Electroweak physics at the LHCb experiment

Marek Sirendi

Emmanuel College,
University of Cambridge

A dissertation submitted to the University of Cambridge
for the degree of Doctor of Philosophy
September 2016

Abstract

This thesis presents a measurement of the cross-section for $W \to e\nu$ production in pp collisions using data corresponding to an integrated luminosity of $2\,\text{fb}^{-1}$ collected by the LHCb experiment at a centre-of-mass energy of $\sqrt{s} = 8\,\text{TeV}$. Differential cross-sections as a function of the electron pseudorapidity are measured. The W^+/W^- cross-section ratio and production charge asymmetry are also reported. Results are compared with theoretical predictions at next-to-next-to-leading order in perturbative quantum chromodynamics. Finally, in a precise test of lepton universality, the ratio of W boson branching fractions is determined to be

$$\mathcal{B}(W \to e\nu)/\mathcal{B}(W \to \mu\nu) = 1.020 \pm 0.002 \pm 0.019,$$

where the first uncertainty is statistical and the second is systematic.

Also described is a study of matching and merging schemes and weak showering, as applied to the production of electroweak bosons in association with jets at the LHC. These advanced theoretical tools are seen to provide a good description of event shapes in the central region when compared to measurements performed by the ATLAS and CMS collaborations at a centre-of-mass energy of 7 TeV. Matching and merging schemes also provide a superior description of forward Z+jets production. The study constitutes a test of matching and merging schemes in a novel region of phase space and can be considered as a validation of the universality of these techniques. Finally, it was determined that current measurements of electroweak physics at LHCb do not yet fully probe the effect of weak showers.

Declaration

This dissertation is the result of my own work, except where explicit reference is made to the work of others, and has not been submitted for another qualification to this or any other university. This dissertation does not exceed the word limit for the respective Degree Committee.

Marek Sirendi

Pühendatud minu emale,
Riina Sirendile

Acknowledgements

Undertaking a PhD has been an immensely fulfilling experience and the culmination of a lifelong love of education. I would not have reached this point without the people around me who contributed greatly with support, help, and advice. Firstly, I would like to thank my supervisor David Ward who is quite simply the best supervisor in the world. If I am allowed to call myself a scientist now, it is because of his expert supervision, comprehensive guidance and feedback on all aspects of my work throughout my research.

I also take this opportunity to acknowledge the Cambridge HEP group. In particular, Val Gibson, Susan Haines, Chris Jones, Dave Sutherland, John Hill, and Steve Wotton have been tremendously helpful. The LHCb software framework was made much simpler by expert advice from Chris. Help from John and Steve with the group's computational infrastructure were immensely valuable for my research. Susan and Dave were always a useful sounding board when I found myself getting stuck. Val, as the head of group, provided financial support for many of the conferences and summer schools that were instrumental to my studies.

I am fortunate to have been a member of the LHCb QCD, Electroweak, and Exotica Working Group. Over the past four years, I have had the pleasure of working with some of the brightest and most diligent people I know. I would like to acknowledge the support I received from the conveners Wouter Hulsbergen, Will Barter, Phil Ilten, and Xabier Cid Vidal on many aspects of the $W \to e\nu$ measurement, particularly in the later stages of the analysis. What is more, Will Barter graciously offered to read the thesis and provided many valuable suggestions. Finally, I would like to extend a special thanks to Simone Bifani, Stephen Farry, and Katharina Müller for their contributions to bringing this measurement to fruition.

I had the good fortune of spending an exciting three months with the Lund University particle physics group as a Marie Curie Fellow. During my stay, I benefited greatly from the deep theoretical knowledge of my supervisors, Leif Lönnblad and Torbjörn Sjöstrand. I also had a lot of fun and felt at home almost immediately thanks to the friendship of

Jesper Roy Christiansen, Christine Overgaard Rasmussen, and Jonas Wessen. Finally, I must thank Evi Martinson for welcoming me into her home and making my stay in Lund a pleasure.

A.H. Tammsaare once said that you can survive on love and air. Fortunately, I never had to put this idea to the test thanks to the financial support of the Archimedes Foundation and Emmanuel College, Cambridge.

I would like to thank Weijia Li for her encouragement and patience during the writing of this thesis. Her constant support made what would otherwise have been an arduous undertaking feel effortless.

Soovin tänada ka oma perekonda, kelle toetuseta ma poleks nii kaugele jõudnud. Minu ema Riina, vanavanemad Malle, Mai ja Arvo on alati kaasa elanud minu tegemistele. Minu esmane huvi ja armastus reaalainete vastu tekkis tänu minu isa Aare ja vanaisa Ennu eeskujule. Aitäh!

Contents

Keldrikakand kena kakand,
keldris söönd ja keldris kakand.
Pole keldrist väljas käinud,
pole välisilma näinud.

Kõik, mis kehtib keldri õhus,
kehtib ka ta peas ja kõhus.
Ta ei ole isehakand,
ta on sünnist saati kakand.

Hando Runnel

Chapter 1

Introduction

Since the Large Hadron Collider (LHC) entered into operation in November 2009, its four main collaborations, ALICE, ATLAS, CMS, and LHCb, have produced a veritable bonanza of results. In fact, at the time of writing, roughly 1 600 papers have been published between the experiments and the world is fantastically richer for it. This thesis seeks to make a small, but significant, contribution to the wealth of knowledge that has been created in the form of measurements of the $W \to e\nu$ production cross-sections[1] in the forward region at $\sqrt{s} = 8\,\text{TeV}$ [1], theoretical predictions for measurements of Z production[2] at $\sqrt{s} = 13\,\text{TeV}$ with the LHCb experiment [2], and a study of electroweak (EW) boson production in association with jets with two advanced theoretical tools: matching and merging and weak showering [3]. The latter two contributions are presented together in Chapter 3.

The motivation for performing these measurements and studies is threefold. Firstly, precise measurements of the production cross-sections for W and Z bosons are important tests of the quantum chromodynamic (QCD) and EW sectors of the Standard Model (SM) of particle physics. Secondly, the parton[3] distribution functions (PDFs) of the proton can be better constrained [4]. Finally, these processes are also a ubiquitous background to physics beyond the SM[4] and so must be precisely measured and theoretically understood. The production of EW bosons has therefore been an important benchmark process to measure at current and past colliders.

[1]The decay $W \to e\nu$ denotes both $W^+ \to e^+\nu_e$ and $W^- \to e^-\overline{\nu}_e$ and similarly for the other leptonic decays. The $W \to e\nu$ cross-section denotes the product of the cross-section for W boson production and the branching fraction for $W \to e\nu$ decay.

[2]Throughout this thesis Z denotes the combined Z and virtual photon (γ^*) contribution.

[3]Partons denote both quarks and gluons.

[4]Typically when produced in association with jets.

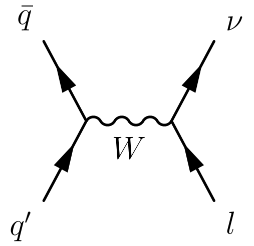


Figure 1.1: The leading order (LO) Feynman diagram for the production, and leptonic decay, of the W boson. Here, l can be any of (e, μ, τ) and ν can be any of $(\nu_e, \nu_\mu, \nu_\tau)$.

The thesis is organised as follows. Sections 1.1 and 1.2 provide theoretical and experimental contexts for the research undertaken. Chapter 2 gives an overview of the theory that forms the basis for the phenomenological studies that are detailed in Chapter 3 and the measurements of Chapter 5. Chapter 4 presents a description of the LHCb detector and Chapter 6 concludes the thesis. Finally, Chapter 7 provides a summary of the research in Estonian.

1.1 The theoretical context

The W and Z bosons are elementary particles that mediate the weak interaction according to the SM. They are extremely short-lived ($\sim 3 \cdot 10^{-25}$ s) and decay before they can be directly detected in an experiment. Together with the massless photon, they compose the bosonic fields of the unified electroweak theory proposed by Glashow, Weinberg, and Salam. They are among the heaviest particles within the SM with masses[5] $M_W = 80.385 \pm 0.015$ GeV and $M_Z = 91.1876 \pm 0.0021$ GeV [5]. They are produced by two weakly interacting charged particles which, in the case of the LHC, will be partons originating from the proton with the leading order process shown in Fig. 1.1 and subdominant contributions from the production of WW, WZ, WH or the decay of $t\bar{t}$, and others[6].

The SM postulates that the W boson couples to all generations of leptons equally (a principle known as "lepton universality") and so has approximately equal branching ratios, $\mathcal{B}(W \to e\nu) \approx \mathcal{B}(W \to \mu\nu) \approx \mathcal{B}(W \to \tau\nu) \approx 11.11\%$, to leptonic final states. This postulate is put to the test in Chapter 5 by determining the ratio $\mathcal{B}(W \to e\nu)/\mathcal{B}(W \to \mu\nu)$ in the forward region. Although the hadronic branching ratio dominates over the leptonic one, the latter is experimentally favoured in hadronic collisions as the cross-section for

[5] Natural units with $\hbar = c = 1$ are used throughout.

[6] Subdominant contributions in perturbative QCD are considered in Chapters 2 and 3.

dijet production swamps the hadronically-decaying W boson signal making it difficult to measure precisely.

1.2 The experimental context

The W and Z bosons were discovered using the UA1 and UA2 detectors on the $Sp\bar{p}S$ collider at CERN in 1983 [6–9], for which experimentalists Carlo Rubbia and Simon van der Meer were awarded the Nobel Prize for physics a year later. Subsequently, their production cross-sections were measured by the DØ and CDF experiments at the Tevatron at centre-of-mass energies of 1.8 TeV [10] and 1.96 TeV [11]. Constructed to perform precision measurements of the EW sector, the Large Electron-Positron Collider (LEP) studied W and Z bosons in detail [12]. More recently, measurements have been made by the CMS [13–15] and ATLAS [16–18] experiments at central rapidities at the LHC. Measurements tend to agree well with theoretical expectations and validate the theoretical framework outlined in Chapter 2.

1.2.1 EW measurements at LHCb

EW measurements in the forward region are complementary to those performed centrally for two reasons. Firstly, these measurements probe perturbative QCD (pQCD) and EW theory in a novel region of phase space as functions of the Bjorken-x of the proton and energy scale, Q, of the interaction. As a consequence, they are able to constrain PDFs at low Bjorken-x and high Q^2, where PDF uncertainties become particularly large as discussed in more detail in Section 2.3.

LHCb has performed measurements of Z boson production in the di-electron [2, 19, 20], di-muon [2, 21, 22] and di-tau [23] final states as a function of rapidity[7]. Additionally, W production cross-sections have been measured in the muon final state [22, 24] and, now, in the electron[8] final state [1] as a function of lepton pseudorapidity[9]. LHCb has also studied jet production in association with EW bosons [25–28].

[7]Rapidity is defined as $y \equiv \ln \left(\frac{E+p_L}{E-p_L}\right)^{\frac{1}{2}}$, where E is the energy and p_L is the longitudinal momentum of the particle.

[8]When referred to generically, "electron" denotes both e^+ and e^- and similarly for other leptonic states.

[9]Pseudorapidity η is defined as $\eta \equiv \ln \tan \frac{\theta}{2}$, where the polar angle θ is measured with respect to the beam axis. Pseudorapidity approximates rapidity for relativistic particles.

In addition, Ref. [24] compares measurements of W production cross-sections made by ATLAS [16] and CMS [29] to those of LHCb [24]. As Fig. 1.2 illustrates, good agreement between the experiments is observed in the overlap region and the results are internally consistent across the LHC.

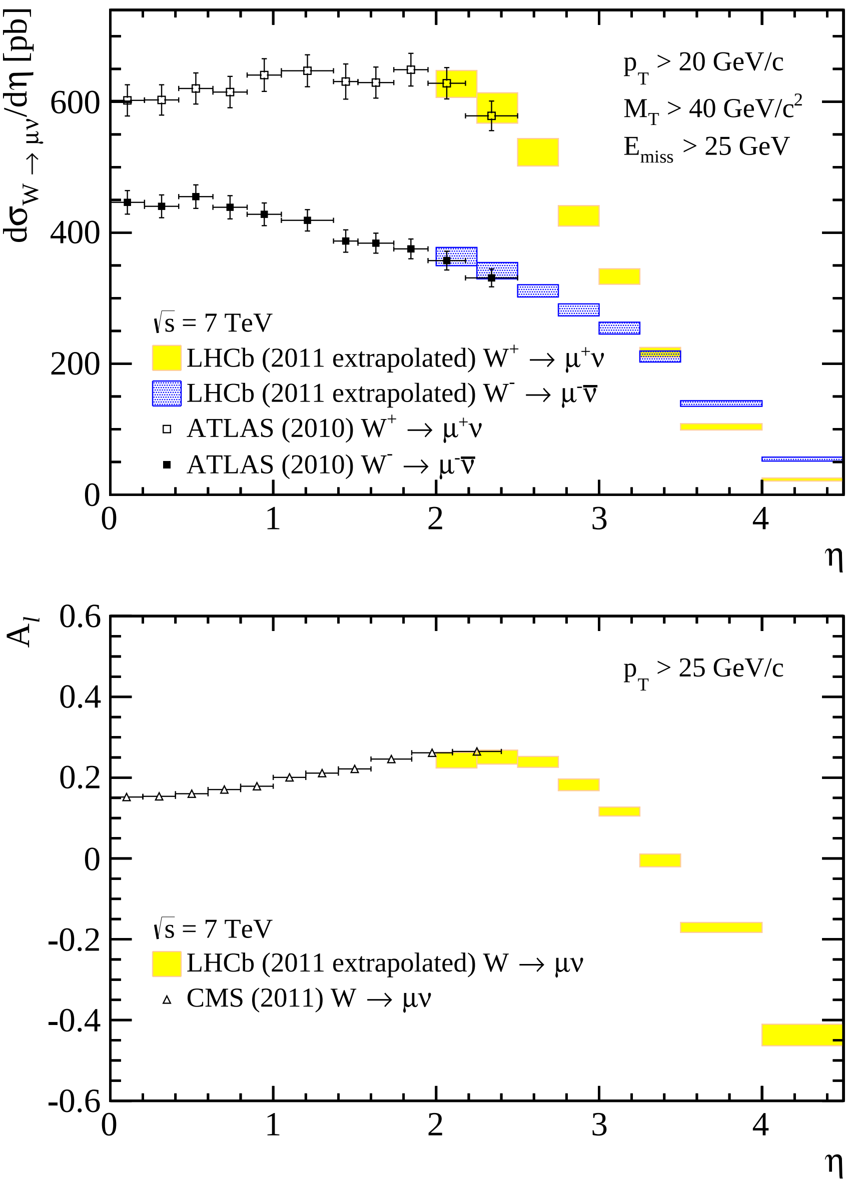


Figure 1.2: Top: Differential W cross-sections in bins of muon pseudorapidity. Measurements, represented as bands corresponding to the total uncertainty are extrapolated [30] to the ATLAS fiducial volume and compared to the ATLAS determinations (markers). Bottom: Lepton charge asymmetry in bins of muon pseudorapidity, where A_l is the difference in W^+ and W^- cross-sections divided by the sum. Measurements, represented as bands corresponding to the total uncertainty, are extrapolated to the CMS fiducial volume and compared to the CMS determinations (markers).

Chapter 2

Theoretical overview

This chapter begins by giving a broad overview of the SM with emphasis on QCD. It then proceeds by discussing the production of EW bosons in hadronic collisions in Section 2.2. Here, the perturbative treatment of QCD meets the non-perturbative model-driven description of the proton. Section 2.3 proceeds with an emphasis on the proton itself and the role of EW measurements in the forward region in probing the distributions of momenta inside the proton. This chapter is a review of the work of others and is largely based on Refs. [31–33].

2.1 The Standard Model of particle physics

The Standard Model of particle physics is a collection of quantum field theories describing the elementary constituents of matter and three of the fundamental interactions between them, the strong, weak, and electromagnetic (EM) force, in terms of an exchange of force carrying mediators. The quantum field theories are QCD, the weak interaction and quantum electrodynamics (QED), respectively. The weak and electromagnetic interactions have been successfully integrated to a single, electroweak, theory. The fourth known force, gravity, lacks an experimentally verified quantum field theory and so has yet to be incorporated into the SM.

Among other quantum numbers, particles carry different types of charge which specify whether or not they participate in a certain interaction. The matter particles are composed of quarks which carry colour charge, and leptons which do not. A hierarchy of masses leads one to separate them into three generations. There are six quarks, each carrying colour charge, fractional electric charge and weak charge. The three charged

Name	Symbol	EM charge (e)	Possible interactions			Generation
			Strong	Weak	EM	
up quark	u	+2/3	✓	✓	✓	I
down quark	d	−1/3	✓	✓	✓	
electron	e^-	−1		✓	✓	
electron neutrino	ν_e	0		✓		
charm quark	c	+2/3	✓	✓	✓	II
strange quark	s	−1/3	✓	✓	✓	
muon	μ^-	−1		✓	✓	
muon neutrino	ν_μ	0		✓		
top quark	t	+2/3	✓	✓	✓	III
bottom quark	b	−1/3	✓	✓	✓	
tau	τ^-	−1		✓	✓	
tau neutrino	ν_τ	0		✓		

Table 2.1: The matter particles of the SM and their properties.

leptons, the electron, muon and tau lepton, all have electric charge of −1. The three electrically neutral leptons are the electron neutrino (ν_e), the muon neutrino (ν_μ) and tau neutrino (ν_τ). All leptons carry weak charge and therefore participate in weak interactions. The matter particles of the SM and their properties are summarised in Table 2.1. Each particle listed in the table has an antiparticle with opposite quantum numbers such as electric and colour charges. Finally, all matter particles have intrinsic angular momentum, or "spin", of 1/2 and are known as "fermions".

On the other hand, the force carrying mediators of the SM have integer spin and are known as "bosons". Bosons are further divided into spin-1 particles, the photon, the gluon, W and Z bosons, and the scalar Higgs boson with a spin of 0. In the SM, the photon, the gluon and EW bosons are introduced to the theory by requiring the Lagrangian of the respective quantum field to be invariant under certain transformations. More specifically, it has been determined that the condition of local (*i.e.* varying in space-time) "gauge" invariance under the unitary groups U(1), SU(2) and SU(3) leads one to introduce additional bosonic fields (known as "gauge" bosons) to the Lagrangian that reproduce the fundamental interactions observed in nature. Within the SM, gauge bosons interact with all generations of leptons with the same strength and this principle is referred to as "lepton universality". This postulate of the theory is put to the test in Chapter 5. Finally, the Higgs boson that was discovered recently by the ATLAS [34]

Name	Symbol	Charge (e)	Spin	Mediator for			Mass [GeV]
				Strong	Weak	EM	
Photon	γ	0	1			✓	$< 10^{-27}$
W boson	W^-	-1	1		✓		80.385 ± 0.015
	W^+	$+1$	1		✓		
Z boson	Z	0	1		✓		91.1876 ± 0.0021
Gluon	g	0	1	✓			$< \mathcal{O}(10^{-3})$
Higgs	H	0	0				125.09 ± 0.24

Table 2.2: The fundamental bosons of the SM and their properties. The masses of the photon and the EW bosons are taken from Ref. [5]. For a discussion of bounds on the gluon mass see Ref. [42]. The mass of the Higgs boson is from a recent combination of ATLAS and CMS results [43].

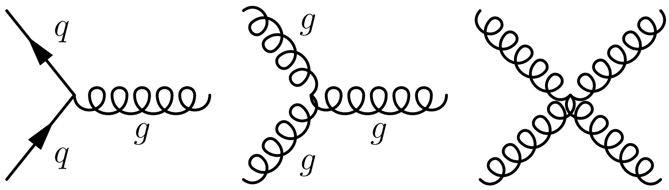


Figure 2.1: The fundamental interactions of QCD.

and CMS [35] collaborations plays a special role within the SM as a manifestation of the Higgs mechanism [36–41]. It gives mass to all matter particles via the so-called "Yukawa" interaction and, more importantly, provides a mechanism for the W and Z bosons to obtain mass without violating local gauge invariance under the groups listed above. The fundamental bosons of the SM and their properties are summarised in Table 2.2.

Unlike other matter particles, quarks have never been observed as free particles. Rather they are always detected as composite states with no net colour charge[1]. This observation is explained by the hypothesis of "colour confinement" which is believed to originate from gluon-gluon self-interactions. Feynman diagrams for gg self-interactions along with the fundamental vertex with quarks is illustrated in Fig. 2.1. These self-interactions lead to an attractive force between quarks with the energy stored in the field being proportional to distance. Consequently, as quarks are separated from one another, the energy of their bond is sufficient to produce additional quarks that then arrange themselves into colourless combinations in a process known as "hadronisation" as illustrated in Fig. 2.2. At hadron colliders this process occurs in a frame that is highly

[1]Technically, these composite particles are colour singlet states.

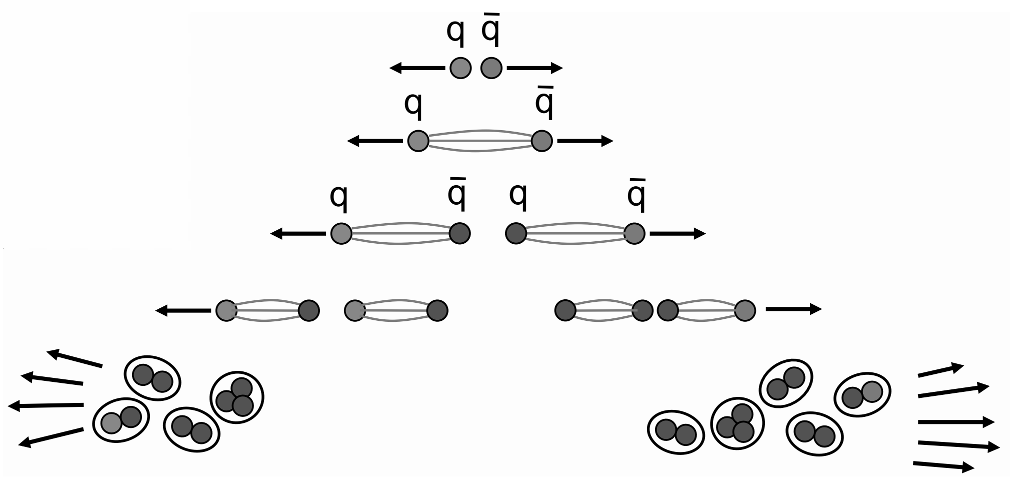


Figure 2.2: A qualitative picture of hadronisation. The diagram is adopted from Ref. [32].

boosted with respect to the lab frame and leads to a collimated beam of hadrons that are referred to as "jets".

Finally, it should be mentioned that QCD has a property known as "asymptotic freedom" whereby the strength of the strong coupling, α_s, decreases asymptotically as a function of the energy scale, Q, of the interaction as

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + \alpha_s(\mu^2)\frac{1}{12\pi}(33 - 2n_f)\ln(\frac{Q^2}{\mu^2})}, \tag{2.1}$$

where n_f is the number of quark flavours and μ is a reference scale. Eq. 2.1 exhibits asymptotic freedom as long as $n_f < 17$. The variation in α_s is shown both theoretically and experimentally in Fig. 2.3. As a consequence, QCD calculations can be performed perturbatively in the high-energy regime as discussed in Section 2.2 and Chapter 3. In contrast, non-perturbative models must be used to describe hadronisation and to model the distributions of momenta in the proton as presented in 2.3. The application of QCD to computing cross-sections for EW boson production in hadron-hadron collisions is described subsequently.

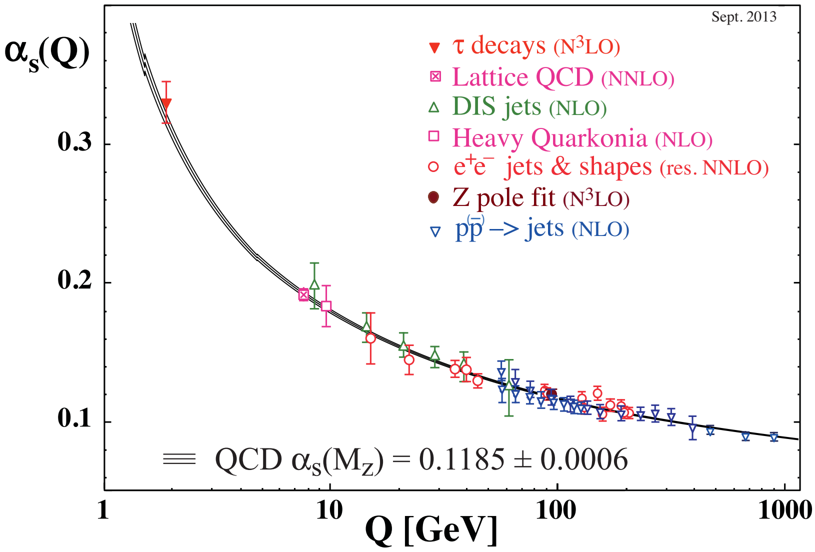


Figure 2.3: Summary of measurements of α_s as a function of the energy scale Q in different processes. Also provided is the world average evaluated at the Z boson mass. The graph is taken from Ref. [5].

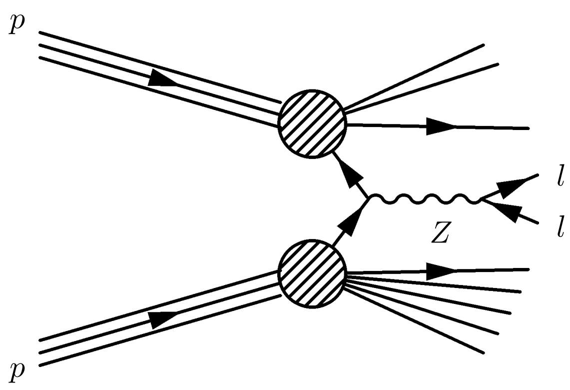


Figure 2.4: The Drell-Yan process for producing a Z boson in a pp collision.

2.2 The hadroproduction of weak gauge bosons

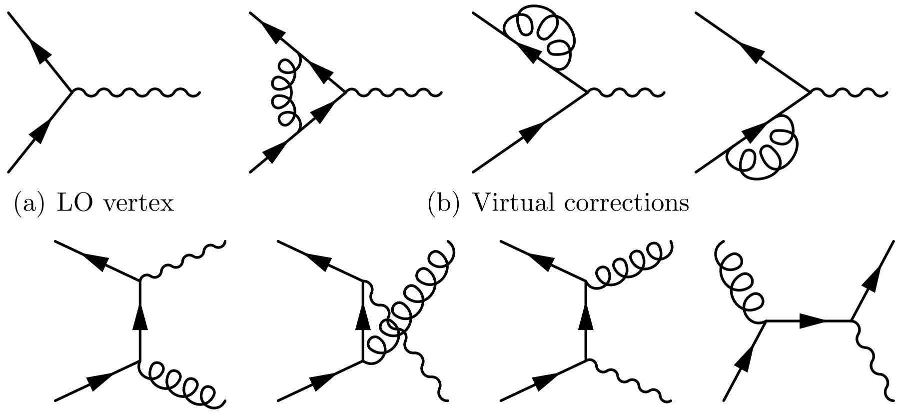
Drell and Yan first proposed that a scattering cross-section in hadronic collisions can be treated as a convolution of the partonic cross-section, calculable in pQCD, and the distributions of parton momentum in the proton, given by a PDF [44]. The archetypal example for the procedure is provided by the Drell-Yan (DY) interaction, illustrated in Fig. 2.4. However, it has been successfully extended to other processes and, crucially, enables higher order QCD corrections to be incorporated. The cross-section, $\sigma_{pp\to Z/W}$, summed over initial state partons a and b with momentum fractions x_a and x_b, with PDFs $f_a(x_a)$ and $f_b(x_b)$, and at a given energy scale Q is

$$\sigma_{pp\to Z/W} = \text{PDF} \otimes \hat{\sigma}_{ab\to Z/W} = \sum_{a,b} \int dx_a\, dx_b\, f_a(x_a, Q^2)\, f_b(x_b, Q^2)\, \hat{\sigma}_{ab\to X}(Q^2). \quad (2.2)$$

The partonic cross-section, $\hat{\sigma}_{ab\to Z/W}$, is calculated as a power series expansion in the strong coupling constant with next-to-leading-order (NLO) and next-to-next-to-leading-order (NNLO) terms corresponding to higher order emissions:

$$\hat{\sigma}_{ab\to Z/W} = [\hat{\sigma}_{\text{LO}} + \alpha_s \hat{\sigma}_{\text{NLO}} + \alpha_s^2 \hat{\sigma}_{\text{NNLO}} + ...]_{ab\to Z/W}. \quad (2.3)$$

Feynman diagrams for leading virtual and real (*i.e.* NLO) corrections are illustrated in Fig. 2.5. The higher order terms increase the LO cross-section for W and Z bosons by about $20-30\%$ [33] and are needed to reproduce the boson transverse momentum, $p_{\text{T}}^{Z/W}$, spectrum observed in data. Generally, the precision of the measurement dictates the order of the perturbative calculation that must be performed to make a valid comparison.



(a) LO vertex (b) Virtual corrections

(c) Real corrections

Figure 2.5: The LO interaction vertex with first order corrections.

For the DY process specifically, terms up to $\mathcal{O}(\alpha_s^2)$ have been computed [45–47] with recent progress on the N^{3}LO calculation [48]. However, as the number of Feynman diagrams increases roughly factorially with each order in perturbation theory, it becomes increasingly difficult to extend the precision of these calculations. What is more, QED and EW corrections become comparable in size at this order in α_s and must also be computed.

There is an additional difficulty with this approach. Namely, when the mass of the boson, $M_{Z/W}$, is significantly greater than its $p_{\mathrm{T}}^{Z/W}$ (*i.e.* close to threshold production), higher-order terms become large and spoil the convergence of the series. In particular, the emission of multiple soft gluons must be considered. The leading contributions at each order are proportional to

$$\alpha_s \ln \frac{M_{Z/W}^2}{(p_{\mathrm{T}}^{Z/W})^2}. \tag{2.4}$$

Clearly these logarithms are significant when $p_{\mathrm{T}}^{Z/W}$ is small. In fact, after taking prefactors into account, these terms become $\mathcal{O}(1)$ for $p_{\mathrm{T}}^{Z/W}$ values of less than $10 - 15\,\mathrm{GeV}$ [33]. Fortunately, these terms can be resummed to all orders. This is particularly advantageous near the boundaries of phase space where fixed order calculations break down. The method of resumming logarithms has been implemented in the ResBos [49–51] event generator where next-to-next-to-leading log (NNLL) corrections are computed for NLO differential cross-sections.

The inclusive Z or W cross-section, $\sigma_{Z/W}$, can also be decomposed into its multijet components, formulated as

$$\sigma_{Z/W} = \sigma_{Z/W+0\,\mathrm{jets}} + \sigma_{Z/W+1\,\mathrm{jets}} + \sigma_{Z/W+2\,\mathrm{jets}} + \sigma_{Z/W+n\,\mathrm{jets}} + \ldots, \tag{2.5}$$

where n is the number of jets and also sets the lowest power of α_s in the expansion of each term. In Chapter 3 cross-sections for $\sigma_{Z/W+0\,\mathrm{jets}}$ and $\sigma_{Z/W+1\,\mathrm{jets}}$ are computed to NLO, while LO matrix elements (ME) are used for $\sigma_{Z/W+2\,\mathrm{jets}}$. Provided the jets are energetic and well-separated, this serves as a reasonable approximation to the multijet cross-section. However, a parton shower is needed to extend the jet multiplicity beyond that, to describe collinear and infrared emissions, and to provide a natural entry point into hadronisation models. The basics of the parton shower formalism is outlined in Sect. 3.1. We instead now turn to the other key component of the factorisation theorem, the parameterisation of the distributions of momenta within the proton.

2.3 Probing the structure of the proton

When the proton is probed at energies higher than 1 GeV, it ceases to behave like a point particle and must be treated as a composite object. Kinematically, the proton is parameterised by PDFs that give the likelihood for a given parton to carry fraction x of the proton's momentum at a particular energy scale Q. The momentum fraction is also known as Bjorken-x [52] and was first introduced in deep inelastic scattering[2] (DIS) experiments with the definition

$$x \equiv Q^2/2p \cdot q, \tag{2.6}$$

where p is the hadron's 4-momentum and q is the 4-momentum transfer in the interaction being considered. Note that the relation $-q^2 = Q^2$ holds. In the high energy limit, Bjorken-x corresponds to the fraction of the proton's momentum carried by the parton [32].

PDFs are not derived perturbatively, but obtained from global fits to experimental data from processes such as DIS, DY, and jet production and then evolved to higher energy scales using Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [53–55]. Extensive and precise DIS data have been collected at fixed-target experiments at SLAC, FNAL, CERN, and from the HERA collider at the DESY complex in Hamburg [56].

[2] Deep inelastic scattering refers to hadron-lepton collisions at high momentum transfer.

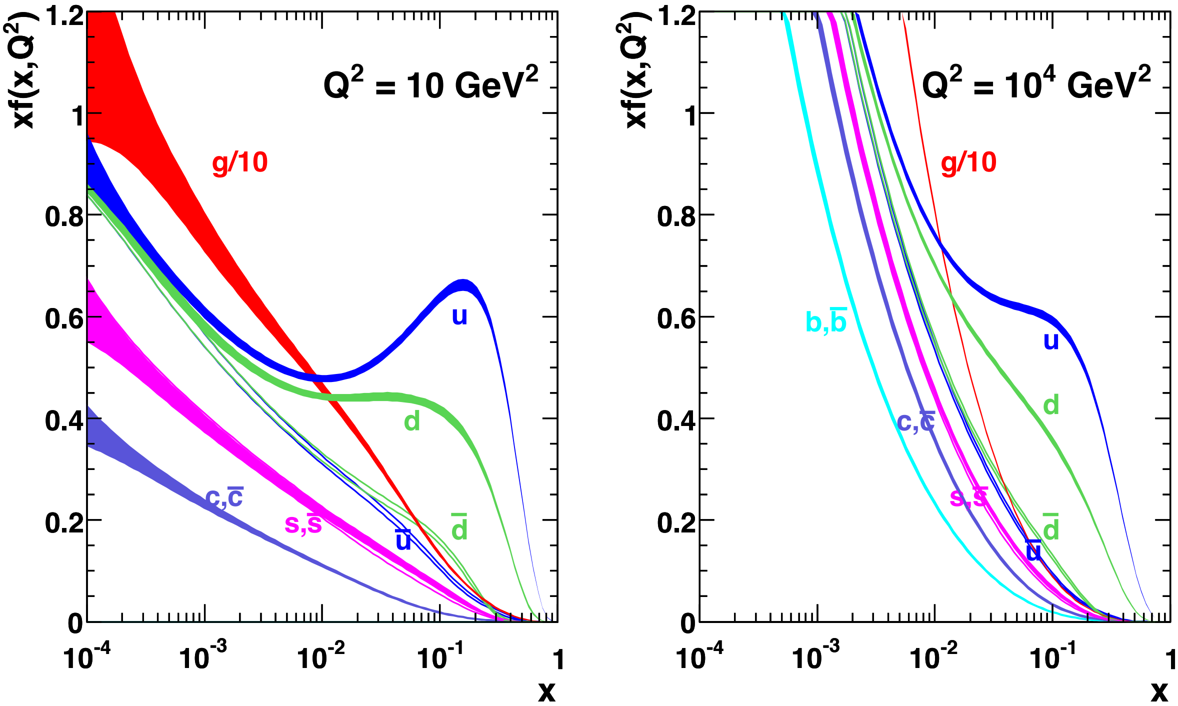


Figure 2.6: NLO proton PDFs produced by the MSTW collaboration in 2008 at two energy scales. The coloured areas correspond to the 68% confidence-level [56]. These demonstrate that the proton is mostly comprised of low x gluons with local maxima for u and d at high x.

More recent DY and jet data are provided by the Tevatron and the LHC experiments [4]. Fig. 2.6 provides an example of these distributions as determined by the MSTW collaboration [56].

As Eq. 2.2 makes clear, the precision of a $pp \to X$ cross-section calculation is dependent on the precision of the partonic cross-section and the knowledge of PDFs. $\hat{\sigma}_{Z/W}$ is known to NNLO in QCD and NLO in EW theory with an accuracy at the level of 1% [45, 57–59]. Therefore, these measurements act as "standard candles" that can be used to probe PDFs.

At large rapidities, *e.g.* $y \approx 4$, the $pp \to W$ process is dominated by valence-sea scattering with

$$u\bar{d} \to W^+ \qquad \bar{u}d \to W^-, \tag{2.7}$$

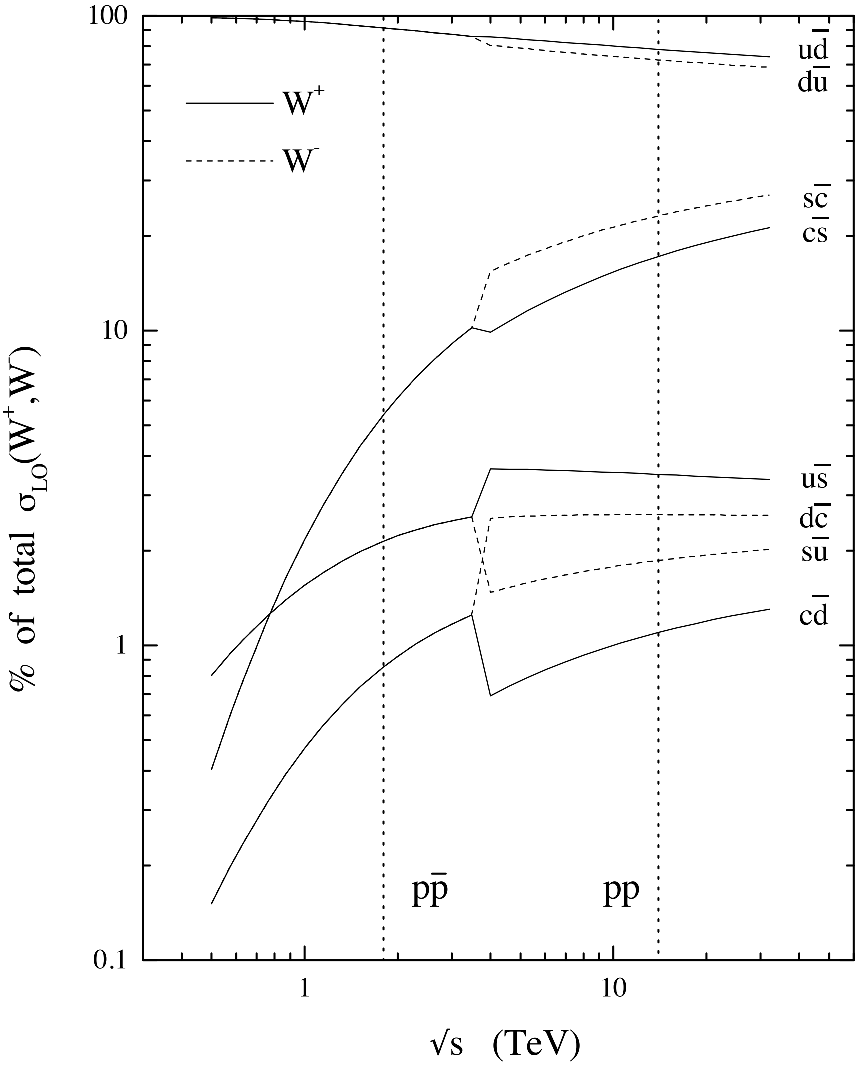


Figure 2.7: Flavour decomposition of the W boson cross-section as a function of $\sqrt{s}$ [TeV]. The vertical dashed lines correspond to the approximate centre-of-mass energies of the Tevatron (a $p\bar{p}$ collider) and the LHC (a pp collider). Figure taken from Ref. [60].

while at low rapidities, even at $y \approx 2$, scattering between pairs of sea quarks also contributes significantly. Furthermore, $\sim 10\%$ of the total is associated with

$$c\bar{s} \to W^+ \qquad \bar{c}s \to W^-. \tag{2.8}$$

These sea-sea processes dominate over Cabibbo suppressed inter-generational modes that contribute at the percent level as illustrated in Fig. 2.7.

Since $\sigma_{Z/W}$ is sensitive to the flavour composition of the quark sea, it can be used in constraining the strange PDF. These measurements are particularly valuable as the

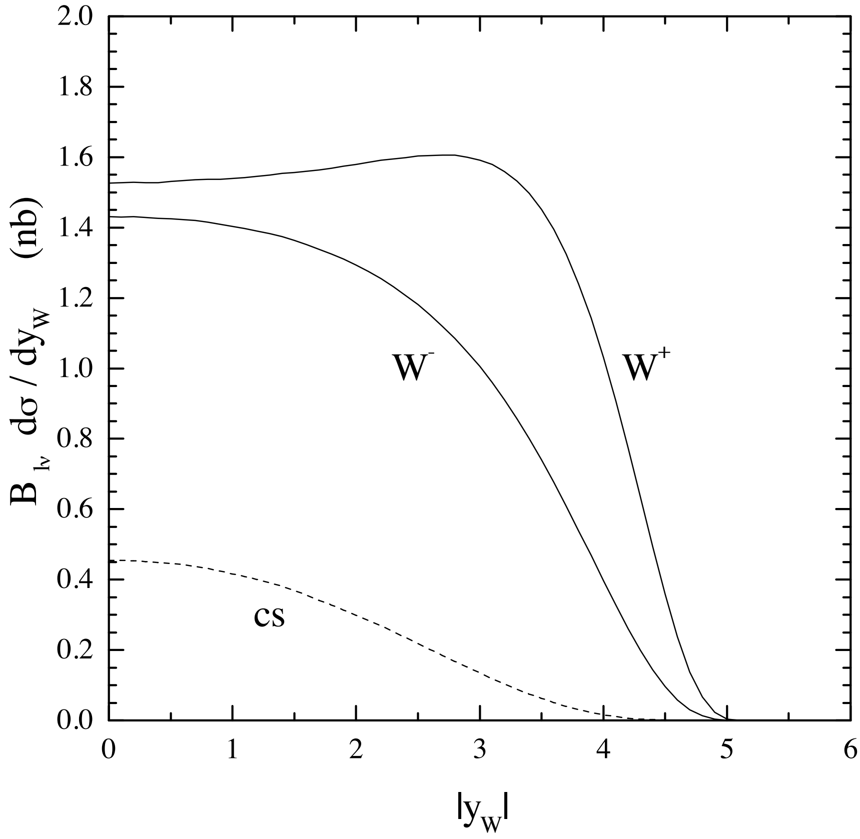


Figure 2.8: The product of the leptonic branching fraction and the differential cross-section for W production as a function of $|y_W|$ with the dashed line demonstrating the contribution from charm-strange scattering. Figure taken from Ref. [60].

strange PDF is not well-known at high energy scales of $Q \sim M_{Z/W}$. Flavour SU(3) symmetry suggests that the contributions from u, d and s should be roughly equal (this is known as "light flavour democracy"). However, the strange quark may be suppressed due to its larger mass. In particular, Fig. 2.8 illustrates how the contribution from cs impacts the differential cross-section for W production as a function of rapidity. For example, Ref. [61] determined the strange-to-down sea quark density from differential measurements of $\sigma_{Z/W}$. As an alternative probe of the strange PDF, the NuTeV collaboration measured the process $\nu_\mu N \to \mu c + X$, where N is a nucleon and X are any additional particles [62]. The measurement was used to study the strange quark distributions in Refs. [63, 64] and conclude that there is an asymmetry present between s and $\bar{s}$ PDFs within nucleons.

Insights about PDFs can be gained both from the differential cross section measurements as well as from derived measurements. Firstly, one can consider the ratio of W

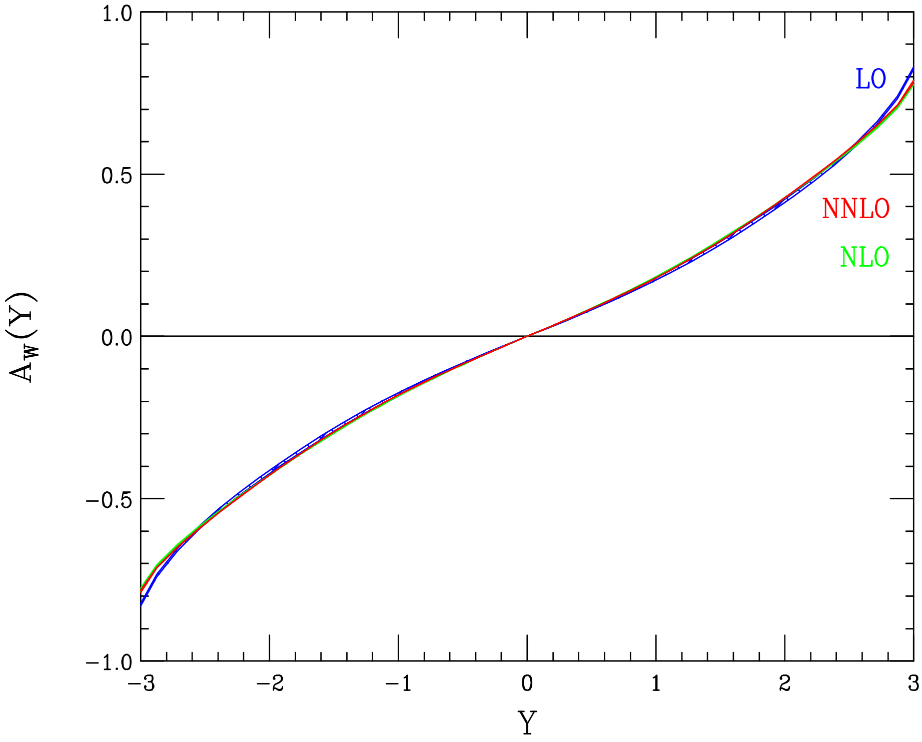


Figure 2.9: $\mathcal{A}(y_W)$ computed for $\sqrt{s} = 1.96\,\text{TeV}$. Figure taken from Ref. [65].

cross-sections defined as

$$R_W \equiv \frac{\sigma_{W^+}}{\sigma_{W^-}} \approx \frac{u}{d}, \tag{2.9}$$

where the approximate equality expresses the dominant PDFs probed. A similar quantity is the charge asymmetry that is sensitive to the sea quark distributions in the proton as previous data do not strongly constrain the ratio of u over d at low x. It is defined as

$$\mathcal{A} \equiv \frac{\sigma_{W^+} - \sigma_{W^-}}{\sigma_{W^+} + \sigma_{W^-}} \approx \frac{u-d}{u+d} \approx \frac{u_{\text{val}} - d_{\text{val}}}{u_{\text{val}} + d_{\text{val}} + 2u_{\text{sea}}}. \tag{2.10}$$

Since the u and d distributions are dominated by the valence quark contributions u_{val} and d_{val} and $u_{\text{sea}} \approx d_{\text{sea}}$, the sea quark contribution approximately cancels in the numerator, but adds constructively in the denominator. The charge asymmetry is a particularly valuable quantity as it is almost completely insensitive to higher order QCD corrections as seen in Fig. 2.9. Therefore, a measurement of this quantity is a powerful probe of PDFs.

A similar, but complementary, measurement is given by the ratio of W to Z cross-sections that is defined as

$$R_{W/Z} \equiv \frac{\sigma_{W^+} + \sigma_{W^-}}{\sigma_Z} \approx \frac{u(x_1)\bar{d}(x_2) + d(x_2)\bar{u}(x_2)}{u(x_1')\bar{u}(x_2') + d(x_1')\bar{d}(x_2')}, \tag{2.11}$$

where the PDFs in the numerator and the denominator are evaluated at different values of Bjorken-x due to the different masses of the W and Z bosons. What is more, as there are subdominant contributions to W boson production in the form of $\bar{s}c \to W^+$ and $s\bar{c} \to W^-$ which differ from the corresponding contributions to Z boson production ($s\bar{s}/c\bar{c} \to Z$), the second generation quark PDFs can be probed as well. Finally, from an experimental standpoint, measuring cross-section ratios is advantageous due to the cancellation of many systematic uncertainties as discussed in Section 5.9, most notably that of luminosity, which is the leading uncertainty for most analyses.

While ATLAS and CMS are largely limited to central pseudorapidities of $\eta \leq 2.5$, LHCb detects Z/W daughters in the forward region which are formed in a highly boosted system with respect to the lab frame. There is a correspondence between rapidity and the momentum fractions of the quarks probed within the proton that is given by

$$y_W \simeq \ln \left(\frac{x_a}{x_b}\right)^{\frac{1}{2}} \Longrightarrow x_{a,b} \simeq \frac{M_W}{\sqrt{s}} e^{\pm y_W}. \tag{2.12}$$

The pseudorapidity range probed by ATLAS and CMS corresponds to Bjorken-x in the range of $10^{-3} \leq x \leq 0.1$, while LHCb probes a system where one parton is at $x_a \sim 0.1$ and the other at $x_b \sim 10^{-4}$.

Consequently, because of the large asymmetry in the momentum fractions of the interacting partons and the high centre-of-mass energies of the LHC, LHCb is unique in its ability to probe PDFs at very low Bjorken-x and high energy scale as demonstrated by Fig. 2.10. In particular, PDF uncertainties on measurements of cross-section ratios become large at high rapidities as seen in Fig. 2.11 making LHCb ideally placed to contribute. In fact, LHCb data of W production provide the largest impact on valence d and sea light-quark PDFs as is evident from Fig. 2.12.

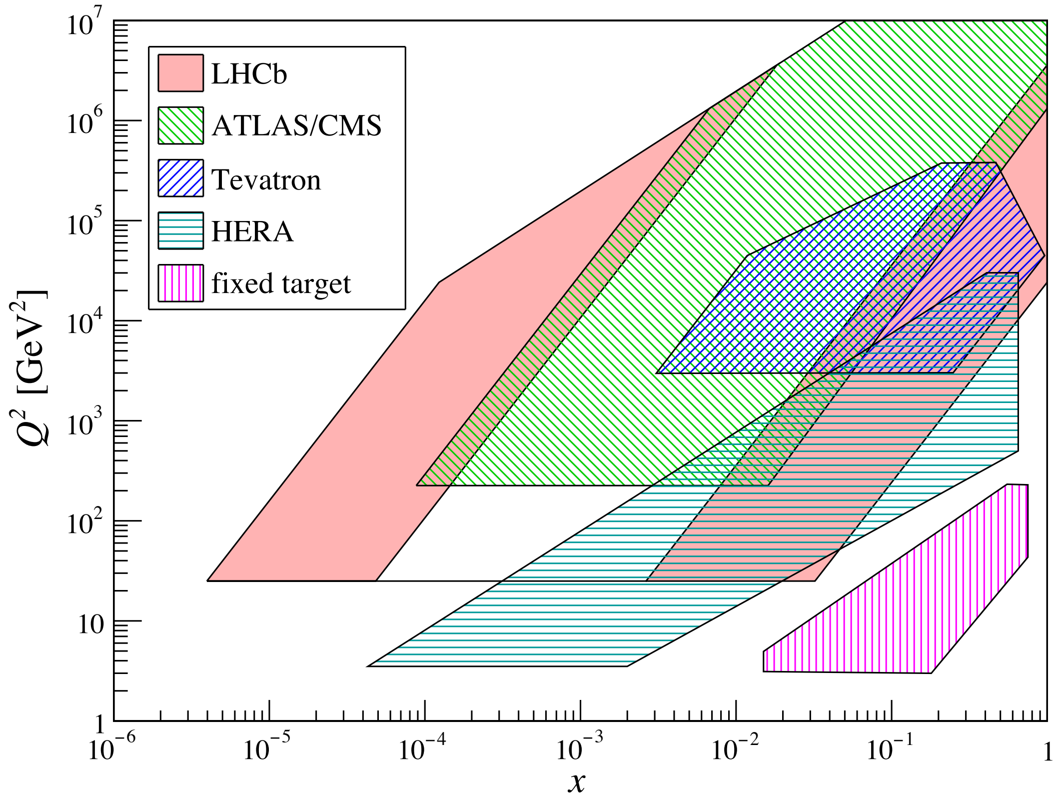


Figure 2.10: LHCb acceptance in $x - Q^2$ phase space [66].

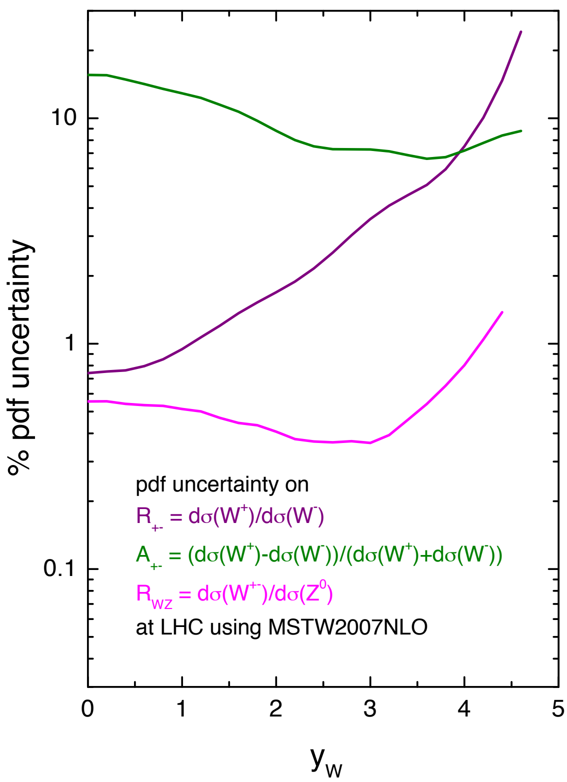


Figure 2.11: 90% confidence limit PDF uncertainties on measurements of cross-section ratios as a function of the rapidity of the W boson using the MSTW 2007 PDF set [67].

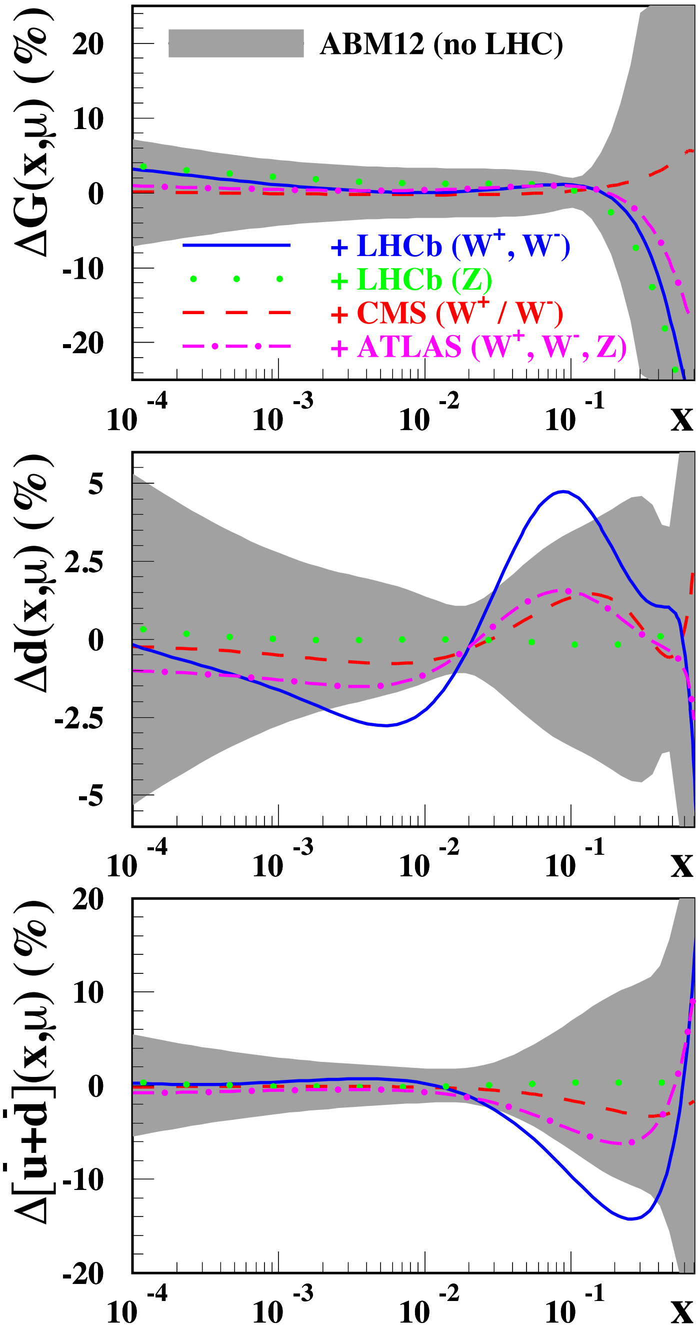


Figure 2.12: The 1σ band for the 4-flavour NNLO ABM12 PDFs versus Bjorken-x (shaded area) with the relative difference to the PDF obtained when various LHC EW measurements are included [68]. ΔG is the difference in the gluon PDF. Δd and $\Delta[\bar{u} + \bar{d}]$ probe changes in light quark PDFs.

