

# Multiplicity-Dependent Modification of $X(3872)$ and $\psi(2S)$ Production in $pp$ Collisions at $\sqrt{s} = 8$ TeV

Matt Durham, Jana Crkovska, Cameron Dean, Eliane Epple, Gerd Kunde, Cesar da Silva<sup>1</sup>.

<sup>1</sup>*Los Alamos National Laboratory, Los Alamos, NM USA 87545*

## Abstract

This analysis note describes the measurement of prompt and non-prompt  $X(3872)$  and  $\psi(2S)$  yields from the 8 TeV  $pp$  data recorded in 2012. The fraction of prompt  $X(3872)$  and  $\psi(2S)$  and the ratio of the cross sections  $\sigma_{X(3872)}/\sigma_{\psi(2S)}$  are measured as a function of produced particle multiplicity, which is represented by the number of tracks reconstructed in the VELO. We observe that the prompt fraction of  $X(3872)$  and  $\psi(2S)$  decreases in events with higher charged particle multiplicity, relative to lower multiplicity collisions. The ratio of cross sections of prompt  $\sigma_{X(3872)}/\sigma_{\psi(2S)}$  decreases, which may be due to final state interactions with co-moving hadrons that inhibit formation of weakly-bound states. Within uncertainties, the measured ratio of non-prompt  $\sigma_{X(3872)}/\sigma_{\psi(2S)}$  has no significant dependence on the charged particle multiplicity, as is expected for hadrons produced in vacuum which do not interact with other produced particles. This is an entirely new metric for studying the production of the exotic  $X(3872)$  state.



# Contents

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Introduction</b>  | <b>1</b>  |
| <b>2</b> | <b>Data Selection</b>  | <b>2</b>  |
| 2.1      | Stripping and Cuts . . . . .   | 3         |
| 2.2      | $J/\psi$ Selection . . . . .   | 4         |
| 2.3      | Event Activity Determination . . . . .                                 | 4         |
| <b>3</b> | <b>Signal Extraction</b>   | <b>9</b>  |
| 3.1      | Background shape studies . . . . .                                     | 9         |
| 3.2      | Prompt/Non-prompt Separation . . . . .                                 | 10        |
| 3.3      | Bin Width Corrections . . . . .  | 16        |
| 3.4      | Primary Vertex Reconstruction . . . . .                                | 19        |
| <b>4</b> | <b>Acceptance and Efficiency Corrections</b>                           | <b>23</b> |
| 4.1      | sWeighted signal . . . . .   | 25        |
| 4.2      | Comments on Decay Kinematics . . . . .                                 | 26        |
| 4.3      | LHCb Acceptance Correction . . . . .                                   | 28        |
| 4.4      | Reconstruction and Selection Efficiencies . . . . .                    | 31        |
| 4.4.1    | Rapidity Distribution Check . . . . .                                  | 36        |
| 4.4.2    | PVZ Distribution . . . . .   | 36        |
| 4.4.3    | Multiplicity dependent effects . . . . .                               | 39        |
| 4.5      | Particle Identification Efficiencies . . . . .                         | 43        |
| <b>5</b> | <b>Results</b>   | <b>47</b> |
| 5.1      | $f_{prompt}$ . . . . .   | 47        |
| 5.2      | Cross Section Ratio . . . . .  | 48        |
|          | <b>References</b>  | <b>51</b> |
| <b>A</b> | <b>Data Tables</b>   | <b>56</b> |
| <b>B</b> | <b>Fits to Mass and Proper Time</b>                                    | <b>58</b> |
| B.1      | $\psi(2S)$ with Fit as Proper Time Background . . . . .                | 58        |
| B.2      | $\psi(2S)$ with Sideband Histogram as Proper Time Background . . . . . | 64        |
| B.3      | X(3872) with Fit as Proper Time Background . . . . .                   | 71        |
| B.4      | X(3872) with Sideband Histogram as Proper Time Background . . . . .    | 74        |

# <sup>1</sup> 1 Introduction

<sup>2</sup> In recent years, multiple new resonances containing heavy quarks have been observed  
<sup>3</sup> which do not follow the patterns expected of typical hadrons (see [1] for a recent review).  
<sup>4</sup> The most well studied of these exotic hadrons is the  $X(3872)$ . It was first discovered in  
<sup>5</sup> the mass spectrum of  $J/\psi\pi^+\pi^-$  coming from  $B$  hadron decays at Belle [2], and has since  
<sup>6</sup> been confirmed by other experiments in multiple decay channels at other colliders [3–6].  
<sup>7</sup> Despite intense experimental and theoretical scrutiny, the exact nature of the  $X(3872)$  is  
<sup>8</sup> still not clear.

<sup>9</sup> Multiple explanations of the  $X(3872)$  structure have been proposed. Shortly after  
<sup>10</sup> its discovery, assignment as one of several possible charmonium states was considered  
<sup>11</sup> [7]. LHCb measurements have since confirmed the quantum numbers to be  $J^{PC} = 1^{++}$   
<sup>12</sup> [8], thereby disfavoring assignment as conventional charmonium because no compatible  
<sup>13</sup> charmonium states near the measured mass are expected to exist [9]. Other models  
<sup>14</sup> consider the  $X(3872)$  to be a tetraquark, which may have further substructure composed  
<sup>15</sup> of a diquark-antidiquark bound state [10–12], or a hadrocharmonium state where two  
<sup>16</sup> light quarks orbit a charmonium core [13]. Mixtures of various exotic and conventional  
<sup>17</sup> states have also been studied [14–16].

<sup>18</sup> The remarkable similarity of the  $X(3872)$  mass and the sum of the masses of the  
<sup>19</sup>  $D^0$  and  $\bar{D}^{*0}$  mesons led to the consideration of the structure as a hadronic molecule, a  
<sup>20</sup> state consisting of these two mesons bound via pion exchange [17, 18]. In this case, the  
<sup>21</sup> binding energy of the  $D^0$  and  $\bar{D}^{*0}$  would be very small, as the mass defect, defined as  
<sup>22</sup>  $M_{X(3872)} - (M_{D^0} + M_{\bar{D}^{*0}}) = 0 \pm 0.27$  MeV, is consistent with zero. Consequently, the  
<sup>23</sup> radius of this state is very large, of order  $\sim 7$  fm [16, 19].

<sup>24</sup> Weakly-bound quarkonia states have been studied extensively in proton-nucleus and  
<sup>25</sup> nucleus-nucleus collisions. Measurements of charmonia production in  $pA$  collisions at fixed  
<sup>26</sup> target experiments [20, 20] and colliders [21–25] have shown that the  $\psi(2S)$  is preferentially  
<sup>27</sup> suppressed more than the  $J/\psi(1S)$  in rapidity regions where a relatively large number of  
<sup>28</sup> charged particles are produced. Similarly, measurements of  $\Upsilon$  production at the LHC have  
<sup>29</sup> shown that the  $\Upsilon(2S)$  and  $\Upsilon(3S)$  states are suppressed more than the more tightly bound  
<sup>30</sup>  $\Upsilon(1S)$  [26, 27]. As the effects governing heavy quark production and transport through the  
<sup>31</sup> nucleus are expected to be similar for states with the same quark content, the mechanism  
<sup>32</sup> for the suppression of excited states is expected to occur in the late stages of the collision,  
<sup>33</sup> after the heavy  $Q\bar{Q}$  pair has projected onto a final state. Models incorporating **final state**  
<sup>34</sup> **effects**, such as breakup via interactions with co-moving hadrons, are able to describe  
<sup>35</sup> the sequential suppression of excited quarkonia states in  $pA$  collisions [28–31]. Given its  
<sup>36</sup> **small binding energy and large radius**, the  $X(3872)$  could be sensitive to similar final-state  
<sup>37</sup> effects that occur in the late stages of collisions, such as breakup via interactions with  
<sup>38</sup> pions produced in the underlying event [32].

<sup>39</sup> High-multiplicity  $pp$  collisions provide a hadronic environment that is similar to nuclear  
<sup>40</sup> collisions. Recently, phenomena typically thought only to occur in collisions of large nuclei  
<sup>41</sup> has been observed in high multiplicity  $pp$  collisions, including strangeness enhancement [33]  
<sup>42</sup> and collective flow [34]. It is possible that the final-state effects observed on quarkonia  
<sup>43</sup> in  $pA$  collisions can be manifest in  $pp$  collisions as well. As such, high-multiplicity  $pp$   
<sup>44</sup> collisions provide a testing ground for examining these effects.

<sup>45</sup> This Analysis Note describes a measurement of  $X(3872)$  and  $\psi(2S)$  production in  $pp$   
<sup>46</sup> collisions at 8 TeV, through their decays to  $J/\psi\pi^+\pi^-$ . We find that, in  $pp$  collisions where

47 a relatively large number of charge particles are produced, the fraction of  $X(3872)$  and  
48  $\psi(2S)$  hadrons that are produced at the collision vertex decreases relative to collisions  
49 with a lower number of produced particles. We also find that the ratio of the cross sections  
50 of prompt  $\sigma_{X(3872)}/\sigma_{\psi(2S)}$  decreases with increasing charged particle multiplicity, showing  
51 that prompt  $X(3872)$  production is suppressed relative to prompt  $\psi(2S)$  production. The  
52 same ratio for  $X(3872)$  and  $\psi(2S)$  produced in  $b$ -decays remains constant as a function of  
53 charged particle multiplicity.

54 The suppression of the promptly produced  $X(3872)$  and  $\psi(2S)$ , as compared to the  
55 non-suppression of non-prompt component, may suggest that interactions at the primary  
56 vertex may be disrupting formation of these bound states. These interactions with co-  
57 moving hadrons produced in the collision can prevent coalescence of the  $c\bar{c}$  pairs [35, 36],  
58 and, in the scenario where the  $X(3872)$  is a hadronic molecule, these interactions can  
59 break up the formed hadron before it can decay or prevent the coalescence of charmed  
60 mesons into a bound state [37, 38]. The effects of these interactions on various hadronic  
61 states is expected to be highly dependent on the state's binding energy [39], so the fact  
62 that the  $X(3872)$  is preferentially suppressed more than the  $\psi(2S)$  in high multiplicity  
63 collisions suggests that its constituent quarks are very weakly bound. Qualitatively, this  
64 supports the interpretation of  $X(3872)$  structure as a weakly bound hadronic molecule or  
65 a charmonium state with very low binding energy, as opposed to a compact tetraquark.  
66 The fact that the non-prompt ratio  $\sigma_{X(3872)}/\sigma_{\psi(2S)}$  remains constant with the number  
67 of produced charged particles is expected, given that these state which are produced  
68 away from the collision point and decay in vacuum will not be affected by interactions  
69 with co-moving hadrons. This also provides an important cross check, showing that the  
70 observed suppression is not merely the result of some loss in tracking efficiency or other  
71 detector effect.

72 The analysis proceeds as follows: Samples of  $X(3872)$  and  $\psi(2S)$  mesons are selected  
73 from the high statistics 8 TeV  $pp$  data recorded in 2012, by fitting peaks found in the  
74  $J/\psi\pi^+\pi^-$  mass spectrum. Prompt and non-prompt components are quantified by a  
75 simultaneous fit of the mass and proper time spectra in multiple bins of charged particle  
76 multiplicity, which is represented by the number of tracks that are reconstructed in the  
77 VELO. The measured yields are corrected for the relative acceptance and efficiency effects  
78 and divided to give the ratio of cross sections  $\sigma_{X(3872)}/\sigma_{\psi(2S)}$  in this decay channel.

## 79 2 Data Selection

80 We begin with an overview of the data samples and steps of the analysis: this analysis  
81 uses data from  $pp$  collisions recorded in 2012 at  $\sqrt{s} = 8$  TeV. The sampled luminosities  
82 are  $914.352 \text{ pb}^{-1}$  in the Magnet Down configuration and  $923.561 \text{ pb}^{-1}$  in the Magnet  
83 Up configuration, with a relative uncertainty of 1.16% [40]. The decays of interest are  
84  $X(3872) \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\pi^+\pi^-$  and  $\psi(2S) \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\pi^+\pi^-$ . Davinci v44r6 is  
85 used to select the candidates. The  $J/\psi\pi^+\pi^-$  mass spectrum is constructed from stripped  
86 data with preselected dimuons near the  $J/\psi$  mass by combining  $J/\psi$  candidates with  
87 two charged pions from the same event. Signal events are found in the  $J/\psi\pi^+\pi^-$  mass  
88 spectrum, which is constructed using pions with opposite charges, while the background  
89 (which is dominated by combinatorial backgrounds from the large number of charged  
90 pions produced in collisions with nuclei) is studied by constructing the like-sign mass

91 spectra  $J/\psi\pi^+\pi^+$  and  $J/\psi\pi^-\pi^-$  mass spectra using pions with the same charge.

92 Once a candidate is identified, the `DecayTreeFitter` TupleTool is used to refit the  
93 decay product tracks with the constraints of a common vertex, and kinematic constraints  
94 that set the dimuon mass to the  $J/\psi$  mass reported in the PDG. Prompt and non-prompt  
95 contributions are separated by a simultaneous fit to the mass and proper time spectra.  
96 In a future analysis, this  $pp$  data will provide a baseline at similar  $\sqrt{s}$  to measure the  
97 modification of cross sections in  $p\text{Pb}$  and  $\text{Pb}p$  collisions.

## 98 2.1 Stripping and Cuts

99 The  $pp$  data is preselected using the `StrippingFullDSTDiMuonJpsi2MuMuTOSLine` line  
100 from stripping 21, which requires muon candidates to have  $p_T > 650 \text{ MeV}/c$ ,  $P > 10$   
101  $\text{GeV}/c$ , and a track  $\chi^2$  per degree of freedom  $\text{TRCHI2DOF} < 5$ . The reconstructed pair  
102 is required to have a pair  $p_T > 3 \text{ GeV}/c$  and the combined vertex fit  $\chi^2$  per degree of  
103 freedom, `VFASPF(VCHI2PDOF)`, is required to be less than 20.

104 The cuts applied to the decay products used to construct the  $J/\psi\pi^+\pi^-$  mass spectrum  
105 are summarized in Tab. 1. These selection criteria largely follow those used in the  
106 previously published LHCb analysis of inclusive  $X(3872)$  production in 7 TeV  $pp$  collisions  
107 [41]. In addition to the cuts required by stripping, the reconstructed muon tracks are  
108 required to have a  $\chi^2$  per degree of freedom that is less than 3, and also must have a ghost  
109 track probability `TRACK_GhostProb`  $< 0.3$  and satisfy the `IsMuon` particle identification  
110 cut. Pairs made from these muons are required to have a distance of closest approach of  
111 less than 0.5 mm. Pion tracks are required to have a  $\chi^2$  per degree of freedom that is  
112 less than 3, transverse momentum  $> 500 \text{ MeV}/c$ , and a total momentum greater than 3  
113  $\text{GeV}/c$ , where the RICH detectors are at full efficiency for identifying pions. Pion pairs are  
114 required to have a distance of closest approach to each other that is less than 0.5 mm, and  
115 also must have a distance of closest approach within 0.5 mm of the muon pair. Dimuon  
116 pairs are required to have a mass within  $\pm 3\sigma$  of the mean of the  $J/\psi$  peak. Finally, the  
117 reconstructed  $J/\psi\pi^+\pi^-$  candidate is required to have a transverse momentum above 5  
118  $\text{GeV}/c$ .

119 The main source of background in this measurement is due to combinations of charged  
120 pions which do not originate from the same decay. To limit this, a cut on the reaction  
121  $Q$  value is applied which requires  $m_{J/\psi\pi^+\pi^-} - m_{J/\psi} - m_{\pi^+\pi^-}$  to be less than  $300 \text{ MeV}/c^2$ ,  
122 following what was used in the previous analysis at  $\sqrt{s} = 7 \text{ TeV}$ . Previous measurements  
123 from CDF [42], ATLAS [43], CMS [44], as well as LHCb [45, 46] have studied the  
124 mass spectra  $m_{\pi^+\pi^-}$  for pion pairs produced in the decays of interest here. These  
125 papers (and others) established that the pion pair produced in the  $X(3872)$  decay comes  
126 from the decay of an intermediate  $\rho^0(770)$  meson, *i.e.* the complete decay chain is  
127  $X(3872) \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\rho^0(770)(\rightarrow \pi^+\pi^-)$ . Fig. 1 shows measurements of the dipion  
128 mass from LHCb [45], CDF [42], and ATLAS [43], where we see that the mass spectrum  
129 is consistent with expectations from the  $\rho$  decay. Similar measurements of the pion pair  
130 from the decay  $\psi(2S) \rightarrow J/\psi\pi^+\pi^-$  (see Fig. 2) have shown that these pairs fall within the  
131 mass range  $300 < m_{\pi^+\pi^-} < 600$ . The cut on reaction  $Q$  effectively requires  $m_{\pi^+\pi^-} > 476$   
132  $\text{MeV}/c^2$  for  $X(3872)$  decays, and  $m_{\pi^+\pi^-} > 290 \text{ MeV}/c^2$  for  $\psi(2S)$  decays, which retains  
133 the signal while reducing combinatorial backgrounds. The  $J/\psi\pi^+\pi^-$  mass spectrum that  
134 is prepared using these cuts is shown in Fig. 3.

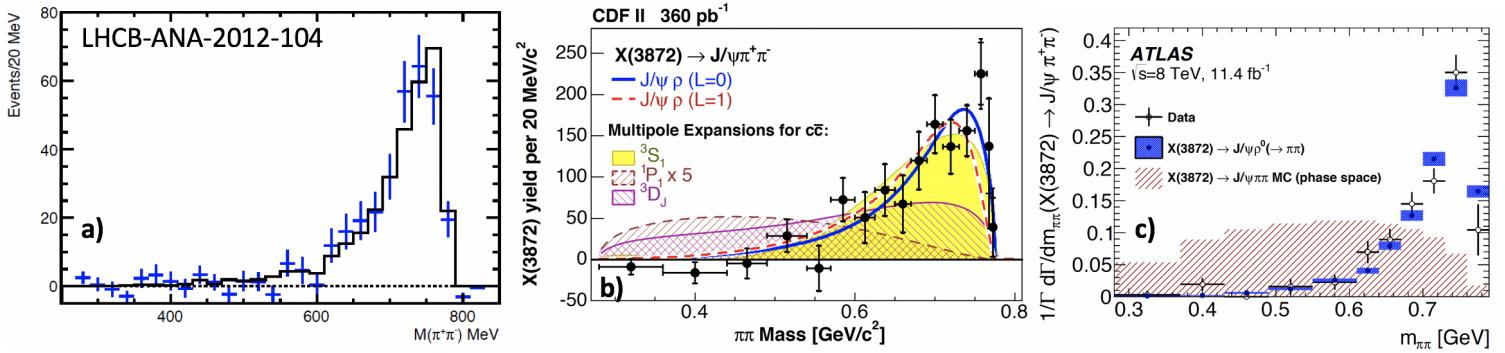


Figure 1: Measurement of the invariant mass spectrum of pion pairs produced in the decay  $X(3872) \rightarrow J/\psi\pi^+\pi^-$ , from a) LHCb [45], b) CDF [42], and c) ATLAS [43]. These measurements motivate our cut on the pion pair mass  $600 < m_{\pi^+\pi^-} < 780$  MeV/ $c^2$  for the  $X(3872)$  analysis.

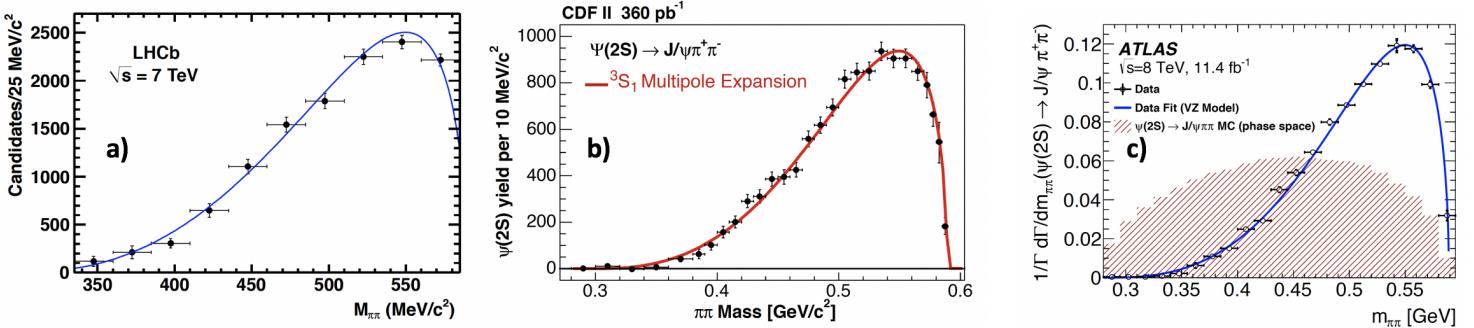


Figure 2: Measurement of the invariant mass spectrum of pion pairs produced in the decay  $\psi(2S) \rightarrow J/\psi\pi^+\pi^-$ , from a) LHCb [46], b) CDF [42], and c) ATLAS [43]. These measurements motivate our cut on the pion pair mass  $300 < m_{\pi^+\pi^-} < 600$  MeV/ $c^2$  for the  $\psi(2S)$  analysis.

## 135 2.2 $J/\psi$ Selection

136 The  $J/\psi$  candidates used to form the  $J/\psi\pi^+\pi^-$  mass spectra are selected within a  $\pm 3\sigma$   
 137 window around the mean. This selection window is determined by a fitting the  $\mu^+\mu^-$   
 138 mass spectra that is prepared using the same cuts as the  $X(3872)$  and  $\psi(2S)$  candidates,  
 139 excluding the cuts on the  $\mu^+\mu^-$  mass, with a Gaussian to represent the peak and a third  
 140 order Chebyshev polynomial to represent the background. The  $J/\psi$  peaks from the 2012  
 141  $pp$  data are shown in Fig. 4, for the Mag Down (left) and Mag Up (right) running periods.  
 142 The sample contains  $\sim 1.3 \times 10^6$   $J/\psi$  counts, and there is good agreement between the  
 143 widths measured in these two run periods.

## 144 2.3 Event Activity Determination

145 This analysis relies on tracks reconstructed in the VELO to serve as the measure of event  
 146 activity. It is therefore crucial to select events that occur in a  $z$  vertex range where the  
 147 acceptance of the VELO is stable to avoid biases in the track reconstruction that can  
 148 induce non-physical detector effects. First we examine the distributions for the number of  
 149 tracks reconstructed in the VELO, for different number of primary vertices nPV, see Fig.

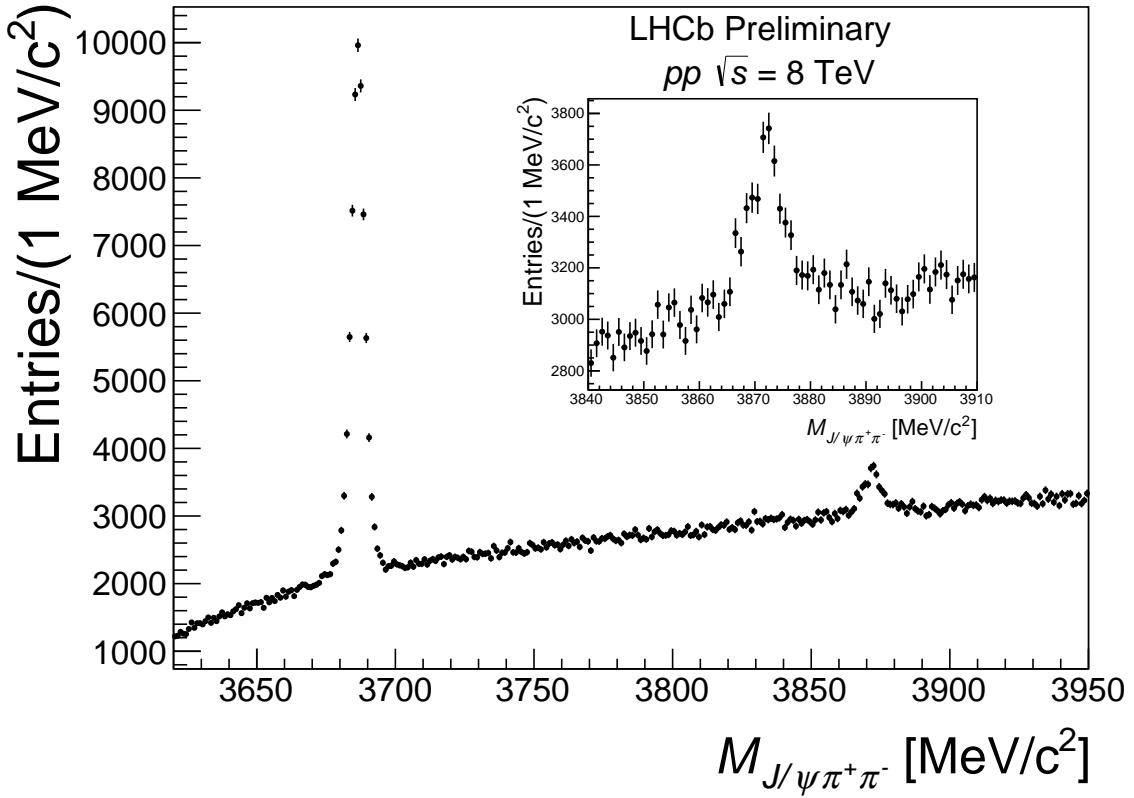


Figure 3: The invariant mass spectrum of  $J/\psi\pi^+\pi^-$  prepared using the selection criteria as described in this Note.

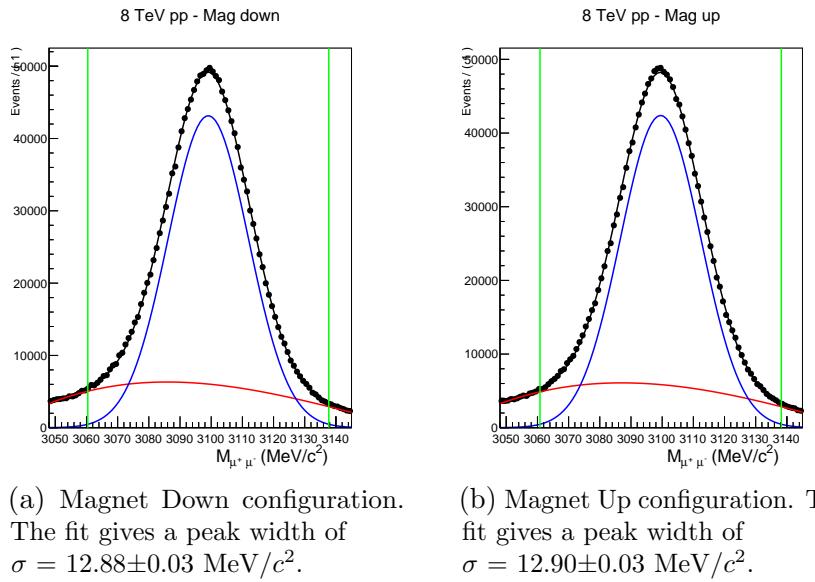


Figure 4: Fits to the dimuon mass spectra in the 2012  $pp$  data sample, for the Magnet Down (left) and Magnet Up (right) configurations. The  $\pm 3\sigma$  cut around the mean, denoted by the green vertical lines, is used to select candidates when constructing the  $J/\psi\pi^+\pi^-$  mass distribution.

| DecayTree Particle   | Variable   | Accepted Value  |
|----------------------|--|---|
| Individual $\mu^\pm$ | $p_T$<br>$p$<br>TRCHI2DOF<br>TRACK_GhostProb<br>IsMuon             | >650 MeV<br>>3 GeV<br>< 3<br>< 0.3<br>TRUE  |
| $\mu^+\mu^-$ pair    | $ M_{\mu^+\mu^-} $<br>$p_T$<br>ACUTDOCA                            | < $3\sigma$ around $J/\psi$ peak<br>> 3 $\text{GeV}/c$<br>0.5 mm                        |
| Individual $\pi^\pm$ | $p_T$<br>$p$<br>TRCHI2DOF<br>TRACK_GhostProb<br>IsMuon<br>ProbNNpi | >500 $\text{MeV}/c$<br>>3 $\text{GeV}/c$<br>< 3<br>< 0.3<br>FALSE<br>> 0.9              |
| $\pi^+\pi^-$ pair    | Cut on reaction $Q$<br>ACUTDOCA                                    | $m_{J/\psi\pi^+\pi^-} - m_{J/\psi}$<br>$m_{\pi^+\pi^-} < 300 \text{ MeV}/c^2$<br>0.5 mm |
| $J/\psi\pi^+\pi^-$   | ACUTDOCA between $\mu^+\mu^-$ and $\pi^+\pi^-$<br>$p_T$            | 0.5 mm<br>>5 $\text{GeV}/c$   |
| Global cuts          | nPV==1<br>-60mm<PVZ<120mm  |   |

Table 1: Candidate selection cuts.

5, for events containing a  $J/\psi\pi^+\pi^-$  candidate selected by the cuts previously discussed. As expected, the number of tracks increases with the number of primary vertices. For a single vertex, there is a maximum of 250 tracks per event. We compare this with the VELO tracking efficiency from the JINST paper in Fig. 6 [47], which shows that the VELO has a constant efficiency in this range of track multiplicity. This analysis only considers candidates produced in events with a single primary vertex, in order to avoid counting tracks which originate from a different collision than the candidate. The distribution of the number of primary vertices for all candidates is shown in Fig. 7. With the nPV=1 cut we retain  $\sim 34\%$  of the candidates, while nPV=2 has 38%, nPV = 3 has 20%, nPV=4 has 6%, and the fraction continues to decrease with nPV.

Figure 8 shows the distribution of the number of reconstructed tracks selected from events with one primary vertex, for different ranges of the primary vertex  $z$  location PVZ, selected from data which contain  $J/\psi\pi^+\pi^-$  candidates which pass our selection criteria. We see a common trend for most of the range of PVZ selections, indicating that in this PVZ range, the VELO acceptance is equivalent. However, at the edges of the PVZ

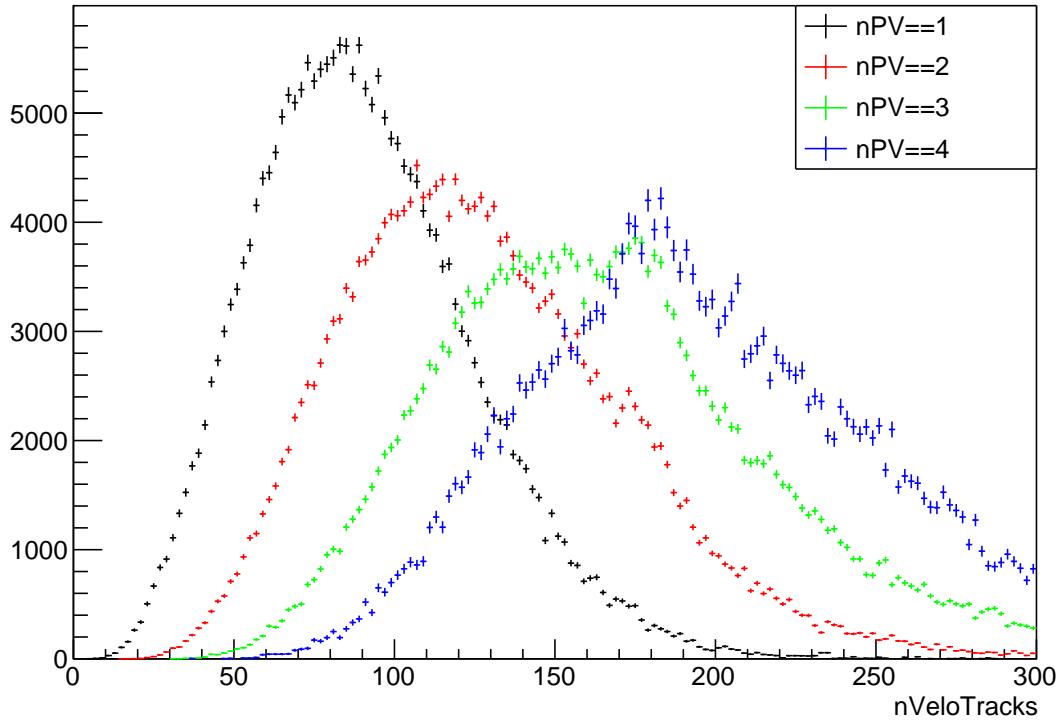


Figure 5: Distribution of the number of tracks reconstructed in the VELO, for different numbers of reconstructed primary vertices, from the 2012 pp data considered here. In this analysis we chose  $nPV == 1$ .

range ( $PVZ < -60$  mm, as indicated by the black and red points), there is a clear deviation from the other curves towards lower track multiplicity. This is due to events producing tracks that do not enter the VELO acceptance. Therefore, in this analysis, we restrict our primary vertex  $z$  range to  $-60 \text{ mm} < PVZ < 120 \text{ mm}$ .

