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

The ATLAS Collaboration

DOI
10.1007/JHEP01(2017)117

Publication date
2017

Document Version
Final published version

Published in
JHEP

License
CC BY

Citation for published version (APA):
The ATLAS Collaboration (2017). Measurements of $\psi(2S)$ and $X(3872) J/\psi \pi^+\pi^-$ production in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. JHEP, 2017(1), [118].
https://doi.org/10.1007/JHEP01(2017)117
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

The ATLAS collaboration

E-mail: atlas.publications@cern.ch

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 $R_B = \frac{B(B \to X(3872) + \text{any})B(X(3872) \to J/\psi\pi^+\pi^-)}{B(B \to \psi(2S) + \text{any})B(\psi(2S) \to J/\psi\pi^+\pi^-)} = (3.95 \pm 0.32(\text{stat}) \pm 0.08(\text{sys})) \times 10^{-2}$, while separating short- and long-lived contributions, assuming that the short-lived component is due to $B_c$ decays, gives $R_B = (3.57 \pm 0.33(\text{stat}) \pm 0.11(\text{sys})) \times 10^{-2}$, 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 theoretical predictions.

KEYWORDS: B physics, Hadron-Hadron scattering (experiments)

ArXiv ePrint: 1610.09303
1 Introduction

The hidden-charm state \( X(3872) \) was discovered by the Belle Collaboration in 2003 [1] through its decay to \( J/\psi \pi^+\pi^- \) in the exclusive decay \( B^\pm \rightarrow K^\pm J/\psi \pi^+\pi^- \). Its existence was subsequently confirmed by CDF [2] through its production in \( p\bar{p} \) collisions, and its production was also observed by the BaBar [3] and D0 [4] experiments shortly after. CDF determined [5] that the only possible quantum numbers for \( X(3872) \) were \( J^{PC} = 1^{++} \) and \( 2^{++} \). At the LHC, the \( X(3872) \) was first observed by the LHCb Collaboration [6], which finally confirmed its quantum numbers to be \( 1^{++} \) [7]. A particularly interesting aspect of the \( X(3872) \) is the closeness of its mass, \( 3871.69 \pm 0.17 \text{ MeV} \) [8], to the \( D^0 \bar{D}^{*0} \) threshold, such that it was hypothesised to be a \( D^0 \bar{D}^{*0} \) molecule with a very small binding energy [9]. A cross-section measurement of promptly produced \( X(3872) \) was performed by CMS [10] as a function of \( p_T \), and showed the non-relativistic QCD (NRQCD) prediction [11] for prompt \( X(3872) \) production, assuming a \( D^0 \bar{D}^{*0} \) molecule, to be too high, although the shape of the \( p_T \) dependence was described fairly well. 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, was adopted in conjunction with the next-to-leading-order (NLO) NRQCD model and fitted to CMS data, showing good agreement [12].
ATLAS previously observed the $X(3872)$ state while measuring the cross section of prompt and non-prompt $\psi(2S)$ meson production in the $J/\psi\pi^+\pi^-$ decay channel with 2011 data at a centre-of-mass energy $\sqrt{s} = 7$ TeV \cite{13}. ATLAS later performed cross-section measurements for $J/\psi$ and $\psi(2S)$ decaying through the $\mu^+\mu^-$ channel at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV \cite{14}.

In this analysis, a measurement of the differential cross sections for the production of $\psi(2S)$ and $X(3872)$ states in the decay channel $J/\psi\pi^+\pi^-$ is performed, using 11.4 fb$^{-1}$ of proton-proton collision data collected by the ATLAS experiment at the LHC at $\sqrt{s} = 8$ TeV. The $J/\psi\pi^+\pi^-$ final state allows good invariant mass resolution through the use of a constrained fit, and provides a straightforward way of comparing the production characteristics of $\psi(2S)$ and $X(3872)$ states, which are fairly close in mass. The prompt and non-prompt contributions for $\psi(2S)$ and $X(3872)$ are separated, based on an analysis of the displacement of the production vertex. Non-prompt production fractions for $\psi(2S)$ and $X(3872)$ are measured, and the $X(3872)/\psi(2S)$ production ratios are measured separately for prompt and non-prompt components. The non-prompt results show that while the non-prompt $\psi(2S)$ data is readily described by a traditional single-effective-lifetime fit, there are indications in the non-prompt $X(3872)$ data which suggest introducing a two-lifetime fit with both a short-lived and long-lived component. Results are presented here based on both the single- and two-lifetime fit models. In the two-lifetime case, assuming that the short-lived non-prompt component of $X(3872)$ originates from the decays of $B_c$ mesons, the best-fit fractional contribution of the $B_c$ component is determined. The distributions of the dipion invariant mass in $\psi(2S) \rightarrow J/\psi\pi^+\pi^-$ and $X(3872) \rightarrow J/\psi\pi^+\pi^-$ decays are also measured. Comparisons are made with theoretical models and available experimental data.

2 The ATLAS detector

The ATLAS detector \cite{15} is a cylindrical, forward-backward symmetric, general-purpose particle detector. The innermost part of the inner detector (ID) comprises pixel and silicon microstrip (SCT) tracking technology for high-precision measurements, complemented further outwards by the transition radiation tracker (TRT). The inner detector spans the pseudorapidity$^1$ range $|\eta| < 2.5$ and is immersed in a 2 T axial magnetic field. Enclosing the ID and the solenoidal magnet are the electromagnetic and hadronic sampling calorimeters, which provide good containment of the electromagnetic and hadronic showers in order to limit punch-through into the muon spectrometer (MS). Surrounding the calorimeters, the MS covers the rapidity range $|\eta| < 2.7$ and utilises three air-core toroidal magnets, each consisting of eight coils, generating a magnetic field providing 1.5–7.5 T⋅m of bending power. The MS consists of fast-trigger detectors (thin-gap chambers and resistive plate chambers) as well as precision-measurement detectors (monitored drift tubes and cathode strip chambers).

---

$^1$ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the $z$-axis along the beam pipe. The $x$-axis points from the IP to the centre of the LHC ring, and the $y$-axis points upward. Polar coordinates $(r, \phi)$ are used in the transverse plane, $\phi$ being the azimuthal angle around the $z$-axis. The pseudorapidity $\eta$ is defined in terms of the polar angle $\theta$ as $\eta = -\ln\tan(\theta/2)$, and the transverse momentum $p_T$ is defined as $p_T = p \sin \theta$. The rapidity $y$ is defined as $y = 0.5 \ln[(E + p_z)/(E - p_z)]$, where $E$ and $p_z = p \cos \theta$ refer to energy and longitudinal momentum, respectively.
The ATLAS detector uses a three-level trigger system in order to select 300 Hz of interesting events to be written out from the 20 MHz of proton bunch collisions. This analysis uses a dimuon trigger with the lowest available transverse momentum threshold of 4 GeV for each muon. The level-1 muon trigger finds regions-of-interest (RoIs) by searching for hit coincidences in layers of the muon trigger detectors inside predefined geometrical windows. The software-based two-stage high-level trigger (HLT) is seeded by the level-1 RoIs, and uses more precise MS and ID information to reconstruct the final muon trigger objects with a resolution comparable to the full offline reconstruction.

