Measurements of $\psi(2S)$ and $X(3872) \to J/\psi\pi^+\pi^-$ production in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector

Michael David Beattie

Department of Physics

Lancaster University

A thesis submitted to Lancaster University for the degree of Doctor of Philosophy in the Faculty of Science and Technology

August 2016

Declaration

I declare that the work in this thesis is my own, except where clearly stated in the text, and that no portion of this work has been submitted for consideration for any other degree.

Michael David Beattie

August 2016

Acknowledgements

I would firstly like to thank my supervisor, Vato Kartvelishvili, for his patience and guidance throughout my studies. I would also like to thank Lee Allison and James Walder, for helping with my many coding and ATLAS software issues, and also to Cameron Cuthbert and Bruce Yabsley for providing invaluable help and expertise. I also want to more generally thank the Lancaster University Experimental Particle Physics group, who have all helped me in one way or another during my time at Lancaster. I am grateful to the B-physics and Onia group convenors, particularly Darren Price and Andy Chisholm, for their interest and input throughout my research. I am also grateful to all those who made my LTA at CERN such a memorable experience. I wish to thank the Science and Technology Facilities Council (STFC) for their funding, without which it would not have been possible for me to work on this research. Lastly, I want to thank my parents, whose support and belief in me throughout my studies has been absolute.

Abstract

Differential cross sections are presented for the prompt and non-prompt production of the hidden-charm states $X(3872)$ and $\psi(2S)$, in the decay mode $J/\psi\pi^+\pi^-$, measured using 11.4 fb^{-1} of pp collisions at $\sqrt{s} = 8$ TeV by the ATLAS detector at the LHC. The ratio of cross sections $X(3872)/\psi(2S)$ is also given, separately for prompt and non-prompt components, as well as the non-prompt fractions of $X(3872)$ and $\psi(2S)$. Assuming independent single effective lifetimes for non-prompt $X(3872)$ and $\psi(2S)$ production gives $R_B = \frac{Br(B\to X(3872)\text{ + any})Br(X(3872)\to J/\psi\pi^+\pi^-)}{Br(B\to\psi(2S)\text{ + any})Br(\psi(2S)\to J/\psi\pi^+\pi^-)} = (3.95 \pm 0.32(\text{stat}) \pm 0.08(\text{sys}))\%$, while separating short and long-lived contributions, assuming that the short-lived component is due to B_c decays, gives $R_B = (3.57 \pm 0.33(\text{stat}) \pm 0.11(\text{sys}))\%$, with the fraction of non-prompt $X(3872)$ produced via B_c decays for $p_T(X(3872)) > 10$ GeV being $(25 \pm 13(\text{stat}) \pm 2(\text{sys}) \pm 5(\text{spin}))\%$. The distributions of the dipion invariant mass in the $X(3872)$ and $\psi(2S)$ decays are also measured and compared to expectations.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

The J/ψ meson and its excited state $\psi(2S)$ were among the first heavy quark-antiquark bound states, known as quarkonium, to be discovered in the 1970s. Since then, their production measurements have acted as excellent 'standard candles' to compare theoretical predictions of quarkonium production. The exotic state $X(3872)$ was later discovered in 2003 [1], and its existence was subsequently confirmed by several other experiments [2–4]. This was followed by the discoveries of a variety of other X, Y, Z states, which cannot be explained as 'traditional' quarkonium due to the combination of their measured quantum numbers and masses [5].

The construction of the Large Hadron Collider (LHC) along the French-Swiss border at CERN allows measurements to be performed at world-record breaking collision energies. The ATLAS experiment, one of the general purpose detectors at the LHC, has already provided a wide variety of quarkonium measurements, and previously observed the $X(3872)$ state during production measurements of the $\psi(2S)$ meson in 2011 [6]. Measurements of $X(3872)$ production have also been performed by the LHCb [7] and CMS [8] experiments at the LHC.

In this analysis, production measurements of the $\psi(2S)$ and $X(3872)$ states in the decay channel $J/\psi\pi^+\pi^-$ are performed, using 11.4 fb^{-1} of proton-proton collision data collected by the ATLAS experiment at the LHC at a centre-of-mass energy $\sqrt{s} = 8$ TeV. Chapter 2 provides a theoretical overview of quarkonium and the current status of $X(3872)$. In Chapter 3 there is a summary of the LHC, followed by a description of the ATLAS detector and its various components. Chapter 4 provides a more detailed description of the B-physics triggers, and details a specific B-physics trigger efficiency measurement. The main analysis method, explaining the $\psi(2S)$ and $X(3872)$

production measurements, is in Chapter 5, with the corresponding systematic error evaluation and results shown in Chapter 6. We describe measurements of the dipion invariant mass spectra for $\psi(2S)$ and $X(3872)$ decays in Chapter 7. Finally, a summary of the results is in Chapter 8. Supplementary material is shown in Appendices A - I.

The main analysis presented in this thesis has been published as an ATLAS conference note [9], and has been presented at the Quarkonium Working Group International Workshop QwG2016, and at the ICHEP2016 international conference. The dimuon efficiency measurement also presented in this thesis was used for various published ATLAS quarkonia analyses using 2011 data, for instance [10].

Chapter 2

Theory and Background

This chapter first provides a brief overview of the Standard Model of particle physics in Section 2.1. The theory of the strong interaction between quarks, known as Quantum Chromodynamics, is described in Section 2.2. In Section 2.3 the charmonium spectrum is described, and the current status of charmonium production, focusing on the J/ψ and $\psi(2S)$ states, is discussed. Finally, the physical interpretation and experimental results of the $X(3872)$ are summarised in Section 2.4.

2.1 The Standard Model

The Standard Model (SM) is the most successful theory to describe the way in which sub-atomic particles behave, through the electromagnetic, weak and strong nuclear interactions. The fundamental particles that make up matter in the Universe exist as either *quarks* or *leptons*. These are known as *fermions*, which have a half-integer spin. For each elementary matter particle, there also exists its corresponding anti-matter particle. There are three generations of quarks and leptons, which are shown in Tables 2.1 and 2.2, respectively. Quarks are confined to bound states, for instance mesons ($q\bar{q}$), baryons (qqq) and also newly hypothesised tetraquark [11] and pentaquark states [12], whereas leptons are free to exist individually.

The interactions between the fermions are governed by fundamental forces mediated by *gauge bosons*, which have an integer spin. The three fundamental forces in the SM are: electromagnetism, and the strong and weak nuclear forces, each mediated by respective gauge bosons shown in Table 2.3. A fourth force, which is not described by

the SM, is gravity. The strength of the gravitiational force is measured to be many orders of magnitude lower than the three fundamental forces considered in the SM, and so its omission presents no practical effect in pp collisions.

The way in which particles acquire mass in the SM is by spontaneous symmetry breaking through the *Brout-Englert-Higgs* mechanism [13–15], which requires the existence of a Higgs boson with zero spin, which was discovered in 2012 jointly by the ATLAS and CMS experiments at the LHC [16, 17]. The SM is described by a non-abelian gauge theory containing the unitary composite symmetry group $SU(3) \bigotimes SU(2) \bigotimes U(1)$. The $SU(2) \bigotimes U(1)$ group describes the electroweak interaction, while $SU(3)$ describes the strong interaction between quarks and gluons.

Generation I	Generation II	Generation III
up (u)	charm (c)	top (t)
$Q_u = +\frac{2}{3}$	$Q_c = +\frac{2}{3}$	$Q_t = +\frac{2}{3}$
$m_u = 2.3^{+0.7}_{-0.5}$ MeV	$m_c = 1.275 \pm 0.025$ GeV	$m_t = 173.21 \pm 0.51 \pm 0.71$ GeV
down (d)	strange (s)	bottom (b)
$Q_d = -\frac{1}{3}$	$Q_s = -\frac{1}{3}$	$Q_b = -\frac{1}{3}$
$m_d = 4.8^{+0.5}_{-0.3}$ MeV	$m_s = 95 \pm 5$ MeV	$m_b = 4.18 \pm 0.03$ GeV

Table 2.1 Selected properties of the quarks. The electric charge Q is shown in units of absolute electron charge. Values are taken from [18].

2.2 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is a non-abelian gauge theory, which describes the strong interaction and is described by the $SU(3)$ symmetry group. QCD initially appears very similar to Quantum Electrodynamics (QED), which describes the electromagnetic interaction. Instead of the interaction being mediated by a photon, which corresponds to the single generator of the $U(1)$ symmetry group, it is mediated by eight gluons corresponding to the eight generators of the $SU(3)$ symmetry group. Compared to the single electric charge in QED, there are three 'colour' charges in QCD : *red*, *green* and *blue*, which are the orthogonal states in $SU(3)$ colour space. The $SU(3)$ colour symmetry is exact, and so the strength of the QCD interaction

Generation I	Generation II	Generation III
electron (e)	muon (μ)	tau (τ)
$Q_e = -1$	$Q_\mu = -1$	$Q_\tau = -1$
$m_e = 0.51100$ MeV	$m_\mu = 105.66$ MeV	$m_\tau = 1776.86 \pm 0.12$ MeV
electron neutrino (ν_e)	muon neutrino (ν_μ)	tau neutrino (ν_τ)
$Q_{\nu_e} = 0$	$Q_{\nu_\mu} = 0$	$Q_{\nu_\tau} = 0$
$m_{\nu_e} < 2$ MeV	$m_{\nu_\mu} < 0.19$ MeV	$m_{\nu_\tau} < 18.2$ MeV

Table 2.2 Selected properties of the leptons. The electric charge Q is shown in units of absolute electron charge. Though neutrinos are expected to be massless in the SM, experiment suggests they have mass. The neutrino mass upper limits are shown at a 95% confidence level. Values are taken from [18].

Name	Mass [GeV]	Spin	Charge	Force
γ (photon)	0	1	0	Electromagnetic
Z^0	91.1876 ± 0.0021	1	0	Weak
$W^\pm$	80.385 ± 0.015	1	± 1	Weak
g (gluon)	0	1	0	Strong
H (Higgs)	$125.09 \pm 0.21 \pm 0.11$	0	0	-

Table 2.3 Selected properties of the bosons in the SM. The electric charge is shown in units of absolute electron charge. Values are taken from [18].

does not depend on the colour charge [19]. Similarly to QED, where anti-particles carry opposite electric charge, in QCD anti-quarks carry opposite colour charge. The symmetry group $SU(3)$ is non-abelian, meaning a key difference between QED and QCD is that the gluons themselves have colour charge, so unlike the photon they are self-coupling. Another important property of QCD is asymptotic freedom. That is, the binding between quarks becomes asymptotically weaker at high energy scales or small distances such that they can be considered as free particles, which subsequently allows the use of perturbation theory for calculations and predictions of production in high-energy collisions. This arises due to the strong coupling constant α_S having an explicit dependence on the energy scale Q as

$$\alpha_S(Q^2) = \frac{4\pi}{\beta_0 \log(Q^2/\Lambda^2_{\text{QCD}})}, \tag{2.1}$$

at leading order, where β_0 is the leading-order coefficient of the beta-function for QCD, which describes the behaviour of α_S at various energy scales and is equal to $11 - 2n_f/3$ [19], where n_f is the number of quark flavours. For the six flavours $n_f = 6$ in the SM, β_0 is positive, leading to asymptotic freedom primarily due to anti-screening effects from gluon self-interactions. The QCD scale Λ_{QCD} defines the scale at which α_S can be considered small enough for perturbation theory and is determined experimentally to be around 100 MeV.

2.3 Quarkonium

Quarkonium denotes the bound state of a $q\bar{q}$ pair by the strong interaction. They are the subset of a group of particles called mesons, which are bosons consisting of a bound $q\bar{q'}$ pair which may have different flavours. Quarkonium can be the bound state of any of the (u, d, c, s, b) quarks with their corresponding anti-quark. However, it is not possible for the top and anti-top quarks to form a quarkonium state because the very heavy top mass gives it a large decay width, with a correspondingly short lifetime, so it decays weakly before quarkonium can form. The quarkonium states can be further categorised: the large mass of the c and b quarks forming 'heavy quarkonium' allow for significantly different properties than quarkonium made from lighter quarks. The bound $c\bar{c}$ state is named 'charmonium', while the $b\bar{b}$ state is named 'bottomonium'.

2.3.1 Spectroscopy

The properties of heavy quarkonium states have been measured in high energy collider experiments for decades. The first heavy quarkonium state to be observed in experiment was the J/ψ in 1974, jointly by experiments at the Brookhaven National Laboratory (BNL) [20] and the Stanford Linear Accelerator Centre (SLAC) [21], through decays to e^+e^- and $\mu^+\mu^-$, as a narrow resonance with a mass around 3.1 GeV. The J/ψ is a charmonium vector state with the quantum numbers $J^{PC} = 1^{--}$ and a very narrow decay width, leading it to become an ideal 'standard candle' for quarkonium studies since its discovery. The bottomonium analogue to the J/ψ, named $\Upsilon(1S)$, was discovered a few years later in the decay channel $\Upsilon(1S) \to \mu^+\mu^-$ at Fermilab [22], with a mass around 9.5 GeV.

The quantum numbers typically assigned to quarkonium are: the total spin of the $q\bar{q}$ system S, the orbital angular momentum L, and the total angular momentum J (where $\vec{J} = \vec{L} + \vec{S}$). In quarkonium spectroscopy, the notation $n^{2S+1}L_J$ is usually used to describe the quarkonium states, where n is defined as the principal quantum number. The P-parity quantum number of the $q\bar{q}$ system is defined as $P = (-1)^{L+1}$, and the charge conjugation, or C-parity quantum number, is defined as $C = (-1)^{L+S}$. A diagram showing the spectrum of charmonium mass states and their typical decays to lower mass states is shown in Figure 2.1. The charmonium spectrum includes the pseudoscalar 0^{-+} state η_c and the vector ground state J/ψ, their radially excited counterparts η' and $\psi(2S)$, and the C-even, orbital excitations $\eta_{c0,1,2}$.

We also see in this diagram the 'open charm' threshold. It is defined as the addition of the mass of the lightest meson consisting of a charm quark, which is the D^0 meson ($\bar{u}c$), and its corresponding anti-particle $\bar{D^0}$ ($u\bar{c}$). Charmonium states above the open charm threshold have broad resonances, and decay into other, lighter charmonium states.

The masses of the states in the charmonium spectrum, ignoring any dependence on spin, can be estimated using a potential model, which has a dependence on the distance r between the quark and anti-quark. At small distances, a (leading-order) potential with the same form as the Coulomb potential can be used:

$$V(r) = -\frac{4}{3}\frac{\alpha_s}{r}, \tag{2.2}$$

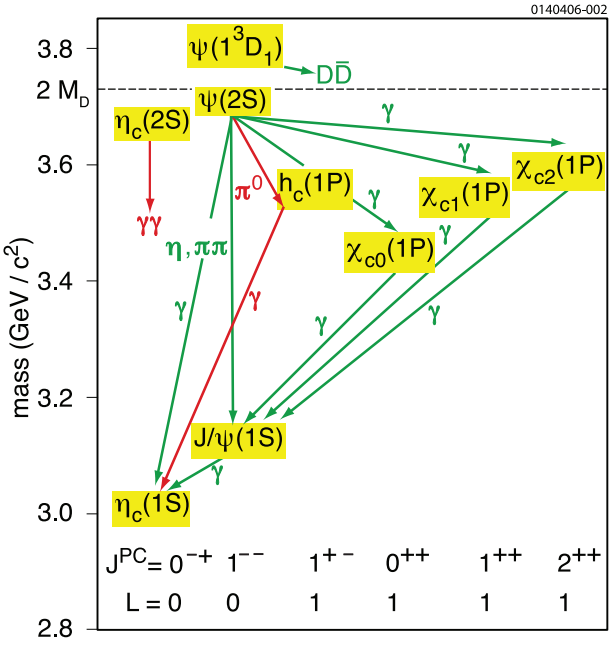


Fig. 2.1 The charmonium spectrum, including transitional decays between states. Also shown is the 'open-charm' threshold, above which resonances are expected to be broad [23].

and at large distances a confining potential dominates,

$$V(r) = Kr, \tag{2.3}$$

where K is the QCD string tension, equal to roughly 1 GeV/fm [19]. This is known as the Cornell potential [24] and is commonly used to model quarkonium states, but other models have also been used. The relatively large mass of the c quark in comparison to the Λ_{QCD} scale allows the subsequent use of non-relativistic quantum mechanics to model the charmonium state and estimate the mass of the system. Corrections can subsequently be made to account for spin-spin interactions [25]. The masses of the charmonium states can also be calculated directly using QCD, through the use of lattice QCD with a non-relativistic QCD (NRQCD) effective field theory, with generally good agreement to the measurements [26]. The heaviest charmonium state measured to be below the open charm threshold is the $\psi(2S)$ meson, with the result that it contains no or little feed-down from higher mass states, and so is seen as a cleaner state for production studies.

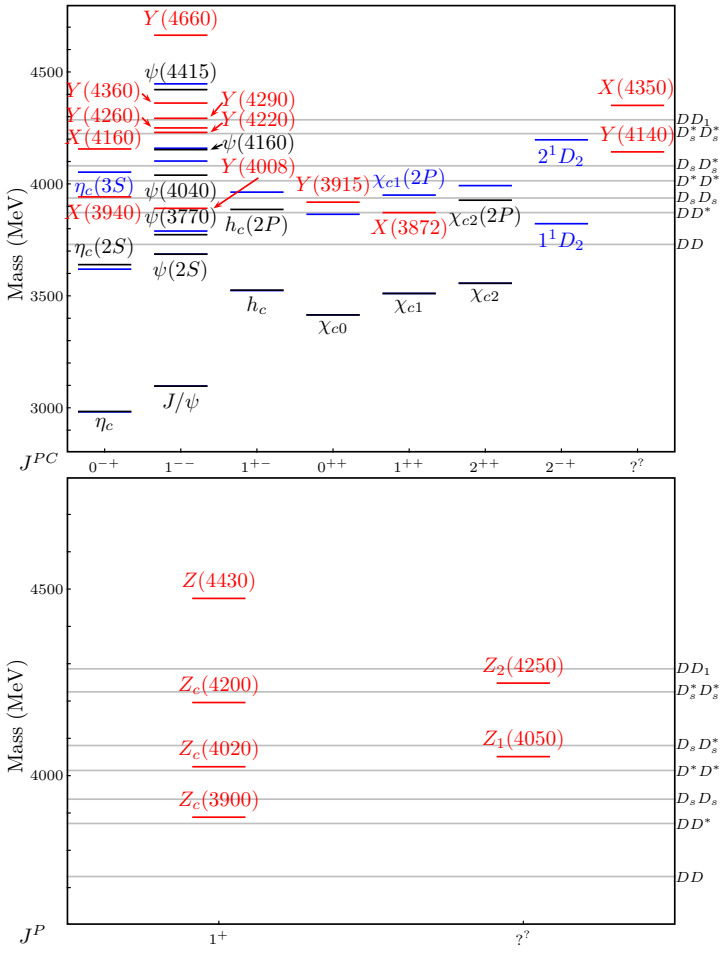


Fig. 2.2 The upper panel shows the charmonium spectrum, including exotic states indicated in red. The blue lines represent quark mode predictions. The lower panel shows the charged exotic states [27]. More recently, the exotic states $X(4274)$, $X(4500)$ and $X(4700)$ have been observed by LHCb [28].

In 2003, an exotic resonance was discovered by the Belle collaboration [1] in the $J/\psi\pi^+\pi^-$ invariant mass spectrum, and was later confirmed by several other experiments [2–4]. More details of the $X(3872)$ can be found in Section 2.4. This was followed by the discovery of a variety of other so-called X, Y and Z states which cannot be explained as 'traditional' quarkonium due to the combination of their measured quantum numbers and masses [5]. The charmonium spectrum, now including the exotic states and decay patterns, can be seen in Figure 2.2. Also shown are the various DD thresholds, which are relevant as some of these exotic states are often theorised to be loosely bound DD molecules of various types.

2.3.2 Production

Quarkonium can be produced through hadron-hadron interactions in two main ways. Prompt quarkonium are produced from QCD sources, and from the subsequent feed-down from higher mass states, which are produced from QCD sources. Non-prompt quarkonium are produced from the decay of long-lived sources such as b-hadrons produced from the hadronic collisions. Theoretical models have been produced which describe the cross sections and spin-alignments of the quarkonium, separately for prompt and non-prompt production processes, which are described in the following sections.

2.3.3 Prompt Quarkonium Production

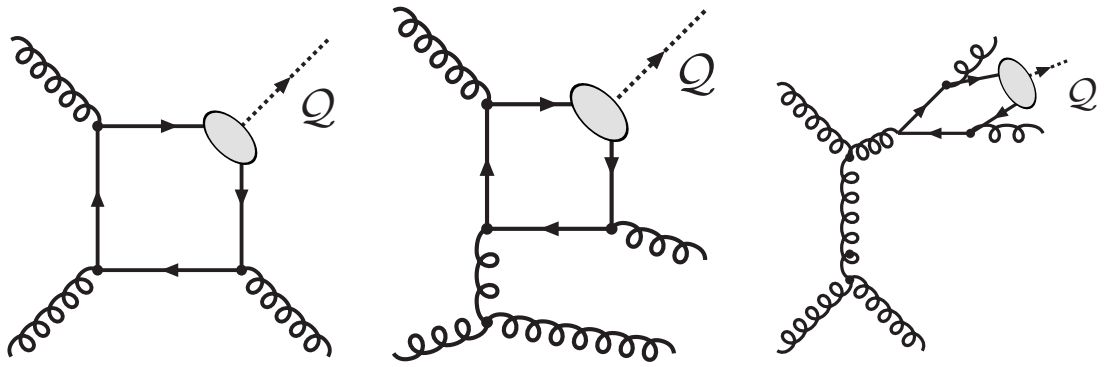
Colour Singlet Model

The Colour Singlet Model (CSM) is one of the earliest theories to model quarkonium production. It can be factorised into two independent processes [29]. The first is the initial production of the $q\bar{q}$ pair, which happens at short relative distances, so can be calculated with perturbative QCD. The second, is the formation of the $q\bar{q}$ pair into a bound state. This is a low energy, non-perturbative process, but can be parameterised into a constant according to the initial wavefunction of the quarkonium state, assuming the static approximation [30] (i.e. the constituent quarks are at rest in the quarkonium frame). The partonic cross section of a heavy quarkonium state H in the CSM can therefore be expressed as

$$\sigma_{\mathrm{CSM}}[ij \to H + X] = \sigma[ij \to q\bar{q} + X] \left| \frac{d\Psi_{nL}(0)}{dr^L} \right|^2 , \tag{2.4}$$

where $\sigma[ij \to q\bar{q} + X]$ is the partonic cross section of the $q\bar{q}$ pair produced from interactions of partons i and j, and $d\Psi_{nL}(0)$ is the quarkonium wavefunction evaluated at the origin, which can be extracted either from potential models (detailed in Section 2.3.1) or through experiment. Aside from the parton distribution function [31], which governs the internal structure of the proton at a given energy scale, the quarkonium wavefunction is the only free term in the CSM, meaning it is strongly predictive. The main assumption of the CSM is that the $q\bar{q}$ pair is produced with the same colour and spin as the quarkonium in the final state. This means that a $q\bar{q}$ pair must be produced in a

colour singlet state, due to the requirement of the final state quarkonium to be colour singlet. In the high-energy conditions of modern collider experiments, quarkonium production is dominated by gluon fusion processes [32]. Gluon fusion diagrams for producing ${}^3S_1$ states at various orders of α_S, consistent within the CSM framework, are seen in Figure 2.3.

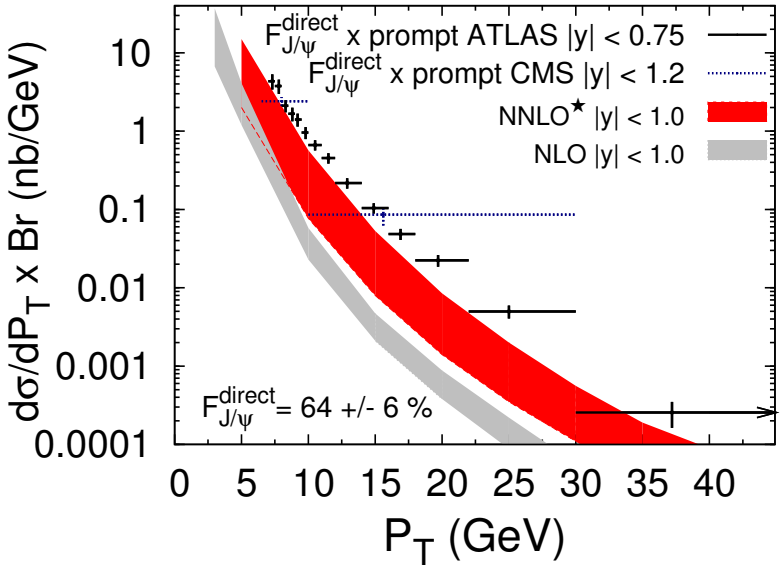


(a) (b) (c)

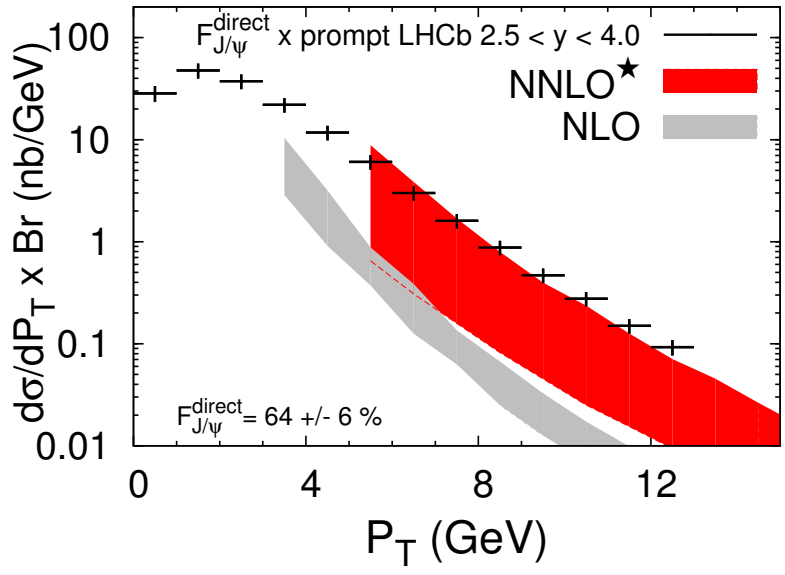
Fig. 2.3 Production of ${}^3S_1$ quarkonium states through gluon fusion processes at (a) LO (α_S^3), (b) NLO (α_S^4) and (c) NNLO (α_S^5) consistent with the Colour Singlet Model [33].

By the early 1990's, measurements of cross sections for large-p_T prompt J/ψ and $\psi(2S)$ samples using data from the Tevatron showed stark differences in both overall normalisation and as a function of p_T to the CSM predictions at leading order [34]. Further attempts have been made to explain the data within the CSM framework by including higher order contributions (at NLO, NNLO*). Although increasing the power of α_S, some of these higher-order contributions need not be small, for instance the NNLO gluon fragmentation channel at order α_S^5 (see Figure 2.3(c)) has a p_T^4 enhancement compared to leading order diagrams [33], so is expected to dominate over leading order production at high-p_{T}. However, whilst NLO and NNLO contributions have improved the p_{T} dependence and normalisation of the CSM with respect to data, it is still not in agreement. The comparison of the CSM at NNLO* with J/ψ production measurements from ATLAS, CMS and LHCb at $\sqrt{s} = 7$ TeV for a variety of rapidity regions is in Figure 2.4, and shows the CSM still underestimates direct J/ψ production. In Figure 2.4(b), it is seen that the cross section measured by LHCb

decreases at very low p_{T}. This is due to the non-zero initial transverse momenta of the colliding partons.



(a)



(b)

Fig. 2.4 The Colour Singlet Model at NNLO* order, compared to direct J/ψ cross sections measured differentially in p_{T} by (a) ATLAS and CMS for the central rapidity regions and (b) LHCb at a more forward rapidity region [32].

Colour Evaporation Model

Along with the CSM, another early model which attempts to explain quarkonium production is the Colour Evaporation Model (CEM), which was proposed in the late 1970's [35]. Like the CSM, the CEM had reasonable success in explaining measurements until the mid-1990's when disagreement had emerged at higher collision energies. The key difference is that the initial production of the $q\bar{q}$ pair in the CEM can be produced in a coloured state, and in some versions of the theory a spin state [36], different to that of the quarkonium state it will eventually form through soft interactions with the colour field [37], which is the origin of the term 'colour evaporation'. These soft gluon interactions are expected to have a negligible effect on the kinematics of the $q\bar{q}$ pair. Hence, in the CEM, every heavy $q\bar{q}$ pair with an invariant mass below the open-flavour threshold will form some quarkonium state. The leading-order cross section of a quarkonium state H from pp collisions is given in the CEM as

$$\sigma_{\mathrm{CEM}}[pp \to H + X] = F_H \int_{4m_q^2}^{4M^2} dm_{q\bar{q}}^2 \frac{d\sigma}{dm_{q\bar{q}}^2}[pp \to q\bar{q} + X]. \tag{2.5}$$

The probability of a produced $q\bar{q}$ pair forming a heavy quarkonium state H, commonly expressed as F_H, is assumed to be production-process and kinematically independent, meaning once F_H for a given quarkonium state is measured, it can be used to predict other production processes in different kinematic regions [38]. M is the mass of the lightest meson which contains the heavy quark q, m_q is the mass of the heavy quark, $m_{q\bar{q}}$ is the invariant mass of the $q\bar{q}$ pair, and $\frac{d\sigma}{dm_{q\bar{q}}^2}$ is the inclusive differential cross section for the $q\bar{q}$ pair to be produced from the pp collision. It is in this term where the assumptions of the CEM i.e. colour evaporation and spin randomisation are made [37] and it can typically be calculated using perturbation theory.

The Colour-Octet-Mechanism (COM) and NRQCD Factorisation

The Colour-Octet-Mechanism (COM) was developed in the early 1990's [39], and contains elements of both the CSM and CEM in its construction. It expands upon the CSM idea of factorising the production of quarkonium into high-energy/short-distance (typically in the hard-scattering processes of the $q\bar{q}$ pair production) and low-energy/long-distance (typically in the transition of the $q\bar{q}$ pair into the quarkonium state) components. The factorisation can be explained using an effective field theory

known as NRQCD, which approximates full QCD in a non-relativistic limit. It is based on the momentum transfer scale $m_q v$ (the relative momentum of the heavy quark in the centre-of-mass frame). Other important energy-momentum scales include the mass of the heavy quark m_q and the kinetic energy $m_q v^2$. For the non-relativistic approximation to be valid, one requires:

$$\left(m_q v^2\right)^2 << \left(m_q v\right)^2 << m_q^2 \tag{2.6}$$

which is satisfied when $v^2 \approx 0.25$ and $v^2 \approx 0.1$ for charmonium and bottomonium, respectively.

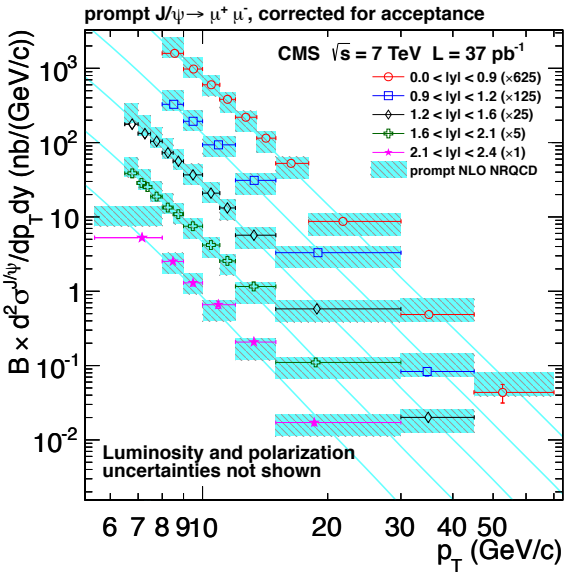
Similarly to the CEM, the $q\bar{q}$ pair need not be produced with the same quantum numbers as the final quarkonium state, but can end up in the final state through the non-perturbative emission of soft gluons. The inclusive cross section for the direct production of a heavy quarkonium state H can be expressed as a sum over short-distance coefficients and long distance matrix elements (LDMEs):

$$\sigma_{\mathrm{COM}}[pp \to H + X] = \sum_n \sigma_n[pp \to q\bar{q} + X](\Lambda) \left\langle \mathcal{O}_n^H(\Lambda) \right\rangle \tag{2.7}$$

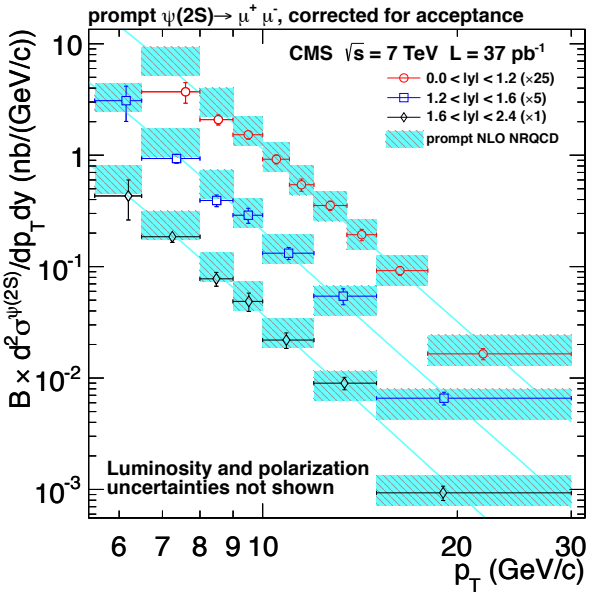
where Λ is the ultra-violet cut-off of the NRQCD effective theory, $\sigma_n[pp \to q\bar{q} + X]$ are the short-distance production coefficients for producing a $q\bar{q}$ with a certain set of quantum numbers and can usually be calculated perturbatively, and $\left\langle \mathcal{O}_n^H(\Lambda) \right\rangle$ are the LDMEs, which describe the probability of a $q\bar{q}$ pair with a certain set of quantum numbers to transition into the quarkonium final state. They are expected to be phenomological constants, determined through fits to experimental data, and once found can be used to predict production rates in other processes. Typical values for the matrix elements, found from fits to CDF data, can be found in [40].

The NRQCD factorisation framework is currently the most successful model for describing quarkonium production, possibly due to the large number of fittable parameters. Comparison with data from CMS [41] is shown in Figure 2.5 for J/ψ and $\psi(2S)$ production at $\sqrt{s} = 7$ TeV across a wide range of p_{T} and rapidity. Good agreement is seen across the whole p_{T} and rapidity range. Comparison to ATLAS data [10] in Figure 2.6, also at $\sqrt{s} = 7$ TeV but with a wider p_{T} range, again shows overall good agreement across the whole p_{T} and rapidity region, although there is a possible slight overestimation of the data in the largest p_{T} bins. The NRQCD theory is also compared

to LHCb data in a more forward rapidity region $2.0 < y < 4.5$ for prompt J/ψ [42] and $\psi(2S)$ [43] production, shown in Figure 2.7, where again reasonable agreement is seen.

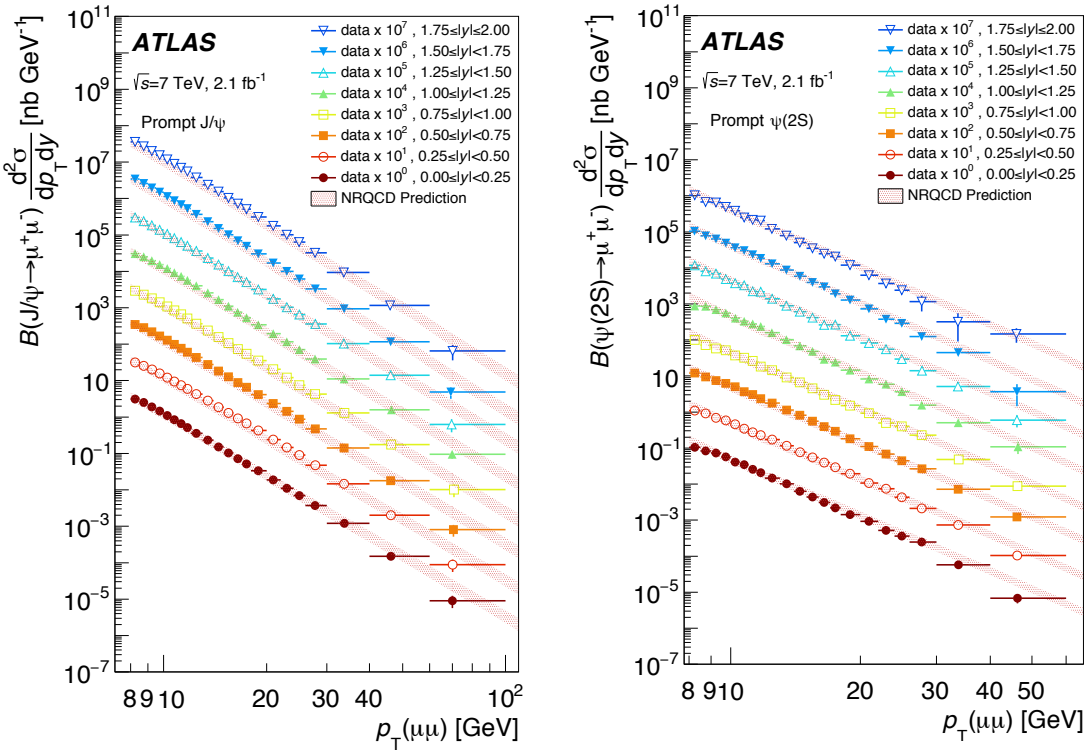


(a)



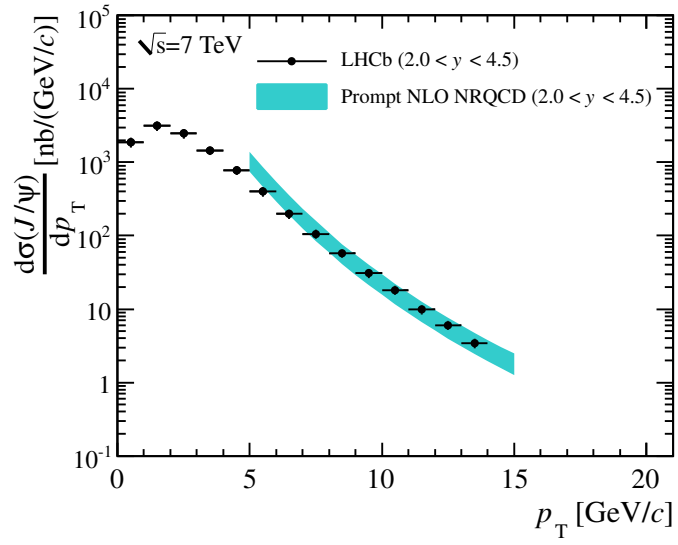
(b)

Fig. 2.5 p_{T} differential cross sections measured by CMS compared to NRQCD for (a) prompt J/ψ and (b) prompt $\psi(2S)$ production at $\sqrt{s} = 7$ TeV for a range of rapidity regions [41].

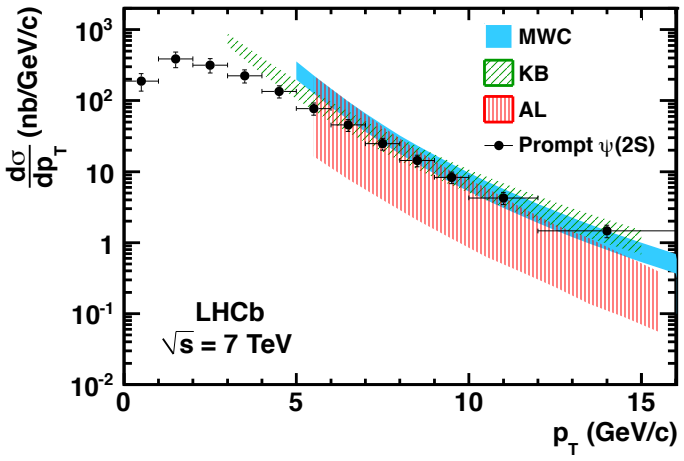


(a) (b)

Fig. 2.6 p_{T} differential cross sections measured by ATLAS compared to NRQCD for prompt (a) J/ψ and (b) $\psi(2S)$ production at $\sqrt{s} = 7$ TeV for a range of rapidity regions [10].



(a)



(b)

Fig. 2.7 p_{T} differential cross sections measured by LHCb compared to NRQCD for (a) prompt J/ψ [42] and (b) prompt $\psi(2S)$ [43] production at $\sqrt{s} = 7$ TeV for the forward rapidity region $2.0 < y < 4.5$. MWC [44] and KB [45] are NLO NRQCD calculations including colour-singlet and colour-octet contributions, while AL [32] is a colour-singlet model including the dominant NNLO terms.

k_T Factorisation

The k_T factorisation model uses an alternative approach to calculate hadronic cross sections. The transverse momentum k_T of the initial partons is much smaller than their longitudinal momentum, hence it is often assumed to be zero. This is called collinear factorisation. Instead, the k_T factorisation approach uses parton-level cross sections calculated with the CSM [46], or non-perturbative colour-octet matrix elements deduced from fits to data [47], and supplements these with transverse momentum dependent (unintegrated) PDFs to include the effects of initial gluon radiation and intrinsic gluon transverse momentum. These transverse momentum dependent (TMD) gluon densities are not well constrained phenomenologically, however, and numerous sets of distributions exist (see [47]).

2.3.4 Charmonium Production from b-hadron Decays

Charmonium states are produced in significant fractions through the decays of long-lived b-hadrons produced in high-energy hadronic collisions. This is due to a combination of the large b-hadron cross section, and branching fractions to charmonium states of the order 1% [18]. Typically, b-hadrons such as $B^{\pm}$, B^0, B_s^0 and Λ_b, along with excited B^* states which decay into these, are produced in the pp collisions, which then decay into final states associated with charmonium. Although the B_c cross section is measured to be small [18] in comparison to other b-hadrons, charmonium states produced from B_c decays have been observed and studied by experiment, for instance in the decay channel $B_c \to J/\psi D_s^+$ [9, 48]. Charmonium states produced through b-hadron decays are separable experimentally from promptly produced charmonium (from QCD sources) by a displaced vertex relative to the primary collision point, due to the long lifetime of most b-hadrons at around 10^{-12} seconds [18].

FONLL

In the Fixed Order + Next-to-Leading Log (FONLL) model, the b-hadron production cross section is found by matching NLO QCD with an all-order resummation to next-to-leading log accuracy in the limit where the heavy quark p_{T} is much greater than its mass. The logarithms resummed to NLL accuracy are of the form $\alpha_S^n \log^n(p_{\mathrm{T}}/m)$ and $\alpha_S^n \log^{(n-1)}(p_{\mathrm{T}}/m)$ [49].

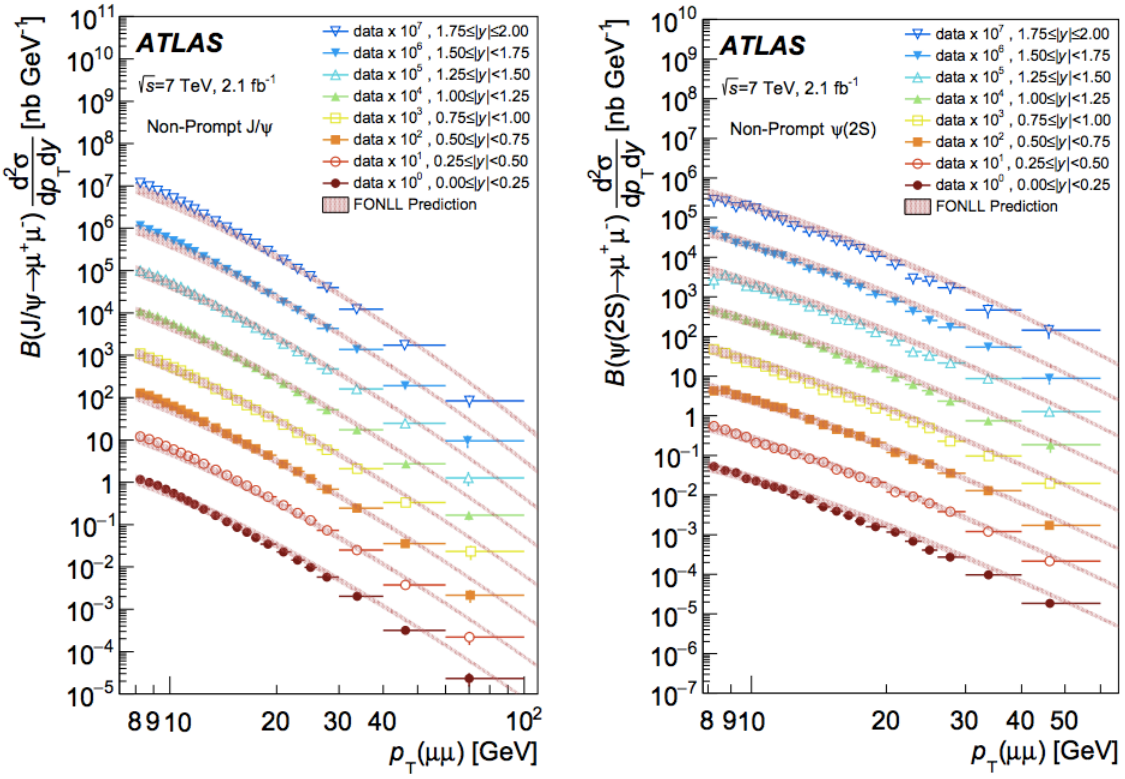
The b-hadron cross section is found from two processes. Firstly, the production cross section of the b quark $d\sigma_b^{FONLL}$ is calculated, which is then convoluted by a non-perturbative fragmentation function $f(b \to H_b)$ describing the probability of a b quark to fragment into a heavy hadron H_b. A single inclusive distribution, for instance p_{T} or rapidity, of a charmonium state C produced from b-hadron decays $d\sigma_C^{FONLL}$ can then be obtained with

$$d\sigma_C^{FONLL} = d\sigma_b^{FONLL} \otimes f(b \to H_b) \otimes D(H_b \to C), \tag{2.8}$$

where $D(H_b \to C)$ is a function describing the weak decay of the b-hadron into the charmonium state. The fragmentation functions $f(b \to H_b)$ depend on the fraction of momentum x that H_b receives from the b quark during fragmentation. One such parameterisation is the Kartvelishvili et al. function [50]

$$f(x) = x^{\alpha}(1-x), \tag{2.9}$$

which contains a free, phenomological parameter α that can be determined from experiment. The FONLL model has been shown to have good agreement with heavy hadron production measurements, highlighted in Figure 2.8 for non-prompt J/ψ and $\psi(2S)$ production at $\sqrt{s} = 7$ TeV.



(a) (b)

Fig. 2.8 p_{T} differential cross sections compared to FONLL for non-prompt (a) J/ψ and (b) $\psi(2S)$ production at $\sqrt{s} = 7$ TeV for a range of rapidity regions [10].

2.3.5 Spin Alignment and the 'Polarisation Puzzle'

While there is reasonably good agreement between the NRQCD model and the direct production of J/ψ and $\psi(2S)$ states differentially in p_{T} and rapidity (although the colour-octet and colour-singlet contributions are determined by fits to data, thus lowering predictive power), this historically has not translated into agreement between predicted spin alignment and that observed in experiment [51]. The spin alignment, or polarisation, of a particle can be defined as its preferential production with a certain subset of angular momentum eigenstates J_z along a given quantisation axis, from a superposition of various production mechanisms. For NRQCD, at higher energies the colour-octet components are expected to dominate production through gluon fragmentation processes, leading to an expected strong *transverse* polarisation $J_z = \pm 1$

for the ψ mesons in the helicity frame (i.e. the z axis aligned with their momentum direction). Meanwhile, the CSM predicts strong *longitudinal* polarization $J_z = 0$ [51].

For vector quarkonium ($J^{PC} = 1^{--}$), for instance the J/ψ and $\psi(2S)$, an individual production mechanism would produce the vector state $|V\rangle$ with angular momentum eigenstates $J_z = +1, 0, -1$ with respect to a given polarisation axis z,

$$|V\rangle = a_{+1}\,|{+1}\rangle + a_0\,|0\rangle + a_{-1}\,|{-1}\rangle\,, \tag{2.10}$$

where common reference axes include: the momentum direction of the quarkonium, named the *helicity* (HX) frame; the momentum direction of one of the colliding beams in the collision, named the *Gottfried-Jackson* (GJ) frame; and the bisector of the direction of one of the beams and the opposite of the other, named the *Collins-Soper* (CS) frame [52].

In most experiments, measuring the polarisation of vector quarkonia V is performed by analysing the angular distribution of the dilepton daughters of the quarkonium decay, which usually proceeds through the Drell-Yan process whereby the V annihilates into a virtual photon, which subsequently decays into a pair of leptons $V \to \gamma^* \to \ell^+\ell^-$. Helicity is defined as the projection of spin into momentum space $h = \vec{S} \cdot \vec{p}/|\vec{p}|$, and in the massless approximation the fermions must have opposite helicities due to the coupling to the virtual photon, leading to $J_{z'} = \pm 1$ for the lepton pair, where the z' axis is the direction of travel of the positively charged muon in the V rest frame. This means the dilepton system is a $J_{z'}$ eigenstate $|\ell^+\ell^-; J = 1; l = \pm 1\rangle$. The massless assumption is a fair one to make; the $J_z = 0$ component for J/ψ produced through the same mechanism from e^+e^- collisions is found from precise QED calculations to be of the order 10^{-4} [52].

The dilepton system eigenstates of $J_{z'}$ can then be expressed in terms of a superposition of eigenstates of J_z through a change in quantisation axis $z' \to z$ with complex rotations involving Wigner D-matrices [52]

$$|J, M'\rangle = \sum_{M=-J}^{+J} D^J_{MM'}(\alpha, \beta, \gamma)\,|J, M\rangle\,, \tag{2.11}$$

where (α, β, γ) are the Euler angles describing the rotation and $D^J_{MM'}$ is defined as $e^{-iM\alpha} d^J_{MM'}(\beta) e^{-iM'\alpha}$, where $d^J_{MM'}(\beta)$ is a reduced Wigner D rotation matrix for particular J and MM', and as an example $d^1_{0,\pm 1} = \pm \sin\theta/\sqrt{2}$, where θ is the polar

angle between the momentum direction of the positively charged lepton in the V rest frame and the arbitrary quantisation axis z, hereby referred to as θ^*. The angular distribution of the dilepton system can subsequently be described by [53]

$$\frac{dN}{d\cos\theta^* d\phi^*} = \frac{1}{(3+\lambda_\theta)}\left(1+\lambda_\theta\cos^2\theta^* + \lambda_\phi \sin^2\theta^*\cos 2\phi * + \lambda_{\theta\phi}\sin 2\theta * \cos\phi^*\right), \tag{2.12}$$

where ϕ^* is defined as the azimuthal angle between the V decay plane and its production plane, and λ_i are coefficients directly related to the helicity amplitudes, often known as the polarisation parameters. A diagram visualising the θ^* and ϕ^* variables in dilepton decays is shown in Figure 2.9.

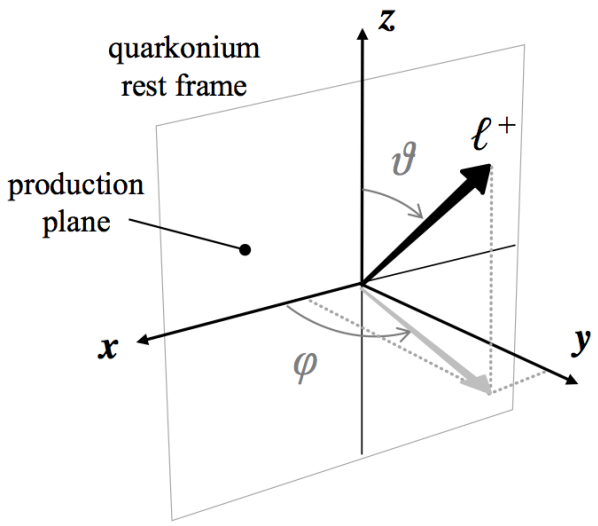


Fig. 2.9 Coordinate system for the measurement of a dilepton decay angular distribution, showing the polarisation angles θ^* and ϕ^* [52].

A study of J/ψ and $\psi(2S)$ decays to the dimuon final state by CMS [54] at $\sqrt{s} = 7$ TeV measured the polarisation parameters to be consistent with zero for both the J/ψ and $\psi(2S)$ (see Figure 2.10). Also shown is the comparison to the NLO NRQCD prediction, which predicts transverse polarisation for both cases and therefore disagrees with the data, exhibiting a stronger disagreement with the $\psi(2S)$ polarisation. The LHCb experiment has also measured the polarisation of J/ψ [55] and $\psi(2S)$ [56] decaying to the dimuon final state in the forward rapidity region

$2.5 < y < 4.0$, measuring the polarisation as close to zero for both states. The results for the polarisation parameter λ_θ are shown in Figure 2.11.

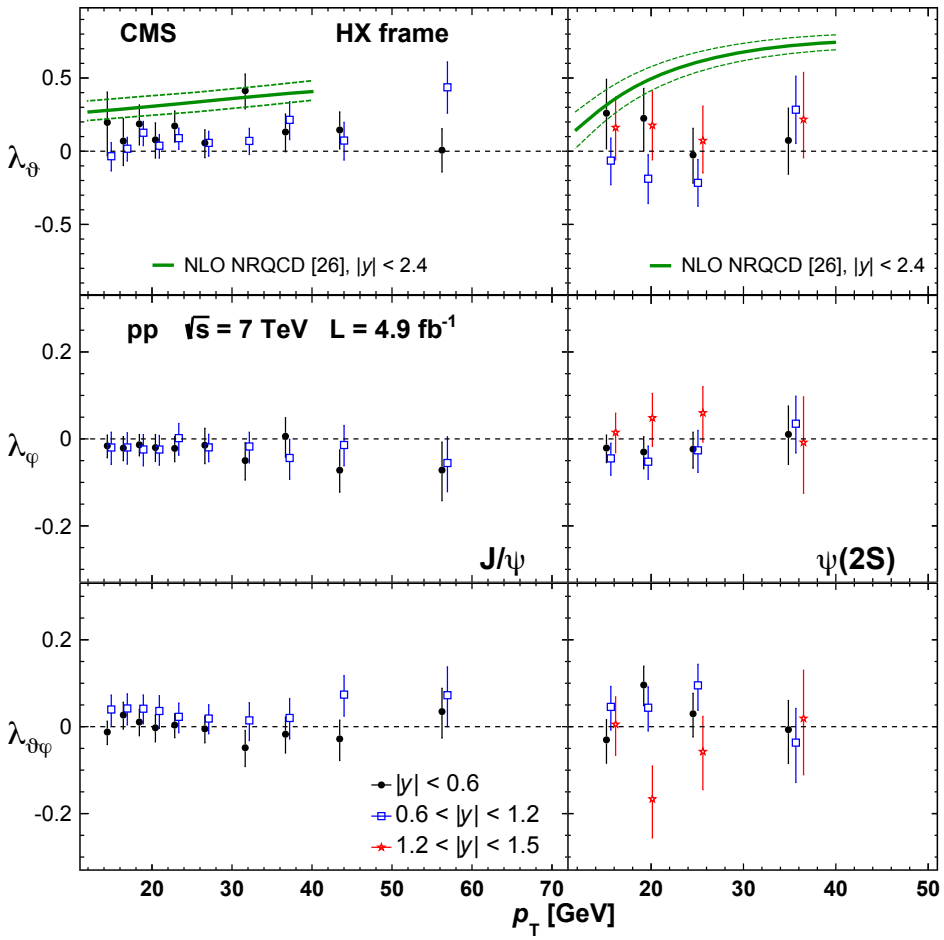
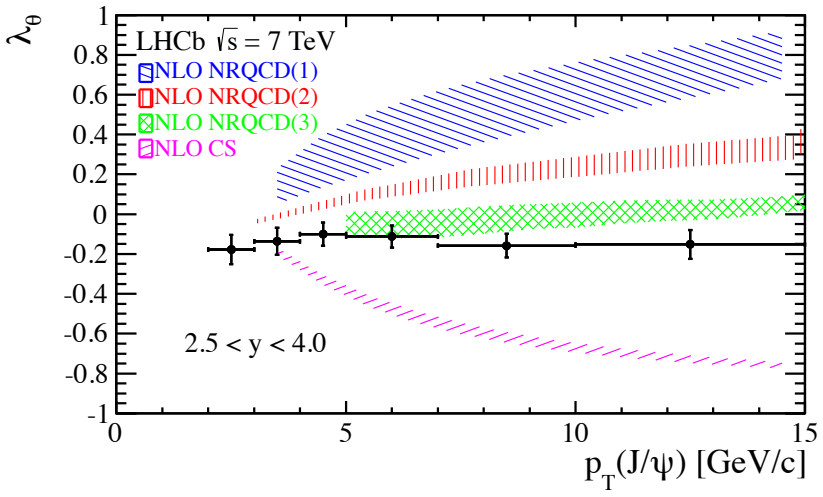
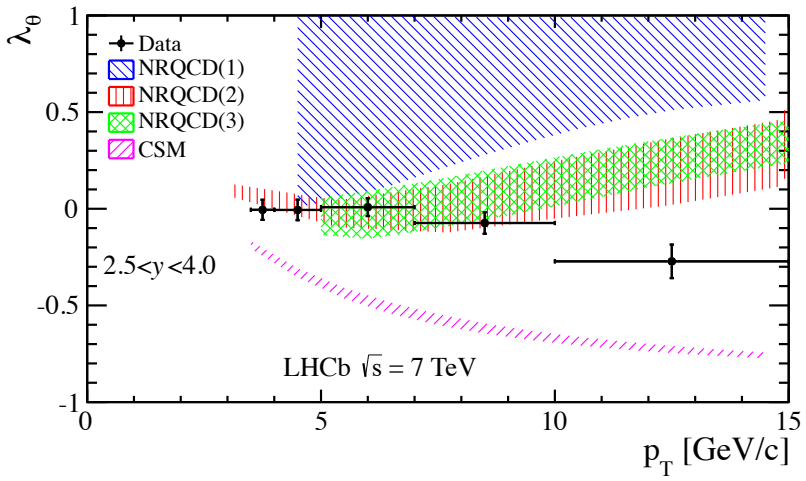


Fig. 2.10 The polarisation parameters measured by CMS for J/ψ (left) and $\psi(2S)$ (right) compared to NLO NRQCD predictions [54].



(a)



(b)

Fig. 2.11 The polarisation parameter λ_θ measured by LHCb for (a) J/ψ [55] and (b) $\psi(2S)$ [56] compared to theoretical predictions.

2.4 The $X(3872)$

2.4.1 Background and Physical Interpretation

The $X(3872)$ is an exotic resonance first discovered by the Belle collaboration in 2003 [1] through its decay to $J/\psi\pi^+\pi^-$ in the exclusive decay $B^\pm \to K^\pm J/\psi\pi^+\pi^-$. The resonance in the $J/\psi\pi^+\pi^-$ invariant mass spectrum can clearly be seen in Figure 2.12.

Its existence was subsequently confirmed by BaBar [2], and it was also measured to be produced predominantly through prompt QCD processes in $p\bar{p}$ collisions by the CDF [3] and D0 [4] experiments. It was first observed at the LHC by the LHCb experiment [7], and LHCb have subsequently confirmed the quantum numbers of $X(3872)$ to be $J^{PC} = 1^{++}$ [57]. A particularly interesting aspect of the $X(3872)$ is the closeness of its mass of 3871.69 ± 0.17 MeV to the $D^0\bar{D}^{*0}$ threshold, such that it is often speculated to be a loosely bound charm-meson molecule with a very small binding energy [58] and particle content

$$X(3872) = \frac{1}{\sqrt{2}}\left(D^{*0}\bar{D}^0 + D^0\bar{D}^{*0}\right) = \left(D^{*0}\bar{D}^0\right)_+ . \tag{2.13}$$

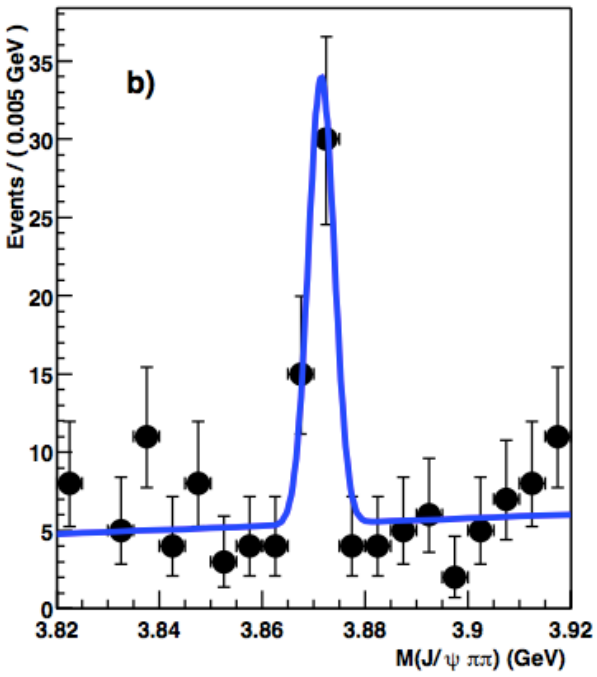


Fig. 2.12 The discovery of the exotic resonance $X(3872)$ in the $J/\psi\pi^+\pi^-$ invariant mass spectrum by Belle [1].

The typical relative momentum between the D^0 and $\bar{D}^{*0}$ in high-p_{T} pp collisions is around 1.5 GeV, compared to a relative momentum of around 28 MeV estimated for the $X(3872)$ [59]. An upper bound was therefore placed on the prompt production cross section of such a molecule [60], and it was concluded that the measured prompt

$X(3872)$ production rate from the Tevatron exceeded this upper bound by orders of magnitude.

In a more recent interpretation of $X(3872)$ production as a $D^0\bar{D}^{*0}$ molecule, it was assumed that the $X(3872)$ could also be produced through rescattering processess [61], shown in Figure 2.13. Using the NRQCD factorisation framework, with a factorisation formula of the form

$$\sigma[X(3872)] = \sum_n \hat{\sigma}[c\bar{c}_n] \left\langle \mathcal{O}_n^X \right\rangle, \tag{2.14}$$

various simplifying assumptions can be made to reduce NRQCD matrix elements $\left\langle \mathcal{O}_n^X \right\rangle$ to a smaller set. The assumptions considered were:

- S-wave dominance: The $X(3872)$ has an equal probability to be formed from any $c\bar{c}$ pair that is created in a S-wave state with small relative momentum, regardless of the colour and spin of the $c\bar{c}$. This assumption uses the same pattern of matrix elements as the Colour Evaporation Model [61].
- Colour-octet $^3S_1$ dominance: The $X(3872)$ can only be formed from a $c\bar{c}$ pair produced in the $^3S_1$ colour-octet state with small relative momentum [61].

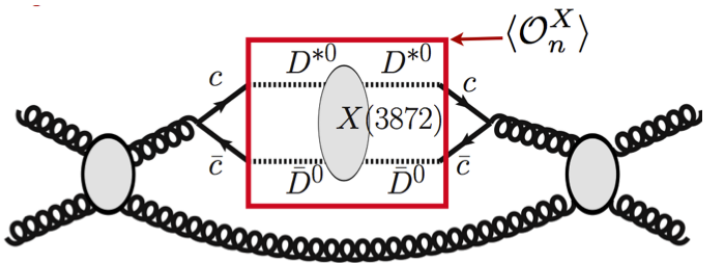
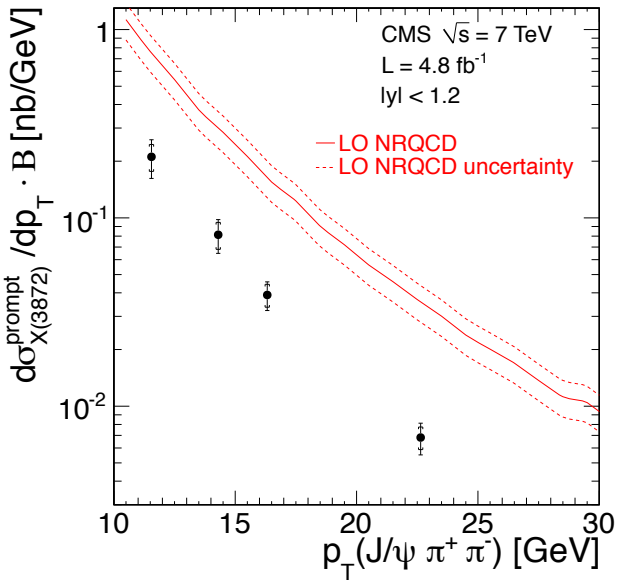


Fig. 2.13 Diagram showing the production of the $X(3872)$ as a charm-meson molecule through rescattering process [59].

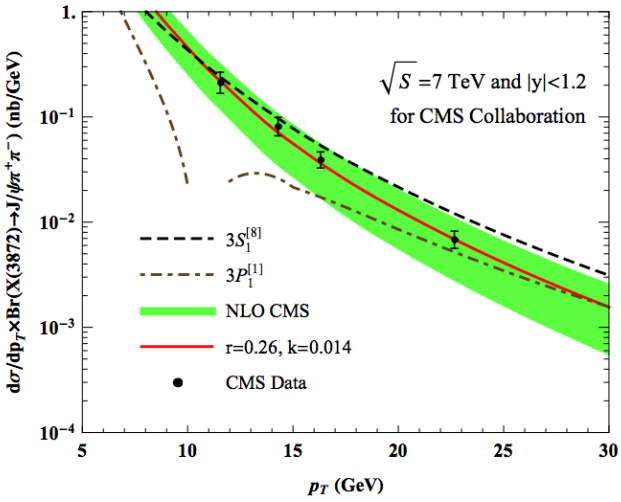
For $p_{\mathrm{T}} > 4$ GeV, it was found that each of these simplifying assumptions give the same differential cross section result. The production rate was found to increase by several orders of magnitude relative to the previously determined upper bound, providing a possible explanation of the discrepency between theory and experiment.

However, a production cross section measurement of promptly produced $X(3872)$ was performed by CMS [8] as a function of p_T (see Figure 2.14(a)), and shows the NRQCD prediction for the prompt $X(3872)$ production rate (where $X(3872)$ is assumed

to be a loosely bound $D^0\bar{D}^{*0}$ molecule) with rescattering to be significantly too large, although it describes the p_T dependence reasonably well.



(a)



(b)

Fig. 2.14 (a) Measurement of the prompt $X(3872)$ production cross section differentially in p_{T} by CMS [8]. The NRQCD factorisation framework with the $X(3872)$ modelled as a loosely bound $D^0\bar{D}^{*0}$ molecule with rescattering shows an overprediction. (b) The same measurement fitted to an NLO NRQCD factorisation theory of $X(3872)$ modelled as mixed $\chi_{c1}(2P) - D^0\bar{D}^{*0}$ state, produced through the $\chi_{c1}(2P)$ component [62].