$pp_{lsf}it$

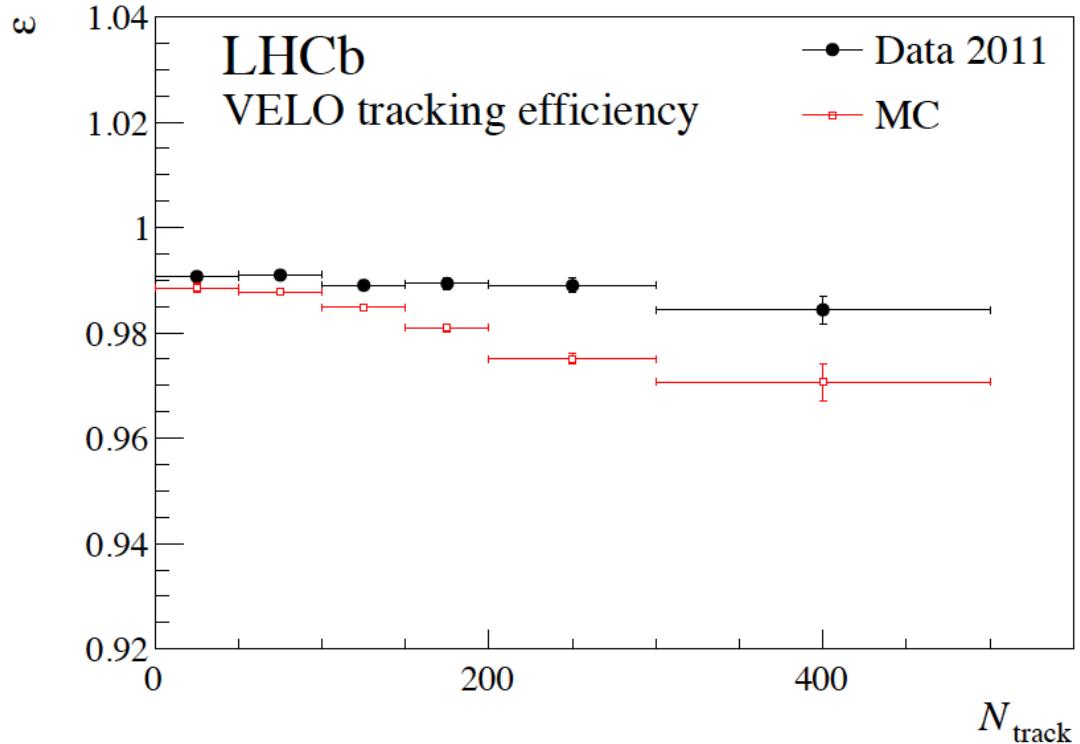


Figure 6: VELO tracking efficiency [47].

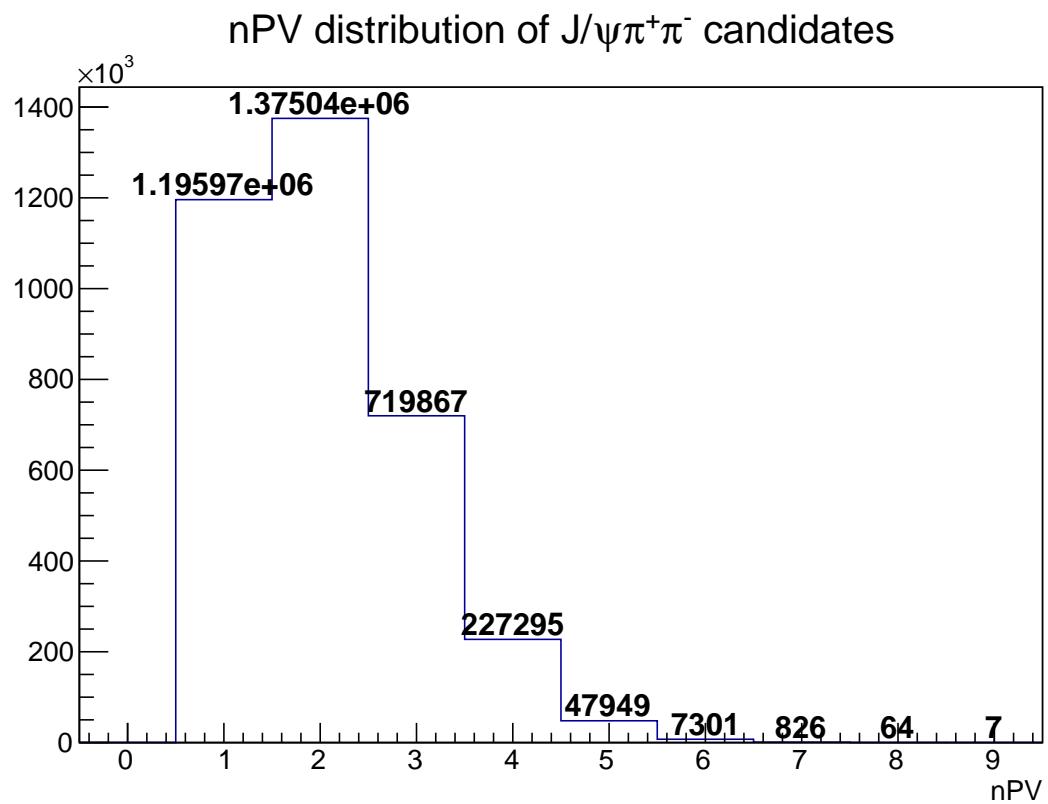


Figure 7: Distribution of the number of primary vertices. In this analysis we chose  $nPV==1$ .

nPV==1

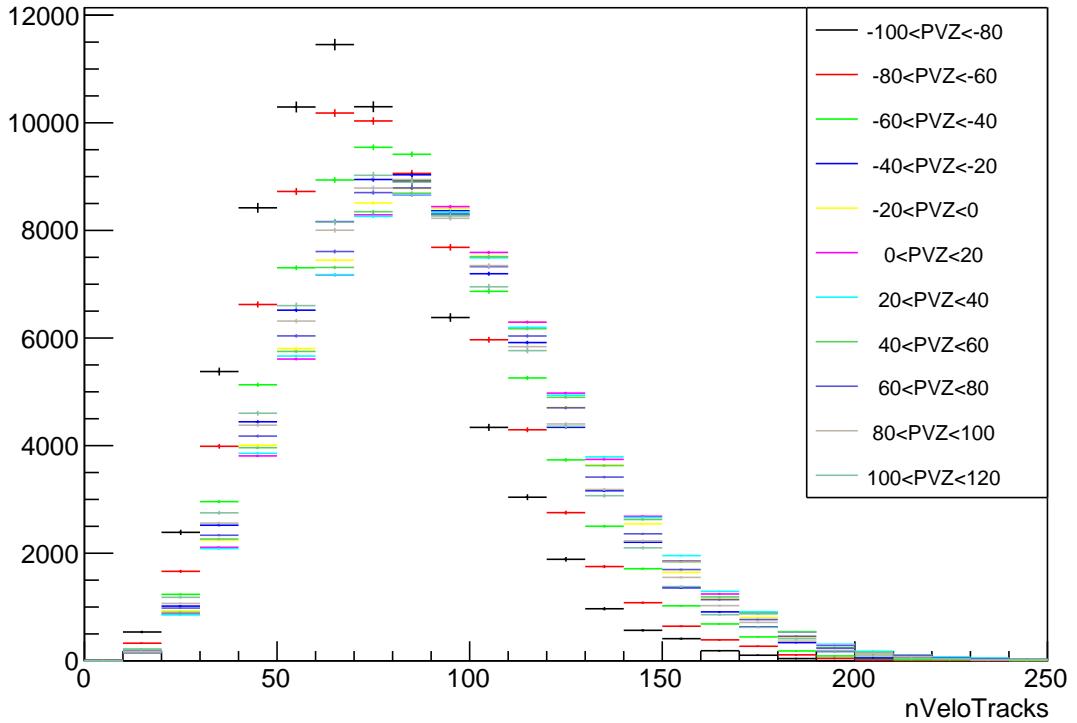


Figure 8: Distribution of the number of tracks reconstructed in the VELO, for different  $z$  ranges of the primary vertex, from data containing a decay candidate. To maintain constance VELO acceptance in that analysis we chose  $-60 \text{ mm} < z < 120 \text{ mm}$ .

### 169 3 Signal Extraction

170 The  $X(3872)$  and  $\psi(2S)$  signals are extracted from the data by simultaneously fitting the  
 171  $J/\psi\pi^+\pi^-$  mass and proper time spectra obtained as described in the preceding section of  
 172 this Note. Following previous LHCb analyses of  $\psi(2S)$  production in  $pp$  collisions through  
 173 the same decay channel [46], the  $\psi(2S)$  peak is fit with a sum of two single-sided Crystal  
 174 Ball functions that are constrained to have the same central value. The natural width of  
 175 the  $X(3872)$  is not known, but the PDG value is constrained to be less than 1.2 MeV at  
 176 the 90% CL [48]. Therefore, we expect the width to be dominated by the mass resolution  
 177 of the LHCb spectrometer, which should be described by a Gaussian, as was done in the  
 178 previous LHCb analysis at 7 TeV [41]. The background shape is studied by examining  
 179 the mass spectra produced by combing  $J/\psi$  candidates with like-sign pion pairs. Various  
 180 fitting procedures are varied to determine a systematic uncertainty on the yields extracted  
 181 from the fits.

#### 182 3.1 Background shape studies

183 The background under the  $X(3872)$  and  $\psi(2S)$  peaks is expected to be dominated by  
 184 uncorrelated combinatorial background from charged pions that do not originate from  
 185 the same parent particle decay. The shape of this background is studied by constructing  
 186 the like-sign mass spectra from  $J/\psi\pi^+\pi^+$  and  $J/\psi\pi^-\pi^-$ . All the same selection criteria

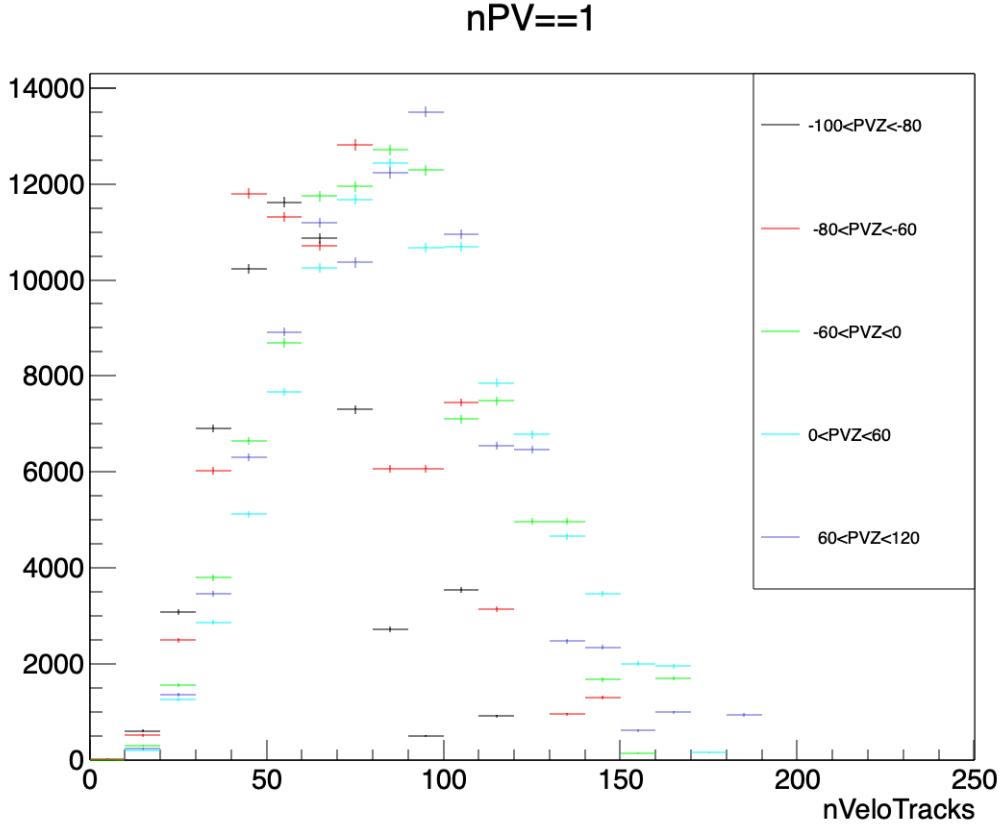


Figure 9: The same distribution as Fig.8, but from  $\psi(2S)$  simulation data.

187 applied to the candidates found in the unlike-sign  $J/\psi\pi^+\pi^-$  mass spectra are applied to  
 188 create the like-sign mass spectra. The like-sign mass distribution (summed over both  
 189 the MagUp and MagDown running periods and the  $\pi^+\pi^+$  and  $\pi^-\pi^-$  dipion selections) is  
 190 shown in Fig. 10, and is found to be well described by a third order Chebychev polynomial,  
 191 which is of the form  $ax^3 - bx$ . We therefore take this shape to represent the background  
 192 in the fit to the unlike-sign mass spectra.

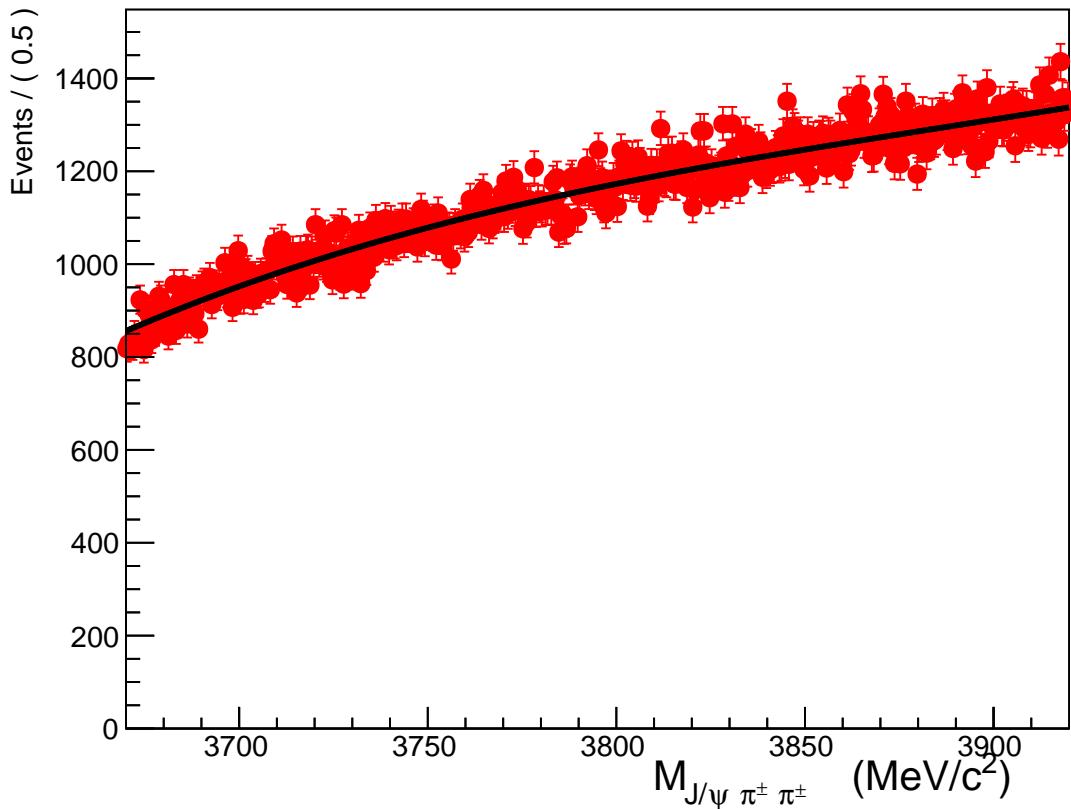
### 193 3.2 Prompt/Non-prompt Separation

194 Separation of the prompt and non-prompt components of the signals is accomplished by  
 195 performing a simultaneous fit of the  $J/\psi\pi^+\pi^-$  mass spectrum and the pseudo proper time  
 196 spectra, as was performed in multiple previous LHCb analyses (c.f. [49, 50]). The pseudo  
 197 proper time  $t_z$  is defined as

$$198 \quad t_z = \frac{(z_{decay} - z_{PV})M}{p_z}, \quad (1)$$

199 where  $z_{decay} - z_{PV}$  is the difference between the  $z$  positions of the reconstructed  
 200 vertex of the  $J/\psi\pi^+\pi^-$  and the collision's primary vertex,  $M$  is the nominal mass of  
 201 the reconstructed hadron (3871.69 MeV/ $c^2$  for the X(3872) and 3686.097 MeV/ $c^2$  for  
 202 the  $\psi(2S)$ , as reported in the PDG [48]), and  $p_z$  is the candidate's momentum along  
 203 the  $z$ -axis. When doing a simultaneous fit, the mass spectrum uniquely constrains the signal  
 to background ratio, while the  $t_z$  spectrum uniquely constrains the prompt and

8 TeV pp



### Pull Distribution

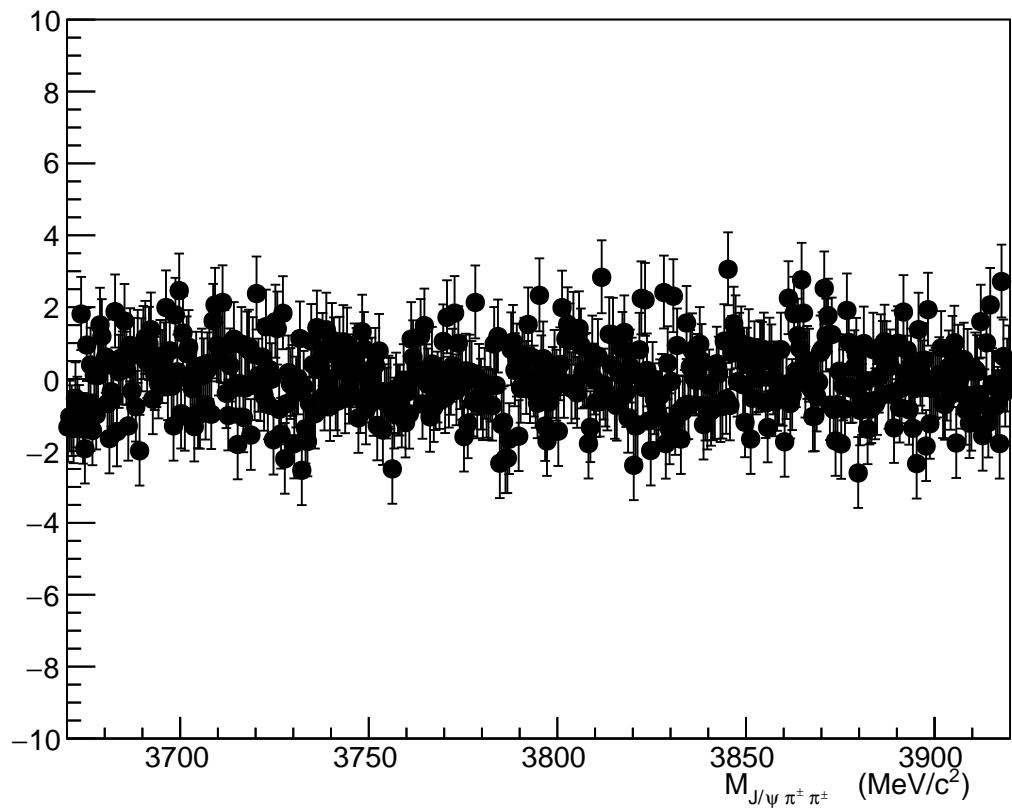


Figure 10: The sum of the  $J/\psi \pi^+ \pi^+$  and  $J/\psi \pi^- \pi^-$  mass spectra constructed from  $pp$  data, fit with a 3rd order Chebychev polynomial.

204 non-prompt yields, allowing both the total number of signal candidates and the fraction  
205 which are promptly produced  $f_{prompt}$  to be determined. The data sample is divided into  
206 bins of  $N_{tracks}^{VELO}$  to examine the modifications in  $f_{prompt}$  as a function of event multiplicity.  
207 As previously discussed, only events with exactly one reconstructed primary vertex are  
208 considered in this part of the analysis, in order to avoid biases in  $N_{tracks}^{VELO}$  caused by tracks  
209 from more than one primary vertex.

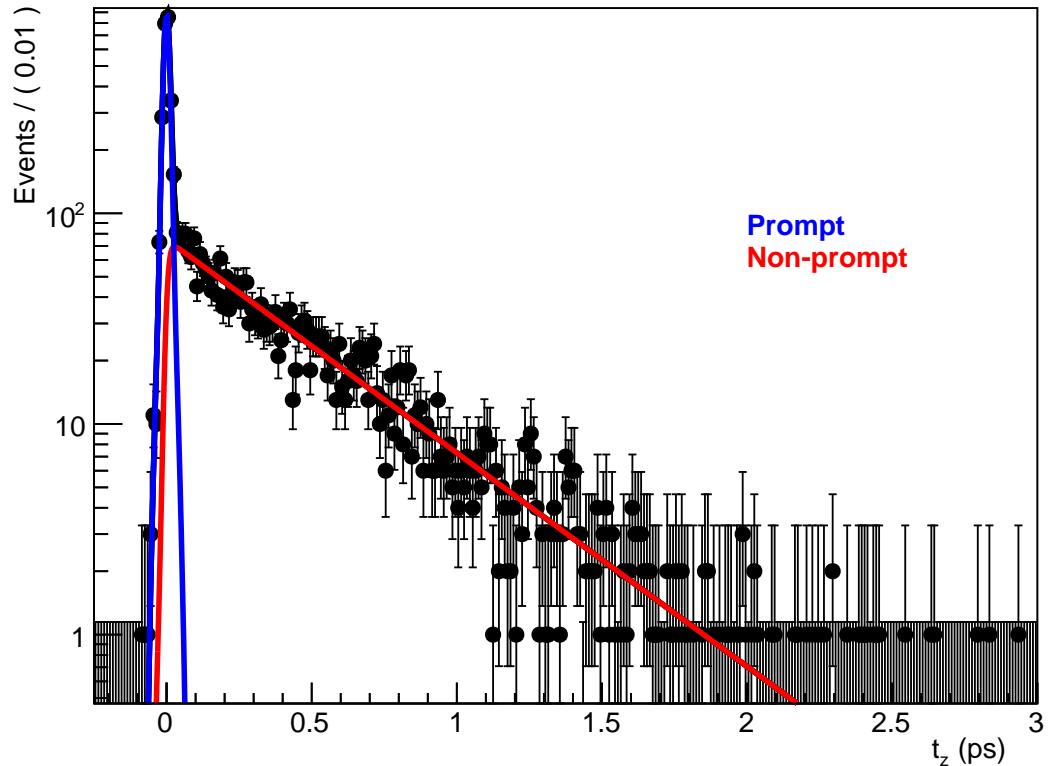
210 The fit to the mass spectra proceeds using peak and background shapes that are  
211 discussed in the previous section of this Note. The function fit to the proper time  
212 spectrum is the sum of a prompt signal, a non-prompt signal, and prompt and non-prompt  
213 backgrounds. The prompt signal is represented by a Dirac  $\delta$  function, convolved with the  
214 sum of three Gaussian resolution functions to represent the  $t_z$  resolution of the LHCb  
215 spectrometer, as was performed in previous analyses. The non-prompt signal is represented  
216 by an exponential decay function that is also convolved with the resolution function. This  
217 is studied in simulation by selecting truth-matched  $\psi(2S)$  to create a background-free  
218 sample of candidates satisfying our selection criteria. The sample is fit with the sum of  
219 prompt and non-prompt functions that are convolved with the resolution functions as  
220 previously described and shown in Fig. 11. We see that this choice of functional forms  
221 provides a very good description of the simulated data.

222 Two methods are used to determine the shape of the background  $t_z$  distribution. The  
223 first is similar to what is used in previous analyses: a sum of empirical functions are  
224 used to represent the backgrounds in the fit to the  $t_z$  spectrum. Following the procedure  
225 in [49], the prompt background is represented by a Dirac delta function convolved with a  
226 triple Gaussian resolution function, which is the same shape as the prompt signal function.  
227 The non-prompt background is represented in the fit with one decay function that is  
228 symmetric about zero, plus two decay functions at positive times and two decay functions  
229 at negative times. The time constants of these decays are determined by fitting the  
230  $t_z$  spectrum prepared from the mass sidebands without any selection on  $N_{tracks}^{VELO}$ , with  
231 the signal components set to zero. The decay time constants are fixed to the values  
232 determined from this fit when performing the fit in the individual bins of event activity.  
233 The normalization of the background components are allowed to float individually, as  
234 the relative contributions of the various backgrounds is expected to change between the  
235 different bins of  $N_{tracks}^{VELO}$ . An example of such a fit is shown in Fig. 12, which is the  $pp$   
236 data in the multiplicity class  $60 < N_{tracks} < 70$ . All of the fits in all bins for both the  
237  $\psi(2S)$  and  $X(3872)$  can be found in the Appendix of this Note.

238 A different method that is more data-driven is also studied, where we use histograms  
239 of the mass sideband region directly in the fit to represent the background. First, the  $t_z$   
240 spectra in the upper and lower sideband regions (defined as  $3700 < m < 3730 \text{ MeV}/c^2$   
241 and  $3640 < m < 3700 \text{ MeV}/c^2$  for the  $\psi(2S)$ , and  $3820 < m < 3850 \text{ MeV}/c^2$  and  
242  $3890 < m < 3920 \text{ MeV}/c^2$  for the  $X(3872)$ ) are prepared. The shape taken from the  
243 sum of these two histograms is used to represent the proper time background when  
244 the simultaneous fit is performed to the mass and  $t_z$  spectra. The delta function and  
245 exponential decay functions, convolved with the resolution function, are used represent  
246 the prompt and non-prompt decay contributions, as is done in the previously described  
247 method. An example of this fit is shown in Fig. 13, which uses the same data as shown in  
248 Fig. 12.

249 The measured fraction of prompt  $\psi(2S)$   $f_{prompt}^{measured}$  extracted using both of these methods  
250 is shown as a function of event activity for the  $pp$  data for the  $\psi(2S)$  in Fig. 14. We

### Fit to simulation



### Pull Distribution

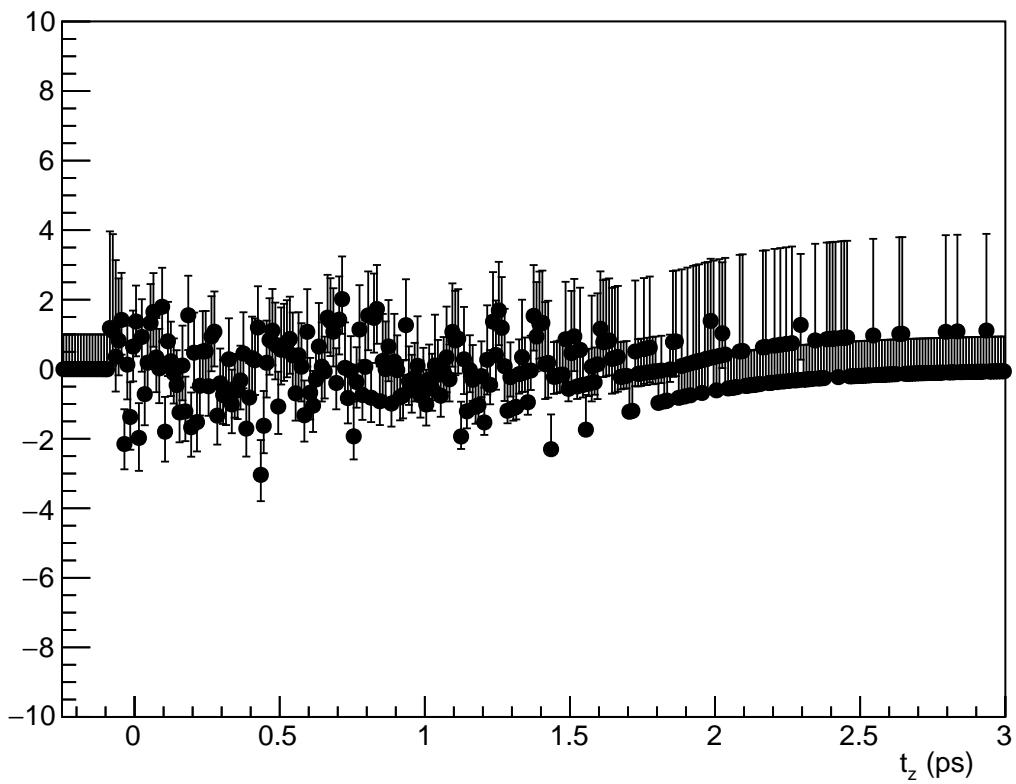


Figure 11: Fit to simulated  $\psi(2S)$  candidates, showing that our choice of functions for prompt and non-prompt components provides an accurate description.

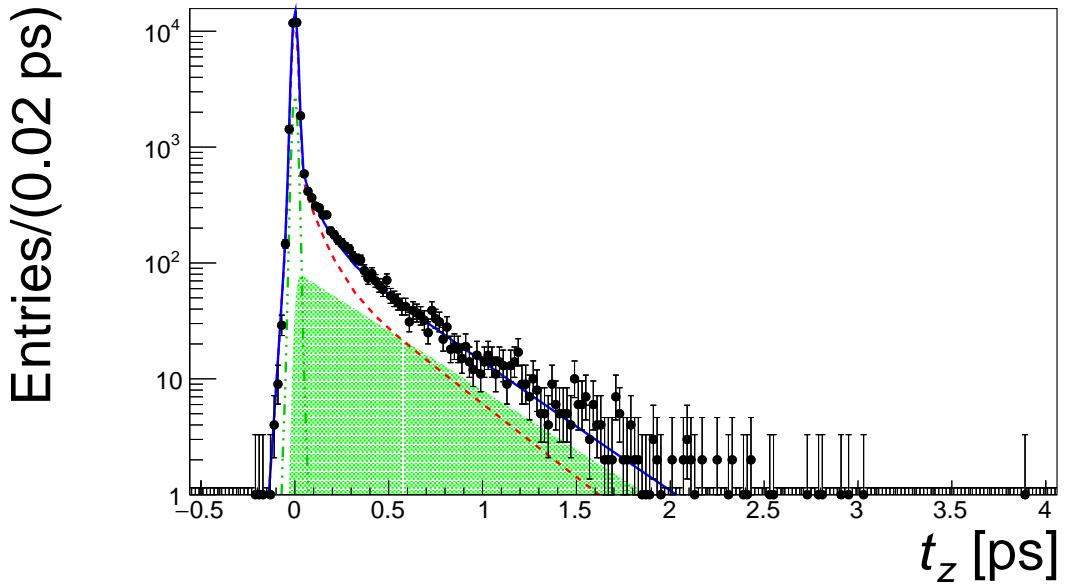
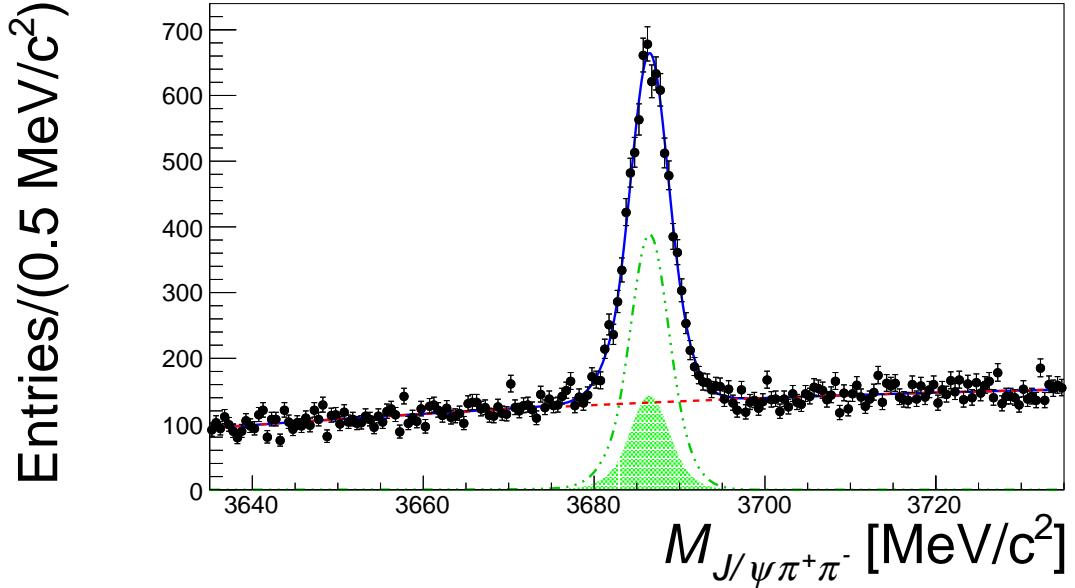


Figure 12: Simultaneous fit to the mass and  $t_z$  spectra, for the  $\psi(2S)$  mass region in the  $pp$  data in the  $60 < N_{\text{tracks}}^{\text{VELO}} < 70$  event activity class. This simultaneous fit uses empirical functions to model the  $t_z$  backgrounds. In these fits, the data is the black points, the prompt and non-prompt signal components are shown in green, the background is red, and the total fit is shown in blue.

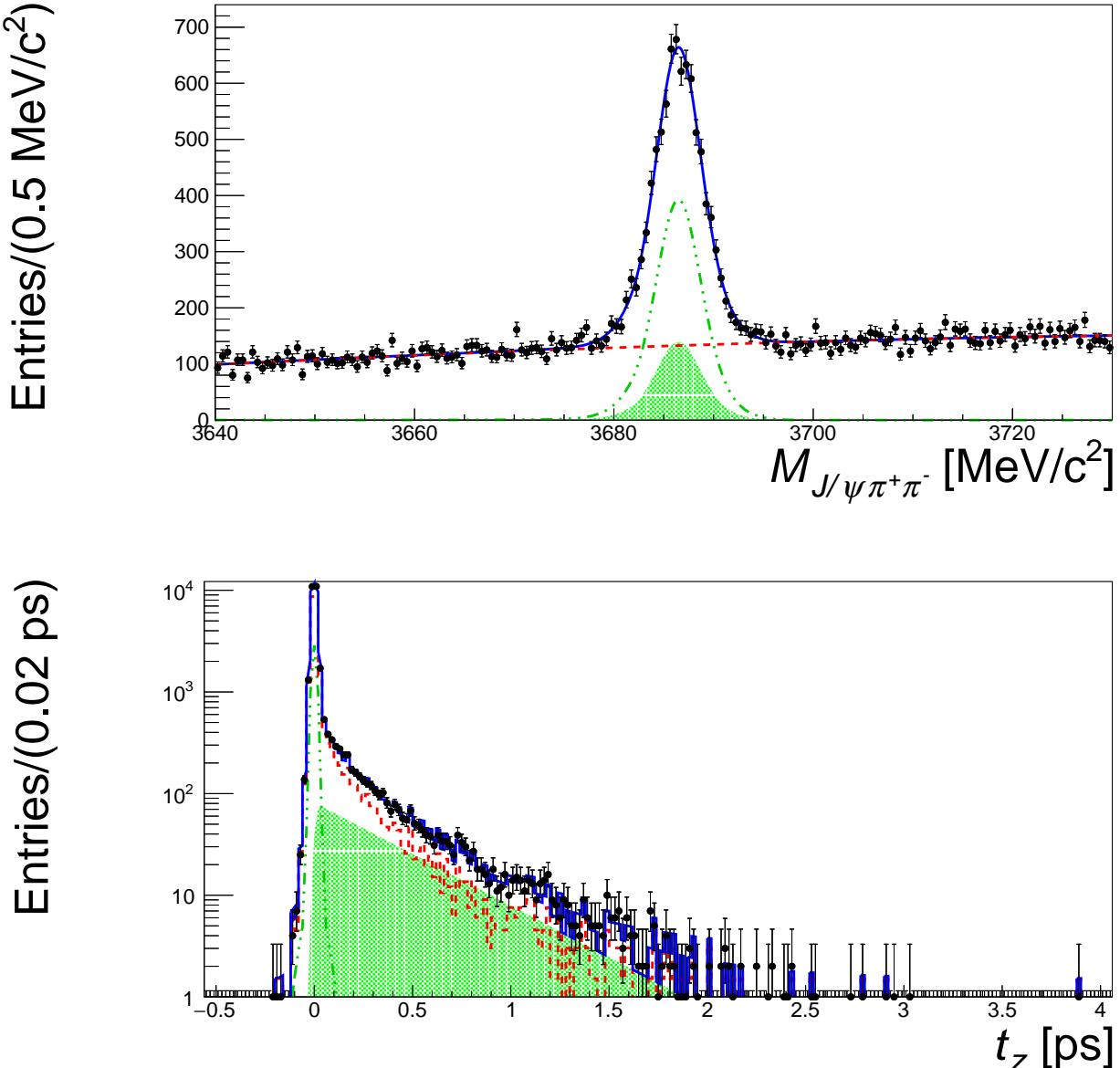


Figure 13: Simultaneous fit to the mass and  $t_z$  spectra, for the  $\psi(2S)$  mass region in the  $pp$  data in the  $60 < N_{\text{tracks}}^{\text{VELO}} < 70$  event activity class. This simultaneous fit uses the shape of histograms of the  $t_z$  spectra from the mass sidebands to represent the  $t_z$  backgrounds.