3 Event selection

Events used in this analysis are triggered by a pair of muons successfully fitted to a common vertex. The data sample corresponds to an integrated luminosity of 11.4 fb$^{-1}$ [16], collected at a proton-proton collision energy $\sqrt{s} = 8$ TeV. Each muon candidate reconstructed offline is required to have good spatial matching to a trigger object, satisfying

$$R_q^2 + (\Delta \phi)^2 < 0.01.$$  

Events where two oppositely charged muon candidates are reconstructed with pseudorapidity $|\eta| < 2.3$ and transverse momenta $p_T^\mu > 4$ GeV are kept for further analysis only if the invariant mass of the dimuon system falls within ±120 MeV of the mass of the $J/\psi$ meson, $m(J/\psi) = 3096.916 \pm 0.011$ MeV [8].

The two muon tracks are fitted to a common vertex with a loose cut on fit quality, $\chi^2 < 200$. The dimuon invariant mass is then constrained to the $J/\psi$ mass, and the four-track vertex fit of the two muon tracks and pairs of non-muon tracks is performed to find $J/\psi \pi^+\pi^-$ candidates. The two non-muon tracks are assigned pion masses, and are required to have opposite charges and to satisfy the conditions $p_T^\pi > 0.6$ GeV, $|\eta|$ < 2.4. Four-track candidates with fit $\chi^2$ probability $P(\chi^2) < 4\%$ are discarded.

Only $J/\psi \pi^+\pi^-$ combinations with rapidity $y$ within the range $|y| < 0.75$ are considered in this analysis, with most of the contributing tracks measured within the barrel part of the detector $|\eta| \lesssim 1$ where the tracking resolution is optimal. Then the transverse momenta of the $J/\psi \pi^+\pi^-$ candidates are required to be within the range $10$ GeV < $p_T < 70$ GeV.

Further selection requirements are applied to the remaining $J/\psi \pi^+\pi^-$ combinations:

$$\Delta R(J/\psi, \pi^\pm) < 0.5, \quad Q < 0.3 \text{ GeV},$$  

where $\Delta R(J/\psi, \pi^\pm)$ is the angular distance between the momenta of the dimuon system and each pion candidate, while $Q \equiv m(J/\psi \pi^+\pi^-) - m(J/\psi) - m(\pi^+\pi^-)$. Here $m(J/\psi \pi^+\pi^-)$ and $m(\pi^+\pi^-)$ are the fitted invariant masses of the $\mu^+\mu^-\pi^+\pi^-$ and the dipion system, respectively. These requirements are found to be > 90% efficient for the signal from $\psi(2S)$ and $X(3872)$ decays, while significantly suppressing the combinatorial background.

The invariant mass distribution of the dimuons contributing to the selected $J/\psi \pi^+\pi^-$ combinations is shown in figure 1(a) between the dashed vertical lines. The distribution is fitted with the sum of a second-order polynomial background and a double-Gaussian function, which contains about 3.6 M $J/\psi$ candidates. The invariant mass distribution of the $J/\psi \pi^+\pi^-$ candidates selected for further analysis is presented in figure 1(b). The fitted
function is the sum of a fourth-order polynomial background and two double-Gaussian functions. The double-Gaussian functions for $\psi(2S)$ and $X(3872)$ peaks and a fourth-order polynomial for the background. The $X(3872)$ mass range is highlighted in the inset.

Monte Carlo (MC) simulation is used to study the selection and reconstruction efficiencies. The MC samples with $b$-hadron production and decays are generated with 

Pythia 6.4 [17], complemented, where necessary, with a dedicated extension for $B_c$ production based on calculations from refs. [18–21]. The decays of $b$-hadrons are then simulated with EvtGen [22]. The generated events are passed through a full simulation of the detector using the ATLAS simulation framework [23] based on Geant4 [24, 25] and processed with the same software as that used for the data.

4 Analysis 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, 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 [13, 14] a per-candidate weight $\omega$ was calculated as

$$\omega = \left[ A(p_T, y) \cdot \epsilon_{\text{trig}}(p_T^\pi, \eta^\pi, y_{J/\psi}) \cdot \epsilon^\mu(p_T^{\mu^+}, \eta^{\mu^+}) \cdot \epsilon^\mu(p_T^{\mu^-}, \eta^{\mu^-}) \cdot \epsilon^\pi(p_T^{\pi^+}, \eta^{\pi^+}) \cdot \epsilon^\pi(p_T^{\pi^-}, \eta^{\pi^-}) \right]^{-1}. \tag{4.1}$$

Here, $p_T$ and $y$ stand for the transverse momentum and rapidity of the $J/\psi \pi^+\pi^-$ candidate, $y_{J/\psi}$ is the rapidity of the $J/\psi$ candidate, while $p_T^\pi$, $p_T^{\mu^\pm}$ and $\eta^\pi$ are transverse

![Figure 1](attachment:image.png)

**Figure 1.** (a) The invariant mass distribution of the $J/\psi$ candidates satisfying all selection criteria except the $\pm 120$ MeV $J/\psi$ mass window requirement indicated here by the dotted vertical lines. The curve shows the result of a fit with a double-Gaussian function 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 requirements. The curve shows the results of the fit using double-Gaussian functions 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.
momenta and pseudorapidities of the respective pions and muons. The trigger efficiency $\epsilon_{\text{trig}}$ and the muon reconstruction efficiency $\epsilon_{\mu}$ were obtained using data-driven tag-and-probe methods described in refs. [14, 26]. The pion reconstruction efficiency $\epsilon_{\pi}$ is obtained through MC simulations using the method described in ref. [13].

The acceptance $A(p_T, y)$ 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 3. The acceptance map is created using generator-level simulation, with small reconstruction-level corrections applied at a later stage (see ref. [14] for more details). The different quantum numbers of the $\psi(2S)$ and $X(3872)$ ($J^{PC} = 1^- -$ and $1^{++}$, respectively) cause a difference in the expected dependence of the acceptance on the spin-alignments of the two states. The cross sections measured in this paper are obtained assuming no spin-alignment, but appropriate sets of correction factors for a number of extreme spin-alignment scenarios are calculated and presented in appendix A for each $p_T$ bin, separately for $\psi(2S)$ and $X(3872)$.

The efficiencies of the reconstruction-quality requirements and the background-suppression requirements described in section 3 are determined using MC simulations, and the corrections are applied in each of the $p_T$ bins, separately for $\psi(2S)$ and $X(3872)$. These efficiencies are found to vary between 84% and 95%. The simulated distributions are reweighted to match the data, and values with and without reweighting are used to estimate systematic uncertainties (see section 6).

