Probing $b\bar{b}$ production with the ATLAS detector at the LHC

Peter Davison
University College London

Submitted to University College London in fulfilment
of the requirements for the award of the
degree of **Doctor of Philosophy**

January 19, 2016

Declaration

I, Peter Davison confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Peter Davison _____

Abstract

This thesis describes the measurement of $b\bar{b}$ production with two different techniques, both using the ATLAS detector at the LHC.

Firstly $Z + b\bar{b}$ events are measured using $\sqrt{s} = 7$ TeV data collected in 2011. The Z-boson was reconstructed using pairs of leptons ($Z \to e^+e^-$ or $Z \to \mu^+\mu^-$), with the b-jets used to tag the presence of B-hadrons. The cross-section was measured as a function of four variables and compared to theoretical predictions. NLO predictions from MCFM and aMC@NLO best reproduce the data with a discrepancy seen at low angular separation of the b-jets.

Secondly inclusive $b\bar{b}$ production is probed using the decay of B-hadrons to muons using the $\sqrt{s} = 8$ TeV data collected by ATLAS in 2012. Differential cross-sections were measured as a function of six variables and compared to predictions from PYTHIA8 and HERWIG++. In general the shape of data was well reproduced by the HERWIG++ prediction.

Acknowledgements

Firstly I would like to thank my supervisor Gavin Hesketh. You have been a great help over the last four years, and a genuine pleasure to work with.

I would like to thank my friends for the many distractions over the course of the PhD, especially to the *Mudchuters* for a barnstorming first year.

Fay, you have been a constant over the last four years, which have been infinitely more enjoyable with you in my life.

Finally a thank you for my family, and especially Mum and Dad. Literally would not have been able to do the PhD without your unquestioning support and encouragement.

Author's Contributions

ATLAS is a collaborative experiment, with contributions from thousands of individuals. It is common practice for measurements performed within the collaboration to be the work of several people. The two analyses discussed in this thesis (Chapter 4 and Chapter 5) were performed as part of a team effort. The majority of work presented in this thesis was produced by the author and where it has not been it should be clear from references given in the text. All the figures not produced by the author for this thesis are referenced in the figure description. A brief list of the main contributions from the author and others is given below:

Chapter 4: The author contributed to the majority of the $Z + b\bar{b}$ analysis including the fitting procedure, unfolding and the systematic analysis. Major contributions from others include the derivation of the b-template re-weighting, QCD background determination, template shape systematics and the generation of theoretical predictions. The published analysis can be found in [1].

Chapter 5: The author contributed to all parts of this analysis with the exception of the production of the data files used in the measurement, the production of the trigger and muon reconstruction maps and the generation of the theoretical predictions.

Contents

List of figures

List of tables

Chapter 1.

Introduction

Quantum chromodynamics (QCD), introduced formally in Section 2.3, is the theory of the strong force which is felt by coloured objects such as quarks and gluons. QCD has been probed extensively over the years [2–4]. However QCD production of heavy flavour (i.e. b-quarks) is less well tested at hadron colliders due to the smaller cross-sections for b-jet production[1], which is about 2-3% of the inclusive jet production cross-section [5], and theoretical predictions suffer from larger uncertainties. The Large Hadron Collider (LHC) has opened up a whole new era of particle physics measurements with its higher centre of mass energies and large luminosity. It allows for more precise investigations into b-quark production at hadron colliders, which can help tune the calculations of heavy flavour production.

From a theoretical perspective b-type quarks are interesting due to their large mass. This leads to different approximations in calculations which can result in contrasting behavior compared to light quarks. The modeling of b-quark density in the proton parton density functions is one example with different theoretical predictions using different approximations [6]. Gluon splitting ($g \to b\bar{b}$) to heavy flavour quarks also uses different splitting functions to light quarks, so it's important to have data measurements to constrain different theoretical modeling.

The increased collision energy at the LHC also gives the possibility to see new physics. Lots of beyond the Standard Model (BSM) physics models have enhanced couplings to the third generation quarks [7]. So it is important that Standard Model b-quark production is well understood to decrease uncertainties on backgrounds to

[1] A *b-jet* is a cluster of particles originating from a b-quark

these models. Finally an increasingly common analysis technique at the LHC is to measure resonances in the high transverse momentum or *boosted* regime [8]. Boosted topologies often lead to close-by jet production or even jet merging. If the jets are expected to be b-jets then close-by $b\bar{b}$ production needs to be well understood as a background to these boosted physics processes.

This thesis describes two measurements involving heavy flavour production. Initially an introduction to the relevant theory is presented in Chapter 2 followed by a description of the experimental apparatus in Chapter 3. The first measurement of B-hadrons produced in association with a Z-boson is detailed in Chapter 4. This is followed by a second measurement using muons to detect the heavy flavour production in Chapter 5.

Chapter 2.

Theoretical Framework

"I have always had more dread of a pen, a bottle of ink, and a sheet of paper than of a sword or pistol"

— Alexandre Dumas

2.1. Standard Model

The theory describing fundamental particles and their interactions is called the Standard Model. The Standard Model was developed during the 1970s and describes the most basic constituents of the universe and the forces mediated between these particles through the framework of a relativistic quantum field theory. The theory has been tested rigorously and has been very successful, for example predicting the existence of many particles, the W-boson [9, 10] and Z-boson [11], top-quark [12] and more recently the Higgs boson [13, 14], before being detected in experiments. However the theory is not complete and experiments in the neutrino sector have proved the existence of neutrino oscillations [15, 16] (changing of neutrino flavour) which indicates that neutrinos are massive, something not predicted by the Standard Model. The most striking irregularity with nature though is the Standard Model only being a theory of three out of the four fundamental forces, gravity has so far not been incorporated into the theory.

Fundamentally the Standard Model is a theory of particles, it describes two distinct sets; fermions and bosons which are differentiated by particle spin. Fermions are characterized by being spin-$\frac{1}{2}$. There is a natural grouping within the fermions into leptons and quarks. There are six leptons: electron, muon, tau and a neutrino for each of these flavours. The quarks mirror the leptons in that there are also six types: up, down, charm, strange, top and bottom. Bosons are integer spin particles, which act as force carriers for the fundamental forces described by the Standard Model, these include the strong-force, electromagnetic-force and weak-force. The Standard Model also predicts that for every particle there is an anti particle which has the same mass but opposite quantum numbers. Experimentally this is seen as opposite charge.

The electromagnetic and weak forces are actually two components of a single force, the electroweak force [17], that through spontaneous symmetry breaking by the Brout-Englert-Higgs mechanism [18, 19] become the electromagnetic and weak forces which have the force carriers γ and ($W^{\pm}$, Z^0) respectively. The strong force is the force responsible for keeping nuclei together at the centre of an atom. The strong force is governed by the laws quantum chromodynamics (QCD) and mediated by massless gluons, g. The final piece of the Standard Model is the Higgs boson. The Higgs particle interacts with all massive particles (which includes self-interaction), with the mass of the particle proportional to the strength of the interaction. The fermions predicted by the Standard Model and the bosons they interact with are listed in Table 2.1.

In group theoretic terms the Standard Model is a quantum field theory based on a gauge invariant $SU(3)_C \times SU(2)_L \times U(1)_Y$, with $SU(3)_C$ corresponding to the strong interaction and $SU(2)_L \times U(1)_Y$ being the group for the electroweak interaction.

2.2. Electroweak Theory

From Table 2.1 we see that quarks can interact weakly via a W or Z-boson. An example of Z-boson production is shown in Figure 2.1 which shows a Drell-Yan [20] like process. The propagator also includes contributions from a virtual photon, γ^*, due to the unification of electroweak theory (for the rest of this thesis the mixing of

Fermion	Charge, q	Mass [GeV]	Boson Interactions
u	$\frac{2}{3}$	2×10^{-3}	$\gamma, g, W/Z, H$
d	$-\frac{1}{3}$	5×10^{-3}	$\gamma, g, W/Z, H$
c	$\frac{2}{3}$	1.3	$\gamma, g, W/Z, H$
s	$-\frac{1}{3}$	0.1	$\gamma, g, W/Z, H$
t	$\frac{2}{3}$	173	$\gamma, g, W/Z, H$
b	$-\frac{1}{3}$	4.2	$\gamma, g, W/Z, H$
e	-1	5.11×10^{-4}	$\gamma, W/Z, H$
μ	-1	0.106	$\gamma, W/Z, H$
τ	-1	1.78	$\gamma, W/Z, H$
ν_e	0	$<2 \times 10^{-9}$	W/Z
ν_μ	0	$<1.7 \times 10^{-4}$	W/Z
ν_τ	0	$<1.55 \times 10^{-2}$	W/Z

Table 2.1.: Standard Model particle properties.

Z and virtual photons is assumed). The branching fractions of Z-boson decays have been measured extensively [21] with the branching ratio to a same pair of oppositely charge leptons of $3.363 \pm 0.004\%$.

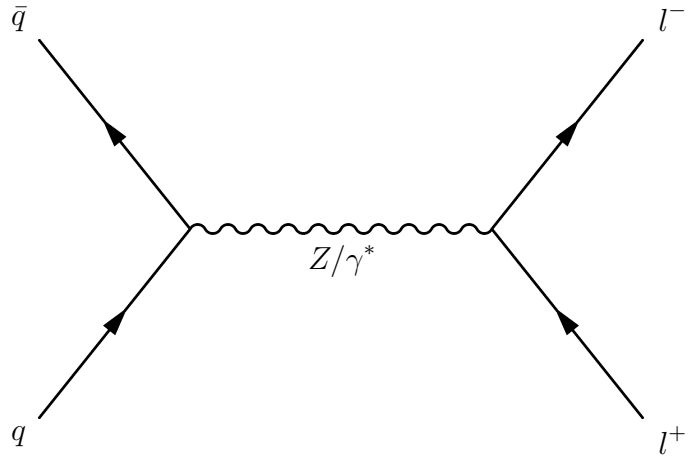


Figure 2.1.: Drell-Yan production of a lepton pair via Z/γ^* production.

2.3. Quantum Chromodynamics

Quantum chromodynamics (QCD) is a quantum field theory represented by the non-abelian $SU(3)$ group that describes the interactions of the strong force. QCD describes the force felt by coloured objects, where colour comes in three variates, red, green and blue. Gauge invariance of the theory is apparent as invariance under colour changes. The force carriers of QCD, the gluons, also carry colour charge. As coloured objects feel the strong force this results in gluon self-interaction. The fact that gluons carry colour themselves leads to two distinct properties of QCD, namely colour confinement and asymptotic freedom.

- *Colour Confinement:* Yet to be proved mathematically, colour confinement states that a coloured object cannot be seen in isolation, and observed particles are grouped into colourless objects. Consider an isolated coloured object; as this particle is coloured it can interact via QCD. In analogy to QED, where quantum fluctuations can produce a charge shielding effect, QCD quantum fluctuations produce gluons which contain charge themselves. This results in an increase in colour density. There is no mechanism to stop this so it would result in an infinite colour density. This problem is solved by coloured objects grouping into colourless hadrons.

- *Asymptotic Freedom:* As the energy scale increases (or probing distance decreases) the strong coupling becomes weaker, this is known as the *running of the coupling* of the strong coupling constant, α_s. This means for high enough energies ($\sim$ 1GeV) the coupling becomes weak enough for perturbation theory to be valid. This is in keeping with the colour confinement analogy, where at increasing energy or equivalently decreasing probing scales less of the *sea* quarks and gluons, produced via quantum fluctuations around coloured objects, are discernible, decreasing the observed colour density. This results in less colour charge and a smaller interaction coupling.

2.4. Perturbation Theory

The probability for a particle interaction is proportional to the square of a transitional probability from an initial state of particles to a final state. This transitional

probability is often called a *matrix element*. Matrix elements are formed from a perturbation expansion in increasing orders of the coupling constant, α, relating to the type of interaction. For interactions of photons and matter (quantum electrodynamics, or QED) α is small (approximately $\frac{1}{137}$) which means higher orders of α contribute more and more negligibly to the cross section. For QCD however, the coupling constant α_s can only be considered small above some cut-off energy and perturbative techniques can only be considered above this scale. For theoretical calculations this often means that the hard-interaction between quarks in a proton-proton scattering experiment can be calculated perturbatively, but the formation of hadrons, which occurs at a much lower energy scale, cannot.

When calculating the probability of a particle interaction the term in the perturbation with the lowest order of α represents the leading order term (LO). Terms with higher α orders can then be constructed by including extra interaction vertices. A common prescription for calculating matrix elements and visualising processes was developed by Richard Feynman called the Feynman rules and Feynman diagrams respectively. Feynman diagrams can be thought of as terms in the perturbative series, with increasing orders of α represented by more complicated diagrams. In Figure 2.2 quark pair production is shown at LO and also an example of a next-to-leading order (NLO) correction is given, which is at the second lowest order of the coupling constant. A full cross-section calculation has to sum over all orders of α which is equivalent to including all diagrams with increasing complexity.

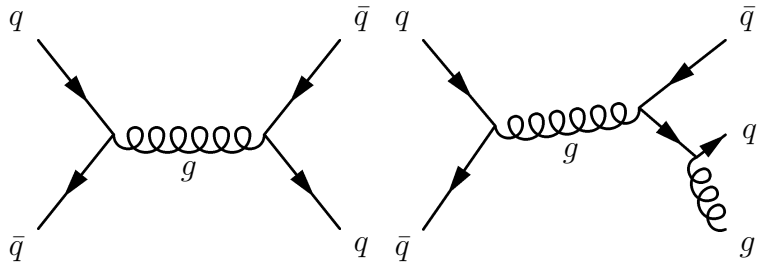


Figure 2.2.: Quark pair production Feynman diagrams for LO (left) and NLO (right).

For QCD radiative emissions, such as the NLO correction in Figure 2.2, you find in the limit of the emission being collinear or equivalently having zero momentum these perturbative terms can diverge: this corresponds to *infrared divergences*. At higher orders you can also have virtual loop corrections. As the particles in these loops are virtual there is no limit on their momentum. Infrared divergences in loop diagrams

also give divergences when considered individually. Fortunately order-by-order these complimentary divergences cancel resulting in finite predictions.

2.5. Factorisation

Perturbation theory is only applicable above a certain energy scale. Full calculations of particle interactions need to include non-perturbative physics. A full cross-section calculation can be factorized into perturbative and non-perturbative parts as shown for a proton-proton interaction to a final state X in Equation 2.1.

$$\sigma_{pp\to X} = \sum_{ij} \int dx_1 dx_2 f_{i,p}(x_1, Q^2) f_{j,p}(x_2, Q^2) \times \sigma_{ij\to X}(x_i, x_j, s, Q^2). \tag{2.1}$$

Here the summation is over the partons i and j that can produce the final state X, $\sigma_{ij\to X}$ is the perturbative cross section for partons i and j with momentum fraction of the proton x and is measured at a centre of mass energy $\sqrt{s}$, Q^2 is a momentum scale typically chosen as the energy scale of the interaction and $f_{x,p}(x_i, Q^2)$ are the parton density functions (PDFs) which describe the parton distributions within the proton. PDFs are not perturbatively calculable and have been measured in data at electron-proton deep-inelastic scattering experiments [22, 23]. Despite not being able to analytically calculate PDFs, they are universal and the evolution to different Q^2 scales is calculable. The PDFs scale with Q^2 as the small scale structure of the proton can be probed with higher energies, and vice versa. The Q^2 dependence is described by the DGLAP equations [24–26] which relates the parton density at one scale, via splitting functions, to the partons at a different scale. Figure 2.3 shows example proton PDFs at two varying momentum scales.

To separate the perturbative and non-perturbative schemes in calculations, an energy scale is chosen so that any emission below that scale is included in the PDF and an emission above that scale is counted as part of the hard interaction, this is known as the *factorisation scale*, μ_F. The renormalisation scale, μ_R, is another scale set for QCD calculations and is the scale at which α_s is calculated. A standard practice is to set $\mu_F = \mu_R$. The renormalisation and factorisation scales are artificial scales that are introduced to the calculation. Different choice of scales can result in different predictions at a fixed order, so it is important to estimate the impact of

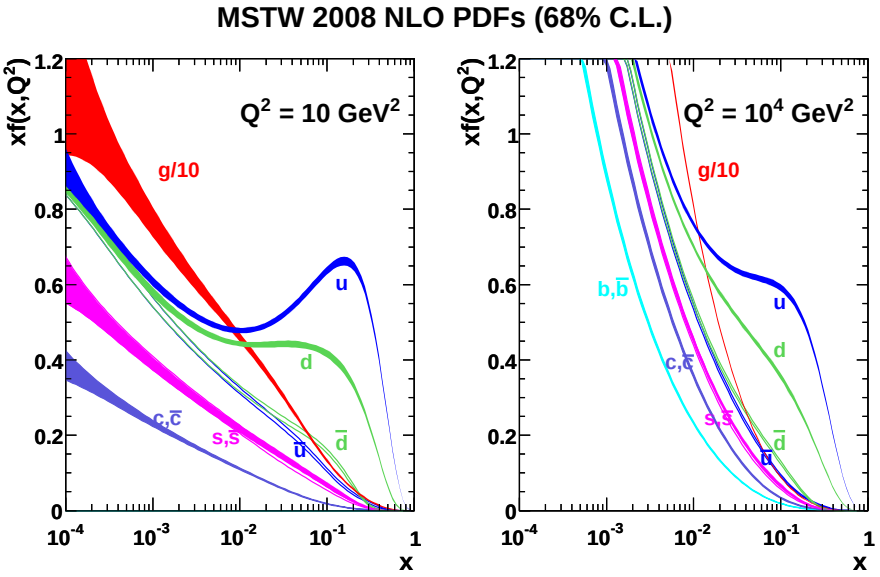


Figure 2.3.: Example MSTW2008 PDFs for the proton at two different energy scales [27].

the choice of scales when calculating cross sections for particle interactions. This is typically done by manually altering the scales in the calculation and taking a theoretical uncertainty on the scale choice from the deviation.

$\sigma_{ij\to X}$ is the perturbative expansion described in Section 2.4. As you go to higher orders of α_s you can include more radiative emissions or internal particle loops. Higher order calculations are increasingly complex and difficult to calculate so matrix elements are typically calculated at a fixed order (often LO or NLO, with some processes having calculations at even higher orders). The radiation of a soft or collinear parton can contribute significantly to a cross section which can result in significant deviations form the LO prediction when considering higher orders.

As the mass of the b-quark is larger than the typical QCD scale (1 GeV) it introduces a new scale in calculations. This means there are two methods for treating heavy flavour quarks in PDFs [28]. Firstly, fixed flavour number schemes (FFNS), where only quark densities for a fixed number of flavours are considered part of the proton PDF. The *4FNS* is an example of a fixed flavour number scheme where only the lightest four quarks are considered. With the b-quarks not present in the initial state they must be produced via gluon splitting in the matrix element. The other approach is the variable flavour number scheme (VFNS). The *5FNS* is an example of a variable number scheme, where the number of PDF flavours depends on the scale,

where a b-parton density is included for scales above the b-quark mass. Typically all quarks are treated as massless in variable number scheme approaches.

2.6. Monte Carlo Techniques

The Monte Carlo method [29] describes a computational technique using a random sampling of probability distribution functions to calculate the probability of particle interactions. In particle physics the theoretical predictions are often fed into a Monte Carlo program that then can produce a stochastic set of simulated data which can be compared to the real data gathered by a detector. The comparison can yield information about how well the theory is predicting real data. Generally Monte Carlo predictions can be split into two parts:

- *Event Generation*: Involves calculating the hard-process and includes initial and final state radiation. Parton radiation is modeled by the use of parton showers and when partons fall below a certain energy they are hadronised using non-perturbative methods. Non-stable particles[1] are then decayed leaving a set of stable particles with defined four-vectors. Underlying event contributions and pile-up events can also be added to the simulation at this stage.

- *Detector Simulation*: The particles are then propagated through a simulated detector (for ATLAS this is done using GEANT4 [30][2]) and the detector response is saved. The output from detector simulation is in the same format as you would have for the real data so that physics object reconstruction (see Section 3.8) can be performed identically for both simulation and data.

Event generation can vary depending on which Monte Carlo program you use. Different Monte Carlo generators differ in the modeling of underlying event, parton shower evolution and hadronisation. Hadronisation is applied at the end of parton showering when the partons fall below the QCD perturbative energy scale to take the partonic final state to a hadronic final state.

[1] A stable particle has a lifetime, $t > 30$ ps.

[2] Full event reconstruction using GEANT4 uses large amounts of computing power. An alternative computationally quicker approach is to use AtlFast-II [31] which is a faster simulation method that only uses full simulation for certain parts of the detector systems, reducing the simulation time by about an order of magnitude. In general the analysis presented in Chapter 4 uses AtlFast-II simulated samples and the Monte Carlo used in Chapter 5 has full detector simulation.

The parton shower models the radiation from a coloured parton produced in the hard scatter. As the gluon is itself coloured it can also radiate gluons and the result can be a cascade of partons. The shower calculates the probability of a parton to radiate (or conversely not to). Parton showers are evolved from the hard scatter down to the non-perturbative scale, at which point hadronisation takes over. Evolution of parton shower is ordered in a phase space variable, typically this is the p_{T} or angle of emissions, where the p_{T} of emissions is evolved from hardest downwards and the angle is ordered from wide to collinear angular emissions. The parton shower effectively sums the leading terms for higher order corrections to the matrix element, which results in the summation of collinear and soft emissions. This results in the parton shower appropriately describing low angle emissions but is less reliable for wide-angle topologies.

Two models of hadronisation are commonly used, the Lund string model and cluster models. The Lund string model [32] is best described by considering a quark anti-quark pair moving in opposite directions. Due to colour confinement the space between the quark anti-quark pair will be filled with a colour flux tube with radius of the order 1 fm. The tube is parameterised as constant along its length and results in a linear potential, so as the quarks move further apart the potential rises. As the potential energy increases quantum fluctuations can convert the potential energy into a quark anti-quark pair breaking the string and creating two colourless quark anti-quark pairs, resulting in a colourless hadronic system. Heavy quarks are highly suppressed in these quantum fluctuations resulting in a production ratio of different quark types: $u : d : s : c \approx 1 : 1 : 0.3 : 10^{-11}$ [33] resulting in negligible production of charm and beauty hadrons during hadronisation.

The cluster model is another approach to hadronisation based on the property of preconfinement of parton showers [34]. As colour lines are propagated through the parton shower colour singlet clusters of partons end up close in phase space. The clusters are then treated as excited hadrons and are decayed to lowest state hadrons to complete the hadronisation.

2.7. Jet Formation

As has already been mentioned single coloured objects are not observed in isolation. Instead partons produced in particle interactions are observed in a detector as a stream of hadrons due to parton showering followed by hadronisation. The sum of the four-vectors of the stream of particles should be equal to the four-momentum of the original parton due to momentum conservation. So combining the stream of particles into *jet* to some extent reflects the kinematics of the original parton. For comparison between Monte Carlo and data a jet reconstruction method needs to be flexible enough to use various inputs: energy deposits in the detector, individual particles or partons for Monte Carlo. The main class of jet algorithms used are iterative combination algorithms. Jets mentioned in this thesis are reconstructed with the anti-k_t algorithm [35]. The anti-k_t algorithm successively combines objects based on a distance criterion, d_{ij}:

$$d_{ij} = min\big(k_{T,i}^{-2}, k_{T,j}^{-2}\big) \cdot \frac{(\Delta R_{ij})^2}{R^2}, \tag{2.2}$$

where k_T is the transverse momentum, ΔR is the angular separation between objects i and j[3] and R is a radial parameter, where $R = 0.4$ used for jets in this thesis. A second quantity:

$$d_{iB} = k_{T,i}^{-2}, \tag{2.3}$$

where d_{iB} is used as a reference for the distance criterion between particle i and the beam, B. The algorithm works by calculating the d_{ij} and d_{iB} for all inputs and then if d_{iB} is the smallest then object i is a jet, and is removed from the collection of objects, however if one of the d_{ij} is the smallest then the two objects are merged (their four-vectors combined), this is repeated until all objects are either classed as a jet or merged into a jet. For a jet to be theoretically stable it needs to be both infra-red and collinear safe. For the former, the jet defined by the algorithm should be unaffected by the emission of very soft radiation, for the latter the jet should be unaffected by a parton splitting into two with small angular separation. These are both true for the anti-k_t algorithm. For detectors at particle physics experiments there is an inbuilt resolution and physically an emission of a soft parton or close

[3] $\Delta R = \sqrt{(\Delta y)^2 + (\Delta \phi)^2}$

by splitting would not affect the jet measured, so infra-red and collinear safety is mimicked experimentally. The advantage of the anti-k_t algorithm can be seen in Figure 2.4 where it results in circular jets in (ϕ, y) space. Other jet algorithms have more irregular shapes due to how softer radiation is included in the jets.

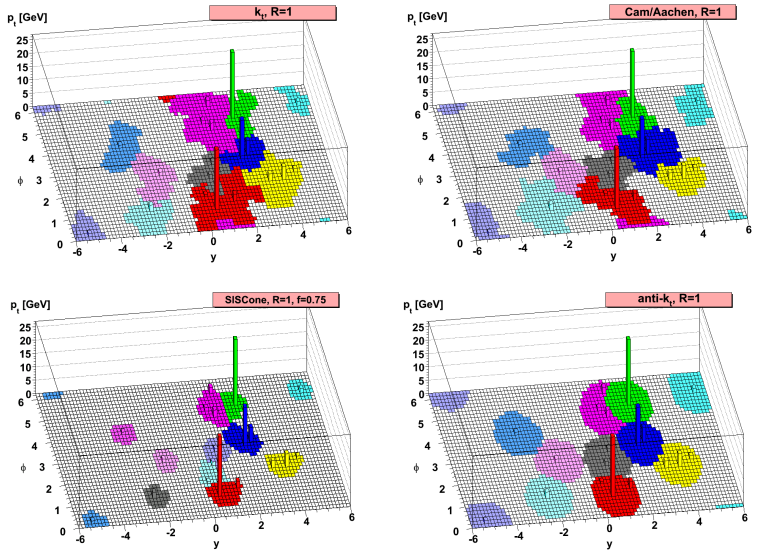


Figure 2.4.: Comparison of jet algorithm shapes in (ϕ, y) [35].

2.8. Truth Lepton Definition

Leptons produce final state radiation (FSR) after production in a hard process, this results in several ways to define a lepton at a particle level in Monte Carlo. A Born-level lepton is defined as the lepton straight from the hard interaction, i.e. before any QED radiation effects. A Bare-lepton is the lepton after all final state radiation. Finally a Dressed-lepton is defined as the lepton which is summed with any QED FSR in a cone, of typical size $\Delta R = 0.1$, to form a single physics object. Historically leptons were considered at the Born-level however due to interference effects between photon radiation in an initial state and a final state in a process producing a electroweak vector boson decaying leptonically, [36] care has to be taken on how a final state lepton is defined. Truth leptons are now defined at the Dressed-level to remove any ambiguity. Electron dressing mimics what is measured in an experimental detector as any final state photon radiation would be measured

in the calorimeter along with the electron. It is less suitable for muons however as the photons will not be measured as part of the muon object at a detector. Dressing muons is done to define a consistent approach between different Monte Carlos, which might differ in how QED radiation is modeled for a muon.

Chapter 3.

The Large Hadron Collider and the ATLAS Detector

"Later For a date than the Hadron Collider, and costs more"
— Daniel Dumile

3.1. Overview

CERN is one of the largest scientific institutions in the world. Based on the Swiss-French border close to Geneva it was founded in 1953 and initially consisted of 25 member states. CERN has been the setting for lots of scientific experiments since its formation. Although having contributions in several scientific fields it is most famous for being a particle physics laboratory. Over the years a wide range of particles have been discovered at a multitude of different experiments, such as the W and Z at the UA1 and UA2 detectors in the 1980s [9, 10] and more recently the Higgs particle [13, 14] using the ATLAS [37] and CMS [38] experiments. The Large Hadron Collider (LHC) [39] is the latest collider experiment located at CERN. The LHC is primarily a proton-proton (p-p) collider with a design centre of mass energy of 14 TeV and luminosity of $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The LHC is also capable of running proton-heavy ion or purely heavy ion collisions, although only p-p runs are considered in this thesis.

The LHC consists of a circular tunnel approximately 27 kilometers in circumference. The tunnel is at a depth of around 100 m underground so that the experiments at the LHC are shielded from cosmic radiation incident on the Earth's surface. Two highly collimated beams consisting of proton bunches, each containing billions of protons, travel in opposing directions around the tunnel. Before reaching the LHC the proton bunches have been through a series of accelerators to increase the beam energy to 450 GeV which is the injection energy for the LHC accelerator. The whole accelerator and pre-accelerator layout can be seen in Figure 3.1. In several positions around the LHC the beams are crossed so that protons in each beam can interact. Detectors are placed at these crossing points that record the debris from these collisions. At the LHC the main detectors are ATLAS, CMS, LHCb [40] and ALICE [41]. ATLAS (A large Toroidal LHC ApparatuS), pictured in Figure 3.2, is the physically largest detector at the LHC. It is a general purpose detector with the capability to measure a wide range of high energy physics processes. It is shaped to completely envelope the crossing point of the proton beams so that most of the spray of particles from an interaction of two colliding protons will propagate into the detector mass. The detector is cylindrically shaped and composed of several systems each designed to help with a certain part of particle identification and/or measurement of particle direction and energy.

The ATLAS detector uses a right-handed co-ordinate system with the origin at the centre of the detector with the x-axis pointed towards the centre of the LHC ring, the y-axis pointing directly up towards the Earth's surface and the z-axis pointing along the beam axis. This is also expressed in cylindrical co-ordinates, with the azimuthal angle, ϕ, defined as the angle around the beam axis starting at the x-axis and the polar angle, θ, the angle from the z-axis. Commonly used variables include the transverse momentum p_{T} defined as the momentum in the transverse $x-y$ plane and rapidity defined as:

$$y = \frac{1}{2}\left(\ln\left(\frac{E+p_z}{E-p_z}\right)\right). \tag{3.1}$$

These quantities are useful as the p_{T} of particles, and the difference in rapidity between particles is invariant under a Lorentz boost in the z-direction. For massless

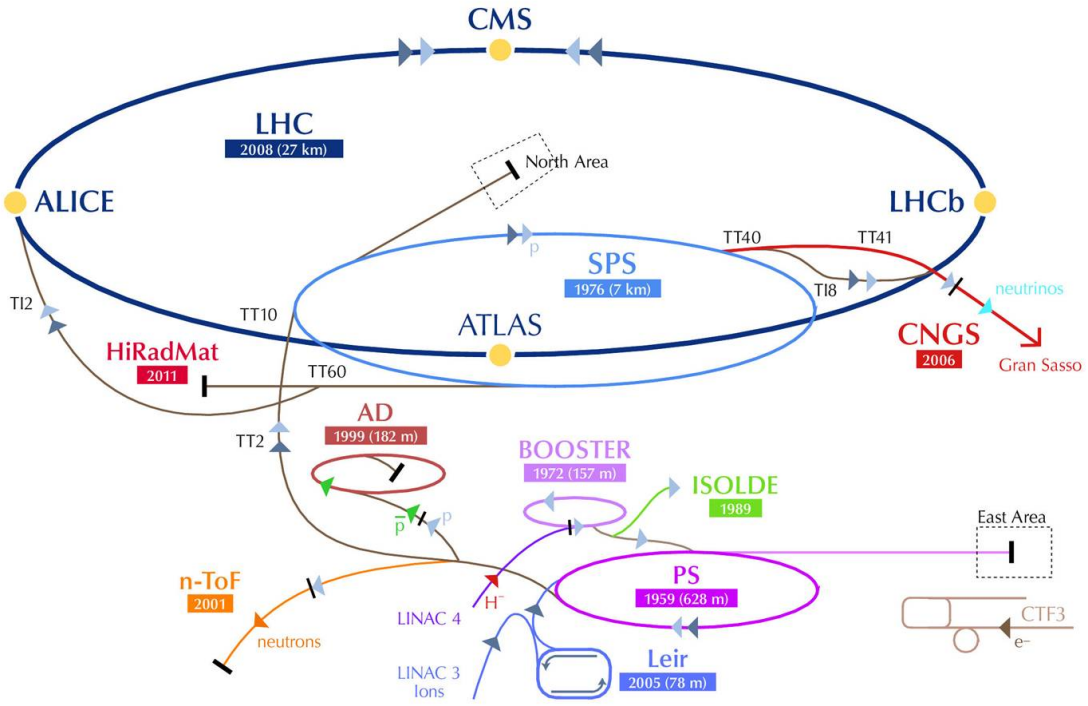


Figure 3.1.: The accelerator complex located at CERN [42].

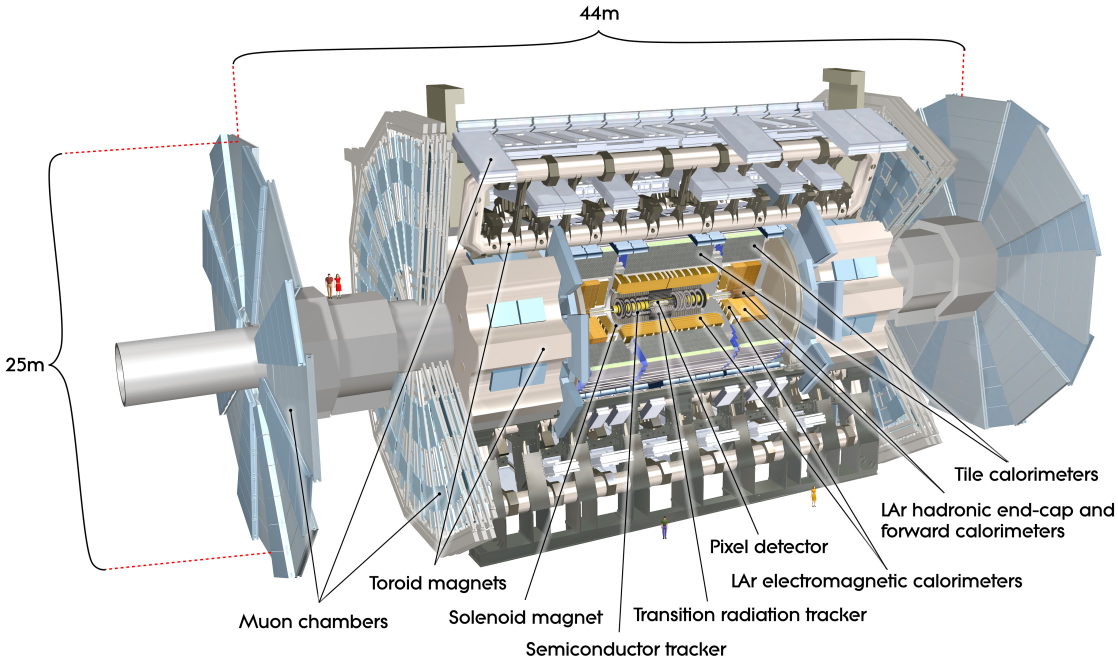


Figure 3.2.: A computer generated image of the ATLAS detector [43].

objects rapidity becomes equal to pseudo-rapidity, η:

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right). \tag{3.2}$$

The different detector systems are concentrically positioned around the proton beam axis to ensure complete coverage in ϕ around the beams. This is then supported by end cap detectors which extend the coverage at high rapidities. The innermost system is the Inner Detector (ID), after that comes the calorimetry systems and finally the outermost part of the detector is the muon system (MS). Interleaved with these sub-systems are a series of magnets which supply approximately uniform magnetic field separately for the ID and MS. The ID is contained within a thin superconducting solenoid magnet system, whereas the MS uses superconducting toroids.

3.2. Inner Detector

The inner detector is the first part of the detector that particles travel through after an interaction at the beam line. It starts at a radial position of approximately 5 cm with an outer radius at a distance of 1.15 m from the beam axis. The role of the ID is to measure the trajectory of charged particles. A charged particle will leave a signal (hit) in different sub-detectors within the ID. These hits are then combined into tracks which represent the trajectory of the particle. The ID system is immersed in a 2 T axial magnetic field, which bends charged particles and allows momentum measurements, as the degree of bending is inversely proportional to the momentum. The direction of bending also gives the charge of the track. The ID itself consists of three separate systems in order of increasing radial distance from the beam line: the pixel detector, the semiconductor tracker (SCT) and the transition radiation tracker (TRT) as shown in a schematic of the ID in Figure 3.3. Table 3.1 documents the main features of each of the sub-detectors.

The pixel detector consists of three barrel layers and three endcap disks on both sides of the barrel section. The barrel and end-cap layers consist of a total of 1744 identical modules each containing $\sim 50,000$ pixels. The pixel detector is designed to measure accurately charged particle trajectory and has highest granularity of the

Subsystem	$\|\eta\|$ range	resolution ($\perp$, z) [μm]	Typical hits	Radial position [mm]
Pixel Detector	<2.5	(10, 115)	3	50.5, 88.5, 120.5
SCT	<2.5	(17,580)	8	299, 371, 443, 514
TRT	<2.5	(130)	30	554 $\rightarrow$ 1082

Table 3.1.: Various properties of the ID components including the coverage, resolution in perpendicular and transverse directions, typical number of track hits and the radial distance to the start of the barrel region layers of each subsystem.

ID systems to achieve this. The vertex resolution when combining several tracks to an origin is approximately 30 μm and 50 μm for the transverse and longitudinal directions respectively. The high intrinsic resolution of the pixel detector allows the reconstruction of secondary vertices from, for example B or D-hadrons. The next ID system a particle would pass through is the SCT. The SCT is composed of eight silicon microstrip layers in the barrel and nine end-cap disks either side of the detector. The SCT is used to measure momentum, track impact parameters and vertex positions. The outer sub-system is the TRT. The TRT is a straw drift tube detector and complements the precision of the pixel and SCT detectors by increasing the number of hit measurements in track formation. It is made up of 300,000 cylindrical drift tubes measuring 4 mm in diameter and a length of 144 cm in the barrel and 37 cm in the end-cap wheels. The barrel tubes are parallel to the beam line so provide only accurate position measurements in the transverse plane.

3.3. Calorimeters

Radially outside the ID are the calorimeter systems shown in Figure 3.4. Housed outside the solenoid magnet, the calorimeter consists of two sub-systems, the electromagnetic and hadronic calorimeters. Both are sampling calorimeters that have alternating layers of absorbing and active materials. The absorbing layers are made of dense materials which the passing particles interact with, resulting in a cascade of particles. The active material then detects the cascade decays and infers the energy of the incident particles. By using multiple layers the calorimeters are designed to absorb all of the energy of the incoming particles.

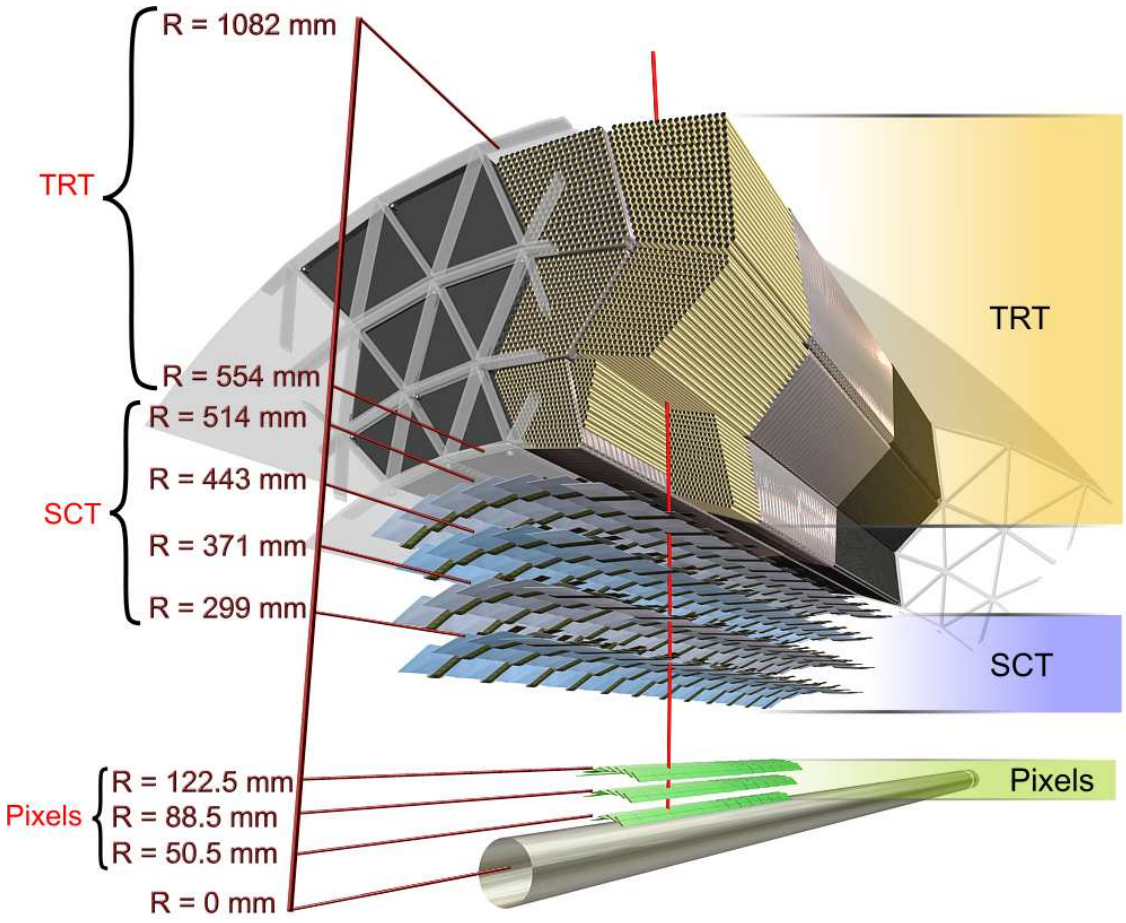


Figure 3.3.: Diagrammatic of the barrel sections of the ID systems [44].

The electromagnetic calorimeter (ECAL) is a high angular granularity calorimeter, with liquid argon (LAr) used as the active material and lead for the absorption layers. The calorimeter has an accordion-like layout so there is complete coverage in ϕ without any gaps. The ECAL is split into a barrel and two end-caps. The barrel region extends to $|\eta| < 1.475$, with the end-caps extending the η coverage $1.375 < |\eta| < 3.2$. Over the high granularity section of the system ($|\eta| < 2.5$) the ECAL is split into three layers, the first layer has the finest granularity but makes up only 20% of the radiation depth and is called the strip layer. The middle layer makes up the bulk of the system in terms of radiation lengths with the outer layer used to help with discriminating electromagnetic and hadronic showers. The primary goal of the ECAL is to accurately determine the energy information of electrons and photons.

Surrounding the outside of the ECAL is the hadronic calorimeter (HCAL). The HCAL is split into three sections with differing η coverage; $0 < |\eta| < 1$ for the barrel, $0.8 < |\eta| < 1.7$ for the extended barrel and $1.5 < |\eta| < 3.2$ for the end-caps. The barrel and extended-barrel consist of absorbing layers of steel alternating with layers of the scintillating plastic tiles which act as the active material. The end-caps use LAr as the active material with Iron used as an absorber. The HCAL is positioned outside the ECAL due to the stronger penetration of hadronic particles.

