used in event generators. To make events more realistic multi parton interactions can be taken into account, as when two protons collide more than one partonic interaction can happen. We refer to the combination of parton shower, multi parton interactions and hadronisation as the underlying event.

Event generators are a great way to study proton-proton collisions on a event basis. These predictions are complementary to analytical predictions, as they offer a easier way to study exclusive observables. Moreover, these simulations are heavily used for measurements for a better understanding of the measured data and to correct the measurements for detector effects. In the following we mention some of the event generators and parton showers available in the market:

- **Madgraph** [45] is an event generator which computes the generation step described above. The generated events then can be fed to a wide range of parton shower Monte-Carlo event generators. In Madgraph the events can be calculated at leading order and up to next-to-leading order in QCD interactions and electroweak interactions [46]. However, for electroweak interactions only virtual corrections are included (see chapter 3).

- **Pythia8** [47] is a multipurpose event generator where event generation, parton showers, hadronisation and multi parton interactions are available. The events are generated at leading order in perturbation theory. External events at higher orders can be fed to Pythia8. The shower is a transverse momentum ordered shower, which relates the splitting scale μ', with the transverse momentum of the emission p_T as $\mu' = p_T$ and the hadronisation is modelled with the Lund string model. The modelling of multi parton interactions is described in [48].

- **Herwig7** [49,50] is a multipurpose event generator which includes event generation, parton showers, hadronisation and multi parton interactions. Events are generated at leading order in perturbation theory. Herwig7 offers an interface to event generators, such as, Madgraph to generate events at higher orders. Besides, external events can be used for later showering, etc. The parton shower is an angular ordered shower and the hadronisation is modelled by the Clustering method. Information on the modelling of multiparton interactions can be found in [49].

- **CASCADE3** [51] is a multipurpose generator with a parton shower based on TMDs, which follows exactly the PB-TMD distribution for the initial state shower. For the final state the Pythia6 [52] shower is used. Both initial and final state showers are angular ordered. Events generated by Madgraph at next-to-leading order accuracy in QCD can be fed to CASCADE3, as well as events generated according to processes calculated within the k_T-factorisation.

CHAPTER 3

Electroweak corrections

In the previous chapter we have studied the different theoretical aspects to describe particle interactions. However, at the time of this thesis the precision of measurements by the Large Hadron Collider experiments is unprecedented. The uncertainties of precision measurements is smaller than the ones from theoretical calculations. For a better understanding of the measurements we need an improvement of theoretical calculations. Let us consider the collinear factorisation theorem in equation 2.25. The theoretical prediction is composed by three functions: two PDFs and one hard-function. To improve the theoretical calculations we must improve these functions.

PDFs are non-perturbative objects. The starting distribution of the PDF has to be determined from measurements, while the evolution of the PDF is done perturbatively. The hard-function is a perturbative object. We can improve the PDF evolution and the hard-function by including more terms to their perturbative expansion. The additional terms to the leading order term of the expansion are known as corrections. In chapter 2 we have discussed the different areas where QCD higher order corrections can be applied to achieve a better description of data, e.g., in PDFs, resummation and hard-functions. Due to the coupling strength of QCD interactions the perturbative expansion is done in terms of the strong coupling. However, it is known that electroweak corrections become large at high energies [53]. In this thesis we have measured the contribution of real electroweak corrections to multijet events, thus it is important to have solid theoretical understanding of electroweak corrections. In this chapter we introduce electroweak corrections from a theoretical point of view.

Corrections are divided in two categories: virtual and real corrections. Virtual corrections are loop interactions, a particle is emitted and absorbed again. Real corrections are radiations of particles. Corrections for massless particles are infrared divergent. These

are known as mass singularities. Mass singularities from virtual corrections cancel with the corresponding singularities of the real corrections. For Abelian theories the cancellation of the divergences is understood by the Bloch-Norsviek theorem [54]. For non-Abelain theories, such as QCD, it is understood by the Kinisoshita-Lee-Neuenberg (KLN) theorem [55, 56]. The KLN theorem states that when summing over all degenerate states, mass singularities cancel. For QED and QCD theories, where the bosons are massless, we include virtual and real corrections in the calculation to cancel the mass divergences. For massive particles the mass acts as a cut-off for the mass singularity. This is the case for electroweak theory, where the bosons are massive. Due to the mass of the electroweak bosons virtual and real corrections are finite. Hence, we do not need to include virtual and real corrections in the calculation to achieve a finite result.

This is an advantage for theoretical predictions when looking at exclusive observables. A real electroweak correction changes the defined final state of the exclusive observable. Thus, only virtual corrections are introduced in the calculation of exclusive observables. In this chapter we explain the behaviour of electroweak corrections at high energies for virtual and real corrections, the violation of the BN and KLM theorems in inclusive observables. In section 3.3 we will investigate electroweak corrections in hadron colliders, their impact in the hard-function, PDFs, resummation and how phase space restrictions in measurements can lead to enhanced electroweak effects.

3.1 Virtual electroweak corrections

Virtual electroweak corrections become important at high energies. To understand their behaviour at high energies we study the vertex correction (see figure 3.1). We define k and M as the momentum and mass of the particle in the loop and p_i and m_i as the momentum of the outgoing particles, where $i = 1, 2$. We follow the calculation of [53]. The vertex correction can contain infrared divergences ($k \sim 0$). Integrals of the vertex correction containing k in the numerator are infrared safe by power-counting. Then, the only integral that can contain an infrared divergence is C_0:

$$C_0(m_1, m_2, M_V, p_1, p_2) = \frac{i}{2\pi} \int \frac{d^4k}{[(p_1+k)^2 - m_1^2 + i\epsilon][k^2 - M_V^2 + i\epsilon][(k-p_2)^2 - m_2^2 + i\epsilon]}. \tag{3.1}$$

For the massless case, such as QCD and QED, where $m_1 = m_2 = M = 0$

$$C_0(0, 0, 0, p_1, p_2) = \frac{i}{2\pi} \int \frac{d^4k}{[(p_1+k)^2 + i\epsilon][k^2 + i\epsilon][(k-p_2)^2 + i\epsilon]}, \tag{3.2}$$

and we find the integral is divergent for three cases: soft divergence ($k \sim 0$) and two collinear divergences ($k \sim p_2$, $k \sim -p_1$). For purely electroweak corrections, where

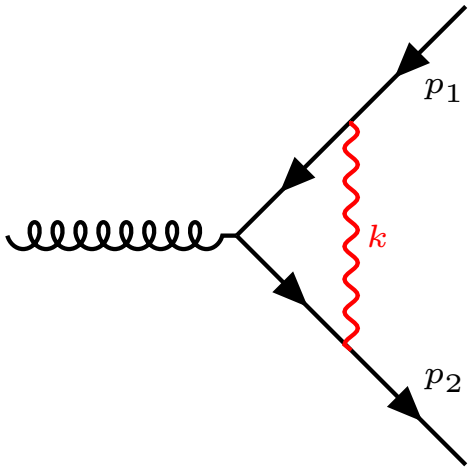

Figure 3.1: Vertex correction induced by a electroweak boson with momentum k and mass M.

$M \neq 0$, C_0 does not contain soft divergence. However, in the large energy limit, $s \equiv (p_1 + p_2)^2 \gg M^2$, the real part of C_0 becomes:

$$\mathrm{Re}[C_0(0, 0, M_V, p_1, p_2)] \propto \frac{1}{2s} \log^2 \left(\frac{s}{M^2} \right). \tag{3.3}$$

The logarithmic behaviour at high energies is known as a Sudakov logarithm. Note the difference with the Sudakov logarithm arising from QCD radiation (see equation 2.42). The Sudakov logarithm arising from EW radiation is a double logarithm compared to the single logarithm of QCD radiation.. The leading order vertex, $\mathscr{V}_0$, is modified by the electroweak vertex correction as

$$\mathscr{V} \sim -\frac{\alpha_{EW}}{4\pi} \mathscr{V}_0 \log^2 \left(\frac{s}{M^2} \right). \tag{3.4}$$

We see that at high energies the contribution of virtual electroweak correction becomes large and negative.

The interest of this thesis lies on the contribution of electroweak corrections to dijet process. The contribution of virtual electroweak corrections for dijet processes in hadronic collisions have been studied in [57,58]. We follow the study in [58]. Two ways of producing a dijet final state are considered: QCD production and electroweak production. For the QCD production the dijet is produced through interactions between gluons and quarks. The electroweak production of the dijet consists of the production of an electroweak boson that decays into two quarks. The calculation is done at $\mathscr{O}(\alpha_s^2 \alpha_{EW})$, QED processes are not included. The cross-section at leading order, σ_0, is calculated at $\mathscr{O}(\alpha_S^2, \alpha_S \alpha_{EW}, \alpha_{EW}^2)$. Due to the small contribution of the electroweak production compared to the QCD correction,

the electroweak production is treated as a corrections such that

$$\sigma_0 = \sigma_{QCD} \times (1 + \delta^{EW}_{\text{tree}}) \tag{3.5}$$

where σ_{QCD} denotes the leading order cross-section of the QCD production of the dijet and δ_{EW} the cross-section of the electroweak production. The next-to-leading cross-section is then defined as:

$$\sigma^{NLO} = \sigma_0 \times (1 + \delta^{EW}_{\text{1-loop}}) \sim \sigma_0^{QCD}(1 + \delta^{EW}_{\text{tree}} + \delta^{EW}_{\text{1-loop}}) \tag{3.6}$$

where $\delta^{EW}_{\text{1-loop}}$ are the one-loop virtual electroweak corrections to the QCD diagrams. The electroweak dijet production does not contain electroweak virtual corrections. For the total cross-section it is found that virtual electroweak corrections are negligible. However, for some observables the contribution of the virtual electroweak corrections becomes important. The observables that are studied are the invariant mass of the two leading jets and the transverse momentum of the two leading jets.

In figures 3.2, 3.3 we show the contribution of virtual electroweak corrections (δ^{1-loop}_{EW}), the electroweak production of dijets (δ^{tree}_{EW}) and the sum of both contributions to the leading order QCD calculation for the different observables obtained in [58]. For the three observables it is shown that the contributions of virtual electroweak corrections are negative and increasing with increasing values of the observables. For the dijet invariant mass (M_{12}) the effect is below -6% for invariant masses of 5 TeV. The contribution of the virtual electroweak corrections is enhanced for the transverse momentum of the two leading jets. For a leading jet with transverse momentum of 3 TeV the virtual electroweak correction has an effect of -10%, while for a sub-leading jet it goes up to -15%. At large invariant masses of the dijet the transverse momentum of the dijet system is small. This makes virtual electroweak corrections smaller at high invariant masses compared with the transverse momentum of the leading jets.

The contribution of the electroweak production of the dijet is positive. The total cross-section in equation 3.6 is calculated by summing the contributions from the virtual electroweak corrections and the electroweak production. Hence, there is a cancellation between both contributions [1]. For the invariant mass of the dijet system the contribution of virtual electroweak corrections is larger than the electroweak production. Hence, the total contribution to the QCD calculation is negative -3% at $M_{12} \sim 6$ TeV. For the transverse momentum of the two leading jets the contribution from the electroweak production is larger than the virtual electroweak corrections at large transverse momentum. The sum of both contributions is positive and around 5% at high transverse momentum.

[1] The cancellation between the virtual electroweak corrections and the electroweak production does not refer to the cancellation between real and virtual electroweak corrections. We will discuss this effect in more detail in the following section.

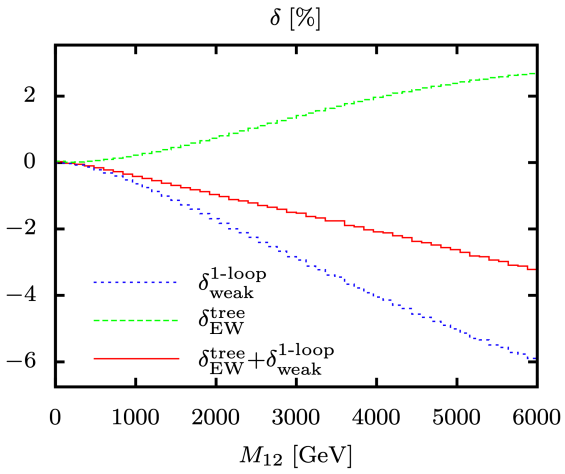

Figure 3.2: Relative differential cross-section to the leading order QCD cross-section for dijet production at a centre-of-mass energy of 14 TeV in the invariant mass of the two leading jets. The electroweak production of the dijet (δ^{tree}_{EW}) is shown with green dashed lines. The virtual electroweak corrections (δ^{1-loop}_{EW}) are shown with the dotted blue line. The figure is taken from [58].

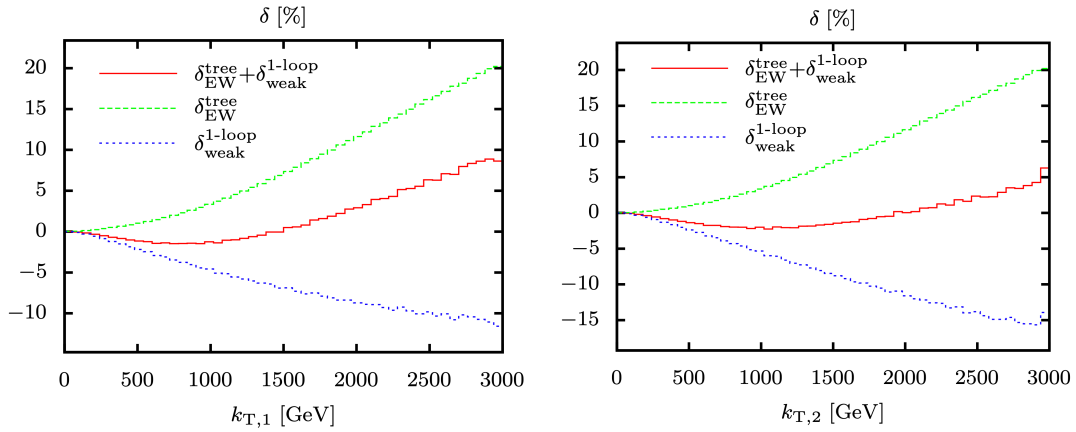

Figure 3.3: Relative differential cross-section to the leading order QCD cross-section for dijet production at a centre-of-mass energy of 14 TeV. Left panel: differential in the leading jet transverse momentum $k_{T,1}$. Right panel: differential in the sub-leading jet transverse momentum $k_{T,2}$. The electroweak production of the dijet (δ^{tree}_{EW}) is shown in green dashed lines. The virtual electroweak corrections (δ^{1-loop}_{EW}) are shown in dotted blue line. The figure is taken from [58].

3.2 Real electroweak corrections

We have argued that it is possible to separate real and virtual electroweak corrections. Virtual corrections are not infrared divergent, hence we do not need the corresponding real emissions to cancel a divergence. For exclusive observables, such as dijet events, a real radiation would change the final state.

For inclusive observables, such multijet states, we need to include real electroweak corrections. The sum of real and virtual contributions leads to a partial cancellation. The partial cancellation between real and virtual corrections are known as Bloch-Nordsieck (BN) violation. The Bloch-Nordsieck theorem states that for Abelian theories the infrared divergences cancel with the corresponding real correction. However, this is not true for non-Abelian theories. For QCD the cancellation is true due to the KLN theorem. Since hadrons are symmetric under SU(3) we can average over all the possible color charges of the constituents cancelling the mass singularities. For electroweak corrections it is not always possible to average over all isospin charges. Hence, the NB violations arise. In the following we investigate the NB violations, their origin and the scale of the violation. Then, we present the real electroweak corrections and the emissions of electroweak bosons in hadronic collisions.

3.2.1 Bloch-Nordsieck violation

The BN violations have been studied for electron-positron collisions in [59–63] and for hadronic collisions in [64]. To understand the origin of BN violations, let us follow the work in [63]. Here three cases are presented to pinpoint the origin of BN violations: an Abelian theory with massive bosons, a spontaneously broken SU(2) symmetry and the SM. Moreover, the effect of different phase space cuts on BN violations are studied.

The process under study is the four fermion process, $f_1\bar{f}_2 \to f_3\bar{f}_4$. For the Abelian theory with massive bosons the Born cross-section for the four-fermion process (σ_B) is modified by the virtual corrections as:

$$\sigma^{(V)} = \frac{\alpha}{4\pi}\left\{-4\left[\ln^2\frac{s}{M^2} - 3\ln\frac{s}{M^2}\right] + \cdots\right\}\sigma_B. \tag{3.7}$$

In the four-fermion process the radiation of the massive boson can happen from the initial and final state. To compute the contribution of the real emission a cut is applied to the minimum invariant mass of the fermion pair, $Q^2 \leq Q_c^2 \gg M^2$, where Q is the invariant mass of the lepton pair, Q_c is the minimum cut on Q and M is the mass of the massive boson. The cross-section for the four-fermion process including real radiation to the Born

process is:

$$\sigma^{(R)} = \frac{\alpha}{4\pi}\left\{4\left[\ln^2\frac{s}{M^2} - 3\ln\frac{s}{M^2}\right] - 4\left[\ln\frac{z_c}{(1-z_c)^4} + \cdots\right]\ln\frac{s}{M^2} + \cdots\sigma_B \right. \quad \text{with } z_c = Q_c^2/s \tag{3.8}$$

By summing 3.8 and 3.7 we get the total cross section:

$$\sigma_{Tot} \propto \ln\frac{z_c}{(1-z_c)^4}\ln\frac{s}{M}\sigma_B \tag{3.9}$$

The surviving logarithm corresponds to a radiation of the boson from an initial state fermion.

For a spontaneously broken SU(2) theory, the theory for weak interactions, we have three massive gauge bosons: $W^\pm$ and W^3. The isospin charges are $t^\pm$ and t^3. For simplicity the corrections that affect one of the outgoing quarks are studied. More precisely on f_3. The virtual corrections on the f_3 leg modify the Born cross-section in the collinear limit as:

$$\sigma^V = -\frac{\alpha}{4\pi}\left[\ln^2\frac{s}{M^2} - 3\ln\frac{s}{M^2}\right]\left((t^3)^2 + (t^\pm)^2\right)\sigma_B^{f_1\bar{f}_2\to f_3\bar{f}_4} \tag{3.10}$$

For the real correction we have to take into account that the radiation of a $W^\pm$ boson from the final state leg f_3 will change the flavour of the final state of f_3 to it's isospin conjugate. Then, to get a f_3 fermion in the final state we need to separate the W^3 and the $W^\pm$ interactions. The Born cross-section for the W^3 radiation is $\sigma_B^{f_1\bar{f}_2\to f_3\bar{f}_4}$. For the $W^\pm$ emission we need the isospin conjugate of f_3 in the final state, $f_3^\pm$, of the Born cross-section, $\sigma_B^{f_1\bar{f}_2\to f_3^\pm\bar{f}_4}$. The emission of a $W^\pm$ will change $f_3^\pm$ into f_3. The real emission cross-section is

$$\sigma^R = \frac{\alpha}{4\pi}\left[\ln^2\frac{s}{M^2} - 3\ln\frac{s}{M^2}\right]\left((t^3)^2\sigma_B^{f_1\bar{f}_2\to f_3\bar{f}_4} + (t^\pm)^2\sigma_B^{f_1\bar{f}_2\to f_3^\pm\bar{f}_4}\right) \tag{3.11}$$

We get the total cross-section by summing equation 3.10 and equation 3.11 :

$$\sigma_{Tot} = \frac{\alpha}{4\pi}\left[\ln^2\frac{s}{M^2} - 3\ln\frac{s}{M}\right](t^\pm)^2\left(\sigma_B^{f_1\bar{f}_2\to f_3^\pm\bar{f}_4} - \sigma_B^{f_1\bar{f}_2\to f_3\bar{f}_4}\right), \tag{3.12}$$

As we can see, only the $W^\pm$ bosons contribute to the non cancellation between real and virtual corrections. In comparison with the Abelian theory with a massive boson in the broken SU(2) theory, a double logarithm and a single logarithm survive. We can note that in a final state where we average over the isospin charges, the cancellation between the real and virtual corrections will be complete. In the soft limit we have a single logarithm from the $W^\pm$ real radiation.

From this analysis we understand that only the $W^\pm$ contribute to BN violations. We conclude that collinear radiation contributes with a double logarithm to the BN violations, while the soft radiation with a single logarithm. Hence, phase space cuts will affect to the size of the BN violations.

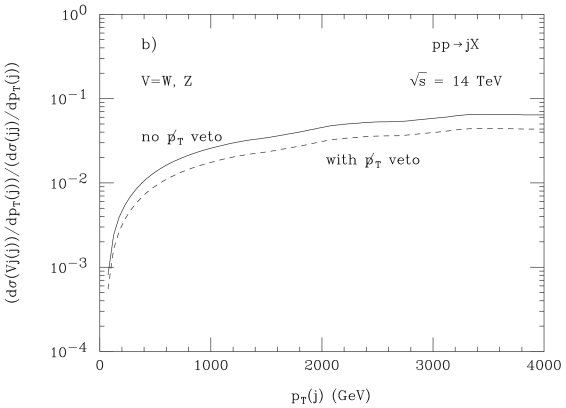

Figure 3.4: Ratio between the jet transverse momentum cross-section from leading order V+1,2 jets (V+j(j)) and the next-to leading order QCD calculation of dijet production. Dashed lines represent cuts on missing transverse energy. V represents the electroweak bosons Z and $W^{\pm}$. Dashed (solid) line (no) veto on missing transverse energy, $\not{p}_T$ [64].

3.3 Real electroweak corrections in hadronic collisions

We have presented the origins of the BN violations, where the cancellation between real and virtual electroweak corrections is partial. In inclusive measurements in proton-proton collisions we can average over the isospin charges in the final state. However, we can not average over the incoming particles in the initial state. The incoming protons are not symmetric under SU(2). Therefore, only electroweak corrections in the initial state will contribute to BN violations, and moreover, only corrections involving $W^{\pm}$ will contribute. In [65, 66] we can find a detailed study on BN violations for dijet processes and which diagrams are involved in BN violations.

In [64] the contribution of real electroweak corrections is studied in hadronic collisions at a centre-of-mass energy of 14 TeV. The real electroweak corrections are treated by calculating V+1,2 jets, where V represents an electroweak bosons Z or $W^{\pm}$ at leading order. To study the contribution of real electroweak corrections to inclusive jet production, the V+1,2 jets is compared to a next-to-leading order QCD calculation. In figure 3.4 the ratio between the QCD and V+1,2 calculation is shown depending on the jet transverse momentum. The ratio between the two calculations increases logarithmically with the transverse momentum of the jets. The ratio at high jet-transverse-momentum, $p_T^j = 4$ TeV is around $6 \cdot 10^{-2}$.

From [58, 64] we have learned that the effect of electroweak production of dijet, virtual and real electroweak corrections are important at high energies. While virtual electroweak

corrections are negative, the electroweak production and real emissions are positive. This leads to a cancellation of the contributions. From figures 3.3 3.4 we could expect that by including the real electroweak corrections on top of the electroweak dijet production and the virtual electroweak corrections, the dijet cross-section will increase around $1-5\%$ for high transverse momentum jets.

3.3.1 Electroweak parton distribution functions

So far we have studied electroweak corrections in the context of the hard-function. For a complete treatment of electroweak corrections we need to expand electroweak corrections to the whole factorisation formula. Collinear and soft radiations below the factorisation scale are not treated by the hard-function, but by the PDFs. Hence, we will need electroweak PDFs for the colliding hadrons to factorise the collinear and soft singularities we observed in the BN violations studies.

Electroweak PDFs are not widely studied, the first approximations for the electroweak boson PDFs dates back to 1980s [67, 68]. In these approximations the electroweak boson PDFs are generated dynamically, the electroweak boson is generated from a radiation of a quark. There is no intrinsic content of the electroweak bosons in the proton. Electroweak bosons have longitudinal and transverse polarisation. For each polarisation a different PDF is needed as they involve different interactions. From [68] the PDFs for the electroweak bosons are defined as:

$$
\begin{aligned}
f_V^{\perp}(x, m_v, p_T) &= \frac{g_V^2 + g_A^2}{8\pi^2} \frac{1+(1-x)^2}{x} \ln\left(\frac{p_T^2 + (1-x)m_V^2}{(1-x)m_V^2}\right) \\
f_V^{L}(x, m_v, p_T) &= \frac{g_V^2 + g_A^2}{4\pi^2} \frac{(1-x)}{x} \frac{p_T^2}{p_T^2 + (1-x)m_V^2} \\
&\text{with:} \\
&\text{For: } W^{\pm} \rightarrow g_A = g_V = g_2\sqrt{2} \\
&\text{For: } Z^0 \rightarrow g_V = \frac{1}{2}\left(T_3^f - 2Q_f \sin^2\Theta_W\right) \frac{g_2}{\cos\Theta_W} \text{ and } g_A = \frac{1}{2}\frac{T_3^f g_2}{cos\Theta_W}
\end{aligned}
\tag{3.13}
$$

where V is the vector boson, m_V is the mass of the boson, p_T is the maximum allowed transverse momentum, Q_f is the electromagnetic charge of the fermion f, T_3^f the weak isospin charge and Θ_W is the weak mixing angle.

In a more recent calculation the full SM PDFs were derived [69]. For a full SM evolution of the PDFs the evolution is divided in two regions by an energy scale $q_0 \sim m_V = 100$ GeV. For energy scales below q_0 a QCD and QED evolution of the PDFs is done. For energy scales above q_0 electroweak interactions emerge. Electroweak interactions are

chiral, hence the fermion PDFs are divided into left and right handed PDFs. Besides, electroweak bosons PDFs are divided into longitudinal and transverse PDFs. For energy scales above q_0 the SU(2)⊗U(1) symmetry is taken as unbroken theory, which makes the electroweak bosons massless. The real splitting functions for the electroweak evolution are defined as

$$\begin{aligned} P^R_{ff}(z) &= \frac{1+z^2}{1-z} \\ P^R_{Vf}(z) &= P_{ff}(1-z) \\ P^R_{fV}(z) &= \frac{1}{2}[z^2+(1-z)^2] \\ P^R_{VV}(z) &= 2\left[\frac{z}{1-z}+\frac{1-z}{z}+z(1-z)\right] \end{aligned} \tag{3.14}$$

where V represents spin 1 bosons and f the spin 1/2 fermions (taken from [69]). Note the similarity of with the QCD splitting functions in equation 2.32.

In figure 3.5 the full SM evolution of the up quark (left panel) and the gluon (right panel) PDFs are divided by the standard QCD evolution of the up quark and gluon respectively. The right handed up quark PDF has a similar behaviour as the QCD evolved PDF at low x. Right handed quarks only couple to Z bosons and photons besides gluons. At high x there is a deviation from the QCD evolved PDF. The deviation is caused by the valence quarks. Valence quarks are more likely to radiate an electroweak boson as they carry large momentum fractions. The left handed distribution is smaller that the right handed distribution. The left handed fermions can radiate $W^{\pm}$ bosons, which is a flavour changing interaction. The flavour change leads to a smaller distribution. The gluon PDF becomes smaller for increasing x. With the full SM evolution the probability of emitting and EW boson increases with increasing x.

In figure 3.6 we show the electroweak boson PDFs divided by the gluon PDF. The behaviour and the contribution of the four bosons is similar. At low x the contribution of the electroweak bosons is around 1% of the one from the gluon. At large x the contribution is similar, except for the photon which is 2% higher than the contribution of the gluon, as for the photon PDF the PDF derived in [70, 71] is used, which assumes an intrinsic content of photons in the proton.

A newer derivation of the electroweak boson PDFs was done in [72]. The authors propose the same method as in [70] for the derivation of the electroweak boson PDFs.

3.3.2 Electroweak resummation

We have explored the electroweak corrections in the hard-function and the PDFs. In section 2.4 we saw that calculations involving two different energy scales where one of the

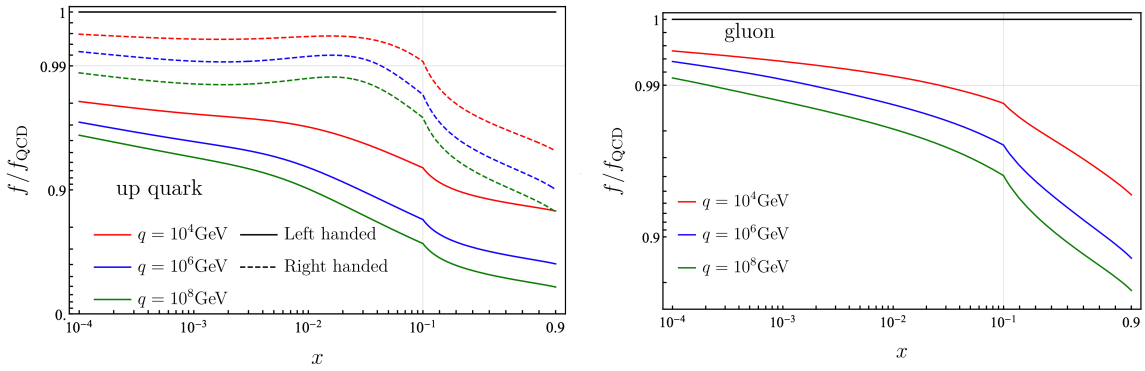

Figure 3.5: Up quark and gluon Standard Model PDFs divided by their corresponding QCD+QED PDF for left and right handed chiralities [69].

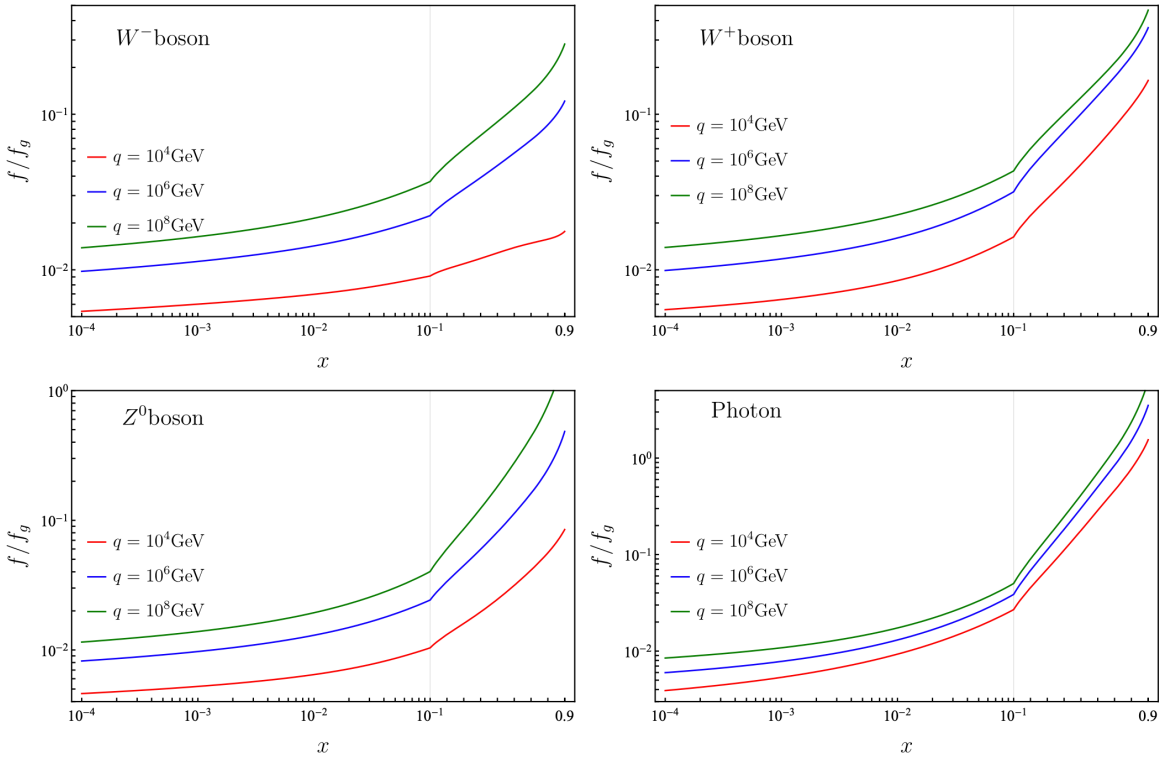

Figure 3.6: Electroweak boson PDF evolved with a SM model evolution divided by the QCD evolved gluon PDF for different energy scales [69].

scales is much smaller than the other one, resummation to all orders is needed for a correct description of data. Resummation is well understood in the context of QCD and QED. The collinear and soft radiation of massless particles produce single logarithms of the type $\ln(Q/\mu)$, where Q is the hard scale and μ is the scale of the radiation. The summation of these logarithms leads to the well-known Sudakov form factors. In collinear weak emissions apart from the single-logarithms, double-logarithms of the type $\ln^2(Q^2/M^2)$ arise, where M is the mass of the emitted electroweak boson. The resummation of electroweak emissions has to deal with the double and single logarithms. In [73] the resummation of the double logarithms is done at leading logarithm accuracy for the initial state for proton proton collisions with a lepton pair in the final state. In [74] the resummation of the double logarithms is done for any $2 \to 2$ process at next-to-leading logarithmic accuracy for both initial and final state.

We can also resum the electroweak logarithms numerically via parton showers. A first attempt to include electroweak parton showers was done in [75] for the Pythia8 parton shower. More recently a more complete electroweak parton shower was introduced in Vincia [76]. In [77] a comprehensive derivation of the electroweak splitting functions is performed, both for an unbroken SU(2)xU(1) symmetry, where the electroweak bosons are massless, and for a broken SU(2)xU(1) symmetry, where the electroweak bosons are massive. The splitting functions for an unbroken SU(2)xU(1) are those in equation 3.14.

3.4 Phenomenology of real electroweak corrections

In the previous sections we have introduced electroweak corrections and their different effects in proton-proton collisions for multijet processes. In this section we study real electroweak corrections at a centre of mass energy of 13 TeV in multijet events. These phenomenological studies will help us understand the measurement of real electroweak radiations in chapter 5.