Chapter 3

Modelling QCD alongside EW boson production

The study of EW boson production in association with jets ("Z/W+jets") in pp collisions serves as a testing ground for theoretical predictions derived from pQCD and EW theory. These processes are also a ubiquitous background to physics beyond the SM and so must be precisely measured and theoretically understood.

Two conventional theoretical tools employed in high energy physics to make QCD predictions are fixed order calculations and Monte Carlo (MC) event generators. Fixed order calculations are limited in their ability to describe event topologies with many jets. Conversely, MC event generators commonly employ the parton shower (PS) formalism which approximates parton emissions to all orders and multiplicities. This approach, however, can mismodel the number, and hardness, of jets due to its approximate nature.

This chapter concerns itself with the study of two novel theoretical approaches for rectifying the above issues: matching and merging (M&M) and weak showering (WS) [3]. M&M seeks to unite a matrix element (ME) calculation at a given order in pQCD with the PS. In contrast, WS improves the PS description of jet formation by allowing for the emission of EW bosons within the jet.

For the first time multiple M&M schemes are compared to one another to determine whether a coherent picture of QCD phenomenology emerges. As another novel result, this chapter compares these theoretical approaches to measurements performed by the LHCb experiment in the forward region at the LHC. The LHCb detector, which is instrumented in the pseudorapidity range $2.0 < \eta < 5.0$, is in a unique situation to validate the universality of these techniques and act as a complementary test to measurements performed by the ATLAS and CMS collaborations.

There are several other theoretical effects that affect the performance of these predictions such as higher order QED and EW corrections, their interplay and effect on pQCD corrections. Within the realm of MC event generators, the tuning of the PS and multiparton interactions (MPI) becomes important. Finally, uncertainties due to PDFs enter and can become particularly large at high and low values of Bjorken-x, which corresponds to the kinematic acceptance of the LHCb detector. While this chapter seeks to contrast different M&M schemes and test the WS formalism, other studies have devoted considerable effort to understanding the other effects mentioned [4, 69].

In this chapter, we present an introduction to the PS formalism in Sect. 3.1, a comparison of M&M schemes in Sect. 3.2 and an explanation of WS in Sect. 3.4. We compare M&M schemes and WS to measurements performed by ATLAS [70], CMS [71, 72], and LHCb [25] at a centre-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$. We also compare M&M schemes to a measurement of inclusive Z production by LHCb at $\sqrt{s} = 13\,\text{TeV}$ [2].

3.1 The parton shower framework

While Section 2.2 discusses the issues involved with performing a precise pQCD calculation at fixed order in perturbation theory, this section presents an alternative approach where an approximate calculation is carried out to all orders in α_s. A calculation that encompasses higher orders is particularly relevant for collinear branchings and the emission of soft partons. Splittings of the type $q \to qg$, $\bar{q} \to \bar{q}g$, $g \to q\bar{q}$ will be considered, all with QED counterparts, along with the $g \to gg$ branching which stems from the non-Abelian nature of QCD[1]. The tree-level MEs for these branchings are divergent unless virtual corrections are computed, with the exception of $g \to q\bar{q}$ which does not have an infrared divergence (but does still have a collinear divergence).

To begin, consider a hard scattering interaction that leads to $n+1$ outgoing partons. Suppose that within this process particle a branches to b and c at an angle θ. Letting $\theta \to 0$, a becomes nearly on-shell with a time scale for the branching that is considerably longer than the one characteristic of the hard subprocess. In this limit, the interaction can be described as a product of the ME for the production of n particles and the $a \to bc$

[1] While the formalism outlined in this section deals with QCD emissions, it should be emphasised that QED emissions can be handled in a similar manner.

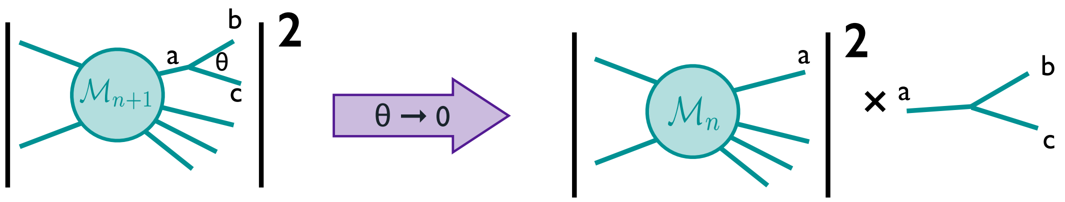


Figure 3.1: The production of $n+1$ outgoing particles can be described as the product of the n-particle ME and the $a \to bc$ branching provided $\theta \to 0$. Source: O. Mattelaer

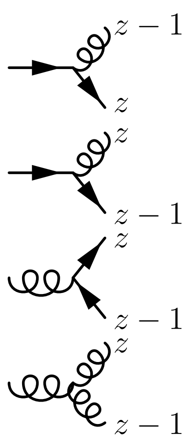
branching as illustrated in Fig. 3.1. Formally, the collinear factorisation of the interaction can be expressed as

$$|\mathcal{M}_{n+1}|^2 d\Phi_{n+1} \approx |\mathcal{M}_n|^2 d\Phi_n \frac{dt}{t} dz \frac{d\xi}{2\pi} \frac{\alpha_s}{2\pi} P_{a\to bc}(z), \tag{3.1}$$

where $\mathcal{M}$ is a ME, Φ denotes phase space, and other terms are explained in what follows.

The branchings are characterised by the energy fraction $z = E_b/E_a$ carried by one of the emerging partons, and an ordering (or evolution) variable t such that each subsequent emission has decreasing t. There is some freedom in the choice of the physical form of t. One possibility is to use the virtuality of the mother parton while other common choices include transverse momentum and the energy-weighted opening angle of emission. Here, ξ is the angle between the polarisation of a and the plane of the branching. The functions $P_{a\to bc}(z)$ are known as the DGLAP splitting kernels. The spin-averaged kernels are

$$\begin{aligned}
\hat{P}_{q\to qg}(z) &= C_F \left[\frac{1+z^2}{1-z}\right], \\
\hat{P}_{q\to gq}(z) &= C_F \left[\frac{1+(1-z)^2}{z}\right], \\
\hat{P}_{g\to qq}(z) &= T_R \left[z^2 + (1-z)^2\right], \\
\hat{P}_{g\to gg}(z) &= C_A \left[\frac{z}{1-z} + \frac{1-z}{z} + z(1-z)\right],
\end{aligned} \tag{3.2}$$



where $C_F = \frac{4}{3}$, $C_A = 3$, and $T_R = \frac{1}{2}$. Eqs. 3.2 illustrate that gluons radiate the most, and that the $g \to qq$ splitting does not have a soft divergence.

Figure 3.2: Virtual corrections and unresolvable emissions are accounted for via a Sudakov form factor which gives the probability for no branching to occur between two points along the PS evolution.

Proceeding by considering $\mathcal{M}_{n+1}$ as the core process with an additional $b \to de$ branching, the differential cross-section becomes

$$\begin{aligned} |\mathcal{M}_{n+2}|^2 d\Phi_{n+2} \approx |\mathcal{M}_n|^2 d\Phi_n \frac{dt}{t} dz \frac{d\xi}{2\pi} \frac{\alpha_s}{2\pi} P_{a\to bc}(z) \\ \times \frac{dt'}{t'} dz' \frac{d\xi'}{2\pi} \frac{\alpha_s}{2\pi} P_{b\to de}(z). \end{aligned} \tag{3.3}$$

This procedure can be extended to an arbitrary number of emissions, each with an independent branching probability.

Having established the machinery for real emissions, virtual corrections and non-resolvable emissions (as in Fig. 3.2) must also be accounted for such that unitarity is ensured and divergences are counterbalanced. This is achieved by computing a non-emission probability, known as a Sudakov form factor, that is conventionally labelled as Δ. The differential probability, $\mathcal{P}$, for a branching to not occur at scale t is

$$d\mathcal{P}_{non-emission}(t) = 1 - d\mathcal{P}_{emission}(t) = 1 - \frac{dt}{t} \int dz \, \frac{\alpha_s}{2\pi} \hat{P}(z), \tag{3.4}$$

where the singularities at $z = 0$ and $z = 1$ are conventionally regularised with explicit cut-offs. Therefore, the probability for no branchings to occur between scales t_1 and t_2 split into N infinitesimal subranges, *i.e.* the Sudakov form factor, is

$$\begin{aligned} \Delta(t_1, t_2) &= \lim_{N\to\infty} \prod_{i=0}^{N} \left(1 - \frac{dt}{t_i} \int dz \, \frac{\alpha_s}{2\pi} \hat{P}(z)\right), \\ &\simeq \lim_{N\to\infty} \exp\left(\sum_{i=0}^{N} \left(-\frac{dt}{t_i} \int dz \, \frac{\alpha_s}{2\pi} \hat{P}(z)\right)\right), \\ &\simeq \exp\left(-\int_{t_1}^{t_2} \frac{dt}{t} \int dz \, \frac{\alpha_s}{2\pi} \hat{P}(z)\right). \end{aligned} \tag{3.5}$$

With the emission and non-emission probabilities specified, the cascade of partons can be evolved down to a certain virtuality, conventionally $\mathcal{O}(1\,\text{GeV})$, and a hadronisation model is then applied. A MC event generator implements this model for final-state showers as a Markov chain *i.e.* with independent randomly generated splittings. Initial-state showers are treated in a similar fashion working backwards from the hard interaction to earlier branchings with the added complication that a PDF weight must be applied to each branching.

The PS framework thus offers a robust, physical way of treating multiple splittings to an arbitrary number of partons. The formalism can be applied to generate a timelike shower to describe jets in the final state or evolved backward from the hard interaction to model initial-state radiation. Crucially, it can be shown that the effect of a PS is unitary *i.e.* it leaves the inclusive cross-section unchanged [73]. The PS also gives an approximate estimate of the effect of higher order corrections that can become large at edges of phase space. In addition, it is a natural gateway to the non-perturbative regime of QCD where hadronisation models are employed. Finally, the method is well suited for computer implementation and forms the basis for several MC event generators.

As the PS approximation is strictly true only in the collinear limit and does not fully account for quantum interference, input from a precise ME calculation is typically needed to improve the description of well-separated jets and measures of event shape. This marriage of a ME calculation to a PS is the focus of the next section.

Matrix element	Parton shower
Good description of well-separated hard partons	Resums to all orders
Exact to given order in perturbation theory	Arbitrary particle multiplicity ⇒ ideal for description of jet formation
Quantum interference correct	Hadronisation model easily applied
Limited number of particles	Computationally cheap
Computationally expensive	Valid in collinear limit
Hadronisation model difficult to implement	Quantum interference not fully accounted for

Table 3.1: Pros and cons of ME and PS descriptions of jet formation.

3.2 Matching matrix elements to parton showers

As previous sections illuminate, ME and PS calculations have different virtues and ranges of applicability. The strengths and weaknesses of both approaches are summarised in Table 3.1. Ultimately, it is desirable to accurately describe both the hard process and the development of partons into jets. This achieved by interfacing a ME calculation and a PS. The combination would use a $Z/W + (0, 1, ..., n)$ parton ME for events with up to n hard jets and dress up the outgoing partons with radiation using the PS. Given that n is finite, additional jets must then arise from PS emissions. In this manner, a more accurate description of events with up to n jets is achieved and the precision of the inclusive cross-section is improved.

Specifically, a given $(m + 1)$-jet event, where $0 \leq m < n$, can be obtained in two ways: from the PS evolution of a $(m + 1)$-parton ME without branchings hard enough to produce a jet; or from a m-parton ME, where the PS leads to the formation of an additional jet. A particular M&M scheme defines which of the two paths should be followed on an event-by-event basis.

The potential pitfalls with this approach are double counting and the biased filling of phase space. Double counting arises when an identical final state is allowed to be generated by both a $(m + 1)$-parton ME and a m-parton ME with an extra jet from the PS. This is illustrated for increasing multiplicities in Fig. 3.3. The solution is to introduce a phase-space separation, referred to as a matching scale, that vetoes PS

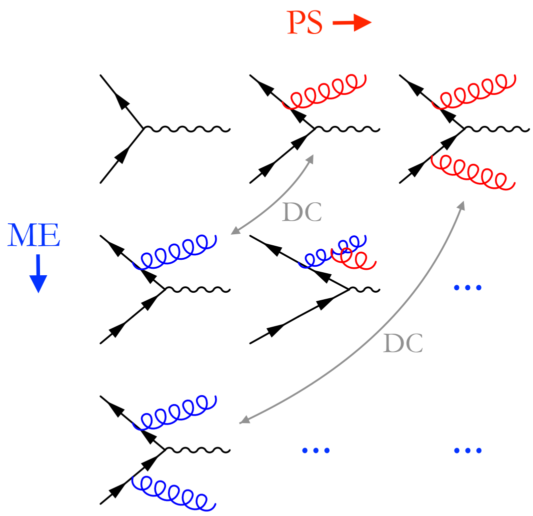


Figure 3.3: Branchings in the PS (in red) and real emissions in a ME calculation (in blue) can lead one to double count (DC) configurations with the same number of outgoing particles.

emissions above a certain threshold. In other words, the PS is only allowed to generate unresolved emissions for multiplicities for which there is a ME calculation available.

In the study presented in this chapter, the ME-based sample was required to have jets with transverse momentum, $p_{\mathrm{T}}^{\mathrm{jet}}$, in excess of 10 GeV, where the jets are clustered using the longitudinally-invariant k_{T}-clustering algorithm for hadron-hadron collisions [74]. The k_{T}-scheme defines two final-state particles i and j as belonging to separate jets if their relative transverse momentum squared,

$$k_{\mathrm{T}ij}^2 = 2 \min\left(p_{\mathrm{T}i}, p_{\mathrm{T}j}\right)^2 \frac{(\cosh(\eta_i - \eta_j) - \cos(\phi_i - \phi_j))}{R^2}, \tag{3.6}$$

is larger than $(10\,\mathrm{GeV})^2$, where R is a parameter that controls the size of the jet, η is the pseudorapidity and ϕ is the azimuthal angle. The threshold effectively controls the relative weights given to the ME calculation and the PS. A low threshold enhances the contribution from the ME while a high scale gives a bigger role to the shower approximation.

One must also make sure that there is a smooth transition along the phase space boundary between the ME calculation and the PS. This is achieved by two modifications to the ME. Firstly, the ME must be reweighted for the value of α_s that would have been used had the configuration been generated by a PS, which uses a running value of α_s. Secondly, the PS employs Sudakov factors for the non-emission probability. This requires

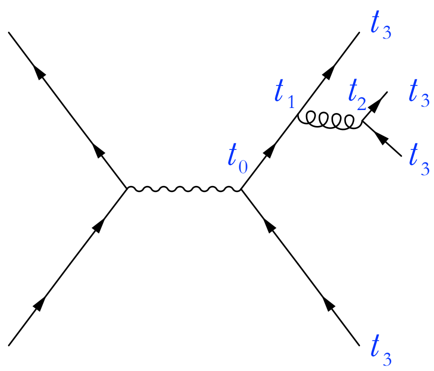


Figure 3.4: A $2 \to 4$ process with gluon emission and subsequent splitting to illustrate the relevant scales for α_s reweighting and Sudakov suppression of the ME.

a pseudo-PS history to be created for each ME-based event and a Sudakov reweighting applied to it. This history interprets a particular final state as the result of a sequence of branchings in a PS. Clearly, there is no unique history for a given event so all histories must be generated and one chosen probabilistically. For the configuration shown in Fig. 3.4 with a starting scale of t_0 and emissions at t_1 and t_2, the ME becomes

$$|\mathcal{M}|^2 \to |\mathcal{M}|^2 \frac{\alpha_s(t_1)}{\alpha_s(t_0)} \frac{\alpha_s(t_2)}{\alpha_s(t_0)} \left(\Delta_q(t_3, t_0)\right)^2 \Delta_g(t_2, t_1) \left(\Delta_q(t_3, t_2)\right)^2 , \tag{3.7}$$

where t_3 is the endpoint of the evolution in this example. Sudakov suppression also serves to make MEs exclusive such that different parton multiplicities can be added *i.e.* merged.

Several prescriptions have been proposed to resolve the issues of avoiding double counting and ensuring a smooth transition between hard and soft regimes. Additionally, these approaches differ in their treatment of higher orders in the ME calculation. One strategy is to extend real emissions to higher orders while approximating virtual/loop corrections with shower Sudakov factors. Alternatively, NLO MEs can be used for the first few emissions. The CKKW [75], CKKW-L [76–79], MLM [80], and UMEPS [73] schemes employ LO MEs while NLO MEs are used in MINLO [81], FxFx [82] and UNLOPS [83]. Finally, NNLO MEs have been employed for DY and Higgs physics in the UN^{2}LOPS scheme [84, 85].

What follows is an overview of the different M&M schemes that were included in this study with emphasis on the differences between them.

3.2.1 The MLM scheme

Two earlier M&M schemes, CKKW and CKKW-L, examine individual branchings within a PS and reject those with $t_i > t_{\rm MS}$ for branching i and matching scale $t_{\rm MS}$ to establish a phase space division. However, this form of truncated showering requires a modification of the showering algorithm. The innovation of the MLM scheme is to evolve the PS from t_0 to the hadronisation scale, form jets and reject the event if a shower-initiated jet exceeds the matching scale. The scheme is attractively simple by enabling matching to any shower algorithm without requiring modifications. However, the Sudakov suppression of the ME is, as a consequence, imprecise since a product of Sudakov factors that the jet represents is substituted for the factor associated to a single branching.

3.2.2 Unitarised matrix element + parton shower merging

As previously mentioned, it can be shown that the PS does not modify the total cross-section - a property known as PS unitarity. However, since there is a only a partial correspondence between m-jet MEs and an equivalent sequence of approximate splitting kernels, the total cross-section is modified in the CKKW-L M&M scheme. Specifically, the PS is generally correct to LL or NLL accuracy. Therefore, a residual mismatch remains for further subleading terms.

UMEPS resolves this mismatch by explicitly integrating over the phase space of a jet in the m-jet sample to arrive at a virtual correction to the $(m-1)$-jet sample. This enables the scheme to better describe unresolvable emissions at any multiplicity and enforces a cancellation between real and virtual terms for any multiplicity. In this manner the unitarity of the inclusive cross section is enforced. The cancellation of unresolved and resolved contributions also significantly reduces the scheme's dependence on the choice of matching scale.

3.2.3 The FxFx scheme

If all jets in a process are the result of PS emissions, one obtains an inclusive cross-section which is LO+LL accurate with LL accurate exclusive observables[2]. The purpose of

[2]If the definition of a given observable explicitly involves m jets, it is considered exclusive in m jets and inclusive in $n-m$ jets with $0 \leq m \leq n$. Otherwise, the observable is typically called fully inclusive.

tree-level M&M schemes is to improve the description of exclusive observables with LO MEs. The aim of the FxFx scheme is to promote exclusive observables to NLO+LL accuracy. This is advantageous as the precision of collider phenomenology is improved and the scale uncertainty is greatly reduced. In the FxFx scheme, the m-parton cross-section is calculated using the MC@NLO procedure [86] and the matching to the PS is handled according to the MLM procedure. Unlike other schemes, the matching scale can be replaced by a smooth monotonic function (effectively, a matching range).

3.2.4 Merging NLO matrix elements with parton showers

Multi-jet tree-level merging improves event shapes but cannot give a correct overall normalisation or decrease the uncertainty deriving from scale variations. The UNLOPS method aims to have a NLO accurate description of exclusive m-jet final states while also ensuring the unitarity of the overall cross-section in a manner inspired by the UMEPS scheme. In this spirit, once a one-jet NLO calculation is added, its integrated version is used to correct zero-jet events and similarly for all multiplicities that MEs are available for. The precision of the method is further improved by including tree-level MEs for multiplicities beyond those for which an NLO calculation is available. These multiplicities are handled in a manner similar to that of UMEPS. Analogously to UMEPS, variations caused by a dependence on the matching scale are minimised due to the "subtract what you add" principle. Finally, it is also worth pointing out that the scheme can be straightforwardly promoted to NNLO accuracy (termed UN^{2}LOPS) as is established in Refs. [84, 85].

3.2.5 The appropriate choice of PDF

A NLO ME calculation gives an enhancement of a factor $\ln(1/x)$ that is compensated by a commensurate term in a NLO PDF. It is crucial that the order of the ME and the PDF match, particularly for small x where $\ln(1/x)$ is large. Therefore, a LO PDF is used for the PS and MPI.

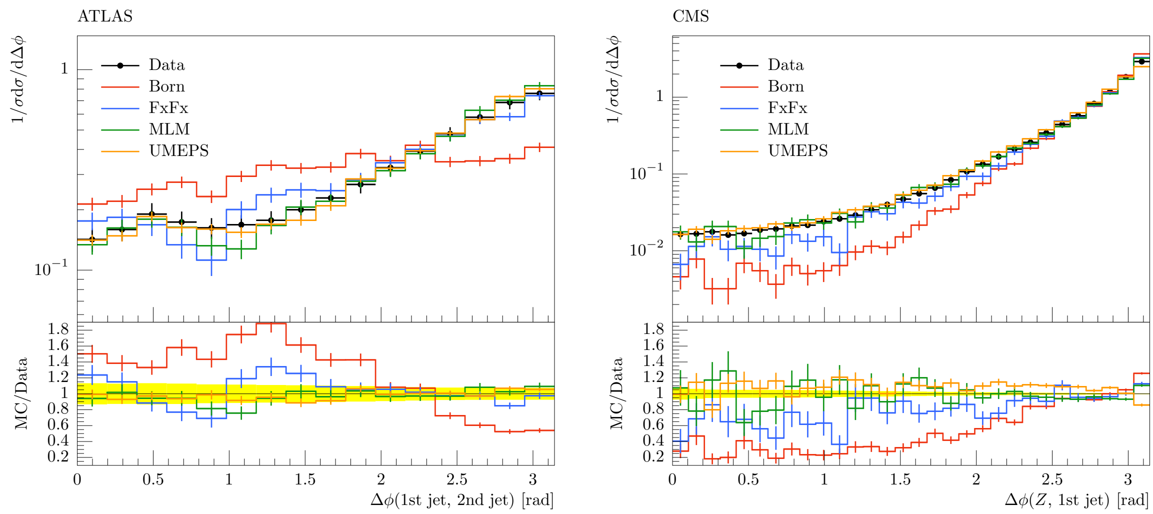
3.3 Results on matching and merging

This study is carried out using PYTHIA 8.201 [87]. Events based on MEs are generated externally using the aMC@NLO framework as implemented in MADGRAPH 5.2.3.2.2 [88] and interfaced to PYTHIA using LHE-files [89]. The NLO M&M schemes have a NLO CT10 PDF set [90] in the ME-level calculation, while the LO CTEQ6L1 PDF set [91] is employed for LO M&M schemes. The CTEQ6L1 PDF set is also used for the PS and for MPI within PYTHIA. α_s is set to 0.118 at the Z mass in the ME calculation and a value of 0.1365 is used for modelling the remainder of the event. The matching scale is set at 10 GeV. All comparisons to LHC measurements are carried out within the RIVET framework [92].

The standard LO description of single weak gauge boson production is referred to as the "Born description" in what follows. MLM is LO up to 2 jets; FXFX is NLO to 1 jet; UNLOPS is NLO to 1 jet and LO up to 2 jets. Higher jet multiplicities are generated by the PS. In the central region the predictions are compared to measurements by ATLAS [70] and CMS [71] at $\sqrt{s} = 7$ TeV. At the same centre-of-mass energy, a LHCb measurement of Z+jets [25] is used to gauge agreement in the forward region. Finally, the predictions are compared to a measurement of inclusive Z production in the forward region at $\sqrt{s} = 13$ TeV [2].

Fig. 3.5 illustrates that M&M schemes clearly outperform the Born description in modelling event shapes such as the angle between jets or between the Z boson and the jets. No significant differences between the M&M schemes are observed apart from some tension between FXFX and the other schemes.

Figs. 3.6 and 3.7 (a) illustrate that certain regions of phase space are poorly modelled in the forward region by the Born description. Discrepancies are seen in the first bin of Fig. 3.6 (a), which corresponds to events with multiple hard jets, but low Z boson p_{T}, p_{T}^{Z}. Also, the slope of the differential cross-section as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ is incorrect. Due to momentum conservation, mismodelling of $p_{\mathrm{T}}^{\mathrm{jet}}$ also has an impact on p_{T}^{Z} and, by extension, the p_{T} of the daughter leptons. Finally, the Born description predicts a larger number of pencil-like event topologies when compared to the measurement and to the predictions from M&M schemes as seen in Fig. 3.7 (a). Fig. 3.7 (b) gives the difference in rapidity between the Z boson and the leading jet with no significant differences seen between the predictions. Overall, it is clear that M&M schemes offer a superior description of event topologies with multiple hard jets in the forward region.

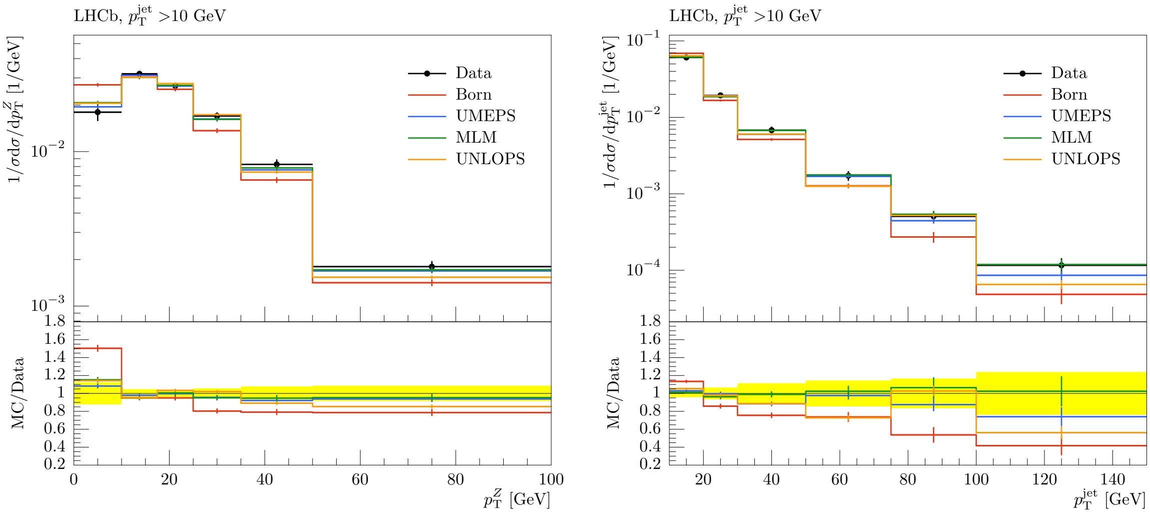


(a) (b)

Figure 3.5: The normalised differential cross-section for Z+jets production in leptonic final states as a function of the (a) azimuthal angle between the leading and subleading jet and (b) the Z boson and the leading jet as measured by (a) ATLAS [70] and (b) CMS [71]. Data points with full uncertainties are compared to the Born description in red and a selection of M&M schemes. Simulation uncertainties are statistical only.

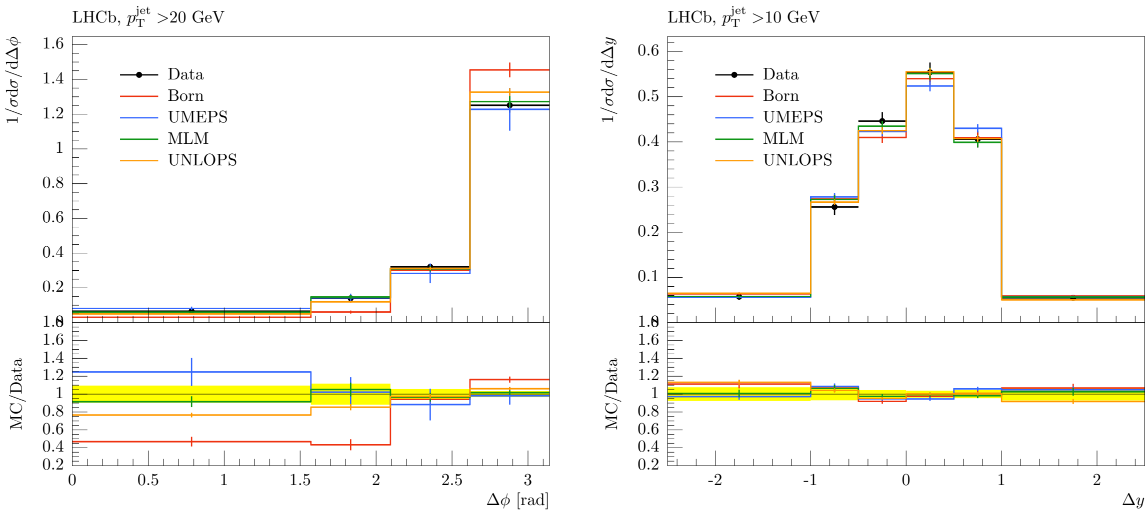
Fig. 3.8 illustrates that M&M schemes model the ϕ^*_η distribution[3] and p_{T}^Z well. UNLOPS is slightly discrepant due to the lack of a dedicated tune for NLO M&M methods in Pythia. Finally, it is worth noting that none of the M&M schemes have the correct overall cross-section in the forward region with discrepancies at the 5% level. This suggests corrections beyond NLO are crucial and UN^{2}LOPS should ideally be used.

[3] Relying solely on angles, ϕ^*_η acts as a proxy for transverse momentum [93]. It is defined as $\phi^*_\eta \equiv \tan\frac{\pi-|\Delta\phi|}{2}/\cosh\frac{\Delta\eta}{2} \simeq p_{\mathrm{T}}/M$, where M and p_{T} refer to the lepton pair, $\Delta\eta$ and $\Delta\phi$ are the differences in pseudorapidity and azimuthal angles respectively between the leptons.



(a) (b)

Figure 3.6: The normalised differential cross-section for Z+jets production in the muon final state as a function of the (a) transverse momentum of the Z and (b) the transverse momentum of the leading jet as measured by LHCb [25]. Data points with full uncertainties are compared to the Born description in red and a selection of M&M schemes. Simulation uncertainties are statistical only.



(a) (b)

Figure 3.7: The normalised differential cross-section for Z+jets production in the muon final state as a function of the (a) azimuthal angle and (b) the rapidity between the Z boson and the leading jet as measured by LHCb [25]. Data points with full uncertainties are compared to the Born description in red and a selection of M&M schemes. Simulation uncertainties are statistical only.

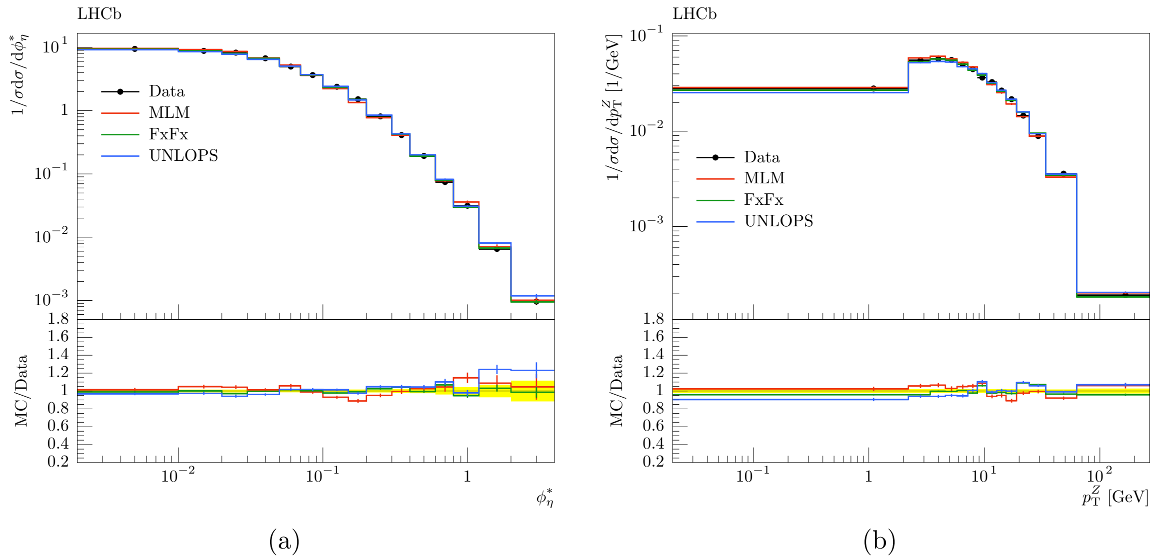


Figure 3.8: The normalised differential cross-section for inclusive Z boson production in the muon final state as a function of (a) ϕ^*_η and (b) the transverse momentum of the Z as measured by LHCb [2]. Data points with full uncertainties are compared to a selection of M&M schemes. Simulation uncertainties are statistical only.

3.4 Weak gauge boson emission in parton showers

As an alternative to the previously presented methods, WS approaches Z and W production from a novel viewpoint. In a Born description of weak gauge boson production the hardest scale in the event is associated with creating the bosons and any jet activity is modelled by means of a PS originating from initial- or final-state radiation, and MPI. This, however, will typically underestimate the number, and hardness, of associated jets. WS instead starts from a QCD process, most commonly dijet production, and allows for the emission of Z and W bosons within the development of the PS [94]. For this purpose, appropriate splitting kernels must be implemented and the helicity of the weak boson daughters accounted for. An additional complication is the relatively high mass of EW bosons which leads to kinematic constraints. In PYTHIA, energy-momentum conservation is achieved by treating MPI, initial- and final-state radiation together such that all features of the event are in direct competition for phase space [95].

The Born description and WS can be added to populate the phase space more fully and to achieve a better description of EW physics in an hadronic environment with centre-of-mass energies that can greatly exceed the EW scale. While these two production paths are expected to preferentially populate different regions of phase space, it is clear that their combination can result in double counting. Since both production modes can lead to topologies with the same number of final state jets, care must be taken to veto events which are better described by the competing production path. This is achieved by employing the k_{T}-clustering algorithm on all final-state particles, including weak bosons. Events with bosons lying close to jets in a Born description and, correspondingly, events with bosons that are separated from jets in the QCD-initiated path are vetoed.

Extending the PS machinery to include EW boson emissions brings several improvements to the simulation of high-energy pp collisions. Experimentally, weak bosons are seen to be present in jets from their leptonic signatures and the emission probability increases as a function of the centre-of-mass energy of the collision. Weak emission, therefore, can lead to a more realistic description of jet substructure which complements the many ongoing experimental efforts to study the internal structure of jets and formulate discriminating variables based on this [96].

Secondly, weak corrections are required to reproduce the jet rate suppression expected at $p_{\mathrm{T}}^{\mathrm{jet}}$ values in excess of 1 TeV. What is more, the WS machinery complements efforts in developing M&M schemes. As explained previously, M&M schemes require PS histories to be generated for Sudakov suppression terms. Allowing for EW emission in the PS leads to more realistic histories and also enables histories to be generated for processes which currently lack one.

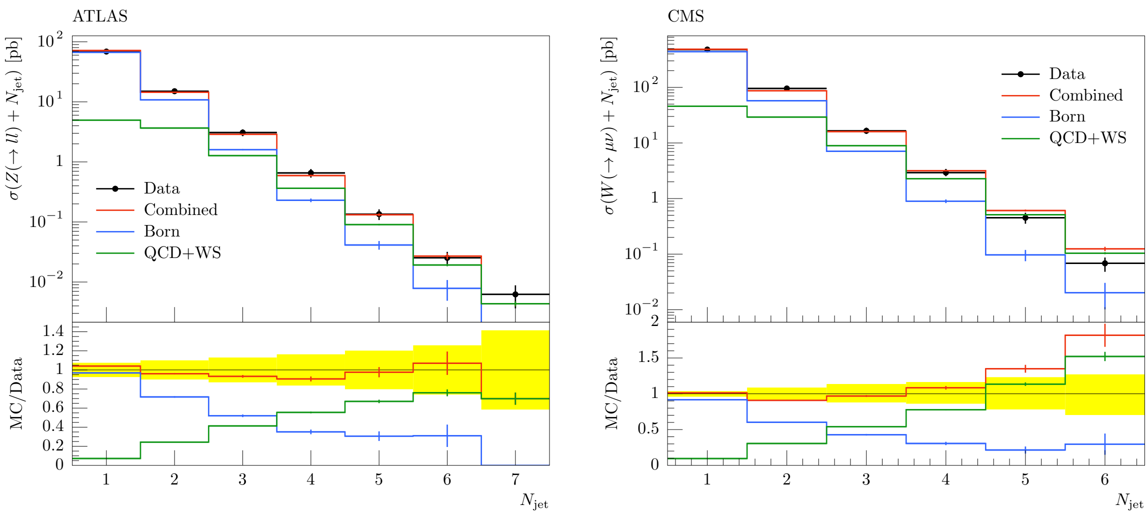
Finally, this study seeks to test if a better description of EW boson production in association with jets can be achieved with WS. In particular, WS is expected to be relevant for events with hard jets and high jet multiplicities.

3.5 Results on weak showering

Here, the Born description of Z or W production as defined in the previous section and simulation of dijet production along with weak emissions within the PS are compared to measurements made at ATLAS [70], CMS [72], and LHCb [25] at $\sqrt{s} = 7\,\text{TeV}$. PDFs are parameterised by the CTEQ6L1 PDF set. The Monash tune [97] is used which sets $\alpha_s = 0.1365$ within the PS. The Z mass is required to be above 40 GeV and PS emissions are allowed up to the kinematic limit of $p_{\text{T}} = \sqrt{s}/2$ to avoid gaps in phase space (known as "power showers").

The Born description does not yield the correct normalisation of the cross-section and, consequently, an overall scaling is applied (the so-called "K-factor"). K-factors are chosen empirically by scaling the Born contribution to agree with the data. Centrally, the K-factor for Z+jets is taken to be 1.37 based on the zero-jet bin of the exclusive jet multiplicity distribution, N_{jet}. The K-factor for W+jets is determined to be 1.28 based on the one-jet bin of the N_{jet} distribution. In the forward region, the Born contribution is scaled by 1.10 to enforce agreement with the $[10 - 20]$ GeV bin of the Z+jet cross-section as a function of $p_{\text{T}}^{\text{jet}}$. Empirically, it is seen that a LO prediction of QCD dijet production does not need a K-factor but rather an enhancement of α_s can compensate for missing higher orders [98]. Here, $p_{\text{T}}^{Z/W}$ is required to be above 1 GeV with respect to the parton in the shower to remove divergences. Double counting between the Born and WS contributions is removed as described in Sect. 3.4 with a k_{T} algorithm R parameter of 0.6.

The comparisons in Figs. 3.9-3.12 demonstrate that a significant improvement in the modelling of the data is achieved with the inclusion of WS. The number of events with a high jet multiplicity is increased and a better description of $p_{\text{T}}^{\text{jet}}$ is achieved. What is more, Figs. 3.10 and 3.11 demonstrate that the weak shower contribution becomes increasingly



(a) (b)

Figure 3.9: Differential cross-section for (a) Z+jets production and (b) W+jets production in leptonic final states as a function of the number of jets as measured by (a) ATLAS [70] and (b) CMS [72]. Data points with full uncertainties are compared to the Born description in blue, QCD dijet production with WS in green, and the sum of the two contributions in red. Simulation uncertainties are statistical only.

dominant for subleading jets. The description of event shapes, while not perfect, is also improved as seen in Fig. 3.12. This is to be expected from a LO description.

Figs. 3.9 and 3.13 demonstrate that the high N_{jet} and high $p_{\text{T}}^{\text{jet}}$ regions, respectively, are overestimated by the combined prediction. This can be attributed to the use of power showers and a value for the strong coupling, α_s^{PS}, that is significantly higher than the best-fit NLO value at the Z mass pole, α_s^{NLO}. Together these serve to compensate for missing higher orders in the calculations. However, the former involves a stretching of the PS beyond its formal region of validity and the effect of the latter grows as $\left(\alpha_s^{\text{PS}}/\alpha_s^{\text{NLO}}\right)^{N_{\text{jet}}}$. It is, therefore, not surprising that the modelling of high N_{jet} events is poor as seen in Fig. 3.9 (b) and as corroborated by Ref. [94].

It is evident that LHCb does not yet probe regions of phase space where WS becomes a significant contribution as is illustrated in Figs. 3.14 and 3.15. The Born description does not fully account for the data at high $p_{\text{T}}^{\text{jet}}$ and low $\Delta\phi$, but neither does the WS component significantly improve agreement with data. Overall, a good description of the LHCb measurement is not achieved in this case.

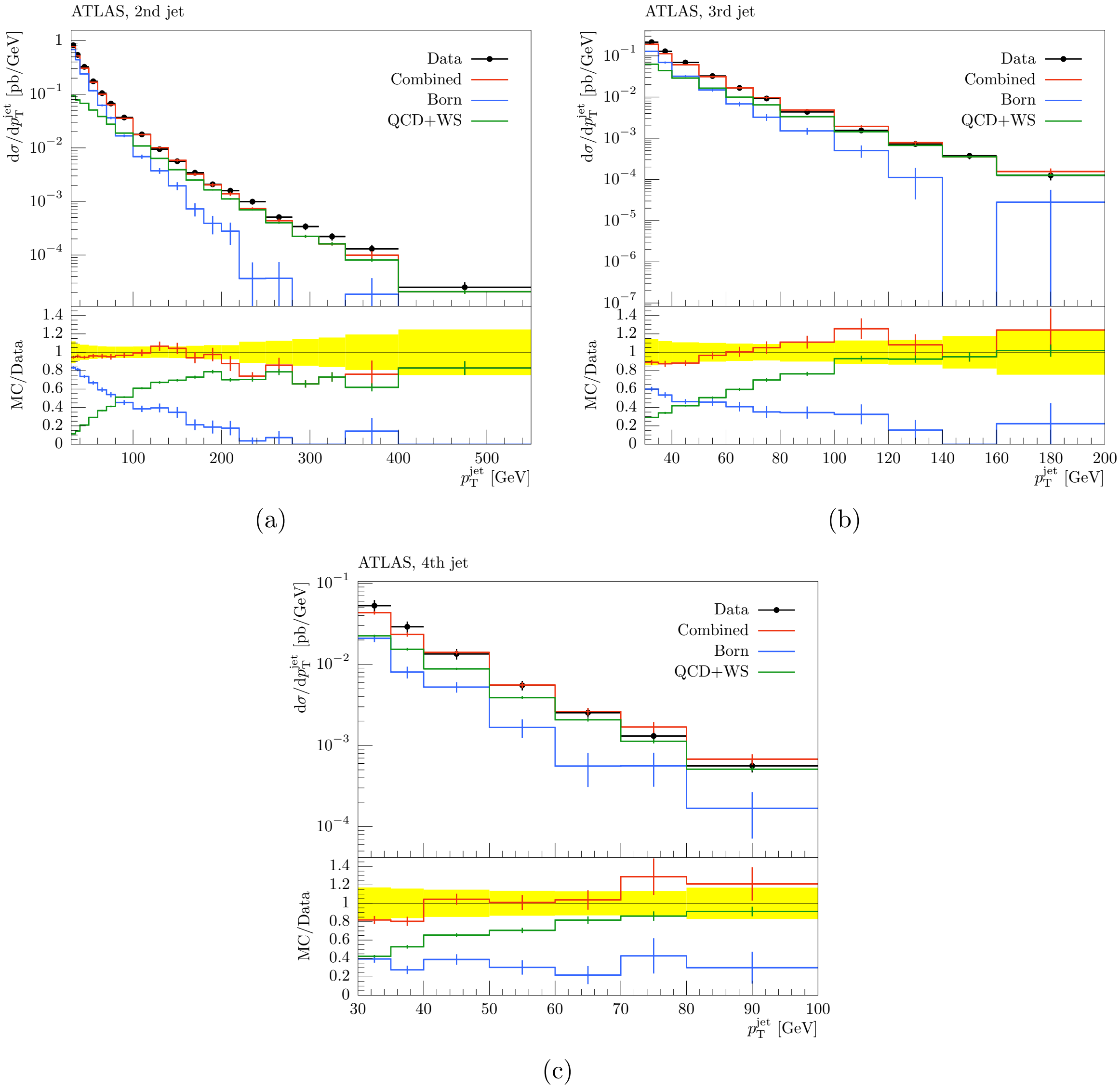


Figure 3.10: Differential cross-section for Z+jets production in leptonic final states as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ for subleading jets as measured by ATLAS [70]. Data points with full uncertainties are compared to the Born description in blue, QCD dijet production with WS in green, and the sum of the two contributions in red. Simulation uncertainties are statistical only.

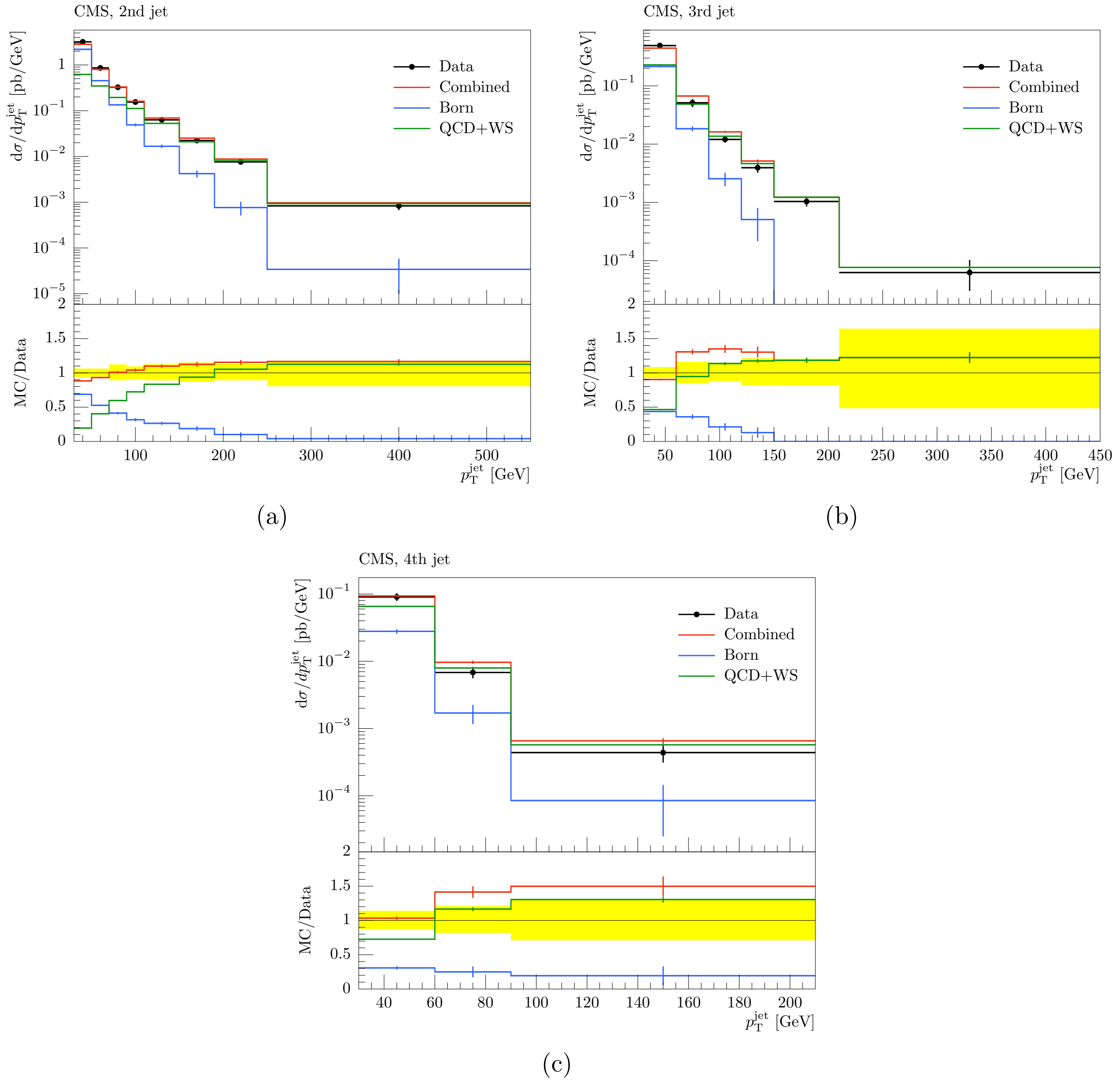


Figure 3.11: Differential cross-section for W+jets production in leptonic final states as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ for subleading jets as measured by CMS [72]. Data points with full uncertainties are compared to the Born description in blue, QCD dijet production with WS in green, and the sum of the two contributions in red. Simulation uncertainties are statistical only.

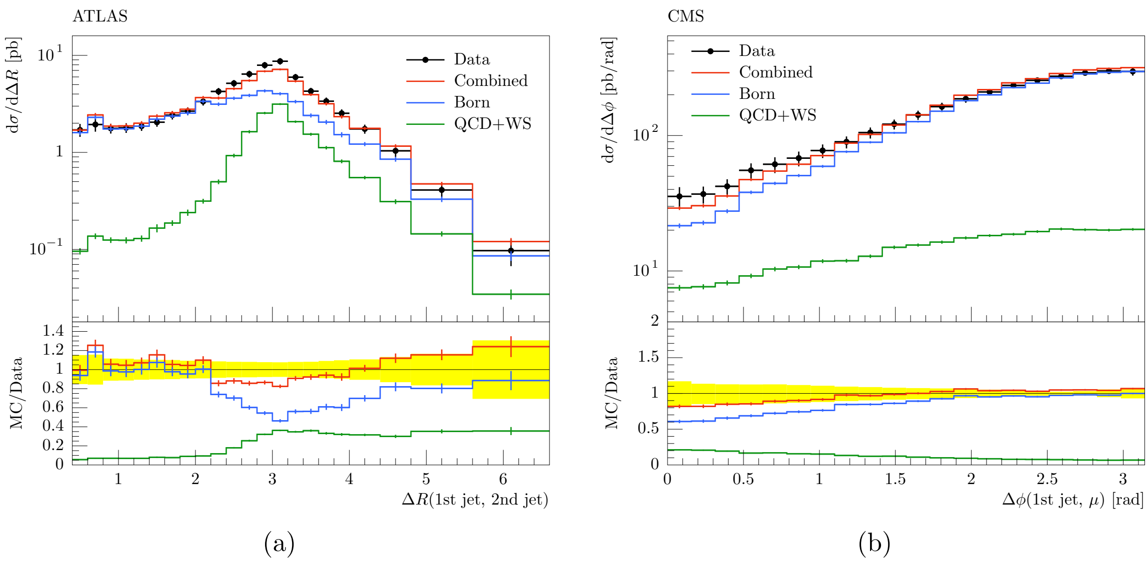
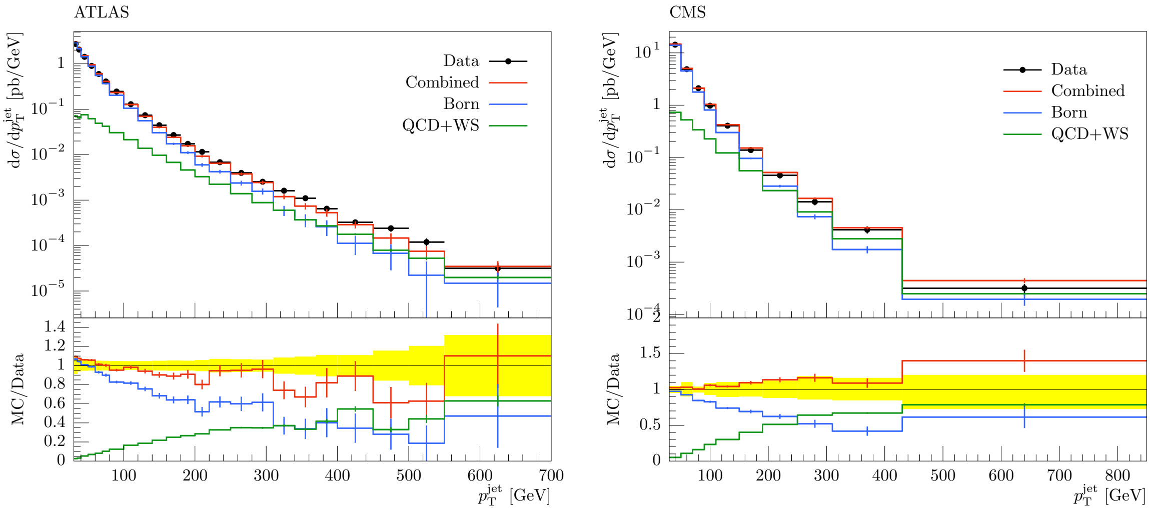


Figure 3.12: Differential cross-section for (a) Z+jets production as a function of the radial distance between jets as measured by ATLAS [70] and (b) W+jets production as a function of the azimuthal angle between the muon and the leading jet as measured by CMS [72]. Data points with full uncertainties are compared to the Born description in blue, QCD dijet production with WS in green, and the sum of the two contributions in red. Simulation uncertainties are statistical only.

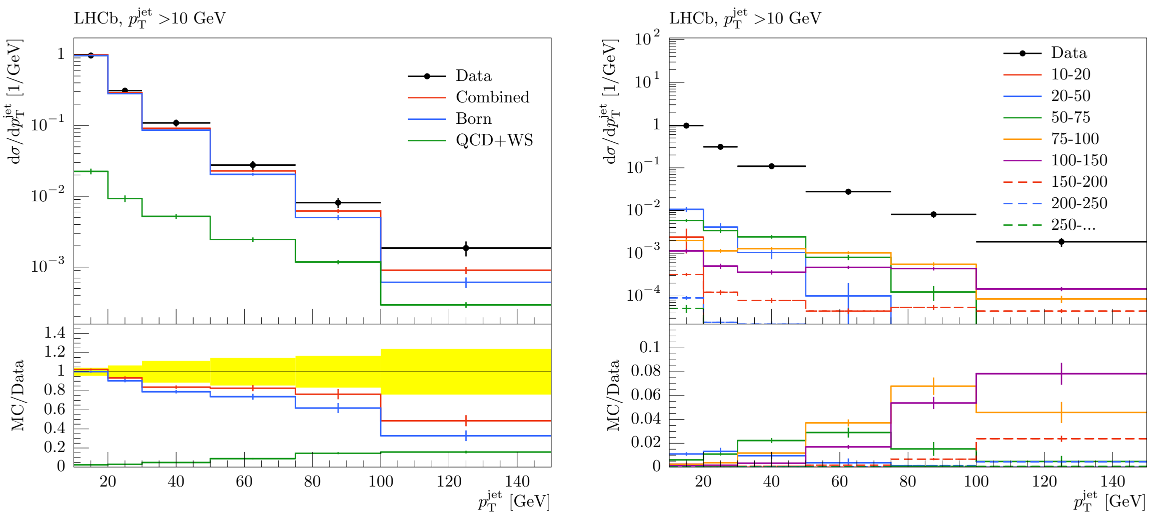
As a side note, Fig. 3.14 (b) shows individual contributions to the weak shower from ranges of outgoing partonic p_{T} in the centre-of-mass frame of the underlying $2 \to 2$ process. The figure demonstrates that relatively small values of parton transverse momentum can lead to significant subleading contributions to the WS. Specifically, the $[10-20]$ GeV range is the third most important contribution to the first bin of $p_{\mathrm{T}}^{\mathrm{jet}}$. This fact is surprising given that W and Z are significantly heavier than the scale of the process and very little phase space is expected to be available for the emission of a massive gauge boson. Emissions are made possible, however, due to the interplay between the large dijet production cross-section at these values of p_{T} and the fact that Pythia treats the phase space of the event holistically.