In the forward regions, at large values of $|\eta|$, are positioned the forward calorimeters (FCAL). The FCAL covers the range $3.1 < |\eta| < 4.9$ and has both electromagnetic and hadronic calorimeters. Both use LAr as the active material with the inner electromagnetic part using copper as an absorber, whereas the outer layers for hadronic calorimetry are using tungsten.

The calorimeters were designed to have large coverage in solid angle as well as the ability to absorb all of the energy of electrons, photons and hadrons, as this is needed to detect missing energy in the transverse plane, $E_{\mathrm{T}}^{\mathrm{miss}}$. $E_{\mathrm{T}}^{\mathrm{miss}}$ is indicative of particles not interacting electromagnetically or strongly, for instance a neutrino, so this is an important feature of the detector.

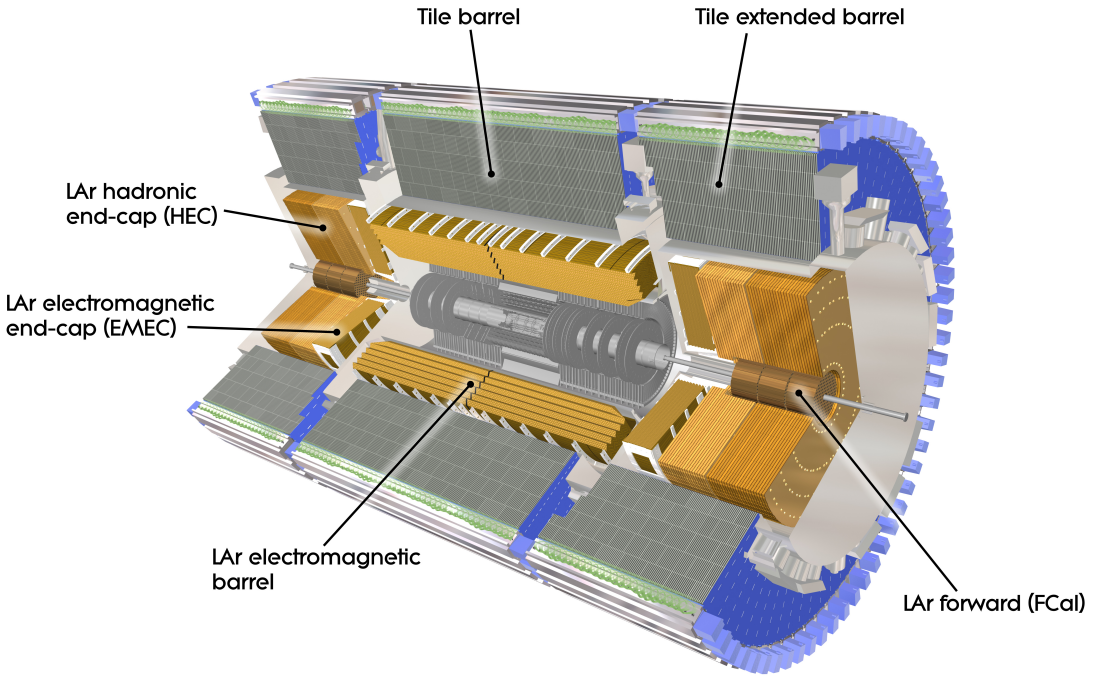


Figure 3.4.: Diagram of calorimeter systems [45].

3.4. Muon System

The furthest detector component from the beam axis is the muon system, responsible for most of the detector volume. Muons are the only directly detectable particle from the Standard Model that are expected to consistently reach the MS. The MS is embedded in a toroidal magnetic field of 0.5 T, angled to be perpendicular to the traversing muon momentum, which allows for muon position, momentum and charge measurements. The MS comprises of four parts, two used for precision tracking and two used for triggering.

The two precision tracking components cover the combined range $|\eta| < 2.7$. Most of the precision tracking is done using Monitored Drift Tubes (MDTs) for $|\eta| < 2.7$ with Cathode Strip Chambers (CSCs) used in the first layer of the high multiplicity region $2.0 < |\eta| < 2.7$. Tracks in the MS are matched to tracks in the ID for the formation of reconstructed muons.

The two trigger sub systems cover the range $|\eta| < 2.4$. In the central $|\eta| < 1.05$ region resistive plate chambers (RPCs) are used, whereas in the higher multiplicity

region $1.05 < |\eta| < 2.4$ thin gap chambers (TGCs) are used for their capability at higher rates.

3.5. Trigger System

In 2011 the LHC was delivering proton bunches with a spacing of 50 ns (twice the design specification of 25 ns) or a crossing rate of 20 MHz. For 2012 running the instantaneous luminosity was increased by a factor of two whilst keeping the same bunch spacing. Proton bunches contain more than 10^{11} protons. On average in 2012 there were approximately 20 p-p interactions per bunch crossing (see Section 3.7), which results in p-p interaction rates of 400 MHz (twice that of 2011). This interaction rate is far too fast to record all interactions with today's computing and digital storage abilities. Most of the events will not result in interesting, usually high p_{T}, event topologies. A trigger system is implemented to bring the event rate down and look for only relevant event signatures.

The ATLAS trigger has a three tier system, Level 1 (L1), Level 2 (L2) and the event filter (EF), listed in order of decreasing output rate. L2 and EF are referred collectively as the high level trigger (HLT). L1 is a hardware based trigger, the target output event rate is around 75 MHz. L1 has only about 2 μs latency so has to reduce the detector granularity in order to process the large number of incoming events. L1 uses the calorimeter and MS systems in regions of interest (RoIs) in (η, ϕ)-space. If the energy measurements of a physics object in a RoI is larger than a set of thresholds the event is flagged by the L1 system. L1 is able to search for potential muons, electrons, photons, jets and tau leptons. For muons the L1 trigger looks for a series of hits in the MS consistent with a muon track coming from the luminous region of the detector which is passing one of the pre-determined energy thresholds. For electrons the L1 trigger looks for deposits in the ECAL in RoIs. If an event is flagged by L1 the RoIs are sent as seeds to the next trigger system L2.

The first of the software based HLT systems, L2 takes the RoI information from L1 and uses the full detector granularity in these regions to better resolve object information. The calorimeter system and MS information in these regions is now also supported by other detector information such as tracking. The output rate for L2 is about 4 kHz.

Finally events passing the L2 trigger will then seed the final part of a trigger chain, the EF. The input rate to the EF allows for full detector granularity and the use of more sophisticated algorithms similar to that used in offline data analysis. The output events selected by the EF are written to disk with an output rate of below 400 Hz.

Offline analyses use trigger chains, which is a sequence of L1 and HLT triggers. Unfortunately, due to the high level data output and the need to keep the event output rate of each trigger level down, not all trigger chains can be written to disk every time they are satisfied. Some triggers are pre-scaled which results in only a fraction of the triggered events for that trigger chain being saved. For example, for a trigger chain with a prescale factor of 10 only 1 in 10 of the triggered events is saved, picked by a random sampler.

3.6. Luminosity

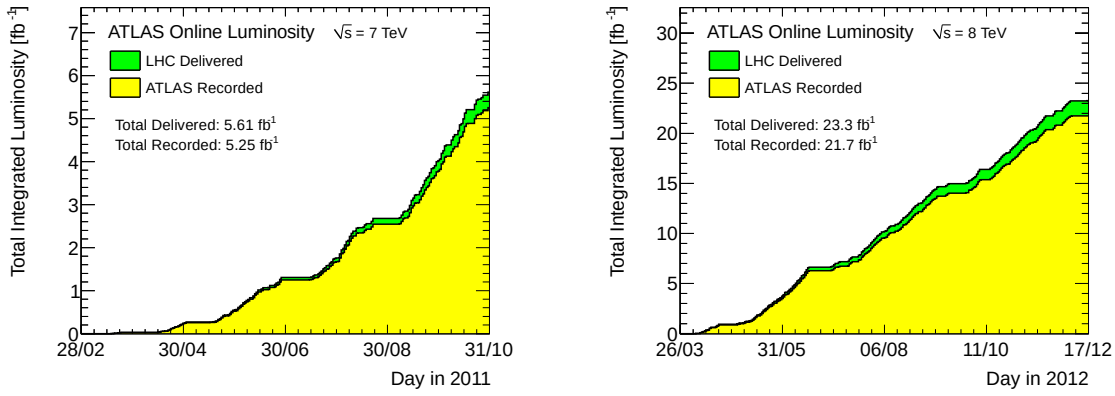
The amount of data recorded by the ATLAS detector is given in terms of integrated luminosity, L. The number of events produced for a given process, p-$p \rightarrow X$, is proportional to L:

$$N_{events} = \sigma_{pp\rightarrow X} L = \sigma_{pp\rightarrow X} \int \mathcal{L} dt, \tag{3.3}$$

where $\sigma_{pp\rightarrow X}$ is the cross-section for that process and $\mathcal{L}$ is the instantaneous luminosity. The total integrated luminosity which was delivered by the LHC and recorded by the ATLAS detector for the 2011 and 2012 data taking periods are shown in Figure 3.5. The measured luminosity is recorded with specialised sets of detectors placed up and down stream from the main ATLAS detector volume [46].

3.7. Pile-up And Underlying Event

Due to the large number of protons in each bunch it is possible to have more than one proton-proton interaction within a single bunch crossing. These multiple interactions are termed *in-time* pile-up. Due to the low cross-section of interactions resulting in high p_{T} particles most of the in-time pile-up interactions are soft in nature and are



(a) (b)

Figure 3.5.: Integrated luminosity recorded per day for a) 2011 data taking and b) 2012 data taking [47].

termed *minimum bias* interactions. Due to the extra soft collisions pile-up events tend to produce a small increase in the energy deposited in the calorimeter systems due to the spray of soft particles, thus increasing the energy measurements of other physics objects. Another form of pile-up comes from the finite operation cycle of the detector hardware. This *out-of-time* pile-up is the result of p-p interactions in neighbouring bunch crossings, and hardware components that have an operational cycle longer than the bunch spacing. The average number of interactions per bunch crossing, μ, is a measure of pile-up and can be seen for both 2011 and 2012 data in Figure 3.6.

The *underlying event* is defined as the interactions between remnants in a p-p collision that are not involved in the hard scatter. This usually results in the production of extra soft particles. However when there are multiple hard interactions this is termed *multi-parton interactions* (MPI).

3.8. Object Reconstruction

The ATLAS detector sub-systems are used to measure properties of several types of particles. Muons are detected by tracks in the MS, which are matched to tracks in the ID. Electrons are found by a track in the ID with a matching energy deposit in the ECAL. Tau leptons are not observed directly due to their short lifetime,

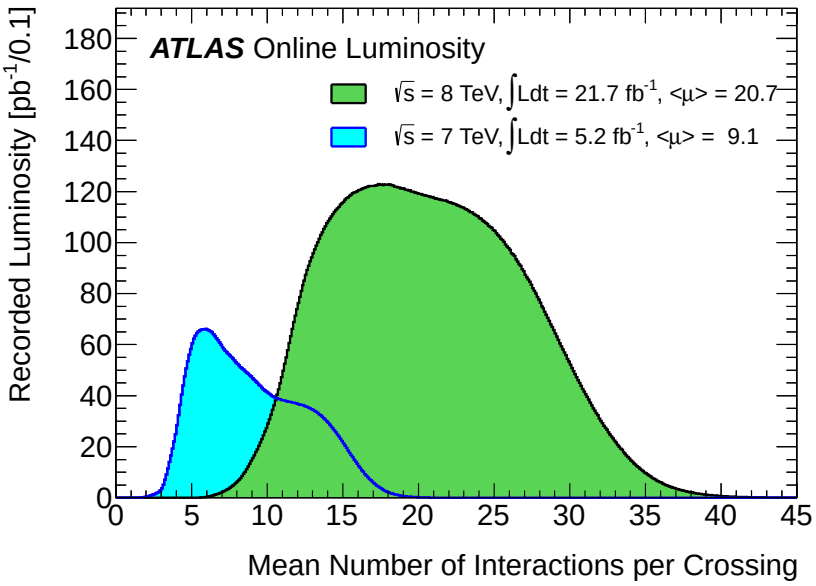


Figure 3.6.: Luminosity weighted distributions of the mean average number of bunch crossings for the 2011 (blue) and 2012 (green) data taking [47].

(2.9×10^{-13} s), but will be detected by their decay products. Photons leave only a deposit in the ECAL with no corresponding track in the ID. Jets are constructed from energy deposits in the calorimeters. Finally neutrinos are not observed directly but are inferred using energy conservation in the transverse plane of the detector. The analyses in this thesis use muons, electrons and jets as the signatures for signal events so these objects are described in more detail in the following sections.

3.8.1. Muon Reconstruction

Muons can be detected in both the ID and MS systems. This allows for different muon reconstruction algorithms which use different parts of the detector. For both analyses in this thesis only *combined* muons are used [37]. Combined muons use independent tracks formed in both the ID and MS combined into a single muon object. The track segments are combined only if they are consistent with each other in direction and p_{T}. The requirement of tracks in both sub-detectors gives a higher purity of real muons compared to muon algorithms using the MS information alone and results in an improved momentum resolution.

3.8.2. Electron Reconstruction

Electron identification is performed using information from the ECAL and ID. Jets can leave tracks in the ID and deposit energy in the ECAL and due to the much larger hadronic production cross-section at the LHC it is important for electron reconstruction algorithms to reject jets faking electrons. Electron reconstruction uses sliding window algorithms that find the maximum energy deposit in the ECAL within a (η, ϕ) window by scanning (η, ϕ) space. A track is then matched to the group of ECAL cells, or clusters. In the case of multiple matched tracks the track with the highest number of hits in the innermost ID layers is selected. For 2011 running a series of working points defined at certain electron selection efficiency and jet rejection power were defined [48]. These working points use shower shapes, penetration depth as well as tightening up the match quality between the track and calorimeter deposits and track quality for ID tracks.

3.8.3. Jet Reconstruction

As mentioned in Chapter 2 we do not see bare partons in the detector, after the process of showering and hadronisation the original partons are observed as a spray of collimated hadrons with a direction and total energy indicative of the original parton. These jets of particles are detected as energy deposits in the ATLAS calorimeter systems. The energy deposits are formed into 3-dimensional energy cluster groupings, *topo-clusters*, which form the physics objects acting as inputs into the jet algorithms described in Section 2.7. Topo-clusters are seeded by a calorimeter cell with an energy deposit four times larger than the expected noise in the cell. All neighboring cells with energy deposits of at least twice the background noise are then added to the topo-cluster. This is repeated until no neighboring cells satisfy the energy threshold. The total energy is formed from the energy of all the topo-cluster cells. The direction of the topo-cluster is an energy-weighted average of all the constituent cells and the mass is assumed to be zero. This information is then used for the objects entering jet reconstruction algorithms. The energy as measured from the topo-clusters is referred to as jet energy at the electromagnetic (EM) scale due to the calibration of the calorimeter response to EM showers. Topo-clusters that have been calibrated before passing to the jet finding algorithms result in jets at the local cell weighting (LCW) scale.

After jet formation the jet energy and positional information needs to be adjusted for detector effects. This calibration takes jets from the EM (LCW) scale to the final jet energy scale (JES), or for the two initial jet calibration states: EM+JES (LCW+JES) scale. The calorimeter has a lower response for hadrons than EM particles, so jets need to be corrected from the EM scale to the hadronic energy scale. Energy deposited in passive detector material, energy leakage from the calorimeters and energy not reconstructed in the jet are other examples of sources of missing jet energy that are calibrated for. The JES correction consists of several terms to reproduce the jet energy at the hadronic level. These are:

- *Pile-up correction*: Corrects for the average soft particle flux from pile-up events, parameterised using number of reconstructed vertices and the average number of interactions. Energy is removed from jet to account for the pile-up energy.
- *Direction correction*: Initial topo-clusters where formed under the assumption the jet originated from the origin of the detector. This correction changes the direction of the jet to point towards the primary vertex (PV). The PV is defined as the reconstructed vertex in an event with the highest scalar sum p_{T} of constituent tracks.
- *Energy correction*: The energy correction term corrects the energy of the jets due to the lower response of hadrons in the calorimeters.
- *Data in-situ correction*: Only for data events, the in-situ correction is a data-driven correction looking at the p_{T} imbalance of a jet and a reference object e.g. Z+jet events.

3.8.4. *b*-Jet Reconstruction

Jets originating from B-hadrons have several features that allow their identification from other jets. B-hadrons are relatively long lived, with a lifetime of $\approx 1.5 \times 10^{-12} s$ [49]. With sufficient p_{T} the B-hadrons will decay at a position displaced from the primary vertex (order of a few mm). The decay products of the B-hadron can then be reconstructed as tracks in the inner detector and the properties of these tracks can be used to assess the likelihood the jet originates from a B-hadron, or is a b-jet.

The algorithms used to determine if a jet is a b-jet are called b-taggers. There are several b-taggers used at ATLAS. The two main groups of b-taggers are those using displaced ID tracks, lifetime taggers, and those that look for soft muons from semi-leptonic B-hadron decays associated to the jet. The b-taggers used in Chapter 4 are of the first type and are explained below [50]:

- *IP3D* : Uses track parameters at the distance of closest approach to the primary vertex, or *impact parameters*, see Figure 3.7. IP3D uses the impact parameter for tracks in both the longitudinal and transverse planes. The tracks considered by the IP3D algorithm first have to pass a series of quality cuts to remove fake tracks. The impact parameter significance is used which is the impact parameter divided by its measurement error, so that badly measured tracks are weighted down accordingly.

- *SV1*: SV1 is a b-tagger that looks to reconstruct the secondary vertex of a long lived particle from tracks associated to the jet, see Figure 3.7. An iterative procedure is used, starting with 2-track vertices that are combined into a single vertex. This single vertex is then stripped of tracks that fit poorly to the vertex until finally the fit quality of the vertex passes a requirement. Cuts are applied on the mass of the secondary vertices to remove vertices from long lived particles, such a the K_L and material interactions, such as photon conversion. The SV1 algorithm then combines several properties of the secondary vertex and associated jet to determine the likelihood the jet is a b-jet.

- *JetFitter*: The JetFitter algorithm looks for a cascade decay of a B-hadron to a long lived D-hadron (a hadron containing a charm quark) by looking for two secondary vertices along the direction of the jet. Jet flavour is then determined from a likelihood fit based on the track and vertex properties.

- *JetFitterCombNN*: JetFitterCombNN is neural network based tagger that uses as inputs the output from other b-taggers, namely IP3D and JetFitter. A per-jet output of the jet flavour hypothesis for b-jets, c-jets and *light*-jets (p_b, p_c and p_l respectively) is given by the JetFitterCombNN algorithm. Higher values of p_x indicate a stronger likelihood of the jet being of flavour x.

- *MV1*: The MV1 algorithm is another neural network b-tagger taking JetFitterCombNN, SV1 and IP3D as inputs.

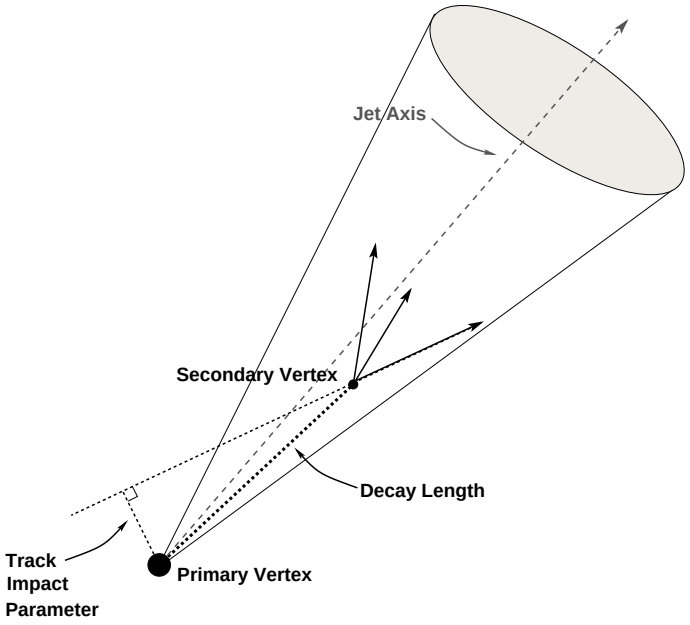


Figure 3.7.: Diagram of typical b-jet topology including the variables used by b-tagging algorithms [51].

b-tagging performance is judged on two criteria; b-jet tagging efficiency and light-jet rejection which is evaluated on simulated Monte Carlo samples. Figure 3.8 shows the performance of the MV1 algorithm as a function of these two criteria. The widely used b-tagging algorithms at ATLAS have *working points* defined with these two performance parameters tuned so a cut on the b-tagger output has a set efficiency.

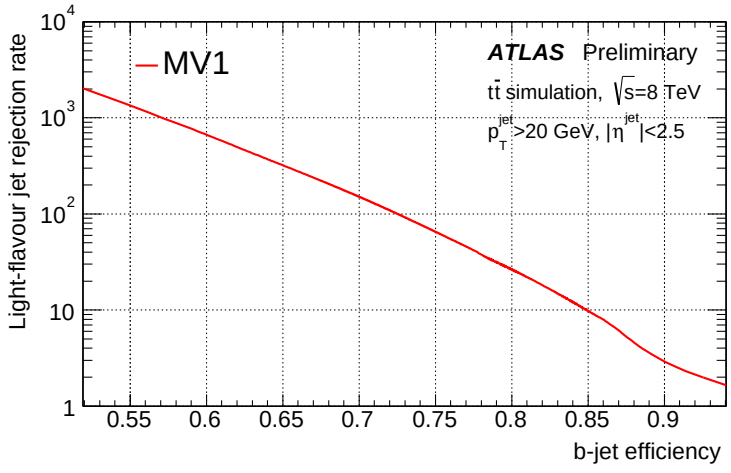


Figure 3.8.: The MV1 algorithm b-tag efficiency versus the light-jet rejection [52].

3.8.5. Missing Energy

Particles that do not interact with the detector material are reconstructed by the missing energy in the transverse plane to the beam, $E_{\mathrm{T}}^{\mathrm{miss}}$. As the proton beams are highly collimated the protons contain little momentum in the plane transverse to the beam. Momentum must be conserved after an interaction so it is expected that the outgoing particles from a hard-scatter will have zero total momentum in this transverse plane. A deviation from this is indicative of a particle that has not interacted with the detector such as a neutrino. $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated as the magnitude of the vector sum of all reconstructed physics objects in an event, and is commonly used to detect the presence of these invisible particles.

Chapter 4.

$Z + b\bar{b}$ Production

"We shall not cease from exploration, and the end of all our exploring will be to arrive where we started and know the place for the first time."

— T.S. Eliot

4.1. Introduction

The LHC is a proton-proton collider which produces copious amounts of jets. This results in a very challenging environment to make measurements of fully hadronic decay products. Physics signatures with jet-only final states, whether light quark, gluon initiated or heavy flavour jets, are swamped by backgrounds. A leptonically decaying Z-boson provides a clean signature in the detector which is easily triggered and reconstructed. The leptons are produced on average with high p_{T} and can be triggered on with high efficiency. Backgrounds to Z production can also be reduced significantly by requiring di-lepton pairs have invariant masses around the Z mass. The analysis presented in this Chapter measures events where b-jets are produced in association with a leptonically decaying Z-boson, $Z + b\bar{b}$ production. The work presented here was part of a larger measurement including the more inclusive $Z + b$ production.

Figure 4.1 shows example LO Feynman diagrams for $Z + b$ and $Z + b\bar{b}$ processes, where there is a distinction depending on the number of b-jets in the final state. Figure 4.1 (a) is an example of $Z + b$ production at LO which has been produced in the 5FNS approach where the initial b-quark is taken from the b-parton density in the PDF. This is not included in the 4FNS approach. This means there is potential sensitivity to the b-parton density in the proton by measuring $Z + b$ final states. The other diagrams seen in Figure 4.1 (b) and (c) are examples of LO $Z + b\bar{b}$ production. These occur in both the 4FNS and 5FNS approaches, which means that for $Z + b\bar{b}$ production there is little sensitivity to different b-parton density schemes.

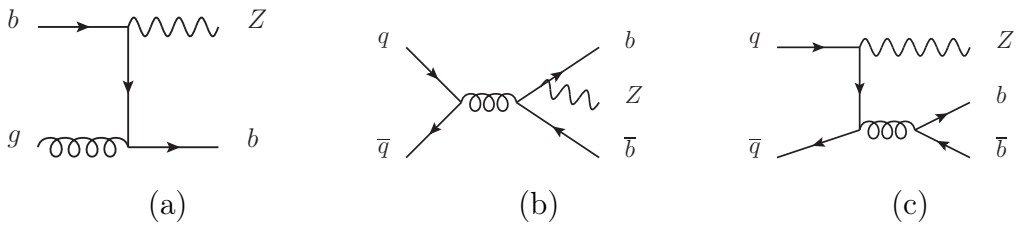


Figure 4.1.: Example Feynman diagrams for b-quark production in association with a Z-boson.

Events with a Z-boson and two b-jets are also a background to other physics processes. An important example is Higgs production in association with a Z-boson, with the subsequent $H \to b\bar{b}$ decay [53]. $Z + b\bar{b}$ is an irreducible background to this process, with the same final state particles, and cannot be constrained from data. As the Higgs is expected to decay to $b\bar{b}$-pairs about 50% of the time [54], it is important to have an accurate Monte Carlo description of $Z + b\bar{b}$ production to reduce background uncertainties and increase sensitivity to measuring the Higgs coupling in this channel. For the ZH, $H \to b\bar{b}$ analysis performed on data from 2011 [55] the background from Z+jets (of which $Z + b\bar{b}$ is the main contribution) was expected to account for over 90% of the data in some measured kinematic bins. The analysis uses the shape of the $m_{b\bar{b}}$ distribution to look for $H \to b\bar{b}$ decays so it is important that the Monte Carlo used to describe the $m_{b\bar{b}}$ shape for $Z + b\bar{b}$ events is accurate so a small Higgs signal (1% in some observable bins) is measurable. The ZH analysis also relies on the Monte Carlo description of the relative fraction of Z+b-jets that populate the background Z+jets sample. Measuring the $Z + b$ and $Z + b\bar{b}$ cross section can thus confirm or improve the Monte Carlo description of the rate of heavy flavour jets produced in association of a Z-boson.

The production of a Z-boson and b-jets has been measured previously at the Tevatron [56, 57] and the LHC [58, 59]. The Tevatron measurements were done in a smaller kinematic region and calculated differential cross-sections for $Z + b$-jets as a ratio to either inclusive Z or Z+jet production with a generally poor overall description by the available Monte Carlo at the time. The previous $Z + b$-jet ATLAS analysis [58] measured the inclusive $Z + b$-jet cross-section using the small dataset from 2010, with the CMS analysis [59] measuring the cross-section for exclusive $Z + b$ and $Z + b\bar{b}$ production. The analysis presented in this chapter significantly extends the scope of these previous measurements by investigating differential cross-sections for $Z + b\bar{b}$ production in a large fiducial region.

The delivered luminosity during the 2011 running at the LHC allowed for the first time a measurement of differential $Z + b\bar{b}$ cross-sections. The measurement is performed for both $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$ decays and requires at least two jets originating from B-hadrons as well as the di-lepton pair. Differential cross-sections are measured as a function of four observables listed in Table 4.1 and compared to several theoretical predictions.

Variable	Definition	Range
$p_{\mathrm{T}}(Z)$	Z boson transverse momentum	0–250 GeV
$\lvert y(Z)\rvert$	Z boson absolute rapidity	0.0–2.5
$m(b, \bar{b})$	$b\bar{b}$ invariant mass	20–350 GeV
$\Delta R(b, \bar{b})$	$b\bar{b}$ angular separation	0.4–5.0

Table 4.1.: Definitions of variables for which differential production cross-sections are measured and the ranges over which those measurements are performed.

4.2. Data and Monte Carlo Selection

4.2.1. Dataset

Data collected by the ATLAS detector between 13th March and 30th October 2011 at a centre of mass energy $\sqrt{s} = 7$ TeV are used for this analysis. Only data which are taken in stable beam conditions and have passed quality requirements which

satisfies the requirement of the *GoodRunsList* (GRL) are considered[1]. The analysis used only un-prescaled triggers which differ for the electron and muon channel, with a di-electron trigger used for electrons and a single muon trigger used for muons. Within each lepton channel the triggers changed during the year as the instantaneous luminosity was increased and correspondingly the trigger rate went up. This meant the trigger requirements were tightened to keep the output trigger rate the same. Table 4.2 details the triggers used and the recorded luminosity. After triggering both the $Z \to \mu^+\mu^-$ and $Z \to e^+e^-$ channels have a dataset totaling 4.5 fb^{-1}of integrated luminosity.

Decay channel	Trigger	Description	luminosity $[fb^{-1}]$
$Z \to e^+e^-$	`EF_2e12_medium`	Two HLT electrons of threshold energy 12 GeV seeded from two L1 7 GeV electrons.	1.64
$Z \to e^+e^-$	`EF_2e12T_medium`	L1 electron threshold increased from 7 to 10 GeV.	0.57
$Z \to e^+e^-$	`EF_2e12Tvh_medium`	Extra quality cut on leakage into HCAL included for HLT electrons.	2.37
$Z \to \mu^+\mu^-$	`EF_mu18_MG`	Single HLT muon of threshold energy 18 GeV seeded from 10 GeV L1 muon.	1.42
$Z \to \mu^+\mu^-$	`EF_mu18_MG_medium`	L1 muon threshold increased to 11 GeV.	3.16

Table 4.2.: Triggers used and integrated luminosities collected during 2011 data taking periods.

To increase signal purity the data to be analysed has to pass a series of event level cuts. The event must contain at least one primary vertex with at least three distinct

[1]The GRL is a list of data running periods when the detector was fully operational.

tracks. Events should have exactly two oppositely charged leptons of the same flavour (either muon or electron) which match to the trigger objects that fired the trigger. The leptons are further required to pass track cuts on the longitudinal and transverse impact parameters, z_0 and d_0 respectively. The muons have an additional requirement of being isolated from other event activity, where all of the tracks in a cone of $\Delta R < 0.2$ around the muon are required to have less than 10% of the muon p_{T}. The di-lepton pair must have an invariant mass within the Z-boson mass window, $76 < m_{l^+l^-} < 106$ GeV. Events must also have a $E_{\mathrm{T}}^{\mathrm{miss}} < 70$ GeV to remove backgrounds from events containing t-quarks with subsequent $t \to W^{\pm}(\to l^{\pm}\nu_l) + b$ decays. Finally the event must have at least two b-tagged jets which have been b-tagged using the MV1 algorithm at the 75% b-tagging efficiency point. Jets are reconstructed using the anti-k_t algorithm and must have $p_{\mathrm{T}} > 20$ GeV and $\eta < 2.4$. To suppress jets from pile-up, a cut on the *jet vertex fraction* (JVF) is applied, JVF>0.75, which requires 75% of the scalar p_{T} sum of tracks associated to the jet to come from the PV. The full event selection is summed up in Table 4.3.

4.2.2. Monte Carlo Samples

After the event level cuts the data will still contain several sources of events, both signal and background. Monte Carlo samples were used to model these.

Signal Monte Carlo

ALPGEN version 2.13 [60] was used to generate Z+jet samples. ALPGEN is a LO tree-level matrix element generator which produces up to five final state partons. The event generation used the CTEQ6L1 [61] PDF set with the 4FNS approach and used parameter tune AUET2-CTEQ6L1 [62]. The ALPGEN matrix elements were combined with HERWIG version 6.520 [63] for parton showering and hadronisation and JIMMY version 4.31 [64] for underlying event and MPI. A dedicated $Z + b\bar{b}$ sample was also produced using ALPGEN +HERWIG +JIMMY, which used the same PDF set and parameter tune as the Z+jet sample.

It is possible to get a $Z + b\bar{b}$ event from both the Z+jet sample (via $g \to b\bar{b}$ splitting in the parton shower) and from the dedicated $Z + b\bar{b}$ sample via matrix

Primary Vertex	≥ 3 tracks
Electron selection criteria	$\lvert\eta_e\rvert < 2.47$
	$p_T > 20$ GeV
	Crack region $1.37 < \lvert\eta\rvert < 1.52$ excluded
	e track $z_0 < 1\,\text{mm}$ from PV
	$\lvert d_0/\sigma(d_0)\rvert < 10$
$Z \to ee$ selection	Two oppositely charged electrons
	Veto on events with further leptons
	$76 < M_{ee} < 106$ GeV
Muon selection criteria	$\lvert\eta_\mu\rvert < 2.4$
	$p_T > 20$ GeV
	μ ID track $\lvert z_0(\mu)\rvert < 1\,\text{mm}$ from PV
	$\lvert d_0/\sigma(d_0)\rvert < 3$
	Isolation: $\Sigma p_{\text{T}}(ID)/p_{\text{T}} < 0.1$
$Z \to \mu\mu$ selection	Two oppositely charged muons
	Veto on events with further leptons
	$76 < M_{\mu\mu} < 106$ GeV
$E_{\text{T}}^{\text{miss}}$	$E_{\text{T}}^{\text{miss}} < 70$ GeV
Jet selection criteria	Anti-k_T, $R = 0.4$
	$p_T > 20$ GeV
	$\lvert y\rvert < 2.4$
	$\Delta R(jet, lepton) > 0.5$
	$JVF > 0.75$
	Two leading tagged jets are used

Table 4.3.: Object and event selection applied to data and Monte Carlo.

element production. During combination of these samples care has to be taken so that double counting of $Z + b\bar{b}$ events does not occur, as both samples can produce the same final state of particles. The samples are combined under the assumption that the $b\bar{b}$ pairs produced by the parton shower will model the production at low angular separation more accurately than the matrix element prediction. The reverse is also assumed; the matrix element prediction will model the $b\bar{b}$ pair production with large angular separation better. The cut-off of $\Delta R = 0.4$ is used so that in the

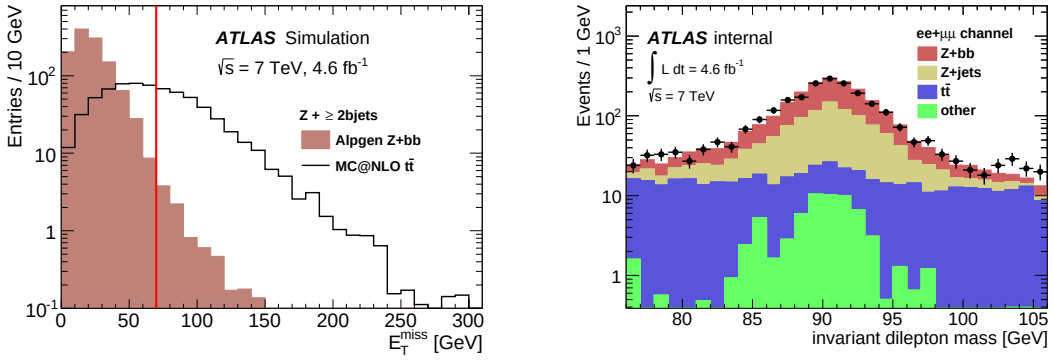
combined sample $b\bar{b}$ pairs produced with $\Delta R < 0.4$ come from the Z+jet sample and $b\bar{b}$ pairs produced with $\Delta R > 0.4$ come from the $Z + b\bar{b}$ sample.

Background Monte Carlo

The largest background comes from Z+jet events where light jets (jets originating from a $u/d/s$ type quarks or gluon) and c-jets (originating from a c type quark) are misidentified as b-jets. This background is taken from the ALPGEN Z+jet Monte Carlo samples. The next largest background comes from top-quark pair production, with both W-bosons decaying to same flavour leptons. The Monte Carlo sample used to model $t\bar{t}$ events was produced using the MC@NLO [65] generator. MC@NLO is a NLO matrix element generator interfaced to a parton shower simulation. Double counting of final states between the matrix element and parton shower are removed internally and can result in negative event weights. The sample was produced using the CT10 PDF set [66] and was interfaced to HERWIG and JIMMY for showering, hadronisation and underlying event modeling. Single-top production (including in association with a W-boson) as well as di-boson production are also backgrounds to the analysis, the simulated samples for these were also produced with MC@NLO interfaced with HERWIG and JIMMY. The exception being single-top t-channel production which was produced using a specialist Standard Model background processes LO generator, ACER-MC 3.7 [67] which used the CTEQ6L1 PDF set and was interfaced to PYTHIA version 6.425 [68] for showering, hadronisation and underlying event simulation. Table 4.4 details the main MC samples used with the production cross-section. Even with large production rates at the LHC, backgrounds from multi-jet QCD events were found to be consistent with zero due to the event selection requirements of two isolated leptons and two b-tagged jets. The QCD background estimation was taken from a data driven method looking at events in a control region expected to contain more QCD events. A fit to the m_{ll} distribution found that in the fiducial region of this analysis this background was negligible [1]. The cuts on di-lepton mass and $E_{\mathrm{T}}^{\mathrm{miss}}$ are applied to remove a large proportion of backgrounds in the dataset. This can be seen in Figure 4.2 which shows for the Monte Carlo events passing the selection criteria the distribution of $E_{\mathrm{T}}^{\mathrm{miss}}$ and the di-lepton invariant mass.

Description	MC generator	$\epsilon_F \cdot \sigma$ [pb]
$Zb\bar{b}$+0 part., $Z \to e^+e^-$	ALPGEN	6.57
$Zb\bar{b}$+1 part., $Z \to e^+e^-$	ALPGEN	2.48
$Zb\bar{b}$+2 part., $Z \to e^+e^-$	ALPGEN	0.89
$Zb\bar{b}$+3 part., $Z \to e^+e^-$	ALPGEN	0.39
Z+0 part., $Z \to e^+e^-$	ALPGEN	668.32
Z+1 part., $Z \to e^+e^-$	ALPGEN	134.36
Z+2 part., $Z \to e^+e^-$	ALPGEN	40.54
Z+3 part., $Z \to e^+e^-$	ALPGEN	11.16
Z+4 part., $Z \to e^+e^-$	ALPGEN	2.88
Z+5 part., $Z \to e^+e^-$	ALPGEN	0.83
$Zb\bar{b}$+0 part., $Z \to \mu^+\mu^-$	ALPGEN	6.56
$Zb\bar{b}$+1 part., $Z \to \mu^+\mu^-$	ALPGEN	2.47
$Zb\bar{b}$+2 part., $Z \to \mu^+\mu^-$	ALPGEN	0.89
$Zb\bar{b}$+3 part., $Z \to \mu^+\mu^-$	ALPGEN	0.39
Z+0 part., $Z \to \mu^+\mu^-$	ALPGEN	668.68
Z+1 part., $Z \to \mu^+\mu^-$	ALPGEN	134.14
Z+2 part., $Z \to \mu^+\mu^-$	ALPGEN	40.3
Z+3 part., $Z \to \mu^+\mu^-$	ALPGEN	11.19
Z+4 part., $Z \to \mu^+\mu^-$	ALPGEN	2.75
Z+5 part., $Z \to \mu^+\mu^-$	ALPGEN	0.77
$t\bar{t}$	MC@NLO	79.01
Wt inclusive	MC@NLO	14.59
Single top s-channel	MC@NLO	0.47
Single top t-channel	ACER-MC	7.12
ZZ, $\ell\ell qq$	MC@NLO	0.559
W^+Z, $qq\ell\ell$	MC@NLO	0.5415
W^-Z, $qq\ell\ell$	MC@NLO	0.2944

Table 4.4.: List of Monte Carlo samples used in the $Z + b\bar{b}$ analysis including the generator and filter efficiency (ϵ_F) times cross-section.



(a) (b)

Figure 4.2.: Monte Carlo signal and background distributions for (a) $E_{\mathrm{T}}^{\mathrm{miss}}$, and (b) di-lepton mass [1].

Corrections to MC

Various corrections are applied to the Monte Carlo samples so that the simulation is modeling data as accurately as possible. All the Monte Carlo samples listed in the previous sections that contained b-quarks and used HERWIG for hadronisation were reweighted so that the decay tables for B-hadrons was given by the EVTGEN package [69]. The EVTGEN package has decay tables based on Particle Data Group [49] data and is expected to reproduce the decay spectrum more accurately than the internal B-hadron decay tables used in HERWIG. This correction was applied after discovering the number of charged particles from secondary vertices was mis-modeled in Monte Carlo samples showered and hadronised with HERWIG [1]. Other corrections applied to the Monte Carlo samples, which are derived by the various ATLAS performance groups, are listed below:

- *Pile-up correction:* The number of pile-up interactions included in the Monte Carlo is re-weighted to that of data. The re-weighting is performed as a function of the average number of interactions, $< \mu >$.
- *Primary vertex position correction:* The Monte Carlo z-position of the primary vertices was not representative of data. Monte Carlo events are re-weighted to bring alignment with data.
- *Trigger scale-factor correction:* The trigger efficiency in Monte Carlo is corrected so that it agrees with data. The efficiency in data is measured across the full

p_{T} threshold spectrum using a tag and probe method using $Z \to \mu^+\mu^-$ and $J/\Psi \to \mu\mu$ events for muons [70, 71]. The di-electron trigger again uses tag and probe methods on $Z \to e^+e^-$ events to find the single electron trigger efficiency, with the di-electron efficiency being the product of two single electron trigger efficiencies [72].

- *Lepton resolution correction:* The measured lepton energy is smeared so that it resembles the data distribution.
- *Lepton reconstruction efficiency correction:* The Monte Carlo reconstruction efficiency for leptons is corrected. The correction is derived as the ratio of reconstruction efficiency in data over Monte Carlo and is measured using tag and probe methods [73]. The correction is parameterised by lepton kinematics.
- *Jet-resolution correction:* The jet momentum resolution in Monte Carlo is smeared so that it is in agreement with data.
- *B-tagging correction:* The b-tagging efficiency for b-tagging algorithms is determined for various operating points defined by the efficiency to tag a real b-jet. Mismodeling of the b-tagging efficiency in Monte Carlo is corrected by using scale factors derived from the difference in data and Monte Carlo [74].
- *Jet mis-tag correction:* The operating points of b-tagging algorithms also are defined based on the mis-tag rate, explicitly the rate at which a non b-jet is tagged as a b-jet. The mis-tag rate in MC for light and charm-jets are scaled so that they match data [75, 76].

Systematic uncertainties for all the Monte Carlo corrections are estimated in the analysis to account for any possible under or over compensation.

4.3. Signal Extraction

To extract differential cross-sections, detector level $Z + b\bar{b}$ event yields are obtained by a series of extended maximum likelihood fits performed on data in bins of the differential observables. Each fit uses several templates describing the different signal and background components to the dataset. Each contribution, either signal or background, is modeled by Monte Carlo, using the samples detailed in Table 4.4.

The templates are fitted to the data and the number of signal events, $N_{b\bar{b}}$ can be extracted.

The fitted distribution is a variable that is designed to discriminate between b-jets and non-b-jets. The variable uses a combination of outputs from the JetFitter-CombNN b-tagging algorithm, p_x, defined in Section 3.8.4. For each tagged jet in an event the variable $\ln({}^{p_b}/_{p_c})$ is formed. As each p_x can be interpreted as the probability that a jet is of flavour x, this variable has on average higher values for b-jets. Figure 4.3 shows a shape comparison of the Monte Carlo distributions of $\ln({}^{p_b}/_{p_c})$ for jets which have been either classed as a b-jet, c-jet or light-jet. It can be seen that the distribution for b-jets is considerably different for this variable compared to the other jet flavours, which gives good discriminating power. For $Z + b\bar{b}$ measurements, where two b-jets are expected, an event level variable is used for fitting. This combines $\ln({}^{p_b}/_{p_c})$ for each tagged jet: $\sum \ln({}^{p_b}/_{p_c})$, where the summation is over both tagged jets in an event. In the case of more than two tagged jets in an event (approximately 3% of the data sample) the two jets with the largest p_{T} are used in the summation.

