Gluon splitting to b-quark pairs in proton-proton collisions at $\sqrt{s} = 8\,\text{TeV}$ with ATLAS

Giles C. Strong*
MSc. Physics and Astronomy
University of Glasgow

Primary Supervisor: Dr Aidan Robson†
Secondary Supervisor: Dr Andy Buckley†
External Supervision: Dr Gavin Hesketh‡and Dr Josh McFayden‡

5th September 2015

Abstract

Gluon splitting to $b\bar{b}$ is shown to be an important process which is currently not well modelled by Monte Carlo event generators. In order to understand the $g \to b\bar{b}$ process in Monte Carlo generation, Pythia 8 and Sherpa are compared in terms of their production methods of $b\bar{b}$ pairs, and the parton shower approximation of gluon splitting is compared to its matrix element calculation using Sherpa. The parton shower in Sherpa shows reasonable agreement with the matrix element, but large differences are seen between the predictions of Pythia 8 and Sherpa, particularly in their $b\bar{b}$ pair production through the parton shower. In order to compare the generators to data, a non-prompt di-J/ψ sample is prepared via a four-dimensional simultaneous binned-fit to data collected by the ATLAS detector in the Large Hadron Collider at CERN. The Monte Carlo result distributions are then smeared based on $B \to J/\psi$ correlation functions derived from the analysis of a Pythia 8B J/ψ sample. The generator and collider samples are then compared in terms of the distribution of the angular separations between J/ψ pairs. It is found that of the two generators: Pythia 8 shows the best agreement with the data, and Sherpa primarily disagrees in the region of low angular-separation, where parton shower production is seen to dominate.

*giles.c.strong@gmail.com
†University of Glasgow
‡University College London

Contents

1 Introduction

1.1 Background and motivation

The hadronic collisions which occur at high-energy particle colliders, such as the Large Hadron Collider (LHC) at CERN, invariably produce large quantities of hadronic final states resulting from quark-pair production through quantum chromodynamic (QCD) interactions. In perturbative QCD (pQCD[1]), gluons are regarded as being responsible for the dominant mode of QCD interactions, corresponding to the generators of the theory. If a gluon has enough energy it can split to form a pair of bottom quarks providing a background process for Higgs boson searches in the $H \to b\bar{b}$ channel. Because of this, it is important to make sure that the properties of the $g \to b\bar{b}$ process are well modelled by Monte Carlo event generators in order that this background might be correctly accounted for in such searches. The $g \to b\bar{b}$ process also has a considerable contribution in $W/Z + b\bar{b}$.

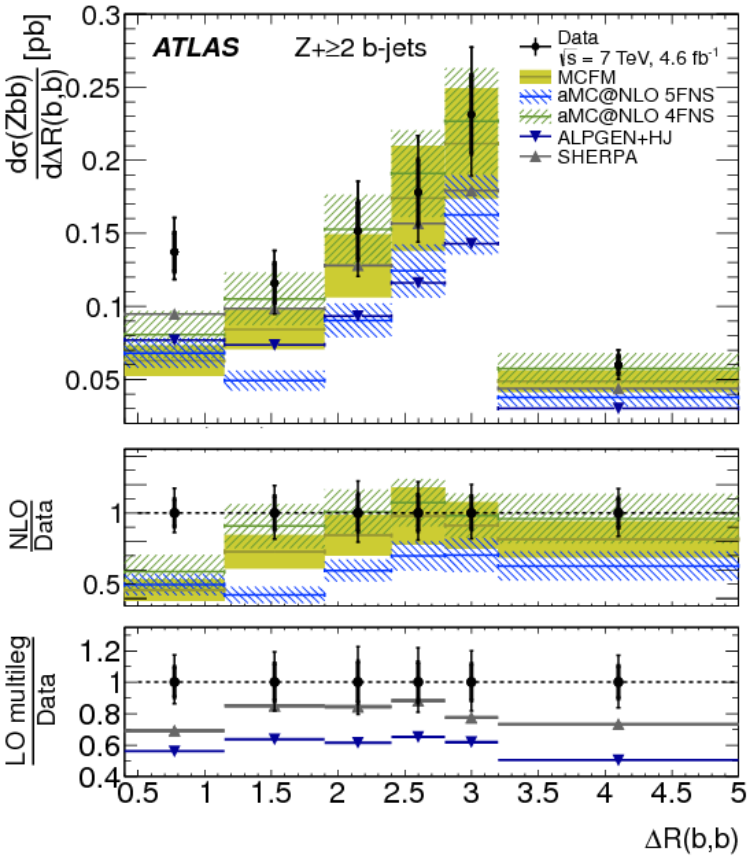
Studies [1, 2] by the two multipurpose detectors at the LHC, ATLAS [3] and CMS [4], investigating associated production of Z bosons with b quarks indicate that there are discrepancies between collider data and Monte Carlo predictions, particularly when the b quarks have a narrow separation as shown in Figure 1.

The LHC has recently undergone upgrades to perform at higher luminosities and centre-of-mass energies. As data taking begins it is expected that there will be large quantities of gluons produced in the collisions which are highly boosted (have a high energy or momentum relative to the detector's rest frame). The large boost will mean that if they split, their products will have a narrow angular separation, placing them in the low ΔR region. Given the importance of Monte Carlo data in high-energy particle-physics analysis it is essential that the generators used accurately reflect the real data.

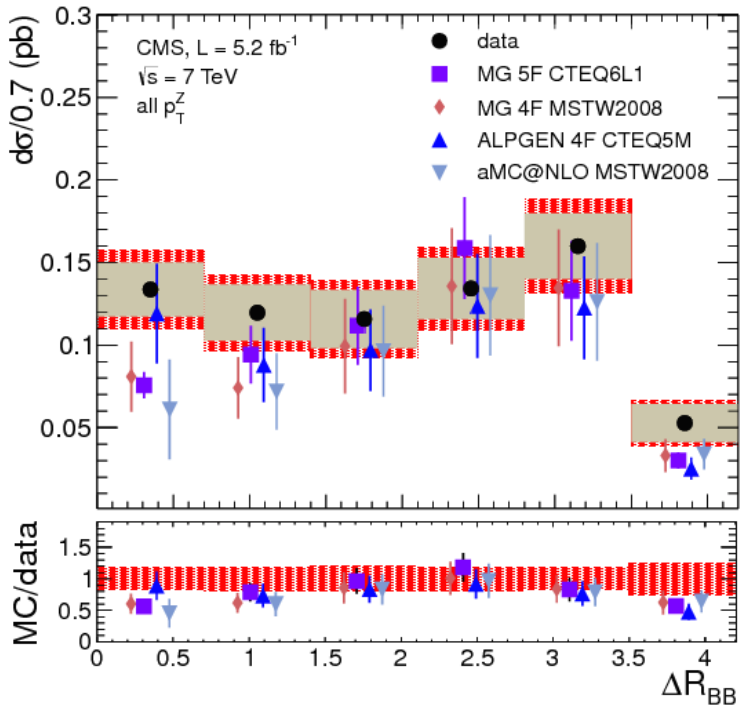
Because of the small angles involved in the problem, it is difficult to construct an analysis using jets. In the ATLAS study jets are used and, as shown in Figure 1a, resolution of $\Delta R_{b,b}$ only extends down to 0.5. The CMS detector's inner tracker has a high enough resolution that it is able to reconstruct b hadron vertices via an inclusive vertex finding technique, and using these as proxies to b quarks extends sensitivity down to $\Delta R_{B,B} = 0$, as shown in Figure 1b.

The study presented in this thesis makes use of the full 2012 $\sqrt{s} = 8\,\text{TeV}$ ATLAS dataset and uses non-prompt, i.e. produced by a b hadron, J/ψs as proxies to b hadrons, which are in turn used as proxies to b quarks. The angular variables of di-J/ψ events are analysed and compared to the leading-order predictions of two Monte Carlo generators: Pythia 8 [5, 6] and Sherpa [7, 8, 9, 10]. J/ψs are chosen because of the clean dimuon signal they can produce in the ATLAS detector. Additionally, an in depth investigation into the characteristics of the production mechanisms for b-quark pairs in Monte Carlo is presented.

[1] An approach to QCD calculations utilising perturbative techniques applicable over short distances or at high energy, where the strong coupling constant is small



(a) ATLAS: $\Delta R_{b,b}$ distribution for the hardest two b jets produced in association with a Z boson compared to several Monte Carlo predictions. Data selected from 2011 $\sqrt{s} = 7\,\text{TeV}$ ATLAS dataset. Plot taken from Ref. [1]



(b) CMS: $\Delta R_{B,B}$ distribution for b-hadron pairs produced in association with a Z boson compared to several Monte Carlo predictions. Data selected from 2011 $\sqrt{s} = 7\,\text{TeV}$ CMS dataset. Plot taken from Ref. [2]

Figure 1

1.2 Thesis overview

Initially, the currently accepted model for particle interactions, the Standard Model, is described and followed by a series of descriptions and definitions of various particle-physics variables, tools, and instruments (Sections 2 and 3).

Next, the production channels for $b\bar{b}$ pairs in Monte Carlo generators are investigated with a focus on generator comparison (Section 4).

In Section 5 SHERPA's parton shower approximation of $g \to b\bar{b}$ is compared to the explicitly calculated matrix-element for the process by forcing the simulation into a phase-space-enhanced region.

In order to compare the Monte Carlo results, which use the angular variables of b hadrons, to the J/ψ data, the investigation in Section 6 tests the suitability of J/ψs as b-hadron proxies and derives $B \to J/\psi$ correlation functions which are used later to smear the Monte Carlo distributions to account for the $B \to J/\psi$ decay.

The non-prompt di-J/ψ comparison sample is prepared in Section 7 using a four-dimensional simultaneous binned-fit to reweight an ATLAS data sample according to the probability of a given event containing two non-prompt J/ψs.

Comparison takes place in Section 8 where Monte Carlo samples for inclusive production of $b\bar{b}$ pairs are analysed and smeared using the correlation functions derived in Section 6 and then compared to the non-prompt di-J/ψ sample distributions.

The thesis then concludes with a short summary of the investigation in Section 9.

1.3 Notes

In this investigation, b hadrons are used as proxies to b quarks and no attempt is made to extrapolate angular distributions back to parton level as this would involve modelling heavy-quark fragmentation and hadronisation which is beyond the scope of this thesis.

Preparation of the di-J/ψ comparison sample made great use of the fit constructed and developed by Dr Gavin Hesketh and Dr Josh McFayden of University College London; their work, and the work of others will be explicitly noted when discussed in the write up.

A lot of code for data preparation and analysis was written by the author for this investigation and parts of this, as well as the exact runcards used for the Monte Carlo sample generation are available on request at giles.c.strong@gmail.com.

2 The Standard Model

The Standard Model (SM) is a $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge theory[2] developed to describe the interactions between fundamental particles [11]. The Lagrangian density for this theory according to Glashow-Weinberg-Salam is given in Equation 2.1 [12]:

$$\mathcal{L}_{SM} = \mathcal{L}_{QCD} + \mathcal{L}_{EW} = \mathcal{L}_W + \mathcal{L}_B + \sum_\ell \mathcal{L}_\ell + \sum_\ell \mathcal{L}_Y(\ell) + \mathcal{L}(q) + \mathcal{L}_Y(q) + \mathcal{L}_\Phi + \mathcal{L}_{QCD} + ..., \tag{2.1}$$

where "..." are gauge-fixing and ghost terms, $\mathcal{L}_{QCD}$ and $\mathcal{L}_{EW}$ are the Lagrangian densities for the QCD and electroweak (EW) sectors respectively, the remaining terms are explained in Table 1 below.

[2]The unitary group U(n) is the group of all $n \times n$ unitary-matrices, and the special unitary group SU(n) is the group of all $n \times n$ unitary-matrices with determinant one

$\mathcal{L}_W$	Kinetic-energy terms for weak bosons
$\mathcal{L}_B$	Kinetic-energy terms for the electromagnetic boson
$\sum_\ell \mathcal{L}_\ell$	Lepton kinematics and EW interactions
$\sum_\ell \mathcal{L}_Y(\ell)$	Yukawa mass-terms for leptons
$\mathcal{L}(q)$	Quark kinematics and EW interactions
$\mathcal{L}_Y(q)$	Yukawa mass-terms for quarks
$\mathcal{L}_\Phi$	Higgs-boson kinematics and EW interactions

Table 1: Summary of terms contributing to the EW sector [12].

As can be seen in Equation 2.1, the Standard Model consists of two main parts: QCD, which describes the interactions caused by the strong force, and the EW sector, which describes the interactions caused by the weak and electromagnetic forces.

QCD has the greatest influence on the effects investigated in this thesis, however, for completeness, the EW sector will also be described.

2.1 Electroweak sector

The Glashow-Weinberg-Salam model for EW interactions takes the form of a $SU(2)_L \times U(1)_Y$ gauge theory and follows the Lagrangian density shown in Equation 2.2:

$$\mathcal{L}_{EW} = \mathcal{L}_W + \mathcal{L}_B + \sum_\ell \mathcal{L}_\ell + \sum_\ell \mathcal{L}_Y(\ell) + \mathcal{L}(q) + \mathcal{L}_Y(q) + \mathcal{L}_\Phi, \tag{2.2}$$

which consists of several sectors described in Table 1.

2.1.1 Electroweak gauge bosons

There are four gauge bosons in the EW sector: the W^+ and W^- bosons, which mediate charge-current interactions; and the Z and γ bosons, which mediate neutral-current interactions. The W and Z bosons propagate the weak force, and the photon, (γ), propagates the electromagnetic force. They are all vector bosons having spin 1.

Whilst the photon is massless, the experimental results indicate that weak bosons are massive and current measurements are: $M_W = 80.385\,\text{GeV}$ and $M_Z = 91.1876\,\text{GeV}$ [13]. The presence of massive gauge bosons initially led to problems as the addition of mass terms caused the Lagrangian density to become non-gauge-invariant meaning that the physical results of using it depended on the choice of gauge, (redundant degrees of freedom, such as choice of inertial frame). The suggested solution "fixing" this was via spontaneous symmetry-breaking (SSB) through the Higgs-Kibble-Guralnik-Hagen-Brout-Englert mechanism (the Higgs mechanism) [14, 15, 16].

2.1.2 The Higgs mechanism

The Higgs mechanism involves specifying the vacuum state with a scalar, (spin-0), field ϕ. The mass term for a scalar field is $-\mu^2\phi\phi^*$ [12] and for $-\mu^2 < 0$ the vacuum state is no longer located at $\phi = 0$ but occupies a set of degenerate minima in a circle about $\phi = 0$.

ϕ can be rewritten in terms of two scalar fields: the massive Higgs field, corresponding to radial excitations around the circle of minima, and the massless Goldstone fields, corresponding to excitations along the circle of minima. Picking a specific vacuum state from the set of degenerate minima causes spontaneous symmetry breaking and by working in the unitary gauge

it can be seen that the Goldstone fields are "eaten" by the W and Z bosons, allowing them to acquire mass without violating gauge invariance [12]. The vacuum state is chosen such that the photon remains massless.

The presence of SSB in the SM also allows quarks and leptons to have mass terms through the Yukawa interaction without violating gauge invariance [12].

The Higgs boson couples to particles according to their mass, with a stronger coupling to heavier particles [11]; this means that it can also self-couple.

A particle with a mass of 126 GeV corresponding to a SM Higgs boson was recently discovered with a significance level of 5.9 σ at the LHC [17].

2.1.3 Leptons

Leptons are spin-$1/2$ elementary particles with a lepton number of $+1$ and can be represented as a chiral weak isospin doublet:

$$\chi_L^f = \begin{pmatrix} \nu_f \\ e_f \end{pmatrix}_L,$$

where ν_f are electrically-neutral particles known as neutrinos, and e_f are the charged leptons, which carry an electric charge of $-1e$, where e is the charge of the electron, which is the lightest charged-lepton. The f indicates the generation, and three generations have so far been discovered [13]. Transformations between charged leptons and neutrinos are known as charge-current interactions and are mediated by the W^+ and W^- bosons. Observations show that the W bosons only couple to left-handed fermions (spin opposite to momentum) hence the L in the chiral doublet [11].

Scatterings of leptons, and creation and annihilation of lepton-antilepton[3] pairs are known as neutral-current interactions and are mediated by the photon and the Z boson, though the photon couples to electric charge and so only the Z boson can mediate neutral-current interactions involving neutrinos.

The second-generation charged-lepton is known as the muon (μ) and it is used in this investigation to detect the decays of J/ψ particles. It has a mass of 105.658 371 5 MeV and a mean lifetime of 2.196 981 1 µs [13].

A lepton of generation f carries a generation number of $+1f$. In all EW interactions in the lepton sector, lepton number and lepton generation number must be conserved.

2.1.4 Quarks

All quarks are of spin $1/2$ and carry a baryon number of $+1/3$. Quarks have two types: up-type and down-type. Up-type quarks carry an electrical charge (Q) of $+2/3$ and an isospin (I_3) of $+1/2$. Down-type quarks have opposite isospin and Q $= -1/3$ [13]. Antiquarks carry charge, baryon number, and isospin opposite to their quark counterparts, but are of the same mass and spin. The two types of quarks can be represented as an isospin doublet:

$$\chi_L^f = \begin{pmatrix} U_f \\ D'_f \end{pmatrix},$$

where f denotes the generation number of the quark, and D'_f has been rotated in flavour-space:

$$D'_f = \sum_{f'} V_{ff'} D_{f'},$$

[3]Antileptons, the antiparticle to leptons, are of the same mass and spin as their corresponding particle, but carry opposite charge, lepton number, and lepton generation number

Quark	Charge [$e/3$]	Generation	Mass [MeV]
Up	2	1	2.3
Down	-1	1	4.8
Charm	2	2	1275
Strange	-1	2	95
Top	2	3	173 210
Bottom	-1	3	4180

Table 2: Summary of quark properties [13].

where V is the Cabbibo-Kobayashi-Maskawa (CKM) matrix:

$$V_{CKM} = \begin{pmatrix} |V_{ud}| = 1 - \frac{\lambda^2}{2} & |V_{us}| = \lambda & |V_{ub}| = A\lambda^3\phi \\ |V_{cd}| = -\lambda & |V_{cs}| = 1 - \frac{\lambda^2}{2} & |V_{cb}| = A\lambda^2 \\ |V_{td}| = A\lambda^3\left(1-\phi\right) & |V_{ts}| = -A\lambda^2 & |V_{tb}| = 1 \end{pmatrix} + \mathcal{O}(\lambda^4),$$

where $\phi = \rho - i\eta, \lambda = 0.22537, A = 0.814, \rho = 0.117$, and $\eta = 0.353$ [13].

Currently three generations have been confirmed, resulting in six quark flavours. The properties of these are summarised in Table 2. In this thesis quark flavours are sometimes abbreviated to just the first letter, e.g. b quark, and uud.

In the EW sector quarks undergo interactions with all four EW bosons conserving charge (electric and colour) but not necessarily flavour and indeed it is the EW interactions which allow the rotations in flavour space mentioned above.

2.2 Quantum chromodynamics

QCD is a non-Abelian SU(N_c) gauge theory whose interactions are described by the Lagrangian density shown in Equation 2.3 [18]:

$$\begin{aligned} \mathcal{L}_{QCD} &= \bar{\psi}\left(i\not{\partial} - m\right)\psi - \frac{1}{4}\left(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu\right)^2 + gA^a_\mu\bar{\psi}\gamma^\mu T^a\psi \\ &\quad -gf^{abc}\left(\partial_\mu A^a_\nu\right)A^{\mu b}A^{\nu c} - \frac{1}{4}g^2\left(f^{eab}A^a_\mu A^b_\nu\right)\left(f^{ecd}A^{\mu c}A^{\nu d}\right), \end{aligned} \tag{2.3}$$

where g is the gluon coupling-constant, f^{abc} and T^a are the structure constants and generators of the group respectively, the matter fields ψ are referred to as quarks, and the gluon fields (A^a_μ) act as propagators for the strong nuclear force between quarks,

N_c is the number of colours in QCD. Colour charge was introduced to explain the detection of the Δ^{++} baryon, a double-charged, spin-$^3/_2$ particle consisting of three up quarks. The quantum numbers of the up quarks would all have to be equal for this particle to exist, however this would violate Fermi-Dirac statistics which require that non-integer-spin particles (fermions) do not exist in the same state. To resolve this issue, an additional quantum number, colour charge, was assigned to quarks by Han and Nambu, and Greenberg, in 1965 [18]. Experimental results [13] suggest that $N_c = 3$ resulting in $N_c^2 - 1 = 8$ generators of the form $T_a = \lambda_a/2$, where $a \in \{1, 2, 3, 4, 5, 6, 7, 8\}$, and λ_a are the Gell-Mann λ matrices. The generators of SU(3) satisfy the Lie algebra: $[T_a, T_b] = if_{abc}T_c$ [12]

$\frac{1}{\sqrt{2}}\left(r\bar{b}+b\bar{r}\right)$	$\frac{-i}{\sqrt{2}}\left(r\bar{b}+b\bar{r}\right)$
$\frac{1}{\sqrt{2}}\left(r\bar{g}+g\bar{r}\right)$	$\frac{-i}{\sqrt{2}}\left(r\bar{g}+g\bar{r}\right)$
$\frac{1}{\sqrt{2}}\left(g\bar{b}+b\bar{g}\right)$	$\frac{-i}{\sqrt{2}}\left(g\bar{b}+b\bar{g}\right)$
$\frac{1}{\sqrt{2}}\left(r\bar{r}-b\bar{b}\right)$	$\frac{1}{\sqrt{6}}\left(r\bar{r}+b\bar{b}-2g\bar{g}\right)$

Table 3: SU(3)'s eight possible gluons in colour-octet representation [19].

2.2.1 Quarks

The quark fields in QCD can be formed, according to colour charge, into triplets of the form:

$$\psi(\tilde{x}) = \begin{pmatrix} \psi_{Red}(\tilde{x}) \\ \psi_{Green}(\tilde{x}) \\ \psi_{Blue}(\tilde{x}) \end{pmatrix}.$$

Requiring $\mathcal{L}_{QCD}$ to be invariant under local SU(3) gauge-transformations means that $\psi(\tilde{x})$ must transform as shown by:

$$\psi(\tilde{x}) \rightarrow U(\tilde{x})\,\psi(\tilde{x}) = e^{iT^a\alpha_a(\tilde{x})}\psi(\tilde{x})\,.$$

2.2.2 Gluons

The gauge fields A_μ in SU(N_c) are built from linear combinations of the generators of SU(N_c) and the gluon fields: $A_\mu = \sum_a^{N_c^2-1} A_\mu^a T_a$. So for $N_c = 3$ there are eight possible gluons. Each of these gluons carries a unique combination of colour and anticolour added in a quantum superposition, e.g. red-antiblue + blue-antired. A full listing of the possible gluon states as a colour octet is shown in Table 3. If QCD were formed using the group U(N_c) then there could also exist a colour singlet state for gluons: $1/\sqrt{3}\left(r\bar{r}+b\bar{b}+g\bar{g}\right)$. This state is colourless and would allow for long-range gluon interaction, however since it is observed that there are no long-range gluon interactions, QCD is chosen to be formed using SU(N_c) which does not contain the colour singlet state for gluons.

Unlike photons, which couple to electric charge, but are themselves uncharged, gluons, carrying colour charge, can self-couple as evidenced by the terms

$$-gf^{abc}\left(\partial_\mu A_\nu^a\right)A^{\mu b}A^{\nu c} - \frac{1}{4}g^2\left(f^{eab}A_\mu^a A_\nu^b\right)\left(f^{ecd}A^{\mu c}A^{\nu d}\right)$$

in Equation 2.3. These allow for 3- and 4-point interactions between gluons. Since both quarks and gluons can radiate gluons, the presence of a parton (collective name for quarks and gluons) quickly leads to a parton shower in which radiated gluons subsequently radiate more gluons or produce quarks pairs. Parton showers will be discussed further in Section 3.7.4.

Interactions between quarks and gluons must conserve charge (both electric and colour) and flavour.

In terms of fundamental properties, gluons appear in the theory as massless and have spin 1 so follow bosonic statistics.

2.2.3 Hadrons

QCD exhibits a property called colour confinement, which prevents the existence of isolated colour-charged particles. The current understanding is that, unlike an electric field between a pair of particles, the strength of which diminishes with distance, the gluon field's strength

Particle	Constituents	Charge	Mass [MeV]	Mean Lifetime [ps]
B^+	$u\bar{b}$	+1	5279.26	1.638
B^0	$d\bar{b}$	0	5279.58	1.519
B_s^0	$s\bar{b}$	0	5366.77	1.512
B_c^+	$c\bar{b}$	+1	6275.6	0.452

Table 4: Summary of B meson properties [13].

between a quark pair remains constant with distance [18] and so more energy must be put in to separate the pair further. This continues until enough energy has been put into the system that another quark pair can be formed to bind to the previous quark pair at a smaller separation.

The lack of isolated quarks results in the formation of particles called hadrons which consist of: mesons (two quarks in a bound state) and baryons (three quarks in a bound state). Perhaps the most well-known baryons are the proton and neutron, consisting of *uud*, and *udd*, respectively.

Mesons, as stated earlier, are a bound state of two quarks, and are categorised into flavour families according to the heaviest quark in the pair. The family of primary importance in this investigation are the B mesons, which contain a bottom quark bound to either an up, down, strange, or charm quark [13]. Their properties are summarised in Table 4.

Another two-quark particle of importance in this investigation is the J/ψ, a form of charmonium, consisting of a bound state of charm and anticharm quarks ($c\bar{c}$). The J/ψ has a mass of 3.096 916 GeV [13] and a mean lifetime of 7.2×10^{-9} ps. Of particular use is its high branching fraction (relative probability for a certain decay to occur) to a pair of muons of 5.961% [13], since, as will be seen in Section 3.6, the ATLAS experiment has detectors and algorithms available which are geared towards the identification of muons.

3 Definitions

3.1 Natural units

Natural units will be used in this thesis. These are defined by setting c, the speed of light in vacuo, and $\hbar$, the reduced Planck constant, to be equal to one. Kinematic properties of particles can then be quoted in terms of electron volts (eV).

In this scheme mass, energy, and momentum are measured in eV, length and time in eV^{-1}, force in eV^2, and velocity, angular momentum, and charge are dimensionless. Whilst cross section can be measured in terms of eV^{-2}, it is more common to use the units of barns (b).

3.2 Coordinate system and common variables

The ATLAS detector operates on a three-dimensional Cartesian coordinate system, the origin of which is defined to be at the interaction point. The z-axis points along to the beam axis, the x-axis points perpendicular to the beam axis towards the centre of the LHC ring, and the y-axis points perpendicularly-upwards [3].

Spherical coordinates are also used, with ϕ being an azimuthal angle measured in the plane perpendicular to the beam axis, and θ being the angle from the beam axis.

Pseudorapidity, η, defined as:

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right), \tag{3.1}$$

it is a purely geometrical measure of position from the beam axis in the bounds of $\pm\infty$, with 0 being perpendicular to the beam axis, and ∞ being directly along the beam axis [3].

Similar to pseudorapidity is rapidity, y, defined as:

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}, \tag{3.2}$$

which takes into account the particle's energy [3].

The measure ΔR used in this thesis is the separation of two points in $\eta - \phi$ space:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{3.3}$$

A common kinematic variable used in particle physics analysis is transverse momentum, ($p_\perp$), which is the momentum of a particle in the plane transverse to the beam axis. Particles with a high $p_\perp$ are referred to as 'hard' and those with a low $p_\perp$ as 'soft'. The momentum of a particle is calculated using $p = \gamma m_0 v$, where m_0 is the particle's rest mass, v is its velocity, and γ is its Lorentz factor defined as:

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}},$$

where c is the speed of light in vacuo: $299\,792\,458\,\mathrm{m\,s^{-1}}$.

3.3 Pseudoproper time

The pseudoproper lifetime, τ, of a particle p is defined as:

$$\tau = \frac{L_{xy} M_p}{p_{\perp,\mathrm{p}} c}, \tag{3.4}$$

where L_{xy} is the separation in the transverse plane between the primary interaction vertex and the decay vertex of the particle, i.e. its total transverse decay-length [20].

It is a powerful tool for determining where a J/ψ was produced: as their mean lifetime is very short ($7.2 \times 10^{-9}\,\mathrm{ps}$), if $\tau_{J/\psi}$ is close to 0, then it is likely to be prompt (produced in the primary interaction), however, if $\tau_{J/\psi}$ is higher then it is likely to be non-prompt (produced by the decay of an intermediate particle).

Negative values for τ are expected and accepted. They are caused by the finite resolution of the detector, which can sometimes lead to the J/ψ momentum vector being constructed in an opposite direction to the vector between the two vertices [21].

3.4 Luminosity

For particle colliders, instantaneous luminosity (L) is the number of interactions per unit time and can be calculated using:

$$L = \text{const.} \times \text{intensity} \times \text{focusing} = \frac{f\gamma}{4\pi\epsilon_n} N^2 k_b \frac{F(\beta^*)}{\beta^*}, \tag{3.5}$$

where N is the number of particles per bunch, k_b is the number of bunches, F and β^* account for the fact that the beams do not collide head on but at an angle, ϵ_n is the beam emittance[4], f is the revolution frequency of the particles, and γ is the Lorentz gamma factor [22, 23]. In CGS units $[L] = \mathrm{cm^{-2}s^{-1}}$, however L can also be expressed in terms of $\mathrm{b^{-1}s^{-1}}$.