A later interpretation of $X(3872)$ as a mixed $\chi_{c1}(2P) - D^0\bar{D}^{*0}$ state, where the $X(3872)$ is produced predominantly through its $\chi_{c1}(2P)$ component has also been proposed [62]. As previously discussed, the very small binding energy of a hypothesised $D^0\bar{D}^{*0}$ molecule makes it unlikely that this alone could have such a high rate of production through prompt QCD processes found at the Tevatron, and including rescattering processes overpredicts the data. Additionally, properties of the $X(3872)$ such as the p_{T} distribution also have similarities to the $\psi(2S)$. This interpretation might also explain the observed large isospin violation in $X(3872) \to J/\psi\rho$ decays through rescattering of the intermediate $D^0\bar{D}^{*0}$ component [63]. The mixed $\chi_{c1}(2P) - D^0\bar{D}^{*0}$ hypothesis is used with the NLO NRQCD model and fitted to CMS data (see Figure 2.14(b)), showing good agreement [62]. An additional motivation for this interpretation is that one of the free parameters of the fit, the colour-octet matrix element ratio, is almost the same as for $\chi_{c1}(1P)$ production.

2.4.2 Other Theoretical Interpretations

Charmonium State

The closest charmonium state in mass to the $X(3872)$ which also satisfies quantum number requirements is the 2^3P_1 state, known as the $\chi_{c1}(2P)$. This could explain observed decays to $J/\psi\omega$, however the isospin violating decay to $J/\psi\rho$ at around the same rate as $J/\psi\omega$ is difficult to explain with this pure charmonium state, although the large isospin violation could be explained through a virtual coupling to the $D^0\bar{D}^{*0}$, which could also have the effect of lowering the bare mass of the $\chi_{c1}(2P)$ down to the $X(3872)$ mass.

Tetraquark

A tetraquark is the binding of a coloured diquark and anti-diquark pair to form a colourless state. They are similar to $q\bar{q}$ mesons, but instead there are diquark pairs at the end of the colour string instead of quarks. Tetraquarks are therefore often expected to decay into pairs of baryons, for example the tetraquark candidate $Y(4660)$ is found to decay into a $\Lambda_c^+\Lambda_c^-$ final state [64]. However, due to limited phase space from its relatively low mass, the $X(3872)$ is forced to decay through $J/\psi\rho$ and $J/\psi\omega$.

In the interpretation of $X(3872)$ as a tetraquark, the $X(3872)$ can be considered as a mixture of two neutral mass eigenstates:

$$X_u = [cu][\bar{c}\bar{u}], \qquad X_d = [cd][\bar{c}\bar{d}], \tag{2.15}$$

where the mass difference is on the order of $m_d - m_u$. Mixing occurs from annihilation-creation of the light quark pairs, leading to pure isospin mixtures of the form $(X_u \pm X_d)/\sqrt{2}$ [65]. Isospin symmetry is predicted to be close to maximally violated, and correspondingly so are the subsequent strong decays. This may explain the observed isospin-violation from $J/\psi\omega$ and $J/\psi\rho$ final state measurements of the $X(3872)$. However, a study by CDF found no evidence for the existence of two nearby mass states, limited by their experimental resolution, and set an upper limit on the mass difference of 3.6 MeV at a 95% confidence level, assuming an equal contribution of the two states [66].

Additional predictions of the tetraquark model are the existence of charged states. The charged tetraquark candidate $Z_c(3900)^{\pm}$ has been observed in decays to $J/\psi\pi^{\pm}$ [67], and its neutral partner $Z_c(3900)^0$ was later observed in decays to $J/\psi\pi^0\pi^0$ [11]. However, searches for the charged partner of the $X(3872)$ have been performed, for instance in the $J/\psi\rho^+$ final state [68], and no evidence of its existence has been found.

Chapter 3

The ATLAS Experiment

3.1 The LHC

The Large Hadron Collider (LHC) is the world's highest energy particle accelerator, with beams of protons colliding at a maximum design centre-of-mass energy of $\sqrt{s} = 14$ TeV, and an instantaneous luminosity of $L = 10^{34}$ cm^{2}s^{-1}. The LHC was constructed at the CERN accelerator complex, close to the French-Swiss border, inside the tunnel previously used for the LEP experiments (which used an electron-positron collider). This circular tunnel is 26.7 km in circumference and approximately 100 m underground to shield from cosmic radiation and to provide protection from the high energy radiation produced in the beam collisions.

A diagram of the CERN accelerator complex can be seen in Figure 3.1. The protons are initially accelerated by the LINACS, a linear accelerator which boosts their energy to 50 MeV. They then travel through to the BOOSTER, a synchrotron accelerator which increases their energy to 1.4 GeV. Next, they are accelerated through the Proton Synchrotron (PS) and Super Proton Synchrotron (SPS) to energies of 25 GeV and 450 GeV, respectively. After they reach an energy of 450 GeV the protons can be injected into the main ring, where they are further accelerated to maximum energy for the beam collisions.

There are four main experiments situated along the main ring, which are: CMS [69], ATLAS [70], LHCb [71] and ALICE [72]. The ATLAS and CMS general purpose experiments are designed to study a wide range of physics. These include: Higgs boson searches and measurements, supersymmetry searches, Standard Model measurements

and B-physics studies. The LHCb and ALICE experiments are more specialised to study B-physics and heavy ion collisions, respectively. There are also several smaller experiments, for instance LHCf [73].

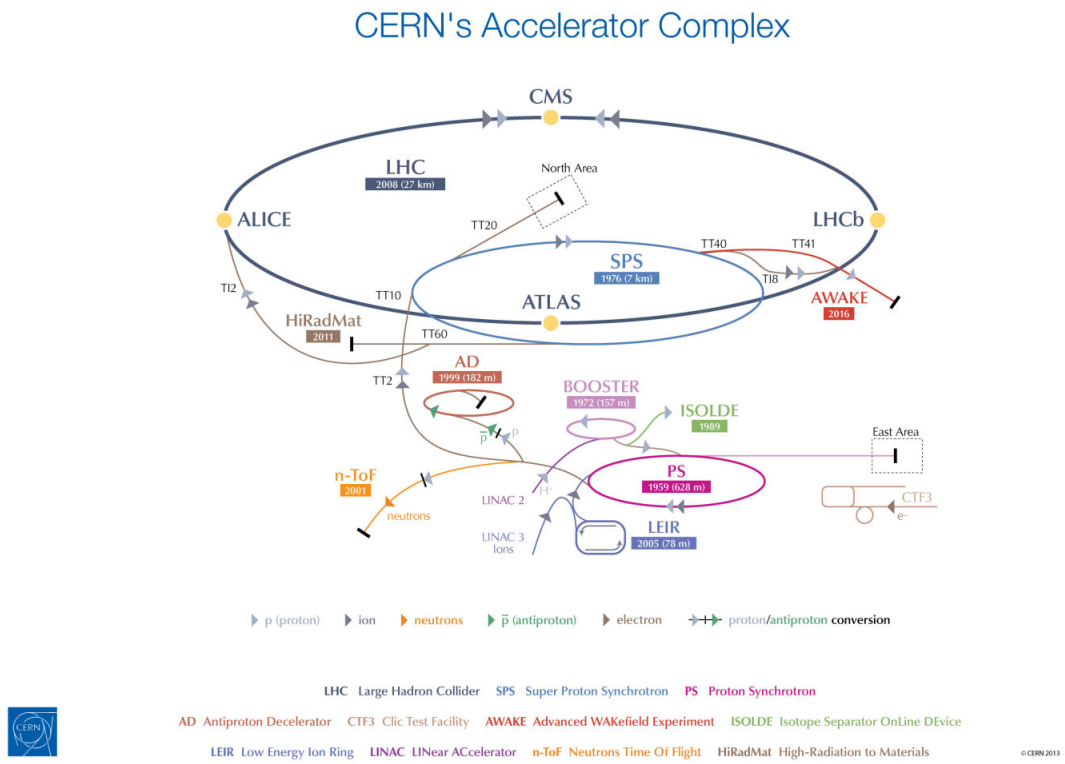


Fig. 3.1 Diagram showing the group of accelerators and experiments which form the CERN Accelerator complex [74].

The LHC started running in 2010, with each proton beam accelerated to an energy of 3.5 TeV, and a centre-of-mass energy of $\sqrt{s} = 7$ TeV. By the end of 2011 the LHC had delivered a total integrated luminosity of over 5 fb^{-1} to each of the ATLAS and CMS detectors, and by mid-June in 2012 had produced enough collisions for both the ATLAS and CMS collaborations to jointly make the historic discovery of a boson consistent with the Standard Model Higgs at a mass of around 125 GeV [16, 17].

After the first long shutdown (LS1) from spring 2013 to the beginning of 2015 for upgrades, the LHC has resumed data taking with beam collisions at a record-breaking $\sqrt{s} = 13$ TeV (6.5 TeV per beam).

3.2 Overview of the ATLAS Detector

The ATLAS detector [75] is a cylindrical, forward-backward symmetric, general purpose particle detector designed to study physics at the TeV energy scale. It is located at Point 1 on the LHC ring, approximately 93 m underground to shield the outside from the high energy radiation produced in the collisions, and to prevent cosmic rays reaching the detector. The detector consists of numerous sub-detector components. A schematic diagram of the detector showing the various sub-detector components in a cutaway view is shown in Figure 3.2.

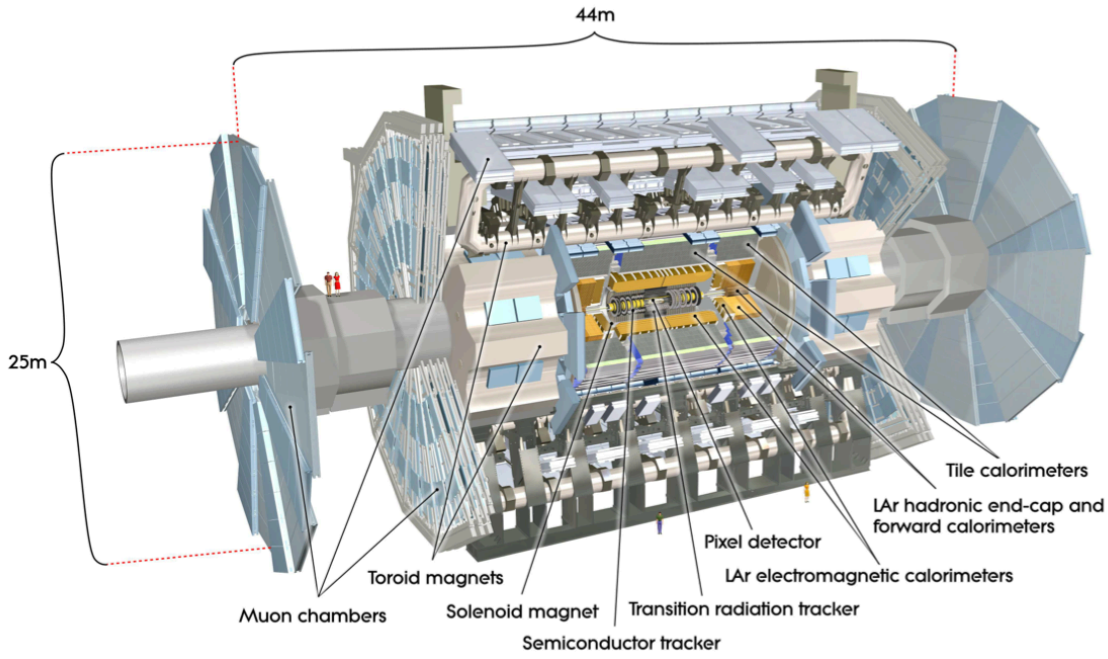
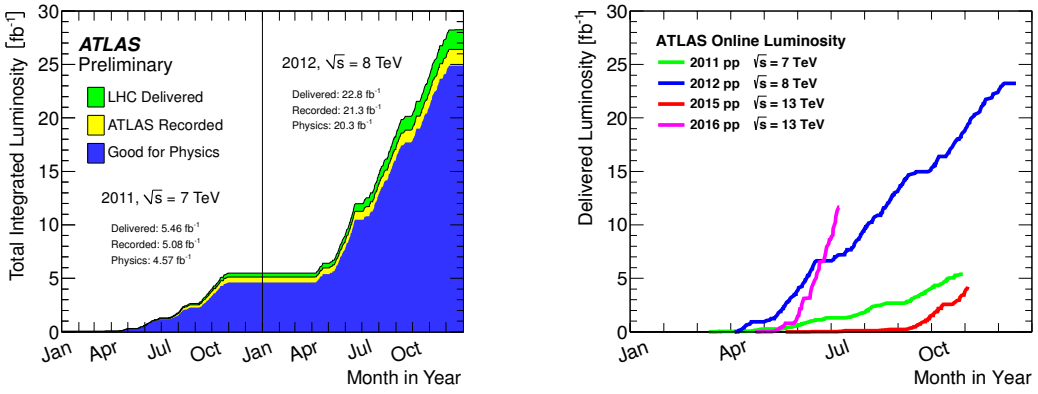


Fig. 3.2 Schematic diagram showing a cutaway view of the ATLAS detector and its sub-detector components [75].

The innermost part of the inner detector (ID) is comprised of pixel and silicon microstrip (SCT) tracking technology for high-precision measurements, and further outwards also the transition radiation tracker (TRT). The inner detector is immersed in a 2 T solenoidal magnetic field. Enclosing the ID and solenoidal magnet are the electromagnetic (ECAL) and hadronic (HCAL) sampling calorimeters, which provide good containment of the electromagnetic and hadronic showers in order to limit punch-through into the muon spectrometer (MS). The ECAL uses liquid argon, while the

HCAL is an arrangement of steel and scintillator tiles. Surrounding the calorimeters, the MS consists of three air-core toroidal magnets, generating a magnetic field capable of 1.5 - 7.5 Tm of bending power, dependent on rapidity. The MS consists of both precision measurement detectors (monitored drift tubes and cathode strip chambers) and fast-trigger detectors (thin-gap chambers and resistive plate chambers). The ATLAS detector uses a three-level trigger system in order to reduce the frequency of events written to disk to around 200 Hz. Figure 3.3 shows the integrated luminosity of *pp* collisions delivered to and collected by the ATLAS detector from 2011 onwards.



(a) (b)

Fig. 3.3 (a) Total integrated luminosity of *pp* collisions delivered, recorded and good-for-physics from the ATLAS detector for the 2011 and 2012 run periods [76]. (b) Delivered luminosity to the ATLAS detector for run periods from 2011 - 2016 [77].

This analysis uses a dimuon trigger. The hardware-based Level 1 (L1) muon trigger finds regions-of-interest (RoIs) by searching for hit coincidences in layers of the muon trigger detectors inside pre-defined geometrical windows, and uses one of a range of fixed p_T thresholds to bind the path of the muon to give a rough position ($\Delta\eta \times \Delta\phi \approx 0.1 \times 0.1$). Overlap between two or more closely separated muons whereby the trigger only finds a single RoI at L1 is possible, and is accounted for in the data-driven method of measuring the dimuon trigger efficiency. The software-based two-stage high level trigger (HLT) is seeded by the L1 RoI, and can then use more precise MS and ID information in its algorithms at each stage in order to reconstruct the final muon trigger object with a resolution comparable to the full offline reconstructed muons.

The coordinate system widely used to describe the ATLAS detector is as follows. The origin of the coordinate system is defined by the nominal interaction point. A right-handed coordinate system is used, with the z-axis defined by the beam direction. Side-A of the detector represents the positive z direction. The $x-y$ plane is orthogonal to the beam direction, where the positive x-axis direction points from the interaction point to the centre of the LHC ring, and the positive y-axis direction points upwards along the zenith. The azimuthal angle ϕ is measured around the beam axis, and the polar angle θ is measured from the z axis. The pseudorapidity η is defined as

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

which is equal to rapidity when a physics object is approximately massless. For the case where there is an appreciable mass, for instance jets, rapidity y is more appropriate,

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

where E is the energy and p_z is the momentum projection in the z-direction of the physics object. The distance ΔR in pseudorapidity-azimuthal space is defined as

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}, \tag{3.3}$$

where $\Delta\phi$ and $\Delta\eta$ are defined as the separation in azimuthal and pseudorapidity space, respectively.

3.3 The Inner Detector

The inner detector (ID) is located closest to the beam pipe and the interaction point (IP), providing high-precision momentum measurements of charged particles and also vertex identification, with a fine detector granularity needed for the large track density. The ID consists of the pixel and silicon microstrip (SCT) trackers, which are surrounded by the straw tubes of the transition radiation tracker (TRT). The layout of the various sub-components during Run 1 are shown in Figure 3.4. During Long Shut Down 1 (LS1), the ID was upgraded to include an additional pixel layer closest to the beam pipe, known as the Insertable B-layer (IBL) [78], compensating for the eventual deterioration

of the current pixel layers from radiation effects. This also provides the benefit of improved vertex resolution, particularly of secondary vertices necessary for B-physics studies. The ID is immersed in a 2 T magnetic field produced by the central solenoid magnet in order to measure the transvere momentum of the charged particles. These precision tracking components cover the region $|\eta| < 2.5$.

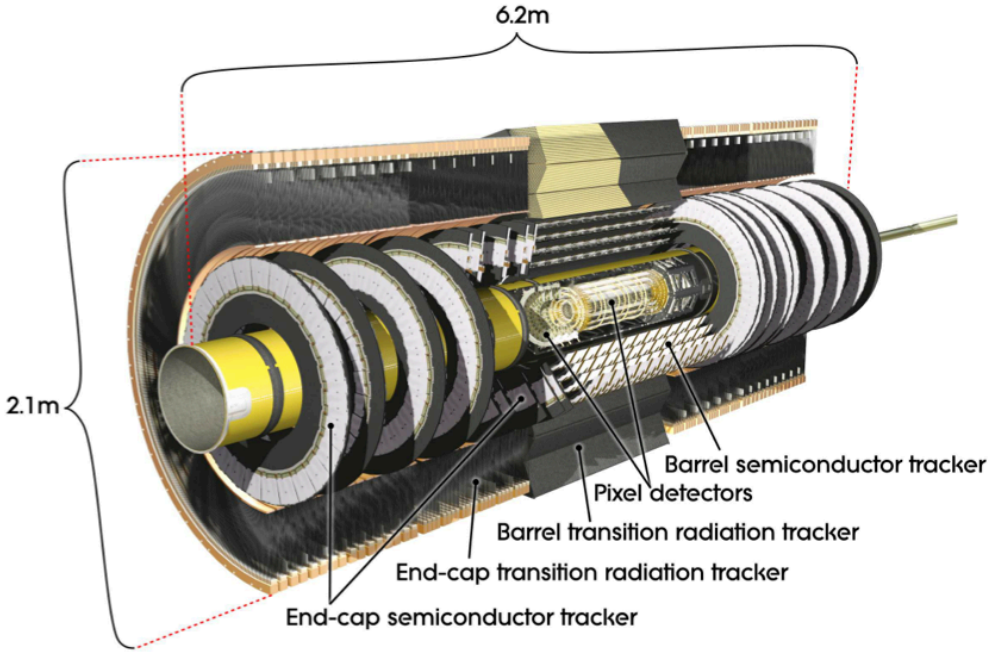


Fig. 3.4 Schematic diagram showing a cutaway view of the ATLAS inner detector (ID) and its sub-components during Run 1 [75]. An additional pixel layer known as the IBL was inserted for Run 2.

3.3.1 Pixel Detector and Semiconductor Tracker

During Run 1

The silicon pixel detector offers the highest granularity, and is located closest to the beam pipe. It provides the first spatial measurement of charged particles produced from the IP. It is arranged as three concentric cylindrical layers around the beam pipe in the barrel region $|\eta| < 1$, and as disks of three layers perpendicular to the beam axis in the endcap regions. The pixel layer closest to the beam pipe is critical for B-physics studies, in order to best resolve the secondary vertices from long-lived b-hadrons, and is known as the B-layer. The pixel detector contains around 80 million

read-out channels: around 67 million in the cylindrical barrel layers and 13 million in the endcap layers, with each pixel sensor having size $R-\phi \times z = 50 \times 400$ μm^2 with resolution $R-\phi \times z = 10 \times 115$ μm^2 [75]. The pixel detector covers the radial region $50.5 < r < 150.0$ mm.

The SCT consists of 4088 silicon strip modules. These are arranged as four concentric cylindrical layers around the beam pipe in the barrel region. Each of the two endcaps contain nine disk layers perpendicular to the beam direction, which can be seen in Figure 3.4. The SCT provides four spatial measurements for traversing charged particles, from eight strip measurements. In the barrel region, each module consists of two 6.4 cm silicon strip sensors daisy-chained together with a pitch of 80 μm, on each side of the module [75]. The strips on each side of the module have a small 40 mrad angle between them, to perform a stereo angle measurement. The typical resolution of each module in the SCT is $(R-\phi\,,\ z) = (17\,,\ 580)$ μm in the barrel, and $(R-\phi\,,\ R) = (17\,,\ 580)$ μm in the endcaps, covering the radial region $299 < r < 514$ mm.

Run 2 Upgrade

The high-luminosity, high-radiation environment close to the IP has the effect of degrading the pixel detector, particularly the B-layer. The lifetime in integrated luminosity of the present B-layer due to these radiation effects is estimated to be around 300 fb^{-1} [78], with the onset of tracking efficiency degradation at an even lower integrated luminosity. For Run 2, an additional pixel layer was inserted closest to the beam pipe, known as the Insertable B-Layer (IBL). A comprehensive technical report of the IBL can be found in [78]. The IBL is expected to maintain robust tracking until the replacement of the inner detector in 2025. The IBL also provides improved precision for vertexing and b-tagging.

3.3.2 Transition Radiation Tracker

The TRT is the outermost component of the ID, and consists of roughly $298,000$ proportional drift tubes. The tubes have a 4 mm diameter, with a wall made from polyimide, offering good electrical and mechnical properties with minimal thickness at around 35 μm. The walls are kept at a voltage of -1.5 kV [75]. The tubes are filled with a gas mixture consisting of 70% Xe, 27% CO_2 and 3% O_2. An anode sense wire,

kept at ground potential, made from gold-plated tungsten, runs through the centre of each straw tube. The tubes are around 1.5 m in length in the barrel, arranged along the beam axis. They are around 0.4 m in length in the endcaps, and are arranged perpendicular to the beam axis [79]. The spaces between the straws are filled with polymer fibres, which create the transition radiation as the highly relativistic charged particles traverse the material boundary. TRT measurements are only performed in the $R-\phi$ plane, and have a precision of 130 μm for each tube, across the radial region $563 < r < 1066$ mm.

3.4 The Calorimeters

The sampling calorimeters cover the range $|\eta| < 4.9$, and are designed to measure the energy (and position/direction) of both charged and neutral particles and jets. The calorimeter system must also contain electromagnetic and hadronic showers to prevent punch-through into the muon system (MS) [75]. A cutaway schematic diagram of the calorimeter system is shown in Figure 3.5.

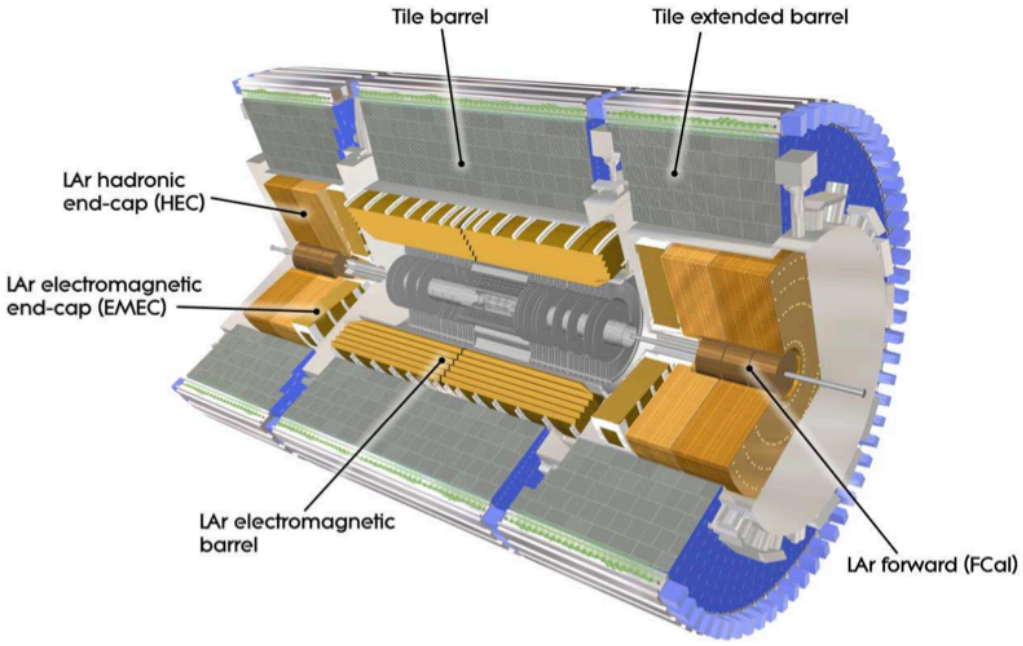


Fig. 3.5 Schematic diagram showing a cutaway view of the ATLAS calorimeter system [75].

3.4.1 Liquid Argon Calorimeter

The Liquid Argon (LAr) calorimeter system is used for both electromagnetic (EM) and hadronic calorimetry, and shares a vacuum vessel with the solenoidal magnet to reduce dead material. The LAr EM calorimeter in the barrel region ($|\eta| < 1.475$) provides fine granularity ($\Delta\eta \times \Delta\phi = 0.0031 \times 0.0245$ in the first layer, known as the strip layer, becoming coarser $\Delta\eta \times \Delta\phi = 0.05 \times 0.0245$ in the back layer) for precision measurements of photons and electrons [75], with a thickness between 24 X_0 and 33 X_0 dependent upon $|\eta|$. The EM calorimeter in the endcaps (EMEC) is divided into two coaxial wheels, with the outer wheel covering $1.375 < |\eta| < 2.5$ and the inner wheel covering $2.5 < |\eta| < 3.2$. The EM calorimeter uses lead absorber plates and an accordion geometry to provide full ϕ symmetry without any azimuthal cracks. The electrodes used to read out the signal are located in between the gaps of the absorber plates, and consist of three copper layers separated by polyimide insulating sheets [75]. The two outer layers are placed under a high-voltage, leaving the inner layer as the readout of the signal by capacitive coupling from the induced current in the LAr gap from drifting charges due to particle interactions [80].

The LAr calorimeter system is also used for the Hadronic End-cap Calorimeters (HEC). The HEC is a copper-LAr sampling calorimeter which uses a flat-plate design [75], covering the range $1.5 < |\eta| < 3.2$. The HEC consists of a front (HEC1) and back (HEC2) wheel in each of the end-cap cryostats, which they share with the EMEC. The gaps between the copper plates are also occupied by electrodes, with the central electrode being the readout. The size of the readout cell is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for $|\eta| < 2.3$ and 0.2×0.2 for $|\eta| > 2.3$ [75].

3.4.2 Tile Calorimeter

Surrounding the EM calorimeter in the barrel region is the tile calorimeter, covering the region $|\eta| < 2.7$. The tile calorimeter is sub-divided into a central barrel and two forward barrels. Each barrel contains 64 steel-scintillator modules, with the modules divided longitudinally into three layers. The steel is used as the absorber material, while the scintillator acts as the active medium. The signal is then read out through fibre optic cables into photomultiplier tubes. The inner two layers have the resolution $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$, while the outer layer has a slightly poorer granularity in pseudorapidity $\Delta\eta \times \Delta\phi = 0.2 \times 0.1$.

3.5 The Muon Spectrometer

The Muon Spectrometer (MS), the outermost part of ATLAS, detects charged particles which have travelled through the calorimeters. The MS uses air-core toroid magnets to measure the momentum of charged particles in the range $|\eta| < 2.7$ by measuring track trajectories of charged particles bent by the toroidal magnetic field, and triggers on particles in the range $|\eta| < 2.4$. The MS contains a variety of sub-detector components, providing a mixture of high precision tracking measurements and fast read-outs for triggering. The sub-detector components comprising the MS are shown in Figure 3.6.

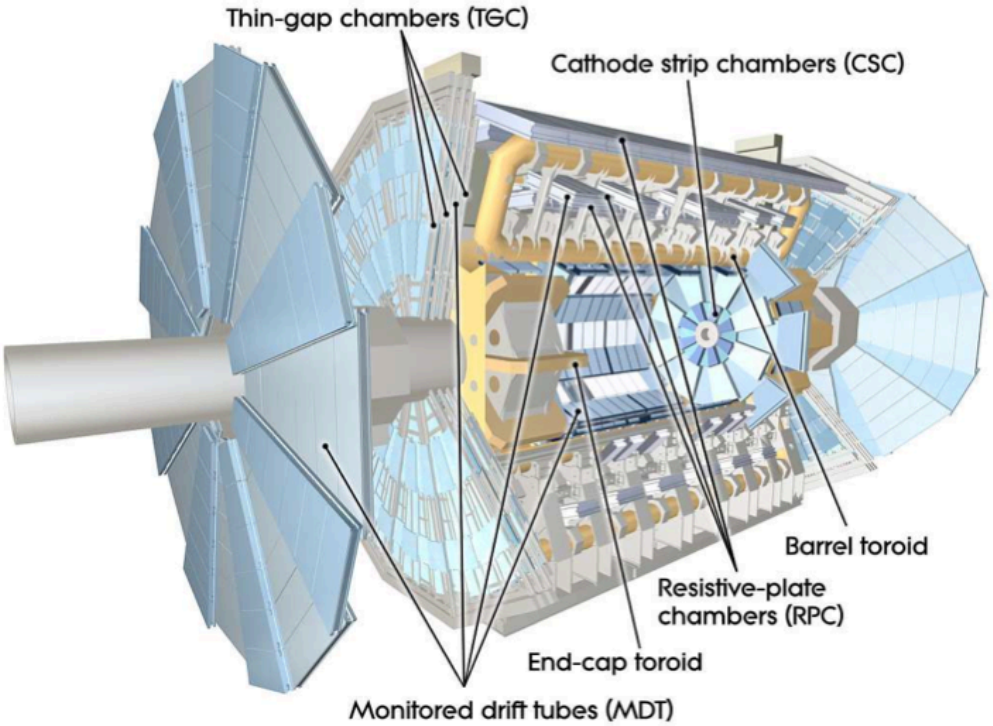


Fig. 3.6 Schematic diagram showing a cutaway view of the ATLAS muon spectrometer [75].

In the barrel region, the muon chambers are arranged in three concentric cylinders around the beam axis. In the end-caps, the chambers are placed in planes perpendicular to the beam axis, again as three layers. The sub-detector components are:

- **Monitored Drift Tubes (MDT)** provide precision measurements. They are pressurised drift tubes consisting of an Ar/CO_2 gas mixture at 3 bar. The free

electrons resulting from ionisation are collected at a central tungsten-rhenium wire at a potential of around 3 kV. The maximum drift time is around 700 ms, with an average resolution per tube of around 80 μm.

- **Cathode-Strip Chambers (CSC)** provide precision measurements. The safe-operation rate limit for the MDTs is 150 Hz/cm^2, which is exceeded in the region $|\eta| > 2$ in the first end-cap layer, so here they are replaced by the CSCs. The CSCs are multiwire proportional chambers with wires oriented in the radial direction [75], with a resolution of 60 μm.

- **Resistive Plate Chambers (RPC)** are part of the trigger system in the barrel region. They consist of three cylindrical layers around the beam axis, offering up to six measurements in η and ϕ. The RPC is a gaseous parallel electrode plate detector, where to resistive plates are sepearated by a distance of 2 mm by insulating gas mainly consisting of $C_2H_2F_4$. An electric field across the plates of 4.9 kV/mm causes avalanches to form along the traversing, ionising track, which drift towards the anode. The signal is read out by capacitive coupling. [75].

- **Thin Gap Chambers (TGC)** are part of the trigger system in the end-cap region. They are multiwire proportional chambers, similar to the CSC, to cope with the high interaction rate in the $|\eta| > 2$ region.

3.5.1 Muon Reconstruction

Muons with a wide range of transverse momenta from 3 GeV to around 1 TeV can be reconstructed with the ATLAS detector, with high efficiency within the range $|\eta| < 2.3$. Momentum resolution is typically between 4 % - 10 %, and is largely limited by energy loss as particles traverse the detector before reaching the MS, and is also limited at very high-p_{T} (> 300 GeV) as particles traverse the detector in much straighter lines. High precision measurements in the ID can also help with muon identification, particularly for $p_{\mathrm{T}} < 20$ GeV. There are three main types of muon identification, which each use two main types of reconstruction algorithm, named 'Staco' and 'Muid':

- **Standalone muons:** the muons are identified from hits solely in the MS by linking track segments in each of the three layers, with no information used from the ID. The muon trajectory is subsequently extrapolated back to the IP, taking

into account multiple scattering and energy loss in the calorimeters. The Staco algorithm assigns energy loss based on the material traversed by the muon, while Muid can also use calorimeter energy measurements.

- **Combined muons:** MS tracks are paired with ID tracks, using a χ^2 match from the covariance matrices of the inner and outer tracks, to identify combined muons. The Staco and Muid algorithms perform a recombination of the inner and outer tracks to obtain the combined track vector, both taking into account energy losses from traversing the detector material [81]. Staco performs a statistical recombination of the inner and outer track vectors, while Muid does a partial refit by starting from the inner track vector and covariance matrix and then adding measurements from the outer track [81].

- **Tagged muons:** MS algorithms propagate all ID tracks with enough momentum into the first layer of the MS, where nearby track segments are searched for. If a segment is found to be close enough to the predicted track, then that ID track is tagged as a muon. However, the ID track and MS segment tracks are not combined to form a new track [81].

3.6 Data Aquisition and Trigger System

The nominal pp bunch crossing rate delivered by the LHC of 40 MHz would lead to far too many events for the ATLAS detector to read out with current technology. A three-level trigger system is therefore used to filter events, named Level 1 (L1), Level 2 (L2) and Event Filter (EF). At L1, the trigger system uses reduced granularity information from the calorimeters and muon trigger chambers described previously, which deliver fast read-outs. The two higher trigger levels (HLT) subsequently refine the decisions of the previous level using full granularity and precision detector measurements [75] at the software level. The L1 trigger reduces the event rate from the bunch crossing rate to around 75 kHz, which is further reduced to around 200 Hz after the L2 and EF trigger selections.

The data aquisition system (DAQ) receives event information from the read-out electronics at rates up to the L1 trigger rate. The L1 trigger searches for specific signatures, for example high-p_{T} muons, electrons and jets. From these it defines one or more Regions of Interest (RoIs), which are regions in (η, ϕ) space where the system has

identified interesting physics processes. These are determined by the central trigger processor, which implements the menu of L1 trigger selections. The L1 trigger decision on whether or not to keep the event must be made within 2.5 μs. The L2 trigger is then seeded by the L1 RoI data, which uses full granularity and precision detector information within the RoI, which is approximately 2 % of the total event data. The L2 trigger selection reduces the event rate to below 3.5 kHz, with a processing time of around 40 ms [75]. Finally, the EF trigger selection uses the full event information to perform the final trigger decision, reducing the event rate to around 200 Hz with a processing time of the order 4 seconds. More detail on the algorithms used for B-physics triggers can be found in Chapter 4.

Events selected by the EF are then sent to the output nodes (called SFOs) of the DAQ/HLT system [75], and the event files can subsequently be transferred to the CERN central data-recording facility for processing and storage. The SFOs can transfer event files at rates of up to 400 Hz, and have a local storage capacity buffer of 24 hours before data transfer. At the transfer stage, the sets of files map to ATLAS-defined data streams, for instance B-physics and τ-leptons, depending on which triggers fired in the event.

The amount of data collected by the ATLAS experiment, along with the other LHC experiments, in addition to simulated data, is too large to be managed at a single site. Data processing is therefore shared using the Worldwide LHC Computer Grid (WLCG), which is usually referred to as just the 'Grid'. The Grid sites are split into three different 'tiers'. The CERN data centre represents the first tier, known as 'Tier 0', which processes and stores the raw data files. This data can be distributed to one of roughly thirteen 'Tier 1' sites, which are usually located at national laboratories of participating countries and provide storage of raw and reconstructed data, as well as large-scale reprocessing. The 'Tier 2' sites are usually located in participating research institutions, and can be used to perform physics analyses on reconstructed data and provide storage of the processed data.

Chapter 4

B-Physics Triggers

In this chapter, a brief overview of the B-Physics triggers is first provided in Section 4.1. The B-Physics trigger algorithms are explained, and the various B-Physics triggers used in ATLAS analyses are described. In Section 4.2, we describe an efficiency measurement of a dimuon B-Physics trigger performed using 2011 data, which formed part of the author's service task for ATLAS. The method used to extract this trigger efficiency is similar to the method used for 2012 data.

4.1 B-Physics Triggers Overview

Most B-physics analyses from the ATLAS experiment use decays containing at least one dimuon pair in the final state. Therefore, dimuon triggers are used around the invariant mass range of the ψ, Υ and B mesons, and also across the full mass range, spanning from below the J/ψ mass to just above the Υ mass (called DiMu triggers). These are known as the B-physics triggers. The B-physics triggers can be separated into two categories according to the algorithms used to select the dimuon pair [82]. These are highlighted in Figure 4.1.

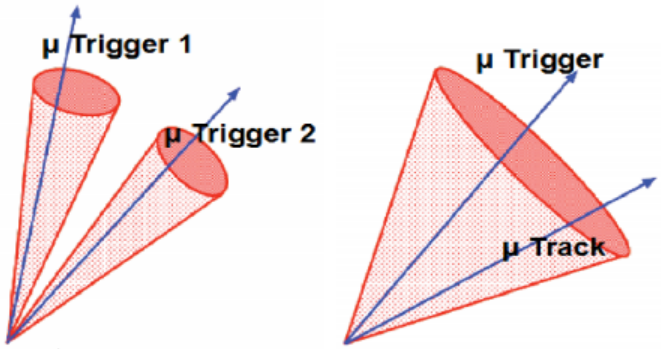


Fig. 4.1 Diagrams showing the two different B-physics trigger algorithms. The left figure shows the topological trigger, seeded by two Level 1 muon RoIs. On the right shows the TrigDiMuon trigger, seeded by a single L1 muon RoI [82].

Either the trigger is seeded at Level 1 by a single muon RoI, and at Level 2 the HLT algorithm searches for a second track within a wide (η, ϕ) region around the triggered muon candidate; or the trigger is seeded at L1 by two muon RoIs and the muons are confirmed at the HLT, known as topological triggers. The muon trigger objects are typically required to have opposite sign charge and a loose vertex χ^2 match, assuring the muons originate from a common vertex, to reduce the trigger rate while maintaining high signal yields. The invariant mass ranges for the various triggers are:

- ψ mass range: (2.5 - 4.3) GeV, trigger names ending with Jpsimumu.
- B mass range: (4.0 - 8.5) GeV, trigger names ending with Bmumu.
- Υ mass range: (8 - 12) GeV, trigger names ending with Upsimumu.
- DiMu mass range: (1.5 - 14.0) GeV, trigger names ending with DiMu.

The beginning of the trigger names denote the trigger level: either Level 1 (L1), Level 2 (L2), or 'Event Filter' (EF). The trigger name also contains information on the minimum p_{T} of the pairs of muons. For example, an Event Filter level trigger firing on two muons each with minimum $p_{\mathrm{T}} = 4$ GeV in the J/ψ mass range is named `EF_2mu4T_Jpsimumu`, where the additional 'T' denotes a tighter selection requirement on the muons. As an illustrative example, the spectrum of dimuon candidates collected during the first half of the 2011 run period is shown in Figure 4.2. Also shown are

the invariant mass ranges of some of the B-physics triggers. There is a large increase in signal yield for the ψ and Υ mesons using the B-physics triggers (red, green and blue regions), rather than a high-p_{T} single muon trigger, which is shown in grey. The B-physics triggers covering the narrower ψ and Υ mass ranges also produce slightly higher signal yields than the DiMu trigger, which covers a wider mass range and can be seen as the black line.

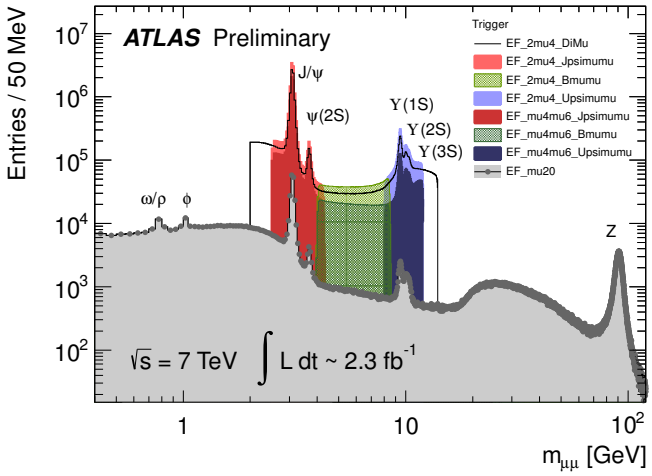


Fig. 4.2 Invariant mass spectrum of dimuon candidates from the first half of 2011 data taking. The mass range and measured yields are shown for various triggers [83].

4.2 Dimuon Trigger Efficiency Measurement

The `EF_2mu4T_Jpsimumu` dimuon trigger efficiency is extracted for data collected from the ATLAS detector during the latter half of 2011, using offline reconstructed dimuon events from $\mathrm{J}/\psi \to \mu^+\mu^-$ and $\Upsilon(nS) \to \mu^+\mu^-$ decays. For this measurement we use the triggers:

- `EF_mu18_medium`: single muon trigger with minimum p_{T} threshold of 18 GeV.
- `EF_2mu4T_DiMu_voVtx_noOS`: dimuon trigger with minimum p_{T} threshold of 4 GeV for both muons but no vertex quality or opposite sign charge requirements.
- `EF_2mu4T_Jpsi(Upsi)mumu`: dimuon trigger with minimum p_{T} threshold of 4 GeV for both muons, with vertex quality and opposite charge requirements

consistent with being produced from a J/$\psi(\Upsilon(nS))$. For events where the `EF_2mu4T_DiMu_voVtx_noOS` trigger has fired, the only difference in the `EF_2mu4T_Jpsi(Upsi)mumu` trigger are the additional vertex and opposite sign requirements.

To extract the efficiency we use a method similar to the early 2011 trigger efficiency measurement [84], which employs a data-driven *tag-and-probe* technique. We use the efficiency formula:

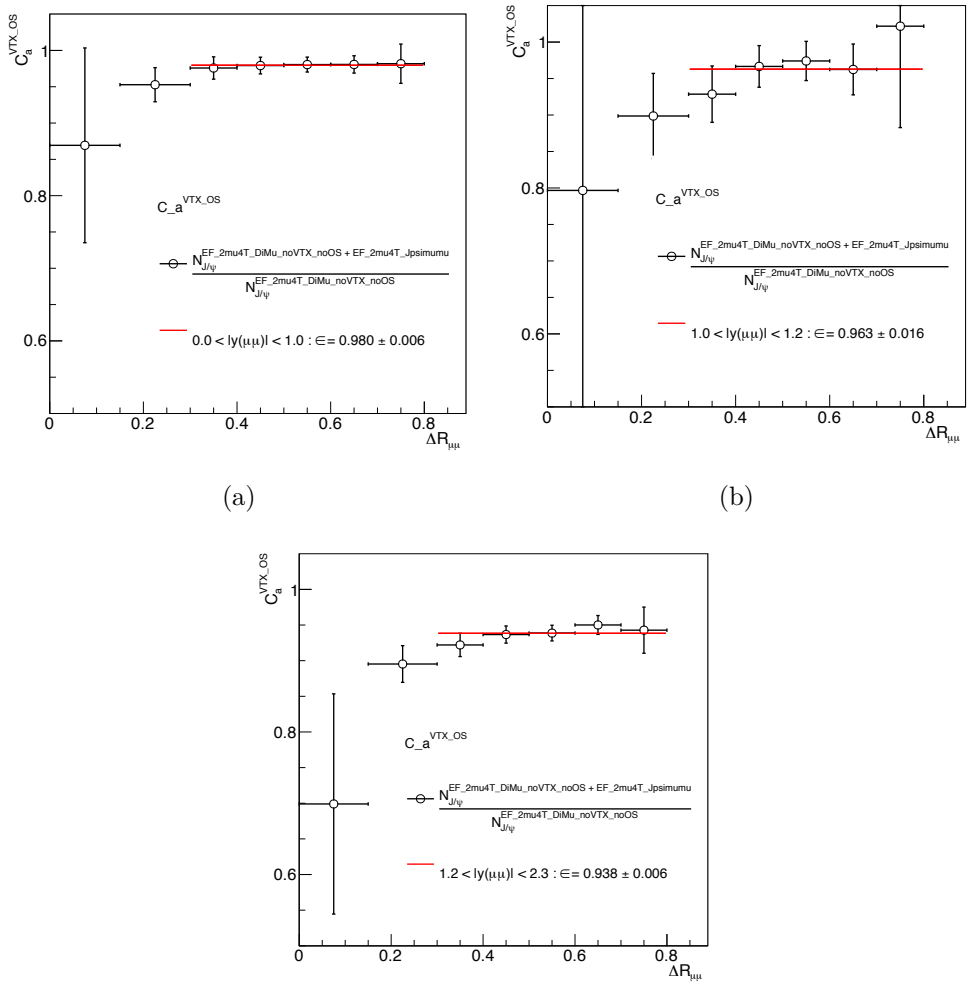
$$\epsilon_{\mu_2}^{\texttt{EF\_mu4T}} = \frac{N_{J/\psi}^{\texttt{EF\_mu18\_medium \& EF\_2mu4T\_Jpsimumu}}}{N_{J/\psi}^{\texttt{EF\_mu18\_medium}}} \times \frac{1}{c_{\Delta R}} \times \frac{1}{c_a^{VTX\_OS}} \tag{4.1}$$

where $c_{\Delta R}$ and $c_a^{VTX\_OS}$ are the correction factors for the effect of vertex and opposite sign requirements, and overlapping RoIs, respectively, for the the `EF_2mu4T_Jpsimumu` dimuon trigger. These two correction factors are subsequently combined to create the overall dimuon correction $c_{\mu\mu}$, which should be used in addition to each single muon efficiency for a pair of muons to find the overall `EF_2mu4T_Jpsimumu` trigger efficiency.

4.2.1 The Dimuon Correction Factor $c_a^{VTX\_OS}$

$c_a^{VTX\_OS}$ is a correction to the dimuon trigger efficiency due to signal loss from the vertex χ^2 cut and opposite sign requirements. This correction is measured in three separate rapidity bins each covering distinct regions of the detector. These are: $|y(\mu\mu)| < 1.0$ (barrel), $1.0 < |y(\mu\mu)| < 1.2$ (transition region) and $1.2 < |y(\mu\mu)| < 2.3$ (endcaps) for $\Delta R(\mu\mu) > 0.3$. $c_a^{VTX\_OS}$ has been shown to be roughly constant across $\Delta R(\mu\mu)$ above this value (see Figure 4.3). For $\Delta R(\mu\mu) < 0.3$, the signal efficiency further descreases due to the closeness of the muons in ΔR, and is investigated in the next Section 4.2.2. To calculate $c_a^{VTX\_OS}$ we find the ratio between those reconstructed J/$\psi \to \mu^+\mu^-$ events passing the `EF_2mu4T_DiMu_voVtx_noOS` trigger and those subsequently passing the `EF_2mu4T_Jpsimumu` trigger across a further subset of $|y(\mu\mu)|$ bins for the barrel, transition and endcap regions. This is done by fitting the invariant mass of the dimuons with a double Gaussian for the signal region and a second-order polynomial for the background. The $\psi(2S)$ signal region is ignored by the fit. Examples of the fits in each rapidity subregion are shown in Figure 4.4. All dependencies cancel out in this ratio besides the effect of the vertex quality and opposite sign requirements, and the

subsequent ratios are fitted in the three rapidity regions with zeroth order polynomials (see Figure 4.5).



(a) (b)

(c)

Fig. 4.3 $c_a^{VTX\_OS}$ measured across $\Delta R(\mu\mu)$ for the three rapidity sub-regions of the detector: (a) the barrel region, (b) transition region and (c) endcap region. We see only slight variation of $c_a^{VTX\_OS}$ when $\Delta R(\mu\mu) > 0.3$. This is therefore used as a cut when measuring $c_a^{VTX\_OS}$ as a function of J/ψ rapidity.

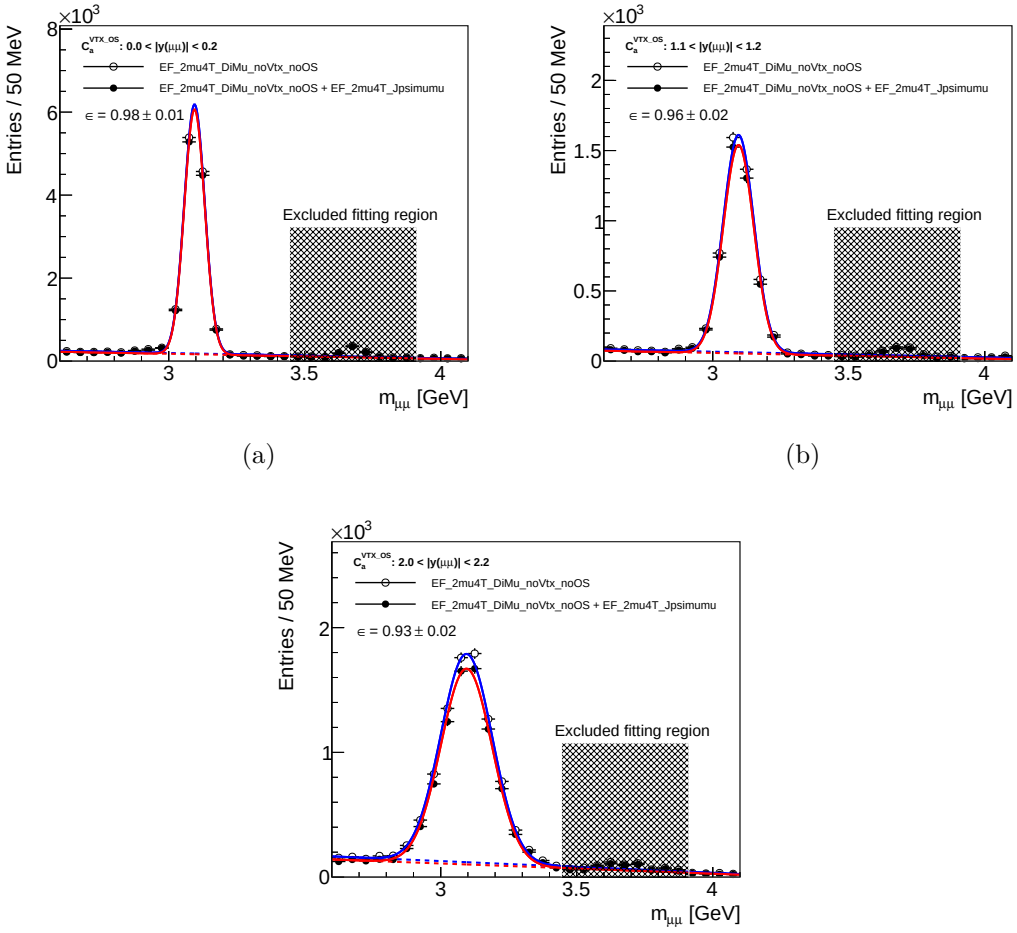


Fig. 4.4 Examples of the tag-and-probe fits used to extract J/ψ signal in bins of $|y(\mu\mu)|$ for $\Delta R(\mu\mu) > 0.3$, measuring $c_a^{VTX\_OS}$ in each bin. These highlighted fits show a range of $|y(\mu\mu)|$. The fit model is described in the text.

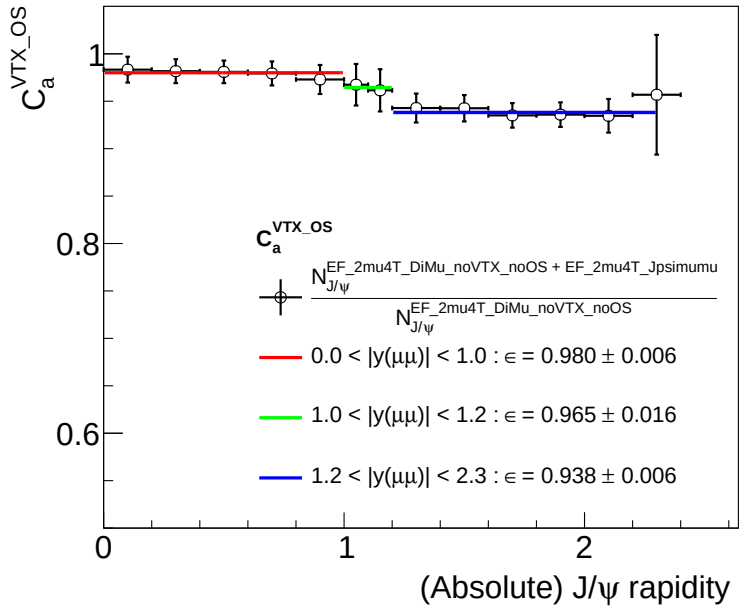


Fig. 4.5 $c_a^{VTX\_OS}$ measured across multiple $|y(\mu\mu)|$ rapidity sub-regions, fitted with zeroth order polynomials for the three rapidity regions of the detector: barrel (red), transition region (green) and endcaps (blue), to reduce statistical error.

4.2.2 The Dimuon Correction Factor $c_{\Delta R}$

The dimuon trigger efficiency is also dependent on the spatial separation of the two muons, with small separation leading to overlapping RoIs and hence a reduced efficiency. We measure the ratio of the signal yields in $\Delta R_{\mu\mu}$ bins:

$$\rho_{\Delta R}(\Delta R_{\mu\mu}, |y(\mu\mu)|) = \frac{N_{\Delta R}(\texttt{EF\_mu18\_medium \& EF\_2mu4T\_Xmumu})}{N_{\Delta R}(\texttt{EF\_mu18\_medium})} \tag{4.2}$$

for the same three rapidity regions as the $c_a^{VTX\_OS}$ measurement. The specific dimuon trigger used depends in which ΔR bin the signals are being measured. Due to phase space limitations, above $\Delta R_{\mu\mu} \approx 0.35$ the $J/\psi \to \mu^+\mu^-$ signal dies out. For ΔR values larger than this (up to 0.9) the subsequent dimuon trigger used is `EF_2mu4T_Upsimumu` and $\Upsilon(nS) \to \mu^+\mu^-$ $(n = 1, 2, 3)$ yields are measured. For smaller ΔR (less than 0.35) the `EF_2mu4T_Jpsimumu` trigger is used. Examples of the fits are shown in Figure 4.6. For the J/ψ signal region, we fit the invariant mass of the dimuon candidates with

a double Gaussian for the signal and a second-order polynomial for the background, where the nearby $\psi(2S)$ signal region is ignored by the fit. For the $\Upsilon(nS)$ signal region, double Gaussians are used to parameterise the $\Upsilon(1S)$, $\Upsilon(2S)$ and $\Upsilon(3S)$ mass peaks. The widths are left free to be determined by the fit, while the means are fixed to the world-average values. The background is described by an exponential function.

The subsquent $\rho_{\Delta R}$ values are fitted with an error function (which describes the shape of the ΔR turn-on reasonably well). The $\rho_{\Delta R}$ fits, in addition to describing the ΔR turn-on, also describe the effect of $c_a^{VTX\_OS}$ *and* the effect of the single muon trigger on the lower p_T muon. To ensure this dependence does not influence the shape of the turn-on, we impose a minimum p_T cut on the second muon of 8 GeV, which places it in the plateau region of the single muon efficiency. For the final $c_{\Delta R}$ corrections, we must normalise the plateau value for each of the $\rho_{\Delta R}$ fits to one. The results of the $\rho_{\Delta R}$ fits and $c_{\Delta R}$ are shown in Figure 4.7.

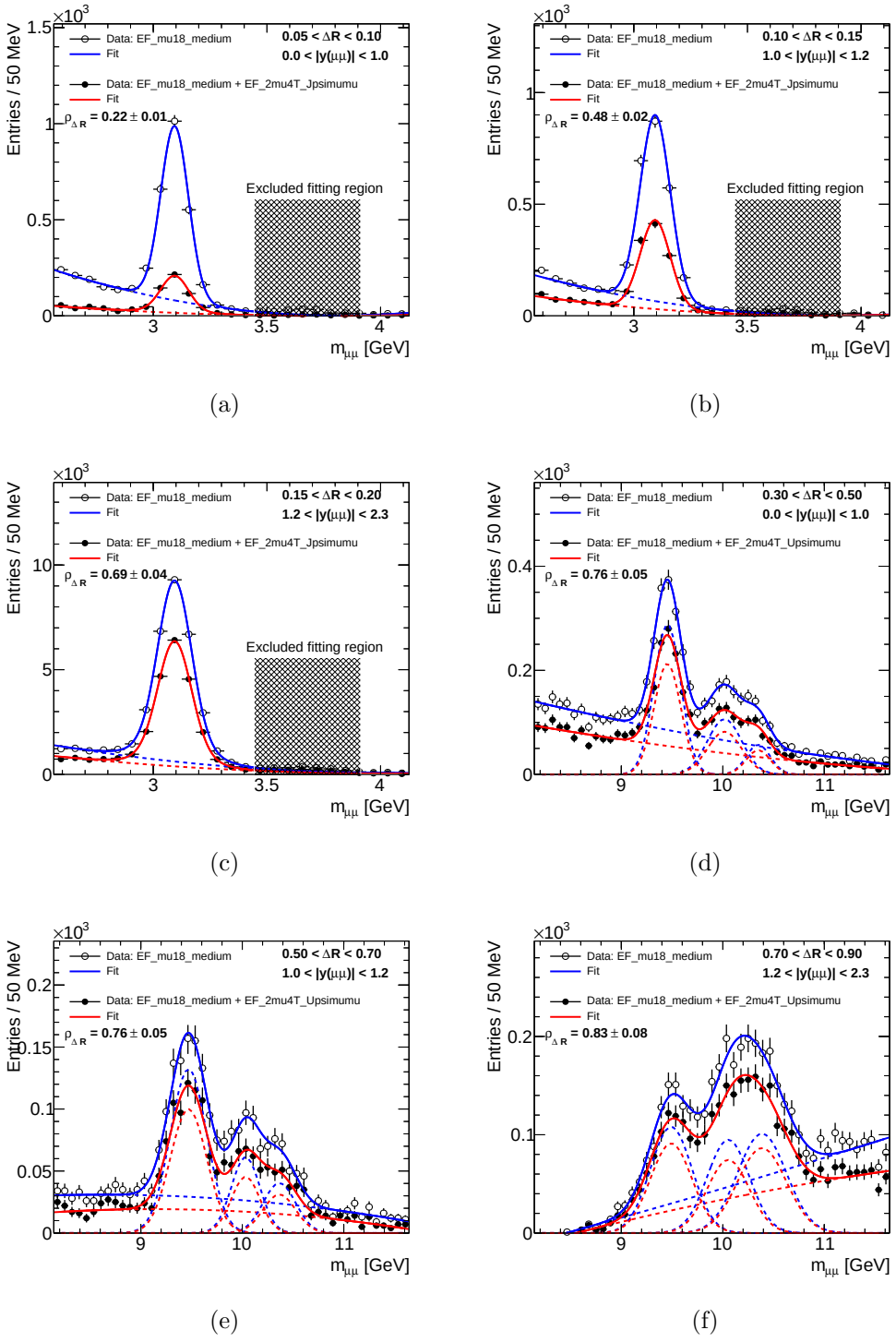


Fig. 4.6 (a) - (c) Fits to the dimuon invariant mass spectrum to measure J/ψ signal, using a tag-and-probe technique to calculate $\rho_{\Delta R}$, for a variety of $\Delta R(\mu\mu)$ and $|y(\mu\mu)|$ bins. For larger $\Delta R(\mu\mu)$, the J/ψ signal runs out of phase space, and so fits to the heavier Υ mesons are performed instead. Examples of these are shown in (d) - (f).

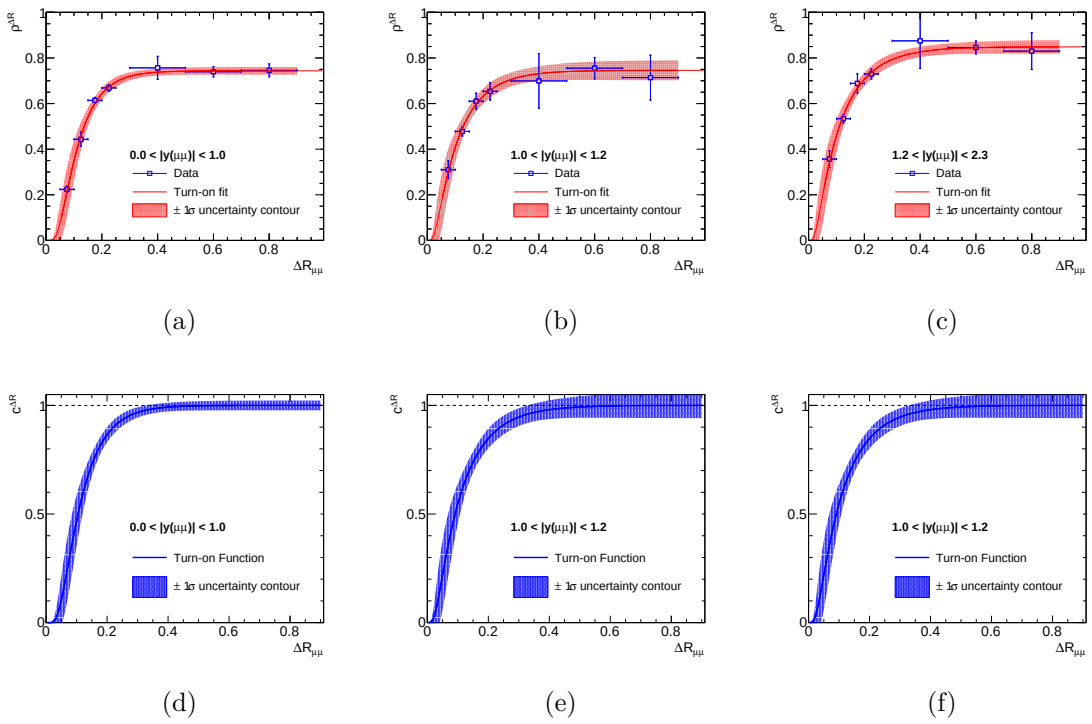


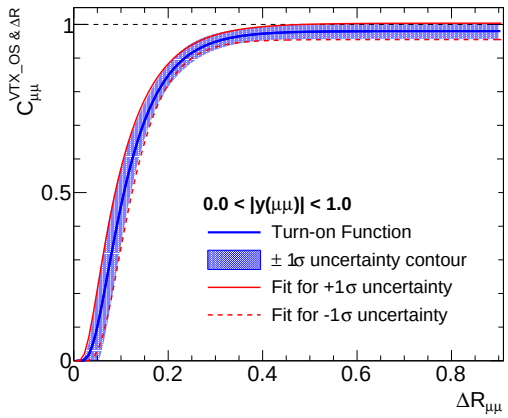
Fig. 4.7 The upper row shows $\rho_{\Delta R}$ ratios plotted as a function of ΔR for each of the three rapidity regions: (a) barrel, (b) transition region, (c) endcaps, fitted with an error function. Also shown are the $\pm 1\sigma$ variations in the fits. The lower row shows the error functions with their plateau regions normalised to 1.

4.2.3 The Combined Correction Factor $c_{\mu\mu}$

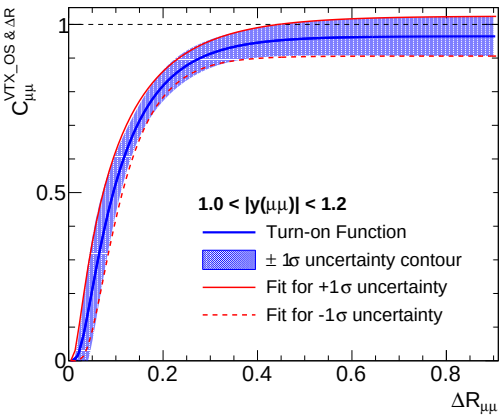
The total dimuon correction $c_{\mu\mu}$ is the combined effect of $c_a^{VTX\_OS}$ and $c_{\Delta R}$,

$$c_{\mu\mu}(\Delta R, |y(\mu\mu)|) = c_a^{VTX\_OS}(|y(\mu\mu)|) \times c_{\Delta R}(\Delta R, |y(\mu\mu)|). \tag{4.3}$$

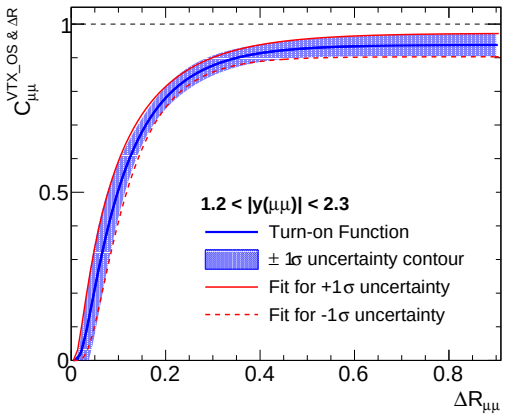
In Figure 4.8 we therefore scale each $c_{\Delta R}$ plot by its respective $c_a^{VTX\_OS}$ to achieve the total dimuon correction.



(a)



(b)



(c)

Fig. 4.8 Total dimuon efficiency corrections $c_{\mu\mu}(\Delta R, |y(\mu\mu)|)$ for each of the three $|y(\mu\mu)|$ rapidity regions. Also shown are the error bounds generated from $\pm 1\sigma$ variations of the $\rho_{\Delta R}$ fits combined with the errors from the $c_a^{VTX\_OS}$ measurements.

4.2.4 Single Muon Efficiency Map

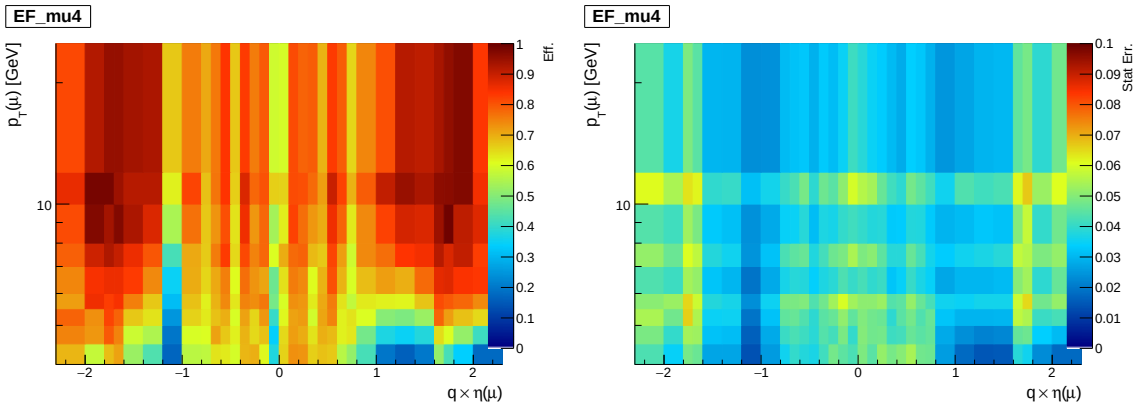
We measure the single muon efficiency following:

$$\epsilon(p_T^{\mu_2}, q^{\mu_2} \times \eta^{\mu_2}) = \frac{N_{J/\psi}(\texttt{EF\_mu18\_medium \& EF\_2mu4T\_Jpsimumu})}{N_{J/\psi}(\texttt{EF\_mu18\_medium}) \times c_{\mu\mu}(\Delta R, |y(\mu\mu)|)}. \tag{4.4}$$

To do this we measure the ratio between the number of $J/\psi \to \mu^+\mu^-$ signal events which pass the `EF_mu18_medium` trigger and the additional subset of events passing the `EF_2mu4T_Jpsimumu` trigger, however each event belonging to the subset is also scaled by its corresponding $1/c_{\mu\mu}(\Delta R, |y(\mu\mu)|)$ to correct for loss in dimuon trigger efficiency. The single muon efficiency is measured in $(p_T^{\mu_2}, q^{\mu_2} \times \eta^{\mu_2})$ bins, and the resulting map is shown in Figure 4.9(a).