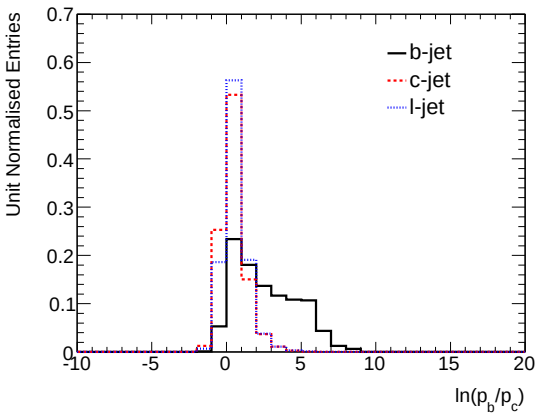


Figure 4.3.: Template Shape comparison for b-jets, c-jets and all other jets (l-jets). See Section 4.4 for jet flavour definitions in Monte Carlo.

For fitting to data, all the Monte Carlo templates are organised into three categories, $Z + b\bar{b}$, Z+*non-*$b\bar{b}$ and *background*:

- The Z+$b\bar{b}$ template models signal events. It is derived from the combined Z+jet and $Z + b\bar{b}$ ALPGEN Monte Carlo sample, for events passing event selection with the two b-tagged jets matching different B-hadrons, see Section 4.4 for the matching procedure.

- The Z+non-$b\bar{b}$ component is modeled by ALPGEN. It is a combination of the five irreducible Z+jet backgrounds distinguishable by the di-jet flavour combinations. Labeling jet flavours: l=light-jet, c=c-jet and b=b-jet, this template consists of a Z-boson in association with di-jet flavour combinations bc, bl, cc, cl, ll. The ratio of these templates in the Z+non-$b\bar{b}$ template can not be taken straight from Monte Carlo and must be scaled due to an underestimation of heavy-flavour jet production with ALPGEN.
- The background template consists of the combined top and di-boson backgrounds, of which $t\bar{t}$ is dominant. The template is taken directly from the background Monte Carlo samples passing event selection, with the normalisation fixed to the NLO predictions scaled to the data luminosity.

To help with the limited statistics of the data and Monte Carlo samples the electron and muon channel events are combined into a single dataset for fitting. This allows for fit results with higher statistical precision and allows for higher granularity in the measured kinematic observables. Figure 4.4 shows the $\sum \ln({}^{p_b}/_{p_c})$ distribution for the combined ALPGEN Z+jet and $Z + b\bar{b}$ Monte Carlo for both lepton decay channels. In Figure 4.4 (a) all events passing selection are plotted, where Figure 4.4 (b) shows the distribution of the Z+$b\bar{b}$ template where the tagged jets are matched to B-hadrons. It can be seen that the template shapes are compatible and no fit bias from combining the Monte Carlo samples is expected.

4.3.1. Z+non-$b\bar{b}$ Template Construction

The fit groups di-jet flavour combinations: bc, bl, cc, cl and ll into a single Z+non-$b\bar{b}$ template. Figure 4.5 shows the shape in the fit variable for each of these contributions; it also includes the Z+$b\bar{b}$ template shape for comparison. The limited statistics and degeneracy evident between the shapes of the individual components does not allow for a separation of the Z+non-$b\bar{b}$ into individual di-jet flavour combinations. Simply adding these contributions to form the non-$b\bar{b}$ template assumes that the Monte Carlo correctly describes the relative proportion of each of these elements. However, results from a preceding analysis [77] indicate that ALPGEN +HERWIG samples underestimate the production of $Z + b$ events, which suggests the bc and bl contributions could be underestimated. To have an accurate representation of the data the relative fraction

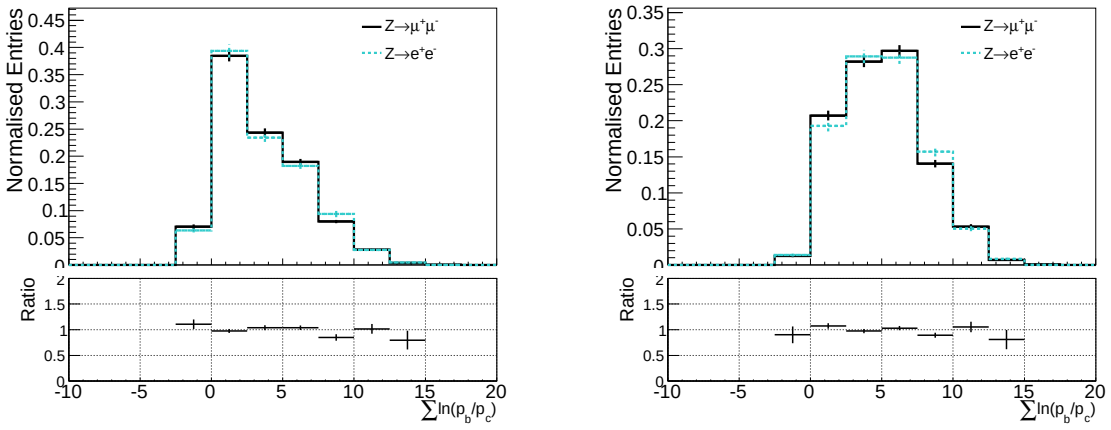


Figure 4.4.: Unity normalised distributions of the fit variable $\sum \ln({}^{p_b}/_{p_c})$ in Monte Carlo for the two lepton channels, for (a) all events passing selection, and (b) the signal component only.

of single b-jet events in the template need to be determined. The relative charm and light contributions to the non-$b\bar{b}$ template also need verification.

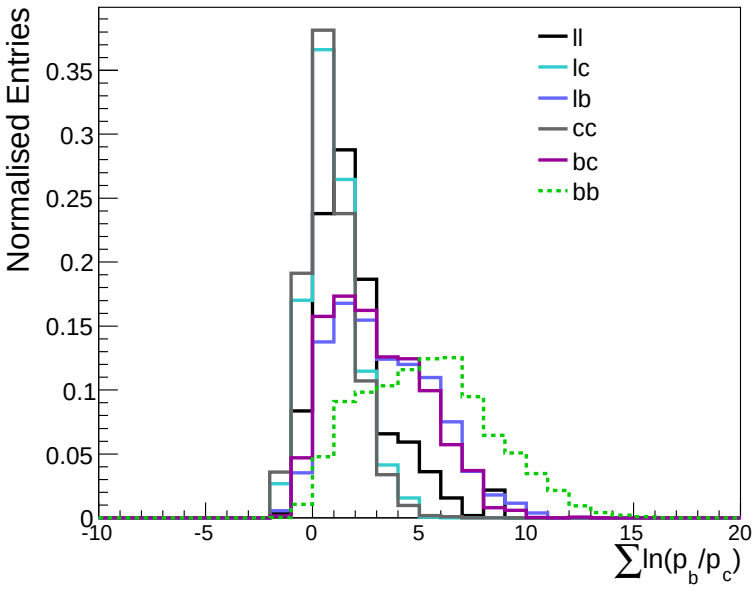


Figure 4.5.: Shape comparison of all di-jet flavour combinations for $\sum \ln({}^{p_b}/_{p_c})$ distribution.

To derive scale factors for the bc, bl, cc and cl templates, and to check the Monte Carlo prediction for the relative fraction of the light contribution, the following maximum likelihood fits to data are performed with differing jet selection criteria.

1. Fit to $ln\,(p_b/p_l)$ distribution for the b-tagged jet in a sample of $Z+ \geq 2$ jets with exactly one b-tag, the extra non-tagged jet helps mimic the event topology of a $Z + b\bar{b}$ event. For the $ln\,(p_b/p_l)$ variable, charm and light jet response differs so the fit is used to derive the default single-b scale factor as well as a charm scale factor.

2. Fit to $ln\,(p_b/p_l)$ distribution for the b-tagged jet in the inclusive $Z+ \geq 1$ tagged jet sample, to check the kinematic dependence of the charm and single-b scale factors and set a systematic uncertainty.

3. Fit to $ln\,(p_b/p_c)$ distribution for the b-tagged jet in a sample of $Z+ \geq 2$ jets with exactly one tag as a cross-check on the fit to $ln\,(p_b/p_l)$ to validate the single-b-jet scale factor derived in 1. The charm scale factor cannot reliably be obtained in this fit, due to the degeneracy of the charm and light templates in the $ln\,(p_b/p_c)$ distribution.

4. Fit to $ln\,(p_b/p_c)$ distribution for the b-tagged jet in the inclusive $Z+ \geq 1$ tagged jet sample, again to validate the single-b scale factor derived in the fit to $ln\,(p_b/p_l)$.

The results of these fits can be seen in Figure 4.6 for $ln\,(p_b/p_l)$, which in addition includes the templates before fitting to data to highlight the missing contribution in the Alpgen Monte Carlo samples. Figure 4.7 shows the validation fits in $ln\,(p_b/p_c)$.

In the nominal scale factor extraction fit, shown in Figure 4.6 (a), a single-b scale factor of 1.35 ± 0.03 is obtained. The validation check for this scale factor, taken from Figure 4.7 (a), and gives a consistent 1.33 ± 0.03. The more inclusive fit region without the requirement of at least another jet in the event shown in Figure 4.6 (b), gives a single-b scale factor of 1.48 ± 0.02. Therefore a conservative 15% systematic uncertainty to this single-b scale factor is assigned which covers the difference in the two regions. The inclusive Z+1-tagged jet fits are in good agreement regarding the single-b scale factor when using either $ln\,(p_b/p_c)$ or $ln\,(p_b/p_l)$ as the fit variable.

In Figure 4.6, the default scale factor fits using $ln\,(p_b/p_l)$, the charm and light scale factors are consistent with unity. The $ln\,(p_b/p_c)$ shape is degenerate for c-jets

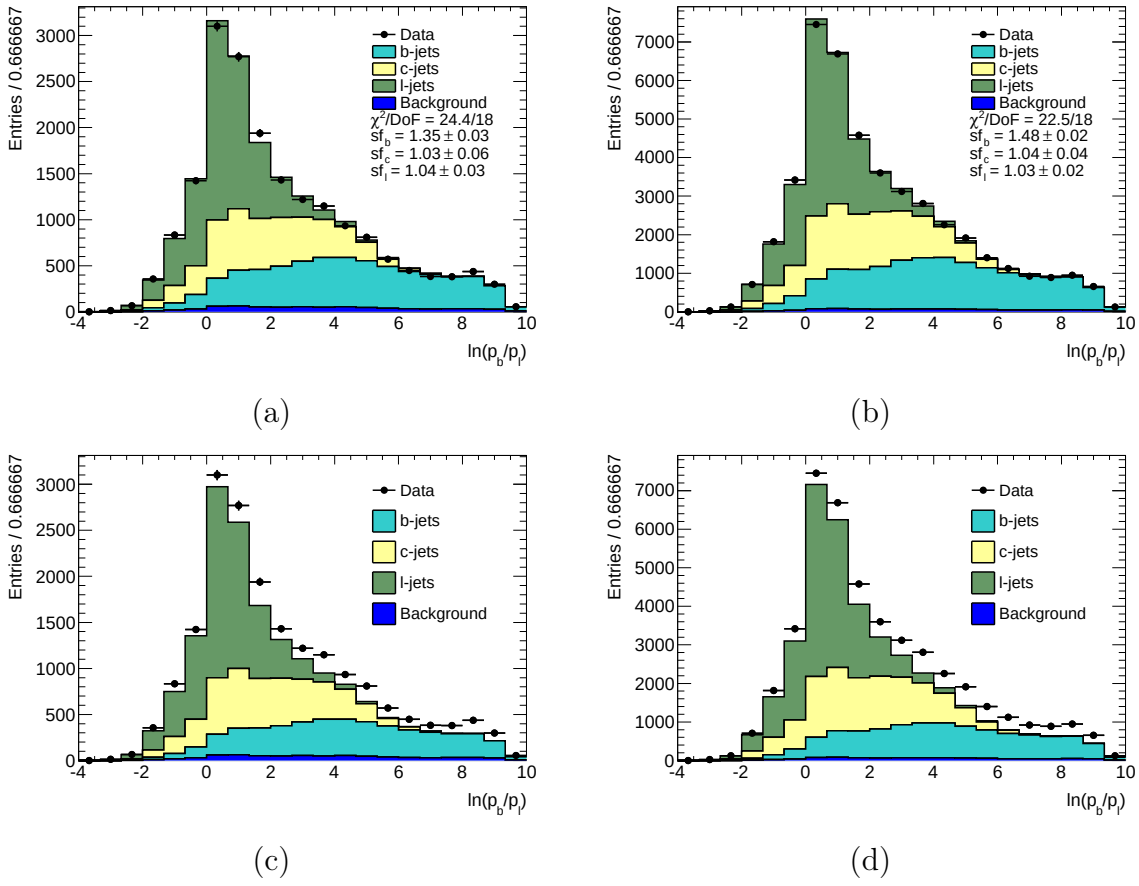
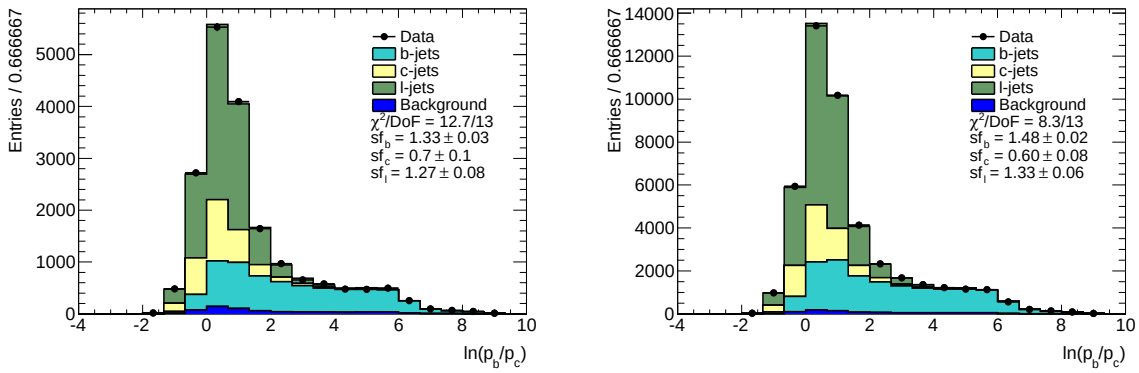


Figure 4.6.: Extracting b and c jet scale factors from fits to the $ln\,(p_b/p_l)$ distribution in (a) the Z+≥2-jet data sample with exactly one b-tagged jet, and (b) the inclusive Z+1-tagged jet data sample. In (c) and (d) the templates directly taken from ALPGEN before fitting are compared with data for the Z+≥2-jet data sample with exactly one b-tagged jet (c) and the inclusive Z+1-tagged jet data sample (d) cases.



(a) (b)

Figure 4.7.: Fits used for validating the results from Figure 4.6. The b-jet scale factors are derived by fitting the $ln\,(p_b/p_c)$ distribution in (a) the Z+$\geq$2-jet data sample with exactly one b-tagged jet, and (b) the inclusive Z+1-tagged jet data sample. As charm and light response is the same in $ln\,(p_b/p_c)$ this is used only to check the consistency of the single-b scale factor derived by fitting the $ln\,(p_b/p_l)$ distribution.

and l-jets so scaling up/down the charm contribution to the non-$b\bar{b}$ template has little effect. The charm scale factor is taken as unity with a 15% systematic uncertainty cover any possible mismodeling effects. The Z+non-$b\bar{b}$ template used to fit the nominal double-tagged sample is thus constructed using the following scaling:

1. ll, lc and cc taken directly from Monte Carlo prediction
2. lb and cb are scaled up by 1.35.

4.3.2. Differential Fit Results

The data $\sum \ln({}^{p_b}/_{p_c})$ distribution is fitted in bins of four kinematic distributions: $p_T(Z), |y(Z)|, m(b,\bar{b})$ and $\Delta R(b,\bar{b})$. The binning for each variable was defined so that there was approximately equal data statistics in each bin, and is listed in Table 4.5. Each differential bin is fitted individually with potential statistical correlations between the observables ignored throughout.

The three Monte Carlo templates describing both the signal and background events are constructed in each of the bins in Table 4.5. All three template shapes are

Bin	$\Delta R(b,\bar{b})$	$p_T(Z)$ [GeV]	$\lvert y(Z)\rvert$	$m(b,\bar{b})$ [GeV]
1	0.4-1.15	0-25	0-0.2	10-45
2	1.15-1.9	25-40	0.2-0.4	45-85
3	1.9-2.4	40-55	0.4-0.6	85-115
4	2.4-2.8	55-70	0.6-0.8	115-165
5	2.8-3.2	70-95	0.8-1.2	165-350
6	3.2-3-5	95-250	1.2-1.6	
7			1.6-2.5	

Table 4.5.: Bins of the $Z + b\bar{b}$ differential observables.

kept fixed to the Monte Carlo predictions during the fit. The normalisation of the background template is also constant, which leaves the normalisation of the $Z + b\bar{b}$ and Z+non-$b\bar{b}$ templates as the only free parameters in the fit. Some example fit results for each differential bin in the measured $p_T(Z)$ distribution are included in Figure 4.8. The full set of fits for the other observables are included in Appendix A. The fit quality is generally good with the total fitted Monte Carlo templates in reasonable agreement with data in all bins.

4.3.3. Fit Performance

The fit method needs to be validated so that the fit is not imposing any bias onto the result and is accurately distinguishing heavy flavour jets. Three independent checks were performed to assess the fit performance.

1. *Linearity Test*: For each bin in the $Z + b\bar{b}$ differential observables, the three post-fit templates (Z+$b\bar{b}$, Z+non-$b\bar{b}$ and background) are used as a basis to form a set of pseudo-templates. The Z+$b\bar{b}$ template is further scaled by a factor X, (X=0.5, 0.70, 1.0, 1.5, 2.0). The scaled Z+$b\bar{b}$ template is then combined with the Z+non-$b\bar{b}$ and background templates to form a parent sample. From each parent sample 500 pseudo-datasets are formed by sampling from each bin in the parent template distribution assuming Poisson statistics. Each of the pseudo data-sets are then fitted with the original pre-fit templates. The number of fitted $Z + b\bar{b}$ events, $N_{b\bar{b}}^{toy}$ is compared to the original post-fit value, $N_{b\bar{b}}$, and the value of X is then estimated by taking the ratio $N_{b\bar{b}}^{toy}/N_{b\bar{b}}$. The X vs

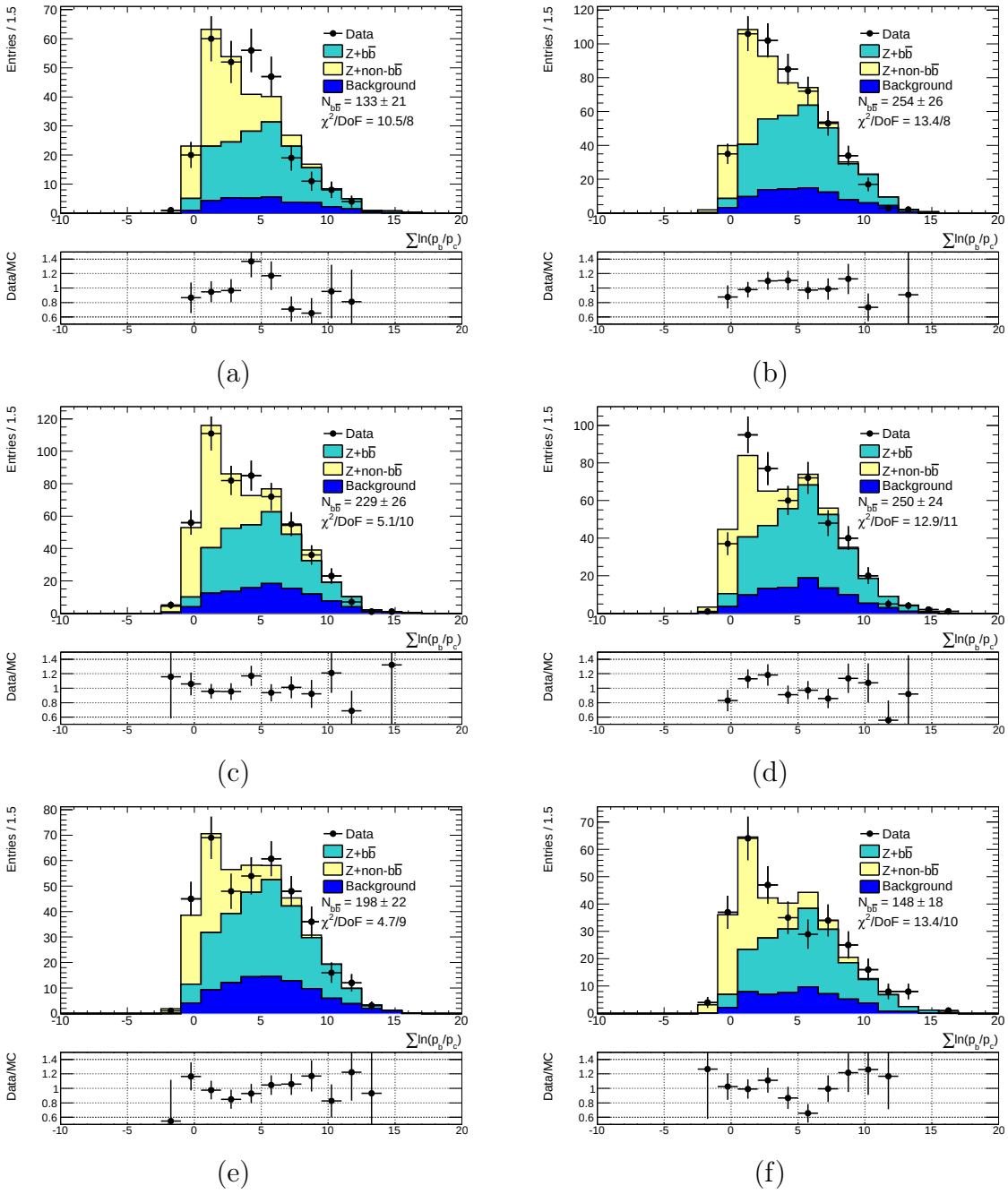


Figure 4.8.: Result of fitting the number of true $b\bar{b}$ pairs, $N_{b\bar{b}}$, in bins of Z transverse momentum for (a) $p_T(Z) < 25$GeV, (b) $25 < p_T(Z) < 40$GeV, (c) $40 < p_T(Z) < 55$GeV, (d) $55 < p_T(Z) < 70$GeV, (e) $70 < p_T(Z) < 95$GeV and (f) $95 < p_T(Z) < 250$GeV.

$N^{toy}_{b\bar{b}}/N_{b\bar{b}}$ distribution is then fitted with a straight line, this is then compared to the line $y = x$ in Figure 4.9. The good consistency with the $y = x$ function shows the fit is accurately measuring the number of signal events. The same method is also repeated with the Z+non-$b\bar{b}$ background being scaled instead of the signal component the linearity results shown in Figure 4.10, again excellent performance.

2. *Pull Test*: Post-fit templates are combined into a single parent template. The number of fitted $Z + b\bar{b}$ events $N_{b\bar{b}}$ is taken from the normalisation of the Z+$b\bar{b}$ template. 1000 pseudo datasets are drawn from the parent template by fluctuating each bin in the $ln\,(p_b/p_c)$ distribution in the parent sample around the bin content using a Poisson distribution. The pseudo datasets are then fitted with the pre-fit templates using the standard fit procedure. The number of Z+$b\bar{b}$ events, $N^P_{b\bar{b}}$, is taken from the fit along with its error, $\sigma^P_{b\bar{b}}$. The pull is defined as $(N^P_{b\bar{b}} - N_{b\bar{b}})/\sigma^P_{b\bar{b}}$, it should be a Gaussian with mean of zero (with a deviation from this indicative of a bias in the heavy flavour normalisation estimator) and a width of unity (a pull distribution narrower or wider than this means the errors could be over or under estimated respectively). The pull distributions are fitted with Gaussian functions which are compared to a normal distribution with mean zero and standard deviation of 1 in Figure 4.11. The same procedure is repeated with the pull distributions for the other free parameter in the fit, $N_{non-b\bar{b}}$. The results are shown in Figure 4.12. Both figures show that the fit is performing as expected.

3. *Bin Number Stability*: To test the stability of the fit procedure it is important to check that the results are independent of the binning chosen for the fit variable. The default fit is binned in 20 bins of $\sum \ln\left({}^{p_b}/_{p_c}\right)$ for the range: $-10 < \sum \ln\left({}^{p_b}/_{p_c}\right) < 20$. The fits are repeated in this range using 5, 10, 30 and 60 bins. The fitted number of $N_{b\bar{b}}$ and $N_{non-b\bar{b}}$ are compared for the different bin multiplicities. It is found that across all the differential observables the fit results are consistent within the statistical uncertainty from the fit. Figure 4.13 shows the distributions of $N_{b\bar{b}}$ and $N_{non-b\bar{b}}$ derived using the different binnings of the fit variable for each of the differential $p_T(Z)$ fits.

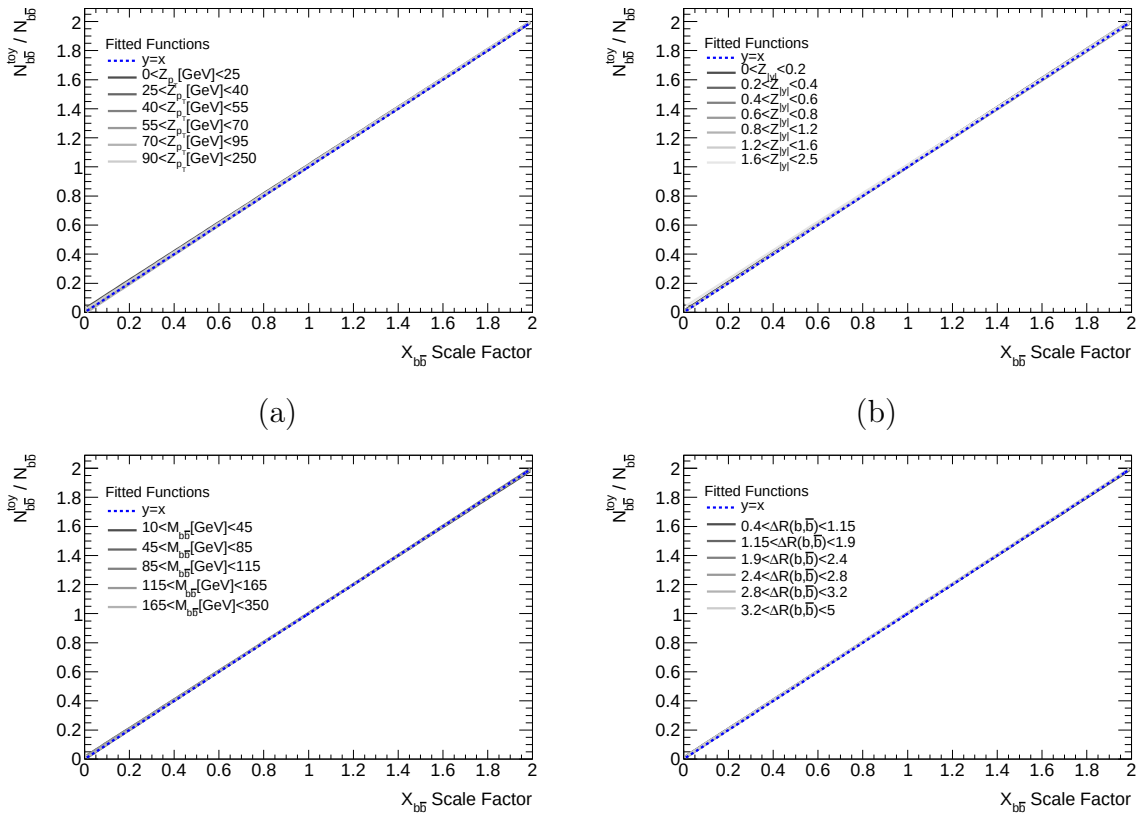


Figure 4.9.: Linearity tests for the $Z + b\bar{b}$ template in all fitted differential bins of, (a) $p_T(Z)$, (b) $|y(Z)|$, (c) $m(b,\bar{b})$, (d) $\Delta R(b,\bar{b})$.

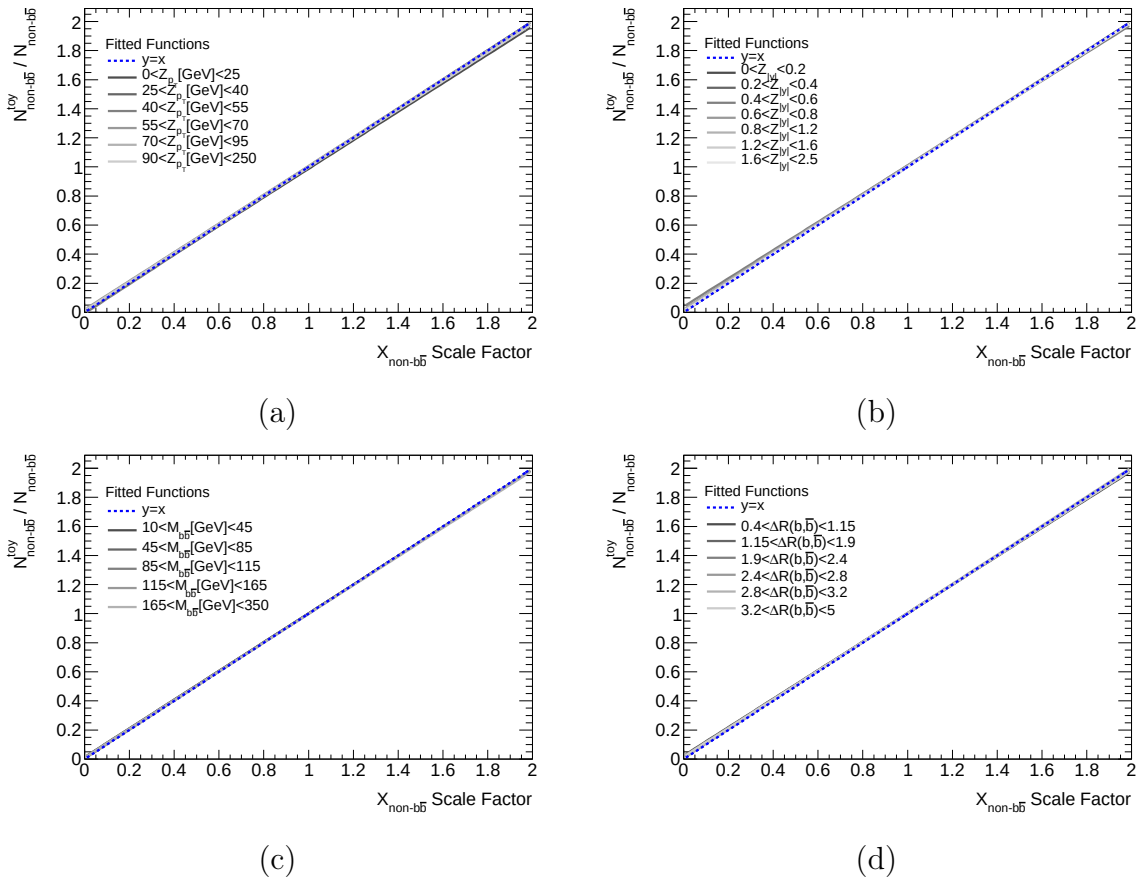


Figure 4.10.: Linearity tests for the Z++non-$b\bar{b}$ template in all fitted differential bins of, (a) $p_T(Z)$, (b) $|y(Z)|$, (c) $m(b,\bar{b})$, (d) $\Delta R(b,\bar{b})$.

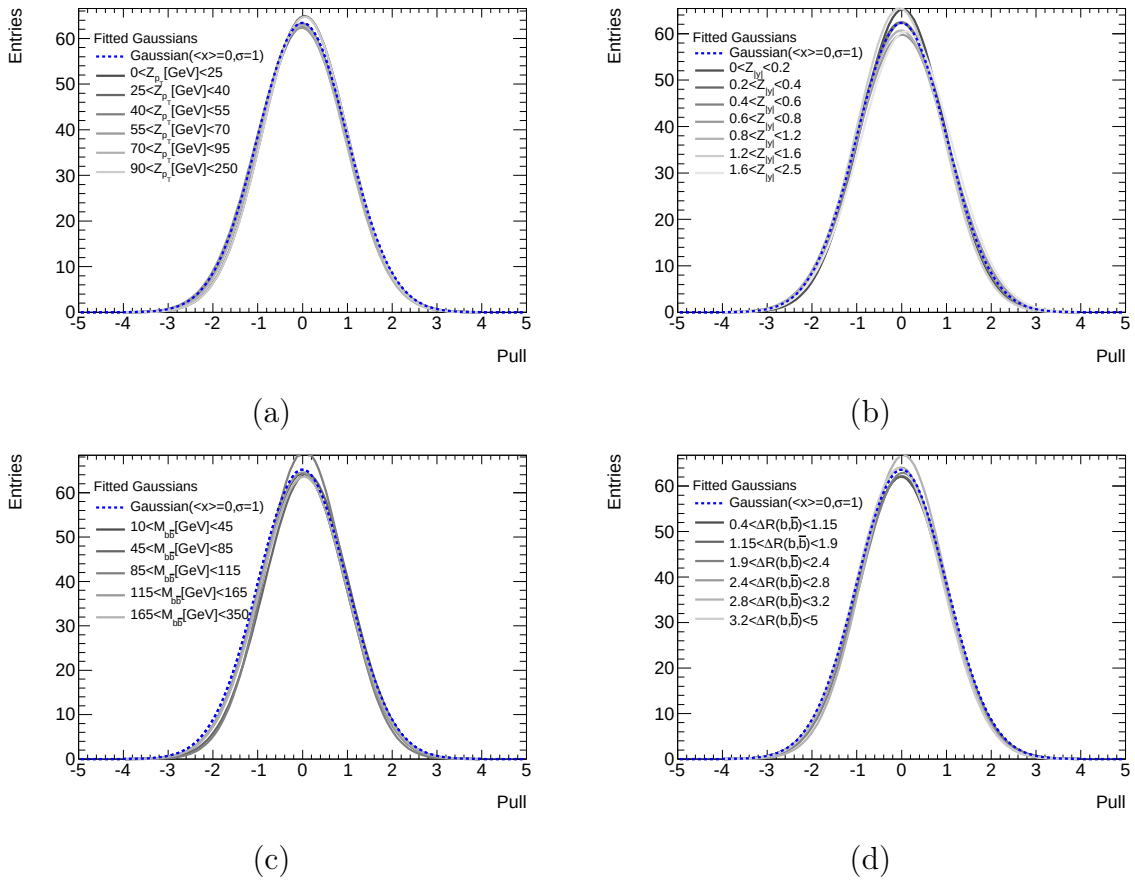


Figure 4.11.: Pull tests for the $Z + b\bar{b}$ template in all fitted differential bins of, (a) $p_T(Z)$, (b) $|y(Z)|$, (c) $m(b,\bar{b})$, (d) $\Delta R(b,\bar{b})$.

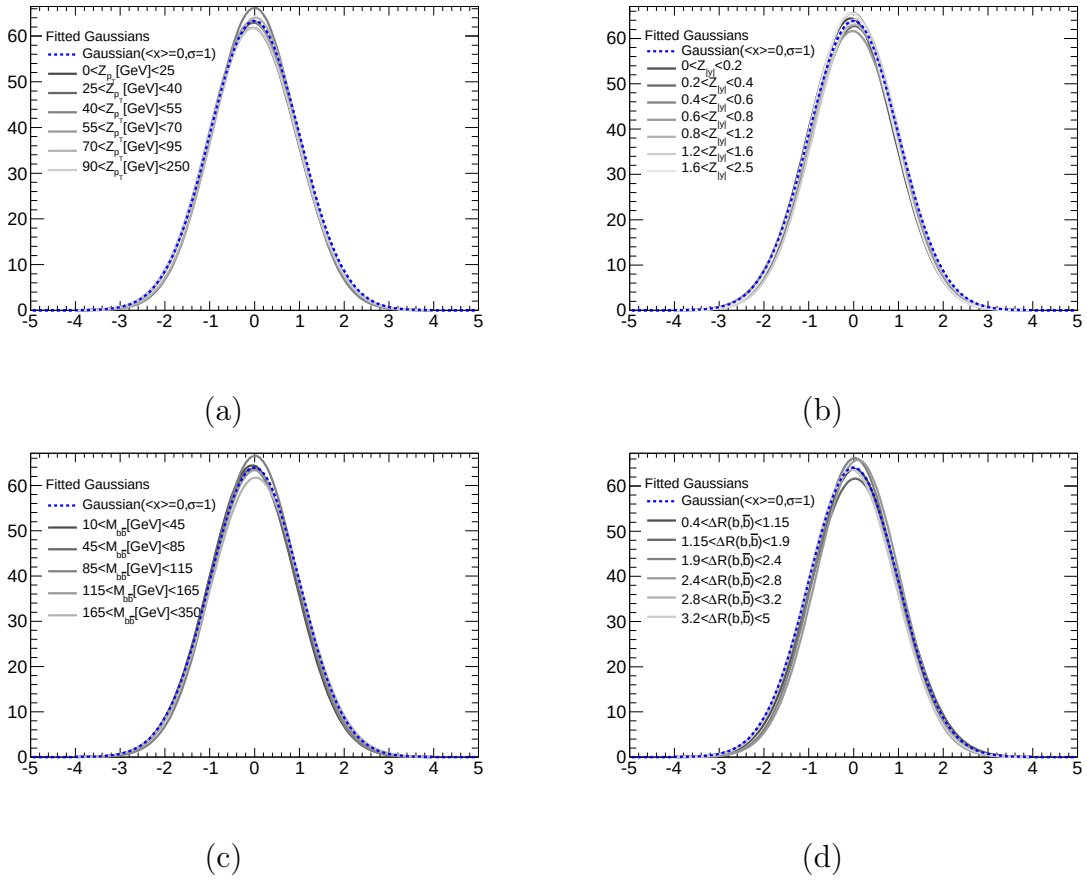


Figure 4.12.: Pull tests for the Z++non-$b\bar{b}$ template in all fitted differential bins of, (a) $p_T(Z)$, (b) $|y(Z)|$, (c) $m(b,\bar{b})$, (d) $\Delta R(b,\bar{b})$.

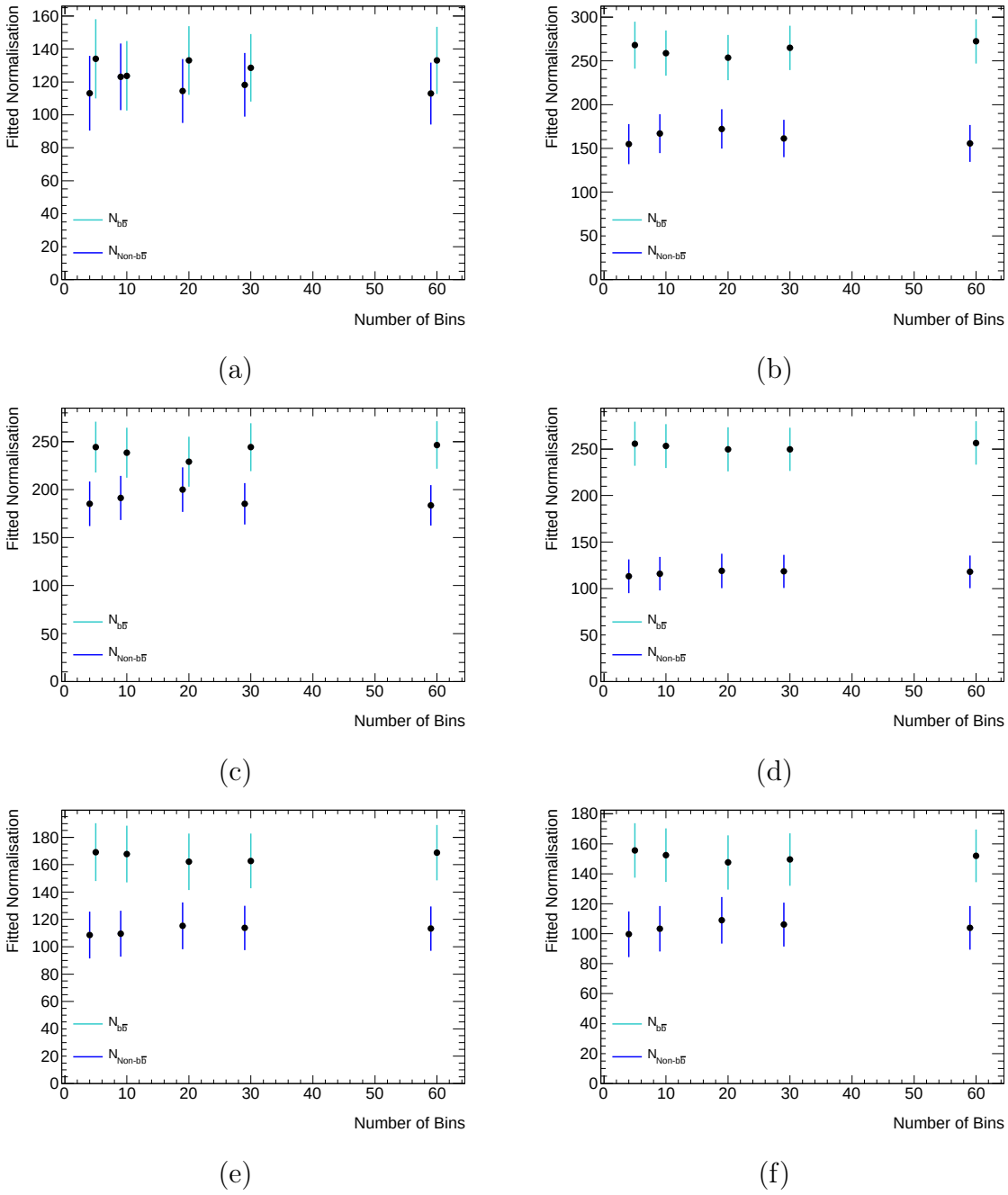


Figure 4.13.: Fitted normalisations of floating parameters $N_{b\bar{b}}$ and $N_{non-b\bar{b}}$ as a function of number of bins of $\sum \ln({}^{p_b}/_{p_c})$ used during the fit. For (a) $p_T(Z) < 25$GeV, (b) $25 < p_T(Z) < 40$GeV, (c) $40 < p_T(Z) < 55$GeV, (d) $55 < p_T(Z) < 70$GeV, (e) $70 < p_T(Z) < 95$GeV and (f) $95 < p_T(Z) < 250$GeV. The $N_{non-b\bar{b}}$ distribution is offset by -1 on the x-axis to avoid overlapping distributions.

