Vertex counting as a luminosity measure at ATLAS and determination of the electroweak Zjj production cross-section

A thesis submitted to the University of Manchester for the degree of
Doctor of Philosophy
in the Faculty of Engineering and Physical Sciences

Julia Mariana Iturbe Ponce

School of Physics and Astronomy

2016

Contents

Vertex counting as a luminosity measure at ATLAS and determination of the electroweak *Zjj* production cross-section

A thesis submitted to the University of Manchester for the degree of
Doctor of Philosophy in the Faculty of Engineering and Physical Sciences

Julia Mariana Iturbe Ponce
School of Physics and Astronomy, 2016

Abstract

This thesis presents two analyses of data recorded by the ATLAS detector during proton-proton collisions at the LHC. The first is the implementation of a vertex counting algorithm to measure the luminosity recorded by ATLAS during collisions at a centre-of-mass energy of $\sqrt{s} = 8\,\text{TeV}$ in 2012. This comprises a Monte Carlo closure test for validation of the method and its corrections, the calibration of the method using the van der Meer scans performed in 2012 and the application of the method to physics runs. It also includes tests of the internal and external consistency of the algorithm and the potential to use this algorithm to measure the luminosity of data collected during proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$.

The second analysis is the measurement of the inclusive and purely electroweak production of dijets in association with a Z boson, performed using the $3.2\,\text{fb}^{-1}$ of data collected during collisions at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ in 2015. Cross-section measurements are presented for five fiducial regions, each of which has a different sensitivity to the electroweak component of the *Zjj* production. Data and Monte Carlo predictions are compared and found to be in reasonable agreement for most cases. The electroweak *Zjj* production cross-section is then extracted in a fiducial region where this contribution is enhanced. This measurement is also in good agreement with the Monte Carlo prediction. These first 13 TeV measurements will set the scene for studies of weak boson fusion, both within the Standard Model and in new phenomena searches, which will become even more important in Run 2 and the future of the LHC due to the electroweak sector not being as constrained yet, compared to the strong sector, and due to the larger enhancements as a result of a higher $\sqrt{s}$, where electroweak physics can be most easily extracted.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright statement

1. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

2. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

3. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

4. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy[1], in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations[2] and in The University's policy on Presentation of Theses.

---

[1] see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487

[2] see http://www.manchester.ac.uk/library/aboutus/regulations

Acknowledgements

I want to thank CONACyT for giving me the opportunity and financial support to come all the way to Manchester to study in this very prestigious University. I also want to thank Terry Wyatt, first, for agreeing to be my supervisor before I even got here and without knowing much about me, and second, for continuing to be my supervisor once he did get to know me. Also, I want to thank Darren Price for his support and supervision during the last year of my PhD, injecting that much needed enthusiasm for physics towards the end of what turned out to be a very hard, but rewarding journey.

I would also like to thank everyone in the Manchester HEP group because they were incredibly welcoming when I first arrived and very patient when I couldn't understand their British accent. For the next four years they continued to be friendly and helpful, and patient when I still struggled to understand what they were saying. They make Manchester a great place to work.

I want to thank all the friends I've made since I came to Manchester because without them I would have probably missed Mexico and my family too much and gone back, but you made it worth staying (and also, the PhD).

Lastly, I want to thank my friends and family back home for their love and support, especially my mum and dad, Carmen and Angel, and my sisters Angela and Cristalia, who have always believed in me, maybe too much and definitely more than I do. Everything I do is in the hope that someday your faith in me will be justified.

Finalmente, quiero agradecer a mi familia y amigos en México, por su apoyo y cariño, especialmente a mi mamá y papá, Carmen y Ángel, y a mis hermanas Ángela y Cristalia, que siempre han creído en mí, tal vez demasiado y definitivamente más de lo que yo creo que mí. Todo lo que hago es con la esperanza de que algún día su fe en mí sea justificada.

Chapter 1

Introduction

This thesis presents the implementation of a vertex counting algorithm to measure the luminosity recorded by the ATLAS detector during the LHC proton-proton (*pp*) collisions at a centre-of-mass energy of $\sqrt{s} = 8\,\text{TeV}$ in 2012. This particular algorithm was used before to measure the luminosity recorded by the ATLAS detector during 2011 and the work presented here builds upon that analysis. When the algorithm was used in 2011 it was found to suffer from "pile-up" effects, which means that the measured luminosity did not increase linearly with pile-up, where pile-up refers to the average number of *pp* interactions in the same bunch crossing. The pile-up seen in 2012 was almost double the pile-up seen during 2011 and the work presented here illustrates how this vertex counting algorithm had to be modified to cope with the characteristics of the 2012 data.

This thesis also presents the first measurement of the cross-section of the electroweak production of a Z boson in association with two jets (Zjj) at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$. This was performed using the $3.2\,\text{fb}^{-1}$ of data collected by the ATLAS detector during the LHC *pp* collisions in 2015. The first measurement of the cross-section of said events at $\sqrt{s} = 8\,\text{TeV}$ was performed by ATLAS with the $20.3\,\text{fb}^{-1}$ of data collected during Run-1 of the LHC and the work presented here builds upon that analysis. The change in ATLAS dataset formats between Run-1 and Run-2 of the LHC required a complete re-write of the analysis code in order to perform this measurement and the parton luminosity enhancement allowed the 13 TeV measurement presented here to be statistically comparable to the 8 TeV one. These measurements are shown to be consistent.

The structure of this thesis is as follows. Chapter 2 describes the basics of the Standard Model of particle physics, the theoretical framework in which the analyses presented in this thesis are set. Chapter 3 introduces the LHC and the ATLAS detector, the experimental apparatus used to produce and collect the data for the analyses presented in this thesis. This chapter describes the ATLAS design in detail and the different steps followed in order to obtain the data to be analysed. Chapter 4 introduces the concept of luminosity and describes the detectors and algorithms that ATLAS employs to measure it. Chapter 5 focuses on one of these algorithms, the vertex counting algorithm. This chapter describes in detail the implementation of this algorithm to measure the luminosity delivered to ATLAS during 2012. It also discusses the pile-up effects from which this algorithm suffers and the corrections developed to cope with them. The consistency of the algorithm is tested with a Monte Carlo closure test and then the algorithm is calibrated using the 2012 van der Meer scans. Finally, the vertex counting algorithm is used to measure the luminosity of 2012 physics data and the results are tested for internal consistency, and also compared to those obtained by other luminosity detectors and algorithms. Finally, Chapter 6 describes the method followed to measure the fiducial cross-section of the inclusive and electroweak production of a Z boson and two jets using data collected by ATLAS during 2015. The measurements are compared to the theory prediction provided by the SHERPA, MADGRAPH and ALPGEN generators and also to a similar analysis performed with data collected during pp collisions at a centre-of-mass-energy of $\sqrt{s} = 8\,\text{TeV}$.

Unless otherwise stated, this thesis uses *natural units*, the standard in high energy particle physics. In these units, the speed of light, c, and the reduced Planck's constant, $\hbar$, are set equal to unity ($c = \hbar = 1$) and energies are expressed in electron-volts, eV.

Chapter 2

The Standard Model

Elementary particle physics is the study of the basic constituents of matter and how these interact with one another. The best description available today of these elements and their interactions is the Standard Model (SM). The SM describes all of the known elementary particles (shown in Figure 2.1) and three of the four fundamental forces of nature, namely, the electromagnetic, the strong and the weak forces. It makes use of the theories of Quantum Electrodynamics (QED) and Quantum Chromodynamics (QCD) to describe the electromagnetic and strong interactions, respectively. At high energies, the electromagnetic and weak forces may be unified and described by the Glashow-Weinberg-Salam (GWS) theory of electroweak processes. Gravity is the only force the SM cannot yet describe but this force is too weak to play any significant role in ordinary high energy physics processes.

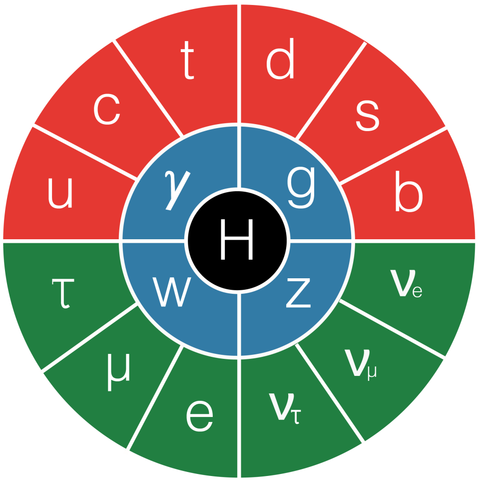


Figure 2.1: The fundamental particles of the SM. The outer ring contains the *fermions*, the particles that make up all matter. The inner ring contains the *bosons*, the force mediators and in the centre is the *Higgs* boson. Adapted from the diagram in [1].

Since the SM was proposed in the 1970s, the majority of its predictions have been verified by extremely precise measurements and continue to do so as new data is acquired. One of the biggest successes of the model, theoretically, was the unification of the electromagnetic and weak forces into the electroweak force and the generation of mass through spontaneous symmetry breaking (SSB). Experimentally, it's most recent success came with the observation of a Higgs boson. Nevertheless, the SM still has some limitations. As previously mentioned, the SM describes only three of the four forces in nature, excluding gravity. Also, it predicts a smaller matter-antimatter asymmetry than that observed in the universe and it does not include any particle that could account for the large amount of dark matter known to exist in the universe. It is clear then that the SM is not the complete picture for elementary particle physics but its many successes suggest that future, more complete models will likely be extensions to it.

This chapter gives an overview of the SM, paying special attention to the electroweak force since this is an important aspect of an analysis later presented in this thesis.

2.1 Constituents of Matter

Fermions are particles with spin $\frac{1}{2}$ and they make up all matter. They can be further classified into *leptons* (bottom half of outer ring in Figure 2.1) and *quarks* (top half of outer ring in Figure 2.1). There are six leptons in the SM. The *electron* (e^-), the *muon* (μ^-) and the *tau* (τ^-) all have charge $Q = -1$ (in units of electron charge, e) and similar properties, except for their masses, shown in Table 2.1. Each of these charged leptons has an associated neutral particle ($Q = 0$) called a *neutrino* (ν), therefore the three neutral leptons are called the electron neutrino (ν_e), the muon neutrino (ν_μ) and the tau neutrino (ν_τ). Each fermion has an anti-particle with almost the same properties as its fellow particle but with opposite charge. For the case of neutrinos, there are ongoing experiments probing whether they have a distinct anti-particle or they are their own anti-particle [2].

There are six quarks in the SM: the *up* (u), *charm* (c), and *top* (t) quarks have charge $Q = +{}^2/_3$, while the *down* (d), *strange* (s), and *bottom* (b) quarks have charge $Q = -{}^1/_3$. In addition to electric charge, quarks also carry *colour* charge. Quarks are never seen individually in nature but only in colourless bound states with integer electric charge, known as *hadrons*. Hadrons consisting of three quarks (or anti-quarks) are known as *baryons* (or *anti-baryons*), while hadrons consisting of a quark and an anti-quark are known as *mesons*. The most famous baryons are the proton and the neutron, found in the atomic nucleus. The fermions are often classified in three generations depending on their masses. These generations and some more properties of the fermions are shown in Table 2.1.

Generation	Particle		Electric Charge [e]	Mass [MeV]
1	electron	e^-	-1	0.5110
	electron neutrino	ν_e	0	$< 2 \times 10^{-6}$
	up quark	u	$+^2/_3$	$2.3^{+0.7}_{-0.5}$
	down quark	d	$-^1/_3$	$4.8^{+0.5}_{-0.3}$
2	muon	μ^-	-1	105.7
	muon neutrino	ν_μ	0	$< 2 \times 10^{-6}$
	charm quark	c	$+^2/_3$	1275 ± 25
	strange quark	s	$-^1/_3$	95 ± 5
3	tau	τ^-	-1	1776.82 ± 0.16
	tau neutrino	ν_τ	0	$< 2 \times 10^{-6}$
	top quark	t	$+^2/_3$	$173210 \pm 510 \pm 710$
	bottom quark	b	$-^1/_3$	4180 ± 30

Table 2.1: Summary of the SM fermions and their properties. Values taken from Ref. [3].

2.2 Forces and their mediators

Formally, the SM is a quantum field theory with a local $SU(3) \times SU(2) \times U(1)$ symmetry, where each symmetry corresponds roughly to the strong, weak and electromagnetic forces, respectively. Each symmetry gives rise to a spin 1 boson (inner ring in Figure 2.1) which couples to particles with the associated symmetry charge, hence, bosons are known as the forces mediators.

The *gluon (g)* is the mediator of the strong interactions occurring between particles carrying colour charge, i.e. quarks and gluons. The strong force is so-called because it is the strongest of the forces in nature. It is responsible for binding the quarks together to form hadrons and also, on a larger scale, for binding protons and neutrons together inside atomic nuclei. The gluon is a massless and electrically neutral particle but it does carry colour charge. The *photon* (γ) is the mediator of the electromagnetic (EM) interactions occurring between charged particles, i.e.

charged leptons and quarks. The photon is a massless particle with zero electric and colour charge. The W and Z bosons mediate the weak interactions that can occur between all types of fermions, making it the only force by which neutrinos can interact. The Z boson is electrically and colour neutral, like the photon, but it has mass; it mediates the neutral weak interactions. The W boson also has mass and is colour neutral but carries electric charge, which can be positive (W^+) or negative (W^-). It mediates the charged weak interactions. The final piece of the SM is the *Higgs* boson (centre of diagram in Figure 2.1), the only spin 0 particle in the SM and crucial in the proof of electroweak unification and subsequent symmetry breaking. An overview of the SM bosons can be seen in Table 2.2.

Particle		Electric Charge [e]	Mass [GeV]	Spin
Photon	γ	0	0	1
Z		0	91.1876 ± 0.0021	1
$W^\pm$		± 1	80.385 ± 0.015	1
Gluon	g	0	0	1
Higgs	H	0	125.7 ± 0.4	0

Table 2.2: Summary of the SM bosons and their properties. Values taken from Ref. [3].

2.3 The Electroweak Force

The electromagnetic and weak forces have been unified into a single theory in the Glashow-Weinberg-Salam (GWS) model [4, 5]. This theory incorporates local gauge invariance and SSB.

The electroweak Lagrangian derived from the $SU(2) \times U(1)$ symmetry group can be expressed as:

$$\mathcal{L} = \mathcal{L}_{\text{bosons}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{fermions}} + \mathcal{L}_{\text{Yukawa}}, \tag{2.1}$$

where the different terms correspond to the gauge boson kinetic and self-interaction terms, the Higgs field kinetic and potential terms which generate the gauge boson masses and gauge couplings to the Higgs boson, the fermion kinetic term which is responsible for the fermion interactions with the gauge bosons and lastly, the Yukawa term which generates the fermion masses and their coupling to the Higgs boson. The exact form of these terms will now be discussed.

In the original proposal by GWS, the electroweak Lagrangian contains four massless gauge bosons, W_i (i = 1, 2, 3) and B, associated with the gauge groups $SU(2)$ and $U(1)$, respectively. Their kinetic term is given by:

$$\mathcal{L}_{\text{bosons}} = -\frac{1}{4} W^i_{\mu\nu} W^{i\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \tag{2.2}$$

where the gauge field strength can be written as:

$$X^i_{\mu\nu} = \partial_\mu X^i_\nu - \partial_\nu X^i_\mu - g f^{ijk} X^j_\mu X^k_\nu, \tag{2.3}$$

where g is the coupling constant and f^{ijk} are the structure constants of the group considered. The third term in Equation 2.3 generates gauge boson self-interactions and it is therefore present for all non-Abelian groups, in this case $SU(2)$, but vanishes for Abelian groups, in this case $U(1)$. The exact form of the $SU(2)$ coupling constant, g, will be defined later and the $SU(2)$ structure constants, f^{ijk}, are simply equal to ϵ_{ijk}, the fully antisymmetric tensor.

Another two important properties in the electroweak theory are hypercharge (Y) and weak isospin (I). They are related via the electromagnetic charge (Q) as:

$$Y = 2(Q - I). \tag{2.4}$$

Left-handed fermions have $I = \pm\frac{1}{2}$, the sign depending on Q, and right-handed fermions have $I = 0$. Only left-handed fermions transform under the $SU(2)$ symmetry.

SSB and the Higgs field allow the transition from these four massless gauge bosons to the four gauge bosons that have been observed experimentally, three of which are not massless. The Higgs field is defined as a doublet of complex scalar fields invariant under $SU(2)$ transformations with hypercharge $Y = 1$ and weak isospin $I = 1$ and can be expressed as:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \phi^\dagger = \begin{pmatrix} \phi^- & \overline{\phi^0} \end{pmatrix}. \tag{2.5}$$

The corresponding Lagrangian can be written as:

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi^\dagger \phi), \tag{2.6}$$

where the covariant derivative that would make the system invariant under local gauge transformations is:

$$D_\mu = \partial_\mu - igW^i_\mu T^i - i\frac{Y}{2}g'B_\mu, \tag{2.7}$$

where the $SU(2)$ generators are $T^i = \sigma^i/2$ and σ^i are the Pauli matrices.

The Higgs potential can be written as:

$$V(\phi^\dagger\phi) = \lambda(\phi^\dagger\phi)^2 - \mu^2\phi^\dagger\phi \tag{2.8}$$

which clearly depends on the choice of the λ and μ parameters. If $\mu^2 < 0$, the potential is purely positive with only a minimum at the origin. However, if $\mu^2 > 0$ then $\phi = 0$ is an unstable maximum. In the latter case, the degenerate minima of the potential energy are described by a circle of radius

$$v = \sqrt{\frac{\mu^2}{\lambda}}. \tag{2.9}$$

Choosing a specific minimum, for example at

$$\phi_{min} = \frac{1}{\sqrt{2}}\phi\begin{pmatrix}0\\ v\end{pmatrix}, \tag{2.10}$$

gives the vacuum a preferred direction in weak isospin space – this is known as the SSB mechanism. Inserting the example value of ϕ_{min} from Equation 2.10 into Equation 2.6 and remembering that $\partial_\mu\phi_{min} = 0$, then $\mathcal{L}_{\text{Higgs}}$ gives:

$$\begin{aligned}\left| -i\left(\frac{g}{2}W^i_\mu\sigma^i + \frac{g'}{2}B_\mu\right)\phi\right|^2 &= \frac{1}{8}\left|\begin{pmatrix} gW^3_\mu + g'B_\mu & g(W^1_\mu - iW^2_\mu)\\ g(W^1_\mu + iW^2_\mu) & -gW^3_\mu + g'B_\mu\end{pmatrix}\begin{pmatrix}0\\ v\end{pmatrix}\right|^2\\ &= m^2W^+_\mu W^{-\mu} + \frac{1}{2}m^2_Z Z_\mu Z^\mu\end{aligned} \tag{2.11}$$

where the last equality arises using the following relations:

$$W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu), \tag{2.12}$$

$$Z^0_\mu = \frac{1}{\sqrt{g^2 + g'^2}}(gW^3_\mu - g'B_\mu), \tag{2.13}$$

$$A_\mu = \frac{1}{\sqrt{g^2 + g'^2}}(g'W^3_\mu + gB_\mu). \tag{2.14}$$

The symmetry of this group is now spontaneously broken by the Higgs mechanism and the original four massless bosons W_i ($i = 1, 2, 3$) and B associated to the $SU(2)$ and $U(1)$ gauge groups, respectively, are recombined to give rise to the four well known gauge bosons of the electroweak theory. The three weak gauge bosons that have now acquired mass through SSB are the W^+, W^- and Z bosons, while the photon (γ) remains massless.

The photon vector field, A_μ, is orthogonal to Z^0_μ and this can be seen by expressing the new fields after SSB as a rotation of the old fields before SSB by a certain angle, known as the weak mixing angle θ_W, which acts as a change of basis:

$$\begin{pmatrix} Z^0 \\ A \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W^3 \\ B \end{pmatrix}. \tag{2.15}$$

The coupling constants are:

$$g = \frac{e}{\sin\theta_W} \quad \text{and} \quad g' = \frac{e}{\cos\theta_W}, \tag{2.16}$$

and the masses are:

$$m_W = \frac{1}{2} g v \quad \text{and} \quad m_Z = \frac{1}{2} v \sqrt{g^2 + g'^2}. \tag{2.17}$$

It has been experimentally observed that W bosons only couple to left-handed fermions. This implies that right-handed fermions must occur in singlets even if left-handed fermions occur in doublets in the $SU(2)$ group. The lepton and quark fields can therefore be defined as:

$$\psi_L = \gamma_L \begin{pmatrix} \nu_L \\ \ell^- \end{pmatrix}, \gamma_L \begin{pmatrix} u \\ d' \end{pmatrix}, \gamma_L \begin{pmatrix} c \\ s' \end{pmatrix}, \gamma_L \begin{pmatrix} t \\ b' \end{pmatrix} \quad \text{and} \quad \psi_R = \gamma_R \ell^-, \gamma_R q, \tag{2.18}$$

where $\gamma_{R,L} = \frac{1}{2}(1 \pm \gamma^5)$ are the projection operators which give the right- or left-handed components of a fermion spinor, ℓ represents the charged leptons e, μ and τ, and the prime on the quark doublets is related to the CKM matrix explained below. Therefore, the Lagrangian that describes the fermion interactions with the gauge bosons can be written as:

$$\mathcal{L}_{fermions} = \overline{\psi^L} i \not{D} \psi_L + \overline{\psi_R} i \not{D} \psi_R, \tag{2.19}$$

using the appropriate hypercharge, either Y_L or Y_R, in the covariant derivative. Since the isospin of right-handed fields is 0, the second term in Equation 2.7 vanishes for these. The boson-fermions coupling strengths are:

$$g_{\gamma ff} = eQ, \tag{2.20}$$

$$g_{Wff} = \frac{g}{2\sqrt{2}}(1-\gamma^5), \tag{2.21}$$

$$g_{Zff} = \frac{g}{2\cos\theta_W}(c_V - c_A\gamma^5), \tag{2.22}$$

where $c_V = I - 2Q\sin{\theta_W}^2$ and $c_A = I$. From this it follows that the weak interaction does not conserve parity as there are vector-like terms ($\propto \gamma^\mu$) added to axial-like terms ($\propto \gamma^\mu\gamma^5$).

Lastly, the Yukawa term creates the fermion masses after the Higgs acquires a vacuum expectation value. Ordinary mass terms cannot simply be introduced into the Lagrangian because the left- and right-handed components of the fermion fields have different quantum numbers and would violate gauge invariance. However, the hypercharge difference between the left- and right-handed fermions of a certain flavour will always be ±1, which is the hypercharge of the Higgs field. Therefore, the gauge invariant Lagrangian can be written as:

$$\mathcal{L}_{Yukawa} = -g_e\bar{E}_L\phi g_R - f_d\bar{Q}_L\phi d_R - g_u\bar{Q}_L\phi^c u_R + h.c., \tag{2.23}$$

where the g's are the Yukawa couplings, $\phi^c = i\sigma^2\phi^\dagger$ and the Q_L are the quark doublets defined in Equation 2.18. A Yukawa interaction of the form $g_f\bar{\psi}_f\phi\psi_f$ generates a fermion mass through SSB equal to:

$$m_f = g_f v/\sqrt{2}. \tag{2.24}$$

In Equation 2.18 the down-type quarks were denoted with a prime in the quark doublets. This was done because, since there is more than one quark generation, there can be extra coupling terms which mix the generations. Therefore, the probability of a u quark coupling to a single d type quark does not make sense and instead one is concerned with the probability of coupling to the physical eigenstates that are a linear superposition of d, s and b quarks with coefficients determined by the 3×3 Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix, defined as:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.25}$$

2.4 Gauge Boson Couplings

Since the $SU(2) \times U(1)$ electroweak gauge group is non-Abelian, it requires the existence of triple (TGC) and quadratic gauge couplings, which are vertices in which three or four electroweak gauge bosons couple to each other, just like the gluon self-coupling vertices in QCD. As was shown in the previous section, one massless and three massive electroweak gauge bosons remain after SSB. However, due to certain symmetries and conservation laws, not every combination of these bosons is allowed. Focusing only on TGCs, since charge has to be conserved at a vertex, four out of the ten combinations are not allowed. Additionally, photons cannot couple to each other because the $U(1)$ group is Abelian, hence $\gamma\gamma\gamma$ and $Z\gamma\gamma$ vertices are also not allowed. ZZZ and $ZZ\gamma$ vertices are also prohibited in the SM and this can be seen by examining Equation 2.3. Here, the fully anti-symmetric ϵ^{ijk} only allows a TGC vertex of the form $W^1W^2W^3$. Recalling Equations 2.12-2.14, W^1 and W^2 make up the $W^{\pm}$ bosons and W^3 contributes to forming the Z and γ bosons. Hence, the only possible TGCs are W^+W^-Z an $W^+W^-\gamma$.

A TGC Lagrangian can be constructed by choosing the terms which give rise to the interaction of three fields in the expansion of Equation 2.2 and reordering the fields so that they are expressed in terms of the fields after SSB, i.e. W^+, W^-, Z and A. This yields the expression:

$$\begin{aligned} \mathcal{L}_{TGC} = & \, ig_{WW\gamma}[A_\mu(W^-_\nu W^{+\mu\nu} - W^- W^+_\nu) + W^-_\mu W^+_\nu A^{\mu\nu}] \\ & + ig_{WWZ}[Z_\mu(W^-_\nu W^{+\mu\nu} - W^- W^+_\nu) + W^-_\mu W^+_\nu Z^{\mu\nu}], \end{aligned} \tag{2.26}$$

where the field tensor is $X_{\mu\nu} = \partial_\mu X_\nu - \partial_\nu X_\mu$ and the coupling strengths are:

$$\begin{aligned} g_{WW\gamma} &= g \sin\theta_W = e \\ g_{WWZ} &= g \cos\theta_W = e \cot\theta_W. \end{aligned} \tag{2.27}$$

2.4.1 aTGCs

As it was mentioned before, the SM has been verified to astounding precision by different experiments but it is not the complete picture. There are many proposals for different theories that would complement the SM and these are known as beyond the standard model (BSM) theories. As an example which is related to one of the analysis in this thesis, a few proposals for BSM physics would change the observed rate at which bosons couple to one another and these are known as anomalous triple gauge couplings (aTGCs) [6, 7].

These theories can be tested in a model independent way by measuring the rate at which bosons couple to each other and comparing to the SM prediction. Significant deviation from the SM prediction would allow the aTGC values to be measured, whereas precise agreement with the SM prediction places upper limits on the aTGCs and can help exclude BSM models that predict a higher value for the coupling.

Chapter 3

Experimental Apparatus

3.1 The LHC

The Large Hadron Collider (LHC) [8] is a particle accelerator located at CERN, the European Centre for Nuclear Research, in Geneva. The tunnel in which it resides was originally constructed for the Large Electron-Positron (LEP) collider [9]. The LHC is located 100 m underground and its 26.7 km circumference spans parts of France and Switzerland. The LHC holds the record for highest centre-of-mass energy ($\sqrt{s}$), colliding particles at 13 TeV during its 2015 running.

The LHC was designed to accelerate protons to $\sqrt{s} = 14$ TeV but this goal has not yet been reached. During Run 1, which took place from 2009 to 2013, the LHC collided protons first and briefly at a centre-of-mass energy of 900 GeV, in November 2009 and then at 7 TeV, during 2010 and 2011, delivering in total 5.6 fb^{-1} of data to ATLAS. The centre-of-mass energy was then increased to 8 TeV in 2012 and during this running the LHC delivered 22.8 fb^{-1} of data. Starting in February 2013, the LHC was shut down for repairs and upgrades and in 2015 it was turned back on, delivering the first proton-proton collisions at a centre-of-mass energy of 13 TeV and delivering 3.2 fb^{-1} of data. This was the first phase of Run 2 of the LHC.

In order to reach such high centre-of-mass energies at the LHC, the accelerator complex at CERN is used (Figure 3.1). This is a chain of accelerators each of which boosts the energy of a beam of particles and then injects the beam into the next machine in the sequence. The LHC is the last element in this chain. The source protons come from hydrogen atoms which are stripped of their electrons using an electric field. These protons are accelerated to an energy of 50 MeV by Linac 2, the first accelerator in the chain. Then the beam is injected into the Proton Synchrotron Booster (PSB), which accelerates the protons to 1.4 GeV. Afterwards, the Proton Synchrotron (PS) increases the energy of the beam to 25 GeV and then the Super

Proton Synchrotron (SPS) accelerates the beam to 450 GeV. Finally, the proton beams are transferred to the two beam pipes at the LHC, where one beam circulates clockwise and the other beam anticlockwise. After a few minutes of circulating in the LHC, the proton beams reach their maximum energy of 6.5 TeV.

The LHC employs radio-frequency (RF) cavities to accelerate the bunches of protons and it uses several different types of magnets to force the beams into the desired path. The primary bending of the beams is done by superconducting dipole magnets which can generate magnetic fields of up to 8.4 T. Quadrupole magnets are used to squeeze the beams either horizontally or vertically throughout the LHC ring. At the collision points, steering dipoles direct the beams into collision and a system of three quadrupoles, called an inner triplet, is used on each side of the interaction point to focus the beams even more, reducing its size from $\approx 1\,\text{mm}$ to $\approx 10\,\mu\text{m}$ across.

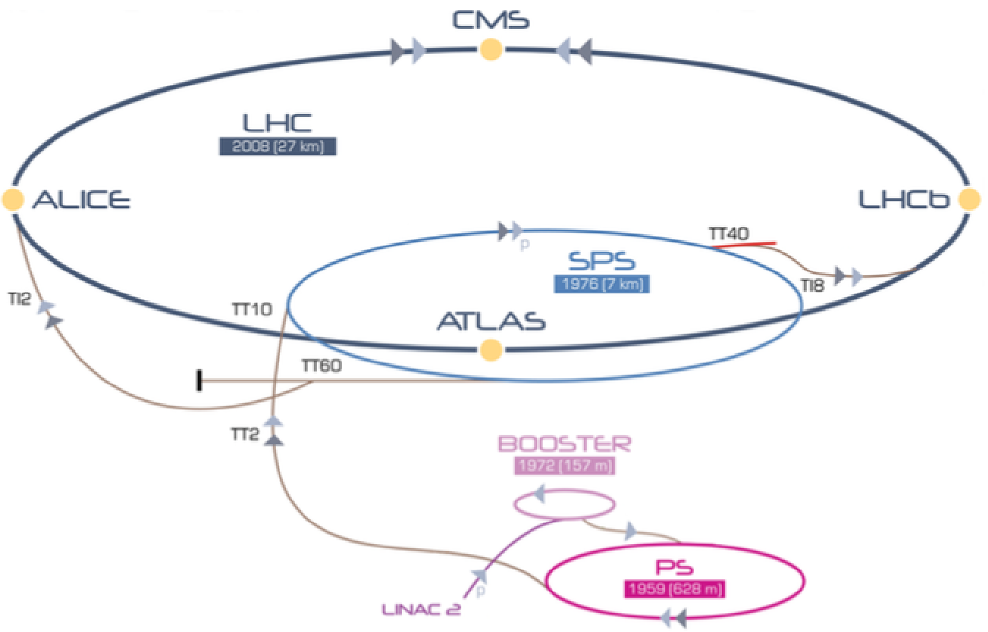


Figure 3.1: The CERN accelerator complex used to inject protons to the LHC. Taken from Ref. [10].

The circulating proton beams are brought to collision at four interaction points (IP) at the LHC, which correspond to the location of four particle detectors: ATLAS, CMS, ALICE and LHCb. LHCb (Large Hadron Collider beauty) [11] is an asymmetric forward detector that specialises in B-hadron physics to study CP-violation and matter-antimatter asymmetry. ALICE (A Large Ion Collider Experiment) [12] focuses on heavy-ion collisions (lead-lead and proton-lead) and it is designed to study the physics of strongly interacting matter at extreme energy densities, where the phase of matter quark-gluon plasma forms. CMS (Compact Muon Solenoid) [13] and ATLAS (A Toroidal LHC ApparatuS) [14] are general purpose detectors (GPD) which means they have a broad physics programme. They are designed to make precision measurements to test and constrain the Standard Model but also to look for new physics beyond the Standard Model (BSM). One of the main goals of these detectors was to find the long-sought after Higgs boson; this goal was achieved in July 2012 when both experiments announced they had discovered a Higgs-like particle [15, 16].

3.2 The ATLAS detector

ATLAS is one of the two GPDs at the LHC and the biggest particle detector ever built. It is 46 m long, approximately cylindrical with a diameter of 25 m, and it weighs 7000 tonnes. It is located at IP 1 of the LHC, in a cavern 100 m underground near the main CERN site, close to the commune of Meyrin, in Switzerland.

The coordinate system used by ATLAS is right-handed with the origin set to be the interaction point. The z-axis is defined along the beam line with the positive direction being anti-clockwise if looking at the LHC ring from above. The positive direction of the x-axis is pointing towards the centre of the LHC ring and the positive direction of the y-axis points upwards, away from the centre of the Earth. The ATLAS detector is designed to be symmetric with respect to the plane at $z = 0$, with the side on the positive z-axis direction being called side A and the other side, side C. The x-y plane, referred to as the transverse plane, is where the variables transverse momentum, p_{T}, transverse energy, E_{T} and missing transverse momentum, $E_{\mathrm{T}}^{\mathrm{miss}}$, are defined, where $E_{\mathrm{T}}^{\mathrm{miss}}$ is the momentum carried by particles which do not interact with the detector, such as neutrinos.

Because of the geometry of the detector, a polar coordinate system is preferred when describing the trajectories of particles in the detector from the interaction point. For this system, the azimuthal angle, ϕ, is measured around the beam axis while the polar angle, θ, is measured from the beam axis. A useful variable to measure is the *rapidity*, y, defined as:

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \tag{3.1}$$

where E is the energy of the particle and p_z its momentum along the z-direction. This is a helpful variable since differences in rapidity are invariant with respect to Lorentz boosts along the z-axis. Most of the particles observed in ATLAS have energy considerably larger than their mass, therefore they can be approximated as massless and their energy can be equated to their momentum, simplifying the rapidity equation and yielding the definition of pseudo-rapidity, η:

$$\eta = -\ln \tan \frac{\theta}{2}. \tag{3.2}$$

A diagram of the ATLAS detector is shown in Figure 3.2. The ATLAS design can be divided into three parts: a central region called "the barrel", which is a cylinder providing coverage in the region $|\eta| < 1.4$, and two end-caps, one at either side of the barrel, covering the range $1.4 < |\eta| < 4.9$. Each of these regions can be further split into several sub-detectors. At the core of ATLAS is the Inner Detector (ID) which enables measurements of the momentum of charged particles and accurate reconstruction of vertices produced by primary proton-proton collisions, as well as those that arise from the decay of long-lived particles. Surrounding the ID are the electromagnetic and hadronic calorimeters, used to determine the energy of all charged and neutral particles, except for neutrinos. The outermost sub-detector is the Muon Spectrometer (MS), which enables precision measurements of the momentum of muons.

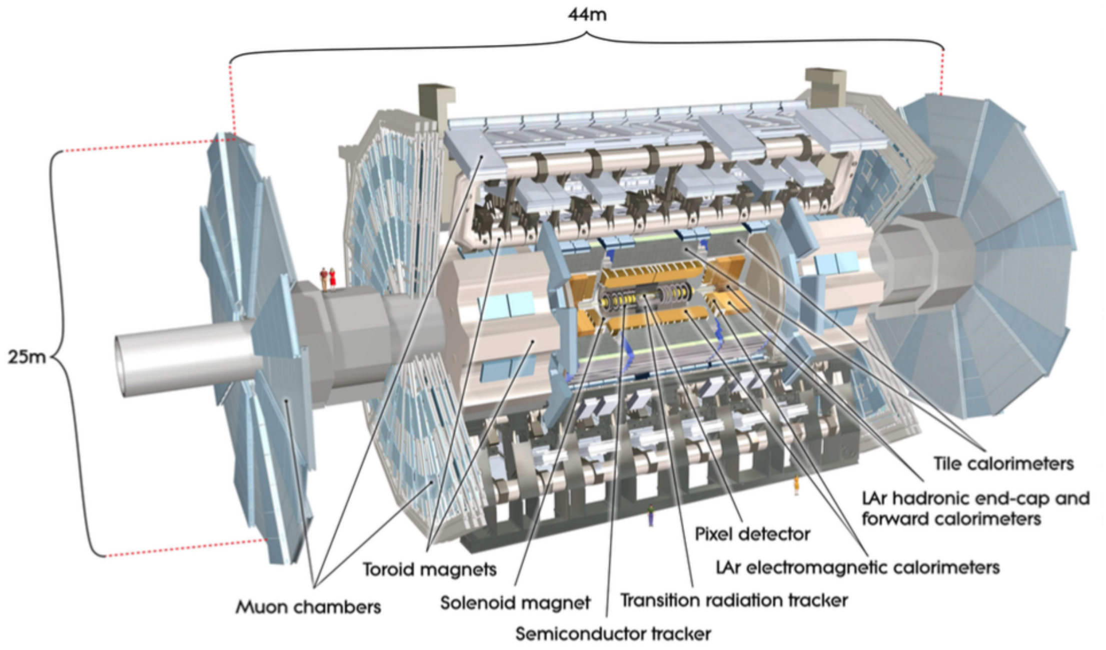


Figure 3.2: Diagram of the ATLAS detector, taken from Ref. [14]. The most important sub-detectors have been highlighted.

ATLAS uses both solenoidal (in the barrel) and toroidal (in barrel and end caps) magnet systems. The solenoid is aligned along the beam axis and provides a 2 T magnetic field. It completely surrounds the ID but has been designed to ensure that the material thickness in front of the calorimeters is as low as possible. Charged particles travelling in the solenoidal field have their trajectory curved in the ϕ-direction. The barrel and end-cap toroids produce 0.5 and 1 T magnetic fields, respectively. The toroid magnets enable measurements of the momentum of muons and cause the tracks produced by them to be curved in the η-direction. Measurements of the momentum of all charged particles are performed by measuring the curvature of the tracks produced by the particles as they traverse the detector subject to these magnetic fields.

3.2.1 The Inner Detector

The ATLAS inner detector is the closest detector to the interaction point, surrounding the beam pipe. It is located inside a solenoid magnet that provides a 2 T axial-symmetric field and it is used to measure the trajectories and momentum of charged particles in the region $|\eta| < 2.5$. The ID system constituents, shown in Figure 3.3, are a silicon pixel detector, a silicon microstrip detector (SCT) and the straw tubes of the transition radiation tracker (TRT).

Between Run 1 and Run 2, ATLAS went through a series of upgrades which included improvements to the ID. In order to have better track and vertex reconstruction performance at the higher luminosities expected during Run 2 and to mitigate the impact of radiation damage to the innermost layer of the pixel detector, a fourth layer was added to the ID, the insertable B-layer (IBL).

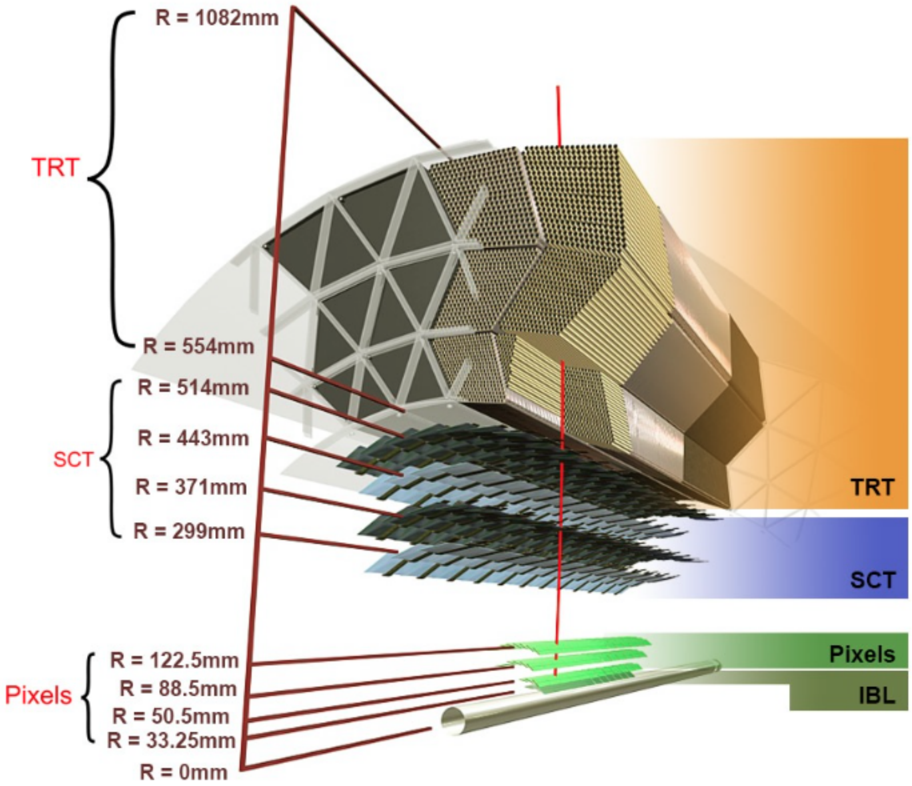


Figure 3.3: Sketch of the ATLAS inner detector showing all its components, including the new insertable B-layer (IBL). The distances to the interaction point are also shown. Taken from Ref. [17].

The precision tracking detectors, the pixel and SCT sensors, are arranged in concentric cylinders around the beam axis in the barrel region, and on disks perpendicular to the beam axis in the end-cap region. The pixel detectors are arranged in a way that typically three pixel layers are crossed by each track. The pixel sensor sizes in R-ϕ × z are 50×250 µm^2 in the IBL and 50×400 µm^2 or 50×600 µm^2 in the rest of the pixel layers. The tracking precision in the barrel is ∼10 µm in R-ϕ and ∼ 115 µm in z, while in the disks it is ∼10 µm in R-ϕ and ∼115 µm in R. The pixel detector has approximately 80.4 million readout channels and provides the highest granularity, achieved around the vertex region. Unfortunately, because of the cost of these sensors, the entire ATLAS detector cannot be constructed using only these and therefore the SCT and TRT detectors are used as a complement.

In the case of the SCT, eight strip layers (four space points) are crossed by each track. This detector uses small-angle (40 mrad) stereo strips in the barrel region, with one set of strips in each layer parallel to the beam direction, measuring R-ϕ. Each strip consists of two 6.4 cm long daisy-chained sensors with a strip pitch of 8 µm. In the end-cap region, the detectors have a set of strips running radially and a set of stereo strips at an angle of 40 mrad. The mean pitch of the strips is also approximately 80 µm. The intrinsic accuracies per module in the barrel are ∼17 µm in R-ϕ and ∼580 µm in z, while in the disks are ∼17 µm in R-ϕ and ∼580 µm in R. The total number of readout channels in the SCT is approximately 6.3 million.

The TRT provides a large number of hits, typically 36 per track, using its 4 mm diameter straw tubes. This enables track reconstruction up to $|\eta| = 2.0$. The TRT provides only R-ϕ information with an intrinsic accuracy of ∼130 µm per straw. In the barrel region, the straws are parallel to the beam axis and are 144 cm long, with their wires divided into two halves, approximately at $\eta = 0$. In the end-cap region, the 37 cm long straws are arranged radially in wheels. The total number of TRT readout channels is approximately 351,000.

3.2.2 Calorimeters

ATLAS contains a number of calorimeters that measure the energy of incident particles through absorption and they cover the wide range $|\eta| < 4.9$. Over the η region covering the ID, the fine granularity of the electromagnetic (EM) calorimeter is ideal for precision measurements of electrons and photons. For the rest of the calorimeter, the granularity is coarser but sufficient to satisfy the physics requirements for jet reconstruction and $E_{\mathrm{T}}^{\mathrm{miss}}$ measurements. An overview of ATLAS calorimeters is shown in Figure 3.4.

The next detector in a particle's path after the ID is the electromagnetic calorimeter, the main purpose of which is to measure the energy of electrons and photons. The EM calorimeter design alternates layers of lead absorbers and kapton electrodes organised in an accordion-like shape and immersed in liquid argon (LAr). This design is the same for both the barrel region and the end-caps, with each housed in their own cryostat kept at a temperature of 87 K. The regions covered by the EM calorimeter are $|\eta| < 1.475$ in the barrel region, $1.375 < |\eta| < 2.5$ in the outer end-cap wheel and $2.5 < |\eta| < 3.2$ in the inner end-cap wheel. Also, the accordion geometry of the EM calorimeter provides complete ϕ symmetry without any cracks.

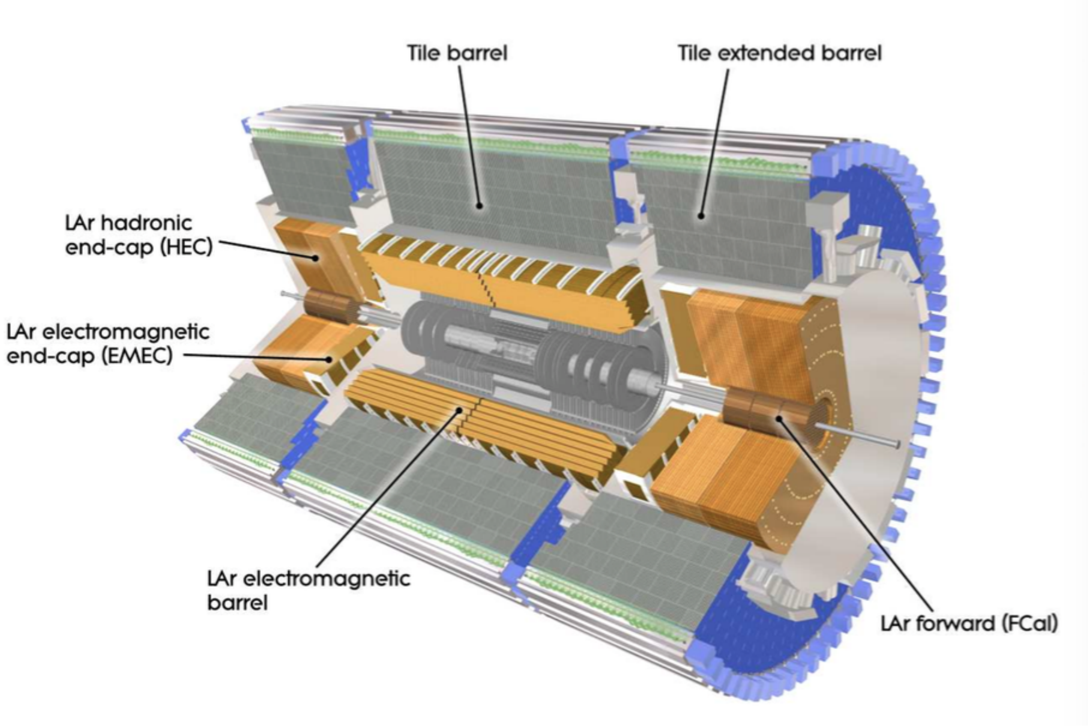


Figure 3.4: Cut-away view of the ATLAS calorimeter system. Taken from Ref. [14].

When electrons pass through the layers of lead in the EM calorimeter they emit bremsstrahlung photons, which in turn can produce electron-positron pairs. This results in showers of electromagnetic particles which ionise the active LAr layers and the electrodes collect the charge. The signal produced is proportional to the energy of the particle. The conversion factor is obtained using test-beam measurements and the behaviour of well-understood decays such as $Z \to e^+e^-$.

The depth of the EM calorimeter is an important design consideration since it must contain the electromagnetic showers of electrons and photons. The EM calorimeter is more than 22 radiation lengths thick in the barrel and more than 24 thick in the end-cap regions. In the range devoted to precision measurements, $|\eta| < 2.5$, the EM calorimeter has three active layers in depth and two layers in the range $2.5 < |\eta| < 3.2$. Most of the energy of an electromagnetic shower is absorbed by the second layer. A diagram of a barrel module indicating the granularity in η and ϕ of each of the layers is shown in Figure 3.5. The energy resolution in the range $1.37 < |\eta| < 1.52$ is worse due to the transition between the barrel and end-cap cryostats. This region is therefore not used for electron or photon identification.

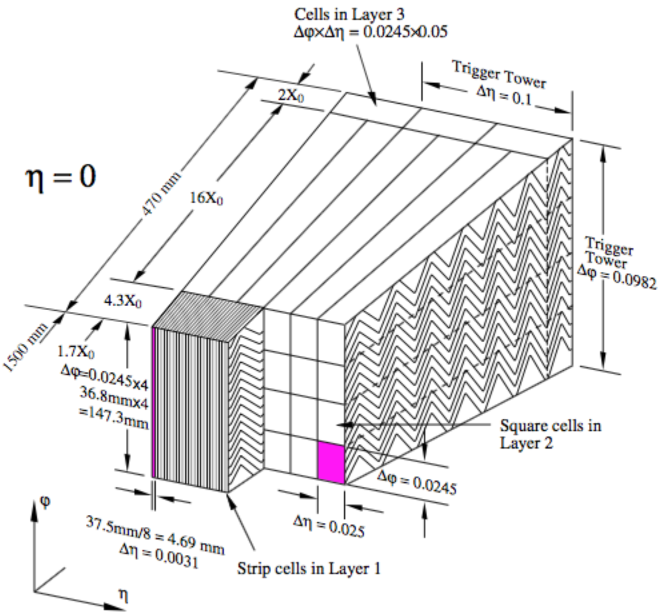


Figure 3.5: A diagram of a barrel EM calorimeter module indicating the granularity in η and ϕ in each of the three layers. Taken from Ref. [14].