Also shown in Figure 4.9(b) is the statistical error map for the single muon efficiency. Systematic uncertainties were also measured by varying: between single and double Gaussians for the signal, and 1^{st} and 2^{nd} order polynomials for the background, the fit range, and the bin widths of the fits. The subsequent systematic errors were found to be much smaller than the statistical uncertainty.

For comparison, we also show single muon efficiency maps for early 2011 data in Figure 4.10. This map had to be split into two separate samples due to a change in trigger matching criteria from data period **H** onwards.



(a) (b)

Fig. 4.9 *a*) Single muon efficiency map for late 2011 data periods in $(p_T, q \times \eta)$ bins of the lower p_T muon, derived from $J/\psi \to \mu^+\mu^-$ signal events. *b*) the corresponding statistical error map.

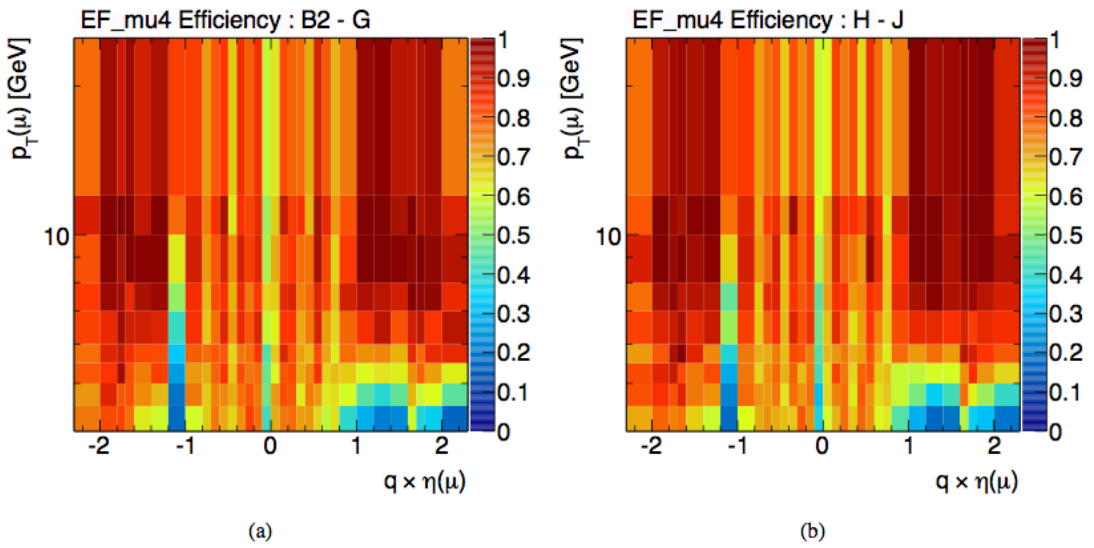


Fig. 4.10 Single muon efficiency maps for early 2011 data periods (*a*) **B2** - **G** and (*b*) **H** - **J**, previously determined in [84].

In summary, we use the formula shown in Equation 4.1 to measure single muon efficiency in bins of muon $(p_{\mathrm{T}}, q \times \eta)$ for late 2011 data. The dimuon trigger efficiency is lowered due to its vertex and opposite sign requirements on the dimuons, and also when the ΔR between the muons is small (due to RoI overlap). Dimuon trigger correction factors are derived for the vertex and opposite sign requirements, $c_a^{VTX\_OS}(|y(\mu\mu)|)$, in Section 4.2.1. The correction for the muon RoI overlap, $c_{\Delta R}(\Delta R, |y(\mu\mu)|)$, is described in Section 4.2.2. The single muon efficiency map is then created, shown in Figure 4.9(a).

Chapter 5

$\psi(2S)$ and $X(3872)$ Analysis

In this chapter, we describe the analysis performed to extract prompt and non-prompt $\psi(2S)$ and $X(3872)$ cross sections, along with their relevant ratios and fractions, in the decay channel $J/\psi\pi^+\pi^-$. In Section 5.1, we describe the selection requirements for the $J/\psi\pi^+\pi^-$ candidates, and measure the unweighted $\psi(2S)$ and $X(3872)$ signal candidates after the selection requirements. We briefly outline the analysis method used to extract the prompt and non-prompt signal in Section 5.2, and in Section 5.3 we detail the weights used to correct for signal loss due to: the detector acceptance, trigger and reconstruction efficiencies, and also the additional selection requirements. Section 5.4 provides a detailed description of the signal extraction method, for two separate descriptions of the non-prompt $\psi(2S)$ and $X(3872)$ signal. The first, where the non-prompt signal originates purely from long-lived b-hadrons, and the second where we hypothesise an additional, shorter-lived B_c component.

5.1 Event Selection

Events used in this analysis are triggered by a pair of muons successfully fitted to a common vertex, using the dimuon trigger `EF_2mu4T_Jpsimumu_L2StarB`, which was active between the data taking periods **C6** - **L** in 2012. This trigger differs slightly from the `EF_2mu4T_Jpsimumu` trigger used in most 2011 ATLAS quarkonia analyses, described in Section 4.2, by using a new software framework for the tracking trigger at Level 2, named L2Star [85]. The `EF_2mu4T_Jpsimumu` trigger was found to have a lifetime bias during the 2012 run period, so the majority of 2012 quarkonia analyses

use the L2Star alternative trigger. The data sample corresponds to an integrated luminosity of 11.4 fb^{-1}, collected at a proton-proton collision energy $\sqrt{s} = 8$ TeV. Detail on the specific datasets used for the analysis can be found in Appendix A.

The dimuon trigger requires two oppositely-charged muons to originate from a common vertex through a constrained fit with a very loose χ^2 cut on the event selection, giving negligible signal loss. Events are chosen containing ≥ 2 muons, where each of the muons is required to have formed a combined track through the inner detector and muon spectrometer as identified through the *Staco* algorithm [86], which is described in Section 3.5.1. These combined muons must each have $p_T \geq 4$ GeV and $|\eta| < 2.3$. Additionally, the muon candidates are subject to the Muon Combined Performance (MCP) requirements detailed in [87].

We perform a constrained vertex fit of the J/ψ and pairs of the remaining tracks, for which we use the pion mass hypothesis and require the charge of these reconstructed tracks to have opposite signs and to also satisfy $\mathrm{p}_T > 0.6$ GeV, $|\eta| < 2.4$. We also impose a loose quality cut on the $J/\psi\pi^+\pi^-$ vertex, requiring the difference between the longitudinal impact parameter, z_0, of the reconstructed J/ψ vertex and the pion candidate tracks to be less than 50mm, which is imposed before the four-track vertexing takes place and has been shown with simulation samples to cause negligible impact on signal yield. Each reconstructed muon is required to have good spatial matching to a trigger object of $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} < 0.01$, in order to accurately correct for the efficiency of the trigger.

Missing trigger objects were discovered in the data for around 4% of the reconstructed events (due to a known bug in the B-physics code), so for these cases we remove the matching requirement. This way of dealing with the missing trigger objects is consistent with other analyses [88]. The loss of events due to the trigger object requirement is already small ($< 1\%$) and so the effect of skipping this requirement on a small amount of the dataset is expected to be negligible.

Significance studies were performed on the data by introducing additional selection criteria, to suppress background while retaining high signal yields, and it was found that a χ^2 probability cut on the $J/\psi\pi^+\pi^-$ vertex, $P(\chi^2) > 4\%$, gave the greatest $X(3872)$ signal signficance defined as $S = N_{\mathrm{signal}}/\sqrt{N_{\mathrm{signal}} + N_{\mathrm{background}}}$. Requiring an opening angle $\Delta R < 0.5$ between the J/ψ and each pion candidate was found to significantly boost $X(3872)$ signal significance, particularly at low p_T, by suppressing

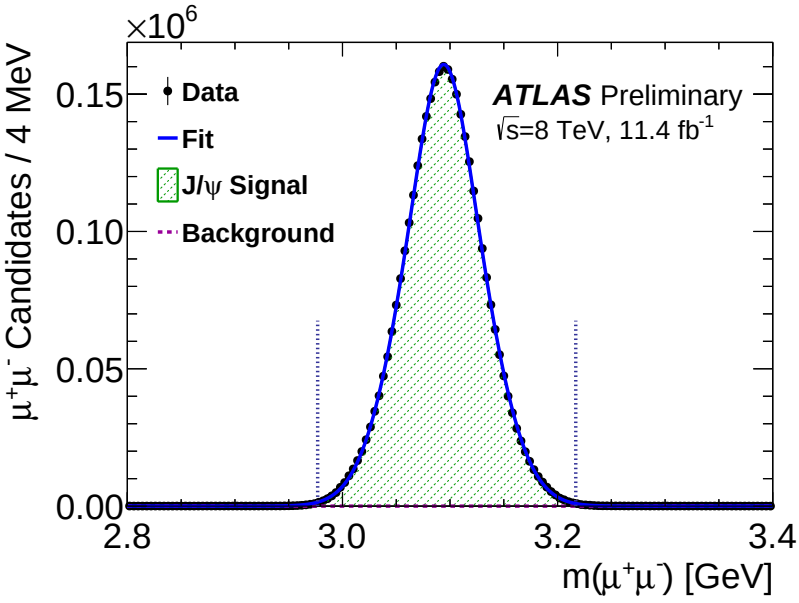
combinatoral background especially at larger $m_{J/\psi\pi^+\pi^-}$. A constraint on the value of $m_{J/\psi\pi^+\pi^-} - m^{\rm PDG}_{J/\psi} - m_{\pi^+\pi^-}$ to be less than 0.3 GeV was also imposed. This cut is similar to the ones imposed in previous $X(3872)$ analyses [89][7][8]. More detail on these signficance studies can be found in Appendix B.1. We summarise the selection criteria used for the analysis in Table 5.1. The efficiencies of these selections were studied differentially in p_T with Monte Carlo, with more detail in Appendix B.2.

A fit to the invariant mass of the selected candidates is shown in Figure 5.1(b). We model the $\psi(2S)$ and $X(3872)$ signal with double Gaussians, and the background with a fourth order Chebyshev polynomial. We measure total uncorrected $\psi(2S)$ and $X(3872)$ yields of roughly 470,000 and 30,000 events, respectively, for the total phase space in which the measurement takes place ($|y(J/\psi\pi^+\pi^-)| < 0.75$, 10 GeV $< p_T(J/\psi\pi^+\pi^-) <$ 70 GeV).

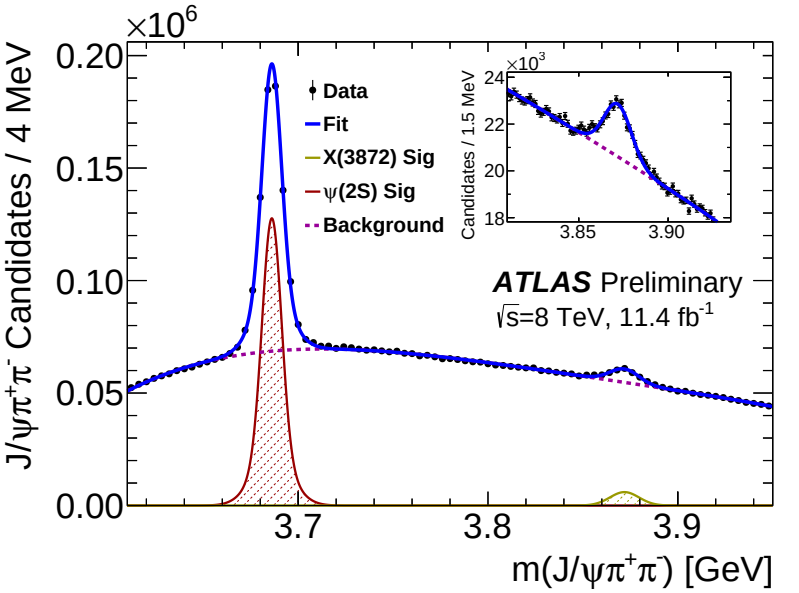
The total J/ψ sample for the $J/\psi\pi^+\pi^-$ candidates, including selection cuts, is shown in Figure 5.1(a). We observe an extremely clean J/ψ sample with very small background. A double Gaussian is used to fit the signal, which is required due to the large span in rapidity and p_T of the J/ψ candidates, and a second order Chebyshev polynomial parameterises the background. We measure approximately 3.6 M J/ψ candidates. The dotted vertical lines in Figure 5.1(a) show the cut in J/ψ mass of the $J/\psi\pi^+\pi^-$ candidates, requiring $\left|m\left(J/\psi\right) - m\left(J/\psi^{\rm PDG}\right)\right| < 120$ MeV.

cut description	cut values
reco quality	$q_{\mu_1}\cdot q_{\mu_2} < 0 \quad q_{\pi^+}\cdot q_{\pi^-} < 0$ $\chi^2\left(J/\psi\right) < 200$ $p_T\left(J/\psi\right) > 8$ GeV $\quad \left\|y\left(J/\psi\right)\right| < 2.3$ $\left\|m\left(J/\psi\right) - m\left(J/\psi^{\rm PDG}\right)\right\| < 120$ MeV $P\left(\chi^2\right)\left(J/\psi\pi\pi\right) > 0.005$
$P\left(\chi^2\right)\left(J/\psi\pi\pi\right)$	> 0.04
$\Delta R\left(J/\psi,\pi\right)$	< 0.5
$Q\left(J/\psi\pi\pi\right) = m_{J/\psi\pi^+\pi^-} - m^{\rm PDG}_{J/\psi} - m_{\pi^+\pi^-}$	< 300 MeV

Table 5.1 List and description of selection criteria applied to candidates to suppress the background.



(a)



(b)

Fig. 5.1 (a) The invariant mass distribution of the J/ψ candidates satisfying all selection criteria except the ± 120 MeV J/ψ mass window requirement indicated here by the dotted vertical lines. The curve shows the result of a fit with a double Gaussian for signal and a second order polynomial for background. (b) Invariant mass of the selected $J/\psi\pi^+\pi^-$ candidates collected over the full p_{T} range $10-70$ GeV and the rapidity range $|y| < 0.75$ after selection cuts. The curve shows the results of the fit using double Gaussians for the $\psi(2S)$ and $X(3872)$ peaks and a fourth-order polynomial for the background. The X(3872) mass range is highlighted in the inset.

5.2 Outline of the Method

The production cross sections of the $\psi(2S)$ and $X(3872)$ states decaying to $J/\psi\pi^+\pi^-$ are measured in five bins of $J/\psi\pi^+\pi^-$ transverse momentum p_{T}, with bin boundaries $(10, 12, 16, 22, 40, 70)$ GeV. The selected $J/\psi\pi^+\pi^-$ candidates are weighted in order to correct for signal loss at various stages of the selection process. Following previous similar analyses [6, 10] a per-candidate weight ω is calculated as

$$\omega = \left[\mathcal{A}(p_{\mathrm{T}}, y) \cdot \epsilon_{\mathrm{trig}}(p_{\mathrm{T}}, y) \cdot \epsilon^{\mu}(p_{\mathrm{T}}^{\mu+}, \eta^{\mu+}) \cdot \epsilon^{\mu}(p_{\mathrm{T}}^{\mu-}, \eta^{\mu-}) \cdot \epsilon^{\pi}(p_{\mathrm{T}}^{\pi+}, \eta^{\pi+}) \cdot \epsilon^{\pi}(p_{\mathrm{T}}^{\pi-}, \eta^{\pi-})\right]^{-1}. \tag{5.1}$$

The trigger efficiency ϵ_{trig} and the muon reconstruction efficiency ϵ^{μ} were obtained using data-driven tag-and-probe methods consistent with [10, 90] and are detailed in Sections 5.3.1 and 5.3.2, respectively. The pion reconstruction efficiency ϵ^{π} was obtained through MC simulations using a method consistent with previous ATLAS analyses [6], with detail in Section 5.3.3.

The acceptance, $\mathcal{A}(p_{\mathrm{T}}, y)$, is defined as the probability that the muons and pions comprising a $J/\psi\pi^+\pi^-$ candidate with transverse momentum p_{T} and rapidity y fall within the fiducial limits described in Section 5.3.4. The acceptance correction maps were created using the same generator-level simulation method as the majority of previous ATLAS quarkonium production measurements [6, 10]. The difference in the quantum numbers of $\psi(2S)$ and $X(3872)$ ($J^{PC} = 1^{--}$ and 1^{++}, respectively) shows up as a difference in the expected dependence of acceptance on the spin alignment of the two states. This leads us to create two different sets of acceptance correction maps for the $\psi(2S)$ and $X(3872)$, with respective explanations in Sections 5.3.5 and 5.3.6.

In addition to the per-event weights described above, in each of the p_{T} bins the efficiencies of the other background suppression cuts (shown in Table 5.1) are determined using MC simulation and are corrected for, separately for $\psi(2S)$ and $X(3872)$. The simulated distributions are reweighted to match the data, and the subsequent efficiencies are found to be between $(84 - 95)\%$. The details of the MC generation and efficiency determination are found in Section 5.3.8.

In order to separate prompt production of the $\psi(2S)$ and $X(3872)$ states and the non-prompt production from the decays of long-lived particles such as b-hadrons, the data sample in each p_{T} bin is further divided into intervals of pseudo-proper lifetime τ,

defined as

$$\tau = \frac{L_{xy} m}{c\, p_T}, \tag{5.2}$$

where m is the invariant mass, p_T is the transverse momentum and L_{xy} is the transverse decay length of the $J/\psi\pi^+\pi^-$ candidate. L_{xy} is defined as

$$L_{xy} = \frac{\vec{L} \cdot \vec{p}_T}{p_T}, \tag{5.3}$$

where $\vec{L}$ is the vector pointing from the primary pp collision vertex to the $J/\psi\pi^+\pi^-$ vertex, while $\vec{p}_T$ is the transverse momentum vector of the $J/\psi\pi^+\pi^-$ system. The coordinates of the primary vertices (PV) are obtained from charged tracks not used in the decay vertices, and are transversely constrained to the luminous region of the colliding beams. The matching of a $J/\psi\pi^+\pi^-$ candidate to a PV is made by finding the one with the smallest three-dimensional impact parameter, calculated between the $J/\psi\pi^+\pi^-$ momentum and each PV.

Based on the analysis of the lifetime resolution and lifetime dependence of the signal, four lifetime intervals were defined:

$$w_0 : -0.3\,\text{ps} < \tau(J/\psi\pi\pi) < 0.025\,\text{ps},$$

$$w_1 : 0.025\,\text{ps} < \tau(J/\psi\pi\pi) < 0.3\,\text{ps},$$

$$w_2 : 0.3\,\text{ps} < \tau(J/\psi\pi\pi) < 1.5\,\text{ps},$$

$$w_3 : 1.5\,\text{ps} < \tau(J/\psi\pi\pi) < 15.0\,\text{ps}.$$

In each of these intervals, and for each p_T bin, the invariant mass distribution of the $J/\psi\pi^+\pi^-$ system is built using fully corrected weighted events. Fits are then performed to the invariant mass distribution to simultaneously extract $\psi(2S)$ and $X(3872)$ signal for each p_T and pseudoproper lifetime bin. Subsequently, lifetime fits are performed on these signal yields to extract the prompt and non-prompt signal yields Y^ψ, Y^X. Once these yields are determined in each p_T bin, the double differential cross sections (times the product of the relevant branching fractions) can be calculated:

$$Br(i \to J/\psi\pi^+\pi^-) Br(J/\psi \to \mu^+\mu^-) \frac{d^2\sigma(i)}{dp_T dy} = \frac{Y^i}{\Delta p_T \Delta y \int \mathcal{L} dt}, \tag{5.4}$$

where i stands for $\psi(2S)$ or $X(3872)$, $\int \mathcal{L}dt$ is the total integrated luminosity, while Δp_T and Δy are widths of the relevant transverse momentum and rapidity bins, with $\Delta y = 1.5$. $Br(i \to J/\psi\pi^+\pi^-)$ and $Br(J/\psi \to \mu^+\mu^-)$ are the branching fractions of these respective decays, and their values are shown in Table 5.2.

Decay channel	Branching Fraction [%]
$J/\psi \to \mu^+\mu^-$	5.93 ± 0.06
$\psi(2S) \to J/\psi\pi^+\pi^-$	34.46 ± 0.30
$X(3872) \to J/\psi\pi^+\pi^-$	$4.2 - 9.3$

Table 5.2 Branching fractions of decay channels relevant to this analysis. The first two values are taken from [18]. The last value is estimated in [61].

We also measured the yields of non-prompt $\psi(2S)$, $X(3872)$ as a fraction of the inclusive samples:

$$f^i_{\mathrm{NP}} = \frac{Y^i_{\mathrm{NP}}}{Y^i_{\mathrm{P}} + Y^i_{\mathrm{NP}}} \tag{5.5}$$

which benefits from a reduction in systematic uncertainty due to the partial cancellation of the errors in the numerator and denominator terms.

The cross sections measured in this thesis are obtained under the assumption of no spin alignment, but appropriate sets of correction factors for a number of extreme spin alignment scenarios are calculated and presented in Section 6.2 for each p_{T} bin, separately for $\psi(2S)$ and $X(3872)$.

5.3 Event Weights

5.3.1 Trigger Efficiency

The efficiency of the dimuon trigger `EF_2mu4T_Jpsimumu_L2StarB` for the 2012 run period was previously measured for other quarkonia analyses [10]. The method follows similar steps to previous measurements of the dimuon trigger efficiency, which was developed for 7 TeV data, and is described in detail in Section 4.2.

The dimuon trigger efficiency ϵ_{trig} can be expressed as the product of the two single muon efficiencies ϵ_{RoI} as

$$\epsilon_{\text{trig}} = \epsilon_{\text{RoI}}\left(p_T^1, q^1 \cdot \eta^1\right) \cdot \epsilon_{\text{RoI}}\left(p_T^2, q^2 \cdot \eta^2\right) \cdot c_{\mu\mu}\left(\Delta R, |y_{\mu\mu}|\right), \tag{5.6}$$

with some correlation correction factor $c_{\mu\mu}$, which is the combined correction for the effect of overlapping RoI's, and vertex and opposite sign requirements, which reduce the dimuon trigger efficiency. The single muon efficiency map $\epsilon_{\text{RoI}}\left(p_T, q\right)$ for 2012 data, measured by following Equation 4.1, is shown in Figure 5.2. The typical shapes of $c_{\mu\mu}\left(\Delta R, |y_{\mu\mu}|\right)$ can be found in Section 4.2.3.

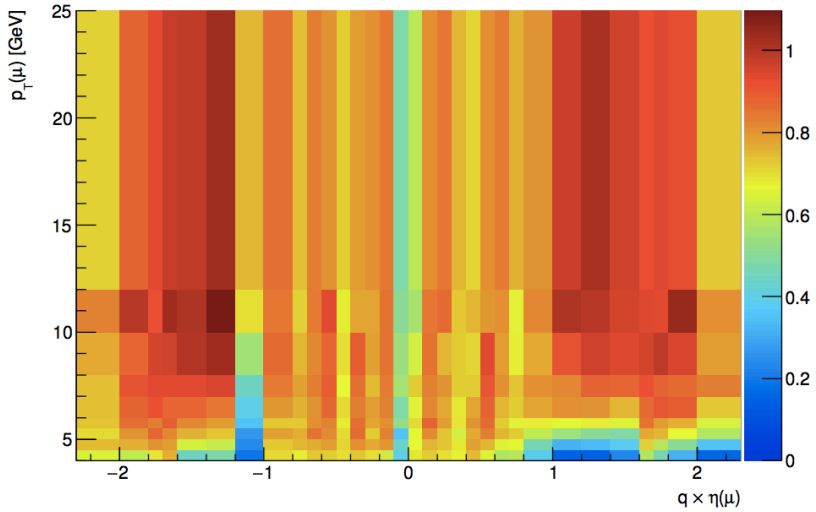


Fig. 5.2 Effective trigger efficiency map for single muons in bins of $(p_{\text{T}}, q \times \eta)$ [91].

5.3.2 Muon Reconstruction Efficiency

We use the same muon reconstruction maps as described in [87], which used a sample of approximately 2M $J/\psi \to \mu\mu$ decays that are unbiased by dimuon trigger requirements due to the use of single muon and 'muon + track' triggers. A tag-and-probe method was used to measure the efficiency. For $J/\psi \to \mu\mu$ events, the tag is defined as a combined muon, and the probes are ID tracks which satisfy various constraints without significant loss of J/ψ reconstruction efficiency. The reconstruction efficiency is measured in fine bins of probe muon (p_T, η) by finding the ratio of events where the tag-and-probe pairs are found but the probe was not successfully reconstructed in the MS, to events where

both the tag and probe are found, and subtracting this 'inefficiency' from unity. There is an additional efficiency of muon track reconstruction from the inner detector, which is taken to be $99 \pm 1\%$ [10, 90].

5.3.3 Pion Reconstruction Efficiency

The pion reconstruction efficiency is measured using pions from $J/\psi \to \mu^+\mu^-$ Monte Carlo, simulated under ATLAS 2012 run conditions. We require the closest reconstructed track in an event to be within an opening angle $\Delta R < 0.05$ of truth pions. If the track condition is satisified then the pion is counted as being reconstructed. The number of pions which are successfully matched are then divided by the total number of truth pions in bins of $(p_T, q \cdot \eta)$ to generate the pion reconstruction efficiency map, seen in Figure 5.3.

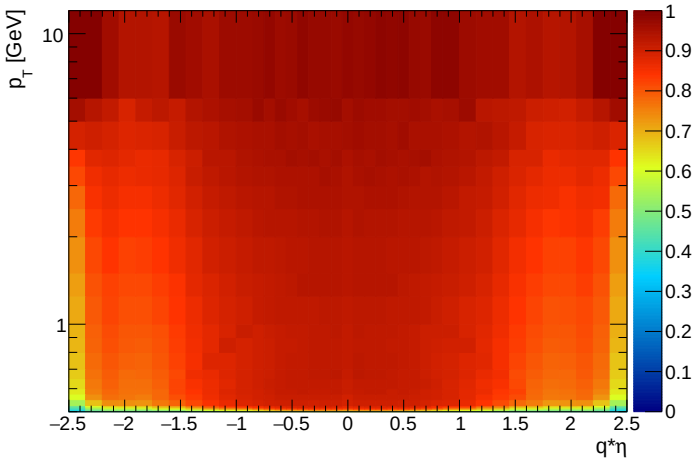


Fig. 5.3 Charged pion reconstruction efficiency map versus p_T and $q{\cdot}\eta$ from $J/\psi \to \mu^+\mu^-$ MC under 2012 run conditions.

The effect of pile-up on pion reconstruction has been investigated and was found to cause small ($< 1\%$) differences in the efficiency map, which have been included as a systematic. Small differences in efficiency for various ATLAS geometry tags (which model the ATLAS detector in simulation) have also been investigated, and also the effect of tightening the successful match criteria between the truth pion and reconstructed track to an opening angle of less than 0.03. Additionally, the simulated events are reweighted to match the data distribution of z vertex displacement, and the

difference in reconstruction efficiency is found to be $< 0.5\%$ in each bin. More detail on these studies can be found in Appendix C, where we conclude the overall track reconstruction efficiency uncertainty of each pion to be $\pm 1.25\%$, which is conservatively added for both pions to give a dipion reconstruction efficiency uncertainty $\pm 2.5\%$.

5.3.4 Fiducial Region and Acceptance

The acceptance is defined as the probability that the $\mu^+\mu^-\pi^+\pi^-$ final decay products from the mother $\psi(2S)$ or $X(3872)$ travel through a fiducial region of the detector. This fiducial region is defined as:

- $p_T(\mu^\pm) > 4$ GeV and $|\eta(\mu^\pm)| < 2.3$
- $p_T(\pi^\pm) > 0.6$ GeV and $|\eta(\pi^\pm)| < 2.4$

The acceptance strongly depends on spin alignment, and acceptance maps for several polarisation hypotheses are created by means of re-weighting 'central' unpolarised event samples. High statistics generator-level Monte Carlo is produced separately for $\psi(2S)$ and $X(3872)$ due to their different masses, and also due to the difference in calculating the event weights for the different polarisation scenarios. In the case of $\psi(2S) \to J/\psi\pi^+\pi^-$, the dipion system is known to be produced in a spatially isotropic state with respect to the J/ψ, so the J/ψ directly inherits its spin alignment from the mother $\psi(2S)$. In contrast, $X(3872)$ has been measured to decay to $J/\psi\pi^+\pi^-$ through an intermediate $\rho^0 \to \pi^+\pi^-$ decay, causing the dipion system to be no longer spatially isotropic. The acceptance maps are produced in fine bins of $[p_T,\ |y|,\ m_{\pi^+\pi^-}]$, with 100,000 events generated in each bin.

5.3.5 Acceptance Maps for $\psi(2S)$

As mentioned above, spin alignment of the J/ψ is directly inherited from the mother $\psi(2S)$, due to the dipion system primarily being produced in an S-wave state relative to the J/ψ. We express the angular dependence of the $J/\psi \to \mu^+\mu^-$ decay as:

$$\frac{\mathrm{d}^2\mathbf{N}}{\mathrm{d}\cos\theta^*\mathrm{d}\phi^*} \propto 1 + \lambda_\theta \cos^2\theta^* + \lambda_\phi \sin^2\theta^* \cos 2\phi^* + \lambda_{\theta\phi} \sin 2\theta^* \cos\phi^* \qquad (5.7)$$

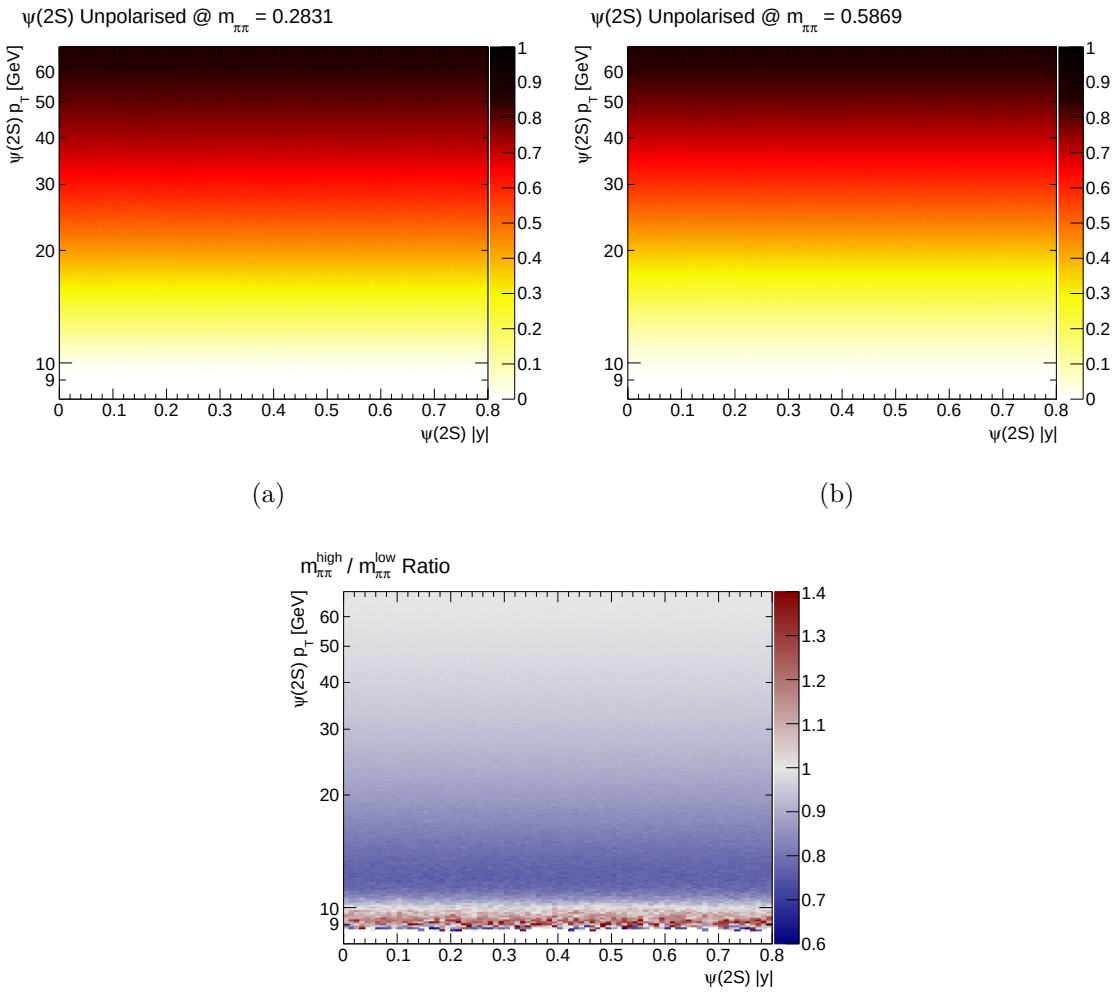
where λ_i are the three separate polarisation parameters and $\theta^*(\phi^*)$ are the polar(azimuthal) angles between the positive muon momentum in the $J/\psi \to \mu^+\mu^-$

decay frame and the direction of $\psi(2S)$ in the lab frame. The acceptance maps are built in fine bins across the ranges: 8 GeV $< p_T^{\psi(2S)} <$ 70 GeV, $|y|^{\psi(2S)} < 0.8$, $2m_\pi < m_{\pi^+\pi^-} < m_{\psi(2S)} - m_{J/\psi}$. The acceptance can vary substantially between various polarisation scenarios, and several extreme polarisation hypotheses are considered:

- *Unpolarised* - An isotropic distribution which is independent of the polarisation parameters *i.e.* $\lambda_\theta = \lambda_\phi = \lambda_{\theta\phi} = 0$, which we label UNPOL. This is used as the central hypothesis.
- *Transversely* polarised with $\lambda_\theta = +1, \lambda_\phi = \lambda_{\theta\phi} = 0$, which we label T_{+0}
- *Transversely* polarised with $\lambda_\theta = +1, \lambda_\phi = +1, \lambda_{\theta\phi} = 0$, which we label T_{++}
- *Transversely* polarised with $\lambda_\theta = +1, \lambda_\phi = -1, \lambda_{\theta\phi} = 0$ which we label T_{+-}
- *Longitudinal* - with $\lambda_\theta = -1, \lambda_\phi = \lambda_{\theta\phi} = 0$, which we label LONG
- *Off-Plane Positive* - with $\lambda_\theta = 0, \lambda_\phi = 0, \lambda_{\theta\phi} = +0.5$, which we label OFFP+
- *Off-Plane Negative* - with $\lambda_\theta = 0, \lambda_\phi = 0, \lambda_{\theta\phi} = -0.5$, which we label OFFP−

It is of note that there exists no individual production process that can lead to an unpolarised state, but only a fortunate cancellation of the spin alignments from a superposition of different production subprocesses [52]. The polarisation of $\psi(2S)$ has been measured by CMS [54] and LHCb [56], and it was found that each of the λ parameters were close to zero, justifying our choice of unpolarised production for the central hypothesis.

The acceptance maps generated for $\psi(2S) \to J/\psi\pi^+\pi^-$ decays with the lowest and highest dipion invariant mass, $m_{\pi\pi}$, along with their ratio, are shown in Figure 5.4. The acceptance increases as a function of p_T, and is largely independent of the rapidity range considered in this analysis. Acceptance decreases as a function of $m_{\pi\pi}$, as can be seen in Figure 5.4(c). In Figure 5.5 we show the acceptance maps for different polarisation scenarios, each divided by the unpolarised acceptance map, for some central $m_{\pi\pi} = 0.4381$ GeV bin. We see for the LONG case that the acceptance increases with respect to unpolarised acceptance. For the T_{+0} , T_{++} and T_{+-} scenarios, the acceptance decreases with respect to the unpolarised map, and for OFFP+ and OFFP− the acceptance is very similar to the unpolarised case.



(a) (b)

(c)

Fig. 5.4 Acceptance maps for unpolarised $\psi(2S)$ decays for (a) low $m_{\pi\pi}$, (b) high $m_{\pi\pi}$, and (c) the ratio between the two.

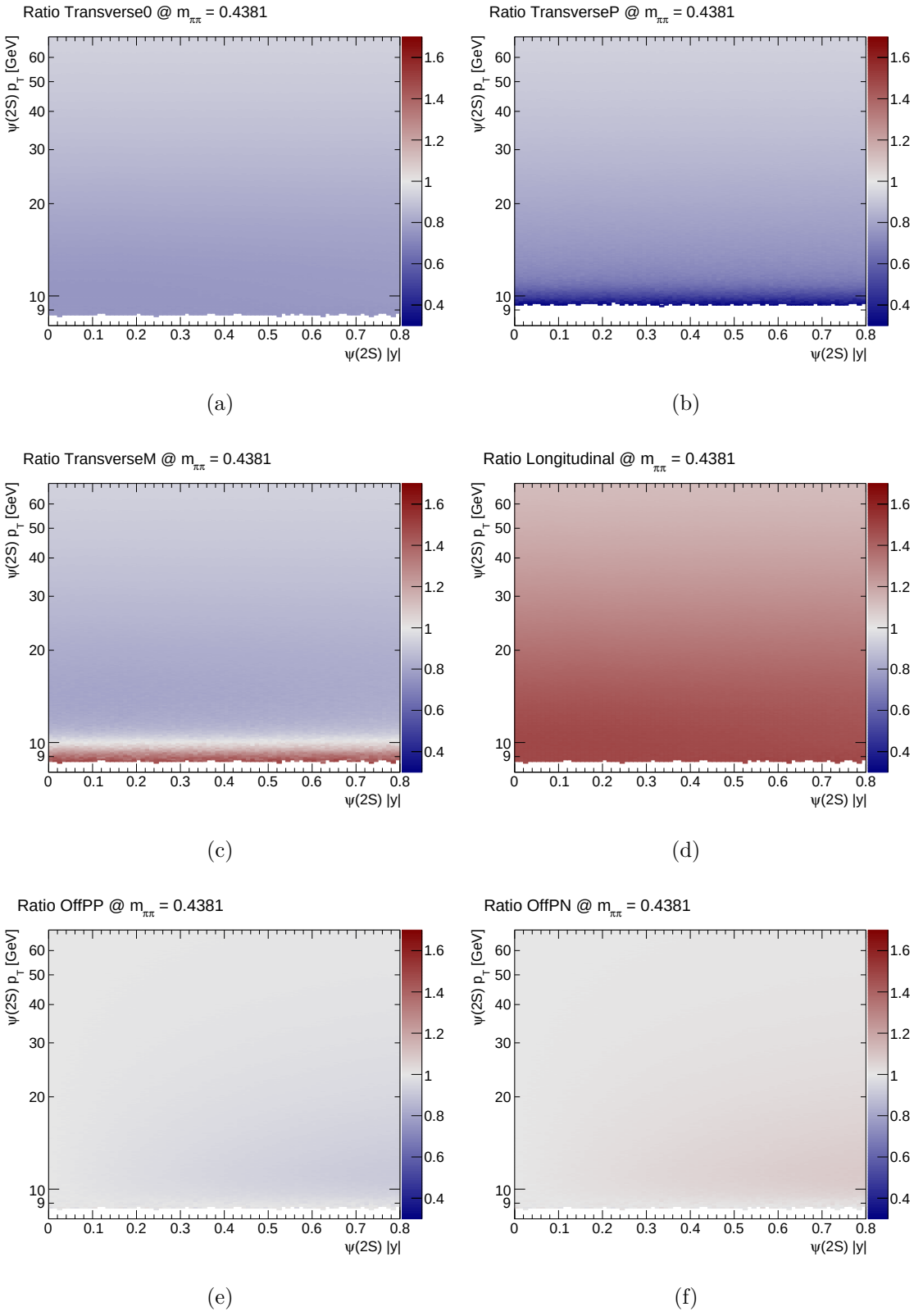


Fig. 5.5 Acceptance maps for $\psi(2S)$ as a ratio to the unpolarised acceptance for (a) T_{+0}, (b) T_{++}, (c) T_{+-}, (d) LONG, (e) OFFP+, (f) OFFP−.

5.3.6 Acceptance Maps for $X(3872)$

The angular distribution of $X(3872)$ in the decay $B^+ \to K^+ X(3872)_{\to J/\psi(\to\mu^+\mu^-)\rho^0(\to\pi^+\pi^-)}$ has been previously determined using the helicity formalism [92] and is expressed as the product of Wigner D-matrices. This distribution is dependent on the six parameters $\left(\theta_X, \theta_{J/\psi}, \theta_\rho, \phi_X, \phi_{J/\psi}, \phi_\rho\right)$, which are defined relative to the relevant helicity decay frames. A diagram detailing these parameters is shown in Figure 5.6.

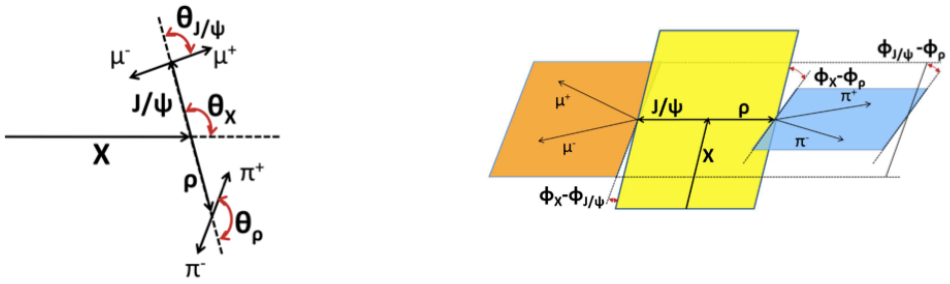


Fig. 5.6 Definitions of the helicity angles $\left(\theta_X, \theta_{J/\psi}, \theta_\rho\right)$ (left) and $\left(\phi_X, \phi_{J/\psi}, \phi_\rho\right)$ (right) [92].

The angles θ_X, ϕ_X are defined by the momentum of the J/ψ in the $X(3872)$ rest frame, with θ_X defined as the polar angle between the J/ψ in the $X(3872)$ rest frame and the $X(3872)$ in the lab frame, and ϕ_X being the angle of the plane created between the J/ψ in the $X(3872)$ frame and the $X(3872)$ in the lab frame, where $\phi_X = 0$ when aligned with the $X(3872)$ production plane. The angle $\theta_{J/\psi}$ is defined as the polar angle between the momentum of the positive muon in the J/ψ decay frame and the J/ψ momentum in the $X(3872)$ rest frame. $\phi_{J/\psi}$ is the angle of the plane formed between the positive muon in the J/ψ rest frame and the J/ψ in the $X(3872)$ rest frame. The angles (θ_ρ, ϕ_ρ) are calculated completely analogously to $\left(\theta_{J/\psi}, \phi_{J/\psi}\right)$, but instead taking π^+ instead of μ^+ and ρ^0 instead of J/ψ. We take the LHCb angular distribution result [92] and generalise to $X(3872)$ produced with any subset of angular momentum eigenstates as:

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}\Omega} \propto \sum_j \left| \sum_m \sum_{\lambda_{J/\psi}\lambda_\rho} A_m D^{1*}_{\lambda_{J/\psi} j}\left(\phi_{J/\psi}, \theta_{J/\psi}, -\phi_{J/\psi}\right) D^{1*}_{-\lambda_\rho 0}\left(\phi_\rho, \theta_\rho, -\phi_\rho\right) \times D^{J_X *}_{m(\lambda_{J/\psi}-\lambda_\rho)}\left(\phi_X, \theta_X, -\phi_X\right) A_{\lambda_{J/\psi}\lambda_\rho} \right|^2 \tag{5.8}$$

where j is the angular momentum projection of the dimuon system along its z' (the direction of travel of the positive muon in the J/ψ rest frame), and must be equal to ± 1 to good approximation to satisify helicity conservation from the decay $J/\psi \to \gamma^* \to \mu^+\mu^-$ (where this photon is approximately on-shell). $\lambda_{J/\psi}$ and λ_ρ are the helicities of the J/ψ and ρ^0, respectively. A_m is the amplitude of each $X(3872)$ angular momentum eigenstate $m = -1, 0, +1$. $A_{\lambda_{J/\psi}\lambda_\rho}$ is defined as the helicity amplitude for the pair of helicity states $\lambda_{J/\psi}\lambda_\rho$.

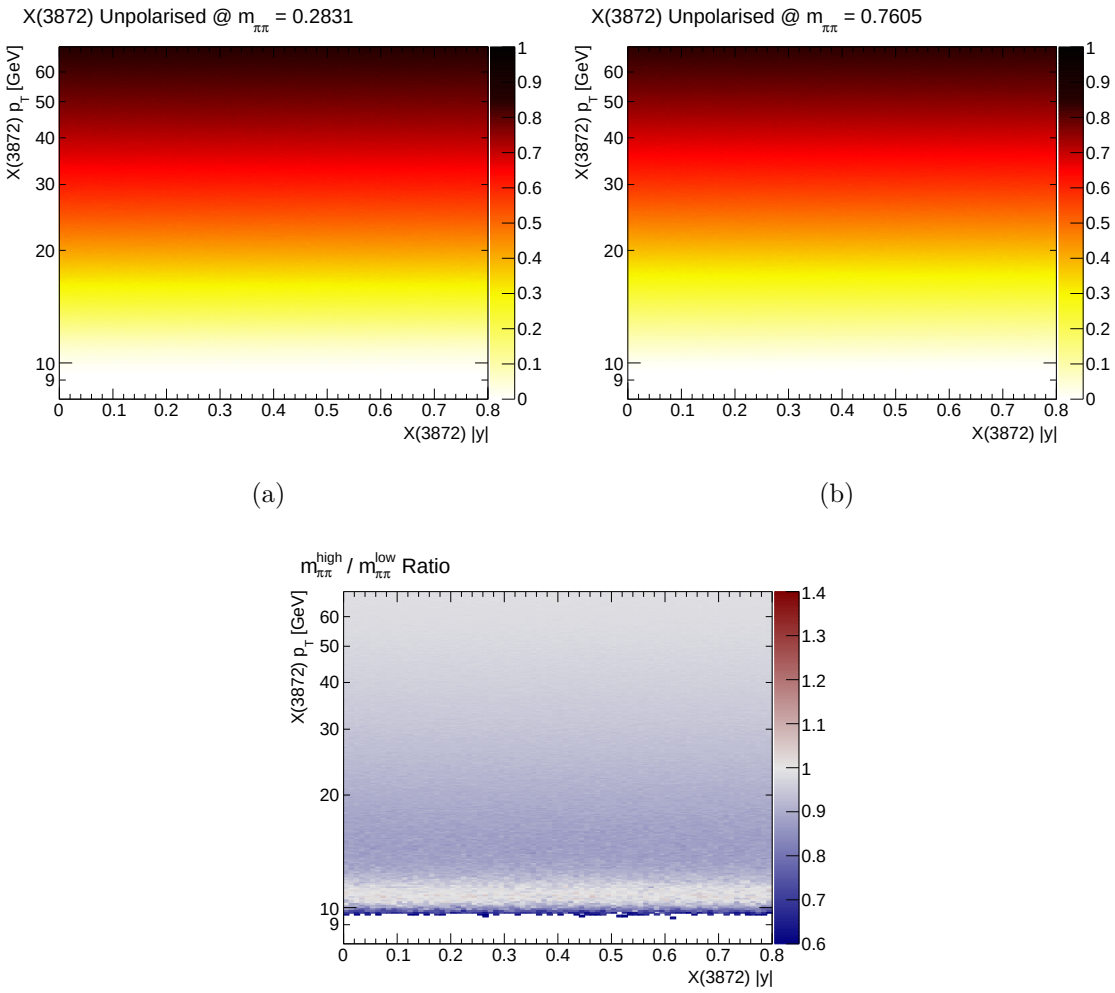
We use the same p_T and $|y|$ binning as with $\psi(2S)$, and additionally extend the dipion invariant mass range to $2m_\pi < m_{\pi^+\pi^-} < m_{X(3872)} - m_{J/\psi}$. We use the same polarisation hypotheses as for the $\psi(2S)$ case, but redefine the polarisation scenarios to be in terms of $X(3872)$ angular momentum eigenstate amplitudes:

- T_{+0}: $A_{m=+1} = +1, A_{m=0} = A_{m=-1} = 0$ OR $A_{m=+1} = A_{m=0} = 0, A_{m=-1} = +1$
- T_{++}: $A_{m=+1} = +1, A_{m=0} = 0, A_{m=-1} = +1$
- T_{+-}: $A_{m=+1} = -1, A_{m=0} = 0, A_{m=-1} = +1$
- LONG: $A_{m=+1} = 0, A_{m=0} = +1, A_{m=-1} = 0$
- OFFP+: $A_{m=+1} = -\sqrt{6}/3, A_{m=0} = +1/\sqrt{3}, A_{m=-1} = 0$
- OFFP−: $A_{m=+1} = +\sqrt{6}/3, A_{m=0} = +1/\sqrt{3}, A_{m=-1} = 0$

As for the $\psi(2S)$ case, there exists no individual sub-process which produces an unpolarised distribution of $X(3872)$. Instead, it should be an *incoherent* sum of the angular distributions for pure $m = -1$, $m = 0$ and $m = +1$ states to produce the unpolarised distribution. We see the acceptance map of $X(3872)$ decays in p_T and $|y|$ for low $m_{\pi\pi} = 0.2831$ GeV and high $m_{\pi\pi} = 0.7605$ GeV in Figure 5.7, along with the ratio of the two acceptance maps. Similarly to the $\psi(2S)$ decays, the acceptance is found to decrease as a function of $m_{\pi\pi}$.

The acceptance maps for other polarisation hypotheses are shown in Figures 5.8 as ratios with respect to the unpolarised acceptance. In contrast to the $\psi(2S)$, we measure for the LONG case that the acceptance decreases with respect to unpolarised decays. For the T_{+0} , T_{++} and T_{+-} scenarios, the acceptance increases with respect to unpolarised, and for OFFP+ and OFFP− the acceptance is very similar to the unpolarised case. The magnitude of the difference from unpolarised for $X(3872)$ decays is found to be smaller than for $\psi(2S)$ decays.

Additionally, we show the ratio of UNPOL acceptance for $X(3872)$ compared to $\psi(2S)$ for the lowest dipion mass ($m_{\pi\pi} \approx 2m_{\pi}$) and the largest possible mass $\left(m_{\pi\pi} \approx m_{\psi(2S)} - m_{J/\psi}\right)$ in Figure 5.9. These differences are discussed further in Section 5.3.7.



(a) (b)

(c)

Fig. 5.7 Acceptance maps for unpolarised $X(3872)$ decays for (a) low $m_{\pi\pi}$, (b) high $m_{\pi\pi}$, and (c) the ratio between the two.

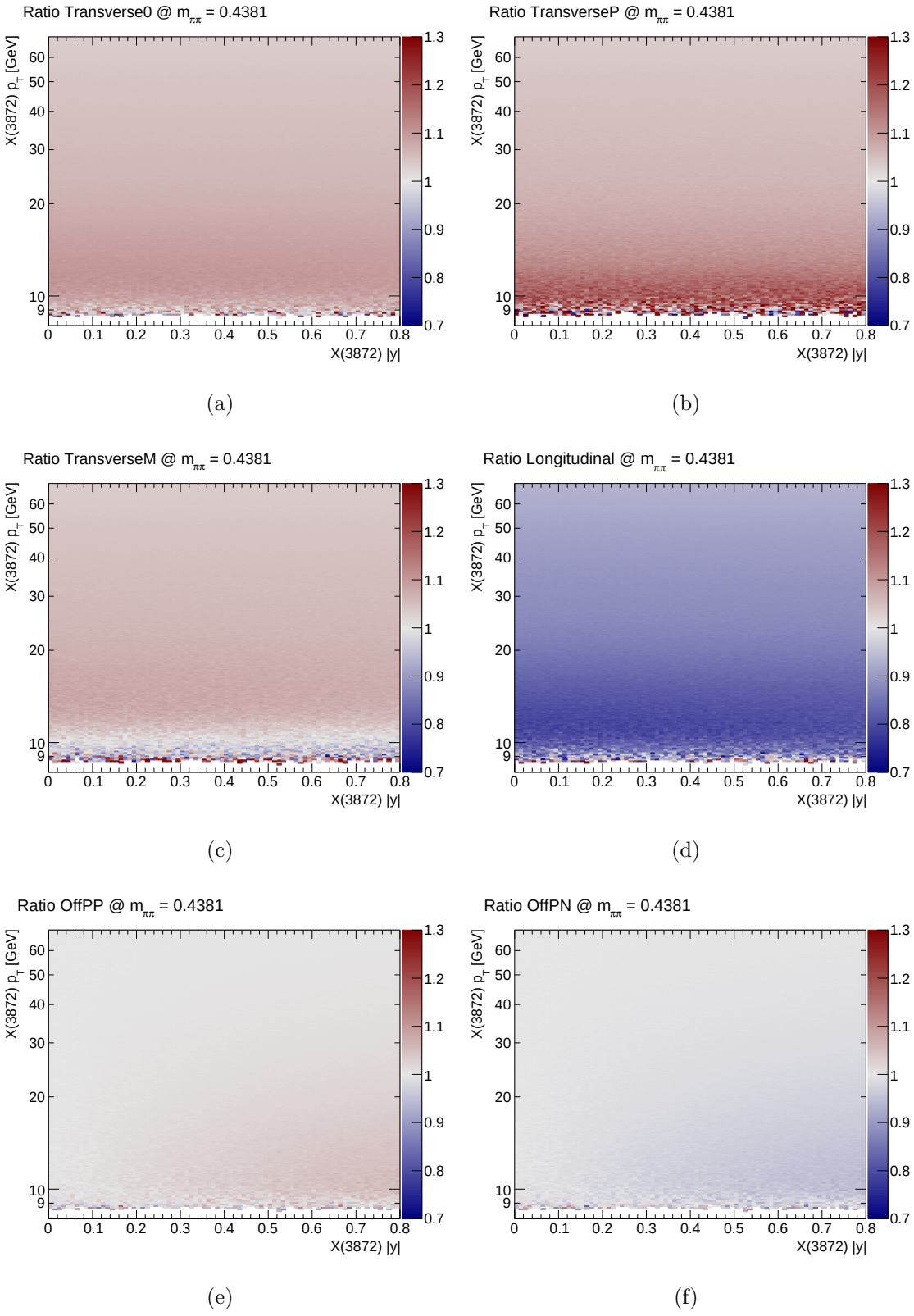
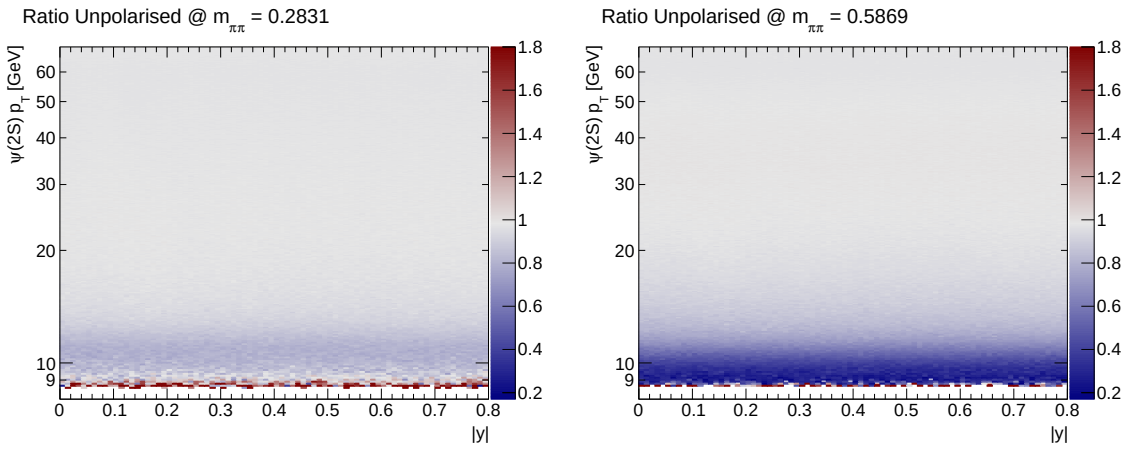


Fig. 5.8 Acceptance maps for $X(3872)$ as a ratio to the unpolarised acceptance for (a) T_{+0}, (b) T_{++}, (c) T_{+-}, (d) LONG, (e) OFFP+, (f) OFFP−.



(a) (b)

Fig. 5.9 Acceptance maps for unpolarised $X(3872)$ as a ratio of the unpolarised $\psi(2S)$ acceptance for (a) low ($m_{\pi\pi} \approx 2m_{\pi}$) and (b) high $\left(m_{\pi\pi} \approx m_{\psi(2S)} - m_{J/\psi}\right)$ dipion invariant mass.

5.3.7 J/ψ and Dipion Polarisation in $X(3872)$ and $\psi(2S)$ Decays

We see from Figures 5.5 and 5.8 that the acceptance for the different polarisation scenarios compared to the unpolarised case is opposite for $\psi(2S)$ and $X(3872)$ decays. For example, the transverse polarisation acceptance is larger than unpolarised for $X(3872)$ decays, but lower than unpolarised for $\psi(2S)$. For the $\psi(2S) \to J/\psi\pi^+\pi^-$ decays, the daughter J/ψ is expected to inherit the polarisation of the $\psi(2S)$. The dipion system is assumed to have spin-0, and thus remains unpolarised regardless of the polarisation of the mother $\psi(2S)$. However, in $X(3872) \to J/\psi\rho^0_{\to\pi^+\pi^-}$ decays the picture is more complicated because the dipion system is in a vector state. These effects are investigated in Appendix D.

5.3.8 Monte Carlo Generation for Selection Efficiency

We use Monte Carlo simulation to determine the efficiency of the background suppression selection criteria detailed in Table 5.1 in bins of analysis p_{T}. We generate $\psi(2S) \to J/\psi\pi^+\pi^-$ events separately for promptly and non-promptly produced $\psi(2S)$

using Pythia 8.1 [93]. The decays of b-hadrons are then simulated with EvtGen [94]. For the $X(3872)$ decay, we use $\chi_{c1}(1P) \to J/\psi\pi^+\pi^-$, as $X(3872)$ is unavailable in Pythia. The $\chi_{c1}(1P)$ is chosen as having the same quantum numbers ($I^G(J^{PC}) = 0^+(1^{++})$) as the $X(3872)$, and its mass is set to the $X(3872)$ PDG mass. The generated events are then passed through full simulation of the detector using the ATLAS simulation framework [95], based on Geant 4 [96, 97], and then processed with the same software as used for data.

Independent $\chi_{c1}(1P) \to J/\psi\rho^0$ MC is generated at truth level, and is used to weight the $X(3872)$ simulation across the dipion invariant mass range to account for the intermediate ρ^0 decay in $X(3872) \to J/\psi\pi^+\pi^-$, which has been measured as the dominant decay process [8], by finding the ratio of the fitted functions shown in Figures 5.10(a) and 5.10(b). The $m_{\pi\pi}$ distribution of the non-resonant $X(3872) \to J/\psi\pi^+\pi^-$ simulation in Figure 5.10(a) is fitted with a function describing the phase space (PS),

$$\mathrm{PS} = N\sqrt{\frac{(m_{\pi\pi}^2 - 4m_\pi^2)[m_{J/\psi}^4 + m_X^4 + m_{\pi\pi}^4 - 2(m_{J/\psi}^2 m_{\pi\pi}^2 + m_X^2 m_{\pi\pi}^2) + m_{J/\psi}^2 m_X^2]}{4m_X^2}}, \tag{5.9}$$

where N is the normalisation, $m_{\pi\pi}$ is the dipion invariant mass, and: m_π, $m_{J/\psi}$ and m_X are the world-average pion, J/ψ and $X(3872)$ masses, respectively. The $m_{\pi\pi}$ distribution of the $X(3872) \to J/\psi\rho^0$ simulation is fitted with the same phase space function combined with an additional term describing the ρ^0 mass resonance,

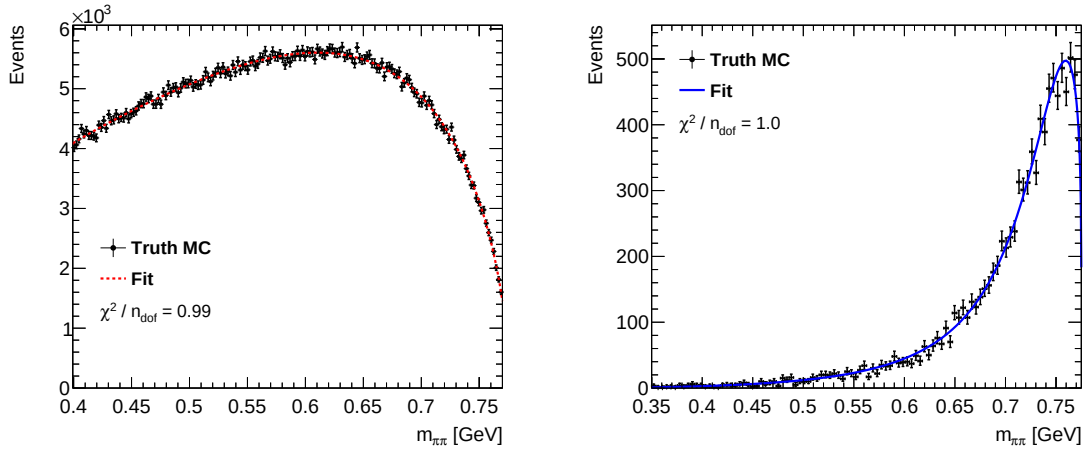
$$f(m_{\pi\pi}) = \mathrm{PS} \times \frac{p_1}{(m_{\pi\pi} - m_\rho)^2 + p_2}, \tag{5.10}$$

where PS is the phase space function described in Equation 5.9, $m_{\pi\pi}$ is the dipion invariant mass, and m_ρ is the world-average mass of the ρ^0 meson. The free parameters p_1 and p_2 are determined from the fit, which is shown in Figure 5.10(b).

Differences are also seen between the dipion mass distribution of the $\psi(2S)$ in simulation and data, with the dipion invariant mass in data following a Voloshin-Zakharov curve [98] (see Equation 7.2) to a good approximation, while following a phase-space distribution in simulation consistent with Equation 5.9. The dipion mass distributions are fitted for data and simulation in Figure 5.11, and the ratio of the fitted functions is used to weight the $\psi(2S)$ simulation at truth level. The method

used to extract the dipion invariant mass distribution for $\psi(2S)$ and $X(3872)$ signal is expanded upon in Chapter 7.

We observe the selection efficiency for prompt and non-prompt produced signal to agree within statistical error for both $\psi(2S)$ and $X(3872)$, and so the prompt and non-prompt samples were combined in order to improve the statistical precision of this study. Additionally, the simulated events are weighted at reconstruction level to correct for any remaining differences in $p_{\mathrm{T}}(J/\psi)$, $p_{\mathrm{T}}(\pi^{\pm})$ and $\Delta R(J/\psi, \pi^{\pm})$ between data and simulation, with the difference from the unweighted case used as a systematic error. More detail on the reweighting can be found in Appendix B.2.



(a) (b)

Fig. 5.10 Fits to the truth-level dipion invariant mass spectrum for (a) $X(3872) \to J/\psi\pi^+\pi^-$ and (b) $X(3872) \to J/\psi\rho^0(\to \pi^+\pi^-)$. The ratio between the two fitted functions, which are described in the text, is used as a per-event weight in the $X(3872) \to J/\psi\pi^+\pi^-$ simulation sample.

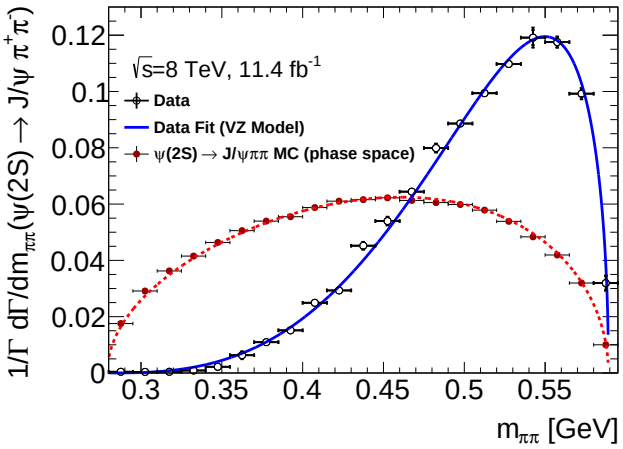


Fig. 5.11 Fit to the truth-level dipion invariant mass spectrum for $\psi(2S) \to J/\psi\pi^+\pi^-$ decays (red) using a function describing the phase-space, and to the data distribution (blue) using the Voloshin-Zakharov function [98] (see Equation 7.2). The ratio between the two fitted functions is used as a per-event weight in the $\psi(2S) \to J/\psi\pi^+\pi^-$ simulation sample.

For each of the simulation samples, we initally require:

- The dimuon trigger `EF_2mu4T_Jpsimumu_L2StarB` to have fired.
- The muons and pions belonging to each $J/\psi\pi^+\pi^-$ candidate to satisfy the acceptance criteria.
- The $J/\psi\pi^+\pi^-$ candidates to consist of two combined muons (details on muon reconstruction definitions can be found in Section 3.5.1).

because each of these is already corrected for in the data using per-event weights. We can then fit the signal mass peaks for each of the subsequent selection criteria, detailed in Table 5.1, to measure the selection efficiency in each analysis p_{T} bin.

The selection criteria were determined using data-driven studies on the $\psi(2S)$ and $X(3872)$ signal, across the full range of p_{T} (due to limited statistics), by performing mass fits simultaneously to the $\psi(2S)$ and $X(3872)$ signal whilst varying the strength of selection criteria. More detail of these studies can be found in Appendix B.1.

The systematic error is evaluated as the efficiency differences with and without these corrections. The selection efficiencies are shown in Table 5.3 for $\psi(2S)$ and Table 5.4 for $X(3872)$, with the corresponding mass fits shown in Appendix B.2.