Finally, Fig. 3.16 shows the p_{T} spectra of W daughters in leptonic decays, p_{T}^{l}, for the Born and weak shower description. No acceptance cuts have been applied in this case. It is clear that leptons which originate from the PS have a wider Jacobian peak and have the effect of smearing the p_{T}^{l} distribution and increasing the contribution from high-p_{T}^{l} events. With higher centre-of-mass energies and increased precision, this will impact EW analyses that use fits to the p_{T}^{l} distribution to extract physical quantities such as event yields or the W boson mass.



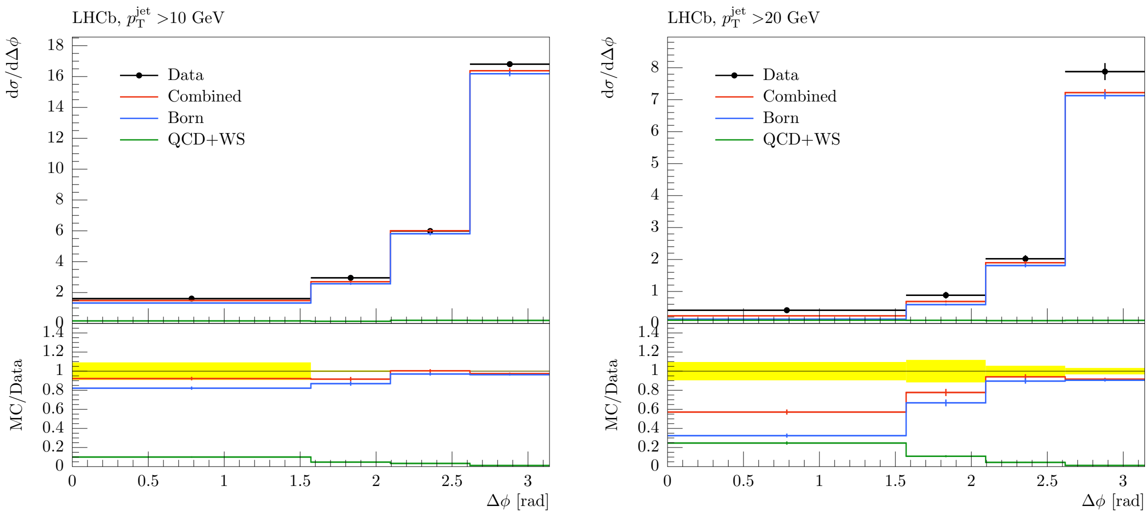
(a) (b)

Figure 3.13: Differential cross-section for (a) Z+jets production and (b) W+jets production in leptonic final states as a function of the transverse momentum of the leading jet as measured by (a) ATLAS [70] and (b) CMS [72]. Data points with full uncertainties are compared to the Born description in blue, QCD dijet production with WS in green, and the sum of the two contributions in red. Simulation uncertainties are statistical only.



(a) (b)

Figure 3.14: Differential cross-section for Z+jets production in the muon final state as a function of p_T^{jet} as measured by LHCb [25]. (a) Data points with full uncertainties are compared to the Born description in blue, QCD dijet production with WS in green, and the sum of the two contributions in red. (b) Individual contributions from ranges of outgoing parton transverse momenta in the centre-of-mass frame in dijet production are illustrated. Simulation uncertainties are statistical only.



(a) (b)

Figure 3.15: Differential cross-section for Z+jets production in the muon final state as a function of the azimuthal angle between the Z and the leading jet for (a) $p_T^{\text{jet}} > 10\,\text{GeV}$ and (b) $p_T^{\text{jet}} > 20\,\text{GeV}$ as measured by LHCb [25]. Data points with full uncertainties are compared to the Born description in blue, QCD dijet production with WS in green, and the sum of the two contributions in red. Simulation uncertainties are statistical only.

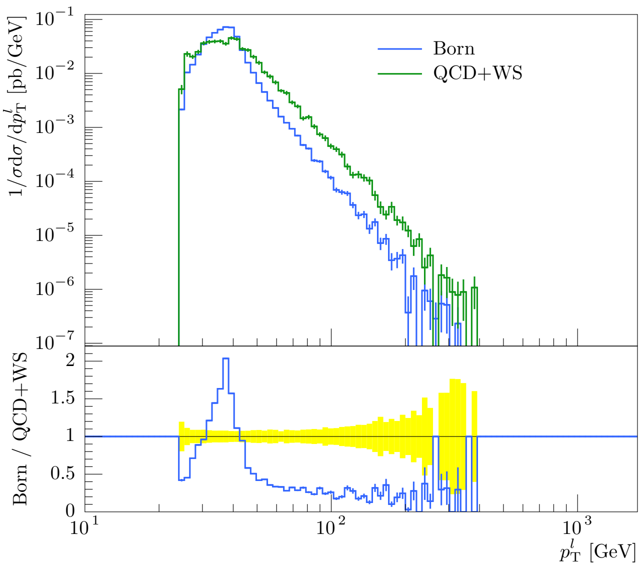


Figure 3.16: The lepton transverse momentum distribution in the Born description in blue is compared to QCD dijet production with WS in red. Uncertainties are statistical only and no acceptance requirements have been applied.

Chapter 4

The LHCb experiment and the LHC

This chapter begins by presenting a brief overview of CERN, the European Organization for Nuclear Research. The LHC, a collider designed to probe the TeV energy scale, is then discussed. This section is largely based on Refs. [99, 100]. Subsequently, one of the experiments on the LHC ring, the LHCb detector, is presented in more detail as the results presented in this thesis are largely based on data collected at LHCb and the author of this thesis has contributed to the LHCb collaboration for a period of 4 years as the EW simulation liaison[1] and as a member of the EW working group. This section is mostly based on Ref. [101].

4.1 CERN

CERN has been the leading European particle physics laboratory since its formation in 1954. Located on the Franco-Swiss border next to the city of Geneva, it consists of two large sites near Meyrin, a suburb of Geneva, and near the villages of Prévessin-Moëns and Ferney-Voltaire along with other experimental sites. While CERN's main function is to operate the particle accelerators and related infrastructure, it also has data processing facilities for experimental data analysis.

CERN has been at the forefront of particle physics since its inception. In 1974, neutral currents were discovered with the Gargamelle bubble chamber [102–104]. A decade on, the W and Z bosons were first observed in the UA1 and UA2 experiments [6–9]. The

[1]The liaison's duties are to coordinate and validate the submission of decay files, submit production requests, produce generator statistics for productions, validate new versions of code, contribute to tuning, and attend and report in simulation meetings.

number of light neutrino families were measured in 1989 with the LEP collider at an energy corresponding to the Z boson mass resonance [12]. Antihydrogen was first created in the PS210 experiment in 1995 [105] and, more recently, 38 atoms of antihydrogen were successfully isolated in 2010 [106]. Finally, in 2012 the discovery of a boson with mass around 125 GeV [34, 35], consistent with the Higgs boson, was announced which led to a Nobel prize being awarded to theorists Peter Higgs and François Englert for the development of the Higgs mechanism [36–41].

4.2 The Large Hadron Collider

CERN's premier facility, the LHC [99, 100], is housed in a 3.8-metre wide concrete-lined tunnel up to 175 m deep underground with a circumference of 27 km. The tunnel is bordered by Geneva on one side and the Jura Mountains on the other as is illustrated on a map of the area in Fig. 4.1. "LEP/LHC" indicates the fact that the tunnel was previously used to house the LEP collider. While primarily a proton-proton collider, the machine also collides lead ions for a period of time each year for the study of quark-gluon plasma among other experimental goals.

The LHC is the largest constituent of the CERN accelerator complex that is illustrated in Fig. 4.2. The purpose of the other components of the accelerator complex is to successively increase the energy of the proton beams. A linear particle accelerator, LINAC 2, is the first system to accelerate protons and is followed by the Proton Synchrotron Booster (PSB). At this point, protons have an energy of 1.4 GeV and are then injected into the Proton Synchrotron (PS) with the Super Proton Synchrotron (SPS) acting as the final pre-accelerator before the main ring. The accelerators that prepare the proton beams for the LHC were themselves used to make some of the important discoveries in particle physics over the course of the 20$^{\text{th}}$ century.

The LHC ring consists of two counter-rotating beams of protons that are accelerated using radio-frequency (RF) cavities. 1 232 dipole magnets are used to keep the beams on circular paths while 392 quadrupole magnets prevent the beams from becoming diffuse with the aim of maximising the number of collisions at the interaction points. A schematic of the cross-sectional view of the LHC beamline is shown in Fig 4.3. The magnets are kept at an operating temperature of 1.9 K using roughly 96 tonnes of superfluid helium-4. The accelerator is designed to accelerate protons to an energy of 7 TeV but is currently

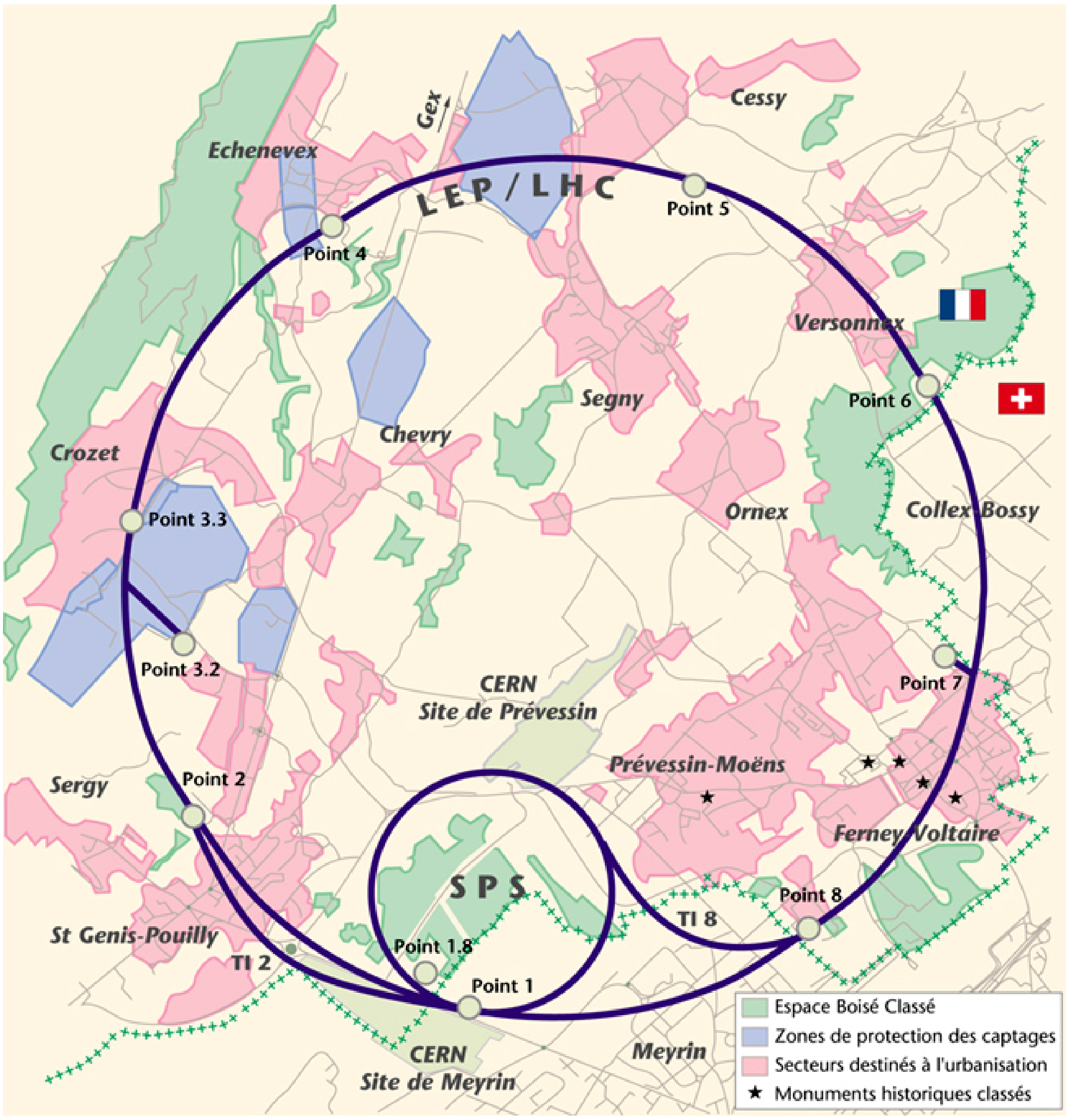


Figure 4.1: The CERN site on the Franco-Swiss border next to the suburb of Meyrin with the LHC and SPS tunnels superimposed [107].

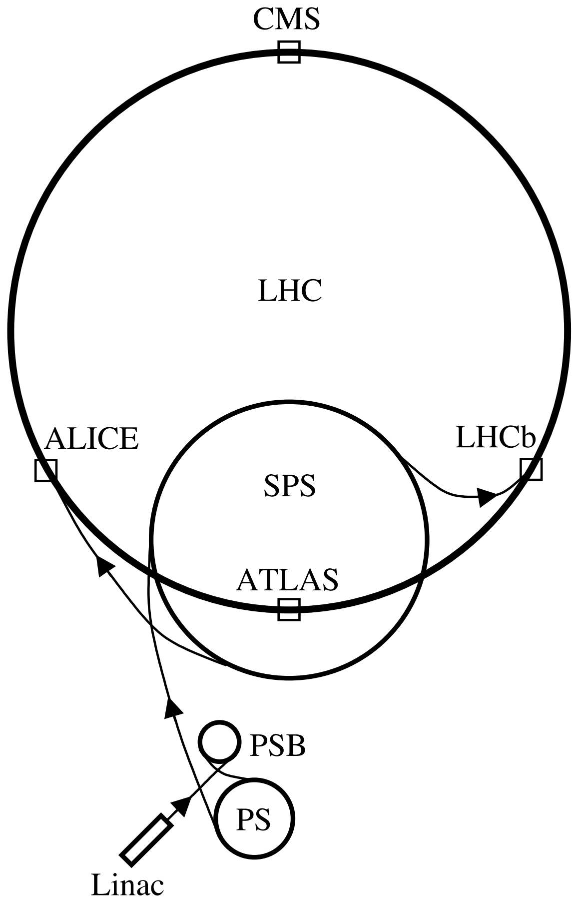


Figure 4.2: The CERN accelerator complex with the main experiments ATLAS, CMS, LHCb, and ALICE along with injector rings [108].

operating at 6.5 TeV. The results presented in this thesis are based on data collected at beam energies of 3.5, 4 and 6.5 TeV.

The protons are kept in up to 2 808 bunches with ~$1.15 \cdot 10^{11}$ protons per bunch such that collisions take place at discrete intervals of 25 ns at the nominal configuration. In 2011 and 2012, when the beam energy was 3.5 and 4 TeV, respectively, the beam spacing was set to 50 ns and around 1 400 bunches were typically filled with protons. The LHC is designed to operate at an instantaneous luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, which was first reached in June 2016.

During the 2012 data taking period the LHCb experiment operated at a reduced instantaneous luminosity of roughly $4 \cdot 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. At LHCb the luminosity is intentionally levelled by locally de-focusing the beams. The aim of luminosity levelling is to allow for the more efficient reconstruction of secondary vertices by reducing the number of

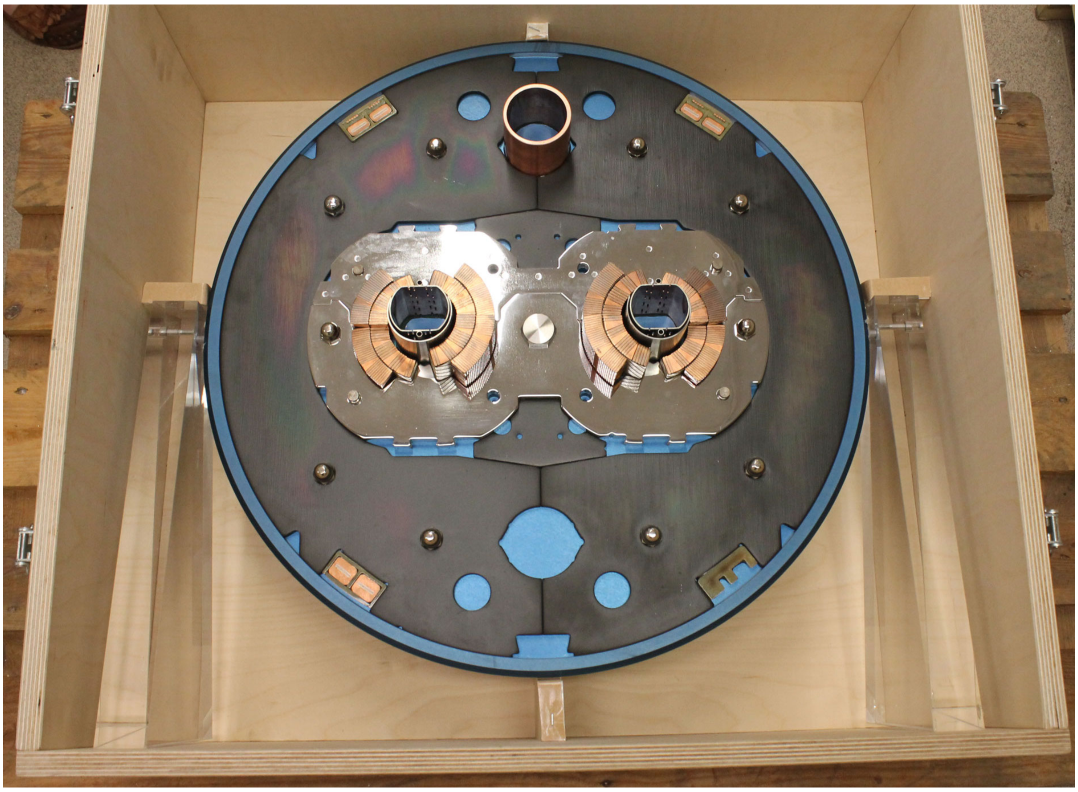


Figure 4.3: The cross-section of a LHC dipole. The beam pipe is enclosed by superconducting coils, an iron yoke and a thermal shield. The instrumentation is kept in a vacuum vessel denoted by the outer blue line. Source: H. Cliff

interactions per bunch crossing and to minimise radiation damage. Luminosity levelling results in a far more stable instantaneous luminosity delivered to LHCb than to the other experiments as illustrated in Fig. 4.4.

The beams are brought together at four interaction points with each surrounded by one of the four main detectors. Two of these, ATLAS [110] and CMS [111], are multipurpose detectors designed to search for, among other aims, New Physics such as supersymmetric particles and to measure the Higgs boson. The other two detectors are more specialised. ALICE [112] primarily studies lead ion collisions and LHCb [101, 113] was designed to carry out precise measurements in the heavy flavour sector, but has, in fact, realised a more comprehensive research programme as this thesis attests. The LHCb detector is described in more detail in the following section.

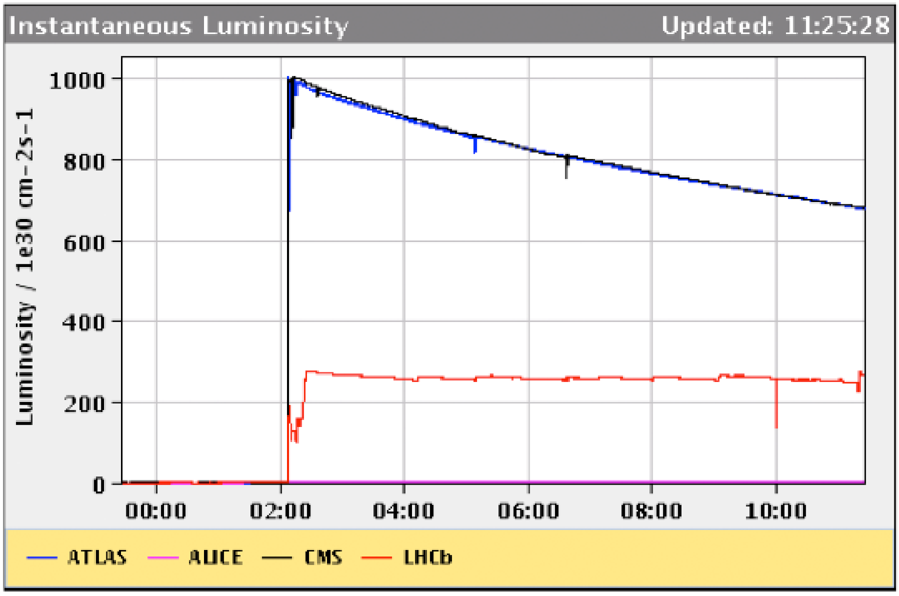


Figure 4.4: Typical evolution of the luminosity during a LHC fill in Spring 2011, with the delivered luminosity to the general purpose detectors ATLAS and CMS shown in blue and black, and the delivered luminosity to LHCb in red. The observed near constant luminosity at LHCb is obtained through "luminosity levelling" [109].

4.3 The Large Hadron Collider Beauty Experiment

The LHCb detector [101, 113] is a single-arm forward spectrometer instrumented in the forward region in the range $2.00 < \eta < 5.00$ and is shown in Fig. 4.5. While primarily designed for the study of particles containing b or c quarks, it is a versatile instrument that also affords measurements of QCD, EW theory, exotica and central exclusive production, for instance. The detector consists of multiple planar subdetectors: the Vertex Locator (VELO) is closest to the interaction point followed by the first ring-imaging Cherenkov detector (RICH) detector, the Tracker Turicensis (TT), a dipole magnet, three tracking stations (T1, T2, T3), the second RICH detector, the first muon station (M1), the calorimeter system that is made up of a scintillating pad detector (SPD), a pre-shower detector (PRS, but denoted as "PS" in Fig. 4.5), an electromagnetic calorimeter (ECAL) and a hadronic calorimeter (HCAL), and muon stations (M2-M5).

Although the acceptance of the detector is limited to ~2.5% of the solid angle, a significant fraction of B hadrons produced at the interaction point will be within its acceptance as shown in Fig. 4.6. This is due to the fact that the parton distribution functions of the initial states lead to a system that is frequently boosted with respect to the lab frame. As it is crucial for LHCb to measure B hadron lifetimes accurately, precision vertexing and track reconstruction is necessary. A resolution of $(15 + 29/p_{\mathrm{T}})\,\mu\mathrm{m}$, where p_{T} is measured in GeV, is achieved for reconstructing the minimum distance of a track to a primary vertex (known as the impact parameter). Momentum is measured with a relative uncertainty that varies from 0.5% at low momentum to 1.0% at 200 GeV.

The detector uses a right-handed coordinate system with the origin at the nominal interaction point, the x-axis pointing to the centre of the LHC ring, the y-axis pointing up, and the z-axis pointing along the beamline toward the muon system. The coordinate system is superimposed onto the detector schematic in Fig. 4.5.

What follows is a discussion of LHCb detector components. Subsequently, the LHCb trigger system and luminosity determination is presented. Finally, a discussion of the software framework and event simulation concludes the chapter.

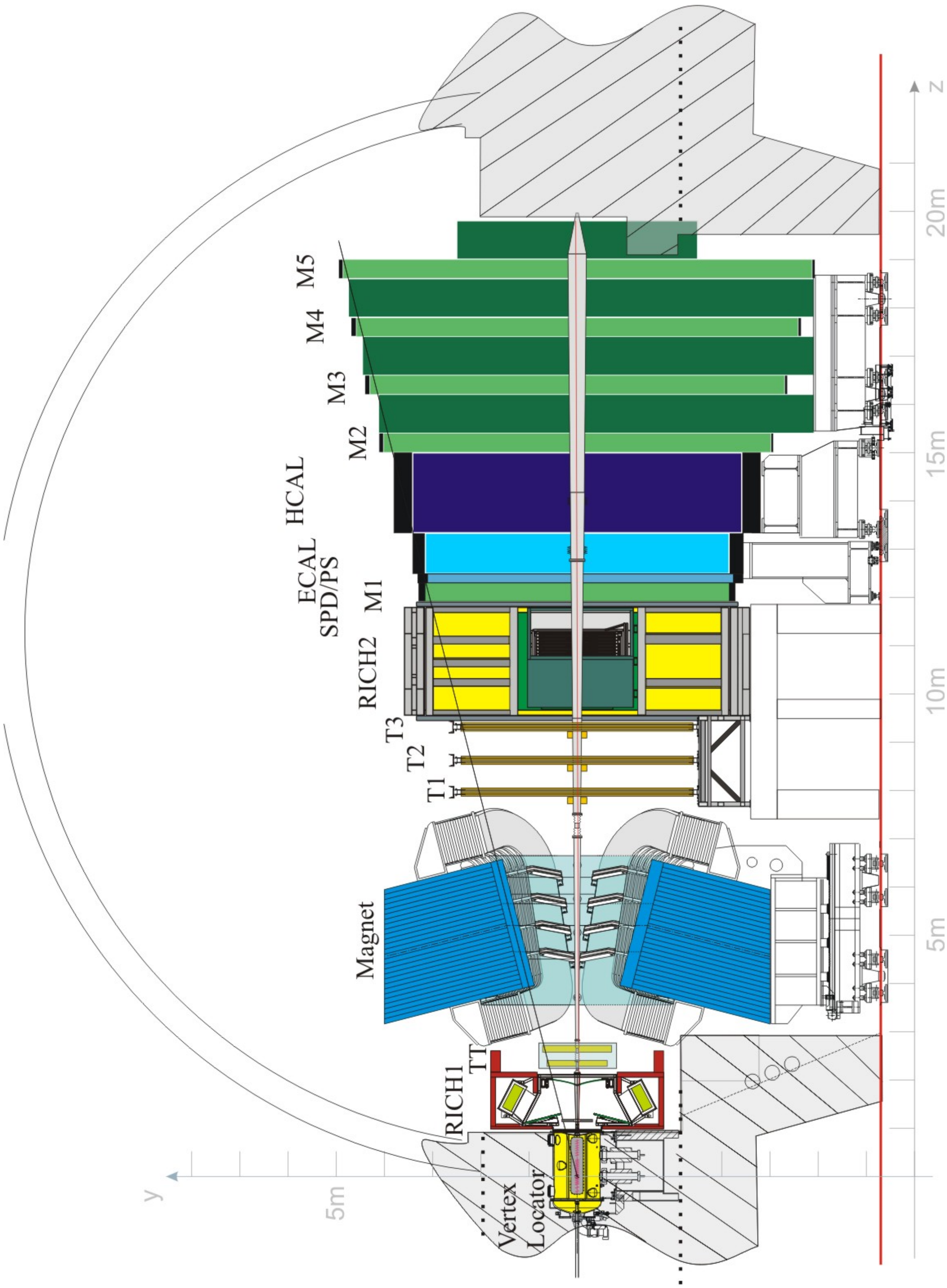


Figure 4.5: The LHCb detector [101]. The various sub-detectors are colour-coded and are described in the text.

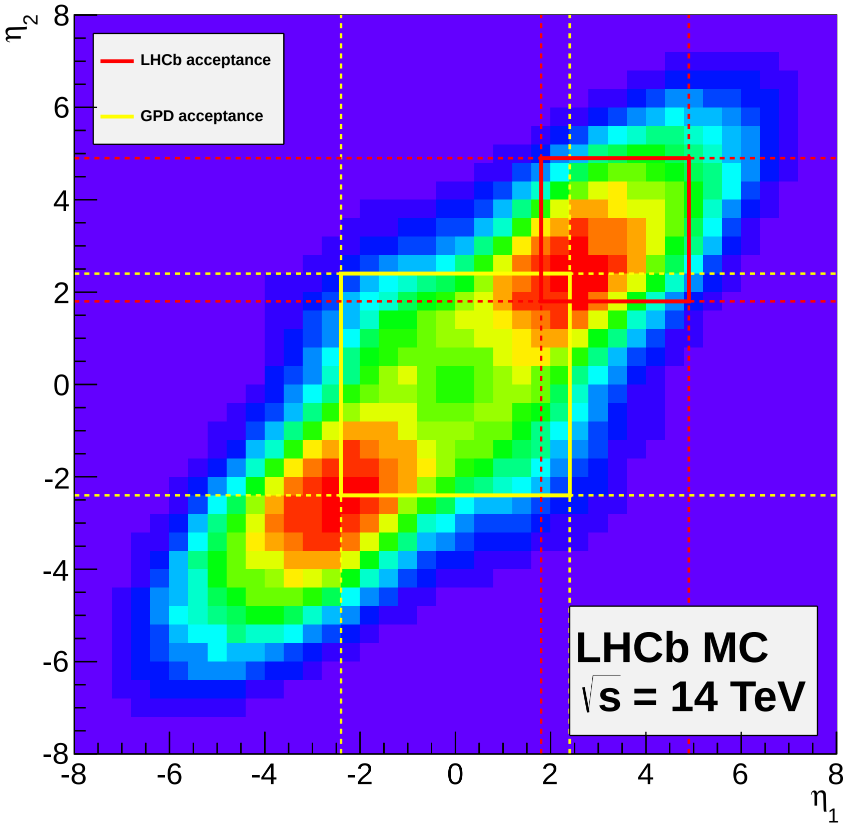


Figure 4.6: The pseudorapidities of the b quarks in $b\bar{b}$ production in 14 TeV collisions at the LHC simulated using Pythia (referred to as "MC"). The number of quarks is designated by the colour scheme with red indicating a higher prevalence. LHCb acceptance is given by the red square, while ATLAS and CMS (collectively referred to as "GPD") acceptance is given by the yellow square [114].

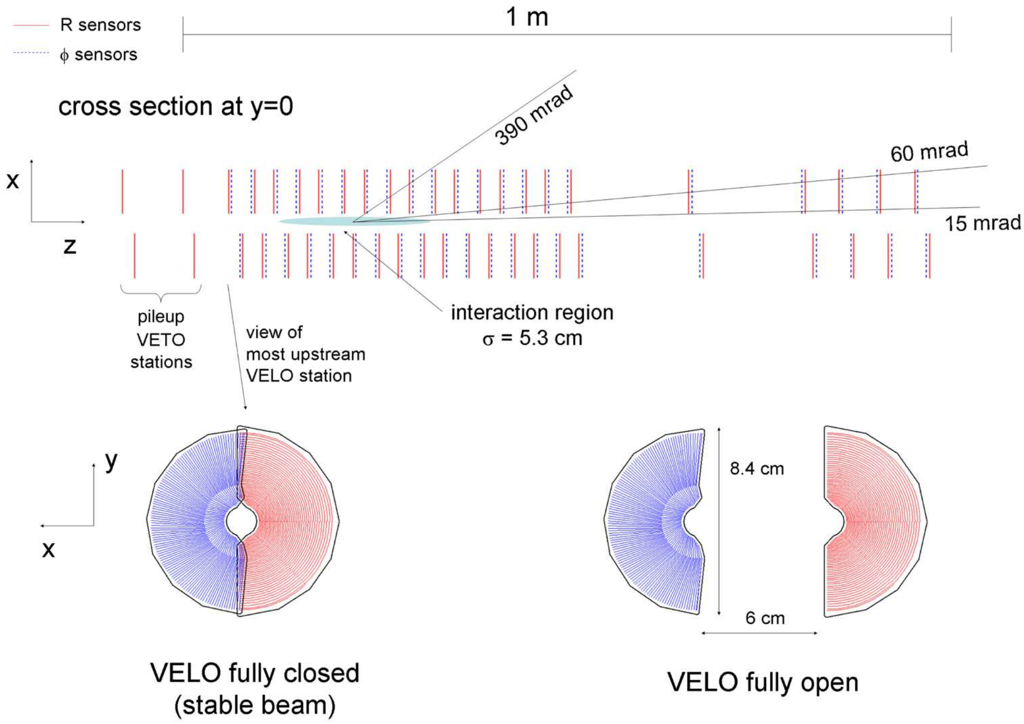


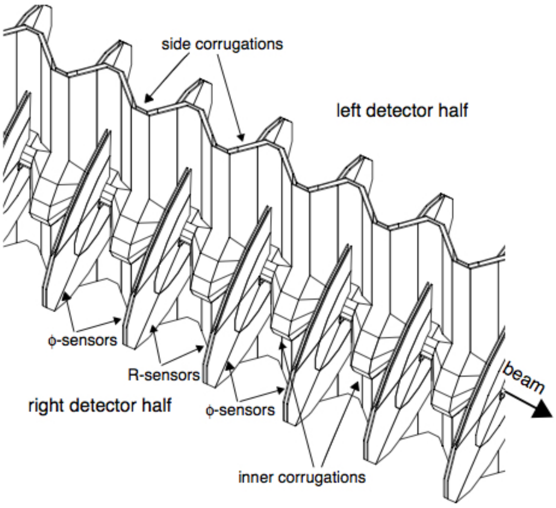
Figure 4.7: Top: The arrangement of sensor plains is illustrated along the x- and z-axes [101]. Bottom: A cross-sectional view of the VELO in the closed and open position in the (x,y) plane [101].

4.3.1 The LHCb tracking system

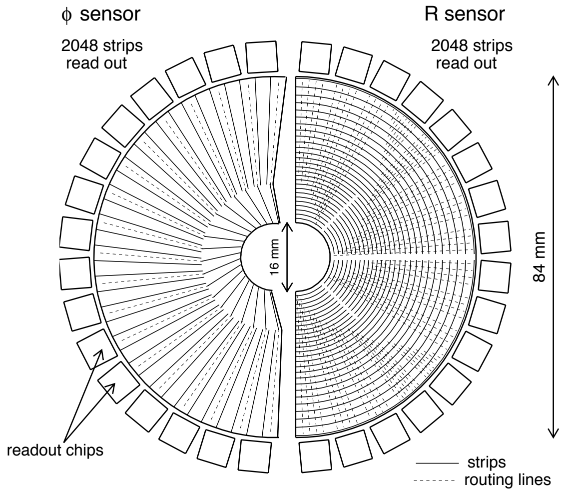
The LHCb tracking system consists of the VELO, a warm dipole magnet, the TT, and tracking stations T1 to T3. The collective aims of these subdetectors are precise vertexing, momentum measurement, and particle identification by matching tracks to hits in the RICH, calorimeter system, and muon stations.

VELO

The primary objectives of the VELO are to provide precise vertexing, an accurate measurement of decay lifetimes, and to measure impact parameters. It does so by the use of 21 silicon R- and ϕ-modules in a cylindrical geometry over a distance of ~1 m along the z-axis. When the beam is defocused, both halves are retracted by 3 cm to prevent radiation damage. However, during running conditions the two halves of the VELO are located a mere 8 mm from beam. Fig. 4.7 provides an overview of the VELO along the x- and z-axes and also illustrates the VELO in the open and closed position. There are



(a)



(b)

Figure 4.8: (a) The RF foil, the R-sensors and ϕ-sensors. The corrugations in the foil allow for an overlap between the left and right detector halves. A cross-sectional view of ϕ-sensors and R-sensors is shown in (b). The R-sensors consist of azimuthal strips, while the ϕ-sensors consist of radial strips. Both schematics are taken from Ref. [115].

two pileup veto stations upstream of the interaction region that consist of R-sensors only and are used to estimate the number of collisions per bunch crossing.

The R-sensors are concentric semi-circles that have centres at the beam position with the aim of measuring the radial distance from the beam axis. The ϕ-sensors are a collection of strips running radially outward and measure the azimuthal coordinate. Finally, the displacement of the sensors along the z-axis enables the third dimension to be determined. A resolution of 4 μm is achieved and this enables B and D hadron production and decay vertices to be reconstructed. This information is used in the high level trigger to select events of interest.

The VELO operates in a secondary vacuum to the LHC system and needs to be shielded against the RF electromagnetic fields the beams are subjected to. This is achieved by placing the VELO modules inside an aluminium box known as the RF foil that is illustrated in Fig. 4.8 (a). The foil is corrugated to enable the left and right modules to overlap, which reduces undesired edge effects in track reconstruction. Fig. 4.8 (b) illustrates the R-sensors and ϕ-sensors. Each strip of the R-sensors is subdivided into four 45-degree regions to minimise occupancy and strip capacitance. The strips in the ϕ-sensors have a skew with respect to the radial of between 10 and 20 degrees to improve

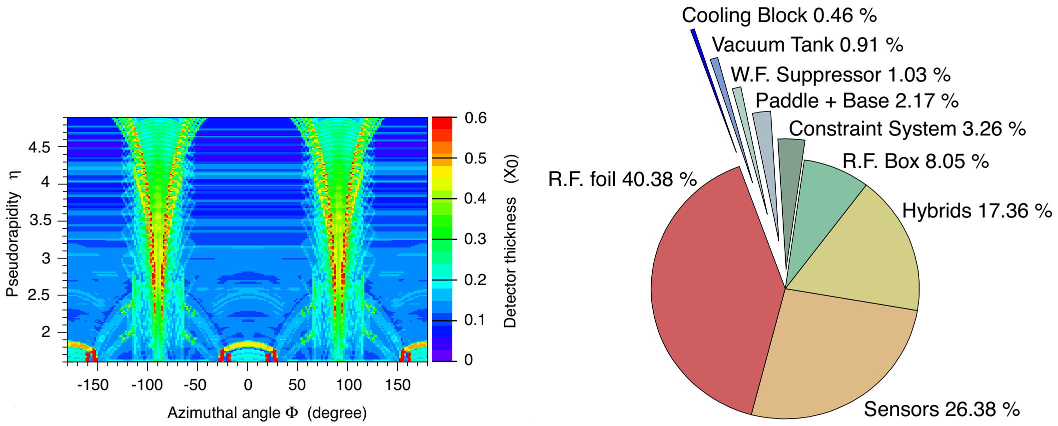


Figure 4.9: Left: Detector thickness in units of radiation length as a function of azimuthal angle and pseudorapidity [101]. Right: The average fractional contributions to the radiation length [101].

pattern recognition. Both sensors have a thickness of 300 μm and were developed for high radiation tolerance. Despite a measurable degree of radiation damage over the first two years of data taking, the tracking efficiency has not degraded significantly [116].

The material budget was investigated in simulation by generating straight tracks and extrapolating them through the VELO [101]. The VELO as a whole corresponds to 17.5% of a radiation length. A major fraction of the VELO material budget is represented by the RF foil as is made clear by Fig. 4.9. The contribution from other components is also shown. A precise understanding of the material seen by a particle is particularly important for the measurement of $W \to e\nu$ production due to bremsstrahlung. As the electron traverses the detector material, its momentum is continually decreased by the emission of photons. This also impacts the reconstructed transverse momentum distribution that is used to determine event yields. Studies detailed in Sect. 5.9 demonstrate that the knowledge of the momentum scale is the leading uncertainty in this measurement.

The Tracker Turicensis

Located between the first RICH and the dipole magnet, the TT is a planar silicon microstrip detector that is 150 cm wide and 130 cm high with an active area of $8.4\,\mathrm{m}^2$. It consists of four layers of silicon strip sensors with alternating layers rotated by an angle with respect to the y-axis to enable better spatial resolution. This subdetector is extensively used in the trigger because it enables fast, though relatively imprecise, track reconstruction. It is also useful for low momentum tracks that can bend out of acceptance

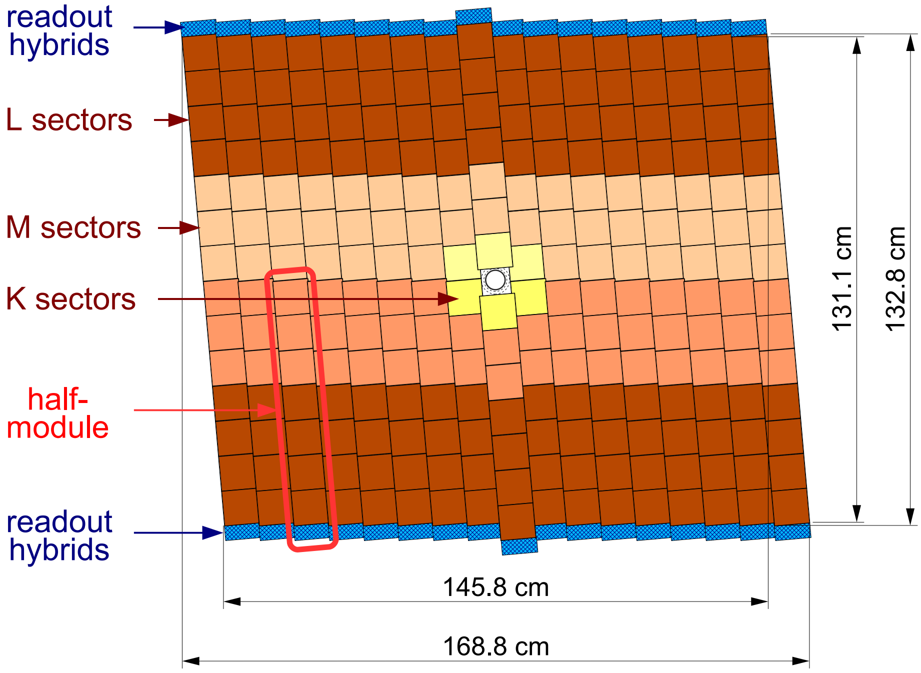


Figure 4.10: The layout of the TT with different shadings for the various readout sectors [101].

without leaving hits in the T-stations. For the measurement of $W \to e\nu$ production, requiring tracks to have hits in the TT serves as a powerful discriminant against ghost tracks where track segments in the VELO and T-stations have been wrongly assigned to the same track.

The layout of one of the four detection layers is illustrated in Fig. 4.10. A detection layer consists of vertical half-modules with seven silicon sensors organised into two or three readout sectors. For half-modules close to the beampipe, where the expected particle density is highest, the sensors are organised into three readout sectors. Two readout sectors are used for the other half-modules. Each sensor is 500 μm thick, 9.64 cm wide and 9.44 cm long.

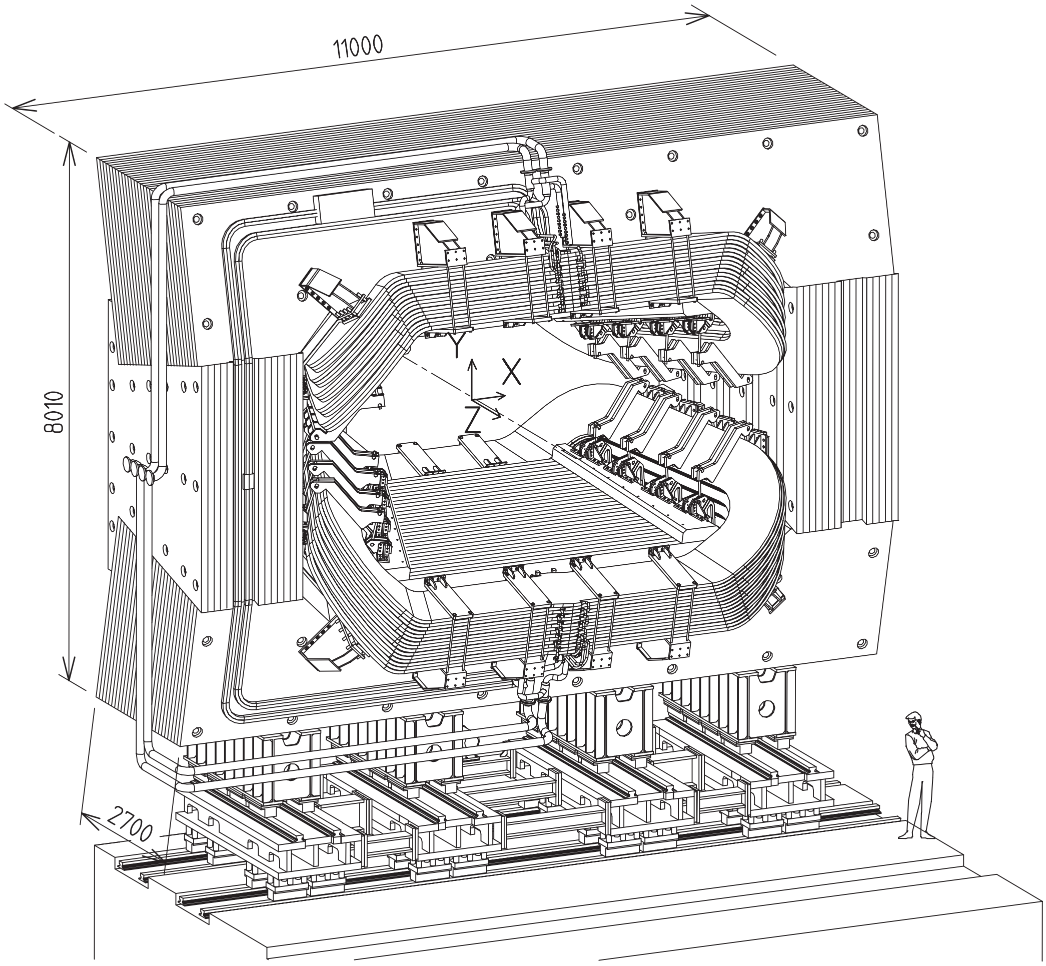


Figure 4.11: A perspective view of the LHCb dipole magnet [101]. The interaction point lies behind the magnet.

The magnet

The purpose of the LHCb dipole magnet is to measure the momentum of charged particles. It consists of two aluminium coils inside an iron yoke; a perspective view of the magnet is shown in Fig. 4.11. With a bending power of $\int B\mathrm{d}l \sim 4\,\mathrm{Tm}$, it creates a magnetic field in the positive, or negative, y-direction bending charged particles in the plane perpendicular to the field by the Lorentz force. As the bending angle determined from measuring track segments before and after the magnet is proportional to $1/p$, the momentum of the particle can be measured.

In order to achieve the required momentum resolution, the magnetic field must be measured with a relative precision of a few parts per 10^{-4}. The measurement is carried

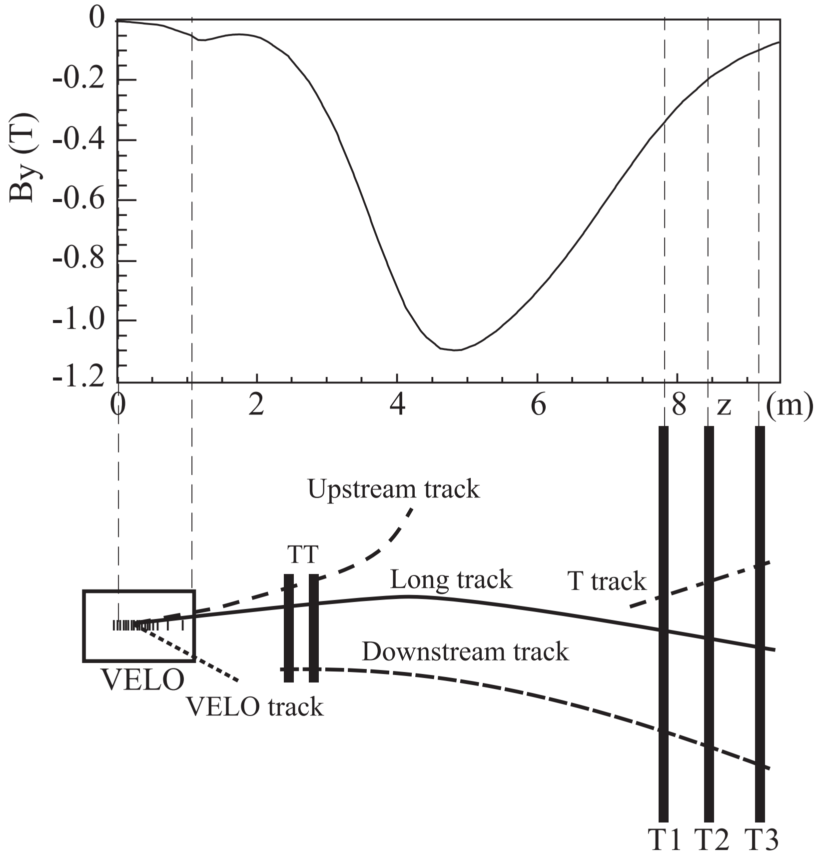


Figure 4.12: Top: The strength of the magnetic field along the z-axis. Bottom: An illustration of various track types according to the tracking systems they traverse. The figure is taken from Ref. [101].

out on a fine grid of $8 \times 8 \times 10\,\text{cm}^3$ covering most of the LHCb acceptance region. The main component of the magnetic field, B_y, is shown in Fig. 4.12.

The tracking stations

While the VELO and the TT consist of silicon strips, the tracking stations ("T-stations") are combinations of a silicon strip inner tracker (IT) and a gas straw tube outer tracker (OT). The inner trackers follow a design similar to the TT with alternating layers rotated slightly with respect to their adjacent layers. The outer tracker consists of drift tubes which are filled with an inert gas and contain a wire held under voltage. When a charged particle traverses the gas, atoms are ionised and electrons drift toward the wire leading to a current in the wire.

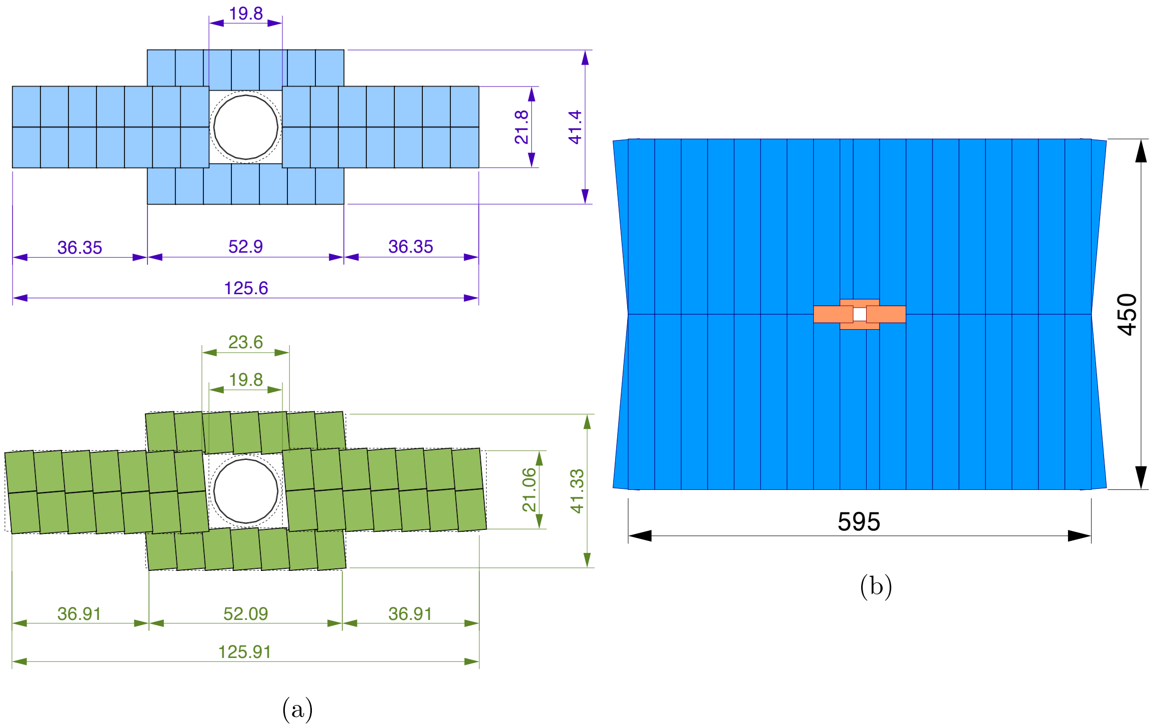


Figure 4.13: The layout of the IT detection layers is illustrated in (a) [117]. (b) shows how the IT and OT work together [117]. Dimensions are given in cm.

The IT covers a 120 cm wide and 40 cm high cross-shaped region around the beamline with a layout illustrated in Fig. 4.13 (a). It consists of multiple detection layers rotated by a stereo angle of 5 degrees. It has an active area of 4.0 m^2 with 129024 strips of either 11 cm or 22 cm in length. Both the IT and TT are required to have a single-hit resolution of 50 μm with a signal-to-noise ratio in excess of 12:1. The OT, shown in Fig. 4.13 (b), is a drift-time detector designed for the tracking of charged particles over a large acceptance area. It consists of an array of drift tubes with an inner diameter of 4.9 mm. A mixture of Argon (70%) and CO_2 (30%) is used to ensure a fast drift time of less than 50 ns and a spatial resolution of 200 μm is achieved.

Track reconstruction

The process of building tracks at LHCb begins by examining hits in the tracking systems in the (x,z) and (y,z) projections to form track stubs (also referred to as "seeds") within each subsystem. Subsequently, track stubs are joined to create tracks across all the tracking systems as illustrated in Fig. 4.14. Tracks that span the VELO and the T-stations are referred to as "long tracks". These are also shown in Fig. 4.12. Finding

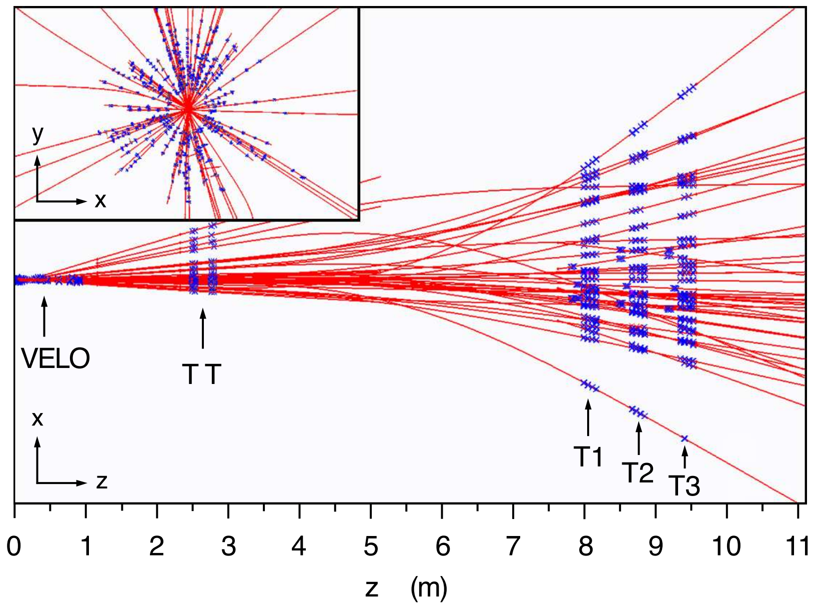


Figure 4.14: Hits in the various tracking subsystems along with reconstructed tracks. The insert shows the (x, y) plan in the VELO region [101].

appropriate combinations of track segments is accomplished by using two methods. One option is to propagate a track segment in the VELO to the T-stations and attempt to find a corresponding track segment in that system. Alternatively, one can proceed in the opposite direction by beginning in the T-stations and matching in the VELO. Thereafter, tracks that are probably due to the same particle are checked for. Specifically, an algorithm based on the Kullback-Liebler distance [118, 119] is employed to identify clones. These are removed prior to fitting the tracks to reduce the computational load of track reconstruction.