The hadronic calorimeters are used to measure the energy of baryons and mesons. Directly outside the EM calorimeter is the tile calorimeter (TileCal), a sampling calorimeter that uses steel as the absorber and scintillating tiles as the active material. The tile calorimeter is composed of a barrel that covers the region $|\eta| < 1.0$ and two extended barrels which cover the region $0.8 < |\eta| < 1.7$. Each of these sections is divided azimuthally into 64 modules. The ultraviolet light produced by the scintillating tiles is converted to visible light by wavelength-shifting fibres which transmit the signal to photomultiplier tubes (PMTs). Radially, the TileCal extends from an inner radius of 2.28 m to an outer radius of 4.25 m. This thickness combined with the wide η coverage enables precise measurements of the $E_{\mathrm{T}}^{\mathrm{miss}}$.

Located directly behind the end-cap EM calorimeter and sharing the same LAr cryostats, is the Hadronic End-cap Calorimeter (HEC). The HEC covers the region $1.5 < |\eta| < 3.5$ and uses copper as the absorbing material and LAr as the sampling material. The HEC consists of two independent wheels per end-cap, each of which are divided into 32 wedge-shaped modules. Finally, the forward calorimeters (FCal) extend the coverage of the calorimeters to $3.1 < |\eta| < 4.9$. Each FCal is made up of three modules: an EM module that uses copper as its main absorber material, and two hadronic modules which use tungsten. Again LAr is used as the active material in all three modules and they all make use of the same cryostat systems as the other end-cap calorimeters, which reduces gaps in the coverage.

3.2.3 Muon System

The next and final part of the ATLAS detector, working outwards from the IP, is the muon system, the components of which are shown in Figure 3.6. This is designed to detect any particles that were not absorbed by the calorimeters, which are mostly muons. Muons are about 200 times heavier than electrons, therefore they loose energy to bremsstrahlung radiation much less than electrons do, making it possible for them to travel through the calorimeters without loosing much of their energy.

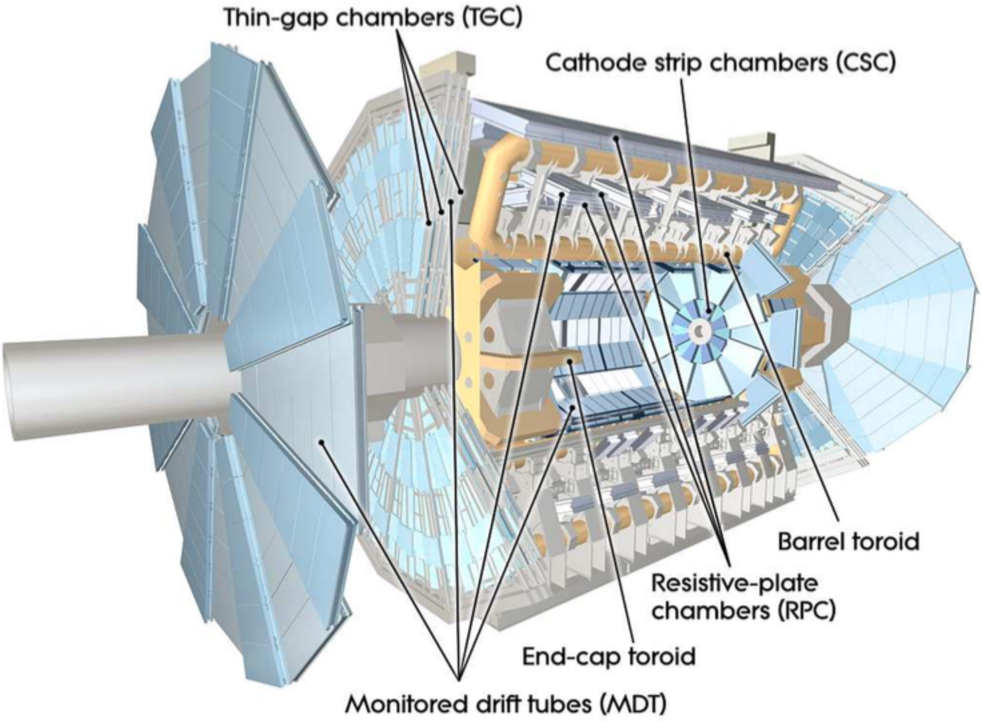


Figure 3.6: Cut-away view of the ATLAS muon system. Taken from Ref. [14].

The ATLAS muon system is composed of the Muon Spectrometer (MS) and the toroidal magnet system. The MS is used for both triggering and for precision measurements. These functions are performed using both the barrel and end-caps but triggering on muons is done in the region $|\eta| < 2.4$ while precision tracking on muons is done in the region $|\eta| < 2.7$. Precision measurements are performed using Monitored Drift Tubes (MDT), which are proportional drift tubes that cover the range $|\eta| < 2.7$ and provide tracking to a precision of approximately 80 µm, and Cathode Strip Chambers (CSC), which are multi-wire proportional chambers with cathode strips and anode wires that provide R-ϕ coordinates for tracking in the forward region, $2.0 < |\eta| < 2.7$, to an accuracy of approximately 60 µm.

Faster but coarser tracking information for use in the trigger system (section

3.2.4) is provided in the barrel region by Resistive Plate Chambers (RPCs), which provide position measurements for $|\eta| < 1.05$ when a muon ionises the gas held between two charged resistive plates, and in the end-cap regions by the Thin Gap Chambers (TGCs), which are multi-wire proportional chambers that provide measurements in the region $1.05 < |\eta| < 2.7$.

Surrounding the MS is the toroidal magnet system, which produces a magnetic field of between 0.5 and 1 T. This magnetic field causes the muons to curve in the y-z plane and allows the charge of the muon to be identified for muons with transverse momentum up to approximately 3 TeV.

3.2.4 Trigger System and data acquisition

It is impossible to store the output of every proton-proton collision produced at ATLAS since the technology to read out and write such large amounts of data at the necessary speed is not yet available. Additionally, the majority of the events produced do not contain objects of interest for physics analysis and physics searches, such as hard jets or high-p_T leptons. Because of this, ATLAS employs a trigger system designed to quickly decide whether an event is of interest for physics analysis or not. This trigger system has three steps: the Level-1 (L1) trigger, the Level-2 (L2) trigger and the event filter (EF). The last two are known collectively as the High Level Trigger (HLT).

The L1 trigger is a hardware base trigger which reduced the event rate to 70 kHz during Run 1 and to 100 kHz during Run 2 [18, 19]. This was possible because of upgrades to the L1 hardware and improvements to the detector readouts during the LHC shut down in 2013-2015. The L1 trigger makes the decision to keep an event within 2.5 µs. It uses the reduced granularity information from a subset of the detectors, such as the RPC and TGC sections of the MS. Using these, it can check whether muons with p_T above a certain threshold are present in the event. It also uses parts of the EM and hadronic calorimeters to place requirements on each event, such as the number and energy of electrons, photons, jets and τ-leptons, the amount of $E_\mathrm{T}^\mathrm{miss}$ and E_T.

If the L1 trigger finds a physics object, it records its position and a region-of-interest (RoI) containing the object, and passes this information on to the L2 trigger system. This trigger is software-based and further reduced the event rate to approximately 3.5 kHz during Run 1, making the decision to keep the event within 40 ms. This is possible because the L2 algorithms use the full granularity of all detector systems contained within the η-ϕ RoI passed by the L1 trigger. The L2 trigger uses tracks from the ID and more precise calorimeter and muon spectrometer measurements, all of which provides better p_{T}-resolution and isolation information.

The last part of the trigger system is the EF. This trigger reduces the event rate to 400 Hz during Run 1, making the decision whether to keep an event within 4 s. It reconstructs the objects of interest using algorithms similar to those used when reconstructing events offline. If an event passes the EF, then it is written to permanent storage at a rate of approximately 300 MB/s.

One of the upgrades to the ATLAS trigger system for Run 2 was the merging of the different farms used for the L2 and EF triggers, into one farm which runs unified HLT processes. This increased the event output rate of the HLT from 400 Hz to 1 kHz.

The naming of the triggers usually follows the format: LEVEL_(object)(p_{T})_(ExtraID). For example, an event passing the trigger `L1MU15` must contain a muon identified at L1, with $p_{\mathrm{T}} > 15\,\mathrm{GeV}$. An event passing the trigger `HLT_e60_lhmedium` must contain an electron identified at HLT, with $p_{\mathrm{T}} > 60\,\mathrm{GeV}$ and satisfying the "lhmedium" identification criteria, a collection of cuts on the cluster, track and combined cluster-track quantities used to identify an electron [20].

The huge amounts of data collected by ATLAS are organised into periods and runs, for easy management. A run begins once the Data Acquisition (DAQ) infrastructure, detectors and other sub-systems are configured correctly, and once the conditions of the beam provided by the LHC are stable. A run finishes either cleanly, when the instantaneous luminosity decreases below a certain threshold, or is aborted when a problem occurs, for example if the LHC beams are lost. A period is defined as a succession of DAQ runs with similar running conditions.

Once data has been collected and organised, a series of further data quality checks are performed. If all the triggers and detector systems during a certain run were working optimally and the data in this run is deemed of sufficient quality, then this run is added to the Good Runs List (GRL). Only those runs which appear in a GRL can be used for physics analysis.

Chapter 4

Luminosity

4.1 Luminosity Formalism

The luminosity $\mathscr{L}$ of a particle collider quantifies its ability to produce useful interactions and it can be expressed as:

$$\mathscr{L} = \frac{R}{\sigma}, \tag{4.1}$$

where R is the rate of events produced per unit of time for a process with cross-section σ. From this equation it can be seen that the luminosity units are cm^{-2} s^{-1}.

For proton-proton colliders with bunched beams, such as the LHC, $\mathscr{L}$ is equal to the sum of the luminosity of each bunch crossing, $\mathscr{L}_b$. This bunch luminosity can be expressed as:

$$\mathscr{L}_b = \frac{\mu f_r}{\sigma_{inel}}, \tag{4.2}$$

where μ is the average number of inelastic collisions per bunch crossing, also known as the pile-up parameter, f_r is the machine frequency and σ_{inel} is the pp inelastic cross-section. Each bunch crossing has an identifier named BCID.

The absolute luminosity can be determined from Equation 4.1 by counting the rate of a process for which the associated cross-section is well known, such as a Z boson decaying to two muons. However, the most common methodology is to determine the luminosity by measuring the *observed* interaction rate per bunch crossing, μ_{vis}, for a particular detector and algorithm. Equation 4.2 can be expressed in terms of μ_{vis} as:

$$\mathscr{L}_b = \frac{\mu_{vis} f_r}{\sigma_{vis}}, \tag{4.3}$$

where $\sigma_{vis} = \epsilon\sigma_{inel}$ is the total inelastic cross-section multiplied by the efficiency, ϵ, of the particular detector and algorithm. Therefore, to calibrate the luminosity scale of a particular detector and algorithm, σ_{vis} needs to be determined. This is done during dedicated beam-separation scans, called van der Meer scans, in which the luminosity is calculated from direct measurements of the beam parameters. This will be explained in more detail in the following section.

4.2 Luminosity Calibration: van der Meer Scans

To be able to use the measured interaction rate, μ_{vis}, as an *absolute* luminosity monitor, each detector and algorithm needs to be calibrated by determining its visible cross-section, σ_{vis}. The preferred technique to achieve this is the van der Meer method which employs dedicated scans, known as van der Meer scans (vdM scans), where the delivered luminosity can be inferred at one point in time from measurable parameters of the colliding beams. The known luminosity delivered in a vdM scan can be compared to the visible interaction rate, μ_{vis}, and the visible cross-section can be determined using Equation 4.3.

The bunch luminosity, $\mathscr{L}_b$, can be expressed in terms of various beam parameters as:

$$\mathscr{L}_b = f_r n_1 n_2 \int \rho_1\left(x,y\right)\rho_2\left(x,y\right)\mathrm{d}x\mathrm{d}y, \tag{4.4}$$

where the beams are assumed to be colliding with zero crossing angle, $n_{1(2)}$ is the number of particles (population) in bunch 1(2), and $\rho_{1(2)}\left(x,y\right)$ is the normalised particle density in the x-y plane for bunch 1(2) at the interaction point. An important component of the van der Meer method is the assumption that these particle densities can be factorised into independent horizontal and vertical components and this assumption allows Equation 4.4 to be written as:

$$\mathscr{L}_b = f_r n_1 n_2 \Omega_x\left(\rho_{x1}\rho_{x2}\right)\Omega_y\left(\rho_{y1}\rho_{y2}\right), \tag{4.5}$$

where:

$$\Omega_x\left(\rho_{x1}\rho_{x2}\right) = \int \rho_{x1}\left(x\right)\rho_{x2}\left(x\right)\mathrm{d}x \tag{4.6}$$

is the beam-overlap integral in the x-direction with an equivalent definition for the y-direction.

To calculate $\Omega_x\left(\rho_{x1}\rho_{x2}\right)$, van der Meer proposed measuring a quantity proportional to the luminosity, for example μ_{vis}, as the beams were separated in the horizontal (x) direction [21]. Therefore, $\Omega_x\left(\rho_{x1}\rho_{x2}\right)$ can be expressed as:

$$\Omega_x\left(\rho_{x1}\rho_{x2}\right)=\frac{R_x(0)}{\int R_x(\delta)\mathrm{d}\delta}, \tag{4.7}$$

where $R_x(\delta)$ is a quantity proportional to the luminosity measured during a horizontal scan when the beams are separated by a distance δ, and $\delta=0$ represents zero beam separation. Defining a parameter Σ_x as:

$$\Sigma_x=\frac{1}{\sqrt{2\pi}}\frac{\int R_x(\delta)\mathrm{d}\delta}{R_x(0)} \tag{4.8}$$

and a similar expression for Σ_y, Equation 4.5 can be rewritten as:

$$\mathscr{L}_b=\frac{f_r n_1 n_2}{2\pi\Sigma_x\Sigma_y}, \tag{4.9}$$

which allows one to determine the absolute bunch luminosity from the revolution frequency, f_r, the bunch population product, n_1n_2, and the product $\Sigma_x\Sigma_y$ which can be directly measured during a pair of orthogonal vdM scans.

When the luminosity curve $R_x(\delta)$ is a Gaussian function, Σ_x coincides with the standard deviation of that distribution. However, the van der Meer method works for any functional form of $R_x(\delta)$: the quantities Σ_x and Σ_y can be determined for any curve $R_x(\delta)$ using Equation 4.8 and then used with Equation 4.9 to determine the absolute luminosity at $\delta=0$.

To calibrate a particular detector and algorithm, Equations 4.9 and 4.3 for the bunch luminosity using beam parameters can be compared, resulting in the expression:

$$\sigma_{vis}=\mu_{vis}^{MAX}\frac{2\pi\Sigma_x\Sigma_y}{n_1n_2}, \tag{4.10}$$

where μ_{vis}^{MAX} is the visible interaction rate per bunch crossing at the peak of the scan curve, i.e. when the beams are colliding head-on, measured by that particular detector and algorithm. Therefore, Equation 4.10 provides an absolute calibration of the visible cross-section, σ_{vis}, for any detector and algorithm in terms of the peak visible interaction rate, μ_{vis}^{MAX}, the product of the convolved beam widths, $\Sigma_x\Sigma_y$, and the bunch population product, n_1n_2.

Another useful variable that can be extracted from the vdM scans is the *specific* luminosity, $\mathscr{L}_{sp}$:

$$\mathscr{L}_{sp}=\frac{\mathscr{L}}{n_1n_2}=\frac{f_r}{2\pi\Sigma_x\Sigma_y}. \tag{4.11}$$

This is a purely geometrical quantity because it depends only on the transverse beam sizes. Therefore, comparing $\mathscr{L}_{sp}$ measured by different detectors and algorithms during the same scan is a direct check on the consistency of the absolute luminosity scale provided by these methods.

4.3 Measuring Luminosity in ATLAS

Many physics analyses in ATLAS require an accurate determination of the delivered luminosity. For example, beyond the Standard Model searches rely on a precise measurement of the delivered luminosity to evaluate background levels and determine the sensitivity to the signatures of new phenomena. Also, one of the major systematic uncertainties on cross-section measurements is usually the uncertainty on the delivered luminosity. For these reasons, ATLAS employs a wide range of detectors and algorithms to measure the luminosity and by comparing the measurements made by these, it controls the systematic uncertainty on the result.

ATLAS monitors the delivered luminosity by measuring the observed interaction rate per bunch crossing, μ_{vis}, independently with a variety of detectors and algorithms and it determines σ_{vis} using dedicated vdM scans. The detectors and algorithms used in 2012 by ATLAS, as well as the 2012 dedicated vdM scans, are described in the following section.

Most bunch-by-bunch luminometers are composed of two symmetric detectors, one located on side A of the ATLAS detectors, and another located on side C.

4.3.1 Luminosity Algorithms

There are three main techniques used to determine the luminosity: *event counting*, where the number of bunch crossings in which a detector registers an event is counted, *hit counting*, where the number of hits per bunch crossing is counted, and *particle counting*, where the distribution of the number of particles per bunch crossing is determined.

ATLAS employs two *event counting* algorithms to determine μ_{vis}: the *EventOR* algorithm, where a BCID is said to contain an event if there was at least one hit on one side of the detector, and the *EventAND* algorithm, where a BCID is said to contain an event if there was at least one hit on both sides of the detector.

Under the assumption that the number of interactions in a bunch crossing obeys a Poisson distribution, the probability of a BCID passing the EventOR algorithm criteria, i.e. there being at least one hit, is equal to one minus the probability of there being zero hits:

$$P_{EventOR}\left(\mu_{vis}^{OR}\right) = \frac{N_{OR}}{N_{BC}} = 1 - P_0\left(\mu_{vis}^{OR}\right) = 1 - e^{-\mu_{vis}^{OR}}, \tag{4.12}$$

where N_{OR} is the number of bunch crossings, in a given time interval, in which the EventOR algorithm was satisfied, and N_{BC} is the total number of bunch crossings in the same time interval. This equation can be solved for μ_{vis} as a function of the event counting rate:

$$\mu_{vis}^{OR} = -\ln\left(1 - \frac{N_{OR}}{N_{BC}}\right). \tag{4.13}$$

A similar expression for μ_{vis}^{AND} can be derived: the probability of a BCID satisfying the EventAND algorithm criteria, i.e. there being at least one hit on both sides of the detector, is equal to one minus the probability of no hits on side A, minus the probability of no hits on side C plus the probability of no hits on either side. This can be expressed as:

$$P_{EventAND}\left(\mu_{vis}^{AND}\right) = \frac{N_{AND}}{N_{BC}} = 1 - e^{-\mu_{vis}^{A}} - e^{-\mu_{vis}^{C}} + e^{-\mu_{vis}^{OR}}. \tag{4.14}$$

This expression can be simplified using the relation between the visible interaction rates of each algorithm: $\mu_{vis}^{OR} = \mu_{vis}^{A} + \mu_{vis}^{C} - \mu_{vis}^{AND}$ and also assuming that the detectors in side A and C have approximately equal acceptances, which would imply $\mu_{vis}^{A} \approx \mu_{vis}^{C}$. Therefore, Equation 4.14 can be written as:

$$\frac{N_{AND}}{N_{BC}} = 1 - 2\exp\left[-\left(\frac{\mu_{vis}^{AND} + \mu_{vis}^{OR}}{2}\right)\right] + \exp\left[-(\mu_{vis}^{OR})\right]. \tag{4.15}$$

Finally, using Equation 4.3, Equation 4.15 can be rewritten in terms of μ_{vis}^{AND} and the cross-sections σ_{vis}^{OR} and σ_{vis}^{AND}, measured in a vdM scan. This is shown in Equation 4.16. Unfortunately, in this case the equation for μ_{vis}^{AND} cannot be analytically solved so instead μ_{vis}^{AND} is calculated numerically.

$$\frac{N_{AND}}{N_{BC}} = 1 - 2\exp\left[-\left(1 + \frac{\sigma_{vis}^{OR}}{\sigma_{vis}^{AND}}\right)\frac{\mu_{vis}^{AND}}{2}\right] + \exp\left[-\left(\frac{\sigma_{vis}^{OR}}{\sigma_{vis}^{AND}}\right)\mu_{vis}^{AND}\right] \tag{4.16}$$

ATLAS also employs the *hit counting* algorithm, where the number of hits registered by a certain detector is counted, instead of just the total number of events. With the assumptions that the number of hits in a *pp* interaction follows a binomial distribution and that the number of interactions per bunch crossing follows a Poisson distribution, the probability of having a hit per bunch crossing in one of the detector channels is:

$$P_{HIT}\left(\mu_{vis}^{HIT}\right) = \frac{N_{HIT}}{N_{BC}N_{CH}} = 1 - e^{-\mu_{vis}^{HIT}}, \tag{4.17}$$

where N_{HIT} is the total number of hits in the detector in a given time interval, N_{BC} is the total number of bunch crossing in that same interval and N_{CH} is the number of detector channels [22]. It is also assumed that each channel has an equal probability to be hit. This equation can be solved for μ_{vis}^{HIT}:

$$\mu_{vis}^{HIT} = -\ln\left(1 - \frac{N_{HIT}}{N_{BC}N_{CH}}\right). \tag{4.18}$$

ATLAS employs a further two algorithms for measuring the luminosity: *track counting* and *vertex counting*. These methods count the tracks of charged particles reconstructed by the ATLAS ID or the pp interaction vertices reconstructed from said tracks. There are two track counting algorithms used by ATLAS: *allTracks*, where all reconstructed tracks per event, satisfying certain criteria on p_{T} and number of hits (good tracks), are counted, and *vtxTracks*, where only good tracks matched to a vertex are counted. In the case of the vertex counting algorithm, the number of vertices per event satisfying certain criteria are counted. The vertex counting algorithm is one of the main topics of this thesis; therefore more details on this method such as triggers, track quality criteria and corrections for non-linear behaviour, will be explained in Chapter 5.

4.3.2 Bunch-by-bunch Luminosity Detectors

BCM, the Beam Conditions Monitor, is a device designed mainly to monitor background levels and send a signal to abort the beam in case beam losses threaten to damage the ATLAS detector. In addition to this, its fast readout provides a bunch-by-bunch low-acceptance luminosity measurement for $|\eta| = 4.2$. BCM consists of two stations located along the beam pipe at $z = \pm\, 1.84\,\mathrm{m}$ and each station has four diamond sensors distributed in a cross pattern. Diamond was chosen for its radiation hardness and fast signal formation [23]. The horizontal and vertical pairs of BCM detectors are read out separately which yields two luminosity measurements, BCMH and BCMV, respectively, which are treated as coming from two different detectors.

LUCID is a Cherenkov light detector specially designed to monitor the luminosity in ATLAS. It consists of sixteen aluminium tubes located around the beam pipe on each side of the ATLAS IP, at a distance of approximately 17 m and it covers the pseudorapidity range $5.6 < |\eta| < 6.0$. Originally, the tubes were filled with

C_4F_{10} gas, which increased the flux of Cherenkov radiation but, for most of 2012 LUCID was operated without the gas to reduce the pile-up effects by reducing the sensitivity of the detector [24]; this allowed measurements of a wider range of luminosity values. In this configuration, the Cherenkov photons were produced only in the quartz windows separating the gas volumes from the PMTs located at the back of the detector.

4.3.3 Bunch-integrating Luminosity Detectors

The hadronic tile calorimeter, TileCal, and the forward calorimeter, FCal, are used to determine the luminosity under the assumption that the currents drawn in particular regions of the calorimeter should be proportional to the particle flux in that region. In the case of TileCal, the measurements are based on the currents drawn by the PMTs from selected cells in a region near $|\eta| \approx 1.25$, where the largest variations in current as a function of the luminosity are observed [22]. In the case of FCal, the measurements are based on the currents that provide a stable field across the liquid argon cells in the modules closest to the IP [25].

It is not possible to calibrate TileCal or FCal directly, since they are not sensitive enough to the low luminosity of the vdM scans. Instead, σ_{vis} for these detectors is obtained by comparing the currents measured to the luminosity measured by either LUCID or BCM. In the case of TileCal, the comparison is made at the peak of a vdM scan and in the case of FCal, the comparison is made using a high-luminosity reference run.

The MPX system consists of 13 Medipix pixel detectors distributed throughout the ATLAS detector and its main goal is to measure the radiation field and its composition within ATLAS [26]. It is used for relative luminosity measurements since the number of particles detected by each MPX detector is assumed to be proportional to the luminosity. Each MPX detector is read-out independently and, like TileCal and FCal, the MPX system measures the luminosity summed over all colliding bunch pairs and cannot be calibrated during a vdM scan. Instead, the total number of pixel clusters observed in each detector during the same reference run as FCal is counted and compared to the BCMHEventOr luminosity to provide a scale factor for each detector.

4.3.4 ATLAS vdM Scans

ATLAS employs the van der Meer method to calibrate its luminosity detectors and algorithms. This method was explained in Section 4.2 and in the case of ATLAS, the quantity R_x proportional to the luminosity is μ_{vis}.

Ideally for this method, μ_{vis} would be measured continuously as a function of beam separation, allowing $\int \mu_{vis}(\delta)\,\mathrm{d}\delta$, and therefore Σ_x and Σ_y, to be determined exactly. However, to be able to obtain a precise measurement of μ_{vis}, the beams need to be held at a constant separation for a period of around 45 seconds, known as a pseudolumiblock (pLB)[1]. Because of time constraints, μ_{vis} is measured at a limited number of beam separation steps, usually 25 during the 2012 vdM scans.

What is in fact measured during a vdM scan, is the *specific* visible interaction rate, μ_{vis}^{sp}, defined as:

$$\mu_{vis}^{sp} = \frac{\mu_{vis}}{n_1 n_2}. \tag{4.19}$$

The values of Σ_x and Σ_y are determined by fitting the curve of μ_{vis}^{sp} as a function of beam separation; Σ_x is obtained from the x-scan and Σ_y from the y-scan. The functions used to fit the μ_{vis}^{sp} versus beam separation curves for the 2012 vdM scans were a Gaussian plus a constant “background” and the sum of two Gaussians (double Gaussian) plus a constant “background”, where the background terms are subtracted before determining the values of the Σ’s. However, these functions do not perfectly describe the luminosity curves and there are dedicated studies conducted by the ATLAS luminosity task force to find a better fit [25]. For the purpose of this thesis, however, the fits mentioned before will be used.

---

[1] *Lumiblocks* (LB) typically last 60 seconds and are used during physics runs.

Chapter 5

Measuring Luminosity with the Vertex Counting Algorithm

5.1 Introduction

The vertex counting algorithm has been developed and studied before in ATLAS [27, 28]. This algorithm was used to measure the luminosity delivered to ATLAS during 2011 [27] and the method described in this chapter builds upon that work. The purpose of this chapter is to explain the implementation of the vertex counting algorithm to measure the luminosity delivered to ATLAS during 2012. This represented an interesting challenge since the vertex counting algorithm was known to suffer from “pile-up effects”, which means that the luminosity measured by the vertex counting algorithm does not increase linearly with pile-up and, as can be seen in Figure 5.1, the pile-up seen in 2012 was more than twice the pile-up seen in 2011.

Nevertheless, there are some advantages to using the vertex counting algorithm. One advantage is that the backgrounds associated are smaller than for other luminosity detectors and algorithms and they can be further reduced by being more selective with the definition of the vertices to be counted, for example requiring more and well-identified tracks associated to them. Another advantage is that this method can be *event-based*, counting events with or without at least one vertex, or *hit-based*, counting every vertex in the event. This provides an internal consistency check, although this thesis only focuses on the hit-based method. Another internal consistency check can be done by using more than one definition of a vertex, depending, for example, on the number of tracks associated to them. Additionally, vertex counting is relatively easy to study in simulated events produced using the Monte Carlo (MC) method.

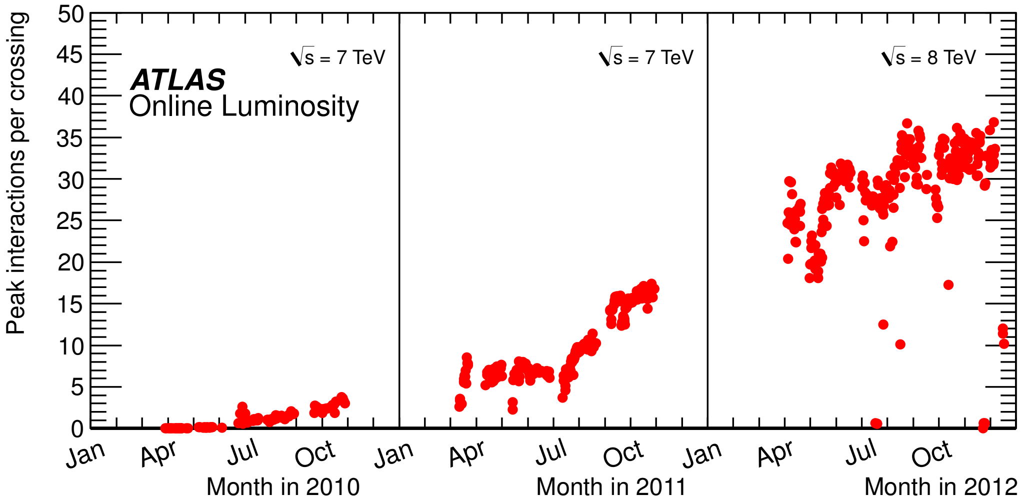


Figure 5.1: The maximum mean number of events per beam crossing versus day during the pp runs of 2010, 2011 and 2012. This figure shows the average value for all bunch crossings in a lumi-block. The online luminosity measurement is used for this calculation. Only the maximum value during stable beam periods is shown. Taken from Ref. [29].

To measure the luminosity using the vertex counting algorithm, Equation 4.3 is used, where μ_{vis} represents the measured number of vertices per bunch crossing and σ_{vis} is determined using the centred vdM scans performed during 2012.

Vertices are reconstructed in ATLAS using an iterative algorithm which first finds a vertex seed and then considers the corresponding tracks for consistency with the vertex seed [30, 31]. A vertex seed is found by looking for the global maximum in the distribution of z-coordinates of the reconstructed tracks. Then, the position of the vertex is determined by an adaptive vertex fitting algorithm [32], which uses the position of the seed and the tracks around it as input. This is a robust χ^2-based fit which suppresses the contribution from outlier tracks. Tracks that are incompatible with the vertex by more than 7σ are reused to seed a new vertex and this process is repeated until no more vertices can be found.

For general purposes, the tracks used to reconstruct vertices are required to have $p_{\mathrm{T}} > 400\,\mathrm{MeV}$, 7 silicon hits and at most two holes in the pixel detector, where a hole is defined when a track traverses an active module of the detector but the module does not report a hit. However, for the vertex counting algorithm, tracks are required to pass a new set of quality cuts which was developed in 2012 and is referred to as *VtxLumi* cuts. These tighter cuts, compared to the normal track finding in ATLAS, require tracks to have higher p_{T}, above 900 MeV, zero pixel holes and no more than two SCT holes, all of which reduces CPU timing drastically.

In the vertex counting algorithm, a vertex is counted if it satisfies certain quality criteria. One such vertex is called a *tight* vertex, while a *loose* vertex is one that fails to pass the quality cuts. In this particular analysis there is one quality cut, the number of tracks associated to the vertex. Therefore, a *tight* vertex is one which has at least “NTrkCut” tracks associated it. Varying the value of NTrkCut provides a cross-check on the consistency of the method and helps control the background.

The data and MC samples used for the vertex counting algorithm are described in Sections 5.2 and 5.3. The pile-up effects and the corrections associated to them are described in Section 5.4. A MC closure test was performed to validate the vertex counting algorithm and the results are presented in Section 5.5. Section 5.6 shows the calibration of the method using the 2012 vdM scans. Section 5.7 presents the results of implementing the vertex counting algorithm to measure the luminosity during 2012 physics runs. Section 5.8 addresses the systematic uncertainties affecting this method. Finally, Section 5.10 presents the conclusions and future plans for the vertex counting algorithm. Throughout this chapter there are also subsections describing particular changes that had to be implemented to the method used for the 2011 running to be able to cope with the Run 1 conditions; these are labelled “change to original method”.

5.2 Data samples

The data for the vertex counting algorithm comes from different triggers and streams, depending on the run type. For special runs like the vdM scans, the stream `calibration_vdM` provides data for a small set of BCIDs at a high rate of approximately 15 kHz. There were three vdM scan sets in 2012, in April, July and November (Table 5.2). The triggers used by this stream were a random trigger at Level-1, while at Level-2 the triggering was done on a few selected BCIDs. During the July and November vdM scan sets, the triggers also required that the event had at least two hits in any MBTS (Minimum Bias Trigger Scintillators). [22].

For physics runs, the stream `calibration_PixelBeam` is used. This data stream was implemented in 2012 and stores only Pixel and SCT information for fast readout and low bandwidth, and allows the rate to go up to approximately 100 Hz with random triggers, as opposed to a rate of about 5 Hz in 2011. Some important features of the data analysed is shown in Table 5.1. The data and MC samples used in this chapter use the *VtxLumi* cuts mentioned before.

Run	Number of events	Average peak μ	Beam spot width [mm]
vdM scans 1, 2, 3	95,398,099	4.2	44
vdM scans 4, 5, 6	92,521,905	0.489	59
vdM scan 8	54,260,835	0.538	56
vdM scans 10, 11, 14	118,405,868	0.645	60
vdM scan 15	60,867,255	0.627	60
2012 Physics Runs	$1-6\times10^6$ (per run)	~32	40-55

Table 5.1: Data samples used in the vertex counting method. The vdM scan samples come from the stream `calibration_vdM` and the 2012 physics runs samples come from the stream `calibration_PixelBeam`. The number of events, peak value of the average number of observed interactions, μ, and the beam spot width is listed in every case.

vdM Scan Set	Scans	Run Number
April	1, 2, 3	201351
July	4, 5, 6	207216
	8	207219
November	10, 11, 14	214984
	15	215021

Table 5.2: vdM Scan sets performed in 2012 with their corresponding month, scan numbers and run numbers.

5.3 Monte Carlo samples

Two MC samples, both generated with Pythia 8 and tune A2M [33, 34], were used in this analysis. From here onwards they will be referred to as “low μ” and “high μ” MC samples because of the μ range they span, shown in Table 5.3. μ is the average number of interactions simulated in a set of events, while the actual number of generated interactions in each simulated event is denoted by n_{gen}.

Sample ID	Energy	μ range	Average beam spot width
Low μ	8 TeV	1 to 22	47 mm
High μ	8 TeV	38 to 71	48 mm

Table 5.3: MC samples used for calibration of the vertex counting method. Both samples were generated using the *VtxLumi* reconstruction settings and they have a similar beam spot width, however, they span a different range of μ, a low μ range and a high μ range.

Previous studies show that the most important feature of the MC when it comes to extracting the correction for fake vertices is the luminous region width, which for 2012 physics runs varies from 40 to 55 mm. The MC samples used in this analysis have a beam spot width of approximately 47 mm. Figure 5.2 shows the distribution of the z coordinate (z-distribution) for vertices with NTrkCut 3, for the two MC samples, two vdM scans and a physics run, for comparison. The RMS of this distribution is equivalent to the luminous region width.

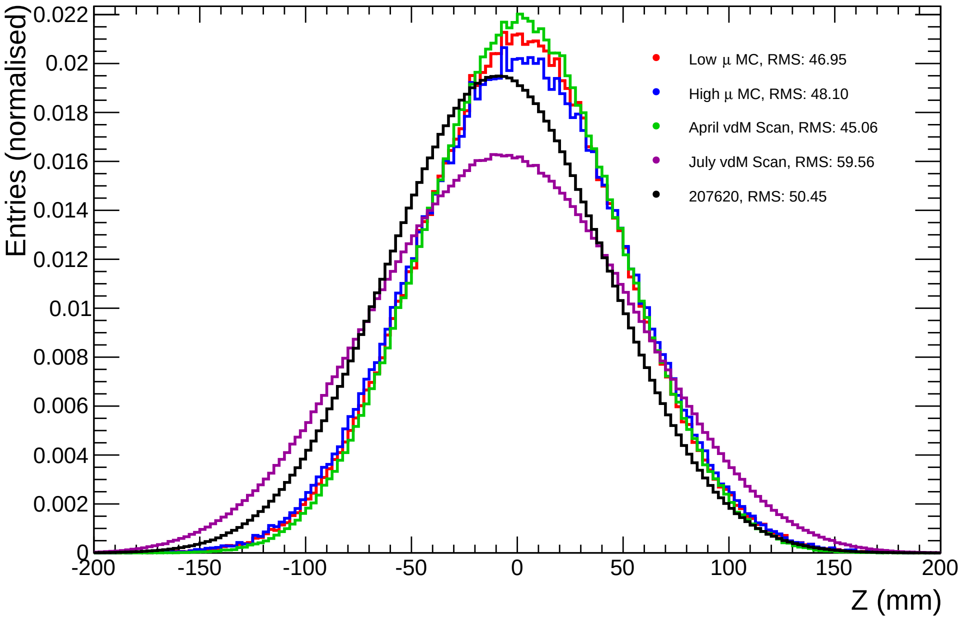


Figure 5.2: Normalised z-distribution of NTrkCut 3 vertices for the two MC samples, the April and July vdM scans and the physics Run 207620.

5.3.1 Change to original method: μ vs. n_{gen}

When the vertex counting algorithm was used in 2011, the corrections extracted from MC were calculated as a function of the number of generated proton-proton inelastic collisions in a particular bunch crossing, n_{gen}, which in principle is a random number from a Poisson distribution with mean μ. As can be seen in Table 5.3, the available MC samples do not span a continuous range of μ and the gap missing is where physics runs most typically occur. The variable n_{gen} does span the whole range but unfortunately, when a cross-check was performed, it was found that the mean of the n_{gen} distribution coming from events with a particular value of μ, was not μ. Instead, the mean value of the distribution differed by 2% from μ in the high μ sample, as can be seen in Figure 5.3.

The MC samples are generated with particular values of μ set as input parameters and n_{gen} is calculated, for a sample with a particular μ, by counting all of the in-time pile-up generated vertices. The classification of a vertex as in-time pile-up by the particular event collection available in the MC samples is not 100% reliable and some in-time pile-up vertices could be failing the cuts to be classified as such. For this reason, μ is a more reliable variable than n_{gen} and the MC corrections in this analysis are calculated as a function of μ, and then parametrised so they can be applied in all relevant μ ranges. Also, n_{gen} is not available in physics runs whereas μ is calculable over a LB.

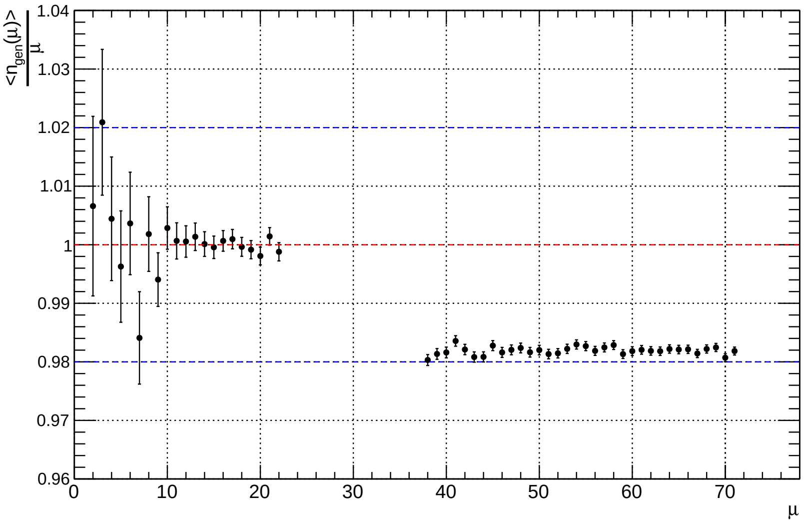


Figure 5.3: Ratio of the mean of the n_{gen} distributions and μ, as a function of μ for both MC samples. This figure shows that for the high μ MC sample, the mean of the n_{gen} distributions coming from events with a particular μ differs by 2% of μ.

5.4 Pile-up effects

The presence of multiple interactions per bunch crossing, or pile-up, affects the vertex counting algorithm resulting in a non-linear relationship between the luminosity and the number of reconstructed vertices. The pile-up effects that arise in this method are the following:

Split vertices

When two tight vertices are reconstructed from a single interaction.

Fake vertices

When a tight vertex is reconstructed from loose vertices combined with other tracks.

Vertex masking

When a tight vertex fails to be reconstructed because some of its tracks were used in the reconstruction of another close-by vertex.

The analysis begins with the measurement of μ_{rec}, the average number of reconstructed vertices per event. This variable is strongly affected by pile-up effects; therefore the aim is to calculate μ_{vis}, the average number of vertices that would have been reconstructed if there were no pile-up effects. In order to do so, two corrections are applied. First, the contribution from fake vertices, μ_{fake}, is subtracted from μ_{rec} to obtain the real number of reconstructed vertices per event, μ_{real} (Equation 5.1) and then a scaling function of μ_{real} is applied to account for the fraction of real vertices that were masked, $f_{mask}(\mu_{real})$ (Equation 5.2):

$$\mu_{real} = \mu_{rec} - \mu_{fake}(f_{mask}(\mu_{rec}) \times \mu_{rec}), \tag{5.1}$$

$$\mu_{vis} = f_{mask}(\mu_{real}) \times \mu_{real}. \tag{5.2}$$

As was shown in Figure 5.2, the beam spot width in data and MC is different, which means vertex masking and fake vertices affect data and MC differently. In the case of the beam spot being narrower in MC than in data, vertex masking would be a greater effect in MC than in data. Therefore, for a given level of activity in the inner detector, fewer vertices will be reconstructed in MC than in data. If the bare number of reconstructed vertices in data and MC are matched, the fake correction would be wrongly determined. To account for this, an initial masking correction is applied to data and MC before evaluating the fake correction. This is why, in Equation 5.1, μ_{fake} is a function of $\mu_{rec} \times f_{mask}(\mu_{rec})$.

A correction for the split vertices is not applied since their contribution is quite small, as is shown in the next section.

5.4.1 Split vertices

The effect of split vertices is studied using truth matching in simulation and can be seen in Figure 5.4. For vertices with NTrkCut 3, 4 and 5, the contribution from vertex splitting increases slightly with pile-up but is generally below 0.25%, and therefore no correction is applied. For NTrkCut 2 vertices, the split vertices contribution is quite large, therefore this vertex definition is not used in this analysis.

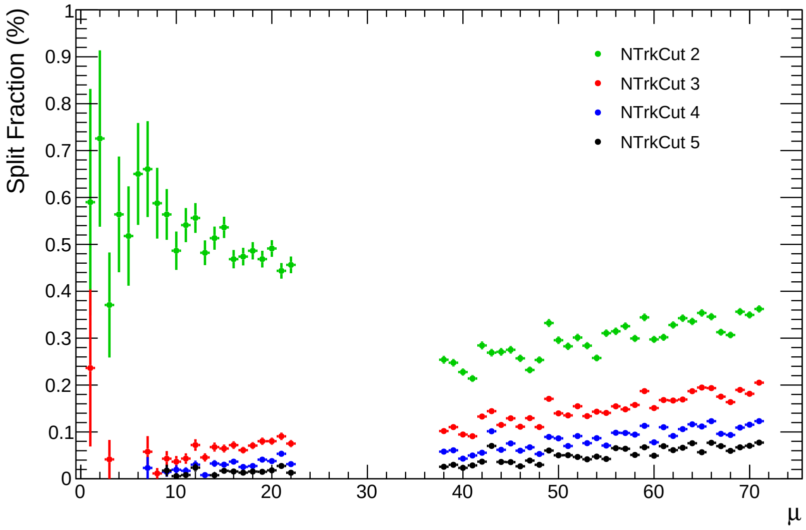
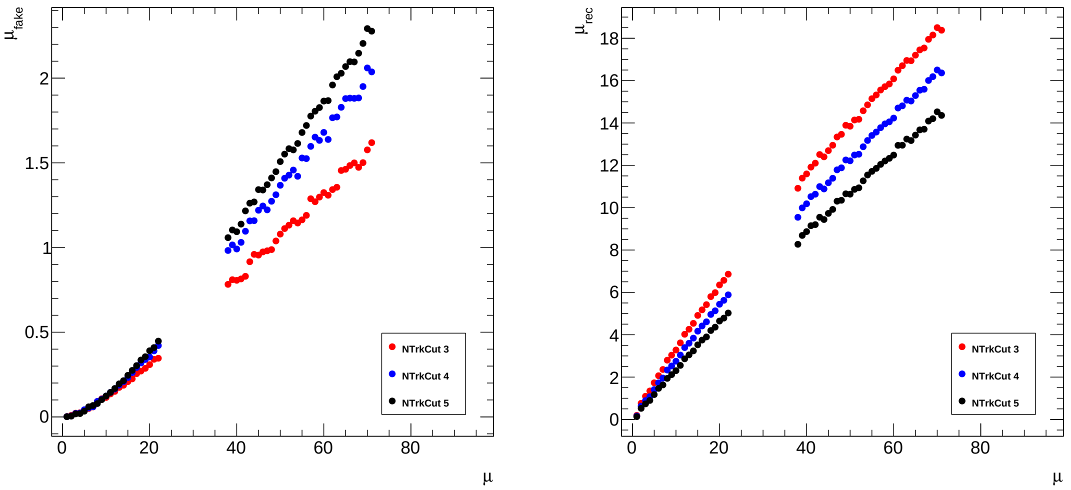


Figure 5.4: Fraction of reconstructed vertices labelled as split by MC truth matching, as a function of μ. For vertices with NTrkCut 3, 4 and 5, this fraction is generally below 0.25% and no correction is applied. However, for vertices with NTrkCut 2, this fraction is quite large; therefore this vertex definition is not used in the analysis.

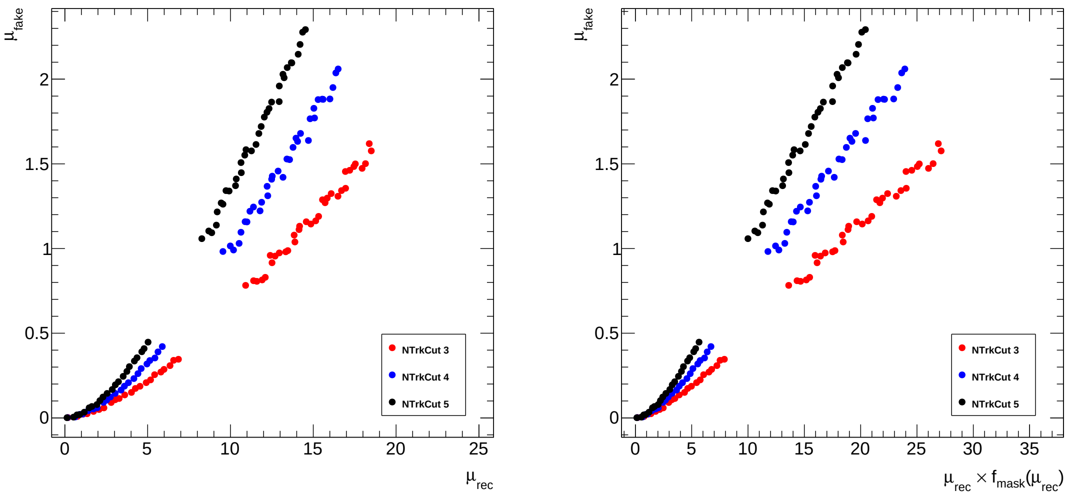
5.4.2 Fake vertices

A fake vertex is a vertex that would normally fail to pass the cut on minimum number of tracks but does pass it because it gathers extra tracks from nearby pile-up interactions. Fake vertices and their effects are analysed using truth matching in MC simulation. The definition of a fake vertex depends on the cut on the minimum number of tracks per vertex, i.e. if we require m tracks per vertex, a reconstructed vertex will be considered *real* if at least m of its tracks are found to be matched to charged particles originating from the same interaction, and *fake* otherwise.



(a) μ_{fake} vs μ.

(b) μ_{rec} vs μ.



(c) μ_{fake} vs μ_{rec}.

(d) μ_{fake} vs μ_{rec} after masking correction.

Figure 5.5: Calibration of the fake correction using both MC samples together. Plots 5.5a and 5.5b show μ_{fake} and μ_{rec} as a function of μ, respectively. Using these, plot 5.5c is constructed. Then, the masking correction is applied to every value of μ_{rec} to construct plot 5.5d. This last plot is parametrised to extract the fake correction.

