THE UNIVERSITY OF SYDNEY

COPYRIGHT AND USE OF THIS THESIS

This thesis must be used in accordance with the provisions of the Copyright Act 1968.

Reproduction of material protected by copyright may be an infringement of copyright and copyright owners may be entitled to take legal action against persons who infringe their copyright.

Section 51 (2) of the Copyright Act permits an authorized officer of a university library or archives to provide a copy (by communication or otherwise) of an unpublished thesis kept in the library or archives, to a person who satisfies the authorized officer that he or she requires the reproduction for the purposes of research or study.

The Copyright Act grants the creator of a work a number of moral rights, specifically the right of attribution, the right against false attribution and the right of integrity.

You may infringe the author's moral rights if you:

- fail to acknowledge the author of this thesis if you quote sections from the work
- attribute this thesis to another author
- subject this thesis to derogatory treatment which may prejudice the author's reputation

For further information contact the University's Director of Copyright Services

sydney.edu.au/copyright

Search for the X_b and other hidden-beauty states in the $\pi^+\pi^-\Upsilon(1S)$ channel at ATLAS

Cameron Cuthbert

Faculty of Science

University of Sydney

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

August 2014

Acknowledgements

Firstly, and most importantly, I would like to express my sincere gratitude to Bruce Yabsley. As my supervisor, he has provided me with invaluable academic guidance and advice throughout my PhD — sometimes in passionate arguments at the whiteboard, and often through popular culture references. He has brought me to a point of competency where I felt that in the latter parts of the X_b analysis we worked more as teammates than as student and sensei. I greatly appreciate all of our on- and off-topic discussions.

I also wish to thank Kevin Varvell for inspiring me to persue particle physics with his undergraduate and honours lecture courses, and for answering my questions along the way. To Aldo Saavedra: thank you for introducing me to the world of physics analysis at ATLAS during my honours year, and for assisting me in my service task project on the SCT detector.

Many people at CERN provided me with extremely useful technical assistance and analysis advice. In particular, I would like to thank Darren Price, James Catmore, Anthony Morley, the B-physics convenors between 2011 and 2014, the e/γ convenors during 2011, and the editorial board on the X_b analysis.

To all my friends from the office — especially Mark, Ian, Nik, Curtis — many thanks for helping me overcome the (surely several thousands of) issues we have had over the past five years, and I hope I was occasionally of help to you too. I thank Alex and Ian for tolerating the ridiculous questions on theoretical physics and pure mathematics with which I often bugged you.

Finally, I thank my friends and family for the ongoing support throughout my postgraduate years. My parents deserve a special mention of gratitude; without their financial assistance and encouragement this PhD simply would not have come together.

Abstract

This thesis primary concerns a search for a hypothetical state — referred to throughout as the X_b — using the ATLAS detector at the LHC. Specifically, the $\pi^+\pi^-\Upsilon(1S)$ channel was reconstructed in the mass ranges between 10.05–10.31 GeV and 10.40–11.00 GeV using the $\Upsilon(1S) \to \mu^+\mu^-$ decay. In the analysis of the 2011 datasets, the method of *candidate selection*, whereby only a single dipion combination was used for each $\Upsilon(1S)$, was adopted. Fits were performed every 10 MeV to test for the presence of a signal peak, but nothing statistically significant was found. The analysis was re-optimised for the 2012 dataset, with the candidate selection method being abandoned in favour of a binning approach. As for the results of the search in the 2011 dataset, nothing statistically significant was observed. This was also the case for dedicated searches for the $\Upsilon(1^3D_J)$ triplet, the $\Upsilon(10860)$ and the $\Upsilon(11020)$, which were performed in both the 2011 and 2012 analyses. Consequently, CLs upper limits were calculated at the 95% confidence level, with values for the relative production rate, with respect to the $\Upsilon(2S)$, of between 0.8% and 4.0%.

Also included are a technical-based study of the performance and operation of the SCT sub-detector, and an investigation of the feasibility of using the dielectron final state for quarkonium studies at ATLAS. The latter of these had a direct influence on the X_b search methodology.

Preface

The thesis that follows is the culmination of studies conducted toward the PhD of the author, between March 2010 and February 2014. The primary focus is a search for a hypothetical hidden-beauty state — the so-called X_b — through dipion transitions to the $\Upsilon(1S)$. The motivation behind the existence of such a particle is based on the symmetry between the $c\bar{c}$ and $b\bar{b}$ quarkonium systems and the experimental characteristics of the $X(3872)$. If the X_b were to be observed, it could have a profound impact on the theoretical understanding of such *exotic* states and, in turn, of the QCD-based dynamics underpinning the structure of all hadrons. Chapter 1 explores this in further detail, providing an introduction to the current theoretical and experimental status of the field of conventional and exotic quarkonium states.

The search itself was conducted with the ATLAS detector [1] at the LHC. For context, both of these are described in Chapter 2, with particular attention paid to the most relevant aspects. The development of the analysis is explained in Chapter 5, with the subsequent chapter dedicated to the results obtained from its application to the full 2011 dataset. Shortly after this was completed, the CMS Collaboration released their results of an analogous study [2], exposing an opportunity for sensitivity improvements that had not yet been fully exploited here. Consequently, the analysis was re-optimised on the 2011 dataset before being applied to the (much larger) 2012 dataset. The results of this, which are the basis for the final outcomes of the search, are explained in Chapter 7.

The procedures involved from the creation of the proton beams to their eventual collision, detection, reconstruction and analysis are extremely involved, requiring the cooperative actions of literally thousands of people in a number of complementary roles (physicists, but also engineers, technicians, collider and detector experts, administrators etc). In the first year of collisions, for example, great importance was placed in developing a detailed understanding of the operation of each of the ATLAS sub-detectors. This was achieved with the help of a mandatory technical-based project for all new members of ATLAS, in which the author participated. This work is presented in Chapter 3. Chapter 4 is another example of an investigation conducted by the author designed to add to the collective understanding of the detectors' capabilities, rather than to produce a specific physics measurement or observation. In fact, the findings of this chapter led to the decision to exclude the dielectron final state, $\Upsilon(1S) \to e^+e^-$, from the X_b search.

Whilst a certain group of people may perform a physics analysis or measurement, no individual can truly claim sole recognition for the project. On the contrary, the *physics analysts* are generally handed the basic physics objects (electrons, muons etc) with all their associated information, and pursue their studies by building up more complicated combined objects or event signatures. This was the case for the analyses presented in Chapters 4, 5, 6 and 7. With this caveat conceded, however, the choices of techniques and procedures are the original work of the author,[1] under the guidance of Dr. Bruce Yabsley and the influence of other similar studies at ATLAS.[2]

It is the intention that the results of the X_b search be summarised as a journal publication — at the time of writing, the process of doing so is well under way and on track to be completed within 2014.

---

[1] With the exception of Section 4.2 — a summary of the $J/\psi \rightarrow \mu^+\mu^-$ cross section measurement peformed by other members of ATLAS.

[2] In particular, the $J/\psi \rightarrow e^+e^-$ cross-section measurement (Chapter 4) was deliberately chosen to be as similar as possible to the corresponding $J/\psi \rightarrow \mu^+\mu^-$ measurement (performed by other members of ATLAS) to allow for a fair comparison.

Contents

List of Figures

List of Tables

Chapter 1

Exotic Quarkonia

The Standard Model of particle physics has been remarkably successful in describing a large number of microscopic phenomena with incredible precision. It is an interacting relativistic quantum field theory exhibiting (internal) symmetry under local $SU(3) \times SU(2) \times U(1)$ transformations. The $SU(3)$ component refers to the *strong* or *Quantum Chromodynamics* (QCD) sector of the model, whose non-Abelian nature leads to difficulties in producing predictions for observables. Quarkonium systems lie on the border between the (calculable) perturbative regime and the non-perturbative regime, providing an ideal environment with which to investigate effective techniques and approaches for QCD.

This chapter begins with an overview of the theory of QCD, leading to an introduction to the heavy quarkonium systems. After a brief discussion on attempts to model these states, a few of the so-called *exotic* quarkonium states are reviewed in Section 1.3, including some of the leading proposals for their structure. Theoretical models and experimental results for the production of both the conventional and exotic variants at the LHC are provided in Section 1.5, followed by a description of their subsequent dipion transitions to the S-wave ground states. This naturally leads into the possibility of observing a new hidden beauty exotic state — the X_b — at the LHC, which is the main focus of this thesis.

1.1 Introduction to QCD and Quarkonium

By 1960, hadron physics had become chaotic; a “zoo” of mesons and baryons had been experimentally observed, whose production and decay could only be understood by introducing the *ad hoc* principles of baryon number and strangeness conservation. The first step to gaining an understanding of the situation, often likened to that of chemistry in the days before the periodic table was established to make sense of the plethora of elements, was taken by Murray Gell-Mann through his so-called *Eightfold way* in 1961 [18]. Under his scheme, the baryons and mesons were organised into striking geometrical octets and decuplets based on their charge and strangeness. As well as organising the particles into orderly groups, it also lead to the prediction of a new particle with a charge of -1 and strangeness -3 (the Ω^-), which was confirmed shortly thereafter in 1964 [18].

The explanation for the origin of this classification was also provided by Gell-Mann[1] a few years later through the *quark model.* He proposed that all of the known hadrons, including the proton and neutron, were combinations of more fundamental objects he called *quarks* — the mesons as quark-antiquark pairs and the baryons as either three quarks or three antiquarks. Though the three quarks were distinguishable based on their *flavour* (isospin and strangeness quantum numbers), the groups of the Eightfold way could be explained as J^{PC} supermultiplets of a $SU(3)$ symmetry amongst them.

Over the next decade, all of the newly discovered resonances were successfully placed into one of these supermultiplets. Even so, there was still some unease over the quark model because

1. free quarks had not been observed,
2. the Δ^{++}, for example, is supposedly comprised of three u quarks in identical states, which violates the Pauli exclusion principle, and
3. it did not explain why other combinations of quarks, e.g. $qq\overline{q}$, were not observed.

Proponents of the quark model introduced the notion of *quark confinement* to explain the first of these, while the latter two were explained by introducing an additional *colour* degree of freedom to the quarks. The modern explanation, *Quantum Chromodynamics* (QCD), is quantum field theory where each of the quark flavours, f, are additionally labelled by a *colour charge*, c, (where c =red, green or blue), and are described by Dirac spinors, $\psi_{f,c}$. The dynamics of the theory come from demanding symmetry under local $SU(3)$ gauge transformations of colour, with the quarks in the defining **3** representation, i.e.

$$\psi_f = \begin{pmatrix} \psi_{f,r} \\ \psi_{f,g} \\ \psi_{f,b} \end{pmatrix} \tag{1.1}$$

and the antiquarks in the conjugate $\overline{\mathbf{3}}$ representation. The eight resulting gauge fields, the quanta of which are called *gluons*, participate in three-point colour-changing interaction terms, and thus, to obey colour conservation, must also carry colour. This leads to the

---

[1] Simultaneously and independently by Zweig, also.

phenomenon of *asymptotic freedom* (see [19]), whereby the effective strength of the QCD interaction tends to zero at large momenta (small distance scales). For low momenta (large distance scales), the strength of the interaction increases without bound, providing a qualitative explanation for the confinement hypothesis. This states that only colour singlet states can exist as free particles — as one attempts to separate a colourless state into two coloured particles, e.g. to dissociate a meson into an individual quark and antiquark, the energy cost of doing so grows linearly. Eventually it becomes energetically favourable to create new particles in-between to establish a new set of colourless states.

The simplest colour singlets can be constructed from the inner product of an element from the $\overline{\mathbf{3}}$ and one from the $\mathbf{3}$ (i.e. $q_c\overline{q}_c$). Alternatively, one can construct a completely antisymmetric three quark state, $\epsilon_{ijk}\psi_i\psi_j\psi_k$, which transforms under a local $SU(3)$ gauge transformation, U, as

$$\begin{aligned}\epsilon_{ijk}\psi_i\psi_j\psi_k \to& \left(\epsilon_{ijk}U_{li}U_{mj}U_{nk}\right)\psi_l\psi_m\psi_n \\ &= \det U \epsilon_{lmn}\psi_l\psi_m\psi_n \\ &= \epsilon_{lmn}\psi_l\psi_m\psi_n. \end{aligned} \tag{1.2}$$

The last equality follows since the determinant of the transformation matrix is 1. A similar argument shows that $\epsilon_{ijk}\overline{\psi}_i\overline{\psi}_j\overline{\psi}_k$ is also an allowed combination. This, then, explains how colour allows for the existence of particles such as the Δ^{++}; the colour part of the baryon wavefunction is anti-symmetric under particle exchange. Notice that the colour singlet requirement does not preclude the possibility for more complicated combinations; in particular, $qq\overline{q}\overline{q}$ combinations are allowable.

In 1970, Glashow, Illiopoulos and Maiani proposed the possibility for a fourth *charm* quark [20] as a solution to an issue involving strangeness-changing neutral currents, and Gaillard and Lee predicted its mass to be in the range $1.5-2$ GeV [21]. Politzer and Applequist suggested that if such a heavy quark were to exist, it should form non-relativistic bound states with a spectrum analogous to positronium [22]. So, when the observation of a new heavy meson was reported simultaneously by a group from Brookhaven [23] and another from SLAC [24] at a mass of ~ 3100 MeV, it was rapidly identified as the 1^3S_1 state[2] of *charmonium*. This was supported by the fact that it was electrically neutral, had the quantum numbers of the photon, $J^{PC} = 1^{--}$, was much heavier than any of the other known mesons, and had a comparatively small total decay width. Within two weeks, the radial excitation, $\psi(2\text{S})$ or ψ', was also discovered in the same SLAC experiment. Figure 1.1a shows the masses and quantum numbers of all the charmonium states

---

[2] The specification given here is in spectroscopic notation, $N^{2S+1}L_J$, where N, S, L and J are the principle, spin, orbital angular momentum and total angular momentum quantum numbers, respectively.

observed to date [3], which is now essentially complete (at least below the threshold for decay into a pair of $D-$mesons, which are the lightest states containing a charm quark).

Following the discovery of the narrow charmonium states, similar signals were observed in 1977 in the mass range $9.5 - 10.5$ GeV [25, 26], and were immediately claimed as bound states of a new, even heavier, *beauty* (or *bottom*) quark. The bottomonium spectrum, Figure 1.1b, is richer in the region below the open flavour threshold due to the higher quark mass. It has been well mapped out over the past few decades, but still has some conspicuous missing pieces (the dashed lines show the expected positions) or states which have only been observed with marginal statistical significance and are still not well understood. ATLAS contributed to this in 2012 with the observation [27] of a new state interpretted as the $\chi_b(3P)$. Of particular interest for this thesis is the $\Upsilon(1^3\mathrm{D_J})$ triplet, which has only recently been seen in e^+e^- collisions in the $J = 2$ state [28, 29].

States in which the heavy quark flavours of the two partons are different also exist — in fact, the B_c^+, with quark content $c\bar{b}$ and a mass of ~ 6.3 GeV, was recently discovered at the CDF detector [30]. Though it is now known that there is a sixth quark much heavier than all the others, the *top*, formation of toponium is generally thought to be excluded by the very short lifetime of this particle [31].

1.2 Models for Quarkonium Spectroscopy

Heavy quarkonium systems sit at the borderline of two commonly used regimes of QCD. The velocity of the quarks can be estimated using the virial theorem of quantum mechanics, which states that

$$\langle T\rangle = \frac{1}{2}\langle \mathbf{X}\cdot\nabla V\rangle \tag{1.3}$$

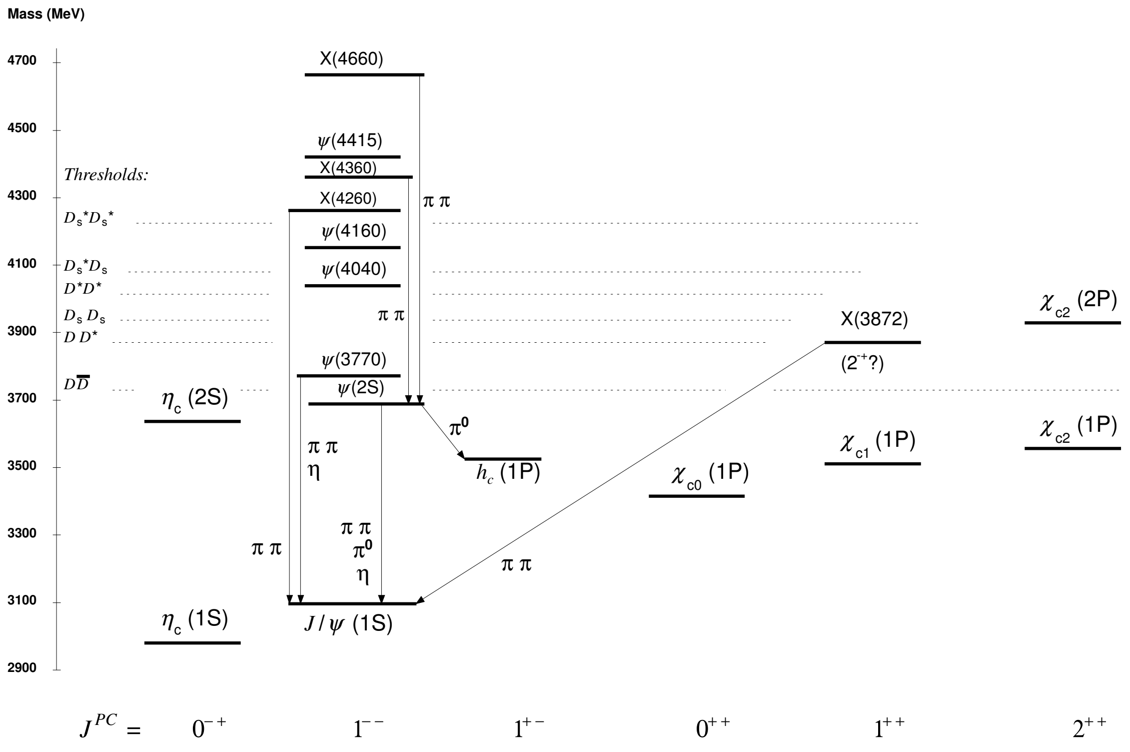
for stationary states, where T is the kinetic energy, $\mathbf{X}$ is the position operator, and V is the potential energy. For power-law potentials, $V = Ar^n$, this takes the simple form

$$\langle T\rangle = n/2\langle V\rangle, \tag{1.4}$$

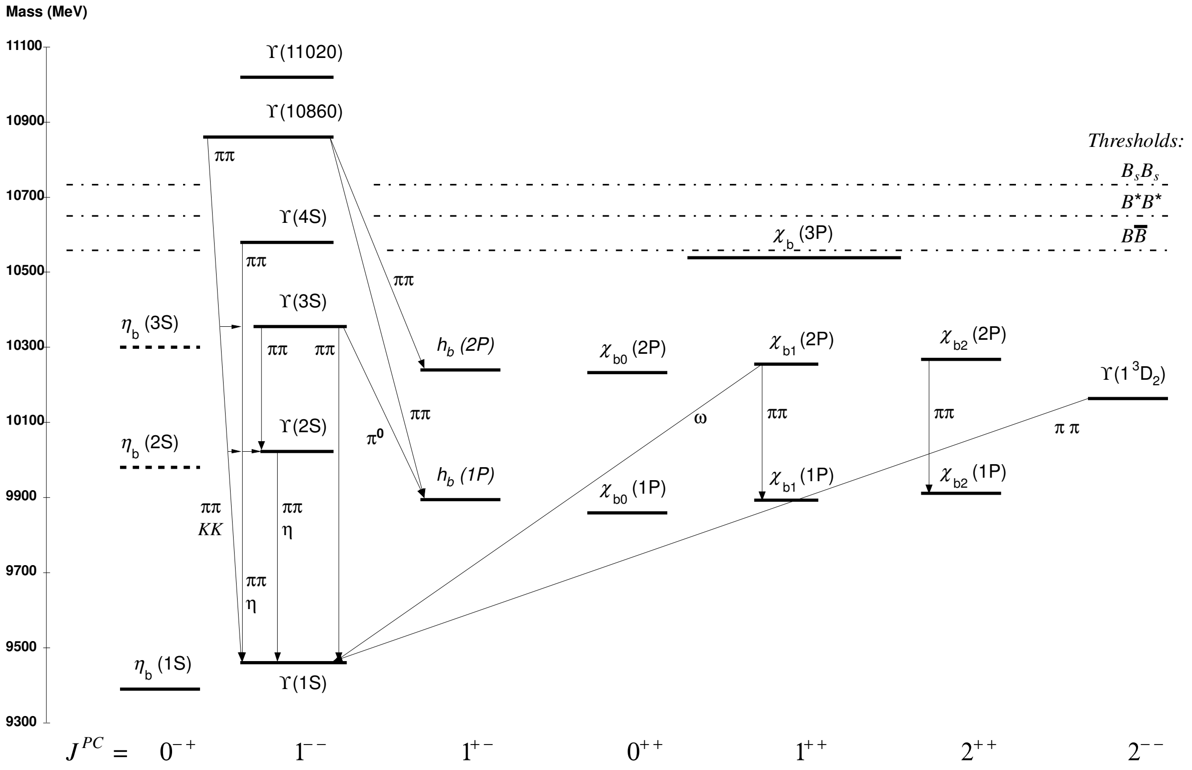
allowing the mean kinetic energy to be written in terms of the binding energy, E_b:

$$\begin{aligned} E_b &= \langle T\rangle + \langle V\rangle \\ &= (1 + 2/n)\langle T\rangle. \end{aligned} \tag{1.5}$$

As explained below, the strong potential is usually modeled as a combination of a coloumb-like part at short separations and a linear part at larger separations. Since the binding energies for quarkonia are all positive, the latter can be assumed to give



(a) Charmonium



(b) Bottomonium

Figure 1.1: The currently observed $c\bar{c}$ and $b\bar{b}$ spectra [3].

the dominant contribution. In this case, the velocity of each quark can be estimated as $v \sim \sqrt{E_b/3m}$, where m is the mass of the quark, giving values of $v \sim 0.37$ and $v \sim 0.3$ for the J/ψ and $\Upsilon(1\mathrm{S})$, respectively. The quark velocity for other states within the charmonium or hidden-beauty families will be different, but modern authors generally quote values of $v^2 = 0.3$ and $v^2 = 0.1$, respectively (see [5, p. 30], for example). Because of this, there exists a hierarchy of energy scales between the energy of the state itself, $\sim m$, the typical momentum of the quark, $p \sim mv$, and the energy of the quark, $\sim mv^2$:

$$m >> mv >> mv^2. \tag{1.6}$$

The first two of these are large in comparison to the QCD confinement scale, $\Lambda_{QCD} \sim 200$ MeV, leading to values of the strong coupling constant $\alpha_S << 1$, where perturbative methods can be used. For the lowest scale, it is typically the case that $mv^2 \sim \Lambda_{QCD}$, meaning that non-perturbative techniques must be used. To complicate matters, Feynman diagrams relavant to heavy quark systems will typically involve momenta at all three of these energy levels. This makes heavy quarkonium an ideal testing ground for the interplay between perturbative and non-perturbative regimes of QCD in a controlled environment.

Based on the non-relativistic nature of quarkonium and the similarity in their spectra to positronium (see [32]), states are often labelled in spectroscopic notation — i.e. as $n^{2S+1}L_J$, where n is the principle quantum number, S is the spin of the system, L the orbital angular momentum and J the total angular momentum. Though most interactions preserve the values of L, it is technically not a good quantum number and mixing can occur between states with $\Delta L = \pm 2$ (due to parity conservation - see below). More fundamentally, states are also labelled by their J^{PC} quantum numbers. The parity, P, can be determined from the intrinsic parity of the quark-antiquark pair, -1, and the effect of a parity transformation on the spherical harmonics describing non-relativistic eigenstates of L^2:

$$P = (-1) \cdot (-1)^L = (-1)^{L+1}. \tag{1.7}$$

For $f\overline{f}$ (f = fermion) systems, a charge conjugation transformation is equivalent to a parity transformation plus a switching of spin states. The latter of these two can be explicitly calculated to have an effect of $(-1)^{S+1}$, combining with the former to give $C = (-1)^{L+S}$.

The simplest and earliest attempts to model the $Q\overline{Q}$ quarkonium systems were based on solving the Schrödinger equation with a potential, $V(\mathbf{x})$, representing the strong interaction between quarks. At very short separations, r, QCD becomes asymptotically free and the potential is expected to be dominated by a Coulomb-like dependence:

$$V(r \to 0) = -\frac{4}{3}\frac{\alpha_S}{r}, \tag{1.8}$$

where $-4/3$ is a colour factor. In this equation, the running value of $\alpha_S(r)$ is given by [33]

$$\alpha_S(r) = \frac{2\pi}{9\ln\frac{1}{r\Lambda_{QCD}}}. \tag{1.9}$$

At larger separations, the confinement hypothesis dictates that the potential should grow without bound, for example

$$V(r) = -\frac{4}{3}\frac{\alpha_S}{r} + \frac{r}{a}, \tag{1.10}$$

for some unknown constant, a. The form of $V(r)$ in Equation 1.10 was first used by Eichten *et. al.* in 1975 [34], and is referred to as the *Cornell Potential Model.* Fine and hyperfine splittings are then achieved by adding relativistic, spin-orbit, spin-spin, and tensor perturbation terms [33]. The free parameters of the model are then used to perform a fit of the resulting spectrum to the observed masses of the known charmonium or bottomonium states. Recent examples of such calculations (see, e.g. [35]) are reviewed in [31] with an updated comparison to the currently observed states. Most of the models are able to reproduce the observed spectra to a good level of accuracy — their true test comes in their ability to predict the masses and mass splittings of states that have not yet been observed (e.g. the splitting in the $\Upsilon(1^3\mathrm{D_J})$ triplet).

The quarkonium spectra can also be predicted from a more fundamental standpoint. For example, lattice QCD can be used to calculated the 2-point Green's Function for the creation of a heavy $Q\overline{Q}$ pair in a zero-momentum state with definite J, P and C and destruction of the state a time t later. For large t, the particle should decay with a characteristic dependence of $\sim \exp^{-mt}$, where m is the mass of the ground state [31]. The masses of the excited states can be extracted by assuming a multi-exponential decay and performing a fit on precise data. The QCD sum rules method is based on a similar principle [33]. Unfortunately, the accuracy of the lattice calculations is still limited by computing power, which leads to large uncertainties on these predictions. The computational demands can be decreased by using an anisotropic lattice[3] or by using an effective field theory (EFT) in the place of QCD. The most commonly used EFTs include Non-Relativistic QCD (NRQCD) and potential NRQCD (pNRQCD), which both exploit the hierarchy of energy scales mentioned above by integrating out high energy modes (see, e.g., [5] for more information). These EFTs can also be used to more accurately

---

[3] In lattice QCD, space-time is discretised into points with a characteristic separation. An anistropic lattice is one in which the spacing in the space dimensions is different to that in the time dimension.

determine the correction terms of the potential models, as well as calculations involved in predicting production and decay of heavy quarkonium (see Section 1.5).

1.3 Discovery of Exotic Quarkonium

Until about a decade ago, it seemed as though the theoretical models for quarkonium systems could account for the observed spectra of the charmonium and bottomonium systems; though new states were still being discovered, they all had properties (mass, width, decays etc) consistent with *conventional* $Q\overline{Q}$ states. In recent times, a large number of new states have been observed at the B factories (Belle and BaBar) which, for various reasons, are inconsistent with the simple quarkonium picture. This section focuses on just a few of these *exotic* states; a more complete review can be found in Section 2.9 of [5]. In particular, the $X(3872)$ is the oldest and most well-studied exotic state, and is the main motivation for the analysis reported in this thesis.

1.3.1 The $X(3872)$

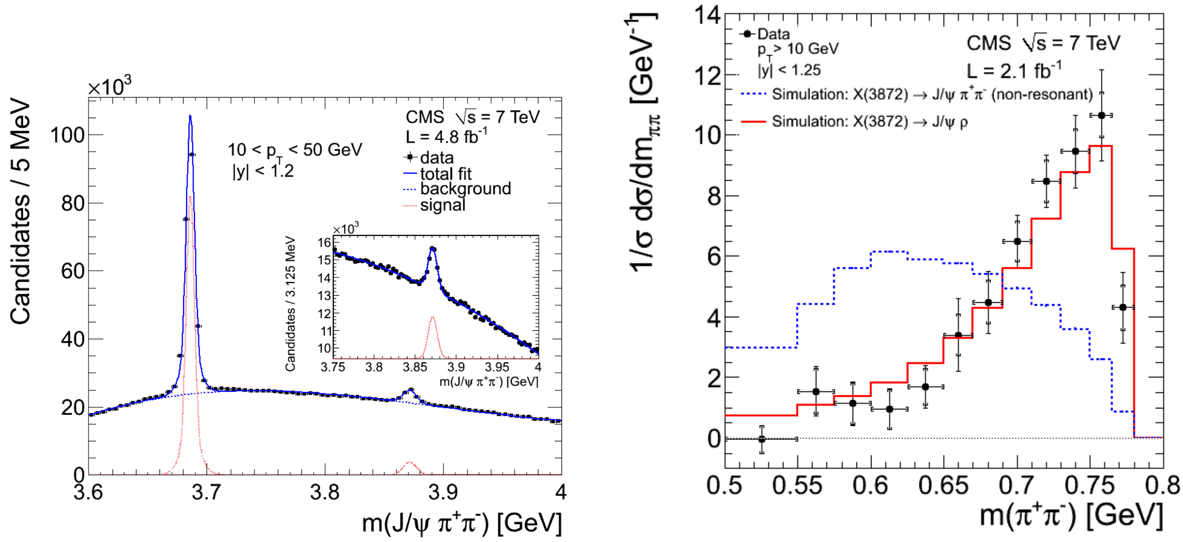
In 2003, Belle reported an enhancement in the $\pi^+\pi^- J/\psi$ spectrum of $B^+ \to K^+\pi^+\pi^- J/\psi$ decays near $m = 3872$ MeV [36]. This was later confirmed by BaBar [37] in the same decay mode, then by CDF [38] and DØ[39] in $p\overline{p}$ collisions, and more recently by CMS [4] and LHCb [40] in pp collisions at the LHC. A recent mass distributon from CMS is shown in Figure 1.2a.

Its proximity to the $D^{*0}\overline{D}^0$ threshold was immediately conspicuous, leading to the interpretation of $X(3872)$ as a molecular state (more detail is given in the next section). For the molecular hypothesis to be possible, it is required that the binding energy be positive, $E_b > 0$. The expression for the binding energy which exploits the most precise measurements currently available is [41]

$$\begin{aligned} E_b &= M(D^0) + M(D^{*0}) - M(X(3872)) \\ &= 2M(D^{*0}) - (M(D^{*0}) - M(D^*)) - M(X(3872)) \\ &= 0.16 \pm 0.32 \text{ MeV}, \end{aligned} \tag{1.11}$$

which is still inconclusive.

Although the state is above the $D^0\overline{D}^0$ threshold, decays through this channel were not observed. This is consistent with the current upper limit for the width, $\Gamma < 1.2$ MeV [42]. This decay is presumably disallowed on the grounds of parity conservation, which immediately limited the $X(3872)$ to the so-called unnatural quantum numbers 0^-, 1^+,



(a) $X(3872)$ mass distribution in $\pi^+\pi^- J/\psi$ decays at CMS.

(b) Dipion mass distribution in $X(3872) \to \pi^+\pi^- J/\psi$ decays.

Figure 1.2: The latest $X(3872)$ and dipion mass distributions from CMS [4].

2^- etc. Some years later, CDF performed an angular analysis of $X(3872) \to \pi^+\pi^- J/\psi$ decays which excluded all but two of these options — $J^{PC} = 1^{++}$ or 2^{-+} [43]. They also demonstrated [44], from the dipion invariant mass, that it was favoured that the $\pi^+\pi^-$ pair originate from a ρ^0. This is now well-established, and the latest version of this result from CMS is shown in Figure 1.2b. The question of the correct quantum numbers assignment was finally settled by the LHCb experiment very recently, who conclusively excluded 2^{-+} in favour of 1^{++} by studying the angular distributions in $B^\pm \to K^\pm X(3872)$ decays [40].

With this now established, the only plausible conventional charmonium assignment available to the $X(3872)$ is as the $\chi_{c1}(2P)$. From the outset, this seemed unlikely given the mass estimates for this state of $m \sim 3930 - 3990$ MeV [31]. The predicted widths for such a state are much too large — $\mathcal{O}(10$ MeV$)$ [45]. Furthermore, the ratio of radiative $\gamma J/\psi$ decays to $\pi^+\pi^- J/\psi$ decays would be expected to be ~ 10 based on potential models and other similar decays [31], but the combined Belle and BaBar measurements for this ratio are currently 0.31 ± 0.08 [5].

At this point, then, it seems that the $X(3872)$ cannot be interpreted merely as a $c\bar{c}$ charmonium state, but its decays imply that it contains hidden[4] charm. The possibility

---

[4] I.e. it contains equal numbers of charm and anti-charm quarks, hence the charm quantum number for the state is zero.

that this is a molecular state has already been mentioned, but there are also a number of other competing options. Some of these will be discussed in Section 1.4.

1.3.2 Exotic Substructure in $\Upsilon(5\mathrm{S}) \to \pi^+\pi^-\Upsilon(1\mathrm{S})$ Decays

In a recent study [46], it was found that the partial width for the decay $\Upsilon(5\mathrm{S}) \to \pi^+\pi^-\Upsilon(1\mathrm{S})$ is two orders of magnitude larger than the corresponding values for the $\Upsilon(1\mathrm{S})$, $\Upsilon(2\mathrm{S})$ and $\Upsilon(3\mathrm{S})$. The Belle collaboration has subsequently reported the observation of substructure in the $\Upsilon(5\mathrm{S}) \to \pi^+\pi^-\Upsilon(1\mathrm{S})$ transition which could explain this enhancement [15]. Specifically, they claimed the discovery of two new charged states, $Z_b^{\pm}(10610)$ and $Z_b^{\pm}(10650)$, in the decay chain $\Upsilon(5\mathrm{S}) \to Z_b^{\pm}\pi^{\mp}$, where $Z_b^{\pm} \to \pi^{\pm}\Upsilon(nS)$. They also showed a consistent observation of such states in $\Upsilon(5S) \to \pi^+\pi^- h_b(mP)$ decays, and since then have gone on to claim significant evidence for $Z_b^{\pm}(10610) \to B\overline{B}^*$ and $Z_b^{\pm}(10610) \to B^*\overline{B}^*$ [47]. By inspecting $\Upsilon(5\mathrm{S}) \to \pi^0\pi^0\Upsilon(\mathrm{nS})$ decays, they have also found evidence for a neutral isotriplet partner to the charged states $Z_b(10610)$.

In contrast to the $X(3872)$, the Z_b states are *manifestly* exotic by virtue of their charge; their production in $\Upsilon(5\mathrm{S})$ decays indicates that the valence quark content must include $b\bar{b}$, but the charge indicates the need for an additional $u\overline{d}$ or $\overline{u}d$ pair, for example. Again, the mass of these states [15] seems to provide a strong clue that the structure of these states may be predominantly that of a molecular state:

$$\begin{aligned} m(Z_b^{\pm}(10610)) &= 10608.4 \pm 2.0 \approx m(B^{\pm}) + m(B^*) = 10604.5 \pm 0.4 MeV \\ m(Z_b^{0}(10610)) &= 10609^{8}_{-6} \pm 6 \approx m(B^{0}) + m(B^*) = 10604.8 \pm 0.4 MeV \\ m(Z_b(10650)) &= 10653.2 \pm 1.5 \approx 2M(B^*) = 10650.4 \pm 0.6 MeV. \qquad (1.12) \\ & \qquad (1.13) \end{aligned}$$

The most favoured spin-parity assignment for the states, according to an angular analysis from Belle [15], is $J^P = 1^+$, which is also consistent with the molecular interpretation.

These Z_b states are the first indications of exotic structures containing hidden beauty. The discussion of the possibility for such states is continued in Section 1.7.

1.4 Models for the $X(3872)$

In the time since its discovery, many models have emerged which claim to account for the exotic properties discussed in the previous section. With increasing understanding of its production and decay properties, most of these have been excluded. Two models — the molecular and tetraquark interpretations — have emerged as the favoured options and are discussed in some detail below. Some of the other alternatives to these are also briefly mentioned.

1.4.1 Molecular Models

The discovery of the $X(3872)$ was immediately of interest due to its closeness in mass to the $D^0\overline{D}^{*0}$ threshold. The notion of mesons forming loosely bound molecules in a manner akin to that of protons and neutrons in deuteron is not a new one. A decade before the discovery, Tornqvist [48] used a single pion exchange model based on a well-established procedure for deuteron to predict a $D^0\overline{D}^{*0}$ molecular state with a mass of $m \approx 3870$ MeV.

Since the Belle discovery, interest in these molecular states has been reignited [45]. Swanson [49] extended the Tornqvist analysis by assuming a short distance interaction dominated by quark exchange processes, where an effective potential $V(\mathbf{r})$ is extracted by comparing the scattering amplitude with that from point-like mesons interacting via a generic S-wave potential. These quark exchange processes naturally lead to mixing of the $D^0\overline{D}^{*0}$ with hidden charm combinations, such as $\omega J/\psi$ and $\rho J/\psi$. Swanson also found that the most likely bound state is S-wave, which is encouraging in light of the quantum number assignment 1^{++}.

The molecular interpretation leads to some specific predictions on decay characteristics:

- As it is expected to have $\omega J/\psi$ and $\rho J/\psi$ admixture components, decays to these states should be common; in contrast, decays to the latter would be isospin-violating for a conventional charmonium state and would be suppressed. Assuming $\rho \to \pi^+\pi^-$, which was shown to be the case by CDF [38] (see Figure 1.2b for the latest CMS result), both decays have been observed by Belle and BaBar [5]. The measured relative decay rate of $1.0 \pm 0.4 \pm 0.3$ is expected based on the contributions to the molecular wave function from these two modes in the Swanson model [50, 49].

- The components of the molecule are weakly interacting and should decay almost independently (i.e. as though they were free). Consequently, decays to $D^0\overline{D}^0\pi^0$ and$D^0\overline{D}^0\gamma$ should occur in the ratio 62:38.

Another avenue for investigating the molecule hypothesis is the production of this state in $p\overline{p}$ and pp collisions at the Tevatron and the LHC, respectively. The relative production rate with respect to the ψ' was estimated (see [5]) at $4.7 \pm 0.8\%$ at the Tevatron (under some reasonable assumptions), which appeared too high in comparison to predictions based on MC for molecular production. However, a tunable parameter (the maximum allowed momentum k_{max} for $D^0\overline{D}^{*0}$ formation) and some underlying assumptions about the uncertainty principle behaviour built into the MC generator were found to have a large impact on the production rate. Various papers were released arguing for and against certain choices, but the matter remains unsolved (see p.44 of [5] for a review of this issue). The recent differential production cross section measurement from CMS [4] provides a more detailed and precise comparison point. This result will be discussed in Section 1.5.3.

The above discussion was restricted to the meson combinations in the proximity of the $D^0\overline{D}^0\pi^0$ threshold, but in Tornqvist's original paper [48] he found six viable molecular states in the charm section and an additional six in the beauty sector. Presumably, these should also be observable using similar experimental searches, which is an important factor in Section 1.7, where potential X_b candidates are considered.

1.4.2 Tetraquark Models

As noted above, there is nothing in QCD which prohibits the formation of colour singlet $qq\overline{qq}$ states. The possibility that the $X(3872)$ is an example of such a state was proposed by Maiani *et. al.* [51] about a year after its discovery. The central assumption of their model is that the state is a diquark-antidiquark pair containing hidden charm — i.e. $[qc][\overline{qc}]$ — with one in a colour antitriplet scalar state and the other a colour triplet vector. The mass of the tetraquark is derived from a constituent quark model with dynamics based on spin-spin interactions, where the coefficients are inferred from the known meson and baryon spectra or from educated guesses from one-gluon exchange processes. Under the symmetric arrangement of singlet and vector components, $[qc]_S[\overline{qc}]_V + [qc]_V[\overline{qc}]_S$, the $J^{PC} = 1^{++}$ quantum number assignment can be achieved.

Because of the mass of the $X(3872)$ and the decay characteristics, the additional quarks have to be either u or d, corresponding to $X_u = uc\overline{uc}$ and $X_d = dc\overline{d}\overline{c}$. These states should mix with an angle θ, leading to two neutral X candidates with a mass difference of $\sim 7 \pm 2$ MeV. This is inconsistent with a study from CDF, who placed an upper limit on such a splitting of 3.6 MeV for a relative production fraction of unity [52].

The main issue with tetraquark models is that they generate a large number of additional new states which have not yet been observed. In particular, the $X(3872)$ should have charged partners of the form $[cu][\overline{c}\overline{d}]$ at a similar mass and with a similar production rate in B decays [45], and their should also be open charm tetraquarks. So far, there is no evidence that these extra states exist.

Putting these issues aside, Maiani claims that the tetraquark model for the $X(3872)$ is able to explain several of its notable characteristics; its narrow width, the non-observation of the $D\overline{D}$ decay and the decays into $\rho J/\psi$ and $\omega J/\psi$.

1.4.3 Other Alternatives

A number of additional explanations have been offered for the $X(3872)$ over the past decade. These include [45]:

- **Hybrids:** bound states with gluonic content (i.e. $c\bar{c}g$) are allowed by QCD and may be investigated with many of the same techniques as for conventional quarkonium (potential models, lattice QCD, QCD sum rules etc). The predicted masses for these states tends to in the range 4200–4400 MeV [45], making them unlikely to be the underlying structure of the $X(3872)$.

- **Glueball with $c\bar{c}$ admixture:** along the same lines as above, a glue ball is a state consisting solely of gluons. This model was excluded many years ago by the observation of radiative decays (gluons are electrically neutral and cannot directly couple to photons).

- **Tetraquark variant:** whereby the four quarks in X_u or X_d above are not arranged into diquark-antidiquark pairs but instead are all loosely bound. This model still produces a splitting of the neutral states into two mass eigenstates and charged parters.

- **Charmonium with coupled channel effects:** the $c\bar{c}$ state is directly coupled to the $D\overline{D}$ meson continua. This still produces a $\chi_{c1}(2P)$ state at $\sim$ 4000 MeV, but additionally leads to a virtual bound state just above the $D\overline{D}^*$ threshold. In practice this would be indistinguishable from a weakly bound molecular state [45]. This model can also spoil some of the good agreement in the remainder of the $c\bar{c}$ spectrum.

- **Cusps:** due to threshold effects have been excluded by the narrow width of the peak.

1.5 Production in *pp* Collisions at the LHC

So far, only the spectroscopy of the quarkonium states and models for their structure have been explored. In this section, the *production* of these states in *pp* collisions is discussed with the specific aim of collecting together experimental observations from Run I of the LHC which will be of importance for the search for the X_b. For context and physical intuition, a basic overview of the popular theoretical methods for estimating differential cross sections is also provided (see Section 4 of [5] for more details).

1.5.1 Theoretical Approaches to Calculations

At a naive level, the amplitude for the production of a quarkonium state can be estimated as follows (see also Section 5.3 of [19]). Consider the Schrödinger wavefunction, $\psi(\mathbf{x})$, of the non-relativistic potential model for the state of interest. This can be converted to a wavefunction in momentum space by performing a Fourier transformation,

$$\psi(\mathbf{k}) = \int d^3x \exp^{i\mathbf{k}\cdot\mathbf{x}} \psi(\mathbf{x}). \tag{1.14}$$

The bound state can then be represented in the centre-of-mass frame as a weighted integral over final states, $|Q_k\overline{Q}_{-k}X\rangle$, containing a free heavy quark (Q) and antiquark ($\overline{Q}$):

$$|B\rangle \propto \int \frac{d^3k}{(2\pi)^3}\psi(\mathbf{k})|Q_k\overline{Q}_{-k}X\rangle. \tag{1.15}$$

The amplitude for the production of this state can then be calculated from the amplitude for production of free two-particle final states, $\mathcal{M}(pp \to Q_k\overline{Q}_{-k}X)$, as

$$\mathcal{M}(pp \to BX) \propto \int \frac{d^3k}{(2\pi)^3}\psi(\mathbf{k})\mathcal{M}(pp \to Q_k\overline{Q}_{-k}X), \tag{1.16}$$

where normalisation proportionality factors have been omitted.

The state-of-the-art theoretical approaches to calculating production rates are more sophisticated, but are generally based on the same principles. As explained in Section 1.2, quarkonium systems naturally lead to a heirarchy of energies — $m >> mv >> mv^2$, where m is the mass of the heavy quark. The initial production of the heavy quark antiquark pair, which go on to form the bound state of interest, must occur at energies $E > 2m >> \Lambda_{QCD}$, making this step inherently perturbative. On the other hand, the subsequent hadronisation will involve a mixture of perturbative and non-perturbative processes at energy scales $E < m$. The central assumption of all of the popular theoretical approaches is that these two stages are decoupled, and can be calculated separately. If H_q is the quarkonium state of interest with quantum numbers collectively represented by q, this is equivalent to the *factorisation* of the production amplitude:

$$\mathcal{M}(pp \to XX'H_q) = \sum_n \mathcal{M}(pp \to X(Q\overline{Q})_n)\mathcal{M}((Q\overline{Q})_n \to X'H_q), \tag{1.17}$$

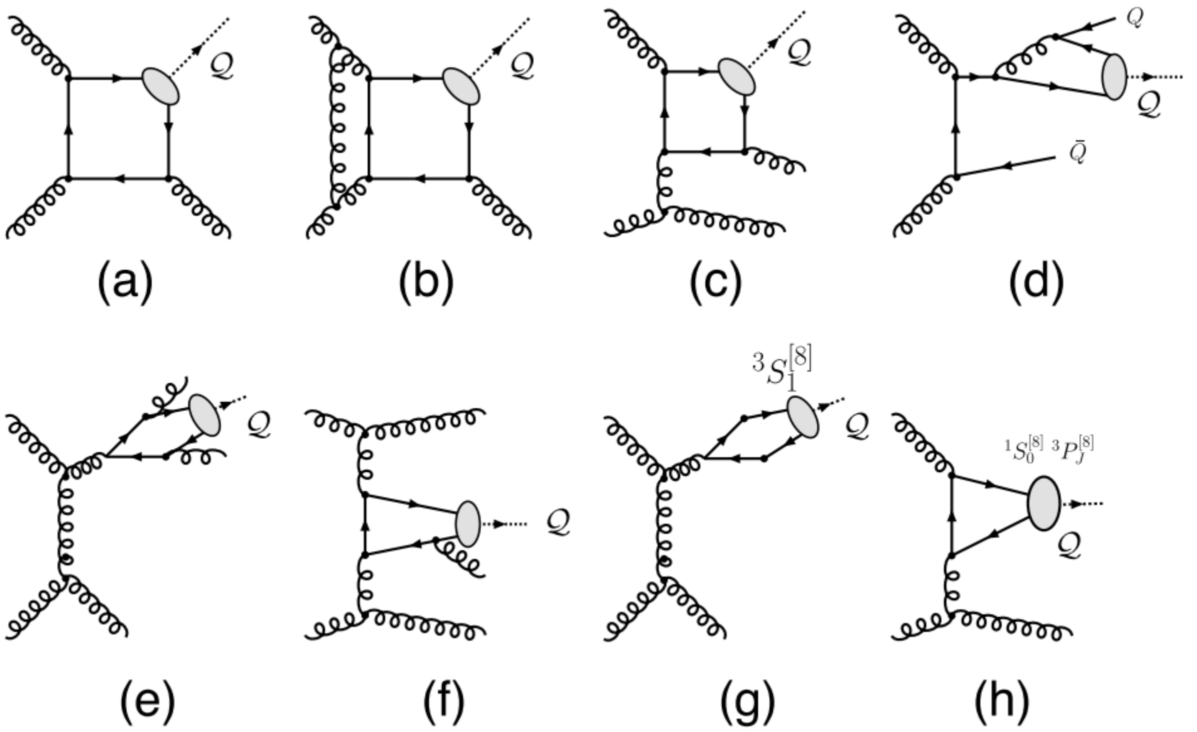


Figure 1.3: Example diagrams contributing to the amplitude for hadroproduction of $^3S_1$ quarkonium states via colour-singlet channels (a-f) and colour-octet channels (g,h) [5]. Where the heavy quarks (solid lines) are attached to the ellipses (representing the quarkonium) they are taken to be on shell and have zero relative velocity. The labels on the ellipses give the different possible spectroscopic assignments.

where n represents the quantum numbers of the intermediate heavy quark-antiquark pair. The first of the factors represents the "short distance" amplitude for partons from two protons to interact and form a $Q\overline{Q}$ state in a particular colour, spin and orbital angular momentum state, n. This can be calculated using perturbative QCD in powers of α_S by convolving the relevant Feynman diagrams (dominated by gluon fusion process — see Figure 1.3) with parton distribution functions for the protons. In these diagrams, the quarks and antiquarks are assumed to be on shell with zero relative momentum.

In constrast, the "long distance" matrix elements (LDMEs), $\mathcal{M}((Q\overline{Q})_n \to X'H_q)$, describing the ensuing hadronisation of the $Q\overline{Q}$ pair into the quarkonium bound state of interest (and any by-products, X'), are inherently non-pertubative and must be obtained by other means. These are presumed to be *universal* — that is, they take the same value independent of the initial production mechanism — allowing them to be extracted from results at one experiment and then used to make predictions at another (see below for more details on this).

In the simplistic analysis above, the long distance factors are effectively encompassed in the Schrödinger wavefunction together with the condition that the two sets of quantum numbers are the same — $n = q$. This is not dissimilar to the assumptions made in the *Colour Singlet Model* (CSM), which was proposed shortly after the discovery of the J/ψ (e.g. [53]). The CSM assumes that the $Q\overline{Q}$ pair created in the hard process is in a colour singlet state with the same spin and angular momentum quantum numbers as the final quarkonium. The hadonisation factors are related to the absolute values of the wavefunction of the quarkonium state and its derivatives, evaluated at zero separation, but are generally extracted by comparing theoretical predictions with experimental observations.

The Non-Relativistic QCD (NRQCD) factorisation approach is a more theoretically well-motivated extension of the CSM. NRQCD is an effective theory of QCD for heavy quarks, where energy scales above m are integrated out of the action to reduce the Lagrangian to an *effective* lagrangian, $\mathcal{L}_{NRQCD}$ [54]. This effective lagrangian describes the dynamics of the heavy quarks at energy scales $E < m$, and so is well suited to the conditions under which the $Q\overline{Q}$ pair hadronise. The NRQCD factorisation is usually written as

$$\mathcal{M}(pp \to XX'H) = \sum_n \mathcal{M}(pp \to X(Q\overline{Q})_n)\langle \mathcal{O}_n^H \rangle, \tag{1.18}$$

where the LDMEs, $\langle \mathcal{O}_n^H \rangle$, are matrix elements of NRQCD operators taking the input state $Q\overline{Q}_n$ into the final quarkonium, H, plus any additional by-products, X'. These operators have simple scaling characteristics with v, the relative velocity of the quark-antiquark pair, so that the NRQCD amplitude is a double series in α_S and v. Phenomenologically, this is usually truncated at an order allowing adequate precision, with the relevant LDMEs extracted from existing experimental data. In contrast to the CSM, the NRCQD approach allows for the possibility that the $Q\overline{Q}$ pair is produced in a colour-octet state — if these are excluded, and only the leading order in v is used, then one recovers the CSM from Equation 1.18.

Another popular model is the so-called *Colour Evaporation Model* (CEM), which assumes that every $Q\overline{Q}$ pair produced below the open-flavour threshold will evolve into a quarkonium state. The probability for this pair to form the state H is specified by a constant, F_H, which is energy-momentum and process independent, and is generally extracted from data. As in the NRQCD factorisation method, colour-octet production is permissible — the excess colour supposedly *evaporates* away in the hadronisation procedure through, for example, the emission of a gluon. This model is usually disfavoured in comparison to the two above because of its lack of theoretical motivation and rigour.

NRQCD is distinct from the CSM and CEM in that it is *not* a model — given that the assumptions regarding energy scales are reasonable, NRQCD is equivalent to QCD for the processes of interest. To the extent that the factorisation formula is correct and the perturbation series in α_S and v converges, the predictions of NRQCD *should* be equivalent to those of the full theory. In fact, whereas the factorisation of the production amplitudes was taken as an *Ansatz* above, it should be possible to *prove* this as a feature of NRQCD. Currently, this has only been shown up to two loops — proving this relation for all orders of perturbation theory is one of the key problems in quarkonium theory at the current time.

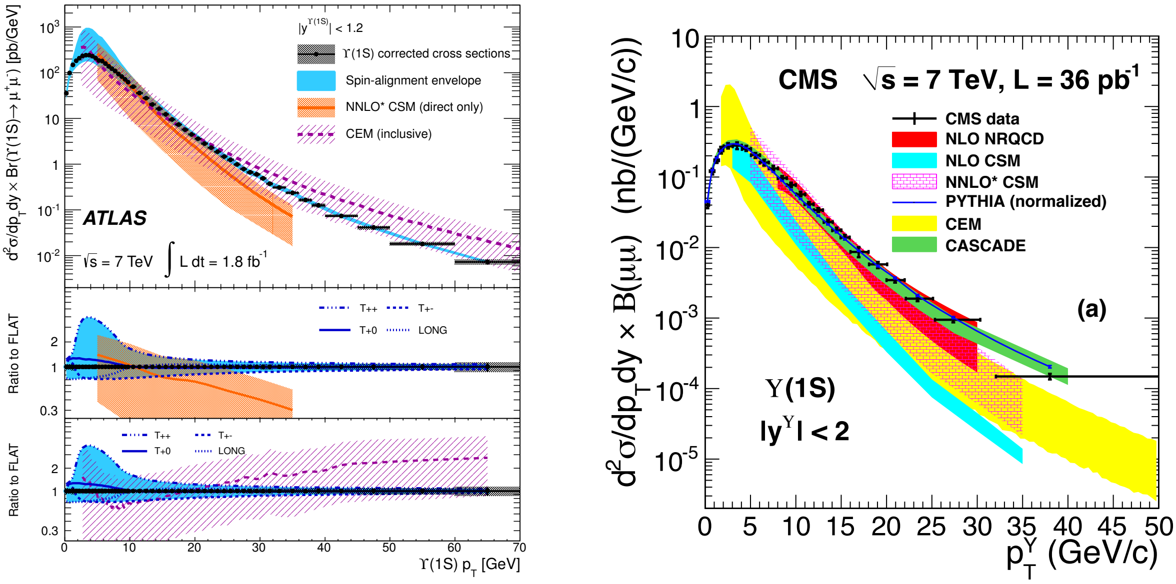
The prediction of differential cross-sections based on the above methods is further complicated by the effects of feed-down from higher states in the quarkonium system of interest, which can be significant for the lowest lying states (e.g. J/ψ). The dependence on the modeling of this feed-down can be reduced by performing theory/data comparisons of the excited states, such as the ψ'.

1.5.2 Υ(1S), Υ(2S) and Υ(3S) Production

When the first measurements [55] of J/ψ and ψ' production in $p\overline{p}$ collisions at the CDF detector were released, they were an order of magnitude higher than the predicted leading-order (LO) CSM predictions. This originally led some to believe that the colour-octet mechanisms may play an important role in production, particularly at high p_T. Since then, the next-to-next-to-leading-order (NNLO) corrections to the CSM predictions have been shown to be large, accounting for most of the original discrepancy in the CDF results [56].

With the emergence of production cross section results from the LHC experiments, the Tevatron data now serves the purpose of determining the unknown LDMEs. The latest results for Υ production from the LHC seem to indicate that none of the models are capable of providing accurate predictions for both the magnitude and shape of the differential cross sections in p_T [6, 57]. Figure 1.4 shows recent measurements of Υ(1S) production from ATLAS (left) and CMS (right), both of which indicate that the CSM predictions perform reasonably well at intermediate p_T but, even at NNLO, underestimate the high p_T tail. This tends to again point to the importance of including colour-octet production mechanisms, which have a slower drop-off in p_T [5]. In fact, the NLO NRQCD and PYTHIA (where the production of quarkoinia is simulated based on NRQCD) predictions are in good agreement with the CMS data.

CMS later updated their measurement to use the full 2011 dataset (~ 4.9 fb^{-1}) [58], allowing them to probe higher in p_T and with smaller binning. Though they have not yet



(a) ATLAS [6]. (b) CMS [57].

Figure 1.4: Recent measurements of the differential production cross sections for $\Upsilon(1S)$ as a function of p_T, with comparisons to various theoretical models. Explanations of these models are given in the papers cited in each figure.

made comparison to specific theoretical predictions, they do show that the differential cross section can be well-described by an exponential decay for $p_T < 20$ GeV and a power law of the form

$$\frac{d\sigma}{dp_T} = \frac{A}{C + \left(\frac{p_T}{20}\right)^{\alpha}} \tag{1.19}$$

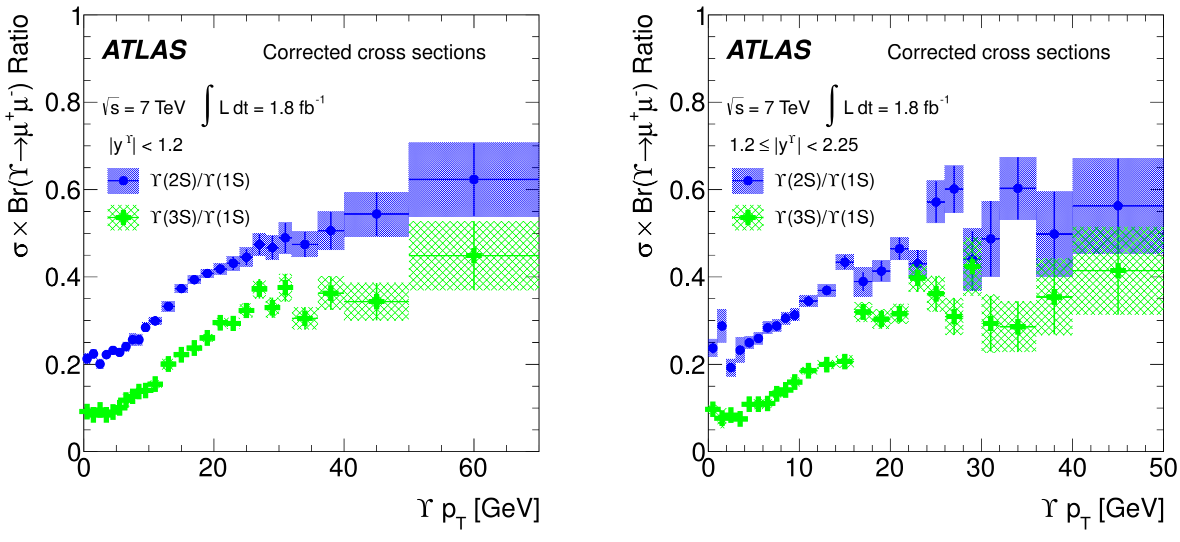
for higher p_T.

Another feature of Υ production evident from the CMS [58] and ATLAS [6] measurements is the increasing hardness in the spectrum with increasing principal quantum number. That is, although the overall cross section is higher for $\Upsilon(1S)$ production, the tail at high p_T for the normalised distributions is significantly higher for the $\Upsilon(2S)$ and even more so for the $\Upsilon(3S)$ (see Figure 1.5a). Some of this is due to the softening effect of feed-down, but there may also be an intrinsic difference in the production shapes for excited states.

The ATLAS and CMS analyses also reported the differential cross section as a function of rapidity.[5] Figure 1.6 gives the $\Upsilon(1S)$ result for ATLAS in the range $|y| < 2.25$ [6] and for a CMS-LHCb combination spanning $|y| < 4.5$, showing that the cross-section is

---

[5] The *rapidity* is defined as $y = \frac{1}{2}\ln\frac{E+p_z}{E-p_z}$. Its usefulness comes from the property that differences in rapidity and differential cross section shapes with respect to rapidity are both preserved under boosts along the *z-axis* [3].



(a) $|y| < 1.2$.

(b) $1.2 < |y| < 2.25$.

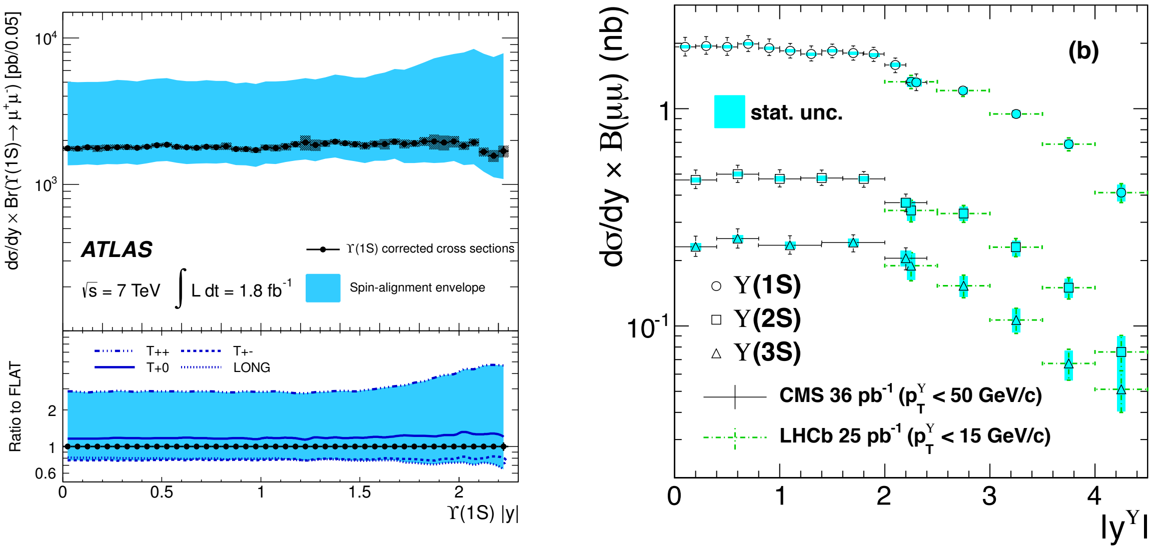
Figure 1.5: The differential production cross section ratios measured by ATLAS [6] for the $\Upsilon(2S)$ (blue) and $\Upsilon(3S)$ (green) with respect to the $\Upsilon(1S)$ in bins of transverse momentum. The results are given for two different rapidity regions.

essentially flat until $|y| \sim 2$, where it slowly decreases thereafter. This trend will be used in Chapter 7 to extrapolate the ATLAS measurement.

The production spectra for quarkonium states played an important role in the analysis described later in this thesis (Chapters 5-7). Simulation samples were produced using PYTHIA, in which modeling for the production of the parent state is based on NRQCD with tuneable LDME parameters; the default values are based on fits to previous measurements [59]. Various issues arose from differences between the data and simulation — these, and the solutions implemented to circumvent them, are discussed in Sections 5.2, 6.4.1, 7.1.1 and 7.6.8.

Aside from the production *rate* of quarkonia, another key observable is the polarisation, or spin-alignment, of these states. Any theoretical approach hoping to accurately describe the dynamics at play in producing quarkonium should consistently be able to account for both aspects. The Υ states, with $J = 1$, can be expressed in terms of spin eigenstates with respect to a particular z axis, $|\Upsilon\rangle = a_1|1\ \ 1\rangle + a_0|1\ \ 0\rangle + a_{-1}|1\ \ -1\rangle$. Assuming the Υ decays to a pair of muons, the helicity formalism [60, 61] can be used to completely determine the angular distribution as

$$\frac{d^2N}{d\cos\theta^* d\phi^*} \propto 1 + \lambda_\theta \cos^2\theta^* + \lambda_\phi \sin^2\theta^* \cos 2\phi^* + \lambda_{\theta\phi} \sin 2\theta^* \cos\phi^*, \tag{1.20}$$



(a) ATLAS [6].

(b) CMS [57].

Figure 1.6: The differential cross section for Υ(1S) production as a function of rapidity from ATLAS (left) and CMS (right).

where ϕ^* and θ^* are the decay angles of the μ^+ in the rest frame for the Υ. The λ variables in Equation 1.20 are simple functions of the a_i which depend on the direction of the polarisation or z-axis. A common choice is the *helicity* frame, where the polarisation axis coincides with the Υ momentum vector in the lab frame (more details on this can be found in [62]). There is also another variable,

$$\tilde{\lambda} = \frac{\lambda_\theta + 3\lambda_\phi}{1 - \lambda_\phi}, \tag{1.21}$$

which provides a frame-independent measurement of polarisation.[6]

The measured angular distributions of dilepton decays from Υ produced in pp collisions will depend on the processes which created them. For example, the colour-singlet mechanisms at NNLO lead to mostly longitudinal polarisations ($a_0 = 1, a_{\pm1} = 0$ in the helicity frame), whereas the colour-octet processes that get included in the NRQCD approach predict strong transverse polarisations ($J_z = \pm 1$ in the helicity frame) at high transverse momenta. Contrary to both of these, CMS has recently reported values [7] for the λ paramters of the Υ(1S), Υ(2S) and Υ(3S) which are consistent with zero — i.e. uniform or flat angular distributions — for the entire p_T range considered (see Figure 1.7

---

[6] Regardless of the convention used, the polarisation axis always lies within the production plane. As such, the angular distributions for each can be related to one another by a rotation involving a single angular variable. $\tilde{\lambda}$ is an invariant of this rotation.

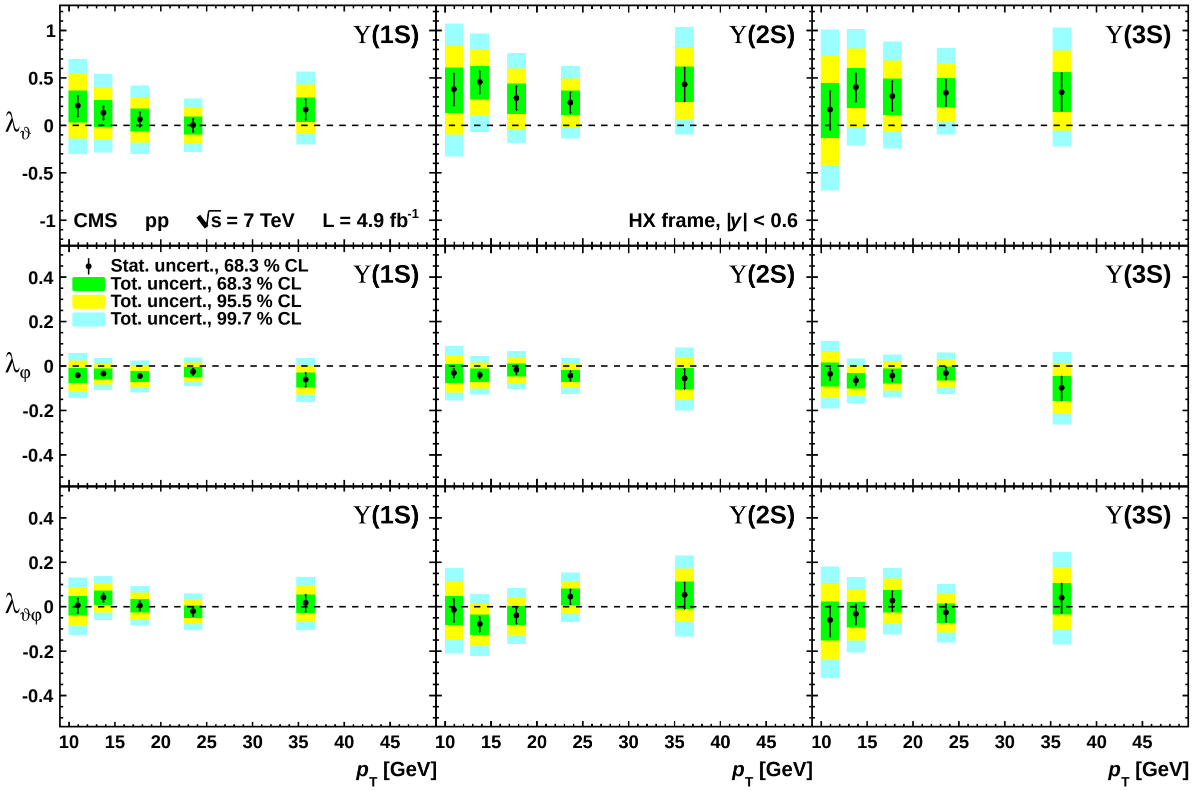


Figure 1.7: The measured angular distribution parameters for the Υ(1S), Υ(2S) and Υ(3S) in the helicity frame as measured by CMS [7]. All results are consistent with uniform angular distributions.

for the helicity frame results). This agrees well with a similar measurement from CDF for Υ produced in $p\bar{p}$ collisions at the Tevatron [63].[7] Obviously, the stark disagreement with CSM and NRQCD predictions is a significant additional issue in understanding the mechanisms at work in producing these states.

1.5.3 X(3872) Production Cross Section

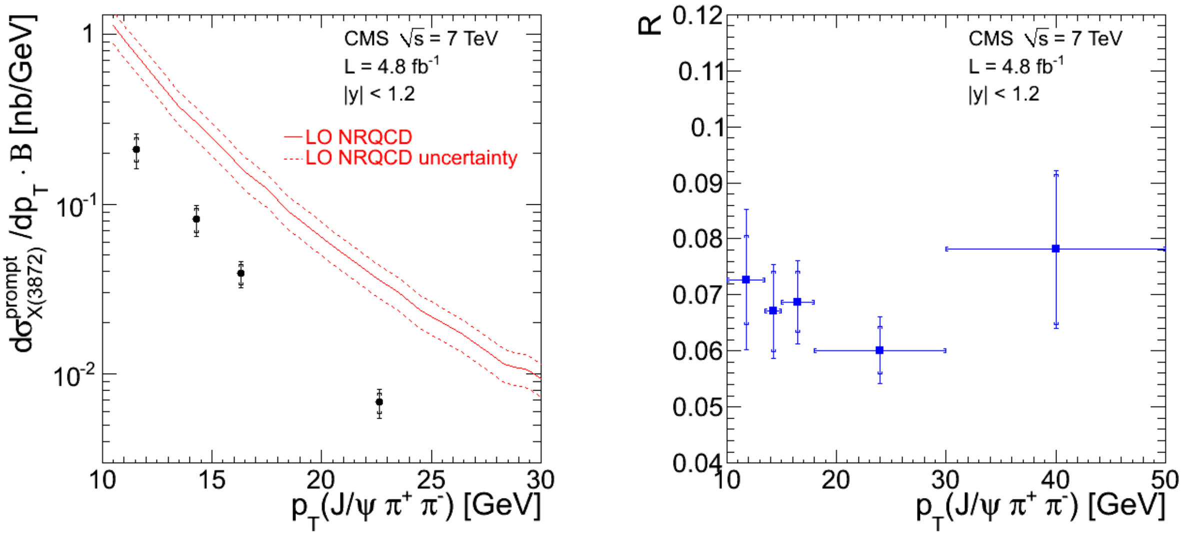
Neither CDF nor DØ made a measurement of the production cross-section of the $X(3872)$ at the Tevatron, but from publicly available documents and reasonable assumptions on selection efficiencies it can be estimated as

$$\sigma(p\bar{p} \to X + \text{ anything}) \cdot \mathcal{B}(X \to \pi^+\pi^+ J/\psi)|_{|y|<0.6, p_T>5\text{ GeV}} = (3.1 \pm 0.7)\text{ nb}, \qquad (1.22)$$

which is ~ 5% of the equivalent value for the ψ'. This is the rate for *prompt* production, which, importantly, demonstrates that the $X(3872)$ is created in the hard collision as

---

[7] There is, however, a disagreement between the Υ polarisation measurements from CDF and DØexperiments — see Figure 5 in [63].



(a) Differential production cross section

(b) Production rate with respect to the ψ'

Figure 1.8: Result of the CMS cross section measurement for the $X(3872)$ [4].

well as in B decays (in fact, the majority of $X(3872)$ at CDF were produced promptly [64]).

In 2012, CMS extended this by performing a measurement [4] of the differential cross section as a function of p_T with the $\pi^+\pi^-J/\psi$ decay channel. This is shown in Figure 1.13, alongside a measurement of the relative production rate, R, with respect to the ψ'. The theoretical prediction [65] given in Figure 1.8a is due to an NRQCD-based estimate for production of an $D^{*0}\overline{D}^0$ molecule, where simplifying assumptions reduce the number of LDMEs to a single value (which the authors extracted from the Tevatron estimate of the prompt cross section). Clearly this does not agree well with the observed result, but from previous experience (see Section 1.5.2) NLO and NNLO corrects may be important.

For future reference, the average value of R from this measurement is $(6.56 \pm 0.29 \pm 0.65)\%$.

1.6 $\pi^+\pi^-$ Transitions Amongst Quarkonia

Once produced, quarkonia inevitably decay through mostly electromagnetic and strong interactions, with characteristic lifetimes of $\mathcal{O}(10^{-20}$ s). In terms of strong decays, the obvious low-order process of $q\overline{q} \to g$ is blocked by colour conservation (the quarkonium is in a singlet state but the gluons are are in octet representation). To annihilate the heavy quark-antiquark pair, then, requires at least two gluons and sometimes a third, depending on the J^{PC} quantum numbers of the quarkonium state. As an example, $J/\psi \to ggg$ and $\Upsilon(1\text{S}) \to ggg$ are the dominant decays for the 1^3S_1 states,[8] leading to a rich array of hadronic final states [3]. In general, though, these strong annihilation decays neccessarily involve Feynman diagrams with s-channel gluons, so are suppressed by the OZI rule. In fact, electromagnetic annihilation into dilepton pairs, $q\bar{q} \to \gamma^* \to \ell^+\ell^-$, is a competitive[9] decay for the S-wave quarkonia. Dilepton decays provide a clean detector signature, and are particularly useful for triggering on (see Chapter 2).

For the excited states, transitions *within* the quarkonium system are more common — for example, the branching fraction for $\psi' \to J/\psi X$ is $60.3 \pm 0.7\%$. Radiative decays, such as the recently observed [27] $\chi_b(3\text{P}) \to \gamma\Upsilon(1\text{S})$, are completely analogous to the manner in which atoms and positronium transition between energy levels through the release of a photon. More common are the *hadronic transitions*, arising from the radiation of gluons, the most important of which for the $c\bar{c}$ and $b\bar{b}$ systems are shown in Figure 1.1. The gluons go on to hadronise into one or more lighter particles with u, d, s quark content, controlled mostly by the conservation of angular momentum, parity and $C-$parity. It should be obvious from the previous sections of this chapter that transitions involving a pair of pions have played a crucial role in discovering of some of the new exotic states; their prevalence, strength and distinct detector signatures makes them an ideal search channel.

In this section, the kinematics of three-body decays are discussed. In particular, the $\pi^+\pi^-$ mass distributions is shown to be a key experimental observable, providing a key insight into the dynamics at play in dipion transitions.

1.6.1 Kinematic Considerations

In Appendix 1.A, it is shown that the differential width, $d\Gamma$, for the decay $A \to B, C, D$ is given by

$$d\Gamma = \frac{1}{(2\pi)^5}\frac{1}{64M^3}|\mathcal{M}(A \to B, C, D)|^2 dm_{13}^2 dm_{23}^2 d\Omega_1 d\phi_2, \tag{1.23}$$

where M is the mass of particle A and $\mathcal{M}$ is the amplitude for the transition. Labelling the final state particles with numbers 1–3 in any manner, m_{13} and m_{23} are combined invariant masses, Ω_1 is the solid angle of particle 1 and ϕ_2 the azimuthal angle of particle 2 defined in the rest frame of the decaying state.

---

[8] Analogous to the decay of the 1^3S_1 positronium state into three photons.

[9] The branching fraction for hadronic and electromagnetic decays of J/ψ is $\sim 88\%$ and $\sim 12\%$, respectively.

If, for whatever reason (e.g. the original particle is spin-less or produced in an unpolarised manner), the transition amplitude is independent of the angles in Equation 1.23 then their integral is trivial, and

$$d\Gamma = \frac{1}{(2\pi)^3}\frac{1}{34M^3}|\mathcal{M}(A \to B, C, D)|^2 dm_{13}^2 dm_{23}^2. \tag{1.24}$$

For a dipion transition of the form $X \to \pi^+\pi^- Y$ the mass variables are $m_{\pi\pi}^2$ and $m_{\pi+Y}^2$. For a given value of $m_{\pi\pi}^2$, the energy of each pion, E_π, and the third particle, E_Y, are fixed to the following values in the rest frame of the dipion system:

$$E_\pi = \frac{m_{\pi\pi}}{2} \tag{1.25}$$

$$E_Y = \frac{m_X^2 - m_{\pi\pi}^2 - m_Y^2}{2m_{\pi\pi}}, \tag{1.26}$$

but the orientation of the Y momentum vector is unconstrained. The maximum and minimum of $m_{\pi+Y}$ therefore correspond to the cases where the momentum of the Y and π^+ are aligned or anti-aligned,

$$(m_{\pi+Y}^{\pm})^2 = (E_{\pi^+} + E_Y)^2 - \left(\sqrt{E_\pi^2 - m_\pi^2} \mp \sqrt{E_Y^2 - m_Y^2}\right)^2, \tag{1.27}$$

and the total available range for $m_{\pi+Y}$ is

$$\begin{aligned}\Delta m_{\pi+Y}(m_{\pi\pi}) =& 4\sqrt{(E_{\pi^+}^2 - m_\pi^2)(E_Y^2 - m_Y^2)} \\ =& \frac{1}{4m_{\pi\pi}}\sqrt{(m_{\pi\pi}^2 - 4m_\pi^2)\left[(m_X^2 - m_{\pi\pi}^2 - m_Y^2)^2 - 4m_{\pi\pi}^2 m_Y^2\right]}.\end{aligned} \tag{1.28}$$

The allowed values for $m_{\pi\pi}$ are those for which this range is ≥ 0, which can be solved from Equation 1.28 as $2m_\pi \leq m_{\pi\pi} \leq m_X - m_Y$ (alternatively, this follows from energy conservation and the definition of $m_{\pi\pi}$). This, together with Equation 1.28, define a region in the $(m_{\pi\pi}, m_{\pi+Y})$ plane. In the case of pure phase space only — i.e. $\mathcal{M}$ is a constant for any decay satisfy energy-momentum conservation — a scatter-plot in this plane, called a *Dalitz plot*, will be uniformly filled. Conversely, distinct patterns within the Dalitz plot point to non-trivial dynamics in the decay.

If the hypothesis of pure phase space is correct, then the projected distribution in $m_{\pi\pi}$ alone can be found by setting $\mathcal{M}$ constant and integrating Equation 1.24 with respect to $m_{\pi+Y}$:

$$\frac{d\Gamma}{dm_{\pi\pi}^2} = \frac{1}{(2\pi)^3}\frac{1}{34m_X^3}|\mathcal{M}(X \to \pi^+\pi^- Y)|^2 \Delta m_{\pi+Y}(m_{\pi\pi}). \tag{1.29}$$

Finally, substituting in the formula for $\Delta m_{\pi+Y}$ and changing variables to $m_{\pi\pi}$ gives the pure phase space distribution for the dipion mass,

$$\implies \frac{d\Gamma}{dm_{\pi\pi}} \propto \sqrt{(m_{\pi\pi}^2 - 4m_\pi^2)\left[(m_X^2 - m_{\pi\pi}^2 - m_Y^2)^2 - 4m_{\pi\pi}^2 m_Y^2\right]}. \tag{1.30}$$

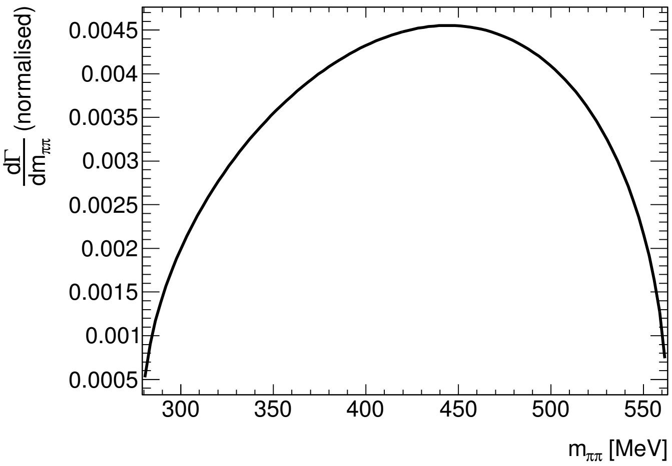


Figure 1.9: The dipion mass distribution which would be expected for the decay $\Upsilon(2S) \to \pi^+\pi^-\Upsilon(1S)$ under the hypothesis of pure phase space.

As an illustrative example, the shape of this distribution for the $\Upsilon(2S) \to \pi^+\pi^-\Upsilon(1S)$ decay is shown in Figure 1.9. The function is concave down with the maximum somewhere just above the midpoint. As for the Dalitz plot, deviations from this phase space distribution would prove that the dynamics of this decay are non-trivial. Though the $\eta_c' \to \pi^+\pi^-\eta_c$ decay data fits reasonably well to a phase space distribution [8], strong disagreements have been observed in almost all other known dipion transitions. The most important of these for this thesis are discussed in the next section, with a mention of some of the theoretical explanations offered to explain the nature of these decays.

1.6.2 Observed $\pi^+\pi^-$ Mass Distributions

When the dipion mass spectrum in $\psi' \to \pi^+\pi^- J/\psi$ transitions was first measured [66], it was considerably different to that expected based on phase space alone. More recently, the distribution for $\Upsilon(2S) \to \pi^+\pi^-\Upsilon(1S)$ decays has been observed at CLEO [8] with a similar shape (Figure 1.10).

The challenge of developing an adequate understanding of the origin of this functional dependence attracted theoretical attention (see below). The popular models often proceed by factorising the dipion transition into two distinct steps (see Figure 1.11):

1. the *emission* of gluons from the quarkonium, which is left in its final state, and

2. the *hadronisation* of the gluons into the $\pi^+\pi^-$ pair.

The emission of gluons occurs at soft energy scales given by the difference in masses between the initial state quarkonium, Φ_i, and the final state, Φ_f, making peturbative

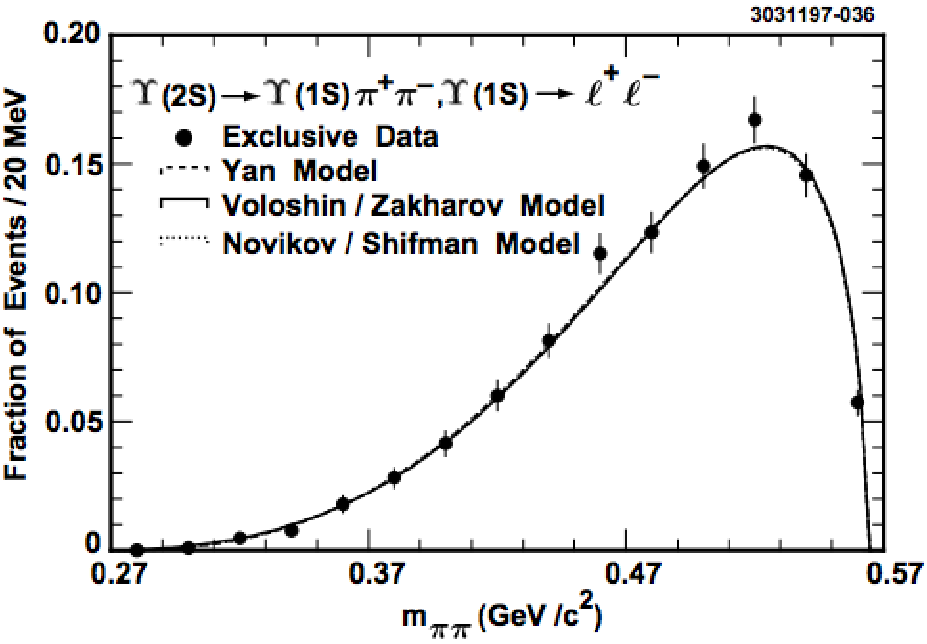


Figure 1.10: The dipion mass distribution in $\Upsilon(2S) \to \pi^+\pi^-\Upsilon(1S)$ decays at CLEO [8]. The fitted shapes are based on the leading theoretical models at the time.

QCD an inappropriate approach. However, this process does naturally lend itself well to the so-called *QCD multipole expansion*, analogous to the multipole expansion of radiating sources in electromagnetism. The central principle of the expansion in the latter is based on the fact that the source, d, is small compared to the wavelength of the radiation, or $kd << 1$. Potential models give a typical size of $\sqrt{\langle r^2 \rangle} \sim 10^{-1}$ fm for quarkonium systems [67] and, taking the energy difference in the $\Upsilon(2S) \to \pi^+\pi^-\Upsilon(1S)$ decay as an example, these transitions have a typical radiation wavelength of $k = E/\hbar c = 2.7 \times 10^{15}$ m^{-1}; combining the two gives $kd \sim 0.27$. An overview of the formalism used to write down an effective Lagrangian for the quarkonium system and then expand it in the usual multipole manner is described in [67]. For the $n_I{}^3S_1 \to hn_F{}^3S_1$ transition, where h represents final state light hadrons, a sequential two-gluon $E1 - E1$ transition is predicted to be the most significant.

The hadronisation stage is what determines the dipion mass distribution. This can be incorporated and made consistent with the multipole expansion using, for example, the Quark Confining String (QCS) model or bag models [67]. Alternatively, one can take the phenomenological approach of using an effective soft-pion lagrangian together with the hypothesis of Partial Conservation of Axial Current (PCAC) to write down the generic form of the hadronisation matrix element [68]

$$\mathcal{M} = \mathcal{A}(\epsilon' \cdot \epsilon)(q^2 - 2M_\pi^2) + \mathcal{B}(\epsilon' \cdot \epsilon)E_1E_2 + \mathcal{C}((\epsilon' \cdot q_1)(\epsilon \cdot q_2) + (\epsilon' \cdot q_2)(\epsilon \cdot q_1)), \quad (1.31)$$

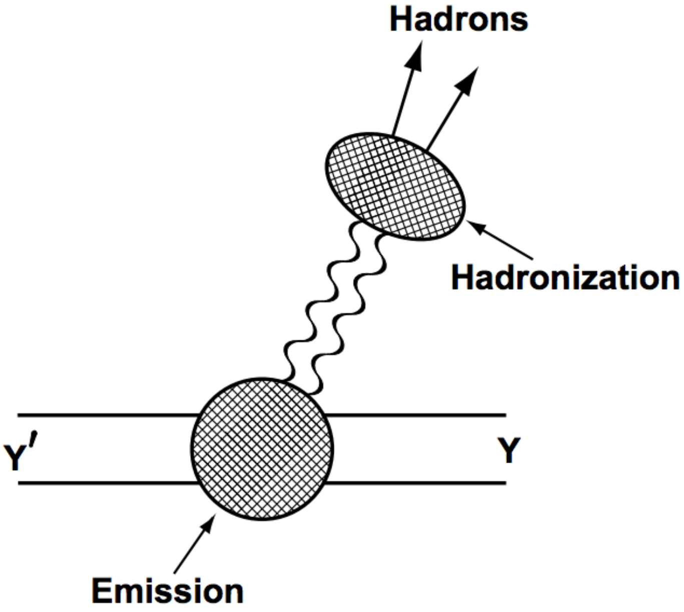


Figure 1.11: The schematic representation of the dipion transition process assumed in most theoretical models [8]. The decay is factorised into an *emission* stage and a *hadronisation* stage (see text).

where ϵ and ϵ' are the polarisation vectors of the parent and final state, $q_{1,2}$ are the four-momenta of the pions, $E_{1,2}$ are the energies of the pions in the parent rest frame and A, B and C are complex-valued form factors. When this is forced to be consistent with two-gluon multipole emission, simplifications can be made. For example, Voloshin and Zakharov [33] found that

$$\mathcal{M} \propto m_{\pi\pi}^2 - \lambda m_\pi^2, \tag{1.32}$$

where λ is an unknown parameter. This, together with Equation 1.30, leads to a total decay rate of

$$\frac{d\Gamma}{dm_{\pi\pi}} \propto (m_{\pi\pi}^2 - \lambda m_\pi^2)^2 \sqrt{(m_{\pi\pi}^2 - 4m_\pi^2)\left[(m_X^2 - m_{\pi\pi}^2 - m_Y^2)^2 - 4m_{\pi\pi}^2 m_Y^2\right]}. \tag{1.33}$$

This functional form, provided a good description of the dipion mass in $\Upsilon(2S) \to \pi^+\pi^-\Upsilon(1S)$ and $\psi' \to \pi^+\pi^-\psi$ decays when fitted to the CLEO data (see Figure 1.10). Similar approaches by others (see [8]) describe the data equally as well, but equation 1.33 is the simplest and will be used to model these $\Upsilon(2S)$ decays in the X_b search presented in the latter half of this thesis. Each of the models was general in the sense that they should apply equally well to any $n_I^3S_1 \to \pi^+\pi^- n_F^3S_1$ transition, so at this point the situation seemed to be fairly well understood.

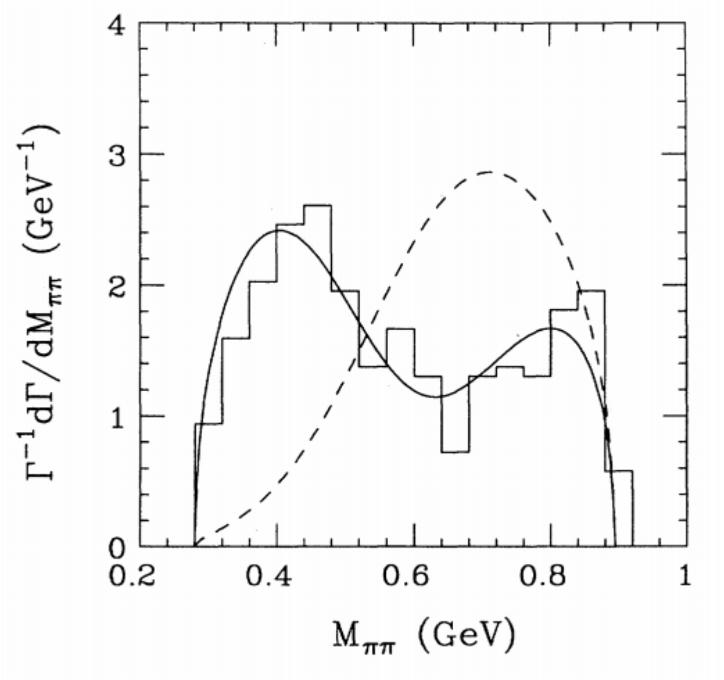


Figure 1.12: The dipion mass distribtion observed by CLEO in $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)$, showing the successful fit using the Moxhay model (solid line) [9] and the unsuccessful description by the multipole/PCAC model with final state interactions.

Since this time, some of the subsequently observed dipion mass shapes, for example that of the $\Upsilon(4S) \to \pi^+\pi^-\Upsilon(1S)$ transition [69, 70], have also been successfully described within this framework. For $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)$ decays [71, 13], the mass shape exhibits a double-humped structure very different to either the pure phase space or $\Upsilon(2S) \to \pi^+\pi^-\Upsilon(1S)$ distributions. This could be an indication of the break-down of either the multipole expansion, PCAC, or the factorisation of the problem into the two stages described by these models. CLEO later succesfully performed a fit to a number of different mass distributions from transitions in the Υ system, including $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)$, based simply on the generic PCAC amplitude (Equation 1.31), suggesting the multipole expansion may be the issue.

Other possibilities included (1) the need to take into account $\pi^+\pi^-$ final-state interactions and (2) assuming there to be exotic or otherwise resonant intermediate states. A successful explanation asserted that $B\overline{B}$ and other open flavour coupled-channels effects are important, which Moxhay [9] modelled by adding to the multipole and PCAC-based amplitude an unknown complex constant. The fit using this altered amplitude, Figure 1.12, reproduces the observed shape by CLEO reasonably well.

The dipion distributions for $X(3872)$ decays and in the $\Upsilon(5S) \to \pi^+\pi^-\Upsilon(1S)$ transition have already been discussed; they show different behaviour still.

In summary, a variety of shapes have been observed in $\pi^+\pi^-$ transitions; the challenge of explaining and predicting them in a satisfactory manner demonstrates the insight they give into the underlying dynamics of the decay. In Chapter 7, the unknown dipion mass distribution in $X_b \to \pi^+\pi^-\Upsilon(1S)$ decays (see below) enters the search as a key systematic uncertainty.

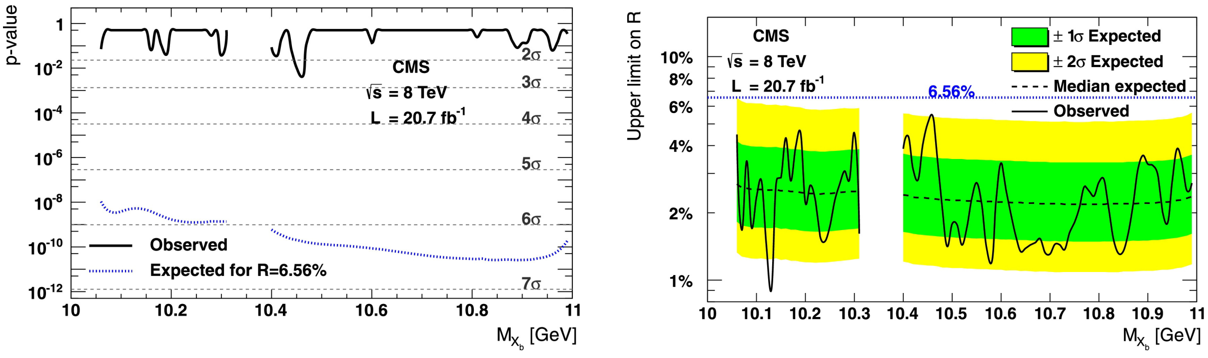
1.7 The Hypothetical X_b State

All of the ingredients are now in place to explain the rationale behind the search reported in this thesis. The symmetry between the beauty and charm quarks due to their heavy masses, and the correspondence between the $c\bar{c}$ and $b\bar{b}$ spectra, leads naturally to the hypothesis that there exists a hidden beauty analogue to the $X(3872)$ with the following properties:

1. quantum numbers $J^{PC} = 1^{++}$,
2. it is a narrow state, either because it is below threshold or becaue open-beauty decays are suppressed or forbidden, with Γ less than the experimental resolution at ATLAS (~ 5 MeV),
3. it is produced promptly in pp collisions,
4. it decays to $\pi^+\pi^-\Upsilon(1S)$ in analogy with tranisition of the $X(3872)$ to $\pi^+\pi^- J/\psi$ (however, there exist isospin-based arguments that the branching fraction for the former may be much smaller than for the latter [72, 73]), and
5. a mass within the range of known $b\bar{b}$ states — roughly 10–11.1 GeV.

For the remainder of this thesis, such a hypothetical state will be referred to as the X_b. The mass of the state and more detailed predictions for its decay properties and rates are model-dependent, for example

- the Swanson model predicts the mass for a 1^{++} $B\overline{B}^*$ molecular state of $m = 10561$ MeV, hereafter the *Swanson mass*, and
- a relativistic tetraquark model predicts $m = 10492$ for $bq\bar{b}\bar{q}$, $m = 10682$ MeV for $bs\bar{b}\bar{s}$ or $m = 10593$ MeV for mixed s/q content, where each has $J^{PC} = 1^{++}$ [74].



(a) Local p-values vs. mass

(b) Upper limits vs. mass

Figure 1.13: Results of the search for $X_b \to \pi^+\pi^-\Upsilon(1\mathrm{S})$ decays performed by CMS. The local p-values (left) show no evidence for any new states in the mass ranges considered, and upper limits on the relative production rate $R = (\sigma\mathcal{B})/(\sigma\mathcal{B})_{2\mathrm{S}}$ are shown on the right [2].