251 see reasonable agreement between the two methods in all bins, with typical differences  
 252 of 1 to 2 % percent, and the difference is taken as a systematic uncertainty. The same  
 253 procedure is followed for the  $X(3872)$  data, and is shown in Fig. 15. As expected, due to  
 254 the smaller signal/background ratio in the  $X(3872)$  region, we see a larger effect when  
 255 choosing different parameterizations of the  $t_z$  background. The central values of  $f_{\text{prompt}}$   
 256 are determined by averaging the data points from the two methods, and the difference  
 257 between this central value and the value obtained from the individual methods is taken as  
 258 a systematic uncertainty.

259 As an additional crosscheck, we examine the prompt fraction of pure background

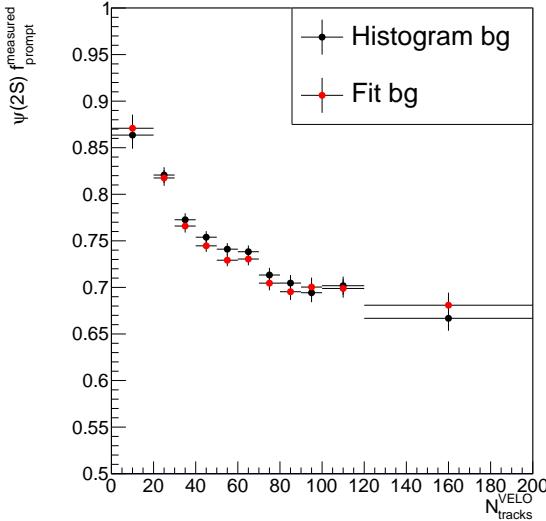


Figure 14: The measured fraction of promptly produced  $\psi(2S)$ .

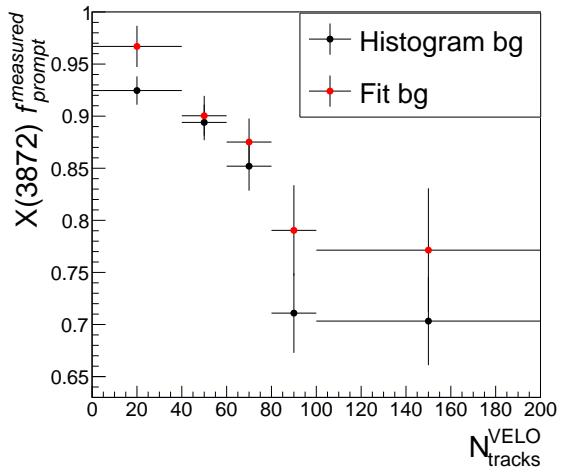


Figure 15: The measured fraction of promptly produced  $X(3872)$ .

260 events. Using the previously discussed histograms selected from the mass sidebands, we  
 261 define a prompt region as  $t_z < 0.1\text{ps}$  and a non-prompt region as  $t_z > 0.1 \text{ ps}$ . We can  
 262 then study the prompt fraction of background events as a function of Nvelo tracks by  
 263 calculating

$$f_{prompt}^{BG} = \frac{N_{prompt}^{BG}}{N_{prompt}^{BG} + N_{non-prompt}^{BG}} = \frac{\int_{t_{min}}^{0.1} H_{BG} dt}{\int_{t_{min}}^{t_{max}} H_{BG} dt} \quad (2)$$

264 where  $H_{BG}$  is our  $t_z$  histogram constructed from the mass sidebands. This quantity is  
 265 shown in Fig. 16.

266 We see a slight increase in the prompt fraction in the sideband regions as the  $N_{tracks}^{VELO}$   
 267 increases, but this behavior is drastically different from what is observed in the signal  
 268 region. Therefore we can conclude that our observations in the signal region are not due  
 269 to influences of increasing backgrounds.

### 270 3.3 Bin Width Corrections

271 The fits performed in the preceding section, and thereby the physics results of interest here,  
 272 are binned in ranges of  $N_{tracks}^{VELO}$ . Over the range covered by individual bins, particularly the  
 273 large bins required for the  $X(3872)$  signals, the  $N_{tracks}^{VELO}$  distribution can vary significantly  
 274 (see Fig. 5). Therefore, the center of the bin may not represent the ideal horizontal  
 275 coordinate for plotting the results.

276 To study these effects, the method of sWeights is used to extract the distribution of  
 277  $N_{tracks}^{VELO}$  for prompt and non-prompt  $\psi(2S)$  and  $X(3872)$  signals (see the discussion in Sec.  
 278 4.1 for details on the sWeight procedure). These distributions are shown in Fig. 17, along  
 279 with fits to a negative binomial distribution. We can determine the weighted average  
 280 value  $\langle N_{tracks}^{VELO} \rangle$  in each bin by solving

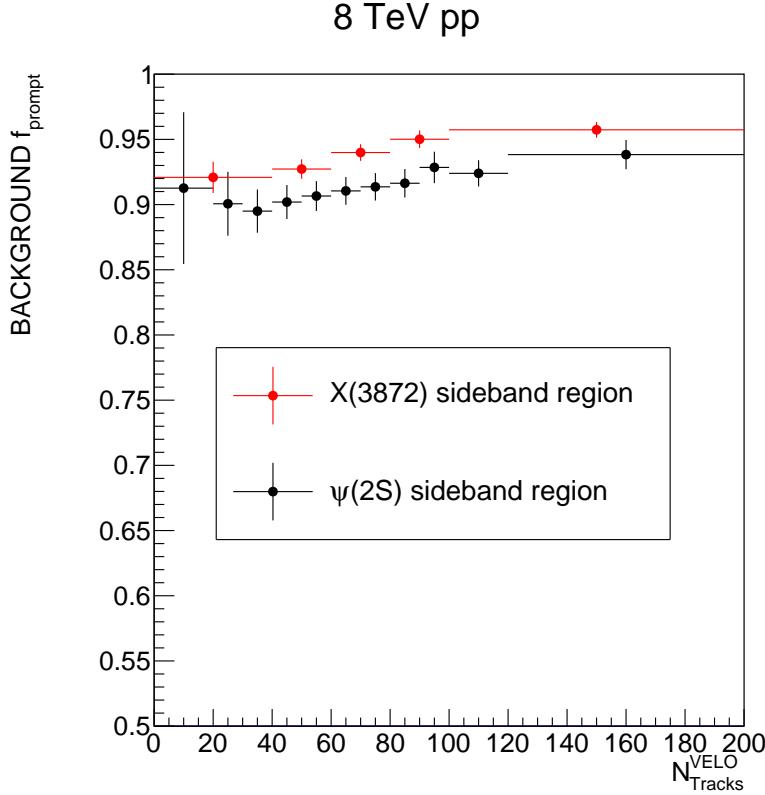


Figure 16: The fraction of background that is prompt for the  $X(3872)$  and  $\psi(2S)$  sideband regions.

$$\langle N_{\text{tracks}}^{\text{VELO}} \rangle = \int_{N_{\text{min}}}^{N_{\text{max}}} N f(N) dN / \int_{N_{\text{min}}}^{N_{\text{max}}} f(N) dN \quad (3)$$

where  $N_{\text{min}}$  and  $N_{\text{max}}$  are the low edge and high edge of the bin and  $f(N)$  is the negative binomial function fit to the sWeighted data.

While the  $\psi(2S)$  data is well constrained, the  $X(3872)$  data shows significant statistical uncertainties. To determine an uncertainty on the extracted values of  $\langle N_{\text{tracks}}^{\text{VELO}} \rangle$ , 10000 fitting trials are done in Monte Carlo. In each trial, the values of  $N_{\text{tracks}}^{\text{VELO}}$  in each bin of the plots shown in Fig. 17 are varied by randomly sampling from a Gaussian that is constructed by setting the mean to the value of  $N_{\text{tracks}}^{\text{VELO}}$  in that bin, and a width that is set to the error bar. The resulting plots are fit and the value of  $\langle N_{\text{tracks}}^{\text{VELO}} \rangle$  is calculated from each trial. The resulting values are collected in histograms and fit with a gaussian, and the mean and width of that gaussian are taken as the central value and uncertainty on  $\langle N_{\text{tracks}}^{\text{VELO}} \rangle$ , respectively. This is done for each of the four data sets considered: prompt and non-prompt  $\psi(2S)$  and  $X(3872)$ . An example of the distributions from the trials is shown in Fig. 18, and the results are collected in Tab.2.

We observe that there are some differences between the values for prompt and non-prompt, and between the two species. This is an effect of the different modifications that the various species undergo as a function of multiplicity (which is the point of this analysis). To choose a point at which to plot the data, we make the average value of  $\langle N_{\text{tracks}}^{\text{VELO}} \rangle$  across the four species in each bin, weighted by the inverse of the square of

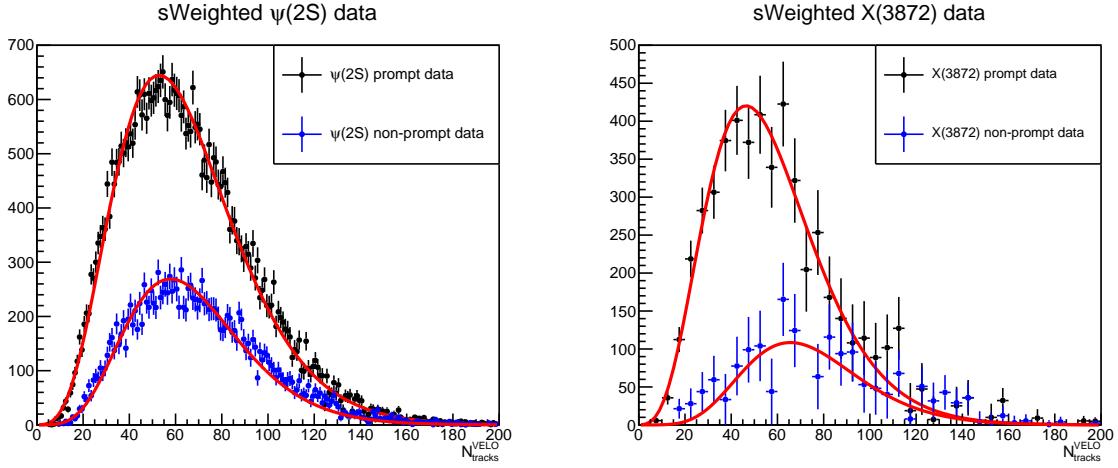


Figure 17: The distributions of  $N_{\text{tracks}}^{\text{VELO}}$  for prompt and non-prompt  $\psi(2S)$  and  $X(3872)$  signals, fit with negative binomial distributions.

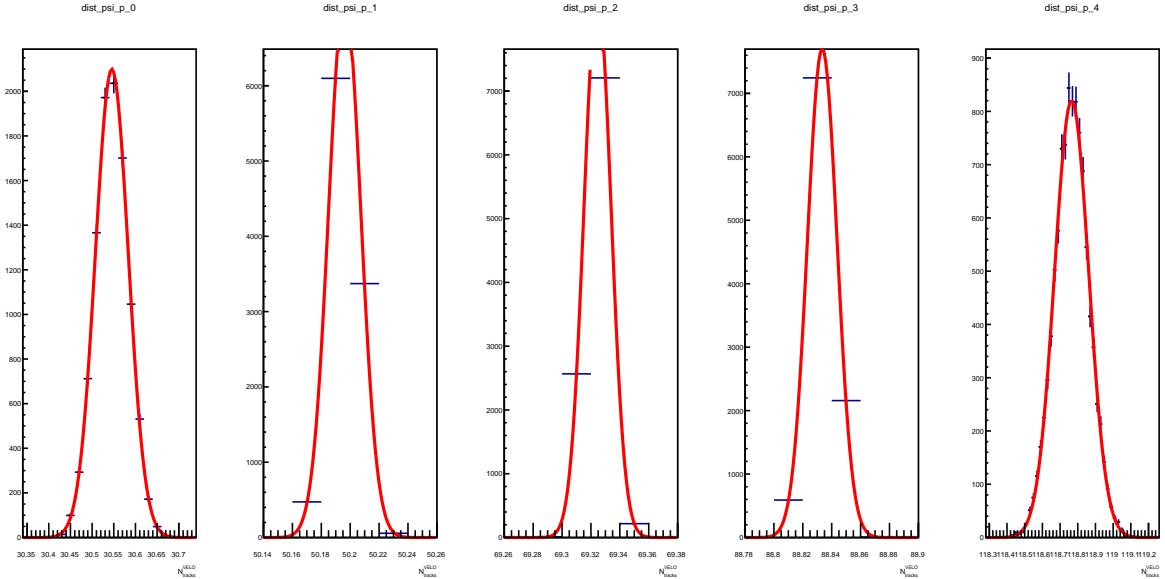


Figure 18: The distribution of  $N_{\text{tracks}}^{\text{VELO}}$  for prompt  $\psi(2S)$  signals, obtained by varying the data points from Fig. 17 within uncertainties in 10000 toy experiments.

299 the uncertainty (following the PDG weighted average procedure). This gives the final  
300 points that are used when plotting, collected in Tab. 3.

301 This procedure is repeated for the relatively fine bins that are used when plotting the  
302 prompt fraction of  $\psi(2S)$ , following the binning scheme as in Fig. 14. The results are  
303 collected in Tab. 4. We see similar results, where the only appreciable differences from  
304 the center of the bin occur in the first and last bins, which are relatively wide.

| Bin Size | Center | Prompt $\psi(2S)$ | Non-prompt $\psi(2S)$ | Prompt X(3872)  | Non-prompt X(3872) |
|----------|--------|-------------------|-----------------------|-----------------|--------------------|
| 0-40     | 20     | $30.55 \pm 0.04$  | $31.74 \pm 0.07$      | $29.7 \pm 0.2$  | $34 \pm 1.9$       |
| 40-60    | 50     | $50.19 \pm 0.01$  | $50.54 \pm 0.02$      | $49.8 \pm 0.1$  | $51.4 \pm 0.8$     |
| 60-80    | 70     | $69.32 \pm 0.01$  | $69.47 \pm 0.02$      | $69.0 \pm 0.1$  | $69.5 \pm 0.3$     |
| 80-100   | 90     | $88.83 \pm 0.01$  | $88.87 \pm 0.01$      | $88.5 \pm 0.1$  | $88.4 \pm 0.2$     |
| 100-200  | 150    | $118.7 \pm 0.1$   | $118.0 \pm 0.1$       | $116.4 \pm 0.8$ | $112 \pm 2$        |

Table 2: The weighted average  $\langle N_{\text{tracks}}^{\text{VELO}} \rangle$  for prompt and non-prompt  $\psi(2S)$  and X(3872) signals in each bin.

| Bin Size | Center | Weighted average $\langle N_{\text{tracks}}^{\text{VELO}} \rangle$ |
|----------|--------|--|
| 0-40     | 20     | 30.8   |
| 40-60    | 50     | 50.3   |
| 60-80    | 70     | 69.3   |
| 80-100   | 90     | 88.8   |
| 100-200  | 150    | 118.3  |

Table 3: The weighted average  $\langle N_{\text{tracks}}^{\text{VELO}} \rangle$ , used as the horizontal axis values when plotting the results.

### 3.4 Primary Vertex Reconstruction

In order to characterize each event, a selection on the number of reconstructed primary vertices  $n_{\text{PV}}=1$  is made, to avoid track contributions from multiple PVs that occur in the same beam crossing (pile up). However, the average LHCb primary vertex reconstruction efficiency has been determined to be  $\sim 94\%$  for minimum bias events [51], so in some events with multiple PVs, only a single PV will be reconstructed, contaminating the sample of actual single PV events. This effect is studied in Monte Carlo by preparing a set of simulated PVs, assigning each a number of tracks and  $z$  vertex, and applying efficiency curves from [51], as shown in Fig. 22, to these simulated events. Thereby the amount of contamination in each bin of nVeloTracks from neighboring bins with an unreconstructed PV can be quantified.

The simulation samples consist of  $10^8$  events. To construct these simulated events, first a number of PVs is chosen by randomly sampling from a Poisson distribution with a mean of 1.7, which is the pile-up from the 2012 data set considered in this analysis. Each PV in an event is assigned a  $z$  coordinate by randomly sampling from a fit the PVZ distribution obtained from the data as shown in Fig. 19, which is well described by a Gaussian of mean 22 mm and width 53 mm, to obtain the generated PV distributions shown in Fig. 20. Similarly, each vertex is also assigned a track multiplicity. One vertex (representing the signal event) is assigned a multiplicity by randomly sampling from the binomial fit to the prompt  $\psi(2S)$  data shown in Fig. 17, which is the most well constrained distribution extracted from the data (although there is little difference between the various distributions). The other vertices in the event are assigned a multiplicity by randomly sampling from a similar negative binomial function with a most probable value of 35, which is what is observed in Min Bias collisions. These distributions are shown in Fig. 21. Following the definitions in [51], the generated primary vertices are characterized as “isolated” if their assigned  $z$  position is greater than 10 mm from any other PV in the

| Bin Size | Center | Prompt $\psi(2S)$  | Non-prompt $\psi(2S)$ | Weighted average $\langle N_{\text{tracks}}^{\text{VELO}} \rangle$ |
|----------|--------|--------------------|-----------------------|--|
| 0-20     | 10     | $15.98 \pm 0.02$   | $16.48 \pm 0.03$      | 16.1   |
| 20-30    | 25     | $25.748 \pm 0.007$ | $25.98 \pm 0.02$      | 25.8   |
| 30-40    | 35     | $35.359 \pm 0.004$ | $35.52 \pm 0.01$      | 35.4   |
| 40-50    | 45     | $45.128 \pm 0.003$ | $45.236 \pm 0.009$    | 45.1   |
| 50-60    | 55     | $54.975 \pm 0.002$ | $55.048 \pm 0.008$    | 55.0   |
| 60-70    | 65     | $64.868 \pm 0.002$ | $64.914 \pm 0.009$    | 64.9   |
| 70-80    | 75     | $74.788 \pm 0.002$ | $74.815 \pm 0.009$    | 74.8   |
| 80-90    | 85     | $84.726 \pm 0.002$ | $84.74 \pm 0.01$      | 84.7   |
| 90-100   | 95     | $94.678 \pm 0.002$ | $94.68 \pm 0.01$      | 94.7   |
| 100-120  | 110    | $108.52 \pm 0.01$  | $108.48 \pm 0.02$     | 108.5  |
| 120-200  | 160    | $136.97 \pm 0.08$  | $136.1 \pm 0.1$       | 136.7  |

Table 4: The values of  $\langle N_{\text{tracks}}^{\text{VELO}} \rangle$  for prompt and non-prompt  $\psi(2S)$  signals, and the weighted average used when plotting the results.

331 event, and characterized as “close” if they are less than 10 mm from any other PV. The  
 332 reconstruction efficiencies for close and isolated PVs from [51] shown in Fig. 22 is applied  
 333 to each PV, depending on its isolation status.

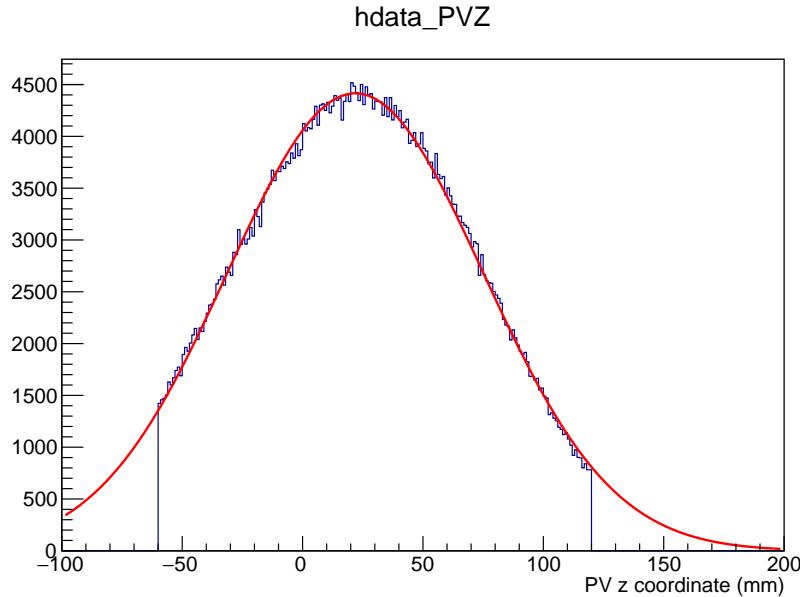


Figure 19: The primary vertex  $z$  distribution from the data considered in this analysis. Random sampling from this distribution is used to set the  $z$  vertex (PVZ) values of the generated PVs.

334 From these simulated samples, events are selected which have exactly one reconstructed  
 335 primary vertex. These selected events are binned according to the track multiplicity  
 336 assigned to this reconstructed vertex, and the total track multiplicity of the event, defined  
 337 as the sum of tracks from all PVs in the event (reconstructed plus non-reconstructed), is  
 338 plotted in Fig. 23 for each bin. We see that the majority of events are correctly identified,  
 339 however there are some events with at least one unreconstructed primary vertex that  
 340 would be incorrectly identified as a higher multiplicity single PV event and placed in the

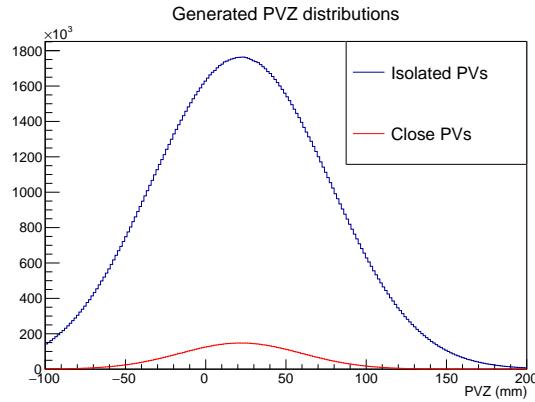


Figure 20: Primary vertex  $z$  (PVZ) values of the close and isolated generated PVs.

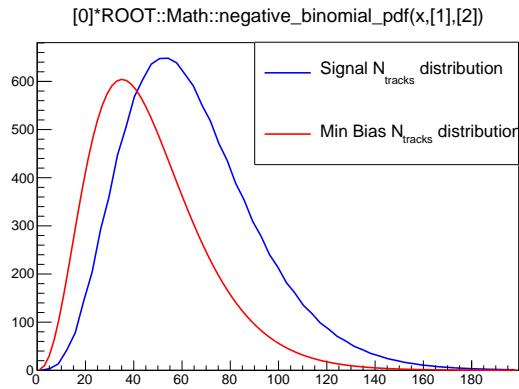


Figure 21: Distributions used to assign multiplicities to the generated PVs (see text for details).

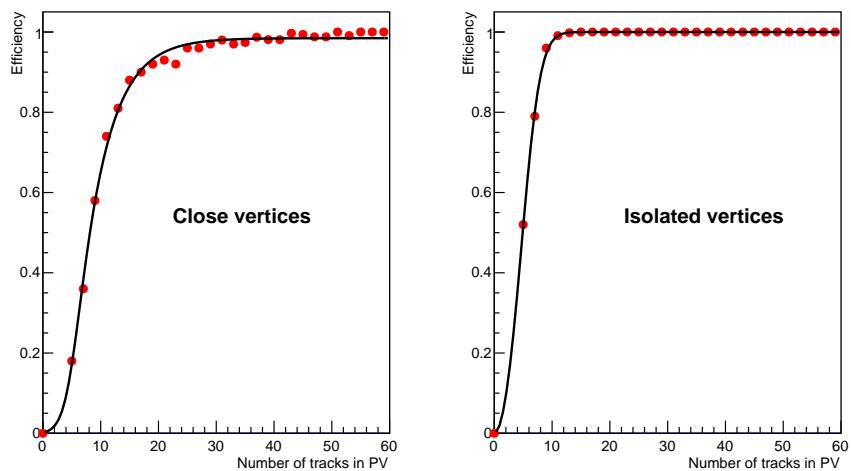


Figure 22: Efficiencies for reconstructing “close” and “isolated” primary vertices, from [51].

<sup>341</sup> wrong bin. Most of this bin migration moves data into the neighboring higher bin, due to  
<sup>342</sup> the relatively low efficiency for reconstructing vertices with < 10 tracks. The fraction of  
<sup>343</sup> truth events in each measured bin is given in Tab. 5.

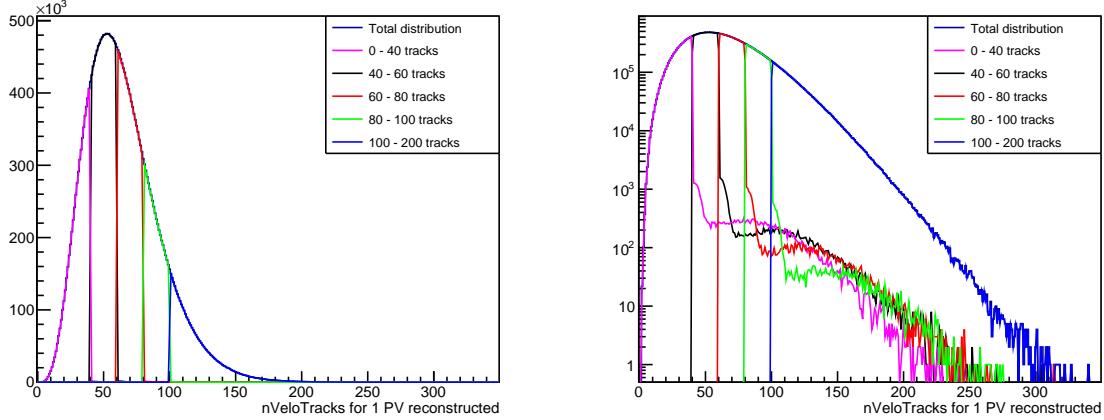


Figure 23: Distribution of number of tracks when selecting 1 reconstructed PV, with the specified number of tracks from that reconstructed PV. Both panels show the same plot in a linear (left) and log (right) vertical axis scale.

| Measured Bin | Frac of 0-40 | Frac of 40-60 | Frac of 60-80 | Frac of 80-100 | Frac of 100-200 |
|--------------|--------------|---------------|---------------|----------------|-----------------|
| 0-40         | 1.00         | 0             | 0             | 0              | 0               |
| 40-60        | 0.023        | 0.977         | 0             | 0              | 0               |
| 60-80        | 0.0007       | 0.0306        | 0.9687        | 0              | 0               |
| 80-100       | 0.0011       | 0.0008        | 0.0350        | 0.9631         | 0               |
| 100-200      | 0.0020       | 0.0024        | 0.0017        | 0.0243         | 0.9696          |

Table 5: Relative contributions in each measured multiplicity bin from truth multiplicity bins.

<sup>344</sup> The true values of, for example,  $f_{prompt}$  in each multiplicity bin can be unfolded  
<sup>345</sup> from the data as follows: A measurement  $M_i$  in the  $i^{th}$  bin will consist of a linear  
<sup>346</sup> combination of data which correctly is identified as coming from that multiplicity bin plus  
<sup>347</sup> contributions from other bins which are incorrectly placed due to an unreconstructed PV,  
<sup>348</sup> which are weighted by their relative contributions to the  $i^{th}$  bin. For example, in the bin  
<sup>349</sup> corresponding to the 100-200 track multiplicity, the measurement in that bin is described  
<sup>350</sup> by

<sup>351</sup>

$$M_4 = f_{0in4}Y_0 + f_{1in4}Y_1 + f_{2in4}Y_2 + f_{3in4}Y_3 + f_{4in4}Y_4 \quad (4)$$

<sup>352</sup>

<sup>353</sup> where  $f_{jin4}$  is the fraction of events placed in bin 4 that actually come from bin  $j$ , given  
<sup>354</sup> in Tab. 5, and  $Y_i$  is the true measurement in bin  $i$ . Equivalently, this can be expressed in  
<sup>355</sup> matrix form as

356

$$(5) \quad \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ f_{0in1} & f_{1in1} & 0 & 0 & 0 \\ f_{0in2} & f_{1in2} & f_{2in2} & 0 & 0 \\ f_{0in3} & f_{1in3} & f_{2in3} & f_{3in3} & 0 \\ f_{0in4} & f_{1in4} & f_{2in4} & f_{3in4} & f_{4in4} \end{pmatrix} \times \begin{pmatrix} Y_0 \\ Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix} = \begin{pmatrix} M_0 \\ M_1 \\ M_2 \\ M_3 \\ M_4 \end{pmatrix}$$

357 This therefore gives a system of equations which can be solved for the quantities of  
 358 interest  $Y_i$  via matrix inversion. We note that in the case of the first bin, corresponding  
 359 to the lowest number of tracks,  $M_0 = Y_0$ , since there can be no migration to bins of lower  
 360 multiplicity ( $f_{0in0} = 1$ ). The diagonal row of Tab. 5 gives the fraction of true events in  
 361 each bin, which is greater than 96% in all cases. Therefore the possible bin migration  
 362 effects are expected to be small.

363 This formalism is applied to the fraction of prompt X(3872) measured as a function of  
 364 multiplicity as shown in Tab. 6. In each bin, the effect is less than 1/10th the magnitude  
 365 of the uncertainty on the measurement of  $f_{prompt}$ , and is therefore negligibly small.

| Bin     | $f_{prompt}$ before correction ( $M_i$ ) | $f_{prompt}$ after correction ( $Y_i$ ) | change in value |
|---------|--|---|-----------------|
| 0-40    | 0.939                                    | 0.939                                   | 0               |
| 4-60    | 0.885                                    | 0.884                                   | 0.1%            |
| 60-80   | 0.848                                    | 0.847                                   | 0.1%            |
| 80-100  | 0.727                                    | 0.722                                   | 0.6%            |
| 100-200 | 0.713                                    | 0.712                                   | 0.2%            |

Table 6: Effect of multiplicity bin migration due to missing PV reconstruction on  $f_{prompt}$  of the X(3872).

366 The identical procedure is applied to the finer binning scheme used for the  $\psi(2S)$   
 367 prompt fraction, and yields similar results. The bin migration is shown in Fig. 24, and  
 368 the resulting changes between the measured and unfolded prompt fraction in each bin are  
 369 shown in Tab. 7. We see that in all bins, the effect is 0.5% or smaller. Again, this effect  
 370 is negligible.