Figure 5.5 shows the steps on the evaluation of the fake vertices correction. First, two histograms are constructed, the average number of reconstructed vertices labelled as fake, μ_{fake}, as a function of μ (Figure 5.5a) and the average number of reconstructed vertices, μ_{rec}, as a function of μ (Figure 5.5b). The next step is to combine these two histograms to create a histogram of μ_{fake} as a function of μ_{rec} (Figure 5.5c). Then, the plot shown in Figure 5.5d is constructed by taking every value of μ_{rec} and multiplying it by its masking correction factor $f_{mask}(\mu_{rec})$[1] and plotting it against its corresponding μ_{fake} value. As can be seen in these figures, in every case there is a large dependence on μ and on the minimum number of tracks per vertex. Finally, since the values in which the fake correction will be evaluated when applying this correction to physics runs are mostly where this plot (Figure 5.5d) has a gap, the plot is parametrised. Several fit functions were tried and χ^2/ndf was used as a figure of merit to choose between them. The resulting fit functions, χ^2/ndf values and comparisons between MC data and fits for different values of NTrkCut are shown in Figure 5.6.

---

[1] A detailed description of f_{mask} can be found in the next section.

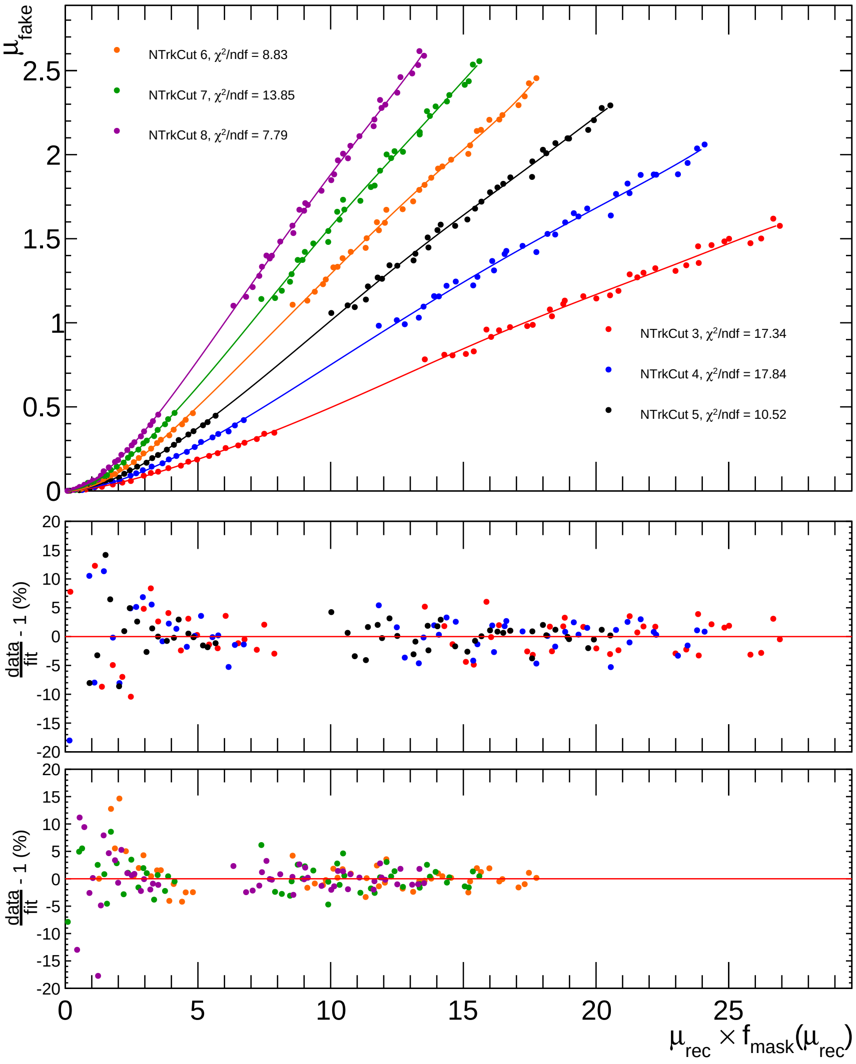


Figure 5.6: Parametrised fake correction plots for all considered values of NTrkCut. Shown on the top plot are the μ_{fake} as a function of $\mu_{rec} \times f_{mask}$ graphs for different values of NTrkCut fitted with a polynomial of order 7 constrained to pass through the point (0,0). The percentage difference between the MC data and the fit for NTrkCut values of 3, 4 and 5 is shown in the middle plot. The percentage difference between the MC data and the fit for NTrkCut values of 6, 7 and 8 is shown in the the bottom plot.

5.4.3 Vertex masking

Vertex masking is the inability of the vertex counting algorithm to distinguish between vertices close to one another. Vertex masking arises from two sources:

1. Tracks coming from two very close-by interactions may all be compatible with coming from the same point, i.e. the reconstructed vertex.

2. The vertex seeding mechanism used to decide the starting point for a new vertex is highly inefficient in finding a seed nearby to an already reconstructed vertex.

Out of these two effects, the second one has the largest contribution and that is why the seeding algorithm was modified for Run 2 of the LHC.

The vertex masking effect can be analysed using the distribution of Δz between pairs of vertices in the same event. The Δz-distribution is generated taking all vertices in an event, pairing them and plotting $\Delta z = \text{vertex}_1(z) - \text{vertex}_2(z)$. For *truth* interaction pairs, the distribution of Δz should depend on the vertex z-distribution. For example, if the vertex z-distribution is a Gaussian with width σ_z, then the Δz-distribution of vertex pairs should also be a Gaussian but with width $\sqrt{2}\sigma_\text{z}$. Vertex masking produces a prominent absence near $\Delta z = 0$. This is illustrated in Figure 5.7.

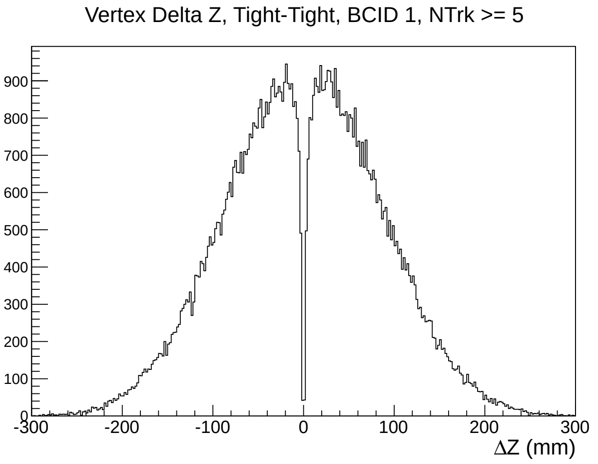


Figure 5.7: Δz-distribution for pairs of tight vertices in the same event. Data is taken from the July 2012 van der Meer scan. The absence of events near $\Delta z = 0$ is due to vertex masking.

The vertex masking correction is data-driven and it works as follows:

1. Using samples with low pile-up, vdM scans in the case of data and event samples with μ values between 1 and 10 for MC, the two-vertex masking

probability, $p_{mask}(\Delta z)$, is calculated. This is the probability that only one of two *tight* vertices separated by Δz gets reconstructed. This function is assumed to be a universal property of the vertex counting algorithm, therefore independent of μ.

To do this, the shape of the expected Δz-distribution, $f_{exp}(\Delta z)$, is derived by randomly sampling pairs of points from the observed vertex z-distribution. Then, the observed Δz-distribution, $f_{obs}(\Delta z)$ (black histogram in Figures 5.8a, 5.8c and 5.8e) is fitted using the shape of $f_{exp}(\Delta z)$ as a template to obtain the correct normalisation of $f_{exp}(\Delta z)$ to the data (shown as the red histogram in Figures 5.8 and 5.9). This is done in the range 20 mm $\leq |\Delta z| \leq$ 300 mm, where vertex masking is negligible. Lastly, the masking probability is defined as:

$$p_{mask}(\Delta z) = \frac{f_{exp}(\Delta z) - f_{obs}(\Delta z)}{f_{exp}(\Delta z)} \tag{5.3}$$

2. In order to apply the masking correction to a particular data sample with a given μ and width of the luminous region in z, an *expected* Δz-distribution is derived, $f_{exp}(\Delta z)$, by sampling pairs of vertices in a random way from the observed vertex z-distribution of that particular sample. Then, the total probability that only one of two tight vertices is reconstructed, $\mathcal{P}_{mask}$, is derived:

$$\mathcal{P}_{mask} = \int_{-\infty}^{\infty} p_{mask}(\Delta z) f_{exp}(\Delta z) d\Delta z \tag{5.4}$$

   (The actual limits are ± 20 mm and $f_{exp}(\Delta z)$ is normalised before calculating the integral.)

3. Using the total masking probability, $\mathcal{P}_{mask}$, a function is generated relating the number of reconstructible vertices per event, N_{vis}, and the average number of reconstructed vertices per event given the vertex masking effect, μ_{real}, as follows. The generated vertices are labelled v_i, with $1 \leq i \leq N_{gen}$, in the order in which they were reconstructed by the iterative vertex finding algorithm. Analogously, the probability of vertex v_i to be reconstructed is labelled p_i, with $1 \leq i \leq N_{vis}$. Proceeding vertex by vertex, the p_i follow the recursion

relation:

$$
\begin{aligned}
p_1 &= 1 \\
p_2 &= p_1 \times (1 - \mathcal{P}_{mask}) \\
&\dots \\
p_k &= \prod_{i=1}^{k-1} (p_i \times (1 - \mathcal{P}_{mask}) + (1 - p_i) \times 1) \\
&= \prod_{i=1}^{k-1} (1 - p_i \mathcal{P}_{mask}) \\
&= p_{k-1} \times (1 - p_{k-1} \mathcal{P}_{mask})
\end{aligned}
\tag{5.5}
$$

Then, the average number of real vertices is:

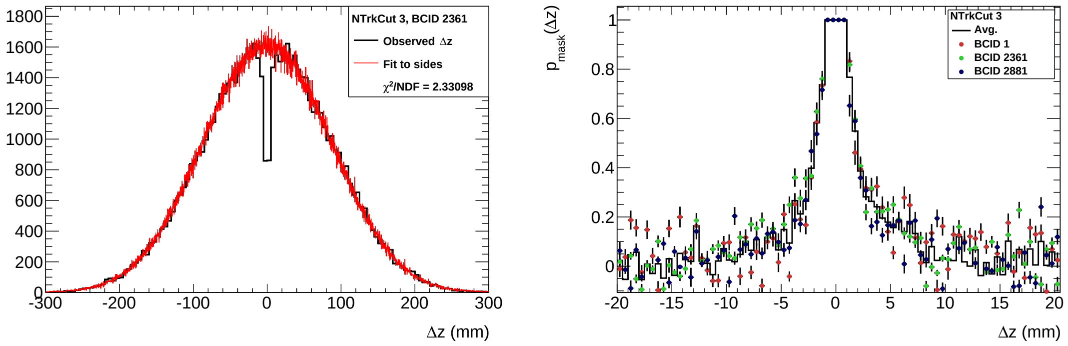
$$\mu_{real}(N_{vis}) = \sum_{i=1}^{N_{vis}} p_i \tag{5.6}$$

4. Then, an expression for μ_{real} as a function of μ_{vis} is computed by convolving $\mu_{real}(N_{vis})$ with a Poisson distribution:

$$\mu_{real}(\mu_{vis}) = \sum_{N_{vis}=1}^{\infty} P(N_{vis}; \mu_{vis}) \mu_{real}(N_{vis}) \tag{5.7}$$

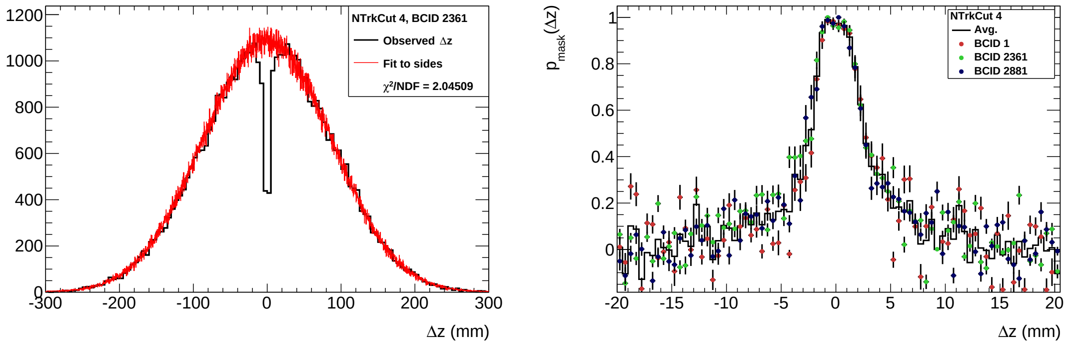
5. Finally, a plot of $f_{mask} = \mu_{vis}/\mu_{real}$ as a function of μ_{real} is generated, from which the masking correction factor, f_{mask}, will be extracted to be used in Equation 5.2. This plot is generated every time the masking correction is evaluated in data or MC, since it is a function of the size of the luminous width.

Figures 5.8 and 5.9 show the Δz-distribution fitted with the “expected” Δz-distribution, and the p_{mask} vs Δz distributions calculated for data (vdM scan 15, BCID 2361) and MC, for different values of NTrkCut. Figure 5.10 shows the plots of the masking correction factor for all the vdM scans and the MC, for different values of NTrkCut. The figures shown for NTrkCut 3 include a “technical correction” that is explained on the next section and is a change to the original method.



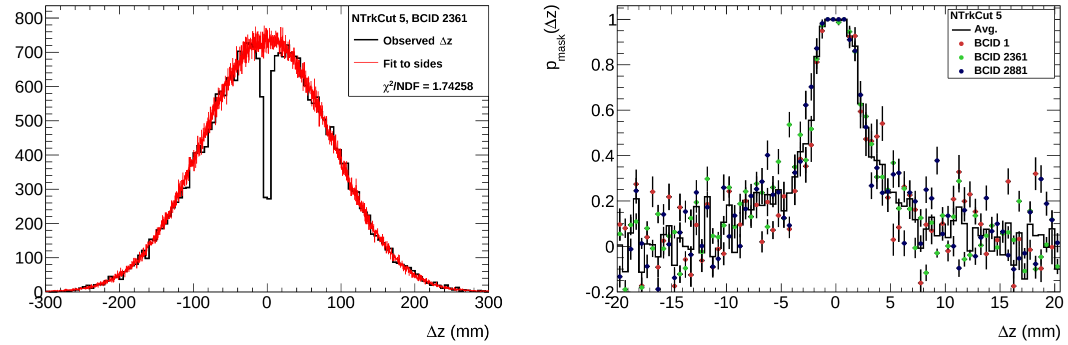
(a) Δz distribution and template fit, NTrkCut 3.

(b) p_{mask} vs. Δz, NTrkCut 3.



(c) Δz distribution and template fit, NTrkCut 4.

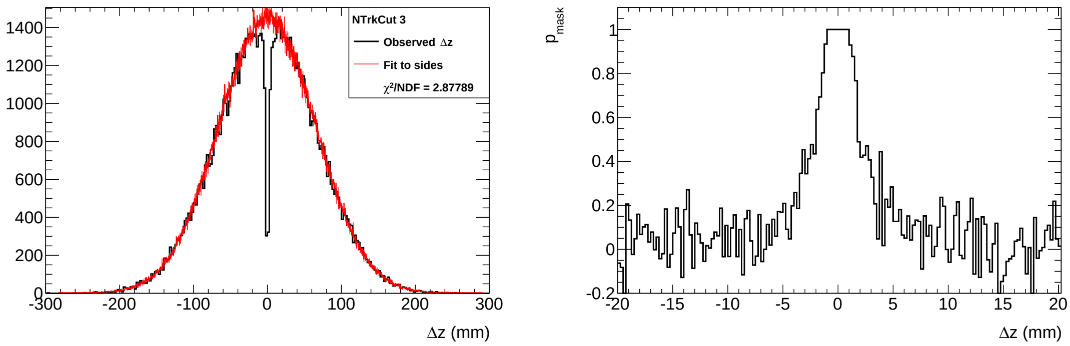
(d) p_{mask} vs. Δz, NTrkCut 4.



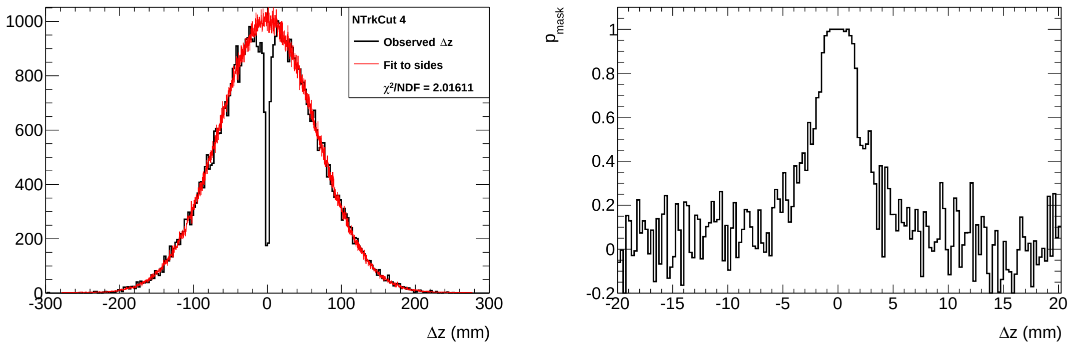
(e) Δz distribution and template fit, NTrkCut 5.

(f) p_{mask} vs. Δz, NTrkCut 5.

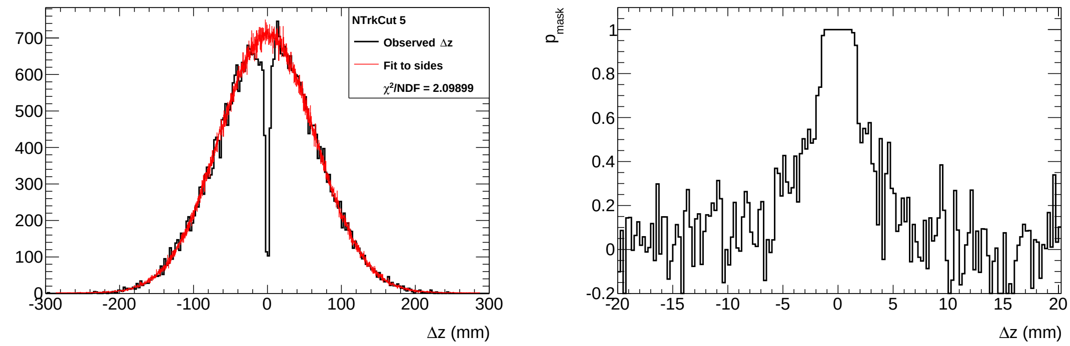
Figure 5.8: Calibration of the masking correction method in data, using the November vdM scan 15. The black histograms on the left correspond to the Δz-distributions for BCID 2361 and different values of NTrkCut, and the red "curves" are the template fits using the expected Δz-distribution in the range $20\,\text{mm} \leq |\Delta z| \leq 300\,\text{mm}$. The p_{mask} as a function of Δz is shown on the right for all BCIDs in the scan.



(a) Δz distribution and template fit, NTrkCut3. (b) p_{mask} vs. Δz, NTrkCut3.

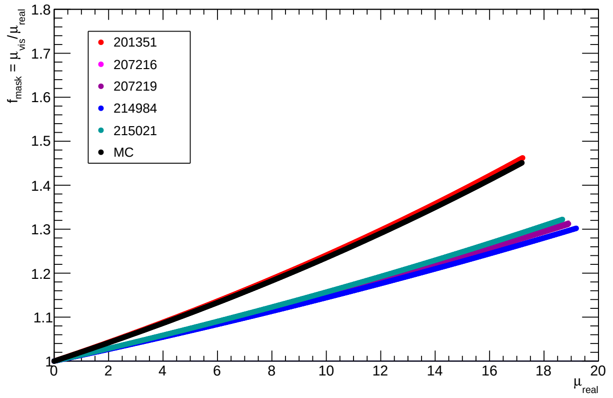


(c) Δz distribution and template fit, NTrkCut4. (d) p_{mask} vs. Δz, NTrkCut4.

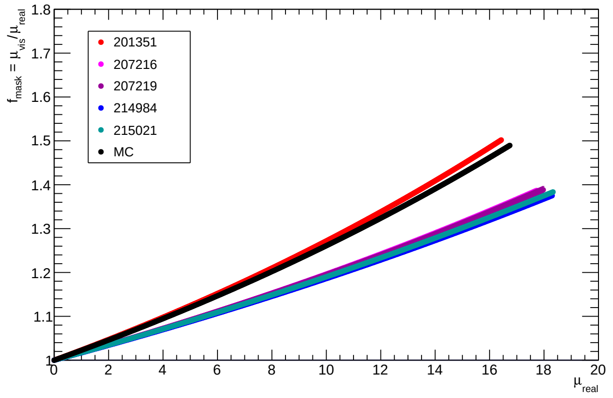


(e) Δz distribution and template fit, NTrkCut5. (f) p_{mask} vs. Δz, NTrkCut5.

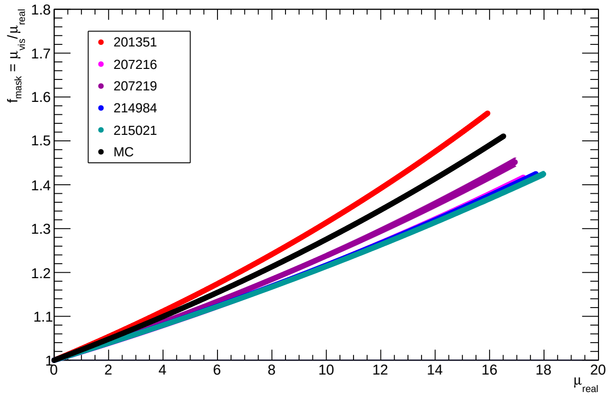
Figure 5.9: Calibration of the masking correction method in MC. On the left, the black histograms correspond to the Δz-distribution and the red “curves” are the template fits using the expected Δz-distribution in the range $20\,\text{mm} \leq |\Delta z| \leq 300\,\text{mm}$. On the right p_{mask} as a function of Δz is shown for different values of NTrkCut.



(a) NTrkCut 3.



(b) NTrkCut 4.

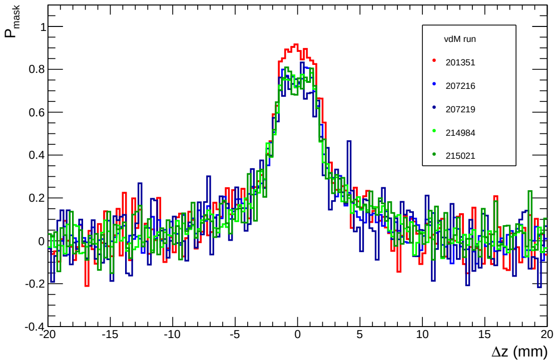


(c) NTrkCut 5.

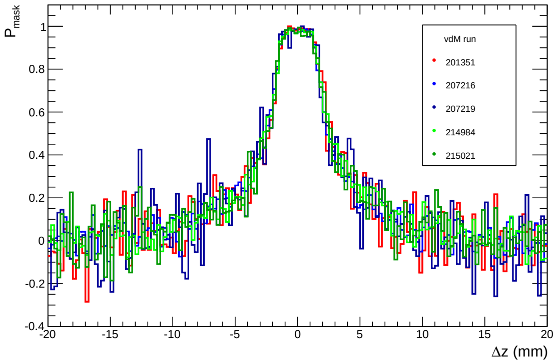
Figure 5.10: Calibration of the masking correction for all vdM scans and MC for NTrkCut values of 3, 4 and 5. Plotted is the masking correction factor, $f_{mask} = \mu_{vis}/\mu_{real}$, vs. μ_{real}.

5.4.4 Change to original method: Vertex masking correction for NTrkCut 3 and low μ

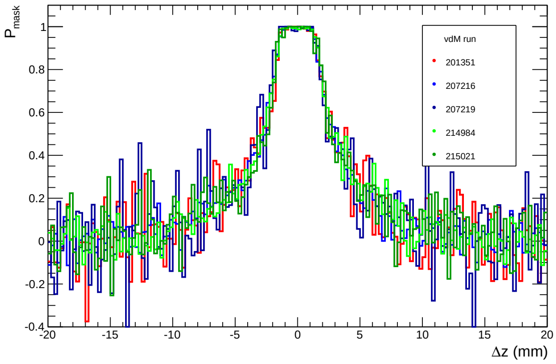
While studying the masking correction, it became clear that for every vdM run, the p_{mask} vs Δz distribution for vertices with NTrkCut 3 was different from that for vertices with NTrkCut 4 and 5, as can be seen in Figure 5.11.



(a) NTrkCut 3.



(b) NTrkCut 4.



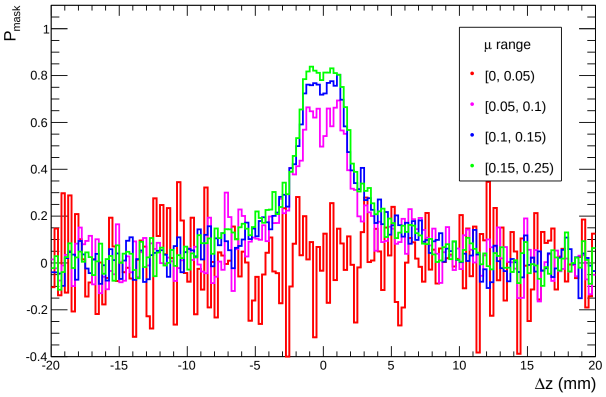
(c) NTrkCut 5.

Figure 5.11: p_{mask} vs Δz for all vdM scans and for vertices with NTrkCut 3, 4 and 5.

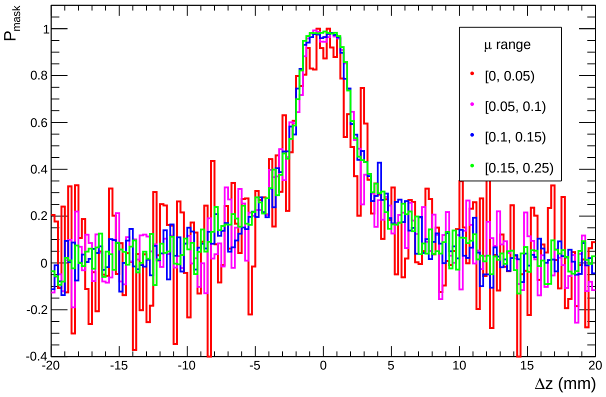
Figure 5.11 shows that, for vertices with NTrkCut 4 and 5 , the probability of a vertex masking another very close-by vertex ($\Delta z \approx 0$) is practically 1, as it would be expected. However, this is not the case for NTrkCut 3 vertices, where the histogram suggests that it is possible to distinguish between two vertices that are very close together, which should not be possible with the current experimental resolution.

To investigate further, four artificial runs were created by taking all the z-distributions and Δz-distributions from the July and November 2012 vdM scans and grouping them together in ranges of μ_{rec}. Then, for each artificial run, a masking correction and a p_{mask} vs Δz distribution were calculated. The resulting p_{mask} vs Δz distributions are shown in Figure 5.12 and it shows a few features. First, for vertices with NTrkCut 4 and 5, the p_{mask} vs Δz distributions are very similar irrespective of the μ_{rec} range, except for the lowest one. Therefore, after this study, the pLBs with $\mu_{rec} \leq 0.05$ were removed from the masking correction calculation. Second, for vertices with NTrkCut 3, the p_{mask} vs Δz distributions show a clear dependency on the μ_{rec} range and for the lowest range, the distribution is definitely not what would be expected.

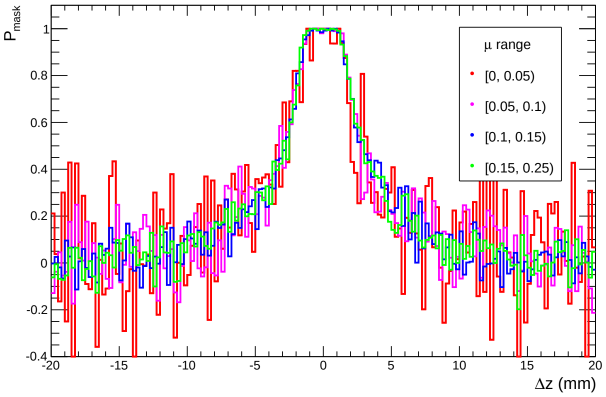
Only the data from the July and November vdM scans were grouped together since these scans have similar beam spot width. A similar test was attempted with the April vdM scan data but there was not enough data at low μ_{rec} to extract any significant results.



(a) NTrkCut 3.



(b) NTrkCut 4.



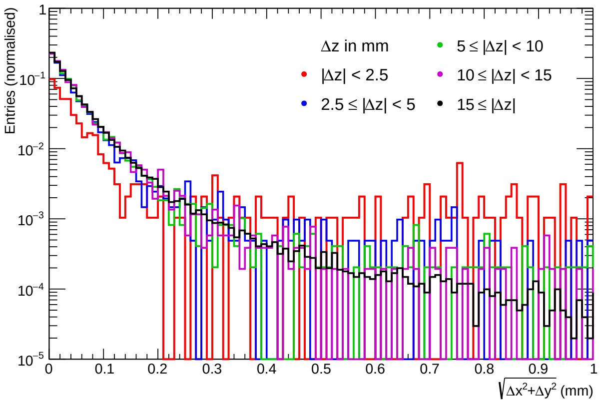
(c) NTrkCut 5.

Figure 5.12: p_{mask} vs Δz for four different μ_{rec} ranges for vertices with NTrkCut 3, 4 and 5.

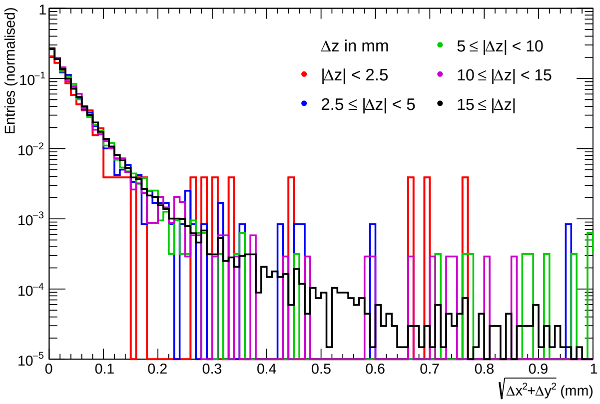
Another test was performed where the distribution of $\sqrt{\Delta x^2 + \Delta y^2}$ for pairs of vertices with different Δz were calculated. These distributions are shown in Figure 5.13. For vertices with NTrkCut 4 and 5, the distribution of the pairs of vertices distance in the x-y plane is very similar regardless of their distance in the z-axis. This means that the vertices are evenly distributed in the x-y-z space. However, this is not the case for vertices with NTrkCut 3, where the distribution of $\sqrt{\Delta x^2 + \Delta y^2}$ for vertices very close together in z is different from that for vertices that are separated in z, with a much more pronounced tail to large values of $\sqrt{\Delta x^2 + \Delta y^2}$.

In conclusion, these studies suggest that there is a problem with the evaluation of the masking correction for vertices with NTrkCut 3. There appears to be a mechanism producing two vertices that are very close together in z, from a single collision. This effect is most visible at very low luminosity, when the probability of events containing two genuine collisions is very low. Without further study it would be difficult to say whether the origin of this effect is split vertices, secondary interactions, or some other mechanism. For this reason, in this analysis p_{mask} is forcibly set to 1 when $|\Delta z| < 1\,\mathrm{mm}$, for vertices with NTrkCut 3.

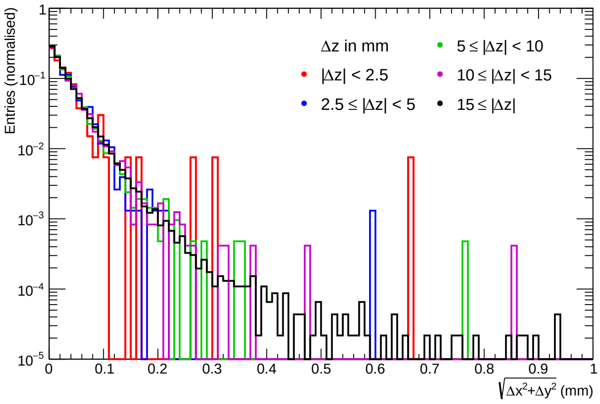
These corrections, i.e. eliminating low μ_{rec} pLBs from the masking correction calculation of vdM scans, and setting p_{mask} to 1 when $|\Delta z| < 1\,\mathrm{mm}$ for NTrkCut 3 vertices, are already applied in Figures 5.8 to 5.10.



(a) NTrkCut 3.



(b) NTrkCut 4.

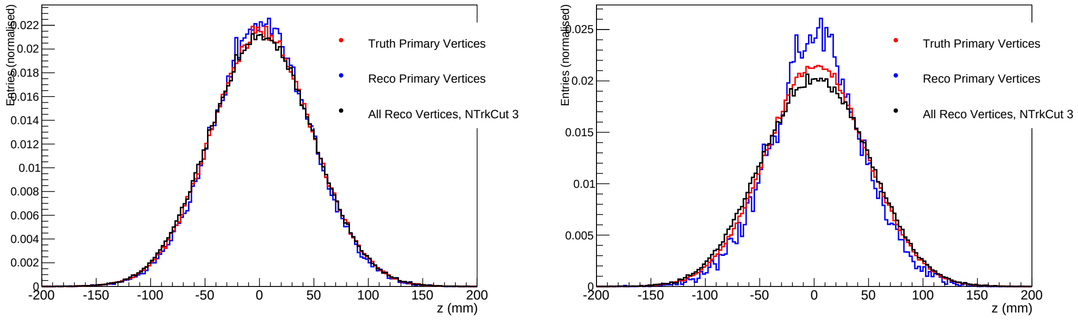


(c) NTrkCut 5.

Figure 5.13: Distribution of $\sqrt{\Delta x^2 + \Delta y^2}$ for pairs of vertices with Δz in four different μ_{rec} ranges for vertices with NTrkCut 3, 4 and 5 in BCID 1 of vdM Run 214984, which corresponds to scans 10, 11 and 14 in November.

5.4.5 Change to original method: Primary vertices distribution vs. all vertices distribution

Another cross-check was performed while studying the masking correction, this time using the MC samples. In the original method, the z-distribution used in the evaluation of the masking correction was coming from the reconstructed primary vertices, where a primary vertex is the reconstructed vertex with the highest sum p_{T}^2. [2] Figure 5.14 shows the comparison of the z-distribution from reconstructed primary vertices and from *generated* vertices (truth vertices) for both MC samples. It is clear that for the high μ sample, these distributions are quite different. Also shown in Figure 5.14 and in Figure 5.15 is the z-distribution of all reconstructed vertices. These distributions are in better agreement with the ones obtained with the truth vertices. Therefore, this analysis uses the z-distribution of all reconstructed vertices satisfying the NTrkCut requirement to evaluate the masking correction in data and MC.



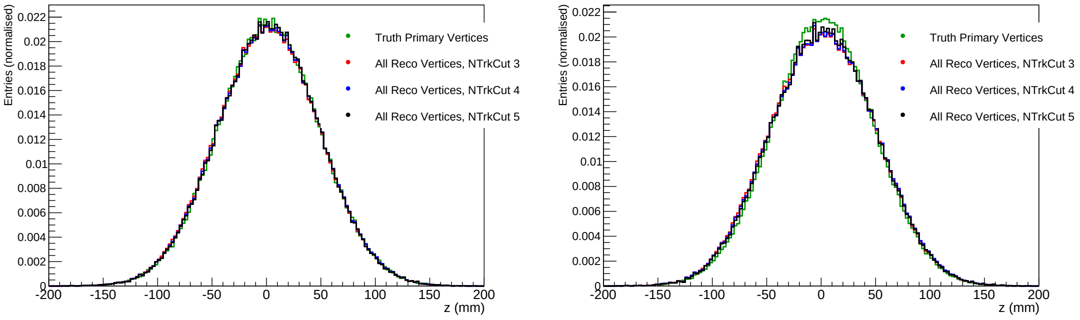
(a) Low μ sample.

(b) High μ sample.

Figure 5.14: Comparison of the z-distributions coming from truth primary vertices, reconstructed primary vertices and all reconstructed primary vertices with NTrkCut 3, for the low and high μ MC samples. The distributions have been normalised.

---

[2] The sum p_{T}^2 of a vertex is defined as the sum of the p_{T}^2 of all the tracks associated to it.



(a) Low μ sample.

(b) High μ sample.

Figure 5.15: Comparison of the z-distributions coming from truth primary vertices and all reconstructed primary vertices with different values of NTrkCut, for the low and high μ MC samples. The distributions have been normalised.

5.5 MC Closure test

A closure test was performed on the MC samples to validate the masking and fake corrections. Vertices were counted in bins of μ and the low and high μ MC samples were analysed together. The same MC samples were used to perform the vertex counts and to extract the corrections. Ideally, statistically independent MC samples would have been used to evaluate the corrections and then test them. Unfortunately, the available MC statistics were not sufficient to allow this.

For the masking correction, the p_{mask} vs. Δz distribution was created using vertices in the MC event samples corresponding to a μ in the range [2,10] (shown in Figure 5.9) which is a range in which there is low pile-up but also adequate statistics. An “expected” Δz-distribution and a masking correction were generated for the low μ and the high μ samples separately.

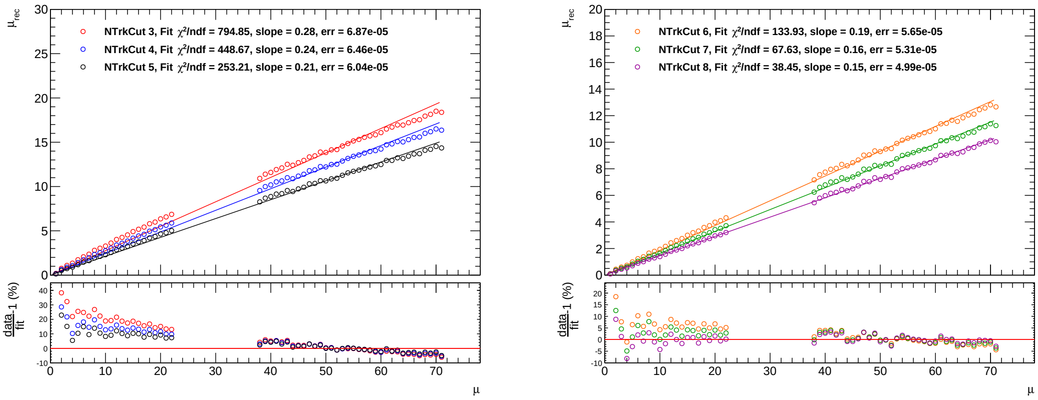
For the fake correction, both samples were analysed together but here too a masking correction (to go from histogram 5.5c to histogram 5.5d) was calculated for the two MC samples separately.

The next sections show the steps to go from μ_{rec} to μ_{vis} for different values of NTrkCut as a function of μ, and in each case a straight line passing through zero is fitted to the curves to show the progress of the corrections. The MC closure test results are shown for the standard procedure followed in this analysis but also for two special cases. In case 1, the z-distribution of *truth* primary vertices is used to calculate the masking correction in every step, instead of the z-distribution of reconstructed vertices. In case 2, the true value of μ_{real} is obtained from the MC sample instead of calculating it using Equation 5.1. The results from the MC closure test with the standard procedure show a very linear relation between μ_{vis} and μ and also they are comparable to the results from cases 1 and 2, where parts of the procedure used "cheat" information. Therefore, it can be concluded that the vertex counting algorithm with its pile-up corrections works well and can be applied to study 2012 data.

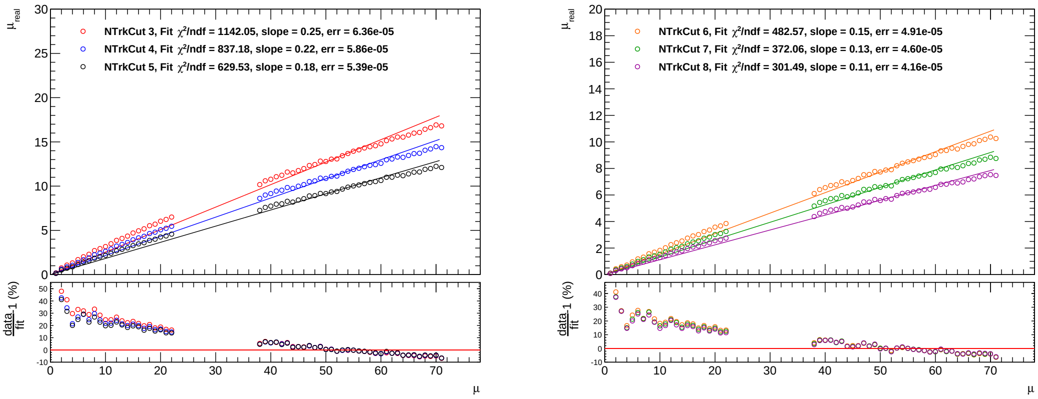
The results of the MC closure test before the changes to the original method were applied are also shown in Figure 5.19. This figure shows that the previous method, when applied to the 2012 MC samples, did not close.

5.5.1 Results

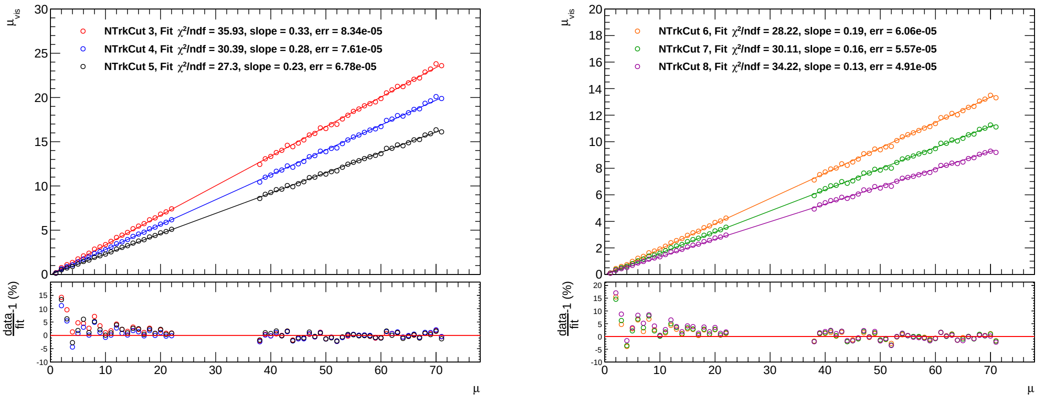
Figure 5.16 shows the results of the MC closure test using the standard procedure in this analysis.



(a) μ_{rec} vs. μ (b) μ_{rec} vs. μ



(c) μ_{real} vs. μ (d) μ_{real} vs. μ

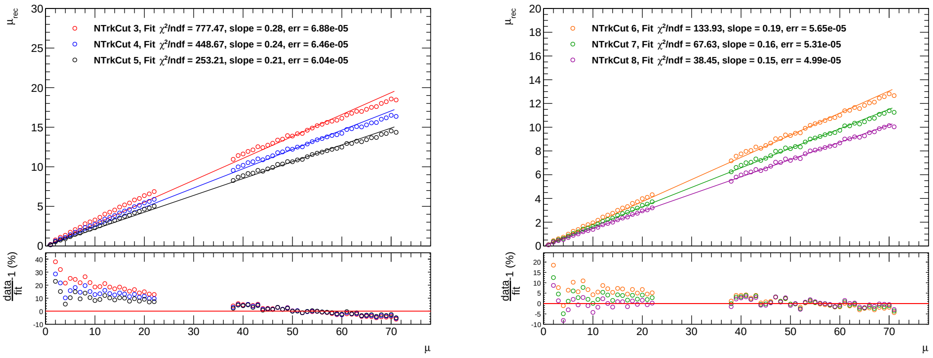


(e) μ_{vis} vs. μ (f) μ_{vis} vs. μ

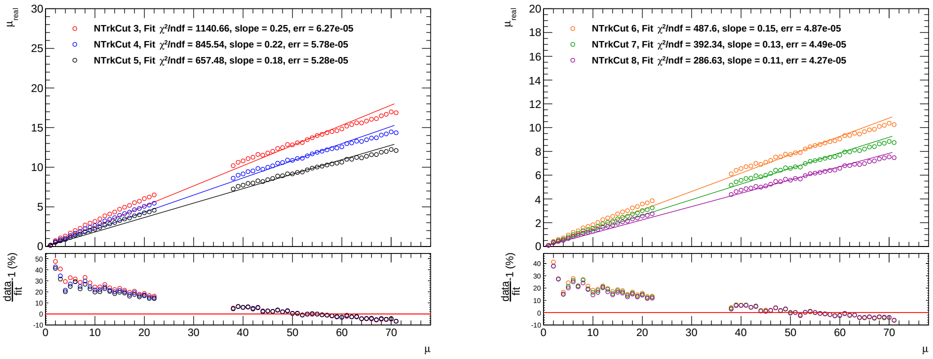
Figure 5.16: MC closure test results in steps for vertices with NTrkCut 3, 4 and 5, fitted with a straight line passing through 0.

5.5.2 Results: Special Case 1

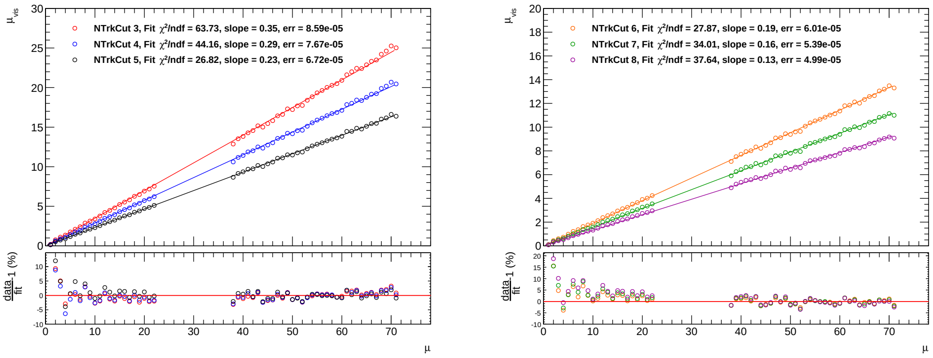
Figure 5.17 shows the results of the MC closure test using the z-distributions of *truth* primary vertices when calculating the masking correction.



(a) μ_{rec} vs. μ (b) μ_{rec} vs. μ



(c) μ_{real} vs. μ (d) μ_{real} vs. μ

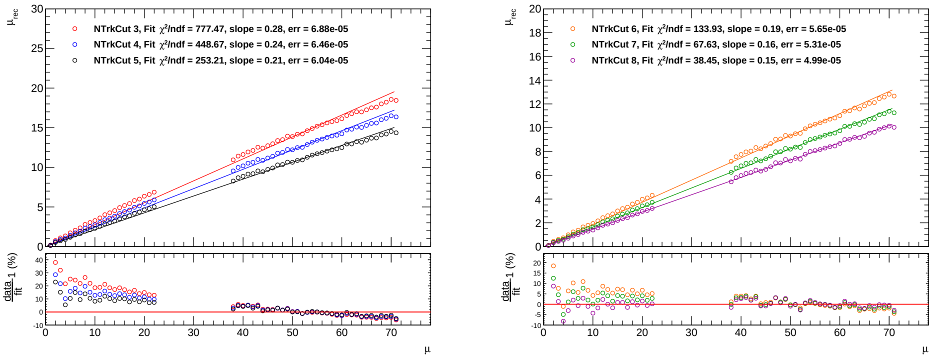


(e) μ_{vis} vs. μ (f) μ_{vis} vs. μ

Figure 5.17: MC closure test results in steps for vertices with NTrkCut 3, 4 and 5, fitted with a straight line passing through 0. z-distributions used throughout the calculation of the masking correction correspond to the *truth* primary vertices.

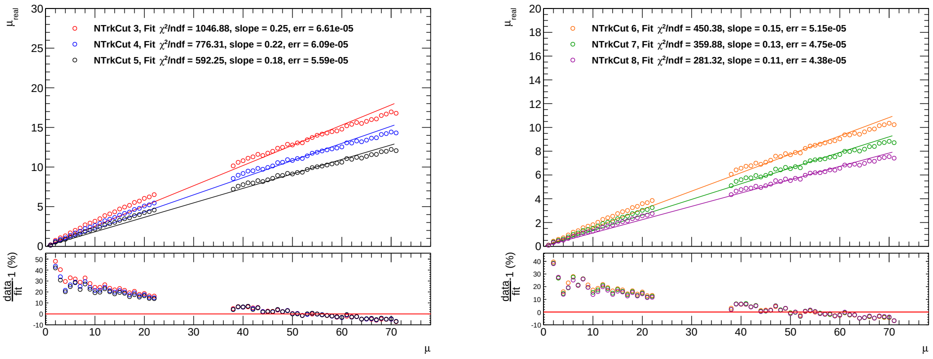
5.5.3 Results: Special Case 2

Figure 5.18 shows the results of the MC closure test using the z-distributions of *truth* primary vertices when calculating the masking correction and also reading in the value of μ_{real} from the MC instead of calculating it using equation 5.1.