The purpose of the track fit is to determine track parameters along the trajectory of the particle. Specifically, the track fit returns the momentum of the particle based on a deflection in the magnetic field, the position and direction of the track in various subdetectors, and a covariance matrix along with a χ^2 for the fit quality. As LHCb is a forward spectrometer with subdetectors at planes of constant z, it is natural to define the track parameters

$$\vec{s} = (x, y, \mathrm{d}x/\mathrm{d}z, \mathrm{d}y/\mathrm{d}z, q/p), \tag{4.1}$$

where x and y are transverse coordinates, q is the charge of the particle and p is the momentum.

At LHCb track fitting is based on the Kalman filter [120,121] formalism whereby a state vector ($\vec{s}$ in Eq. 4.1) is defined and subsequently predicted and updated in an alternating fashion. While the original formulation of the Kalman filter considered evolution in time with measurements affected by random noise, here the z-coordinate takes the analogous role to time and "noise" results from multiple scattering and bremsstrahlung. Knowledge of track parameters is updated iteratively across the subsystems such that the best estimate of track parameters is obtained at the last measurement plane.

The equation of motion for a charged particle in a magnetic field $\vec{B}$ is

$$\frac{\mathrm{d}p}{\mathrm{d}t} = q\vec{v} \times \vec{B}, \tag{4.2}$$

where t is time and $\vec{v}$ is the particle's velocity. In reality, however, the magnetic field will be inhomogeneous and the particle will be subject to multiple scattering and energy loss. Consequently, numerical methods must be used in predicting the location of hits along a track. The two main contributions to the uncertainty in determining the momentum are due to the position measurement with $\delta p/p \propto p$ and due to multiple scattering, which displays the inverse dependence. As a result, the overall momentum resolution is surprisingly insensitive to momentum with an uncertainty of 0.5% for tracks with a low momentum and 1.0% for 200 GeV tracks at LHCb.

4.3.2 Particle identification

At LHCb particle identification (PID) is of fundamental importance as many of the goals of the experiment necessitate pions, kaons and protons to be distinguished in B hadron decays. The soft momentum spectrum at larger polar angles is covered by RICH1 while the harder spectrum is covered by the RICH2 detector downstream of the magnet. For the measurement of $W \to e\nu$ production, the ECAL and HCAL (and the SPD and PRS to a lesser extent) provide PID, while the five muon stations, M1-M5, are used in measurements of EW production in the muon final state such as Ref. [2] shown in Sect. 3.3.

The RICH detectors

At LHCb ring imaging Cherenkov detectors are chiefly employed to distinguish pions from kaons in selected B hadron decays for precise $\mathcal{CP}$ measurements. Fig. 4.15 illustrates how

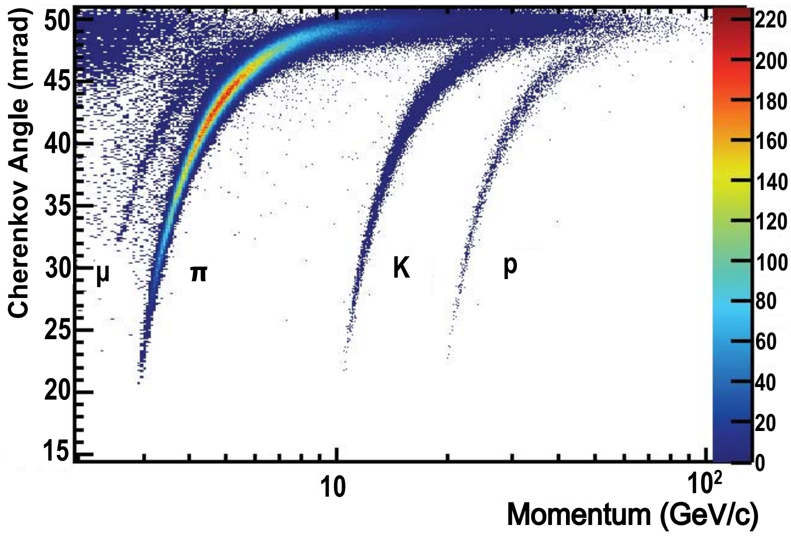


Figure 4.15: The Cherenkov angle as a function of track momentum in the C_4F_{10} radiator. Cherenkov bands for muons, pions, kaons and protons are clearly visible [101].

the different particles form clearly distinct bands in the RICH. Physically, the detectors operate by the Cherenkov effect whereby charged particles, whose speed surpasses the phase velocity of light in that medium, create a disturbance in the field causing photons to be emitted. The cosine of the angle at which Cherenkov photons are emitted with respect to the direction of motion of the charged particle is inversely proportional to the particle's speed with

$$\cos\theta_c = \frac{1}{n\beta}, \tag{4.3}$$

where θ_c is the Cherenkov angle, n is the refractive index of the material, and β is the fractional speed of the charged particle. Together with momentum information from the tracking system, a mass hypothesis can be assigned.

From an instrumental point of view, this is achieved with two detectors: RICH1 that is located upstream of the magnet and is used for low momentum charged particles with ~1-60 GeV and RICH2 that is located downstream and employed for particles with ~15 GeV to beyond 100 GeV of momentum. The detectors are illustrated in Fig. 4.16. The medium in RICH1 is aerogel and C_4F_{10}, whilst CF_4 is used in RICH2. While RICH1 extends to ±300(250) mrad horizontally (vertically), RICH2 is limited to ±120(100) mrad horizontally (vertically). This is acceptable as the majority of high momentum particles are produced closer to the beam axis. For both subdetectors a combination of spherical

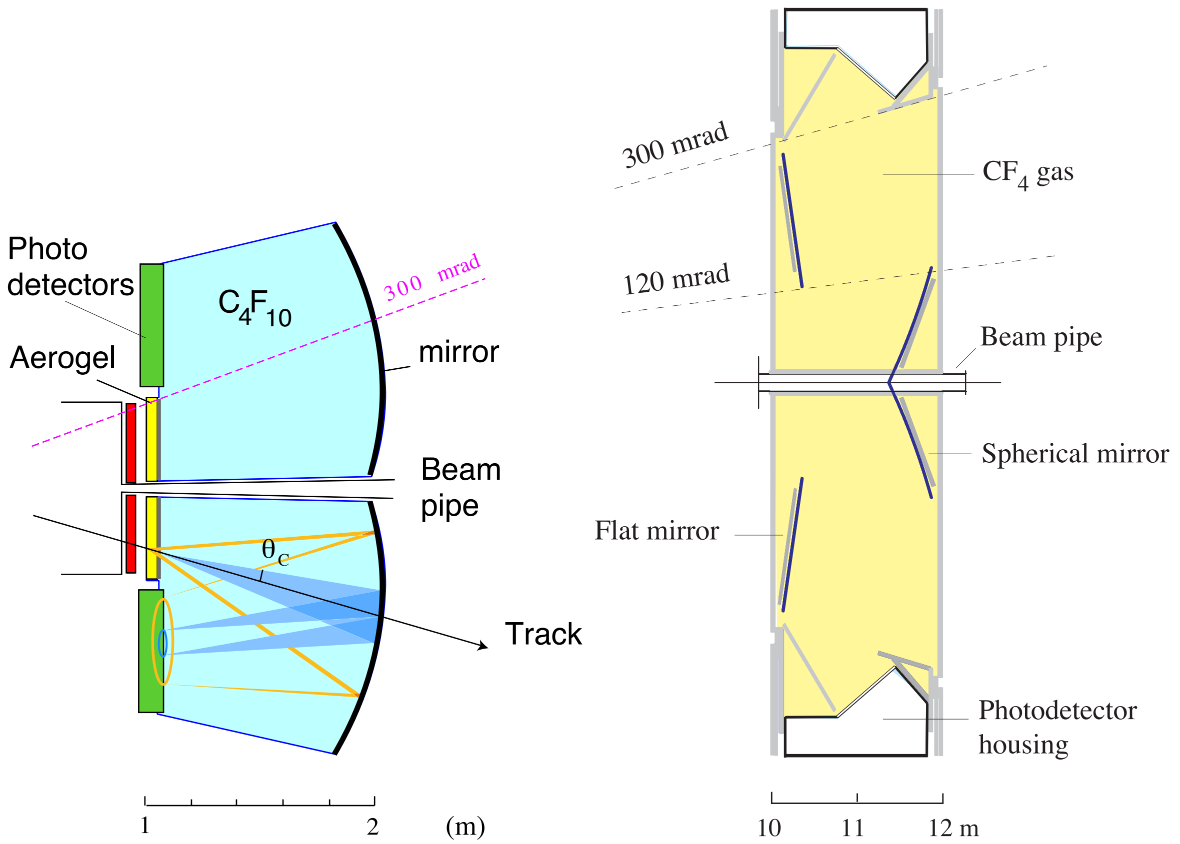
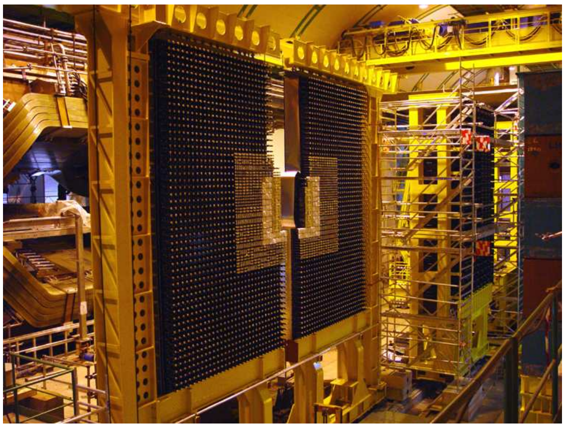


Figure 4.16: Schematic layout of the RICH detectors with RICH1 on the left and RICH2 on the right [122]. The flat mirrors deflect the light onto photodetectors that are placed outside of LHCb acceptance.

and flat mirrors are used to reflect the light onto hybrid photon detectors. These detect photon hits and convert it to current by the photoelectric effect.

The calorimeter system

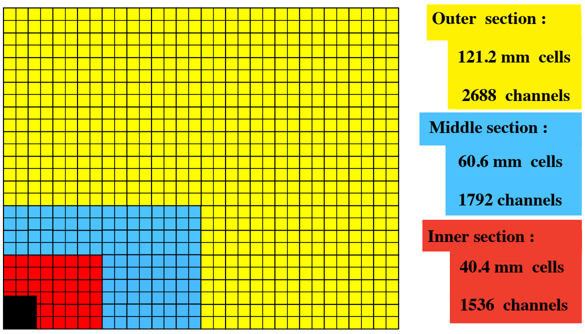
The calorimeter system is used for energy measurement, hardware-level triggering, reconstruction and PID. The first stage of the measurement is performed by the SPD/PRS system which consists of a tile of lead sandwiched between two scintillating pads. Before traversing the lead absorber, charged, rather than neutral, particles will mostly start to shower and be detected in the SPD enabling photons to be differentiated from electrons. There is, however, a background from photons due to pair production in the material and from backward moving particles that originate from the lead absorber. The lead



(a)

(b)

Figure 4.17: Photos of the (a) ECAL and (b) HCAL [101].



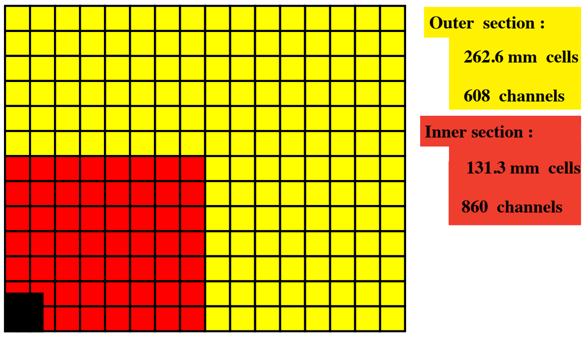


Figure 4.18: The lateral segmentation of the ECAL (left) and the HCAL (right) [101]. One quarter of the detector plane is shown.

absorber will cause an electromagnetic shower but is insufficient for most hadrons to interact, thus differentiating charged hadrons from electrons.

The ECAL and HCAL consist of alternating layers of scintillator and absorber (lead and iron, respectively) and are shown in Fig. 4.17. They have a rectangular design and result in an acceptance that varies as a function of the azimuthal angle. The calorimeter system adopts a variable lateral segmentation that is illustrated in Fig. 4.18 as the hit density varies by several orders of magnitude along the polar angle. The thickness of the ECAL is 25 radiation lengths to contain most electromagnetic showers while the HCAL is limited to 5.6 interaction lengths due to space limitations.

The behaviour of particles traversing the material is dependent on mass and particle type. Electrons are light enough to undergo bremsstrahlung and cause an electromagnetic shower to develop which eventually dissipates the energy of the electrons. Muons pass

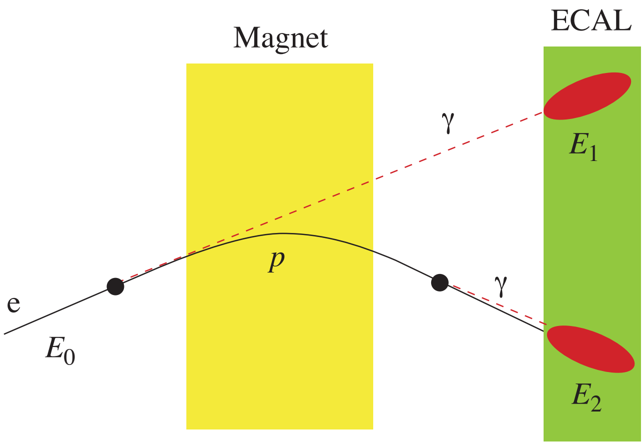


Figure 4.19: A schematic of bremsstrahlung recovery [101].

through the system without significant showering in most cases and are instead observed in the dedicated muon stations. Hadrons are also significantly more massive than electrons but interact principally in the HCAL and produce secondaries which enable an energy measurement to be made. In all cases, the energy is recorded by the scintillators and the resultant light is transported via wavelength-shifting fibres to photomultiplier tubes.

The energy resolution of the ECAL is $1\% + 10\%\sqrt{E[\mathrm{GeV}]}$. However, the energy measurement of electrons from the decay of W bosons requires additional considerations. Firstly, electrons pass through $\sim\frac{1}{4}X_0$ of material[2] and so their momenta are decreased by bremsstrahlung ahead of the magnet, before the momentum measurement is made using information from T-stations. For low energy particles, bremsstrahlung photons can be recovered in the calorimeter system as illustrated in Fig. 4.19. Leptons originating from W and Z boson decay, however, are less likely to deflect by a significant amount. Consequently, the bulk of the energy carried by the particle is commonly deposited in a single calorimeter cell. As LHCb calorimeters are designed for the detection of particles originating from heavy flavour decays, the cells measure transverse energy[3] accurately only up to about 10 GeV. As a result, a typical electron from the decay of a EW boson will saturate the cell and lead to an unknown, variable reduction in energy.

[2] X_0 is the radiation length for a material. $\frac{1}{4}X_0$ corresponds to a mean energy loss of 22%.

[3] Transverse energy, E_T, is defined as

$$E_\mathrm{T} = \sum_{i=1}^{4} E_i \sin\theta_i,$$

where E_i is the energy deposited in cell i and θ_i is the angle between the z-axis and a neutral particle assumed to be coming from the mean position of the interaction envelope hitting the centre of the cell.

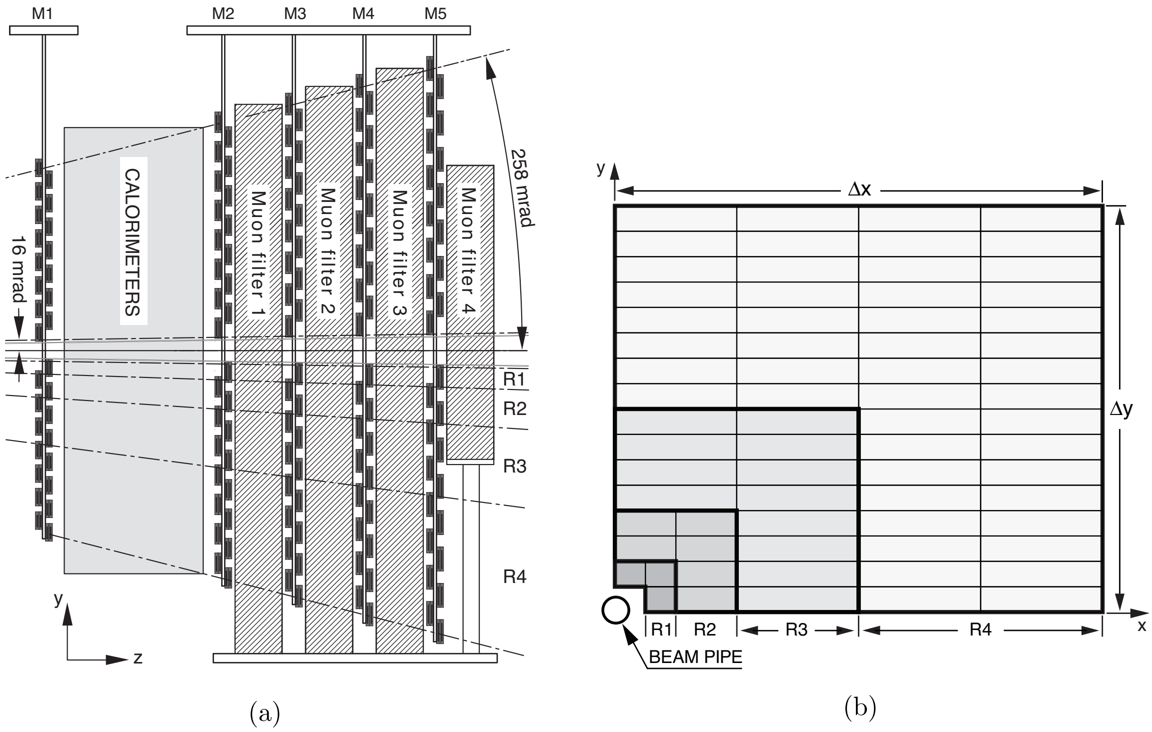


Figure 4.20: (a) Side view of the muon system. (b) Front view of a quadrant of a muon station [101].

For this reason, the application of the calorimeter system to the measurement of $W \to e\nu$ production is limited to PID and the energy is largely extracted from the momentum measurement.

Muon system

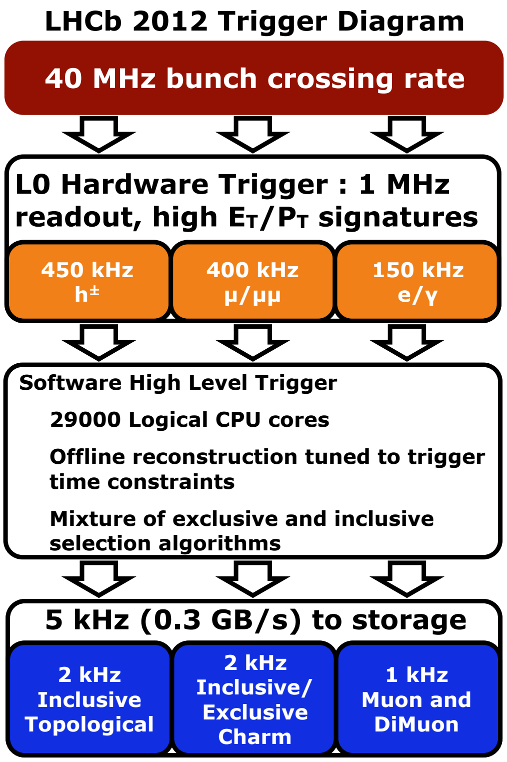
The muon system is composed of five subsystems (M1-M5) that are rectangular in shape and placed along the beam axis as illustrated in Fig. 4.20 (a). The system covers a total area of $435\,\mathrm{m}^2$ and has an inner angular acceptance of 20(16) mrad and an outer one of 306(258) mrad in the bending (non-bending) plane. Stations M2 to M5 are interleaved with iron absorbers with a total interaction length of 20 including the calorimeter system. M1 is upstream of the calorimetry to improve the transverse momentum measurement. M1-M3 have a high spatial resolution to calculate track directions and transverse momenta while stations M4 and M5 have been designed primarily for identification and have limited spatial resolution. To maintain a roughly constant particle flux and channel occupancy, each station is subdivided into four regions that scale in the ratio 1 : 2 : 4 : 8 as illustrated in Fig. 4.20 (b).

While multi-wire proportional chambers (MWPC) are used for muon detection in most stations, a triple gas electron multiplier is employed in the inner region of M1 due to high particle rates of up to 500 kHz/cm^2. What is more, the chambers were also designed to be radiation hard to withstand the large particle flux without exhibiting significant ageing effects in 10 years of operation. A total of 1 368 MWPCs are used, each with a gas mixture of $Ar/CO_2/CF_4$ in the ratio 40 : 55 : 5. Each chamber has roughly $3 \cdot 10^6$ wires with a spacing of 2.0 mm.

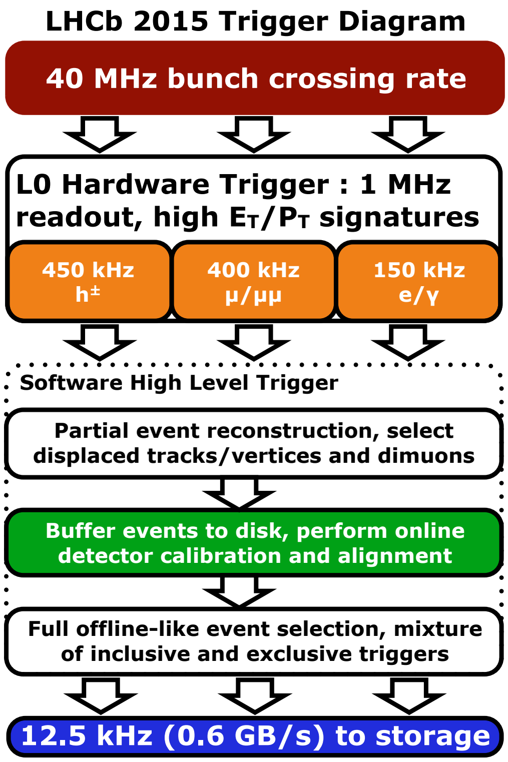
4.3.3 The LHCb trigger system

All LHC experiments employ a trigger to filter events for two reasons: offline storage of all reconstructed events is unfeasible and the cross-section for events of interest is small with respect to the total inelastic *pp* cross-section. Triggering is performed in stages where each subsequent stage is designed to deal with a reduced rate of information but with an increased degree of complexity and processing time.

At LHCb the online event selection is performed by a trigger [124, 125], which consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, which applies a full event reconstruction as illustrated in Fig. 4.21. A set of global event cuts (GEC) is applied, which prevents events with high occupancy from dominating the processing time of the software trigger. The relevant requirement for EW analyses is based on the number of SPD hits and this is discussed at length in Sect. 5.7.4.



(a)



(b)

Figure 4.21: The trigger scheme for the 2012 data taking period is illustrated in (a) while (b) shows modifications made to the scheme in 2015 such as online calibration and alignment, and offline event selection [123]. "Inclusive Topological" refers to triggers that select partially reconstructed B hadron decays based on the properties of combinations of 2, 3, or 4 tracks.

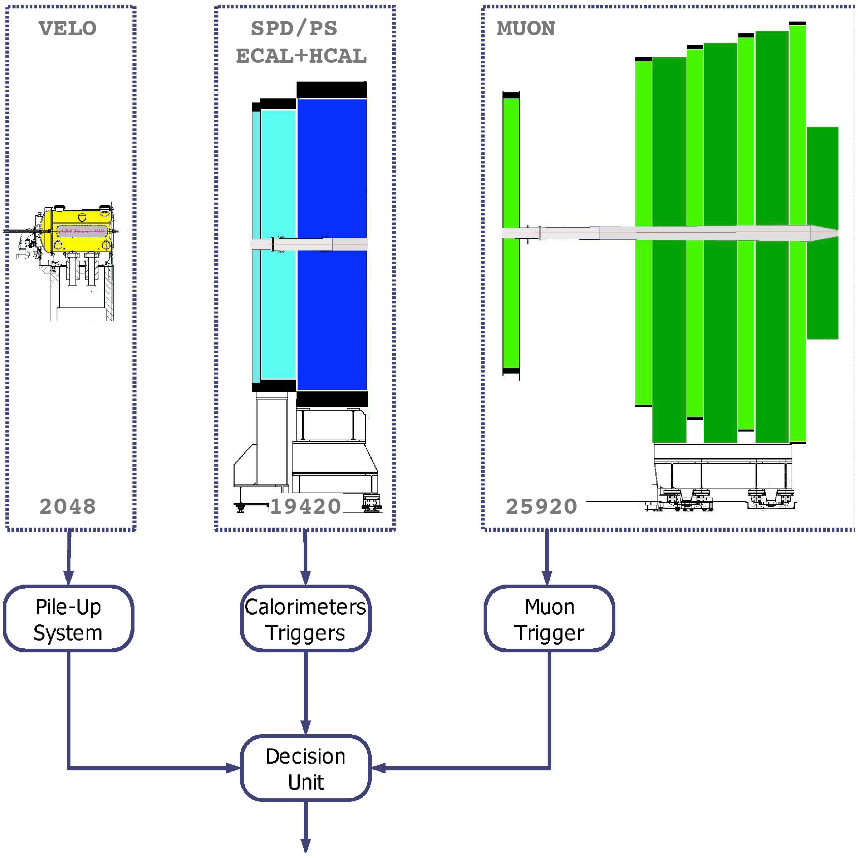


Figure 4.22: Overview of the Level-0 trigger [101]. Also listed is the number of channels per subdetector.

The hardware trigger

The LHCb hardware ("Level-0") triggering system consists of custom electronics operated synchronously with bunch crossings and is subdivided into three components: the pile-up system, the calorimeter trigger and the muon trigger as shown schematically in Fig. 4.22. The trigger identifies the largest transverse energy electron, photon, charged hadron, and neutral pion in the event, if any, and the two highest p_{T} muons in the muon chambers reducing the rate from roughly 11 MHz to around 1 MHz in the process. Further requirements are then applied as shown in Table 4.1.

The purpose of the pile-up system is to distinguish between crossings with single and multiple visible interactions using four silicon sensors upstream of the VELO measuring radial positions of tracks. It consists of two planes perpendicular to the beam-line with two overlapping R-sensors in each. The system is used to determine the number of vertices at the interaction point. If multiple interactions are found, the crossing is vetoed.

<table>
<tr><th>Trigger</th><th>Requirements</th><th>SPD multiplicity</th><th>Rate [kHz]</th></tr>
<tr><td>L0Muon</td><td>$p_{\mathrm{T}}^{\mu} > 1.76\,\mathrm{GeV}$</td><td>$< 600$</td><td rowspan="2">400</td></tr>
<tr><td>L0DiMuon</td><td>$\sqrt{p_{\mathrm{T}}^{\mu_1,\,\text{largest}} \times p_{\mathrm{T}}^{\mu_2,\,\text{2nd largest}}} > 1.6\,\mathrm{GeV}$</td><td>$< 900$</td></tr>
<tr><td>L0Hadron</td><td>$E_{\mathrm{T}}^{h} > 3.7\,\mathrm{GeV}$</td><td>$< 600$</td><td>490</td></tr>
<tr><td>L0Electron</td><td>$E_{\mathrm{T}}^{e} > 3\,\mathrm{GeV}$</td><td>$< 600$</td><td rowspan="2">150</td></tr>
<tr><td>L0Photon</td><td>$E_{\mathrm{T}}^{\gamma} > 3\,\mathrm{GeV}$</td><td>$< 600$</td></tr>
</table>

Table 4.1: The main Level-0 trigger lines for physics analysis with their selection requirements and typical rates in 2012. The rates for `L0Muon` and `L0DiMuon` are specified together as are those for `L0Electron` and `L0Photon`.

The calorimeter trigger collects information from the SPD, PRS, ECAL and HCAL and forms clusters by adding the E_{T} of 2×2 cells and selecting the largest E_{T} clusters. 2×2 cells are used as they are large enough to contain most of the energy while small enough to avoid overlaps between particles. Information from the subdetectors is combined to identify the type of electromagnetic candidate: electron, photon or π^0. Only the highest E_{T} cluster per candidate type is selected.

The muon trigger looks for two muon tracks with the largest transverse momenta by forming a tower of hits through the five muon stations and pointing towards the interaction point with a total of 192 possible towers. The seeds for track finding are hits in M3. A straight line is passed through stations M2, M4 and M5 that defines the extrapolated positions where hits are searched for. A field of interest is created around these extrapolated positions and the size of these is defined by the station being considered and other factors.

The components are connected to the Level-0 decision unit which collects all the information and makes a decision. Information from all subdetectors is used in the process and a decision is made for each bunch crossing within 4μs. As this time includes the time-of-flight of the particles, cable delays and all other delays in the electronics, then 2μs is left for processing data.

Subsequently, the entire detector is read out and the second level, the software trigger, is employed to reduce the rate to 5 kHz enabling events to be written to storage for offline analysis.

The software trigger

The software *i.e.* high-level trigger (HLT) is a C++ application running on a computing farm consisting of roughly 29 000 commercially available cores operated asynchronously with respect to beam crossings. The HLT is subdivided into HLT1 and HLT2 with the former reconstructing particles in the VELO and T-stations based on Level-0 objects and the latter building on existing information to reconstruct composite particles and their decays.

A "trigger line" is a collection of reconstruction algorithms and selection requirements that returns an accept or reject decision. The LHCb physics goals along with processing and bandwidth constraints dictate the trigger programme. During 2011 running, the HLT contained 38 HLT1 and 131 HLT2 lines.

All Level-0 objects *i.e.* calorimeter clusters and muon tracks are passed to HLT1. The main purpose of HLT1 is to confirm Level-0 decisions using more sophisticated event information reducing the input rate from 1 MHz to roughly 80 kHz in the process.

The rate of events passed onto HLT2 is sufficiently low to allow for offline track reconstruction. Tracks in the HLT are different from offline tracks in that they have been fitted with a simplified geometry and with fewer iterations. Track segments in the VELO are matched to those in the T-stations in a procedure known as forward tracking. Minimum momentum and transverse momentum requirements of $p > 5\,\text{GeV}$ and $p_{\text{T}} > 0.5\,\text{GeV}$ are imposed to reduce the search windows in the T-stations.

While an HLT2 output rate of 3 kHz was maintained during 2011 running, an increased rate of 5 kHz was achieved in 2012. This was largely due to storing a fraction of events to hard disk and buffering these between runs whilst there is no beam in the LHC. Specifically, the LHC delivers stable beams for about 30% of the time. In an approach called "deferred triggering", roughly 20% of events accepted at L0 level are stored and analysed at a later time. A further improvement was achieved from the 2015 data taking period onwards whereby the deferral of events occurred after HLT1 decisions enabling online detector calibration and alignment to be performed followed by offline reconstruction and selection requirements at HLT2 level. Finally, an increase in computing resources allowed the output rate to be increased to 12.5 kHz as illustrated in Fig. 4.21.

Single electron trigger lines across the L0, HLT1 and HLT2 stages are used for the measurement of $W \to e\nu$. In addition, a veto on the HLT2 di-electron line is applied to reduce $Z \to ee$ candidates with both daughters in acceptance. The triggering strategy is discussed in more detail in Section 5.2.1 and the procedure to determine the trigger efficiency is explained in Section 5.7.5.

4.3.4 Luminosity determination at LHCb

The results presented in this thesis take the form of cross-section measurements, which rely on a determination of integrated luminosity. Since the precision of theory predictions in pQCD is of the order of 1-2%, then the luminosity must also be determined with similar precision [126]. The size of other systematic uncertainties that LHCb EW measurements are subject to also support a similar target for precision [1, 22].

At the LHC the instantaneous luminosity can be measured using the beam overlap integral expressed as

$$\mathcal{L} = N_1 N_2 f \sqrt{(\vec{v_1} - \vec{v_2})^2 - (\vec{v_1} \times \vec{v_2})^2/c^2} \int \rho_1(x,y,z,t)\rho_2(x,y,z,t)\mathrm{d}x\mathrm{d}y\mathrm{d}z\mathrm{d}t, \qquad (4.4)$$

where f is the revolution frequency, N_1 and N_2 are the numbers of protons in the two bunches, $\vec{v_1}$ and $\vec{v_2}$ are the velocities of the particles in the bunches, and the particle densities for the beams, normalised over all space, are given by $\rho_{1,2}(x,y,z,t)$. For highly relativistic beams colliding with a small half crossing-angle ϕ, the Møller factor $\sqrt{(\vec{v_1} - \vec{v_2})^2 - (\vec{v_1} \times \vec{v_2})^2/c^2}$ reduces to $2\cos^2\phi \simeq 2$.

Methods for determining luminosity can be classified as either direct or indirect. Indirect methods make a measurement of the elastic or the total cross-section of a process that is known from theory or by a previous measurement. Direct measurements, on the other hand, use Eq. 4.4 and employ several strategies to determine the various parameters of the equation. LHCb relies on two direct methods for luminosity measurement: the "van der Meer scan" (VDM) and the "beam-gas imaging" method (BGI) [127].

The BGI method reconstructs beam-gas interaction vertices to measure the properties of the beam such as angles, positions, and shapes. The method relies on the LHCb vertex detector to make high precision measurements of interaction vertices. The VDM method proceeds by shifting the beams in transverse coordinates with high precision and performs a scan of the colliding beams. The luminosity measurements are used to evaluate the visible cross-section of a reference process. The rate of the reference process is then used to determine the luminosity during main data-taking periods. While these

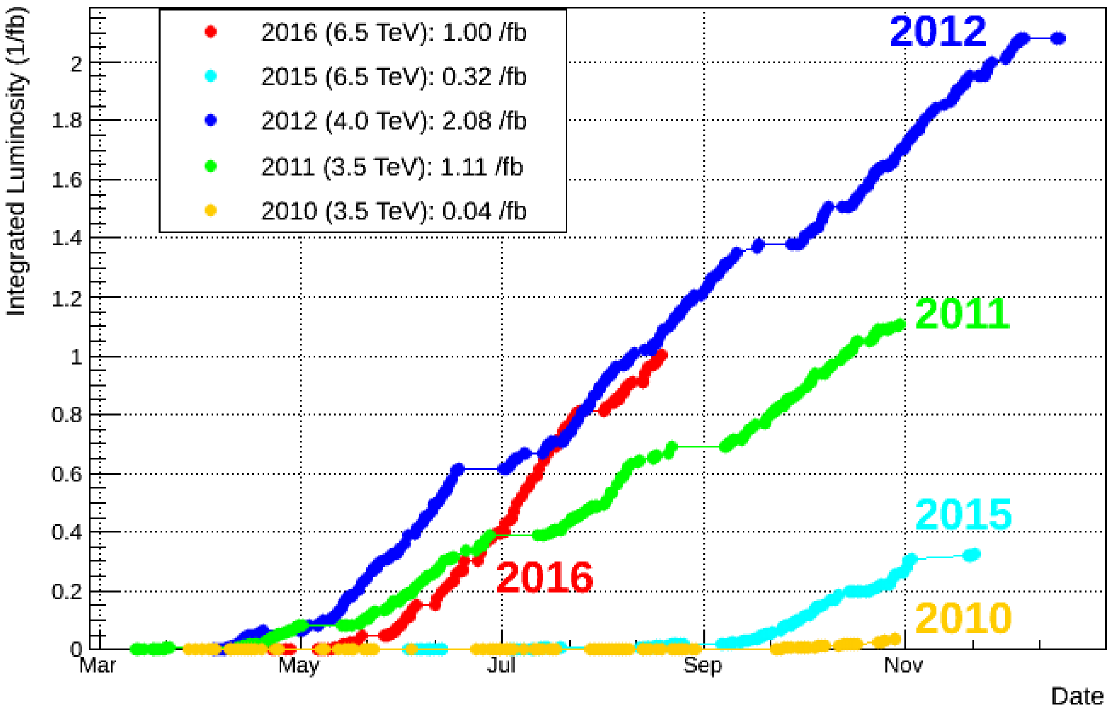


Figure 4.23: LHCb integrated luminosity in *pp* collisions from 2010 to 2016. Source: LHCb collaboration

methods are employed during short periods of data taking only, they are performed at regular intervals throughout the full period of running.

The measurements of luminosity at LHCb using the two methods are found to be consistent and an average is made for each dataset [127]. A precision of 1.16% is achieved for the 1990 pb^{-1} of data collected at 8 TeV and the 13 TeV measurement presented in this thesis is based on 295 fb^{-1} of data with a precision of 3.9% [2]. Currently, the precision of the luminosity calibration is limited by the systematic uncertainties of the beam overlap determination. However, these uncertainties are largely different for the VDM and BGI methods and so improved precision is achieved in the combination. The integrated luminosity LHCb has collected since the start of running in 2010 is shown in Fig. 4.23.

4.3.5 LHCb software and event simulation

LHCb software is implemented in an object-oriented framework called Gaudi [128]. The software chain works as follows: Gauss [129] handles the generation of events and simulates their behaviour in the detector; Boole [130] digitises the output of Gauss so that simulated data mimic real data; BRUNEL [131] performs event reconstruction; MOORE [132] runs the triggering software; and DAVINCI [133] is the analysis toolkit used to perform measurements. Computing is typically performed on the LHC worldwide computing grid [134] managed by a software framework called DIRAC [135] interfaced to LHCb software using GANGA [136].

The purpose of GAUSS is to create simulated data that mimic real data for a better understanding of the experimental conditions in the detector and the detector's performance. Simulated data are also used to optimise the event selection, estimate the background contamination and determine efficiencies in some cases. Simulation consists of a generator phase, where *pp* collisions are generated and decayed, and a simulation phase where particles are tracked in the detector and their interactions with the material are modelled. At LHCb *pp* collisions are generated using PYTHIA 8 [137] with a specific LHCb configuration [138]. The momentum distributions of the partons inside the proton are parameterised by the leading-order CTEQ6L1 [139] PDF set.

The interactions of the generated particles with the detector such as multiple scattering, bremsstrahlung, photon conversion, *etc.*, and its response, are implemented using the GEANT4 toolkit [140] as described in Ref. [141].

Chapter 5

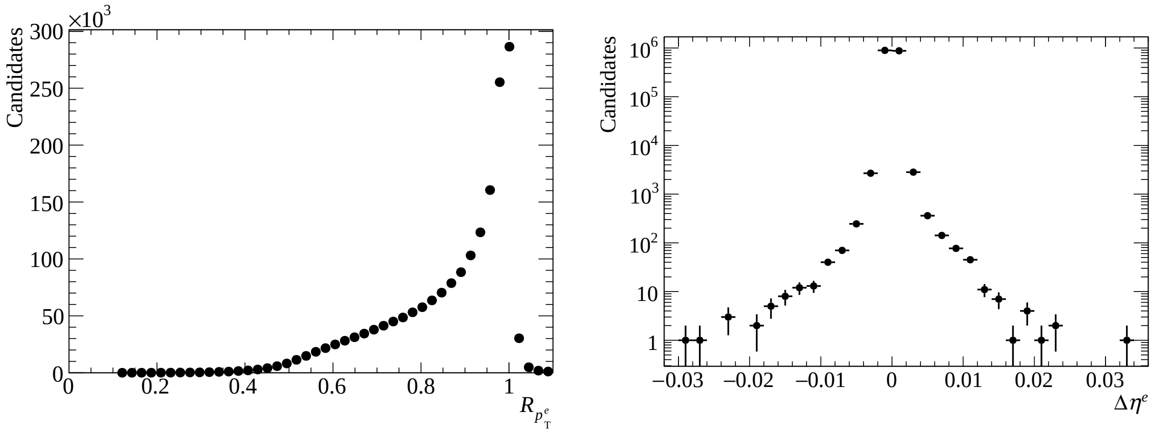
Measurement of the *W* boson production cross-section at LHCb

This chapter presents measurements of W boson production cross-sections in the electron final state at $\sqrt{s} = 8\,\text{TeV}$ using data corresponding to an integrated luminosity of $2\,\text{fb}^{-1}$ collected by the LHCb detector [1]. Measurements are made in eight bins of lepton pseudorapidity and for each charge. In addition, cross-section ratios and the production charge asymmetry are presented. Finally, the ratio of branching fractions $\mathcal{B}(W \to e\nu)/\mathcal{B}(W \to \mu\nu)$ is determined in the forward region. The electrons are required to have more than 20 GeV of transverse momentum and to lie between 2.00 and 4.25 in pseudorapidity. The results are corrected for QED final-state radiation (hereinafter denoted as "Born level"). These requirements define the fiducial region of the measurements.

5.1 General considerations

Analyses of the W boson are complicated by the fact that neutrinos are not detected nor can they be reconstructed from missing momentum in the transverse plane at LHCb as they are in ATLAS and CMS due to the fact that, as a forward detector, it only covers ~2.5% of the solid angle. This has the additional consequence that measurements of $W \to e\nu(\mu\nu)$ have an irreducible background from $Z \to ee(\mu\mu)$ when only one lepton is in acceptance.

The $W \to e\nu$ analysis presents different challenges to the μ final state. Compared to muons, the measurement of electrons has an additional experimental difficulty arising



(a) (b)

Figure 5.1: (a) The ratio of measured and true transverse momentum, $R_{p_T^e}$. (b) The difference in measured and true pseudorapidity. The uncertainties are statistical only.

from the bremsstrahlung emitted when traversing the detector material. While the emitted photon energy can often be recovered for low-energy particles, electrons from W and Z boson decays tend to have high momentum and transverse momentum, with bremsstrahlung photons that are not generally well-separated from the lepton. Coupled with the fact that individual LHCb ECAL cells saturate by design at a transverse energy of approximately 10 GeV, this leads to a poor energy measurement and a reconstructed distribution of transverse momentum, p_T^e, which differs significantly from the true transverse momentum of the electrons as shown in Fig. 5.1 (a).

By contrast, the electron direction is measured well as seen in Fig. 5.1 (b), so that the differential cross-sections in lepton pseudorapidity have negligible bin-to-bin migrations. Specifically, the resolution of the LHCb detector in reconstructed electron pseudorapidity, η^e, is found to be two parts per mille with respect to the size of the bins used in the differential measurements. Therefore, unfolding is not performed in pseudorapidity. For these measurements, the electron candidates are required to lie within $2.00 < \eta^e < 4.25$. An upper limit of $\eta^e < 4.25$ is imposed due to the limited acceptance of the calorimetry. The $W \to \mu\nu$ measurement does not suffer from this limitation and extends to $\eta^\mu = 4.50$.

Despite the challenges, there are strong justifications for carrying out measurements of $W \to e\nu$ production. Firstly, these act as independent cross-checks to measurements of $W \to \mu\nu$ [22, 24] with largely independent systematics. Secondly, they provide further input in constraining proton PDFs and, finally, enable a precise test of lepton universality to be made at a hadron collider.

Trigger	Requirements
`L0Electron`	$E_{\mathrm{T}}^{e} > 3\,\mathrm{GeV}$, $N_{\mathrm{SPD}} < 600$
`Hlt1SingleElectronNoIP`	$\chi^2/\mathrm{ndf} < 3$, $p^e > 20\,\mathrm{GeV}$, $p_{\mathrm{T}}^{e} > 10\,\mathrm{GeV}$
`Hlt2SingleTFVHighPtElectron`	$\chi^2/\mathrm{ndf} < 20$, $p_{\mathrm{T}}^{e} > 15\,\mathrm{GeV}$, $E_{\mathrm{PRS}} > 50\,\mathrm{MeV}$, $E_{\mathrm{ECAL}}/p^e > 0.1$, $E_{\mathrm{HCAL}}/p^e < 0.05$

Table 5.1: The trigger lines used in the measurement of $W \to e\nu$ [142]. N_{SPD} is the SPD multiplicity.

5.2 Candidate selection

The production of $W \to e\nu$ is characterised by a single, isolated high-p_{T} charged particle originating from a PV and with a large energy deposit in the electromagnetic calorimeter. However, several other physics processes can mimic this signature. Significant EW backgrounds include $Z \to ee$ with one daughter in the LHCb acceptance, $Z \to \tau\tau$ and $W \to \tau\nu$, where the τ decays to a final state containing an electron. Prompt photon+jets contributes in cases where the photon converts to an ee pair and only one daughter is reconstructed and selected. Hadronic backgrounds stem from four sources: hadron mis-identification (hereinafter denoted as "fake electrons"), semileptonic heavy flavour (HF) decay, decay in flight, and $t\bar{t}$ production. Some of the contributing processes are illustrated in Fig. 5.2. The procedure for determining the sizes of the various contributions is explained in Sect. 5.5 and fractions as a function of η^e are given in Fig. 5.19.

5.2.1 Trigger and pre-selection

Electron candidates are required to be triggered through the Level-0, HLT1, and HLT2 single electron lines listed in Table 5.1. To be identified as electrons, the candidates are required to deposit energy $E_{\mathrm{ECAL}} > 0.1p^e$ in the ECAL while depositing relatively little energy $E_{\mathrm{HCAL}} < 0.05p^e$ in the HCAL, where p^e is the momentum of the electron. The candidates are also required to have deposited energy, E_{PRS}, of more than 50 MeV in the PRS. The background formed by $Z \to ee$ events with both electrons in the LHCb acceptance is largely removed using a dedicated di-electron software trigger.

In addition, global event cut (GEC) requirements are applied at Level-0 to prevent high-multiplicity events from dominating the processing time. The requirement relevant for this analysis is a veto on SPD hits being equal to or greater than 600. This is

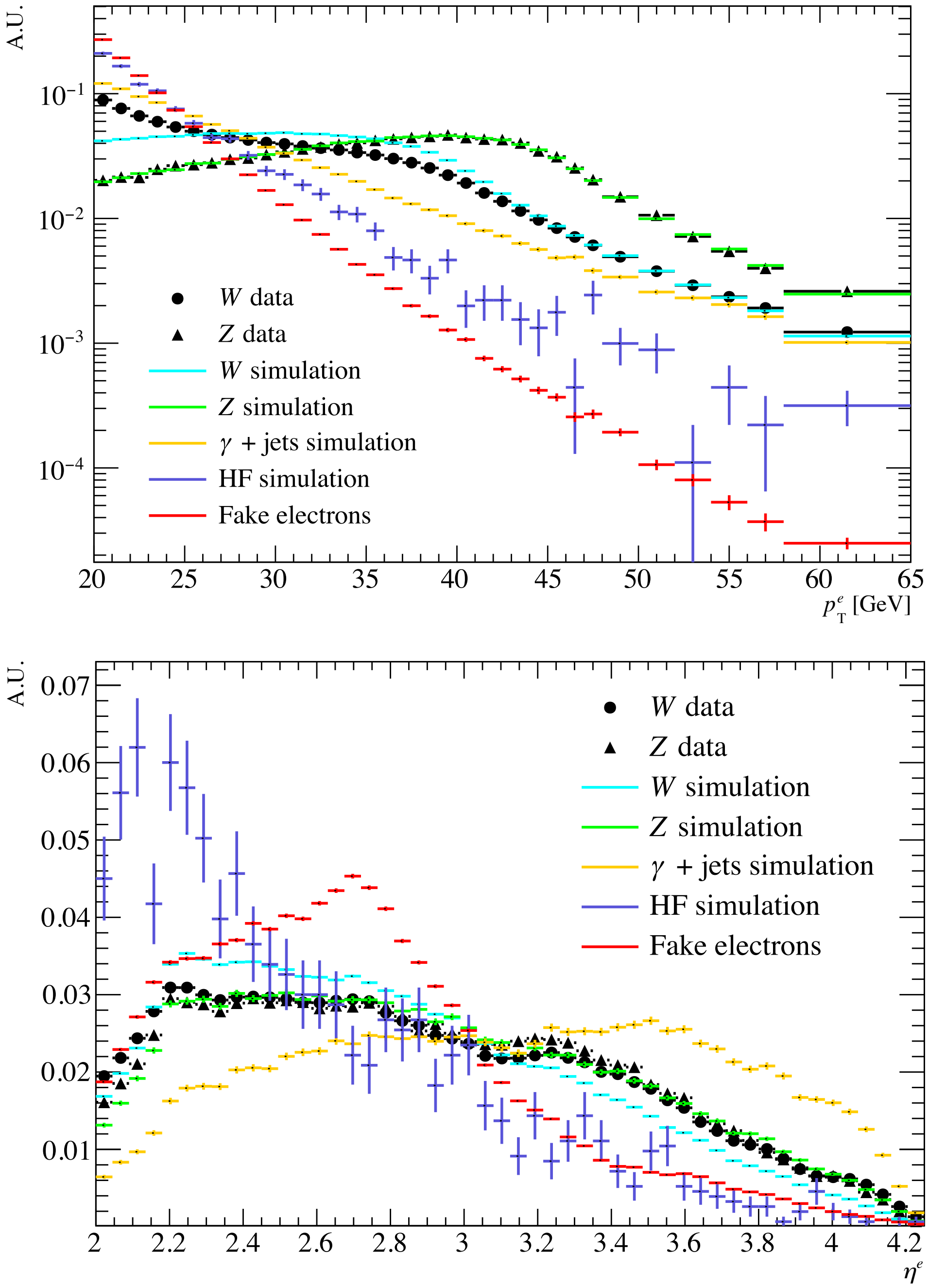


Figure 5.2: A selection of processes (after requirements) that populate the signal region, normalised to unity, as a function p_T^e (top) and η^e (bottom). Z data is included to illustrate the level of agreement with simulation. Fake electrons are extracted from a control region in data as explained in Sect. 5.5.4.

Description	Requirements
An energy deposit in the PRS	$E_{\rm PRS} > 50\,{\rm MeV}$
An energy deposit in the ECAL	$E_{\rm ECAL}/p^e > 0.1$
An energy deposit in the HCAL	$E_{\rm HCAL}/p^e < 0.05$
Transverse momentum requirement	$p_{\rm T}^e > 20\,{\rm GeV}$
Relative momentum uncertainty	$\delta p^e/p^e < 0.1$

Table 5.2: The $W \to e\nu$ candidate pre-selection.

corrected for by using di-muon candidates since these have a less stringent requirement of $N_{\rm SPD} < 900$ imposed on them. Sect. 5.7.4 explains how the efficiency is determined and corrected for $W \to e\nu$ production.

The starting point for the analysis is the collection of long tracks with a track fit quality of $\chi^2/{\rm ndf} < 3$ and rejection of clones according to the Kullback-Leibler divergency as discussed in the previous chapter. A pre-selection[1,2] given in Table 5.2 is applied to the electron candidates. Here, the relative uncertainty on the momentum, δp^e, is required to be less than 10% to ensure that the charge is measured well.

Additionally, a sample of $Z \to ee$ candidates is used to determine efficiencies and perform other studies. These electron candidates are subjected to identical requirements, but have a looser transverse momentum requirement of $p_{\rm T}^e > 10\,{\rm GeV}$ and a requirement on the reconstructed di-electron invariant mass of $M_Z > 40\,{\rm GeV}$. Furthermore, a sample of same-sign, $e^{\pm}e^{\pm}$, combinations with the same requirements as above are used to provide a data-based estimate for hadronic backgrounds in the Z sample as explained in Sect. 5.7.5.

[1] The pre-selection is referred to as “stripping” in LHCb jargon.

[2] Some of the requirements at the trigger level and in the pre-selection, while nominally equivalent, are based on different calibrations of the detector and so do not have identical efficiencies.

5.2.2 Further requirements

While the pre-selection is applied to both W and Z candidate events and so efficiencies must be extracted from simulation, further requirements (referred to as the "tight selection" subsequently) are imposed on the single electron sample to achieve increased background suppression. The efficiency for these requirements, however, can be determined in data using the Z sample.

The PID requirements are made more stringent with $E_{\mathrm{ECAL}}/p^e > 0.15$ and $E_{\mathrm{HCAL}}/p^e < 0.0075$. These distributions are illustrated in Fig. 5.3. Further tracking requirements are applied by demanding hits in the TT with $N_{\mathrm{TT}} > 0$ and by requiring the χ^2 of the match between VELO and T-station track segments, χ^2_{match}, to be better than 10. The momentum resolution and χ^2_{match} distributions are shown in Fig. 5.4. The tracking quality requirement is loosened to 15 in the overlap region of the detector, $1.2 < |\phi| < 1.8$, where ϕ is the azimuthal angle. This is motivated by inferior agreement between data and simulation in this region.

The remainder of the tight selection exploits other physical features of the process. Firstly, electrons that originate from the W boson are prompt. This is in contrast to leptons that come from decays of HF mesons or τ leptons which tend to have displaced vertices measurable in the detector. This motivates the requirement of the impact parameter (IP) to be less than 0.04 mm. A loose requirement on χ^2_{IP}, where χ^2_{IP} is the modulus of the difference between the χ^2 of the PV reconstructed with and without the candidate electron, is applied to remove a poorly modelled region.

Another discriminant against hadronic processes is the fact that electrons from the W boson tend to be isolated: there is relatively little p_{T} in a cone surrounding the lepton. On the other hand, leptons originating from hadronic decays, or fake electrons, tend to have hadrons travelling alongside them. The isolation requirement is set to be $I^e_{\mathrm{T}} > 0.9$, where I^e_{T} is defined as

$$I^e_{\mathrm{T}} \equiv \frac{p^e_{\mathrm{T}}}{p^e_{\mathrm{T}} + E^{\gamma}_{\mathrm{T}} + p^{\mathrm{ch}}_{\mathrm{T}}}. \tag{5.1}$$

Here, E_{T} is the sum of the transverse component of neutral energy in the annular cone with $0.1 < R < 0.5$, where $R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2}$ and $\Delta\eta$ and $\Delta\phi$ are the differences in the pseudorapidity and azimuthal angle between the candidate and the particle being considered and $p^{\mathrm{ch}}_{\mathrm{T}}$ is the scalar sum of the transverse momenta of charged tracks in the same annulus.