## 371 4 Acceptance and Efficiency Corrections

372 One of the goals of this analysis is to determine the ratio of the cross sections X(3872) to  
 373  $\psi(2S)$  in our decay channel. This ratio is given by:

$$\frac{\mathcal{B}[X(3872) \rightarrow J/\psi \pi^+ \pi^-]}{\mathcal{B}[\psi(2S) \rightarrow J/\psi \pi^+ \pi^-]} \times \frac{\sigma_{X(3872)}}{\sigma_{\psi(2S)}} = \frac{N_{X(3872)}}{N_{\psi(2S)}} \frac{\epsilon_{\psi(2S)}}{\epsilon_{X(3872)}} \quad (6)$$

374 where  $N_{\psi(2S)}$  and  $N_{X(3872)}$  are the number of  $\psi(2S)$  and X(3872) counts extracted by  
 375 fitting the peaks in the  $J/\psi \pi^+ \pi^-$  mass spectra and  $\frac{\epsilon_{\psi(2S)}}{\epsilon_{X(3872)}}$  is the ratio of the efficiencies

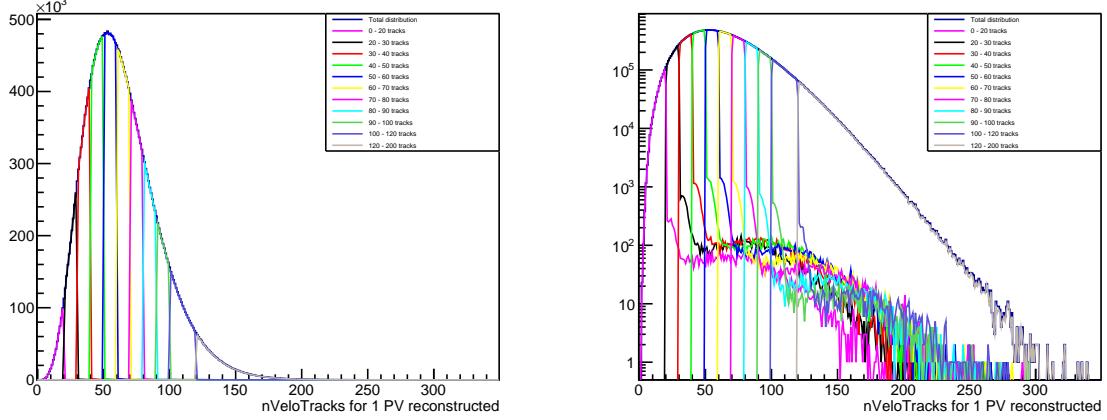


Figure 24: Distribution of number of tracks when selecting 1 reconstructed PV, with the specified number of tracks from that reconstructed PV, for the finer binning scheme used for the  $\psi(2S)$  prompt fraction. Both panels show the same plot in a linear (left) and log (right) vertical axis scale.

| Bin     | $f_{prompt}$ before correction ( $M_i$ ) | $f_{prompt}$ after correction ( $Y_i$ ) | change in value |
|---------|--|---|-----------------|
| 0-20    | 0.868                                    | 0.868                                   | 0               |
| 20-30   | 0.821                                    | 0.820                                   | 0.1%            |
| 30-40   | 0.771                                    | 0.769                                   | 0.26%           |
| 40-50   | .751                                     | 0.750                                   | 0.1%            |
| 50-60   | 0.737                                    | 0.736                                   | 0.1%            |
| 60-70   | 0.736                                    | 0.735                                   | 0.1%            |
| 70-80   | 0.711                                    | 0.709                                   | 0.28%           |
| 80-90   | 0.702                                    | 0.701                                   | 0.1%            |
| 90-100  | 0.699                                    | 0.698                                   | 0.1%            |
| 100-120 | 0.770                                    | 0.767                                   | 0.4%            |
| 120-200 | 0.676                                    | 0.673                                   | 0.5%            |

Table 7: Effect of multiplicity bin migration due to missing PV reconstruction on  $f_{prompt}$  of the  $\psi(2S)$ .

for reconstructing those states. These efficiencies can in turn be factorized into several parts:

$$\frac{\epsilon_{\psi(2S)}}{\epsilon_{X(3872)}} = \frac{\epsilon_{\psi(2S)}^{acc} \epsilon_{\psi(2S)}^{rec} \epsilon_{\psi(2S)}^{PID}}{\epsilon_{X(3872)}^{acc} \epsilon_{X(3872)}^{rec} \epsilon_{X(3872)}^{PID}}, \quad (7)$$

where  $\epsilon^{acc}$  is the acceptance of the LHCb spectrometer for the given decay,  $\epsilon^{rec}$  is the efficiency for selecting and reconstructing the parent particle, and  $\epsilon^{PID}$  is the total particle ID efficiency for the daughters of the decay.

In this section we describe each of these efficiencies that are required to determine the ratio of the  $X(3872)$  and  $\psi(2S)$  cross sections in the transverse momentum range  $p_T > 5$  GeV/ $c$  and rapidity range  $2 < y < 4.5$ . Monte Carlo simulations using the PYTHIA8 event generators coupled to full GEANT4 simulations of the LHCb detector

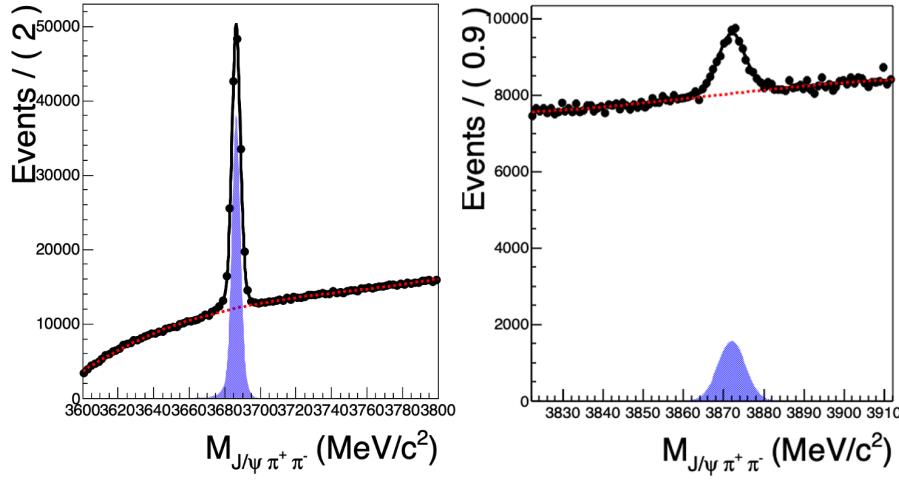


Figure 25: Fits to the  $J/\psi\pi^+\pi^-$  invariant mass spectrum in the  $\psi(2S)$  (left) and  $X(3872)$  (right) signal regions, used as the discriminating variable to determine sWeights.

are used to determine the acceptance, reconstruction, and selection efficiencies. The momentum distributions used in the simulations are reweighted to match the momentum distributions extracted from the data by the method of sWeights [52], which uses the previously described fits to the  $J/\psi\pi^+\pi^-$  invariant mass spectrum as the discriminating variable between signal and background. Corrections are applied to the efficiencies determined from the Monte Carlo simulations in order to reflect differences measured between the MC and data. Calibrated data samples from the PIDCalib package (available at <https://twiki.cern.ch/twiki/bin/view/LHCb/PIDCalibPackage>) are used to determine the muon and pion identification efficiencies. Systematic uncertainties on these corrections are discussed in the relevant subsections.

## 4.1 sWeighted signal

The method of sWeights is commonly used to statistically unfold the distribution of signal and background events in a data sample (see Ref. [52] and <https://indico.cern.ch/event/257864/contributions/1587661/attachments/453353/628570/sweff.pdf> for details). Here we use the invariant mass of the  $J/\psi\pi^+\pi^-$  candidates as the discriminating variable for separating signal and background. The previously described functions are used in a unbinned maximum likelihood fit to the mass regions around the  $\psi(2S)$  and  $X(3872)$  signals, as shown in Fig. 25.

With this formalism, the transverse momentum distributions of signal candidates can be extracted from the unbinned data sets using the signal sWeights. The  $p_T$  distributions for the prompt and non-prompt  $\psi(2S)$  and  $X(3872)$  signal candidates are shown in Figs. 26 and 27, respectively. For the purposes of determining the sWeights, the prompt samples are defined by a cut on the pseudo proper time  $t_z < 0.1$  ps, while the non-prompt sample is defined by  $t_z > 0.1$  ps. Since this analysis is binned in multiplicity, and multiplicity dependent breakup effects may vary with  $p_T$ , the shapes of the  $p_T$  spectra may change depending on the multiplicity of the selected events. To study this possible effect the  $p_T$  spectra are prepared for three different multiplicity classes: all selections of nVeloTracks as shown in black, a high multiplicity sample (defined as candidates from events with

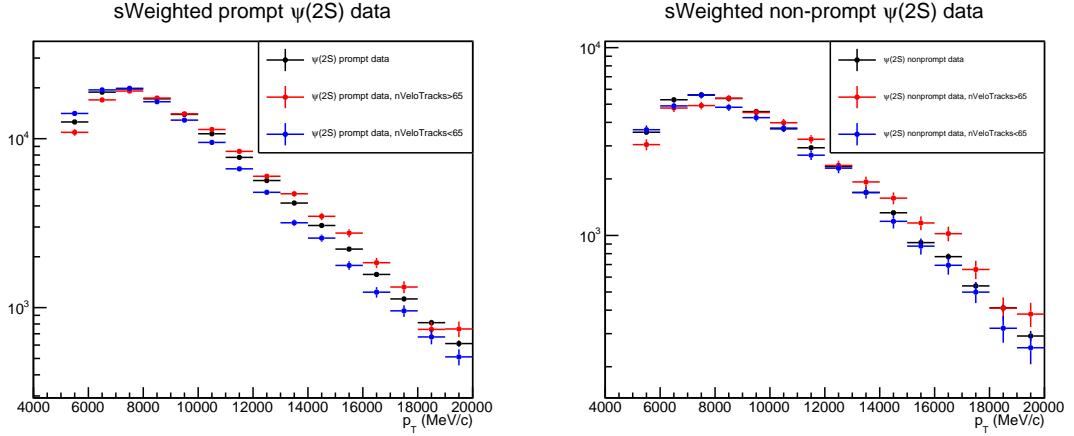


Figure 26: Signal  $p_T$  distributions for  $\psi(2S)$ , determined via sWeights.

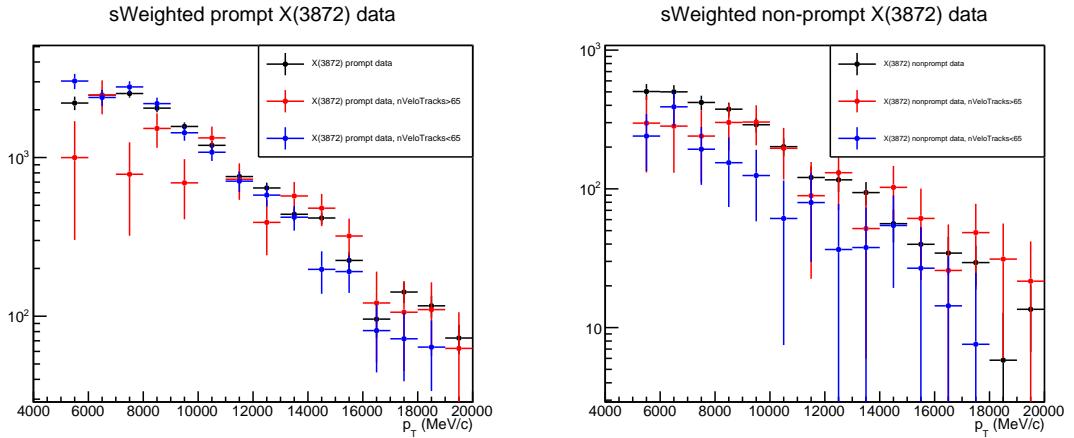


Figure 27: Signal  $p_T$  distributions for  $X(3872)$ , determined via sWeights.

413 greater than the mean number of nVeloTracks 65) in red, and a low multiplicity sample  
 414 (nVeloTracks<65) in blue. The difference between the efficiencies determined using data  
 415 from the various multiplicity classes are used to determine systematic uncertainties on  
 416 the relative efficiency corrections as discussed in the following subsections.

## 417 4.2 Comments on Decay Kinematics

418 The decays of interest here are  $\psi(2S) \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \pi^+ \pi^-$  and  $X(3872) \rightarrow J/\psi (\rightarrow$   
 419  $\mu^+ \mu^-) \rho (\rightarrow \pi^+ \pi^-)$ . While both have a  $\mu^+ \mu^- \pi^+ \pi^-$  final state, the differences in the decay  
 420 chains (namely the intermediate  $\rho$  state in the  $X(3872)$  decay) lead to differences in the  
 421 final state kinematics that do not cancel out when making the ratio of efficiencies. Here  
 422 we look at details of the Monte Carlo samples to get an idea of the differences, before  
 423 proceeding to the exact determination efficiencies using the full simulation chain described  
 424 below.

425 First we examine the mass spectrum of the pion pairs produced in  $\psi(2S)$  and  $X(3872)$   
 426 decays to ensure they represent reality. Using Monte Carlo truth information, the mass  
 427 spectrum  $m_{\pi^+ \pi^-}$  from pions originating from the simulated  $\psi(2S)$  and  $X(3872)$  decays are

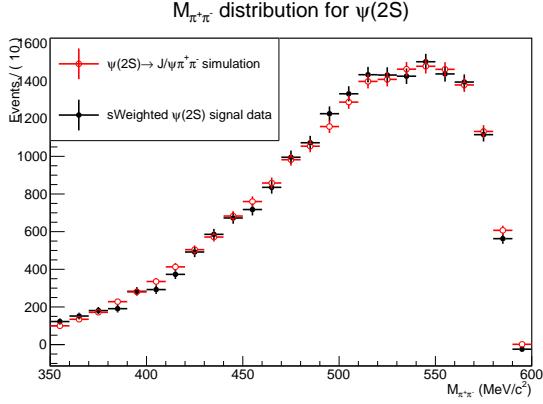


Figure 28: The measured mass spectra of dipions produced in the decays of  $\psi(2S)$ , compared to the MC simulation.

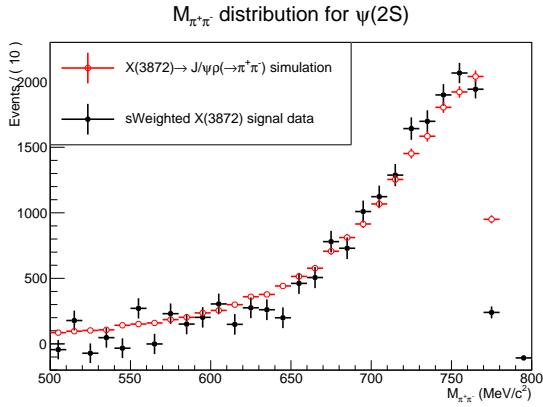


Figure 29: The measured mass spectra of dipions produced in the decays of  $X(3872)$ , compared to the MC simulation.

428 constructed and shown in Fig. 28 and Fig. 29, and compared to the  $m_{\pi^+\pi^-}$  distributions  
 429 extracted from the data using the method of sWeights as described above. As previously  
 430 discussed, the pions from  $X(3872)$  decays come from an intermediate  $\rho(770)$  state, so  
 431 their mass is constrained to follow the kinematics of the decay  $\rho(770) \rightarrow \pi^+\pi^-$ . The  
 432 simulations are in good agreement with the measured data.

433 From the Monte Carlo simulation samples, we take truth-matched  $X(3872)$  and  $\psi(2S)$   
 434 decay daughters, as generated by PYTHIA8. The total momentum  $P$  and transverse  
 435 momentum  $P_T$  of the daughter muons and pions are shown in Figs. 30 and 31, respectively,  
 436 from parents with  $P_T > 5$  GeV/ $c$  as required by this analysis. The kinematic cuts listed  
 437 in Tab. 1 are applied individually to the daughters: we find the cut on muon  $P > 10$   
 438 GeV/ $c$  retains 80% of muons from both decays, and that the cut on muon  $P_T > 650$   
 439 MeV/ $c$  retains 92% of muons from  $\psi(2S)$  decays and 91% of muons from  $X(3872)$  decays.  
 440 No significant differences in efficiency are expected due to the muon cuts.

441 Similarly, the cuts on the daughter pions are applied individually. We find that the  
 442 cut on pion  $P > 3$  GeV/ $c$  retains 70% of pions from  $\psi(2S)$  decays and 79% of pions from  
 443  $X(3872)$  decays, and the cut on pion  $P_T > 500$  MeV/ $c$  retains 54% of pions from  $\psi(2S)$   
 444 decays and 63% of pions from  $X(3872)$  decays. These differences are significant and arise

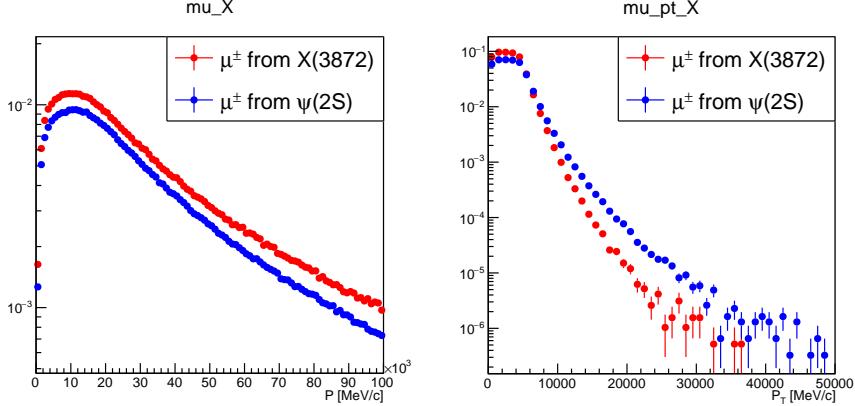


Figure 30: Total momentum  $P$  and transverse momentum  $P_T$  distributions of muons from  $X(3827)$  and  $\psi(2S)$  decays produced by PYTHIA8.

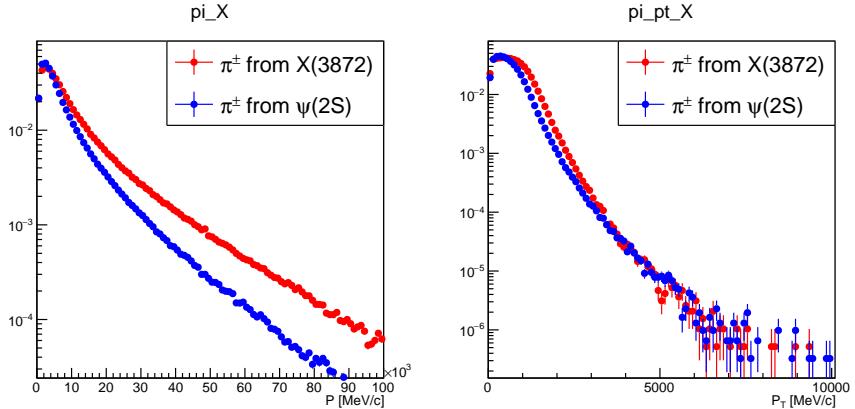


Figure 31: Total momentum  $P$  and transverse momentum  $P_T$  distributions of pions from  $X(3827)$  and  $\psi(2S)$  decays produced by PYTHIA8.

<sup>445</sup> due to the differences in the soft pion decay kinematics as discussed above. Therefore,  
<sup>446</sup> we expect the ratio of total efficiencies  $\frac{\epsilon_{\psi(2S)}}{\epsilon_{X(3872)}}$  to be less than unity. Since two pions  
<sup>447</sup> are required, the differences in efficiency are effectively squared, so from these simple  
<sup>448</sup> considerations we expect  $\frac{\epsilon_{\psi(2S)}}{\epsilon_{X(3872)}} \approx (0.70/0.79)^2 \times (0.54/0.63)^2 = 0.58$ .

### <sup>449</sup> 4.3 LHCb Acceptance Correction

<sup>450</sup> First, we consider the efficiency for particles which are produced in our selected range of  
<sup>451</sup> transverse momentum, rapidity, and primary vertex  $z$  range to produce daughters that  
<sup>452</sup> are within the LHCb spectrometer acceptance, defined as  $10\text{mrad} < \theta < 400\text{ mrad}$ . This  
<sup>453</sup> efficiency is defined as

$$\epsilon_{acc} = \frac{\text{Number of parents in selected pt, y, PVZ range with all daughters in LHCb acceptance}}{\text{Number of parents generated in selected pt, y, PVZ range}} \quad (8)$$

<sup>454</sup> where the selected transverse momentum range is  $p_T > 5\text{ GeV}/c$ , the rapidity range is

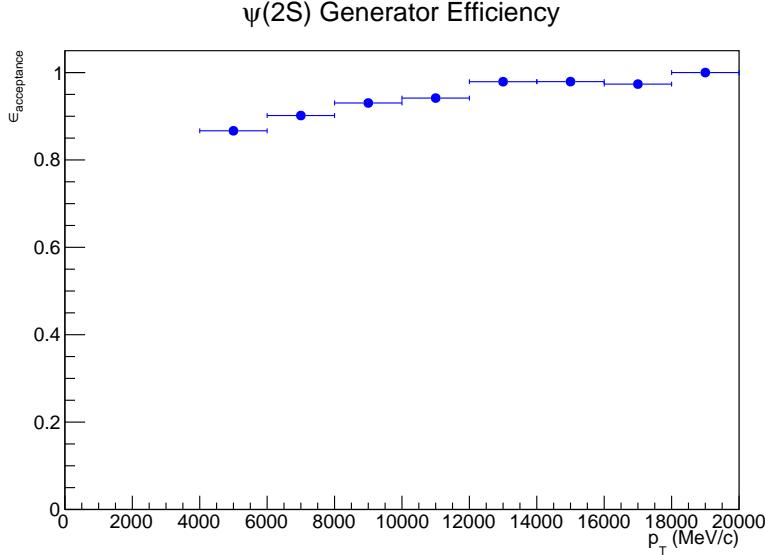


Figure 32: The acceptance  $\epsilon_{acc}$  for  $\psi(2S) \rightarrow J/\psi \pi^+ \pi^-$  decays as a function of  $p_T$ .

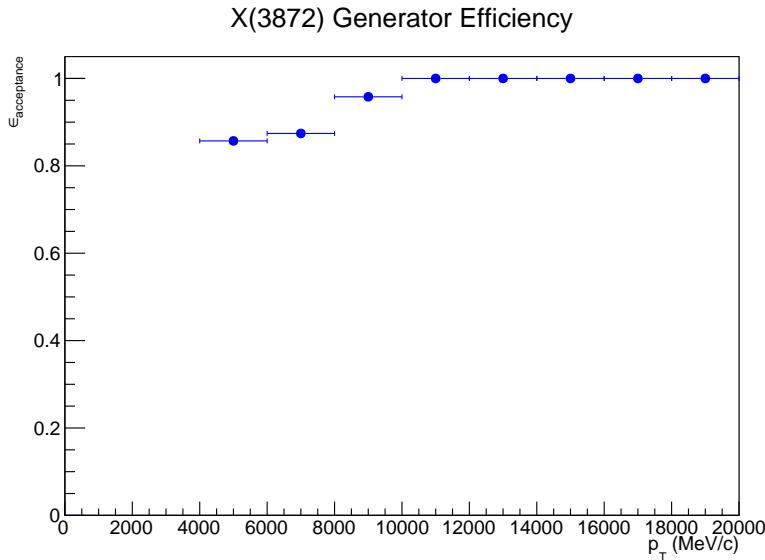


Figure 33: The acceptance  $\epsilon_{acc}$  for  $X(3872) \rightarrow J/\psi \pi^+ \pi^-$  decays as a function of  $p_T$ .

455     $2 < y < 4.5$ , and the primary vertex  $z$  range is  $-60\text{mm} < \text{PVZ} < 120\text{mm}$ . This is evaluated  
 456    using generator-level simulations of event types 28144001 for  $\psi(2S)$  and 28144012 for  
 457     $X(3872)$ , where the  $p_T$  distributions are weighted to match the distributions obtained from  
 458    sWeights. The acceptance correction factors for  $X(3872)$  and  $\psi(2S)$  decays to  $J/\psi \pi^+ \pi^-$ ,  
 459    obtained using the simulations weighted to match the  $p_T$  distributions of the prompt  
 460    signal from the inclusive multiplicity sample, are shown as a function of  $p_T$  in Figs. 32 and  
 461    33, respectively. This analysis is only concerned with parent particles that have  $p_T > 5$   
 462     $\text{GeV}/c$ , where the efficiencies are determined to be  $\epsilon_{acc} = 0.895$  for  $X(3872)$ , and  $\epsilon_{acc} =$   
 463    0.897 for  $\psi(2S)$  produced in the rapidity window  $2 < y < 4.5$ .

464    To determine the affect of the varying  $p_T$  distributions, the acceptance calculation is

<sup>465</sup> repeated, by varying the weights to give  $p_T$  distributions which match the high and low  
<sup>466</sup> multiplicity samples for both prompt and non-prompt. Since this analysis is concerned  
<sup>467</sup> with measurements of the ratios of cross sections, the ratio  $\epsilon_{\psi(2S)}^{acc}/\epsilon_{X(3872)}^{acc}$  is given for each  
<sup>468</sup> case. The resulting efficiencies are shown in Tab. 8.

| $p_T$ parameterization | $\epsilon_{\psi(2S)}^{acc}/\epsilon_{X(3872)}^{acc}$ |
|------------------------|--|
| Prompt                 | $0.897/0.895 = 1.002$                                |
| Prompt, hi mult.       | $0.900/0.905 = 0.994$                                |
| Prompt, low mult.      | $0.895/0.888 = 1.008$                                |
| Non-prompt             | $0.900/0.877 = 1.026$                                |
| Non-prompt, hi mult.   | $0.903/0.885 = 1.020$                                |
| Non-prompt, low mult.  | $0.899/0.876 = 1.026$                                |

Table 8: Variations of the LHCb acceptance calculation for the ratio of  $\psi(2S)$  and  $X(3872)$  efficiencies.

<sup>469</sup> For both prompt and non-prompt, we take the average of the three cases as the  
<sup>470</sup> central value for  $\epsilon_{\psi(2S)}^{acc}/\epsilon_{X(3872)}^{acc}$ . A systematic uncertainty is determined by calculating  
<sup>471</sup> the standard deviation between these samples, defined as

$$\sigma = \sqrt{\frac{1}{N-1} \sum (\epsilon_i - \bar{\epsilon})^2} \quad (9)$$

<sup>472</sup> where  $N$  is the sample size (3 in this case),  $\epsilon_i$  is the value determined from each trial,  
<sup>473</sup> and  $\bar{\epsilon}$  is the mean value. For both prompt and non-prompt, this is rounded up to 1%  
<sup>474</sup> An additional conservative systematic uncertainty of 1% is assigned to account for the  
<sup>475</sup> finite statistics of the Monte Carlo sample used to determine this efficiency, and added in  
<sup>476</sup> quadrature. We thereby determine the relative correction factor for prompt production to  
<sup>477</sup> be

$$\epsilon_{prompt\psi(2S)}^{acc}/\epsilon_{promptX(3872)}^{acc} = 1.00 \pm 0.014 \quad (10)$$

<sup>478</sup> and for non-prompt production

$$\epsilon_{non-prompt\psi(2S)}^{acc}/\epsilon_{non-promptX(3872)}^{acc} = 1.02 \pm 0.014 \quad (11)$$

<sup>479</sup> For the extraction of the prompt fraction  $f_{prompt}$ , the ratio of efficiencies for prompt  
<sup>480</sup> and non-prompt production of the same species are required. These ratios are shown in  
<sup>481</sup> Tab. 9.

<sup>482</sup> Following the same procedure as described above, we find

$$\epsilon_{prompt\psi(2S)}^{acc}/\epsilon_{non-prompt\psi(2S)}^{acc} = 1.00 \pm 0.014 \quad (12)$$

<sup>483</sup> and

$$\epsilon_{promptX(3872)}^{acc}/\epsilon_{non-promptX(3872)}^{acc} = 1.02 \pm 0.014 \quad (13)$$

| Species    | $p_T$ parameterization | $\epsilon_{prompt}^{acc}/\epsilon_{non-prompt}^{acc}$ |
|------------|------------------------|---|
| $\psi(2S)$ | Inclusive              | 0.897/0.900 = 0.997                                   |
| $\psi(2S)$ | High mult.             | 0.900/0.903 = 0.997                                   |
| $\psi(2S)$ | Low mult.              | 0.895/0.899 = 0.996                                   |
| $X(3872)$  | Inclusive              | 0.895/0.877 = 1.021                                   |
| $X(3872)$  | High mult.             | 0.905/0.885 = 1.023                                   |
| $X(3872)$  | Low mult.              | 0.888/0.876 = 1.014                                   |

Table 9: Variations of the LHCb acceptance calculation for the ratio of  $\psi(2S)$  and  $X(3872)$  efficiencies.

#### 484 4.4 Reconstruction and Selection Efficiencies

485 The next step is determining the efficiency for reconstructing and selecting the parent  
 486 particles from the daughters that are in the LHC acceptance, given the selection criteria  
 487 imposed by the requirements imposed by the stripping and analysis cuts. This efficiency  
 488  $\epsilon_{rec}$  is defined as

$$\epsilon_{rec} = \frac{\text{Number of reconstructed and selected parents in selected pt, y, PVZ range}}{\text{Number of parents in selected pt, y, PVZ range with all daughters in LHCb acceptance}}. \quad (14)$$

489 We use the Gauss package with the full GEANT4 simulation of the LHCb  
 490 detector to create simulated samples of events containing our particles of inter-  
 491 est. The same stripping selections used in the data are applied to the sim-  
 492 ulation samples, specifically, the `StrippingFullDSTDiMuonJpsi2MuMuTOSLine` selec-  
 493 tions from stripping 21, which includes trigger requirements fo L0 Muon TOS,  
 494 `Hlt1DiMuonHighMassDecision` TOS, and `Hlt2DiMuonJPsiHighPTDecision` TOS (see  
 495 [http://lhcbdoc.web.cern.ch/lhcbdoc/stripping/config/stripping21/  
 496 dimuon/strippingfullstdimuonjpsiz2mumutosline.html](http://lhcbdoc.web.cern.ch/lhcbdoc/stripping/config/stripping21/dimuon/strippingfullstdimuonjpsiz2mumutosline.html) for a full description of the  
 497 stripping line requirements) . The  $X(3872)$  simulations for the 2012  $pp$  data set use  
 498 event type 28144011, which is the same event type as used for the previously discussed  
 499 acceptance correction, but with the decay product cuts applied to minimize computation  
 500 time. PYTHIA8 is the underlying event generator, and the simulations performed for the  
 501 MagUp and MagDown field configurations each contain  $10^6$  events. The MagDown data  
 502 can be found in the bookkeeping at:

503 `/MC/2012/Beam4000GeV-2012-MagDown-Nu2.5-Pythia8/Sim09d/Trig0x409f0045/  
 504 Reco14c/Stripping21NoPrescalingFlagged/28144011/ALLSTREAMS.DST`

505 and the MagUp at:

506 `/MC/2012/Beam4000GeV-2012-MagUp-Nu2.5-Pythia8/Sim09d/Trig0x409f0045/  
 507 Reco14c/Stripping21NoPrescalingFlagged/28144011/ALLSTREAMS.DST`

508 Similarly, the full GEANT4 simulations of  $\psi(2S)$  production use event type 28144002,  
 509 which consists of PYTHIA8 events containing  $\psi(2S)$  that decay to  $J/\psi\pi^+\pi^-$  with all four  
 510 daughters in the LHCb acceptance. The MagDown and MagUp data samples each contain  
 511  $1.6 \times 10^6$  events, and can be found in the bookkeeping at

512 `/MC/2012/Beam4000GeV-2012-MagDown-Nu2.5-Pythia8/Sim09e/Trig0x409f0045/  
 513 Reco14c/Stripping21NoPrescalingFlagged/28144002/ALLSTREAMS.DST`

514 and

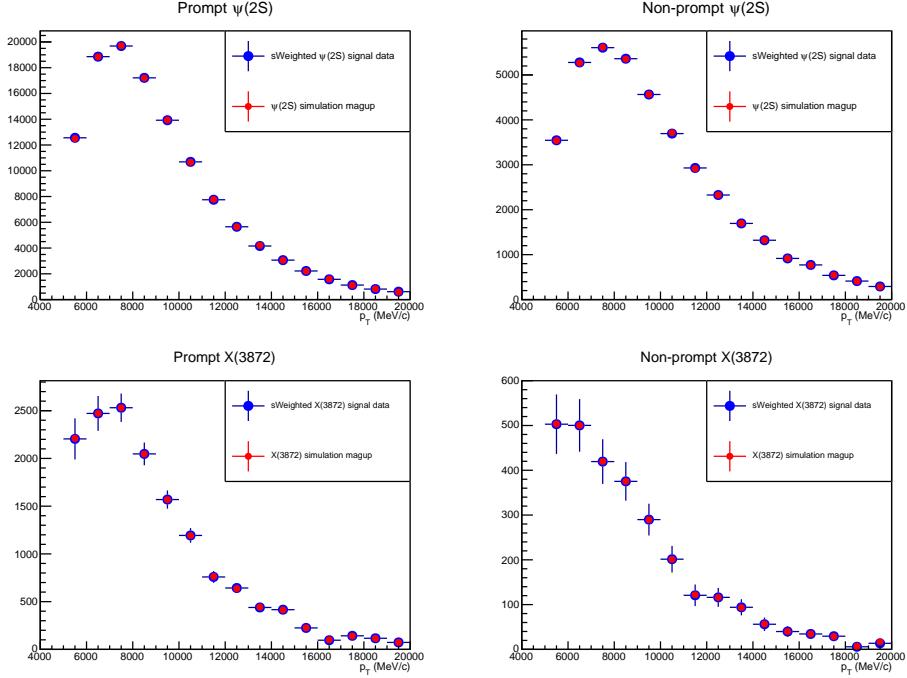


Figure 34: Transverse momentum distributions from the reconstructed  $\psi(2S)$  and  $X(3872)$  from simulation (red), which have been weighted to match the distributions extracted from the sWeighted data (blue).

515 /MC/2012/Beam4000GeV-2012-MagUp-Nu2.5-Pythia8/Sim09e/Trig0x409f0045/  
 516 Reco14c/Stripping21NoPrescalingFlagged/28144002/ALLSTREAMS.DST ,  
 517 respectively.

518 As was done for the previously described acceptance correction, the  $p_T$  distributions  
 519 in the simulated samples are weighted to match the distributions extracted from the  
 520 sWeighted signal data sets. Figure. 34 shows the  $p_T$  spectra from reconstructed  $\psi(2S)$   
 521 and  $X(3872)$  simulations and the multiplicity inclusive data signal data from sWeights.  
 522 We see there is good agreement, by construction.

523 It is known that there are slight differences between the reconstruction efficiency deter-  
 524 mined directly from Monte Carlo simulations, and reconstruction efficiencies determined  
 525 via tag-and-probe methods applied to the data. These differences have been extensively  
 526 studied and quantified for this data set, as documented on the Tracking Efficiencies Twiki:  
 527 <https://twiki.cern.ch/twiki/bin/view/LHCbInternal/LHCbTrackingEfficiencies>. By com-  
 528 paring the Monte Carlo and data, a map of correction factors for the 2012  $pp$  data set has  
 529 been prepared, as shown in Fig. 35. These corrections are applied track-by-track when  
 530 calculating the reconstruction efficiencies, in order to correct the Monte Carlo simulations  
 531 to the measured efficiencies.