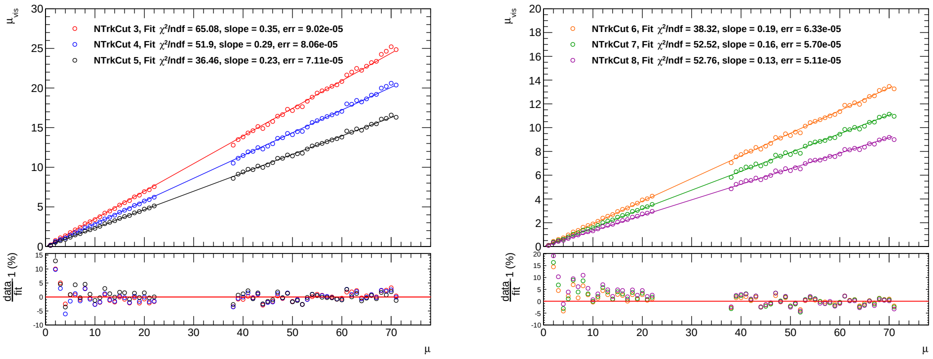
(a) μ_{rec} vs. μ

(b) μ_{rec} vs. μ



(c) μ_{real} vs. μ

(d) μ_{real} vs. μ



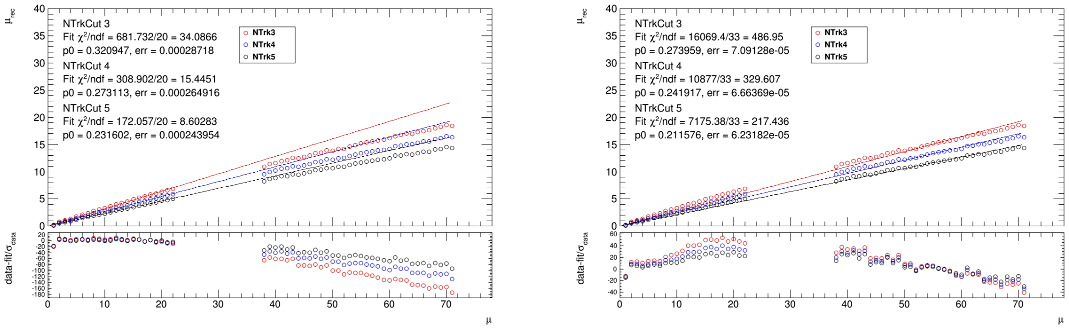
(e) μ_{vis} vs. μ

(f) μ_{vis} vs. μ

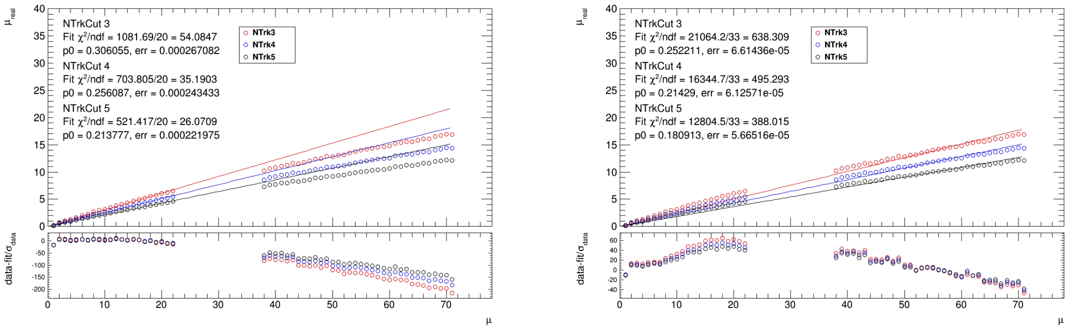
Figure 5.18: MC closure test results in steps for vertices with NTrkCut 3, 4 and 5, fitted with a straight line passing through 0. z-distributions used throughout the calculation of the masking correction correspond to the *truth* primary vertices and the value of μ_{real} was read in from the MC.

5.5.4 Results: Original Corrections

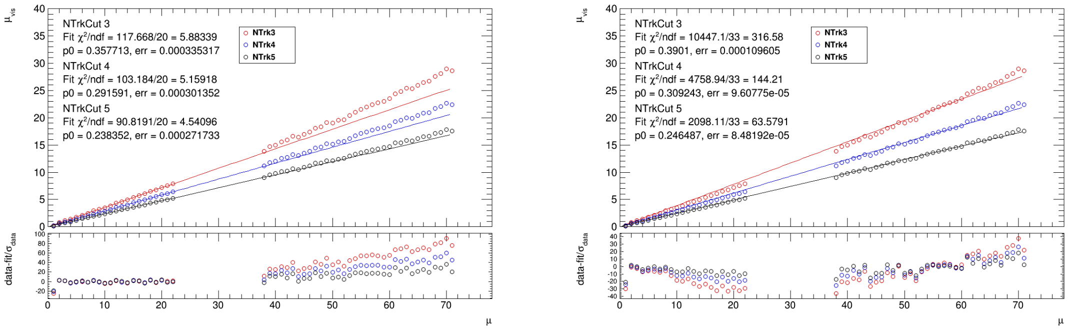
Figure 5.19 shows the results of the MC closure test using the original vertex counting algorithm, without the updates to the method this analysis has implemented. The most important difference is that the corrections were calculated as a function of n_{gen} instead of as a function of μ. It is clear from this picture that the original method, when applied to the 2012 MC samples, did not close.



(a) μ_{rec} vs. μ, Fitting low μ (b) μ_{rec} vs. μ, Fitting high μ



(c) μ_{real} vs. μ, Fitting low μ (d) μ_{real} vs. μ, Fitting high μ



(e) μ_{vis} vs. μ, Fitting low μ (f) μ_{vis} vs. μ, Fitting high μ

Figure 5.19: MC closure test results with the original method and fitting the low and high μ range separately with a straight line passing through 0.

5.6 Calibration Results: 2012 vdM Scans

In 2012 there were three sets of van der Meer scans performed. The first one was on April 16th (ATLAS Run 201351, LHC Fill 2520) and consisted of three centred scans labelled 1, 2 and 3. The second one was on July 19th (ATLAS runs 207216 and 207219, LHC Fills 2855 and 2856) and consisted of three centred scans labelled 4, 5 and 6, then an offset scan labelled 7, then again a centred scan labelled 8 and finally an offset scan labelled 9. The third set of vdM scans took place on November 22nd and 23rd (ATLAS runs 214984 and 215021, LHC Fills 3311 and 3316) and consisted of two consecutive centred scans labelled 10 and 11, then two offset scans labelled 12 and 13, and finally two centred scans labelled 14 and 15.

The centred vdM scans were analysed using the vertex counting algorithm and σ_{vis} was calculated in each case. Every scan was fitted with two functions: a single Gaussian plus a constant and a double Gaussian plus a constant. The results presented in the following sections correspond to the fit function that returned the best χ^2/ndf of the two. It is important to mention that sometimes data is not described perfectly by either of these functions and the luminosity task force has performed studies to find which function might describe data to a better degree [25].

The *centring correction* was also applied to the vdM scans. This corrects for the fact that sometimes the peak μ_{vis}, i.e. when the two beams are colliding head on, does not occur at precisely $\delta = 0$. The scan in the horizontal direction is always performed first and the beams are centred in the LHC control room before this scan begins. However, unless the beam position is too far from the centre before the vertical scan begins, the beams are not re-centred. This correction accounts for this, even though it is a relatively small effect.

The results obtained for every vdM scan set are presented in Sections 5.6.1, 5.6.2 and 5.6.3. The results are then compared among the different vdM scan sets and to the results obtained by other luminosity detectors and algorithms in Section 5.6.4.

5.6.1 April 2012

Figures 5.20 and 5.21 show the performed fits using a double Gaussian plus constant fit function for the four recorded BCIDs for vertices with NTrkCut 3, 4 or 5. The two points with largest separation (at $-137.8\,\mathrm{mm}$ and $135.8\,\mathrm{mm}$) were not used when performing the fit. The σ_{vis} values are shown in Table 5.4 for all scans and BCIDs. The consistency of the obtained σ_{vis} values among the different BCIDs and scans will be discussed later.

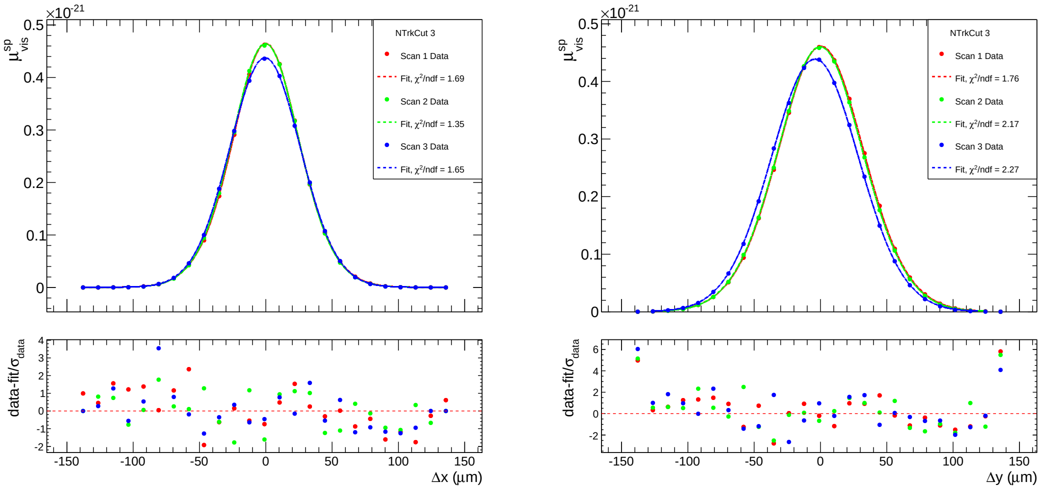
Figure 5.22 shows the corrections applied to the April vdM scans for three different values of NTrkCut; the absolute correction value on the left and the correction percentage on the right. It can be seen in these figures that the fake correction can be as big as 3.5%, the masking correction can be as big as 4.5% and the overall correction as big as 1.9%.

		σ_{vis} [mb]		
Scan	BCID	NTrkCut 3	NTrkCut 4	NTrkCut 5
1	1	24.447 ± 0.041	20.03 ± 0.038	16.595 ± 0.034
	241	24.323 ± 0.045	19.891 ± 0.041	16.505 ± 0.037
	2881	24.391 ± 0.037	20.033 ± 0.034	16.621 ± 0.031
	3121	24.238 ± 0.039	19.854 ± 0.035	16.454 ± 0.032
2	1	24.397 ± 0.041	19.977 ± 0.037	16.573 ± 0.034
	241	24.426 ± 0.049	19.985 ± 0.044	16.553 ± 0.04
	2881	24.532 ± 0.037	20.091 ± 0.033	16.655 ± 0.03
	3121	24.455 ± 0.041	20.029 ± 0.037	16.602 ± 0.034
3	1	24.867 ± 0.043	20.348 ± 0.038	16.876 ± 0.035
	241	24.951 ± 0.048	20.41 ± 0.043	16.926 ± 0.039
	2881	24.886 ± 0.037	20.395 ± 0.034	16.903 ± 0.031
	3121	24.863 ± 0.043	20.361 ± 0.039	16.886 ± 0.037

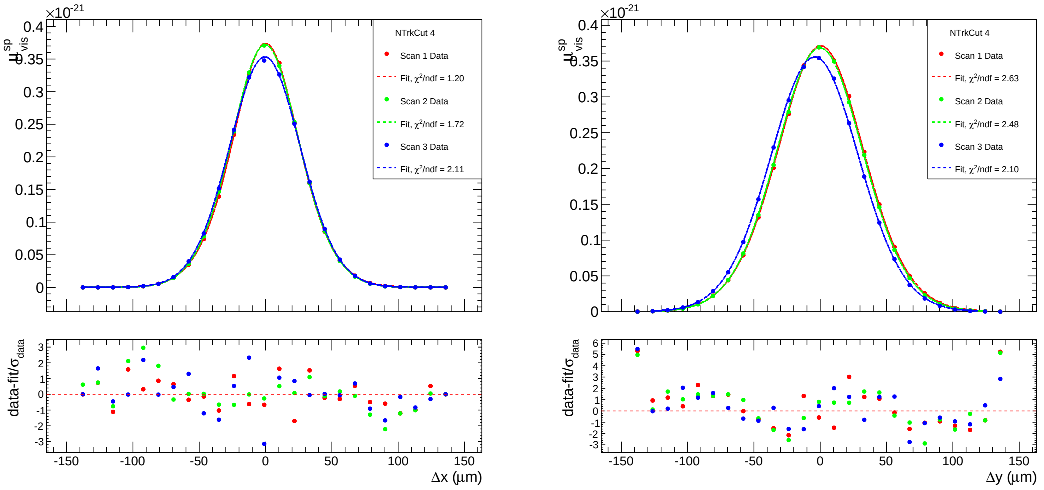
Table 5.4: Visible cross sections for vdM scan 201351, determined using a double Gaussian plus a constant fit function. The uncertainties are statistical only.

	σ_{vis} [mb]		
Scan	NTrkCut 3	NTrkCut 4	NTrkCut 5
1	24.349 ± 0.021	19.95 ± 0.019	16.543 ± 0.017
2	24.449 ± 0.022	20.017 ± 0.02	16.592 ± 0.018
3	24.894 ± 0.022	20.379 ± 0.02	16.898 ± 0.018

Table 5.5: Visible cross sections for vdM scan 201351 averaged over BCIDs, determined using a double Gaussian plus a constant fit function. The uncertainties are statistical only.

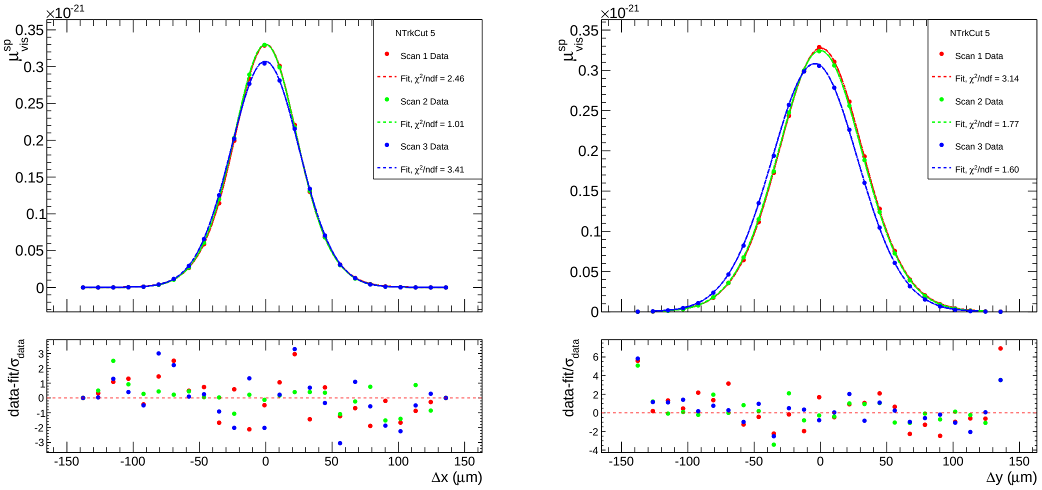


(a) x-scan, BCID 1 (b) y-scan, BCID 1

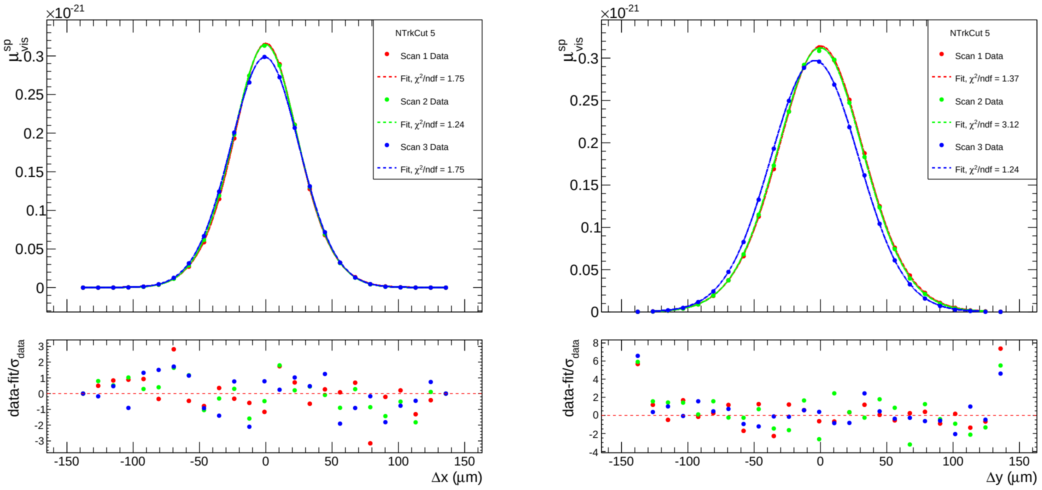


(c) x-scan, BCID 241 (d) y-scan, BCID 241

Figure 5.20: $\mu_{\text{vis}}^{\text{sp}}$ vs. beam separation with double Gaussian fits and residuals for BCIDs 1 and 241 of the April vdM scans.

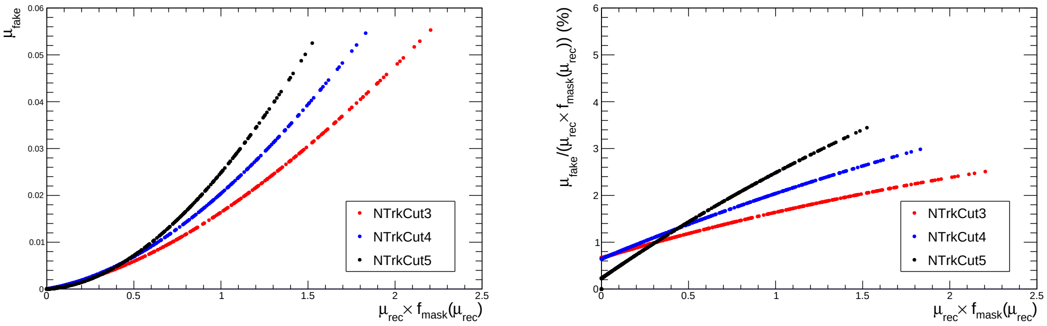


(a) x-scan, BCID 2881 (b) y-scan, BCID 2881



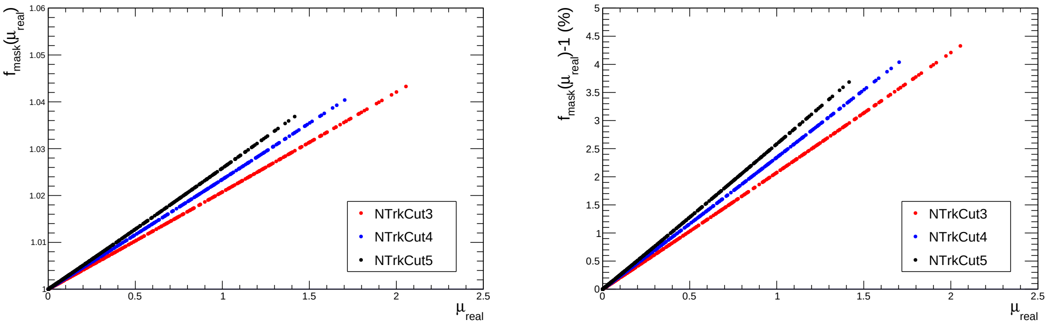
(c) x-scan, BCID 3121 (d) y-scan, BCID 3121

Figure 5.21: $\mu_{\text{vis}}^{\text{sp}}$ vs. beam separation with double Gaussian fits and residuals for BCIDs 2881 and 3121 of the April vdM scans.



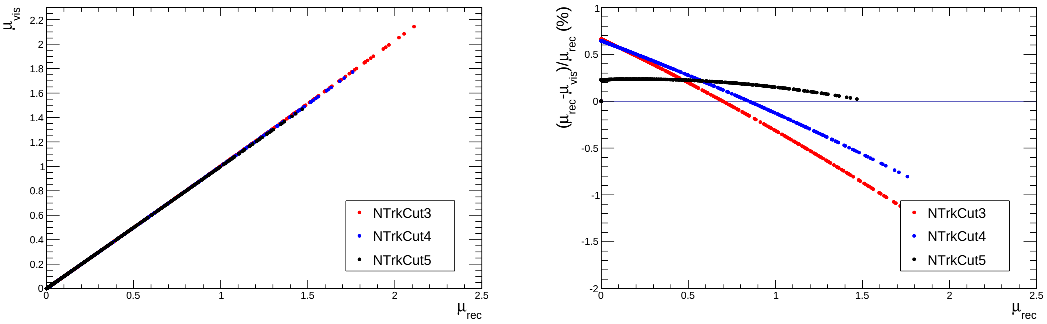
(a) Fake correction.

(b) Fake correction as a percentage.



(c) Masking correction.

(d) Masking correction in percentage.



(e) Overall correction.

(f) Overall correction in percentage.

Figure 5.22: Pile-up corrections applied to the April vdM scans, for vertices with NTrkCut 3, 4 and 5.

5.6.2 July 2012

Figure 5.23 shows the performed fits using a double Gaussian plus constant fit function for the three recorded BCIDs using vertices with NTrkCut 3, 4 or 5. The σ_{vis} values are shown in Table 5.6 for all scans and BCIDs.

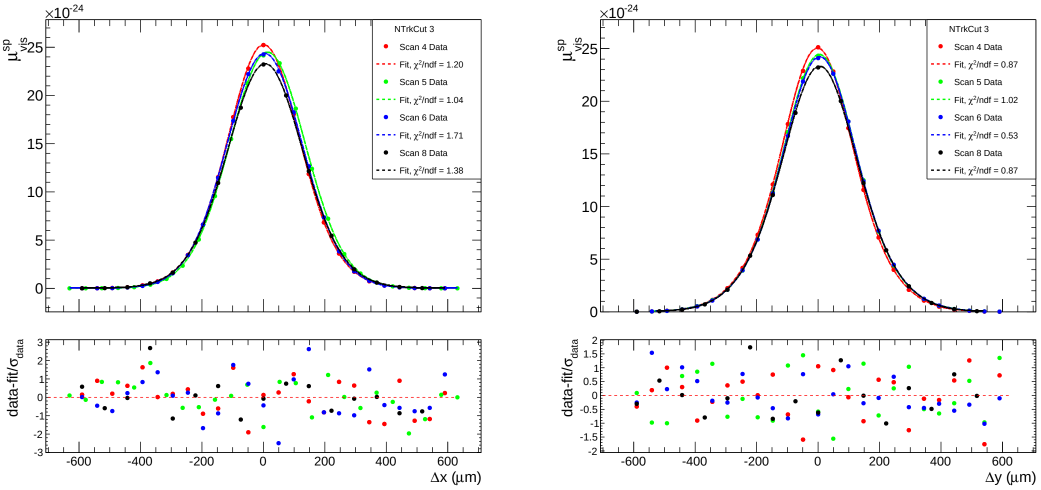
Figure 5.24 shows the corrections applied to the July vdM scans for the three different values of NTrkCut; the absolute correction value on the left and the correction percentage on the right. This figure shows that the fake correction can be as big as 0.9%, the masking correction can be as big as 0.45% and the overall correction as big as 0.5%.

		σ_{vis} [mb]		
Scan	BCID	NTrkCut 3	NTrkCut 4	NTrkCut 5
4	1	24.238 ± 0.045	19.783 ± 0.041	16.388 ± 0.037
	721	24.075 ± 0.037	19.635 ± 0.033	16.292 ± 0.03
	1821	24.176 ± 0.045	19.727 ± 0.041	16.36 ± 0.037
5	1	24.864 ± 0.048	20.285 ± 0.043	16.814 ± 0.039
	721	24.743 ± 0.04	20.204 ± 0.037	16.748 ± 0.033
	1821	24.743 ± 0.048	20.164 ± 0.044	16.72 ± 0.039
6	1	24.639 ± 0.048	20.068 ± 0.043	16.627 ± 0.039
	721	24.67 ± 0.04	20.148 ± 0.037	16.693 ± 0.033
	1821	24.618 ± 0.048	20.059 ± 0.043	16.629 ± 0.039
8	1	24.241 ± 0.058	19.791 ± 0.053	16.419 ± 0.048
	721	24.228 ± 0.054	19.8 ± 0.05	16.421 ± 0.045
	1821	24.131 ± 0.068	19.737 ± 0.062	16.387 ± 0.056

Table 5.6: Visible cross sections for vdM scans 207216 and 207219, determined using a double Gaussian plus a constant fit function. The uncertainties are statistical only.

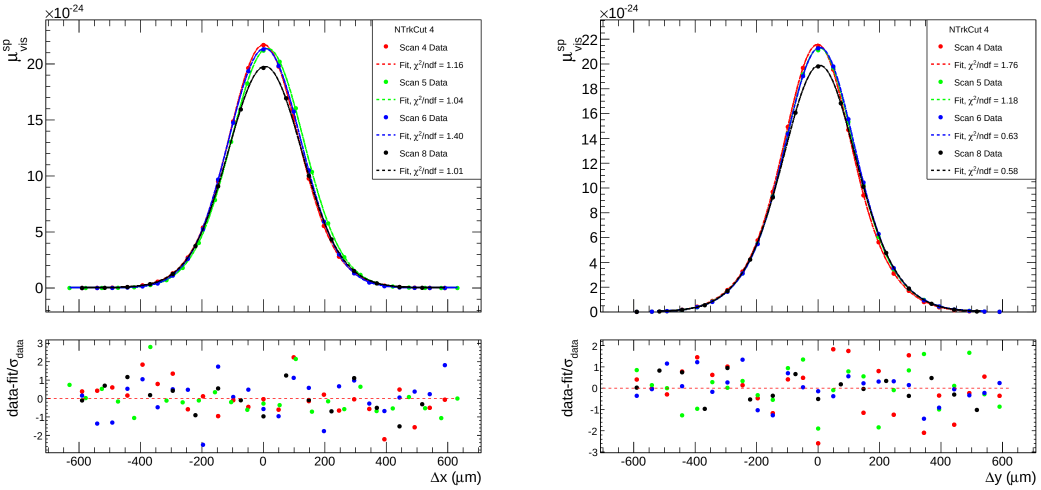
	σ_{vis} [mb]		
Scan	NTrkCut 3	NTrkCut 4	NTrkCut 5
4	24.169 ± 0.025	19.721 ± 0.023	16.35 ± 0.021
5	24.786 ± 0.027	20.218 ± 0.024	16.761 ± 0.022
6	24.641 ± 0.027	20.089 ± 0.024	16.647 ± 0.022
8	24.196 ± 0.036	19.773 ± 0.032	16.408 ± 0.029

Table 5.7: Visible cross sections for vdM scans 207216 and 207219 averaged over BCIDs, determined using a double Gaussian plus a constant fit function. The uncertainties are statistical only.



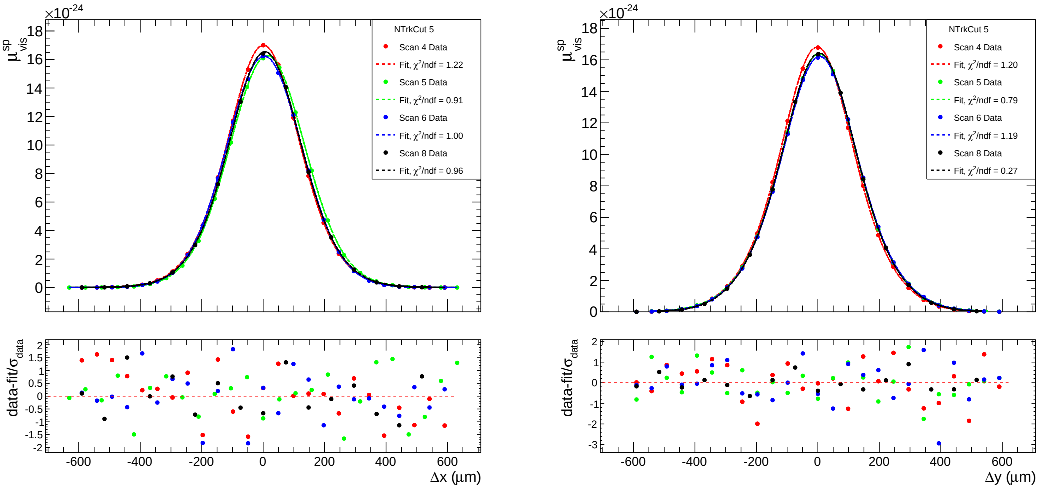
(a) x-scan, BCID 1

(b) y-scan, BCID 1



(c) x-scan, BCID 721

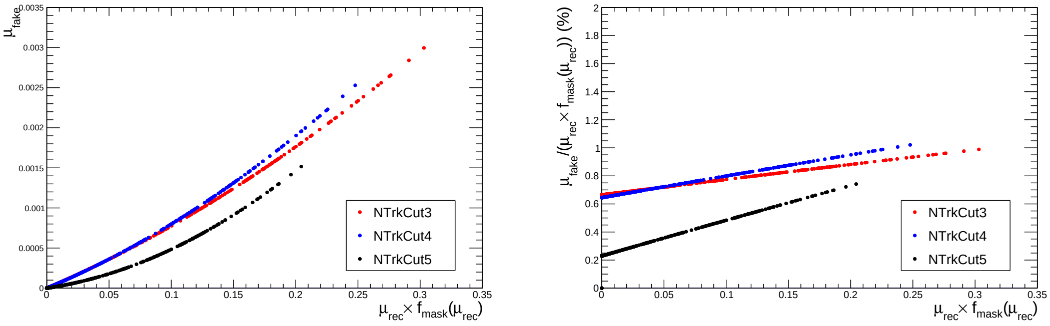
(d) y-scan, BCID 721



(e) x-scan, BCID 1821

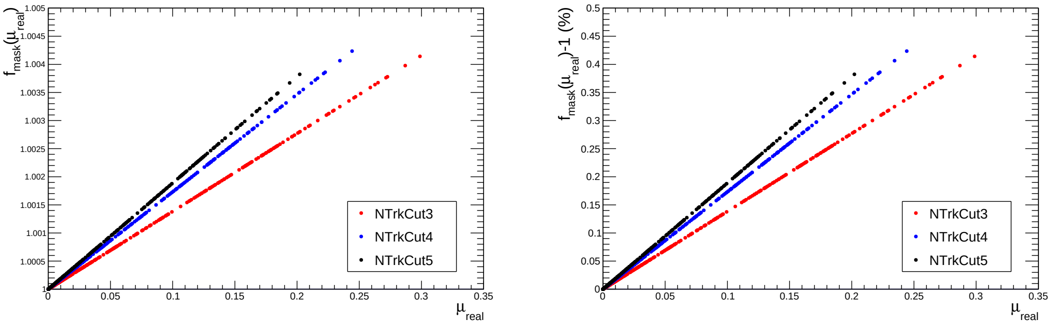
(f) y-scan, BCID 1821

Figure 5.23: $\mu_{\text{vis}}^{\text{sp}}$ vs. beam separation with double Gaussian fits and residuals for BCIDs 1, 721 and 1821 of the July vdM scans.



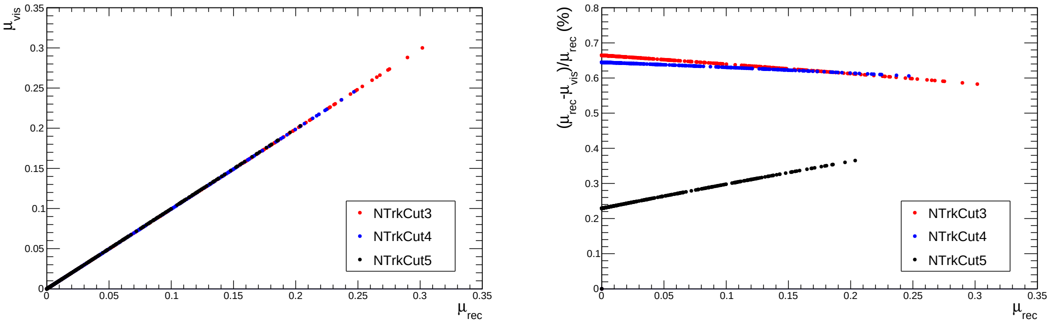
(a) Fake correction.

(b) Fake correction as a percentage.



(c) Masking correction.

(d) Masking correction in percentage.



(e) Overall correction.

(f) Overall correction in percentage.

Figure 5.24: Pile-up corrections applied to the July vdM scans, for vertices with NTrkCut 3, 4 and 5.

5.6.3 November 2012

Figure 5.25 shows the performed fits using a single Gaussian plus constant fit function for the three recorded BCIDs using vertices with NTrkCut 3, 4 or 5. The σ_{vis} values are shown in Table 5.8 for all scans and BCIDs.

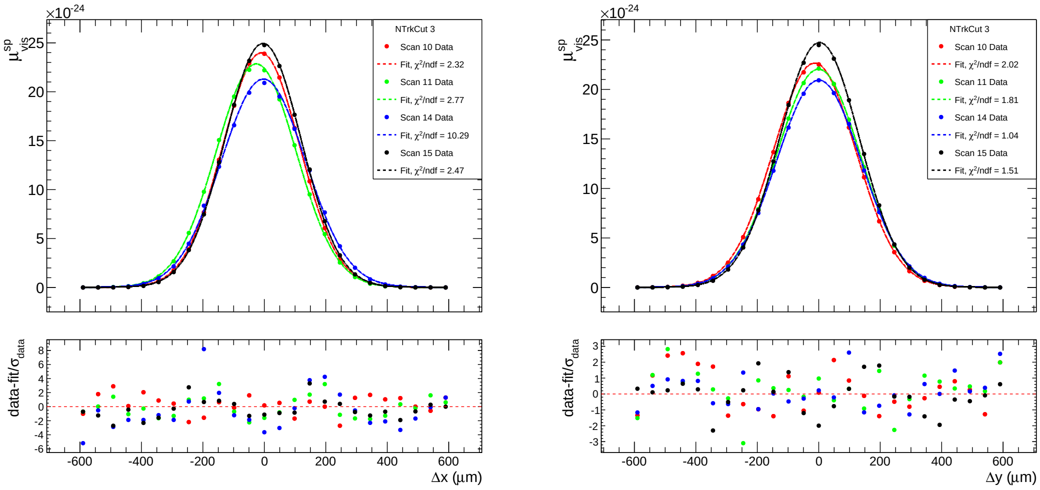
Figure 5.26 shows the corrections applied to the November vdM scans for the three different values of NTrkCut; the absolute correction value on the left and the correction percentage on the right. This figure shows that the fake correction can be as big as 0.85%, the masking correction can be as big as 0.45% and the overall correction as big as 0.6%.

		σ_{vis} [mb]		
Scan	BCID	NTrkCut 3	NTrkCut 4	NTrkCut 5
10	1	24.66 ± 8e-05	20.105 ± 9.9e-05	16.669 ± 0.00012
	2361	24.632 ± 5.7e-05	20.127 ± 6.9e-05	16.672 ± 8.4e-05
	2881	24.698 ± 0.00012	20.166 ± 4.4e-05	16.72 ± 0.00017
11	1	25.075 ± 0.0003	20.456 ± 0.00038	16.954 ± 0.00013
	2361	24.983 ± 6.9e-05	20.384 ± 0.00026	16.904 ± 9.7e-05
	2881	24.936 ± 4e-05	20.366 ± 4.9e-05	16.876 ± 0.00019
14	1	25.184 ± 0.00034	20.556 ± 0.00014	17.045 ± 0.00016
	2361	25.306 ± 0.0003	20.649 ± 0.00036	17.113 ± 0.00044
	2881	25.159 ± 5.4e-05	20.558 ± 0.0002	17.044 ± 0.00024
15	1	25.217 ± 7.4e-05	20.589 ± 9e-05	17.06 ± 0.00033
	2361	25.151 ± 4.1e-05	20.531 ± 5e-05	17.016 ± 6.1e-05
	2881	25.16 ± 0.00017	20.531 ± 0.00021	17.037 ± 7.8e-05

Table 5.8: Visible cross sections for vdM scans 214984 and 215021, determined using a single Gaussian plus a constant fit function. The uncertainties are statistical only.

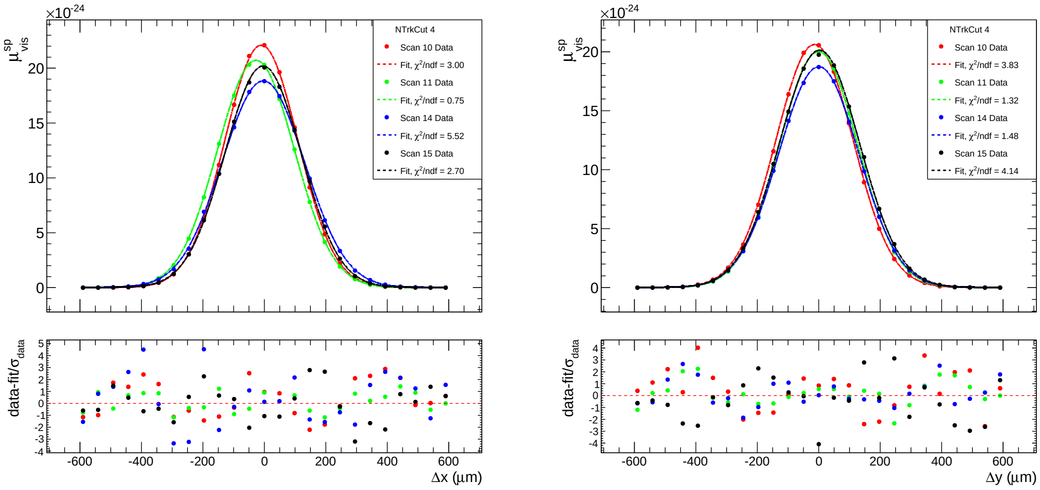
		σ_{vis} [mb]	
Scan	NTrkCut 3	NTrkCut 4	NTrkCut 5
10	24.671 ± 6e-05	20.125 ± 5.2e-05	16.693 ± 8.8e-05
11	25.046 ± 0.00022	20.423 ± 0.00023	16.907 ± 9.6e-05
14	25.234 ± 0.00021	20.605 ± 0.0002	17.081 ± 0.00025
15	25.174 ± 0.0001	20.546 ± 0.00013	17.05 ± 0.00023

Table 5.9: Visible cross sections for vdM scans 214984 and 215021 averaged over BCIDs, determined using a single Gaussian plus a constant fit function. The uncertainties are statistical only.



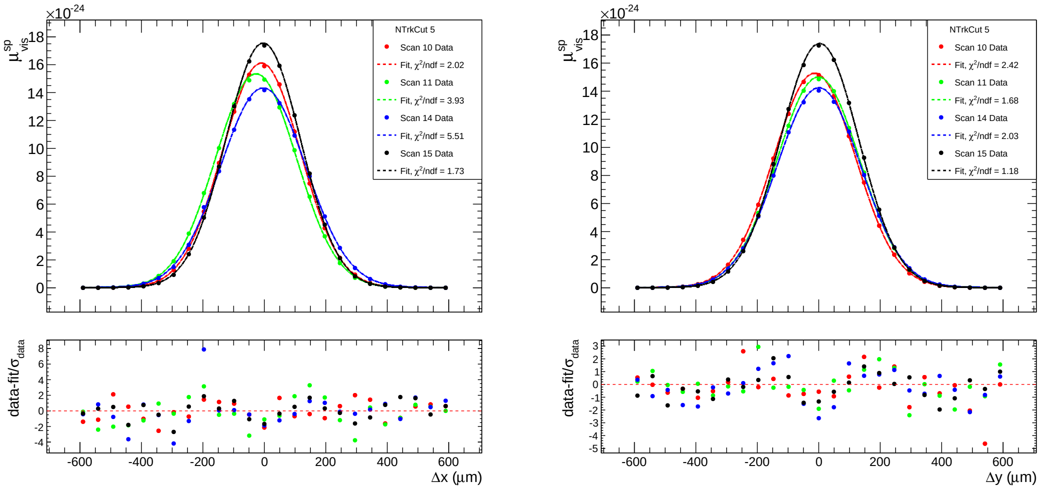
(a) x-scan, BCID 1

(b) y-scan, BCID 1



(c) x-scan, BCID 2361

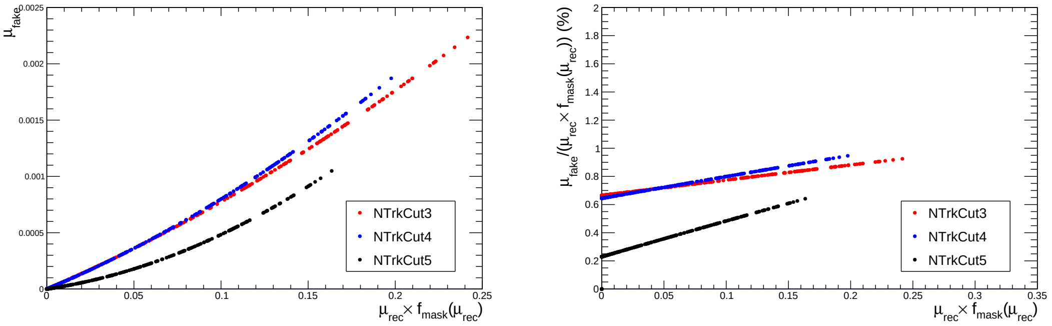
(d) y-scan, BCID 2361



(e) x-scan, BCID 2881

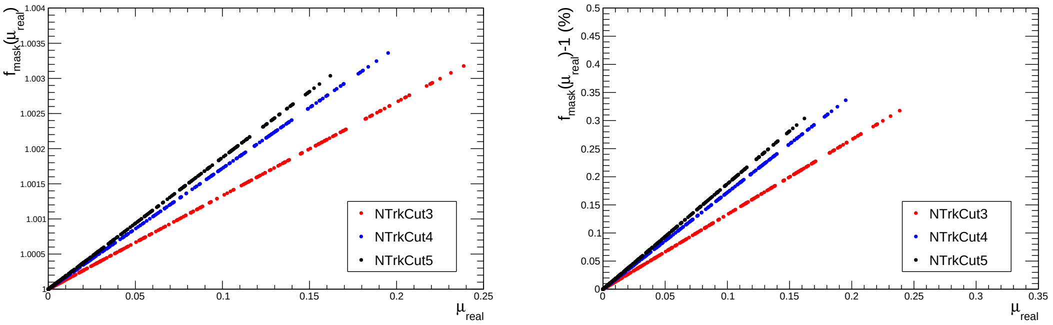
(f) y-scan, BCID 2881

Figure 5.25: $\mu_{\mathrm{vis}}^{\mathrm{sp}}$ vs. beam separation with single Gaussian fits and residuals for BCIDs 1, 2361 and 2881 of the November vdM scans.



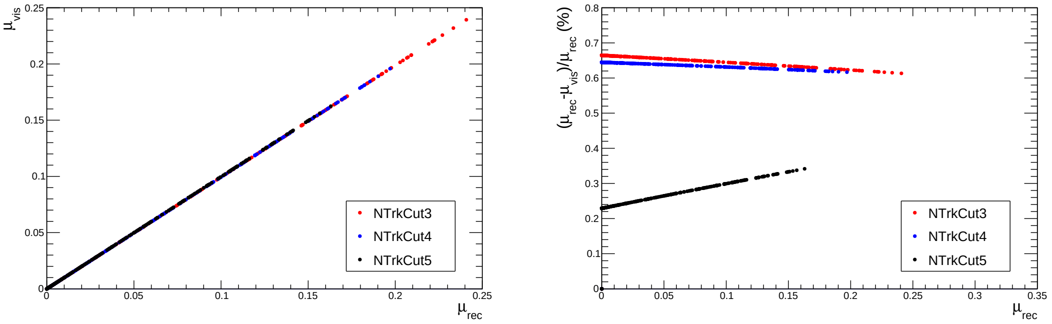
(a) Fake correction.

(b) Fake correction as a percentage.



(c) Masking correction.

(d) Masking correction in percentage.



(e) Overall correction.

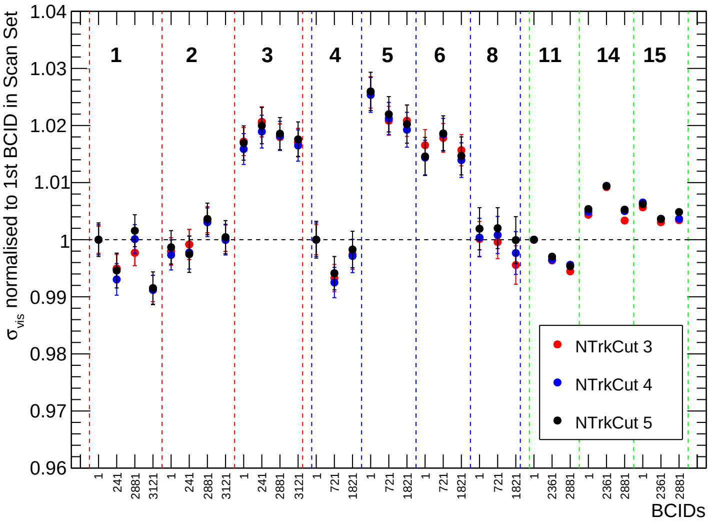
(f) Overall correction in percentage.

Figure 5.26: Pile-up corrections applied to the November vdM scans, for vertices with NTrkCut 3, 4 and 5.

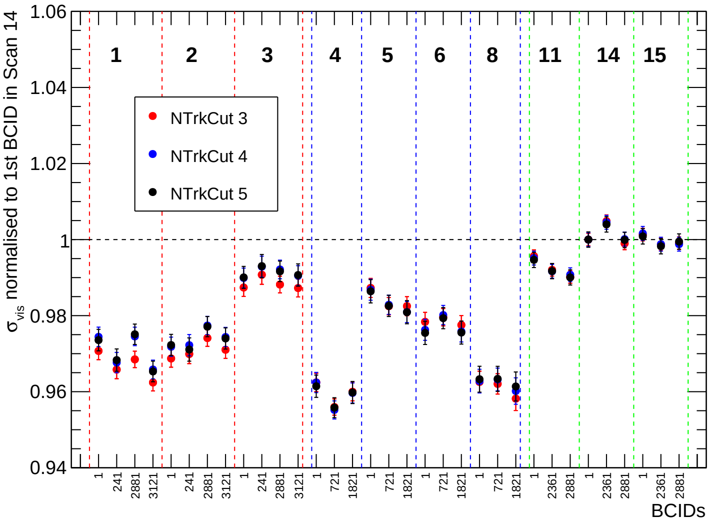
5.6.4 Comparison within scan sets and to other algorithms

Figure 5.27 shows the σ_{vis} values for every BCID of the 2012 vdM scans obtained using the vertex counting algorithm with three different values of NTrkCut. Scan 10 has been excluded from this figure since this scan has been found to suffer from some technical issues not related with the vertex counting algorithm, namely a large orbit-drift and emittance growth due to the long delay between the performance of the x- and y-scans [35]. On the top plot, the values are normalised to the σ_{vis} value of the first BCID in every scan set, i.e., to BCID 1 of scan 1, BCID 1 of scan 4 and

BCID 1 of scan 11. On the bottom plot, the values are normalised to the σ_{vis} value of the first BCID of scan 14, which is the scan that has been chosen as a reference by the ATLAS luminosity task force. The red vertical lines denote the April scan set, the blue vertical lines the July scan set and the green vertical lines the November scan set.



(a) Normalised to scan set.



(b) Normalised to 1st BCID in scan 14.

Figure 5.27: σ_{vis} values for all 2012 vdM scans using the vertex counting algorithm and vertices with NTrkCut 3, 4 and 5. On the top plot the values have been normalised to the first BCID of every scan set and on the bottom plot the values have been normalised to the first BCID of scan 14.

Figure 5.27 shows that the results obtained using different values of NTrkCut are consistent with each other. It also shows that the variation of the σ_{vis} values within a scan is around 0.5%, within a scan set is around 2% and there is a 2% difference between scan sets.

Figure 5.28 shows a comparison of the visible cross-section of the 2012 vdM scans extracted using different algorithms. The only correction applied to these values is the centring correction. Scan 10 has been added to this plot since the technical issues affecting this scan are expected to cancel out when looking at the ratios from different algorithms. The values presented for the vertex counting algorithm are those extracted using vertices with NTrkCut 5. Similar plots for vertices with NTrkCut 3 and 4 can be seen in the Appendix A.1.