[4] A measure of the spread of particles in position-momentum space.

The amount of data collected by a detector can be expressed in terms of integrated luminosity by integrating over the runtime, T, of the experiment:

$$L_{int} = \int_0^T L \, \mathrm{d}t.$$

This is normally expressed in inverse-barns (b^{-1}). Due to the inverse nature of the unit, measurements with lower SI prefixes (e.g. inverse femto-barn, fb^{-1}) indicate a larger integrated luminosity than measurements with higher prefixes (e.g. inverse micro-barn, $\mu\mathrm{b}^{-1}$).

3.5 The Large Hadron Collider

The LHC is a ring accelerator based at CERN which began operation in 2008. In normal operation it collides proton beams, but it also has the capacity to collide lead ions. Its design specifications are to collide bunches of 7 TeV protons at a centre-of-mass energy ($\sqrt{s}$) of 14 TeV, with around 1.15×10^{11} protons per bunch. The peak instantaneous luminosity will be $1.0 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ [24].

The above specifications are the LHC's design specifications but these are being reached in two stages, and the specifications for the data taken during 2012, which this thesis uses, are: 4 TeV proton beams at $\sqrt{s} = 8$ TeV, with around (1.6 to 1.7) $\times 10^{11}$ protons/bunch and a peak instantaneous luminosity of $7.7 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ [22].

3.5.1 Pile-up

As mentioned above, the LHC collides bunches of protons rather than single protons; this is to increase the chance of a collision occurring when the beams are crossed, but it also means that there can be multiple collisions occurring in the same beam crossing involving separated pairs of protons.

Whilst it doesn't matter which particular pair produces the interaction of interest, it is important to make sure that the final states being used in analysis all came from the same initial interaction. The presence of final states in the detector which came from other proton-proton interactions is referred to as pile-up.

The effect of pile-up can be partially removed by examining the vertex separation of particle tracks; if two particles have a large separation then it is likely that they came from different hadronic collisions. For jets (described in Section 3.8), techniques involving the jet vertex fraction (probability that a jet originated from a particular vertex) or the jet areas may be used [25]. Further removal of pile-up requires accounting for detector effects, which will often introduce large systematic uncertainties on results.

3.6 The ATLAS detector

The ATLAS (A Toroidal Lhc ApparatuS) detector [3] is a large multipurpose detector used in conjunction with the LHC. Detection is performed in three main areas: the inner detector, the calorimeters, and the muon spectrometer. A three-stage trigger system is used to accept or reject events. ATLAS also has a forward detector which is used to monitor beam luminosity and collect data for heavy-ion collisions. Figure 2 shows the overall layout of the detector.

3.6.1 Inner detector

The inner detector (ID), as shown in Figure 3, consists of three main sections: the semiconductor tracker (SCT), the transition-radiation tracker (TRT), and the pixel detectors. The central

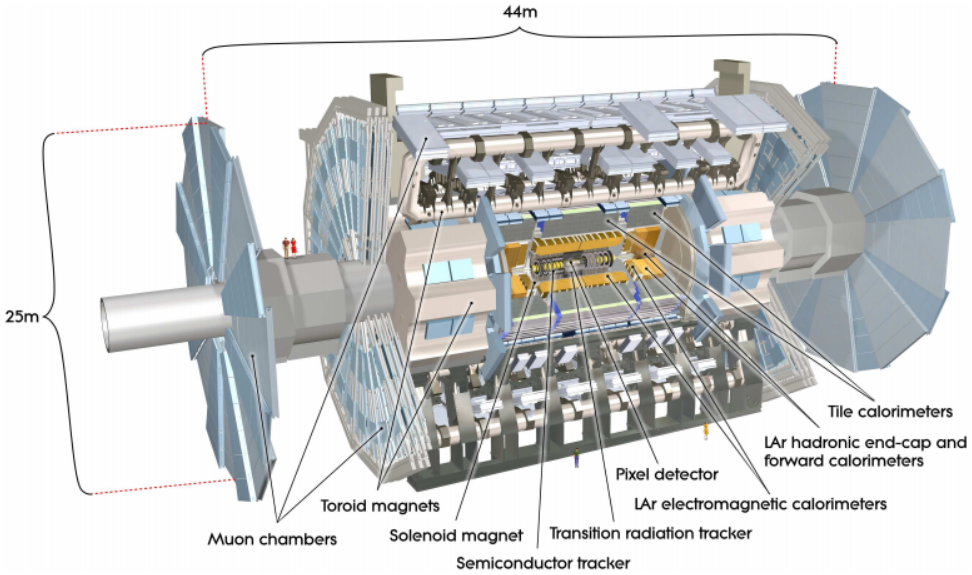


Figure 2: Layout of ATLAS detector. Image from Ref. [3]

solenoid creates a 2 T magnetic field across the ID [3] which causes the paths of electrically-charged particles to be curved. The direction of curvature allows the charge of the particle to be known, and the amount of curvature allows for their charge-momentum ratio to be calculated.

Precision trackers The SCT and pixel detectors are classed as precision trackers and together provide coverage of $|\eta| < 2.5$. They both consist of barrel and end-cap sections with the barrel section arranged in concentric cylinders about the beam, and the end-cap sections arranged in discs in the transverse plane. The pixel detectors offer the highest granularity with intrinsic accuracies of: 10 µm $(R-\phi)$, 115 µm (z) in the barrel region and 10 µm $(R-\phi)$, 115 µm (R) in the end-cap region, with three pixel layer being crossed by a typical track [3]. The intrinsic accuracies of the SCT are: 17 µm $(R-\phi)$, 580 µm (z) in the barrel region and 17 µm $(R-\phi)$, 580 µm (R) in the end-cap region with four space points being crossed by a typical track [3].

Transition-radiation tracker Whilst only measuring $(R-\phi)$ to an intrinsic accuracy of 130 µm the TRT provides the highest amount of track hits at an average of 36 per track [3]. It consists of straws filled with a xenon-based gas mixture [3] arranged parallel to the beam in its barrel region, and in radial disks in its end-cap region with a total coverage of $|\eta| \leq 2$.

3.6.2 Calorimeters

Figure 4 shows the layout of the ATLAS calorimetry section which consists of electromagnetic (EM) and hadronic calorimeters. Calorimeters are designed to measure EM and hadronic showers, however in order to measure any missing energy (which can signal neutrinos or exotic physics), it is required that the showers lose the majority of their energy in their respective calorimeter. To this end, absorber plates are used to reduce the amount of particles which escape from the calorimeter.

Electromagnetic calorimeter The electromagnetic calorimeter (ECal) is designed to provide high-precision measurements of electrons and photons over a large η range using liquid argon interspaced with lead absorption plates. Its barrel section covers the range $|\eta| < 1.475$ and its

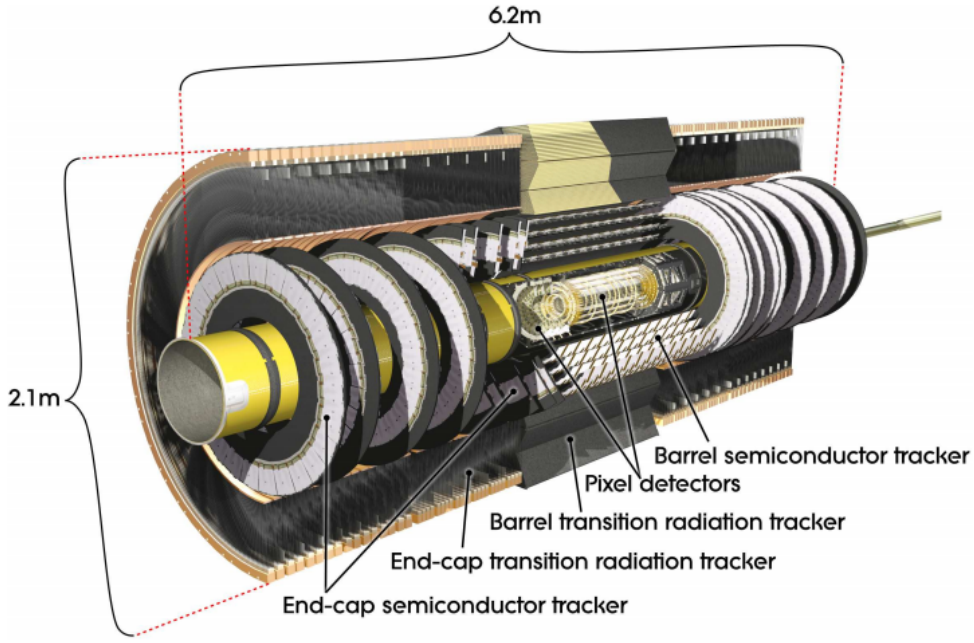


Figure 3: Layout of ATLAS inner detector. Image from Ref. [3]

end-cap region covers $1.375 < \eta < 3.2$. The thickness of the absorber plates varies according to η as does the granularity of the detector with $|\eta| < 2.5$ having the highest granularity. A presampling detector is used in the region $|\eta| < 1.8$ to correct for energy losses upstream by electron and photons [3].

Hadronic calorimeter The hadronic calorimeter (HCal) is made up of the tile calorimeter, the hadronic end-cap calorimeter (HEC), and the forward calorimeter.

The tile calorimeter covers the eta range $|\eta| < 1.7$, and consists of steel absorber plates and scintillating tiles which provide a readout to photomultiplier tubes [3].

The HEC sits behind the end-cap ECal and uses the same cryostats for its own liquid argon active material which, combined with copper absorption plates, it uses to cover $1.5 < |\eta| < 3.2$, slightly overlapping with both the tile calorimeter and the FCal [3].

The FCal covers $3.1 < |\eta| < 4.9$ and is a dual-purpose calorimeter, using copper plates in its innermost part to measure EM showers, and tungsten plates in the outermost layers to measure hadronic showers. It again uses liquid argon as an active medium. [3]

3.6.3 Muon spectrometer

The muon spectrometer (MS), shown in Figure 5, is the outermost area of detection the ATLAS detector offers. Using three air-core toroid magnets [3] it bends the paths of muons escaping from the HCal through four different types of muon chambers. The MS is divided up into three sections: the barrel region ($0 < |\eta| < 1.4$), the transition region ($1.4 < |\eta| < 1.6$), and the end-cap region ($1.6 < |\eta| < 2.7$) [3].

Muon chambers Monitored drift tubes are used over most of the MS ($|\eta| < 2.7$), with cathode strip chambers being used at larger pseudorapidities ($2.0 < |\eta| < 2.7$), [3].

For $|\eta| < 1.05$, resistive plate chambers are used and for $1.05 < |\eta| < 2.7$ thin gap chambers are used [3]. These two chamber types are used in the ATLAS trigger system, however triggering only takes place in the region: $|\eta| < 2.4$ [3].

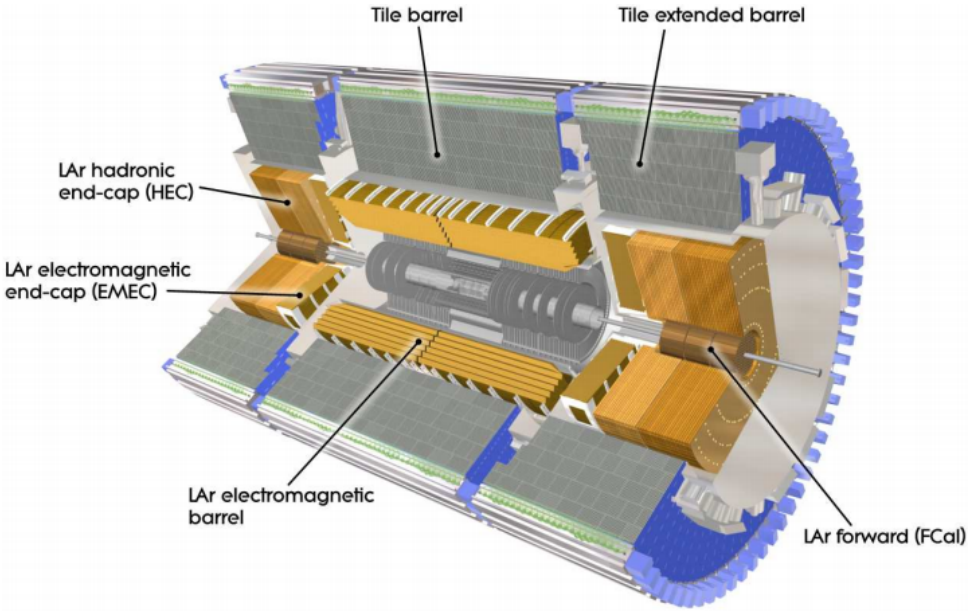


Figure 4: Layout of ATLAS calorimeters. Image from Ref. [3]

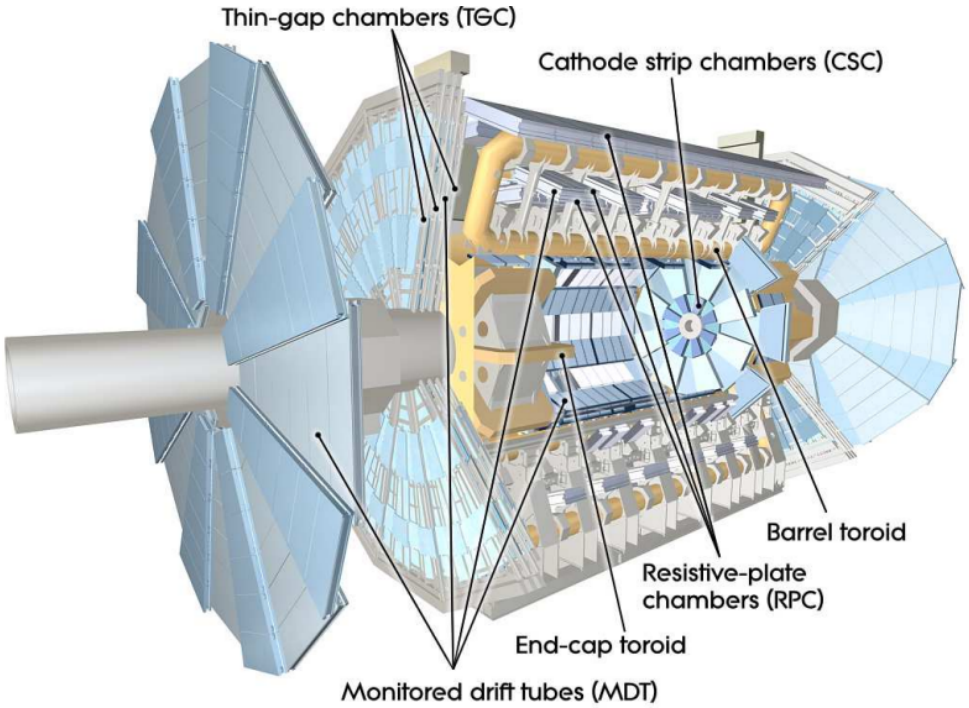


Figure 5: Layout of ATLAS muon spectrometer. Image from Ref. [3]

3.6.4 Trigger system

As mentioned in Section 3.5, the LHC, running at its design specifications, will reach a bunch-crossing rate of 40 MHz and recording all of the data for all of these events each second would be non-trivial. However a lot of these events will be uninteresting, either because there is nothing special going on or because the event does contain an interesting process, but it is obscured by contamination from other processes. Instead, a trigger system is used to decide at runtime which events to record, and afterwards which to pass into storage for further analysis.

For data taking in 2012 the ATLAS detector used a three-part trigger system consisting of the level 1 trigger (L1), the level 2 trigger (L2), and the event filter (EF). The aim of this system is to reduce the event rate to a more manageable 200 Hz [26].

Level 1 trigger The L1 trigger is a hardware-based trigger which makes use of fast, custom electronics to lower the event rate to 75 kHz nominal, 50 kHz 2012 [27]. It does this by identifying regions of interest (RoIs) in the event from course-granularity data from the calorimeters, muon chambers, and the forward detectors [27].

High-Level trigger The L2 trigger and the EF are often grouped together as the high-level trigger (HLT). The HLT is a software-based trigger system run in a processor farm consisting of, in 2010, 800 nodes for L2 and EF usage, and 300 nodes purely for EF usage, each node consisting of an eight-core processor running at a modal speed of 2.4 GHz [26].

The L2 trigger starts from the RoIs identified by the L1 trigger and examines these at full granularity, along with tracking information from the ID, through various algorithms to see if the event really does contain features of interest. If the L2 trigger decides to keep the event, then the rest of the event data is requested and stored to be passed to the EF.

The EF has access to the full event record and enough CPU time to run more-advanced algorithms to identify signal in the event.

3.7 Monte Carlo simulation

3.7.1 General principle

Monte Carlo (MC) generators use stochastic methods to simulate particle collisions and interactions according to particle theory. They are an important tool in particle physics with a variety of uses such as: calculating the quantity of background[5] in a measurement, tuning analysis algorithms such as top taggers, and in designing particle detectors.

The self-interacting nature of QCD means that the calculation of QCD processes is more computationally intensive than QED processes, however QCD can be approximated using perturbation theory (pQCD). This can only be used to describe partons with high momenta, and so once momentum drops below a certain level, empirical models have to be employed. Colour confinement, described in Section 2.2.3, requires that the partons are bound into hadrons, some of which may be unstable.

The hadronic cross section (σ) can be thought of as the probability that an interaction will occur, and consists of the sum of the cross sections of all the possible scattering subprocesses. The cross section of the scattering subprocess $ab \to n$, in which partons a and b scatter to produce final state n, can be calculated by Monte Carlo generators using factorisation formulas

[5] Results from processes which are not considered signal.

of the form:

$$\sigma_{ab\to n} = \int_0^1 \mathrm{d}x_a \int_0^1 \mathrm{d}x_b \int \mathrm{d}\Phi_n \, f_a^{h_1}(x_a, \mu_F) \, f_b^{h_2}(x_b, \mu_F) \times \frac{1}{2\hat{s}} \left|\mathcal{M}_{ab\to n}\right|^2(\Phi_n; \mu_F, \mu_R) \,, \tag{3.6}$$

where: x_j is the light-cone[6] momentum fraction of parton j; μ_F and μ_R are the factorisation and renormalisation scales respectively; Φ_n is the final-state phase-space; $\hat{s} = x_a x_b s$ is the parton flux (s is the squared centre-of-mass energy of the incoming hadrons); $\left|\mathcal{M}_{ab\to n}\right|^2(\Phi_n; \mu_F, \mu_R)$ is the square of the matrix element (effectively the sum over all possible Feynman diagrams for the process $ab \to n$); $f_j^{h_i}(x_j, \mu_F)$ is the parton density function (PDF) of hadron i [28].

The Monte Carlo simulation of an event is therefore split into four main sections: PDF sampling, the matrix element, the parton shower, and hadronisation.

3.7.2 Parton density function

Hadrons are defined by the two or three quarks which carry the majority of their momentum (their 'valence quarks'), however other partons such as sea quarks (virtual quark pairs) and gluons also carry some fraction of the total momentum of the hadron [18]. The fractions of momentum carried by the different partons in a hadron are summarised as parton density functions (PDFs), which give the probability density of finding a given parton with a fraction x of the total momentum of the hadron at an energy scale Q [29]. Since pQCD can only describe partons and the LHC collides protons, PDFs are used to provide the parton input to the Monte Carlo simulation.

3.7.3 The matrix element

The matrix element (ME) is the analytic calculation in pQCD of a specified process (the hard process), such as: gluon gluon $\to$ quark anti-quark ($g\ g \to q\ \bar{q}$) with the gluons being sampled from the PDF. As mentioned in Section 2.2, partons carry colour charge and so can radiate gluons, or form quark pairs, so the process could easily become $g\ g \to q\ \bar{q}\ g$ or $g\ g \to q\ \bar{q}\ q\ \bar{q}$, where the incoming or outgoing partons radiate gluons. The cross sections of these processes would be calculated using pQCD and Monte Carlo integration over the phase space.

The ME calculations may be performed at leading order (LO) using tree-level[7] MEs, or at higher orders, such as next-to-leading order (NLO), by including the contribution of loop corrections[8]. One could expect to continue to NNLO and so forth, increasing the accuracy of the simulation by including greater numbers of loop corrections, however the associated work in calculating these corrections causes the simulation time to increase dramatically.

3.7.4 The parton shower

Instead of explicitly calculating the ME to higher and higher orders and suffering the associated increase in runtime, modelling methods are used to approximate the calculation.

In a process labelled "the parton shower" (PS), the hard process in the ME is dressed with emitted partons and probabilistic functions are used to determine the evolution of these

[6] Coordinate system which transforms easily along the collision axis

[7] ME only includes contributions from the set of simplest Feynman diagrams for a process

[8] Process in which a particle either emits and subsequently reabsorbs another particle, or splits into a particle-antiparticle pair which subsequently annihilate to reform the original particle

Process	Function
$q \to q\ g$	$P_{qq}(z) = C_F \frac{1+z^2}{1-z}$
$q \to g\ q$	$P_{gq}(z) = C_F \frac{1+(1-z)^2}{z}$
$g \to g\ g$	$P_{gg}(z) = C_A \frac{z^4+1+(1-z)^4}{z(1-z)}$
$g \to q\ \bar{q}$	$P_{qg}(z) = T_R \left(z^2 + (1-z)^2\right)$

Table 5: Summary of spin-averaged light-parton splitting functions. $C_F = \frac{N_c^2-1}{2N_c}$, $C_A = N_c$, and T_R is defined to be 1/2. Here, z is the energy fraction of the emitted parton with respect to the initial parton: E_j/E_i [28].

extra partons through further splittings to light partons. For each parton, a random number is generated in the range $[0, 1]$ and Equation 3.7,

$$\Delta_i\left(X^2, x^2\right) = \exp\left(-\frac{\alpha_s}{2\pi}\int_{x^2}^{X^2}\frac{\mathrm{d}k^2}{k^2}\int_{X_0^2/k^2}^{1-X_0^2/k^2}\mathrm{d}z\, P_{ji}(z)\right), \tag{3.7}$$

is set equal to this and solved for x^2. If x^2 is greater than X_0^2 (a cutoff value which determines the point at which branchings become resolvable as two distinct partons instead of one) then a branching is generated at a scale of x^2 and evolution continues on the branching products. If, however, $x^2 \leq X_0^2$ then evolution of that parton ceases [28].

$\Delta_i(X^2, x^2)$ in Equation 3.7 is the probability of no branching between x^2 and X^2, where X^2 is the maximum value of variable x^2. The Sudakov form factor, $\Delta_i(X^2, X_0^2)$, where X_0^2 is the lowest value of x^2, is the probability of no branching at all.

$P_{ji}(z)$, where $i, j \in \{g, q, \bar{q}\}$, are the splitting functions defined in Table 5, and z is the energy fraction of parton j with respect to parton i. x in $\Delta_i(X^2, x^2)$ is the evolution variable for the parton shower, such as transverse momentum or virtuality. In the collinear limit (where θ, the angle between branching products, equals zero) all possible choices of z and x give the same result; it is only once θ increases that the differences in variable choice become apparent [28]. There is no *a priori* prescription for which variables to use and different PS models vary in their variable choice and their X_0^2 cutoff value.

3.7.5 Matrix element and parton shower combination

The ME and PS methods are both viable ways of simulating the evolution of particles, however each has its own advantages and disadvantages: The ME method is well suited for simulating hard, separated partons, but becomes difficult to calculate in the cases of high multiplicity or soft, collinear partons; The PS method, instead, describes soft, collinear partons well, and does not become overly complex as the number of partons is increased [28].

Contemporary Monte Carlo generators use both methods by combining their simulations through matching or merging procedures. When generating LO MEs with fixed partonic final-state multiplicities, such as $2 \to 2$ or $2 \to 3$ MEs, in Pythia 8 and Sherpa, this matching takes the form of the addition of a PS which takes the kinematics of the ME partons as a starting condition.

Sherpa, however, is able to generate MEs with a variable number of final-state partons, e.g. $2 \to 2 + \{2\}$ where $\{2\}$ indicates that upto two extra partons may generated by the ME. Including more of the partonic evolution within the ME improves the accuracy of the simulation as it allows the hard, wide-angle splittings, which the PS simulates poorly, to be explicitly calculated in pQCD. Attaching a PS to these kinds of MEs, however, requires a

merging technique to account for the possibility of over- or under-counting regions of phase space and to resolve inherent differences in the nature of each methods' simulation; the fact that the ME produces the probability of at least n final states, whereas the PS produces the probability of exactly n final states [28]. In SHERPA this is performed by the algorithm detailed in Ref. [30], in which the emission phase-space is divided into two domains, one for the ME and the other for the PS, with the cut defined by the "jet criterion". The PS is then truncated such that it can only emit within its assigned phase-space domain.

When generating NLO MEs, a more advanced form of ME+PS matching is required, two such schemes are MC@NLO [31] and POWHEG [32]. Currently, the most advanced forms of simulation involve the merging of NLO MEs with varying final-state parton multiplicity, and their subsequent matching to a PS in approaches such as MEPS@NLO [33], which is implemented in SHERPA.

3.7.6 Hadronisation

Once the PS finishes, the particles are passed to the hadronisation stage of simulation. This is a non-perturbative process which is modelled according to observations of QCD [34]. There are two models which are commonly adopted: the string model, based on observations in lattice QCD[9] where the gluon field between partons collapses into a thin tube; and the cluster model, based on the observation of pre-confinement where groups of partons can be formed into clusters with no net colour-charge [34]. The hadronisation stage serves to bind the partons into colourless particles (hadrons), thereby reconciling the simulation with the observed lack of free colour-charge.

Particles exiting the hadronisation stage and their subsequent decay products are the final states one can observe in a detector.

3.7.7 Light partons

Light partons are partons which are approximated in Monte Carlo generators to be massless in the ME. In PYTHIA 8 these are normally gluons and u, d, and s quarks. In SHERPA the c and b quarks are normally included in the massless approximation, however in order to allow for accurate comparison with PYTHIA 8 these have been explicitly set to be treated as massive.

The term heavy flavour (HF), is used to refer to c, b, and t quarks, which are treated as massive in the ME by the generators.

Once the generators move into the PS phase, all quarks are treated as massive.

3.7.8 Multi-parton interactions

The composite nature of the hadrons collided by accelerators, such as the LHC at CERN, mean that in a collision the signal process can be accompanied by softer interactions resulting from the underlying event (UE). The primary production mode of the UE is multi-parton interaction (MPI), in which other partons in the colliding hadrons interact with one another and produce final states which can become mixed in with the signal process's final states. It doesn't matter if the signal one is looking for comes from one of these softer interactions, provided the signal only comes from one interaction and is not a mixture of MPI.

The effect of MPI is difficult to remove as the interactions take place at an experimentally indistinguishable distance from the signal process; however, by simulating MPI in Monte Carlo generation, one can see how the addition of MPI affects ones measurements, and including MPI in simulation allows for more accurate comparison with collider data. This is done in PYTHIA 8

[9] A non-perturbative approach to QCD calculations based on discretising space-time.

and SHERPA by including a number of $2 \to 2$ QCD MEs, with options to restrict the possible outgoing flavours.

3.8 Jet clustering

Jets are an important tool in the study of hadronic collisions such as those which take place at the LHC. They are a collection of final state particles from a collision which have been grouped together to be treated as a single object with its own kinematic and geometric variables: $p_\perp$, $E_\perp$, *et cetera*. There are numerous jet algorithms which dictate the way in which the final states are clustered into jets, but they can be divided into two categories: cone and sequential recombination.

Cone algorithms are possibly the most intuitive means of defining a jet: a circle is drawn in $(\eta - \phi)$ space and any particles within the circle are clustered into a jet. Cone algorithms vary in how they choose where to centre the circles and how they deal with overlapping cones, but in general they are infrared and collinear (IRC) unsafe[10]. This means that the way in which they cluster particles can be greatly changed by the addition of soft or collinear particles, which makes them undesirable for theoretical application where calculations rely on perturbation theory. Their naïve clustering method does, however, give them well defined boundaries meaning that their shape is resilient to the effects of soft radiation such as that resulting from the underlying event. This property of soft-resilience made them an attractive choice for experimental application.

Sequential recombination algorithms can be generalised using Equations 3.8 and 3.9,

$$d_{ij} = \min(k_{t,i}^{2p}, k_{t,j}^{2p}) \frac{\Delta R_{ij}^2}{R^2}, \tag{3.8}$$

$$d_i = k_{t,i}^{2p}, \tag{3.9}$$

in which $k_{t,i}$ and $k_{t,j}$ are the transverse momenta of two objects i and j, ΔR_{ij} is the separation of the two objects in $(\eta - \phi)$ space, and R is the radius parameter of the algorithm [35]. For each iteration, d_{ij} and d_i are calculated for each object and each pair of objects. The minimum value of d_{ij} and d_i is then found and if it is a member of d_{ij} then that pair of objects are combined into a single object, if instead it is a member of d_i then that object is defined as a jet and is removed from consideration when calculating subsequent d_{ij} and d_i values. This continues until all particles have been clustered or some stop condition is met.

The value of p in Equations 3.8 and 3.9 determines the amount of weight the algorithm places on the transverse momentum of the objects when clustering: For $p = 0$ objects are clustered in order of closest pairs (the Cambridge/Aachen algorithm); for $p = 1$ objects are clustered in order of closest, softest pairs (the k_t algorithm).