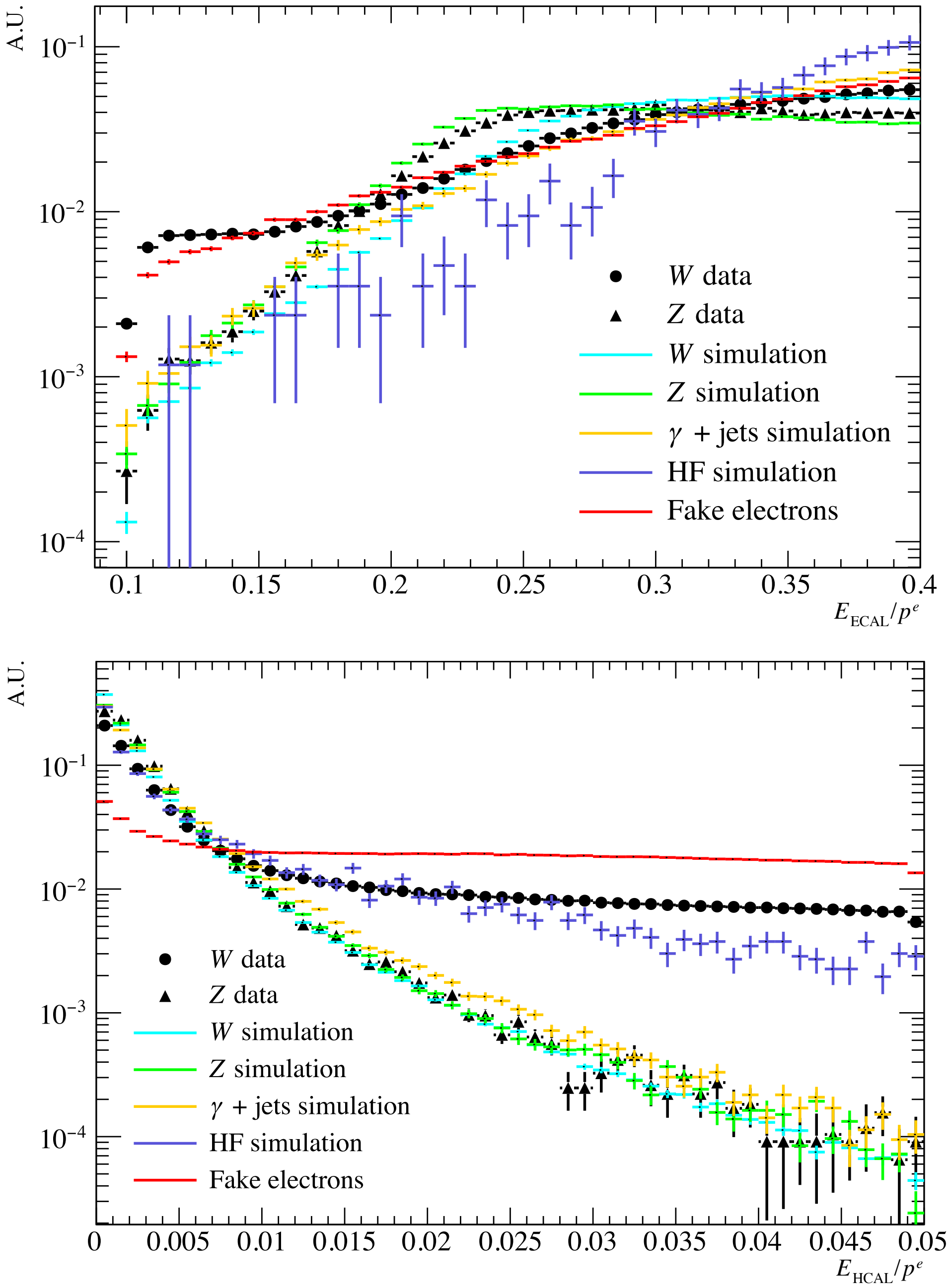


Figure 5.3: The relative energy deposition in the ECAL (top) and the HCAL (bottom) for a selection of processes. All other selection requirements are applied. Same-sign subtraction is performed on the Z samples to remove hadronic backgrounds (in the case of data) and charge mis-identification (for data and simulation).

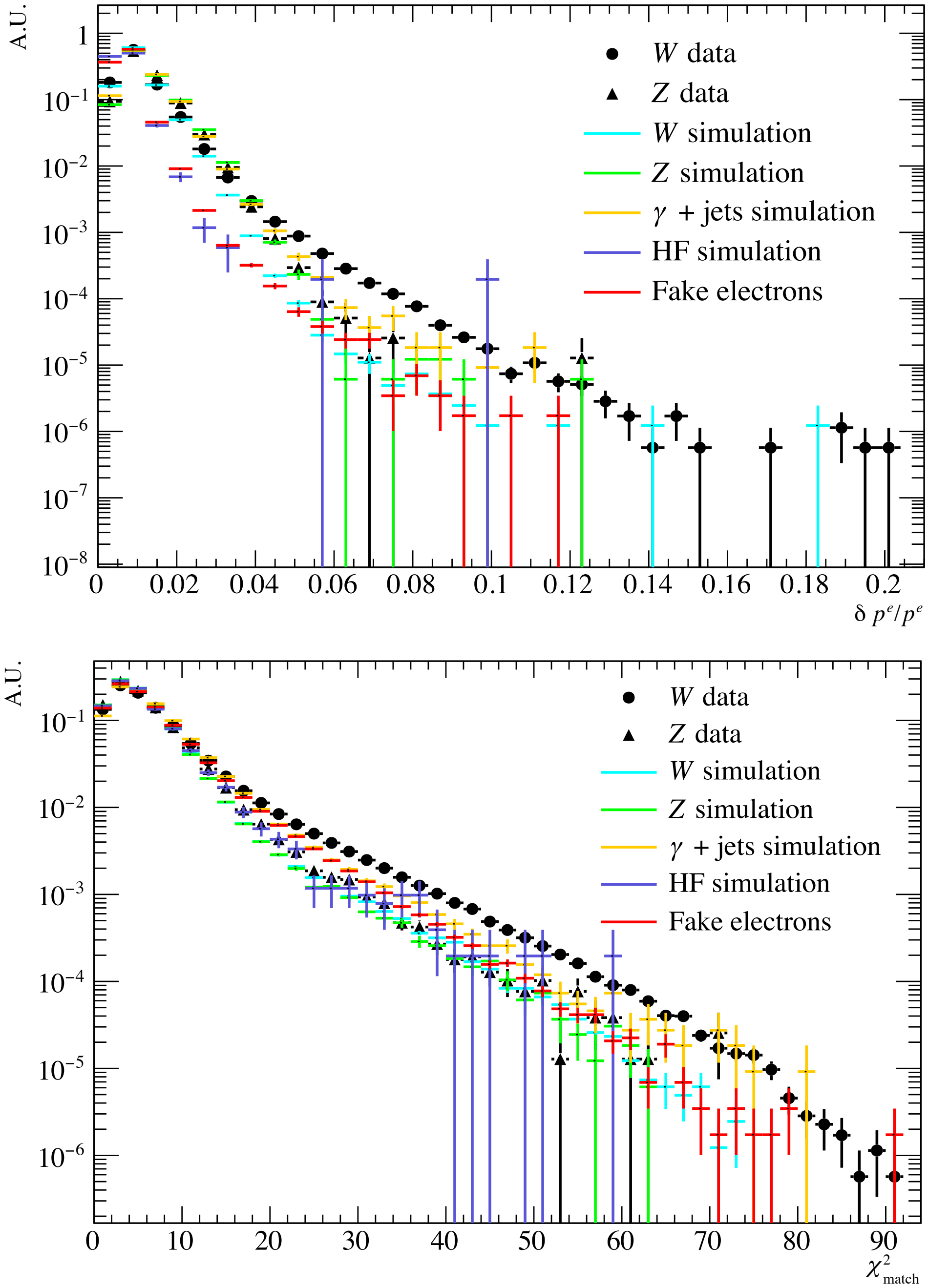


Figure 5.4: The fractional momentum resolution (top) and track quality (bottom) for a selection of processes. All other selection requirements are applied. Same-sign subtraction is performed on the Z samples to remove hadronic backgrounds (in the case of data) and charge mis-identification (for data and simulation).

Requirement	Description
$N_{\rm TT} > 0$	Hits in the Tracker Turicensis
$\chi^2_{\rm match} < 10$ in $\lvert\phi\rvert < (>) 1.2 (1.8)$	Tight track quality requirement
$\chi^2_{\rm match} < 15$ in $1.2 < \lvert\phi\rvert < 1.8$	Adjusted tight track quality requirement
$E_{\rm ECAL}/p^e > 0.15$	ECAL PID requirement
$E_{\rm HCAL}/p^e < 0.0075$	HCAL PID requirement
IP $< 0.04\,$mm	Require the electron to be prompt
$\log_{10}(\chi^2_{\rm IP}) < 1.6$	Remove poorly modelled $\chi^2_{\rm IP}$ region
$I^e_{\rm T} > 0.9$	Require the electron to be isolated

Table 5.3: The $W \to e\nu$ tight selection.

An annular cone, rather than a cone, is used as bremsstrahlung photons contribute to the energy in the cone, but are mostly contained in the range $0.0 < R < 0.1$ and so can be excluded. A ratio incorporating $p^e_{\rm T}$ instead of a simple absolute requirement on energy in the cone is used for two reasons. Firstly, the requirement is most stringent at low $p_{\rm T}$ where hadronic backgrounds dominate. Secondly, the $p_{\rm T}$ in the cone is correlated to $p^e_{\rm T}$ for hadrons misidentified as electrons since these are typically jets, while being uncorrelated for true electrons as the particles in the cone originate from the underlying event in that case. The IP and $I^e_{\rm T}$ distributions are plotted in Fig. 5.5. In summary, the tight selection summarised in Table 5.3 is applied to the dataset which increases the contribution of signal in the selected sample to ~61%.

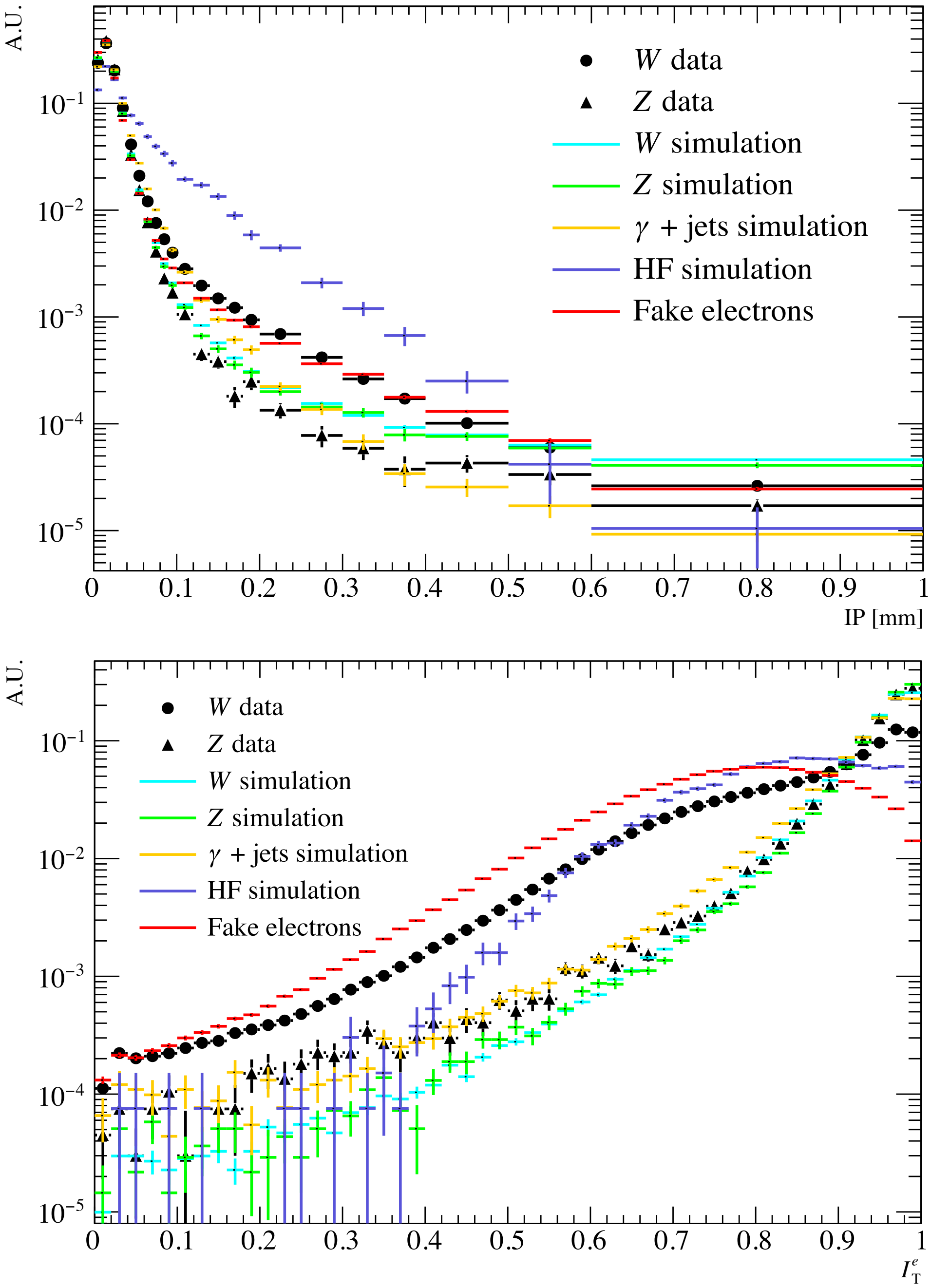


Figure 5.5: The IP (top) and I^e_{T} (bottom) distributions for a selection of processes. All other selection requirements are applied. Same-sign subtraction is performed on the Z samples to remove hadronic backgrounds (in the case of data) and charge mis-identification (for data and simulation).

Decay mode	Acceptance requirement(s)	N_{ev}
$W \to e\nu$	$p_{\mathrm{T}}^{e} > 4\,\mathrm{GeV}$	5M
$W \to \tau\nu$	$p_{\mathrm{T}}^{e} > 10\,\mathrm{GeV}$	1M
$Z \to ee$	$M_Z > 40\,\mathrm{GeV}$, $p_{\mathrm{T}}^{e} > 4\,\mathrm{GeV}$	2M
Low-mass DY	$M_Z > 5\,\mathrm{GeV}$, $p_{\mathrm{T}}^{e} > 1\,\mathrm{GeV}$	1M
$Z \to \tau\tau$	$p_{\mathrm{T}}^{\tau} > 4\,\mathrm{GeV}$	4M
$WW \to lX$	$p_{\mathrm{T}}^{l} > 15\,\mathrm{GeV}$	200k
$WZ \to lX$	$p_{\mathrm{T}}^{l} > 15\,\mathrm{GeV}$	200k
$W\gamma \to lX$	$p_{\mathrm{T}}^{l} > 4\,\mathrm{GeV}$	200k
$Z\gamma \to lX$	$p_{\mathrm{T}}^{l} > 4\,\mathrm{GeV}$	200k
$gg \to t\bar{t}$	$p_{\mathrm{T}}^{l} > 17\,\mathrm{GeV}$	200k
$q\bar{q} \to t\bar{t}$	$p_{\mathrm{T}}^{l} > 17\,\mathrm{GeV}$	200k
$\gamma(\to ee)$+jets	$p_{\mathrm{T}}^{\gamma} > 20\,\mathrm{GeV}$	10M
$c\bar{c} \to eX$	$p_{\mathrm{T}}^{e} > 18\,\mathrm{GeV}$	1M
$b\bar{b} \to eX$	$p_{\mathrm{T}}^{e} > 18\,\mathrm{GeV}$	1M
Minimum bias	$p_{\mathrm{T}}^{K/\pi} > 18\,\mathrm{GeV}$	1M

Table 5.4: Overview of the simulated samples used in the analysis. Here, l refers to an electron or a muon and X indicates any additional particles. The last column lists the number of events, N_{ev}, generated. "M" denotes a million and "k" denotes a thousand. All samples have an acceptance requirement of $\theta_l < 400\,\mathrm{mrad}$ applied to them, where θ_l is the polar angle of the lepton.

5.3 Simulated samples

An overview of the simulated samples used in the analysis is provided by Table 5.4. $Z \to ee$ production is modelled using two samples: one with a $M_Z > 40\,\mathrm{GeV}$ requirement, which is the dominant contribution, and another with a $M_Z > 5\,\mathrm{GeV}$ requirement referred to as the low mass Drell-Yan (DY) sample in what follows. For the latter, events with $M_Z > 40\,\mathrm{GeV}$ are vetoed to avoid double counting. A high-p_{T} e/μ is required to be in LHCb acceptance for the di-boson samples. However, the bosons are not otherwise forced to decay into a particular mode.

Production of prompt photons in association with jets is modelled using a leading-order matrix element. The photon is required to have at least $20\,\mathrm{GeV}$ of p_{T} and to be in LHCb acceptance. An on-shell photon is only able to convert to an ee pair in the presence of material. What is referred to as "Minimum bias" simulation in this chapter is Pythia 8 simulation of di-jet production. However, an additional generator requirement, whereby only events with a high-p_{T} charged kaon or pion are stored, is applied to increase the probability that these candidates satisfy the W measurement requirements.

5.4 Cross-section determination

The production cross-section for $W \to e\nu$ is measured in each bin of lepton pseudorapidity and for each charge with electron transverse momentum in excess of 20 GeV. The cross-sections are determined from

$$\sigma_i^{W\to e\nu} = \frac{N_i^W}{\mathcal{A}_i\,\mathcal{L}\,\epsilon_i^{\mathrm{tot}}\,f^{\mathrm{FSR}}}, \tag{5.2}$$

where N_i^W is the signal yield in the range $20 < p_{\mathrm{T}}^e < 65$ GeV obtained from a fit to p_{T}^e in bin i of η^e, $\epsilon_i^{\mathrm{tot}}$ is the total efficiency in that bin, and $\mathcal{L}$ is the integrated luminosity as discussed in Sect. 4.3.4. The signal yields are corrected for excluded candidates with $p_{\mathrm{T}}^e > 65$ GeV by computing a charge-dependent acceptance factor, $\mathcal{A}_i$. The extraction of the signal yield, N_i^W, is explained in Sect. 5.5. Subsequently, the acceptance factors, $\mathcal{A}_i$, are discussed in Sect. 5.6. The determination of the total efficiency, $\epsilon_i^{\mathrm{tot}}$, is presented in Sect. 5.7.

The results of the measurement are quoted at Born level to enable comparisons to theoretical predictions that do not incorporate the effect of QED final-state radiation (FSR). Correcting to Born level also enables a comparison to be made with the measurement of $W \to \mu\nu$. Corrections due to FSR, f^{FSR}, are computed separately using PYTHIA 8 and HERWIG++ [143] and then averaged. The corrections are listed in Tables 5.10 and 5.11 in Sect. 5.10.2 so that the measurement can be compared to a prediction that incorporates the effect of FSR.

Process	Cross-section [pb]
$W\gamma$	11.4 ± 0.1
$Z\gamma$	2.07 ± 0.3
WW	5.771 ± 0.007
WZ	1.0132 ± 0.0005
$t\bar{t}$	11.50 ± 0.09

Table 5.5: Cross-sections for rare processes with statistical uncertainties. These are computed using MCFM in the fiducial volume defined in Table 5.4.

5.5 Determination of the event yield

In total, 1 368 539 $W \to e\nu$ candidates in data fulfil the selection requirements. The signal yields are determined in eight bins of lepton pseudorapidity and for each charge. Binned maximum likelihood template fits to p_{T}^{e} are performed following Ref. [144]. The extraction of templates and their normalisations are described in Sects. 5.5.1-5.5.5. The results of the fits are then presented in Sect. 5.5.6. Corrections and cross-checks to the fitting procedure are discussed in Sects. 5.5.7 and 5.5.8.

5.5.1 Single W and Z yields

Templates for $W \to e\nu$, $W \to \tau\nu$, $Z \to ee$ and $Z \to \tau\tau$ are taken from simulation. The known ratio of branching fractions [5] is used to constrain $W \to \tau\nu$ to $W \to e\nu$. The measured LHCb cross-section for $Z \to \mu\mu$ production [22] is used to constrain $Z \to ee$ and $Z \to \tau\tau$ in the fit and the uncertainty on the ratio of branching fractions in Ref. [5] is taken into account. The low-mass DY component is also taken from simulation and constrained to the corresponding MCFM 6.8 [145] NLO cross-section.

5.5.2 Rare processes

Contributions from $W\gamma$, $Z\gamma$, WW, WZ, and $t\bar{t}$ are included in the fits to the lepton p_{T} spectrum. Templates for these processes are obtained from Pythia 8 simulation and normalised to the MCFM NLO cross-sections listed in Table 5.5. These processes account for $(0.46 \pm 0.01)\%$ of the selected candidates and are denoted as “rare processes”.

5.5.3 Prompt photon+jets

Prompt photon production in association with jets has a relatively large cross-section of around 50 nb at 8 TeV for a $p_{\mathrm{T}} > 20$ GeV photon in LHCb acceptance as computed using MCFM at NLO. The process mimics the signal in cases where the photon converts into an ee pair in the detector material and only one daughter is reconstructed, triggered on, and selected by the $W \to e\nu$ selection.

The probability for conversion in the VELO is roughly 15% based on the radiation length and the mean free path for pair production. Simulation indicates that after requiring the electron to be reconstructed and the `Hlt2SingleTFVHighPtElectron` trigger to have fired in the event, the fraction of selected candidates is reduced to ~2%.

A sample of photon+jets is extracted from data by searching for an ee pair with mass below 50 MeV, where one electron is required to satisfy the pre-selection. The transverse momentum of the photon is required to be greater than 20 GeV. The sample is then further cleaned with a dedicated selection listed in Table 5.6.

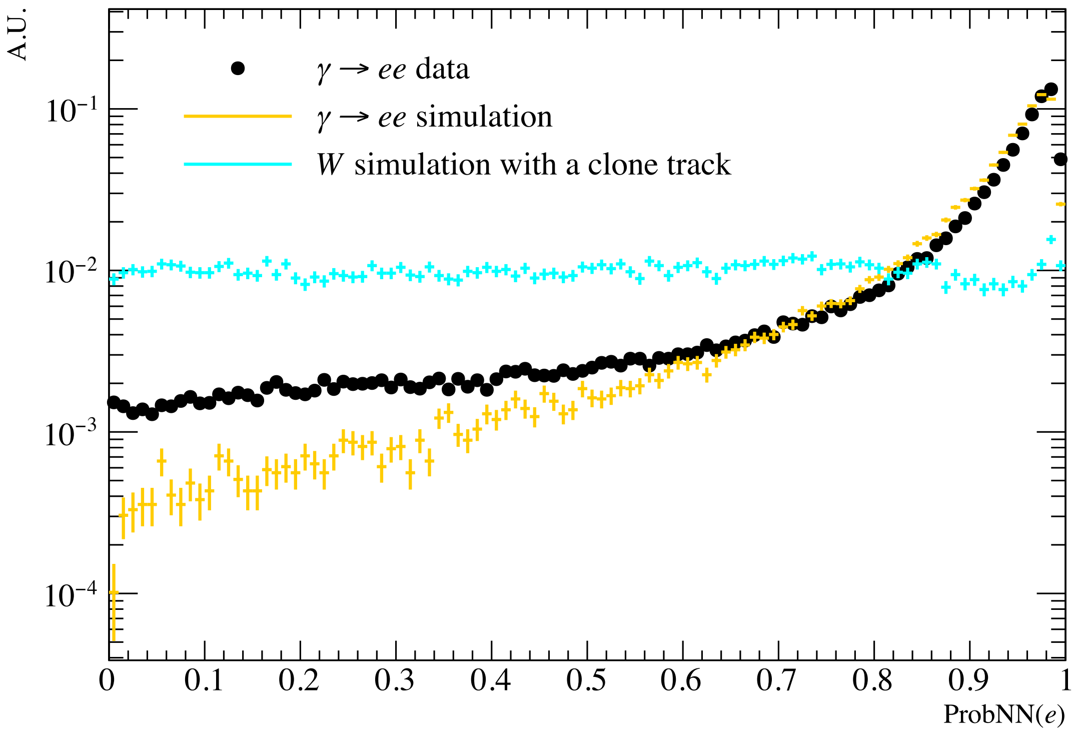
While several variables such as the position of the conversion vertex have weak discriminating power over backgrounds, the $\gamma \to ee$ selection was chosen to match the $W \to e\nu$ selection closely. This ensures biases in the modelling of efficiencies do not affect the normalisation of the sample in the fits to p_{T}^{e}. A less stringent selection is applied to the second electron from the conversion.

As an additional nuance, multivariate PID variables, ProbNN(e) and ProbNN(π), are used to suppress hadrons faking electrons in the low momentum range with $p^{e} < 100$ GeV. These variables have values between 0 and 1 representing a probability for a particular particle hypothesis. The variables have been trained using a Bayesian Neural Network algorithm combining information from all subdetectors [113, 146]. The distributions for the electron and pion hypotheses are shown in Fig. 5.6.

As in the case of the $W \to e\nu$ selection, isolation, shown in Fig. 5.7, is seen to have the largest discriminating power. It is observed that the selected data sample agrees reasonably well with simulation of photon+jets as shown in Fig. 5.8.

Requirement	Description
$M^{\gamma} < 50\,\text{MeV}$	Require the photon to be roughly on-shell
$I_{\text{T}}^{\gamma} > 0.9$	Require the photon to be isolated
(*) ProbNN(e) > 0.5	LHCb e PID variables for $p^e < 100\,\text{GeV}$
(*) ProbNN(π) < 0.01	LHCb π PID variables for $p^e < 100\,\text{GeV}$
(*) $p_{\text{T}}^{e} > 3\,\text{GeV}$	Loose p_{T} requirement
(*) $N_{\text{TT}} > 0$	Hits in the Tracker Turicensis
(*) $E_{\text{PRS}} > 50\,\text{MeV}$	Energy in the Pre-Shower Detector
(*) $E_{\text{ECAL}}/p^e > 0.1$	Loose ECAL identification requirement
(*) $E_{\text{HCAL}}/p^e < 0.05$	Loose HCAL identification requirement
$p_{\text{T}}^{e} > 20\,\text{GeV}$	Tight p_{T} requirement
`L0Electron`	Trigger requirement at Level-0 level
`Hlt1SingleElectronNoIP`	Trigger requirement at HLT1 level
`Hlt2SingleElectronTFHighPt`	Trigger requirement at HLT2 level
$\chi^2_{\text{match}} < 10$ in $\lvert\phi\rvert < (>) 1.2 (1.8)$	Tight track quality requirement
$\chi^2_{\text{match}} < 15$ in $1.2 < \lvert\phi\rvert < 1.8$	Adjusted tight track quality requirement
$E_{\text{ECAL}}/p^e > 0.15$	ECAL PID requirement
$E_{\text{HCAL}}/p^e < 0.0075$	HCAL ID requirement
IP$< 0.04\,\text{mm}$	Require the electron to be prompt
$\log_{10}(\chi^2_{\text{IP}}) < 1.6$	Remove poorly modelled χ^2_{IP} region

Table 5.6: The $\gamma(\rightarrow ee)$+jets selection. Requirements marked with an asterisk, (*), are applied to both daughters while others must be met by at least one of the daughters.



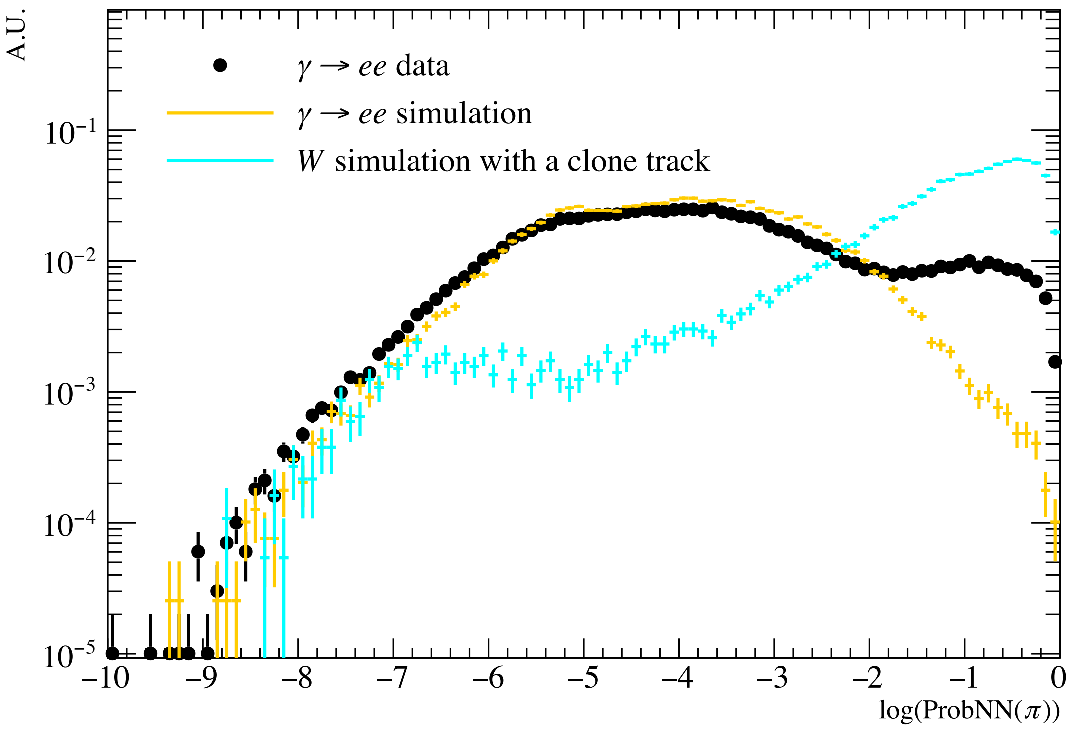


Figure 5.6: The ProbNN distribution for electron (top) and pion (bottom) hypotheses for the daughters in photon conversions. Only candidates with less than 100 GeV of momentum are considered. The distribution from data (black points) is compared to γ+jets simulation, and the dominant EW background, $W \to e\nu$ combined with a clone track. All other selection requirements are applied. Distributions are normalised to unity.

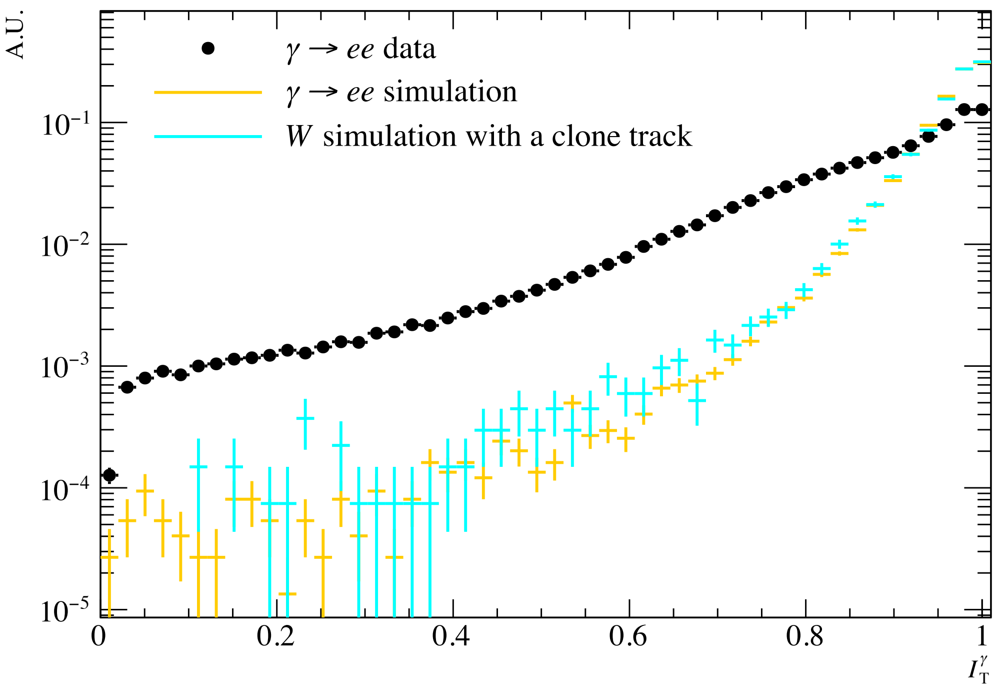
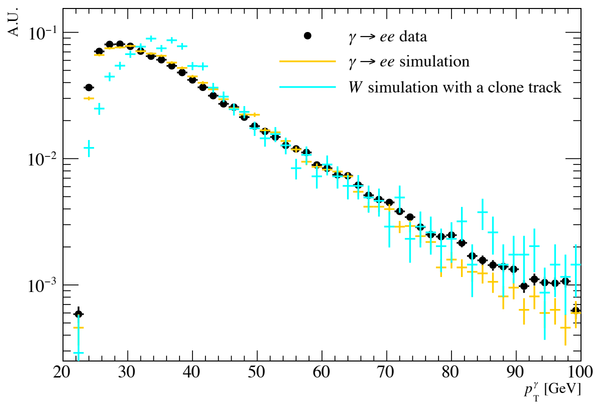
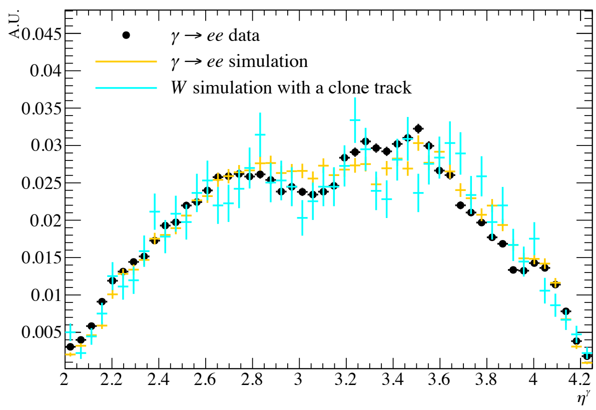


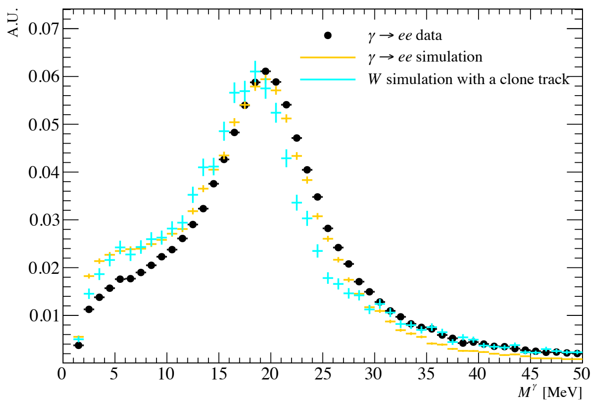
Figure 5.7: The I_{T}^{γ} distribution of the photon in conversions. The distribution from data (black points) is compared to γ+jets simulation, and the dominant EW background, $W \to e\nu$ combined with a clone track. All other selection requirements are applied. Distributions are normalised to unity.



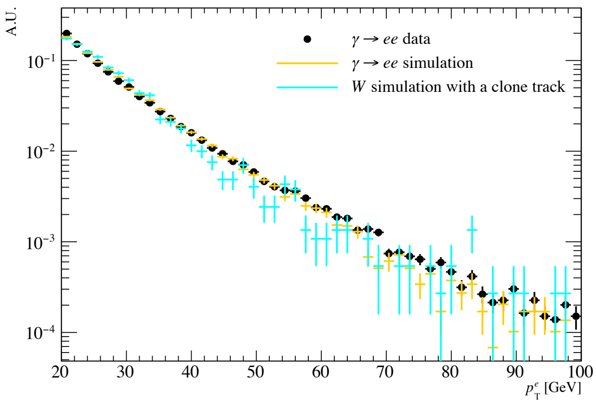
(a)



(b)

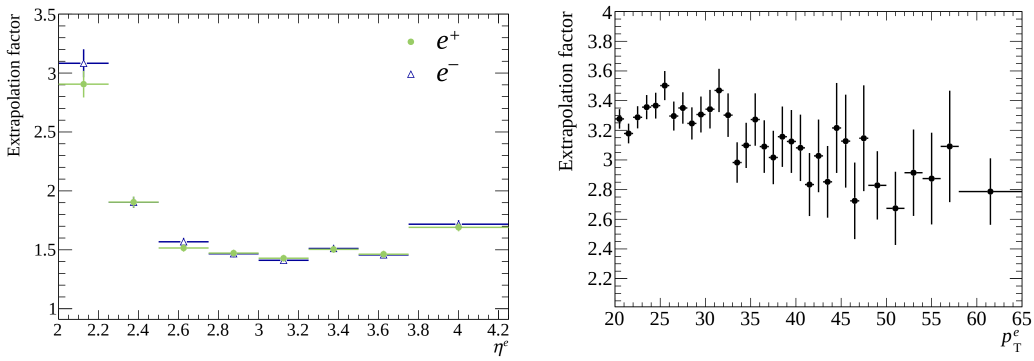


(c)



(d)

Figure 5.8: (a) p_{T}, (b) η, and (c) mass of the γ in photon conversions. (d) shows the p_{T} of the leading daughter in the conversion. The distribution from data (black points) is compared to γ+jets simulation, and the dominant EW background, $W \to e\nu$ combined with a clone track. All selection requirements are applied. Distributions are normalised to unity.



(a) (b)

Figure 5.9: $\gamma(\to ee)$+jets extrapolation factors. (a) shows the factors as a function of η^e and split by charge. (b) plots them as a function p_T^e but inclusively in charge and η^e.

Pythia 8 simulation is used to extrapolate the p_T^e distribution of the leading[3] electrons in data to the signal region (SR) defined by the selection listed in Table 5.3. The extrapolation is performed as a function of p_T^e, η^e, and charge. Eq. 5.3 shows how the photon+jets contribution to the $W \to e\nu$ SR is obtained by extrapolating from the $\gamma \to ee$ SR using simulation.

$$[\gamma + jets]_{\text{data}}^{W\to e\nu\ \text{SR}} = [\gamma + jets]_{\text{data}}^{\gamma\to ee\ \text{SR}} \cdot \frac{[\gamma + jets]_{\text{sim}}^{W\to e\nu\ \text{SR}}}{[\gamma + jets]_{\text{sim}}^{\gamma\to ee\ \text{SR}}} \tag{5.3}$$

The extrapolation factors in η^e are shown in Fig. 5.9 (a) while Fig. 5.9 (b) shows these as a function p_T^e, but inclusively in η^e and charge.

In reality, the data are an admixture of different processes. Most crucially, a $(W \to)e$ track coupled with a clone track comprises a significant fraction of selected events (roughly 10%). Leakage from $W \to e\nu$, $W \to \tau\nu$, $Z \to ee$, and $Z \to \tau\tau$ is removed using a simulation-based estimate. Another potential background is high-p_T π^0's from QCD processes. The process $\pi^0 \to \gamma\gamma$ is largely removed by the requirement on I_T^γ due to the fact that hadrons predominantly travel in close proximity in $\eta - \phi$ space with a p_T distribution that differs from EW processes. Their presence can be inferred from Fig. 5.7. Subdominant hadronic processes on the other hand, such as Dalitz decays, are largely removed by the requirement on mass.

[3]The leading electron must satisfy all the requirements detailed in Table 5.6 while its partner must only satisfy those marked with an asterisk (*).

5.5.4 Fake electron background

Hadron mis-ID occurs when hadrons begin to shower early in the ECAL, giving a shower profile similar to that of electrons. These hadrons, however, will tend to deposit fractionally more energy in the HCAL than genuine electrons. They will also frequently travel alongside other hadrons so will be less isolated on average. This motivates one to extract a p_{T} template for this process by examining the isolation and HCAL energy requirements of the selection. Specifically, if a region in the I_{T}^{e}-E_{HCAL}/p^{e} plane can be found that is predominantly populated by fake electrons then the p_{T} spectrum of candidates in this region can serve as a template in the fits.

Different physical processes that contribute to the SR are illustrated in Fig. 5.10. Minimum bias simulation is seen to populate the full hadronic deposition range with I_{T}^{e} mostly in the range from 0.5 to 1. In contrast, $W \to e\nu$ simulation and prompt photon data are seen to lie predominantly in the high-I_{T}^{e} and low-E_{HCAL}/p^{e} region of the I_{T}^{e}-E_{HCAL}/p^{e} plane. However, Fig. 5.10 (a) and (b) illustrate that residual amounts of $W \to e\nu$ and prompt photon contaminate the full spectrum of E_{HCAL}/p^{e} values in the medium-I_{T}^{e} region. This is problematic as these processes have a significantly higher mean p_{T}^{e} than fake electrons, thus potentially biasing the fake electron template. Finally, $b\bar{b}$ is less isolated than signal, but mostly restricted to lie close to $E_{\mathrm{HCAL}}/p^{e} = 0$.

These observations lead to selecting a control region defined by $[0.035, 0.05]$ in E_{HCAL}/p^{e} and $[0.55, 0.75]$ in I_{T}^{e} as providing a relatively clean sample of fake electrons. Remaining EW contamination in the selected region is estimated using simulation and explicitly subtracted. The SR is defined as $[0, 0.0075]$ in E_{HCAL}/p^{e} and $[0.9, 1]$ in I_{T}^{e}.

Figure 5.11 (top) indicates that the mean p_{T} of fake electrons does not vary in this plane by a discernible degree, so the fake electron p_{T} spectrum in the control region is expected to be representative of the signal region. To quantify the uncertainty associated with the choice of requirements, two other regions are defined and used as alternative templates,

- A larger region defined by $[0.03, 0.05]$ in E_{HCAL}/p^{e} and $[0.5, 0.8]$ in I_{T}^{e};
- A smaller region defined by $[0.04, 0.05]$ in E_{HCAL}/p^{e} and $[0.6, 0.7]$ in I_{T}^{e}.

The p_{T} spectra of the three regions are shown inclusively in Fig. 5.11 (bottom).

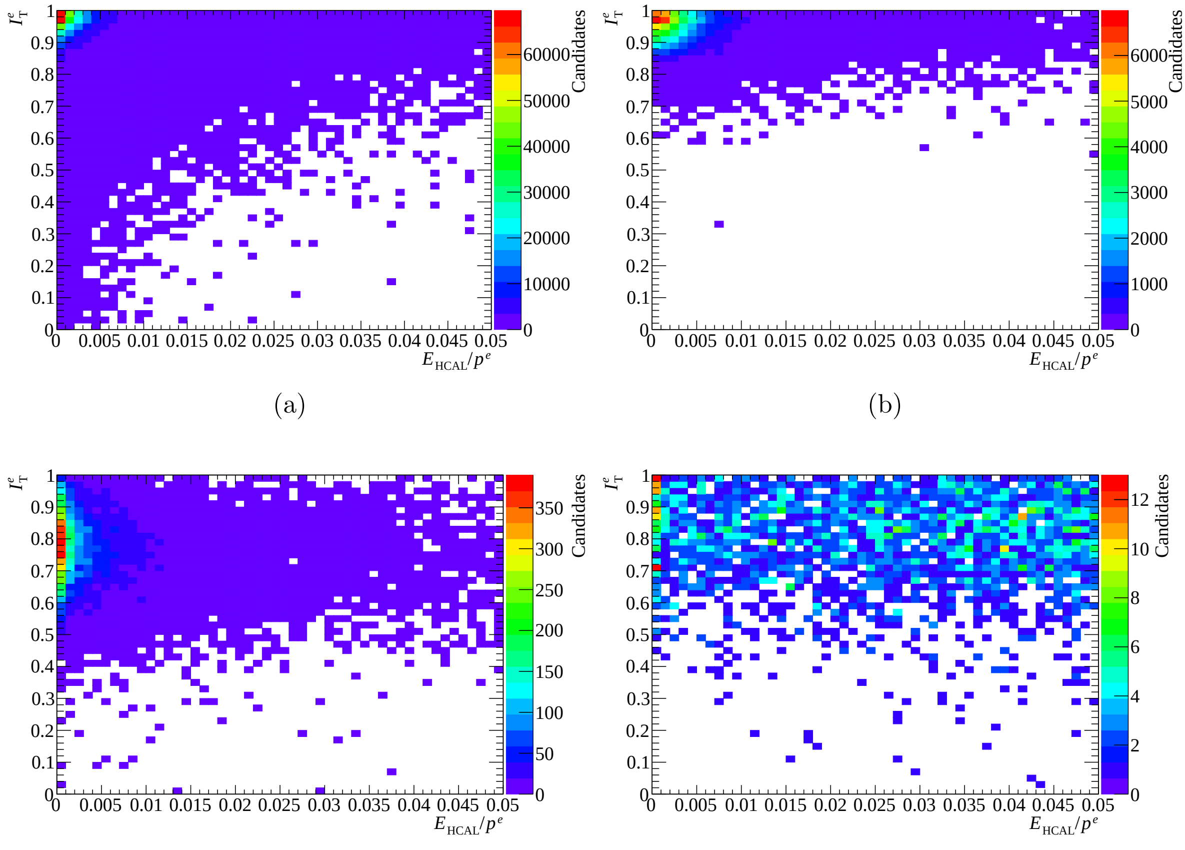


Figure 5.10: The number of candidates in the I_{T}^{e}-E_{HCAL}/p^{e} plane for (a) $W \to e\nu$ simulation, (b) prompt photon data, (c) $b\bar{b}$ simulation, and (d) minimum bias simulation.

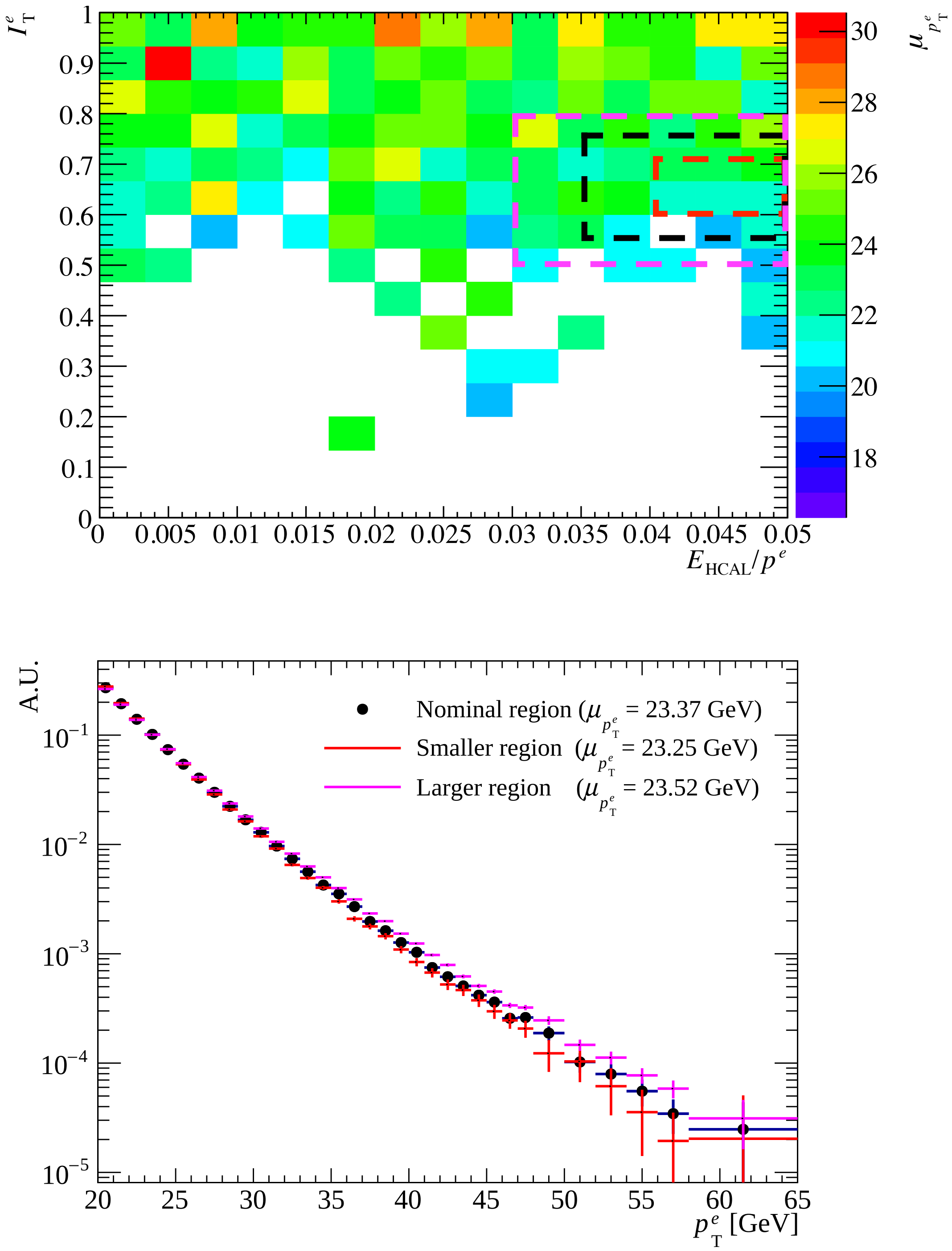


Figure 5.11: Top: The mean p_T of candidates, $\mu_{p_T^e}$, in the I_T^e-E_{HCAL}/p^e plane for minimum bias simulation. The nominal (black), smaller (red), and larger (purple) regions used to extract the fake electron template are indicated by dashed lines. Bottom: The p_T spectra of candidates extracted from three regions in the I_T^e-E_{HCAL}/p^e plane. All distributions are normalised to unit area and uncertainties are statistical only.

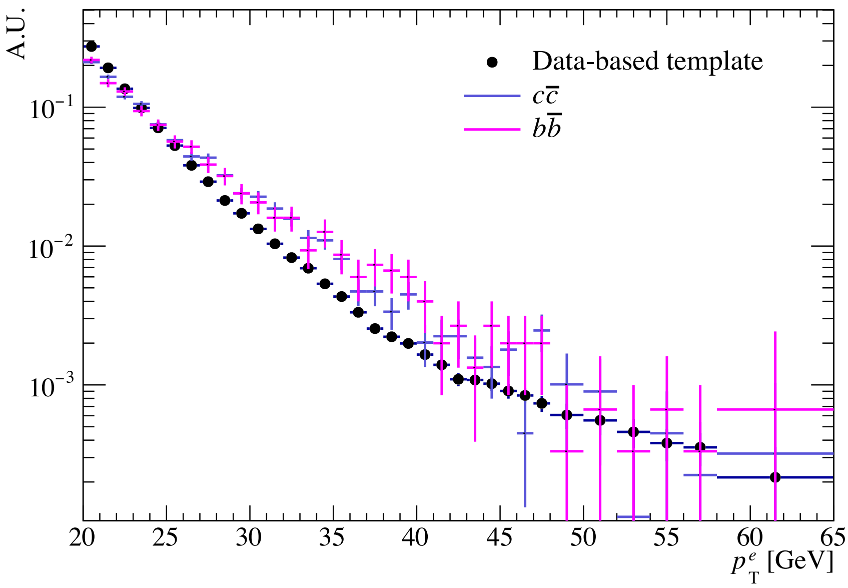


Figure 5.12: The data-driven heavy flavour template is compared to $c\bar{c}$ and $b\bar{b}$ simulation. EW contamination is subtracted using a simulation-based estimate. All distributions are normalised to unit area and uncertainties are statistical only.

5.5.5 Heavy flavour

The semileptonic decay of heavy flavour (HF) hadrons gives rise to genuine electrons. This background is suppressed using the IP requirement to exploit the long lifetimes of hadrons containing b and c quarks. The IP and χ^2_{IP} variables are first unbiased by refitting the primary vertex without the candidate electron and the partner electron in the case of Z boson production.

The HF component is described by a data-driven template obtained by applying the standard selection but requiring a significant impact parameter of more than 0.6 mm. Contamination from EW processes is estimated using simulation and explicitly subtracted. The shape of the template is compared to $c\bar{c}$ and $b\bar{b}$ simulation in Fig. 5.12.

The normalisation of the HF template is determined by a separate template fit to the $\log_{10}(\chi^2_{\mathrm{IP}})$ distribution and constrained in the fit to p_{T}^{e}. The contributions to the heavy flavour fit are listed below along with an explanation of how templates are extracted, normalised, and corrected.

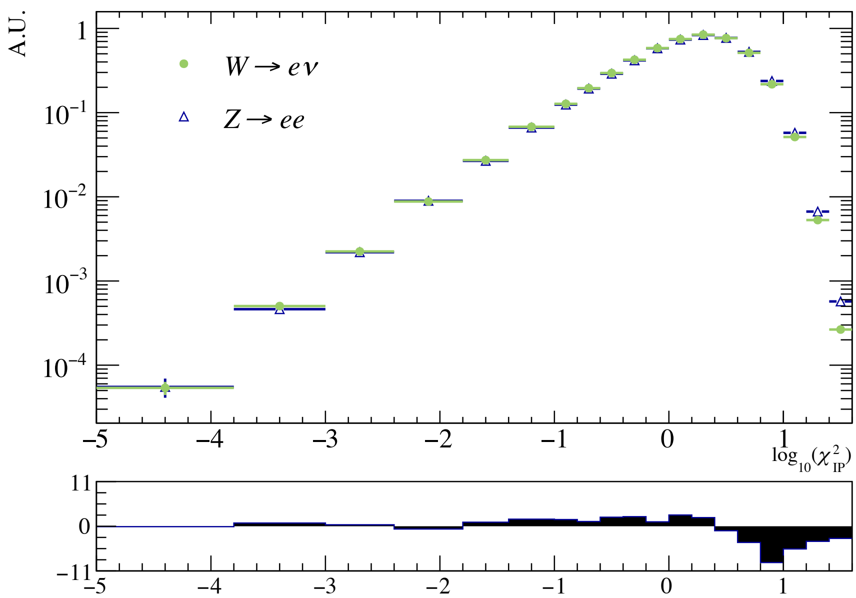


Figure 5.13: A comparison of the $\log_{10}(\chi^2_{\mathrm{IP}})$ distributions for $W \to e\nu$ and $Z \to ee$ in simulation. Pulls are shown underneath. All distributions are normalised to unit area and uncertainties are statistical only.

Templates to the HF fit

Prompt EW processes are modelled using simulation. Fig. 5.13 illustrates that, after unbiasing the impact parameter variables, the template shapes for W and Z in simulation become reasonably similar across most of the χ^2_{IP} range. However, discrepancies are seen in the range $0.5 < \log_{10}(\chi^2_{\mathrm{IP}}) < 1.6$, which is the most sensitive region in determining the heavy flavour contribution. Therefore, separate templates are used for tracks originating from W and Z boson decay.

EW bosons decaying to the τ final state also contribute at the few-percent level with a significantly different χ^2_{IP} distribution, as expected from the non-negligible τ lifetime. This component is taken from simulation and added to the $W \to e\nu$ and $Z \to ee$ templates, respectively, using the known ratios of branching fractions.

A separate template for the prompt photon+jets component is needed as the photon traverses ~17 mm from the primary vertex on average before the conversion occurs and thus results in tracks with an IP distribution different from EW processes. The template is taken from data with a data-driven normalisation that has a simulation-based acceptance correction applied to it in bins of χ^2_{IP} analogously to the procedure outlined in Sect. 5.5.3.

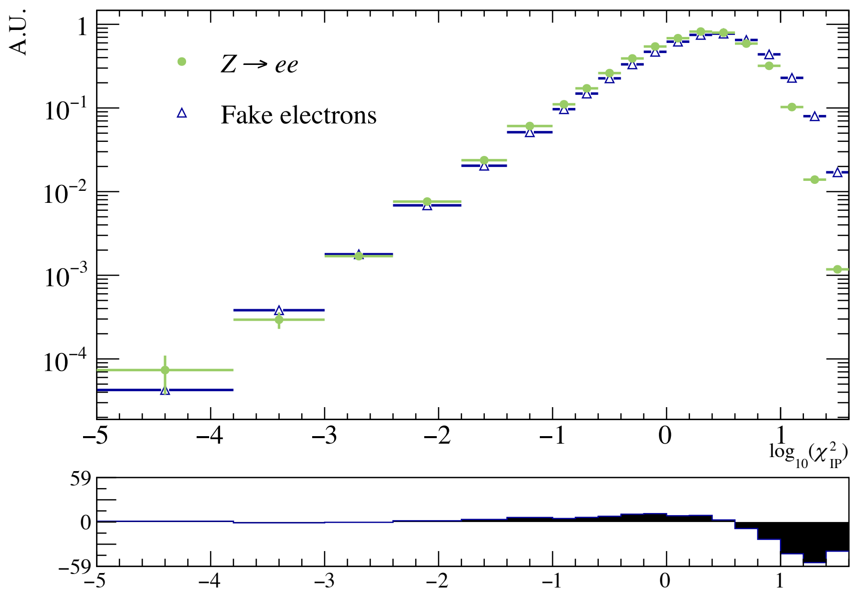


Figure 5.14: A comparison of the $\log_{10}(\chi^2_{\rm IP})$ distributions for $Z \to ee$ data and fake electrons. Pulls are shown underneath. All distributions are normalised to unit area and uncertainties are statistical only.

Hadrons that fake electrons also have a significant non-prompt component as can be seen from Fig. 5.14. This contribution is taken from data as described in Sect. 5.5.4. Minimum bias simulation is used to derive a constraint of $(21 \pm 7)\%$ on the fractional contribution of the fake electron component.