	$\psi(2S)$ selection efficiency [%] across p_T range [GeV]				
Selection	$10-12$	$12-16$	$16-22$	$22-40$	$40-70$
reco quality	$94.0 \pm 1.1 \pm 0.8$	$95.7 \pm 0.7 \pm 0.2$	$96.2 \pm 0.7 \pm 0.2$	$95.7 \pm 0.6 \pm 0.1$	$97.8 \pm 3.0 \pm 1.5$
$P\left(\chi^2\right)(J/\psi\pi\pi)$	$91.4 \pm 1.1 \pm 0.8$	$92.9 \pm 0.7 \pm 0.1$	$94.0 \pm 0.7 \pm 0.3$	$93.4 \pm 0.7 \pm 0.1$	$94.9 \pm 3.0 \pm 1.1$
$\Delta R(J/\psi, \pi)$	$91.0 \pm 1.1 \pm 1.4$	$92.9 \pm 0.7 \pm 0.1$	$94.0 \pm 0.7 \pm 0.3$	$93.4 \pm 0.7 \pm 0.1$	$94.9 \pm 3.0 \pm 1.0$
$Q(J/\psi\pi\pi)$	$91.7 \pm 1.1 \pm 1.4$	$93.4 \pm 0.7 \pm 0.1$	$94.3 \pm 0.7 \pm 0.3$	$93.7 \pm 0.7 \pm 0.1$	$94.7 \pm 2.9 \pm 1.1$

Table 5.3 Cumulative efficiency of selection requirements in p_T analysis bins for $\psi(2S)$ simulation. The first column of errors are statistical. The second are systematic.

	$X(3872)$ selection efficiency [%] across p_T range [GeV]				
Selection	$10-12$	$12-16$	$16-22$	$22-40$	$40-70$
reco quality	$98.4 \pm 3.8 \pm 2.1$	$98.2 \pm 2.1 \pm 2.1$	$96.6 \pm 1.8 \pm 1.1$	$97.1 \pm 2.5 \pm 1.3$	$88.3 \pm 5.7 \pm 1.4$
$P\left(\chi^2\right)(J/\psi\pi\pi)$	$96.3 \pm 3.8 \pm 2.1$	$96.0 \pm 2.1 \pm 2.3$	$93.6 \pm 1.8 \pm 0.7$	$94.4 \pm 2.5 \pm 1.3$	$82.7 \pm 5.8 \pm 2.8$
$\Delta R(J/\psi, \pi)$	$86.5 \pm 3.5 \pm 0.4$	$92.8 \pm 2.1 \pm 1.3$	$93.1 \pm 1.8 \pm 0.4$	$94.4 \pm 2.5 \pm 1.3$	$82.7 \pm 5.8 \pm 2.8$
$Q(J/\psi\pi\pi)$	$85.6 \pm 3.6 \pm 0.0$	$91.8 \pm 2.1 \pm 1.6$	$92.7 \pm 1.9 \pm 0.6$	$94.2 \pm 2.6 \pm 1.5$	$84.3 \pm 5.5 \pm 2.3$

Table 5.4 Cumulative efficiency of selection requirements in p_T analysis bins for $X(3872)$ simulation. The first column of errors are statistical. The second are systematic.

The effect of each selection requirement on the dipion invariant mass distribution $m_{\pi\pi}$, separately for $\psi(2S)$ and $X(3872)$ decays, is shown in Appendix B.3 for a range of p_{T}. It is seen that certain selection criteria, particularly $Q(J/\psi\pi\pi)$ and $\Delta R(J/\psi, \pi)$, tend to bias the dipion invariant mass spectrum.

5.4 Signal Extraction

Our approach is to measure the yields of $\psi(2S)$ and $X(3872)$ by performing 1-dimensional binned χ^2 fits to the $J/\psi\pi^+\pi^-$ invariant mass distributions, in windows of pseudo-proper lifetime for each p_{T} bin. The extracted yields are then fitted with lifetime PDFs to separate prompt and non-prompt components. Modelling of the lifetime PDFs requires knowledge of the lifetime resolution function for each p_{T} bin.

5.4.1 Lifetime Resolution Determination

The lifetime resolution function $F_{\mathrm{res}}(\tau)$ is parameterised as a weighted sum of three normalised Gaussians with a common mean, with respective width parameters $\sigma_1 = \sigma_\tau$, $\sigma_2 = 2\sigma_\tau$ and $\sigma_3 = 4\sigma_\tau$. $F_{\mathrm{res}}(\tau)$ has a width dependence on the rapidity and p_{T} of the $J/\psi\pi^+\pi^-$ candidates, and it was found that this triple Gaussian appropriately

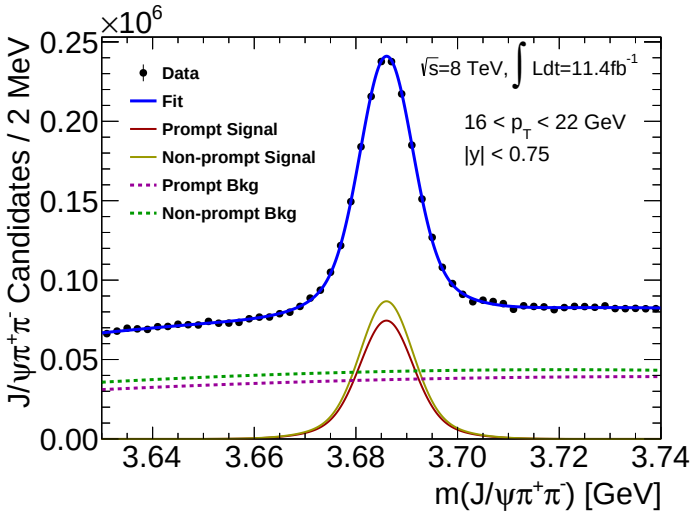
models the variation of $F_{\rm res}(\tau)$ in the rapidity and $p_{\rm T}$ bins used in this analysis. The resolution parameter σ_τ and the relative weights of the three Gaussians are determined separately for each analysis $p_{\rm T}$ bin, using 2-dimensional mass-lifetime unbinned maximum likelihood fits on the subset of data which contains a narrow range of masses around the $\psi(2S)$ peak. A full description of the PDF used to perform the fits, along with fit projections and results, can be found in Appendix E.1. The fit results are highlighted in Table 5.5. The fitted values for σ_τ are within the range of $(32-52)\times10^{-3}$ ps, with the weight of the narrowest Gaussian $(1-f_{23}^\tau)$ steadily increasing with $p_{\rm T}$ from 6% to about 50%. Also shown in the table is the fraction of the narrower Gaussian f_1 in the double Gaussian used to parameterise the $\psi(2S)$ mass peak, with the average used as a fixed parameter throughout the analysis.

	$p_{\rm T}$ interval [GeV]				
	$10-12$	$12-16$	$16-22$	$22-40$	$40-70$
	Lifetime Resolution Parameters				
$\tau_{\rm mean}$[ps] $\times 10^{-3}$	-1.7 ± 0.5	-0.6 ± 0.3	-0.5 ± 0.2	-0.3 ± 0.2	-1.0 ± 0.9
σ_τ[ps] $\times 10^{-3}$	44.3 ± 6.2	45.9 ± 3.3	52.0 ± 1.5	40.9 ± 1.5	32.2 ± 6.9
f_3^τ	0.34 ± 0.08	0.29 ± 0.03	0.19 ± 0.01	0.22 ± 0.01	0.20 ± 0.13
f_{23}^τ	0.94 ± 0.07	0.85 ± 0.05	0.59 ± 0.03	0.61 ± 0.04	0.49 ± 0.22
	Mass Peak Parameters				
$m_{\rm mean}$[GeV]	3.686 ± 0.000	3.686 ± 0.000	3.686 ± 0.000	3.686 ± 0.000	3.686 ± 0.000
m_σ[MeV]	5.4 ± 0.1	5.1 ± 0.1	5.0 ± 0.1	5.0 ± 0.1	5.4 ± 0.8
f_1	0.81 ± 0.03	0.73 ± 0.02	0.72 ± 0.03	0.74 ± 0.04	0.72 ± 0.24

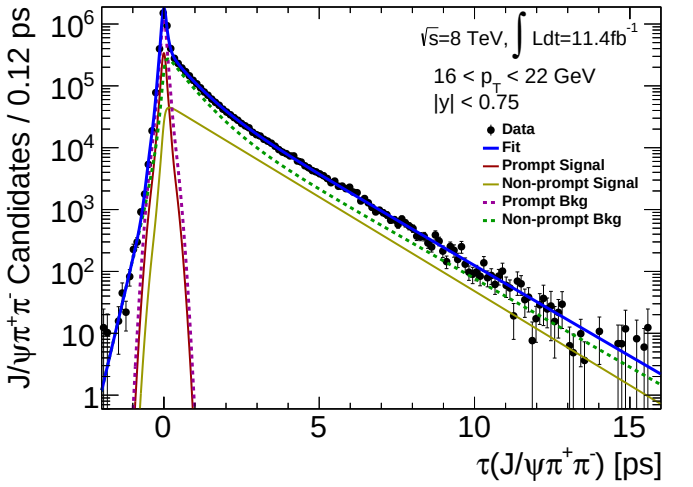
Table 5.5 Fit parameters determined from unbinned maximum likelihood fits of the $\psi(2S)$ signal region, to determine the mass and lifetime resolution functions in bins of $p_{\rm T}$.

The same lifetime resolution function can also be applied to the $X(3872)$ signal in each $p_{\rm T}$ bin - studies were performed on $\psi(2S)$ and $X(3872)$ simulation showing that the measured lifetime resolution using unbinned maximum likelihood fits are equal to a good approximation for $\psi(2S)$ and $X(3872)$ signal, seen in Appendix G.

An example of the 2D unbinned mass-lifetime maximum likelihood fit for the $p_{\rm T}$ bin $16-22$ GeV is shown in Figure 5.12, where the fit is projected into the mass and pseudo-proper lifetime of the $J/\psi\pi^+\pi^-$ candidates.



(a)



(b)

Fig. 5.12 (a) Mass projection over the full lifetime range and (b) lifetime projection over the full mass range, for the 2D unbinned maximum likelihood fit in the $\psi(2S)$ signal region in the p_T bin $[16, 22]$ GeV to establish lifetime resolution.

5.4.2 $X(3872)$ Mass Resolution

The ratio between $X(3872)$ and $\psi(2S)$ mass peak widths in fits to the invariant mass spectrum of reconstructed $J/\psi\pi^+\pi^-$ candidates, which we define as κ, is expected to be roughly constant in the p_{T} and rapidity range considered in this analysis. κ can therefore be fixed in simultaneous fits to the $\psi(2S)$ and $X(3872)$ mass peaks, so the $X(3872)$ mass peak width does not need to be determined by the fit. This is necessary for stable fits due to the small signal-to-background ratio and low statistics of the $X(3872)$ signal at high p_{T}. Due to constraining the invariant mass of the dimuon candidates to the J/ψ world-average mass in the constrained four track vertex fit, we can expect the mass signal resolution of the mother $\psi(2S)$ or $X(3872)$ decaying into $J/\psi\pi^+\pi^-$ to increase roughly linearly with distance from threshold $m - 2m_\pi - m^{PDG}_{J/\psi}$. We would therefore expect the width of the $X(3872)$ mass peak to be wider than the $\psi(2S)$ peak by a factor

$$\kappa \simeq \frac{m_{X(3872)} - 2m_\pi - m^{PDG}_{J/\psi}}{m_{\psi(2S)} - 2m_\pi - m^{PDG}_{J/\psi}} = 1.6. \tag{5.11}$$

We perform fits to the invariant mass distribution of the $J/\psi\pi^+\pi^-$ candidates to simultaneously measure $\psi(2S)$ and $X(3872)$ signal parameters, and so determine κ. Double Gaussians are used to fit each signal peak, where the width of the wider Gaussian is fixed as twice the width of the narrower. Due to the wide variation of weights in the lower p_{T} bins and the lower signal-to-background ratio, the fit is quite unstable when we attempt to fit the weighted $J/\psi\pi^+\pi^-$ invariant mass distribution in the full p_{T} range. Instead, we fit the range $16 < p_{\mathrm{T}} < 70$ GeV, which offers a better signal-to-background ratio and a more stable fit. The mass resolution was observed in simulation to be largely independent of p_{T}, justifying the decision to remove lower p_{T} candidates from this study. We see this fit in Figure 5.13.

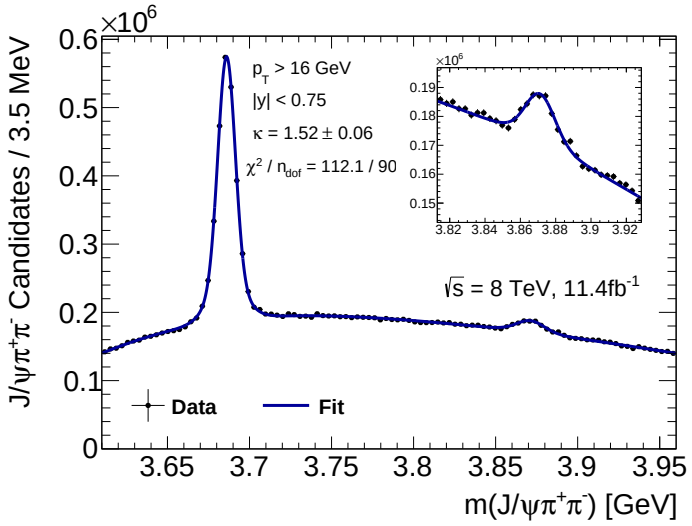


Fig. 5.13 Fit to the weighted $J/\psi\pi^+\pi^-$ invariant mass distribution in the range 16 GeV $< p_{\mathrm{T}} <$ 70 GeV.

This fit gives a good $\chi^2/n_{dof} = 112.1/90$ and $\kappa = 1.52 \pm 0.06$. As discussed earlier in this section, for simulation the value of κ does not vary much as a function of p_{T}, and is certainly within the $\pm$ 0.06 error we use as a systematic variation. It was decided to use this value of $\kappa = 1.52 \pm 0.06$ as central for the main analysis.

5.4.3 Invariant Mass Fits in Pseudo-Proper Lifetime Windows

Our approach is to measure the yields of $\psi(2S)$ and $X(3872)$ by performing 1-dimensional binned χ^2 fits in $J/\psi\pi^+\pi^-$ invariant mass. We perform these fits in four pseudo-proper lifetime windows $w_0 - w_3$:

w_0 : $-0.3\,\mathrm{ps} < \tau(J/\psi\pi\pi) < 0.025\,\mathrm{ps}$

w_1 : $0.025\,\mathrm{ps} < \tau(J/\psi\pi\pi) < 0.3\,\mathrm{ps}$

w_2 : $0.3\,\mathrm{ps} < \tau(J/\psi\pi\pi) < 1.5\,\mathrm{ps}$

w_3 : $1.5\,\mathrm{ps} < \tau(J/\psi\pi\pi) < 15.0\,\mathrm{ps}$

The fitting function used to determine the yields consists of two double Gaussian peaks for the two signals and a smooth background:

$$f(m) = Y^{\psi}\left(f_1 G_1^{\psi}(m) + (1-f_1)\, G_2^{\psi}(m)\right) + Y^{X}\left(f_1 G_1^{X}(m) + (1-f_1)\, G_2^{X}(m)\right) + N_{\mathrm{bkg}}(m-m_{\mathrm{th}})^{p_1} e^{p_2(m-m_{\mathrm{th}})} P(m-m_{\mathrm{th}}), \tag{5.12}$$

where the threshold mass $m_{\mathrm{th}} = m_{J/\psi} + 2m_{\pi} = 3376.06$ MeV. The $\psi(2S)$ and $X(3872)$ signal yields Y^{ψ} and Y^{X}, coefficients of the second-order polynomial P, and parameters N_{bkg}, p_1 and p_2 are determined from the fits. Signal peaks for $\psi(2S)$ and $X(3872)$ are described by normalised double Gaussians with common means: $G_1^{\psi}(m)$ and $G_1^{X}(m)$ are the narrower Gaussians with respective widths σ_{ψ} and σ_X, while $G_2^{\psi}(m)$ and $G_2^{X}(m)$ are wider Gaussians with widths $2\sigma_{\psi}$ and $2\sigma_X$. The fraction of the narrower Gaussian f_1 is fixed to be the same for $\psi(2S)$ and $X(3872)$, while the widths σ_{ψ} and σ_X are related by $\sigma_X = \kappa\sigma_{\psi}$. The parameters f_1 and κ are fixed for the main fits to the values $f_1 = 0.76 \pm 0.04$, $\kappa = 1.52 \pm 0.06$ (see Sections 5.4.1 and 5.4.2 for their determination), and are varied within these uncertainties in the systematic studies. Differences in κ as a function of p_{T}, and differences in f_1 between $\psi(2S)$ and $X(3872)$ signal are found to be well within these systematic variations using simulation. The fits and their corresponding pull distributions are shown in Appendix F. The fit quality is found to be good throughout the range of transverse momenta and lifetimes.

The yields extracted from the fits are shown in Table 5.6 for the $\psi(2S)$ and Table 5.7 for the $X(3872)$, and are subsequently fitted with PDFs describing the lifetime distributions, to separate prompt and non-prompt signal components.

	ψ(2S) Yields [$\times 10^5$] vs. p_{T} [GeV]				
τ window	$10-12$	$12-16$	$16-22$	$22-40$	$40-70$
w_0	17.48 ± 0.36	11.03 ± 0.11	3.53 ± 0.03	1.14 ± 0.01	0.078 ± 0.004
w_1	14.07 ± 0.37	9.04 ± 0.10	2.94 ± 0.03	1.01 ± 0.01	0.071 ± 0.003
w_2	9.13 ± 0.29	7.04 ± 0.09	2.97 ± 0.03	1.27 ± 0.01	0.104 ± 0.004
w_3	6.74 ± 0.16	5.21 ± 0.06	2.22 ± 0.02	0.94 ± 0.01	0.081 ± 0.003

Table 5.6 Invariant mass fit results in pseudoproper lifetime and p_{T} bins for the ψ(2S). Uncertainties are statistical only.

	$X(3872)$ Yields [$\times 10^4$] vs. p_{T} [GeV]				
τ window	$10-12$	$12-16$	$16-22$	$22-40$	$40-70$
w_0	10.77 ± 2.32	10.55 ± 0.76	3.53 ± 0.26	1.19 ± 0.11	0.09 ± 0.03
w_1	9.27 ± 2.70	8.21 ± 0.71	2.60 ± 0.24	0.72 ± 0.11	0.04 ± 0.02
w_2	4.03 ± 1.74	3.83 ± 0.63	1.29 ± 0.21	0.45 ± 0.10	0.04 ± 0.02
w_3	2.06 ± 0.81	2.09 ± 0.34	0.98 ± 0.13	0.30 ± 0.06	0.02 ± 0.01

Table 5.7 Invariant mass fit results in pseudoproper lifetime and p_{T} bins for the $X(3872)$. Uncertainties are statistical only.

5.4.4 Lifetime Fitting

The probability density function (PDF) describing the dependence of $\psi(2S)$ and $X(3872)$ signal yields on the pseudo-proper lifetime τ is a superposition of prompt (P) and non-prompt (NP) components:

$$F^i(\tau) = (1 - f^i_{NP})F^i_P(\tau) + f^i_{NP}F^i_{NP}(\tau), \tag{5.13}$$

where f_{NP} is the non-prompt fraction, while i stands for either $\psi(2S)$ or $X(3872)$. The prompt components of $\psi(2S)$ and $X(3872)$ production should not have any observable decay length, hence $F_P(\tau)$ is effectively described by the lifetime resolution function $F_{\mathrm{res}}(\tau)$, assumed to be the same for $\psi(2S)$ and $X(3872)$ signals. This was verified using simulated data samples, seen in Appendix G. The lifetime resolution function was determined for each p_{T} bin in Section 5.4.1.

Single-Lifetime Fits

The simplest description of the non-prompt components of the signal PDF is given by a single one-sided exponential smeared with the resolution function, with the effective lifetime τ_{eff} of the exponential determined from the fit. This model, referred to as a 'single-lifetime fit', is applied to the $\psi(2S)$ and $X(3872)$ yields from Tables 5.6 and 5.7, and the results of the corresponding binned minimum-χ^2 fits are shown in Tables 5.8 and 5.9.

	p_{T} interval [GeV]				
	$10-12$	$12-16$	$16-22$	$22-40$	$40-70$
	$\psi(2S)$ Yields [$\times 10^5$] vs. p_{T}				
Total Yield	47.82 ± 0.62	32.55 ± 0.18	11.72 ± 0.06	4.37 ± 0.03	0.33 ± 0.01
Prompt Yield	28.76 ± 0.54	17.73 ± 0.15	5.39 ± 0.05	1.64 ± 0.02	0.10 ± 0.01
Non-Prompt Yield	19.06 ± 0.46	14.82 ± 0.14	6.33 ± 0.05	2.73 ± 0.02	0.23 ± 0.01
	Fractions vs. p_{T}				
f_{NP}	0.40 ± 0.01	0.46 ± 0.00	0.54 ± 0.00	0.62 ± 0.00	0.69 ± 0.01
	Fit quality in each p_{T} bin				
χ^2/n_{dof}	$0.62/1$	$1.21/1$	$0.10/1$	$0.28/1$	$1.60/1$

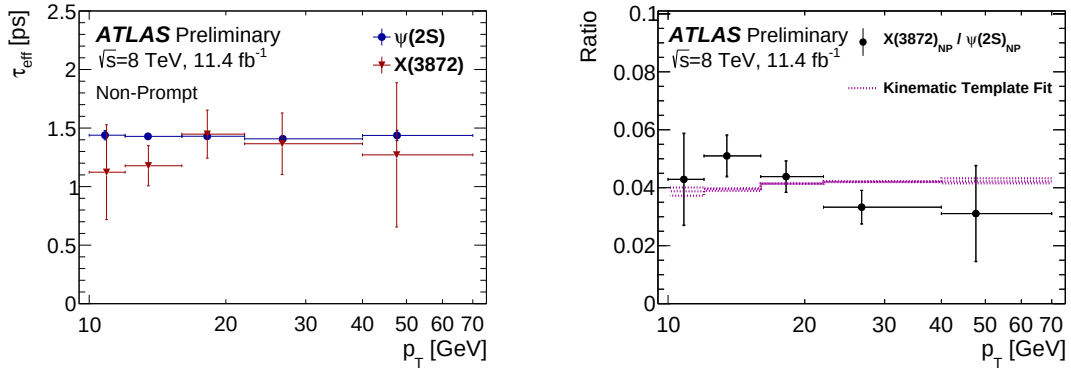
Table 5.8 Fit results in p_{T} bins for the $\psi(2S)$ for the single-lifetime fit. Uncertainties are statistical only.

	p_{T} interval [GeV]				
	$10-12$	$12-16$	$16-22$	$22-40$	$40-70$
	$X(3872)$ Yields [$\times 10^4$] vs. p_{T}				
Total Yield	26.11 ± 3.99	24.92 ± 1.27	8.44 ± 0.43	2.68 ± 0.20	0.19 ± 0.05
Prompt Yield	18.36 ± 3.57	17.50 ± 1.11	5.69 ± 0.37	1.77 ± 0.16	0.12 ± 0.04
Non-Prompt Yield	7.74 ± 2.83	7.42 ± 1.01	2.75 ± 0.33	0.92 ± 0.16	0.07 ± 0.03
	Fractions vs. p_{T}				
f_{NP}	0.30 ± 0.09	0.30 ± 0.03	0.33 ± 0.03	0.34 ± 0.05	0.35 ± 0.16
	Fit quality in each p_{T} bin				
χ^2/n_{dof}	$0.13/1$	$0.21/1$	$0.23/1$	$0.23/1$	$0.18/1$

Table 5.9 Fit results in p_{T} bins for the $X(3872)$ for the single-lifetime fit. Uncertainties are statistical only.

p_T bin [GeV]	$\tau_{\text{eff}}(\psi(2S))$ [ps]	$\tau_{\text{eff}}(X(3872))$ [ps]
$10-12$	1.44 ± 0.04	1.12 ± 0.40
$12-16$	1.43 ± 0.02	1.18 ± 0.17
$16-22$	1.43 ± 0.01	1.45 ± 0.21
$22-40$	1.41 ± 0.01	1.37 ± 0.26
$40-70$	1.44 ± 0.04	1.27 ± 0.62

Table 5.10 Effective pseudo-proper lifetimes for $\psi(2S)$ and $X(3872)$ obtained with the single-lifetime fit model.



(a) (b)

Fig. 5.14 (a) Measured effective pseudo-proper lifetimes for non-prompt $X(3872)$ and $\psi(2S)$. (b) The ratio of non-prompt $X(3872)$ and $\psi(2S)$ production. The measured distribution is fitted to the kinematic template described in the text. The fit quality is acceptable, with $\chi^2/\text{dof} = 5.4/4$.

Figure 5.14(a) and Table 5.10 show the effective pseudo-proper lifetimes for non-prompt $\psi(2S)$ and $X(3872)$ signals in bins of p_{T}. While for $\psi(2S)$ the fitted values of τ_{eff} are measured to be around 1.45 ps in all p_{T} bins, the signal from $X(3872)$ at low p_{T} tends to have shorter lifetimes, possibly hinting on a different production mechanism at low p_{T}.

In Figure 5.14(b) the ratio of non-prompt yields of $X(3872)$ and $\psi(2S)$ is plotted as a function of transverse momentum. The measured distribution is compared to the kinematic template, which was calculated as a ratio of the simulated p_{T} distributions

of non-prompt $X(3872)$ and non-prompt $\psi(2S)$, assuming that the same mix of the parent b-hadrons contributes for both signals. The shape of the template reflects the kinematics of the decay of a b-hadron into $\psi(2S)$ or $X(3872)$, with the width of the band showing the range of variation for extreme values of the invariant mass of the recoiling hadronic system. The creation of this template is detailed in Appendix H. A fit of the measured ratio to this template allows us to determine the ratio of the average branching fractions

$$\begin{aligned} R_B^{1L} &= \frac{Br(B \to X(3872) + \text{any})Br(X(3872) \to J/\psi\pi^+\pi^-)}{Br(B \to \psi(2S) + \text{any})Br(\psi(2S) \to J/\psi\pi^+\pi^-)} \\ &= (3.95 \pm 0.32(\text{stat}) \pm 0.08(\text{sys}))\%. \end{aligned} \tag{5.14}$$

However, the observed variation of this ratio with p_{T} shows a somewhat falling trend (Figure 5.14(b)) which does not completely agree with the shape of the template, possibly suggesting the presence of an additional contribution to the non-prompt X(3872) yield in the low-p_{T} bins, which is unforeseen within the single-lifetime model.

Two-Lifetime Fits

An alternative fit model is also implemented in this analysis, which allows for two non-prompt contributions with distinctly different effective lifetimes (the 'two-lifetime fit'). The statistical power of the data sample is insufficient for determining two free lifetimes, especially in the case of $X(3872)$ production, so in this fit model the non-prompt PDF is represented in each p_{T} bin by a sum of two contributions with different lifetimes, and a relative weight determined by the fit:

$$F_{NP}^i(\tau) = (1 - f_{SL}^i)F_{LL}(\tau) + f_{SL}^i F_{SL}(\tau), \tag{5.15}$$

where the indices SL and LL refer to short-lived and long-lived non-prompt components, respectively, and f_{SL}^i are the short-lived non-prompt fractions, for $i = \psi(2S), X(3872)$. The PDFs $F_{SL}(\tau)$ and $F_{LL}(\tau)$ are parameterised as single-sided exponentials with fixed lifetimes, smeared with the lifetime resolution function $F_{\mathrm{res}}(\tau)$ described above. According to this fit model, the long-lived parts of the non-prompt contributions originate from the usual mix of $B^{\pm}, B^0, B_s$ mesons and baryons, while the short-lived part is assumed to be due to the contribution of $B_c^{\pm}$ mesons.

The B_c cross section is predicted to have a steeper dependence in p_{T} than other b-hadrons [99], for instance the B^+, at low and intermediate transverse momenta such as those seen at the LHC. This is primarily due to a dominant non-fragmentation mechanism, where the $\bar{c}$ quark is produced in the interaction of the b quark with the rest of the hadronic system. This production mechanism is expected to dominate at low p_{T}, but is suppressed at high p_{T} by an extra factor of $1/p_{\mathrm{T}}^2$ [99]. Signficant production of $X(3872)$ from B_c decays may also explain the measured effective pseudoproper lifetimes of $X(3872)$ at low p_{T} for the single-lifetime fits being smaller than that for the $\psi(2S)$ (see Table 5.10), because the lifetime of the B_c, $\tau(B_c) = (0.507 \pm 0.009)$ ps [18], is much smaller than the lifetimes of other b-hadrons.

The effective pseudo-proper lifetime of the short-lived component τ_{SL} is determined from MC simulation studies using B_c decays. It is found to be $\tau_{SL} = 0.40 \pm 0.05$ ps, somewhat shorter than the world average lifetime $\tau(B_c) = 0.507$ ps of the parent B_c [18]. This takes into account the non-negligible energy release in the B_c decay, with the uncertainty in τ_{SL} reflecting its probable range of variation. The effective pseudo-proper lifetime of the long-lived component τ_{LL} is determined from the two-lifetime test fits to the $\psi(2S)$ mass range, with τ_{LL} free and allowing for an unknown contribution of a short-lived component with lifetime τ_{SL}. Across the p_{T} bins, τ_{LL} is found to be within the range 1.45 ± 0.05 ps. The effective pseudo-proper lifetimes τ_{LL} and τ_{SL} are fixed to the above values for the main fits, and are varied within the quoted uncertainties during systematic studies. More detail on the methods used to estimate the pseudo-proper lifetimes of the short-lived and long-lived components can be found in Appendices E.2 and E.3, respectively.

The results of the two-lifetime fits are presented in Tables 5.11 and 5.12 for the $\psi(2S)$ and $X(3872)$, respectively, where the errors are purely statistical and determined by the fits. The final results of the two-lifetime fits are presented in Section 6.3 after the discussion of systematic uncertainties in Section 6.1. The two-lifetime fits are used for $\psi(2S)$ and $X(3872)$ to obtain all subsequent results, unless specified otherwise, with the relatively small differences between the results of the single-lifetime and two-lifetime fits being highlighted alongside all other sources of systematic uncertainty.

	p_{T} interval [GeV]				
	$10-12$	$12-16$	$16-22$	$22-40$	$40-70$
	$\psi(2\mathrm{S})$ Yields [$\times 10^5$] vs. p_{T}				
Total Yield	47.85 ± 0.62	32.55 ± 0.18	11.72 ± 0.06	4.37 ± 0.03	0.33 ± 0.01
Prompt Yield	28.66 ± 0.60	17.64 ± 0.17	5.37 ± 0.06	1.63 ± 0.02	0.10 ± 0.01
Non-Prompt Yield	19.19 ± 0.58	14.92 ± 0.17	6.35 ± 0.06	2.74 ± 0.03	0.23 ± 0.01
Short-Lived NP Yield	0.35 ± 0.75	0.36 ± 0.23	0.10 ± 0.08	0.10 ± 0.04	0.01 ± 0.01
Long-Lived NP Yield	18.85 ± 0.48	14.56 ± 0.17	6.25 ± 0.06	2.64 ± 0.03	0.22 ± 0.01
	Fractions vs. p_{T}				
f_{NP}	0.40 ± 0.01	0.46 ± 0.00	0.54 ± 0.00	0.63 ± 0.00	0.69 ± 0.01
f_{SL}	0.02 ± 0.04	0.02 ± 0.02	0.02 ± 0.01	0.04 ± 0.01	0.03 ± 0.03
	Fit quality in each p_{T} bin				
χ^2/n_{dof}	0.48/1	0.45/1	0.67/1	0.39/1	1.00/1

Table 5.11 Fit results in p_{T} bins for the $\psi(2\mathrm{S})$ for the two-lifetime fit. Uncertainties are statistical only.

	p_{T} interval [GeV]				
	$10-12$	$12-16$	$16-22$	$22-40$	$40-70$
	$X(3872)$ Yields [$\times 10^4$] vs. p_{T}				
Total Yield	26.27 ± 4.03	24.92 ± 1.28	8.45 ± 0.44	2.68 ± 0.20	0.19 ± 0.05
Prompt Yield	17.76 ± 3.81	17.02 ± 1.23	5.66 ± 0.41	1.77 ± 0.17	0.12 ± 0.04
Non-Prompt Yield	8.52 ± 3.52	7.90 ± 1.22	2.79 ± 0.40	0.91 ± 0.19	0.06 ± 0.04
Short-Lived NP Yield	3.00 ± 4.37	2.19 ± 1.56	0.08 ± 0.53	0.03 ± 0.24	0.00 ± 0.05
Long-Lived NP Yield	5.50 ± 2.42	5.71 ± 1.00	2.71 ± 0.37	0.88 ± 0.17	0.06 ± 0.04
	Fractions vs. p_{T}				
f_{NP}	0.32 ± 0.12	0.32 ± 0.04	0.33 ± 0.04	0.34 ± 0.06	0.34 ± 0.18
f_{SL}	0.35 ± 0.39	0.28 ± 0.16	0.03 ± 0.19	0.03 ± 0.26	0.03 ± 0.63
	Fit quality in each p_{T} bin				
χ^2/n_{dof}	0.05/1	0.01/1	0.20/1	0.31/1	0.23/1

Table 5.12 Fit results in p_{T} bins for the $X(3872)$ for the two-lifetime fit. Uncertainties are statistical only.

Chapter 6

Systematics and Results

6.1 Systematics

6.1.1 Muon Reconstruction and Trigger Efficiency

The procedure used to determine systematic errors for the trigger and muon reconstruction efficiencies closely follows the procedures adopted in [6]. Each bin of the muon reconstruction and trigger efficiency maps is randomly fluctuated according to a normal distribution, with the statistical error of each bin used as the width. This process is repeated 1000 times, to create 1000 trigger and muon reconstruction efficiency map variations. The average weight from each of the map variations is found in the $\psi(2S)$ and $X(3872)$ signal regions, for prompt $|\tau\left(J/\psi\pi^+\pi^-\right)/\Delta\tau\left(J/\psi\pi^+\pi^-\right)| < 2$ and non-prompt $|\tau\left(J/\psi\pi^+\pi^-\right)/\Delta\tau\left(J/\psi\pi^+\pi^-\right)| > 2.5$ signal. The resulting distribution of average weights is fitted with a Gaussian, the width of which is used as the systematic error. Examples of the fitted distributions are shown in Appendix I.1 for both the muon and trigger reconstruction efficiencies. We measure the systematic error of the trigger efficiency as $1-3\%$ and the muon reconstruction efficiency as less than 1% for both the $\psi(2S)$ and $X(3872)$ signal regions.

The efficiency of the muon tracking in the ID is assigned an error of $\pm 1\%$ [88], primarily due to a dependence on the total number of pp collisions per event. This error is used directly as a systematic error for each p_T bin.

6.1.2 Pion Reconstruction Efficiency

A number of effects were considered as sources of systematic error in pion reconstruction efficiency corrections. These include: corrections due to differences in pile-up and vertex z displacement between data and MC, material uncertainties in the inner detector, uncertainty in truth matching, and statistical errors in the Monte Carlo used to create the pion efficiency maps. The overall systematic uncertainty of these effects was found to be a flat 2.5% for the dipion system. Details of the uncertainty determination can be found in Appendix C.

6.1.3 Selection Cuts

Details of the selection criteria are shown in Table 5.1. The efficiencies of the cuts were studied with Monte Carlo simulation. The simulated events were weighted at the reconstruction level to match the data $p_{\mathrm{T}}(J/\psi)$, $p_{\mathrm{T}}(\pi^{\pm})$ and $\Delta R(J/\psi, \pi^{\pm})$ signal distributions. The efficiency of the cuts, along with their statistical and systematic errors, can be seen in Table 5.3 for $\psi(2S)$ signal and Table 5.4 for $X(3872)$. The statistical and systematic errors of the efficiency when all the cuts are applied are summed in quadrature and used directly as a systematic in each p_{T} bin.

6.1.4 Lifetime Resolution

We vary the lifetime resolution PDF by the uncertainty of its width, and separately the uncertainty of its mean, in each p_{T} bin (shown in Table 5.5), and repeat the lifetime fits as a systematic variation. Additionally, we re-measure the lifetime resolution using unbinned fits with a small short-lived (0.4 ps) non-prompt signal component, fixed to a fraction of 15% of the non-prompt signal, and re-perform the lifetime fits.

6.1.5 Non-prompt Signal Lifetimes

We vary the slope of the non-prompt signal lifetime PDFs by their uncertainties, individually for short-lived and long-lived signal components. The maximum variations from the central model for each variation are used as systematic errors.

6.1.6 Mass Fit Model Systematics

Several different fit models are considered for systematic study, each varying by one component from the central fit model by an amount considered reasonable from the errors of unweighted mass-lifetime fits and binned χ^2 fits to the total data sample. The weighted events are refitted for each model, and the maximum difference in yields returned by the fits from the central fit for each p_T bin are used as a systematic error. A summary of each of the systematic fit models is shown in Table 6.1.

Variation	
	Signal variation
1	(fit $\sigma_1^{\psi(2S)}, \sigma_2^{\psi(2S)} = 2.0 \times \sigma_1^{\psi(2S)}$) $\rightarrow$ (fit $\sigma_1^{\psi(2S)}, \sigma_2^{\psi(2S)} = \mathbf{1.8} \times \sigma_1^{\psi(2S)}$)
2	(fit $\sigma_1^{\psi(2S)}, \sigma_2^{\psi(2S)} = 2.0 \times \sigma_1^{\psi(2S)}$) $\rightarrow$ (fit $\sigma_1^{\psi(2S)}, \sigma_2^{\psi(2S)} = \mathbf{2.2} \times \sigma_1^{\psi(2S)}$)
3	(fit $\sigma_1^{\psi(2S)}, \sigma_2^{\psi(2S)} = 2.0 \times \sigma_1^{\psi(2S)}$) $\rightarrow$ (fit $\sigma_1^{\psi(2S)}, \sigma_2^{\psi(2S)} = \mathbf{free} \times \sigma_1^{\psi(2S)}$)
4	narrow Gauss fraction $f_1 = 0.76 \rightarrow$ narrow Gauss fraction $f_1 = \mathbf{0.72}$
5	narrow Gauss fraction $f_1 = 0.76 \rightarrow$ narrow Gauss fraction $f_1 = \mathbf{0.80}$
6	narrow Gauss fraction $f_1 = 0.76 \rightarrow$ narrow Gauss fraction $f_1 = \mathbf{free}$
7	(fit $\sigma_1^{\psi(2S)}, \sigma^{X(3872)} = 1.52 \times \sigma_1^{\psi(2S)}$) $\rightarrow$ (fit $\sigma_1^{\psi(2S)}, \sigma^{X(3872)} = \mathbf{1.47} \times \sigma_1^{\psi(2S)}$)
8	(fit $\sigma_1^{\psi(2S)}, \sigma^{X(3872)} = 1.52 \times \sigma_1^{\psi(2S)}$) $\rightarrow$ (fit $\sigma_1^{\psi(2S)}, \sigma^{X(3872)} = \mathbf{1.58} \times \sigma_1^{\psi(2S)}$)
9	(fit $\sigma_1^{\psi(2S)}, \sigma^{X(3872)} = 1.52 \times \sigma_1^{\psi(2S)}$) $\rightarrow$ (fit $\sigma_1^{\psi(2S)}, \sigma^{X(3872)} = \mathbf{free} \times \sigma_1^{\psi(2S)}$)
	Background variation
10	2^{nd} order polynomial term $\times$ exponential $\rightarrow$ 3^{rd} order polynomial term $\times$ exponential

Table 6.1 Summary of mass fit model variations for systematic error studies.

6.1.7 z-displacement of Primary Vertices

The distribution of primary vertices along the z-axis of the detector could have an effect on cross section measurements in several ways. The z-displacement of the vertices is measured to have some difference between data and simulation. This effect is already considered on pion reconstruction, studied in Appendix C.5, which is included as a systematic error in the pion reconstruction efficiency. In addition, the z-displacement will also cause a smearing of the acceptance criteria at the boundaries $|\eta(\mu^{\pm})| < 2.3$ and $|\eta(\pi^{\pm})| < 2.4$, and smearing of the rapidity selection boundary $|y(J/\psi\pi^+\pi^-)| < 0.75$. In the case of acceptance, the smearing is found to make no difference to the cross section measurement, due to the muons and pions lying far inside acceptance boundaries, along

with the smearing effect being small. For the rapidity selection, the effect of smearing causes a cross section uncertainty limited to no more than about 1%. Details of how the z-displacement effects were estimated can be found in Appendix I.3.

6.1.8 Luminosity

The uncertainty of the integrated luminosity measurement during the 2012 data-taking period is 1.9%, using a methodology described in [100] by the ATLAS Luminosity Group.

6.1.9 Combined Uncertainties

We produce uncertainty plots showing the contribution of each of the systematic and statistical errors, and combine by adding in quadrature to find the total error of the cross section measurement for $\psi(2S)$ (Figure 6.1) and $X(3872)$ (Figure 6.2) for each p_T bin. Additionally, we create uncertainty maps for the non-prompt $\psi(2S)$ and $X(3872)$ fractions (Figure 6.3), and also for the ratio of the measured $X(3872)$ and $\psi(2S)$ cross sections for prompt and non-prompt production (Figures 6.4 and 6.5).

The sources of various uncertainties and their smallest (Min), median (Med) and largest (Max) values across the p_{T} bins are summarised in Table 6.2 for the differential cross sections of $X(3872)$ and $\psi(2S)$ states, and in Table 6.3 for the measured fractions. The statistical and individual systematic uncertainties are added in quadrature to form the total error shown in the tables. In general, the results for $X(3872)$ are dominated by statistical errors, while for $\psi(2S)$ statistical and systematic uncertainties are of comparable size. The last row in Tables 6.2 and 6.3 shows the relative differences between the values obtained using the single- and the two-lifetime fits, labelled as '1L-fit' and '2L-fit', respectively. For the quantities listed in Tables 6.2 and 6.3, these differences were found to be generally fairly small, compared to the combined systematic uncertainty from other sources.

6.1.10 Short-lived Fractions and other Ratios

In the case of F_{SL} for $\psi(2S)$, we use the systematic errors from the non-prompt $\psi(2S)$ fraction, shown at the top of Figure 6.3. Accordingly, in the case of F_{SL} for $X(3872)$, the systematic errors obtained for the non-prompt $X(3872)$ fraction, shown at the

bottom of Figure 6.3 are used. As for the ratios of various lifetime components of $X(3872)$ and $\psi(2S)$, the systematic errors derived for the non-prompt $X(3872)/\psi(2S)$ ratio, shown in Figure 6.5, are used.

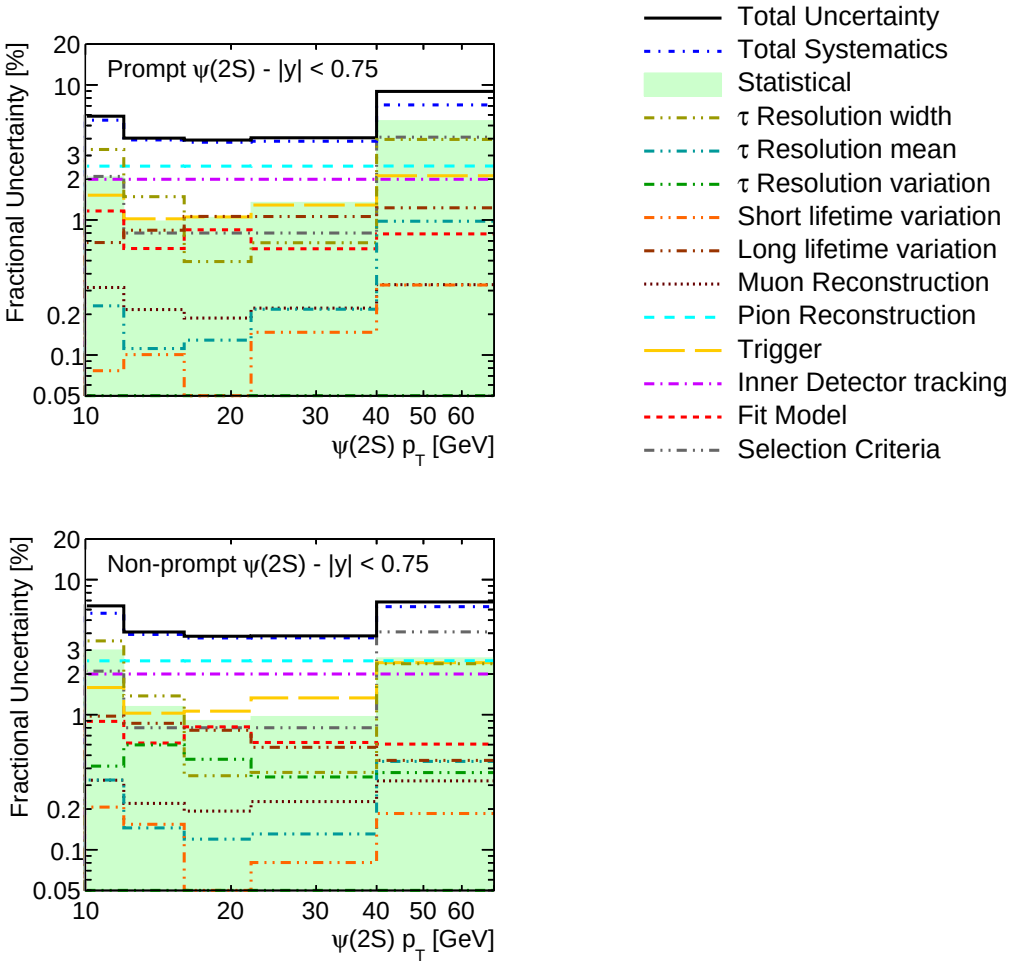


Fig. 6.1 Summary of uncertainties for the measured prompt and non-prompt $\psi(2S)$ cross sections. The 1.9% luminosity uncertainty is not included in the plot.

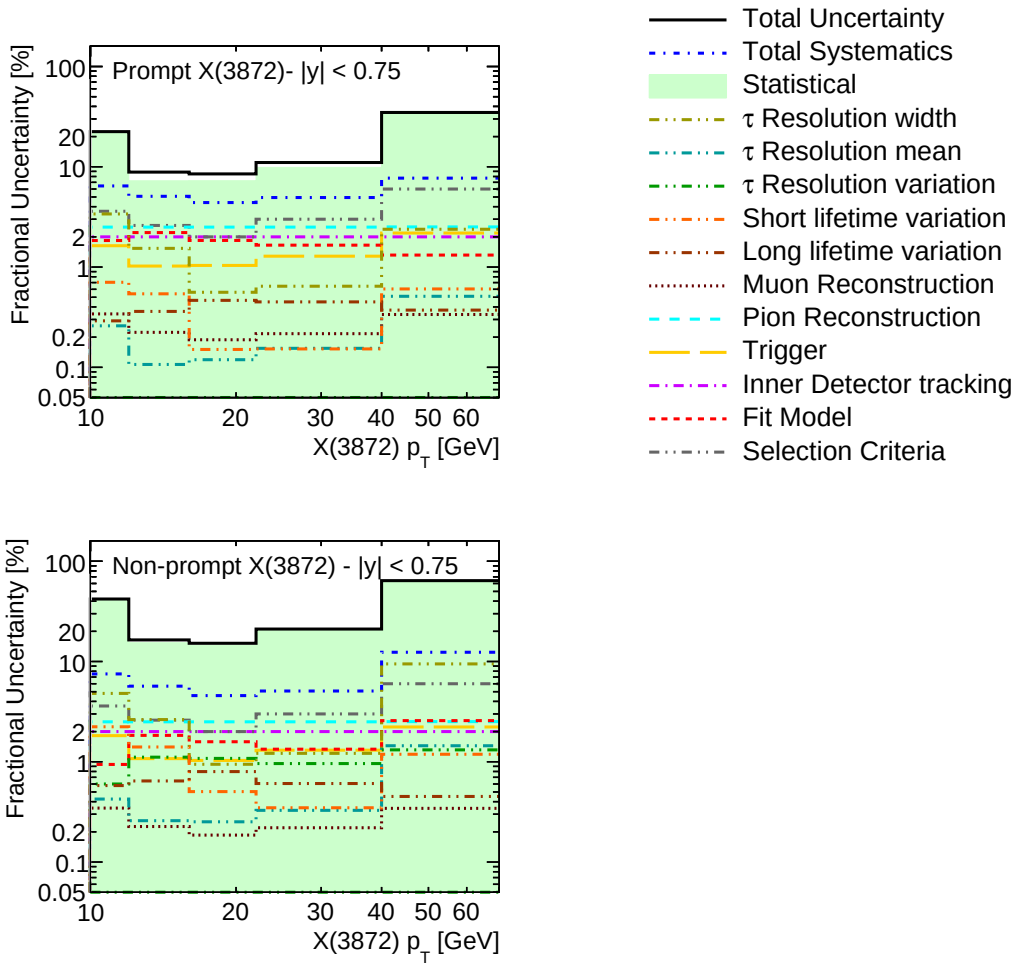


Fig. 6.2 Summary of uncertainties for the measured prompt and non-prompt $X(3872)$ cross sections. The 1.9% luminosity uncertainty is not included in the plot.

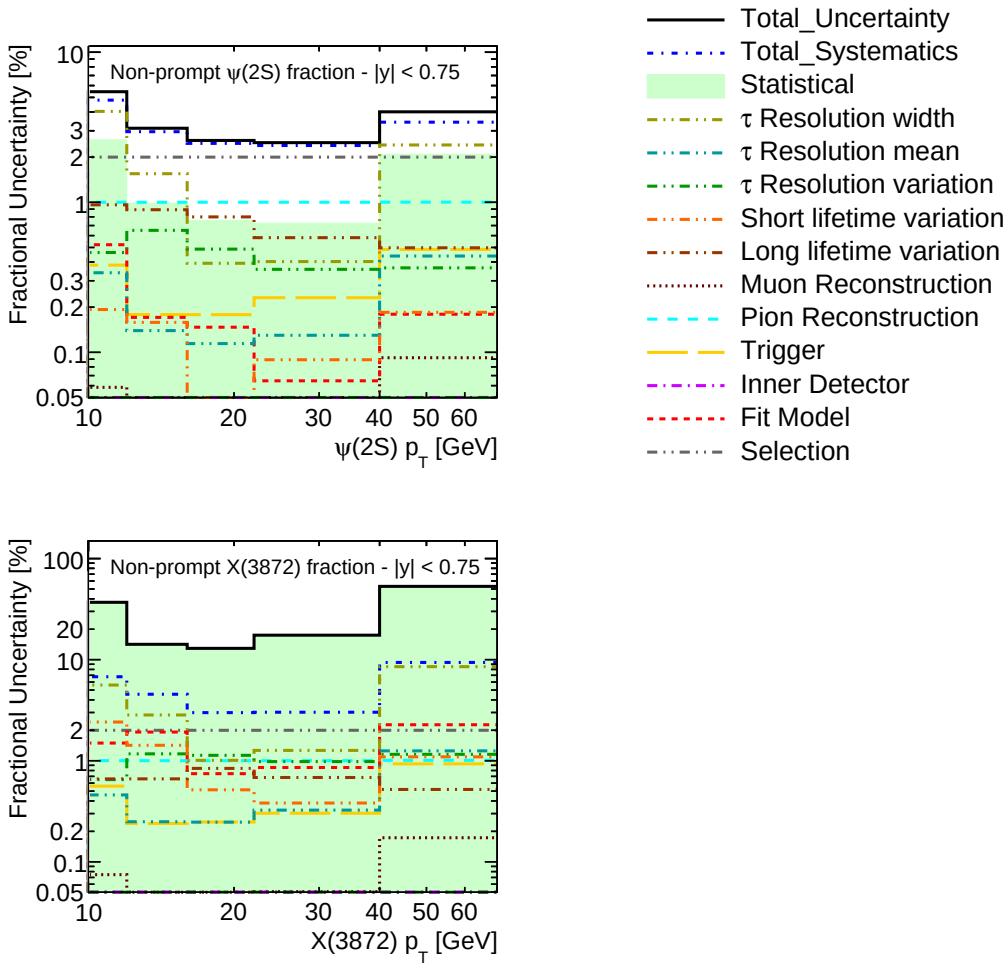


Fig. 6.3 Summary of uncertainties for the measured non-prompt fractions of $\psi(2S)$ (top) and $X(3872)$ (bottom).

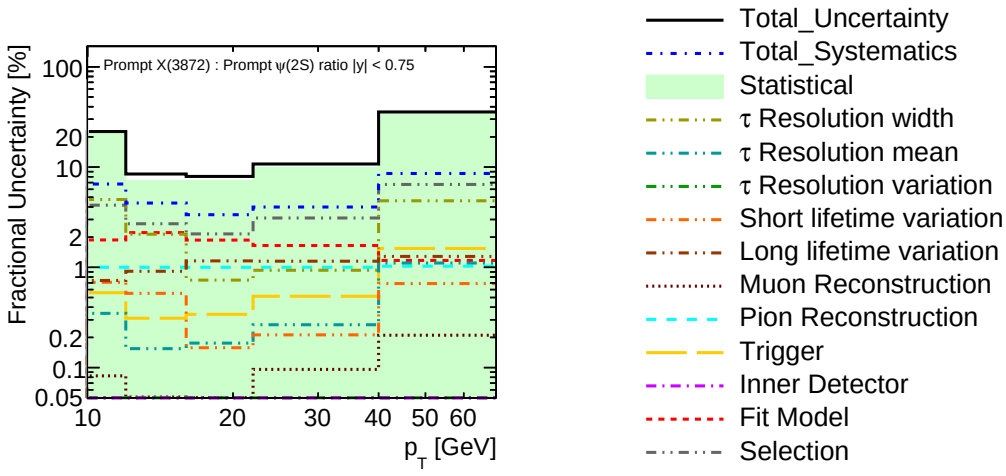


Fig. 6.4 Summary of uncertainties for the ratio of measured cross sections between prompt $X(3872)$ and $\psi(2S)$.

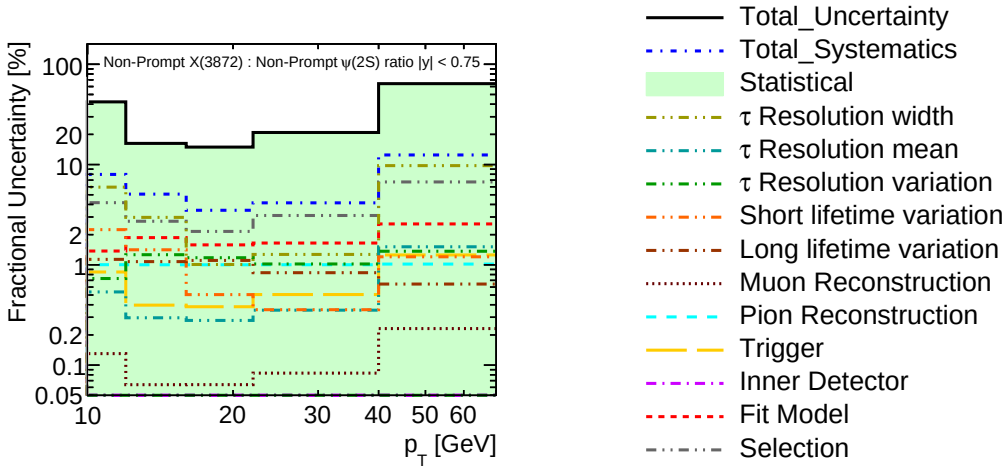


Fig. 6.5 Summary of uncertainties for the ratio of measured cross sections between non-prompt $X(3872)$ and $\psi(2S)$.

	$\psi(2S)$[%]			$X(3872)$[%]		
Source of uncertainty	Min	Med	Max	Min	Med	Max
Statistical	0.9	1.4	5.4	7.3	9.9	63
Trigger eff.	1.0	1.3	2.5	1.1	1.3	2.6
Muon tracking	2.0	2.0	2.0	2.0	2.0	2.0
Muon reconstruction eff.	0.2	0.2	0.3	0.2	0.2	0.4
Pion reconstruction eff.	2.5	2.5	2.5	2.5	2.5	2.5
Bkgd suppression cuts	0.8	0.8	3.0	2.0	3.0	6.0
Mass fit model variation	0.6	0.8	1.2	0.9	1.6	2.6
Short-lifetime variation	0.1	0.2	0.3	0.2	0.7	1.7
Long-lifetime variation	0.6	1.0	1.2	0.3	0.6	0.9
Lifetime resolution model	0.4	1.5	4.0	0.6	2.6	3.4
Total Systematic	3.5	3.6	6.4	4.1	4.9	7.5
(2L-fit − 1L-fit) / 2L-fit (prompt)	−0.1	−0.4	−0.6	−0.3	−0.5	−3.4
(2L-fit − 1L-fit) / 2L-fit (non-prompt)	+0.1	+0.4	+0.7	+0.1	+1.4	+9.8

Table 6.2 Summary of uncertainties for the $\psi(2S)$ and $X(3872)$ cross section measurements showing the smallest (Min), median (Med) and largest (Max) values across the p_{T} bins. The last two rows are described in the text. The uncertainty of the integrated luminosity 2.8% is not included.

	Absolute uncertainty [%]								
	$f_{\rm NP}^{\psi}$			$f_{\rm NP}^{X}$			$f_{\rm SL}^{X}$		
Source of uncertainty	Min	Med	Max	Min	Med	Max	Min	Med	Max
Statistical	0.4	0.5	1.4	4.2	5.8	17.8	16.4	25.8	63
Trigger eff.	0.1	0.1	0.3	0.1	0.1	0.4	0.0	0.1	0.1
Muon tracking eff.	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Muon reconstruction eff.	0.0	0.0	0.1	0.0	0.0	0.1	0.0	0.0	0.1
Pion reconstruction eff.	0.4	0.5	0.7	0.3	0.3	0.4	0.0	0.3	0.4
Bkgd suppression cuts	0.8	1.1	1.4	0.6	0.7	0.7	0.1	0.1	0.7
Mass fit model variation	0.1	0.1	0.2	0.2	0.6	1.8	1.0	1.3	2.4
Lifetime resolution variation	0.2	0.7	1.7	0.4	1.0	2.9	1.8	3.6	12.1
Short-lifetime variation	0.0	0.1	0.1	0.1	0.4	0.8	0.3	0.7	2.8
Long-lifetime variation	0.3	0.4	0.4	0.2	0.2	0.3	3.3	4.0	4.4
Total Systematic	1.3	1.5	2.4	1.0	1.4	3.6	4.1	4.9	13.5
(2L-fit − 1L-fit) / 2L-fit	+0.4	+0.6	+0.9	+0.9	+3.1	+9.1	–	–	–

Table 6.3 Summary of uncertainties for $\psi(2S)$ and $X(3872)$ non-prompt fractions, and short-lived non-prompt fraction for $X(3872)$ production, showing the smallest (Min), median (Med) and largest (Max) values across the $p_{\rm T}$ bins. For the fractions in this table the luminosity uncertainty cancels out. The last row is described in the text.

6.2 Polarisation Variation

Various extreme polarisation scenarios considered in this analysis were described in Sections 5.3.5 and 5.3.6. We find the difference in yields for each $\psi(2S)$ and $X(3872)$ polarisation scenario by comparing the mean of the distribution of weights for the respective signal bands for each polarisation hypothesis, and dividing the mean of the distribution by the mean weight in the unpolarised case. The correction factors for the different polarisation scenarios are shown separately for $\psi(2S)$ (Figure 6.6) and $X(3872)$ (Figure 6.7) across the $p_{\rm T}$ bins, and in tabular form in Tables 6.4 and 6.5.

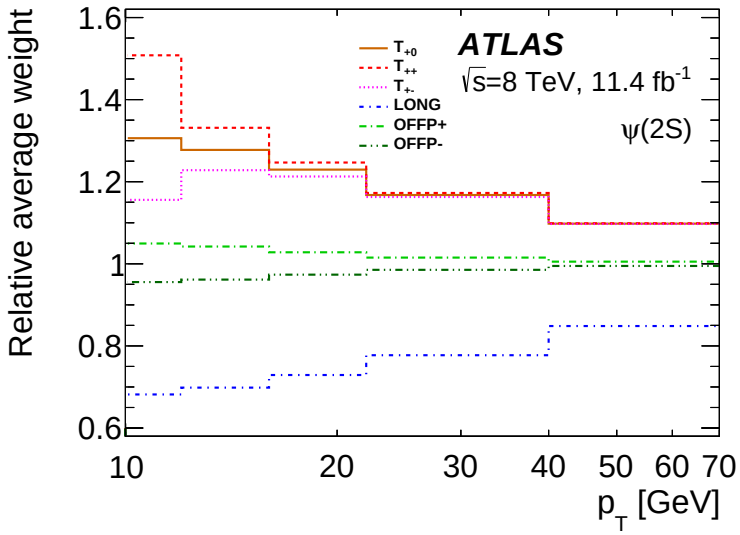


Fig. 6.6 Difference in $\psi(2S)$ yield as a factor of the unpolarised yield for polarisation hypotheses.

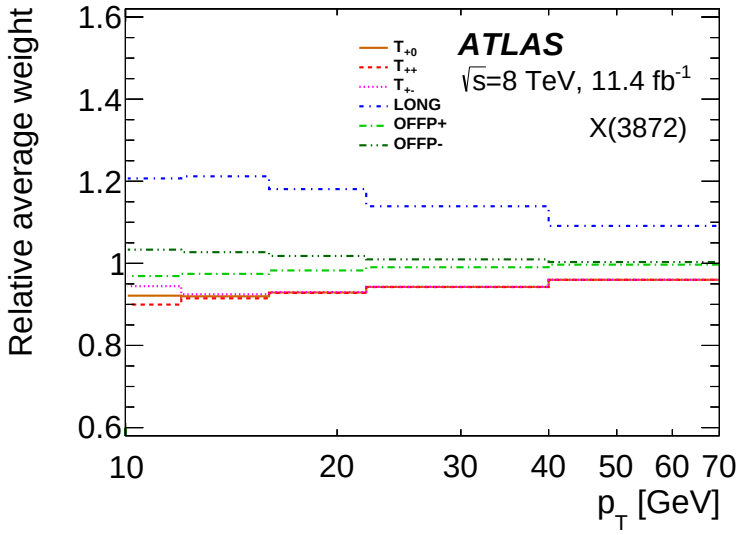


Fig. 6.7 Difference in $X(3872)$ yield as a factor of the unpolarised yield for polarisation hypotheses.

	p_{T} [GeV]				
Polarisation Hypothesis	$10-12$	$12-16$	$16-22$	$22-40$	$40-70$
T_{+0}	1.306	1.277	1.229	1.168	1.098
T_{++}	1.508	1.331	1.247	1.173	1.099
T_{+-}	1.156	1.228	1.213	1.163	1.097
LONG	0.682	0.698	0.729	0.777	0.848
OFFP+	1.049	1.042	1.028	1.015	1.005
OFFP-	0.956	0.962	0.974	0.985	0.995

Table 6.4 Correction factors for various polarisation hypotheses in p_{T} bins for $\psi(2\mathrm{S})$ production.

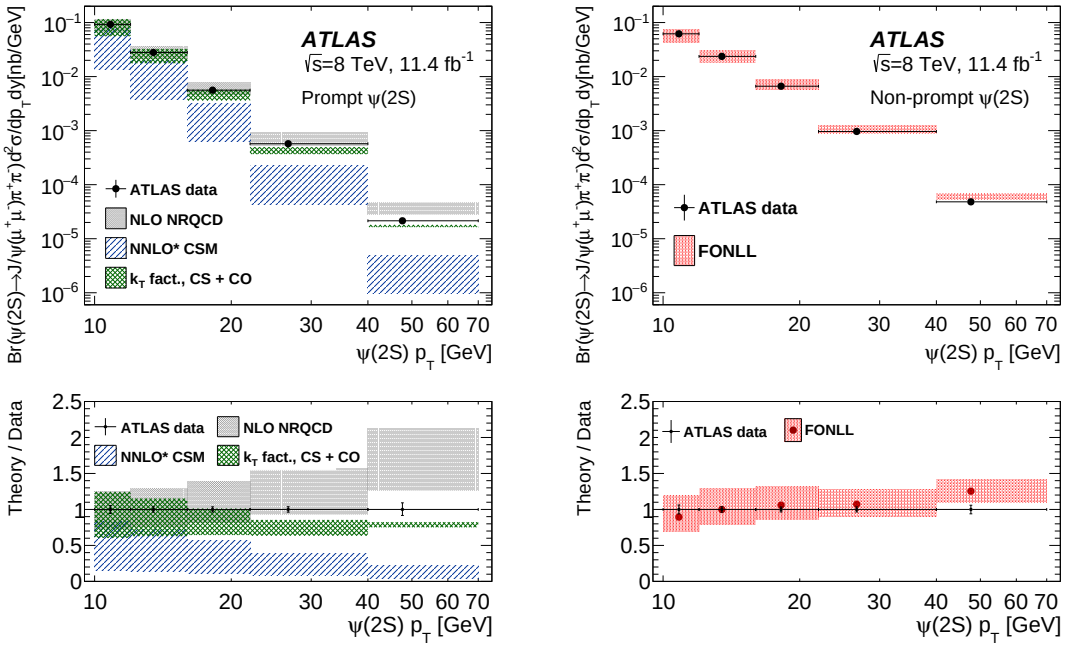
	p_{T} [GeV]				
Polarisation Hypothesis	$10-12$	$12-16$	$16-22$	$22-40$	$40-70$
T_{+0}	0.921	0.920	0.929	0.943	0.960
T_{++}	0.900	0.915	0.928	0.942	0.960
T_{+-}	0.944	0.925	0.930	0.943	0.960
LONG	1.207	1.212	1.181	1.139	1.091
OFFP+	0.969	0.974	0.983	0.990	0.997
OFFP-	1.033	1.027	1.018	1.010	1.003

Table 6.5 Correction factors for various polarisation hypotheses in p_{T} bins for the $X(3872)$ production.

6.3 Results and Discussion

The measured differential cross section (times the product of the relevant branching fractions) for prompt production of $\psi(2S)$ is shown in Figure 6.8(a). It is described reasonably well by the NLO NRQCD model [101] with long-distance matrix elements (LDMEs) determined from the Tevatron data, although some overestimation is observed at the highest p_{T} values. The k_T factorisation model [47] with LDMEs determined from fits to 7 TeV CMS data [102] describes ATLAS data reasonably well with a slight underestimation at higher p_{T}. The NNLO* Colour-Singlet Model (CSM) predictions [33] are also shown. These are close to the data points at low p_{T}, but significantly underestimate them at higher p_{T} values.

The measured differential cross section for non-prompt $\psi(2S)$ production is presented in Figure 6.8(b), compared with the predictions of the FONLL calculation [49] with no free parameters. The calculation describes the data well over the whole range of transverse momenta.