Sequential recombination has the advantage that it is inherently IRC-safe; since all possible object pairs are considered at any stage, collinear particles will combined together early, and soft radiation will be clustered with other particles rather than seeding their own jet. For $p \geq 0$ they do, however, feature soft-adaptable boundaries meaning that soft radiation can cause their shapes to become irregular, which made that less useful in experiments.

The anti-k_t algorithm, defined by setting $p = -1$, was introduced in 2008 by Cacciari, Salam, and Soyez [35]. This algorithm clusters in order of closest, hardest pairs giving it a soft-resilient boundary whilst retaining IRC-safety. Since it desirable for both theoretical and experimental application, it has been adopted by both ATLAS and CMS as their main standard jet algorithm.

[10] Some clustering algorithms such as SISCone are designed to be IRC-safe cone algorithms

4 $b\bar{b}$ pair production in Monte Carlo

4.1 Methods of heavy-flavour production

Production of heavy-flavour quarks can take place in the the ME or the PS of an event simulation. Direct production in the ME would require a hard process of the form $\mathscr{P}\,\mathscr{P} \to b\,\bar{b}\,X$, where $\mathscr{P}$ is a parton (light or heavy - see Section 3.7.7) and X are any other products of the hard process, such as extra light-partons (ρ). A fully-light ME can still lead to heavy-flavour production in its subsequent PS: if a radiated gluon has enough virtuality then it can split to a $b\,\bar{b}$ pair.

It can be expected that ME production will result in a large angle between the quarks as momentum conservation means that with no initial state radiation (ISR), the outgoing partons from a $2 \to 2$ ME must be back-to-back in the transverse plane of the laboratory rest-frame. If, however, the incoming partons are able to radiate, or the ME is able to emit extra light-partons, then the outgoing partons do not necessarily have to be back-to-back. As mentioned in Section 3.7, the ME production method is an explicit calculation of the process, so it can be expected that the two generators will not show a large difference in their results due to there only being one correct way to calculate the ME.

For massless partons, the gluon-splitting function P_{qg}, in Table 5, shows maxima at $z = 0$ and 1, so it can be expected that majority of massless quarks would be produced in the low-angle region. b quarks, however, are treated are massive and Pythia 8 and Sherpa employ different methods to model massive splittings: Pythia 8 uses ME information to apply process-dependant corrections to massive splittings [36], where as Sherpa uses generalised forms of the splitting functions to account for parton mass-effects [37].

In current theoretical particle physics there is no *a priori* way of choosing the correct evolution variable for the PS approximation as there are several variables and which would be a reasonable choice; all choices give the same result if calculated using an infinite series, but differences arise when only using a finite number of perturbation terms. In some cases the choice of variable can be supported by showing that it provides a good approximation when using explicitly-calculated higher-order perturbation terms, or if other choices produce nonsensical results, such as negative cross-sections. It can therefore be expected that the varying choices of scale and the different modelling methods of massive partons used by the generators should lead to noticeable deviations in their PS production.

Including extra light-partons as outgoing particles from the ME reduces the amount of approximation in the event generation, and so reduces the effect of PS scale choice. To this end, samples of the form LO+$\{N_e\}$ will also be generated, where LO is the leading order production process (ME or PS) and $\{N_e\}$ indicates the maximum number of extra light-partons exiting the explicitly calculated ME. Since a PS provides probabilities for any number of splittings, both the LO+$\{0\}$ and LO+$\{N_e\}$ samples require a PS to correct them to all orders, however if the results of the LO+$\{N_e\}$ samples are the same as those for LO+$\{0\}$, then the PS for the LO+$\{0\}$ sample must be providing a good approximation to the splittings which are calculated explicitly in the LO+$\{N_e\}$ samples.

4.2 Monte Carlo samples

Monte Carlo production was divided into two main categories: ME and PS. The titles refer to where the b-quark production primarily takes place. Two generators were used to produce LO samples: Sherpa and Pythia 8. Sherpa was also used to create samples with extra possible outgoing light partons in the ME. The number of possible extra partons is indicated in the summary tables in the braces in the sample name and following the `93` in the process

description, e.g. `93{1}`, where `93` is SHERPA's light-parton container, p. When generating these samples, `CKKW` (the jet criterion for SHERPA's merging algorithm) was set to $\sqrt{20/E_{\text{CM}}}$.

All production, except where noted, was performed using the *MSTW2008lo68cl_nf4* PDF [38], which uses a fixed-flavour scheme to prevent the sampling of b and t quarks for incoming partons in MEs. When using SHERPA, c and b quarks were set to be massive such that they were not included in its `93` light-parton container and the number of EW interactions in the ME was set to zero to ensure the propagators were partons. The minimum $p_\perp$ for outgoing particles ($\hat{p}_{\perp,min}$) for the ME was set to 10 GeV and the colliding beams were set to 4 TeV protons. MPI simulation was performed by the generators' in-built handlers, with AMISIC [39] being selected for SHERPA. In both cases the MPI was set to produce $2 \to 2$ QCD interactions.

4.2.1 Matrix-element production

The ME samples were designed such that b quarks were primarily produced in the ME of the simulation and used the process of $p\ p \to b\bar{b}$, with additional outgoing light partons in the extra-parton samples. Where MPI was simulated, indicated by "MPI" in the sample name, the handler was set to be able to produce HF quarks in the ME of the MPI.

The ME samples are summarised in the Table 6.

4.2.2 Parton-shower production

The PS samples were designed such that b quarks were primarily produced in the PS of the simulation and used the process of $p\ p \to p\ p$, with additional outgoing light partons in the extra-parton samples. The processes switched on in PYTHIA 8 were, from the `HardQCD` processes: `gg2gg` ($g\ g \to g\ g$), `gg2qqbar` ($g\ g \to q\ \bar{q}$, only light quarks exiting), `qg2qg` ($q\ g \to q\ g$), `qq2qq` ($q\ q' \to q\ q'$, $q\ \bar{q}' \to q\ \bar{q}'$, and $\bar{q}\ \bar{q}' \to \bar{q}\ \bar{q}'$, where outgoing and incoming flavours are equal and q and q' may be equal), `qqbar2gg` ($q\ \bar{q} \to g\ g$), `qqbar2qqbarNew` ($q\ \bar{q} \to q'\ \bar{q}'$, incoming HF allowed but only light quarks exiting). This collection of processes is hereafter referred to as `HardQCD:Light`. Note that the `qg2qg` and `qq2qq` processes would usually be able to have outgoing b quarks, however the fixed-flavour PDF used for generation here prevents this. The fixed-flavour PDF also prevents initial state b quarks for the `qqbar2gg` and `qqbar2qqbarNew` processes. A justification for these restrictions follows in Section 4.5. Where MPI was simulated, indicated by "MPI" in the sample name, the handler was set to be unable to produce new quark flavours in its ME, but was able to in the subsequent showering of the MPI ME. This ensured that b quarks were only generated through gluon splitting in parton showering.

The PS samples are summarised in the Table 7.

4.3 Sherpa and Pythia 8 analyses

Analysis of the SHERPA and PYTHIA 8 MC samples was performed using RIVET [40]. b hadrons (denoted "B") were used as proxies for the b quarks in order to relate the results to post-hadronisation particles, which can be measured in a particle detector. This also allowed the same analysis code to be used with both generators; SHERPA and PYTHIA 8 differ in their population of the event record and a separate analysis would be required for each if pre-hadronisation particles were used. The b hadrons were found using the `HeavyHadrons` projection in the range: $-2.5 \leq \eta \leq 2.5$. In order for an event to be accepted, at least two b hadrons had to be found with $p_{\perp,B} \geq 10$ GeV. When more than two b hadrons were found, the hardest two were used. The angular and kinematic properties of the selected b-hadron pairs are then plotted.

Sample Name	Process	σ [nb]
Sherpa ME LO	`93 93` $\to$ `5 -5`	$16\,828 \pm 4$
Sherpa ME LO+MPI	`93 93` $\to$ `5 -5`	$16\,829 \pm 3$
Sherpa ME LO+{1}	`93 93` $\to$ `5 -5 93`{1}	$17\,097 \pm 3$
Sherpa ME LO+{1}+MPI	`93 93` $\to$ `5 -5 93`{1}	$17\,097 \pm 5$
Sherpa ME LO+{2}	`93 93` $\to$ `5 -5 93`{2}	$17\,133 \pm 3$
Sherpa ME LO+{2}+MPI	`93 93` $\to$ `5 -5 93`{2}	$17\,132.4 \pm 4.7$
Pythia 8 ME LO	`HardQCD:hardbbbar`	$16\,371 \pm 2$
Pythia 8 ME LO+MPI	`HardQCD:hardbbbar`	$16\,365 \pm 2$

Table 6: Summary of ME sample production. `93` *and* `5` *indicate light partons and b quarks respectively.* 1×10^7 *events were generated for all samples except for* Sherpa *ME LO+{2}+MPI, where* 9 999 982 *events were generated.*

Sample Name	Process	σ [nb]
Sherpa PS LO	`93 93` $\to$ `93 93`	$5\,209\,900 \pm 600$
Sherpa PS LO+MPI	`93 93` $\to$ `93 93`	$5\,212\,000 \pm 1000$
Sherpa PS LO+{1}	`93 93` $\to$ `93 93 93`{1}	$5\,216\,200 \pm 700$
Sherpa PS LO+{1}+MPI	`93 93` $\to$ `93 93 93`{1}	$5\,216\,000 \pm 1000$
Sherpa PS LO+{2}	`93 93` $\to$ `93 93 93`{2}	$5\,218\,000 \pm 1000$
Sherpa PS LO+{2}+MPI	`93 93` $\to$ `93 93 93`{2}	$5\,218\,000 \pm 1000$
Pythia 8 PS LO	`HardQCD:Light`	$5\,135\,700 \pm 500$
Pythia 8 PS LO+MPI	`HardQCD:Light`	$5\,137\,300 \pm 600$

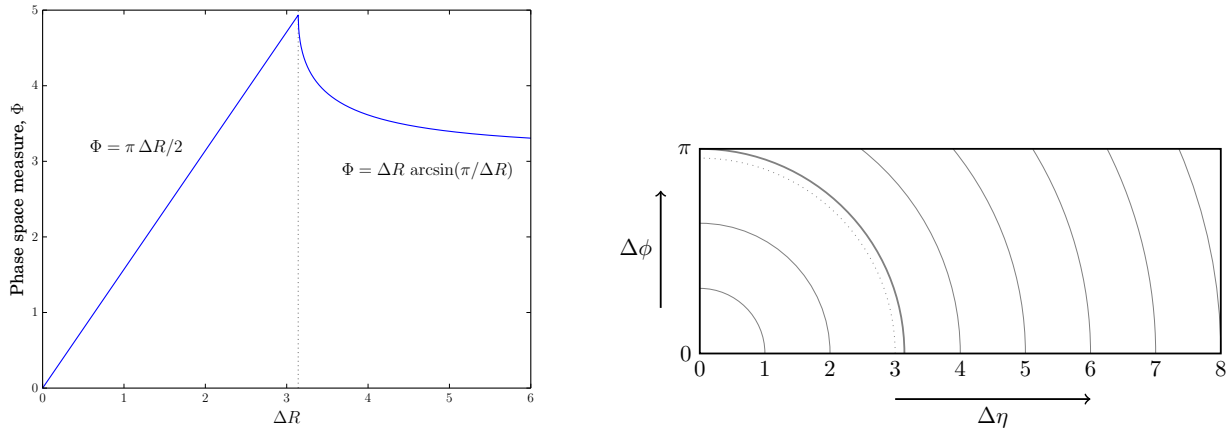
Table 7: Summary of PS sample production. `93` *indicates light partons.* 1×10^7 *events were generated for all samples.*

4.4 Accounting for ΔR phase space

The measure ΔR is calculated using:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}.$$

The phase-space measure, Φ, of a state is the probability of a system being found in that state. Because $\Delta\phi$ is bound between 0 and π but $\Delta\eta$ can be anywhere between 0 and ∞, the phase-space measure for ΔR forms a peak at $\Delta R = \pi$. The shape of $\Phi_{\Delta R}$ is shown in Figure 6. A varying Φ means that even for a uniform distribution of $\Delta\phi$ and $\Delta\eta$, ΔR will be uneven displaying a peak at $\Delta R = \pi$. To avoid confusing shapes in the ΔR distribution caused by underlying physics with those due to the phase-space measure, each entry in the ΔR plots will be weighted by $\Phi^{-1}_{\Delta R}$. The equations for $\Phi_{\Delta R}$ used in this investigation and Figure 6 were derived and created by Andy Buckley.



(a) *(b)*

Figure 6: Figure (a) shows the growth of $\Phi_{\Delta R}$ as a function of ΔR. Figure (b) shows pictorially the evolution of $\Phi_{\Delta R}$: The grey lines indicate the possible combinations of $\Delta\phi$ and $\Delta\eta$ values required to produce a given ΔR. The length of each line is the phase space for that ΔR, $\Phi_{\Delta R}$. The arc-length can be seen to increase linearly with ΔR upto $\Delta R = \pi$, after which the upper bound of π on $\Delta\phi$ begins to reduce the phase space available to produce higher ΔR values. Plot credits: Buckley.

4.5 Justification for fixed-flavour scheme

In an initial investigation, the *CTEQ6L1* PDF [41] was used which meant that the `qg2qg` and `qq2qq` processes in Pythia 8's PS production were able to have outgoing b quarks if there were b quarks in the incoming partons. As this was found to occur in 1.8% of events[11], it was assumed that their presence would have a negligible effect, however it was later found that, due to the very low acceptance of events in PS production, these events were likely to be accepted when they occurred. By placing a cut on events with incoming b-quarks Figure 7 was created. As can be seen, including events with incoming b-quarks increases acceptance by about 60 %, with the additional events being mostly in the high-angle region (particularly noticeable in the $\Delta\phi_{B,B}$ distribution (Figure 7b)).

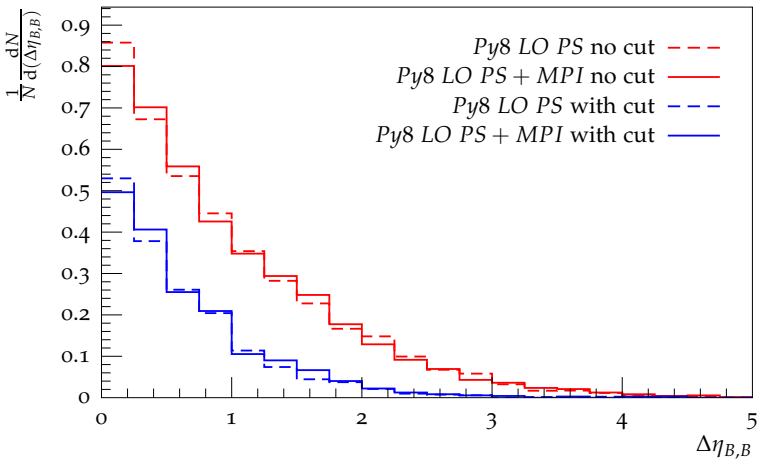
[11] Measurement taken from searching through HepMC output of 10 000 events.

Further investigation of $\Delta\phi_{B,B}$ revealed that for $\Delta\phi_{B,B} \geq \pi/2$ about 70 % of events were generated though flavour excitation[12] of the form: $p\ b \to p\ b$. As shown in Figure 8, Pythia 8 would generate a $b\bar{b}$ quark pair from gluon splitting in backwards evolution from the ME and the hard scattering of one of the b-quarks is likely to be what causes the large angle between the b hadrons.

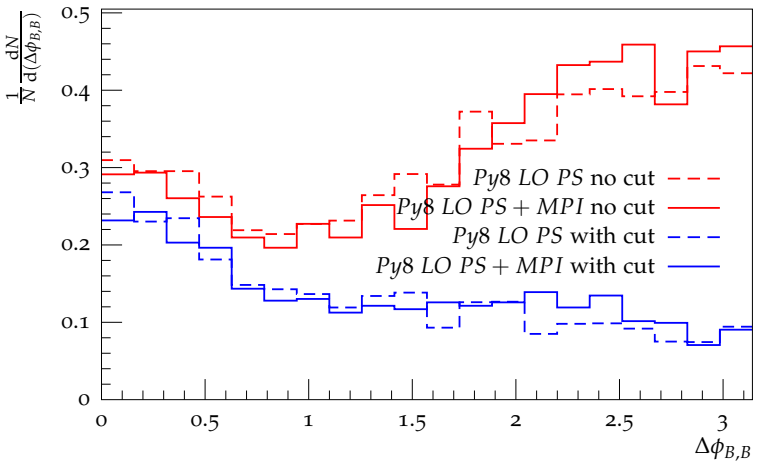
This is gluon splitting, however this section focuses on accurate comparison between MC generators for HF production. The Sherpa run card could have been set to include the same b-quarks processes as Pythia 8 by including the following processes: $b\bar{b} \to p\ p$, $b\ p \to b\ p$, $\bar{b}\ p \to \bar{b}\ p$, $b\bar{b} \to b\bar{b}$, $b\ b \to b\ b$, and $\bar{b}\ \bar{b} \to \bar{b}\ \bar{b}$. However the b quarks in Sherpa must be treated as massive in order to prevent its MPI handler from generating them at ME level, and so the b quarks produced by including these processes would also be massive in the ME, whereas Pythia 8 treats them as massless.

For the sake of accurate generator comparison in this section, it was decided to restrict the presence of b quarks in the ME for the PS production samples, but rather than using an analysis-based cut, production was moved to use a fixed-flavour PDF, where the number of active flavours was less than five (*MSTW2008lo68cl_nf4*). This prevented the sampling of b quarks from the PDF which meant that the production cross sections of the samples would be correct. In order to accurately compare the generator predictions to collider data, the restriction on initial-state b quarks must be removed as these extra processes will contribute to the final prediction. The restriction on initial-state b quarks will be removed in Section 8 where the MC results are compared to collider data.

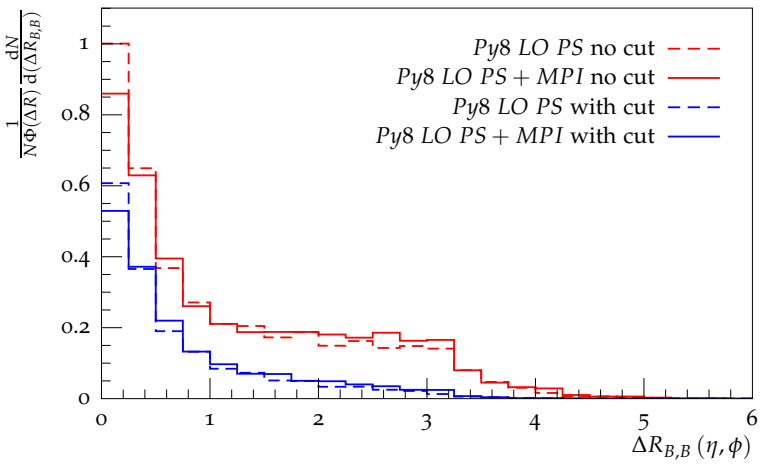
[12]Process in which a HF quark, either sampled from the PDF or created in the initial state shower, is scattered by a light parton in the hard process through virtual gluon exchange in either the t or u channel



(a) $\Delta\eta_{B,B}$



(b) $\Delta\phi_{B,B}$



(c) $\Delta R_{B,B}$*. Plotting accounts for phase space.*

Figure 7: Plots showing the effect of including initial-state b-quarks in ME for PS production. The "no cut" results are normalised to one, and the "with cut" results are normalised to $f_{with\ cut}/f_{no\ cut}$ *to show the decrease in acceptance from their corresponding "with-cut" sample. Samples were created with CTEQ6L1 PDF. The cut prevents acceptance of an event if it has a b quark as an incoming parton to its ME.*

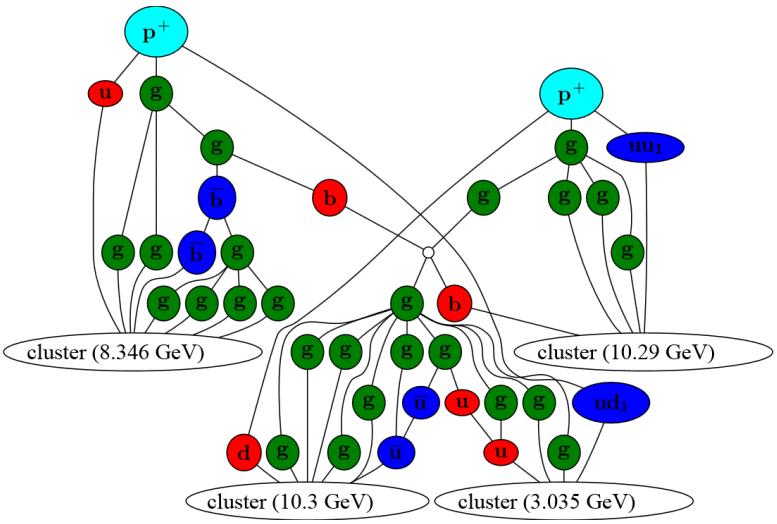


Figure 8: Diagram showing the evolution of one of the Pythia *8 events from the PS production sample where* $\Delta\phi_{B,B} \geq \pi/2$*. The hard process, indicated by the white circle, is* $b\ g \to b\ g$*, where the* b *quark is created in gluon splitting during backwards evolution from the ME. Image credit: author.*

4.6 Results and discussion

4.6.1 LO production

The results of the LO+{0} sample analysis are presented here as plots of ΔR, $\Delta\eta, \Delta\phi, m_{B,B}, p_{\perp B}$ and η_B for b-hadron pairs in ME production using the LO process: $p\ p \to b\,\bar{b}$, and PS production using the LO process: $p\ p \to p\ p$. The non-MPI results are normalised to one, and the results with MPI are normalised to $f_{\text{MPI}}/f_{\text{NMPI}}$, where f_{MPI} and f_{NMPI} are the fraction of events accepted from the sample with MPI and without MPI, respectively; they are scaled to show the increase in acceptance MPI causes. The error bars and yellow error bands show the statistical uncertainty on bin fills.

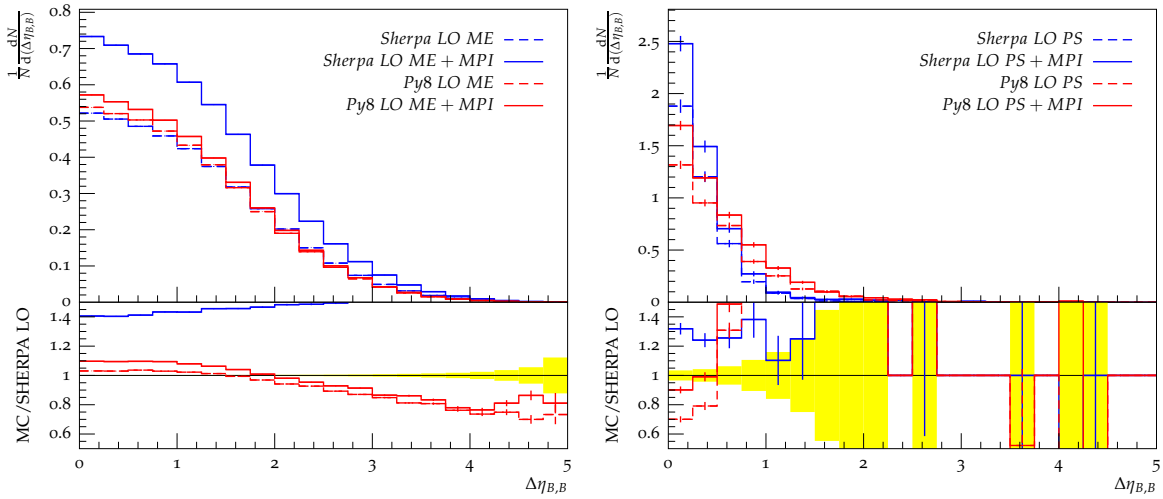
ME production As was expected, the left-hand-side plots in Figures 9, 10, and 11 show that ME production of HF in Sherpa and Pythia 8 produces similar shapes for all plotted variables.

The effect of MPI differs significantly between the generators: for Pythia 8 the results with and without MPI almost overlap, whereas for Sherpa the addition of MPI causes an acceptance increase of $(44.0 \pm 0.2)\,\%$. The process used for ME production of b quarks ensures that there will be at least two b hadrons after hadronisation[13] so there are two ways in which the addition of MPI can increase acceptance:

1. One or both of the hard-process b hadrons are outside the η range required by the analysis, but the MPI produces enough b hadrons which are in the η range resulting in the event being accepted.

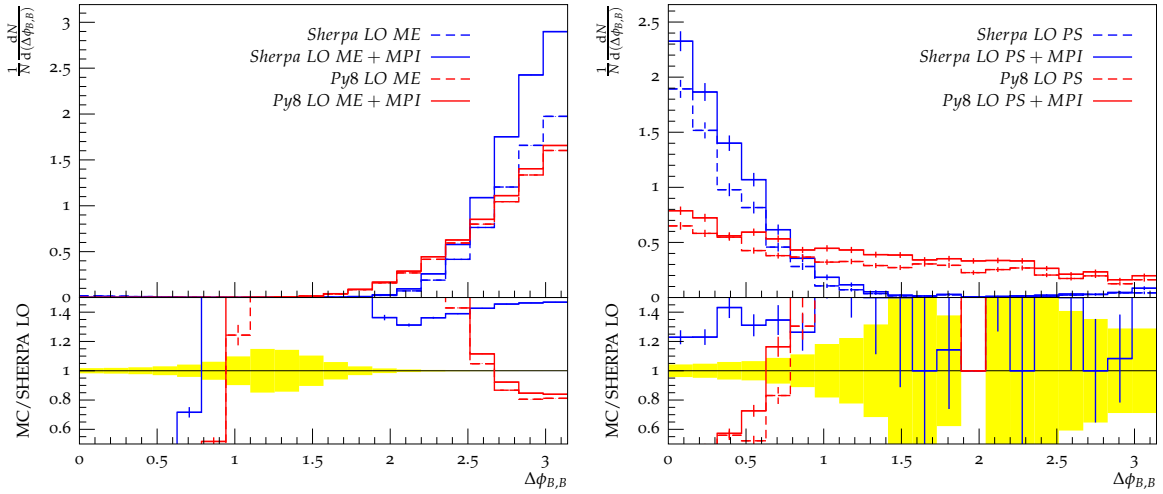
2. One or both of the hard-process b hadrons fail the $p_\perp$ cut required by the analysis, but the MPI produces enough b hadrons pass the $p_\perp$ cut resulting in the event being accepted.

[13]Whilst a bound state of the two b quarks, (bottomonium), could in theory be formed, the b quarks would have to exit the hard process at a very small angle which is extremely unlikely due to momentum conservation, (see Section 4.1)



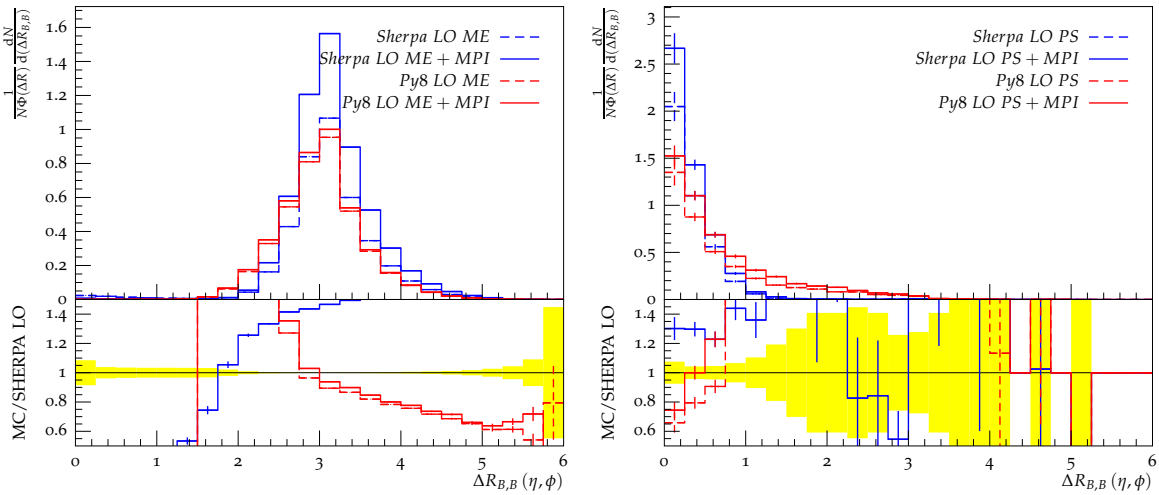
(a) ME Production. *(b) PS Production.*

Figure 9: Comparison of $\Delta\eta_{B,B}$ for ME and PS b-quark pair-production. PS production can be seen to produce closer b-hadron pairs. MPI addition has a much greater effect in Sherpa *than it does in* Pythia *8, particularly in ME production. As expected, both generators produce similar results for ME production (see Section 4.1). Normalisation for the plots is detailed in Section 4.6.1. Error bars and bands show the statistical uncertainty on bin fills.*



(a) ME Production. *(b) PS Production.*

Figure 10: Comparison of $\Delta\phi_{B,B}$ for ME and PS b-quark pair-production. Both generators, again, produce similar results for ME production, but show a very large difference in PS production; Pythia *8's distribution is very flat in $\Delta\phi$. MPI addition is again seen to have a much greater effect in* Sherpa *than in* Pythia *8.*



(a) ME Production. *(b) PS Production.*

Figure 11: Comparison of $\Delta R_{B,B}$ for ME and PS b-quark pair-production taking into account the phase space $\Phi(\Delta R)$. The plots are the combination of the $\Delta\phi$ and $\Delta\eta$ distributions and, as expected, ME production is seen to be centred around back-to-back production. PS production can be seen to produce closer b-hadron pairs.