In order to separate prompt production of the $\psi(2S)$ and $X(3872)$ states from the non-prompt production occurring via 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 $\tau$, defined as

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

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

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

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-particle tracks with $p_T > 0.4$ GeV 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 an analysis of the lifetime resolution and lifetime dependence of the signal, four lifetime intervals were defined:

- $w_0 : -0.3 \, \text{ps} < \tau(J/\psi\pi\pi) < 0.025 \, \text{ps},$
- $w_1 : 0.025 \, \text{ps} < \tau(J/\psi\pi\pi) < 0.3 \, \text{ps},$
- $w_2 : 0.3 \, \text{ps} < \tau(J/\psi\pi\pi) < 1.5 \, \text{ps},$
- $w_3 : 1.5 \, \text{ps} < \tau(J/\psi\pi\pi) < 15.0 \, \text{ps}.$
Figure 2. 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) $[12, 16]$ GeV and (b) $[22, 40]$ GeV. Shown underneath the fits are the corresponding pull distributions, with respective values of $\chi^2$ per degree of freedom for each fit.

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. These distributions are shown in figure 2 for representative $p_T$ bins.
The yields extracted from the fits are shown in table 1.

\[
\begin{array}{|c|c|c|c|c|}
\hline
\tau \text{ window} & 10-12 & 12-16 & 16-22 & 22-40 & 40-70 \\
\hline
w_0 & 17.48 \pm 0.36 & 11.03 \pm 0.11 & 3.53 \pm 0.03 & 1.14 \pm 0.01 & 0.078 \pm 0.004 \\
w_1 & 14.07 \pm 0.37 & 9.04 \pm 0.10 & 2.94 \pm 0.03 & 1.01 \pm 0.01 & 0.071 \pm 0.003 \\
w_2 & 9.13 \pm 0.29 & 7.04 \pm 0.09 & 2.97 \pm 0.03 & 1.27 \pm 0.01 & 0.104 \pm 0.004 \\
w_3 & 6.74 \pm 0.16 & 5.21 \pm 0.06 & 2.22 \pm 0.02 & 0.94 \pm 0.01 & 0.081 \pm 0.003 \\
\hline
\end{array}
\]

Table 1. Fitted yields of \( \psi(2S) \) in bins of pseudo-proper lifetime and \( p_T \). Uncertainties are statistical only.

\[
\begin{array}{|c|c|c|c|c|}
\hline
\tau \text{ window} & 10-12 & 12-16 & 16-22 & 22-40 & 40-70 \\
\hline
w_0 & 10.8 \pm 2.3 & 10.55 \pm 0.76 & 3.53 \pm 0.26 & 1.19 \pm 0.11 & 0.093 \pm 0.030 \\
w_1 & 9.3 \pm 2.7 & 8.21 \pm 0.71 & 2.60 \pm 0.24 & 0.72 \pm 0.11 & 0.039 \pm 0.023 \\
w_2 & 4.1 \pm 1.7 & 3.83 \pm 0.63 & 1.29 \pm 0.21 & 0.45 \pm 0.10 & 0.036 \pm 0.023 \\
w_3 & 2.06 \pm 0.81 & 2.09 \pm 0.34 & 0.98 \pm 0.13 & 0.30 \pm 0.06 & 0.020 \pm 0.014 \\
\hline
\end{array}
\]

Table 2. Fitted yields of \( X(3872) \) in bins of pseudo-proper lifetime and \( p_T \). Uncertainties are statistical only.

In order to determine the yields of the \( \psi(2S) \) and \( X(3872) \) signals, the distributions are fitted in each lifetime interval to the function:

\[
f(m) = Y^\psi \left( f_1 G_1^\psi(m) + (1 - f_1) G_2^\psi(m) \right) + Y^X \left( f_1 G_1^X(m) + (1 - f_1) G_2^X(m) \right) + N(m - m_{th})^{P_1 e^{P_2(m-m_{th})}} P(m - m_{th}),
\]

where the threshold mass \( m_{th} = m_{J/\psi} + 2m_\pi = 3376.06 \text{ MeV} \). The \( \psi(2S) \) and \( X(3872) \) signal yields \( Y^\psi \) and \( Y^X \), coefficients of the second-order polynomial \( P \), parameters \( P_1 \) and \( P_2 \), and the normalisation of the background term \( N \), are determined from the fits. Signal peaks for \( \psi(2S) \) and \( X(3872) \) are described by normalised double-Gaussian functions with common means: \( G_1^\psi(m) \) and \( G_1^X(m) \) are the narrower Gaussian functions with respective widths \( \sigma_\psi \) and \( \sigma_X \), while \( G_2^\psi(m) \) and \( G_2^X(m) \) are wider Gaussian functions with widths \( 2\sigma_\psi \) and \( 2\sigma_X \). The fraction of the narrower Gaussian function \( f_1 \) is assumed to be the same for \( \psi(2S) \) and \( X(3872) \), while the widths \( \sigma_\psi \) and \( \sigma_X \) are related by \( \sigma_X = \kappa \sigma_\psi \). The parameters \( f_1 \) and \( \kappa \) are fixed for the main fits to the values \( f_1 = 0.76 \pm 0.04, \kappa = 1.52 \pm 0.05 \) as determined from a fit applied in the range \( 16 \text{ GeV} < p_T < 70 \text{ GeV} \), which offers a better signal-to-background ratio than the full range, and is varied within these errors in the systematic uncertainty studies. 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 1 for the \( \psi(2S) \) and table 2 for the \( X(3872) \).
Once the corrected yields $Y^\psi$ and $Y^X$ are determined in each $p_T$ bin, the double differential cross sections (times the product of the relevant branching fractions) can be calculated:

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

(4.5)

where $i$ stands for $\psi(2S)$ or $X(3872)$, $\int \mathcal{L} dt$ is the integrated luminosity, while $\Delta p_T$ and $\Delta y$ are widths of the relevant transverse momentum and rapidity bins, with $\Delta y = 1.5$. $B(i \rightarrow J/\psi \pi^+ \pi^-)$ and $B(J/\psi \rightarrow \mu^+ \mu^-)$ are the branching fractions of these respective decays.

### 5 Lifetime fits

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

$$
F^i(\tau) = (1 - f_{NP}^i)F_P^i(\tau) + f_{NP}^iF_{NP}^i(\tau),
$$

(5.1)

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, and hence $F_P(\tau)$ is effectively described by the lifetime resolution function $F_{\text{res}}(\tau)$, assumed to be the same for $\psi(2S)$ and $X(3872)$ signals. This was verified with simulated data samples. The resolution function $F_{\text{res}}(\tau)$ is parameterised as a weighted sum of three normalised Gaussian functions with a common mean, with respective width parameters $\sigma_1 = \sigma_\tau$, $\sigma_2 = 2\sigma_\tau$ and $\sigma_3 = 4\sigma_\tau$. The resolution parameter $\sigma_\tau$ and the relative weights of the three Gaussian functions are determined separately for each analysis $p_T$ bin, using two-dimensional mass-lifetime unbinned maximum-likelihood fits on the subset of data which contains a narrow range of masses around the $\psi(2S)$ peak. The fitted values for $\sigma_\tau$ are within the range of 32–52 fs, with the weight of the narrowest Gaussian function steadily increasing with $p_T$ from 6% to about 50%.