The heavy flavour template is taken from simulation of $c\bar{c}$ and $b\bar{b}$. MCFM NLO cross-sections corrected for the probability of producing a high-$p_{\rm T}$ electron in acceptance are used to enforce a relative normalisation for the processes and a single, combined template is employed in the fit.

HF fit results

The fit to the $\log_{10}(\chi^2_{\rm IP})$ distribution shown in Fig. 5.15 gives a HF component consistent with zero, with an upper limit of 0.8% at 68% confidence level based on a study of systematic uncertainties outlined in Sect. 5.9.1. The χ^2/ndf of the fit is 1.7.

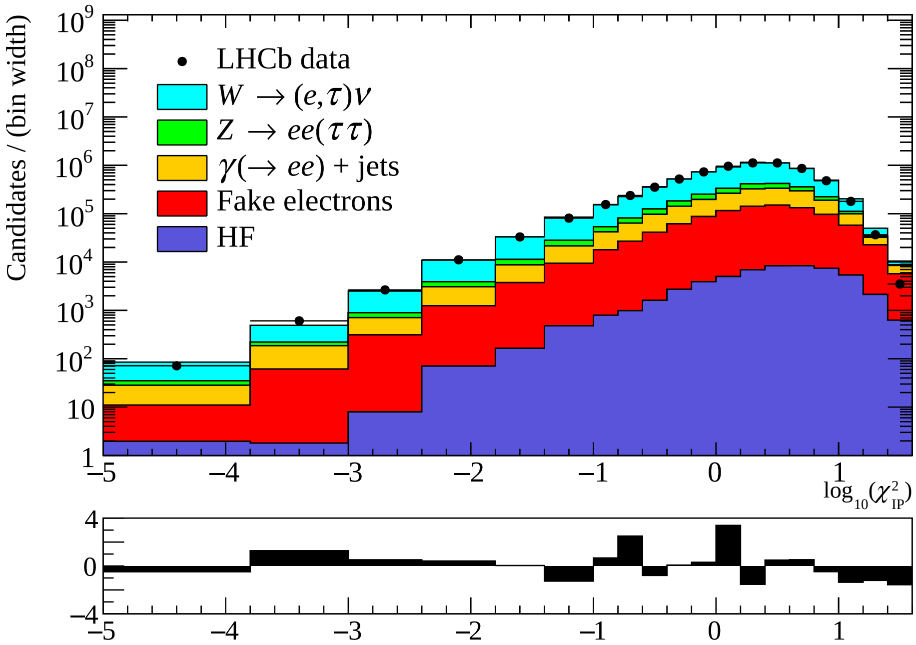


Figure 5.15: The nominal fit to the $\log_{10}(\chi^2_{\rm IP})$ distribution. The best-fit normalisation of the HF component is consistent with zero so a normalisation of 0.8% is enforced here to illustrate the shape of the template with respect to the other processes.

Track impact parameter correction

As detector misalignment is not modelled in simulation, a bin-by-bin correction is applied to the simulation-based inputs to the fit obtained by taking the ratio of Z data to simulation. The effect of the reweighting is to shift the distributions to higher values of $\chi^2_{\rm IP}$ as expected.

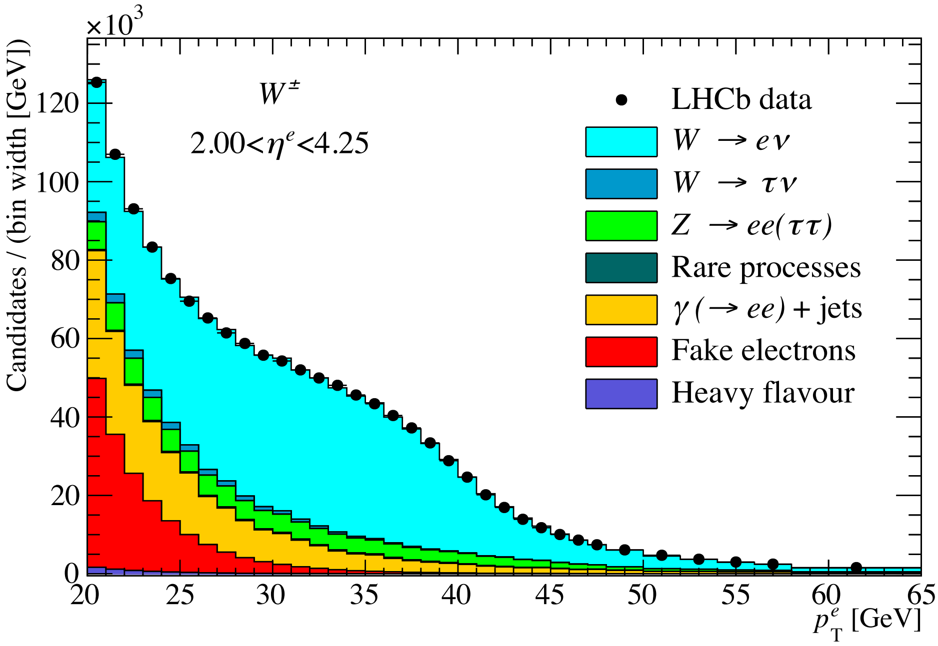


Figure 5.16: The inclusive fit to the p_{T}^{e} distribution of the full dataset. The χ^2/ndf of the fit is 1.1 with 33 degrees of freedom.

5.5.6 Fits to the lepton p_{T} spectrum

The signal yield is extracted from binned maximum likelihood template fits to p_{T}^{e} in eight bins of η^{e} and for the two charges separately. The $W \to (e, \tau)\nu$ and fake electron fractions float freely in the fit while the remaining components are constrained as explained in previous sections. The validity of the SM is implicit in the constraints based on theoretical cross-sections obtained from MCFM and in extracting template shapes from Pythia 8 simulation.

The p_{T}^{e} distribution of the full dataset with the result of the fit overlaid is shown for illustration in Fig. 5.16 and is used in the estimation of systematic uncertainties. The p_{T}^{e} spectra in the 16 bins of pseudorapidity and charge with the results of the fits superimposed are shown in Figs. 5.17 and 5.18. The pulls shown underneath each fit are statistical only. The $W \to \tau\nu$ component has been separated from $W \to (e, \tau)\nu$ and the low-mass DY component has been added to the $Z \to ee(\tau\tau)$ component for presentational purposes.

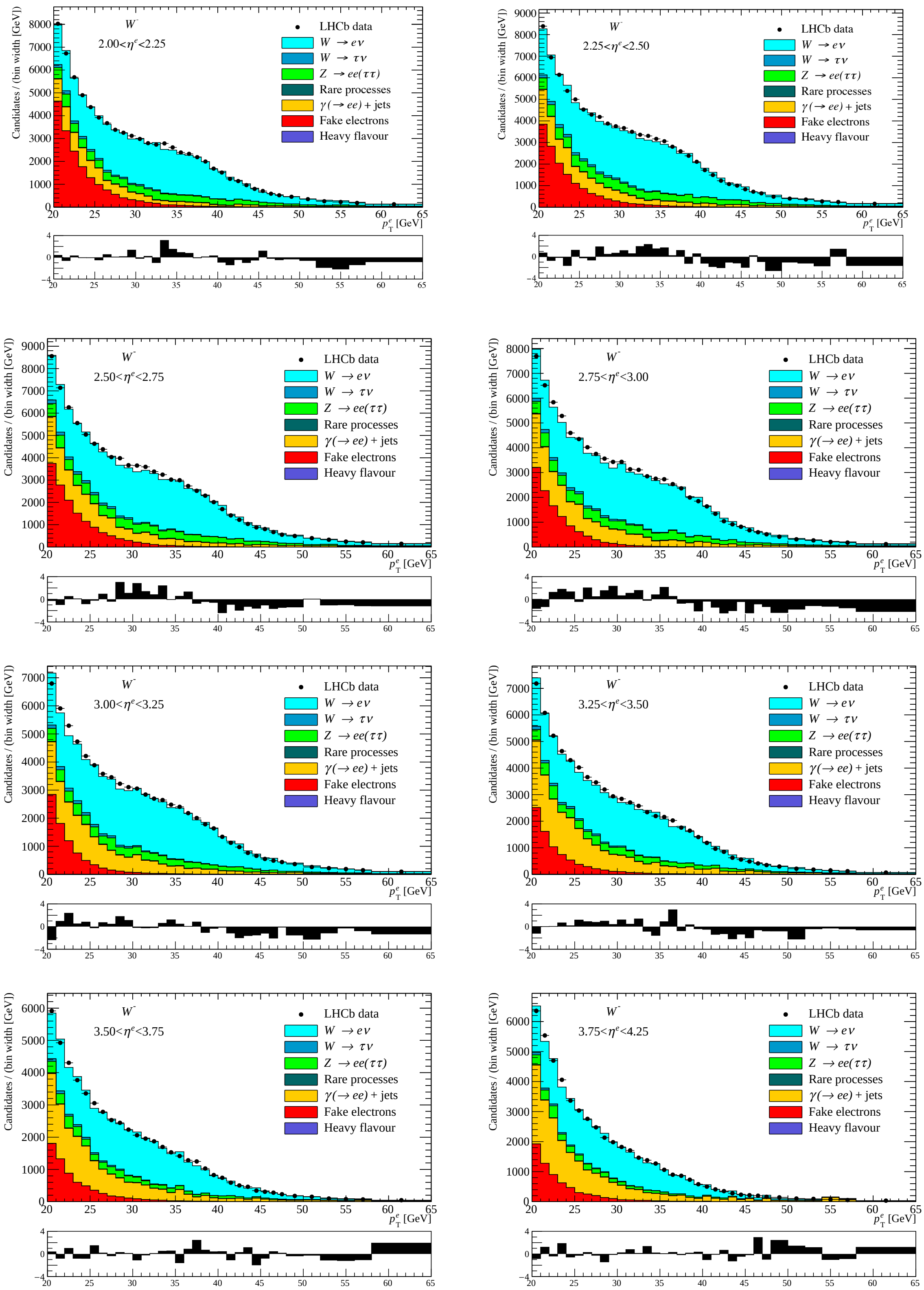


Figure 5.17: Fits to p_{T}^{e} for e^{-} in bins of η^{e}. Pulls are shown underneath.

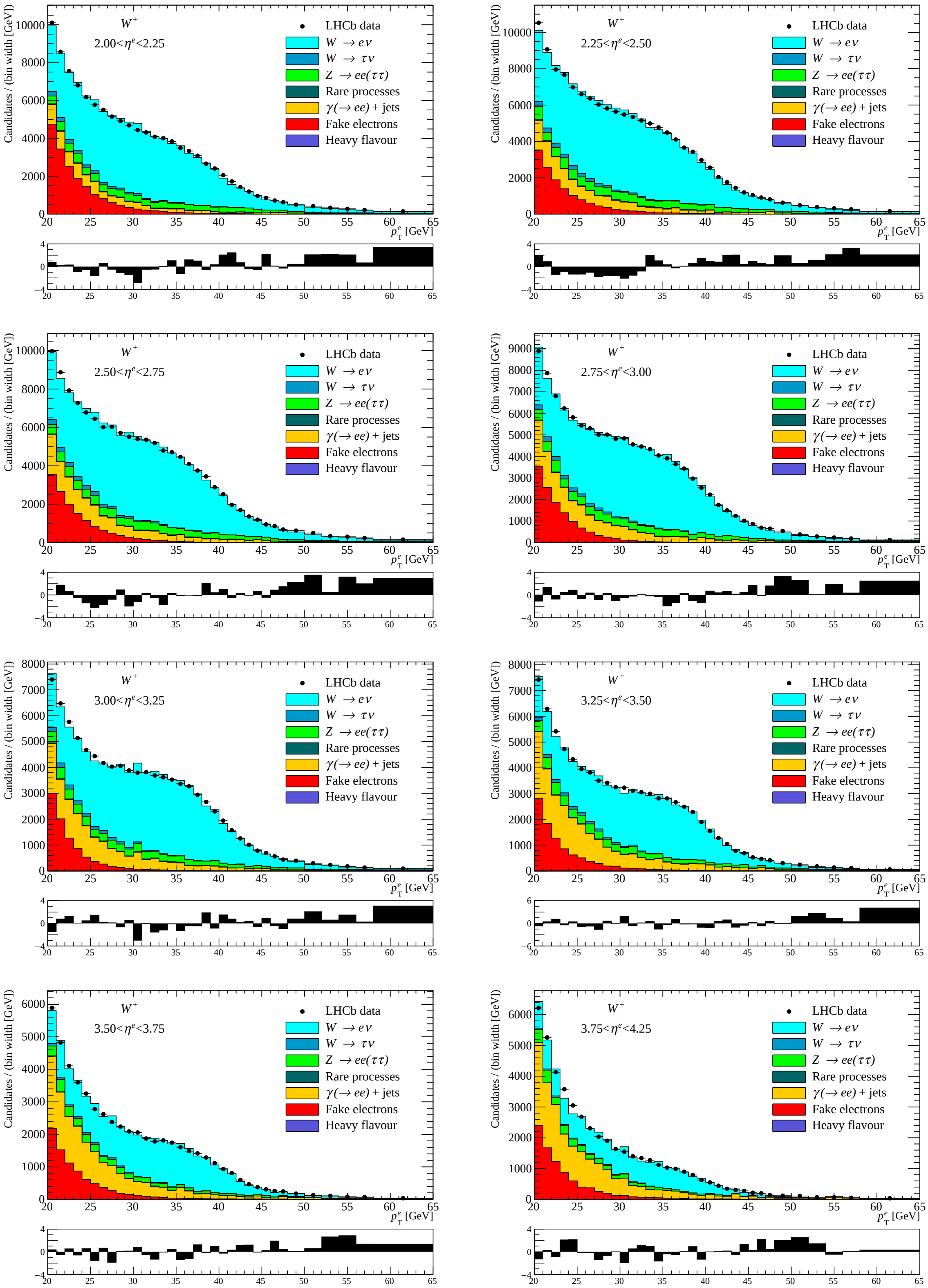


Figure 5.18: Fits to p_T^e for e^+ in bins of η^e. Pulls are shown underneath.

η^e	Fractional signal contribution [%]	
	e^+	e^-
2.00 - 2.25	67.5 ± 0.7	56.2 ± 1.0
2.25 - 2.50	70.4 ± 0.6	59.0 ± 0.8
2.50 - 2.75	68.6 ± 0.5	57.1 ± 0.8
2.75 - 3.00	66.8 ± 0.6	57.9 ± 0.8
3.00 - 3.25	66.2 ± 0.6	59.6 ± 0.8
3.25 - 3.50	56.5 ± 0.9	53.9 ± 1.2
3.50 - 3.75	49.8 ± 1.3	52.5 ± 1.5
3.75 - 4.25	37.1 ± 2.4	47.8 ± 1.9

Table 5.7: The fractional signal contribution in bins of lepton pseudorapidity. The uncertainties are statistical only.

η^e	χ^2/ndf	
	e^+	e^-
2.00 - 2.25	1.9	1.0
2.25 - 2.50	2.1	1.9
2.50 - 2.75	2.2	1.8
2.75 - 3.00	1.5	2.3
3.00 - 3.25	1.5	1.5
3.25 - 3.50	1.5	1.3
3.50 - 3.75	1.3	0.9
3.75 - 4.25	1.4	1.1

Table 5.8: The χ^2/ndf of the fits in bins of lepton pseudorapidity.

The fractional contributions of $W \to e\nu$ to the SR in bins of η^e are listed in Table 5.7. The corresponding fit qualities are given in Table 5.8. The fractions of each of the processes that contribute to the signal sample are shown as a function of η^e in Fig. 5.19. $W \to (e, \tau)\nu$ and fake electrons are shown alone in Fig. 5.20.

The fractional signal contribution in the W^+ (W^-) sample varies from ~70%(~60%) near $\eta^e = 2$ to ~40%(~50%) at the largest pseudorapidity. The values of χ^2/ndf for the fits range between 0.9 and 2.3, based on statistical uncertainties only. The systematic uncertainties in the event yields presented in Section 5.9 are found to cover the uncertainty that arises from imperfect fit quality.

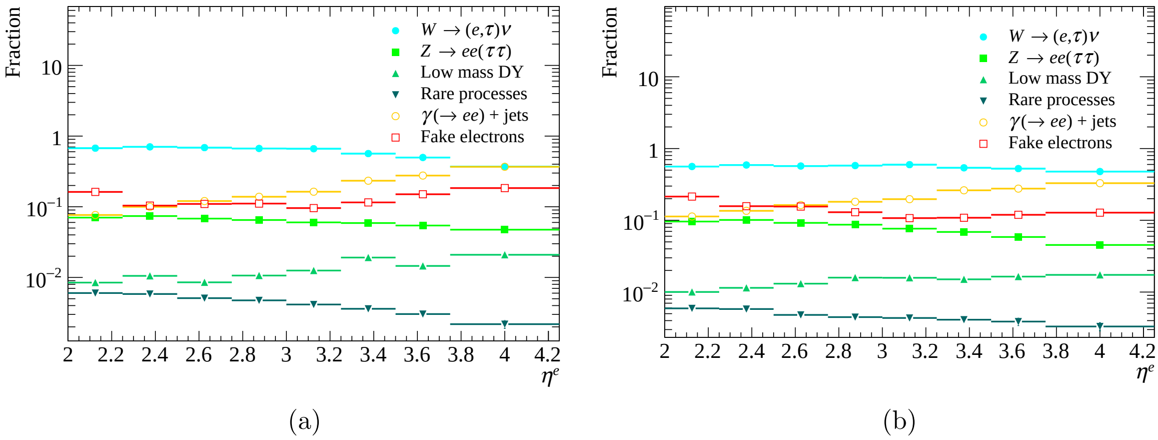


Figure 5.19: The fractions of the processes as a function of η^e for (a) e^+ and (b) e^-.

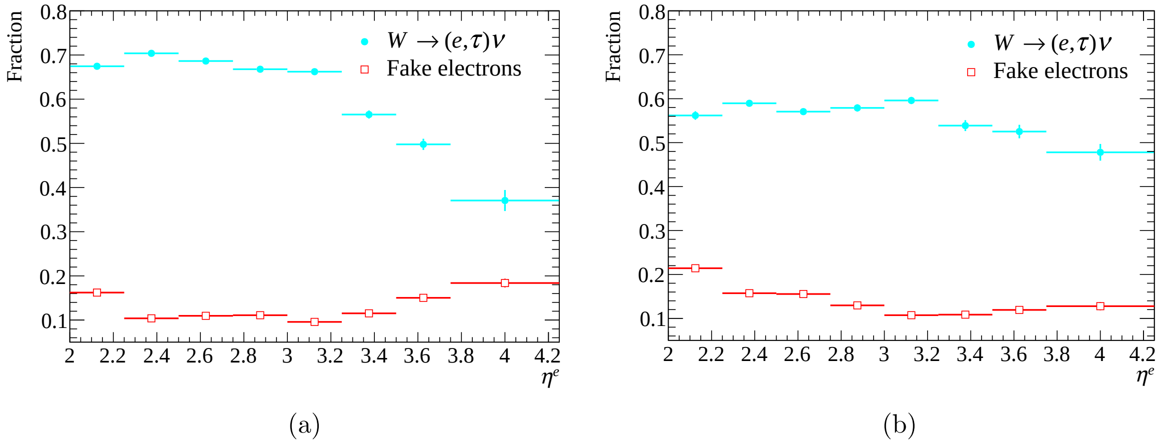
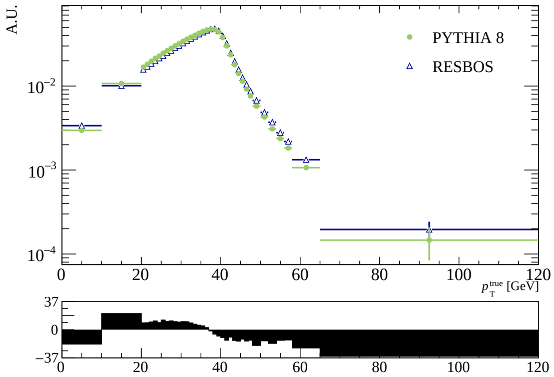
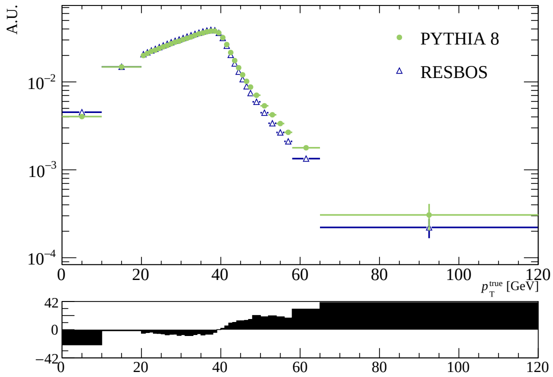


Figure 5.20: The fractions of the two components, $W \to (e, \tau)\nu$ and fake electrons, that are unconstrained in the fit as a function of η^e for (a) e^+ and (b) e^-. The uncertainties are obtained from repeated pseudoexperiments as explained in Sect. 5.9.2 and are statistical in nature.



(a)



(b)

Figure 5.21: PYTHIA 8 simulation of true transverse momentum, $p_{\mathrm{T}}^{\mathrm{true}}$, is compared to RESBOS simulation, which incorporates NLO MEs. The comparison is performed inclusively in η^e for (a) e^+ and (b) e^-.

5.5.7 Corrections to the fits to p_{T}^{e}

Correction for higher orders in perturbative QCD

The standard simulation employed in this analysis has LO MEs for EW boson production with the effect of higher orders modelled using the parton shower approach. This can underestimate the multiplicity of jets as well as their p_{T} due to the fact that it is only accurate to leading logarithms at all orders. Since the boson will recoil against jet(s) in the event, then its p_{T} is also affected. As electron p_{T} is determined from the p_{T} of the parent and any additional momentum imparted to it by the mass of the boson, then biases can result as can be seen in Fig. 5.21. Here, the true transverse momentum distribution of electrons in PYTHIA 8 simulation is compared to that of RESBOS. Since detector-level simulation of RESBOS is absent, PYTHIA 8 simulation must be corrected.

Jet activity in the event can be considered as a proxy for missing higher orders and the following procedure is adopted to correct $W \to e\nu$ and $Z \to ee$ simulation. Anti-k_{T} jets with a radius parameter of $R = 0.5$ and $p_{\mathrm{T}}^{\mathrm{jet}} > 10\,\mathrm{GeV}$ are reconstructed in LHCb acceptance using the standard particle flow algorithm [25, 26, 147]. Jets that lie within one unit from the candidate electron in η-ϕ space are vetoed to avoid storing jets that are formed around the electrons.

The candidates are then separated into two samples: events with no jets in acceptance (the "0-jet sample") and events with one or more jets in acceptance (the "n-jet sample").

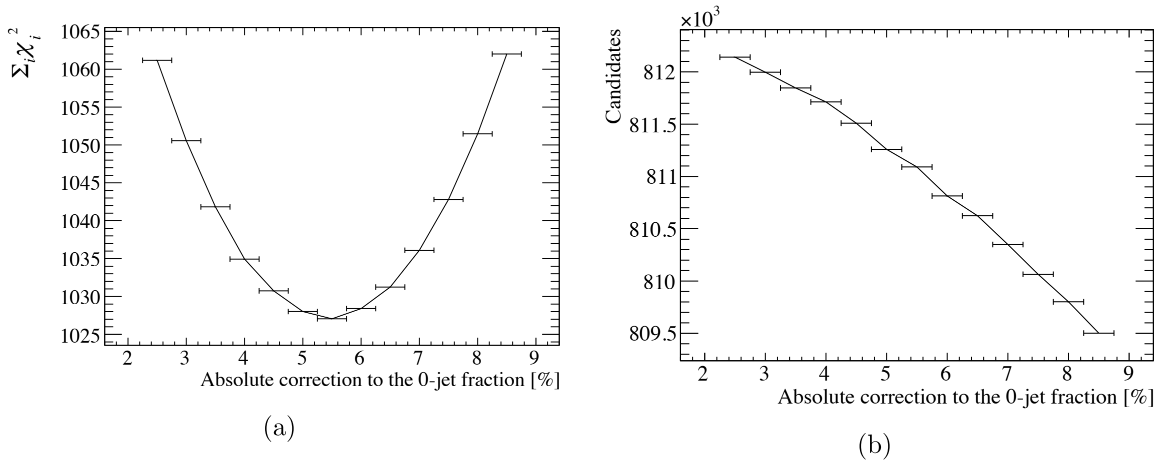


Figure 5.22: (a) The summed χ^2 of the fits as a function of the fractional correction to the nominal 0-jet fraction. The number of degrees of freedom is 528 in total. (b) demonstrates the effect of the jet fraction correction on the W boson event yield.

Nominally, the "n-jet" sample is 23.5% of the inclusive ("0-jet"+"n-jet") sample. The fits to $p_{\mathrm{T}}^{\mathrm{e}}$ are carried out for a range of different mixtures of these samples. It is determined that an increase in jet activity of $(5.5 \pm 1.0)\%$ from the nominal fraction to 29% minimises the summed χ^2 of the fits as shown in Fig. 5.22 (a). The correction has a weak influence on the measured W boson yield as can be seen in Fig. 5.22 (b). The shift of 5.5% is applied in all bins of η^e when performing the differential measurement. The procedure to determine the systematic uncertainty is explained in Sect. 5.9.2.

As a cross-check to the above procedure, a fit to Z data is carried out with 0-jet and n-jet templates floating freely. The $p_{\mathrm{T}}^{\mathrm{e}}$ distribution with the result of the fit overlaid is shown in Fig. 5.23. A template for fake electrons is obtained from the same-sign sample and is constrained in the fit. The fit favours a correction to the jet fraction of 4.7% and is, therefore, consistent with the one determined using W data.

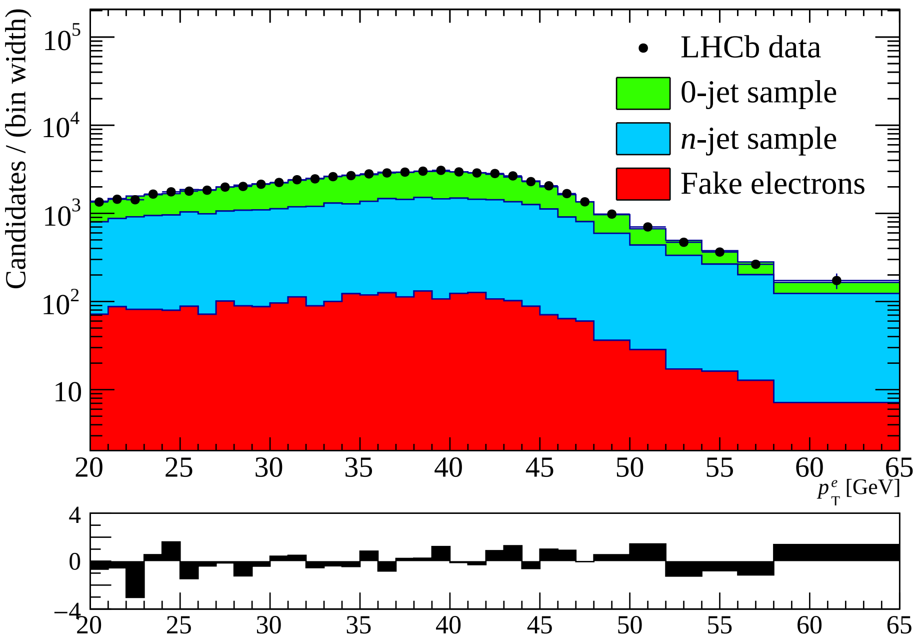
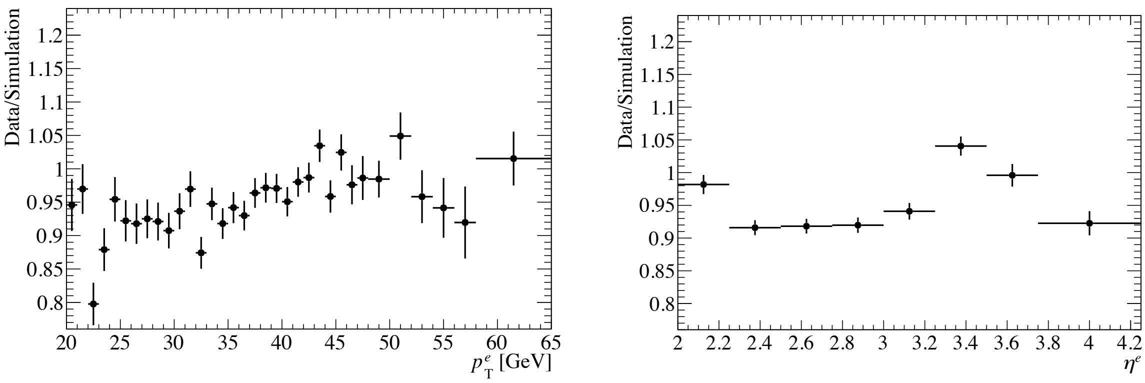


Figure 5.23: The fit to the p_{T}^{e} distribution of Z daughters. Pulls are shown underneath. The quality of the fit is excellent with $\chi^2/\mathrm{ndf} = 0.99$.



(a) (b)

Figure 5.24: The efficiency seen in Z data divided by the efficiency in simulation as a function of (a) p_{T}^{e} and (b) η^{e}.

Correction for the modelling of efficiencies

Shortcomings present in the modelling of trigger and tight selection efficiencies may result in biases in the p_{T}^{e} spectra of simulation-based inputs. $W \to (e, \tau)\nu$ templates are sensitive only to biases in the shape of the p_{T}^{e} spectrum as they are free to vary in the fits. Other simulation-based EW templates are sensitive to both the mismodelling of shape as well as normalisation since they are constrained. Fig. 5.24 (a) demonstrates that the ratio is consistent with unity in p_{T}^{e}. However, Fig. 5.24 (b) makes it clear that the normalisation of the efficiency is incorrect in simulation. Therefore, a correction is applied to the simulation-based templates, Z/W, low-mass DY, and rare processes, as a function of pseudorapidity.

Special consideration must be given to γ+jets simulation as there is a partial cancellation of efficiency mismodelling in the ratio in Eq. 5.3. However, requirements that are applied to one of the templates, but not the other, might bias the template. Specifically, both daughters in the $[\gamma + jets]_{\mathrm{sim}}^{\gamma\to ee\ \mathrm{SR}}$ sample are required to pass the N_{TT} requirement and be reconstructed and so will be affected by the mismodelling of the efficiency for these requirements. The normalisations of the $[\gamma + jets]_{\mathrm{sim}}^{\gamma\to ee\ \mathrm{SR}}$ templates are corrected accordingly.

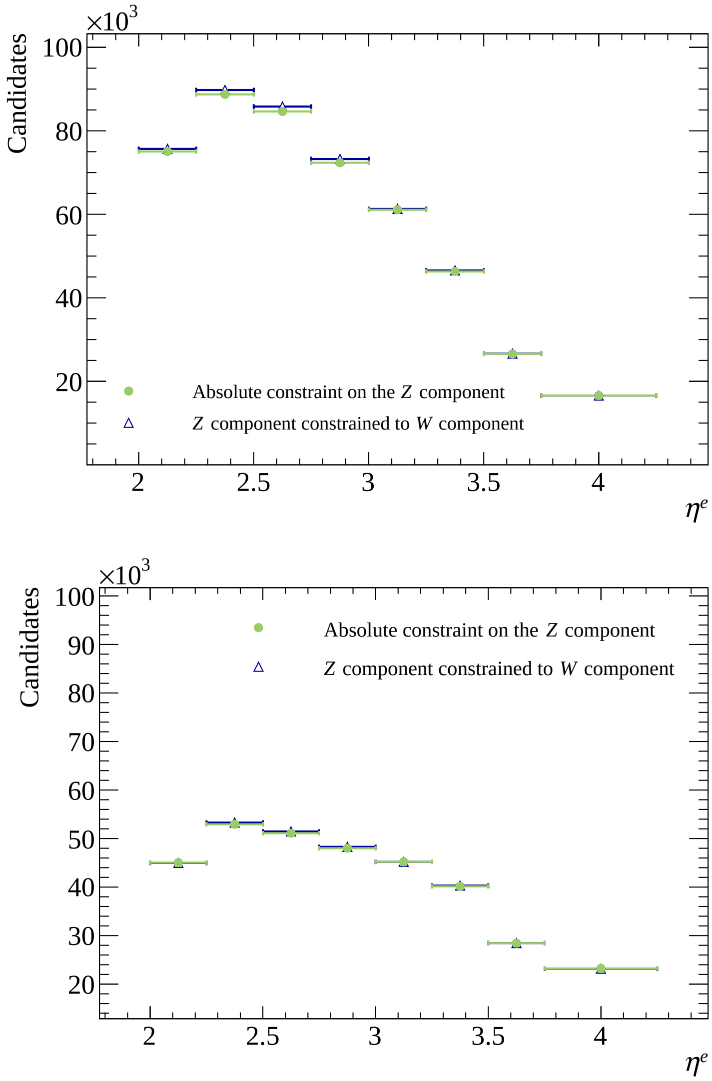


Figure 5.25: Top: The signal yields for e^+ with Z constrained absolutely or with a relative normalisation to the W component. Bottom: The same comparison for e^-.

5.5.8 Fitting W and Z components together

Here, a cross-check for the procedure used to constrain the Z component in the fits to p_{T}^{e} is considered. The Z component is alternatively constrained to the W component i.e. a relative, rather than an absolute, constraint is enforced. The combined template is allowed to float freely in the fit. The change in signal yields is seen to be negligible as shown in Fig. 5.25. This can be considered to be a test of the robustness of the fitting technique.

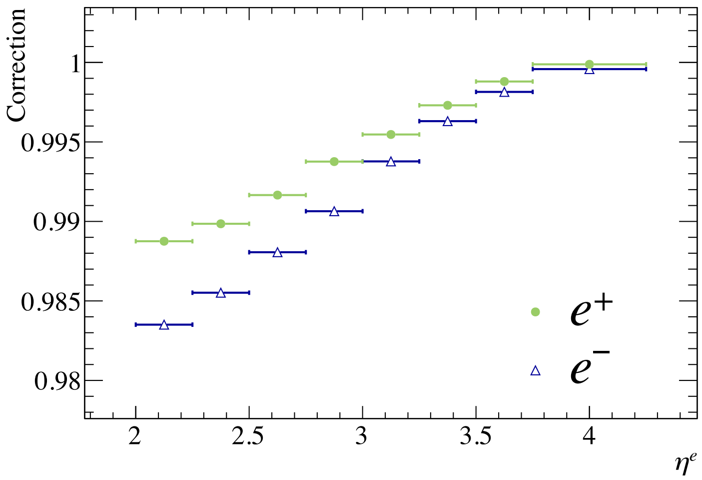


Figure 5.26: The acceptance factors, $\mathcal{A}(\eta^e)$, obtained from ResBos for excluded candidates with $p_{\mathrm{T}}^{\mathrm{true}} > 74\,\mathrm{GeV}$. The uncertainties (shadowed by the datapoints) are statistical only.

5.6 Correction for the range of the fit to p_{T}^{e}

The fits to p_{T}^{e} are carried out in the range $20 < p_{\mathrm{T}}^{e} < 65\,\mathrm{GeV}$. The candidate yields are corrected for excluded candidates with $p_{\mathrm{T}}^{e} > 65\,\mathrm{GeV}$ by computing a charge-dependent acceptance factor, $\mathcal{A}(\eta^e)$, for all bins of η^e using a ResBos [49–51] simulation. Since ResBos simulation is not available at detector-level, it cannot be used to evaluate the contribution of reconstructed tracks with $p_{\mathrm{T}}^{e} > 65\,\mathrm{GeV}$ in the SR. Consequently, Pythia 8 simulation is used to determine the equivalent, truth-level and pre-FSR transverse momentum requirement in the fiducial volume that reproduces the acceptance factors obtained at detector-level. The truth-level requirement is determined to be $p_{\mathrm{T}}^{\mathrm{true}} < 74.25\,(73.75)\,\mathrm{GeV}$ for e^- (e^+). A value of $p_{\mathrm{T}}^{\mathrm{true}} < (74.00 \pm 0.25)\,\mathrm{GeV}$ is therefore adopted. The acceptance factors are shown in Fig. 5.26. The method to set a systematic uncertainty for the procedure is discussed in Sect. 5.9.5.

5.7 Efficiencies

The total efficiency used to correct the candidate yield can be written as the product

$$\epsilon^{\text{tot}} \equiv \epsilon^{\text{track}} \cdot \epsilon^{\text{kin}} \cdot \epsilon^{\text{PID}} \cdot \epsilon^{\text{GEC}} \cdot \epsilon^{\text{trigger}} \cdot \epsilon^{\text{tight}}. \tag{5.4}$$

Each subsequent efficiency is determined in a subset of events defined by the preceding requirements in order to ensure that correlations between the requirements are correctly accounted for.

Tracking, kinematic, and PID efficiencies are extracted from $W \to e\nu$ simulation with data-driven cross-checks in all cases. The efficiency for the di-electron software trigger veto is also determined from simulation. The GEC efficiency is measured from $Z \to \mu\mu$ data. Single-electron trigger and the tight selection efficiencies are determined from $Z \to ee$ data with a simulation-based correction for the difference in p_{T}^{e} distributions for Z and W boson daughters.

A tag-and-probe method [20] is employed for efficiencies determined from $Z \to ee$ data. This is a valid procedure as the $Z \to ee$ and $W \to e\nu$ datasets correspond to the same luminosity, have identical trigger lines and reconstruction, and an identical pre-selection apart from a di-electron invariant mass requirement of 40 GeV for the former dataset.

The components that contribute to the total efficiency are outlined in more detail in Sects. 5.7.1 to 5.7.6.

5.7.1 Track reconstruction efficiency

The track reconstruction efficiency, ϵ^{track}, accounts for candidates that fail to be reconstructed as long tracks satisfying the track quality criteria that form a part of the pre-selection. The efficiency is extracted from Pythia 8 simulation of $W \to e\nu$.

This efficiency is cross-checked using a study that looks for high-p_{T} photons accompanied by an e^- or e^+ in data and $Z \to ee$ simulation. For true $Z \to ee$ events, these cases represent instances where one daughter has failed to be reconstructed in the tracking system and offer a basis for comparison to the simulation. The ratio of the efficiency for such events in data divided by the efficiency in simulation is found to be (0.995 ± 0.005). The correction is applied to the simulation-based efficiency and its uncertainty is propagated to the measurement. The method is described in more detail in Refs. [19, 20].

5.7.2 Kinematic efficiency

An electron with $p_{\mathrm{T}}^{\mathrm{true}}$ of more than 20 GeV can be reconstructed as having $p_{\mathrm{T}}^{e} < 20\,\mathrm{GeV}$. This is predominantly due to bremsstrahlung. For high-p_{T} candidates, the photons tend to lie close to the electron and are often not correctly identified by bremsstrahlung recovery. The correction for this effect, ϵ^{kin}, is determined using simulation and is cross-checked in data by scaling the p_{T}^{e} distribution of simulation-based templates and computing the χ^2 for a range of scales to probe the uncertainty associated with the modelling of bremsstrahlung. The method is discussed in more detail in Sect. 5.9.2.

5.7.3 PID efficiency

Simulation of $W \to e\nu$ is used to extract an efficiency, ϵ^{PID}, for the PID requirements that form a part of the pre-selection. While the acceptance requirement is dictated by the geometry of the HCAL and should be reliably modelled, the other efficiencies may be subject to biases when extracted from simulation. A data-driven cross-check following the procedure used for Refs. [19, 20] is adopted to assign a systematic uncertainty of 0.6% on the efficiency. Here, agreement (or lack thereof) between data and simulation is investigated near the cut region and observed discrepancies determine the level of uncertainty on the simulation-based efficiency.

The overall efficiency taken from simulation, $\epsilon_{\mathrm{sim}} \equiv \epsilon^{\mathrm{track}} \cdot \epsilon^{\mathrm{kin}} \cdot \epsilon^{\mathrm{PID}}$, is shown in Fig. 5.27. The efficiency displays a significant p_{T}^{e} dependence due to bremsstrahlung, which has the effect of shifting particles below the transverse momentum threshold of $p_{\mathrm{T}}^{e} = 20\,\mathrm{GeV}$. By contrast, the η^{e} dependence is largely determined by edge effects of the geometric acceptance of the calorimetry due to its rectangular shape.

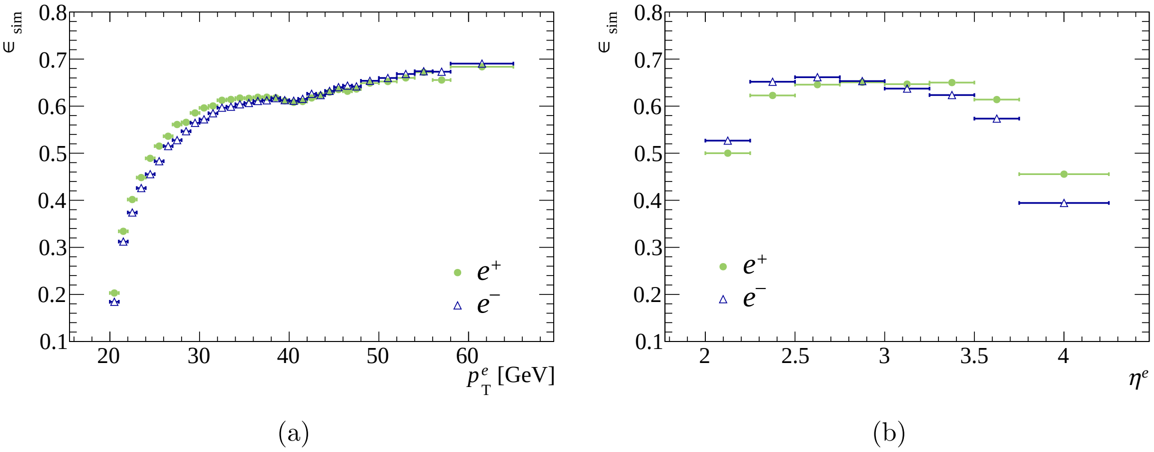
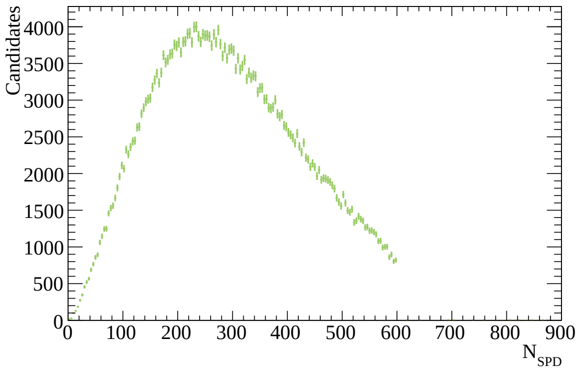
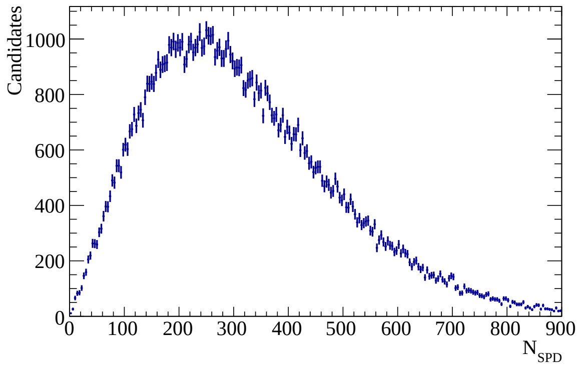


Figure 5.27: The overall efficiency that is extracted from simulation as a function of (a) p_T^e and (b) η^e.



(a)



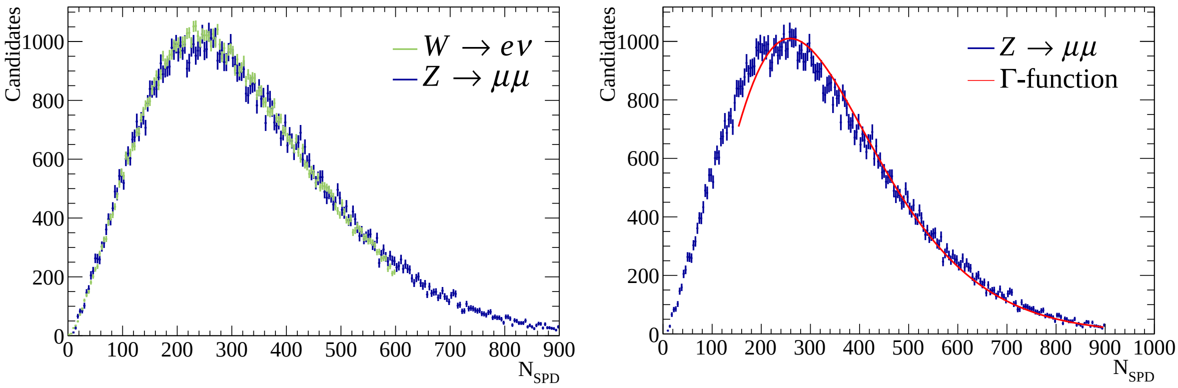
(b)

Figure 5.28: The SPD multiplicity distribution for (a) $W \to e\nu$ candidates and (b) $Z \to \mu\mu$ candidates.

5.7.4 Global event cut efficiency

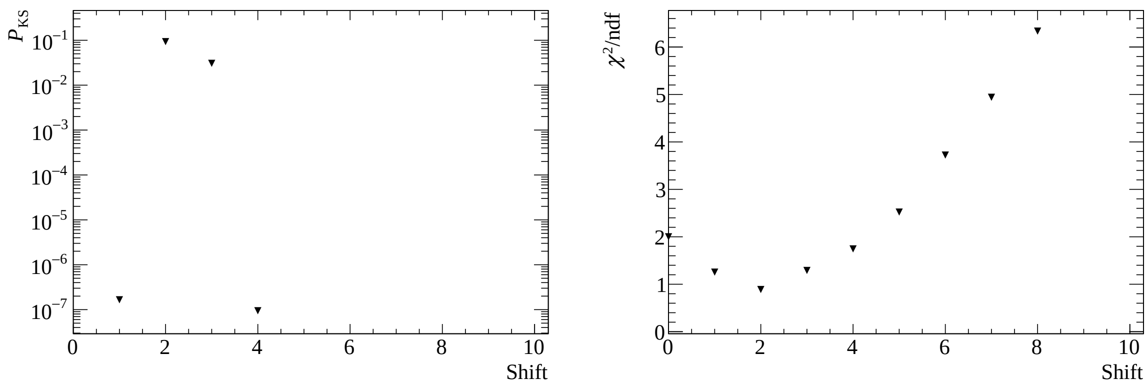
The hardware trigger incorporates a global event cut (GEC) on the number of SPD hits, $N_{\mathrm{SPD}} < 600$, to prevent high-multiplicity events from dominating the processing time at trigger-level as shown in Fig. 5.28 (a). As the SPD multiplicity distribution is not well modelled in simulation, a data-driven method for determining the GEC efficiency is necessary. Fortunately, the di-muon dataset has a less stringent $N_{\mathrm{SPD}} < 900$ requirement applied to it at Level-0 level as seen in Fig. 5.28 (b) and can be employed in determining the fraction of events beyond $N_{\mathrm{SPD}} = 600$.

However, di-muon events are not entirely comparable to $W \to e\nu$ as electrons will shower in the detector and lead to more hits in the SPD. Nevertheless, Fig. 5.29 (a) demonstrates that after a suitable shift of the di-muon distribution, good agreement is observed with candidates obtained with the $W \to e\nu$ selection. The optimal shift is determined by performing a test of compatibility for every shift of the di-muon SPD multiplicity distribution. Two methods, a Kolmogorov-Smirnov test [148, 149] and a χ^2 test, are used for this and demonstrated in Fig. 5.30. The remaining fraction of events beyond $N_{\mathrm{SPD}} = 900$ ($< 1\%$ of the candidates) is then extracted by fitting the tail of the di-muon distribution with a Γ-function as shown in Fig. 5.29 (b).



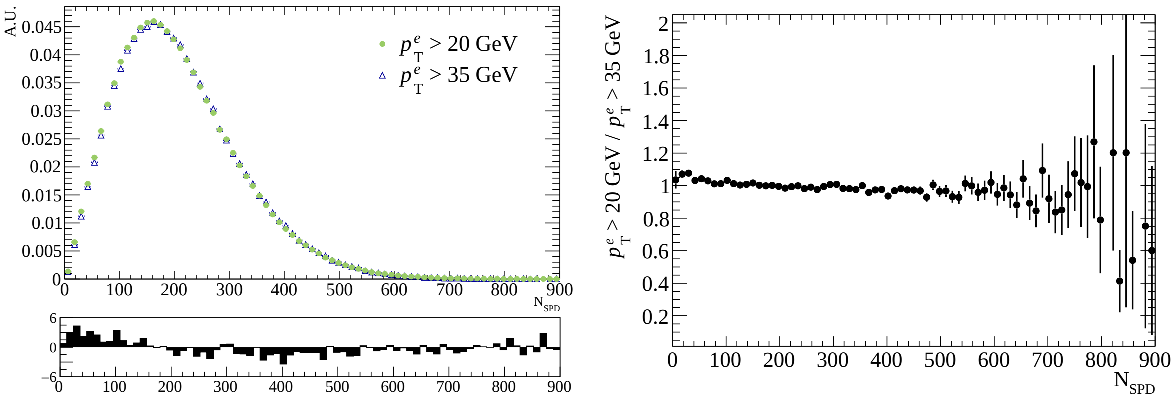
(a) (b)

Figure 5.29: (a) The di-muon SPD multiplicity distribution shifted to agree with that of $W \to e\nu$ candidates. (b) The Γ-function used to extract the fraction of events beyond 900 hits.



(a) (b)

Figure 5.30: (a) The Kolmogorov-Smirnov probability, P_{KS}, for compatibility as a function of the shift applied to the di-muon distribution. Similarly, (b) shows the χ^2/ndf for every shift. Here, all values of N_{PV} are considered inclusively.



(a) (b)

Figure 5.31: Comparison of the SPD hits distribution for (a) $p_{\mathrm{T}}^{e} > 20$ GeV and $p_{\mathrm{T}}^{e} > 35$ GeV with pulls are shown underneath. (b) The ratio of the SPD hits distributions.

A possible concern with the method detailed above is that the $W \to e\nu$ candidate sample is contaminated by backgrounds. This is mitigated with a stringent requirement of $p_{\mathrm{T}}^{e} > 35$ GeV applied to the W sample as this removes the majority of the fake electron background. This does not significantly bias the SPD hits distribution as can be seen from Fig. 5.31. However, the isolation requirement used in the selection preferentially selects low multiplicity events and so must be removed when determining the efficiency.

Nevertheless, a slight bias is introduced as a result of the other tight selection requirements and the increased p_{T}^{e} threshold. The bias is primarily due to the stringent HCAL requirement. A correction of $(0.53 \pm 0.04)\%$, determined from simulation, is subtracted from the estimate. A SPD requirement that reproduces the efficiency seen in data is used when determining this correction.

Further to this, residual backgrounds are removed by explicitly subtracting the $Z \to ee$, γ+jets, and fake electron components. All of these are extracted from data with an acceptance and efficiency correction taken from simulation for the $Z \to ee$ and $\gamma(\to ee)$+jets components. A dedicated fit to p_{T}^{e} is carried out to quantify the normalisation of the fake electron component within this dataset and it is determined to be (0.073 ± 0.004).

The overall GEC efficiency is determined to be (0.9133 ± 0.0031). The procedure to determine the uncertainty is explained in Sect. 5.9.3. The requirement on the number of SPD hits that reproduces the above efficiencies in simulation is determined to be (380 ± 3).

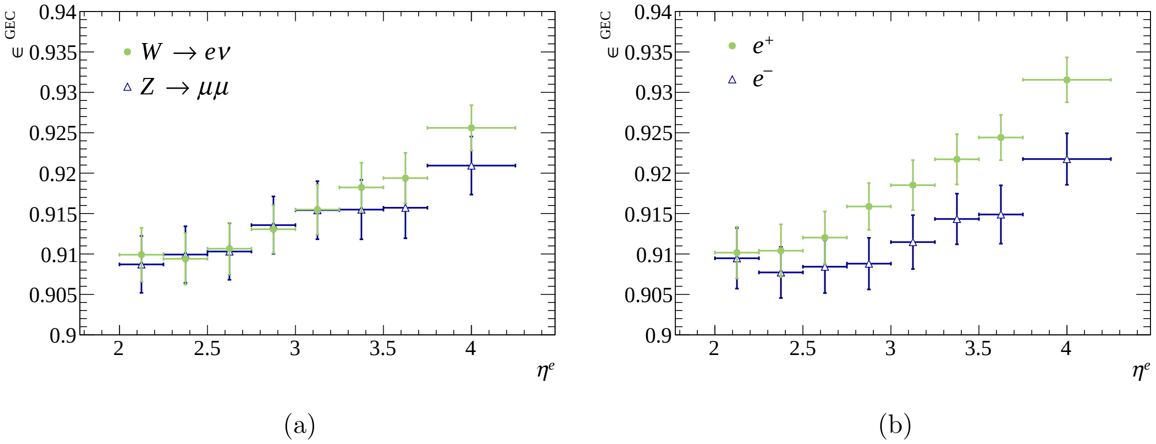


Figure 5.32: (a) The η^e-dependence of ϵ^{GEC} in simulation and in $Z \to \mu\mu$ data where both have been corrected to the efficiency determined in $W \to e\nu$ data. (b) demonstrates the charge dependence of the efficiency in $W \to e\nu$ simulation.

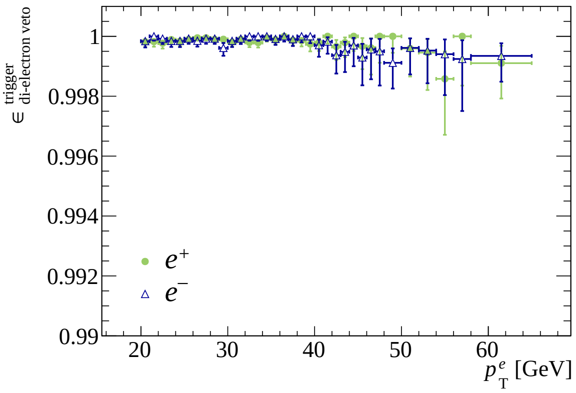
This, more stringent, requirement is applied to the simulation samples throughout the analysis to reproduce the effect of the requirement in data.

After correcting $W \to e\nu$ simulation and $Z \to \mu\mu$ data to reproduce the efficiency determined on $W \to e\nu$ data, simulation is seen to model the η^e-dependence of the efficiency well as shown in Fig. 5.32 (a) and so is employed in the analysis. Furthermore, the efficiency is found to also depend on W daughter charge as is demonstrated in Fig. 5.32 (b). However, the charge-independent simulation-based efficiency is used in the analysis and an additional uncertainty due to the observed differences in charge is applied to the efficiency.

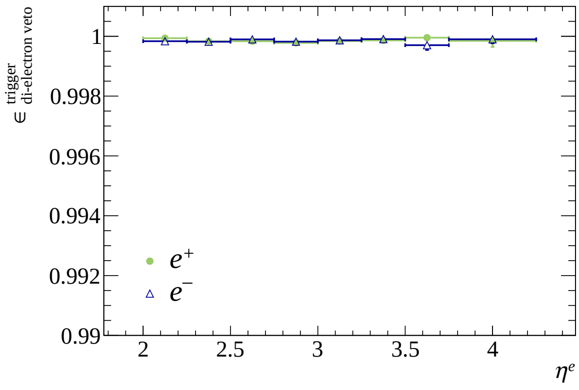
5.7.5 Trigger efficiency

The trigger efficiency, $\epsilon^{\text{trigger}}$, consists of two contributions: (a) the efficiency for satisfying single-electron triggers across Level-0, HLT1, HLT2, $\epsilon^{\text{trigger}}_{\text{single } e}$; (b) the efficiency for the di-electron software trigger veto, $\epsilon^{\text{trigger}}_{\text{di-electron veto}}$.