4.4. Unfolding

The fitted data yields are unfolded back to the particle level. This allows consistent comparison to particle level Monte Carlo predictions and also removes the effects of the detector resolution and efficiency so that the result can be compared to other experiments within the defined fiducial region. The particle level event selection is kept as close as possible to that of the data selection, so that the theoretical extrapolation to unmeasured phase space is kept to a minimum. The particle level selection is outlined in Table 4.6. The main differences to the data selection is that there is no cut on $E_{\mathrm{T}}^{\mathrm{miss}}$ and the lepton rapidity cuts for both lepton channels are the same at the particle level, $|\eta| < 2.5$.

The particle level cross-section is defined as:

$$\sigma_{Z+b\bar{b}} = \frac{N_{b\bar{b}} \cdot \mathcal{C}}{L}, \tag{4.1}$$

with $N_{b\bar{b}}$ the number of fitted signal events, $\mathcal{C}$ the unfolding correction factor and L the measured integrated luminosity. $\mathcal{C}$ is used to correct for detector efficiencies and resolution and is comprised of two terms:

$$\mathcal{C} = \epsilon_{b\bar{b}} \cdot \mathcal{C}_f, \tag{4.2}$$

which are applied sequentially:

1. $\epsilon_{b\bar{b}}$: The correction for the efficiency to tag b-jets.
2. $\mathcal{C}_f$: A bin-by-bin correction factor corrects for detector resolution effects, reconstruction and trigger inefficiencies.

These are described in more detail in the subsequent sections.

4.4.1. $\epsilon_{b\bar{b}}$ Correction

The largest single contribution to the unfolding correction is from the efficiency to tag two b-jets, $\epsilon_{b\bar{b}}$. Using the MV1 75% efficiency working point the efficiency to tag two jets will be approximately 50%. $\epsilon_{b\bar{b}}$ is separated from the other parts of the

Lepton dressing	All stable leptons are dressed so that all photon 4-momentum in a cone $\Delta R < 0.1$ around the lepton is added to the lepton 4-momenta
Lepton selection	In the event of more than two leptons, the highest p_{T} dressed, same flavour opposite signed leptons are selected. They are required to have $p_{\mathrm{T}} > 20$ GeV, $\lvert\eta\rvert < 2.5$ and have a di-lepton mass $76 < M < 106$ GeV.
Jet Selection	All other stable particles are passed to the jet finding algorithm. Jets are required to have $p_{\mathrm{T}} > 20$ GeV and $\lvert y\rvert < 2.4$. Any jets within $\Delta R < 0.5$ of a signal lepton are discarded.
Jet flavour	Jets are assigned a flavour sequentially. First b-jets are identified by matching to weakly decaying B-hadron with hadron $p_{\mathrm{T}} > 5$ GeV using a spatial match of $\Delta R < 0.3$. Next c-jets are identified with the leftover jets matched to charm quarks with the same $\Delta R < 0.3$ matching. Any jet not labeled as a b-jet or c-jet is then labeled as a light jet.
Event Selection	Events require a di-lepton pair passing the above requirements and at least two b-jets. In the case of more than two b-jets, the b-jets with the highest p_{T} are selected for construction of the differential observables.

Table 4.6.: Particle and event selection at the particle level.

unfolding as it is a correction derived solely from reconstruction-level objects. The correction is derived from the combined ALPGEN Z+jet and $Z + b\bar{b}$ Monte Carlo samples which have passed the full event selection. $\epsilon_{b\bar{b}}$ is calculated from the ratio of events with two reconstructed jets matched to separate B-hadrons over the number of the events which have two b-tagged jets. $\epsilon_{b\bar{b}}$ is calculated separately for each differential bin which are shown in Figure 4.14 as a function of each distribution.

4.4.2. $\mathcal{C}_f$ Correction

Data yields need further correction for detector and selection inefficiencies. Due to the large bin widths of the differential distributions the migration of events, from particle level to reconstruction, between bins is expected to be small, so by default a simple bin-by-bin correction factor, $\mathcal{C}_f$, is used. As a cross check a Bayesian unfolding method has also been implemented. For both of these methods a 2D histogram, or *migration matrix*, is produced by filling the underflow bin of the y-axis with all events passing the event selection. The underflow of the x-axis is filled by all events passing the particle level selection. An event that passes both reconstruction and particle level cuts and has the leading reconstruction level jets match to the leading particle level jets, with a spatial matching criterion of $\Delta R < 0.4$, populates the main body of the matrix. The matrix is binned for both the x and y-axis in the bins of the differential observable. A diagonal migration matrix represents observables where the resolution is narrow compared to the binning used.

$\mathcal{C}_f$ is derived as the ratio of the number of events filling the corresponding bin in the particle level underflow to the number of events populating the reconstruction level underflow bin. Figure 4.14 includes the $\mathcal{C}_f$ corrections as a function of each measured differential distribution. The Bayesian unfolding method uses the main body of the migration matrix as it accounts for bin migrations. The migration matrices are shown in Figure 4.15 where the matrices have been normalised so that the bins show the probability of particle level events being reconstructed in each bin. Due to the statistical limitation of the dataset, the binning defined in Table 4.5 is chosen to give equal event numbers in each bin which allows for a stable fit. As the default choice is bin-by-bin unfolding it is important that the bin widths are also larger than the resolution of the kinematic variable of interest. The matrices are diagonal which justifies the use of bin-by-bin unfolding, which ignores migration of events between observable bins, and means the chosen binning is wider than the variable resolution. Figure 4.16 shows the unfolded cross-section comparing the default unfolding method with the Bayesian alternative with any difference between the two methods yielding a result well within the statistical uncertainty of the fitted signal yield.

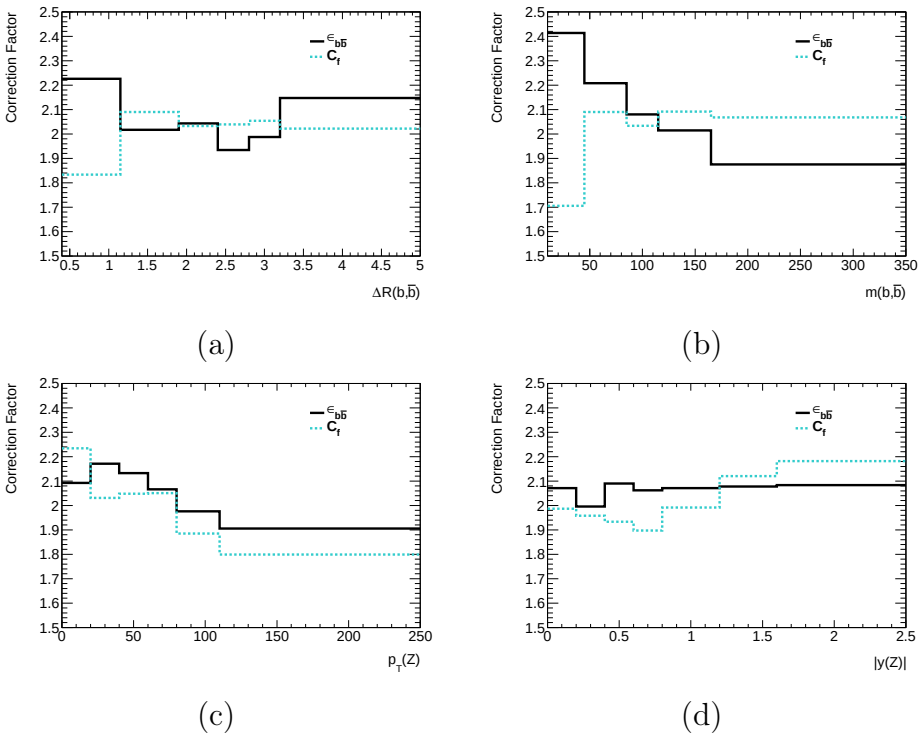


Figure 4.14.: Unfolding correction factors for the b-tagging efficiency, $\epsilon_{b\bar{b}}$, and the bin-by-bin correction factors, $\mathcal{C}_f$. Shown as a function of (a) $\Delta R(b, \bar{b})$, (b) $m_{b\bar{b}}$, (c) $p_T(Z)$ and (d) $|y(Z)|$.

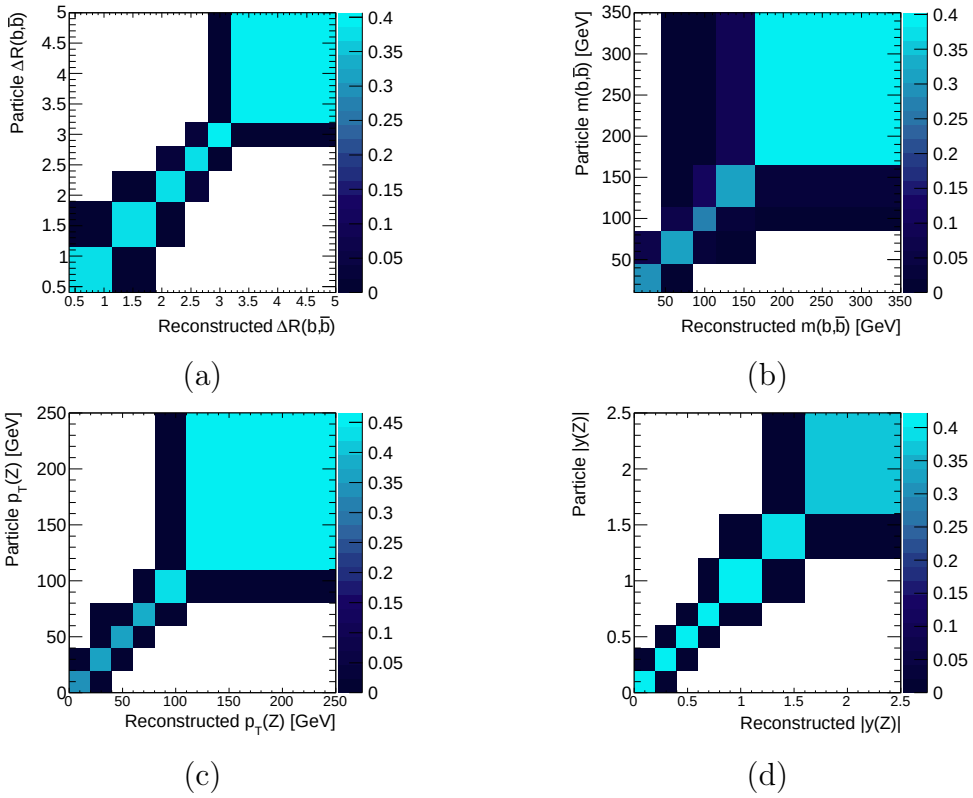


Figure 4.15.: Migration matrices for (a) $\Delta R(b, \bar{b})$, (b) $m_{b\bar{b}}$, (c) $p_T(Z)$ and (d) $|y(Z)|$. The matrices have been normalised so that the bins show the probability of particle level events being reconstructed in each bin.

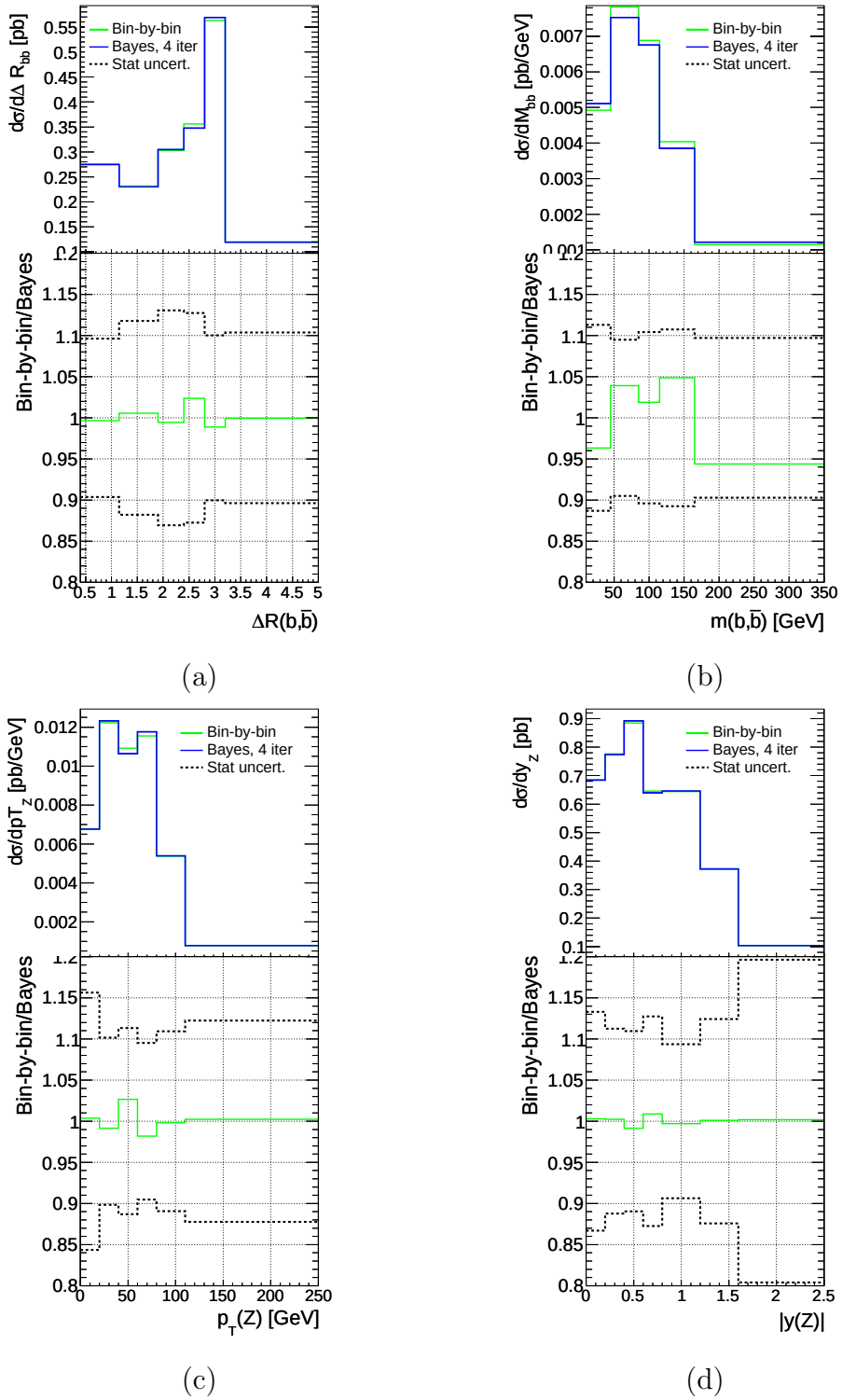


Figure 4.16.: Comparison between different unfolding methods at particle level for (a) $\Delta R(b, \bar{b})$, (b) $m_{b\bar{b}}$, (c) $p_T(Z)$ and (d) $|y(Z)|$.

4.5. Systematic Uncertainties

Systematics are accounted for by repeating the entire analysis chain with a change to the Monte Carlo samples that reflects a reasonable uncertainty on a part of the analysis. This means that systematic sources can change either the fit result (by altering template shapes), the unfolding to particle level (by taking events in or out of acceptance), or both of these simultaneously. The largest systematics are summarised in Table 4.7. These and all other sources of systematic uncertainty considered are detailed below.

Source of uncertainty	Relative fractional range [%]
b-jet template shape	4.2-20.4
Total b-jet tagging efficiency	8.5-12.0
c-jet tagging efficiency	1.0-5.1
Z+non-$b\bar{b}$ template b-jet scale factor	1.3-5.7
Jet energy scale	3.1-7.1
Background normalisation	1.1-5.5
$t\bar{t}$ modeling	3.0
MC template statistics	2.9-6.0

Table 4.7.: Summary of the largest systematic uncertainties determined for the cross-section measurements. The range covers all differential observables.

4.5.1. Jet-Tagging Efficiency

This systematic represents the uncertainty on the Monte Carlo scale factors that were applied to bring the data and Monte Carlo b-tagging efficiencies into agreement. The efficiency to tag a b, c or l-jet is parameterised as a function of jet-p_{T} and η. The uncertainty on a particular jet flavour tagging efficiency is assessed by individually varying the scale factors by the associated uncertainty. The efficiency to tag a jet is anti-correlated to the tagging inefficiency. For consistency when increasing an efficiency the corresponding inefficiency scale factors are decreased and vice versa. The uncertainty on the b-tagging efficiency scale factors dominates the c/l-tagging efficiency and is further broken down into 10 eigenvector components. The scale

factors are derived in 10 p_{T} bins and the breakdown into eigenvectors accounts for correlation between the different p_{T} bins. The breakdown into uncorrelated eigenvectors reduces the number individual sources of uncertainty from over 50 to 10 and is derived by the ATLAS b-tagging performance group [74]. Due to the relative size of the uncertainty on c-jet and l-jet components these have not been broken down into their eigenvector contributions. Figure 4.17 shows the fractional uncertainty of each of the tagging scale-factor uncertainties. The statistical error from the fit result is shown for comparison as well as the combined jet-tagging uncertainties, defined as the sum in quadrature of all the individual components.

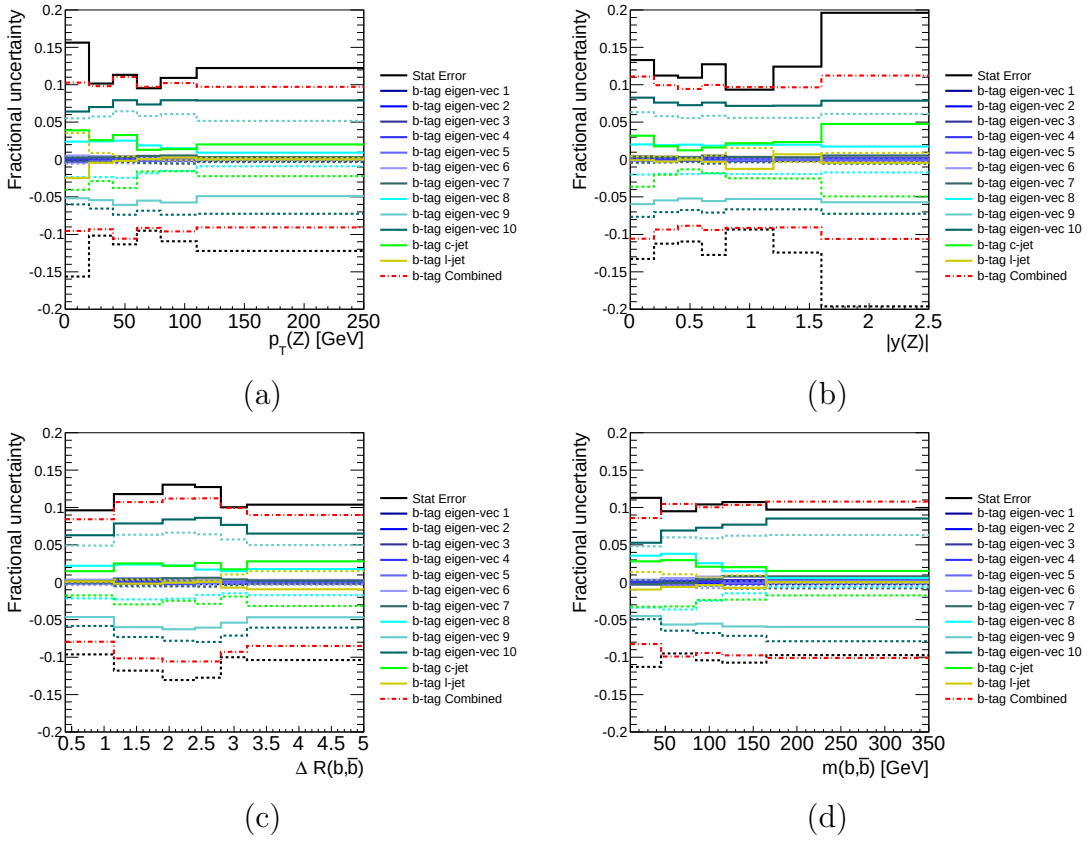


Figure 4.17.: Fractional uncertainty for jet-tagging systematic uncertainties for (a) $p_T(Z)$, (b) $|y(Z)|$, (c) $\Delta R(b,\bar{b})$ and (d) $m(b,\bar{b})$ distributions. Plots include the statistical uncertainty on the measurement and the combined uncertainty from all jet-tagging systematic sources. A shift in a downwards direction (i.e. -1σ) is represented by a dashed line. Solid lines represent upward variations.

4.5.2. Jet Energy Scale and Jet Energy Resolution

The uncertainty on the jet energy scale (JES) has 16 different components [78]. Figure 4.18 shows the fractional uncertainty of each of the JES uncertainties, the statistical error from the fit result is shown for comparison as well as the sum in quadrature of all the individual uncertainty components. Six of the components are again a set of eigenvectors from correlated sources of uncertainty taken from in-situ measurements, these are referred to as "NP" in Figure 4.18. The other sources are attributable to jet flavour, jet kinematics, pile-up environment, presence of close by jets and different responses in the Monte Carlo samples used for JES determination. Each JES component is parameterised as a function of p_{T} and η and is individually perturbed by $\pm 1\sigma$ and propagated through the analysis chain, with the difference to the nominal result taken as the systematic uncertainty.

The jet energy resolution in Monte Carlo is further smeared to account for any possible underestimation. This is achieved through an additional smearing of jet p_{T} [79]. For an estimate of the effect of the JER being overestimated in Monte Carlo the systematic is symmetrised and included in Figure 4.21.

4.5.3. Lepton Uncertainties

The lepton scale factors for identification and reconstruction efficiency are independently varied, within their associated uncertainty, for each lepton channel by $\pm 1\sigma$ of the default scale factor. The uncertainty due to the electron scale-factors are around the 2% level with the muon scale factor consistently below 1%. The lepton energy scale and resolution is also individually varied by $\pm 1\sigma$ with the caveat that the muon momentum resolution has been separated into momentum resolution variations in the inner detector and muon spectrometer separately. These have negligible effect, with uncertainties below 1% in the full fiducial region.

4.5.4. $E_{\mathrm{T}}^{\mathrm{miss}}$ Uncertainty

$E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated from several components which can loosely be thought of as the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculated using the reconstructed physics objects. When an uncertainty is propagated on a particular physics object, e.g. the JES, the corresponding $E_{\mathrm{T}}^{\mathrm{miss}}$

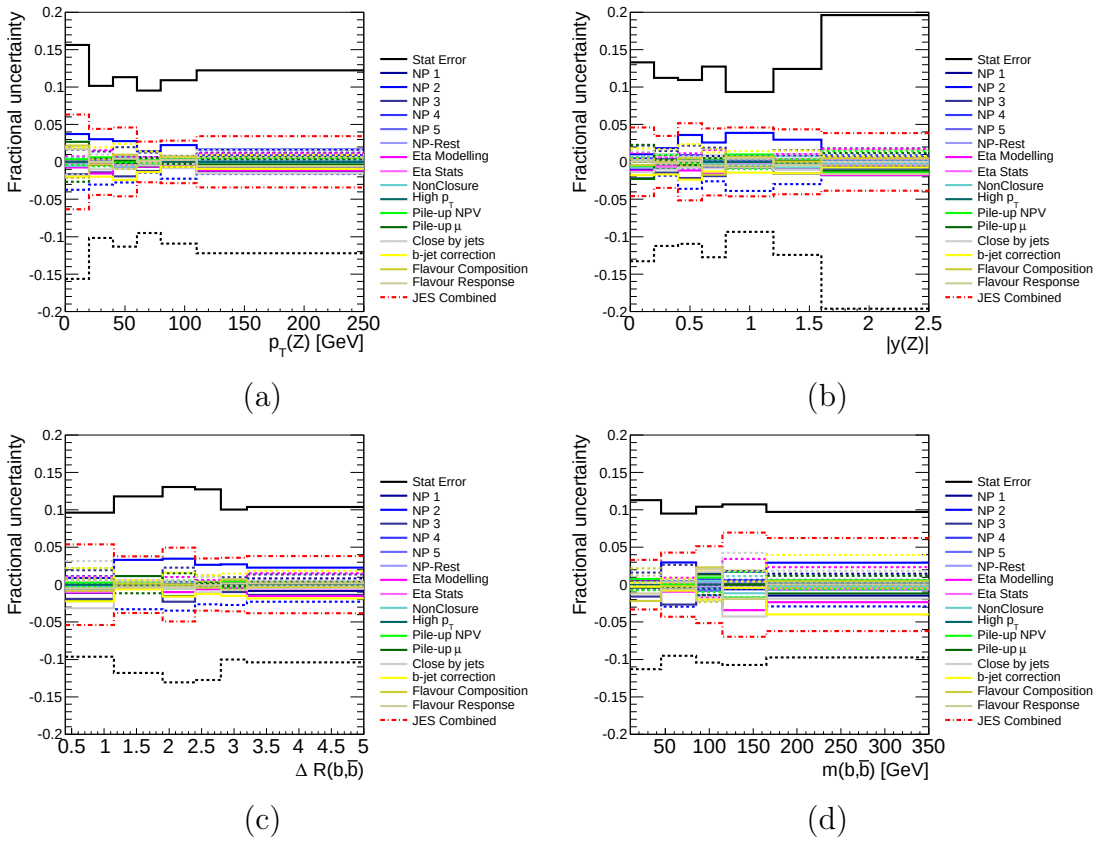


Figure 4.18.: Fractional uncertainty for JES systematic uncertainties for (a) $p_T(Z)$, (b) $|y(Z)|$, (c) $\Delta R(b,\bar{b})$ and (d) $m(b,\bar{b})$ distributions. Plots include the statistical uncertainty on the measurement and the combined uncertainty from all JES components. A shift in a downwards direction (i.e. -1σ) is represented by a dashed line. Solid lines represent upward variations.

component, here jets, is consistently altered also. However there are parts of the total $E_{\mathrm{T}}^{\mathrm{miss}}$ that are not covered by any of the physics objects in this analysis. A mis-modeling of the $E_{\mathrm{T}}^{\mathrm{miss}}$ of topo-clusters in the Monte Carlo that are not part of reconstructed physics objects is covered by simultaneously varying the scale and resolution uncertainties on this term [80]. Separately the $E_{\mathrm{T}}^{\mathrm{miss}}$ term for jets within a $7 < p_{\mathrm{T}} < 20$ GeV range also has the scale and resolution varied. The result is minimal with uncertainties well below the 1% level.

4.5.5. z-Position Re-Weighting

Monte Carlo samples were re-weighted to correct for the z-position of primary vertices. This re-weighting was turned off and the analysis repeated as a conservative uncertainty on the z-position re-weighting. The resulting uncertainty is typically around the 2% level but can be larger in certain differential bins.

4.5.6. Luminosity Uncertainty

An uncertainty on the delivered integrated luminosity of 1.8% is used [46].

4.5.7. b-Jet Template Shape

The re-weighting of events based on the decay tables in EvtGen, as outlined in Section 4.2.2, changes the shape of the $\sum \ln({}^{p_b}/_{p_c})$ for events containing at least one jet containing a B-hadron. To assess the uncertainty of this template shape correction or any residual mismodeling of b-jet response in Monte Carlo an alternative data-driven jet re-weighting procedure is applied based on a $t\bar{t}$ control region. The control region is defined for events with an isolated lepton with at least four jets with exactly two of them b-tagged, with all jets passing the same cuts detailed in Table 4.3. Expected backgrounds in this control region from W+jets and single top events are estimated from simulation. This control region is expected to be dominated by real b-jets from $t\bar{t}$ decays. The jet flavour variable $ln\,(p_b/p_c)$ shape is then compared to a MC@NLO $t\bar{t}$ sample which is interfaced to the same HERWIG version used for the ALPGEN Z+jet samples but without the EvtGen weights applied.

The Data/MC ratio is then fit in the b-jet dominated region $2.0 < ln\,(p_b/p_c) < 9.0$ with a 3rd order polynomial, Figure 4.19 shows the function in the fit range. Jets are re-weighted by this function with jets falling outside the fit range given a weight of 1. The analysis is repeated with templates using this alternative re-weighting scheme and the difference compared to the default case is taken as a systematic.

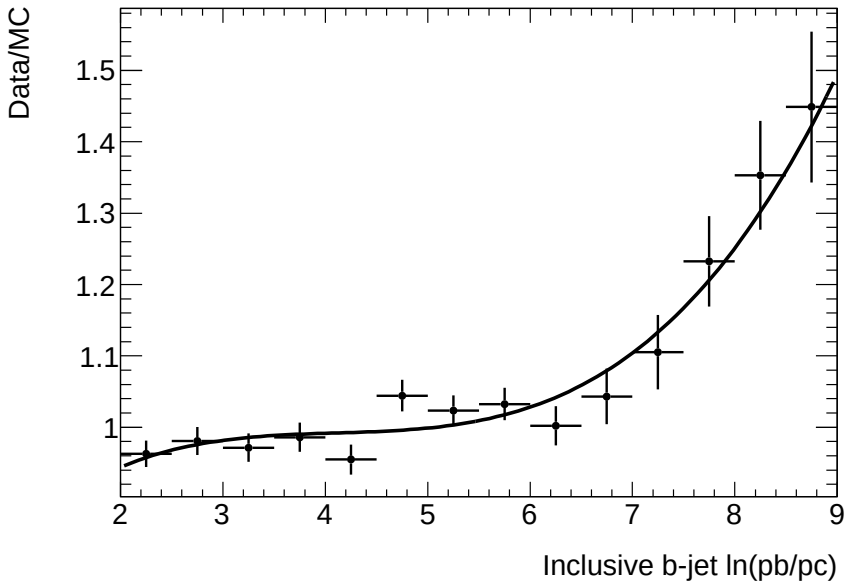


Figure 4.19.: Data/MC ratio of $\ln(\frac{p_b}{p_c})$ fitted with a third order polynomial in $t\bar{t}$ control region [1].

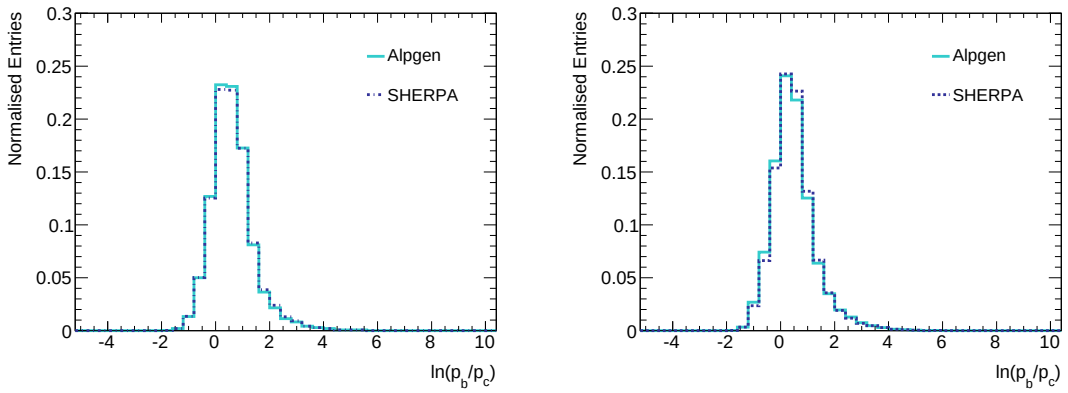
This results in one of the largest systematics as the number of fitted $Z + b\bar{b}$ events is heavily dependent on the $ln\,(p_b/p_c)$ b-jet shape. The fractional uncertainty due to the b-template re-weight is included in Figure 4.21.

4.5.8. Non-b Template Shapes

An uncertainty on the template shape of $ln\,(p_b/p_c)$ for c-jets comes from re-weighting this distribution to the shape of an alternative Monte Carlo sample. A simulated SHERPA [81] sample of $Z + c$ events, where c-jets are defined as jets matched to a charm hadron which is not from a cascade decay of a B-hadron, is used to produce the alternative templates. The ratio of template shapes for the default and alternative sample is fit with a 2^{nd} order polynomial function which defines the systematic c-jet weights as a function of $ln\,(p_b/p_c)$.

The light-jet shape uncertainty is calculated in an analogous way to the c-jet uncertainty. Again a SHERPA sample is used as an alternative, a Z+jet sample is produced with a veto on events with B-hadrons to ensure the light-jet template is described accurately. A per l-jet weight function is derived by taking the ratio of the SHERPA and ALPGEN samples so that the reweighted jets have the SHERPA $ln\,(p_b/p_c)$ distribution.

Figure 4.20 includes the default ALPGEN and the systematic SHERPA $ln\,(p_b/p_c)$ distribution for both c and l-jets. The fractional uncertainty from these alternative template shapes is included in Figure 4.21.



(a) (b)

Figure 4.20.: $ln\,(p_b/p_c)$ distribution for ALPGEN and SHERPA for (a) l-jets and (b) c-jets.

4.5.9. Z+non-$b\bar{b}$ Template Construction Uncertainty

The uncertainty on the Z+non-$b\bar{b}$ template composition as described in Section 4.3.1, is assessed by separately scaling the single-b components and the c components as follows:

- The single-b components of the Z+non-$b\bar{b}$ template is scaled up and down by 15%.
- Each c-jet component is scaled up and down by 15%.

The results are included in Figure 4.21

4.5.10. Monte Carlo Statistics

The limited statistics of the Monte Carlo sample used can impact the result in two places, the fit and the unfolding. A systematic is assigned for both of these. To assess the impact of the available statistics of the Monte Carlo templates used in the fit, a series of pseudo-templates are formed by varying the populations of each bin in a given $\sum \ln ({}^{p_b}/_{p_c})$ template by a random sampling from a Poisson distribution with the mean of the original bin occupancy. This is repeated for the two floating fit templates, Z+$b\bar{b}$ and Z+non-$b\bar{b}$, so that a set of 5000 pseudo-templates are formed. These are each used to fit the data and the distribution of the difference of fitted signal events to the default fit is fitted with a Gaussian function. The width of the Gaussian is taken as the uncertainty on the Monte Carlo template statistics. For unfolding, the statistical uncertainty on the number of events in the migration matrices is propagated through to the bin-by-bin correction factors and the unfolding step is repeated varying the default factor by the corresponding uncertainty. Both sources of Monte Carlo statistical uncertainties are included in Figure 4.21.

4.5.11. Background Uncertainties

The backgrounds containing real b-jets have template shapes that are similar to the signal Z+$b\bar{b}$ template. The normalisation of these predictions are thus fixed in the fits to the NLO prediction of the cross-sections. A conservative $\pm 10\%$ scale factor on the normalisation of the background template is used and the fits are repeated, with the difference from the default fit taken as the uncertainty. The size of this systematic reflects the uncertainty on the $t\bar{t}$ cross-section [82], which is the dominant contribution to the background template. The fractional uncertainty on the background normalisation is included in Figure 4.21.

The $t\bar{t}$ background dominates the background template so another separate uncertainty estimate to assess modeling of this background is performed. The normalisation of an alternative $t\bar{t}$ Monte Carlo sample generated with POWHEG [83], which is matched for showering and hadronisation to PYTHIA, passing event selection was found to give a 3% difference to the fitted yield. Unfortunately, the statistics available in this sample did not allow a differential comparison so an additional flat 3% uncertainty for the $t\bar{t}$ modeling is included.

Finally a data-driven study found that the background from purely QCD jet events, determined by fitting to QCD enhanced control regions, was negligible. A flat uncertainty of 0.5% is assumed to cover any residual QCD events passing event selection and is taken from the uncertainty on the QCD data control region fits [1].

4.5.12. MPI Model Systematic

The amount of MPI events, where the Z-boson and b-jets are from different hard-interactions within the same p-p interaction, could affect the differential distributions of signal events. The Monte Carlo truth-record flags events originating from MPI and these events have been re-weighted by $\pm 50\%$ and the analysis repeated. Systematically changing the MPI contribution by 50% was found to cover the expected MPI from combining the inclusive Z [84] and b-jet [85] cross-sections with the effective p-p cross-section [86] for the $Z + b$ analysis [1]. The systematic difference from the nominal cross-section is typically at the 1% level across the differential distributions but rises to 5% at low $p_T(Z)$.

4.6. Theoretical Predictions

For comparison to the unfolded data several particle level theoretical predictions have been produced both at NLO and LO:

- *MCFM:* A fixed order perturbative QCD prediction at parton level, MCFM [87] is a NLO prediction of $Z + b\bar{b}$ production calculated using the 5FNS. The calculation uses a massless b-quark approximation, with $\mu_F = \mu_R = \sqrt{m(Z)^2 + p_T(Z)^2}$. For an uncertainty on the factorization and renormalisation scales, each are varied up and down by a factor of two to estimate the missing contributions from higher order effects. The calculation is done with three different PDF sets CT10 [61, 88], NNPDF2.3 [89] and MSTW2008 [27], with the uncertainty on each PDF propagated through the calculation. The uncertainty on the value of the strong coupling constant is assessed by using PDFs where the value $\alpha_s(m(Z))$ is varied by $\pm 1\sigma$ around the default value. The envelope of separately altering each theoretical uncertainty with respect to the default setup is taken as the theoretical uncertainty on the MCFM prediction.

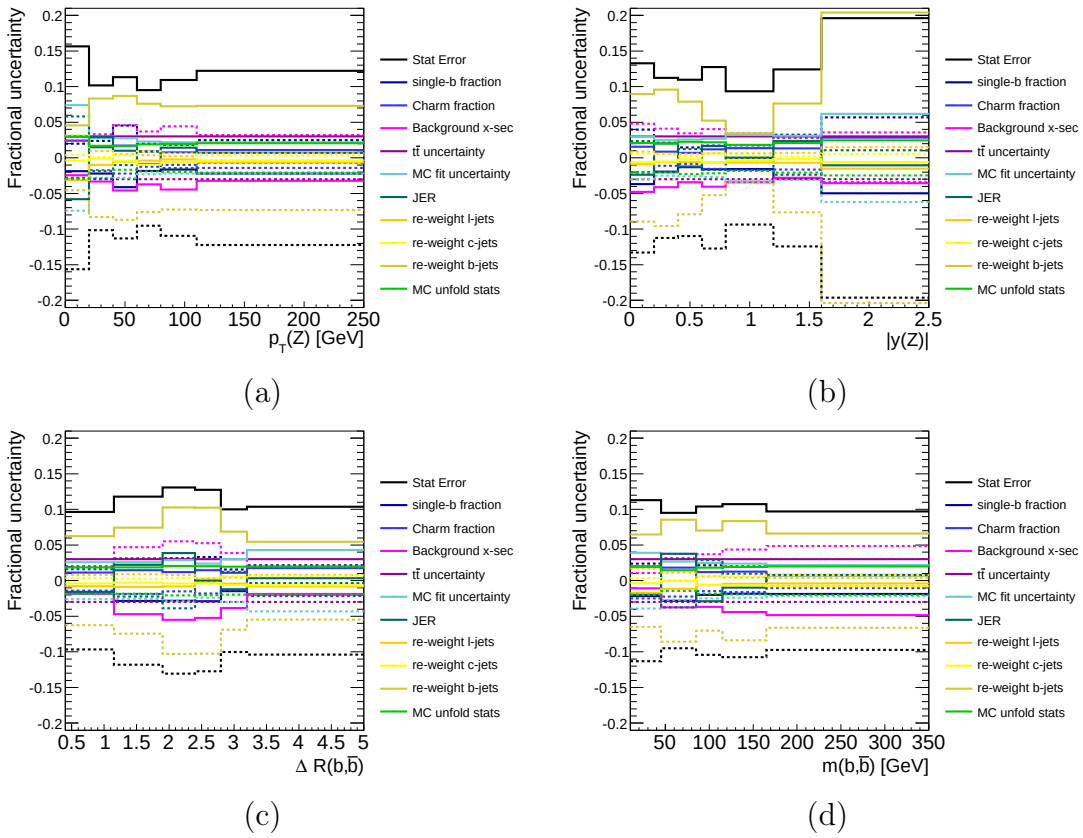


Figure 4.21.: Fractional uncertainty for non-negligible systematic uncertainties, not including jet-tagging and JES for clarity, for (a) $p_T(Z)$, (b) $|y(Z)|$, (c) $\Delta R(b, \bar{b})$ and (d) $m(b, \bar{b})$ distributions. Plots include the statistical uncertainty on the measurement for comparison. A shift in a downwards direction (i.e. -1σ) is represented by a dashed line. Solid lines represent upward variations.

As the MCFM prediction is at the parton level it needs to be corrected to the particle level for a fair comparison with the unfolded data and thus needs corrections for non-perturbative QCD effects and final state QED radiation. The non-perturbative effects, which include hadronisation, underlying event activity and MPI are estimated using SHERPA and PYTHIA. A $Z + b\bar{b}$ sample generated with SHERPA v1.4.1 using the CT10 PDF set is used where the non-perturbative effects can be switched on and off. The correction is derived in each differential bin by taking the ratio of the number of events passing the particle level selection with the effects turned on to the events passing the particle level selection with the corrections turned off, with the caveat that a b-quark is matched to a jet rather than a B-hadron for the parton level prediction. An identical prescription is followed with a PYTHIA 6.427 sample using the CTEQ5L PDF set. The corrections derived from the both predictions agree very well, with the average of the two used to correct the MCFM prediction. For the QED final state radiation the correction is derived from the ALPGEN Monte Carlo sample, described in Section 4.2.2, where PHOTOS [90] has been used to simulate lepton FSR. A comparison of dressed leptons to the Born level leptons without any final state radiation gives the correction to the MCFM sample. This correction was calculated from the combined electron and muon channel and not found to be dependent on the differential kinematics so was applied as a flat 0.9739 correction to all observable bins.

- *aMC@NLO:* Another NLO calculation is performed at particle level using the 4FNS with the MSTW2008NLO_nf4 PDF set. NLO matrix elements are showered with HERWIG++ [91] and include all non-perturbative QCD corrections as well as QED final state radiation. The b-quark mass effects are included in the calculation, with the mass fixed at the pole-mass of 4.5 GeV. The factorisation and renormalisation scales are set to the MCFM choices for consistency. A second aMC@NLO sample is also used for comparison and was calculated in the 5FNS. This alternative sample produced was for $Z+\geq$ 1-jet prediction at NLO, which results in a LO prediction for $Z + b\bar{b}$ production. The 5FNS sample is produced ignoring the b-quark mass and using the MSTW2008NLO PDF set and again is matched to HERWIG++ for showering and the inclusion of non-perturbative effects.

The aMC@NLO predictions are missing a component when the Z-boson and $b\bar{b}$-pair are produced in separate partonic interactions within the same p-p interaction. This MPI contribution is included in the signal definition so needs to be included in the theory predictions. The correction is taken from ALPGEN where at truth-level events can be tagged as coming from MPI events. The predictions are scaled to include this missing component. As the two aMC@NLO samples are constructed with different matrix elements (Z+jet for 5FNS and $Z + b\bar{b}$ for 4FNS) these corrections have been derived from the corresponding ALPGEN sample.

- *LO Predictions:* The ALPGEN and SHERPA predictions have been used to obtain LO multi-leg predictions in the 4FNS and 5FNS respectively. ALPGEN has up to 5 partons produced from the matrix element compared with SHERPA's 4. Both matrix element calculations are internally matched to a parton shower simulation and include hadronisation and MPI modeling. These predictions have no k-factors applied so are expected to underestimate the data due to missing higher-order contributions.