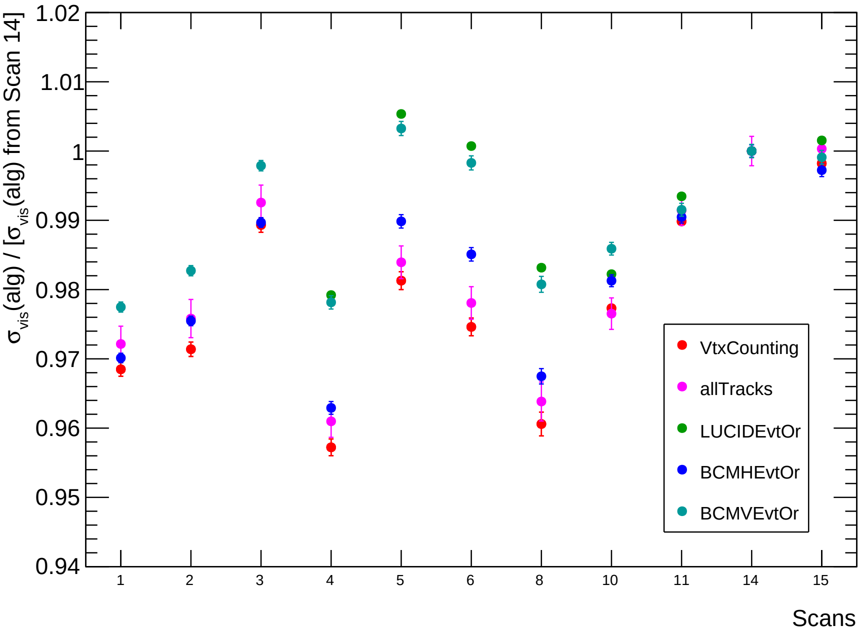
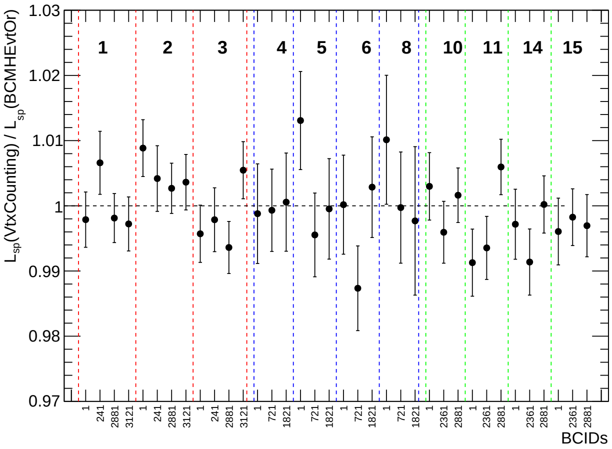


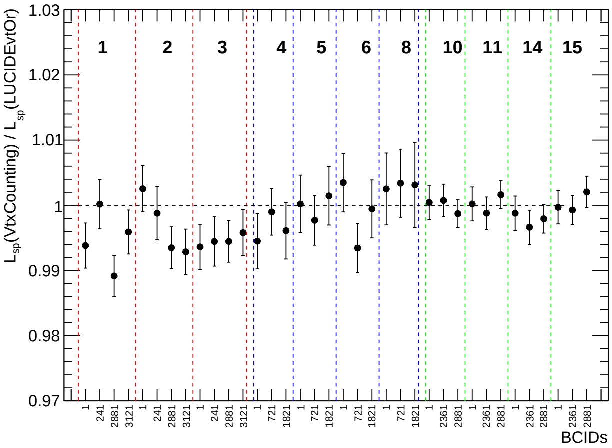
Figure 5.28: σ_{vis} values for all 2012 vdM scans from different algorithms, normalised to scan 14 (vertex counting value using vertices with NTrkCut 5).

In this comparison, all the σ_{vis} values have been normalised to the value from scan 14, in November, to study the long-term stability. It is clear that the vertex counting algorithm agrees well with other algorithms, especially with track counting (*allTracks*) and all of them are quite stable within a scan set. As pointed out before for the vertex counting algorithm, there is a difference of around 2% between the July and November scan sets and also between the April and November scan sets. It can be seen in this figure that this is also the case for the other algorithms. This difference is largely due to the non-factorisation correction that has to be applied to the σ_{vis} values obtained from the April and July scan sets; the van der Meer method assumes that the particle densities in each bunch can be factorised into independent horizontal and vertical components but this assumption was not entirely correct for some of the 2012 vdM scans sets, especially for April and July. The ATLAS luminosity task force performed different studies to assess to what extent this assumption of factorisation held, and it was found that neglecting non-factorization effects in the vdM calibration leads to overestimating the absolute luminosity scale (or equivalently underestimating the visible cross-section) by up to 3% in April and 4.5% in July. For November, however, the non-factorization biases remain below 0.8%, thanks to bunch-tailoring in the LHC injector chain [36]. On top of this effect, subsequent analysis showed that BCM and LUCID suffered a drift at the 1 to 2% level over the year due to, respectively, radiation-induced lattice defects and PMT ageing.

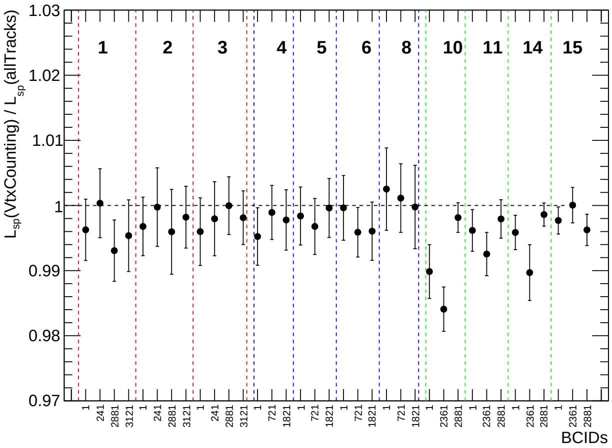
Finally, Figure 5.29 shows a comparison of the specific luminosity from vertex counting and three other selected algorithms. Given that the specific luminosity is a purely geometrical quantity of the beams, it provides a useful comparison between luminosity algorithms. This figure shows that the consistency is mostly within 1%.



(a) Comparison to BCMHEvtOR.



(b) Comparison to LUCIDEvtOR.



(c) Comparison to allTracks.

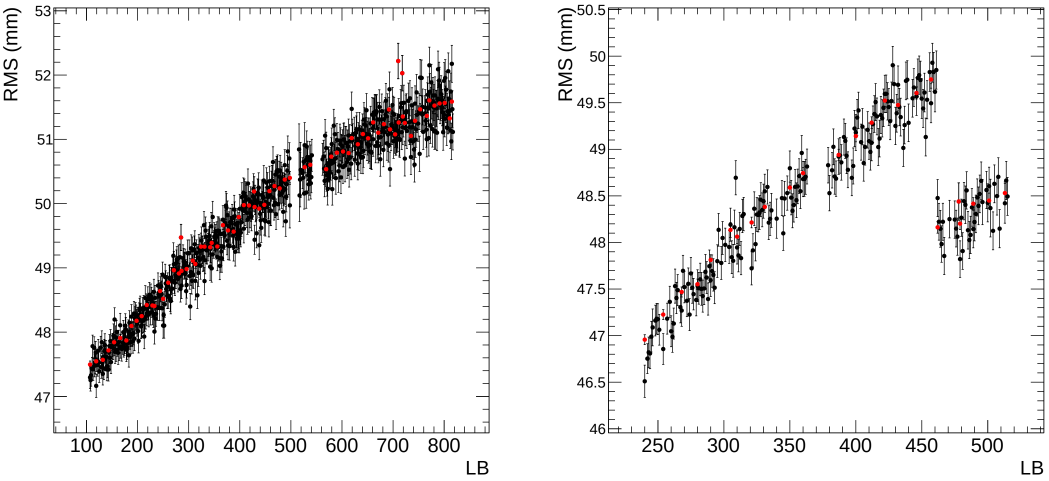
Figure 5.29: Specific luminosity comparison between vertex counting (NTrkCut 5) and three selected algorithms, for all BCIDs and all scans. The values are consistent within 1% for most cases.

5.7 Results for the 2012 Physics Runs

The vertex counting algorithm was used to analyse several physics runs performed in 2012. All the runs have 1368 BCIDs and span different ranges of μ. The first run analysed was Run 206962, which occurred on the 14th of July 2012, and the last run analysed was Run 215541, performed on the 2nd of December 2012. The vdM Run 207616, which corresponds to the vdM scans 4, 5 and 6 performed in July 2012, was used for calibration.

It is important to notice that within a single physics run, the luminous region width changes and this will affect the masking correction evaluation. The change in the luminous region width for two particular physics runs can be seen in Figure 5.30, where the black points correspond to the RMS of the z-distribution for every LB. This effect has been addressed by generating an "expected" Δz-distribution and masking correction look-up graphs for every group of 10 consecutive LBs. There is also a requirement that the LBs to be combined have a roughly consistent Δz-distribution RMS.

The red points in Figures 5.30 correspond to the first LB of the subset of LBs grouped together and the RMS of the z-distribution resulting from grouping those LBs together, for physics runs 206962 and 215541. The same procedure was applied to all the analysed physics runs.

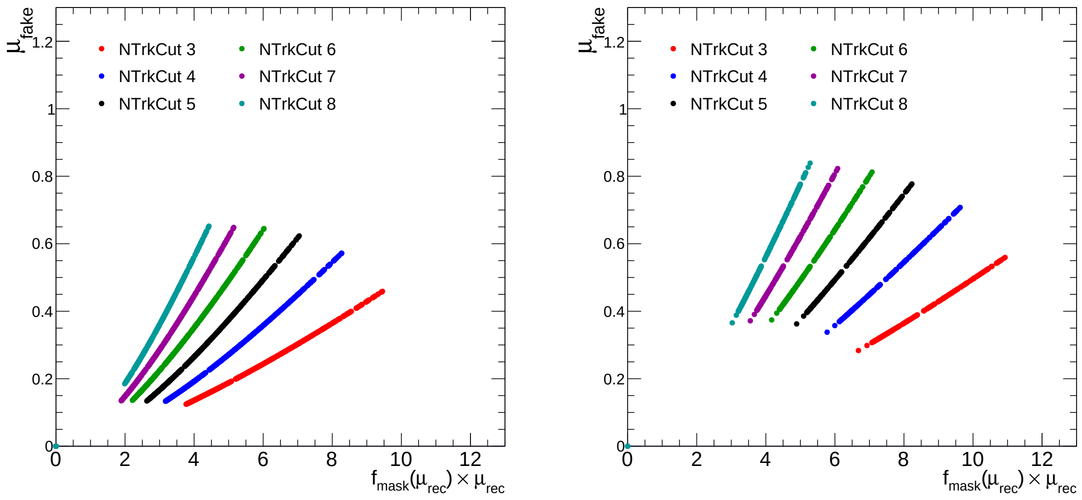


(a) Run 206962, NTrkCut 5.

(b) Run 215541, NTrkCut 5.

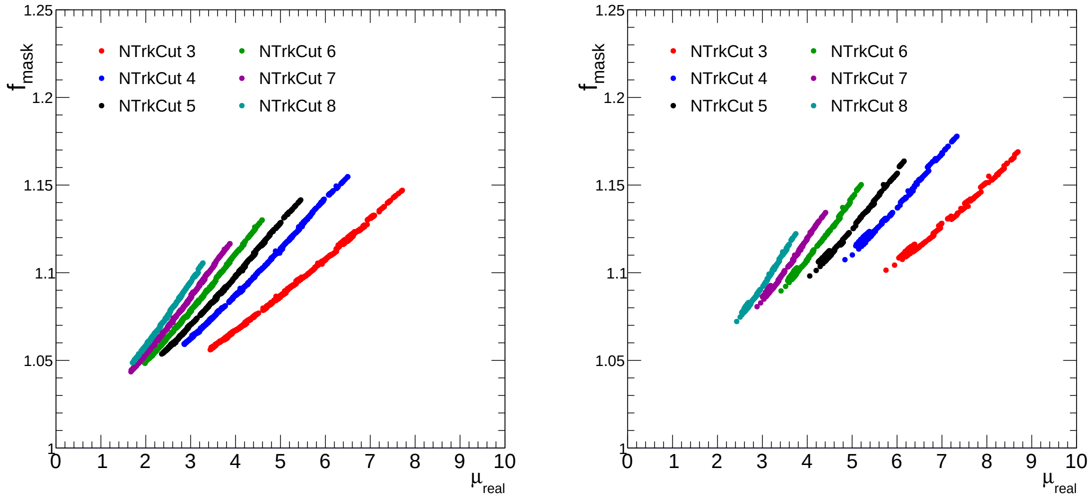
Figure 5.30: Black points correspond to the RMS of z-distribution vs LB for physics runs 206962 and 215541 and NTrkCut 5. Red points correspond to the first LB of the subset of LBs grouped together and the RMS of the z-distribution resulting from grouping those LBs together.

Figure 5.31 shows the fake and masking corrections for physics runs 206962 and 215541. These figures show that the fake correction ranges from 5 to 16% and the masking correction ranges from 10 to 20%. They also show how these corrections vary with μ and with NTrkCut values. Similar behaviour was shown by the rest of the analysed physics runs.



(a) Fake correction for Run 206962.

(b) Fake correction for Run 215541.



(c) Masking correction for Run 206962.

(d) Masking correction Run 215541.

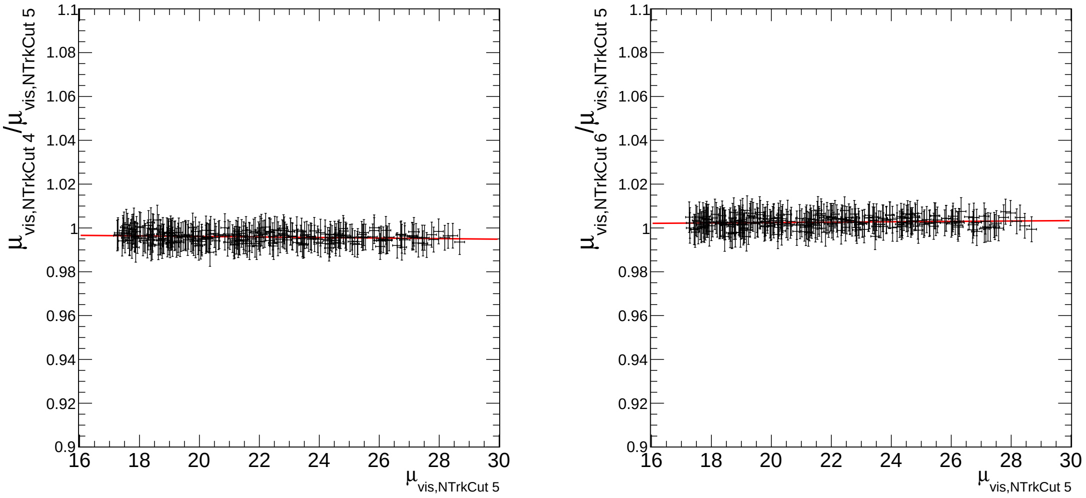
Figure 5.31: Fake and masking correction plots for physics runs 206962 and 215541 for different values of NTrkCut. The top plots are the fake correction plots and they show μ_{fake} as a function of $f_{mask}(\mu_{rec}) \times \mu_{rec}$. The bottom plots are the masking correction plots and they show the masking correction factor f_{mask} as a function of μ_{real}. The fake correction can be as big as 16% and the masking correction can be as big as 20%.

The next sections present a study on the internal consistency of the vertex counting method by comparing the results obtained using different cuts on the minimum number of tracks per vertex. The external consistency is also studied by comparing the luminosity measurements performed with the vertex counting algorithm to the luminosity measurements obtained by other detectors and algorithms.

5.7.1 Internal Consistency

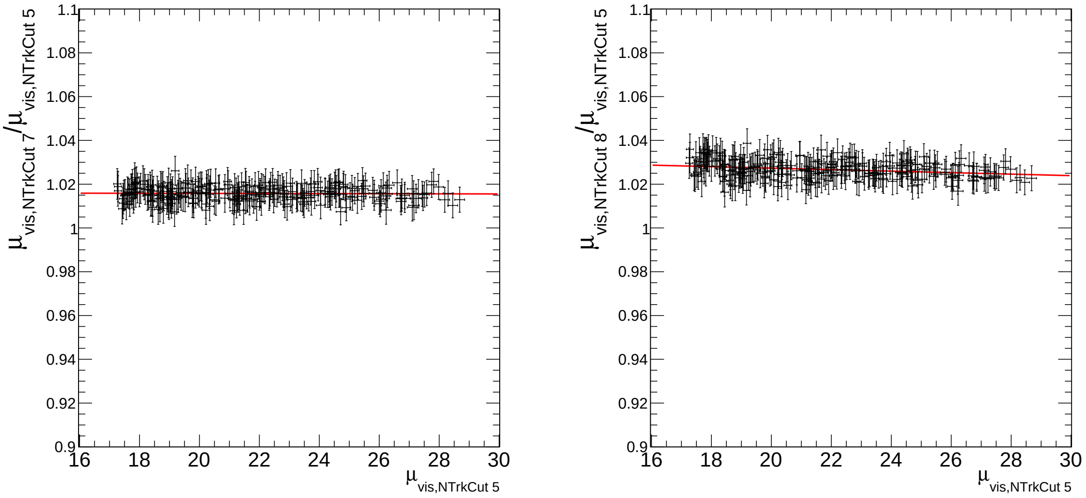
The next figures show the ratio of μ_{vis} from different values of NTrkCut to NTrkCut 5 as a function $\mu_{vis,NTrkCut5}$ for the physics run 207589. A straight line is fitted to the curves to help visualise the consistency of the results. Figure 5.32 shows that the results for vertices with NTrkCut 4, 6, 7 and 8 are all very consistent with the results for vertices with NTrkCut 5; it is easy to see that the plots are very flat; the slopes are practically zero and the intercept is very close to 1. The rest of the analysed runs show a similar behaviour. Given that the pile-up corrections are very different for the different values of NTrkCut, this is a very encouraging result and proves that the vertex counting algorithm has internal consistency.

Unfortunately, the case for NTrkCut 3 is different. Figure 5.33a shows that the ratio of μ_{vis} from vertices with NTrkCut 3 to μ_{vis} from vertices with NTrkCut 5 has a dependency on μ_{vis}. This means that fixing the masking correction evaluation for vertices with NTrkCut 3 is not enough and there must be other mechanisms affecting this particular definition of a vertex. Further and more detailed studies would be necessary to fully understand this NTrkCut. However, as part of the systematic uncertainties study, the analysis was performed using the p_{mask} vs. Δz distribution from NTrkCut 4 vertices on NTrkCut 3 vertices and it was found that this improves the consistency between NTrkCut 3 and 5, as can be seen in Figure 5.33b, so this would be a study worth developing in any subsequent vertex counting analysis.



(a) NTrkCut 4 vs NTrkCut 5.

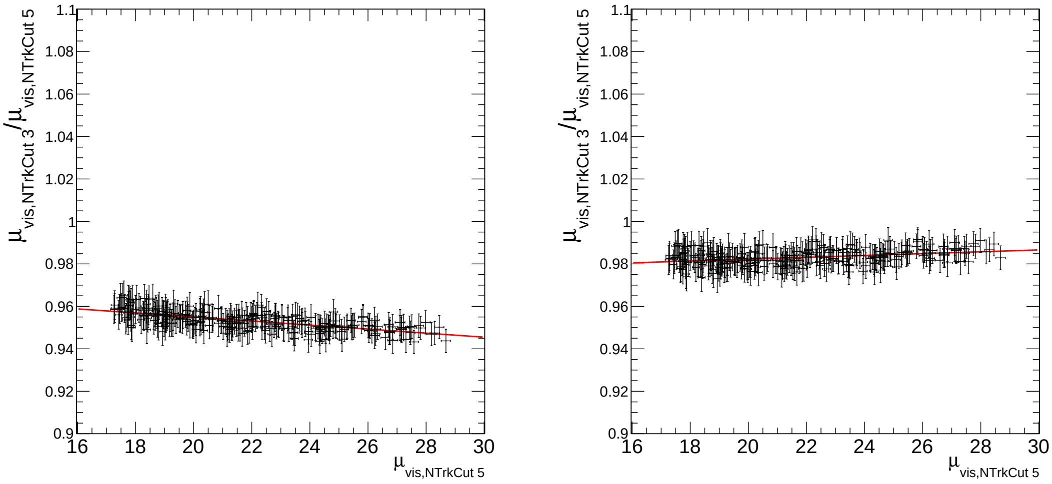
(b) NTrkCut 6 vs NTrkCut 5.



(c) NTrkCut 7 vs NTrkCut 5.

(d) NTrkCut 8 vs NTrkCut 5.

Figure 5.32: Ratio of μ from NTrkCut 4, 6, 7 and 8 to NTrkCut 5 for physics run 207589.



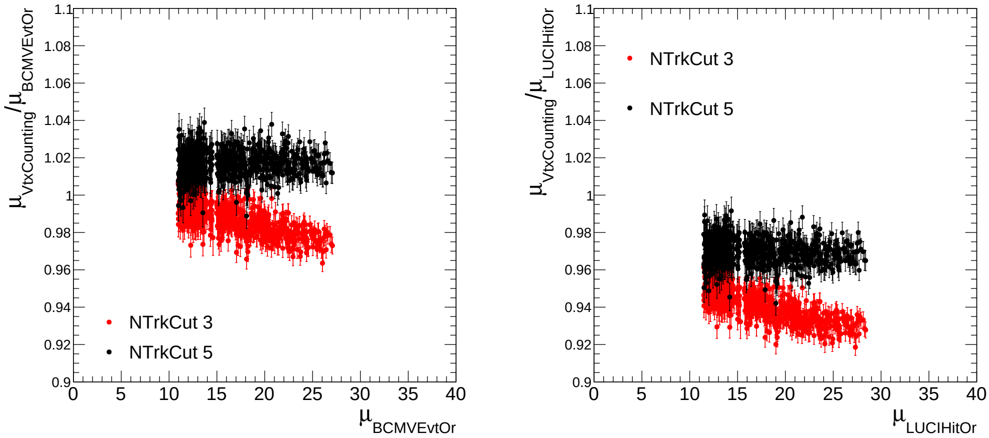
(a) (b)

Figure 5.33: Ratio of μ_{vis} from vertices with NTrkCut 3 to vertices with NTrkCut 5 for physics Run 207589. (a) The p_{mask} vs. Δz distribution from NTrkCut 3 was used on NTrkCut 3. (b) The p_{mask} vs. Δz distribution from NTrkCut 4 was used on NTrkCut 3.

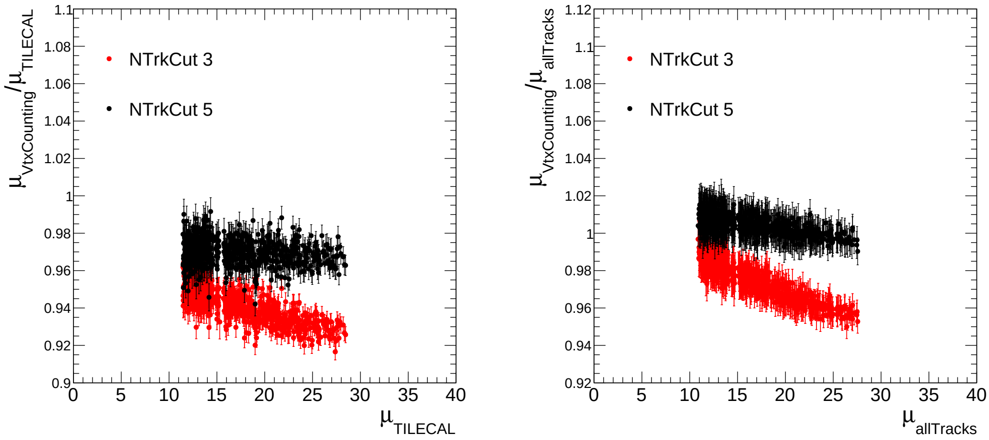
5.7.2 External Consistency

The next figures show the ratio of μ_{vis} obtained using the vertex counting algorithm to the μ_{vis} obtained by other detectors and algorithms for the physics runs 206962 and 215541. The calibration used by the other algorithms is the one quoted by the luminosity group in preparation for the Moriond 2013 conference. Only the results from the vertex counting algorithm using NTrkCut 3 and 5 are being used for comparison, since every other NTrkCut is consistent with NTrkCut 5.

Figures 5.34 and 5.35 show the comparison of the vertex counting algorithm to BCMVEvtOR, LUCIDHitOR, TileCal and allTracks for the physics runs 206962 and 215541, the first and last runs on the list of analysed runs; the rest of the analysed physics runs present a similar behaviour. These figures show that a μ dependency still exists but, for NTrkCut 5 vertices, it is quite small, around 0.1%, which indicates a great improvement for the vertex counting algorithm. A similar behaviour was found on the other runs analysed, therefore the conclusion is that the vertex counting algorithm is consistent with other detectors and algorithms when using an NTrkCut value in the range 4 to 8. This is not the case when using NTrkCut 3. However, many reasons have been found that indicate that this particular value of NTrkCut would not yield results consistent with other algorithms.

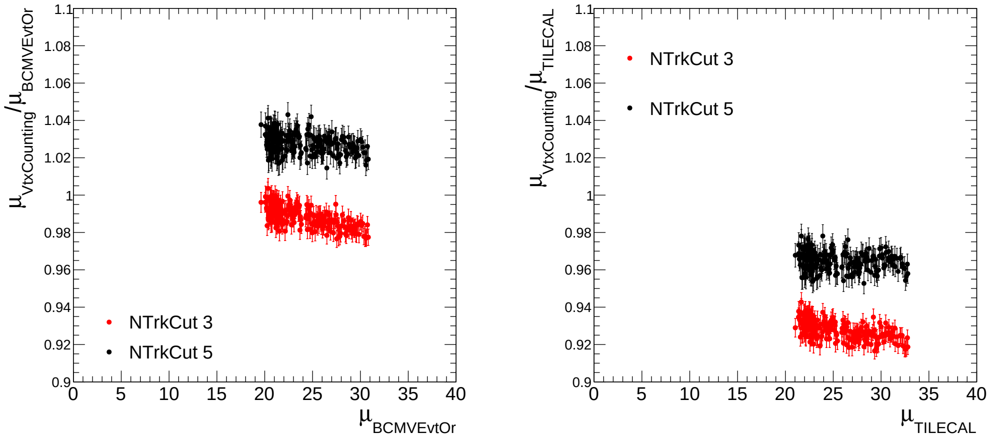


(a) Vertex Counting vs. BCMVEvtOR. (b) Vertex Counting vs. LUCIDHitOR.

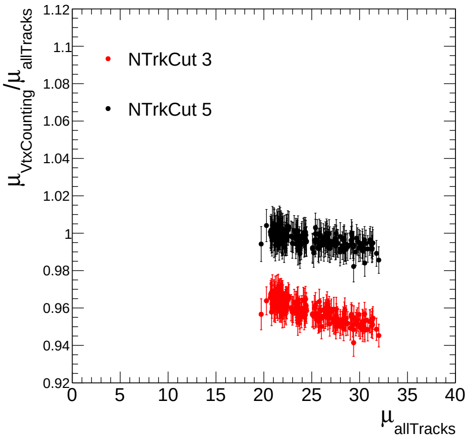


(c) Vertex Counting vs. TILECAL. (d) Vertex Counting vs. vtxTracks.

Figure 5.34: Ratio of μ_{vis} from different detectors and algorithms to vertex counting with NTrkCut 3 and 5, for physics Run 206962.



(a) Vertex Counting vs. BCMVEvtOR. (b) Vertex Counting vs. TILECAL.



(c) Vertex Counting vs. vtxTracks.

Figure 5.35: Ratio of μ_{vis} from different detectors and algorithms to vertex counting with NTrkCut 3 and 5, for physics Run 215541.

5.8 Systematic uncertainties

The systematic uncertainties listed below are internal to the vertex counting algorithm. However, there are other systematic uncertainties which affect the calibration of the absolute luminosity of every luminosity algorithm that uses the van der Meer method to be calibrated. These systematic uncertainties affect these algorithms equally and are due mainly to instrumental effects and conditions of the beams. They can be found in Table 5.10.

- **Split vertices:** One of the pile-up effects affecting the vertex counting algorithm is the one due to split vertices. However, it was found that their contribution is quite small (see Figure 5.4) and therefore no correction is applied to account for this effect. Given the size of their contribution, a 0.1% systematic uncertainty is assign to the vertex counting results due to not applying any correction for split vertices.

- **p_{mask} vs Δz for NTrkCut 3:** When evaluating the masking correction for vertices with NTrkCut 3, it was found that the standard procedure was failing for very close-by vertices. Therefore, for pairs of vertices with $|\Delta z| < 1\,\mathrm{mm}$, the masking probability, p_{mask}, was set to 1. This cut on Δz was somewhat arbitrary and, to account for this, a comparison was made of the results of applying a bigger cut, $|\Delta z| < 1.75\,\mathrm{mm}$. The average difference on the luminosity measurements for the physics runs coming from these two methods is 0.24%, which is assigned as a systematic uncertainty to vertices with NTrkCut 3.

- **p_{mask} vs Δz on NTrkCut 4:** The effects of manually setting p_{mask} to 1 were also tested by applying this procedure to vertices with NTrkCut 4. In this case, p_{mask} was set to 1 for pairs of vertices with $|\Delta z| < 1.5\,\mathrm{mm}$. The average difference on the luminosity measurements for the physics runs coming from using or not this procedure on NTrkCut 4 is 0.14%, which is assigned as a systematic uncertainty to the measurements for every NTrkCut, except NTrkCut 3.

- **Fake correction and the luminous region width:** Is it clear from Figures 5.5a to 5.5d that the contribution from fake vertices depends on NTrkCut. However, the fake correction should also depend on the width of the luminous region. Given that the MC samples used in this analysis were generated with a particular beam spot width, the error due to different beam spot lengths between data and MC, which can be seen in Figure 5.2, was assessed by applying an extra correction factor to μ_{fake}:

  $$\mu'_{fake} = \frac{RMS(MC)}{RMS(data)}\mu_{fake} \tag{5.8}$$

  Applying this extra correction factor yields an average 0.23% difference with respect to the standard method, which is assigned as a systematic uncertainty of the method.

- **μ dependence:** The comparisons of the vertex counting to other algorithms during physics runs suggest that there may be some residual μ dependence in the vertex counting algorithm. For vertices with NTrkCut 4 and above, a 1%

systematic uncertainty is assigned using the typical change from beginning to end of a physics run, due to the slope relative to TileCal. For vertices with NTrkCut 3, a systematic uncertainty of 3% is assigned in a similar way but also taking into account the absolute discrepancy in physics runs with respect to the vdM calibration.

Source	Uncertainty [%]
Reference specific luminosity	0.50
Noise and background subtraction	0.30
Length-scale calibration	0.40
Absolute ID length scale	0.30
Subtotal, instrumental effects	0.77
Orbit drifts	0.10
Beam-position jitter	0.20
Beam–beam corrections	0.28
Fit model	0.50
Non-factorization correction	0.50
Emittance-growth correction	0.10
Bunch-by-bunch σ_{vis} consistency	0.23
Scan-to-scan consistency	0.31
Subtotal, beam conditions	0.89
Bunch-population product	0.24
Total	1.20

Table 5.10: Fractional systematic uncertainties affecting the visible cross-section averaged over vdM scan sets 11-15 (November 2012). Taken from Ref. [35].

5.9 Integrated luminosity comparison

The list of physics runs analysed with the vertex counting algorithm and the two track counting algorithms is: 206962, 207589, 207620, 207664, 215433, 215456, 215473 and 215541. This is a small subset of the physics runs recorded in 2012 but contains runs taking place between July and December of 2012. The integrated luminosity has been calculated for these runs and for the three different algorithms mentioned before and the ratio of the vertex counting results to the track counting results has been calculated to be:

$$\frac{\int \mathcal{L}_{\text{vertex counting, NtrkCut5}}}{\int \mathcal{L}_{\text{vtxTracks}}}(\text{test runs}) = 1.016 \pm 0.009(\text{stat.}) \pm 0.011(\text{vertex syst.}) \tag{5.9}$$

$$\frac{\int \mathcal{L}_{\text{vertex counting, NtrkCut5}}}{\int \mathcal{L}_{\text{allTracks}}}(\text{test runs}) = 1.004 \pm 0.009(\text{stat.}) \pm 0.010(\text{vertex syst.}) \tag{5.10}$$

Therefore, the integrated luminosity from these test runs calculated using the vertex counting algorithm is consistent with the integrated luminosity calculated using either of the track counting algorithms.

5.10 Conclusion and future prospects

The vertex counting algorithm has been updated to the requirements of the 2012 data recorded by the ATLAS detector. This method and its corrections for pile-up effects were successfully validated through a MC closure test for different definitions of *tight* vertices. The vertex counting algorithm was calibrated using the 2012 vdM scans and the values of the visible cross-section for these scans were presented in this chapter. It was shown that the method is stable within a scan set at the 0.5% level, and the difference between scan sessions is around 2%. The method was also found to be consistent with other detectors and algorithms; when comparing the specific luminosity, the consistency between the vertex counting and other algorithms was at the 1% level. Finally, this algorithm was used to measure the luminosity in 2012 physics runs and excellent internal consistency was found, by comparing the results obtained with different definitions of vertices, and also good external consistency, by comparing with other detectors and algorithms. In particular, the integrated luminosity from a few test physics runs from 2012 calculated with the vertex counting algorithm was found to be consistent within errors with the integrated luminosity calculated using both track counting algorithms. Systematic uncertainties affecting this method were also calculated and shown in this chapter. Given that the vertex counting algorithm has been proven to yield good results when analysing both vdM scans and physics runs during 2012, the plan is to use this algorithm again for 2015 data.

There are some outstanding issues with the vertex counting algorithm but there already are some ideas to further improve the method before it is applied to 2015 data. One such issue is the dependency of the fake rate on the beam spot width and the fact that the beam spot widths in MC and data are different. This was briefly addressed in Section 5.8, by calculating a new value of μ_{fake} which is the original μ_{fake} multiplied by the ratio of the RMS of the z-distributions in data and MC. This is a very simple and linear approach to this correction but with further studies, this correction could be more precise.

Another issue is the failure of the masking correction evaluation for vertices with NTrkCut 3. It was shown in this chapter that the p_{mask} vs. Δz distribution behaviour for NTrkCut 3 suggests that there is a mechanism producing two very close-by vertices in z from a single collision. It is yet unknown if this effect is due to split vertices, secondary interactions, or some other mechanism but an idea to help decrease the impact of this effect is cutting on the x-y position of each vertex, requiring it to be consistent with the beam spot.

Chapter 6

Strong and Electroweak production of *Zjj* at 13 TeV

6.1 Introduction

This chapter presents the cross-section measurement for the production of two jets in association with a Z boson (Zjj) due to electroweak (EWK) and strong processes, and also due to purely electroweak processes. This was performed using the data collected by ATLAS during pp collisions at $\sqrt{s} = 13\,\text{TeV}$ in 2015, which corresponds to an integrated luminosity of $3.2\,\text{fb}^{-1}$. The uncertainty on the integrated luminosity is $\pm 2.1\%$. It is derived following a methodology similar to that detailed in [22], from a preliminary calibration of the luminosity scale using x-y beam-separation scans performed in August 2015.

The production of Zjj events at the LHC is predominantly a result of the strong interaction; the production mechanisms are quark-quark, gluon-quark or gluon-gluon fusion, with the jets arising from additional quark/gluon radiation, an example of which is shown in Figure 6.1. The events produced by these kinds of processes will be referred to as QCD Zjj events. Zjj events can also be produced as a result of quark-quark scattering mediated by a vector gauge boson. However, this is a purely electroweak process which makes it a rare occurrence at the LHC.

Electroweak Zjj production, with the Z boson decaying to two charged leptons, is defined as all the contributions to $\ell^+\ell^- jj$ production involving a t-channel exchange of an electroweak gauge boson [37, 38]; these correspond to the diagrams shown in Figures 6.2a-6.2g. One of these contributions is the vector boson fusion (VBF) process, shown in Figure 6.2a, which is of particular interest due to its similarity with the VBF production of a Higgs boson, as well as its sensitivity to new physics via the anomalous triple gauge couplings (aTGCs) in the WWZ vertex [39]. Another contribution comes from the ZV (V=W,Z) diboson production with one boson decaying hadronically, since it contains purely electroweak couplings. Examples of this kind of process for the t-channel and the s-channel can be seen in Figure 6.3.

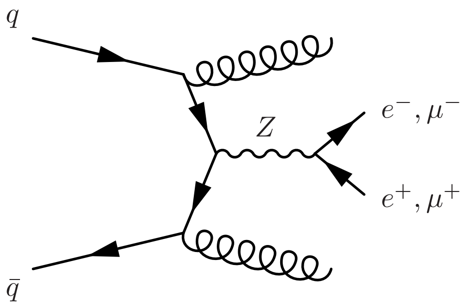
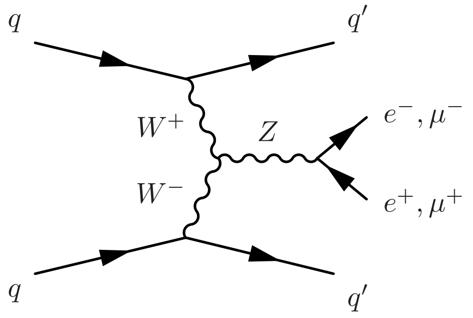


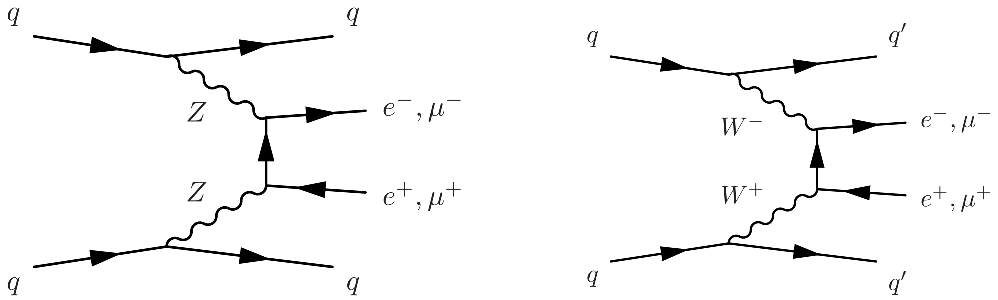
Figure 6.1: Example Feynman diagram for QCD Zjj production at the LHC.

The first measurement of the electroweak Zjj cross-section was performed by ATLAS with data collected during pp collisions at $\sqrt{s} = 8\,\text{TeV}$, which corresponded to an integrated luminosity of $20.3\,\text{fb}^{-1}$ [40]. The 8 TeV analysis cross-section measurements were in good agreement with the theory prediction from the Powheg generator [41, 42, 43]. This study also placed limits on aTGCs through the VBF diagram containing the WWZ vertex. The aim of the analysis presented here is to repeat these measurements but with $\sqrt{s} = 13\,\text{TeV}$ data. This would allow to perform a study on the dependence of these results on $\sqrt{s}$. Additionally, a higher $\sqrt{s}$ translates to more data at high dijet invariant mass, which is the region where the electroweak signal is more easily extracted. Finally, the results from this analysis would be important to a range of new phenomena searches which have the electroweak Zjj production as a background [44, 45].

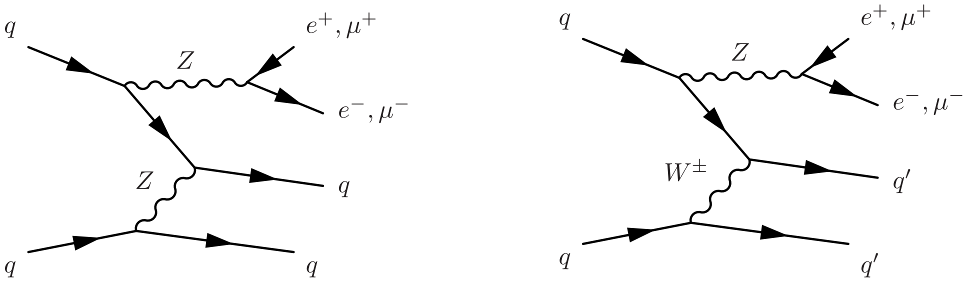
As previously mentioned, the 8 TeV analysis was performed with $20.3\,\text{fb}^{-1}$ of data, which is approximately 6 times more data than that available for the 13 TeV analysis. However, the parton luminosity enhancement produces an increase in the production rate of the events of interest for this analysis, which have a large dijet invariant mass, of between 5 and 10, as is shown in Figure 6.4. This makes the statistical precision from both analyses comparable in the high dijet invariant mass region. Comparing the luminosities of parton-parton collisions as a function of the centre-of-mass energy of the colliding partons is a good way of learning about the general issues of energy, luminosity and relative merits of proton-proton collisions. In essence, a high-energy proton is a broadband unseparated beam of quarks, antiquarks, and gluons [46].



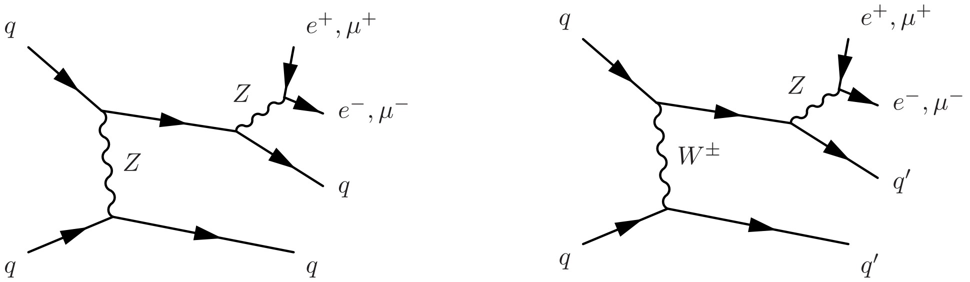
(a) Vector boson fusion.



(b) Non-resonant (via Z). (c) Non-resonant (via W).

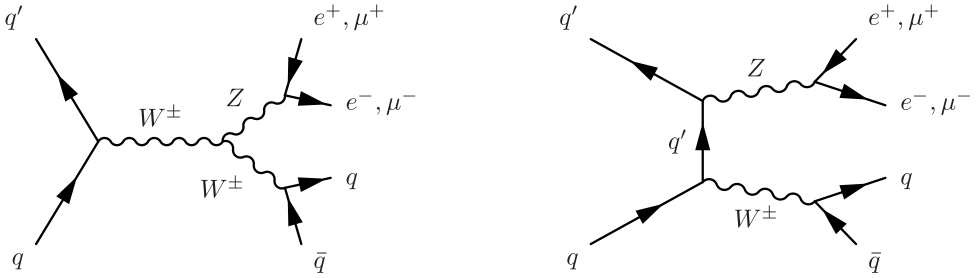


(d) Z and t-channel (via Z). (e) Z and t-channel (via W).



(f) t-channel (via Z) and Z. (g) t-channel (via W) and Z.

Figure 6.2: Feynman diagrams for EWK *Zjj* production at the LHC.



(a) s-channel diboson. (b) t-channel diboson.

Figure 6.3: Feynman diagrams for EWK *Zjj* diboson production at the LHC.

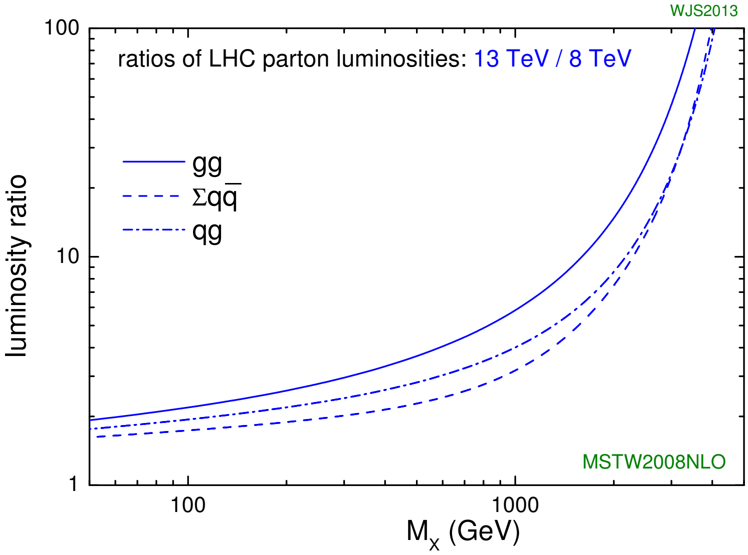


Figure 6.4: Ratio of parton luminosities at $\sqrt{s} = 13\,\text{TeV}$ compared to $\sqrt{s} = 8\,\text{TeV}$ at the LHC, using the MSTW 2008 NLO PDFs [47]. Taken from Ref. [48].

The structure of this chapter is as follows. First, the backgrounds affecting this analysis and the data and MC samples used to extract the desired signal are detailed in Section 6.2. Section 6.3 describes the event and object selections used for this analysis. Section 6.4 presents comparisons to an earlier analysis of Z+jets production at 13 TeV by ATLAS [49]. For brevity, this will be referred to below as the "13 TeV Z+jets analysis". The purpose of this comparison was to confirm that the analysis presented in this chapter provided a good description of general Z+(n)jets production and relevant backgrounds in a well-studied fiducial region, before moving on to studying QCD and electroweak production in the extreme kinematic region of high dijet invariant mass for the extraction of the electroweak signal. After confirming a good data and MC agreement in this region, the different fiducial regions and variables designed to study the different components of the *Zjj* events are introduced in Section 6.5. These are defined in the same way as in the 8 TeV analysis to allow for direct comparisons of the results. Section 6.6 presents comparisons of data and MC for the fiducial regions and a few of the variables defined for this analysis. Then, Section 6.7 presents the inclusive *Zjj* production cross-section measurement and finally, Section 6.8 presents the electroweak *Zjj* cross-section measurement, including the data-driven technique used to correct for the mismodelling of QCD *Zjj* production in the MC samples.

6.2 Data and MC samples

This analysis uses data collected by ATLAS between August and November of 2015. During this period, proton beams with an energy of 6.5 TeV, were circulated in the LHC, with a bunch spacing of 25 ns. The peak delivered instantaneous luminosity was $\mathrm{L} = 5 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, and the average number of additional *pp* interactions per bunch crossing was $\langle\mu\rangle = 13.5$ [50]. Quality cuts are applied to this data to remove events with activity in problematic regions of the detector or collected during periods in which a particular sub-detector was not functioning properly[1]. Data which passed these quality criteria correspond to a total integrated luminosity of $3.2\,\mathrm{fb}^{-1}$.

Monte Carlo simulations are used to compare the data to the SM predictions and to estimate the contribution from background events. Electroweak *Zjj* production has been simulated using the Sherpa v.2.1.1 generator [51]. Sherpa is a matrix-element plus parton-shower generator that provides Z+n parton predictions (n = 0,1,2,...) at leading-order (LO) accuracy in perturbative QCD. The CKKW method is used to combine the various final-state topologies and match to the parton shower [52]. Matrix elements were calculated for up to two partons at next-to-leading-order (NLO) and up to four additional partons at LO using the Comix [53] and OpenLoops [54] matrix element generators, and were merged with the Sherpa parton shower [55] using the ME+PS@NLO prescription [56].

Unlike the electroweak *Zjj* sample used by the 8 TeV analysis [40], the electroweak *Zjj* sample used in this analysis includes *ZV* diboson events. However, a request has been made to produce a new electroweak *Zjj* sample without these events.

QCD *Zjj* production has been simulated using three different generators: Sherpa v.2.1.1, MadGraph5_aMC@NLO v2.2.2 [57] and Alpgen. The MadGraph5 _aMC@NLO v2.2.2 generator uses explicit matrix elements for up to four partons at LO. The parton shower for this sample was obtained from Pythia v8.186 [58],

---

[1] This was done by requiring events to belong to runs in the good runs list: `data15_13TeV.periodAllYear_DetStatus-v63-pro18-01_DQDefects-00-01-02_PHYS_StandardGRL_All_Good.xml`.

with the A12 tune and NNPDF23LO PDF set [59]. Having different samples for the same process allowed cross-checks to be performed of the behaviour of the QCD contribution to Zjj production and helped assess the theoretical modelling uncertainty.

The backgrounds affecting this analysis come from WW semileptonic decays, $t\bar{t}$ and single-top production, and W+jets events. WW diboson production was simulated using SHERPA v2.1.1. The samples for $t\bar{t}$ and single-top production were generated with the POWHEG-BOX v2 generator and PYTHIA v6.428 [60], with the Perugia 2012 tune [61]. Finally, W+jets events were simulated with SHERPA v2.1.1. All the SHERPA samples were produced using the CT10 [62] parton distribution functions (PDFs) and the default generator tune for underlying event activity.

All the MC samples above are passed through the GEANT 4 simulator [63] for a full simulation of the ATLAS detector [64]. The effect of pile-up interactions in the same or nearby bunch crossings is also simulated, using PYTHIA v8.186 with the A2 tune [34] and the MSTW2008LO PDFs [47].

The cross-sections for the different processes used in this analysis are normalised to the results of the highest order calculations available. The MC samples used are reweighted so that the distribution of $\langle \mu \rangle$ matches that observed in data.