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 $\tau_{\text{eff}}$ 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 1 and 2, and the results of the corresponding binned minimum-$\chi^2$ fits are shown in figure 3.

Figure 3(a) shows the effective pseudo-proper lifetimes $\tau_{\text{eff}}$ for non-prompt $\psi(2S)$ and $X(3872)$ signals in bins of $p_T$ (see also table 3). While for $\psi(2S)$ the fitted values of $\tau_{\text{eff}}$ 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 at a different production mechanism at low $p_T$.

In figure 3(b) the ratio of non-prompt production cross sections of $X(3872)$ and $\psi(2S)$, times respective branching fractions, for the single-lifetime fit is plotted as a function of transverse momentum. The measured distribution is compared to the kinematic template, which is 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 to both.
Table 3. Effective pseudo-proper lifetimes for non-prompt $\psi(2S)$ and $X(3872)$ obtained with the single-lifetime fit model.

<table>
<thead>
<tr>
<th>$p_T$ bin [GeV]</th>
<th>$\tau_{\text{eff}}(\psi(2S))$ [ps]</th>
<th>$\tau_{\text{eff}}(X(3872))$ [ps]</th>
</tr>
</thead>
<tbody>
<tr>
<td>10–12</td>
<td>$1.44 \pm 0.04$</td>
<td>$1.12 \pm 0.40$</td>
</tr>
<tr>
<td>12–16</td>
<td>$1.43 \pm 0.02$</td>
<td>$1.18 \pm 0.17$</td>
</tr>
<tr>
<td>16–22</td>
<td>$1.43 \pm 0.01$</td>
<td>$1.45 \pm 0.21$</td>
</tr>
<tr>
<td>22–40</td>
<td>$1.41 \pm 0.01$</td>
<td>$1.37 \pm 0.26$</td>
</tr>
<tr>
<td>40–70</td>
<td>$1.44 \pm 0.04$</td>
<td>$1.27 \pm 0.62$</td>
</tr>
</tbody>
</table>

Figure 3. (a) Measured effective pseudo-proper lifetimes for non-prompt $X(3872)$ and $\psi(2S)$. (b) Ratio of non-prompt production cross sections times branching fractions, $X(3872)/\psi(2S)$, in the single-lifetime fit model. The measured distribution is fitted to the kinematic template described in the text.
and a relative weight determined by the fit:

$$F_{\text{NP}}(\tau) = (1 - f_{\text{SL}}^i)F_{\text{LL}}(\tau) + f_{\text{SL}}^iF_{\text{SL}}(\tau).$$  \tag{5.3}

Here, the labels SL and LL refer to short-lived and long-lived non-prompt components, respectively, and $f_{\text{SL}}^i$ are the short-lived non-prompt fractions for $i = \psi(2S), X(3872)$. The PDFs $F_{\text{SL}}(\tau)$ and $F_{\text{LL}}(\tau)$ are parameterised as single one-sided exponential functions with fixed lifetimes, smeared with the lifetime resolution function $F_{\text{res}}(\tau)$ described above. Any long-lived part of the non-prompt contribution is assumed to originate from the usual mix of $B^\pm, B^0, B_s$ mesons and $b$-baryons, while any short-lived part would be due to the contribution of $B_c^\pm$ mesons.

Simulations show that the observed effective pseudo-proper lifetime of $\psi(2S)$ or $X(3872)$ from $B_c$ decays depends on the invariant mass of the hadronic system recoiling from the hidden-charm state. Within the kinematic range of this measurement, it varies from about 0.3 ps for small masses of the recoiling system to about 0.5 ps for the largest ones. The majority of the decays are expected to have masses of the recoiling system between these values, therefore $\tau_{\text{SL}}$ is taken as the mean of the two extremes, 0.40±0.05 ps.

The effective pseudo-proper lifetime of the long-lived component, $\tau_{\text{LL}}$, is determined from the two-lifetime test fits to the $\psi(2S)$ mass range, with $\tau_{\text{LL}}$ free and allowing for an unknown contribution of a short-lived component with lifetime $\tau_{\text{SL}}$. Across the $p_T$ bins, $\tau_{\text{LL}}$ is found to be within the range 1.45 ± 0.05 ps. The effective pseudo-proper lifetimes $\tau_{\text{LL}}$ and $\tau_{\text{SL}}$ are fixed to the above values for the main fits, and are varied within the quoted errors during systematic uncertainty studies.

Figure 4 shows the $p_T$ dependence of the ratio of $X(3872)$ to $\psi(2S)$ cross sections (times respective branching fractions), 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 non-prompt $\psi(2S)$ production is found to be not significant (see table 6 below). The measured ratio of long-lived $X(3872)$ to long-lived $\psi(2S)$, shown in figure 4(b) with blue triangles, is fitted with the MC kinematic template described before to obtain

$$R_{\text{BL}}^{2L} = \frac{\mathcal{B}(B \to X(3872) + \text{any})\mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-)}{\mathcal{B}(B \to \psi(2S) + \text{any})\mathcal{B}(\psi(2S) \to J/\psi\pi^+\pi^-)} = (3.57 \pm 0.33(\text{stat}) \pm 0.11(\text{sys})) \times 10^{-2},$$  \tag{5.4}

with $\chi^2$/dof = 2.3/4, corresponding to the confidence level of 68%. This value of $R_{\text{BL}}$ is somewhat lower than the corresponding result in equation (5.2) obtained from the same data with the single-lifetime fit model. Either is significantly smaller than the value 0.18 ± 0.08 obtained by using the estimate for the numerator, $(1.9 \pm 0.8) \times 10^{-4}$ [11], obtained from the Tevatron data, and the world average values for the branching fractions in the denominator: $\mathcal{B}(B \to \psi(2S)) = (3.07 \pm 0.21) \times 10^{-3}$, $\mathcal{B}(\psi(2S) \to J/\psi\pi^+\pi^-) = (34.46 \pm 0.30)\%$.

Production of $B_c$ mesons in high-energy hadronic collisions at low transverse momentum is expected to be dominated by non-fragmentation processes [27]. These processes are expected to have $p_T$ dependence $\propto p_T^{-2}$ relative to the fragmentation contribution, while it is the fragmentation contribution which dominates the production of long-lived $b$-hadrons [28].
Figure 4. Ratio of cross sections times branching fractions, $X(3872)/\psi(2S)$, for (a) prompt and (b) non-prompt production, in the two-lifetime fit model. 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 data points are slightly shifted horizontally for visibility.

So the ratio of short-lived non-prompt $X(3872)$ to non-prompt $\psi(2S)$, shown in figure 4(b) with red squares, is fitted with a function $a/p_T^2$ to find $a = 2.04 \pm 1.43\text{(stat)} \pm 0.34\text{(sys)}\text{ GeV}^2$, with $\chi^2/\text{dof} = 0.43/4$. This value of $a$, and the measured non-prompt yields of $X(3872)$ and $\psi(2S)$ states, are used to determine the fraction of non-prompt $X(3872)$ from short-lived sources, integrated over the $p_T$ range ($p_T > 10$ GeV) covered in this measurement, giving:

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

where the last uncertainty comes from varying the spin-alignment of $X(3872)$ over the extreme scenarios discussed in appendix A. Since $B_c$ production is only a small fraction of the inclusive beauty production, this value of the ratio could mean that the production of $X(3872)$ in $B_c$ decays is enhanced compared to its production in the decays of other $b$-hadrons.