To increase the probability of electroweak boson radiation we look at high transverse momentum multijet events. From equation 3.11 we know that the probability of an electroweak radiation is

$$\alpha_{EW} \times \ln^2 \frac{Q^2}{m_V^2}, \tag{3.15}$$

where α_{EW} is the coupling strength of electroweak interactions, Q is the energy scale of the hard interaction and m_V is the mass of the radiated electroweak bosons, with $V = Z, W^{\pm}$. Hence, to be sensitive to real electroweak radiations we define the phase space for multijet events as

1. At least two central jets with transverse momentum larger than 100 GeV. We define a central jet by requiring a maximum absolute rapidity of 2.5.

2. We only consider the muon decay channel of the Z boson for an easier reconstruction of the Z boson. We require two opposite charged muons with transverse momentum larger than 20 GeV, with an invariant mass of the muon pair in between 76 and 106 GeV.

3. We remove jets in a ΔR distance below 0.2 of any of the two muons, to avoid the misidentification of the muon as a jet (see appendix B).

In this phase space we study the following observables:

- The transverse momentum ratio of the Z boson and the leading jet. It is likely that one of the leading jets has radiated a Z boson. Hence, its transverse momentum is largely modified and the average transverse momentum of the two leading jets is not a good approximation for the scale of the hard interaction. For that, we use the transverse momentum of the leading jet as scale of the hard process. With the ratio we can compare the hardness of the emission of the Z boson to the hard scale.

- The difference in the ϕ angle between the two leading jets ($\Delta\phi_{j_1,j_2}$). In QCD interactions a pair of quark anti-quarks, a pair of gluons or a quark and a gluon are produced, hence two jets will be the hardest objects in the event. From momentum conservation of the whole event these two jets have to be in a back-to-back configuration, the ϕ angle between the two jets is 180°. However, a radiation modifies the ϕ angle of the parent particle so that $\Delta\phi_{j_1,j_2} < 180°$.

- The ϕ angle between the dijet system composed by the two leading jets and the Z boson ($\Delta\phi_{Z,dijet}$). We expect the two leading jets to be in a back-to-back configuration, as they are a product of the hard interaction. Then, the ϕ of the dijet system in an ideal dijet event should be zero. Hence, with any radiation the ϕ angle of the dijet system should be nearly zero. In a pure dijet event, the Z boson can be radiated from either of the jets or from any of the incoming partons, and thus, its ϕ angle should be "random", leading to a no correlation between the Z boson and the dijet system.

- The minimum ΔR distance between the Z boson and a jet ($\Delta R^{Min}_{Z,j}$). The Z boson can be radiated either in the initial or in the final state. The Z boson is radiated by one of the jets and by measuring ΔR we can infer the angle of the radiation.

To understand the radiation of electroweak bosons we compare two different ways to produce an electroweak boson in association with at least two jets with the PYTHIA8 event generator:

- **QCD+EWShower**: QCD events supplemented with a QCD and electroweak shower. QCD events are $2 \to 2$ processes where the two outgoing partons are quarks and gluons, which initiate two jets. We supplement these processes with the

PYTHIA8 simple electroweak shower [75], where we only allow Z boson radiation and we only consider the muon decay channel.

- **DY+1Jet**: DY+1 jet events are supplemented with the QCD shower. In this process the Z boson is calculated in the matrix element together with a jet. Any second jet in the event is product of a radiation in the parton shower. Hence, the second jet corresponds to the radiated Z boson in the QCD+EWShower scenario.

By comparing these two scenarios we aim to understand the differences between an electroweak boson radiation, QCD+EWShower scenario, and an electroweak boson generated as a product of the partonic interaction, the DY+1Jet scenario. We have studied this events with the RIVET analysis tool [78].

The transverse momentum ratio of the Z boson and the leading jet in figure 3.7 shows that Z bosons coming from a radiation (QCD+EWShower) are softer than the hard energy scale. However, we observe that we can find Z bosons with larger transverse momentum than the one of the leading jet. For the DY+1Jet, the Z boson is the hardest object in the event, where the transverse momentum of the Z boson is more than two times larger than the one of the leading jet. In figure 3.8 the $\Delta\phi$ angle between the two leading jets is shown. While for the QCD+EWShower the two leading jets are in a back-to-back configuration, the DY+1Jet scenario is not able to reproduce a back to back scenario. The $\Delta\phi_{Z,dijet}$ in figure 3.9 shows that in a DY+1jet scenario the Z boson is back-to-back with the dijet system, as the Z boson is recoiling against the radiation. In QCD+EWShower scenario the correlation in $\Delta\phi_{Z,dijet}$ is not as strong. We can observe from the $\Delta R^{Min}_{Z,j}$ in figure 3.10 that radiated Z bosons are more likely to be closer to a jet in comparison to a Z boson produced in a DY process.

From these distributions in figures 3.7-3.10 we learn that there is a difference in the kinematics of the event depending the production mode of the Z boson at leading order. Radiated Z bosons are soft in comparison to the scale of the hard scattering and they are likely to be found close to a jet. In addition, the two leading jets show a correlation in the $\Delta\phi$ observable. For Z bosons produced in the DY process, the Z boson is likely to be the hardest particle in the event and there will be no correlation between the leading and subleading jet in the $\Delta\phi$ distribution.[2].

As we have seen from the comparison in figures 3.7-3.10 a DY+1 jet calculation at leading order is not able to reproduce real electroweak radiations. In [75] the authors discuss that DY+1 jet events at next-to-leading order are able to describe real electroweak radiations. Thus, a DY+2 jets at leading order accuracy will be needed to describe real electroweak radiations. In the DY+2 jets processes together with the Z boson two jets are included in the calculation of the matrix element, which leads to diagrams where the Z boson is

[2]It is important to note that electroweak showers are not yet implemented in the simulations used by the experiments

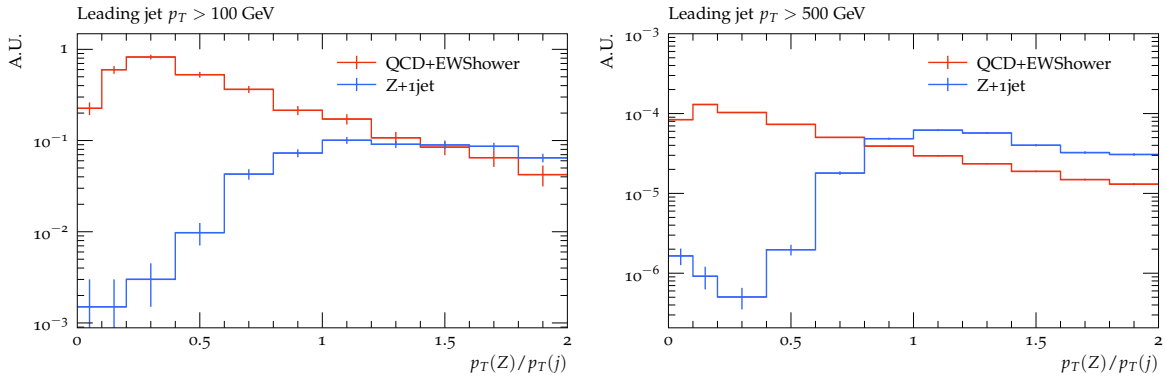

Figure 3.7: The transverse momentum ratio between the Z boson and the leading jet for transverse momentum of the leading jet above 100 GeV (left panel) and above 500 GeV (right panel).

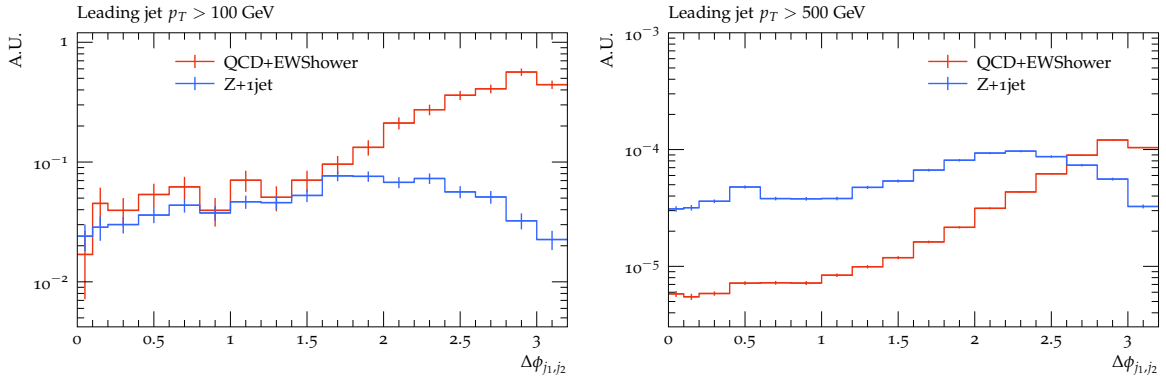

Figure 3.8: The $\Delta\phi$ between the two leading jets for transverse momentum of the leading jet above 100 GeV (left panel) and above 500 GeV (right panel).

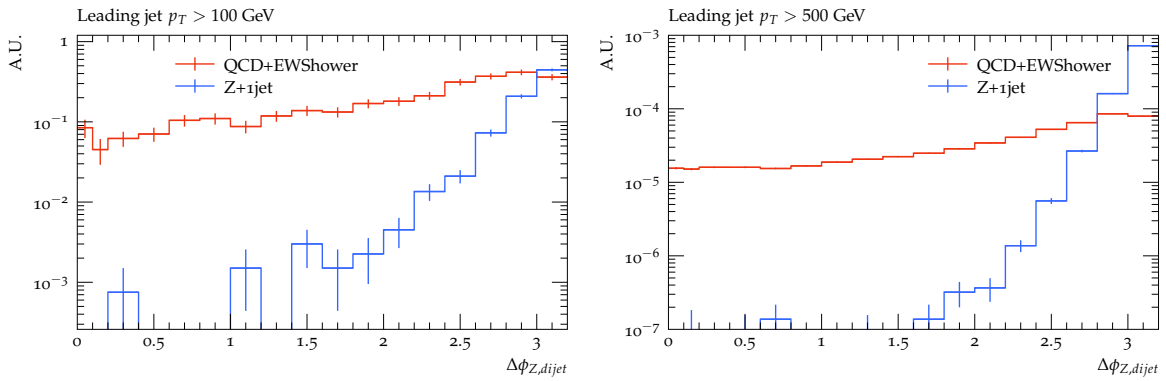

Figure 3.9: The $\Delta\phi$ between the two leading jets for transverse momentum of the leading jet above 100 GeV (left panel) and above 500 GeV (right panel).

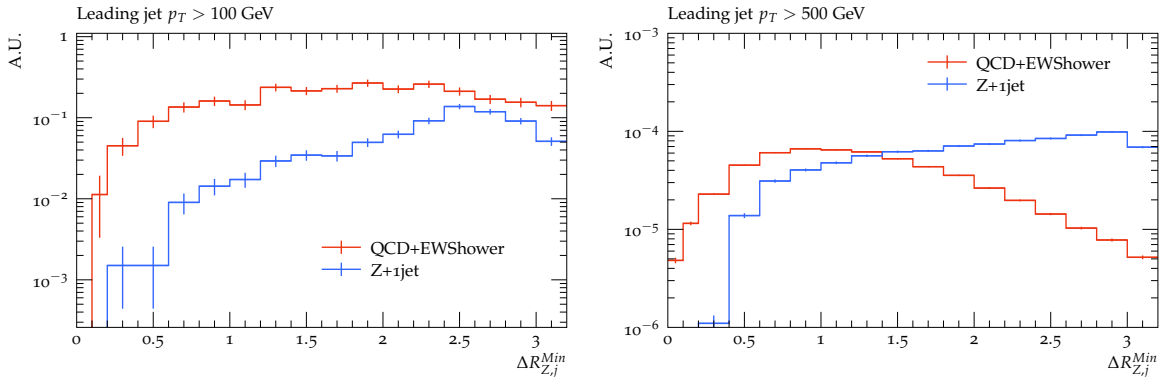

Figure 3.10: The minimum ΔR between the Z boson and a jet for transverse momentum of the leading jet above 100 GeV (left panel) and above 500 GeV (right panel).

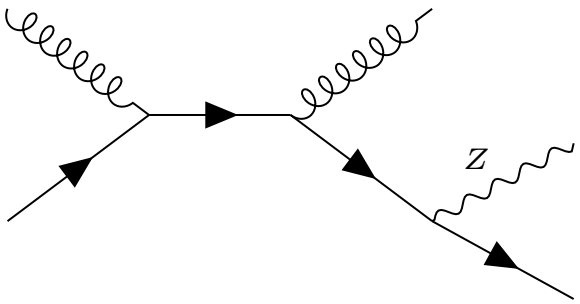

Figure 3.11: Leading order Drell-Yan + 2 jets Feynman diagram where the Z boson is a radiation of one of the outgoing partons.

radiated from one of the outgoing partons. In figure 3.11 we show Feynman diagrams for the DY+2jets where the Z boson is radiated from one of the outgoing quarks. Hence a DY+2 jets leading order calculation includes diagrams corresponding to real electroweak radiations. Moreover, with the phase space definition we are able to disentangle the different production channels of the Z boson.

CHAPTER 4

The Large Hadron Collider and the Compact Muon Solenoid

The proton-proton collision data used for the measurement of the real electroweak corrections presented in this thesis were collected by the Compact Muon Solenoid (CMS) experiment in 2016 at the Large Hadron Collider (LHC). In this chapter we present the experimental set-up for our measurement. We give a general overview of the LHC and the machinery to bring protons into collision. We describe the CMS detector, its different components and the particle reconstruction from energy deposits and tracks in the detector. We define jets, which are collimated sprays of hadrons product of the QCD bremshtralung of the outgoing partons. We explain the different reconstruction algorithms for the jet reconstruction. Further, we give an overview of the jet calibration, where during my time as a Ph.D. student I have contributed in the derivation of the corrections factors, by deriving the relative η dependent corrections for Run2 [79] and Run3 [80] of the LHC. Last, we present the common corrections applied for the physics analysis in the CMS experiment.

4.1 The Large Hadron Collider

The LHC is a circular hadron collider with a radius of 27 km located in Geneva, Switzerland. The LHC is the most powerful collider, which at the time of this thesis is running at a centre-of-mass energy of 13.6 TeV. Besides protons, it is able to collide heavy ions. For the following descriptions let us consider the collision of protons.

To reach such large energies the protons are accelerated in different steps, where in each step a different accelerator is involved. In figure 4.1 we show a sketch of the accelerator

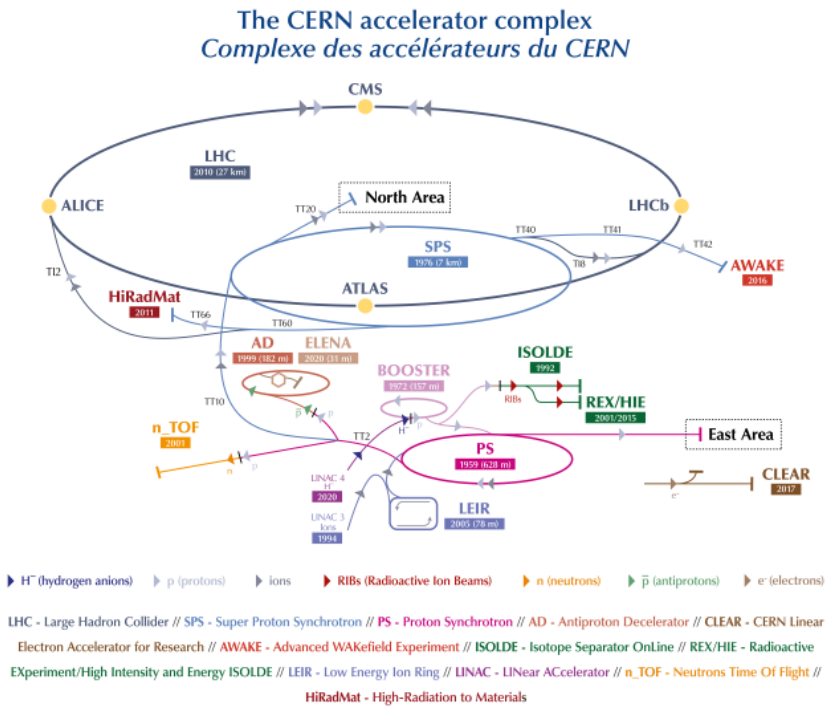

Figure 4.1: The accelerator complex at CERN. Figure taken from [81]

complex at CERN. In order to accelerate protons, the first step is to have protons. For that, hydrogen atoms with an extra electron (H^-) are used, as these atoms are formed by a proton and two electrons. Hence, we need to strip the electrons from the H^- atoms. The LINAC accelerator accelerates the H^- atoms to 160 MeV. These atoms are then injected in the Proton Synchrotron (PS). In the injection the electrons are stripped from the hydrogen atom with an electrical field. In the PS the protons reach 26 GeV and then they are injected into the Super Proton Synchrotron (SPS). The SPS is the last step before the protons are injected in two different directions into the LHC. The SPS accelerates the protons from 26 to 450 GeV. These two beams are accelerated to the corresponding energy in the LHC ring.

The beams are formed by bunches of protons. The characteristics of the beams are important as they affect directly the luminosity. The luminosity represents the number of collisions in a given time:

$$\mathscr{L} = \int \mathscr{L}(t)dt \tag{4.1}$$

where $\mathscr{L}$ and $\mathscr{L}(t)$ are the luminosity and the instantaneous luminosity, and t is time.

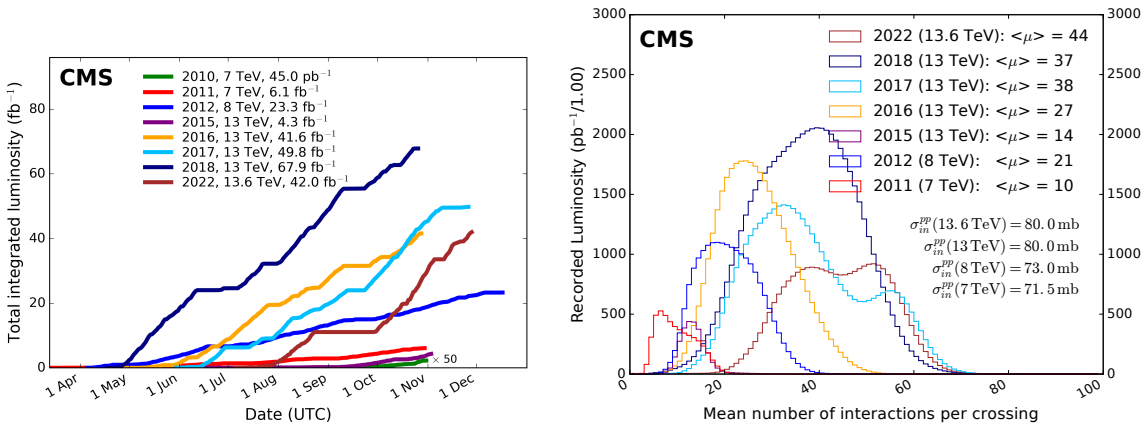

Figure 4.2: The delivered luminosity by the LHC in different years and different centre of mass energies (left). The recorded luminosity by the CMS experiment against the mean number of interactions per bunch crossing (right) [82].

The instantaneous luminosity is defined as

$$\mathscr{L}(t) = \frac{N_b n_1 n_2 f}{A_{eff}}, \tag{4.2}$$

where N_b is the number bunches in the beams, n_1 and n_2 are the number of protons in the respective bunch of the incoming beams and A_{eff} is the effective overlap between the two incoming beams. We can increase or decrease the luminosity by varying the characteristics of the beams. With the luminosity we can calculate the expected number of events that happen in a given time for a given process:

$$N = \sigma \mathscr{L} \tag{4.3}$$

where N is the number of events of the given process and σ is the cross-section of the given process. In figure 4.2 we show the delivered luminosity by the LHC in the different years and the mean of the number of interaction per bunch crossing. We can observe that the different configuration of the beams throughout the years changes the number of interactions per bunch crossing.

There are four interactions points at the LHC where the beams are able to collide. In each interaction point we find a different experiment: ALICE, ATLAS, CMS and LHCb. While ATLAS (A Toroidal LHC ApparatuS) and CMS are general purpose detectors, ALICE (A Large Ion Collider Experiment) and LHCb (LHC beauty) have specific goals. ALICE is designed for the study of matter state in the early universe. For that, they measure the collisions of heavy ions, such as, lead, which generate quark-gluon plasma. LHCb's goal is to understand the matter-antimatter asymmetry through bottom quarks.

4.2 The Compact Muon Solenoid

The Compact Muon Solenoid (CMS) is a multi-purpose experiment located in interaction point five of the LHC ring. It aims to detect and measure with high precision all charged particles. For this, the CMS detector was designed as a cylindrical detector wrapping the interaction point. It consists of different sub detectors designed for different purposes, in total four detectors constitute the CMS experiment: the tracker, the electromagnetic calorimeter, the hadronic calorimeter and the muon system. Besides these four detectors, a solenoid is placed in between the hadronic calorimeter and the muon system to bend the trajectories of charged particles. To enclose the magnetic field produced by the solenoid, iron yokes are placed around the solenoid. Together with all the different components and layers, the CMS experiment amounts to a total of 21 m long and 15 m high cylinder with a mass around 14 tons. Here we present the coordinate system of the CMS detector, the different sub-detectors that form the CMS experiment and the reconstruction of particles.

4.2.1 The CMS coordinate system

For the measurement of particles it is necessary to define a coordinate system of the detector. The z axis represents the direction along the beams. The x and y axes form the transverse plane to the beam direction. The x axis points to the centre of the LHC ring and the y axis is the vertical component. The polar coordinate system is composed by the θ and ϕ angles and the radius, r, in the xy plane. The ϕ angle measures the angle in the xy plane and θ in the zx plane.

In electron-positron colliders the energy of the incoming particles is known and measurements can be done in the Cartesian or the polar coordinate systems. However, for proton-proton collisions the energy of the incoming particles, the constituents of the protons, varies in each collision. Then, it is convenient to measure observables invariant under Lorentz boosts in the z axis, e.g.: observables that are invariant under different energies of the incoming partons. The polar angle θ is not invariant under Lorentz boosts in z axis. Thus, for measuring the momentum in the beam axis the rapidity relative to the beam axis,

$$y = \ln \frac{E + p_z}{E - p_z}, \tag{4.4}$$

where y is the rapidity, E is the energy of the particle and p_z is the momentum of the particle in the beam axis, is a better observable. While the rapidity is not invariant under boosts in the z axis, the difference in rapidity between two particles is invariant. The rapidity can be understood as boost from the detector frame to a frame where the particle is only moving in the transverse plane. For massless particles or in the high energy regime where the mass of the particle can be neglected, the rapidity converges to

the pseudorapidity:

$$\eta = -\ln \tan \frac{\theta}{2}. \tag{4.5}$$

which directly relates to the θ angle. Now that the longitudinal components (θ, p_z) of the particles are embedded in the pseudorapidty, the transverse components can be represented by the transverse momentum of the particle,

$$p_T = \sqrt{p_x^2 + p_y^2} \tag{4.6}$$

where p_x and p_y are the momentum in the x and y axes respectively. The transverse momentum is invariant under boosts along the z axis. The transverse mass is also invariant under boosts along the z axis, and it is defined as

$$m_T = \sqrt{m^2 + p_T^2} \tag{4.7}$$

where m is the invariant mass of the particle. Another important quantity is the missing transverse energy, $\not{E}_T$, as neutrinos or other weakly interacting particles cannot be measured, where $\not{E}_T = \sum_i \vec{p}_{T,i}$ and we sum over the transverse momentum of all the measured particles in the collision. We can infer the missing transverse energy by extracting the total measured energy from the expected energy.

4.2.2 The CMS detector

Due to the high rate of collisions, the detector is exposed to a large amount of radiation. Moreover, the refresh rate of the detector has to be large to record the high rate of collisions. Hence, we need materials resistant to radiation with a high refresh rate. As mentioned before the CMS detector is divided in four sub-detectors, where each detector has different goals. In this subsection we present the different sub-detectors and their basic working principle. We start from the inner most detector, which is closest to the beam pipe, and finish with the outer most detector, the muon system.

The tracking system

The tracker is the closest detector to the interaction point. Its goal is to measure the trajectory of charged particles. The tracker is immersed in the magnetic field of the solenoid. This will make the trajectory of the charged particles to bend. From the curvature of the trajectory we can infer the momentum and charge of charged particles crossing the tracker system.

To measure the trajectory of the charged particles two detectors form the tracker system: the pixel and the strip detectors. Both detectors are based on Silicon, which offers a fast recovery time. During the 2016 data taking period the pixel detector consisted of three

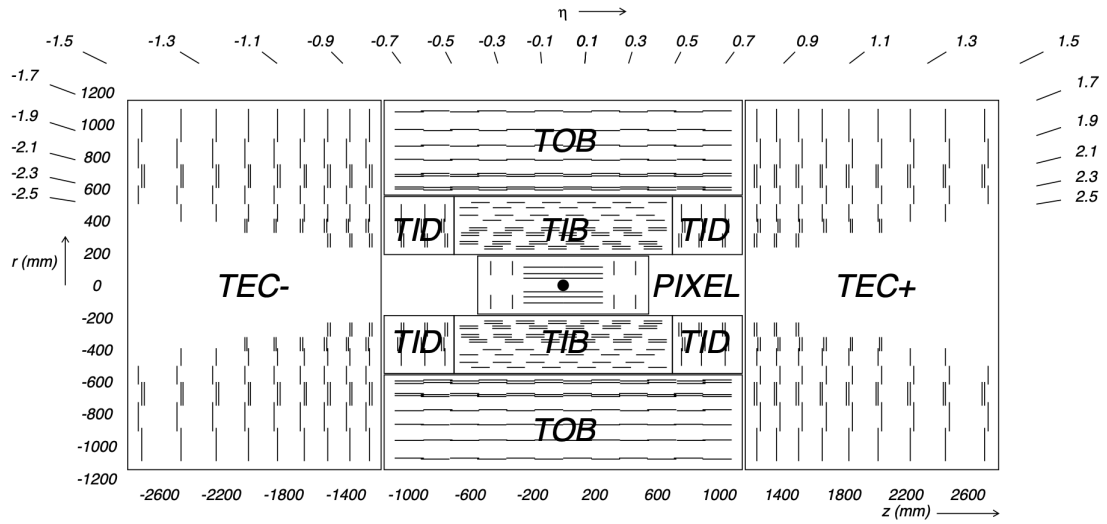

Figure 4.3: Cross-section of the CMS tracking system in the zy plane [84].

cylindrical layers of radius of 44, 73 and 102 mm wrapping the beam pipe. The cylinders cover the -270 mm to 270 mm region of the beam axis around the interaction point. Two end-cap discs are placed on each side of the cylinders. The whole pixel detector consists of 66 million pixels of $100 \times 150 \mu\text{m}^2$. It offers a resolution in $r - \phi$ of 10 μm and 20 μm in z. The pixel detector was upgraded during the end of the year shut-down in 2016/2017 to meet the luminosity requirements of the increase of energy from Run1 to Run2 [83].

The strip detector is divided in four parts: the inner barrel (TIB), the inner disk (TID), the outer barrel (TOB) and the outer end-caps (TEC). The inner strip structure, the TIB and TID, consist of four layers and three disks on each side. The outer strip, the TOB and TEC, is conformed by six layers and nine disks on each side. In total 15148 modules, where each module consists of one or two silicon sensors, form the strip detector. The size of the sensors vary with the distance to the interaction point. Overall, the strip detector offers a 20-50 μm resolution in $r - \phi$ and 500 μm in z. The pixel and the strip detector together cover the $|\eta| < 2.5$ region. For a schematic view of the tracker in figure 4.3 we show a scheme of the zy plane of the tracking system.

The electromagnetic calorimeter

The goal of the electromagnetic calorimeter (ECAL) is to measure the energy of photons and electrons. The ECAL is a fast detector and offers a high granularity. To measure the energy, the ECAL is build with lead tungsten crystals, a highly transparent material with scintillation properties. A photon or a electron passing through a crystal will interact with the atoms of the crystal generating an electromagnetic shower. The shower will make the

crystals to scintillate. The photons produced in the scintillation process pass trough a photomultiplier located in the back of the crystal, generating an electric signal.

The ECAL is divided in two regions: the barrel region (EB) and the end-caps (EE), with around 15000 crystals. In the EB the cross-section of the crystal reaches $\Delta\eta \times \Delta\phi = 0.0174 \times 0.0174$ and of 230 mm in r. At the EE the cross-section of the crystals is of 2.68×2.68 cm^2 with a radial length of 220 mm.

The spatial resolution of the ECAL in the EE is not fine enough. Low momentum photon pairs produced as a decay of pions will be detected as single photons. To increase the spatial resolution and resolve the pair of photons the pre-shower (ES) detector is placed between the tracker system and the EE. The ES is composed by a lead absorber and a silicon strip.

The ECAL covers the $|\eta| < 3.0$ region. More precisely, the $|\eta| < 1.48$ region is covered by the EB and $|\eta| < 3.0$ by the EE. The pre-shower detector covers the $1.65 < |\eta| < 2.6$. In figure 4.4 we show the a sketch of the ECAL with the corresponding η regions and sub-detectors.

The relative energy resolution of the ECAL is given by

$$\frac{\sigma}{E} = \frac{2.8\%}{\sqrt{E[\text{GeV}]}} \oplus \frac{12\%}{E[\text{GeV}]} \oplus 0.3\%, \tag{4.8}$$

where the first fraction corresponds to the stochastic term, the second to the noise in the ECAL, which is composed by electronic noise and by event pileup. The event pileup refers to the overlap of the signal of previous collisions. The third term refers to non-linearities in the ECAL response [85].

The hadronic calorimeter

Photons and electrons deposit most of their energy in the ECAL. However, the hadrons, due to a larger mass, have less ionisation power and cross the ECAL without a large deposition of energy. To measure the energy of hadrons an hadronic calorimeter (HCAL) is placed after the ECAL.

Due to the limited space between the ECAL and the solenoid, the barrel region of the HCAL is divided in two regions: The inner barrel (HB), situated between the ECAL and the solenoid, and the outer barrel (HO), situated outside the solenoid. The HB is flanked by end-caps (HE) and in the forward direction ($|\eta| > 2.5$) forward calorimeters (HF) are placed. The HCAL is an hermetic detector as it covers the $|\eta| < 5.2$ range, this is important to infer the $\not{E}_T$. For hadrons not to leave the HCAL, as they do in the ECAL, plastic scintillating layers are interleaved with brass plates and enclosed by two steel layers. The brass and steel layers act as stoppers, the hadrons will interact with

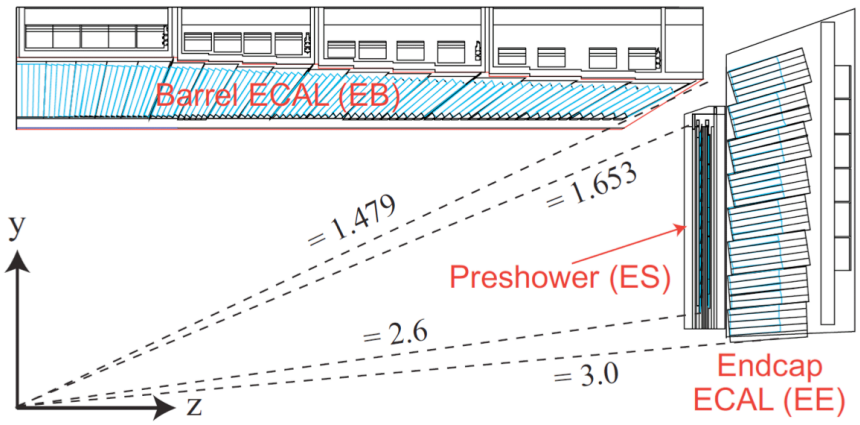

Figure 4.4: Sketch of the geometrical view of a quarter of the electromagnetic calorimeter [86].

them generating a shower, which is measured by the scintillating material. In figure 4.5 we show a scheme of a quadrant of the HCAL.

The different regions in the HCAL offer different cross-sections of the segmentation of the calorimeter towers in $\eta \times \phi$. For the central region, where $|\eta| < 1.6$, a cross-section of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ and in the forward region, where $|\eta| > 1.6$, a cross-section of $\Delta\eta \times \Delta\phi = 0.017 \times 0.017$ is used.

The muon system

The muon system is the outermost detector of the CMS detector. Muons are minimum ionising particles, they are able to leave the ECAL and HCAL with a minimum loss of energy. To measure the energy of the muons gas detectors are used. The muon system is interleaved with the iron yokes and it is composed by three different gas detectors. The working principle of gas detectors is ionisation. A muon crossing a gas detector will ionise the gas atoms creating an avalanche of electrons and ions. In the gas detector an anode and a cathode are placed, where electrons will be absorbed by the anode and the ions by a cathode generating an electric signal.