4.7. Results

The results for the integrated cross-sections for both data and the theoretical predictions are shown in Table 4.8 and Figure 4.22. The data measurement is quoted with the statistical uncertainty first and the systematic uncertainty second. The systematic uncertainties are combined with all downward or upward fluctuations added in quadrature. The total downward and upward shifts are consistent in the full fiducial region and are symmetrised by taking the average of the two giving the total systematic uncertainty on the measurement. Only statistical uncertainty on the predictions from ALPGEN and SHERPA are calculated as the theoretical uncertainties on these LO predictions are expected to be large. The NLO predictions having both the statistical and theoretical uncertainties quoted. The theoretical uncertainties are formed from all of the theoretical uncertainty sources (scale variations envelope uncertainty, QED FSR uncertainty and non-perturbative correction) combined in quadrature.

The theoretical predictions generally underestimate the data with the LO predictions, particularly ALPGEN and the 5FNS aMC@NLO prediction, poorly describing the inclusive cross-section. However, within uncertainties the NLO $Z + b\bar{b}$ predictions do agree with the data.

For completeness the inclusive cross-section for the $Z + b$ production is included in Figure 4.23. It seems to suggest that 5FNS predictions describe the data more accurately.

	$\sigma_{Z+b\bar{b}}$ [fb]
Data	$520 \pm 20^{+74}_{-72}$
MCFM⊗MSTW2008	$410 \pm 10^{+60}_{-60}$
MCFM⊗CT10	$386 \pm 5^{+55}_{-50}$
MCFM⊗NNPDF23	$420 \pm 10^{+70}_{-50}$
aMC@NLO 4FNS⊗MSTW2008	$485 \pm 7^{+80}_{-70}$
aMC@NLO 5FNS⊗MSTW2008	$314 \pm 9^{+30}_{-30}$
SHERPA ⊗CT10	422 ± 2
ALPGEN ⊗CTEQ6L1	317 ± 2

Table 4.8.: The integrated cross-sections for the data measurement and the theory predictions described in Section 4.6.

The differential cross-section distributions are shown in Figure 4.24. In all observables the measurements are statistically limited and the two NLO predictions agree reasonably well when considering the large uncertainties on the data. The LO predictions generally underestimate the data, with ALPGEN prediction consistently below the prediction from SHERPA. Due to the large uncertainties it is difficult to gauge the shape of the Monte Carlo predictions compared to the data distributions. However there is a possible excess of data with respect to theory at low $\Delta R(b, \bar{b})$ (and correspondingly low $m(b, \bar{b})$), which is in the region final state gluon splitting starts to dominate, so could indicate a mismodeling in the gluon splitting component of the Monte Carlo predictions.

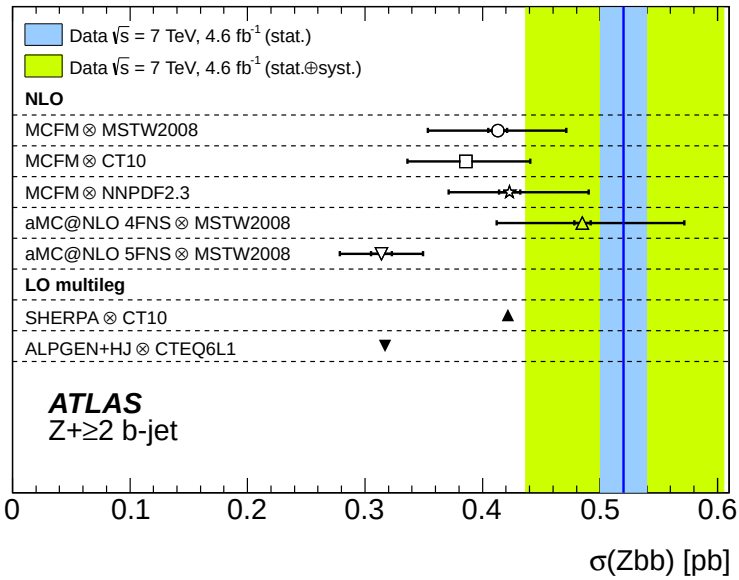


Figure 4.22.: Inclusive cross-sections for $Z + b\bar{b}$ analysis. Compares data and MC predictions [1].

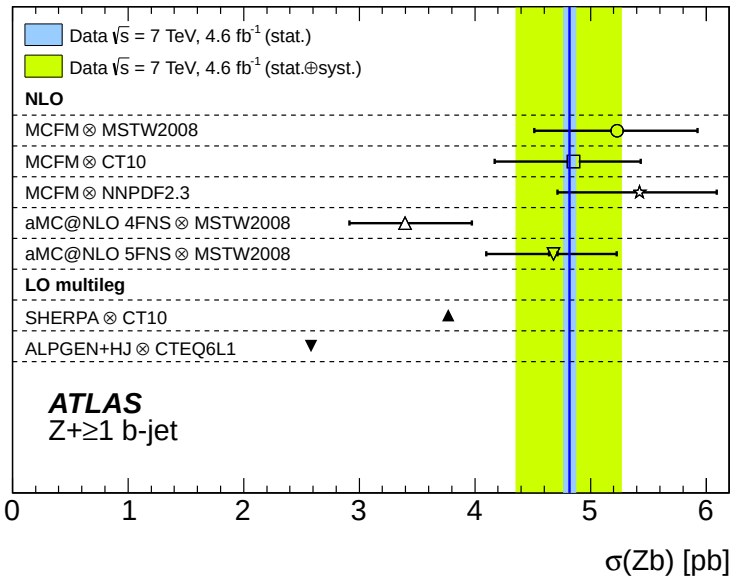


Figure 4.23.: Inclusive cross-sections for $Z + b$ analysis. Compares data and MC predictions [1].

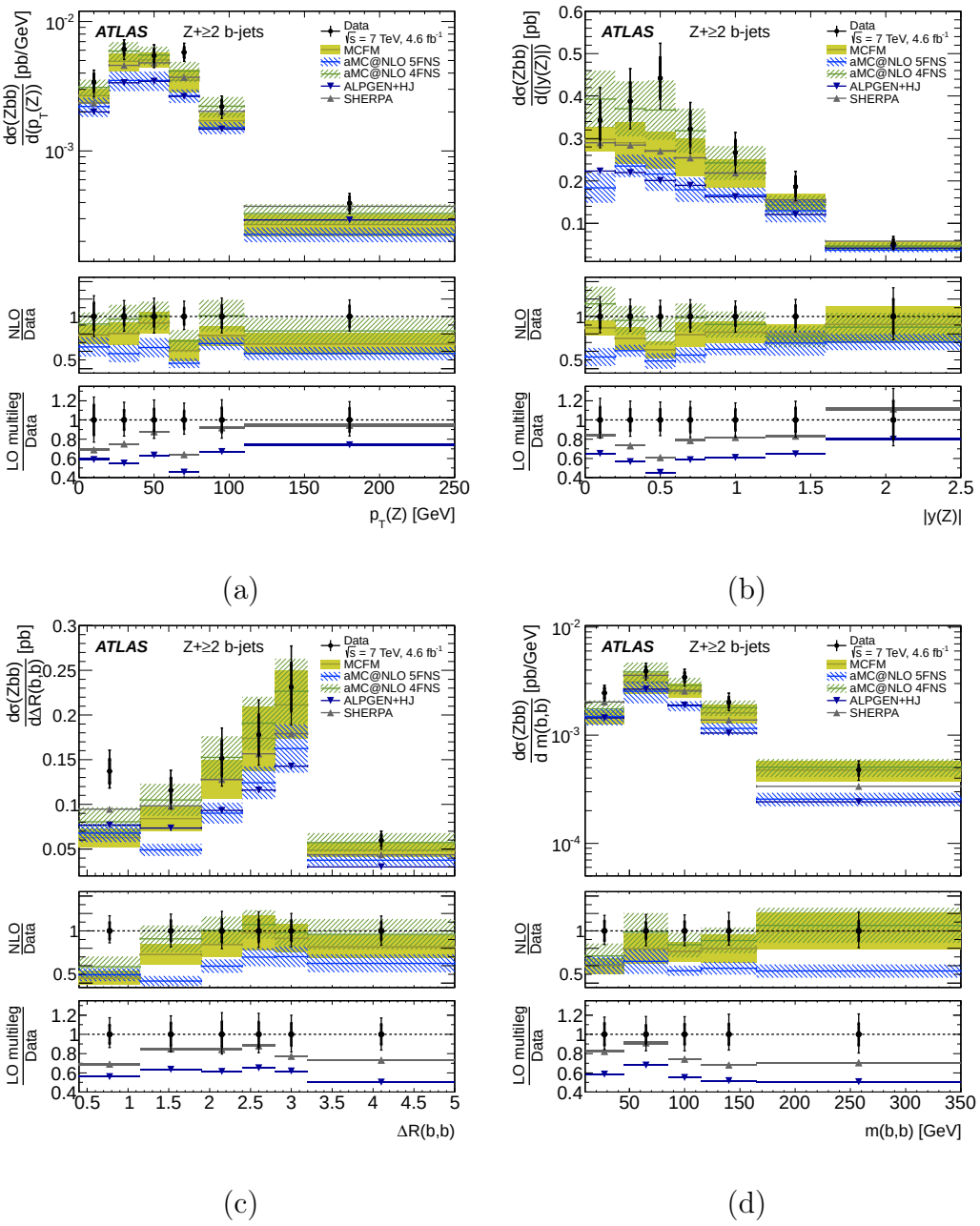


Figure 4.24.: Differential cross-sections for $Z + b\bar{b}$ production as a function of (a) $p_T(Z)$, (b) $|y(Z)|$, (c) $\Delta R(b, \bar{b})$ and (d) $m(b, \bar{b})$ [1].

4.8. Conclusion

The first ever differential $Z + b\bar{b}$ cross-section measurement was performed using the 2011 $\sqrt{s} = 7$ TeV data collected by ATLAS. The data was compared to several theoretical predictions at both LO and NLO accuracy. Within the uncertainties the NLO predictions provided by aMC@NLO and MCFM predict the normalisation and shape of the distributions, with tension at low $\Delta R(b, \bar{b})$ and $m(b, \bar{b})$. The LO multi-leg prediction from SHERPA describes the data more accurately than ALPGEN which severely underestimates the data. The 4FNS and 5FNS approaches gave consistent predictions for $Z + b\bar{b}$ final states with the data giving little discrimination between the two schemes. However, the 5FNS predictions better described the $Z + b$ final state.

Chapter 5.

$b\bar{b}$ Production via $J/\psi + \mu$ Decays

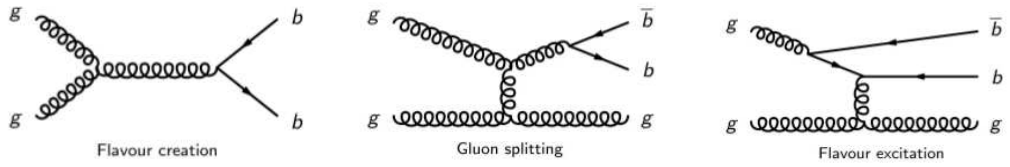
"Once more unto the breach, dear friends, once more"
— William Shakespeare

5.1. Introduction

As previously motivated, heavy flavour production is an important test of QCD. Beauty production has been measured previously at hadron colliders, starting with measurements at the $Sp\bar{p}S$ [92] and Tevatron [93, 94]. These early measurements were not described well by theoretical predictions at the time which motivated improvements in heavy flavour modeling. New beauty measurements are important to test current theories and to constrain backgrounds to other processes involving b-jets. Inclusive $b\bar{b}$ production has been analysed more recently at the LHC using b-jets to indicate the presence of B-hadrons [95].

Figure 5.1 shows some example Feynman diagrams for $b\bar{b}$ production. LO $b\bar{b}$ production from a $2 \to 2$ process is shown in Figure 5.1 (a), this is often called *flavour creation*. Examples of higher order diagrams are shown in Figures 5.1 (b) and (c), known respectively as gluon splitting and *flavour excitation*. These processes can be included in NLO calculations or from final and initial showers in LO generators matched to a parton shower simulation. Each of these diagrams can contribute to

particular regions of phase space, for example flavour creation is characterised by event topologies with the b-quarks produced back-to-back in ϕ.



(a) (b) (c)

Figure 5.1.: Example Feynman diagrams for $b\bar{b}$ production [96]

In the previous Chapter for $Z + b\bar{b}$ production there was suggestion that theoretical predictions underestimate the gluon splitting component of $b\bar{b}$ production which mainly contributes at low angular separation of the two b-jets. Due to the finite radius of jet reconstruction algorithms, analyses using jets to find the presence of B-hadrons are limited in measuring small angular separations, with the jet radius as the lower boundary. An alternative which does not have the angular resolution limitations is using physics objects that have a very high angular resolution at ATLAS, namely muons. Previous measurements at the CMS collaboration [97, 98] have used an alternative approach of inclusive vertex finding to look at the low angular separation region but have suffered from large statistical and systematic uncertainties.

B-hadrons decay to a J/ψ, a bound charmonium ($c\bar{c}$) state, approximately 1% of the time with the J/ψ decaying to a pair of opposite sign muons in $\approx 6\%$ of cases. J/ψs decaying to muons are easily identifiable at ATLAS with the di-muon pair giving a clean signature for reconstruction and triggering. The presence of a reconstructed J/ψ can therefore be evidence for a B-hadron. However this is not the only source of J/ψ mesons, which can be produced in two distinct ways at ATLAS:

1. *non-prompt* production via the decay of a B-hadron,
2. *prompt* production via a hard parton-parton interaction or feed down from higher mass charmonium states.

Experimentally these two production methods are distinguishable by the position of the J/ψ decay vertex. The J/ψ has a short lifetime (7.2×10^{-21} s) so will not travel noticeably from where it was produced. For non-prompt production the J/ψ decay

vertex will on average be displaced from the PV (assuming this is where the B-hadron was produced) due to the non-negligible B-hadron flight length, where prompt J/ψ production will result in a decay vertex at the PV. The B-hadron lifetime in the plane transverse to the beam, t, can be calculated:

$$t = \frac{L_{xy} \times m_B}{p_T^B}, \tag{5.1}$$

where L_{xy} is the distance between the PV and the B-hadron decay vertex in the transverse plane, m_B is the mass of the B-hadron and p_T^B is the transverse momentum of the B-hadron. However when considering only the muons from a $J/\psi \to \mu^+\mu^-$ decay as indication of a B-hadron, not all of the B-hadron is reconstructed and it is not p_T^B that is measured. Thus another variable based on J/ψ information is used to infer the lifetime, the *pseudo-proper lifetime*, τ:

$$\tau = \frac{L_{xy} \times m_{J/\psi}}{p_T^{J/\psi}}, \tag{5.2}$$

where L_{xy} is the distance between the primary vertex and the J/ψ decay vertex in the transverse plane projected onto the J/ψ transverse momentum direction:

$$L_{xy} = \vec{L_{xy}} \cdot \frac{\vec{p_T^{J/\psi}}}{|p_T^{J/\psi}|}, \tag{5.3}$$

$m_{J/\psi}$ is the world average measured J/ψ mass (3097 MeV [49]) and $p_T^{J/\psi}$ is the transverse momentum of the J/ψ. To remove any potential bias the two J/ψ tracks are not considered for the PV formation when calculating L_{xy}.

The presence of a non-prompt J/ψ, with positive non-zero values of τ, can be used as a proxy for a B-hadron. Due to the relatively small branching ratios for $B \to J/\psi \to \mu^+\mu^-$ decays it was not possible to reconstruct enough di-J/ψ events for meaningful analysis when investigating $b\bar{b}$ events. To increase the data statistics the semi-leptonic decay of a B-hadron, e.g. $B^+ \to \mu^+ + X$, is used as a proxy for the additional B-hadron expected in $b\bar{b}$ production. The semi-leptonic branching ratio to muons is $\approx 11\%$ for B-hadrons, raising to $\approx 20\%$ when including $B \to D + X \to \mu + X + Y$ cascade decays yielding a much larger dataset. Muons can be produced in many ways at ATLAS, but looking for *non-prompt* muons can be indicative of muons from B-hadron decays. The transverse distance of the muon

track to the beam axis at the position of closest approach to the PV, d_0, is used distinguish non-prompt muons from muons of other sources. The non-negligible B-hadron flight length results in a wider d_0 distribution for muons from a B-hadron decay. The d_0 significance, defined as d_0/σ_{d_0}, is used to reduce the tails in the d_0 distribution from poorly measured tracks. Figure 5.2 shows a shape comparison of the d_0 significance distributions for non-prompt and prompt muons taken from Monte Carlo simulation, where the wider distribution for non-prompt muons can easily be observed.

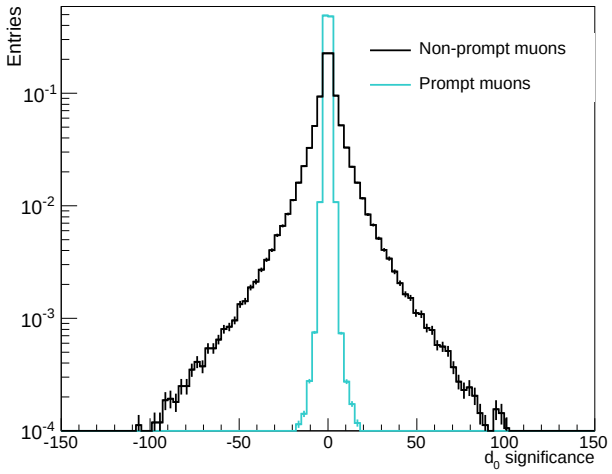


Figure 5.2.: Unit normalised histograms of d_0 significance distributions of non-prompt and prompt muons taken from Monte Carlo.

This Chapter probes QCD production of $b\bar{b}$ pairs by looking for three muon events. Two of the muons form a J/ψ candidate, with the other muon, termed *third muon*, a non-prompt muon representative of the another B-hadron. The advantage of this approach is that the low angular separation of the two B-hadrons can be probed without any lower cut-off due to intrinsic jet sizes. As well as the angular separation other interesting kinematic observables are also measured leading to differential cross-sections.

The full analysis is presented in this Chapter, starting with Section 5.2 introducing the dataset used for this measurement. The dataset is corrected for trigger and reconstruction efficiencies derived from data using methods detailed in Section 5.3. The Monte Carlo samples used to model signal and background processes in the data are detailed in Section 5.4. The goal of the analysis is to measure differential

cross-sections which are calculated from the number of signal events seen in data. The data are analysed with a series of fits to the J/ψ and third muons which are introduced in Section 5.5. First the J/ψ candidates in the dataset are fitted, as discussed in Section 5.6. Next in Section 5.7, the third muon in events is fitted with signal and background models. The results of the J/ψ and third muon fits need to be combined to find the number of signal $J/\psi + \mu$ events, which is discussed in Section 5.7.8. The reliability of the fitting procedures, for both the J/ψ and third muon fits, are tested in Section 5.8. The various sources of systematic uncertainty are motivated and assessed in Section 5.9. Finally, the differential cross-sections measured in data are compared to those from two LO Monte Carlo generators. An introduction of the theoretical predictions and the comparison with data can be seen in Section 5.11 and Section 5.12 respectively. The work presented in this Chapter is from an analysis that is yet to be published and is ongoing. The results presented are up to date at the time of writing. The state of the analysis and the planned extensions to the work presented here are documented in the conclusion in Section 5.13.

5.2. Dataset and Event Selection

The analysis uses the high luminosity runs of the 2012 data-set of p-p collisions at a centre of mass collision energy $\sqrt{s} = 8$ TeV collected from July to December 2012. All data analysed are on the GRL and has fired the trigger: EF_2MU4T_Jpsimumu_L2starB. EF_2MU4T_Jpsimumu_L2starB is a di-muon trigger requiring two L1 muons with p_{T} greater than 4 GeV. The trigger muons need to be of opposite charge and have a invariant mass consistent with the J/ψ mass: $2.5 < m_{\mu\mu} < 4.2$ GeV. A common vertex formed by the trigger muon ID tracks is also required to pass a quality cut. Data used in this analysis was only collected from July as before this (April to July) the corresponding trigger was found to have a bias on the measured lifetime of the di-muon system and could not be relied on to accurately represent the J/ψ decay lifetime. After accounting for the trigger, the GRL and pre-scales in some high instantaneous luminosity runs, the dataset has an integrated luminosity of 11.45 fb^{-1}.

Events passing the trigger must contain at least three muons. All muons are required to be *combined*, with matching tracks reconstructed in both the ID and MS, and have a p_{T} above 6 GeV and $|\eta| < 2.5$. Muon kinematic and quality

requirements are listed in Table 5.1 along with the recommended ATLAS cuts on the Inner Detector track segment. J/ψ candidates are formed from selected muons that must be oppositely charged, with tracks reconstructed to a common vertex, and pass a tighter $|\eta| < 2.3$ cut to lie within the kinematic range of the trigger. In the event of the selected muons forming more than one J/ψ candidate, where for example one muon is formed into a J/ψ candidate with two muons of opposite charge, the J/ψ candidate with an invariant mass closest to world average measured J/ψ mass is selected as the J/ψ candidate for the event. The di-muons forming the J/ψ candidate are further required to spatially match the trigger objects associated to EF_2MU4T_Jpsimumu_L2starB and to have an invariant mass in the range of the J/ψ mass (2.6-3.5 GeV). For events with more than three muons, 2% of events, the hardest muon not used in the selected J/ψ candidate is selected as the third muon.

Due to a software error in the data-file processing approximately 2% of events firing the trigger did not have the corresponding trigger objects saved to disk. Rejecting these events would require a re-calculation of the dataset luminosity. As this is unavailable it is assumed that any kinematic bias on the dataset from including the events with missing trigger objects would be negligible. Hence these events are included in the dataset.

Muon p_T	>6 GeV
Muon η	-2.5<η <2.5
Muon is combined	true
Further requirements for ID tracks forming the muon candidate	
Number of pixel hits and crossed dead pixels	>0
Number of SCT hits+number of crossed dead SCT sensors	>4
Number of pixel holes + number of SCT holes	<3
If track 0.1 <η <1.9: number of TRT hits and number of TRT outliers	>5
Fraction of TRT hits and TRT outliers that are TRT hits	>90%

Table 5.1.: Selected muon kinematic and ID track quality requirements.

5.2.1. Differential Binning

Differential cross-sections are extracted by finding event yields in bins of observables designed to probe the Monte Carlo predictions for different production mechanisms of B-hadron pairs. These are: $\Delta\phi(J/\psi, \mu)$; $\Delta R(J/\psi, \mu)$; $\Delta y(J/\psi, \mu)$; $p_T(J/\psi, \mu)$, the transverse momentum of the three-muon system; $m(J/\psi, \mu)$, the mass of three-muon system; y_{boost}, defined as the magnitude of the average rapidity of the J/ψ and third muon. The number of the bins as well as the bin dimensions for each differential observable is listed in Table 5.2.

Bin	$\Delta\phi$	ΔR	Δy	$p_T(J/\psi, \mu)[GeV]$	$m(J/\psi, \mu)[GeV]$	y_{boost}
1	0-0.1	0-0.2	0-0.1	0-10	3-8	0-0.1
2	0.1-0.25	0.2-0.4	0.1-0.2	10-15	8-15	0.1-0.2
3	0.25-0.5	0.4-0.8	0.2-0.4	15-20	15-22	0.2-0.4
4	0.5-1.0	0.8-1.3	0.4-0.7	20-25	22-25	0.4-0.7
5	1.0-2.0	1.3-2.2	0.7-1.2	25-40	25-35	0.7-1.2
6	2.0-2.3	2.2-2.7	1.2-1.9	40-60	35-50	1.2-1.7
7	2.3-2.6	2.7-3.1	1.9-2.7	60-150	50-75	1.7-2.5
8	2.6-2.9	3.1-3.8	2.7-4.8		75-200	
9	2.9-π	3.8-7.0				

Table 5.2.: $J/\psi + \mu$ analysis bins of the differential observables.

5.3. Muon Reconstruction and Trigger Efficiency Corrections

The inclusive cross-section is calculated as:

$$\sigma_{J/\psi+\mu} = \frac{N_{corr}}{L}, \tag{5.4}$$

where N_{corr} is the number of measured signal events after correction for detector inefficiencies and L is the measured luminosity. The correction from the number of signal events in the dataset to N_{corr}, uses a data-driven event-by-event weight,

which is applied to the data to account for trigger and muon reconstruction inefficiencies. The event weight can be factorised into separate trigger and muon reconstruction components which are described in the following sections. Each event is multiplicatively corrected by the the weight, w, defined as:

$$w = (\epsilon_{reco} \cdot \epsilon_{trig})^{-1}, \tag{5.5}$$

where ϵ_{reco} is the efficiency to reconstruct the three muons in the event and ϵ_{trig} is the efficiency of the trigger.

5.3.1. Trigger Efficiency Correction

To correct for any inefficiencies associated with the EF_2MU4T_Jpsimumu_L2starB trigger, a correction is applied as part of the event weight defined in Equation 5.5. The trigger efficiency correction can be further factorised:

$$\epsilon_{trig} = \epsilon^1_{ROI}(p_T, q \cdot \eta) \times \epsilon^2_{ROI}(p_T, q \cdot \eta) \times C_{\mu\mu}(\Delta R, y_{\mu,\mu}), \tag{5.6}$$

where ϵ^x_{ROI} is the efficiency of the muon trigger to identify a single ROI and $C_{\mu\mu}$ is required for the di-muon trigger, which corrects for overlapping of ROIs and vertex and opposite muon charge requirements. The ϵ^x_{ROI} efficiency depends on the kinematics of the trigger muons so is parameterised in bins of p_T and $q \cdot \eta$ and takes the form of a 2-D map in these variables. $C_{\mu\mu}$ is a multiplicative correction which is a function of the angular separation of the trigger muons.

The single muon trigger efficiency maps are formed by using a tag and probe approach on J/ψ candidate data. One of the J/ψ muons is required to fire a single muon trigger. In bins of $(p_T, q \cdot \eta)$, for the probe muon, the ratio of fitted J/ψs for events firing the single muon trigger and EF_2MU4T_Jpsimumu_L2starB, to the ratio of fitted J/ψs just firing the single muon trigger gives $\epsilon^x_{ROI}(p_T, q \cdot \eta)$. Due to the high prescale on low p_T-threshold single muon triggers during 2012 data taking, the tag muon must be matched to a single muon trigger with a threshold of at least 18 GeV. Due to this high threshold the J/ψ candidates in events passing both the EF_2MU4T_Jpsimumu_L2starB and single muon trigger tend to be at high p_T. This boosting of the J/ψ candidates is not reflective of the data collected by the EF_2MU4T_Jpsimumu_L2starB trigger as this only has a 4 GeV muon p_T threshold

for each muon ROI and can thus include events with wider opening angles of the J/ψ decay products. To un-bias the data-driven trigger efficiency map, a correction is applied based on a Monte Carlo sample of inclusive J/ψ production. A trigger efficiency map is reproduced in this Monte Carlo sample using exactly the same procedure for the data-driven map. In addition a separate map using a 4 GeV p_T threshold single muon trigger is produced. The difference between the low and high threshold maps from Monte Carlo is applied to the data driven map to correct for the map to the full kinematic range of J/ψ decay muons in the dataset. The original data driven ϵ_{ROI} map, the Monte Carlo based correction and the corrected map used for the analysis can be seen in Figure 5.3. The $C_{\mu\mu}$ term is also corrected by a Monte Carlo factor using the same procedure.

To apply the efficiency correction, a trigger weight is applied to each event in the dataset, explicitly this is the inverse of the efficiency for the trigger to fire. As taking the inverse of a variable is a non-linear transformation, care must be taken when propagating the uncertainty (which is defined symmetrically around each given efficiency map bin) on the efficiency to the uncertainty of the event yield. Considering the distribution $1/x$, the rate of change of gradient as you go to lower x values is more pronounced than for values greater than x. This is more evident the closer you get to $x = 0$, or in this case, lower values of efficiency. With the symmetric error on efficiency maps this translates to a larger contribution for downward fluctuations of the efficiency.

Each bin in the trigger efficiency map has an associated uncertainty, resulting from the statistics of the data control regions used in the derivation. A Gaussian probability density function (pdf) can be formed for each map bin, representing the range of possible efficiency values. The Gaussian pdfs are constructed using the measured trigger efficiency (for a given map bin) as the mean and associated uncertainty as the width of the Gaussian. A pseudo-efficiency map is formed by sampling randomly, in each map bin, from the Gaussian pdf. A set of 200 pseudo-efficiency maps is created by repeating this procedure 200 times. The data are corrected for trigger efficiency using each pseudo-map in turn to find the total number of events in the re-weighted dataset after applying the trigger efficiency correction. The distribution of the number of weighted events in the dataset, from using each pseudo-map, is used to assess the impact of the average efficiency correction and its uncertainty.

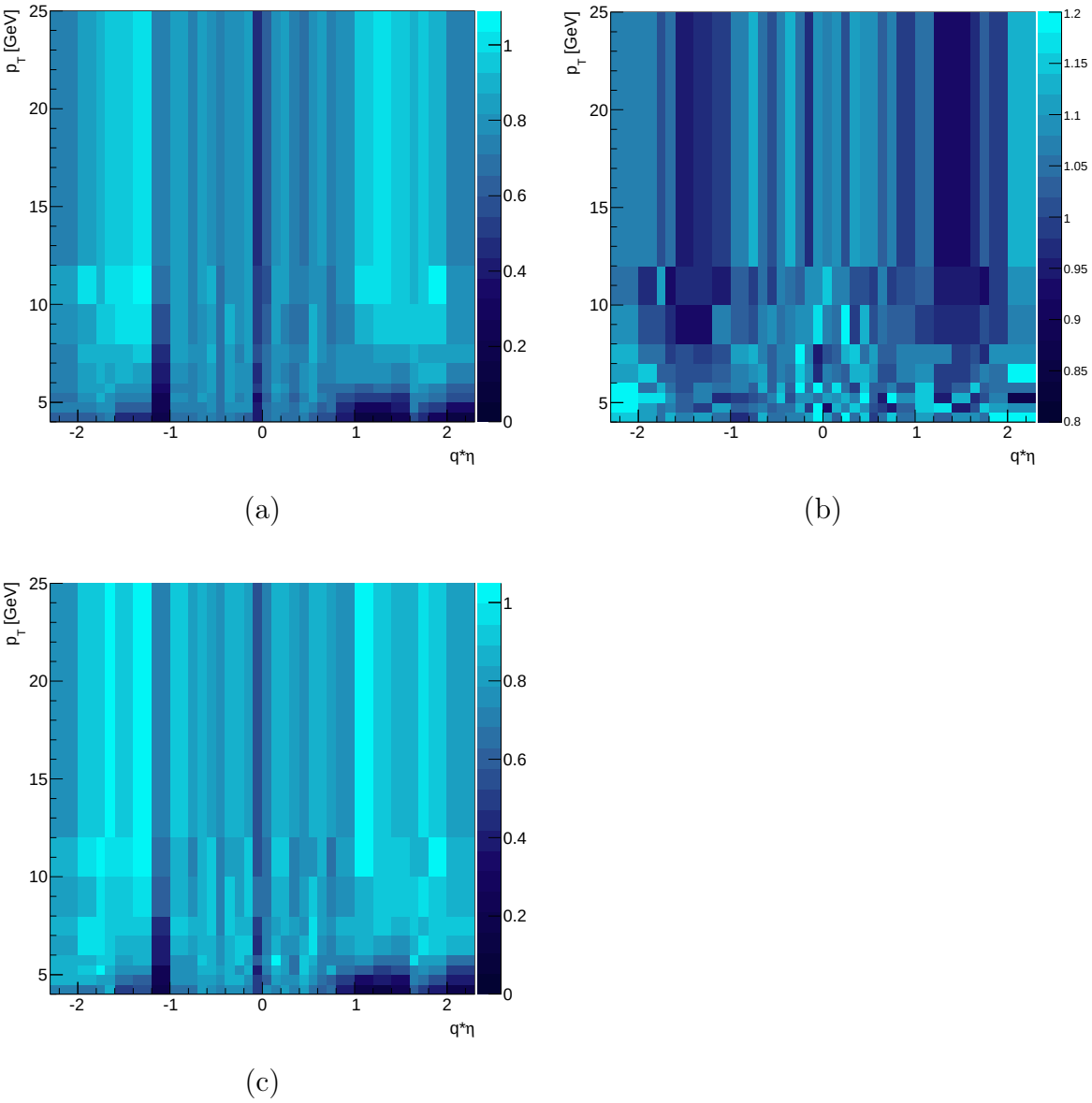


Figure 5.3.: Single muon trigger ROI maps , a) Data-Driven, b) MC based correction factors and c) Data-Driven with MC correction.

The total weighted event yield for the dataset after correction from each pseudo-map is fitted with a Gaussian function, $\mathcal{G}_{tot}$. The mean of $\mathcal{G}_{tot}$ gives the nominal weighted event yield after corrections for trigger efficiency and the width is taken as a systematic uncertainty on the trigger efficiency weighting procedure. The mean of $\mathcal{G}_{tot}$ is not equal to the the event yield using the nominal efficiency map due to the non-linear transformation from efficiency to event weight, although the difference is small (<1%). The weighted event yield after re-weighting the inclusive dataset by each pseudo-map and the fitted $\mathcal{G}_{tot}$ function can be seen in Figure 5.4, the nominal map event yield is indicated with the vertical arrow.

The analysis is performed using the nominal reconstruction map, the difference between the mean of $\mathcal{G}_{tot}$ and the result using the nominal map is applied as a correction to the event yield used to determine the cross-section. This correction is calculated and applied separately to each differential observable bin.

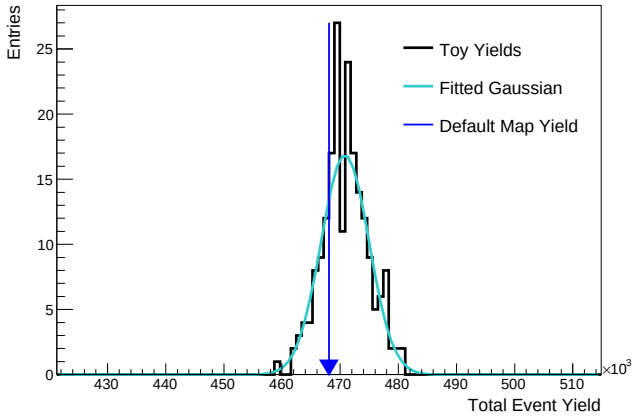


Figure 5.4.: Total weighted event yield (black histogram) when using the toy trigger efficiency maps. The distribution is fitted with a Gaussian which corrects the event yield.

5.3.2. Third Muon Trigger Efficiency Corrections

When the third muon is close to one of the trigger muons a slight drop in efficiency occurs. This is more pronounced when the muons are of opposite charge. A Monte Carlo based correction parameterised as a function of ΔR and comparative muon charge between the third muon and the closest, in ΔR, trigger muon is applied as an additional correction to account for this. The correction is derived by fitting two

functions to the ratio of signal Monte Carlo events firing the trigger to all true signal events as a function of the separation of the third muon to the closest J/ψ muon. The first function is a flat 0^{th} order polynomial and is fitted in the plateau region where the third muon is well separated from the J/ψ muons, $2 < \Delta R(\mu_{J/\psi}, \mu_{3^{rd}}) < 4$. The second function is a linear 1^{st} order polynomial which is fit in low efficiency region $0 < \Delta R(\mu_{J/\psi}, \mu_{3^{rd}}) < 0.5$. A further condition is that the linear fit has a maximum value of the y-intercept of the 0^{th} order polynomial. The data is corrected for events where $0 < \Delta R(\mu_{J/\psi}, \mu_{3^{rd}}) < 0.5$, with events outside of this range receiving no correction. Figure 5.5 shows the trigger efficiency correction as a function of ΔR between the closest J/ψ muon to the selected third muon in the event.

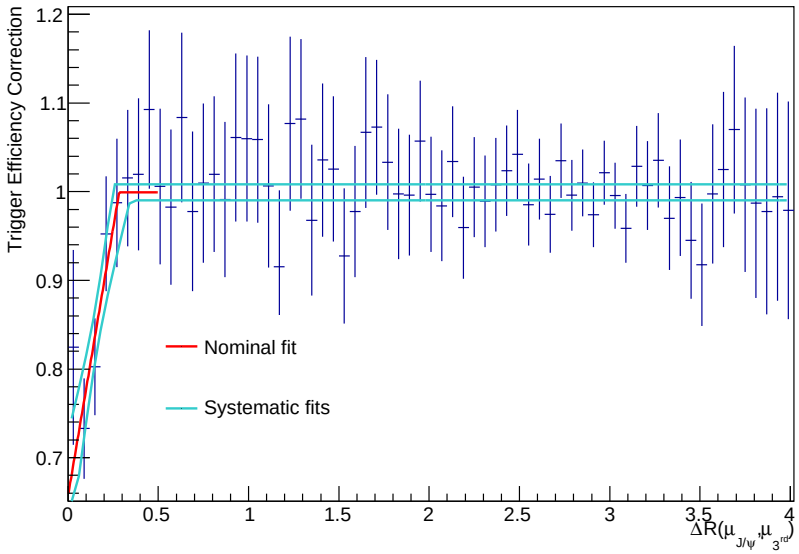


Figure 5.5.: Trigger correction for opposite sign muons as a function of ΔR between the closest J/ψ muon to the selected third muon in the event. The red function is the nominal fit to the efficiency drop with the blue functions the systematic variation used by varying the fitted parameters from the nominal fit within errors. The nominal correction is shown in the range used to correct data, $0 < \Delta R(\mu_{J/\psi}, \mu_{3^{rd}}) < 0.5$, where the systematic fits show the full fitted range.

5.3.3. Muon Reconstruction Efficiency

Data is corrected for muon reconstruction efficiency effects for each of the three muons in an event, with the total efficiency taken as the product of the three individual

efficiencies. The efficiency for a single muon is given by:

$$\epsilon_{reco} = \epsilon_{trk}(p_T, \eta) \times \epsilon_{\mu}(p_T, q \cdot \eta), \tag{5.7}$$

where q is the charge of the muon, ϵ_{trk} is the ID efficiency of track reconstruction for real muons, measured to be $99.5 \pm 0.5\%$ over the full kinematic range used in this analysis [73]. A 0.5% uncertainty on this track selection is assigned to each muon as a systematic. $\epsilon_{\mu}(p_T, q \cdot \eta)$ is the efficiency for reconstructing a muon given there is an ID track, and was assessed using a tag and probe method on $Z \to \mu\mu$ data for high p_{T} muons and $J/\psi \to \mu\mu$ data for low p_{T} muons [73]. The correction is applied using a two dimensional map parameterised in bins of muon $(p_T, q \cdot \eta)$ and can be seen in Figure 5.6.

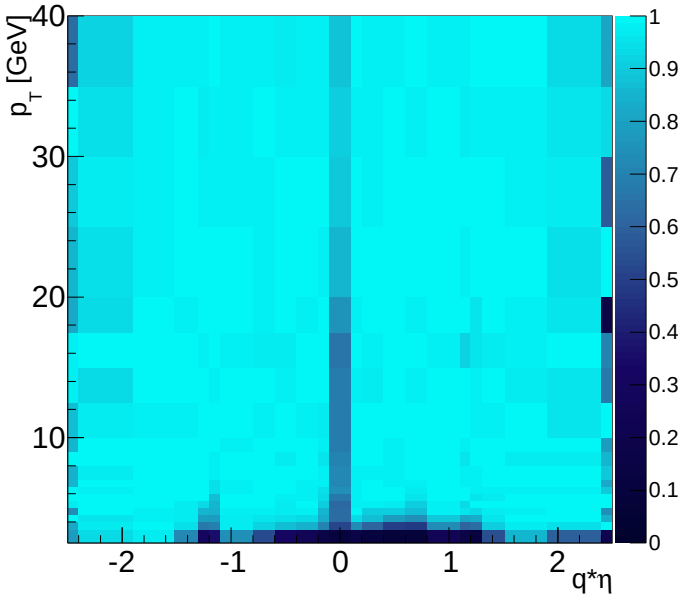


Figure 5.6.: Muon reconstruction efficiency map as a function of (p_T vs $q \cdot \eta$).

An identical procedure outlined for the trigger efficiency is used to correct the data yield for muon reconstruction inefficiencies. A set of pseudo single muon efficiency maps is produced by sampling in each bin from a Gaussian pdf, with mean taken from the nominal muon reconstruction map value and width from the error on this. The data are weighted using each pseudo map and output for all the total weighted event yields is fitted with a Gaussian function, shown in Figure 5.7. The difference between the mean of this distribution and the result using the nominal map is applied

as a correction for the event yield in each differential analysis bin, the width is taken as a systematic uncertainty on the muon reconstruction efficiency. As the muon reconstruction efficiency is generally close to unity over the kinematic range of muons in the analysis, the effect of the non-linear behavior when considering the errors on the nominal map is less pronounced here. Thus the correction taking the result using the nominal map to the mean of the toy maps is small, typically <0.2%, although it is still derived and applied for each differential observable bin.

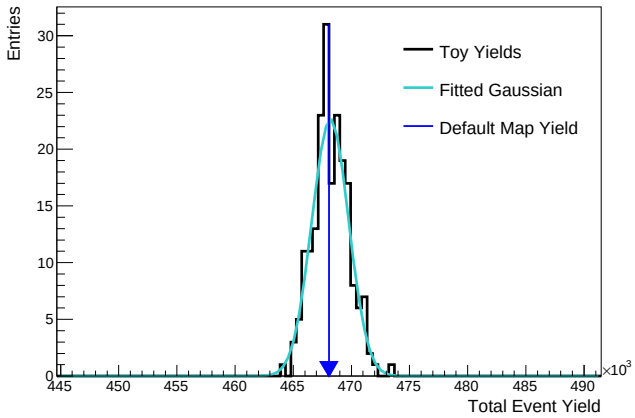


Figure 5.7.: Total weighted event yield when using the pseudo muon reconstruction efficiency maps (black histogram). The distribution is fitted with a Gaussian which corrects the event yield.

5.4. Monte Carlo Samples

A PYTHIA8 [99, 100] sample is used throughout the analysis to model the signal and backgrounds to $J/\psi + \mu$ events. The sample is for inclusive $b\bar{b}$ production, $pp \to b\bar{b}$, which is filtered for $BB \to J/\psi(\to \mu\mu) + X$ decays and was produced with PYTHIA v8.185 using the CTEQ6L1 PDF set [61]. Further event filtering required at least two muons with $p_{\mathrm{T}} > 6$ GeV in the final state. All J/ψs were forced to decay to a pair of muons, so there is an overestimate of the number of events containing four muons from two simultaneous J/ψ decays. To correct for this, events containing four muons produced from two different J/ψs are weighted by the $J/\psi \to \mu\mu$ branching fraction (5.93%).

An additional PYTHIA8 sample of inclusive prompt $pp \to J/\psi$ production is used in the analysis for prompt J/ψ and muon modeling and various cross checks. This sample was produced with the same PYTHIA8 version and again used the CTEQ6L1 PDF set.

In addition an inclusive $pp \to b\bar{b}$ sample was produced using HERWIG++ v2.7.1 [91]. The prediction used the CTEQ6L1 PDF set, with UE-EE5 [101] tune of the underlying event parameters. HERWIG++ uses angular ordered parton showers and the cluster hadronisation model. This prediction was not put through the detector simulation and was used as alternative model for backgrounds in data which had been corrected for detector effects.