6.3 Event and object selection

This analysis was performed in both the electron and muon channels. Events containing a Z candidate in the muon channel are required to have passed either the isolated single muon trigger `HLT_mu20_iloose_L1MU15` or the non-isolated higher-p_{T} trigger `HLT_mu50`. Events containing a Z candidate in the electron channel are required to have passed at least one of the single electron triggers `HLT_e120_lhloose`, `HLT_e60_lhmedium` or `HLT_e24_lhmedium_L1EM20VH`. These triggers follow the nomenclature described in Section 3.2.4 and, for the case of the L1 electromagnetic trigger, H indicates that a hadronic core isolation cut was added, and V indicates that the thresholds were varied with η to account for energy loss [65].

All events are required to have a reconstructed primary vertex with at least two tracks with $p_{\mathrm{T}} > 400\,\mathrm{MeV}$ associated to it.

6.3.1 Muons

ATLAS defines different sets of requirements to identify muons, the aim of which is to reject fake muons coming mainly from pion and kaon decays and provide a good momentum measurement [66]. This analysis uses muons passing the “Medium” identification criteria, which is the default muon selection for ATLAS. These criteria minimise the systematic uncertainties on the muon reconstruction and calibration, and use tracks from stand-alone (SA) muons and from combined (CB) muons. A SA muon is a muon whose trajectory is reconstructed only by the MS and a CB muon is a muon whose trajectory is a successful combination of a track in the ID with a track in the MS.

Candidate muons must have $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.4$. They must also satisfy a set of isolation requirements based on the activity in the tracking sub-detectors and in the calorimeters [66]. This set of isolation cuts has an efficiency of 90% for muons and electrons with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$, and 99% for muons and electrons with $p_{\mathrm{T}} > 60\,\mathrm{GeV}$.

Muons are also required to satisfy the following lepton-to-vertex association cuts:

- $|\Delta \mathrm{z}_0 \times \sin(\theta)| < 0.5$ mm,
- $|\mathrm{d}_0$ significance$| < 3$,

where d_0 and z_0 are the transverse and longitudinal impact parameters with respect to the beamspot.

The inner detector tracks associated with each muon must pass the following quality requirements, defined by the ATLAS Muon Combined Performance group [67]:

- Number of pixel hits + number of crossed dead pixel sensors > 0.
- Number of SCT hits + number of crossed dead SCT sensors > 4.
- Number of pixel holes + number of SCT holes < 3.
- If n_{TRT}^{hits} denote the number of TRT hits in the muon track, $\mathrm{n}_{TRT}^{outliers}$ the number of TRT outliers[2] on the muon track, and $\mathrm{n} = \mathrm{n}_{TRT}^{hits} + \mathrm{n}_{TRT}^{outliers}$, then the requirement is:

$$\text{If } 0.1 < |\eta| < 1.9,\ \mathrm{n} > 5 \text{ and } \mathrm{n}_{TRT}^{outliers} < 0.9\,\mathrm{n}.$$

6.3.2 Electrons

Electrons in the central region of the ATLAS detector are triggered and reconstructed from energy deposits in the EM calorimeter that are matched to a track

---

[2] The TRT outliers are usually a set of hits in the TRT that fail to form a smooth trajectory with the pixel and SCT hits.

in the ID. Electrons are identified using different sets of requirements with different background rejection levels and signal efficiencies. The variables used in these identification criteria are the shape of the EM showers in the calorimeter and the tracking and track-to-cluster matching quantities. The identification requirements can be based on either single cuts on these variables or a single cut on the resulting likelihood function which takes these variables as input. This analysis uses electrons passing the "Medium Likelihood" identification criteria [68], which provides a good signal efficiency while rejecting light and heavy-flavour jets and background electrons coming from photon conversions.

Candidate electrons must have $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.47$, excluding the *crack* region[3] between the barrel and end-cap calorimeters at $1.37 < |\eta| < 1.52$. These cuts on the electron η are perfomed using the measurement of the cluster position in the second calorimeter layer. This is the same variable as the one used for the η binning of the electron selection [69]. Additionally, electrons must satisfy the same set of isolation requirements as previously mentioned in the muon selection.

Electrons are also required to pass the following lepton-to-vertex association cuts:

- $|\Delta \mathrm{z}_0 \times \sin(\theta)| < 0.5\,\mathrm{mm}$.
- $|\mathrm{d}_0$ significance$| < 5$.

6.3.3 Jets

Candidate hadronic jets are reconstructed from calorimeter clusters [70] using the anti-k$_T$ algorithm [71] with radius parameter $R = 0.4$. To remove jets that do not come from the primary interaction, known as pile-up jets, a multivariate algorithm referred to as Jet Vertex Tagger (JVT) is used [72], designed to remove jets in which a large fraction of the tracks are not associated to a primary vertex.

Jets are required to have $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|y| < 4.5$, and must be well separated from any of the selected leptons. This last requirement translates to removing any jets in the event if $\Delta\mathrm{R} \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ between the jet and a lepton is less than 0.2.

---

[3] The crack region refers to a poorly instrumented region of the calorimeter where the measurements have a large energy uncertainty.

6.4 Background modelling test

Before moving on to studying QCD and electroweak *Zjj* production in a high dijet invariant mass and more selective kinematic region, a test was performed to confirm that this analysis provided a good description of general Z+(n)jets production, with the relevant backgrounds, in a well-studied fiducial region. The test consisted in comparing this analysis to the results of an earlier ATLAS Z+jets study performed using $3.2\,\text{fb}^{-1}$ of data collected in 2015 [49].

The 13 TeV Z+jets analysis is concerned with events in which a Z boson is produced in association with any number of jets. In Ref. [49], this analysis presents different variables distributions in a fiducial region in which a Z boson is produced in association with at least one jet. These same distributions were produced with the analysis of this note.

In order to compare this analysis to the 13 TeV Z+jets analysis, a fiducial region was defined requiring events to contain exactly two opposite sign leptons and at least one jet. This fiducial region will be labelled as Z+1 jet. The candidate leptons and jets are selected as described in Section 6.3 and are also required to pass the following quality cuts:

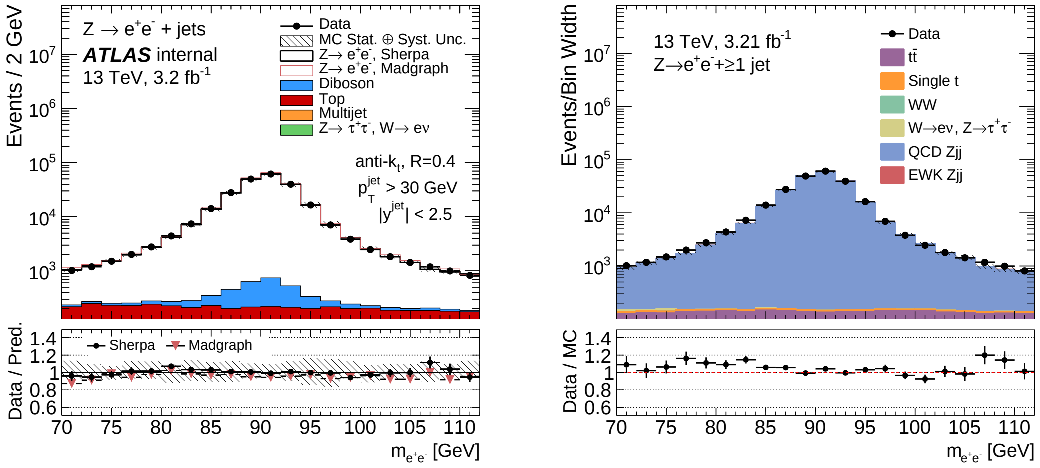
- Transverse momentum of the leptons: $p_T^\ell > 25\,\text{GeV}$,
- Dilepton invariant mass: $66\,\text{GeV} < m_{\ell\ell} < 116\,\text{GeV}$,
- Transverse momentum of the jets: $p_T^{jet} > 30\,\text{GeV}$,
- Rapidity of the jets: $|y^{jet}| < 2.5$.

Table 6.1 shows the contribution of the different MC samples to the total number of Z+1 jet events in percent for the electron and muon channels. These values are consistent with the equivalent table presented in the 13 TeV Z+jets analysis [49]. QCD multijet production has been found to be negligible and so is not considered any further in this analysis. The electroweak *Zjj* sample composition includes ZV diboson events in the signal definition.

	Z+1 jet	
Process	e–channel	μ–channel
QCD Zjj	97.74 %	97.93 %
EWK Zjj	0.72 %	0.68 %
$t\bar{t}$	1.38 %	1.25 %
single–t	0.11 %	0.10 %
$W \to e\nu$, $Z \to \tau\tau$	0.05 %	0.03 %

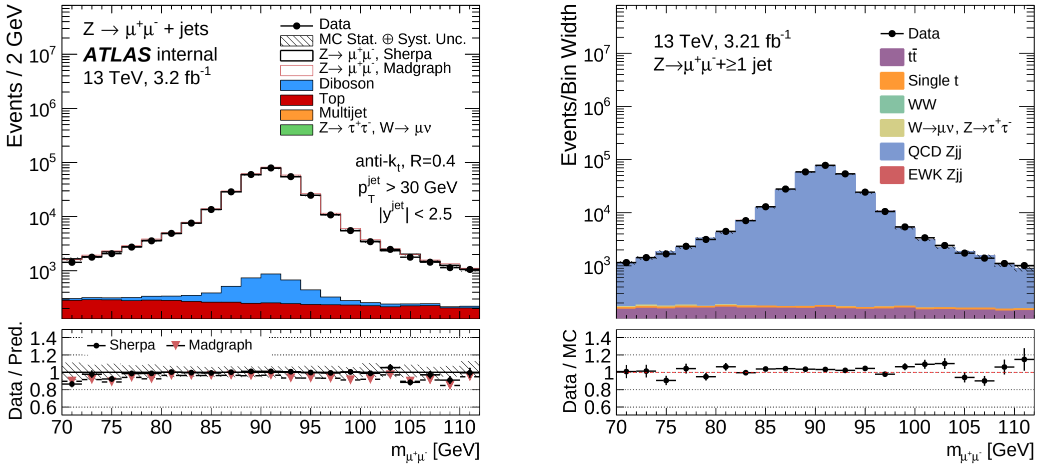
Table 6.1: Predicted events composition of the fiducial region requiring a Z candidate and at least one jet, for the electron and muon channel. The electroweak Zjj sample includes ZV events. The QCD Zjj sample comes from the SHERPA generator. WW are not presented in the table since their contributions is negligible.

Figures 6.5–6.8 show the distributions presented by the 13 TeV Z+jets analysis on the left, and the equivalent distributions obtained with the analysis from this note on the right. From these figures it is clear that the data and MC comparisons behave similarly in both analyses, from which it can be concluded that the analysis in this note provides a good modelling of the general Z+(n)jets production in this well-studied region of fiducial region.



(a) 13 TeV Z+jets analysis.

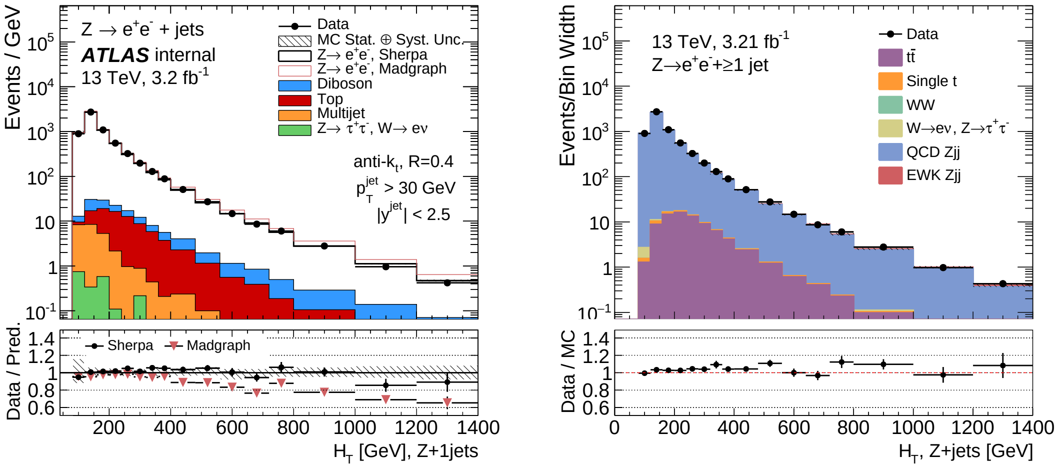
(b) This analysis.



(c) 13 TeV Z+jets analysis.

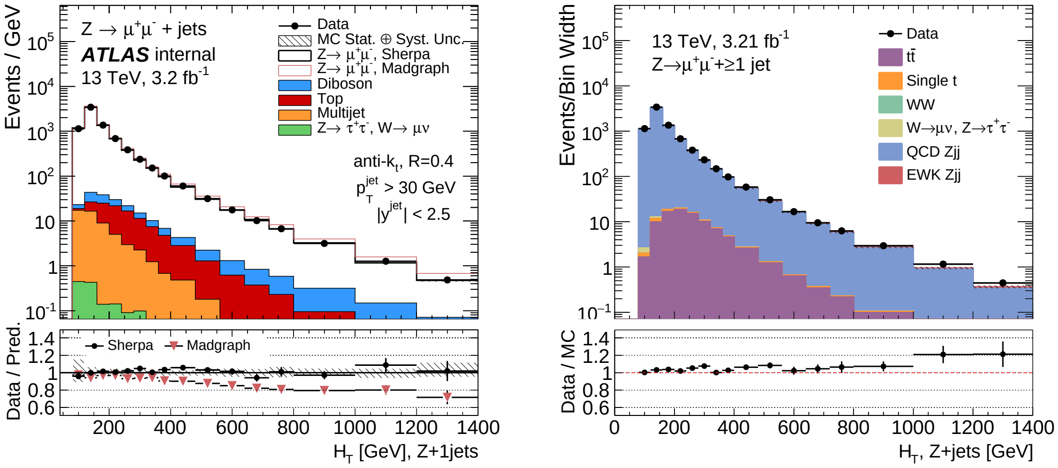
(d) This analysis.

Figure 6.5: $m_{\ell\ell}$ distribution in the electron and muon channel measured by the 13 TeV Z+jets analysis [49] (left) and the analysis from this note (right). The electroweak Zjj sample in this analysis includes ZV events and the QCD Zjj sample is coming from the Sherpa generator. The upper plots are for electron pair and the lower plots are for muon pair events.



(a) 13 TeV Z+jets analysis.

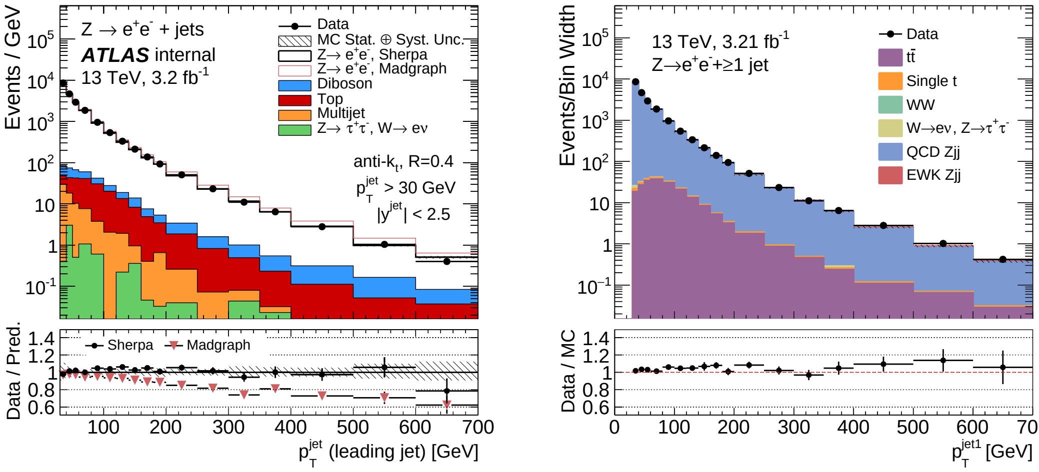
(b) This analysis.



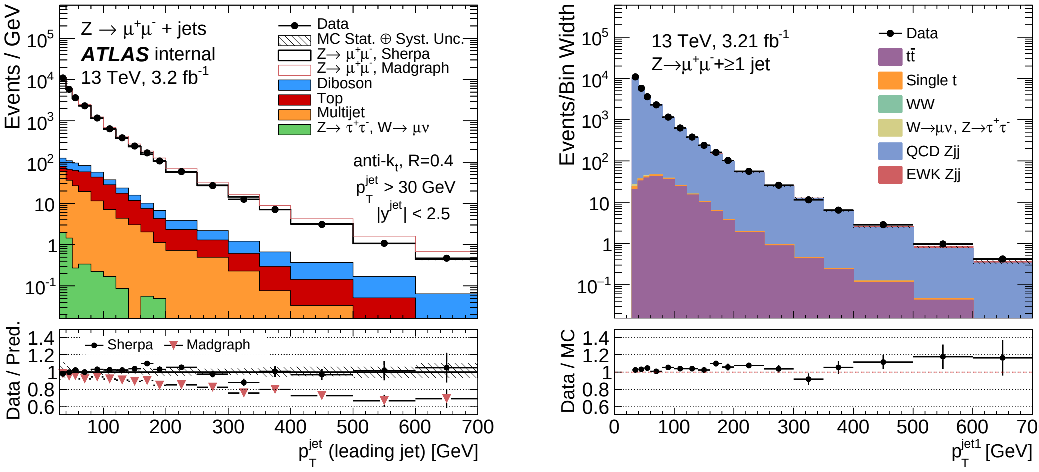
(c) 13 TeV Z+jets analysis.

(d) This analysis.

Figure 6.6: $\mathrm{H_T}$ distribution of the Z +jets system in the electron and muon channel measured by the 13 TeV Z+jets analysis [49] (left) and the analysis from this note (right). The electroweak *Zjj* sample in this analysis includes ZV events and the QCD *Zjj* sample is coming from the SHERPA generator. The upper plots are for electron pair and the lower plots are for muon pair events.

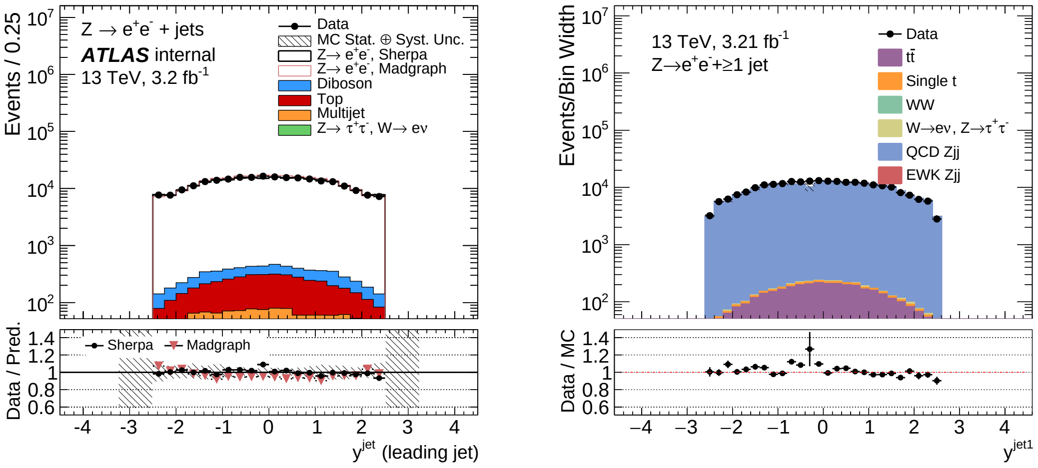


(a) 13 TeV Z+jets analysis. (b) This analysis.



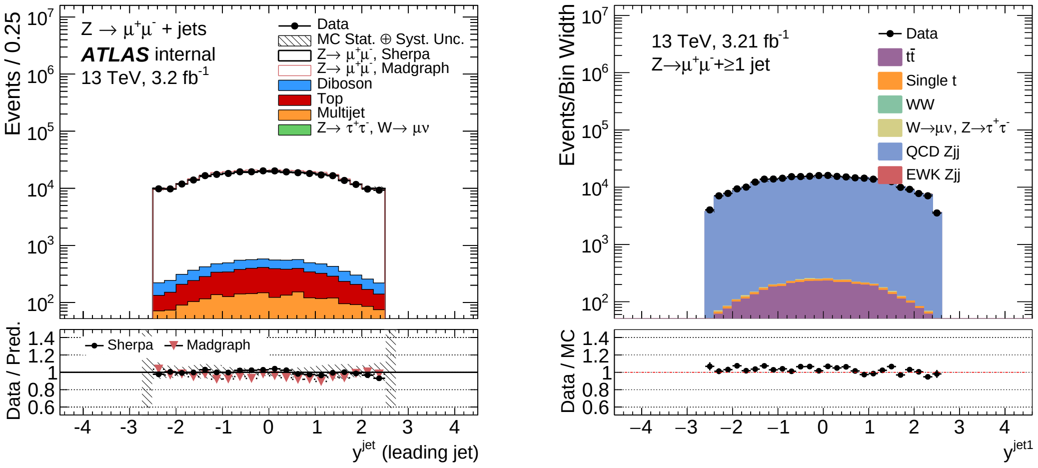
(c) 13 TeV Z+jets analysis. (d) This analysis.

Figure 6.7: p_{T}^{jet1} distribution in the electron and muon channel measured by the 13 TeV Z+jets analysis [49] (left) and the analysis from this note (right). The electroweak Zjj sample in this analysis includes ZV events and the QCD Zjj sample is coming from the SHERPA generator. The upper plots are for electron pair and the lower plots are for muon pair events.



(a) 13 TeV Z+jets analysis.

(b) This analysis.



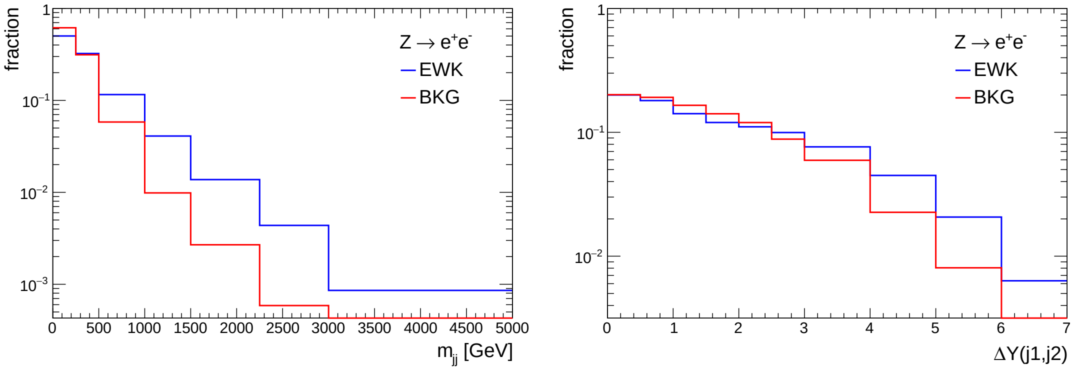
(c) 13 TeV Z+jets analysis.

(d) This analysis.

Figure 6.8: Leading jet rapidity distribution in the electron and muon channel measured by the 13 TeV Z+jets analysis [49] (left) and the analysis from this note (right). The electroweak Zjj sample in this analysis includes ZV events and the QCD Zjj sample is coming from the SHERPA generator. The upper plots are for electron pair and the lower plots are for muon pair events.

6.5 Fiducial regions and variables definitions

In order to study electroweak Zjj events, it is essential to distinguish these events from the more common QCD Zjj events. Fortunately, electroweak Zjj events have very characteristic configurations which can be exploited to differentiate the electroweak signal events from other production mechanisms. For example, in the electroweak production of Zjj via VBF, shown in Figure 6.2a, the two outgoing quarks which recoil against the emitted W boson go on to form *tagging* jets, which are typically produced with a large separation in rapidity (Δy) and relatively large transverse momenta (relative to the QCD jets). Clearly, this pair of jets will also have a large dijet invariant mass (m_{jj}). Additionally, the VBF diagram has no colour flow between the incoming partons because the interaction is produced via the exchange of colourless electroweak bosons. Therefore, the two jets in the final state are not colour connected which means very little additional quark and gluon radiation is expected in the rapidity interval between the two tagging jets. This translates experimentally to detecting very few jets in this rapidity gap (N^{gap}_{jets}). The discriminative power of these variables can be observed in Figures 6.9 and 6.10. These and other properties can be used to define fiducial regions with different sensitivity to the electroweak component of the *Zjj* cross-section.



(a) m_{jj} distribution, electron channel.

(b) $|\Delta y|$ distribution, electron channel.

Figure 6.9: Shape comparisons of m_{jj} and $|\Delta y|$ between EWK *Zjj* and the Monte Carlo backgrounds, which includes QCD *Zjj*, $t\bar{t}$, single–t, WW and W+jets, in the *baseline* region. Both curves are normalised to unit area.

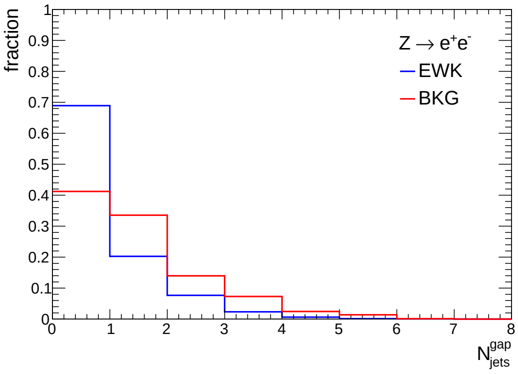


Figure 6.10: Shape comparisons of N^{gap}_{jets} between EWK *Zjj* and the Monte Carlo backgrounds, which includes QCD *Zjj*, $t\bar{t}$, single–t, WW and W+jets, in the *high-mass* region. Both curves are normalised to unit area.

This analysis uses the same fiducial regions defined by the 8 TeV analysis, each with a different sensitivity to the electroweak component of the *Zjj* production. Using the same fiducial regions allows a direct comparison of the 8 TeV and the 13 TeV analysis results. A detailed description is presented below and an overview of the cuts in each of these fiducial regions can be seen in Table 6.2.

First, the most inclusive possible fiducial region is defined, the *baseline* region. Events in the *baseline* region are required to contain two opposite-charge same-flavour leptons (e^+e^- or $\mu^+\mu^-$) with an invariant mass in the window $81 \leq m_{\ell\ell} \leq 101$ GeV; this is called a *Z* candidate. These events must also contain at least two jets, with $|\mathrm{y}| < 4.4$, and $p_{\mathrm{T}}^{jet1} > 55$ GeV and $p_{\mathrm{T}}^{jet2} > 45$ GeV. The jets in the event are ordered by their p_{T}, starting with the highest p_{T} jet (leading jet). Then, two subsets of this fiducial region are defined where the energy scale of the dijet system is increased, either by increasing the minimum p_{T} cut on the leading and sub-leading jets (*high-p_T* region) or by applying a minimum cut on the dijet invariant mass (*high-mass* region). Both of these regions, especially the *high-mass* region, are designed to examine the impact of the electroweak *Zjj* processes, since these produce a harder jet transverse momentum and greater dijet invariant mass than the QCD *Zjj* processes, as can be seen in Figures 6.11 and 6.9.

The next fiducial region is the *search* region, where the number of electroweak *Zjj* signal events is counted. The *search* region was designed to maximise the electroweak component of the *Zjj* processes while at the same time suppress the contributions from QCD *Zjj* and other kinds of processes. Events in this region are required to pass the cuts of the *baseline* region in addition to the following cuts:

$m_{jj} > 250$ GeV

This cut removes most of the contribution from diboson events where a *Z* boson decays leptonically and the other boson decays to two jets.

$\mathrm{N}_{jets}^{gap} = 0$

This is a veto on the additional jets in the rapidity gap between the two leading jets. As mentioned before, in the electroweak process there is less QCD radiation between the two leading jets than in the background process, as can be seen in Figure 6.10. Therefore, applying this veto helps suppress the QCD component of the *Zjj* events with respect to the electroweak component.

$p_T^{\ell\ell} > 20$ GeV and $p_T^{\mathrm{balance}} < 0.15$

p_T^{balance} is the normalised transverse momentum balance between the two leptons and the two highest transverse momentum jets, or

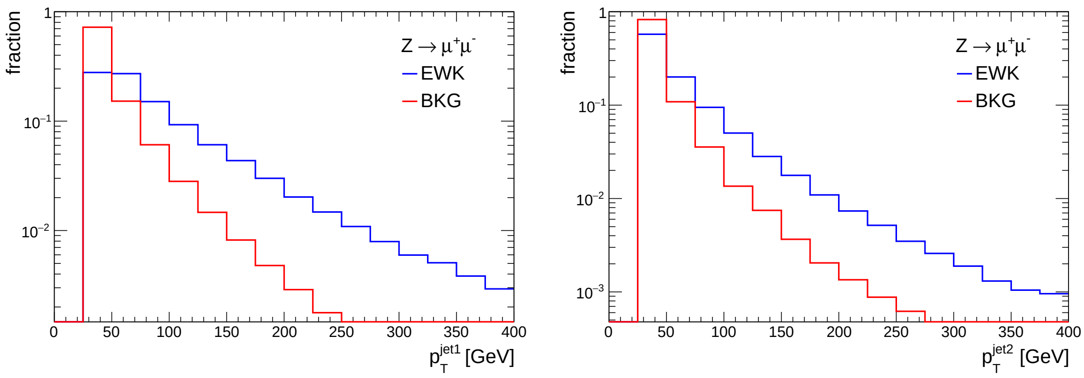
$$p_{\mathrm{T}}^{\mathrm{balance}} = \frac{|\vec{p}_{\mathrm{T}}^{\,\ell_1} + \vec{p}_{\mathrm{T}}^{\,\ell_2} + \vec{p}_{\mathrm{T}}^{\,j_1} + \vec{p}_{\mathrm{T}}^{\,j_2}|}{|\vec{p}_{\mathrm{T}}^{\,\ell_1}| + |\vec{p}_{\mathrm{T}}^{\,\ell_2}| + |\vec{p}_{\mathrm{T}}^{\,j_1}| + |\vec{p}_{\mathrm{T}}^{\,j_2}|}, \tag{6.1}$$

where $\vec{p}_{\mathrm{T}}^{\,i}$ is the transverse momentum vector of object i, and ℓ_1 and ℓ_2 represent the two leptons that define the Z boson candidate. These cuts help remove events where the jets arise from pile-up or multiple parton interactions (MPI). The p_T^{balance} cut also helps remove events in which the p_{T} of one or more jets was badly measured and it enhances the VBF *Zjj* contributions, where the lack of additional radiation causes the Z boson and the dijet system to be very well balanced. The discriminative power of this variable can be appreciated in Figure 6.12.

Finally, the *control* region was designed to suppress the electroweak contribution and maximise the QCD contribution to the *Zjj* production in order to study the modelling of the QCD *Zjj* production. Events in the *control* region must pass similar cuts to the ones in the *search* region with two modifications: (*i*) they must have *at least one* jet with $p_{\mathrm{T}} > 25$ GeV in the rapidity interval between the two leading jets (referred to as a “gap” jet), (*ii*) the transverse momentum balance variable is redefined to use the two leptons, the two highest p_{T} jets and the highest p_{T} “gap” jet. This new variable, $p_{\mathrm{T}}^{\mathrm{balance,3}}$, is defined analogously to the $p_{\mathrm{T}}^{\mathrm{balance}}$ variable in Equation 6.1 but including the additional jet in the numerator and denominator.

object	*baseline*	*high-mass*	*search*	*control*	*high-p_T*
leptons	$\|\eta^\ell\| < 2.47$, $p_{\mathrm{T}}^\ell > 25$ GeV				
dilepton pair	$81 \leq m_{\ell\ell} \leq 101$ GeV				
	—		$p_{\mathrm{T}}^{\ell\ell} > 20$ GeV		—
jets	$\|y^j\| < 4.4$, $\Delta R_{j,\ell} \geq 0.2$				
	$p_{\mathrm{T}}^{jet1} > 55$ GeV				$p_{\mathrm{T}}^{jet1} > 85$ GeV
	$p_{\mathrm{T}}^{jet2} > 45$ GeV				$p_{\mathrm{T}}^{jet2} > 75$ GeV
dijet system	—	$m_{jj} > 1$ TeV	$m_{jj} > 250$ GeV		—
gap jets	—		$N_{jet} = 0$	$N_{jet} \geq 1$	—
Zjj system	—		$p_{\mathrm{T}}^{\mathrm{balance}} < 0.15$	$p_{\mathrm{T}}^{\mathrm{balance,3}} < 0.15$	—

Table 6.2: Summary of the selection cuts for every fiducial region.



(a) Leading jet p_T distribution.

(b) Subleading jet p_T distribution.

Figure 6.11: Shape comparisons of the p_{T} of the leading and subleading jets between EWK *Zjj* and the Monte Carlo backgrounds, which includes QCD *Zjj*, $t\bar{t}$, single–t, *WW* and *W*+jets, in a region requiring only a *Z* candidate. Both curves are normalised to unit area.

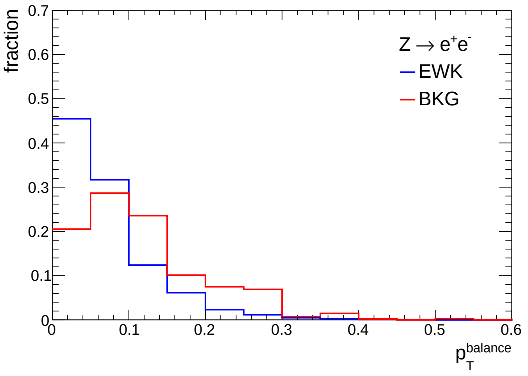


Figure 6.12: Shape comparisons of $p_{\mathrm{T}}^{\mathrm{balance}}$ between EWK *Zjj* and the Monte Carlo backgrounds, which includes QCD *Zjj*, $t\bar{t}$, single–t, *WW* and *W*+jets, in the *high-mass* region. Both curves are normalised to unit area.

6.6 Comparisons of Data and MC

This section presents the distributions of a few observables in data and MC at reconstruction level, for all of the fiducial regions relevant to this analysis defined in Section 6.5. The Monte Carlo signal simulations are compared to the data, with the various backgrounds added.

Tables 6.3 and 6.4 show the expected *Zjj* process composition in percent predicted by the MC simulation for all of the fiducial regions defined in this analysis, for the electron and muon channel, and using the QCD *Zjj* sample from the Alpgen, the MadGraph or the Sherpa generator. It is clear that the event sample is dominated by real *Zjj* events, mainly due to the tight cut on the Z mass peak. These are the ones above the horizontal line. This table also confirms that the main contribution to *Zjj* production is through strong interactions. The WW dibosons and W+jets backgrounds are not shown since their contribution is practically zero. The main non-*Zjj* background is $t\bar{t}$ but its contribution is mostly below 2.5%.

The fact that the electroweak sample used in this note includes ZV events makes the contribution of electroweak *Zjj* to the *search* and *control* regions quite similar in percentage terms. However, in the *control* region the electroweak *Zjj* contribution is mainly through ZV-like events, which contribute predominantly to the low m_{jj} region. In contrast, as will be shown below, the electroweak *Zjj* contribution to the *search* region is mainly in the high m_{jj} region. Therefore, the electroweak *Zjj* events in this region are mainly VBF-like.

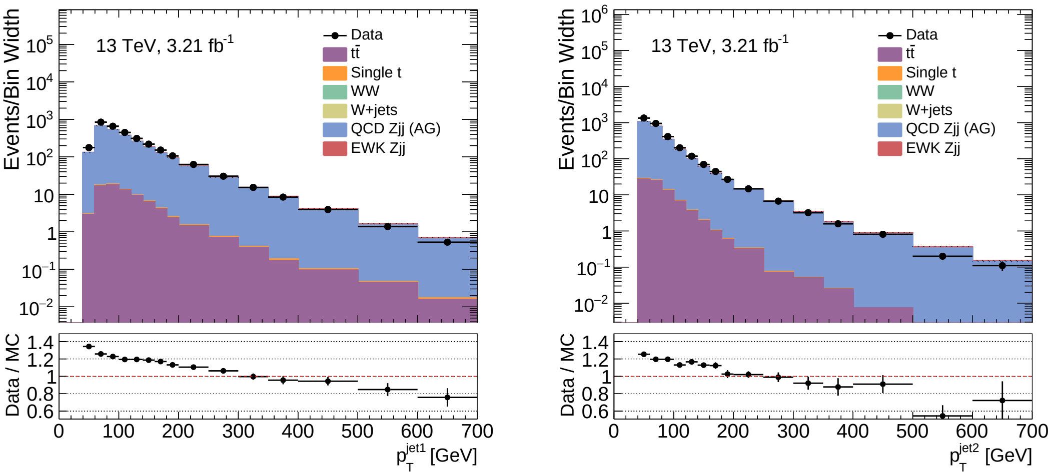
		Composition (%)				
Process		*baseline*	*high-p_T*	*high-mass*	*search*	*control*
QCD *Zjj*	e	93.85	92.25	82.46	93.97	94.12
	μ	94.20	92.67	83.40	94.22	94.56
EWK *Zjj*	e	2.96	4.05	15.63	5.20	4.18
	μ	2.79	3.86	14.56	4.93	3.86
$t\bar{t}$	e	3.04	3.55	1.81	0.79	1.64
	μ	2.88	3.33	1.94	0.80	1.52
single-t	e	0.14	0.15	0.10	0.04	0.06
	μ	0.13	0.14	0.10	0.04	0.06

Table 6.3: *Zjj* production composition in all of the fiducial regions for the electron and muon channel, using the QCD *Zjj* sample from the ALPGEN generator. The electroweak samples contains ZV diboson events; samples without these are being processed.

		Composition (%)				
Process		*baseline*	*high-p_T*	*high-mass*	*search*	*control*
QCD *Zjj*	e	95.26 (94.71)	94.15 (92.58)	88.46 (88.68)	95.98 (94.47)	95.50 (95.45)
	μ	95.61 (95.01)	94.46 (93.16)	88.94 (88.56)	96.20 (94.61)	95.93 (95.90)
EWK *Zjj*	e	2.28 (2.55)	3.06 (3.88)	10.28 (10.09)	3.47 (4.77)	3.21 (3.24)
	μ	2.11 (2.40)	2.92 (3.60)	9.70 (10.04)	3.24 (4.61)	2.89 (2.91)
$t\bar{t}$	e	2.35 (2.62)	2.68 (3.40)	1.19 (1.17)	0.53 (0.72)	1.25 (1.27)
	μ	2.18 (2.48)	2.52 (3.11)	1.29 (1.34)	0.53 (0.75)	1.13 (1.14)
single-t	e	0.11 (0.12)	0.11 (0.14)	0.06 (0.06)	0.03 (0.04)	0.04 (0.04)
	μ	0.10 (0.11)	0.11 (0.13)	0.06 (0.07)	0.03 (0.04)	0.04 (0.04)

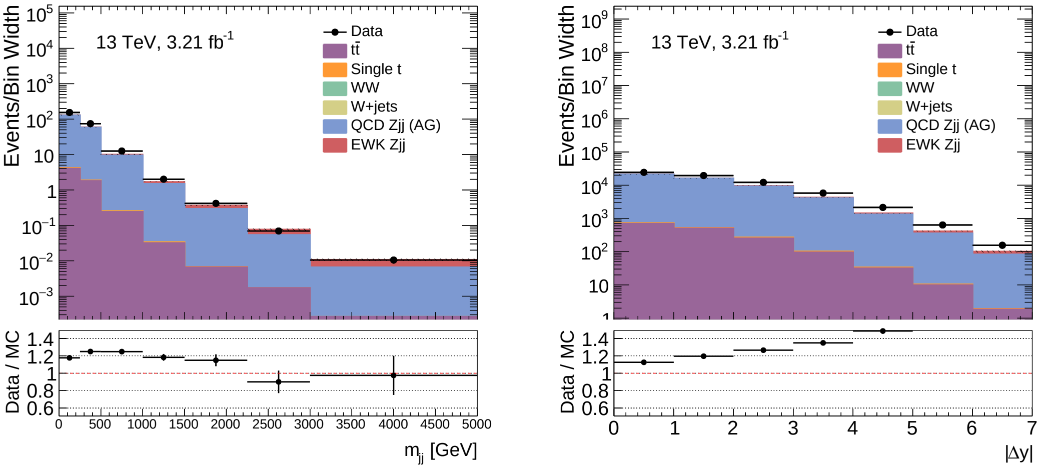
Table 6.4: *Zjj* production composition in all of the fiducial regions for the electron and muon channel, using the QCD *Zjj* sample from the MadGraph (Sherpa) generator. The electroweak samples contains ZV diboson events; samples without these are being processed.

Figures 6.13, 6.14 and 6.15 show the comparisons between data and MC of the leading and subleading jet p_{T} (p_{T}^{jet1} and p_{T}^{jet2}, respectively), m_{jj} and $|\Delta y|$ between the two leading jets distributions in the *baseline* region. These are shown for the electron and muon channels combined, and for the cases where the QCD *Zjj* sample comes from the Alpgen generator (Figure 6.13), the MadGraph generator (Figure 6.14) or the Sherpa generator (Figure 6.15). These figures show that, in general, data is approximately simulated by the MC samples, although there are clear indications of generator mismodelling at high jet transverse momentum and high dijet invariant mass.



(a) p_{T}^{jet1} (*baseline* region).

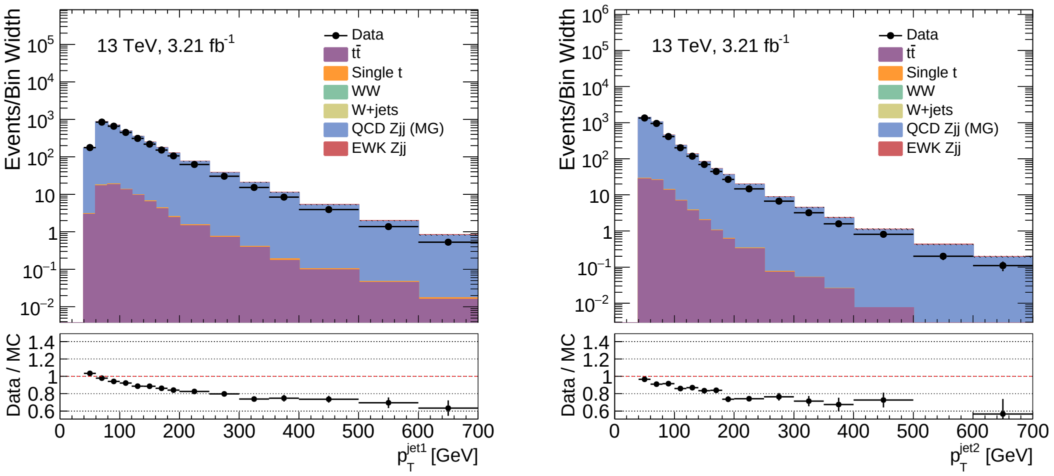
(b) p_{T}^{jet2} (*baseline* region).



(c) m_{jj} (*baseline* region).

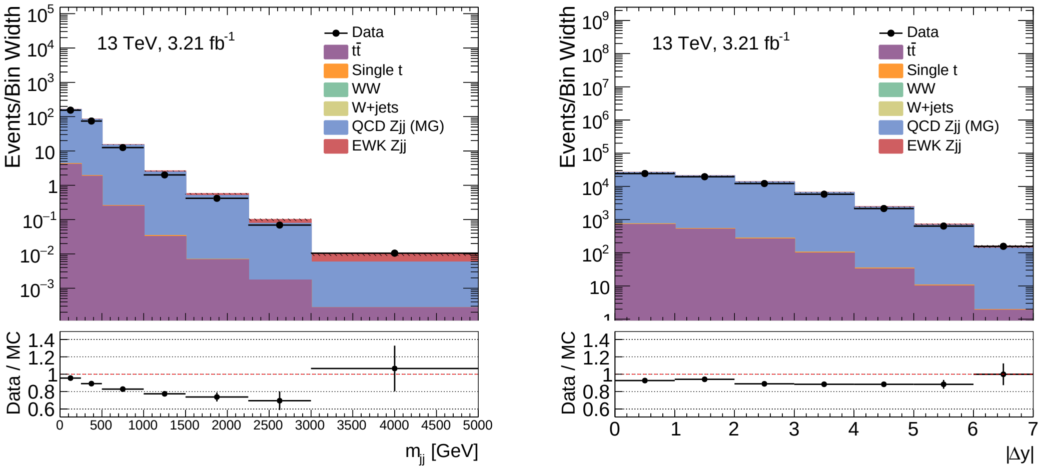
(d) $|\Delta y|$ (*baseline* region).

Figure 6.13: Reconstruction-level comparisons of the leading and sub-leading jets p_{T}, m_{jj} and $|\Delta y|$ between the leading jets distributions from data and Monte Carlo, in the *baseline* region, for the electron and muon channel combined. The QCD *Zjj* MC sample comes from the ALPGEN generator. The hashed area corresponds to the stacked MC samples uncertainty. Uncertainties shown are statistical only.



(a) p_{T}^{jet1} (*baseline* region).

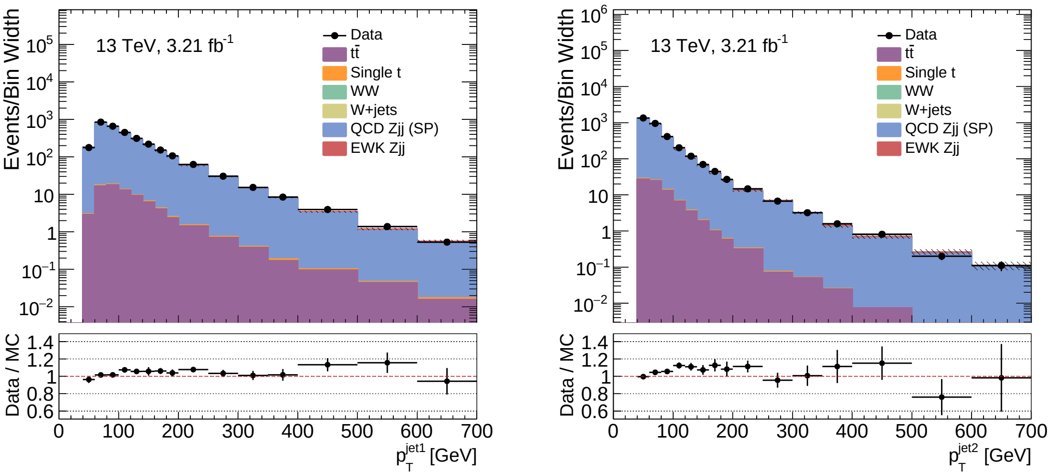
(b) p_{T}^{jet2} (*baseline* region).



(c) m_{jj} (*baseline* region).

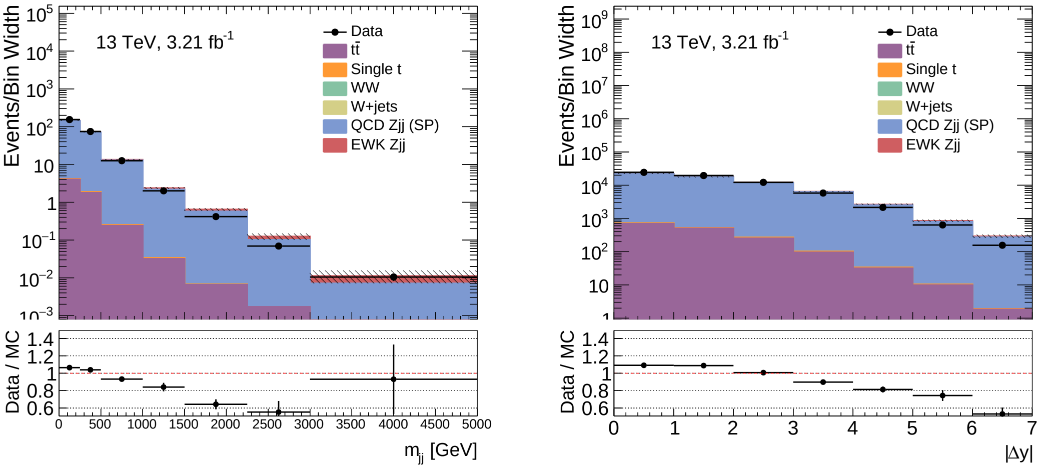
(d) $|\Delta y|$ (*baseline* region).

Figure 6.14: Reconstruction-level comparisons of the leading and sub-leading jets p_{T}, m_{jj} and $|\Delta y|$ between the leading jets distributions from data and Monte Carlo, in the *baseline* region, for the electron and muon channel combined. The QCD *Zjj* MC sample comes from the MadGraph generator. The hashed area corresponds to the stacked MC samples uncertainty. Uncertainties shown are statistical only.



(a) p_{T}^{jet1} (*baseline* region).

(b) p_{T}^{jet2} (*baseline* region).