For the design of the muon system the muon flux and the magnetic field of the solenoid are taken into account. In the central region, $|\eta| < 1.2$, the muon flux and the magnetic fields are small. For $|\eta| > 1.2$ the muon flux is larger and the magnetic field is not homogeneous. Hence, in the forward region a better space and time resolution is needed. Four layers of drift tubes (DTs) cover the $|\eta| < 1.2$ region. For an increase in time and space resolution

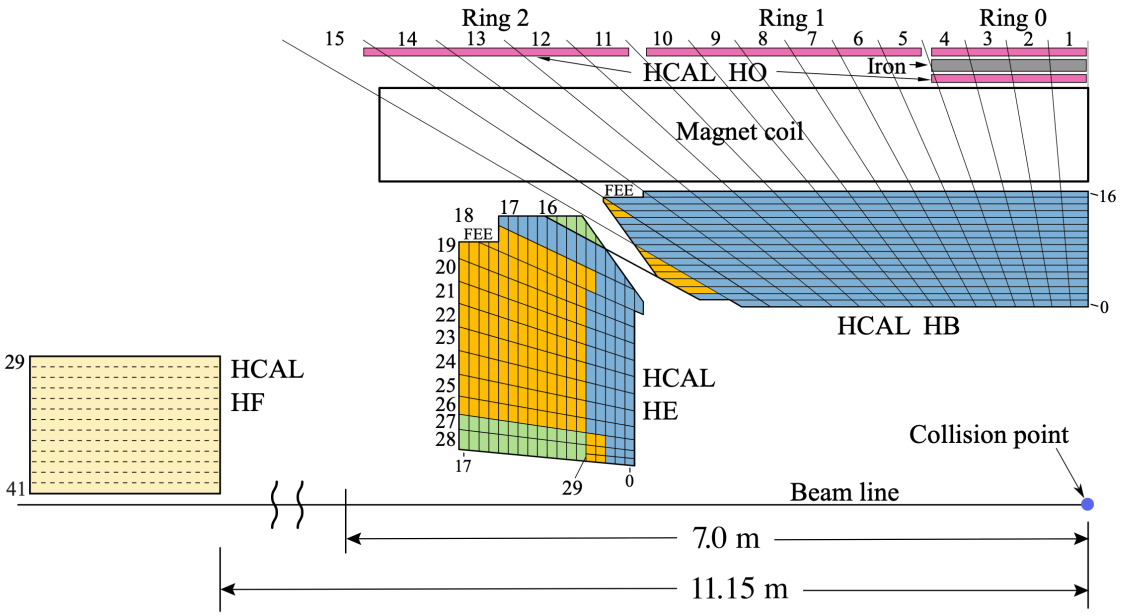

Figure 4.5: A quadrant of the hadronic calorimeter [87].

cathode strip chambers (CSC) are used in the $0.9 < |\eta| < 2.4$ region. Both the CSCs and the DTs are supported by resistive plate chambers (RPCs) in the $|\eta| < 1.9$ region. The RPCs have a fast response, smaller than the bunch crossing frequency. This allows the RPCs to time-resolve muons produced in the same bunch crossing and thus are used as part of the trigger system of the CMS detector. In figure 4.6 we show quadrant of the muon system in the zy plane, where the different gas detectors and the η coverage is shown.

The trigger system

With the high frequency of the LHC, a bunch crossing every 25 ns, the amount of data produced cannot be stored. With an implementation of a trigger system we are able to decide which collisions produce an interesting event worth storing. The trigger system has to be relatively fast to be able to read the following collision. The trigger system is composed by two levels, level-1 (L1) and high level triggers (HLT).

The L1 trigger is a hardware system and decides if the collision is further analysed by the HLTs. The L1 uses information of the muon system and both calorimeters to reconstruct leptons and jets. In addition, the L1 trigger is able to compute the missing transverse energy. To the objects reconstructed by the L1 minimum requirements are applied to decide for a collision to be further analysed.

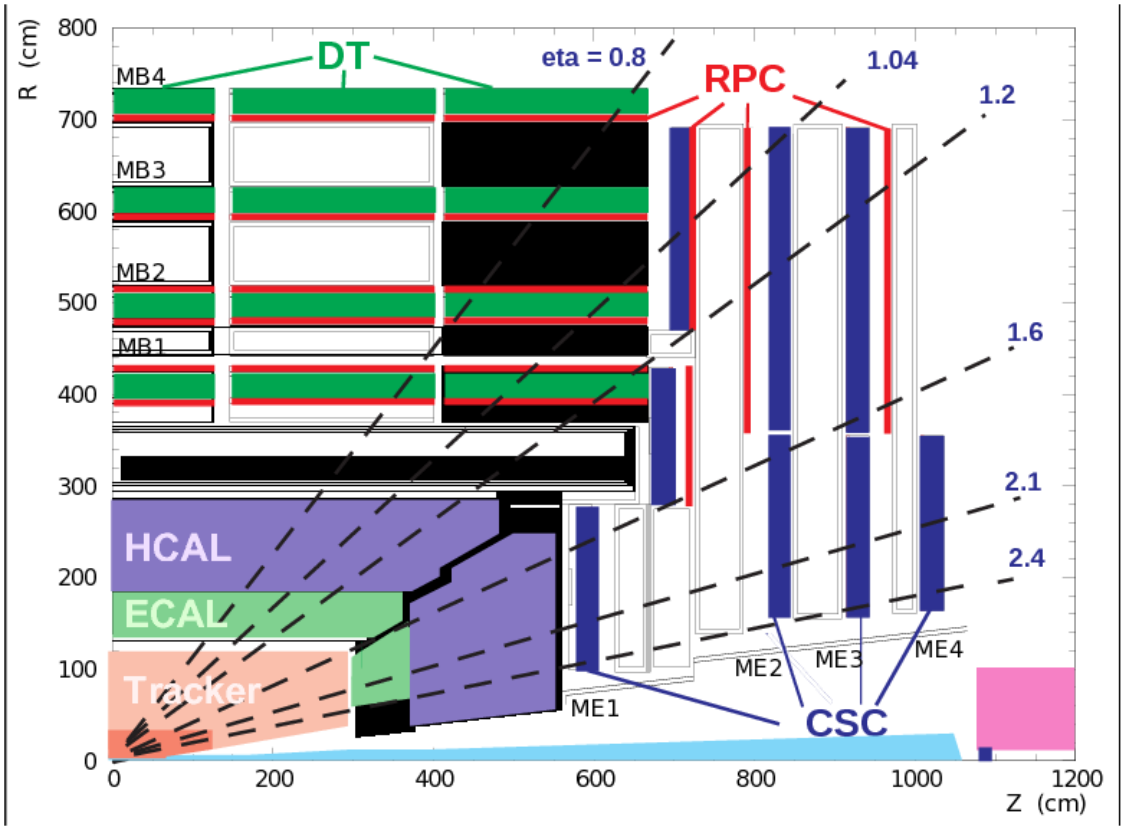

Figure 4.6: Quadrant of the Muon system, drift tubes (DT), cathode strip chambers (CSC) and resistive plate chambers (RPC) [88]

The HLT system is a more complex triggers system. It uses information from all the sub-detectors to reconstruct different particles. There are different HLTs for the different interesting particles produced in the collision. Hence, a HLT of a given particle will be fired when the HLT reconstruction of the particle fulfils some minimum requirements. Collisions that fire any HLT trigger are then stored. Moreover, the event will be stored in a dataset corresponding to the HLT trigger that has fired. A collision can fire different HTLs and the same collision is saved in different datasets.

Simulation of the CMS detector

To compare the measured data with simulations in the same footing, we need to simulate the detector response. The simulation of CMS detector is performed with the GEANT4 software [89–91]. GEANT4 simulates the geometry and the interactions of the particles with the different materials of the detector.

The final state particles of the simulated events are fed to GEANT4. GEANT4 will simulate the tracks and energy deposits that such final state will leave in the detector. Then, particles are reconstructed with the same procedure as in the measurement. This generates two levels of final state in the simulation: the *particle level*, which corresponds to the final state before the detector simulation, and the *detector level*, which corresponds to the particle level final state that has undergone the simulation of the CMS detector. These two levels are important later in the measurement procedure for a first comparison between measured data and simulation and the latter correction to particle level.

4.3 Event reconstruction

In this section we discuss how particles are reconstructed from the information collected by the CMS detector. The Particle-Flow (PF) algorithm [92] collects the information of the different sub-detectors and combines them into PF candidates, which are the reconstructed particles. In a first step the PF algorithm creates the PF elements: tracks, vertices, muons, electrons and hadrons.

To reconstruct the tracks of charged particles information of the tracker and the muon chambers is used. As a charged particle crosses the layers of the tracker it will leave a *hit* in the different layers. As the tracker is submerged in a magnetic field the trajectories of charged particles are bent. To reconstruct the tracks in the tracking detector the *hits* in each layer are fitted to tracks, for the *hits* in the muon chamber the same procedure is used. From the curvature of the tracks the momentum and charge of the charged particle is computed.

From the tracks the vertices are reconstructed. A vertex is the location where protons have collided. They are infered by clustering adjacent reconstructed tracks to their closest

point in the beam axis. The sum of the transverse momentum of the tracks defines the energy of the vertex. The most energetic vertex is defined as the primary vertex. Other vertices in the event are then considered as coming from pileup collisions.

Then, muons are reconstructed, and the PF algorithm uses the reconstructed tracks both from the tracker and muon chamber. First *standalone muons* are reconstructed, which correspond to tracks in the muon chamber. If a track from the tracker matches the *standalone muon*, the *standalone muon* becomes a *global muon*. Then, *tracker muons* are reconstructed, which are tracks from the tracker with a corresponding hit in the muon chamber. For the analysis in chapter 5 we make use of *global* and *tracker muons*, as *standalone muons* can arise from cosmic rays crossing the detector. The tracks used for the muon reconstruction are removed from the work-flow of the PF algorithm.

The QED shower generated by electrons and photons creates many close-by clusters in the ECAL. These are grouped in larger clusters, called superclusters. If a supercluster has a matching track in the tracker it is reconstructed as an electron. When no track is compatible with the supercluster it is reconstructed as a photon.

With the remaining tracks the hadron reconstruction is performed. A charged hadron is reconstructed when a cluster in the HCAL has a matching track in the tracker. HCAL clusters with no matching track are then reconstructed as neutral hadrons.

4.3.1 Jets

Quarks and gluons produced in the partonic interaction evolve producing a collimated spray quarks and gluons that later hadronise. We refer to the initial parton producing the collimated spray as the parent parton and to the collimated spray of hadrons as a jet. To cluster the measured hadrons into jets different jet algorithms exist. Depending on the goal of the measurement a jet clustering algorithm will be favoured. One of the most important aspects is that the jet clustering algorithm is infra-red and collinear safe. Jet clustering algorithms that are not infra-red and collinear safe, spoil the perturbative calculation of theoretical predictions, and the comparison to data becomes difficult. For this reason sequential clustering algorithms are favoured nowadays in experimental measurements. These algorithms suppose that particles arising from the radiation of the same parent parton are close to each other. Thus, the definition of the distance between particles is important. We denote the distance between particle i and j as d_{ij}, and the distance between the particle i and the beam axis as d_{iB}. The definition of d_{ij} and d_{iB} differs between the different algorithms. Two particles are clustered when $d_{ij} < d_{iB}$ to form a particle k. Then, all particles with $d_{ik} < d_{iB}$ are used to reconstruct the jet. Once $d_{ik} > d_{iB}$ the jet is defined and the reconstruction of other jets starts. Three main sequential clustering algorithms exist, the Cambridge-Aachen [93], the k_T [94] and the anti-k_T [95] algorithms:

- The Cambridge-Aachen (CA) algorithm defines the distance parameters as

$$\begin{aligned} d_{ij} &= \frac{\Delta R_{ij}}{R} \text{ with } \Delta R_{ij} = \sqrt{\Delta\phi_{ij}^2 + \Delta\eta_{ij}^2} \\ d_{iB} &= 1 \end{aligned} \tag{4.9}$$

  where R is jet radius parameter defined by the user. The CA algorithm does a spatial clustering, favouring the clustering of close-by particles. Jets reconstructed with the CA algorithm are susceptible to particles originating from pileup events. Thus, making it difficult to relate the energy and momentum of the jet to the parent parton. However, it is the preferred method to study the jet substructure.

- The k_T-algorithm defines the distance parameters as

$$\begin{aligned} d_{ij} &= \min\left(k_{T,i}^2, k_{T,j}^2\right) \frac{\Delta R_{ij}}{R} \text{ with } \Delta R_{ij} = \sqrt{\Delta\phi_{ij}^2 + \Delta\eta_{ij}^2}, \\ d_{iB} &= k_{T,i}^2, \end{aligned} \tag{4.10}$$

  where $k_{T,i}$ is the transverse momentum of the particle i. This algorithm favours the clustering of soft particles, making the reconstructed jets prone to include particles coming from pileup events.

- The anti-k_T algorithm can be seen as an improvement of cone reconstruction algorithms. Cone algorithms assume that jets are confined in conical shapes. However, cone algorithms are not infrared and collinear safe. The anti-k_T algorithm defines the distance parameters as

$$\begin{aligned} d_{ij} &= \min\left(\frac{1}{k_{T,i}^2}, \frac{1}{k_{T,j}^2}\right) \frac{\Delta R_{ij}}{R} \text{ with } \Delta R_{ij} = \sqrt{\Delta\phi_{ij}^2 + \Delta\eta_{ij}^2}, \\ d_{iB} &= \frac{1}{k_{T,i}^2}, \end{aligned} \tag{4.11}$$

  In comparison to the k_T-algorithm, it favours the clustering of large transverse momentum particles. In this way, reconstructed jets are not as exposed to particles arising from pileup events. For these characteristics the anti-k_T algorithm is preferred in hadronic collisions, as the reconstructed jets are closer to the energy and transverse momentum of their corresponding parent, and it is suitable for theoretical comparisons.

In the CMS experiment jets are reconstructed with the anti-k_T algorithm, where all particles reconstructed by the PF algorithm are used for clustering. We can further remove pileup contamination from the jets with the charged hadron substraction (CHS) method. The CHS method removes the charged particles that do not originate from the primary

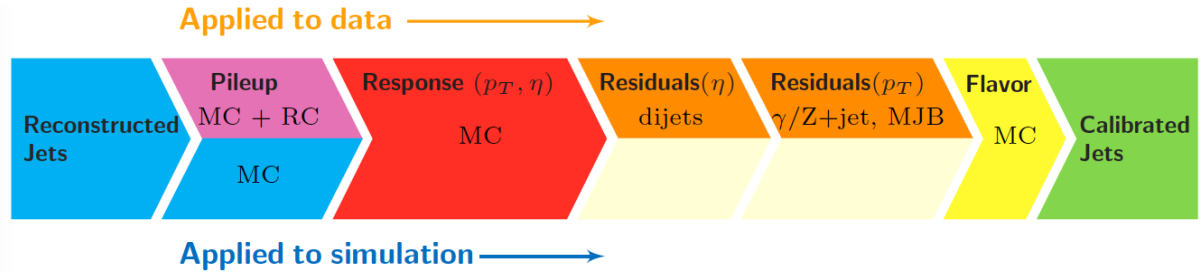

Figure 4.7: Ilustration of JEC derivation [98]

vertex. The CHS method was the preferred method for pileup removal for the Run2 measurements, while for Run3 the *Puppi* (Pileup Per Particle Identification) method [96] is favoured, as with the increased pileup events in Run3, the *Puppi* method performs better [97]. Another important aspect of the jet reconstruction is the jet radius parameter R. With a large value of R soft radiation will be included in the jet reconstruction, which allows for a better estimation of the mass and content of the jet. However, it makes the jet prone to include particles from pileup events. Hence, a smaller value of R will reduce the amount of pileup particles in the jet. The CMS experiment by default reconstructs jets with $R = 0.4$ and $R = 0.8$. For our analysis in chapter 5 we make use of reconstructed jets with $R = 0.4$, as we are not interested in the studies of the jet substructure, with the CHS method for pileup cleaning. In CMS jargon these jets are known as AK4PFchs.

Jet energy calibration

The reconstructed jets at the CMS experiment are clusters of many different particles reconstructed from different sub-detectors. Thus, they are susceptible to the response and resolution of detector, which makes the measured jet energy differ from its true energy. We can calibrate the measured jet energy to its real value with the jet energy corrections.

To derive the jet energy corrections a factorisation approach is used by the CMS experiment. The factorisation helps to simplify the task and derive correction factors for different sources. The derivation is organised in consecutive steps, as illustrated in figure 4.7. These corrections are applied as a factor, C_{JEC}, to the momentum of the reconstructed jet, p_{Raw}, to recover the true momentum of the jet, p_{True},

$$p_{\text{True}} = C_{JEC} \times p_{Raw} \tag{4.12}$$

In this section we use $[A]$ as the bins of a given observable A, $\langle A \rangle$ the average of the observable A and (A) the dependence of a function on the observable A.

First, the pileup offset corrections are derived. Here we differentiate between two different sources of pileup events, multiple proton-proton interactions within a bunch crossing and

consecutive bunch crossings can leave a signal in the same detector region. These pileup events will induce an incorrect measurement of the jet energy. The two sources of pileup modify the measured momentum of the jet. Even though with the CHS we have removed most of the pileup contamination, there are still remnants. To derive the pileup offset correction, vertices are not used, as they might be affected by tracking inefficiencies. Instead the energy density, ρ, is used, which is the median of the energy density calculated in $\eta - \phi$ grids.

To calculate the offset produced by pileup, two samples of simulated events are used. The origin of the events in the two samples are the same, but one of the samples includes the simulation of pileup events. Another important variables is μ, the number of pileup interactions in a bunch crossing, which is known in simulated events.

The offset of the transverse momentum is calculated as

$$\langle p_{T,part,offset} \rangle \left(\langle \rho \rangle, [\eta], p_{T,Raw} \rangle \right) = \langle p_{T,Raw}^{\text{No pileup}} - p_{T,Raw}^{\text{With pileup}} \rangle \left[\mu, \eta, p_{T,part} \right] \tag{4.13}$$

where $p_{T,part}$ and $p_{T,Raw}$ are the transverse momentum of the jet at particle level and the reconstructed jet (after detector simulation) respectively. For this measurement only reconstructed jets with a matching jet at particle level are considered. The matching criteria used is $\Delta R_{j_{part},j_{reco}} < R/2$ where j_{part} and j_{reco} are the jets at particle level and detector level. The transverse momentum offset cannot be applied to data as μ and $p_{T,part}$ are not available in real data. The correction factor for pileup-offset is then mapped to ρ, $p_{T,Raw}$, A and η, where A is the area of the jet:

$$C_{\text{hybrid}}(p_{T,}, \eta, A, \rho) = 1 - \frac{\langle p_{T,part,offset} \rangle}{p_{T,Raw}}. \tag{4.14}$$

The next derived corrections are the simulated response corrections which it is also based on simulation. It is derived after applying the pileup offset correction to the jets. The response of the detector simulation is the defined as

$$R_{part} \left(\langle p_T \rangle, \eta \right) = \frac{\langle p_T \rangle}{p_{T,part}} \left[p_{T,part}, \eta \right] \tag{4.15}$$

where, $p_T = C_{\text{hybrid}} \times p_{T,raw}$. Then, the momentum of the jet is corrected as $R_{part} \times p_T$. At this stage the jets in the simulated samples are considered as fully calibrated. However, the simulation of the detector response in GEANT4 is not time dependent. In reality the radiation damage worsens the detector response with time. To remove this difference between simulation and measured data, residual corrections are derived for data, using both measured data and simulation.

Different event topologies are used to derive residual corrections to cover the full phase space in η and transverse momentum p_T. At this stage both the jets from simulated events

and measured data are corrected with the simulation based corrections. In all the different topologies the same procedure is used to derive the response of the jets, a reference and a probe objects is defined, where the probe object corresponds to the jet we measure the response from. Two main methods are used to study the jet response, the p_T-balance and the missing transverse momentum fraction (MPF) methods. The response by these methods is defined as

$$R_{jet}^{p_T\text{-balance}} = \frac{p_{T,jet}}{p_{T,ref}}, \qquad R_{jet}^{\text{MPF}} = 1 + \frac{\vec{p}_{T,miss} \cdot \vec{p}_{T,jet}}{p_{T,ref}^2} \tag{4.16}$$

where $p_{T,ref}$, $p_{T,jet}$ and $p_{T,miss}$ are the transverse momenta of the reference object, the probe jet and the missing transverse momentum. The p_T-balance method only uses the information of the probe and reference objects, while the MPF method uses information from the whole event, as $p_{T,miss}$ is the sum of all particles in the final state. Both definitions of the jet balance are sensitive to radiation and besides the response we measure the extra radiation. To only consider the jet response we measure the response in terms of α, which measures the activity of the extra radiation. Depending on the topology, α is differently defined. Then, to measure purely the jet response, α is extrapolated to a value of 0. Two types of residual corrections are derived:

- **Relative η dependent corrections**
  To derive relative η corrections the measurements of the response are done in p_T and η, and hence a process with large event sample is needed. For that, dijet event topologies are used due to their large cross-section in hadronic collisions. To define the reference and probe object, the two leading jets are used. At least one of the two leading jets must have $|\eta| < 1.3$ to be considered in the event. The jet with $|\eta| < 1.3$ will be defined as the reference object and the other leading jet as the probe. When the two jets have $|\eta| < 1.3$ the reference and probe objects will be randomly assigned. The extra activity in the dijet events is defined as

  $$\alpha = \frac{p_{T,3rdjet}}{p_{T,ref} + p_{T,probe}} \tag{4.17}$$

  Due to the bad resolution of the reference jet, the measurement of the response is performed by an average of the transverse momenta in the dijet system: $p_{T,ave} = 0.5(p_{T,ref} + p_{T,probe})$. Hence the response in the dijet system is measured as

  $$\begin{aligned} R_{jet}^{p_T\text{-balance}} &= \frac{1 - \langle A \rangle}{1 + \langle A \rangle} \text{ with } A = \frac{p_{T,probe} - p_{T,ref}}{p_{T,probe} - p_{T,ref}} \\ R_{jet}^{\text{MPF}} &= \frac{1 - \langle B \rangle}{1 + \langle B \rangle} \text{ with } B = 1 + \frac{\vec{p}_{T,miss} \cdot \vec{p}_{T,jet}}{p_{T,ref}^2} \end{aligned} \tag{4.18}$$

  and when $\alpha \to 0$, these definitions converge to the ones in equation 4.16. The relative η dependent corrections are known, in CMS jargon, as Level 2 residuals

(L2Res), and the correction factor is defined as

$$C_{L2Res} = R_{jet}^{\text{Simulation}} / R_{jet}^{\text{Data}} \tag{4.19}$$

- **Absolute corrections**
  Absolute corrections are p_T dependent corrections, and in order to cover the whole p_T range, from 30 GeV to 1 TeV, different event topologies are used. To determine the response of jets at low transverse momentum we need a precisely reconstructed object, as the reconstruction of low transverse momentum jets is not precise enough. The Z boson can be precisely reconstructed through its leptonic decay. Hence, in the Z+1 jet event topology we can measure the jet response through the Z boson. The Z boson is defined as the tag object to probe the response of the jet. To cover the central transverse momentum region Z+ 1 jet and γ + 1 jet event topologies are combined to increase the number of events. However, at large jet transverse momentum the amount of Z/γ events is too small to measure the jet response, and multijet events are used. The extra radiation acitivity for Z/γ + 1 jet events is defined as

$$\alpha = \frac{p_{T,jet}}{p_{T,Z/\gamma}} \tag{4.20}$$

  Absolute corrections are also known as Level 3 residuals (L3Res). From each event topology the correction factor $C_{L3Res}^{(i)} = R^{i}_{\text{simulation}} / R^{(i)}_{\text{Data}}$ is derived, as a function of the transverse momentum of the probe jet, where i refers to a different event topology. Then, the different $C_{L3Res^{(\text{i})}}$ are fitted, resulting in an absolute correction C_{L3Res}.

Once all the corrections have been derived sequentially we can define the jet energy calibration correction factor for simulations and real data as:

$$\begin{aligned} C_{JEC}^{\text{Simulation}} &= C_{Offset} \times C_{\text{hybrid}} \\ C_{JEC}^{\text{Data}} &= C_{Offset} \times C_{\text{hybrid}} \times C_{L2Res} \times C_{L3Res} \end{aligned} \tag{4.21}$$

Jet energy resolution

After the reconstructed jets are calibrated both in data and simulations, it turns out the transverse momentum resolution of the jets is better in simulations than in data. To reduce this disagreement the resolution of the simulated jets is smeared. In simulations we can define the resolution of a jet as

$$\Delta = \frac{p_{T,rec} - p_{T,part}}{p_{T,part}} \tag{4.22}$$

where $p_{T,rec}$ and $p_{T,part}$ are the transverse momenta of the jet at detector and particle level. The distribution of Δ follows a Gaussian like distribution and it is fitted to a Gaussian

distribution of a width σ_{JER}. The width of the fitted Gaussian distribution is also known as the resolution. Then, we can relate the $p_{T,rec}$ with $p_{T,part}$ as

$$p_{T,rec} = p_{T,gen} \times (1 + \Delta). \tag{4.23}$$

The transverse momentum of the reconstructed jet in the simulation is smeared to match the resolution of the reconstructed jets in data, by applying the following correction

$$p_{T,rec} = p_{T,gen} \times (1 + \Delta_{\text{Data/Sim}} \times \Delta), \tag{4.24}$$

where $\Delta_{\text{Data/Sim}}$ is the data-to-simulation core resolution scale factor, which is derived with data base methods.

However, in simulation not all reconstructed jets at detector level have a corresponding jet at particle level, due to the detector response. Thus, depending on the availability of a matching jet two methods are defined: the scaling method and the stochastic method. The matching criteria of a jet at detector level and particle level is defined as

$$\Delta R < \frac{R_{cone}}{2}, |\Delta| < 3\sigma_{JER}, \tag{4.25}$$

where R_{cone} is the jet radius parameter of the reconstruction algorithm. For our measurenent AK4PFChs jets we take $R_{cone} = 0.4$.

When there is a matching jet at particle level the scaling method is applied. The four-momentum of the jet is smeared with a factor

$$c_{JER} = 1 + \Delta_{\text{Data/Sim}} \times \Delta, \tag{4.26}$$

However, when there is no matching particle level jet the stochastic method is used. The four-momentum is scaled with

$$c_{JER} = 1 + \mathcal{N}(0, \sigma_{JER}) \sqrt{\max(\Delta^2_{\text{Data/Sim}} - 1, 0)}, \tag{4.27}$$

where $\mathcal{N}(0, \sigma_{JER})$ is a random number generated according to a normal distribution with a width of σ_{JER}. If c_{JER} is negative, it is set to zero, as the transverse momentum of a jet cannot be negative. Thus after, jets are calibrated and the resolution of the reconstructed jets at detector level in simulations are smeared, jets are fully corrected and ready to be analysed for the measurement. In [99] information on the jet energy scale and resolution in the CMS experiment can be found.

4.3.2 Common corrections for physics analysis

Due to detector malfunctions, the amount of simulated data, the incorrect simulation of the detector and the resolution offered by the detector, apart from the jet energy

and resolution, differences between the simulated samples and the measured data persist. In this subsection we present the corrections we apply both in the measured data and the simulated data to remove the remaining discrepancies for a precise measurement in chapter 5.

Jet veto maps

Due to radiation damage, some of the detector components are not longer reliable for the reconstruction. This is the case for the HCAL, where some of the towers are damaged and the jet reconstruction in this area is not realiable. Hence, any jet reconstructed in the damaged region has to be removed.

The jet veto maps are boolean $\eta - \phi$ grids covering the whole detector $\eta - \phi$ phase space. Hence, if a jet falls into a cell of the grid with a False value the jet is removed from the event. The veto maps are derived by the JetMET group.

Muon momentum calibration

The measurement of the momentum of the reconstructed muon is affected by misalignments of the detector, the reconstruction software and uncertainties of the magnetic field. These effects are included in the simulation of the detector response in GEANT4. However, the inputs for the simulation of the effects are not completely correct and differences in the momentum between the measured muon and the simulated muon arises. To remove the discrepancy, the momentum of the muons is calibrated both in data and simulation. The correction factors where originally derived by the CMS group at the University of Rochester [100], and are better known as "Rochester corrections". The "Rochester corrections" used for the measurement in this thesis correspond to the roccor.Run2.v5 package.

Cross-section scale factor

The number of events in the simulated samples and recorded data do not match. To compare simulated events to data, we need to scale them using a weight. We derive the cross-section weight for each simulated event. For a given event, i, of a sample with N events, the weight is:

$$w_{XS}^{i} = \frac{\mathscr{L}\sigma_{sample}w_i}{\sum_N w_i}, \tag{4.28}$$

where σ_{sample} is the cross-section of the simulated process and w_i is the weight of the given event i. At leading order the weight, w_i, of an event is 1, however for events calculated at NLO the weight is no longer one. Events generated with a virtual correction will have a destructive contribution to the total number of events.

Multiple proton-proton interactions (pileup)

The proton beams at the LHC are not continuous, but they are organised in proton packets or bunches. These proton packets cross at the collision points producing proton-proton collisions. From these crossings more than one proton-proton interactions can happen. These multiple proton-proton interactions are known as pileup interactions.

The number of pileup interactions is determined from the instantaneous luminosity of a given bunch, $\mathscr{L}_{inst}$, the cross-section of inelastic processes, $\sigma_{inelastic}$, and the orbit frequency of the LHC, f_{LHC}:

$$N_{PU} = \frac{\mathscr{L}_{inst}\sigma_{inelastic}}{f_{LHC}} \tag{4.29}$$

The luminosity at CMS is measured in luminosity sections. The computed number of pileup interactions is the average of pileup interactions within a luminosity section. Since the bunch crossings are independent of each other we can describe the number of pileup interactions with a Poisson distribution.

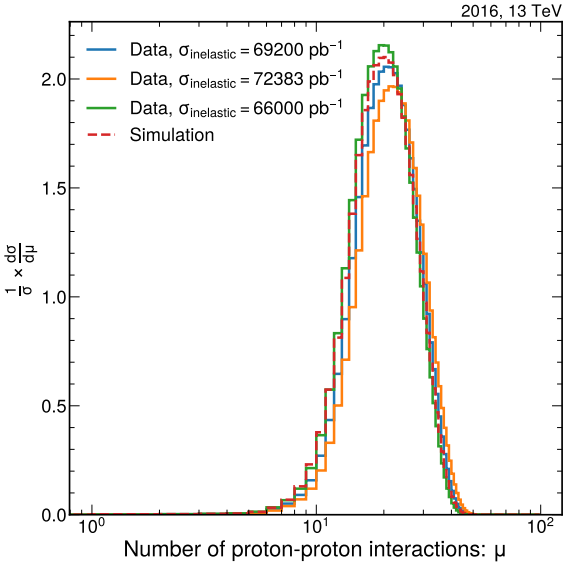

Figure 4.8: Pileup profile for data and simulation. The cross-section for inelastic QCD process at 13 TeV corresponds to $\sigma_{inelastic} = 69200$ pb^{-1}, with a 4.6% variation of $\sigma_{inelastic}$ the uncertainty on the number of pileup interactions in data is estimated. The profiles were derived using the pileupcalc tool from the Luminosity Physics Object Group.

The pileup interactions are simulated as QCD events following a Poisson distribution. The amount of pileup interactions are overestimated, as the beam configuration at the LHC is changed per era, the simulation is constant for the whole year. From the difference

between the pileup profiles in data and simulation we derive the weight for correcting for pileup, w_{PU}. In figure 4.8 we show the pileup profile for the 2016 data taking period for simulation and data, for a 4.6% variation in $\sigma_{inelastic} = 69200$ pb^{-1}. These profiles are then used to determine the pileup weights for the measurement in chapter 5.

Level-1 prefiring issue

In the years 2016 and 2017 there was a timing shift of the ECAL with respect to the level-1 triggers. This made physics objects mistakenly identified from the previous bunch crossing. The Level-1 prefiring mainly affects the large η region of muons and jets.

The level-1 prefiring issue is corrected in simulations with weights provided centrally: one weight for the HCAL(jets) and another one for muons. We combine both weights since both jets and muons are used in the analysis,

$$w_{L1} = w_{L1}^{jets} \cdot w_{L1}^{\mu} \tag{4.30}$$

The level-1 prefiring weights are centrally produced in the CMS experiment.

CHAPTER 5

Measurement of real electroweak corrections

In chapter 3 we have presented a theoretical overview on electroweak corrections. We have learned that these become large at high energies, at TeV scales. While virtual electroweak corrections are widely studied for different process, real electroweak radiations have not received much attention. In an era where new physics might hide in the realm of precision, with small disagreements between measurements and theoretical predictions, a proper understanding and modelling of Standard Model processes is necessary.

From the work in [58, 64] we know that for dijet events in hadronic collisions at high jet transverse momentum the cross-section deviates $1 - 2\%$ from the QCD calulation when different electroweak processes are included: real and virtual electroweak corrections and the electroweak production of the dijet.

The contribution of the electroweak production of dijets was measurend in proton-proton collisions at $\sqrt{s} = 13$ GeV in [101]. The cross-section for this process at high transverse momentum of the electroweak boson is around 10^{-4} pb^{-1}/GeV (see left panel in figure 5.1), whereas previous measurements of the inclusive jet cross-section [102] at high transverse momentum amounts to 10^{-2} pb^{-1}/GeV (see right panel in figure 5.1).

In this chapter we present the measurement of the contribution of real electroweak corrections to high transverse momentum jets, especially, the contribution of the real emission of the Z boson, with $Z \to \mu^+\mu^-$. Besides measuring the contribution of the radiated Z boson, this measurement can be used to study the BN violations we have presented in chapter 3, as the Z boson radiation must cancel with the corresponding Z boson virtual correction.

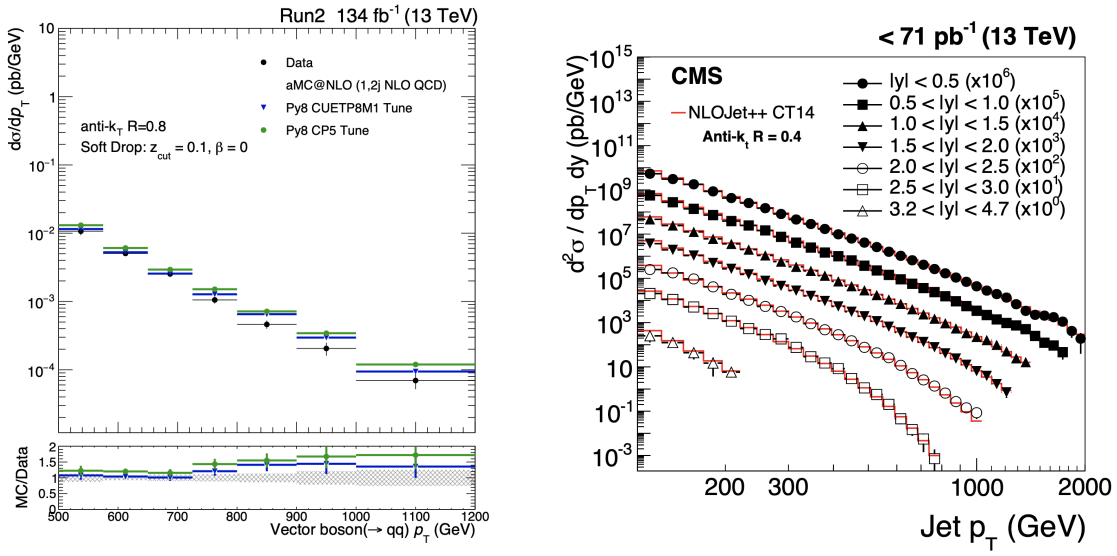

Figure 5.1: Left panel: differential cross-section of the electroweak boson production as a function of the vector boson transverse momentum (figure taken from [101]). Right panel: double differential cross-section of multijet events as a function of the transverse momentum and pseudorapidity of jets (figure taken from [102]).