From the properties above, the X_b should be produced at a considerable rate in pp collisions at the LHC, and may appear in the $\pi^+\pi^-\Upsilon(1\mathrm{S})$ mass spectrum as a distinct, narrow peak. In 2013, the CMS collaboration presented the results of a search they performed [2] for the X_b in the $\pi^+\pi^-\Upsilon(1\mathrm{S})$ channel. The analysis, which used $\mathcal{L} = 20.7$ fb^{-1} at $\sqrt{s} = 8$ TeV, found no evidence for any new narrow states in the mass ranges 10.06-10.31 GeV and 10.40-10.99 GeV (Figure 1.13a). Upper limits were set on the relative production rate $R = (\sigma\mathcal{B})/(\sigma\mathcal{B})_{2\mathrm{S}}$ at the 95% confidence level using the CL$_S$ approach, with values between 0.9% and 5.4% (Figure 1.13b).

This thesis presents an analogous search using the ATLAS detector which was carried out during a similar time period to the CMS analysis.[10] The next chapter describes the collider and detector apparatus needed to perform this analysis, which is later described in Chapters 5-7.

Appendix 1.A Three-body Phase Space

The differential rate for a three-body decay, such as $\Upsilon(2\mathrm{S}) \to \pi^+\pi^-\Upsilon(1\mathrm{S})$, is given by

$$d\Gamma = \frac{1}{2M}|\mathcal{M}(A \to B, C, D)|^2 d\Phi, \tag{1.34}$$

---

[10] The study presented in Chapter 6 of this thesis began in late 2011.

where M is the mass of the decaying particle and $d\Phi$ is the Lorentz-invariant phase-space factor specifying the kinematics of the final state particles:

$$d\Phi := (2\pi)^4\delta^4(p^\mu - p_1^\mu - p_2^\mu - p_3^\mu)\prod_{i=1}^{3}\frac{d^3p_i}{(2\pi)^3}\frac{1}{2E_i}. \tag{1.35}$$

Because of the energy-momentum conservation enforced by the delta function, there are only $3 \times 3 - 4 = 5$ degrees of freedom in the decay which enables a considerable simplification of Equation 1.35. Using the vector component of the delta function allows the integral over d^3p_3 to be performed, giving:

$$d\Phi := \frac{1}{8(2\pi)^5E_1E_2E_3}\delta(E - E_1 - E_2 - E_3)d^3p_1d^3p_2, \tag{1.36}$$

where it is now understood that $\mathbf{p_3} = \mathbf{p} - \mathbf{p_1} - \mathbf{p_2}$. Now consider the rest frame of the decaying particle, and express the differentials in terms of polar coordinates. Using the fact that $E_1dE_1 = p_1dp_1$ leads to

$$d\Phi = \frac{1}{8(2\pi)^5E_3}\delta(M - E_1 - E_2 - E_3)p_1dE_1d\phi_1d\cos\theta_1p_2dE_2d\phi_2d\cos\theta_2 \tag{1.37}$$

Performing a rotation so that $\mathbf{p_2}$ is expressed with respect to $\hat{\mathbf{p}}_\mathbf{1}$, the energy of the third particle can be written as

$$E_3^2 = m_3^2 + p_3^2 = m_3^2 + (\mathbf{p_1} + \mathbf{p_2})^2 = m_3^2 + p_1^2 + p_2^2 + 2p_1p_2\cos\theta_2. \tag{1.38}$$

The remaining δ function is therefore equivalent to

$$\begin{aligned}\delta(M - E_1 - E_2 - E_3) &= \delta(M - E_1 - E_2 - \sqrt{m_3^2 + p_1^2 + p_2^2 + 2p_1p_2\cos\theta_2})\\ &=: \delta(f(\cos\theta_2))\\ &= \frac{\delta(\cos\theta_2 - \cos\widetilde{\theta}_2)}{\left|\frac{df}{d\cos\theta_2}\left(\cos\widetilde{\theta}_2\right)\right|},\end{aligned} \tag{1.39}$$

where $\cos\widetilde{\theta}_2$ is the value giving zero for the δ function argument. Evaluating the denominator gives

$$\delta(M - E_1 - E_2 - E_3) = \delta(\cos\theta_2 - \cos\widetilde{\theta}_2)\frac{E_3}{p_1p_2} \tag{1.40}$$

and, using this, the phase space factor can be reduced to

$$d\Phi = \frac{1}{8(2\pi)^5}dE_1dE_2d\Omega_1d\phi_2. \tag{1.41}$$

Without further assumptions, the phase space factor cannot be simplified any further (though the choice of 5 variables above is not unique). Finally, changing variables to

$$m_{13}^2 = (E_1 + E_2)^2 + (p_1 + p_2)^2 = M^2 + m_2^2 - 2ME_2 \tag{1.42}$$

$$m_{23}^2 = (E_2 + E_3)^2 + (p_2 + p_3)^2 = M^2 + m_1^2 - 2ME_1 \tag{1.43}$$

and substituting back into Equation 1.34 gives the following form for the differential rate:

$$d\Gamma = \frac{1}{(2\pi)^5}\frac{1}{64M^3}|\mathcal{M}(A \to B, C, D)|^2 dm_{13}^2 dm_{23}^2 d\Omega_1 d\phi_2. \tag{1.44}$$

Chapter 2

The ATLAS Experiment at the LHC

The Large Hadron Collider (LHC), and its associated detectors, comprise one of the largest and most intricate scientific experiments ever built. A complete description of the LHC and the ATLAS detector is well beyond the scope of this chapter, but can be found in [10] and [1], respectively. Instead, a brief overview of the two is provided, with particular focus on elements which will be referred to later in the thesis. Event triggering, data acquisition and the subsequent algorithms used for track-finding and calorimeter cluster-formation are also discussed, leading to an explanation of how this information is combined to form electron and muon candidates. The muon objects, in particular, are crucial ingredients in almost every quarkonium analysis performed at ATLAS, including that being presented in this thesis.

2.1 Overview of the LHC

The LHC was constructed primarily to search for the Higgs boson and other physics beyond the Standard Model (BSM). The rate of production of these types of events is given by $Ł\sigma$, where Ł is the *instantaneous luminosity* and σ is the *cross-section* for the process of interest. For rare events, the latter of these is typically an increasing function of energy. With this in mind, the LHC was designed to operate at record-breaking energies and greater luminosities than have ever been reached by a hadron collider. Both goals were achieved over the first four years of running (2009-2012), culminating in the observation of the Higgs boson[1] in July, 2012 [75, 76].

To capitalise on the conditions provided by the LHC, several complementary experiments were installed at various points around the ring. The two largest of these, ATLAS and CMS, are general purpose detectors capable of a broad range of physics projects,

---

[1] Technically, the newly observed resonance is a particle with properties consistent with a Higgs boson.

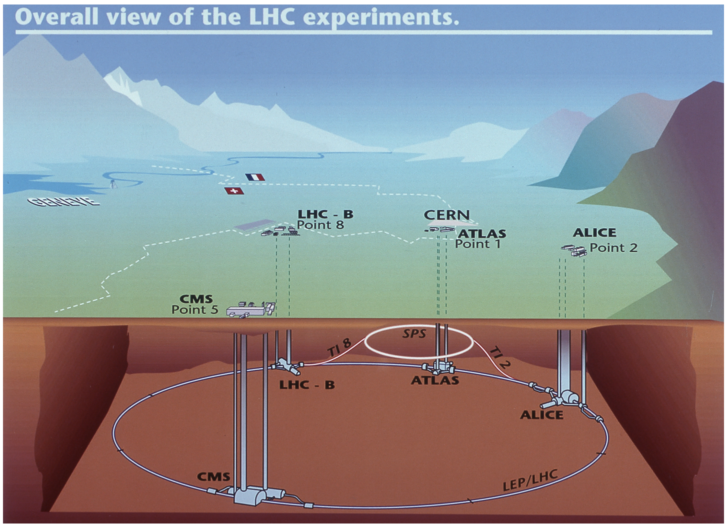


Figure 2.1: A schematic view of the LHC and the four main detectors. ATLAS Experiment © 2014 CERN.

while the smaller LHCb and ALICE experiments were specially designed for $B-$physics and heavy ion collisions, respectively. A schematic view of the detectors and their positions on the LHC ring is shown in Figure 2.1.

The collider itself is housed within the previously-built CERN Large Electron Positron (LEP) facility in Geneva, Switzerland — a 26.7 km tunnel situated 45-150 m below the surface. The ring consists of eight straight sections, each approximately 528 m long, joined by eight arcs (Figure 2.2). Each of the straight sections are labelled as Point N, where N is an integer between 1 and 8, and can serve as an experiment or accelerator utility insertion. For example, ATLAS is located at Point 1, CMS at Point 5, LHCb at Point 2 and ALICE at Point 8.

Unlike LEP, the LHC accelerates and collides two counter-rotating *proton* beams. Each is pre-accelerated to 450 GeV by a series of older facilities, and injected into the main ring in the vertical plane at Points 2 and 8. As the beams pass into each arc section, they are bent by a series of 138 dipole magnets, arranged in 23 *cells*. Because of the limited space within the tunnel, the two beams travel within twin-bore magnets consisting of two oppositely wound coils within the same mechanical structure and cryostat. For an energy of 7 TeV, a dipole field of 8.33 T is required to maintain the beams within the ring and, for this reason, superconductor magnet technology was employed.

The main acceleration of the beams occurs in the straight section at Point 4, where two RF systems are installed — one independent system for each beam, containing eight cavities each generating an oscillating potential difference. Initially, each beam contains a spectrum of energy states, certain of which are *synchronous* in the sense that their

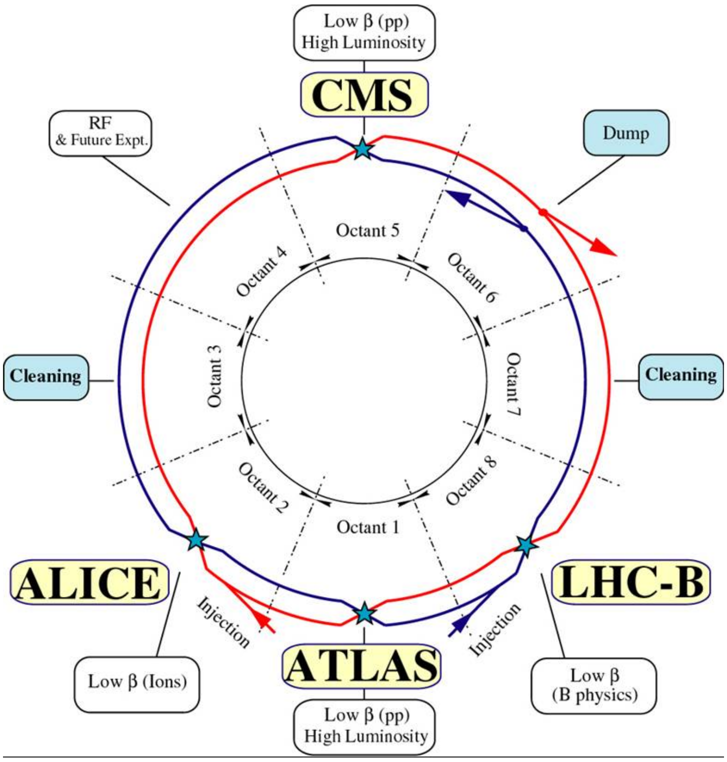


Figure 2.2: An overview of the LHC ring, showing the four main experiments and accelerator utilities [10].

arrival at the RF cavities coincides with a zero in the voltage. Higher energy states arrive slightly earlier and experience a decelerating potential, while those with lower energy are accelerated. In other words, the potential from the RF cavities establishes a restoring force for the beams. This *synchrotron* operating principle has been used by all major particle accelerators in recent times.

If the circulation frequency of the particles around the ring is f and the oscillation frequency of the RF cavities is f_{RF}, then one condition for beam stability is $f_{RF} = hf$, for some integer h. When this is satisfied, the beams naturally arrange themselves into the h available pockets of stable phase space, referred to as *buckets*. The circulation frequency at the LHC is approximately 400 MHz, accommodating approximately 35640 buckets. Under normal running conditions, only a subset of these buckets is filled with *bunches* of protons.

Acceleration occurs when the dipole field strength, B, is adiabatically *ramped up*. This slowly raises the synchronous momentum, $p = eBR$, where R is the radius of curvature in the arc region. When the beams reach their target energy, they are brought into collision at each of the experiments by focusing quadrupoles, sextupoles and octupoles. Each such *bunch-crossing* typically gives rise to several pp collisions, the hardest of which is referred

Table 2.1: Summary of the typical LHC running conditions in 2011 and 2012, compared to their design values. The 2012 results are based on [16] and [17].

	Design	**Typical 2011 Values** [77]	**Typical 2012 Values**
Beam Energy [TeV]	7	3.5	4
Peak Ł [cm^{-2}s^{-1}]	$1{\times}10^{34}$	$3.6{\times}10^{33}$	$7.73{\times}10^{33}$
Protons per bunch, n_p	$1.15{\times}10^{11}$	$1.2{\times}10^{11}$	$1.9{\times}10^{14}$
Number of bunches, N_b	2808	1331	1380
Bunch spacing [ns]	25	≥50	50

to as the primary vertex and the others as *pile-up* vertices. At the end of a run, or if the beams become unstable, their trajectories are (independently) diverted out of the ring in the straight section at Point 6, and the beam is said to have been *dumped.*

The typical operating parameters of the LHC in 2011 and 2012 running are summarised in Table 2.1. In general, the performance exceeded expectations, delivering 5.32 fb^{-1} of integrated luminosity to ATLAS in 2011 and 23.26 fb^{-1} in 2012. It is currently in a shutdown period, as upgrades are made ahead of *Run II.* Data-taking is scheduled to resume in early 2015 at a centre-of-mass energy of $\sqrt{s} = 14$ TeV.

2.2 The ATLAS Detector

The ATLAS detector is the result of over fifteen years of design, construction and installation by thousands of members of the ATLAS Collaboration; a collection of physicists, engineers, technicians and students. Though it is commonly referred to as a *general purpose* detector, its primary task has always been the search for the Higgs boson, and other new physics. In fact, it is stated in [1] that "The search for the Standard Model Higgs boson has been used as a benchmark to establish the performance of important subsystems of ATLAS". For example, one of the most important decay modes is $H \to \gamma\gamma$, which drove the need for extremely good resolution in the electromagnetic calorimeter. Similarly, efficient reconstruction of the the $H \to 4\ell$ channel depends on the tracking capabilities for low-p_T muons and electrons, and understanding the large QCD background processes required good reconstruction of jets by the hardonic calorimeter. Consideration for generic BSM signatures, such as missing E_T, also impacted the design of the detector.

Other physics-motivated design considerations for ATLAS are quite standard, including: the need for close-to hermetic coverage, efficient particle identification, and good measurement resolution in general. Practical and technical issues also placed restrictions on the implementation of the detector elements; they had to be radiation hard, very fast and have high granularity to cope with the immense particle fluxes produced by the collisions.

Despite the design being tailored to Higgs and BSM physics, many of the features of ATLAS make it an excellent experiment for the study of quakonium production and decay in a hadronic environment. Run I of the LHC saw many important results published, including the observation of a new state [27]. Of course, there are also limitations — particularly effect of the trigger on the acceptance for certain decays — which will be pointed out in the appropriate places throughout this thesis.

Probably the most striking feature of the ATLAS detector is its physical extent; installed within the pit at Point 1, it is about 5 storeys high (25 m), 44 m in length, and has a weight of around 7,000 tonnes (c.f. the CDF detector with dimensions 12×12×12 m, or the CMS detector at 15 m tall, 25 m long and 12,000 tonnes [78]). Its over-arching structure is, however, much the same as other symmetric particle detectors at colliders such as CDF, DØ, CMS, and Belle. Placed close to the interaction point is the *Inner Detector* (ID) tracker system, responsible for the identification and measurement of charged particles of all types over a large range of momenta. Consistent with current trends, the ID is comprised mostly of silicon pixel and strip detectors, though additionally features a large number of drift tubes. Facilitating the operation of the ID is the solenoid magnet, which provides a 2 T longitudinal magnetic field and sits immediately outside the tracker. The Electromagnetic and Hadronic Calorimeters are the next subsystems, followed by a dedicated *Muon Spectrometer* (MS), making up the bulk volume of ATLAS. In the central regions the sub-detectors are generally organised into concentric cylindrical layers (the *barrel*), while in the forward and backward regions they are placed on wheels perpendicular to the beam direction (the *endcaps*). A schematic overview of this arrangement is provided in Figure 2.3.

In what follows, the sub-systems particularly important for B-physics and quarkonium studies are described in some detail. It should be kept in mind that the detector is presented as it was during Run I — subsequent to this, many sub-detectors have been (or are planned to be) upgraded to ensure that the performance does not suffer under the increased luminosity and energy planned for the future runs. The intended changes to the detector for the next two phases of LHC running can be found in [79] and [80], respectively.

For convenience and future reference, the coordinate system of ATLAS is briefly introduced. The x-axis points toward the centre of the ring, with the y-axis nearly

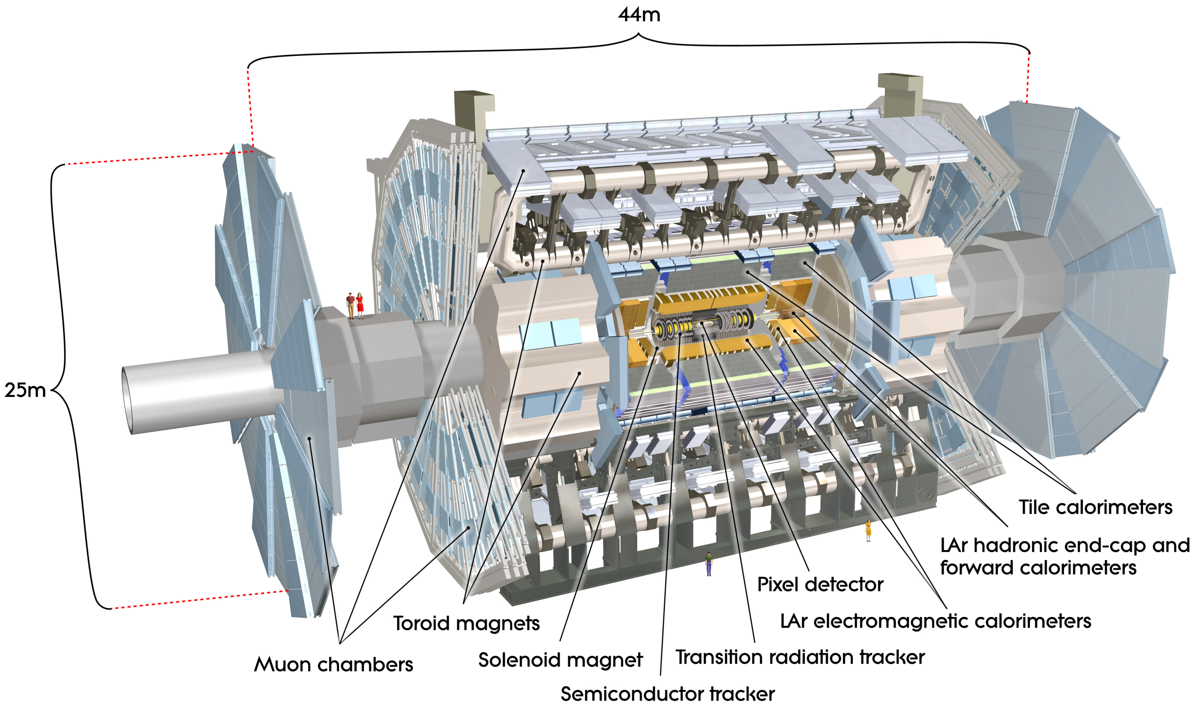


Figure 2.3: A schematic overview of the arrangement of the ATLAS detector, showing the main sub-detector systems [1].

vertical and the z-axis along the beam line. The azimuthal angle ϕ, and the *pseudo-rapidity*, $\eta = -\ln\tan\frac{\theta}{2}$ are more commonly used coordinates, along with the projection of the vector onto the transverse $x-y$ plane. For example, one may talk of a track along (ϕ, η) with a transverse momentum, $\mathbf{p_T}$. Because of the approximate[2]) azimuthal symmetry of the detector, usually the magnitude of the transverse momentum is of most interest. Similarly, because the detector is mostly[3] forward-backward symmetric, often only the absolute value of the pseudo-rapidity is quoted.

2.2.1 Inner Detector

The Inner Detector tracker is contained within a cylindrical envelope of length 7.024 m and radius 1.15 m, centred at the origin and within a solenoidal (that is, longitudinal) magnetic field of strength 2 T. It was designed to provide coverage over a pseudo-rapidity range of $|\eta| < 2.5$, with the primary task of pattern recognition for the reconstruction

---

[2] Some of the ATLAS detector systems are not quite azimuthally symmetric. Notable examples are the Pixel and SCT detectors, whose modules are tilted with respect to $\hat{\phi}$

[3] Again, this is not quite true for some detector systems. In particular, the magnetic field of the Muon Spectrometer means that it is more symmetric in the product $q\eta$, where q is the charge of the muon.

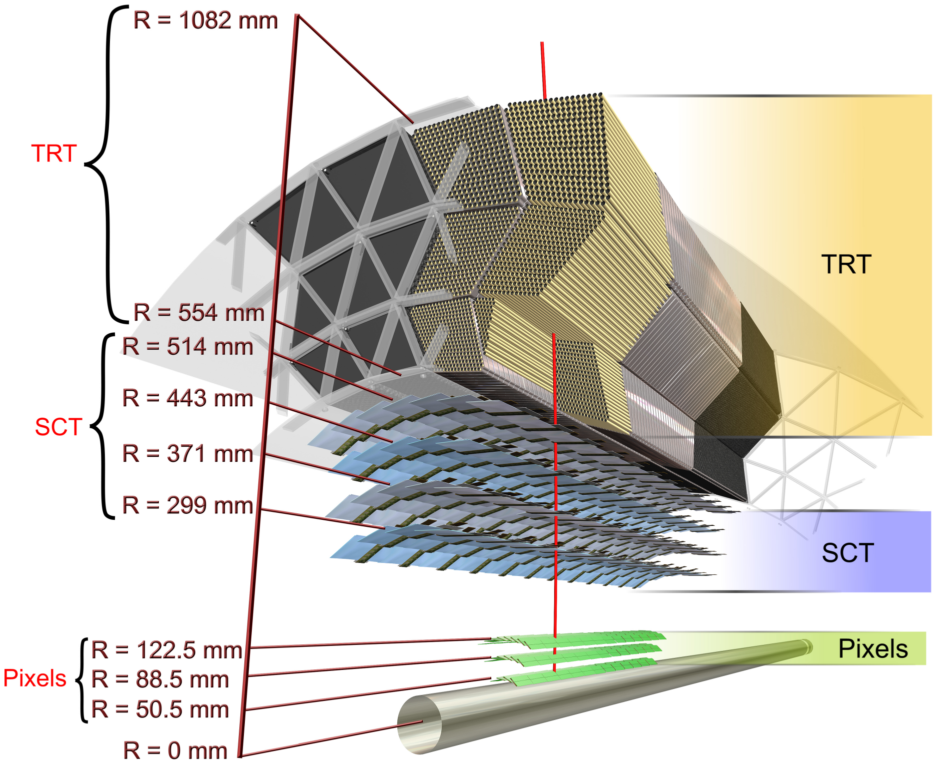


Figure 2.4: A computer-generated perspective on the barrel section of the Inner Detector. The red line shows an example trajectory for a $p_T = 10$ GeV charged track with $\eta = 0.3$ [1].

and measurement of particle tracks with transverse momentum $p_T > 100$ MeV, and the subsequent primary and secondary vertex measurements. An early study [81] using cosmic rays measured the relative momentum resolution of tracks reconstructed by the ID as ∼1-2% over a large momentum range. This is of crucial importance to the main study presented in this thesis — the resolution on the measurement of the track parameters of the pions and muons in the $\pi^+\pi^-\Upsilon(1S)$ final state determine the width of the observed mass peak, which, in turn, has a strong influence on the sensitivity to new states (see Chapters 5, 6 and 7).

The ID itself consists of three further detector subsystems — three layers of silicon-based pixel detectors, four layers of silicon strip detectors and a combined transition radiation and drift tube array (see Figure 2.4) — arranged in order of decreasing resolution and granularity.

2.2.1.1 Pixel Detector

As the closest detector system to the interaction point, the operational conditions for the Pixel Detector are very demanding. The sensor, electronics and mechanical supports were designed to withstand extreme particle fluxes and integrated radiation dose over a ten-year lifetime, with little or no degradation in measurement resolution. As with the other elements of the ID, this all needed to be achieved with as small as possible contribution to the material budget (the importance of this for electron reconstruction is discussed in Chapter 4). This being said, it was always planned that the innermost layer (B-layer) would be replaced after the first three years of running.

The fundamental unit of the Pixel Detector is the sensor; a 256±3 μm thick n-type silicon wafer with a p^+ backpane and n^+ pixel implants, operating fully-depleted under an initial negative bias voltage of 150 V (and up to 600 V later in their lifetime). This, rather than the classic $p-n-n^+$ type (used, for example, in the SCT), was chosen for its expected ability to maintain efficient operation after long-term particle irradiation. On each such sensor, the pixels are arranged in 144 columns and 328 rows, giving a total of 47232 pixels, with a typical[4] size of 50×400 μm^2. To reduce the effects of radiation, and minimise leakage current, the sensors are operated at a temperature of $-5°$ C to $-10°$ C.

A set of *front-end electronics* is attached to the pixel side of each wafer, along with a flexible circuit board on the backpane — together, this is referred to as a *module* of the Pixel Detector. Each n^+ pixel implant is coupled to a single readout cell of the front-end electronics. Charged particles traversing the bulk of the wafer create electrons and holes in the depleted region, which then drift towards the pixels and backpane, respectively. A signal is induced on pixels in the vicinity of the *cloud* of drifting electrons, which is then digitised into an effective amplitude, given by the time the signal was greater than a preset threshold (the *time-over-threshold* (ToT)).

In the barrel, the modules are arranged into three concentric cylinders with the long axis of the sensor parallel to the beam axis. The Lorentz angle for the barrel pixel sensors implies that the optimal incident angle, in terms of resolution, is 10-15°, but due to geometric constraints the modules are actually mounted at tilt angles of 20° (Figure 2.5). This also allows for slight overlap between the modules to ensure good azimuthal coverage. In each of the endcaps, the modules are affixed to both sides of three coaxial disks.

---

[4] Approximately 10% of the pixels are 50×600 μm^2.

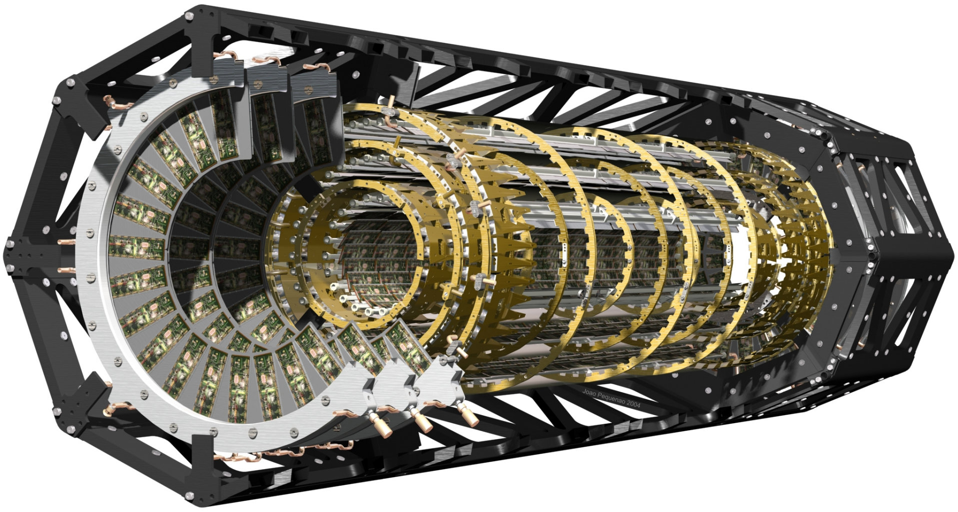


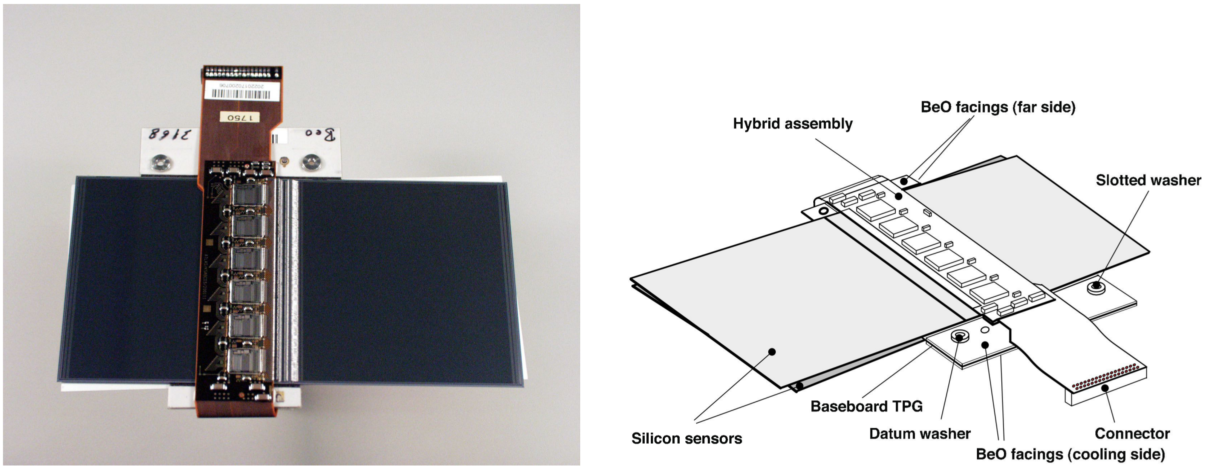
Figure 2.5: A cut-away view of the Pixel detector and support structures, showing the barrel cylinders, endcap disks and the attached modules [1].

The total number of pixel channels is approximately 80.4 million, allowing the detector system to meet its low occupancy and high-granularity demands, with initial test beam studies showing that the spatial resolution at normal incidence is 12 μm. All tracks produced at the origin and within the physical acceptance of the Pixel Detector, $|\eta| < 2.5$, are expected to pass through at least three layers of sensors.

2.2.1.2 Semi-Conductor Tracker (SCT)

The Semi-Conductor Tracker (SCT) is similar to the Pixel Detector, with some notable differences:

- most importantly, the read-out elements of the sensors and modules are *strips* rather than pixels.
- the SCT sensors are also the more standard $p-n-n^+$ type, chosen mainly because of their lower cost. This means that the induced signal is due to the flow of holes, rather than electrons, and that the Lorentz Angle is smaller and in the opposite direction. For this reason, the tilt angle for the SCT barrel modules is 11° (11.25 in the outer two layers) and in the opposite sense to the Pixel modules.
- the barrel modules are made up of four sensors arranged in back-to-back pairs; two pairs are glued either side of a support board, with the front-end electronics running roughly down the centre (see Figure 2.6). Each back-to-back pair is arranged at a



(a) (b)

Figure 2.6: A photograph of an actual barrel SCT module (left), alongside a diagrammatic representation (right) labelling the key elements [1].

stereo angle of ±20 mrad to allow for measurement of the hits in the direction of the strips. The strip pitch for the barrel modules is 80 μm.

- the endcap module sensors are trapezoidal to allow attached to the endcap wheels in a gap-less fashion. Their strips run radially, with a mean pitch of 80 μm.
- the read-out of each strip is binary — that is, hit or no hit (controlled by a tuneable threshold) but no time over threshold information.
- there are 4 barrel cylinders, and 9 endcap disks (of varying size). Each track with $|\eta| < 2.5$ should traverse at least eight sensor layers (i.e. four double layers).

The total number of channels in the SCT from the 2112 barrel and 1976 endcap modules is approximately 6.3 million, providing an intrinsic resolution of 17 μm $(r - \phi)$ × 580 μm (z) in the barrel and 17 μm $(r - \phi)$ × 580 μm (r) in the endcap.

The next chapter discusses a case study of a particular aspect of the calibration of the SCT, and includes further details on the read-out electronics and their integration with the level-1 trigger system.

2.2.1.3 Transition Radiation Tracker (TRT)

The outermost detector of the inner tracker, the TRT, is a departure from the silicon-based Pixel and SCT detectors. It consists of a large number of drift tubes, or *straws*, 4 mm in diameter and with an intrinsic resolution of 130 μm. In the barrel, the 144 cm long straws are aligned with the beam direction, with their wires split in the middle (i.e. at

$\eta = 0$) and are read out from the ends. The endcap TRT straws are shorter (37 cm), and attached radially to two sets of independent wheels in bundles, which are also referred to as *modules*. In both barrel and endcap regions, the space between the straws is filled with a *radiator* material to encourage the production of transition radiation (TR). TR is electromagnetic radiation emitted by charged particles traversing an inhomogeneous medium — for example, at the boundary between two materials with different dielectric constants. The amount of radiation produced is greater for particles with higher Lorentz factors (γ), making TR a useful tool for identifying particles of similar energies but different masses. In particular, electrons are expected to produce a much larger amount of TR than, say, pions (the dominant track type in the ID), meaning that the TRT signals from the former are typically stronger (in which case they are labelled as *high threshold* hits). This provides the TRT an additional capacity for particle identification.

The total number of channels in the TRT is approximately 351,000, with up to 73 layers in any one direction (from the origin). A typical charged particle traversing the TRT is expected to produce hits in at least 36 straws, except in the transition region between the barrel and endcap regions ($0.8 < |\eta| < 1.0$) where this number is decreased to 22. For electrons with a transverse momentum larger than 2 GeV, 7-9 of these are expected to be high-threshold hits. This large number of hits is one of the main reasons the TRT system was implemented at ATLAS; it provides an important input to the pattern recognition used for track reconstruction.

In contrast to the Pixel and SCT detectors, the coverage of the TRT only extends to $|\eta| = 2$, and there is no longitudinal (radial) information available from the hits in the barrel (endcap).

2.2.2 The Calorimeters

A calorimeter is a detector element intended primarily to absorb and measure the energy of incident particles (see Chapter 8 of [82] for a full description). The operation principle for electromagnetic calorimeters (used for electrons and photons) is based on electron-photon *cascades*. At energies of $E \sim 100$ MeV or higher, electrons passing through material will suffer energy losses almost exclusively through bremsstrahlung,[5] with a rate of energy loss described by

$$\frac{dE}{dx} = -\frac{E}{X_0}, \tag{2.1}$$

---

[5] Bremsstrahlung is electromagnetic radiation produced by the deceleration of a charged particle as it passes through a medium.

where x is the distance traversed and X_0 is called the *radiation length*. Photons at these energies are most likely to undergo electron-positron *pair production*, with a probability

$$\frac{dw}{dx} = \frac{7}{9X_0} e^{-x\frac{7}{9X_0}}. \tag{2.2}$$

For an incident high-energy photon or electron, the combination of these two processes leads to a *shower* of electrons, positrons and photons. After a certain number of iterations, the energy of the secondary particles drops below a critical value and absorption processes (ionisation, the photoelectric effect etc) begin to dominate. The detector signal is based on the ionisation due to the electrons and positrons participating throughout the shower development.

For hadronic calorimeters, the shower develops as follows; the incident particle interacts with the nucleons within the material via the strong force to produce secondary hadrons, which themselves go on to interact with other nuclei. The relevant unit of length here is the *interaction length*, which is analogous to the raditation length of electrons. At the end of the shower, the energy is released through the production of electrons and photons (e.g. from $\pi^0 \to \gamma\gamma$, β-decay and γ-decay of excited nuclei, fission) where it is absorbed to produce the detector signal.

The ATLAS Electromagnetic Calorimeter (ECAL) is a sampling calorimeter, with layers of lead (the absorbing material) interleaved with liquid argon (LAr) (the active material) in an accordion geometry (Figure 2.7), ensuring seamless coverage in ϕ. The folds of the layers lie within the $R-\phi$ plane in the barrel, and are projected (i.e. they *run*) longitudinally, whereas in the endcap the folds are in the $z-\phi$ plane and run radially. At the centre of each LAr strip are two high voltage (HV) copper sheets, containing a third copper sheet separated from the former two by insulating polyimide sheets. The electrons released by ionisation in the LAr move towards the HV sheets, with a total drift time of 450 ns, where their signal is extracted with a central read out strip using capacitive coupling. A complicated calibration process, involving shaping and amplification of the signal, is then performed to determine the amount of energy this signal represents.

In the barrel, the ECAL covers the pseudorapidity range $|\eta| < 1.475$ and is structurally two half-units joined at $\eta = 0$. A barrel *module* (shown in Figure 2.7) demonstrates the segmentation principle of the ECAL:

- in depth (that is, r), the module is split into an inner, middle and outer layer. Each electrode is etched to reflect this sectioning, with the inner layer read out at the front of the calorimeter and the middle and outer at the rear.

- the electrodes are likewise etched into η *towers* whose axis points towards the origin. This *pointing* allows for a better determination of the incident particle's trajectory. The granularity in η is greatest in the innermost layer to provide information on the shower position and shape, and decrease toward the outer layers.
- in ϕ, the electrodes are *ganged* together — that is, the signal is summed and analysed together.

In each of the endcaps, the ECAL consists of two coaxial wheels. The outermost of these covers the region $1.375 < |\eta| < 2.5$ (projectively aligned with the edge of the ID), with a similar level of granularity to the barrel modules, while the inner wheel extends from $\eta = 2.5$ to $\eta = 3.2$. The combined number of channels in the barrel and endcap for the ECAL is approximately 164,000.

The transition region between the barrel and endcap detectors is filled with cables and services for the ID and barrel calorimeter, amounting to several radiation lengths. To try to mitigate some of the issues this, and the other material in front of the barrel calorimeter, causes, a separate 11 mm *pre-sampler* layer of LAr is placed in the region $|\eta| < 1.8$ (despite this, most physics analyses using photons and electrons still place a veto on $1.37 < |\eta| < 1.52$).

In the most recent test beam study of a barrel module, the stochastic term[6] was measured as $10.1 \pm 0.4\% \times \sqrt{\text{GeV}}$ and the constant term as $0.2 \pm 0.1\%$, consistent with the results obtained from a GEANT4 simulation [1].

Outside of the ECAL is a dedicated hadronic calorimeter (HCAL), which is described in detail elsewhere [1]. Additionally, a combined electromagnetic (first layer) and hadronic (second and third layers) forward calorimeter (FCAL) services the very high rapidity region: $3.1 < |\eta| < 4.9$.

Together, the calorimeter systems have a secondary role of preventing the *punch-through* of any particles other than muons and invisibles (neutrinos etc.). Simulations have shown that the > 22 radiation lengths of the ECAL and approximately 9.7 interaction lengths (10 in the endcaps) of the HCAL are sufficient for this purpose.

2.2.3 Muon Spectrometer

The initial step in electromagnetic shower production, bremsstrahlung, is suppressed by a factor of $(m_e/m_\mu)^2 \approx 2.5 \times 10^{-5}$ for muons relative to the rate for electrons. To muons,

---

[6] The resolution of calorimeters is of the form $\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E[GeV]}} \bigoplus b$, where a is the *stochastic* term and b is the constant term (which is due to non-uniformities in the detector response).

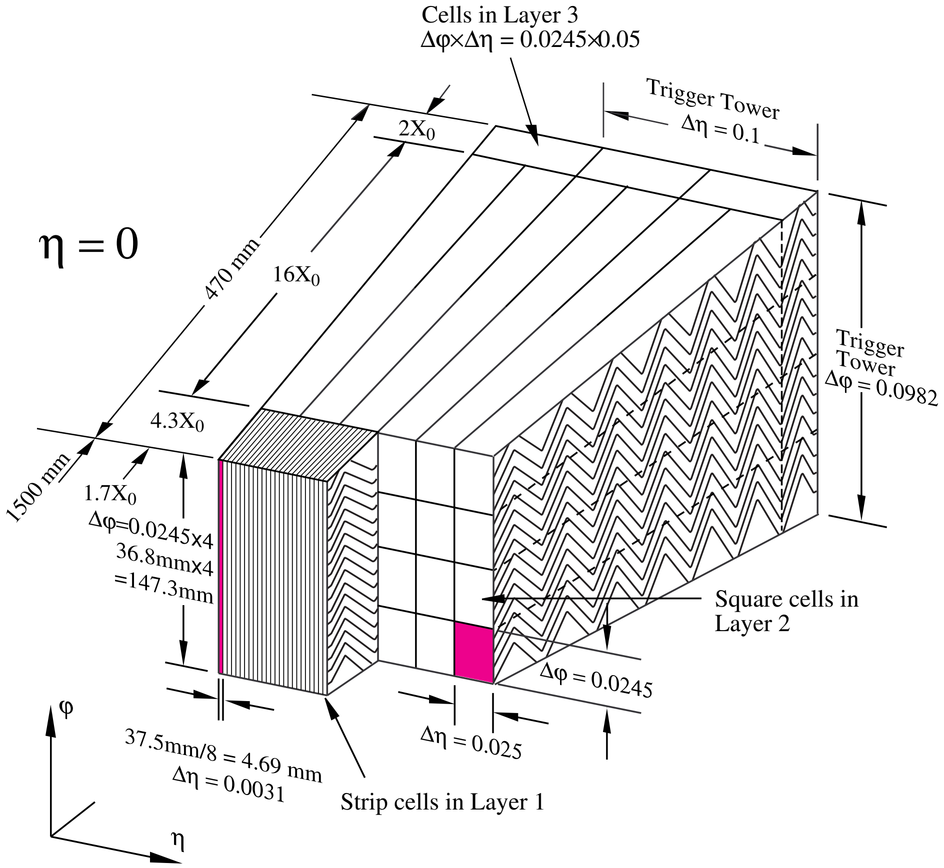
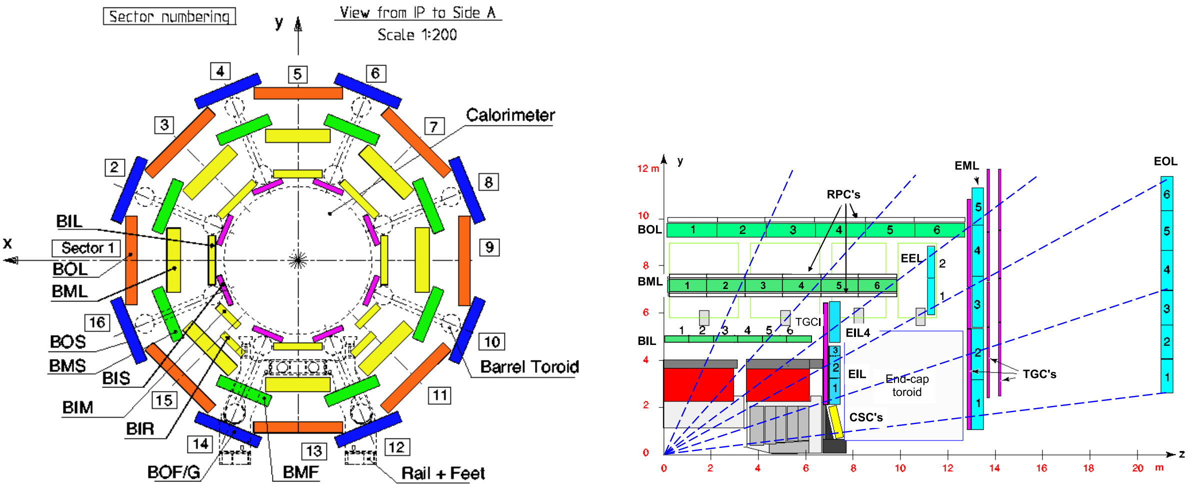


Figure 2.7: A sketch of a barrel module of the ECAL, demonstrating the accordion geometry and the segmentation in (r, ϕ, θ) [1].

then, the total amount of material in the calorimeters is effectively zero radiation lengths and they pass through with only small ionisation losses. This isolation makes muons an exceptionally clean object for particle detectors, which, at ATLAS, is exploited by a dedicated *Muon Spectrometer* (MS). The MS performs two complementary roles — triggering and tracking — using two separate specialised sub-detector systems.

For reasons of cost and practicality, a set of eight air-core toroids are used in the barrel and endcaps to produce a magnetic field in the radial direction over the large volume which encompasses the muon detector elements. The bending plane is therefore along the η direction in the MS, rather than in ϕ as in the ID.

Precision tracking is provided mostly by a set of *Monitored Drift Tubes* (MDTs) arranged into *chambers* — 3-4 layers of tubes mounted on either side of a rectangular prism-shaped support structure. These chambers are, themselves, organised into three cylindrical shells in a manner reflecting the octagonal ϕ symmetry of the toroid arrangement; between two neighbouring toroids is a large wedge of MDT chambers and within each toroid a small wedge, with the size of the chambers projective in ϕ and η (see Figure



(a) (b)

Figure 2.8: The arrangement of the MDT chambers in the $r-\phi$ plane of the barrel (left) and in the longitudinal plane (right). The naming scheme is as follows: the first letter refers to the barrel (B) or endcap (E), the middle letter to the inner (I), middle (M) or outer (O) layers, and the third to the large (L) and small (S) sector types. The right-hand figure also shows the position of the muon trigger elements and the CSCs [1].

2.8). In the endcaps, the chambers form disks, again made from small and large overlapping wedges. In both cases, the direction of the tubes is perpendicular to the straws of the TRT, because the bending plane for the MS is in η. This provides a measurement resolution in this direction of about 35 μm per chamber.

In the forward regions ($|\eta| > 2$) of the first layer of the endcap, the flux rate exceeds the limit for the MDTs. For this part of the MS, the MDTs were replaced by cathode-strip chambers (CSCs) (see Figure 2.8b), which can handle the higher rates and provide high spatial and time resolution. Their resolution is similar to the MDTs in the bending plane (40 μ m), and also provide a second transverse coordinate, albeit with much lower precision (about 5 mm). Together, the MDTs and CSCs provide momentum measurement for $|\eta| < 2.7$.

The ability to trigger upon muons of specified transverse momenta was an essential requirement of the MS. The drift time for the MDTs is of the order of 700 ns, making them unsuitable for triggering purposes. Instead, the precision tracking elements are accompanied by Resistive Plate Chambers (RPCs) in the barrel ($|\eta| < 1.05$) and Thin Gap Chambers (TGCs) in the endcaps ($1.05 < |\eta| < 2.4$) for their higher-rate capability. The three layers of RPCs are located either side of the middle cylindrical shell, and outside the last shell. A TGC wheel was placed in front of each of the first precision tracking endcap disks, another two sandwich the outer MDT wheel, and a final wheel lies outside of all of this. The reason for this non-uniform placement of trigger chambers is to accommodate low-p_T and high-p_T muons with the same set of detectors. In the RPCs, for example, the low-p_T muon trigger operates on coincidence between hits in the first and second layers only, while for high-p_T thresholds coincidence is also required with the final layer.

The coverage of the RPCs and TGCs extends to $\eta = 2.4$, and provides triggering over a large range of transverse momenta. The also serve a secondary role of providing a measurement of the ϕ coordinate of the muon tracks, complementing the η measurement from the MDTs.

2.2.4 The ATLAS Trigger System

Current computational and electronics technology is not capable of transmitting and processing the full event data at the nominal beam-crossing rate of 40 MHz. The interesting event types (e.g. Higgs boson production) are generally quite rare and are only produced in a small fraction of these collisions. ATLAS therefore implements a *trigger* system to select only events containing certain types of decay products, which makes the rate manageable. This process is performed in three stages — the hardware-based level-1 (L1) trigger, the intermediate level-2 (L2) trigger and the event filter (EF), with the latter two commonly referred to as the High-Level Trigger (HLT).

The L1 system includes all the components of the ECAL, HCAL and FCAL calorimeters as well as the trigger chambers (RPCs and TGCs) of the MS. The L1 calorimeter trigger is based on the integrated signal across *trigger towers*, which are typically $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ cells (see Figure 2.7). A *sliding window* algorithm is used to determine regions of the calorimeter containing deposits of (transverse) energy, which are then compared to preset thresholds. As explained in the previous section, the muon triggers are based on a coincidence of hits in the trigger chambers within a *road* along the path of a particle, whose width determines the transverse momentum threshold. The information from both of these is then sent to the *Central Trigger Processor*, which decides whether or not the event is to be discarded. Examples of L1 trigger selections may include the presence of an ECAL cluster with some minimum transverse energy, one or more muons with some transverse momenta, missing energy requirements, or τ/jet clusters. If the event is to be kept, then a *L1 accept* signal is sent to all of front-end electronics of the ATLAS sub-detectors, instructing them to transmit the data from this event to a set of readout systems, located in an adjacent room (USA15). The finite length of the memory buffers in the on-detector electronics sets the timing for these decisions to be within 2.5 μs from the bunch-crossing.

As well as the initial trigger decision, the L1 hardware also sends information on the (η, ϕ) location of the calorimeter clusters or muon hits, which are used to build *Regions of Interest* (RoI). At level 2 (L2), the event fragments within the RoIs (about 1-2% of the total event data) is read out in full granularity and processed by a computing farm of ∼500 nodes. At this point, the additional accuracy allows tighter thresholds on the muon momenta and calorimeter energies, but can also include selections based on tracks within the ID. This allows, for example, separation of electron and photon objects in the L2 trigger and explains why each L1 trigger can seed several L2 trigger *chains*.

If the decision is made to keep the event at L2, the complete set of data for the event is collected and sent to the *Event Filter*, which is another computing farm of some 1800 nodes. By this stage, the software used to reconstruct particles is almost identical to that used in the actual physics analyses (see the next section) and much more sophisticated requirements can be placed on the event. For example, in the triggers used for the X_b search of this thesis, a vertex fit is performed and the mass derived from this is used as a selection criterion (see below).

Events which pass L1,L2 and the EF stages of the ATLAS trigger are organised into data *streams* based on their physics content. For example, the data streams of interest in this thesis were the *egamma* stream (in Chapter 4), for events passing electron and/or photon triggers, and the *muons/bphysics* streams, which are fairly self-explanatory[7]. All of the detector information for these events is promptly sent to the CERN computing centre at a rate of ∼400 Hz for further processing. In 2012, a so-called *delayed* stream (which included the *bphysics* stream) was also sent directly to storage, with processing deferred until later in the year (and early in 2013).

The entire process, from signals in the calorimeters and the MS to the CERN computing centre, is shown in the flow diagram of Figure 2.9.

After L1, the event rate must be reduced to less than 75 kHz, and after L2 it must be less than 3.5 kHz. To help achieve this, a *prescale* can additionally be applied at each stage of the ATLAS trigger system; a prescale of N implies that only a fraction $1/N$ of events passing a particular trigger are used.

2.2.4.1 Triggers for Quarkonium in 2011 and 2012

When one talks of a particular ATLAS *trigger*, he or she is usually referring to an L1–L2–EF trigger *chain*; that is, a specification of the event selection requirements at each of the trigger stages. The collection of triggers operating at any point of time is said

---

[7] The *bphysics* stream encompasses any triggers intended for use in quarkonium analyses or studies involving B-mesons.

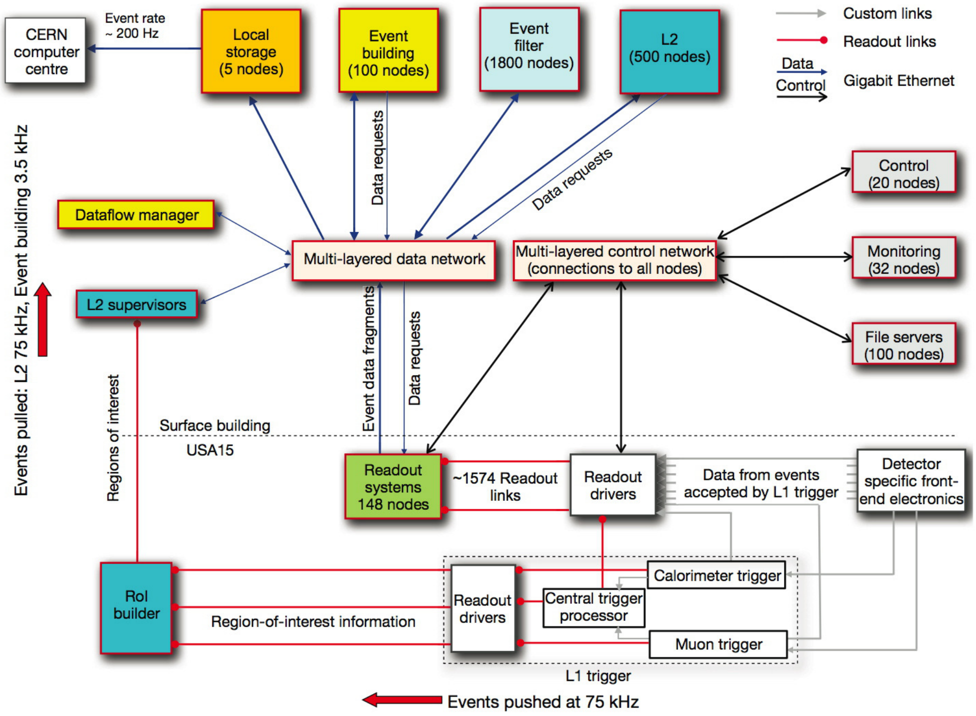


Figure 2.9: The triggering process — from detector signals in the calorimeters and muon trigger chambers to event transmission to the CERN computing centre [1].

to be the trigger *menu*. Each physics sub-group of the ATLAS collaboration receives a rate allotment of this menu, which is used mostly on their main (un-prescaled) analysis triggers. Additionally to this, there are also *supporting* triggers which might be used for efficiency or background studies, for example.

In early 2011, the main un-prescaled triggers for quarkonia focused on collecting large numbers of $J/\psi \rightarrow \mu^+\mu^-$ and $\Upsilon \rightarrow \mu^+\mu^-$ decays. At L1, only the presence of two muons was needed (i.e. L1_2MU0)[8] whereas in the EF the muons were required to have opposite sign and be successfully fitted to a common vertex, with a mass in a window around the J/ψ or Υ(1,2,3S) states. Due to the increased instantaneous luminosity later in the year, the transverse momentum threshold at L1 was increased to 4 GeV (L1_2MU4) to control the rate. The invariant mass distribution for muon pairs collect by these, and other, *bphysics* triggers is shown in Figure 2.10. The 2011 dataset alone represents millions of

---

[8] The number after the "MU" refers to the explicit transverse momentum threshold imposed on the muons. The actual response of the trigger has a turn-on shape centred slightly above this. In the case of the $MU0$ triggers, the effective threshold is about 3 GeV.

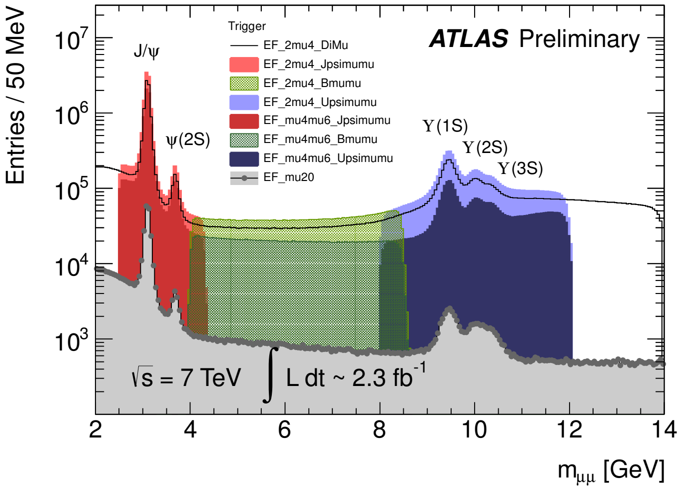


Figure 2.10: The invariant mass distribution for dimuon candidates selected by the main B-physics and quarkonia triggers in 2011 [11].

$J/\psi \to \mu^+\mu^-$ and $\Upsilon \to \mu^+\mu^-$, and was the basis for all spectroscopic studies performed by ATLAS using this data.

In 2012 the change in the centre-of-mass energy and the instantaneous luminosity resulted in a further increase in the muon thresholds of the primary quarkonia triggers to $p_T > 6$ GeV. Other supporting triggers with asymmetric 4,6 GeV thresholds, or barrel-only muons, were used as part of a delayed stream.

The 2011 trigger menu also included some supporting $J/\psi \to e^+e^-$ triggers, intended to be used for low-p_T electron efficiency calculations, but, in this thesis, were used to perform the J/ψ production cross-section measurement in Chapter 4.

2.3 Particle Reconstruction and Identification

For events passing a trigger, the digitised electronic signals from all of the ATLAS detector's sensitive elements (i.e. the *hits*) are transferred to the CERN computing Centre to be processed offline[9]. The purpose of this processing is to *reconstruct* and *identify* the particles produced in the collision(s) for each such event, and is performed within a C++ package-based framework referred to as ATHENA [83]. The resultant data structures retain their stream separation (see above) and are available to be further analysed by individual users or groups.

Reconstruction of the event is carried out in a step-wise fashion, starting with track fitting in the ID/MS and cluster formation in the calorimeters. These basic objects can

---

[9] As opposed to the *online* analysis of the EF trigger stage.

then be used to construct more complex structures which require information from several sub-detectors, such as electron or muon candidates, dimuon decays (e.g. $J/\psi \to \mu^+\mu^-$ or $\Upsilon(1S) \to \mu^+\mu^-$), jets, missing energy and so on. A number of selection criteria on the combined quantities can be applied along the way to attempt to identify them as particular physics objects. For example, electrons are required to have a track pointing at a cluster in the ECAL within some directional tolerance.

The aspects of particle reconstruction and identification relevant to quarkonia decays are briefly summarised below. As always, the full details can be found in the corresponding references.

2.3.1 Track Reconstruction

The task of recovering tracks left by charged particles from hits in detectors is classically separated into two stages; *pattern recognition*, where hits are grouped into track candidates, and *track fitting*, where the hits belonging to a particular track are used to determine the best estimate of the underlying particle momentum and perigee. In high track multiplicity environments such as ATLAS, however, there is often considerable interplay between the two.

Prior to pattern recognition, the raw data from the detector elements is transformed into three-dimensional *space points* (SPs) or, in the case of the TRT, *drift circles*[10] in a step referred to as *data preparation.* For the Pixel, this is fairly trivial; the centre-of-gravity of groups or *clusters* of adjacent pixels defines local 2-dimensional coordinates on the module, which can then be transformed into a global 3-dimensional SP by using the known position of the module within the ATLAS coordinate system. Corrections for the Lorentz angle and time-over-threshold for each cluster are incorporated into this process. The analogous process in the SCT is slightly more involved, and takes into account the double-sided nature of the modules.

The default ID pattern recognition algorithm is called *New Tracking*, or NEWT [84], though others are also available. NEWT is seeded by SPs in the three pixel layers and the first layer of the SCT, which are then extended through the remaining three SCT layers to form track candidates. These are then fitted, outliers are removed, and any ambiguities in the track–cluster association resolved. Quality selections on the number of hits, holes (detector elements traversed by the fitted track which do not have a matching hit), and number of shared hits of these silicon track segments are applied to decrease the fake rate.

---

[10] The straws of the TRT only provide information on the drift time and hence distance of the traversing particle from the central wire. As such, the resultant measurement is a radius (i.e. a *drift circle*) rather than a space point.

The remaining tracks are further extrapolated into the TRT and associated to the drift circles, followed by a refit of these extended using all of the measurement information available from the three detector systems.

The track *fitting* can be performed using several different techniques, the simplest of which is the *Least-Squares Method* (LSM). The quantities of interest from the tracks are generally the initial momentum and position (or other equivalent variables, collected into the initial state vector $\mathbf{p_0}$) of the charged particle at its production vertex. In the LSM, these variables are used to perform a (linear) extrapolation of the track through the layers of detector elements, taking into account the local magnetic field. The residuals between these predictions and each measurement, weighted by the associated measurement covariance matrix, are used to construct a χ^2 variable which is subsequently minimised over $\mathbf{p_0}$. While the LSM produces track parameters with desirable statistical properties (e.g. they are unbiased and have minimum variance), it involves the inversion of an $n \times n$ matrix, where n is the number of measurements made.

To reduce demands on CPU-time, a recursive track-fitting technique, based on the Kalman-filter (KF) approach, is used instead of the LSM [85]. The first two stages of the KF involve extrapolating the state vector, $\mathbf{p}$, from one detector layer to the next, and combining it with the measurement at that layer to give an updated (*filtered*) state vector and associated covariance matrix. The final step, *smoothing*, is essentially the KF run in reverse to ensure that the initial state vector takes into account all of the measurements. The key property of the KF is that it only requires the inversion of matrices of the same order as the state vector, and hence is much faster in a high n environment.

Both the KF and LSM methods assume that the measurement noise is Gaussian. For electrons, bremsstrahlung is the dominant energy loss mechanism, characterised by the Bethe-Heitler functional form (an example distribution and more complete explanation of this is given in Appendix A.1.3 of [85]). Because this is significantly different to the Gaussian shape, track reconstruction for low-p_T electrons by these fitting techniques tends to be far poorer than for other particles. The Gaussian Sum Filter (GSF) algorithm, a generalisation of the KF using a mixture of different Gaussian components to model the noise, provides large improvements for electrons and, from mid 2011, became the standard track fitter used for electrons. See [86] for technical details on the implementation of the GSF algorithm.

Additional to the KF and LSM approaches, there is a complementary method, *back-tracking*, which searches for unused track segments in the TRT and extrapolate them back through the silicon layers. This helps to improve the track reconstruction efficiency and recover tracks not originating from the origin (such as those from $\gamma \to e^+e^-$ conversions).

The final stage of tracking is the reconstruction of the primary decay vertices using the fitted tracks, as well as dedicated algorithms which reconstruct secondary and conversion vertices.

2.3.2 Muon Track Reconstruction and Identification

Muons are identified by their presence in the Muon Spectrometer — either as a full track, passing through several layers of the MS, or a segment within a particular precision tracking chamber. In contrast to the ID, there are two competing MS track reconstruction algorithms in widespread use:

- ***Muonboy*** starts by identifying *Regions of Activity* (RoA) in the RPC/TGC systems. In the precision tracking chambers adjacent to these RoAs, track segments are created and combined together to form track candidates. A global fit is then performed on all the measurements associated to a given track candidate.

- ***Moore*** is a very similar, but independent, pattern recognition and track fitting procedure.

Tracks reconstructed by Muonboy or Moore alone are referred to as *Standalone* muons. If these originated from the primary vertex, they should also have a corresponding ID track. The matching and combination of ID and MS tracks or segments is also carried out with a number of different tools:

- ***STAtistical COmbination (STACO)*** combines the track parameters of an ID track and a matched MuonBoy track using their best-fit values and associated covariance matrices. The result is referred to as a (STACO) *combined* muon.

- ***MuTag*** identifies muons by associating ID tracks to segments in the MS reconstructed by MuonBoy. Only segments which do not belonging to a combined STACO muon track are used. The result is a *tagged* muon.

- ***Muid*** is an algorithm which matches ID tracks to Moore tracks, and performs a global refit to all of the measurements of each such pairing (that is, from the ID and MS detector elements). These objects are also referred to as a *combined* (Muid) muon.

- ***MuGirl*** searches for segments and tracks in the MS by using ID tracks as the seeds. If MuGirl manages to find a full track in the MS, a global refit is performed to produce a *combined muon.* If it finds only a segment, then a *tagged muon* is created.

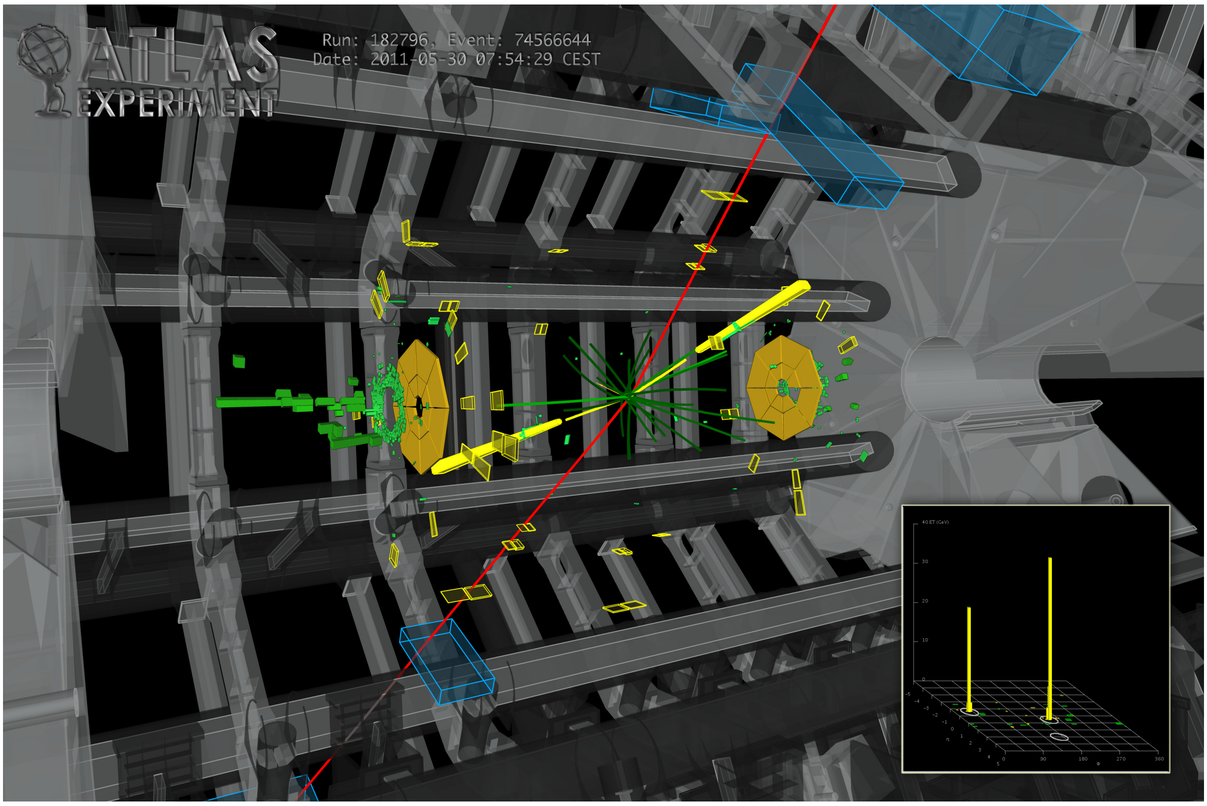


Figure 2.11: An example event display for a candidate $H \rightarrow 2\mu 2e$ decay used in the search for the Standard Model Higgs boson [12].

The chain {Muonboy segments/tracks, STACO combined muons, MuTag tagged muons} is collectively referred to as *STACO muons*, while the remainders are grouped as *Muid muons*. A hybrid of the two chains, *Muons*, was later developed; the segment-finding stage of track reconstruction is based on the MuonBoy algorithm, while the track builder is most like Moore. Combined, tagged and standalone muons are then constructed in an analogous manner to above.

The muon objects described above are demonstrated in Figure 2.11, the event display from a candidate $H \rightarrow 2\mu 2e$ decay used in the search for the Standard Model Higgs boson [12]. Two reconstructed muons (red lines) emerge from the primary vertex and produce a track in the ID, then deposit a small amount of energy in the calorimeter (yellow cells) before passing through several precision tracking chambers in the MS. These are hence both *combined* muons.

2.3.2.1 Muon Combined Performance (MCP) Recommendations

The Muon Combined Performance (MCP) group recommends the following selections on the associated ID track of any of the muon types defined above:[11]

- A B-layer pixel hit,
- ≥ 2 pixel hits,
- ≥ 6 SCT hits,
- Less than 3 silicon (SCT+Pixel) holes
- A *successful* TRT extension, when in the acceptance of the TRT detector ($|\eta| < 2$). Defining the number of TRT hits as n_{hits}, the number of outlier hits of the muon track as $n_{outliers}$, and their addition as n, the TRT extension is successful if:
  - $|\eta| < 1.9$, $n > 5$ and $n_{outliers} < 0.9n$, or
  - $|\eta| > 1.9$, $n \leq 5$ or $n_{outliers} < 0.9n$.

Each of these is applied in a manner which is safe against detector malfunction:

- a B-layer hit is only required when it is expected — that is, when the track passes through a fully operational pixel module in the B-layer,
- likewise, any time the track traverses a pixel or SCT module whose sensor is, for some reason or other, *dead*, it counts as a hit in that layer. So, the second requirement could be read as *number of pixel hits plus number of dead pixel sensors along the track greater than 1*, for example.

For the 2012 data-taking, the MCP recommendations were loosened to recover efficiency losses. The minimum number of pixel and SCT hits was changed to 1 and 5, respectively, and the second of the two TRT criteria also applied to $|\eta| < 0.1$, rather than the first.

---

[11] Obviously, standalone muons are an exception.

2.3.3 Electron Reconstruction and Identification

Within the tracking region, $|\eta| < 2.5$, electron reconstruction begins with energy deposits in the ECAL. A *sliding window* algorithm, using 3×5 cells of size $0.025\eta \times 0.025\phi$ (corresponding to the granularity of the second ECAL layer — see Figure 2.7), is used to find *pre-clusters* with at least 2.5 GeV of transverse energy. Tracks reconstructed in the ID are then extrapolated to the middle layer of the ECAL, and matched to the closest pre-clusters if the distance between them satisfies $\Delta\eta < 0.05$ and $-0.1 < q\Delta\phi < 0.05$. An electron is reconstructed if there is at least one track to the pre-cluster of interest (if not, the cluster automatically becomes a photon candidate). The same $H \to 2\mu 2e$ decay shown in Figure 2.11 above also provides a visual representation of electron reconstruction — the two yellow tracks pointing at two separate clusters in the ECAL (also see the inset figure, which gives a $\eta - \phi$ map of energy deposits in the ECAL).

If a particular pre-cluster is associated to several tracks, those with silicon hits are preferred and the best match in $\Delta R := \sqrt{\Delta\eta^2 + \Delta\phi^2}$ is chosen. The pre-cluster is then rebuilt using 3×7 (5×5) towers in the barrel (endcaps) to account for the larger lateral shower shape of electrons compared to photons. The energy of these *clusters* is calculated as the addition of four terms:

1. the estimated energy deposited in front of the ECAL,
2. the energy deposited in the cluster,
3. the lateral leakage (energy deposited outside of the cluster),
4. the longitudinal leakage (energy deposited beyond the ECAL),

with each parameterised in terms of the measured energy deposited in the pre-sampler and three layers of the cluster, based on Monte Carlo (MC) detector simulations.

The four-vectors for these electron candidates can be calculated from the track and cluster information in several ways. The recommended method, referred to as the *uncombined* four-vector, uses the directional information of the tracks at the primary vertex (i.e. η and ϕ) and the measured energy of the cluster. If the track does not have an adequate number of silicon hits (defined as less than 4), the direction is taken from the cluster, instead.

In the forward region not covered by the tracker, $|\eta| > 2.5$, the FCAL can still be used to reconstruct electrons using so-called *topological* clusters. Rather than the fixed size used in the sliding window algorithm, topological clusters are constructed by grouping together neighbouring cells with large signal-to-noise ratio [87]. The energy is calculated by summing the cluster measurements and taking into account the material in front of the detector. If the hadronic leakage is small and the transverse energy is > 5 GeV, the resultant object is called a *forward electron.*

Of course, the above electron candidates will include a reasonably large number of *fakes* — i.e. cases where these detector signatures were not caused by an electron. To increase the purity of real electrons, a number of discriminating, or *identification*, variables from the track and cluster are used. These include the following

- **Cluster variables**: hadronic leakage, lateral shower width in the strip and middle layers, ratios of energy deposited in certain cells of the cluster.
- **Track variables**: number of B-layer, pixel, pixel+SCT, TRT, high-threshold TRT hits, and transverse impact parameter.
- **Track/Cluster combined variables**: matching in $\Delta\eta$, $\Delta\phi$, E/p (energy-momentum ratio).

In version 16.6.X of ATHENA (used for the majority of analyses of the 2011 dataset), the *egamma* group provided three working-points for electron identification — *loose*, *medium* and *tight* — based on these variables, which were optimised in 10 bins of η and 11 bins of E_T. Medium is built upon the loose selections, and tight upon the medium selections (i.e. they are inclusive). For forward electrons, there are analogous loose and tight identifications available. Further details on electron reconstruction and identification can be found in [88].

In mid 2011, the identification menu was re-optimised to include '++' options — i.e. loose++, medium++, and tight++. These are basically a tightening of the specific cut values used in the regular menu, designed to reduce the trigger rate at the event filter stage. The '++' menu was further revised for 2012 data taking by shifting the emphasis of the selections away from pile-up sensitive variables. The original '++' menu was available from ATHENA v. 17.0.0 onwards, and both 2011 and 2012 '++' menus from 17.2.X onwards.

Chapter 3

Excursus I: Timing Calibration of the SCT

The last chapter described the elements of the ATLAS detector in a mostly idealised context, with little mention of the actual implementation – i.e. the structural and electrical engineering aspects, as well as the provision of other services (e.g. the cryogenic cooling systems). In reality, a great emphasis was placed on understanding and calibrating the detector sub-systems within this setting over the first few years of Run I, and will be ongoing for the lifetime of the detector.

In this chapter, one particular study related to the timing calibration of the SCT is presented. In this case, the time-of-flight of particles from the interaction point to each individual module and the length of the read-out cables to the underground service caverns, adjacent to ATLAS, are the main factors affecting the optimal timing settings.

This work was conducted as part of the author's ATLAS qualification project in 2010, and is presented below in the format of an ATLAS *Internal Note* (ATL-COM-INDET-2011-020). Subsequent to this note being written, the study was extended to include data collected in 2011 – this is described at the end of the chapter.

Draft version x.y

ATLAS NOTE

May 31, 2011

SCT Cluster Size Variation With Respect to Clock Phase Offsets

Cameron Cuthbert[a], Aldo Saavedra[a], Takahiko Kondo[b]

[a]*School of Physics, University of Sydney*
[b]*High Energy Accelerator Research Organisation (KEK)*

Abstract

The SCT timing scan conducted during run 152777 allowed for the measurement of mean cluster size variations as a function of the arrival time of the clock signal. The results indicate that each individual module has an approximately sinusoidal dependency with a period equal to that of the clock (40MHz) and an amplitude of ~0.1 in units of SCT strips. A hypothesis was proposed to explain this in terms of the operation of the SCT in latching signals to the clock, and was tested both with real data and through simulations. The findings provided strong direct and indirect evidence of its validity. An alternative set of SCT module optimum time settings were obtained based on the maximum cluster size and were found to be systematically ~3 ns earlier than those of the original scheme.

1 Introduction

The Semi-Conductor Tracker (SCT) is comprised of silicon strip detector modules, located in the barrel and endcap regions of the ATLAS detector between the pixel and Transition Radiation Tracker (TRT) subdetectors. These three components, collectively referred to as the Inner Detector (ID), lie within a 2 T magnetic field aligned parallel to the beam line. The ID is responsible for tracking the trajectory (which is a helix under the influence of the **B** field) of charged particles, providing both momentum measurements and vertex reconstruction. The SCT is designed to provide at least 8 high-precision *hits* (4 space-points) over a range of $\eta \leq |2.5|$ and is therefore an integral part of the operation of the ID [1].

There are many quantities (for example error rates, efficiencies, noise occupancy, hit patterns etc) which are used to probe the operation and performance of the SCT, and characterise the hits being produced. A *cluster* refers to a collection of hits in adjacent strips (channels), with the *cluster size* simply defined as the number of strips therein. The size of a cluster helps to reveal features such as the incident angle of the track producing it and, if >1, provides good evidence that it was not a product of noise. The relationship between cluster size and incident angle is also commonly used to measure the Lorentz Angle[1] [2].

As charged particles traverse the silicon of the SCT they produce electron-hole pairs. The *signal* induced by the movement of these charges is detected, processed and compared to a nominal threshold by the front end electronics connected to each strip. Upon receiving a level 1 (L1) trigger, this signal is then read out and transmitted upstream to the higher level DAQ [3]. This process strongly relies on the arrival time of the combined clock/L1 Trigger signal, which controls when the output from the discriminator is registered by the detector and which bunch crossing it is assigned to. It is therefore very important that this be properly tuned. Optimum arrival times will vary between modules, due to differences in both the time of flight (TOF) of particles emerging from the interaction point and lengths of optical fibres joining them to the ATLAS DAQ. In April, 2010, a timing scan was performed in order to determine these times for each individual SCT module. The scan involved (globally) delaying the timing information being transmitted to the SCT, and calibrating each of the modules by maximising the ratio of in-time hit patterns. Interestingly, the mean cluster size was observed to be sensitive in a periodic manner to the time delays being applied. Features such as this are important to investigate, as they can serve to either reinforce the current understanding of the influence of timing on the SCT or reveal inadequacies with it. This is particularly relevant to the modelling and simulation of the SCT used in the production of ATLAS Monte Carlo data sets.

In the following section, the operation of the SCT in terms of signal collection and manipulation is reviewed, along with a discussion of the timing control of individual modules and the timing calibrations performed prior to the scan. Then we describe how the timing scan was performed and present the initial observation of mean cluster size variation. A hypothesis explaining this fluctuation will then be proposed, with some results from specific tests designed to assess its validity. Given this hypothesis, a new technique for optimising the module timing settings will be compared to the standard approach for doing so. This will be followed with a discussion of the meaning of these results, their implications, and some suggested future research directions. Lastly, some concluding remarks will be made.

2 The ATLAS Semi-Conductor Tracker

The SCT is a tracking detector which occupies the volume between the pixel and TRT subdetectors, in both the barrel ($|\eta| \leq 1.6$) and the endcap ($1.2 \leq |\eta| \leq 2.5$) regions. The basic units of the SCT are the modules, which are arranged into four coaxial cylindrical layers in the barrel and nine coaxial disks in

---

[1] The Lorentz Angle is defined as the angle which the holes drift under the influence of the combined **E** and **B** fields (with respect to $\hat{\mathbf{E}}$) in the silicon active material, and is approximately 4.2 degrees for the SCT.

Table 1: Conventionally, integral coordinates are used to uniquely assign coordinates to each module of the SCT. This includes a label for the disk or layer and a 'ϕ' and 'η' label. It is assumed that the region (barrel, endcap A or C) is to be provided along with these numbers.

	Barrel				Endcap	
Layer/Disk	0	1	2	3	[0,8]	
η	[-6,-1],[1,6]				0	[1,2]
ϕ	[0,31]	[0,39]	[0,47]	[0,55]	[0,51]	[0,39]

each of the endcaps. The barrel section is 12 modules deep longitudinally, and so each module is conventionally labelled by an integer 'η' index in the ranges [-6,-1] and [1,6]. Increasing from inner to outer layer, there are 32, 40, 48 and 56 modules around in the ϕ direction, providing a second 'coordinate'[2]. In the endcaps, each of the disks contains up to three concentric wheels, the inner two with 40 modules in ϕ and the outer with 52 [1]. In this case, the ϕ label is defined as in the barrel but η labels the inner, middle and outer wheels with 2, 1 and 0 respectively. Combining these ϕ and η numbers with a layer (0–3) or disk (0–8) index provides a unique set of coordinates for every module of a given region (Barrel, Endcap A or C) the SCT. These coordinate assignments will appear in plots presented in subsequent sections and are summarised in table 1.

The modules of the barrel are slightly different to those in the endcaps. We will describe the barrel module here, most of which applies equally well to all three different types of endcap modules. Each barrel module has two sides (each composed of two sensors), back to back and offset by stereo angle of 40 mrad with respect to each other, to provide resolution in the axial direction. All the on-detector electronics are located on a *hybrid* which, in the case of the barrel module, wraps around the two sides. Each side of the hybrid is populated with six ABCD3TA readout and control chips, each possessing 128 channels, resulting in a total of 1536 strips per module [1].

Charged particles traversing the depletion zone of silicon in the SCT wafers produce large numbers of free electrons and holes, which drift under the bias potential toward the strips and backplane, respectively. The motion of these particles induces a current in the strips, which is integrated and amplified by a trans-impedance preamplifier. Following this, the amplified signal is shaped and processed by a discriminator set to some tunable threshold voltage V_{Th} (nominally corresponding to 1 fC). The output from the discriminator is binary – signals above the threshold produce a '1' and those below the threshold a '0'. This output is then latched to a 40 MHz clock by the *input register* in one of two ways. In *edge sensing* mode, the input register detects a high to low change in the discriminator output, and produces a '1' of well-defined width for the bunch crossing in which the transition occurred. In *level sensing* mode, the input register produces a '1' for each coincidence of the discriminator output and the rising edge of the clock pulse. The edge sensing circuit is desirable in terms of occupancy (only a single '1' is produced per pulse) and its ability to detect every signal, however it leads to an increase in noise with respect to the level sensing mode [4]. Currently the SCT is running in level mode, but may eventually change to the edge sensing mode.

The preamplifier, shaper and discriminator are collectively referred to as the front end electronics of the SCT. Following this stage of processing, the signal produced from the input register is placed into a FIFO binary pipeline 128 channels across and 132 memory cells deep, which services an entire ASIC chip. In the pipeline, the entries are shifted along every 25 ns so that the original signal appears at the end after 132 clock cycles. If an L1 trigger signal is received, the last three cells of each channel (representing the triggered, previous and following time bins) are transferred to the readout buffer. At this point the output from each channel can be suppressed based on its hit pattern (i.e. XXX) through

---

[2] The ϕ and η labels described here are for module assignments, and do not correspond to the usual definitions of these quantities.

a configurable criterion. For example, *hit mode* accepts any hit pattern XXX with a ‘1’ in at least one time bin and *edge compression mode*[3] only accepts 01X hits. The content of the readout buffer for each chip on a single module side is then combined into a bit-stream, which is converted to an optical signal and transmitted via the Back of Crate (BOC) card to the higher level ATLAS DAQ [3]. This process is schematically shown in figure 1.

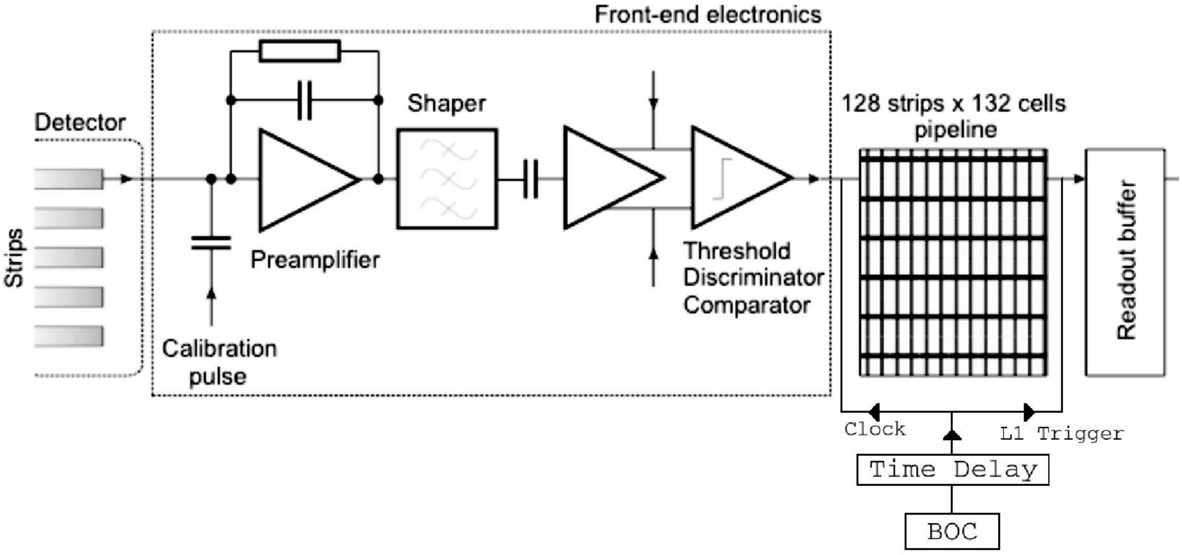


Figure 1: A schematic diagram showing how a signal from the silicon is processed, compared to a voltage threshold and transmitted from the readout buffer to the higher level ATLAS DAQ (via a Back of Crate card) [5].

2.1 Timing Control of the SCT

Timing influences the operation of each SCT module at two points (as shown in figure 1). Firstly, it sets the phase of the 40 MHz clock arriving at the input register. In level mode, this is critical as it determines when the output from the discriminator is sampled by the input register. At the edge of a signal, a small shift in the phase of the clock can mean the difference between whether or not a hit is recorded. This aspect of timing is obviously not so important in the edge sensing mode (which simply looks for a high-to-low transition). Secondly, the arrival time of the L1 trigger pulse determines which three time bins are read from the pipeline into the readout buffer. Since hits take 132 clock cycles to progress through the pipeline, the phase difference between the clock and the L1 trigger should reflect this.

Timing and trigger information is sent to each of the modules along the following pathway. Firstly, the global ATLAS timing, trigger and control (TTC) information is sent to several Timing Interface Modules (TIMs) attached to different VME crates in underground service cavern 15 (US15). Each crate can house up to 16 Back of Crate(BOC)/Read Out Driver (ROD) pairs, and each of these is responsible for the control and read out of 48 modules. From the TIM, the clock signal is sent directly to a BOC card whilst the remaining information (including the L1 trigger) is routed through the corresponding ROD to the BOC. From the BOC, the clock and trigger are encoded into a single combined signal using the bi-phase marking (BPM) system [6]. This signal, like the data transmission *from* the modules, is converted to an optical form and sent along fibres to each of the modules. There are several opportunities to offset the timing in this process. At the TIM level, the L1 trigger can be delayed in steps of 1 clock cycle (25

---

[3] This is not to be confused with edge-sensing mode.

ns) and the entire TTC signal in 104 ps steps in the range [0, 25 ns]. These are both global settings which are transmitted to all the modules under the control of the crate to which the TIM resides. There are also coarse (32 steps of 25 ns) and fine (128 steps of 280 ps) settings available in the BOC at the module level, which shift the BPM signal [3].

2.2 Current Timing Settings of the SCT Modules

Each module of the SCT needs to be *individually* timed-in with the rest of ATLAS, due to differences amongst the modules in (1) optical fibre lengths and (2) time of flight (TOF) for particles emerging from the interaction point. This is the reason for the availability of module-level time adjustments.

Before the April 2010 timing scan was performed, the relative optimum time settings between modules was estimated using their known individual fibre lengths. During the cosmic ray data taking period the overall timing of all modules was adjusted to a point of maximum occupancy of hits from tracks, and then a finer offset was applied in an attempt to maximise entries in the central time bin. Prior to collisions, the TOF differences between modules were calculated and incorporated into the timing settings [3].

These adjustments were made in order to have the timing reasonably accurate prior to first collisions. The purpose of the timing scan performed in April 2010, 2010, was to further fine tune these settings to their final optimum values. The methodology (see section 7 for more details) was to maximise the ratio of clusters with all of its strips having the hit pattern 01X, which are those expected to be produced by particles from the interaction point (i.e. as opposed to noise), to all other clusters with their strips in a single hit pattern. The following section will describe how this scan was performed.

3 The Timing Scan

On the morning of the 10^{th} of April, 2010, an SCT timing scan was performed during run 152777 in level mode detection and hit mode compression. The timing offsets were achieved by carrying out the following steps:

1. Shift the L1 trigger signal forward 25 ns.
2. Delay the global clock/L1 trigger signal in steps of 5 ns from 5-20 ns.
3. Reset the L1 trigger to its original setting and delay the global clock/L1 trigger signal in steps of 5 ns from 5-20 ns.

The central time bin being read out is always that which is 132 clock cycles ahead of the L1 trigger, so it is easiest just to consider the shift in the clock with respect to this point. Then, if we call t_0 the original time of the rising edge of that clock cycle, the first step set $t = t_0 - 25$. The second step then produced the following timing settings [t_0-20 ns, t_0-15 ns, t_0-10 ns, t_0-5 ns], whilst the final step produced [t_0+5 ns, t_0+10 ns, t_0+15 ns, t_0+20 ns]. The timing scan therefore produced time delays of [-20 ns, 20 ns] relative to t_0 in 5 ns steps (we include zero here as it is the setting before and after the scan). This is summarised in terms of event number in table 2.

4 Observation of Mean Cluster Size Variation

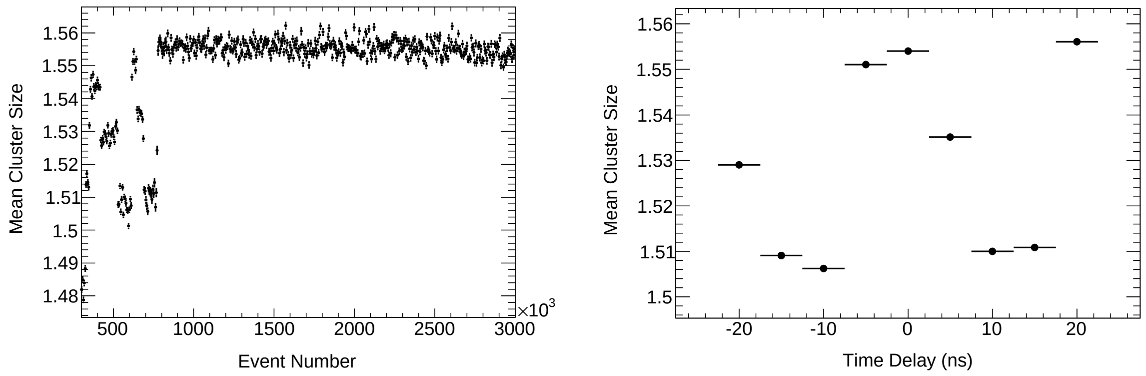
The results presented in this report were obtained from the Minimum Bias trigger stream of the experimental run corresponding to the SCT timing scan. In our study, we only select clusters associated to (on) tracks to eliminate noise clusters, whose features and presence are expected to be time independent.

During the scan, as expected, the mean hit pattern shifted from 001 to $1XX$ and the fraction of $01X$ (in-time) clusters peaked at specific times for each module. However, a variation in the mean cluster size

Table 2: The SCT time delays in run 152777 on April 10^{th}, 2010, showing the event numbers corresponding to specific delays and the total number of Minimum Bias events captured for each given delay. Note that events outside this range of values (~ 2.6 x 10^6 events) have ‘delays’ of zero.

Time Delay (ns)	**Event Range (x10^3)**	**Total Number of Events (x10^3)**
-20	430–525	95
-15	528–575	47
-10	580–613	33
-5	618–645	27
5	650–685	35
10	689–723	33
15	727–771	34
20	775–827	52

was also observed (see figure 2). The immediately obvious features of the fluctuation with respect to time delay (figure 2b) are a peak at or around zero with a symmetric decrease to a minimum on either side at approximately ±10 ns and another increase towards the endpoints. In other words, the mean cluster size seems to vary in an approximately sinusoidal manner with a period of ~25 ns and amplitude of ~0.025.



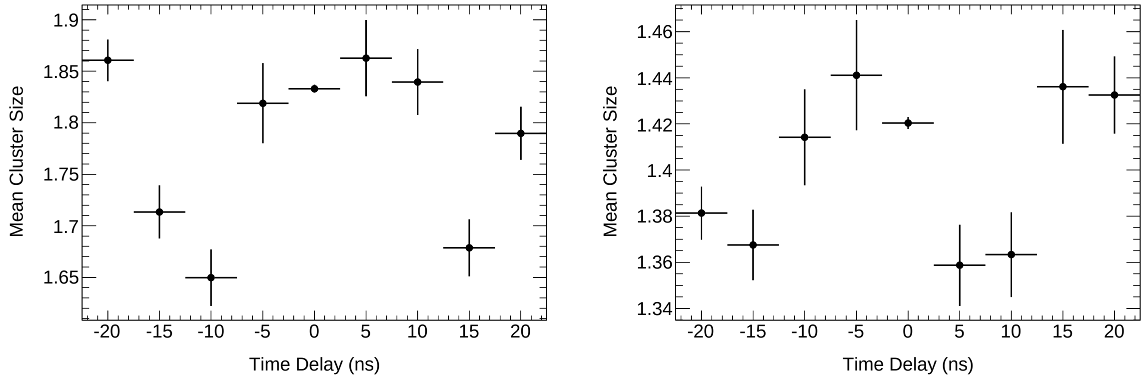
(a) Mean cluster size vs. event number.

(b) Mean cluster size vs. time delay.

Figure 2: Mean cluster size averaged over all modules during the timing scan. The left plot shows the variation as a function of event number, where the scan took place between 4.30x10^5 and 8.27x10^5 , and the right plot shows the same results as a function of the time offset of the combined clock/L1 trigger signal from its original value t_0.

Keeping in mind that each module will have different timing dependencies, a more relevant result is the variation in mean cluster size for a particular module (examples of which are shown in figure 3 for a typical endcap and barrel module). In comparison to the combined plot for all modules, these distributions have a similar shape and period, but the range in mean cluster size is about double. The reason for this lies in the fact that the peak in the distribution for each individual module occurs at different time values. This is clear from figure 3, where the barrel module peaks at ~5 ns whilst the endcap module peaks around -5 ns. In averaging over all modules then, there will be some destructive interference and hence a lower amplitude. The individual module plots also sit on different pedestals – a reflection of the fact that modules in different geometrical positions have inherently different mean cluster sizes[4].

To gain insight into the extent of cluster size variation across the SCT, the range (that is, the highest value of the mean subtracted by the lowest) was calculated for each module. Figure 4 shows the result for the innermost and outer two layers of the barrel and the second disk of endcap A. In all cases, the range was consistent along the ϕ direction, as was expected due to the inherent axial symmetry of the SCT. The range increased (decreased) towards higher values of η for the barrel (endcaps). These trends generally hold for all of the other SCT layers and disks, though they are not presented here. The range was also observed to progressively increase towards the outer layers of the barrel, which is clear from figures 4a, 4c and 4d.



(a) Barrel module at (layer, η, ϕ) = (3, 2, 4).

(b) Endcap C module at (disk, η, ϕ) = (1, 0, 11).

Figure 3: Variation in mean cluster size for specific barrel (left) and endcap (right) modules. The distribution, in both cases, approximately resembles a sinusoidal shape with a period of ~25 ns. The times of the central peaks (i.e. the phases of the distributions) are not the same, due to the differences in each module's geometric position and the length of optical fibre joining it to the rest of ATLAS.

5 Hypothesis for Cluster Size Variation

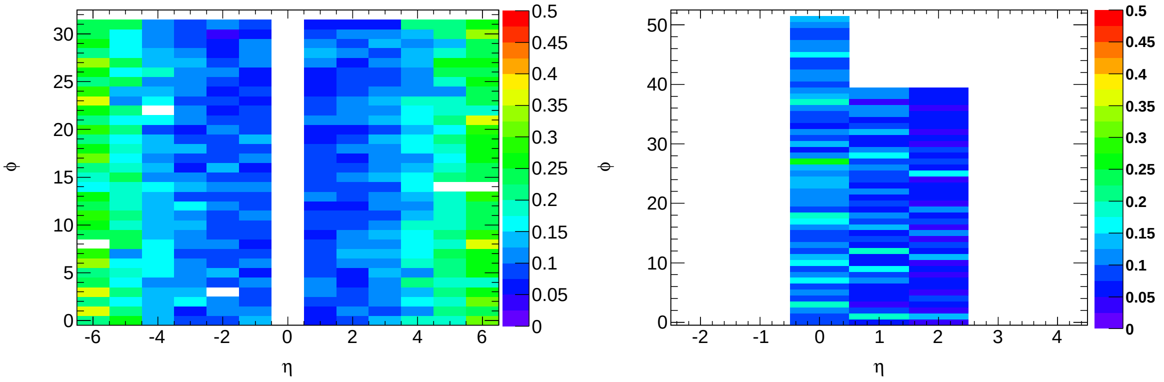
A successful explanation of the cluster size variation during the timing scan should be able to account for the features in the distributions shown in section 4, in particular:

- The reason for the peak and the minima adjacent to it.
- The periodic nature of the fluctuation ($T \sim 25$ ns).
- The differences in the phase of the distributions for different modules (i.e. the position of the central peak).