5.4.1. Monte Carlo Corrections

A correction for the muon reconstruction efficiency is applied to muons in the PYTHIA8 Monte Carlo samples so that they are consistent with the re-weighted dataset. The correction was parameterised in the same $(p_T, q \cdot \eta)$ bins as the correction in data. It is calculated as the ratio of number of reconstructed muons to truth muons, for reconstructed muons that have been matched spatially to a truth muon, $\Delta R(\mu^{truth}, \mu^{reco}) < 0.02$, and is derived only for reconstructed muons passing the muon selection criteria outlined in Table 5.1. The efficiency map is shown in Figure 5.8. An event-by-event re-weighting corrects for each muon reconstructed in the Monte Carlo samples.

5.5. Signal Extraction

To extract the number of signal events in the dataset after event selection has been applied, a series of two simultaneous maximum likelihood fits are performed on the J/ψ candidate and the selected third muon in the event. The physics of prompt J/ψ production is quite different to that of non-prompt J/ψ production. For prompt J/ψ production the third muon can come from many sources, resulting in an admixture of third muon sources: prompt muons, non-prompt muons and hadrons being reconstructed as muons (*fake muons*, detailed in Section 5.7.1). Whereas third muons produced in association with a non-prompt J/ψ contain a much purer

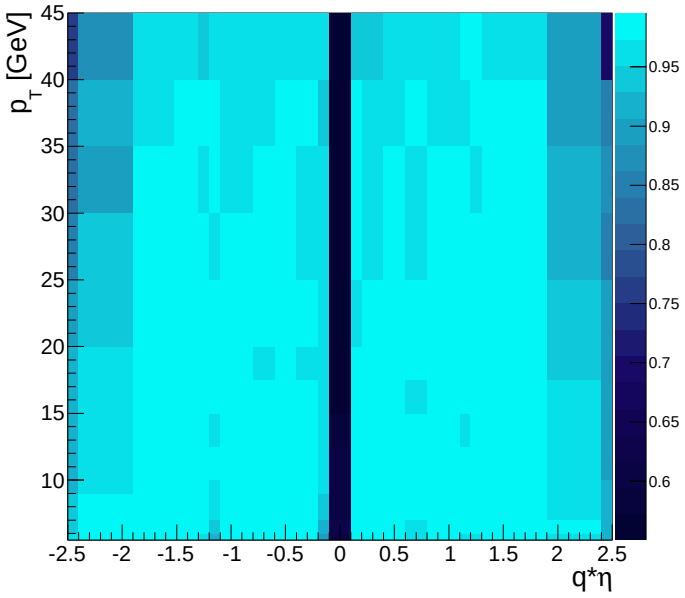


Figure 5.8.: 2-D Monte Carlo muon reconstruction efficiency map parametrised in bins of $(p_T, q \cdot \eta)$.

sample of non-prompt muons. To allow for a cleaner environment the third muon fits are performed separately from the J/ψ fits for events with the J/ψ candidate having $\tau > 0.25$ ps, which is in the non-prompt J/ψ regime. The number of signal events resulting in non-prompt J/ψs and non-prompt third muons is extracted by combining the results of the two independent fits. The following sections will outline the fits to the J/ψ and third muon system and the resulting combination.

5.6. J/ψ Fitting

After triggering the dataset has a high purity of real J/ψs. There does however remain a residual combinatoric di-muon background contribution forming fake J/ψ candidates. To extract the signal component of real non-prompt J/ψs two variables are used:

- *Di-muon invariant mass*: used to distinguish between real J/ψs and combinatoric muon backgrounds. The J/ψ has a narrow width (~ 93 keV) which

allows easy discrimination between peaked signal and a continuum background, referred to as the *fake* J/ψ background.

- *Pseudo-proper lifetime,* τ: was previously defined in Section 5.1, allows the discrimination of prompt and non-prompt J/ψs.

In order to extract the number of non-prompt J/ψs in the data, a 2-D unbinned maximum likelihood fit of these two variables is performed. The fake J/ψ background is split into three components which have different τ distributions. With prompt and non-prompt J/ψs there is in total five individual components for the fit, each with a distribution for di-muon mass and τ. The following five functional fit models are used to describe the data:

1. *Non-prompt* J/ψ*:* The di-muon invariant mass is modeled using a combined Crystal Ball (CB_1) [102] and Gaussian function (G). The τ distribution is modeled using a single-sided exponential (E), convolved with a detector resolution function.

2. *Prompt* J/ψ*:* The di-muon invariant mass is modeled using the same Crystal Ball and Gaussian function used for the non-prompt J/ψ. The τ distribution is modeled using a delta function at $\tau = 0$ (δ), convoluted with a detector resolution function.

3. *Prompt fake* J/ψ *background:* The di-muon invariant mass for this background is modeled using a 0^{th} order polynomial distribution ($P^{(0)}$). The τ distribution is modeled using a delta function at zero (δ), convoluted with a detector resolution function.

4. *Single-sided fake* J/ψ *background:* The di-muon invariant mass distribution is modeled by an exponential function (E). The τ distribution is modeled using a single-sided exponential function (E) convoluted with the detector resolution function.

5. *Double-sided fake* J/ψ *background:* The di-muon invariant mass distribution is modeled by an exponential (E). The τ distribution is modeled using a double-sided exponential decay ($E(|\tau|)$) convoluted with the detector resolution function.

The resolution function is common to all models an is constructed as the sum of two Gaussian functions ($G_1^r(\tau) \oplus G_2^r(\tau)$). It is used to smear the τ distributions to mimic the effect of the finite measurement resolution of the detector. The total fit model, $F(m_{\mu\mu}, \tau)$, is summarised by Equation 5.8:

$$F(m_{\mu\mu}, \tau) = \sum_{i=1}^{5} f_i(m) * (h_i(\tau) \otimes g(\tau)), \tag{5.8}$$

where $\otimes$ represents the convolution of two functions. The individual di-muon invariant mass and τ functions are summarised in Table 5.3, where $\oplus$ represents the normalised weighted average.

Component	$f_i(m)$	$h_i(\tau)$	$g(\tau)$
Non-Prompt J/ψ	$CB_1 \oplus G$	E	$G_1^r(\tau) \oplus G_2^r(\tau)$
prompt J/ψ	$CB_1 \oplus G$	δ	$G_1^r(\tau) \oplus G_2^r(\tau)$
Prompt fake J/ψ	$P^{(0)}$	δ	$G_1^r(\tau) \oplus G_2^r(\tau)$
Single-sided fake J/ψ	E	E	$G_1^r(\tau) \oplus G_2^r(\tau)$
Double-sided fake J/ψ	E	$E(\lvert\tau\rvert)$	$G_1^r(\tau) \oplus G_2^r(\tau)$

Table 5.3.: J/ψ fit component models given by Equation 5.8, each function is defined in the text.

For the real J/ψ components the parameters determining the mean position and shape of the J/ψ mass peak, i.e. the relative fractions of the Crystal Ball and Gaussian functions, are the same. This leaves a total of 19 free fit parameters. The most important parameters used for the analysis are the five component normalisations (including the uncertainties taken from the covariance matrix) and the slope of the non-prompt J/ψ τ exponential function.

5.6.1. J/ψ Fit Results

The same composite di-muon model is fit to data in each differential observable bin. Figure 5.9 shows examples of the 2-D J/ψ fits for a bin taken from random differential variable bins, as defined in Table 5.2, with each of the five post-fit model components included, as well as the full model. The normalisation of the non-prompt J/ψ component, taken from the fitted data, is used in combination with the results

of the third muon fits to calculate signal event yields, N_{corr}, in Section 5.7.8. The normalisation of the fake J/ψ components is used to define backgrounds for the third muon fits in Section 5.7.2.

5.7. Third Muon Fitting

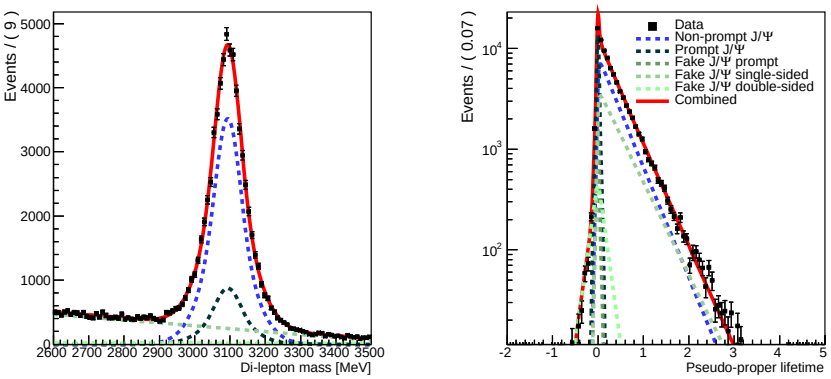
With the non-prompt J/ψ contribution fully determined, attention is focused on extracting the yield of non-prompt third muons in these events. Another 2-D maximum likelihood fit is performed using variables attributed to the third muon. A fit to a boosted decision tree (BDT) output, trained to separate muons from fake muons, and d_0 significance used for non-prompt muon extraction.

The dataset comprises third muons composed of the following components:

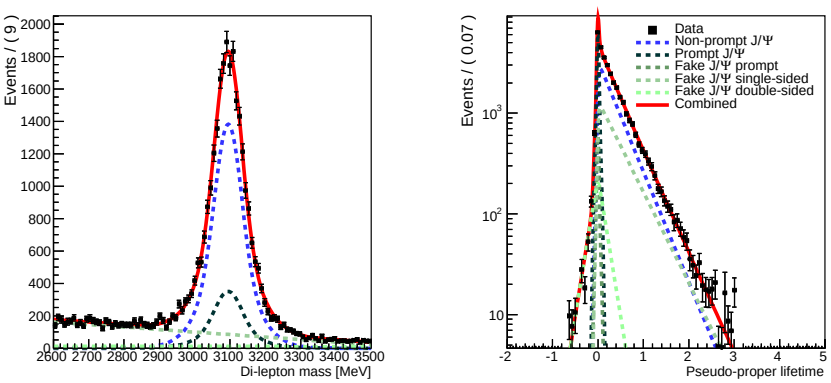
- non-prompt muons (the signal), from semi-leptonic decays of B-hadrons or cascade $B \to D \to \mu + X$ decays,
- prompt muons, produced at the PV,
- fake muons, fully described in Section 5.7.1.

In addition the third muon could be produced in background events containing a fake J/ψ, see Section 5.7.2, or *pile-up* events, where the J/ψ and third muon produced in different p-p interactions (see Section 5.7.3). This results in a model of third muons composed of five individual components. Each of these are modeled by templates taken from either Monte Carlo (non-prompt, prompt and fake muon) or data-driven techniques (fake J/ψ and pile-up) which are used to fit the data.

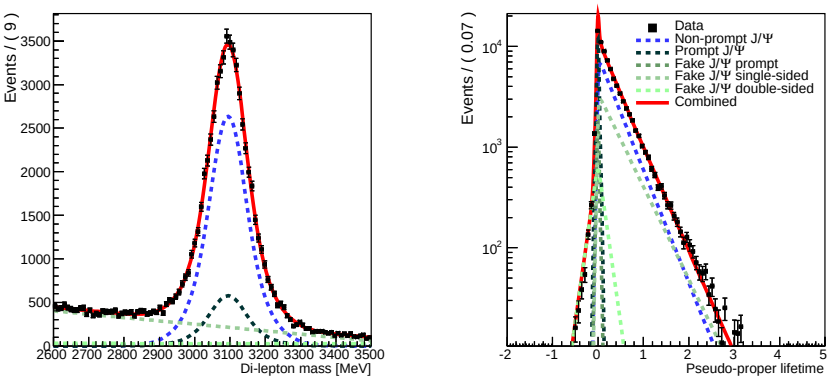
To increase non-prompt third muon purity only events with a non-prompt J/ψ candidate, $J/\psi\ \tau > 0.25$, are used for the third muon fits. The data is significantly purer after the cut, however the cut also rejects a proportion of non-prompt J/ψ events, an efficiency loss which must be corrected for. This is achieved by extrapolating the fit result to the full $J/\psi\ \tau$ spectrum and is described in Section 5.7.8. The τ distribution for non-prompt J/ψs taken from Monte Carlo is shown in Figure 5.10, J/ψs above the $\tau > 0.25$ cut contribute to $\approx 52\%$ of the J/ψs in Monte Carlo. In addition a cut is placed on the di-muon mass, $2.95 < m_{\mu^+\mu^-} < 3.25$ GeV, to reduce the background component of fake J/ψ events. This cut was found to contain $\approx 98\%$



(a)



(b)



(c)

Figure 5.9.: Example 2-D simultaneous J/ψ fit results, for differential observable bins: a) $10<p_T(J/\psi,\mu)<15$ GeV, b) $8<m(J/\psi,\mu)<15$ GeV, c) $1.2<y_{boost}<1.7$.

of real J/ψ events. The inefficiency of this cut is corrected for once the signal yield is extracted from the fits.

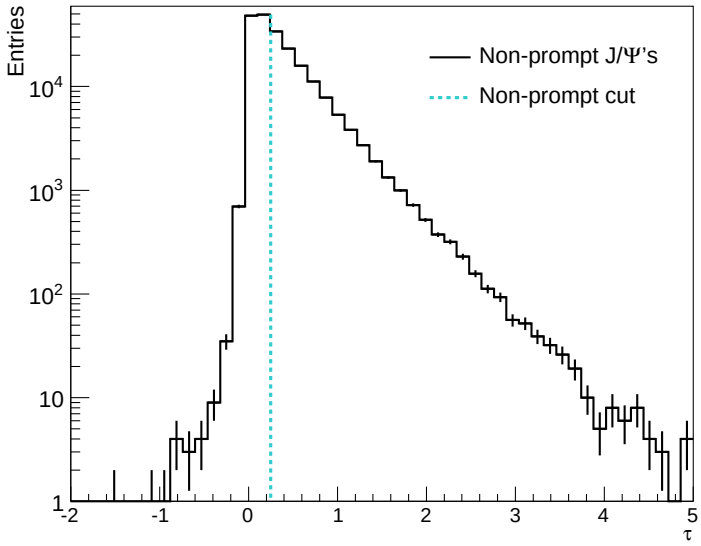


Figure 5.10.: Monte Carlo pseudo-proper lifetime distribution of non-prompt J/ψs. The figure also includes the cut on pseudo-proper lifetime defining the non-prompt regime used for third muon fits.

5.7.1. Fake Muon Background

One of the most challenging backgrounds is from charged pions or kaons decaying to muons while traveling through the detector, e.g. $\pi/K \to \mu\nu$. The charged pions and kaons leave a track in the ID with the daughter muon leaving a track in the MS; this is known as decay-in-flight (DIF). Due to the small mass difference between pions, kaons and muons often the angular deflection is small so the muon continues in the direction of the hadron. The muon reconstruction algorithms then combine the resultant tracks and form a muon candidate. Decays such as $B^+ \to J/\psi + K^+$, result in a peak in this background at low angular separation between the J/ψ and third muon. Unfortunately this is a region of particular interest so it is important to have an accurate measurement of these DIFs so they do not contribute to the cross-section calculations.

Another background prevalent to the analysis are muon fakes from leakage of charged hadrons into the MS. Charged hadrons leave a track in the ID and shower in the HCAL. If charged particles from the shower leak through to the MS this can leave tracks in the MS, which are combined into muon candidates with the ID track. This background is also peaked at low angular separation and can again be caused by decays such as $B^+ \to J/\psi + K^+$.

To discriminate between signal muons and muons coming from a DIF or hadronic shower leakage, henceforth collectively termed *fake muons*, a BDT is used. From the PYTHIA Monte Carlo simulation it was found that the fake muon candidates were approximately 75% DIF muons and 25% hadronic shower leakage muons. A BDT is a method used for event classification. Given a muon candidate, it assigns a weight based on the muon properties. Signal objects, in this case real muons, have on average higher BDT output weights. Using a set of variables associated with the muon candidate a BDT is trained to create a set of binary splits of the data based on the input variables. Each variable is considered one at a time with a yes/no (signal/background) decision taken for each until a stop criterion based on signal purity is satisfied. To stabilise the BDT a process termed *boosting* is applied where a set of multiple trees is used with the classification based on the majority decision of the set of trees.

The BDT is trained on a set of signal (real muons) and background (fake muons) taken from Monte Carlo, the PYTHIA8 Monte Carlo sample described in Section 5.4 is used to source both. Signal muons are defined as reconstructed muons matched to a muon from the truth muon collection, $\Delta R(\mu_{true}, \mu_{reco}) < 0.02$, where the truth muon comes from a semi-leptonic decay of a B-hadron. Background muons are defined as reconstructed muons matched to a truth muon that has a charged pion or kaon as a parent, $\Delta R(\mu_{true}, \mu_{reco}) < 0.15$ for DIF, or is matched to a charged pion or kaon if there is no matching truth muon for the hadronic shower leakage, $\Delta R(\pi/K_{true}, \mu_{reco}) < 0.15$. All reconstructed muons are required to pass the muon quality cuts used in the analysis detailed in Table 5.1. The BDT used is developed in the TMVA framework [103], which splits Monte Carlo into two independent samples. One is used for training the BDT and the other is used for testing the performance.

The BDT uses four input variables, each with some discriminating power between real and fake muons:

- *Momentum balance significance* - The difference between the ID track transverse momentum and the MS track transverse momentum. The MS track momentum has been extrapolated back to the ID, to account for energy loss in the calorimeter systems. This variable should discriminate when a fake muon candidate decays outside of the ID and inside the calorimeter, with higher values for fake candidates.

- *Scattering significance* - Maximum difference in track curvature as computed upstream or downstream of an ID detection plane. DIF muons decaying in the ID tend to have higher values of scatter significance.

- *Scattering neighbouring significance* - Track segment angles $(\delta\phi, \delta\theta)$ either side of a ID scattering centre compared to look for kinks. DIF muons decaying in the ID again populate higher values.

- *Absolute* η - Fake muons tend to be produced more centrally.

The BDT incorporates each of these variables and gives a much larger discriminating power than any of the individual variable alone. The input variables are plotted for both real and fake muons in Figure 5.11. The choice of variables means there is potential discrimination of fake muons in both the inner-detector and calorimeter systems. The number of input variables is kept relatively small to minimise any potential mis-modeling of the BDT performance between Monte Carlo and data. This does not affect the performance of the BDT as the goal is not to maximise the separation between real and fake muons but to have enough separation to use the BDT output distributions, for real and fake muons, as fit templates that accurately describe the data. The inputs were thus chosen as the Monte Carlo accurately describes data for each of the four variables selected. The BDT output for both real and fake muons is shown in Figure 5.12.

5.7.2. Fake J/ψ Background

The backgrounds from fake J/ψ events, where the third muon is not produced in association with a real J/ψ, are removed in a two step process. Firstly taking the full di-muon mass range of 2.60-3.50 GeV, a tighter window around the peak, 2.95-3.25 GeV, is used to define the signal region. Then residual background events underneath the mass peak are removed using a side-band subtraction method.

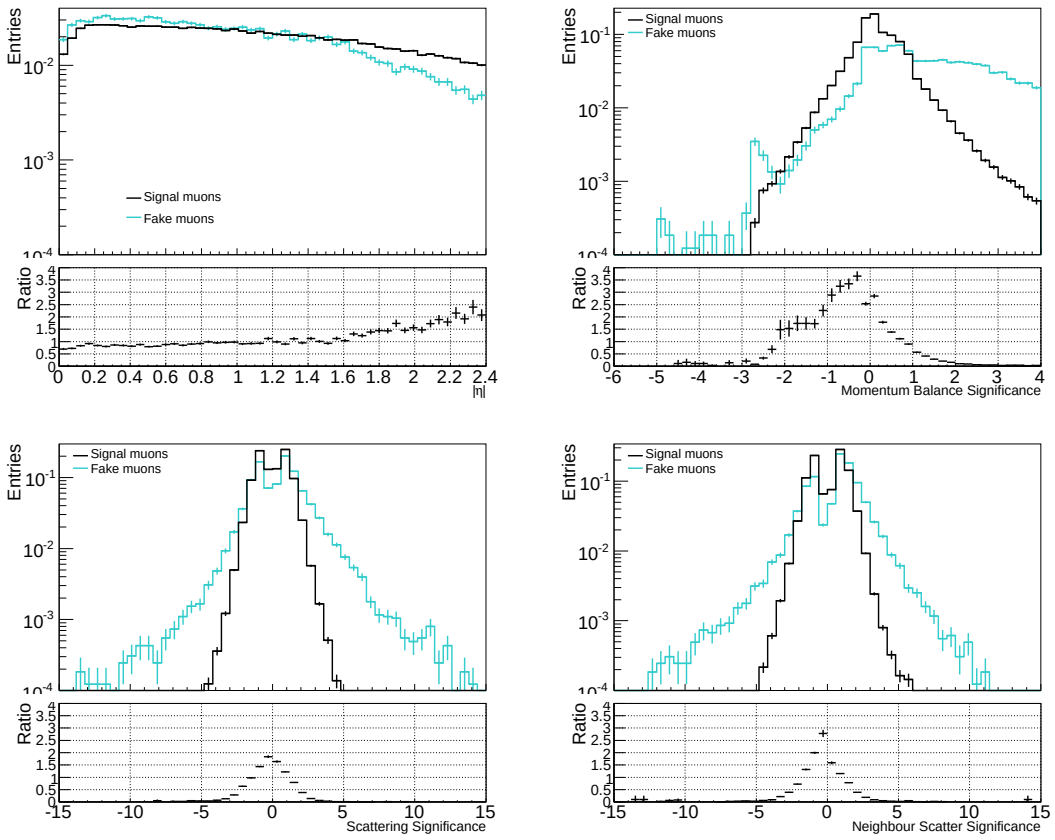


Figure 5.11.: Unit normalised BDT training input distributions for signal muons (black) and background fake muons (blue).

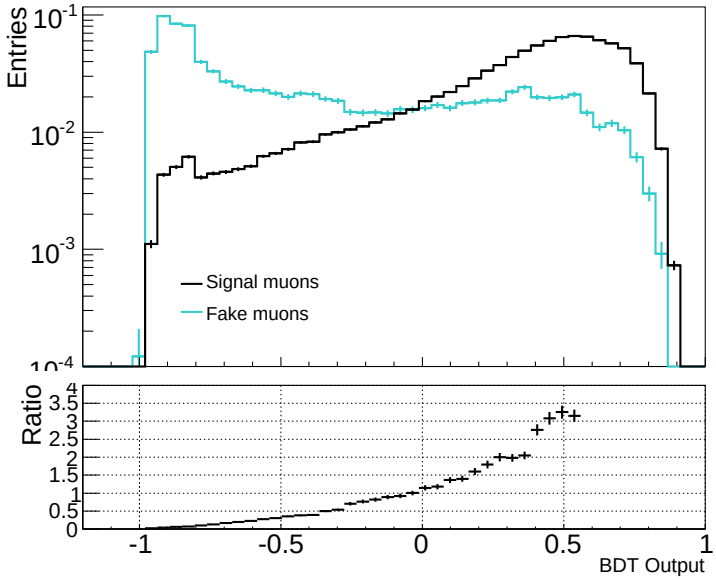
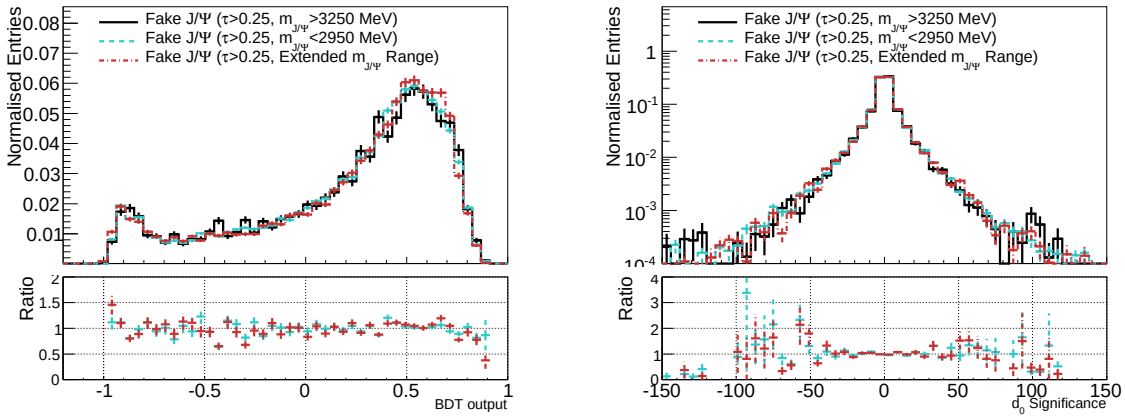


Figure 5.12.: Shape of the BDT output distribution for signal muons (black) and background fake muons (blue).

To model the BDT and d_0 significance distributions of third muons in fake J/ψ events, the di-muon mass side-bands are used along with the result of the 2-D J/ψ mass-τ fit. BDT and d_0 significance templates from data in the mass side-band regions, $2.60 < m_{\mu^+\mu^-} < 2.95$ GeV and $3.25 < m_{\mu^+\mu^-} < 3.50$ GeV, are formed to define the shape of third muon fit distributions. The normalisation of the templates is taken from the fitted 2-D J/ψ model, where the three fake J/ψ background models are integrated over the phase space used for the third muon fits ($2.95 < m_{\mu^+\mu^-} < 3.25$ GeV, $\tau > 0.25$ ps). The fake J/ψ background is fully constrained from this method so the templates representing this background are kept fixed when fitting the third muon distributions in data.

It was assumed that the three fake J/ψ background components have the same third muon BDT output and d_0 significance distributions. To check this data in the upper and lower mass side-bands are compared. Due to the different di-muon mass parameterisations for each of the fake J/ψ backgrounds the relative contribution of each is different in the two side-band regions. The BDT and d_0 significance distributions for each are compatible as shown in Figure 5.13 showing that there is no di-muon τ dependence to the shape of the third muon fit variables in fake J/ψ events. In addition data in an extended mass range, $2.50 < m_{\mu^+\mu^-} < 2.6$ GeV and

$3.5 < m_{\mu^+\mu^-} < 4.30$ GeV, was used as an additional cross-check. Extensions to even lower or higher masses is excluded as the EF_2MU4T_Jpsimumu_L2starB trigger has a mass cut included on the di-muon system ($2.50 < m_{\mu^+\mu^-} < 4.3$ GeV). The shape of this extended mass range for both the BDT and d_0 significance is included in Figure 5.13. No shape dependence is observed for any di-muon mass range.



(a) (b)

Figure 5.13.: Unit normalised histograms of third muon fit variable distributions in high, low and extended J/ψ mass side-bands which are described in the text for, a) BDT output, b) d_0 significance. The lower plot includes the ratio to the high mass sideband ($3.25 < m_{\mu^+\mu^-} < 3.50$ GeV).

5.7.3. Pile-up Background

Due to the large number of interactions per bunch crossing, see Figure 3.6, another background comes from pile-up. This background is defined where the J/ψ and third muon are produced from different hard scatters in the same bunch crossing. In order to fully determine this background, the difference in the reconstructed z-position, Δz_0:

$$\Delta z_0 = z_0^{\mu_{J/\psi}} - z_0^{\mu}, \tag{5.9}$$

is used, where $z_0^{\mu_{J/\psi}}$ is the z-position, at point of closest approach to the beam axis, of a muon used in the reconstructed J/ψ, and z_0^{μ} is the z-position taken from the position of the third muon track at the point of closest approach to the beam axis.

The J/ψ muon that maximises Δz_0 is used on an event by event basis. To remove the majority of the pile-up events a cut is placed at $|\Delta z_0| < 40$ mm, with only events passing this cut used for fitting the third muon distributions. The residual pile-up background that passes this cut is removed using a side-band subtraction method.

Figure 5.14 shows the Δz_0 distribution for data after all event selection cuts. The distribution consists of two components, a peaked structure centred on zero represents events where the J/ψ candidate and third muon are produced in the same p-p interaction and a Gaussian distributed pileup background spanning across a wide Δz_0 range. The $|\Delta z_0| < 40$ mm cut was chosen as it removes most of the pile-up background and contains all of the signal events.

To remove the pile-up events within the $|\Delta z_0| < 40$ mm signal region the events outside this cut are fit with a Gaussian model. The nomalisation of the Gaussian within $|\Delta z_0| < 40$ mm cut gives the number of residual pile-up events in the signal region. The shape of third muon BDT and d_0 distributions for pile-up events is taken from the purely pile-up region outside the cut. With the normalisation from under the signal peak and the shape from outside the peak third muon distributions for pile-up events is fully constrained. Figure 5.14 includes the Gaussian fit to the Δz_0 distribution for the inclusive dataset.

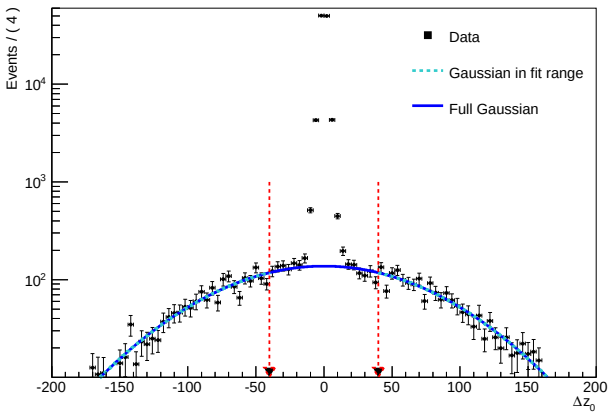
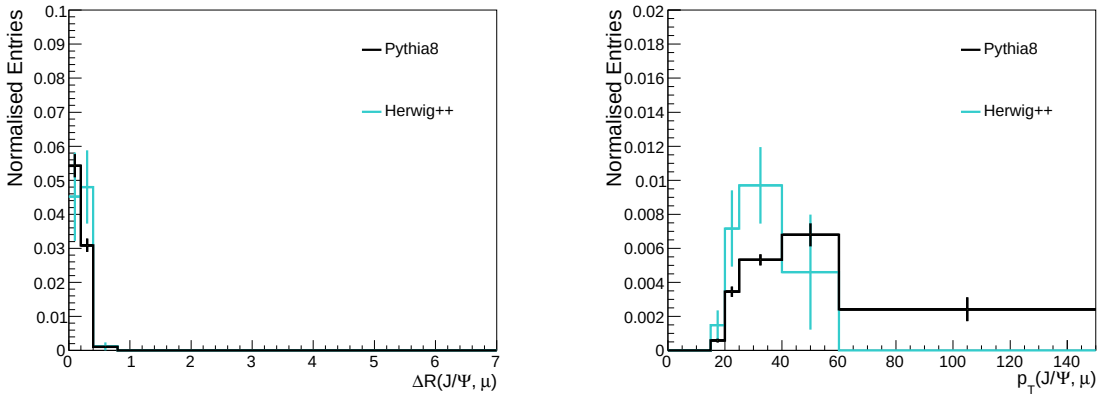


Figure 5.14.: Data Δz_0 distribution for the inclusive dataset including the Gaussian pile-up background fit. The vertical arrows indicate the boundaries of the signal and background regions.

.

5.7.4. B_c-meson Background

$B_c \to J/\psi + \mu + X$ decays are an irreducible background to this analysis. Although a small background when considered inclusively, the J/ψ and muon from a B_c decay are produced close together, in the low angular separation region which is of particular interest, so needs to be fully accounted for. A Monte Carlo prediction from the PYTHIA8 sample for the B_c contribution passing event selection is calculated in each differential observable bin. An alternative prediction from HERWIG++ is also used to decrease the reliance of the modeling in PYTHIA8. The average of the two predictions is subtracted from the fitted signal yield to remove the background from B_c decays. As detailed in Section 5.9 the difference between the two predictions is used as a systematic uncertainty. Figure 5.15 shows the B_c distribution from PYTHIA8 and HERWIG ++ as a function of $\Delta R(J/\psi, \mu)$ and $p_T(J/\psi, \mu)$, where the histograms have been normalised by the total signal yield.



(a) (b)

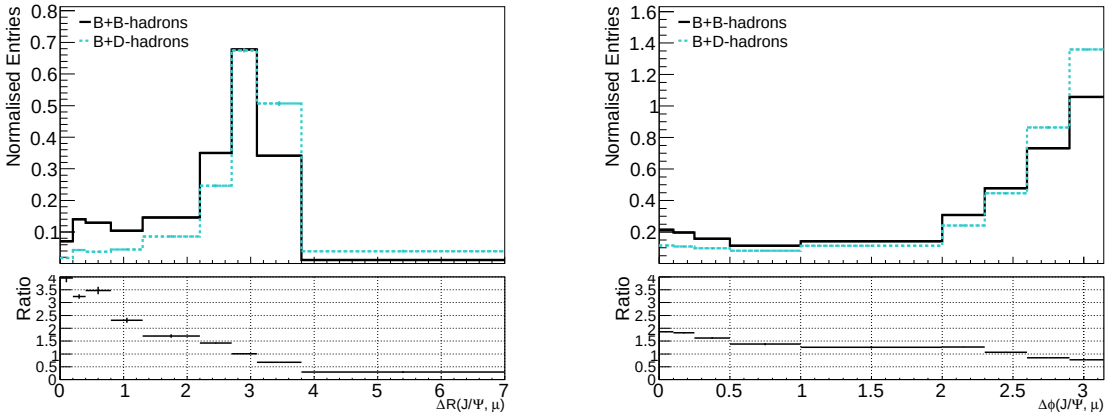
Figure 5.15.: B_c-meson contribution to signal events from PYTHIA8 and HERWIG++. The amount of B_c events (as a % of signal events) is shown for the differential distributions: a) $\Delta R(J/\psi, \mu)$ and b) $p_T(J/\psi, \mu)$.

5.7.5. B+D-hadron Background

In PYTHIA8 Monte Carlo studies, it was found that approximately 4% of third muons originated from the decays of D-hadrons which were not from a B-hadron cascade decay. D-hadrons, similarly to B-hadrons, have a relatively long lifetime,

1.04×10^{-12} s for $D^{\pm}$-mesons. Experimentally this means it is difficult to distinguish third muons from D and B-hadron decays, as both will have a wider d_0 significance distribution indicating production away from the beam line. Due to the similarity of the d_0 significance distributions it is not possible to disentangle these two sources of non-prompt muons from fitting to data.

A Monte Carlo correction removes the expected B+D-hadron contribution from the fitted number of signal events. The correction is derived from an average estimate of this background taken from the PYTHIA8 and HERWIG++ Monte Carlo samples. Figure 5.16 shows the distribution of B+D-hadron and B+B-hadron events passing the event selection for two of the kinematic distributions in the PYTHIA8. $J/\psi + \mu$ events from $B + D$-hadrons have different correlations between the J/ψ and third muon compared to double-B hadron events, so the B+D event background is determined separately for each differential observable bin. The $B + D$ contribution is removed from the signal yield after the third muon fits by subtracting the fraction of B+D events compared to B+B events using the average Monte Carlo prediction.



(a) (b)

Figure 5.16.: Unit normalised muon level distributions comparing Monte Carlo events from two B-hadrons to events with a B and D hadron, a) $\Delta R(J/\psi, \mu)$ and b) $\Delta\phi(J/\psi, \mu)$.

5.7.6. Punch-Through Background

The final background considered is punch-through. Punch-through here is defined as a charged pion or kaon having sufficient p_{T} to traverse the detector to the MS without interacting with the detector material or decaying. This background will look like real muons as you have a well defined track in the ID and a well matching track in the MS. The punch-through estimate is taken from Monte Carlo, where a reconstructed muon has no matching truth muon but is matched to a charged kaon or pion which has no decay vertex inside the detector. The distribution of punch-through is similar to the main fake muon background as it is again a background of pions and kaons, although there are higher contributions in regions of the detector with less material in the particle flight path. The ratio of the number of punch-through to fake muons is taken from PYTHIA8 ($\approx 1\%$) and is used to normalise the background from the number of fake muons fitted in data. As punch-through is such a small background any mis-modeling of this by the simulation, most likely coming from the modeling of material in the ATLAS detector, is unlikely to have any effect on the final results. For example at low angular separation between the third muon and J/ψ, where there is the highest density of fake muon and punch-through backgrounds from decays such as $B^+ \to J/\psi + K^+$, the estimated punch-through is less than 0.5% of the signal. The third muon yield in each differential bin is corrected, post fitting, by removing the expected number of punch-through events.

5.7.7. Third Muon Fits

The third muon BDT and d_0 significance distributions are fitted simultaneously to determine the non-prompt real muon component. The fit is performed on a subset of the data passing the event selection, in a region with less backgrounds. The phase space cuts have been documented in the preceding sections but are summarised here. Events must be in the non-prompt J/ψ region ($\tau > 0.25$ ps), the J/ψ candidate invariant mass must be within the J/ψ mass peak, $2.95 < m_{J/\psi} < 3.25$ GeV and finally events must be in the signal Δz_0 region, $|\Delta z_0| < 40$ mm. The fit is an extended maximum log-likelihood fit, fitting data events using templates derived for each expected third muon component. The following list details the five fitted third muon components as well as the source of the templates used for each.

- Non-prompt: The signal component. BDT and d_0 significance fit templates are taken from the PYTHIA8 Monte Carlo sample from reconstructed muons matched to a truth muon which derives from a B-hadron. One expects this component to populate the high values of the BDT output signifying real muons and have a wide d_0 significance distribution indicating production away from the interaction point. The shape of the templates are fixed but the normalisation floats in the fit. The d_0 significance template has had a small correction (-0.030) so that the mean is aligned to the mean of the d_0 significance distribution in the dataset.

- Prompt: Both BDT and d_0 significance templates taken from J/ψ muons in the inclusive $pp \to J/\psi$ PYTHIA8 Monte Carlo where J/ψ production is dominated by prompt production. These muons will be real and should thus occupy the high values in the BDT output distribution and will have a narrow d_0 significance distribution as they are produced at the interaction point. The shape of the templates are fixed but the normalisation floats in the fit. The d_0 significance template has had a small correction (-0.052) so that the mean is aligned to data.

- Fake Muon: Both the BDT and d_0 significance fake muon templates are taken from the PYTHIA8 Monte Carlo sample. The fake muons are defined in the same way as for BDT training in Section 5.7.1. The BDT template contains a large contribution at low values. The d_0 significance template is a mix of prompt and non-prompt shapes as fake muons can have both prompt and non-prompt sources. The shape of the templates are fixed but the normalisation floats in the fit. The d_0 significance template has had a small correction (-0.12) so that the mean is aligned to data.

- Fake J/ψ: Data driven templates described in 5.7.2. The templates are fixed in the fit as the data driven method to determine this contribution should account for it completely.

- Pile-up: Data driven background templates described in 5.7.3. The templates are fixed in the fit as the data driven method to determine the pile-up contribution should account for it completely.

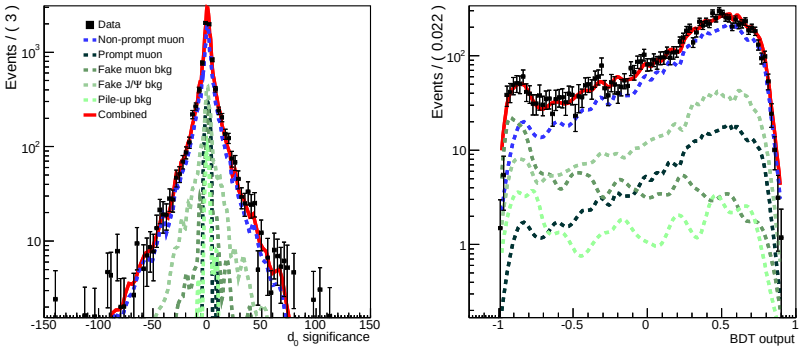
The fit is executed in bins of the various observables listed in Table 5.2 with the non-prompt, fake muon, fake J/ψ and pile-up templates derived in the same

observable bin as fitted to in data. The prompt template was filled using the $pp \to J/\psi$ PYTHIA8 Monte Carlo but with a modified event selection. Events were required to have at least two muons, as only a small fraction contained three reconstructed muons passing the quality cuts, all the other cuts were the same as the nominal selection. Without three muons in every event the differential variables cannot be calculated so the same template formed on the inclusive Monte Carlo sample is used for the prompt template in each differential fit. To minimise the impact of statistical fluctuations the differential templates are smoothed using a kernel estimation procedure [104] before being fit to data.

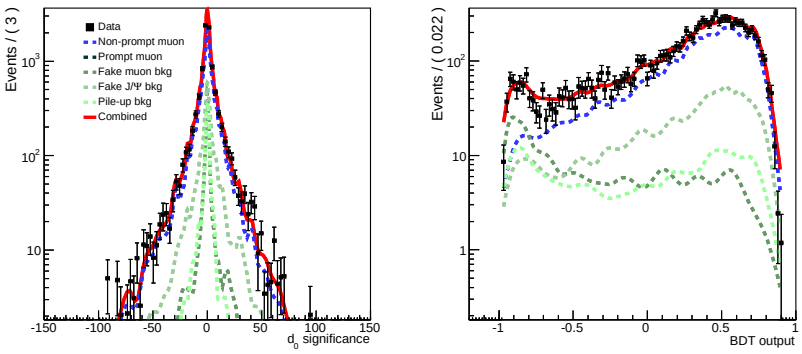
Example third muon fit results for a bin in the three angular correlation variables can be seen in Figure 5.17.

5.7.8. Extrapolation to Full τ Region

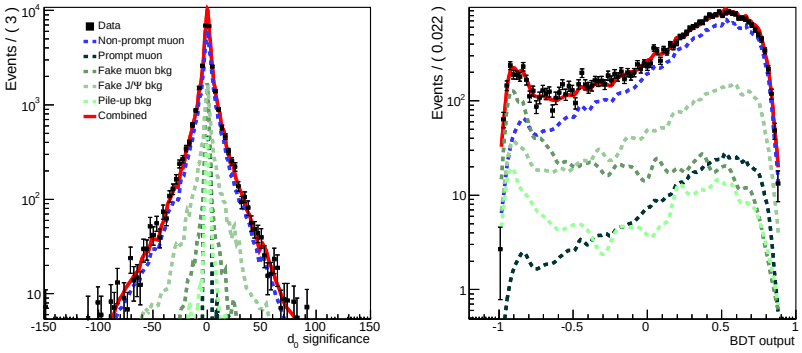
To account for the signal events below the J/ψ $\tau < 0.25$ ps cut, the number of fitted non-prompt third muons needs to be extrapolated over the full pseudo-proper lifetime range of the J/ψ. The assumption used is that the composition of third muons is unchanging in τ for a non prompt J/ψ, i.e. for an event with a non-prompt J/ψ the probability of the third muon being from a B-hadron is independent of τ. This was confirmed in data by looking at inclusive fits in different τ-slices and looking at how the fitted third muon composition changes. Figure 5.18 shows for bins of J/ψ pseudo-proper lifetime, the ratio of fitted number of non-prompt J/ψs to fitted number of third muons types, for each of the three floating third muon components, non-prompt, prompt and fake muon. It shows that above the lifetime cut used in the analysis ($\tau > 0.25$ ps) the third muon composition with respect to non-prompt J/ψs is constant as a function of τ. Below values of $\tau < 0.1$ ps deviation from flatness is observed. This region has a much larger background contamination for both the J/ψ and the third muon fits. Non-prompt J/ψ events are no longer the dominant contribution for the di-muon fits, with larger contributions from prompt J/ψs and fake J/ψs. The reduced number of non-prompt J/ψs in the low τ region explains the deviation in Figure 5.18, as the di-muon background events can also contain non-prompt third muons, changing the ratio of fitted non-prompt J/ψs to non-prompt muons.