This chapter is organised as follows, in section 5.1 we present the datasets and simulations used for the measurement. In section 5.2 we define the phase space of the measurement and the different measured observables. In section 5.3 we define the event selection. Then, in section 5.7 we present the uncertainties affecting our measurement. We show the measured cross-sections in section 5.8 and we extract the contribution of electroweak corrections in QCD events in section 5.9.

5.1 Data and simulated samples

In this section we present the data and simulated samples used in the measurement of the contribution of electroweak corrections:

5.1.1 Data samples

After the event reconstruction by the HLT trigger system, the event is stored by HLT triggers, in a dataset corresponding to the HLT trigger it has fired. To measure the contribution of real electroweak radiation we measure two types of events, QCD events and Z+jets events. For the analysis of QCD events we are interested in events firing jet HLT triggers, which are stored in the JetHT dataset. For the Z+jets events, we could

make use of the JetHT dataset, however the jet triggers are highly prescaled. Due to the cross-section of QCD processes, storing all events triggering jet triggers will not be viable, as too much storage space will be needed. For that, jet triggers do not save all events but only a fraction of them. This fraction is called the prescale factor of the trigger. Thus, for Z+jets events we use the double muon dataset.

The data used in this measurement was recorded during RunII of the LHC in 2016 at a centre-of-mass energy of 13 TeV. The recorded proton-proton collisions during 2016 corresponds to 36.3 fb^{-1}.

5.1.2 Simulated samples

We make use of simulated events for a better understanding of the physics scenario we are studying. From simulated samples we derive efficiency and background corrections, as well as, systematic uncertainties that affect our measurement, and we can correct the measured data from the detector response. Background processes are processes that produce the same final state as the process under study. We refer to the process under study as the signal process. For each type of event, QCD and Z+jets events, we have a signal process.

All the simulated samples are generated using the NNPDF3.1 PDF set at next-to next-to leading order accuracy in QCD [103]. Besides, the simulated samples used for the measurement include the simulation of the underlying event with the Pythia8 event generator. The free parameters of the simulation of the underlying event are extracted from measurements, this process is called the tuning of the parameters. The tune used for the simulation of the underlying event in Pythia8 is the CP5 tune [30].

For the QCD sample events are generated at leading order. As QCD processes have the largest cross-section in hadronic collisions no background processes are expected.

For the electroweak radiation of the Z boson we use the DY+Jets sample, where the DY decays into a lepton pair. It is calculated at next-to-leading order with up to two partons using Madgraph [104]. The matching of the next-to-leading order events with the Pythia8 parton shower is done with the FXFX matching procedure [105]. We have discussed in chapter 3 that there are no dedicated simulations of electroweak radiations available in the CMS experiment, however, the DY+2 Jets process contains Feynman diagrams corresponding to the radiation of the Z boson.

Other processes are able to produce the same final state as our signal: two or more jets with two opposite charge muons. The background processes we consider are: top pair production, W boson production in association with jets, single top production and double boson scattering. In table 5.1 we summarize the simulated sample used for signal and background processes.

Simulations

Process	**Phase space cuts**	**Generator**	**Cross-section [pb]**	**Label**
QCD	$80 < p_T < 120$	Pythia8	$2.346 \cdot 10^6$	QCD
	$120 < p_T < 170$		$4.077 \cdot 10^5$	
	$170 < p_T < 300$		$1.037 \cdot 10^5$	
	$300 < p_T < 470$		$6.826 \cdot 10^3$	
	$470 < p_T < 600$		$5.512 \cdot 10^2$	
	$600 < p_T < 800$		$1.567 \cdot 10^2$	
	$800 < p_T < 1000$		$2.625 \cdot 10^2$	
	$1000 < p_T < 1400$		7.465	
	$1400 < p_T < 1800$		0.649	
	$1800 < p_T < 2400$		$8.734 \cdot 10^{-2}$	
	$2400 < p_T < 3200$		$5.237 \cdot 10^{-3}$	
	$3200 < p_T < \infty$		$1.352 \cdot 10^{-4}$	
$Z \to ll + 0, 1, 2$jets	$M_{ll} > 50$ GeV	aMC@NLO FxFx	5765.4	Z+jets
$t\bar{t} \to 2l + 2\nu$		Powheg	87.3	$t\bar{t}$
Single top	s-channel	aMC@NLO	3.549	ST
	t-channel	Powheg+Madspin	69.09	
	tw antitop	Powheg	34.97	
	tw top	Powheg	34.91	
$W \to l\nu + 0, 1, 2$jets		MadgraphMLM	61526.7	W+jets

Table 5.1: Simulation samples used in the analysis.

5.2 Definition of the phase space and observables

Now we define the phase space where multijet events are sensitive to electroweak radiations. This will be used as a base for the phase space for our measurement. From the studies on electroweak corrections in chapter 3, the probability of radiating an electroweak boson reads as

$$\alpha_{EW} \ln^2 \frac{Q^2}{m_V^2}, \tag{5.1}$$

where α_{EW} is the coupling strength of electroweak interactions, Q is the energy scale of the partonic interaction and m_V is the mass of the radiated electroweak boson, with $V = Z, W^{\pm}$. In a pure QCD dijet process the two outgoing jets are in a back-to-back configuration, their opening angle between the jets is $\Delta\phi = \pi$, for the momentum conservation. Besides, they will be produced in the central region of the detector. Then, the energy scale of the dijet production, Q, can be defined as

$$Q = p_{T,j_1} = p_{T,j_2} = \frac{1}{2}\left(p_{T,j_1} + p_{T,j_2}\right), \tag{5.2}$$

where p_{T,j_1} and p_{T,j_2} are the transverse momentum of the two jets. Thus, to radiate an electroweak boson p_{T,j_1} and p_{T,j_2} have to be larger than m_V. We need to be in a phase space where the transverse momentum of the jets is larger than the mass of the Z boson. For that, we require a minimum transverse momentum of the jets, $p_{T,j}$, larger than 100 GeV. In addition, we take jets in the central region of the detector, with $|\eta| < 2.4$, to make sure the jets are produced in the partonic interaction and not in a radiation from the initial state.

In this region of the phase space we are sensitive to QCD dijet process where a Z boson has been radiated. Moreover, with a proper definition of the observables we are able to disentangle between radiated Z bosons, from Z bosons produced in the hard process.

5.2.1 Observables

In figure 5.2 we show a sketch where a Z boson has been radiated from a jet in a dijet event. Due to the radiation, the two jets are not in a back-to-back configuration. The radiation of the Z boson modifies the transverse momentum of the parent jet, thus equation 5.2 is no longer valid. However, we can use the transverse momentum of the leading jet to determine the energy scale of the hard-process.

From the work in [58,64] we know that the cross-section as a function of the jet transverse momentum is affected by electroweak corrections (see figures 3.4). Hence, to measure the contribution of real electroweak corrections, we measure jet transverse momentum spectrum for QCD events and Z+jets events.

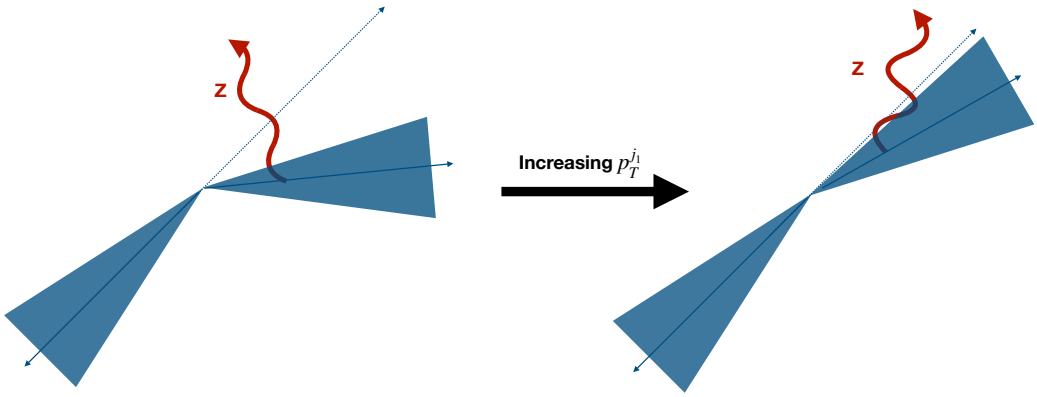

Figure 5.2: Sketch of a dijet event where a Z boson has been radiated in the final state. At low leading jet transverse momentum we expect a separation between the Z boson and the radiated jet. This separation is reduced for increasing leading jet transverse momentum.

Besides, to understand the behaviour of real electroweak radiations, we measure the observables presented in section 3.4, which are sensitive to the radiation of an electroweak boson: the azimuthal correlation between the two leading jets, the minimum ΔR between the Z boson and a jet, the transverse momentum of the Z boson and the azimuthal correlation between the dijet system and the Z boson. We measure these observables in different regions of the leading jet transverse momentum, which we have related to the energy scale of the hard-scattering. We investigate in the following leading jet transverse momentum regions:

$$100 < p_{\mathrm{T}}^{j_1} < 250 \text{ GeV}, \; 250 < p_{\mathrm{T}}^{j_1} < 500 \text{ GeV}, \; p_{\mathrm{T}}^{j_1} > 500 \text{ GeV}.$$

5.3 Event selection

In the previous section we have defined the phase space where dijet events are sensitive to electroweak radiations. We are measuring the contribution of the Z boson radiation, where we only consider the $Z \to \mu^+\mu^-$ decay channel. In this analysis we make use of jets and muons.

We use anti-k$_T$ PF jets with a distance parameter $R = 0.4$, with charge hadron subtraction. For the analysis we consider central jets, $|\eta| < 2.4$, with tight identification criteria, and a minimum transverse momentum of 100 GeV. The tight identification criteria are given in table 5.2. To consider the event as a dijet event we require a minimum of two jets.

Jet identification criteria for AK4PFCHS jets

PF Jet ID	Tight ID
Neutral hadron fraction	<0.90
Neutral EM fraction	<0.90
Number of Constituents	>1
Muon fraction	-
Charged hadron fraction	>0
Charged multiplicity	>0
Charged EM fraction	-
Number of neutral particles	-

Table 5.2: Jet identification criteria.

Once the dijet event has been selected, we consider muons in the tracker acceptance region, $|\eta| < 2.4$, with a minimum transverse momentum of 20 GeV and satisfying medium identification criteria. The medium identification criteria are defined as global or tracker muon with additional requirements on track and muon quality. Besides, we apply a tight PF isolation criterion to remove muons contaminated by pileup that might affect the reconstruction of the muon. The PF isolation is defined as

$$I_{PF} = \frac{1}{p_{T,\mu}} \left(\sum^{charged} p_T + \max(0, \sum^{neutral} p_T + \sum^{EM} p_T - \rho EA) \right), \tag{5.3}$$

where $p_{T,\mu}$ is the transverse momentum of the muon and we sum over the p_{T} of PF candidates inside a cone of radius 0.4 around the direction of the muon: charged hadrons coming form the primary vertex (charged), neutral hadrons (neutral), photons (EM) and charged particles from pileup (ρEA). A tight PF isolation corresponds to $I_{PF} < 0.15$.

We require exactly two muons in the event with opposite charge to proceed with the Z boson reconstruction. The invariant mass of the muon pair has to be in $76 < m_{\mu^+\mu^-} < 106$ GeV. To fulfil the requirement of the transverse momentum threshold of the muon trigger the transverse momentum of the leading muon has to be larger than 25 GeV. In the following section we present the different triggers used in the analysis.

At this stage we categorise the selected events as QCD or Z+jets events. We define a

QCD event as a dijet event where no Z boson has been reconstructed. When a Z boson is reconstructed in the dijet event we define the event as Z+jets.

In Z+jets events we perform a ΔR cleaning between the decay products of the Z boson (the muons) and the jets. In the jet reconstruction all PF candidates are used (muons, electrons, hadrons). Whenever extra activity is found near a muon a jet will be reconstructed, where most of its momentum is coming from the muon. Hence, the muon is faking a jet. This leads to misinterpretation of the event shape. To remove these fakes jets we apply a minimum ΔR distance between a muon and the jets of 0.2 (see appendix B).

5.4 Trigger strategy

The decision to store an event is done by the HLT trigger system. The online event reconstruction needs to be fast, as the decision has to be done in-situ. After the event is stored, the offline reconstruction is performed with the PF algorithm, which is slower and more precise. The difference in precision between the HLT and the PF algorithm leads to a mismatch in the momentum reconstruction of the physics objects. The HLT triggers are fired depending on transverse momentum thresholds. However, these thresholds correspond to the HLT reconstruction, and thus, for the PF reconstruction new thresholds are needed. This causes the HLT trigger not to be fully efficient at PF level. In this section we present the trigger strategy for QCD and Z+jets events.

5.4.1 QCD events

For QCD events we make use of single jet triggers, which are named as HLT_PFJet$X$, where X is the transverse momentum threshold of the HLT reconstruction. Hence, a single jet trigger is fired when in the HLT reconstruction there is a jet with transverse momentum larger than X. To do the analysis using the PF reconstruction, we need to compute the transverse momentum thresholds of the single jet triggers. The new thresholds are determined by computing the turn-on point of the trigger. The turn-on point is the transverse momentum in the PF level where the trigger becomes fully efficient, with an efficiency above 99%.

Due to the large number of QCD events, not all the events that fire single-jet triggers are recorded, these triggers are pre-scaled. The lower the transverse momentum threshold the larger the pre-scale is. In table 5.3 we present the triggers used for the analysis of QCD events, with the corresponding thresholds, turn-on points and the effective recorded luminosity. Each trigger probes a different leading jet transverse momentum region. For the HLT_PFJetX$_i$ trigger probes the region corresponds to

$$p_T^{j_1} \in (Y_i, Y_{i+1}), \tag{5.4}$$

where Y_i and Y_{i+i} are the turn-on points for the HLT_PFJetX$_i$ and HLT_PFJetX$_{i+1}$ triggers respectively. For QCD events the triggers are only applied in data. With the definition of the turn-on point, the triggers are fully efficient in their corresponding region. Hence, we do not need to apply triggers in the simulation.

Trigger path	HLT threshold [GeV]	Effective luminosity [fb^{-1}]	Turn-on point [GeV]
HLT_PFJet60	60	0.003	97
HLT_PFJet80	80	0.003	133
HLT_PFJet140	140	0.024	172
HLT_PFJet200	200	0.104	242
HLT_PFJet260	260	0.592	304
HLT_PFJet320	320	1.765	366
HLT_PFJet400	400	5.171	446
HLT_PFJet450	450	35.9	548
HLT_PFJet500	500	35.9	592

Table 5.3: HLT triggers used for the QCD event selection with their corresponding p_{T} threshold of the HLT reconstruction, effective recorded luminosity and turn-on point where the trigger becomes fully efficient for a jet reconstructed with the PF algorithm.

As mentioned before HLT_PFJetX triggers are prescaled, not all the events that trigger single jet triggers are recorded. To correct for this we reweight the events. We define the weight corresponding for an event triggering the HLT_PFJetX trigger as

$$w_{HLT} = \frac{\mathscr{L}_{total}}{\mathscr{L}_{recorded}}, \tag{5.5}$$

where $\mathscr{L}_{total}$ is the total recorded luminosity during the run and $\mathscr{L}_{recorded}$ is the recorded luminosity by the trigger. In the following we explain the determination of the turn-on points for the HLT_PFJetX triggers.

Computation of turn-on points

To calculate the turn-on point of a trigger, first the efficiency of the trigger has to be determined. Different methods can be used to determine the trigger efficiency, for example: the tag and probe method, the reference trigger method and the trigger emulation method.

With the reference trigger method we compute the efficiency of the HLT_PFJetX$_i$ trigger relative to the HLT_PFJetX$_{i-1}$. To compute the efficiency both triggers need to be fired. This will affect the statistical uncertainty of the efficiency distribution, since both triggers can be prescaled. The trigger emulation method is an improved method of the reference trigger method. In this case, the trigger whose turn-on point is computed, HLT_PFJetX$_i$,

is not used, but emulated and HLT_PFJetX$_{i-1}$ is used as a reference. In this way the statistical uncertainty of the efficiency distribution is reduced in comparison to the reference trigger method.

In this analysis we use the trigger emulation method to determine the trigger efficiency. The emulation of the trigger consists on selecting events that should fire the trigger. The efficiency is defined as

$$\epsilon = \frac{N(\mathrm{T}^{\mathrm{fired}}_{\mathrm{emulated}}|\mathrm{T}^{\mathrm{fired}}_{\mathrm{reference}})}{N(\mathrm{T}^{\mathrm{fired}}_{\mathrm{reference}})}, \tag{5.6}$$

where T stands for trigger. To emulate the single-jet triggers the following cuts are applied:

- A dijet final state
- Back to back topology of the two leading jets: $\Delta\phi(j_1 j_2) > 2.7$
- For events with more than two jets, the p_T of the leading jets has to be larger compared to rest of the jets: $\frac{1}{2}\frac{p_T(j_i)}{p_T(j_1)+p_T(j_2)} < 0.3 : i > 2$
- The two leading jets have to match in ΔR with the two leading HLT jets: $\Delta R(j^{HLT} j^{PF}) < 0.4$
- The p_T of the leading HLT jet has to be larger than the p_T threshold of the jet: $p_T(j_1^{HLT}) > \mathrm{X}$

To determine the turn-on point from the efficiency a fit is done. The efficiency is fitted with a modified version of the error function:

$$\epsilon(p_T(j_1)) = a + 0.5(1-a)\left(1 + \mathrm{erf}\left(\frac{p_T^{PF} - \mu}{\sigma}\right)\right), \tag{5.7}$$

a, μ and σ are the parameters to fit. The turn-on point is defined as the p_T^{PF} value where the efficiency is 99%. Table 5.3 shows a summary of the turn-on points.

5.4.2 Z+jets events

The muon triggers HLT_IsoTKMu24 and HLT_IsoMU24 are applied both to data and simulation. In principle, we could follow the trigger strategy of QCD events and compute the corresponding turn-on points for the triggers. However, in CMS the efficiency of the single muon triggers are centrally calculated by the Muon physics object group. These efficiencies are applied in the simulations as a weight to the event depending on the transverse momentum and rapidity of the leading muon.

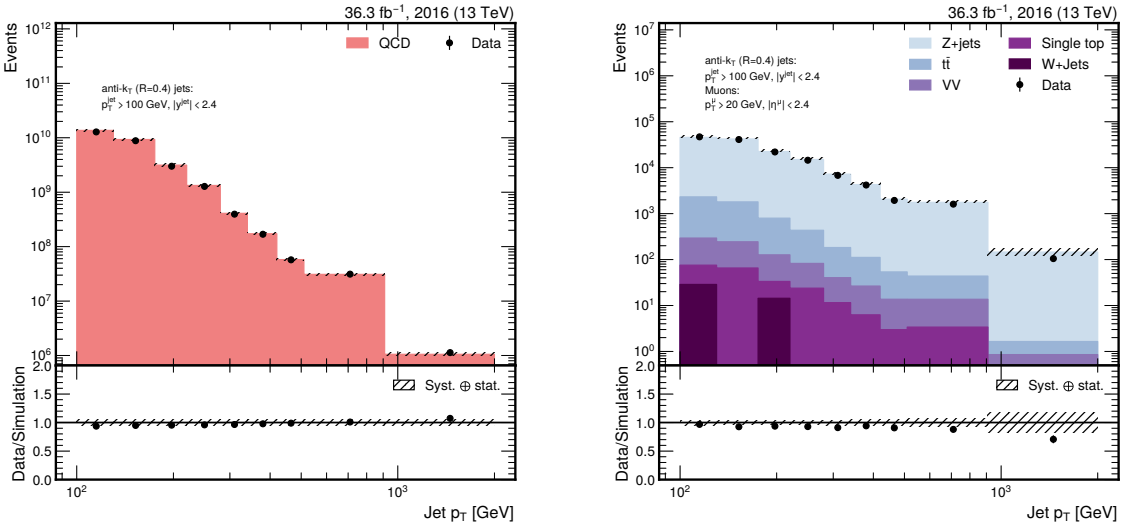

Figure 5.3: The jet transverse momentum comparison between data and simulation after event selection for QCD events (left panel) and Z+jets events (right panel).

5.5 Measurement at detector level

In this section we describe the measurement of the observables presented is section 5.2 and compare the measured data to the theoretical predictions in section 5.1 after the event selection at detector level. In this way we make sure that theoretical predictions are able to describe the data and our assumptions on background processes are correct.

It is important to note that for the systematic uncertainty we include jet energy scales and resolution, pileup and luminosity uncertainties for QCD events. The systematic uncertainty for Z+jets events includes jet energy scale and resolution, pileup, luminosity, muon identification and isolation and single muon trigger uncertainties. We represent the systematic uncertainty together with the statistical uncertainty of the simulation with hashed lines and the statistical uncertainty of data with errorbars.

In figure 5.3 we show the measured jet transverse momentum distribution and compare it to theoretical predictions. For QCD events (left panel) we compare the data to the QCD simulation. We observe that QCD events are well described by the QCD simulation at leading order accuracy. For Z+jets events (right panel) we compare the data with a stack of the different simulations contributing to the final state of Z+jets events. We observe that the simulation is able to describe the measured data. However, at large jet transverse momentum $\mathcal{O}$(1 TeV) the simulation tends to overshoot the measurement.

In figures 5.4-5.7 we present distributions for the different observables for Z+jets events in

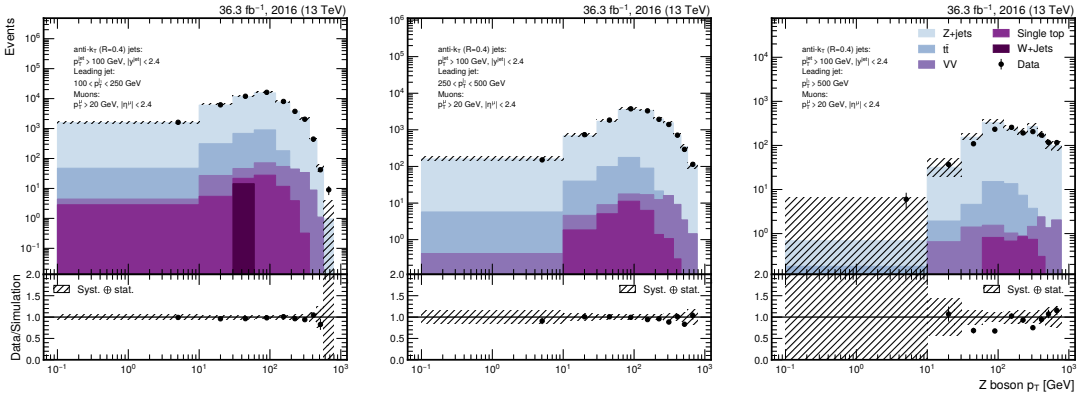

Figure 5.4: The Z boson transverse momentum comparison between data and simulation after event selection for Z+jets events in bins of leading jet transverse momentum, $p_{\mathrm{T}}^{j_1}$: $100 < p_{\mathrm{T}}^{j_1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j_1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j_1} > 500$ GeV (right panel).

different regions of the leading jet transverse momentum. Overall, in the $p_{\mathrm{T},j_1} > 500$ GeV region the simulations are statistically limited, which makes the uncertainty band large. The transverse momentum of the Z boson distribution is centred around 100 GeV for the three different p_{T,j_1} regions (see figure 5.4). In the $100 < p_{\mathrm{T},j_1} < 250$ GeV region most of the events have a p_{T}^{Z} similar or larger than $p_{\mathrm{T}}^{j_1}$. This region is dominated mainly by Drell-Yan events, rather than radiated Z bosons. As we increase $p_{\mathrm{T}}^{j_1}$ the p_{T}^{Z} becomes smaller than $p_{\mathrm{T}}^{j_1}$ making the Z boson radiation more dominant.

The measurement of the minimum ΔR distance between the Z boson and a jet, $\Delta R_{Z,j}^{Min}$, agrees with the simulation (see figure 5.7). We can observe that for increasing $p_{\mathrm{T}}^{j_1}$ the minimum ΔR distance decreases. The azimuthal correlation of the two leading jets (see figure 5.5) increases with increasing p_{T,j_1}. From $\Delta R_{Z,j}^{Min}$ and $\Delta\phi_{j_1,j_2}$ we can infer an event configuration depicted in figure 5.2, where at large p_{T,j_1} the emitted Z boson is collinear to the emitter jet and dominates the large p_{T,j_1} region.

5.6 Correction to particle level

The response of the detector creates a mismatch between the measured and the real distribution. Even though, we correct all the biases produced by the detector response in the measurement, we need to remove the detector response for a precision measurement and correct the measured distribution to particle level. This procedure is called unfolding. The unfolding can be formulated mathematically as a transformation, from detector level,

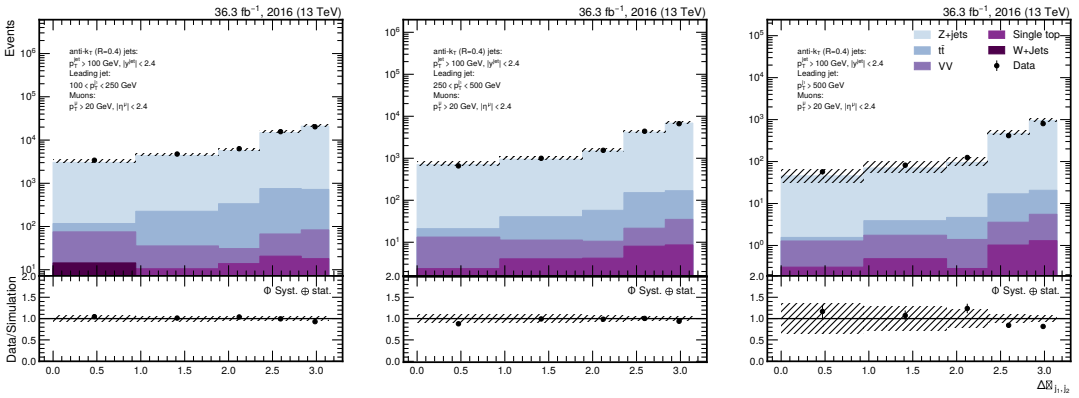

Figure 5.5: Data and simulation comparison of the minimum $\Delta\phi$ between the two leading jets after event selection for Z+jets events in bins of leading jet transverse momentum, $p_{\mathrm{T}}^{j_1}$: $100 < p_{\mathrm{T}}^{j_1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j_1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j_1} > 500$ GeV (right panel).

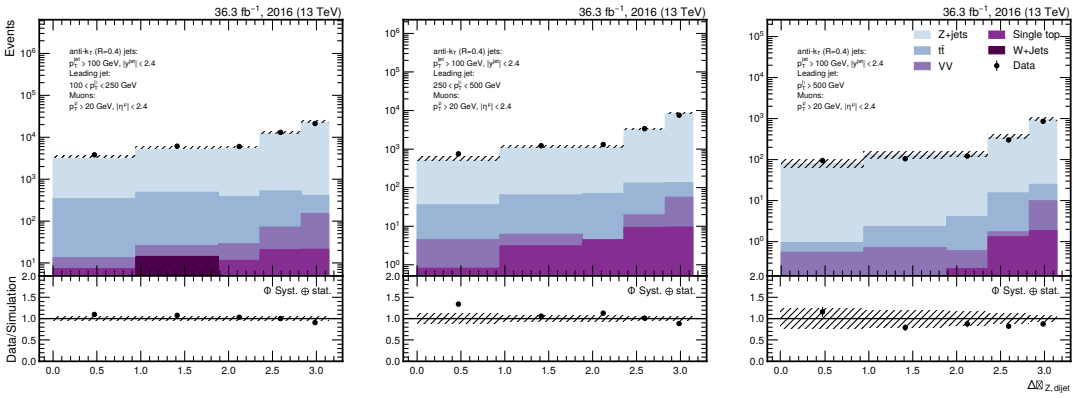

Figure 5.6: Data and simulation comparison of the minimum $\Delta\phi$ between the Z boson and the dijet after event selection for Z+jets events in bins of leading jet transverse momentum, $p_{\mathrm{T}}^{j_1}$: $100 < p_{\mathrm{T}}^{j_1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j_1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j_1} > 500$ GeV (right panel).

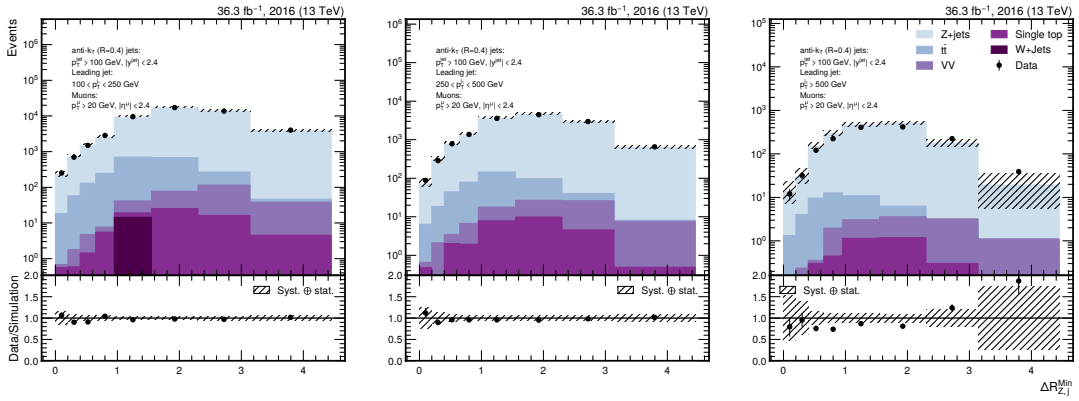

Figure 5.7: Data and simulation comparison of the minimum ΔR between the Z boson and a jet after event selection for Z+jets events in bins of leading jet transverse momentum, $p_{\mathrm{T}}^{j_1}$: $100 < p_{\mathrm{T}}^{j_1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j_1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j_1} > 500$ GeV (right panel).

x, to particle level, y:

$$\bar{x} = A\bar{y}, \tag{5.8}$$

where $\bar{x}$ and $\bar{y}$ represent the observable at the corresponding level and A is the transformation matrix, which we will refer to as probability matrix. Hence, we retrieve $\bar{y}$ with:

$$\bar{y} = A^{-1}\bar{x}, \tag{5.9}$$

making the unfolding an inversion problem. However, sometimes the probability matrix is ill posed. We determine if a matrix is ill posed by calculating the condition number of the matrix: the condition number estimates how much the result of a function changes by a small variation in the input parameters. The condition number of a matrix is defined as:

$$\mathscr{K} = \frac{\sigma_{max}(A)}{\sigma_{min}(A)}, \tag{5.10}$$

where $\sigma_{max}(A)$ and $\sigma_{min}(A)$ are the maximum and minimum eigenvalues of the matrix A. With a large condition number, the matrix is ill posed, the result will have a large variation with a small variation of the input parameters. With a small condition number, the matrix is well conditioned. With the condition number we can do a rough estimation on how the accuracy of the result can vary due to variation in the input

$$\mathscr{K} = 10^k \tag{5.11}$$

where k is a rough estimate of the digits of accuracy we may loose. Thus, to define if a matrix is ill posed will depend on the accuracy of the measurement. The statistical

committee of the CMS experiments recommends a condition number below 10 to consider the probability matrix well posed. However, this is an arbitrary value, as measurements looking at 1% may introduce another 1% inaccuracy in the measurement. For our measurement we observe that for the jet transverse momentum distribution in QCD events our uncertainty is around 5%, while for Z+jets observables at large leading jet transverse momentum we are statistically limited by the simulation sample. Thus, while a condition number of 10 might be too large for the jet transverse momentum in QCD events, for Z+jets observables might be appropriate. Whenever, the condition number is too large for a given observable the probability matrix is ill posed and a different unfolding method should be applied. Our uncertainties at detector level are larger than 5%, thus a condition number smaller than 10 is desired.

We divide the correction to particle level into three steps. First we define the probability matrix and then we study migration effects that affect the observable. Lastly, we unfold the measured data correcting it to particle level.

5.6.1 Probability matrix

We derive the probability matrix, A, in equation 5.8 with the response matrix:

$$A_{ij} = R_{ij}/\bar{y_j} \tag{5.12}$$

where R is the response matrix and $\bar{y}$ is the distribution at particle level. It is important to remember that the simulation of events contains a simulation of the detector, which introduces two levels of final state: a particle and a detector level. As the name suggest the response matrix is a measurement of the simulated detector response for a given observable. The response matrix is a two dimensional histogram and for each measured observable we define a response matrix. One axis of the response matrix corresponds to the detector level and another for the particle level. We define i and j as the detector and particle level bin indexes respectively.

We can interpret the probability matrix as the probability of finding a measured event in bin i with a corresponding event at particle level in bin j. In the following we introduce the response matrices for the different observables and their condition number.

We fill the response matrix using the simulation sample of the signal. For the QCD events we use the QCD sample (Pythia8 with NNPDF3.1), and for Z+jets events we use the Z+jets sample (MC@NLO with NNPDF3.1). The background samples are used to subtract the background contribution from data before correcting to particle level.