The timing scan was conducted with the input register set to level mode, and so the action of shifting the clock should be evaluated in this context. Consider a pulse which is only over the threshold for a short period of time – to be sure, let this time over threshold (ToT) be less than a single clock cycle (ToT < 25 ns). For simplicity, say that the signal was over threshold in the time range $[t_0 - \epsilon, t_0 + \epsilon]$, where $\epsilon < 12.5$ ns and t_0 is again the rising edge of the central clock cycle. The signal will only be recorded if it coincides with a rising clock edge (i.e. t_0 or $t_0 \pm 25$). Without a time shift, for example, the hit

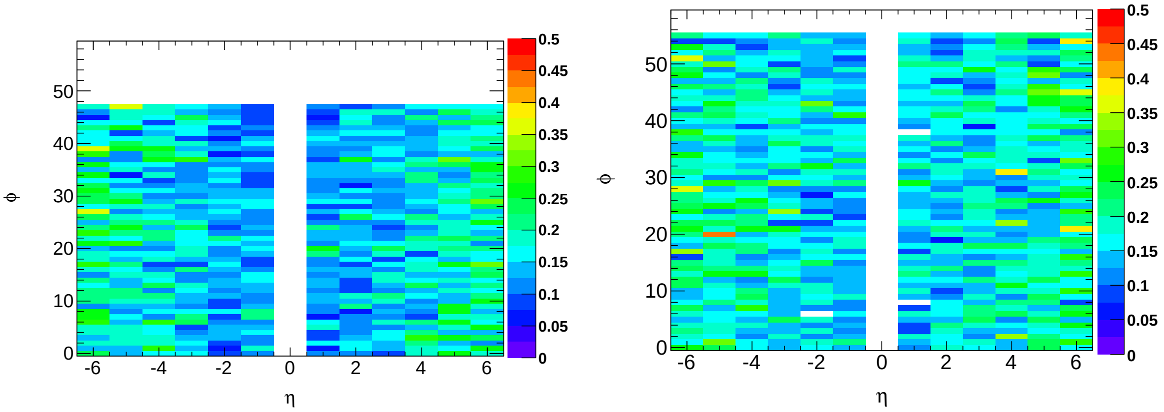
---

[4] For example, particles travelling to outer layers of the barrel are bent more by the **B** field and therefore enter these layers at higher incident angle. Typical cluster sizes are therefore larger in outer modules.



(a) Barrel layer 0.

(b) Endcap A disk 1.



(c) Barrel layer 2.

(d) Barrel layer 3.

Figure 4: The range of mean cluster size values throughout the timing scan, plotted per module in units of SCT strips. In all plots the range is consistent in ϕ as expected. There is a general increase towards higher (lower) values of η in the barrel (endcaps), and towards the outer layers of the barrel. The values of η and ϕ numbers here correspond to the module assignments and not the usual ATLAS coordinate system. The 'white' modules in the figures above represent those that were out of the configuration (i.e. excluded due to some form of malfunction) from run 152777.

pattern will therefore be 010. As the clock is shifted towards $t > t_0 + \epsilon$ the hit pattern will eventually become 000 and the signal on that strip will no longer be recorded. Once the time is shifted enough (to $t > t_0 + 25 - \epsilon$) the later time bin will start sampling the signal and the hit pattern will become 001. The same will happen as the clock is shifted to earlier times (010 → 000 → 100). The important point of the above discussion is that *signals over threshold for less than a clock cycle will be missed under some time settings in level mode* (see figure 5).

The hypothesis on the cluster size variation is that as the timing is shifted, the mean cluster size rises and falls because some fraction of these strips with ToT < 25 ns are being missed by the input register. For example, a cluster that is three strips wide in optimum timing could become 2, 1 or 0 strips wide in bad timing. In general, the outputs from the discriminator will have different widths and will not necessarily be centred on t_0 – in fact, their centres will also not coincide (an effect referred to as timewalk [4]). Under the hypothesis, then, the peak in mean cluster size corresponds to a 'best-case scenario' time which captures the most number of strips, and the the total variation (the range) is related to the fraction of strips missed by shifting the clock. This clearly addresses the first of the above requirements for a

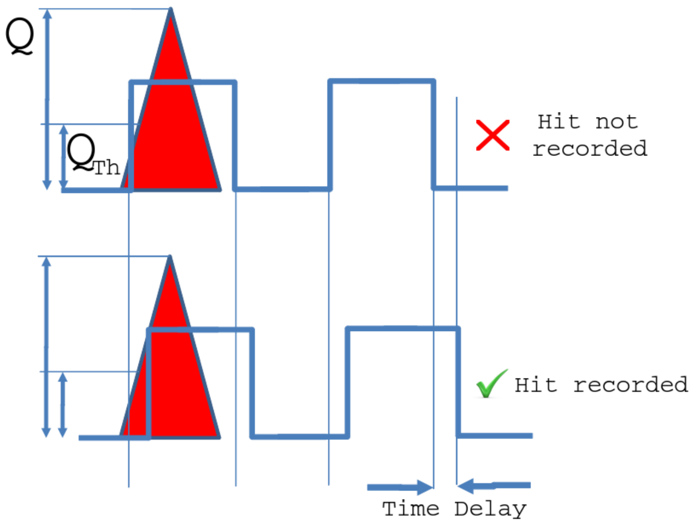


Figure 5: A schematic representation of the consequences of shifting the clock on the recording of signals over threshold for < 25 ns. To be recorded, the signal must be over threshold at the time of the rising clock edge. In some configurations the signal is missed (upper figure), whilst in others it is registered (lower figure). Here, Q refers to the signal charge and Q_{Th} is the threshold charge.

successful explanation. It also solves the second, since if we shift the clock by exactly 25 ns the signal will still be recorded, but just with a different hit pattern (001 or 100). The third is accounted for by the fact that the time in which the most strips are captured (i.e. the peak) changes from module to module in response to optical fibre lengths and time of flight of particles, as discussed in section 2.2.

This explanation assumes that these signals with ToT < 25 ns exist, and that a significant fraction of them occur around a similar time. This is what makes it a hypothesis and necessitates testing it to evaluate its validity. The following section concentrates on some of these tests, and briefly describes on an attempt to reproduce the sensitivity of cluster size to timing using a simulation of the SCT front end electronics.

6 Testing the Hypothesis

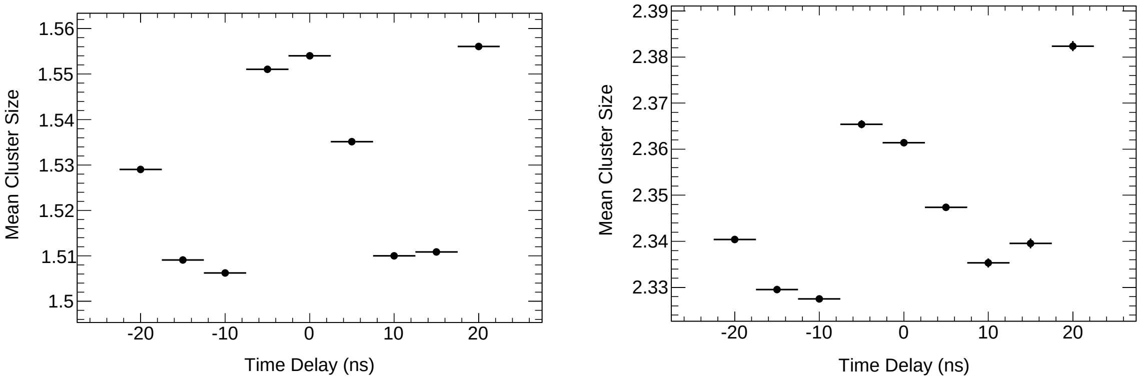
After formulating the hypothesis discussed in section 5, attempts were made to test the predictions it made and therefore assess its correctness and validity. Initially, this involved analysing the Minimum Bias data sample produced during the scan. In later stages, the sensitivity of the front end electronics to time offsets was investigated using several simulation models. Results from both of these approaches are described below.

6.1 Isolated Clusters from Particles Only

The hypothesis predicts that the fluctuation of cluster size occurs due to a loss of strips in the signal clusters[5] themselves as opposed to, say, the signal-to-noise ratio (a decrease of which could also lead to a drop in mean cluster size, since noise clusters characteristically have a size of 1). The aim of the first test was to isolate clusters which had a high probability of containing real hits (i.e. not noise) and check that they demonstrate a similar fluctuation in mean size as was originally observed. A simple way to do

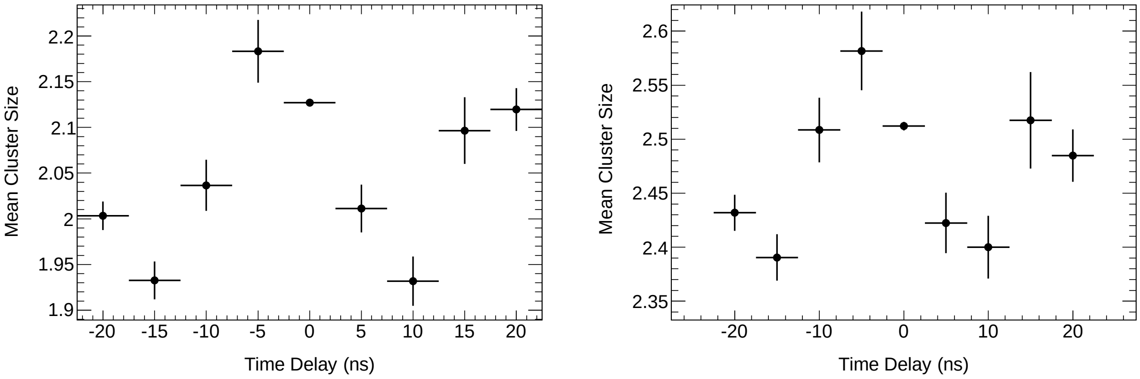
---

[5] In this context, the term *signal clusters* refer to those produced by particles, rather than noise.



(a) All clusters for all modules.

(b) Only clusters with size ≥2 (all modules).



(c) All clusters for barrel SCT module at (layer, η, ϕ) = (0, 6, 25).

(d) Only clusters with size ≥2 (the same barrel module).

Figure 6: Variation in mean cluster size with time shifts, showing the comparison between all clusters (left) and clusters almost completely from real hits (right). The top plots are the combined results from all modules, whilst the bottom are for a specific example barrel module.

this was to demand that clusters were selected with size greater than one, which greatly diminishes the fraction of noise clusters[6]. Figure 6 shows the distribution of mean cluster size with time offset using all clusters (6a and 6c) and just clusters with size greater than one (6b and 6d). The top plots are the combined results for all modules across the entire SCT, and the bottom is for a specific example module in the barrel. In both cases, the original fluctuation (using all clusters) is still clearly visible in the case of signal-only clusters. Both the position of the peak (i.e. the phase of the distribution) and the amplitude of fluctuation are well correlated. This is particularly striking in the case of a single module, where the complexity associated with averaging across all modules is removed.

This by no means proves the hypothesis to be correct, but it does provide some good evidence for it. At the very least, it demonstrates that the fluctuation is due to changing characteristics of the signal clusters leading to a downward shift in the mean cluster size and not just a change in the signal-to-noise ratio.

---

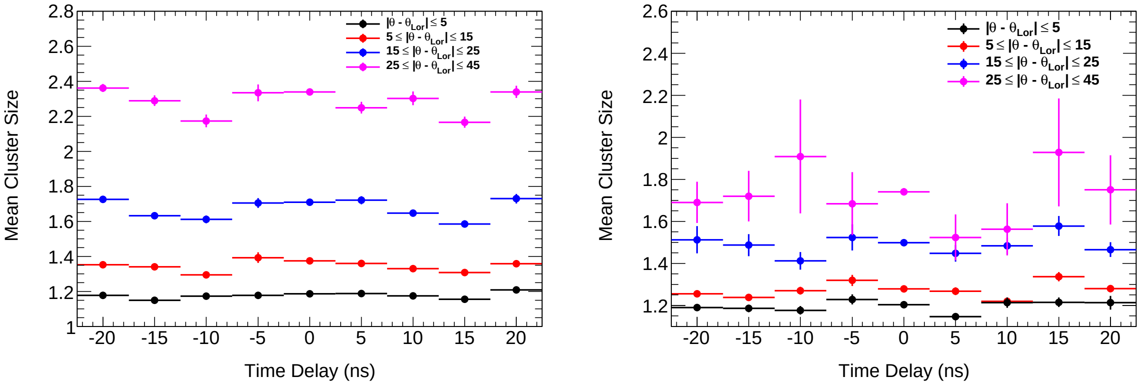
[6] Noise occupancy (NO = μ/N where μ is the mean number of noise hits on a particular strip for a given number of time bins, N) has been measured to be $\sim 10^{-5}$ [4]. From binomial statistics, this is also the probability of an individual strip having a noise hit in a specific time bin. Assuming the noise is not correlated to itself, the probability of at least one noise cluster with size 2, 3, 4 etc across the entire SCT is then $NP^2 \sim 10^{-4}$, $NP^3 \sim 10^{-9}$, $NP^4 \sim 10^{-14}$ respectively (the factor N approximately accounts for combinatorics)

6.2 Isolating Narrow Signals

The next test was designed to more specifically target the hypothesis. It involved attempting to isolate clusters which were rich in hits produced from signals with ToT < 25 ns (‘narrow’ hits). If the hypothesis were true, it would be expected that the fluctuation in mean size during the timing scan would be more pronounced in these clusters.

To produce a set of clusters with a greater concentration of these narrow hits, a selection criterion based on the incident angle of the particle to the module (obtained from the track to which the cluster was assigned) was used. The motivation behind this was that the holes produced from tracks with higher incident angles are shared amongst more strips, and so the signal on each of these strips would, on average, be smaller and therefore over threshold for a shorter time. For several randomly chosen modules, the mean cluster size plot was produced separately for four different incident angle bins (measured about the Lorentz Angle). Two examples are shown in figure 7.

For some modules (figure 7a, for example), it was clear that clusters with increasingly larger incident angles (from black to purple in the plot) show larger variation in mean size. This is just what we would expect from the hypothesis, and so constitutes clear support for it. However, consider figure 7b – the total variations in the first three angle bins (black, red, blue) are all quite similar and whilst the purple seems to show a larger fluctuation it also has bigger error bars, making it difficult to be certain. In cases such as this, the result is more ambiguous (commonly due to a lack of statistics) and conclusions with respect to the hypothesis are harder to draw.



(a) Barrel module at (layer, η, ϕ) = (0, 1, 10).

(b) Endcap A module at (disk, η, ϕ) = (0, 1, 10).

Figure 7: Variation in mean cluster size with time offset, shown for four different ranges of track (to which the cluster belongs) incident angles to the modules, measured with respect to the lorentz angle.

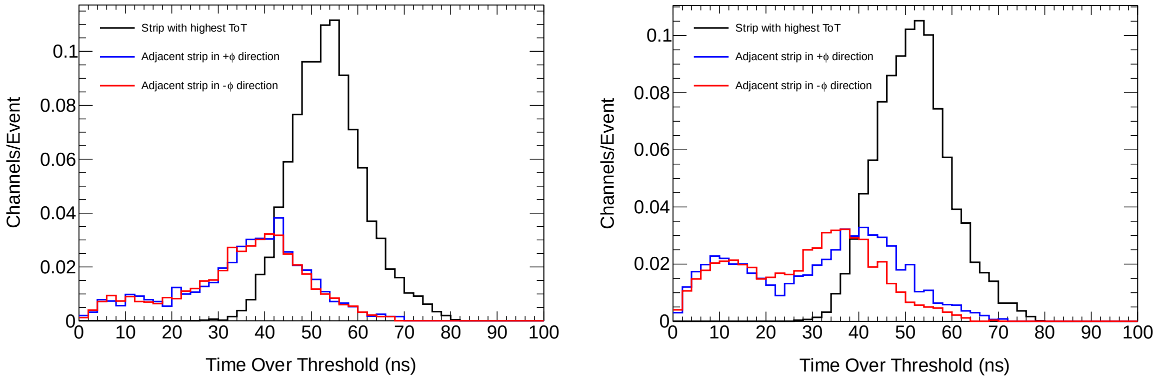
6.3 Testing with Simulation

The results from collision data are essentially binary outputs for each channel in each time bin. In contrast, a toy Monte Carlo model simulating the operation of the SCT can provide information regarding the ToT of the hits. This is extremely useful for directly testing the hypothesis, rather than looking for the signatures it predicts as done in sections 6.1 and 6.2.

The initial step in developing the toy Monte Carlo data was to simulate the creation of electron/hole pairs, their propagation through the silicon and the subsequent signal formation on the SCT strips. Two different models were considered – the first being the ballistic model (see [7] for details) currently employed as the standard for Monte Carlo studies and the second a recent development based on induced current at the strips due to the movement of the charges (the *Induced Current Model*). The signal processing by the front-end electronics was then modelled to produce a simulated pulse shape. At particular 'times', the pulse shape was compared to the threshold voltage, and the resulting hits grouped into clusters. There are several parameter values that must be provided to the simulation algorithms – for example the (layer/disk, η) of the module and the type and p_T of the incident particles. For simplicity, a barrel module at (layer, η) = (0,6) was chosen as a particular case study, with $\pi^{\pm}$ particles incident having a p_T spectrum matching the distribution from the Minimum Bias stream of run 152777.

Using the toy Monte Carlo model, several important measurements and distributions were produced. The existence and abundance of hits with ToT < 25 ns was a particularly critical assumption in the hypothesised explanation of mean cluster size variation. Figure 8 shows the distribution of ToT for the strip with the highest ToT value, and the two adjacent strips. Both the standard simulation and the Induced Current Model (ICM) predict that a substantial fraction (4.1% and 7.7%, respectively) of hits will have ToT<25 ns. It also shows that these narrow hits will generally occur on the 'edge' of the clusters – i.e. adjacent to the strongest signal. As far as we can trust the simulation, this is very good evidence for the hypothesis, but is still not conclusive as it only proves the existence of signals with ToT<25 ns rather than directly implicating them in the mean cluster size variation. To do this, the fraction of narrow hits recorded by the input register was plotted for time offsets corresponding to those used in the timing scan. The change in this fraction was then compared to that of the mean cluster size and is displayed in figure 9. For both models there is a very clear direct positive correspondence between the two quantities, which is essentially proof of the original statement of the hypothesis – that the rising and falling mean cluster size is due to a respective fluctuation in the fraction of narrow signals recovered.

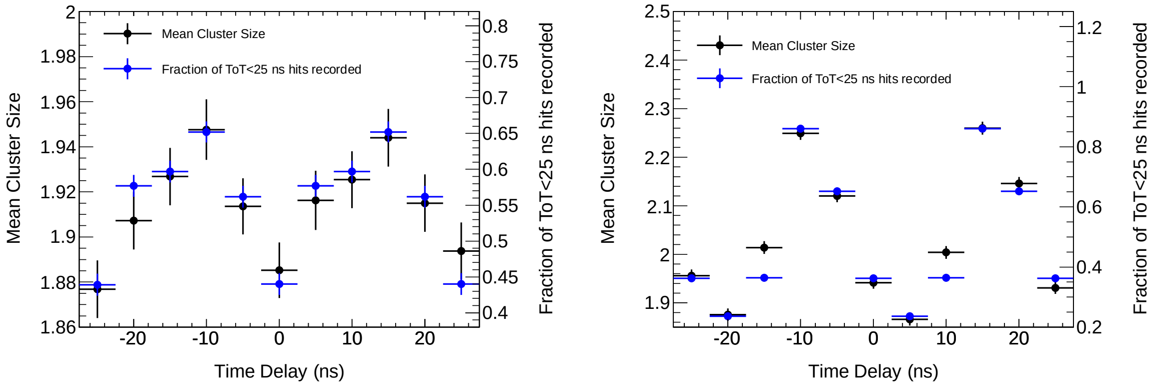
It is important to note that both models also reproduced the observed sinusoidal-like periodic variation of mean cluster size with timing offset, which itself is a good indication that they are performing the simulation adequately. However, the models also show distinct differences in their results – notably, the ICM has a greater number of (adjacent) strips in the 0–20 ns region of ToT and also predicts a significantly larger level of fluctuation in mean cluster size. Features such as this can be used, to some degree, as a tool for determining the relative correctness of the models by comparing to equivalent results in real collision data.



(a) The standard simulation model

(b) The Induced Current Model

Figure 8: The distribution of ToTs for hits generated using the standard simulation model and the recently developed Induced Current Model. The distribution is presented separately for the hit with the highest ToT and the two adjacent strips. For both models, there is a significant fraction of hits with ToT< 25 ns which is strong support for the hypothesis.



(a) The standard simulation model

(b) The induced current model

Figure 9: The variation of mean cluster size overlaid with the corresponding variation in the fraction of hits with ToT<25 ns recorded.

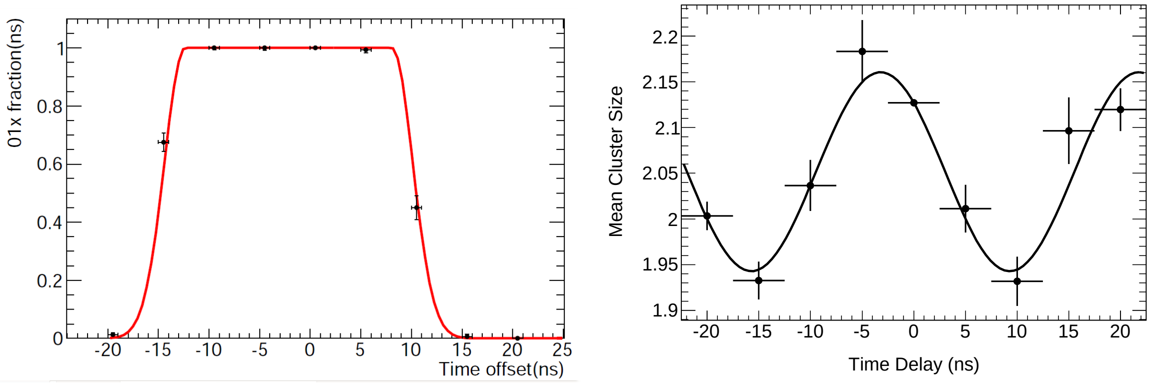
7 Implications for Optimum Timing

The tests presented in the previous section seem to strongly indicate that the hypothesis is correct (a discussion of this can be found in the section 8). For the moment, consider what the implications of this being true are – specifically, in regard to the optimum time settings of the modules. The results indicate that, at certain time settings, hits resulting from the passage of particles are being missed by the input register. This is clearly undesirable, and it appears that our aim should be to pick the time offset which minimises the number of these hits that are lost. This is simply a matter of picking the time offset for each module that gives a maximum in the mean cluster size. The number of data points in the mean cluster size distributions are not enough to obtained a detailed picture of their exact shapes, but they are clearly periodic and contain a fluctuation of order $\sim$0.1 strips. To obtain the time of maxiumum cluster size, we simply needed to determine the phase of each of the distributions. The simplest function capable of doing so and containing the general features of the distributions, a sinusoidal curve, was therefore used as the fitting function for each individual module. The period was fixed at 25 ns (which we know from the clock frequency) while the pedestal and amplitude were floated. An example is shown in figure 10b for a barrel module at (layer, η, ϕ) = (0, 6, 25). The optimum time was then extracted as the point of the central maximum (-3.21 ns for this particular module).

The most relevant quantities to compare these optimum times to are those extracted using the original scheme. This approach selects clusters on tracks with $p_T > 500$ MeV and all their strips having the same hit pattern, then plots the ratio of 01X clusters to all others. A plateau (of width $\sim$25 ns) is then fitted (see figure 10a for an example), and the optimal time is chosen as its centre to ensure that the maximum numbers of clusters are in the correct time bin (i.e. 01X). Figure 11a shows the sets of optimum times obtained from each of the two methods, and figure 11b shows the differences[7] in times on a module-by-module basis. The times here are measured with respect to the time setting prior to the scan (t_0 in the earlier notation). The distribution from maximising the 01X ratio is double peaked, which is a result of two different populations of modules (the barrel modules tend to the values around the +5 ns peak and the endcap modules to the -2 ns peak). In comparison, the times from the maximum cluster size scheme show this feature to a much smaller extent, and the entire distribution appears negatively shifted by $\sim$3 ns. This is confirmed by the plot of differences (figure 11b), which is peaked around -3 ns with a small spread about this value.

---

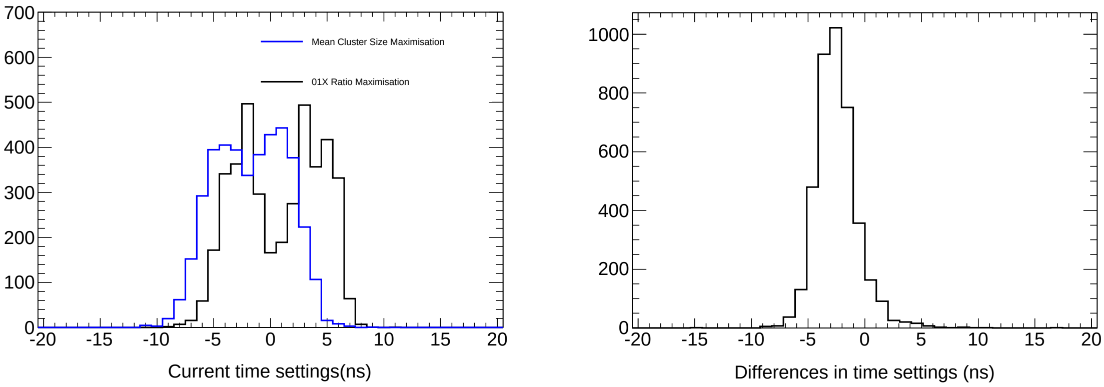
[7] The difference here is defined as maximum cluster size time – maximum 01X ratio time.



(a) Endcap C module at (disk, η, ϕ) = (0, 0, 25)

(b) Barrel module at (layer, η, ϕ) = (0, 6, 25)

Figure 10: Two different approaches to optimising the clock timing of the SCT modules. On the left is an example of the current scheme, whereby the clusters with all strips having the same hit pattern are selected, and the ratio of 01X clusters is plotted. The optimum time is then set as the centre of the maximum ratio plateau. On the right is an example of a sinusoidal fit to the mean cluster size plot, with the optimum time extracted as the point of the central peak.



(a)

(b)

Figure 11: A comparison between the two different approaches to finding the optimum timing for each module. On the left are the distributions of times found by both methods and on the right are the differences in these times calculated module-by-module, defined here as maxiumum cluster size time – original time setting.

Currently, the modules in the SCT are set to the times using the original 01X ratio method. The results presented here raise several important questions, including whether this is indeed the correct approach and if not should the cluster size results be used instead? In fact, since the 01X ratio is maximised (very close to 1) across an entire plateau of ~25 ns, the criteria from both approaches can be simultaneously satisfied – so perhaps this is the best solution. However, before any changes can be made, further study is required to understand this more completely – in particular, what the reason behind the -3 ns difference is.

7.1 Results Utilising Maximum Mean Cluster Size Times

Due to the axial symmetry of the SCT, modules with a given η value (but different in ϕ) are expected to be inherently similar. In particular, figure 4 shows that these modules have a similar total mean cluster size variation. As such, it seems plausible that all the modules from a specific η 'ring' can be analysed together. However, we also know that different modules have peaks in mean cluster size at different time offsets (see figure 3). This issue can be overcome by measuring the time offsets with respect to the peak times, found using the above method for each particular module. This effectively 'shifts' all the distribution so their peaks are aligned.

In combining results from different modules, we get a quasi-continuous distribution of time offsets because the times of the mean cluster size peaks are not necessarily discrete[8]. This allows the use of finer binning and therefore more detailed distribution shapes. Using this combination technique also lowers the uncertainty in the mean cluster size, since more results are being used. Example plots using this technique are shown in figure 12. The top plots are the usual mean cluster size distributions for two example eta rings, both of which are supportive of the original observation of a 25 ns periodic fluctuation with a central peak and two adjacent side peaks. However, with the increased detail it appears that the precise shape has a broader structure with, from inspection, more of a |cosine| dependence (shown fitted to plots) rather than simply cosine. Since the period of the two functions would still be chosen to be equal, the optimum times extracted by fitting with either function should be very similar. To confirm this, a second fit over all the modules was performed using a $\sim$ |cosine| function, and the optimum times (the positions of the central peaks) obtained were compared to the original set. An average shift of 0.05 ns (with an RMS of 0.6 ns) was observed, which was considered insignificant with respect to the $\sim$1 ns uncertainty on the original values [9]. Importantly, the $\sim$3 ns difference also remains stable. The bottom plots are the equivalent of figure 7, and again clearly demonstrate strong evidence in favour of the hypothesis. They also greatly eliminate the ambiguity due to low statistics present in the case of many of the individual modules.

8 Discussion

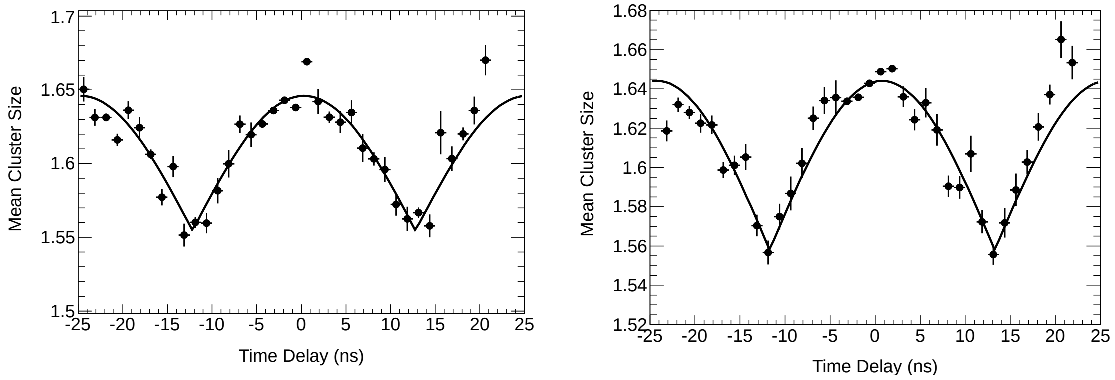
It is difficult to determine directly from real data whether the hypothesis presented in section 5 is indeed correct, mainly because the outputs from the SCT are binary with no information regarding the ToT of hits stored. Instead, the analysis focused on signatures or predictions derived from the hypothesis (other than the original observation) that could be investigated using real data. The two main tests conducted involved attempting to isolate hits either from real particles or from signals that had ToT < 25 ns. Both provided a reasonable amount of positive evidence for the hypothesis, in that some modules (usually those with a larger number of clusters recorded) showed the predicted trends well whilst others were more ambiguous (again, generally due to statistical uncertainty). This alone provides reasonable evidence for the hypothesis, but for completeness we showed direct evidence in results obtained from simulations of the SCT.

For several of the results, a lack of adequate statistics lead to a significant uncertainty in the shapes of distributions and the values extracted from them. For example, the sinusoidal distribution shape and periodicity for some modules was greatly obscured or absent, which consequently introduced a large uncertainty into the time of the peak mean cluster size for that module. In the test of the hypothesis regarding the fluctuations of clusters belonging to different incident angle tracks, it was also remarked that trends were ambiguous and obscured. These issues were somewhat overcome by combining results throughout specific η 'rings', but this procedure itself depends on extracting the times of maximum

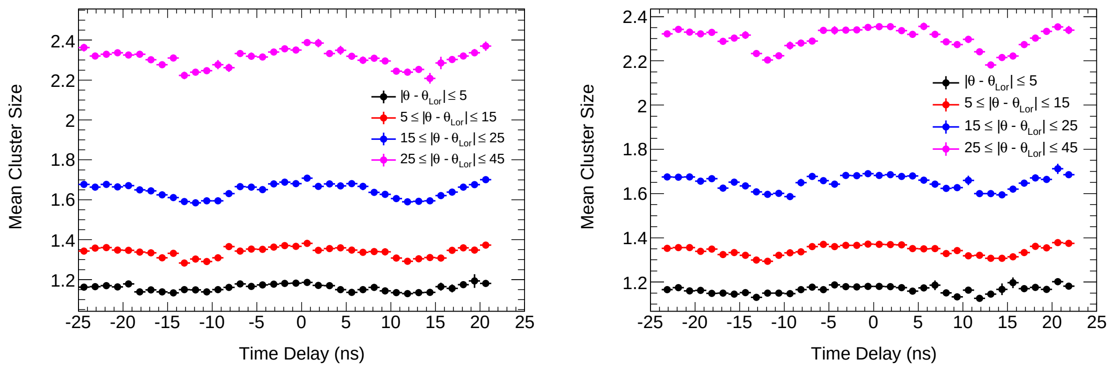
---

[8] For example, if the peak occurs at -3.4 ns then the possible time offsets are $-3.4 \pm 5, -3.4 \pm 10, -3.4 \pm 15, -3.4 \pm 20$ ns.

[9] There are a small number of modules in the tails with up to $\sim$2 ns differences. These may require further investigation.



(a) Barrel η ring at (layer, eta) = (1,1). (b) Barrel η ring at (layer, eta) = (1,-1).



(c) Barrel η ring at (layer, eta) = (1,1). (d) Barrel η ring at (layer, eta) = (1,-1).

Figure 12: Results obtained from combining several modules with a given η value (i.e. an η 'ring'), by measuring the time offsets with respect to the time of peak mean cluster size. The tops plots show the mean cluster size distributions as a function of time offset (fitted with ~ |cosine|), and the bottom plots are equivalent to figure 7. In both cases, combining the results allows finer binning and more detailed distribution shapes.

cluster sizes for individual modules. Ultimately, further timing scans conducted with more time offsets and more statistics would be the most helpful in revealing the exact shapes of these distributions. This would also allow a p_T cut on the tracks belonging to the clusters without an unacceptable loss in statistics. This is desirable as it would remove the smearing effect on the optimum time due to the differences in the time of flights of particles with different p_T.

At some point the SCT may begin operating in edge sensing mode. In this mode, the signal is not latched to the rising edge of the clock, but rather a high to low change in the discriminator output is detected at any point within each time bin (see section 2 for more details). As a result, the recording of a hit is far less dependent on the arrival time of the clock signal and under the hypothesis presented here there should subsequently be almost no variation in mean cluster size. This would therefore provide a futher test of the validity of the hypothesis. Another future avenue would involve determining the reason for the systematic ~3 ns difference in optimum timing approaches, as well as the impact of changing the timing settings to those obtained from the cluster size results. In a preliminary attempt to understand this, the cluster selection criteria used for the original optimisation scheme were applied to the maximum mean cluster size technique. The results showed that the effect of this was very slight and that the -3 ns shift remained. In section 6.3 it was mentioned that these cluster size results are also of interest to the digitisation community for use as a discrimination tool between competing models. Aside from the mean cluster size distribution, other quantities such as just the 001, 01X, 1X0 cluster sizes and ratios can be measured and compared to simulation predictions. This would most likely necessitate the combination of results technique described in section 7.1 to remove limitations due to statistics.

The results shown here are important for many reasons. Firstly, they represent an opportunity to further understand the detailed operation of the SCT and in particular the method of latching the signals to the clock in level sensing mode. This is especially relevant to the digitisation and simulation of the SCT and their ability to emulate reality. Secondly, it gives an insight into how the timing settings at the module level affect the ability of the SCT to detect signals and to correctly measure cluster size. Finally, assuming the hypothesis is correct, it is intuitively clear that maximising the cluster size for each individual module provides an alternative definition of optimum timing offset.

9 Conclusion

The SCT timing scan performed in April 2010 showed that, when operating in level mode, the mean cluster size for each module is sensitive to time offsets. The dependency appears to be periodic and, at least to some extent, sinusoidal with a frequency equal to that of the ATLAS clock (40 MHz) and an amplitude of order ~0.1 strips (~ 7%). A hypothesis based on the operation of the front end electronics and their dependence on trigger, timing and control signals was proposed as an explanation for this behaviour. The phase of each distribution, which determines the central maximum in mean cluster size, was found to be different for each module and was interpreted as a reflection of the differences in particle time of flights and optical fibre lengths.

This hypothesis was tested indirectly by analysing distributions related to specifically selected clusters, and directly by investigating the action of time offsets on simulated clusters. Both of these seem to strongly indicate that the hypothesis is indeed correct and valid. Following this, the times corresponding to maximum mean cluster size were determined and interpreted as the optimum time settings for each module. These were compared to the original optimisation scheme, with a systematic difference of ~3 ns being revealed. There is therefore contention over which is fundamentally the correct approach to take. The original scheme maximises the 01X ratio and is the current system instated. However, the nature of this approach allows for simultaneous maximisation with the cluster size, so this seems to be the best solution. Further work may be needed to confirm no adverse outcomes of this action, and would involve understanding the nature of the 3 ns difference.

References

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

[2] ATLAS Collaboration, G. Aad et al., *The ATLAS Inner Detector commisioning and calibration,* arXiv:1004.5293v2 [physics.ins-det] (2010) .

[3] A. Abdesselam et al., *The data aquisition and calibration system for the ATLAS Semiconductor Tracker*, JINST **3** (2008) P01003.

[4] F. Campabadal et al., *Design and performance of the ABCD3TA ASIC for readout of silicon strip detectors in the ATLAS semiconductor tracker*, Nucl. Instrum. Meth. A **552** (2005) 292–328.

[5] C. A. Magrath and S. C. J. J. Kortmann, *The heart of ATLAS: Commissioning and performance of the ATLAS silicon tracker. oai:cds.cern.ch:1193348*. PhD thesis, Nijmegen, University, Nijmegen, 2009. Presented on 12 Jun 2009.

[6] M.-L. Chu et al., *The off-detector opto-electronics for the optical links of the ATLAS SemiConductor Tracker*, Nucl. Instrum. Meth. A **530** (2004) 293–310.

[7] S. Gadomski, *Model of the SCT detectors and electronics for the ATLAS simulation using Geant4*, Tech. Rep. ATL-SOFT-2001-005, CERN, Geneva, May, 2001. revised version number 1 submitted on 2001-06-18 18:53:11.

3.1 Extension to 2011 Timing Scans

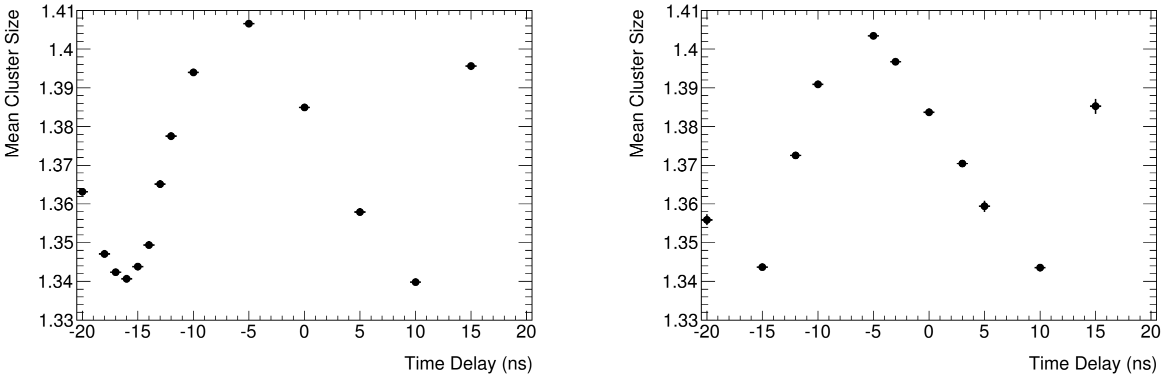
Subsequent to the 2010 timing scan, the time offsets for each module were set to the central value of the 01X ratio plateau (see Section 7 of the previous note). Given the result in Figure 11.b, this implied that the central maxima in the cluster size vs. timing offsets distribution should be fairly well-aligned for each module, at approximately -3 ns. To validate the 01X ratio calibration procedure (or, alternatively, to determine if it was, in some way, time-dependent) another two timing scans were performed in level-mode during 2011, and it was found that, though the endcap modules were generally well-calibrated, the timing offsets of the barrel modules tended to be a few nanoseconds too low. For these modules, then, the central maximum of the mean cluster size should be expected closer to -5 ns.

The conditions for the 2011 timing scans were significantly different to the 2010 scan – in particular, the minimum bias data stream no longer had the highest bandwidth and the number of tracks per event was significantly larger, due to the types of events being triggered on and the increase incidence of pile-up events. Consequently, the *JetTauEtMiss* stream was used for the 2011 scans, with the clusters required to be hits on a reconstructed track with ≥ 1 pixel hits and ≥ 24 TRT hits.

Because all of the modules were now in phase with one another, they could be analysed collectively with no 'destructive interference', eliminating the need for the combination procedure used in Section 7.1. This allowed for a large increase in statistics, which, together with the extra timing offsets used in the 2011 scans, revealed extra structure in the mean cluster size distributions not previously seen (Figure 3.1). Though the periodicity and position of the central peak are as expected, the shape is clearly not sinusoidal; the rise on the left of the peaks is far sharper than that on the right. From the discussion in Section 5, and the details of the level-mode operation in Section 2, this implies that the signal pulse shapes produced by the preamplifier and shaper must also be asymmetric in a similar way. In fact, as explained in Ref. 7 of the above note, the theoretical amplifier output of the nearby strips for a delta pulse of total charge Q can be written as:

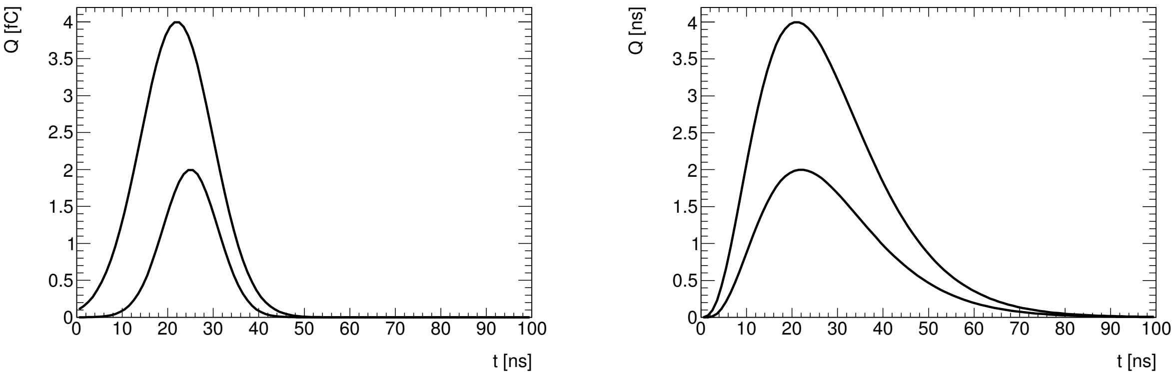
$$f(t) = \frac{Q}{3^3 e^{-3}} \left(\frac{3t}{t_{peak}} \right)^3 \exp \left(\frac{-3t}{t_{peak}} \right), \tag{3.1}$$

where t_{peak} is the peaking time for the amplifier. For the SCT modules, the total charge, Q, is expected to be of the order of fC and the peaking time for the ABCD chip is close to 21 ns. Examples of this pulse shape, and symmetric Gaussian pulse shapes with similar parameters, are given in Figure 3.2. As the charge builds up on the readout electrodes, a voltage is induced on the adjacent strips proportional to the derivative of



(a) Scan 1, 2011 (b) Scan 2, 2011

Figure 3.1: The mean cluster size vs. time delay distribution for the first (left) and second (right) 2011 timing scans, using the results from all modules.



(a) Symmetric (Gaussian) Pulse Shapes (b) Asymmetric Pulse Shapes

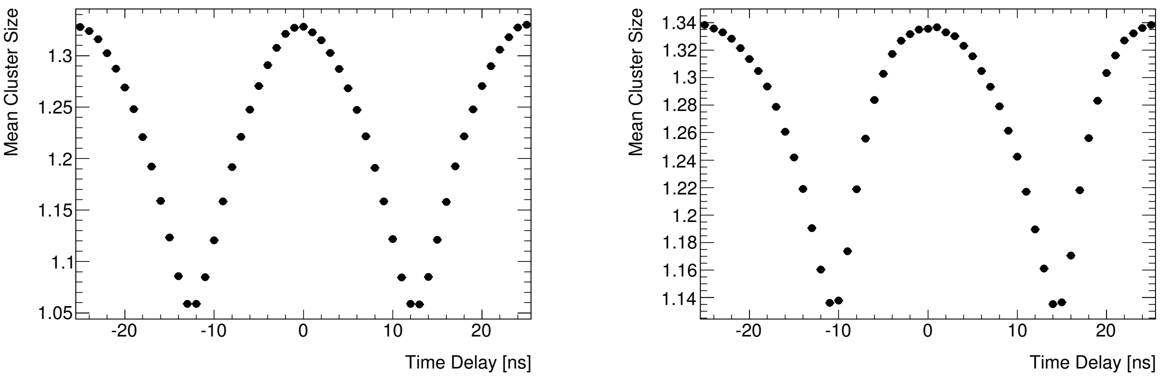
Figure 3.2: Example pulse shapes under a symmetric, Gaussian amplifier output hypothesis (left), and the expected asymmetric output for a delta pulse (see text).

the original signal shape, in a process referred to as cross-talk. For the SCT electronics, the proportionality constant is approximately 10%, and almost all of the induced signal appears on the immediate neighbouring strips.

Assuming these pulse shapes, a simple toy MC can be used to simulate the effect of timing offsets on mean cluster size. The procedure is as follows:

1. Sample the number, N_{hits} of *central* strips from Uniform(1,2). The mean cluster size of ≈1.4 implies that contributions of more than two central strips can be neglected.

2. Sample the total charge from a Gaussian distribution centred at 4 fC, and distribute this evenly amongst N_{hits}.



(a) Symmetric Amplifier Shapes (b) Asymmetric Amplifier Shapes

Figure 3.3: The mean cluster size vs. time delay distributions for the toy study assuming symmetric (left) and asymmetric (right) amplifier pulse shapes. For the former, the width of the pulse was sampled from $\sigma \sim \mathcal{N}(10, 4)$ ns, while both used $Q \sim \mathcal{N}(4, 2)$ fC and $t_{peak} \sim \mathcal{N}(21, 1)$ ns.

3. Construct each of the N_{hits} amplifier pulse shapes. In the case of the asymmetric pulse shapes of Equation 3.1, the peak time is assumed to be Gaussian distributed with a mean of 21 ns and a tuneable width. For the symmetric case, the width, σ, is also sampled from a Gaussian distribution.

4. Simulate the cross-talk on the closest-neighbour strips as the derivative of the shapes in the above step.

5. Pick 100 random time delays in the range $-25 < t_{delay} < 25$ ns. For each, check whether the pulse was over the threshold at either t_{delay} or $t_{delay} \pm 25$ ns.

Figure 3.3 shows the resultant mean cluster size distribution for symmetric (left) and asymmetric (right) pulse shapes. As expected, the former leads to a symmetric mean cluster size dependence on timing offset while the latter produces similar behaviour to that observed in the 2011 scans (c.f. Figure 3.1). This adds significant weight to the hypothesis for the mean cluster size variation outlined in Section 5. These results are for $\sigma \sim \mathcal{N}(10, 4)$ ns, $Q \sim \mathcal{N}(4, 2)$ fC and $t_{peak} \sim \mathcal{N}(21, 1)$ ns, though these parameters can be tuned to produce different amplitudes, ‘sharpness’ and asymmetry in the mean cluster size distribution.

In light of the above discussion, the mean cluster size distributions observed in the timing scans are a probe for the underlying amplifier shapes. In fact, the toy MC procedure above could be inverted to test the validity of any simulation and digitisation models for the SCT sensor and front-end electronics.

The 2011 scans also showed that some of the modules were badly calibrated in terms of their 01X ratio. The mean cluster size of distribution of these modules also showed poor calibration, in a consistent way, again validating the ability of the cluster size maximisation as a useful diagnostic tool.

Chapter 4

Excursus II: The Feasibility of the e^+e^- Channel for Quarkonia Studies

4.1 Introduction and Motivation

Despite the fact that vector quarkonia partial decay widths to the e^+e^- channel are generally very similar to those of the $\mu^+\mu^-$ channel[1], B-hadron physics measurements at ATLAS tend to be conducted primarily with the latter. For example, the official ATLAS[2] J/ψ inclusive cross-section measurement was performed solely with muons [14].

The large track-multiplicity and luminosity environment at the LHC demands high granularity in the ATLAS tracking detectors — the Pixel, SCT and TRT detectors contain 80.4×10^6, 6.3×10^6 and 3.51×10^5 channels, respectively [1]. The accompanying on-detector electronics, power distribution, cooling and mechanical support contribute significantly to the detector material budget [86]. The interaction of electrons with this material at typical quarkonium decay energies is dominated by the emission of photons under the process of bremsstrahlung, which cause deviations in the their trajectory as they pass through the detector. For muons, the rate of bremsstrahlung is lower by a factor of $\sim (m_e/m_\mu)^2 \simeq (1/207)^2 = 2.3 \times 10^{-5}$ [82]. Consequently, the tracking of electrons is a more challenging task, with the momentum resolution worse than that for muons. This allows greater accuracy for the $\mu^+\mu^-$ channel in the measurement of parent energy-momentum four-vectors and other derived quantities, such as invariant mass. The lower rate of bremsstrahlung also allows muons to pass through the calorimeter layers

---

[1] In the case of the J/ψ, for example, $\Gamma(J/\psi \to e^+e^-)/\Gamma(J/\psi) = (5.94 \pm 0.06) \times 10^{-2}$ and $\Gamma(J/\psi \to \mu^+\mu^-)/\Gamma(J/\psi) = (5.93 \pm 0.06) \times 10^{-2}$ [3]. Any slight differences are due to the slightly larger phase space available to the dielectron final state.

[2] This is also the case for other detectors at high energy hadronic colliders, including the CMS [89] and CDF detectors [90]. In contrast, for detectors such as Belle and BaBar which operate at lower energy e^+e^- colliders, electrons and muons are on equal footing. See [91] and [92] for examples.

with little energy loss, where they are essentially isolated from all other charged particles and can easily be further tracked and identified. Consequently, studies using muons tend to suffer considerably less from backgrounds associated with "fakes" (wrongly identified objects) than do those involving electrons (compare Figures 4.1a and 4.16a, for example). These differences obviously have significant carry-on effects for the feasibility and precision of physics measurements, such as cross-sections and spin-alignment, conducted with the dielectron final state.

Putting these difficulties aside for the moment, if the electron channel was able to be adequately reconstructed and utilised it could provide valuable complementary contributions to the corresponding muon analyses, beyond the obvious doubling in statistics. In the study of rare decays with a final state involving two quarkonia, for example, the increase in statistics is (naively speaking) four-fold. Electrons also generally have different acceptance properties to muons, since these objects are reconstructed in rather different ways (e.g. the ATLAS forward calorimeters could potentially be used to reconstruct electrons in an $|\eta|$ range out to 4.9, whereas muon tracking is restricted to $|\eta| < 2.5$). This being the case, quarkonium candidates reconstructed using electrons would populate the phase space with a different distribution to those using muons, which would be useful if one wanted to perform differential cross-section measurements in these regions. As another possible motivation for using electrons, consider the irreducible systematic uncertainties in a typical muon analysis. Simply increasing the statistics will be of little help here, but electrons will have different systematic properties, which may complement those of the muon channel.

In light of these possible advantages, an analysis was carried out by the author of this thesis to determine the feasibility of the dielectron final state for quarkonium analyses at ATLAS. The J/ψ meson was chosen as a particular case study because it has well established properties (mass, spin, decay widths etc.) and the corresponding analysis of $J/\psi \to \mu^+\mu^-$ had already been completed (by other members of ATLAS) and provided a basis with which to draw comparisons. The main goal was to mimic the inclusive differential cross-section measurement to produce some reasonable measurements with a basic idea of the main systematic uncertainties. Any difficulties or possible advantages of using the electron channel would then emerge as a natural consequence of conducting the analysis, with the assumption that these may generalise to other quarkonium studies. The philosophy of the approach to the analysis was to implement existing tools, techniques and guidelines for the use of electron objects as much as possible, and to use reasonable placeholders where necessary. This chapter provides a detailed description of the methodology and results of this investigation, beginning with a review of the

$J/\psi \to \mu^+\mu^-$ cross-section analysis. It has also been supplied to the ATLAS community as an internal communication (ATL-COM-PHYS-2013-386).

4.2 Overview of $J/\psi \to \mu^+\mu^-$ Inclusive Cross-Section Analysis

The following provides a brief summary of the ATLAS $J/\psi \to \mu^+\mu^-$ inclusive cross-section analysis. For further details on any aspect of the analysis, refer to the full paper [14].

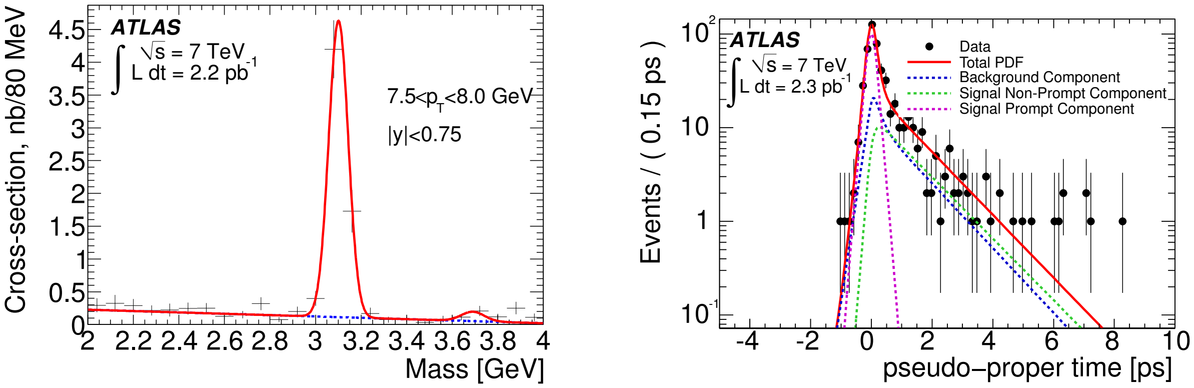
The ATLAS collaboration conducted a measurement of the J/ψ inclusive p_T-rapidity differential production cross-section at a $p-p$ centre-of-mass energy of $\sqrt{s}$ =7 TeV, using collision data from 2010 with a total integrated luminosity of 2.2 pb^{-1}. The production fraction of non-prompt J/ψ produced from the decay of a B-hadron,

$$f_B := \frac{\sigma(pp \to B + X \to J/\psi X')}{\sigma(pp \xrightarrow{\text{Inclusive}} J/\psi X')}, \tag{4.1}$$

was also measured in the same differential bins.

The analysis made use of a L1 muon trigger during periods of lower luminosity running, and two Event Filter (EF) single muon triggers (with p_T thresholds of 4 and 6 GeV) as the instantaneous luminosity increased. For the $B \to J/\psi$ fraction measurement, two extra triggers were employed, making the total integrated luminosity for this measurement 2.3 pb^{-1}. For an event to be selected for the analysis, one of the muons used in constructing a J/ψ candidate was required to match the muon trigger object. Further event-level selections were made to ensure that (1) the data was of sufficient quality (i.e. stable beams from the LHC and the Muon Spectrometer (MS), Inner Detector (ID) and magnet systems operational) and (2) the muons did not originate from cosmic rays. The latter was achieved by requiring there be a primary vertex with at least three tracks, each of which with at least one hit in the Pixel detector and six in the Semi-Conductor Tracker (SCT).

Within the remaining events, pairs of oppositely charged *combined* muons with ≥ 1 pixel hit and ≥ 6 SCT hits were designated as J/ψ candidates, and their inner detector tracks fitted to a common vertex. For the cross-section measurement, only a very loose requirement on the vertex quality was used which retained over 99% of candidates. In the case of the $B \to J/\psi$ fraction, where lifetime information was important, an additional cut on the probability of the vertex fit (being greater than 0.005) was necessary. Events were also rejected if the two muons were used in the construction of two different primary vertices. Collectively, these cuts excluded less than 0.2% of J/ψ candidates.



(a) The acceptance-and-efficiency-corrected invariant mass distribution in the $7.5 < p_T < 8.0$ p_T and $|y| < 0.75$ bin, weighted by the integrated luminosity. The fit shown is a Gaussian for the J/ψ and $\psi(2S)$ signal peaks with a linear background, and the errors are systematic and statistical combined in quadrature.

(b) The pseudo-proper time distribution in the $9.5 < p_T < 10$ GeV and $|y| < 0.75$ bin. The data is displayed as black points with error bars, and the solid red line is the total resultant PDF from the simultaneous fit to the corresponding invariant mass distribution. The signal prompt and non-prompt components, as well as the background component, are plotted separately as dashed curves.

Figure 4.1: Examples of a weighted invariant mass distribution and a pseudo-proper time plot for particular $p_T - y$ bins.

The probability that a particular $J/\psi \to \mu^+\mu^-$ decay is reconstructed by ATLAS and passes all the above selections depends on the kinematics of the J/ψ and of each muon. As such, to obtain the true number of $J/\psi \to \mu^+\mu^-$ decays that occurred, each candidate was assignment a weight, w, defined as the inverse of this probability. The probability itself was calculated in the following manner:

$$P(= w^{-1}) = \mathcal{A} \cdot \mathcal{M} \cdot \varepsilon_{trk}^2 \cdot \varepsilon_\mu^+(p_T^+, \eta^+) \cdot \varepsilon_\mu^-(p_T^-, \eta^-) \cdot \varepsilon_{trig} \tag{4.2}$$

where $\mathcal{A}$ is the kinematic acceptance, $\mathcal{M}$ is a correction factor for bin migrations (due to finite detector resolution), ε_{trk} is the inner detector tracking efficiency and $\varepsilon_\mu^\pm$ is the muon reconstruction efficiency (the ± here refers to the muon charge). The efficiency for one of the muons to be triggered is given by $\varepsilon_{trig} = 1 - (1 - \varepsilon_{trig}^-(p_T^-, \eta^-)) \cdot (1 - \varepsilon_{trig}^+(p_T^+, \eta^+))$, where $\varepsilon_{trig}^\pm$ now refers to the single muon trigger efficiency. This weighting method and the terms in Eqn. 4.2 will be explained and discussed in much greater detail for the case of electrons in Section 4.6 below.

The J/ψ candidates were organised into various bins of p_T and rapidity y — for example, in the region $|y| < 0.75$ fifteen p_T bins of variable width between 7 and 70 GeV were used. A (binned) minimum-χ^2 fit was performed for the weighted invariant mass distributions in each $p_T - y$ bin, with the J/ψ and $\psi(2S)$ signal peaks each described by a single Gaussian, and the background treated as linear. The total yield after background subtraction $N_{corr}^{J/\psi}$ was then extracted and used to determine the inclusive differential cross-section as

$$\frac{d^2\sigma(J/\psi)}{dp_T dy} Br(J/\psi \rightarrow \mu^+\mu^-) = \frac{N_{corr}^{J/\psi}}{\mathcal{L}\Delta p_T \Delta y} \tag{4.3}$$

where $\mathcal{L}$ is the total integrated luminosity of the data sample and Δp_T and Δy are the bin widths. For the $B \rightarrow J/\psi$ fraction, a more complicated procedure was required. A variable referred to as the 'pseudo-proper time' was calculated using the transverse displacement of the $\mu^+\mu^-$ vertex, L_{xy}, and the J/ψ transverse momentum, $p_T^{J/\psi}$:

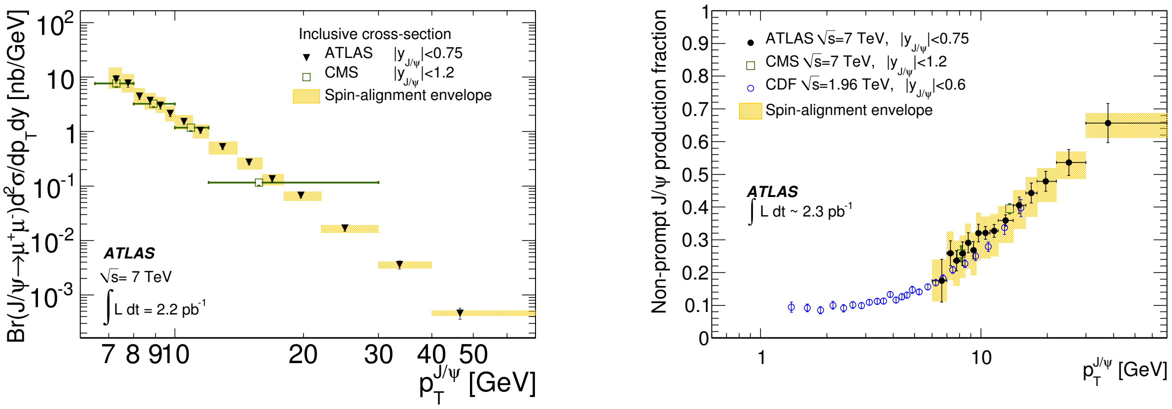
$$\tau = \frac{L_{xy} m_{PDG}^{J/\psi}}{p_T^{J/\psi}} \tag{4.4}$$

where $m_{PDG}^{J/\psi}$ is the world average value of the J/ψ mass. At large values of $p_T^{J/\psi}$, non-prompt J/ψ carry away a large fraction of the parent momentum, and the value of τ approximates the B-hadron lifetime (which follows an decaying exponential distribution). For promptly produced J/ψ the value of τ should be consistent with zero within the resolution of L_{xy} and $p_T^{J/\psi}$. By fitting the pseudo-proper time with a delta distribution at zero and an exponential (both convolved with appropriate resolution functions), the relative fraction of prompt to non-prompt J/ψ could be determined. In practice, unbinned maximum likelihood fits were simultaneously performed on the lifetime and invariant mass to ensure a proper treatment of the background component of the pseudo-proper time distribution. As illustrative examples, the fitted and weighted invariant mass distribution for the bin $7.5 \leq p_T \leq 8$ GeV, $|y| < 0.75$ and the pseudo-proper time distribution for $9.5 \leq p_T \leq 10$ GeV, $|y| < 0.75$ are shown in Figures 4.1a and 4.1b, respectively.

For both measurements, a number of systematic uncertainties were considered. The most significant of these is due to the fact that the spin-alignment of the J/ψ is unknown, and in fact must be measured[3]. The J/ψ polarisation determines the angular distribution of the decaying muons and can therefore have a strong impact on the acceptance of the ATLAS detector to these decays (this will be discussed in more detail in Section 4.6.1 for the case of electrons). Other important systematic studies included those pertaining to the tag and probe method used to determine the reconstruction and trigger efficiencies, the luminosity, bin migration effects, final state radiation, fitting methods and the effect of pile up on the accuracy of L_{xy}.

---

[3] This polarisation measurement was started after the initial cross-section analysis and when available, the spin-alignment uncertainty systematic can be lifted.



(a) The inclusive J/ψ production cross-section as a function of transverse momentum in the rapidity range $|y| < 0.75$. The yellow bands represent the systematic uncertainties associated with the unknown J/ψ spin alignment and the error bars represent the remaining (combined) systematic and statistical uncertainties. Comparable results from the CMS detector are overlaid

(b) The J/ψ non-prompt fraction f_B plotted as a function of transverse momentum in the rapidity range $|y| < 0.75$. Again, the yellow bands represent the spin-alignment uncertainty and the overlaid results are similar measurements from the CMS and CDF collaborations.

Figure 4.2: Examples of the inclusive differential J/ψ production cross-section and $B \to J/\psi$ fraction measurements as a function of J/ψ P_T for $|y| < 0.75$.

The evolution of the inclusive differential cross-section and non-prompt fraction were plotted for each of the rapidity regions $|y| < 0.75$ (Figure 4.2), $0.75 < |y| < 1.5$, $1.5 < |y| < 2.0|$, $2.0 < |y| < 2.4$, alongside comparable measurements made by the CMS and CDF collaborations (where appropriate). Finally, the non-prompt fraction was used to determine the separate prompt and non-prompt differential cross-sections, which were subsequently compared to various theoretical models. The predictions for the non-prompt J/ψ cross-sections were found to be consistent with experimental results, but those for prompt showed significant discrepancies, particularly at high $p_T^{J/\psi}$.

4.3 Data, Trigger and Monte Carlo Selections

4.3.1 Data and Triggers

At the beginning of LHC running in 2011, a new trigger menu was implemented for ATLAS which included several dedicated $J/\psi \to e^+e^-$ triggers. The motivation and design of these triggers was primarily for performance studies (in particular, for measurement of electron efficiencies at low E_T), but they nevertheless were also useful for physics analyses[4]. Two triggers in particular, *EF_2e5_tight_Jpsi* and *EF_e9_tight_e5_tight_Jpsi* , were chosen for the large number of J/ψ candidates they provided, the complementary p_T range they populated, and the selections they placed on the electrons. In both triggers, the electrons were required to pass the tight identification cuts at the event filter level with an E_T above threshold and a combined invariant mass in the 1–5 GeV range. There was no requirement on the sign of the two electrons, as same-sign combinations provide a valuable tool in the study of the (significant) combinatorial background.

During the early data-taking periods of 2011 (Periods A-C), most runs were dedicated to commissioning or special conditions and ATLAS was often not fully operational (e.g. for Period A a large fraction of the data was taken with the magnetic fields switched off). Physics runs at 7 TeV began in earnest in Period D with 50 ns bunch trains and between two and four-hundred colliding bunches. During Period E, a problem occurred in a subset of the Liquid Argon (LAr) calorimeter which caused six front-end boards (FEBs) to stop functioning. Unfortunately, this issue persisted until the beginning of Period I, where the FEBs were recovered and the LAr calorimeter became fully functional again. From Period L onward a new Athena release was implemented at Tier 0, which provided a natural end point for the data to be used in this analysis (significant changes in tracking, for example, would introduce complications regarding efficiencies).

Given the various conditions described above, only Periods D, I[5], J and K were included in this analysis. Though the data taken with the LAr FEBs non-operational was not useless it would have required careful treatment, particularly in the calculation of the acceptance of the detector under these conditions. Since statistics was not a limiting factor here, runs during this time were not included. As in the muon case, requirements were placed on functioning of individual detector components during these runs to ensure that the data used was of an adequate quality. Along with global criteria common to all physics objects, elements of ATLAS particularly important to electron and photon reconstruction (e.g. the electromagnetic calorimeter) were required to be operating at a high standard. The number of good events that each trigger supplied for these periods, and their total integrated luminosity, is given in Table 4.1.

---

[4] Ideally, one would use a single electron trigger for a physics analysis as these generally have more desirable acceptance properties (amongst other advantages). In a high instantaneous luminosity regime, however, this is not possible with a low E_T threshold without large prescaling.

[5] Due to technical problems in reading and processing a particular data file, Run 186179 was excluded from the analysis. This loss was conceded as it caused a reduction of <0.3% in the total integrated luminosity of the *EF_2e5_tight_Jpsi* trigger, and even less for the *EF_e9_tight_e5_tight_Jpsi* trigger.

Table 4.1: The number of good events and the corresponding total integrated luminosity shown for each of the four periods of interest, for the two triggers employed.

Period	***EF_2e5_tight_Jpsi***	***EF_e9_tight_e5_tight_Jpsi***
D	9204 (4.838 pb^{-1})	71566 (90.96 pb^{-1})
I	6378 (2.982 pb^{-1})	36539 (40.57 pb^{-1})
J	2132 (0.980 pb^{-1})	8435 (8.97 pb^{-1})
K	5856 (2.597 pb^{-1})	23167 (25.16 pb^{-1})
Total	23570 (11.396 pb^{-1})	139705 (165.65 pb^{-1})

4.3.2 Monte Carlo Samples

Monte Carlo (MC) data samples were primarily necessary for calculating various efficiencies and investigating the validity of the weighting method for electrons (though they were also used to determine other effects such as the loss of signal events in enforcing certain cuts). Separate Monte Carlo data samples of ∼5 million prompt and ∼2 million non-prompt J/ψ events were used. In both cases, the PYTHIA 6 [93] generator (which uses the MRST LO* [94] parton distribution functions) was used with the MC10 tune [95], and the passage of the particles through the ATLAS detector was simulated with GEANT4 [96].

The prompt sample includes direct production of the J/ψ in the primary (hard) interaction, as well as from the decay of excited charmonium states (e.g. $\chi_{c1} \to \gamma J/\psi$). The J/ψ candidates in the non-prompt sample originated from a B-hadron decay. In both cases, the decay of the electrons was assumed to be isotropic in the azimuthal and polar angles, since the true angular distribution (which is determined by the spin-alignment of the J/ψ) was unknown and represented a theoretical uncertainty (see Section 4.6.1 for more details).

All of the real and MC datasets were processed with the ATLAS reconstruction software in Athena Release 16.6.4.3 which, in particular, corresponds to the *loose, medium* and *tight* electron identification algorithms outlined in chapter 2.3.

4.4 Event and Candidate Selection

The event selection for the $J/\psi \to e^+e^-$ analysis was chosen to be as analogous as possible to the $J/\psi \to \mu^+\mu^-$ case. Accordingly, the events were required to:

- have been recorded while the detector was operating to a high standard and providing good quality data. This judgement was based on a set of criteria recommended by the e/γ combined performance working group and was evaluated for each lumiblock. Aside from global flags common to all physics objects (e.g. LHC providing stable beams), more specific requirements are made on detectors specifically implemented in the reconstruction and measurement of electrons, such as the Inner Detector (ID) and calorimeter systems. As well as this, a flag is checked on an event-by-event basis to exclude events in which there were any significant noise bursts or data integrity issues in the LAr calorimeter.

- pass the nominated trigger, *EF_2e5_tight_Jpsi* or *EF_e9_tight_e5_tight_Jpsi* (note that separate analyses were conducted for each individual trigger).

- have at least one primary vertex with at least three tracks, to veto cosmic events. Unlike in the muon case, no quality cuts (such as number of Pixel/SCT hits) were placed on these tracks as this information was not easily accessible in the data format employed and was not deemed to be of vital importance. This is particularly true in 2011 data, where the increased instantaneous luminosity means that there are frequently many primary vertices in each bunch crossing.

Within events which passed the above selection, a search was conducted for electron objects which:

1. had author 1 or 3, which is to say that they were created with the standard electron reconstruction algorithm rather than those for soft (low E_T) or forward ($|\eta|$ >2.47) electrons.

2. fell within the *acceptance* of the detector. This term will be properly defined in Section 4.6.1, but for now note that this means the electrons must have

   - (a) $|\eta| < 2.47$ (be in the barrel of the Inner Detector), $|\eta| \notin [1.37, 1.52]$ (be out of the transition region of the electromagnetic calorimeter between barrel and endcap), and
   - (b) have E_T above the threshold of the particular trigger in question[6].

   In principle, the acceptance defines where the electrons can be reconstructed and triggered with reasonable efficiency.

---

[6] This is slightly more complicated for asymmetric triggers such as *EF_e9_tight_e5_tight_Jpsi* , and requires looking at both electrons simultaneously.

3. had a track with at least 1 hit in the Pixel detector and 7 hits in the SCT.
4. passed the *Object Quality* (OQ) flag. This ensures that the individual detector elements employed in the reconstruction of the electron in question were operational during the event under analysis.

At this stage the electrons were tagged as being *reconstructed*. Finally, to ensure a low fake rate[7] they were required to be

5. *identified* as *tight* electrons (see chapter 2.3 for an explanation of these terms).

Any electrons remaining were at this point paired into J/ψ candidates, each of which with:

6. opposite sign electrons.
7. the electrons matching the trigger electron objects. The main electron triggers at the EF level are track-based, and as such a match to the tracks of the selected (offline) electron objects was appropriate. This was implemented by ensuring that the $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$ between the tracks being compared was < 0.05 (typically, this value is 0.15 for single electron triggers, but in this case the two electrons from the J/ψ can have an opening angle of this order and a more stringent match was necessary. The value of 0.05 was chosen as a compromise between the recommended single electron value and that used in the muon analysis, 0.005).
8. a successfully fitted e^+e^- vertex with $\chi^2 < 200$. This was expected to have very little effect on real $J/\psi \to e^+e^-$ decays, but should exclude some false candidates.

Electron objects are reconstructed using information from both the ID and the LAr (see chapter 2.3) and so the calculation of the electron four-vector can be carried out in several ways. Unless otherwise stated, the *uncombined* four-vector was used for the electrons as (again) this was the recommendation of the e/γ group for physics analyses. Since the electrons used here necessarily pass a track quality selection, this essentially means that the energy was taken to be that measured by the LAr calorimeter and the (η, ϕ) were taken from the track. This choice for the definition of the electron four-vector was important as it potentially affected whether or not a particular electron passed the acceptance cuts, and strongly influenced the shape of resulting invariant mass distributions.

---

[7] The fake rate is defined as the fraction of electron objects that do *not* correspond to real electrons.

Table 4.2: The number of events (top section) and J/ψ candidates (bottom section) that passed each particular selection criterion for the two triggers used in this analysis. An explanation of each of these requirements is given in Section 4.4.

	EF_2e5_tight_Jpsi	***EF_e9_tight_e5_tight_Jpsi***
	Event Selection	
Detector Operation + Trigger	23570	139705
Primary Vertex	23570	139705
	Candidate Selection	
Total Electron Pairs	248162	1668178
Electron Author	82800	573381
Electron Acceptance	40427	211540
Electron Track Quality	31137	164562
Electron Object Quality	30945	163635
Electron Tight Id.	10448	55617
Electrons Opposite Sign	9577	52291
Electrons Trigger-Matched	9454	51735
Vertex fit with $\chi^2 < 200$	9375	51539

4.5 Observation of $J/\psi \rightarrow e^+e^-$ in Data

The number of events and, subsequently, candidates after each of the selection cuts outlined in the previous section is shown for each trigger in Table 4.2. As expected, the effect of the primary vertex requirement was insignificant (in fact, every triggered event under good detector conditions passed), as was the restriction on the χ^2 of the $J/\psi \rightarrow e^+e^-$ vertex fit and the electron object quality, which each removed less than 0.9% and 0.6% of candidates, respectively. Of the events remaining after the selection, only a very small fraction ($\sim 0.2\%$) contained two or more candidates.

The invariant mass distributions of the final $J/\psi \rightarrow e^+e^-$ candidates in the range 1–5 GeV are shown in Figure 4.3 for each trigger. In comparison with a similar result from the muon analysis (e.g. Figure 4.1a), the shape of the peak for electrons is similar but the width is far larger (which is a result of the finite resolution of energy measurements made by the LAr calorimeter and the uncertainty of the track direction parameters, which are used to construct the electron four-vectors). The position of the peak is also lower than the PDG value of the J/ψ mass (3096.9 MeV), and this difference is quantified in Section 4.8 using fits to the distributions. As in the muon case, there is also evidence for a secondary peak attributed to the $\psi(2S)$ at ∼3700 MeV, though it is less distinct in this case because of the poorer resolution.

In contrast with the $\mu^+\mu^-$ channel, the background here is more pronounced and complicated by the tail on the low-mass side of the peak, due to energy losses of the electrons through bremsstrahlung. The main sources of this background are semi-leptonic c and b quark decays, and the combinatorial background which results from J/ψ candidates constructed using two electrons unassociated to one another or that did not correspond to real particles[8]. This *fake rate* is considerably larger for electrons than for muons, and explains the larger level of background in the former case.

Figure 4.4 displays the distribution of $J/\psi \to e^+e^-$ candidates in the signal region ($2.5 < M_{e^+e^-} < 3.5$) in terms of their rapidity, y, and transverse momentum, p_T. At J/ψ transverse momenta above about 10 GeV, the decaying electrons tend to be well collimated with one another and the p_T vector of the J/ψ, meaning the structure in these distributions can be understood in terms of the electron acceptance cuts. For example, the lack of candidates in a strip about $y = 1.4$ and the end point at $y = 2.5$ correspond to the restrictions of $\eta \notin [1.37, 1.52]$ and $\eta < 2.47$, respectively, placed on electrons. Likewise, the threshold on the p_T reflects the E_T thresholds imposed on the electrons by the respective triggers. As the production spectrum peaks far below this (in fact, below 5 GeV), the J/ψ collected by each trigger strongly populate regions directly above this threshold. This allowed a natural separation in the kinematic range with which each trigger was implemented for the cross-section measurement; *EF_2e5_tight_Jpsi* for the range $p_T < 16$ GeV and *EF_e9_tight_e5_tight_Jpsi* for $p_T > 16$ GeV[9].

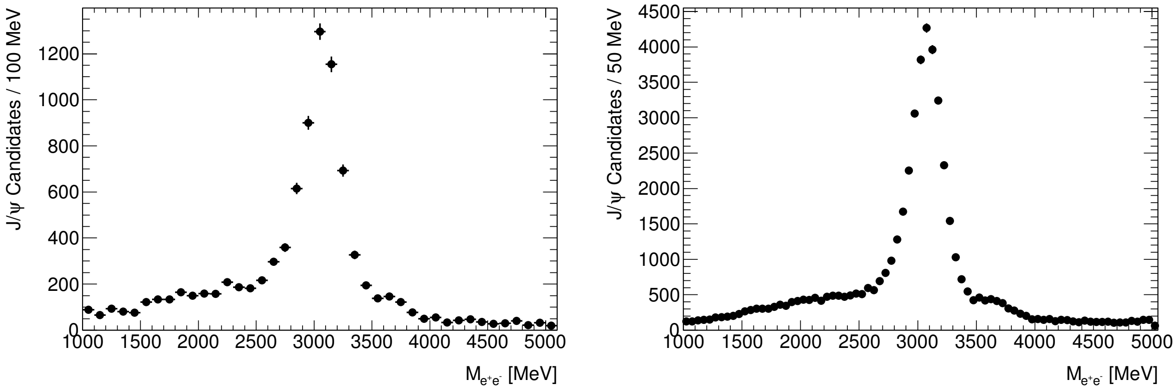
4.6 Calculation of the $J/\psi \to e^+e^-$ Candidate Weight

Of the large number of J/ψ that were created and decayed through the electron channel during the data taking periods of this analysis, only a small fraction were detected and passed all the event and candidate selection requirements. However, the interesting physical quantity being measured here is the *production* cross-section, so an accurate knowledge of this fraction is essential. The event selection was accounted for in the calculation of the total integrated luminosity, which only included times when the detector was operating to an appropriate standard and was modulated by the prescale of the trigger in question (as shown in the previous section, the primary vertex constraint removed no events at all and need not be considered). It was assumed that the most important of the remaining criteria were the reconstruction, identification and triggering of the two electrons, which each have an associated inefficiency. This essentially amounts to presuming that the e^+e^- vertex cuts removed an insignificant number of

---

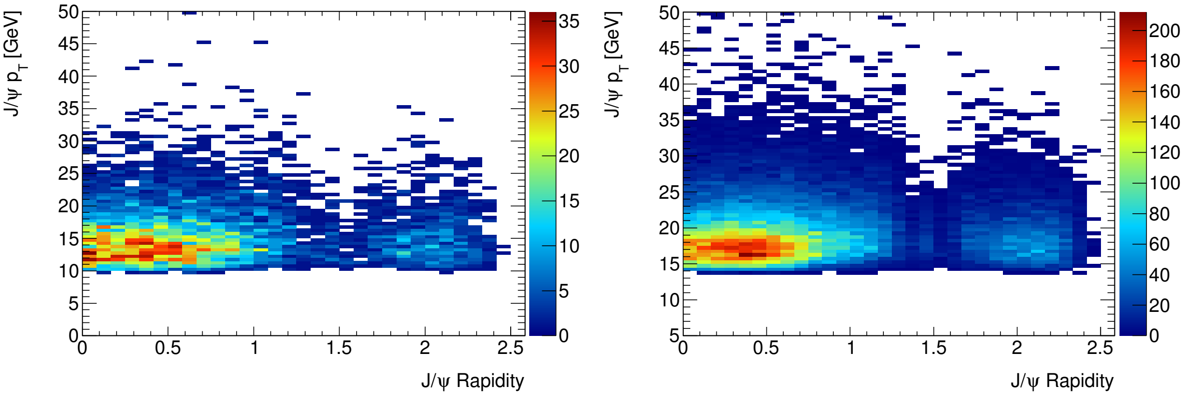
[8] For example, photons can be misidentified as electrons when their calorimeter cluster is matched to an unrelated track.

[9] This complementarity was briefly mentioned earlier, and justifies the use of these two triggers in particular.



(a) *EF_2e5_tight_Jpsi* (b) *EF_e9_tight_e5_tight_Jpsi*

Figure 4.3: The invariant mass distribution of the $J/\psi \to e^+e^-$ candidates for the two triggers. The width is far larger than for the $J/\psi \to \mu^+\mu^-$ channel, and the peak is slightly lower than the PDG value. In both cases there is evidence for the appearance of the $\psi(2S)$ at ∼3700 MeV, as expected.



(a) *EF_2e5_tight_Jpsi* (b) *EF_e9_tight_e5_tight_Jpsi*

Figure 4.4: The distribution of J/ψ candidates (with $2.5 < M_{e^+e^-} < 3.5$ GeV) in the $y - p_T$ plane for each trigger. There is an obvious lower p_T cut-off caused by the trigger thresholds, and the development in y is due to the pseudo-rapidity acceptance cuts placed on the electrons. The distributions are mainly populated in central regions at low transverse momentum.

real $J/\psi \rightarrow e^+e^-$ candidates. Since this has already been shown to remove $< 1\%$ of all candidates, this assumption is justified (especially since the study was not framed as a precision measurement).

The efficiencies, and therefore the probability that a particular $J/\psi \rightarrow e^+e^-$ decay is recorded, depends on the kinematics of the J/ψ and of both electrons. It can be calculated in many ways, but the following form was chosen in analogy to the $J/\psi \rightarrow \mu^+\mu^-$ approach:

$$P = \mathcal{A} \cdot \varepsilon_{\text{reco.}} \cdot \varepsilon_{\text{id.}} \cdot \varepsilon_{\text{trig.}} \tag{4.5}$$

The first term, $\mathcal{A}$ (the *acceptance*), represents the fraction of J/ψ with a particular (p_T, y) that have both electrons *reconstructable* as defined in terms of the electron pseudo-rapidity and transverse energy requirements outlined above. The efficiency terms ($\varepsilon_{\text{reco.}}$, $\varepsilon_{\text{id.}}$ and $\varepsilon_{\text{trig.}}$) are the probability that these electrons will furthermore be reconstructed, identified and triggered, where each efficiency is defined with respect to the subset of electrons passing the previous step (e.g. $\varepsilon_{\text{id.}}$ is the identification efficiency for J/ψ in which both electrons have already been reconstructed). For each J/ψ candidate, this probability (Equation 4.5) was calculated and its inverse applied as a weight to account for those not recovered by the analysis.

In the following sections, each of the terms above are discussed in more detail and the methods used to evaluate them are described. As will be shown, the problems and challenges associated with this *weighting* method present the greatest difficulties to the use of electrons in quarkonium physics.

4.6.1 Acceptance

Some of the electrons from J/ψ decays will inevitably be emitted with trajectories and momenta that make them difficult or even impossible to detect. The simplest example occurs when either of the electrons is produced along the beam pipe (i.e. with very high η) and does not cross any of the ATLAS detector elements. Other acceptance conditions may be a result of the electron reconstruction and identification software, or the trigger requirements. To some extent these latter conditions are arbitrary, but are generally chosen to exclude regions where the corresponding efficiencies or resolutions become very poor. For electrons, the acceptance is most dependent on the pseudo-rapidity and transverse energy for which the following conditions were enforced in this analysis:

- **Pseudo-rapidity.** To be reconstructed and identified as *tight*, an electron must have a corresponding track with a certain number of silicon hits. Given that the inner detector extends to $|\eta| = 2.47$, this was an obvious restriction. The second condition arose as a result of the physical transition or 'crack' region of $1.37 < |\eta| < 1.52$ between the barrel and end-cap sections of the LAr calorimeter, where the electron efficiencies are known to be poor. In summary, it was required that

$$|\eta| < 2.47, |\eta| \notin [1.37, 1.52]$$

- **Transverse Energy.** Electrons from the standard reconstruction algorithm can (technically) have transverse energies as low as 2.5 GeV, though in the case of this analysis the electrons were required to pass dielectron triggers which placed cuts at 5 and/or 9 GeV. The trigger thresholds therefore provided natural and logical acceptance conditions on transverse energy.

In the rest frame of the J/ψ, the electrons decay back-to-back with an energy of $E = 1/2 \cdot M_{J/\psi} = 1.55$ GeV and an angular distribution which is dependent only on the spin and polarisation of all three particles[10]. Given this information, the angular distribution and energy spectrum of electrons in any other reference frame can be determined by performing a Lorentz boost along the direction of J/ψ momentum. This allows the fraction (or *acceptance*) of J/ψ at a given rapidity and transverse momentum with both daughter electrons in the above acceptance regions to be calculated.

The invariance of helicity[11] under rotations and boosts along $\hat{\mathbf{p}}$ provides a simple formalism for extracting the (relative) quantum mechanical transition amplitude for the electrons from the J/ψ to decay in a particular direction within the rest frame (for details on the *helicity formalism* see [60, 61]). The resultant angular distribution is given by:

$$\frac{d^2N}{d\cos\theta^* d\phi^*} \propto 1 + \lambda_\theta \cos^2\theta^* + \lambda_\phi \sin^2\theta^* \cos 2\phi^* + \lambda_{\theta\phi} \sin 2\theta^* \cos\phi^* \tag{4.6}$$

where θ^* is the polar angle of the positron with respect to the J/ψ momentum vector, and ϕ^* is the azimuthal angle measured from the J/ψ production plane in the lab frame. The λ coefficients are related to the components of the polarisation density matrix for the J/ψ,

$$|J/\psi\rangle = A_0|0\rangle + A_+|+1\rangle + A_-|-1\rangle \tag{4.7}$$

The polarisation fractions are not known *a priori* and so constitute a theoretical uncertainty. Of the large number of physically possible combinations of these fractions, the $J/\psi \to \mu^+\mu^-$ analysis concluded that the following five scenarios represented the extrema in acceptance:

---

[10] The polarisation of the electrons is typically not measured, and so this degree of freedom is averaged over.

[11] Helicity is defined as the projection of the spin vector onto the momentum direction. Calling $\mathbf{S}$ the spin and $\mathbf{p}$ the momentum, the helicity is given by $h = \mathbf{S} \cdot \hat{\mathbf{p}}$.

- **FLAT.** Isotropic distribution with $\lambda_\theta = \lambda_\phi = \lambda_{\theta\phi} = 0$. In keeping with the muon case, this was chosen as the central hypothesis.
- **LONG.** Longitudinal alignment with $A_0 = 1, A_+ = A_- = 0$ or $\lambda_\theta = -1, \lambda_\phi = \lambda_{\theta\phi} = 0$.
- **TRP0.** Transverse alignment with $A_0 = A_\mp = 0, A_\pm = 1$ or $\lambda_\theta = 1, \lambda_\phi = \lambda_{\theta\phi} = 0$.
- **TRPP.** Transverse alignment with $A_0 = 0, A_+ = A_- = \pm 1/\sqrt{2}$ or $\lambda_\theta = 1, \lambda_\phi = 1, \lambda_{\theta\phi} = 0$.
- **TRPM.** Transverse alignment with $A_0 = 0, A_+ = -A_- = \pm 1/\sqrt{2}$ or $\lambda_\theta = 1, \lambda_\phi = -1, \lambda_{\theta\phi} = 0$.

In practice, the acceptance fractions were calculated using a *Monte Carlo* (MC) computation. For each simulated decay, a value of θ^* and ϕ^* were randomly chosen in the rest frame, and an appropriate event weight was determined from Equation 4.6. The entire system was then boosted and rotated to the desired rapidity and transverse momentum, and the final electrons' $(E_T, |\eta|)$ were tested against the acceptance conditions. The fractions with both electrons passing these conditions were determined by repeating this 10,000 times in each bin (widths $\Delta y = 0.0125$ and $\Delta p_T = 0.167$ GeV).

Figure 4.5 shows the results for the *EF_e9_tight_e5_tight_Jpsi* trigger in each of the different polarisation scenarios, along with the *FLAT* and *LONG* acceptance maps for the *EF_2e5_tight_Jpsi* trigger for comparison. The most obvious feature here is that the trends are similar to those observed in the corresponding distribution of candidates in data, which justifies earlier claims that these were a result of the electron acceptance conditions. In particular, the lower p_T threshold is clearly determined by the trigger and there is an obvious dip in acceptance around the crack region. The *LONG* and *TRPP* or *TRPM* polarisation scenarios present the greatest differences in acceptance from the *FLAT* distribution, and are later used to calculate the spin alignment systematic uncertainty.

There were a few assumptions made in this calculation, some of them having important consequences. The ATLAS solenoidal magnetic field causes electron trajectories to be essentially helical, such that the original momentum vector (which is what is being used in the MC simulation to test against the acceptance conditions) is not the same as the final momentum vector. The effect of this was measured and found to be insignificant, since the bending occurs in the ϕ plane and the change in η is only due to the increased time of flight as compared to a linear trajectory. Another minor issue concerns the assumption that all decays occur at $x = y = z = 0$. In reality, a significant fraction of decays will be non-prompt and therefore will not occur there (the primary hard interaction is also typically slightly offset from the origin). If the distance of the decay from the origin is δ and the distance to the edge of the Inner Detector is ℓ, then to first order the perturbation in θ is $\sim \delta/\ell$, which is always small and can be ignored for our purposes. The remaining issues centre around the presumption that the reconstructed electron four-vectors are an accurate representation of their true (or *truth*) values. This is typically incorrect, since electrons tend to lose energy through material interactions as they traverse the detector (the strongest mechanism at these energies being bremsstrahlung). In fact, the measurement of electron kinematic quantities such as energy and direction tend to have large resolutions and be biased. These comments are equally valid for the J/ψ four-vectors, which are simply the addition of the electron four-vectors. *Migration* effects such as these have a significant impact on the acceptance, and the problems they cause are explored in Section 4.7.

4.6.2 Reconstruction Efficiency

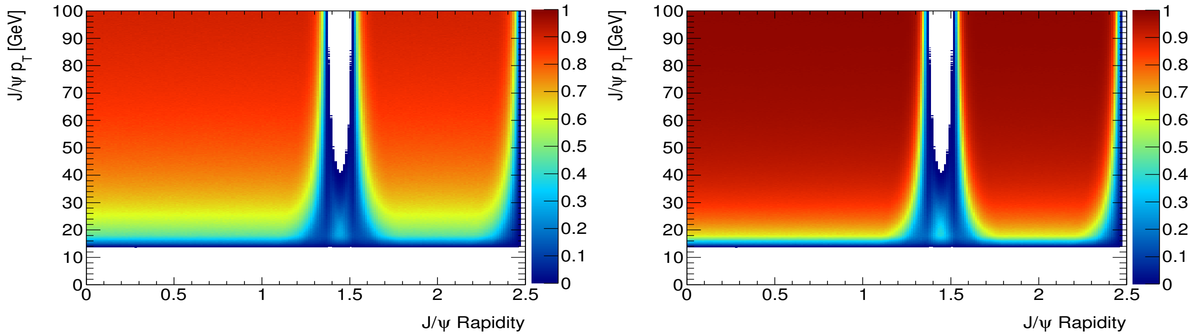
To calculate the reconstruction efficiency, it was necessary to establish a method to match generated electrons with corresponding reconstructed electron objects. The *MCTruthClassifier* algorithm examines the simulation information of each hit on a reconstructed track to ascertain which (generated) particle was most probable to have contributed to the majority of them. A generated electron was said to be matched if there was a reconstructed track linked to it in such a way. In the case that there were two matches, the one closest in terms of $dR = \sqrt{d\phi^2 + d\eta^2}$ was chosen. Very rarely, there were cases where the same track was used to construct two different electrons (i.e. two different calorimeter clusters), meaning the dR were equal and it was necessary to select the one with transverse energy closest to the truth value.

For this analysis, $\varepsilon_{\text{reco.}}$ was defined as the probability for both electrons decaying from a J/ψ to have a matching electron object reconstructed by the standard algorithm (i.e. have author 1 or 3) with at least 1 Pixel and 6 SCT hits. The electrons were also required to be reconstructed with the appropriate sign, fall into the acceptance regions defined above, and pass the Object Quality flag.

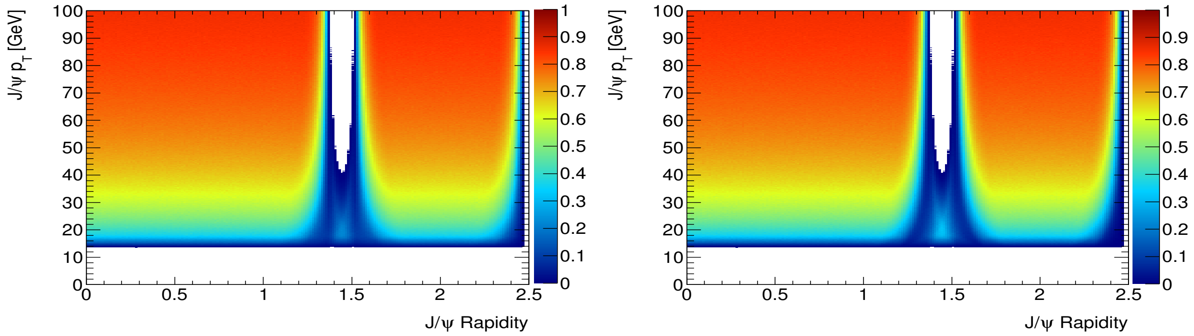
The value of the reconstruction efficiency was evaluated separately for prompt and non-prompt[12] J/ψ using the appropriate MC samples, and binned in terms of J/ψ ra-

---

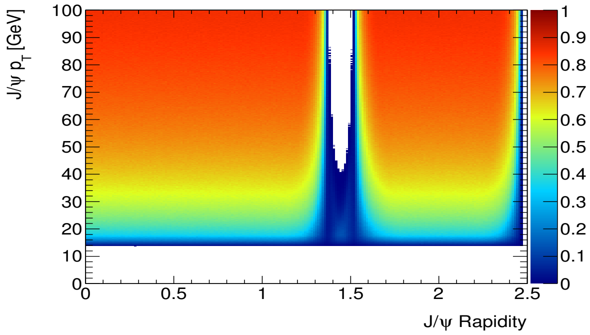
[12] To ensure that only prompt (non-prompt) J/ψ were used for the calculation of the prompt (non-prompt) efficiencies, the rare events with more than one $J/\psi \to e^+e^-$ decay were simply excluded (these additional decays can be either prompt or non-prompt).