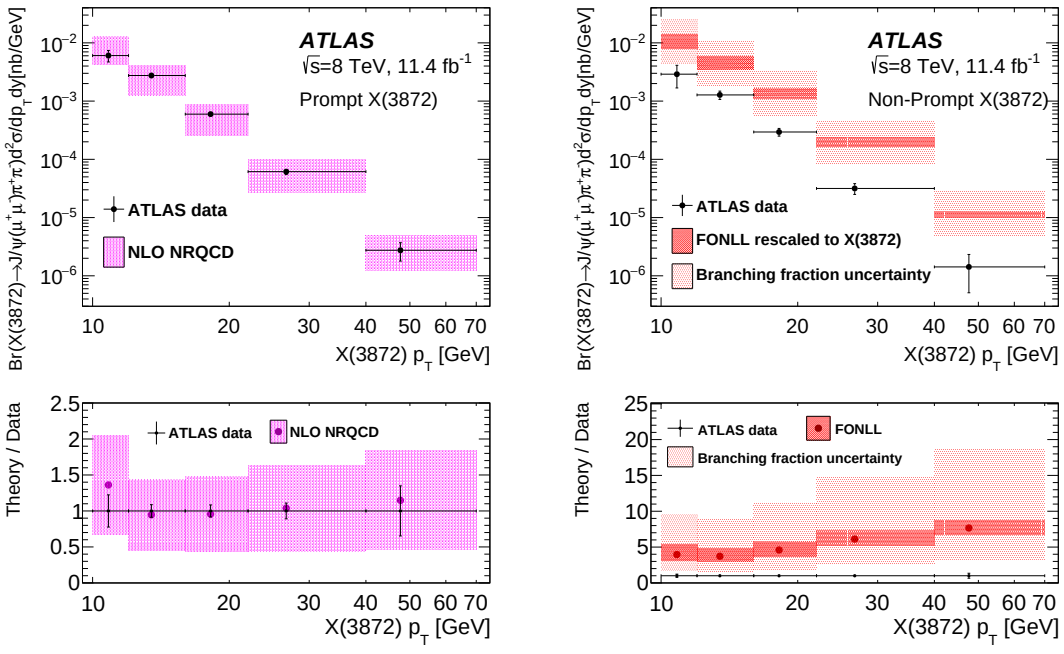


(a) (b)

Fig. 6.8 Measured cross section times branching fractions as a function of p_{T} for (a) prompt $\psi(2S)$ production compared to NLO NRQCD [101] and the NNLO* CSM [33], and (b) non-prompt $\psi(2S)$ production compared to FONLL [49] model predictions.

Similarly, the differential cross section for prompt production of $X(3872)$ is shown in Figure 6.9(a). It is well described by the prediction of the NRQCD model which considers $X(3872)$ to be a mixture of $\chi_{c1}(2P)$ and a $D^0\bar{D}^{*0}$ molecular state [62], with the production being dominated by the $\chi_{c1}(2P)$ component and the normalisation fixed through the fit to CMS data [8].

The measured differential cross section for non-prompt production of $X(3872)$ is shown in Figure 6.9(b). This is compared to a calculation based on the FONLL model prediction for $\psi(2S)$, recalculated for $X(3872)$ using the kinematic template for the non-prompt $X(3872)/\psi(2S)$ ratio shown in Figure 5.14(b) and the effective value of the product of the branching fractions $Br(B \to X(3872) + \text{any})Br(X(3872) \to J/\psi\pi^+\pi^-) = (1.9 \pm 0.8) \times 10^{-4}$ estimated in [61] based on the Tevatron data [103]. This calculation overestimates the data by a factor increasing with p_{T} from about 4 to about 8 over the p_{T} range of this measurement.



(a) (b)

Fig. 6.9 Measured cross section times branching fractions as a function of p_{T} for (a) prompt $X(3872)$ compared to NLO NRQCD predictions with the $X(3872)$ modelled as a mixture of $\chi_{c1}(2P)$ and a $D^0\bar{D}^{*0}$ molecular state [62], and (b) non-prompt $X(3872)$ compared to the FONLL [49] model prediction, recalculated using the branching fraction estimate from [61] as described in the text.

We compare the prompt and non-prompt $\psi(2S)$ differential production cross sections to the CMS result obtained at $\sqrt{s}$ =7 TeV [104] in Figures 6.10(a) and 6.10(b), respectively. Good consistency is seen between each of the measurements for both prompt and non-prompt production. We also compare the prompt $X(3872)$ cross section to the CMS result obtained at $\sqrt{s}$ =7 TeV [8] in Figure 6.10(c). Again, excellent consistency is seen between the measurements, and our result extends the p_{T} range of the measurement from around 30 GeV to 70 GeV.

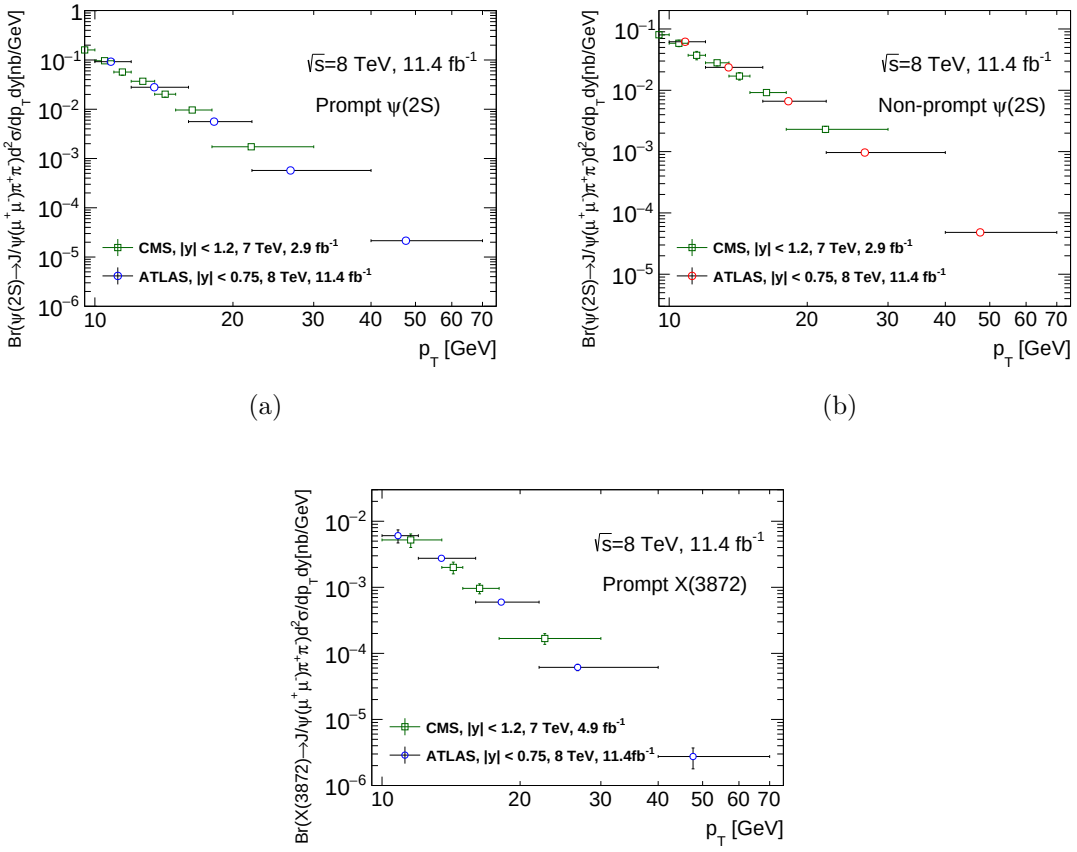


Fig. 6.10 Measured cross section times branching fractions as a function of p_{T} for (a) prompt $\psi(2S)$ production, (b) non-prompt $\psi(2S)$ production, and (c) prompt $X(3872)$ production compared to the CMS results at $\sqrt{s}$ =7 TeV [104, 8]. Good consistency between the results is seen throughout.

The non-prompt fractions of $\psi(2S)$ and $X(3872)$ production are shown in Figure 6.11. In the case of $\psi(2S)$, f_{NP} increases with p_{T}, in good agreement with measurements

obtained with dimuon decays of $\psi(2S)$ from ATLAS [10] and CMS [104]. The non-prompt fraction of $X(3872)$ is essentially independent of $p_{\rm T}$. This measurement agrees within errors with the CMS result obtained at $\sqrt{s}$ =7 TeV [8].

Figure 6.12 shows the $p_{\rm T}$ dependence of the ratio of $X(3872)$ to $\psi(2S)$ cross sections, separately for prompt and non-prompt production contributions. The non-prompt production cross section of $X(3872)$ is further split into short-lived and long-lived components. The short-lived contribution to the non-prompt $\psi(2S)$ production is found to be not significant, see Table 6.6. The measured ratio of long-lived $X(3872)$ to long-lived $\psi(2S)$, shown in Figure 6.12(b) with blue triangles, is fitted with the MC kinematic template described in Section 5.4.4 to obtain

$$\begin{aligned} R_B^{2L} &= \frac{Br(B \to X(3872) + \text{any})Br(X(3872) \to J/\psi\pi^+\pi^-)}{Br(B \to \psi(2S) + \text{any})Br(\psi(2S) \to J/\psi\pi^+\pi^-)} \\ &= (3.57 \pm 0.33(\text{stat}) \pm 0.11(\text{sys}))\%. \end{aligned} \tag{6.1}$$

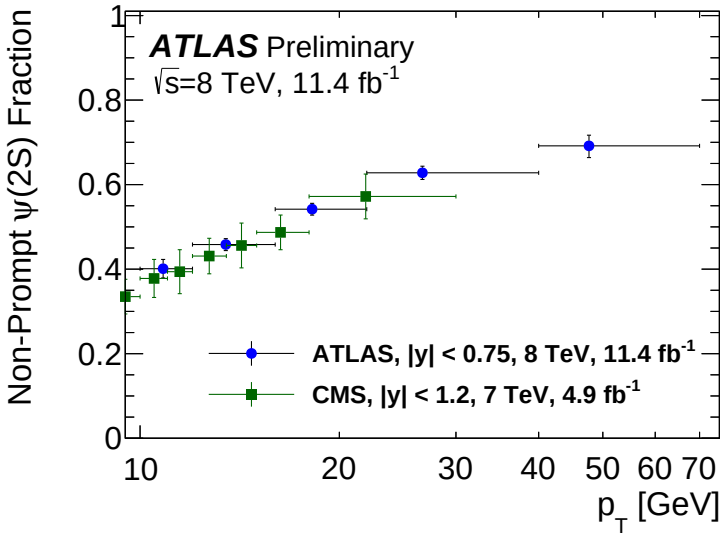
This is somewhat lower than the corresponding result (Equation 5.14) obtained from the same data with the single-lifetime fit model, and both are considerably smaller than the value $(18 \pm 8)\%$ obtained by combining the estimate for the numerator, $(1.9\pm0.8)\times10^{-4}$ [61], obtained from the Tevatron data, with the world average branching fractions from the denominator, $Br(B \to \psi(2S) + \text{any}) = (3.07 \pm 0.21) \times 10^{-3}$ and $Br(\psi(2S) \to J/\psi\pi^+\pi^-) = (34.46 \pm 0.30)\%$.

B_c production in high energy hadronic collisions at low transverse momentum is expected to be dominated by non-fragmentation processes [99]. These processes are expected to have $p_{\rm T}$ dependence $\propto p_{\rm T}^{-2}$ relative to the fragmentation contribution, while it is the fragmentation contribution which dominates the production of long-lived b-hadrons [49]. So the ratio of short-lived non-prompt $X(3872)$ to long-lived non-prompt $\psi(2S)$, shown in Figure 6.12(b) with red squares, is fitted with a function $a/p_{\rm T}^2$ to find $a = 2.04 \pm 1.43(\text{stat}) \pm 0.34(\text{sys})$ GeV^2. This value, and the measured non-prompt yields of $X(3872)$ and $\psi(2S)$, are used to determine the fraction of non-prompt $X(3872)$ from short-lived sources, integrated over the $p_{\rm T}$ range ($p_{\rm T} > 10$ GeV) covered in this measurement:

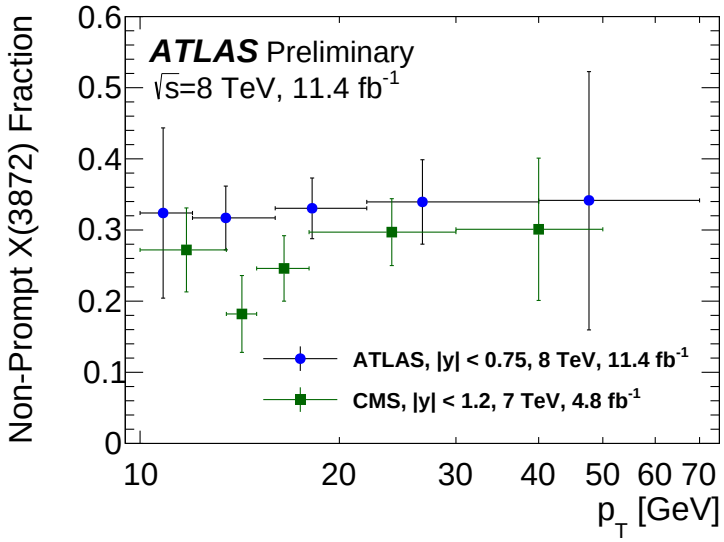
$$\frac{\sigma(pp \to B_c + \text{any})Br(B_c \to X(3872) + \text{any})}{\sigma(pp \to \text{non-prompt } X(3872) + \text{any})} = (25 \pm 13(\text{stat}) \pm 2(\text{sys}) \pm 5(\text{spin}))\%, \tag{6.2}$$

where the last uncertainty comes from the variation of the spin alignment of $X(3872)$ over the extreme scenarios discussed in the Section 6.2. Since the cross section for B_c production makes up only a small fraction of the inclusive beauty production, this value of the ratio would mean that the production of $X(3872)$ in B_c decays is strongly enhanced compared to its production in the decays of other b-hadrons, although we report our result as consistent with zero at 1.8σ.

The numerical values of all cross sections and fractions shown in Figures 6.8 - 6.12 are presented in Table 6.6.

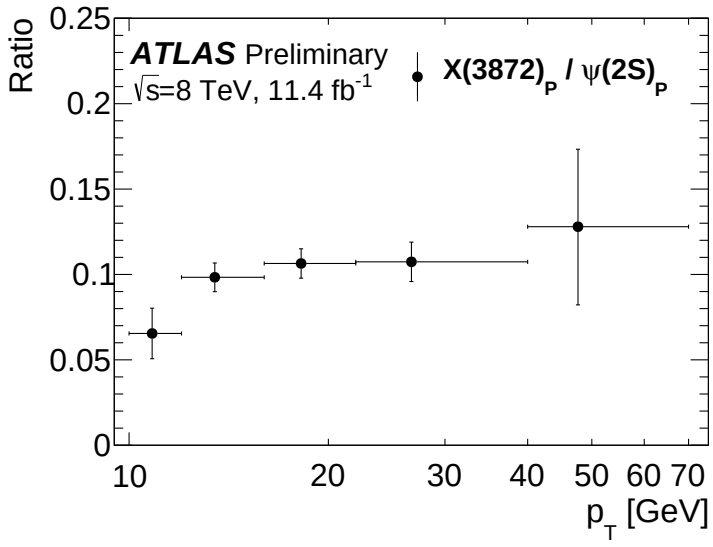


(a)

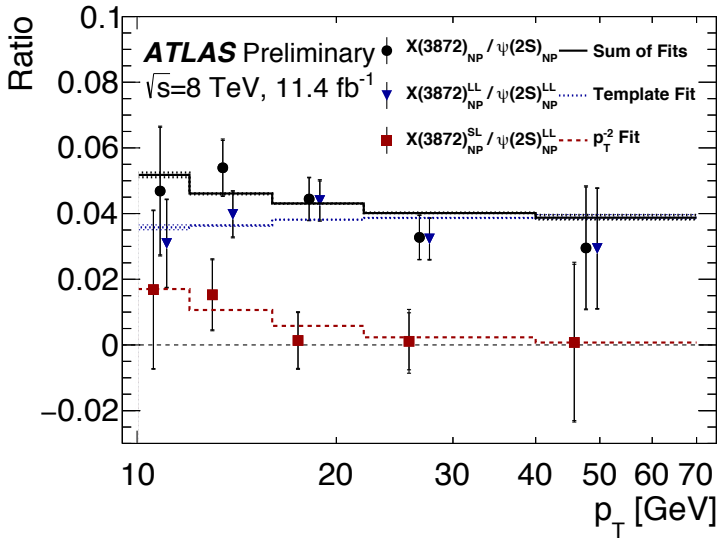


(b)

Fig. 6.11 Measured non-prompt fractions for (a) $\psi(2S)$ and (b) $X(3872)$ production, compared to CMS results at $\sqrt{s} = 7$ TeV. The blue circles are the results shown in this paper, while the green squares show CMS results [104, 8].



(a)



(b)

Fig. 6.12 Ratio of cross section times branching fraction between $X(3872)$ and $\psi(2S)$ for (a) prompt and (b) non-prompt production. In (b), the total non-prompt ratio (black circles) is separated into short-lived (red squares) and long-lived (blue triangles) components for the $X(3872)$, shown with respective fits described in the text. The quality of all three fits is good, with $\chi^2/\text{dof} = 0.43/4$, $2.3/4$ and $2.2/4$ for SL, LL and total NP components, respectively. The data points are slightly shifted horizontally for visibility.

	p_{T} range [GeV]				
	$10-12$	$12-16$	$16-22$	$22-40$	$40-70$
	Cross Sections [pb/GeV/Δy]				
$\psi(2S)_P$	$92.41 \pm 1.93 \pm 4.84$	$27.97 \pm 0.27 \pm 1.02$	$5.61 \pm 0.06 \pm 0.19$	$0.57 \pm 0.01 \pm 0.02$	$0.021 \pm 0.001 \pm 0.001$
$\psi(2S)_{NP}$	$61.88 \pm 1.86 \pm 3.35$	$23.66 \pm 0.27 \pm 0.85$	$6.63 \pm 0.06 \pm 0.22$	$0.97 \pm 0.01 \pm 0.03$	$0.048 \pm 0.001 \pm 0.003$
$\psi(2S)_{NP}^{LL}$	$60.76 \pm 1.56 \pm 3.98$	$23.09 \pm 0.27 \pm 1.46$	$6.53 \pm 0.06 \pm 0.41$	$0.93 \pm 0.01 \pm 0.06$	$0.047 \pm 0.002 \pm 0.003$
$\psi(2S)_{NP}^{SL}$	$1.12 \pm 2.43 \pm 3.85$	$0.56 \pm 0.37 \pm 1.14$	$0.11 \pm 0.08 \pm 0.29$	$0.04 \pm 0.01 \pm 0.04$	$0.001 \pm 0.002 \pm 0.002$
$X(3872)_P$	$6.05 \pm 1.30 \pm 0.38$	$2.75 \pm 0.20 \pm 0.13$	$0.60 \pm 0.04 \pm 0.02$	$0.06 \pm 0.01 \pm 0.00$	$0.003 \pm 0.001 \pm 0.000$
$X(3872)_{NP}$	$2.90 \pm 1.20 \pm 0.21$	$1.28 \pm 0.20 \pm 0.07$	$0.29 \pm 0.04 \pm 0.01$	$0.03 \pm 0.01 \pm 0.00$	$0.001 \pm 0.001 \pm 0.000$
$X(3872)_{NP}^{LL}$	$1.87 \pm 0.82 \pm 0.14$	$0.92 \pm 0.16 \pm 0.06$	$0.29 \pm 0.04 \pm 0.02$	$0.03 \pm 0.01 \pm 0.00$	$0.001 \pm 0.001 \pm 0.000$
$X(3872)_{NP}^{SL}$	$1.02 \pm 1.49 \pm 0.20$	$0.35 \pm 0.25 \pm 0.06$	$0.01 \pm 0.06 \pm 0.02$	$0.00 \pm 0.01 \pm 0.00$	$0.000 \pm 0.001 \pm 0.000$
	Fractions				
$F_{NP}^{\psi(2S)}$	$0.40 \pm 0.01 \pm 0.02$	$0.46 \pm 0.00 \pm 0.01$	$0.54 \pm 0.00 \pm 0.01$	$0.63 \pm 0.00 \pm 0.01$	$0.69 \pm 0.01 \pm 0.02$
$F_{SL}^{\psi(2S)}$	$0.02 \pm 0.04 \pm 0.06$	$0.02 \pm 0.02 \pm 0.05$	$0.02 \pm 0.01 \pm 0.04$	$0.04 \pm 0.01 \pm 0.04$	$0.03 \pm 0.03 \pm 0.05$
$F_{NP}^{X(3872)}$	$0.32 \pm 0.12 \pm 0.02$	$0.32 \pm 0.04 \pm 0.01$	$0.33 \pm 0.04 \pm 0.01$	$0.34 \pm 0.06 \pm 0.01$	$0.34 \pm 0.18 \pm 0.03$
$F_{SL}^{X(3872)}$	$0.35 \pm 0.39 \pm 0.05$	$0.28 \pm 0.16 \pm 0.04$	$0.03 \pm 0.19 \pm 0.05$	$0.03 \pm 0.26 \pm 0.05$	$0.03 \pm 0.63 \pm 0.13$
	Ratios				
$X(3872)_P/\psi(2S)_P$	$0.065 \pm 0.014 \pm 0.004$	$0.098 \pm 0.007 \pm 0.004$	$0.106 \pm 0.008 \pm 0.004$	$0.107 \pm 0.011 \pm 0.004$	$0.128 \pm 0.044 \pm 0.012$
$X(3872)_{NP}/\psi(2S)_{NP}$	$0.047 \pm 0.019 \pm 0.004$	$0.054 \pm 0.008 \pm 0.003$	$0.044 \pm 0.006 \pm 0.002$	$0.033 \pm 0.007 \pm 0.001$	$0.030 \pm 0.019 \pm 0.003$
$X(3872)_{NP}^{LL}/\psi(2S)_{NP}^{LL}$	$0.031 \pm 0.014 \pm 0.002$	$0.040 \pm 0.007 \pm 0.003$	$0.044 \pm 0.006 \pm 0.003$	$0.033 \pm 0.006 \pm 0.002$	$0.030 \pm 0.019 \pm 0.003$
$X(3872)_{NP}^{SL}/\psi(2S)_{NP}^{LL}$	$0.016 \pm 0.024 \pm 0.003$	$0.015 \pm 0.011 \pm 0.003$	$0.001 \pm 0.008 \pm 0.002$	$0.001 \pm 0.009 \pm 0.004$	$0.001 \pm 0.024 \pm 0.005$

Table 6.6 Summary of $\psi(2S)$ and $X(3872)$ production measurements, fractions and ratios. The indices P and NP denote prompt and non-prompt components, while indices SL and LL stand for short-lived and long-lived non-prompt components, respectively. The first uncertainty is statistical, the second is systematic. Uncertainties from integrated luminosity (2.8%) and those due to unknown spin alignment are not included.

Chapter 7

Dipion Invariant Mass Spectra

The distributions of the dipion invariant mass $m_{\pi\pi}$ in the $\psi(2S) \to J/\psi\pi^+\pi^-$ and $X(3872) \to J/\psi\pi^+\pi^-$ decays are measured by determining the corrected yields of $\psi(2S)$ and $X(3872)$ signals in narrow bins of $m_{\pi\pi}$. Some of the selection criteria aimed at the reduction of combinatorial background in the cross section measurement, namely

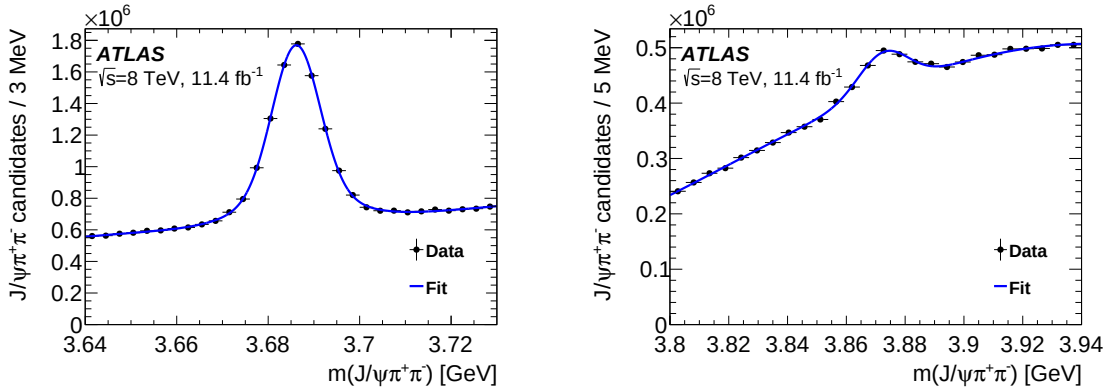
- $Q\left(J/\psi\pi^+\pi^-\right) = m_{J/\psi\pi^+\pi^-} - m^{\mathrm{PDG}}_{J/\psi} - m_{\pi^+\pi^-} < 300$ MeV
- $\Delta R\left(J/\psi, \pi^\pm\right) < 0.5$

are found to bias the $m_{\pi\pi}$ distributions and are removed for this analysis. For this study further requirements are placed on the pseudo-proper lifetime significance, $\tau/\Delta\tau < 2.5$, and the transverse momentum of the $J/\psi\pi^+\pi^-$ candidates, $p_{\mathrm{T}} > 12$ GeV, aimed at reducing the combinatorial background.

The invariant mass distributions of the corrected $J/\psi\pi^+\pi^-$ candidates selected for this analysis are shown in Figure 7.1(a) for the mass range around the $\psi(2S)$ peak and in Figure 7.1(b) for $X(3872)$. The interval of allowed $m_{\pi\pi}$ values is subdivided into 21 and 11 bins for $\psi(2S)$ and $X(3872)$, respectively. In each $m_{\pi\pi}$ bin, the signal yield is extracted using a fit to the function

$$f(m) = f_1 G_1(m) + (1 - f_1) G_2(m) + N_{\mathrm{bkg}} \left(\frac{m - p_0}{m_0 - p_0}\right)^{p_1} e^{-p_2(m-p_0) - p_3(m-p_0)^2}, \quad (7.1)$$

where m is the invariant mass of the $J/\psi\pi^+\pi^-$ system, N_{bkg} is the normalisation factor of the background PDF, m_0 is the world average mass [18] of the parent resonance, and $p_{0,1,2,3}$ are free parameters. The signals are described by the same double-Gaussian



(a) (b)

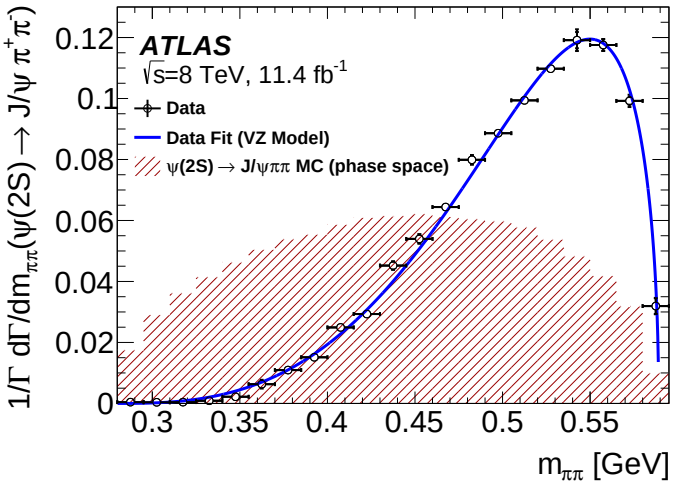
Fig. 7.1 The invariant mass distributions of the $J/\psi\pi^+\pi^-$ candidates to extract (a) $\psi(2S)$ and (b) $X(3872)$ signal integrated over a wide range of $m_{\pi\pi}$.

PDFs $f_1 G_1(m) + (1 - f_1)G_2(m)$ as the ones used in the cross section analysis described in Section 5.2. In most $m_{\pi\pi}$ bins the position of the signal peak is determined from the fit, however in some bins with small signal yields it is necessary to fix the centre and the width of the signal peak to the values obtained from the fits over the whole $m_{\pi\pi}$ range shown in Figure 7.1. As in the cross section analysis, the fraction of the narrow Gaussian f_1 is fixed to 0.76 ± 0.04, varied within the range of ± 0.04 during systematic studies. In another systematic variation a first order polynomial is added as a factor to the PDF in Equation 7.1. For both the $\psi(2S)$ and $X(3872)$ samples, the errors from the fits in $m_{\pi\pi}$ bins are found to be statistically dominated.

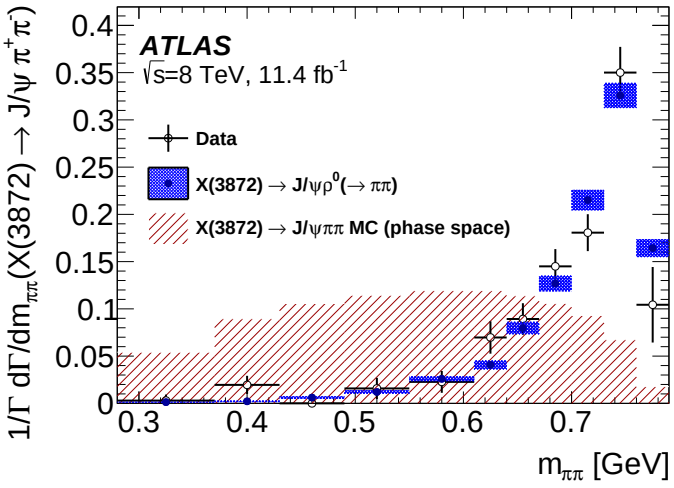
The resulting normalised differential distributions in $m_{\pi\pi}$ are shown in Figure 7.2(a) for the $\psi(2S) \to J/\psi\pi^+\pi^-$ and in Figure 7.2(b) for $X(3872) \to J/\psi\pi^+\pi^-$ decays. The solid blue curve in Figure 7.2(a) represents a fit to the data points with the Voloshin-Zakharov distribution [98]

$$\frac{1}{\Gamma}\frac{d\Gamma}{dm_{\pi\pi}} \propto \left(m_{\pi\pi}^2 - \lambda m_\pi^2\right)^2 \times \mathrm{PS}, \tag{7.2}$$

where PS stands for the dipion phase space. The fitted value of the parameter λ is found to be $\lambda = 4.16 \pm 0.06(\mathrm{stat}) \pm 0.03(\mathrm{sys})$, in agreement with $\lambda = 4.35 \pm 0.18$ measured by BES [105], and $\lambda = 4.46 \pm 0.25$ measured by LHCb [106]. The shaded blue histogram in Figure 7.2(b) is obtained from straightforward simulations, assuming the



(a)



(b)

Fig. 7.2 (a) Normalised differential decay width of $\psi(2S) \to J/\psi(\to \mu^+\mu^-)\pi^+\pi^-$ in bins of dipion invariant mass over the range 0.28 GeV $< m_{\pi\pi} <$ 0.595 GeV, fitted with the Voloshin-Zakharov model. Also shown in red is the normalised distribution of $m_{\pi\pi}$ phase space. (b) Normalised differential decay width of $X(3872) \to J/\psi(\to \mu^+\mu^-)\pi^+\pi^-$ in bins of dipion invariant mass over the range 0.28 GeV $< m_{\pi\pi} <$ 0.79 GeV. Also shown is the MC prediction for the decay $X(3872) \to J/\psi(\to \mu^+\mu^-)\rho^0(\to \pi^+\pi^-)$ (blue histogram) and the normalised distribution of $m_{\pi\pi}$ phase space (red histogram).

dipion system in the decay $X(3872) \to J/\psi\pi^+\pi^-$ is produced purely via the ρ^0 meson, and appears to be in good agreement with the data. In both decays the measured $m_{\pi\pi}$ spectrum strongly disfavours the dipion phase space distribution, shown in Figures 7.2(a) and 7.2(b) with the red shaded areas.

Chapter 8

Summary

The measurement of the differential production cross section of $\psi(2S)$ and $X(3872)$ states in the $J/\psi\pi^+\pi^-$ final state is carried out using 11.4 fb^{-1} of pp collision data recorded by the ATLAS detector at a centre-of-mass energy $\sqrt{s} = 8$ TeV. The prompt and non-prompt production of $\psi(2S)$ and $X(3872)$ is studied separately, as a function of transverse momentum in the rapidity region $|y| < 0.75$ and transverse momentum range 10 GeV $< p_{\mathrm{T}} <$ 70 GeV.

The prompt $\psi(2S)$ cross section measurements show good consistency with theoretical predictions based on NLO NRQCD, and k_T factorisation with the colour-octet component tuned to 7 TeV CMS data, at low p_{T}, with slight deviation apparent at high p_{T}. The predictions from the NNLO* colour-singlet model calculations underestimate the data, especially at higher p_{T}. The non-prompt $\psi(2S)$ measurements show excellent agreement with FONLL predictions across the whole p_{T} range. Both prompt and non-prompt $\psi(2S)$ cross section measurements show good agreement with previous measurements.

The prompt $X(3872)$ cross section measurement shows good agreement with the CMS result [8] in the area of transverse momenta 10 GeV $< p_{\mathrm{T}} <$ 30 GeV where they overlap, and extends the range of transverse momenta up to 70 GeV. Good agreement is found with theoretical predictions within the NLO NRQCD model, which considers $X(3872)$ to be a mixture of $\chi_{c1}(2P)$ and a $D^0\bar{D}^{*0}$ molecular state, with the production being dominated by the $\chi_{c1}(2P)$ component and the normalisation fixed through the fit to CMS data [62]. The non-prompt production of $\psi(2S)$ is described well by the FONLL predictions, but the same predictions, recalculated for $X(3872)$ using the

branching fraction extracted from the Tevatron data, overestimate the non-prompt production of $X(3872)$, especially at large transverse momenta.

Two models of lifetime dependence of the non-prompt production are considered: a model with a single effective lifetime, and an alternative model with two distinctly different effective lifetimes. The two models give virtually identical results for the prompt and non-prompt differential cross sections of $\psi(2S)$ production and for the prompt cross section of $X(3872)$, however the non-prompt production of $X(3872)$ is slightly enhanced in the two lifetime model.

Within the single-lifetime model, under the assumption that non-prompt $\psi(2S)$ and $X(3872)$ come from the same mix of parent b-hadrons, the following result is obtained for the ratio of the branching fractions:

$$\begin{aligned} R_B^{1L} &= \frac{Br(B \to X(3872) + \text{any})Br(X(3872) \to J/\psi\pi^+\pi^-)}{Br(B \to \psi(2S) + \text{any})Br(\psi(2S) \to J/\psi\pi^+\pi^-)} \\ &= (3.95 \pm 0.32(\text{stat}) \pm 0.08(\text{sys}))\%. \end{aligned} \tag{8.1}$$

In the two-lifetime model, the two lifetimes are fixed to expected values for $\psi(2S)$ and $X(3872)$ originating from the decays of B_c and from long-lived b-hadrons, with their relative weight determined from the fits to the data. The ratio of the branching fractions R_B is determined from the long-lived component alone:

$$\begin{aligned} R_B^{2L} &= \frac{Br(B \to X(3872) + \text{any})Br(X(3872) \to J/\psi\pi^+\pi^-)}{Br(B \to \psi(2S) + \text{any})Br(\psi(2S) \to J/\psi\pi^+\pi^-)} \\ &= (3.57 \pm 0.33(\text{stat}) \pm 0.11(\text{sys}))\%. \end{aligned} \tag{8.2}$$

In this two-lifetime model, the fraction of the short-lived non-prompt component in $X(3872)$ production, for $p_{\text{T}} > 10$ GeV, is found to be

$$\frac{\sigma(pp \to B_c + \text{any})Br(B_c \to X(3872) + \text{any})}{\sigma(pp \to \text{non-prompt } X(3872) + \text{any})} = (25 \pm 13(\text{stat}) \pm 2(\text{sys}) \pm 5(\text{spin}))\%. \tag{8.3}$$

The invariant mass distributions of the dipion system in $\psi(2S) \to J/\psi\pi^+\pi^-$ and $X(3872) \to J/\psi\pi^+\pi^-$ decays are also measured. For the $\psi(2S)$, the dipion spectrum is fitted with the Voloshin-Zakharov model to obtain $\lambda = 4.16 \pm 0.08$, in agreement with previous results: $\lambda = 4.35 \pm 0.18$ from BES [105] and $\lambda = 4.46 \pm 0.25$ from LHCb [106]. For $X(3872) \to J/\psi\pi^+\pi^-$, the dipion mass spectrum shows a peak consistent with the

dipion system being produced dominantly through an intermediate ρ^0 meson decay, in agreement with previous measurements from CMS [8], Belle [89] and CDF [107].

Appendix A

Data and Simulation Samples

The input data for this analysis are from periods **C** - **L** of the $\sqrt{s} = 8$ TeV 2012 B-physics group DAODs. We use the JpsiFinder algorithm to find common vertices of offline reconstructed muons, and subsequently perform four-track vertex fits of the muons and pairs of hadronic tracks, with the J/ψ mass constained to the world-average and the hadronic tracks given a pion mass hypothesis, to extract the $J/\psi\pi^+\pi^-$ candidates.

- `data12_8TeV.periodC.physics_Bphysics.PhysCont.DAOD_JPSIMUMU.grp14_v04_p1425/`
- `data12_8TeV.periodD.physics_Bphysics.PhysCont.DAOD_JPSIMUMU.grp14_v04_p1425/`
- `data12_8TeV.periodE.physics_Bphysics.PhysCont.DAOD_JPSIMUMU.grp14_v03_p1425/`
- `data12_8TeV.periodG.physics_Bphysics.PhysCont.DAOD_JPSIMUMU.grp14_v03_p1425/`
- `data12_8TeV.periodH.physics_Bphysics.PhysCont.DAOD_JPSIMUMU.grp14_v04_p1425/`
- `data12_8TeV.periodI.physics_Bphysics.PhysCont.DAOD_JPSIMUMU.grp14_v03_p1425/`
- `data12_8TeV.periodJ.physics_Bphysics.PhysCont.DAOD_JPSIMUMU.grp14_v03_p1425/`
- `data12_8TeV.periodL.physics_Bphysics.PhysCont.DAOD_JPSIMUMU.grp14_v03_p1425/`

We use a group 2012 MC sample in order to measure the pion reconstruction efficiency

- `mc12_8TeV.208001.Pythia8B_AU2_CTEQ6L1_pp_Jpsimu4mu4.merge.AOD.e1331_a159_a173_r3549/`

which was created with the geometry version ATLAS-GEO-20-00-01. The pion reconstruction efficiency is dependent on the simulation of the detector material, and the resulting ntuples are reweighted according to the current *"best"* geometry tag ATLAS-GEO-20-01-00, which has an improved ID material description.

Simulation samples of prompt and non-prompt produced $\psi(2S) \to J/\psi\pi^+\pi^-$ and $X(3872) \to J/\psi\pi^+\pi^-$ are required to measure certain selection criteria efficiencies. Events are generated using Pythia8B, and are reconstructed under ATLAS 2012 run conditions. Pythia8B does not contain $X(3872)$, so we use $\chi_{c1}(1P)$ instead, as having the same ($I^G(J^{PC}) = 0^+(1^{++})$) quantum numbers, setting its mass to the $X(3872)$ world-average.

- mc12_8TeV.208022.Pythia8B_AU2_CTEQ6L1_pp_Psi2S_JpsiPiPi.merge.AOD.e3263_a188_a180_r3549/
- mc12_8TeV.208230.Pythia8B_AU2_CTEQ6L1_bb_Psi2S_JpsiPiPi.merge.AOD.e3263_a188_a180_r3549/
- mc12_8TeV.208023.Pythia8B_AU2_CTEQ6L1_pp_X3872_JpsiPiPi.merge.AOD.e3263_a188_a180_r3549
- mc12_8TeV.208231.Pythia8B_AU2_CTEQ6L1_bb_X3872_JpsiPiPi.recon.AOD.e3263_a188_a180/

We also require a B_c simulation sample. We use a sample of $B_c^+ \to J/\psi\pi^+$ decays, which was previously produced and used for [48], where more information can be found.

- mc12_8TeV.108601.PythiaBc_Bc_JPsi_mu2p5mu2p5_Pi.merge.AOD.e1988_a188_a180_r3549

Appendix B

Selection Criteria Studies

B.1 Signal Significance

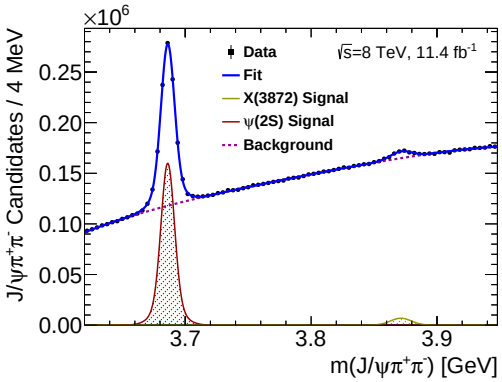
We determine the values of the selection criteria described in Table 5.1 by performing simultaneous binned χ^2 fits of the $\psi(2S)$ and $X(3872)$ signal across $J/\psi\pi^+\pi^-$ invariant mass, and finding the cut values which maximise $X(3872)$ signal significance. The significance is defined as:

$$S = \frac{N_{\text{sig}}}{\sqrt{N_{\text{sig}} + N_{\text{bkg}}}}, \tag{B.1}$$

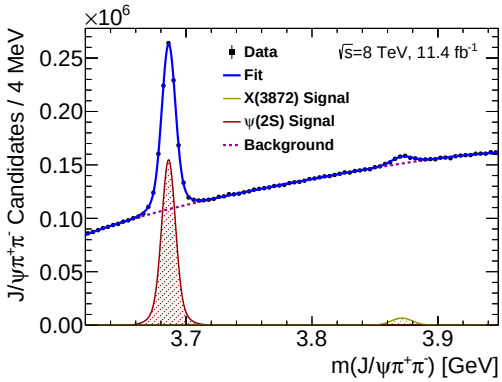
where S is the signal significance and N_{sig} (N_{bkg}) are the number of signal(background) events within a $\pm 3\sigma$ band of the fitted mean. The fits are shown in Figure B.1, with the corresponding cut flow shown in Table B.1.

Selection	N($\psi(2S)$)	N($X(3872)$)	S($\psi(2S)$)	S($X(3872)$)
No additional	588613 ± 3777	35481 ± 2011	478.9 ± 3.1	24.6 ± 1.4
$P(\chi^2)(J/\psi\pi\pi) > 4\%$	568404 ± 1900	34199 ± 1571	477.8 ± 1.6	24.7 ± 1.1
$\Delta R(J/\psi, \pi) < 0.5$	557144 ± 1849	31681 ± 1208	477.3 ± 1.6	27.5 ± 1.0
$Q(J/\psi\pi\pi) < 300$	572764 ± 1746	32726 ± 1717	494.3 ± 1.5	34.8 ± 1.8
$p_T(\pi^{\pm}) > 600$ MeV	465462 ± 1572	30414 ± 883	463.7 ± 1.6	35.5 ± 1.0

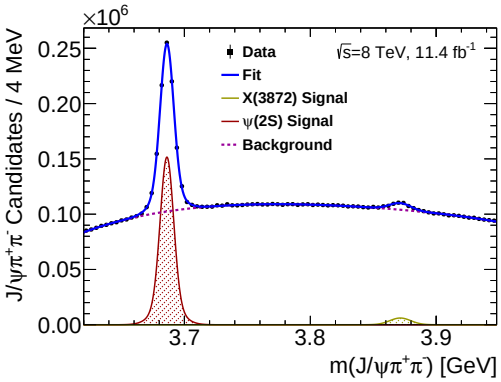
Table B.1 Effect on signal yields and signficance for selection criteria over full p_T range for $\psi(2S)$ and $X(3872)$.



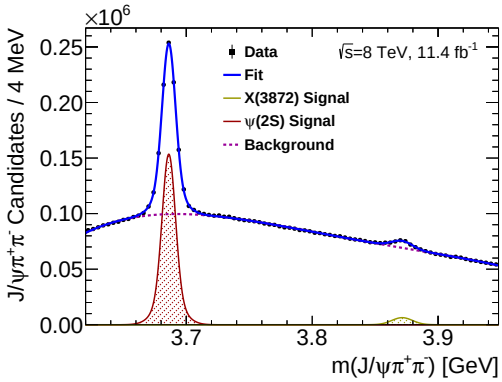
(a)



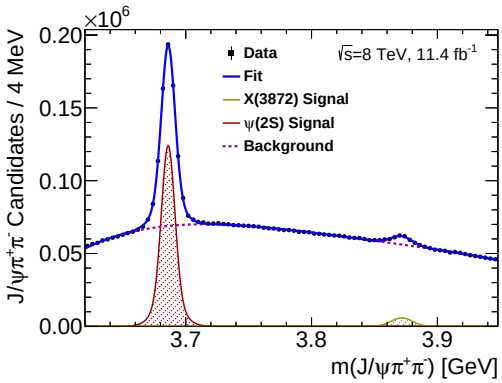
(b)



(c)



(d)



(e)

Fig. B.1 Effect of selection criteria on signal efficiency and significance for $\psi(2S)$ and $X(3872)$ signal for the full data sample. Binned χ^2 fits are performed for (a) no additional cuts, (b) $P(\chi^2)(J/\psi\pi\pi) > 4\%$, (c) $\Delta R(J/\psi, \pi) < 0.5$, (d) $Q(J/\psi\pi\pi) < 300$ MeV, (e) $p_T(\pi^{\pm}) > 600$ MeV.

B.2 Signal Efficiency

We use Monte Carlo samples generated for $\psi(2S)$ and $X(3872)$ production detailed in Section 5.3.8 to measure the efficiency of the selection criteria. The invariant mass of the dipion system for events in the $\psi(2S)$ MC sample are weighted to match that seen in data (the data is fitted with the Voloshin-Sakharov model in Chapter 7) at truth level. Similarly, the $X(3872)$ events are weighted to match the invariant mass distribution of the dipion system when the pions are produced from an intermediate ρ^0 decay, seen in Section 5.3.8, at truth level.

Additionally, the MC generated events are weighted at the reconstruction level to match the data p_{T} distribution of the J/ψ and pions, and also to match the opening angles between the J/ψ and each pion. We use the method of sideband subtraction on the $J/\psi\pi^+\pi^-$ invariant mass spectrum to measure $p_{\mathrm{T}}(J/\psi)$, $p_{\mathrm{T}}(\pi^{\pm})$ and $\Delta R(J/\psi, \pi^{\pm})$ for $\psi(2S)$ and $X(3872)$ signal regions in the data and Monte Carlo. We estimate an invariant mass width of $\sigma = 5$ MeV for the $\psi(2S)$ and $\sigma = 8$ MeV for $X(3872)$ signal, and the PDG mass values are used for the means. The central band covers $\pm 3\sigma$, whilst the upper and lower sidebands cover $(+7 \to +10)\sigma$ and $(-10 \to -7)\sigma$, respectively. The invariant mass spectra of the $J/\psi\pi^+\pi^-$ candidates in the $\psi(2S)$ and $X(3872)$ regions, along with sidebands are shown in Figure B.2 for data, and Figure B.3 for MC.

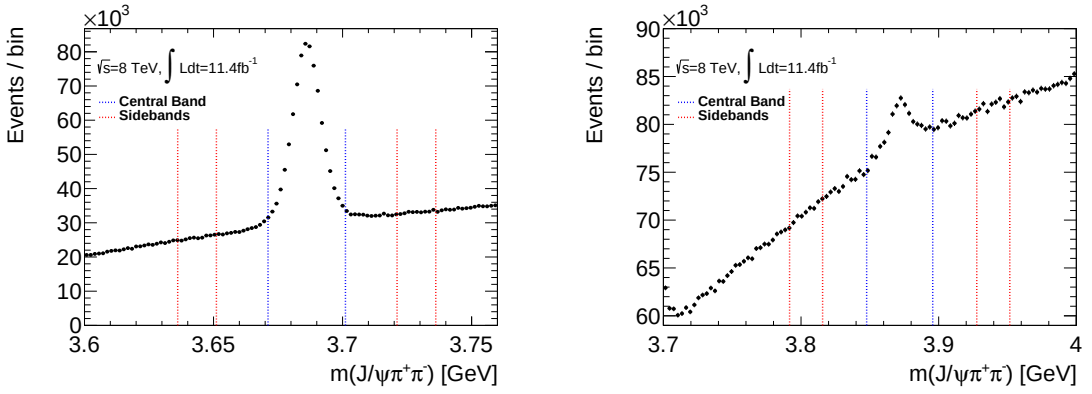
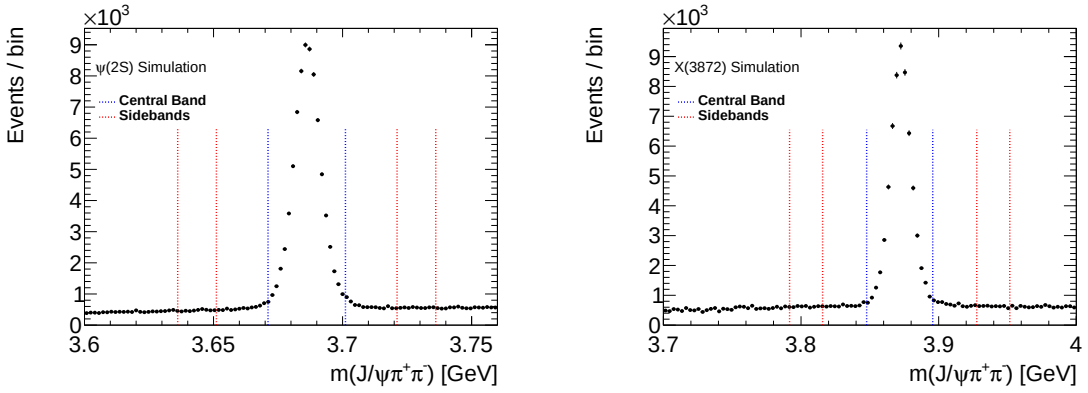


Fig. B.2 Invariant $J/\psi\pi^+\pi^-$ mass distribution in data for (a) $\psi(2S)$ and (b) $X(3872)$ signal regions. Also shown are the central band (blue), and upper/lower sidebands (red).



(a) (b)

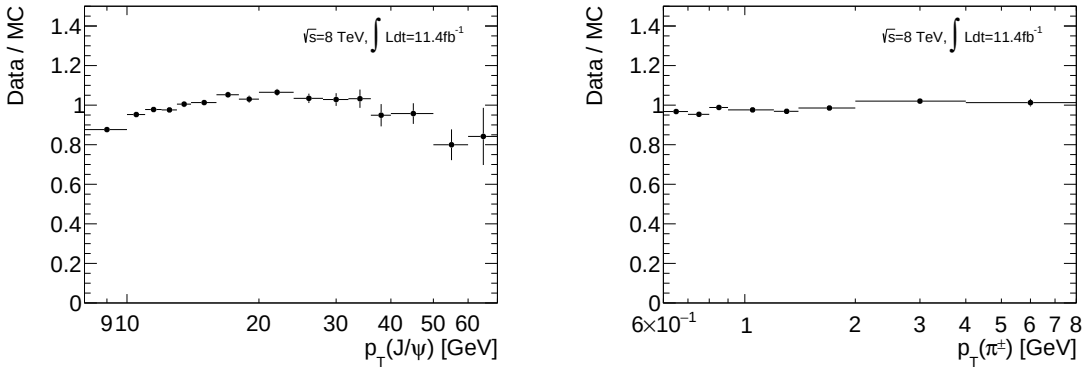
Fig. B.3 Invariant $J/\psi\pi^+\pi^-$ mass distribution in simulation for (a) $\psi(2S)$ and (b) $X(3872)$ signal regions. Also shown are the central band (blue), and upper/lower sidebands (red).

The $p_{\mathrm{T}}(J/\psi)$, $p_{\mathrm{T}}(\pi^{\pm})$ and $\Delta R(J/\psi, \pi^{\pm})$ distributions measured in the sidebands are subtracted from those distributions measured in the central band. The subsequent sideband-subtracted distributions found for data are divided by the ones found for MC to establish data/MC corrections used to reweight the MC on a per-event basis to measure the efficiencies of the selection criteria. The weights are shown in Figure B.4(a) for $p_{\mathrm{T}}(J/\psi)$ and Figure B.4(b) for $p_{\mathrm{T}}(\pi^{\pm})$ for the $\psi(2S)$. The correction weights in those p_{T} distributions are not expected to differ largely for the $\psi(2S)$ and $X(3872)$ signal, so due to the limited statistics of the $X(3872)$ data sample it was decided to use correction weights derived from the $\psi(2S)$ to correct for the $p_{\mathrm{T}}(J/\psi)$ and $p_{\mathrm{T}}(\pi^{\pm})$ of the $X(3872)$.

A larger difference is observed for the $\Delta R(J/\psi, \pi^{\pm})$ correction weights, shown in Figure B.5(a) for $\psi(2S)$ and Figure B.5(b) for $X(3872)$, so here we use each respective correction weight distribution.

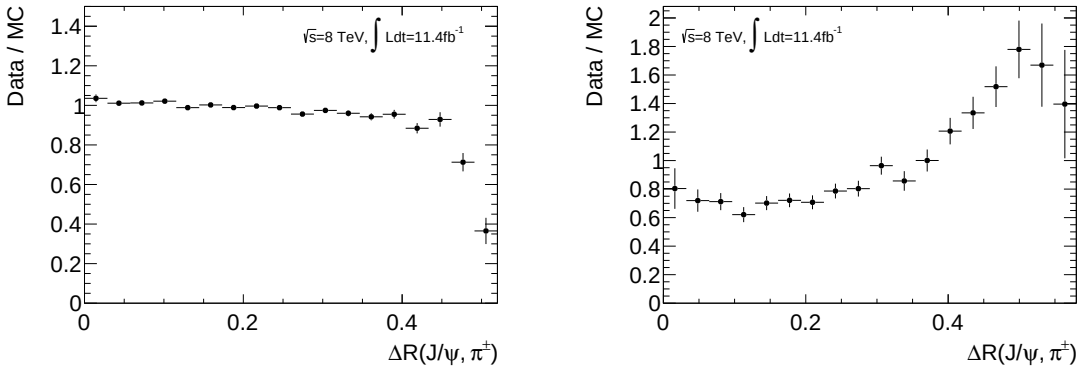
Selection criteria efficiencies are then measured by fitting the $\psi(2S)$ and $X(3872)$ MC signal using the weighted events, examples of which can be seen in Figures B.6 and B.7 in the $p_{\mathrm{T}} = 12 - 16$ GeV bin for $\psi(2S)$ and $X(3872)$, respectively. As an assessment of the systematic error on the reweighting, we then re-perform the fits with the events unweighted by the $p_{\mathrm{T}}(J/\psi)$, $p_{\mathrm{T}}(\pi^{\pm})$ and $\Delta R(J/\psi, \pi^{\pm})$ corrections, with the

difference in each efficiency used as the systematic error. The error is measured to be small (within 1% in most bins).



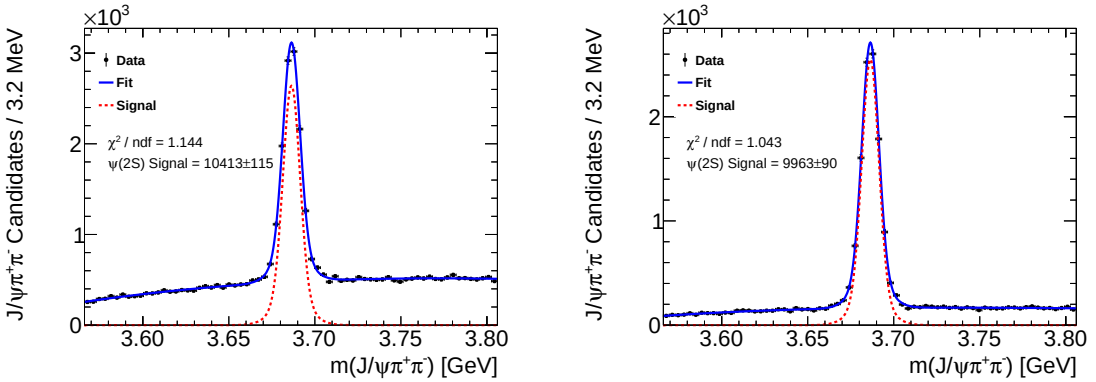
(a) (b)

Fig. B.4 Weights to correct for differences between data and simulation in (a) $p_{\mathrm{T}}(J/\psi)$ and (b) $p_{\mathrm{T}}(\pi^{\pm})$ for the $\psi(2S) \to J/\psi\pi\pi$ signal.



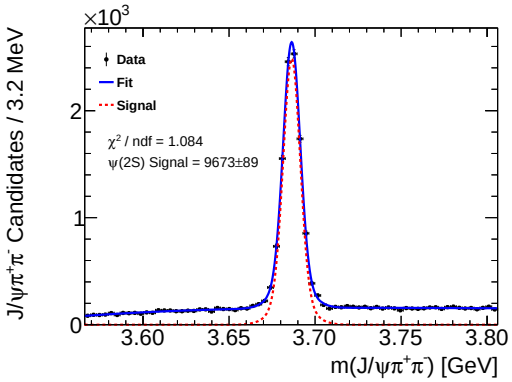
(a) (b)

Fig. B.5 Weights to correct for differences between data and simulation in $\Delta R(J/\psi, \pi^{\pm})$ for (a) $\psi(2S) \to J/\psi\pi\pi$ decays and (b) $X(3872) \to J/\psi\pi\pi$ decays.

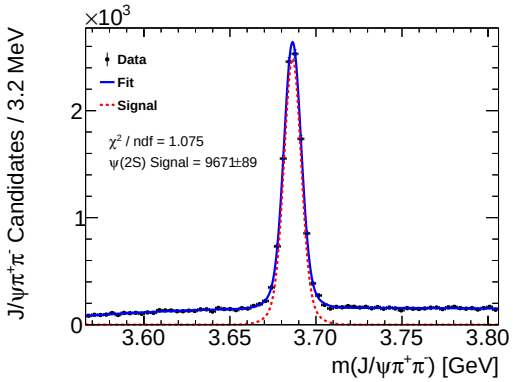


(a)

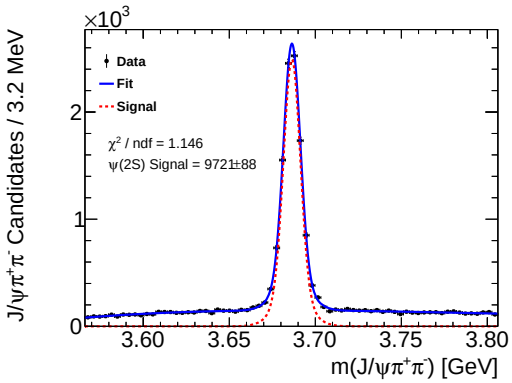
(b)



(c)

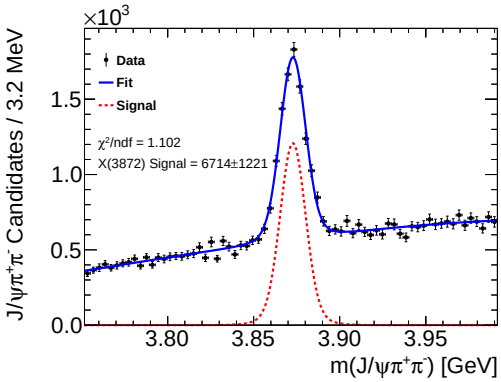


(d)

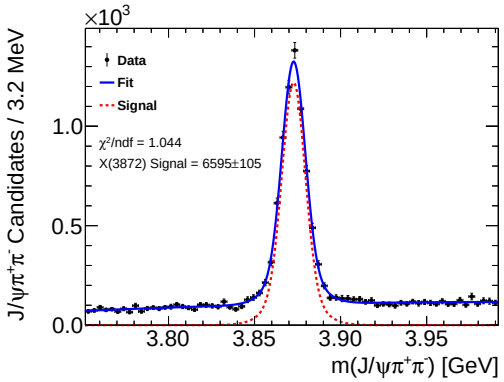


(e)

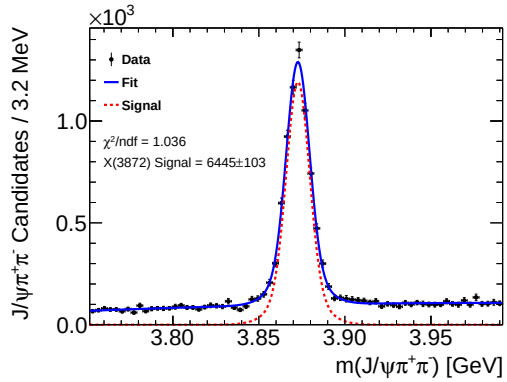
Fig. B.6 $\psi(2S)$ signal yields for selection criteria using MC12 simulation for the p_T bin $[12, 16]$ GeV. (a) The case with only trigger and combined muon requirements. Subsequent selection criteria are cumulatively added in (b) reco quality, (c) $P(\chi^2)(J/\psi\pi\pi)$, (d) $\Delta R(J/\psi, \pi)$ and (e) $Q(J/\psi\pi\pi)$.



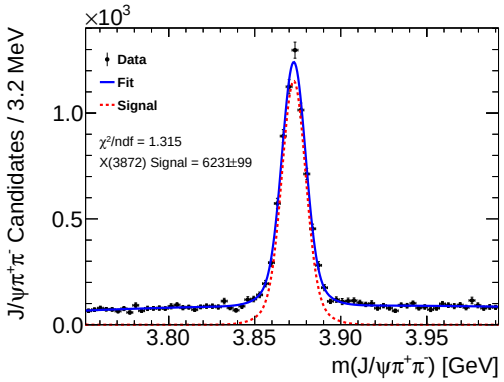
(a)



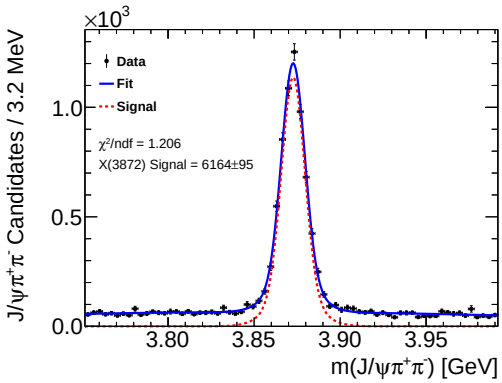
(b)



(c)



(d)

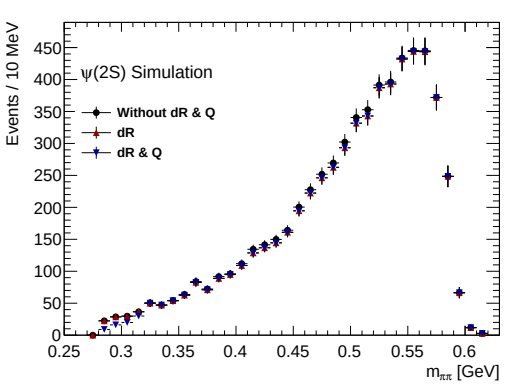


(e)

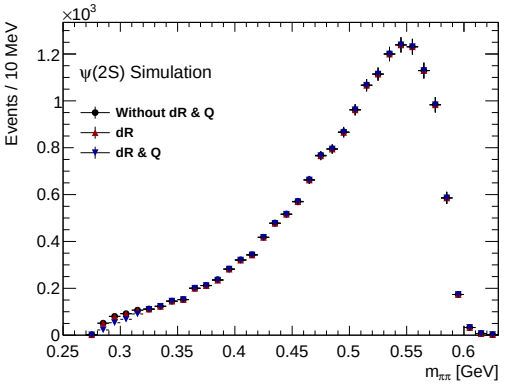
Fig. B.7 $X(3872)$ signal yields for selection criteria using MC12 simulation for the p_T bin $[12, 16]$ GeV. (a) The case with only trigger and combined muon requirements. Subsequent selection criteria are cumulatively added in (b) reco quality, (c) $P(\chi^2)(J/\psi\pi\pi)$, (d) $\Delta R(J/\psi, \pi)$ and (e) $Q(J/\psi\pi\pi)$.

B.3 Effect of Selection Criteria on Dipion Invariant Mass Spectrum

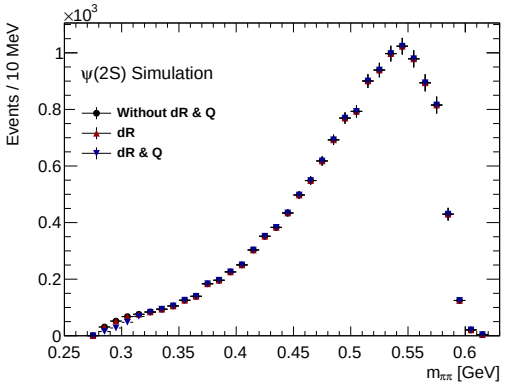
Some of the selection criteria we use to maximise signal efficiency in the main cross-section analysis i.e. the opening angle requirement between the J/ψ and each pion $\Delta R(J/\psi, \pi^{\pm}) < 0.5$, and $Q = m_{J/\psi\pi^+\pi^-} - m^{\mathrm{PDG}}_{J/\psi} - m_{\pi^+\pi^-} < 0.3$ GeV are likely to bias the invariant mass distribution of the dipion system. We investigate this in Figures B.8 and B.9 for $\psi(2S)$ and $X(3872)$ decays, respectively, by plotting the dipion invariant mass distributions with each of the selections cumulatively added. We confirm the selection bias, so omit these selection criteria when measuring dipion invariant mass in Chapter 7.



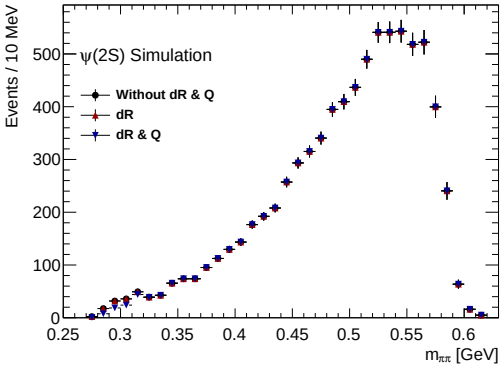
(a)



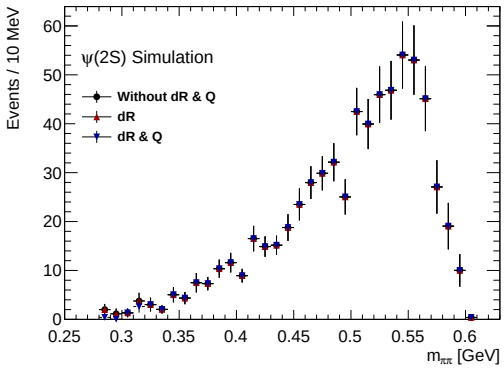
(b)



(c)

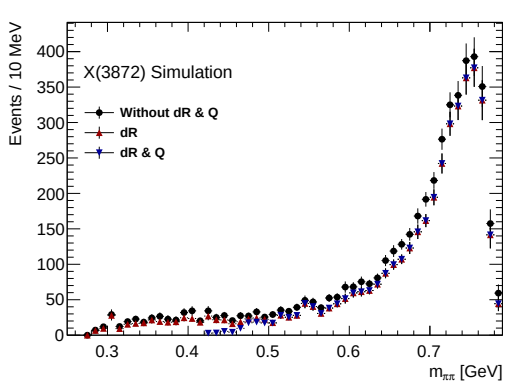


(d)

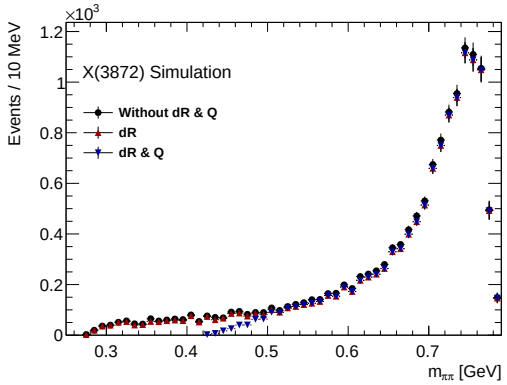


(e)

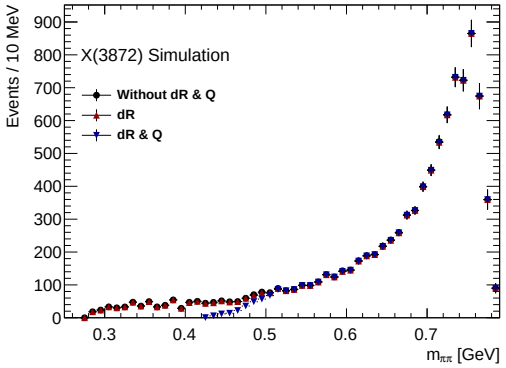
Fig. B.8 Cumulative effect of selection criteria on $m_{\pi\pi}$ in $\psi(2S) \to J/\psi\pi\pi$ simulation for (a) $10 < p_T(J/\psi\pi\pi) < 12$ GeV, (b) $12 < p_T(J/\psi\pi\pi) < 16$ GeV, (c) $16 < p_T(J/\psi\pi\pi) < 22$ GeV, (d) $22 < p_T(J/\psi\pi\pi) < 40$ GeV and (e) $40 < p_T(J/\psi\pi\pi) < 70$ GeV.