To fill the response matrix for the jet transverse momentum we follow the procedure of previous jet transverse momentum procedures explained in [106]. To avoid jets coming from pileup contaminating our response matrix and to identify the corresponding jets

between the particle and detector level, we apply a matching criterion between the jets at particle and detector level. We apply a ΔR matching, with a maximum distance between the jets of $\Delta R = 0.2$. If more than one particle level jet is matched to the detector level jet, we match it to the one with highest transverse momentum. In [106] it was found that there was no sensitivity for different values of the maximum $\Delta R = 0.15, 0.4$.

We show the response matrix for QCD and Z+jets events for the jet transverse momentum in figure 5.8. For QCD events (left panel) and Z+jets events (right panel) the response matrices show a correlation between particle and detector level, with anti diagonal matrices. The condition number for QCD and Z+jets response matrices are 2.9 and 2.4 respectively. Due to our uncertainties in the jet transverse momentum spectrum we can consider that the response matrices are well conditioned.

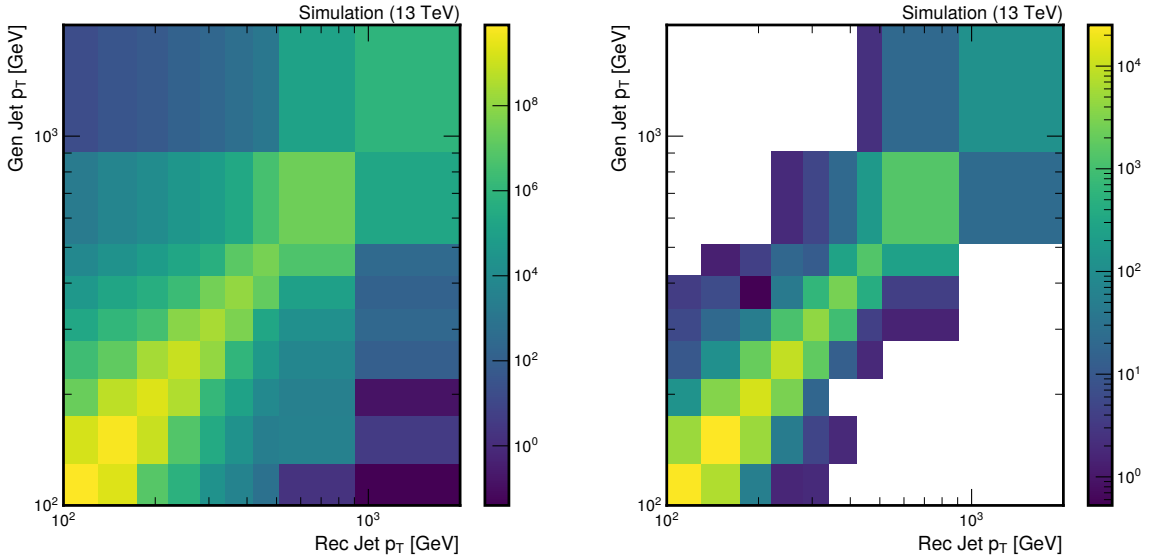

Figure 5.8: Response matrices for the jet transverse momentum for QCD events (left panel) and Z+jets events (right panel).

The jet transverse momentum distribution is sensitive to jets coming from pileup. However, the Z boson is only reconstructed when two muons are in the event. Thus, the Z boson reconstruction is not sensitive to muons arising from pileup, and muons do not need to be matched at detector and particle level.

In figure 5.9 we show the response matrix of the Z boson transverse momentum. Most of the off diagonal elements are zero due to the good spatial and energy resolution of muon reconstruction. This does not cause a large disagreement between the energy of the detector and particle level Z bosons. The condition numbers of the three different regions

of the leading jet transverse momentum for the Z boson transverse momentum response matrices are 1.7, 1.5 and 2.11 respectively.

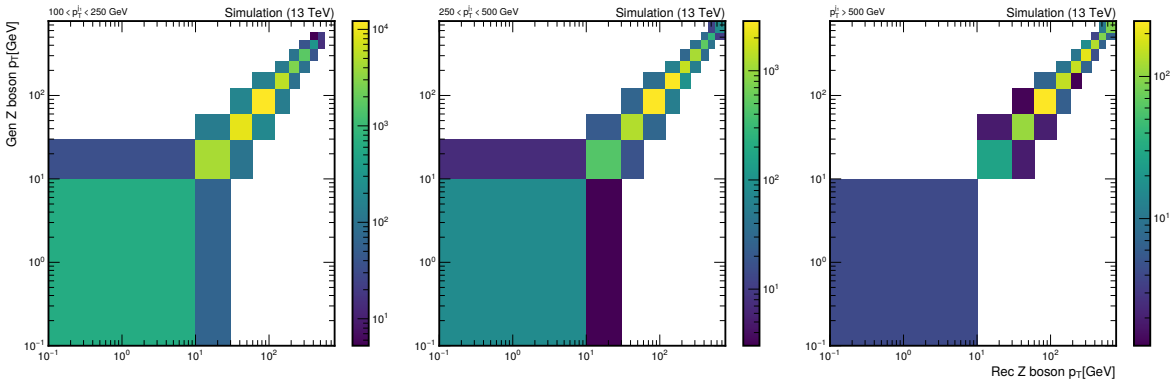

Figure 5.9: Response matrices for the Z boson transverse momentum in bins of leading jet transverse momentum, p_{T}^{j1}: $100 < p_{\mathrm{T}}^{j1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j1} > 500$ GeV (right panel).

For observables measured in Z+jets events the two leading jets are well defined both at detector and particle level. The contamination from pileup is rather small. Thus, we do not need to apply a matching criteria for the jets for the $\Delta\phi$ and ΔR distributions. In figures 5.10-5.12 we present the response matrices for the remaining distributions, where we observe an anti-correlation for the response matrices, where the off diagonal elements are zero or almost zero. In table 5.4 we summarise the condition numbers for all the different response matrices. Note that all condition numbers are below 5, thus making the probability matrix well posed for a matrix inversion.

5.6.2 Migrations

After defining the response matrices we can proceed with the study of the migrations. A migration has happened when the particle level observable and the detector level observable do not belong in the same bin or phase space. These migrations can be caused due to the finite resolution of the detector. We define two types of migrations:

- **Migrations inside the phase space (M1):** the events at detector and particle level pass all the cuts. However, they fall in different bins of the observable. These type of migrations are largely affected by the finite resolution of the detector. Larger bin widths avoids large migrations.

  We study the M1 migrations with the **purity and stability**. The purity of a given bin, i, is defined as the fraction of matched events, N_{matched}, where both particle

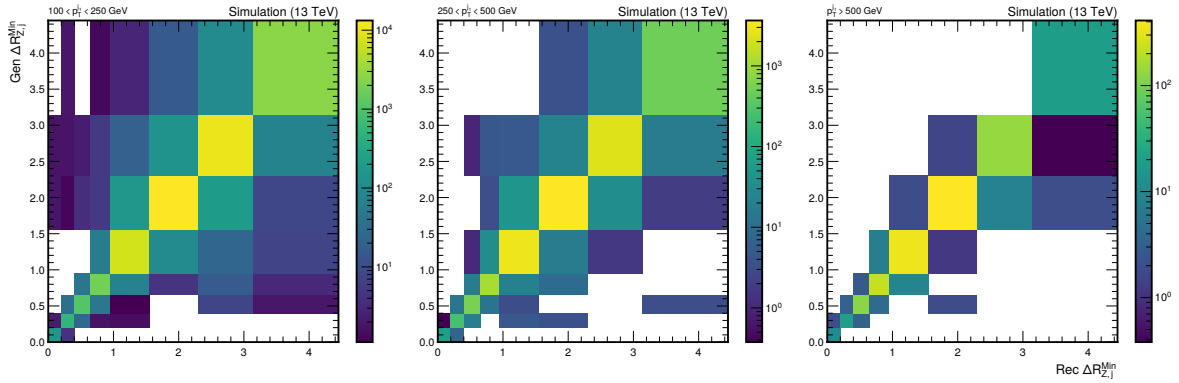

Figure 5.10: Response matrices for the minimum ΔR distance between the Z boson and a jet, $\Delta R_{Z,j}^{Min}$, in bins of leading jet transverse momentum, $p_{\mathrm{T}}^{j_1}$: $100 < p_{\mathrm{T}}^{j_1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j_1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j_1} > 500$ GeV (right panel).

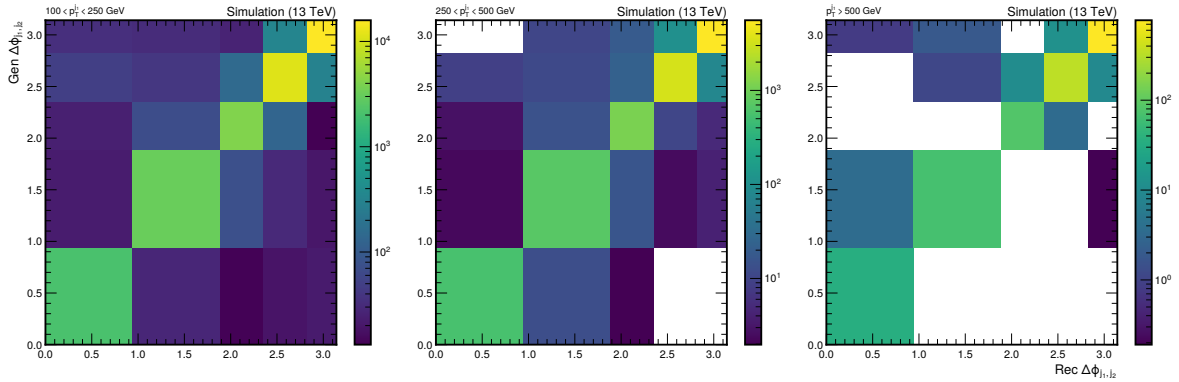

Figure 5.11: Response matrices for the $\Delta\phi$ between the two leading jets, $\Delta\phi_{j_1,j_2}$, in bins of leading jet transverse momentum, $p_{\mathrm{T}}^{j_1}$: $100 < p_{\mathrm{T}}^{j_1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j_1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j_1} > 500$ GeV (right panel).

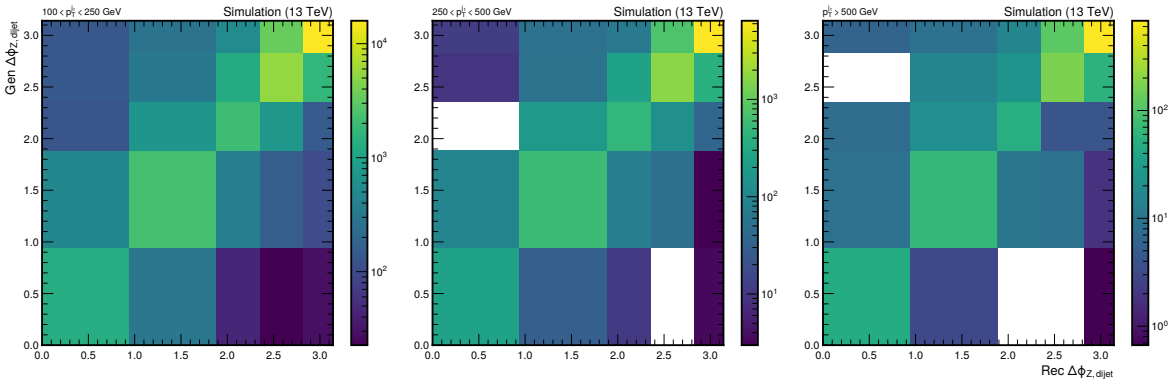

Figure 5.12: Response matrices for the $\Delta\phi$ between the Z boson and the dijet system, $\Delta\phi_{Z,dijet}$, in bins of leading jet transverse momentum, $p_{\mathrm{T}}^{j_1}$: $100 < p_{\mathrm{T}}^{j_1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j_1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j_1} > 500$ GeV (right panel).

Observable	**Event**	**Condition number**
Jet transverse momentum	QCD	2.86
	Z+jets	2.45

Observable	**Event**	**Leading jet p_{T}**	**Condition number**
$\Delta\phi_{j_1,j_2}$	Z+jets	$100 < p_{\mathrm{T}} < 250$ GeV	1.15
		$250 < p_{\mathrm{T}} < 500$ GeV	1.12
		$p_{\mathrm{T}} > 500$ GeV	1.43
$\Delta\phi_{Z,dijet}$	Z+jets	$100 < p_{\mathrm{T}} < 250$ GeV	2.77
		$250 < p_{\mathrm{T}} < 500$ GeV	2.62
		$p_{\mathrm{T}} > 500$ GeV	2.29
$p_{\mathrm{T}Z}$	Z+jets	$100 < p_{\mathrm{T}} < 250$ GeV	1.34
		$250 < p_{\mathrm{T}} < 500$ GeV	1.65
		$p_{\mathrm{T}} > 500$ GeV	4.31
$\Delta_R^{Min}{}_{Z,j}$	Z+jets	$100 < p_{\mathrm{T}} < 250$ GeV	1.74
		$250 < p_{\mathrm{T}} < 500$ GeV	1.52
		$p_{\mathrm{T}} > 500$ GeV	2.11

Table 5.4: Condition numbers of the response matrices for the measured observables.

and detector level are inside the same bin, and the matched events which are at detector level inside that bin:

$$P_i = \frac{N_{\text{matched}}(E_i^{det.} \wedge E_i^{part.})}{N_{matched}(E_i^{det.})}. \tag{5.13}$$

The stability of a given bin, i, is the fraction of matched events where both particle and detector level are inside the same bin, and total of matched events which are at particle level inside that bin:

$$S_i = \frac{N_{\text{matched}}(E_i^{det.} \wedge E_i^{part.})}{N_{matched}(E_i^{gen.})}. \tag{5.14}$$

- **Migrations outside the phase space (M2):** the events at detector and generator level are not in the same phase space. We define two types of events for M2 migrations: misidentified events and efficiency losses. Misidentified events are detector level events with no matching particle level event, we will refer to them as "fake" events. Efficiency losses are particle level events with no matching detector level events, which we will refer as "miss" events.

  These migrations are studied with the **acceptance and background**. The acceptance of a given bin, i, is the fraction of matched events in a bin i and the total events in that bin at generator level:

$$A_i = \frac{N_{matched}(E_i^{det.} \wedge E_i^{part.})}{N_{matched}(E_i^{part.}) + N_{Miss}(E_i)}. \tag{5.15}$$

  The acceptance measures the migration of events outside the phase space. The background in a given bin, i, is the migration of events into the phase space:

$$B_i = 1 - \frac{N_{matched}(E_i^{det.} \wedge E_i^{part.})}{N_{matched}(E_i^{det.}) + N_{Fake}(E_i)}, \tag{5.16}$$

  Next we study the migrations for each observable separately.

Migration study: jet transverse momentum

For the jet transverse momentum in QCD and Z+jets events the migration effects are similar. We observe that migrations inside the phase space are considerable in the lowest jet p_T region, with a purity and stability of around 50-60% (see figures 5.13 and 5.14, left panel). For larger jet p_T ($p_T > 300$ GeV), the purity and stability rise above 60%.

Overall, the migration of jets outside the phase space is stable throughout the whole jet p_T range, with an acceptance of 60% (see figures 5.13 and 5.14, right panel). The migrations of jets into the phase space is small. It mainly affects the lowest jet p_T region, where the cut on the minimum p_T is done.

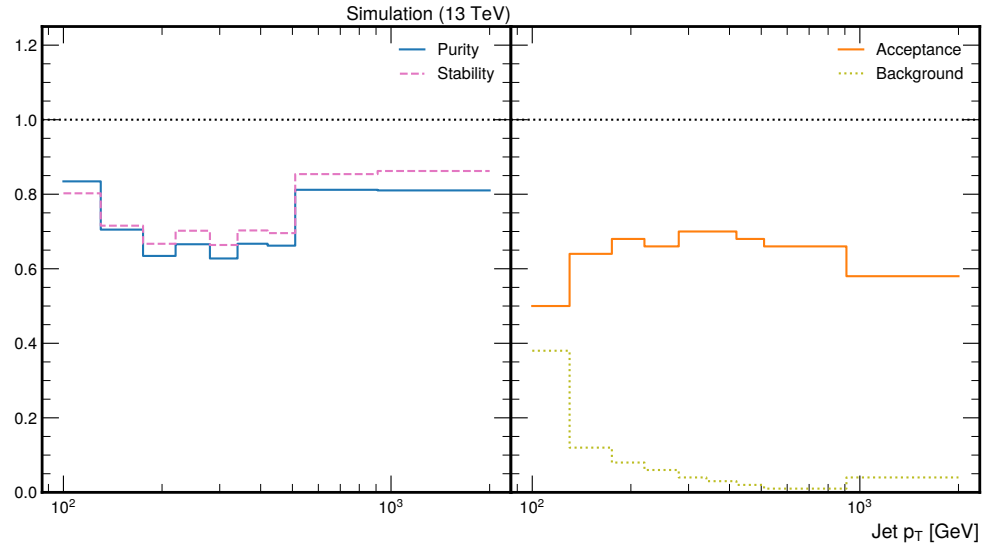

Figure 5.13: Purity and stability distributions (left) and acceptance and background distributions (right) in QCD events as a function of the jet transverse momentum. The results are obtained using the PYTHIA8 QCD simulation.

Migration study: Z boson transverse momentum

The migrations inside the phase space are small for the different leading jet p_T regions (see figure 5.15, upper panel). Overall, the purity and stability are above 80%. For the lowest leading jet p_T region the purity drops to 60% in the largest Z p_T region.

Migrations into the phase space are large in the lowest Z p_T regions for leading jet $p_T \in (100, 500)$ GeV, with background of 60% (see figure 5.15, lower panel). For the largest leading jet p_T region the migration into the phase space is small, with backgrounds smaller than 20%. We observe large migration outside the phase space for the largest leading jet p_T region in the lowest Z p_T. In this region the acceptance is 30%. For the rest of the regions the acceptance is around 60%.

Migration study: minimum ΔR between the Z boson and a jet

The purity and acceptance are above 90% in the whole $\Delta R_{Z,j}^{Min}$ for leading jet $p_T \in (100, 500)$ GeV (see figure 5.16, upper panel). For the largest leading jet p_T region the purity and stability show a different trend. This corresponds to the overlapping region between the jet and the muon. The purity and stability are around 80%. For larger $\Delta R_{Z,j}^{Min}$ the purity and stability are again above 90%. Thus, migrations inside

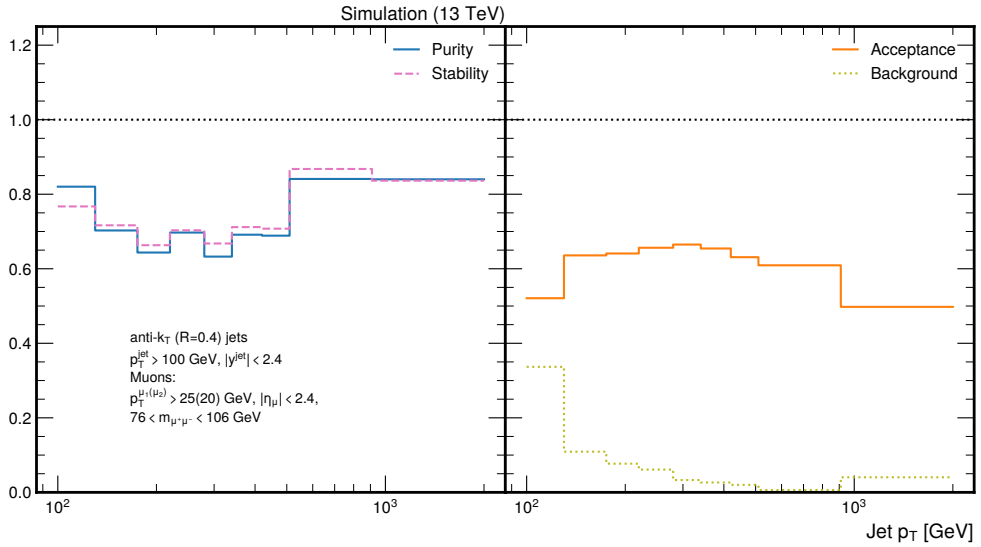

Figure 5.14: Purity and stability distributions (left) and acceptance and background distributions (right) in Z+jets events as a function of the jet transverse momentum. The results are obtained using the DY+jets MC@NLO calculation supplemented with the PYTHIA8 parton shower.

the phase space are small.

For increasing leading jet p_T the migrations outside the phase space are increasing in the lowest $\Delta R_{Z,j}^{Min}$ region (see figure 5.16, lower panel). For leading jet $p_T >$ 500 GeV the acceptance is of 20%. This is a sensitive region, here the jet and the muon are overlapping. For the rest of the regions the acceptance is around 40-60%. The migrations into the phase space are decreasing with increasing leading jet p_T. Overall, the background is smaller than 30%.

Migration study: $\Delta\phi$ between the two leading jets

Migrations inside the phase space are small for $\Delta\phi$ between the two leading jets. The purity and stability are above 90% (see figure 5.17, upper panel).

Migrations into the phase space are decreasing with increasing leading jet p_T. For leading jet $p_T \in (100, 500)$ GeV the background is in between 20-30%, while for leading jet $p_T > 500$ GeV is below 20% (see figure 5.17, lower panel). The acceptance for the different leading jet p_T regions is around 60%.

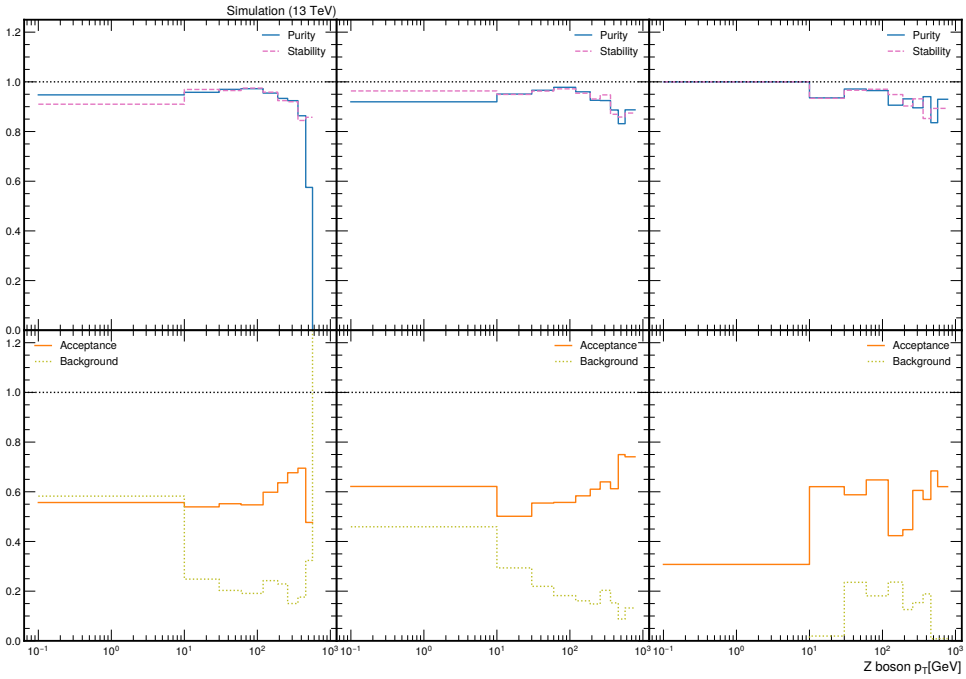

Figure 5.15: Purity and stability distributions (upper row) and acceptance and background distributions (lower row) in Z+jets events as a function of the Z boson transverse momentum, in bins of leading jet transverse momentum, $p_{\mathrm{T}}^{j_1}$: $100 < p_{\mathrm{T}}^{j_1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j_1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j_1} > 500$ GeV (right panel). The results are obtained using the DY+jets MC@NLO calculation supplemented with the PYTHIA8 parton shower.

Migration study: $\Delta\phi$ between the Z and the dijet

The purity and stability are above 60%, thus, migrations inside the phase space are considerably larger than for rest of observables (see figure 5.17, upper panel). This might be an effect of the ill defined ϕ angle of the dijet system. In our phase space the two leading jets are in a back-to-back configuration, $\Delta\phi_{j_1,j_2} \sim \pi$, hence, the ϕ angle of the dijet system is around 0. Thus, calculating the $\Delta\phi$ between the Z boson and the dijet system is sensitive to migrations inside the phase space.

The migrations into the phase space decrease for increasing leading jet transverse momentum (figure 5.17, lower panel). At low leading jet transverse momentum the background is around 30%, while for leading jet $p_T > 500$ GeV the background is

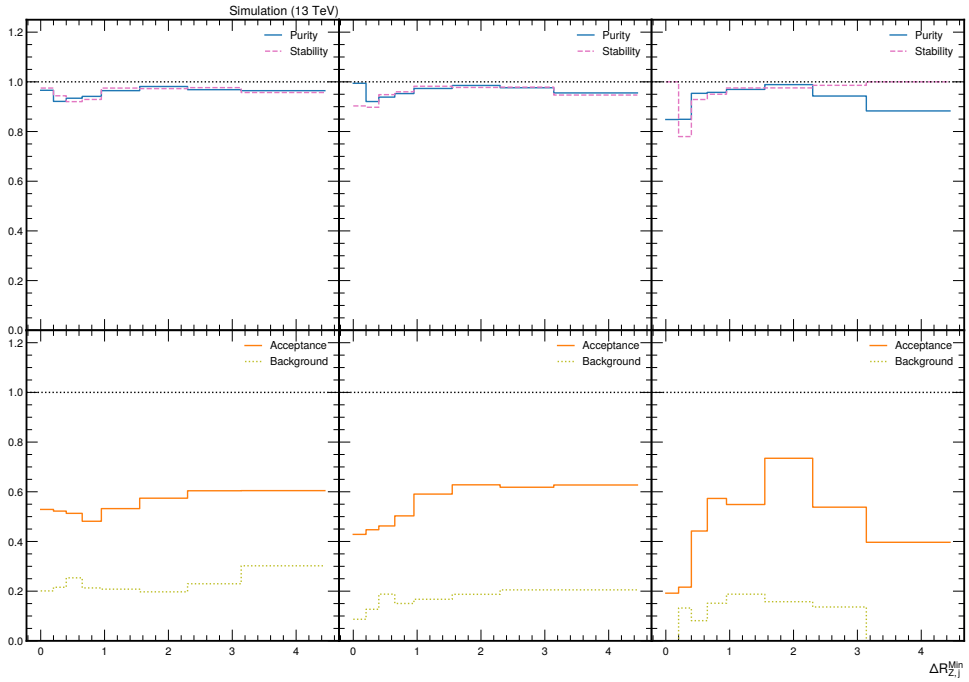

Figure 5.16: Purity and stability distributions (upper row) and acceptance and background distributions (lower row) in Z+jets events as a function of the minimum ΔR between the Z boson and a jet, in bins of leading jet transverse momentum, p_{T}^{j1}: $100 < p_{\mathrm{T}}^{j1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j1} > 500$ GeV (right panel). The results are obtained using the DY+jets MC@NLO calculation supplemented with the PYTHIA8 parton shower.

below 20%. Migrations outside the phase space are constant for all the different leading jet transverse momentum regions, where the acceptance is around 60%.

5.6.3 Unfolding

For the correction to particle level we use the ROOUNFOLD package [107]. The ROOUNFOLD package takes the migrations into account when unfolding. The migrations into the phase space and misidentified events are subtracted from the measured distribution. For the migrations outside the phase space, efficiencies are derived from the distribution of efficiency loss distribution. Each bin of the response matrix is scaled by these efficiencies. It is important to note that migrations inside the phase space are already taken into account in the response matrix. For Z+jets events we

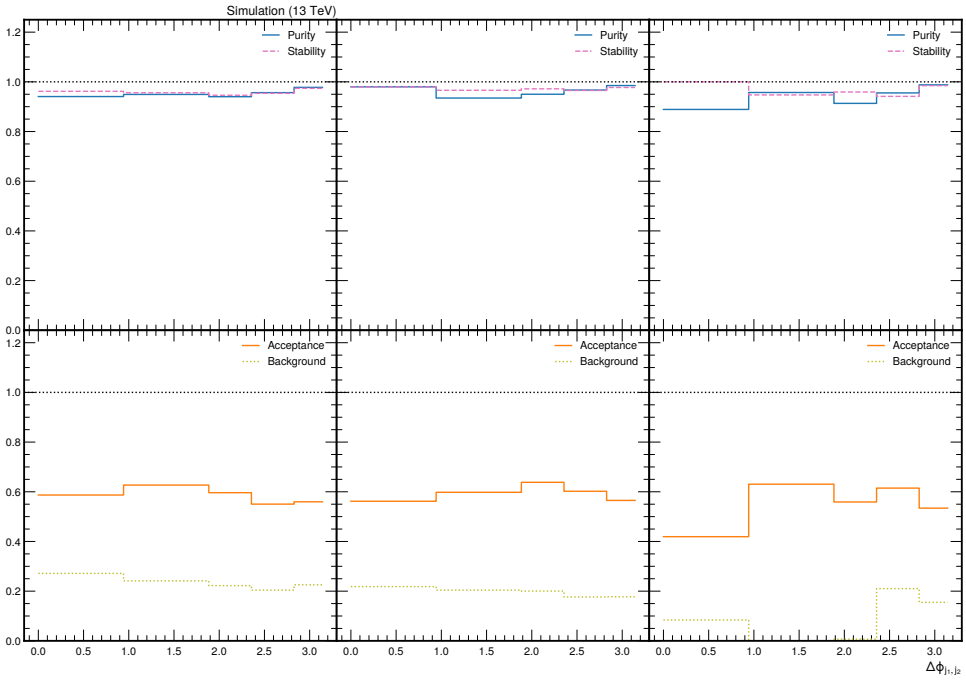

Figure 5.17: Purity and stability distributions (upper row) and acceptance and background distributions (lower row) in Z+jets events as a function of the $\Delta\phi$ between the two leading jets, in bins of leading jet transverse momentum, $p_{\mathrm{T}}^{j_1}$: $100 < p_{\mathrm{T}}^{j_1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j_1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j_1} > 500$ GeV (right panel). The results are obtained using the DY+jets MC@NLO calculation supplemented with the Pythia8 parton shower.

perform a simulated background subtraction to the measured distributions, as the backgrounds are well modelled by the simulations in Z+jets events [108].

We have seen that the condition number for our response matrices is below 5. In table 5.4 we summarise the condition numbers for all the observables. The condition numbers are small enough so that our precision in the measurement is not affected by numerical errors in the inversion of the probability matrix. Thus, it is safe to apply the matrix inversion method to correct the measured distributions to particle level. Another advantage of the matrix inversion over other methods, such as, the iterative Bayesian unfolding (D'Agostini method), is that we do not need any regularisation parameters. Which for the D'Agostini method is the number of iterations.

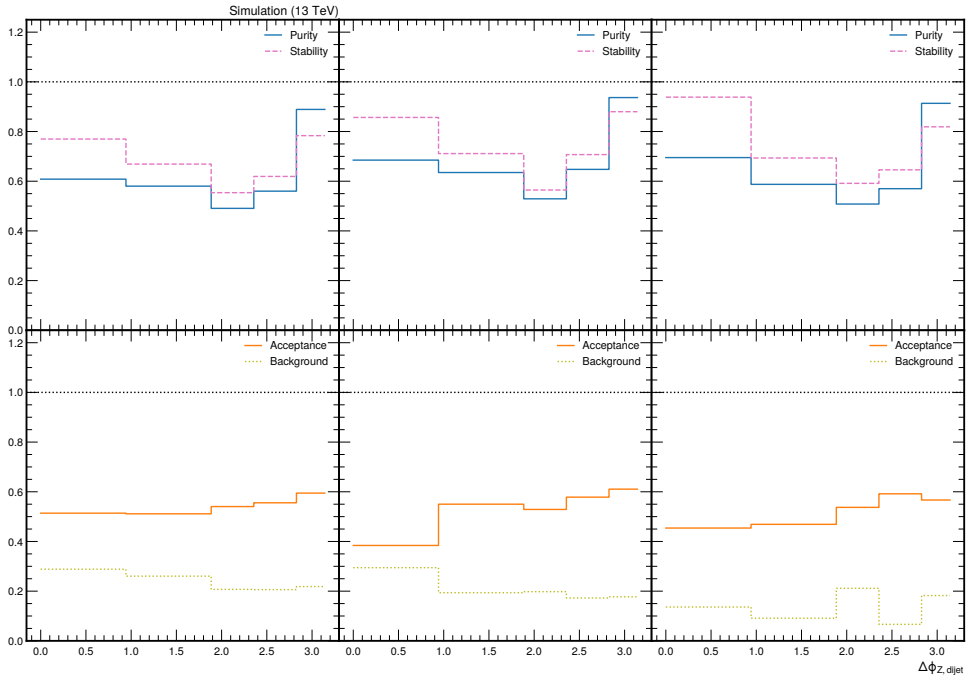

Figure 5.18: Purity and stability distributions (upper row) and acceptance and background distributions (lower row) in Z+jets events as a function of the $\Delta\phi$ between the Z and the dijet, in bins of leading jet transverse momentum, $p_{\mathrm{T}}^{j_1}$: $100 < p_{\mathrm{T}}^{j_1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j_1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j_1} > 500$ GeV (right panel). The results are obtained using the DY+jets MC@NLO calculation supplemented with the Pythia8 parton shower.

5.7 Uncertainties

In this section we present the systematic uncertainties that affect our measurement, together with the theoretical uncertainties of the theoretical predictions.

5.7.1 Systematic uncertainties

In the event reconstruction (section 4.3) we applied several corrections to overcome detector inefficiencies. Each correction that we apply introduces an uncertainty to the measurement, a systematic uncertainty. In this section we present the impact of each systematic uncertainty after the correction to particle level. In this analysis we consider the following sources of systematic uncertainties: luminosity, pileup,

trigger efficiencies, muon reconstruction efficiencies, jet energy scale and resolution. Overall we observe that theses systematic uncertainties have the same impact for all of the observables. In figure 5.19 we show the impact of the uncertainties in the jet transverse momentum spectrum for QCD and Z+jets events. In appendix C we display the uncertainties for the rest of observables.