532 The total reconstruction and selection efficiency  $\epsilon_{rec}$  is constructed in several steps  
 533 we describe here. For this illustration, we choose the  $p_T$  parametrization determined  
 534 by weighting the simulation to match the data extracted from the multiplicity-inclusive  
 535 prompt  $X(3872)$  and  $\psi(2S)$  signal. First, Monte Carlo truth information from the  
 536 MCTruthTupleTool is used to construct the spectrum of generated particles that were  
 537 used in the Monte Carlo simulation of the LHCb detector, as shown in blue in Figs. 36 and

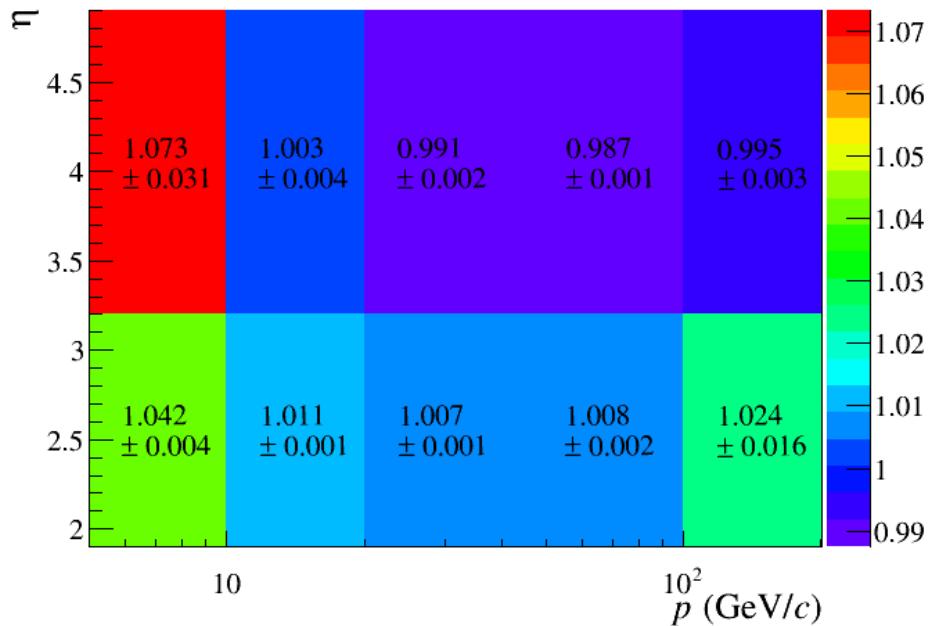


Figure 35: Map of differences between tracking efficiencies determined via data and Monte Carlo. These corrections are applied track-by-track when calculating the reconstruction efficiencies.

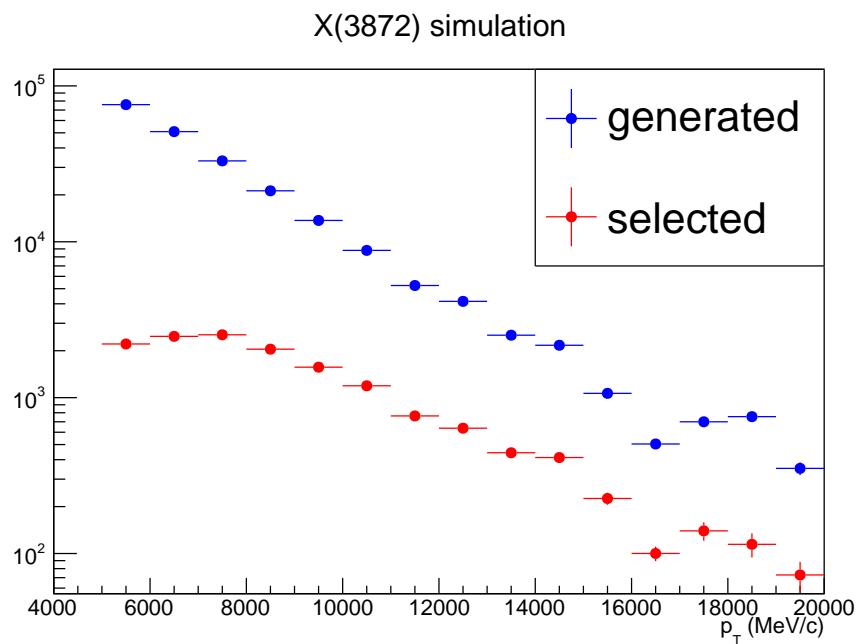


Figure 36: The generated and accepted simulated X(3872) particles as a function of  $p_T$ .

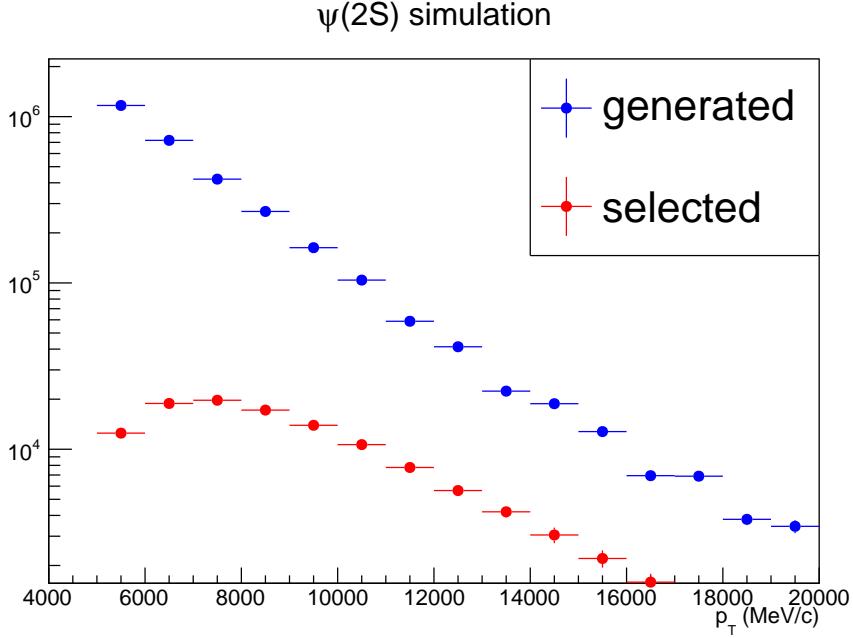


Figure 37: The generated and accepted simulated  $\psi(2S)$  particles as a function of  $p_T$ .

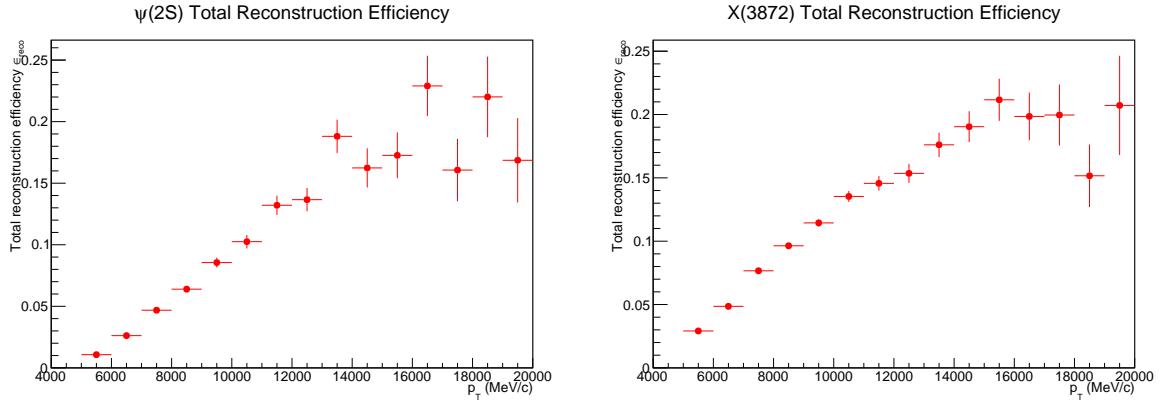


Figure 38: Total efficiency for reconstructing simulated  $\psi(2S)$ .

Figure 39: Total efficiency for reconstructing simulated  $X(3872)$ .

538 37. The `StrippingFullDSTDiMuonJpsi2MuMuTOSLine` line from stripping 21 is applied  
 539 to the generated events, and the selection criteria discussed previously is applied to decay  
 540 candidates which have passed the stripping requirements. Integrating over the  $p_T$  range  
 541 relevant to this measurement ( $p_T > 5$  GeV/c) we find that the reconstruction-selection  
 542 efficiency  $\epsilon_{rec}$  is equal to 6.75%(6.73%) for prompt  $X(3872)$  and 3.97% (3.92%) for prompt  
 543  $\psi(2S)$  in the Mag Up (Mag Down) configurations in our phase space of interest,  $p_T > 5$   
 544 GeV/c and  $2 < y < 4.5$ . These efficiencies are shown as a function of transverse momentum  
 545 in Figs. 38 and 39. The ratio of these efficiencies is shown as a function of  $p_T$  in Fig.  
 546 40. We see that the ratio approaches one for large values of parent  $p_T$ . At relatively low  
 547 values of  $p_T$ , the differences in efficiency arise due to different kinematics of the daughter  
 548 products. However, in the range where  $p_T >>$  parent mass, the kinematics of the decay  
 549 products are dominated by the large relativistic boost inherited from the parents, which

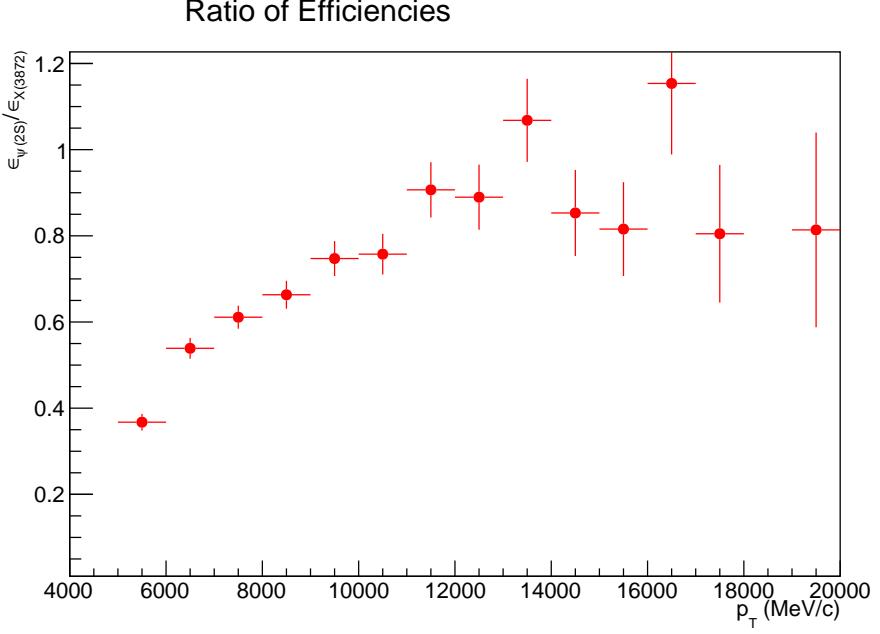


Figure 40: The ratio of efficiencies  $\epsilon_{prompt\psi(2S)}^{reco}/\epsilon_{promptX(3872)}^{reco}$  for promptly produced particles as a function of  $p_T$ . The ratio approaches one for  $p_T \gg M$ , where the kinematics of the daughter products are dominated by inherited relativistic boost.

leads to similar efficiencies.

As was done for the acceptance correction discussed in the previous section, we re-weight the simulations to match the distributions extracted from the sWeighted data and repeat the entire simulation chain for prompt and non-prompt  $\psi(2S)$  and  $X(3872)$ . These simulations are done in both the Magnet Up and Magnet Down configurations. The uncertainties on the data/MC corrections shown in Fig. 35 are taken into account by increasing or decreasing all the efficiency correction factors by their quoted uncertainty, and re-running the simulation in order to determine maximum bounds on the possible effects. This gives three different parameterizations of the momentum spectra (from high multiplicity, low multiplicity, or the multiplicity inclusive data sample), three different values for the Monte Carlo/data corrections (central value,  $+1\sigma$ , and  $-1\sigma$ ), and two different configurations for the magnet (up and down), for a total of 18 different simulation conditions for prompt and non-prompt  $X(3872)$  and  $\psi(2S)$ . The resulting reconstruction efficiencies are calculated for each case and given in Tabs. 10 through 14. We take the mean of these 18 different trials as the central value for  $\epsilon^{reco}$  for each species, and assign the standard deviation between the 18 trials as a systematic uncertainty, which is given at the bottom of each table. We see that the  $\psi(2S)$  reconstruction efficiency is lower than the  $X(3872)$  reconstruction efficiency, as expected from our previous considerations of the decay kinematics. The prompt and non-prompt reconstruction efficiencies for each species agree within uncertainties.

Since the goal of this analysis is the measurement of the ratio of cross sections of  $X(3872)$  and  $\psi(2S)$ , the ratio of efficiencies  $\epsilon_{\psi(2S)}/\epsilon_{X(3872)}$  is the relevant quantity. This quantity is calculated for each case of  $p_T$  shape, magnet configuration, and MC/data correction and collected in Tabs. 16 and 16 for prompt and non-prompt production,

| Species           | $p_T$ parameterization | Magnet Configuration | MC correction    | $\epsilon$        |
|-------------------|------------------------|----------------------|------------------|-------------------|
| Prompt $\psi(2S)$ | Inclusive              | Up                   | Central          | 0.0397            |
| Prompt $\psi(2S)$ | Inclusive              | Up                   | +1 $\sigma$      | 0.0402            |
| Prompt $\psi(2S)$ | Inclusive              | Up                   | -1 $\sigma$      | 0.0391            |
| Prompt $\psi(2S)$ | Inclusive              | Down                 | Central          | 0.0392            |
| Prompt $\psi(2S)$ | Inclusive              | Down                 | +1 $\sigma$      | 0.0397            |
| Prompt $\psi(2S)$ | Inclusive              | Down                 | -1 $\sigma$      | 0.0386            |
| Prompt $\psi(2S)$ | High Multiplicity      | Up                   | Central          | 0.0426            |
| Prompt $\psi(2S)$ | High Multiplicity      | Up                   | +1 $\sigma$      | 0.0432            |
| Prompt $\psi(2S)$ | High Multiplicity      | Up                   | -1 $\sigma$      | 0.0420            |
| Prompt $\psi(2S)$ | High Multiplicity      | Down                 | Central          | 0.0420            |
| Prompt $\psi(2S)$ | High Multiplicity      | Down                 | +1 $\sigma$      | 0.0426            |
| Prompt $\psi(2S)$ | High Multiplicity      | Down                 | -1 $\sigma$      | 0.0415            |
| Prompt $\psi(2S)$ | Low Multiplicity       | Up                   | Central          | 0.0388            |
| Prompt $\psi(2S)$ | Low Multiplicity       | Up                   | +1 $\sigma$      | 0.0372            |
| Prompt $\psi(2S)$ | Low Multiplicity       | Up                   | -1 $\sigma$      | 0.0362            |
| Prompt $\psi(2S)$ | Low Multiplicity       | Down                 | Central          | 0.0382            |
| Prompt $\psi(2S)$ | Low Multiplicity       | Down                 | +1 $\sigma$      | 0.0367            |
| Prompt $\psi(2S)$ | Low Multiplicity       | Down                 | -1 $\sigma$      | 0.0357            |
|                   |                        |                      | $\mu \pm \sigma$ | 0.040 $\pm$ 0.002 |

Table 10: Variations of the prompt  $\psi(2S)$  reconstruction and selection efficiency, for different  $p_T$  spectrum parameterizations, magnetic field configurations, and MC/data corrections.

574 respectively.

#### 575 4.4.1 Rapidity Distribution Check

576 Fig. 41 shows a comparison between the rapidity distributions of  $\psi(2S)$  from the 13 and  
 577 7 TeV data sets published in [50], and the PYTHIA8 simulation used here. We see a  
 578 reasonable agreement. A full study of the sensitivity of our corrections to the rapidity  
 579 distribution is performed by reweighting the simulation to match the data, similar to what  
 580 was done previously for the  $p_T$  distributions. In this case, the uncertainties on the data  
 581 are dominated by a fully correlated systematic uncertainty that arises from the trigger  
 582 and reconstruction efficiency corrections. Thus, all data points can move up or down  
 583 together, but there is no variation in the shape of the distribution.

584 We construct weights that are applied to the simulation to match the distributions  
 585 measured at 7 and 13 TeV. Fig. 42 shows the distributions after weighting, where there  
 586 is good agreement (by construction). With these weights applied, the simulations are  
 587 re-run and the resulting  $\epsilon_{\psi(2S)}^{reco}/\epsilon_{X(3872)}^{reco}$  shows a maximum variation of 0.1%. This is totally  
 588 negligible compared to the systematic uncertainty on the relative acceptance correction  
 589 discussed in the preceding section, so no additional uncertainties are assigned.

#### 590 4.4.2 PVZ Distribution

591 As previously discussed, events considered in the analysis are limited to the primary vertex  
 592  $z$  range of -60 cm  $< PVZ <$  120 cm, to ensure a stable acceptance for the measurement of

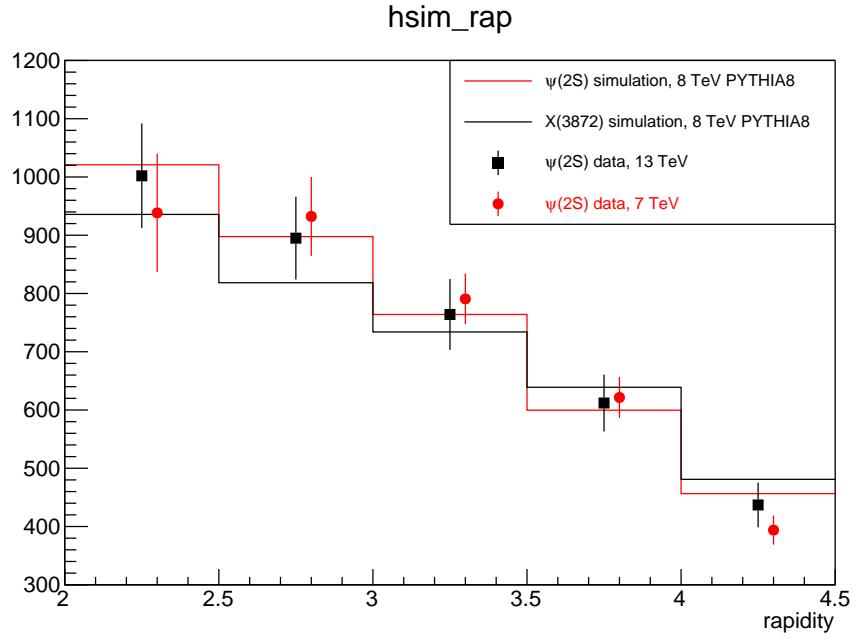


Figure 41: Rapidity distributions of  $\psi(2S)$  from the 13 and 7 TeV data sets [50] and the PYTHIA8 simulation. The 7 TeV data is slightly displaced from the bin center for clarity.

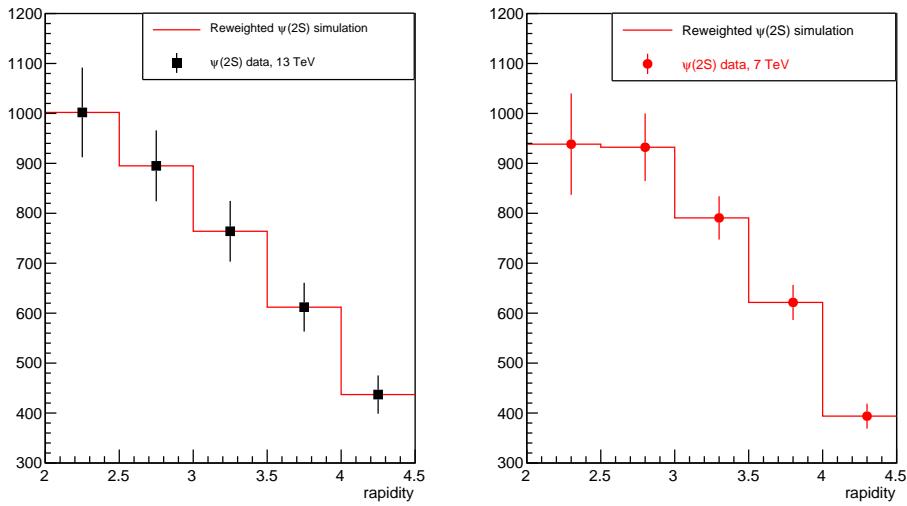


Figure 42: The simulated rapidity distributions, reweighted to match the data from [50]. This gives a negligibly small change in the reconstruction efficiencies.

| Species               | $p_T$ parameterization | Magnet Configuration | MC correction    | $\epsilon$        |
|-----------------------|------------------------|----------------------|------------------|-------------------|
| Non-prompt $\psi(2S)$ | Inclusive              | Up                   | Central          | 0.0406            |
| Non-prompt $\psi(2S)$ | Inclusive              | Up                   | +1 $\sigma$      | 0.0411            |
| Non-prompt $\psi(2S)$ | Inclusive              | Up                   | -1 $\sigma$      | 0.0400            |
| Non-prompt $\psi(2S)$ | Inclusive              | Down                 | Central          | 0.0379            |
| Non-prompt $\psi(2S)$ | Inclusive              | Down                 | +1 $\sigma$      | 0.0384            |
| Non-prompt $\psi(2S)$ | Inclusive              | Down                 | -1 $\sigma$      | 0.0374            |
| Non-prompt $\psi(2S)$ | High Multiplicity      | Up                   | Central          | 0.0439            |
| Non-prompt $\psi(2S)$ | High Multiplicity      | Up                   | +1 $\sigma$      | 0.0445            |
| Non-prompt $\psi(2S)$ | High Multiplicity      | Up                   | -1 $\sigma$      | 0.0434            |
| Non-prompt $\psi(2S)$ | High Multiplicity      | Down                 | Central          | 0.0413            |
| Non-prompt $\psi(2S)$ | High Multiplicity      | Down                 | +1 $\sigma$      | 0.0418            |
| Non-prompt $\psi(2S)$ | High Multiplicity      | Down                 | -1 $\sigma$      | 0.0407            |
| Non-prompt $\psi(2S)$ | Low Multiplicity       | Up                   | Central          | 0.0398            |
| Non-prompt $\psi(2S)$ | Low Multiplicity       | Up                   | +1 $\sigma$      | 0.0404            |
| Non-prompt $\psi(2S)$ | Low Multiplicity       | Up                   | -1 $\sigma$      | 0.0366            |
| Non-prompt $\psi(2S)$ | Low Multiplicity       | Down                 | Central          | 0.0371            |
| Non-prompt $\psi(2S)$ | Low Multiplicity       | Down                 | +1 $\sigma$      | 0.0376            |
| Non-prompt $\psi(2S)$ | Low Multiplicity       | Down                 | +1 $\sigma$      | 0.0366            |
|                       |                        |                      | $\mu \pm \sigma$ | $0.040 \pm 0.002$ |

Table 11: Variations of the non-prompt  $\psi(2S)$  reconstruction and selection efficiency, for different  $p_T$  spectrum parameterizations, magnetic field configurations, and MC/data corrections.

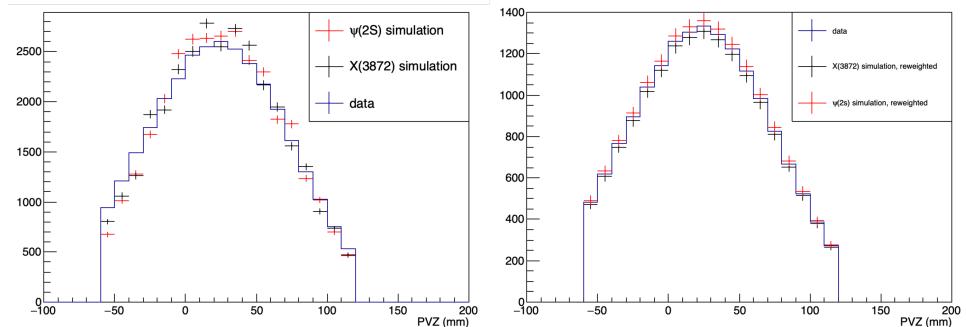


Figure 43: Left: PVZ distributions from simulation and data. Right: simulated PVZ distributions reweighted to match the data (slightly offset on the vertical axis for clarity). This results in a negligibly small change in the efficiency of < 0.3% .

593 the number of tracks in the VELO detector. A comparison of the PVZ distributions in  
 594 the data and Monte Carlo samples shows some slight differences, see left panel of 43. We  
 595 investigate the effects these differences by re-weighting the simulations to precisely match  
 596 the PVZ distributions found in the data as shown in the right panel of 43. The efficiency  
 597 simulations are run with this weighting applied, and a negligibly small change of < 0.3%  
 598 is found. This is not unexpected: since the PVZ range is limited to the center of the  
 599 VELO where tracking is at full efficiency, slight changes should not have any appreciable  
 600 effect.

| Species    | $p_T$ parameterization | Magnet Configuration | MC correction    | $\epsilon_{\psi(2S)}^{prompt}/\epsilon_{\psi(2S)}^{non-prompt}$ |
|------------|------------------------|----------------------|------------------|---|
| $\psi(2S)$ | Inclusive              | Up                   | Central          | 0.978   |
| $\psi(2S)$ | Inclusive              | Up                   | +1 $\sigma$      | 0.978   |
| $\psi(2S)$ | Inclusive              | Up                   | -1 $\sigma$      | 0.978   |
| $\psi(2S)$ | Inclusive              | Down                 | Central          | 1.034   |
| $\psi(2S)$ | Inclusive              | Down                 | +1 $\sigma$      | 1.034   |
| $\psi(2S)$ | Inclusive              | Down                 | -1 $\sigma$      | 1.032   |
| $\psi(2S)$ | High Mult              | Up                   | Central          | 0.970   |
| $\psi(2S)$ | High Mult              | Up                   | +1 $\sigma$      | 0.971   |
| $\psi(2S)$ | High Mult              | Up                   | -1 $\sigma$      | 0.968   |
| $\psi(2S)$ | High Mult              | Down                 | Central          | 1.017   |
| $\psi(2S)$ | High Mult              | Down                 | +1 $\sigma$      | 1.019   |
| $\psi(2S)$ | High Mult              | Down                 | -1 $\sigma$      | 1.020   |
| $\psi(2S)$ | Low Mult               | Up                   | Central          | 0.975   |
| $\psi(2S)$ | Low Mult               | Up                   | +1 $\sigma$      | 0.921   |
| $\psi(2S)$ | Low Mult               | Up                   | -1 $\sigma$      | 0.989   |
| $\psi(2S)$ | Low Mult               | Down                 | Central          | 1.030   |
| $\psi(2S)$ | Low Mult               | Down                 | +1 $\sigma$      | 0.976   |
| $\psi(2S)$ | Low Mult               | Down                 | -1 $\sigma$      | 0.975   |
|            |                        |                      | $\mu \pm \sigma$ | $0.99 \pm 0.03$   |

Table 12: Variations of the ratio of reconstruction and selection efficiencies for non-prompt  $\psi(2S)$ /prompt  $\psi(2S)$ , for different  $p_T$  spectrum parameterizations, magnetic field configurations, and MC/data corrections.

#### 4.4.3 Multiplicity dependent effects

It is well known that there are changes in the absolute tracking efficiency with high numbers of tracks in the LHCb detector, which become especially important in heavy ion collisions and high-multiplicity  $pp$  data with multiple interaction vertices per event. This has been studied in detail for the data set used here, and published in JINST [53]. Over the range of tracks considered in this analysis, from Fig. 44 we see that the absolute tracking efficiency (measured via tag-and-probe methods using actual data) varies from 97.5% to 96%. The final state we consider here requires four long tracks to be reconstructed in the spectrometer, so this maximum variation of 1.5% per track gives a maximum reduction in absolute efficiency of  $1 - (0.985)^4 = 6\%$ , as we go from low multiplicity events to events with 200 long tracks. However, the point of this analysis is measuring ratios of the cross sections of  $X(3872)$  to  $\psi(2S)$ , so our concern here is not the absolute tracking efficiency, but rather the relative efficiency  $\epsilon_{\psi(2S)}^{reco}/\epsilon_{X(3872)}^{reco}$ . To a first approximation, the dependence on number of tracks should cancel out when making this ratio, as both final states require four long tracks to be reconstructed.

To investigate this more fully, we consider the distribution of nVeloTracks for events which contain candidate  $X(3872)$  and  $\psi(2S)$  decays, from the data. The mass distribution for candidates which pass all selection criteria versus the number of Velo tracks in the event the candidate in is shown in Fig. 45. Three regions are selected, as shown by the black lines. Two of the regions are with  $\pm 5$  MeV of the nominal masses of the  $\psi(2S)$  and  $X(3872)$ , while the third, wider region is in between the peaks. This background region

| Species        | $p_T$ parameterization | Magnet Configuration | MC correction    | $\epsilon$        |
|----------------|------------------------|----------------------|------------------|-------------------|
| Prompt X(3872) | Inclusive              | Up                   | Central          | 0.0675            |
| Prompt X(3872) | Inclusive              | Up                   | +1 $\sigma$      | 0.0684            |
| Prompt X(3872) | Inclusive              | Up                   | -1 $\sigma$      | 0.0667            |
| Prompt X(3872) | Inclusive              | Down                 | Central          | 0.0673            |
| Prompt X(3872) | Inclusive              | Down                 | +1 $\sigma$      | 0.0681            |
| Prompt X(3872) | Inclusive              | Down                 | -1 $\sigma$      | 0.0664            |
| Prompt X(3872) | High Multiplicity      | Up                   | Central          | 0.0751            |
| Prompt X(3872) | High Multiplicity      | Up                   | +1 $\sigma$      | 0.0760            |
| Prompt X(3872) | High Multiplicity      | Up                   | -1 $\sigma$      | 0.0742            |
| Prompt X(3872) | High Multiplicity      | Down                 | Central          | 0.0755            |
| Prompt X(3872) | High Multiplicity      | Down                 | +1 $\sigma$      | 0.0764            |
| Prompt X(3872) | High Multiplicity      | Down                 | -1 $\sigma$      | 0.0746            |
| Prompt X(3872) | Low Multiplicity       | Up                   | Central          | 0.0618            |
| Prompt X(3872) | Low Multiplicity       | Up                   | +1 $\sigma$      | 0.0626            |
| Prompt X(3872) | Low Multiplicity       | Up                   | -1 $\sigma$      | 0.0611            |
| Prompt X(3872) | Low Multiplicity       | Down                 | Central          | 0.0613            |
| Prompt X(3872) | Low Multiplicity       | Down                 | +1 $\sigma$      | 0.0664            |
| Prompt X(3872) | Low Multiplicity       | Down                 | -1 $\sigma$      | 0.0605            |
|                |                        |                      | $\mu \pm \sigma$ | 0.068 $\pm$ 0.006 |

Table 13: Variations of the prompt X(3872) reconstruction and selection efficiency, for different  $p_T$  spectrum parameterizations, magnetic field configurations, and MC/data corrections.

622 contains only combinatorial background from random combinations of  $J/\psi\pi^+\pi^-$ , and  
 623 thereby has a distribution of nVeloTracks that represents a pure sample of the underlying

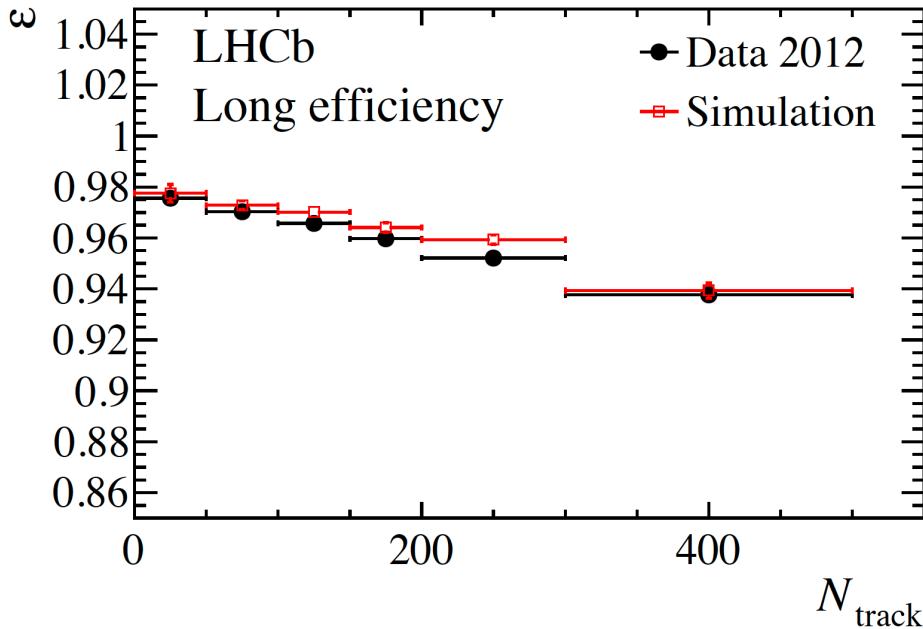


Figure 44: Track reconstruction efficiency for long tracks as a function of event multiplicity [53].

| Species            | $p_T$ parameterization | Magnet Configuration | MC correction    | $\epsilon$        |
|--------------------|------------------------|----------------------|------------------|-------------------|
| Non-prompt X(3872) | Inclusive              | Up                   | Central          | 0.0579            |
| Non-prompt X(3872) | Inclusive              | Up                   | +1 $\sigma$      | 0.0587            |
| Non-prompt X(3872) | Inclusive              | Up                   | -1 $\sigma$      | 0.0572            |
| Non-prompt X(3872) | Inclusive              | Down                 | Central          | 0.0588            |
| Non-prompt X(3872) | Inclusive              | Down                 | +1 $\sigma$      | 0.0595            |
| Non-prompt X(3872) | Inclusive              | Down                 | -1 $\sigma$      | 0.0580            |
| Non-prompt X(3872) | High Multiplicity      | Up                   | Central          | 0.0689            |
| Non-prompt X(3872) | High Multiplicity      | Up                   | +1 $\sigma$      | 0.0697            |
| Non-prompt X(3872) | High Multiplicity      | Up                   | -1 $\sigma$      | 0.0681            |
| Non-prompt X(3872) | High Multiplicity      | Down                 | Central          | 0.0697            |
| Non-prompt X(3872) | High Multiplicity      | Down                 | +1 $\sigma$      | 0.0705            |
| Non-prompt X(3872) | High Multiplicity      | Down                 | -1 $\sigma$      | 0.0688            |
| Non-prompt X(3872) | Low Multiplicity       | Up                   | Central          | 0.0570            |
| Non-prompt X(3872) | Low Multiplicity       | Up                   | +1 $\sigma$      | 0.0578            |
| Non-prompt X(3872) | Low Multiplicity       | Up                   | -1 $\sigma$      | 0.0563            |
| Non-prompt X(3872) | Low Multiplicity       | Down                 | Central          | 0.0574            |
| Non-prompt X(3872) | Low Multiplicity       | Down                 | +1 $\sigma$      | 0.0581            |
| Non-prompt X(3872) | Low Multiplicity       | Down                 | -1 $\sigma$      | 0.0566            |
|                    |                        |                      | $\mu \pm \sigma$ | $0.062 \pm 0.006$ |

Table 14: Variations of the non-prompt X(3872) reconstruction and selection efficiency, for different  $p_T$  spectrum parameterizations, magnetic field configurations, and MC/data corrections.

events that pass our cuts. The regions around the peaks have a sum of combinatorial background and signal events, so the nVeloTrack distributions in these regions are a sum of background events and the nVeloTracks distribution in signal events. The projections from these regions onto the y-axis give the nVeloTracks distributions, as shown in Fig. 46. There are some differences in the nVelotacks distributions projected form these three regions, presumably due to different mixtures of signal and background events.