A tag-and-probe method is employed using $Z \to ee$ data to determine the efficiency for the single-electron triggers. Here, the tag is a Z daughter that satisfies the pre-selection and the single-electron triggers. The probe is then used to determine the fraction of candidates that satisfy the trigger requirements.



(a)



(b)

Figure 5.33: The efficiency for the di-electron software trigger veto as a function of (a) p_{T}^{e} and (b) η^e.

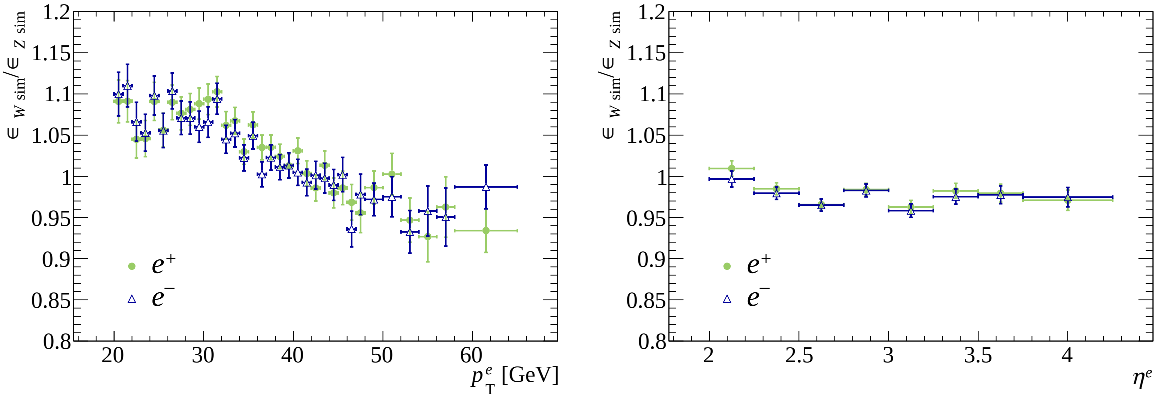
Efficiency determinations using data are liable to be biased due to the presence of background processes and this needs to be accounted for. The hadronic background is estimated from a dataset with same-sign, $e^{\pm}e^{\pm}$, events with the same selection applied to them. This is sensible as hadronic events are expected to contribute approximately equally to same-sign and opposite sign events. Consequently, this contribution is subtracted from the dataset. An uncertainty of 5.5% is assigned for the procedure as described in Section 5.9.3. Contamination from rare EW backgrounds such as di-boson and $t\bar{t}$ was found be to be small and is not corrected for.

The efficiency for the di-electron software trigger veto cannot be determined from Z data. W simulation is used in this case and the efficiency is found to be extremely close to unity as seen in Fig. 5.33.

5.7.6 Tight selection efficiency

Tight selection requirements consist of more stringent track quality requirements and PID requirements, as well as ensuring the track is prompt and isolated. The requirements are discussed in more detail in Sect. 5.2.2. The efficiency for these requirements, $\epsilon^{\mathrm{tight}}$, is determined using Z data analogously to the procedure for determining the trigger efficiency.

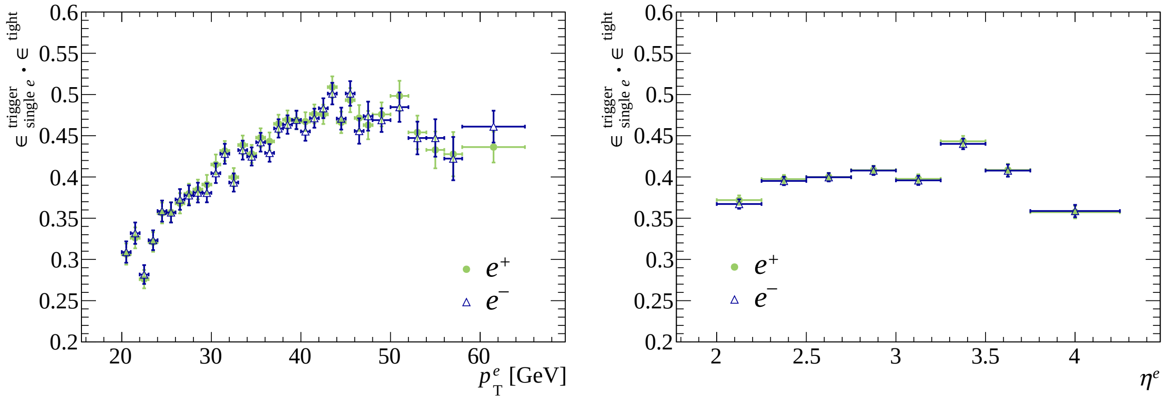
Efficiencies determined from $Z \to ee$ cannot be directly used for W production due to the different couplings at the production and decay vertices, a different mixture of



(a) (b)

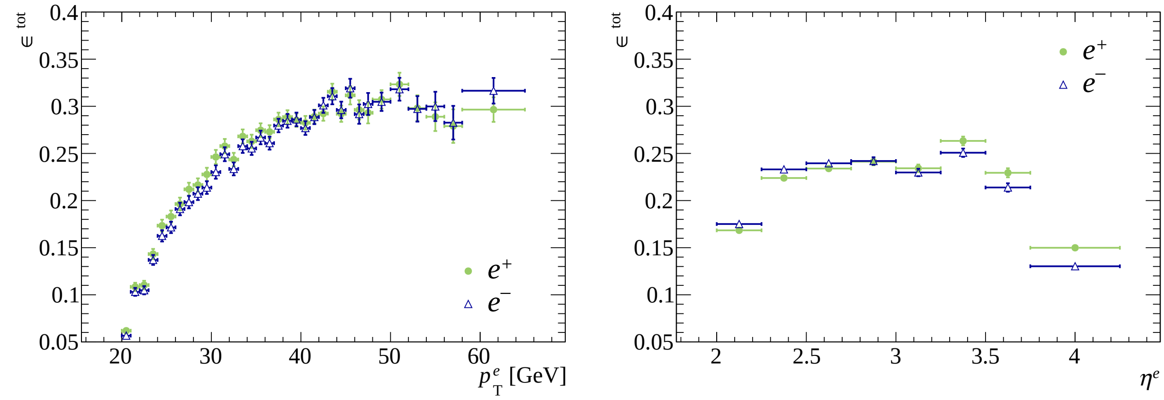
Figure 5.34: Tight selection efficiency corrections obtained from simulation as a function of (a) p_{T}^{e} and (b) η^{e}.

interacting quarks, and, most importantly, the difference in mass. This results in a p_{T}^{e} spectrum that is harder for electrons from the Z boson. Consequently, efficiencies that show a dependence on p_{T}^{e} are liable to be biased. This is corrected for in each bin of η^{e} using W and Z simulation with multiplicative corrections shown in Fig. 5.34. The efficiency for the tight selection requirements is shown in Fig. 5.35 and the total efficiency is shown in Fig. 5.36.



(a) (b)

Figure 5.35: The efficiency for requirements obtained from Z data, $\epsilon^{\text{trigger}}_{\text{single } e} \cdot \epsilon^{\text{tight}}$, as a function of (a) p^e_{T} and (b) η^e.



(a) (b)

Figure 5.36: The total efficiency as a function of (a) p^e_{T} and (b) η^e. The total efficiency is a product of simulation and Z data based efficiencies.

5.8 Correction for charge mis-identification

Track reconstruction for electrons is complicated by the fact that they are continually radiating photons as they traverse the detector material. The photons may then convert to electrons and so the candidate is frequently surrounded by a shower of secondary particles. This can mislead the pattern recognition that tracking algorithms rely on. Consequently, the particle could be assigned the wrong charge and, in the case of simulation, fail to be matched to a true particle.

This will have the effect of diluting the charge asymmetry and also biasing the cross-section ratio measurement. If $f(\eta^e)$ is the fraction of unmatched candidates in simulation, of which $g(\eta^e)$ has been assigned the wrong charge, then the measured charge asymmetry can be written as

$$\begin{aligned}\mathcal{A}_e^{\text{meas}} &= \frac{(1-f)\sigma_{W^+} + f(1-g)\sigma_{W^+} + fg\sigma_{W^-} - \left[(1-f)\sigma_{W^-} + f(1-g)\sigma_{W^-} + fg\sigma_{W^+}\right]}{\sigma_{W^+} + \sigma_{W^-}} \\ &= (1-2fg)\mathcal{A}_e^{\text{true}},\end{aligned} \tag{5.5}$$

where $\sigma_{W^+}(\sigma_{W^-})$ are the W^+ (W^-) production cross-sections and $\mathcal{A}_e^{\text{true}}$ is the true charge asymmetry. All quantities are binned in η^e in this measurement.

The charge mis-identification correction, $1/(1-2fg)$, is shown in Fig. 5.37 and differs from unity by $(0.58 \pm 0.05)\%$ overall. The cross-section ratio and the charge asymmetry are corrected accordingly. The uncertainty in the correction is propagated to the measurement.

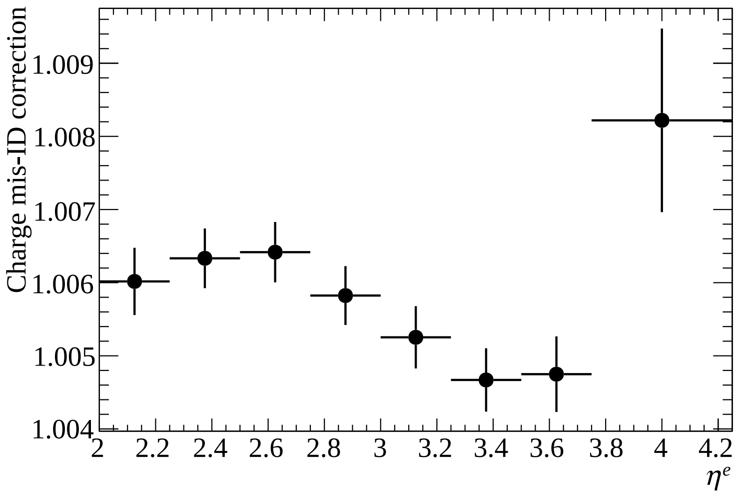


Figure 5.37: The charge mis-identification correction for the charge asymmetry measurement.

Source	Uncertainty [%]		
	$\sigma_{W^+\to e^+\nu_e}$	$\sigma_{W^-\to e^-\overline{\nu}_e}$	$R_{W^\pm}$
Statistical†	0.19	0.24	0.30
Yield (statistical)†	0.28	0.40	0.48
Yield (systematic)	1.42	1.79	0.51
Efficiency (statistical)†	0.55	0.55	0.21
Efficiency (systematic)	1.11	1.14	0.54
FSR corrections†	0.05	0.07	0.09
Acceptance corrections (statistical)†	0.00	0.01	0.01
Acceptance corrections (systematic)	0.15	0.15	0.00
Charge mis-identification†	—	—	0.02
Systematic	1.91	2.23	0.91
Beam energy	1.00	0.86	0.14
Luminosity	1.16	1.16	—
Total	2.46	2.67	0.97

Table 5.9: Summary of the relative uncertainties on the W^+ and W^- boson cross-sections and on the cross-section ratio. Uncertainties marked with † are assumed to be uncorrelated between bins; all others are taken to be correlated.

5.9 Systematic uncertainties

Several sources of systematic uncertainty, both statistical and systematic in nature, enter into the measurement. These are discussed in detail in what follows and summarised in Table 5.9 for the total cross-sections and the ratio measurement with $R_{W^\pm} \equiv \sigma_{W^+\to e^+\nu_e}/\sigma_{W^-\to e^-\overline{\nu}_e}$.

5.9.1 The heavy flavour fit

Both statistical and systematic contributions are considered when determining the uncertainty on the HF fraction in the SR. The statistical component is evaluated using 10 000 pseudoexperiments where each bin of each template is varied by drawing from a Poisson distribution with a mean equal to the number of entries in that bin. The histograms used to correct for alignment and extrapolate photon+jets are also varied in this manner. Using this procedure, 10 000 copies of each template are generated and the fit is repeated for each set. The distribution of HF fractions can be seen in Fig. 5.38. The

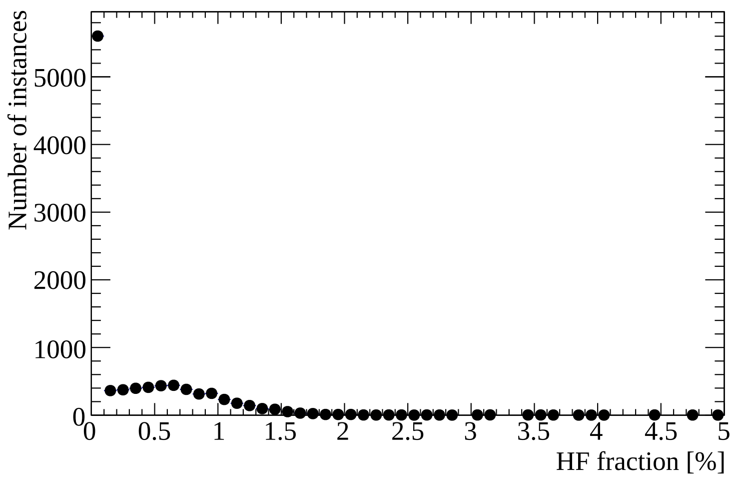


Figure 5.38: The distribution of HF fractions as a result of repeated pseudoexperiments.

variation in the HF fraction is not a symmetric distribution because the nominal fraction is very close to zero. However, higher fractions of around 0.65% are also favoured in a minority of cases. The mean of the distribution is $\mu_{\mathrm{HF}} = 0.31\%$. The standard deviation of the distribution, $\sigma_{\mathrm{HF}}^{\mathrm{stat}} = 0.46\%$, is a gauge of the effect of finite statistics in the fitting templates and is taken as the statistical uncertainty in the HF fraction.

The systematic uncertainty is determined by considering many effects. The $W \to \tau\nu$ and $Z \to \tau\tau$ contributions are varied within their uncertainties. As an estimate of the uncertainty related to the relative normalisations of $c\bar{c}$ and $b\bar{b}$, only the $c\bar{c}$ component is fitted. The background contamination to fake electrons and photon+jets is varied by its uncertainty. Finally, alternative templates for the fake electron component are considered. Overall, the HF fraction is seen to vary in the range $[0\%, 0.52\%]$. The statistical and systematic uncertainties are combined and the uncertainty in the normalisation of the HF component is determined to be 0.8%, which forms a component of "Yield (systematic)" in Table 5.9. This uncertainty is taken to be fully correlated between the bins of the measurement.

5.9.2 Fits to p_{T}^{e}

The fits to p_{T}^{e} are affected by a statistical uncertainty that derives from the finite simulation samples used to carry out the fits and by several sources of systematic uncertainty. The statistical uncertainty in the fits is estimated analogously to the procedure outlined previously. For each fit, 3 330 pseudoexperiments are carried out. The standard deviation of the distribution is taken as the statistical component of the systematic uncertainty and is denoted as "Yield (statistical)" in Table 5.9. The distribution of yields for each bin of η^e is shown in Appendix B. All other sources of uncertainty in the fits are considered systematic in nature (denoted as "Yield (systematic)" in Table 5.9) and are described subsequently.

Templates for contributions from photon+jets, fake electrons and HF, determined using data, contain a mixture of physical processes. A simulation-based estimate for EW contamination is subtracted and a 50% systematic uncertainty is assigned for the procedure. It should be noted that contamination from HF and fake electrons cannot be estimated using this procedure. Photon+jets is not expected to contaminate the fake electron template as it does not populate the region in the I_{T}^{e}-E_{HCAL}/p^{e} plane that is used to extract a sample.

A charge-independent correction for the modelling of tight selection requirements is applied to the simulation-based templates as a function of pseudorapidity. The correction factors are varied up and down by one standard deviation based on the statistical uncertainty with which they have been determined.

Two alternative regions of phase space are considered for extracting the fake electron component as detailed in Sect. 5.5.4. The fits are repeated with the alternative definitions to ascertain the uncertainty associated with the fake electron template. The HF component is constrained in the fit based on a fit to the $\log_{10}(\chi^2_{\mathrm{IP}})$ distribution. The template is varied by 1σ according to the uncertainty.

An uncertainty of 1% is assigned on the jet fraction correction, which corresponds to a ~5σ shift about the minimum of the distribution in Fig. 5.22 (a). Examining the yields in Fig. 5.22 (b) indicates a corresponding variation of ~0.1% in the number of candidates. The uncertainty on the differential measurements is determined by varying the jet fraction correction up and down by 1% in every fit.

The transverse momentum of the candidate in simulation is sensitive to both the potential mismodelling of track reconstruction as well as to the description of the material traversed by the candidate. The latter affects the number of bremsstrahlung photons emitted and thus has an impact on p_{T}^{e} and, by extension, on the fits. Any potential mismodelling can be described by a scaling of the transverse momentum, *i.e.* a multiplicative factor, denoted as α. The effect of varying the momentum scale on all simulation-based templates on the quality of the inclusive fit is shown in Fig. 5.39. The momentum scale is seen to be consistent with unity suggesting that material in the

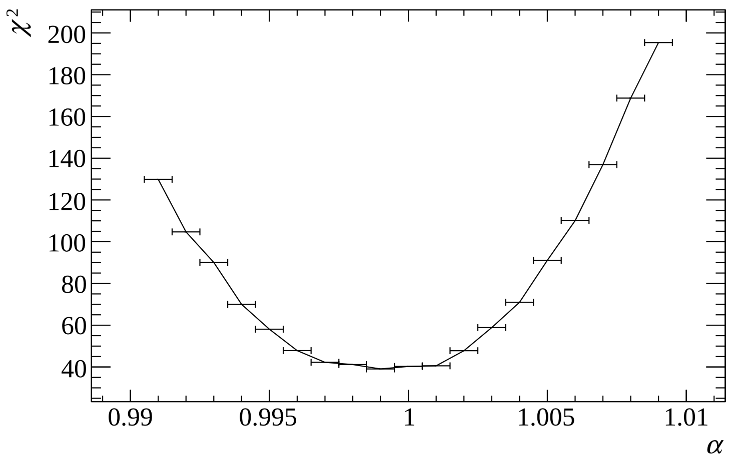


Figure 5.39: The χ^2 of the inclusive fit is shown as a function of the momentum scale, α.

detector is well-modelled. An uncertainty of 0.5% assigned on the momentum scale in Ref. [20] is found to be appropriate for the measurement. Varying the momentum scale by its uncertainty in the fits binned in η^e leads to an uncertainty of 1.3% on the total cross-section which is the largest contribution to "Yield (systematic)".

The known ratio of branching fractions is used to constrain $W \to \tau\nu$ to $W \to e\nu$ in the kinematic region of the generator-level requirements. The uncertainty on the ratio, 2.7% [5], is propagated to the measurement. Additionally, $W \to \tau\nu$ is scaled to agree in the fiducial volume of the measurement. The difference in scaling factors between these approaches is used to set an additional systematic uncertainty.

The constraints on $Z \to ee(\tau\tau)$ are subject to an uncertainty on the measured cross-section. The low-mass DY component is constrained in the fit using a NLO cross-section computed using MCFM and the estimate for the luminosity. The constraint is varied according to their respective uncertainties. Rare processes are also constrained using theoretical cross-sections. The templates are subject to an uncertainty on the cross-sections and an uncertainty on the luminosity.

5.9.3 Efficiencies

The statistical uncertainty on the combined efficiency is taken as a contribution to the uncertainty on the measurement and is denoted as "Efficiency (statistical)" in Table 5.9. In the case of cross-sections, the uncertainties from the finite statistics of the Z data and

Z/W simulated samples all contribute. For the determination of the cross-section ratio and the charge asymmetry, only the uncertainty due to the simulation of the W must be accounted for. All other sources of uncertainty in the efficiencies are collectively denoted as "Efficiency (systematic)" in Table 5.9 and are described in the next paragraphs.

Data-driven cross-checks as described in Ref. [20] are performed on the efficiencies determined using simulation. The cross-checks lead to an uncertainty of 0.5% on the track reconstruction efficiency, an uncertainty of 0.6% on the kinematic efficiency due to the modelling of bremsstrahlung in simulation, and an uncertainty of 0.6% on PID requirements.

Frequentist 1σ Clopper-Pearson intervals [150] are computed for each efficiency. In the case of efficiencies determined from $Z \to ee$ data, same-sign events enter with a weight of -1, which effectively make the histograms weighted. Clopper-Pearson intervals are not used in this case and the efficiency is taken to be normally distributed.

A study that formed electron and charged pion combinations and counted opposite- and same-sign pairs [20] leads to a systematic uncertainty of 0.25% on the $W \to e\nu$ cross-section due to the normalisation of hadronic contamination in the sample of $Z \to ee$ candidates.

Global event cut efficiency

The statistical component of the uncertainty on the overall GEC efficiency is found to be 0.0008. Several sources of systematic uncertainty are considered in the determination of the GEC efficiency. Since ϵ^{GEC} is dependent on the number of primary vertices, the efficiency is measured separately for $N_{\mathrm{PV}} = \{1, 2, 3, \geq 4\}$, and statistically combined. The matching procedure for events separated by the number of primary vertices is illustrated in Fig. 5.40. This is compared with the estimate of the efficiency obtained inclusively for all numbers of primary vertices and an uncertainty is assigned based on the difference between the two methods. This difference can be thought of as quantifying the robustness of the matching method. This is the dominant uncertainty on the efficiency and determined to be 0.0030.

The normalisations of the background components are varied within their uncertainty and the matching procedure is repeated. No shift is found in the efficiency and so no systematic uncertainty is assigned. Additionally, the difference between the Kolmogorov-Smirnov test of compatibility and the χ^2 test is considered. The methods are found

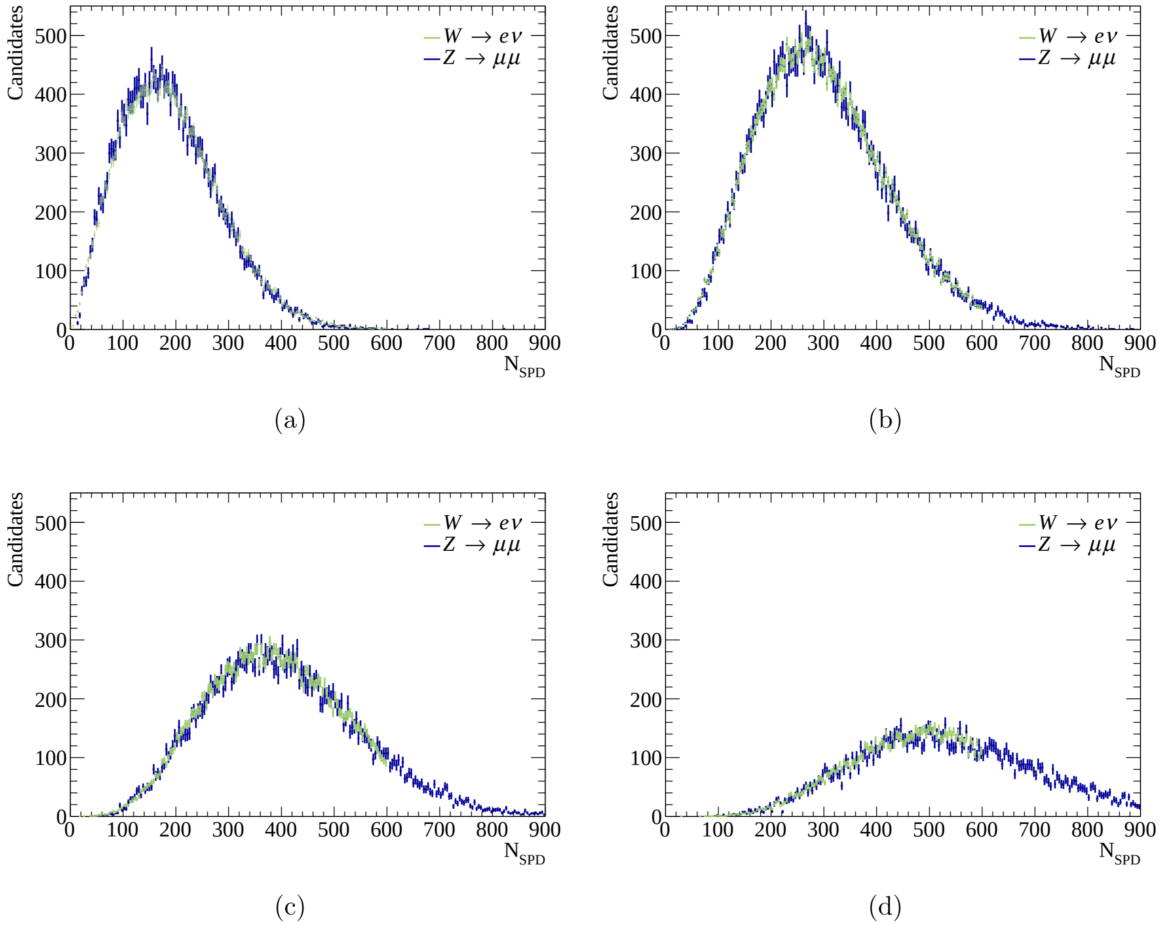
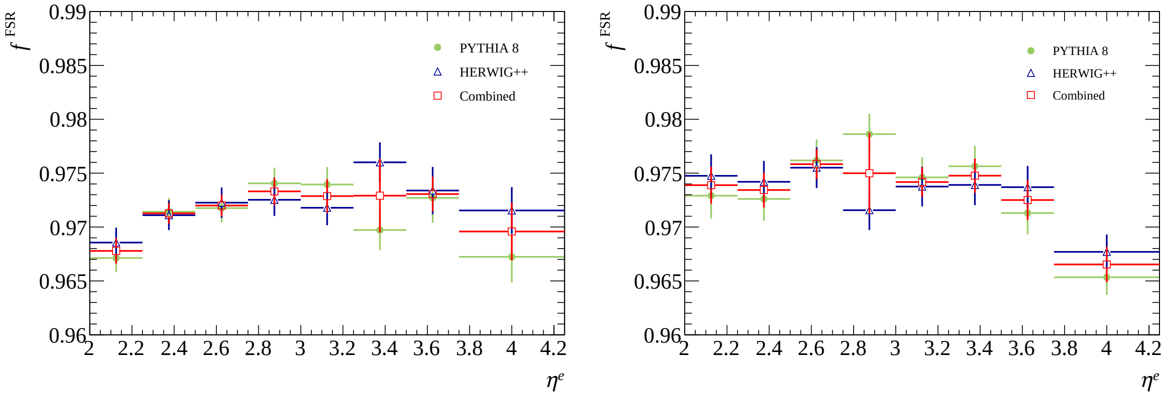


Figure 5.40: The matching procedure to $Z \to \mu\mu$ data performed for events with (a) one primary vertex, (b) two primary vertices, (c) three primary vertices, and (d) more than four primary vertices.

to yield identical results so no systematic uncertainty is assigned. Furthermore, the statistical uncertainty on the bias resulting from the tighter selection is propagated to the uncertainty. Finally, an exponential is considered instead of a Γ-function to determine the fraction of events beyond $N_{\text{SPD}} = 900$ and the fitting parameters are varied. The efficiency is found to vary by 0.0003 as a result and this is propagated to the uncertainty.

Overall, a systematic uncertainty of 0.0031 is assigned for the procedure to determine the efficiency in data which corresponds to a 0.34% uncertainty on the efficiency. An additional systematic uncertainty is assigned on the cross-sections, the cross-section ratio, and the charge asymmetry to account for the differences observed between e^- and e^+ as shown in Fig. 5.32 (b).



(a) (b)

Figure 5.41: FSR corrections obtained from Pythia 8 are compared to those obtained from a Herwig++ event generation for (a) e^+ and (b) e^-. The uncertainties on the combined corrections reflect both the statistics available as well as differences between the generators.

5.9.4 Final-state radiation

Corrections due to FSR are computed using Pythia 8 and Herwig++ as illustrated in Fig. 5.41 and combined. Half the difference between Pythia 8 and Herwig++ predictions is taken as the systematic component of the uncertainty on FSR corrections. The uncertainty is taken to be uncorrelated between bins of η^e and between the e and μ final states.

5.9.5 Acceptance corrections

The statistical uncertainty on the acceptance corrections arises from the ResBos W simulated sample. Half the difference between Pythia 8 and ResBos is taken as a systematic uncertainty on a bin-by-bin basis and is assumed to be correlated between bins. The disagreement between the two simulations is illustrated in Fig. 5.42.

Pythia 8 simulation is used to determine the true transverse momentum that corresponds to $p_T^e = 65\,\text{GeV}$ (*i.e.* as reconstructed in the detector). This procedure yields a difference of $0.25\,\text{GeV}$ between e^- and e^+ that is assigned as a systematic uncertainty. In addition, the procedure is repeated in bins of η^e and found to be consistent apart from in the two most forward bins, $[3.50, 3.75]$ and $[3.75, 4.25]$. An additional systematic uncertainty is assigned to account for the discrepancy seen in those bins.

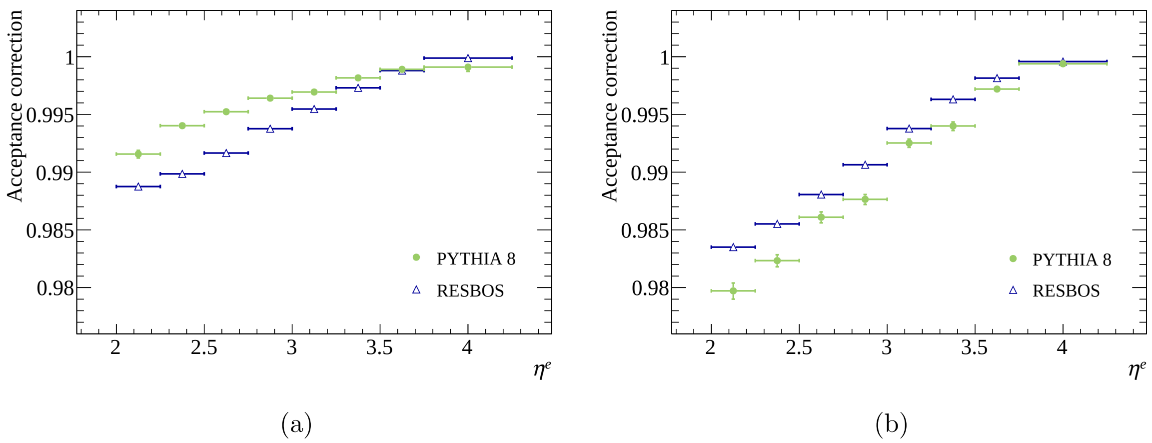


Figure 5.42: The acceptance factors obtained from ResBos are compared to those obtained from Pythia 8 simulation for (a) e^+ and (b) e^-. The uncertainties are statistical only.

5.9.6 The LHC beam energy and luminosity

An uncertainty of 0.7% on the LHC beam energy at 8 TeV [151] leads to a relative uncertainty on the W^+ (W^-) cross-section of 1.00 (0.86)% determined using DYNNLO [152]. The uncertainty on the luminosity is 1.16% for the 8 TeV dataset [153].

5.9.7 Propagation of uncertainties

When computing derived quantities such as the total cross-section, cross-section ratios, and the charge asymmetry, correlations between the 16 measurements of $W \to e\nu$ in bins of η^e must be accounted for. Uncertainties marked with † in Table 5.9 are statistical in nature and are assumed to be uncorrelated between charges and bins of η^e. All other sources of systematic uncertainty, excluding beam and luminosity, are varied by one standard deviation around their nominal value for each measurement and the correlation between each pair of measurements is computed. Correlation matrices between bins of η^e for W^+, W^-, and W^+ against W^- are reported in Appendix A. A consequence of the sizeable positive correlations is that many of the systematic uncertainties add coherently when integrating over bins, but partially cancel in determining W^+/W^- ratios.

Sect. 5.10.4 presents the ratio of the $W \to e\nu$ and $W \to \mu\nu$ branching fractions. Here, the systematic uncertainties of the respective measurements are taken to be uncorrelated between the two final states apart from the uncertainties on the beam, luminosity, GEC efficiency and the acceptance correction, which are taken to be fully correlated.

5.10 Results

5.10.1 Inclusive results

Total inclusive cross-sections for $W \to e\nu$ production are obtained by summing the cross-sections in bins of η^e. The Born level cross-sections in the fiducial region defined as $2.0 < \eta^e < 4.25$ and more than 20 GeV of transverse momentum are measured to be

$$\begin{aligned}
\sigma_{W^+\to e^+\nu_e} &= 1124.4 \pm 2.1 \pm 21.5 \pm 11.2 \pm 13.0\,\text{pb},\\
\sigma_{W^-\to e^-\overline{\nu}_e} &= 809.0 \pm 1.9 \pm 18.1 \pm 7.0 \pm 9.4\,\text{pb},\\
\sigma_{W\to e\nu} &= 1933.3 \pm 2.9 \pm 38.2 \pm 18.2 \pm 22.4\,\text{pb},
\end{aligned}$$

where the first uncertainties are statistical, the second are systematic, the third are due to the knowledge of the LHC beam energy and the fourth are due to the luminosity determination.

The W^+ to W^- cross-section ratio is determined to be

$$R_{W^\pm} = 1.390 \pm 0.004 \pm 0.013 \pm 0.002,$$

where uncertainties are statistical, systematic and due to the LHC beam energy measurement, respectively.

5.10.2 Cross-sections as a function of electron pseudorapidity

Born level cross-sections as a function of electron pseudorapidity are given in Tables 5.10 and 5.11 along with corresponding FSR corrections. The differential cross-sections as a function of η^e are also determined and shown in Fig. 5.43. Measurements are compared to theoretical predictions calculated with the Fewz [154, 155] generator at NNLO for the six PDF sets: ABM12 [68], CT14 [156], HERA1.5 [157], MMHT14 [158], MSTW08 [56], and NNPDF3.0 [159]. Satisfactory agreement is observed apart from in the far forward region of the W^+ differential measurement, where the PDF uncertainties are also greatest.

η^e	$\sigma_{W^+\to e^+\nu_e}$[pb]	f^{FSR}
2.00 - 2.25	$229.9 \pm 1.0 \pm 6.5 \pm 2.3 \pm 2.7$	0.9671 ± 0.0013
2.25 - 2.50	$210.1 \pm 0.8 \pm 4.7 \pm 2.1 \pm 2.4$	0.9714 ± 0.0013
2.50 - 2.75	$191.7 \pm 0.8 \pm 4.9 \pm 1.9 \pm 2.2$	0.9718 ± 0.0013
2.75 - 3.00	$156.3 \pm 0.7 \pm 3.4 \pm 1.6 \pm 1.8$	0.9741 ± 0.0015
3.00 - 3.25	$132.0 \pm 0.7 \pm 3.1 \pm 1.3 \pm 1.5$	0.9739 ± 0.0016
3.25 - 3.50	$87.6 \pm 0.6 \pm 2.2 \pm 0.9 \pm 1.0$	0.9697 ± 0.0019
3.50 - 3.75	$59.1 \pm 0.5 \pm 2.1 \pm 0.6 \pm 0.7$	0.9727 ± 0.0023
3.75 - 4.25	$57.8 \pm 0.7 \pm 2.7 \pm 0.6 \pm 0.7$	0.9672 ± 0.0024

Table 5.10: The Born level cross-section for W^+ boson production in bins of electron pseudorapidity. The first uncertainties are statistical, the second are systematic, the third are due to the knowledge of the LHC beam energy and the fourth are due to the luminosity measurement. The rightmost column gives values of the additional factor, f^{FSR}, by which the results should be multiplied in order to give the cross-sections after FSR.

η^e	$\sigma_{W^-\to e^-\overline{\nu}_e}$[pb]	f^{FSR}
2.00 - 2.25	$132.8 \pm 0.8 \pm 4.1 \pm 1.1 \pm 1.5$	0.9729 ± 0.0021
2.25 - 2.50	$120.8 \pm 0.7 \pm 3.1 \pm 1.0 \pm 1.4$	0.9726 ± 0.0020
2.50 - 2.75	$113.0 \pm 0.7 \pm 2.9 \pm 1.0 \pm 1.3$	0.9762 ± 0.0020
2.75 - 3.00	$103.3 \pm 0.6 \pm 2.7 \pm 0.9 \pm 1.2$	0.9786 ± 0.0019
3.00 - 3.25	$99.3 \pm 0.6 \pm 2.7 \pm 0.9 \pm 1.2$	0.9746 ± 0.0019
3.25 - 3.50	$78.8 \pm 0.6 \pm 2.2 \pm 0.7 \pm 0.9$	0.9756 ± 0.0019
3.50 - 3.75	$67.0 \pm 0.6 \pm 2.8 \pm 0.6 \pm 0.8$	0.9713 ± 0.0020
3.75 - 4.25	$94.0 \pm 0.9 \pm 4.2 \pm 0.8 \pm 1.1$	0.9653 ± 0.0016

Table 5.11: The Born level cross-section for W^- boson production in bins of electron pseudorapidity. The first uncertainties are statistical, the second are systematic, the third are due to the knowledge of the LHC beam energy and the fourth are due to the luminosity measurement. The rightmost column gives values of the additional factor, f^{FSR}, by which the results should be multiplied in order to give the cross-sections after FSR.

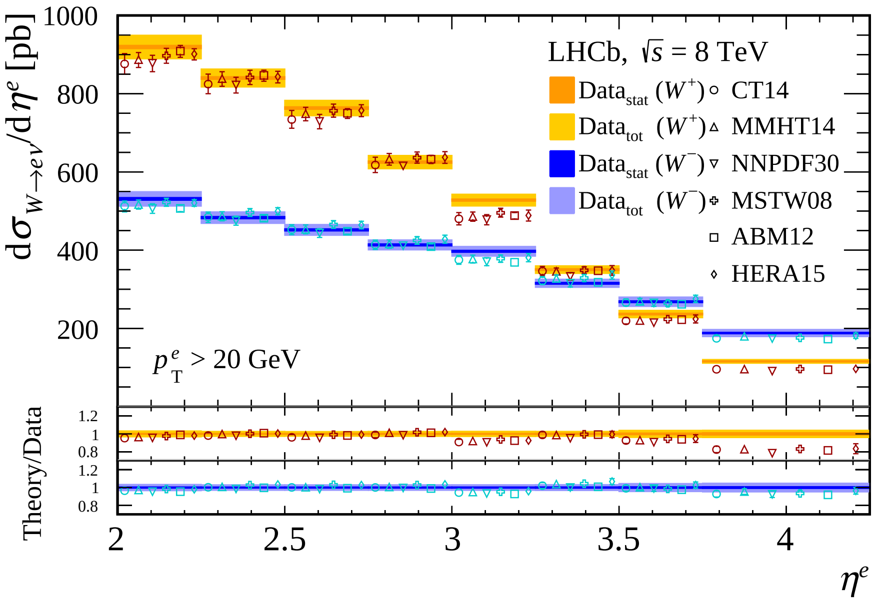


Figure 5.43: The differential W^+ and W^- cross-sections in bins of η^e. Measurements, represented as bands, are compared to NNLO predictions with different parameterisations of the PDFs (markers are displaced horizontally for presentation). The bottom panel displays the theory predictions divided by the measured cross-sections.

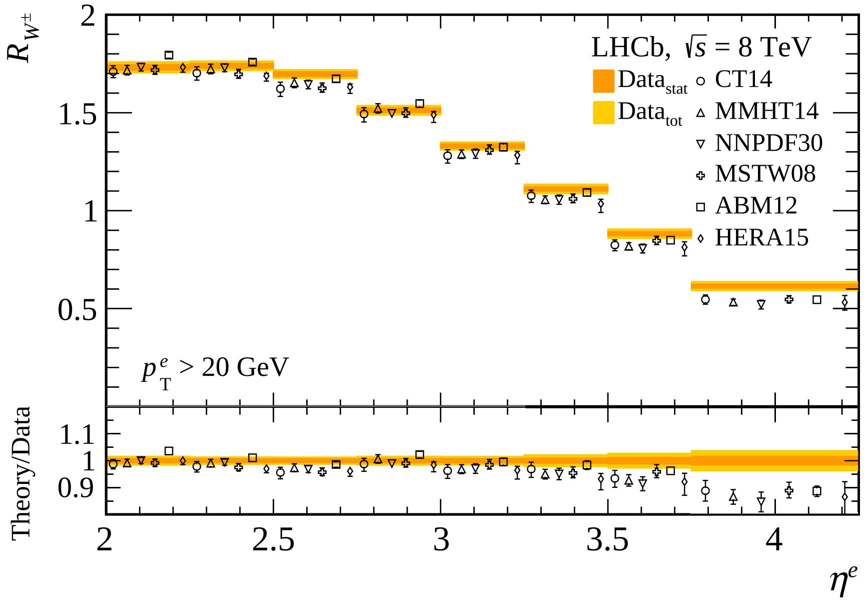


Figure 5.44: The W^+ to W^- cross-section ratio in bins of η^e. Measurements, represented as bands, are compared to NNLO predictions with different parameterisations of the PDFs (markers are displaced horizontally for presentation). The bottom panel displays the theory predictions divided by the measured cross-section ratios.

5.10.3 Cross-section ratio and charge asymmetry

Cross-section ratios as a function of η^e are compared to theory predictions in Fig. 5.44 and listed separately in Table 5.12. Overall the measurements are in agreement with theory predictions, with the exception of the far forward region. In this region the measured ratio is higher than the expectation as a consequence of the discrepancy seen in the W^+ cross-section in that region.

A commonly used alternative to the cross-section ratio is the W boson production charge asymmetry, defined as

$$A_e \equiv \frac{\sigma_{W^+\to e^+\nu_e} - \sigma_{W^-\to e^-\overline{\nu}_e}}{\sigma_{W^+\to e^+\nu_e} + \sigma_{W^-\to e^-\overline{\nu}_e}}. \tag{5.6}$$

The asymmetry measurements are given in Table 5.13 and compared to theory predictions in bins of η^e in Fig. 5.45. The change in sign of the charge asymmetry in the far forward region is a consequence of the axial vector coupling in the weak interaction which introduces preferred directions for the W boson decay products. As a consequence, positrons are less likely to be emitted in the far forward region than electrons are.

η^e	$R_{W^\pm}$
2.00 - 2.25	$1.731 \pm 0.013 \pm 0.026 \pm 0.003$
2.25 - 2.50	$1.739 \pm 0.012 \pm 0.025 \pm 0.003$
2.50 - 2.75	$1.697 \pm 0.012 \pm 0.022 \pm 0.003$
2.75 - 3.00	$1.512 \pm 0.011 \pm 0.023 \pm 0.002$
3.00 - 3.25	$1.330 \pm 0.011 \pm 0.019 \pm 0.002$
3.25 - 3.50	$1.111 \pm 0.010 \pm 0.025 \pm 0.002$
3.50 - 3.75	$0.882 \pm 0.011 \pm 0.023 \pm 0.001$
3.75 - 4.25	$0.615 \pm 0.010 \pm 0.022 \pm 0.001$

Table 5.12: The W^+ to W^- cross-section ratio in bins of electron pseudorapidity. The first uncertainties are statistical, the second are systematic and the third are due to the knowledge of the LHC beam energy.

η^e	$A_e(\%)$
2.00 - 2.25	$26.78 \pm 0.36 \pm 0.70 \pm 0.07$
2.25 - 2.50	$26.98 \pm 0.32 \pm 0.66 \pm 0.07$
2.50 - 2.75	$25.84 \pm 0.33 \pm 0.60 \pm 0.07$
2.75 - 3.00	$20.39 \pm 0.36 \pm 0.74 \pm 0.07$
3.00 - 3.25	$14.15 \pm 0.39 \pm 0.70 \pm 0.07$
3.25 - 3.50	$5.25 \pm 0.47 \pm 1.11 \pm 0.07$
3.50 - 3.75	$-6.25 \pm 0.60 \pm 1.28 \pm 0.07$
3.75 - 4.25	$-23.85 \pm 0.75 \pm 1.72 \pm 0.07$

Table 5.13: The W boson production charge asymmetry in bins of electron pseudorapidity. The first uncertainties are statistical, the second are systematic and the third are due to the knowledge of the LHC beam energy.

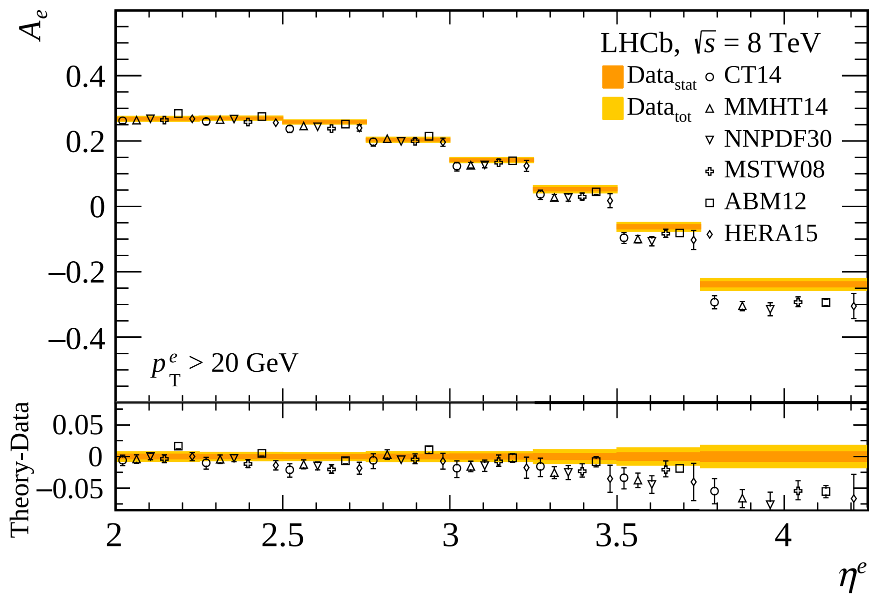


Figure 5.45: The W boson production charge asymmetry in bins of η^e. Measurements, represented as bands, are compared to NNLO predictions with different parameterisations of the PDFs (markers are displaced horizontally for presentation). The bottom panel displays the difference between theory predictions and the measured charge asymmetry.

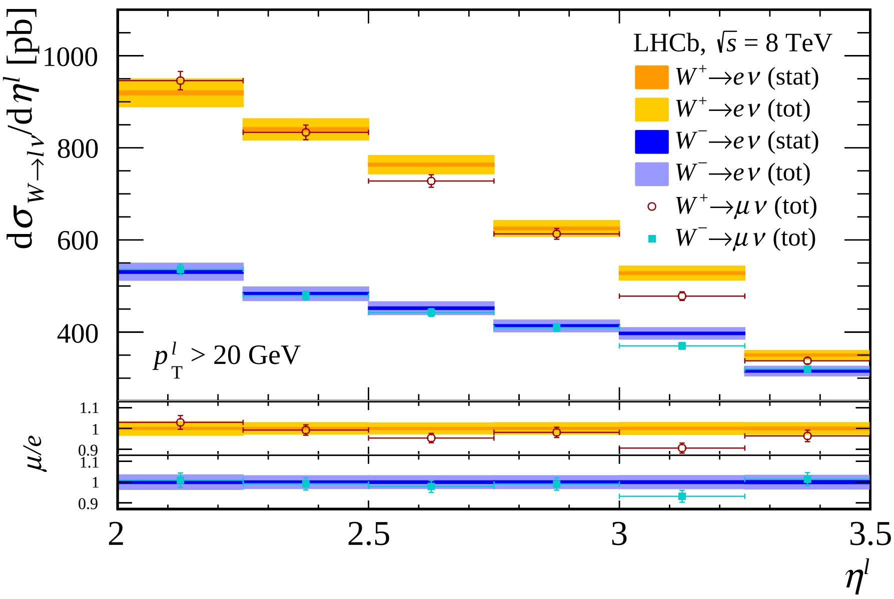


Figure 5.46: The differential W^+ and W^- cross-sections in bins of η^l. The measurement using electrons, represented as bands, is compared to the measurement in the muon final state. The bottom panel displays the muon results divided by the measurements in the electron final state.

5.10.4 Lepton universality

Production of W bosons in the forward region has also been studied in the muon final state [22]. The muon measurement had a different upper kinematic limit in pseudorapidity, and consequently the bin boundaries only coincide with the present measurement for $\eta^l < 3.50$. The results are therefore compared in the range $2.00 < \eta^l < 3.50$ as is shown in Figs. 5.46, 5.47, and 5.48. The results of these measurements are seen to be consistent with the $W \to \mu\nu$ measurements and no significant deviation from lepton universality is observed once uncertainties and correlations between measurements are taken into account. Fig. 5.46 shows good agreement, apart from the bin $3.00 < \eta^l < 3.25$ for W^+, where the difference is approximately 3 standard deviations.

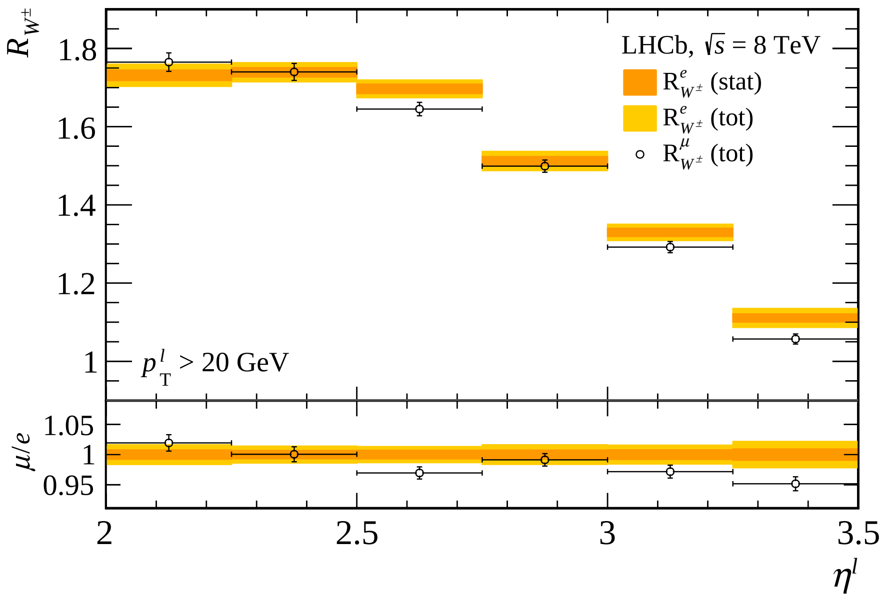


Figure 5.47: The W^+ to W^- cross-section ratio in bins of η^l. The measurement using electrons, represented as bands, is compared to the measurement in the muon final state. The bottom panel displays the muon results divided by the measurements in the electron final state.

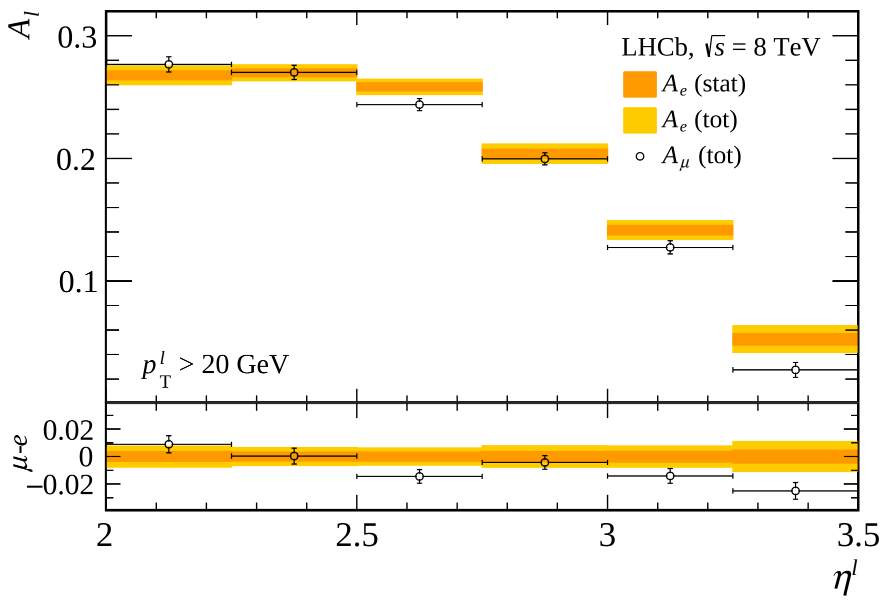


Figure 5.48: The W boson production charge asymmetry in bins of η^l. The measurement using electrons, represented as bands, is compared to the measurement in the muon final state. The bottom panel displays the difference between the muon and electron final states.

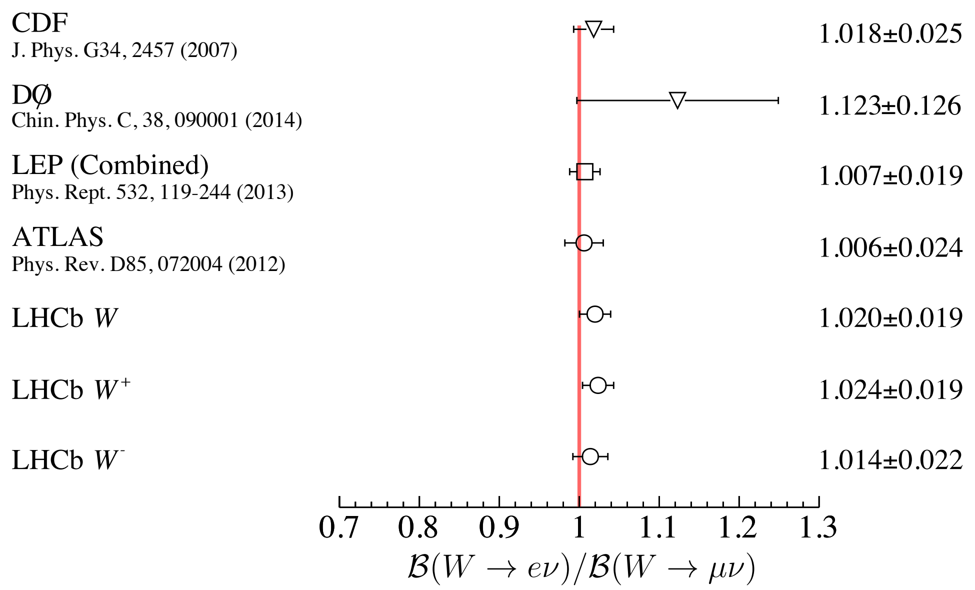


Figure 5.49: The ratio of branching fractions for the electron and muon final states determined for W, W^+, and W^- is compared to hadron collider and LEP results. The theory expectation is represented by the red line.

The consistency with lepton universality is quantified by computing a ratio of W branching fractions using cross-sections determined in the range $2.00 < \eta^l < 3.50$. In this limited acceptance, the ratios of W branching fractions are determined to be

$$\begin{aligned}
\mathcal{B}(W^+ \to e^+\nu_e)/\mathcal{B}(W^+ \to \mu^+\nu_\mu) &= 1.024 \pm 0.003 \pm 0.019, \\
\mathcal{B}(W^- \to e^-\overline{\nu}_e)/\mathcal{B}(W^- \to \mu^-\overline{\nu}_\mu) &= 1.014 \pm 0.004 \pm 0.022, \\
\mathcal{B}(W \to e\nu)/\mathcal{B}(W \to \mu\nu) \qquad &= 1.020 \pm 0.002 \pm 0.019,
\end{aligned}$$

where the first uncertainties are statistical and the second are systematic. The result is compared to past measurements [5, 11, 12, 16] in Fig. 5.49 and its precision is seen to surpass previous individual determinations of the ratio and to be comparable to the combined LEP result.