Luminosity

The total event yield in simulation needs to be normalised to the recorded luminosity of the measured data. Moreover, it is essential for the determination of the cross-section. The uncertainty of the recorded luminosity in 2016 is 1.2% [109]. We scale up and down the distributions of the simulations taking into account the uncertainty of the recorded luminosity.

Pileup

The determination of the pileup profile is also attached to uncertainties. The largest uncertainty is the total proton-proton inelastic cross-section (see equation 4.29). We derive the pileup profile for data for the up and down variation of the inelastic cross-section (see figure 4.8). From there the up and down variations are extracted for the pileup reweighting. We apply the up and down variations to determine the impact of the pileup reweighting to our measurement.

Overall, the pileup uncertainty does not have a large impact in our measurement. At large transverse momentum, as our measurement, pileup processes have small contributions. However, we find observables in our measurement that are sensitive to pileup. The contribution of pileup at low ΔR_{Zj}^{Min} is around 8% (see figure C.4).

Trigger efficiencies

For QCD events the triggers are used only when they are fully efficient. Thus, we do not take the trigger efficiency as a systematic uncertainty in our measurement. For Z+jets events, the single muon triggers are not fully efficient. The efficiency of the single muon trigger is derived centrally in CMS by the Muon Physics Object Group together with their statistical and systematic uncertainties.

We apply the trigger efficiency as a scale factor to the simulated samples. Thus, to derive the systematic uncertainty in our measurement we vary the scale factor up and down with its corresponding systematic and statistical uncertainties. Due to the high efficiency of the single muon triggers and their small uncertainties, the impact of the single muon trigger efficiencies in our measurement are below 2%.

Muon related uncertainties

The application of muon isolation and identification criteria leads to inefficiencies and hence, they introduce an uncertainty. The uncertainties of the identification criteria are computed centrally in CMS by the Muon Physics Object Group.

In our measurement we combine the systematic and statistical uncertainty of the muon isolation and identification in quadrature, which we refer as MuIDISO uncertainty. We vary the muon related scale factors up and down to derive the impact in our measurement. The MuIDISO uncertainty is below 2% for all our observables in Z+jets events. Even though the muon isolation criteria are not fully efficient, their efficiency is close to one, having a small impact on the measurement.

The muons are also affected by the muon momentum scale correction, or "Rochester corrections". However, the uncertainties of this corrections are below 1%. We expect a negligible contribution to our measurement. Thus, we do not take the muon momentum scale as a systematic uncertainty in the measurement.

Jet energy scale and resolution

After the jet reconstruction we correct the jet energy with the jet energy scale factors. The jet energy scale factor is obtained by combining different corrections in a sequential fashion. Each step in the jet energy scale determination introduces different uncertainties. Thus, the jet energy scale itself has many sources of uncertainties. To determine the impact of the jet energy scale factor in our measurement we make use of its total uncertainty. We use the jet energy scale values and uncertainties which are centrally produced in the CMS JetMET group. We vary the jet energy scale up down with the total uncertainty. The jet energy scale introduces an uncertainty of a 5% uncertainty for our observables.

Correction to particle level

When we correct the measured distributions to particle level we find different sources of uncertainties: the model uncertainty and the uncertainty related to the limited statistical precision of the simulation sample.

The **model uncertainty** refers to the uncertainty of using a specific simulation to correct the measured distribution to particle level. In our case, to correct the Z+jets distributions to particle level we use the MC@NLO model supplemented with Pyhia8 parton shower. However, we can model our process with different simulations. Thus, we introduce our choice of model as an uncertainty.

Ideally, to determine the model uncertainty, we correct the measured distribution using two different models. The difference between the corrected data with the

different models is taken as the total model uncertainty. Unfortunately, for Z+jets events no other models are available with a full simulation of the detector within the CMS experiment. To obtain the model uncertainty in Z+jets events we modify the MC@NLO model event by event. We apply a weight to the event depending on the leading jet p_{T}. To determine the weight, first, we compute the difference of the leading jet p_{T} distribution between the measured data and the simulation at detector level.

For QCD events, we use the PYTHIA8 p_{T}-binned sample to correct to particle level. The limited statistical precision at high jet p_{T} by the available models is not sufficient to determine the model uncertainty for QCD events. Thus, we follow the same procedure as for Z+jets events to determine the model uncertainty, we reweight the QCD sample as a function of the leading jet transverse momentum. The weights are determined by computing the difference in the leading jet p_{T} distribution between QCD events and the QCD sample at detector level.

The **uncertainty related to the limited number of simulated events** arises from the limited number of events available in the simulated sample used for the particle level correction. We vary the distributions of the simulation up with the statistical uncertainty. Then, we unfold the measured data. The uncertainty related to the limited number of events in the simulation is one of the largest uncertainties in our measurement. In the largest p_{T} region, the uncertainty is around 8%. Moreover, for the observables measured in the $p_{\mathrm{T}}^{j1} > 500$ GeV range, the statistical precision is the dominant uncertainty.

5.7.2 Theoretical uncertainties

Theoretical predictions are based in factorisation theorems and they are composed by different factors or functions. Each function in the theoretical prediction introduces various uncertainties.

The hard-function in the cross-section is computed in perturbation theory. In the calculation we find ultraviolet divergences, that are absorbed and introduce a scale, the renormalisation scale, μ_{R}. On the other hand, the factorisation formula, introduces the factorisation scale, μ_{F}. The PDFs in the factorisation formula are determined by fits to data, which introduce experimental uncertainties. The strong coupling, which is used for the determination of the PDFs, as well as, for the calculation of the hard-function, needs to be measured, introducing an experimental uncertainty. The impact of these uncertainties are shown in section 5.8.

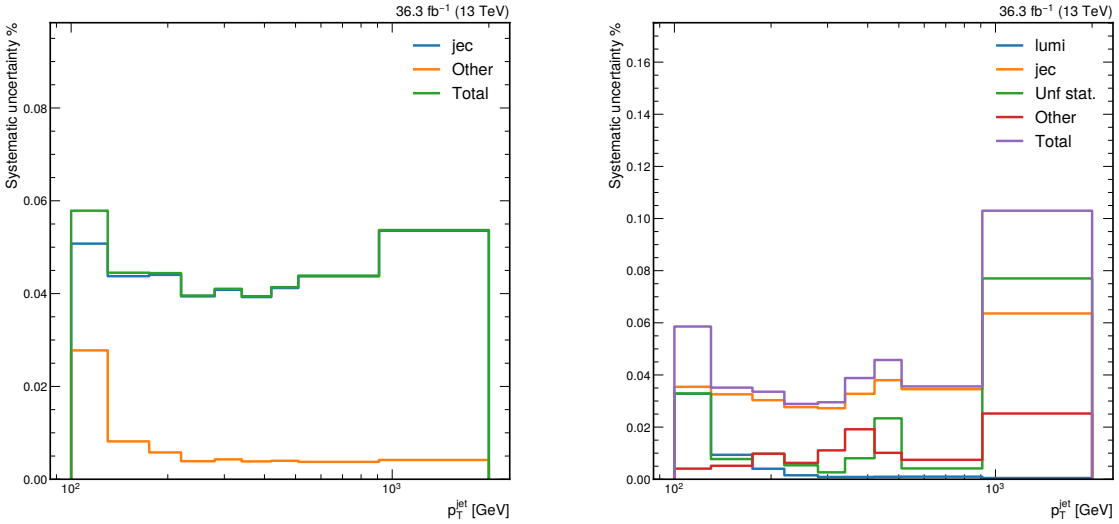

Figure 5.19: Impact of systematic uncertainties in the cross-section as a function of the jet transverse momentum. Systematic uncertainties below 2% are grouped in the category Other. Left panel: QCD events, where Other includes JER, luminosity, pileup and unfolding related uncertainties. Right panel: Z+jets events, where Other includes JER, muon related uncertainties, model uncertainties and trigger uncertainties.

Variations	1	2	3	4	5
$\mu_{\mathbf{R}}$	$\frac{1}{2}\mu_{\mathrm{R}}$	μ_{R}	$\frac{1}{2}\mu_{\mathrm{R}}$	μ_{R}	$2\mu_{\mathrm{R}}$
$\mu_{\mathbf{F}}$	$\frac{1}{2}\mu_{\mathrm{F}}$	$\frac{1}{2}\mu_{\mathrm{F}}$	μ_{F}	$2\mu_{\mathrm{F}}$	μ_{F}

Table 5.5: Variations of renormalisation, μ_{R}, and factorisation, μ_{F}, scales to estimate the missing contribution of higher order corrections.

Scale uncertainty

The scale uncertainty is composed by two scales: the renormalisation and the factorisation scales. The factorisation scale, μ_{F}, is introduced by the factorisation of the cross-section. This scale is set to divide the soft and hard contributions to the cross-section. Processes below μ_{F} are treated in the PDFs, while processes above μ_{F} belong to the hard-function. In this way, infrared divergences are removed from the calculation of the hard-function. Thus, μ_{F} is used as a cut-off for infrared divergences. The renormalisation scale arises from computing the hard-function. Virtual corrections can diverge for large momentums, these divergences are known as ultraviolet divergences. To absorb the divergence a renormalisation scale is introduced.

The theoretical calculation should not depend on the arbitrary parameters μ_{F} and μ_{R}. However, the hard-function is perturbatively expanded in α_{S} and calculated up to a given order in α_{S}. The truncation of the expansion makes the hard-function depend on μ_{F} and μ_{R}. The variation of the two scales can be used to estimate the impact of the missing higher orders. Given the arbitrariness of the two scales, for the calculations $\mu_{\mathrm{F}} = \mu_{\mathrm{R}} = Q$ is used, where Q is the hard scale of the process.

To estimate the contribution of the missing higher order corrections in the calculation, we do seven simultaneous variations of μ_{F} and μ_{R} (see table 5.5). We make an envelope of the variations to estimate the scale uncertainty.

$\alpha_{\mathbf{S}}$ uncertainty

The strong coupling, α_{S}, both used in the determination of the PDF and the hard-function, is calculated using the β function, the running of the coupling. However, an initial value of α_{S} is needed. The initial value needs to be extracted from measurements at a given scale. Usually, α_{S} is measured at the scale of the mass of the Z boson, which corresponds to a value of $\alpha_{\mathrm{S}}(m_Z) = 0.118$. To estimate the impact of α_{S}, $\alpha_{\mathrm{S}}(m_Z)$ is varied 0.001 up and down in the theoretical calculation.

PDF uncertainty

The initial condition of the PDF cannot be calculated from first principles, it is parametrised and then fitted to data. This parametrisation introduces an uncer-

tainty in the PDF determination. The measurements used to fit the PDF are affected by systematic and statistical uncertainties. The uncertainties from the fitted data are the largest contribution to the PDF uncertainty. After the determination of the initial condition, PDFs are evolved using perturbative calculations, which depend on α_S and the scale μ.

The uncertainties introduced by the determination of α_S and μ_F are not included in the PDF uncertainty. In the PDF uncertainty we only consider the uncertainties that affect the initial state condition, which are the uncertainties of the measurements used to fit the PDF. The NNPDF3.1 PDF set includes the variation of the uncertainties, which we use to reweight the simulation up and down to derive the impact on the theoretical prediction.

5.8 Results

In this section we present the differential cross-sections corrected to particle level corresponding to the fiducial phase space sensitive to real electroweak radiations defined in table 5.6. We compare them to different theoretical predictions. For the matrix element calculation we make use of MadGraph5_aMC@NLO and we interface it with different Monte-Carlo event generators. In all the calculations the renormalisation and factorsation scales are set to $\mu_R = \mu_F = 1/2 \sum_i H_{T,i}$, where $H_{T,i}$ is the scalar sum of the transverse momentum of all the final state particles in the matrix element calculation. We label the MadGraph5_aMC@NLO calculations as MG5_aMC. We compare the measurements with the following calculations:

- MG5_aMC+Py8 (jj): QCD calculation of 2 jet final state at next-to leading order accuracy, supplemented with the Pythia8 parton shower. For the calculation the NNPDF3.1 PDF at next-to-next-to leading order accuracy in QCD is used with $\alpha_S(m_Z) = 0.118$, where m_Z is the mass of the Z boson. We make use of this simulation to compare the cross section as a function of the jet transverse momentum in QCD events.

- MG5_aMC+CA3 (jj): QCD calculation of 2 jet final at next-to-leading order accuracy, supplemented with the PB-NLO-HERAI+II-Set2 integrated TMD at next-to-leading order accuracy in QCD with $\alpha_S(m_Z) = 0.118$. The transverse momentum of the initial state evolution is included with the corresponding TMD with CASCADE3, where the initial state parton shower follows the TMD. The final state shower and hadronisation are treated with Pythia6. No multi parton interactions are considered in this calculation.

- MG5_aMC+Py8 (Z+0+1+2j): DY+0,1,2 jet calculation at next-to-leading order accuracy in QCD supplemented with the Pythia8 parton shower and

multi parton interactions. The tuning of the underlying event corresponding to the CUETP8M1 is applied. The NNPDF3.0 PDFs at next-to-next-to leading order in QCD accuracy are used with $\alpha_S(m_Z) = 0.118$. We make use of this simulation to compare the cross section as a function of the jet transverse momentum in Z+jets events, estimate the effect of multi parton interactions and to test the modelling of real electroweak corrections with a DY+0,1,2 jet calculation.

- MG5_aMC+CA3 ($Z + 2j$): DY+2 jet calculation at next-to-leading order accuracy in QCD, supplemented with the PB-NLO-HERAI+II-Set2 integrated TMD at next-to-leading order accuracy in QCD with $\alpha_S(m_Z) = 0.118$. The transverse momentum of the initial state evolution is included with the corresponding TMD with CASCADE3, where the initial state shower follows the TMD. The final state shower and hadronisation are treated with Pythia6. No multi parton interactions are considered in this calculation. We make use of this calculation to account for TMD effects and to test the modelling of real electroweak corrections with a DY+2 jets calculation.

- MG5_aMC+Py8 (jj)+EWshower: QCD calculation of 2 jet final state at next-to leading order accuracy, supplemented with the Pythia8 parton shower and the Pythia8 simple electroweak shower. For the calculation the NNPDF3.1 PDF at next-to-next-to leading order accuracy in QCD is used with $\alpha_S(m_Z) = 0.118$, where m_Z is the mass of the Z boson. We make use of this simulation to compare the effect of treating the real electroweak boson emission in the parton shower and not in the matrix element.

The theoretical uncertainties corresponding to the calculations include the same uncertainties in section 5.7.2 and are represented with a band around the central value of the calculation. The systematic uncertainties of the measurement cross-section are labelled as correlated and the uncertainties due the limited statistical precision of the measured data as uncorrelated.

In figure 5.20 we show the cross-section as a function of the jet transverse momentum for QCD (left panel) and Z+jets (central and right panels) events. In QCD events we are dominated by systematic uncertainties. In Z+jets events for $p_T^j < 1$ TeV the measurement is dominated by systematic uncertainties, and for $p_T^j > 1$ TeV it is limited due to the limited amount of the measured events. In QCD events (left panel), the MG5_aMC+Py8 (jj) calculation agrees with the measurement for most of the jet transverse momentum, however, at large jet transverse momentum underestimates the measurement. The MG5_aMC+CA3 (jj) calculation on the other hand, has a 20-30% normalisation issue, which was observed already in other jet measurements [110]. In Z+jets events (central and right panels), the MG5_aMC+Py8 ($Z + 0 + 1 + 2j$) calculation is in agree-

ment with the measurement in the whole jet transverse momentum range. The MG5_aMC+CA3($Z+2j$) calculation agrees with the measurement within the uncertainties, while the MG5_aMC+Py8 (jj)+EWshower shows a disagreement of 30-40% at large transverse momentum.

The double differential cross-sections in figures 5.21-5.24 are dominated by systematic uncertainties at low leading jet transverse momentum $p_{\mathrm{T}}^{j_1} < 500$ GeV. However, in the highest leading jet transverse momentum bin, $p_{\mathrm{T}}^{j_1} > 500$ GeV, the measured number of Z+jets events are the limiting factor. The distributions are overall well described by the MG5_aMC+Py8 (Z 0+1+2j) and MG5_aMC+CA3 (Z+2j) predictions. We observe that MG5_aMC+Py8 (Z 0+1+2j) has a better agreement with the measured data, however, the differences between the two calculations is not large. It has been checked that multi parton interactions do not play any role in the phase space of the measurement (checked with MG5_aMC+Py8 switching the multi parton interactions on/off). In the phase space of the measurement, at large jet transverse momentum, we are not sensitive to multi parton interaction and TMD effects. The MG5_aMC+Py8 (jj)+EWshower calculation however, fails to describe most of the measured observables. It is not able to describe the shapes of the measured distribution, and underestimates the measured data.

The ϕ angle between the two leading jets, $\Delta\phi_{j_1,j_2}$, (see figure 5.22) increases with increasing jet transverse momentum. At high transverse momentum is more likely to find a back-to-back configuration of the two leading jets, which points to the enhancement of the probability of the Z boson radiation.

In the ϕ angle between the Z boson and the dijet, $\Delta\phi_{Z,j}$, (see figure 5.23) we observe the opposite behaviour, the decorrelation between in $\Delta\phi_{Z,j}$ increases with increasing jet transverse momentum. With higher jet transverse momentum the probability of the dijet system to be in a back-to-back configuration increases. Thus, the ϕ angle of the dijet system is likely to be around zero. The Z boson is radiated from either of the jets or from the initial state, which leads to a random ϕ angle of the Z boson and to a decorrelation in the $\Delta\phi_{Z,j}$.

It is also interesting the behaviour of the minimum ΔR distance between the Z boson and a jet, $\Delta R_{Z,j}^{Min}$, (see figure 5.24). With increasing jet transverse momentum the minimum distance between the Z boson and a jet decreases. We should expect that at higher energies the Z boson could be reconstructed as a part of the jet with larger jet radius in the jet reconstruction, and thus, affecting the jet substructure, where three subjets could be observed in a single jet: the parent quarks and the two decay products of the Z boson.

It is worth noting that at high large jet transverse momentum Z+jets events Z boson candidates are reconstructed in the QCD predictions (MG5_aMC+Py8 (jj)

Fiducial phase space	Z+jets events	QCD events
Muons	$p_{\mathrm{T}}^{\mu_1(\mu_2)} > 25(20)$ GeV, $\|\eta\| < 2.4$	-
Jets	$p_{\mathrm{T}} > 100$ GeV, $\|y\| < 2.5$	$p_{\mathrm{T}} > 100$ GeV, $\|y\| < 2.5$

Table 5.6: Fiducial phase space of the measured cross-section for QCD and Z+jets events

and MG5_aMC+CA3 (jj)). At high jet transverse momentum the probability of producing a $b\bar{b}$ quark pair in QCD processes increases. The corresponding hadrons of the b quarks, B, can then decay semileptonically,

$$B \to \mu + \nu_\mu + X, \tag{5.17}$$

producing an opposite charged muon pair, which are reconstructed as a Z boson. This background has never been considered in Z+jets measurements in CMS. For future measurements of Z boson production in association with high transverse momentum jets the contribution of QCD processes should be studied. In appendix D we show the Z+jets distributions affected by the semileptonic decays of the B $\bar{B}$ pairs.

In summary, we find that a DY+0,1,2 jets at next-to-leading order accuracy in α_{S} are able to describe regions of the phase space were we are sensitive to real electroweak radiations.This finding is supporting the approach taken in [64] to study real electroweak radiations through the calculation of V+1,2 jets processes, where $V = Z, W^{\pm}$, to estimate the contribution of real electroweak radiations in QCD events. Besides, we find that the Pythia8 electroweak shower is not able to describe the measured distributions, which could point out to the missmodelling of the electroweak shower. A comparison with other electroweak showers, Antenna [111] and Herwig7.3 [112], would let us study the consistency of these treatments.

5.9 Contribution of real electroweak radiations

In this section we extract the contribution of real electroweak corrections in multijet events and we discuss the implication of real electroweak corrections at the LHC.

In figure 5.25 we extract the contribution of the radiation of the Z boson to QCD events as a function of the jet transverse momentum. First, we measure the contribution of the $Z \to \mu^+\mu^-$ contribution. We divide the cross-section as a function of the jet transverse momentum of Z+jets events, by the cross-section as function of the jet transverse momentum in QCD events:

$$\delta^{QCD}_{Z\to\mu^+\mu^-}(p_T^j) = \frac{d\sigma^{Z+jets}/dp_T^j}{d\sigma^{QCD}/dp_T^j}, \tag{5.18}$$

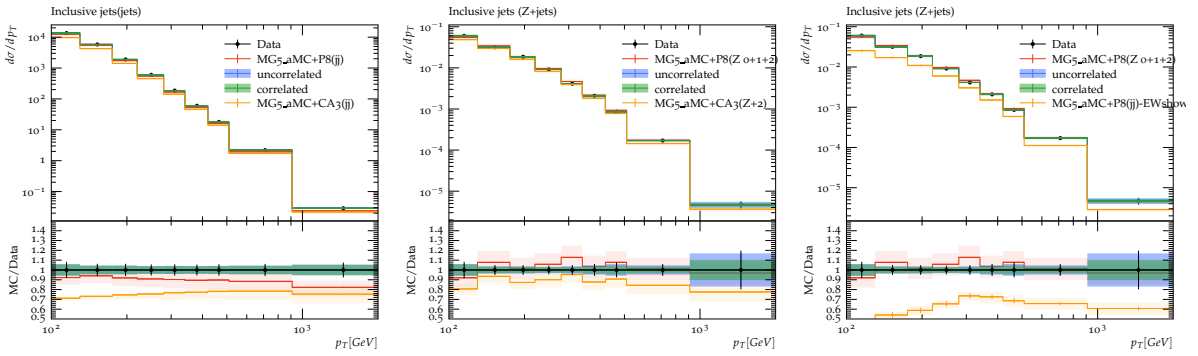

Figure 5.20: Cross-section as a function of the jet transverse momentum for QCD (left panel) and Z+jets events (central and right panel). Systematic (correlated) and statistical (uncorrelated) uncertainties are shown in bands around the measured data. The total uncertainty is given by the errorbars. We compared the measurement to different theoretical predictions. Left panel: MG5_aMC+Py8 (jj) and MG5_aMC+CA3 (jj). Central panel: MG5_aMC+Py8$(Z+0+1+2j)$ and MG5_aMC+Py8 $(Z+0+1+2j)$. Right panel: MG5_aMC+Py8$(Z+0+1+2j)$ and MG5_aMC+Py8(jj) supplemented with electroweak shower (EWshower).

where we define $\delta_{Z\to\mu^+\mu^-}(p_T^j)$ as the contribution of the $Z \to \mu^+\mu^-$ channel to QCD events as a function of the jet transverse momentum, p_T (see figure 5.25). We can observe that the contribution is negligible at low jet transverse momentum but increases rapidly with the increasing jet transverse momentum. Let us consider the $p_T^j \sim 1$ TeV, the bin with a larger contribution, where the contribution of the $Z \to \mu^+\mu^-$ channel is

$$\delta^{QCD}_{Z\to\mu^+\mu^-}(1\text{ TeV}) = (1.63 \pm 0.13(\text{syst.}) \pm 0.21(\text{stat.})) \times 10^{-4} \tag{5.19}$$

where we consider systematic (syst.) and statistical precision (stat.) uncertainties. The $Z \to \mu^+\mu^-$ branching fraction amounts to the 3.36% of the total branching of the Z boson. In table 5.7 we show the Z boson decay modes with their corresponding decay ratio. We extrapolate the total contribution of the Z boson to QCD events:

$$\delta^{QCD}_{Z}(1\text{ TeV}) = 0.0045 \pm 0.0006(\text{stat.}), \tag{5.20}$$

where we only consider the statistical uncertainty, as we expect the leptonic decay channel to have the smallest statistical precision. In this simple extrapolation we do not consider the systematic uncertainties, as the different decay channels suffer from different systematic uncertainties.

We perform a rough estimation of the total contribution of real electroweak corrections by including the W boson. The coupling strength of the W boson is larger

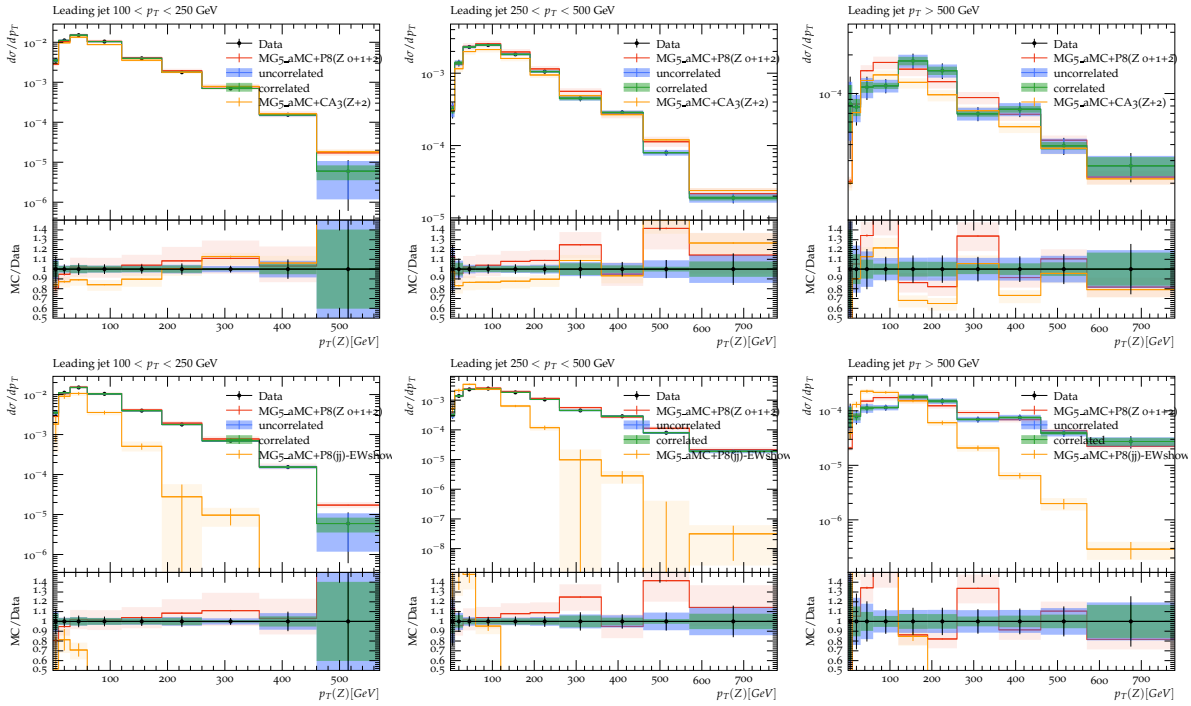

Figure 5.21: Cross-section as a function of the Z boson transverse momentum in Z+jets events in bins of leading jet transverse momentum, p_{T}^{j1}: $100 < p_{\mathrm{T}}^{j1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j1} > 500$ GeV (right panel). Systematic (correlated) and statistical (uncorrelated) uncertainties are shown in bands around the measured data. The total uncertainty is given by the errorbars. We compare the result to different theoretical predictions. Upper row: MG5_aMC+Py8($Z + 0 + 1 + 2j$) and MG5_aMC+Py8($Z + 0 + 1 + 2j$). Lower row: MG5_aMC+Py8($Z + 0 + 1 + 2j$) and MG5_aMC+Py8(jj) supplemented with electroweak shower (EWshower).

Decay mode	**Branching fraction (%)**
e^+e^-	3.3632 ± 0.0042
$\mu^+\mu^-$	3.3662 ± 0.0066
$\tau^+\tau^-$	3.3696 ± 0.0038
Invisible ($\nu\nu$)	20.000 ± 0.055
Hadrons ($q\bar{q}$)	69.911 ± 0.056

Table 5.7: Z boson decay modes with their corresponding branching ratio [113]

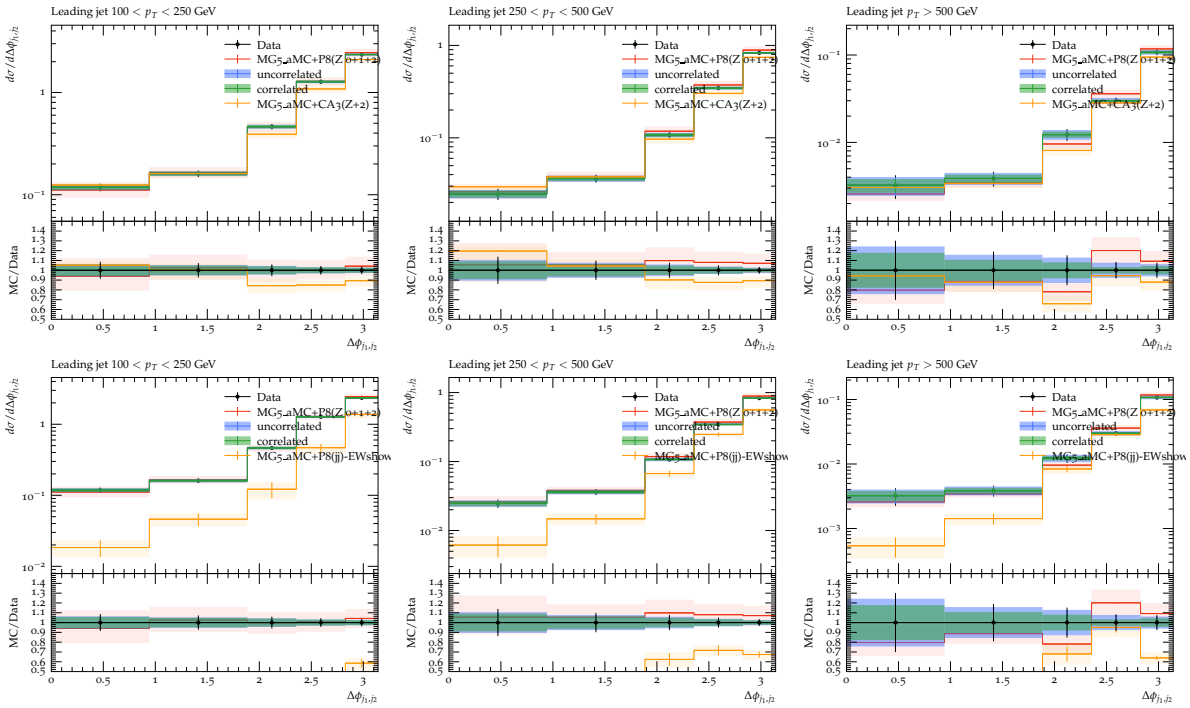

Figure 5.22: Cross-section as a function of the $\Delta\phi_{j_1,j_2}$ in Z+jets events in bins of leading jet transverse momentum, $p_{\mathrm{T}}^{j_1}$: $100 < p_{\mathrm{T}}^{j_1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j_1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j_1} > 500$ GeV (right panel). Systematic (correlated) and statistical (uncorrelated) uncertainties are shown in bands around the measured data. The total uncertainty is given by the errorbars. We compare the result to different theoretical predictions. Upper row: MG5_aMC+Py8($Z + 0 + 1 + 2j$) and MG5_aMC+Py8($Z + 0 + 1 + 2j$). Lower row: MG5_aMC+Py8($Z + 0 + 1 + 2j$) and MG5_aMC+Py8(jj) supplemented with electroweak shower (EWshower).

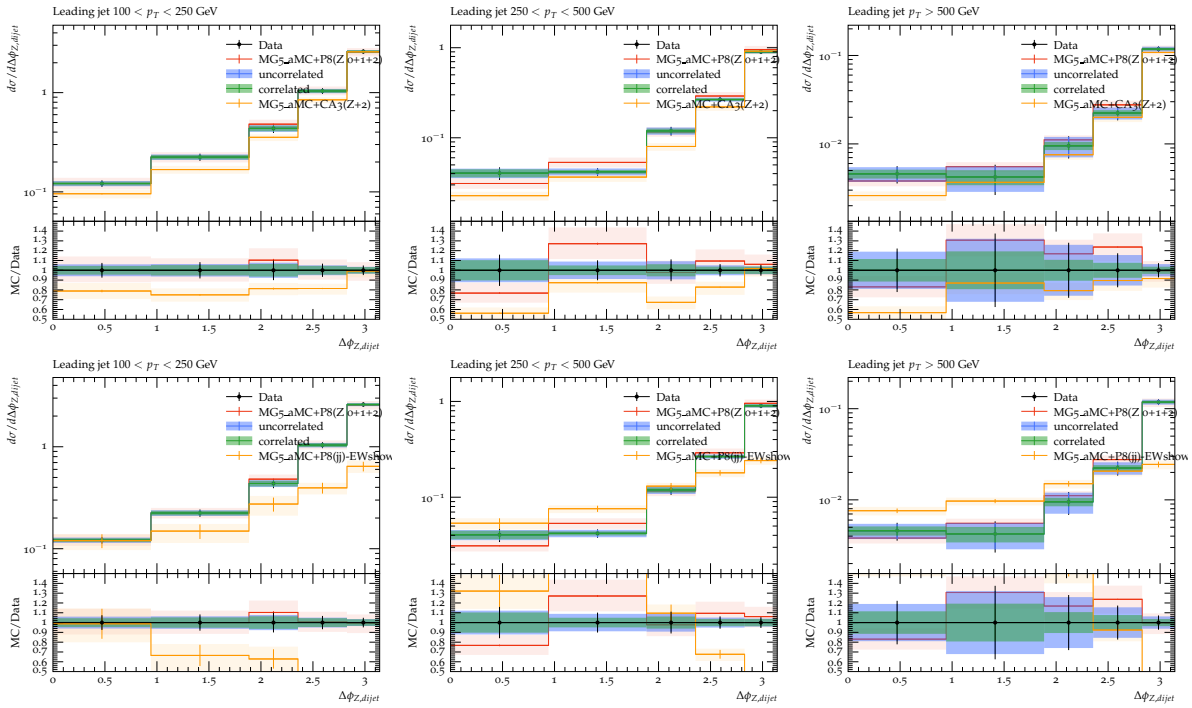

Figure 5.23: Cross-section as a function of the $\Delta\phi_{Z,dijet}$ in Z+jets events in bins of leading jet transverse momentum, p_{T}^{j1}: $100 < p_{\mathrm{T}}^{j1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j1} > 500$ GeV (right panel). Systematic (correlated) and statistical (uncorrelated) uncertainties are shown in bands around the measured data. The total uncertainty is given by the errorbars. We compare the result to different theoretical predictions. Upper row: MG5_aMC+Py8($Z + 0 + 1 + 2j$) and MG5_aMC+Py8($Z + 0 + 1 + 2j$). Lower row: MG5_aMC+Py8($Z + 0 + 1 + 2j$) and MG5_aMC+Py8(jj) supplemented with electroweak shower (EWshower).