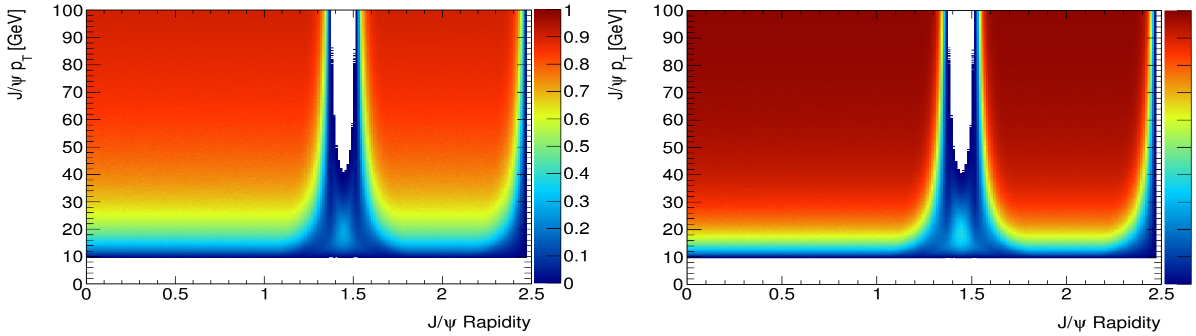
(a) *FLAT*, *EF_e9_tight_e5_tight_Jpsi* (b) *LONG*, *EF_e9_tight_e5_tight_Jpsi*



(c) *TRP0*, *EF_e9_tight_e5_tight_Jpsi* (d) *TRPP*, *EF_e9_tight_e5_tight_Jpsi*

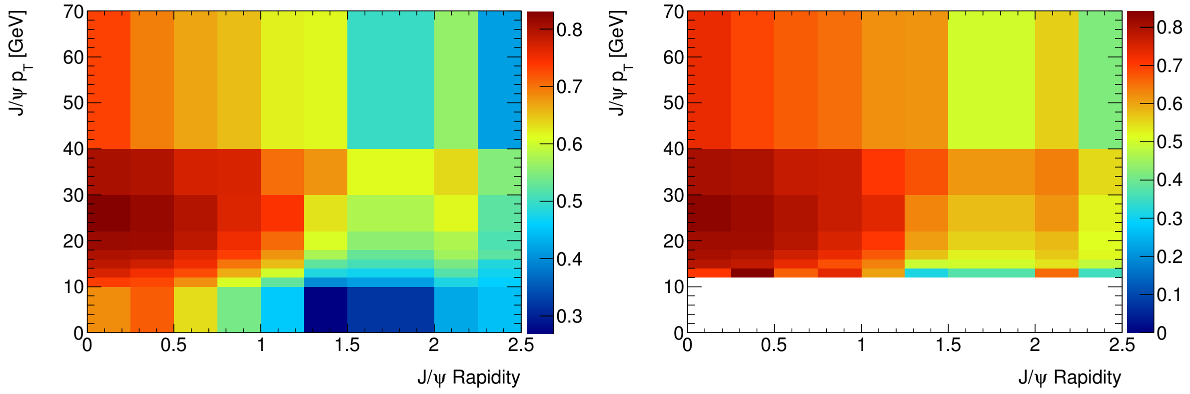


(e) *TRPM*, *EF_e9_tight_e5_tight_Jpsi*

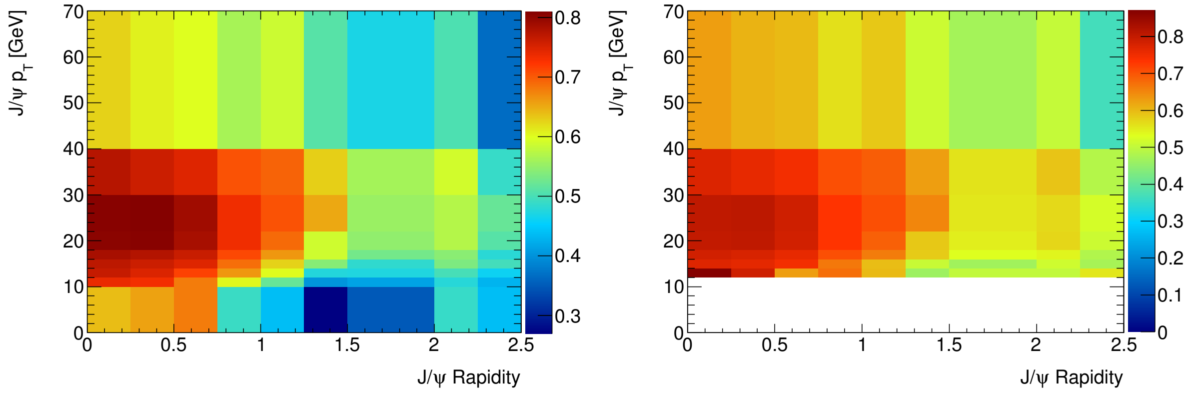


(f) *FLAT*, *EF_2e5_tight_Jpsi* (g) *LONG*, *EF_2e5_tight_Jpsi*

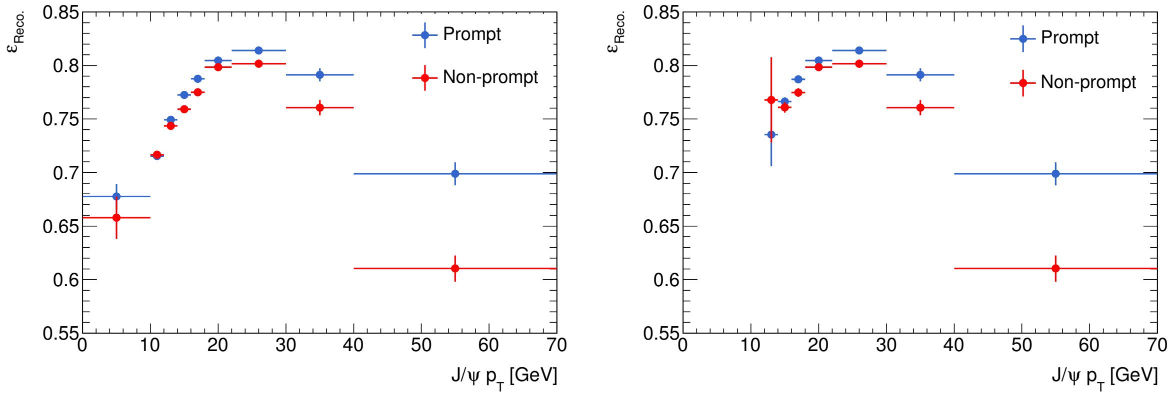
Figure 4.5: Acceptance maps for each of the five extreme polarisation scenarios for the *EF_e9_tight_e5_tight_Jpsi* trigger transverse energy thresholds (top), and for *FLAT* and *LONG* polarisations with the *EF_2e5_tight_Jpsi* trigger thresholds (bottom).



(a) Prompt, *EF_2e5_tight_Jpsi* (b) Prompt, *EF_e9_tight_e5_tight_Jpsi*



(c) Non-prompt, *EF_2e5_tight_Jpsi* (d) Non-prompt, *EF_e9_tight_e5_tight_Jpsi*



(e) *EF_2e5_tight_Jpsi* (f) *EF_e9_tight_e5_tight_Jpsi*

Figure 4.6: The J/ψ reconstruction efficiency maps for the *EF_2e5_tight_Jpsi* and *EF_e9_tight_e5_tight_Jpsi* trigger conditions (left and right, respectively), with a comparison of prompt and non-prompt efficiencies as a function of E_T (bottom). The differences between the prompt and non-prompt populations of J/ψ is a significant issue (see text).

pidity and transverse momentum for consistency with the acceptance calculations. As the definition of *reconstructed* depends on the acceptance conditions, different efficiencies were required for each of the two triggers being used. The results (Figure 4.6) show that the reconstruction efficiencies generally reach a maximum of ~0.8 for J/ψ produced in a central direction with a p_T of ~22-30 GeV. At higher or lower transverse momenta, or in more forward regions, the efficiency is significantly lower. The variable binning here was chosen to match that for which the differential cross-section measurement was performed in the muon analysis and also this analysis (Section 4.8).

Whilst the slight variations in efficiencies observed between the two trigger scenarios were expected (the electron reconstruction efficiency is dependent on E_T), the significant difference between the prompt and non-prompt (Figures 4.6e and 4.6f) J/ψ production channels was concerning. The non-prompt efficiencies are generally lower because of the contamination of the surrounding regions of the Inner Detector and LAr calorimeter by the other hadronic decay products (e.g. $\pi^\pm$, ϕ, $K^\pm$ etc.) of the B-hadron parent. To understand why this is an issue, consider the question of which efficiency (prompt or non-prompt) to apply on a candidate-by-candidate basis; there is no easy way of determining the answer to this (though one could attempt to make a decision based on the pseudo-proper time, for example). This is clearly an area in which the electron analysis presents a significant challenge that did not exist in the muon case. This is discussed in more detail in the following sections.

An alternative approach taken to evaluating $\varepsilon_{\text{reco.}}$ relies on the assumption that the reconstruction efficiency of each electron is independent of the other. This being the case, the probability of reconstructing both electrons is simply the product of each of the individual efficiencies (they are said to have factorised):

$$\begin{aligned} P(\text{both electrons reconstructed}) &= P(e^+\text{reco.})P(e^-\text{reco.}|e^+\text{reco.}) \\ &= P(e^+)P(e^-) \qquad\qquad \text{(Independence)} \\ \Rightarrow \varepsilon_{\text{reco.}} &= \varepsilon_{\text{reco.},e^+}\varepsilon_{\text{reco.},e^-} \end{aligned} \tag{4.8}$$

The individual electron efficiencies averaged over electrons and positrons[13] are shown in Figure 4.7. As expected, there is a sharp ‘turn on’ curve with increasing E_T (figure 4.7c) which reaches a plateau of ~0.85 at around 10 GeV. The second clear feature is that the efficiency is highest in the central regions and poor in the LAr calorimeter transition zone and beyond the extent of the Inner Detector (η=2.47), particularly for low E_T. These

---

[13] There were likely to be only very slight differences between e^+ and e^- reconstruction efficiencies and so this was ignored. In a full analysis this difference would need to be accounted for by either using separate efficiencies or incorporating it as a systematic uncertainty.

features justify the choice of acceptance conditions placed on the electrons. Again, there is an obvious difference in the prompt and non-prompt efficiencies, which favours the prompt J/ψ at high E_T (presumably due to the increased collimation of the B-hadron decay products in the non-prompt sample). The η projection in Figure 4.7d seems to indicate that the non-prompt efficiency is generally higher, but this must be interpreted with caution as the average in the profile distribution has been taken over the entire transverse energy range, which is dominated by the low end of the spectrum.

Many of these trends help to understand those in the case of the total J/ψ reconstruction efficiencies (Figure 4.6), such as the greater efficiency in the barrel region and the general increase with E_T. However, the fact that the J/ψ efficiencies peak at an intermediate p_T and then drop off at higher values is a strong indication that the factorisation assumption is not valid for this region of phase space.

Unfortunately, the large variation in individual electron reconstruction efficiencies with η and E_T implies that the J/ψ reconstruction efficiencies are dependent on the spin-alignment[14]. Unlike the acceptance, where the difference in J/ψ polarisation can determine whether or not the candidate was included in the analysis, the effect here is to change the efficiency by some factor. Since this is (in some sense) a variation-on-a-variation, it was treated as a second-order effect and the reconstruction efficiencies were only calculated for the *FLAT* angular distribution hypothesis.

As a final comment, it is worth noting that (as for the acceptance) the reconstruction efficiencies were necessarily[15] parameterised in terms of *truth* kinematic variables, which is problematic due to the *migration* effects explained above.

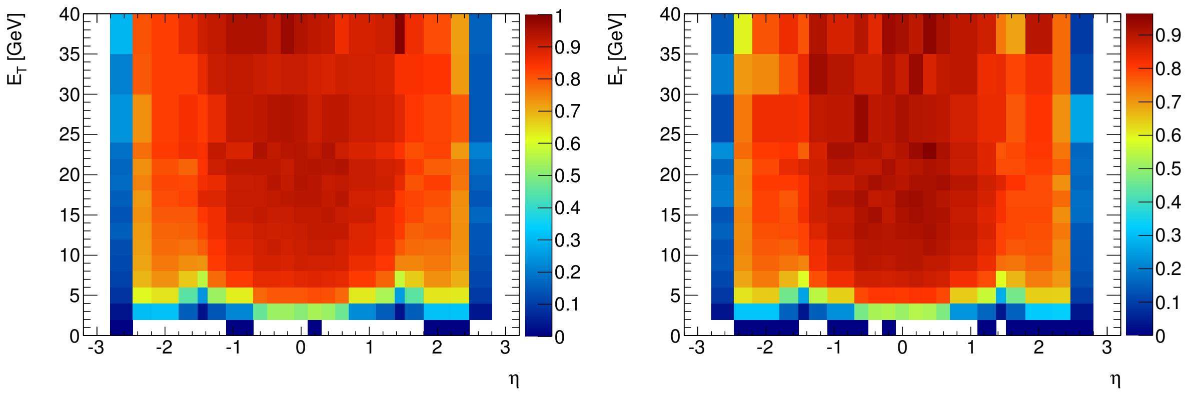
4.6.3 Identification Efficiency

The J/ψ identification efficiency is the probability that, given the daughter electrons were reconstructed, they furthermore both pass the *tight* identification selection. The value of $\varepsilon_{\text{id.}}$ was computed in an obvious manner for the prompt and non-prompt J/ψ under each trigger condition using the MC samples (the difference in what constitutes a reconstructed J/ψ for each trigger is propagated through to the identification efficiencies). The efficiencies calculated in this manner are, as for the equivalent reconstruction efficiencies, slightly dependent on the J/ψ spin-alignment. Again, only the values assuming

---

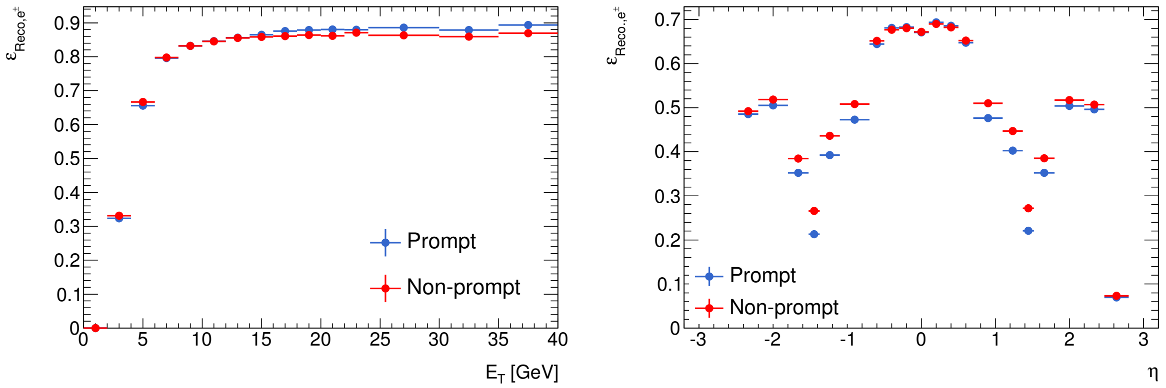
[14] Electrons contributing to the J/ψ reconstruction efficiency of a particular $p_T - |y|$ efficiency bin will have polarisation-dependent angular and momentum distributions. Given that the electron reconstruction efficiency is significantly different across the available phase space, differences in polarisation will therefore lead to differences in the efficiency calculated for each particular bin.

[15] In the case that the electron was not reconstructed at all, the truth values are the only ones available to use.



(a) Prompt

(b) Non-prompt



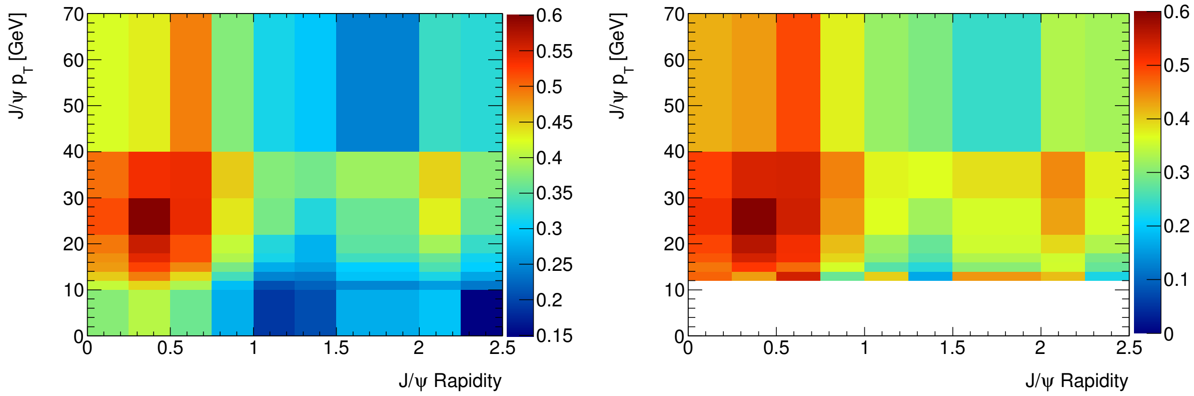
(c) Prompt vs. Non-prompt (E_T)

(d) Prompt vs. Non-prompt (η)

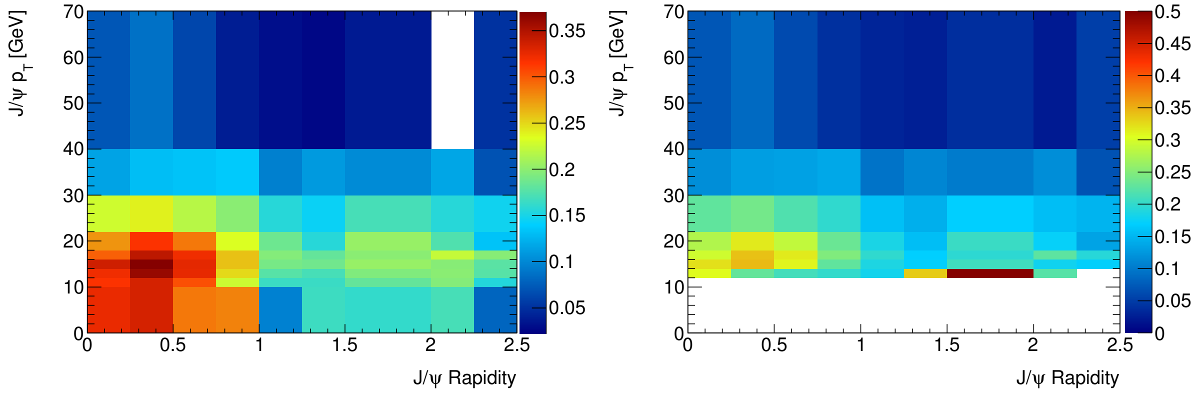
Figure 4.7: The individual electron reconstruction efficiency maps for those from prompt and non-prompt J/ψ (top), and the projection of these samples onto E_T and η (bottom). Again, there is a noticeable difference in the efficiencies of prompt and non-prompt which favours the former at high transverse energies.

an isotropic distribution were produced, and differences between polarisation scenarios was assumed negligible.

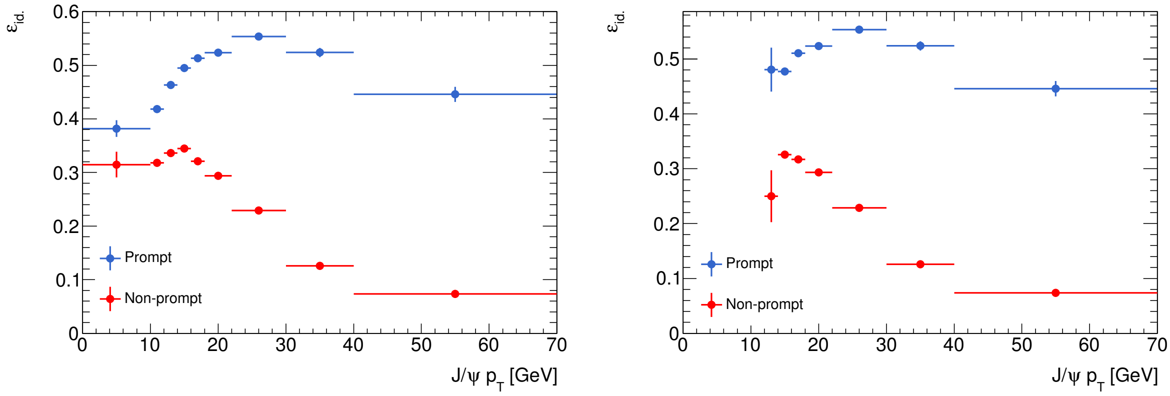
For the prompt sample, the identification efficiencies (Figures 4.8a and 4.8b) rise to a peak of ~0.6 between 22-30 GeV and fall off in the forward regions in a similar way to the reconstruction efficiencies (though the latter are higher in scale). On the other hand, the non-prompt J/ψ identification efficiencies (Figures 4.8c and 4.8d) are much lower and peak at a smaller values of p_T. This can largely be understood by considering the impact of the by-products of the B-hadron decay on the isolation requirement of the tight identification algorithm. For highly boosted decays the trajectories of these particles lie close to those of the electrons, which explains the poor efficiency at high



(a) Prompt, *EF_2e5_tight_Jpsi* (b) Prompt, *EF_e9_tight_e5_tight_Jpsi*



(c) Non-prompt, *EF_2e5_tight_Jpsi* (d) Non-prompt, *EF_e9_tight_e5_tight_Jpsi*



(e) *EF_2e5_tight_Jpsi* (f) *EF_e9_tight_e5_tight_Jpsi*

Figure 4.8: The J/ψ identification efficiency maps for the *EF_2e5_tight_Jpsi* and *EF_e9_tight_e5_tight_Jpsi* trigger conditions (left and right, respectively), with a comparison of prompt and non-prompt efficiencies as a function of E_T (bottom). The differences between the two samples of J/ψ is much more pronounced in comparison with the reconstruction efficiencies.

transverse momentum.

As in the case of reconstruction efficiency, the factorisation assumption can be made for identification efficiencies so that $\varepsilon_{\text{id.}} = \varepsilon_{\text{id.},e^+}\varepsilon_{\text{id.},e^-}$. The individual tight electron efficiencies for prompt and non-prompt J/ψ decays are shown in Figure 4.9. The prompt efficiencies reach a plateau of around 70% and remain essentially at this value with a slight drop at larger E_T. In contrast, the non-prompt efficiencies are very low and become increasingly worse at higher transverse energies. These observations demonstrate a good correlation to the J/ψ identification efficiencies described above. The distribution with η shows the expected drop in the transition region of the calorimeter and the cut off at $|\eta| = 2.5$. It is not as obvious whether or not the factorisation assumption is valid for here by simply comparing the efficiencies in Figures 4.8 and 4.9, and hence required further investigation (see Section 4.7).

The magnitude of the differences between prompt and non-prompt efficiencies is, again, a crucial issue, but in this case there are at least no *migration* effects; the identification efficiencies were calculated with respect to the subset of electrons that were reconstructed, allowing the efficiencies to be binned by the measured kinematic values (for example, in Figure 4.8a the rapidity and p_T of the J/ψ are the values obtained from the measured electron four-vectors).

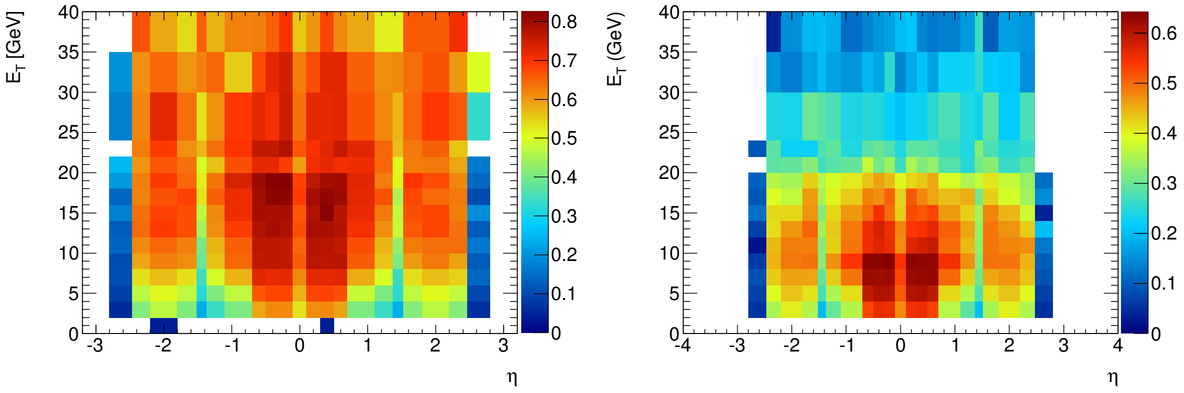
These identification efficiencies should ideally be replaced by the values measured from data to eliminate any dependence on the MC simulation. However, this measurement faces the same prompt/non-prompt issue[16] and at the time of writing the separation of the two efficiencies had not been achieved (instead, a weighted average reflecting the mixture of prompt/non-prompt in data has been measured).

4.6.4 Trigger Efficiencies

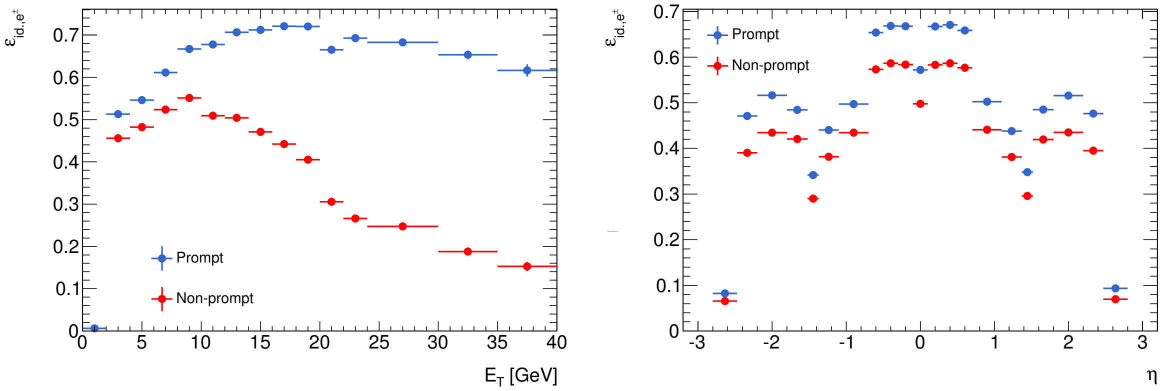
The trigger efficiency was defined with respect to identified J/ψ and was also measured with the MC samples. In this case, no individual electron trigger efficiencies were calculated. From the results of the previous section, particularly reconstruction efficiencies, evidence was already mounting against the validity of the factorisation assumption. Aside from this, dielectron triggers have far poorer acceptance and efficiency properties compared to single electron triggers, and foreseen applications to quarkonia analyses (Section 4.9.1) would generally involve cases where muons (or muons with a single electron) would be the main objects used for triggering. As such, although measuring the total trigger efficiency of the two electron objects as a combined system in real data presents technical

---

[16] i.e. the issue of how one determines, on a case-by-case basis, whether or not a particular J/ψ is prompt or non-prompt



(a) Prompt (b) Non-prompt



(c) Prompt vs. Non-prompt (E_T) (d) Prompt vs. Non-prompt (η)

Figure 4.9: The individual electron tight identification efficiency maps for those from prompt and non-prompt J/ψ (top), and the projection of these samples onto E_T and η (bottom). In comparison to the reconstruction efficiencies, the difference between the two populations of electrons is far greater here and always favours those from prompt J/ψ.

difficulties, the values from MC were considered a suitable placeholder for the purposes of this analysis.

The results (Figure 4.10) are much more uniform in rapidity than the previous efficiencies, and are peaked at lower transverse momenta (between 12−14 GeV for *EF_2e5_tight_Jpsi* and 18−22 GeV for *EF_e9_tight_e5_tight_Jpsi*). All the efficiencies here were binned by the reconstructed kinematic values to avoid migration issues. Considering that the calculation was with respect to J/ψ for which both electrons had already been tightly identified, the trigger efficiencies are very low for most of the p_T spectrum (even the peak values of ∼0.7-0.8 are poor). This inefficiency almost certainly occurs in the level two (L2) trigger stage where regions of interest are defined and the invariant mass requirement is enforced. The falling tail at high transverse momentum suggests that close proximity or overlap between two calorimeter clusters may be responsible, but further study would be required to confirm this. The lack of a plateau in this distribution is also consistent with the predicted lack of factorisation in the individual electron trigger efficiencies (where the values are known to approach a plateau above the thresholds). The prompt/non-prompt dependence is again present, though is not as statistically significant as for the identification efficiencies.

4.7 Closure Tests on the Weighting Method

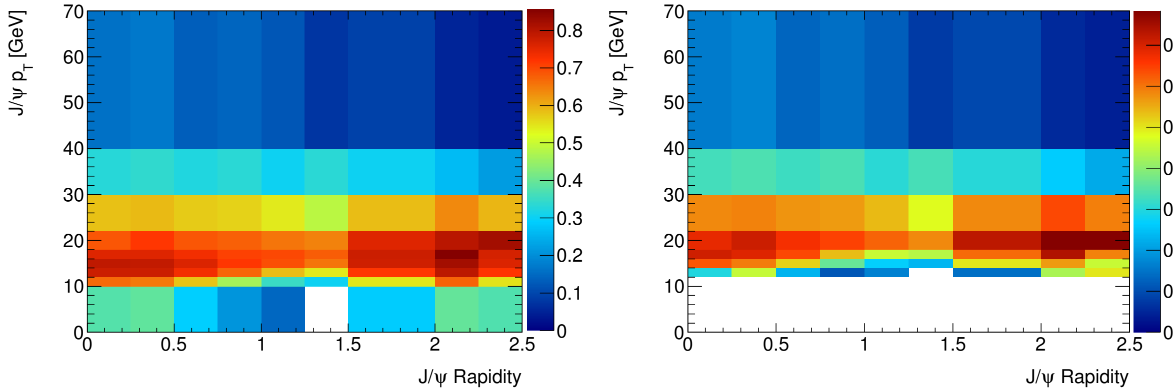
Before blindly applying the results of the previous section, it was important to confirm their validity and ensure that the weighting method did indeed return the true number of J/ψ produced. To do so, several *closure tests* were designed in order to compare the predictions made using this method with the corresponding true known values from the MC samples. As the efficiencies were vastly different for the prompt and non-prompt J/ψ, the test was performed separately for each. For all the closure tests, the FLAT polarisation hypothesis was assumed.

4.7.1 Testing the Weighting Method Logic

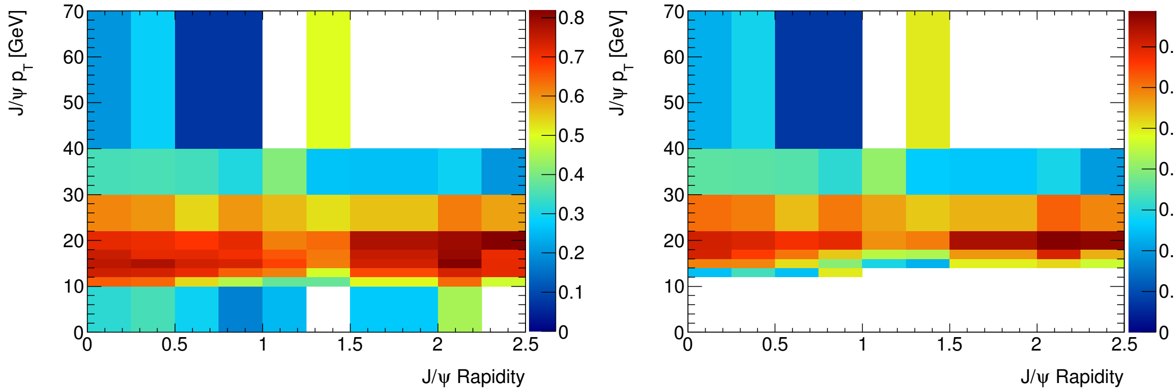
The first test was to determine whether the logic behind the weighting method was accurate and appropriate for electrons, and identify areas where this was not the case and the method failed. It was also intended to examine the validity of the electron efficiency factorisation assumptions introduced in the previous section.

The initial step was to reduce the analysis cuts to the subset which the terms in the weight should completely account for, which excludes the primary and e^+e^- vertex requirements. Then, the following *truth* quantities were defined and extracted from the MC samples:

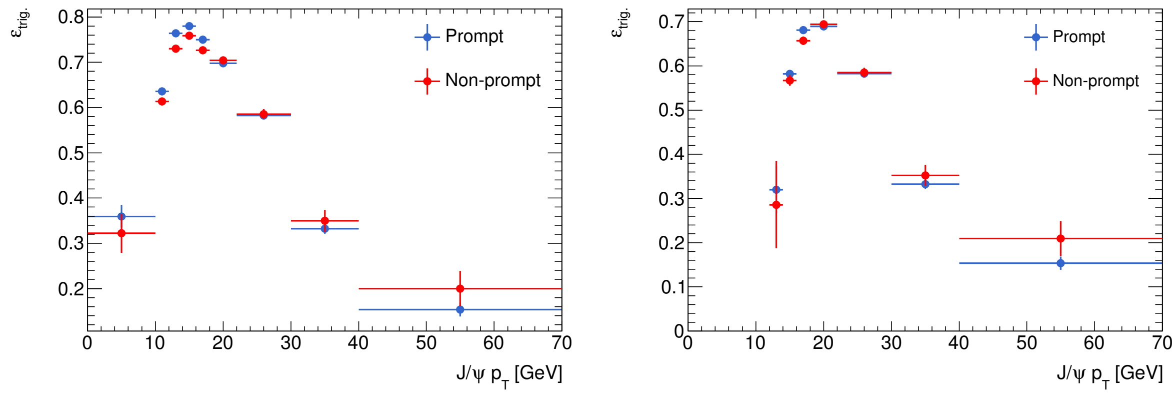
- $T_{\text{trig.}}$ = the true number of J/ψ which pass all of the reduced analysis selection criteria.
- $T_{\text{id.}}$ = the true number of J/ψ which pass the reduced analysis selections with the trigger requirement removed.
- $T_{\text{reco.}}$ = the true number of J/ψ which pass the reduced analysis selections with both the trigger and identification requirements removed.
- $T_{acc.}$ = the true number of J/ψ which have both electrons in the acceptance region.



(a) Prompt, *EF_2e5_tight_Jpsi* (b) Prompt, *EF_e9_tight_e5_tight_Jpsi*



(c) Non-prompt, *EF_2e5_tight_Jpsi* (d) Non-prompt, *EF_e9_tight_e5_tight_Jpsi*



(e) *EF_2e5_tight_Jpsi* (f) *EF_e9_tight_e5_tight_Jpsi*

Figure 4.10: The J/ψ trigger efficiency maps for *EF_2e5_tight_Jpsi* and *EF_e9_tight_e5_tight_Jpsi* (left and right, respectively), with a comparison of prompt and non-prompt efficiencies as a function of E_T (bottom). In this case the prompt efficiencies peak higher but have worse values at high E_T.

- $T_{prod.}$ = the total true number of J/ψ produced.

As the MC samples were being produced, events were filtered to ensure that they contained at least two leptons with transverse energies above 3 GeV and a pseudo-rapidity within the range $[-2.7, 2.7]$. The leptons in this case need not necessarily be from the J/ψ decay, and may even be muons. As such, the J/ψ sample is biased in a rather complicated way. To simplify matters for this test, an additional filter was applied to ensure that the two leptons were indeed the electrons from the J/ψ, which defines what is meant by $T_{prod.}$. Accordingly, the acceptance needed to be adapted such that it was calculated with respect to J/ψ which pass this initial filter. This will be referred to as the *biased* acceptance.

Now, the result of applying a weight to each of the N real[17] J/ψ passing all of the reduced analysis selections is given by:

$$N_{prod.} = \sum_{i=1}^{N} w_i = \sum_{i=1}^{N} (\mathcal{A}_i \cdot \varepsilon_{\text{reco.},i} \cdot \varepsilon_{\text{id.},i} \cdot \varepsilon_{\text{trig.},i})^{-1} \tag{4.9}$$

If the weighting method is performing perfectly, this number should precisely match $T_{prod.}$, the true number of produced J/ψ. The deviation of the closure fraction, defined as $C_{prod.} = N_{prod.}/T_{prod.}$, from unity then provides a measure of the extent of the failure of the weighting. Likewise, consider the following quantities:

$$\begin{aligned} N_{acc.} &= \sum_{i=1}^{N} (\varepsilon_{\text{reco.},i} \cdot \varepsilon_{\text{id.},i} \cdot \varepsilon_{\text{trig.},i})^{-1} \\ N_{\text{reco.}} &= \sum_{i=1}^{N} (\varepsilon_{\text{id.},i} \cdot \varepsilon_{\text{trig.},i})^{-1} \\ N_{\text{id.}} &= \sum_{i=1}^{N} (\varepsilon_{\text{trig.},i})^{-1} \end{aligned} \tag{4.10}$$

Each term in the weight was designed to *unfold* the effect of a particular selection criteria. For example, $(\varepsilon_{\text{trig.},i})^{-1}$ unfolds the inefficiency due to the trigger, so that ideally $N_{\text{id.}}$ should be consistent with $T_{\text{id.}}$. The deviation of the fraction $C_{\text{id.}} = N_{\text{id.}}/T_{\text{id.}}$ from one is therefore a closure fraction which gives an indication of the performance of the trigger efficiency term. The remaining fractions[18], $C_{\text{reco.}}$ and $C_{acc.}$, likewise allow an inspection of the validity of the identification and reconstruction efficiencies on a term-by-term basis.

---

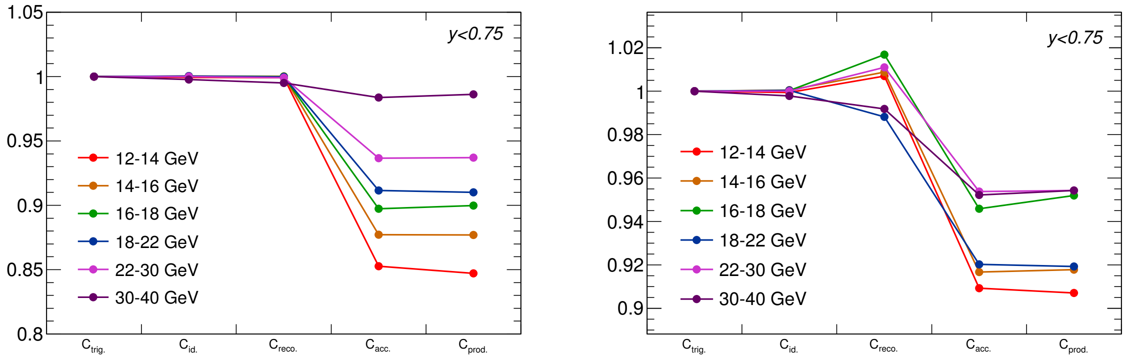
[17] The efficacy of the background subtraction is a separate issue to the weighting method, and as such the impacts of this were removed by requiring that the electrons from the candidates be the best matches of the real J/ψ electrons from that event.

[18] $C_{\text{trig.}}$ could also be included, but by design is identically equal to one ($T_{\text{trig.}}$ and $N_{\text{trig.}}$ are calculated in exactly the same way).

The most relevant and useful closure fractions are those calculated in the same bins for which the cross-section measurement was be carried out. The measurement (Section 4.8) was performed in regions which provided adequate statistics to perform an invariant mass fit, with the specific binning chosen to coincide with a subset of that of the $J/\psi \to \mu^+\mu^-$ analysis. From Figure 4.4, the most populated regions are at central ($|y| < 0.75$) rapidities with $12 < p_T < 40$. This is driven by the acceptance of the triggers (the lower limit) and the nature of the underlying production spectrum, for which the high p_T regions are in the low-statistics tail. The *EF_2e5_tight_Jpsi* trigger was used to cover the 12–14 GeV and 14–16 GeV bins, with the *EF_e9_tight_e5_tight_Jpsi* trigger used for those above this — 16–18, 18–22, 22–30, and 30–40 GeV. These will be referred to as the *analysis* bins.

For the truth quantities T_{acc} and T_{prod}, the generated values of electron η and E_T were used to determine whether or not each decay was in the acceptance and biased production regions, as well as to calculate the y and p_T of the J/ψ parent. This is not by choice, but rather out of necessity — at this point (before reconstruction) these are the only variables available. Obviously, this will have influenced whether or not a particular J/ψ is counted towards these totals, and indeed to which bin they are allocated. The remaining truth variables ($T_{\text{reco.}}, T_{\text{id.}}$ and $T_{\text{trig.}}$) used the reconstructed kinematic values for consistency with the N candidates, whose selection was based on measured values alone (for example, in the acceptance cut). Due to the migration effects alluded to above, this distinction has several important implications which have a large impact on the closure fraction results.

The results for the prompt sample are shown in Figure 4.11, for both the standard definition of the efficiencies and the factorised individual electron efficiencies. As has been mentioned previously, $C_{\text{trig.}}$ is identically 1 by definition. For the collective J/ψ efficiencies (Figure 4.11a), $C_{\text{id.}}$ and $C_{\text{reco.}}$ are very close to unity, as expected. In fact, because the sample used to determine the efficiencies is the same used to conduct this closure study, the only reason that these fractions are not exactly correct is due the fact that events with more than one J/ψ were excluded from the efficiency calculation (whereas the totals $T_{\text{id.}}$ and T_{reco} include these decays). The bias this introduced was expected to be very minor and can be calculated as follows. If the efficiencies do not depend on the number of J/ψ present in the event, then applying the weight to each of the small number, $N_{>1}$, of cases with > 1 $J/\psi \to e^+e^-$ decay the each event will return a total which is consistent with the true value to within the statistical uncertainty on the weight. From the figures in Section 4.6, this uncertainty is of the order of a few per cent. Given that the fraction of such events ($N_{>1}/N_{tot.}$) was measured from the MC samples to be $\sim$1%, the overall affect on the closure fractions would therefore be of the



(a) Using J/ψ efficiencies

(b) Using electron efficiencies (assuming factorisation).

Figure 4.11: The closure fractions for the prompt J/ψ sample for $|y| < 0.75$ in various p_T bins. The left figure shows the results when overall J/ψ efficiencies are used, whilst the right assumes that the individual electron efficiencies can be used under the factorisation assumption. In both cases there is a significant deviation from one in $C_{acc.}$ indicating a failure in the reconstruction efficiency term. Note that each subfigure has a different vertical scale.

order of $\sim 0.01 \times 0.01 = 1 \times 10^{-4}$, and therefore negligible. Since this is not the case for these results, events with > 1 J/ψ must have different efficiencies to those with only one. To find the maximum expected deviation, let the difference in the efficiencies of the J/ψ between single decay and multiple decay events assume the boundary values of ± 1, i.e. $\varepsilon_{>1} = \varepsilon_1 \pm 1$ (where variables with a subscript of 1 indicate those from single $J/\psi \rightarrow e^+e^-$ events). The combined efficiency ε is then given by:

$$\begin{aligned}\varepsilon &= \frac{N_1\varepsilon_1 + N_{>1}\varepsilon_{>1}}{N_1 + N_{>1}} \\ &= \frac{\varepsilon_1 + N_{>1}/N_1(\varepsilon_1 \pm 1)}{1 + N_{>1}/N_1} \\ &= \varepsilon_1 \pm \frac{N_{>1}/N_1}{1 + N_{>1}/N_1} \\ &= \varepsilon_1 \pm 0.01. \end{aligned} \tag{4.11}$$

Consequently, the greatest expected difference between $N_{\text{id./reco./}acc.}$ and $T_{\text{id./reco./acc.}}$ (and therefore the deviation of $C_{\text{id.}/reco/acc.}$ from one) is also $\sim$1%.

The results for the fractions $C_{\text{id.}}$ and $C_{\text{reco.}}$ are clearly consistent with these predictions, being well within 1% for all p_T bins. On the other hand, all the values of $C_{acc.}$ in Figure 4.11a are outside these allowed bounds, and indicate that there is an issue with the reconstruction efficiency term. These latter closure fractions show a strong dependence on J/ψ p_T (being worse at lower transverse momenta), which implies that whatever is causing this failure should explain this behaviour.

It has already been noted that, due to processes such as bremsstrahlung, electrons tend to be reconstructed with underestimated energy and momenta, which of course flows through to the J/ψ p_T as well. Given this problem is also known to be worse at lower transverse energies[19], this was thought to be the most likely reason for the lack of closure seen here. This systematic effect, referred to here as a *migration*, impacts the closure in several competing ways:

- Firstly, consider a J/ψ with a true p_T of 14.5 GeV that is reconstructed with a measured value of 13.9 GeV, say. The reconstruction efficiency is binned by the true p_T and so the correct value should be taken from the 14–16 GeV bin, whereas in the analysis the reconstructed value of 13.9 is all that is available and the efficiency for the 12–14 GeV bin will be that which is used. The efficiency for lower values is generally lower (at least below 30 GeV), and so the weight will tend to be overestimated. Since the trigger and identification efficiencies were binned by the reconstructed value of p_T, this migration effect does not effect $C_{\text{id.}}$ or $C_{\text{reco.}}$.
- In the analysis, the reconstructed values are used to enforce the acceptance conditions, but in the closure test the acceptance conditions used in determining $T_{acc.}$ are (by necessity) applied to the true kinematic values. So whilst an electron generated with a transverse energy of 5.5 GeV may have contributed to $T_{acc.}$, it is quite likely to have been reconstructed with $E_T < 5$ and be excluded from the analysis. This is compounded by the fact that two electrons are involved in the acceptance conditions. Since this tends to push $T_{acc.}$ up with respect to $N_{acc.}$, the net result would be be a drop in $C_{acc.}$.

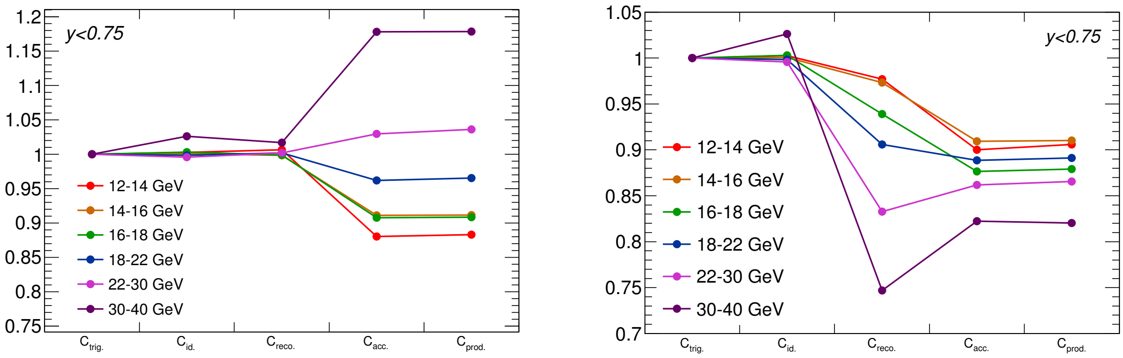
Though these two effects move $C_{acc.}$ in opposite directions, the second should have a larger influence than the first — rather than simply changing the weight of a candidate, it can exclude the candidate from the count altogether. Hence migration should cause a net downward shift in $C_{acc.}$, which is larger at lower transverse momenta. This is precisely the trend the results show. Though this seemed to be the correct explanation, it was important to confirm this by demonstrating that if the migration was removed the weighting method would perform as expected. This was the purpose of the second closure test, which is explained below.

---

[19] At lower values of transverse energy, electrons tend to be bent more by the magnetic field. Any photons produced by bremsstrahlung will therefore tend to be fairly well separated from the electron in the detector and the energy they carry lost. Conversely, at higher energies the electron and photon are more collimated and their clusters merge in the calorimeter.

The flat line between $C_{acc.}$ and $C_{prod.}$ seems to indicate that the acceptance term is operating correctly, but to be sure this too will also needed to be re-examined in the second test.

For the closure fractions in Figure 4.11b, the identification and reconstruction efficiencies were calculated from the individual electron efficiencies under the factorisation assumption (see Equation 4.8). The results for the $C_{\rm reco.}$ term indicate that the electron identification efficiencies do not factorise completely, but do to some extent (if factorisation was strongly violated, these fractions would be farther from unity). This implies that the identification of one electron depends on the environment around it, and in particular on whether or not the second electron from the parent J/ψ was identified. One would expect that this too would show some J/ψ p_T dependence, as this is a key factor determining the opening angle between the two electrons. However, this is not obvious in these results. The 5%–10% drop to the $C_{acc.}$ fractions is presumably also due to migration and again displays a clear dependence on transverse momentum (it is the difference between $C_{\rm reco.}$ and $C_{acc.}$ that is important here, not the absolute value).



(a) Using J/ψ efficiencies

(b) Using electron efficiencies (assuming factorisation).

Figure 4.12: The closure fractions for the non-prompt J/ψ sample for $|y| < 0.75$ and several p_T bins. As in the prompt case, there is a conspicuous failure of the reconstruction efficiency term causing a lack of closure in the $C_{acc.}$ fractions. The results when using individual identification electron efficiencies show that they do not factorise. Note that each subfigure has a different vertical scale.

The equivalent plots are shown for the non-prompt sample in Figure 4.11, and have many of the same trends and features as for the prompt sample. In this case, the allowed bounds for the variables $C_{\rm id.}$ and $C_{\rm reco.}$ in Figure 4.12a are larger because the fraction, $N_{>1}/N_{tot.}$, was measured from the MC sample to be ~8%, rather than 1%. As such, the outlier points of the 30–40 GeV line are acceptable. From earlier considerations, the size of the deviation of these fractions within the allowed ± ∼8% provides insight into the difference in efficiencies for events with one $J/\psi \to e^+e^-$ decay and those with more than one. Aside from the 30–40 GeV bin, the results therefore indicate the efficiency is reasonably independent of the number of J/ψ decays in the event. The result for the 30–40 GeV bin is reflective of the fact that for highly boosted events, the J/ψ (presumably produced in the same interaction) may be closer in proximity within the detector.

A major difference between these results and those for the prompt sample is that the factorisation of electron identification efficiencies appears to be strongly violated and p_T-dependent in this case, with $C_{\text{reco.}}$ falling as low as 0.75 for the highest bin.

4.7.2 Closure Testing With Migration Effects Removed

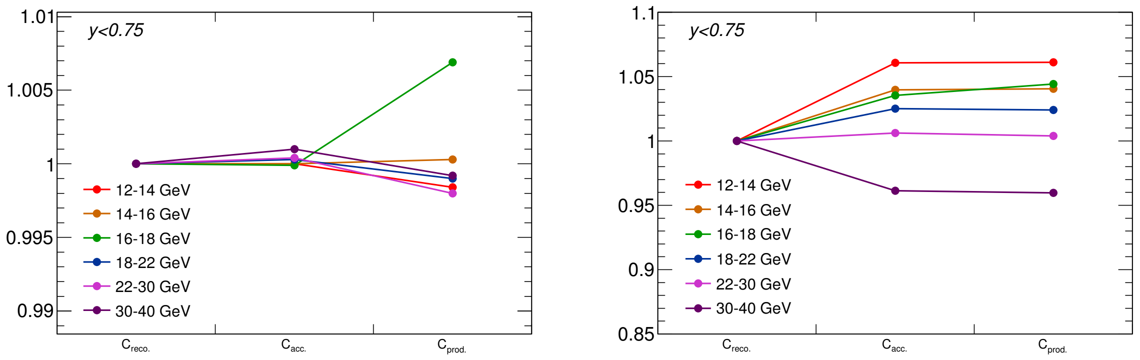
As was mentioned in the previous section, a second closure test was necessary to confirm that, when migration effects are removed, the closure fractions for $C_{acc.}$ and $C_{prod.}$ are be correct (i.e. consistent with one). Furthermore, it was important to ascertain whether the electron reconstruction efficiencies factorised, and whether the acceptance correction was performing as expected (or, if not, identify the cause of any problems).

To negate the effects of migration, the analysis used only the generated values of kinematic variables corresponding to the reconstructed objects of interest[20]. As discussed in the previous section, the most significant differences arising from this change occur when enforcing the acceptance conditions and in looking up reconstruction and acceptance efficiencies. Since the fractions $C_{\text{id.}}$ and $C_{\text{reco.}}$ are not being tested in this study, no trigger or identification selections were performed and therefore the respective efficiencies for these selections were set to one in the calculation of the J/ψ weights.

The results of this second closure test for both samples are shown in figure 4.13 (since there were no trigger or identification criteria applied, $C_{\text{reco.}}$ is trivially equal to one and is only shown as a starting point). When the individual electron efficiencies are used assuming factorisation (figures 4.13b and 4.13d), the closure fractions $C_{acc.}$ are offset from one by ±5% and, in similar fashion to the previous results, there is clearly a strong dependence on J/ψ p_T. Again, this indicates that the J/ψ reconstruction efficiency does not factorise into individual electron efficiencies. The correlation with p_T (which essentially determines the opening angle between the two electrons) provides good evidence that the reconstructability of each electron depends on proximity to the other electron. This may arise, for example, from the merging of the electron clusters in the

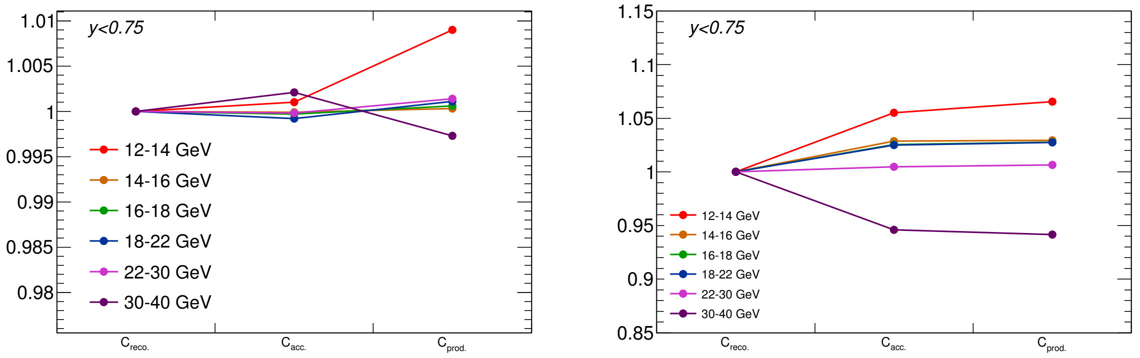
---

[20] In an analysis on real data, this is obviously not possible and is the reason migration is an issue



(a) Using J/ψ efficiencies (prompt)

(b) Using electron efficiencies (prompt).



(c) Using J/ψ efficiencies (non-prompt)

(d) Using electron efficiencies (non-prompt).

Figure 4.13: The second closure test fractions for the prompt and non-prompt J/ψ sample in the same rapidity and p_T bins as the first test. In this case, the effects of migration have been removed. Note that each subfigure has a different vertical scale.

calorimeter or the mis-assignment of hits to the two corresponding tracks. Either of these effects generally result in poorly reconstructed objects which are more likely to fail the electron selection criteria.

The deviations in the closure fractions $C_{acc.}$ when using overall J/ψ efficiencies (figures 4.13a and 4.13c) are extremely small and are again easily explained as being due to the exclusion of multi-J/ψ events in the efficiency calculations. This confirms that migration was the cause of the problem associated with the reconstruction efficiency term of the weight. The values for $C_{prod.}$ are very close to one, and are consistent with the expected statistical fluctuations. These can be approximated in a simple manner if it is assumed that

1. the acceptance fractions themselves have no associated uncertainty (the actual value of their uncertainty is $< 10^{-4}$, and so this is a reasonable assumption),

2. the acceptance within each analysis bin is constant, and
3. the closure fractions $C_{acc.}$ were exactly one.

In this case, the formula for $C_{prod.}$ reduces to $C_{prod.} = \mathcal{A} \cdot T_{acc.}/T_{prod.}$, where $\mathcal{A}$ is the representative value for the acceptance in the bin. The statistical uncertainty is then given by splitting up $T_{prod.}$ as $T_{prod.} = T_{acc.} + T_{n.acc}$, and using Poisson errors:

$$\sigma_C = \sqrt{\left(\frac{\partial C}{\partial T_{acc.}}\right)^2 \cdot T_{acc.} + \left(\frac{\partial C}{\partial T_{n.acc.}}\right)^2 \cdot T_{n.acc.}} = C\sqrt{\frac{1}{T_{acc.}} - \frac{1}{T_{prod.}}} \tag{4.12}$$

The values for $T_{prod.}$ and $T_{acc.}$ depend on the MC data set and the bin of interest, being larger in the prompt sample and at lower p_T. Assuming $C \sim 1$, the extreme values for the two numbers give uncertainties in the range $0.004 < \sigma_C < 0.006$. The fluctuations observed in the results give $\chi^2 = 2.45 (p = 0.87)$ and $\chi^2 = 3.90 (p = 0.69)$ for the prompt and non-prompt samples, respectively. These values are consistent with a χ^2 distribution with 6 degrees of freedom, and demonstrates that the acceptance term is performing precisely as intended.

In summary, the outcomes of this secondary closure test confirm that, in the absence of migration effects, the logic of the weighting methodology is sound and should provide accurate predictions for the number of J/ψ produced given the number found in the analysis[21]. However, the results of the primary closure test (in the previous section) show that the migration of electron E_T results in a deviation of the weighted value of $N_{prod.}$ up to 20% from its expected number. Whilst there are possibly methods of *unfolding* some of the simpler effects of migration[22], it is undesirable to have to do so and may be quite complex in some cases. For example, the acceptance conditions are applied on reconstructed kinematic variables, and therefore so should they be in the acceptance fraction calculation. However, it is not clear that it is even sensible to ask what the measured energy of a generated electron is, particularly since at this stage of the calculation it is not even known whether that electron will actually be reconstructed at all.

Another way to deal with the migration is to restrict the analysis to high p_T J/ψ (and therefore generally higher energy electrons). This has two main benefits —

---

[21] At least so far as the MC samples can be relied upon for their accurate description of the physics and the operation of the detector.

[22] For example, given a reconstructed E_T for an electron, it might be possible to determine the distribution of the average relative number of corresponding generated electrons in various true E_T bins. This could be used to construct a more accurate, weighted reconstruction efficiency.

1. Firstly, reconstruction of higher energy electrons suffers far less from energy loss mechanisms, meaning that the measurement of transverse energy and momentum is more accurate. In turn, this lessens the extent of the impacts of migration.

2. Since the reconstruction efficiency and acceptance both plateau above a certain level, even if the transverse momentum is under-estimated, the correct values will still be used for the weight. Unfortunately, this method is fairly restrictive on the phase space to which the analysis is sensitive.

Obviously, the best solution to the migration problems is to improve the accuracy of the measured values of electron energies and trajectories. This would involve developing software specifically to account for bremsstrahlung, and more specifically to improve the track fitting of the electrons. At the time of writing (late 2012), these tools have started to be implemented for use in ATLAS, and significant improvements have already been observed. Any future analyses using low E_T electrons would almost certainly benefit from using these improved reconstruction algorithms.

Both closure tests also showed that the reconstruction and identification efficiencies for $J/\psi \to e^+e^-$ decays cannot be simply factorised into individual electron efficiencies. This is unfortunate, since there is a large effort within the ATLAS community to measure and provide these individual efficiencies for analyses (for studies such as those for $Z \to e^+e^-$, the electrons are generally far more separated within the detector and factorisation of efficiencies is a valid assumption). A complete analysis would therefore entail the added task of measuring the J/ψ identification and trigger efficiencies on collision data to remove any bias introduced in Monte Carlo (MC) simulations.

Given that the main aim of this study was to identify and quantify areas of difference or difficulty in dielectron quarkonium final states as compared to dimuon states, no attempts were made to remedy any of these problems. Rather, the focus was shifted to the second main purpose of this analysis; to attempt a measurement of the J/ψ production cross-section and compare to the result from the muon analysis. In light of this, the final closure fractions ($C_{prod.}$) from Figures 4.11 and 4.12 were adopted and used as scale factors for the cross-section measurement. To ensure that this would correctly take into account migration, a further (and final) closure test was required — this time by splitting the MC samples into two independent halves.

4.7.3 Closure Test on Independent Samples

The final closure test was intended to confirm that the migration issues in the real data sample could be dealt with by using the closure fractions derived from the MC samples as scale factors. In other words, if the number of J/ψ detected in a particular bin was N, the corrected yield would be $N \cdot C_{prod}^{-1}$. By definition, this will produce the exact correct numbers trivially when applied to the entire MC data sets, which necessitated a test on a statistically independent sample.

Each of the MC samples were separated into two sub-samples of roughly half the size of originals. The first half of the sample was used to extract the various efficiencies. These efficiencies were then used to conduct a (self-)closure test on that same half in the original manner described above (Section 4.7.1). The reciprocals of the resulting closure fractions were relabelled as migration *scale factors* for each bin. Finally, the efficiencies were used to perform a subsequent closure test, this time on the second half of the sample (a proxy for the real data sample), with the final values for $N_{prod.}$ multiplied by the migration scale factors. If primes denote the second half of the sample, the equation describing this is:

$$C \equiv C'_{prod.} C_{prod.}^{-1} = \frac{N'_{prod.}}{T'_{prod.}} \frac{T_{prod.}}{N_{prod.}} \tag{4.13}$$

The measured closure fractions are only meaningful when interpreted in the context of a reasonable understanding of the statistical fluctuations. An exact calculation of the uncertainties was not performed, as migration effects made this a complicated problem and full rigour was not deemed necessary. Instead, some reasonable assumptions were made to simplify the problem — in particular, that there is no migration (uncertainties due to migration will be discussed below). The resulting formula for the uncertainties on C (derived in Appendix 4.A) is given by:

$$\begin{aligned}\frac{\sigma_C^2}{C^2} = \frac{1}{T'_{prod.}} + \frac{1}{T_{prod.}} + \sum_{i=1}^{3}\left(1+\frac{\overline{N}_i}{N_i}\right)\Bigg[D_i^2\frac{\overline{N}_i}{N_i^2} + \sum_{r=1}^{M}\Bigg[\left(1+\frac{\overline{N}_i}{N_i}\right)\left(\frac{q'_{r,i}}{S'}-\frac{q_{r,i}}{S}\right)^2\frac{\sigma^2_{\mathcal{A}_{r,i}}}{\mathcal{A}^4_{r,i}} \\ +\mathcal{A}_{r,i}^{-2}\left(1+\frac{\overline{N}_i}{N_i}\right)\left(\frac{q_{r,i}}{S^2}+\frac{q'_{r,i}}{S'^2}\right) - \frac{2}{S\mathcal{A}_{r,i}}\left(\frac{1}{T_{prod.}}+\frac{\overline{N}_i D_i}{N_i^2}\right)q_{r,i}\Bigg]\Bigg]. \end{aligned} \tag{4.14}$$

In this equation, primes again refer to the second half of the sample and the i subscripts label the efficiency sub-bin (of which there are three for each analysis bin — $0.0 < |y| < 0.25$, $0.25 < |y| < 0.5$, $0.5 < |y| < 0.75$ with the same p_T range). The variables S and S' are the sum of weights (equivalent to $N_{prod.}$ and $N'_{prod.}$, respectively), N_i is the subset of candidates passing all selections in the i^{th} sub-bin and $\overline{N}_i$ is its complement. For each of the M acceptance bins within an efficiency bin, the occupancy of candidates is labelled by $q_{r,i}$, the acceptance by $\mathcal{A}_{r,i}$ and the uncertainty on the acceptance by $\sigma_{\mathcal{A}_{r,i}}$. Lastly, D_i is the difference of the fraction $\sum_{r=1}^{M} q_{r,i}\mathcal{A}_{r,i}^{-1}/S$ between the first and second half of the sample.

To try to understand the typical magnitude of the terms in this equation, several reasonable approximations can be made. Firstly, so long as the statistics of the sample are not limiting (i.e. the N_i are not too small), the variables D_i and the difference $q'_{r,i}/S' - q_{r,i}/S$ should be close to zero. Then, presuming that the acceptance is reasonably constant across each efficiency sub-bin ($\mathcal{A}_{r,i} \approx \mathcal{A}_i$), the relative uncertainty on C is

$$\frac{\sigma_C^2}{C^2} \approx \frac{1}{T'_{prod.}} + \frac{1}{T_{prod.}} + \sum_{i=1}^{3}\sum_{r=1}^{M}\left(1+\frac{\overline{N}_i}{N_i}\right)\left[\mathcal{A}_i^{-2}\left(1+\frac{\overline{N}_i}{N_i}\right)\left(\frac{q_{r,i}}{S^2}+\frac{q'_{r,i}}{S'^2}\right) - \frac{2q_{r,i}}{S\mathcal{A}_i T_{prod.}}\right] \tag{4.15}$$

Furthermore, it is sensible to assume that $T'_{prod.} \approx T_{prod.}$, $S' \approx S$, $q'_{r,i} \approx q_{r,i}$, and $S\mathcal{A}_i \approx 3S_i\mathcal{A}_i \approx 3N_iT_{a,i}/N_i = 3T_{a,i}$ ($T_{a,i}$ being the true number of candidates passing the acceptance cut in the i^{th} efficiency sub-bin), from which Equation 4.16 reduces to

$$\frac{\sigma_C^2}{C^2} \approx \frac{1}{T'_{prod.}} + \frac{1}{T_{prod.}} + \frac{2}{N} - \frac{2}{T_{prod.}} \tag{4.16}$$

$$\Rightarrow \sigma_C \approx C\sqrt{\frac{2}{N}} \tag{4.17}$$

This result is much simpler and more intuitive; $\sqrt{2/N}$ is precisely the uncertainty expected for the quotient (or product) two independent variables N and N' in the case that $N \approx N'$. In other words, the largest contribution to the statistical uncertainty is driven by the uncertainty on the total number of candidates in each analysis bin, in each half of the total sample.

If the effects of migration are included, the uncertainties are only slightly inflated. In Equation 4.13, the number of J/ψ candidates (N and N') in the sum of weights (S and S') and the individual values for the weights are all altered. The sum of weights in a given efficiency sub-bin becomes

$$\begin{aligned} S_i &= \sum_{r=1}^{N_i} w_{r,i} \\ &= \frac{T_{a,i}}{N_i}\sum_{r=1}^{N_i}\mathcal{A}_{r,i}^{-1} \\ \xrightarrow{Migration}\ S_{mig,i} &= \frac{T_{a,i}}{N_i+M_i}\sum_{r=1}^{N_i+M_i}\mathcal{A}_{r,i}^{-1} \\ &\approx \frac{T_{a,i}}{N_i}\left(1-\frac{M_i}{N_i}\right)\sum_{r=1}^{N_i+M_i}\mathcal{A}_{r,i}^{-1} \qquad (4.18) \end{aligned}$$

(4.19)

Taking $\overline{w_i} \equiv \frac{T_{a,i}}{N_i^2}\sum_{r=1}^{N_i}\mathcal{A}_{r,i}^{-1} = S_i/N_i$ and assuming that $\frac{T_{a,i}}{N_i}\sum_{r=N_i}^{N_i+M_i}\mathcal{A}_{r,i}^{-1} \approx M_i\overline{w_i}$, Equation 4.18 can be written as

$$\begin{aligned} S_{mig,i} &= \left(1-\frac{M_i}{N_i}\right)(N_i+M_i)\overline{w_i} \\ &= N_i\overline{w_i} - \frac{M_i^2}{N_i}\overline{w_i} \\ &= S_i\left(1-\frac{M_i^2}{N_i^2}\right) \end{aligned} \tag{4.20}$$

A similar expression can be obtained for $S'_{mig,i}$, but with independent M'_i, S'_i and N'_i. The resultant expression for C in the presence of migration is

$$C = \frac{T_{prod.}}{T'_{prod.}}\frac{\sum_{i=1}^{3} S'_i\left(1-\frac{M_i'^2}{N_i'^2}\right)}{\sum_{i=1}^{3} S_i\left(1-\frac{M_i^2}{N_i^2}\right)}. \tag{4.21}$$

The uncertainty induced in C from fluctuations in the M_i and M'_i (all other variables are accounted for in formula 4.14) are therefore given by:

$$\sigma_C = \sqrt{\sum_{i=1}^{3}\left[\left(\frac{\partial C}{\partial M_i}\right)^2 M_i + \left(\frac{\partial C}{\partial M'_i}\right)^2 M'_i\right]} \tag{4.22}$$

$$= C\sqrt{\sum_{i=1}^{3}\left[\frac{4\left(\frac{M_i}{N_i^2}\right)^2 M_i}{\left(1-\frac{M_i^2}{N_i^2}\right)^2} + \frac{4\left(\frac{M'_i}{N_i'^2}\right)^2 M'_i}{\left(1-\frac{M_i'^2}{N_i'^2}\right)^2}\right]} \tag{4.23}$$

Finally, using the known migration fractions measured from MC samples of $M_i/N_i \sim 0.1$ and assuming that the M_i, M'_i are all similar, the magnitude of this contribution to the uncertainty on C scales like

$$\frac{\sigma_C}{C} \approx \frac{0.05}{\sqrt{M}} \approx \frac{0.16}{\sqrt{N}} \tag{4.24}$$

which shows that this is $\sim 1/13^{th}$ the size of the uncertainty with no migration (see 4.16), which in combination (by adding in quadrature) is only a 0.3% increase. This justifies the assumption that migration can essentially be ignored in the error calculation.

If the approach of using the original closure fractions as scale factors to negate the effects of migrations is valid, this test should return numbers consistent with one, i.e. compared to statistical uncertainties. The results, displayed in table 4.3, confirm that this is in fact the case. The z-values (i.e. the pulls of the closure fractions with respect to unity) reflect a statistical scatter, with the overall χ^2 for each sample at typical values for

Table 4.3: The closure fractions using independent samples, with the fractions from the first sample serving as scale or correction factors for those in the second.

	Prompt Sample		**Non-prompt Sample**	
Analysis Bin [GeV]	$C_{prod.,corr.}$	z	$C_{prod.,corr.}$	z
12–14	0.98 ± 0.02	1.1	0.97 ± 0.03	1.3
14–16	0.99 ± 0.02	0.43	1.02 ± 0.03	-0.77
16–18	1.02 ± 0.03	-0.76	0.97 ± 0.04	0.69
18–22	0.96 ± 0.02	1.6	0.96 ± 0.03	1.3
22–30	1.00 ± 0.03	-0.048	1.00 ± 0.04	0.019
30–40	0.96 ± 0.06	0.72	0.99 ± 0.10	0.080
		$\chi^2 = 5.17$		$\chi^2 = 4.37$

six degrees of freedom. There may be some hint that the fractions are underestimated, since there are just three greater than one and the other nine are less than one. Under the hypothesis of no bias, a simple sign test gives a p-value of 0.33 and 0.09 for the prompt and non-prompt samples, respectively. The best fit to the overall closure fractions returned 0.987 ± 0.009 and 0.98 ± 0.01 for the prompt and non-prompt samples, with corresponding significances (compared to unity) of 1.35 and 1.18. Under the hypothesis that there is an underlying closure fraction common to *both* samples, the significance is slightly higher, at 1.78. With this level of information there is no evidence for concern that the fractions are skewed to lower values.

With the weighting method, amended with migration scale factors, validated on an independent sample, these results were applied to the real data samples. The extraction of the weighted yield of $J/\psi \to e^+e^-$ decays in each bin was then subsequently used to calculate the differential cross-sections, which is explained in the following section.

4.8 Cross-section Measurement and Comparison to $J/\psi \to \mu^+\mu^-$ Results

The product of the J/ψ production cross-section and branching fraction to the dielectron final state was measured in a manner analogous to the muon analysis (Equation 4.3). The bins for the closure tests were specifically chosen to match those used for the cross-section measurement such that the results were directly applicable — in particular, the migration scale factors.

Consider a particular bin with width Δp_T and migration scale factor $1/C_{prod.}$. Then from the considerations of the previous section,

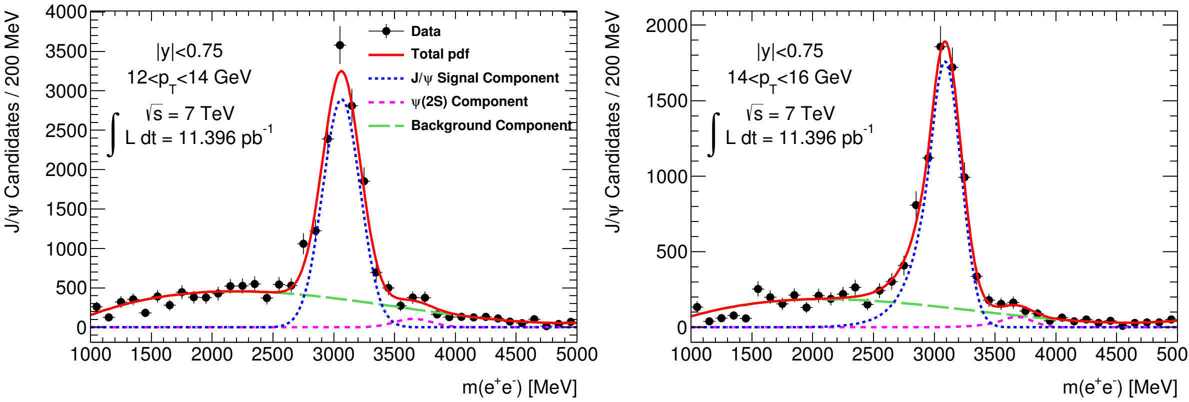
$$\frac{d\sigma(J/\psi)}{dp_T} Br(J/\psi \to e^+e^-) = \frac{N_{corr}^{J/\psi}}{\mathcal{L}\Delta p_T C_{prod.}} \tag{4.25}$$

where $N_{corr}^{J/\psi}$ is the weighted yield of $J/\psi \to e^+e^-$ decays in that bin, and $\mathcal{L}$ is the total integrated luminosity. As in the muon case, the extraction of $N_{corr}^{J/\psi}$ was performed by fitting a signal and background curve to the weighted invariant mass distributions. However, the application of weights to the $J/\psi \to e^+e^-$ candidates is not straight-forward, as it is for muons — it is not known *a priori* whether a particular candidate was produced through the prompt or non-prompt mechanism. Whilst there were only small differences in the trigger and reconstruction efficiency between the two scenarios, the identification efficiency was far greater for prompt J/ψ, in every bin (see Section 4.6 for details). Consequently, weights for prompt and non-prompt J/ψ are quite different, meaning the decision of which to use for a particular candidate can have a strong impact on the final cross-section measurement. Upper and lower bounds on the true (mixed) cross-section were determined by performing binned minimum-χ^2 fits on weighted distributions assuming (1) all J/ψ were produced promptly and (2) all were produced non-promptly, respectively.

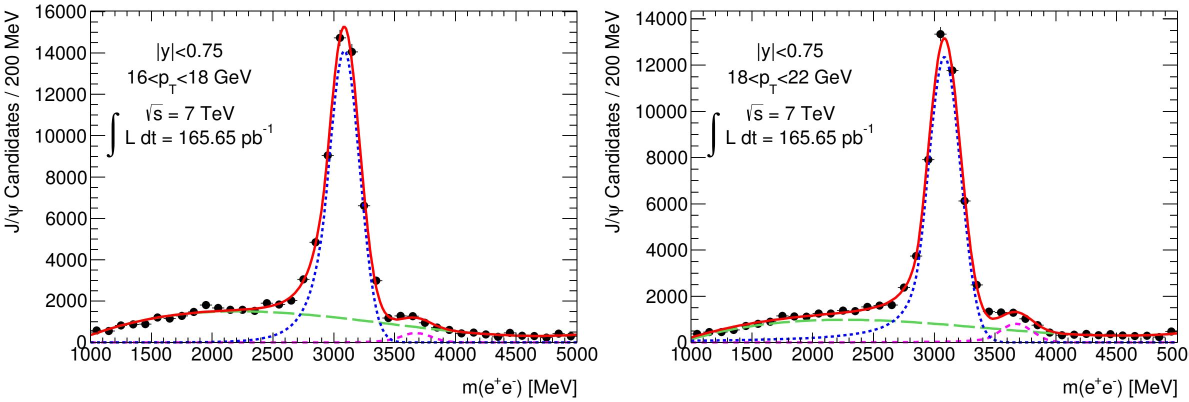
The J/ψ polarisation also has a large affect on the weights (through the acceptance term). The following results all assume the FLAT spin-alignment hypothesis (Section 4.6.1), but a theoretical uncertainty is calculated below to recognise that the polarisation is actually unknown.

The fitting function p.d.f. included two Crystal Ball functions [97] for the signal and $\psi(2S)$ peaks, and a Chebychev polynomial to account for the background. For the $\psi(2S)$, the Crystal Ball parameters α and n were constrained to be equal to those of the J/ψ (which were floated), and the mass difference to the J/ψ was set to the PDG value. In the muon case the range of the fit was chosen to be 2–4 GeV. The width of the mass peak for electrons was considerably larger, meaning a larger range (1–5 GeV) was necessary. The fits under the all-prompt hypothesis are shown in Figure 4.14, whilst those for the non-prompt are given in Figure 4.15. Tables 4.4–4.5 summarise the results of these fits, and the subsequent yields and differential cross-sections for each bin.

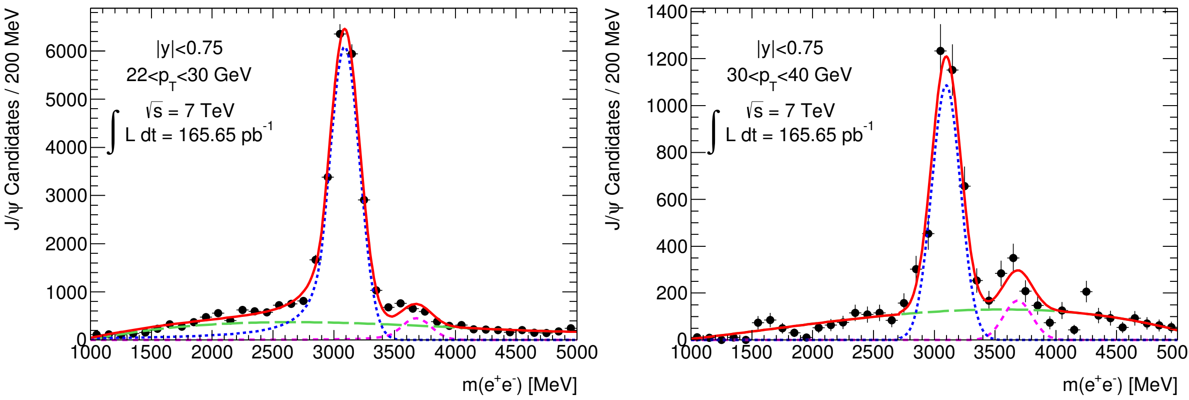
The shape of the distributions and the fitting functions are similar in both the prompt and non-prompt hypothesis cases, but the overall scale is significantly higher for the latter. Consequently, the yields and cross-sections for the non-prompt assumption are larger than the prompt, particularly in the higher p_T bins (as expected, since it is here that the identification efficiencies are most different). The instability of the α and n parameters



(a) 12–14 GeV (b) 14–16 GeV

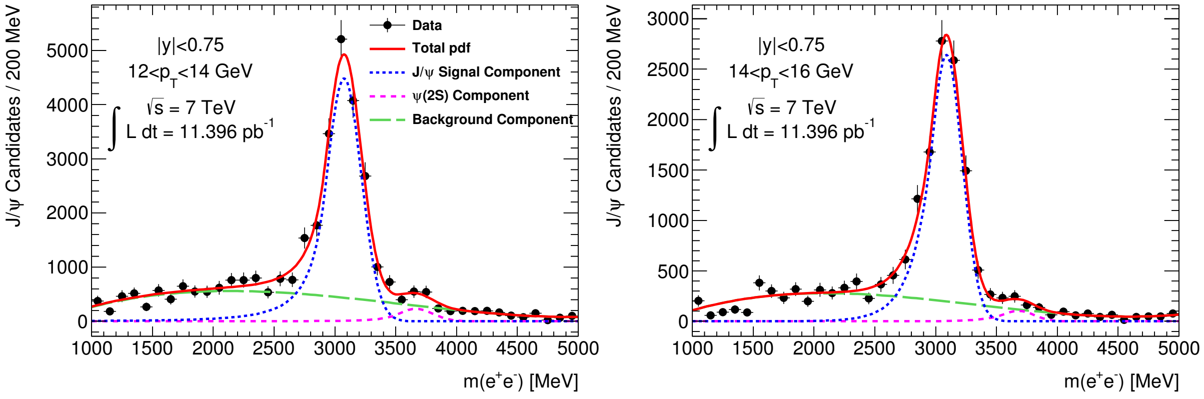


(c) 16–18 GeV (d) 18–22 GeV

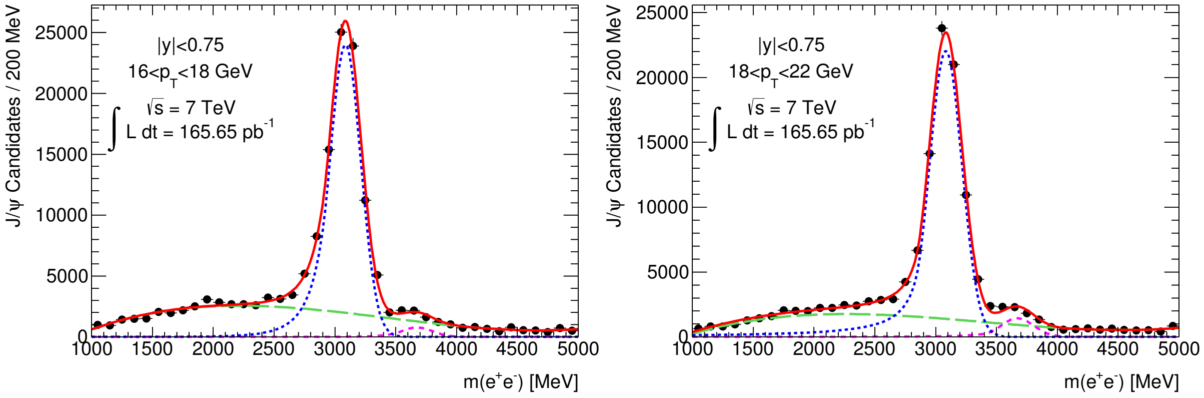


(e) 22–30 GeV (f) 30–40 GeV

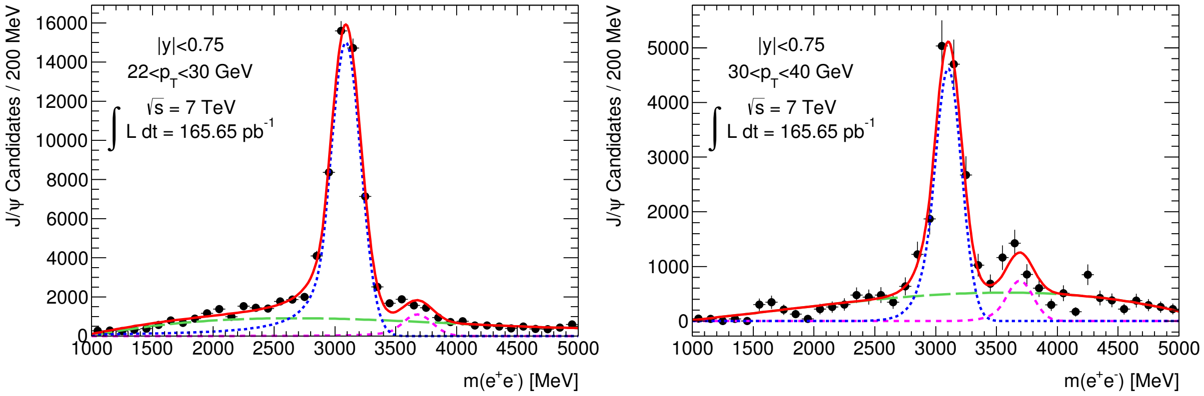
Figure 4.14: Weighted (prompt hypothesis) invariant mass distributions for $|y| < 0.75$ in various p_T bins.



(a) 12–14 GeV (b) 14–16 GeV



(c) 16–18 GeV (d) 18–22 GeV



(e) 22–30 GeV (f) 30–40 GeV

Figure 4.15: Weighted (non-prompt hypothesis) invariant mass distributions for $|y| < 0.75$ in various p_T bins.

Table 4.4: Summary of fit results for $|y| < 0.75$ (prompt hypothesis).

p_T bin [GeV]	$< p_T >$ [GeV]	Mean [GeV]	σ	α	n	$\chi^2/n.d.o.f.$	J/ψ Yield	$\psi(2S)$ Yield	$\frac{d^2\sigma}{dp_T dy} Br(J/\psi \to e^+e^-)$ [pb/GeV]
12–14	12.9	3064 ± 2	154 ± 2	65 ± 1	0 ± 1000	1.9	11400 ± 100	420 ± 40	392 ± 4
14–16	15	3088 ± 3	131 ± 3	0.85 ± 0.05	100 ± 200	2	6600 ± 100	260 ± 30	221 ± 3
16–18	16.9	3086.2 ± 0.9	126.7 ± 0.9	1.11 ± 0.04	8 ± 3	1.8	48900 ± 400	1600 ± 100	109.4 ± 0.9
18–22	19.6	3083 ± 1	133 ± 1	1.52 ± 0.05	1.4 ± 0.2	2.1	46600 ± 900	3100 ± 200	51.6 ± 0.9
22–30	24.6	3090 ± 1	124 ± 2	1.48 ± 0.05	1.4 ± 0.2	2.8	21800 ± 500	1600 ± 100	11.7 ± 0.2
30–40	32.9	3098 ± 2	118 ± 4	0 ± 500	167.8549 ± 0.0008	2.5	3300 ± 200	520 ± 30	1.36 ± 0.07

Table 4.5: Summary of fit results for $|y| < 0.75$ (non-prompt hypothesis).

p_T bin [GeV]	$< p_T >$ [GeV]	Mean [GeV]	σ	α	n	$\chi^2/n.d.o.f.$	J/ψ Yield	$\psi(2S)$ Yield	$\frac{d^2\sigma}{dp_T dy} Br(J/\psi \to e^+e^-)$ [pb/GeV]
12–14	12.9	3076 ± 2	141 ± 2	1.16 ± 0.06	3 ± 1	1.6	18100 ± 500	930 ± 90	600 ± 15
14–16	15	3089 ± 3	131 ± 2	0.85 ± 0.04	137.7 ± 0.2	2	10000 ± 100	390 ± 40	320 ± 4
16–18	16.9	3086.2 ± 0.7	126.8 ± 0.7	1.11 ± 0.03	8 ± 2	1.8	83200 ± 500	2700 ± 100	184 ± 1
18–22	19.6	3083.4 ± 0.8	133.2 ± 0.9	1.52 ± 0.04	1.4 ± 0.2	2.1	83000 ± 1000	5500 ± 200	87 ± 1
22–30	24.6	3089.5 ± 0.8	123.6 ± 1	1.47 ± 0.03	1.4 ± 0.1	2.8	53600 ± 700	4000 ± 200	26.0 ± 0.3
30–40	32.9	3103 ± 1	110 ± 1	1.23 ± 0.04	143.67 ± 0.02	2.4	13800 ± 100	2200 ± 80	4.70 ± 0.05

(for example, in the 12–14 GeV bin for prompt, or the 14–16 GeV bin for non-prompt) is not surprising; large α indicates that the signal peak is essentially Gaussian, whilst for large values of n the crystal ball function has an exponential form. Provided α is not too small (say > 1), this will only affect the fitting function shape in the tail on the low-mass side of the peak, where the background is dominant. The remaining values for α are $\sim 1 - 1.5$, indicating that the non-Gaussian tail generally begins about one standard deviation to the left of the mean.

The other signal shape parameters evolve with p_T in a predictable way:

- The means of the signal function are generally underestimated at low p_T and over-estimated in the highest p_T bin. The former is a result of bremsstrahlung (which is more pronounced at lower transverse momenta), whilst the latter is likely to be due to a mis-calibration of the electromagnetic calorimeter[23].
- The width decreases steadily in the higher p_T bins, due to the improved resolution of the electromagnetic calorimeter (see Section 2.2.2) at higher energies.

The $\chi^2/n.d.o.f.$ values for the fits are somewhat larger than what would ideally be expected (fluctuations about 1), which indicates that it may have been possible to improve the fitting procedure by making it more sophisticated. As an example, since the width of the peak is a function of p_T, using several different crystal ball functions (each with different width) may improve the description of the signal. However, for the purposes of this study the accuracy of the fits and subsequent yields was acceptable.

Before comparing these cross-section values to those from the $J/\psi \to \mu^+\mu^-$ analysis, the results of a strategy to measure the correct *mixed* prompt/non-prompt cross-section is described. The idea is simple — instead of using *either* the prompt or non-prompt weight for each candidate, a mixture of the two (based on the fraction of non-prompt J/ψ for that bin) is used. If the fraction of non-prompt J/ψ candidates in a particular acceptance bin is f_{np}, the appropriate weight is:

$$\tilde{w} \equiv (1 - f_{np})w_p + f_{np}w_{np} \tag{4.26}$$

where w_p and w_{np} are the prompt and non-prompt weights, respectively. Obviously, this prescription is only correct on average and will introduce an uncertainty if the statistics are too low. It also relies on knowing the non-prompt fractions, which have only been measured for the analysis bins (in the muon analysis — see Section 4.2). So long as these fractions do not vary considerably within each bin, this is still a valid approach.

---

[23] The J/ψ peak is typically used as a *standard candle* for calibration of the tracking and calorimeter elements of detectors.

To demonstrate this, consider a particular acceptance bin, with particular values for w_p and w_{np} and an occupancy $q = q_p + q_{np}$ (where q_p is the prompt and q_{np} the non-prompt occupancy). Then, the sum of weights using Equation 4.26 in this acceptance bin is

$$\sum \tilde{w} = q((1 - f_{np})w_p + f_{np}w_{np}). \tag{4.27}$$

On average, $\langle q(1 - f_{np})\rangle = q_p$ and $\langle qf_{np}\rangle = q_{np}$, so,

$$\langle \sum \tilde{w}\rangle = q_p w_p + q_{np}w_{np}. \tag{4.28}$$

In other words, the expectation value for the sum of mixed weights is equal to the desired value for that bin, with a relative scatter which improves with increasing statistics as $1/\sqrt{q}$.

To determine f_{np}, one could perform a simultaneous fit to the pseudo-proper time distribution (with separate prompt and non-prompt components) and the invariant mass distribution. For electrons, this process is known to be beset with several issues [98], mostly centred around the poor treatment of bremsstrahlung by the standard electron reconstruction algorithm. Not wishing to address these issues in this study, the measured fractions from the $J/\psi \to \mu^+\mu^-$ analysis, F_{np}, were used instead. However, this measurement was of the *production* fraction, which for electrons will most definitely not match the fraction remaining after all the analysis cuts (since, for example, prompt J/ψ more readily pass the identification requirement). Fortunately, the prompt and non-prompt weights themselves can be used to deduce the reconstructed fractions from the production fractions. If there are P_p prompt and P_{np} non-prompt produced within a particular acceptance bin, then the expected number passing all analysis cuts is $< q_p >= P_p/w_p$ and $< q_{np} >= P_{np}/w_{np}$. So, the correct fraction to use is:

$$f_{np} = \frac{P_p/w_p}{P_p/w_p + P_{np}/w_{np}} \tag{4.29}$$

Equation 4.26 then gives, for some particular bin,

$$\begin{aligned}
\tilde{w} &= \frac{P_p/w_p}{P_p/w_p + P_{np}/w_{np}}w_p + \frac{P_{np}/w_{np}}{P_p/w_p + P_{np}/w_{np}}w_{np} \\
&= \frac{P_p + P_{np}}{P_p/w_p + P_{np}/w_{np}} \\
&= \left(\frac{1}{w_p}\frac{P_p}{P_p + P_{np}} + \frac{1}{w_{np}}\frac{P_{np}}{P_p + P_{np}}\right)^{-1} \\
&= \left(\frac{F_p}{w_p} + \frac{F_{np}}{w_{np}}\right)^{-1}.
\end{aligned} \tag{4.30}$$

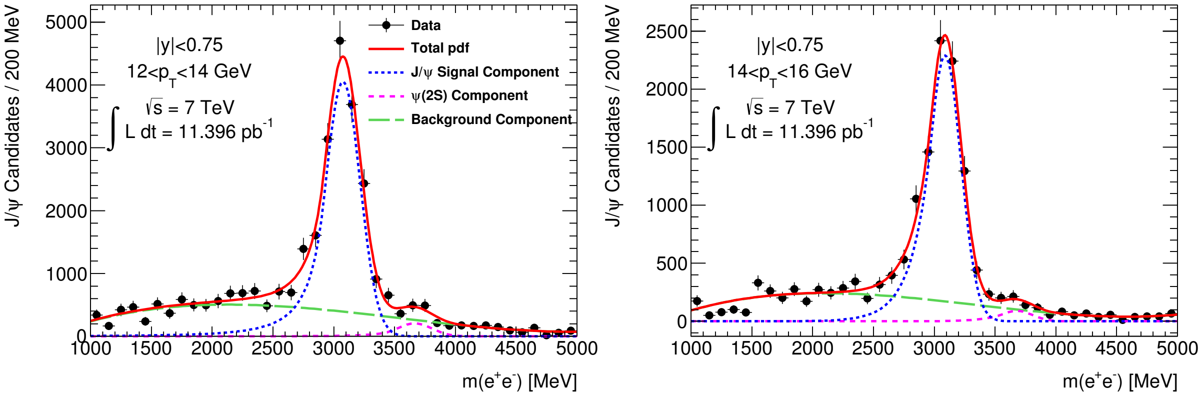
The results of the fits to the dielectron mass distribution using these mixed-weights are given in Figure 4.16 and summarised in Table 4.6. All the trends in the fitting parameters are similar to before, but the yields are (as expected) between that of the prompt and non-prompt.

With this technique for producing the mixed prompt/non-prompt cross-section values established, the spin-alignment uncertainty systematic was calculated. For the muon analysis, this was the most significant uncertainty on the cross-section measurement, so it is prudent to see how this affects the electron results. To do so, the entire analysis was re-done under each of the five extreme polarisation scenarios (Section 4.6.1) and the highest and lowest values for the subsequent cross-section measurements in each bin were adopted as upper and lower bounds on the systematic. Unsurprisingly (see Figure 4.5), the longitudinal polarisation led to the lower limit on the cross-sections, whilst the values for the transverse spin-alignments were generally quite similar and provided the upper limits. A summary of the mixed-weight cross-sections and their statistical and systematic uncertainties are given in Table 4.7. The corresponding results for the muon analysis are shown in parallel (in this case, there is an additional uncertainty term accounting for other major systematics — a calculation which was not conducted for the electron case).