The signal and background distributions of nVeloTracks can be separated by using the pure background distribution as a template for the background, and the signal to background ratio found by fitting the mass spectra in these regions. This can be represented conceptually by

$$nVeloTracks(signal+bkground) = N_{signal}[nVeloTracks(signal)] + (N_{bkground})[nVeloTracks(bkground)] \quad (15)$$

In this equation we know nVeloTracks(signal + background) from the red and blue curves shown in Fig. 46, nVeloTracks(background) from the black curve in Fig. 46, and the fraction of signal events from fitting the mas spectrum (as shown in the appendix). We can therefore solve for the distribution of VELO tracks in signal events nVeloTracks(signal), which is shown in Fig. 47. These distributions show good agreement within uncertainties, with mean values of  $66.5 \pm 1.8$  for X(3872) signal events and  $67.1 \pm 0.2$  for the  $\psi(2S)$  signal events. Since both X(3872) and  $\psi(2S)$  events display similar dependence on nVeloTracks, we conclude that there cannot be drastically different efficiency effects on the two states, however this conclusion is tempered by the uncertainties on the nVeloTrack

| Species | $p_T$ parameterization | Magnet Configuration | MC correction    | $\epsilon_{X(3872)}^{prompt}/\epsilon_{X(3872)}^{non-prompt}$ |
|---------|------------------------|----------------------|------------------|---|
| X(3872) | Inclusive              | Up                   | Central          | 1.165   |
| X(3872) | Inclusive              | Up                   | +1 $\sigma$      | 1.165   |
| X(3872) | Inclusive              | Up                   | -1 $\sigma$      | 1.166   |
| X(3872) | Inclusive              | Down                 | Central          | 1.144   |
| X(3872) | Inclusive              | Down                 | +1 $\sigma$      | 1.144   |
| X(3872) | Inclusive              | Down                 | -1 $\sigma$      | 1.145   |
| X(3872) | High Mult              | Up                   | Central          | 1.089   |
| X(3872) | High Mult              | Up                   | +1 $\sigma$      | 1.090   |
| X(3872) | High Mult              | Up                   | -1 $\sigma$      | 1.090   |
| X(3872) | High Mult              | Down                 | Central          | 1.083   |
| X(3872) | High Mult              | Down                 | +1 $\sigma$      | 1.084   |
| X(3872) | High Mult              | Down                 | -1 $\sigma$      | 1.084   |
| X(3872) | Low Mult               | Up                   | Central          | 1.084   |
| X(3872) | Low Mult               | Up                   | +1 $\sigma$      | 1.083   |
| X(3872) | Low Mult               | Up                   | -1 $\sigma$      | 1.085   |
| X(3872) | Low Mult               | Down                 | Central          | 1.068   |
| X(3872) | Low Mult               | Down                 | +1 $\sigma$      | 1.143   |
| X(3872) | Low Mult               | Down                 | -1 $\sigma$      | 1.069   |
|         |                        |                      | $\mu \pm \sigma$ | 1.11 $\pm 0.04$   |

Table 15: Variations of the ratio of reconstruction and selection efficiencies for non-prompt X(3872)/prompt X(3872), for different  $p_T$  spectrum parameterizations, magnetic field configurations, and MC/data corrections.

distributions extracted using this method. Therefore we assign a conservative systematic uncertainty of 3% (which is half the maximum absolute difference of 6% found above) on the relative correction factor  $\epsilon_{\psi(2S)}^{reco}/\epsilon_{X(3872)}^{reco}$ , to account for any differences in the nVeloTracks distributions that may induce difference reconstruction efficiencies for the two states.

Therefore, taking the mean value determined from Tab. 16 as the central value for  $\epsilon_{\psi(2S)}^{reco}/\epsilon_{X(3872)}^{reco}$ , the standard deviation from that tables as a systematic uncertainty, and an additional 3% uncertainty to account for any possible differences in the reconstruction due to differences in the track reconstruction efficiencies, we find the relative reconstruction efficiency for prompt production to be

$$\epsilon_{\psi(2S)}^{reco}/\epsilon_{X(3872)}^{reco} = 0.58 \pm 0.02 \quad (16)$$

We follow the same procedure for non-prompt production, using data from Tab. 17, and find the relative reconstruction efficiency for non-prompt production to be

$$\epsilon_{\psi(2S)}^{reco}/\epsilon_{X(3872)}^{reco} = 0.65 \pm 0.04 \quad (17)$$

We note that these values are consistent with the value of  $\frac{\epsilon_{\psi(2S)}}{\epsilon_{X(3872)}} \approx 0.58$  obtained from the simple considerations of decay kinematics discussed in Sec.4.2 above.

Similarly, the ratio of reconstruction efficiencies for prompt and non-prompt  $\psi(2S)$  production from Tabs. 12 is

| Species             | $p_T$ parameterization | Magnet Configuration | MC correction    | $\epsilon_{\psi(2S)}/\epsilon_{X(3872)}$ |
|---------------------|------------------------|----------------------|------------------|--|
| Prompt $\psi(2S)/X$ | Inclusive              | Up                   | Central          | 0.588                                    |
| Prompt $\psi(2S)/X$ | Inclusive              | Up                   | +1 $\sigma$      | 0.588                                    |
| Prompt $\psi(2S)/X$ | Inclusive              | Up                   | -1 $\sigma$      | 0.586                                    |
| Prompt $\psi(2S)/X$ | Inclusive              | Down                 | Central          | 0.582                                    |
| Prompt $\psi(2S)/X$ | Inclusive              | Down                 | +1 $\sigma$      | 0.583                                    |
| Prompt $\psi(2S)/X$ | Inclusive              | Down                 | -1 $\sigma$      | 0.581                                    |
| Prompt $\psi(2S)/X$ | High Mult              | Up                   | Central          | 0.567                                    |
| Prompt $\psi(2S)/X$ | High Mult              | Up                   | +1 $\sigma$      | 0.568                                    |
| Prompt $\psi(2S)/X$ | High Mult              | Up                   | -1 $\sigma$      | 0.566                                    |
| Prompt $\psi(2S)/X$ | High Mult              | Down                 | Central          | 0.556                                    |
| Prompt $\psi(2S)/X$ | High Mult              | Down                 | +1 $\sigma$      | 0.558                                    |
| Prompt $\psi(2S)/X$ | High Mult              | Down                 | -1 $\sigma$      | 0.556                                    |
| Prompt $\psi(2S)/X$ | Low Mult               | Up                   | Central          | 0.628                                    |
| Prompt $\psi(2S)/X$ | Low Mult               | Up                   | +1 $\sigma$      | 0.594                                    |
| Prompt $\psi(2S)/X$ | Low Mult               | Up                   | -1 $\sigma$      | 0.592                                    |
| Prompt $\psi(2S)/X$ | Low Mult               | Down                 | Central          | 0.623                                    |
| Prompt $\psi(2S)/X$ | Low Mult               | Down                 | +1 $\sigma$      | 0.553                                    |
| Prompt $\psi(2S)/X$ | Low Mult               | Down                 | -1 $\sigma$      | 0.590                                    |
|                     |                        |                      | $\mu \pm \sigma$ | $0.58 \pm 0.02$                          |

Table 16: Variations of the ratio of reconstruction and selection efficiencies for prompt  $\psi(2S)/X(3872)$ , for different  $p_T$  spectrum parameterizations, magnetic field configurations, and MC/data corrections.

$$\epsilon_{prompt\psi(2S)}^{reco}/\epsilon_{non-prompt\psi(2S)}^{reco} = 0.99 \pm 0.03 \quad (18)$$

and for X(3872) from Tab. 15

$$\epsilon_{promptX(3872)}^{reco}/\epsilon_{non-promptX(3872)}^{reco} = 1.11 \pm 0.04 \quad (19)$$

## 4.5 Particle Identification Efficiencies

The PIDCalib/Urania package is used to determine the efficiencies for identifying the pions and muons required to reconstruct the  $J/\psi\pi^+\pi^-$  mass spectra. For the 2012  $pp$  data, we select MC12TuneV2\_ProbNN prob variables. The PIDCalib efficiencies for muon and pion identification are shown in Figs. 48 and 49.

These efficiencies must be folded with the momentum spectra of the decay particles of interest in order to get a total efficiency for decays in our phase space of interest. Using the MC simulation samples, the momentum spectra for pions and muons from X(3872) and  $\psi(2S)$  decay that fall within our selected phase space and enter the LHCb spectrometer are prepared with the same binning as the PIDCalib plots. For illustration, a projection onto the  $P$  axis of daughters from X(3872) decays is shown in Fig. 50, note the low momentum cut-off at 10 GeV/c for muons and 3 GeV/c for pions: these are the result of our stripping and selection criteria. The integral of these spectra gives the number of generated pions or muons  $N_{generated}$  in the sample. These generated muon

| Species                 | $p_T$ parameterization | Magnet Config | MC correction    | $\epsilon_{\psi(2S)}/\epsilon_{X(3872)}$ |
|-------------------------|------------------------|---------------|------------------|--|
| Non-Prompt $\psi(2S)/X$ | Inclusive              | Up            | Central          | 0.701                                    |
| Non-Prompt $\psi(2S)/X$ | Inclusive              | Up            | +1 $\sigma$      | 0.700                                    |
| Non-Prompt $\psi(2S)/X$ | Inclusive              | Up            | -1 $\sigma$      | 0.699                                    |
| Non-Prompt $\psi(2S)/X$ | Inclusive              | Down          | Central          | 0.645                                    |
| Non-Prompt $\psi(2S)/X$ | Inclusive              | Down          | +1 $\sigma$      | 0.645                                    |
| Non-Prompt $\psi(2S)/X$ | Inclusive              | Down          | -1 $\sigma$      | 0.645                                    |
| Non-Prompt $\psi(2S)/X$ | High Mult              | Up            | Central          | 0.637                                    |
| Non-Prompt $\psi(2S)/X$ | High Mult              | Up            | +1 $\sigma$      | 0.638                                    |
| Non-Prompt $\psi(2S)/X$ | High Mult              | Up            | -1 $\sigma$      | 0.637                                    |
| Non-Prompt $\psi(2S)/X$ | High Mult              | Down          | Central          | 0.593                                    |
| Non-Prompt $\psi(2S)/X$ | High Mult              | Down          | +1 $\sigma$      | 0.593                                    |
| Non-Prompt $\psi(2S)/X$ | High Mult              | Down          | -1 $\sigma$      | 0.591                                    |
| Non-Prompt $\psi(2S)/X$ | Low Mult               | Up            | Central          | 0.698                                    |
| Non-Prompt $\psi(2S)/X$ | Low Mult               | Up            | +1 $\sigma$      | 0.699                                    |
| Non-Prompt $\psi(2S)/X$ | Low Mult               | Up            | -1 $\sigma$      | 0.650                                    |
| Non-Prompt $\psi(2S)/X$ | Low Mult               | Down          | Central          | 0.646                                    |
| Non-Prompt $\psi(2S)/X$ | Low Mult               | Down          | +1 $\sigma$      | 0.647                                    |
| Non-Prompt $\psi(2S)/X$ | Low Mult               | Down          | -1 $\sigma$      | 0.647                                    |
|                         |                        |               | $\mu \pm \sigma$ | $0.65 \pm 0.04$                          |

Table 17: Variations of the ratio of reconstruction and selection efficiencies for non-prompt  $\psi(2S)/X(3872)$ , for different  $p_T$  spectrum parameterizations, magnetic field configurations, and MC/data corrections.

and pion spectra are then multiplied by their respective efficiency histograms, and again integrated to obtain the number which pass the particle ID requirement  $N_{\text{passed}}$ . The ratio  $N_{\text{passed}}/N_{\text{generated}}$  then gives the particle ID efficiency. In this analysis, we are concerned with the ratio of the particle ID efficiencies between daughters from our decays of interest, which may be different due to the parent's decay kinematics. As an example, the ratio of the particle ID efficiencies for  $\mu^+$  is found by evaluating

$$\epsilon_{\psi(2S)}^{\mu^+ \text{PID}} / \epsilon_{X(3872)}^{\mu^+ \text{PID}} = \frac{N_{\text{passed}}^{\mu^+ \text{ from } \psi(2S)} / N_{\text{generated}}^{\mu^+ \text{ from } \psi(2S)}}{N_{\text{passed}}^{\mu^+ \text{ from } X(3872)} / N_{\text{generated}}^{\mu^+ \text{ from } X(3872)}}, \quad (20)$$

with similar expressions for  $\mu^-$ ,  $\pi^+$ , and  $\pi^-$ .

These efficiency ratios are evaluated in sets of 5000 toy experiments. In each of these toy experiments, each bin in the particle ID efficiency histogram assigned a value that is randomly sampled from a Gaussian distribution that is constructed by setting the mean to the value of the efficiency in that bin, and by setting the width to the value of the error bar. The ratio of particle ID efficiencies is calculated using the resulting efficiency sets, and the results are collected into histograms as shown in Fig. 51. This is done separately for  $\mu^+$ ,  $\mu^-$ ,  $\pi^+$ , and  $\pi^-$  using simulations and PIDCalib tables for both the Mag Up and Mag Down configurations.

Each of these histograms is fit with a Gaussian, the results of which are summarized in Tab. 18. The mean of each Gaussian is taken as the value of the particle ID efficiency ratio, which as expected is near unity. The width, which is quoted in the table as the uncertainty,

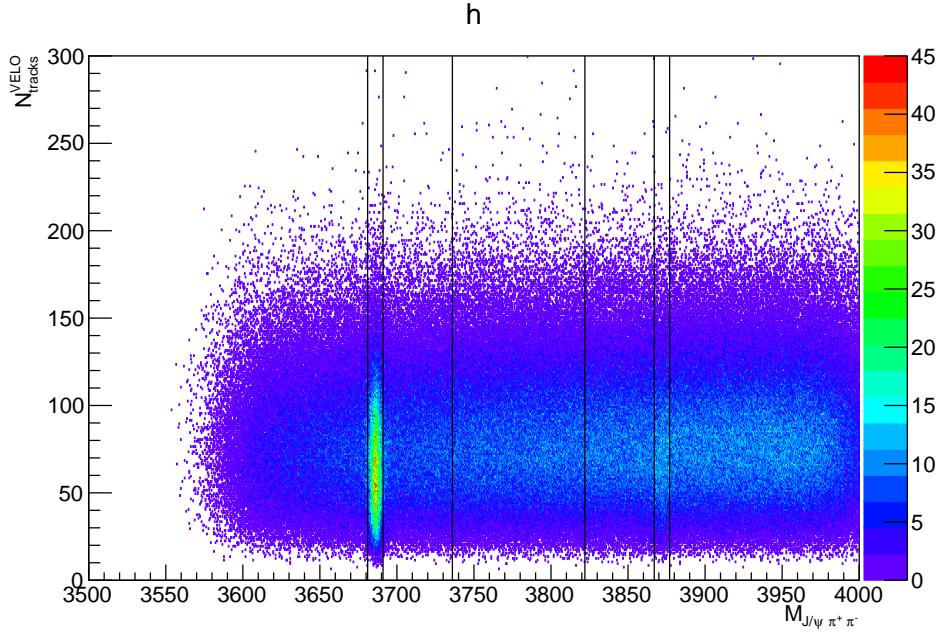


Figure 45: Mass distribution vs nVelotacks of candidates from the data.

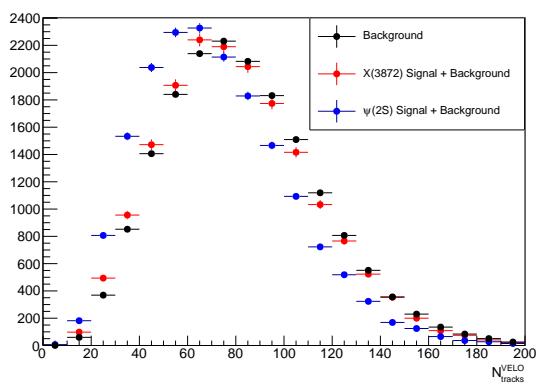


Figure 46: Raw nVelotacks distributions projected from selected areas shown in Fig. 45.

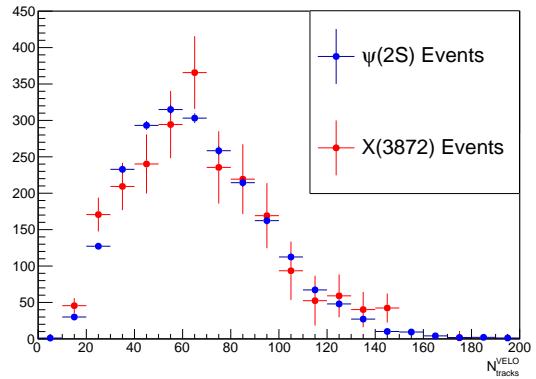


Figure 47: Background-subtracted nVelotacks distributions for events containing X(3872) and  $\psi(2S)$  signal events.

could be assigned as a systematic uncertainty due to the statistical uncertainty on the PIDCalib sample, however, we find that it is negligibly small (from 0.01 to 0.03%). We do assign a conservative systematic uncertainty of 1% on the particle ID efficiency ratio due to the difference between what is measured for positive and negative charges of the same daughter species. The values for positive and negative charges are averaged, and then the resulting values for Mag Up and Mag Down are averaged to give the final corrections:

$$\epsilon_{\psi(2S)}^{\mu^\pm PID} / \epsilon_{X(3872)}^{\mu^\pm PID} = 0.981 \pm 0.010 \quad (21)$$

and

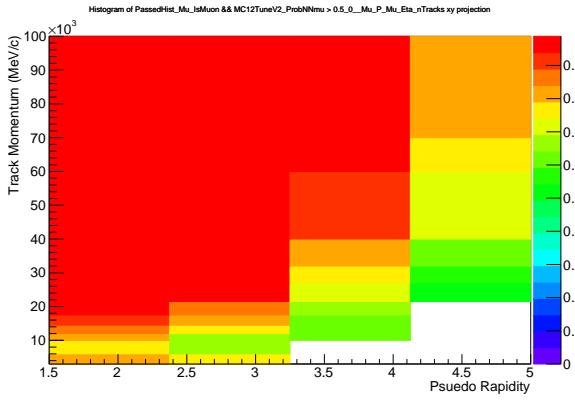


Figure 48: Muon ID efficiency.

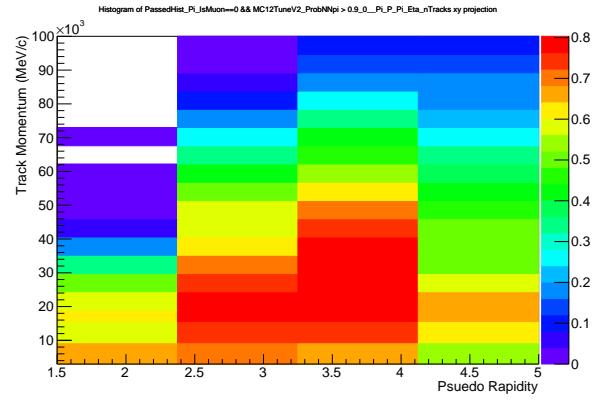


Figure 49: Pion ID efficiency.

### X(3872) daughters

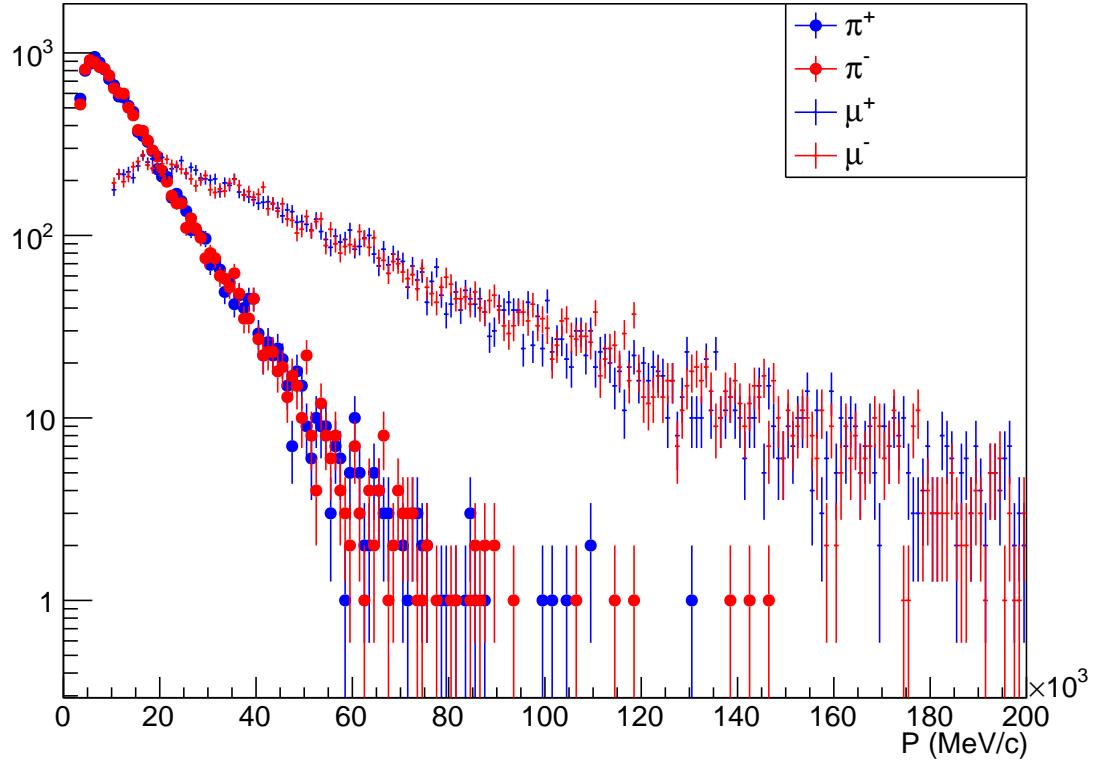


Figure 50: Momentum spectrum of daughters from  $X(3872) \rightarrow J/\psi \pi^+ \pi^-$  decays.

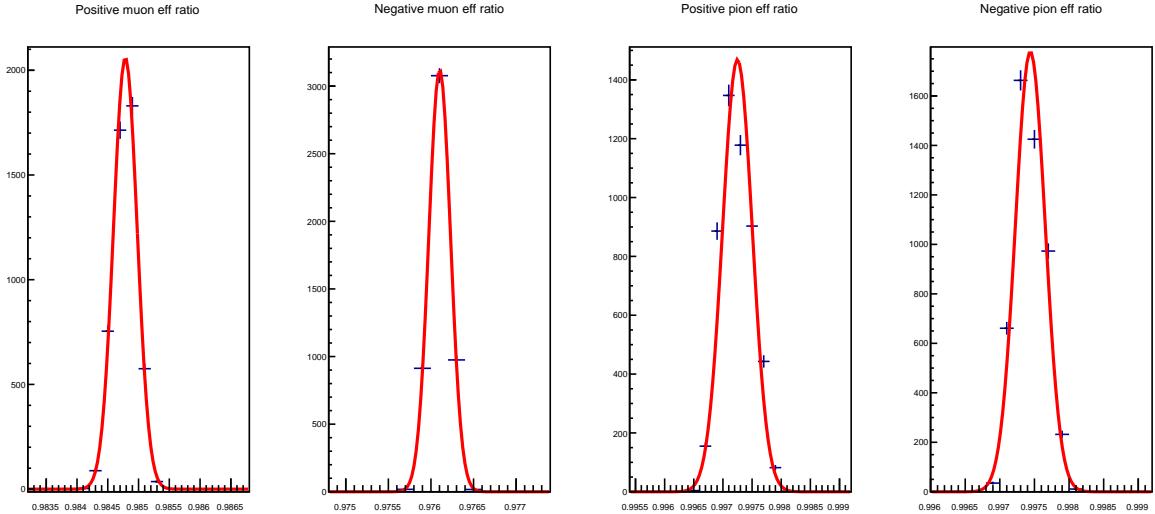


Figure 51: Distribution of ratio of particle ID efficiencies extracted from 5000 toy experiments. The width is assigned as a systematic uncertainty.

| Magnet Config | $\epsilon_{\psi(2S)}^{\mu^+ PID} / \epsilon_{X(3872)}^{\mu^+ PID}$ | $\epsilon_{\psi(2S)}^{\mu^- PID} / \epsilon_{X(3872)}^{\mu^- PID}$ | $\epsilon_{\psi(2S)}^{\pi^+ PID} / \epsilon_{X(3872)}^{\pi^+ PID}$ | $\epsilon_{\psi(2S)}^{\pi^- PID} / \epsilon_{X(3872)}^{\pi^- PID}$ |
|---------------|--|--|--|--|
| Mag Up        | $0.9842 \pm 0.0002$  | $0.9771 \pm 0.0001$  | $0.9985 \pm 0.0003$  | $0.9983 \pm 0.0003$  |
| Mag Down      | $0.9848 \pm 0.0002$  | $0.9761 \pm 0.0001$  | $0.9972 \pm 0.0003$  | $0.9974 \pm 0.0002$  |

Table 18: Summary of the ratio of particle ID efficiencies and their distributions.

$$\epsilon_{\psi(2S)}^{\pi^\pm PID} / \epsilon_{X(3872)}^{\pi^\pm PID} = 0.998 \pm 0.010 \quad (22)$$

## 699 5 Results

700 With the tools developed in the preceding sections of this Note, we can now determine  
 701 the physics quantities of interest: the fraction of prompt X(3872) and  $\psi(2S)$  and the ratio  
 702 of their cross sections, as a function of event multiplicity.

703 **5.1  $f_{prompt}$**

704 The fraction of promptly produced X(3872) and  $\psi(2S)$  hadrons is defined as

$$f_{prompt} = \frac{N_{prompt}}{N_{prompt} + N_{non-prompt}} \quad (23)$$

705 where  $N_{prompt}$  and  $N_{non-prompt}$  are the numbers of prompt and non-prompt  $\psi(2S)$   
 706 or X(3872) hadrons produced in a given bin of  $N_{tracks}^{VELO}$ , respectively. As discussed in  
 707 Section 3.2, two different fitting methods were used to determine the measured prompt  
 708 fraction  $f_p^{meas}$ . The values returned by the fit to the data are corrected for efficiencies for  
 709 measuring prompt and non-prompt hadrons and  $\epsilon_{prompt}$  and  $\epsilon_{non-prompt}$ :

$$f_{prompt} = \frac{N_{prompt}^{meas}/\epsilon_{prompt}}{N_{prompt}^{meas}/\epsilon_{prompt} + N_{non-prompt}^{meas}/\epsilon_{non-prompt}} \quad (24)$$

710 The number of prompt hadrons reconstructed is equivalent to the measured inclusive  
 711 number of hadrons  $N_{inc}$  times the prompt fraction extracted from the fit  $f_p^{meas}$ . Similarly,  
 712 the number of non-prompt hadrons is  $(1 - f_p^{meas}) \times N_{inc}$ . Thus the fully corrected physics  
 713 quantity  $f_{prompt}$  can be expressed as

$$f_{prompt} = \frac{N_{inc}f_p^{meas}/\epsilon_{prompt}}{N_{inc}f_p^{meas}/\epsilon_{prompt} + N_{inc}(1 - f_p^{meas})/\epsilon_{non-prompt}} \quad (25)$$

714 Here we see that  $N_{inc}$  cancels out, and we can multiply through by  $\epsilon_{prompt}/\epsilon_{prompt}$  to  
 715 our physics result:

$$f_{prompt} = \frac{f_p^{meas}}{f_p^{meas} + (1 - f_p^{meas})(\epsilon_{prompt}/\epsilon_{non-prompt})} \quad (26)$$

716 where:

- 717 •  $f_p^{meas}$  comes from the fit values shown in Figs. 14 and 15 for the  $\psi(2S)$  and X(3872),  
 718 respectively. The values of  $f_{prompt}$  determined by using the two different values of  
 719  $f_p^{meas}$  from the fits are averaged to give the central value. The difference between  
 720 this central value and the value determined by the individual methods is assigned  
 721 as a systematic uncertainty on  $f_{prompt}$ , which is uncorrelated between the bins.
- 722 •  $(\epsilon_{prompt}/\epsilon_{non-prompt})$  is the product of the solid angle acceptance corrections given in  
 723 Eqns. 12 and 13, and the selection/reconstruction efficiencies given in Eqns. 18 and  
 724 19. The value of  $(\epsilon_{prompt}/\epsilon_{non-prompt})$  is increased and decreased by its uncertainty,  
 725 and the resulting change in  $f_{prompt}$  is taken as a systematic uncertainty on  $f_{prompt}$ ,  
 726 which is correlated between the bins.

727 The resulting  $f_{prompt}$  values are shown in Fig. 52. The two sources of point-to-point  
 728 uncorrelated uncertainty from 1) the statistical uncertainty on  $f_p^{meas}$  and 2) the systematic  
 729 uncertainty from the two different fitting methods are added in quadrature and shown as  
 730 an error bar on each point. The correlated uncertainty on  $f_{prompt}$  due to the uncertainty  
 731 on  $\epsilon_{prompt}/\epsilon_{non-prompt}$  is shown as a box around the data points. The values for X(3872)  
 732 and  $\psi(2S)$  are given in the appendix in Tabs. 19 and 20, respectively.

733 We observe a clear decrease in the prompt fraction as the event activity increases.  
 734 There are two possible effects that could contribute to this decrease: first, collisions where  
 735  $b$  hadrons are produced will necessarily have a higher average number of produced charged  
 736 particles than those where  $b$  hadrons are not produced, simply because  $b$  hadrons will  
 737 decay into multiple charged tracks. The second effect is of more interest here, which could  
 738 involve suppression of promptly-produced X(3872) and  $\psi(2S)$  due to interactions with  
 739 the other particles in the event, which will occur with higher probability when the density  
 740 of produced particles is higher.

## 741 5.2 Cross Section Ratio

742 The ratio of cross sections for prompt  $\sigma_X(3872)/\sigma_{\psi(2S)}$  times their branching ratios  $\mathcal{B}$  to  
 743  $J/\psi \pi^+ \pi^-$  is found by evaluating

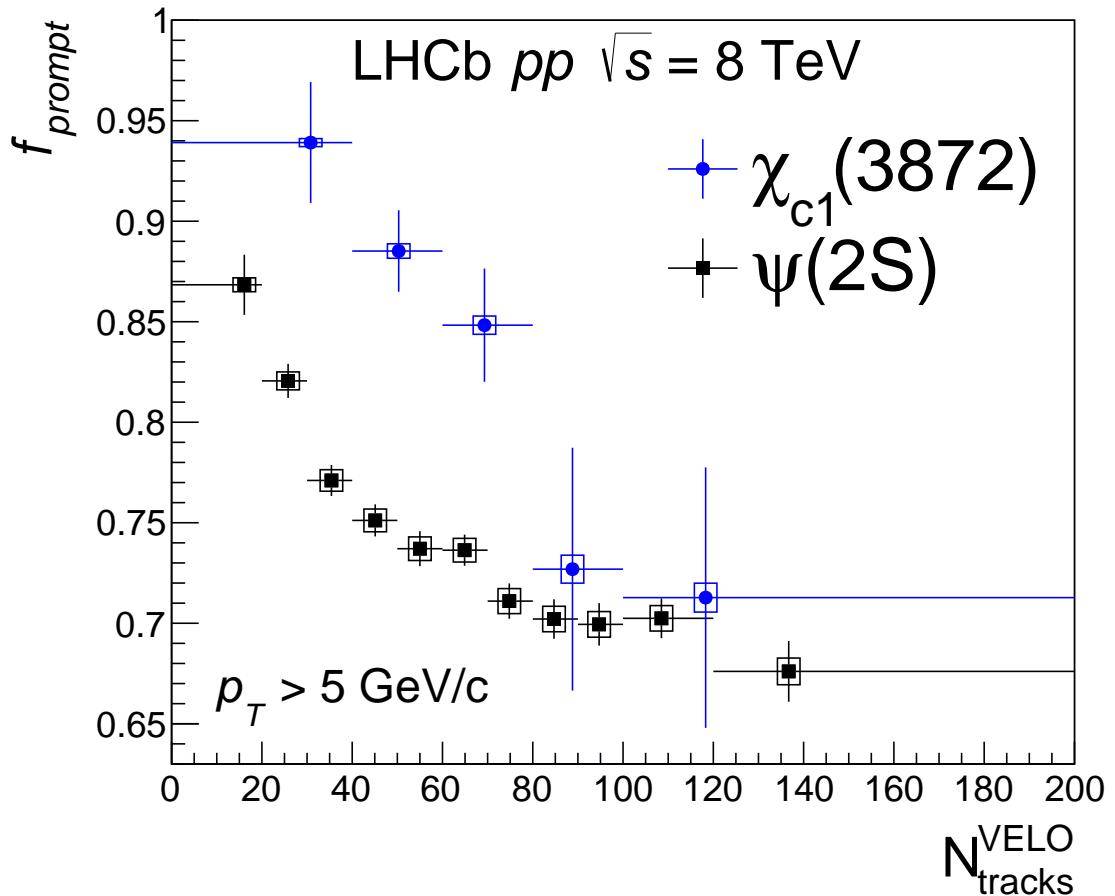


Figure 52: The fraction  $f_{prompt}$  of promptly-produced  $X(3872)$  and  $\psi(2S)$  hadrons, as a function of the number of tracks reconstructed in the VELO.

$$\frac{\mathcal{B}[X(3872) \rightarrow J/\psi \pi^+ \pi^-]}{\mathcal{B}[\psi(2S) \rightarrow J/\psi \pi^+ \pi^-]} \times \frac{\sigma_{X(3872)}}{\sigma_{\psi(2S)}} = \frac{N_{X(3872)}^{inclusive} f_{prompt}^{X(3872)}}{N_{\psi(2S)}^{inclusive} f_{prompt}^{\psi(2S)}} \times \frac{\epsilon_{\psi(2S)}^{acc}}{\epsilon_{X(3872)}^{acc}} \times \frac{\epsilon_{\psi(2S)}^{reco}}{\epsilon_{X(3872)}^{reco}} \times \left[ \frac{\epsilon_{\psi(2S)}^{\mu^\pm PID}}{\epsilon_{X(3872)}^{\mu^\pm PID}} \right]^2 \times \left[ \frac{\epsilon_{\psi(2S)}^{\pi^\pm PID}}{\epsilon_{X(3872)}^{\pi^\pm PID}} \right]^2 \quad (27)$$

744 where:

- 745 •  $N^{inclusive}$  is the yield of  $X(3872)$  or  $\psi(2S)$  determined in the simultaneous fit to the  
746  $J/\psi \pi^- \pi^+$  mass spectrum and proper time distribution, as discussed in Sec. 3.2
- 747 •  $f_{prompt}$  is the prompt fraction of  $X(3872)$  or  $\psi(2S)$  as shown in Fig. 52
- 748 •  $\frac{\epsilon_{\psi(2S)}^{acc}}{\epsilon_{X(3872)}^{acc}}$  is the ratio of the  $\psi(2S)$  and  $X(3872)$  solid angle acceptance of the LHCb  
749 spectrometer, determined by generator-level Monte Carlo simulation to be  $1.00 \pm 0.014$   
750 for prompt production and  $1.02 \pm 0.014$  for non-prompt production, as discussed in  
751 Sec. 4.3

- $\frac{\epsilon_{\psi(2S)}^{reco}}{\epsilon_{X(3872)}^{reco}}$  is the ratio of the stripping/reconstruction/selection efficiencies, determined by the full GEANT4 simulation of the LHCb detector to be  $0.58 \pm 0.03$  for prompt production and  $0.65 \pm 0.05$  for non-prompt production, as discussed in Sec. 4.4
- $\frac{\epsilon_{\psi(2S)}^{\pm PID}}{\epsilon_{X(3872)}^{\pm PID}}$  is the ratio of efficiencies for identifying muons from  $\psi(2S)$  and X(3872) decays, determined via PIDCalib to be  $0.981 \pm 0.01$  as discussed in Sec 4.5
- $\frac{\epsilon_{\psi(2S)}^{\pi^{\pm} PID}}{\epsilon_{X(3872)}^{\pi^{\pm} PID}}$  is the ratio of efficiencies for identifying pions from  $\psi(2S)$  and X(3872) decays, determined via PIDCalib to be  $0.998 \pm 0.01$  as discussed in Sec 4.5

The ratio for non-prompt cross sections is found using the same equation, but replacing  $f_{prompt}$  with  $(1 - f_{prompt})$ .