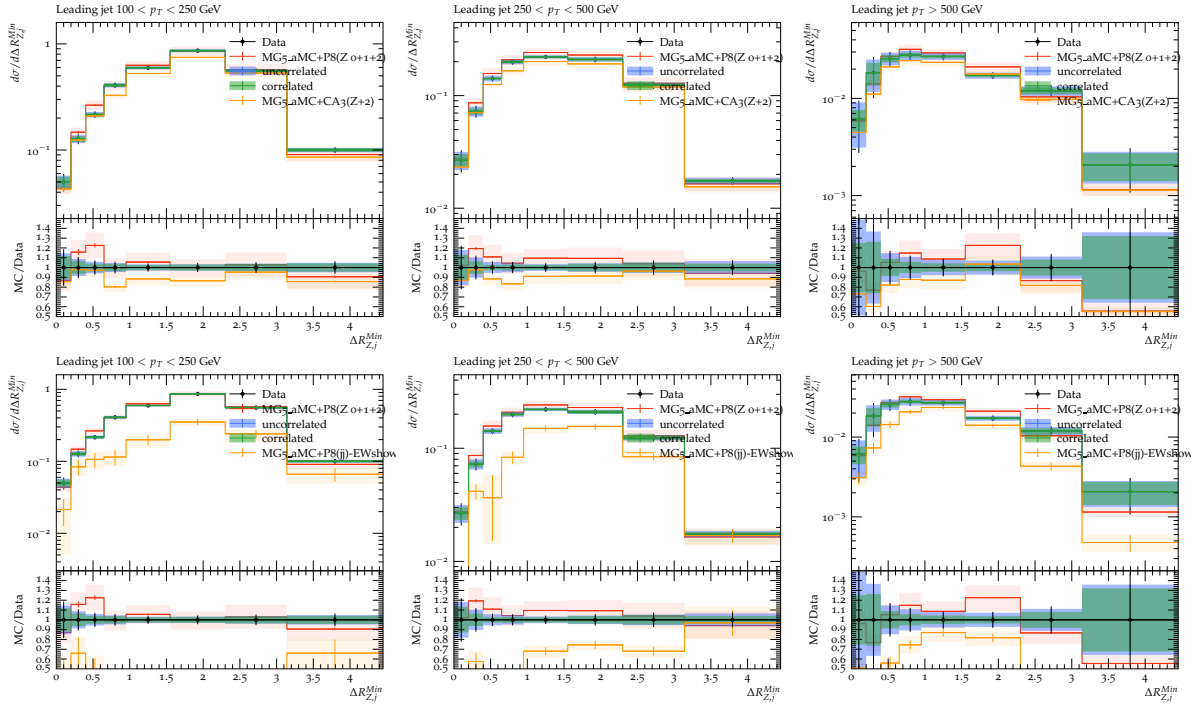

Figure 5.24: Cross-section as a function of the $\Delta R^{Min}_{Z,j}$ in Z+jets events in bins of leading jet transverse momentum, $p_{\mathrm{T}}^{j_1}$: $100 < p_{\mathrm{T}}^{j_1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j_1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j_1} > 500$ GeV (right panel). Systematic (correlated) and statistical (uncorrelated) uncertainties are shown in bands around the measured data. The total uncertainty is given by the errorbars. We compare the result to different theoretical predictions. Upper row: MG5_aMC+Pythia8($Z + 0 + 1 + 2j$) and MG5_aMC+Pythia8($Z + 0 + 1 + 2j$). Lower row: MG5_aMC+Pythia8($Z + 0 + 1 + 2j$) and MG5_aMC+Pythia8(jj) supplemented with electroweak shower (EWshower).

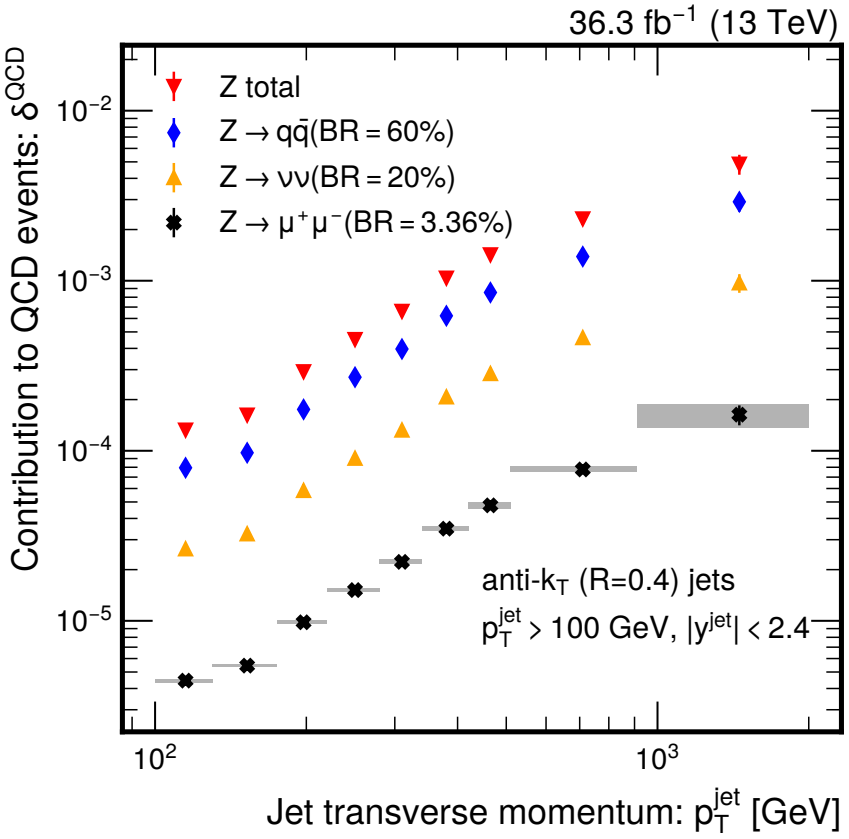

Figure 5.25: Contribution of Z+jets events to QCD events as a function of the jet transverse momentum. The contribution of the $Z \to \mu^+\mu^-$ and the extrapolation of the total branching fraction of the Z boson (Z total) are shown.

to the one of the Z boson, and moreover its mass is smaller. Thus, the radiation probability is enhanced for the W boson (see equation 5.1). We can estimate the contribution of real electroweak corrections to QCD events, $\delta^{QCD}_{\mathrm{EW}^{(R)}}$, at jet transverse momentum of 1 TeV to be

$$\delta^{QCD}_{\mathrm{EW}^{(R)}}(1\ \mathrm{TeV}) = 1 - 2\%. \tag{5.21}$$

Our estimation coincides with the theoretical calculation in [64], where real electroweak radiations are modelled with $V + 1, 2$ jet calculations, where $V = W^{\pm}, Z$.

In our QCD cross-section measurement as a function of the jet transverse momentum, at high jet transverse momentum the jet energy calibration is the largest uncertainty with an impact of 5% (see figure 5.19, left panel). In consequence the contribution of real electroweak radiations can make a significant contribution the measured cross-sections. In previous measurements [114] these contributions are controlled by restricting the amount of missing transverse momentum in the event, which removes the contribution of semi-leptonic decays of the W boson ($W \to l\nu$) and the $Z \to \nu\nu$ decay channel. From the calculation [64] such cuts reduce the contribution of real electroweak radiations. To control the electroweak production of dijet events ($q\bar{q} \to V \to q\bar{q}$, with $V = W^{\pm}, Z$) the calculations in [58] are used as correction factors, as well as, for virtual electroweak radiations. However, real electroweak radiations where the Z boson decays leptonic and hadronically and the $W\pm$ decays hadronically are not taken into account. The contribution of the hadronic decay channel for the Z boson at high jet transverse momentum to QCD events amounts to

$$\delta^{QCD}_{Z\to q\bar{q}}(1\ \mathrm{TeV}) \sim 0.3\%. \tag{5.22}$$

The branching ratio of the hadronic decay channel of the W boson is $(67.41 \pm 0.27)\%$ [113], 7% larger than the one of the Z boson. Thus, on top of the corrections applied in [114], real electroweak corrections could contribute 1% to QCD events at jet transverse momentum of the order of 1 TeV.

CHAPTER 6

Conclusions

The production of the Z boson is a key process both for experimental measurements and theoretical predictions. The leptonic decay offers a high precision reconstruction and measurement of the Z boson. It is present in many aspects of high energy physics: in the calibration of the energy of the measured jets, as a main background in many physiscs analysis, e.g., searches for new physiscs and top qaurk production, in studies of the proton structure, etc. Moreover, its transverse momentum is an important element for the measurement of the W boson mass. Hence, a good understanding of the Z boson is crucial for many aspects in high energy physics. We have investigated two production mechanism of the Z boson: the Drell-Yan (DY) process and real electroweak radiations.

The transverse momentum of the DY lepton pair, q_T, is of interest for the hadron structure. At low DY transverse momentum it is sensitive to the 3D structure of the hadron. To describe the transverse momentum of the DY lepton pair resummation to all orders is needed, where we have presented three methods: transverse momentum dependent factorisations (TMD factorisation, high energy factorisation), parton showers and the Parton Branching method, a hybrid method between the two previous approaches. However, the $q_T \ll m_{DY}$ region is non-perturbative and it has to be modelled, where m_{DY} is the mass of the DY lepton pair.

In the TMD factorisation the b_* prescription is introduced to separate the non-perturbative from the perturbative region, where b is the Fourier transformation of the transverse momentum. This separation gives rise to two non-perturbative functions: the intrinsic transverse momentum of incoming partons and the non-perturbative region of the rapidity evolution. The rapidity evolution is commonly known as the Collins Soper (CS) Kernel. These two functions are then fitted to data.

In the PB method only the intrinsic transverse momentum of the incoming partons is modelled. We have shown that the PB method offers a natural explanation to the non-perturbative region of the CS kernel. In the PB method we are able to reproduce the CS kernel with very soft emissions contained in the $z \to 1$ region of the momentum transfer. Moreover, together with an angular ordering condition and $\alpha_{\mathrm{S}}(p_{\mathrm{T}})$, where p_{T} is the transverse momentum of the emission, the PB method is able to reproduce the $q_T \ll m_{DY}$ region for measurement with different DY masses and different centre-of-mass energies, with an intrinsic transverse momentum independent from the DY mass and the centre-of-mass energy [33].

In the parton shower of Monte-Carlo event generators (PYTHIA8 and HERWIG7) the very soft gluon radiations corresponding to the $z \to 1$ region are removed by introducing a dynamical resolution scale. Thus, the $q_T \ll m_{DY}$ region is only modelled by the intrinsic transverse momentum. Our studies indicate that neglecting very soft radiation leads to a centre-of-mass energy dependence of the intrinsic transverse momentum.

We see that understanding the DY processes is important to understand the hadron structure and the impact of mismodelling the parton shower in SM processes. Half a century after the proposal of the DY process by S. D. Drell and T. -M. Yan, we can still improve our theoretical knowledge on Z boson production. With higher precision measurements at low DY transverse momentum we could gain more insights to the structure of the hadrons. Such a measurement could be possible at the Run3 of the LHC, where double muon triggers with lower transverse momentum thresholds have been implemented in the HLT trigger system.

Apart from the DY production of the Z boson, at high energies, the Z boson can be produced in a radiation, as a real electroweak corrections. Electroweak corrections become large at high energies, and can affect the different agents involved in proton-proton collisions: the hadron structure, the matrix element and the initial and final state radiation. Due to the broken SU(2) symmetry, an important effect in proton-proton collisions is the NB violations, where real corrections do not cancel with the corresponding virtual corrections. While for the Z boson the cancellation is total, for the $W^{\pm}$ boson the cancellation is not complete, as the incoming proton is not invariant under SU(2) transformations. Thus, an understanding of the impact of electroweak corrections is crucial for precision measurements.

High energetic QCD processes are sensitive to electroweak corrections at high jet transverse momentum. Virtual electroweak corrections are negative and at large jet transverse momentum, at TeV scales, their contribution is of the order of -10% [58], while real electroweak corrections are positive with a contribution of the order of 6% [64]. In this thesis we have measured the contribution of real electroweak corrections in QCD process as a function of jet transverse momentum, with the $Z \to \mu^+\mu^-$

decay channel for a precise reconstruction of the Z boson. We analyse proton-proton collisions measured by the CMS detector at the LHC in 2016 corresponding to an integrated luminosity of 36.6 fb^{-1}.

For the measurement we have studied a phase space sensitive to electroweak boson radiation: a multijet scenario with jet transverse momentum larger than 100 GeV. Here, we have compared a DY+1 Jet calculation with a QCD calculation supplemented with an electroweak shower (QCD+EW shower) at leading order accuracy. With the observables we have defined we are able to disentangle the DY+1 Jet calculation from the QCD+EW shower calculation.

At present, our measurement is limited by the small amount of events available in the simulated sample at large jet transverse momentum in Z+Jets events. The limited number of events affects the correction to particle level with an impact of 8%. At low jet transverse momentum ($<$ 1 TeV) the measurement is dominated by jet energy scale uncertainties with an impact of $4-5\%$. In QCD events the jet energy scale is the driving uncertainty with an impact of $4-5\%$. The double differential cross-sections in Z+jets events, the jet energy scale is the dominant uncertainty for the leading jet transverse momentum bins below 500 GeV. In the highest leading jet transverse momentum bin, on the other hand, the limiting factor is the number of measured events. Thus, our measurement at large leading jet transverse momentum will be benefited by a full Run2 analysis, which corresponds to an integrated luminosity of 138 fb^{-1}.

We have compared the measured observables with different theoretical predictions at next-to-leading order accuracy in QCD. We have found that the DY+2 jets calculations agree with the measurement, and that multi-parton interactions and TMD effects do not play a role at high jet transverse momentum. We can conclude DY+2 jets calculations can be used for the modelling of real electroweak radiations. However, we find that the Pythia8 electroweak shower is not able to describe the measurement.

We extract the contribution of the Z boson radiation as a function of the jet transverse momentum. We divide the cross-section as a function of jet transverse momentum of Z+jets events with the one from QCD events. The measured contribution of $Z \to \mu^+\mu^-$ to QCD events at jet transverse momentum of 1 TeV is of

$$\delta^{QCD}_{Z\to\mu^+\mu^-}(1\text{ TeV}) = (1.63 \pm 0.13(\text{syst.}) \pm 0.21(\text{stat.})) \times 10^{-4}.$$

The $Z \to \mu^+\mu^-$ decay channel correspond to 3.6% of the total branching ratio. We extrapolate the contribution to the whole Z boson branching ratio to extract the total contribution:

$$\delta^{QCD}_{Z} = 0.0045 \pm 0.0006(\text{stat.}),$$

where we only have considered the statistical precision, as systematics vary depending on the final state measured. To estimate the total contribution of real electroweak radiations we need to take into account the W boson. Its coupling strength (mass) is larger (smaller) than the one of the Z boson, hence its contribution is larger. We estimate that at jet transverse momentum of 1 TeV the total contribution of real electroweak radiations in QCD events is around 1-2%. This results agree with the theoretical calculation in [64]. Thus, for any precision QCD measurement real electroweak radiations should be taken into account.

Bibliography

[1] "The CERN Experimental Programme: Grey Book database." `https://greybook.cern.ch/experiment/detail?id=CMS` Accessed: June 2023.

[2] S. L. Glashow, "The renormalizability of vector meson interactions." *Nucl. Phys.* **10** (1959) 107–117.

[3] A. Salam and J. C. Ward, "Weak and electromagnetic interactions." *Nuovo Cim.* **11** (1959) 568–577.

[4] S. Weinberg, "A model of leptons." *Phys. Rev. Lett.* **19** (Nov, 1967) 1264–1266.

[5] UA1 Collaboration, G. Arnison et al., "Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider." *Phys. Lett. B* **126** (1983) 398–410.

[6] S. D. Drell and T.-M. Yan, "Massive Lepton Pair Production in Hadron-Hadron Collisions at High-Energies." *Phys. Rev. Lett.* **25** (1970) 316–320. [Erratum: Phys.Rev.Lett. 25, 902 (1970)].

[7] E. Noether, "Invariante Variationsprobleme." *Nachr. D. König. Gesellsch. D. Wiss. Zu Göttingen, Math-phys. Klasse* (1918) 235–257.

[8] E. Fermi, "An attempt of a theory of beta radiation.." *Z. Phys.* **88** (1934) 161–177.

[9] N. Cabibbo, "Unitary symmetry and leptonic decays." *Phys. Rev. Lett.* **10** (Jun, 1963) 531–533.

[10] M. Kobayashi and T. Maskawa, "CP-Violation in the Renormalizable Theory of Weak Interaction." *Progress of Theoretical Physics* **49** (02, 1973) 652–657, [`https://academic.oup.com/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf`].

[11] J. Goldstone, "Field Theories with Superconductor Solutions." *Nuovo Cim.* **19** (1961) 154–164.

[12] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons." *Phys. Rev. Lett.* **13** (1964) 321–323.

[13] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons." *Phys. Rev. Lett.* **13** (1964) 508–509.

[14] R. P. Feynman, "Very high-energy collisions of hadrons." *Phys. Rev. Lett.* **23** (1969) 1415–1417.

[15] V. N. Gribov and L. N. Lipatov, "Deep inelastic e p scattering in perturbation theory." *Sov. J. Nucl. Phys.* **15** (1972) 438–450.

[16] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language." *Nucl. Phys. B* **126** (1977) 298–318.

[17] Y. L. Dokshitzer, "Calculation of the Structure Functions for Deep Inelastic Scattering and e^+ e^- Annihilation by Perturbation Theory in Quantum Chromodynamics.." *Sov. Phys. JETP* **46** (1977) 641–653.

[18] A. Bermudez Martinez, P. Connor, H. Jung, A. Lelek, R. Žlebčík, F. Hautmann, and V. Radescu, "Collinear and TMD parton densities from fits to precision DIS measurements in the parton branching method." *Phys. Rev. D* **99** (2019) 074008, [arXiv:1804.11152].

[19] F. Hautmann, H. Jung, M. Krämer, P. J. Mulders, E. R. Nocera, T. C. Rogers, and A. Signori, "TMDlib and TMDplotter: library and plotting tools for transverse-momentum-dependent parton distributions." *Eur. Phys. J. C* **74** (2014) 3220, [arXiv:1408.3015].

[20] J. Collins and T. C. Rogers, "Connecting Different TMD Factorization Formalisms in QCD." *Phys. Rev. D* **96** (2017) 054011, [arXiv:1705.07167].

[21] I. Bubanja et al., "The small k_T region in Drell-Yan production at next-to-leading order with the Parton Branching Method." *in preparation.*

[22] F. Hautmann, H. Jung, A. Lelek, V. Radescu, and R. Zlebcik, "Collinear and TMD Quark and Gluon Densities from Parton Branching Solution of QCD Evolution Equations." *JHEP* **01** (2018) 070, [arXiv:1708.03279].

[23] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and collider physics*, vol. 8. Cambridge University Press (2, 2011).

[24] H. Jung, "Lectures on QCD and Monte Carlo." (2015).

[25] J. C. Collins, D. E. Soper, and G. F. Sterman, "Transverse Momentum Distribution in Drell-Yan Pair and W and Z Boson Production." *Nucl. Phys. B* **250** (1985) 199–224.

[26] J. Collins, *Foundations of perturbative QCD*, vol. 32. Cambridge University Press (11, 2013).

[27] S. Catani, M. Ciafaloni, and F. Hautmann, "High energy factorization and small-x heavy flavour production." *Nuclear Physics B* **366** (1991) 135–188.

[28] A. S. Ito, R. J. Fisk, H. Jöstlein, D. M. Kaplan, S. W. Herb, D. C. Hom, L. M. Lederman, H. D. Snyder, J. K. Yoh, B. C. Brown, C. N. Brown, W. R. Innes, R. D. Kephart, K. Ueno, and T. Yamanouchi, "Measurement of the continuum of dimuons produced in high-energy proton-nucleus collisions." *Phys. Rev. D* **23** (1981) 604–633.

[29] A. Buckley, H. Hoeth, H. Lacker, H. Schulz, and J. E. von Seggern, "Systematic event generator tuning for the LHC." *Eur. Phys. J. C* **65** (2010) 331–357, [arXiv:0907.2973].

[30] CMS Collaboration, A. M. Sirunyan et al., "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements." *Eur. Phys. J. C* **80** (2020) 4, [arXiv:1903.12179].

[31] A. M. Sirunyan et al., "Development and validation of herwig 7 tunes from cms underlying-event measurements." *The European Physical Journal C* **81** (2021) 312.

[32] W. Jin et al., "Non-perturbative effects: lessons from fixed target, Tevatron, and LHC data." *in preparation.*

[33] A. Bermudez Martinez et al., "The transverse momentum spectrum of low mass Drell–Yan production at next-to-leading order in the parton branching method." *Eur. Phys. J. C* **80** (2020) 598, [arXiv:2001.06488].

[34] A. Bermudez Martinez et al., "The TMD Parton Branching Sudakov form factor in the context of TMD factorization." *in preparation.*

[35] S. Catani, B. R. Webber, and G. Marchesini, "QCD coherent branching and semiinclusive processes at large x." *Nucl. Phys. B* **349** (1991) 635–654.

[36] A. Bermudez Martinez and A. Vladimirov, "Determination of the Collins-Soper kernel from cross-sections ratios." *Phys. Rev. D* **106** (2022) L091501, [arXiv:2206.01105].

[37] I. Scimemi and A. Vladimirov, "Non-perturbative structure of semi-inclusive deep-inelastic and Drell-Yan scattering at small transverse momentum." *JHEP* **06** (2020) 137, [arXiv:1912.06532].

[38] MAP (Multi-dimensional Analyses of Partonic distributions) Collaboration, A. Bacchetta, V. Bertone, C. Bissolotti, G. Bozzi, M. Cerutti, F. Piacenza, M. Radici, and A. Signori, "Unpolarized transverse momentum distributions from a global fit of Drell-Yan and semi-inclusive deep-inelastic scattering data." *JHEP* **10** (2022) 127, [arXiv:2206.07598].

[39] A. Bacchetta, V. Bertone, C. Bissolotti, G. Bozzi, F. Delcarro, F. Piacenza, and M. Radici, "Transverse-momentum-dependent parton distributions up to N^3LL from Drell-Yan data." *JHEP* **07** (2020) 117, [arXiv:1912.07550].

[40] A. Bacchetta, F. Delcarro, C. Pisano, M. Radici, and A. Signori, "Extraction of partonic transverse momentum distributions from semi-inclusive deep-inelastic scattering, Drell-Yan and Z-boson production." *JHEP* **06** (2017) 081, [arXiv:1703.10157]. [Erratum: JHEP 06, 051 (2019)].

[41] M. Schlemmer, A. Vladimirov, C. Zimmermann, M. Engelhardt, and A. Schäfer, "Determination of the Collins-Soper Kernel from Lattice QCD." *JHEP* **08** (2021) 004, [arXiv:2103.16991].

[42] Lattice Parton Collaboration Collaboration, Q.-A. Zhang, J. Hua, Y. Huo, X. Ji, Y. Liu, Y.-S. Liu, M. Schlemmer, A. Schäfer, P. Sun, W. Wang, and Y.-B. Yang, "Lattice QCD Calculations of Transverse-Momentum-Dependent Soft Function through Large-Momentum Effective Theory." *Phys. Rev. Lett.* **125** (2020) 192001.

[43] P. Shanahan, M. Wagman, and Y. Zhao, "Lattice QCD calculation of the Collins-Soper kernel from quasi-TMDPDFs." *Phys. Rev. D* **104** (2021) 114502, [arXiv:2107.11930].

[44] Y. Li, S.-C. Xia, C. Alexandrou, K. Cichy, M. Constantinou, X. Feng, K. Hadjiyiannakou, K. Jansen, C. Liu, A. Scapellato, F. Steffens, and J. Tarello, "Lattice QCD Study of Transverse-Momentum Dependent Soft Function." *Phys. Rev. Lett.* **128** (2022) 062002.

[45] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations." *JHEP* **07** (2014) 079, [arXiv:1405.0301].

[46] R. Frederix, S. Frixione, V. Hirschi, D. Pagani, H. S. Shao, and M. Zaro, "The automation of next-to-leading order electroweak calculations." *JHEP* **07** (2018) 185, [arXiv:1804.10017]. [Erratum: JHEP 11, 085 (2021)].

[47] C. Bierlich et al., "A comprehensive guide to the physics and usage of PYTHIA 8.3." arXiv:2203.11601.

[48] T. Sjostrand and M. van Zijl, "A Multiple Interaction Model for the Event Structure in Hadron Collisions." *Phys. Rev. D* **36** (1987) 2019.

[49] M. Bahr et al., "Herwig++ Physics and Manual." *Eur. Phys. J. C* **58** (2008) 639–707, [arXiv:0803.0883].

[50] J. Bellm et al., "Herwig 7.0/Herwig++ 3.0 release note." *Eur. Phys. J. C* **76** (2016) 196, [arXiv:1512.01178].

[51] CASCADE Collaboration, S. Baranov et al., "CASCADE3 A Monte Carlo event generator based on TMDs." *Eur. Phys. J. C* **81** (2021) 425, [arXiv:2101.10221].

[52] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual." *JHEP* **05** (2006) 026, [hep-ph/0603175].

[53] P. Ciafaloni and D. Comelli, "Sudakov enhancement of electroweak corrections." *Phys. Lett. B* **446** (1999) 278–284, [hep-ph/9809321].

[54] F. Bloch and A. Nordsieck, "Note on the radiation field of the electron." *Phys. Rev.* **52** (1937) 54–59.

[55] T. Kinoshita, "Mass singularities of Feynman amplitudes." *J. Math. Phys.* **3** (1962) 650–677.

[56] T. D. Lee and M. Nauenberg, "Degenerate systems and mass singularities." *Phys. Rev.* **133** (1964) B1549–B1562.

[57] A. Scharf, "Electroweak corrections to b-jet and di-jet production." in *Meeting of the Division of Particles and Fields of the American Physical Society (DPF 2009)*, 10, 2009. arXiv:0910.0223.

[58] S. Dittmaier, A. Huss, and C. Speckner, "Weak radiative corrections to dijet production at hadron colliders." *JHEP* **11** (2012) 095, [arXiv:1210.0438].

[59] M. Ciafaloni, P. Ciafaloni, and D. Comelli, "Electroweak Bloch-Nordsieck violation at the TeV scale: 'Strong' weak interactions?." *Nucl. Phys. B* **589** (2000) 359–380, [hep-ph/0004071].

[60] M. Ciafaloni, P. Ciafaloni, and D. Comelli, "Electroweak double logarithms in inclusive observables for a generic initial state." *Phys. Lett. B* **501** (2001) 216–222, [hep-ph/0007096].

[61] M. Ciafaloni, P. Ciafaloni, and D. Comelli, "Bloch-Nordsieck violating electroweak corrections to inclusive TeV scale hard processes." *Phys. Rev. Lett.* **84** (2000) 4810–4813, [hep-ph/0001142].

[62] M. Ciafaloni, P. Ciafaloni, and D. Comelli, "Bloch-Nordsieck violation in spontaneously broken Abelian theories." *Phys. Rev. Lett.* **87** (2001) 211802, [hep-ph/0103315].

[63] G. Bell, J. H. Kuhn, and J. Rittinger, "Electroweak Sudakov Logarithms and Real Gauge-Boson Radiation in the TeV Region." *Eur. Phys. J. C* **70** (2010) 659–671, [arXiv:1004.4117].

[64] U. Baur, "Weak Boson Emission in Hadron Collider Processes." *Phys. Rev. D* **75** (2007) 013005, [hep-ph/0611241].

[65] S. Moretti, M. R. Nolten, and D. A. Ross, "Weak corrections to four-parton processes." *Nucl. Phys. B* **759** (2006) 50–82, [hep-ph/0606201].

[66] W. J. Stirling and E. Vryonidou, "Electroweak corrections and Bloch-Nordsieck violations in 2-to-2 processes at the LHC." *JHEP* **04** (2013) 155, [arXiv:1212.6537].

[67] S. Dawson, "The Effective W Approximation." *Nucl. Phys. B* **249** (1985) 42–60.

[68] G. L. Kane, W. W. Repko, and W. B. Rolnick, "The Effective W+-, Z0 Approximation for High-Energy Collisions." *Phys. Lett. B* **148** (1984) 367–372.

[69] C. W. Bauer, N. Ferland, and B. R. Webber, "Standard Model Parton Distributions at Very High Energies." *JHEP* **08** (2017) 036, [arXiv:1703.08562].

[70] A. Manohar, P. Nason, G. P. Salam, and G. Zanderighi, "How bright is the proton? A precise determination of the photon parton distribution function." *Phys. Rev. Lett.* **117** (2016) 242002, [arXiv:1607.04266].

[71] A. V. Manohar, P. Nason, G. P. Salam, and G. Zanderighi, "The Photon Content of the Proton." *JHEP* **12** (2017) 046, [arXiv:1708.01256].

[72] B. Fornal, A. V. Manohar, and W. J. Waalewijn, "Electroweak Gauge Boson Parton Distribution Functions." *JHEP* **05** (2018) 106, [arXiv:1803.06347].

[73] C. W. Bauer, N. Ferland, and B. R. Webber, "Combining initial-state resummation with fixed-order calculations of electroweak corrections." *JHEP* **04** (2018) 125, [arXiv:1712.07147].

[74] A. V. Manohar and W. J. Waalewijn, "Electroweak Logarithms in Inclusive Cross Sections." *JHEP* **08** (2018) 137, [arXiv:1802.08687].

[75] J. R. Christiansen and T. Sjöstrand, "Weak Gauge Boson Radiation in Parton Showers." *JHEP* **04** (2014) 115, [arXiv:1401.5238].

[76] H. Brooks, P. Skands, and R. Verheyen, "Interleaved resonance decays and electroweak radiation in the Vincia parton shower." *SciPost Phys.* **12** (2022) 101, [arXiv:2108.10786].

[77] J. Chen, T. Han, and B. Tweedie, "Electroweak Splitting Functions and High Energy Showering." *JHEP* **11** (2017) 093, [arXiv:1611.00788].

[78] C. Bierlich et al., "Robust Independent Validation of Experiment and Theory: Rivet version 3." *SciPost Phys.* **8** (2020) 026, [arXiv:1912.05451].

[79] CMS Collaboration, "Jet energy scale and resolution measurement with Run 2 Legacy Data Collected by CMS at 13 TeV." https://cds.cern.ch/record/2792322" (2021). DP note.

[80] CMS Collaboration, "Jet Energy Scale and Resolution Measurements Using Prompt Run3 Data Collected by CMS in the First Months of 2022 at13.6 TeV." https://cds.cern.ch/record/2841534 (2022). DP note.

[81] E. Mobs, "The CERN accelerator complex in 2019. Complexe des accélérateurs du CERN en 2019." https://cds.cern.ch/record/2684277 (2019).

[82] "Public CMS Luminosity Information." https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults

[83] CMS Tracker Group Collaboration, W. Adam et al., "The CMS Phase-1 Pixel Detector Upgrade." *JINST* **16** (2021) P02027, [arXiv:2012.14304].

[84] CMS Collaboration, S. Chatrchyan et al., "The CMS experiment at the CERN LHC." *Journal of Instrumentation* **3** (2008) S08004.

[85] CMS Collaboration, A. M. Sirunyan et al., "Particle-flow reconstruction and global event description with the CMS detector." *JINST* **12** (2017) P10003, [arXiv:1706.04965].

[86] A. Benaglia, "The CMS ECAL performance with examples." *Journal of Instrumentation* **9** (2014) C02008.

[87] CMS Collaboration, A. M. Sirunyan et al., "Calibration of the CMS hadron calorimeters using proton-proton collision data at $\sqrt{s} = 13$ TeV." *JINST* **15** (2020) P05002, [arXiv:1910.00079].

[88] CMS Collaboration, G. L. Bayatian et al., "CMS Physics: Technical Design Report Volume 1: Detector Performance and Software.".

[89] GEANT4 Collaboration, S. Agostinelli et al., "GEANT4–a simulation toolkit." *Nucl. Instrum. Meth. A* **506** (2003) 250–303.

[90] J. Allison et al., "Geant4 developments and applications." *IEEE Trans. Nucl. Sci.* **53** (2006) 270.

[91] J. Allison et al., "Recent developments in Geant4." *Nucl. Instrum. Meth. A* **835** (2016) 186–225.

[92] CMS Collaboration, A. M. Sirunyan et al., "Particle-flow reconstruction and global event description with the CMS detector." *JINST* **12** (2017) P10003, [arXiv:1706.04965].

[93] CMS Collaboration, "A Cambridge-Aachen (C-A) based Jet Algorithm for boosted top-jet tagging.".

[94] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, "Longitudinally invariant K_t clustering algorithms for hadron hadron collisions." *Nucl. Phys. B* **406** (1993) 187–224.

[95] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm." *JHEP* **04** (2008) 063, [arXiv:0802.1189].

[96] D. Bertolini, P. Harris, M. Low, and N. Tran, "Pileup Per Particle Identification." *JHEP* **10** (2014) 059, [arXiv:1407.6013].

[97] CMS Collaboration, A. M. Sirunyan et al., "Pileup mitigation at CMS in 13 TeV data." *JINST* **15** (2020) P09018, [arXiv:2003.00503].

[98] CMS Collaboration, "Jet Energy Resolution and Corrections (JERC) Subgroup." https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetEnergyScale Accessed: June 2023.