We now proceed to the conclusions and discuss prospective future measurements that have the potential to further the aims of this thesis by probing the SM with unprecedented precision in the unique kinematic region that LHCb has access to.

Chapter 6

Conclusions

This thesis presents measurements of the cross-sections for W boson production in pp collisions at a centre-of-mass energy of $\sqrt{s} = 8\,\text{TeV}$ in the electron final state. The cross-section ratio and the charge asymmetry are also determined. The measurements are found to be in agreement with NNLO calculations in perturbative QCD.

These results represent the first measurements of $W \to e\nu$ production in the forward region at the LHC and are complementary to the previously published measurements of $W \to \mu\nu$ production [22]. The measurements have been performed using statistically independent datasets with largely independent systematic uncertainties. The measurements reported here are found to be consistent with the $W \to \mu\nu$ results.

Comparable precision to the $W \to \mu\nu$ results is achieved in the measurements of the cross-sections and the cross-section ratio has been determined with sub-percent precision. Due to the unique kinematic acceptance of the LHCb detector, these results will be valuable in constraining the parton distribution functions of the proton at low and high values of the Bjorken-x variable.

Finally, the measurements of W boson production in the electron and muon final states are consistent with lepton universality and the ratio of branching fractions has been determined with precision that surpasses all past determinations at hadron colliders as well as measurements made at the LEP collider.

This thesis also describes a study of matching and merging (M&M) schemes and weak showering (WS), as applied to the production of EW bosons in association with jets at the LHC. These novel and advanced theoretical tools are compared to measurements performed at the LHC by the ATLAS, CMS, and LHCb collaborations at $\sqrt{s} = 7\,\text{TeV}$ and $\sqrt{s} = 13\,\text{TeV}$.

M&M schemes are seen to provide a good description of event shapes in the central and forward regions and to model $p_{\text{T}}^{\text{jet}}$ and p_{T}^{Z} distributions accurately in the forward region. WS, while accurate to LO only, is able to significantly improve the description of the jet multiplicity and $p_{\text{T}}^{\text{jet}}$ distributions centrally. WS also yields an improvement in the modelling of event shapes.

Finally, measurements performed at LHCb do not yet access regions of phase space where WS is expected to be dominant. WS becomes a significant, or even the dominant, production mechanism of EW bosons for event topologies with two, or more, jets and scales beyond 100 GeV in $p_{\text{T}}^{\text{jet}}$. It is clear, however, that the reach of measurements performed at LHCb will improve with larger datasets at higher centre-of-mass energies. In parallel, WS effects such as the suppression of the jet rate, as well as multiple EW boson emissions, are expected to become more common at these energies.

This study, presented in Chapter 3, constitutes a test of precise theoretical tools in a novel region of phase space. It proves that a variety of M&M schemes can be employed with some confidence in modelling EW boson production in association with jets. Finally, it is clear that with the extraordinary success of the SM and with beyond-the-Standard-Model (BSM) physics proving elusive, increased precision is required in both the experimental measurements' programme at the LHC as well as in developing theoretical tools such as matching and merging and weak showering.

We now turn to prospective future measurements that follow from the research programme described in this thesis.

6.1 Prospective future measurements

A natural extension of the measurements of forward $W \to e\nu$ production at $\sqrt{s} = 8\,\text{TeV}$ would be an analogous measurement using Run II data from the LHC collected at $\sqrt{s} = 13\,\text{TeV}$. As the $Z \to ee$ and $Z \to \mu\mu$ measurements have been performed [2], a comparable analysis of W boson production in either leptonic final state is a notable absence from the LHCb measurements' programme. Furthermore, as the measurement of $W \to e\nu$ necessitated extensive study of jet production in association with the EW boson, a standalone measurement of $W(\to e\nu)$+jets could be performed. This would be a valuable extension to the LHCb jet physics programme that already includes a measurement of $W(\to \mu\nu)$+jets [26].

Further to the above, measurements of EW boson production in association with b/c-identified jets are advantageous[1]. Specifically, W+c-jets probes the s-quark PDFs, while a measurement of Z+c-jets, as proposed in Ref. [66], would establish whether the proton wave function contains an intrinsic charm (IC) component. The existence of IC has attracted considerable interest [161] and would illuminate the role of non-perturbative dynamics in the nucleon sea. IC would also have an impact on the cross-sections of many processes at the LHC such as W and H boson production and would be a vital component in determining the charm Yukawa coupling.

An excellent extension to the above would be a measurement of γ+c-jets which would probe the proton at lower energy scales. As a methodology to extract γ+jets from data was developed for the measurement of $W \to e\nu$, measurements of γ+jets would be a natural extension of the techniques presented in this thesis. These measurements would enable the gluon PDF to be constrained and also improve our knowledge of the c-quark PDF, if flavour identification is applied to the jets.

As discussed in previous chapters, measurements of W boson production in multiple leptonic final states constitute tests of lepton universality. This thesis has shown that the precision LHCb is able to reach in the ratio of the $W \to e\nu$ to $W \to \mu\nu$ branching fractions is world-class. Performing a ratio of the $W \to e\nu$ to $W \to \tau\nu$ branching fractions would be particularly interesting as the current world average departs from the SM expectation at the level of ~2σ [5].

[1] LHCb has performed measurements of W+b/c-jets [28] and top quark production [160] using Run I data at centre-of-mass energies of 7 and 8 TeV.

Finally, the work carried out to develop a template fit to the transverse momentum distribution of leptons from W bosons presented in this thesis can be used in performing a measurement of W boson mass. This measurement would be performed in the muon final state due to the improved precision over the electron final state. However, the methods developed for the fits to p_{T}^{e} would still be applicable.

A significant systematic uncertainty in the determination will be due to PDFs. The unique kinematic acceptance of LHCb means that these uncertainties will be anti-correlated with measurements performed by ATLAS and CMS. In fact, a weighted average of ATLAS, CMS and LHCb results, based only on PDF uncertainties, would be 30% more precise than a combination of ATLAS and CMS alone [162].

A combination of this kind is particularly interesting due to the fact that the indirect constraint on the W boson mass, based on EW parameters, is roughly a factor two more precise than the direct measurement, which has a precision of 15 MeV [162]. This makes it possible for BSM physics to affect the mass at higher orders in perturbation theory.

The LHC collaborations are expected to continue their research programmes for another 20 years. The intention is to both increase the centre-of-mass energy of the collisions and the instantaneous luminosity as well as improving the data acquisition capabilities of the detectors. The LHC experimental physics community, in cooperation with theorists, will continue to probe the SM at ever higher energy scales with ever greater precision. The aim, as it has always been, is to both be a builder of successful theories as well as a destroyer in pursuit of an ever greater understanding of Nature.

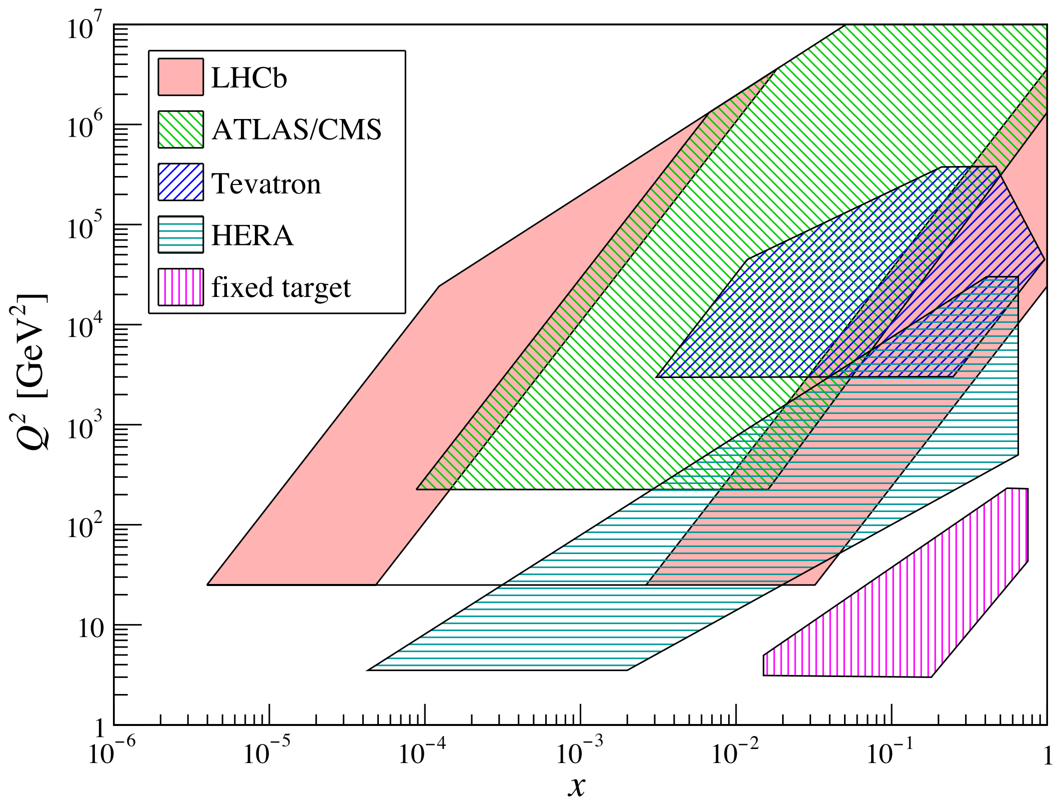
Chapter 7

Kokkuvõte

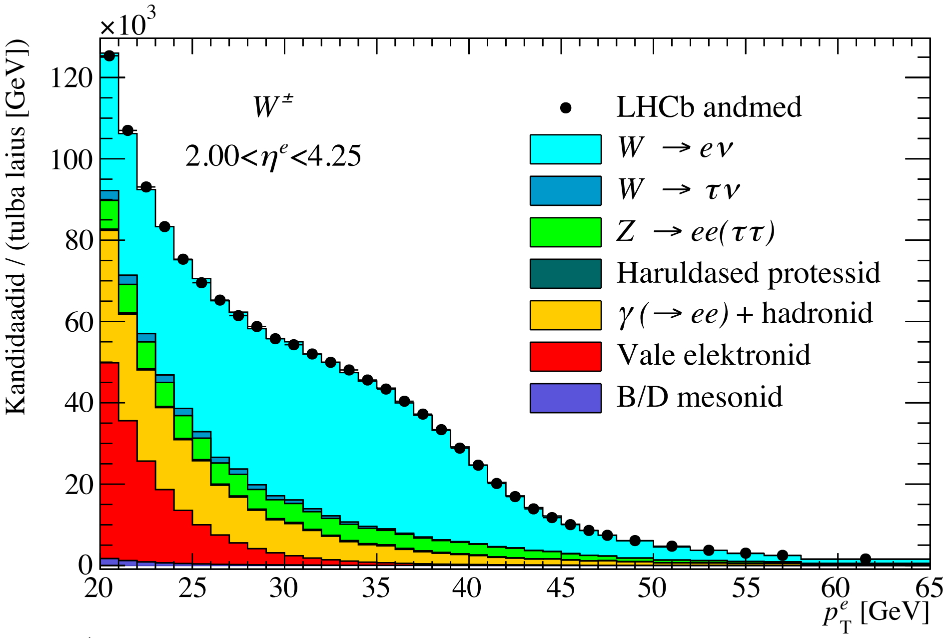
Inimkonna üks suurimaid saavutusi on olnud eduka teooria väljatöötamine mikromaailma kirjeldamiseks. Nähtav aine, mis on üks mikromaailma koostisosadest, koosneb elementaarosakestest: kuut tüüpi kvargist ja samajagu leptonitest, mis on jaotatud kolme põlvkonda. Elementaarosakeste standardmudel (SM) koosneb kvantväljateooriatest, mis kirjeldavad osakestevahelisi vastastikmõjusid.

Üks nendest vastastikmõjudest on nn nõrk interaktsioon, mille kandjateks on W ja Z bosonid. Need on erakordselt haruldased ning lühikese elueaga ($\sim 10^{-25}$ s) osakesed, mis avastati UA1 ja UA2 eksperimentide poolt 1983. aastal CERN'is [6–9]. Nõrk vastastikmõju mängib rolli näiteks aatomituumade lagunemisel ning samuti päikese põlemise reguleerimisel. Üks selle teooria põhipostulaatideks on nn leptoni universaalsus, mis väidab, et W ja Z bosonite interaktsiooni tugevus on võrdne kõigi leptonite põlvkondadega.

LHCb eksperiment LHC kiirendil on unikaalne oma võimes mõõta prootonite omadusi ja testida nõrka interaktsiooni uudses regioonis $x - Q^2$ faasiruumis nagu Joonis 7.1 näitab. Selles doktoritöös esitatakse kolme mõõtmistulemust: W bosoni ristlõige elektroni lagunemiskanalis [1], Z bosoni ristlõige müüonite ja elektronide kanalis [2] ning Monte Carlo simulatsiooni uuring kvantkromodünaamika mudeldamiseks koos W ja Z bosonitega [3].



Joonis 7.1: LHCb detektori mõõtmisvõimekus $x - Q^2$ faasiruumis [66]. Bjorkeni-x muutuja kirjeldab prootonisisest impulsijaotust ja Q on energiaskaala.



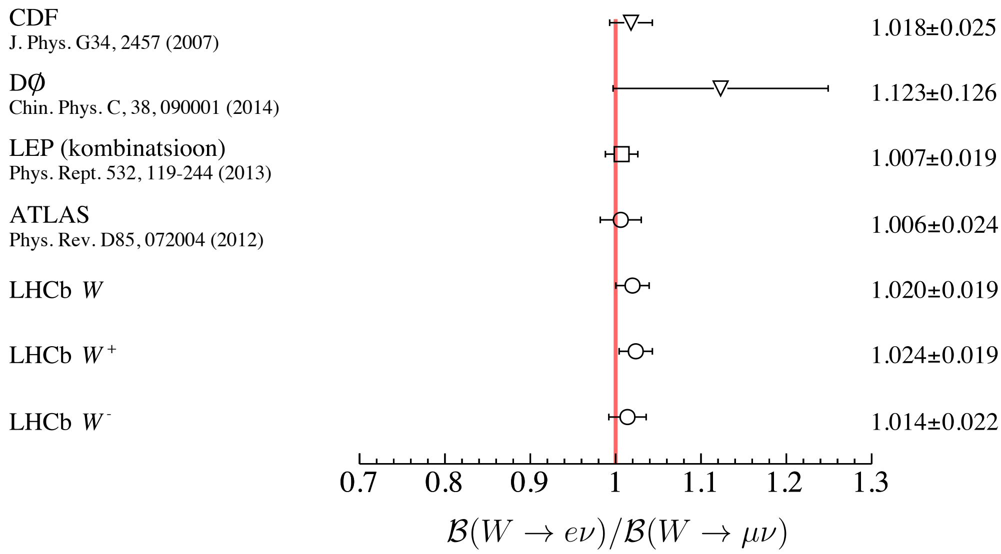
Joonis 7.2: Elektronide impulsijaotus mõõdetuna gigaelektronvoltides. Katseandmete taustal (mustad punktid) on näidatud erinevad protsessid, mis esinevad andmehulgas.

Ülalmainitud uuringute läbiviimiseks on kolm põhjust. Esiteks, täpsed mõõtmised võimaldavad testida SM'i kvantväljateooriaid. Teiseks, nende mõõtmiste abil on võimalik uurida prootoni sisestruktuuri ning seeläbi mõista ka teisi füüsikalisi protsesse paremini. Kolmandaks, üks LHC mõõtmisprogrammi peamisi ülesandeid on nn Uue Füüsika (UF) avastamine, mis võimaldaks uute teooriate väljatöötamist ja testimist. W ja Z bosonid on tihtipeale taustaprotsessid UF'i otsingutes ning seega vajavad uurimist.

W bosoni ristlõike mõõtmise suurim keerukus on taustaprotsesside määramine, mida tehakse regressioonianalüüsi abil elektronide implusijaotusele nagu näitab Joonis 7.2. Üks mõõtmise erakordseid tulemusi on W^+ ja W^- ristlõigete murru määramine üheprotsendilise täpsusega:

$$R_{W^\pm} = 1.390 \pm 0.004 \pm 0.013 \pm 0.002,$$

kus esimene mõõtemääramatus on statistilist laadi, teine on süstemaatiline ning kolmas on seotud kiire energiaga.



Joonis 7.3: Elektroni ja müüoni lagunemismurdude suhe W, W^+ ja W^- bosonitele. LHCb tulemust võrreldatakse hadronite kiirendite ja LEP'i tulemustega. Teoreetiline ennustus vastab punasele joonele.

Teine edukas tulemus on maailma täpseim individuaalne leptonite universaalsuse test müüonite ja elektronide vahel, mida mõõdeti järgnevate murdudega:

$$
\begin{aligned}
\mathcal{B}(W^+ \to e^+\nu_e)/\mathcal{B}(W^+ \to \mu^+\nu_\mu) &= 1.024 \pm 0.003 \pm 0.019, \\
\mathcal{B}(W^- \to e^-\overline{\nu}_e)/\mathcal{B}(W^- \to \mu^-\overline{\nu}_\mu) &= 1.014 \pm 0.004 \pm 0.022, \\
\mathcal{B}(W \to e\nu)/\mathcal{B}(W \to \mu\nu) \qquad &= 1.020 \pm 0.002 \pm 0.019,
\end{aligned}
$$

kus esimene mõõtemääramatus on statistilist laadi ning teine on süstemaatiline. Tulemust saab võrrelda eelnevate mõõtmistega [5, 11, 12, 16] Joonisel 7.3.

7.1 Tulevikuplaanid

Kuna antud doktoritöös esitatud $W \to e\nu$ mõõtmine on esmakordne LHCb eksperimendis, on ka mitmed edasiarendused võimalikud. Üks huvipakkuv idee on mõõta W bosoni massi sarnase analüüsi abil nagu on esitatud Joonisel 7.2. Ka ATLAS'e ja CMS'i kollaboratsioonid plaanivad läbi viia sarnase mõõtmise ning kolme tulemuse kombinatsioon võimaldaks saavutada hea täpsuse. LHCb mõõtmine oleks eriti väärtuslik lisand, kuna detektori mõõteruum on unikaalne ATLAS'ele ja CMS'ile.

W bosoni massi otsene mõõtmine on teoreetilisest seisukohast huvipakkuv seetõttu, et olemasolevad kaudsed piirangud massile, mis tulenevad elektronõrgast teooriast, on kaks korda rangemad, kui praegune parim otsene mõõtmine [162]. Seega on võimalik, et UF protsessid mõjutavad massi kõrgematel energiaskaaladel.

LHC kollaboratsioonid plaanivad jätkata oma uurimisprogramme järgneva kahekümne aasta jooksul. Samuti kavatsetakse tõsta kiire energiat ja prootonite kimpude tihedust ning parendada detektorite võimekust. LHC eksperimentalistid jätkavad SM'i mõõtmist uutel energiaskaaladel ning suurema täpsusega, kaasates töösse ka teoreetikuid. Selle töö lõplik eesmärk ei ole olemasolevate teooriate kinnitamine vaid nende ümberlükkamine, sest ainult sedasi jõuame paremale arusaamisele universumi seadustest.

Appendix A

Correlation coefficients

The correlation coefficients of the systematic uncertainties between bins of η^e for the W^+ (W^-) cross-sections are given in Table A.1 (A.2) while those between bins for W^+ and W^- are given in Table A.3. The LHC beam energy and luminosity uncertainties, which are fully correlated between cross-section measurements, are excluded.

Bin index	1	2	3	4	5	6	7	8
1	1.00							
2	0.93	1.00						
3	0.84	0.80	1.00					
4	0.95	0.94	0.84	1.00				
5	0.95	0.93	0.87	0.99	1.00			
6	0.74	0.79	0.70	0.86	0.85	1.00		
7	0.87	0.86	0.84	0.93	0.94	0.81	1.00	
8	0.82	0.82	0.75	0.88	0.92	0.78	0.86	1.00

Table A.1: Correlation coefficients of the systematic uncertainties for the differential W^+ cross-section measurement between bins of η^e.

Bin index	1	2	3	4	5	6	7	8
1	1.00							
2	0.99	1.00						
3	0.99	0.99	1.00					
4	0.98	0.99	0.99	1.00				
5	0.98	0.97	0.98	0.99	1.00			
6	0.72	0.75	0.72	0.77	0.76	1.00		
7	0.88	0.89	0.87	0.93	0.93	0.81	1.00	
8	0.84	0.82	0.82	0.87	0.90	0.83	0.95	1.00

Table A.2: Correlation coefficients of the systematic uncertainties for the differential W^- cross-section measurement between bins of η^e.

Bin index	1	2	3	4	5	6	7	8
1	0.94	0.95	0.91	0.90	0.86	0.47	0.77	0.67
2	0.87	0.85	0.85	0.81	0.79	0.36	0.65	0.59
3	0.91	0.90	0.93	0.89	0.90	0.56	0.74	0.74
4	0.93	0.92	0.90	0.88	0.87	0.45	0.81	0.71
5	0.95	0.93	0.92	0.90	0.91	0.53	0.85	0.81
6	0.61	0.64	0.65	0.62	0.60	0.67	0.69	0.60
7	0.84	0.82	0.81	0.81	0.83	0.53	0.86	0.82
8	0.84	0.79	0.81	0.79	0.84	0.53	0.81	0.84

Table A.3: Correlation coefficients of the systematic uncertainties for the differential W^+ and W^- cross-section measurements between bins of η^e. The horizontal bin indices label bins of η^e for electrons while vertical indices label bins for e^+.

Appendix B

Pseudoexperiments for fits to p_{T}^{e}

The variation in yields as a result of repeated pseudoexperiments is shown in Fig. B.1 for e^{-} and in Fig. B.2 for e^{+}.

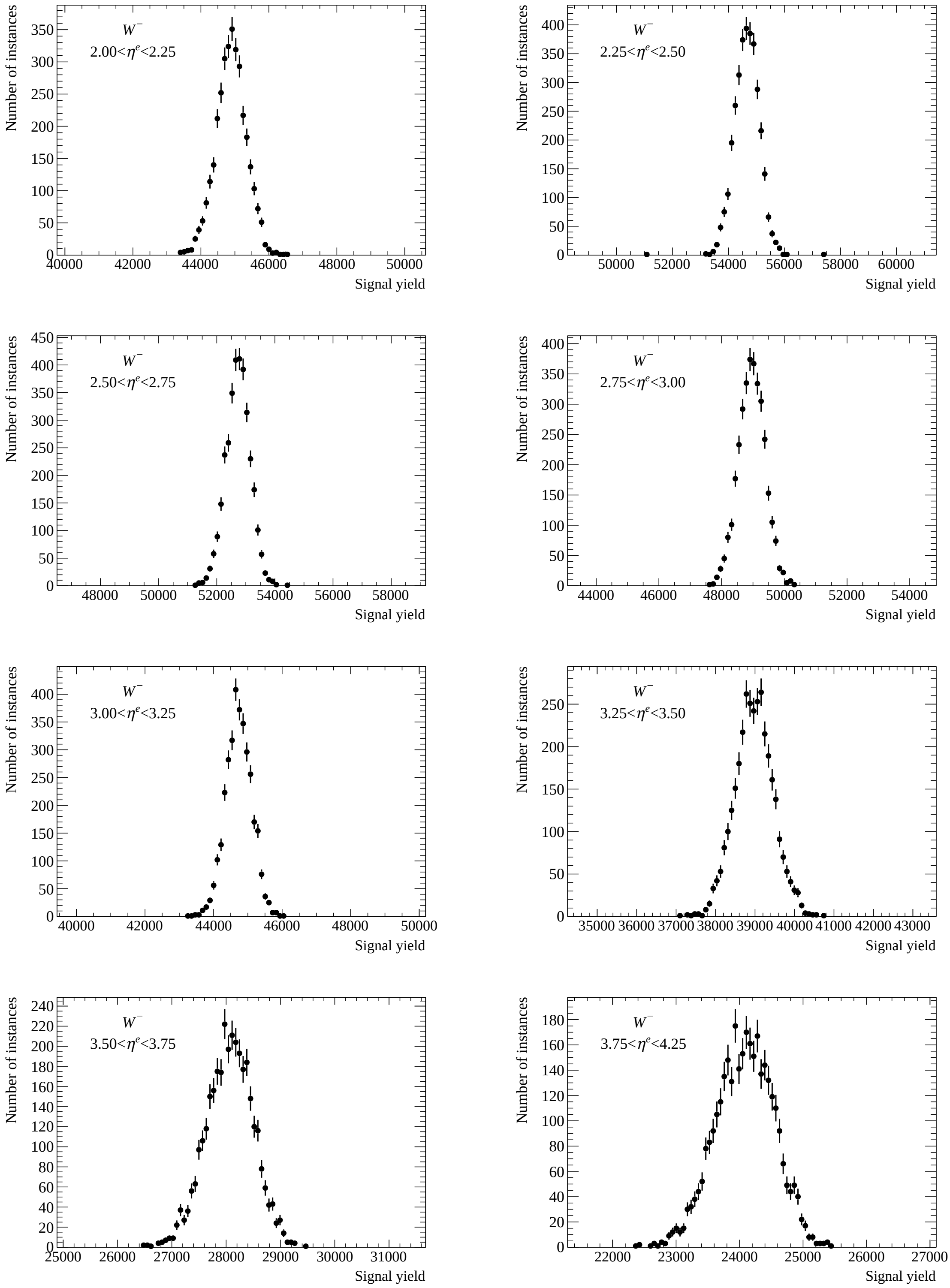


Figure B.1: The variation in yields as a result of repeated pseudoexperiments for e^-.

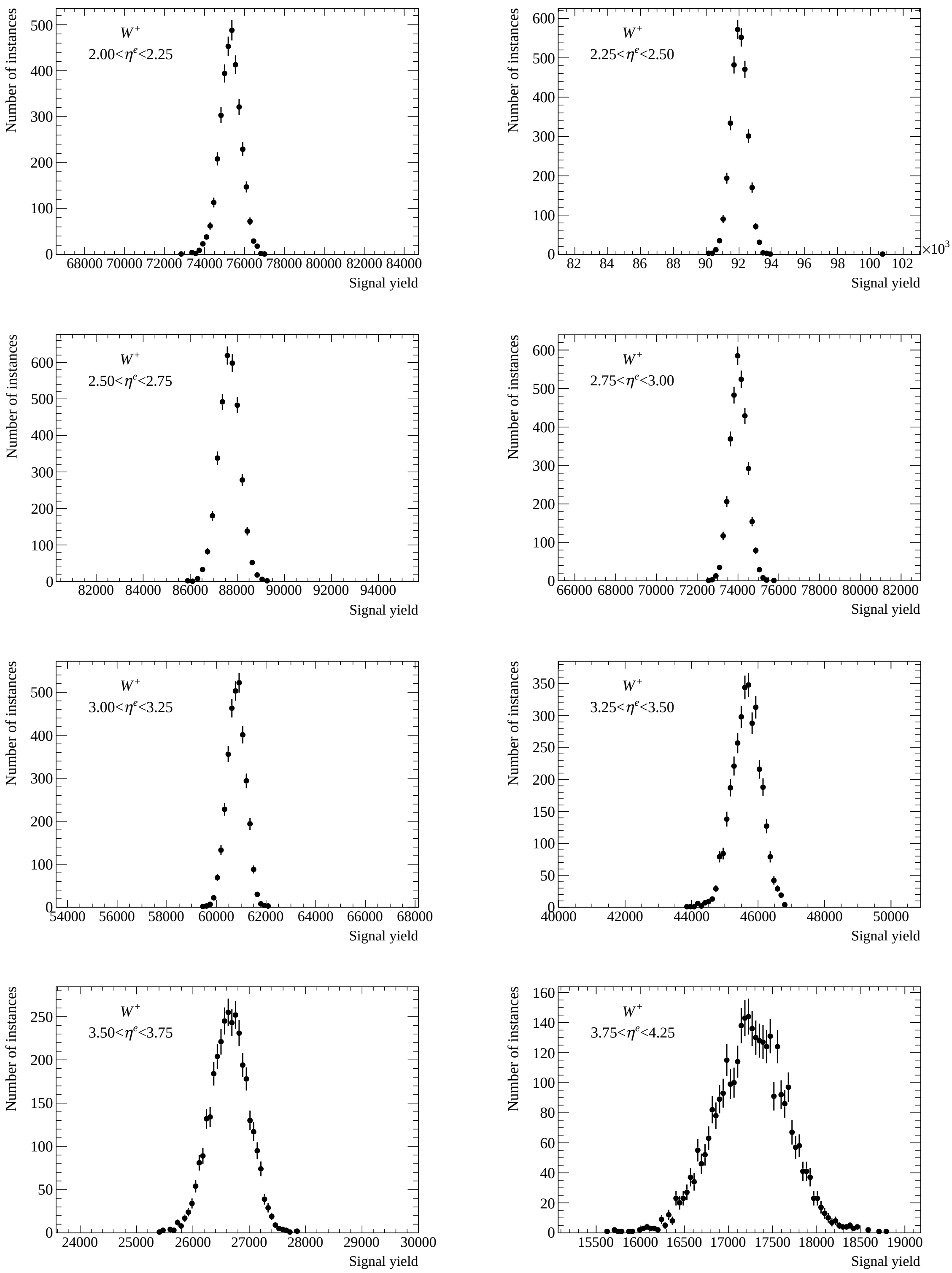


Figure B.2: The variation in yields as a result of repeated pseudoexperiments for e^+.

Colophon

This thesis was made in $\LaTeX\,2_\varepsilon$ using the "hepthesis" class [163].

Bibliography

[1] LHCb collaboration, R. Aaij *et al.*, *Measurement of forward $W \to e\nu$ production in pp collisions at $\sqrt{s} = 8$ TeV*, JHEP **10** (2016) 030, arXiv:1608.01484.

[2] LHCb collaboration, R. Aaij *et al.*, *Measurement of the forward Z boson production cross-section in pp collisions at $\sqrt{s} = 13$ TeV*, JHEP **09** (2016) 136, arXiv:1607.06495.

[3] M. Sirendi, *Modelling electroweak physics for the forward region*, MCnet-16-37 (2016) arXiv:1608.06906.

[4] J. Butterworth *et al.*, *PDF4LHC recommendations for LHC Run II*, J. Phys. **G43** (2016) 023001, arXiv:1510.03865.

[5] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38** (2014) 090001, and 2015 update.

[6] UA1 collaboration, G. Arnison *et al.*, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV*, Phys. Lett. **B122** (1983) 103.

[7] UA2 collaboration, M. Banner *et al.*, *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN $\bar{p}p$ Collider*, Phys. Lett. **B122** (1983) 476.

[8] UA1 collaboration, G. Arnison *et al.*, *Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV at the CERN SPS Collider*, Phys. Lett. **B126** (1983) 398.

[9] UA2 collaboration, P. Bagnaia *et al.*, *Evidence for $Z \to e^+e^-$ at the CERN $p\bar{p}$ Collider*, Phys. Lett. **B129** (1983) 130.

[10] DØ collaboration, B. Abbott *et al.*, *Extraction of the width of the W boson from measurements of $\sigma(p\bar{p} \to W + X) \times B(W \to e\nu)$ and $\sigma(p\bar{p} \to Z + X) \times B(Z \to ee)$ and their ratio*, Phys. Rev. **D61** (2000) 072001, arXiv:hep-ex/9906025.

[11] CDF collaboration, A. Abulencia *et al.*, *Measurements of inclusive W and Z cross sections*

in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV, J. Phys. **G34** (2007) 2457, `arXiv:hep-ex/0508029`.

[12] ALEPH, DELPHI, L3, OPAL collaborations, LEP Electroweak working group, S. Schael *et al.*, *Electroweak measurements in electron-positron collisions at W-boson-pair energies at LEP*, Phys. Rept. **532** (2013) 119, `arXiv:1302.3415`.

[13] CMS collaboration, V. Khachatryan *et al.*, *Measurement of the transverse momentum spectra of weak vector bosons produced in proton–proton collisions at $\sqrt{s} = 8$ TeV*, `arXiv:1606.05864`, submitted to JHEP.

[14] CMS collaboration, V. Khachatryan *et al.*, *Measurement of the differential cross section and charge asymmetry for inclusive $pp \to W + X$ production at $\sqrt{s} = 8$ TeV*, Eur. Phys. J. C (2016) , `arXiv:1603.01803`.

[15] CMS collaboration, S. Chatrchyan *et al.*, *Measurement of inclusive W and Z boson production cross sections in pp collisions at $\sqrt{s} = 8$ TeV*, Phys. Rev. Lett. **112** (2014) 191802, `arXiv:1402.0923`.

[16] ATLAS collaboration, G. Aad *et al.*, *Measurement of the inclusive $W^{\pm}$ and Z/γ^* cross sections in the electron and muon decay channels in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. **D85** (2012) 072004, `arXiv:1109.5141`.

[17] ATLAS collaboration, G. Aad *et al.*, *Measurement of the double-differential high-mass Drell-Yan cross section in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **08** (2016) 009, `arXiv:1606.01736`.

[18] ATLAS collaboration, G. Aad *et al.*, *Measurement of $W^{\pm}$ and Z-boson production cross sections in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. **B759** (2016) 601, `arXiv:1603.09222`.

[19] LHCb collaboration, R. Aaij *et al.*, *Measurement of the cross-section for $Z \to e^+e^-$ production in pp collisions at $\sqrt{s} = 7$ TeV*, JHEP **02** (2013) 106, `arXiv:1212.4620`.

[20] LHCb collaboration, R. Aaij *et al.*, *Measurement of forward $Z \to e^+e^-$ production in pp collisions at $\sqrt{s} = 8$ TeV*, JHEP **05** (2015) 109, `arXiv:1503.00963`.

[21] LHCb collaboration, R. Aaij *et al.*, *Measurement of the forward Z boson production cross-section in pp collisions at $\sqrt{s} = 7$ TeV*, JHEP **08** (2015) 039, `arXiv:1505.07024`.

[22] LHCb collaboration, R. Aaij *et al.*, *Measurement of forward W and Z boson production in pp collisions at $\sqrt{s} = 8$ TeV*, JHEP **01** (2016) 155, `arXiv:1511.08039`.

[23] LHCb collaboration, R. Aaij *et al.*, *A study of the Z production cross-section in pp collisions at $\sqrt{s} = 7$ TeV using tau final states*, JHEP **01** (2013) 111, `arXiv:1210.6289`.

[24] LHCb collaboration, R. Aaij *et al.*, *Measurement of the forward W boson production cross-section in pp collisions at $\sqrt{s} = 7$ TeV*, JHEP **12** (2014) 079, arXiv:1408.4354.

[25] LHCb collaboration, R. Aaij *et al.*, *Study of forward Z + jet production in pp collisions at $\sqrt{s} = 7$ TeV*, JHEP **01** (2014) 033, arXiv:1310.8197.

[26] LHCb collaboration, R. Aaij *et al.*, *Measurement of forward W and Z boson production in association with jets in proton-proton collisions at $\sqrt{s} = 8$ TeV*, JHEP **05** (2016) 131, arXiv:1605.00951.

[27] LHCb collaboration, R. Aaij *et al.*, *Measurement of the Z+b-jet cross-section in pp collisions at $\sqrt{s} = 7$ TeV in the forward region*, JHEP **01** (2015) 064, arXiv:1411.1264.

[28] LHCb collaboration, R. Aaij *et al.*, *Study of W boson production in association with beauty and charm*, Phys. Rev. **D92** (2015), no. 5 052001, arXiv:1505.04051.

[29] CMS collaboration, S. Chatrchyan *et al.*, *Measurement of the muon charge asymmetry in inclusive $pp \to W + X$ production at $\sqrt{s} = 7$ TeV and an improved determination of light parton distribution functions*, Phys. Rev. **D90** (2014), no. 3 032004, arXiv:1312.6283.

[30] LHCb collaboration, *Graphical comparison of the LHCb measurements of W and Z boson production with ATLAS and CMS*, LHCb-CONF-2013-005.

[31] D. Griffiths, *Introduction to elementary particles*, Wiley-VCH, 2008.

[32] M. Thomson, *Modern particle physics*, Cambridge University Press, 2013.

[33] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and collider physics*, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **8** (1996).

[34] ATLAS collaboration, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, arXiv:1207.7214.

[35] CMS collaboration, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, arXiv:1207.7235.

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

[37] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132.

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

[39] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585.

[40] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. **145** (1966) 1156.

[41] T. W. B. Kibble, *Symmetry breaking in non-Abelian gauge theories*, Phys. Rev. **155** (1967) 1554.

[42] S. Nussinov and R. Shrock, *Upper Limits on a Possible Gluon Mass*, Phys. Rev. **D82** (2010) 034031, `arXiv:1005.0850`.

[43] ATLAS and CMS collaborations, G. Aad *et al.*, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at* $\sqrt{s} = 7$ *and 8 TeV*, JHEP **08** (2016) 045, `arXiv:1606.02266`.

[44] S. D. Drell and T.-M. Yan, *Partons and their Applications at High-Energies*, Annals Phys. **66** (1971) 578.

[45] R. Hamberg, W. L. van Neerven, and T. Matsuura, *A complete calculation of the order* α_s^2 *correction to the Drell-Yan K-factor*, Nucl. Phys. **B359** (1991) 343, [Erratum: Nucl. Phys. B644, 403 (2002)].

[46] S. Catani *et al.*, *Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO*, Phys. Rev. Lett. **103** (2009) 082001, `arXiv:0903.2120`.

[47] K. Melnikov and F. Petriello, *Electroweak gauge boson production at hadron colliders through* $O(\alpha(s)^2)$, Phys. Rev. **D74** (2006) 114017, `arXiv:hep-ph/0609070`.

[48] Y. Li, A. von Manteuffel, R. M. Schabinger, and H. X. Zhu, *N^3LO Higgs boson and Drell-Yan production at threshold: The one-loop two-emission contribution*, Phys. Rev. **D90** (2014), no. 5 053006, `arXiv:1404.5839`.

[49] G. A. Ladinsky and C. P. Yuan, *The non-perturbative regime in QCD resummation for gauge boson production at hadron colliders*, Phys. Rev. **D50** (1994) 4239, `arXiv:hep-ph/9311341`.

[50] C. Balazs and C. P. Yuan, *Soft gluon effects on lepton pairs at hadron colliders*, Phys. Rev. **D56** (1997) 5558, `arXiv:hep-ph/9704258`.

[51] F. Landry, R. Brock, P. M. Nadolsky, and C. P. Yuan, *Tevatron Run-1 Z boson data and Collins-Soper-Sterman resummation formalism*, Phys. Rev. **D67** (2003) 073016, `arXiv:hep-ph/0212159`.

[52] J. D. Bjorken, *Current algebra at small distances*, Conf. Proc. **C670717** (1967) 55.

[53] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, Nucl. Phys. **B126** (1977) 298.

[54] 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, [Zh. Eksp. Teor. Fiz. 73, 1216 (1977)].

[55] V. N. Gribov and L. N. Lipatov, *Deep inelastic ep scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972) 438, [Yad. Fiz. 15, 781 (1972)].

[56] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189, `arXiv:0901.0002`.

[57] L. Barze *et al.*, *Neutral current Drell-Yan with combined QCD and electroweak corrections in the POWHEG BOX*, Eur. Phys. J. **C73** (2013), no. 6 2474, `arXiv:1302.4606`.

[58] P. Richardson *et al.*, *QCD parton showers and NLO EW corrections to Drell-Yan*, JHEP **06** (2012) 090, `arXiv:1011.5444`.

[59] Y. Li and F. Petriello, *Combining QCD and electroweak corrections to dilepton production in FEWZ*, Phys. Rev. **D86** (2012) 094034, `arXiv:1208.5967`.

[60] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, *Parton distributions and the LHC: W and Z production*, Eur. Phys. J. **C14** (2000) 133, `arXiv:hep-ph/9907231`.

[61] ATLAS collaboration, G. Aad *et al.*, *Determination of the strange quark density of the proton from ATLAS measurements of the $W \to \ell\nu$ and $Z \to \ell\ell$ cross sections*, Phys. Rev. Lett. **109** (2012) 012001, `arXiv:1203.4051`.

[62] NuTeV collaboration, G. P. Zeller *et al.*, *A Precise determination of electroweak parameters in neutrino nucleon scattering*, Phys. Rev. Lett. **88** (2002) 091802, `arXiv:hep-ex/0110059`, [Erratum: Phys. Rev. Lett.90,239902(2003)].

[63] Y. Ding, R.-G. Xu, and B.-Q. Ma, *Effect of asymmetric strange-antistrange sea to the NuTeV anomaly*, Phys. Lett. **B607** (2005) 101, `arXiv:hep-ph/0408292`.

[64] J. Alwall and G. Ingelman, *Strange quark asymmetry in the nucleon and the NuTeV anomaly*, Phys. Rev. **D70** (2004) 111505, `arXiv:hep-ph/0407364`.

[65] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, *High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO*, Phys. Rev. **D69** (2004) 094008, `arXiv:hep-ph/0312266`.

[66] T. Boettcher, P. Ilten, and M. Williams, *Direct probe of the intrinsic charm content of the proton*, Phys. Rev. **D93** (2016), no. 7 074008, arXiv:1512.06666.

[67] R. S. Thorne, A. D. Martin, W. J. Stirling, and G. Watt, *Parton Distributions and QCD at LHCb*, in *Proceedings, 16th International Workshop on Deep Inelastic Scattering and Related Subjects (DIS 2008): London, UK, April 7-11, 2008*, p. 30. arXiv:0808.1847. doi: 10.3360/dis.2008.30.

[68] S. Alekhin, J. Blumlein, and S. Moch, *The ABM parton distributions tuned to LHC data*, Phys. Rev. **D89** (2014), no. 5 054028, arXiv:1310.3059.

[69] S. Alioli *et al.*, *Precision studies of observables in $pp \to W \to l\nu$ and $pp \to \gamma/Z \to l^+l^-$ processes at the LHC*, arXiv:1606.02330.

[70] ATLAS collaboration, G. Aad *et al.*, *Measurement of the production cross section of jets in association with a Z boson in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **07** (2013) 032, arXiv:1304.7098.

[71] CMS collaboration, S. Chatrchyan *et al.*, *Event shapes and azimuthal correlations in Z+jets events in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. **B722** (2013) 238, arXiv:1301.1646.

[72] CMS collaboration, V. Khachatryan *et al.*, *Differential cross section measurements for the production of a W boson in association with jets in proton-proton collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. **B741** (2015) 12, arXiv:1406.7533.

[73] L. Lönnblad and S. Prestel, *Unitarising Matrix Element + Parton Shower merging*, JHEP **1302** (2013) 094, arXiv:1211.4827.

[74] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, *Longitudinally invariant K_t clustering algorithms for hadron hadron collisions*, Nucl. Phys. **B406** (1993) 187.

[75] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, *QCD matrix elements + parton showers*, JHEP **0111** (2001) 063, arXiv:hep-ph/0109231.

[76] L. Lönnblad, *Correcting the color dipole cascade model with fixed order matrix elements*, JHEP **0205** (2002) 046, arXiv:hep-ph/0112284.

[77] N. Lavesson and L. Lönnblad, *W+jets matrix elements and the dipole cascade*, JHEP **0507** (2005) 054, arXiv:hep-ph/0503293.

[78] N. Lavesson and L. Lönnblad, *Merging parton showers and matrix elements: Back to basics*, JHEP **0804** (2008) 085, arXiv:0712.2966.

[79] L. Lönnblad and S. Prestel, *Matching Tree-Level Matrix Elements with Interleaved*

Showers, JHEP **1203** (2012) 019, arXiv:1109.4829.

[80] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, *Matching matrix elements and shower evolution for top-quark production in hadronic collisions*, JHEP **0701** (2007) 013, arXiv:hep-ph/0611129.

[81] K. Hamilton, P. Nason, and G. Zanderighi, *MINLO: Multi-Scale Improved NLO*, JHEP **10** (2012) 155, arXiv:1206.3572.

[82] R. Frederix and S. Frixione, *Merging meets matching in MC@NLO*, JHEP **1212** (2012) 061, arXiv:1209.6215.

[83] L. Lönnblad and S. Prestel, *Merging Multi-leg NLO Matrix Elements with Parton Showers*, JHEP **1303** (2013) 166, arXiv:1211.7278.

[84] S. Höche, Y. Li, and S. Prestel, *Drell-Yan lepton pair production at NNLO QCD with parton showers*, Phys. Rev. **D91** (2015), no. 7 074015, arXiv:1405.3607.

[85] S. Höche, Y. Li, and S. Prestel, *Higgs boson production through gluon fusion at NNLO QCD with parton showers*, Phys. Rev. **D90** (2014), no. 5 054011, arXiv:1407.3773.

[86] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **0206** (2002) 029, arXiv:hep-ph/0204244.

[87] T. Sjöstrand *et al.*, *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv:1410.3012.

[88] J. Alwall *et al.*, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **1407** (2014) 079, arXiv:1405.0301.

[89] J. Alwall *et al.*, *A Standard format for Les Houches event files*, Comput. Phys. Commun. **176** (2007) 300, arXiv:hep-ph/0609017.

[90] J. Gao *et al.*, *CT10 next-to-next-to-leading order global analysis of QCD*, Phys. Rev. **D89** (2014) 033009, arXiv:1302.6246.

[91] P. M. Nadolsky *et al.*, *Implications of CTEQ global analysis for collider observables*, Phys. Rev. **D78** (2008) 013004, arXiv:0802.0007.

[92] A. Buckley *et al.*, *Rivet user manual*, Comput. Phys. Commun. **184** (2013) 2803, arXiv:1003.0694.

[93] A. Banfi *et al.*, *Optimisation of variables for studying dilepton transverse momentum distributions at hadron colliders*, Eur. Phys. J. **C71** (2011) 1600, arXiv:1009.1580.

[94] J. R. Christiansen and T. Sjöstrand, *Weak Gauge Boson Radiation in Parton Showers*, JHEP **1404** (2014) 115, `arXiv:1401.5238`.

[95] T. Sjöstrand and P. Z. Skands, *Transverse-momentum-ordered showers and interleaved multiple interactions*, Eur. Phys. J. **C39** (2005) 129, `arXiv:hep-ph/0408302`.

[96] ATLAS and CMS collaborations, M. Negrini, *Review of physics results using jet substructure techniques in LHC Run 1*, EPJ Web Conf. **90** (2015) 09001.

[97] P. Skands, S. Carrazza, and J. Rojo, *Tuning PYTHIA 8.1: the Monash 2013 Tune*, Eur. Phys. J. **C74** (2014), no. 8 3024, `arXiv:1404.5630`.

[98] T. Sjöstrand, Private communication.

[99] L. Evans and P. Bryant, *LHC Machine*, JINST (2008) S08001.

[100] LHC Study Group, T. S. Pettersson and P. Lefvre, *The Large Hadron Collider: conceptual design*, CERN-AC-95-05-LHC (1995), unpublished.

[101] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[102] Gargamelle collaboration, F. J. Hasert *et al.*, *Search for Elastic ν_μ Electron Scattering*, Phys. Lett. **B46** (1973) 121.

[103] Gargamelle collaboration, F. J. Hasert *et al.*, *Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment*, Phys. Lett. **B46** (1973) 138.

[104] Gargamelle collaboration, F. J. Hasert *et al.*, *Observation of Neutrino Like Interactions without Muon or Electron in the Gargamelle Neutrino Experiment*, Nucl. Phys. **B73** (1974) 1.

[105] G. Baur *et al.*, *Production of anti–hydrogen*, Phys. Lett. **B368** (1996) 251.

[106] ALPHA collaboration, G. B. Andresen *et al.*, *Search For Trapped Antihydrogen*, Phys. Lett. **B695** (2011) 95, `arXiv:1012.4110`.

[107] S. Dailler, *Urbanistic map of the LHC area*, AC Collection. Pictures from 1992 to 2002, 1997.

[108] R. M. Van der Eijk, *Track reconstruction in the LHCb experiment*, PhD thesis, Amsterdam University, Geneva, 2002.

[109] LHCb collaboration, P. Collins, *The LHCb Upgrade*, in *Flavor physics and CP violation. Proceedings, 9th International Conference, FPCP, Maale HaChamisha, Israel, May 23-27,*

2011. arXiv:1108.1403.

[110] ATLAS collaboration, G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[111] CMS collaboration, S. Chatrchyan *et al.*, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004.

[112] ALICE collaboration, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.

[113] LHCb collaboration, R. Aaij *et al*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.

[114] C. Elsasser, *The production angles of $b\bar{b}$ pairs at the LHC*, LHCb Speakers' Bureau, https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html (visited: 2016-08-19).

[115] LHCb collaboration, *LHCb VELO TDR: Vertex locator*, Tech. Rep. CERN-LHCC-2001-011, CERN, Geneva.

[116] A. Affolder *et al.*, *Radiation damage in the LHCb Vertex Locator*, JINST **8** (2013) P08002, arXiv:1302.5259.

[117] LHCb collaboration, *LHCb: Inner tracker technical design report*, Tech. Rep. CERN-LHCC-2002-029, CERN, Geneva.

[118] S. Kullback and R. A. Leibler, *On information and sufficiency*, Ann. Math. Statist. **22** (1951) 79.

[119] M. Needham, *Clone Track Identification using the Kullback-Liebler Distance*, Tech. Rep. LHCb-2008-002, CERN, Geneva.

[120] R. E. Kalman, *A new approach to linear filtering and prediction problems*, Transactions of the ASME–Journal of Basic Engineering **82** (1960), no. Series D 35.

[121] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. **A262** (1987) 444.

[122] LHCb collaboration, *LHCb: RICH technical design report*, Tech. Rep. CERN-LHCC-2000-037, CERN, Geneva.

[123] LHCb collaboration, *Trigger Schemes*, LHCb Speakers' Bureau, https://lhcb.web.cern.ch/lhcb/speakersbureau/html/TriggerScheme_110714.html (visited: 2016-08-30).

[124] LHCb collaboration, R. Aaij *et al.*, *The LHCb Trigger and its Performance in 2011*,

JINST **8** (2013) P04022, arXiv:1211.3055.

[125] LHCb HLT project, J. Albrecht, V. V. Gligorov, G. Raven, and S. Tolk, *Performance of the LHCb High Level Trigger in 2012*, J. Phys. Conf. Ser. **513** (2014) 012001, arXiv:1310.8544.

[126] M. Mangano, *Motivations and precision targets for an accurate luminosity determination at the LHC*, LHC Lumi Days: LHC Workshop on LHC Luminosity Calibration (2011).

[127] LHCb collaboration, R. Aaij *et al.*, *Precision luminosity measurements at LHCb*, JINST **9** (2014), no. 12 P12005, arXiv:1410.0149.

[128] LHCb collaboration, *The Gaudi Project*, 2016, http://proj-gaudi.web.cern.ch/proj-gaudi/.

[129] LHCb collaboration, *The Gauss Project*, 2016, http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/gauss/.

[130] LHCb collaboration, *The Boole Project*, 2016, http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/.

[131] LHCb collaboration, *The Brunel Project*, 2016, http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/.

[132] LHCb collaboration, *The Moore Project*, 2016, http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/moore/.

[133] LHCb collaboration, *The Moore Project*, 2016, http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/.

[134] LHC Experiments Committee, *Worldwide LHC Computing Grid*, 2016, http://wlcg-public.web.cern.ch/.

[135] A. Tsaregorodtsev *et al.*, *DIRAC: A community grid solution*, J. Phys. Conf. Ser. **119** (2008) 062048.

[136] K. Harrison *et al.*, *GANGA: a user-grid interface for ATLAS and LHCb*, CoRR **cs.SE/0306085** (2003).

[137] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820; T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

[138] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb*

simulation framework, J. Phys. Conf. Ser. **331** (2011) 032047.

[139] P. M. Nadolsky *et al.*, *Implications of CTEQ global analysis for collider observables*, Phys. Rev. **D78** (2008) 013004, `arXiv:0802.0007`.

[140] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270; Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[141] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

[142] A. Puig, *The LHCb trigger in 2011 and 2012*, Tech. Rep. LHCb-PUB-2014-046, CERN, Geneva.

[143] M. Bahr *et al.*, *Herwig++ physics and manual*, Eur. Phys. J. **C58** (2008) 639, `arXiv:0803.0883`.

[144] R. J. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, Comput. Phys. Commun. **77** (1993) 219.

[145] J. M. Campbell and R. K. Ellis, *MCFM for the Tevatron and the LHC*, Nucl. Phys. Proc. Suppl. **205-206** (2010) 10, `arXiv:1007.3492`.

[146] J.-H. Zhong *et al.*, *A program for the Bayesian Neural Network in the ROOT framework*, Comput. Phys. Commun. **182** (2011) 2655, `arXiv:1103.2854`.

[147] W. Barter *et al.*, *Jets reconstruction and performances at LHCb*, Tech. Rep. LHCb-INT-2012-015, CERN, Geneva.

[148] A. N. Kolmogorov, *Sulla determinazione empirica di una legge di distribuzione*, Giornale dellIstituto Italiano degli Attuari **4** (1933) 83.

[149] N. Smirnov, *Table for estimating the goodness of fit of empirical distributions*, Ann. Math. Statist. **19** (1948) 279.

[150] C. J. Clopper and E. S. Pearson, *The use of confidence or fiducial limits illustrated in the case of the binomial*, Biometrika **26** (1934), no. 4 404.

[151] J. Wenninger, *Energy calibration of the LHC beams at 4 TeV*, Tech. Rep. CERN-ATS-2013-040, CERN, Geneva.

[152] S. Catani *et al.*, *Vector boson production at hadron colliders: A fully exclusive QCD calculation at NNLO*, Phys. Rev. Lett. **103** (2009) 082001, `arXiv:0903.2120`.

[153] LHCb collaboration, R. Aaij *et al.*, *Precision luminosity measurements at LHCb*, JINST

9 (2014) P12005, arXiv:1410.0149.

[154] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, *FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order*, Comput. Phys. Commun. **182** (2011) 2388, arXiv:1011.3540.

[155] Y. Li and F. Petriello, *Combining QCD and electroweak corrections to dilepton production in FEWZ*, Phys. Rev. **D86** (2012) 094034, arXiv:1208.5967.

[156] S. Dulat *et al.*, *New parton distribution functions from a global analysis of quantum chromodynamics*, Phys. Rev. **D93** (2016) 033006, arXiv:1506.07443.

[157] ZEUS and H1 collaborations, A. M. Cooper-Sarkar, *PDF Fits at HERA*, PoS **EPS-HEP2011** (2011) 320, arXiv:1112.2107.

[158] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, Eur. Phys. J. **C75** (2015) 204, arXiv:1412.3989.

[159] NNPDF collaboration, R. D. Ball *et al.*, *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, arXiv:1410.8849.

[160] LHCb collaboration, R. Aaij *et al.*, *First observation of top quark production in the forward region*, Phys. Rev. Lett. **115** (2015), no. 11 112001, arXiv:1506.00903.

[161] S. J. Brodsky *et al.*, *A review of the intrinsic heavy quark content of the nucleon*, Adv. High Energy Phys. **2015** (2015) 231547, arXiv:1504.06287.

[162] G. Bozzi, L. Citelli, M. Vesterinen, and A. Vicini, *Prospects for improving the LHC W boson mass measurement with forward muons*, Eur. Phys. J. **C75** (2015), no. 12 601, arXiv:1508.06954.

[163] A. Buckley, hepthesis - a LaTeX document class for academic theses.

List of figures

List of tables