Sample Name	ME Acceptance Fraction	PS Acceptance Fraction
Sherpa LO	0.1099 ± 0.0001	$0.000\,181 \pm 0.000\,004$
Sherpa LO+MPI	0.1583 ± 0.0001	$0.000\,236 \pm 0.000\,005$
Sherpa LO+{1}	$0.116\,93 \pm 0.000\,06$	$0.000\,197 \pm 0.000\,004$
Sherpa LO+{1}+MPI	0.1652 ± 0.0001	$0.000\,253 \pm 0.000\,006$
Sherpa LO+{2}	$0.117\,70 \pm 0.000\,09$	$0.000\,196 \pm 0.000\,004$
Sherpa LO+{2}+MPI	0.1662 ± 0.0001	$0.000\,246 \pm 0.000\,005$
Pythia 8 LO	0.2687 ± 0.0001	$0.000\,424 \pm 0.000\,005$
Pythia 8 LO+MPI	$0.283\,31 \pm 0.0001$	$0.000\,44 \pm 0.000\,008$

Table 8: Fraction of events accepted by analysis for ME and PS production samples.

The addition of MPI in the Pythia 8 sample causes an acceptance increase of $(5.42 \pm 0.06)\,\%$, which is not as significant as that seen in Sherpa. This lack of increase could be due to two factors which are not necessarily mutually exclusive:

1. The MPI ME tends to produce very low $p_\perp$ partons.
2. The outgoing partons from the MPI ME are normally at high pseudorapidities.

As expected, the Figure 11a shows a peak around $\Delta R = \pi$ meaning that the b hadrons are normally produced back to back, with some deviation from this configuration due to ISR or the PS.

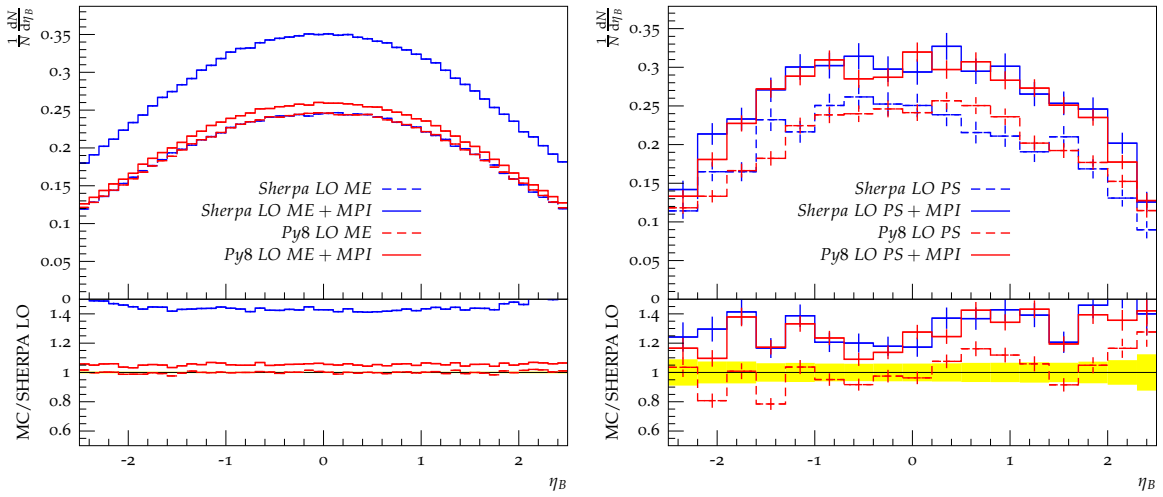
Without MPI, the acceptance fractions of the generators show a large difference, with Pythia 8's fraction being $(144.6 \pm 0.3)\,\%$ greater than Sherpa's. This suggests that Pythia 8 normally produces b hadrons which are at lower $|\eta|$ values, or are at $p_\perp$s higher than those of Sherpa. Figures 13a and 14a show that the generators' η distributions are approximately equal, but that Pythia 8 does produce more b-hadrons with $p_\perp$s just above the cut threshold than Sherpa.

PS production The plots for PS production of b quarks, and in particular Figure 10b, show that the two generators differ greatly. Pythia 8's results show a very slight linear decrease with $\Delta\phi$ and a moderate exponential decay slope in $\Delta\eta$ which peaks towards $\Delta\eta = 0$. Sherpa's results, on the other hand, show a strong peak towards $\Delta\eta = 0$ and $\Delta\phi = 0$ which decays exponentially towards higher angles. As can be seen in Figure 11b, both generators produce b-hadron pairs deviating from the zero-angle configuration, but Pythia 8, due to it slow dependence on $\Delta\phi$, extends farther out to higher ΔR values than Sherpa does.

The two generators also show a large difference in the mass (Figure 12b) of the b-hadron pairs which are accepted with Sherpa's pairs peaking in the range 10 GeV to 20 GeV and Pythia 8's mass distribution being smeared out to higher masses.

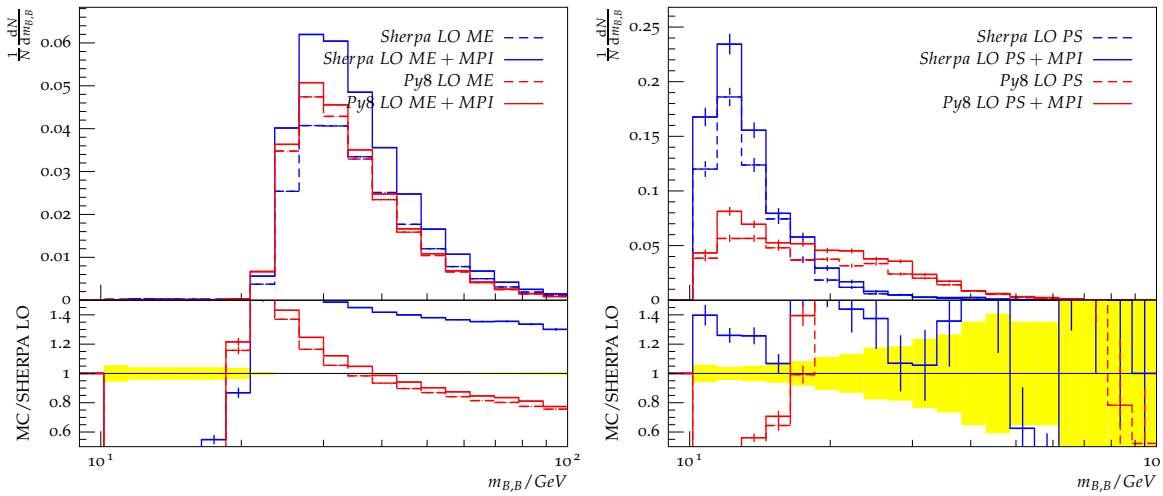
The addition of MPI here causes even less of an acceptance increase for Pythia 8, of $(3 \pm 3)\,\%$ (cf. $(5.42 \pm 0.06)\,\%$ for ME production). The acceptance increase for Sherpa is also lower at $(30 \pm 4)\,\%$ (cf. $(44.0 \pm 0.2)\,\%$ for ME production). The ways in which the addition of MPI can increase are the same as those listed in Section 4.6.1, however not all events are guaranteed to contain a $b\bar{b}$ pair, hence the very low acceptance fractions.

The generators also differ greatly in their acceptance fraction for their non-MPI samples with Pythia 8's acceptance fraction being larger than Sherpa's by $(134 \pm 6)\,\%$, which is a similar figure to the difference seen in ME production ($(144.6 \pm 0.3)\,\%$). Figures 14b and 13b show that the accepted b-hadrons from both generators are produced with an approximately equal η and $p_\perp$ distributions, so for Pythia 8 to have such high acceptance fractions must mean that it's $g \to b\bar{b}$ splitting fraction is higher than Sherpa's.



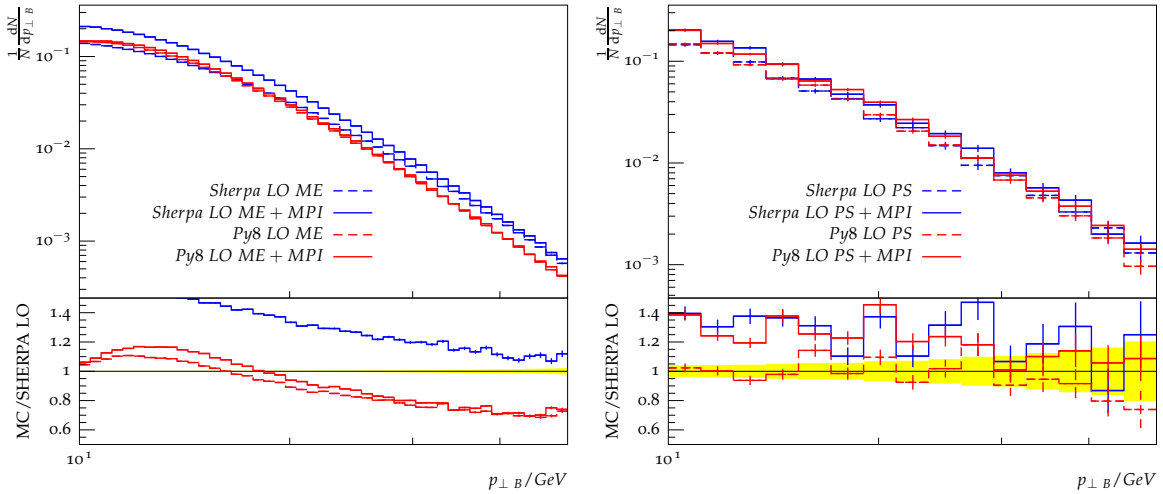
(a) ME Production. *(b) PS Production.*

Figure 14: Comparison of η_B for ME and PS b-quark pair-production. No great deviation can be seen between the two generators' distribution shapes.



(a) ME Production. *(b) PS Production.*

Figure 12: Comparison of $m_{B,B}$ for ME and PS b-quark pair-production. The ME production distribution is seen to peak at a higher mass than the PS production distribution. This is due to the back-to-back configuration of ME production shown in Figure 11, where the momenta of the outgoing products cancel meaning that more of the energy of the system is transferred into mass. NB: The distributions "switch on" at around 10 GeV due to the minimum mass of a two b-hadron system.



(a) ME Production. *(b) PS Production.*

Figure 13: Comparison of $p_{\perp B}$ for ME and PS b-quark pair-production. Sherpa *is seen to produce a slightly higher proportion of harder b-hadrons than* Pythia *8 in ME production, and approximately the same proportion in PS production.*

Acceptance-ratio comparison From Table 9 it can be seen that for both generators show approximately the same production ratios of about two thirds ME to one third PS. The addition of MPI serves to slightly shift the ratio towards ME production for both generators.

Sample Set	Acceptance Ratio [ME : PS]
Sherpa LO	0.662 ± 0.006 : 0.338 ± 0.009
Sherpa LO+MPI	0.684 ± 0.005 : 0.316 ± 0.007
Sherpa LO+{1}	0.660 ± 0.005 : 0.340 ± 0.008
Sherpa LO+{1}+MPI	0.681 ± 0.005 : 0.319 ± 0.007
Sherpa LO+{2}	0.663 ± 0.009 : 0.337 ± 0.009
Sherpa LO+{2}+MPI	0.690 ± 0.005 : 0.310 ± 0.007
Pythia 8 LO	0.669 ± 0.003 : 0.331 ± 0.004
Pythia 8 LO+MPI	0.673 ± 0.006 : 0.327 ± 0.009

Table 9: Acceptance ratios for ME and PS production results. The acceptance ratio is the normalised ratio of a sample's production cross section times its acceptance fraction, i.e. it describes the relative contributions from the ME and PS production methods.

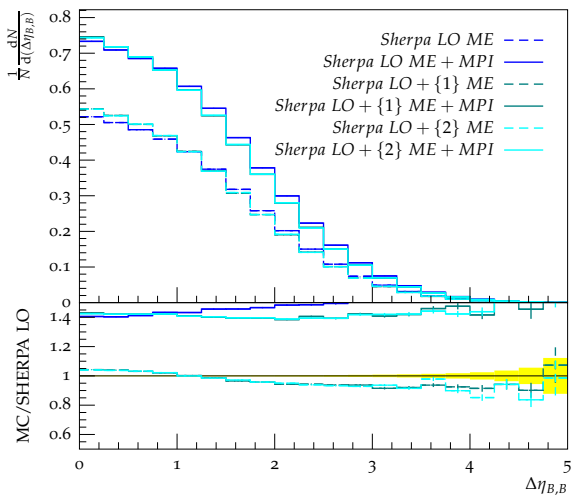
4.6.2 LO production with extra partons

The results of the LO+$\{N_e\}$ sample analysis are presented in Figures 15 to 20 as plots of ΔR, $\Delta\eta, \Delta\phi, m_{B,B}, p_{\perp B}$ and η_B for b-hadron pairs in ME production using the LO process: $p\ p \to b\,\bar{b}\ p\{N_e\}$, and PS production using the LO process: $p\ p \to p\ p\ p\{N_e\}$, where N_e is the maximum number of extra light-partons exiting the ME. The non-MPI results are normalised to one, and the results with MPI are normalised to $f_{\text{MPI}}/f_{\text{NMPI}}$, where f_{MPI} and f_{NMPI} are the fraction of events accepted from the sample with MPI and without MPI, respectively; they are scaled to show the increase in acceptance MPI causes. The error bars and yellow error bands show the statistical uncertainty on bin fills.

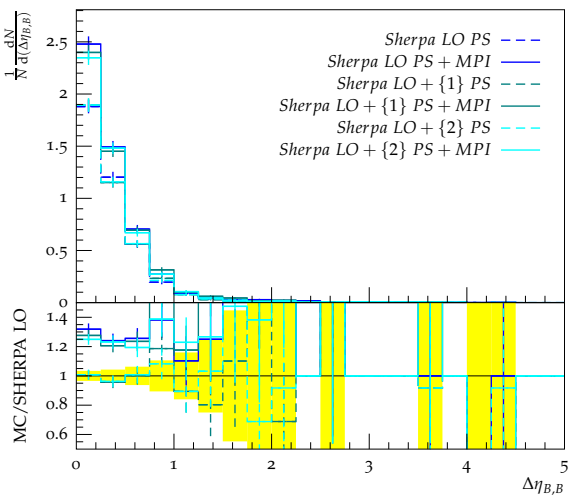
ME production As can be seen in the extra-parton angular-distribution plots, there is a small deviation from LO+{0} to LO+{1}, and only a very slight deviation from this to LO+{2}. This deviation appears to be in favour of lower values of $\Delta\phi$ leading to a bump above LO+{0} in ΔR between 0.5 and 2.2. This is particularly noticeable in Figure 17a. It is possible that production in this region is due to the $b\,\bar{b}$ pair recoiling against a hard light parton. The $m_{B,B}$ distribution in Figure 18a shows no significant deviation, nor does the η distribution in Figure 20a. The $p_\perp$ distribution in Figure 19a, shows agreement between the samples up to around 20 GeV, after which the LO+$\{N_e\}$ samples show increased production of higher-$p_\perp$ b-hadrons.

As can be seen in Table 10, the LO+$\{N_e\}$ samples offer slight deviations, in the range (5 to 7.1) %, for event acceptance.

The lack of significant deviation between LO+{0} and LO+$\{N_e\}$ indicates that Sherpa's PS produces a good approximation to the higher-order ME calculation, however the higher-order MEs are necessary to provide the increased lower-angle production seen in the $\Delta\phi$ and ΔR plots.

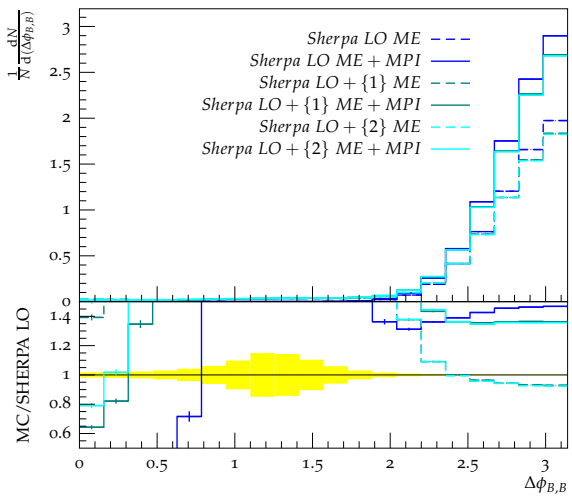


(a) ME Production.

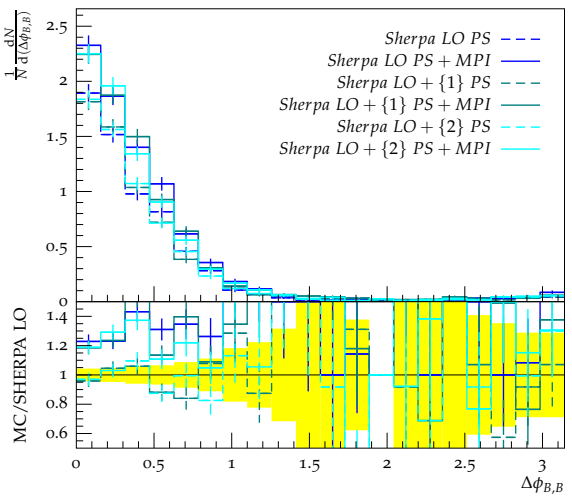


(b) PS Production.

Figure 15: Comparison of $\Delta\eta_{B,B}$ for ME and PS b-quark pair-production.

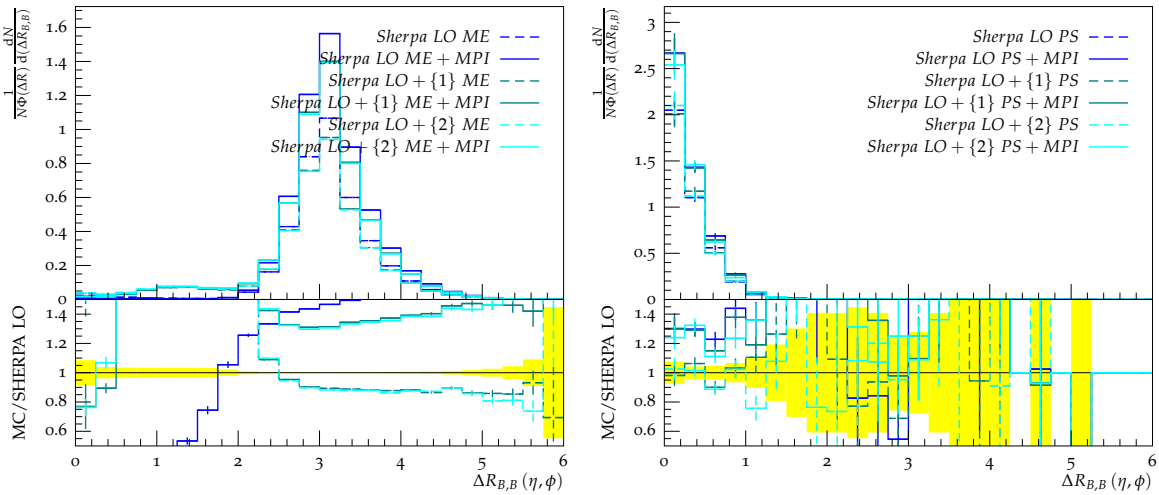


(a) ME Production.



(b) PS Production.

Figure 16: Comparison of $\Delta\phi_{B,B}$ for ME and PS b-quark pair-production.



(a) ME Production. *(b) PS Production.*

Figure 17: Comparison of $\Delta R_{B,B}$ for ME and PS b-quark pair-production taking into account the phase space $\Phi(\Delta R)$.

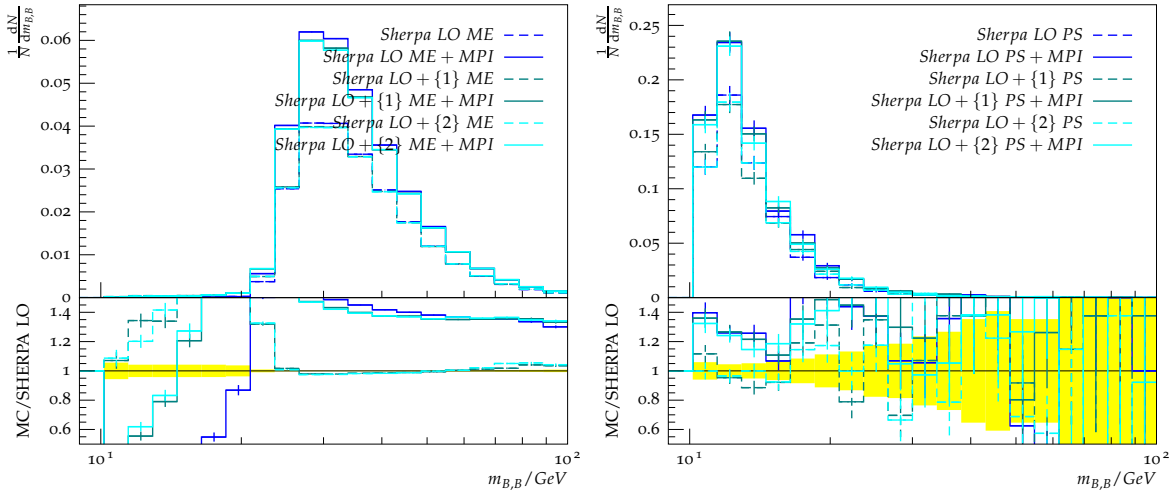
PS production From the extra-parton plots for PS production it can be seen that the LO+$\{N_e\}$ distributions offer no significant deviations in terms of distribution shapes from the LO+$\{0\}$ distributions. Again, this indicates that Sherpa's PS does a good job of approximating the higher-order ME calculation.

The acceptance deviations shown in Table 10 do suggest that the extra-parton samples have a higher acceptance, up to $(9 \pm 4)\,\%$ for LO+$\{1\}$, however the large associated uncertainty of these deviations mean that larger samples would need to be run in order to pin down exactly how much of an effect the addition of extra partons has on event acceptance. As this investigation is primarily interested in the shape of the variable distributions, the difference in acceptance will not be investigated further at the moment.

Acceptance-Ratio Comparison No large deviation can be seen either in the acceptance ratios shown in Table 9.

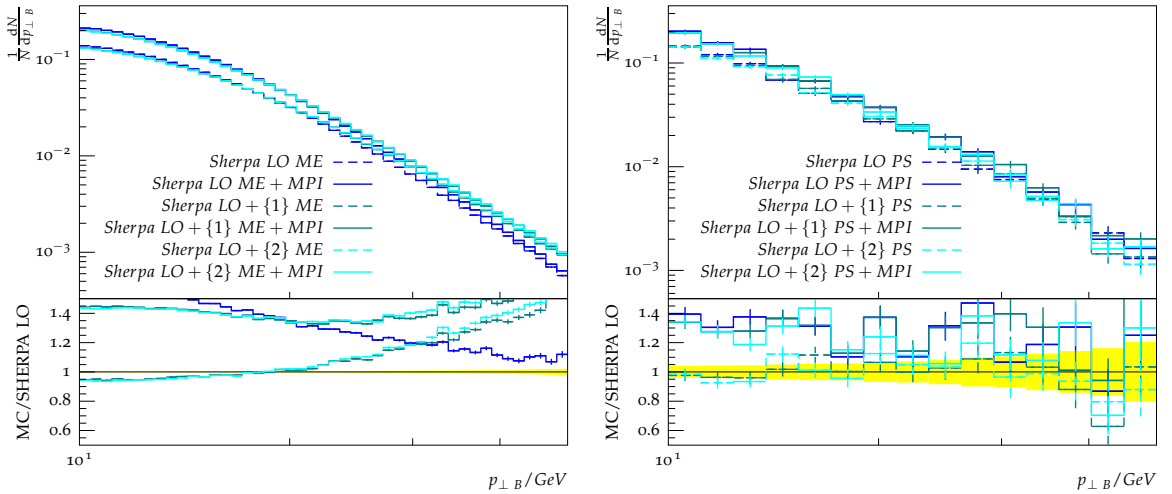
Sample Set	Percentage Deviation [%]	
	ME	PS
Sherpa LO+{1}	6.4 ± 0.1	9 ± 4
Sherpa LO+{1}+MPI	4.4 ± 0.1	7 ± 3
Sherpa LO+{2}	7.1 ± 0.1	8 ± 4
Sherpa LO+{2}+MPI	5.0 ± 0.1	4 ± 3

Table 10: Percentage deviation in acceptance of LO+$\{N_e\}$ from LO+$\{0\}$ for ME and PS production.



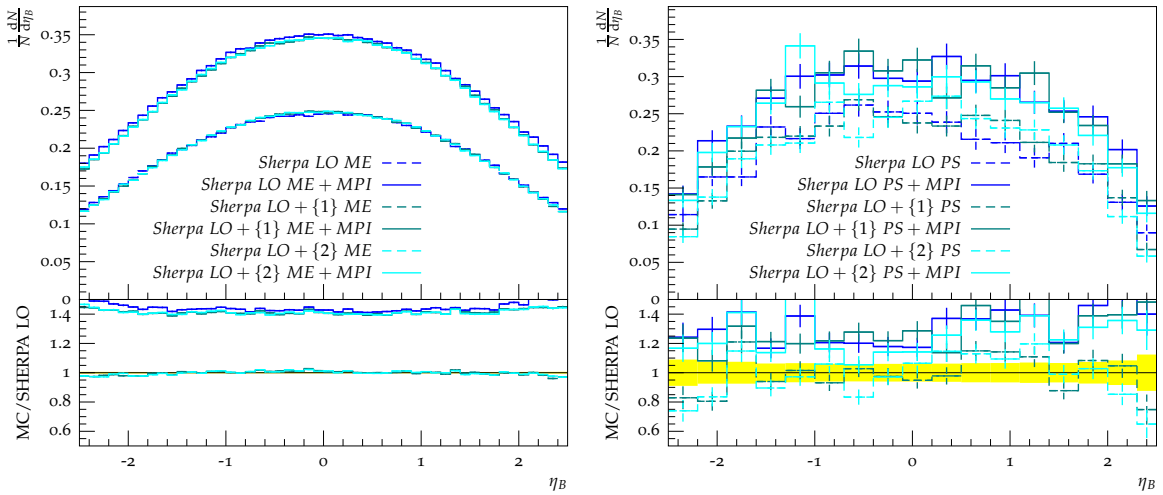
(a) ME Production. *(b) PS Production.*

Figure 18: Comparison of $m_{B,B}$ for ME and PS b-quark pair-production.



(a) ME Production. *(b) PS Production.*

Figure 19: Comparison of $p_{\perp B}$ for ME and PS b-quark pair-production.



(a) ME Production. *(b) PS Production.*

Figure 20: Comparison of η_B for ME and PS b-quark pair-production.

4.7 Conclusion

In summary it can be seen that:

- The two generators approximately agree in their cross section for the samples.
- The addition of extra light-partons in SHERPA causes no significant change to the PS-production variable-distribution shapes, indicating that LO $p\,p \to p\,p$ MEs are sufficient for simulating HF production in SHERPA's PS.
- The addition of extra light-partons in SHERPA causes a slight change to the ME-production variable-distribution shapes, indicating that LO+$\{N_e\}$ MEs are required for complete simulation of HF production in SHERPA's ME.
- In PYTHIA 8 at least, the inclusion of flavour excitation processes has a large effect on both the shape of the angular distributions, and the fractional acceptance.
- The two generators show large differences in their modelling of PS production.
- The addition of MPI has a much larger effect in SHERPA than it does in PYTHIA 8, whose acceptance increase is very small for ME production.
- PYTHIA 8's PS production has a much larger acceptance that SHERPA's suggesting than PYTHIA 8 has a higher $g \to b\,\bar{b}$ branching ratio than SHERPA.
- PS production is seen to primarily take place at low ΔR.
- ME production is seen to primarily take place around $\Delta R = \pi$.

Since the generators show such large differences in their predictions it is necessary to compare their results to real data to determine how well they are modelling the $g \to b\,\bar{b}$ process. This requires the addition of the initial-state b-quark processes which were excluded in this section. Since the investigation will focus on comparison to data, the difference between the generators' treatment of the b quark's mass will not be an issue.

The PS is seen to provide a significant contribution to the overall production of b quarks and, since the PS uses approximations rather than explicit calculations, it is important to make sure that PS modelling is being done accurately. This can be tested by comparing it to the explicit calculation of $g \to b\,\bar{b}$ in the ME.