[99] CMS Collaboration, V. Khachatryan et al., "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV." *JINST* **12** (2017) P02014, [arXiv:1607.03663].

[100] A. Bodek, A. van Dyne, J. Y. Han, W. Sakumoto, and A. Strelnikov, "Extracting Muon Momentum Scale Corrections for Hadron Collider Experiments." *Eur. Phys. J. C* **72** (2012) 2194, [arXiv:1208.3710].

[101] J. Lidrych, *Measurement of the differential cross section of W and Z bosons in the boosted region in proton-proton collisions with the CMS experiment*, dissertation, Universität Hamburg, Hamburg (2022). Dissertation, Universität Hamburg, 2021.

[102] CMS Collaboration, V. Khachatryan et al., "Measurement of the double-differential inclusive jet cross section in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$." *Eur. Phys. J. C* **76** (2016) 451, [arXiv:1605.04436].

[103] NNPDF Collaboration, R. D. Ball et al., "Parton distributions from high-precision collider data." *Eur. Phys. J. C* **77** (2017) 663, [arXiv:1706.00428].

[104] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations." *JHEP* **07** (2014) 079, [arXiv:1405.0301].

[105] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO." *JHEP* **12** (2012) 061, [arXiv:1209.6215].

[106] P. Connor, *Precision measurement of the inclusive b jet production in proton proton collisions with the CMS experiment at the LHC at* $\sqrt{s} = 13\ TeV$, dissertation, Universität Hamburg, Hamburg (2018). Dissertation, Universität Hamburg, 2018.

[107] L. Brenner, R. Balasubramanian, C. Burgard, W. Verkerke, G. Cowan, P. Verschuuren, and V. Croft, "Comparison of unfolding methods using RooFitUnfold." *Int. J. Mod. Phys. A* **35** (2020) 2050145, [arXiv:1910.14654].

[108] CMS Collaboration, "Azimuthal correlations in Z+jets events in proton-proton collisions at $\sqrt{s} = 13$ TeV." arXiv:2210.16139.

[109] CMS Collaboration, A. M. Sirunyan et al., "Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS." *Eur. Phys. J. C* **81** (2021) 800, [arXiv:2104.01927].

[110] CMS Collaboration, "Measurements of jet multiplicity and jet transverse momentum in multijet events in proton-proton collisions at $\sqrt{s} = 13$ TeV." arXiv:2210.13557.

[111] R. Kleiss and R. Verheyen, "Collinear electroweak radiation in antenna parton showers." *Eur. Phys. J. C* **80** (2020) 980, [arXiv:2002.09248].

[112] M. R. Masouminia and P. Richardson, "Implementation of angularly ordered electroweak parton shower in Herwig 7." *JHEP* **04** (2022) 112, [arXiv:2108.10817].

[113] Particle Data Group Collaboration, R. L. Workman et al., "Review of Particle Physics." *PTEP* **2022** (2022) 083C01.

[114] CMS Collaboration, A. Tumasyan et al., "Measurement and QCD analysis of double-differential inclusive jet cross sections in proton-proton collisions at $\sqrt{s}$ = 13 TeV." *JHEP* **02** (2022) 142, [arXiv:2111.10431]. [Addendum: JHEP 12, 035 (2022)].

[115] NuSea Collaboration, J. C. Webb et al., "Absolute Drell-Yan dimuon cross-sections in 800 GeV/c pp and pd collisions." hep-ex/0302019.

[116] J. C. Webb, *Measurement of continuum dimuon production in 800-GeV/C proton nucleon collisions*, Ph.D. thesis, New Mexico State U. (2003). hep-ex/0301031.

[117] D. Antreasyan et al., "Dimuon Scaling Comparison at 44-GeV and 62-GeV." *Phys. Rev. Lett.* **48** (1982) 302.

[118] PHENIX Collaboration, C. Aidala et al., "Measurements of $\mu\mu$ pairs from open heavy flavor and Drell-Yan in $p+p$ collisions at $\sqrt{s} = 200$ GeV." *Phys. Rev. D* **99** (2019) 072003, [arXiv:1805.02448].

[119] D0 Collaboration, B. Abbott et al., "Measurement of the inclusive differential cross section for Z bosons as a function of transverse momentum in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV." *Phys. Rev. D* **61** (2000) 032004, [hep-ex/9907009].

[120] CDF Collaboration, T. Affolder et al., "The transverse momentum and total cross section of e^+e^- pairs in the Z boson region from $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV." *Phys. Rev. Lett.* **84** (2000) 845–850, [hep-ex/0001021].

[121] D0 Collaboration, V. M. Abazov et al., "Measurement of the Normalized $Z/\gamma^* -> \mu^+\mu^-$ Transverse Momentum Distribution in $p\bar{p}$ Collisions at $\sqrt{s} =$ 1.96 TeV." *Phys. Lett. B* **693** (2010) 522–530, [arXiv:1006.0618].

[122] CDF Collaboration, T. Aaltonen et al., "Transverse momentum cross section of e^+e^- pairs in the Z-boson region from $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV." *Phys. Rev. D* **86** (2012) 052010, [arXiv:1207.7138].

[123] CMS Collaboration, S. Chatrchyan et al., "Study of Z production in PbPb and pp collisions at $\sqrt{s_{\rm NN}} = 2.76$ TeV in the dimuon and dielectron decay channels." *JHEP* **03** (2015) 022, [arXiv:1410.4825].

[124] ATLAS Collaboration, G. Aad et al., "Measurement of the transverse momentum and ϕ^*_η distributions of Drell–Yan lepton pairs in proton–proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector." *Eur. Phys. J. C* **76** (2016) 291, [arXiv:1512.02192].

[125] CMS Collaboration, A. M. Sirunyan et al., "Study of Drell-Yan dimuon production in proton-lead collisions at $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV." *JHEP* **05** (2021) 182, [arXiv:2102.13648].

[126] CMS Collaboration, "Measurement of mass dependence of the transverse momentum of Drell Yan lepton pairs in proton-proton collisions at $\sqrt{s} = 13$ TeV." *arXiv preprint arXiv:2205.04897* (2021).

[127] LHCb Collaboration, R. Aaij et al., "Precision measurement of forward Z boson production in proton-proton collisions at $\sqrt{s} = 13$ TeV." *JHEP* **07** (2022) 026, [arXiv:2112.07458].

[128] M. Cacciari and G. P. Salam, "Dispelling the N^3 myth for the k_t jet-finder." *Phys. Lett. B* **641** (2006) 57–61, [hep-ph/0512210].

[129] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual." *Eur. Phys. J. C* **72** (2012) 1896, [arXiv:1111.6097].

APPENDIX A

Tune of the intrinsic transverse momentum

In this appendix we present the major aspects of the tuning of the intrinsic transverse momentum. The DY measurements used for the tuning are summarised in table A.1. For the tuning of the intrinsic transverse momentum we consider the 0 to 10 GeV range in the Drell-Yan transverse momentum spectrum. Above that range the DY transverse momentum is no longer sensitive to the intrinsic transverse momentum. We consider this choice of the tuning range as an uncertainty, which will we refer to as range uncertainty. We estimate the uncertainty by comparing the tuned result obtained with the 0 to 10 GeV range with another tune with a 0 to 15 GeV range. The order of the polynomial used in the interpolation is also considered an uncertainty. We determine the interpolation of the intrinsic transverse momentum with an order-5 polynomial. To determine the uncertainty the difference with an order-3 polynomial is used. We summarise the tuned values of the intrinsic transverse momentum for the different DY measurements in table A.2.

Experiment	**Process**	$\sqrt{s}$ **[GeV]**	Q **[GeV]**
E866/NuSea [115, 116]	pp/pd, fixed target	38.8	4-12.85
R209 [117]	pp	62	5-8
PHENIX [118]	pp	200	4.8-8.2
D0/CDF [119, 120]	$p\bar{p}$	1800	Z mass
D0/CDF [121, 122]	$p\bar{p}$	1960	Z mass
CMS [123]	pPb	2760	Z mass
ATLAS [124]	pp	8000	46-150
CMS [125]	pPb	8160	15-120
CMS [126]	pp	13000	50-1000
LHCb [127]	pp	13000	Z mass

Table A.1: Analysed Drell-Yan measurements for the tune of the intrinsic transverse momentum in Pythia8 and Herwig7 Monte-Carlo event generators. Where the experiments consist in different collision processes at different centre-of-mass energies, $\sqrt{s}$, and different DY lepton pair masses, Q.

Energy	Generator setup	Tune result ± stat ± range ± int
38.8 GeV	Pythia 8 CP5	0.988 ± 0.026 ± 0.022 ± 0.015
	Pythia 8 CP4	0.993 ± 0.026 ± 0.017 ± 0.009
	Pythia 8 CP3	0.990 ± 0.027 ± 0.017 ± 0.020
	Herwig 7 CH2	0.829 ± 0.016 ± 0.010 ± 0.06
62 GeV	Pythia 8 CP5	1.24 ± 0.06 ± 0.0015 ± 0.06
	Pythia 8 CP4	1.24 ± 0.06 ± 0.0012 ± 0.006
	Pythia 8 CP3	1.15 ± 0.05 ± 0.0010 ± 0.010
	Herwig 7 CH2	0.94 ± 0.04 ± 0.0012 ± 0.024
200 GeV	Pythia 8 CP5	1.47 ± 0.11 ± 0.005 ± 0.06
	Pythia 8 CP4	1.54 ± 0.12 ± 0.003 ± 0.004
	Pythia 8 CP3	1.54 ± 0.12 ± 0.003 ± 0.022
	Herwig 7 CH2	1.14 ± 0.08 ± 0.003 ± 0.018
1.8 TeV	Pythia 8 CP5	1.93 ± 0.07 ± 0.02 ± 0.015
	Pythia 8 CP4	1.94 ± 0.08 ± 0.04 ± 0.0005
	Pythia 8 CP3	2.09 ± 0.08 ± 0.03 ± 0.007
	Herwig 7 CH2	1.52 ± 0.09 ± 0.03 ± 0.0024
1.96 TeV	Pythia 8 CP5	1.88 ± 0.06 ± 0.10 ± 0.009
	Pythia 8 CP4	1.93 ± 0.07 ± 0.03 ± 0.009
	Pythia 8 CP3	2.03 ± 0.07 ± 0.03 ± 0.008
	Herwig 7 CH2	1.41 ± 0.08 ± 0.08 ± 0.0019
2.76 TeV	Pythia 8 CP5	2.36 ± 0.26 ± 0.005 ± 0.005
	Pythia 8 CP4	2.39 ± 0.3 ± 0.024 ± 0.013
	Pythia 8 CP3	2.49 ± 0.3 ± 0.004 ± 0.007
	Herwig 7 CH2	1.87 ± 0.44 ± 0.06 ± 0.003
8 TeV	Pythia 8 CP5	2.50 ± 0.07 ± 0.02 ± 0.03
	Pythia 8 CP4	2.55 ± 0.03 ± 0.016 ± 0.04
	Pythia 8 CP3	2.60 ± 0.04 ± 0.002 ± 0.010
	Herwig 7 CH2	1.95± 0.05 ± 0.04 ± 0.017
8.16 TeV	Pythia 8 CP5	2.66 ± 0.12 ± 0.02 ± 0.015
	Pythia 8 CP4	2.63 ± 0.12 ± 0.023 ± 0.013
	Pythia 8 CP3	2.62 ± 0.12 ± 0.015 ± 0.007
	Herwig 7 CH2	1.96 ± 0.16 ± 0.06 ± 0.027
13 TeV (CMS)	Pythia 8 CP5	2.673 ± 0.024 ± 0.028 ± 0.04
	Pythia 8 CP4	2.654 ± 0.026 ± 0.08 ± 0.004
	Pythia 8 CP3	2.619 ± 0.027 ± 0.05 ± 0.009
	Herwig 7 CH2	2.05 ± 0.04 ± 0.03 ± 0.035
13 TeV (LHCb)	Pythia 8 CP5	2.66 ± 0.05 ± 0.13 ± 0.0007
	Pythia 8 CP4	2.67 ± 0.05 ± 0.10 ± 0.003
	Pythia 8 CP3	2.62 ± 0.05 ± 0.11 ± 0.00007
	Herwig 7 CH2	1.99 ± 0.07 ± 0.05 ± 0.04

Table A.2: The tune results for the BeamRemnants:primordialkThard parameter in Pythia 8 and the ShowerHandler:IntrinsicPtGaussian parameter in Herwig 7, taking into account the uncertainty from tune ranges and the functions for interpolation.

APPENDIX B

Phenomenological studies in ΔR cleaning

The Z boson might be radiated from one of the outgoing partons of the hard-scattering. This radiation might be collinear and the decay products of the Z boson can be found in the reconstructed jet.

In the phase space of our measurement we are sensitive to electroweak radiations, and hence, we expect to find muons inside the jets.

In the CMS experiment for the jet reconstruction all the Particle Flow candidates are used. For this, any muon with extra hadronic activity near by will be reconstructed as a jet, and the muon will fake a jet. To remove these fake jets, jets closer than a give ΔR distance from any muon in the event are removed from the event selection. In Drell-Yan measurements the minimum distance is set to $\Delta R = 0.4$. However, for our measurement this might be too conservative, as we expect the decay products of the Z boson to be found inside a jet.

To find an optimal minimum ΔR distance between the decay products of the Z boson and jets we simulate QCD events supplemented with the simple electroweak shower with the Pythia8 Monte-Carlo event generator. In a first step we estimate the impact of fake jets by performing two different jet reconstructions with FastJet [128, 129]:

- *Muons in*: all particles in the final state are included in the jet reconstruction.
- *Muons out*: all particles in the final state except muons are included in the jet reconstruction

We define a Rivet routine to measure the minimum ΔR distance between a muon

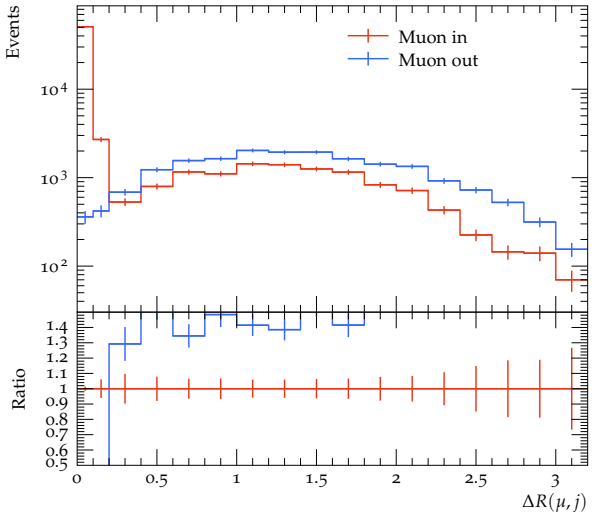

Figure B.1: Minimum ΔR distance between a jet and a muon in QCD events supplemented with an electroweak shower for two different jet reconstruction scenarios. Muons in: all particles in the final state are included in the jet reconstruction. Muons out: all particles in the final state except muons are included in the jet reconstruction.

(decay product of the Z boson) and a jet in the event, $\Delta R(\mu, j)$. In figure B.1 we show the $\Delta R(\mu, j)$ for both scenarios. At low $\Delta R(\mu, j)$ in the *Muons in* scenario we observe an excess of events compared to the *Muons out* scenario. Many fake jets are reconstructed when muons are used for the jet reconstruction. Thus, a $\Delta R(\mu, j)$ cleaning is necessary to remove the fake jets.

In figure B.2 we introduce a new scenario, which corresponds to the *Muons in* scenario where jets with a $\Delta R(\mu, j) < 0.2$ are removed from are removed from the event selection. With this extra requirement on the *Muons in* scenario we are able to recover the *Muons out* for $\Delta R(\mu, j) > 0.2$. With this cut we are able to remove the contribution of fake jets without removing much of our signal process.

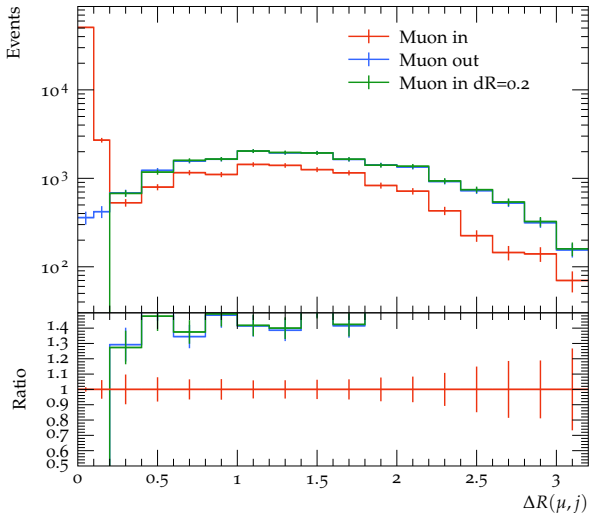

Figure B.2: Minimum ΔR distance between a jet and a muon in QCD events supplemented with an electroweak shower for three different jet reconstruction scenarios. Muons in: all particles in the final state are included in the jet reconstruction. Muons out: all particles in the final state except muons are included in the jet reconstruction. Muons in dr=0.2: Muons in scenario where jets with a $\Delta R(\mu, j) < 0.2$ are removed from are removed from the event selection.

APPENDIX C

Break down of uncertainties

In this appendix we present the systematic uncertainties affecting our measurement of the double differential cross-section in Z+jets events. In figures C.1-C.4 we can observe that at low jet transverse momentum jet energy calibration (jec) is the largest uncertainty, while for increasing jet transverse momentum the statistical precision of the simulated sample is the leading uncertainty (Unf. stat.).

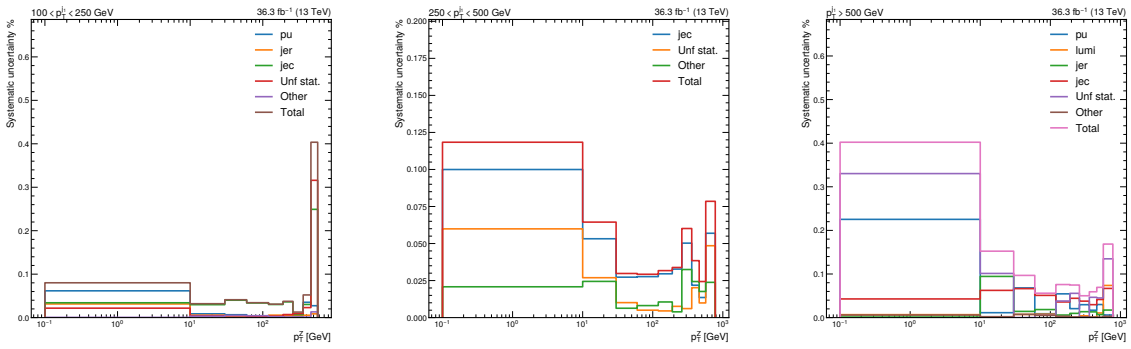

Figure C.1: Impact of systematic uncertainties in the cross-section as a function of the Z boson transverse momentum in Z+jets events in bins of leading jet transverse momentum, $p_{\mathrm{T}}^{j_1}$: $100 < p_{\mathrm{T}}^{j_1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j_1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j_1} > 250$ GeV (right panel). Systematic uncertainties below 2% are grouped in the category Other.

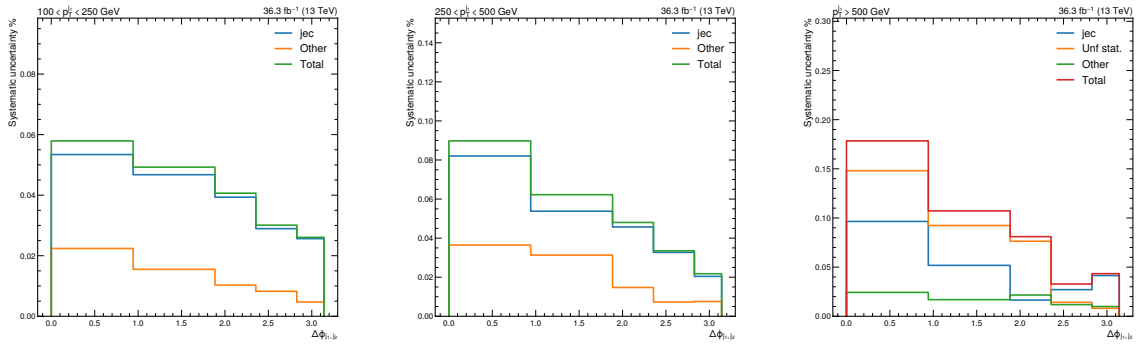

Figure C.2: Impact of systematic uncertainties in the cross-section as a function of $\Delta\phi_{j_1,j_2}$ in Z+jets events in bins of leading jet transverse momentum, $p_{\mathrm{T}}^{j_1}$: $100 < p_{\mathrm{T}}^{j_1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j_1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j_1} > 250$ GeV (right panel). Systematic uncertainties below 2% are grouped in the category Other.

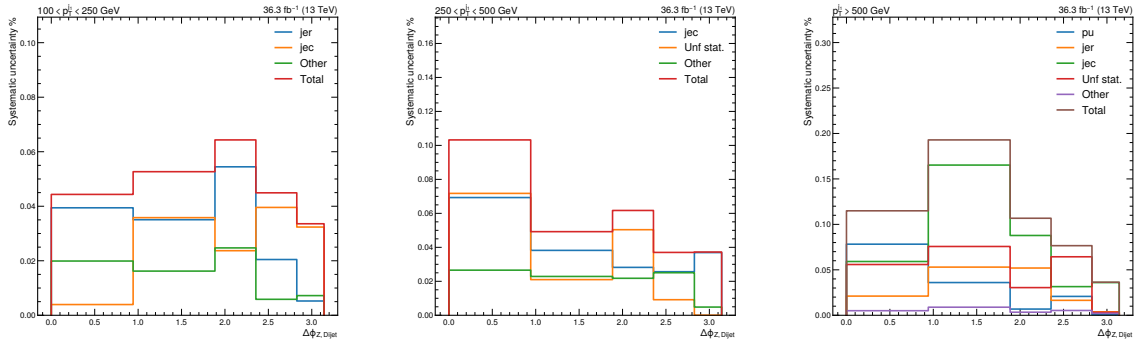

Figure C.3: Impact of systematic uncertainties in the cross-section as a function of $\Delta\phi_{Z,Dijet}$ in Z+jets events in bins of leading jet transverse momentum, $p_{\mathrm{T}}^{j_1}$: $100 < p_{\mathrm{T}}^{j_1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j_1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j_1} > 250$ GeV (right panel). Systematic uncertainties below 2% are grouped in the category Other.

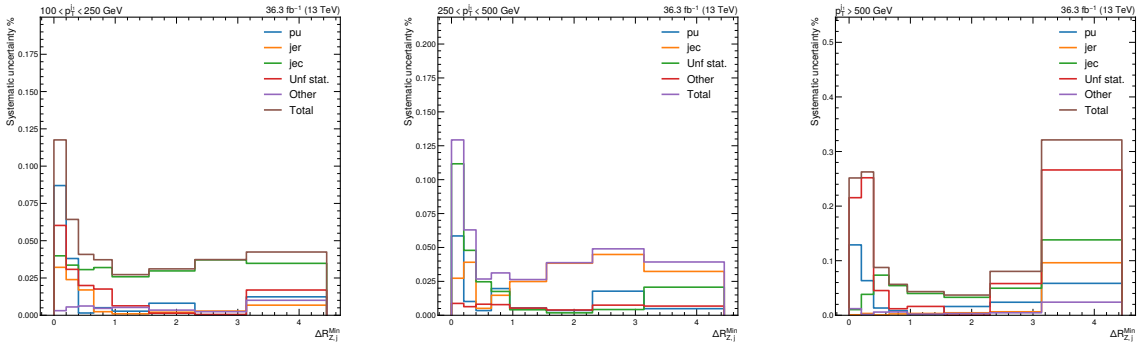

Figure C.4: Impact of systematic uncertainties in the cross-section as a function of $\Delta R^{Min}\phi_{Z,j}$ in Z+jets events in bins of leading jet transverse momentum, $p_{\mathrm{T}}^{j_1}$: $100 < p_{\mathrm{T}}^{j_1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j_1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j_1} > 250$ GeV (right panel). Systematic uncertainties below 2% are grouped in the category Other.

APPENDIX D

Contribution of b quark decays to Z+jets

In this appendix we present the background produced by QCD calculations in Z+jets events. We compare the MG5_aMC+Py8 (jj) and MG5_aMC+CA8 (jj) calculations presented in section 5.8 to the measured Z+jets events. With increasing jet transverse momentum the probability of b quark production increases, as a $b\bar{b}$ pair production. The corresponding B hadrons can then decay semileptonically,

$$B \to \mu + \nu_\mu + X, \tag{D.1}$$

producing muon pairs with opposite charge, which later can be reconstructed as a Z boson. We observe in figures D.1-D.4 that at small jet transverse momentum the contribution is negligible, however, for increasing jet transverse momentum the contribution becomes seizable. Thus, for future Z+jets analysis at high jet transverse momentum the contribution of such decays should be considered.

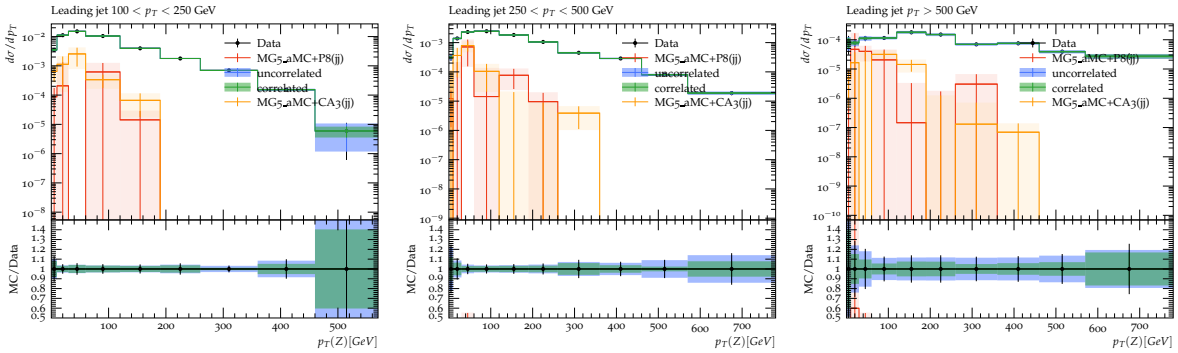

Figure D.1: Cross-section as a function of the Z boson transverse momentum in Z+jets events in bins of leading jet transverse momentum, p_{T}^{j1}: $100 < p_{\mathrm{T}}^{j1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j1} > 500$ GeV (right panel). Systematic (correlated) and statistical (uncorrelated) uncertainties are shown in bands around the measured data. The total uncertainty is given by the errorbars. We compare the result to different theoretical predictions: MG5_aMC+Py8(jj) and MG5_aMC+CA8(jj).

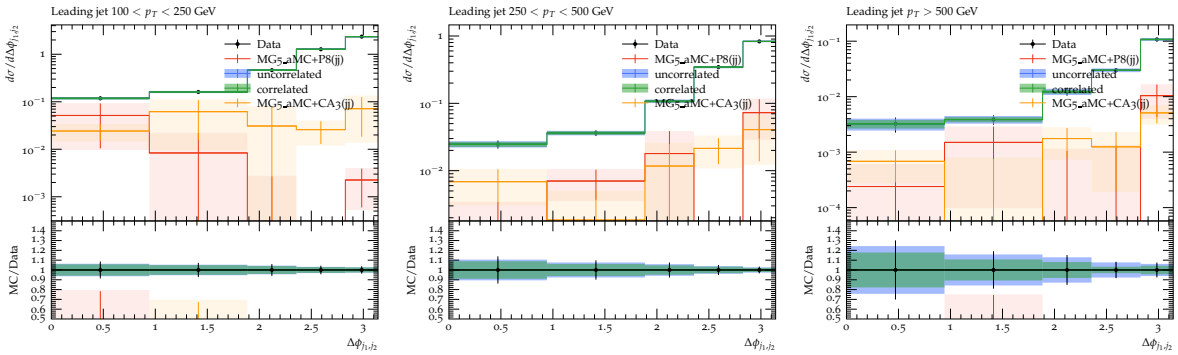

Figure D.2: Cross-section as a function of the $\Delta\phi_{j_1,j_2}$ in Z+jets events in bins of leading jet transverse momentum, p_{T}^{j1}: $100 < p_{\mathrm{T}}^{j1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j1} > 500$ GeV (right panel). Systematic (correlated) and statistical (uncorrelated) uncertainties are shown in bands around the measured data. The total uncertainty is given by the errorbars. We compare the result to different theoretical predictions: MG5_aMC+Py8(jj) and MG5_aMC+CA8(jj).

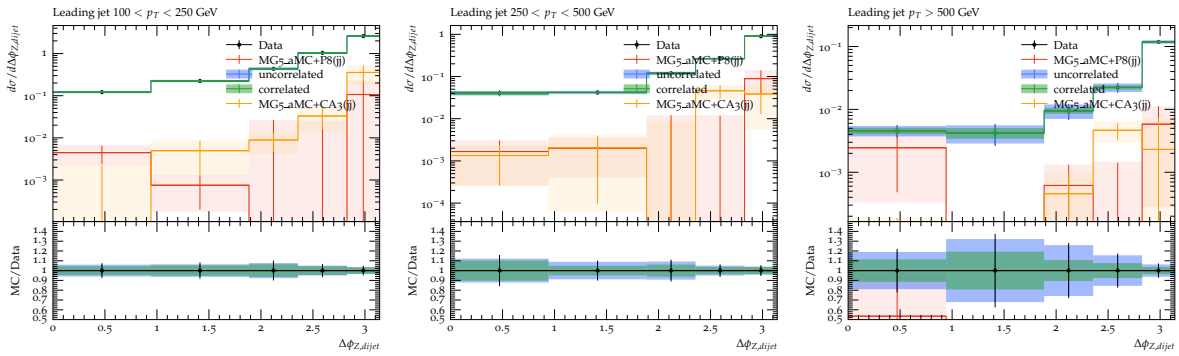

Figure D.3: Cross-section as a function of the $\Delta\phi_{Z,dijet}$ in Z+jets events in bins of leading jet transverse momentum, $p_{\mathrm{T}}^{j_1}$: $100 < p_{\mathrm{T}}^{j_1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j_1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j_1} > 500$ GeV (right panel). Systematic (correlated) and statistical (uncorrelated) uncertainties are shown in bands around the measured data. The total uncertainty is given by the errorbars. We compare the result to different theoretical predictions:MG5_aMC+Py8(jj) and MG5_aMC+CA8(jj).

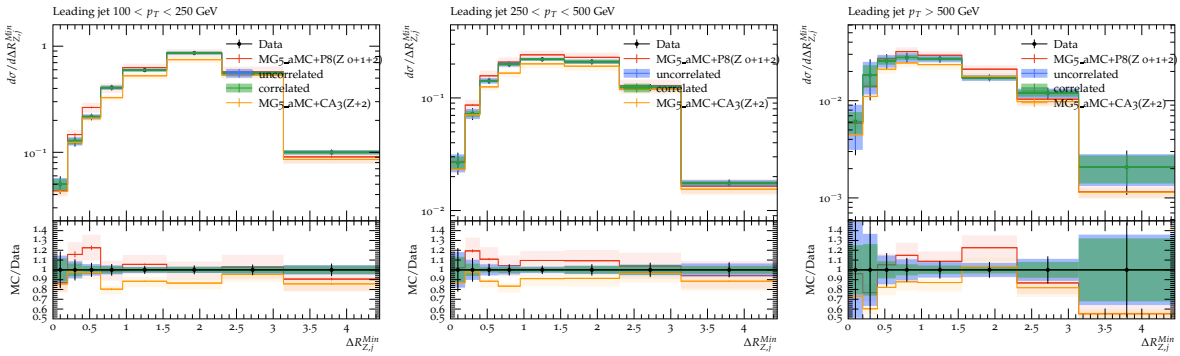

Figure D.4: Cross-section as a function of the $\Delta R_{Z,j}^{Min}$ in Z+jets events in bins of leading jet transverse momentum, $p_{\mathrm{T}}^{j_1}$: $100 < p_{\mathrm{T}}^{j_1} < 250$ GeV (left panel), $250 < p_{\mathrm{T}}^{j_1} < 500$ GeV (central panel) and $p_{\mathrm{T}}^{j_1} > 500$ GeV (right panel). Systematic (correlated) and statistical (uncorrelated) uncertainties are shown in bands around the measured data. The total uncertainty is given by the errorbars. We compare the result to different theoretical predictions: MG5_aMC+Py8(jj) and MG5_aMC+CA8(jj).

Acknowledgements

I extend my deepest gratitude to Hannes Jung, thank you for accepting me in the QCD group as a PhD student. Your guidance and support have not only contributed significantly to my understanding of physics but have also shaped my growth beyond the academic realm.

A special acknowledgment goes to Elisabetta Gallo for her warm welcome to the CMS DESY group and her supervision throughout my research.

My sincere appreciation goes to the people in the particle physics department at University of Antwerp, where six months were made remarkable by the kindness of its people. Pierre van Mechelen, your insightful supervision and innovative ideas greatly enhanced the measurement. Francesco Hautmann, for helping me with a deeper understanding of QCD. To the QCD theory group in Antwerp – Ola, Mees, and Lissa – collaborating with you again has been a pleasure. Nick, you made Antwerp not just a place of work but a cheerful experience beyond physics.

Jindrich and David, your friendship and guidance during the challenging times of the pandemic were pivotal. Your advice and support were unwavering, and I am truly grateful. Dominic, I appreciate the many walks and Frisbee games during the pandemic. Thanks to the entire QCD group, Sara, Qun, Luis Ignacio, Armando, Patrick and Keila, for making the journey enjoyable.

To the Italian Office – Federico, Valentina, and Alberto – you made office hours memorable, with the unexpected bonus of learning Italian. I am indebted to the PhDs and postdocs in the DESY CMS group for delightful lunchtimes, coffee breaks and standing together.

Lastly, my deepest appreciation goes to Thea. Thank you from the bottom of my heart.