(a)



(b)



(c)

Figure 5.17.: Third muon simultaneous fit results. The data are compared to the fitted model, with each individual model component included, for differential bins a) $0.4<\Delta R<0.8$, b) $0.5<\Delta\phi<1.0$, c) $0.7<\Delta y<1.2$.

As a cross check the τ dependence of the third muon composition has also been checked in Monte Carlo. For events containing a J/ψ, which comes from a B-hadron decay, the fraction of non-prompt muons and fraction of fake muons is plotted as a function of τ in Figure 5.19. The non-prompt third muon composition was found to be extremely consistent over the whole range including below $\tau = 0.1$ ps which validates the extrapolation assumption.

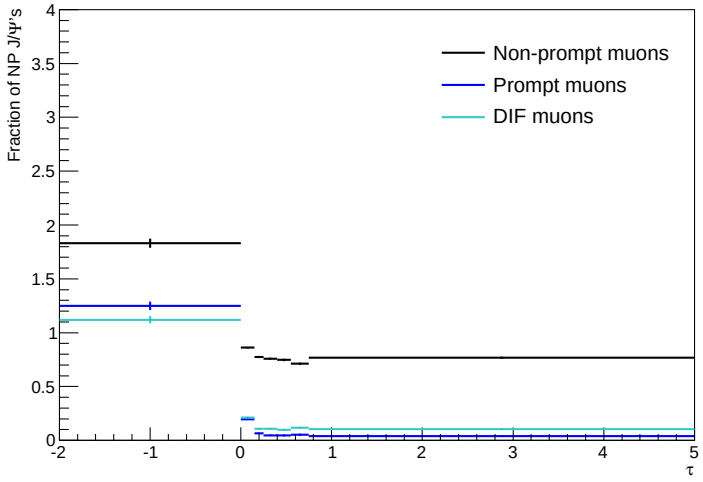


Figure 5.18.: In bins of J/ψ lifetime, the ratio of numbers of non-prompt J/ψs to number of third muons determined from fits to data, for each of the three floating third muon components, non-prompt, prompt and fakes.

The extrapolation to the full τ-spectrum is then performed by simply correcting the third muon yield found in the τ>0.25 ps region by an extrapolation factor taken as the ratio of all non-prompt J/ψs over the full τ range to the number of J/ψs found above the τ>0.25 ps cut. This correction is taken from the 2-D J/ψ fit results and accounts for all of the non-prompt J/ψ events below the $\tau < 0.25$ ps cut. The correction is derived individually for each differential observable bin.

5.7.9. Propagation of Statistical Uncertainty

The statistical uncertainty on the number of signal events, and therefore the cross-section, comes from two sources. Firstly uncertainty on the number of fitted non-prompt third muons.

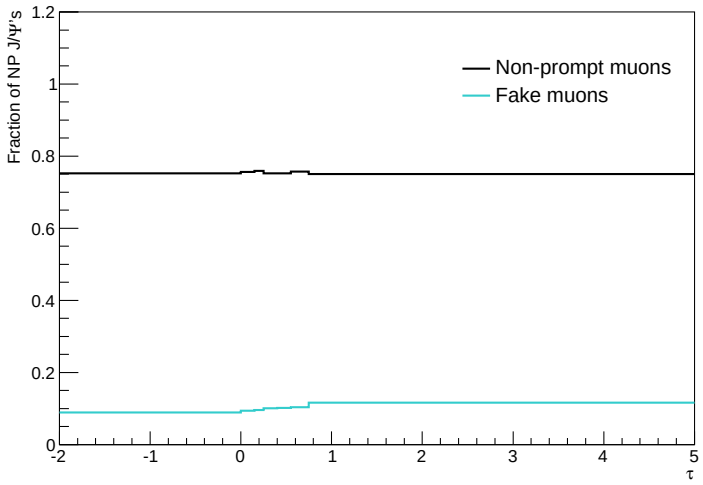


Figure 5.19.: In bins of J/ψ lifetime, the ratio of numbers of non-prompt J/ψs to number of third muons determined from Monte Carlo, for non-prompt and fakes muons.

Secondly a source of statistical uncertainty from the $\tau < 0.25$ ps cut on the dataset for the third muon fits is included, as this limits the number of events going into the third muon fit. This is derived from the fitted uncertainty on the slope parameter of the non-prompt J/ψ lifetime component. The slope parameter was found to be uncorrelated with any of the other J/ψ model parameters and is solely responsible for how many non-prompt J/ψs populate the high τ region. The number of non-prompt J/ψs above the lifetime cut was varied by changing the slope parameter by $\pm 1\sigma$ the uncertainty from the non-prompt J/ψ model, with the uncertainty taken from the 2-D J/ψ fit. The extrapolation factor, from Section 5.7.8, is recalculated for the varied non-prompt J/ψ model with the largest difference from either the plus or minus variation as compared to the default extrapolation factor taken as the error. This error estimate is then combined with the third muon fit uncertainty on the number of non-prompt third muons to define the total statistical uncertainty on the number of signal events.

5.8. Fit Verification

As the J/ψ and third muon fits determine the final differential cross-sections it is important that they are performing accurately. Several tests were performed to validate them.

To test the stability and performance of the J/ψ fit a Monte Carlo closure test was performed. Toy datasets of 20000 events were produced by combining random amounts of non-prompt J/ψs and prompt J/ψs. With non-prompt J/ψ contribution coming from the PYTHIA8 $bb \to J/\psi + X$ Monte Carlo and the prompt from the PYTHIA8 $pp \to J/\psi$ sample. Both samples have passed the data selection as outlined in Section 5.2 before sampling to form the toy datasets. The 2-D J/ψ model is then fit to each toy dataset and the fraction of prompt and non-prompt J/ψs extracted from the fit is compared to the fractional composition of prompt and non-prompt Monte Carlo used in it's creation. Figure 5.20 shows the fractional difference between the true and fitted number of prompt and non-prompt J/ψs. It can be seen that the fit model is performing very well with deviations from the true composition consistently below 2%.

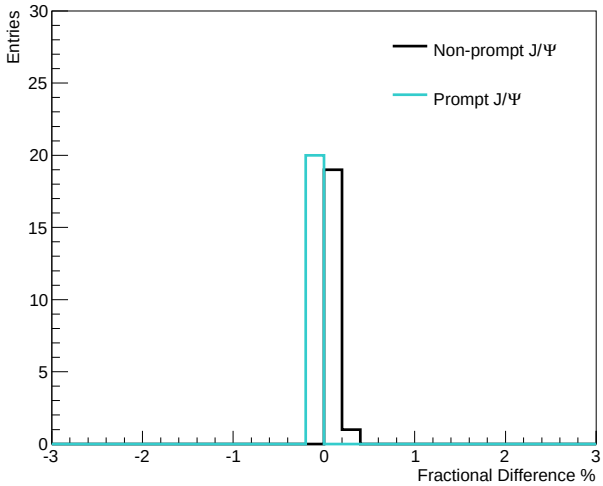


Figure 5.20.: Fractional difference (%) of the number of fitted J/ψ components, both prompt and non-prompt, when compared to known fraction from Monte Carlo.

.

To test the stability and performance of the third muon d_0 significance fit another Monte Carlo closure test was implemented. Toy datasets of 20000 events were produced by combining random amounts of events with expected non-prompt third muons and events with expected prompt third muons. The non-prompt third muon contribution comes from events in the PYTHIA8 $BB \to J/\psi + X$ Monte Carlo, where the third muon contribution is dominated by muons from a decay of a B-hadron. The prompt third muon contribution comes from events in the PYTHIA8 $pp \to J/\psi$ sample, where the third muon composition is mostly prompt muon candidates. Both samples are required to pass the data selection cuts, as outlined in Section 5.2, before sampling. A 1-D maximum likelihood fit of the d_0 significance distribution was performed on the third muons in the toy datasets. The normal fitting procedure was performed using the same Monte Carlo templates used in the data analysis to model the prompt and non-prompt d_0 significance templates. The fake J/ψ and pile-up backgrounds were removed using the usual procedure, however fake muon background was set to zero for the fits. The number of fitted (non-)prompt third muons are compared to fractions of the non-prompt and prompt Monte Carlo sources used in the toy dataset construction. Figure 5.21 shows the fractional difference for both the prompt and non-prompt third muon components compared to the true number. It can be seen that the model is performing well with only small deviations for the prompt distribution.

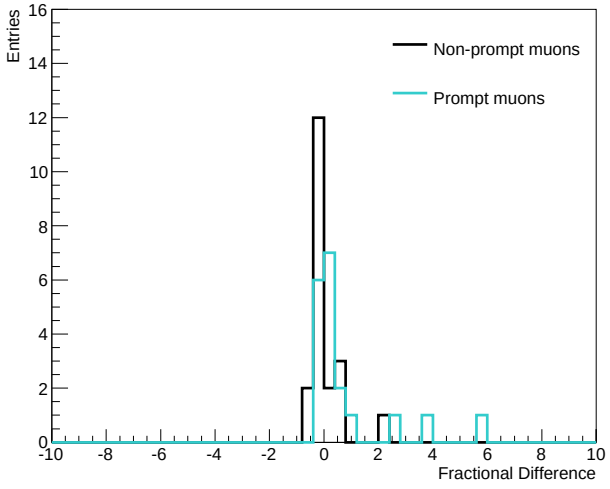


Figure 5.21.: Fractional difference (%) of the number of fitted muon components, both prompt and non-prompt, when compared to fraction from Monte Carlo.

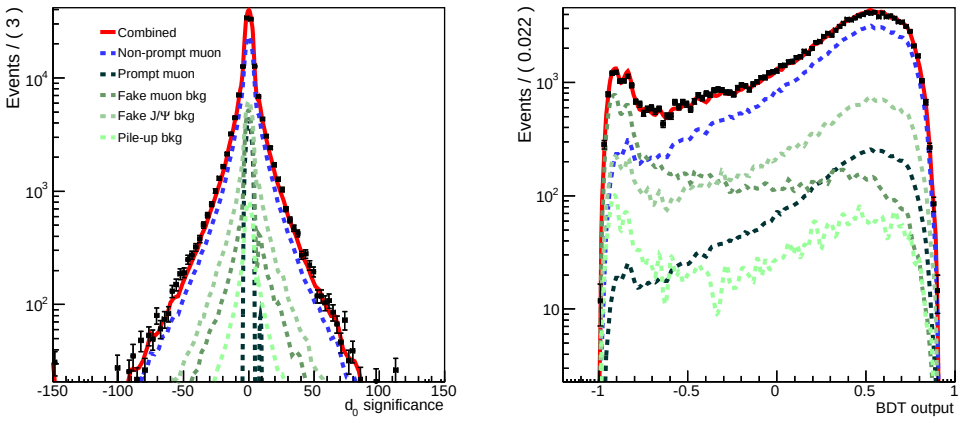
.

5.8.1. Fake Muon Fit Test

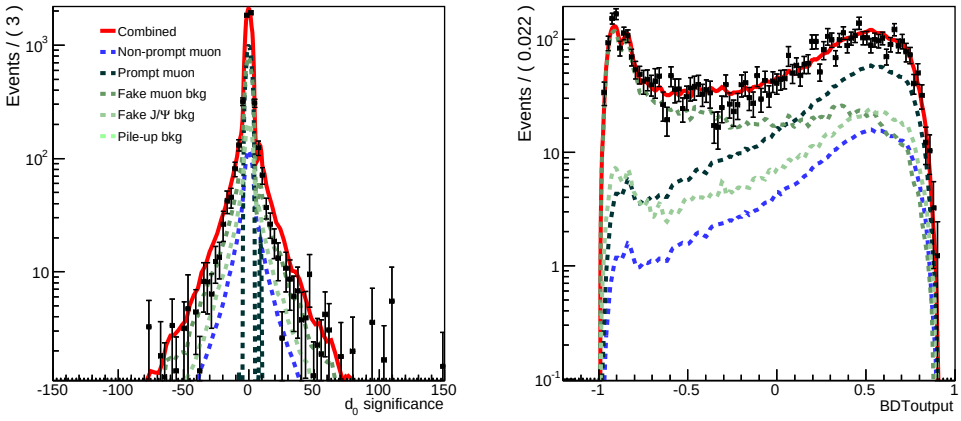
The fake muon component peaks in the region of most interest to this analysis, so it is important that the amount of fake muons is accurately modeled in the third muon fits. Two qualitative cross checks on the modeling of the fake muon component were defined by looking at independent orthogonal data control regions expected to contain more fake muons. The first control region is defined by reversing the pile-up cut, so looking at events with $|\Delta z_0| > 40$ mm. This means the J/ψ candidate and third muon have been created in separate p-p collisions and as QCD interactions dominate at a hadron collider the probability of charged pions and kaons faking a third muon increases in this region. The second control region is defined by looking at prompt di-muon events, that is reversing the pseudo-proper lifetime cut, $\tau < 0.25$ ps. In this region prompt J/ψ production is from QCD processes which again increases the probability of charged pions and kaons faking a third muon. The fits in these control regions were performed inclusively as statistics did not allow for a splitting into differential bins. The third muon fit procedure was the same as the default (see Section 5.7.7). The exception being that no pile-up template was included in the fit to the pile-up control region, as by definition the pile-up template is defined by all events with $|\Delta z_0| > 40$ mm. The result of these fits can be seen in Figure 5.22, where Table 5.4 details the amount of fake muons fitted in the two control regions as well as the main inclusive fit. It can be seen that in both control regions the fit is behaving as expected and fitting a higher fraction of fake muons.

Fit Region	Fake muons %	Non-prompt %	Prompt %
Standard	12	81	7
Pile-up	50	11	39
Prompt	20	61	19

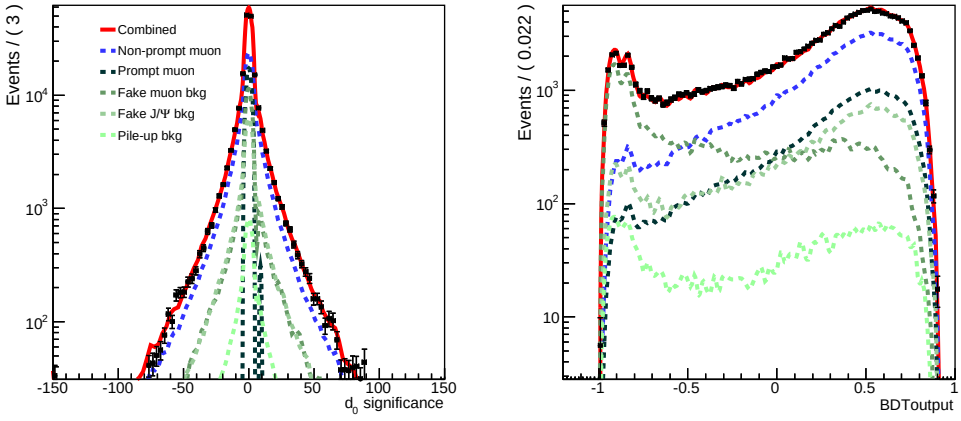
Table 5.4.: Components of third muon fits in nominal and fake muon control regions. The amount of each variable third muon component is listed as a percentage of the total number of fitted floating components, excluding the pile-up and fake J/ψ contributions.



(a)



(b)



(c)

Figure 5.22.: Inclusive 2-D third muon fit result, for a) nominal fit, b) fit in pile-up region and c) fit in prompt region.

5.9. Systematics

A wide range of possible systematic uncertainties are accounted for in this measurement. They broadly fit into three categories: uncertainties associated with the efficiency corrections to data, J/ψ model systematics and uncertainties on the backgrounds in the fits. Each systematic source is considered individually by repeating the differential analysis with systematic change implemented, the difference from the nominal result is then taken as the uncertainty. All of the systematics apart from those concerning J/ψ modeling are double sided, and are varied in both directions. The total systematic uncertainty on the measurement is taken as the sum in quadrature, with all upward/downward fluctuations summed accordingly. The largest deviation for either the combination of upward or downward systematics is symmetrised and used as the total systematic uncertainty on the measurement. Table 5.5 lists the dominant systematic uncertainties to the analysis. The derivation of each of these, as well as all other sources of systematic uncertainty, are described in more detail in the following sections.

Source of uncertainty	Approximate relative fractional range [%]
J/ψ model uncertainty	4-18
$C_{\mu\mu}$ uncertainty	2-10
$B + D$ background	0-10
Monte Carlo template statistics	0-8
Close-by trigger third muon	1-6
Fake J/ψ background normalisation	0-5
Luminosity	2.8

Table 5.5.: Summary of the largest systematic uncertainties determined for the cross-section measurements. The range covers all differential observable bins.

5.9.1. Trigger and Muon Reconstruction Efficiency Uncertainty

The uncertainty on the trigger and muon reconstruction maps was detailed in Section 5.3. The trigger efficiency was factorised into two components in Equation 5.6.

The uncertainty on ϵ^{x}_{ROI} was derived from the spread in dataset yields when using a series of pseudo maps to re-weight the dataset. This spread is propagated to the number of fitted signal events to define the trigger map uncertainty. The uncertainty on $C_{\mu\mu}$ is applied separately by varying the nominal function by $\pm 1\sigma$ of its associated error consistently when re-weighting the dataset. In addition an uncertainty on the efficiency correction is applied for cases when the third muon is close to a trigger muon, as discussed in Section 5.3.2. The parameters of the function describing the correction are varied within errors to define two systematic functions used in lieu of the nominal function to determine the uncertainty. These are included along with the nominal correction in Figure 5.5 for the case where the muons are of opposite charge.

The uncertainties on muon reconstruction efficiencies are defined for the two factorised components of Equation 5.7. A 0.5% uncertainty is included for the efficiency of reconstructing a muon track in the ID, ϵ_{trk}. This is added coherently for each muon of the three muons in an event resulting in a flat 1.5% systematic uncertainty. Similarly as the trigger map efficiency, the uncertainty on the muon reconstruction maps, $\epsilon_{\mu}(p_T, q \cdot \eta)$, is defined by the spread on the dataset yields when using a set of pseudo maps.

Figure 5.23 shows the relative fractional uncertainty for each of these detector efficiency corrections for the $\Delta\phi(J/\psi, \mu)$ distribution. The uncertainty on $C_{\mu\mu}$ in general gives the largest systematic for the efficiency corrections with a fractional uncertainty of $\approx 3\%$.

5.9.2. Luminosity Uncertainty

A flat 2.8% uncertainty is assigned to the delivered integrated luminosity, the method of deriving this is documented in [105].

5.9.3. Template Statistical Uncertainty

To assess the statistical uncertainty on the Monte Carlo templates used for the third muon fits, each fit is repeated 100 times with a set of toy templates. Each of the three Monte Carlo floating templates used in the fit, non-prompt, prompt and fake

muon, are randomly sampled to produce a toy template that has the same number of events as the nominal case. The fit is then repeated using the toy templates and the number of fitted non-prompt muons is recorded. The non-prompt muon yield distribution, from the toy fits, is fitted with a Gaussian function and the width taken as a systematic uncertainty on the statistical fluctuations of the Monte Carlo templates. Figure 5.23 includes the relative fractional uncertainty due to template statistics for the $\Delta\phi(J/\psi, \mu)$ distribution, where it typically contributes at the 1-2% level.

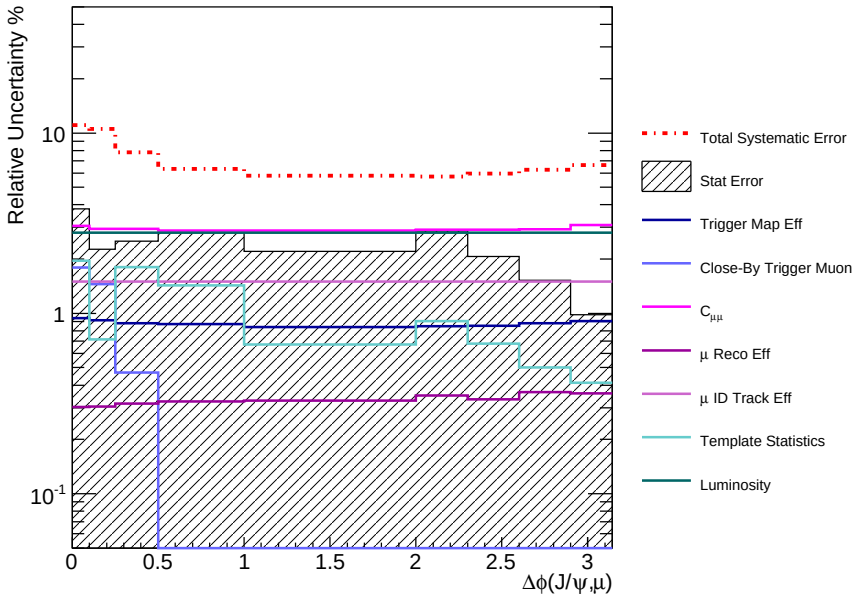


Figure 5.23.: Relative systematic uncertainties for trigger and muon reconstruction efficiencies, luminosity uncertainty and template statistical uncertainty as a function of $\Delta\phi(J/\psi, \mu)$. The statistical uncertainty and total systematic uncertainty is included for comparison. Only the largest relative uncertainty from either the upward or downward systematic is plotted.

5.9.4. Fake Muon Template Uncertainty

The fake muon template contains two types of background with similar behavior: DIF and hadronic leakage, as explained in Section 5.7.1. Both backgrounds are due to the decay or interactions of charged pions and kaons. To assess the robustness of the Monte Carlo modeling of the fake muon background the templates used in the third muon fits were systematically altered.

Firstly the fraction of pions to kaons populating the fake muon templates is changed by ±50%. The BDT response is subtly different for pions and kaons but due to limited statistics of fake muon candidates in Monte Carlo the two sources of fake muons have been combined, changing the ratio should cover any effect of the combination. Secondly the ratio of the number of decays of DIF muons inside the ID and outside the ID is varied by ±50%. As the BDT has variables that specially pick out fake muons in certain parts of the detector it is important to assess a potential mis-modeling of the radial decay position of fake muons in MC. Finally the ratio of DIF muons and hadronic leakage muons in the fake muon template is changed by ±50%. The BDT response is different for the two types of fake muons as two of the BDT variables are based on ID variables which have little discriminative power for hadronic leakage faking muons. The fake muon template from MC is composed of approximately 75% DIF muons, with the available Monte Carlo statistics not allowing separation of the two contributions. Changing the fractional composition of the fake muon template should cover any mis-modeling in Monte Carlo of the composition of the two sources of fake muons.

The effects these systematic shifts have on the BDT template can be inferred from Figure 5.24, where the Monte Carlo fake muon template is broken down into all the individual sources that are varied as part of the fake muon template uncertainties.

5.9.5. B_c Background Uncertainty

The B_c background prediction is taken from the average of PYTHIA8 and HERWIG++ Monte Carlo predictions, as discussed in Section 5.7.4. The difference between the two predictions is assigned as an uncertainty on the number of B_c-mesons in the dataset.

5.9.6. B+D Events Uncertainty

Similarly to the B_c-meson background, and discussed in Section 5.7.5, the number of events estimated to be from $B+D$-hadrons is taken from the average of PYTHIA8 and HERWIG++ Monte Carlo predictions. A systematic uncertainty to this prediction is applied, using the difference between PYTHIA8 and HERWIG++ for the rate of $B + D$ events.

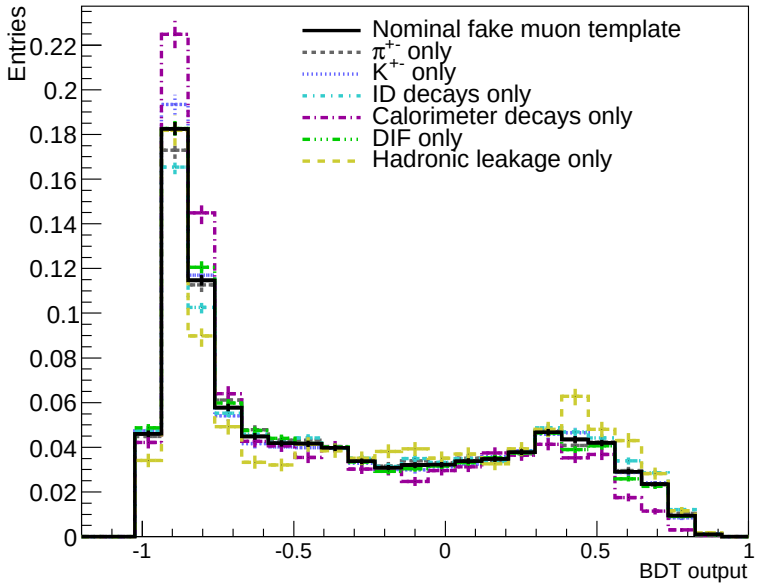


Figure 5.24.: Unit normalised BDT output distribution for fake muons. The nominal fake muon template is split into six contributing components.

5.9.7. Punch-through Background

The punch-through background prediction is varied by ±50% to estimate the uncertainty on this background. This is approximately equivalent to 11 cm of extra (or less) steel for pions to traverse[1]. As punch-through is such a small background to the analysis this systematic is negligible across all differential bins.

5.9.8. Data-Driven Background Uncertainties

The uncertainty on the fake J/ψ background is assessed by changing the normalisation of the fake J/ψ templates in the third muon fits. The number of fake J/ψ events was derived from the 2-D di-muon fits, given by the normalisation of the three fake J/ψ components within the di-muon signal mass window. Due to the di-muon pseudo-lifetime cut in the third muon fit region, the single sided fake J/ψ component is the only background to contribute non-negligibly to the high pseudo-proper lifetime

[1] this assumes a 14.2 cm interaction length for a pion traversing steel [106]

region. The differential fits are repeated with fake J/ψ template normalisations altered by $\pm 1\sigma$ of the uncertainty on the normalisation of the single-sided background.

For the error on the pile-up background a similar procedure is used. The templates used in the third muon fits are changed by altering the normalisation within their uncertainty. The uncertainty is derived from the Gaussian fit to Δz_0, see Figure 5.14, and is applied as a $\pm 1\sigma$ variation to nominal pile-up templates.

Figure 5.25 shows the relative uncertainty for all the background systematic variations for the $\Delta\phi(J/\psi, \mu)$ distribution.

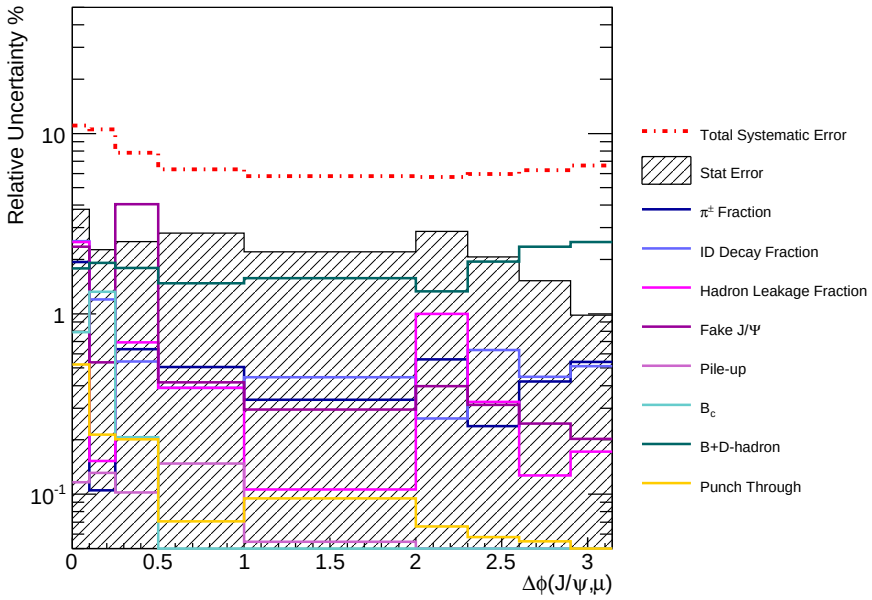


Figure 5.25.: Relative background modeling uncertainties as a function of $\Delta\phi(J/\psi, \mu)$. The individual systematic changes to a background component are plotted. The statistical uncertainty and total systematic uncertainty is included for comparison. Only the largest relative uncertainty from either the upward or downward systematic is plotted.

5.9.9. J/ψ Model Uncertainty

To assess any bias on the fitted number of non-prompt J/ψs taken from the 2-D di-muon fit due to the model choices taken, various changes were made to the functions describing the five fit components. The analysis was repeated for each of the varied J/ψ models, with only one change at a time. To avoid any potential double counting

of errors, the envelope of the largest deviation from the nominal event yield when considering all the individual model changes is taken as the total systematic for the J/ψ model uncertainty. This envelope is calculated separately in each differential bin. The nominal fit model is listed in Table 5.3, with the individual systematic changes to the J/ψ model listed below:

- The J/ψ mass model was switched to two Gaussian functions.
- The non-prompt J/ψ pseudo-proper lifetime model was changed to double exponential function convoluted with the same resolution function.
- The resolution model was changed to a single Gaussian.
- The Crystal Ball function's fixed parameters are varied by $\pm 10\%$
- The prompt fake J/ψ background had its di-muon mass model interchanged with an exponential function.
- The single sided fake J/ψ background had its di-muon mass model interchanged with a first order polynomial function.
- The double sided fake J/ψ background had its di-muon mass model interchanged with a first order polynomial function.
- The single sided fake J/ψ pseudo-proper lifetime model was swapped from a single to a double exponential function.

Figure 5.26 shows the model systematics as a function of $\Delta\phi(J/\psi, \mu)$ including the envelope which is taken as the uncertainty on J/ψ modeling. The dominating model uncertainties are from changing the lifetime parameterisation of the non-prompt J/ψ and mass parameterisation of the single sided fake J/ψ background background.

5.10. Particle Level Definition

The differential cross-sections calculated from data are compared to Monte Carlo predictions in Section 5.11. It is important to have a stable particle level definition for signal events so that different Monte Carlo predictions can be compared consistently. The Monte Carlo predictions are taken at the three-muon level, explicitly after decay

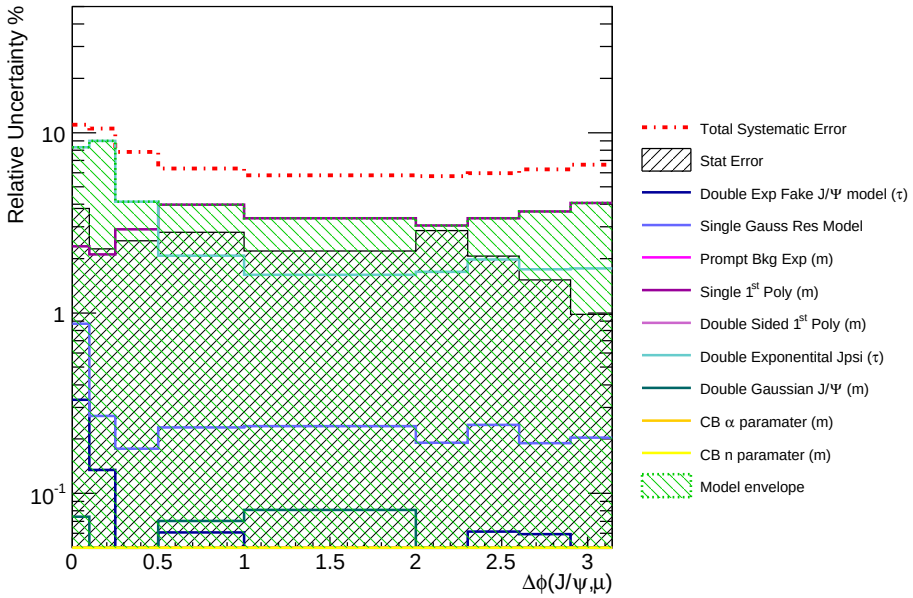


Figure 5.26.: Relative uncertainty for J/ψ model systematics as a function of $\Delta\phi(J/\psi, \mu)$. The individual model changes are plotted including the envelope used to define the model systematic. The statistical uncertainty and total systematic uncertainty is included for comparison.

of the B-hadrons. This *muon level* definition is kept as close as possible to the data selection so as to avoid any theoretical extrapolations into unmeasured phase space.

A particle level signal event requires two muons from a J/ψ meson which itself is the product of a B-hadron decay, which includes feed-down of higher mass charmonium states to a J/ψ if the excited charmonium states are from a B-hadron decay. In addition there must be another muon in the event from the decay of a different B-hadron. This muon can be from a direct semi-leptonic B decay or from a cascade decay with the B-hadron in its decay tree. All three muons are required to have $p_{\mathrm{T}} > 6$ GeV, with the two J/ψ muons in the pseudo-rapidity range $|\eta| < 2.3$ and the third muon in the range $|\eta| < 2.5$.

Truth events can be reconstructed in different analysis bins due to migration effects caused by the finite detector resolutions. However, in the PYTHIA8 Monte Carlo the migrations between particle and reconstruction level were found to be negligible. Figure 5.27 shows the migration matrix for the $\Delta R(J/\psi, \mu)$ distribution.

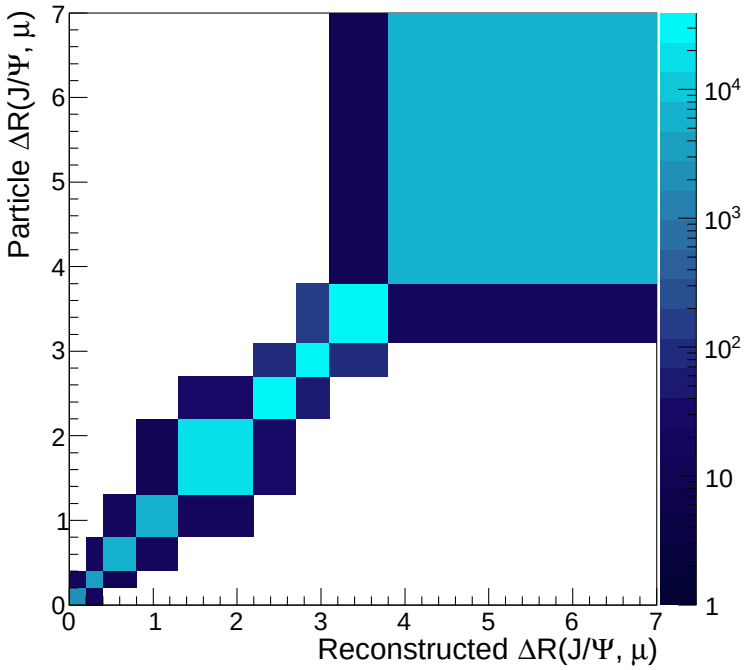


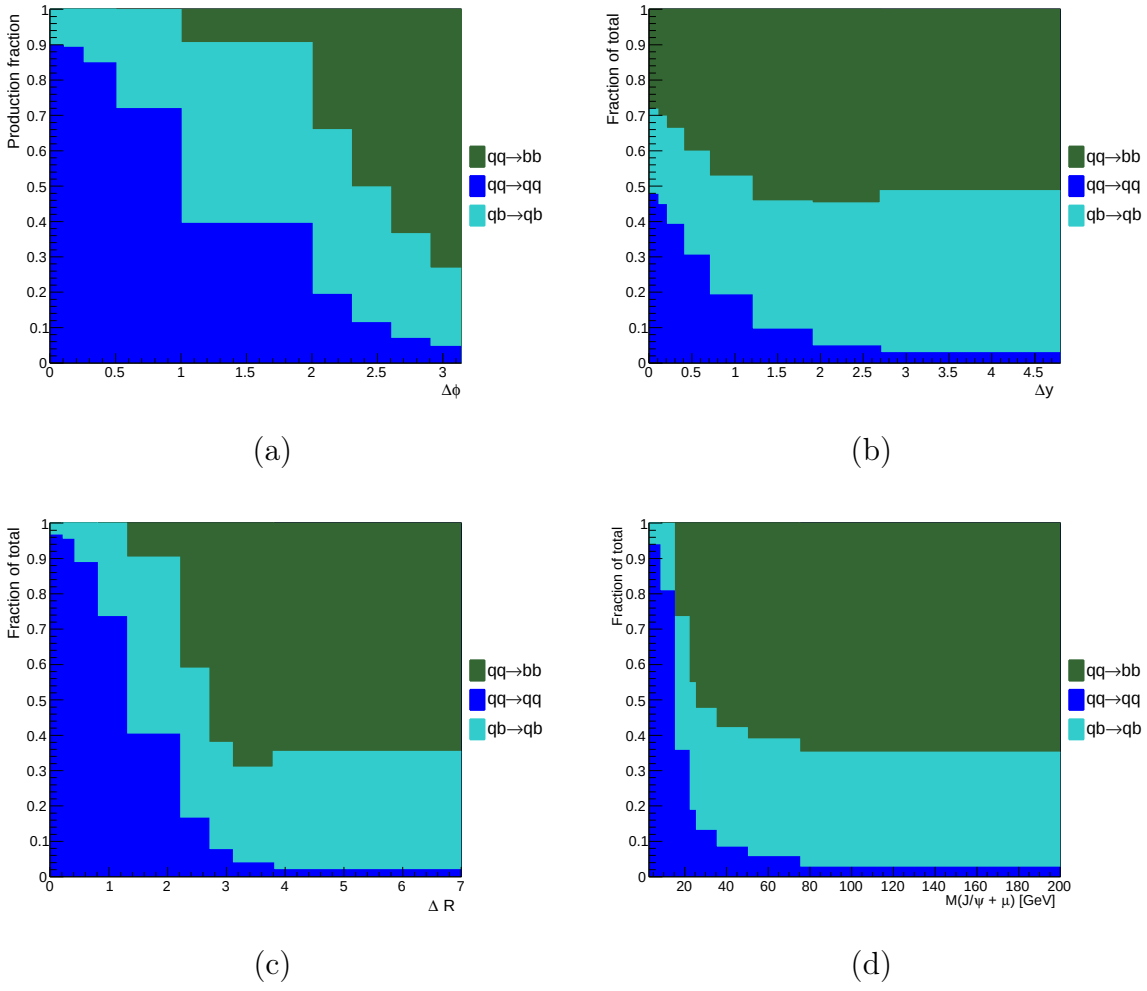
Figure 5.27.: Migration matrix for $\Delta R(J/\psi, \mu)$ distribution.

5.11. Theory Predictions

Two LO $2 \to 2$ multi-purpose generators matched to parton showers are used in comparison to the cross-sections measured in data. A PYTHIA8 prediction is taken from the Monte Carlo sample used in the analysis as described in Section 5.4. In addition a prediction is taken from an inclusive $pp \to b\bar{b}$ sample produced using HERWIG++ v2.7.1. The predictions are taken for the fiducial region defined in Section 5.10.

Both PYTHIA8 and HERWIG++ are LO predictions and are not expected to correctly estimate the normalisation. The predictions have been scaled to that the normalisation is matched to data in the high $\Delta R(J/\psi, \mu)$ region ($\Delta R > 3.0$) where they are expected to be most accurate. In this region the J/ψ and third muon are close to being back-to-back which should be well described by the LO $2 \to 2$ process. Large angle radiative corrections, which could lead to a more collimated J/ψ and third muon system, are not well modeled by parton shower simulations so the region at lower $\Delta R(J/\psi, \mu)$ may be less well described.

Figure 5.28 shows as a fraction of the total production rate in the PYTHIA8 sample the different $2 \to 2$ production processes included in the calculation. It should be noted that these are not observable quantities and so have no physical meaning but does help to envision the event topologies for the different production processes. Referring to Figure 5.1 the flavour creation process is given by $qq \to bb$ (where q can be any light quark or gluon), gluon splitting by $qq \to qq$ and flavour excitation by $qb \to qb$. It can be seen that the gluon splitting component is more concentrated in certain parts of the kinematic distributions. In the high $\Delta R(J/\psi, \mu)$ region the flavour creation process is expected to contribute to the majority of events.



(a) (b)

(c) (d)

Figure 5.28.: Fractional breakdown of the $b\bar{b}$ production processes in PYTHIA8. The distributions a) $\Delta\phi(b, \bar{b})$, b) $\Delta y(b, \bar{b})$, c) $\Delta R(b, \bar{b})$ and d) $m(b, \bar{b})$ are for the outgoing partons from a $2 \to 2$ process.

5.12. Results

Figures 5.29-5.34 show the data compared to the theory predictions outlined in 5.11. The inner error bars on the data points represent the statistical uncertainty, the outer error bars represent the total uncertainty taken as the total systematic and statistical uncertainties added in quadrature. The error on the theoretical predictions is statistical only.

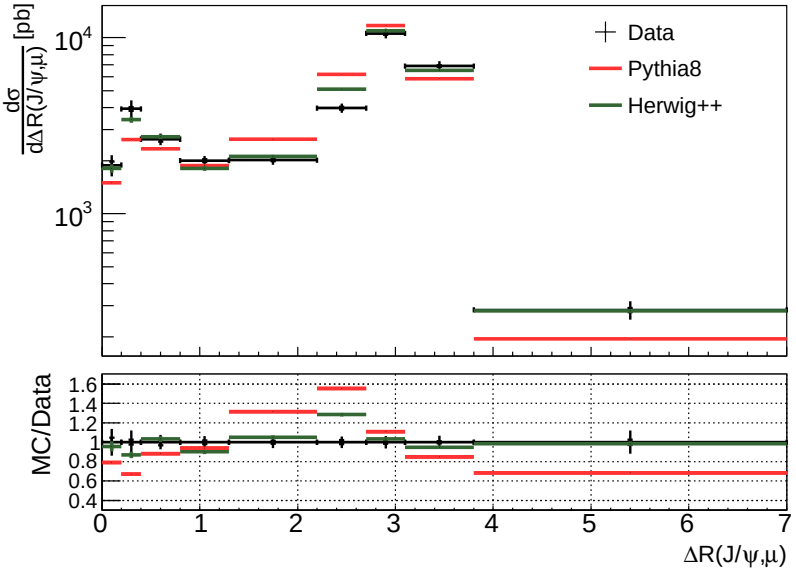


Figure 5.29.: Data differential cross-section for $\Delta R(J/\psi, \mu)$ compared to the Monte Carlo predictions from PYTHIA8 and HERWIG++.

In general HERWIG++ reproduces the shape of data more accurately for the angular distributions. This is especially apparent at low $\Delta R(J/\psi, \mu)$; in this region the PYTHIA8 prediction considerably under predicts the data. However HERWIG++ does not seem to describe the shape of the $p_{\mathrm{T}}(J/\psi, \mu)$ distribution and overestimates events in the low $p_{\mathrm{T}}(J/\psi, \mu)$ region. This distribution is well reproduced by PYTHIA8. Both predictions seem to have a shape dependence for the $m(J/\psi, \mu)$ distribution underestimating events in the low and high $m(J/\psi, \mu)$ regions. The y_{boost} distribution for both predictions is very similar, with deviations from data at high values of y_{boost}.