The two-lifetime fits are used for $\psi(2S)$ and $X(3872)$ to obtain all subsequent results in this paper, 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.

6 Systematic uncertainties

The sources of various uncertainties and their smallest (Min), median (Med) and largest (Max) values across the $p_T$ bins are summarised in table 4 for the differential cross sections of $X(3872)$ and $\psi(2S)$ states, and in table 5 for the measured fractions.
<table>
<thead>
<tr>
<th>Source of uncertainty</th>
<th>$\psi(2S)[%]$</th>
<th>$X(3872)[%]$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min</td>
<td>Med</td>
</tr>
<tr>
<td>Statistical</td>
<td>0.9</td>
<td>1.4</td>
</tr>
<tr>
<td>Trigger eff.</td>
<td>1.0</td>
<td>1.3</td>
</tr>
<tr>
<td>Muon tracking</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Muon reconstruction eff.</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Pion reconstruction eff.</td>
<td>2.5</td>
<td>2.5</td>
</tr>
<tr>
<td>Bkgd suppression req.</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>Mass fit model variation</td>
<td>0.6</td>
<td>0.8</td>
</tr>
<tr>
<td>Short-lifetime variation</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>Long-lifetime variation</td>
<td>0.6</td>
<td>1.0</td>
</tr>
<tr>
<td>Lifetime resolution model</td>
<td>0.4</td>
<td>1.5</td>
</tr>
<tr>
<td>Total systematic</td>
<td>3.5</td>
<td>3.6</td>
</tr>
<tr>
<td>$\frac{1L-\tau}{2L-\tau}$ (prompt)</td>
<td>$-0.1$</td>
<td>$-0.4$</td>
</tr>
<tr>
<td>$\frac{1L-\tau}{2L-\tau}$ (non-prompt)</td>
<td>$+0.1$</td>
<td>$+0.4$</td>
</tr>
</tbody>
</table>

Table 4. Summary of relative 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 (1.9%) is not included.

<table>
<thead>
<tr>
<th>Source of uncertainty</th>
<th>$f_{\psi}^{NP}$ [%]</th>
<th>$f_{X}^{NP}$ [%]</th>
<th>$f_{\psi}^{\tau}$ [%]</th>
<th>$f_{X}^{\tau}$ [%]</th>
<th>$f_{\psi}^{SL}$ [%]</th>
<th>$f_{X}^{SL}$ [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min</td>
<td>Med</td>
<td>Max</td>
<td>Min</td>
<td>Med</td>
<td>Max</td>
</tr>
<tr>
<td>Statistical</td>
<td>0.4</td>
<td>0.5</td>
<td>1.4</td>
<td>4.2</td>
<td>5.8</td>
<td>17.8</td>
</tr>
<tr>
<td>Trigger eff.</td>
<td>0.1</td>
<td>0.1</td>
<td>0.3</td>
<td>0.1</td>
<td>0.1</td>
<td>0.4</td>
</tr>
<tr>
<td>Muon tracking eff.</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Muon reconstruction eff.</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
</tr>
<tr>
<td>Pion reconstruction eff.</td>
<td>0.4</td>
<td>0.5</td>
<td>0.7</td>
<td>0.3</td>
<td>0.3</td>
<td>0.4</td>
</tr>
<tr>
<td>Bkgd suppression req.</td>
<td>0.8</td>
<td>1.1</td>
<td>1.4</td>
<td>0.6</td>
<td>0.7</td>
<td>0.7</td>
</tr>
<tr>
<td>Mass fit model variation</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.6</td>
<td>1.8</td>
</tr>
<tr>
<td>Lifetime resolution variation</td>
<td>0.2</td>
<td>0.7</td>
<td>1.7</td>
<td>0.4</td>
<td>1.0</td>
<td>2.9</td>
</tr>
<tr>
<td>Short-lifetime variation</td>
<td>0.0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.4</td>
<td>0.8</td>
</tr>
<tr>
<td>Long-lifetime variation</td>
<td>0.3</td>
<td>0.4</td>
<td>0.4</td>
<td>0.2</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>Total systematic</td>
<td>1.3</td>
<td>1.5</td>
<td>2.4</td>
<td>1.0</td>
<td>1.4</td>
<td>3.6</td>
</tr>
<tr>
<td>$\frac{1L-\tau}{2L-\tau}$</td>
<td>+0.4</td>
<td>+0.6</td>
<td>+0.9</td>
<td>+0.9</td>
<td>+3.1</td>
<td>+9.1</td>
</tr>
</tbody>
</table>

Table 5. 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 $p_T$ bins. The last row is described in the text.
Uncertainties in the trigger efficiency, and in the muon and pion reconstruction efficiencies are determined using the procedures adopted in ref. [13]. Additional uncertainty of ±2% [14] is assigned to the tracking efficiency of the two muons within the ID, primarily due to its dependence on the total number of \(pp\) collisions per event. The uncertainties in matching generator-level particles to reconstruction-level particles, and in the detector material simulation within the barrel part of the inner detector are found to be the main contributions to the systematic uncertainty of the pion reconstruction efficiency, estimated to be ±2.5%. Such efficiency uncertainties largely cancel in the various non-prompt fractions (table 5).

The uncertainties in the efficiency of the background suppression requirements (see section 4), obtained by combining MC statistical errors and systematic errors in quadrature, are in the range 1%-6%. The uncertainties in the mass fits are estimated by varying the values of parameters that were fixed during the main fit, and by increasing the order of the polynomial \(P\) in the background parameterisation (see equation (4.4)). Similarly, the systematic uncertainties of the lifetime fits are determined by varying the values of the fixed lifetimes and the parameters of the lifetime resolution function within their predetermined ranges.

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 rows in tables 4 and 5 show the relative differences between the values obtained using the single- and two-lifetime fits, labelled as ‘1L-fit’ and ‘2L-fit’, respectively. For the quantities listed in tables 4 and 5, these differences were found to be generally fairly small, compared to the combined systematic uncertainty from other sources.

7 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 5(a). It is described fairly well by the NLO NRQCD model [29] 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 [30], which includes the colour-octet (CO) contributions tuned to 7 TeV CMS data [31] in addition to colour-singlet (CS) production, describes ATLAS data fairly well, with a slight underestimation at higher \(p_T\). The NNLO* Colour-Singlet Model (CSM) predictions [32] 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 5(b), compared with the predictions of the FONLL calculation [28]. The calculation describes the data well over the whole range of transverse momenta.

Similarly, the differential cross section for prompt production of \(X(3872)\) is shown in figure 6(a). It is described within the theoretical uncertainty by the prediction of the NRQCD model which, in this case, considers \(X(3872)\) to be a mixture of \(\chi_{c1}(2P)\) and a
Figure 5. Measured cross section times branching fractions as a function of $p_T$ for (a) prompt $\psi(2S)$ production compared to NLO NRQCD [29], the $k_T$ factorisation model [30] and the NNLO* CSM [32], and (b) non-prompt $\psi(2S)$ production compared to FONLL [28] predictions.