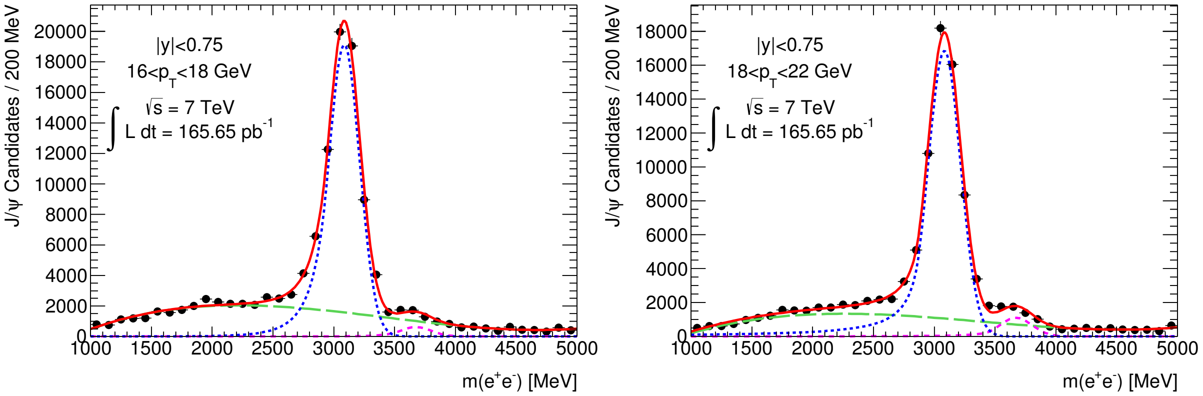
All of the cross-section measurement results are displayed in Figures 4.17a and 4.17b, with the $J/\psi \rightarrow \mu^+\mu^-$ analysis overlaid to provide visual comparison. As expected, the all-prompt and all-non-prompt results lie below and above the muon measurements, respectively. Furthermore, the differential cross-sections derived from the mixed-weights are consistent with the corresponding dimuon values, with a $\chi^2/n.d.o.f.$ of 0.7. In the two lowest p_T bins, the uncertainty associated with spin-alignment is significantly larger for the electron measurements, whereas for the remaining (higher) p_T bins they are of a similar size. The agreement of the results provides some assurance of the validity of the MC, the scaling factors used to negate the effects of migration, and the mixed weights developed above. Possible causes for any residual differences include:

- **Incorrect Spin-Alignment Hypothesis.** Whilst it is certainly true that the electron and muon results should be consistent when the correct spin-alignment is assumed in calculating the weights, this is not true for incorrect hypotheses. If the correct weights are w_e and w_μ for electrons and muons, respectively, then in a particular acceptance bin (where the weights are constant)

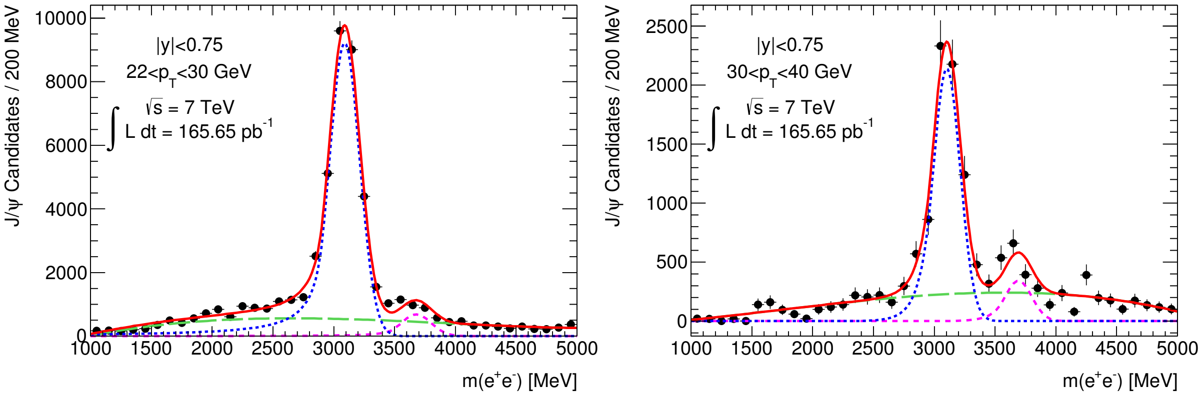
$$w_e N_e = w_\mu N_\mu \tag{4.31}$$



(a) 12–14 GeV (b) 14–16 GeV



(c) 16–18 GeV (d) 18–22 GeV

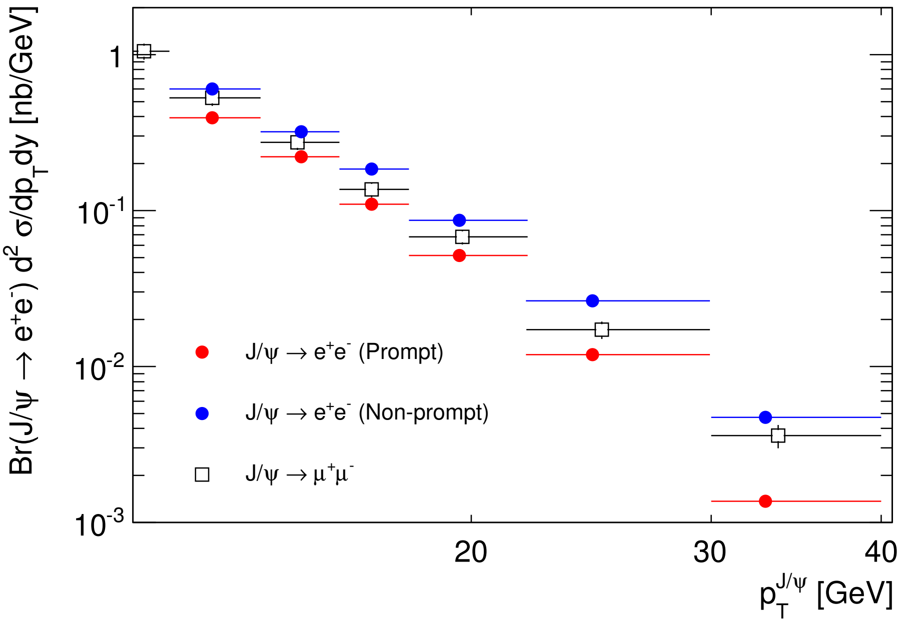


(e) 22–30 GeV (f) 30–40 GeV

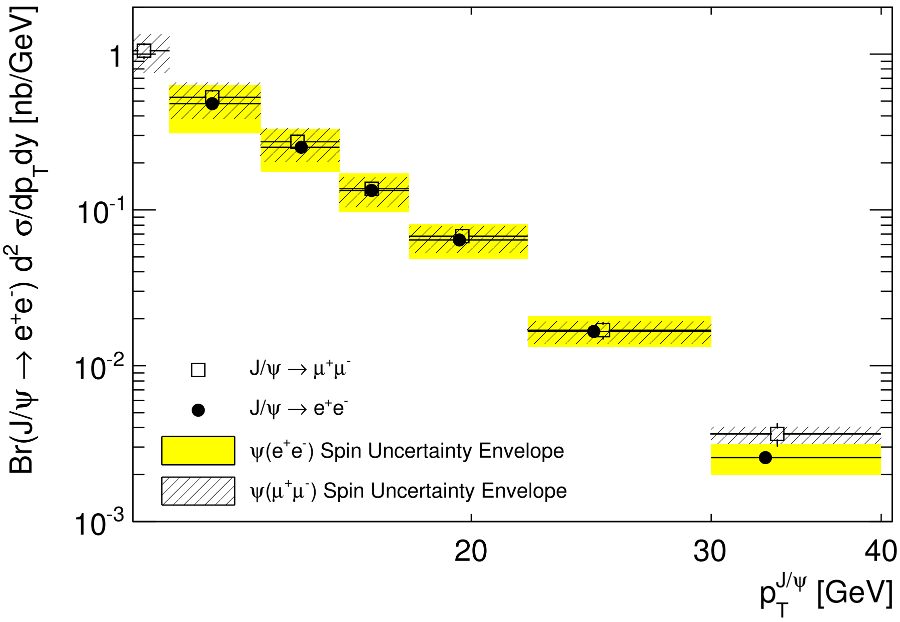
Figure 4.16: Weighted (mixed prompt/non-prompt) invariant mass distributions for $|y| < 0.75$ in various p_T bins.

Table 4.6: Summary of fit results for $|y| < 0.75$ (using non-prompt fractions).

p_T bin [GeV]	$< p_T >$ [GeV]	Mean [GeV]	σ	α	n	$\chi^2/n.d.o.f.$	J/ψ Yield	$\psi(2S)$ Yield	$\frac{d^2\sigma}{dp_T dy} Br(J/\psi \rightarrow e^+e^-)$ [pb/GeV]
12–14	12.9	3076 ± 2	142 ± 2	1.16 ± 0.07	3 ± 1	1.6	16400 ± 400	840 ± 80	480 ± 13
14–16	15	3088 ± 3	131 ± 3	0.85 ± 0.04	137 ± 3	2	8600 ± 100	340 ± 40	253 ± 3
16–18	16.9	3086.2 ± 0.8	126.7 ± 0.8	1.11 ± 0.03	8 ± 2	1.8	66300 ± 500	2100 ± 100	133.4 ± 0.9
18–22	19.6	3083.4 ± 0.9	133.1 ± 1.0	1.52 ± 0.04	1.4 ± 0.2	2.1	64000 ± 1000	4200 ± 200	64 ± 1
22–30	24.6	3090 ± 1	124 ± 1	1.48 ± 0.04	1.4 ± 0.2	2.8	32900 ± 600	2500 ± 100	16.5 ± 0.3
30–40	32.9	3103 ± 1	110.2 ± 0.4	1.2206 ± 0.0004	146.791 ± 0.005	2.4	6390 ± 80	1020 ± 40	2.57 ± 0.03



(a) Comparison of the J/ψ differential production cross-sections in $|y| < 0.75$ from the muon analysis with the upper and lower bounds on the $J/\psi \to e^+e^-$ values, obtained by assuming all J/ψ to be non-prompt and prompt, respectively. Uncertainties shown (though small enough to be obscured by the polymarker) are statistical for the electron results, and statistical and systematic combined in quadrature for the muon results. In all cases the J/ψ decays were assumed isotropic.



(b) Comparison between the J/ψ differential production cross-sections in $|y| < 0.75$ from dimuon and dielectron decays (for which a mixed prompt/non-prompt weight prescription was used). The error bars on the points are statistical (though are small enough to be obscured by the polymarker) for the case of electrons, and combined (in quadrature) statistical and systematic for muons. The main theoretical uncertainty concerning the unknown J/ψ spin-alignment is also shown.

Figure 4.17: Comparison of $J/\psi \to \mu^+\mu^-$ and $J/\psi \to e^+e^-$ cross-section measurements

Table 4.7: The differential cross section measurements for $|y| < 0.75$ using the mixed weights deduced from the prompt/non-prompt production ratio. The first uncertainty is statistical and the second due to the unknown J/ψ polarisation. The equivalent results for $J/\psi \to \mu^+\mu^-$ are shown in parallel, with an extra uncertainty term to account for all other systematic effects (i.e. other than spin-alignment).

p_T bin [GeV]	$< p_T^{e^+e^-} >$ [GeV]	$\frac{d^2\sigma}{dp_T dy} Br(J/\psi \to e^+e^-)$ [pb/GeV]	$< p_T^{\mu^+\mu^-} >$ [GeV]	$\frac{d^2\sigma}{dp_T dy} Br(J/\psi \to \mu^+\mu^-)$ [pb/GeV]
12–14	12.9	$480 \pm 13\ ^{+150}_{-170}$	12.9	$528 \pm 17\ ^{+127}_{-141}\ ^{+56}_{-56}$
14–16	14.9	$252 \pm 3\ ^{+77}_{-78}$	14.9	$274 \pm 12\ ^{+60}_{-70}\ ^{+27}_{-27}$
16–18	16.9	$133.4 \pm 0.9\ ^{+38.0}_{-37.4}$	16.9	$136.2 \pm 7.5\ ^{+26.5}_{-32.1}\ ^{+13.1}_{-13.1}$
18–22	19.6	$64 \pm 1\ ^{+17}_{-15}$	19.7	$67.7 \pm 3.6\ ^{+10.9}_{-14.5}\ ^{+6.4}_{-6.3}$
22–30	24.6	$16.5 \pm 0.3\ ^{+4.0}_{-3.3}$	25.0	$16.9 \pm 1.4\ ^{+2.2}_{-3.0}\ ^{+1.7}_{-1.7}$
30–40	32.9	$2.57 \pm 0.03\ ^{+0.56}_{-0.58}$	33.6	$3.60 \pm 0.48\ ^{+0.43}_{-0.52}\ ^{+0.38}_{-0.39}$

where $N_{e,\mu}$ are the number of electron and muon J/ψ candidates reconstructed in that bin. Under a different polarisation assumption, the weight is approximately[24] given by $w'_{e,\mu} = \mathcal{A}_{e,\mu}/\mathcal{A}'_{e,\mu} w_{e,\mu}$, and so 4.31 can be written

$$w'_e N_e = \frac{\mathcal{A}'_e \mathcal{A}_\mu}{\mathcal{A}_e \mathcal{A}'_\mu} w'_\mu N_\mu =: r_{\mathcal{A}} w'_\mu N_\mu \tag{4.32}$$

with $r_{\mathcal{A}}$ defined by this expression. In general, there is no reason to expect that $r_{\mathcal{A}}$ should be identically equal to one. On the contrary, if $r_{\mathcal{A}} > 1$ the resultant cross-section measurement would be higher for electrons than for muons and vice versa. Physically, this would correspond to a ratio in acceptance between the incorrect and correct polarisation assumptions which is different for electrons than it is for muons.

- **Incorrect MC Efficiencies.** All of the efficiencies used in calculating the weights were derived from the MC samples. Ideally, this would be replaced by data-driven efficiency measurements. Currently, this is only being conducted for individual electron efficiencies which were found to be unusable for analyses such as these (the J/ψ efficiencies do not *factorise* into electron efficiencies). Even if they were, the current methods are unable to separate the prompt and non-prompt electron

---

[24] i.e. ignoring the slight changes that result in the reconstruction, identification and trigger efficiencies.

efficiencies. In any case, it is quite common for the data-driven efficiencies to be systematically offset to those from MC samples. For example, measurements of the tight electron identification efficiencies from $Z \to e^+e^-$ decays using 2010 data for $15 < p_T < 20$ GeV and $|\eta| < 2.5$ were lower than the corresponding MC values by 0.15 [99]. In this analysis, MC was used for reconstruction, identification and trigger efficiencies, and as such even a small offset in each will produce a large difference in the overall weights (and therefore cross-sections). To demonstrate this, if a 3% difference is assumed for each efficiency in the p_T bin of 12–14 GeV, a rough calculation gives a $15 - 20\%$ change in the resultant weight.

- $J/\psi \to e^+e^-$ **Systematic Uncertainties.** The only (though, most significant) systematic considered was that of the unknown polarisation. However, there are likely to be several other contributions which could inflate the uncertainties by a substantial amount. This would include the uncertainty on the luminosity (typically 3–4 % [100]) and the effects of the fitting range and functions chosen for the signal and background.

4.9 Discussion

The purpose of this analysis was determine the feasibility of using electrons in quarkonium physics by carrying out the core steps of a J/ψ production cross-section analysis. It was shown in Section 4.5 that the dielectron final state can be used for the *observation* of particle states such as the J/ψ and $\psi(2S)$. For this to be true for other quarkonium states requires, in particular and with comparison to an equivalent muon analysis, (1) larger separation from other states and (2) larger production rates. The former is driven by the poorer resolution of the e^+e^- invariant mass peak, which is typically ~2-3 times larger than for $\mu^+\mu^-$ in the central rapidity range considered here. The second requirement is due mostly to the low efficiency for electron identification at low p_T, but is also related to the mass resolution — for a given number of particles produced, N, the statistical significance is strongly correlated to the width of the peak (in the limit of extremely poor resolution, the significance is zero, for example). So, for example, the $\Upsilon(1S)$, $\Upsilon(2S)$ and $\Upsilon(3S)$ would be very difficult to discern with electrons.

This is an unfortunate restriction, and is a direct result of the resolution of the electromagnetic calorimeter. The momentum from electron tracks was not used in this analysis because (at the time) there was no adequate accounting for bremsstrahlung losses. However, algorithms to recover these losses and determine more accurate electron track parameters have since been implemented and are an ongoing project within the

ATLAS reconstruction framework. Preliminary results show that invariant mass peak widths using track momenta from electrons reconstructed in this manner are smaller (by a factor of up to ∼2) than when using calorimeter information [86]. These resolutions are still not quite at the level of those for muons, but the gains made lessen the two requirements stated above.

Even if one were able to use electrons to observe a significant peak in an invariant mass distribution, it would not be of great use due to mis-calibration and p_T dependence of the electron energies — if it weren't already well known, how would the mass of the J/ψ observed in this study be determined? In fact, the same bremsstrahlung-aware reconstruction software also produced large improvements in this aspect of the analysis as well [86].

For measurements such as differential cross-sections, prompt/non-prompt production fraction and polarisation it is typically not the number of observed J/ψ that is desired, but rather the number produced. This necessitates the need to *unfold* the acceptance of the detector and the inefficiencies of the reconstruction, identification and trigger algorithms. The current implementation of weighting each observed J/ψ (see Section 4.6), which worked successfully for muons was found to have several significant issues. The most important of these are:

- **Migration.** The standard reconstruction algorithm for electrons fails to account for energy loss processes, the most important of which being bremsstrahlung. As a result, the measured transverse energy of the electrons tends to be systematically underestimated.[25] Since two electrons are used in the construction of a J/ψ candidate, the effect on the combined four-vector will be doubled (e.g. since $E_{J/\psi} = E_1 + E_2$, if $E_{1,2} \to E_{1,2} - \delta$ then $E_{J/\psi} \to E_{J/\psi} - 2\delta$).

  By necessity, the acceptance maps and reconstruction efficiencies are binned according to the generated kinematic variables — in particular, the J/ψ p_T and $|y|$. When carrying out the analysis on data, only the (mis-)measured values of these variables are available, meaning incorrect acceptances and efficiencies are used. Depending on the trends in the region surrounding the bin of interest, this will either lead to an over or under-estimated weight.

  Furthermore, in the calculation of the acceptance, selections are placed on the electron E_T and $|\eta|$ to emulate the effect of the trigger (e.g. for the *EF_2e5_tight_Jpsi* trigger, both electrons are required to have $E_T > 5$ GeV with $|\eta| < 2.47, |\eta| \notin$

---

[25] Bremsstrahlung typically increases the curvature of the fitted track, lowering the measured momentum and therefore also the energy.

$[1.37, 1.52]$). In the analysis this requirement is placed on the measured values. Since the transverse energy cuts are lower thresholds and the measurements are systematically low, this tends to cause an over-estimate of the acceptance and, therefore, a reduction in the weight.

Anything that can improve the measurement of the electron four-vectors would alleviate the problems caused by migration. Although it has not been verified, the bremsstrahlung recovery software mentioned above should provide significant gains.

- **Prompt/Non-prompt weights.** Muons produce tracks in the Inner Detector (ID) and Muon Spectrometer (MS), leaving only a very small signal in the electromagnetic calorimeter. The ID was designed to handle large charged-track multiplicity, and only the muons[26] interact with the MS. As such, the reconstruction and identification efficiencies for $J/\psi \to \mu^+\mu^-$ decays do not depend strongly on whether the J/ψ was produced promptly or non-promptly (i.e. with neighbouring tracks from the parent B-hadron decay). In contrast, electron reconstruction and identification relies heavily on the electromagnetic calorimeter, and explicitly implements isolation criteria. Consequently, the efficiencies for electrons from J/ψ produced non-promptly are significantly lower than for those produced promptly[27] (see Section 4.6 for more details).

  An important question therefore arises as to which weight to assign to a particular J/ψ candidate (prompt or non-prompt), for which there is no straightforward answer. The pseudo-proper time of the $J/\psi \to e^+e^-$ vertex, or the ability to reconstruct a B-hadron involving this J/ψ candidate, could be used to infer one way or the other. However, there will generally always be some level of uncertainty associated to this allocation. For example, one may assume that a pseudo-proper time of $\tau > 1$ ps provides good evidence for non-prompt production, but the resolution on this variable means that there is still a significant number of prompt J/ψ with values at least that large.

  In this study, another method was developed to deal with this issue, whereby a *mixed* prompt/non-prompt weight was determined from the $J/\psi \to \mu^+\mu^-$ non-prompt fraction measurement and assigned to every candidate. However, this removes the independence between the electron and muon channels, which is undesirable.

---

[26] Ignoring rare contributions from particles that 'punch' through the calorimeters, or those produced in cosmic ray events.

[27] This also holds for isolated and non-isolated electrons in general.

Whichever method is used, it is bound to contribute additional (in comparison to the muon analysis) systematic uncertainties. Fortunately, this will only be an issue when the quarkonia in the decay of interest can be produced in a non-prompt fashion.

- **Non-factorisation of individual electron efficiencies.** There are dedicated efforts within the e/γ community of ATLAS to perform data-driven measurements of the efficiencies for electron identification and triggering over a large range of transverse energies and pseudo-rapidities. Unfortunately, the electrons from J/ψ decays are not particularly isolated and the efficiency for one electron generally depends on the other — in other words, they are correlated. As such, the overall J/ψ efficiencies are not simply the product of the individual electron efficiencies and so the above mentioned measurements cannot be applied.
- **Inability to derive data-driven efficiencies.** From the above point and previous considerations, overall (e^+e^-) data-driven efficiencies would need to be measured separately for prompt and non-prompt J/ψ. This once more raises the problem of distinguishing the two. There is already an established study to determine separate prompt/non-prompt individual electron efficiencies, which utilises pseudo-proper time (or equivalent) information. At the time of writing (late 2012), however, this analysis is incomplete and faces significant difficulties.

The last major difficulty concerns the statistics with which the dielectron channel is able to contribute. The efficiencies (in particular, the trigger efficiencies) are lower than the corresponding muon efficiencies (for example, for most of the $p_T - y$ range the efficiency for *EF_e9_tight_e5_tight_Jpsi* trigger is less than 0.5). So, even though the branching fraction to e^+e^- is essentially equal to that of $\mu^+\mu^-$, the number of recorded decays of the former will be far less than that of the latter. Aside from this, as the electron channel is much noisier than the muon channel (many more particles can *fake* an electron than a muon), the electron trigger thresholds are consistently set much higher than for muons. The muon channel consequently provides better coverage of the quarkonium phase space than does the electron channel.

One possible advantage that electrons may hold is in the forward region of $|\eta| > 2.5$, which is not covered by either the ID or MS — i.e. muons are not reconstructable here. The forward electromagnetic calorimeter (FCAL) was specifically designed to cater for $2.5 < |\eta| < 4.9$, and therefore provides an advantage to electrons over the muon channel. Not only does the inclusion of such *forward* electrons extend the range of the electron acceptance to higher rapidities, but it also results in a general increase across the entire $p_T - y$ plane (for low p_T, even if the J/ψ is produced centrally a reasonable fraction of the decay electrons will be in forward directions). However, without the aid of tracking the task of electron reconstruction and identification becomes considerably more difficult.

4.9.1 Applicability to Future Onia Analyses

In light of the issues explained regarding electrons, their use is most suitable in situations where

- only prompt quarkonia participate in the decay channel,
- the electrons are not the main objects being triggered on,

and ideally,

- the use of weights is unnecessary.

Several analyses can been identified that fulfil these criteria. The most applicable analysis are searches for rare events with double-onia decays — specifically, $J/\psi + J/\psi$, $\Upsilon + \Upsilon$ and $J/\psi + \Upsilon$. In this case, the inclusion of electrons (naively) quadruples the statistics[28] and is a valuable addition. Likewise, analysis of Z/W-boson production with associated quarkonia benefits greatly by including electrons. In both of these cases, the quarkonia are produced promptly and if one of the particles decays to muons these can be used as the main triggering objects. The dielectron final state may also be useful in searches for rare decays where weighting is not necessary and statistics are crucial.

To facilitate the use of the dielectron final state in such analyses, a software tool called *JpsiEEFinder*[29] was developed in analogy with the *JpsiFinder* for muons. This is described in Appendix 4.B.

4.10 Conclusion

The motivation for this study centred around the prospect of using the dielectron final state, along with the preferred dimuon final state, for quarkonia analyses. The branching fractions for these two decays are equal (to within experimental uncertainty), and so electrons may provide a valuable addition in terms of statistics, at least. Furthermore, electrons may have advantages over, or be complementary to, muons, in terms of systematic uncertainties or acceptance properties.

---

[28] The electron efficiencies are generally poorer than those of the muons, so this is not quite true.

[29] This is a bit of a misnomer — in fact, there is nothing that makes this tool specific to the J/ψ.

To investigate this idea, the $J/\psi \to \mu^+\mu^-$ inclusive differential production cross-section analysis was repeated with electrons, using standard tools and procedures. Several major difficulties were confronted in the process. The most significant of these was due to bremsstrahlung, which causes the measured kinematic variables of electrons to *migrate* from their true values. Consequently, the weighting method failed and gave incorrect results. This was somewhat recovered by introducing a migration scale factor to account for this, but a more fundamental solution is desirable.

Another problem was met in applying the weights to the candidates — for electrons, the promptly produced J/ψ have different efficiencies to those produced non-promptly. For each specific candidate, however, there is no way of conclusively determining by which mechanism it was produced. A *mixed* weight was used to attempt to account for this, though this is again just a placeholder for a more accurate and robust procedure.

The efficiency for J/ψ reconstruction, identification and triggering were also found to be dependent on the configuration of the two decaying electrons — in other words, the overall J/ψ efficiencies do not factorise into the individual electron efficiencies.

Despite these issues, the differential cross-sections were measured in six p_T bins in the central rapidity region, $|y| < 0.75$, with results that were consistent with the $J/\psi \to \mu^+\mu^-$ values (the differences between the two sets of values gave a $\chi^2/n.d.o.f.$ value of 0.7).

A possible advantage identified for electrons over muons is in the forward region of $2.5 < |y| < 4.9$, where the coverage provided by the forward electromagnetic calorimeters allows for a larger phase space and higher acceptance. This is a possible avenue for future studies.

Many of the above issues concerning electrons could be alleviated through the use of bremsstrahlung recovery techniques, and is a strong suggestion for any future quarkonium analyses using electrons. Other issues may be avoided by choosing appropriate analyses — in particular, those where only prompt quarkonia participate in the final state, the electrons are not the main objects being triggered on, or where the use of weights or efficiency unfolding is not required.

Appendix 4.A Derivation of Systematic Uncertainty for Closure Fractions

In the third closure test (Section 4.7.3) on independent MC samples, the statistical uncertainty (Formula 4.14) for the closure fractions was stated without proof. This is derived below.

Firstly, recall that the individual weights within a particular acceptance bin are calculated as follows:

$$\begin{aligned} w^{-1} &= \mathcal{A} \cdot \varepsilon_{\text{reco.}} \cdot \varepsilon_{\text{id.}} \cdot \varepsilon_{\text{trig.}} \\ &= \mathcal{A} \cdot \frac{T_{\text{reco.}}}{T_{acc.}} \cdot \frac{N_{\text{id.}}}{N_{\text{reco.}}} \cdot \frac{N_{\text{trig.}}}{N_{\text{id.}}} \\ &\approx \mathcal{A} \cdot \frac{N_{\text{trig.}}}{T_{acc.}} \end{aligned} \tag{4.33}$$

where the first and second lines are just the definition of the terms, and the third line contains the crucial approximation of no migration (i.e. $N_{\text{reco.}} \approx T_{\text{reco.}}$). Now, consider the sum over the weights of J/ψ candidates in the same sample used to determine the efficiencies, which appears as the denominator of migration scale factors used in this test. In each analysis bin, the efficiencies were calculated for three equal rapidity strips (i.e. $|y| < 0.25, 0.25 < |y| < 0.5$ and $0.5 < |y| < 0.75$), allowing the sum in this bin to be split up and factorised accordingly:

$$\sum_{j=1}^{N} w_j = \frac{T_{a,1}}{N_1} \sum_{k=1}^{N_1} \mathcal{A}_{k,1}^{-1} + \frac{T_{a,2}}{N_2} \sum_{k=1}^{N_2} \mathcal{A}_{k,2}^{-1} + \frac{T_{a,3}}{N_3} \sum_{k=1}^{N_3} \mathcal{A}_{k,3}^{-1}. \tag{4.34}$$

In terms of notation, the numbered subscripts refer to the three efficiency sub-bins, a stands for acceptance and the *trig* subscript has been dropped from the N for simplicity. In other words, this equation is just obtained by substituting Equation 4.33 into the sum of weights over a typical analysis bin. As it stands, the N_i and $T_{a,i}$ are correlated to each other (N_i is essentially a subset of $T_{a,i}$). This was avoided by expanding $T_{a,i}$ into the number of J/ψ in that bin passing the acceptance and all other analysis cuts, N_i, and the number in the acceptance region but not passing the remaining cuts, $\overline{N}_i$ (again, migration is assumed to be zero here). Equation 4.34 can then be rewritten in terms of these variables as:

$$\begin{aligned} S := \sum_{j=1}^{N_{trig}} w_j &= \frac{N_1 + \overline{N}_1}{N_1} \sum_{k=1}^{N_1} \mathcal{A}_{k,1}^{-1} + \frac{N_2 + \overline{N}_2}{N_2} \sum_{k=1}^{N_2} \mathcal{A}_{k,2}^{-1} + \frac{N_3 + \overline{N}_3}{N_3} \sum_{k=1}^{N_3} \mathcal{A}_{k,3}^{-1} \\ &= \sum_{i=1}^{3} \left[\left(1 + \frac{\overline{N}_i}{N_i}\right) \sum_{k=1}^{N_i} \mathcal{A}_{k,i}^{-1} \right], \end{aligned} \tag{4.35}$$

where the sum is now denoted S. Any two $\mathcal{A}_{k,i}$ are independent of one another so long as k or i are distinct. When this is not the case, the two will be correlated and so it is instructive to split up the sum over the acceptances as

$$\sum_{k=1}^{N_i} \mathcal{A}_{k,i}^{-1} = \sum_{r=1}^{M_i} q_{r,i} \mathcal{A}_{r,i}^{-1}. \tag{4.36}$$

Here, $q_{r,i}$ is the multiplicity of J/ψ candidates in the rth acceptance bin of the ith efficiency sub-bin. The sum over the multiplicities of each of the M acceptance bins has to equal the total number of candidates in the efficiency sub-bin, i.e. $\sum_{r=1}^{M} q_{r,i} = N_i$. Then, substituting Equation 4.36 into 4.35 and rewriting the N_i in terms of the $q_{r,i}$ variables yields

$$S = \sum_{i=1}^{3} \left[\left(1 + \frac{\overline{N}_i}{\sum_{r=1}^{M} q_{r,i}} \right) \sum_{r=1}^{M} q_{r,i} \mathcal{A}_{r,i}^{-1} \right] \tag{4.37}$$

Every variable that appears in this expression is now independent of all others (since the acceptance was not calculated with the MC samples). In a similar way, another equivalent variable for the sum of weights can be defined for the half of the sample in which the test is being performed:

$$S' := \sum_{i=1}^{3} \left[\left(1 + \frac{\overline{N}_i}{\sum_{r=1}^{M} q_{r,i}} \right) \sum_{r=1}^{M} q'_{r,i} \mathcal{A}_{r,i}^{-1} \right], \tag{4.38}$$

the only difference being on the multiplicities in each of the acceptance bins (primes represent the *testing* half of the sample). The migration-scaled closure fraction (Equation 4.13) for a specific analysis bin can be expressed as

$$C = \frac{T_{prod.}}{T'_{prod.}} \cdot \frac{S'}{S} \tag{4.39}$$

Now, $T'_{prod.}$ is independent of both of the sums (remembering that N'_i does not appear directly) and $T_{prod.}$. On the other hand, $T_{prod.}$ itself is correlated to the variables in S and S', and so was split up as $T_{prod.} = \overline{T}_{acc.} + \sum_{i=1}^{3} T_{a,i} = \overline{T}_{acc.} + \sum_{i=1}^{3} \left(N_i + \overline{N}_i \right)$. The new variable introduced here, $\overline{T}_{acc.}$, is the number of J/ψ produced in the analysis bin but not passing the acceptance cuts. Finally, then, the closure fraction is given by:

$$\begin{aligned} C &= \left(\frac{\overline{T}_{acc.} + \sum_{i=1}^{3} \left(\sum_{r=1}^{M} q_{r,i} + \overline{N}_i \right)}{T'_{prod.}} \right) \cdot \frac{S'}{S} \\ &= \left(\frac{\overline{T}_{acc.} + \sum_{i=1}^{3} \left(\sum_{r=1}^{M} q_{r,i} + \overline{N}_i \right)}{T'_{prod.}} \right) \cdot \frac{\sum_{i=1}^{3} \left[\left(1 + \frac{\overline{N}_i}{\sum_{r=1}^{M} q_{r,i}} \right) \sum_{r=1}^{M} q'_{r,i} \mathcal{A}_{r,i}^{-1} \right]}{\sum_{i=1}^{3} \left[\left(1 + \frac{\overline{N}_i}{\sum_{r=1}^{M} q_{r,i}} \right) \sum_{r=1}^{M} q_{r,i} \mathcal{A}_{r,i}^{-1} \right]}. \end{aligned} \tag{4.40}$$

Since each of the variables in this equation are statistically independent, the standard error propagation formula can be used to determine the uncertainty on C:

$$\begin{aligned} \sigma_C^2 = \sum_{i=1}^{3} &\left[\left(\frac{\partial C}{\partial \overline{N}_i} \right)^2 \sigma_{\overline{N}_i}^2 + \sum_{r=1}^{M} \left[\left(\frac{\partial C}{\partial q_{r,i}} \right)^2 \sigma_{q_{r,i}}^2 + \left(\frac{\partial C}{\partial q'_{r,i}} \right)^2 \sigma_{q'_{r,i}}^2 + \left(\frac{\partial C}{\partial \mathcal{A}_{r,i}^{-1}} \right)^2 \sigma_{\mathcal{A}_{r,i}^{-1}}^2 \right] \right] \\ &+ \left(\frac{\partial C}{\partial \overline{T}_{acc.}} \right)^2 \sigma_{\overline{T}_{acc.}}^2 + \left(\frac{\partial C}{\partial T'_{prod.}} \right)^2 \sigma_{T'_{prod.}}^2 \end{aligned} \tag{4.41}$$

The partial derivatives with respect to $\overline{T}_{acc.}$ and $T'_{prod.}$ are simply given by $C/T_{prod.}$ and $-C/T'_{prod.}$, respectively. After some algebra, the remaining derivatives are:

$$\begin{align}
\left(\frac{\partial C}{\partial q_{r,i}}\right) &= C\left[\frac{1}{T_{prod.}} + \frac{\overline{N}_i}{N_i^2}\left(\frac{\sum_{r=1}^{M} q_{r,i}\mathcal{A}_{r,i}^{-1}}{S} - \frac{\sum_{r=1}^{M} q'_{r,i}\mathcal{A}_{r,i}^{-1}}{S'}\right) - \frac{\mathcal{A}_{r,i}^{-1}}{S}\left(1+\frac{\overline{N}_i}{N_i}\right)\right] \nonumber \\
&= C\left[\frac{1}{T_{prod.}} + \frac{\overline{N}_i}{N_i^2}D_i - \frac{\mathcal{A}_{r,i}^{-1}}{S}\left(1+\frac{\overline{N}_i}{N_i}\right)\right] \tag{4.42} \\
\left(\frac{\partial C}{\partial q'_{r,i}}\right) &= C\left[\frac{\mathcal{A}_{r,i}^{-1}}{S'}\left(1+\frac{\overline{N}_i}{N_i}\right)\right] \tag{4.43} \\
\left(\frac{\partial C}{\partial \overline{N}_i}\right) &= C\left[\frac{1}{T_{prod.}} - \frac{D_i}{N_i}\right] \tag{4.44} \\
\left(\frac{\partial C}{\partial \mathcal{A}_{r,i}^{-1}}\right) &= C\left(1+\frac{\overline{N}_i}{N_i}\right)\left(\frac{q'_{r,i}}{S'} - \frac{q_{r,i}}{S}\right) \tag{4.45} \\
& \tag{4.46}
\end{align}$$

where the D_i are defined by Equation 4.42. The uncertainty on any $\mathcal{A}_{r,i}$ is related to that of its reciprocal by $\sigma^2_{\mathcal{A}_{r,i}^{-1}} = \sigma^2_{\mathcal{A}_{r,i}}/\mathcal{A}^4_{r,i}$. Given this, and the fact that the variances of each of the Poisson variables $q_{r,i}$, $q'_{r,i}$, $\overline{N}_i$, $\overline{T}_{acc.}$, $T'_{prod.}$ are just their own value, the final result is:

$$\begin{align}
\frac{\sigma_C^2}{C^2} = \frac{1}{T'_{prod.}} + \frac{1}{T_{prod.}} + \sum_{i=1}^{3}\left(1+\frac{\overline{N}_i}{N_i}\right)\Bigg[D_i^2\frac{\overline{N}_i}{N_i^2} + \sum_{r=1}^{M}\Bigg[\left(1+\frac{\overline{N}_i}{N_i}\right)\left(\frac{q'_{r,i}}{S'} - \frac{q_{r,i}}{S}\right)^2\frac{\sigma^2_{\mathcal{A}_{r,i}}}{\mathcal{A}^4_{r,i}} \nonumber \\
+\mathcal{A}_{r,i}^{-2}\left(1+\frac{\overline{N}_i}{N_i}\right)\left(\frac{q_{r,i}}{S^2} + \frac{q'_{r,i}}{S'^2}\right) - \frac{2}{S\mathcal{A}_{r,i}}\left(\frac{1}{T_{prod.}} + \frac{\overline{N}_iD_i}{N_i^2}\right)q_{r,i}\Bigg]\Bigg] \tag{4.47}
\end{align}$$

which is the formula used to calculate the uncertainties in the closure test.

□

Appendix 4.B The *JpsiEEFinder* Tool

As has been discussed, the default reconstruct of Quarkonia at ATLAS utilises the $\mu^+\mu^-$ decay channel. An ATHENA[30] software tool, *JpsiFinder*, processed events in search of muon pairs which can be fitted to a common vertex. In doing so, several further constraints (e.g. opposite sign, p_T thresholds and so on) can be imposed upon the individual muons, or their combined object. The $\mu^+\mu^-$ combination then serves as the input to further analysis, or can be read out into an ntuple structure. *JpsiFinder* currently underpins every analysis that utilises Quarkonia.

---

[30] The ATLAS (C++) software framework — see Chapter 2.

To enable electrons to be used in parallel with muons, an equivalent piece of software, *JpsiEEFinder*, was created. In analogy to *JpsiFinder*, this tool combines pairs of electrons within a particular event into possible J/ψ candidates. The algorithm proceeds as follows:

1. Select a pair of reconstructed electrons.
2. Retrieve the track associated with each electron (along with the corresponding covariance matrices for track parameters).
3. Perform a minimum-χ^2 common vertex fit to the tracks and update track parameters and covariance matrices accordingly.
4. Return the result as an *ExtendedVxCandidate* C++ object.

All of the variables of interest, such as the invariant mass and its uncertainty or the pseudo-proper time and its uncertainty, can then be deduced from the *ExtendedVxCandidates.*

At certain points in the above process, the tool allows for the user to make particular selections on the electrons or their combination. Some of these include:

- The sign of the electrons —- opposite or the same. The later may, for example, be used for background studies.
- The *author* of each electron — i.e. whether they were reconstructed with the *standard*, *soft* or *forward* electron algorithm.
- The quality of the electron tracks — for example, number of hits in the Pixel, SCT and TRT detectors, impact parameters etc.
- The electron *identification* — loose, medium, tight, or the corresponding '++' options (see chapter 2.3 for more detail on the meaning of these terms).
- Kinematic cuts on the p_T of either electron, the collimation between the two and their combined invariant mass.
- The χ^2 of the vertex fit.

Despite the naming of this tool, it can be used for a variety of masses aside from that of the J/ψ — for example, the Υ states. The electron tracks contained in the input to the *JpsiEEFinder* can easily be switched between several sources, including those reconstructed with the aid of bremsstrahlung-recovery techniques, as well as those from the standard reconstruction. For examples of the results from each of these, see [86].

Chapter 5

Search for $X_b \to \pi^+\pi^-\Upsilon(1S)$ Decays: Development of the Analysis

The previous chapter explored the possibility of using electrons for Quarkonia analyses at ATLAS. Focus is now turned to the main analysis – a search for new particles (e.g. X_b) in the hidden-beauty spectrum using the $\pi^+\pi^-\Upsilon(1S)$ transition. Due to the significant issues discovered in the previous chapter surrounding electrons, only the dimuon decay of the $\Upsilon(1S)$ was utilised. This decision was supported by the trigger menu used for 2011 and 2012 data taking periods, for which un-prescaled, low transverse momenta triggers specifically designed for the purpose of collecting $\Upsilon \to \mu^+\mu^-$ decays were available [1].

This chapter details the crucial first stage of the search; the development of the analysis using Monte Carlo (MC) simulated decays. After a short introduction motivating the analysis, the MC samples are then described in some detail, followed by an explanation of the analysis procedure. With this established, the signal and background fitting models for the $\pi^+\pi^-\Upsilon(1S)$ mass line-shape are discussed. Finally, projections are made as to the sensitivity of the analysis to the observation of new states.

5.1 Monte Carlo (MC) Samples

Monte Carlo (MC) samples were needed to establish and refine the analysis approach. This included determining appropriate

1. quality-ensuring requirements placed on the the muons, pions, $\Upsilon(1S)$ and $\pi^+\pi^-\Upsilon(1S)$ objects, and

---

[1] For example, an un-prescaled dimuon trigger with a p_T threshold of 4 GeV (for each muon) was used throughout 2011.

2. procedures to select particular $\pi^+\pi^-$ combinations for each Υ(1S) candidate.

As well as producing MC events for known or hypothetical particles decaying through the $\pi^+\pi^-\Upsilon$(1S) channel (i.e. *signal* processes), a sample representing a major constituent of the background was also utilised. After establishing the analysis approach, fitting models were developed to describe both the signal and background components of the $\pi^+\pi^-\Upsilon$(1S) mass spectrum. These shapes, along with calculated analysis efficiencies, were then used to make predictions concerning the discovery or exclusion prospects for new particles over a wide range of hypothetical masses. Each of these stages is described in detail throughout this chapter.

5.1.1 Signal Samples

The S-wave radial excitations of the Υ(1S) are known to decay into the $\pi^+\pi^-\Upsilon$(1S) final state. In particular, the Υ(2S) and Υ(3S) have branching fractions of $17.85 \pm 0.26\%$ and $4.37 \pm 0.08\%$ [3], respectively. For the purposes of this study, these serve as calibration points in the $\pi^+\pi^-\Upsilon$(1S) mass spectrum and are a good basis with which the construction and selection of candidates can be optimised. Samples of ∼1 million $\pi^+\pi^-\Upsilon$(1S) decays were produced for both of these parent states, with a filter to ensure the muons had a transverse momentum above 4 GeV and were within a pseudo-rapidity range of $|\eta| < 2.5$ [2].

Similar MC samples were also produced for the recently observed $\Upsilon(1^3\mathrm{D_J})$ states, with half a million decays for each possible value for the total angular momentum ($J = 1, 2, 3$). The two observations, at CLEO [28] and BaBar [29], only provided evidence for observation of the $J = 2$ decay with a combined uncertainty-weighted[3] mass close to 10164 MeV. The splitting of the triplet is a matter of debate within the theoretical community, though most agree that it should be of the order of 3-11 MeV with the $J = 2$ state sitting $0.5 - 1.0$ MeV below the centre-of-gravity [101]. Averaging the predictions from several sources in [101], and rounding to the closest MeV gives $m(\Upsilon(1^3\mathrm{D_1})) = 10156$ MeV and $m(1\Upsilon(1^3\mathrm{D_3})) = 10170$ MeV. Taking this as a reasonable assumption[4], the Υ(1D) triplet was generated with masses of $10156, 10164$ and 10170 MeV.

---

[2] Technically, the filter was placed to ensure that *any* two muons in the event passed these conditions, but the likelihood that these did not come from the signal decay is low.

[3] Specifically, a χ^2-best-fit to the two values of 10161.1 ± 1.71 and 10164.5 ± 0.94 from CLEO and BaBar, respectively. The errors quoted here are the quadrature-combined statistical and systematic uncertainties.

[4] For the purposes of this study, the chosen splitting is not critical. The $J = 1$ and $J = 3$ samples were produced to ascertain the ability of the analysis to separate the three states, for which any close mass splitting is appropriate.

In the first attempt at producing the Υ(2S) MC, a mistake was made in setting the parent mass to 10233 MeV, rather than 10023 MeV. Fortunately, this error meant that another MC mass point, which may be considered that of a hypothetical X_b state, was available for analysis. In particular, this provided another measurement of the fitting shapes and efficiencies in the region between the Υ(1D) and Υ(3S). For the remainder of this chapter, this sample will be referred to as that of the $X(10233)$.

All of the MC samples were generated with PYTHIA 6.4 [93] using the 2011 ATLAS tune [102] and filtered for the presence of the decay of interest. *Atlfast II* [103] was then used to simulate the response of the detector to the particles produced in the collision. This implements Geant4 [96] for some parts of the simulation (namely, the passage of particles through the Inner Detector and Muon Spectrometer) and a fast (i.e. parameterised) approach for the calorimeters. Since no calorimeter information was used in the analysis, little difference is expected between this and a full Geant4 simulation.

The production of the parent states in all the MC samples was modelled on Non-Relativistic Quantum Chromodynamics (NRQCD), while the available phase space for the ensuing three body decay[5] was sampled uniformly. The angular distribution for the Υ(1S) decay into two muons was also taken to be isotropic. These assumptions represent the ignorance as to the quantum numbers and decay dynamics of any potential new states, and so are appropriate for the purposes of this search. Furthermore, recent measurements from the CMS detector demonstrate that Υ(nS) (n $= 1, 2, 3$) states produced in pp collisions are only weakly polarised, if at all [7]. In the case that there is a significant deviation from pure phase-space, as in the $\Upsilon(2S) \to \pi^+\pi^-\Upsilon(1S)$, $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)$ [13] and $X(3872) \to \pi^+\pi^- J/\psi$ [4] dipion mass distributions, a simple reweighting scheme is sufficient to properly model the decays. This is necessary, for example, to form an accurate comparison between observables in the MC samples and collision data (see the next chapter for details). However, for the purposes of this chapter this level of precision was not necessary, and the MC samples were used as-is. For a more complete discussion of the kinematics and dynamics involved in a dipion transition between $b\bar{b}$ states, see Chapter 1.

The signal MC samples are summarised in Table 5.1, including the efficiencies for the muon filter.

---

[5] Three body decays $x \to a, b, c$ have 5 degrees of freedom, which are conventionally taken to be m^2_{ab}, m^2_{bc} and the three Euler angles giving the orientation of the decay plane with respect to the parent rest frame. See Appendix 1.A.

Table 5.1: Summary of the six different signal Monte Carlo (MC) samples used in this analysis.

Parent State	**Mass [MeV]**	**Size**	**Filter Efficiency**
Υ(2S)	10023	999499	0.188
$\Upsilon(1^3D_1)$	10156	500000	0.188
$\Upsilon(1^3D_2)$	10164	499000	0.188
$\Upsilon(1^3D_3)$	10170	499499	0.188
X(10233)	10233	998499	0.187
Υ(3S)	10355	997500	0.184

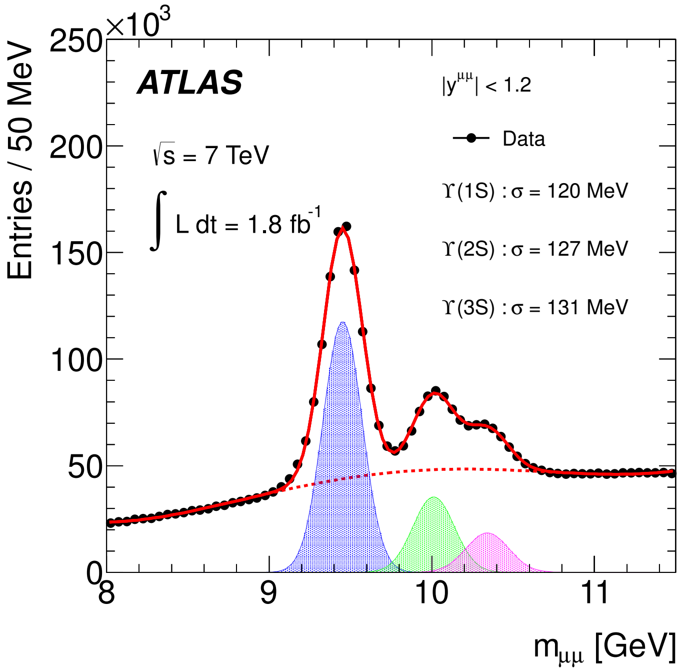


Figure 5.1: The $\mu^+\mu^-$ mass spectrum for central rapidities in the region of the $\Upsilon(1S, 2S, 3S)$, showing the contribution from the real Υ(1S) decays (blue shaded curve) and the background underneath this from the non-1S decays (red dashed line) [6].

5.1.2 Background

Background, in this case, refers to any $\pi^+\pi^-\Upsilon$(1S) candidate in which one or more of the pions or muons did not originate from the same parent particle. These can be split into two categories,

- **Inclusive Υ(1S) background** – in which the $\mu^+\mu^-$ come from a Υ(1S) decay, and
- **Non-1S background** – where the muons are not the product of a Υ(1S) decay.

The contributions from each are shown in Figure 5.1, which is the fitted $\mu^+\mu^-$ mass spectrum for central rapidities as observed in the ATLAS analysis of the Υ cross-section measurement [6]. The blue shaded curve represents the decays in which the $\mu^+\mu^-$ combination came from a real Υ(1S), while the non-1S background can be thought of as that underneath the red dashed curve around 9460.30 MeV. Each of these appears to be of similar magnitude.

The inclusive background could be further split into those cases in which the Υ(1S) was produced from a decay of our signal type (but one or both of the pions were not), and those for which the Υ(1S) was produced by other means (for example, directly in the hard collision). All of the signal MC samples are a source of the former, since events typically contain large numbers of unrelated pions which can be mistaken as from the signal decay. Aside for some possible kinematic differences, these sub-categories should be similar enough to consider them together.

A MC sample, in which an Υ(1S) was produced via any mechanism and decayed directly into two muons, was produced to model this source of background. The same generator, generator tune and simulator were used as for the signal samples, but only a very weak filter of $|\eta| < 10$ was used[6]. This dataset is referred to as the *inclusive* 1S *sample*. The polarisation of the Υ(1S) was again assumed to be such that the decay of the muons was isotropic.

The non-1S background originates from events in which two unrelated muons are combined, and happen to have an invariant mass close to that of the Υ(1S). For such combinations, collision data serves as a far more convenient source for analysis than does Monte Carlo. Typically, one would utilise sidebands to the left and right of the mass of the particle of interest. From Figure 5.1, however, there is significant contamination on the high mass side from the Υ(2S), rendering it useless. The lower mass sideband alone could be used, except in the occasion where there are mass-dependent effects (in which case, the upper mass sideband is useful in determining the trends or averaging these away). Another alternative is provided by *same-sign* muon combinations, though this can potentially introduce other biases due to charge-asymmetries. Both of these approaches are implemented in the next chapter, where collision data from 2011 is analysed.

---

[6] This background sample was not produced exclusively for this analysis, and in seeking generality, filtering was avoided.

5.2 Acceptance Considerations

As in the $J/\psi \to e^+e^-$ analysis of the previous chapter, *acceptance* here refers to the efficiency with which the decay can be reconstructed under certain kinematic constraints on the final state particles. The dependence of the acceptance on these constraints is one of the key guiding principles in determining optimal event and object kinematic selections.

For simplicity, consider the Υ(2S) $\to \pi^+\pi^-\Upsilon$(1S) decay under the same conditions used for the MC production; the p_T and rapidity, y, of the Υ(2S) determined by NRQCD modelling and the phase-space is uniformly sampled. This is not strictly accurate, but is sufficient to capture the qualitative features required here. A simple toy MC tool, which will be referred to as the *acceptance calculator*, was built to randomly generate signal decays under the above assumptions.

The typical energy available to each of the pions in the rest frame of the Υ(2S) is of similar size to half the difference between m_{2S} and m_{1S}, approximately 281 MeV, which is small compared to typical energy scales at ATLAS. This is true even in the lab frame (where the decay is boosted), as shown by the transverse momentum distribution of the two pions for 1 million simulated toy decays (Figure 5.2). It is obvious from these plots (especially that for the sub-leading pion) that it is desirable to use reconstructed pion tracks of the lowest possible p_T. Unfortunately, the minimum at ATLAS (at least in 2011 and 2012 data taking) is 400 MeV.

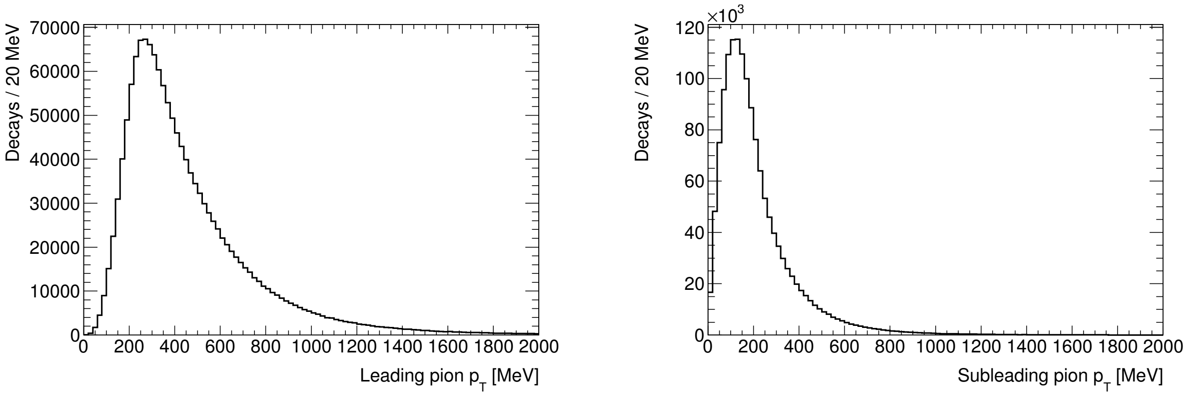
While the same is true for the muon tracks, the more demanding constraint comes from the fact that the muons are the primary objects used for triggering the event. In 2011, the lowest un-prescaled trigger selecting Υ(1S) $\to \mu^+\mu^-$ decays required that each muon have $p_T > 4$ GeV, which effectively sets the threshold for the muons. Fortunately, the relatively large mass of the Υ(1S) guarantees that a significant fraction of decays will satisfy this criterion.

As well as these, there is the physical condition that the four particles lie within the physical extent of the Inner Detector and, in the case of the muons, the muon trigger chambers. In terms of pseudo-rapidity, this corresponds to $|\eta^\pi| < 2.5$ and $|\eta^\mu| < 2.3$.

In summary, the following are the *kinematic acceptance conditions* for the analysis:

- **Muons** – $p_T > 4$ GeV, $|\eta| < 2.3$,
- **Pions** – $p_T > 400$ MeV, $|\eta| < 2.5$.

The acceptance for the muon criteria alone is approximately 17% (with respect to generator-level production), whilst with the additional conditions on the pion it is close to 1.8%.



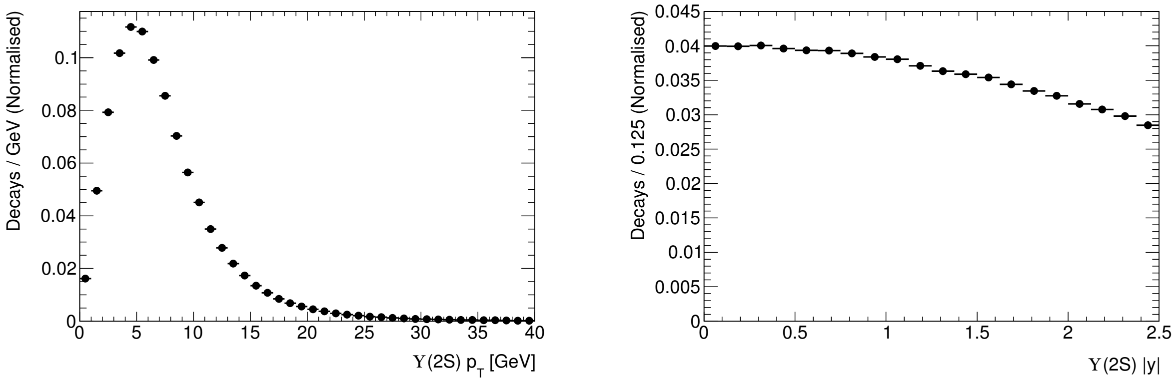
(a) (b)

Figure 5.2: The p_T distribution of the two pions in 1 million simulated Υ(2S) $\rightarrow$ $\pi^+\pi^-\Upsilon$(1S) decays. The harder of the two (leading) is shown on the left, and the softer (leading) on the right.

These selections have follow-on effects on the kinematics of the parent objects. In particular, selecting only the high end of the p_T spectrum coincides with highly boosted decays. Figure 5.3 demonstrates this, by showing the production p_T and y spectra according to NRQCD (top), the acceptance as a function of these variables (middle) and the product of these, which gives the resultant expected reconstructed spectra (bottom). As expected, there is a smooth rise in acceptance with the parent p_T, while it is fairly flat in rapidity, with a smooth decline at the edges of the detector coverage. Because the production spectrum is predominantly at low p_T, most of the reconstructable candidates are right at the threshold of acceptance (between 10-15 GeV).

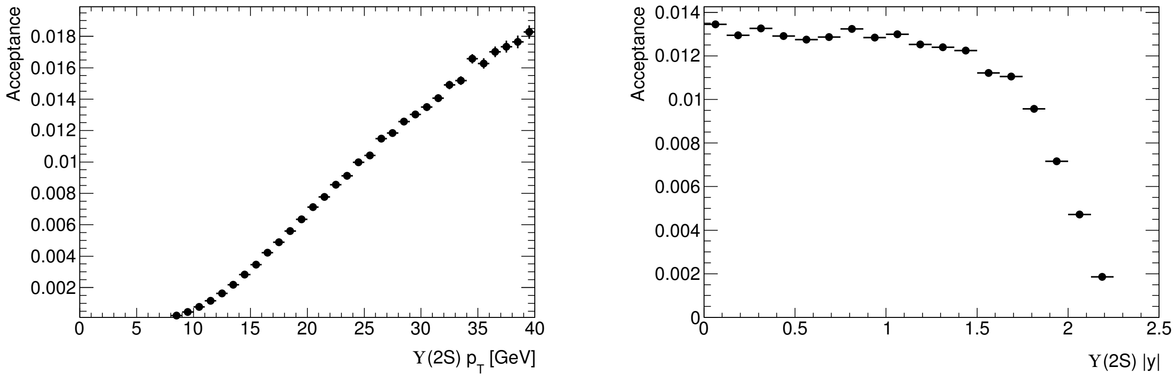
While this special case of the Υ(2S) decay captures most of the acceptance features, there are some which are clearly mass dependent. As was noted early, the natural scale for pion energy is determined primarily by the difference in mass between the parent particle and the Υ(1S). Assuming a similar production spectrum, this would suggest that higher mass states would generally have a larger acceptance. Moreover, the threshold p_T for reconstruction would be expected to drop to zero at some point. To illustrate this, the procedure above was repeated at the mass of the Υ(3S), Υ(5S), and an intermediate hypothetical mass of 10.6 GeV, with the same production p_T and rapidity spectrum. The resulting acceptance is given in Figure 5.4, along with the expected reconstructable spectrum.

As well as a significant increase across-the-board, there is a small pocket of increasing acceptance for the scenario where the parent state is produced at rest. In the parent rest frame, the Υ(1S) will also be close to rest and hence the muons will have similar



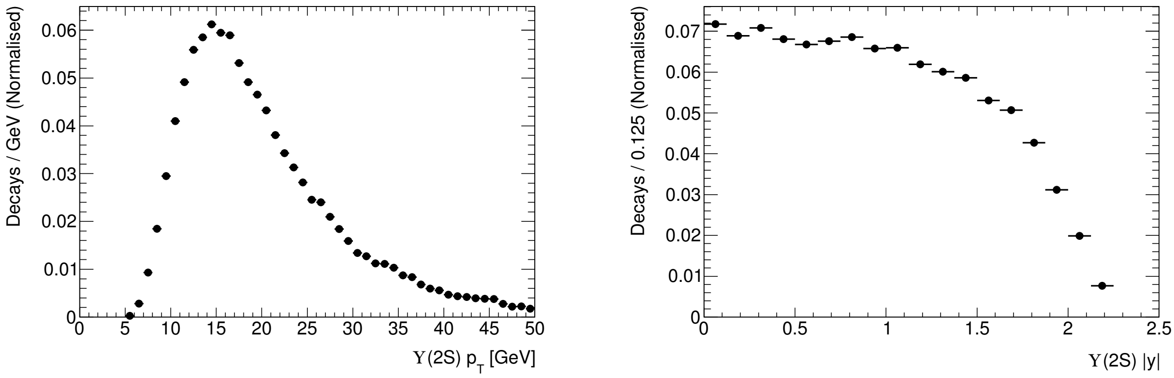
(a) Production vs. p_T.

(b) Production vs. $|y|$.



(c) Acceptance vs. p_T.

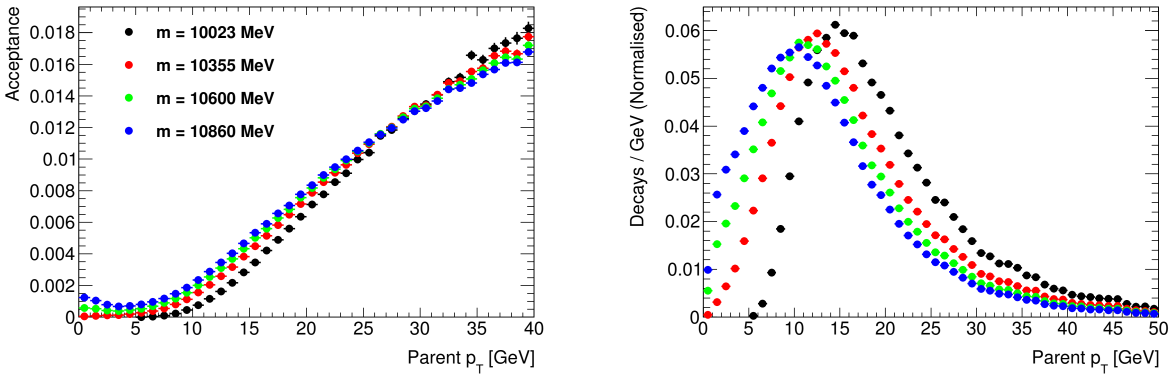
(d) Acceptance vs. $|y|$.



(e) Production × Acceptance vs. p_T

(f) Production × Acceptance vs. $|y|$

Figure 5.3: The NRQCD-based production spectrum in p_T and y for the $\Upsilon(2S)$ (top), the kinematic acceptance as a function of these variables (middle), and the product of these two (bottom).



(a) (b)

Figure 5.4: The acceptance as a function of parent p_T, with the resultant spectrum assuming the same production kinematics as for the Υ(2S) used above.

momenta. When their momentum lies in a direction close to purely transverse, and the parent mass is large enough, they are able to pass the acceptance conditions. If the parent state is mildly boosted (say, with $p_T < 5$ GeV), this symmetry is lost and the acceptance drops.

Even though this acceptance increase for the higher mass states at low p_T is marginal, the number of reconstructable decays in this region is significantly increased. These acceptance properties are exploited to remove a large amount of background, as explained in the next section.

5.3 Reconstruction and Selection of $\pi^+\pi^-\Upsilon(1\mathrm{S})$ Candidates

Keeping the acceptance features in mind, the reconstruction of $\pi^+\pi^-\Upsilon$(1S) candidates was built upon a well-established process developed for similar analysis, such as $\pi^+\pi^- J/\psi$, with similar or equivalent selections on the events and physics objects. The first stage involved organising and fitting all possible combinations of four tracks into vertices.

5.3.1 Reconstruction of $\pi^+\pi^-\Upsilon(1\mathrm{S})$ Vertices

In each event, the Inner Detector tracks of *Staco* muons, passing the *Muon Combined Performance* (MCP) cuts (see Chapter 2.3), and with a $p_T > 400$ MeV, were organised into pairs and subsequently fitted to a common vertex. Both muons were required to be *combined* – that is, have a matching Muon Spectrometer track. If their mass fell in

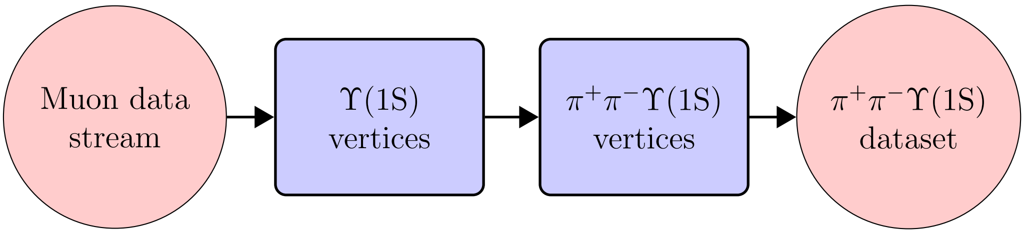


Figure 5.5: The major stages of $\pi^+\pi^-\Upsilon$(1S) candidate reconstruction.

a wide range surrounding m_{1S}, $8.5 < m_{\mu^+\mu^-} < 10.7$ GeV, they were designated *Upsilon Candidates*.

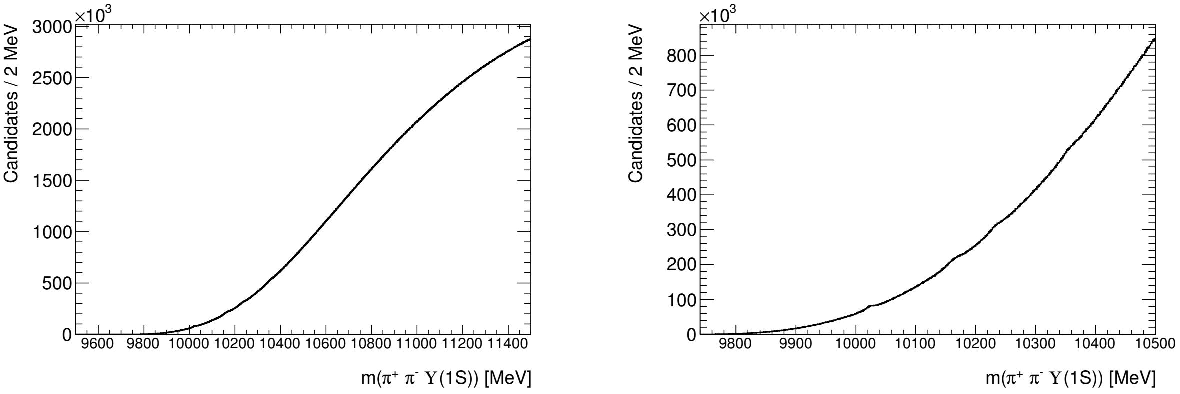
These *Upsilon Candidates* then served as the input in the construction of $\pi^+\pi^-\Upsilon$(1S) four-track vertices. Within the same event, all other pairs of oppositely charged tracks were added to each Υ(1S) candidate, and a vertex was fitted under the constraint that all particles emerged from the same point and $m(\mu^+\mu^-) = 9460.30$ MeV (the PDG world average for m_{1S} [3]). Constraining the Υ(1S) mass in the fit effectively removes the smearing due to the resolution on the Υ(1S) mass measurement, which (see Figure 5.1) is large. This leads to a resolution of the order of 6 MeV at the Υ(2S) mass and 11 MeV for the Υ(3S). Without such an improvement in resolution, the sensitivity of this study would be very low.

As for the muons, some loose selections were also placed on the pion tracks:

- $p_T > 400$ MeV,
- $|\eta| < 2.5$,
- ≥ 1 hits in the Pixel detector, and
- ≥ 2 hits in the SCT detector.

The latter two are common selections to reduce background from fake tracks, while the former two are the acceptance conditions for the pions, described in the previous section. Any four-track ($\mu^+, \mu^-, \pi^+, \pi^-$) vertices constructed in this manner falling in the range $9.5 < m < 11.5$ GeV were stored as $\pi^+\pi^-\Upsilon$(1S) *candidates*. This initial stage of reconstruction is summarised in Figure 5.5.

Of the original number of events in each MC sample, approximately 80% contained at least one $\pi^+\pi^-\Upsilon$(1S) candidate. This is primarily driven by the muon reconstruction efficiency, since all the samples were filtered to ensure the muons met their acceptance requirements (and almost all events contain at least one viable $\pi^+\pi^-$ pair). Figure



(a) (b)

Figure 5.6: The $\mu^+\mu^-\pi^+\pi^-$ mass spectrum for reconstructed candidates across all signal MC samples. The entire spectrum in the interval [9.5,11.5 GeV] is shown on the left, and a close up of the lower mass region on the right. At this stage, the peaks at each of the parent masses are almost completely dominated by the large combinatorial background.

5.6 shows the combined[7] mass distribution for all the signal MC samples at this stage of reconstruction. The low end of the spectrum is bounded by a threshold equal to $m_{1\text{S}} + 2m_\pi = 9740$ MeV [3], which precedes a rising shape at higher masses. This is due in part to the increased acceptance here, but mostly just the statistical tendency for background pions to form larger mass objects. If the distribution is restricted to the region $m < 10.5$ GeV, small bumps at each of the signal MC masses become noticeable. The background level within these signal samples alone is overwhelming, and with a proper luminosity-weighted signal-plus-background combination would not be observable at all.

Though it is uncommon to encounter an event with more than one Υ(1S) candidate, each collision typically produces hundreds of pions. Each Υ(1S) candidate can therefore give rise to hundreds of viable $\pi^+\pi^-\Upsilon$(1S) candidates, which is the cause of the large background levels. To quantify this, Figure 5.7 shows the multiplicity of Υ(1S) (left) candidates per event and $\pi^+\pi^-\Upsilon$(1S) candidates per Υ(1S) (right), for the Υ(2S) MC sample. For this analysis to be effective, the latter clearly needs to be significantly reduced. This is the purpose of the object quality selections.

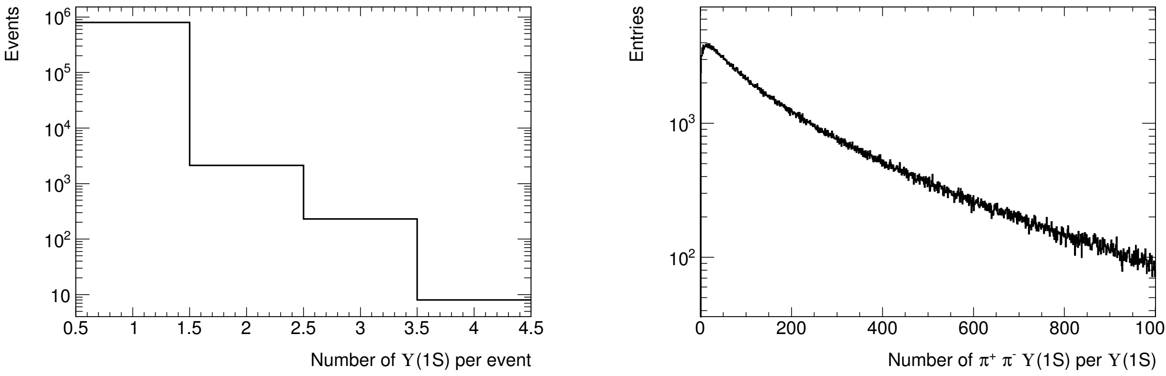
5.3.2 Object Selection Criteria

Of the large number of dipion combinations passing the reconstruction stage in each event, only one[8] (or, more likely, zero) can be from a signal decay – the remainder constitute

---

[7] By a flat addition – in other words, no cross-section or luminosity weighting was performed.

[8] Neglecting the very unlikely occasion that two signal decays occur within the same event.



(a) (b)

Figure 5.7: The number of Υ(1S) candidates reconstructed per event (left), and the number of $\pi^+\pi^-\Upsilon$(1S) candidates per Υ(1S) (right). The large multiplicity of pions typical in each event lead to large numbers of $\pi^+\pi^-\Upsilon$(1S) candidates. These results are for Υ(2S) decays, though are similar for all the MC samples.

the background. The next stage of the analysis involved placing tighter selections on the pions, muons and their combined objects to ensure they were of high quality or in signal-rich regions. These lead to a reduction in the combinatorial background of approximately an order of magnitude, with little decrease in signal efficiency.

5.3.2.1 Muons

Events were selected by demanding that either the *EF_2mu4_Upsimumu* or *EF_2mu4T_Upsimumu* trigger was passed. Both demand the presence of two muons with

- $p_T > 4$ GeV,
- $|\eta| < 2.3$ (the physical extent of the muon trigger detectors),
- opposite charges,

and which can be fitted to a common vertex[9] with a combined mass in the range 8–12 GeV. The *EF_2mu4_Upsimumu* trigger was operational up to Period J of data taking in 2011 (i.e. up to August, 2011), and was seeded by the *L1_MU0* level 1 trigger. The *EF_2mu4T_Upsimumu* trigger ran for the remainder of the year, differing only in the fact it was seeded by the *L1_2MU4* trigger. The latter is more demanding and so slightly less efficient, but was needed to control the L1 rate with the increasing instantaneous luminosity in the later runs of 2011.

---

[9] This is ensured by a very loose selection on the vertex fit χ^2.

Though triggering is unimportant for the MC samples, it is essential in identifying potential signal-containing events in collision data. To ensure the (offline) muons used in reconstructing Υ(1S) candidates corresponded to those being triggered upon, they were required to be matched to the trigger muon objects with a separation, $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$, of less than 0.01.

For consistency, the same acceptance conditions on p_T and η used in the trigger were also enforced on the offline muons.

5.3.2.2 Pions

The only additional selection made to the pions (on top of those in the reconstruction) was an increase in the minimum number of hits in the SCT detectors[10] to 6. Other track variables were considered, such as minimum numbers of TRT hits, but produced no signal to background ratio increase, and so were not pursued.

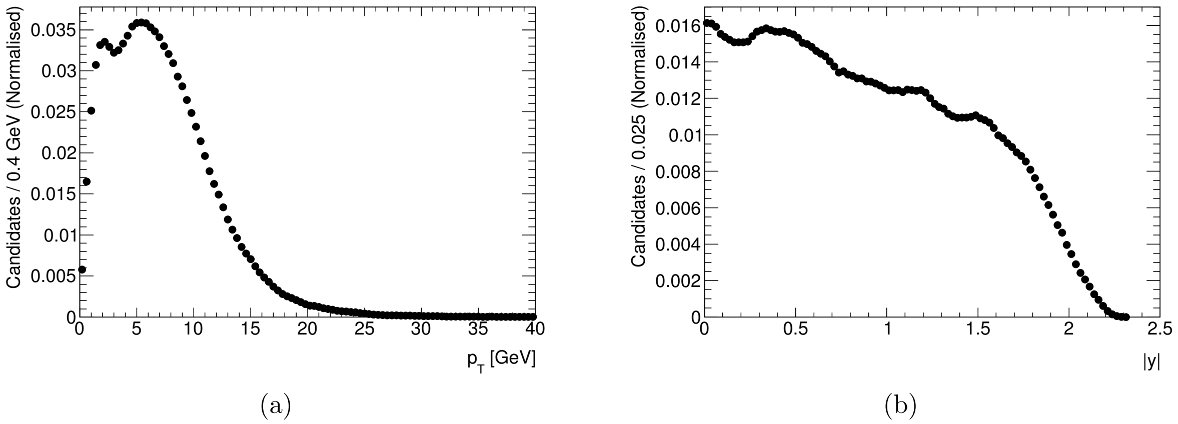
5.3.2.3 Υ(1S) and $\pi^+\pi^-\Upsilon$(1S) Candidates

Until this point, the mass of the Υ(1S) candidates was only required to be in the range $8.5 < m_{\mu^+\mu^-} < 10.7$ GeV. This is very loose (see Figure 5.1) and needed to be tightened to the region immediately surrounding the Υ(1S) peak. In choosing the appropriate mass window, there were two competing desirable outcomes in mind – (1) being as inclusive as possible to maintain signal efficiency, and (2) rejecting as much background as possible. Erring on the side of the former, and relying on the other selections for the latter, $m_{1\text{S}} \pm 350$ MeV was chosen as a conservative mass restriction. Treating the signal Υ(1S) mass peak as approximately Gaussian, this corresponds to roughly ± 1.75 standard deviations.

In an earlier discussion (Section 5.2), it was found that the reconstruction requirement of $p_T^{\pi} > 400$ MeV caused the signal decay acceptance to be appreciably large only when produced in a highly boosted state (see Figure 5.4). In contrast, the background candidates in the inclusive 1S MC sample have a p_T distribution which is peaked below 5 GeV. This is shown for background combinations passing all of the object quality selections discussed so far (and the χ^2 restrictions shown below) in Figure 5.8a. This variable therefore provides an opportunity to separate signal from background – the signal will be mostly confined to the high p_T region due to the acceptance, whilst the background naturally populates the low end of the spectrum. Keeping in mind that

---

[10] To be specific, *hits* here refers to the number of hits in addition to the number of dead sensors traversed.



(a) (b)

Figure 5.8: The transverse momentum and rapidity distributions for reconstructed background candidates from the inclusive 1S MC sample.

- the acceptance increases with the mass of the parent state, causing the expected reconstructed spectrum to have a considerable contribution from the low p_T region (Figure 5.4, and
- the analysis should be sensitive to a range of mass states,

a conservative requirement of $p_T > 5$ GeV was placed. Based on the inclusive 1S MC, this removes close to 40% of the background candidates.

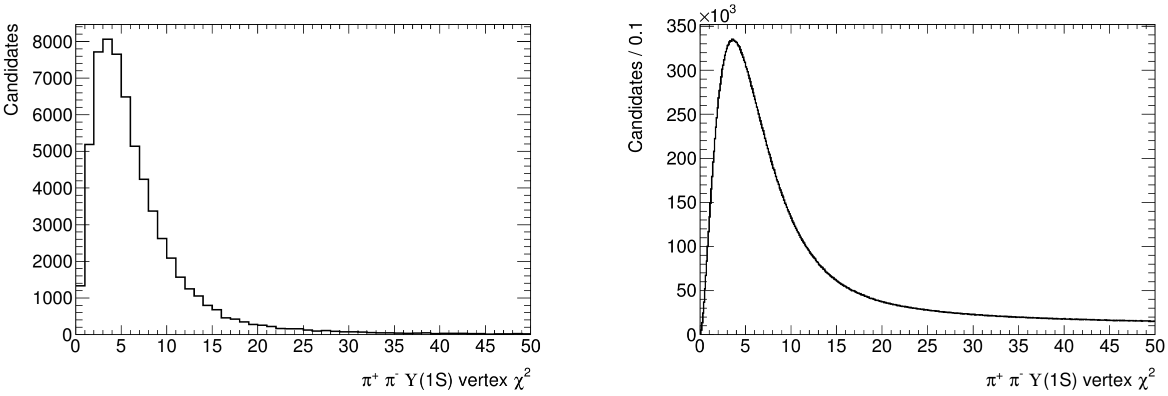
Equivalent considerations (see Figure 5.8b) for the rapidity reveal no signal-to-background advantage in making any restrictions, and so the entire range $|y| < 2.4$ was used (beyond this, the acceptance is essentially zero).

The last variable used in selection of Υ(1S) and $\pi^+\pi^-\Upsilon$(1S) candidates was the χ^2 of their respective vertex fits. The majority of the Υ(1S) background is expected to come from inclusive production, for which the χ^2 of the vertex fit to the two muons will ineffective in removing. As such, only a very loose restriction of $\chi^2 < 200$ was used to exclude very poor candidates.

On the other hand, it is reasonable to assume that the χ^2 of the $\pi^+\pi^+\mu^+\mu^-$ vertex may be able to discriminate signal and background combinations. With all the constraints taken into account, the $\pi^+\pi^-\Upsilon$(1S) vertex fits have six degrees of freedom. The measured χ^2 for signal vertices should therefore follow a χ^2_6 distribution. The actual distribution for the Υ(2S) MC sample[11] is shown in Figure 5.9a, and for the inclusive 1S background sample in Figure 5.9b. The major difference here is that the background has a much longer, slowly falling, tail, whilst the signal distribution is effectively zero for $\chi^2 > 30$.

---

[11] The other MC samples are very similar.



(a) Signal Υ(2S) vertices

(b) Inclusive 1S background vertices

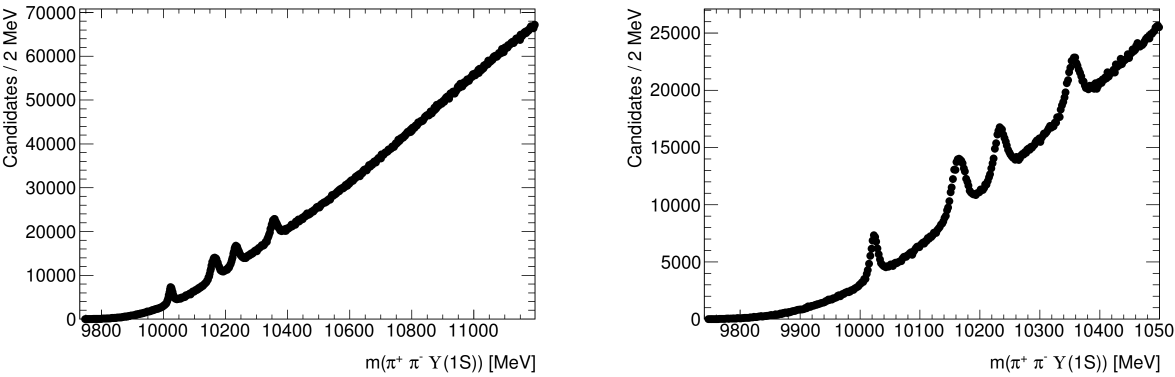
Figure 5.9: The χ^2 distribution for signal vertices in the Υ(2S) MC sample, and background candidates in the inclusive 1S MC sample.

The aim here was to exploit this difference by setting a maximum value for the χ^2 to reject as much background as possible, while retaining a high efficiency for signal decays. After carrying out an optimisation procedure, a cut of $\chi^2 < 20$ was chosen. The details of this process involve the next stage of the analysis, *Candidate Selection*, and will be explained there.

The combined mass distribution for all signal samples after these additional requirements are placed is shown in Figure 5.10. In comparison to the earlier result, Figure 5.6, the peaks for the individual MC samples are now easily visible (though the Υ(1D) triplet appears as one), and the background level has been dramatically reduced. Qualitatively, it also shows that the yield in each of the peaks is a function of their position on the mass spectrum; the efficiency is greater for the higher mass samples. This is unsurprising, given the acceptance considerations discussed earlier.

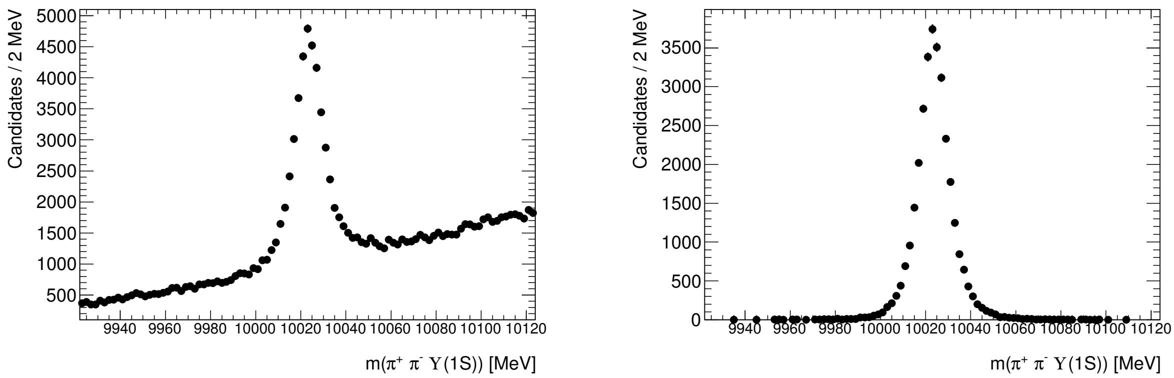
A close-up view of the Υ(2S) sample in the signal region is given in Figure 5.11a, adjacent to the corresponding distribution for signal only. The shape of the signal appears symmetric and fairly Gaussian, and the background smooth and polynomial (perhaps even linear in this region). These observations are equally valid for the other MC samples, with some slight differences in shape parameters (in particular, the width of the signal peak). In Section 5.4, these considerations will be used to develop a generic fitting model valid for any local mass region.

The multiplicities for Υ(1S) and $\pi^+\pi^-\Upsilon$(1S) candidates are shown for the Υ(2S) sample in Figure 5.12. The former now only includes cases where the Υ(1S) candidate was linked to at least one valid $\pi^+\pi^-\Upsilon$(1S) candidate, so that zero is a possibility. The latter is restricted to those Υ(1S) for which the signal decay was a member of its $\pi^+\pi^-\Upsilon$(1S)



(a) (b)

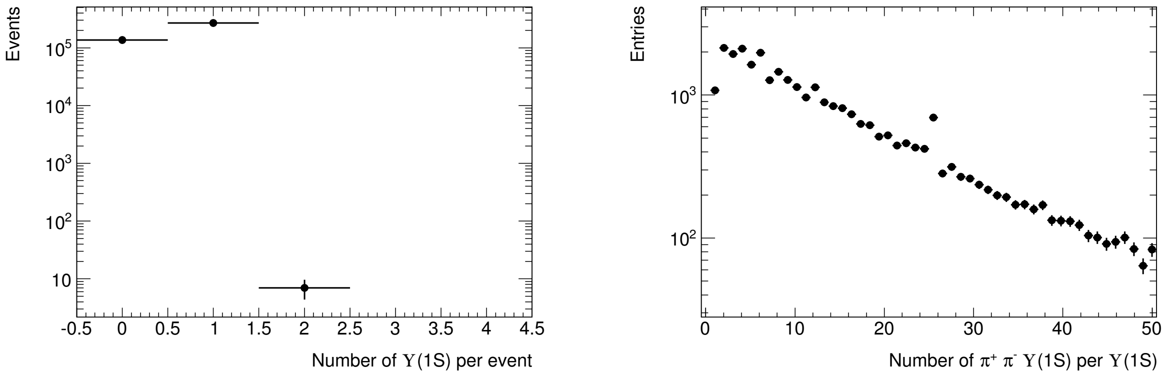
Figure 5.10: The $\mu^+\mu^-\pi^+\pi^-$ mass spectrum for reconstructed candidates after further selections. A close up of the lower mass region is given on the right.



(a) (b)

Figure 5.11: The total mass spectrum for the Υ(2S) sample after reconstruction and selections (left), and the signal-only component (right).

candidates, as this is the most relevant case. In contrast to the equivalent plots before the object quality selections, Figure 5.7, there are now a considerable number of events with zero Υ(1S) candidates, and the multiplicity of $\pi^+\pi^-\Upsilon$(1S) candidates has been greatly reduced; the mean number of candidates per Υ(1S) has dropped from 220 to 16 by using object quality selections. However, Figure 5.12b also shows that the majority of signal-containing events still contain large numbers of background candidates. *Candidate Selection* is a technique designed to handle with this issue.



(a) (b)

Figure 5.12: The number of reconstructed and selected Υ(1S) candidates per event (left) and $\pi^+\pi^-\Upsilon$(1S) candidates per Υ(1S) (right). These results are for the Υ(2S) sample (the others are much the same).

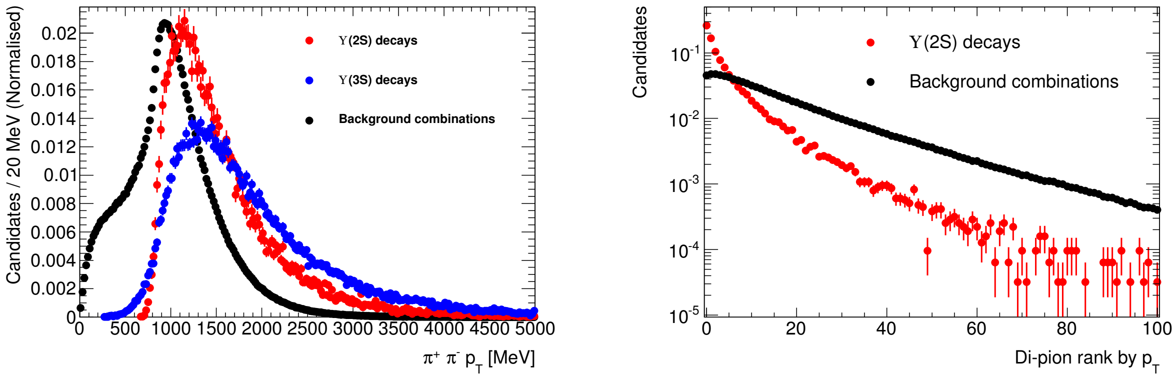
5.3.3 Candidate Selection

To combat the remaining background, a procedure was developed whereby, for all possible $\pi^+\pi^-\Upsilon$(1S) candidates reconstructed for a particular Υ(1S), an attempt was made to pick out only the signal decay. For obvious reasons, this is referred to as *candidate selection*. Any such procedure will have an associated failure rate, leading to an inefficiency in the analysis, but will also bring about a very large reduction in the background level. In events which contain only background candidates, this approach is especially useful.

In essence, any justification for choosing particular candidates can be used, so long as they do not sculpt artificial signal peaks[12]. However, since candidate selection is designed to distinguish between various $\pi^+\pi^-$ combinations, it was natural to concentrate on variables associated to these objects. Most of the individual pion track quantities have already been utilised in existing selections (p_T, η, hits etc.), narrowing the possibilities to quantities associated to the dipion *pairs*. The background candidates are composed of randomly selected pions, which in general have trajectories which are *not* co-linear. This is in contrast to the topology of signal decays, where the pions both recoil against the Υ(1S). This suggests the combined dipion p_T as a discriminating variable, though several other related quantities (e.g. the opening angle between the two pions, or the magnitude of their combined momentum vector) would be similarly effective. The distributions for this variable in the Υ(2S), Υ(3S) and inclusive 1S background samples are shown in Figure 5.13a. Background candidates tend to populate the lower p_T region and peak at

---

[12] Choosing the candidate with the mass closest to 10400 MeV each time would be inappropriate, for example.



(a) (b)

Figure 5.13: The dipion p_T distribution for the Υ(2S), Υ(3S) and inclusive 1S background samples. The rank of signal and background candidates, when ordered by this variable, are given for the Υ(2S) sample as an example. Background here refers only to that in direct competition for best candidate with a signal decay (see text).

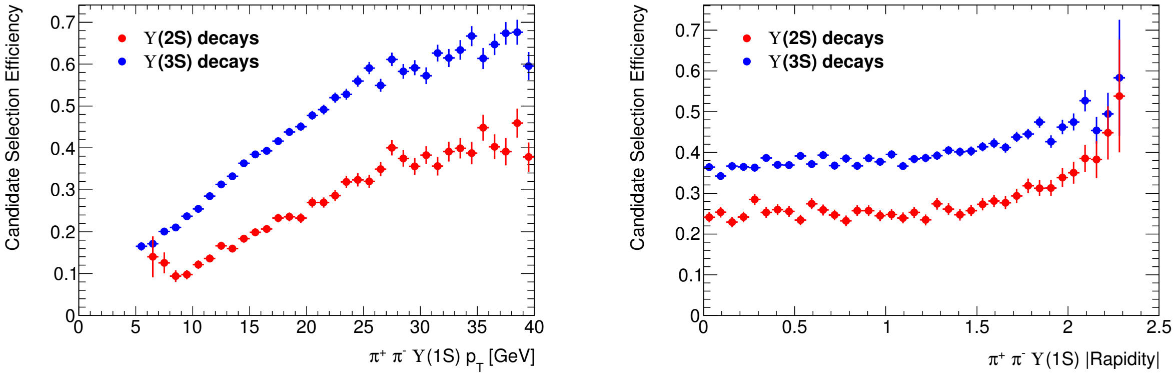
around 1 GeV, whereas this appears to be the threshold for signal combinations. The signal distributions also have longer tails, particularly for the higher-mass Υ(3S) sample.

Proceeding along these lines, each $\pi^+\pi^-\Upsilon$(1S) combination for a particular Υ(1S) candidate were ranked according to their dipion p_T. The distribution for the ranks of signal and background candidates is shown for the Υ(2S) sample in Figure 5.13b. In this instance, *background* refers to those combinations within the Υ(2S) sample in direct competition with a reconstructed signal decay – in other words, in cases where it is possible for the signal decay to be selected as the best candidate. In the candidate selection approach, only the top ranked candidate is retained; all others are rejected. The fraction of Υ(2S) signal decays occupying the top rank is over a quarter, and that for the Υ(3S) sample is close to 40% (the other MC samples have intermediate values). Keeping in mind the typically high number of candidates to choose from (see Figure 5.12b), these success rates are satisfactory.

The candidate selection efficiency is defined as the probability that the signal decay will be selected as the best candidate, given it has been reconstructed and passed all the analysis selections. It is shown as a function of total (that is, $\pi^+\pi^-\Upsilon(1S)$) transverse momentum and rapidity in Figure 5.14, again for the Υ(2S) and Υ(3S) samples. There is no significant development in efficiency with rapidity, but there is a steady increase with p_T.[13] This is expected – if the parent is either highly boosted, the pions tend to be more collimated and energetic. The higher efficiency for the Υ(3S) is also unsurprising,

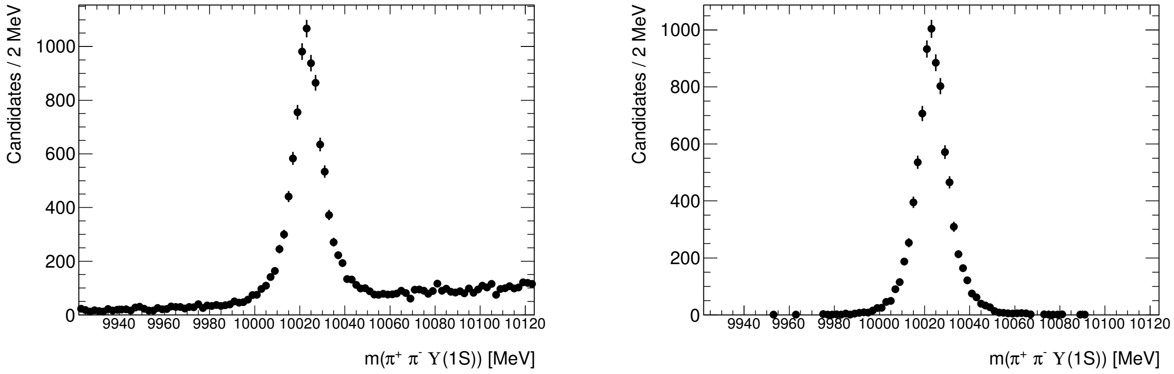
---

[13] The dependence of candidate selection efficiency on dipion mass was also checked, and a weak linear relationship (decreasing at higher values of $m_{\pi^+\pi^-}$) was found.



(a) (b)

Figure 5.14: The efficiency for candidate selection as a function of the $\pi^+\pi^-\Upsilon(1S)$ transverse momentum and rapidity. Samples with masses between that of the $\Upsilon(2S)$ and $\Upsilon(3S)$ have efficiencies intermediate to these.



(a) (b)

Figure 5.15: The invariant mass distribution for the $\Upsilon(2S)$ sample after best-candidate selection. The full distribution is given on the left and the signal component on the right.

given that the energy of the pions is strongly correlated to the mass difference between the $\Upsilon(1S)$ and the decaying parent state.

Figure 5.15a shows the invariant mass distribution that results for the $\Upsilon(2S)$ sample after candidate selection. In comparison to previous distributions (Figure 5.11), the background level is now much lower and has taken on a flatter shape. The signal shape, apart from being reduced in overall normalisation, is much the same and appears to be free from any distortions or biases that may have been introduced.

Given the signal-to-background advantages of candidate selection, it was reasonable to assume that it would allow for a more sensitive search. This is in fact the case, as is demonstrated in Section 5.5, and so candidate selection was adopted as the primary analysis approach. The version of the analysis without any such selection is also retained, and here-forth will be referred to as the *all candidates* method. In a sense, these two are extremes – they correspond to either (1) keeping every $\pi^+\pi^-\Upsilon(1\text{S})$ candidate or (2) picking only one per $\Upsilon(1\text{S})$. A third option, *podium selection*, was also considered whereby the top three ranking $\pi^+\pi^-\Upsilon(1\text{S})$ combinations were used. This was motivated by Figure 5.13b (and those of the other MC samples) as a compromise between efficiency and background suppression. In terms of significance, no advantage was found in using podium selection, and so it was not pursued any further.

With the candidate selection method established, the optimisation procedure used to determine the χ^2 requirement on the $\pi^+\pi^-\Upsilon(1\text{S})$ vertex can now be explained. The goal was to maximise the number of signal candidates selected as the best candidate. Placing a selection on the χ^2 has two competing effects:

1. it diminishes the number of signal decays passing the object quality requirements, and
2. it reduces the multiplicity of background $\pi^+\pi^-\Upsilon(1\text{S})$ candidates, making those signal candidates remaining more likely to be chosen as the best candidate.

Figure 5.16 shows how the total candidate selection efficiency changes with the maximum allowed $\pi^+\pi^-\Upsilon(1\text{S})$ vertex χ^2. The vertical axis is the total fraction of signal decays which would go on to be selected as the best candidate, and is the quantity to be maximised. The efficiency plateaus for both the $\Upsilon(2\text{S})$ and $\Upsilon(3\text{S})$ samples between 10 and 20, then very gradually decreases beyond this. The maximum χ^2 could therefore be chosen to be essentially loss-less for the signal anywhere on this plateau. To remove as much background as possible, it was desirable to choose the strictest such value. Avoiding choices close to the edge of the plateau (a point which seems weakly dependent on mass and may be different in real data), $\chi^2 < 20$ was chosen as the most appropriate requirement.

A summary of the efficiencies for reconstruction, object quality and candidate selection are given for the signal and background samples in Table 5.2. In this table, the various efficiencies for the signal MC samples each refer to the fraction of events which contain a signal decay passing the corresponding analysis stage. For the inclusive 1S background sample, it includes all events in which there is *at least one* background candidate passing the analysis stage. Note, however, that for each $\Upsilon(1\text{S})$ reconstructed in this 3.34% of inclusive 1S events, an average of 14.6 background $\pi^+\pi^-\Upsilon(1\text{S})$ candidates pass the object quality selections. This again hits at the core motivation for using candidate selection – this average of nearly 15 is reduced to just *one*.