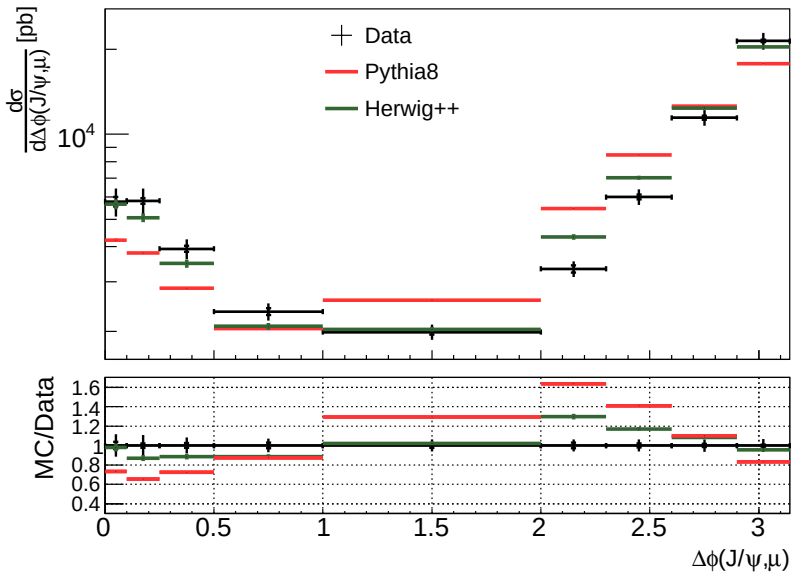


Figure 5.30.: Data differential cross-section for $\Delta\phi(J/\psi, \mu)$ compared to the Monte Carlo predictions from PYTHIA8 and HERWIG++.

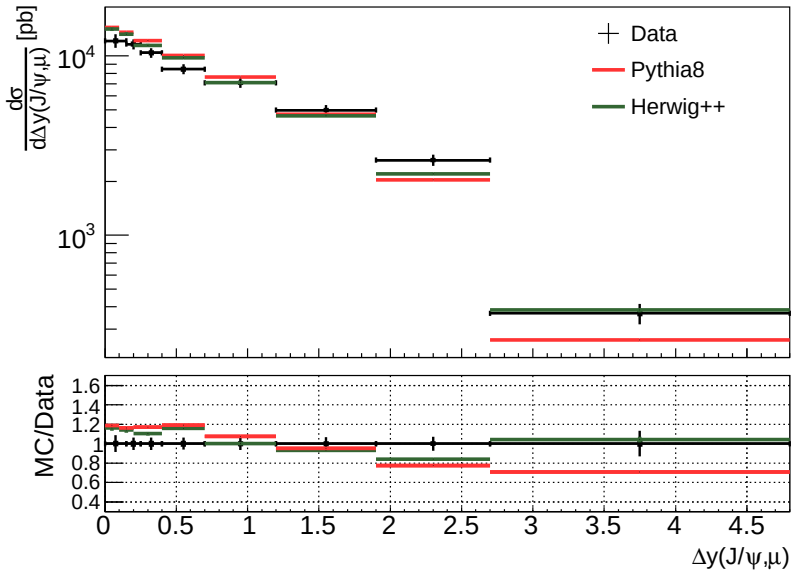


Figure 5.31.: Data differential cross-section for $\Delta y(J/\psi, \mu)$ compared to the Monte Carlo predictions from PYTHIA8 and HERWIG++.

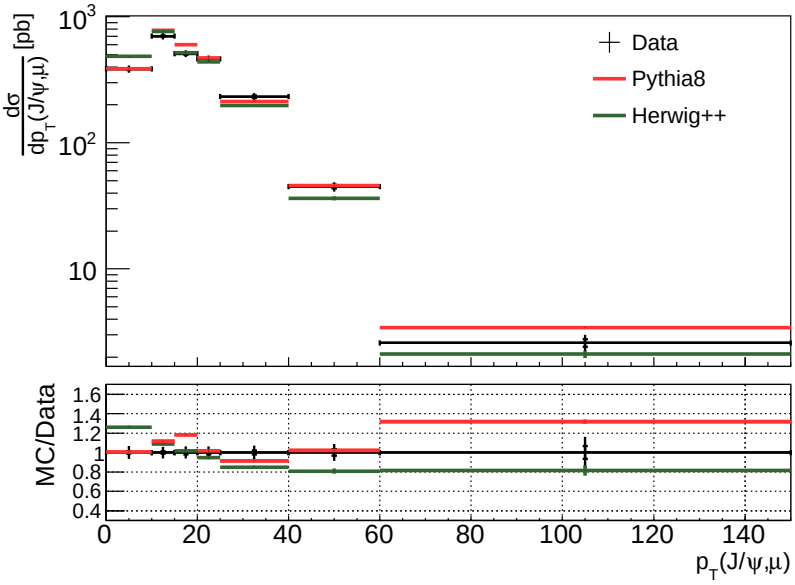


Figure 5.32.: Data differential cross-section for $p_{\mathrm{T}}(J/\psi, \mu)$ compared to the Monte Carlo predictions from PYTHIA8 and HERWIG++.

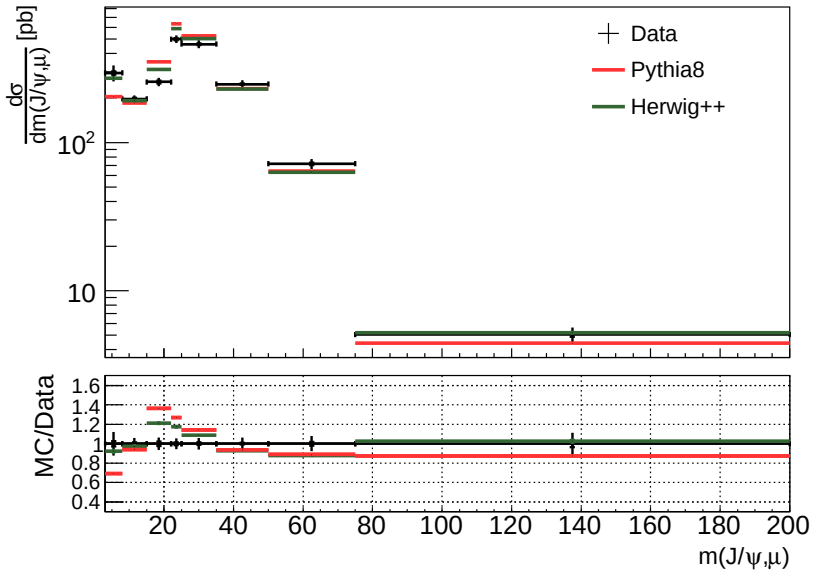


Figure 5.33.: Data differential cross-section for $m(J/\psi, \mu)$ compared to the Monte Carlo predictions from PYTHIA8 and HERWIG++.

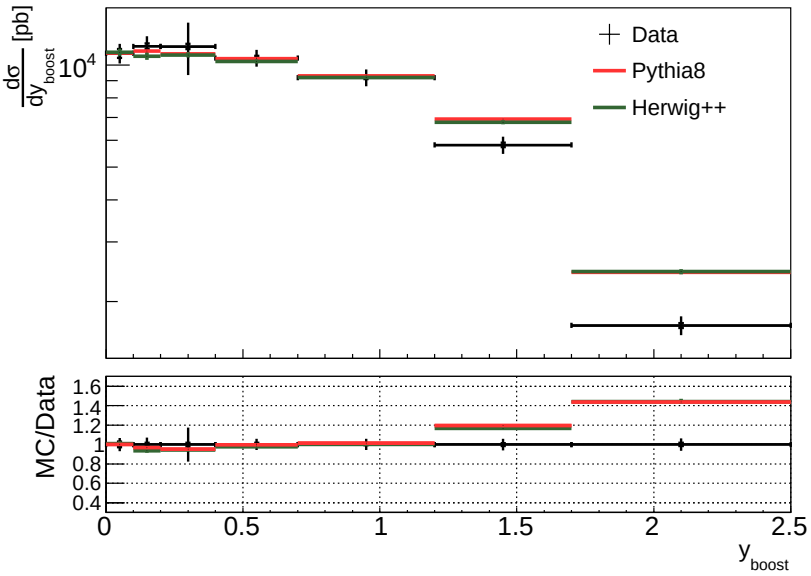


Figure 5.34.: Data differential cross-section for y_{boost} compared to the Monte Carlo predictions from PYTHIA8 and HERWIG++.

5.13. Conclusion

Differential cross-section distributions for muon level predictions of B-hadron decays have been presented. The data distributions were compared to two LO predictions from PYTHIA8 and HERWIG++ which were normalised to data in the high $\Delta R(J/\psi, \mu)$ region. The advantage of using muons rather than jets to tag B-hadrons is the improved angular resolution, which allows measurements in the region of phase space expected to be populated by events from the parton shower. In this low angular separation region, with an enhanced gluon splitting contribution, the prediction from HERWIG++ modeled the data well. While the prediction from PYTHIA underestimated the data in this region.

As mentioned at the start of this Chapter the work presented here is currently ongoing. The future planned extensions to the measurement detailed here mostly concentrate on how the data and theory predictions are compared. The current predictions are at the muon level while it is planned to present the data at the B-hadron level. For this to be achieved, theoretical transfer functions need to be derived to correct the kinematic variables of the J/ψ and the third muon to that of the parent B-hadrons. These are yet to be derived. In future, further Monte

Carlo predictions are to be added including LO multi-leg predictions. In particular a SHERPA prediction is being generated which should allow a comparison to the $Z + b\bar{b}$ analysis presented in Chapter 4 which included a SHERPA prediction (with a suggested deficiency of $b\bar{b}$ pairs at low angular separation).

Chapter 6.

Conclusions

The measurements of objects decaying to b-quarks will be very important for physics processes at ATLAS for Run-2 of data taking at the LHC. Unfortunately these measurements are very challenging in the vastly hadronic environment produced at the LHC. Boosted event topologies will become increasingly important to disentangling the signals from backgrounds. An accurate description of close-by jet production will become more important. For resonances decaying to heavy flavour, backgrounds from QCD $b\bar{b}$ production will be most challenging. Two measurements of $b\bar{b}$ production using Run-1 data produced at the LHC have been presented in this thesis which should help constrain the modeling of these backgrounds.

Firstly in Chapter 4, the production of two b-jets produced in association with a Z-boson was presented. This measurement used 4.5 fb^{-1}of data collected in 2011 at $\sqrt{s} = 7$ TeV. The inclusive fiducial cross-section for $Z + b\bar{b}$ production was found to be $520 \pm 20^{+74}_{-72}$ fb. The data was compared to several theoretical predictions at NLO and LO accuracy. Within large uncertainties the NLO predictions described the data, with tension in the low angular splitting region. The measurement was statistically limited and could be improved with a repeat of the analysis using the larger dataset from 2012. With approximately four times as much collected data in 2012 by the ATLAS detector the statistical precision would improve by a factor of two. The uncertainty on b-tagging was one of the dominant systematics (see Section 4.5.1). Improvements on the performance of b-tagging algorithms is possible for Run-2 of data taking at ATLAS due to the introduction of an extra Inner Detector pixel layer close to the beam pipe [107]. This *Insertable B-Layer* along with new b-tagging algorithms has resulted in improved performance for b-tagging efficiency and light

and c-jet rejection. These would lead to a purer sample of b-jets after b-tagging and would help reduce the jet-tagging efficiency systematics and would lead to a more accurate measurements for $Z + b\bar{b}$ final states.

Secondly in Chapter 5, $b\bar{b}$ production was probed by exploiting the decay of B-hadrons to muons. This allowed the measurement to low angular separations which are unable to be reached using jets. The measurement used data collected in 2012 at $\sqrt{s} = 8$ TeV, with a dataset size of 11.45 fb^{-1}. Differential cross-sections defined at the muon level were compared to two LO Monte Carlo predictions. A prediction from HERWIG++ described the angular correlation distributions in data well. The alternative prediction from PYTHIA8 showed shape differences compared to data across all distributions. The total uncertainty on the differential cross-section measurements was dominated by systematic sources. The potential of a larger dataset collected during Run-2 would not directly increase the accuracy of the measurements due to improvements in the statistical precision. However, the dominant systematic to this measurement came from the model choices used for fitting the di-muon system (see Section 5.9.9). The largest contributions came from changing the model of the two non-prompt components; non-prompt J/ψs and single-sided fake J/ψs. For Run-2 the p-p collision energy is increased (for the first data taking period this was to 13 TeV) which would lead to an increase in high p_{T} B-hadrons populating the non-prompt region (at high τ). The increase in statistics at high τ would help to better constrain the non-prompt models used in the J/ψ fits, potentially improving the overall precision of any cross section measurements. An alternative for measuring low angular separated B-hadrons is to use the improvements in continuous b-tagging [96]. This allows for tagging jets containing multiple B-hadrons. This can thus be used as a complementary technique to probe the low angular separation region below the jet-radius distance and can extend the range of jet measurements of $b\bar{b}$ pairs allowing a comparison of the results presented here.

Appendix A.

$Z + b\bar{b}$ Fit Results

The $Z + b\bar{b}$ fit results for the $|y(Z)|$, $m(b,\bar{b})$ and $\Delta R(b,\bar{b})$ distributions are shown in Figures A.1, A.2 and A.3 respectively.

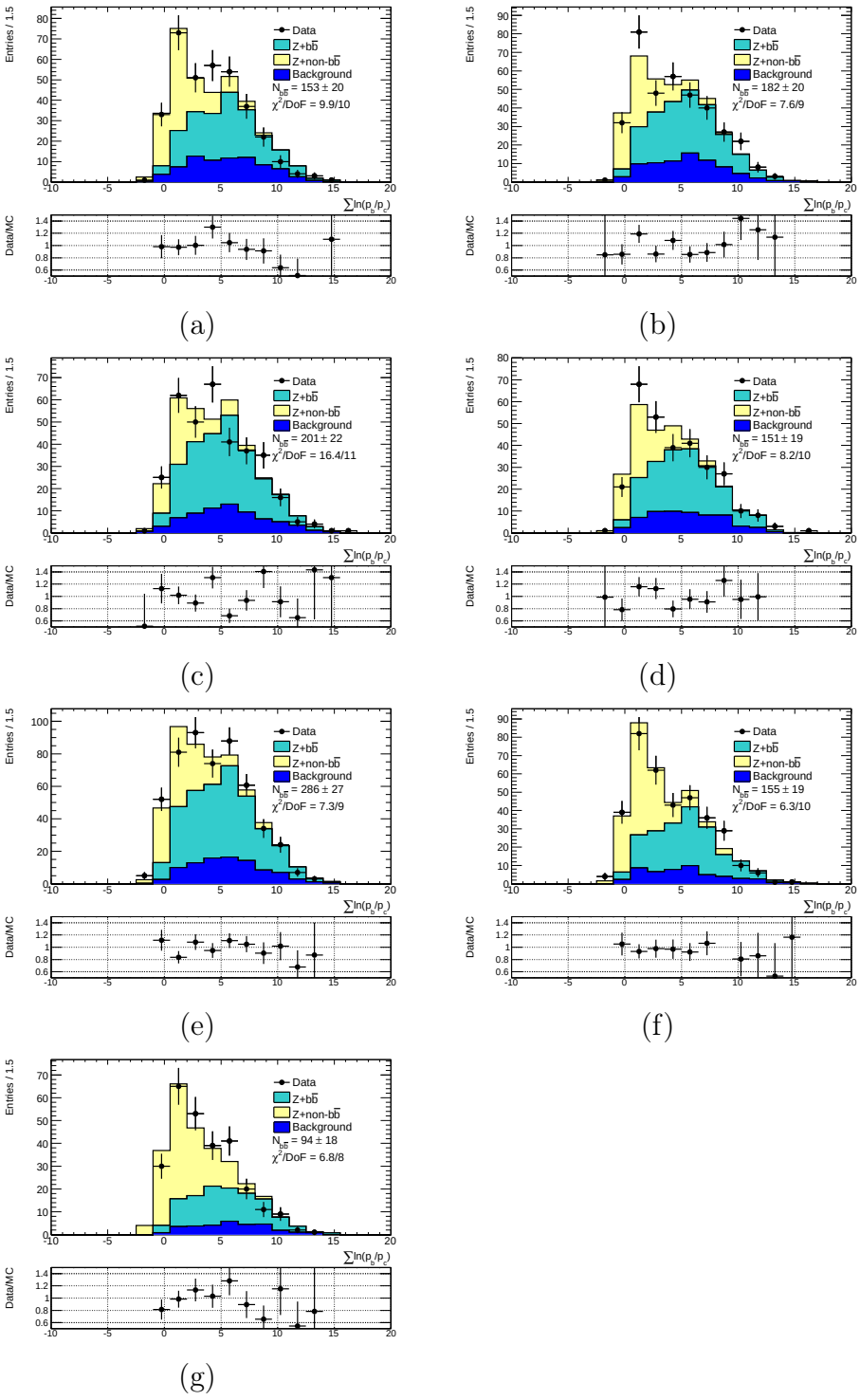


Figure A.1.: Result of fitting the number of true $b\bar{b}$ pairs, $n_{b\bar{b}}$, in bins of Z rapidity for (a) $|y(Z)| < 0.2$, (b) $0.2 < |y(Z)| < 0.4$, (c) $0.4 < |y(Z)| < 0.6$, (d) $0.6 < |y(Z)| < 0.8$, (e) $0.8 < |y(Z)| < 1.2$, (f) $1.2 < |y(Z)| < 1.6$ and (g) $1.6 < |y(Z)| < 2.5$.

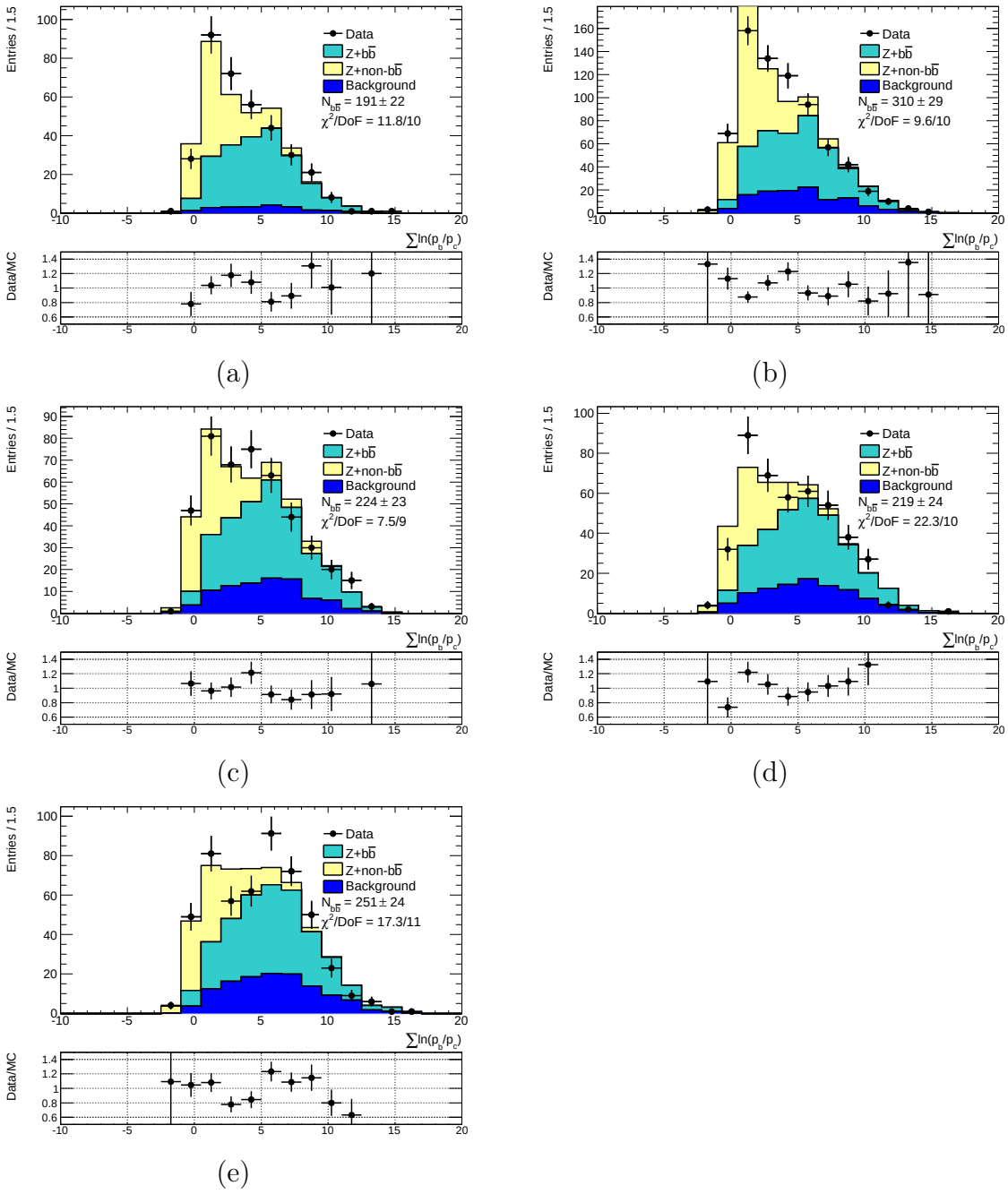


Figure A.2.: Result of fitting the number of true $b\bar{b}$ pairs, $n_{b\bar{b}}$, in bins of the invariant mass of the two tagged b-jets for (a) $10 < m_{b\bar{b}} < 45$GeV, (b) $45 < m_{b\bar{b}} < 85$GeV, (c) $85 < m_{b\bar{b}} < 115$GeV, (d) $115 < m_{b\bar{b}} < 165$GeV and (e) $165 < m_{b\bar{b}} < 350$GeV.

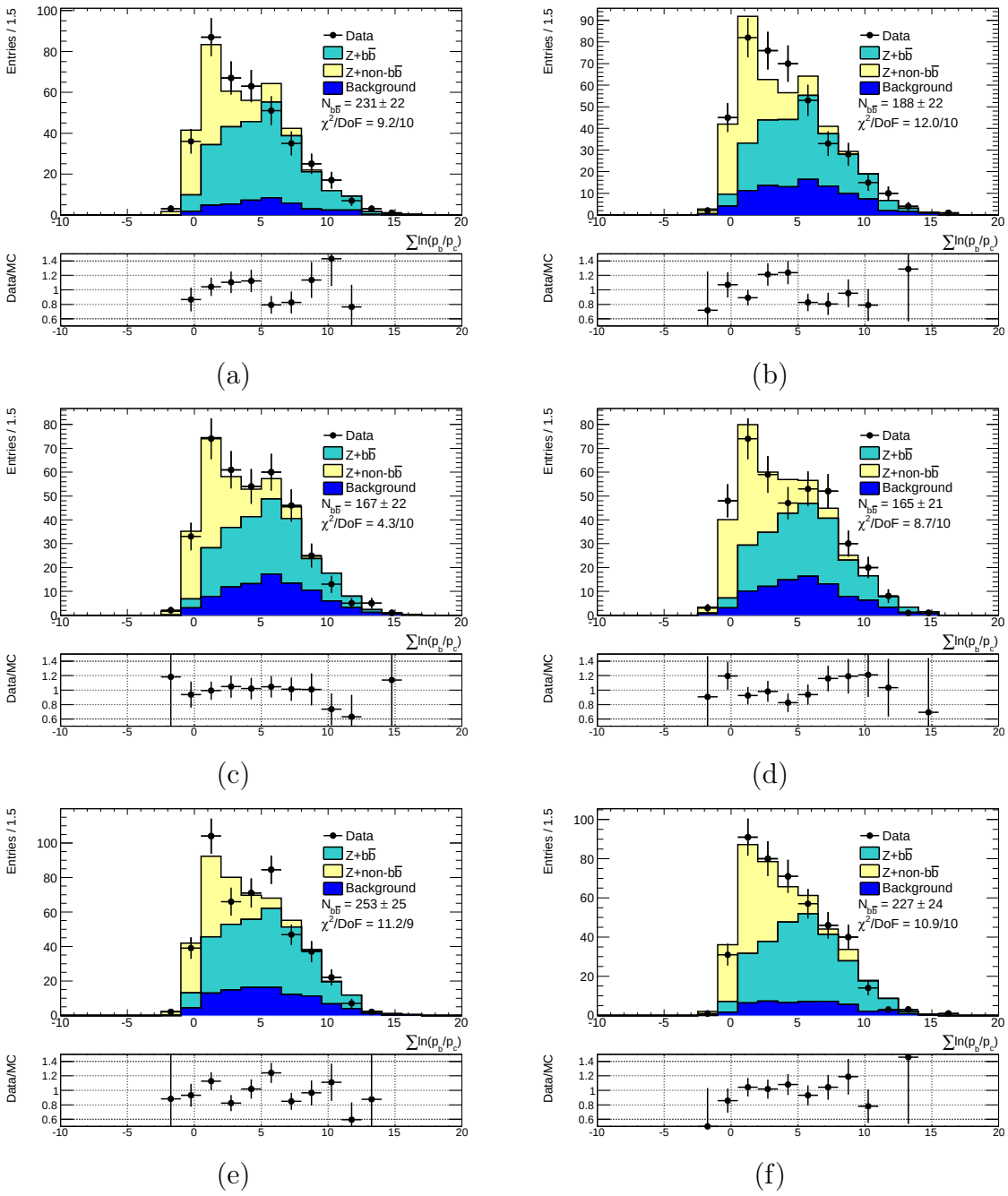


Figure A.3.: Result of fitting the number of true $b\bar{b}$ pairs, $n_{b\bar{b}}$, in bins of the angular separation of the two tagged b-jets for (a) $0.4 < \Delta R_{b\bar{b}} < 1.15$, (b) $1.15 < \Delta R_{b\bar{b}} < 1.9$, (c) $1.9 < \Delta R_{b\bar{b}} < 2.4$, (d) $2.4 < \Delta R_{b\bar{b}} < 2.8$, (e) $2.8 < \Delta R_{b\bar{b}} < 3.2$ and (f) $3.2 < \Delta R_{b\bar{b}} < 5$.

Colophon

This thesis was made in LaTeX 2_ε using the "hepthesis" class [108].

Bibliography

[1] ATLAS Collaboration, *Measurement of differential production cross-sections for a Z boson in association with b-jets in 7 TeV proton-proton collisions with the ATLAS detector*, JHEP **10** (2014) 141, arXiv:1407.3643 [hep-ex].

[2] S. Mikocki, *Precision QCD measurements at HERA: Jet production and determination*, Nuclear Physics B - Proceedings Supplements **245** (2013) 25 – 32.

[3] CDF, D0 Collaboration, D. Bandurin, *QCD measurements at the Tevatron*, in *International Workshop on Future Linear Colliders (LCWS11) Granada, Spain, September 26-30, 2011.*

[4] ATLAS, CMS Collaboration, P. Lenzi, *Precision QCD measurements at the LHC*, PoS **DIS2013** (2013) 007.

[5] CMS Collaboration, S. Chatrchyan et al., *Inclusive b-jet production in pp collisions at* $\sqrt{s} = 7$ *TeV*, JHEP **04** (2012) 084, arXiv:1202.4617 [hep-ex].

[6] F. Maltoni, G. Ridolfi, and M. Ubiali, *b-initiated processes at the LHC: a reappraisal*, JHEP **07** (2012) 022, arXiv:1203.6393 [hep-ph]. [Erratum: JHEP04,095(2013)].

[7] ATLAS Collaboration, *Search for strong production of supersymmetric particles in final states with missing transverse momentum and at least three b-jets at* $\sqrt{s}$*= 8 TeV proton-proton collisions with the ATLAS detector*, JHEP **10** (2014) 24, arXiv:1407.0600 [hep-ex].

[8] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, Phys. Rev. Lett. **100** (2008) 242001, arXiv:0802.2470 [hep-ph].

[9] UA1 Collaboration, G. Arnison et al., *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at s**(1/2) = 540-GeV*, Phys. Lett. **B122** (1983) 103–116.

[10] UA2 Collaboration, M. Banner et al., *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider*, Phys. Lett. **B122** (1983) 476–485.

[11] UA1 Collaboration, G. Arnison et al., *Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider*, Phys. Lett. **B126** (1983) 398–410.

[12] CDF Collaboration, F. Abe et al., *Observation of top quark production in $\bar{p}p$ collisions*, Phys. Rev. Lett. **74** (1995) 2626–2631, arXiv:hep-ex/9503002 [hep-ex].

[13] ATLAS Collaboration, *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].

[14] 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].

[15] R. Davis, D. S. Harmer, and K. C. Hoffman, *Search for Neutrinos from the Sun*, Phys. Rev. Lett. **20** (May, 1968) 1205–1209. http://link.aps.org/doi/10.1103/PhysRevLett.20.1205.

[16] Super-Kamiokande Collaboration, Y. Fukuda et al., *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562–1567, arXiv:hep-ex/9807003 [hep-ex].

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

[18] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (Aug, 1964) 321–323. http://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[19] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys.

Lett. **12** (1964) 132–133.

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

[21] Particle Data Group Collaboration, J. Beringer et al., *Review of Particle Physics (RPP)*, Phys. Rev. **D86** (2012) 010001.

[22] H1 Collaboration, T. Ahmed et al., *A Measurement of the proton structure function f2 (x, Q**2)*, Nucl. Phys. **B439** (1995) 471–502, arXiv:hep-ex/9503001 [hep-ex].

[23] ZEUS, H1 Collaboration, F. D. Aaron et al., *Combined Measurement and QCD Analysis of the Inclusive e+- p Scattering Cross Sections at HERA*, JHEP **01** (2010) 109, arXiv:0911.0884 [hep-ex].

[24] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972) 438–450. [Yad. Fiz.15,781(1972)].

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

[26] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, Sov. Phys. JETP **46** (1977) 641–653. [Zh. Eksp. Teor. Fiz.73,1216(1977)].

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

[28] R. S. Thorne and W. K. Tung, *PQCD Formulations with Heavy Quark Masses and Global Analysis*, in *Proceedings, workshop: HERA and the LHC workshop series on the implications of HERA for LHC physics.*

[29] N. Metropolis and S. Ulam, *The Monte Carlo Method*, Journal of the American Statistical Association **44** (1949) no. 247, pp. 335–341. http://www.jstor.org/stable/2280232.

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

[31] M. Beckingham, M. Duehrssen, E. Schmidt, M. Shapiro, M. Venturi, J. Virzi, I. Vivarelli, M. Werner, S. Yamamoto, and T. Yamanaka, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, Tech. Rep. ATL-PHYS-PUB-2010-013, CERN, Geneva, Oct, 2010. `http://cds.cern.ch/record/1300517`.

[32] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, *Parton Fragmentation and String Dynamics*, Phys. Rept. **97** (1983) 31–145.

[33] A. Buckley et al., *General-purpose event generators for LHC physics*, Phys. Rept. **504** (2011) 145–233, `arXiv:1101.2599 [hep-ph]`.

[34] D. Amati and G. Veneziano, *Preconfinement as a Property of Perturbative QCD*, Phys. Lett. **B83** (1979) 87.

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

[36] J. M. Butterworth et al., *The Tools and Monte Carlo working group Summary Report*, in *Physics at TeV colliders. Proceedings, 6th Workshop, dedicated to Thomas Binoth, Les Houches, France, June 8-26, 2009.*

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

[38] CMS Collaboration, G. L. Bayatian et al., *CMS technical design report, volume II: Physics performance*, J. Phys. **G34** (2007) 995–1579.

[39] O. S. BrÃijning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report*. CERN, Geneva, 2004. `https://cds.cern.ch/record/782076`.

[40] LHCb Collaboration, A. A. Alves, Jr. et al., *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.

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

[42] C. Lefevre, *The CERN accelerator complex*, `https://cds.cern.ch/record/1621894`, 2008.

[43] J. Pequenao, *Computer generated image of the whole ATLAS detector*, 2008.

[44] ATLAS Collaboration, *Alignment of the ATLAS Inner Detector and its Performance in 2012*, Tech. Rep. ATLAS-CONF-2014-047, CERN, Geneva, Jul, 2014. http://cds.cern.ch/record/1741021.

[45] J. Pequenao, *Computer generated image of the ATLAS calorimeter*, 2008.

[46] ATLAS Collaboration, *Updated Luminosity Determination in pp Collisions at* $\sqrt{s} = 7TeV$ *using the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2011-011, CERN, Geneva, 2011.

[47] ATLAS Collaboration. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults.

[48] ATLAS Collaboration, *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, Eur. Phys. J. **C72** (2012) 1909, arXiv:1110.3174 [hep-ex].

[49] J. Beringer et al, *Review of Particle Physics (RPP)*, Phys.Rev. **D86** (2012) 010001.

[50] ATLAS Collaboration, *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*, Tech. Rep. ATLAS-CONF-2011-102, CERN, Geneva, Jul, 2011. https://cds.cern.ch/record/1369219.

[51] ATLAS Collaboration, *Performance of the ATLAS Secondary Vertex b-tagging Algorithm in 7 TeV Collision Data*, Tech. Rep. ATLAS-CONF-2010-042, CERN, Geneva, Jul, 2010. https://cds.cern.ch/record/1277682.

[52] ATLAS Collaboration, *Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data*, Tech. Rep. ATLAS-CONF-2014-046, CERN, Geneva, Jul, 2014. http://cds.cern.ch/record/1741020.

[53] ATLAS Collaboration, *Search for the* $b\bar{b}$ *decay of the Standard Model Higgs boson in associated* $(W/Z)H$ *production with the ATLAS detector*, JHEP **01** (2015) 069, arXiv:1409.6212 [hep-ex].

[54] A. Denner, S. Heinemeyer, I. Puljak, D. Rebuzzi, and M. Spira, *Standard*

Model Higgs-Boson Branching Ratios with Uncertainties, Eur. Phys. J. **C71** (2011) 1753, arXiv:1107.5909 [hep-ph].

[55] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with a vector boson and decaying to a b-quark pair using up to 4.7 fb^{-1} of pp collision data at 7 TeV with the ATLAS detector at the LHC*, Tech. Rep. ATLAS-CONF-2012-015, CERN, Geneva, Mar, 2012. https://cds.cern.ch/record/1429664.

[56] CDF Collaboration, T. Aaltonen et al., *Measurement of cross sections for b Jet production in events with a Z boson in $p\bar{p}$ collisions at $\sqrt{s}$=1.96 TeV*, Phys. Rev. **D 79** (2009) 052008, arXiv:0812.4458 [hep-ex].

[57] D0 Collaboration, V. Abazov et al., *Measurement of the ratio of differential cross sections $\sigma(p\bar{p} \to Z + b\ jet)/\sigma(p\bar{p} \to Z + jet)$ in $p\bar{p}$ collisions at $\sqrt{s}$=1.96 TeV*, Phys. Rev. **D 87** (2013) 092010, arXiv:1301.2233 [hep-ex].

[58] ATLAS Collaboration, *Measurement of the cross-section for b-jets produced in association with a Z boson at $\sqrt{s}$=7 TeV with the ATLAS detector*, Phys. Lett. **B 706** (2012) 295, arXiv:1109.1403 [hep-ex].

[59] CMS Collaboration, *Measurement of the production cross sections for a Z boson and one or more b jets in pp collisions at $\sqrt{s}$=7 TeV*, submitted to JHEP , arXiv:1402.1521 [hep-ex].

[60] M. L. Mangano et al., *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **0307** (2003) 001, arXiv:hep-ph/0206293.

[61] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, et al., *Implications of CTEQ global analysis for collider observables*, Phys.Rev. **D78** (2008) 013004, arXiv:0802.0007 [hep-ph].

[62] ATLAS Collaboration, *New ATLAS event generator tunes to 2010 data*, ATL-PHYS-PUB-2011-008. http://cds.cern.ch/record/1345343.

[63] G. Corcella et al., *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **0101** (2001) 010, arXiv:hep-ph/0011363.

[64] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, *Multiparton*

interactions in photoproduction at HERA, Z. Phys. **C 72** (1996) 637, arXiv:hep-ph/9601371.

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

[66] H. L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. **D 82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[67] B. P. Kersevan and E. Richter-Was, *The Monte Carlo event generator AcerMC versions 2.0 to 3.8 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1*, Comput. Phys. Commun. **184** (2013) 919, arXiv:hep-ph/0405247.

[68] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **0605** (2006) 026, arXiv:hep-ph/0603175.

[69] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A 462** (2001) 152–155.

[70] ATLAS Collaboration, *Performance of the ATLAS muon trigger in 2011*, Tech. Rep. ATLAS-CONF-2012-099, CERN, Geneva, Jul, 2012. https://cds.cern.ch/record/1462601.

[71] ATLAS Collaboration, *A measurement of the ATLAS muon reconstruction and trigger efficiency using J/psi decays*, Tech. Rep. ATLAS-CONF-2011-021, CERN, Geneva, Mar, 2011. https://cds.cern.ch/record/1336750.

[72] ATLAS Collaboration, *Performance of the ATLAS Electron and Photon Trigger in p-p Collisions at sqrts = 7 TeV in 2011*, Tech. Rep. ATLAS-CONF-2012-048, CERN, Geneva, May, 2012. https://cds.cern.ch/record/1450089.

[73] ATLAS Collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data*, Eur.Phys.J. **C74** (2014) , arXiv:hep-ex/1047.3935.

[74] ATLAS Collaboration, *Measuring the b-tag efficiency in a top-pair sample with 4.7 fb^{-1} of data from the ATLAS detector*, Tech. Rep. ATLAS-CONF-2012-097, CERN, Geneva, Jul, 2012.

https://cds.cern.ch/record/1460443.

[75] ATLAS Collaboration, *Measurement of the Mistag Rate with 5 fb^1 of Data Collected by the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2012-040, CERN, Geneva, Mar, 2012. https://cds.cern.ch/record/1435194.

[76] ATLAS Collaboration, *b-jet tagging calibration on c-jets containing D^{*+} mesons*, Tech. Rep. ATLAS-CONF-2012-039, CERN, Geneva, Mar, 2012. https://cds.cern.ch/record/1435193.

[77] ATLAS Collaboration, *Measurement of the cross-section for b-jets produced in association with a Z boson at $\sqrt{s}$=7 TeV with the ATLAS detector*, Phys. Lett. B **706** (Sep, 2011) 295–313. 20 p. https://cds.cern.ch/record/1380892.

[78] ATLAS Collaboration, *Jet energy scale and its systematic uncertainty in proton-proton collisions at sqrt(s)=7 TeV with ATLAS 2011 data*, Tech. Rep. ATLAS-CONF-2013-004, CERN, Geneva, Jan, 2013. https://cds.cern.ch/record/1509552.

[79] ATLAS Collaboration, *Jet energy resolution and selection efficiency relative to track jets from in-situ techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy sqrts = 7 TeV*, Tech. Rep. ATLAS-CONF-2010-054, CERN, Geneva, Jul, 2010. https://cds.cern.ch/record/1281311.

[80] ATLAS Collaboration, *Performance of Missing Transverse Momentum Reconstruction in Proton-Proton Collisions at 7 TeV with ATLAS*, Eur. Phys. J. **C72** (2012) 1844, arXiv:1108.5602 [hep-ex].

[81] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv:0811.4622 [hep-ph].

[82] S. Moch and P. Uwer, *Theoretical status and prospects for top-quark pair production at hadron colliders*, Phys. Rev. **D78** (2008) 034003, arXiv:0804.1476 [hep-ph].

[83] 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].

[84] ATLAS Collaboration, *A measurement of the inclusive $W^{\pm}$ and Z/γ^{*} cross sections in the e and μ decay channels in pp collisions at $\sqrt{s}$=7 TeV with the ATLAS detector*, Phys. Rev. D **85** (2012) 072004, arXiv:1109.5141 [hep-ex].

[85] ATLAS Collaboration, *Measurement of the inclusive and dijet cross-sections of b-jets in pp collisions at $\sqrt{s}$=7 TeV with the ATLAS detector*, Eur. Phys. J. **C 71** (2011) 1846, arXiv:1109.6833 [hep-ex].

[86] ATLAS Collaboration, *Measurement of hard double-parton interactions in $W(\rightarrow l\nu)$ + 2 jet events at $\sqrt{s}$=7 TeV with the ATLAS detector*, New J. Phys. **15** (2013) 033038, arXiv:1301.6872 [hep-ex].

[87] J. M. Campbell and R. Ellis, *MCFM for the Tevatron and the LHC*, Nucl.Phys.Proc.Suppl. **205-206** (2010) 10–15, arXiv:1007.3492 [hep-ph].

[88] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, et al., *New parton distributions for collider physics*, Phys.Rev. **D82** (2010) 074024, arXiv:1007.2241 [hep-ph].

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

[90] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A Precision tool for QED corrections in Z and W decays*, Eur.Phys.J. **C45** (2006) 97–107, arXiv:hep-ph/0506026 [hep-ph].

[91] M. Bahr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. **C58** (2008) 639–707, arXiv:0803.0883 [hep-ph].

[92] UA1 Collaboration, C. Albajar et al., *Beauty Production at the CERN Proton - anti-Proton Collider. 1.*, Phys. Lett. **B186** (1987) 237–246.

[93] D0 Collaboration, B. Abbott et al., *The $b\bar{b}$ production cross section and angular correlations in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Lett. **B487** (2000) 264–272, arXiv:hep-ex/9905024 [hep-ex].

[94] D0 Collaboration, B. Abbott et al., *Cross section for b jet production in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. Lett. **85** (2000) 5068–5073, arXiv:hep-ex/0008021 [hep-ex].

[95] ATLAS Collaboration, *Measurement of the inclusive and dijet cross-sections of b^- jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Eur. Phys. J. **C71** (2011) 1846, arXiv:1109.6833 [hep-ex].

[96] ATLAS Collaboration, *Identification and Tagging of Double b-hadron jets with the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2012-100, CERN, Geneva, Jul, 2012. https://cds.cern.ch/record/1462603.

[97] CMS Collaboration, V. Khachatryan et al., *Inclusive b-hadron production cross section with muons in pp collisions at $\sqrt{s} = 7$ TeV*, JHEP **03** (2011) 090, arXiv:1101.3512 [hep-ex].

[98] CMS Collaboration, S. Chatrchyan et al., *Measurement of the cross section and angular correlations for associated production of a Z boson with b hadrons in pp collisions at $\sqrt{s} = 7$ TeV*, JHEP **12** (2013) 039, arXiv:1310.1349 [hep-ex].

[99] 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].

[100] T. Sjostrand, *PYTHIA 8 Status Report*, in *Proceedings, workshop: HERA and the LHC workshop series on the implications of HERA for LHC physics*.

[101] S. Gieseke, C. Rohr, and A. Siodmok, *Colour reconnections in Herwig++*, Eur. Phys. J. **C72** (2012) 2225, arXiv:1206.0041 [hep-ph].

[102] T. Skwarnicki, *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*. PhD thesis, Cracow, INP, 1986. http://lss.fnal.gov/cgi-bin/find_paper.pl?other/thesis/skwarnicki.pdf.

[103] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039.

[104] K. S. Cranmer, *Kernel estimation in high-energy physics*, Comput. Phys. Commun. **136** (2001) 198–207, arXiv:hep-ex/0011057 [hep-ex].

[105] ATLAS Collaboration, *Improved luminosity determination in pp collisions at*

$\sqrt{s}$ = *7 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **73** (Feb, 2013) 2518. 40 p. https://cds.cern.ch/record/1517411.

[106] P. Speckmayer and C. Grefe, *Comparison of the Performance of Tungsten and Steel Hadronic Sampling Calorimeters*, . https://cds.cern.ch/record/1443830.

[107] *Expected performance of the ATLAS b-tagging algorithms in Run-2*, Tech. Rep. ATL-PHYS-PUB-2015-022, CERN, Geneva, Jul, 2015. https://cds.cern.ch/record/2037697.

[108] A. Buckley, *A class for typesetting academic theses*, 2010.