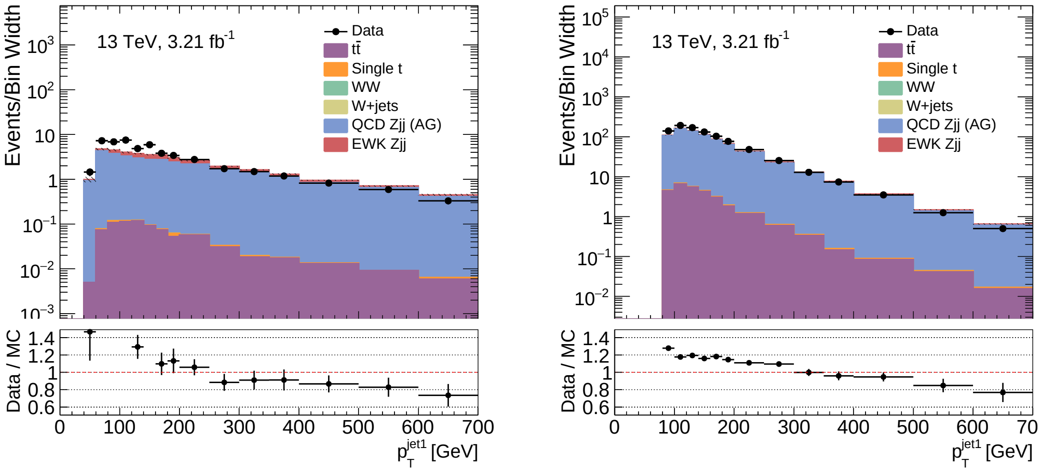


(c) m_{jj} (*baseline* region).

(d) $|\Delta y|$ (*baseline* region).

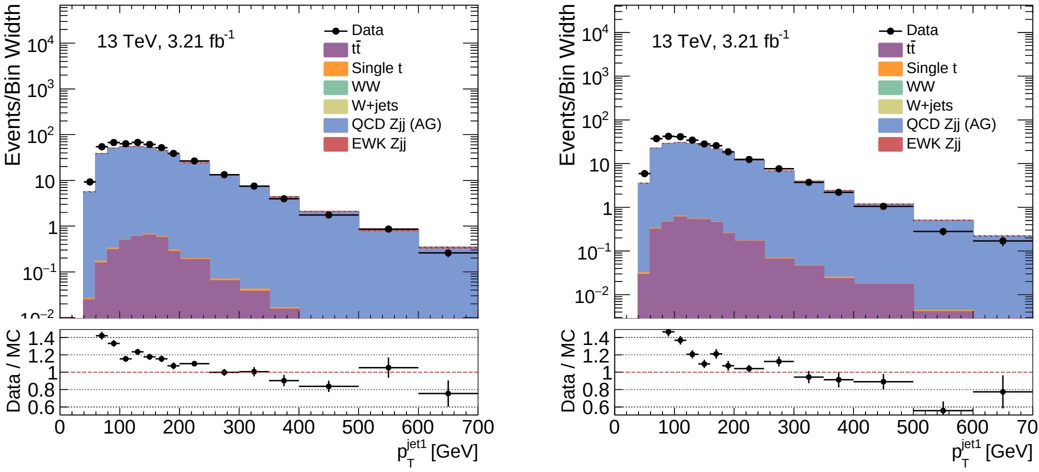
Figure 6.15: Reconstruction-level comparisons of the leading and sub-leading jets p_{T}, m_{jj} and $|\Delta y|$ between the leading jets distributions from data and Monte Carlo, in the *baseline* region, for the electron and muon channel combined. The QCD *Zjj* MC sample comes from the Sherpa generator. The hashed area corresponds to the stacked MC samples uncertainty. Uncertainties shown are statistical only.

Figures 6.16 to 6.24 show similar comparisons for the rest of the studied fiducial regions. The mismodelling in these fiducial regions is enhanced. However, this mismodelling was foreseen because other analysis and experiments have also documented a high mass, wide-angle jet mismodelling in QCD vector boson plus dijet simulations [73, 74, 75], including Ref. [40]. This mismodelling will not affect the extraction of the inclusive *Zjj* cross-section measurement, presented in Section 6.7 but does affect the extraction of the EWK *Zjj* cross-section. A correction is derived to account for this signal simulation issue and it is detailed in Section 6.7.



(a) p_{T}^{jet1} (*high-mass* region).

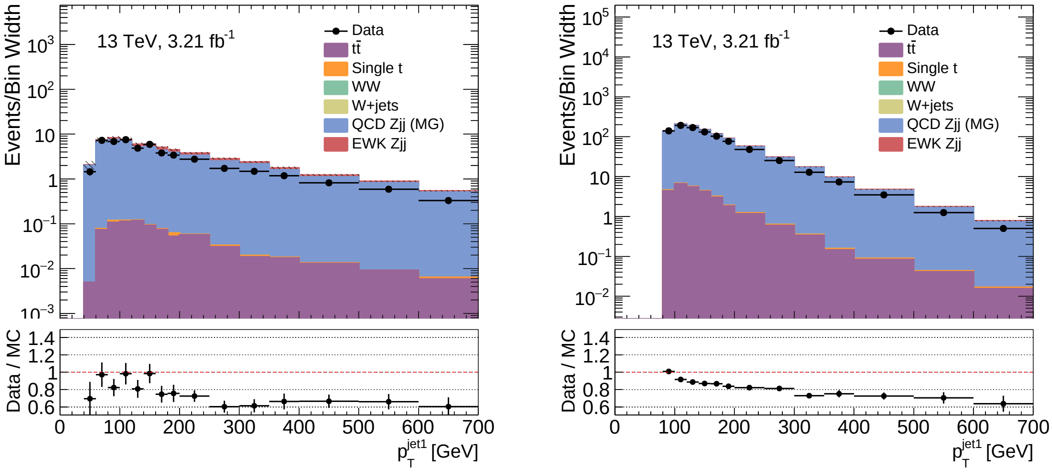
(b) p_{T}^{jet1} (*high-p_T* region).



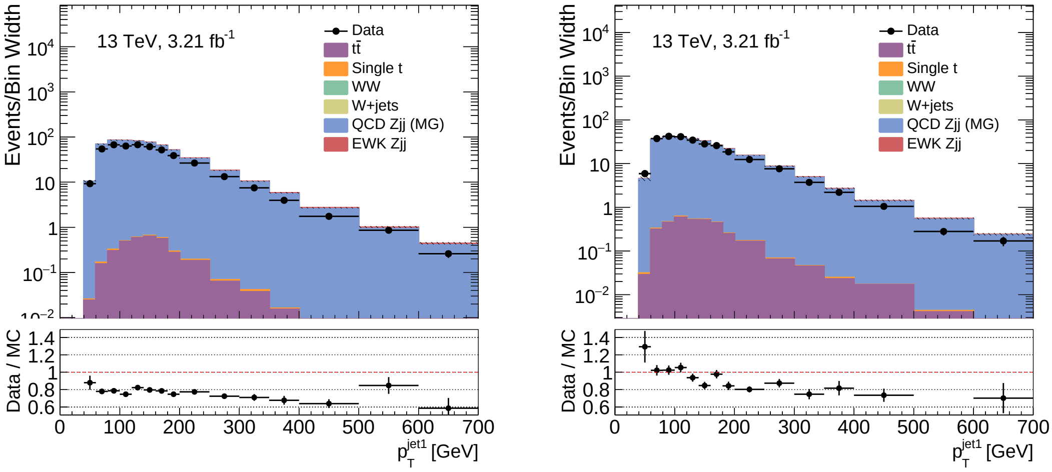
(c) p_{T}^{jet1} (*search* region).

(d) p_{T}^{jet1} (*control* region).

Figure 6.16: Reconstruction-level comparisons of the leading jet p_{T} distribution between data and Monte Carlo in different fiducial regions, for the electron and muon channel combined. The QCD *Zjj* MC sample comes from the ALPGEN generator. The hashed area corresponds to the stacked MC samples uncertainty. Uncertainties shown are statistical only.

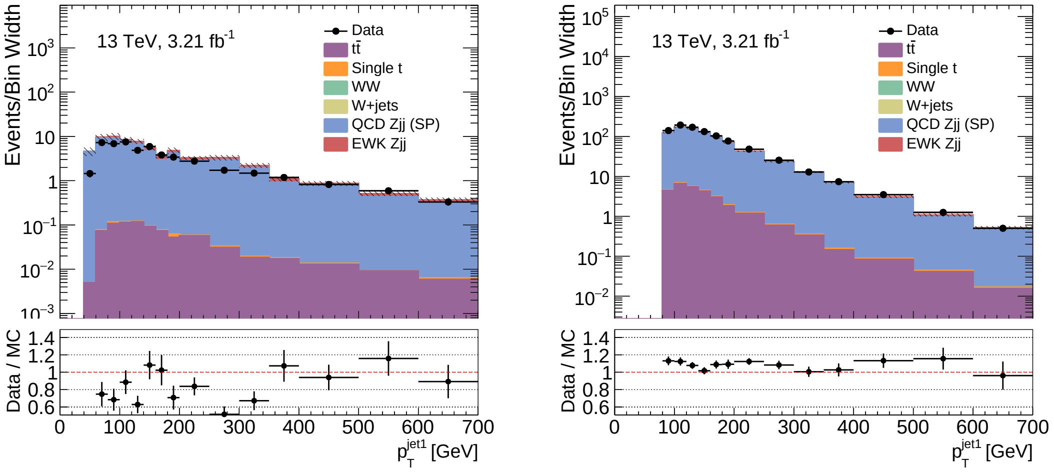


(a) p_{T}^{jet1} (*high-mass* region). (b) p_{T}^{jet1} (*high-p_T* region).



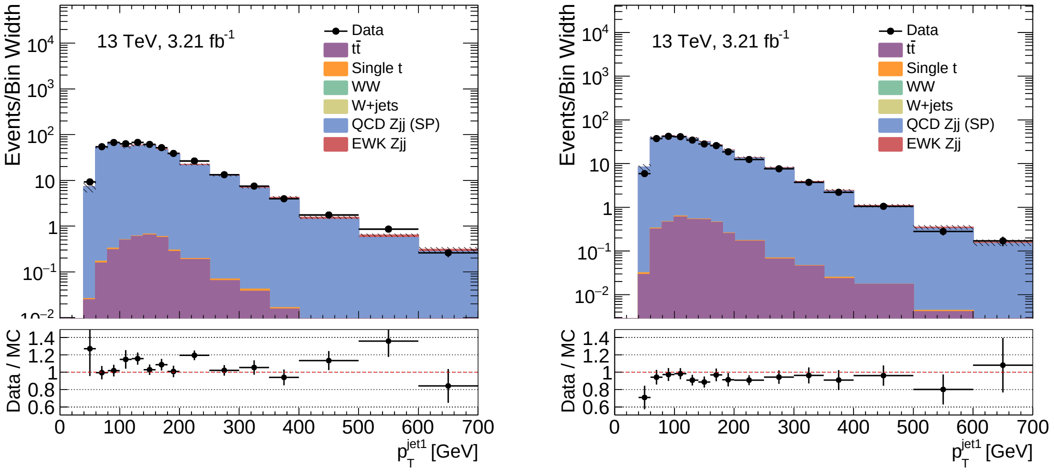
(c) p_{T}^{jet1} (*search* region). (d) p_{T}^{jet1} (*control* region).

Figure 6.17: Reconstruction-level comparisons of the leading jet p_{T} distribution between data and Monte Carlo in different fiducial regions, for the electron and muon channel combined. The QCD *Zjj* MC sample comes from the MadGraph generator. The hashed area corresponds to the stacked MC samples uncertainty. Uncertainties shown are statistical only.



(a) p_{T}^{jet1} (*high-mass* region).

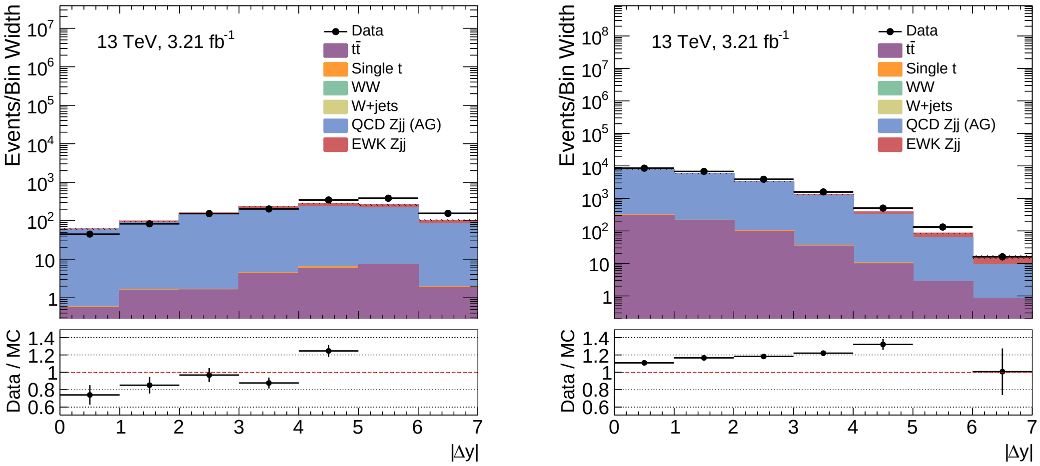
(b) p_{T}^{jet1} (*high-p_T* region).



(c) p_{T}^{jet1} (*search* region).

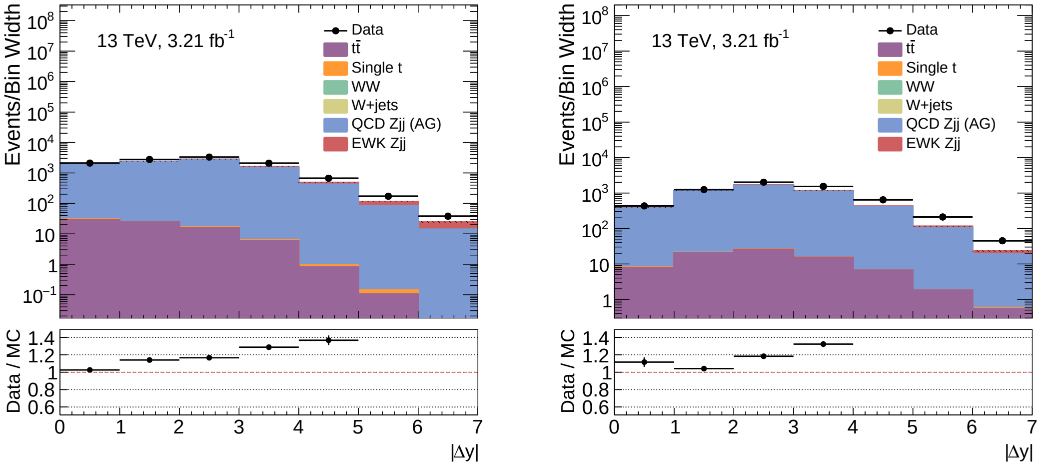
(d) p_{T}^{jet1} (*control* region).

Figure 6.18: Reconstruction-level comparisons of the leading jet p_{T} distribution between data and Monte Carlo in different fiducial regions, for the electron and muon channel combined. The QCD *Zjj* MC sample comes from the SHERPA generator. The hashed area corresponds to the stacked MC samples uncertainty. Uncertainties shown are statistical only.



(a) $|\Delta y|$ (*high-mass* region).

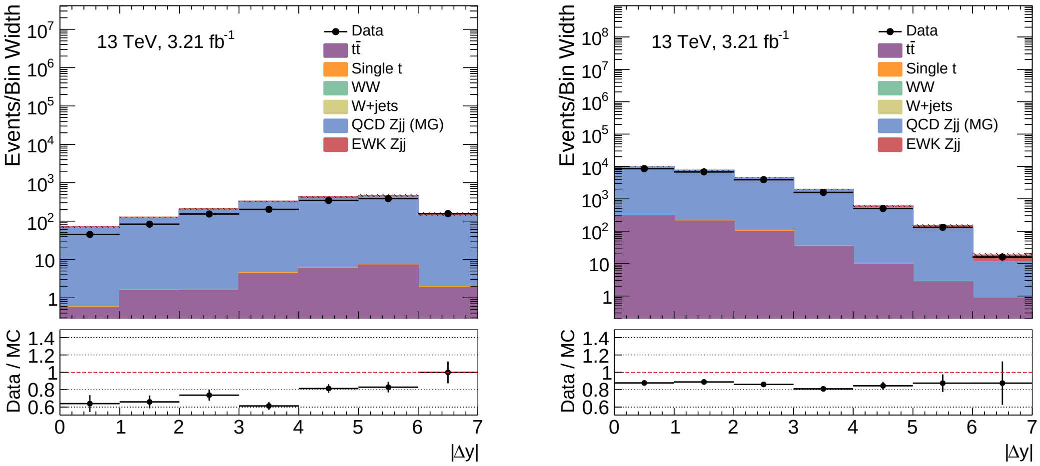
(b) $|\Delta y|$ (*high-p_T* region).



(c) $|\Delta y|$ (*search* region).

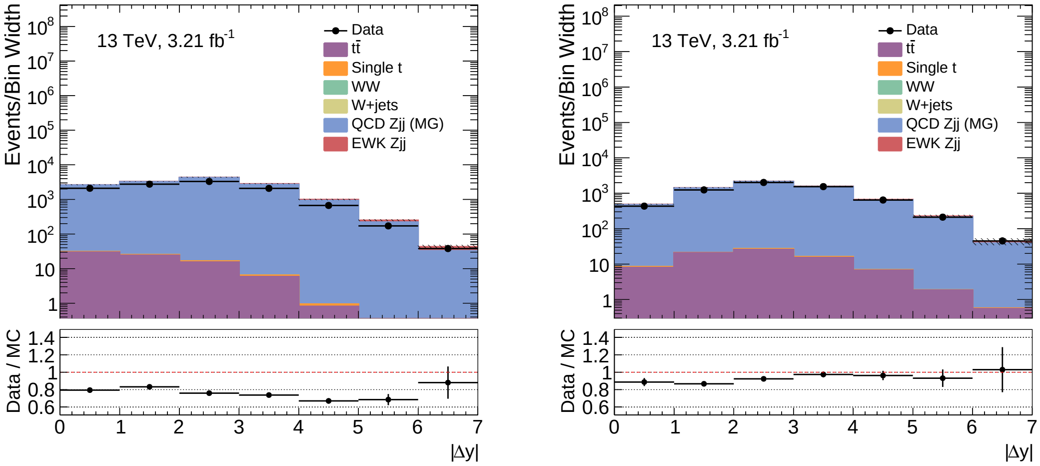
(d) $|\Delta y|$ (*control* region).

Figure 6.19: Reconstruction-level comparisons of the $|\Delta y|$ between the two leading jets distribution between data and Monte Carlo in different fiducial regions, for the electron and muon channel combined. The QCD *Zjj* MC sample comes from the ALPGEN generator. The hashed area corresponds to the stacked MC samples uncertainty. Uncertainties shown are statistical only.



(a) $|\Delta y|$ (*high-mass* region).

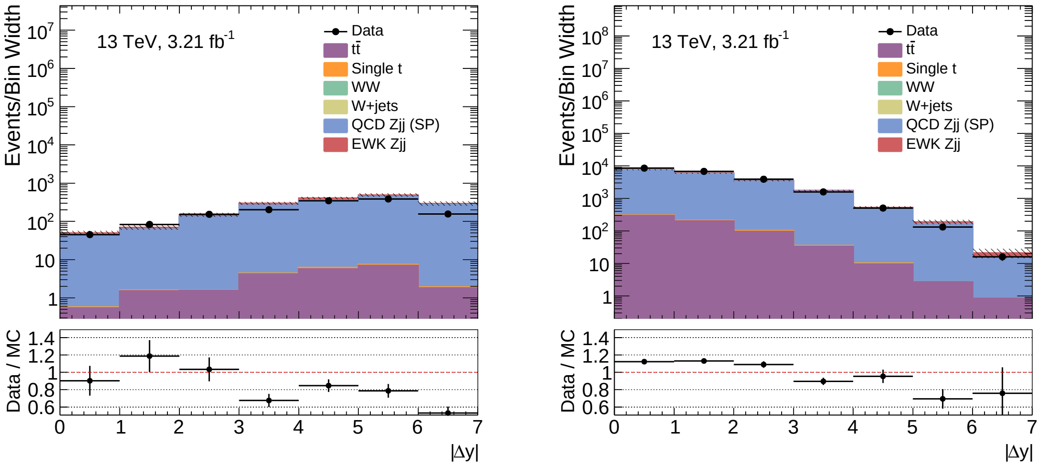
(b) $|\Delta y|$ (*high-p_T* region).



(c) $|\Delta y|$ (*search* region).

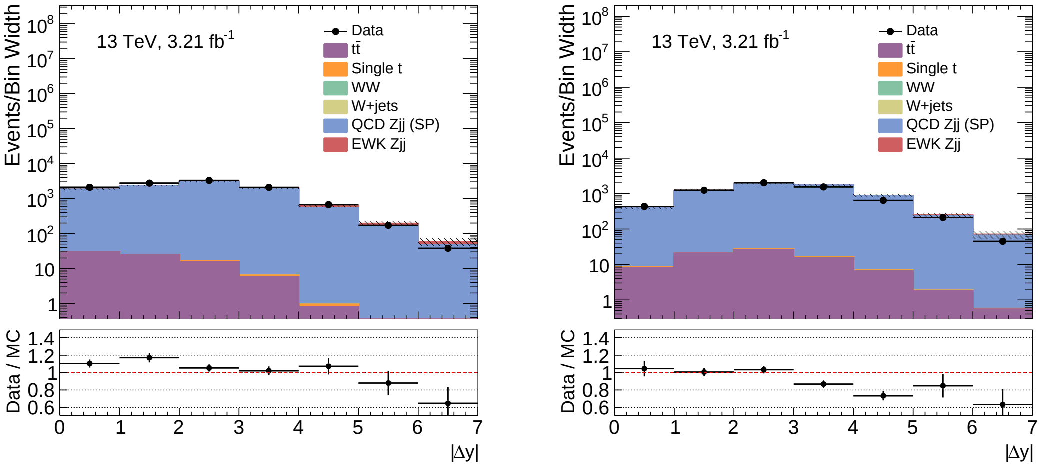
(d) $|\Delta y|$ (*control* region).

Figure 6.20: Reconstruction-level comparisons of the $|\Delta y|$ between the two leading jets distribution between data and Monte Carlo in different fiducial regions, for the electron and muon channel combined. The QCD *Zjj* MC sample comes from the MadGraph generator. The hashed area correMGonds to the stacked MC samples uncertainty. Uncertainties shown are statistical only.



(a) $|\Delta y|$ (*high-mass* region).

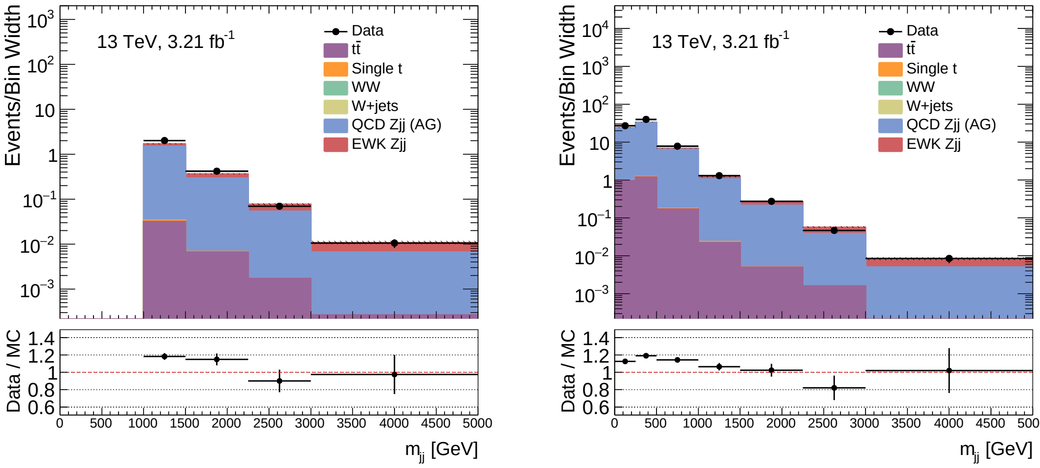
(b) $|\Delta y|$ (*high-p_T* region).



(c) $|\Delta y|$ (*search* region).

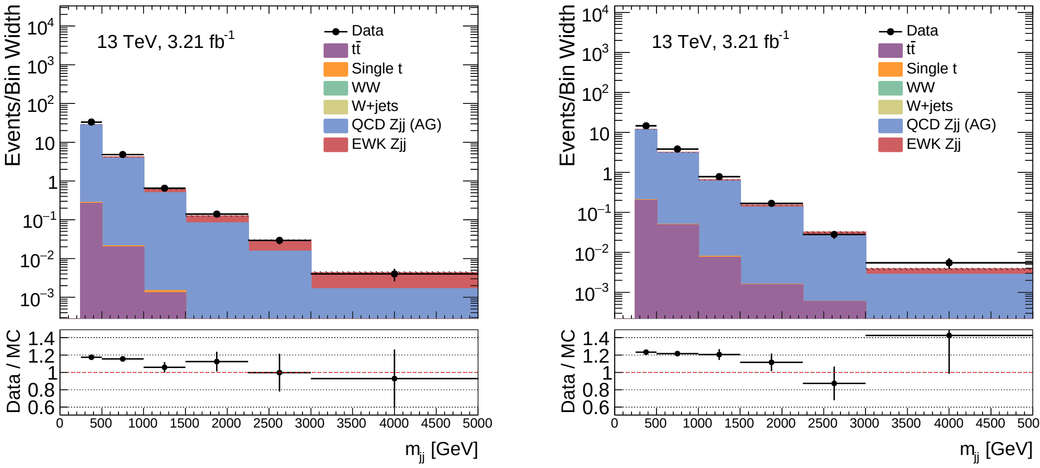
(d) $|\Delta y|$ (*control* region).

Figure 6.21: Reconstruction-level comparisons of the $|\Delta y|$ between the two leading jets distribution between data and Monte Carlo in different fiducial regions, for the electron and muon channel combined. The QCD *Zjj* MC sample comes from the SHERPA generator. The hashed area corresponds to the stacked MC samples uncertainty. Uncertainties shown are statistical only.



(a) m_{jj} (*high-mass* region).

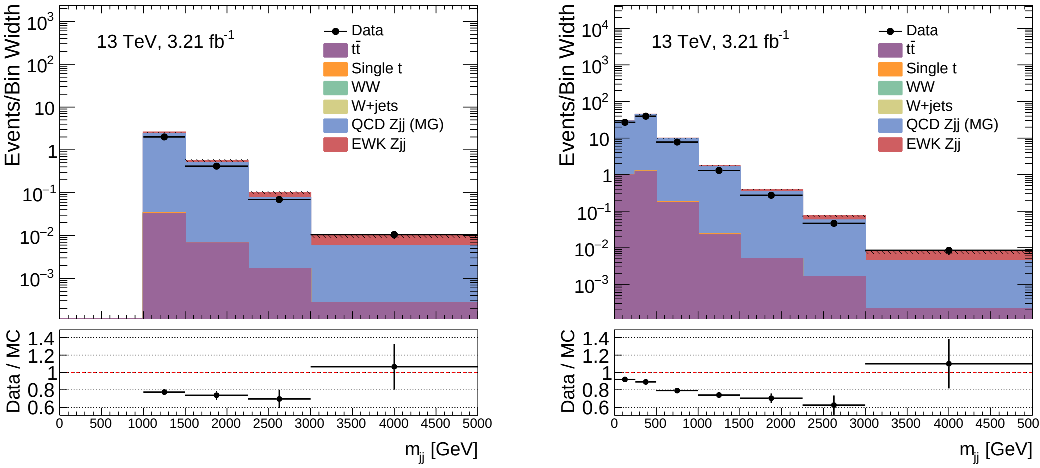
(b) m_{jj} (*high-p_T* region).



(c) m_{jj} (*search* region).

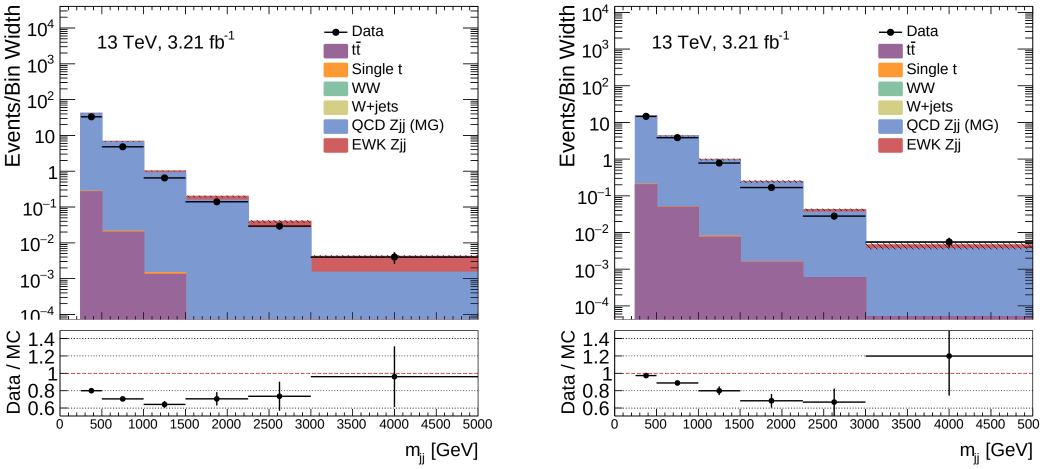
(d) m_{jj} (*control* region).

Figure 6.22: Reconstruction-level comparisons of the m_{jj} distribution between data and Monte Carlo in different fiducial regions, for the electron and muon channel combined. The QCD *Zjj* MC sample comes from the ALPGEN generator. The hashed area corresponds to the stacked MC samples uncertainty. Uncertainties shown are statistical only.



(a) m_{jj} (*high-mass* region).

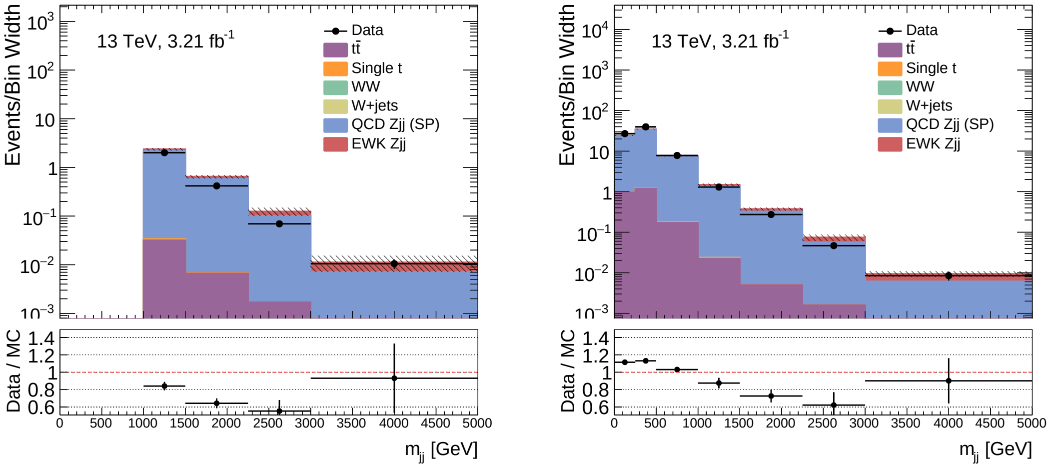
(b) m_{jj} (*high-p_T* region).



(c) m_{jj} (*search* region).

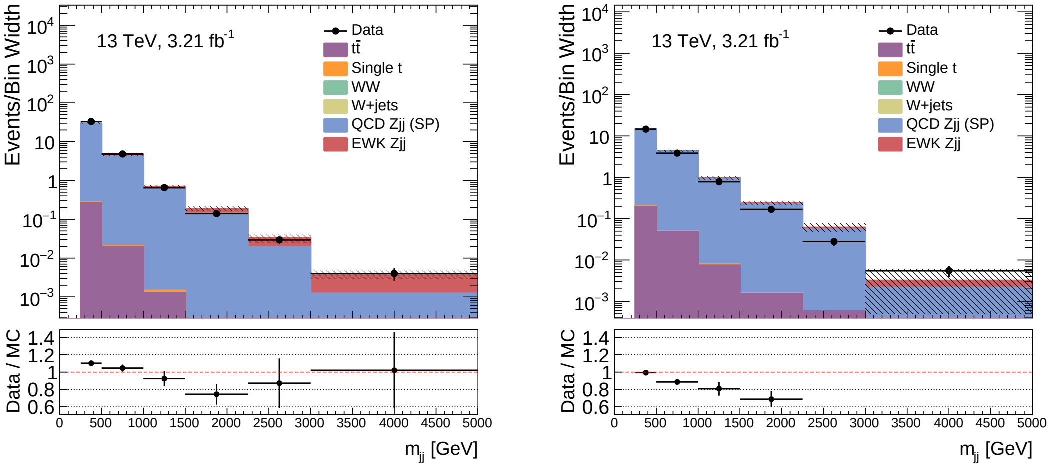
(d) m_{jj} (*control* region).

Figure 6.23: Reconstruction-level comparisons of the m_{jj} distribution between data and Monte Carlo in different fiducial regions, for the electron and muon channel combined. The QCD *Zjj* MC sample comes from the MadGraph generator. The hashed area corresponds to the stacked MC samples uncertainty. Uncertainties shown are statistical only.



(a) m_{jj} (*high-mass* region).

(b) m_{jj} (*high-p_T* region).

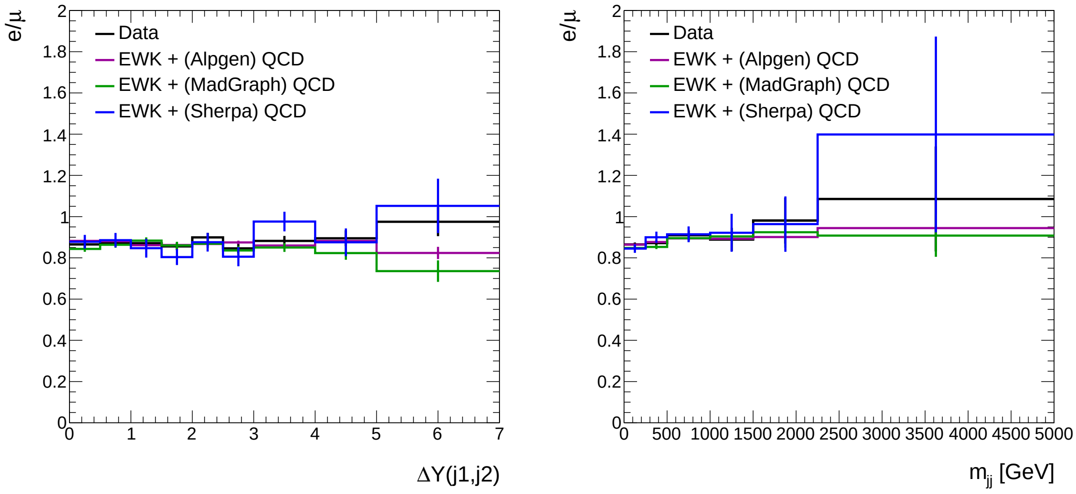


(c) m_{jj} (*search* region).

(d) m_{jj} (*control* region).

Figure 6.24: Reconstruction-level comparisons of the m_{jj} distribution between data and Monte Carlo in different fiducial regions, for the electron and muon channel combined. The QCD *Zjj* MC sample comes from the Sherpa generator. The hashed area corresponds to the stacked MC samples uncertainty. Uncertainties shown are statistical only.

Finally, Figure 6.25 shows the ratio of the electron and muon channel distributions of two observables, m_{jj} and $|\Delta y|$ between the leading jets, for data and for the *Zjj* QCD and electroweak samples together. This comparison shows that both channels behave similarly and therefore can be combined. Additionally, these figures show that the data and MC discrepancies that start to arise at high m_{jj} and high $|\Delta y|$ between the leading jets is not a feature of an individual channel but it is observed in both.



(a) (b)

Figure 6.25: Ratio of the (a) $|\Delta y|$ between the leading jets and (b) m_{jj} for the electron and muons channels, for data and for the QCD and EWK *Zjj* samples together.

6.7 Fiducial cross-section measurements of inclusive *Zjj* production

This section presents the cross-section measurements of inclusive *Zjj* production, which includes QCD and electroweak *Zjj* events, in the five fiducial regions defined in Section 6.5.

The fiducial cross-section, σ_{fid}, for inclusive *Zjj* production is defined as:

$$\sigma_{fid} = \frac{N_{obs} - N_{bkg}}{\int \mathscr{L}\mathrm{d}t \cdot \mathcal{C}} \tag{6.2}$$

where N_{obs} is the number of events observed in the data passing the selection requirements at reconstruction level, N_{bkg} is the number of expected background events ($t\bar{t}$, single-t, WW dibosons and W+jets), $\int \mathscr{L}\mathrm{d}t$ is the integrated luminosity and $\mathcal{C}$ is a correction factor which accounts for differences in event yields at reconstruction and truth level due to detector inefficiencies and resolutions.

The correction factor is calculated as:

$$\mathcal{C} = \frac{N_{reco}}{N_{truth}} \tag{6.3}$$

where N_{reco} is the number of signal events that satisfy the selection criteria at reconstruction level, and N_{truth} is the number of signal events that pass the same selection but at truth level. These numbers are derived for each fiducial region.

In order to calculate the MC statistical uncertainty on the correction factor, it should be rewritten as:

$$\mathcal{C} = \frac{r+b}{t+b}, \tag{6.4}$$

where the MC predictions at reco and truth level have been split into events that pass the selection at both truth and reconstruction levels (b), events that pass the selection only at truth level (t) and events that pass the selection only at reconstruction level (r). The uncertainty on the correction factor can then be calculated as:

$$\sigma_{\mathcal{C}} = \frac{1}{(t+b)^2} \times \sqrt{(\sigma_r \times (t+b))^2 + (\sigma_t \times (r+b))^2 + (\sigma_b \times (t-r))^2}. \tag{6.5}$$

The total statistical uncertainty on the observed cross-section can then be calculated using standard error propagation:

$$\Delta\sigma_{fid} = \sigma_{fid} \times \sqrt{\frac{(\Delta N_{obs})^2 + (\Delta N_{bkg})^2}{(N_{obs} - N_{bkg})^2} + \left(\frac{\sigma_{\mathcal{C}}}{\mathcal{C}}\right)^2}. \tag{6.6}$$

The results obtained are presented in Table 6.5 and in Figure 6.26, for the electron and muon channels combined and all of the fiducial regions. Table 6.5 presents the measured values of σ_{fid} and the factors needed in Equation 6.2 to calculate it; the correction factor $\mathcal{C}$ presented is the weighted average of the values resulting from using each of the three available QCD *Zjj* samples and the systematic uncertainty on this value is equal to the difference between the Alpgen and Sherpa values, since these two were the most different. Tables 6.6 and 6.7 present the theory prediction for σ_{fid} provided by different generators and the ratio of the measured value to the theory prediction.

Additionally, the experimental systematic uncertainties due to the jet energy scale (JES), the jet energy resolution (JER), JVT and other components such as the leptons scale factors and pile-up reweighting were calculated for each fiducial region, and are presented in Table 6.8. These were calculated for the correction factor $\mathcal{C}$ and N_{bkg} and added in quadrature. Finally, Figure 6.26 summarises the results showing the measured fiducial cross-sections including statistical and systematic uncertainties from Tables 6.5 and 6.8. This figure also shows the expected σ_{fid} using the SHERPA sample to simulate electroweak *Zjj* and each of the available QCD *Zjj* samples. The bottom of the figure shows the ratios of data to MC expectation, for all of the fiducial regions.

Fiducial Region	N_{obs}	N_{bkg}	$\mathcal{C}$	$\mathcal{C}$ Syst. Unc.	σ_{fid} [pb]
baseline	64748	1666.26	0.745 ± 0.001	0.022	13.230 ± 0.078
high-p_T	21461	665.67	0.721 ± 0.002	0.015	4.504 ± 0.046
search	11145	80.03	0.661 ± 0.002	0.025	2.615 ± 0.036
control	6127	81.78	0.751 ± 0.003	0.032	1.258 ± 0.023
high-mass	1390	23.76	0.776 ± 0.007	0.053	0.275 ± 0.011

Table 6.5: Inclusive *Zjj* fiducial cross-sections and the elements needed to calculate it. The values presented are for the electron and muon channels combined. N_{obs} is the number of data events observed in the respective fiducial region and N_{bkg} is the expected contribution due to non-*Zjj* backgrounds ($t\bar{t}$, single-t, diboson and W+jets). Statistical and systematic uncertainties are shown for the correction factor $\mathcal{C}$ and only statistical errors are presented for the fiducial cross-section.

	ALPGEN	
Fiducial Region	σ_{Theory} [pb]	$\sigma_{Data}/\sigma_{Theory}$
baseline	10.886 ± 0.014	1.215 ± 0.007
high-p_T	3.867 ± 0.008	1.165 ± 0.012
search	2.236 ± 0.006	1.170 ± 0.016
control	1.022 ± 0.003	1.231 ± 0.023
high-mass	0.235 ± 0.002	1.171 ± 0.047

Table 6.6: Theory predictions of the fiducial cross-sections for inclusive *Zjj* and the ratio to the measured values. Results were calculated using the SHERPA EWK *Zjj* samples and the ALPGEN QCD *Zjj* sample.

	MadGraph		Sherpa	
Fiducial Region	σ_{Theory} [pb]	$\sigma_{Data}/\sigma_{Theory}$	σ_{Theory} [pb]	$\sigma_{Data}/\sigma_{Theory}$
baseline	14.441 ± 0.051	0.916 ± 0.006	13.111 ± 0.068	1.009 ± 0.008
high-p_T	5.185 ± 0.022	0.869 ± 0.010	4.194 ± 0.040	1.074 ± 0.015
search	3.355 ± 0.021	0.779 ± 0.012	2.512 ± 0.026	1.041 ± 0.018
control	1.360 ± 0.012	0.925 ± 0.019	1.404 ± 0.023	0.897 ± 0.022
high-mass	0.360 ± 0.005	0.763 ± 0.032	0.381 ± 0.010	0.723 ± 0.034

Table 6.7: Theory predictions of the fiducial cross-sections for inclusive *Zjj* and the ratio to the measured values. Results were calculated using the Sherpa EWK *Zjj* samples and either the MadGraph or the Sherpa QCD *Zjj* sample.

	Experimental Systematic Uncertainty [%]				
Fiducial Region	JER	JES	JVT	Rest	Total
baseline	+0.98	+8.11/–7.27	+3.71/–3.46	+1.72/–1.18	+9.14/–8.14
high-p_T	+0.70	+6.37/–6.07	+2.71/–2.52	+1.78/–1.30	+7.18/–6.70
search	+1.86	+3.75/–4.05	+1.62/–1.67	+1.85/–1.68	+4.86/–4.69
control	+4.38	+11.76/–13.01	+4.80/–4.55	+1.84/–2.83	+13.56/–14.08
high-mass	+7.67	+10.45/– 8.62	+3.70/–3.15	+4.66/–2.22	+14.27/–9.44

Table 6.8: Experimental systematic uncertainties calculated for every fiducial region. These take into account the systematic uncertainties affecting the correction factor $\mathcal{C}$ and N_{bkg}. The column “Rest” refers to the systematic uncertainties related to the leptons’ identification, isolation and triggers, and pile-up reweighting.

These tables and figures show that the observed σ_{fid} are in reasonable agreement with the expected values provided by the different generators. Additionally, the cross-sections measured are a factor of ~ 2.5 times greater than those measured by the 8 TeV analysis for every fiducial region, except for the *high-mass* region, where the factor is ~ 4.5. These factors are greater than the 13/8 factor coming from the $\sqrt{s}$ increase, showing that the parton luminosity enhancement is also contributing.

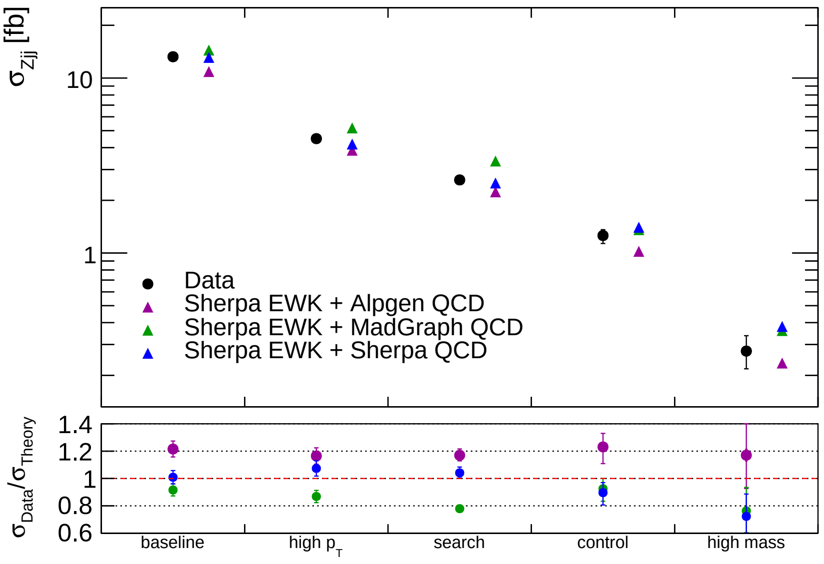


Figure 6.26: Inclusive *Zjj* cross-sections in five fiducial regions. The measurement is compared to theory predictions based on SHERPA for EWK *Zjj* and ALPGEN, MADGRAPH or SHERPA for QCD *Zjj*. Statistical uncertainties are included as well as the systematic uncertainties shown in Tables 6.5 and 6.8.

6.8 Fiducial cross-section measurements of Electroweak *Zjj* production

In the previous section, the QCD plus electroweak *Zjj* production cross-section was measured in QCD and electroweak enhanced fiducial regions. However, it was shown in Section 6.5 that the electroweak component of *Zjj* has a different shape in m_{jj} in the *search* region than the rest of the samples. Using this fact, a purely electroweak *Zjj* cross-section value can be extracted.

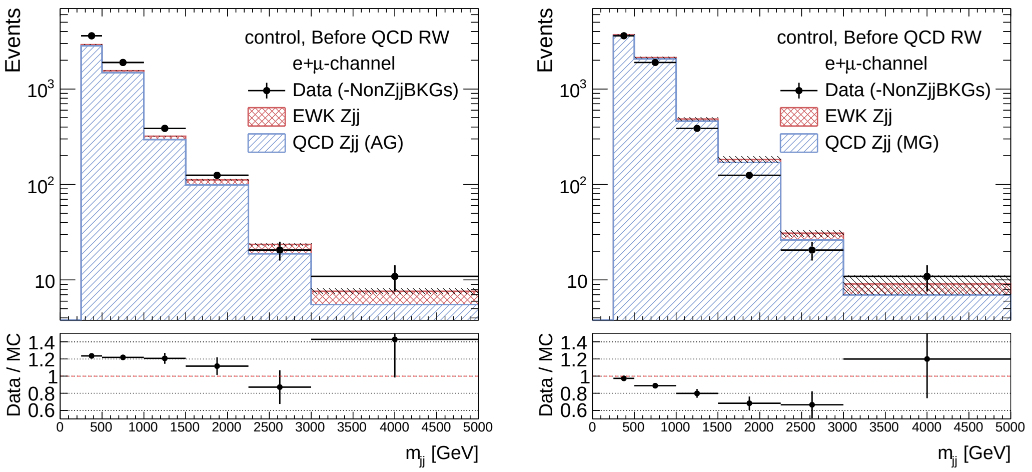
First, non-*Zjj* background estimates are subtracted from the data bin-by-bin in m_{jj} using MC simulations. The size of these subtractions ranges from $< 0.1\%$ to 3.2% across m_{jj}. This background-subtracted m_{jj} data distribution can then, in principle, be fitted with a template for QCD *Zjj* production from Monte Carlo in the low mass region where electroweak production contributions to the total rate are negligible, and the electroweak cross-section extracted from the observed deficit between the predicted QCD rate and the measured data.

This procedure uses the low dijet mass data to constrain the overall QCD production rate, but is reliant on having a good model for the shape of the QCD *Zjj* dijet mass spectrum. However, since the *search* region is a high p_{T}, high dijet mass, wide jet rapidity separation fiducial region, the QCD *Zjj* template is affected by the well-documented QCD vector boson plus dijet mismodelling shown before and seen by previous analyses and experiments. Therefore, this modelling needs to be corrected before the QCD *Zjj* template can be reliably used to extract the electroweak cross-section.