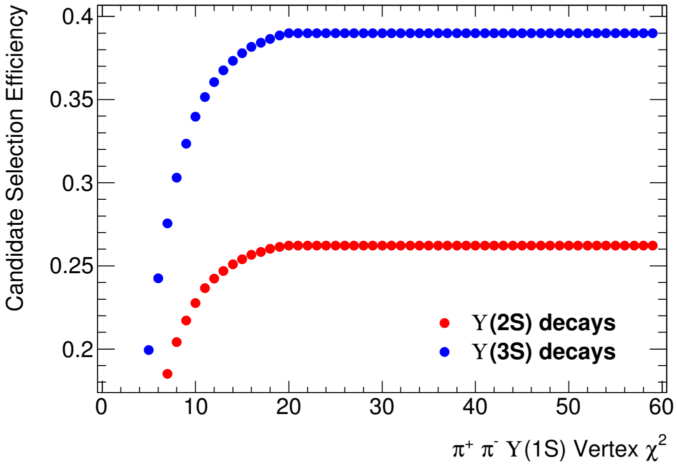


Figure 5.16: The fraction of reconstructed signal decays passing the object quality criteria which would be selected as the best candidate when enforcing various $\pi^+\pi^-\Upsilon(1\mathrm{S})$ vertex maximum χ^2 requirements.

Table 5.2: Summary of the efficiencies for reconstruction, object quality requirements and candidate selection for the Monte Carlo (MC) samples used in this analysis.

Parent State	% Reconstructed	% Passing Object Selection Criteria	% Passing Candidate Selection
$\Upsilon(2\mathrm{S})$	6.54	3.15	0.826
$\Upsilon(1^3\mathrm{D}_1)$	9.86	4.62	1.44
$\Upsilon(1^3\mathrm{D}_2)$	9.95	4.67	1.49
$\Upsilon(1^3\mathrm{D}_3)$	10.2	4.78	1.52
$X(10233)$	12.1	5.59	2.08
$\Upsilon(3\mathrm{S})$	15.0	6.65	2.59
Inclusive 1S Background	24.8	3.34	

5.3.4 Low Mass Analysis

It did not go without notice that some selections were chosen to be conservative, in order to cater for the wide mass range of possible new states. For the purposes of studying just the lower-mass states – i.e. the $\Upsilon(2\mathrm{S})$, $\Upsilon(1\mathrm{D})$, $\Upsilon(3\mathrm{S})$ and any in-between – the following adjustments can be made to the analysis:

- the acceptance-driven minimum on the total transverse momentum of each candidate was increased from 5 to 10 GeV, and
- a constraint of $m_{\pi^+\pi^-} < 1000$ MeV was added.

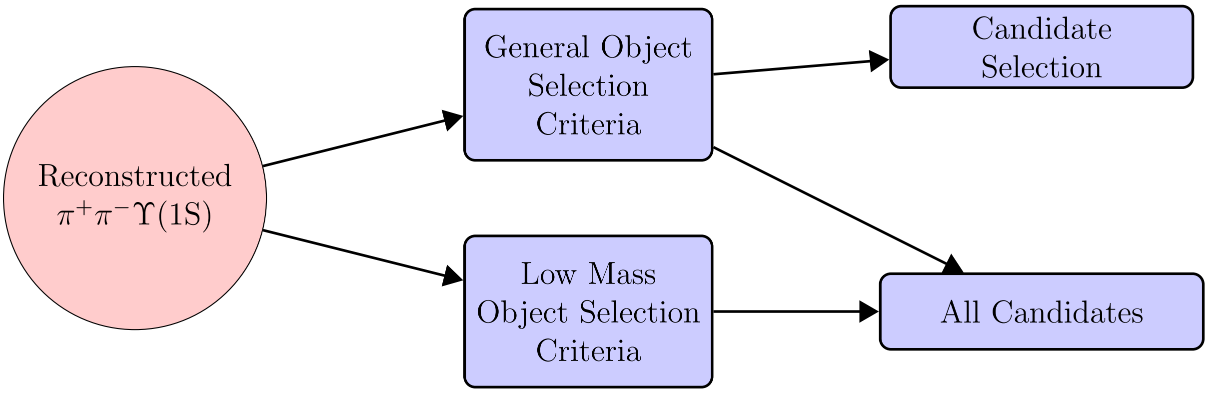


Figure 5.17: A summary of the different analysis approaches in terms of the requirements placed on the reconstructed objects to ensure their quality (middle), and the selection scheme (right).

The latter of these two comes purely from kinematic considerations. In the centre-of-mass frame of the parent,

$$m = \sqrt{m_{1S}^2 + p^2} + \sqrt{m_{\pi^+\pi^-}^2 + p^2}, \tag{5.1}$$

and noting that $p^2 > 0$ gives $m_{\pi^+\pi^-} < m - m_{1S}$. The choice of 1000 MeV covers the range up to 10460.30 MeV, providing plenty of lee-way above the Υ(3S) mass. The set of selection criteria, with these amendments, will be referred to as the *low mass requirements*.

Counter-intuitively, the low mass analysis is actually disadvantageous if candidate selection is being used. To see this, consider the scenario where all of the $\pi^+\pi^-\Upsilon$(1S) candidates in a particular event are background. Forcing $m_{\pi^+\pi^-}$ to be less than 1 GeV generally ensures that the best candidate will be within the low mass region, whereas without this constraint the best-candidate may lie higher up in the mass spectrum. In other words, candidate selection, coupled with the dipion mass constraint, essentially *forces* background candidates into the signal region.

Owing to this, the low mass requirements were reserved only for the situation where a strong signal has been observed in the bottom half of the mass spectrum, and a higher signal purity is necessary than is provided by the standard selections.

The entire process explained in this section – the reconstruction, object quality requirements and selection of $\pi^+\pi^-\Upsilon$(1S) candidates, is schematically represented in Figure 5.17.

5.4 The Fitting Model

The last step in this search will involve

- determining the statistical significance of potential new states, and
- setting upper limits for the product of cross-section and branching fraction to $\pi^+\pi^-\Upsilon(1\mathrm{S})$, of such states.

To perform these calculations, a fitting model capable of simply and successfully describing the (local) shape of the $\pi^+\pi^-\Upsilon(1\mathrm{S})$ invariant mass spectrum was developed. All such fits are binned, and use the maximum likelihood method.

Consider the mass spectrum in some region, $[\overline{m} - L/2, \overline{m} + L/2]$, centred at $\overline{m}$, where L is of the order of 100-200 MeV. From Figure 5.10, it is clear that the background within such a local region is smooth, and can be well-approximated by a polynomial. For example, in the background shape immediately surrounding the $\Upsilon(2\mathrm{S})$ mass (Figure 5.11a) is quite linear. The background term, then, was chosen as a 2^{nd}-order Chebychev polynomial, whose coefficients are floated and dependent on $\overline{m}$.

The signal shape is more complicated; though it is symmetric, it is *not* well-described by a single Gaussian. One reason for this is that the ATLAS tracking resolution is a strong function of $q\eta$, where q is the particle charge. In particular, the measurement of track parameters is better in the barrel region of the detector than in the endcaps. This difference feeds into the $\mu^+\mu^-\pi^+\pi^-$ mass measurement through the vertex fit, which takes into account the covariance matrix for each track. For reasonably boosted decays, the four tracks are well-collimated and their direction can be approximated by that of the total four-vector. It is therefore reasonable to expect that signal decays with higher rapidity will have poorer resolution. This is demonstrated in Figure 5.18, where the $\Upsilon(2\mathrm{S}) \to \pi^+\pi^-\Upsilon(1\mathrm{S})$ mass shape is shown for signal decays with $|y| < 1.2$ and $1.2 < |y| < 2.4$, roughly corresponding to the barrel and endcap regions of the detector. Because of this large difference in resolution, the analysis was performed simultaneously in these two rapidity bins. But, even when the signal mass distributions are split in this way, they are still not Gaussian. Reasons for this include

- the residual rapidity-dependence of the resolution within the barrel/endcap,
- the fact that the decaying tracks are not perfectly collimated – i.e. there will be mixed cases where some tracks end up in the barrel and the others in the endcaps,
- tracks which are broken in some way (e.g. have missing hits, or are associated to incorrect hits),
- broken $\pi^+\pi^-\Upsilon(1\mathrm{S})$ candidates (at least one of the tracks is incorrectly assigned), and

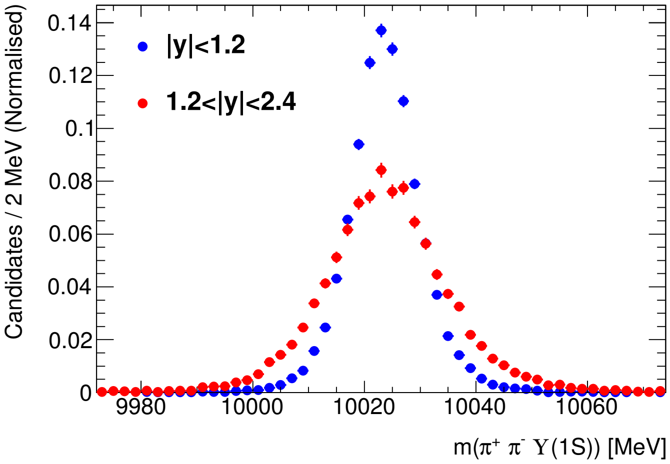


Figure 5.18: The invariant mass distribution for signal decays in the $\Upsilon(2\text{S})$ sample for the two rapidity regions – $|y| < 1.2$ and $|y| < 2.4$. The central region clearly has significantly better resolution than at higher rapidities.

- effects of the detector material (e.g. multiple scattering and energy loss from bremsstrahlung).

It would be ideal to model these effects using a large number of Gaussians, each with a common mean (the mass of the state) but different widths. Unfortunately, this is impractical (this will be apparent later in this chapter, and in those that follow) and in an effort to keep the fitting model simple only two Gaussians were used – these will be called the *narrow* and *broad* components of the signal.

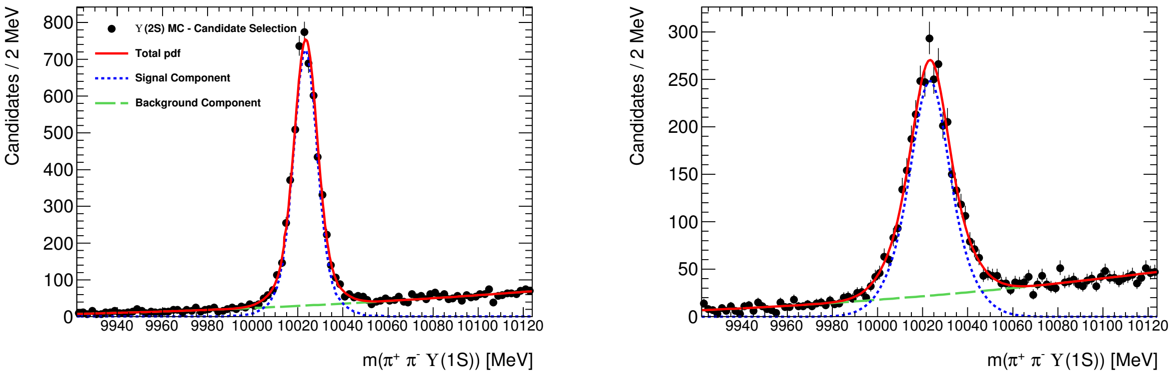
The general fitting model p.d.f. is an addition of these Gaussians and the background polynomial, each with independent normalisations. It is given by:

$$f(m) = N_s \cdot \left(f_n G_n\left(\overline{m}, \sigma_n\right) + \left(1 - f_n\right) G_w\left(\overline{m}, \sigma_w\right)\right) + N_b \cdot P_2\left(c_0, c_1\right), \tag{5.2}$$

where

- N_s and N_b are the signal and background yields, respectively,
- f_n is the *narrow fraction* of the signal curve,
- $\overline{m}$ is the fitted mass of the peak,
- σ_n and σ_w are the widths for the narrow and broad components, and
- c_0 and c_1 are the coefficients of the 2^{nd}-order Chebychev polynomial, P_2.

To demonstrate the efficacy of this model, fits to the barrel and endcap (that is, $|y| < 1.2$ and $1.2 < |y| < 2.4$) bins of the candidate selection mass spectrum in the $\Upsilon(2\text{S})$ MC



(a) Barrel (b) Endcap

Figure 5.19: The fitted mass spectrum in the $\Upsilon(2S)$ MC sample near $m = 10023$ MeV, for the barrel (left) and endcap (right) bins.

sample, centred at $\overline{m} = 10023$ MeV, are shown in Figure 5.19. In these figures (and others containing fits), the total p.d.f. is represented by the solid red line, the background component by the green dashed line, and the signal by the dashed blue line(s).

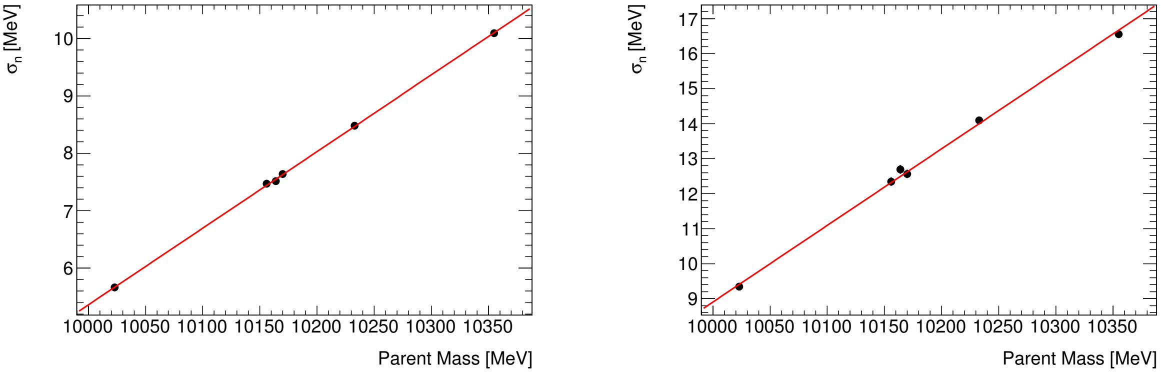
5.4.1 The Signal Shape Parameters

In the fit p.d.f, Equation 7.11, the σ_n and σ_w parameters are both expected to scale roughly with the mass of the parent state. On the other hand, the *ratio*, r, between these two is a result of the processes described above which are generally not a function of mass. In fact, is was found that the values of this ratio in signal-only fits to each of the MC samples were all consistent with one another. Similar comments are applicable to the narrow fraction parameter, f_n. To simplify the signal fitting shape (which is essential in the process of setting upper limits or producing significance scans), r and f_n were hereafter fixed to representative values as determined from the MC samples. Specifically,

1. each of the signal MC samples was fitted with all parameters floated, and the weighted average value for $r := \sigma_w/\sigma_n$ was extracted, then
2. with r fixed to this value, the weighted average value for f_n was determined.

To minimise the statistical uncertainty, the *all candidates* samples were used; because the candidate selection efficiency is essentially flat in y, no difference in signal shape was expected. However, the obvious differences between the barrel and endcap made it essential to compute these averages separately for each. The values thus obtained are

- **Barrel:** $r = 2.24 \pm 0.02$, $f_n = 0.874 \pm 0.004$



(a) Barrel

(b) Endcap

Figure 5.20: The value of σ_n as a function of the parent mass, for the barrel (left) and endcap (right) bins.

- **Endcap:** $r = 2.14 \pm 0.02$, $f_n = 0.824 \pm 0.007$.

As expected, repeating this process with the candidate selection approach produced consistent values, but with larger uncertainties. Note also that for the majority of fits, fixing these parameters actually resulted in a *drop* in $\chi^2/n.d.o.f$ (the remainders had a meagre increase).

For future reference, it was also necessary to understand how the signal shape developed as a function of parent mass. The only remaining free parameter in the signal shape is the width of the narrow component, σ_n. The fitted value for σ_n is shown as a function of mass in Figure 5.20, for the two rapidity regions. As suspected, the trend is linear and can therefore be described by just two parameters – the *y-intercept*, α, and the slope, β.

A summary of the signal shape parameters is given in Table 5.3. Though each of these were extracted from MC samples, they are all independent of the *absolute value* of the parameters. That is, they are all ratios in some sense, and so are expected to be reasonably safe against data/MC differences. This will be re-assessed in the following chapter using the Υ(2S) peak in 2011 data.

Using these parameters, the signal shape for *any* parent mass can be predicted. This is a necessary input into most subsequent statistical computations, such as those described in the following section.

5.4.2 Fitted Mean Bias

In the scenario where a new particle is observed, the most important property to extract initially will be its mass. It is therefore prudent to assess how well the analysis is able

Table 5.3: Summary of the signal shape parameters extracted from the MC samples.

	r	f	α	β
Barrel	2.24±0.02	0.874±0.004	-128±2	0.0134±0.0002
Endcap	2.14±0.02	0.824±0.007	-210±4	0.0219±0.0003

Table 5.4: Generated vs. fitted masses for the signal MC samples.

Generated Mass	Fitted Mass (Barrel) [MeV]	Difference (Barrel) [MeV]	Fitted Mass (Endcap) [MeV]	Difference (Endcap) [MeV]
10023	10023.68±0.04	0.68±0.04	10023.8±0.1	0.8±0.1
10156	10156.78±0.07	0.78±0.07	10157.1±0.2	1.1±0.1
10164	10164.83±0.07	0.83±0.07	10165.0±0.2	1.0±0.2
10170	10170.79±0.07	0.79±0.07	10171.0±0.2	1.0±0.2
10233	10233.82±0.05	0.82±0.05	10234.0±0.1	1.0±0.1
10355	10355.94±0.05	0.94±0.05	10356.1±0.1	1.1±0.1

to recover the mass of a particle decaying into the $\pi^+\pi^-\Upsilon(1S)$ final state. A comparison of the input (that is, generated) MC masses vs. the fitted masses for the signal-only distributions are given for the barrel and endcap bins in Table 5.4. Generally, the fits are successful in recovering the generated mass, though there is evidence for a mildly mass-dependent systematic offset of the order of +1 MeV. The difference also seems to be slightly higher in the $1.2 < |y| < 2.4$ bin. These results can be used to perform an inversion map of this difference for any observed state where the fitted mass is of importance. For example, in the case of the $\Upsilon(2S) \to \pi^+\pi^-\Upsilon(1S)$ and $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)$ decays, the inversion will simply be to subtract 0.68 MeV from the fitted mass. For others, a linear interpolation/extrapolation or linear fit to the differences as a function of mass can be used. This will be revisited when needed.

5.5 Analysis Sensitivity Estimates

With the fitting model established, the expected sensitivities of the *all candidates* and *candidate selection* approaches for the 2011 dataset could be assessed and compared. Of particular interest are

1. the expected significance for a signal at mass m with a given production rate σ and branching fraction to $\pi^+\pi^-\Upsilon(1S)$, $\mathcal{B}$, and
2. the corresponding upper limit that could be set on this rate in the case nothing significant is observed.

Because of the relatively large background, both of these could be computed by utilising the *asymptotic formulae* for test statistics, as explained in [104].

Rather than the absolute production rate, $\sigma\mathcal{B}$, the relative rate with respect to the Υ(2S) is used:

$$\begin{aligned} R &:= \frac{\sigma\mathcal{B}}{(\sigma\mathcal{B})_{2\mathrm{S}}} \\ &= \frac{N}{\mathcal{L}_{\mathrm{int}}\varepsilon}\frac{\mathcal{L}_{\mathrm{int}}\varepsilon_{2\mathrm{S}}}{N_{2\mathrm{S}}} \\ &= \frac{N}{N_{2\mathrm{S}}}\frac{\varepsilon_{2\mathrm{S}}}{\varepsilon}, \end{aligned} \tag{5.3}$$

where N ($N_{2\mathrm{S}}$) is the expected number of observed decays for the new particle (Υ(2S)), and ε ($\varepsilon_{2\mathrm{S}}$) is the total analysis efficiency. This variable is preferable mainly because the systematic uncertainty of the integrated luminosity is completely removed. The analogous ratio $(\sigma\mathcal{B})_{X(3872)}\,/\,(\sigma\mathcal{B})_{\psi(2\mathrm{S})}$ for the $X(3872)$ in the $c\bar{c}$ system has also been measured by CMS as $R = 6.56\%$ [4], which provides a good baseline with which to compare the results of this analysis.

The total analysis efficiency, ε, can be factorised in a number of ways. The most convenient for these purposes is

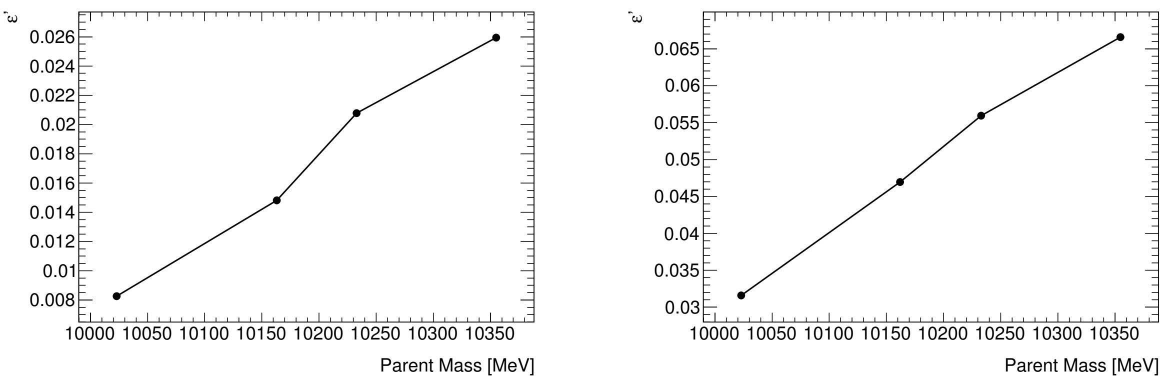
$$\varepsilon = \mathcal{A}_{\mu\mu}\varepsilon', \tag{5.4}$$

where $\mathcal{A}_{\mu\mu}$ is the muon acceptance and ε' is the remaining (conditional) efficiency. The former of these is defined as the fraction of decays for which both muons satisfy the kinematic acceptance conditions – $p_T > 4$ GeV and $|\eta| < 2.5$[14]. The value of this acceptance is determined by the p_T and rapidity spectra of the Υ(1S), and is expected to have a very minor dependence on the parent mass. For example, under the conditions of pure phase space and the Υ(2S) production kinematics determined by NRQCD, $\mathcal{A}_{\mu\mu}(m = 10023\text{ MeV}) = 17.2\%$ and $\mathcal{A}_{\mu\mu}(m = 11000\text{ MeV}) = 16.5\%$ (see Section 5.2). In reality, production at 11 GeV is likely to be marginally harder[15], mitigating this small loss of acceptance. Given this, the ratio of muon acceptance for the Υ(2S) and a state at any mass can be taken as one for now. In the final analysis on the 2012 dataset, a different factorisation was chosen for the total efficiency, and the extrapolation of the acceptance became more involved (see Section 7.6.6.2).

---

[14] In Section 5.2, a slightly different value was used for the pseudo-rapidity restriction. Here, 2.5 was chosen to match the muon filter used in producing the MC samples.

[15] The ATLAS measurement of the differential production cross-section against p_T indicate that the $\Upsilon(\mathrm{nS}|n = 1, 2, 3)$ states tend to have a more prominent tail for increasing n [6].



(a) Candidate selection

(b) All candidates

Figure 5.21: The (conditional - see text) efficiency as a function of the mass of the parent state, for candidate selection (left) and all candidates (right).

Using Equation 5.4, the formula for R is reduced to

$$R = \frac{N}{N_{2\mathrm{S}}} \frac{\varepsilon'_{2\mathrm{S}}}{\varepsilon'}. \tag{5.5}$$

To determine the efficiency ratio in Equation 5.5, ε' is first calculated for each of the MC samples (the $\Upsilon(1^3\mathrm{D_J})$ are analysed as one). Because the MC samples were produced with a muon filter, these efficiencies are just those given in Table 5.2. Figure 5.21 shows how the efficiency develops as a function of the parent mass. The efficiency at intermediate masses (or those below/above the first/last point) is estimated by linear interpolation (extrapolation). The uncertainty associated with this will again be an important systematic to include in the final measurement – currently, only reasonable estimates are needed.

5.5.1 Expected Significance

In the asymptotic regime (i.e. large sample sizes), the significance of a potential signal peak over a background can be determined by Wilks' theorem (see, e.g., [104]) as

$$z = \sqrt{q_0} = \sqrt{-2\frac{\mathcal{L}(0, \theta(0))}{\mathcal{L}(\hat{R}, \hat{\theta})}}, \tag{5.6}$$

where θ represent the nuisance parameters (in this case, the background shape parameters). $\hat{R}$ and $\hat{\theta}$ are the maximum likelihood estimates (MLEs), and $\theta(0)$ is the maximum estimator under the condition that $R = 0$.

Now consider the scenario in which a state with mass m has a non-zero value for R. Of interest is the expected median significance, as calculated by Equation 5.6, for a given data sample of integrated luminosity $\mathcal{L}_{\text{int}}$. This section describes how this expected significance was estimated.

The calculation required the following inputs:

- **The expected Υ(2S) yield**, $N_{2\text{S}}$. This can be calculated from $\mathcal{A}_{\mu\mu}$ using the *acceptance tool* of Section 5.2, the value of ε' in the Υ(2S) MC sample, and the measured value of $(\sigma\mathcal{B})_{2\text{S}}$ by ATLAS [6], as

  $$N_{2\text{S}} = \mathcal{L}_{\text{int}} \, (\sigma\mathcal{B})_{2\text{S}} \, \mathcal{A}_{\mu\mu}\varepsilon', \tag{5.7}$$

  where

  - $\mathcal{L}_{\text{int}}$ is the integrated luminosity (4.58 fb^{-1} for the 2011 dataset)
  - $(\sigma\mathcal{B})_{2\text{S}}$ is calculated as $\sigma\,(pp \to \Upsilon(2\text{S}))\,\mathcal{B}\,(\Upsilon(2\text{S}) \to \mu^+\mu^-)\cdot\mathcal{B}\,(\Upsilon(2\text{S}) \to \mu^+\mu^-)^{-1}\cdot \mathcal{B}\,(\Upsilon(2\text{S}) \to \pi^+\pi^-\Upsilon(1\text{S}))$. The product of the first two factors has been measured by ATLAS in the kinematic range $p_T < 50$ GeV and $|y| < 2.25$ [6], and for the others the world average value is used [3].
  - the muon acceptance, $\mathcal{A}_{\mu\mu} = 0.275$, was determined using the acceptance calculator of Section 5.2. To take into account the fact that the cross-section was only determined in a restricted kinematic range, the acceptance was calculated with respect to these conditions.

  Equation 5.7 is still slightly inconsistent in that (1) ε' is not conditional to the p_T and rapidity restrictions, and (2) $N_{2\text{S}}$ will only give the yield within this range, whereas the analysis allows for any p_T and up to $|y| < 2.4$. Fortunately, both of these effects are small and can be ignored for the purposes of this calculation. For the 2011 dataset, Equation 5.7 gives $N_{2\text{S}} = 18290$ for all candidates and 4800 for candidate selection.

- **The expected background.** In this analysis, the background is categorised as *Inclusive 1S production* or *Non-1S*. Currently, there is only a model for the former - the latter will be studied with collision data in next chapter. Since precision is not important here, only the *Inclusive 1S production* is assumed to constitute the background. The number of expected Υ(1S) contributing to this can be estimated in a similar manner to that for the Υ(2S) above. Namely,

  $$N_{1\text{S}} = \mathcal{L}_{\text{int}} \, (\sigma\mathcal{B})_{1\text{S}} \, \varepsilon, \tag{5.8}$$

  where $(\sigma\mathcal{B})_{1\text{S}}$ is again taken from the ATLAS measurement. In this case, the calculation of ε is much simpler – because the background sample is unfiltered, the acceptance term does not need to be independently determined (it will naturally be incorporated in the analysis efficiency from the MC). However, the efficiency does need to be taken with respect to the restrictions $p_T < 70$ GeV and $|y| < 2.25$ for it to be consistent with the measured $(\sigma\mathcal{B})_{1\mathrm{S}}$. This gives $\varepsilon = 0.0525$, leading to a predicted $\Upsilon(1\mathrm{S})$ contribution of 1.9×10^6 (for an integrated luminosity corresponding to the 2011 dataset). Again, there is the caveat that this ignores the small number of candidates with $|y| > 2.25$ or $p_T > 70$ GeV.

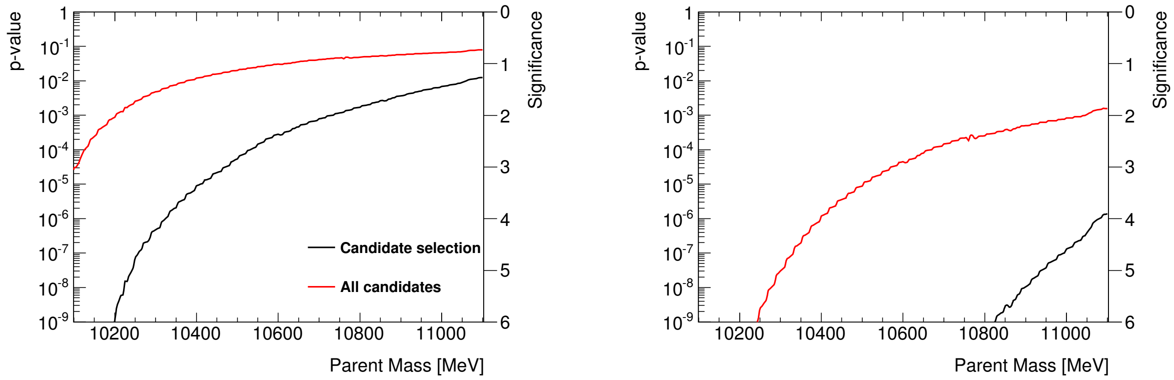
The *shape* of the background in the vicinity of any mass, m, is determined by performing a background-only fit to the mass spectrum in the background MC sample. Using the fraction of the mass spectrum in this region (i.e. the fractional integral of the mass histogram), the expected *amount* of background is found by multiplying by $N_{1\mathrm{S}}$. The shape, scaled by the amount, gives the expected background.

- **Signal Bin Fractions.** The fraction of signal candidates present in each of the two analysis bins, $|y| < 1.2$ and $1.2 < |y| < 2.4$, is denoted F_{yL} and F_{yH} respectively. From the signal MC samples, these fractions were found to be independent of mass, with weighted averages $F_{yL} = 0.66$ for all candidates and $F_{yL} = 0.64$ for candidate selection.

With all the ingredients in place, the procedure for calculating the expected significance, for a fixed value of R, can now be described. Firstly, the following is established:

1. The integrated luminosity, $\mathcal{L}_{\mathrm{int}}$.
2. The mass, m.
3. The signal shapes in the $|y| < 1.2$ and $1.2 < |y| < 2.4$ bins, as prescribed in Section 5.4.
4. Given a value for R, the signal yield in each bin is determined as $N_{2\mathrm{S}} R F_{yi} \varepsilon' \varepsilon_{2\mathrm{S}}^{\prime -1}$, for $i = L$ or H.
5. The expected background in the region $[m - 8\sigma_{yH}, m + 8\sigma_{yH}]$, where σ_{yH} is the width of the narrow signal component in the $1.2 < |y| < 2.4$ bin.

Given this setup, the procedure outlined in [104] is carried out. In short, an *Asimov dataset* is created for the signal-plus-background model, which is then simultaneously fit across the two rapidity bins to extract $q_{0,A}$ (Equation 5.6). According to [104], the median expected significance is then given by $z_{\mathrm{med}} = \sqrt{q_{0,A}}$. This procedure is repeated for the entire mass spectrum, $10.1 - 11.1$ GeV, with a 5 MeV step size.



(a) $\mathcal{L}_{\text{int}} = 4.58$ fb^{-1} (b) $\mathcal{L}_{\text{int}} = 20$ fb^{-1}

Figure 5.22: The expected significance for a state with a relative production rate $R = 0.0656$ for 4.58 fb^{-1} (left) and 20 fb^{-1} (right).

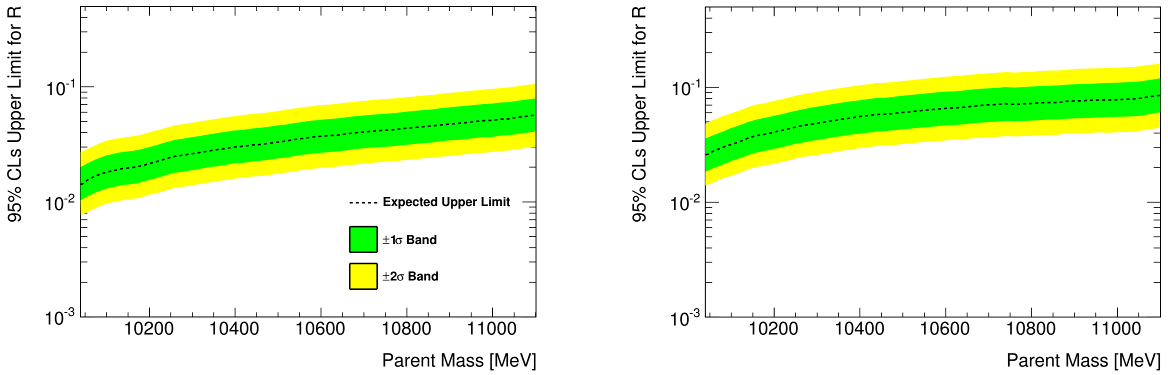
The results for a relative production rate $R = 0.0656$ are shown in Figure 5.22 for $\mathcal{L}_{\text{int}} = 4.58$ fb^{-1}, representing the 2011 dataset, and $\mathcal{L}_{\text{int}} = 20$ fb^{-1}, which roughly corresponds to the 2012 dataset size. This clearly demonstrates the benefit of candidate selection, confirming the extra sensitivity it provides. Based on these results, evidence for an X_b (or other similar state) with such a relative production would be expected up to a mass of 10.4 GeV with the 2011 dataset, but a discovery claim only for $m < 10.2$ GeV. With $\mathcal{L}_{\text{int}} = 20$ fb^{-1}, the discovery potential for this value of R essentially covers the entire mass spectrum. However, this is somewhat misleading - conditions in 2012 data taking were not exactly as they were in 2011 (in particular, the primary trigger would have a p_T threshold of 6 GeV), and so the sensitivity may be different.

The $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)$ has a production rate which is 2.49% of the $\Upsilon(2S) \to \pi^+\pi^-\Upsilon(1S)$ decay (based on [6] and [3]), which, according to Figure 5.22a, indicates there will be little evidence for it in the 2011 dataset.

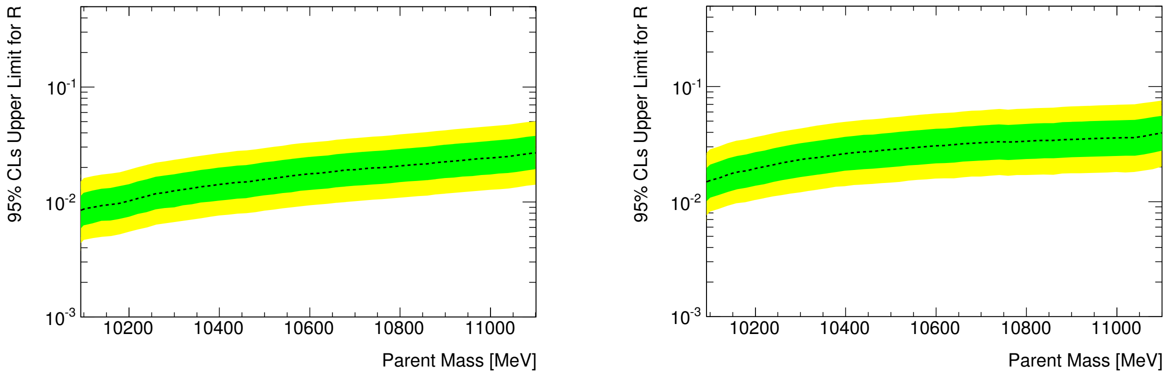
5.5.2 Expected Upper Limits

In the case that no significant signal peak is observed, an upper limit can be set. Using the framework established above, the expected upper limit is extracted using the asymptotic formulas for the test statistic $\tilde{q}_\mu$ (details provided in [104]). For consistency with the current trends, particularly those followed and agreed upon by the ATLAS and CMS collaboration, CLs upper limits are used.

The expected 95% upper limits are given for $\mathcal{L}_{\text{int}} = 4.58$ fb^{-1} and $\mathcal{L}_{\text{int}} = 20$ fb^{-1}, in both the all candidate and candidate selection analysis techniques, in Figure 5.23. Candidate selection provides the strongest upper limits, again confirming it as the favoured



(a) Candidate selection, $\mathcal{L}_{\text{int}} = 4.58\ \text{fb}^{-1}$ (b) All candidates, $\mathcal{L}_{\text{int}} = 4.58\ \text{fb}^{-1}$



(c) Candidate selection, $\mathcal{L}_{\text{int}} = 20\ \text{fb}^{-1}$ (d) All candidates, $\mathcal{L}_{\text{int}} = 20\ \text{fb}^{-1}$

Figure 5.23: The expected 95% CLs upper limits for the candidate selection (left) and all candidates (right) analysis approaches. The top plots are for an integrated luminosity corresponding to the 2011 dataset, whereas the bottom roughly correspond to the 2012 dataset size. The yellow and green bands indicate the $\pm 1\sigma$ and $\pm 2\sigma$ fluctuations around the median expected limit.

approach. According to these results, $R = 0.0656$ can be expected to be excluded for all masses, with the 2011 dataset alone. With the four-times increase to $\mathcal{L}_{\text{int}} = 20\ \text{fb}^{-1}$, the limits improve by a factor of approximately two, as expected.

Chapter 6

Search for $X_b \to \pi^+\pi^-\Upsilon(1S)$ Decays: Analysis of the 2011 Data

In the previous chapter, the analysis approach was established on MC samples. In particular, the sensitivity studies indicated unambiguously that *candidate selection* is the preferable technique to use. Consequently, *all candidates* was reserved only to assist in background modelling and as a cross-check with the MC predictions.

The analysis of the 2011 dataset was carried out in a sequential fashion, beginning with the most well-known parts of the mass spectrum. The threshold region, below the Υ(2S), was initially used to develop a model for the *Non-1S* background component. Following this, the spectrum was studied in the vicinity of the Υ(2S) state to further develop the background modelling and validate the efficiencies derived from MC. The Υ(3S) region was also inspected with the hopes of an observation (though the sensitivity studies of the previous chapter indicated this was not to be expected), then the $\Upsilon(1^3\mathrm{D_J})$ triplet, and finally the entire mass spectrum.

The main emphasis in this chapter was in establishing robust analysis procedures, which could then be directly applied to the 2012 sample. There is therefore no accounting for systematic uncertainties, and only limited discussions of the implication of the results.

6.1 The 2011 Data Sample

During 2011 the LHC delivered ATLAS an estimated 5.61 fb^{-1} of integrated luminosity, of which 5.25 fb^{-1} was recorded. For the first part of the year (until mid-August) the *EF_2mu4_Upsimumu* trigger ran un-prescaled, but was later replaced by the *EF_2mu4T_Upsimumu* trigger to reduce the rate at L1 (see Section 2.2.4.1); this analysis therefore makes use of both of these. To ensure high data quality, events were excluded if any of the detector subsystems crucial for muon and track reconstruction were compromised at the time (this is analogous to similar requirements used in Chapter 4 for electrons). Taking this into account, the total integrated luminosity for these two trigger is 4.58 ± 0.08 fb^{-1}, where the systematic uncertainty of 1.8% is taken from [105].

6.2 Background Modelling

The background can be split into two categories — (1) the *inclusive 1S* $\mu^+\mu^-\pi^+\pi^-$ combinations, where the muons originate from an Υ(1S) decay, and (2) the *non-1S* combinations, where they do not originate from a Υ(1S) decay.

Understanding the $\pi^+\pi^-\Upsilon$(1S) mass background involves two pieces of information:

1. the **total amount**xf of background in each of these categories, and
2. how this is distributed across the $\pi^+\pi^-\Upsilon$(1S) mass spectrum — i.e. the **shape** of the background in each category.

The following sections describes how these features were extracted from the 2011 dataset.

6.2.1 Inclusive and Non-1S $\mu^+\mu^-$ Yields

The $m(\mu^+\mu^-)$ mass distribution of the *Upsilon Candidates* used in $\pi^+\pi^-\Upsilon$(1S) combinations for the 2011 analysis is shown in Figure 6.1. In producing this Figure, the usual restriction of $9.1103 < m(\mu\mu) < 9.8103$ GeV was loosened to the range 8.5-10.7 GeV to allow for a better idea of the overall shape. Within this distribution are each of the two background components, as well as contributions from the $\Upsilon(2S) \to \mu^+\mu^-$ and $\Upsilon(3S) \to \mu^+\mu^-$ decays. The signal decays form an insignificant subset of the inclusive 1S production category, and can be ignored here.

The shape of this distribution is not particularly simple — it is not immediately obvious what the shape or normalisation of each constituent component is. In comparison to that found in the cross section measurement [6] (see Figure 5.1), the distribution here is suppressed for masses above and below $m_{1\text{S}}$. The reason for this lies in the fact that only $\mu^+\mu^-$ combinations which go on to be included in a viable $\pi^+\pi^-\Upsilon$(1S) combination contribute to the distribution. In particular, these $\pi^+\pi^-\Upsilon$(1S) candidates must pass the vertex χ^2 criterion. Given that the $\mu^+\mu^-\pi^+\pi^-$ vertex fit constrains the mass of the muons to $m_{1\text{S}}$, it follows that *Upsilon candidates* whose $m(\mu^+\mu^-)$ was originally far from this will be less likely to satisfy the χ^2 restriction.

Taking a step back, consider the $m(\mu^+\mu^-)$ distribution in the case that no χ^2 maximum is placed. These are shown for the barrel, $|y^{\mu\mu}| < 1.2$, and endcap, $1.2 < |y^{\mu\mu}| < 2.4$,

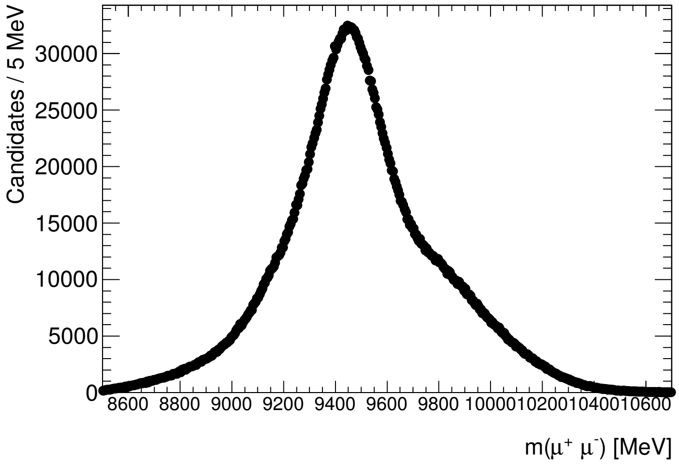
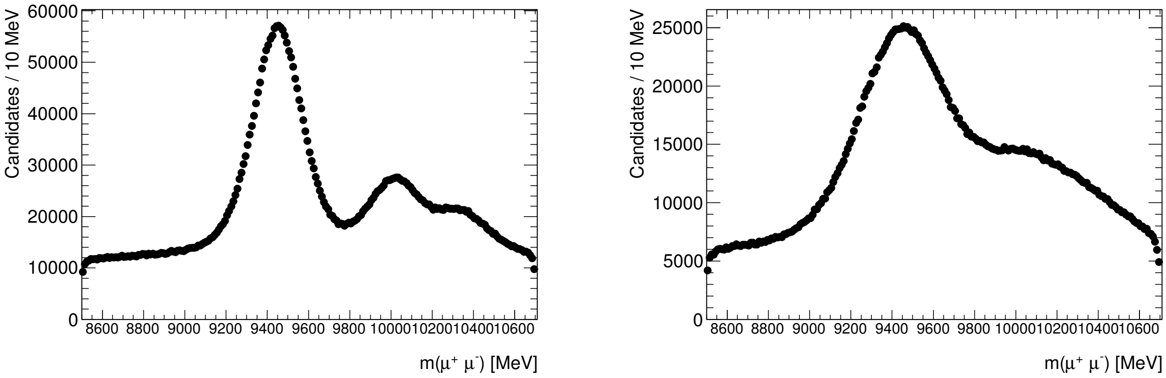


Figure 6.1: The invariant mass distribution of *Upsilon Candidates* in the 2011 dataset, for the range 8500-10700 MeV.

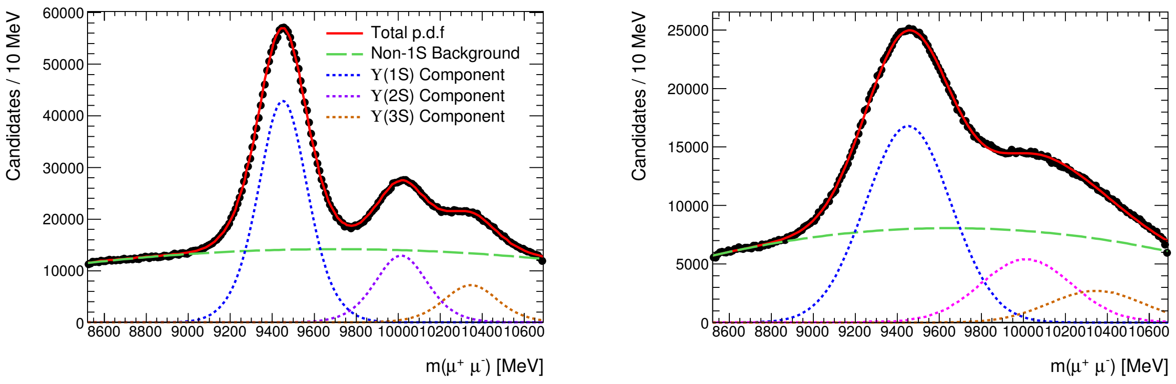


(a) (b)

Figure 6.2: As above, with the χ^2 cut omitted, and the distributions separated into barrel, $|y^{\mu\mu}| < 1.2$, and endcap, $1.2 < |y^{\mu\mu}| < 2.4$, regions.

regions, in Figure 6.2. These shapes are much simpler — in particular, the combinatorial (i.e. non-1S) background can be visualised as the flat part of the distribution in the lower sideband ($m < 9000$ MeV, say). Both distributions can be fitted in a way very similar to that described in [6]. Namely, the p.d.f. is constructed from the following elements:

- **Signal terms**. For each $\Upsilon(\mathrm{nS}, n = 1, 2, 3)$, two Gaussians — a narrow and broad component — are used. The mass of the $n = 1$ state is floated, with the $n = 2$ and $n = 3$ masses fixed by the world average value of the mass differences between these states [3]. The σ parameters for the broad and narrow signal components are assumed to be proportional to mass; these parameters are floated for the $\Upsilon(1\mathrm{S})$ signal term, and this relationship is used to constrain their values for the $\Upsilon(2\mathrm{S})$



(a) $|y^{\mu\mu}| < 1.2$

(b) $1.2 < |y^{\mu\mu}| < 2.4$

Figure 6.3: The fitted $m(\mu^+\mu^-)$ mass distributions for the barrel (left) and endcap (right) regions, for *Upsilon Candidates* used in the analysis with the χ^2 requirement removed.

and Υ(3S) signal peaks. The ratio of narrow to broad yields is constrained to be equal for each of the states.

- **Non-1S background**. This is represented by a 2^{nd}-order Chebychev polynomial.

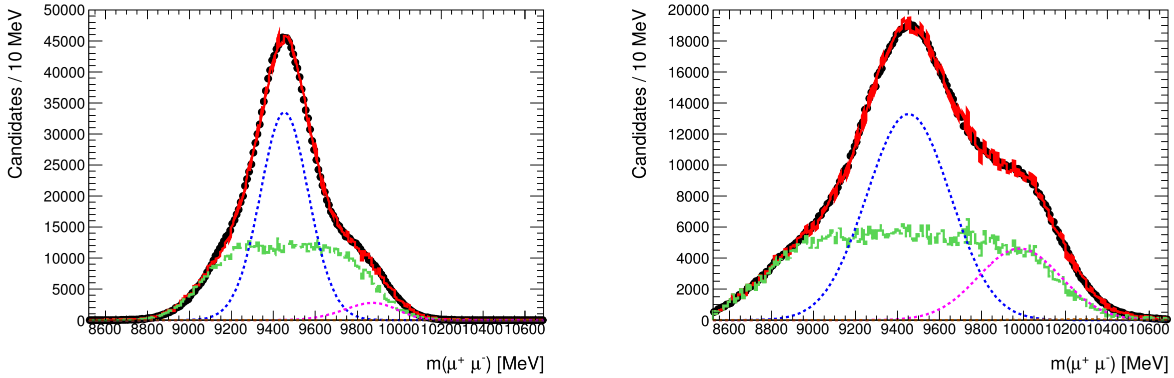
These fits, for each rapidity bin, are given in Figure 6.3.

The approach for fitting the full distributions (i.e. with the χ^2 restriction included) was to introduce a modification to the above fitted shapes. This was done by modelling the effect of the χ^2 selection on the signal terms by a Gaussian *penalty*; that is, a multiplicative Gaussian factor, centred at the vertex-constrained mass, 9460.3 MeV, and with a floated width parameter. Owing to the different inherent resolution in the barrel and endcap, the width of this penalty term was allowed to be different in each. All of the other parameters for the signal p.d.f.s are fixed to their values from the fits in Figure 6.3. The background shape is less well-behaved, and rather than use this method, a template was derived from the same-sign $m(\mu^\pm\mu^\pm)$ shape (see Section 6.2.2 to see how this same-sign sample was selected). The resultant fits are shown in Figure 6.3.

To extract the inclusive and non-1S yield within the allowed window, $9110.3 < m < 9810.3$ MeV, the respective p.d.f.s were integrated within this range. For the purpose of this study, the Υ(2S) $\to \mu^+\mu^-$ component falling in this window is assigned to the inclusive 1S category, for which its characteristics should be the most similar. Performing this integration and adding together the two rapidity bins gives:

$$N_{\mathrm{1S}} = 1.67 \times 10^6$$

$$N_{\text{Non-1S}} = 1.17 \times 10^6$$



(a) $|y^{\mu\mu}| < 1.2$ (b) $1.2 < |y^{\mu\mu}| < 2.4$

Figure 6.4: The resultant fits for the final mass distributions for the *Upsilon Candidates*, where the analysis now includes the χ^2 requirement.

No error is quoted here, because all that is currently available is the statistical uncertainty returned by the fit. The more dominant source of uncertainty comes from the shape of the non-1S template, derived from same sign muon combinations. To estimate the effect of this on the yields, each bin of the template is Poisson-fluctuated and the entire fitting procedure is repeated. For 1000 such fluctuations, the inclusive and non-1S yields were extracted as above and binned-χ^2 Gaussian fits were performed to their distributions. The means of the fits were taken as the best estimates of the yields, with the widths as the uncertainties:

$$N_{\text{1S}} = (1.707 \pm 0.012) \times 10^6$$
$$N_{\text{Non-1S}} = (1.158 \pm 0.011) \times 10^6$$

In Section 5.5.1, a prediction of 1.9×10^6 was made for N_{1S}, based on MC efficiencies and the measured $\Upsilon(\text{1S}) \to \mu^+\mu^-$ production cross section. That this is far too high was an indication that these efficiency estimates were inaccurate — this will be revisited shortly. As suspected, the non-1S yield is of the same order of magnitude as inclusive production.

6.2.2 Same-Sign Muons as a model for the Non-1S Background

The remaining unknown element of the background model was the non-1S $\pi^+\pi^-\Upsilon(\text{1S})$ mass shape (that of the inclusive 1S can be taken from the background MC sample). Since the same-sign muons are certainly not the product of an $\Upsilon(\text{1S})$ decay, it is reasonable to assume that a non-1S model may be constructed from them. That is, the mass distribution for combinations of the type $\mu^\pm\mu^\pm\pi^+\pi^-$ should be similar to the non-1S shape. Unfortunately, there was again the issue that same sign muons are not allowed by the analysis trigger being used. Other similar triggers, which did not require opposite sign muons, were available but with very low statistics. As such, no specific trigger requirement was used and same-sign muon events were the result of a single muon triggers, jet, electron, missing energy triggers etc. — many of which have event topologies very different to the signal-type events. In an attempt to make them as similar as possible, same-sign events were required to contain a low$-p_T$ (defined as < 10 GeV) triggered muon. Furthermore, events where a high-threshold muon trigger fired were vetoed.

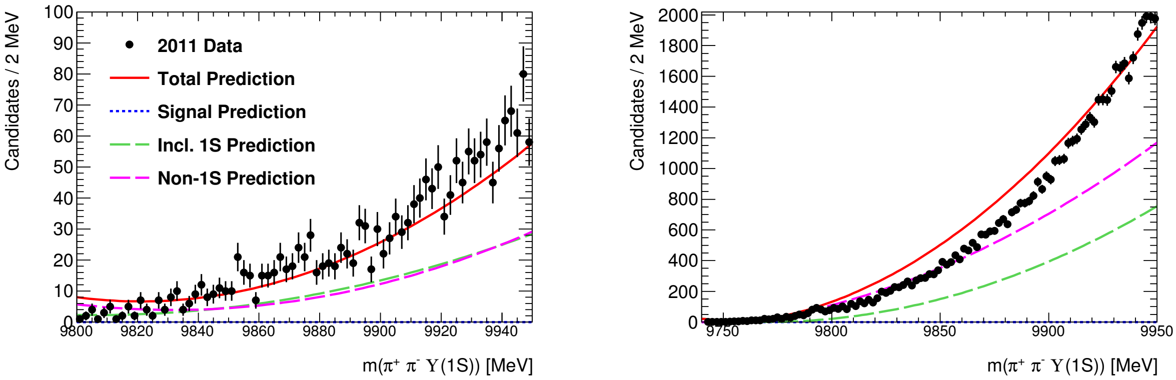
Despite these measures being put in place, at least one major issue remained. In the 2011 dataset, the mean number of $\pi^+\pi^-$ combinations per Υ(1S) candidate was 15.83. Using the known mean from the inclusive 1S MC sample (14.6) and the measured proportion of inclusive to non-1S events, the mean number of candidates for the non-1S sample should be close to 17.2. In the same-sign model constructed here, the mean was found to be 21.94. This is a significant difference, and indicates a fundamental difference between opposite and same-sign events. A number of different possible reasons for this were investigated, including differences in number of primary vertices, number of tracks per event and the effect of most of the analysis selections; none of these provided a satisfactory explanation. The difference in mean number of candidates was found to persist even when restricting to the *EF_2mu4_DiMu_noOS_noVTX* and *EF_2mu4T_DiMu_noOS_noVTX* triggers, which as the name suggests, are essentially those used in the analysis, with the opposite sign requirement removed.

Whether this difference would render the same-sign sample a poor non-1S model remained to be seen. At this point, a simple fix was adopted whereby 17.2 was used (rather than 21.94) wherever the mean number of candidates was needed (in particular, in step 3 below).

6.3 The $\pi^+\pi^-\Upsilon$(1S) Mass Spectrum Below the Υ(2S)

The region of the mass spectrum lying below the Υ(2S) provided an ideal testing ground for the background modelling techniques developed in the previous section. Specifically, predictions of the background for $m < 9950$ MeV were made using the following procedure:

1. Within the inclusive 1S MC sample, a background-only maximum likelihood fit was performed to the $\pi^+\pi^-\Upsilon$(1S) mass distribution in the region of $m < 9950$.



(a) Candidate selection (b) All candidates

Figure 6.5: The background predictions for $m < 9950$ MeV for the candidate selection (left) and all candidates (right) analysis approaches. The 2011 data is overlaid.

2. The fractional integral of this subset of the histogram, with respect to the total sample size, was determined.

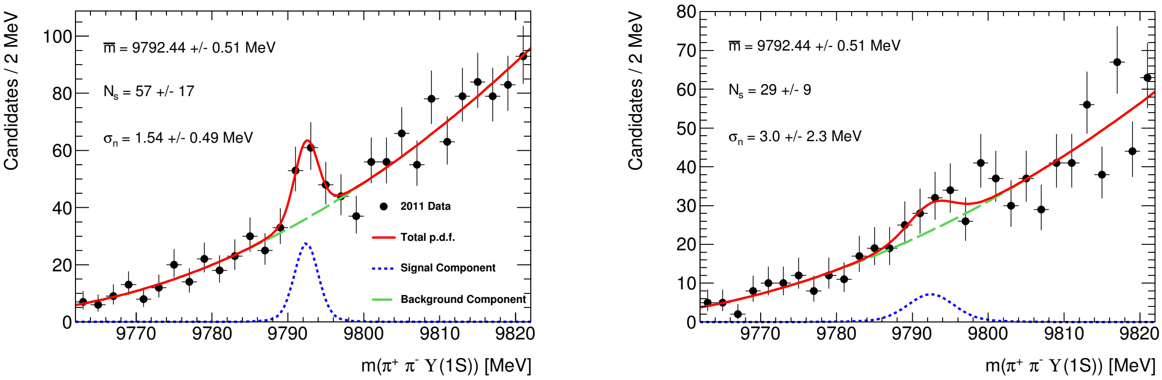
3. This was multiplied by the total inclusive 1S yield, N_{1S}, to get a scaling factor, S_{inc}. In the case of all candidates, S_{inc} was also multiplied by the mean number of candidates per Υ(1S).

4. The process was repeated for the same-sign sample (representing the non-1S component).

5. The two fitted background shapes were scaled by their respective scale factors, and added together to give the total background prediction.

The predictions for all candidates and candidate selection are shown in Figure 6.5 (the process above can be applied to the individual rapidity bins, but at this point the two were analysed together). In the case of candidate selection, the prediction is reasonable, though the limited statistics make it hard to say anything definitive. For all candidates, on the other hand, the predicted shape clearly does not represent the data well.

Rather than launch into a full-scale investigation of why this may be, it was decided that this background prediction apparatus would be tested on the Υ(2S) mass region, which has substantially larger statistics. This also allowed a simultaneous test of the prediction of the observed signal yields based on the MC efficiencies.



(a) Barrel (b) Endcap

Figure 6.6: The simultaneous fit to the structure at $m = 9792$ MeV, with the mean and sigma parameters allowed to float.

6.3.1 The Structure at $m = 9792$ MeV

Before moving on from this mass region, a structure at roughly $m = 9790$ MeV in the all candidates spectrum was noticed. A simultaneous fit to the peak at $m = 9792$ across the two rapidity bins was carried out in a 60 MeV range (Figure 6.5), with the mean and σ_n parameters initially floated. At their best-fit values, the local significance (calculated with the likelihood ratio statistic) of the peak is 4.8σ. Even with the parameters fixed to their predicted values, from the MC signal shape parameters (see 5.4), the significance is still 4.1σ. In other words, it was very unlikely that the bump was simply a statistical anomaly, and as it turns out there is a simple explanation for it.

The hint leading to the explanation for this peak lies in its particular mass. Consider the well-known decay $\Upsilon(3\mathrm{S}) \to \pi^+\pi^-\Upsilon(2\mathrm{S})$, where the $\Upsilon(2\mathrm{S})$ then decays into a dimuon pair. Because a component of the $\Upsilon(2\mathrm{S})$ peak spills over into the allowed $m(\mu^+\mu^-)$ region, a number of these decays will be reconstructed in this analysis. In constraining the value of $m(\mu^+\mu^-)$ to $m_{1\mathrm{S}}$ in the $\pi^+\pi^-\Upsilon(1\mathrm{S})$ vertex fit, however, the $\Upsilon(2\mathrm{S})$ will effectively be allocated a mass which is too low, by $m_{2\mathrm{S}} - m_{1\mathrm{S}}$. Hence, $\Upsilon(3\mathrm{S}) \to \pi^+\pi^-\Upsilon(2\mathrm{S})$decays recorded by the analysis should appear at $m_{3\mathrm{S}} - (m_{2\mathrm{S}} - m_{1\mathrm{S}}) = 9792.24$ MeV, rather than at $m_{3\mathrm{S}}$. To confirm this was contributing to the peak at 9792 MeV, the variable $\widetilde{m}_{2\mathrm{S}} = m(\mu^+\mu^-\pi^+\pi^-) - m(\mu^-\mu^+) + m_{2\mathrm{S}}$ was plotted in the region near the $\Upsilon(3\mathrm{S})$ mass, where $m(\mu^+\mu^-\pi^+\pi^-)$ is the mass obtained only from the four-vectors of each track (i.e. not from the vertex fit). Since this variable effectively fixes the dimuon mass to $m_{2\mathrm{S}}$, rather than $m_{1\mathrm{S}}$, a peak is expected at the correct place, $m_{3\mathrm{S}}$. The distribution, shown in Figure 6.7 (both rapidity bins combined), has an obvious peak and confirms the presence of the $\Upsilon(3\mathrm{S}) \to \pi^+\pi^-\Upsilon(2\mathrm{S})$ in the 2011 dataset.

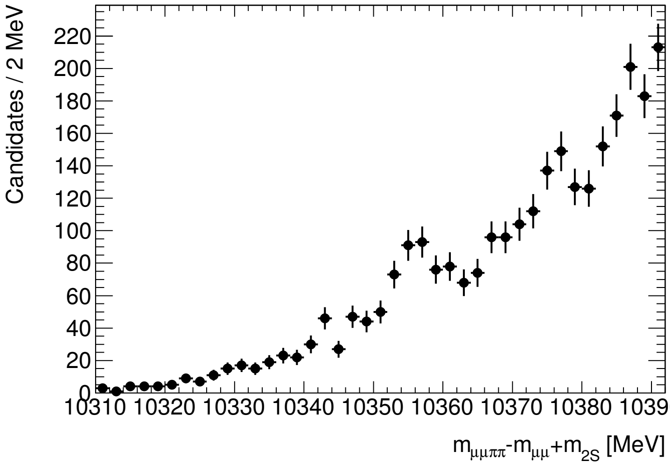


Figure 6.7: The distribution of the mass variable, $\tilde{m}_{2S} = m(\mu^+\mu^-\pi^+\pi^-) - m(\mu^-\mu^+) + m_{2S}$, confirming the presence of the decay $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(2S)$.

One issue still remained — if the allowed dimuon mass range for the analysis was tightened to $9.1103 < m(\mu^+\mu^-) < 9.5$ GeV, the $\Upsilon(2S) \to \mu^+\mu^-$ contamination should be insignificant and hence this structure at 9792 MeV should disappear. This was found not to be the case — in fact, the strength of the peak was only slightly diminished. So, as well as the $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(2S)$ decay chain, there must be other constituents of the 9792 MeV signal peak. Because of the precise mass value, it was natural to assume that this would again involve the $\Upsilon(3S)$.

A possible explanation was in the $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(2S)\,(\to X\Upsilon(1S))$ decay chain, where X represents any decay products and the $\Upsilon(1S)$ decays into a dimuon pair. In the first stage of this decay, the phase-space is very constrained — the $\pi^+\pi^-$ system and $\Upsilon(2S)$ are produced with low momentum (177 MeV). Consequently, it can be shown that the mass distribution of the system comprising these two pions and the final two muons is very narrow and has a mean of 9792 MeV. In fact, in a simple toy MC study assuming pure phase-space, the width of the mass distribution was found to be 2.18 MeV (see Figure 6.8), which is consistent with what was observed. Although it was not possible to demonstrate this second decay chain directly with the data, it is true that this, in principle, will contribute to the 9792 MeV peak to some extent.

Finally, it is noted that there is no such peak in the candidate selection spectrum. In fact, the candidate selection procedure appears to have an effective threshold of ≈ 9.8 GeV — that is, the efficiency for candidate selection below this point is effectively zero.

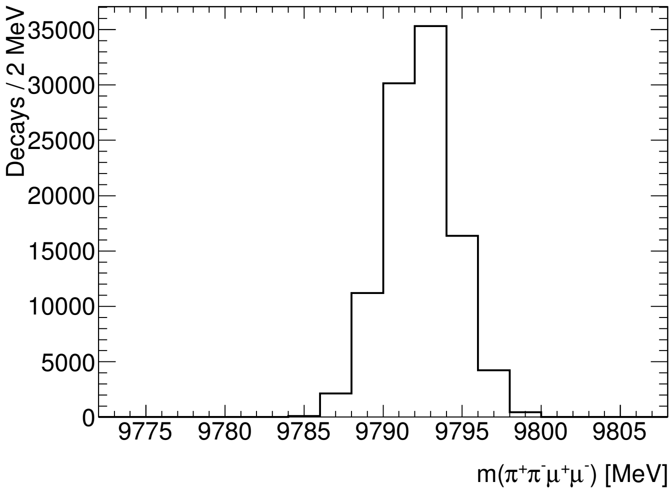


Figure 6.8: The $m(\mu^+\mu^-\pi^+\pi^-)$ distribution in toy $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(2S)\,(\to X\Upsilon(1S))$ decays, where the muons are taken from the $\Upsilon(1S)$ decay and the pions from the initial $\Upsilon(3S)$ decay. Uniform phase space was assumed.

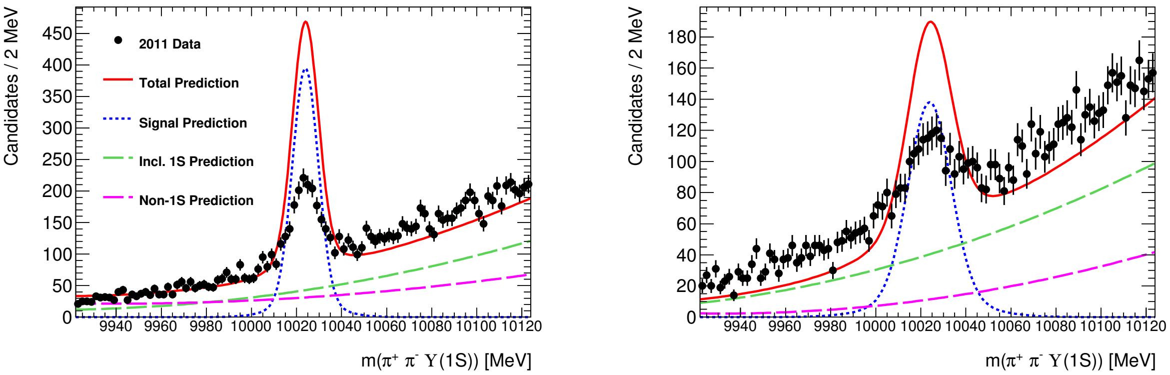
6.4 The $\pi^+\pi^-\Upsilon(1S)$ Mass Spectrum in the $\Upsilon(2S)$ Region

The next region of interest was the vicinity of the $\Upsilon(2S)$ mass. The main objective here was to assess

1. the background predictions based on the modelling approach established in the previous section. Thus far, there has been evidence that the characterisations of the background (in terms of shape and normalisation) may not be particularly accurate. In particular, the background MC sample lead to a prediction of the $\Upsilon(1S)$ yield which was significantly overestimated.
2. to validate the signal MC samples in terms of the efficiencies, and signal shape parameters.

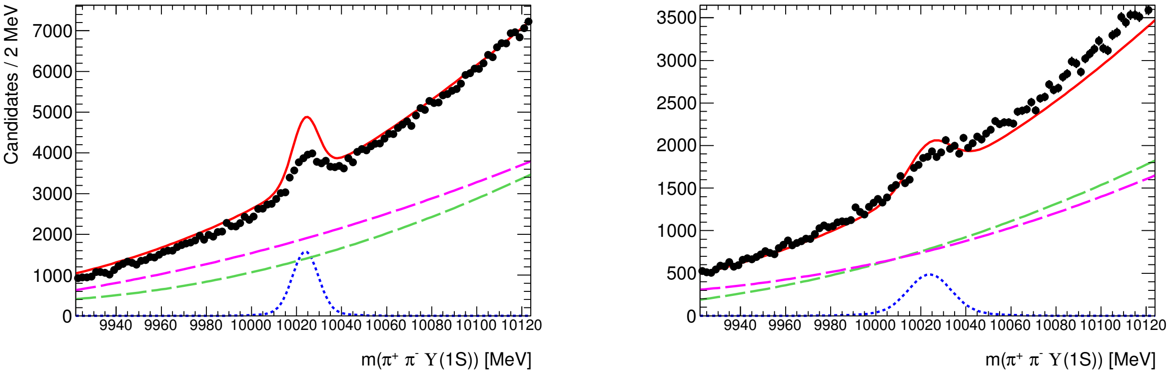
Confidence in the second of these points is essential when performing significance calculations or placing upper limits, and is the main reason for assessing these predictions.

On top of the background prediction (from the procedure outline previously), a signal peak is added. The shape parameters are taken from the $\Upsilon(2S)$ MC sample (see Section 5.4), and the normalisation was estimated using the same method described in 5.5.1. Though the uncertainties on the predicted $\Upsilon(2S)$ normalisations will not be included in what follows (in fact, for the entire chapter), they can be taken as dominated by the luminosity systematic of approximately 1.8% [105] and the uncertainty on the measured



(a) Candidate selection, Barrel

(b) Candidate selection, Endcap



(c) All candidates, Barrel

(d) All Candidates, Endcap

Figure 6.9: The signal-plus-background predictions for the Υ(2S) mass regions in the barrel and endcap bins (left and right, respectively), for the candidate selection and all candidates analysis approach (top and bottom, respectively).

cross section (ignoring the uncertainty of the efficiencies due to MC mis-modelling). The predictions were also split into the analysis rapidity bins, with the yields split according to the bin fractions from the Υ(2S) MC sample. The results for a ±100 MeV window, with the observed data overlaid, are given in Figure 6.9.

There are many issues with these predictions — most obviously, the signal contribution seems to be grossly overestimated in all cases. This pointed to a poor understanding of the analysis efficiency for the $\Upsilon(2\mathrm{S}) \to \pi^+\pi^-\Upsilon(1\mathrm{S})$ decay. The methods by which this was rectified are described in the following section.

6.4.1 Mismodeling in Signal MC Samples

An analysis of the observed Υ(2S) signal peak will be given shortly, but for now it is noted that the fitted yields are $N_{2\mathrm{S}} = 9231 \pm 796$ for all candidates, and $N_{2\mathrm{S}} = 1541 \pm 143$ for candidate selection. The predicted number of $\Upsilon(2\text{S}) \to \pi^+\pi^-\Upsilon(1\text{S})$ decays (18290 and 4796 for all candidates and candidate selection, respectively) are therefore 2-3 times what was observed. This implies that the efficiencies determined from the $\Upsilon(2\text{S})$ MC sample must be overestimated by a similar factor. Recall that the predicted number of $\Upsilon(2\text{S})$ was determined according to the following formula:

$$N_{2\text{S}} = \mathcal{L}_{\text{int}}\,(\sigma\mathcal{B})_{2\text{S}}\,\mathcal{A}_{\mu\mu}\varepsilon', \tag{6.1}$$

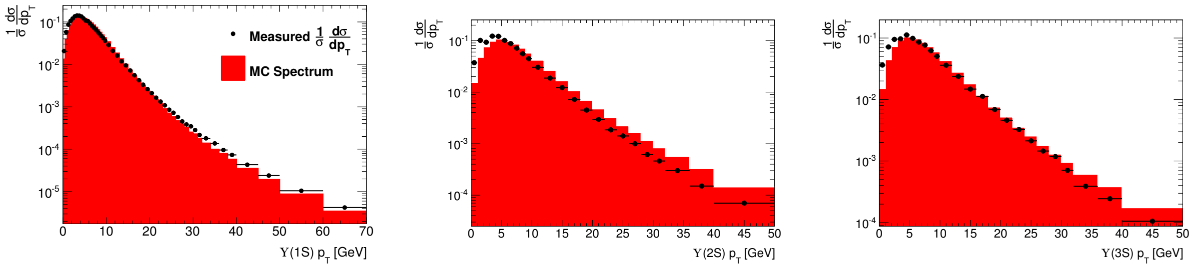
where $\mathcal{L}_{\text{int}}$ is the integrated luminosity for the 2011 dataset, $(\sigma\mathcal{B})_{2\text{S}}$ is derived from the measured production cross section, $\mathcal{A}_{\mu\mu}$ is the muon acceptance as determined by the *acceptance calculator* (see Section 5.2), and ε' is the remaining analysis efficiency.

In Section 5.1, it was pointed out that the decays were generated under the assumption of a parent production spectrum predicted by NRQCD, and uniform phase-space. As it turns out, these assumptions have a significant impact on the analysis efficiencies and the muon acceptance, and are responsible for the poor $\Upsilon(2\text{S})$ yield predictions:

- **Production Spectrum Differences**
  The biggest cause of efficiency loss is the p_T threshold placed on the pions, which are naturally very soft. As was shown in Section 5.2, only relatively highly boosted ($p_T > 10$ GeV) decays are reconstructable with respect to this threshold. Small changes in tail of the underlying production p_T spectrum can have a significant effect on the acceptance and analysis efficiency.

  Fortunately, the differential production cross section for the $\Upsilon(1\text{S})$,$\Upsilon(2\text{S})$ and $\Upsilon(3\text{S})$ states have been measured directly by ATLAS [6], and can be compared with the NRQCD production spectra used by the MC generator. The normalised differential cross sections based on the two are shown in Figure 6.10. For the $\Upsilon(1\text{S})$ and $\Upsilon(3\text{S})$ states, the difference is not particularly pronounced, but for the $\Upsilon(2\text{S})$ production there is a large difference in the tail. Because the decays in this tail are those reconstructed in the analysis, this discrepancy is important.

  To correct this, *production spectrum weights* were developed for each state from the ratio of their measured and MC spectra. In fact, the measured cross section was actually performed across two rapidity bins — $|y^{\mu\mu}| < 1.2$ and $|y^{\mu\mu}| < 2.25$ — allowing for a doubly-differential ratio to be formed. Each signal decay selected in the analysis was then assigned a weight based on the (p_T, y) at which it was generated. For example, $\Upsilon(2\text{S}) \to \pi^+\pi^-\Upsilon(1\text{S})$ decays with a $p_T > 10$ GeV (essentially the whole sample) are assigned weights less than one. For signal decays outside the range of measured cross sections, the weight was taken to be 1. As will be shown



(a) Υ(1S) production (b) Υ(2S) production (c) Υ(3S) production

Figure 6.10: A comparison of the normalised differential production cross sections in the MC samples and those measured in [6], for the Υ(1S), Υ(2S) and Υ(3S) states.

shortly, this reweighting scheme almost completely accounts for the overestimated Υ(2S) yield in 2011 data.

These differences also affect the calculated muon acceptance. The fix for this is simple — rather than have the acceptance calculator sample the decays from the NRQCD prediction, it directly used the measured differential cross section.

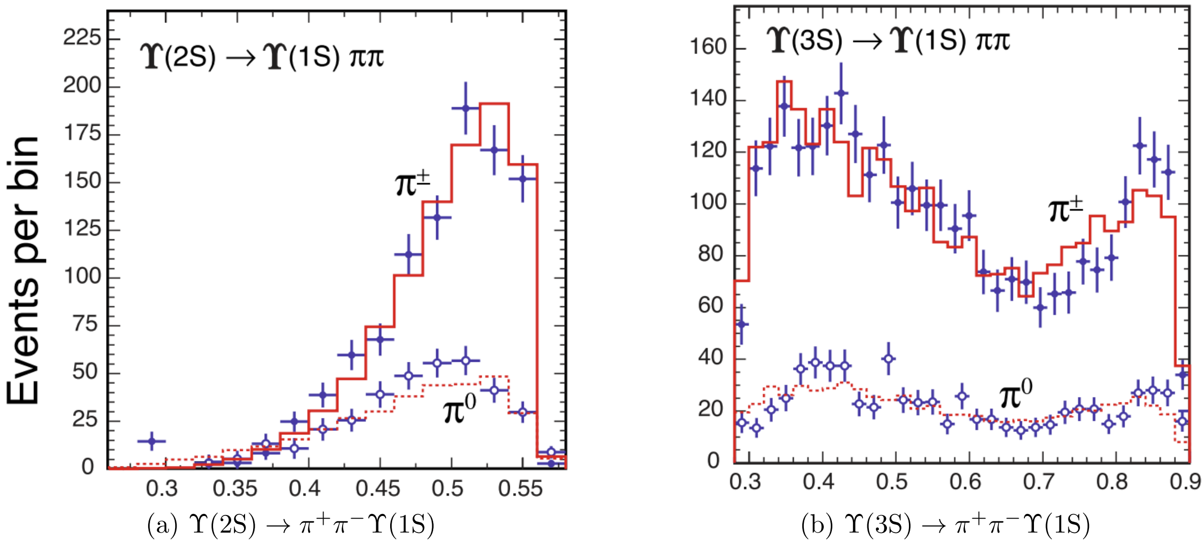
- **Dipion Mass Distribution**

  The phase-space in the MC generator was assumed to be uniform. However, it is well-known that this is not the case — in particular, the $m(\pi^+\pi^-)$ distribution typically has large deviations from the phase-space shape. The dipion mass distributions for the $\Upsilon(2S) \to \pi^+\pi^-\Upsilon(1S)$ and $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)$ transitions, as measured by the CLEO collaboration [13], are given in Figure 6.11. In the Υ(2S) case, the observed CLEO shape can be fitted by a number of simple functions given in an earlier CLEO paper [8]. For simplicity, only the Voloshin and Zakharov model [106] was considered here, and is given by

  $$\frac{d\sigma}{dm_{\pi^+\pi^-}} \propto PS \cdot \left(m^2_{\pi^+\pi^-} - \lambda m^2_\pi\right)^2, \tag{6.2}$$

  where PS is the phase-space distribution, and the best fit value for λ with the CLEO data is 3.11.

  In this form, it was a simple matter to develop a *dipion mass weight* for the Υ(2S) MC sample, in an analogous way to above. Because we want to force the distribution to change from phase space to Equation 7.4, the relevant scale factor is again the ratio of the two. This is essentially $\left(m^2_{\pi^+\pi^-} - \lambda m^2_\pi\right)^2$, up to an overall normalisation factor. This *dipion mass weight* is then assigned to the signal decays selected in the analysis according to their generated dipion mass value. The dipion



(a) $\Upsilon(2S) \to \pi^+\pi^-\Upsilon(1S)$ (b) $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)$

Figure 6.11: The $m(\pi\pi)$ distributions for the $\Upsilon(2S) \to \pi^+\pi^-\Upsilon(1S)$ and $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)$ transitions, as measured by the CLEO collaboration [13].

mass weights were taken to be independent from the production spectra weights, meaning the total weight was just the product of the two.

For the $\Upsilon(3S)$, extracting the correct functional form for the dipion mass distribution took more effort, and required the observed distribution from the 2011 dataset. An explanation of this is deferred to the relevant section below (6.5.1), but for now it should be understood that the weighted $\Upsilon(3S)$ sample includes this correction.

These dipion mass weights improve the efficiency estimates from MC (though do not have as large an effect on the efficiencies as the production spectra weights), and allow for a fair comparison with the observed dipion mass spectrum in ATLAS data.

Analogous changes were also made to the acceptance calculator — rather than pure phase space, the dipion mass was randomly sampled from the above functional form.

- **Spin-alignment Uncertainty**
  The remaining parameters of the phase space are the angles of decay of the $\Upsilon(1S)$ in the rest frame of the parent, θ^* and ϕ^*, and the analogous angular variables for the dipion and dimuon systems. As explained in Chapter 1, the multipole expansion leads to the standard approximation that the dipion system is emitted in an S-wave state. In this case, the emission of the pions within the dipion rest frame is isotropic, and the polarisation state of the $\Upsilon(1S)$ is inherited from the parent (see [13] and references therein). This is a convenient simplification, because the

angular distribution of a vector state into a dilepton pair is well-known (see Section 4.6, for example).

Reweighting the signal decays according to the extreme polarisation scenarios gives shifts in the product $\mathcal{A}_{\mu\mu} \cdot \varepsilon'$ of more than ±50%. Fortunately, a recent measurement from the CMS collaboration [7] has demonstrated that the polarisation of Υ(1S),Υ(2S) and Υ(3S) produced in *pp* collisions is generally consistent with zero. It was therefore justifiable to assume isotropic angular distributions in calculating efficiency and acceptance values.

The above weights also lead to a change in the effective size of each of the samples, equal to the sum of weights of all the generated decays. Taking this into account, the total change in the efficiency and muon acceptance due to these corrections is, for the $\Upsilon(2S) \rightarrow \pi^+\pi^-\Upsilon(1S)$ decay, as follows:

$\mathcal{A}_{\mu\mu}$:	$0.275 \rightarrow 0.261$	
ε' :	$3.15\% \rightarrow 1.69\%$	(All candidates)
	$0.826\% \rightarrow 0.414\%$	(Candidate selection)

leading to predictions of $N_{2S} = 9483$ for all candidates, and $N_{2S} = 2324$. The former of these is consistent with the observed yield, but the latter is still considerably overestimated. This indicated that the candidate selection efficiency (c.f. section 5.3.3) was lower than expected. Because candidate selection is based on dipion p_T, and the kinematics of the signal decays has been corrected for, the issue had to be due to the modelling of the background dipion pairs. In order to determine the cause of this, the background mis-modelling needed to be addressed.

6.4.2 The Lower Sideband Non-1S Model

The production spectrum weights developed above are equally applicable to the Υ(1S) in the inclusive 1S background MC sample. Because the generated p_T spectrum in this sample is reasonably close to the measured distribution (Figure 6.10a), the effect of this is quite minor and the mis-modelling of the background (evident in the predictions for the Υ(2S) region) remained. The same-sign muon sample was therefore abandoned as a model for the non-1S background.

The other source of combinatorial muons is the sidebands of the $m(\mu^+\mu^-)$ distribution. Unfortunately, the upper sideband of the Υ(1S) is contaminated by $\Upsilon(2S) \rightarrow \mu^+\mu^-$ decays, which, in general, do not share the same characteristics as the non-1S muons. As such, only the lower sideband — $8000 < m(\mu^+\mu^-) < 8900$ MeV — was selected, where the upper bound was chosen to ensure minimal contributions from $\Upsilon(1\text{S}) \to \mu^+\mu^-$ decays (see Figure 6.3).

The assumption here is that the combinatorial background in the sidebands is similar in nature to that in the signal region, and because $m(\mu^+\mu^-)$ is fixed to 9460.3 MeV in the four track vertex fit, the $\pi^+\pi^-\Upsilon(1\text{S})$ mass distribution should be similar. Some adjustments were made to the sideband sample to resolve the following differences:

1. **An effective χ^2 maximum.** As discussed above, the χ^2 restriction is harsher in the sidebands than in the signal $m(\mu^+\mu^-)$ region, and so needed to be replaced by an effective value. As the maximum allowed value is increased, the mean number of dipion candidates per $\mu^+\mu^-$ pair increases monotonically. The expected mean for the non-1S component was deduced as ~ 17.2, so an effective χ^2 restriction of was chosen to target this value. The mean was most closely reproduced for $\chi^2 < 50$.

2. **Dimuon kinematic weighting.** With confidence in the $\Upsilon(1\text{S})$ kinematics of the weighted inclusive 1S background sample, the expected dimuon kinematics of the non-1S component could be obtained by subtracting the former from the 2011 dataset. This requires $N_{1\text{S}}$, which was previously determined. Sideband weights were developed in bins of dimuon $(p_T, |y|)$ based on the ratio of the observed sideband spectrum and the known non-1S spectrum, in an analogous manner to the production spectrum weights discussed above. These weights will be referred to as the *dimuon sideband weights*.

3. **Dipion kinematic weighting.** Weighting the sideband dimuons to the correct kinematic distribution tends to 'pull' the associated dipions into a $(p_T, |\eta|)$ distribution which is mostly correct. To account for any residual differences, *pion sideband weights* were formulated in $(p_T, |\eta|)$ bins, based on the mismatch between the all candidates same-sign sample in the sideband (after the above corrections) and signal regions. These weights were then applied to the dipions in the opposite-sign sideband sample.

After these corrections, the mean number of candidates per dimuon pair is 18.5. When used in the predictions below, the mean was actually fixed to the expected value, as before, but with a renewed value due to the reweighting of the inclusive 1S MC samples (which, as it turns out, is 18.3).

This model for the non-1S component of the background, coupled with the inclusive 1S MC sample, produced much better agreement with the observed background near the

Υ(2S) (see Figure 6.13 below). However, there still remains the issue that the candidate selection efficiency is overestimated in the signal MC samples. As discussed, this pointed to the mis-modelling of background dipion pairs in the MC. With a reasonable non-1S model in hand, this could now be dealt with.

6.4.3 Further corrections to the MC Samples

The idea here was to use the non-1S model developed above from the lower sideband to isolate the inclusive 1S component of the 2011 dataset; in fact, the non-1S component could now be subtracted from the 2011 data using the known $N_{\text{Non-1S}}$ yield. The remainder of the distribution (i.e. after subtraction) should match the inclusive MC distributions.

Since the dipion p_T is the main variable of interest for candidate selection, the mean value of this variable per Υ(1S) candidate was studied (other such global-level variables may be equally as good). Because this relies on the sideband non-1S model and the corrections already explained for the MC sample, the distribution of the mean dipion p_T was extracted as follows:

1. For each Υ(1S), the mean p_T of all the associated dipion pairs was calculated as a weighted average (by the candidate-by-candidate $\pi^+\pi^-\Upsilon(1S)$ weights — e.g. the dipion sideband weights), then
2. this contributed to the distribution of means with a weight due to the corrections described above that pertain to the dimuon pair.

The result, shown in Figure 6.12, clearly indicates that the overall dipion p_T spectrum is harder in data than it is in the MC. Once more, this difference could be alleviated by weighting the MC based on the ratio of the two distributions, giving rise to the so-called *mean dipion p_T weights*. These weights were then applied not only to the inclusive 1S MC sample, but all of the signal samples too (since their modelling of the background pions is equivalent). In the latter, the calculated mean did not include the signal dipion pair.

This stage of reweighting was only used in conjunction with candidate-selection; for the all candidates approach, the signal efficiency should not be strongly dependent on the kinematics of the background dipion pairs.

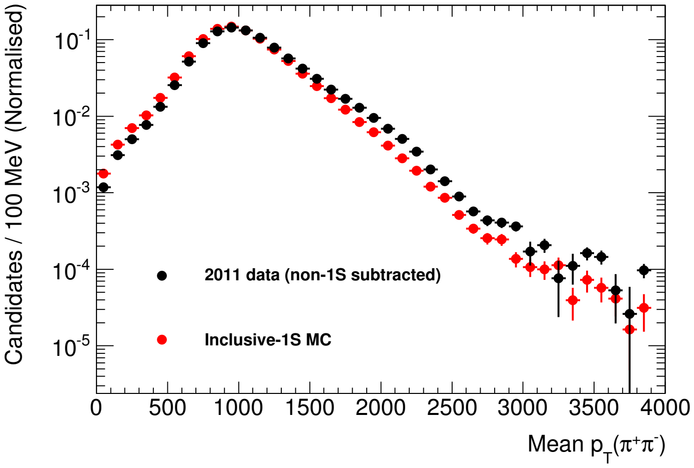


Figure 6.12: Distribution of the mean dipion p_T per Υ(1S) candidate in inclusive MC events, and 2011 data with the non-1S component subtracted.

6.4.4 Analysis of the Υ(2S) Region After Corrections

In the sequence of weighting schemes, each layer of complexity was built on the confidence afforded by the previous step. For clarity, a summary of these procedures is provided here:

1. The MC samples are weighted based on
    - (for inclusive 1S, Υ(2S), Υ(3S)) their production spectrum — to match the measured differential cross section
    - (for Υ(2S), Υ(3S)) their dipion mass spectrum — to match CLEO observations.
2. The *acceptance calculator* was likewise altered to reflect the correct production spectrum and dipion mass spectrum.
3. The lower sideband model was developed for the non-1S background component, with the adjustments:
    - an effective χ^2 maximum of 50.
    - weighting for dimuon $(p_T, |y|)$ to match the shape in 2011 data, with inclusive 1S MC subtracted.
    - weighting the pion $(p_T, |\eta|)$ spectrum based on the discrepancy between same-sign distributions in the sideband and signal regions.

4. When using candidate selection, the MC samples were weighted to account for differences in the mean dipion p_T in the inclusive 1S sample and the 2011 data, with the non-1S component subtracted.

After all of these steps, the analysis efficiencies for the Υ(2S) sample were determined to be $\varepsilon' = 1.69\%$ for all candidates, and $\varepsilon' = 0.274\%$ for candidate selection. This leads to yield predictions of $N_{2S} = 9483$ and $N_{2S} = 1540$, in excellent agreement with the fitted values (9227 ± 500 and 1546 ± 102, respectively). These corrections also lead to an improvement of the predicted N_{1S} yield, from 1.926×10^6 to 1.723×10^6 (which is much closer to the measured value of 1.707×10^6).

The renewed signal-plus-background predictions, with all the inputs (barring the signal shape parameters — see below) recalculated based on the corrected MC samples, are shown in Figure 6.13. In comparison to the earlier predictions (Figure 6.9), each of these is much improved; the only obvious remaining issue are the normalisations of the background, which is not of great importance.

With this established, the Υ(2S) peak was now studied more closely.

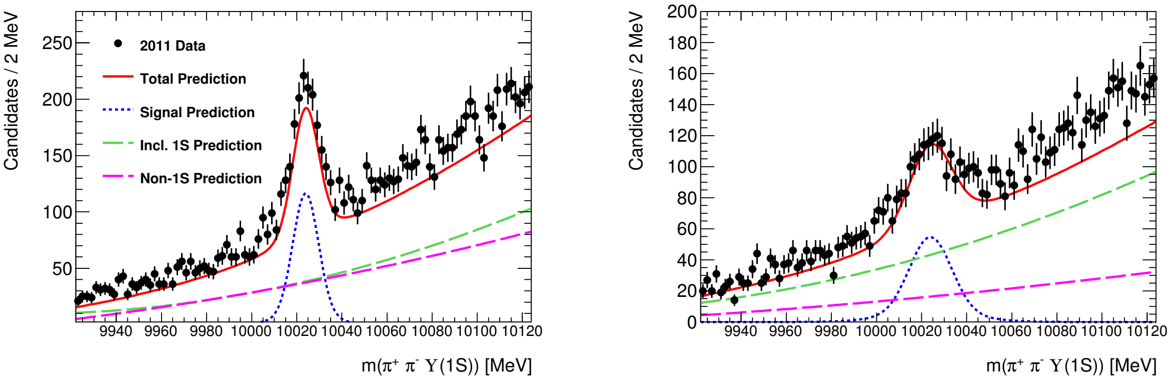
6.4.4.1 Results for Fits to the Υ(2S) Peak

As discussed in the previous chapter, the signal shape parameters for a particular parent mass are dependent mostly on the underlying rapidity distribution of the sample. Because none of the above weighting methods for the signal MC affected the rapidity in a significant way, the signal shape in the weighted samples was expected to be consistent with the original values.[1] The corrected samples also contain weighted data, which is problematic for likelihood fits, and so it was sensible not to recalculate the shape parameters based on these.

The observed Υ(2S) peak in the 2011 dataset was fitted independently (i.e. *not* simultaneously) in each rapidity bin, with r and f_n fixed to the averages determined in the previous chapter (see Section 5.4), and the mean and σ_n parameters floated. The results for candidate selection and all candidates, given in Figure 6.14, shows that the fitted σ_n parameters are all consistent with the values from the Υ(2S) MC sample, and the fitted masses are all consistent with the world average value of 10023.26 MeV [3]. The most accurate measurement of the Υ(2S) mass comes from the central all candidates bin which, after removing the fitted mean bias of Section 5.4.2, is $m_{2S}^{\mathrm{fit}} = 10023.25 \pm 0.29$ MeV.

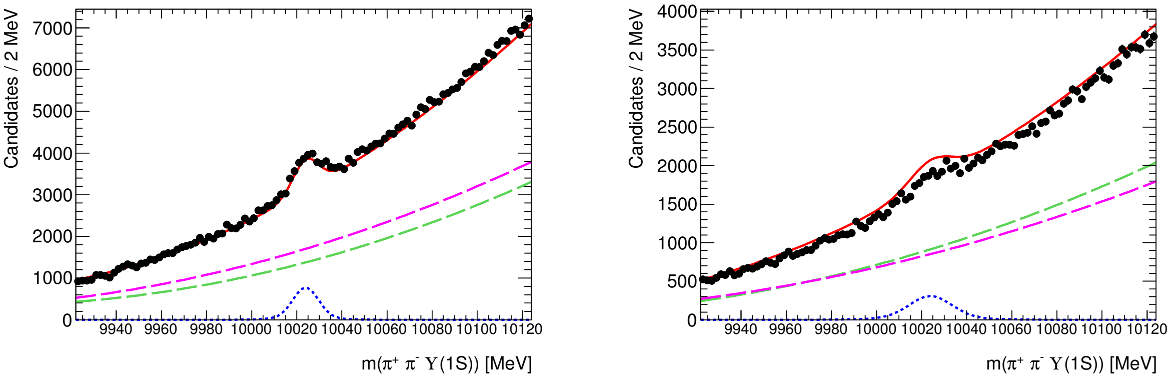
---

[1] This was explicitly confirmed with the Υ(2S) sample.



(a) Candidate selection, Barrel

(b) Candidate selection, Endcap



(c) All candidates, Barrel

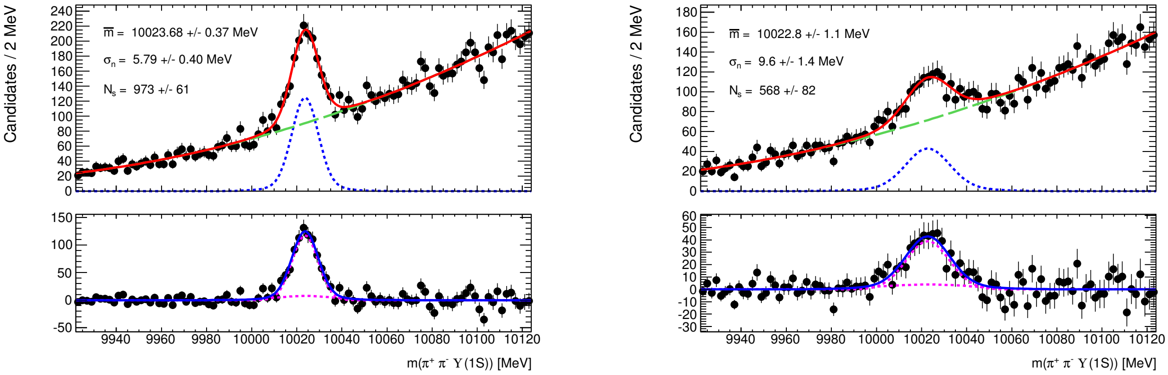
(d) All Candidates, Endcap

Figure 6.13: The signal-plus-background predictions for the Υ(2S) mass regions using the corrections to the MC samples and the sideband model for the non-1S component.