The ratio of the X(3872) and  $\psi(2S)$  cross sections measured in the  $J/\psi\pi^+\pi^-$  channel as a function of the number of tracks reconstructed in the VELO is shown in Fig. 53, for both the prompt and non-prompt selections. The uncertainties on the ratios of efficiencies are added in quadrature and taken as a systematic uncertainty on the ratio of cross sections, as shown by the boxes around the points. The point-to-point uncorrelated uncertainty on the ratio of cross sections due to the uncertainties on  $N^{inclusive}$  and the uncorrelated uncertainty on  $f_{prompt}$  are added in quadrature and shown as the error bar on the points. The correlated uncertainties due to the fully correlated uncertainty on  $f_{prompt}$  and the uncertainties on the efficiency corrections are shown as a box around the points. The horizontal axis values of each point are chosen are discussed in Sec. 3.3, and the horizontal error bars represent the bin width. The values for  $\psi(2S)$  and X(3872) are given in the appendix in Tabs. 21 and 22, respectively.

Here we observe a significant difference between the prompt and non-prompt cross section ratio. The prompt ratio decreases with increasing event activity, showing that X(3872) production is preferentially suppressed relative to prompt  $\psi(2S)$  production in higher-multiplicity collisions. Models that successfully reproduce the difference in  $\psi(2S)$  and  $J/\psi$  suppression in pPb collisions typically rely on interactions with co-moving hadrons as the breakup mechanism, which depends on the binding energy of the hadron. The fact that we observe suppression of X(3872) relative to  $\psi(2S)$  may suggest that X(3872) has an even smaller binding energy and/or larger radius than the  $\psi(2S)$ .

Crucially, we see that the non-prompt ratio of cross sections has no significant change as a function of event activity. Since  $b$  hadrons move some distance from the primary vertex before they decay, their daughter products will not interact with other hadrons. This flat behavior is therefore expected for  $b$  decays in vacuum, and this ratio of cross sections will be a constant that is set by  $b$  decay branching ratios. We note that the values obtained here are in good agreement with the values for the same ratio of non-prompt cross sections reported by ATLAS in [43]. Although the ATLAS measurements cover a higher range of  $p_T$ , this ratio is not expected to show large variations with  $p_T$  as it is set by branching ratios of  $b$  decays. The fact that this ratio is flat also shows that the decrease we observe in the prompt cross section is not merely some detector effect.

As the simplest possible check of significance, we fit a straight line to both the prompt and non-prompt ratios, as shown in Fig. 54, considering only the uncorrelated vertical errors shown by the error bars on the points. The fit to the non-prompt ratio gives a slope of  $3 \pm 1.9 \times 10^{-4}/N_{tracks}^{VELO}$ , which is consistent with zero within  $1.6\sigma$ , with a  $\chi^2/NDF$

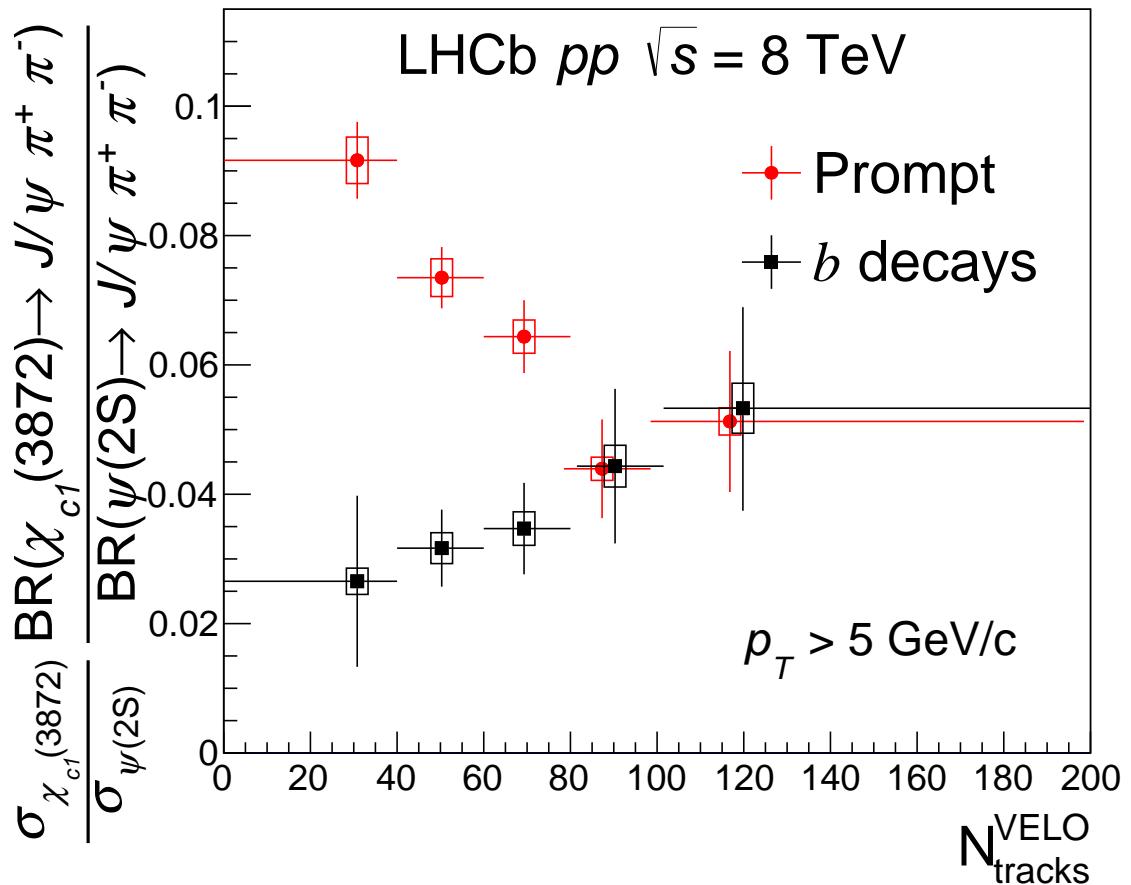


Figure 53: The ratio of the  $X(3872)$  and  $\psi(2S)$  cross sections measured in the  $J/\psi\pi^+\pi^-$  channel as a function of the number of tracks reconstructed in the VELO. The two points at the highest values of  $N_{tracks}^{VELO}$  are slightly offset for clarity.

<sup>795</sup> = 0.14/3 = 0.05. The fit to the prompt ratio returns a slope of  $-5.84 \pm 1.16 \times 10^{-4}/N_{tracks}^{VELO}$ ,  
<sup>796</sup> which is  $5.0\sigma$  away from zero, with a  $\chi^2/NDF = 4.24/3 = 1.4$ . However, we note there is  
<sup>797</sup> no physical reason to expect a simple linear dependence on the multiplicity.

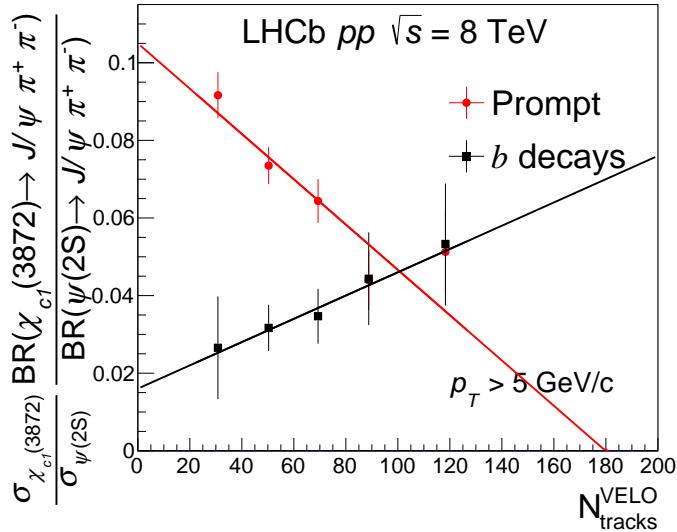


Figure 54: Simple linear fit to the ratios of cross sections.

## 798 References

- 799 [1] S. L. Olsen, T. Skwarnicki, and D. Zieminska, *Nonstandard heavy mesons and baryons: 800 Experimental evidence*, Rev. Mod. Phys. **90** (2018) 015003.
- 801 [2] Belle Collaboration, S.-K. Choi *et al.*, *Observation of a Narrow Charmoniumlike 802 State in Exclusive  $B^\pm \rightarrow K^\pm \pi^+ \pi^- J/\psi$  Decays*, Phys. Rev. Lett. **91** (2003) 262001.
- 803 [3] CDF II Collaboration, D. Acosta *et al.*, *Observation of the narrow state  $x(3872) \rightarrow 804 j/\psi \pi^+ \pi^-$  in  $\bar{p}p$  collisions at  $\sqrt{s} = 1.96 \text{ TeV}$* , Phys. Rev. Lett. **93** (2004) 072001.
- 805 [4] D0 Collaboration, V. M. Abazov *et al.*, *Observation and properties of the  $x(3872)$  806 decaying to  $j/\psi \pi^+ \pi^-$  in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96 \text{ TeV}$* , Phys. Rev. Lett. **93** (2004) 807 162002.
- 808 [5] BABAR Collaboration, B. Aubert *et al.*, *Study of the  $B^- \rightarrow j/\psi K^- \pi^+ \pi^-$  decay and 809 measurement of the  $B^- \rightarrow x(3872) K^-$  branching fraction*, Phys. Rev. D **71** (2005) 810 071103.
- 811 [6] LHCb, R. Aaij *et al.*, *Observation of  $X(3872)$  production in  $pp$  collisions at  $\sqrt{s} = 7 \text{ TeV}$* , Eur. Phys. J. **C72** (2012) 1972, arXiv:1112.5310.
- 813 [7] E. J. Eichten, K. Lane, and C. Quigg, *Charmonium levels near threshold and the 814 narrow state  $x(3872) \rightarrow \pi^+ \pi^- j/\psi$* , Phys. Rev. D **69** (2004) 094019.
- 815 [8] LHCb, R. Aaij *et al.*, *Determination of the  $X(3872)$  meson quantum numbers*, Phys. 816 Rev. Lett. **110** (2013) 222001, arXiv:1302.6269.
- 817 [9] T. Barnes, S. Godfrey, and E. S. Swanson, *Higher charmonia*, Phys. Rev. D **72** (2005) 818 054026.

- [10] L. Maiani, F. Piccinini, A. D. Polosa, and V. Riquer, *Diquark-antidiquark states with hidden or open charm and the nature of  $x(3872)$* , Phys. Rev. D **71** (2005) 014028.
- [11] R. D. Matheus, S. Narison, M. Nielsen, and J.-M. Richard, *Can the  $x(3872)$  be a  $1^{++}$  four-quark state?*, Phys. Rev. D **75** (2007) 014005.
- [12] S. Dubnicka, A. Z. Dubnickova, M. A. Ivanov, and J. G. Körner, *Quark model description of the tetraquark state  $x(3872)$  in a relativistic constituent quark model with infrared confinement*, Phys. Rev. D **81** (2010) 114007.
- [13] P. C. Wallbott, G. Eichmann, and C. S. Fischer,  *$x(3872)$  as a four-quark state in a dyson-schwinger/bethe-salpeter approach*, Phys. Rev. D **100** (2019) 014033.
- [14] R. D. Matheus, F. S. Navarra, M. Nielsen, and C. M. Zanetti, *Qcd sum rules for the  $x(3872)$  as a mixed molecule-charmonium state*, Phys. Rev. D **80** (2009) 056002.
- [15] M. Butenschoen, Z.-G. He, and B. A. Kniehl, *Deciphering the  $x(3872)$  via its polarization in prompt production at the cern lhc*, Phys. Rev. Lett. **123** (2019) 032001.
- [16] S. Coito, G. Rupp, and E. van Beveren,  *$X(3872)$  is not a true molecule*, The European Physical Journal C **73** (2013) 2351.
- [17] N. A. Tornqvist, *Isospin breaking of the narrow charmonium state of Belle at 3872-MeV as a deuson*, Phys. Lett. **B590** (2004) 209, [arXiv:hep-ph/0402237](#).
- [18] E. Braaten and M. Kusunoki, *Exclusive production of the  $x(3872)$  in  $b$  meson decay*, Phys. Rev. D **71** (2005) 074005.
- [19] M. Cardoso, G. Rupp, and E. van Beveren, *Unquenched quark-model calculation of  $\varvec{x}(3872)$  electromagnetic decays*, The European Physical Journal C **75** (2015) 20.
- [20] B. Alessandro et al.,  *$J/\psi$  and  $\psi'$  production and their normal nuclear absorption in proton-nucleus collisions at 400 gev*, The European Physical Journal C - Particles and Fields **48** (2006) 329.
- [21] A. Adare et al., *Phys. Rev. Lett.* **111** (2013) 202301.
- [22] B. Abelev, *Journal of High Energy Physics* **2014** (2014) 1.
- [23] LHCb, R. Aaij et al., *Study of  $\psi(2S)$  production and cold nuclear matter effects in pPb collisions at  $\sqrt{s_{NN}} = 5$  TeV*, *JHEP* **03** (2016) 133.
- [24] PHENIX, A. Adare et al., *Measurement of the relative yields of  $\psi(2S)$  to  $\psi(1S)$  mesons produced at forward and backward rapidity in  $p+p$ ,  $p+Al$ ,  $p+Au$ , and  ${}^3He+Au$  collisions at  $\sqrt{s_{NN}} = 200$  GeV*, *Phys. Rev. C* **95** (2017) 034904, [arXiv:1609.06550](#).
- [25] CMS, A. M. Sirunyan et al., *Measurement of prompt  $\psi(2S)$  production cross sections in proton-lead and proton-proton collisions at  $\sqrt{s_{NN}} = 5.02$  TeV*, *Phys. Lett. B* **790** (2019) 509, [arXiv:1805.02248](#).
- [26] CMS, S. Chatrchyan et al., *Event activity dependence of  $Y(nS)$  production in  $\sqrt{s_{NN}}=5.02$  TeV pPb and  $\sqrt{s}=2.76$  TeV pp collisions*, *JHEP* **04** (2014) 103, [arXiv:1312.6300](#).

- 854 [27] *LHCb*, R. Aaij et al., Study of  $\Upsilon$  production in  $p\text{Pb}$  collisions at  $\sqrt{s_{NN}} = 8.16 \text{ TeV}$ ,  
 855 *JHEP* **11** (2018) 194, [arXiv:1810.07655](#).
- 856 [28] A. Capella, A. Kaidalov, A. K. Akil, and C. Gerschel *Physics Letters B* **393** (1997)  
 857 431 .
- 858 [29] E. G. Ferreiro *Physics Letters B* **749** (2015) 98 .
- 859 [30] Y.-Q. Ma, R. Venugopalan, K. Watanabe, and H.-F. Zhang,  $\psi(2s)$  versus  $j/\psi$  suppression  
 860 in proton-nucleus collisions from factorization violating soft color exchanges,  
 861 *Phys. Rev. C* **97** (2018) 014909.
- 862 [31] E. G. Ferreiro and J. P. Lansberg, Is bottomonium suppression in proton-nucleus  
 863 and nucleus-nucleus collisions at lhc energies due to the same effects?, *Journal of  
 864 High Energy Physics* **2018** (2018) 94.
- 865 [32] E. Braaten, H.-W. Hammer, and T. Mehen, Scattering of an Ultrasoft Pion and the  
 866 X(3872), *Phys. Rev.* **D82** (2010) 034018, [arXiv:1005.1688](#).
- 867 [33] ALICE, J. Adam et al., Enhanced production of multi-strange hadrons  
 868 in high-multiplicity proton-proton collisions, *Nature Phys.* **13** (2017) 535,  
 869 [arXiv:1606.07424](#).
- 870 [34] CMS, V. Khachatryan et al., Evidence for collectivity in pp collisions at the LHC,  
 871 *Phys. Lett. B* **765** (2017) 193, [arXiv:1606.06198](#).
- 872 [35] A. Capella, A. Kaidalov, A. Kouider Akil, and C. Gerschel,  $J/\psi$  and  $\psi'$  suppression  
 873 in heavy ion collisions, *Phys. Lett. B* **393** (1997) 431, [arXiv:hep-ph/9607265](#).
- 874 [36] E. G. Ferreiro, Excited charmonium suppression in protonnucleus collisions as a  
 875 consequence of comovers, *Phys. Lett. B* **749** (2015) 98, [arXiv:1411.0549](#).
- 876 [37] E. Braaten, H.-W. Hammer, and T. Mehen, Scattering of an Ultrasoft Pion and the  
 877 X(3872), *Phys. Rev. D* **82** (2010) 034018, [arXiv:1005.1688](#).
- 878 [38] L. M. Abreu et al., X(3872) production and absorption in a hot hadron gas, *Phys.  
 879 Lett. B* **761** (2016) 303, [arXiv:1604.07716](#).
- 880 [39] H. Satz, Colour deconfinement and quarkonium binding, *J. Phys. G* **32** (2006) R25,  
 881 [arXiv:hep-ph/0512217](#).
- 882 [40] T. L. collaboration, Precision luminosity measurements at LHCb, *Journal of Instrumentation* **9** (2014) P12005.
- 884 [41] LHCb collaboration, R. Aaij et al., Observation of  $X(3872)$  production in  $pp$  collisions  
 885 at  $\sqrt{s} = 7 \text{ TeV}$ , *Eur. Phys. J. C* **72** (2012) 1972, [arXiv:1112.5310](#).
- 886 [42] CDF, A. Abulencia et al., Measurement of the dipion mass spectrum in  $X(3872) \rightarrow$   
 887  $J/\psi\pi^+\pi^-$  decays., *Phys. Rev. Lett.* **96** (2006) 102002, [arXiv:hep-ex/0512074](#).
- 888 [43] ATLAS, M. Aaboud et al., Measurements of  $\psi(2S)$  and  $X(3872) \rightarrow J/\psi\pi^+\pi^-$   
 889 production in  $pp$  collisions at  $\sqrt{s} = 8 \text{ TeV}$  with the ATLAS detector, *JHEP* **01**  
 890 (2017) 117, [arXiv:1610.09303](#).

- 891 [44] CMS, S. Chatrchyan et al., Measurement of the  $X(3872)$  Production Cross Section  
 892 Via Decays to  $J/\psi\pi^+\pi^-$  in  $pp$  collisions at  $\sqrt{s} = 7$  TeV, *JHEP* **04** (2013) 154,  
 893 arXiv:1302.3968.
- 894 [45] B. Gui and T. Skwarnicki, Determination of  $X(3872)$  quantum numbers, *LHCb-  
 895 ANA-2012-104*.
- 896 [46] LHCb collaboration, R. Aaij et al., Measurement of  $\psi(2S)$  meson production in  $pp$   
 897 collisions at  $\sqrt{s} = 7$  TeV, *Eur. Phys. J.* **C72** (2012) 2100, arXiv:1204.1258.
- 898 [47] R. Aaij et al., Performance of the LHCb Vertex Locator, *JINST* **9** (2014) P09007,  
 899 arXiv:1405.7808.
- 900 [48] Particle Data Group, M. Tanabashi et al., Review of Particle Physics, *Phys. Rev.  
 901 D98* (2018) 030001.
- 902 [49] LHCb, R. Aaij et al., Prompt and nonprompt  $J/\psi$  production and nuclear mod-  
 903 ification in  $p\text{Pb}$  collisions at  $\sqrt{s_{\text{NN}}} = 8.16$  TeV, *Phys. Lett.* **B774** (2017) 159,  
 904 arXiv:1706.07122.
- 905 [50] LHCb, R. Aaij et al., Measurement of  $\psi(2S)$  production cross-sections in proton-  
 906 proton collisions at  $\sqrt{s} = 7$  and 13 TeV, *Submitted to: Eur. Phys. J.* (2019)  
 907 arXiv:1908.03099.
- 908 [51] M. Kucharczyk, P. Morawski, and M. Witek, Primary Vertex Reconstruction at  
 909 LHCb, *Tech. Rep. LHCb-PUB-2014-044. CERN-LHCb-PUB-2014-044*, CERN,  
 910 Geneva, 2014.
- 911 [52] M. Pivk and F. R. Le Diberder, SPlot: A Statistical tool to unfold data distributions,  
 912 *Nucl. Instrum. Meth.* **A555** (2005) 356, arXiv:physics/0402083.
- 913 [53] LHCb, R. Aaij et al., Measurement of the track reconstruction efficiency at LHCb,  
 914 *JINST* **10** (2015) P02007, arXiv:1408.1251.

915 **Appendices**

916 **A Data Tables**

| nVeloTracks | $f_{prompt}$ | uncorrelated uncertainty | correlated uncertainty |
|-------------|--------------|--------------------------|------------------------|
| 30.8        | 0.939106     | 0.0301034                | 0.00200062             |
| 50.3        | 0.885164     | 0.0202548                | 0.00359083             |
| 69.3        | 0.848286     | 0.0281662                | 0.00454127             |
| 88.8        | 0.726909     | 0.0604044                | 0.00695124             |
| 118.3       | 0.712743     | 0.0647922                | 0.00718835             |

Table 19: The fraction  $f_{prompt}$  for X(3872), as shown in Fig. 52.

| nVeloTracks | $f_{prompt}$ | uncorrelated uncertainty | correlated uncertainty |
|-------------|--------------|--------------------------|------------------------|
| 16.1        | 0.868376     | 0.0149599                | 0.00346327             |
| 25.8        | 0.820567     | 0.00847784               | 0.00446178             |
| 35.4        | 0.771039     | 0.00772964               | 0.00534955             |
| 45.1        | 0.751159     | 0.00796617               | 0.0056639              |
| 55          | 0.737084     | 0.0087096                | 0.00587177             |
| 64.9        | 0.736346     | 0.00774819               | 0.00588297             |
| 74.8        | 0.711015     | 0.00878657               | 0.00622635             |
| 84.7        | 0.702108     | 0.00982191               | 0.00633782             |
| 94.7        | 0.699442     | 0.0105605                | 0.00637064             |
| 108.5       | 0.70245      | 0.00988378               | 0.00633421             |
| 136.7       | 0.67605      | 0.0151009                | 0.0066357              |

Table 20: The fraction  $f_{prompt}$  for  $\psi(2S)$ , as shown in Fig. 52.

| nVeloTracks | X/ $\psi(2S)$ | uncorrelated uncertainty | correlated uncertainty |
|-------------|---------------|--------------------------|------------------------|
| 30.8        | 0.0265459     | 0.0132296                | 0.00203169             |
| 50.3        | 0.0316677     | 0.00595951               | 0.00238706             |
| 69.3        | 0.0346934     | 0.00706946               | 0.00259032             |
| 88.8        | 0.0443463     | 0.0119505                | 0.00322719             |
| 118.3       | 0.0533022     | 0.0158538                | 0.00386698             |

Table 21: The ratio of cross section for non-prompt X(3872)/ $\psi(2S)$ , as shown in Fig. 53.

| nVeloTracks | X/ $\psi(2S)$ | uncorrelated uncertainty | correlated uncertainty |
|-------------|---------------|--------------------------|------------------------|
| 30.8        | 0.0916358     | 0.00594127               | 0.00359448             |
| 50.3        | 0.0734887     | 0.00473419               | 0.00291431             |
| 69.3        | 0.0643733     | 0.00563091               | 0.00257019             |
| 88.8        | 0.0439398     | 0.00763238               | 0.00179524             |
| 118.3       | 0.0512599     | 0.0109104                | 0.00210371             |

Table 22: The ratio of cross section for prompt X(3872)/ $\psi(2S)$ , as shown in Fig. 53.

## 917 B Fits to Mass and Proper Time

918 Here we show all the simultaneous fits to the proper time and mass spectra for completeness.

### 919 B.1 $\psi(2S)$ with Fit as Proper Time Background

| nVeloTracks bin center | $f_{prompt}$              | Inclusive N $\psi(2S)$ |
|------------------------|---------------------------|------------------------|
| 10                     | $0.870894 \pm 0.0146211$  | $695.373 \pm 29.4444$  |
| 25                     | $0.817512 \pm 0.00839727$ | $2851.81 \pm 61.387$   |
| 35                     | $0.76589 \pm 0.00700425$  | $5068.46 \pm 83.7202$  |
| 45                     | $0.744649 \pm 0.00653348$ | $6381.31 \pm 96.6811$  |
| 55                     | $0.729206 \pm 0.00645427$ | $7045.9 \pm 104.157$   |
| 65                     | $0.730472 \pm 0.00673112$ | $6688.96 \pm 104.554$  |
| 75                     | $0.70458 \pm 0.00768902$  | $5715.64 \pm 99.7098$  |
| 85                     | $0.695353 \pm 0.00874046$ | $4622.91 \pm 91.7485$  |
| 95                     | $0.700369 \pm 0.0101457$  | $3547.68 \pm 82.0708$  |
| 110                    | $0.69887 \pm 0.00988083$  | $4034.55 \pm 92.465$   |
| 160                    | $0.68088 \pm 0.0135079$   | $2480.45 \pm 77.4954$  |

Table 23: Values returned from the simultaneous fit to the  $\psi(2S)$  mass and proper time spectra, using the fit function as background.

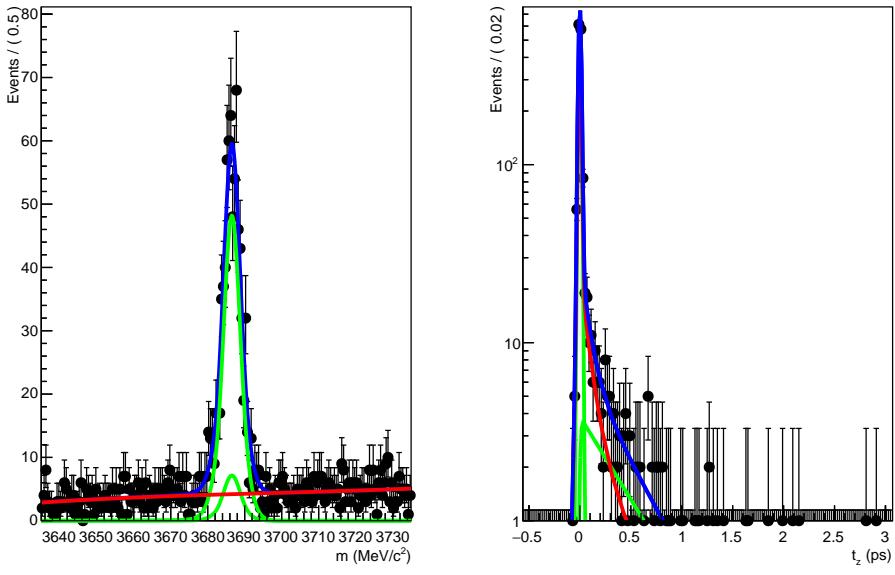


Figure 55: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $0 < N_{tracks}^{VELO} < 20$ .

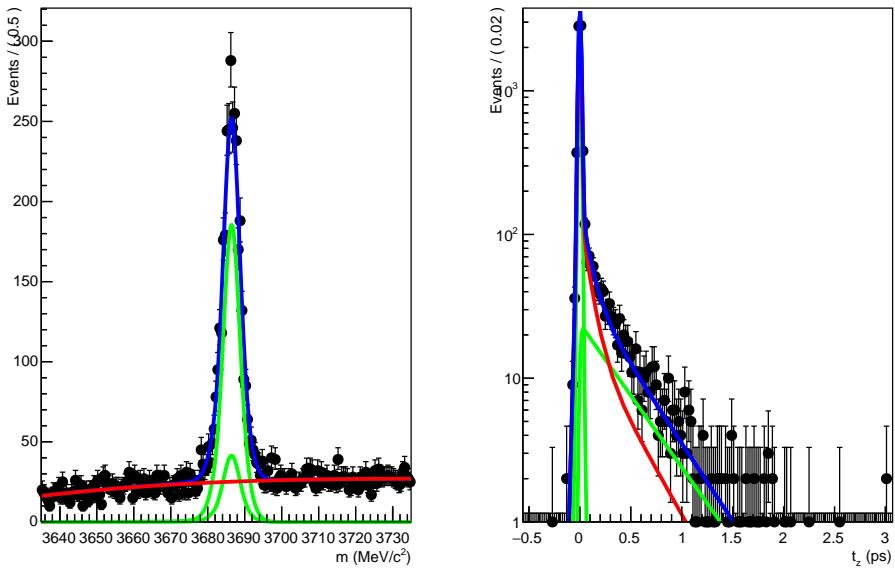


Figure 56: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $20 < N_{tracks}^{VELO} < 30$ .

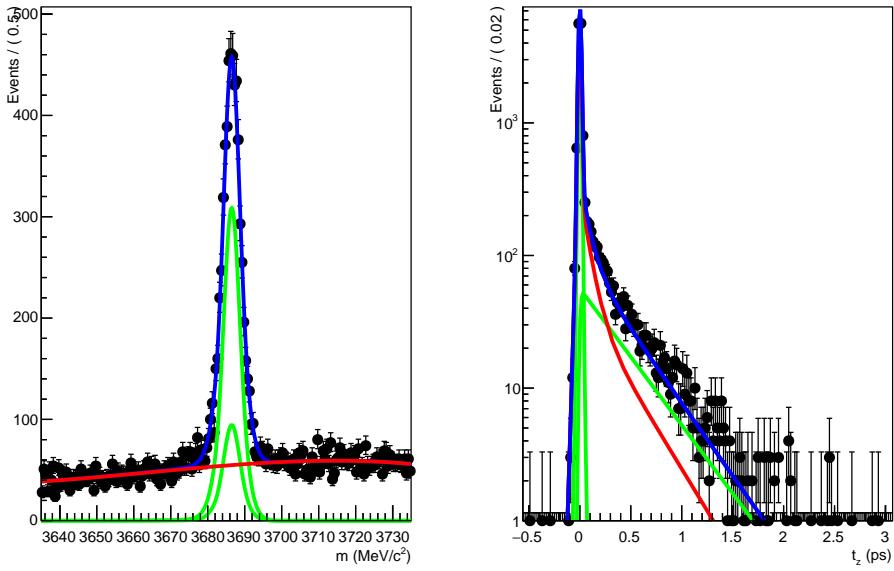


Figure 57: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $30 < N_{tracks}^{VELO} < 40$ .