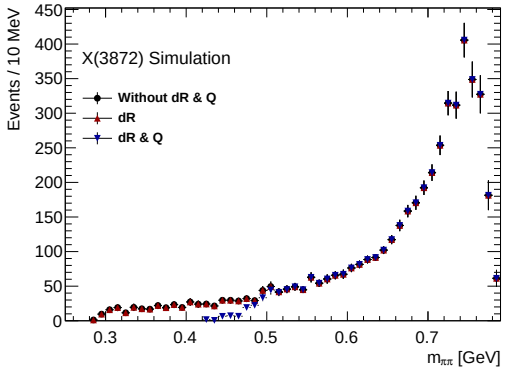
(a)



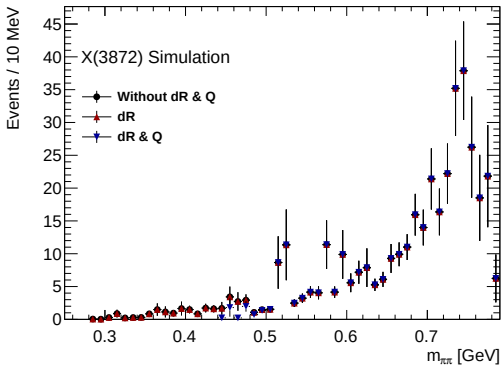
(b)



(c)



(d)



(e)

Fig. B.9 Cumulative effect of selection criteria on $m_{\pi\pi}$ in $X(3872) \to J/\psi\pi\pi$ simulation for (a) $10 < p_T(J/\psi\pi\pi) < 12$ GeV, (b) $12 < p_T(J/\psi\pi\pi) < 16$ GeV, (c) $16 < p_T(J/\psi\pi\pi) < 22$ GeV, (d) $22 < p_T(J/\psi\pi\pi) < 40$ GeV and (e) $40 < p_T(J/\psi\pi\pi) < 70$ GeV.

B.4 Bin Migration Effects

We use a $\psi(2S) \to J/\psi\pi^+\pi^-$ simulation sample to measure the effect that the limited resolution of event reconstruction has on measured signal yields. We initally measure reconstructed $\psi(2S)$ yields in p_{T} bins with all the selection requirements used in the analysis, described in Table 5.1. The results of this study are summarised in Table B.2, where the column entitled 'inital' contains the yields corresponding to the selection criteria used in the main analysis.

To measure the effect of reconstructed muon p_T smearing for the requirement of both muons satisfying $p_T > 4$ GeV, we instead require the truth muons in the event to satisfy $p_T > 4$ GeV. These numbers are presented in the column entitled 'truth muon'. The difference in the signal yields between using the reconstructed or truth muons for the minimum p_T requirement allows us to quantify the effect of the smearing on the selection requirement. The same method is applied for the minimum $p_T(\pi^{\pm})$ requirement - we require the truth pions belonging to the $\psi(2S) \to J/\psi\pi^+\pi^-$ event to have $p_T(\pi^{\pm}) > 600$ MeV instead of the reconstructed pions, with signal yields shown in the column entitled 'truth pion'.

To estimate the effect of bin migration on our analysis, we measure the signal in bins of *truth* $\psi(2S)$ p_T and rapidity, with the number shown in the columns entitled 'truth $\psi(2S)$ p_T' and 'truth $\psi(2S)$ y' . The measured signal yields for each case are shown in Table B.2.

We can conclude from this that the difference in measured yields due to the finite resolution of muon and pion p_T reconstruction in our selection requirements does not exceed 1% in each of the p_T bins. We can make a similar conclusion for the bin migration effect of reconstructed $J/\psi\pi^+\pi^-$ candidates - using the truth p_{T} and rapidity for the analysis bins instead of the reconstructed variables causes a small difference in yields of $< 1\%$, and $< 3\%$ in the highest p_T bin. Similar small bin migration effects were measured in previous quarkonia cross section analyses, and are considered negligible due to their small effect.

	Measured Signal Yields				
$p_T(J/\psi\pi\pi)$ [GeV]	Initial	truth muon	truth pion	truth $\psi(2S)$ p_T	truth $\psi(2S)$ y
$10-12$	3170 ± 65	3150 ± 65	3134 ± 65	3181 ± 65	3168 ± 65
$12-16$	9093 ± 111	9049 ± 111	9036 ± 112	9062 ± 110	9085 ± 111
$16-22$	7251 ± 99	7248 ± 98	7229 ± 99	7215 ± 99	7252 ± 98
$22-40$	3735 ± 71	3720 ± 71	3709 ± 72	3759 ± 71	3731 ± 71
$40-70$	232 ± 19	233 ± 19	238 ± 19	240 ± 19	232 ± 19

Table B.2 Measured signal yields when using truth muon, pion and $\psi(2S)$ variables for selection and bin requirements instead of reconstructed values for simulated data, to determine bin migration effects.

Appendix C

Pion Reconstruction Efficiency

C.1 Pile-up Conditions

We measure the pion reconstruction efficiency using a 2012 $J/\psi \to \mu\mu$ MC sample. The pion reconstruction efficiency depends on the pile-up conditions, which are different for simulation and real data. We investigate the difference in reconstruction efficiency for four distinct cases, where the number of primary vertices in the reconstructed events are: $5-10$, $10-15$, $15-20$, $20-25$. The distribution of number of primary vertices, $N(PV)$, for the MC sample is shown in Figure C.1.

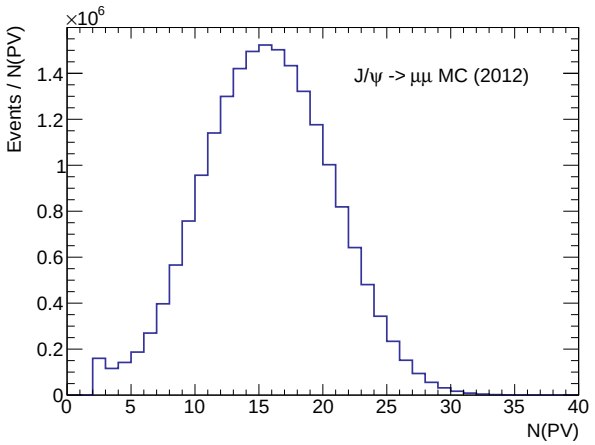


Fig. C.1 Distribution of number of reconstructed primary vertices for the 2012 $J/\psi \to \mu\mu$ MC samples.

Figure C.2 shows the reconstruction efficiency map for the lowest $N(PV)$ bin and the difference in efficiency maps for the higher $N(PV)$ bins. We observe the differences in efficiency increase gradually with increase in pile-up, and find maximal differences of around $\pm 2\%$ around the edges of the $(p_T, q \times \eta)$ space, with agreement more centrally within 1%.

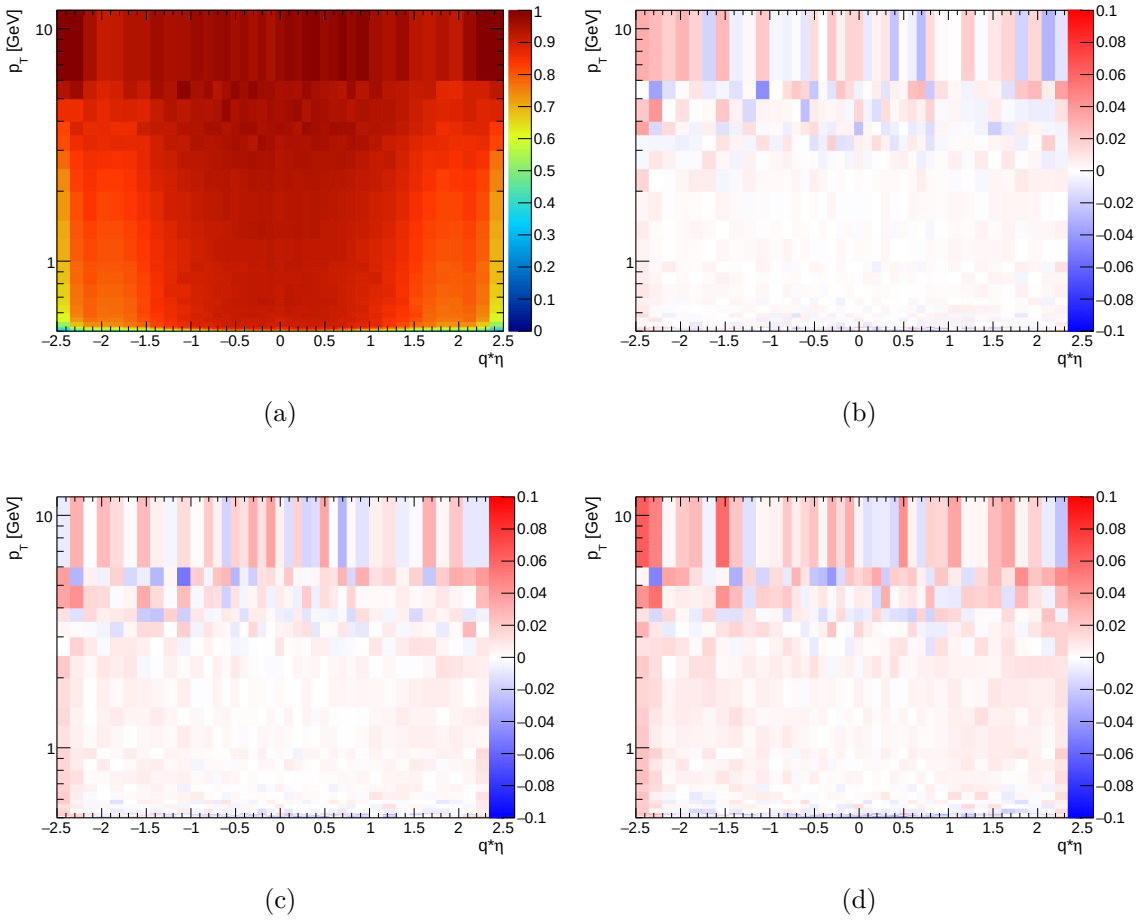
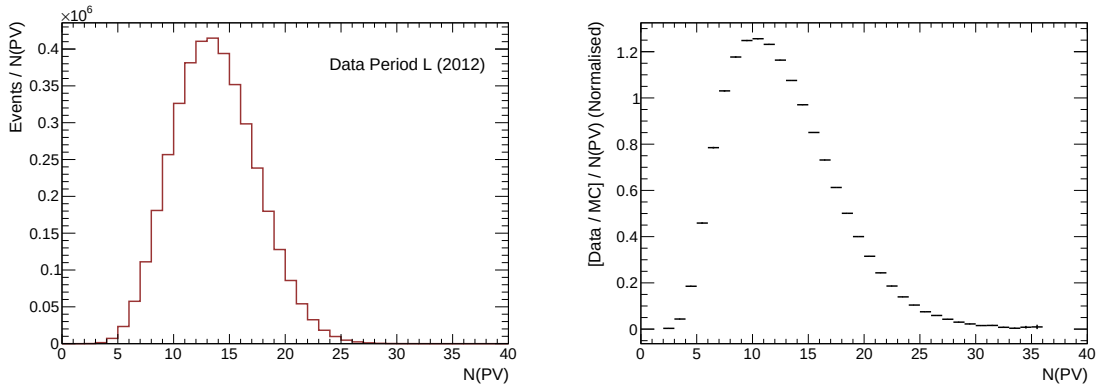


Fig. C.2 (a) Pion reconstruction efficiency map for MC where the number of reconstructed primary vertices is between 5 and 10. Also shown are the differences between this map and the case where $N(PV)$ is between (b) $10-15$, (c) $15-20$, (d) $20-25$.

We now compare the $N(PV)$ distribution of the MC sample with the real data (specifically of portion of Period L 2012 data, which is assumed to be representative of the full 2012 data). We show the $N(PV)$ distribution of the data in Figure C.3(a). The MC can be reweighted according to the data distribution to create a new reconstruction

efficiency map which is more representative of real run conditions. The per-event weights applied to the MC are taken from the normalised ratio between the data and MC distributions, shown in Figure C.3(b). The difference between the central unweighted map and the reweighted map are shown in Figure C.4. We see very small differences ($<< 1\%$) between the reweighted efficiency map and the central efficiency map.



(a) (b)

Fig. C.3 (a) Distribution of number of reconstructed primary vertices for a portion of 2012 Period L data and (b) the normalised ratio between the data and MC distribution.

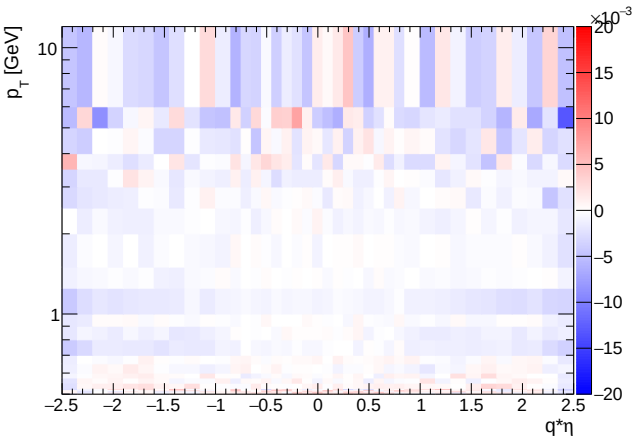
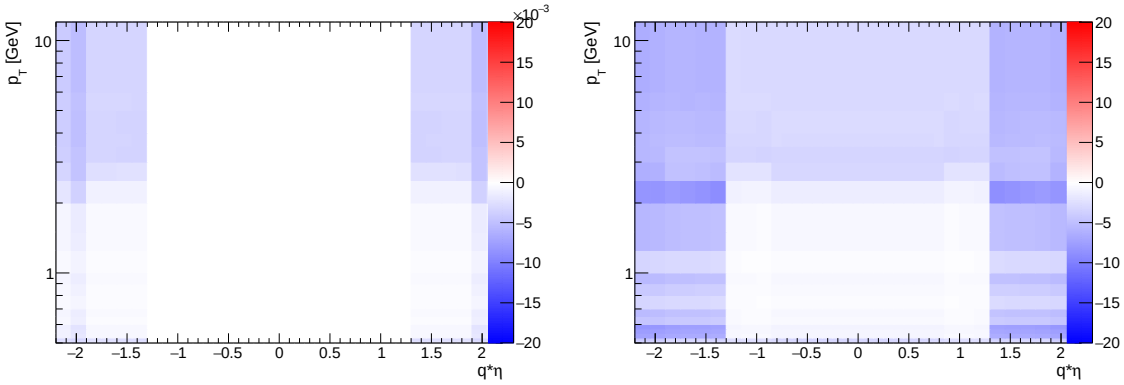


Fig. C.4 Difference between the unweighted pion reconstruction efficiency map and the case where events have been reweighted to match 2012 pile-up conditions.

C.2 Material Uncertainty

We note the geometry tag used for the pion reconstruction efficiency measurement using simulation as `ATLAS-GEO-20-00-01`, which is not the current 'best' description of the detector (which is `ATLAS-GEO-21-01-00`). Studies performed in [108] confirm that the difference in efficiency between `ATLAS-GEO-18-01-03` and the best description is very small ($< 1\%$) and their other studies found the efficiency with our tag is in agreement with `ATLAS-GEO-18-01-03` within statistical error (primarily due to an identical description of the ID for these two tags). We have reweighted the map obtained from our MC sample to match the 'best' geometry tag, with the difference in efficiency shown in Figure C.5(a). The difference in efficiency is seen to be $<< 1\%$. To measure the additional systematic error due to material uncertainty, we create a new map in Figure C.5(b) by finding the difference between the efficiency maps for the 'best' tag `ATLAS-GEO-21-01-00` and the tag `ATLAS-GEO-21-06-01`, where there is 15% more inner detector material. This is obtained using the approriate scale factors from [108]. The difference is seen to be small, limited by a value of 0.35%.



(a) (b)

Fig. C.5 (a) Difference in pion reconstruction efficiency maps between tags `ATLAS-GEO-20-00-01` and `ATLAS-GEO-21-01-00`, and (b) difference in pion reconstruction efficiency maps between tags `ATLAS-GEO-21-01-00` and `ATLAS-GEO-21-06-01`.

C.3 Opening Angle Selection

We decide whether a truth pion has been successfully matched to a reconstructed track if the opening angle between the pion and its closest reconstructed track is less than 0.05. We see in Figure C.6 that this selection is conservative, and that the large majority of pion matches are at much smaller opening angles. However, the amount of flat background, where the truth pion is matched with the wrong reconstructed track, is very small compared to the number of correct matches. We investigate the effect of the maximum opening angle by recreating the efficiency map with a tighter opening angle of 0.03. The difference between the efficiency maps using $\Delta R < 0.05$ and $\Delta R < 0.03$ is shown in Figure C.7. We observe a decrease in the measured efficiency of $< 1\%$ for the central η region, rising to $2 - 3\%$ for larger η.

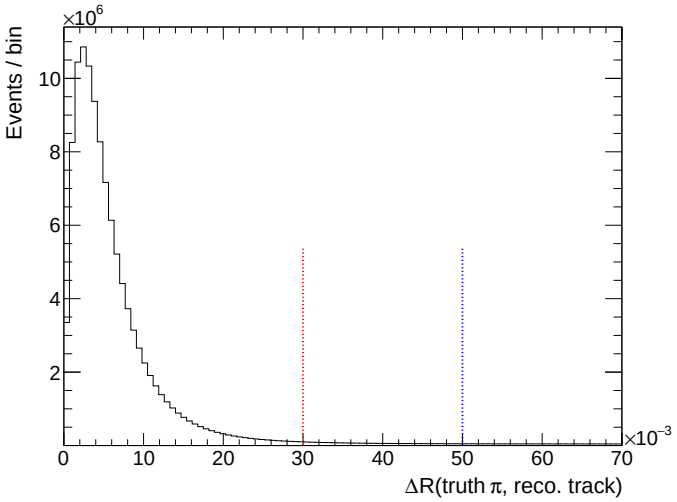


Fig. C.6 Distribution of opening angles between the truth pions and closest reconstructed track in the event for simulation. The blue line shows the maximum ΔR selection for the central efficiency method, and the red line shows a systematic change to $\Delta R < 0.03$.

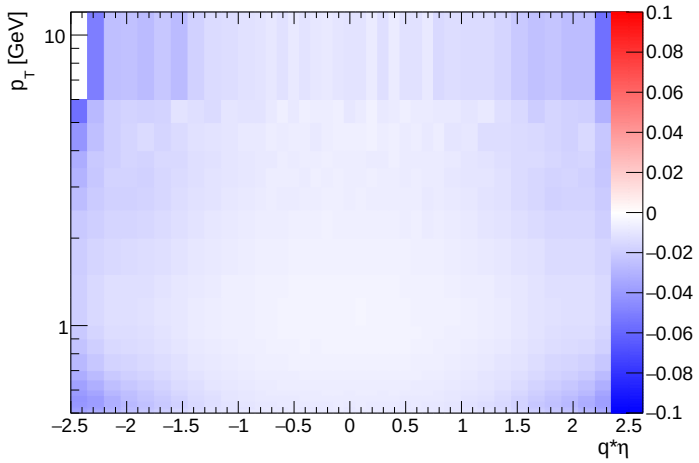


Fig. C.7 Difference in pion reconstruction efficiency for $\Delta R < 0.05$ and $\Delta R < 0.03$ between the truth pion and closest reconstructed track.

C.4 Distribution of Signal Pions

We see the η distribution of the pions associated with $\psi(2S) \to J/\psi\pi^+\pi^-$ and $X(3872) \to J/\psi\pi^+\pi^-$ decays, where $|y(J/\psi\pi^+\pi^-)| < 0.75$ and for our selection criteria, are contained within the range $\eta(\pi^\pm) < 1.5$, as seen in Figure C.8.

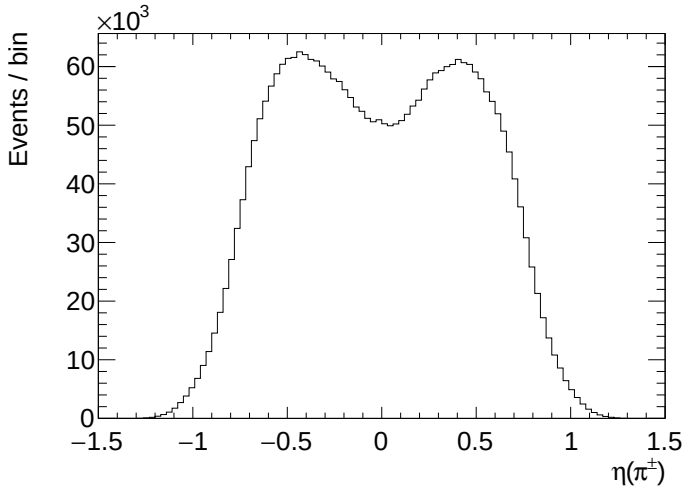
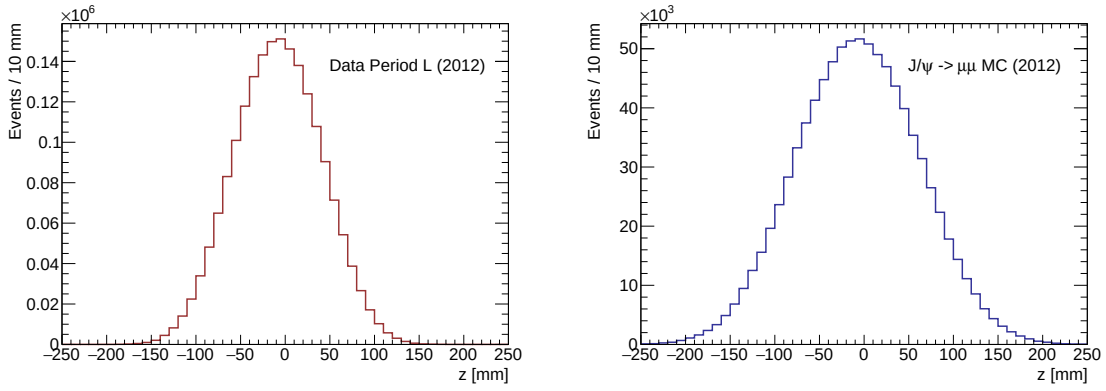


Fig. C.8 Distribution in η of signal pions associated with reconstructed $\psi(2S) \to J/\psi\pi^+\pi^-$ and $X(3872) \to J/\psi\pi^+\pi^-$ decays.

C.5 z-Displacement of Primary Vertices

The z-displacement distribution of primary vertices is measured to have a difference in data and simulation, seen in Figure C.9. We investigate the effect this has on pion reconstruction efficiency determination by reweighting the Monte Carlo simulation according to a [data/MC] scale factor as a function of z-displacement of primary vertex. The difference between the central efficiency map and the z-displacement reweighted map is shown in Figure C.10. We observe only small differences of $< 0.5\%$ between the two maps, which is used as an additional systematic.



(a) (b)

Fig. C.9 z-displacement distribution of primary vertices for (a) data and (b) simulation.

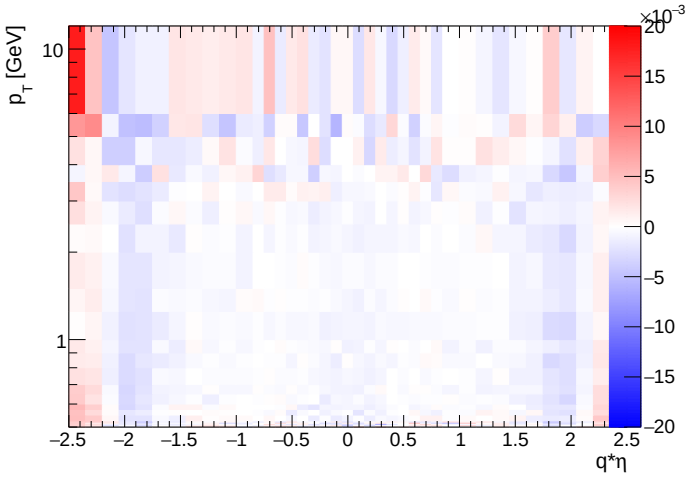


Fig. C.10 Difference between the central pion reconstruction efficiency map and the case where simulation has been weighted to match the data z-displacement of primary vertices. We observe only small differences $< 0.5\%$.

C.6 Total Pion Reconstruction Uncertainty

The sources of the systematic errors on the pion reconstruction efficiency corrections include the following:

1. Corrections of the MC due to pile-up differences (Section C.1).
2. Material uncertainty in the inner detector (Section C.2).
3. Uncertainty on truth matching (Section C.3).
4. Corrections of the MC due to z-displacement differences (Section C.5).
5. Statistical error of the Monte Carlo (see Figure C.11).

As seen from the respective efficiency maps (Figures C.4, C.5(b), C.7, C.10 and C.11), the variations from central are generally small. The uncertainties for each of the above effects are found to be within for the following limits: $(0.5\%, 0.4\%, 1.0\%, 0.3\%, 0.1\%)$, respectively. The overall systematic uncertainty linked to the pion efficiency is then calculated as the quadratic sum of these numbers, equal to 1.25% per pion. Since the pions are likely be close in the detector fiducial volume, a systematic uncertainty of $2 \times 1.25\% = 2.5\%$ was assigned for the dipion system.

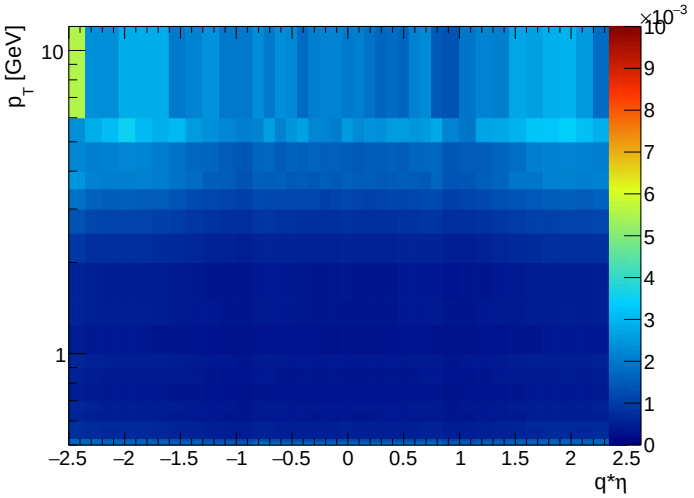


Fig. C.11 Statistical error of the pion efficiency map used in the main analysis. Each bin contains only a small ($<< 1\%$) error due to the high statistics of the MC sample used to measure the efficiency.

Appendix D

Spin Alignment Studies

In this Appendix, we investigate the dependence of the acceptance on different polarisation scenarios, separately for $\psi(2S)$ and $X(3872)$ decays. We plot the $\cos\theta^*$ and ϕ^* distributions of the daughter J/ψ and dipion system for $\psi(2S)$ and $X(3872)$ generator level simulated decays to $J/\psi\pi^+\pi^-$ for each polarisation scenario, where the momentum of $\psi(2S)(X(3872))$ is used as the polarisation axis. We fix the $\psi(2S)(X(3872))$ $p_T = 12$ GeV, $y = 0.05$ and $m_{\pi\pi} = 0.435$ GeV, plotting the distributions after the acceptance criteria are applied.

For the $\psi(2S) \to J/\psi\pi^+\pi^-$ decays (Figures D.1 to D.4), we see the daughter J/ψ inherits the polarisation of the $\psi(2S)$. The dipion system is unpolarised regardless of the polarisation of the mother $\psi(2S)$. However, in $X(3872) \to J/\psi\rho^0(\to \pi^+\pi^-)$ decays (Figures D.5 to D.8) we see that the daughter J/ψ in every case has a polarisation different to that of the $X(3872)$. For example, when $X(3872)$ is longitudinally polarised, the J/ψ is then transversely polarised. The dipion system inherits the same polarisation as the $X(3872)$. The acceptance criteria are harsher for the J/ψ than for the dipion system, and so the acceptance of the $X(3872)$ is dominated by the daughter J/ψ polarisation. This explains why the dependence of acceptance on polarisation is different for $\psi(2S)$ and $X(3872)$.

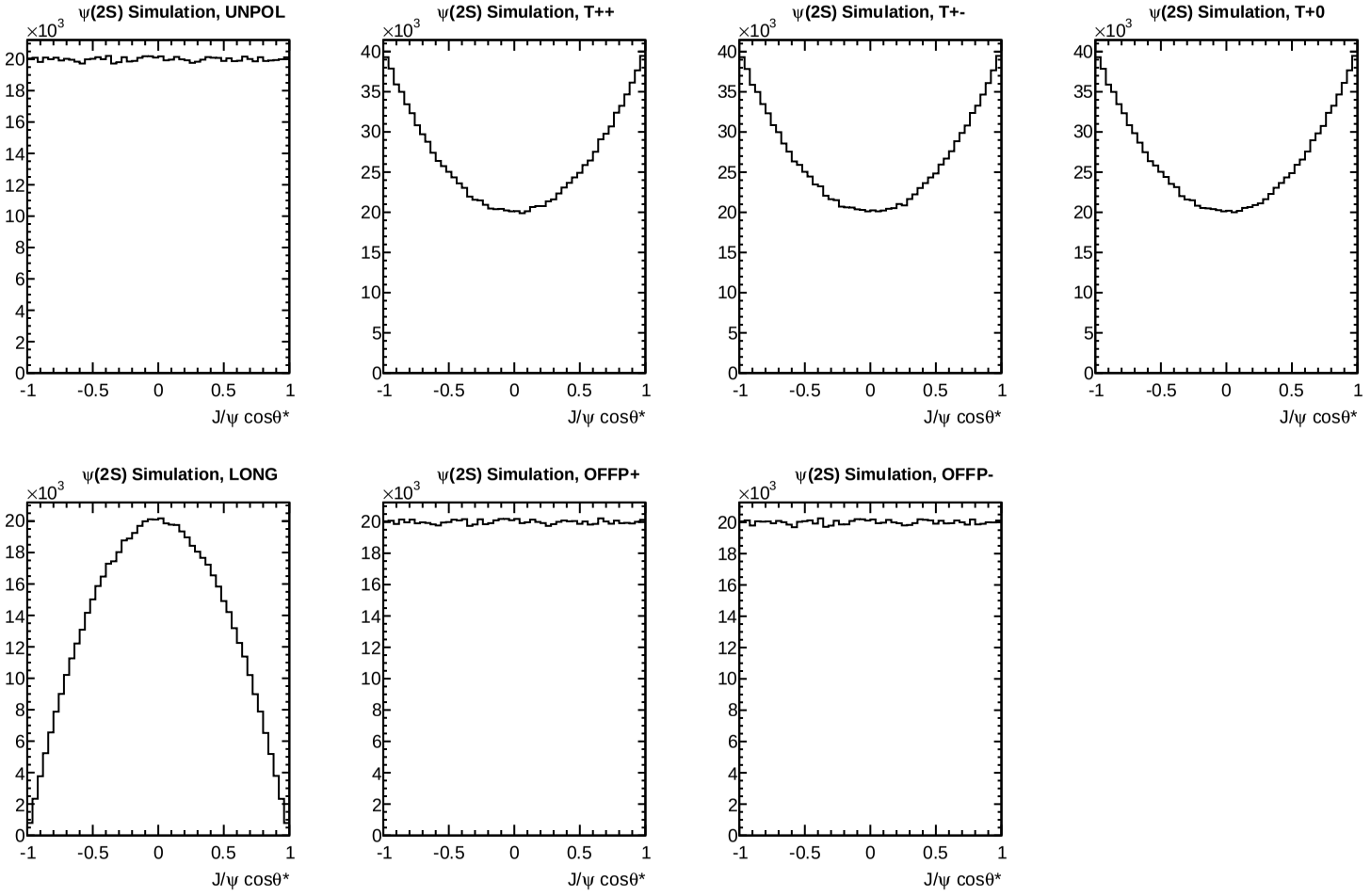


Fig. D.1 $\cos\theta^*$ distributions for J/ψ produced in $\psi(2S) \to J/\psi\pi^+\pi^-$ decays for various polarisation scenarios

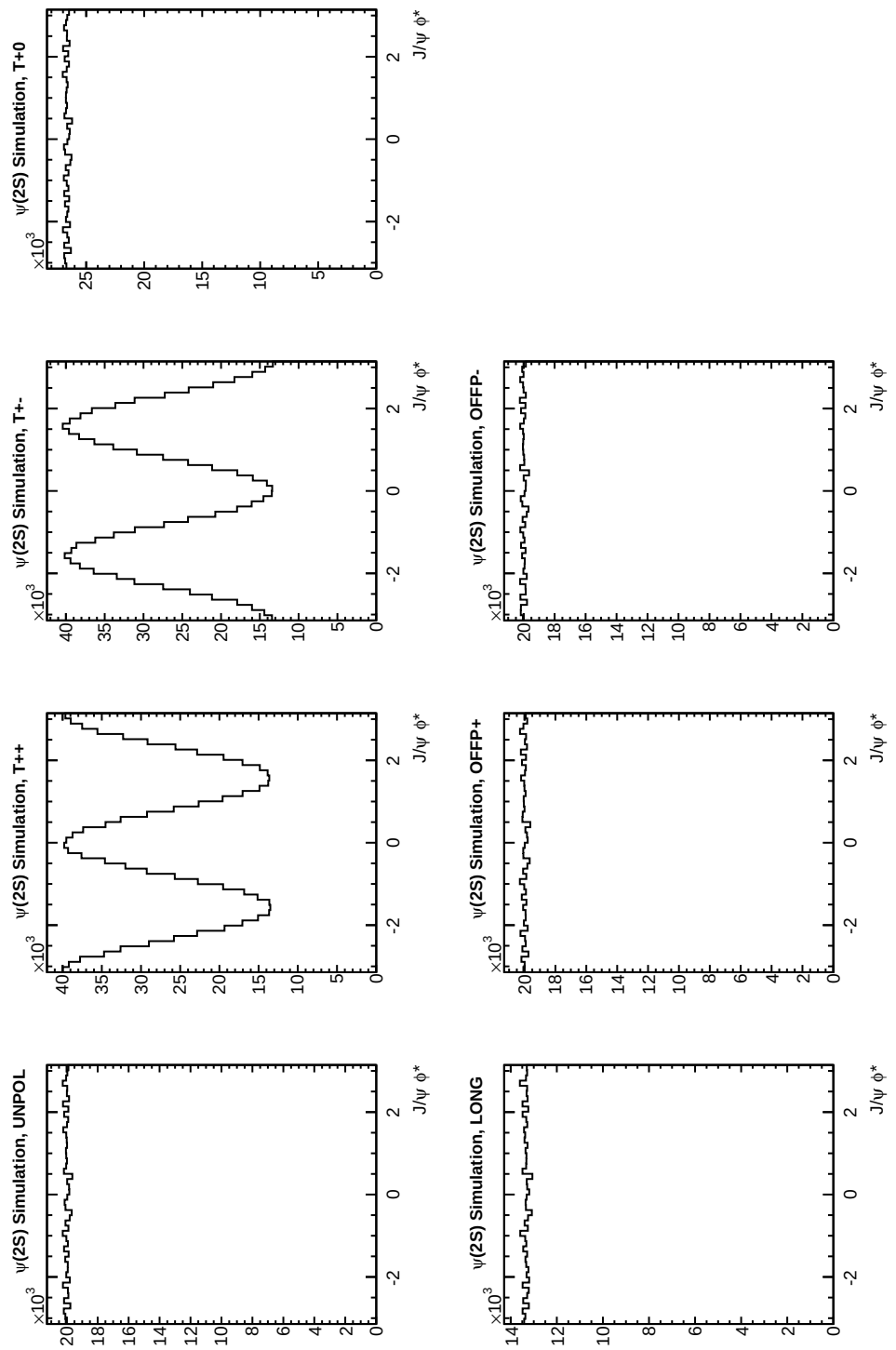


Fig. D.2 ϕ^* distributions for J/ψ produced in $\psi(2S) \to J/\psi\pi^+\pi^-$ decays for various polarisation scenarios

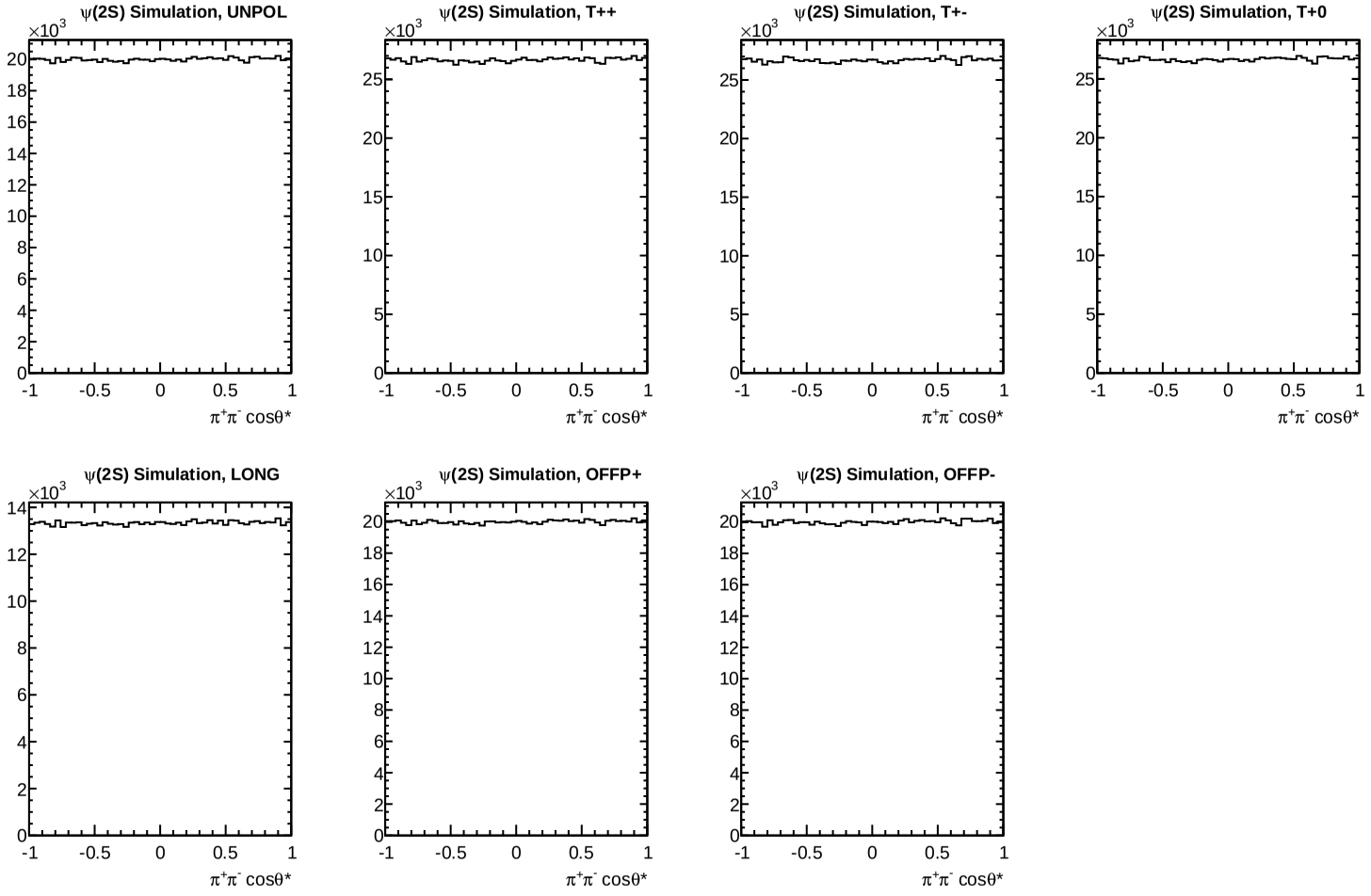


Fig. D.3 $\cos\theta^*$ distributions for dipions produced in $\psi(2S) \to J/\psi\pi^+\pi^-$ decays for various polarisation scenarios

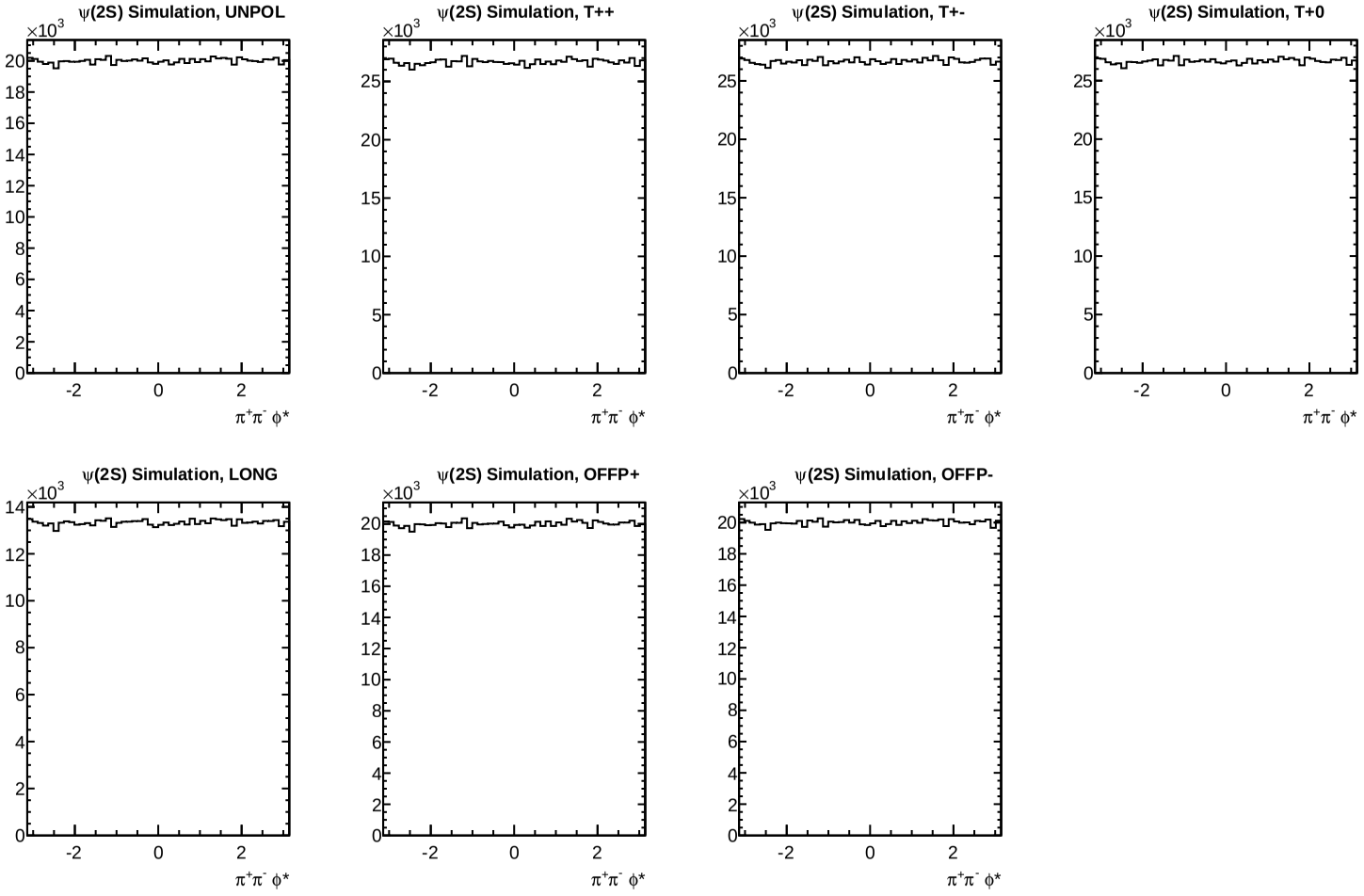


Fig. D.4 ϕ^* distributions for dipions produced in $\psi(2S) \to J/\psi\pi^+\pi^-$ decays for various polarisation scenarios

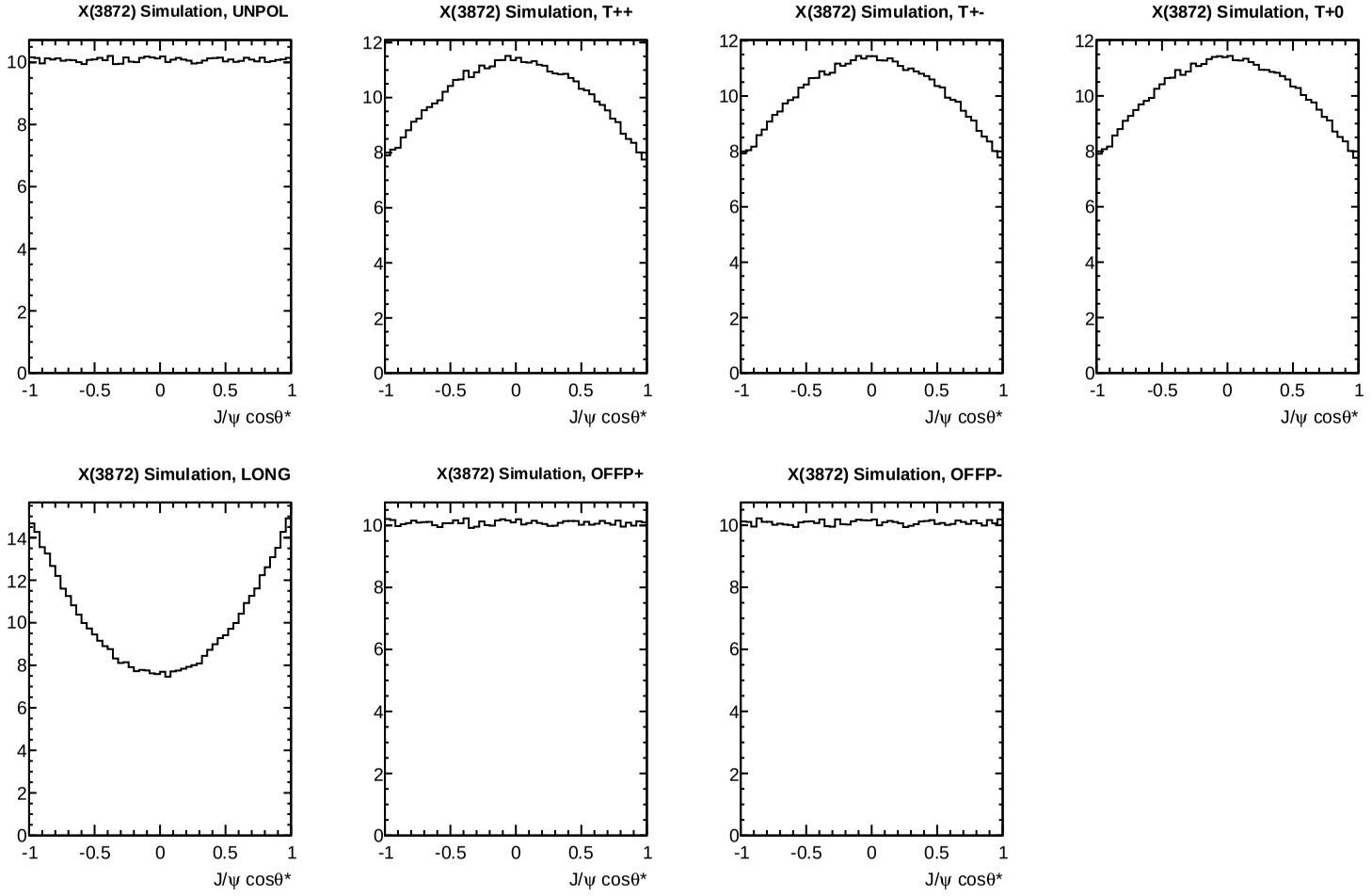


Fig. D.5 $\cos\theta^*$ distributions for J/ψ produced in $X(3872) \to J/\psi\rho^0_{\to\pi^+\pi^-}$ decays for various polarisation scenarios

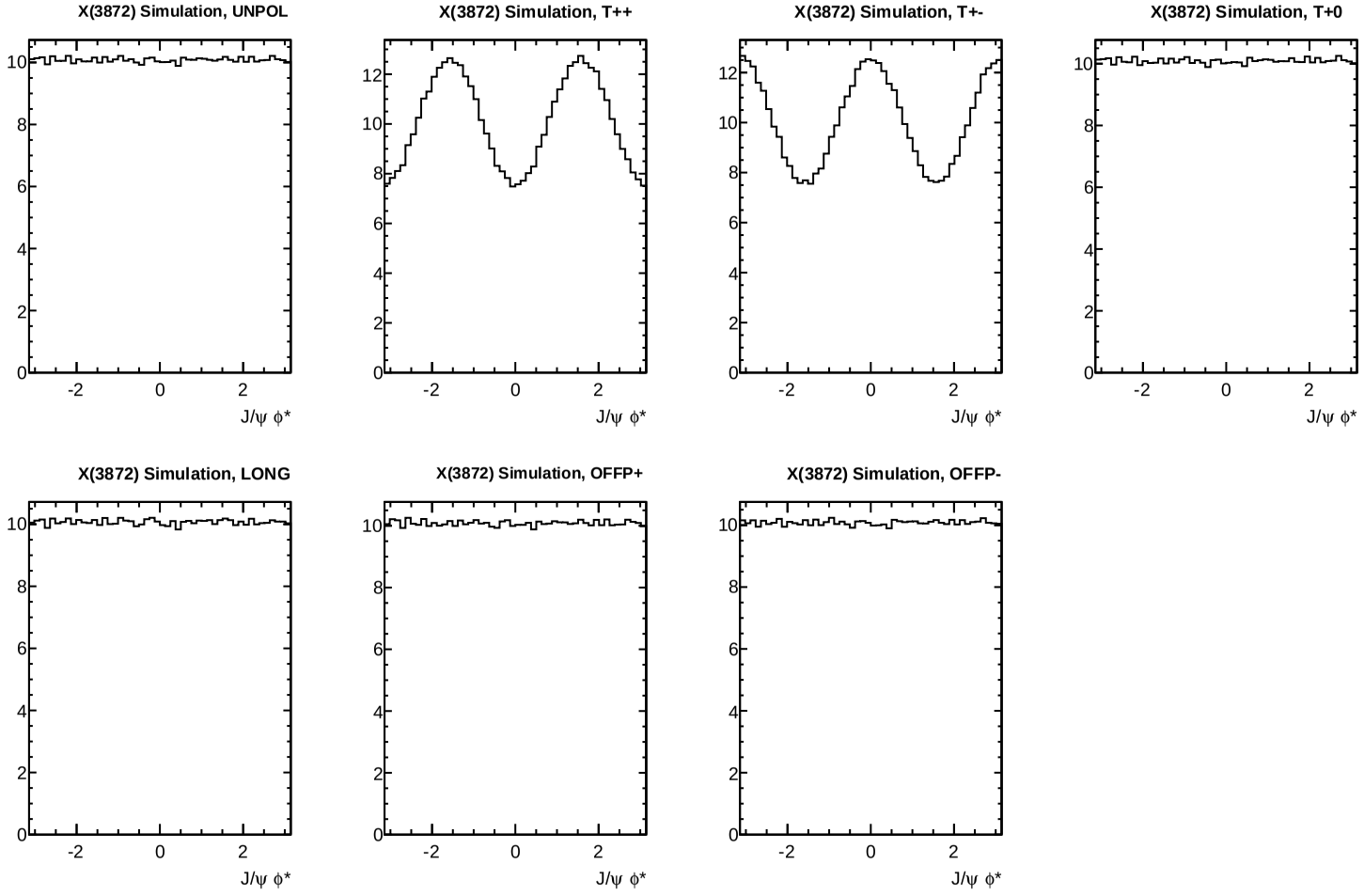


Fig. D.6 ϕ^* distributions for J/ψ produced in $X(3872) \to J/\psi\rho^0_{\to\pi^+\pi^-}$ decays for various polarisation scenarios

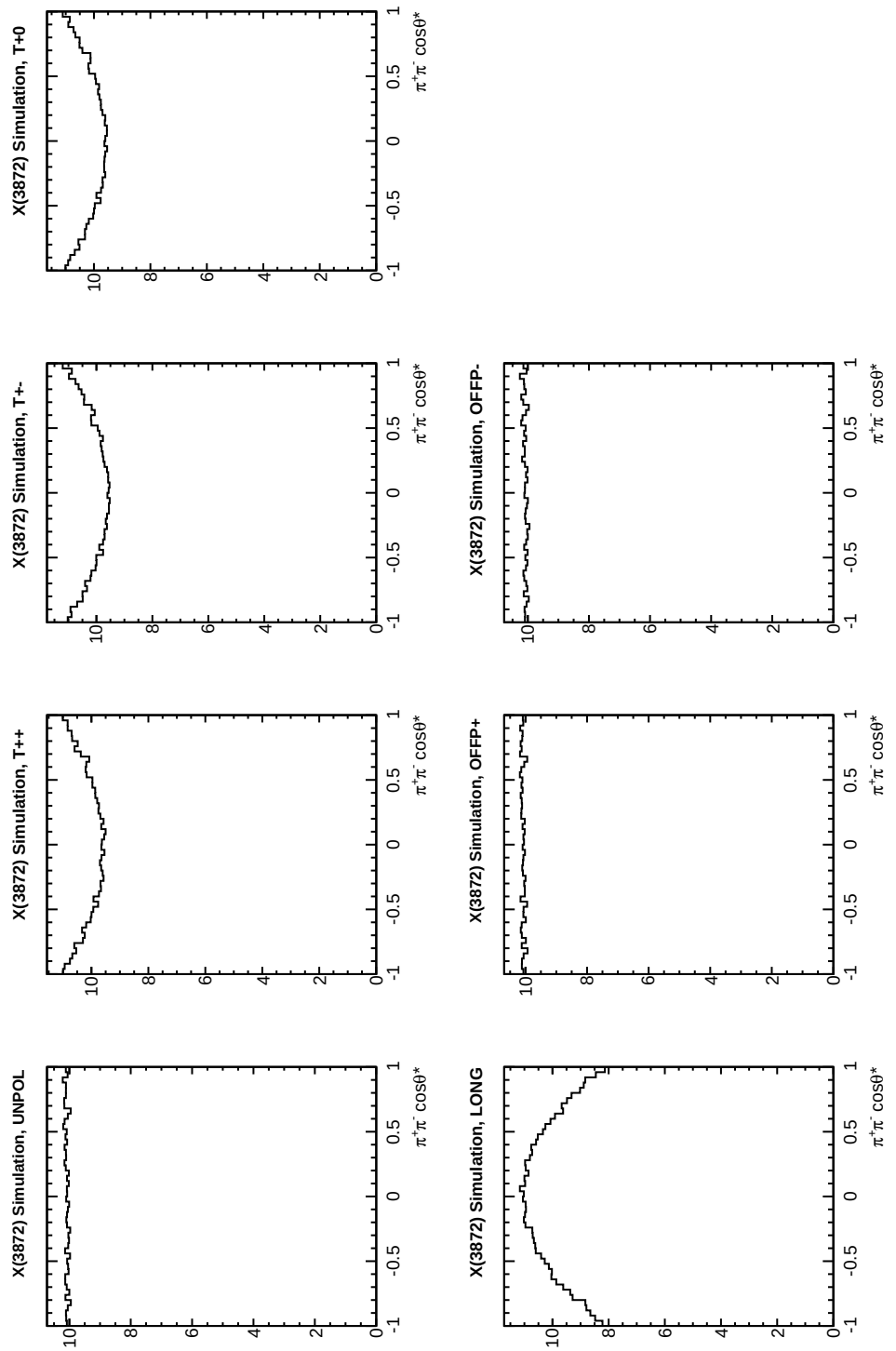


Fig. D.7 $\cos\theta^*$ distributions for dipions produced in $X(3872) \to J/\psi\rho^0_{\to\pi^+\pi^-}$ decays for various polarisation scenarios

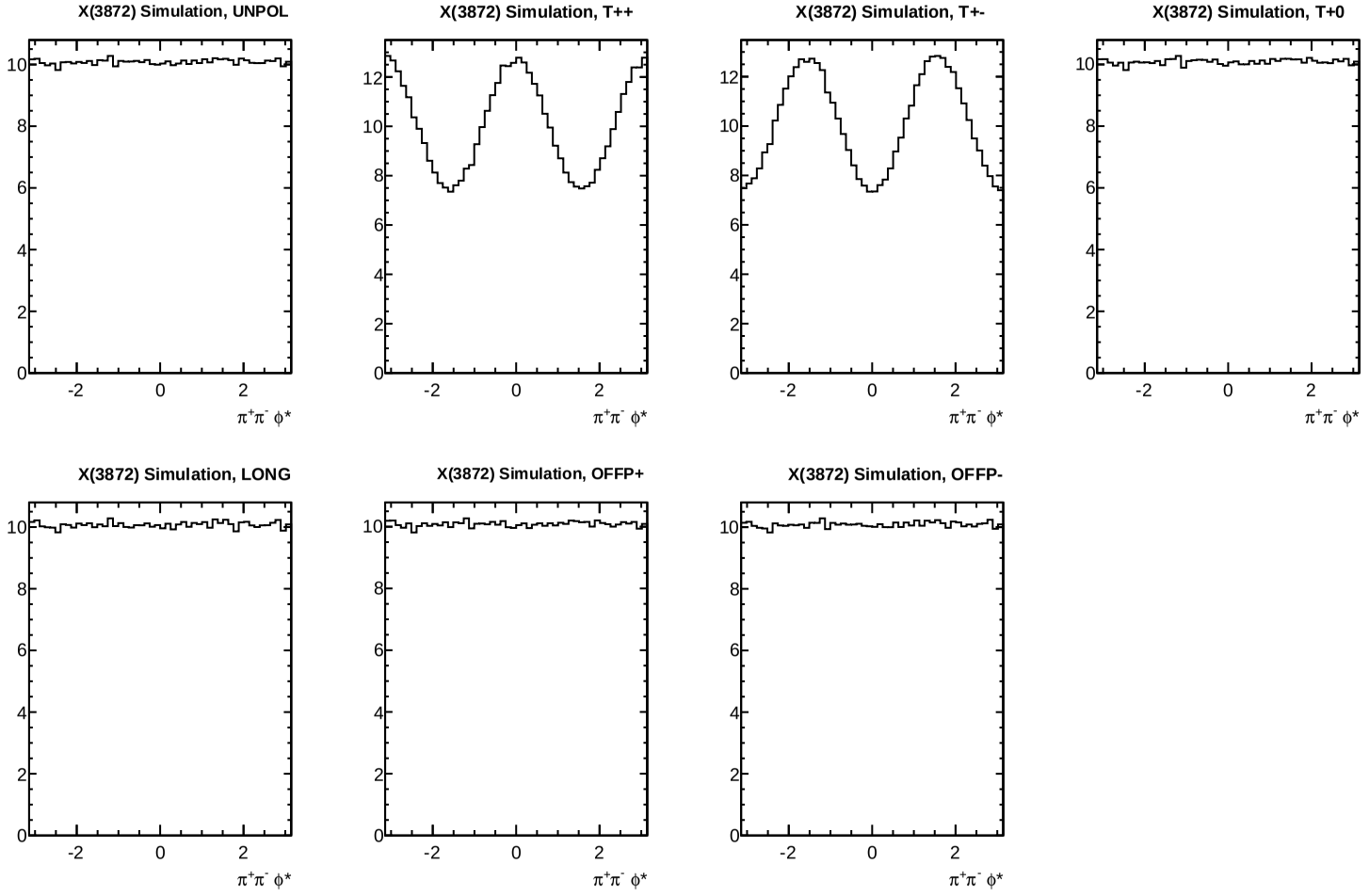


Fig. D.8 ϕ^* distributions for dipions produced in $X(3872) \to J/\psi\rho^0_{\to\pi^+\pi^-}$ decays for various polarisation scenarios

Appendix E

Lifetime Modelling

E.1 Lifetime Resolution Determination

We determine the lifetime resolution function using two-dimensional mass-lifetime unbinned maximum likelihood fits in the $\psi(2S)$ signal region, separately for each of the five p_{T} bins. The signal PDF is defined as:

$$\mathrm{PDF}_{\mathrm{sig}}(m,\tau) = f_{\mathrm{sig}}^{P}\left[M_{\psi}^{P}(m)\cdot T_{\psi}^{P}(\tau)\right] + (1-f_{\mathrm{sig}}^{P})\left[M_{\psi}^{N}(m)\cdot T_{\psi}^{N}(\tau)\right], \tag{E.1}$$

where f_{sig}^{P} is the fraction of the signal which is promptly produced. The PDFs $M_{\psi}^{P}(m), M_{\psi}^{N}(m)$ model the mass distributions of promptly and non-promptly produced $\psi(2S)$, and $T_{\psi}^{P}(\tau), T_{\psi}^{N}(\tau)$ model their respective lifetime distributions. $M_{\psi}^{P}(m), M_{\psi}^{N}(m)$ are modelled as double Gaussians where the wider Gaussian G_2 has twice the width of the narrower Gaussian G_1: $f_1 G_1 + (1-f_1)G_2$. The prompt lifetime signal component $T_{\psi}^{P}(\tau)$ is modelled by the resolution function, defined here to be a triple Gaussian, where the widths of the two wider Gaussians are fixed to be 2 times and 4 times the width of the narrowest. A double Gaussian is first constructed with the widest G_3^{τ} and intermediate G_2^{τ} Gaussians, as $G_{23}^{\tau} = f_3^{\tau}G_3^{\tau} + (1-f_3^{\tau})G_2^{\tau}$. We then construct the triple Gaussian to include the narrowest gaussian G_1 as: $T_{\psi}^{P}(\tau) = f_{23}^{\tau}G_{23}^{\tau} + (1-f_{23}^{\tau})G_1^{\tau}$.

As an example, for the $p_{\mathrm{T}} = 12-16$ GeV bin, the double Gauss constructed from the wider two Gaussians contains a fraction $f_3^{\tau} = 0.29 \pm 0.03$ of the widest Gauss. When the triple Gaussian is then formed by subsequently adding the narrowest Gaussian, the double Gaussian has a fraction $f_{23}^{\tau} = 0.85 \pm 0.05$. The non-prompt signal component

$T^N_\psi(\tau)$ is a single-sided exponential with a free lifetime convoluted with the resolution function $T^P_\psi(\tau)$.

The background PDF is defined as:

$$\mathrm{PDF}_{\mathrm{bkg}}(m,\tau) = f^P_{\mathrm{bkg}}\left[M^P_{\mathrm{bkg}}(m)\cdot T^P_{\mathrm{bkg}}(\tau)\right] + (1-f^P_{\mathrm{bkg}})\left[M^N_{\mathrm{bkg}}(m)\cdot T^N_{\mathrm{bkg}}(\tau)\right], \quad \text{(E.2)}$$

where f^P_{bkg} is the fraction of background which is promptly produced, $M^P_{\mathrm{bkg}}(m)(M^N_{\mathrm{bkg}}(m))$ are the mass components of the prompt(non-prompt) background, and $T^P_{\mathrm{bkg}}(\tau)(T^N_{\mathrm{bkg}}(\tau))$ are the respective lifetime components. The mass components are both modelled as second order Chebyshev polynomials. The prompt lifetime component $T^P_{\mathrm{bkg}}(\tau)$ is modelled by the resolution function, and the non-prompt component $T^N_{\mathrm{bkg}}(\tau)$ is modelled as a mixture of a double-sided exponential and two single-sided exponentials, all convoluted by the resolution function $T^P_\psi(\tau)$.

The total PDF is a combination of the signal and background PDFs:

$$\mathrm{PDF}_{\mathrm{tot}}(m,\tau) = f_{\mathrm{sig}}\mathrm{PDF}_{\mathrm{sig}}(m,\tau) + (1-f_{\mathrm{sig}})\mathrm{PDF}_{\mathrm{bkg}}(m,\tau), \quad \text{(E.3)}$$

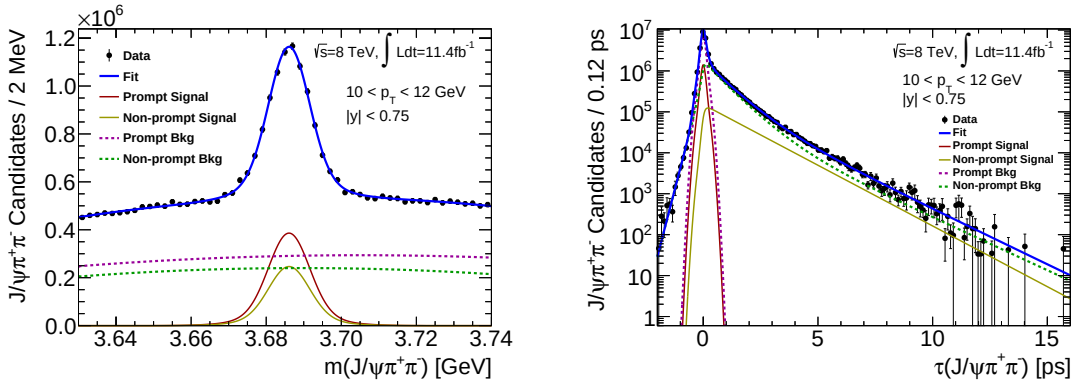
where f_{sig} is the fraction of events which are signal. This PDF was used to fit the mass-lifetime distribution of the data around the $\psi(2S)$ signal. A summary of the resulting fit parameters are shown in Table E.1.

The resolution function $T^P_\psi(\tau)$, with parameters listed in Table E.1, is used as the lifetime resolution PDF for the rest of this analysis, for both $\psi(2S)$ and $X(3872)$ signal and for the background as well. In Appendix G it is shown that indeed the lifetime resolution for $\psi(2S)$ and $X(3872)$ signal is the same within error. Since the fitted values of f_1 are reasonably stable across the p_{T} range, the value $f_1 = 0.76 \pm 0.04$ was adopted for the main fits.

The mass and lifetime projections of the fits are shown in Figures E.1 - E.5 for each p_{T} bin, and fit quality is assessed by plotting the ratio of the data and the fit model for the mass projections in Figure E.6 and the lifetime projections in Figure E.7. Good consistency with unity is seen throughout.