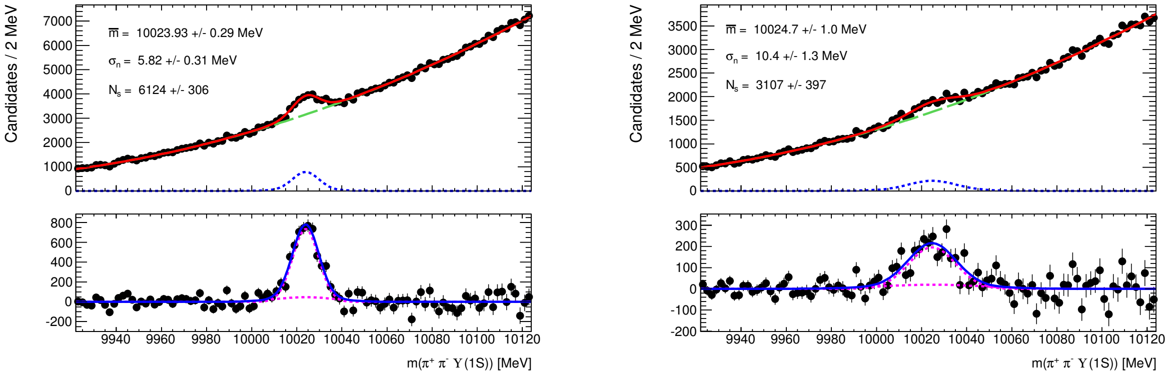
For completeness, fits were also performed with *all* of the signal parameters floated (not shown here). These were found to be consistent with the above fits, but with larger uncertainties on the floated parameters and no improvement in the reduced-χ^2 values.

6.4.4.2 Υ(2S) p_T, Rapidity and $m_{\pi^+\pi^-}$ Distributions

To further validate the (weighted) Υ(2S) signal MC samples, several key distributions for the Υ(2S) $\to \pi^+\pi^-\Upsilon$(1S) decays were checked. Doing so required isolating the signal decays, which was performed using the sideband-subtraction technique. Because of the larger amount of statistics available, the all candidates sample was utilised here.



(a) Candidate selection, Barrel (b) Candidate selection, Endcap



(c) All candidates, Barrel (d) All Candidates, Endcap

Figure 6.14: The fitted mass distributions for the Υ(2S) mass region, with the r and f_n parameters fixed. The lower plot of each sub-figure is the data, with the fitted background shape subtracted.

The following regions were defined:

$$\begin{aligned}
\text{Signal Region: } 10023.9 - 15\text{ MeV } &< m < 10023.9 + 15\text{ MeV} \\
\text{Lower Sideband Region: } 10023.9 - 45\text{ MeV } &< m < 10023.9 - 15\text{ MeV} \\
\text{Upper Sideband Region: } 10023.9 + 15\text{ MeV } &< m < 10023.9 + 45\text{ MeV},
\end{aligned} \tag{6.3}$$

where 10023.9 MeV was chosen as it is the fitted value in the dominant lower rapidity bin. The sideband regions were placed directly adjacent to the signal region to negate as much of the non-linear mass-dependence of the underlying variables as possible.

These bands are shown schematically in Figure 6.15. Because the sidebands are equal width to the signal region, the sideband procedure was simple; for the variable of interest (e.g. p_T), the two sideband distributions were averaged and subtracted from the signal

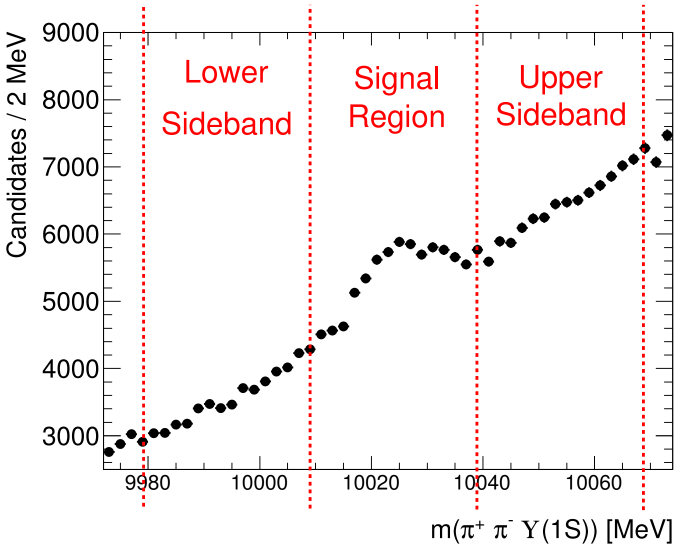
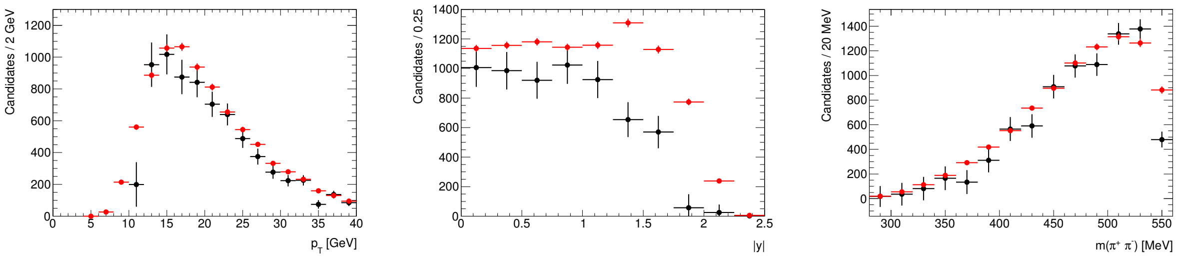


Figure 6.15: The definition of the signal and sideband regions for the sideband-subtraction procedure on the Υ(2S) region.

distribution. In expecting the result to accurately isolate the signal decays, the standard approximation regarding the linearity of any mass dependence must be made. For the $m(\pi^+\pi^-)$ and p_T distributions, where there is a mass dependent threshold involved, this assumption may break down at the edges. No attempts were made to account for this possibility.

The sideband-subtracted p_T, y and $m(\pi^+\pi^-)$ distributions are provided in Figure 6.16. The corresponding signal-only distributions from the Υ(2S) MC, scaled to achieve a normalisation equal to the observed yield, $N_{2\mathrm{S}} = 9231$, are overlaid. The normalisation of the sideband-subtracted signal region was actually 6174, implying that the either (1) the lower sideband contains an upward fluctuation, (2) the background in the signal or upper sideband regions contains a downward fluctuation or (3) a combination of both of these is occurring. Despite this, most of the p_T distribution is reasonably well-matched to the MC spectrum; as expected, though, the threshold region is problematic. The dipion mass spectrum also displays excellent agreement with the expected distribution.

For the rapidity distribution, the situation is less ideal. In the previous two cases, the additional background (responsible for the low normalisation) had obviously distributed itself in signal-poor regions of the respective distribution (for example, in the un-physical dipion mass region above $m_{2\mathrm{S}} - m_{1\mathrm{S}}$). In the case of the rapidity, there is no such signal-poor region, and hence the normalisation is noticeably different to the MC expectation. Importantly, however, the shapes are not dissimilar.



(a) Υ(2S) p_T (b) Υ(2S) $|y|$ (c) Υ(2S) $m(\pi^+\pi^-)$

Figure 6.16: The sideband-subtracted p_T, $|y|$ and $m(\pi^+\pi^-)$ distributions for $\Upsilon(2S) \to \pi^+\pi^-\Upsilon(1S)$ decays in the 2011 dataset. For comparison, the weighted Υ(2S) MC sample is shown overlaid with a normalisation equal to the measured all candidates yield, $N = 9231$. Negative-suppression has been used in the p_T and $m(\pi^+\pi^-)$ distributions.

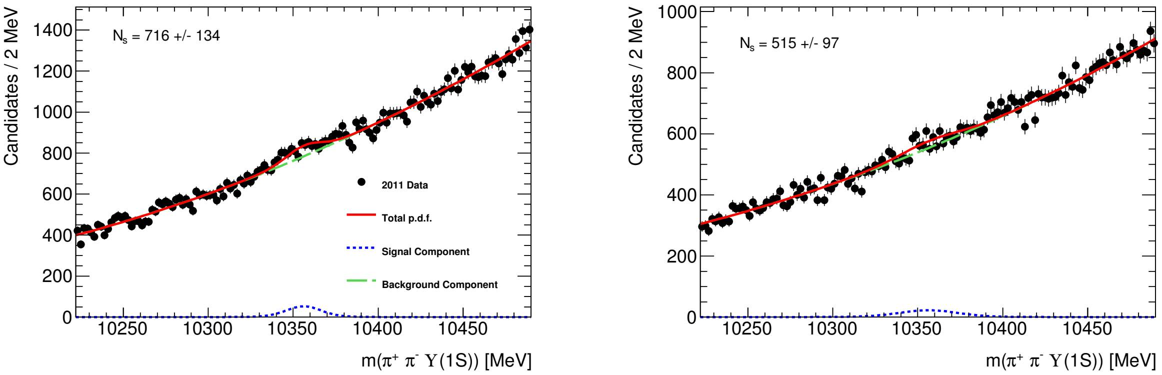
6.5 The $\pi^+\pi^-\Upsilon(1S)$ Mass Spectrum in the $\Upsilon(3S)$ Region

From the studies of the analysis sensitivity in the last chapter (Section 5.5), a significant Υ(3S) peak was not expected. To measure the observed significance at the Υ(3S) mass, the likelihood ratio test was applied to a simultaneous fit, across the two rapidity bins, of the candidate selection sample. The signal shape parameters were fixed in the fit as follows:

- the mass was fixed to the world average value [3] with the average of the low and high-rapidity fitted mean bias added, 10356.2 MeV,
- the σ_n parameters (σ_{yL} and σ_{yH} for the low and high rapidity bins, respectively) were taken from the signal-only fits to the Υ(3S) MC sample, with r and f_n fixed to their usual averages, and
- the yield fractions within each rapidity bin were taken from the weighted Υ(3S) MC sample.

The result for a fitting window of $[10356.2 - 8\sigma_{yH}, 10356.2 + 8\sigma_{yH}]$, projected onto each rapidity bin, is given in Figure 6.17 and has a significance of 5.4σ. In contrast, the equivalent result for all candidates returns a significance of only 3.8σ, reflecting the advantage of candidate selection.

The fitted yield in Figure 6.17 is $N_{3S} = 1227 \pm 165$, consistent with the predicted value of 886.



(a) Barrel

(b) Endcap

Figure 6.17: The projections of a simultaneous fit to the Υ(3S) mass region across the two rapidity bins.

6.5.1 Dipion Mass Distribution

The same sideband-subtraction method used for the Υ(2S) was applied to the candidate-selection sample at the Υ(3S) mass, with slightly wider (50 MeV) signal and sideband regions. The dipion mass distribution, Figure 6.18, follows the same characteristic shape as was observed at CLEO; as well as a concentration at the upper end of the spectrum, there is another near the lower threshold.

Unlike the $m(\pi^+\pi^-)$ distribution for the $\Upsilon(2S) \to \pi^+\pi^-\Upsilon(1S)$ decay, this is not consistent with the multipole expansion model for the dipion transition (see Section 1.6). In the Moxhay model [9], an additional complex constant, B, is added to the amplitude to account for the possibility of $B\bar{B}^*$ and $B^*\bar{B}$ intermediate states. The total decay rate is then written as

$$\frac{d\Gamma}{dm_{\pi\pi}} = |A| \left| m_{\pi\pi}^2 F(m_{\pi\pi}^2) - (B/A) \right|^2 \cdot PS, \qquad (6.4)$$

where $F(m_{\pi\pi}^2)$ is the multipole amplitude, modified by Belanger *et al.* [107] to include $\pi\pi$ $S-$wave final-state interactions. Moxhay found $Re(B/A) = 0.2196$ and $Re(B/A) = -0.2983$ from a fit to an early CLEO dataset, which was later updated by the CLEO collaboration with a larger dataset to $Re(B/A) = 0.097 \pm 0.006$ and $Im(B/A) = 0.284 \pm 0.003$ [71].

Due to some unknown convention, or otherwise, the correct shape could not be reproduced with the CLEO parameters. Instead, a fit of the Moxhay model was performed on the ATLAS Υ(3S) sideband-subtracted data, with best fit values of $Re(B/A) = 0.068 \pm 0.018$ and $Im(B/A) = 0.17 \pm 0.01$. This functional form (red curve in Figure 6.18) was then used as the input to the calculation of the dipion mass weights for the

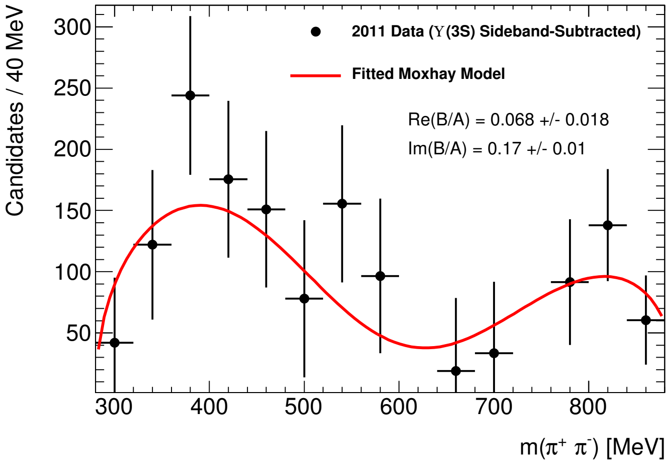


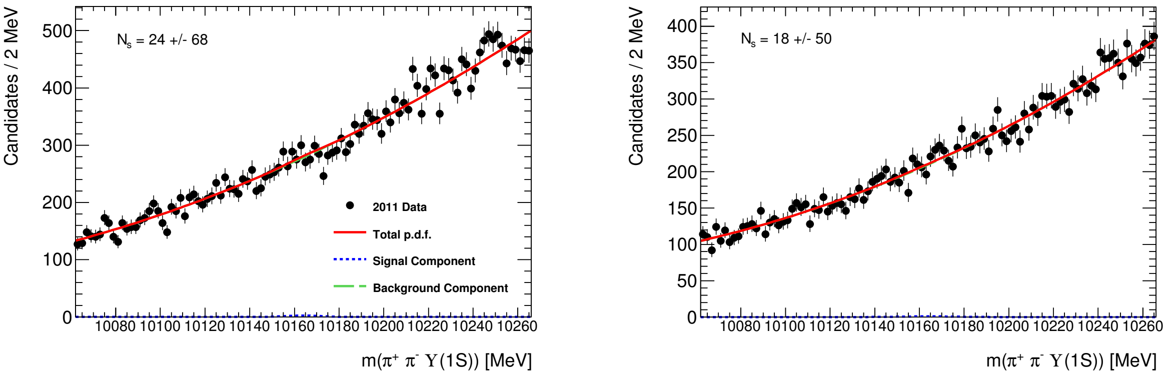
Figure 6.18: The sideband-subtracted Υ(3S) dipion mass distribution. The negative entry in the bin at $m(\pi^+\pi^-) \approx 600$ MeV has been suppressed for clarity.

Υ(3S) sample.

6.6 The $\pi^+\pi^-\Upsilon(1\mathrm{S})$ Mass Spectrum in the $\Upsilon(1^3\mathrm{D}_1)$ Region

Since the $\Upsilon(1^3\mathrm{D}_\mathrm{J})$ triplet has only been observed at e^+e^- colliders [28, 29], there was no evidence as to the strength of pp production to be expected. Furthermore, there appears to be few theoretical predictions on the matter; one example [108] gives a value similar to the Υ(2S) cross section, but is based largely on colour-octet production mechanisms. In the BaBar analysis [29], the feed-down branching fraction $\mathcal{B}(\Upsilon(3\mathrm{S}) \rightarrow \gamma\chi_{b1}(2\mathrm{P}) \rightarrow \gamma\gamma\Upsilon(1^3\mathrm{D}_2) \rightarrow \gamma\gamma\pi^+\pi^-\mu^+\mu^-)$ was measured as 2.5×10^{-7}. Assuming that the emission of the two photons is reasonably soft, the expected yield of $\Upsilon(1^3\mathrm{D}_2)$ can then be approximated in terms of the number of observed Υ(3S). The result for the 2011 data sample is a value less than one, indicating that any observed signal would be almost entirely due to direct pp production.

Both BaBar and CLEO only claim evidence, and have a mass measurement for, the $J = 2$ state of the triplet. As such, the first hypothesis test only includes an individual signal peak fixed at the average measured mass value of $m = 10164$ MeV. A simultaneous fit (Figure 6.19) was carried out in an analogous manner to the Υ(3S) peak, with the omission of the fitted mean correction (since the uncertainty on the mass is larger than this correction), with a resultant significance of 0.4σ. Based on this, there is no evidence for the $\Upsilon(1^3\mathrm{D}_2) \rightarrow \pi^+\pi^-\Upsilon(1\mathrm{S})$ decay in 2011 data.



(a) Barrel (b) Endcap

Figure 6.19: The projections of a simultaneous fit to the $\Upsilon(1^3\mathrm{D_J})$ mass region, with only a single signal term for the $J = 2$ state.

For the purposes of establishing procedures for the 2012 dataset, a second hypothesis was tested whereby a signal term was included for each of the angular momentum states. The σ_n parameters were taken as the averages of the values in the $\Upsilon(1^3\mathrm{D_J})$ samples and the splitting of the states was assumed to be that used in the MC samples (which was motivated by a review of theoretical expectations [101]). The best fit values for the yields of the $J = 2$ and $J = 3$ states were negative, and so the significance was taken to be zero. Similar results were found for various fixed splittings in a range ±5 MeV about their nominal values.

In the case where the splittings are larger than ≈ 10 MeV, each state can be treated individually. This is covered by the significance scan in the next section. Likewise, the upper limit on the production rate for the $\Upsilon(1^3\mathrm{D_J})$ states can also be found there.

6.7 Search for the X_b in the $\pi^+\pi^-\Upsilon(1\mathrm{S})$ Mass Distribution

After stepping through the previous specific mass points, the entire spectrum was inspected for signs of new states. Specifically, a scan of the local significance as a function of mass was performed, as well as an upper limit calculation. As for the $\Upsilon(1^3\mathrm{D_J})$ region, only the candidate selection approach was used and, because the search was later extended to the much larger 2012 dataset (next chapter), no systematic uncertainties were included.

The key input here was again the efficiency ratio, $\varepsilon'/\varepsilon'_{2\mathrm{S}}$, where ε' is the (conditional) analysis efficiency at the mass of interest (see Section 5.5 for more detail). To calculate

ε', the values of the efficiencies in the MC samples are again used to extrapolate or interpolate to the given mass. Due to the corrections to the MC samples discussed above, these efficiencies needed to be recalculated. The dipion mass distribution of a hypothetical new state is not known *a priori*, so only the mean dipion p_T and production spectrum weights were used. For the $\Upsilon(1^3\mathrm{D_J})$ and $X(10233)$ samples, the latter of these was calculated as

$$w(p_T, |y|) = w_{2S}(p_T, |y|) + \left(\frac{M - 10023}{10355 - 10023}\right)(w_{3S}(p_T, |y|) - w_{2S}(p_T, |y|)), \tag{6.5}$$

where M is the mass of the signal decay in the sample, and w_{2S} and w_{3S} are the Υ(2S) and Υ(3S) production spectrum weights, respectively. Since the production weights tend to reduce the efficiency of the MC samples, the expected analysis sensitivity is worse than initially thought.

As in the previous chapter, the expected significances and upper limits are calculated with respect to the relative production ratio to the Υ(2S), R.

6.7.1 Significance vs. Mass

The determination of the local significance as a function of mass proceeded in a similar fashion to that of the expected significance described in the previous chapter. For simplicity, the production of the new state was assumed to be Υ(2S)-like, implying that the value of F_{yL} could be taken from the weighted Υ(2S) MC sample. Then, at each mass, m,

1. The signal model was determined from the MC shape parameters in Section 5.4.
2. The known contributions at the Υ(2S) and Υ(3S), from the fits of the previous sections, were manually inserted into the background model.
3. Simultaneous fits were performed in a range $[m - 8\sigma_{yH}, m + 8\sigma_{yH}]$ across both rapidity bins to extract the significance in the usual manner (σ_{yH} is the narrow width component in the endcap bin).

This was repeated in 5 MeV steps, from 10.1 to 11.1 GeV, to obtain the solid line in Figure 6.20. There is no evidence for any other narrow state decaying to the $\pi^+\pi^-\Upsilon$(1S) final state. The expected significance for $R = 0.0656$ was also calculated and is shown as the dashed line, indication that a significant signal may not actually be expected. With the additional data provided by the 2012 sample, the situation will be much clearer — the significance of genuine signal peaks should scale like $\sqrt{\mathcal{L}_{int}^{2012}/\mathcal{L}_{int}^{2011}} \approx 2$, while the level statistical fluctuations should be largely unchanged.

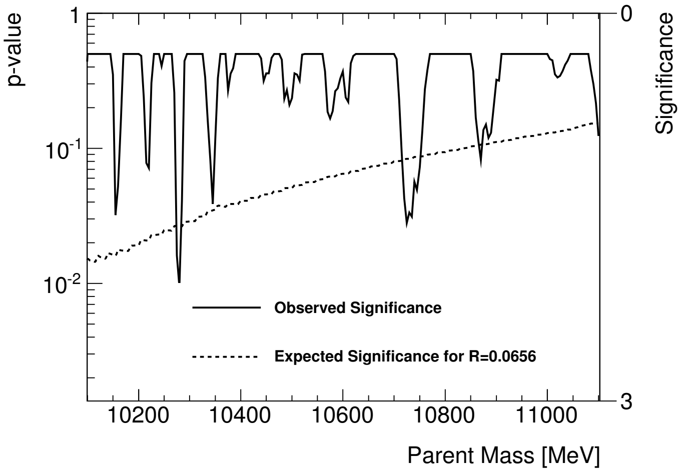


Figure 6.20: The observed local significance as a function of the hypothetical mass of the parent particle. For comparison, the dashed line shows the expected significance for $R = 0.0656$.

Interestingly, the dip at $m \approx 10860$ MeV corresponds to the mass of a broad state which has a well-known $\pi^+\pi^-\Upsilon(1S)$ transition. The prospect of this is explored below.

6.7.2 Upper Limits

At each mass, the same signal model was established as above, with the $\Upsilon(2S)$ and $\Upsilon(3S)$ contributions again built-in to the background term. As in the upper limit expectations from the previous chapter, the $\tilde{q}_\mu$ formalism was used to extract the 95% CLs upper limits from the asymptotic formulas in [104]. Figure 6.21 gives the observed value as well as the expected upper limits and their $\pm 1, 2\sigma$ bands. At this point, a production rate of $R = 0.1$ can be excluded at the 95% confidence level for most of mass range. It is worth keeping in mind that these upper limits will increase when the systematic uncertainties are incorporated.

6.8 The $\Upsilon(10860)$ and $\Upsilon(11020)$

The above formalism was developed for narrow states whose line shape is dominated by the detector resolution. The $\Upsilon(10860)$ and $\Upsilon(11020)$ are both states which have significant branching fractions to $\pi^+\pi^-\Upsilon(1S)$, are presumably produced in *pp* collisions with a similar strength to the other Υ states,[2] but have large natural widths. In order to

---

[2] The cross sections for $\Upsilon(1S)$, $\Upsilon(2S)$, and $\Upsilon(3S)$ production are all of the same order of magnitude [6].

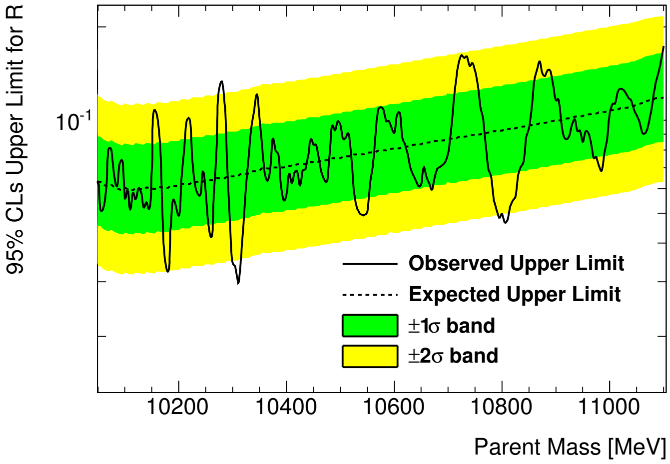


Figure 6.21: The 95% CLs upper limits for the relative production rate as a function of the hypothetical particle mass.

look for these decays, the fitting model was changed such that the signal term was given by the Breit-Wigner line shape[3].

In terms of the relative cross section to the $\Upsilon(2\mathrm{S})$, $\alpha := \sigma/\sigma_{2\mathrm{S}}$, a reasonable estimate of the $\Upsilon(10860)$ is given by

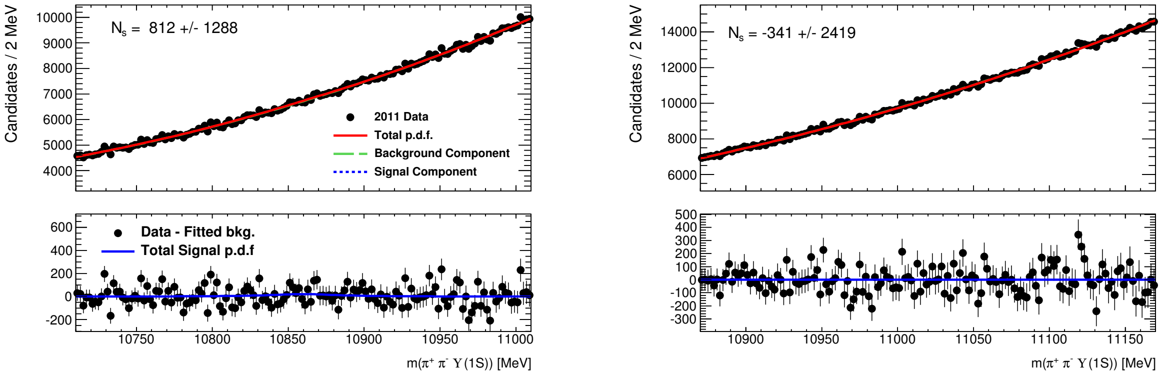
$$N_{\Upsilon(10860)} = 100\alpha. \tag{6.6}$$

In arriving at this, the acceptance and production spectrum was assumed to be the same as the $\Upsilon(2\mathrm{S})$ and the analysis efficiency was determined by the usual extrapolation of the known values. From the decreasing $\Upsilon(n\mathrm{S})$ production rates with $n = 1, 2, 3$, α is expected to be less than one so that the signal should not be significant (given the large background level).

Using the measured natural widths of the $\Upsilon(10860)$ and $\Upsilon(11020)$, fits (Figure 6.22) were performed on the candidate selection sample at the appropriate masses. Because the signal shapes are not a strong function of rapidity in this case, there was no need to split the fit across the two analysis bins. The significance for these fits was found to be 0.6σ and 0 for the $\Upsilon(10860)$ and $\Upsilon(11020)$, respectively, giving no evidence for their decays to $\pi^+\pi^-\Upsilon(1\mathrm{S})$.

---

[3] More accurately, the signal shape should be a Briet-Wigner convolved with a mass-dependent resolution function. For the case of the $\Upsilon(10860)$ and $\Upsilon(11020)$, the uncertainty on the widths are too large to warrant such specificity.



(a) $\Upsilon(10860)$ (b) $\Upsilon(11020)$

Figure 6.22: The fits to the candidate selection mass spectrum at $m = 10860$ MeV and $m = 11020$ MeV, using a Breit-Wigner line shape for the signal and the world average values for the widths [3].

Chapter 7

Search for $X_b \to \pi^+\pi^-\Upsilon(1S)$ Decays: Analysis of the 2012 Data

Towards the end of the analysis of the 2011 dataset, the CMS collaboration released their results for a similar X_b search, exposing an area of unexploited signal/background separation. Consequently, the search methodology used in the previous chapter was reassessed and altered to take advantage of this additional discrimination. The procedure used to determine the optimal way of doing so, described in Section 7.3, found that the best approach was one in which candidate selection was abandoned and the analysis was split into 8 different bins.

The resultant analysis has become the first official ATLAS search for the X_b, to be published in mid-2014. What follows in this chapter is essentially the accompanying internal documentation — "The search for the X_b and other hidden-beauty states in the $\pi^+\pi^-\Upsilon(1S)$ channel at ATLAS" [109] — along with some supplementary material on the optimisation of the analysis. Though very similar to the 2011 analysis (modulo the changes described above), the 2012 analysis is self-contained and the documentation makes little or no reference to the former; the reader is encouraged to compare the results of this chapter with those of the last. For the same reasons, some sections may repeat or summarise aspects of the analysis that have already been explained previously.

As a final remark, the figures in this chapter strictly adhere to the ATLAS style requirements (the aspect ratios, labels etc.) and will appear slightly different to most of those in the previous chapters of this thesis.

7.1 Simulation Samples

This analysis makes use of 8 signal simulation samples, each for a different parent state (and mass) decaying to $\pi^+\pi^-\Upsilon(1S)$:

- The Υ(2S), Υ(3S), the $\Upsilon(1^3\mathrm{D_J})$ triplet and the Υ(10860) — all of which are known to decay through the $\pi^+\pi^-\Upsilon$(1S) channel, and
- two hypothetical X_b states at 10233 MeV and 10561 MeV. The latter corresponds to a particular molecular model for an $X(3872)$ analogue [45].

For all of these, the following were used:

- **Athena version** — 17.2.10.5 (EVNT), 17.2.6.2 (HITS), 17.2.1.4.2 (AOD).
- **Generator** — Pythia 8.170 with AU2 tune [110].
- **Generator PDF tune** — CTEQ6L1 [111] with LO α_s
- **Simulation** — Geant4 + Atlfast II

The production of the parent state was assumed to be Υ(2S)-like as modelled under the Non-Relativistic Quantum Chromodynamics (NRQCD) formalism, with the available phase-space for the ensuing decay into the $\pi^+\pi^-\Upsilon$(1S) sampled uniformly. In particular, the dipion mass-squared $m^2_{\pi^+\pi^-}$ and the decay angles[1] of the Υ(1S) in the parent rest frame were both generated flat. Furthermore, the $\Upsilon(1S) \to \mu^+\mu^-$ decay was also assumed to be isotropic. In the case of the angular distributions, this is well justified by a measurement from CMS [7] indicating that the Υ(1S), Υ(2S), and Υ(3S) are all produced weakly polarised in *pp* collisions, if at all. However, all well-known transitions to $\pi^+\pi^-\Upsilon$(1S) (e.g. from Υ(2S) and Υ(3S) [13]) or the analogous charmonium final state (e.g. $X(3872) \to \pi^+\pi^- J/\psi$ [4]) proceed with distinctive dipion mass distributions which are significantly different to the phase-space-only shape. To account for this, the MC samples were reweighted to reflect this difference where necessary (see below).

To reduce the computational demands in producing these simulated samples, a filter was placed on the kinematics of the muons and pions from the signal decay — the former were required to have $p_T > 4$ GeV and the latter $p_T > 380$ MeV, with all four particles within the physical acceptance of the detector, $|\eta| < 2.5$. For later convenience, the samples were further filtered to ensure that both signal pions had $p_T > 400$ MeV (this reduces the sample size by 5–11% depending on the parent mass). A summary of the MC samples is given in Table 7.1 (the last column will be described below).

---

[1] i.e. the azimuthal angle, ϕ, and the cosine of the polar angle, $\cos\theta^*$.

Table 7.1: Summary of the signal Monte Carlo (MC) samples used for the 2012 analysis. The Size column refers to the number of events generated with the original acceptance conditions on the pions ($|\eta| < 2.5$ and $p_T > 380$ MeV), the Size (full-filter) refers to the number remaining after further-requiring $p_T > 400$ MeV for each pion, and the last column is the total sum of weights for these events (see Section 7.1.1 for details). Section 7.6.5 explains the reasoning for the choice of $\Upsilon(1^3\mathrm{D_J})$ masses.

Parent State	**Mass [MeV]**	**Size**	**Size (full-filter)**	**Weighted+filtered Size**
Υ(2S)	10023	100000	88889.0	55766 ± 189
Υ(2S) ($m_{\pi^+\pi^-}$ weighted)	10023	100000	88889.0	45570 ± 210
$\Upsilon(1^3\mathrm{D}_1)$	10156	50000	45213.0	32559 ± 154
$\Upsilon(1^3\mathrm{D}_2)$	10164	50000	45385.0	32912 ± 155
$\Upsilon(1^3\mathrm{D}_3)$	10170	50000	45461.0	33119 ± 156
X(10233)	10233	100000	91365.0	70067 ± 233
Υ(3S)	10355	100000	91533.0	76853 ± 233
Υ(3S) ($m_{\pi^+\pi^-}$ weighted)	10355	100000	91533.0	83764 ± 341
X(10561)	10561	100000	91778.0	87187 ± 294
Υ(10860)	10860	100000	94476.0	99873 ± 337

7.1.1 Weighting

Two types of weights are used in this analysis — (1) production spectrum weights and (2) dipion mass weights. Additionally, there is the option to reweight the samples to reflect different spin-alignment scenarios, which was used in studying systematic uncertainties.

ATLAS has measured the doubly-differential cross section for Υ(2S) and Υ(3S) production at 7 TeV in two rapidity regions — the *barrel*, $|y| < 1.2$, and the *endcap*, $1.2 < |y| < 2.25$ — in p_T bins up to 70 and 50 GeV, respectively [6]. The same analysis showed that the differential cross section in rapidity is essentially flat up to $|y| = 2.25$, with a slight dip seen at higher rapidity by CMS and LHCb (see [57]). This observation of flatness is used to extend the ATLAS measurement to cover the fiducial region allowed in this search, $|y| < 2.4$ and all p_T:

- In the range $p_T < 50$ GeV, the differential cross section for $2.25 < |y| < 2.4$ is assumed to be equal to the value in the adjacent endcap measurement bin.
- For the bin $50 < p_T < 70$ GeV, the value of the differential cross section in the barrel region is assumed for the entire rapidity range, $|y| < 2.4$.
- For $70 < p_T < 100$ GeV, the differential cross-section is estimated from a CMS measurement [112] in the region $|y| < 0.6$. Once again, this was extended to the full rapidity range assuming flatness in y.

- $p_T > 100$ GeV was assumed to be insignificant (and checked in the MC samples). For the 70–100 GeV bin, the cross section is already 4 orders of magnitude lower than the peak value, and beyond this it is suppressed even further under a power law relationship [112].

The fact that the cross section actually dips slightly (by a factor ~ 0.8) at the edge of the allowed rapidity range will only have a very small effect, because the acceptance for signal decays beyond $|y| = 2.25$ is itself very small. Additionally, only acceptance *ratios* are used for the search, where any difference due to this dip mostly cancels (see Section 7.6.8.1).

The extended, measured cross sections for the Υ(2S) and Υ(3S) described above were compared with the corresponding simulated spectrum of two additionally produced Υ(2S) and Υ(3S) MC samples. In contrast to the signal MC samples described above, these were produced unfiltered and with the 2011 data-taking conditions (in particular, at 7 TeV) to allow a fair comparison. The details for these samples are:

- **Generator** - Pythia 6.4
- **Generator PDF tune** - 2011 ATLAS tune (mcllb) [102].
- **Size** $\sim$ 1 million.

The main observation of this comparison was that the MC spectrum tends to be significantly harder in p_T than the measured cross section, leading to an overestimate in the signal efficiencies. Based on this, a set of weights were derived from the ratio of the measured and generated differential cross sections, with the same binning as the extended measurement described above, to rectify this difference.

Currently, there is no corresponding measurement for Υ(2S) or Υ(3S) production at 8 TeV in the region of phase space relevant to this analysis. If there were, the corresponding 2012 weights (or 8 TeV weights) would be defined as

$$w_{2012} := \left(\frac{\partial^2\sigma}{\partial p_T \partial|y|}\right)_{\text{meas.,2012}} \left(\frac{\partial^2\sigma}{\partial p_T \partial|y|}\right)^{-1}_{\text{mc,2012}} \tag{7.1}$$

in obvious notation. If it is presumed that the NRQCD modelling successfully predicts the *ratio of increase* in the differential cross section between 7 and 8 TeV (a reasonable assumption, in the absence of any measured values), then

$$\begin{aligned} w_{2012} &= \sigma''_{\text{meas.,2012}} \left(\frac{\sigma''_{\text{meas.,2011}}}{\sigma''_{\text{meas.,2012}}} \frac{\sigma''_{\text{mc,2012}}}{\sigma''_{\text{mc.,2011}}} \right) \sigma''^{-1}_{\text{mc,2012}} \\ &= \sigma''_{\text{meas.,2011}} \sigma''^{-1}_{\text{mc,2011}} \\ &=: w_{2011}, \end{aligned} \tag{7.2}$$

where the doubly differential cross section was abbreviated to σ''. In other words, the 7 TeV weights developed above can be applied directly to the 8 TeV MC samples under the assumption mentioned above. Even if this is not strictly true, the change in the shape of the cross section between 7 and 8 TeV is likely to be minor, based on the increase between the Tevatron and LHC energies (see [113]). Given this, the 7 TeV weights were adopted for the $\Upsilon(2\mathrm{S})$ and $\Upsilon(3\mathrm{S})$ signal MC samples, with the other MC samples given linearly extrapolated values (the systematic uncertainty associated with this is evaluated in Section 7.6.8.1) based on the parent mass, m:

$$w_m(p_T, |y|) = w_{2\mathrm{S}}(p_T, |y|) + \frac{m - m_{2\mathrm{S}}}{m_{3\mathrm{S}} - m_{2\mathrm{S}}}(w_{3\mathrm{S}}(p_T, |y|) - w_{2\mathrm{S}}(p_T, |y|)) \tag{7.3}$$

As well as these production spectrum weights, a scale factor is applied to the simulation samples account for data vs. simulation differences in the fraction of $\Upsilon(2\mathrm{S}) \to \pi^+\pi^-\Upsilon(1\mathrm{S})$ signal decays in the barrel region. This is explained in Section 7.6.3.

The dipion mass weights are more straightforward. If the distribution for pure phase-space is PS, then the correct distribution for the $\Upsilon(2\mathrm{S}) \to \pi^+\pi^-\Upsilon(1\mathrm{S})$ transition can be fitted with the Voloshin and Zakharov model [106] as

$$\frac{d\sigma}{dm_{\pi^+\pi^-}} \propto PS \cdot \left(m^2_{\pi^+\pi^-} - \lambda m^2_{\pi}\right)^2, \tag{7.4}$$

where the best fit value for λ with the CLEO data is 3.11 [13]. The dipion mass weights for the $\Upsilon(2\mathrm{S})$ MC sample are then essentially Equation 7.4, divided by the phase-space factor.

The correct $\Upsilon(3\mathrm{S})$ dipion mass distribution is more complicated, but can be successfully fitted with the Moxhay model [9]:

$$\frac{d\Gamma}{dm_{\pi\pi}} = |A| \left|m^2_{\pi\pi} F(m^2_{\pi\pi}) - (B/A)\right|^2 \cdot PS, \tag{7.5}$$

where $F(m^2_{\pi\pi})$ is the multipole amplitude, modified by Belanger *et. al.* [107] to include $\pi\pi$ $S-$wave final-state interactions. Moxhay found $Re(B/A) = 0.2196$ and $Im(B/A) = -0.2983$ from a fit to an early CLEO dataset, which was later updated by the CLEO collaboration with a larger dataset to $Re(B/A) = 0.097 \pm 0.006$ and $Im(B/A) = 0.284 \pm 0.003$ [71].

For unknown reasons, the correct shape could not be reproduced with the CLEO parameters. Instead, a fit of the Moxhay model was performed on the ATLAS $\Upsilon(3\mathrm{S})$ sideband-subtracted 2011 data, with best fit values of $Re(B/A) = 0.068 \pm 0.018$ and $Im(B/A) = 0.17 \pm 0.01$ (see Section 6.5.1). This functional form was then used as the input to the calculation of the dipion mass weights for the $\Upsilon(3\mathrm{S})$ sample.

The dipion mass distribution has a significant impact on the decay acceptance and, to a lesser extent, the reconstruction and trigger efficiencies. In the search for the X_b, the $m_{\pi^+\pi^-}$ shape is assumed to be pure phase-space, reflecting ignorance as to the correct distribution. The Υ(2S) and Υ(3S) distributions are used mainly to allow a sensible comparison with the Υ(2S) and Υ(3S) states observed in the 2012 dataset, and as a tool for studying systematic uncertainties.

The final column of Table 7.1 provides the effective size of the reweighted MC samples.

7.2 Background Model

Rather than construct a background model based on an MC sample for the inclusive-1S production component, and a data-driven (sideband) sample for the non-1S component, a simpler approach was used for the 2012 analysis; the background mass shape and kinematics were assumed to be the same as those of the 2011 dataset. For example, to obtain a sample representing the background in the vicinity of $m = 10561$ MeV, the analysis would be carried out on the 2011 dataset, but with the additional requirement that the masses of the $\pi^+\pi^-\Upsilon$(1S) candidates be in some narrow window about this mass (chosen here as $m \pm 30$ MeV).

In the case where the *EF_2mu6_Upsimumu* trigger was being considered (see below), the *EF_2mu6_DiMu* trigger was used as a proxy as the former was not available during 2011 data-taking. With the trigger specified, the background sample was scaled to the appropriate 2012 luminosity – for the *EF_2mu4T_Upsimumu* (used in the 2011 analysis) and *EF_2mu6_Upsimumu* triggers, the relevant scale factors[2] are 3.53 and 5.92, respectively.

7.3 Changes to the Analysis Methodology

7.3.1 Trigger

Due mainly to the larger instantaneous luminosity in 2012 data-taking, the p_T threshold of the primary trigger responsible for collecting $\Upsilon(1S) \to \mu^+\mu^-$ decays was increased from 4 GeV to 6 GeV and, for obvious reasons, renamed *EF_2mu6_Upsimumu* (referred to as the *2mu6* trigger below). The *EF_2mu4T_Upsimumu* (or *2mu4*) trigger used in the 2011 analysis ran in parallel, but generally with a prescale in operation. Despite this, the difference in total integrated luminosity for the two was mild – 16.2 fb^{-1} for the *2mu4*

---

[2] The ratio between the total integrated luminosity collected in 2011 to that in 2012 for the specific trigger(s).

trigger compared with 20.3 fb^{-1} for the *2mu6*. From this, and MC-based efficiencies, the lower threshold trigger should provide roughly 1.3 times as many $\Upsilon(2\mathrm{S}) \to \pi^+\pi^-\Upsilon(1\mathrm{S})$ decays and so seemingly would be the preferable choice for the 2012 search.

As it turns out, there is an additional source of signal/background separation which was not fully exploited in the 2011 analysis. This becomes stronger at higher p_T, complicating the question of which trigger to use. This extra discrimination is discussed below, including an explanation of how the choice of trigger was arrived at.

7.3.2 Additional discrimination in the $(p_T, \cos\theta^*)$ plane

For $\pi^+\pi^-\Upsilon(1\mathrm{S})$ decays from parent states of mass m, the typical energy available to each of the pions in the rest frame is of the order of $0.5(m - m_{1S})$ (about 280 MeV for the $\Upsilon(2\mathrm{S}) \to \pi^+\pi^-\Upsilon(1\mathrm{S})$, for example). Because of this, the pions will only exceed the 400 MeV p_T reconstruction threshold if they receive a substantial boost. The two most important variables[3] in determining whether this will be the case are

1. $\cos\theta^*$, where θ^* is the angle between the dipion system in the rest frame and the direction of boost to the lab frame. Lower values of θ^*, or $\cos\theta^*$ close to 1, correspond, on average, to larger transverse momenta for the pions in the lab frame.

2. Parent p_T.

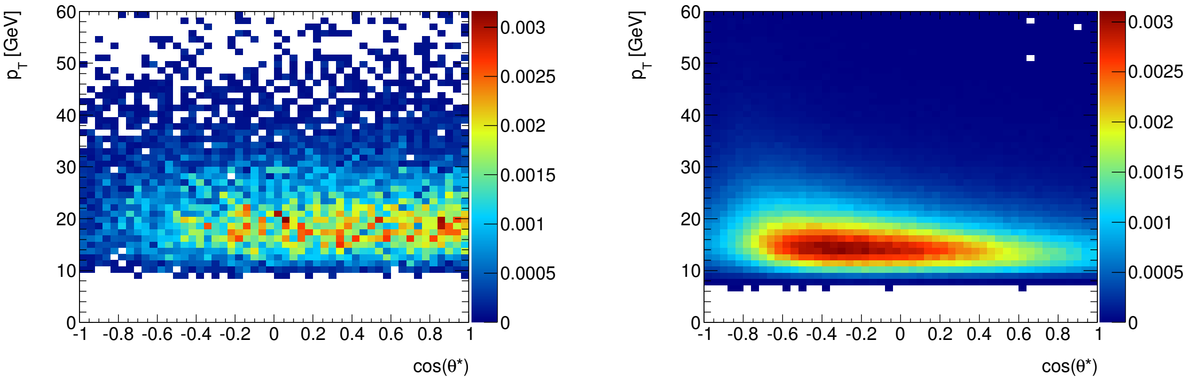
At lower values of p_T, the orientation of the dipion system is critical in determining the acceptance of the decay, whereas for strong enough boosts (i.e. p_T) the acceptance will approach 100% for any value of $\cos\theta^*$. Signal decays reconstructed by this analysis will therefore tend to populate the upper right-hand triangle of the p_T vs. $\cos\theta^*$ plane.

For background candidates, the p_T is still dominated by the dimuon pair but is no longer related to the pion decay angles. In boosting into the rest frame of the $\mu^+\mu^-\pi^+\pi^-$ combination, the dipion trajectory generally tends to be pulled towards values of $\cos\theta^* \sim -1$. The larger the transverse momentum, the more this effect becomes enhanced, meaning background combinations cluster in the lower left triangular region of the p_T vs. $\cos\theta^*$ plane.

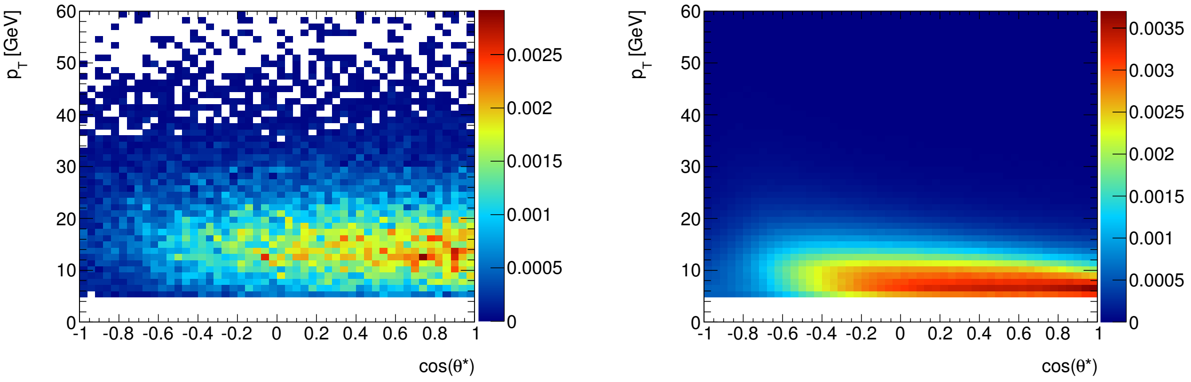
A number of approaches were considered to exploit this difference, assessed in terms of the expected significance they would provide for a weak signal. Because the p_T and $\cos\theta^*$ distributions are mass-dependent, it was convenient to target a specific mass. The $\Upsilon(2\mathrm{S})$ mass was chosen for this purpose in the equivalent CMS analysis [2], but it seemed more appropriate here to do so at a higher mass value, for two main reasons:

---

[3] Others include the (total) rapidity, and the orientation of the pion decay in the dipion rest frame.



(a) Signal, *2mu6* trigger (b) Background, *2mu6* trigger



(c) Signal, *2mu4* trigger (d) Background, *2mu4* trigger

Figure 7.1: The distribution of signal (left) and background (right) candidates in the $(p_T, \cos\theta^*)$ plane at the Swanson mass, $m = 10561$ MeV. The upper plots are for the *2mu6* trigger, and the bottom are for the *2mu4* trigger.

1. Most X_b models predict a mass significantly above the m_{2S}, and
2. the background is already kinematically suppressed in the vicinity of $m = 10023.26$ MeV, but rises strongly as a function of mass.

Taking these into account, the Swanson model [45] mass of $m = 10561$ MeV was chosen as a plausible value in the upper end of the spectrum (close to the $B\overline{B}$ threshold).

The signal/background separation in the $(p_T, \cos\theta^*)$ plane is shown for the Swanson mass in Figure 7.1, with signal decays on the left and background on the right. For the *2mu4* trigger, most of the discrimination is in p_T – the background sitting at the threshold and the signal somewhat above this. To some extent, this difference was already accounted for in the 2011 analysis through the use of the candidate selection technique. For the *2mu6* trigger, on the other hand, the separation is more obvious and it is clear that further discrimination between signal and background is available. The first two options

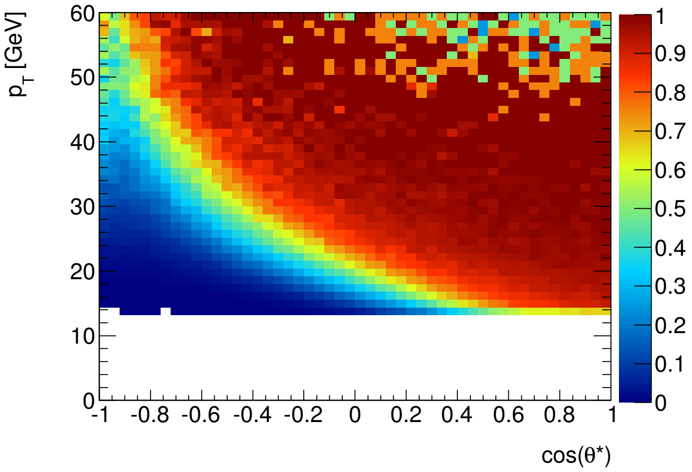


Figure 7.2: The 'imprint' of (i.e. fraction passing) the CMS selections on background candidates near the Swanson mass, $m = 10561$.

discussed below were specifically developed for this higher-threshold trigger, whereas the third is more general and was also considered as an option for the *2mu4* trigger.

7.3.2.1 Option 1: CMS cuts

The first option is a direct copy of that used in the CMS analysis [2], namely:

1. $p_T > 13.5$ GeV
2. $\Delta R(\pi^{\pm}, \Upsilon(1S)) := \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.7$,

where $\Delta\eta$ and $\Delta\phi$ are the angular differences between the $\Upsilon(1S)$ and each of the pions. $\Delta R(\pi^{\pm}, \Upsilon(1S))$ is, more fundamentally, a function of $\cos\theta^*$ and p_T (as well as some other less important variables), and the restriction $\Delta R(\pi^{\pm}, \Upsilon(1S)) < 0.7$ effectively corresponds to selecting a subset of the p_T vs. $\cos\theta^*$ plane. This is demonstrated by Figure 7.2, giving the fraction of background candidates passing the *2mu6* trigger and these CMS cuts, as a function of $\cos\theta^*$ and p_T. The reasoning for this approach is obvious from the discussion above and Figures 7.1 and 7.2; it preferentially selects candidates in the upper right-hand triangle of the $(p_T, \cos\theta^*)$ plane.

The particular values for the selections used by CMS were determined through an optimisation procedure using a genetic algorithm on an $\Upsilon(2S)$ MC sample. Because the CMS analysis used a higher threshold (6.9 GeV on each muon) trigger than the two ATLAS options, the p_T threshold of 13.5 GeV is most probably too high to be optimal in this analysis.

When used in conjunction with candidate selection, the above cuts were applied *post*-selection; i.e. to the $\pi^+\pi^-\Upsilon$(1S) combinations which had already been chosen as the best candidate. The reverse ordering is also a possibility, but because the CMS cuts reduce the mean number of candidates per Υ(1S) from ∼12 to ∼4, much of the original motivation for candidate selection is removed. This is also the case for the diagonal cut method, described below.

7.3.2.2 Option 2: A diagonal cut

At high p_T, the signal and background split into two distinct regions, separated by a diagonal line running from the top left to the bottom right of the p_T vs. $cos\theta^*$ plane. A selection of the form:

$$p_T > -m \cos\theta^* + (p_T^{min} + m), \tag{7.6}$$

where m is the slope and p_T^{min} is the minimum transverse momentum, is a simple way to take advantage of this.

For narrow signal peaks sitting on large backgrounds, the expected significance is naively given by $z = S/\sqrt{B}$, where S and B are the expected signal and background yields, respectively. In the case there are two independent bins, the relevant expression is $z = \sqrt{S_1/B_1^2 + S_2/B_2^2}$. Using this formula, with the barrel and endcaps regions considered as independent bins, the expected significance increase after imposing a diagonal cut (see Figure 7.3) can be easily calculated. The results for candidate selection and all candidates are slightly different, but have large overlapping plateau regions – $m = 10$, $p_T^{min} = 13$ GeV can serve as optimal values for both. This, then, defines the selection used for the *diagonal cut* method:

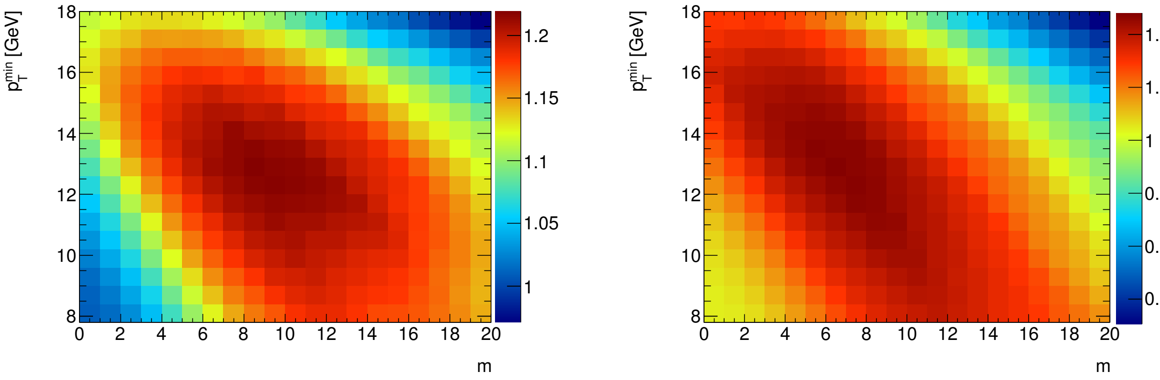
$$p_T > -10\cos\theta^* + 23. \tag{7.7}$$

7.3.2.3 Option 3: Binning method

The third method is different from the above two in that no extra selections are made on the candidates. Rather, a segregation of the candidates into different bins within the p_T vs. $\cos\theta^*$ plane is performed. The motivation for this essentially comes from the n bin generalisation of the naive significance estimate:

$$z = \sqrt{S_1/B_1^2 + S_2/B_2^2 + ... + S_n/B_n^2}. \tag{7.8}$$

Bins with large signal-to-background ratio will dominate the overall significance, but the less sensitive bins still contribute (rather than being removed completely). For simplicity,



(a) All Candidates (b) Candidate Selection

Figure 7.3: The increase in expected statistical significance as a function of the slope, m, and intercept, p_T^{min}, parameters of a diagonal cut through the $(p_T, \cos\theta^*)$ plane. The significance is approximated using a simple formula (see text), and is calculated relative to the baseline (i.e. no diagonal cut – $m = 0, p_T^{min} = 0$) selections.

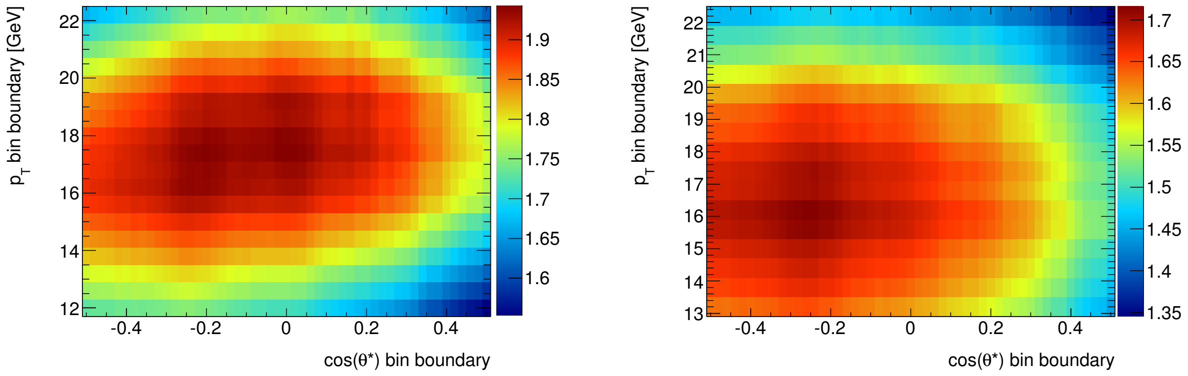
the binning considered was a splitting of the plane into four rectangular quadrants, specified by p_T and $\cos\theta^*$ *bin boundaries*. In combination with the existing split into barrel and endcap rapidities, this gives a total of 8 analysis bins. The choice of the particular values for the bin boundaries was driven by estimating the increase in expected significance as calculated using Equation 7.8.

The expected significance maps, Figure 7.4, show that the binning method is optimised in a fairly wide band across $\cos\theta^*$ for p_T ~15-20 GeV and $p_T \sim 14-18$ GeV for the *2mu6* and *2mu4* triggers, respectively. Zero was selected as the $\cos\theta^*$ boundary because of its simple properties under changes in the presumed spin-alignment. The p_T boundaries were chosen as 18 GeV and 16 GeV for the *2mu6* and *2mu4* triggers, respectively, as they are on the optimisation plateau for both all candidates and candidate selection, and correspond to bin boundaries of the production spectrum weights.

7.3.3 Selecting the optimal analysis approach

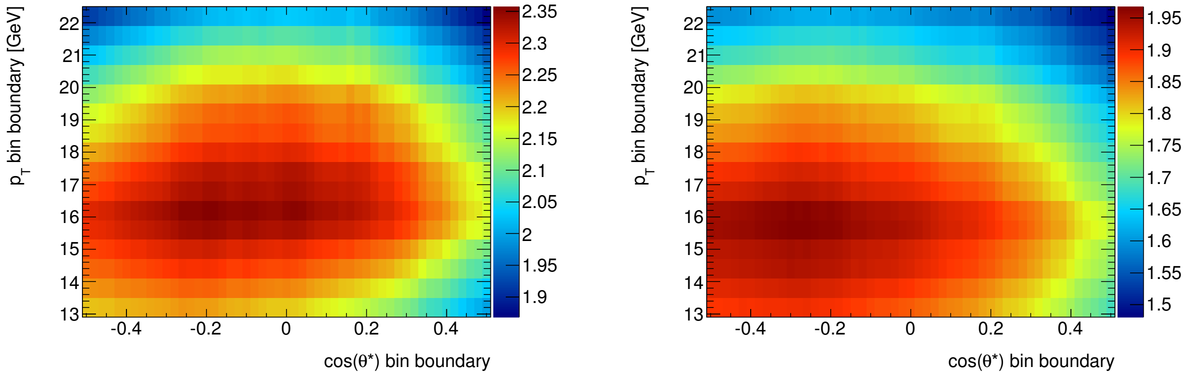
The combination of the trigger and candidate selection choices with the three additional options above gives 8 possible analysis approaches:

- CMS cuts: *2mu6* trigger, with all candidates or candidate selection,
- The diagonal cut: *2mu6* trigger, with all candidates or candidate selection, and
- Binning method: either the *2mu6* or *2mu4* trigger, and either all candidates or candidate selection.



(a) All Candidates, *2mu6* trigger

(b) Candidate Selection, *2mu6* trigger



(c) All Candidates, *2mu4* trigger

(d) Candidate Selection, *2mu4* trigger

Figure 7.4: The increase in expected statistical significance as a function of the bin boundaries for p_T and $\cos\theta^*$ used in the binning method. As for the diagonal cut, the significance is estimated with Equation 7.8 and calculated with respect to the baseline-only selections.

To decide which of these to use, each was run through the full fitting procedure (see Chapter 6) to determine the significance that would be expected for a hypothetical X_b state at $m = 10561$ MeV with a relative production strength of $R = 0.05$. This essentially amounts to injecting a signal peak, with a yield calculated based on the efficiency of the particular analysis approach under study, into the background distribution and performing a simultaneous fit across the relevant number of bins (2 rapidity bins for the CMS/diagonal cut options, or the 8 analysis bins for the binning method). To do so required the following inputs:

- **Signal shape.** The signal was represented by exactly the same model as for the 2011 analysis, with the parameters recalculated using the 2012 samples[4]. The

---

[4] Excluding the $\Upsilon(10860)$, which has a very different shape because of the larger natural width.

signal shapes were still determined separately for the two rapidity regions, but were assumed to be independent of p_T and $\cos\theta^*$. See Section 7.4 for more on this.

- **Signal yield**. The total signal yield was calculated as

  $$N = \mathcal{L} \cdot R \cdot (\sigma\mathcal{B})_{2\mathrm{S}} \cdot \mathcal{A} \cdot \epsilon \tag{7.9}$$

  where:

  - $\mathcal{L}$ is the total integrated luminosity for the trigger under consideration – 20.3 fb^{-1} for the *2mu6* trigger, and 16.2 fb^{-1} for the *2mu4* trigger.
  - $(\sigma\mathcal{B})_{2\mathrm{S}}$ is the production rate for the Υ(2S), with $\sigma_{2\mathrm{S}}$ calculated from the extended, measured cross section described in Section 7.1.1 and $\mathcal{B}_{2\mathrm{S}}$ from the world average values [3].
  - $\mathcal{A}$ is the total acceptance for muons to pass the trigger threshold and pions to be above 400 MeV, with all four particles within $|\eta| < 2.5$. Using the extended, measured cross sections for the Υ(2S) and Υ(3S) states leads to two different values, $\mathcal{A}_{2\mathrm{S}}$ and $\mathcal{A}_{3\mathrm{S}}$, because Υ(3S) production is slightly harder in p_T. Presuming this trend would continue to a hypothetical state at $m = 10561$ MeV, the acceptance was calculated as the linear extrapolation of the two:

    $$\mathcal{A}(10561) = \mathcal{A}_{2\mathrm{S}} + \frac{m - m_{2\mathrm{S}}}{m_{3\mathrm{S}} - m_{2\mathrm{S}}} \left(\mathcal{A}_{3\mathrm{S}} - \mathcal{A}_{2\mathrm{S}}\right). \tag{7.10}$$

    This will be discussed in more detail in Section 7.6.6.2.
  - ϵ is the remaining conditional analysis efficiency. It is calculated as the ratio between the number of signal decays recovered in the particular analysis being studied and the effective sample size of the MC.

- **Bin splitting** The splitting of the signal yield into each of the analysis bins (barrel/endcap, or for the binning method all 8) was taken as that in the $X_b(10561)$ MC sample.

Simultaneous, binned, extended maximum likelihood fits were then performed in a range $m \in [10561 - 8\sigma_{yH}, 10561 + 8\sigma_{yH}]$, where σ_{yH} is the narrow component width for the endcap rapidity bin, with the expected significance estimated using asymptotic formulas [104]. The results are summarised in Table 7.2.

As expected from the previous chapter, candidate selection provides an advantage when only the baseline (i.e. none of the additional discrimination options) selections are used. However, this ceases to be the case when using any of the CMS, diagonal cut or

Table 7.2: The expected significance for a weak signal at $m = 10561$ MeV with a relative production rate of $R = 0.05$.

$\mathbf{p_T}$,$\cos\theta^*$ **discrimination method**	**Trigger**	**All Candidates**	**Candidate Selection**
2011 baseline selections	*2mu4*	2.4	3.3
CMS cuts	*2mu6*	4.9	4.1
Diagonal cut	*2mu6*	4.9	5.0
Binning	*2mu6*	5.6	5.3
Binning	*2mu4*	5.6	5.5

binning options, where all candidates is generally preferable. Of the additional analysis options, the binning method appears to provide the greatest expected significance – in particular, based on these results it out-performs the CMS cuts by a considerable margin.

The particular values for the p_T and $\cos\theta^*$ bin boundaries of the binning method were driven by the approximate formula for expected significance, Equation 7.8. This makes a number of assumptions – in particular, it does not account for either the signal or background shapes – and should only be used as a guide. To find the true optimal choices, the procedure above was repeated for a number of different bin boundary values. With the $\cos\theta^*$ boundary fixed at zero, the p_T boundary giving the highest significance was determined as 20 GeV for both of the trigger options (see Table 7.3). Then, with the p_T boundary fixed at this value, the $\cos\theta^*$ boundary was shifted to the left and right to check the stability around zero. In fact, a small increase in the expected significance was found for a slightly positive value. Despite this, zero was still thought to be a more sensible choice for the following reasons:

- as mentioned above, $\cos\theta^* = 0$ has well-understood properties under changes in the decay angular distributions, which play an important role in determining the spin-alignment uncertainty,
- small changes in the $|y|$ scaling factor (i.e. within its statistical uncertainty) can cause a change in the outcome, making negative $\cos\theta^*$ bin boundaries more optimal,
- the $\cos\theta^*$ distribution of the background is mass-dependent – strongly clustered near -1 for higher masses and more widespread at lower masses – whereas the signal distribution changes far more slowly. The optimal boundary in $\cos\theta^*$ is therefore likely to shift from positive to negative values in moving from lower to higher masses. In this sense, $\cos\theta^* = 0$ represents a compromise that is best for the significance across the mass spectrum as a whole.

Table 7.3: The expected significance for a number of different bin boundary choices for the binning method.

$\cos\theta^*$ bin boundary	p_T bin boundary [GeV]	*2mu4* trigger	*2mu6* trigger
0	16	5.6	-
0	18	5.9	5.6
0	20	5.9	5.7
0	22	5.6	5.6
-0.2	20	5.7	5.4
0.2	20	6.0	5.8

- the increase in expected significance between zero and 0.2 is mild.

Table 7.3 also demonstrates conclusively that the *2mu4* trigger is more optimal than the *2mu6* in general for the binning approach. The former is also more convenient as it allows for a more continuous extension of the 2011 to the 2012 analysis.

In summary, the 2012 analysis makes use of the *2mu4* trigger, accepting all candidates and using binning in $|y|$, p_T and $\cos\theta^*$ for all hypothesis tests and mass fits (unless otherwise stated). The bin boundaries are placed at $|y| = 1.2$, $p_T = 20$ GeV and $\cos\theta^* = 0$, giving 8 bins altogether. In the remainder of this chapter, these bins will be referred to as yL_ptL_ctsL, yL_ptL_ctsH etc., where the L and H refer to 'low' and 'high' regions, respectively, of the preceding variable. For example, yL_ptL_ctsH is the bin with $|y| < 1.2$, $p_T < 20$ GeV and $\cos\theta^* > 0$. The reconstruction of $\Upsilon(1S) \to \mu^+\mu^-$ and $\pi^+\pi^-\Upsilon(1S)$ candidates and the selections made on these composite objects, as well as the muons and pions, is otherwise unchanged with respect to those used in the previous chapter.

7.4 Fitting Procedure

Signal peaks in the $\pi^+\pi^-\Upsilon(1S)$ mass spectrum are fitted with a double Gaussian shape:

$$f(m) = N_s \cdot \left(f_n G_n\left(\overline{m}, \sigma\right) + \left(1 - f_n\right) G_w\left(\overline{m}, r\sigma\right)\right), \tag{7.11}$$

where

- N_s is the signal yield,
- f_n is the *narrow fraction* of the signal curve,
- $G(a, b)$ refers to the Gaussian function with mean a and width b.
- $\overline{m}$ is the fitted mass of the peak,

- σ is the width of the narrow component of the peak, and
- r is the ratio of widths between the broad and narrow components (i.e. $r > 1$ by construction).

The narrow component of the signal peak reflects the inherent detector resolution, whereas the wide component is used to account for signal decays that are, in some way, broken (e.g. tracks with missing hits, effects of detector material such as multiple scattering, and so on).

From studying the MC samples,[5] the narrow fraction, f_n, and the ratio of widths, r, are both independent of mass.[6] These parameters were fixed as follows:

1. f_n was taken as the weighted average value over fits to each of the MC samples, then
2. r was calculated as the weighted average over fits performed with f_n fixed as above.

The reduced χ^2 values of the fits before and after fixing these parameters were found to be stable in each of the MC samples. The remaining signal shape parameter, σ, is assumed to have a linear dependence on mass, $\sigma(m) = \alpha + m\beta$. The parameters α and β were extracted by performing a fit over the σ values in the MC samples, allowing an extrapolation/interpolation of the signal shape to any potential X_b mass (i.e. not just those for which there is an MC sample). The signal shape parameters, including the values of σ for the Υ(2S), are summarised in Table 7.4.

The fits are performed separately in the rapidity ranges $|y| < 1.2$ and $1.2 < |y| < 2.4$ due to the resolution difference of the Inner Detector between the barrel and endcap regions. However, the signal shape was not expected to be strongly dependent on either p_T or $\cos\theta^*$. To check this, the reduced χ^2 for the Υ(2S) signal fits in the analysis sub-bins of each rapidity region were checked before and after fixing the r and f_n parameters in the procedure above; in all cases, the fit quality was found to be stable.

The background shape in the vicinity of a signal peak was described by a second-order Chebychev polynomial, with the yield and value of the two parameters allowed to be independent in each of the analysis bins. In cases where the fitting range is in the vicinity of $m_{2\mathrm{S}}$ or $m_{3\mathrm{S}}$, contributions from the Υ(2S) and Υ(3S) signal peaks are also included in the background model (see Section 7.6.6.3 for more detail).

---

[5] For the purposes of fitting, the un-weighted MC samples are used to avoid issues associated with likelihood fits on weighted datasets. The weights produce little change in the rapidity distribution within each region (barrel/endcap), and so will only have a minor effect on the signal shapes.

[6] In that the fitted values for these parameters were consistent across each of the MC samples

Table 7.4: Summary of the signal shape parameters extracted from the MC samples.

	r	f	α	β	$\sigma_{2\mathrm{S}}$ [MeV]
Barrel	2.26 ± 0.03	0.872 ± 0.005	-132 ± 2	0.0137 ± 0.0002	5.66 ± 0.04
Endcap	2.11 ± 0.03	0.823 ± 0.009	-204 ± 4	0.0212 ± 0.0004	9.37 ± 0.09

Unless otherwise stated, all fits were performed using the binned, extended maximum-likelihood method in a range $[m - 8\sigma_{yH}, m + 8\sigma_{yH}]$, where σ_{yH} is the signal width in the endcap bin for the mass of interest, m. The range varies from $m \pm 72$ MeV at 10 GeV to $m \pm 242$ MeV at 11 GeV. In most cases, the fit will also have been performed simultaneously across each of the 8 analysis bins, with the relative signal yields in each bin fixed to the expected splitting from the MC samples (or an extrapolation thereof — see Section 7.6.6).

7.5 2012 Dataset

Events passing the *EF_2mu4T_Upsimumu* trigger chain were collected by the *Muons* stream during Period A, and the (delayed) *BPhysics* stream for the remainder of the year's data-taking. After applying the GRL

`data12_8TeV.periodAllYear_DetStatus-v61-pro14-02_DQDefects-00-01-00_PHYS_StandardGRL_All_Good`

the total integrated luminosity for the trigger was 16.189 fb^{-1} with the nominal 1.8% uncertainty [105].

The $\pi^+\pi^-\Upsilon(1S)$ n-tuples containers period-by-period are:

`group.phys-beauty.data12_8TeV.periodA.physics_Muons.PhysCont.AOD.repro14_v02.NTUP_UPP.v1/`

`group.phys-beauty.data12_8TeV.periodB.physics_Bphysics.PhysCont.DAOD_UPSIMUMU.grp14_v03_p1425.NTUP_UPP.v1/`

`group.phys-beauty.data12_8TeV.periodC.physics_Bphysics.PhysCont.DAOD_UPSIMUMU.grp14_v04_p1425.NTUP_UPP.v1/`

`group.phys-beauty.data12_8TeV.periodD.physics_Bphysics.PhysCont.DAOD_UPSIMUMU.grp14_v04_p1425.NTUP_UPP.v1/`

`group.phys-beauty.data12_8TeV.periodE.physics_Bphysics.PhysCont.DAOD_UPSIMUMU.grp14_v03_p1425.NTUP_UPP.v1/`

`group.phys-beauty.data12_8TeV.periodG.physics_Bphysics.PhysCont.DAOD_UPSIMUMU.grp14_v03_p1425.NTUP_UPP.v1/`

`group.phys-beauty.data12_8TeV.periodHpatched.physics_Bphysics.PhysCont.DAOD_UPSIMUMU.grp14_v04_p1425.NTUP_UPP.v1/`

`group.phys-beauty.data12_8TeV.periodI.physics_Bphysics.PhysCont.DAOD_UPSIMUMU.grp14_v03_p1425.NTUP_UPP.v1/`

`group.phys-beauty.data12_8TeV.periodJ.physics_Bphysics.PhysCont.DAOD_UPSIMUMU.grp14_v03_p1425.NTUP_UPP.v1/`

`group.phys-beauty.data12_8TeV.periodL.physics_Bphysics.PhysCont.DAOD_UPSIMUMU.grp14_v04_p1425.NTUP_UPP.v1/`

Each of these has a corresponding *.NTUP_UPP.merge.v1 merged containter, which are themselves pooled into the *AllYear* container:

`group.phys-beauty.data12_8TeV.AllYear.NTUP_UPP.AllBphysStream.patchedPeriodH.merge.v1/`

7.6 Results

Initially, only the regions where an existing state is expected (i.e. the Υ(2S) and Υ(3S) mass regions) were used to validate the MC samples and analysis procedures, in terms of

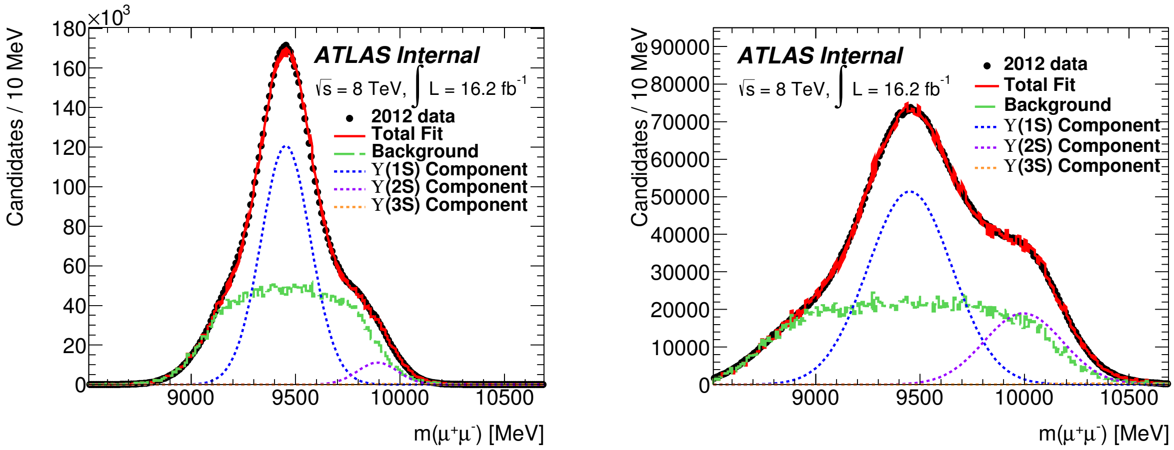
- the signal shape — in particular, the σ parameters and the fitted masses (to check for potential data/MC differences),
- the signal splitting into the 8 analysis bins, and
- the calculation of the analysis efficiency, in turn leading to predicted yields for the Υ(2S) and Υ(3S) peaks.

Additionally, a small peak is expected at $m_{3S} - m_{2S} + m_{1S} \sim 9792$ MeV due to the decays $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(2S)$ and $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(2S)(\to \pi^+\pi^-\Upsilon(1S))$ (see below for an explanation of why these result in peaks at 9792 MeV).

7.6.1 Υ(1S) Yield

Before looking at the $\pi^+\pi^-\Upsilon(1S)$ mass spectrum, the mass distribution for the initial $\mu^+\mu^-$ vertex fit, Figure 7.5, was inspected to extract the total number of $\Upsilon(1S) \to \mu^+\mu^-$ decays used in this analysis. The allowed mass range here was expanded to 8.5–10.7 GeV to show the shape of the distribution outside the nominal selection window. Every $\Upsilon(1S) \to \mu^+\mu^-$ candidate in Figure 7.5 must participate in at least one $\pi^+\pi^-\Upsilon(1S)$ vertex fit, which are conducted under the constraint that the dimuon mass is 9460.30 MeV and restricted to a maximum χ^2 of 20. This implies that those combinations whose mass naturally lies far from the mass constraint will be suppressed because their χ^2 value will generally be too high.

A standard fitting approach (e.g. see [6]) for the Υ(1S), Υ(2S), and Υ(3S) peaks of the dimuon spectrum is to use a double Gaussian shape, with the mass splitting fixed to the world average [3] and the width parameters forced to be proportional to mass. In the present case, an additional multiplicative *penalty* Gaussian with a mean of 9460.3 MeV is used to account for the suppression away from m_{1S} described above.



(a) Barrel, $|y| < 1.2$

(b) Endcap, $1.2 < |y| < 2.5$

Figure 7.5: The fitted $\mu^+\mu^-$ mass distribution in the barrel (left), $|y_{\mu\mu}| < 1.2$, and endcap (right), $1.2 < |y_{\mu\mu}| < 2.4$.

For the background shape, which is usually modelled with a polynomial, the necessary alteration is not as straightforward. Instead, the polynomial was replaced with the same-sign shape as a template, with a floating normalisation.

Even with these measures in place, the fit fails to converge to a sensible distribution. To guide it, the fit is first performed on the distribution obtained when the χ^2 restriction is loosened to 100. In this fit, the relative yields of the narrow and broad components are constrained to be the same for each of the nS states. The masses and widths are then fixed to the resultant best-fit values, and the fit is repeated on the $\chi^2 < 20$ distribution (see Figure 7.5).

Integrating the fitted functions in the allowed range, $9110.3 < m(\mu^+\mu^-) < 9810.3$ MeV, the Υ(1S), Υ(2S), and combinatorial yields are extracted and given in Table 7.5. Only the statistical uncertainties are quoted, which, in particular, do not account for the significant uncertainty present in the same-sign template. A high level of accuracy in these yields is not required here, and so these uncertainties are ignored; more important is that the shape of the dimuon mass spectrum is well-understood, and that ~ 6 million Υ(1S) decays are used in this search.

7.6.2 The Structure at 9792 MeV

Besides the hadronic transitions $\Upsilon(2S) \to \pi^+\pi^-\Upsilon(1S)$ and $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)$, there are more complex cascade decays. In particular, the Υ(3S) can decay to $\pi^+\pi^-\Upsilon(2S)$ followed by $\Upsilon(2S) \to \mu^+\mu^-$ or $\Upsilon(2S) \to X\Upsilon(1S)(\to \mu^+\mu^-)$.

Table 7.5: The yields for the significant contributions to the $m(\mu^+\mu^-)$ spectrum for the 2012 dataset.

	Barrel Yield ($\times 10^6$)	Endcap Yield ($\times 10^6$)	Total Yield ($\times 10^6$)
Υ(1S)	3.611 ± 0.004	2.386 ± 0.010	5.998 ± 0.011
Υ(2S)	0.0537 ± 0.0003	0.146 ± 0.002	0.200 ± 0.002
Combinatorial	3.109 ± 0.004	1.479 ± 0.002	4.588 ± 0.005

If, in the former case, the reconstructed mass of the dimuon pair falls within the allowed range for this analysis ($9110.3 < m(\mu^+\mu^-) < 9810.3$ MeV), the $\mu^+\mu^-\pi^+\pi^-$ combination will be included in the $\pi^+\pi^-\Upsilon$(1S) mass spectrum. Because of the mass constraint on the dimuon pair in the four track vertex fit, the peak for these decays will appear at $m_{3S} - m_{2S} + m_{1S} = 9792.24$ [3], rather than m_{3S}. In the second cascade, there is a possibility that the two muons from the $\Upsilon(1\mathrm{S}) \to \mu^+\mu^-$ decay may be combined with the two pions from the initial stage of the decay. The small mass difference between the Υ(3S) and Υ(2S) leads to a special kinematic scenario where the mass spectrum of such combinations forms a narrow ($\sim$ 2 MeV) peak, also at $m_{3S} - m_{2S} + m_{1S}$.[7]

Consequently, a small peak is expected in the 2012 dataset at ~9792 MeV. This is confirmed by fits in a range ±30 around 9792 MeV, shown in Figure 7.6, performed separately in the barrel and endcap regions with the mass and σ parameters allowed to float (in these, and other similar plots, the bottom plots are the observed mass spectrum with the fitted background subtracted). The local significance for the peak in the barrel and endcap regions was $z = 6.0$ and $z = 2.8$, respectively.

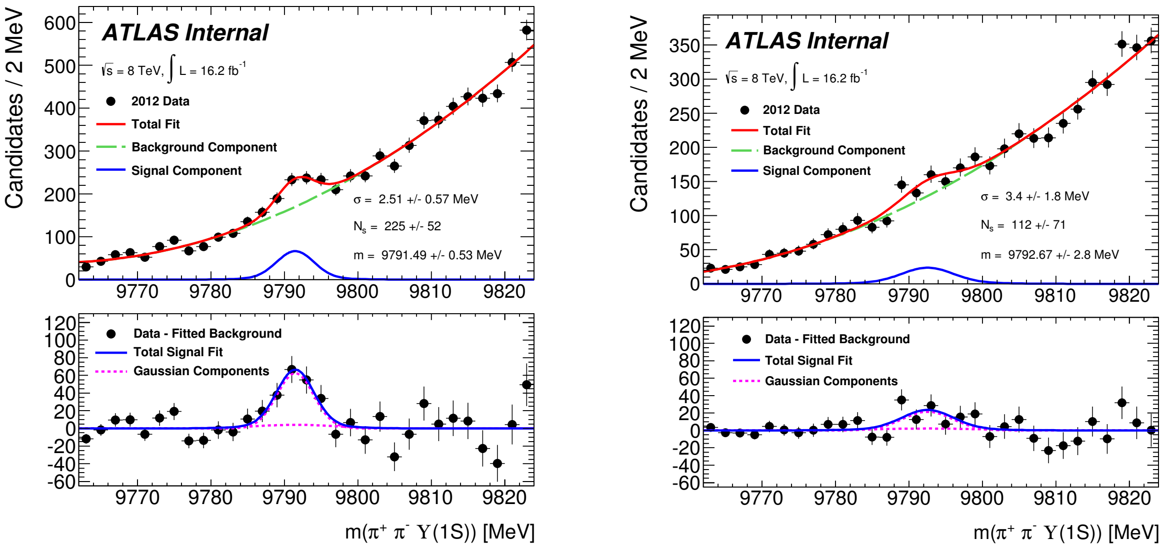
7.6.3 Analysis of Υ(2S) Mass Region

The first check of the Υ(2S) peak was to establish that the peak had a mass, m, close to the world average [3] and σ parameters consistent with those from the MC samples. To do so, fits were performed to the mass spectrum in the vicinity of $m = 10023.26$ MeV, with the above parameters allowed to float (f_n and r were fixed to their average values in Table 7.4). The results, shown in Figure 7.7, confirm that this is indeed the case for both the barrel and endcap rapidity regions.

With this established, focus was turned to validating the relative number of signal decays in each of the analysis bins (i.e. the *bin splitting*), starting with the split into barrel and endcap regions. With the m and σ parameters now fixed to their world average and MC values, respectively, the fits were repeated separately in each of the 8 analysis bins

---

[7] This can be easily shown with a simple toy MC.



(a) Barrel, $|y| < 1.2$ (b) Endcap, $1.2 < |y| < 2.4$

Figure 7.6: The fitted $\pi^+\pi^-\Upsilon(1\mathrm{S})$ mass distribution in the region close to 9792 MeV in the barrel (left) and endcap (right) regions. In both cases, the full p_T and $\cos\theta^*$ range were used.

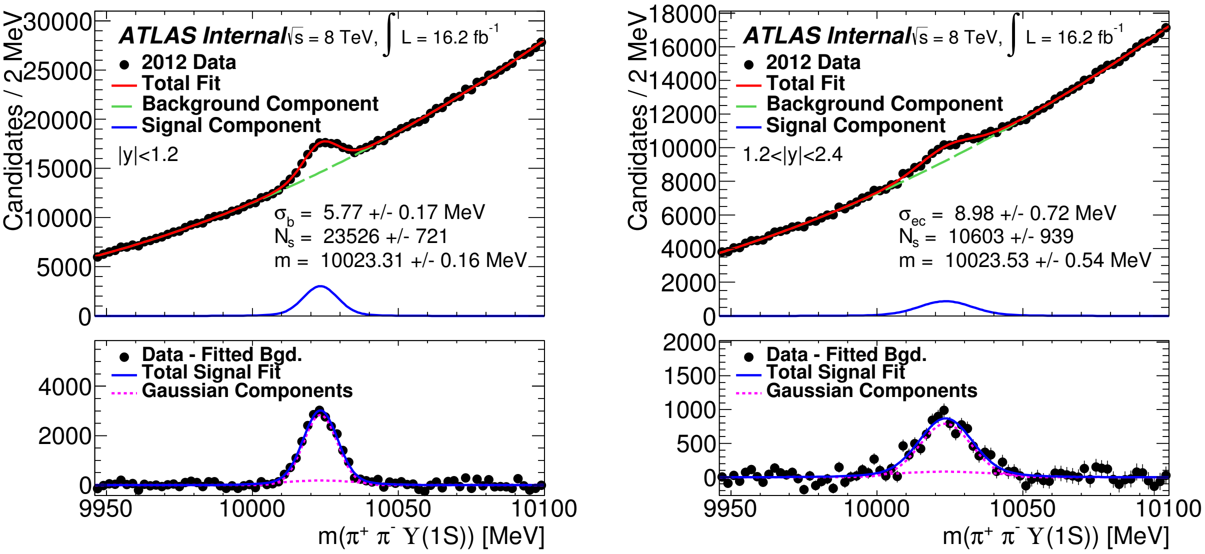
(Figure 7.8). Based on these, the fraction of $\Upsilon(2\mathrm{S})$ signal decays in the barrel is given by $f_b = 0.67 \pm 0.04$, in some tension with the MC value of 0.606 ± 0.004 (cf. the 2011 data and MC values of 0.66 ± 0.03 and 0.597 ± 0.004, respectively). Possible reasons for this include (1) variation of the differential cross section within the barrel and endcap bins, and (2) differences in reconstruction and trigger efficiencies between data and MC. Because the barrel/endcap splitting is mass-independent (see 7.6.6.1), the MC can be forced to be consistent with the data in terms of this fraction by introducing scale factors:

$$S(|y| < 1.2) = 0.67/0.606 = 1.11 \tag{7.12}$$

$$S(|y| < 2.4) = (1 - 0.67)/(1 - 0.606) = 0.838, \tag{7.13}$$

which are universally applied to all the signal MC samples.

With these scaling factors now applied, the relative yields in the 8 bins, from the fits in Figure 7.7, were compared with the corresponding MC splitting values (see Table 7.6). The χ^2/n.d.o.f. value corresponds to a p-value of 0.07, so there is no evidence for inconsistency in this comparison. The main cause for the slightly high reduced-χ^2 value is the first bin, which (see Figure 7.8c) is one with a small signal peak sitting on a large background. As such, the signal yield extracted from the fit to this bin can be sensitive



(a) Barrel, $|y| < 1.2$ (b) Endcap, $1.2 < |y| < 2.4$

Figure 7.7: The fitted $\pi^+\pi^-\Upsilon(1S)$ mass distribution in the region close to m_{2S} in the barrel (left) and endcap (right) regions integrated over the full p_T and $\cos\theta^*$ range. The fitted parameters for m and σ are consistent with the world average [3] and MC samples, respectively.

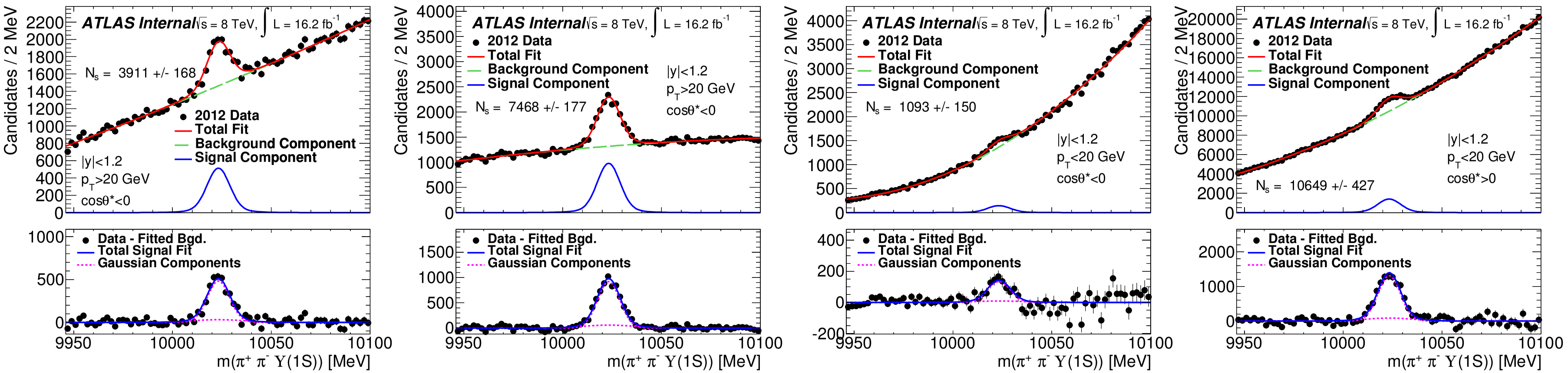
Table 7.6: A comparison between the bin fractions for the $\Upsilon(2S) \to \pi^+\pi^-\Upsilon(1S)$ decay, based on the $\Upsilon(2S)$ MC and the measured values from the 2012 dataset.

Bin	MC Fraction	+/-	Data Fraction	+/-	z-score
yL_ptL_ctsL	0.043	0.002	0.032	0.004	+2.37
yL_ptL_ctsH	0.311	0.004	0.311	0.011	+0.07
yL_ptH_ctsL	0.104	0.003	0.114	0.005	−1.68
yL_ptH_ctsH	0.211	0.003	0.218	0.006	−0.95
yH_ptL_ctsL	0.022	0.001	0.013	0.006	+1.40
yH_ptL_ctsH	0.133	0.003	0.145	0.013	−0.93
yH_ptH_ctsL	0.058	0.002	0.061	0.006	−0.54
yH_ptH_ctsH	0.118	0.003	0.106	0.006	+1.76
			χ^2/n.d.o.f.		1.95

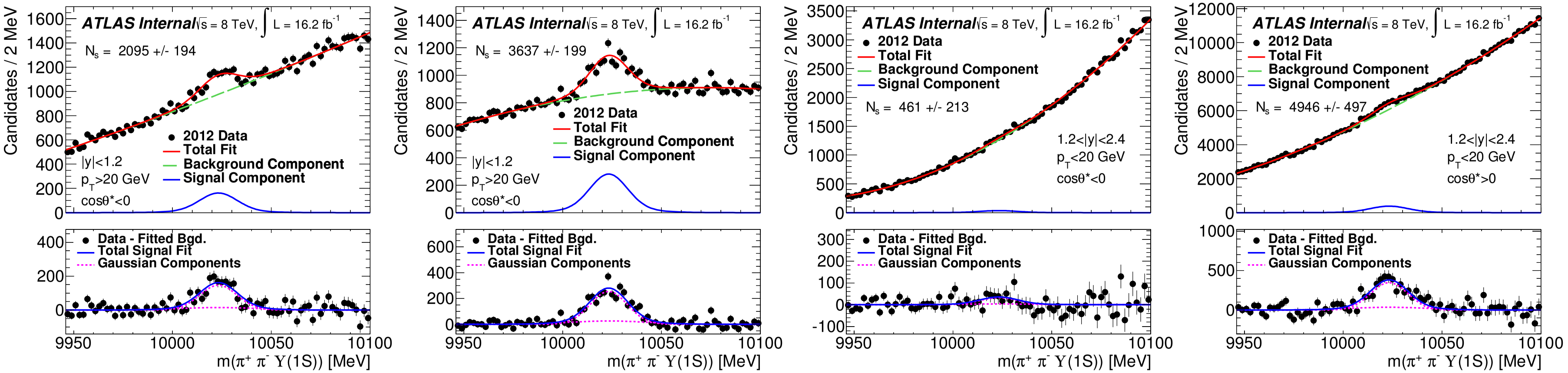
to background modelling. As a check, the fits to each of the bins was repeated with a 3^{rd}-order Chebychev polynomial (rather than the standard 2^{nd}-order), with a resultant χ^2/n.d.o.f. of 1.00.

The last cross-check for the Υ(2S) signal region was the fitted yield of $N_{\Upsilon(2S)\to\pi^+\pi^-\Upsilon(1S)} = 34378 \pm 795$. This is consistent with a prediction of 33300 ± 2500, calculated as:

$$N = \mathcal{L} \cdot \sigma_{2S} \cdot \mathcal{B}(\Upsilon(2S) \to \mu^+\mu^-) \cdot \frac{\mathcal{B}(\Upsilon(2S) \to \pi^+\pi^-\Upsilon(1S)(\to \mu^+\mu^-))}{\mathcal{B}(\Upsilon(2S) \to \mu^+\mu^-)} \cdot \mathcal{A} \cdot \epsilon, \tag{7.14}$$



(a) Barrel, high p_{T}, low $\cos\theta^{*}$ (b) Barrel, high p_{T}, high $\cos\theta^{*}$ (c) Barrel, low p_{T}, low $\cos\theta^{*}$ (d) Barrel, low p_{T}, high $\cos\theta^{*}$



(e) Endcap, high p_{T}, low $\cos\theta^{*}$ (f) Endcap, high p_{T}, high $\cos\theta^{*}$ (g) Endcap, low p_{T}, low $\cos\theta^{*}$ (h) Endcap, low p_{T}, high $\cos\theta^{*}$

Figure 7.8: Individual fits to each of the analysis bins in the $\Upsilon(2S)$ region, with the signal mass fixed to the world average value and simulation-based values used for the signal shape parameters.

where

- $\mathcal{L} = 16.2 \pm 0.3$ fb^{-1} is the total integrated luminosity for 2012,
- $\sigma_{2S} \cdot \mathcal{B}(\Upsilon(2\text{S}) \to \mu^+\mu^-) = 2.20 \pm 0.16$ fb is estimated from the extended differential cross section measurement for pp collisions at 7 TeV.
- $\mathcal{B}(\Upsilon(2\text{S}) \to \pi^+\pi^-\Upsilon(1\text{S})(\to \mu^+\mu^-))/\mathcal{B}(\Upsilon(2\text{S}) \to \mu^+\mu^-) = 0.229 \pm 0.004$ is determined from world-average values [3],
- $\mathcal{A} = 0.01442 \pm 0.00004$ is the *acceptance*, defined as the fraction of $\Upsilon(2\text{S})$ produced which, if they were to decay to the $\pi^+\pi^-\Upsilon(1\text{S})$ final state, would have both muons with $p_T > 4$ GeV, both pions with $p_T > 400$ MeV, and all four particles within the physical extent of the Inner Detector, $|\eta| < 2.5$. This was calculated using a toy MC approach[8] based on the extended differential cross section measurement assuming isotropic angular distributions and the correct $\Upsilon(2\text{S}) \to \pi^+\pi^-\Upsilon(1\text{S})$ dipion mass shape (see Appendix 7.A).
- $\epsilon = 0.283 \pm 0.002$ is the residual (conditional) analysis efficiency — i.e. the probability for the signal decay to be recovered by the analysis given that the muons and pions pass the acceptance conditions.

The uncertainties on $\mathcal{A}$ and ϵ above are statistical only (their systematic uncertainties are assessed in Section 7.6.8.1). Based on this prediction, the increase in the $\Upsilon(2\text{S})$ production between 7 and 8 TeV pp collisions is estimated at the 3% level (assuming that the differential cross section shape changes insignificantly).