5 Comparison of parton shower and matrix-element models for gluon splitting

5.1 Motivation

As was seen in Section 4 the PS plays an important role in the production of heavy-flavour quarks in Monte Carlo generators. It is, however, an approximation, as discussed in Section 3.7. Section 4.6.2 showed that using higher-order MEs resulted in very little deviation from the base $2 \to 2$ MEs, indicating that the PS performs well at approximating the interactions included in the higher-order MEs, however the majority of these interactions would have been light-parton radiation. This section will investigate just how well gluon splitting is modelled in the PS by comparing results of PS gluon splitting to gluon splitting in the ME, which is calculated explicitly. There are two main reasons why it could be expected that the PS and ME results might differ: 1) There is no *a priori* choice for the correct scale for the PS; 2) Gluon splitting is approximated in the same way as gluon radiation, however due to differences in colour-coherence considerations for the two processes, it is possible that using the same approximation could be incorrect.

5.2 Sample

Because MC generators simulate events by generating a hard process (a sum of Feynman diagrams) in the ME, rather than by calculating a single Feynman diagram, it is impossible to explicitly choose the gluon splitting diagrams (Figures 21a and 21b) for the process $p\,p \to b\,\bar{b}$, as opposed to the other possible diagrams (Figures 21c and 21d). What can be done, however, is to force the system into a configuration in which gluon splitting become kinematically favourable.

The ME sample used the process $p\,p \to b\,\bar{b}\,p$ with the intention of selecting events in which the $b\,\bar{b}$ pair recoils against the outgoing p as demonstrated in Figure 22 (note that only the s-channel diagram is shown here, but the selection is equally valid for the t and u channels). In this situation, the $g \to b\,\bar{b}$ is calculated in the ME.

The PS sample is somewhat easier as any production of a $b\,\bar{b}$ pair must have come from gluon splitting for a fully light ME of the form: $p\,p \to p\,p$. However, in order to make sure that the samples are comparable, the analysis of the PS sample will also require the b system to be balanced by a p. As was seen in Section 4, the acceptance efficiency of the PS sample is very low, and placing extra requirements on the event will reduce this even further. Because of this a very large number of events will need to be run in order to make the analysis results statistically useful.

Again, b hadrons are used as proxies to the b quarks. Owing to limitations in Pythia 8's generation options, this investigation will only make use of Sherpa. Generation used the *MSTW2008lo68cl* PDF [42]. Beams were set to collide 4 TeV protons. b and c quarks were set to be massive, excluding them from Sherpa's `93` light parton container. The number of EW interactions in the ME was set to zero. In order to increase the selection efficiency for the ME sample, a weak bias is used in the generation which requires that the pair of outgoing b quarks have a maximum separation of $2\pi/3$. For both samples $\hat{p}_{\perp,min}$ for the light partons was set to 10 GeV, and for the ME sample $\hat{p}_{\perp,min}$ for the b quarks was set to 5 GeV; as the two b

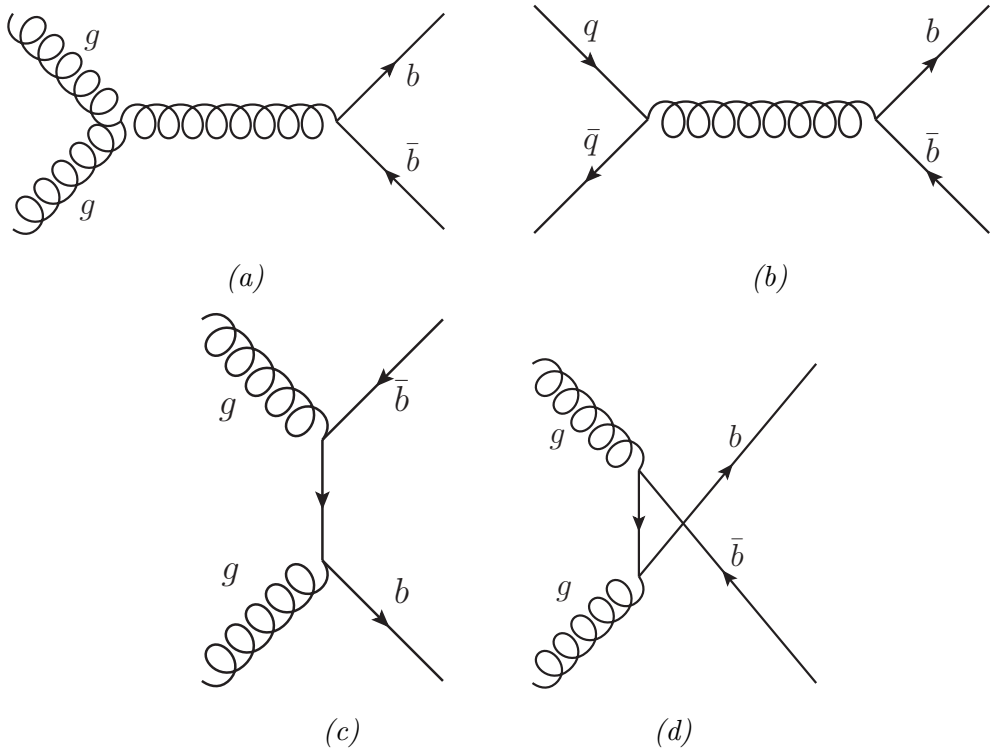


Figure 21: Figures (a) and (b) show the two $2 \to 2$ QCD diagrams for $b\,\bar{b}$ production via gluon splitting. Figures (c) and (d) show the two $2 \to 2$ QCD diagrams for $b\,\bar{b}$ production via virtual-quark exchange.

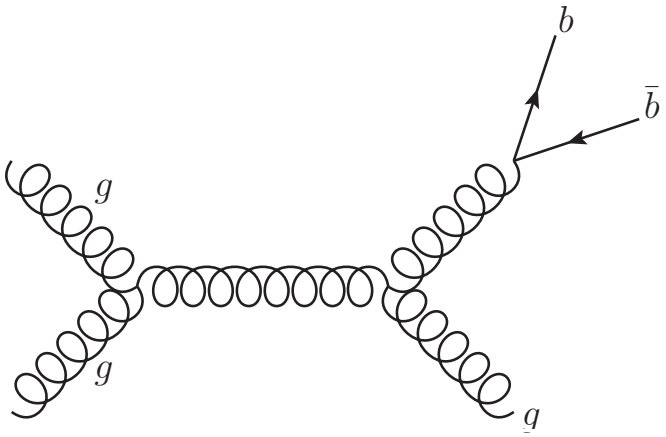


Figure 22: Possible diagram for $p\ p \to b\bar{b}\ p$, where the $g \to b\bar{b}$ splitting is calculated in the ME and the $b\bar{b}$ pair recoils against an outgoing gluon.

quarks outgoing b quarks will be recoiling against a light-parton, their $p_\perp$s should sum to the $p_\perp$ of the outgoing light parton. No MPI simulation was performed for either sample. A total of 4 999 992 and 1 000 000 164 events were produced for the ME and PS samples, respectively.

5.3 Analysis

In order to select events in which a $b\bar{b}$ pair recoils against a light parton, the analysis will look for a momentum balance between final states. This must be done using jets since light partons are likely to seed their own PSs, and the b quarks will fragment during hadronisation. This thesis assumes that b hadrons are good proxies to b quarks only for angular variables; the other products of the b quarks' fragmentations will carry some momentum, which must be accounted for in order to find balanced systems.

The analysis first requires two b hadrons in the range $|\eta| \leq 2.5$ with $p_\perp \geq 5\,\text{GeV}$, selecting the hardest two if a greater number are found. Jets are then formed using the anti-$k_\perp$ algorithm for $R = 0.4$ (single jet) and $R = 0.8$ (fat jet), with the following requirements on the minimum jet momenta: $|p| \geq 20\,\text{GeV}$ (single jet), 40 GeV (fat jet).

Since a range of system momenta will be examined, the angles between b quarks, and the radii of the jets they and the light parton produce, are likely to vary. Hence the analysis looks for several different jet-combinations and requires at least one of following three cases to be found:

- Each b hadron is contained ($\Delta R_{B,jet} \leq R$) in its own single jet and the summed momentum of the two jets is balanced by a 3$^{\text{rd}}$ single jet.
- Each b hadron is contained in its own single jet and the summed momentum of the two jets is balanced by a fat jet.
- Both b hadrons are contained ($\Delta R_{B_1,jet} \leq R$ & $\Delta R_{B_2,jet} \leq R$) in a fat jet whose momentum is balanced by a single jet.

Momentum balance requires that:

$$\frac{|\overline{p_{b\bar{b}}} + \overline{p_{\not p}}|}{|\overline{p_{\not p}}|} \leq 0.1,$$

where $\overline{p_{b\bar{b}}}$ is the momentum vector of the b system (the two b jets or the single b fat-jet) and $\overline{p_{\not p}}$ is the momentum vector of the $\not p$ jet.

If multiple jet-configurations are found which meet these criteria then the configuration with the closest momentum balance between the b system and the $\not p$ jet is selected.

Having selected the best jet-configuration, the kinematic and geometric properties of the b hadrons are plotted in a plot set determined by $|p_{b\bar{b}}|$. The plots sets are divided thusly:

- Bin 0: $40 \leq |p_{b\bar{b}}| < 60$ [GeV]
- Bin 1: $60 \leq |p_{b\bar{b}}| < 80$ [GeV]
- Bin 2: $80 \leq |p_{b\bar{b}}| < 100$ [GeV]
- Bin 3: $100 \leq |p_{b\bar{b}}| < 120$ [GeV]
- Bin 4: $120 \leq |p_{b\bar{b}}|$ [GeV]

5.4 Results and discussion

Figure 23 details the population of each of the $|p_{bb}|$ bins. Geometric and kinematic distributions are shown in Figure 24 and Figure 25 respectively.

Figure 24 shows that the PS does a good job of approximating the ME results, peaking in the same bins as the ME results and following the same shape as the ME distributions. There are a few slight differences, however, as the PS distributions tend to underestimate the tail areas (particularly noticeable in the $\Delta\phi$ plot in Figure 24b) and instead overproduce in the low-angle or peak regions. These differences propagate through to the ΔR plot (Figure 24c), where although both ME and PS distribution peak in the same bin, the PS distributions still underestimate the tail areas.

Figure 25b ($p_{\perp B}$) shows good agreement between the PS and ME results for all $|p_{bb}|$ bins. The shape of the $|p_{bb}|$ distributions for the PS and ME results are also seen to be very similar in Figure 25c.

Figure 25a shows that the mass distributions of the b hadron pairs do differ slightly, with the PS overestimating the production of low mass pairs. It does, however, follow the same trend with $|p_{bb}|$ as the ME results: as $|p_{bb}|$ increases, the distribution shifts to higher-mass pairs.

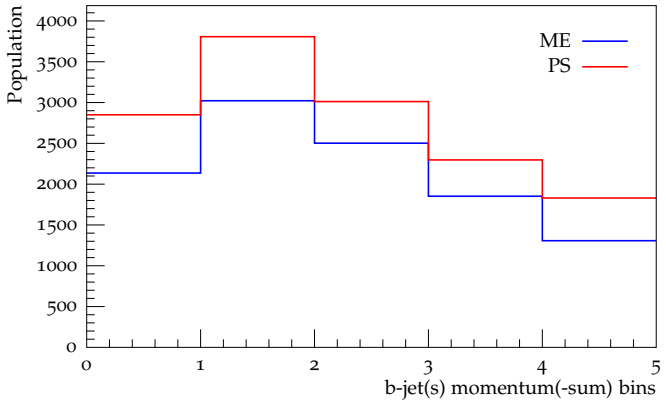
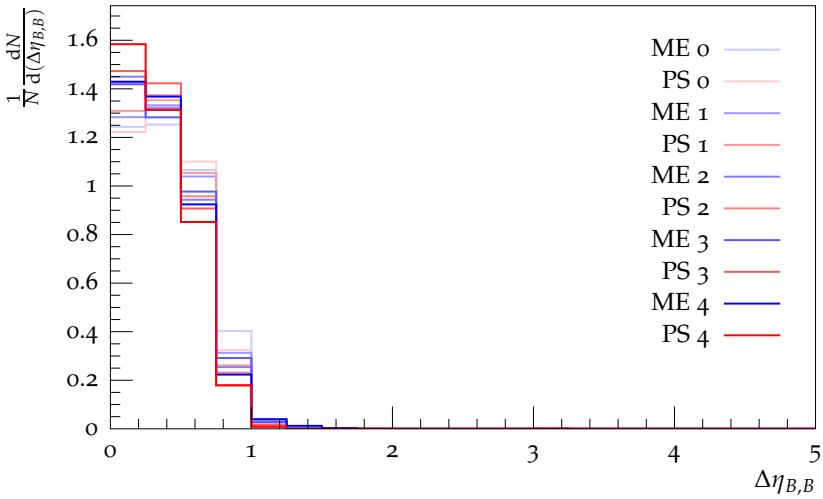
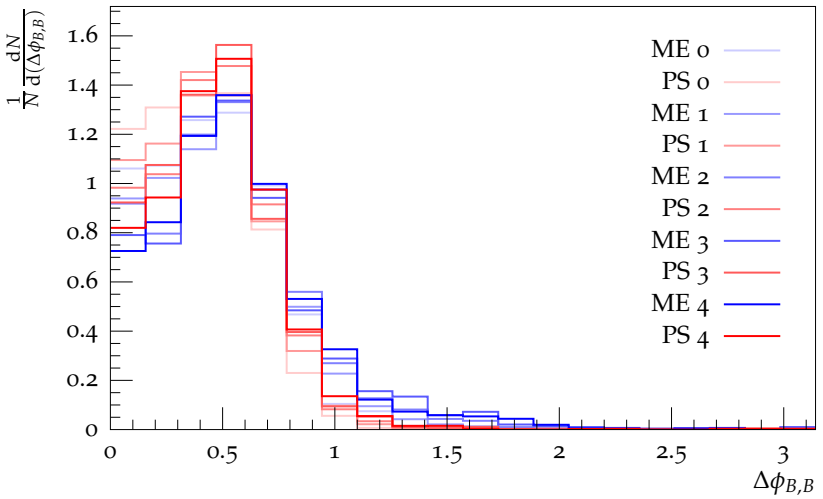


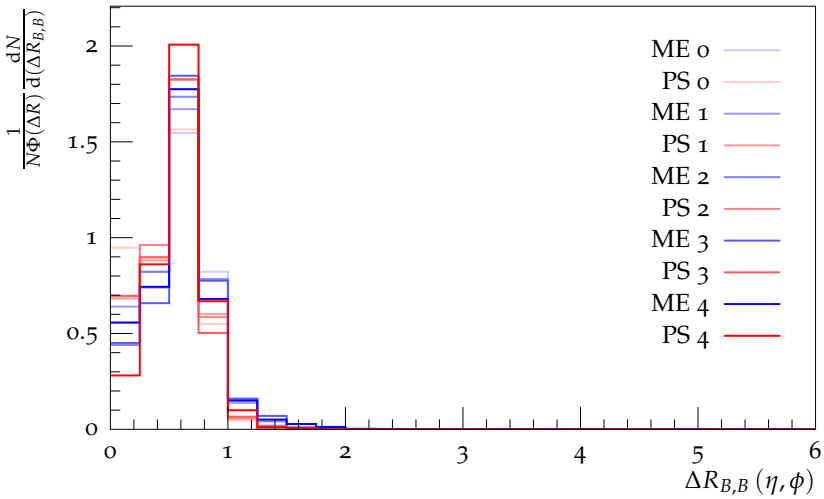
Figure 23: Unnormalised bin-population for ME and PS results. The bin numbers correspond to the list shown in Section 5.3.



(a) $\Delta\eta\,(B,B)$ *distributions for ME and PS results for all bins.*

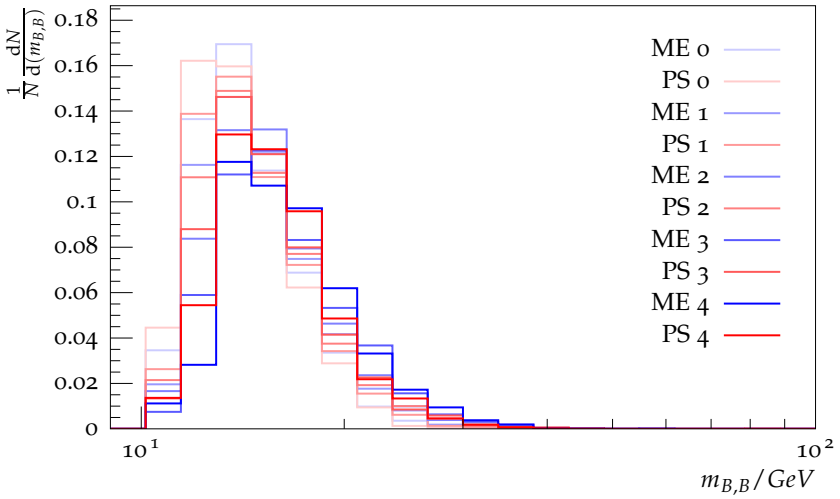


(b) $\Delta\phi\,(B,B)$ *distributions for ME and PS results for all bins.*

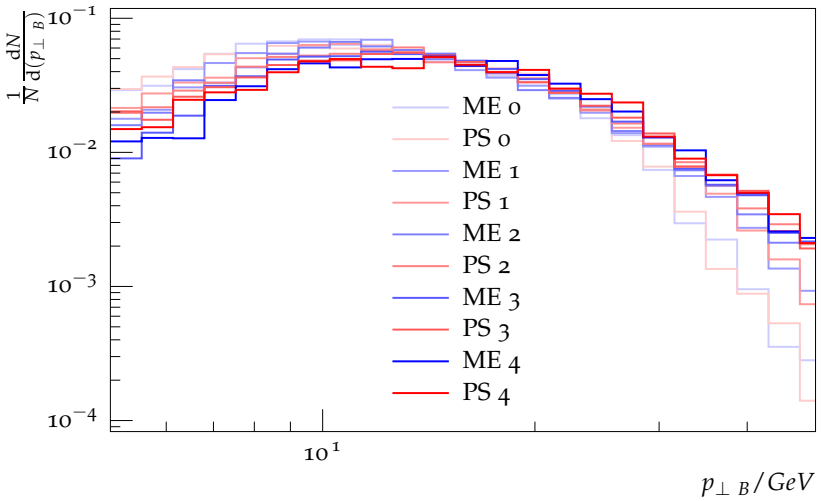


(c) $\Delta R\,(B,B)$ *distributions for ME and PS results for all bins. Plotting accounts for phase space.*

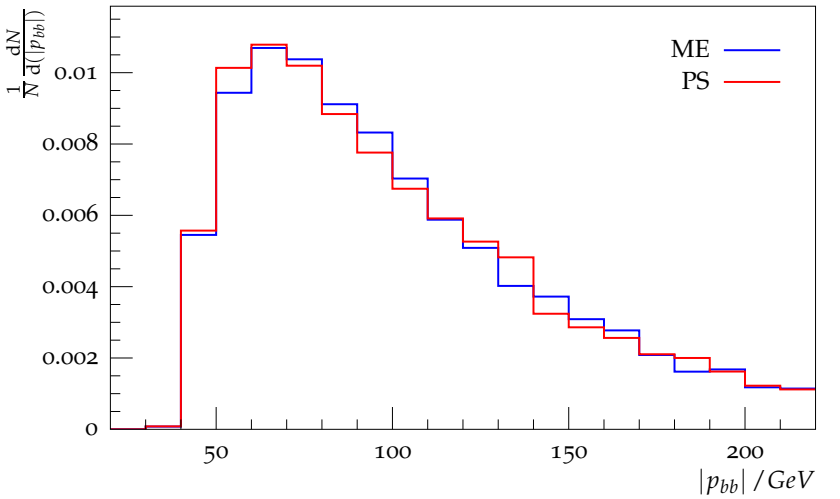
Figure 24: Angular distributions for the ME and PS results for all bins. The bin numbers correspond to the list shown in Section 5.3. All plots are normalised to one.



(a) m_B distributions for ME and PS results for all bins.



(b) $p_{\perp B}$ distributions for ME and PS results for all bins.



(c) $|p_{bb}|$ distributions for b system of ME and PS results for all bins.

Figure 25: Kinematic distributions for the ME and PS results for all bins. The bin numbers correspond to the list shown in Section 5.3. All plots are normalised to one.

5.5 Conclusion

The PS approximation for gluon splitting to b-quark pairs shows a clear tendency to underestimate high-angle production, where the analytical calculation of the ME shows that there should be a significant tail distribution. The PS results do, however, peak at generally the same point as the ME distributions for all three angles, and show momenta distributions which are very similar the ME results.

By altering the scales for the PS simulation, the results could be tuned to show a better match to the ME results, however that is beyond the scope of this thesis, and the PS as it stands is good enough for the investigation being performed here.

6 Using J/ψs as b-hadron proxies

6.1 Motivation

The colour confinement property of QCD means that the b quarks will be bound into b hadrons, however as can be seen in Table 4, the longest b hadron mean lifetime is 1.641 ps. The hadronic calorimeter in the ATLAS detector begins at a radius of 2.28 m [3], so the b hadrons will decay before reaching the calorimeter, preventing them from directly providing accurate measurements of the b quarks. Another intermediary is required.

As mentioned in Section 3.6.3, ATLAS's muon spectrometers provide accurate measurements for muons escaping the inner parts of the detector, and is able to reconstruct them efficiently down to around $p_\perp = 2.5\,\text{GeV}$ [43]. Muons can provide a good means for detecting a signal, especially if the signal is a process which produces two muons, as the tracking information provided by the ID can be used to see if the two muons can be constrained to a common production-vertex. The channel is also very clean since the absorber plates in the calorimeters are designed to prevent other particles from escaping the detector. b hadrons, however, are highly unlikely to decay directly to two muons ($\Gamma_{B\to\mu^+\mu^-+anything}/\Gamma < 3.2\times10^{-4}\,\%$ at 90 % confidence limit for a $B^\pm/B^0/B_s^0/b$-baryon admixture [13]).

As mentioned in Section 2.2.3, the J/ψ particle does have a relatively large branching-fraction to dimuon of 5.961% [13], and the b hadrons are able to decay to J/ψs, ($\Gamma_{B\to J/\psi+anything}/\Gamma$ = 1.16 % for a $B^\pm/B^0/B_s^0$/b-baryon admixture [13]).

The process $b \to B \to J/\psi \to \mu^+\mu^-$ has been measured in previous investigations by ATLAS [44], CMS [45], and LHCb [46], and so can provide a signal which can be detected and measured to a high enough precision in the ATLAS detector, but presents two problems: J/ψs can also be produced in the initial collision, and the production cross-section ($\sigma(pp \to bX \to J/\psi X) \cdot BR(J/\psi \to \mu^+\mu^-)$) is very small:

$$\begin{aligned}
\text{ATLAS}: 23.0 \pm 0.6(\text{stat.}) \pm 2.8(\text{syst.}) \pm 0.8(\text{lumi.}) \pm 0.2(\text{spin})\,\text{nb} \\
\text{for } p_{\perp J/\psi} > 7\,\text{GeV},\ |y_{J/\psi}| < 2.4\ [44], \\
\text{CMS}: 26.0 \pm 1.4(\text{stat.}) \pm 1.6(\text{syst.}) \pm 2.9(\text{lumi.})\,\text{nb} \\
\text{for } 6.5 < p_{\perp J/\psi} < 30\ [\text{GeV}],\ |y_{J/\psi}| < 2.4\ [45].
\end{aligned}$$

The first problem can be overcome by careful analysis of the J/ψ decay vertex's displacement from the primary vertex and will be discussed in detail in Section 7.

Expanding on the second problem: extending the MC analyses from Section 4 to requiring both b hadrons to decay to J/ψs would cause a reduction in acceptance of at least 99.99% which, for a statistically precise result, would require a very long processing time. The fraction $M_{J/\psi}/M_B$ is in the range 0.5 to 0.6 so it is likely that the J/ψ will be the heaviest of the b hadron's decay products and that its direction is unlikely to be largely different from that of

the parent b hadron. Since the investigation is primarily interested in the angular variables of the $g \to b\bar{b}$ process, if it can be shown that $\Delta R_{B,J/\psi}$ is small, then MC results can be compared to real data collected by the ATLAS detector without requiring $B \to J/\psi$ decays in the MC data.

The problem of low J/ψ production still exists, but PYTHIA 8 has been modified to be able to force the decays of hadrons along specified decay chains. This section will use this modified version, called PYTHIA 8B, to investigate how well the J/ψs can be used as proxies to the b hadrons.

6.2 Pythia 8B sample

The ATLAS Monte Carlo sample *Pythia8B_AU2_CTEQ6L1_pp_Jpsimu0mu0*, produced in PYTHIA 8B for the MC12 Production Campaign, was used. This dataset consists of four million events of inclusive proton-proton $\to J/\psi \to \mu\mu$ production. The sample used the LO *CTEQ6L1* PDF with beams set to collide 4 TeV protons. The cross section was $(1.805 \pm 0.003) \times 10^4$ nb.

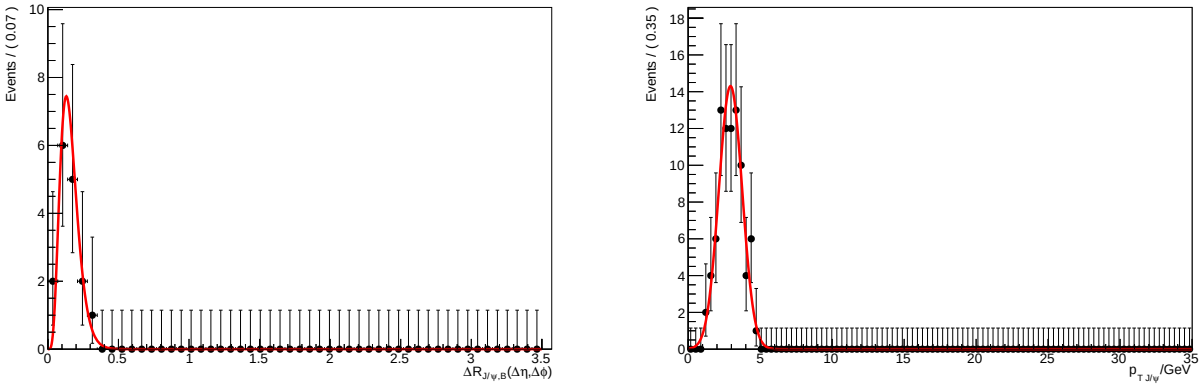
6.3 Pythia 8B analysis

The PYTHIA 8B sample was analysed using RIVET via version 19.1.1.1 of ATLAS's ATHENA framework. The analysis again looked for b hadrons using RIVET's `HeavyHadrons` projection in the range $-2.5 \leq \eta \leq 2.5$, requiring $p_{\perp,B} \geq 1$ GeV. The decay chain of any b hadrons found is then scanned for J/ψs and if one is found then that particle pair is accepted. The kinematic and geometric properties of the pair are then plotted with particular focus on the path deviation of the child J/ψ from the parent b hadron.

6.4 Results

2095 J/ψ b-hadron pairs were selected by the analysis. These pairs were then binned according to the $p_\perp$ of the parent b hadrons and within each bin gamma-functions were fitted to the distribution of data points for the angular variables ($\Delta R_{B,J/\psi}$, $\Delta\phi_{B,J/\psi}$, and $\Delta\eta_{B,J/\psi}$) and Gaussian functions were fitted to the distributions of J/ψ $p_\perp$. Fitting was performed using RooFit [47], which was set to account for the statistical uncertainty of the data points. Figure 26 shows examples of these fits.

The mean (μ) values of the fitted functions were then plotted as a function of the central values of the b hadron $p_\perp$ bins. The y-axis error bar for each point was set to the standard deviation (σ) of the fitted function within that bin. Next, power-law functions were fitted to the mean angular-variable distributions and a linear function was fitted to the mean J/ψ $p_\perp$ distribution. Fitting took place three times per variable to the values of: μ, $\mu - \sigma$, and $\mu + \sigma$. These three fitted functions together allow the mean and standard deviation of distributions to be determined at any point, which will be used in Section 8.4 to construct $B \to J/\psi$ smearing distributions. Fitting again used RooFit, set to account for the statistical uncertainty on points and the systematic uncertainty on the mean values resulting from fitting to the binned distributions. The results of the mean-value fitting are summarised in Table 11 and the fitted plots are shown in Figure 27.



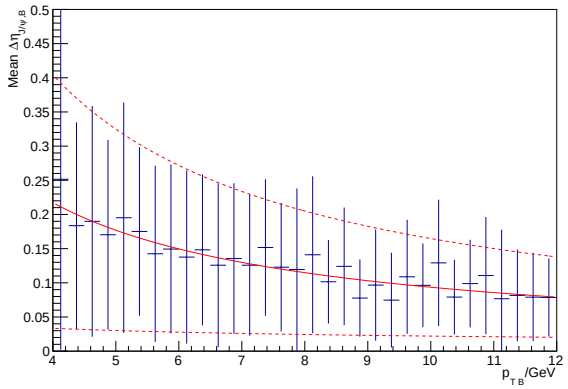
(a) Distribution of $\Delta R_{J/\psi,B}$ for $10.5 \leq p_{\perp B} < 10.75$ [GeV]. Distribution is fitted to with a gamma function, shown in red.

(b) Distribution of $p_{\perp J/\psi}$ for $4.25 \leq p_{\perp B} < 4.5$ [GeV]. Distribution is fitted to with a Gaussian function, shown in red.