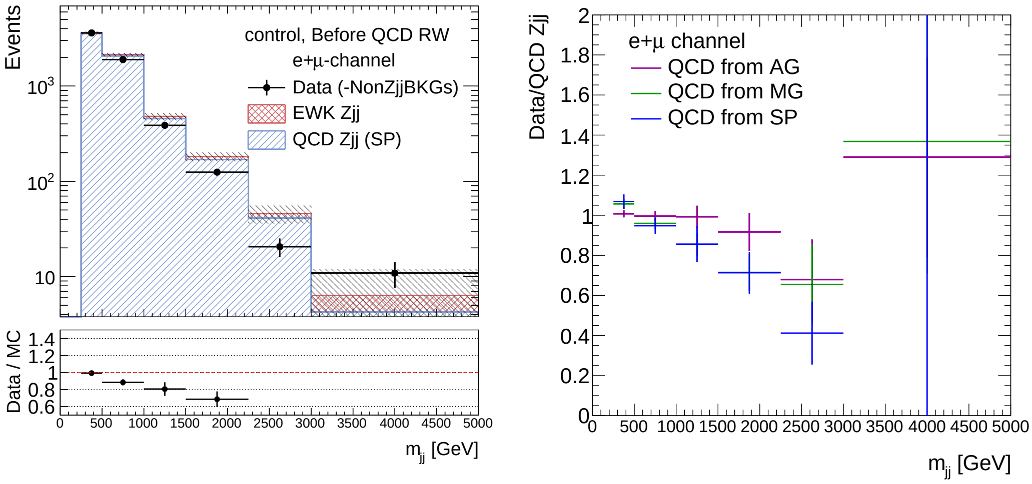
A methodology similar to the 8 TeV analysis [40] is followed to construct a QCD *Zjj* template which is a combination of MC and data. While the data in the *search* region cannot be used to constrain the modelling of the QCD *Zjj* dijet mass shape, due to the presence of the electroweak production sources that this analysis is trying to measure, the *control* region has been designed to have identical kinematics except for the inversion of the jet veto requirement and so can serve as a data-driven constraint on the QCD *Zjj* shape.

Therefore, in the *control* region, where the signal is suppressed, a reweighting function for QCD *Zjj* is extracted by taking the ratio of the data and QCD *Zjj* and fitting it with a polynomial. All other backgrounds and the electroweak *Zjj* events are subtracted from data before performing this ratio, although their contributions are very small, from the definition of the *control* region (see Section 6.5) and as is shown in Tables 6.3 and 6.4.

Figures 6.27a, 6.27b and 6.27c show the m_{jj} distributions in the *control* region of data minus the non-*Zjj* contributions compared to that of QCD *Zjj* and electroweak *Zjj*, for the three different QCD samples. Then, Figure 6.27d shows the ratio of data minus non-QCD *Zjj* to QCD *Zjj* for the three different samples together. These figures show that all generators are consistent with each other within their statistical uncertainties. These figures show the electron and muon channels combined but the same procedure is applied to the electron and muon channel separately, for consistency checks.



(a) Before reweighting, QCD *Zjj* sample from ALPGEN (b) Before reweighting, QCD *Zjj* sample from MADGRAPH

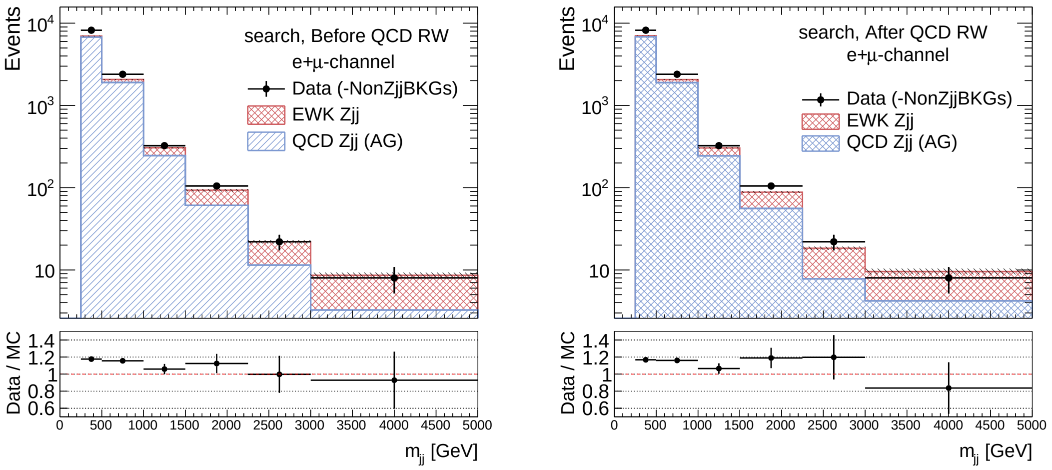


(c) Before reweighting, QCD *Zjj* sample from SHERPA (d) Reweighting functions using different QCD *Zjj* samples

Figure 6.27: Figures (a), (b) and (c) show m_{jj} distributions in the *control* region of data minus non-*Zjj* contributions predicted from MC, QCD *Zjj* and electroweak *Zjj*, for the combined electron and muon channels, for the different available QCD *Zjj* samples. Figure (d) shows the ratio of data minus non-QCD *Zjj* contributions to QCD *Zjj* (normalised to data) for the three different QCD *Zjj* generators together.

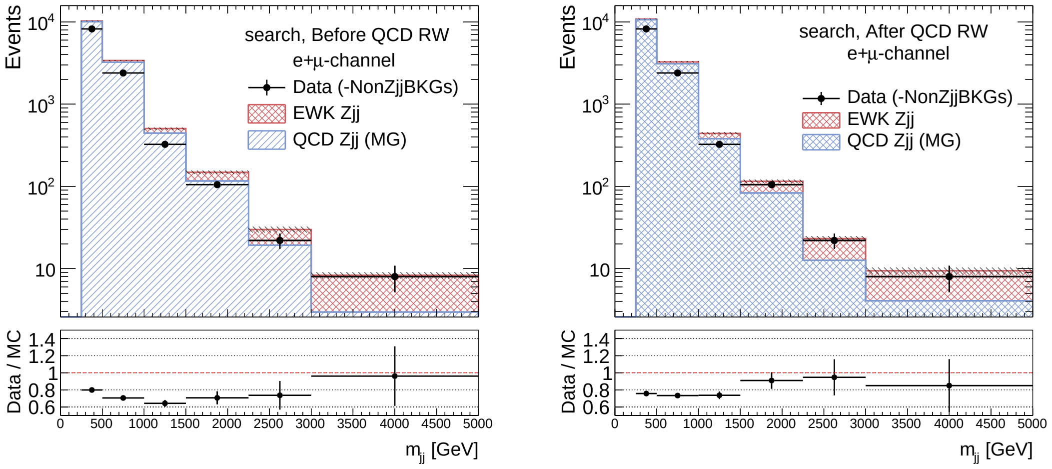
Once the reweighting function is calculated, it is applied directly to the QCD *Zjj* sample in the signal-enhanced *search* region. This is possible due to the fact that the *control* and *search* region differ only by the emission of an additional jet in the rapidity gap between the two leading jets, so the mismodelling of the two leading jets is likely to be similar for the two regions. Figures 6.28-6.30 show the ratio of the electroweak *Zjj* and QCD *Zjj* templates to the data (minus the non-*Zjj* backgrounds) before and after applying the reweighting to the QCD *Zjj* for each available QCD

Zjj sample and Figure 6.31 shows all the available QCD *Zjj* samples reweighted together. The reweighting function was applied by scaling the bin contents of the QCD *Zjj* samples by the value of the ratio histogram extracted in the *control* region, shown in Figure 6.27d.



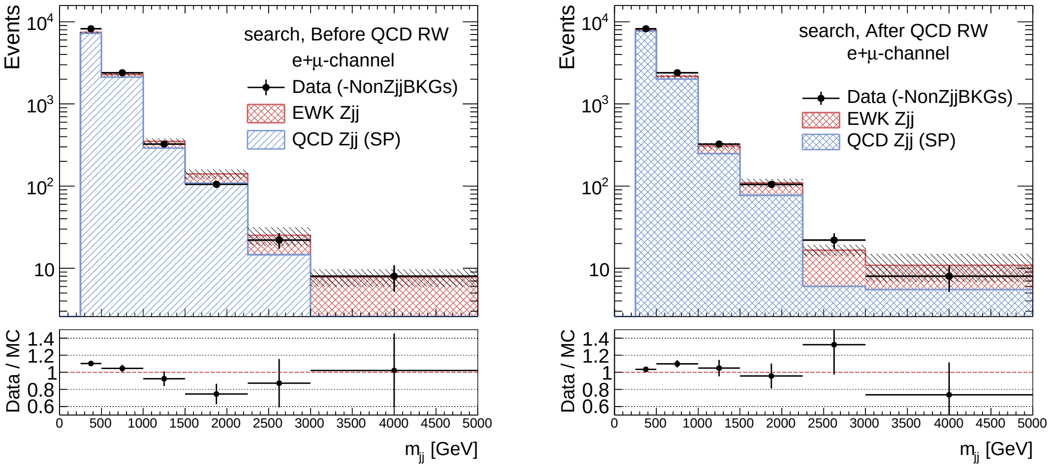
(a) Before reweighting (b) After reweighting

Figure 6.28: Comparison in the *search* region of the EWK *Zjj* and QCD *Zjj* templates to the data (minus the non-*Zjj* backgrounds) (a) before the reweighting to QCD is applied and (b) after the reweighting is applied, using the QCD *Zjj* sample form Alpgen.



(a) Before reweighting (b) After reweighting

Figure 6.29: Comparison in the *search* region of the EWK *Zjj* and QCD *Zjj* templates to the data (minus the non-*Zjj* backgrounds) (a) before the reweighting to QCD is applied and (b) after the reweighting is applied, using the QCD *Zjj* sample form MadGraph.



(a) Before reweighting

(b) After reweighting

Figure 6.30: Comparison in the *search* region of the EWK *Zjj* and QCD *Zjj* templates to the data (minus the non-*Zjj* backgrounds) (a) before the reweighting to QCD is applied and (b) after the reweighting is applied, using the QCD *Zjj* sample form Sherpa.

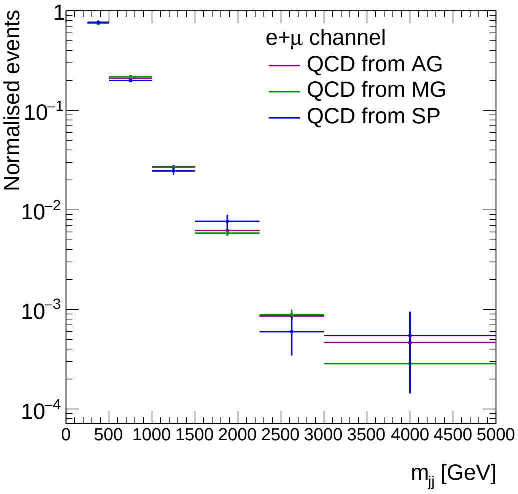
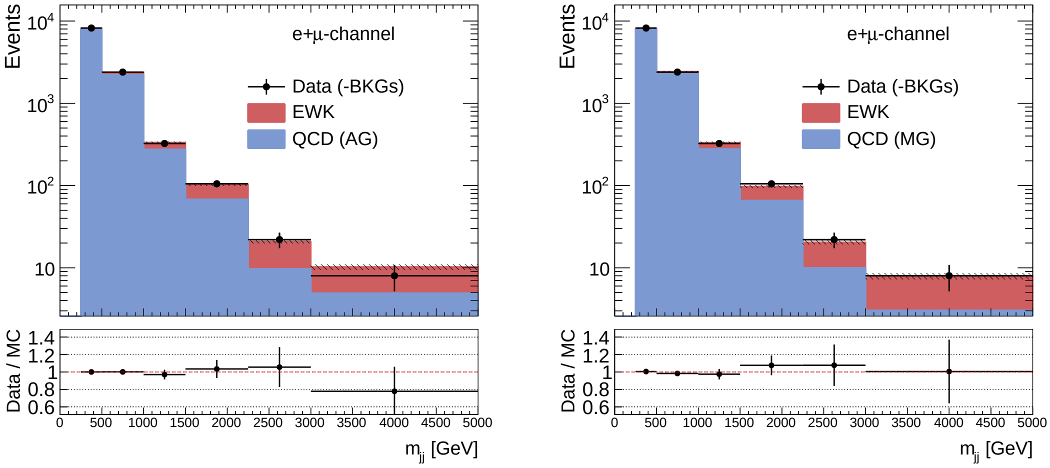


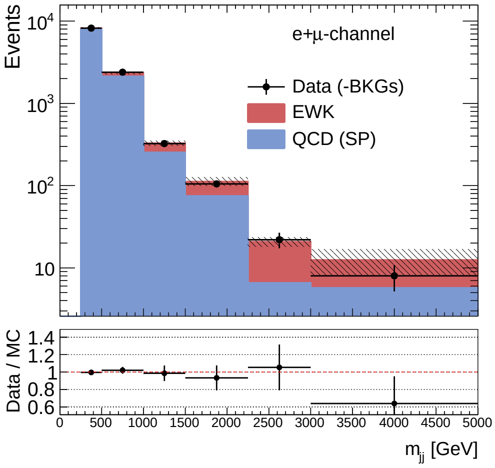
Figure 6.31: Shape comparison of the three available QCD templates after being reweighted using the data/QCD ratios calculated in the *control* region.

Once the QCD *Zjj* template was reweighted, the function TFractionFitter [76] from the ROOT analysis software [77] was used to fit the data minus non-*Zjj* backgrounds with the electroweak and QCD *Zjj* templates. Figure 6.32 shows the results of the fit for the combined electron and muon channels, and using the Alpgen, MadGraph or Sherpa QCD *Zjj* samples. The reweighting of the QCD samples was done using the ratio values derived in the *control* region. Figure 6.31 compares the shape of the reweighted QCD template for the three different available QCD samples.



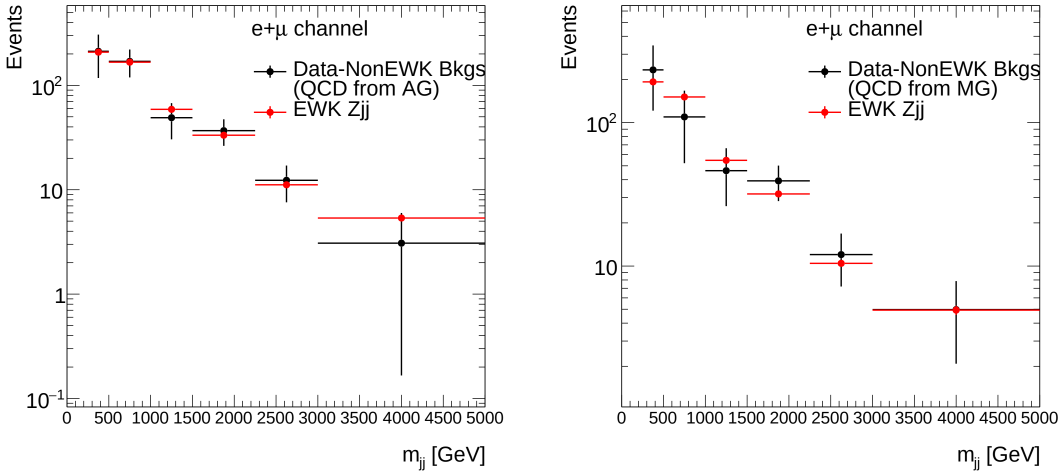
(a) Alpgen

(b) MadGraph

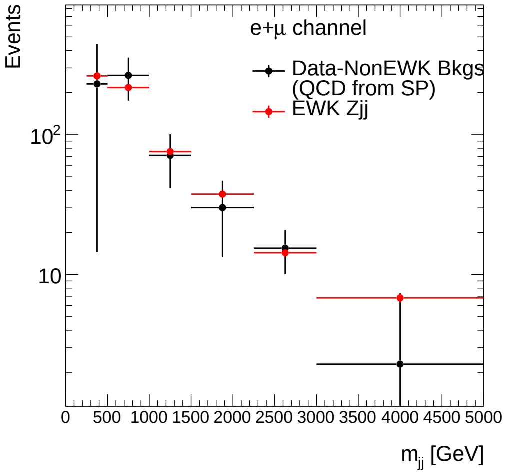


(c) Sherpa

Figure 6.32: Results of fitting the data with the electroweak and QCD Zjj templates, the last one coming from (a) Alpgen, (b) MadGraph or (c) Sherpa, in the *search* region using TFractionFitter and reweighting the QCD Zjj samples using the data/QCD ratios derived in the *control* region.



(a) Alpgen (b) MadGraph



(c) Sherpa

Figure 6.33: Comparison of data minus the fitted QCD template to the fitted EWK template in the *search* region. Results are shown for the cases of using each of the three available QCD templates

The electroweak *Zjj* cross-section was calculated as:

$$\sigma_{\mathrm{EWK}} = \frac{N_{\mathrm{data}} - N_{\mathrm{QCD,rw,sc}}}{\int \mathscr{L}\mathrm{d}t \cdot \mathcal{C}} \tag{6.7}$$

where N_{data} is the number of events in data minus the non-*Zjj* backgrounds, $N_{QCD,rw,sc}$ is the number of events in the QCD *Zjj* template after reweighting and after fitting, $\mathcal{C}$ is the correction factor for electroweak *Zjj* in the *search* region and $\int \mathscr{L}$ is 3.2 fb^{-1}.

These procedure yielded the following EWK cross-section values, each of which uses a different QCD *Zjj* sample:

$$\sigma_{\text{EWK, Alpgen}} = 115.93 \text{ fb} \tag{6.8}$$

$$\sigma_{\text{EWK, MadGraph}} = 106.90 \text{ fb} \tag{6.9}$$

$$\sigma_{\text{EWK, Sherpa}} = 147.65 \text{ fb} \tag{6.10}$$

The theory prediction for the electroweak Zjj cross-section was also calculated using the SHERPA generator and yielded the following value, which includes only an statistical error:

$$\sigma_{\text{EWK,Theory}} = 115.8 \pm 0.8 \text{ fb}. \tag{6.11}$$

In order to assess the uncertainties on the EWK cross-section, different pseudo-experiments were performed. A set of pseudo-experiments were performed to calculate the uncertainty due to the statistics of the QCD Zjj sample in the search region and the results, using the three different generators, are shown in Figure 6.34. This figure shows that the three different values of σ_{EWK} coming from using the three different generators for the QCD Zjj sample are consistent with each other within uncertainties. It is also clear from this figure that ALPGEN has the best statistical precision and therefore we will use $\sigma_{\text{EWK,Alpgen}}$ as a central value.

The statistical uncertainty was also calculated using pseudo-experiments, varying the data sample in the search region, and this resulted in an uncertainty of 19.6%. Another set of pseudo-experiments where the EWK Zjj sample was randomly varied in the search region yielded a systematic uncertainty of 0.8%. The systematic uncertainty due to the reweighting function, calculated using a set of pseudo-experiments where the data/QCD ratio in the control region was randomly varied, is 16.4%. This uncertainty is driven primarily by MC and data statistics in the *control* region, and is expected to improve with new higher-statistics MC requests being processed.

Finally, the experimental systematic uncertainties due to the JES, the JER, the JVT and other components such as the leptons scale factors and pile-up reweighting were calculated and are presented in Table 6.9 together with the rest of the systematic uncertainties.

The measured electroweak Zjj cross-section is:

$$\sigma_{\text{EWK}} \;=\; 115.9 \;\pm\; 22.7 \text{ (stat.) } {}^{+20.8}_{-21.5} \text{ (syst.) fb.} \tag{6.12}$$

Therefore, the measured cross-section is in agreement with the theory prediction, within its errors. Figure 6.35 shows the ratio of the measured EWK *Zjj* cross-section to the MC prediction obtained by this analysis compared to equivalent ratios of the results published by the 8 TeV analysis [40] and a similar analysis performed by CMS [78]. This figure shows that the results from this analysis are in agreement with the 8 TeV results from ATLAS and CMS.

Systematic Uncertainty	
Source	relative uncertainty
QCD stats in search region	± 4.3
EWK stats in search region	± 0.8
Reweighting function / QCD modelling	± 16.4
JER	+2.3
JES	+5.2, –6.4
JVT	+0.2, –0.2
Leptons, pile-up reweightings,...	+1.5, –3.0
Luminosity	±2.1
Total	+18.0, -18.5

Table 6.9: Sources of systematic uncertainty on the measured electroweak *Zjj* cross-section.

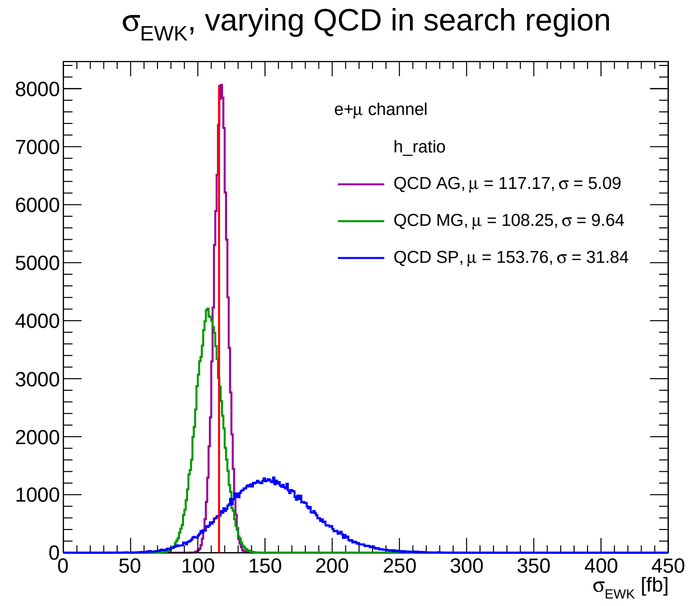


Figure 6.34: Distribution of σ_{EWK} coming from pseudo-experiments where the QCD *Zjj* sample is randomly varied. The results from using the Alpgen (AG), MadGraph (MG) or Sherpa (SP) QCD *Zjj* sample are presented and the red line indicates the theory prediction from the Sherpa generator. All generators are consistent within their uncertainties.

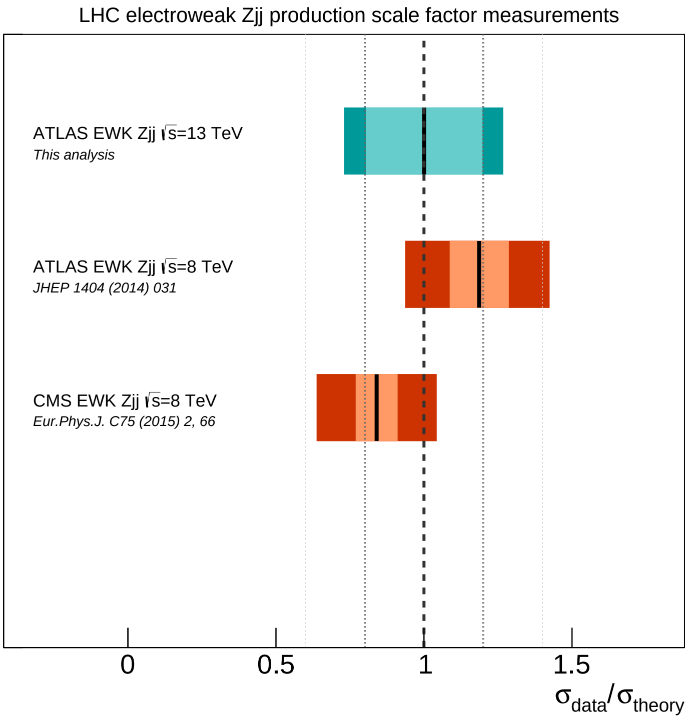


Figure 6.35: Electroweak *Zjj* scale factor measurements at the LHC. The result from this thesis is presented in addition to the results from the 8 TeV analysis and a similar analysis performed by CMS [78].

6.9 Summary and outlook

This chapter presented the fiducial cross-section for inclusive *Zjj* production in five different fiducial regions with varying sensitivity to the electroweak component of the *Zjj* processes. The values measured were consistent with the theory predictions provided by the Alpgen, MadGraph and Sherpa generators. Additionally, a purely electroweak *Zjj* fiducial cross-section was presented in a fiducial region where this component is enhanced. This again was in agreement with the theory prediction provided by the Sherpa generator and the data to theory ratio was consistent with similar analyses performed with 8 TeV data by ATLAS and CMS. The results presented in this chapter represent important steps towards a more complete measurement of the EWK *Zjj* production cross-section.

For both the inclusive and electroweak-only fiducial cross-sections, experimental systematics related to the JER, JES, JVT, lepton scale factors and pile-up reweighting were calculated. The only remaining systematics to be calculated are those due to PDFs and quantum mechanical interference.

One outstanding issue is the statistical limitation of the QCD *Zjj* templates but as mentioned before, large MC samples have been requested and are now being processed. Also, a signal sample has been requested which does not include *ZV* events, to be able to more directly compare the results from this analysis to the 8 TeV one.

The one background that this analysis is missing is the multijet background. There are ongoing studies to produce this. However, its contribution is quite small; in the 8 TeV analysis, this background contributed 0.02% to the *search* region and 0.17% in the *control* region.

The first aim of the analysis presented here is to produce a more complete measurement of electroweak *Zjj* production cross-section with 13 TeV data collected by ATLAS in 2015 and publish a short paper in the summer. In the longer term, this analysis will benefit from a larger 13 TeV dataset, which will be collected in 2016. Therefore, in the future, this analysis could improve its precision on the extraction of the electroweak *Zjj* cross-section. Also, these measurements could improve upon the limits on aTGCs placed by the 8 TeV analysis. This analysis also has the potential to contribute to studies of VBF production of the Higgs boson and in searches for new phenomena, such as dark matter or flavour-violating processes.

Chapter 7

Conclusions

This thesis presented the implementation of a vertex counting algorithm to measure the luminosity delivered to ATLAS during 2012. The algorithm and its corrections for pile-up effects were shown to be robust by means of a Monte Carlo closure test, for different definitions of a *tight* vertex. The algorithm was then calibrated using the 2012 van der Meer scans and the different σ_{vis} values obtained for these scans were presented. It was shown that the method was stable within a scan set at the 0.5% level and the difference between scan sets was around 2%. The results of the calibration were also compared to those obtained by other detectors and algorithms and it was shown that the vertex counting algorithm was consistent with most of them, in particular when comparing the measurements of the specific luminosity, for which the consistency was at the level of 1%. Finally, the vertex counting algorithm was used to measure the luminosity in 2012 physics runs and it was shown that the method had excellent internal consistency, by comparing the results obtained with different definitions of vertices. The method was also shown to have good external consistency, by comparing its luminosity measurements to those obtained with other detectors and algorithms.

This thesis also presented the fiducial cross-section measurement of inclusive *Zjj* production in five different fiducial regions and the results were found to be in agreement with the Standard Model prediction provided by the SHERPA, MADGRAPH and ALPGEN generators, within the uncertainties. Additionally, the electroweak component of the *Zjj* production was extracted by fitting the dijet invariant mass distribution in a region of phase space designed to enhance such a contribution. This cross-section is also in agreement with the theory expectation provided by the SHERPA generator, within the uncertainties. The ratio of the measured cross-section to the theory expectation was compared to the same ratio obtain by the reference analysis performed with $\sqrt{s} = 8\,\text{TeV}$ data and found to be consistent. These early measurements pave the way for a more detailed study of and search for anomalies in production of SM particles through weak boson fusion, which will also drive advances in precision and scope for future new phenomena searches in such production modes with the full LHC Run 2 data.

Bibliography

[1] M. L. David Kaplan, *Particle fever.* U.S.A: Anthos Media, 2013.

[2] SuperNEMO Collaboration, R. Arnold et al., *Probing new physics models of neutrinoless double beta decay with SuperNEMO*, Eur. Phys. J. **C70** (2010) 927–943, `arXiv:1005.1241 [hep-ex]`.

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

[4] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579–588.

[5] J. Goldstone, A. Salam, and S. Weinberg, *Broken Symmetries*, Phys. Rev. **127** (1962) 965–970.

[6] E. Accomando and A. Kaiser, *Electroweak corrections and anomalous triple gauge-boson couplings in W^+W^- and $W^\pm Z$ production at the LHC*, Phys. Rev. **D73** (2006) 093006, `arXiv:hep-ph/0511088 [hep-ph]`.

[7] ATLAS Collaboration, G. Aad et al., *Measurement of W^+W^- production in pp collisions at $\sqrt{s}$=7 TeV with the ATLAS detector and limits on anomalous WWZ and $WW\gamma$ couplings*, Phys. Rev. **D87** no. 11, (2013) 112001, `arXiv:1210.2979 [hep-ex]`. [Erratum: Phys. Rev.D88,no.7,079906(2013)].

[8] L. Evans and P. Bryant, *LHC machine*, JINST **3** (2008) S08001.

[9] D. Brandt, H. Burkhardt, M. Lamont, S. Myers, and J. Wenninger, *Accelerator physics at LEP*, Rept. Prog. Phys. **63** (2000) 939–1000.

[10] CERN, *CERN accelerator complex.* `http://bigscience.web.cern.ch/bigscience/Objects/LHC/accelerator.jpg`. Accessed: 2015-11-15.

[11] LHCb Collaboration, J. Alves et al., *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

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

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

[14] ATLAS Collaboration, G. Aad et al., *The ATLAS experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[15] 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) 1–29, `arXiv:1207.7214 [hep-ex]`.

[16] 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) 30–61, `arXiv:1207.7235 [hep-ex]`.

[17] ATLAS Collaboration, *Track reconstruction performance of the ATLAS inner detector at $\sqrt{s} = 13$ TeV*, Tech. Rep. ATL-PHYS-PUB-2015-018, 2015. `http://cds.cern.ch/record/2037683`.

[18] ATLAS Collaboration, G. Aad et al., *Performance of the ATLAS trigger system in 2010*, Eur. Phys. J. **C72** (2012) 1849, `arXiv:1110.1530 [hep-ex]`.

[19] ATLAS Collaboration, *2015 start-up trigger menu and initial performance assessment of the ATLAS trigger using Run-2 data*, Tech. Rep. ATL-DAQ-PUB-2016-001, 2016. `https://cds.cern.ch/record/2136007`.

[20] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data*, Tech. Rep. ATLAS-CONF-2014-032, 2014. `https://cds.cern.ch/record/1706245`.

[21] S. van der Meer, *Calibration of the effective beam height in the ISR*, Tech. Rep. CERN-ISR-PO-68-31. ISR-PO-68-31, 1968. `https://cds.cern.ch/record/296752`.

[22] ATLAS Collaboration, G. Aad et al., *Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. **C73** no. 8, (2013) 2518, `arXiv:1302.4393 [hep-ex]`.

[23] V. Cindro et al., *The ATLAS beam conditions monitor*, JINST **3** (2008) P02004.

[24] D. Caforio, *The ATLAS forward detectors - LUCID, ALFA and AFP: past, present and future*, ATL-LUM-SLIDE-2013-054 (2013). `https://cds.cern.ch/record/1514203`.

[25] The Luminosity Group, *Preliminary luminosity determination in pp collisions at* $\sqrt{s} = 8$ *TeV using the ATLAS detector in 2012*, Tech. Rep. ATL-COM-LUM-2012-013, 2012. `https://cds.cern.ch/record/1494059`.

[26] A. Sopczak, *Luminosity monitoring in ATLAS with MPX detectors*, JINST **9** no. 01, (2014) C01027, `arXiv:1312.6296 [physics.ins-det]`.

[27] D. Yu, S. Pagan Griso, and B. Heinemann, *Luminosity measurement in pp collisions at* $\sqrt{s} = 7$ *TeV using vertex counting with the ATLAS detector in 2011*, Tech. Rep. ATL-COM-LUM-2013-016, 2013. `https://cds.cern.ch/record/1559846`.

[28] D. Berge, C. Gabaldon, A. Hoecker, A. Messina, and G. Piacquadio, *Determination of the absolute luminosity and the visible cross section from LHC van der Meer scans performed in October 2010, using events with MBTS triggers and with primary vertices*, Tech. Rep. ATL-COM-LUM-2011-003, 2011. `https://cds.cern.ch/record/1326897`.

[29] ATLAS Collaboration, *Luminosity public results*. `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults`. Accessed: 2015-11-15.

[30] ATLAS Collaboration Collaboration, *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*, Tech. Rep. ATLAS-CONF-2012-042, 2012. `https://cds.cern.ch/record/1435196`.

[31] ATLAS Collaboration Collaboration, *Performance of primary vertex reconstruction in proton-proton collisions at* $\sqrt{s}$ *=7 TeV in the ATLAS experiment*, Tech. Rep. ATLAS-CONF-2010-069, 2010. `https://cds.cern.ch/record/1281344`.

[32] R. Fruhwirth, W. Waltenberger, and P. Vanlaer, *Adaptive vertex fitting*, J. Phys. **G34** (2007) N343.

[33] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852–867, `arXiv:0710.3820 [hep-ph]`.

[34] A. Buckley, D. Kar, P. Loch, and J. Monk, *Summary of ATLAS Pythia 8 tunes*, Tech. Rep. ATL-COM-PHYS-2012-738. `https://cds.cern.ch/record/1453344`.

[35] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC,*. In preparation to be submitted to Eur. Phys. J. C.

[36] H. Bartosik and G. Rumolo, *Production of single Gaussian bunches for Van der Meer scans in the LHC injector chain*, Tech. Rep. CERN-ACC-NOTE-2013-0008, 2013. `http://cds.cern.ch/record/1590405`.

[37] C. Oleari and D. Zeppenfeld, *QCD corrections to electroweak $\ell\nu_\ell jj$ and $\ell^+\ell^- jj$ production*, Phys. Rev. **D69** (2004) 093004, `arXiv:hep-ph/0310156 [hep-ph]`.

[38] K. Arnold et al., *VBFNLO: A parton level Monte Carlo for processes with electroweak bosons – Manual for version 2.5.0*, `arXiv:1107.4038 [hep-ph]`.

[39] U. Baur and D. Zeppenfeld, *Measuring three vector boson couplings in $qq \to qqW$ at the SSC*, in *Workshop on Physics at Current Accelerators and the Supercollider Argonne, Illinois*. 1993. `arXiv:hep-ph/9309227 [hep-ph]`.

[40] ATLAS Collaboration, G. Aad et al., *Measurement of the electroweak production of dijets in association with a Z-boson and distributions sensitive to vector boson fusion in proton-proton collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, JHEP **04** (2014) 031, `arXiv:1401.7610 [hep-ex]`.

[41] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, `arXiv:hep-ph/0409146 [hep-ph]`.

[42] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, JHEP **11** (2007) 070, `arXiv:0709.2092 [hep-ph]`.

[43] S. Alioli et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, `arXiv:1002.2581 [hep-ph]`.

[44] ATLAS Collaboration, G. Aad et al., *Search for invisible decays of a Higgs boson using vector-boson fusion in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **01** (2016) 172, `arXiv:1508.07869 [hep-ex]`.

[45] ATLAS Collaboration, G. Aad et al., *Search for the electroweak production of supersymmetric particles in $\sqrt{s}$=8 TeV pp collisions with the ATLAS detector*, Phys. Rev. **D93** no. 5, (2016) 052002, `arXiv:1509.07152 [hep-ex]`.

[46] C. Quigg, *LHC Physics Potential versus Energy*, `arXiv:0908.3660 [hep-ph]`.

[47] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189–285, `arXiv:0901.0002 [hep-ph]`.

[48] W. Stirling, *13/8 TeV LHC luminosity ratios*. `http://www.hep.ph.ic.ac.uk/~wstirlin/plots/lhclumi7813_2013_v1.pdf`. Accessed: 2016-03-09.

[49] U. Blumenschein et al., *Measurement of the production cross section of a Z boson in association with jets at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Tech. Rep. ATL-COM-PHYS-2016-017, CERN. `https://cds.cern.ch/record/2119923`.

[50] ATLAS Collaboration, *Luminosity public results Run 2*. `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2`. Accessed: 2015-11-15.

[51] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, `arXiv:0811.4622 [hep-ph]`.

[52] S. Catani, F. Krauss, B. R. Webber, and R. Kuhn, *QCD matrix elements + parton showers*, JHEP **11** (2001) 063, `arXiv:hep-ph/0109231 [hep-ph]`.

[53] T. Gleisberg and S. Hoeche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039, `arXiv:0808.3674 [hep-ph]`.

[54] F. Cascioli, P. Maierhofer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601, `arXiv:1111.5206 [hep-ph]`.

[55] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **03** (2008) 038, `arXiv:0709.1027 [hep-ph]`.

[56] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, *QCD matrix elements + parton showers: The NLO case*, JHEP **04** (2013) 027, `arXiv:1207.5030 [hep-ph]`.

[57] 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 **07** (2014) 079, `arXiv:1405.0301 [hep-ph]`.

[58] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852–867, `arXiv:0710.3820 [hep-ph]`.

[59] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. **B867** (2013) 244–289, `arXiv:1207.1303 [hep-ph]`.

[60] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, `arXiv:hep-ph/0603175 [hep-ph]`.

[61] P. Z. Skands, *Tuning Monte Carlo generators: the Perugia tunes*, Phys. Rev. **D82** (2010) 074018, `arXiv:1005.3457 [hep-ph]`.

[62] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. **D82** (2010) 074024, `arXiv:1007.2241 [hep-ph]`.

[63] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250–303.

[64] ATLAS Collaboration, G. Aad et al., *The ATLAS simulation infrastructure*, Eur. Phys. J. **C70** (2010) 823–874, `arXiv:1005.4568 [physics.ins-det]`.

[65] T. Hryn'ova and K. Nagano, *Trigger menu strategy for Run 2*, Tech. Rep. ATL-COM-DAQ-2014-054, 2014. `https://cds.cern.ch/record/1703730`.

[66] ATLAS Collaboration, *Muon reconstruction performance in early $\sqrt{s} = 13$ TeV data*, Tech. Rep. ATL-PHYS-PUB-2015-037, 2015. `https://cds.cern.ch/record/2047831`.

[67] ATLAS muon combined performance group, *Muon Selection Tool*. `https://twiki.cern.ch/twiki/bin/view/Atlas/MuonSelectionTool`. Accessed: 2016-02-15.

[68] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data*, Tech. Rep. ATLAS-CONF-2014-032, 2014. `https://cds.cern.ch/record/1706245`.

[69] ATLAS EGamma Combined Performance Group, *Electron and photon selection and identification for Run2*. `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/EGammaIdentificationRun2#Fiducial_region_and_Calorimeter`. Accessed: 2016-02-28.

[70] W. Lampl et al., *Calorimeter clustering algorithms: description and performance*, Tech. Rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, 2008. `https://cds.cern.ch/record/1099735`.

[71] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **04** (2008) 063, `arXiv:0802.1189 [hep-ph]`.

[72] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2014-018, 2014. `https://cds.cern.ch/record/1700870`.

[73] D0 Collaboration, V. M. Abazov et al., *Studies of W boson plus jets production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Rev. **D88** no. 9, (2013) 092001, `arXiv:1302.6508 [hep-ex]`.

[74] 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 [hep-ex]`.

[75] A. Dattagupta et al., *Observation and measurement of electroweak W + 2 jets production at high dijet mass*, Tech. Rep. ATL-COM-PHYS-2014-413, 2014. `https://cds.cern.ch/record/1699872`.

[76] R. J. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, Comput. Phys. Commun. **77** (1993) 219–228.

[77] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. **A389** (1997) 81–86.

[78] CMS Collaboration, V. Khachatryan et al., *Measurement of electroweak production of two jets in association with a Z boson in proton-proton collisions at $\sqrt{s} = 8$ TeV*, Eur. Phys. J. **C75** no. 2, (2015) 66, `arXiv:1410.3153 [hep-ex]`.

Appendix A

2012 vdM Scans

A.1 Vertex Counting External Consistency

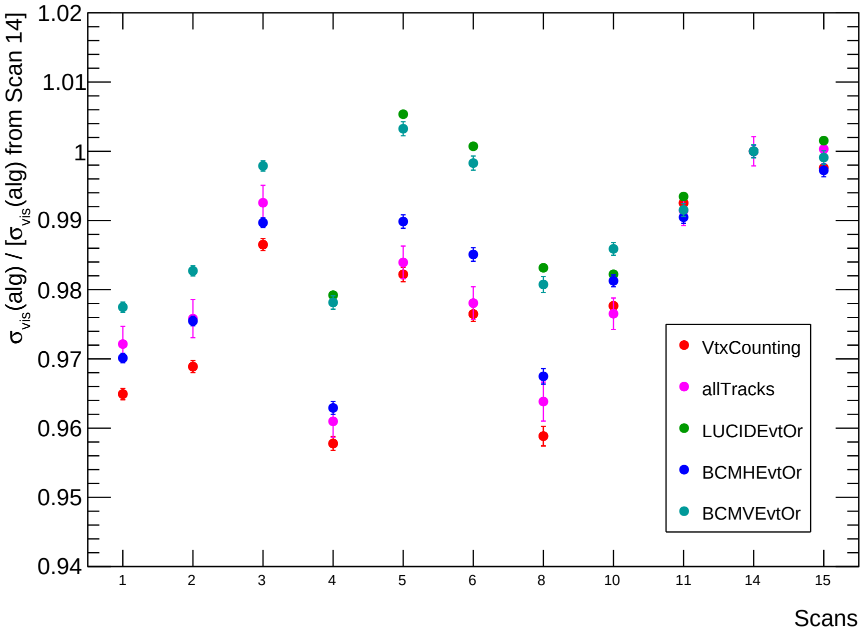


Figure A.1: σ_{vis} values for all 2012 vdM scans from different algorithms, normalised to scan 14 (vertex counting value using NTrkCut 3).

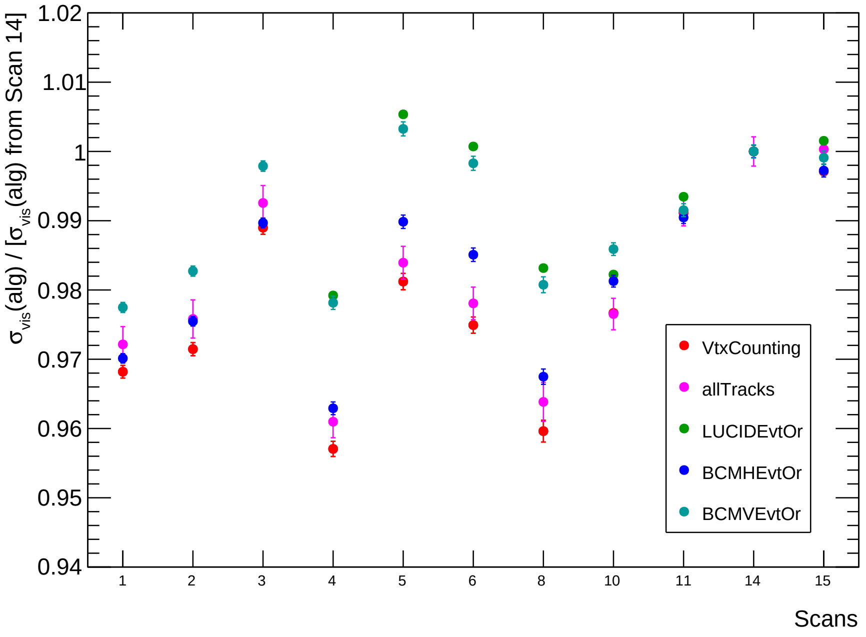
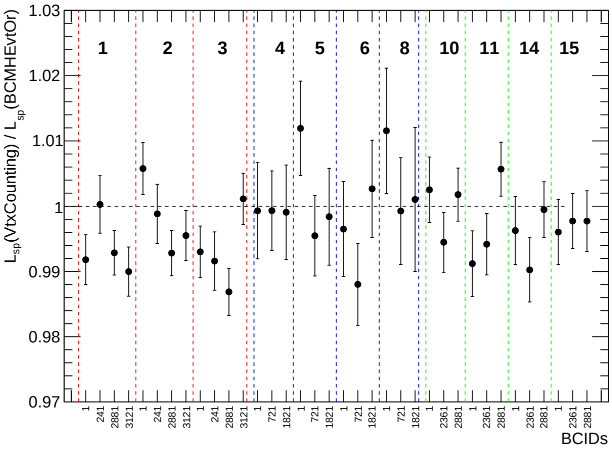
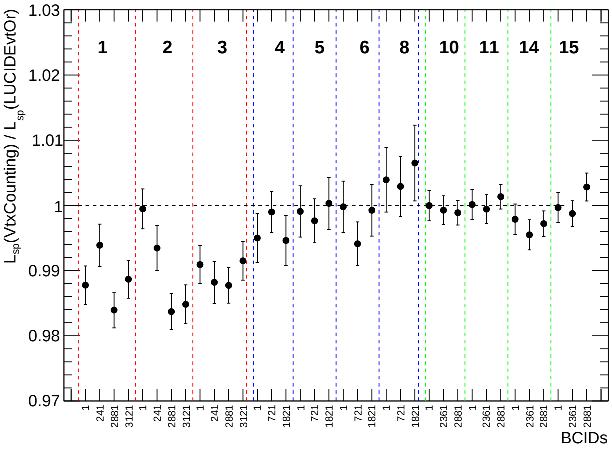


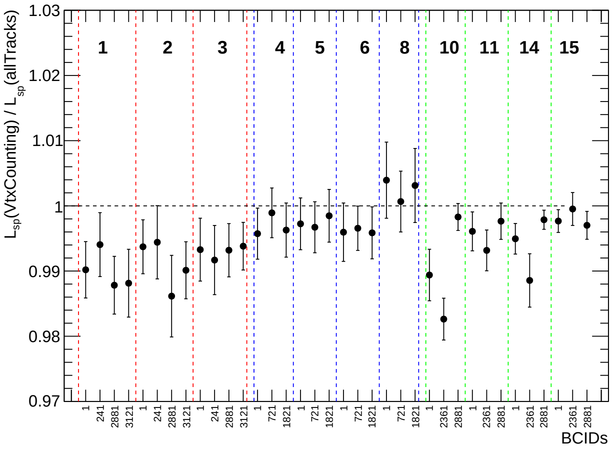
Figure A.2: σ_{vis} values for all 2012 vdM scans from different algorithms, normalised to scan 14 (vertex counting value using NTrkCut 4).



(a) Comparison to BCMHEvtOR.

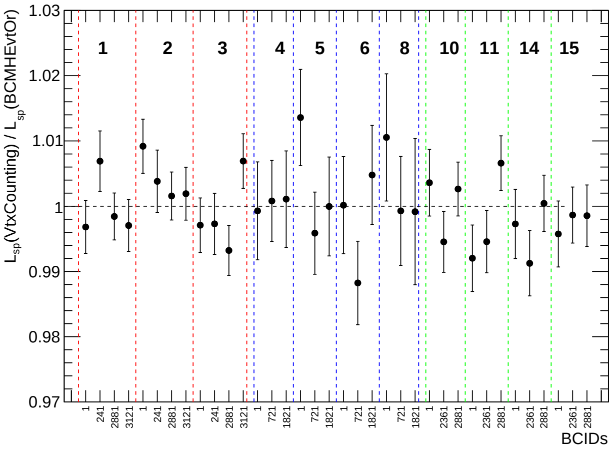


(b) Comparison to LUCIDEvtOR.

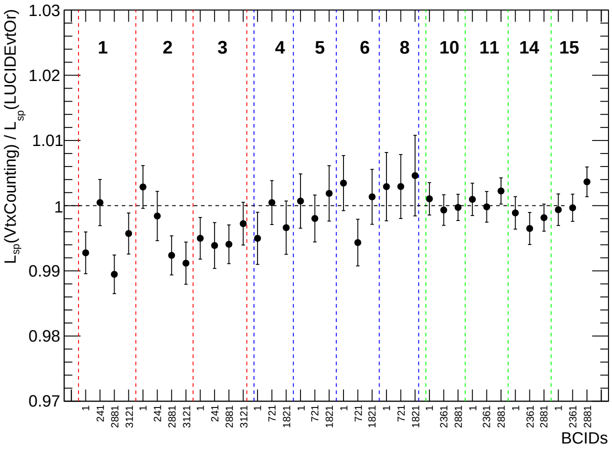


(c) Comparison to allTracks.

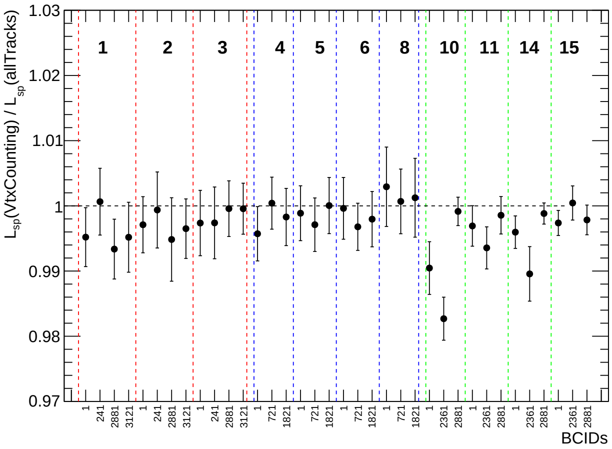
Figure A.3: Specific luminosity comparison between vertex counting (NTrkCut 3) and three selected algorithms, for all BCIDs and all scans.



(a) Comparison to BCMHEvtOR.



(b) Comparison to LUCIDEvtOR.



(c) Comparison to allTracks.

Figure A.4: Specific luminosity comparison between vertex counting (NTrkCut 4) and three selected algorithms, for all BCIDs and all scans.