7.6.4 Analysis of $\Upsilon(3\text{S})$ Mass Regions

In the case of the $\Upsilon(3\text{S}) \to \pi^+\pi^-\Upsilon(1\text{S})$, the production rate is roughly 10% of that for the $\Upsilon(2\text{S})$, it lies in a region with a larger background, and the resolution is significantly worse. Consequently, there is not enough statistical power to perform all of the same checks as were done in the previous section. Instead, the signal shape parameters and bin splittings were *assumed* to be correctly represented by the $\Upsilon(3\text{S})$ MC sample, and were used as the input to a full simultaneous fit across all the analysis bins (in contrast to the fits of Figure 7.8, which were performed individually in each bin). The reduced-χ^2 of the fit is 1.0, indicating that it is not strained by forcing the bin splitting to the MC values, and the significance is $z = 8.8$ (cf. the significance of the highest bin is 6.5). The

---

[8] Very similar to the *OniaGun* algorithm.

best-fit value for the $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)$ yield is 11596 ± 1339, in excellent agreement with the predicted value (made in the same manner as for the $\Upsilon(2S)$) of 11400 ± 1500.

For visual clarity, a further fit was performed at the $\Upsilon(3S)$ mass in the most sensitive bin, yL_ptH_ctsH, with larger binning to emphasise the signal peak (Figure 7.10).

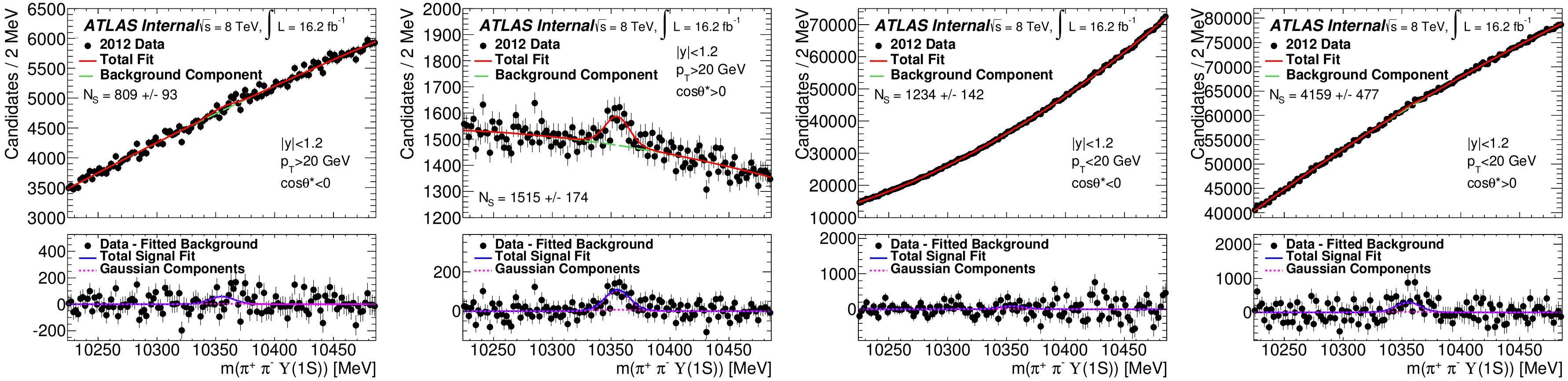
7.6.5 Analysis of the $\Upsilon(1^3D_J)$ Mass Region

With the signal shape, efficiencies and bin splittings validated, the hypothesis test for the production of $\Upsilon(1^3D_J)$ (with subsequent decay into $\pi^+\pi^-\Upsilon(1S)$) was carried out. The $\Upsilon(1^3D_J)$ is a low-lying triplet, sitting at ~ 10164 MeV with a predicted mass splitting of $\sim 6-8$ MeV [101] and the $J = 2$ state 1-2 MeV below the centre-of-gravity. So far, only the $J = 2$ state has been observed [28, 29] in the cascade $\Upsilon(3S) \to \gamma\gamma\Upsilon(1^3D_2)$, with little understanding of how this might transfer to production in *pp* collisions. At least one paper [108] suggests that the production rate could be as high as that of the $\Upsilon(2S)$, but is based entirely on the colour-octet production mechanism. In light of the small branching fraction, $\mathcal{B}(\Upsilon(1^3D_2) \to \pi^+\pi^-\Upsilon(1S)) = (0.66^{+0.15}_{-0.14} \pm 0.06)\%$, even such a high rate would be at the edge of the sensitivity of this analysis.

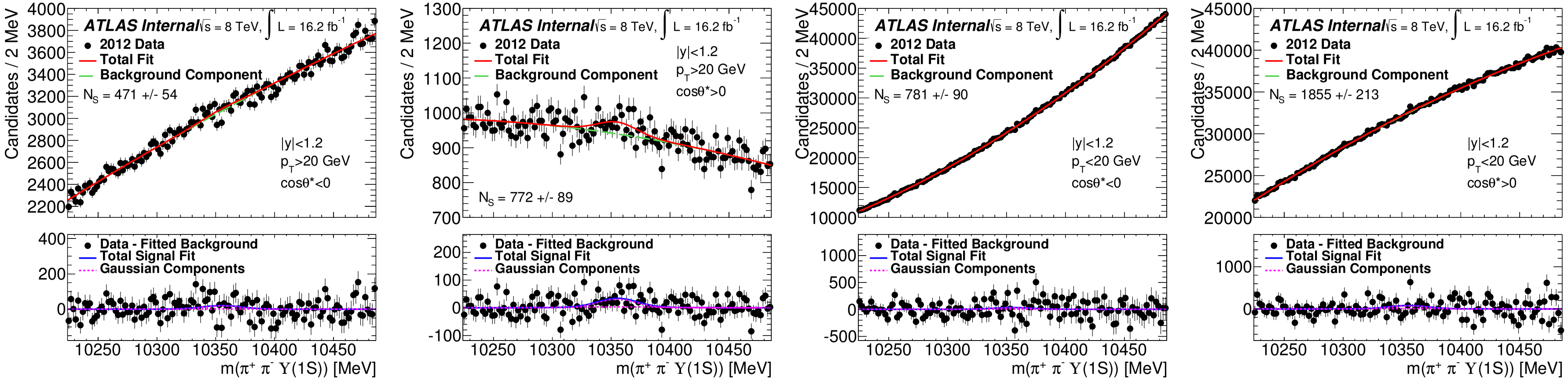
The presence of the $\Upsilon(1^3D_J)$ triplet is tested in two ways — firstly with a simultaneous fit involving a single signal peak, representing the $J = 2$ state, at the world average mass of 10164 MeV, and secondly a triplet fit with the splitting fixed to that used in the MC samples (which itself was determined as a weighted average over a number of theoretical predictions [101]). For both of these, the bin splittings and σ parameters were assumed equal for the three angular momentum states, and were calculated as the weighted averages over the $\Upsilon(1^3D_J)$ MC samples. The significance for the single peak hypothesis was found to be 0.09 with a signal yield of 79 ± 920, whilst the triplet hypothesis lead to a significance of 0.12 and signal yields of -1016 ± 3119, 631 ± 1840 and 781 ± 2313 for $J = 1, 2, 3$ respectively. Therefore, there appears to be no evidence for strong $\Upsilon(1^3D_J)$ production in *pp* collisions, and an upper limit can be set on the rate (see Section 7.6.8.3).

7.6.6 Preparations for the Search for $X_b \to \pi^+\pi^-\Upsilon(1S)$ Decays

The search for new hidden-beauty states decaying through the $\pi^+\pi^-\Upsilon(1S)$ channel was conducted under the assumption that the state is narrow and has a production spectrum which resembles that of the $\Upsilon(2S)$ and $\Upsilon(3S)$ states. Furthermore, the dipion mass and decay angular distributions were assumed to follow that of uniform phase space.



(a) Barrel, high p_{T}, low $\cos\theta^*$ (b) Barrel, high p_{T}, high $\cos\theta^*$ (c) Barrel, high p_{T}, low $\cos\theta^*$ (d) Barrel, Low p_{T}, high $\cos\theta^*$



(e) Endcap, high p_{T}, low $\cos\theta^*$ (f) Endcap, high p_{T}, high $\cos\theta^*$ (g) Endcap, low p_{T}, low $\cos\theta^*$ (h) Endcap, low p_{T}, high $\cos\theta^*$

Figure 7.9: The projections of the simultaneous fit to the Υ(3S) region in each of the analysis bins. The mass was fixed to the world average value, and the remaining signal shape parameters and bin splitting fractions were extracted from the Υ(3S) simulation. The reduced-χ^2 for the fit is 1.0, the significance is $z = 8.7$ standard deviations and the fitted yield is $N_{3\mathrm{S}} = 11596 \pm 1339$. Note the zero suppression on the vertical axes.

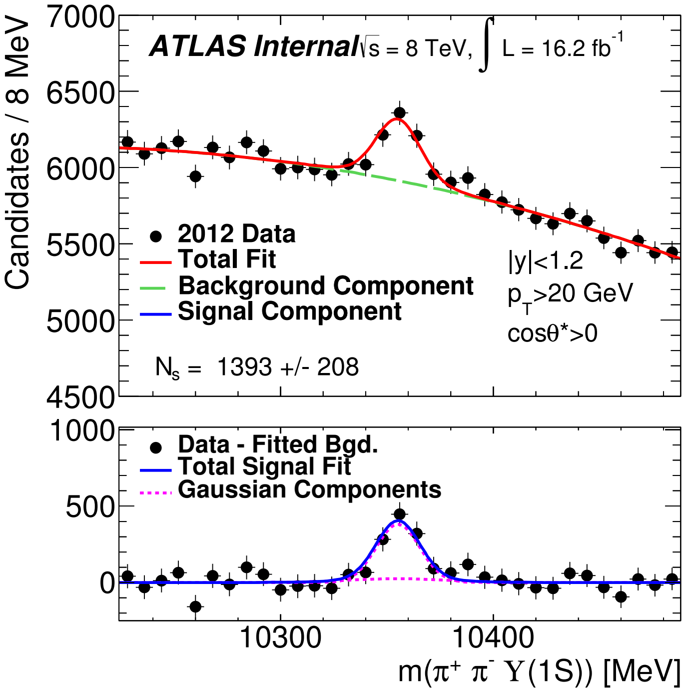


Figure 7.10: An individual fit to the most sensitive bin, yL_ptH_ctsH, with 8 MeV binning chosen to visually emphasise the $\Upsilon(3S)$ signal peak. Note the zero suppression on the vertical axis.

Simultaneous fits were carried out every 10 MeV, between 10 and 11 GeV, in much the same vein as that for the $\Upsilon(3S)$ peak.

To perform this scan requires (1) the signal shape and (2) the bin splitting for any considered mass, m. The former was achieved through a straightforward linear extrapolation of the signal shape parameters using the results of Table 7.4. For the latter, a slightly more involved procedure was needed, and is explained below.

To calculate the expected significance for a signal of given strength and for the upper limit evaluation, the efficiency and acceptance for a state of arbitrary mass additionally needed to be known. The extrapolation of these factors is explained in Section 7.6.6.2.

7.6.6.1 Extrapolation of Bin Splittings

To extrapolate the bin fractions, the splitting was modelled as a three-step process:

1. The separation of signal decays into the barrel ($|y| < 1.2$) and endcap ($1.2 < |y| < 2.4$) rapidity bins, described by a single number $S_{|y|}$.

2. The subsequent splitting of the barrel and endcap bins into high (> 20 GeV) and low (< 20 GeV) p_T regions. This requires two numbers, $S_{p_T}(yR)$, where $R = L, H$ refers to the high and low rapidity bins.

3. A final split into positive and negative $\cos\theta^*$ bins within each of the four regions defined above, with four corresponding numbers $S_{\cos\theta^*}(yR, ptR')$, where $R = L, H$ and $R' = L, H$.

By convention, each of these splitting numbers is defined as the fraction falling into the lower subset of the two. From the weighted MC samples, the barrel/endcap splitting, $S_{|y|}$, was found to be mass-independent and equal to $S_{|y|} = 0.67$. In contrast, the p_T splittings appear to reach a plateau with increasing parent mass (Figure 7.11), matching what one would expect based on physical intuition. To characterise this behaviour, a function of the form

$$S_{p_T}(m; yR) = \frac{a}{1 + be^{-c(m-d)}} \tag{7.15}$$

was fitted[9] through the values corresponding to the signal MC samples (the red curves in Figure 7.11). The evolution of the four different $\cos\theta^*$ splittings with mass is well described by a quadratic function. The corresponding fits are shown in Figure 7.12.

The fraction of signal decays in each analysis bin for a state of any mass, m, can then be evaluated as products of the appropriate three splitting functions. For example,

$$f_{yL\_ptL\_ctsL}(m) = 0.68 \cdot S_{p_T}(m; yL) \cdot S_{\cos\theta^*}(yL, ptL) \tag{7.16}$$

and

$$f_{yH\_ptH\_ctsH}(m) = (1 - 0.68) \cdot (1 - S_{p_T}(m; yH)) \cdot (1 - S_{\cos\theta^*}(yH, ptH)). \tag{7.17}$$

7.6.6.2 Extrapolation of the Acceptance and Efficiency

The acceptance is the fraction of all signal decays which pass the following conditions:

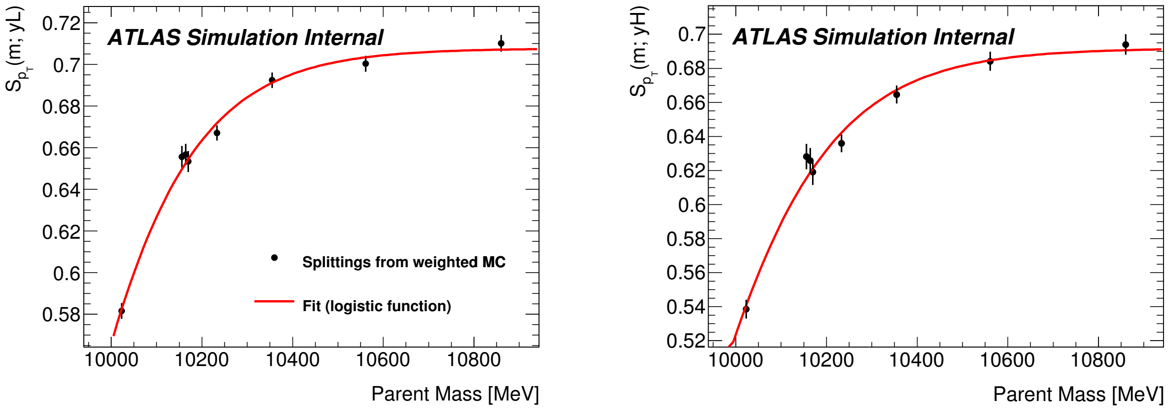
1. Muons with $p_T > 4$ GeV and $|\eta| < 2.5$,
2. Pions with $p_T > 400$ MeV and $|\eta| < 2.5$.

Given a production spectrum in p_T and $|y|$, this can easily be calculated with a toy MC by assuming uniform phase-space sampling. Assuming the extended Υ(2S) or Υ(3S) differential cross sections as inputs leads to different results, because the latter is slightly harder in p_T than the former. Presumably, [10] this indicates a trend of increasing production p_T with increasing mass — at least for S-wave states. To reflect this behaviour,

---

[9] Using the binned, minimum-χ^2 method.

[10] This is also supported by the difference in shape between the Υ(1S) and Υ(2S) spectra. The 'softening' effect of feed-down was also assessed with a rough calculation and determined to be too small to account for the difference.



(a) $S_{p_T}(m; yL)$

(b) $S_{p_T}(m; yH)$

Figure 7.11: The p_T splittings as a function of mass, with the data points from the production-weighted MC samples fitted with the logistic function.

the acceptance values used for masses lying between (outside) $m_{3\mathrm{S}}$ and $m_{2\mathrm{S}}$ were a linear interpolation (extrapolation) of the values based on the two spectra — i.e.

$$A(m) = A_{2\mathrm{S}} + \frac{m - m_{2\mathrm{S}}}{m_{3\mathrm{S}} - m_{2\mathrm{S}}} \left(A_{3\mathrm{S}} - A_{2\mathrm{S}} \right), \tag{7.18}$$

in obvious notation. The acceptance was evaluated every 10 MeV using 10 million simulated decays to give the result in Figure 7.13, which also has the corresponding $\Upsilon(2\mathrm{S})$ and $\Upsilon(3\mathrm{S})$-based values for comparison.

The residual (i.e. conditional on acceptance) efficiency was calculated from the weighted and filtered MC samples. A simple analytic function (essentially the complement of a logistic function) of the form

$$\epsilon(m) = a + \frac{b}{1 + e^{-c(m-d)}} \tag{7.19}$$

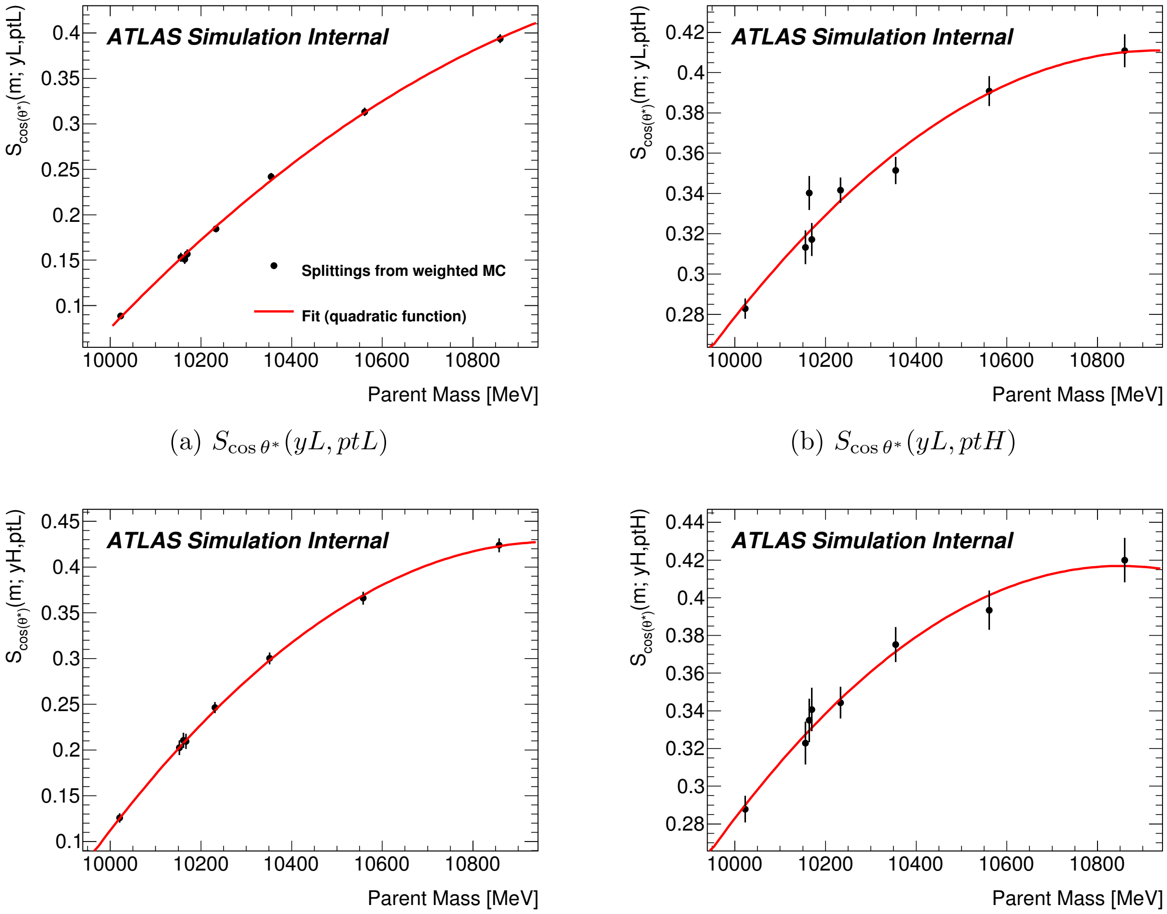
was fitted with the minimum-χ^2 method to the data points[11] in Figure 7.14 to characterise the slight decrease in efficiency with mass. This function then allows a simple extrapolation of the efficiency to any mass within the considered range. Other similar functions could equally well have been used here and the effect of doing so was considered as a systematic uncertainty (Section 7.6.8.1).

7.6.6.3 Incorporating the $\Upsilon(2\mathrm{S})$ and $\Upsilon(3\mathrm{S})$ Peaks

For hypothetical masses close to $m_{2\mathrm{S}}$ or $m_{3\mathrm{S}}$, the fitting range will include part of the $\Upsilon(2\mathrm{S})$ or $\Upsilon(3\mathrm{S})$ signal peaks, respectively. To account for these cases, the $\Upsilon(2\mathrm{S})$ and

---

[11] Each representing one of the MC samples.



(a) $S_{\cos\theta^*}(yL, ptL)$

(b) $S_{\cos\theta^*}(yL, ptH)$

(c) $S_{\cos\theta^*}(yH, ptL)$

(d) $S_{\cos\theta^*}(yH, ptH)$

Figure 7.12: The $\cos\theta^*$ splittings as a function of mass, with the data points from the production-weighted MC samples fitted with a quadratic function.

Υ(3S) signal peaks were included as additional terms in the background model, with the shape parameters again derived from the simulation samples. The normalisations were floated under a Gaussian constraint with mean set to the measured yields found in Sections 7.6.3 and 7.6.4, and widths given by the statistical uncertainties of the corresponding fits.

To avoid any possible bias or circularity in the results, mass windows of $m \pm 4\sigma_{yL}$ around the nominal Υ(2S) and Υ(3S) masses were excluded from any p-value or upper limit calculations, where σ_{yL} is the width of the narrow Gaussian signal component for the barrel region. The barrel width was chosen to minimise the excluded region; the barrel dominates the yield measurement in the simultaneous fit. This splits the analysis into two sub-ranges; 10.05–10.31 GeV and 10.40–11.00 GeV.

To properly include these regions would require independent knowledge of the con-

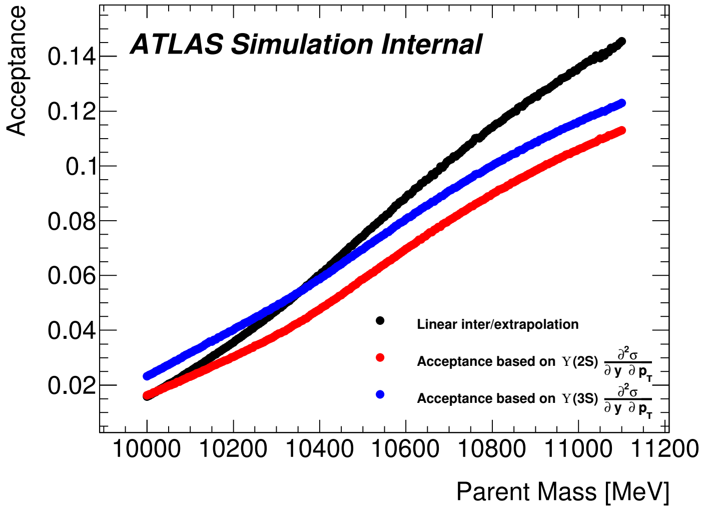
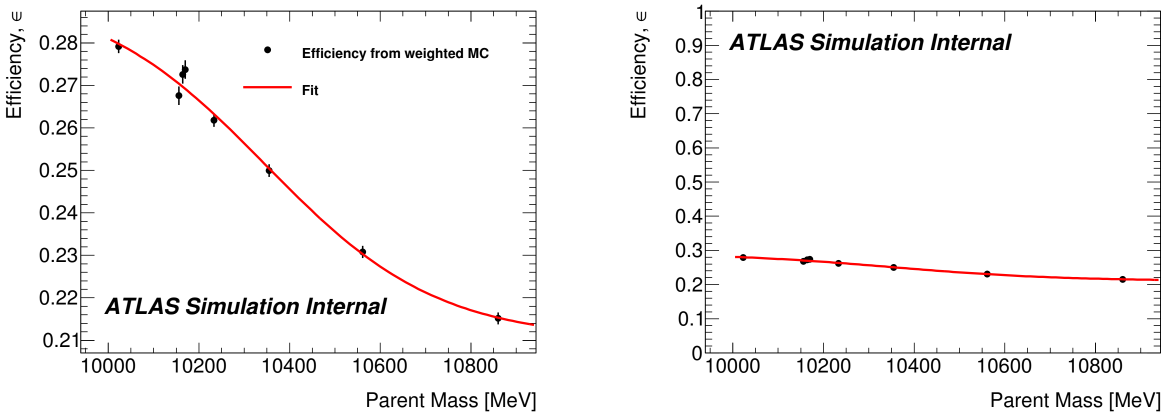


Figure 7.13: The acceptance as a function of mass, showing the values obtained based on the extended, measured $\Upsilon(2S)$ and $\Upsilon(3S)$ differential cross sections, and their linear inter/extrapolation.



(a) (b)

Figure 7.14: The efficiency for each of the weighted MC samples, with a simple fit using an analytic function to represent the mass dependence. The right hand side shows the full vertical scale for perspective.

tributions of the Υ(2S) and Υ(2S) signal peaks to the background, rather than relying on the fitted yields from the 2012 dataset itself. This was attempted using Equation 7.14 and the equivalent for $N_{3\mathrm{S}}$, with additional factors to estimate the increases in the cross sections between 7 and 8 TeV. These factors have not been measured, and instead were estimated using Pythia8 with large uncertainties assigned. At the level of precision provided by these predictions, the sensitivity for new signals close to $m_{2\mathrm{S}}$ or $m_{3\mathrm{S}}$ was very low and, in regions directly adjacent, the expected significance and upper limit curves showed undesirable irregularities. As this would likely lead to exclusion windows similar to those described above, no advantage was seen in adopting this method.

7.6.7 Results of the X_b Search: Local p-values

The local p-values, determined using the likelihood ratio test statistic [104], show no evidence for any new states (Figure 7.15) above the 3σ level — the highest statistical significance (at 10630 MeV) was 2.5σ, which would be lower after taking into account the 'look-elsewhere effect'.

Overlaid on this plot is the median significance one would expect for a signal of relative strength $R = 0.0656$ (blue dashed curve) or $R = 0.03$ (red dashed curved), where R is defined as

$$R := \frac{\sigma\mathcal{B}}{(\sigma\mathcal{B})_{2\mathrm{S}}} \tag{7.20}$$

The value of $R = 6.56\%$, in particular, holds some relevance because it is the analogous rate for $X(3872)$ production with respect to the ψ(2S) [4]. This analysis provides $> 5\sigma$ coverage for a signal of such strength across the entire range, and is significantly more sensitive for the majority of it. This is especially true at higher masses, where the separation in the $(p_T, \cos\theta^*)$ plane between signal and background becomes increasingly greater.

Given this null result, focus was turned to calculating the corresponding upper limits for R.

7.6.8 Upper Limits for X_b Production

7.6.8.1 Systematic Uncertainties

In carrying out the analysis, a number of assumptions were made about the signal production, peak shape, efficiencies, acceptance, bin splittings and so on. The effects of these on the upper limits were assessed by considering their influence on the following two categories of factors:

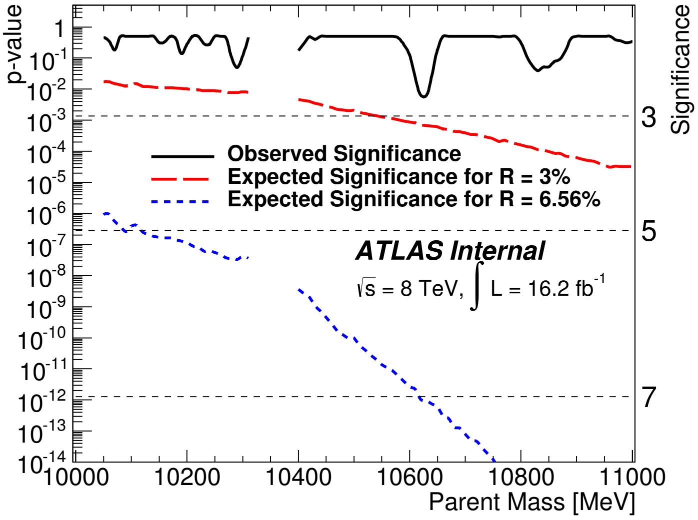


Figure 7.15: The observed local p-value (and significance) as a function of mass, with the expected curves for relative production rates of $R = 6.56\%$ (blue dashed curve) and $R = 3\%$ (red dashed curved).

1. **Fitting-related**
   - Signal shape — the narrow fractions, width ratios, and the widths themselves (f_n, r, σ)
   - Bin splitting — i.e. from the rapidity, transverse momentum and $\cos\theta^*$ splitting functions described above (Section 7.6.6.1).
   - Background shape
   - Fitting procedure — fitting range, binning etc.
2. **Scaling-type**
   The relative production rate, R, can be expressed as

$$R = \frac{N}{N_{2S}} \frac{(\mathcal{A}\epsilon)_{2S}}{\mathcal{A}\epsilon}. \tag{7.21}$$

   Changes to N_{2S} or either of the ratios, $\mathcal{A}/\mathcal{A}_{2S}$ and ϵ/ϵ_{2S}, directly change the upper limit value.

The effects of the systematic uncertainties on each particular factor above were added in quadrature to give a total effective systematic for that particular parameter (Tables 7.7 and 7.8).

The procedures used to extracting the specific values for the uncertainties are discussed in detail below.

Table 7.7: The contribution of the various sources of systematic uncertainty to the fitting-type parameters influencing the upper limit calculation. All values are relative uncertainties, expressed as a percentage.

	σ_{yL}	σ_{yH}	$f_{n,yL}$	$f_{n,yH}$	r_{yL}	r_{yH}	$S_{\lvert y\rvert}$	$S_{p_T}(yL)$	$S_{p_T}(yH)$	$S_{\cos\theta^*}(yL,ptL)$	$S_{\cos\theta^*}(yL,ptH)$	$S_{\cos\theta^*}(yH,ptL)$	$S_{\cos\theta^*}(yH,ptH)$
Extracting f_n, r			0.53	1.08	1.16	1.39							
Extrapolating σ	0.13	0.23											
Data/MC difference in σ	1.94	4.16											
$\lvert y\rvert$ scale factors							5.84						
Production weighting							0.33	8.37	7.02	0.88	2.79	2.14	3.44
Bin splittings: fit							0.18	0.48	0.76	2.35	4.18	2.75	5.98
Bin splittings: parameterisation							1.84	1.04	1.21	0.23	0.18	0.38	0.20
$m_{\pi^+\pi^-}$ shape							0.24	8.01	11.5	34.72	16.21	15.89	15.04
Total	1.95	4.17	0.53	1.08	1.16	1.39	6.14	11.64	13.57	34.81	16.97	16.27	16.55

Table 7.8: The contribution of the various sources of systematic uncertainty to the scaling-type parameters influencing the upper limit calculation. All values are relative uncertainties, expressed as a percentage.

	N_{2S}	ϵ/ϵ_{2S}	$\mathcal{A}/\mathcal{A}_{2S}$	$\epsilon/\epsilon_{2S} \cdot \mathcal{A}/\mathcal{A}_{2S}$
N_{2S} yield	2.31			
ϵ vs. m: fit		0.95		
ϵ vs. m: parameterisation		0.49		
Production weighting		1.03		
Acceptance Extrapolation			11.65	
$m_{\pi^+\pi^-}$ shape				17.25
Total	2.31	1.48	11.65	17.25

Extracting f_n and r

The two fixed shape parameters, f_n and r, were extracted as averages over the MC samples (see Table 7.4). The statistical uncertainties in deducing these averages were taken to be the systematics for these parameters.

Extrapolating σ

The two σ parameters (for the barrel and endcap) were extrapolated to arbitrary masses by assuming a linear relationship (see Table 7.4). The uncertainty on extrapolated values due to the uncertainties on the fitted parameters of the associated linear fit, α and β, was calculated through simple error propagation. The largest relative uncertainty across the considered mass range was taken as the systematic.

Data/MC difference in σ

Following on from the above point, the uncertainty in the *overall scale* of the MC-based estimates for the σ parameters was taken as the data/MC difference in the Υ(2S) peak.

$|y|$ scale factors

The $|y|$ scale factors were introduced to force consistency between data and MC (Section 7.6.3) in the signal splitting into the barrel and endcap regions. This was based on the observed value for this splitting at the Υ(2S) mass and the value in the Υ(2S) MC. The systematic uncertainty on the splitting function $S_{|y|}$ was taken as the statistical uncertainty in the scale factor for the barrel, 5.84%.

N_{2S} yield

The Υ(2S) yield used in the upper limit calculation (see Equation 7.21) was that measured in this analysis: 34400±800. The statistical uncertainty gives a reasonable measure of the level of uncertainty in this factor.

ϵ vs. m: fit

The efficiencies from the MC samples were fitted, as a function of mass, with a logistic function. The uncertainty in this fit leads to a uncertainty of up to 0.95% in the extrapolated efficiencies used in Equation 7.21.

ϵ vs. m: parameterisation

Two other alternative fit models were tried. The first was based on arctan and the second on the algebraic function $x/\sqrt{1+x^2}$. The largest difference in the fitted efficiency across the mass range was 0.49%.

Production weighting

The production weights are based on a 7 TeV measurement, but were used directly on the 8 TeV MC samples. This was based on the assumption that the rate of increase in the cross-section between the two energies is well modelled in the corresponding MC samples. To assess the impact of using these weights on the upper limits, they were removed altogether. The impact on the efficiency ratio was determined to be at most 1.03%, while the impact on the splitting functions is up to 8.37% (see Table 7.7).

Bin splittings: fit

The uncertainties in the fitting parameters for each of the bin splittings were used to determine the largest effect this may have when propagated to the splitting functions.

Bin splittings: parameterisation

The uncertainty associated with fitting the bin splittings with specific models was assessed by considering the following alternative models:

- For $S_{|y|}$, a linear relationship,
- For the $S_{p_T}(yR)$, a turn-on curve based on arctan, and
- For the $S_{\cos\theta^*}(yR, ptR')$, a third-order polynomial.

Acceptance extrapolation

The acceptance at a given mass, m, was calculated as the linear extrapolation of the acceptance values obtained from assuming the Υ(2S) differential cross section and the Υ(3S) differential cross section. This was based on the fact that the latter is harder than the former. In fact, this trend is also true when the Υ(1S) is also included — i.e. $A_{1\mathrm{S}} < A_{2\mathrm{S}} < A_{3\mathrm{S}}$, where A_{nS} is the acceptance based on the differential cross section of the Υ(nS). By default, the Υ(1S) was not included in the extrapolation because its spectrum is strongly influenced by feed-down from higher states. However, as a systematic check a linear extrapolation was performed between the Υ(1S) and Υ(2S) and also between the Υ(1S) and the Υ(3S). This was then compared with the default procedure, with the largest relative difference of 11.65% taken as the systematic uncertainty.

$m_{\pi^+\pi^-}$ shape

The dipion mass distribution of the X_b is unknown, which is reflected in this analysis by assuming a shape based on a uniform sampling of phase-space. However, all of the well known dipion transitions proceed with distinctive $m_{\pi^+\pi^-}$ distributions, viz.:

- the $\Upsilon(2\mathrm{S}) \to \pi^+\pi^-\Upsilon(1\mathrm{S})$ is 'top-heavy' — i.e. is strongest for dipion masses close to the upper threshold, $m_{2\mathrm{S}} - m_{1\mathrm{S}}$,
- the $\Upsilon(3\mathrm{S}) \to \pi^+\pi^-\Upsilon(1\mathrm{S})$ has a double-peaked structure — one peak near the lower threshold and one near the upper,
- the shape for the $X(3872) \to \pi^+\pi^-\Upsilon(1\mathrm{S})$ decay is similar to that for the Υ(2S),
- the shape for the $\psi(2S) \to \pi^+\pi^- J/\psi$ is very similar to that for $\Upsilon(2\mathrm{S}) \to \pi^+\pi^-\Upsilon(1\mathrm{S})$ [114],
- $Y(4260) \to \pi^+\pi^- J/\psi$ produces a similar double-peaked structure in the dipion mass [115], and
- the $\Upsilon(4\mathrm{S}) \to \pi^+\pi^-\Upsilon(1\mathrm{S})$ shape is much like that for $\Upsilon(2\mathrm{S}) \to \pi^+\pi^-\Upsilon(1\mathrm{S})$ [69, 70].

Relative to the phase-space only shape, which is fairly central between the upper and lower thresholds, the Υ(2S) and Υ(3S) options therefore represent two plausible extremes. Shapes based on both of these were tried as alternative hypotheses, and the largest relative difference this caused in the bin splitting parameters and total efficiency were taken as systematics. For the Υ(2S) shape, the parent mass is explicitly included in the analytic formula of the Voloshin and Zakharov model [106], whereas for the Υ(3S) the shape of the distribution was simply *stretched* to cover the available $m_{\pi^+\pi^-}$ range for each of the MC samples.

Background shape uncertainty

The default fitting model for the background shape is a 2^{nd}-order Chebychev polynomial. To allow for some systematic uncertainty in this choice, this was extended to include a small 3^{rd}-order term, Gaussian constrained about zero with a width of 0.05 (the typical fitted value of the 2^{nd}-order parameter is ~ 0.01).

7.6.8.2 Spin-Alignment Uncertainty

The analysis was conducted under the assumption that the states of interest — Υ(2S), Υ(3S), and potentially an X_b — are produced un-polarised in *pp* collisions. For the two *S*-wave states, this is supported by a measurement from CMS [7], but for the X_b the spin-alignment is unknown. In keeping with previous ATLAS analyses involving this issue (e.g. [14, 6]), the unknown polarisation is treated not as a source of systematic uncertainty, but as a parameter of the signal requiring different scenarios to be considered.

The analysis is re-run for each of the four standard extreme alternative polarisations — one 'longitudinal' and three 'transverse' (see [14] or Section 4.6.1 for definitions). The polarisation state of the X_b was assumed to be inherited by the Υ(1S) and the systematic uncertainties were assumed to be similar to their values in the default flat phase space case. Changing the underlying spin-alignment leads to different decay angular distributions, which feeds into the:

- acceptance,
- efficiency, and
- bin splittings.

The differences in the median expected upper limits with respect to the flat polarisation hypothesis is shown in Figure 7.16. Because these are fairly constant in mass, the uncertainty due to the unknown spin alignment is represented in the final result as a single error bar containing representative values for each of the extreme scenarios, conservatively taken as the maximum differences.

7.6.8.3 Result

As for the local p-values, the upper limits[12] were calculated using asymptotic formulae [104] in 10 MeV intervals for masses in the ranges 10.05–10.31 GeV and 10.40–11.00 GeV. Each of the systematic uncertainties were incorporated into the fitting procedure as Gaussian-constrained nuisance parameters, except the unknown polarisation which

---

[12] For consistency with the CMS analysis [2], the CL_S approach was used with a confidence level of 95%.

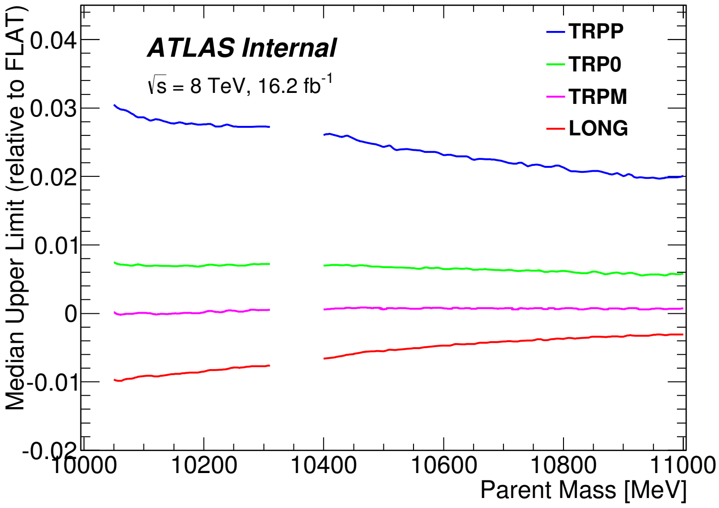


Figure 7.16: The difference in the expected upper limit values between the nominal (flat) polarisation scenario and the four other extreme alternatives ([14] provides definitions for these). In general, the differences are slightly smaller at the top end of the mass spectrum.

was treated separately (see above). The result, given in Figure 7.17, excludes relative production rates (values of R) above 4.0% across the considered range between 10 and 11 GeV. In particular, the observed upper limit in the vicinity of the $\Upsilon(1^3\mathrm{D_J})$ triplet is 2.2% which, given the measured branching fraction for $\Upsilon(1^3\mathrm{D}_2) \to \pi^+\pi^-\Upsilon(1\mathrm{S})$ of $(6.6 \pm 1.6) \times 10^{-3}$ [3], implies an upper limit on $\sigma(pp \to X\Upsilon(1^3\mathrm{D}_2))$ of $0.55 \cdot \sigma(pp \to X\Upsilon(2\mathrm{S}))$.

The inclusion of systematic uncertainties causes an increase in the observed limits of up to 13.2% and an inflation of the $\pm 1\sigma$ band by $9.5 - 25\%$, depending on the X_b mass.

7.6.8.4 Other Systematic Checks

Other systematic checks were carried out to establish the stability of the analysis. This included doubling the binning and changing the fitting range from the default of $m \pm 8\sigma_{yH}$ to $m \pm 7\sigma_{yH}$ and $m \pm 9\sigma_{yH}$. In all cases, only slight differences were seen in the observed upper limits.

The CMS analysis [2] was performed under the central assumption that the dipion distribution for any potential X_b signal would be that of the $\Upsilon(2\mathrm{S}) \to \pi^+\pi^-\Upsilon(1\mathrm{S})$ transition, in contrast to the assumption here that it be uniform in phase space. The analysis was redone under the CMS assumption and found to lead to a relatively small overall increase in the observed upper limits (Figure 7.18), demonstrating no significant advantage in the choice used here.

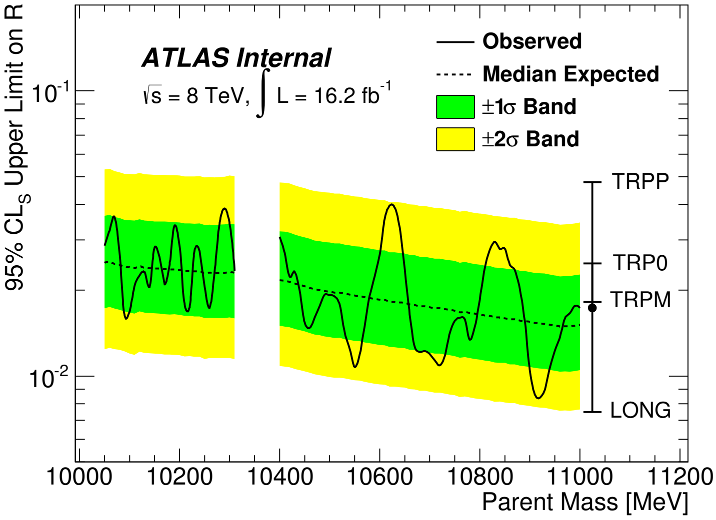


Figure 7.17: The observed and expected 95% CLs upper limits for the relative production rate, R, with the $\pm 1\sigma$ and $\pm 2\sigma$ bands. The single error bar on the right represents the typical systematic shift due to the unknown spin-alignment for the hypothetical state: the heavy dot represents the default (un-polarised) case.

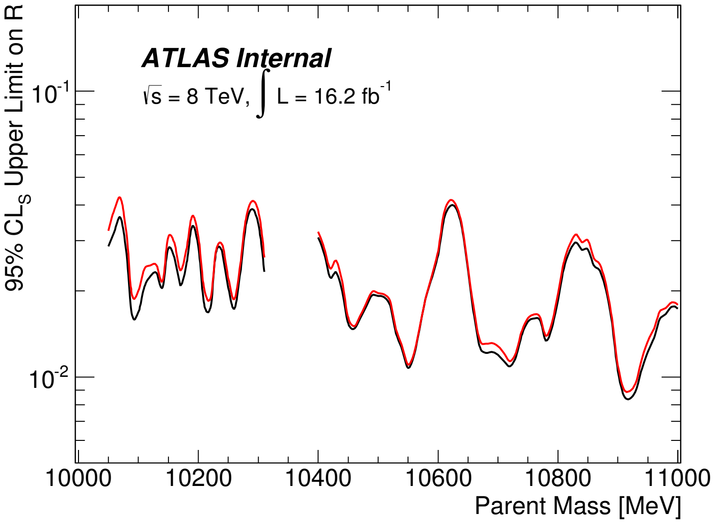


Figure 7.18: The observed upper limits under two different dipion mass distribution hypotheses — the uniform phase space distribution (used in this analysis) and the $\Upsilon(2S) \to \pi^+\pi^-\Upsilon(1S)$ distribution (used by CMS).

7.6.9 Analysis of $\Upsilon(10860)$ and $\Upsilon(11020)$ Mass Regions

The $\Upsilon(10860)$ is also known to decay to the $\pi^+\pi^-\Upsilon(1S)$ final state with a branching fraction of 0.53% [3] — a value unexpectedly high in comparison to that for the $\Upsilon(4S)$, which is two orders of magnitude lower. The Belle collaboration also claims exotic structure in the $\pi\Upsilon(nS)$ mass spectrum of $\Upsilon(10860) \to \pi^+\pi^-\Upsilon(nS)$ decays [15], giving this particular state additional interest. Though not yet observed, the $\Upsilon(11020)$ presumably participates in dipion transitions to the lower-lying S−wave states and could show similar exotic substructure.

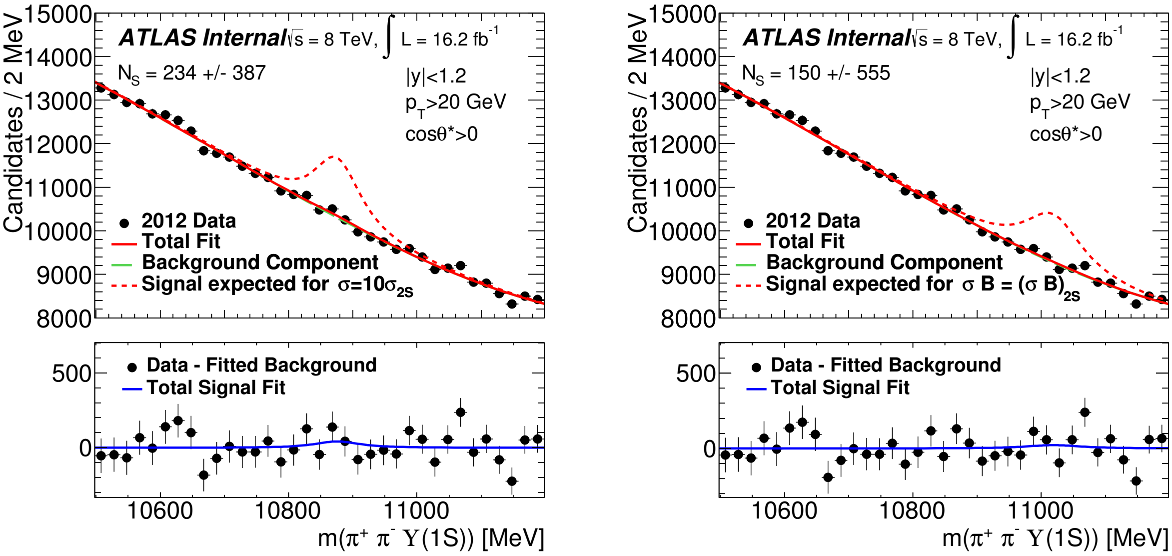
From the trend of decreasing production rate of $\Upsilon(nS)$ with increasing n at ATLAS (see [6]), it might be expected that the production of these two states would be weaker than that for the $\Upsilon(2S)$. Even if the $\Upsilon(10860)$ had a similar rate, the lower branching fraction gives an R value of 0.03, putting it at the edge of the sensitivity of this analysis from the outset. Unlike all of the observed and hypothetical states considered so far in the analysis, though, the $\Upsilon(10860)$ and $\Upsilon(11020)$ have natural widths which are comparable to the resolution of the ATLAS detector — 55 MeV and 79 MeV, respectively [3]. These would therefore appear in the $\pi^+\pi^-\Upsilon(1S)$ mass spectrum with a Breit-Wigner-based shape, convolved with a term representing the mass-dependent detector resolution. The latter of these is just the local signal shape for narrow states as determined in Section 7.4.

The significance for a signal at either of these mass points was extracted by performing fits (see Figure 7.19 for the most sensitive bins) using this modified signal shape in a mass window of 10498 to 11198 MeV.[13] The background was still described by a Chebychev polynomial, though a 3^{rd}-order term was allowed to account for the expanded mass range in the fit. For the $\Upsilon(10860)$, the bin splitting was taken as that in the corresponding MC sample (which takes into account the larger natural width). For the $\Upsilon(11020)$, the fraction of signal in each channel was determined using the bin splitting functions, evaluated at $m = 11020$ MeV.

The local statistical significance was found to be $z = 0.6$ and $z = 0.3$ for the $\Upsilon(10860)$ and $\Upsilon(11020)$, respectively, when the masses and widths of the states were fixed to the world average values [3]. These parameters have large uncertainties, so the significance was recalculated in a grid of $m \pm 20$ MeV and $\Gamma \pm \Delta\Gamma$, where $\Delta\Gamma$ is the world average uncertainty on the width [3]. The largest significance for the $\Upsilon(10860)$ was $z = 1.1$ at $m = 10856$ MeV and $\Gamma = 55$ MeV, and was $z = 0.6$ for the $\Upsilon(11020)$ at $m = 11039$

---

[13] Ideally, the fitting range would be extended much higher (particularly for the $\Upsilon(11020)$), but the the (previously determined) cut-off on the $\pi^+\pi^-\Upsilon(1S)$ mass is 11200 MeV. This affects the sensitivity to these states, but does not exclude the possibility for a significance test.



(a) $\Upsilon(10860)$ (b) $\Upsilon(11020)$

Figure 7.19: The $\Upsilon(10860)$ and $\Upsilon(11020)$ simultaneous fit results projected into the most sensitive bin — $|y| < 1.2$, $p_T > 20$ GeV and $\cos\theta^* > 0$. The dashed red curves give the expected shape of a very strong signal — $\sigma = \sigma_{2\mathrm{S}}$ for the $\Upsilon(10860)$, where the branching fraction to $\pi^+\pi^-\Upsilon(1\mathrm{S})$ is known, and $\sigma\mathcal{B} = (\sigma\mathcal{B})_{2\mathrm{S}}$ for the $\Upsilon(11020)$. The normalisations for the signal peaks under these conditions were estimated using Equation 7.21, where the acceptance and efficiency were calculated at the nominal mass of the state. No attempt was made to account for the observed dipion mass distribution in $\Upsilon(10860) \to \pi^+\pi^-\Upsilon(1\mathrm{S})$ decays [15]. Note that zero has been suppressed on the vertical axis and the bin widths have been increased to 20 MeV to improve the visual clarity of the fits.

MeV and $\Gamma = 95$ MeV. Hence there is also no evidence for $\Upsilon(10860) \to \pi^+\pi^-\Upsilon(1\mathrm{S})$ or $\Upsilon(11020) \to \pi^+\pi^-\Upsilon(1\mathrm{S})$ decays in this analysis.

7.6.10 Comparison to CMS Results

At the energy and mass ranges considered in this analysis, the CMS detector provides superior resolution in two instances:

1. the dimuon resolution is roughly twice as good (see [6] and [112], for instance), and
2. the $\pi^+\pi^-\Upsilon(1\mathrm{S})$ mass resolution is better by a factor $\sim 5.8/3.5 \approx 1.7$ (by comparing Figure 7.7a with the equivalent in [2]).

Assuming a $\mu^+\mu^-$ background which is reasonably flat, the first of these should provide the CMS analysis with a factor ~2 reduction in background when selecting $\Upsilon(1\mathrm{S})$ candidates with respect to an equivalent selection in this analysis. A similar argument based on the second point above applies to the effective background under a potential signal peak in the $\pi^+\pi^-\Upsilon(1\text{S})$ spectrum itself. Labelling this background B and based on the above arguments alone, one may then expect

$$B_{\text{ATLAS}} \approx 3.4 B_{\text{CMS}}. \tag{7.22}$$

Assuming, for simplicity, a single-bin counting experiment, the sensitivities for a weak X_b peak would then be related as $z_{\text{CMS}}/z_{\text{ATLAS}} \approx \sqrt{3.4} = 1.8$; that is, all else (trigger, luminosity, selections etc.) being equal a CMS search for the X_b should be roughly double as sensitive as an equivalent ATLAS search.

A comparison of the results presented in this chapter (specifically, Figures 7.15 and 7.17) with those from the CMS analysis (see Figure 1.13 or [2]) shows this is not the case; this analysis has stronger expected p-values and expected median upper limits for $m > 10.1$ GeV. From the investigation into the sensitivities of various search approaches given in Section 7.3, this is presumably attributable to the following key differences between the two analyses:

- the use of the binning method rather than the ΔR selection,
- the transverse momentum trigger thresholds (4 GeV here and 6.5 GeV for the CMS analysis), and
- the target mass used for optimisation (10561 MeV here and 10023.26 MeV by CMS).

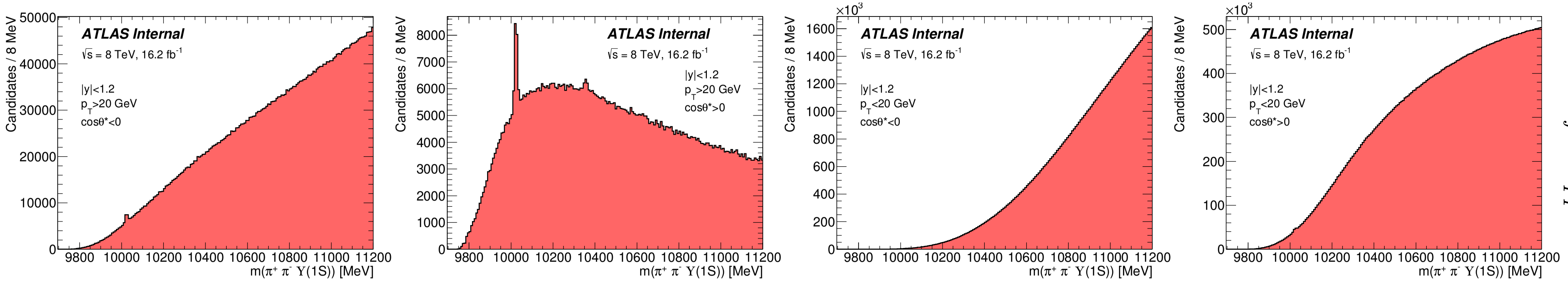
7.7 Conclusions

The recent observation of prompt $X(3872)$ production in pp collisions at CMS [4] through the $\pi^+\pi^- J/\psi$ channel suggests that if a $b\bar{b}$ partner state — the so-called X_b — were to exist, it may be observable in the analogous $\pi^+\pi^-\Upsilon(1\text{S})$ decay channel. In this analysis, a search for such a state was performed using 16.2 fb^{-1} of data collected at $\sqrt{s} = 8$ TeV using the ATLAS detector. $\Upsilon(1\text{S})$ decays were first reconstructed using two identified muons, which were subsequently combined with two charged tracks in a four-track vertex fit to form $\pi^+\pi^-\Upsilon(1\text{S})$ candidates. After validating the kinematics and efficiencies of the MC samples with the $\Upsilon(2\text{S})$ and $\Upsilon(3\text{S})$ states, a search for new narrow states was performed in the mass ranges $10.05 < m < 10.31$ GeV and $10.40 < m < 11.00$ GeV using a binning approach optimised for a hypothetical mass of 10561 MeV. No significant excess over the background was observed, despite an expected statistical sensitivity exceeding 5σ for a relative production rate similar to the analogous rate for the $X(3872)$. A dedicated search for the $\Upsilon(1^3\mathrm{D_J})$ using a triplet signal shape hypothesis, and another for the $\Upsilon(10860)$ and $\Upsilon(11020)$, which used a signal shape reflecting the large natural width of these states, also gave a null result.

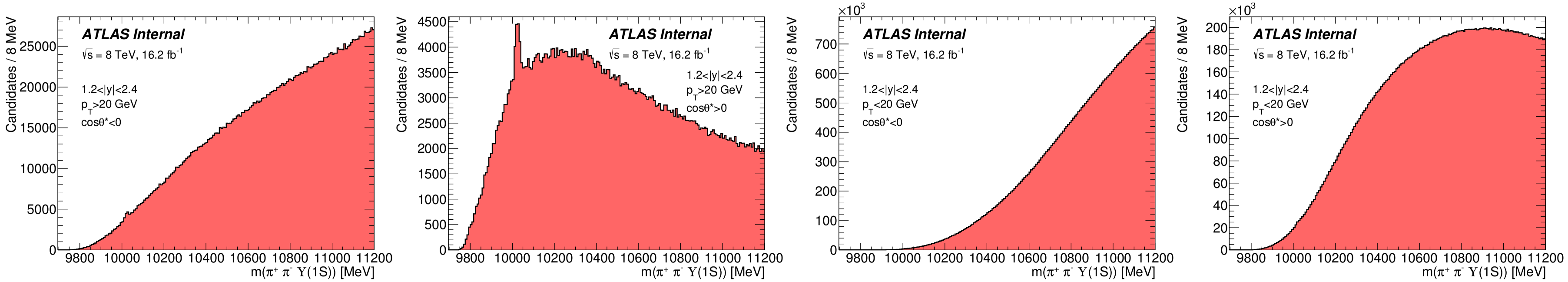
Consequently, an upper limit calculation was performed at the 95% confidence level for the relative production rate $R := \sigma(pp \to X_b \to \pi^+\pi^-\Upsilon(\mathrm{1S}))/\sigma(pp \to \Upsilon(\mathrm{2S}) \to \pi^+\pi^-\Upsilon(\mathrm{1S}))$, which allows a cancellation in many of the systematic uncertainties, with values of between 0.8% and 4.0%. These results are consistent with and, for most of the considered mass range, stronger than a similar analysis carried out using the CMS detector [2]. The uncertainty associated with the unknown production spin-alignment, evaluated for five extreme scenarios, was found to give rise to a maximum shift of $^{+0.03}_{-0.01}$ in the observed upper limits across the mass range considered.

7.8 Additional Material for Approval

The following plots were included in the paper and supporting internal documentation under the heading "Additional Material for Approval", as is standard practice in ATLAS for results which are not directly relevant to the analysis, but nevertheless aid in providing context or background.



(a) Barrel, high p_{T}, low $\cos\theta^*$ (b) Barrel, high p_{T}, high $\cos\theta^*$ (c) Barrel, low p_{T}, low $\cos\theta^*$ (d) Barrel, low p_{T}, high $\cos\theta^*$



(e) Endcap, high p_{T}, low $\cos\theta^*$ (f) Endcap, high p_{T}, high $\cos\theta^*$ (g) Endcap, low p_{T}, low $\cos\theta^*$ (h) Endcap, low p_{T}, high $\cos\theta^*$

Figure 7.20: The $\pi^+\pi^-\Upsilon(1S)$ invariant mass distributions for each of the analysis bins.

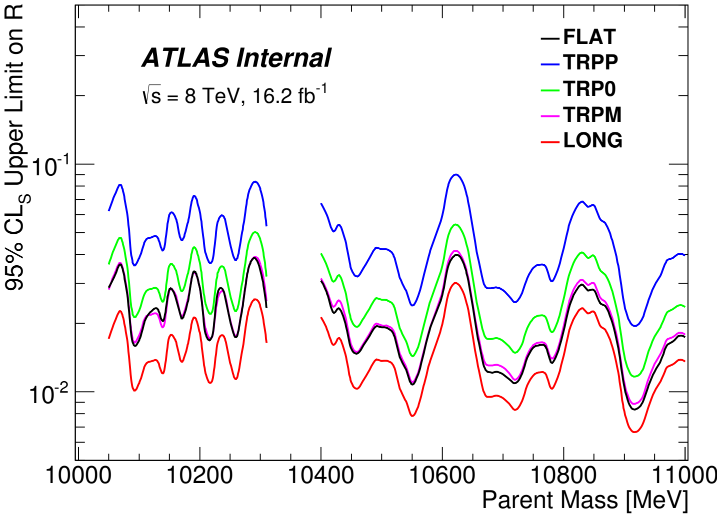


Figure 7.21: The observed 95% confidence level CL$_S$ upper limits for R under each of the extreme spin-alignment scenarios.

Appendix 7.A Acceptance

Figure 7.22 shows the acceptance fraction for Υ(2S) $\to \pi^+\pi^-\Upsilon$(1S) decays as a function of the p_T and rapidity, assuming isotropic angular distributions (i.e. the FLAT spin-alignment hypothesis). The total acceptance $\mathcal{A}$ is essentially calculated as the average value when weighted by the differential cross section.

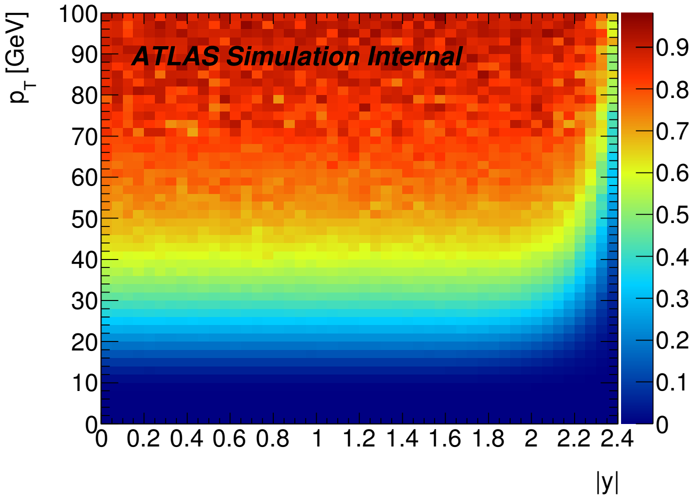


Figure 7.22: The Υ(2S) acceptance as a function of p_T and rapidity, assuming isotropic decay angular distributions (i.e. the flat spin-alignment hypothesis).

Chapter 8

Conclusion

The symmetry between the spectra of the $c\bar{c}$ and $b\bar{b}$ systems naturally leads to the conclusion that there should exist a hidden-beauty partner state to the $X(3872)$, the X_b. The observed properties [4] of the $X(3872)$ indicate that this hypothetical state should be strongly produced in pp collisions and decay through the $\pi^+\pi^-\Upsilon(1\mathrm{S})$ channel (though there are arguments that the branching fraction for this transition may be very low [72, 73]). This thesis set out to discover the X_b in this decay mode using the ATLAS detector at the LHC. Efforts were also made to observe the recently-discovered $\Upsilon(1^3\mathrm{D_J})$ triplet, as well as the $\Upsilon(10860)$ and $\Upsilon(11020)$.

Noting the problems found during the investigation of the feasibility of the dielectron channel in Chapter 4, it was decided that only the $\Upsilon(1\mathrm{S}) \rightarrow \mu^+\mu^-$ final state should be used. The analysis was developed on simulation samples (Chapter 5) using reconstruction and selection techniques based on similar studies performed at ATLAS. Two muons forming an $\Upsilon(1\mathrm{S})$ candidate were used to fire the trigger, then combined with a pair of oppositely charged tracks, representing the pions, in a vertex fit constrained such that $m_{\mu^+\mu^+} = m_{1S} = 9460.3$ MeV. The resulting $\pi^+\pi^-\Upsilon(1\mathrm{S})$ candidates were then subjected to a number of quality-assuring and acceptance selections, which also helped to reduce the background by a factor greater than 10. For the analysis of the 2011 dataset, the method of *candidate selection* was used, whereby for each $\Upsilon(1\mathrm{S})$ combination all $\pi^+\pi^-\Upsilon(1\mathrm{S})$ but a single 'best' candidate were rejected. This took advantage of the difference in the p_T between signal and background dipion pairs.

After checking the $\Upsilon(2\mathrm{S})$ signal yields in the 2011 dataset against predictions based on the simulation sample, it was found that there were some serious mis-modelling issues. Chapter 6 explains the reasons for this; the most important problem was related to the parent production $(p_T, |y|)$ spectrum used in the simulation samples, which was typically too hard and led to overestimated efficiencies. Others included the need to account for the correct dipion mass distributions and the mean p_T of background dipion combinations being incorrectly modelled in the simulation. Reweighting solutions were employed to make the simulation samples consistent with the data observations. The analysis was then applied to the 2011 dataset to perform a search for the X_b. No statistically significant peaks were observed in the $\pi^+\pi^-\Upsilon(1S)$ mass spectrum, so an upper limit calculation was performed with values for the relative production rate, with respect to the $\Upsilon(2S)$, of 3–15%. Dedicated searches for the $\Upsilon(1^3D_J)$, $\Upsilon(10860)$ and $\Upsilon(11020)$ also revealed no significant signal yields.

Soon after this, CMS released their results on an analogous X_b search [2], revealing an opportunity for signal-background separation which had not yet been fully exploited in this analysis. Consequently, the analysis was re-optimised for the much larger 2012 dataset by using the 2011 dataset as the 'background' sample and a new set of simulation samples. The best solution was found to be one in which candidate selection was dropped in favour of segregating the $\pi^+\pi^-\Upsilon(1S)$ candidates into 8 different bins of varying sensitivity, with the bin boundaries chosen to maximise the expected significance at the Swanson mass, $m = 10561$ MeV. With the mis-modelling issues of the simulation samples well-understood from the 2011 analysis, the kinematics of the known $\Upsilon(2S) \to \pi^+\pi^-\Upsilon(1S)$ and $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)$ decays were shown to be consistent with expectations. The search then proceeded much in the same manner as for the 2011 dataset — a dedicated triplet fit at the $\Upsilon(1^3D_J)$, then a scan of the $\pi^+\pi^-\Upsilon(1S)$ mass spectrum and then finally the $\Upsilon(10860)$ and $\Upsilon(11020)$. In all cases, nothing statistically significant was observed.

The final result is a 95% CLs upper limit calculation for masses in the ranges 10.05–10.31 and 10.40–11.00 GeV. Many sources of systematic uncertainties were considered and incorporated into the fits, which extract the upper limit values, as Gaussian-constrained nuisance parameters. The unknown spin-alignment was treated as a special case; each of five extreme scenarios was studied and found to produce shifts in the upper limits which were only a mild function of X_b mass. The uncertainty due to the different possible spin-alignments was represented in the final result with an error bar indicating typical values of these shifts (see Chapter 7). The upper limits for R obtained from this were between 0.8% and 4.0%, which excludes an X_b with a relative production rate equal to that of the $X(3872)$ with respect to the $\psi(2S)$. This is consistent with, and, for most of the mass range, stronger than, the findings of the CMS analysis [2].

In light of this, there are two alternative conclusions that can be made — either (1) the X_b does not exist, or (2) the X_b exists, but is either not produced as strongly as its hidden charm analogue or has a smaller branching fraction to the $\pi^+\pi^-\Upsilon(1S)$ channel.

In any case, the result will help to constrain the theoretical models for the structure, production and decays of the $X(3872)$.

References

[1] ATLAS Collaboration, The ATLAS Experiment at the CERN Large Hadron Collider, Journal of Instrumentation 3 (08) (2008) S08003.
URL http://stacks.iop.org/1748-0221/3/i=08/a=S08003

[2] CMS Collaboration, Search for a new bottomonium state decaying to $\Upsilon(1S)\pi^+\pi^-$ in pp collisions at $\sqrt{s} = 8$ TeV, Phys.Lett. B727 (2013) 57–76. arXiv:1309.0250, doi:10.1016/j.physletb.2013.10.016.

[3] J. Beringer, et al. (Particle Data Group), Review of Particle Physics, Physical Review D 86 (2012) 010001.

[4] CMS Collaboration, Measurement of the $X(3872)$ production cross section via decays to $J/\psi\pi^+\pi^-$ in pp collisions at $\sqrt{(s)} = 7$ TeV, JHEP 1304 (2013) 154. arXiv:1302.3968, doi:10.1007/JHEP04(2013)154.

[5] N. Brambilla, S. Eidelman, B. Heltsley, R. Vogt, G. Bodwin, et al., Heavy quarkonium: progress, puzzles, and opportunities, Eur.Phys.J. C71 (2011) 1534. arXiv:1010.5827, doi:10.1140/epjc/s10052-010-1534-9.

[6] ATLAS Collaboration, Measurement of Upsilon production in 7 TeV pp collisions at ATLAS, Phys.Rev.D. D87 (5) (2013) 052004. arXiv:1211.7255, doi:10.1103/PhysRevD.87.052004.

[7] CMS Collaboration, Measurement of the $\Upsilon(1S)$, $\Upsilon(2S)$, and $\Upsilon(3S)$ Polarizations in pp Collisions at $\sqrt{s} = 7$ TeV, Phys. Rev. Lett. 110 (2013) 081802. doi:10.1103/PhysRevLett.110.081802.
URL http://link.aps.org/doi/10.1103/PhysRevLett.110.081802

[8] CLEO Collaboration, The Hadronic Transitions $\Upsilon(2S) \to \Upsilon(1S)$, Phys. Rev. D 58 (1998) 052004. doi:10.1103/PhysRevD.58.052004.
URL http://link.aps.org/doi/10.1103/PhysRevD.58.052004

[9] P. Moxhay, Coupled-channel Effects in the decay $\Upsilon(3S) \to \Upsilon(1S)\pi^+\pi^-$, Phys.Rev. D39 (1989) 3497. doi:10.1103/PhysRevD.39.3497.

[10] L. Evans, P. Bryant, LHC Machine, Journal of Instrumentation 3 (08) (2008) S08001.
URL http://stacks.iop.org/1748-0221/3/i=08/a=S08001

[11] ATLAS Collaboration, B-physics Public Results (2014).
URL https://twiki.cern.ch/twiki/bin/view/AtlasPublic/BPhysPublicResults#Stand_alone_plots

[12] ATLAS Collaboration, Search for the Standard Model Higgs boson in the decay channel $H \to ZZ^{(*)} \to 4\ell$ with 4.8 fb^{-1} of pp collision data at $\sqrt{s} = 7$ TeV with ATLAS, Phys.Lett. B710 (2012) 383–402. arXiv:1202.1415, doi:10.1016/j.physletb.2012.03.005.

[13] D. Cronin-Hennessy, et al. (CLEO Collaboration), Study of dipion transitions among $\Upsilon(3S)$, $\Upsilon(2S)$, and $\Upsilon(1S)$ states, Phys. Rev. D 76 (2007) 072001. doi:10.1103/PhysRevD.76.072001.
URL http://link.aps.org/doi/10.1103/PhysRevD.76.072001

[14] ATLAS Collaboration, Measurement of the differential cross-sections of inclusive, prompt and non-prompt J/ψ production in proton-proton collisions at $\sqrt{s} = 7$ TeV, Nuclear Physics B 850 (3) (2011) 387 – 444. doi:10.1016/j.nuclphysb.2011.05.015.
URL
http://www.sciencedirect.com/science/article/pii/S0550321311002938

[15] A. Bondar, et al. (Belle Collaboration), Observation of Two Charged Bottomoniumlike Resonances in Υ(5S) Decays, Phys. Rev. Lett. 108 (2012) 122001. doi:10.1103/PhysRevLett.108.122001.
URL http://link.aps.org/doi/10.1103/PhysRevLett.108.122001

[16] CERN, LHC Performance and Statistics (2014).
URL https://lhc-statistics.web.cern.ch/LHC-Statistics/

[17] ATLAS Collaboration, Luminosity Public Results (2014).
URL https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults

[18] V. E. Barnes, et al., Observation of a hyperon with strangeness minus three, Phys. Rev. Lett. 12 (1964) 204–206. `doi:10.1103/PhysRevLett.12.204`.
URL `http://link.aps.org/doi/10.1103/PhysRevLett.12.204`

[19] Peskin, Michael E. and Schroeder, Dan V., An Introduction To Quantum Field Theory (Frontiers in Physics), Westview Press, 1995.

[20] S. L. Glashow, J. Iliopoulos, L. Maiani, Weak interactions with lepton-hadron symmetry, Phys. Rev. D 2 (1970) 1285–1292. `doi:10.1103/PhysRevD.2.1285`.
URL `http://link.aps.org/doi/10.1103/PhysRevD.2.1285`

[21] M. K. Gaillard, B. W. Lee, J. L. Rosner, Search for charm, Rev. Mod. Phys. 47 (1975) 277–310. `doi:10.1103/RevModPhys.47.277`.
URL `http://link.aps.org/doi/10.1103/RevModPhys.47.277`

[22] T. Appelquist, H. D. Politzer, Heavy quarks and e^+e^- annihilation, Phys. Rev. Lett. 34 (1975) 43–45. `doi:10.1103/PhysRevLett.34.43`.
URL `http://link.aps.org/doi/10.1103/PhysRevLett.34.43`

[23] J. J. Aubert, et al., Experimental Observation of a Heavy Particle J, Phys. Rev. Lett. 33 (1974) 1404–1406. `doi:10.1103/PhysRevLett.33.1404`.
URL `http://link.aps.org/doi/10.1103/PhysRevLett.33.1404`

[24] J. E. Augustin, et al., Discovery of a narrow resonance in e^+e^- annihilation, Phys. Rev. Lett. 33 (1974) 1406–1408. `doi:10.1103/PhysRevLett.33.1406`.
URL `http://link.aps.org/doi/10.1103/PhysRevLett.33.1406`

[25] S. Herb, D. Hom, L. Lederman, J. Sens, H. Snyder, et al., Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions, Phys.Rev.Lett. 39 (1977) 252–255. `doi:10.1103/PhysRevLett.39.252`.

[26] W. R. Innes, J. Appel, B. Brown, C. Brown, K. Ueno, et al., Observation of structure in the Υ region, Phys.Rev.Lett. 39 (1977) 1240. `doi:10.1103/PhysRevLett.39.1240`.

[27] ATLAS Collaboration, Observation of a New χ_b State in Radiative Transitions to $\Upsilon(1S)$ and $\Upsilon(2S)$ at ATLAS, Phys. Rev. Lett. 108 (2012) 152001. `doi:10.1103/PhysRevLett.108.152001`.
URL `http://link.aps.org/doi/10.1103/PhysRevLett.108.152001`

[28] G. Bonvicini, et al. (CLEO Collaboration), First observation of a $\Upsilon(1D)$ state, Phys.Rev. D70 (2004) 032001. `arXiv:hep-ex/0404021`, `doi:10.1103/PhysRevD.70.032001`.

[29] P. del Amo Sanchez, et al. (BaBar Collaboration), Observation of the $\Upsilon(1^3\mathrm{D}_J)$ Bottomonium State through Decays to $\pi^+\pi^-\Upsilon(1\mathrm{S})$, Phys.Rev. D82 (2010) 111102. `arXiv:1004.0175`, `doi:10.1103/PhysRevD.82.111102`.

[30] Aaltonen, T. and et al. (CDF Collaboration), Observation of the Decay $B_c^\pm \to J/\psi\pi^\pm$ and Measurement of the $B_c^\pm$ Mass, Phys. Rev. Lett. 100 (2008) 182002. `doi:10.1103/PhysRevLett.100.182002`.
URL `http://link.aps.org/doi/10.1103/PhysRevLett.100.182002`

[31] N. Brambilla, et al. (Quarkonia Working Group), Heavy quarkonium physics`arXiv:hep-ph/0412158`.

[32] Perkins, Donald H., Introduction to High Energy Physics (4^{th} edition), Cambridge University Press, 2000.

[33] M. Voloshin, Charmonium, Prog.Part.Nucl.Phys. 61 (2008) 455–511. `arXiv:0711.4556`, `doi:10.1016/j.ppnp.2008.02.001`.

[34] E. Eichten, K. Gottfried, T. Kinoshita, J. B. Kogut, K. Lane, et al., The Spectrum of Charmonium, Phys.Rev.Lett. 34 (1975) 369–372. `doi:10.1103/PhysRevLett.34.369`.

[35] D. Ebert, R. Faustov, V. Galkin, Properties of heavy quarkonia and B_c mesons in the relativistic quark model, Phys.Rev. D67 (2003) 014027. `arXiv:hep-ph/0210381`, `doi:10.1103/PhysRevD.67.014027`.

[36] S.-K. Choi, et al. (Belle Collaboration), Observation of a narrow charmoniumlike state in exclusive $B^\pm \to K^\pm\pi^+\pi^- J/\psi$ decays, Phys. Rev. Lett. 91 (2003) 262001. `doi:10.1103/PhysRevLett.91.262001`.
URL `http://link.aps.org/doi/10.1103/PhysRevLett.91.262001`

[37] B. Aubert, et al. (BaBar Collaboration), Study of the $B^- \to J/\psi K^-\pi^+\pi^-$ decay and measurement of the $B^- \to X(3872)K^-$ branching fraction, Phys. Rev. D 71 (2005) 071103. `doi:10.1103/PhysRevD.71.071103`.
URL `http://link.aps.org/doi/10.1103/PhysRevD.71.071103`

[38] Acosta, D. and et al. (CDF II Collaboration), Observation of the Narrow State $X(3872) \to J/\psi\pi^+\pi^-$ in $\overline{p}p$ Collisions at $\sqrt{s} = 1.96$ TeV, Phys. Rev. Lett. 93 (2004) 072001. `doi:10.1103/PhysRevLett.93.072001`.
URL `http://link.aps.org/doi/10.1103/PhysRevLett.93.072001`

[39] V. Abazov, et al. (DØ Collaboration), Observation and properties of the $X(3872)$ decaying to $J/\psi\pi^+\pi^-$ in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV, Phys.Rev.Lett. 93 (2004) 162002. `arXiv:hep-ex/0405004`, `doi:10.1103/PhysRevLett.93.162002`.

[40] LHCb Collaboration, Determination of the $X(3872)$ meson quantum numbers, Phys.Rev.Lett. 110 (2013) 222001. `arXiv:1302.6269`, `doi:10.1103/PhysRevLett.110.222001`.

[41] LHCb Collaboration, Precision measurement of D meson mass differences, JHEP 1306 (2013) 065. `arXiv:1304.6865`, `doi:10.1007/JHEP06(2013)065`.

[42] S.-K. Choi, et al. (Belle Collaboration), Bounds on the width, mass difference and other properties of $X(3872) \to \pi^+\pi^- J/\psi$ decays, Phys.Rev. D84 (2011) 052004. `arXiv:1107.0163`, `doi:10.1103/PhysRevD.84.052004`.

[43] A. Abulencia, et al. (CDF Collaboration), Analysis of the quantum numbers J^{PC} of the $X(3872)$, Phys.Rev.Lett. 98 (2007) 132002. `arXiv:hep-ex/0612053`, `doi:10.1103/PhysRevLett.98.132002`.

[44] Abulencia, A. and et al. (CDF Collaboration), Measurement of the dipion mass spectrum in $X(3872) \to J/\psi\pi^+\pi^-$ decays., Phys.Rev.Lett. 96 (2006) 102002. `arXiv:hep-ex/0512074`, `doi:10.1103/PhysRevLett.96.102002`.

[45] E. S. Swanson, The new heavy mesons: A status report, Physics Reports 429 (5) (2006) 243 – 305. `doi:http://dx.doi.org/10.1016/j.physrep.2006.04.003`.
URL
`http://www.sciencedirect.com/science/article/pii/S0370157306001475`

[46] Chen, K.-F. and et al. (Belle Collaboration), Observation of Anomalous $\Upsilon(1S)\pi^+\pi^-$ and $\Upsilon(2S)\pi^+\pi^-$ Production near the $\Upsilon(5S)$ Resonance, Phys. Rev. Lett. 100 (2008) 112001. `doi:10.1103/PhysRevLett.100.112001`.
URL `http://link.aps.org/doi/10.1103/PhysRevLett.100.112001`

[47] A. Bondar, R. Mizuk, Status and new results on the Z_b resonances, PoS ConfinementX (2012) 156. `arXiv:1303.0101`.

[48] N. A. Tornqvist, From the deuteron to deusons, an analysis of deuteron-like meson-meson bound states, Z.Phys. C61 (1994) 525–537.
arXiv:hep-ph/9310247, doi:10.1007/BF01413192.

[49] E. S. Swanson, Short range structure in the X(3872), Phys.Lett. B588 (2004) 189–195. arXiv:hep-ph/0311229, doi:10.1016/j.physletb.2004.03.033.

[50] K. Abe, et al. (Belle Collaboration), Evidence for $X(3872) \to \gamma J/\psi$ and the sub-threshold decay $X(3872) \to \omega J/\psi$ arXiv:hep-ex/0505037.

[51] L. Maiani, F. Piccinini, A. Polosa, V. Riquer, Diquark-antidiquarks with hidden or open charm and the nature of X(3872), Phys.Rev. D71 (2005) 014028.
arXiv:hep-ph/0412098, doi:10.1103/PhysRevD.71.014028.

[52] T. Aaltonen, et al. (CDF Collaboration), Precision Measurement of the $X(3872)$ Mass in $J/\psi\pi^+\pi^-$ Decays, Phys.Rev.Lett. 103 (2009) 152001. arXiv:0906.5218, doi:10.1103/PhysRevLett.103.152001.

[53] M. B. Einhorn, S. D. Ellis, Hadronic production of the new resonances: Probing gluon distributions, Phys. Rev. D 12 (1975) 2007–2014.
doi:10.1103/PhysRevD.12.2007.
URL http://link.aps.org/doi/10.1103/PhysRevD.12.2007

[54] N. Brambilla, NRQCD and Quarkonia, eConf C0610161 (2006) 004.
arXiv:hep-ph/0702105.

[55] Abe, F. and et al. (CDF Collaboration), J/ψ and $\psi(2S)$ Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV, Phys. Rev. Lett. 79 (1997) 572–577.
doi:10.1103/PhysRevLett.79.572.
URL http://link.aps.org/doi/10.1103/PhysRevLett.79.572

[56] P. Artoisenet, J. M. Campbell, J. Lansberg, F. Maltoni, F. Tramontano, Υ Production at Fermilab Tevatron and LHC Energies, Phys.Rev.Lett. 101 (2008) 152001. arXiv:0806.3282, doi:10.1103/PhysRevLett.101.152001.

[57] CMS Collaboration, Measurement of the $\Upsilon(1S)$, $\Upsilon(2S)$, and $\Upsilon(3S)$ cross sections in pp collisions at $\sqrt{s} = 7$ TeV, Physics Letters B 727 (13) (2013) 101 – 125.
doi:http://dx.doi.org/10.1016/j.physletb.2013.10.033.
URL
http://www.sciencedirect.com/science/article/pii/S0370269313008393

[58] CMS Collaboration, Measurement of the $\Upsilon(1S), \Upsilon(2S)$, and $\Upsilon(3S)$ cross sections in pp collisions at $\sqrt{s} = 7$ TeV, Phys.Lett. B727 (2013) 101–125. arXiv:1303.5900, doi:10.1016/j.physletb.2013.10.033.

[59] T. Sjostrand, S. Mrenna, P. Z. Skands, A Brief Introduction to PYTHIA 8.1, Comput.Phys.Commun. 178 (2008) 852–867. arXiv:0710.3820, doi:10.1016/j.cpc.2008.01.036.

[60] Jacob, M. and Wick, G. C., On the general theory of collisions for particles with spin, Annals of Physics 7 (1959) 404–428. doi:10.1016/0003-4916(59)90051-X.

[61] J. D. Richman, An Experimenter's Guide to the Helicity Formalism (CALT-68-1148), Tech. rep. (1984).

[62] P. Faccioli, C. Lourenco, J. Seixas, H. K. Wohri, Towards the experimental clarification of quarkonium polarization, Eur.Phys.J. C69 (2010) 657–673. arXiv:1006.2738, doi:10.1140/epjc/s10052-010-1420-5.

[63] T. Aaltonen, et al. (CDF Collaboration), Measurements of Angular Distributions of Muons From Υ Meson Decays in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV, Phys.Rev.Lett. 108 (2012) 151802. arXiv:1112.1591, doi:10.1103/PhysRevLett.108.151802.

[64] G. Bauer, et al. (CDF Collaboration), The $X(3872)$ at CDF II, Int.J.Mod.Phys. A20 (2005) 3765–3767. arXiv:hep-ex/0409052, doi:10.1142/S0217751X05027552.

[65] P. Artoisenet, E. Braaten, Production of the X(3872) at the Tevatron and the LHC, Phys.Rev. D81 (2010) 114018. arXiv:0911.2016, doi:10.1103/PhysRevD.81.114018.

[66] T. Himel, DECAYS OF THE $\psi'(3684)$ TO OTHER CHARMONIUM STATES, Ph.D. thesis (1979).

[67] Y.-P. Kuang, QCD multipole expansion and hadronic transitions in heavy quarkonium systems, Front.Phys.China 1 (2006) 19–37. arXiv:hep-ph/0601044, doi:10.1007/s11467-005-0012-6.

[68] L. S. Brown, R. N. Cahn, Chiral symmetry and $\psi' \to \pi\pi\psi$ decay, Phys. Rev. Lett. 35 (1975) 1–4. doi:10.1103/PhysRevLett.35.1.
URL http://link.aps.org/doi/10.1103/PhysRevLett.35.1

[69] A. Sokolov, et al. (Belle Collaboration), Observation of the decay $\Upsilon(4S) \to \Upsilon(1S)\pi^+\pi^-$, Phys.Rev. D75 (2007) 071103. arXiv:hep-ex/0611026, doi:10.1103/PhysRevD.75.071103.

[70] B. Aubert, et al. (BaBar Collaboration), Observation of $\Upsilon(4S)$ decays to $\pi^+\pi^-\Upsilon(1S)$ and $\pi^+\pi^-\Upsilon(2S)$, Phys.Rev.Lett. 96 (2006) 232001. arXiv:hep-ex/0604031, doi:10.1103/PhysRevLett.96.232001.

[71] F. Butler, et al. (CLEO Collaboration), Analysis of hadronic transitions in $\Upsilon(3S)$ decays, Phys.Rev. D49 (1994) 40–57. doi:10.1103/PhysRevD.49.40.

[72] F.-K. Guo, U.-G. Meiner, W. Wang, Production of the bottom analogues and the spin partner of the $X(3872)$ at hadron collidersarXiv:1402.6236.

[73] M. Karliner, private communication (2014).

[74] D. Ebert, R. Faustov, V. Galkin, Masses of heavy tetraquarks in the relativistic quark model, Phys.Lett. B634 (2006) 214–219. arXiv:hep-ph/0512230, doi:10.1016/j.physletb.2006.01.026.

[75] ATLAS Collaboration, Observation of a new particle in the search for the standard model higgs boson with the ATLAS detector at the LHC, Physics Letters B 716 (1) (2012) 1 – 29. doi:http://dx.doi.org/10.1016/j.physletb.2012.08.020. URL http://www.sciencedirect.com/science/article/pii/S037026931200857X

[76] CMS Collaboration, Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC, Phys.Lett. B716 (2012) 30–61. arXiv:1207.7235, doi:10.1016/j.physletb.2012.08.021.

[77] ATLAS Collaboration, Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC, Eur.Phys.J. C73 (2013) 2518. arXiv:1302.4393, doi:10.1140/epjc/s10052-013-2518-3.

[78] CMS Collaboration, The CMS experiment at the CERN LHC, Journal of Instrumentation 3 (08) (2008) S08004. URL http://stacks.iop.org/1748-0221/3/i=08/a=S08004

[79] ATLAS Collaboration, Letter of Intent for the Phase-I Upgrade of the ATLAS Experiment, Tech. Rep. CERN-LHCC-2011-012. LHCC-I-020, CERN, Geneva (Nov 2011).

[80] ATLAS Collaboration, Letter of Intent for the Phase-II Upgrade of the ATLAS Experiment, Tech. Rep. CERN-LHCC-2012-022. LHCC-I-023, CERN, Geneva, draft version for comments (Dec 2012).

[81] ATLAS Collaboration, The ATLAS Inner Detector commissioning and calibration, The European Physical Journal C 70 (3) (2010) 787–821. `doi:10.1140/epjc/s10052-010-1366-7`.
URL `http://dx.doi.org/10.1140/epjc/s10052-010-1366-7`

[82] C. Grupen, B. A. Shwartz, Particle Detectors, Cambridge University Press, 2008.

[83] P. Calafiura, W. Lavrijsen, C. Leggett, M. Marino, D. Quarrie, The athena control framework in production, new developments and lessons learned (2005) 456–458.

[84] T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, E. Moyse, A. Salzburger, Concepts, Design and Implementation of the ATLAS New Tracking (NEWT), Tech. Rep. ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002, CERN, Geneva (Mar 2007).

[85] A. Keith Morley, Electron bremsstrahlung studies and track based alignment of the atlas detector, Ph.D. thesis, Melbourne U. (2010).

[86] ATLAS Collaboration, Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung, Tech. Rep. ATLAS-COM-CONF-2012-068, CERN, Geneva (Apr 2012).

[87] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, G. Unal, Calorimeter clustering algorithms: Description and performance, Tech. Rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva (Apr 2008).

[88] ATLAS Collaboration, Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data, The European Physical Journal C 72 (3) (2012) 1–46. `doi:10.1140/epjc/s10052-012-1909-1`.
URL `http://dx.doi.org/10.1140/epjc/s10052-012-1909-1`

[89] CMS Collaboration, Prompt and non-prompt J/ψ production in pp collisions at $\sqrt{s} = 7$ TeV, The European Physical Journal C - Particles and Fields 71 (2011) 1–26, 10.1140/epjc/s10052-011-1575-8.
URL `http://dx.doi.org/10.1140/epjc/s10052-011-1575-8`

[90] D. Acosta, et al. (CDF Collaboration), Measurement of the J/ψ meson and b-hadron production cross sections in $p\overline{p}$ collisions at $\sqrt{s} = 1960$ GeV, Phys. Rev. D 71 (2005) 032001. `doi:10.1103/PhysRevD.71.032001`.
URL `http://link.aps.org/doi/10.1103/PhysRevD.71.032001`

[91] Pakhlov, P. and et al. (Belle Collaboration) , Measurement of the $e^+e^- \rightarrow J/\psi c\bar{c}$ cross section at $\sqrt{s} \sim 10.6$ GeV, Phys. Rev. D 79 (2009) 071101. `doi:10.1103/PhysRevD.79.071101`.
URL `http://link.aps.org/doi/10.1103/PhysRevD.79.071101`

[92] Aubert, B. and et al. (BaBar Collaboration), Measurement of J/ψ Production in Continuum e^+e^- Annihilations near $\sqrt{s} = 10.6$ GeV, Phys. Rev. Lett. 87 (2001) 162002. `doi:10.1103/PhysRevLett.87.162002`.
URL `http://link.aps.org/doi/10.1103/PhysRevLett.87.162002`

[93] T. Sjostrand, S. Mrenna, P. Z. Skands, PYTHIA 6.4 Physics and Manual, JHEP 0605 (2006) 026. `arXiv:hep-ph/0603175`, `doi:10.1088/1126-6708/2006/05/026`.

[94] A. Sherstnev, R. S. Thorne, Parton Distributions for LO Generators, EUR.PHYS.J.C 55 (2008) 553.
URL `doi:10.1140/epjc/s10052-008-0610-x`

[95] A. G. Buckley, ATLAS Monte Carlo tunes to LHC data, Tech. Rep. ATL-PHYS-PROC-2010-122, CERN, Geneva (Nov 2010).

[96] ATLAS Collaboration, The ATLAS Simulation Infrastructure. oai:cds.cern.ch:1267853, Eur. Phys. J. C 70 (arXiv:1005.4568. CERN-PH-EP-2010-044) (2010) 823–874. 53 p.

[97] J. Gaiser, Charmonium Spectroscopy From Radiative Decays of the J/ψ and ψ', Ph.D. thesis, Stanford University, SLAC-R-255 (1982).

[98] S. Cheatham, Electron specific track fitters and J/ψ resonance studies in ATLAS. oai:cds.cern.ch:1334202, Ph.D. thesis, Lancaster, presented 01 Jun 2010 (2010).

[99] ATLAS Collaboration, Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data, The European Physical Journal C - Particles and Fields 72 (2012) 1–46, 10.1140/epjc/s10052-012-1909-1.

[100] ATLAS Collaboration, Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC, Eur.Phys.J. C71 (2011) 1630. `arXiv:1101.2185`, `doi:10.1140/epjc/s10052-011-1630-5`.

[101] S. Godfrey, J. L. Rosner, Production of the D-wave $b\bar{b}$ states, Phys. Rev. D 64 (hep-ph/0105273.) (2001) 097501.

[102] ATLAS Collaboration, ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11, Tech. Rep. ATL-PHYS-PUB-2011-009, CERN, Geneva (Jul 2011).

[103] W. Lukas, Fast Simulation for ATLAS: Atlfast-II and ISF, Tech. Rep. ATL-SOFT-PROC-2012-065, CERN, Geneva (Jun 2012).

[104] G. Cowan, K. Cranmer, E. Gross, O. Vitells, Asymptotic formulae for likelihood-based tests of new physics, Eur.Phys.J. C71 (2011) 1554. `arXiv:1007.1727`, `doi:10.1140/epjc/s10052-011-1554-0`.

[105] ATLAS Collaboration, Improved luminosity determination in pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector at the LHC, Eur.Phys.J. C73 (2013) 2518. `arXiv:1302.4393`, `doi:10.1140/epjc/s10052-013-2518-3`.

[106] M. Voloshin, V. Zakharov, Measuring quantum-chromodynamic anomalies in hadronic transitions between quarkonium states, Phys. Rev. Lett. 45 (1980) 688–691. `doi:10.1103/PhysRevLett.45.688`.
URL `http://link.aps.org/doi/10.1103/PhysRevLett.45.688`

[107] G. Bélanger, T. DeGrand, P. Moxhay, Spectra of the transitions $\Upsilon(nS) \rightarrow \Upsilon(mS)\pi^+\pi^-$, Phys. Rev. D 39 (1989) 257–265. `doi:10.1103/PhysRevD.39.257`.
URL `http://link.aps.org/doi/10.1103/PhysRevD.39.257`

[108] C.-F. Qiao, F. Yuan, K.-T. Chao, Gluon fragmentation to ^3D_J quarkonia, Phys. Rev. D 55 (1997) 5437–5444. `doi:10.1103/PhysRevD.55.5437`.
URL `http://link.aps.org/doi/10.1103/PhysRevD.55.5437`

[109] C. Cuthbert, B. Yabsley, Search for the X_b and other hidden-beauty states using the $\pi^+\pi^-\Upsilon$(1S) channel at ATLAS (Internal Documentation), Tech. Rep. ATL-COM-PHYS-2013-1413, CERN, Geneva (Oct 2013).

[110] ATLAS Collaboration, Summary of ATLAS Pythia 8 tunes, Tech. Rep. ATL-PHYS-PUB-2012-003, CERN, Geneva (Aug 2012).

[111] J. Pumplin, D. Stump, J. Huston, H. Lai, P. M. Nadolsky, et al., New generation of parton distributions with uncertainties from global QCD analysis, JHEP 0207 (2002) 012. `arXiv:hep-ph/0201195`, `doi:10.1088/1126-6708/2002/07/012`.

[112] CMS Collaboration, $\Upsilon(1S)$, $\Upsilon(2S)$ and $\Upsilon(3S)$ cross section measurements in pp collisions at $\sqrt(s) = 7$ TeV, Tech. Rep. CMS-PAS-BPH-12-006, CERN, Geneva (2013).

[113] CMS Collaboration, Upsilon production cross section in pp collisions at $\sqrt{s} = 7$ TeV, Phys. Rev. D 83 (2011) 112004. `doi:10.1103/PhysRevD.83.112004`.
URL `http://link.aps.org/doi/10.1103/PhysRevD.83.112004`

[114] J. Bai, et al. (BES Collaboration), $\psi(2S) \to \pi^+\pi^- J/\psi$ decay distributions, Phys.Rev. D62 (2000) 032002. `arXiv:hep-ex/9909038`, `doi:10.1103/PhysRevD.62.032002`.

[115] Lees, J. P. and et al. (BaBar Collaboration), Study of the reaction $e^+e^- \to J/\psi\pi^+\pi^-$ via initial-state radiation at BaBar, Phys. Rev. D 86 (2012) 051102. `doi:10.1103/PhysRevD.86.051102`.
URL `http://link.aps.org/doi/10.1103/PhysRevD.86.051102`