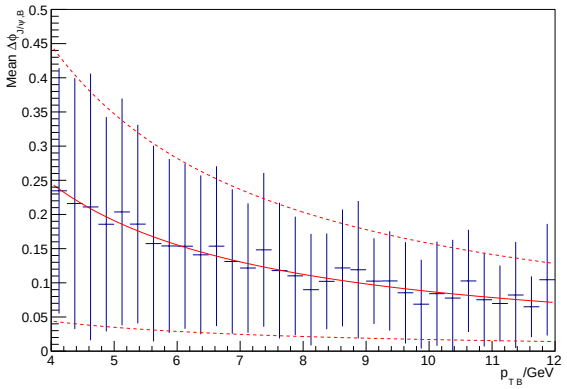
Figure 26: Two examples of the distribution fitting which took place within each $p_{\perp B}$ bin. y-axis Error bars are the statistical uncertainty on each point.

$$\begin{aligned}
\left\langle \Delta R_{J/\psi,B} \right\rangle &= (1.45^{2.38}_{0.54})\, p_{\perp B}^{\left(-1.00^{-1.03}_{-0.90}\right)} \\
\left\langle \Delta \phi_{J/\psi,B} \right\rangle &= (1.17^{2.17}_{0.19})\, p_{\perp B}^{\left(-1.13^{-1.14}_{-1.04}\right)} \\
\left\langle \Delta \eta_{J/\psi,B} \right\rangle &= (0.78^{1.57}_{0.06})\, p_{\perp B}^{\left(-0.92^{-0.98}_{-0.45}\right)} \\
\left\langle p_{\perp J/\psi} \right\rangle &= (0.63^{0.72}_{0.53})\, p_{\perp B} + \left(0.09^{0.55}_{-0.38}\right)
\end{aligned}$$

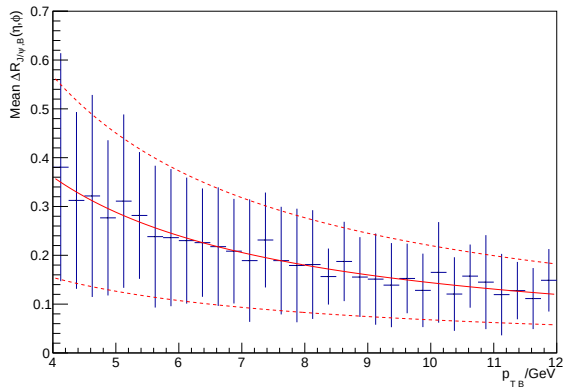
Table 11: Equations for the lines of best fit to the $B \to J/\psi$ correlation data. Parameters are shown in the form $\left(a^b_c\right)$, where a, b, and c are for the fits to the mean values, mean+standard deviation values, and mean-standard deviation values, respectively. Parameters are shown to two decimal places, but will be used at six significant figures.



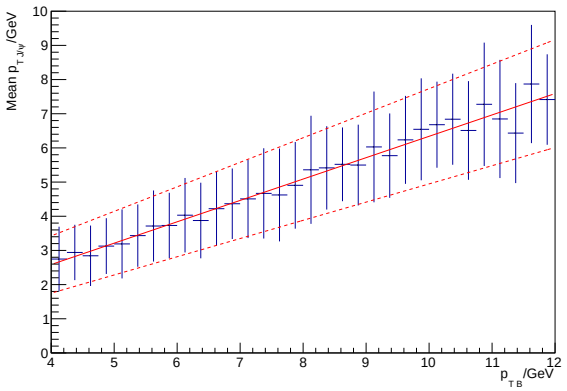
(a) Mean $\Delta\eta_{J/\psi,B}$ as a function of $p_{\perp B}$.



(b) Mean $\Delta\phi_{J/\psi,B}$ as a function of $p_{\perp B}$.



(c) Mean $\Delta R_{J/\psi,B}$ as a function of $p_{\perp B}$.



(d) Mean J/ψ $p_\perp$ as a function of $p_{\perp B}$.

Figure 27: Angular and $p_\perp$ correlations between J/ψs and their parent b-hadrons. The solid lines are the fit to the mean values and the dotted lines are the fit to the mean±standard deviation values. The y-axis error bars indicate the standard deviation of the functions fitted to the distribution of points within each b hadron $p_\perp$ bin.

6.5 Discussion and conclusion

As can be seen in Figure 27, the performance of J/ψs as proxies to b hadrons has a large dependence on the $p_\perp$ of the parent b-hadron. For very soft b-hadrons, the child J/ψ is normally emitted at a large angle to the original direction of travel of the b hadron, however as the b hadron's $p_\perp$ increases, the deviation of the J/ψ's flight path decreases rapidly, and for b-hadron $p_\perp$s greater than 3.7 GeV the mean $\Delta R_{J/\psi,B}$ is less than 0.4, i.e. within a standard ATLAS jet. This indicates that J/ψs may be used as good proxies to b hadrons even down to reasonably low $p_{\perp B}$ values.

The angular equations in Table 11 can be used to construct distributions to smear di-B angular measurements allowing them to be compared to di-J/ψ angular measurements, and the $p_\perp$ relationship equation can be used to calculate any necessary $p_\perp$ cuts on the J/ψs or b hadrons to ensure that the system is in a region where the J/ψs are good proxies to the b hadrons.

It would be useful to be able to run a similar investigation in Sherpa by forcing its b hadron decays, however, as this functionality is currently unavailable, the correlation functions derived here will used with both Pythia 8 and Sherpa data in Section 8.

At the time of investigation, only single J/ψ Pythia 8B samples were available, however it would be of use to analyse a di-J/ψ sample to see if there was any correlation between the

mean separation between a J/ψ and its parent b hadron and the angle between the parent b hadrons. It will be assumed for the remainder of this thesis that there is no direct correlation.

7 Di-J/ψ sample investigation

7.1 Analysis plan

As was shown in Section 6, J/ψs can be used as good proxies to b hadrons. However, as was mentioned in Section 6.1, J/ψs can be produced in either the initial collision (prompt) or through the decay of b hadrons (non-prompt). It is therefore necessary to find some way of preparing a data sample whose distributions reflect the contributions from non-prompt J/ψs.

To do this a di-J/ψ sample is used and a model consisting of functions corresponding to the various prompt and non-prompt contributions is fitted to the mass and pseudoproper lifetime of each J/ψ in the data events. The sample is collected in the di-muon channel ($J/\psi \to \mu^+\mu^-$) and fitting takes place to the properties of each reconstructed di-muon vertex. Fitting to the mass allows for the selection of J/ψs as opposed to some background particle, and fitting to the pseudoproper lifetime allows for discrimination between prompt and non-prompt particles.

Two forms of fitting are used: an inclusive fit, which uses all available data, and a binned fit, where data is divided into bins according to geometric variables. The inclusive fit forms an initial investigation to verify the applicability of the fitting functions and so uses much rougher event selection than the final binned fit method uses.

Having fitted to the data, the probability of a given event containing two non-prompt J/ψs can be calculated. By weighting the di-J/ψ variable distributions by these probabilities, the distributions can be altered to reflect the contributions from J/ψs which originated from the decays of b hadrons.

This section of the thesis builds upon fitting code written and developed by Gavin Hesketh and Josh McFayden of University College London (the UCL researchers). In order to give credit where it is due, work done by the author and that done by the UCL researchers will be explicitly specified. All discussions and conclusions in this section are the words of the author.

7.2 Di-J/ψ sample

The di-J/ψ sample was selected from the 2012 $\sqrt{s} = 8\,\text{TeV}$ ATLAS data for proton-proton collisions. The initial pre-trigger sample corresponded to an integrated luminosity of $20.2814\,\text{fb}^{-1}$. The following event requirements were then applied by Kostas Karakostas [48]:

- Di-muons selected according to the DiMu algorithm:
  - Event contains a $\mu^+\mu^-$ pair, where at least one muon is "combined"[14].
  - Muon pair is in the mass range (2.2 to 4.0) GeV.
  - Muon-pair fits to a common vertex with $\chi^2 < 10\,000$ using `VKalVert` [49].
- Event must contain at least two di-muons.
- Di-muons are ordered by $p_\perp$ and the hardest two selected.
- Di-muon pair must fit to a common vertex with $\chi^2 < 10\,000$ using `VKalVert`.

[14]Detected in both the ID and the MS

The ntuple sample, *flat_v7_user.kara.data12_8TeV.allBPhys_r1v1.DiJpsi.repo3.v9.root*, created by Karakostas consists of 228 071 events. Each event contains two particles which are the reconstructed J/ψ vertices. These are referred to in this section as: the J/ψs, the J/ψ candidates, or the particles. Ideally they are J/ψs, however in order to determine whether they really are J/ψs or some background particle, and whether they are prompt or non-prompt, a fit is used. The vertices are labelled as J/ψ_1 and J/ψ_2 and are ordered by vertex $p_\perp$, with J/ψ_1 being the hardest.

Since the inclusive fit and binned fit use differing event selections, further cuts to the data sample (mass, lifetime, muon and trigger requirements *et cetera*), will be described in each fit's write-up section. Following the event selection described in Ref. [44], no further χ^2 cuts are directly applied to the quality of the di-muon and di-J/ψ vertices, however the trigger requirements applied in event selection for the binned fit do require higher-quality vertices. The χ^2 of a vertex fit functions as a usual χ^2 for small values, but at larger values is a monotonic function of the fit quality [49].

7.3 Inclusive fit

Initially, a four dimensional inclusive fit was performed in which functions were simultaneously fitted to the J/ψ mass and pseudoproper lifetime (τ) for each J/ψ in the event.

7.3.1 Additional cuts

Being an initial investigation, the sample for the inclusive fit underwent very few cuts beyond the initial requirements in Section 7.2, and no account of muon reconstruction efficiency was made.

Mass and pseudoproper-lifetime cuts Data for the fit was required to pass the following cuts on both J/ψs in the event:

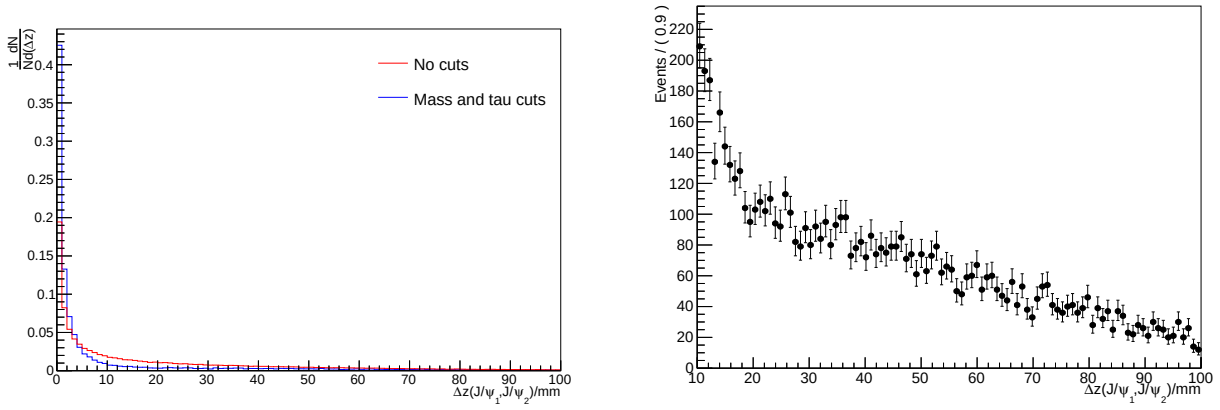
- $2.6 \leq m_{J/\psi} \leq 3.5$ [GeV]
- $-0.5 \leq \tau_{J/\psi} \leq 2.0$ [ps]

The values for these cuts were chosen by the UCL researchers and applying them reduced the sample from 228 071 events to 31 668 events.

Pile-up removal As mentioned in Section 3.5.1, the data sample is likely to contain contributions from pile-up, where there are multiple hard interactions within a single bunch crossing. Looking at the z-axis separation of the J/ψs gives an indication of which events are the product of pile-up. Figure 28 shows the Δz distribution for the J/ψ pairs before and after the mass and pseudoproper-lifetime cuts described above.

Looking at Figure 28b, it can been seen that there is a clear deviation around $\Delta z = 15\,\text{mm}$, this is where the signal data begins. Figure 28a shows that the mass and pseudoproper-lifetime cuts to a good job of reducing the the fraction of pile-up events in the sample, but a sizeable proportion still remains. Placing a cut requiring that $\Delta z(J/\psi_1, J/\psi_2) \leq 15\,\text{mm}$ should remove a lot of the pile-up contribution in the sample. By fitting a linear function to Δz in the region $\Delta z \geq 15\,\text{mm}$, extrapolating it back to $\Delta z = 0\,\text{mm}$, and integrating over the range $0\,\text{mm} \leq \Delta z \leq 15\,\text{mm}$, it was estimated that 6 % of remaining events passing the Δz cut were due to pile-up.

The investigation of pile-up contamination in the sample was performed by the author, who also chose the Δz cut value. Applying this cut on top of mass and pseudoproper-lifetime cuts further reduced the sample to 25 299 events.



(a) *(b)*

Figure 28: Figure (a) shows $\Delta z(J/\psi_1, J/\psi_2)$ before and after the cuts, where the distributions are normalised to one. Figure (b) shows $\Delta z(J/\psi_1, J/\psi_2)$ distribution for $\Delta z > 10\,mm$ after mass and pseudoproper-lifetime cuts. It should be emphasised that, although approximately two thirds of events in the region $10\,mm \leq \Delta z \leq 15\,mm$ appear to be due to pile-up, the overall fraction of events shown in Figure (b) is small compared to the total number of events, and pile-up is estimated to account for just 6 % of remaining events below $\Delta z = 15\,mm$.

7.3.2 The fit

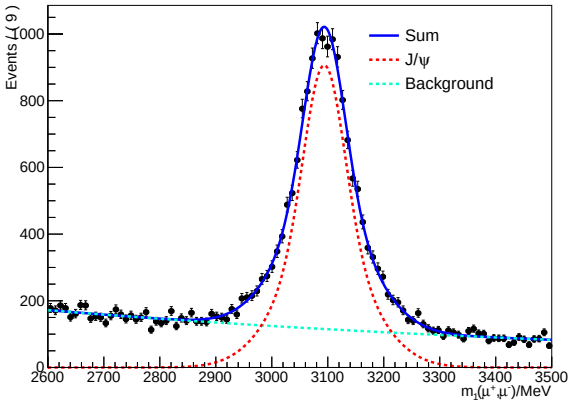
The fit, and associated fitting and plotting code used in this sub-subsection was written and developed by the UCL researchers. The fit consists of models similar to those used in Ref. [44] with some changes to the functions used.

Mass fit The mass distribution will contain contributions from prompt and non-prompt J/ψs (signal) and background. The signal region is fitted with a Gaussian and a Crystal Ball function, whilst the background is modelled with an exponential.

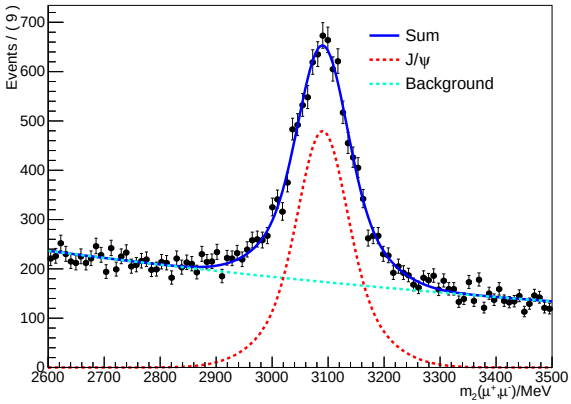
Pseudoproper-lifetime fit The fitting to pseudoproper lifetime uses a resolution model, which takes the form of two Gaussian functions with a shared mean. The signal in the τ distribution are non-prompt J/ψs, which is fitted with a single-sided decay function convolved with the resolution model. The background consists of prompt J/ψs, modelled by the resolution model itself, and other background particles, modelled by a symmetric double-sided decay function convolved with the resolution model.

Fitting With two possibilities for each J/ψ according to their mass, and three possibilities according to their pseudoproper lifetime, each event can be categorised into one of thirty six possible states. The probability density functions (PDFs) of these states are fitted simultaneously to the data by altering the defining variables of the functions which contribute to the mass and pseudoproper-lifetime distributions described above, and then normalising the PDFs to the number of data points in the sample. This is done using RooFit.

Fitting results Figure 29 shows the results of fitting to the mass of candidate J/ψs, and Figure 30 shows the results of fitting to their pseudoproper lifetime. From these figures it can be seen that the functions used appear to fit well to the data.

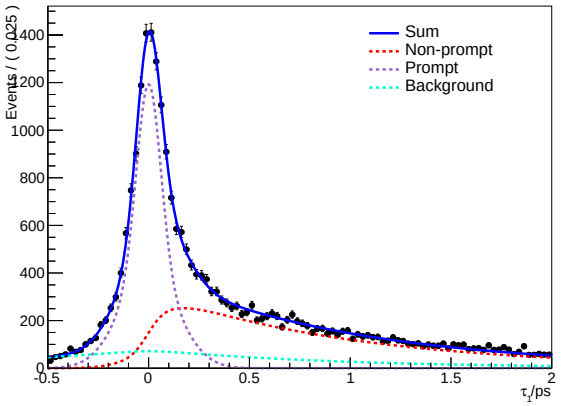


(a) Inclusive fit to the mass of particle 1.

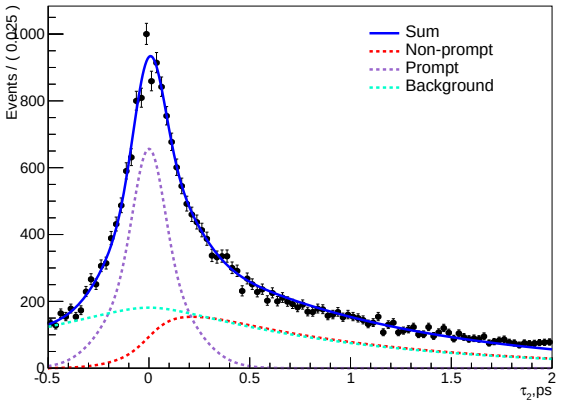


(b) Inclusive fit to the mass of particle 2.

Figure 29: Results of inclusive fit to particles' masses. "J/ψ" is a Gaussian and "Background" is an exponential.



(a) Inclusive fit to the pseudoproper lifetime of particle 1.



(b) Inclusive fit to the pseudoproper lifetime of particle 2.

Figure 30: Results of inclusive fit to particles' pseudoproper lifetimes. "Non-prompt is a single-sided decay convolved with a double Gaussian, "Prompt" is a double Gaussian, and "Background" is a single-sided decay convolved with a double Gaussian.

7.3.3 Data reweighting

The work detailed in this sub-subsection, and associated code, was performed, written, and developed by the author, except for the construction and fitting of the variant fits in the systematic uncertainty investigation, which was adapted by the author from the fitting code written by the UCL researchers. The variant fits were based on work found in Ref. [44].

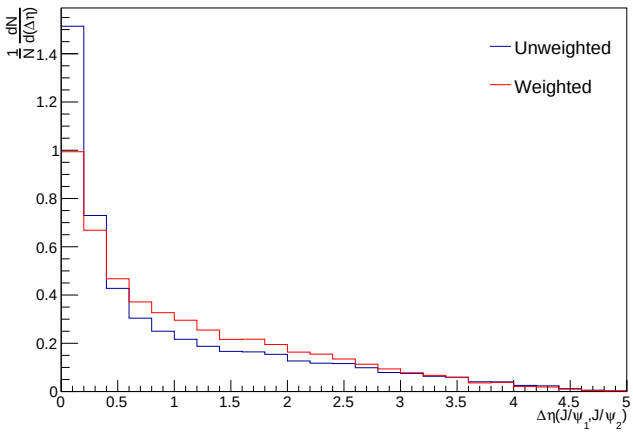
Having built up distributions for the components which make up the overall distributions for the data, it is now possible to extract the probability that a given event contains two non-prompt J/ψs by sampling the PDFs at the points corresponding to the mass and pseudoproper lifetime of the two J/ψ candidates in the event. By weighting events in the data sample by this probability, the contributions from events which do not contain two non-prompt J/ψs are suppressed and the distributions shift towards those of a pure non-prompt di-J/ψ sample. The advantage of using reweighting rather than cutting to get the required events is that it avoids selection inefficiencies: as mentioned in Section 3.3, finite detector resolution leads to a systematic uncertainty in pseudoproper lifetime, so placing a cut at $\tau_{J/\psi} > 0\,\text{ps}$ would still allow through some prompt J/ψs, but placing a higher cut would remove some non-prompt J/ψs; similarly in mass, the J/ψs are found in the peaked region, but it is unclear where exactly this begins and ends.

Probability extraction Each event contains two candidate particles; by sampling the signal and background PDFs at the mass and pseudoproper lifetime of each candidate, the probability that both of the particles are non-prompt J/ψs may be calculated using:

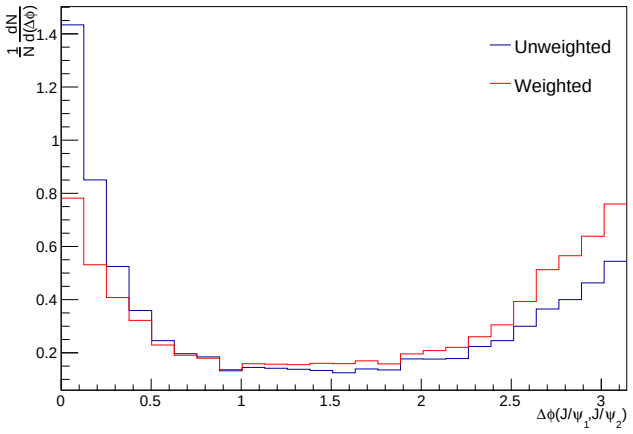
$$\begin{aligned} P_{\text{NP}J/\psi,\text{NP}J/\psi}(m_1, m_2, \tau_1, \tau_2) = &\frac{\text{Mass Signal}_1(m_1)}{\text{Mass Signal}_1(m_1) + \text{Mass Bkg.}_1(m_1)} \\ &\times \frac{\text{Mass Signal}_2(m_2)}{\text{Mass Signal}_2(m_2) + \text{Mass Bkg.}_2(m_2)} \\ &\times \frac{\tau\ \text{Signal}_1(\tau_1)}{\tau\ \text{Signal}_1(\tau_1) + \tau\ \text{Bkg.}_1(\tau_1)} \\ &\times \frac{\tau\ \text{Signal}_2(\tau_2)}{\tau\ \text{Signal}_2(\tau_2) + \tau\ \text{Bkg.}_2(\tau_2)}, \end{aligned} \tag{7.1}$$

where $\text{Mass Signal}_\alpha$ and $\tau\ \text{Signal}_\alpha$ are the PDFs for the signal contributions to the mass and pseudoproper-lifetime distributions, respectively, for candidate α. Mass Bkg._α and $\tau\ \text{Bkg.}_\alpha$ are the corresponding background PDFs.

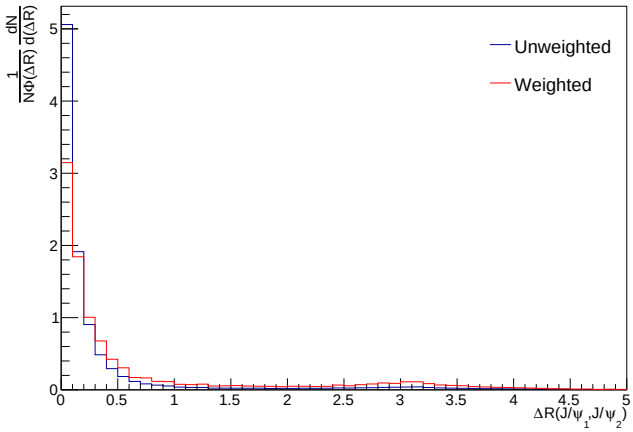
Reweighting results Figure 31 shows the angular distributions for the cut di-J/ψ sample before and after reweighting. The reweighted distributions have been filled with weight equal to the probability of the event containing two non-prompt J/ψs according to the inclusive fit.



(a) $\Delta\eta(J/\psi_1, J/\psi_2)$ *before and after reweighting.*



(b) $\Delta\phi(J/\psi_1, J/\psi_2)$ *before and after reweighting.*



(c) $\Delta R(J/\psi_1, J/\psi_2)$ *before and after reweighting. Distributions account for phase space.*

Figure 31: Angular distributions for the cut di-J/ψ sample before and after reweighting to the probability of the events containing two non-prompt J/ψs. All distributions are normalised to one. In all cases the effect of weighting the sample serves to shift the distributions to higher separations. This is most noticeable in ΔR and $\Delta\phi$, where the zero peak is decreased in favour of the π peak. This indicates that non-prompt J/ψs have a higher likelihood to be produced in a back-to-back scenario than prompt J/ψs or the background in the sample.

Systematic uncertainties The model described in Section 7.3.2 is just one of many possible ways of constructing distributions to fit the data. By using different functions in the fit, slightly different results may be obtained. The deviation of these variant fits from the default fit can be taken as the systematic uncertainty on the default fit.

The following five variant fits were used:

1. The double Gaussian resolution model in the τ fit is replaced by a single Gaussian resolution model.

2. The exponential background function in the mass fit is replaced with a 1^{st}-order polynomial.

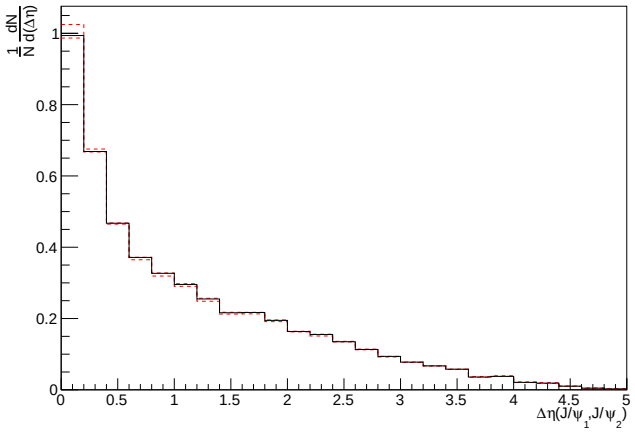
3. The mass range for data is increased to (2.2 to 4.0) GeV, in order to have a better understanding of the background in the mass fit. This variation had up to 67 036 events available for fitting.

4. The symmetric double-sided decay function in the τ fit is allowed to be asymmetric (have different positive and negative slopes).

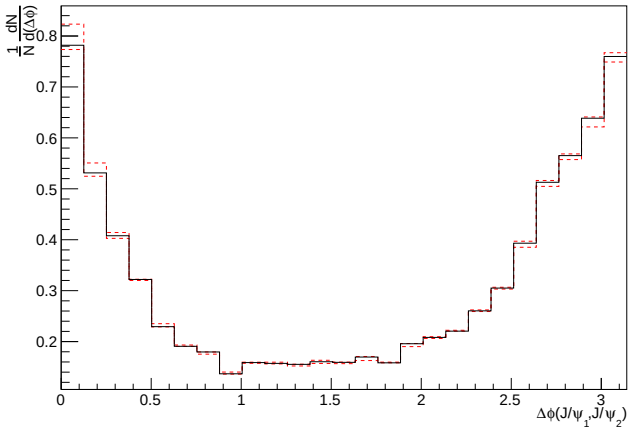
5. The symmetric double-sided decay function in the τ fit is replaced with two asymmetric double-sided decay functions whose positive slopes are the same, but whose negative slopes may differ.

In order to calculate the systematic uncertainties on the reweighted variable distributions, the distributions were weighted according to each of the variant fits in turn and the quadrature sum of the deviations above and below the default-reweighted distributions were calculated.

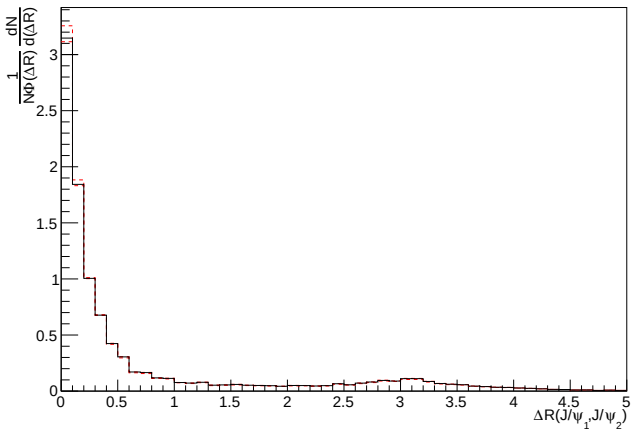
Figure 32 shows the reweighted distributions along with their systematic uncertainties due to the fit.



(a) Reweighted $\Delta\eta(J/\psi_1, J/\psi_2)$.



(b) Reweighted $\Delta\phi(J/\psi_1, J/\psi_2)$.



(c) Reweighted $\Delta R(J/\psi_1, J/\psi_2)$. Distribution accounts for phase space.

Figure 32: Angular distributions for the cut di-J/ψ sample after reweighting to the probability of the events containing two non-prompt J/ψs. All distributions are normalised to one.The systematic uncertainty on the distributions according to the inclusive fit are shown as the dotted lines. It can be seen that the uncertainties are largest in the in the first bin of each distribution, but overall are not large. This indicates that the default fit is robust and models the data well.