$D^0\bar{D}^{*0}$ molecular state [12], with the production being dominated by the $\chi_{c1}(2P)$ component and the normalisation fixed through the fit to CMS data [10]. The measured differential cross section for non-prompt production of $X(3872)$ is shown in figure 6(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 3(b) and the effective value of the product of the branching fractions $B(B \to X(3872))B(X(3872) \to J/\psi\pi^+\pi^-) = (1.9 \pm 0.8) \times 10^{-4}$ estimated in ref. [11] based on the Tevatron data [33]. This calculation overestimates the data by a factor increasing with $p_T$ from about four to about eight over the $p_T$ range of this measurement.

The non-prompt fractions of $\psi(2S)$ and $X(3872)$ production are shown in figure 7. 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 [14] and CMS [34]. The non-prompt fraction of $X(3872)$ shows no sizeable dependence on $p_T$. This measurement agrees within errors with the CMS result obtained at $\sqrt{s}=7$ TeV [10].

The numerical values of all cross sections and fractions shown in figures 4–7 are presented in table 6.
Figure 6. 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 [12], and (b) non-prompt \( X(3872) \) compared to the FONLL [28] model prediction, recalculated using the branching fraction estimate from ref. [11] as described in the text.

Figure 7. 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 [10, 34].
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross sections times branching fractions [pb / GeV]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\psi(2S))</td>
<td>92.4 ± 1.9</td>
<td>± 4.8</td>
<td>27.97 ± 0.27</td>
<td>± 1.02</td>
<td>5.61 ± 0.06</td>
</tr>
<tr>
<td>(\psi(2S)_{NP})</td>
<td>61.9 ± 1.9</td>
<td>± 3.4</td>
<td>23.66 ± 0.27</td>
<td>± 0.85</td>
<td>6.63 ± 0.06</td>
</tr>
<tr>
<td>(\psi(2S)_{NP}^{LL})</td>
<td>60.8 ± 1.6</td>
<td>± 4.0</td>
<td>23.09 ± 0.27</td>
<td>± 1.46</td>
<td>6.53 ± 0.06</td>
</tr>
<tr>
<td>(\psi(2S)_{NP}^{SL})</td>
<td>1.1 ± 2.4</td>
<td>± 3.9</td>
<td>0.56 ± 0.37</td>
<td>± 1.14</td>
<td>0.11 ± 0.08</td>
</tr>
<tr>
<td>(X(3872))</td>
<td>6.05 ± 1.30</td>
<td>± 0.38</td>
<td>2.75 ± 0.20</td>
<td>± 0.13</td>
<td>0.60 ± 0.04</td>
</tr>
<tr>
<td>(X(3872)_{NP})</td>
<td>2.90 ± 1.20</td>
<td>± 0.21</td>
<td>1.28 ± 0.20</td>
<td>± 0.07</td>
<td>0.29 ± 0.04</td>
</tr>
<tr>
<td>(X(3872)_{NP}^{LL})</td>
<td>1.87 ± 0.82</td>
<td>± 0.14</td>
<td>0.92 ± 0.16</td>
<td>± 0.06</td>
<td>0.29 ± 0.04</td>
</tr>
<tr>
<td>(X(3872)_{NP}^{SL})</td>
<td>1.02 ± 1.49</td>
<td>± 0.20</td>
<td>0.35 ± 0.25</td>
<td>± 0.06</td>
<td>0.01 ± 0.06</td>
</tr>
</tbody>
</table>

### Fractions

<table>
<thead>
<tr>
<th></th>
<th>(F_{NP}^{\psi(2S)})</th>
<th>(F_{SL}^{\psi(2S)})</th>
<th>(F_{NP}^{X(3872)})</th>
<th>(F_{SL}^{X(3872)})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(X(3872))</td>
<td>0.40 ± 0.01</td>
<td>± 0.02</td>
<td>0.46 ± 0.00</td>
<td>± 0.01</td>
</tr>
<tr>
<td>(\psi(2S))</td>
<td>0.02 ± 0.04</td>
<td>± 0.06</td>
<td>0.02 ± 0.02</td>
<td>± 0.05</td>
</tr>
<tr>
<td>(X(3872)_{NP})</td>
<td>0.32 ± 0.12</td>
<td>± 0.02</td>
<td>0.32 ± 0.04</td>
<td>± 0.01</td>
</tr>
<tr>
<td>(\psi(2S)_{NP})</td>
<td>0.35 ± 0.39</td>
<td>± 0.05</td>
<td>0.28 ± 0.16</td>
<td>± 0.04</td>
</tr>
</tbody>
</table>

### Ratios

<table>
<thead>
<tr>
<th></th>
<th>(X(3872))</th>
<th>(X(3872)_{NP})</th>
<th>(X(3872)_{NP}^{LL})</th>
<th>(X(3872)_{NP}^{SL})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\psi(2S))</td>
<td>0.065 ± 0.014</td>
<td>± 0.004</td>
<td>0.098 ± 0.007</td>
<td>± 0.004</td>
</tr>
<tr>
<td>(\psi(2S)_{NP})</td>
<td>0.047 ± 0.019</td>
<td>± 0.004</td>
<td>0.054 ± 0.008</td>
<td>± 0.003</td>
</tr>
<tr>
<td>(\psi(2S)_{NP}^{LL})</td>
<td>0.031 ± 0.014</td>
<td>± 0.002</td>
<td>0.040 ± 0.007</td>
<td>± 0.003</td>
</tr>
<tr>
<td>(\psi(2S)_{NP}^{SL})</td>
<td>0.016 ± 0.024</td>
<td>± 0.003</td>
<td>0.015 ± 0.011</td>
<td>± 0.003</td>
</tr>
</tbody>
</table>

**Table 6.** Summary of \(\psi(2S)\) and \(X(3872)\) cross-section measurements, fractions and ratios. The subscripts P and NP denote prompt and non-prompt components, while the labels 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 (1.9%) and those due to unknown spin-alignment are not included.
8 Dipion invariant mass spectra

The distributions of the dipion invariant mass $m_{\pi\pi}$ in the $\psi(2S) \rightarrow J/\psi\pi^+\pi^-$ and $X(3872) \rightarrow J/\psi\pi^+\pi^-$ decays are measured by determining the corrected yields of $\psi(2S)$ and $X(3872)$ signals in narrow bins of $m_{\pi\pi}$. The two additional selection requirements (equation (3.1)) used specifically to reduce combinatorial background in the cross-section measurement, are found to bias the $m_{\pi\pi}$ distributions and are therefore replaced for this study by requirements on the pseudo-proper lifetime significance, $\tau/\Delta\tau < 2.5$, and the transverse momentum of the $J/\psi\pi^+\pi^-$ candidates, $p_T > 12$ GeV.