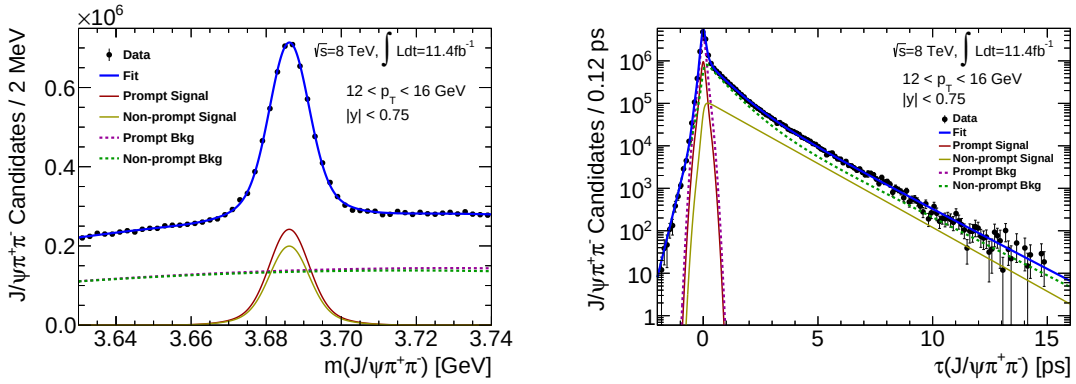
	p_{T} interval [GeV]				
	$10-12$	$12-16$	$16-22$	$22-40$	$40-70$
	Lifetime Resolution Parameters				
$\tau_{\mathrm{mean}}[\mathrm{ps}]\times10^{-3}$	-1.7 ± 0.5	-0.6 ± 0.3	-0.5 ± 0.2	-0.3 ± 0.2	-1.0 ± 0.9
$\tau_{\sigma}[\mathrm{ps}]\times10^{-3}$	44.3 ± 6.2	45.9 ± 3.3	52.0 ± 1.5	40.9 ± 1.5	32.2 ± 6.9
f_3^{τ}	0.34 ± 0.08	0.29 ± 0.03	0.19 ± 0.01	0.22 ± 0.01	0.20 ± 0.13
f_{23}^{τ}	0.94 ± 0.07	0.85 ± 0.05	0.59 ± 0.03	0.61 ± 0.04	0.49 ± 0.22
	Mass Peak Parameters				
$m_{\mathrm{mean}}[\mathrm{GeV}]$	3.686 ± 0.000	3.686 ± 0.000	3.686 ± 0.000	3.686 ± 0.000	3.686 ± 0.000
$m_{\sigma}[\mathrm{MeV}]$	5.4 ± 0.1	5.1 ± 0.1	5.0 ± 0.1	5.0 ± 0.1	5.4 ± 0.8
f_1	0.81 ± 0.03	0.73 ± 0.02	0.72 ± 0.03	0.74 ± 0.04	0.72 ± 0.24

Table E.1 Fit parameters determined from unbinned maximum likelihood fits of the $\psi(2S)$ signal region, to determine the lifetime resolution functions in bins of p_{T}.



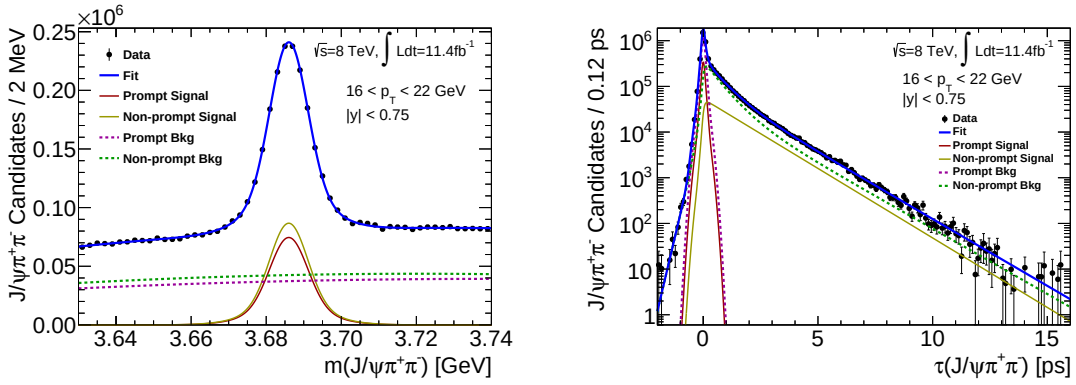
(a) (b)

Fig. E.1 (a) Mass projection over the full lifetime range and (b) lifetime projection over the full mass range, for the 2D unbinned maximum likelihood fit in the $\psi(2S)$ signal region in the p_T bin $[10, 12]$ GeV to establish lifetime resolution.



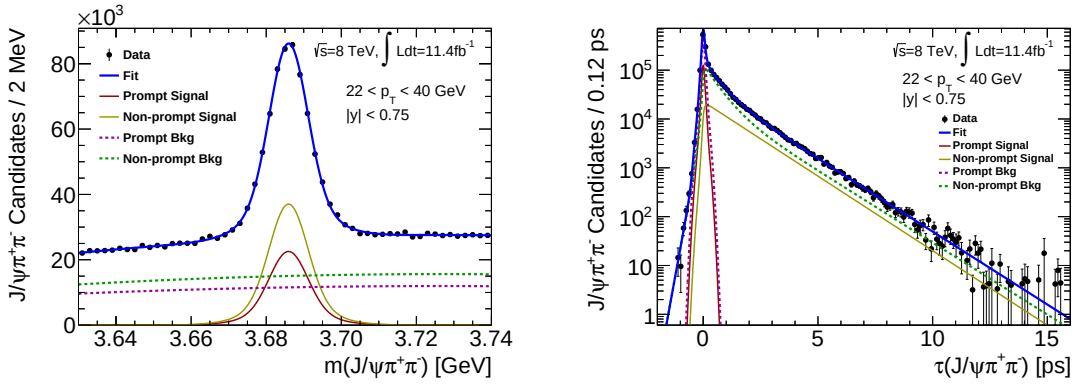
(a) (b)

Fig. E.2 (a) Mass projection over the full lifetime range and (b) lifetime projection over the full mass range, for the 2D unbinned maximum likelihood fit in the $\psi(2S)$ signal region in the p_T bin $[12, 16]$ GeV to establish lifetime resolution.



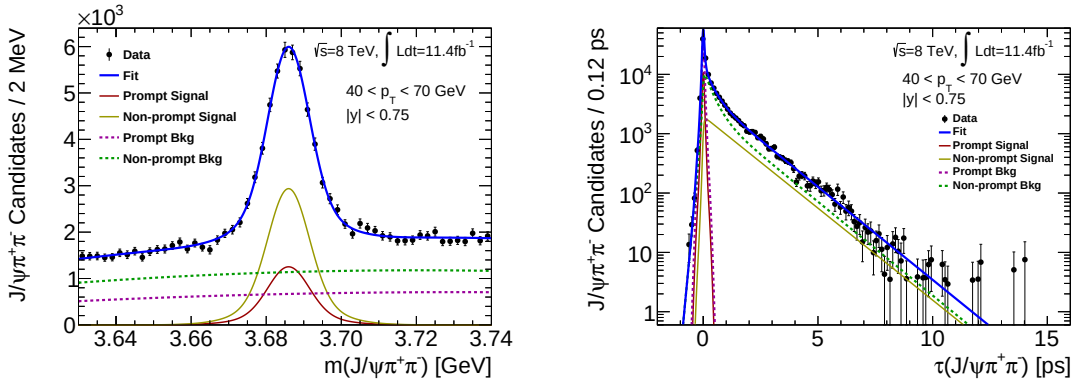
(a) (b)

Fig. E.3 (a) Mass projection over the full lifetime range and (b) lifetime projection over the full mass range, for the 2D unbinned maximum likelihood fit in the $\psi(2S)$ signal region in the p_T bin $[16, 22]$ GeV to establish lifetime resolution.



(a) (b)

Fig. E.4 (a) Mass projection over the full lifetime range and (b) lifetime projection over the full mass range, for the 2D unbinned maximum likelihood fit in the $\psi(2S)$ signal region in the p_T bin $[22, 40]$ GeV to establish lifetime resolution.



(a) (b)

Fig. E.5 (a) Mass projection over the full lifetime range and (b) lifetime projection over the full mass range, for the 2D unbinned maximum likelihood fit in the $\psi(2S)$ signal region in the p_T bin $[40, 70]$ GeV to establish lifetime resolution.

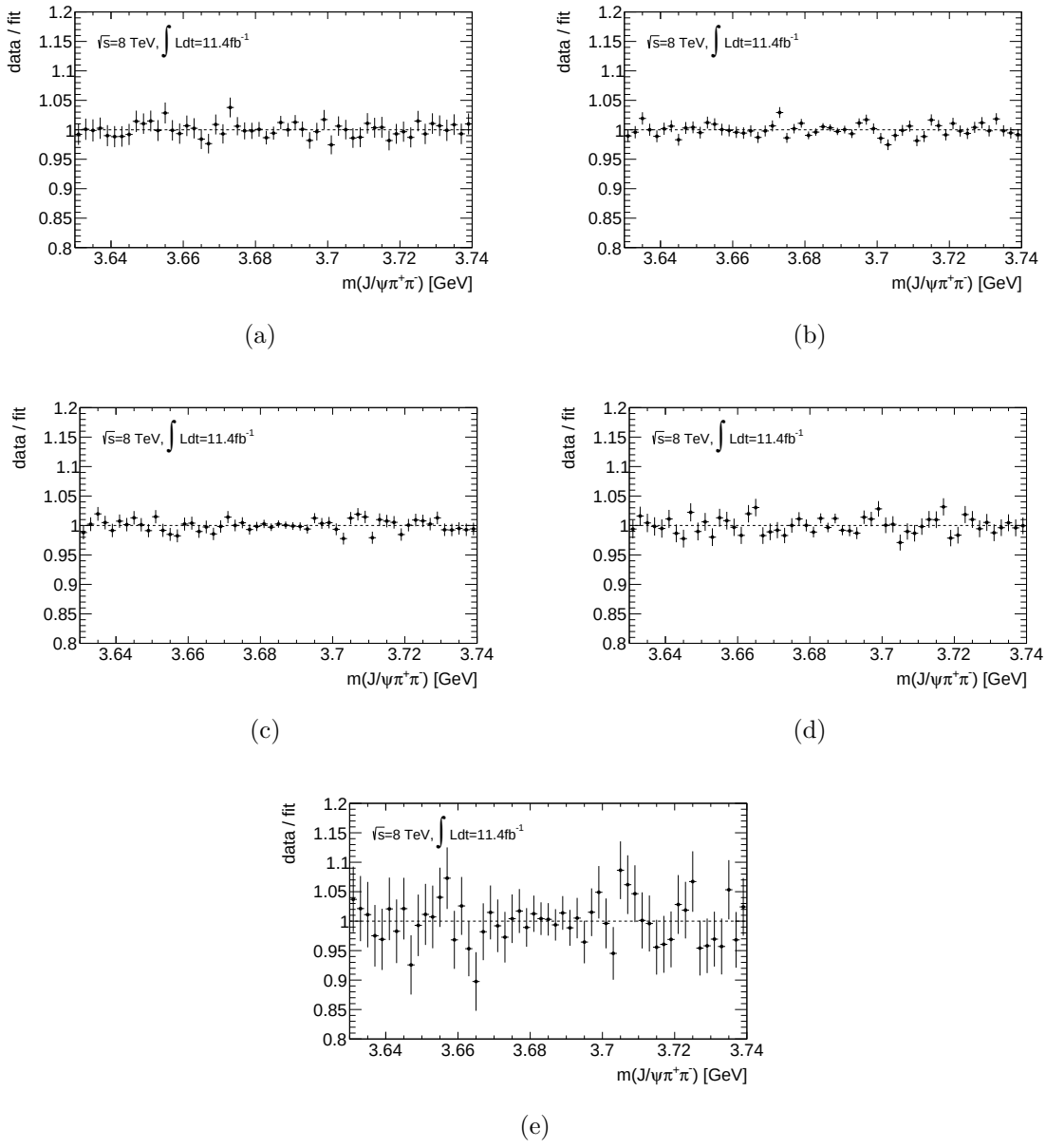


Fig. E.6 Data / fit ratio for the mass projection of the unbinned 2D mass-lifetime likelihood fits performed in the $\psi(2S)$ signal region to determine the lifetime resolution function in analysis bins of p_{T} ascending from (a) $[10, 12]$ GeV to (e) $[40, 70]$ GeV.

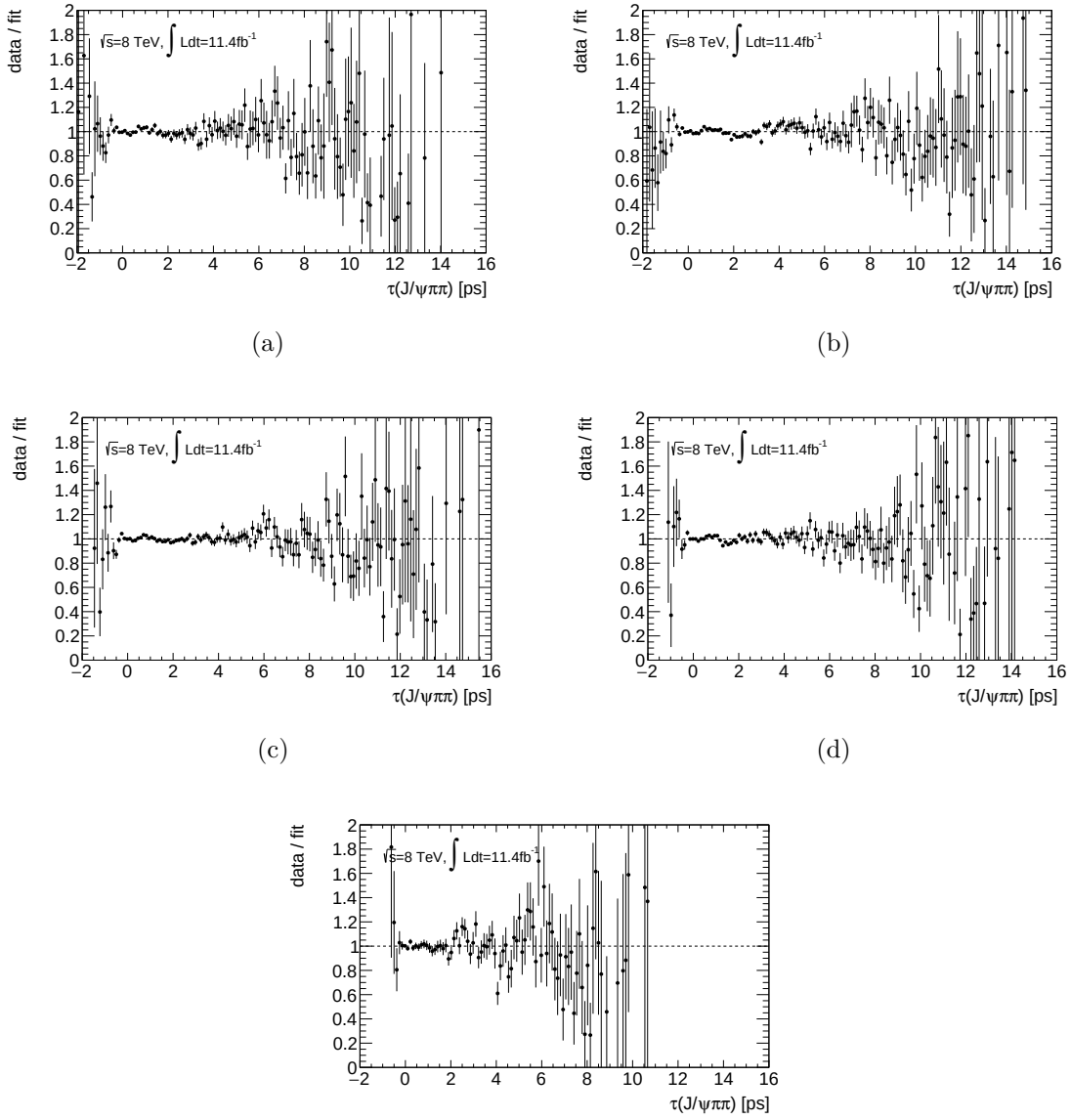


Fig. E.7 Data / fit ratio for the lifetime projection of the unbinned 2D mass-lifetime likelihood fits performed in the $\psi(2S)$ signal region to determine the lifetime resolution function in bins of p_{T} ascending from (a) $[10, 12]$ GeV to (e) $[40, 70]$ GeV.

E.2 Short-lived Non-Prompt Component

We use Monte Carlo simulation to estimate the possible variation of the pseudoproper lifetime of the short-lived component, τ_{SL}, in inclusive $B_c \to i$ $(i = \psi(2S), X(3872))$ decays, for the two-lifetime fits. The largest possible τ_{SL} is close to the lifetime of the B_c, and would correspond to the invariant mass of the associated hadronic particles in the B_c decay being maximal. The upper limit of τ_{SL} is therefore set close to the B_c lifetime, $\tau(B_c) = (0.507 \pm 0.009)$ ps [18]. Oppositely, the shortest possible τ_{SL} arises from the invariant mass of the associated hadronic particles in the B_c decay being minimal, which in this case would be a pion. We use $B_c \to J/\psi\pi$ simulation to estimate the shortest possible τ_{SL}, by fitting the J/ψ pseudoproper lifetime with an exponential in the region in which our analysis is sensitive to B_c signal, shown in Figure E.8.

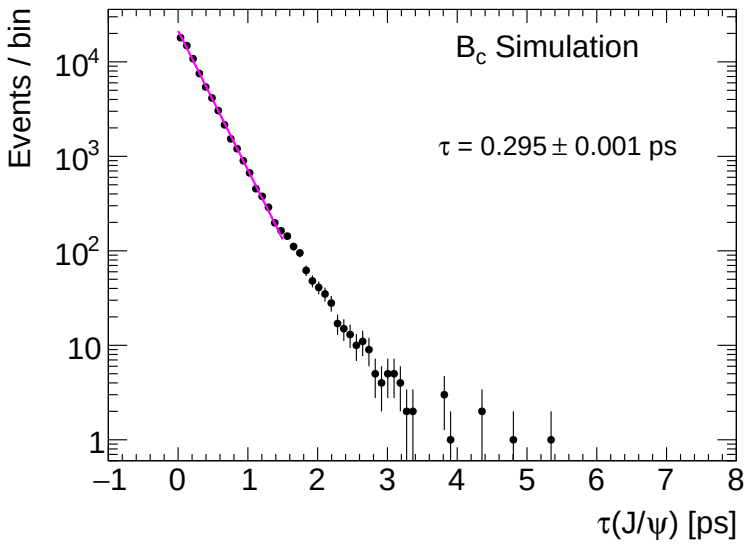


Fig. E.8 Pseudoproper lifetime distribution of the J/ψ in $B_c \to J/\psi\pi$ simulation. The distribution is fitted with an exponential to measure the pseudoproper lifetime in the region our analysis is most sensitive to.

This lighter J/ψ will have an even shorter pseudoproper lifetime than the $\psi(2S)$ and $X(3872)$ from the kinematics, and so this measurement of $\tau_{SL} = 0.295 \pm 0.001$ ps represents by far the shortest possible τ_{SL}. In summary, we set $\tau_{SL} = 0.40 \pm 0.05$ ps,

which appropriately covers the range of possible pseudoproper lifetimes for $\psi(2S)$ and $X(3872)$ produced from B_c decays.

E.3 Long-lived Non-Prompt Component

We use a mix of data and simulation to estimate the pseudoproper lifetime τ_{LL} of $\psi(2S)$ and $X(3872)$ from long-lived B decays, for the two-lifetime fits. A data-driven method is used to measure the central value of τ_{LL} and its statistical error. We perform fits to the $\psi(2S)$ mass signal in windows of pseudo-proper lifetime (shown in Figures E.9 - E.13) far from the resolution threshold such that prompt signal is negligible to first order:

- $0.50 - 0.75$ ps
- $0.75 - 1.5$ ps
- $1.5 - 2.7$ ps
- $2.7 - 5.0$ ps
- $5.0 - 15.0$ ps

The signal yields returned by the fits in each p_{T} bin, and for each pseudoproper lifetime window, are shown in Table E.2.

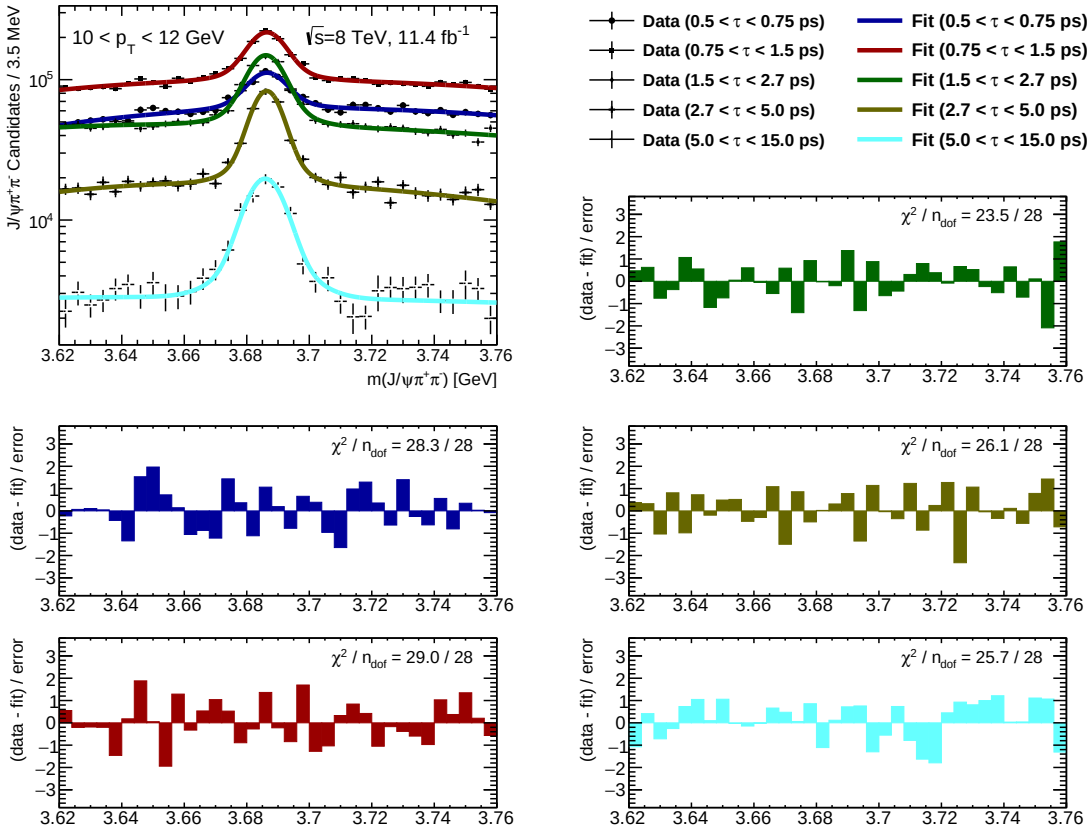


Fig. E.9 Invariant mass fit of the $\psi(2S)$ signal in windows of pseudoproper lifetime for the p_{T} bin $[10, 12]$ GeV.

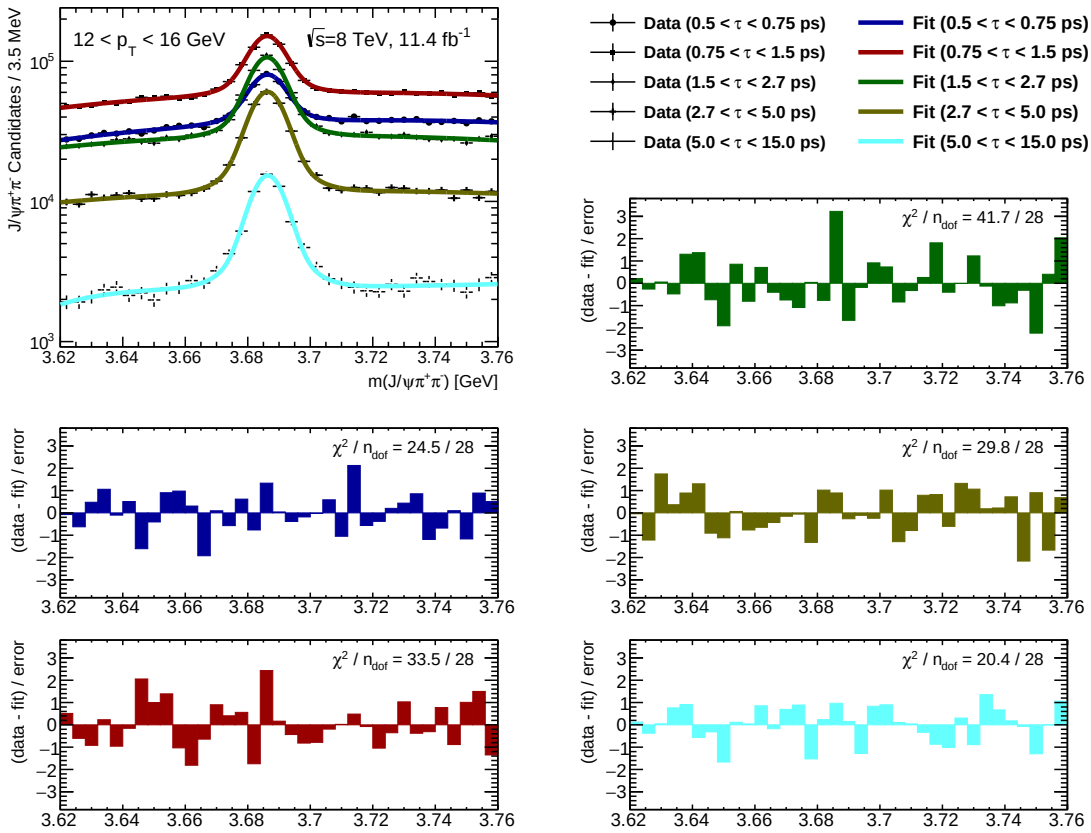


Fig. E.10 Invariant mass fit of the $\psi(2S)$ signal in windows of pseudoproper lifetime for the p_{T} bin $[12, 16]$ GeV.

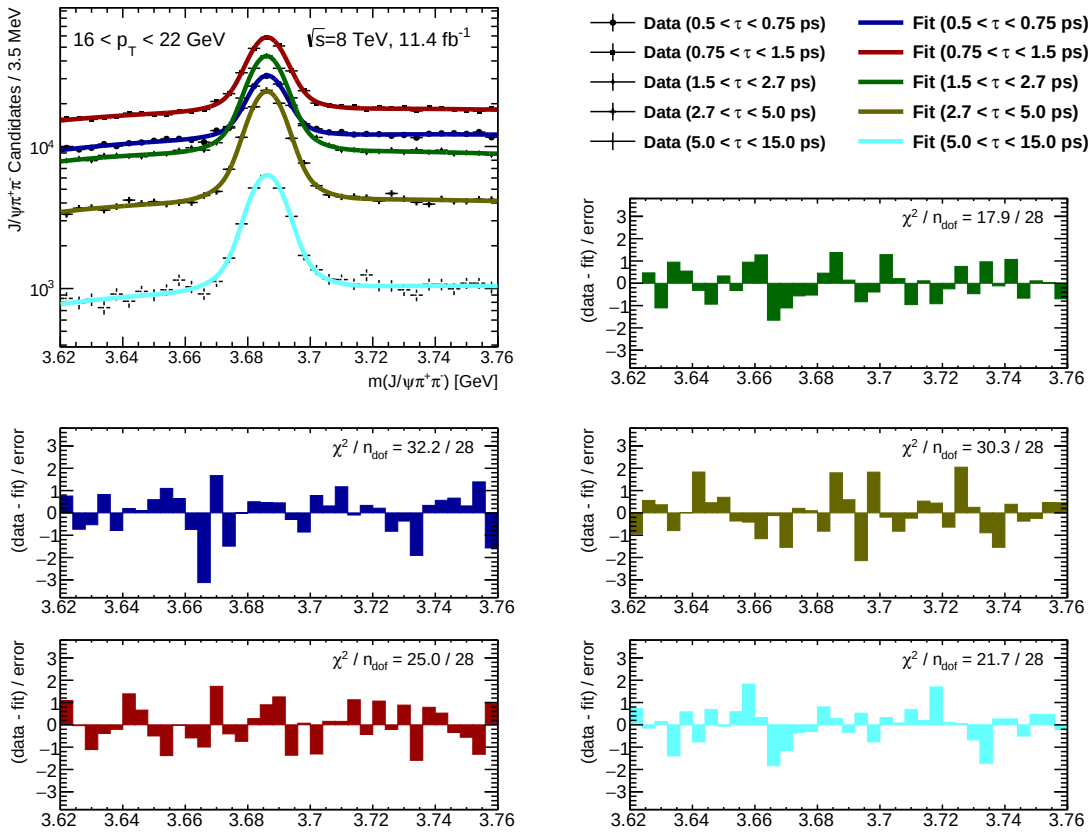


Fig. E.11 Invariant mass fit of the $\psi(2S)$ signal in windows of pseudoproper lifetime for the p_{T} bin $[16, 22]$ GeV.

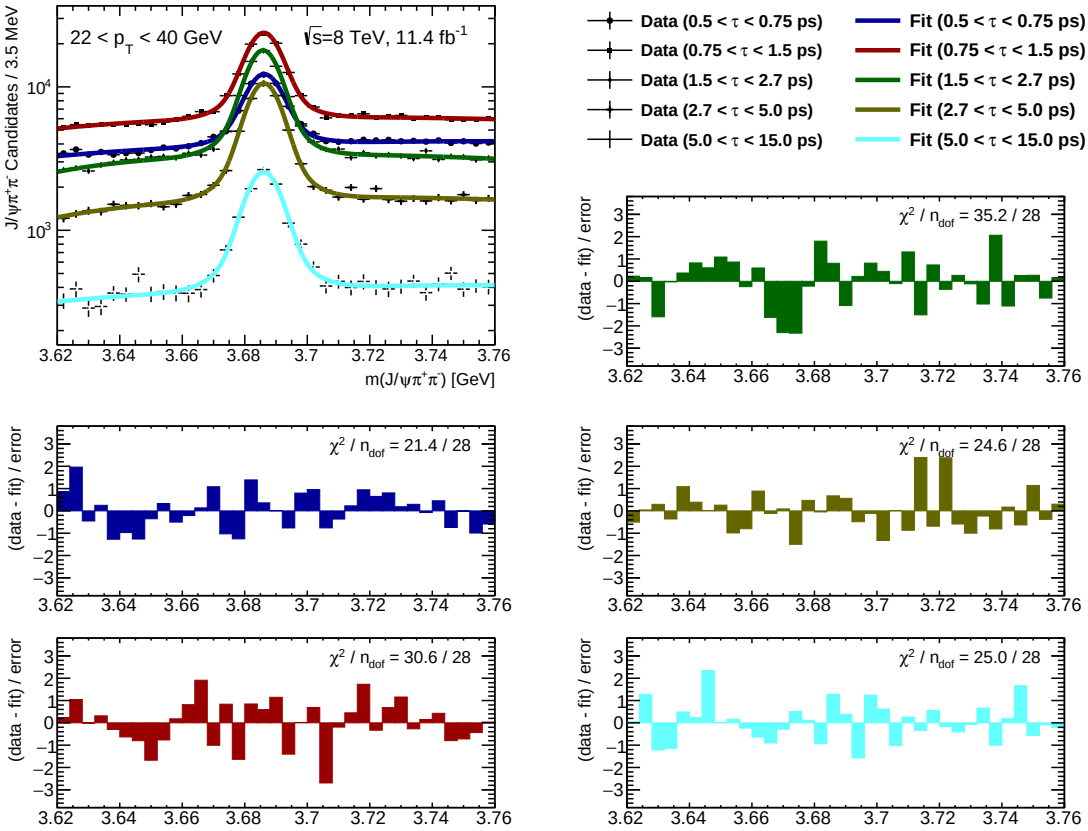


Fig. E.12 Invariant mass fit of the $\psi(2S)$ signal in windows of pseudoproper lifetime for the p_{T} bin $[22, 40]$ GeV.

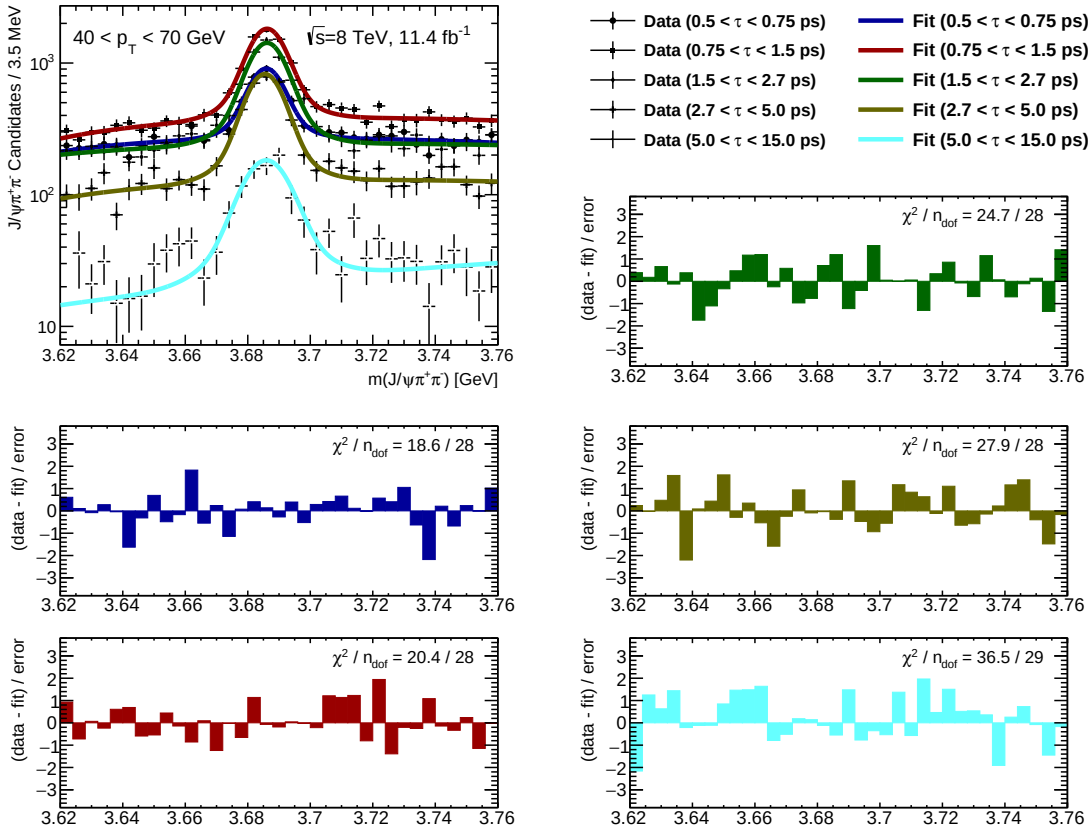


Fig. E.13 Invariant mass fit of the $\psi(2S)$ signal in windows of pseudoproper lifetime for the p_{T} bin $[40, 70]$ GeV.

	$\psi(2S)$ Yields $[\times 10^5]$ vs. p_{T} [GeV]				
τ window	$10-12$	$12-16$	$16-22$	$22-40$	$40-70$
w_0	2.1 ± 0.2	1.7 ± 0.05	0.73 ± 0.02	0.32 ± 0.01	0.023 ± 0.002
w_1	4.7 ± 0.2	3.6 ± 0.07	1.5 ± 0.02	0.67 ± 0.01	0.057 ± 0.002
w_2	4.0 ± 0.2	2.9 ± 0.05	1.3 ± 0.02	0.53 ± 0.01	0.047 ± 0.002
w_3	2.2 ± 0.1	1.8 ± 0.03	0.75 ± 0.01	0.33 ± 0.006	0.028 ± 0.002
w_4	0.7 ± 0.05	0.47 ± 0.02	0.20 ± 0.006	0.082 ± 0.003	0.008 ± 0.0007

Table E.2 Invariant mass fit results in pseudoproper lifetime and p_{T} bins for the $\psi(2S)$. Uncertainties are statistical only.

We then construct two lifetime PDFs. The first describes the short-lived B_c component, modelled as an exponential with a lifetime 0.4 ps convoluted with the resolution function. The long-lived component is an exponential with a free lifetime determined by the fit, convoluted with the resolution function. A minimum-χ^2 fit is performed, where the lifetime PDFs are integrated over the pseudoproper lifetime bins:

$$\chi^2 = \sum_{i=1}^{n} \left[\frac{d_i - Y_{\rm np} \left(F_{sl} f_{sl}^i + (1 - F_{sl}) f_{ll}^i \right)}{\Delta d_i} \right]^2 \tag{E.4}$$

where $Y_{\rm np}$ is the total non-prompt signal yield, F_{sl} is the short-lived non-prompt signal fraction, and f_{sl}^i and f_{ll}^i are the integrals of the short and long non-prompt lifetime PDFs, respectively, over the specific lifetime boundaries of the i^{th} bin.

These fits are performed separately in each of the analysis $p_{\rm T}$ bins. The results of the fits are shown in Table E.3.

	$p_{\rm T}$ interval [GeV]				
	$10-12$	$12-16$	$16-22$	$22-40$	$40-70$
	ψ(2S) Yields [$\times 10^6$] vs. $p_{\rm T}$				
$Y_{\rm np}$	1.95 ± 0.095	1.49 ± 0.029	0.65 ± 0.01	0.28 ± 0.005	0.022 ± 0.001
	Fractions vs. $p_{\rm T}$				
$F_{\rm SL}$	0.051 ± 0.056	0.035 ± 0.023	0.025 ± 0.019	0.053 ± 0.021	-0.077 ± 0.069
	Measured Pseudo-proper Lifetime				
τ [ps]	1.47 ± 0.05	1.47 ± 0.02	1.43 ± 0.02	1.44 ± 0.02	1.45 ± 0.05
	Fit quality in each $p_{\rm T}$ bin				
$\chi^2/n_{\rm dof}$	3.6/3	1.4/3	2.7/3	0.6/3	2.6/3

Table E.3 Fit results in $p_{\rm T}$ bins for the ψ(2S). Uncertainties are statistical only.

The extracted long-lived lifetime for each $p_{\rm T}$ bin is shown in Figure E.14, and fitted with a zeroth order polynomial to determine the average lifetime $\tau_{LL} = 1.45 \pm 0.01$(stat), which is used as the central value for each $p_{\rm T}$ bin in the analysis. The systematic error on this measurement is evaluated using simulation. As explained previously, the average pseudoproper lifetime of the $\psi(2S)$ and $X(3872)$ is dependent on the invariant mass of the associately produced hadronic component, from here on referred to as X'. A large invariant mass of X' causes the pseudoproper lifetime of $\psi(2S)$ or

$X(3872)$ to be much closer to the proper lifetime of the decaying B hadron on average. Oppositely, if X' has a small invariant mass then the pseudoproper lifetime of the $\psi(2S)$ or $X(3872)$ will be less than the proper lifetime of the decaying B hadron. We measure the pseudoproper lifetime from truth Monte Carlo for two extreme cases. The first, where X' is a kaon. The second, where X' is any particle(s) with an invariant mass greater than 1.28 GeV. The pseudo-proper lifetime is measured as $\tau = 1.41 \pm 0.01$ ps where X' is a kaon, and $\tau = 1.56 \pm 0.01$ ps where $m(X') > 1.28$ GeV. These two cases represent the most extreme cases of production, and in reality there is some mixture of all $m(X')$. We consequently assign a total error on the long-lived lifetime measurement of 0.05 ps. In summary, the long-lived lifetime component is fixed as $\tau_{LL} = 1.45 \pm 0.05$ ps, which appropriately covers the range of effective pseudoproper lifetimes for $\psi(2S)$ and $X(3872)$ produced from long-lived B decays.

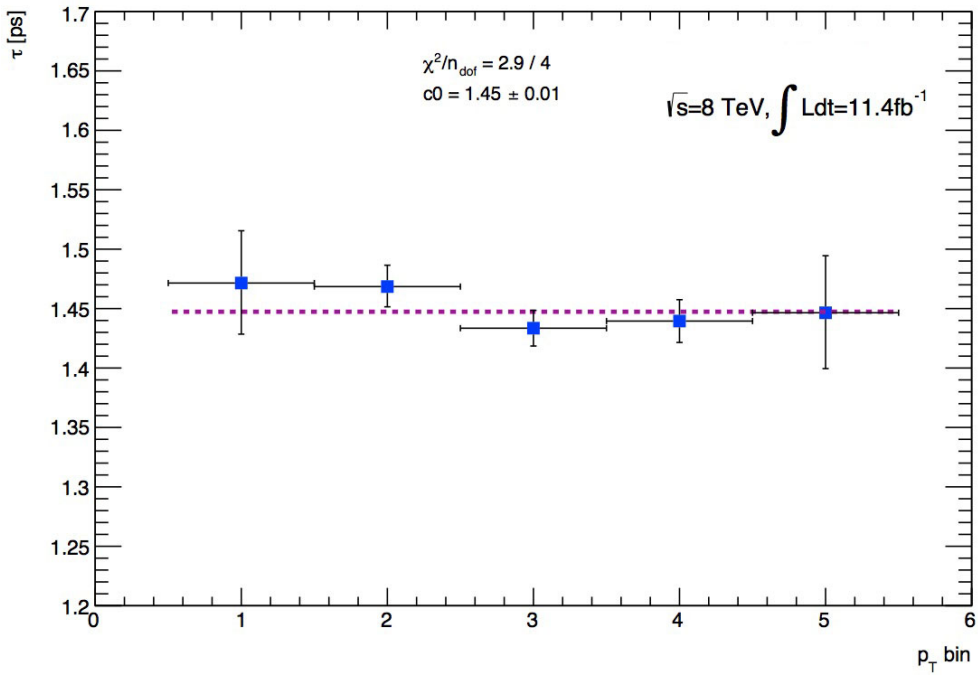
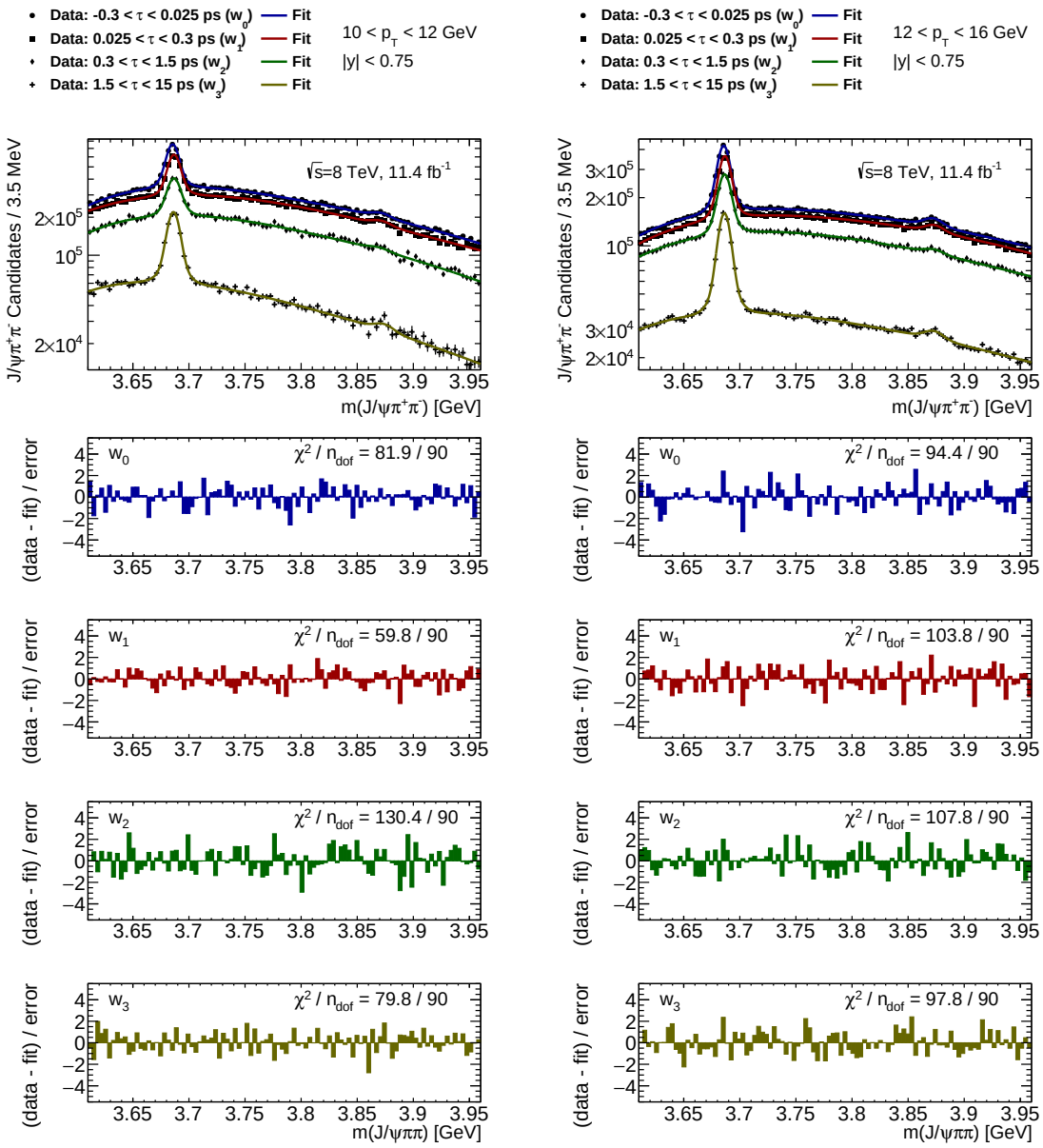


Fig. E.14 Extracted lifetime of the long-lived non-prompt $\psi(2S)$ signal component for each p_{T} bin, fitted with a zeroth order polynomial to measure the average lifetime.

Appendix F

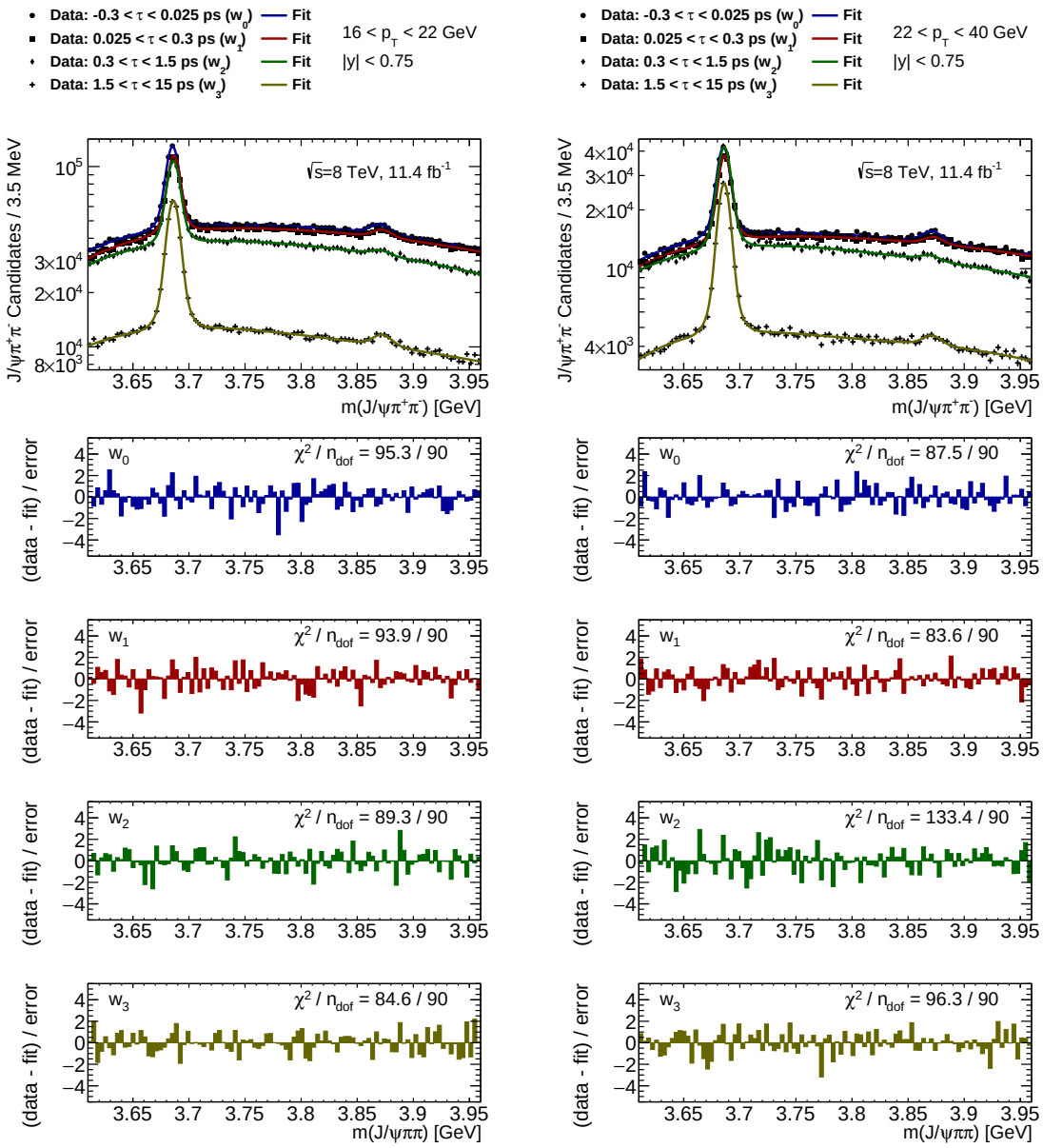
Invariant Mass Fits in Lifetime Windows

We show the the minimum-χ^2 fits to the invariant mass of the $J/\psi\pi^+\pi^-$ candidates in windows of pseudoproper lifetime for each anlysis p_{T} bin, used to extract the $\psi(2S)$ and $X(3872)$ signal yields, in Figures F.1-F.3. We also show the fits with a linear scale on the y-axis instead of a logarithmic scale, in Figures F.4 - F.8.



(a) (b)

Fig. F.1 Fits to the invariant mass spectra of the $J/\psi\pi\pi$ candidates to extract $\psi(2S)$ and $X(3872)$ signal for each pseudo-proper lifetime window in the p_T bin (a) $[10, 12]$ GeV and (b) $[12, 16]$ GeV. Shown underneath the fits are the corresponding pull distributions.



(a) (b)

Fig. F.2 Fits to the invariant mass spectra of the $J/\psi\pi\pi$ candidates to extract $\psi(2S)$ and $X(3872)$ signal for each pseudo-proper lifetime window in the p_T bin (a) $[16, 22]$ GeV and (b) $[22, 40]$ GeV. Shown underneath the fits are the corresponding pull distributions.

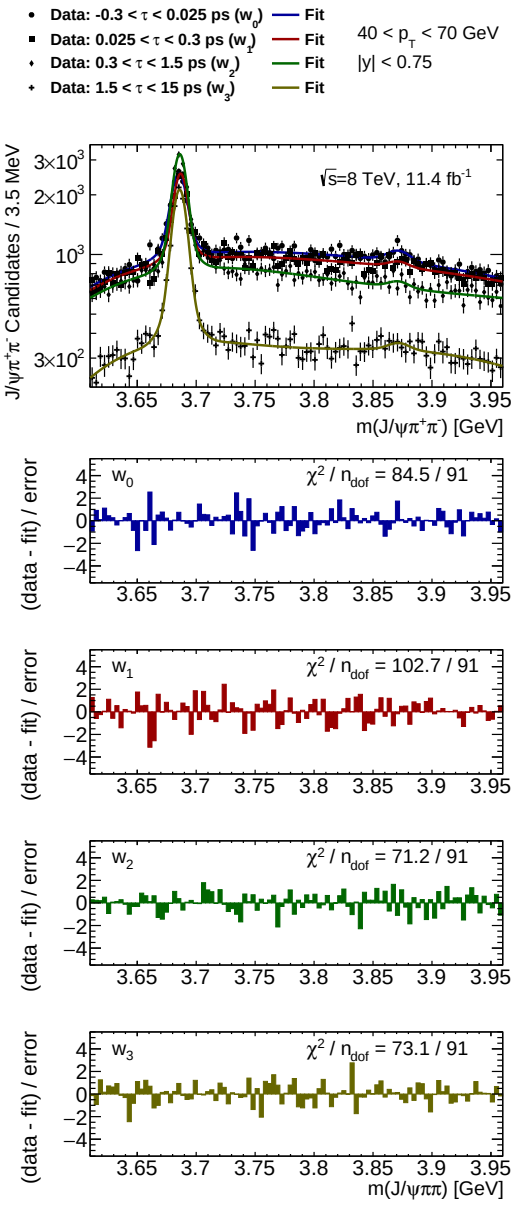


Fig. F.3 Fits to the invariant mass spectra of the $J/\psi\pi\pi$ candidates to extract $\psi(2S)$ and $X(3872)$ signal for each pseudo-proper lifetime window in the p_T bin $[40, 70]$ GeV. Shown underneath the fits are the corresponding pull distributions.

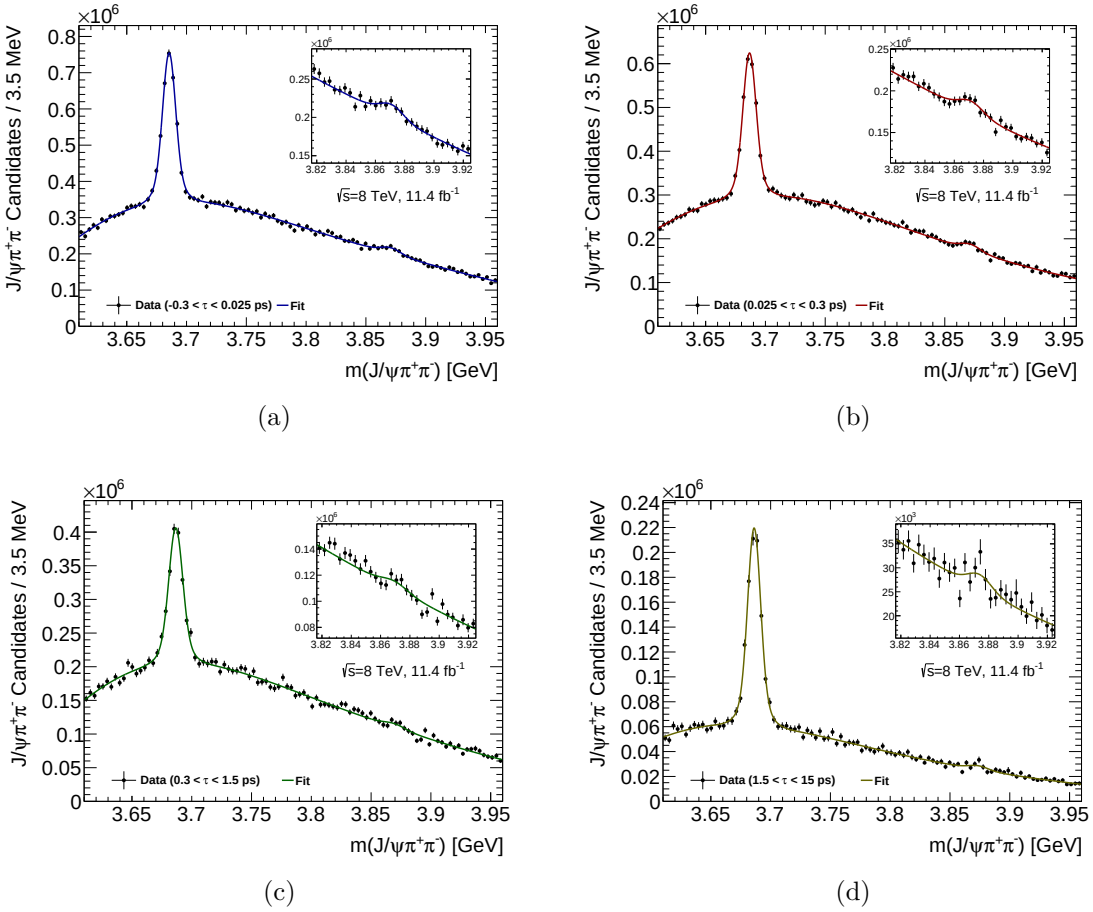


Fig. F.4 Minimum χ^2 mass fits to extract $\psi(2S)$ and $X(3872)$ signal yields displayed with a linear y-axis scale, in windows of pseudoproper lifetime: (a) -0.3 - 0.025 ps, (b) 0.025 - 0.3 ps, (c) 0.3 - 1.5 ps, (d) 1.5 - 15 ps for the $p_{\mathrm{T}} = [10, 12]$ GeV bin.

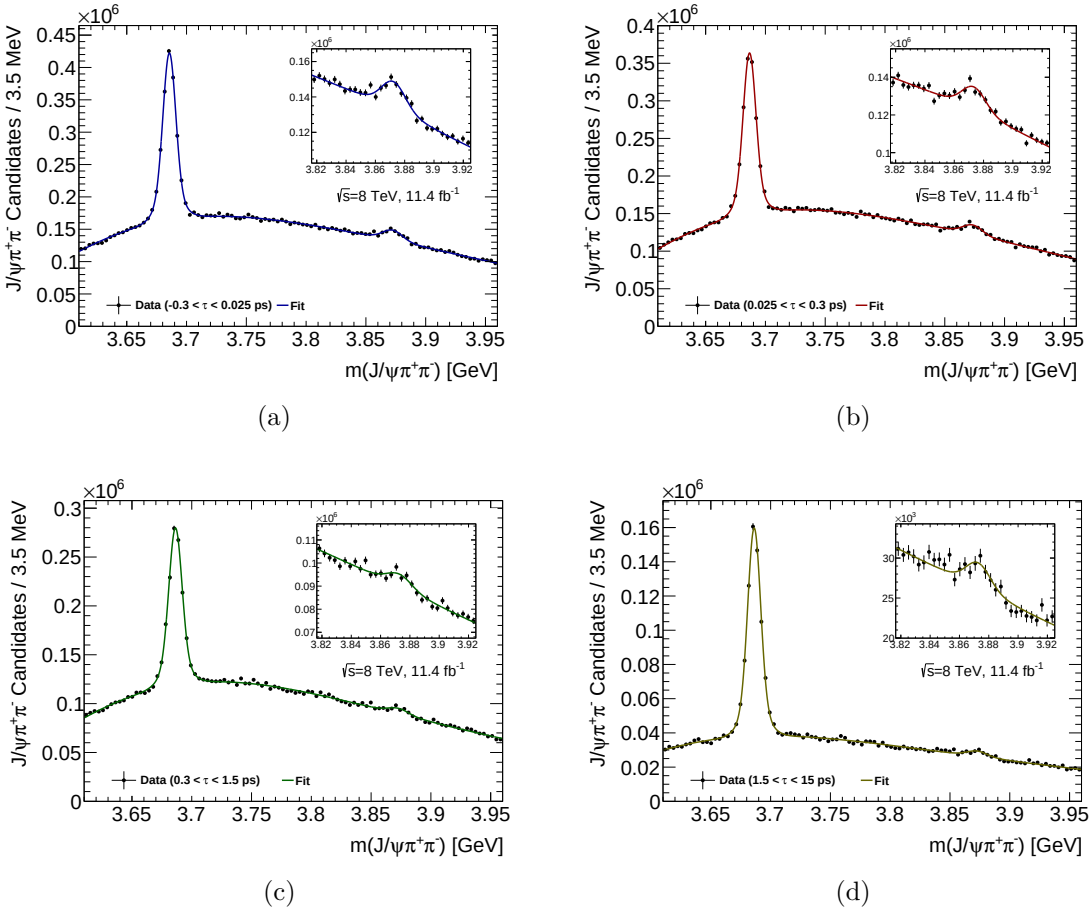


Fig. F.5 Minimum χ^2 mass fits to extract $\psi(2S)$ and $X(3872)$ signal yields displayed with a linear y-axis scale, in windows of pseudoproper lifetime: (a) -0.3 - 0.025 ps, (b) 0.025 - 0.3 ps, (c) 0.3 - 1.5 ps, (d) 1.5 - 15 ps for the $p_{\mathrm{T}} = [12, 16]$ GeV bin.

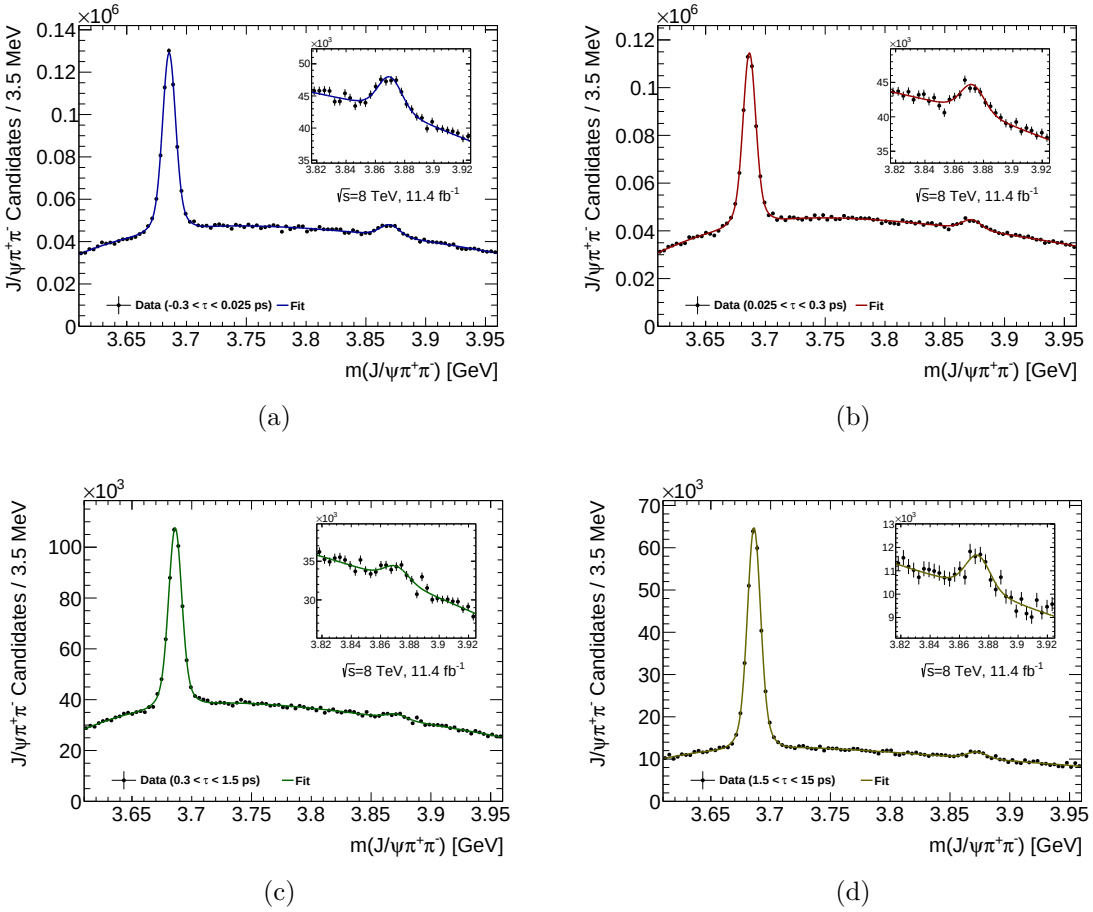


Fig. F.6 Minimum χ^2 mass fits to extract $\psi(2S)$ and $X(3872)$ signal yields displayed with a linear y-axis scale, in windows of pseudoproper lifetime: (a) -0.3 - 0.025 ps, (b) 0.025 - 0.3 ps, (c) 0.3 - 1.5 ps, (d) 1.5 - 15 ps for the $p_{\mathrm{T}} = [16, 22]$ GeV bin.

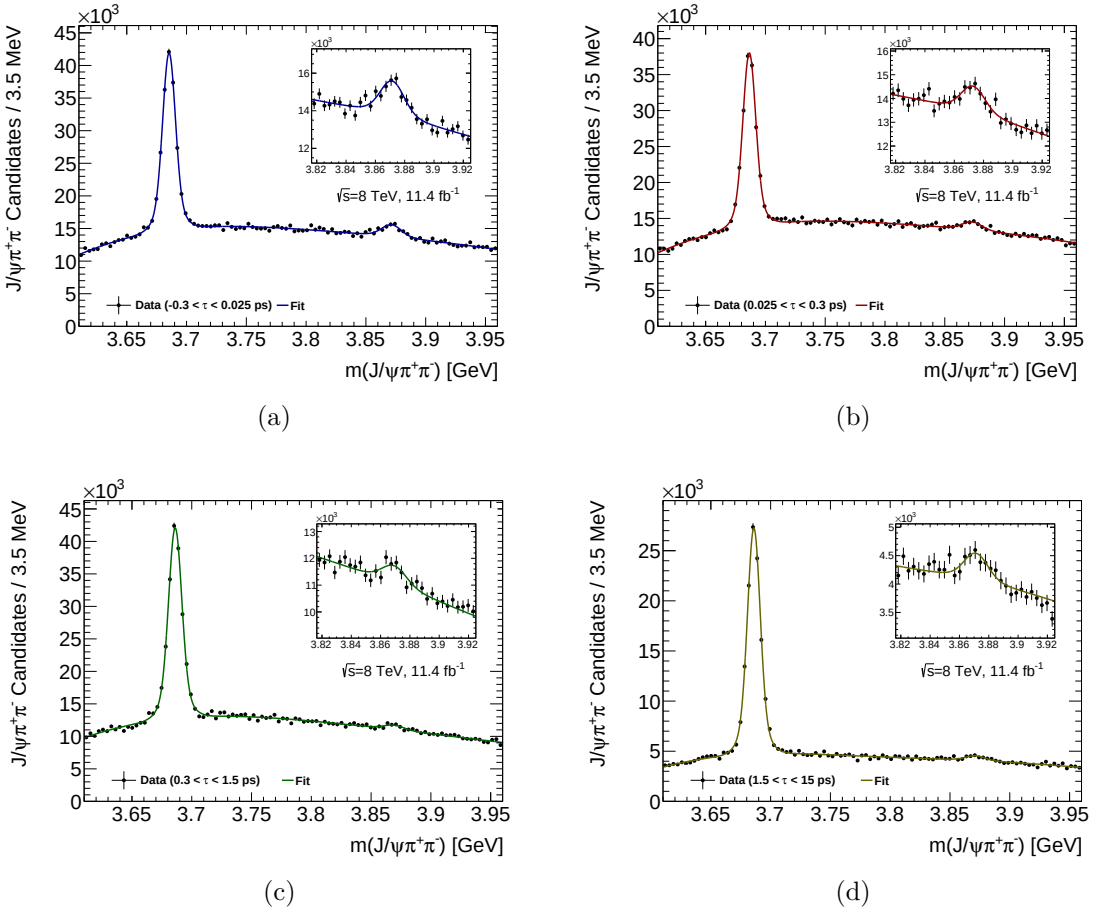


Fig. F.7 Minimum χ^2 mass fits to extract $\psi(2S)$ and $X(3872)$ signal yields displayed with a linear y-axis scale, in windows of pseudoproper lifetime: (a) -0.3 - 0.025 ps, (b) 0.025 - 0.3 ps, (c) 0.3 - 1.5 ps, (d) 1.5 - 15 ps for the $p_{\mathrm{T}} = [22, 40]$ GeV bin.