7.3.4 Discussion

It can be seen in Figures 29 and 30 that the functions used in the fit sum together to match the data distributions well.

Figure 31 shows that the unweighted distribution is highly peaked towards zero in $\Delta\eta$ and ΔR, with $\Delta\phi$ displaying a peak towards zero and a slightly lower one at π. Weighting the data according to the probability of the events containing two non-prompt J/ψs causes a reduction of the zero peaks in the ΔR and $\Delta\phi$ distributions and an enhancement of their π peaks. The reweighted $\Delta\eta$ distribution again features a reduced zero-peak but the enhancement here is in the range: $0.5 \leq \Delta\eta \leq 3$. The reweighting suggests that the production of non-prompt J/ψs can be expected to take place primarily at low angles, but not as low as either the production of prompt J/ψs or whatever background was in the sample. There is also an increased fraction where the J/ψs are produced in a back-to-back scenario.

Figure 32 indicates that there are multiple ways of constructing the fit which are very similar to each other. This suggests that the method of fitting and extracting probabilities is stable and doesn't require that the form of the fit is exactly correct.

7.3.5 Conclusion

In summary:

- Multiple fits can be constructed which match the data distribution well.
- The sample data can be reweighted according to the probability of its events containing two non-prompt J/ψs
- Production of non-prompt J/ψ pairs takes place primarily at low angles, but a small proportion are produced back-to-back.

Comparing the angular distributions of the reweighted data to those of the MC samples in Section 4 suggests that PS production of b quarks contributes the most to non-prompt J/ψs production, but that some ME production is also necessary.

The inclusive fit makes the assumption that the kinematics of the data sample are independent of the geometric variables. In order to account for any changes in kinematics, the UCL researchers moved to a binned fit, where the data is divided into multiple bins according to each variable and fitting takes place to the data in each bin.

7.4 Binned fit

The binned fit is, again, a four dimensional simultaneous fit to the masses and pseudoproper lifetimes of both J/ψs in the events, however, unlike the inclusive fit where all the data were used in a single fit, the data here are binned according to each variable ΔR, $\Delta\phi$, and $\Delta\eta$ and the fit performed once to the data in each bin.

7.4.1 Additional cuts

The requirements for event acceptance for the binned-fit sample were more stringent than those for the inclusive, requiring events to pass certain triggers and $p_\perp$ cuts on the J/ψs and muons.

Mass and pseudoproper-lifetime cuts Data for the fit was required to pass the following cuts on both J/ψs in the event:

- $2.6 \leq m_{J/\psi} \leq 3.5$ [GeV]

- $-2 \leq \tau_{J/\psi} \leq 4$ [ps]

The values for these cuts were chosen by the UCL researchers and applying them reduced the sample from 228 071 events to 63 426 events.

***J/ψ* $p_\perp$ cut** The results of the binned fit will be compared to Monte Carlo data, which will use a b- hadron $p_\perp$ cut of 10 GeV. In order to make sure that the two data samples occupied the same phase-space region, a $p_\perp$ cut of 6.339 678 8 GeV was applied to both J/ψ candidates. This value was calculated using:

$$p_{\perp J/\psi} = 0.625452 p_{\perp B} + 0.0851588,$$

the $B \to J/\psi$ $p_\perp$ correlation function derived in Section 6. This cut was chosen by the author, and applying it reduced the sample from to 29 572 events.

Trigger requirements Events in the data sample were required to pass at least one of the following high-level event-filter triggers:

- `EF_2mu4T_Jpsimumu` (For event numbers below 206 956)
- `EF_2mu4T_Jpsimumu_L2StarB` (For event numbers 206 956 and above)
- `EF_3mu4T`

The choice of triggers was made by the UCL researchers and requiring them reduced the sample to 13 932 events.

Muon requirements All four muons in the events were required to pass the following cuts:

- $|\eta_\mu| \leq 2.5$
- $p_{\perp\mu} > 2.5$ GeV

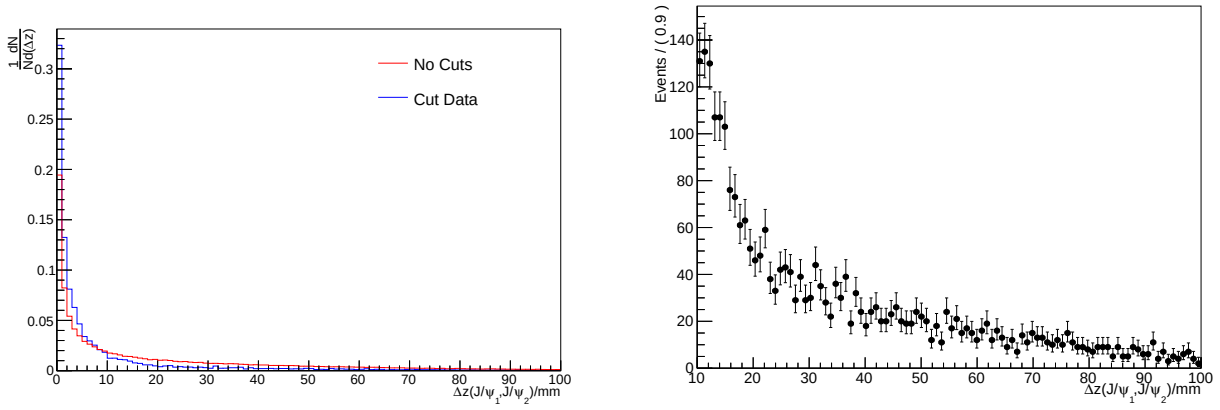
Additionally, at least one J/ψ per event was required to have both its muons with $p_{\perp\mu} > 4$ GeV, have both its muons be combined, and pass either `EF_2mu4T_Jpsimumu` (for event numbers below 206956) or `EF_2mu4T_Jpsimumu_L2StarB` (for event numbers 206956 and above).

These cuts were again chosen by the UCL researchers and applying them reduced the sample to 10 071 events.

Pile-up removal Figure 33 shows the Δz distribution for the J/ψ pairs before and after the cuts described above.

Figure 33b shows that the contribution from pile-up begins to dominate around $\Delta z = 20$ mm. Figure 33a shows that again the cuts heavily reduce the fraction of pile-up events in the sample. A cut requiring $\Delta z(J/\psi_1, J/\psi_2) \leq 20$ mm will be used. By fitting a linear function to Δz in the region $\Delta z \geq 20$ mm, extrapolating it back to $\Delta z = 0$ mm, and integrating over the range 0 mm $\leq \Delta z \leq 20$ mm, it was estimated that 5 % of remaining events passing the Δz cut were due to pile-up.

The investigation of pile-up contamination in the sample was performed by the author, who also chose the Δz cut value. Applying it reduced the sample to 8835 events.



(a) *(b)*

Figure 33: Figure (a) shows $\Delta z(J/\psi_1, J/\psi_2)$ before and after the cuts, where the distributions are normalised to one. Figure (b) shows $\Delta z(J/\psi_1, J/\psi_2)$ distribution for $\Delta z > 10\,mm$ after the cuts.

7.4.2 The fit

The fit, and associated fitting and plotting code used in this sub-subsection was written and developed by the UCL researchers. The fit again consists of models similar to those used in [44], however the fitting functions were revised for the binned fit with the background being split into three components: prompt, prompt DS, and non-prompt SS, where SS and DS refer to the decay function in the pseudoproper-lifetime fit (SS - Single-Sided, DS - Double-Sided). Each background component has its own mass fit.

Mass fit The mass distribution again consists of contributions from prompt and non-prompt J/ψs, modelled by a Gaussian and Crystal Ball function, and the three background components, each modelled by a Chebychev polynomial.

Pseudoproper-lifetime fit Again, a double Gaussian resolution model is used for the pseudoproper lifetime fit and all the component functions are convolved with this. The non-prompt J/ψs and the non-prompt SS background are both modelled with single-sided decay functions, the prompt J/ψs and prompt background are both delta functions, and the prompt DS background is a double-sided decay function.

Data weighting In order to correct for selection inefficiencies, the sample was weighted according to the muon reconstruction efficiencies of the event, and according to the trigger weight. Because there is currently no weight map for the `EF_2mu4T_Jpsimumu` trigger, the `EF_2mu4T_Jpsimumu_L2StarB` weight map was used for all events.

Binning The distribution of each variable (ΔR, $\Delta\phi$, and $\Delta\eta$) was divided up into bins with the aim of having a higher bin density around higher gradient regions. This was, however, not always possible due to the much lower statistics available for fitting. The final values for bin edges used here were decided on by the author, but were adapted from previous values determined by the UCL researchers.

Fitting The five possible states for each particle in an event are:

- Non-prompt J/ψ
- Prompt J/ψ
- Prompt background
- Prompt DS background
- Non-prompt SS background

There are then twenty five possible states for an event. As with the inclusive fit, the sum of the PDFs of these states is fitted to the data using RooFit.

Since fitting in each bin of a variable offers much lower statistics than fitting inclusively, RooFit had to be coaxed into convergence. This was done via a three stage fitting process: First, two two-dimensional fits were performed inclusively to the mass and pseudoproper lifetime of each particle. The result of this initial fit is used to set the shapes of the J/ψ mass peak, the background mass distributions, and the pseudoproper-lifetime resolution model. The two-dimensional fits are then performed in each bin of the variable and the results of these are used to set the position of the J/ψ mass peak and the shapes of the signal and background pseudoproper-lifetime distributions. At this point the only remaining floating parameters are the normalisations of the shapes. These are set by a four dimensional simultaneous fit in each variable bin.

7.4.3 Data reweighting

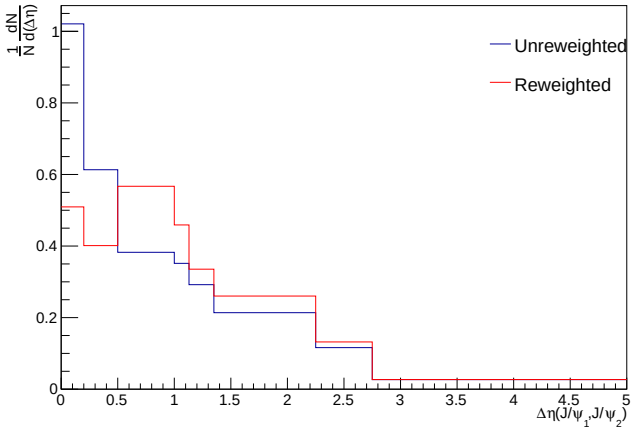
The code for the probability extraction and determination of uncertainties used here was written by the author, but was based on methods and suggestions supplied by the UCL researchers.

Probability extraction Determination of the probability of a given event containing two non-prompt J/ψs again used:

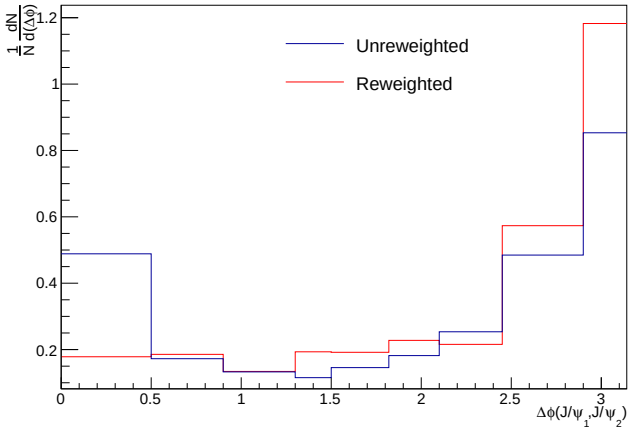
$$\begin{aligned} P_{\mathrm{NP}J/\psi,\mathrm{NP}J/\psi}(m_1, m_2, \tau_1, \tau_2) = &\frac{\text{Mass Signal}_1(m_1)}{\text{Mass Signal}_1(m_1) + \text{Mass Bkg.}_1(m_1)} \\ &\times \frac{\text{Mass Signal}_2(m_2)}{\text{Mass Signal}_2(m_2) + \text{Mass Bkg.}_2(m_2)} \\ &\times \frac{\tau\ \text{Signal}_1(\tau_1)}{\tau\ \text{Signal}_1(\tau_1) + \tau\ \text{Bkg.}_1(\tau_1)} \\ &\times \frac{\tau\ \text{Signal}_2(\tau_2)}{\tau\ \text{Signal}_2(\tau_2) + \tau\ \text{Bkg.}_2(\tau_2)}, \end{aligned} \tag{7.2}$$

where $\text{Mass Signal}_\alpha$ and $\tau\ \text{Signal}_\alpha$ are the PDFs for the signal contributions to the mass and pseudoproper-lifetime distributions, respectively, for candidate α. Mass Bkg._α and $\tau\ \text{Bkg.}_\alpha$ are the corresponding background PDFs.

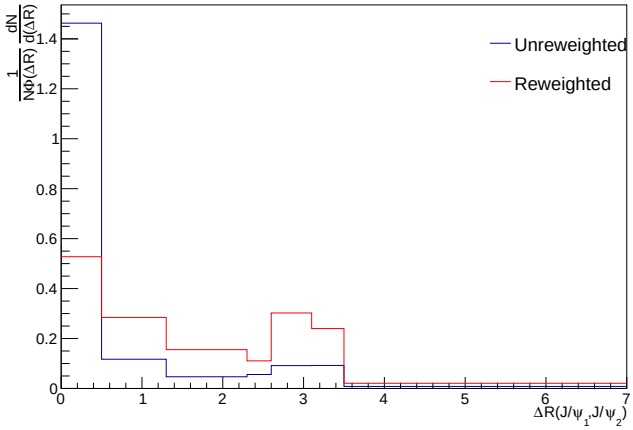
Reweighting results Multiplying the weight of each event by the probability of it containing two non-prompt J/ψs, normalising the distributions, and comparing them to the un-reweighted distributions gave rise to Figure 34.



(a) $\Delta\eta(J/\psi_1, J/\psi_2)$ *before and after reweighting.*



(b) $\Delta\phi(J/\psi_1, J/\psi_2)$ *before and after reweighting.*

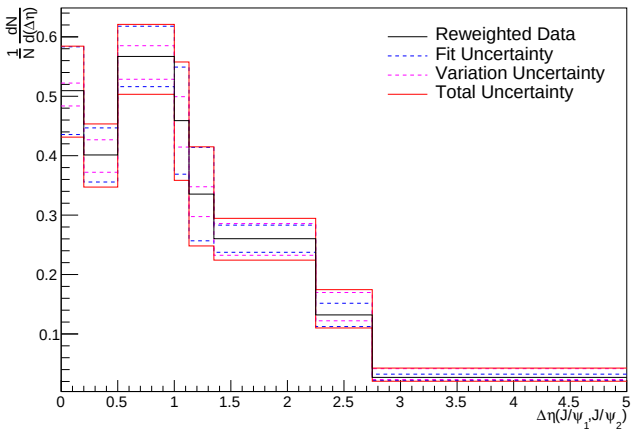


(c) $\Delta R(J/\psi_1, J/\psi_2)$ *before and after reweighting. Distributions account for phase space.*

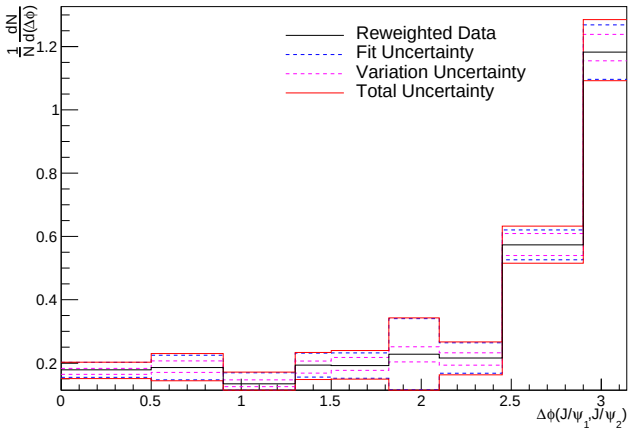
Figure 34: Angular distributions for the cut di-J/ψ *sample before and after weighting with the probability of the events containing two non-prompt* J/ψ*s. All distributions are normalised to one. Reweighting the sample causes a dramatic shift in the distributions to higher separations. This is most noticeable in* ΔR *and* $\Delta\phi$*, where the peaks close to zero are decreased in favour of the peaks around* π*. Reduction of the peak close to zero in the* $\Delta\phi$ *distribution is to the point that the distribution becomes flat upto around 1.3. The reweighted results indicate that non-prompt* J/ψ*s have a higher likelihood to be produced in a back-to-back scenario than prompt* J/ψ*s or the background in the sample.*

Systematic uncertainties To account for the large reduction in the number of degrees of freedom the final four-dimensional fit has, it was repeatedly rerun in each bin, each time allowing one of the previously fixed parameters to be floating. The data distributions were then reweighted according to each of these variant models and the deviations from the original fit summed in quadrature to build upper and lower error bounds on the reweighted fit. RooFit's uncertainty on the non-prompt J/ψ PDF fraction in each bin was also added in quadrature to the bounds.

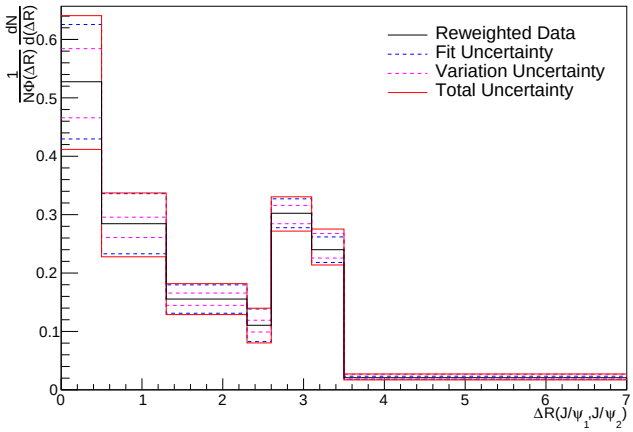
Figure 35 shows the reweighted distributions along with their uncertainty bounds. The fit-uncertainty contributions are the uncertainties on the final four-dimensional fit, and the variation-uncertainty contributions are the quadrature-summed deviations of the variant models from the original fit.



(a) Reweighted $\Delta\eta(J/\psi_1, J/\psi_2)$.



(b) Reweighted $\Delta\phi(J/\psi_1, J/\psi_2)$.



(c) Reweighted $\Delta R(J/\psi_1, J/\psi_2)$. Distribution accounts for phase space.

Figure 35: Angular distributions for the cut di-J/ψ sample after reweighting to the probability of the events containing two non-prompt J/ψs. All distributions are normalised to one. The uncertainty on the distributions according to the binned fit are shown as the red dotted lines, with the sources which contribute to the total uncertainty shown in blue and purple. It can be seen that the uncertainties can be quite large, reaching about 50 % in the 6^{th} bin of the $\Delta\phi$ distribution.

7.4.4 Discussion

Figure 34 shows that, as with the inclusive-fit-reweighted distributions in Figure 31, the reweighting causes a reduction of the peaks close to zero in the ΔR and $\Delta\phi$ distributions and an enhancement of their peaks around π . The reweighted $\Delta\eta$ distribution again features a reduced peak near zero and the enhancement here is in the same range as it was in the inclusive fit: $0.5 \leq \Delta\eta \leq 2.75$. In contrast to the inclusive-fit-reweighted distributions, the suppression of low-angle production here is much more extreme, to the point that $\Delta\phi$ is flat upto around 1.3, and the first two $\Delta\eta$ bins are moved below the value of the third bin. The reweighting again suggests that the production of non-prompt J/ψs can be expected to take place primarily at low ΔR, but not as low as either the production of prompt J/ψs or whatever background was in the sample. There is also an increased fraction where the J/ψs are produced in a back-to-back scenario.

Figure 35 indicates that binned-fit reweighting can carry some large uncertainties (up to 50 % in the sixth bin of Figure 35b), but are generally at acceptable levels. It can be seen that in most cases, the greatest contribution to these uncertainties is the uncertainty on the final four-dimensional fit, due to the extremely small data sample available for fitting, however the variation uncertainty also contributes a non-negligible amount.

7.4.5 Conclusion

The uncertainties on the reweighted variable distributions make it unusable for precision measurements, it should, however, be applicable for shape comparison to Monte Carlo-generated distributions which will give an indication of the agreement between the particle physics models and collider data.

As was discussed in Section 7.4.4, the size of uncertainties are primarily due to the low statistics available in the sample. The requirements for the sample were quite high as the decay chain $b \to B \to J/\psi \to \mu^+\mu^-$ has a very low branching ratio, as was discussed in Section 6.1. By relaxing the requirements on the sample to require a signal of the form $J/\psi + X$, where X is some other particle, the sample statistics could be increased at the cost of lower signal discrimination from the more complex backgrounds which would accompany relaxation. Indeed, it is in this direction that the UCL researchers have proceeded.

8 Di-J/ψ data comparison to Monte Carlo

8.1 Outline

Having prepared variable distributions corresponding to non-prompt J/ψs in Section 7, and investigated Monte Carlo production of $b\bar{b}$ pairs in Section 4 it only remains to compare MC-generated samples to the di-J/ψ data. However, since the analysis of MC samples will be using the properties of b hadrons and the LHC data uses J/ψs, the $B \to J/\psi$ correlation functions derived in Section 6 will be used to move the samples into the same phase-space region and then smear the MC result distributions to simulate the decay of the b hadrons to J/ψs.

8.2 Monte Carlo samples

To allow for accurate comparison with the di-J/ψ data, the MC samples were designed to be as inclusive as possible for heavy-flavour production; the processes included are a combination of the ME and PS processes from Section 4, as well as the flavour excitation and other processes which were excluded from the sample runcards in Section 4. MPI simulation was included in all samples.

Sample Name	σ [nb]	Size [events]
SHERPA+{0}	$6\,002\,000 \pm 2000$	1×10^7
SHERPA+{1}	$6\,011\,000 \pm 2000$	1×10^7
SHERPA+{2}	$6\,011\,000 \pm 1000$	1×10^7
PYTHIA 8	$5\,075\,000 \pm 1000$	1×10^7

Table 12: Production cross sections and sample sizes for MC comparisson data.

Sample Name	Fractional Acceptance
SHERPA+{0}	$0.000\,859 \pm 0.000\,008$
SHERPA+{1}	$0.000\,89 \pm 0.000\,01$
SHERPA+{2}	$0.000\,893 \pm 0.000\,006$
PYTHIA 8	$0.002\,03 \pm 0.000\,01$

Table 13: Fractional acceptance of events by analysis of MC comparison data.

The processes included are: $p\ p \to b\,\bar{b}$, $b\,\bar{b} \to b\,\bar{b}$, $b\ b \to b\ b$, $\bar{b}\ \bar{b} \to \bar{b}\ \bar{b}$, $p\ p \to p\ p$, $b\,\bar{b} \to p\ p$, $p\ b \to p\ b$, and $p\ \bar{b} \to p\ \bar{b}$. SHERPA was also used to produce samples with possible extra outgoing light-partons. These are indicated by $\{N_e\}$ in the sample names, where N_e is the maximum number of extra light-partons. Again, `CKKW` was set to $\sqrt{20/E_{\mathrm{CM}}}$.

The same general settings were used as for the production in Section 4, though production was performed using the *MSTW2008lo68cl* PDF. When using SHERPA, c and b quarks were set to be massive such that they were not included in its `93` light-parton container and the number of EW interactions in the ME was set to zero to ensure the propagators were partons. $\hat{p}_{\perp,min}$ for the ME was set to 10 GeV and the colliding beams were set to 4 TeV protons. MPI simulation was performed by the generators' in-built handlers, with AMISIC being selected for SHERPA. In both cases the MPI was set to produce $2 \to 2$ QCD interactions, and allowed to produce b quarks.

Details of the generation of the samples are given in Table 12

8.3 Monte Carlo analysis

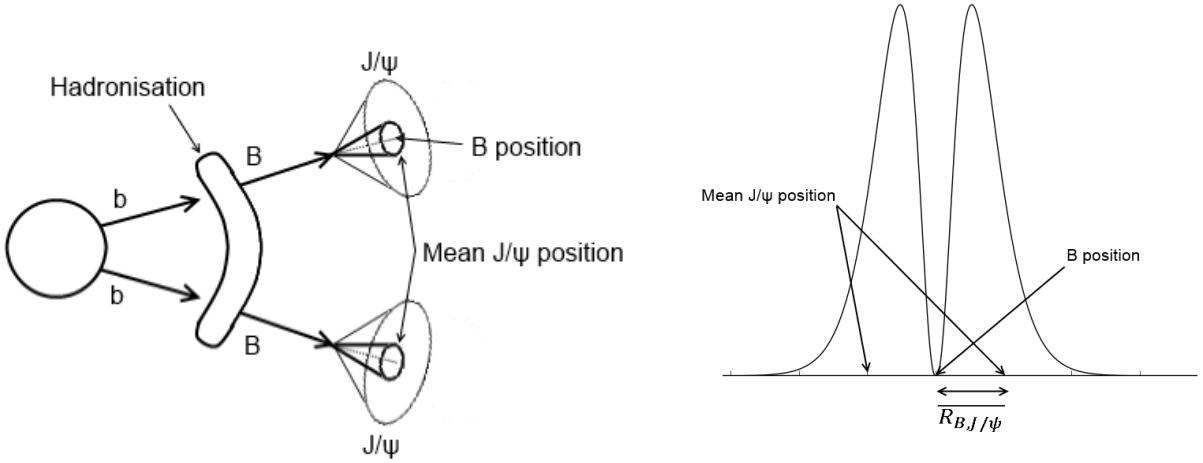
As was discussed in Section 6, the branching ratio of $B \to J/\psi$ is very small, and requiring a pair of $B \to J/\psi$ processes in the analysis would severely reduce acceptance. Instead, the analysis here will simply require a pair of b hadrons and the correlation functions derived in Section 6 will be used to smear the b hadron distributions such that they are comparable to the di-J/ψ distributions.

The RIVET analysis used here differs slightly to that used in Section 4: two b hadrons are required to be found in the range $|\eta| \leq 2.5$, with $p_\perp \geq 10$ GeV, but rather than plotting directly with the analysis, the properties of the hardest two b hadrons in each event are written out to be read in by the smearing program later.

The fractional acceptance for each sample is detailed in Table 13.

8.4 Monte Carlo data smearing

Section 6 showed that the spread of the mean deviation ($\Delta R_{B,J/\psi}$) could be modelled by gamma distributions, and assumed that the shape of $\Delta R_{B,J/\psi}$ was symmetric about the direction of the



(a) *(b)*

Figure 36: Figure (a) illustrates the geometry of the di-$B \to$ di-J/ψ decay. The centre cones indicate the mean position of J/ψs relative to their parent b hadron. Figure (b) shows the $B \to J/\psi$ smearing in one dimension. The smearing consists of a double-gamma distribution which forms the PDF for a J/ψ's position relative to its parent b hadron. The final smearing applied to $\Delta R_{B,B}$ will be the sum of two such smearings; one for each $B \to J/\psi$ decay.

b hadron. Since the smearing will be applied to $\Delta R_{B_1,B_2}$, it should consist of the sum of two double-gamma distributions with means equal to:

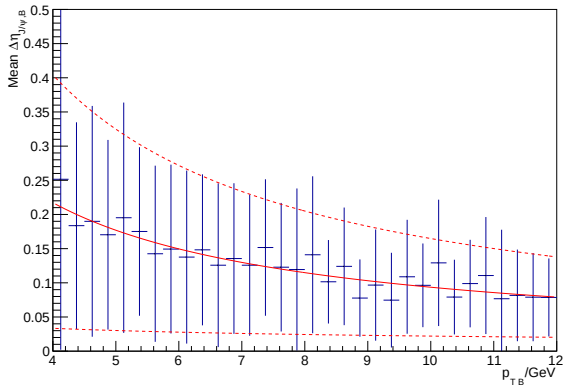
$$\Delta R_{B_1,B_2} \pm \left\langle \Delta R_{B_i,J/\psi_i} \right\rangle,$$

where $i = 1, 2$. Figure 36 illustrates the smear on the $B \to J/\psi$ decays.

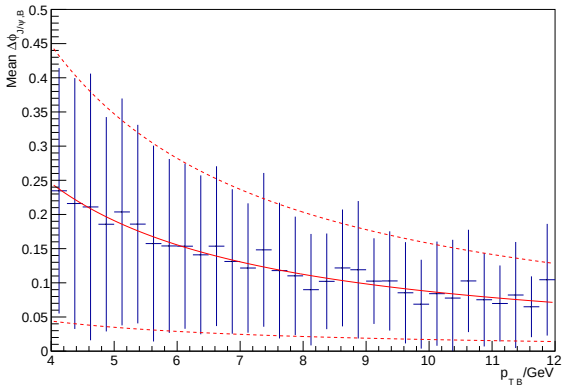
The construction and application of a di-B$\to$di-J/ψ smearing distribution is as follows: The parameters of the double-gamma distributions are derived by sampling the correlation functions derived in Section 6, shown again in Figure 37 and Table 14. For each b hadron, the mean function is sampled to provide $\left\langle \Delta R_{B_i,J/\psi_i} \right\rangle$ and the mean$\pm$standard deviation functions are sampled to provide the standard deviation of its double-gamma distribution (the largest standard deviation is selected). An offset is used to move each double-gamma distribution such that they are centred around $\Delta R_{B_1,B_2}$. Each double-gamma distribution is normalised to one, so that total area of the smeared distribution is two; there are two J/ψs and both must exist somewhere. Similar smearings are constructed for the $\Delta\phi$ and $\Delta\eta$ distributions.