The invariant mass distributions of the corrected $J/\psi\pi^+\pi^-$ candidates selected for this analysis are shown in figure 8(a) for the mass range around $\psi(2S)$ peak and in figure 8(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

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

where $m$ is the invariant mass of the $J/\psi\pi^+\pi^-$ system, $Y$ is the yield of the parent resonance, $N_{\text{bkg}}$ is the normalisation factor of the background PDF, $m_0$ is the world average mass [8] of the parent resonance, and $p_{0,1,2,3}$ are free parameters. The signals are described by the same double-Gaussian PDFs $f_1 G_1(m) + (1 - f_1) G_2(m)$ as the ones used in the cross-section analysis described in section 4. 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 8(b). As in the cross-section analysis, the fraction of the narrow Gaussian function $f_1$ is fixed to $0.76 \pm 0.04$, varied within the range of $\pm 0.04$ during systematic uncertainty studies. In another variation a first-order polynomial is added as a
Figure 9. (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 0.280 GeV < \( m_{\pi\pi} \) < 0.595 GeV, fitted with the Voloshin-Zakharov model. Also shown is the normalised \( m_{\pi\pi} \) phase-space distribution (red shaded histogram). (b) Normalised differential decay width of \( X(3872) \to J/\psi(\to \mu^+\mu^-)\rho^0(\to \pi^+\pi^-) \) (blue histogram) and the normalised distribution of \( m_{\pi\pi} \) phase-space (red shaded histogram).

The resulting normalised differential distributions in \( m_{\pi\pi} \) are shown in figure 9(a) for \( \psi(2S) \to J/\psi\pi^+\pi^- \) and in figure 9(b) for \( X(3872) \to J/\psi\pi^+\pi^- \) decays. The solid blue curve in figure 9(a) represents a fit to the data points with the Voloshin-Zakharov distribution

\[
\frac{1}{\Gamma} \frac{d\Gamma}{dm_{\pi\pi}} \propto (m_{\pi\pi}^2 - \lambda m_\pi^2)^2 \times \text{PS},
\]

where PS stands for the dipion phase-space. The fitted value of the parameter \( \lambda \) is found to be \( \lambda = 4.16 \pm 0.06(\text{stat}) \pm 0.03(\text{sys}) \), in agreement with \( \lambda = 4.35 \pm 0.18 \) measured by BES [36], and \( \lambda = 4.46 \pm 0.25 \) measured by LHCb [37]. The shaded blue histogram in figure 9(b) is obtained from straightforward simulations, assuming the dipion system in the decay \( X(3872) \to J/\psi\pi^+\pi^- \) is produced purely via the \( \rho^0 \) 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 9(a) and 9(b) by the red shaded area), with the data clearly preferring higher masses in either case.

9 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 \( \sqrt{s} = 8 \) TeV pp collision data recorded by the ATLAS detector at the LHC. 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 < \( p_T \) < 70 GeV.
The \( \psi(2S) \) cross-section measurements show good consistency with the theoretical predictions based on NLO NRQCD and FONLL for prompt and non-prompt production, respectively. The predictions from the \( k_T \) factorisation model with the colour-octet component tuned to 7 TeV CMS data describe the prompt \( \psi(2S) \) measurement fairly well, while NNLO* colour-singlet model calculations underestimate the data, especially at higher transverse momenta.

The prompt \( X(3872) \) cross-section measurement shows good agreement with the CMS result for transverse momenta \( 10 \text{ GeV} < p_T < 30 \text{ GeV} \) where they overlap, and extends the range of transverse momenta up to 70 GeV. Good agreement is found with theoretical predictions within the model based on NLO NRQCD, 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.

The non-prompt production of \( \psi(2S) \) is described by the FONLL predictions within the uncertainties. 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 compatible results for the prompt and non-prompt differential cross sections of \( \psi(2S) \) and \( X(3872) \).

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

\[
R_{1L}^B = \frac{\mathcal{B}(B \rightarrow X(3872) + \text{any})\mathcal{B}(X(3872) \rightarrow J/\psi\pi^+\pi^-)}{\mathcal{B}(B \rightarrow \psi(2S) + \text{any})\mathcal{B}(\psi(2S) \rightarrow J/\psi\pi^+\pi^-)} = (3.95 \pm 0.32(\text{stat}) \pm 0.08(\text{sys})) \times 10^{-2}.
\] (9.1)

In the two-lifetime model, the two lifetimes are fixed to expected values for \( X(3872) \) originating from the decays of \( B_c \) and from long-lived \( b \)-hadrons, respectively, 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:

\[
R_{2L}^B = \frac{\mathcal{B}(B \rightarrow X(3872) + \text{any})\mathcal{B}(X(3872) \rightarrow J/\psi\pi^+\pi^-)}{\mathcal{B}(B \rightarrow \psi(2S) + \text{any})\mathcal{B}(\psi(2S) \rightarrow J/\psi\pi^+\pi^-)} = (3.57 \pm 0.33(\text{stat}) \pm 0.11(\text{sys})) \times 10^{-2}.
\] (9.2)

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

\[
\frac{\sigma(pp \rightarrow B_c + \text{any})\mathcal{B}(B_c \rightarrow X(3872) + \text{any})}{\sigma(pp \rightarrow \text{non-prompt } X(3872) + \text{any})} = (25 \pm 13(\text{stat}) \pm 2(\text{sys}) \pm 5(\text{spin}))\%.
\] (9.3)

The invariant mass distributions of the dipion system in \( \psi(2S) \rightarrow J/\psi\pi^+\pi^- \) and \( X(3872) \rightarrow J/\psi\pi^+\pi^- \) decays are also measured. The results disfavour a phase-space distribution in both cases, and point strongly to the dominance of the \( X(3872) \rightarrow J/\psi p^0 \) mode in \( X(3872) \) decays.
Acknowledgments

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently.

We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS, CEA-DSM/IRFU, France; GNSF, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; FOM and NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, the Canada Council, CANARIE, CRC, Compute Canada, FQRNT, and the Ontario Innovation Trust, Canada; EPLANET, ERC, FP7, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex and Idex, ANR, Région Auvergne and Fondation Partager le Savoir, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF; BSF, GIF and Minerva, Israel; BRF, Norway; Generalitat de Catalunya, Generalitat Valenciana, Spain; the Royal Society and Leverhulme Trust, United Kingdom.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (U.K.) and BNL (U.S.A.), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in ref. [38].
A Spin-alignment

The acceptance of the $\mu^+\mu^-\pi^+\pi^-$ final state depends on the spin-alignment of the parent state. Several polarisation hypotheses were considered, based on the measured quantum numbers of the hidden-charm states ($J^P = 1^-$ for $\psi(2S)$ and $J/\psi$ [8], $1^+$ for $X(3872)$ [7]) and of the dipion system ($0^+$ in $\psi(2S) \to J/\psi\pi^+\pi^-$ decay [36], $1^-$ in $X(3872) \to J/\psi\pi^+\pi^-$ [7]). In both decays, the dipion system is assumed to be in $S$-wave with respect to the $J/\psi$.