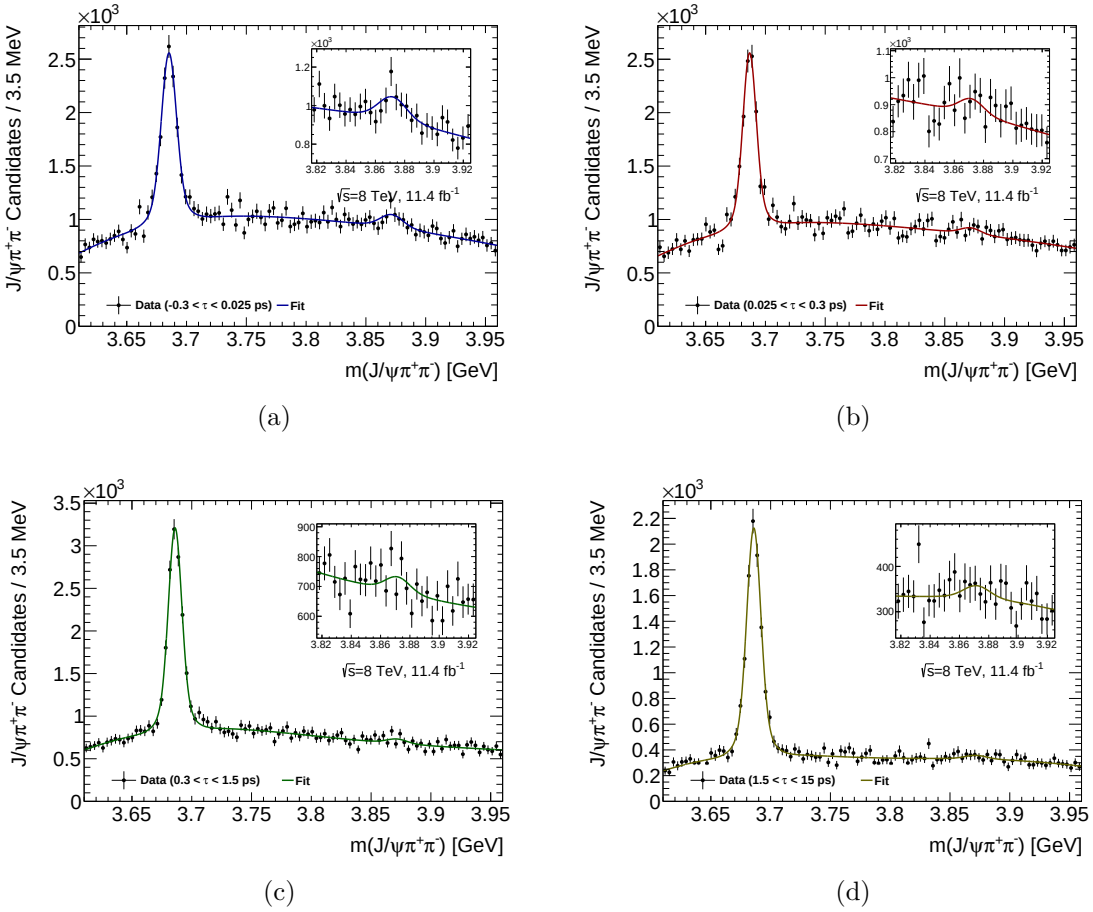


Fig. F.8 Minimum χ^2 mass fits to extract $\psi(2S)$ and $X(3872)$ signal yields displayed with a linear y-axis scale, in windows of pseudoproper lifetime: (a) -0.3 - 0.025 ps, (b) 0.025 - 0.3 ps, (c) 0.3 - 1.5 ps, (d) 1.5 - 15 ps for the $p_{\mathrm{T}} = [40, 70]$ GeV bin.

Appendix G

Verification of Fit Model Assumptions

For good stability of the mass and lifetime fit models when extracting $\psi(2S)$ and $X(3872)$ signal from the $J/\psi\pi^+\pi^-$ candidates, several assumptions were made. When modelling the mass peaks with double Gaussians, the narrower Gauss fraction f_1 was forced to be the same for the two signals at a value of 0.76 ± 0.04 (determined as an average from mass-lifetime fits to the data where this fraction was free). The lifetime resolution model $F_{\text{res}}(\tau)$ for the $\psi(2S)$ and $X(3872)$ signal was also shared.

The larger statistics $\psi(2S)$ signal in data will dominate the shape of $F_{\text{res}}(\tau)$ and f_1. A Monte Carlo study has subsequently been performed to verify that there are no significant differences in $F_{\text{res}}(\tau)$ and f_1 for $\psi(2S)$ and $X(3872)$ signal in each analysis p_{T} bin.

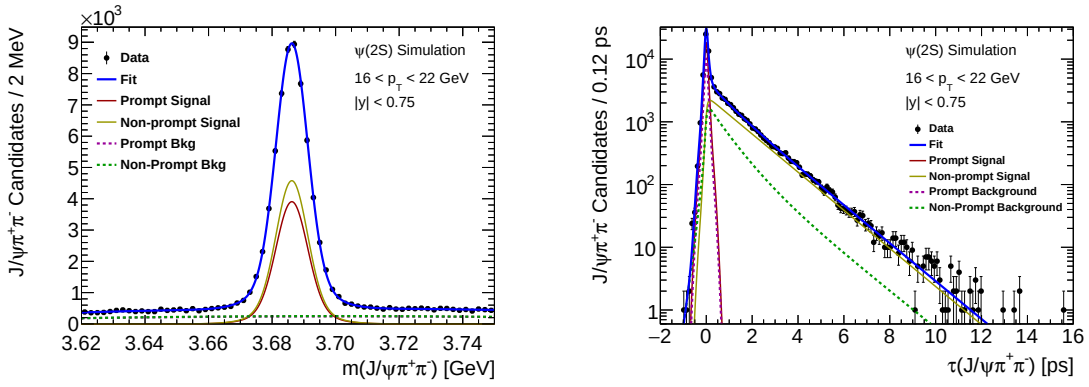
The description of the MC used for this study is detailed in Section 5.3.8. We perform unbinned mass-lifetime maximum likelihood fits separately for $\psi(2S)$ and $X(3872)$ signal regions. To mirror the main analysis, the narrower $X(3872)$ Gaussian describing the mass peak has its width fixed to $1.52\times$ the $\psi(2S)$ Gaussian width. The parameter f_1 is left free for both fits. As with the main analysis, $F_{\text{res}}(\tau)$ is modelled as a triple Gaussian, with the wider two Gaussians $2\times$ and $4\times$ the width of the narrowest. The fractions of these Gaussians are determined from the $\psi(2S)$ fit, and are fixed for the $X(3872)$ fit, however the width and mean of $F_{\text{res}}(\tau)$ are left free in both fits.

Projections of the mass-lifetime fits are shown for the $\psi(2S)$ and $X(3872)$ signal in the $p_{\text{T}} = 12 - 16$ GeV bin in Figures G.1 and G.2, respectively. Table G.1 shows f_1

and the mean and width of $F_{\text{res}}(\tau)$ returned by the fits in each bin of p_{T} for the $\psi(2S)$ and $X(3872)$ samples. We see agreement of f_1 within a few percent for $\psi(2S)$ and $X(3872)$, well within the systematic error of $\pm 4\%$ already applied to the main analysis as a systematic variation. The $F_{\text{res}}(\tau)$ parameters also closely agree with each other, leading us to conclude it is a safe assumption to allow the $\psi(2S)$ and $X(3872)$ signal to share the same lifetime resolution function in the main analysis.

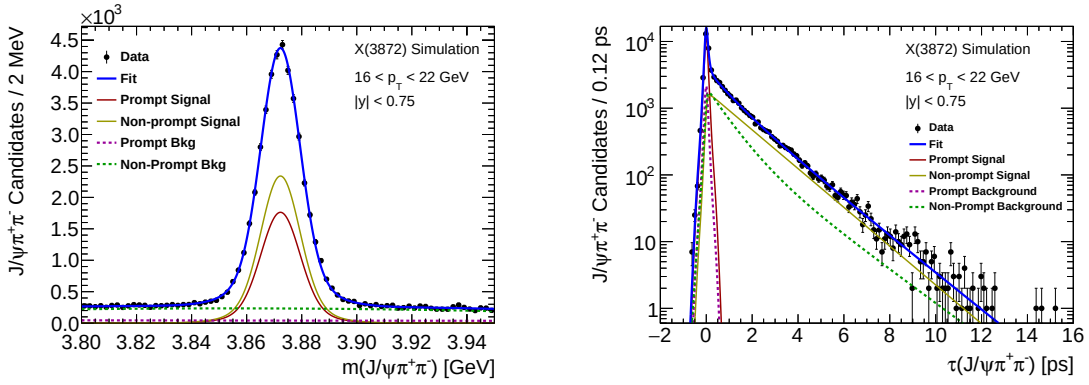
p_{T} [GeV]	$\psi(2S)$			$X(3872)$		
	f_1 [%]	$\tau_{\text{res}}^{\text{mean}}$ [$\times 10^{-3}$ ps]	$\tau_{\text{res}}^{\text{width}}$ [ps]	f_1 [%]	$\tau_{\text{res}}^{\text{mean}}$ [$\times 10^{-3}$ ps]	$\tau_{\text{res}}^{\text{width}}$ [ps]
$10-12$	75.9 ± 2.9	-1.2 ± 0.8	0.20 ± 0.01	78.0 ± 1.8	-1.4 ± 1.3	0.20 ± 0.01
$12-16$	80.1 ± 1.5	-0.7 ± 0.5	0.23 ± 0.01	80.4 ± 0.9	-0.9 ± 0.7	0.23 ± 0.01
$16-22$	82.0 ± 1.5	-0.5 ± 0.5	0.18 ± 0.01	81.0 ± 1.0	-1.7 ± 0.7	0.18 ± 0.01
$22-40$	79.6 ± 2.2	-0.7 ± 0.5	0.15 ± 0.01	78.9 ± 1.3	-0.6 ± 0.9	0.15 ± 0.01
$40-70$	89.9 ± 8.2	-0.3 ± 0.4	0.11 ± 0.01	86.6 ± 5.4	6.8 ± 3.4	0.12 ± 0.01

Table G.1 Summary of the unbinned mass-lifetime fit results to simulated $\psi(2S)$ and $X(3872)$ signal.



(a) (b)

Fig. G.1 (a) Mass projection over the full lifetime range and (b) lifetime projection over the full mass range for the 2D unbinned maximum likelihood fit of the simulated $\psi(2S)$ events in the p_T bin $[16, 22]$ GeV.



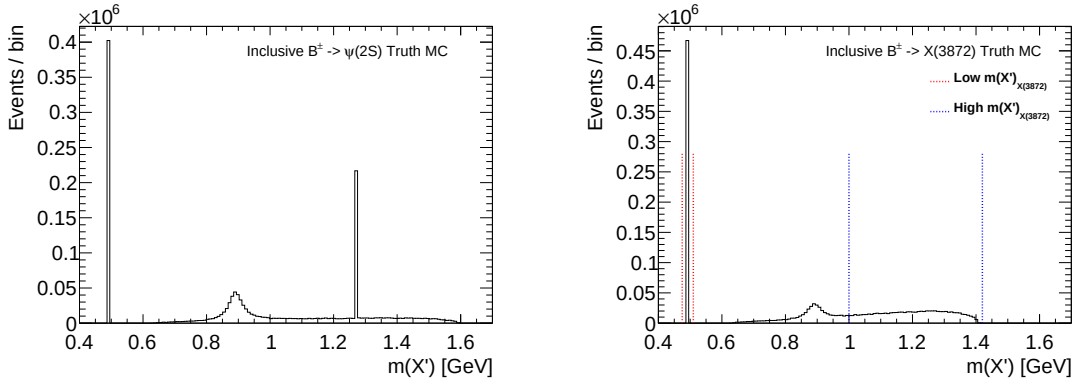
(a) (b)

Fig. G.2 (a) Mass projection over the full lifetime range and (b) lifetime projection over the full mass range for the 2D unbinned maximum likelihood fit of the simulated $X(3872)$ events in the p_T bin $[16, 22]$ GeV.

Appendix H

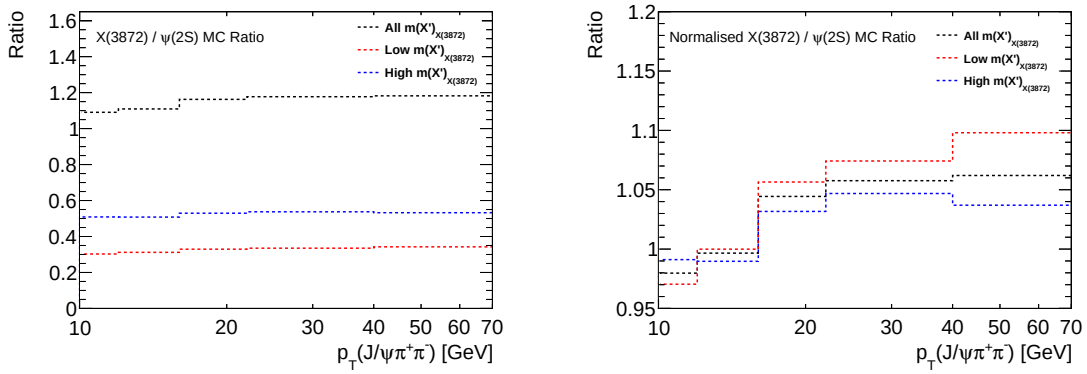
MC Template for Non-Prompt Ratio

We use the truth information from Monte Carlo generated for non-prompt $\psi(2S)$ and $X(3872)$ production detailed in Section 5.3.8 to create templates with which to fit $X(3872)/\psi(2S)$ non-prompt ratio results. The same acceptance criteria are applied to the muons and pions in these samples as in the main analysis. The events are subsequently weighted by acceptance maps as a correction. The templates are produced as a function of $p_T(J/\psi\pi^+\pi^-)$ by measuring the ratio of the produced $X(3872)$ and $\psi(2S)$ in inclusive $B^\pm$ decays. The decay kinematics will vary according to the invariant mass of the associated hadronic particles in the $B^\pm$ decays. We see the distribution of mass of the associated hadronic particles $m(X')$ in the $B^\pm$ decays in Figure H.1 for $\psi(2S)$ and $X(3872)$. From this, we decide to create three templates. The central template uses the ratio between all produced $\psi(2S)$ and $X(3872)$ events which satisfy the acceptance criteria. As a systematic variation, we then split the $X(3872)$ sample into two parts according to $m(X')$, one where X' is a kaon (where $m_K \approx 0.494$ GeV), and the other where $m(X') > 1.0$ GeV. The production ratios as a function of $p_T(J/\psi\pi^+\pi^-)$ are shown in Figure H.2(a). We then normalise the number of events in each of the samples prior to creating the ratio, effectively normalising the cross section times branching ratio of each of the samples to unity, and then re-produce the ratios, shown in Figure H.2(b). These templates can subsequently be used to fit non-prompt $X(3872)/\psi(2S)$ ratios to establish branching fraction ratios, assuming the non-prompt $\psi(2S)$ and $X(3872)$ are produced by a similar combination of b-hadrons.



(a) (b)

Fig. H.1 Invariant mass distribution of associated hadronic particles in inclusive $B^{\pm}$ decays to (a) $\psi(2S)$ and (b) $X(3872)$ in our truth MC samples. As a systematic variation for the production ratios, we split the $X(3872)$ sample according to $m(X')$ at the kaon mass and at > 1.0 GeV.



(a) (b)

Fig. H.2 (a) Ratios of non-prompt $X(3872)$ and $\psi(2S)$ production from $B^{\pm}$ decays as a function $p_T(J/\psi\pi\pi)$ for the three $m(X')$ cases, and (b) normalising the event yields of the samples before producing the ratios.

Appendix I

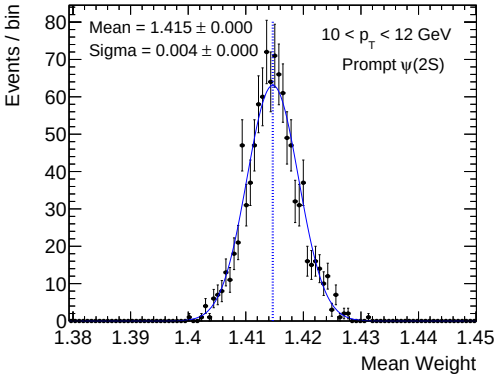
Systematic Studies

I.1 Trigger and Muon Reconstruction Systematics

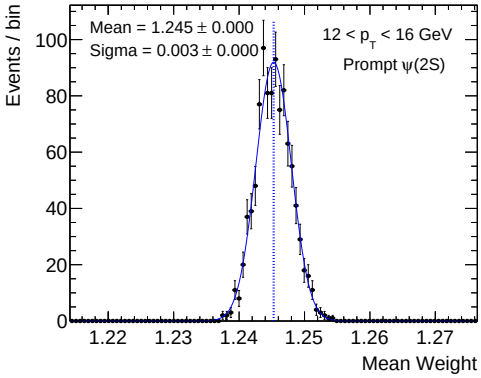
We evaluate the sytematic effects on measured cross sections due to uncertainties of the trigger and muon reconstruction efficiency maps. We build 1000 different trigger and muon reconstruction efficiency maps by varying each bin from its central value, using a normal distribution with a width described by the bin error. Then, separately for the prompt and non-prompt $\psi(2S)$ and $X(3872)$ signal regions for each analysis p_{T} bin, we create distributions of average weights in each signal region, using each of the 1000 maps. The distributions of average weights in each p_{T} bin can then be fitted with a Gaussian, where the width relative to the mean describes the systematic error. These errors are subsequently used in the main analysis.

The fitted average weight distributions for the muon reconstruction map variations are shown in Figure I.1 for prompt $\psi(2S)$ signal, and Figure I.2 for prompt $X(3872)$ signal. Practically identical results were measured for non-prompt signal. The fitted average weight distributions for the trigger efficiency map variations are shown in Figure I.3 for prompt $\psi(2S)$ signal and Figure I.4 for prompt $X(3872)$ signal. Again, results for non-prompt signal are almost identical.

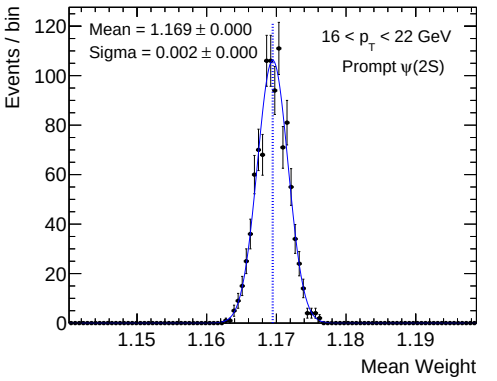
Muon Reconstruction



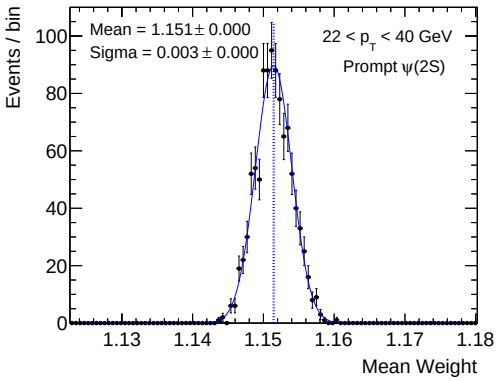
(a)



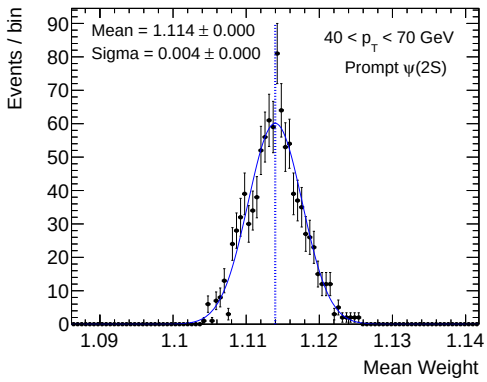
(b)



(c)

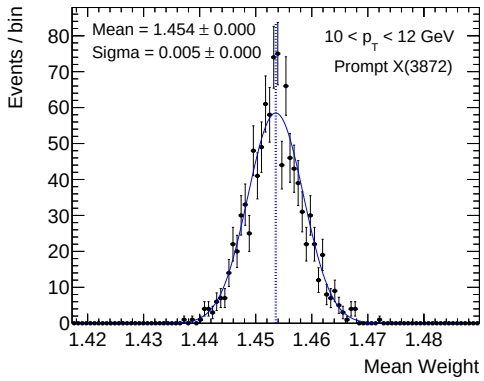


(d)

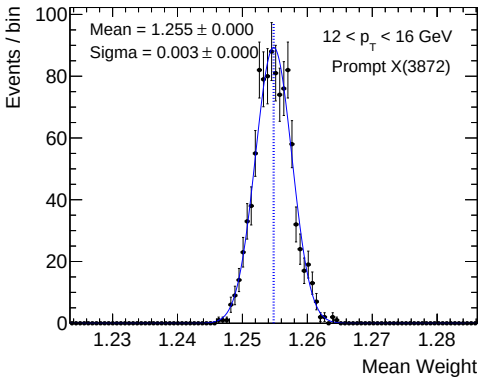


(e)

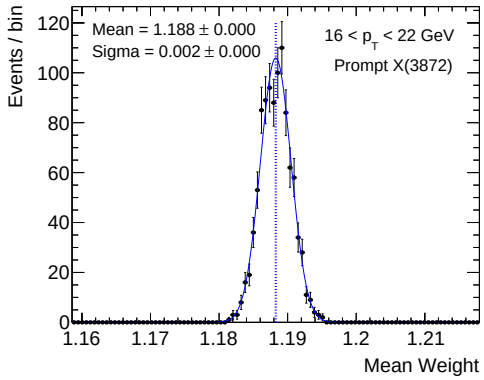
Fig. I.1 Distribution of average muon reconstruction efficiency correction weights fitted with a Gaussian, describing the systematic error for prompt $\psi(2S)$ signal, for the p_{T} bins (a) $10-12$ GeV, (b) $12-16$ GeV, (c) $16-22$ GeV, (d) $22-40$ GeV and (e) $40-70$ GeV.



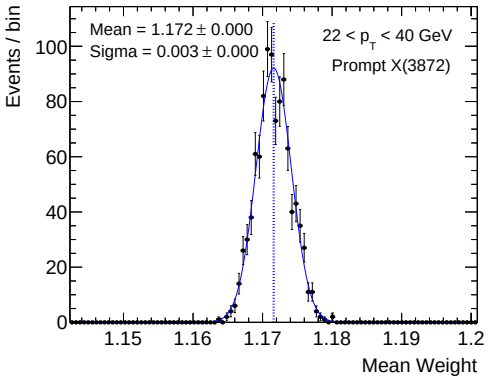
(a)



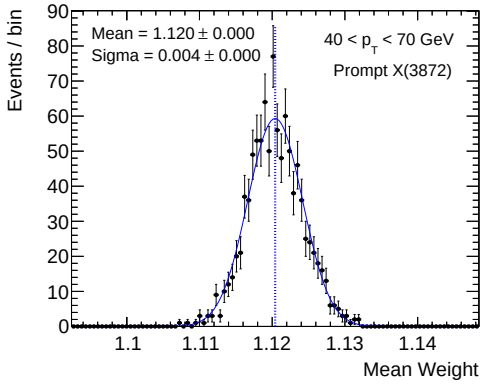
(b)



(c)



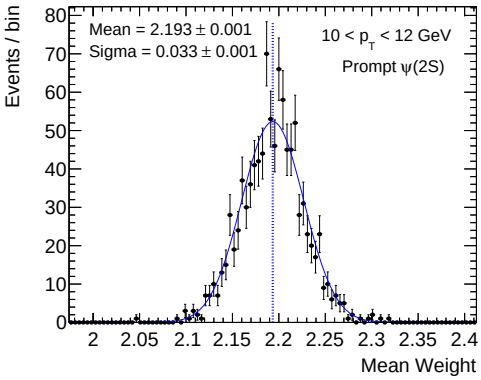
(d)



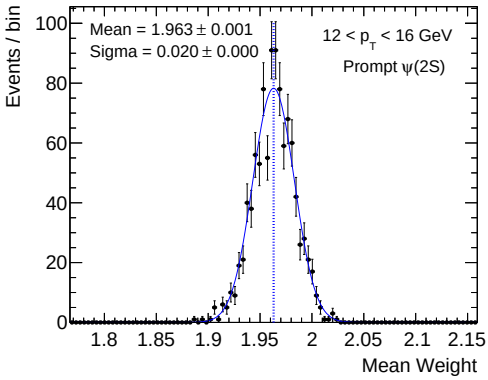
(e)

Fig. I.2 Distribution of average muon reconstruction efficiency correction weights fitted with a Gaussian, describing the systematic error for prompt $X(3872)$ signal, for the p_{T} bins (a) $10-12$ GeV, (b) $12-16$ GeV, (c) $16-22$ GeV, (d) $22-40$ GeV and (e) $40-70$ GeV.

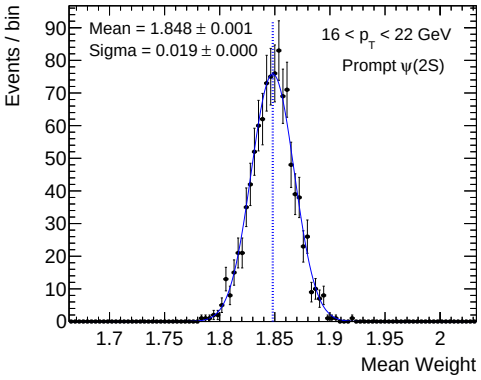
Trigger Efficiency



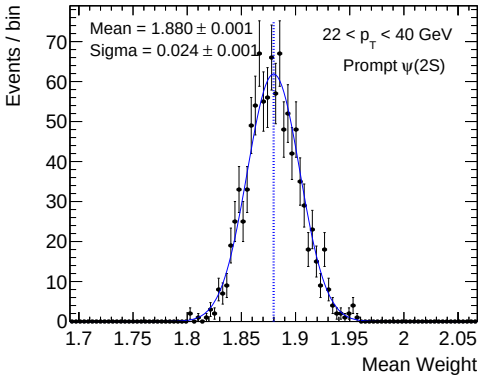
(a)



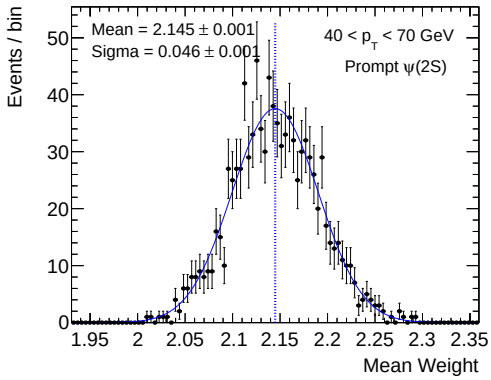
(b)



(c)

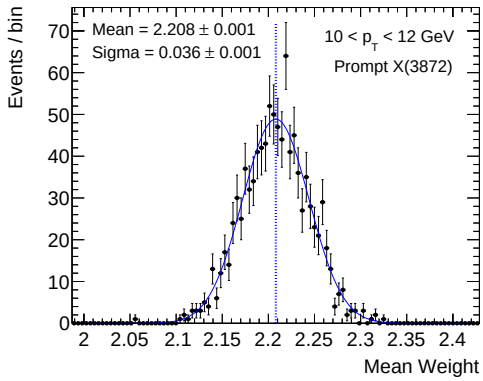


(d)

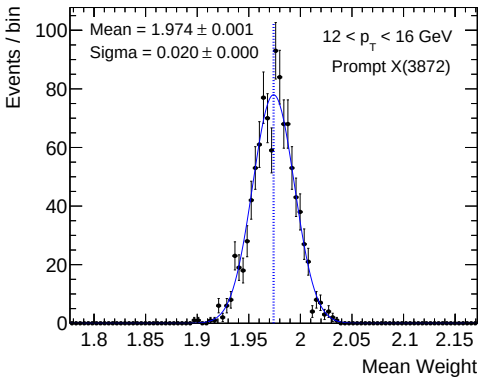


(e)

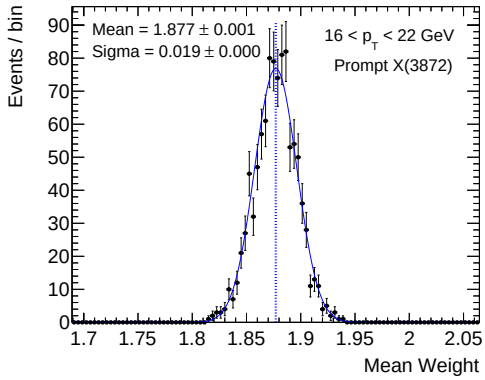
Fig. I.3 Distribution of average trigger efficiency correction weights fitted with a Gaussian, describing the systematic error for prompt $\psi(2S)$ signal, for the p_{T} bins (a) $10-12$ GeV, (b) $12-16$ GeV, (c) $16-22$ GeV, (d) $22-40$ GeV and (e) $40-70$ GeV.



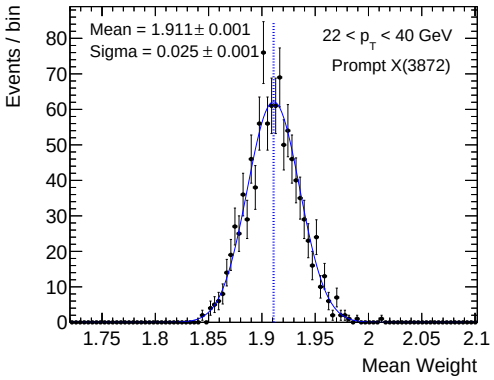
(a)



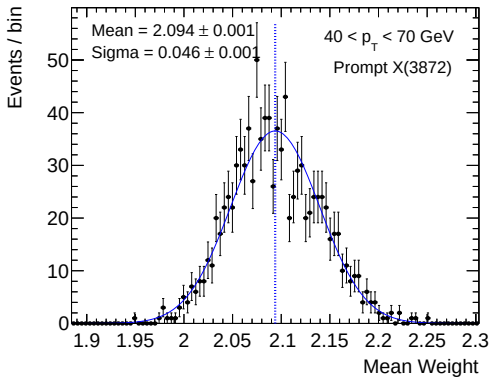
(b)



(c)



(d)



(e)

Fig. I.4 Distribution of average trigger efficiency correction weights fitted with a Gaussian, describing the systematic error for prompt $X(3872)$ signal, for the p_{T} bins (a) $10-12$ GeV, (b) $12-16$ GeV, (c) $16-22$ GeV, (d) $22-40$ GeV and (e) $40-70$ GeV.

I.2 Fit Model Variations

We vary the fit models to the invariant mass distributions of the $J/\psi\pi^+\pi^-$ candidates in windows of pseudoproper lifetime, for each analysis p_{T} bin. The variations of the fit models are described in Section 6.1.6. The p_{T} differential prompt and non-prompt $\psi(2S)$ cross sections according to the fit model variations are shown in Figure I.5, and for $X(3872)$ in Figure I.6. Each variation gives only a small difference from the central model, and the maximum difference in each p_{T} bin is used as the systematic error.

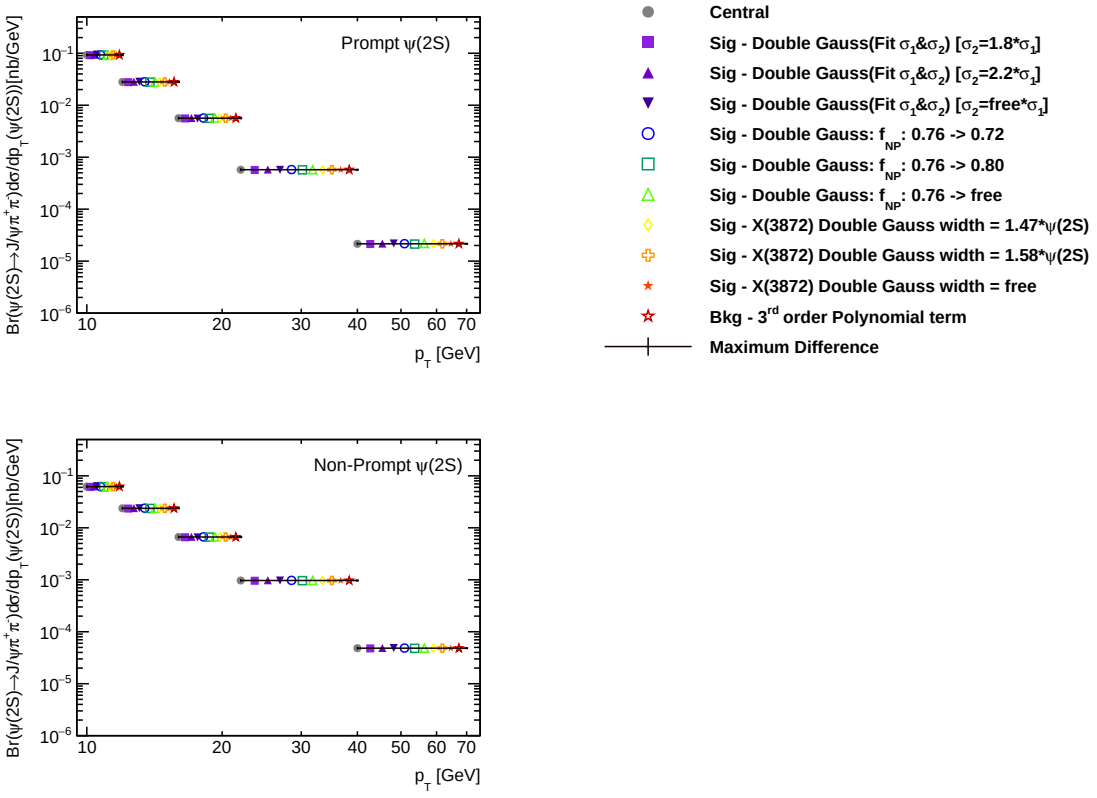


Fig. I.5 Prompt (top left) and non-prompt (bottom left) $\psi(2S)$ p_{T} differential cross sections for variations of fit model (described at the top right). The maximum difference from the central model in each p_{T} bin is used as the systematic error.

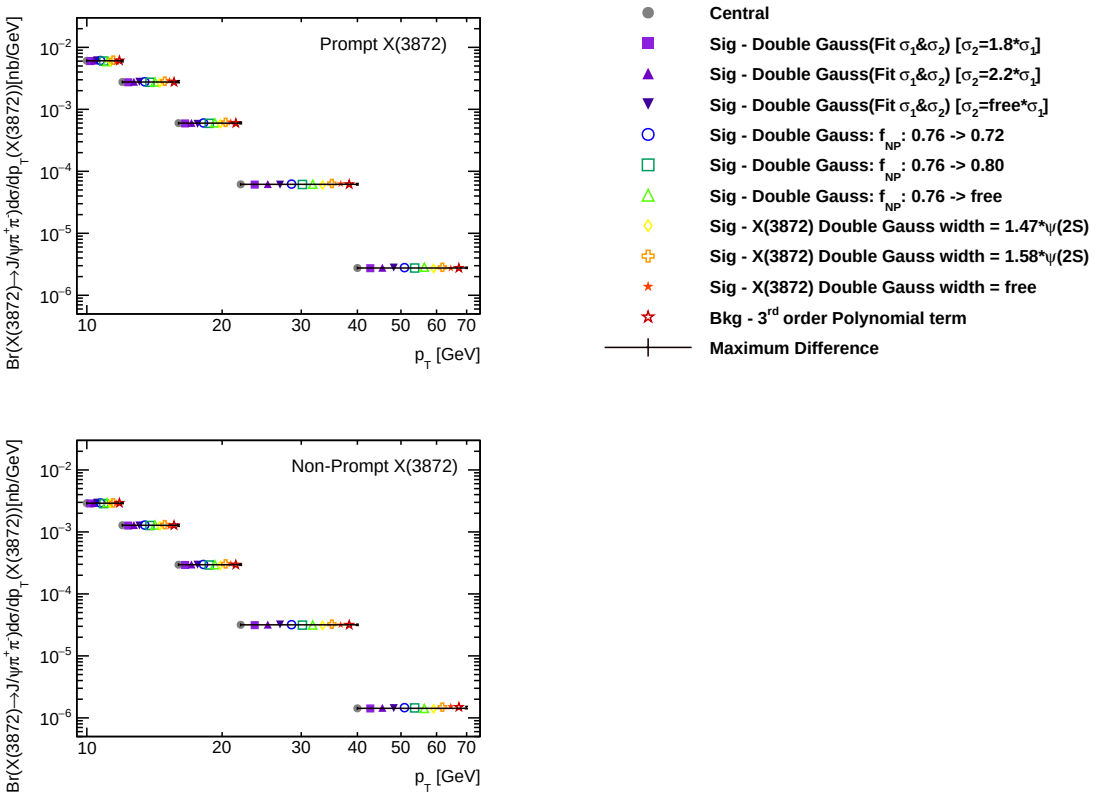


Fig. I.6 Prompt (top left) and non-prompt (bottom left) $X(3872)$ p_{T} differential cross sections for variations of fit model (described at the top right). The maximum difference from the central model in each p_{T} bin is used as the systematic error.

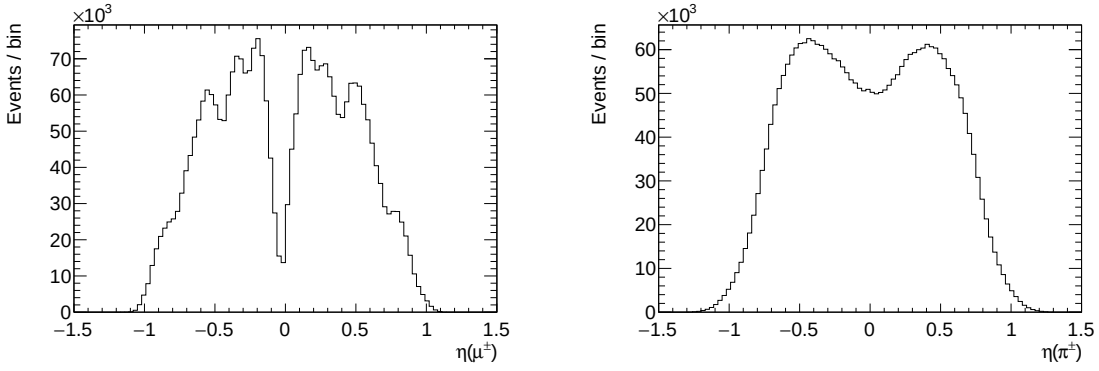
I.3 Effects due to z-displacement of Primary Vertices

The distribution of primary vertices along the z-axis of the detector might be expected to have an effect on the measurement of the cross section in a number of ways. The z-displacement of the vertices is measured to have some difference between data and simulation, see Figure C.9. The magnitude of the effect on pion reconstruction efficiency is measured in Appendix C.5 by reweighting the simulation to match the data z-displacement distribution, and the difference is measured to be within 0.5% of the central efficiency, which is used as an additional systematic.

The z-displacement will also cause a smearing of the acceptance criteria $|\eta_\mu| < 2.3$ and $|\eta_\pi| < 2.4$ of final muon and pion decay products. The degree of the smearing is estimated by modelling one of the decay products traversing the full width of the inner detector (ID), which has a radius $\approx$ 1m, evaluated at the edge of the acceptance $|\eta| < 2.3$. We can express the longitudinal distance z of the traversing particle by $r/\tan\theta$ with θ corresponding to the acceptance edge and r to the radius of the ID. The variation of the acceptance edge in pseudo-rapidity with respect to a change in z can be estimated as

$$\Delta\eta(z) = \left|\frac{\partial\eta(z)}{\partial z}\delta z\right| \tag{I.1}$$

where the error in z is given by δz, which is taken as the RMS of the z-displacement of the primary vertices measured in the data, shown in Figure C.9, and is equal to 48 mm. We estimate the z-displacement to cause only minimal smearing of the acceptance edge $\Delta\eta \approx 0.01$. In Figure I.7 we show η_μ and η_π distributions of the events used in the analysis. None of these events have final decay products close to the edge of the acceptance, and so this small smearing will have negligible effect.



(a) (b)

Fig. I.7 η distribution of (a) muons and (b) pions for $J/\psi\pi^+\pi^-$ candidates passing the selection criteria.

The magnitude of a similar effect is estimated for the smearing of the selection edge $|y(J/\psi\pi^+\pi^-)| < 0.75$. We estimate the smearing in rapidity using a method identical to the acceptance edge smearing estimation, except we use

$$\Delta y(z) = \left| \frac{\partial y(z)}{\partial z} \delta z \right| \tag{I.2}$$

where for this case the rapidity y depends of the invariant mass and p_T of the $J/\psi\pi^+\pi^-$ vertex. We estimate the size of the smearing Δy for four distinct cases. We must separate due to invariant mass of the $J/\psi\pi\pi$ vertices, for which we use the $\psi(2S)$ and $X(3872)$ PDG masses. We also separate into two extreme p_T cases, $p_T = 10$ GeV and 70 GeV. The invariant mass of the $J/\psi\pi^+\pi^-$ vertex is found to have no effect on the smearing to first order. The smearing is estimated as $\Delta y = [0.03, 0.04]$ for $p_T = [10, 70]$ GeV at $y = 0.75$.

The rapidity dependence in the area of interest for both $\psi(2S)$ and $X(3872)$ is expected to be fairly flat. Our estimates are based on Figures 15 and 16 from [91]. The rapidity dependence in each p_T bin can be parameterised as being $\propto (1 + ay^2)$ with $a \lesssim 0.1$. The correction factor due to rapidity interval smearing is then estimated to be $1 - 2a\Delta y^2$, which would mean that this effect is limited to no more than about 1%.

References

[1] Choi, S. K. et al., Belle Collaboration. Observation of a narrow charmonium - like state in exclusive $B^+ \to K^+\pi^+\pi^- J/\psi$ decays. *Phys. Rev. Lett.*, 91:262001, 2003. arXiv:hep-ex/0309032, doi:10.1103/PhysRevLett.91.262001.

[2] 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.*, D71:071103, 2005. arXiv:hep-ex/0406022, doi:10.1103/PhysRevD.71.071103.

[3] D. Acosta et al., CDF Collaboration. Observation of the narrow state $X(3872) \to J/\psi\pi^+\pi^-$ in $\bar{p}p$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev. Lett.*, 93:072001, 2004. arXiv:hep-ex/0312021, doi:10.1103/PhysRevLett.93.072001.

[4] V. M. Abazov et al., D0 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.*, 93:162002, 2004. arXiv:hep-ex/0405004, doi:10.1103/PhysRevLett.93.162002.

[5] Stephen Godfrey and Stephen L. Olsen. The Exotic XYZ Charmonium-like Mesons. *Ann. Rev. Nucl. Part. Sci.*, 58:51–73, 2008. arXiv:0801.3867, doi:10.1146/annurev.nucl.58.110707.171145.

[6] ATLAS Collaboration. Measurement of the production cross-section of $\psi(2S) \to J/\psi(\to \mu^+\mu^-)\pi^+\pi^-$ in pp collisions at $\sqrt{s} = 7$ TeV at ATLAS. *JHEP*, 1409:079, 2014. arXiv:1407.5532, doi:10.1007/JHEP09(2014)079.

[7] R. Aaij et al., LHCb Collaboration. Observation of $X(3872)$ production in pp collisions at $\sqrt{s} = 7$ TeV. *Eur. Phys. J.*, C72:1972, 2012. arXiv:1112.5310, doi:10.1140/epjc/s10052-012-1972-7.

[8] 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*, 04:154, 2013. arXiv:1302.3968, doi:10.1007/JHEP04(2013)154.

[9] ATLAS Collaboration. Production measurements of $\psi(2S)$ and $X(3872) \to J/\psi\pi^+\pi^-$ at $\sqrt{s} = 8$ TeV with the ATLAS detector, 2016. URL: http://inspirehep.net/record/1468552/.

[10] ATLAS Collaboration. Measurement of the differential cross-sections of prompt and non-prompt production of J/ψ and ψ(2S) in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS detector. *Eur. Phys. J.*, C76(5):283, 2016. arXiv:1512.03657, doi:10.1140/epjc/s10052-016-4050-8.

[11] M. Ablikim et al., BESIII Collaboration. Observation of $Z_c(3900)^0$ in $e^+e^- \to \pi^0\pi^0 J/\psi$. *Phys. Rev. Lett.*, 115(11):112003, 2015. arXiv:1506.06018, doi:10.1103/PhysRevLett.115.112003.

[12] R. Aaij et al., LHCb Collaboration. Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays. *Phys. Rev. Lett.*, 115:072001, 2015. arXiv:1507.03414, doi:10.1103/PhysRevLett.115.072001.

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

[14] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, 1964. doi:10.1103/PhysRevLett.13.508.

[15] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global Conservation Laws and Massless Particles. *Phys. Rev. Lett.*, 13:585–587, 1964. doi:10.1103/PhysRevLett.13.585.

[16] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716:1, 2012. arXiv:1207.7214, doi:10.1016/j.physletb.2012.08.020.

[17] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B*, 716:30, 2012. arXiv:1207.7235, doi:10.1016/j.physletb.2012.08.021.

[18] K. A. Olive et al. Review of Particle Physics. *Chin. Phys.*, C38:090001, 2014. doi:10.1088/1674-1137/38/9/090001.

[19] Mark Thomson. *Modern Particle Physics*. Cambridge University Press, Cambridge, 2015.

[20] J. J. Aubert et al., E598 Collaboration. Experimental Observation of a Heavy Particle J. *Phys. Rev. Lett.*, 33:1404–1406, Dec 1974. URL: http://link.aps.org/doi/10.1103/PhysRevLett.33.1404, doi:10.1103/PhysRevLett.33.1404.

[21] J. E. Augustin et al., SLAC-SP-017 Collaboration. Discovery of a narrow resonance in e^+e^- annihilation. *Phys. Rev. Lett.*, 33:1406–1408, Dec 1974. URL: http://link.aps.org/doi/10.1103/PhysRevLett.33.1406, doi:10.1103/PhysRevLett.33.1406.

[22] S. W. Herb et al. Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions. *Phys. Rev. Lett.*, 39:252–255, Aug 1977. URL: http://link.aps.org/doi/10.1103/PhysRevLett.39.252, doi:10.1103/PhysRevLett.39.252.

[23] Z. Metreveli. Charmonium Spectroscopy Below Open Flavor Threshold. *eConf*, C070805:16, 2007. arXiv:0710.1884.

[24] Hee Sok Chung, Jungil Lee, and Daekyoung Kang. Cornell Potential Parameters for S-wave Heavy Quarkonia. *J. Korean Phys. Soc.*, 52:1151–1154, 2008. arXiv:0803.3116, doi:10.3938/jkps.52.1151.

[25] Taichi Kawanai and Shoichi Sasaki. Potential description of the charmonium from lattice QCD. *AIP Conf. Proc.*, 1701:050022, 2016. arXiv:1503.05752, doi:10.1063/1.4938662.

[26] C. T. H. Davies, K. Hornbostel, G. P. Lepage, A. J. Lidsey, J. Shigemitsu, and J. H. Sloan. Precision charmonium spectroscopy from lattice QCD. *Phys. Rev.*, D52:6519–6529, 1995. arXiv:hep-lat/9506026, doi:10.1103/PhysRevD.52.6519.

[27] Angelo Esposito, Andrea L. Guerrieri, Fulvio Piccinini, Alessandro Pilloni, and Antonio D. Polosa. Four-Quark Hadrons: an Updated Review. *Int. J. Mod. Phys.*, A30:1530002, 2015. arXiv:1411.5997, doi:10.1142/S0217751X15300021.

[28] R. Aaij et al., LHCb Collaboration. Observation of $J/\psi\phi$ structures consistent with exotic states from amplitude analysis of $B^+ \to J/\psi\phi K^+$ decays. 2016. arXiv:1606.07895.

[29] Darren David Price and Vato Kartvelishvili. *Studies of quarkonium production and polarisation with early data at ATLAS.* PhD thesis, Lancaster U., Lancaster, 2008. Presented on 13 Nov 2008. URL: https://cds.cern.ch/record/1143624.

[30] Jean-Philippe Lansberg. *Quarkonium production at high-energy hadron colliders: A Systematic gauge-invariant approach to relativistic effects of J/ψ, ψ' and Υ production.* PhD thesis, Liege U., 2005. arXiv:hep-ph/0507175.

[31] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J.*, C63:189–285, 2009. arXiv:0901.0002, doi:10.1140/epjc/s10052-009-1072-5.

[32] J. P. Lansberg. J/ψ production at sqrt(s)=1.96 and 7 TeV: Color-Singlet Model, NNLO* and polarisation. *J. Phys.*, G38:124110, 2011. arXiv:1107.0292, doi:10.1088/0954-3899/38/12/124110.

[33] J. P. Lansberg. On the mechanisms of heavy-quarkonium hadroproduction. *Eur. Phys. J.*, C61:693–703, 2009. arXiv:0811.4005, doi:10.1140/epjc/s10052-008-0826-9.

[34] F. Abe et al., CDF Collaboration. J/ψ and $\psi(2S)$ production in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV. *Phys. Rev. Lett.*, 79:572–577, 1997. doi:10.1103/PhysRevLett.79.572.

[35] Francis Halzen and Satoshi Matsuda. Hadroproduction of quark flavors. *Phys. Rev. D*, 17:1344–1355, Mar 1978. doi:10.1103/PhysRevD.17.1344.

[36] J. F. Amundson, Oscar J. P. Eboli, E. M. Gregores, and F. Halzen. Quantitative tests of color evaporation: Charmonium production. *Phys. Lett.*, B390:323–328, 1997. arXiv:hep-ph/9605295, doi:10.1016/S0370-2693(96)01417-7.

[37] Geoffrey T. Bodwin, Eric Braaten, and Jungil Lee. Comparison of the color-evaporation model and the NRQCD factorization approach in charmonium production. *Phys. Rev.*, D72:014004, 2005. arXiv:hep-ph/0504014, doi:10.1103/PhysRevD.72.014004.

[38] N. Brambilla et al. Heavy quarkonium physics. 2004. arXiv:hep-ph/0412158.

[39] Peter L. Cho and Adam K. Leibovich. Color octet quarkonia production. *Phys. Rev.*, D53:150–162, 1996. arXiv:hep-ph/9505329, doi:10.1103/PhysRevD.53.150.

[40] Michael Kramer. Quarkonium production at high-energy colliders. *Prog. Part. Nucl. Phys.*, 47:141–201, 2001. arXiv:hep-ph/0106120, doi:10.1016/S0146-6410(01)00154-5.

[41] CMS Collaboration. J/ψ and $\psi(2S)$ production in pp collisions at $\sqrt{s} = 7$ TeV. *JHEP*, 02:011, 2012. arXiv:1111.1557, doi:10.1007/JHEP02(2012)011.

[42] R. Aaij et al., LHCb Collaboration. Measurement of J/ψ production in pp collisions at $\sqrt{s} = 7$ TeV. *Eur. Phys. J.*, C71:1645, 2011. arXiv:1103.0423, doi:10.1140/epjc/s10052-011-1645-y.

[43] R. Aaij et al., LHCb Collaboration. Measurement of $\psi(2S)$ meson production in pp collisions at $\sqrt{s}$=7 TeV. *Eur. Phys. J.*, C72:2100, 2012. arXiv:1204.1258, doi:10.1140/epjc/s10052-012-2100-4.

[44] Yan-Qing Ma, Kai Wang, and Kuang-Ta Chao. A complete NLO calculation of the J/ψ and ψ' production at hadron colliders. *Phys. Rev.*, D84:114001, 2011. arXiv:1012.1030, doi:10.1103/PhysRevD.84.114001.

[45] Mathias Butenschoen and Bernd A. Kniehl. Reconciling J/ψ production at HERA, RHIC, Tevatron, and LHC with NRQCD factorization at next-to-leading order. *Phys. Rev. Lett.*, 106:022003, 2011. arXiv:1009.5662, doi:10.1103/PhysRevLett.106.022003.

[46] S. P. Baranov, A. V. Lipatov, and N. P. Zotov. Prompt J/ψ production at LHC: new evidence for the k_t-factorization. *Phys. Rev.*, D85:014034, 2012. arXiv:1108.2856, doi:10.1103/PhysRevD.85.014034.

[47] S. P. Baranov, A. V. Lipatov, and N. P. Zotov. Prompt charmonia production and polarization at LHC in the NRQCD with k_T -factorization. Part I: $\psi(2S)$ meson. *Eur. Phys. J.*, C75(9):455, 2015. arXiv:1508.05480, doi:10.1140/epjc/s10052-015-3689-x.

[48] ATLAS Collaboration. Study of the $B_c^+ \to J/\psi D_s^+$ and $B_c^+ \to J/\psi D_s^{*+}$ decays with the ATLAS detector. *Eur. Phys. J.*, C76(1):4, 2016. arXiv:1507.07099, doi:10.1140/epjc/s10052-015-3743-8.

[49] Matteo Cacciari, Stefano Frixione, Nicolas Houdeau, Michelangelo L. Mangano, Paolo Nason, and Giovanni Ridolfi. Theoretical predictions for charm and bottom production at the LHC. *JHEP*, 10:137, 2012. arXiv:1205.6344, doi:10.1007/JHEP10(2012)137.

[50] V. G. Kartvelishvili, A. K. Likhoded, and V. A. Petrov. On the Fragmentation Functions of Heavy Quarks Into Hadrons. *Phys. Lett.*, B78:615–617, 1978. doi:10.1016/0370-2693(78)90653-6.

[51] Faccioli et al. Quarkonium production in the LHC era: a polarized perspective. *Phys. Lett.*, B736:98–109, 2014. arXiv:1403.3970, doi:10.1016/j.physletb.2014.07.006.

[52] Pietro Faccioli, Carlos Lourenco, Joao Seixas, and Hermine K. Wohri. Towards the experimental clarification of quarkonium polarization. *Eur. Phys. J.*, C69:657–673, 2010. arXiv:1006.2738, doi:10.1140/epjc/s10052-010-1420-5.

[53] Sandro Palestini. Angular distribution and rotations of frame in vector meson decays into lepton pairs. *Phys. Rev.*, D83:031503, 2011. arXiv:1012.2485, doi:10.1103/PhysRevD.83.031503.

[54] CMS Collaboration. Measurement of the prompt J/ψ and $\psi(2S)$ polarizations in pp collisions at $\sqrt{s} = 7$ TeV. *Phys. Lett.*, B727:381–402, 2013. arXiv:1307.6070, doi:10.1016/j.physletb.2013.10.055.

[55] R. Aaij et al., LHCb Collaboration. Measurement of J/ψ polarization in pp collisions at $\sqrt{s} = 7$ TeV. *Eur. Phys. J.*, C73(11):2631, 2013. arXiv:1307.6379, doi:10.1140/epjc/s10052-013-2631-3.

[56] R. Aaij et al., LHCb Collaboration. Measurement of $\psi(2S)$ polarisation in pp collisions at $\sqrt{s} = 7$ TeV. *Eur. Phys. J.*, C74(5):2872, 2014. arXiv:1403.1339, doi:10.1140/epjc/s10052-014-2872-9.

[57] R. Aaij et al., LHCb Collaboration. Quantum numbers of the $X(3872)$ state and orbital angular momentum in its $\rho^0 J/\psi$ decay. *Phys. Rev.*, D92(1):011102, 2015. arXiv:1504.06339, doi:10.1103/PhysRevD.92.011102.

[58] A. Tomaradze, S. Dobbs, T. Xiao, and Kamal K. Seth. Precision Measurement of the Mass of the D^{*0} Meson and the Binding Energy of the $X(3872)$ Meson as a $D^0\overline{D}^{*0}$ Molecule. *Phys. Rev.*, D91(1):011102, 2015. arXiv:1501.01658, doi:10.1103/PhysRevD.91.011102.

[59] Daniele Fasanella and Antonio Rossi. *Study of the X(3872) state with the CMS Experiment at LHC*. PhD thesis, Bologna U., 2012. URL: https://cds.cern.ch/record/1516171.

[60] C. Bignamini, B. Grinstein, F. Piccinini, A. D. Polosa, and C. Sabelli. Is the X(3872) Production Cross Section at Tevatron Compatible with a Hadron Molecule Interpretation? *Phys. Rev. Lett.*, 103:162001, 2009. arXiv:0906.0882, doi:10.1103/PhysRevLett.103.162001.

[61] Pierre Artoisenet and Eric Braaten. Production of the $X(3872)$ at the Tevatron and the LHC. *Phys. Rev.*, D81:114018, 2010. arXiv:0911.2016, doi:10.1103/PhysRevD.81.114018.

[62] Ce Meng, Hao Han, and Kuang-Ta Chao. $X(3872)$ and its production at hadron colliders. 2013. arXiv:1304.6710.

[63] Ce Meng and Kuang-Ta Chao. Decays of the $X(3872)$ and $\chi(c1)(2P)$ charmonium. *Phys. Rev.*, D75:114002, 2007. arXiv:hep-ph/0703205, doi:10.1103/PhysRevD.75.114002.

[64] D. V. Bugg. An Alternative fit to Belle mass spectra for $D\bar{D}$, $D^*\bar{D}^*$ and $\lambda_c\bar{\lambda}_c$. *J. Phys.*, G36:075002, 2009. arXiv:0811.2559, doi:10.1088/0954-3899/36/7/075002.

[65] L. Maiani, F. Piccinini, A. D. Polosa, and V. Riquer. Diquark-antidiquarks with hidden or open charm and the nature of $X(3872)$. *Phys. Rev.*, D71:014028, 2005. arXiv:hep-ph/0412098, doi:10.1103/PhysRevD.71.014028.

[66] T. Aaltonen et al., CDF Collaboration. Precision Measurement of the $X(3872)$ Mass in $J/\psi\pi^+\pi^-$ Decays. *Phys. Rev. Lett.*, 103:152001, 2009. arXiv:0906.5218, doi:10.1103/PhysRevLett.103.152001.

[67] Zhiqing Liu. Observation of $Z_c(3900)$ both by BESIII and Belle. In *Proceedings, 6th International Workshop on Charm Physics (Charm 2013)*, 2013. URL: https://inspirehep.net/record/1263197/files/arXiv:1311.0762.pdf, arXiv:1311.0762.

[68] B. Aubert et al., BaBar Collaboration. Search for a charged partner of the X(3872) in the B meson decay $B \to X^-K$, $X^- \to J/\psi\pi^-\pi^0$. *Phys. Rev.*, D71:031501, 2005. arXiv:hep-ex/0412051, doi:10.1103/PhysRevD.71.031501.

[69] CMS Collaboration. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008. doi:10.1088/1748-0221/3/08/S08004.

[70] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008. doi:10.1088/1748-0221/3/08/S08003.

[71] LHCb Collaboration. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008. doi:10.1088/1748-0221/3/08/S08005.

[72] ALICE Collaboration. The ALICE experiment at the CERN LHC. *JINST*, 3:S08002, 2008. doi:10.1088/1748-0221/3/08/S08002.

[73] LHCf Collaboration. The LHCf detector at the CERN Large Hadron Collider. *JINST*, 3:S08006, 2008. doi:10.1088/1748-0221/3/08/S08006.

[74] Fabienne Marcastel. CERN's Accelerator Complex. La chaîne des accélérateurs du CERN. Oct 2013. General Photo. URL: https://cds.cern.ch/record/1621583.

[75] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008. doi:10.1088/1748-0221/3/08/S08003.

[76] ATLAS 2011 - 2016 Luminosity. https://atlas.web.cern.ch/Atlas/GROUPS/DATAPREPARATION/PublicPlots/2011-2012/Luminosity/intlumivstime2011-2012DQ.png. Accessed: 2016-07-08.

[77] ATLAS 2011 and 2012 Luminosity. https://atlas.web.cern.ch/Atlas/GROUPS/DATAPREPARATION/PublicPlots/2016/DataSummary/figs/intlumivsyear.png. Accessed: 2016-07-08.

[78] ATLAS Collaboration. ATLAS Insertable B-Layer Technical Design Report. Technical Report CERN-LHCC-2010-013. ATLAS-TDR-19, CERN, Geneva, Sep 2010. URL: https://cds.cern.ch/record/1291633.

[79] ATLAS Collaboration. ATLAS Transition Radiation Tracker (TRT): Straw tube gaseous detectors at high rates. *Nucl. Instrum. Meth.*, A732:277–280, 2013. doi:10.1016/j.nima.2013.07.020.

[80] ATLAS Collaboration. Development and Construction of Large Size Signal Electrodes for the ATLAS Electromagnetic Calorimeter. *Nucl. Instrum. Methods Phys. Res., A*, 539(CERN-PH-EP-2004-019):558–594. 43 p, May 2004. URL: https://cds.cern.ch/record/737956.

[81] ATLAS Collaboration. Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics. 2009. arXiv:0901.0512.

[82] ATLAS Collaboration. Measurement of the ATLAS di-muon trigger efficiency in proton-proton collisions at 7 TeV. In *Proceedings, 31st International Conference on Physics in collisions (PIC 2011)*, 2011. URL: https://inspirehep.net/record/946658/files/arXiv:1111.4329.pdf, arXiv:1111.4329.

[83] B-physics trigger public results. http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/BPHYSICS/PUBLIC/DiMu_mas_diff_Triggers.eps. Accessed: 2016-06-29.

[84] ATLAS Collaboration. Measurement of Υ Production in 7 TeV *pp* Collisions at ATLAS. Technical Report ATL-COM-PHYS-2012-208, CERN, Geneva, Feb 2012. URL: https://cds.cern.ch/record/1426297.

[85] ATLAS Collaboration. Architecture and Performance of the Inner Detector Trigger of the ATLAS detector. Geneva, 2012. CERN. URL: http://cdsweb.cern.ch/record/1458060/files/proceedingForCHEP2012.pdf.

[86] ATLAS Collaboration, B. Resende. Muon identification algorithms in ATLAS. *PoS*, EPS-HEP2009:431, 2009.

[87] ATLAS Collaboration. Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data. *Eur. Phys. J.*, C74(11):3130, 2014. arXiv:1407.3935, doi:10.1140/epjc/s10052-014-3130-x.

[88] ATLAS Collaboration. Measurement of the differential cross-sections of prompt and non-prompt production of J/ψ and $\psi(2S)$ in *pp* collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS detector . ATLAS-CONF-2015-024, 2015. URL: https://cds.cern.ch/record/2031453.

[89] Choi, S. K. et al., Belle Collaboration. Bounds on the width, mass difference and other properties of $X(3872) \to \pi^+\pi^- J/\psi$ decays. *Phys. Rev.*, D84:052004, 2011. arXiv:1107.0163, doi:10.1103/PhysRevD.84.052004.

[90] ATLAS Collaboration. Measurement of Upsilon production in 7 TeV *pp* collisions at ATLAS. *Phys. Rev. D*, 87:052004, 2013. arXiv:1211.7255, doi:10.1103/PhysRevD.87.052004.

[91] ATLAS Collaboration. Measurement of J/ψ and $\psi(2S)$ production in $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector at LHC. Technical Report ATL-COM-PHYS-2014-1441, CERN, Geneva, Nov 2014. URL: https://cds.cern.ch/record/1967399.

[92] LHCb Collaboration. A Study of the Angular Properties of the $X(3872) \to J/\psi\pi^+\pi^-$ Decay . LHCb-PUB-2010-003, 2010. URL: https://cds.cern.ch/record/1236331/.

[93] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852–867, 2008. arXiv:0710.3820, doi:10.1016/j.cpc.2008.01.036.

[94] D. J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth.*, A462:152–155, 2001. doi:10.1016/S0168-9002(01)00089-4.

[95] ATLAS Collaboration. The ATLAS Simulation Infrastructure. *Eur. Phys. J. C*, 70:823, 2010. arXiv:1005.4568, doi:10.1140/epjc/s10052-010-1429-9.

[96] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003. doi:10.1016/S0168-9002(03)01368-8.

[97] John Allison et al. Geant4 developments and applications. *IEEE Trans. Nucl. Sci.*, 53:270, 2006. doi:10.1109/TNS.2006.869826.

[98] Mikhail B. Voloshin and Valentin I. Zakharov. Measuring QCD Anomalies in Hadronic Transitions Between Onium States. *Phys. Rev. Lett.*, 45:688, 1980. doi:10.1103/PhysRevLett.45.688.

[99] A. V. Berezhnoy and A. K. Likhoded. Relative Yield of Heavy Hadrons as a Function of the Transverse Momentum in LHC Experiments. *Phys. Atom. Nucl.*, 78(2):292–300, 2015. doi:10.1134/S1063778815020106.

[100] ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC. *Eur. Phys. J.*, C76(12):653, 2016. arXiv:1608.03953, doi:10.1140/epjc/s10052-016-4466-1.

[101] ATLAS Collaboration. Flavour tagged time-dependent angular analysis of the $B_s^0 \to J/\psi\phi$ decay and extraction of $\Delta\Gamma_s$ and the weak phase ϕ_s in ATLAS. *Phys. Rev. D*, 90:052007, 2014. arXiv:1407.1796, doi:10.1103/PhysRevD.90.052007.

[102] CMS Collaboration. Measurement of J/ψ and $\psi(2S)$ Prompt Double-Differential Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV. *Phys. Rev. Lett.*, 114(19):191802, 2015. arXiv:1502.04155, doi:10.1103/PhysRevLett.114.191802.

[103] G. Bauer. The $X(3872)$ at CDF II. *Int. J. Mod. Phys.*, A20:3765–3767, 2005. arXiv:hep-ex/0409052, doi:10.1142/S0217751X05027552.

[104] CMS Collaboration. J/ψ and $\psi(2S)$ production in *pp* collisions at $\sqrt{s} = 7$ TeV. *JHEP*, 02:011, 2012. arXiv:1111.1557, doi:10.1007/JHEP02(2012)011.

[105] J. Z. Bai et al., BES Collaboration. $\psi(2S) \to \pi^+\pi^- J/\psi$ decay distributions. *Phys. Rev.*, D62:032002, 2000. arXiv:hep-ex/9909038, doi:10.1103/PhysRevD.62.032002.

[106] R. Aaij et al., LHCb Collaboration. Measurement of $\psi(2S)$ meson production in pp collisions at $\sqrt{s}$=7 TeV. *Eur. Phys. J.*, C72:2100, 2012. arXiv:1204.1258, doi:10.1140/epjc/s10052-012-2100-4.

[107] A. Abulencia et al., CDF Collaboration. Measurement of the dipion mass spectrum in $X(3872) \to J/\psi\pi^+\pi^-$ decays. *Phys. Rev. Lett.*, 96:102002, 2006. arXiv:hep-ex/0512074, doi:10.1103/PhysRevLett.96.102002.

[108] ATLAS Collaboration. Measurement of the branching fractions for the $B_c^+ \to J/\psi D_s^+$ and $B_c^+ \to J/\psi D_s^{*+}$ decays with the ATLAS detector: supporting note. Technical Report ATL-COM-PHYS-2014-265, CERN, Geneva, Apr 2014. URL: https://cds.cern.ch/record/1693108.