$\left\langle \Delta R_{J/\psi,B} \right\rangle = (1.45^{2.38}_{0.54})\, p_{\perp B}^{\left(-1.00^{-1.03}_{-0.90}\right)}$
$\left\langle \Delta \phi_{J/\psi,B} \right\rangle = (1.17^{2.17}_{0.19})\, p_{\perp B}^{\left(-1.13^{-1.14}_{-1.04}\right)}$
$\left\langle \Delta \eta_{J/\psi,B} \right\rangle = (0.78^{1.57}_{0.06})\, p_{\perp B}^{\left(-0.92^{-0.98}_{-0.45}\right)}$

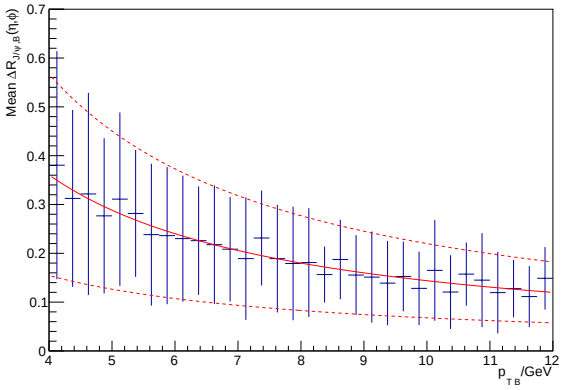
Table 14: Equations for the $B \to J/\psi$ angular correlation functions. Parameters are shown in the form $\left(a^b_c\right)$, where a, b, and c are for the fits to the mean values, mean+standard deviation values, and mean-standard deviation values, respectively. Parameters are shown to two decimal places, but are used at six significant figures.



(a) Mean $\Delta\eta_{J/\psi,B}$ as a function of $p_{\perp B}$.



(b) Mean $\Delta\phi_{J/\psi,B}$ as a function of $p_{\perp B}$.



(c) Mean $\Delta R_{J/\psi,B}$ as a function of $p_{\perp B}$.

Figure 37: Angular correlations between J/ψs and their parent b-hadrons. The solid lines are the fit to the mean values and the dotted lines are the fit to the mean±standard deviation values. The y-axis error bars indicate the standard deviation of the gamma functions fitted to the distribution of points within each b hadron $p_\perp$ bin.

Next, each bin of the histogram for each variable distribution is filled according to the area of the smear distribution which falls in each bin times the weight of the event:

$$A(x)_{i,i+1} = W_j \int_{x_i}^{x_{i+1}} S(x)\ \mathrm{d}x, \tag{8.1}$$

where $A(x)$ is the fill area for the bin with edges at x_i and x_{i+1}, W_j is the weight of event j, and $S(x)$ is the smearing function for angular variable x. $A(x)$ is calculated by finding the difference in the smearing distribution's cumulative distribution function values at x_i and x_{i+1}.

When populating the $\Delta\phi$ distributions, the smear distributions are reflected at $\Delta\phi = 0|\pi$ to account for the bounded nature of the variable. For the ΔR and $\Delta\eta$ distributions, reflection takes place at $\Delta x = 0$.

In order to account for phase space when filling the ΔR distribution, the mean of the smear-weighted phase space is calculated within each bin and used to divide the event-weighted area

for each bin fill:

$$\overline{\Phi(\Delta R)}_{i,i+1} = \frac{\int_{\Delta R_i}^{\Delta R_{i+1}} S(\Delta R)\,\Phi(\Delta R)\ \mathrm{d}(\Delta R)}{\int_{\Delta R_i}^{\Delta R_{i+1}} S(\Delta R)\ \mathrm{d}(\Delta R)}, \tag{8.2}$$

$$A(\Delta R)_{i,i+1} = \frac{W_j}{\overline{\Phi(\Delta R)}_{i,i+1}} \int_{\Delta R_i}^{\Delta R_{i+1}} S(\Delta R)\ \mathrm{d}(\Delta R)\,, \tag{8.3}$$

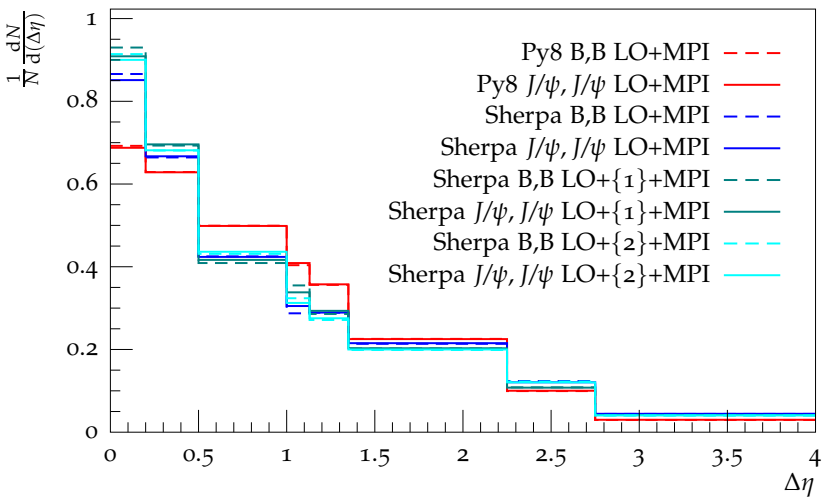
where $\overline{\Phi(\Delta R)}_{i,i+1}$ is the smear-weighted phase space between ΔR_i and ΔR_{i+1}.

The method used for calculating $\overline{\Phi(\Delta R)}_{i,i+1}$ involves numerical integration (Composite Simpson's Rule) and Monte Carlo integration (Vegas algorithm). In order to prevent the runtime for the smearing code becoming prohibitively long, filling of a ΔR bin only takes place if the smear area within that bin is greater than 1×10^{-10}. The calculation of the mean phase space takes place after the smear had been reflected at $\Delta R = 0$.

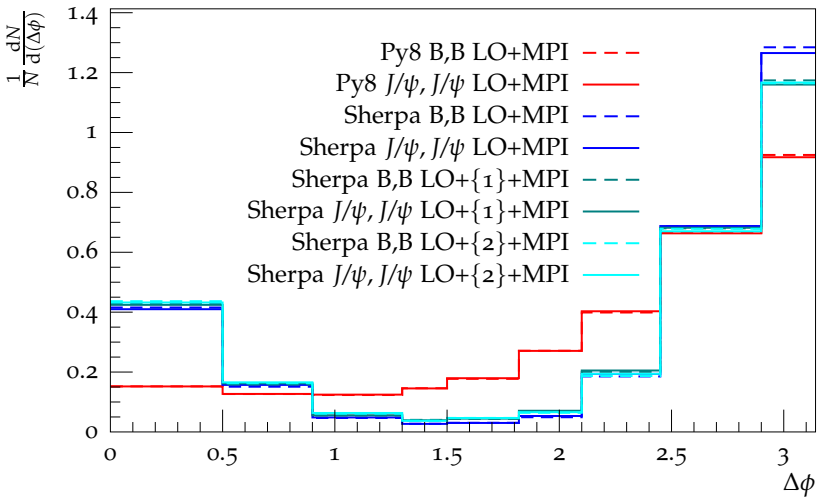
The binning for the histograms was the same as that used in the di-J/ψ data distributions. The effect of applying the smearing can be seen in Figure 38.

8.5 Results

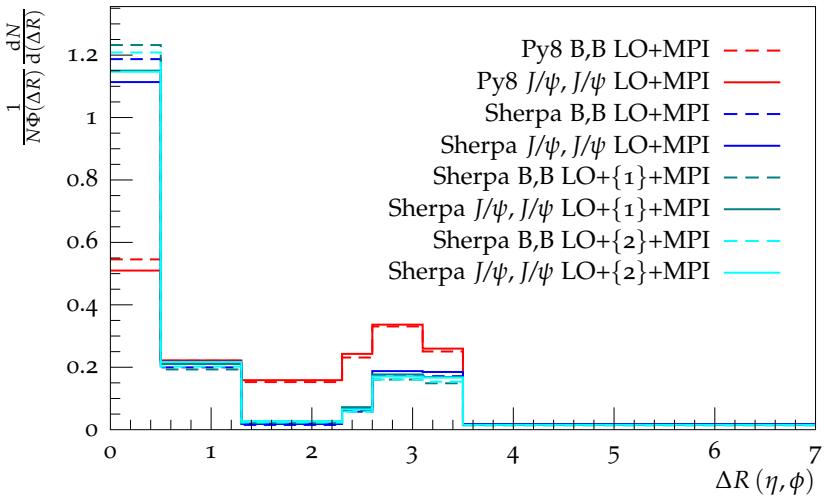
Figure 39 shows the MC distributions, after the di-B $\rightarrow$ di-J/ψ smearing is applied, compared to the binned-fit-reweighted di-J/ψ data. The uncertainty bars on the MC distributions are purely statistical and do not take into account any systematic uncertainty due to the smearing. The uncertainty on the di-J/ψ distributions are those calculated in Section 7.4.3.



(a) $\Delta\eta$.

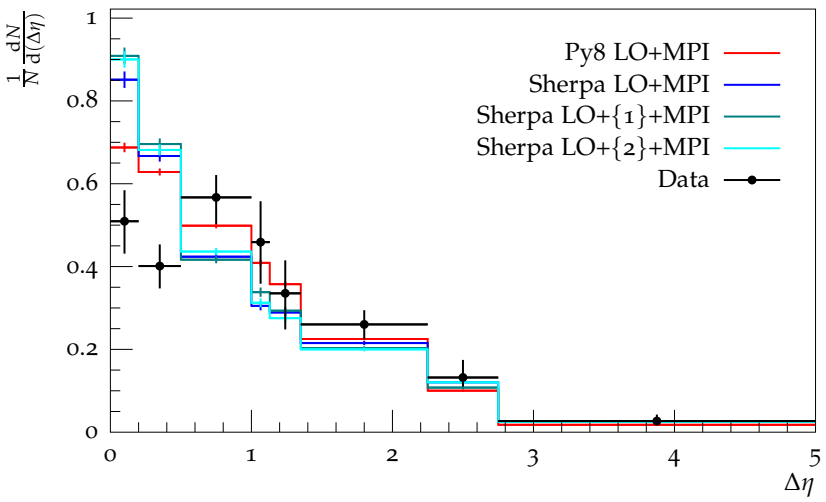


(b) $\Delta\phi$.

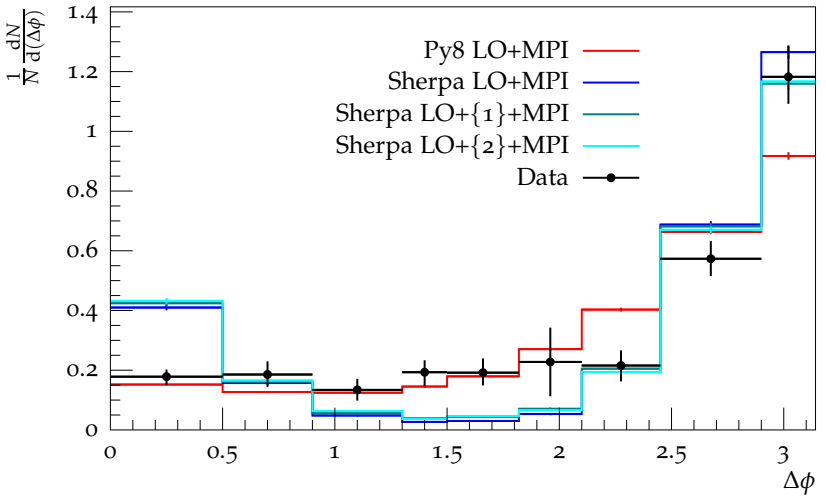


(c) ΔR. *Plotting accounts for phase space.*

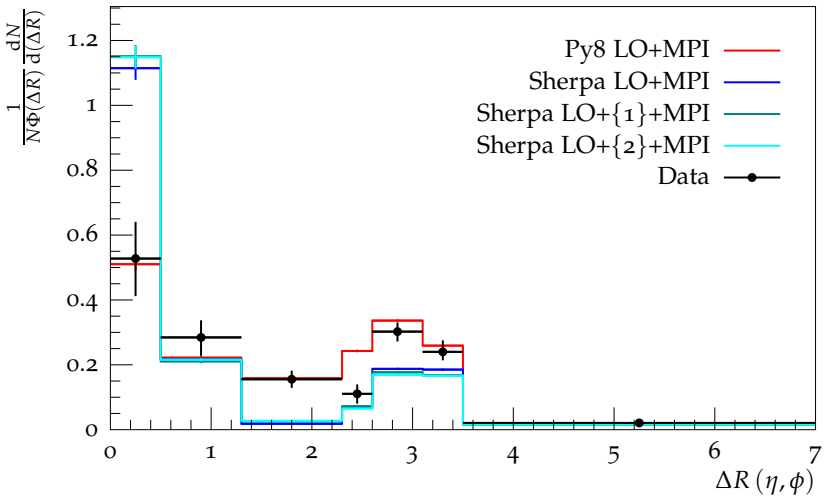
Figure 38: MC data before (B,B) and after $(J/\psi, J/\psi)$ $B \to J/\psi$ *smearing is applied. All distributions are normalised to one.*



(a) $\Delta\eta(J/\psi, J/\psi)$.



(b) $\Delta\phi(J/\psi, J/\psi)$.



(c) $\Delta R(J/\psi, J/\psi)$. *Plotting accounts for phase space.*

Figure 39: Smeared MC data compared to binned-fit-reweighted di-J/ψ data. All distributions are normalised to one.

8.6 Discussion

The generators show a large difference in the production cross section for their samples. Table 12 shows that $\sigma_{\text{SHERPA+\{0\}}}$ is greater than σ_{PYTHIA} by $(18.27 \pm 0.05)\,\%$. This is different to the case in Section 4 where both generators produced at similar cross sections. It is possible that this was caused by the change in PDF, but it could also be due to the difference in mass treatment of HF in backwards evolution from the ME. PYTHIA 8 continues to have a larger acceptance fraction than SHERPA, by $(136 \pm 2)\,\%$.

Figure 38 shows that the application of the $B \to J/\psi$ smearings does not have a large effect on the shape of MC data distribution. This is due to the bin widths being relatively large compared to the mean deviations used in the construction of the smears.

Figure 39 shows that PYTHIA 8 demonstrates better agreement with the data than SHERPA.

In the low-angle region ($0 \leq \Delta x \leq 1.2$ $x \in \{R, \phi, \eta\}$) SHERPA is seen to consistently be too peaked towards zero in all three angles, whereas PYTHIA 8 follows the data much more closely (particularly visible for $\Delta\phi$ in Figure 39b), with the exception of the first two $\Delta\eta$ bins (Figure 39a). Section 4 showed that production in the low-angle region is almost entirely through the PS, so the differences here are most likely due to scale choice and varying treatment of HF quarks.

In the medium-angle region ($1.2 \leq \Delta x \leq 2.1$ $x \in \{R, \phi\}$) SHERPA severely underestimates the data as both PS and ME production in this region is negligible, however PYTHIA 8's PS distribution extends much farther out, and its ME distribution begins at a lower angle than SHERPA's allowing it to be consistent with the data.

At high-angle ($2.1 \leq \Delta x$ $x \in \{R, \phi\}$), where ME production dominates, the results are varied: The ΔR distributions in Figure 39c show better agreement between PYTHIA 8 and the data, with SHERPA underestimating production close to π. SHERPA's underestimation is perhaps due to normalisation, however its sixth bin shows levels approximately equal to those of its fifth bin, unlike the data and PYTHIA 8's distributions where the sixth bin does show a decrease, so SHERPA would still show a slight difference in shape to the data even if its normalisation were adjusted. The $\Delta\phi$ distributions, however, show SHERPA to have better agreement with the data than PYTHIA 8, from $\Delta\phi = 2.1$ onwards. Figure 10a showed that this point was where ME production in SHERPA began and that the slope of SHERPA's distribution was much steeper that PYTHIA 8's. This difference in steepness is reflected in Figure 39b, where PYTHIA 8 overestimates bins seven and eight, and underestimates bin nine. SHERPA, however, only overestimates bin eight. Since PYTHIA 8 and SHERPA LO+MPI use the same ME, differences between the generator distributions at high-angle are likely due to scale choice, event-kinematics calculation, or MPI simulation. As was seen in Section 4, PYTHIA 8's PS causes a non-negligible amount of HF production in the high-$\Delta\phi$ region, so this could also be contributing to the difference in shape.

8.7 Conclusion

Of the two generators, PYTHIA 8 was seen to offer the best agreement with the data. The main source of SHERPA's disagreement with the data appeared to be due to its PS modelling.

The peaked nature of SHERPA's low angle $\Delta\phi$ distribution appears to be inconsistent with the collider data and PYTHIA 8's modelling, both of which show an approximately flat distribution at low angle. It is possible that this inconsistency is due to inadequate modelling by SHERPA's PS and it could be the subject for further investigation to alter its settings and scale choice to see whether better agreement might be reached. A new PS [50] was recently implemented for SHERPA which may also offer improvement in SHERPA's agreement with the

data, however as it is not yet[15] enabled as the default shower, it has not been tested in this investigation.

It would be useful to build a smearing function for the SHERPA results based on $B \to J/\psi$ correlation functions derived from SHERPA data, however, as mentioned in Section 6.5, forced decays of b hadrons is currently not available in SHERPA.

In order to quantify agreement between the MC results and the data it would be necessary calculate the systematic uncertainty on the MC results which arises due to the application of the smearing function. Calculation of this would require using other generators to build up the $B \to J/\psi$ correlation functions to see how much they varied.

9 Conclusion

Section 4 showed that two popular, state-of-the-art Monte Carlo event generators SHERPA and PYTHIA 8 showed very large differences in their simulation of $b\bar{b}$ pair production. The differences were found to be particularly prominent in production through the parton shower. Section 4 also identified regions where production through parton shower or matrix element can be expected to dominate.

The accuracy of SHERPA's parton shower approximation of gluon splitting was investigated in Section 5 where it was seen that it does a reasonable job compared to the explicitly calculated matrix element, but still has room for improvement.

Since the Monte Carlo generators had shown very different predictions it was necessary to compare them to data. Section 6 showed that J/ψs were a viable proxy to the b hadrons used in the previous sections, and the ATLAS detector, described in Section 3.6, with its powerful machinery available for the analysis of high-energy particle collisions, was an ideal instrument to collect the required data.

In Section 7 a sample of non-prompt di-J/ψs was prepared by reweighting a di-J/ψ sample[16] of $\sqrt{s} = 8\,\text{TeV}$ LHC data according to probabilities extracted from a binned fit[17] to the data. It was found that the reweighted data carried uncertainties large enough to make it unsuitable for precision measurements but still usable for qualitative comparison to Monte Carlo results. It was suggested that the uncertainties could be reduced by relaxing the data selection to $J/\psi+X$.

The comparison was constructed in Section 8 where inclusive samples of $b\bar{b}$ production were generated in SHERPA and PYTHIA 8, and the angular distributions of b hadrons were extracted. In order to account for the missing $B \to J/\psi$ decays in the Monte Carlo samples, a smearing function was constructed based on correlation functions derived in Section 6 and applied to the Monte Carlo distributions. It was found that PYTHIA 8 showed the best agreement with the data, and that SHERPA's disagreement was primarily in the low and medium angles. Since this is where production is primarily through the parton shower, it was suggested that further investigation of SHERPA's scale choices and splitting functions could be a way in which data agreement might be improved. It was also noted that in order to quantitatively determine agreement it would be necessary to determine the systematic uncertainty on the smearing function, which would most likely require constructing it using multiple generators to measure how much it varies.

Only two of a plethora of Monte Carlo generators were employed in this thesis, however the investigation could easily be extended to include results from other generators, such as HERWIG++ [51] or MADGRAPH [52]. Additionally, dependence on parton density function

[15] Latest version at time of writing is 2.2.0

[16] Original data sample prepared by Dr Kostas Karakostas of the National Technical University of Athens

[17] The fits used were written and developed by Dr Gavin Hesketh and Dr Josh McFayden of University College London

could be tested by rerunning the investigations with other PDFs such as ones by the CTEQ or NNPDF groups.

Acknowledgements

I would like to thank my supervisors Dr Aidan Robson and Dr Andy Buckley for their great assistance and valued support during this investigation. I am grateful to Dr Gavin Hesketh and Dr Josh McFayden for their work in constructing the fitting code, and for all their suggestions and help. Lastly, I would like to thank my family and friends for their encouragement and support throughout the year.

Software details

Software	Version	References	Use/Notes
Sherpa	2.1.1	[7]	MC generation
Catani-Seymour shower	-	[8]	Included in Sherpa
Amegic++	-	[9]	Included in Sherpa
Comix	-	[10]	Included in Sherpa
Pythia 8	8.186	[6]	MC generation
		[5]	
Rivet	2.1.2	[40]	MC analysis
Yoda	1.1.0 & 1.3.0	[53]	Plot creation
FastJet	3.0.6	[54]	Jet clustering
		[55]	
Root	5.34/11	[56]	Di-J/ψ analysis and plot creation
RooFit	3.58	[47]	Data fitting
Athena	19.1.1.1	-	Running Rivet over Pythia 8B data
Lhapdf	6.1.3	[57]	PDF data for MC generation
MCViz	2013-04-25	[58]	Creation of Figure 8

References

[1] **ATLAS**, G. Aad *et al.*, *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] **CMS**, 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]`.

[3] **ATLAS Collaboration**, G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, J. Instrum. **3** (2008) S08003. 437. Also published by CERN Geneva in 2010.

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

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

[6] T. Sjöstrand, 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].

[7] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, *et al.*, *Event generation with SHERPA 1.1*, JHEP **0902** (2009) 007, arXiv:0811.4622 [hep-ph].

[8] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **0803** (2008) 038, arXiv:0709.1027 [hep-ph].

[9] F. Krauss, R. Kuhn, and G. Soff, *AMEGIC++ 1.0: A Matrix element generator in C++*, JHEP **0202** (2002) 044, arXiv:hep-ph/0109036 [hep-ph].

[10] T. Gleisberg and S. Hoeche, *Comix, a new matrix element generator*, JHEP **0812** (2008) 039, arXiv:0808.3674 [hep-ph].

[11] M. Spannowsky, *Particle phenomenology,*. Lecture series given as a sub-module of the particle theory lecture course at Durham University in 2012-13.

[12] C. Maxwell, *The Standard Model*, Lectures presented at the School for Young High Energy Physicists, Somerville College, Oxford (Sep., 2012). Lecture notes used as supplementary material for the gauge field theory sub-module of the particle theory lecture course at Durham University in 2012-13.

[13] **Particle Data Group**, K. Olive *et al.*, *Review of Particle Physics*, Chin.Phys. **C38** (2014) 090001.

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

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

[16] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (Nov, 1964)585–587. http://link.aps.org/doi/10.1103/PhysRevLett.13.585.

[17] **ATLAS Collaboration**, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys.Lett. **B716** (2012)1–29, arXiv:1207.7214 [hep-ex].

[18] M. Peskin and D. Schroeder, *An Introduction To Quantum Field Theory*, Levant (2005).

[19] D. Griffiths, *Introduction To Elementary Particles*, John Wiley & Sons (1987).

[20] S. Cheatham, *Quarkonia studies with early ATLAS data*, arXiv:0906.0308v1 [hep-ex].

[21] **CMS Collaboration**, *J/Psi prompt and non-prompt cross sections in pp collisions at sqrt(s) = 7 TeV*, Tech. Rep. CMS-PAS-BPH-10-002, CERN, Geneva, 2010.

[22] M. Lamont, *Status of the LHC*, J. Phys.: Conf. Ser. **455** (2013) 012001.

[23] G. de Rijk, *LHC long term upgrade plans and the CERN accelerator R&D programs*, LHC upgrades & R&D, CERN (Oct., 2010).

[24] O. S. Brüning *et al.*, *LHC Design Report*. CERN, Geneva, 2004.

[25] *Pile-up subtraction and suppression for jets in ATLAS*, Tech. Rep. ATLAS-CONF-2013-083, CERN, Geneva, Aug, 2013.

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

[27] C. Gabaldon, *Performance of the ATLAS Trigger System*, Tech. Rep. ATL-DAQ-PROC-2011-042, CERN, Geneva, Nov., 2011.

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

[29] R. Placakyte, *Parton Distribution Functions*, arXiv:1111.5452 [hep-ph].

[30] S. Hoeche, F. Krauss, S. Schumann, and F. Siegert, *QCD matrix elements and truncated showers*, JHEP **05** (2009) 053, arXiv:0903.1219 [hep-ph].

[31] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **06** (2002) 029, arXiv:hep-ph/0204244 [hep-ph].

[32] 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].

[33] T. Gehrmann, S. Hoeche, F. Krauss, M. Schonherr, and F. Siegert, *NLO QCD matrix elements + parton showers in $e^+e^- \to$ hadrons*, JHEP **01** (2013) 144, arXiv:1207.5031 [hep-ph].

[34] B. Webber, *Parton shower Monte Carlo event generators*, Scholarpedia **6** no. 12, (2011) 10662. Revision #128236.

[35] M. Cacciari, G. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **0804** (2008) 063, arXiv:0802.1189 [hep-ph].

[36] E. Norrbin and T. Sjostrand, *QCD radiation off heavy particles*, Nucl. Phys. **B603** (2001)297–342, arXiv:hep-ph/0010012 [hep-ph].

[37] S. Catani, S. Dittmaier, M. H. Seymour, and Z. Trocsanyi, *The Dipole formalism for next-to-leading order QCD calculations with massive partons*, Nucl. Phys. **B627** (2002)189–265, arXiv:hep-ph/0201036 [hep-ph].

[38] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Heavy-quark mass dependence in global PDF analyses and 3- and 4-flavour parton distributions*, Eur. Phys. J. **C70** (2010)51–72, arXiv:1007.2624 [hep-ph].

[39] T. Sjostrand and M. van Zijl, *A Multiple Interaction Model for the Event Structure in Hadron Collisions*, Phys. Rev. **D36** (1987) 2019.

[40] A. Buckley, J. Butterworth, L. Lönnblad, D. Grellscheid, H. Hoeth, *et al.*, *Rivet user manual*, Comput.Phys.Commun. **184** (2013)2803–2819, arXiv:1003.0694 [hep-ph].

[41] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. M. Nadolsky, and W. K. Tung, *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012, arXiv:hep-ph/0201195 [hep-ph].

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

[43] **ATLAS**, G. Aad *et al.*, *Muon reconstruction efficiency and momentum resolution of the ATLAS experiment in proton-proton collisions at $\sqrt{s}$ = 7 TeV in 2010*, Eur. Phys. J. **C74** no. 9, (2014) 3034, arXiv:1404.4562 [hep-ex].

[44] **ATLAS Collaboration**, G. Aad *et al.*, *Measurement of the differential cross-sections of inclusive, prompt and non-prompt J/ψ production in proton-proton collisions at $\sqrt{s} = 7$ TeV*, Nucl.Phys. **B850** (2011)387–444, arXiv:1104.3038 [hep-ex].

[45] **CMS**, V. Khachatryan *et al.*, *Prompt and non-prompt J/ψ production in pp collisions at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. **C71** (2011) 1575, arXiv:1011.4193 [hep-ex].

[46] **LHCb**, R. Aaij *et al.*, *Measurement of J/ψ production in pp collisions at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. **C71** (2011) 1645, arXiv:1103.0423 [hep-ex].

[47] D. Kirkby and W. Verkerke, *The RooFit toolkit for data modeling*, (jun., 2003), arXiv:physics/0306116.

[48] K. Karakostas, *DiJpsi 2012 ntuple status*, Onia Subgroup Meeting (Nov., 2013).

[49] V. Kostyukhin, *VKalVrt - package for vertex reconstruction in ATLAS.*, Tech. Rep. ATL-PHYS-2003-031, CERN, Geneva, Aug, 2003. http://cds.cern.ch/record/685551. revised version number 1 submitted on 2003-09-24 11:10:53.

[50] S. Höche and S. Prestel, *The midpoint between dipole and parton showers*, arXiv:1506.05057 [hep-ph].

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

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

[53] A. Buckley, D. Grellscheid, H. Hoeth, and D. Mallows. http://yoda.hepforge.org, accessed 22/12/14.

[54] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet User Manual*, Eur.Phys.J. **C72** (2012) 1896, arXiv:1111.6097 [hep-ph].

[55] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, Phys.Lett. **B641** (2006)57–61, arXiv:hep-ph/0512210 [hep-ph].

[56] R. Brun and F. Rademakers, *ROOT - An Object Oriented Data Analysis Framework*, Nucl. Inst. & Meth. in Phys. Res. A **389** (1997)81–86. Proceedings AIHENP'96 Workshop, Lausanne, Sep. 1996 See also http://root.cern.ch/.

[57] A. Buckley, *New access to PDF data via LHAPDF6*, Les Houches 2013: Physics at TeV Colliders: Standard Model Working Group Report (2014), arXiv:1405.1067 [hep-ph].

[58] T. Brooks, J. Ebke, and P. Waller. http://mcviz.net, accessed 09/06/15.