The spin-alignment scenarios considered in this paper were derived using the helicity formalism [39–41], and are conveniently classified in terms of the various helicity amplitudes of the parent state, $A_m$, with $m = -1, 0, +1$:

- **Unpolarised** — an incoherent superposition of $A_+ = 1, A_0 = 1$ and $A_- = 1$, which is labelled UNPOL. This is used as the central hypothesis.
- **Transversely polarised** with either $A_+ = +1, A_0 = 0, A_- = 0$, or $A_+ = 0, A_0 = 0, A_- = +1$, which is labelled $T_{+0}$.
- **Transversely polarised** with $A_+ = +1/\sqrt{2}, A_0 = 0, A_- = +1/\sqrt{2}$, which is labelled $T_{++}$.
- **Transversely polarised** with $A_+ = -1/\sqrt{2}, A_0 = 0, A_- = +1/\sqrt{2}$, which is labelled $T_{+-}$.
- **Longitudinally** polarised with $A_+ = 0, A_0 = +1, A_- = 0$, which is labelled LONG.
- **Off-Plane Positive** — with $A_+ = -\sqrt{6}/3, A_0 = +\sqrt{3}/3, A_- = 0$, which is labelled OFFP+.
- **Off-Plane Negative** — with $A_+ = +\sqrt{6}/3, A_0 = +\sqrt{3}/3, A_- = 0$, which is labelled OFFP−.

Average acceptance weights are calculated for each of these scenarios in each of the analysis $p_T$ bins. The ratios of the average weights for each polarisation scenario to those of the unpolarised case are shown in figure 10(a) for $\psi(2S)$ and figure 10(b) for $X(3872)$, with the values tabulated in tables 7 and 8, respectively.

No individual production process can lead to an unpolarised vector state, but an unpolarised vector state can be observed due to a superposition of several production subprocesses with different spin-alignments [42]. The polarisation of prompt $\psi(2S)$ has been measured by CMS [43] and LHCb [44] and it was found that the angular dependence was close to isotropic, justifying the choice of unpolarised production for the central hypothesis. The non-prompt $\psi(2S)$ and $X(3872)$ are unlikely to show significant spin-alignment, since they are produced from a large number of different incoherent exclusive decays of parent $b$-hadrons.
Figure 10. Correction factors for the (a) \(\psi(2S)\) and (b) \(X(3872)\) yields for various polarisation hypotheses.

Table 7. Correction factors for various polarisation hypotheses in \(p_T\) bins for \(\psi(2S)\) production.

<table>
<thead>
<tr>
<th>Polarisation hypothesis</th>
<th>(p_T) [GeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10–12</td>
</tr>
<tr>
<td>(T_+^0)</td>
<td>1.306</td>
</tr>
<tr>
<td>(T_{++})</td>
<td>1.508</td>
</tr>
<tr>
<td>(T_{+-})</td>
<td>1.156</td>
</tr>
<tr>
<td>LONG</td>
<td>0.682</td>
</tr>
<tr>
<td>OFFP+</td>
<td>1.049</td>
</tr>
<tr>
<td>OFFP−</td>
<td>0.956</td>
</tr>
</tbody>
</table>

Table 8. Correction factors for various polarisation hypotheses in \(p_T\) bins for \(X(3872)\) production.

<table>
<thead>
<tr>
<th>Polarisation hypothesis</th>
<th>(p_T) [GeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10–12</td>
</tr>
<tr>
<td>(T_+^0)</td>
<td>0.921</td>
</tr>
<tr>
<td>(T_{++})</td>
<td>0.900</td>
</tr>
<tr>
<td>(T_{+-})</td>
<td>0.944</td>
</tr>
<tr>
<td>LONG</td>
<td>1.207</td>
</tr>
<tr>
<td>OFFP+</td>
<td>0.969</td>
</tr>
<tr>
<td>OFFP−</td>
<td>1.033</td>
</tr>
</tbody>
</table>
References


[13] 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 09 (2014) 079 [arXiv:1407.5532] [inSPIRE].


[29] ATLAS collaboration, Flavor tagged time-dependent angular analysis of the $B_s \rightarrow J/\psi\phi$ decay and extraction of $\Delta\Gamma_s$ and the weak phase $\phi_s$ in ATLAS, *Phys. Rev. D* 90 (2014) 052007 [arXiv:1407.1796] [INSPIRE].


[34] CMS collaboration, $J/\psi$ and $\psi_{2S}$ production in pp collisions at $\sqrt{s} = 7$ TeV, *JHEP* 02 (2012) 011 [arXiv:1111.1557] [INSPIRE].


The ATLAS collaboration

Centro de Investigaciones, Universidad Antonio Narino, Bogota, Colombia

INFN Sezione di Bologna; Dipartimento di Fisica e Astronomia, Università di Bologna, Bologna, Italy

Physikalisches Institut, University of Bonn, Bonn, Germany

Department of Physics, Boston University, Boston MA, United States of America

Centro de Investigaciones, Universidad Antonio Narino, Bogota, Colombia

INFN Sezione di Bologna; Dipartimento di Fisica e Astronomia, Università di Bologna, Bologna, Italy

Physikalisches Institut, University of Bonn, Bonn, Germany
Also at Department of Physics, The University of Texas at Austin, Austin TX, United States of America

Also at CERN, Geneva, Switzerland

Also at Georgian Technical University (GTU), Tbilisi, Georgia

Also at Ochanomizu University, Tokyo, Japan

Also at Manhattan College, New York NY, United States of America

Also at Academia Sinica Grid Computing, Institute of Physics, Academia Sinica, Taipei, Taiwan

Also at School of Physics, Shandong University, Shandong, China

Also at Department of Physics, California State University, Sacramento CA, United States of America

Also at Moscow Institute of Physics and Technology State University, Dolgoprudny, Russia

Also at section de Physique, Université de Genève, Geneva, Switzerland

Also at Eotvos Lorand University, Budapest, Hungary

Also at Departments of Physics & Astronomy and Chemistry, Stony Brook University, Stony Brook NY, United States of America

Also at International School for Advanced Studies (SISSA), Trieste, Italy

Also at Department of Physics and Astronomy, University of South Carolina, Columbia SC, United States of America

Also at Instituut de Física d’Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Barcelona, Spain

Also at School of Physics and Engineering, Sun Yat-sen University, Guangzhou, China

Also at Institute for Nuclear Research and Nuclear Energy (INRNE) of the Bulgarian Academy of Sciences, Sofia, Bulgaria

Also at Faculty of Physics, M.V.Lomonosov Moscow State University, Moscow, Russia

Also at Institute of Physics, Academia Sinica, Taipei, Taiwan

Also at National Research Nuclear University MEPhI, Moscow, Russia

Also at Department of Physics, Stanford University, Stanford CA, United States of America

Also at Institute for Particle and Nuclear Physics, Wigner Research Centre for Physics, Budapest, Hungary

Also at Flensburg University of Applied Sciences, Flensburg, Germany

Also at University of Malaya, Department of Physics, Kuala Lumpur, Malaysia

Also at CPPM, Aix-Marseille Université and CNRS/IN2P3, Marseille, France

* Deceased