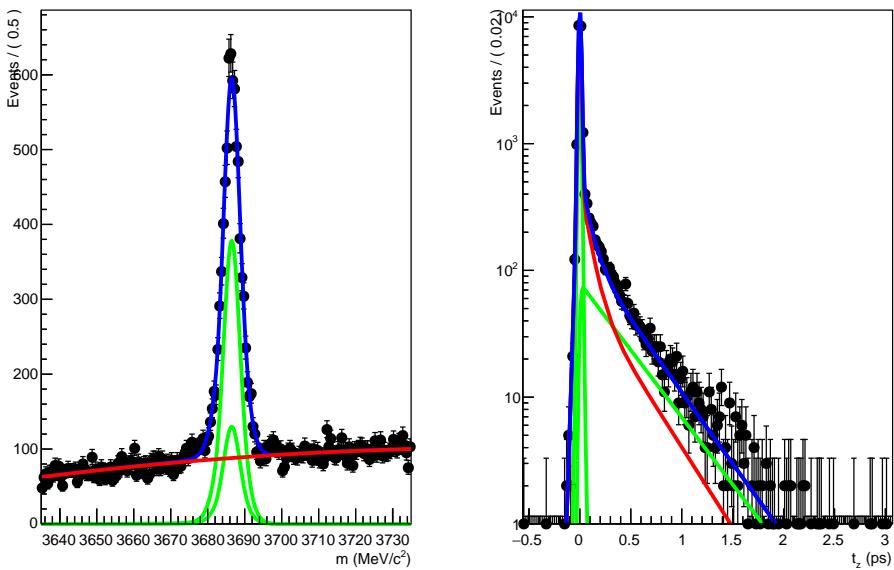


Figure 58: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $40 < N_{tracks}^{VELO} < 50$ .

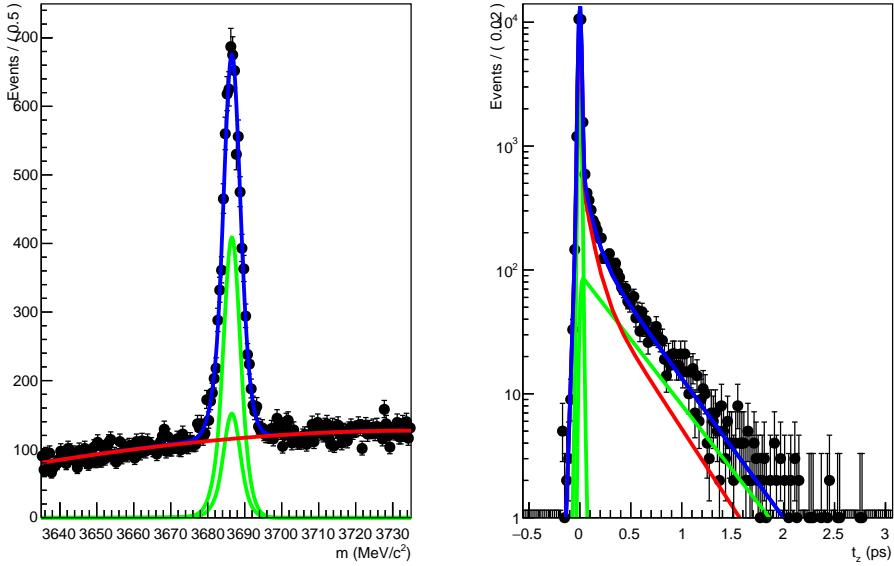


Figure 59: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $50 < N_{tracks}^{VELO} < 60$ .

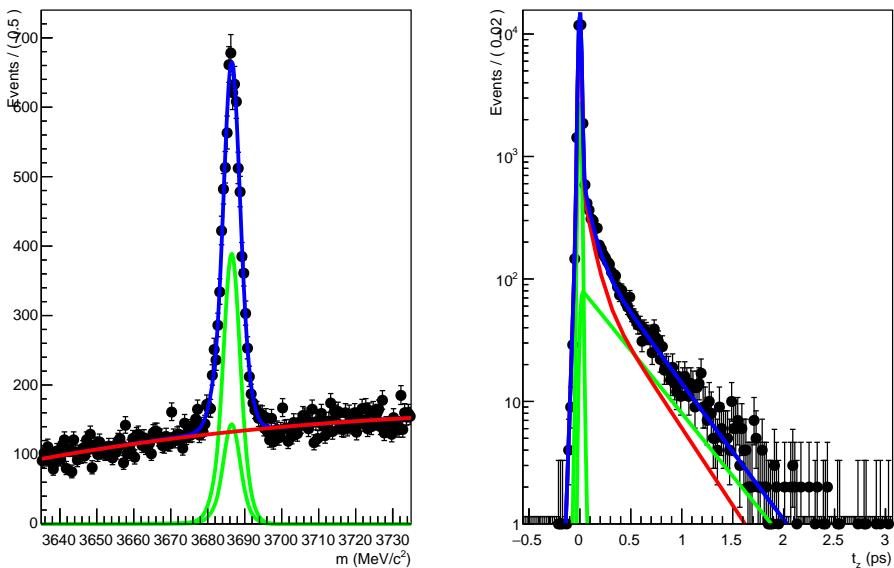


Figure 60: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $60 < N_{tracks}^{VELO} < 70$ .

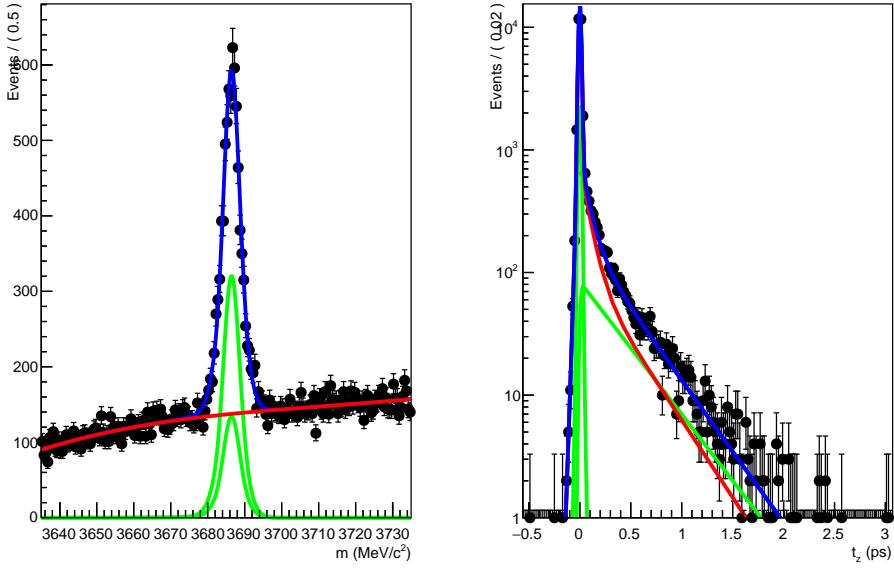


Figure 61: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $70 < N_{tracks}^{VELO} < 80$ .

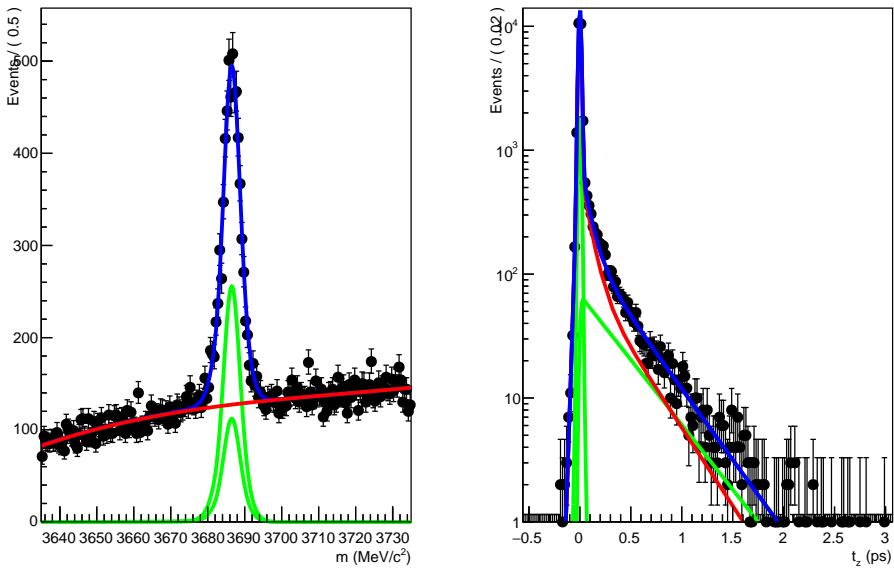


Figure 62: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $80 < N_{tracks}^{VELO} < 90$ .

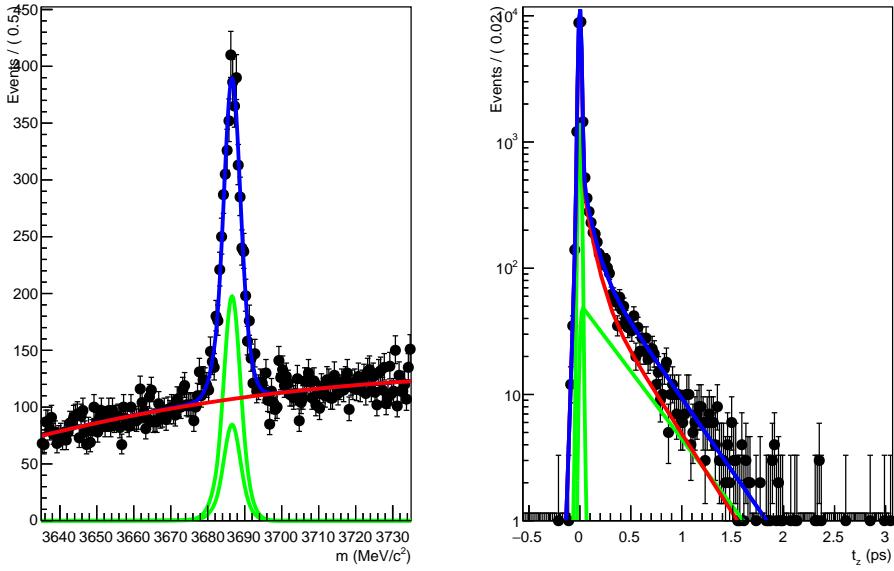


Figure 63: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $90 < N_{tracks}^{VELO} < 100$ .

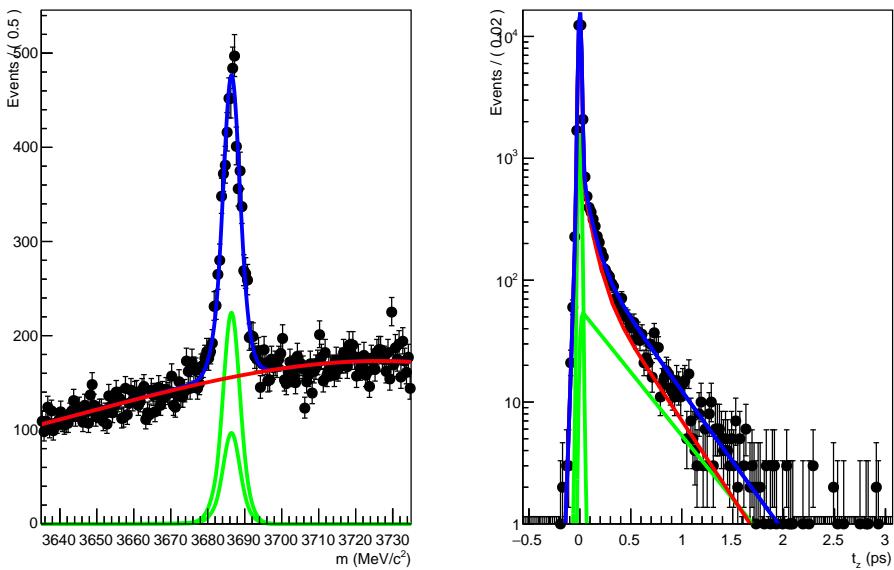


Figure 64: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $100 < N_{tracks}^{VELO} < 120$ .

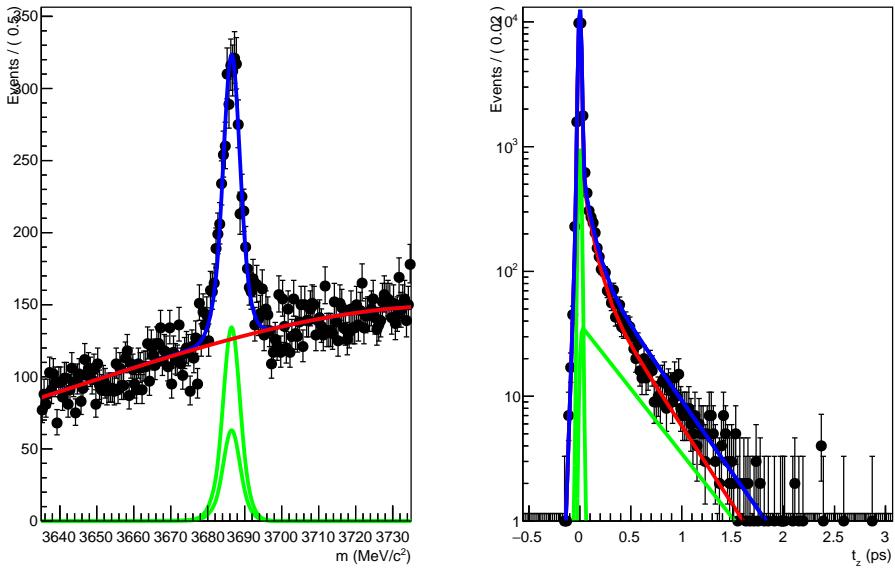


Figure 65: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $120 < N_{tracks}^{VELO} < 200$ .

920    **B.2     $\psi(2S)$  with Sideband Histogram as Proper Time Back-**  
 921    **ground**

| nVeloTracks bin center | $f_{prompt}$              | Inclusive N $\psi(2S)$ |
|------------------------|---------------------------|------------------------|
| 10                     | $0.863552 \pm 0.0146129$  | $709.642 \pm 29.5947$  |
| 25                     | $0.820653 \pm 0.00837621$ | $2872.63 \pm 61.8012$  |
| 35                     | $0.77263 \pm 0.00700187$  | $5071.17 \pm 84.1887$  |
| 45                     | $0.753903 \pm 0.0065344$  | $6364.18 \pm 97.3129$  |
| 55                     | $0.741059 \pm 0.00642318$ | $7043.67 \pm 104.96$   |
| 65                     | $0.738308 \pm 0.00672316$ | $6679.86 \pm 105.401$  |
| 75                     | $0.713312 \pm 0.00764674$ | $5725.9 \pm 100.553$   |
| 85                     | $0.704651 \pm 0.00865373$ | $4648.42 \pm 92.5711$  |
| 95                     | $0.694282 \pm 0.0101672$  | $3560.67 \pm 82.7733$  |
| 110                    | $0.70182 \pm 0.00974645$  | $4079.8 \pm 93.2139$   |
| 160                    | $0.666812 \pm 0.0133375$  | $2528.75 \pm 77.8071$  |

Table 24: Values returned from the simultaneous fit to the  $\psi(2S)$  mass and proper time spectra, using the mass sideband histogram as background.

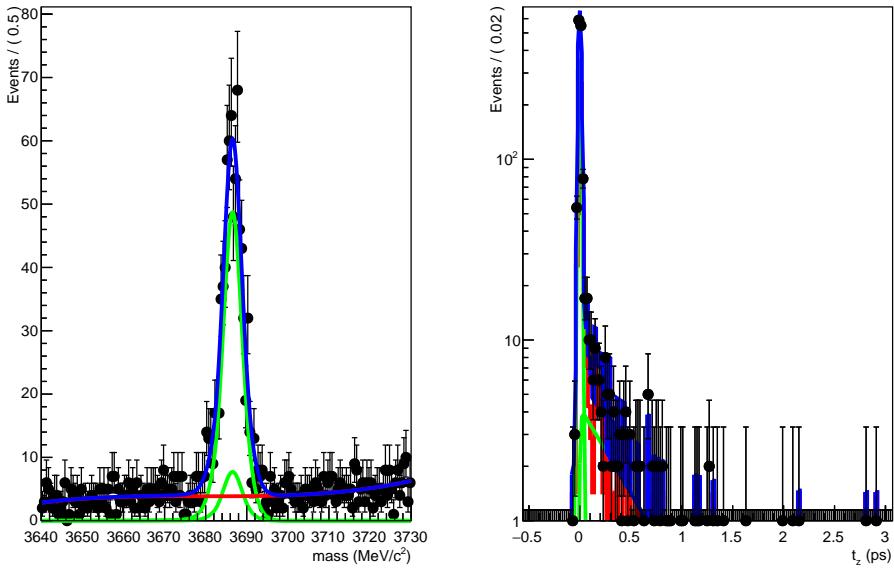


Figure 66: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $0 < N_{tracks}^{VELO} < 20$ .

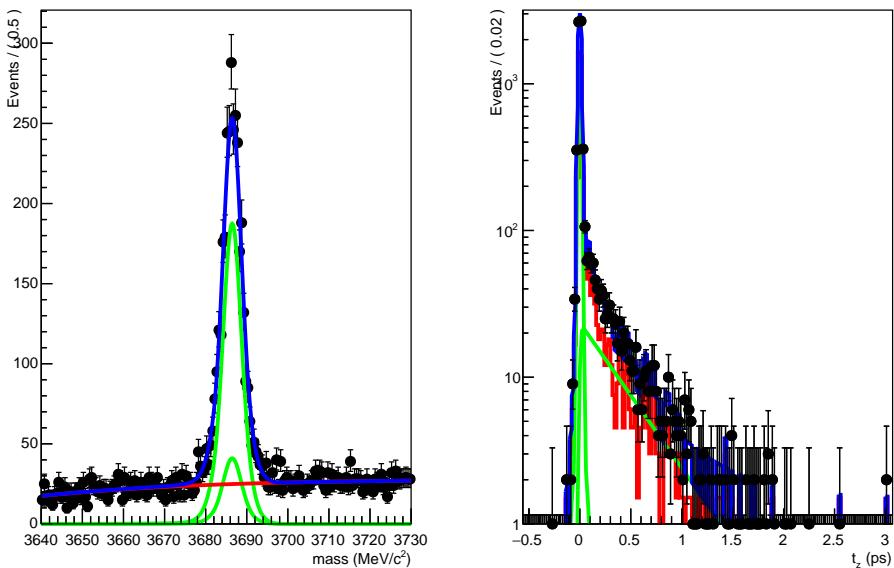


Figure 67: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $20 < N_{tracks}^{VELO} < 30$ .

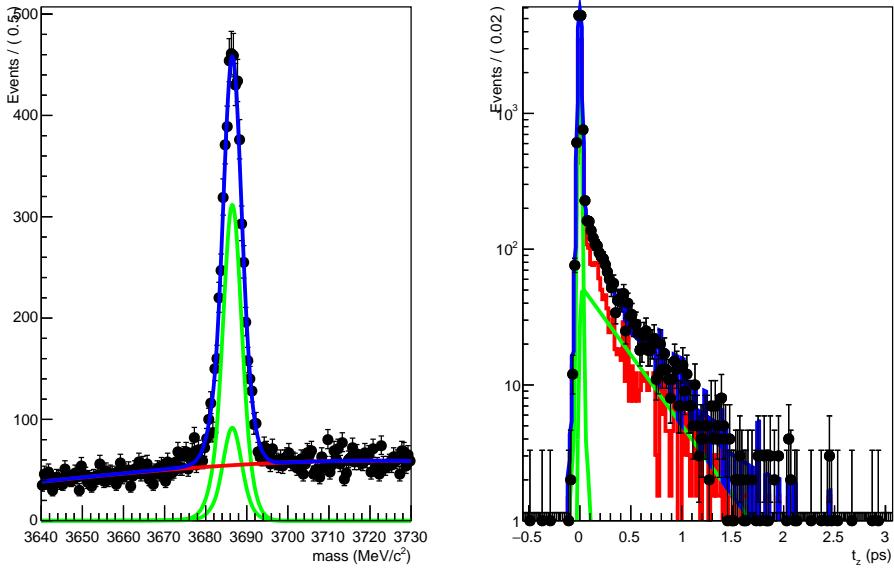


Figure 68: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $30 < N_{\text{tracks}}^{\text{VELO}} < 40$ .

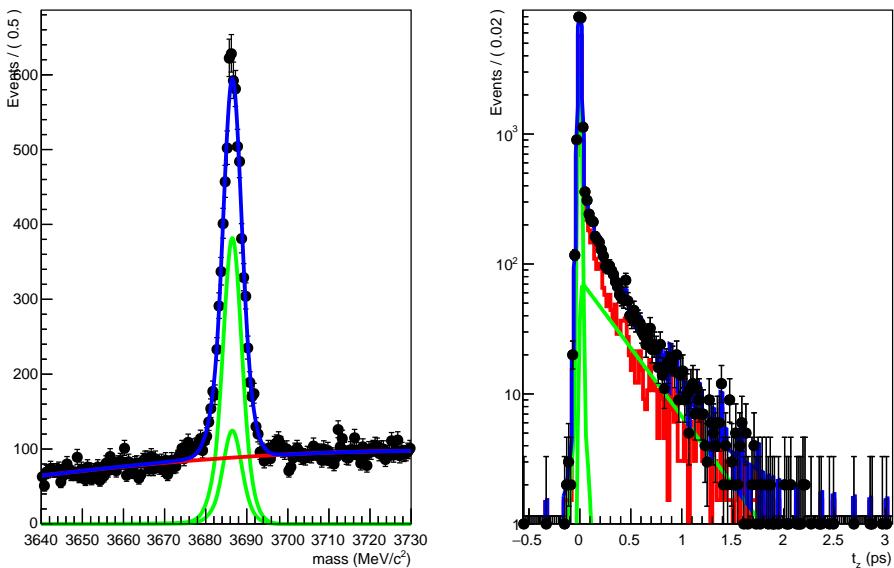


Figure 69: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $40 < N_{\text{tracks}}^{\text{VELO}} < 50$ .

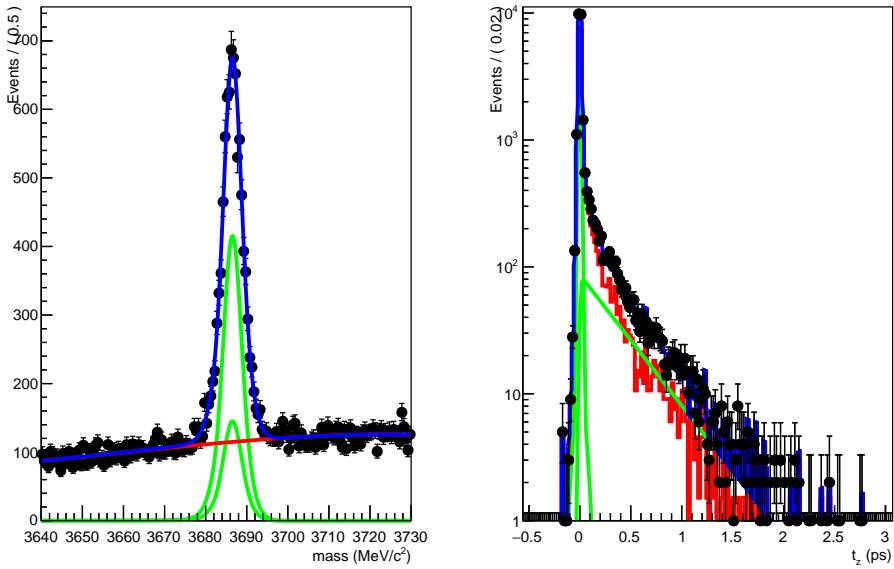


Figure 70: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $50 < N_{tracks}^{VELO} < 60$ .

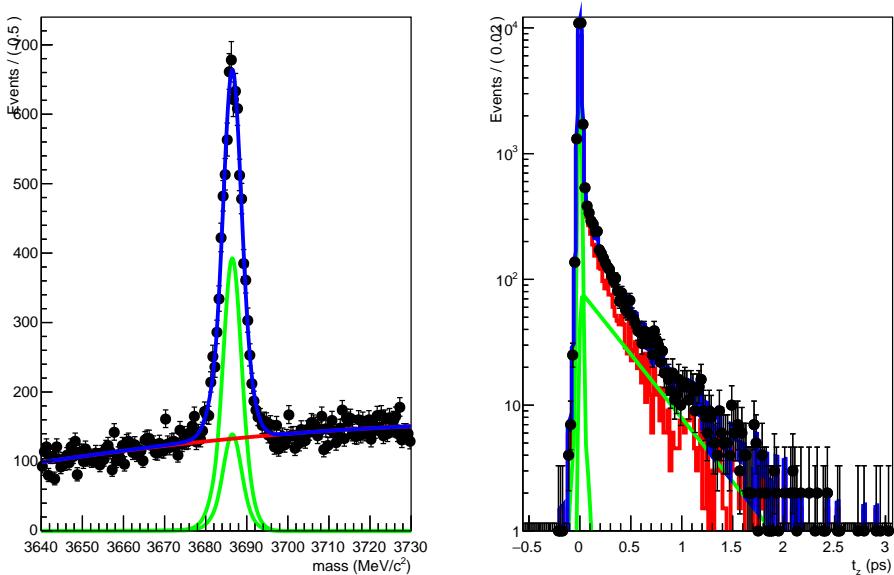


Figure 71: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $60 < N_{tracks}^{VELO} < 70$ .

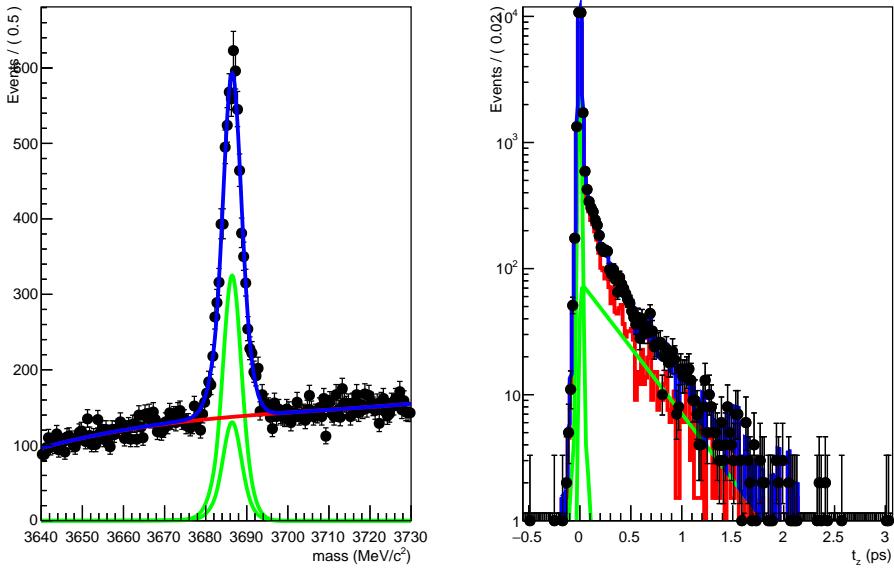


Figure 72: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $70 < N_{tracks}^{VELO} < 80$ .

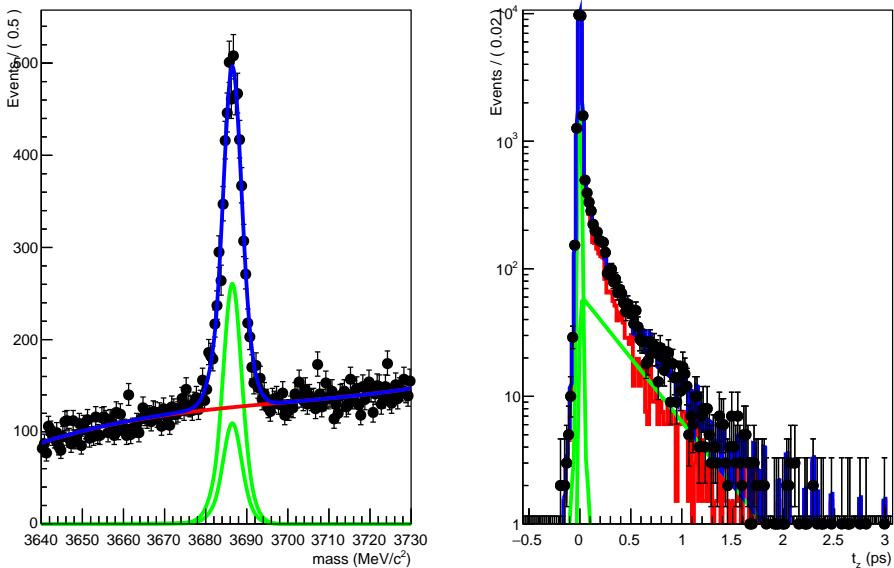


Figure 73: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $80 < N_{tracks}^{VELO} < 90$ .

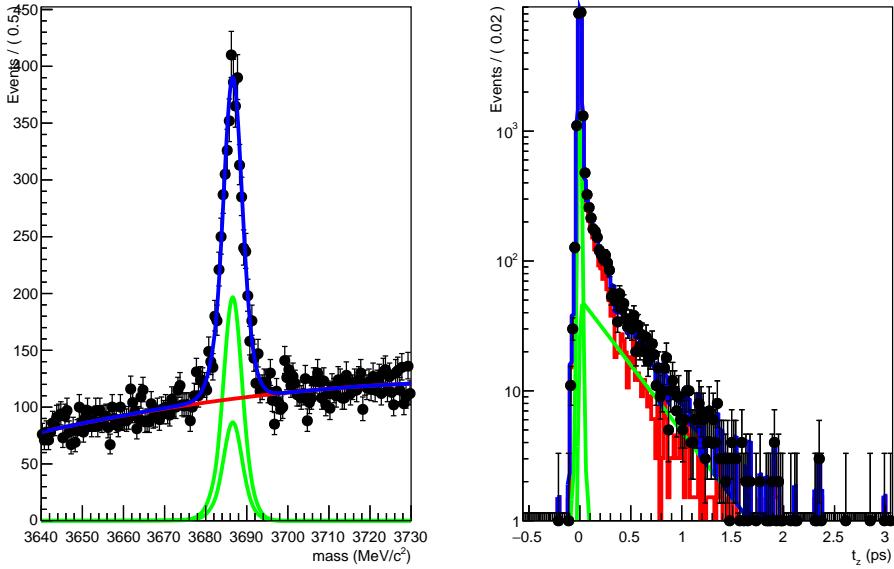


Figure 74: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $90 < N_{\text{tracks}}^{\text{VELO}} < 100$ .

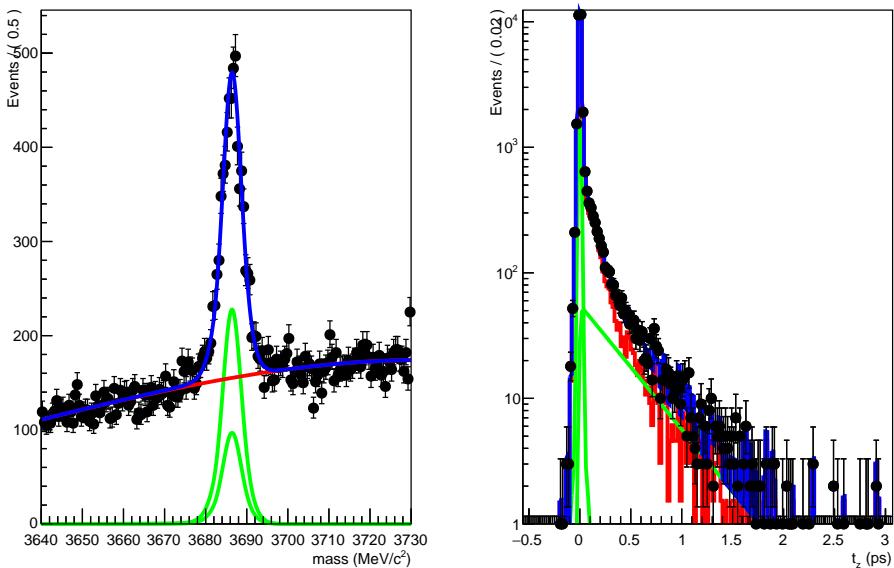


Figure 75: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $100 < N_{\text{tracks}}^{\text{VELO}} < 120$ .

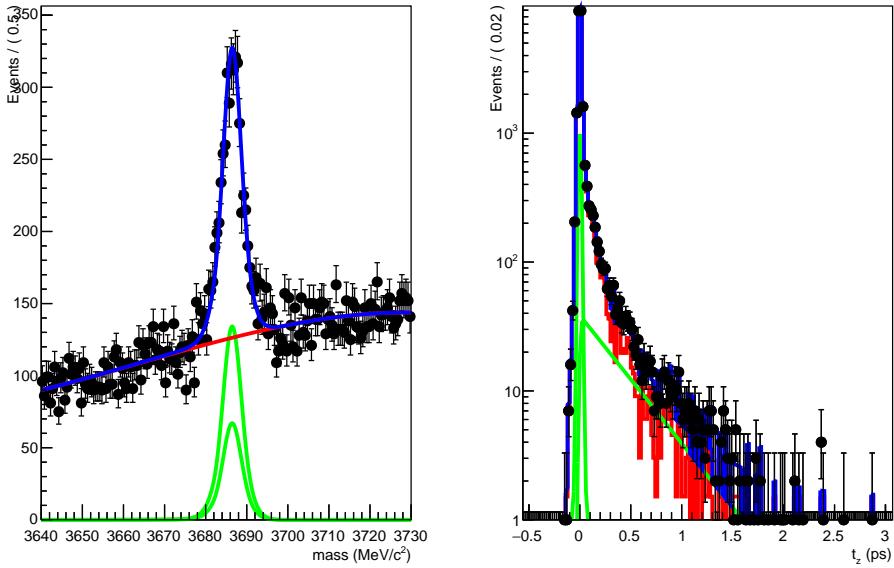


Figure 76: Simultaneous fit to the mass and proper time spectra for  $\psi(2S)$ , in the event activity range  $120 < N_{tracks}^{VELO} < 200$ .

<sub>922</sub> **B.3 X(3872) with Fit as Proper Time Background**

| nVeloTracks bin center | $f_{prompt}$             | Inclusive N <sub>X</sub> |
|------------------------|--------------------------|--------------------------|
| 20                     | $0.966951 \pm 0.0197666$ | $1202.49 \pm 69.8081$    |
| 50                     | $0.900378 \pm 0.0191107$ | $1501.04 \pm 84.904$     |
| 70                     | $0.875151 \pm 0.0225717$ | $1239.79 \pm 94.0938$    |
| 90                     | $0.790374 \pm 0.0433606$ | $626.898 \pm 94.8135$    |
| 150                    | $0.771371 \pm 0.059429$  | $590.995 \pm 120.247$    |

Table 25: Values returned from the simultaneous fit to the X(3872) mass and proper time spectra, using the fit as background.

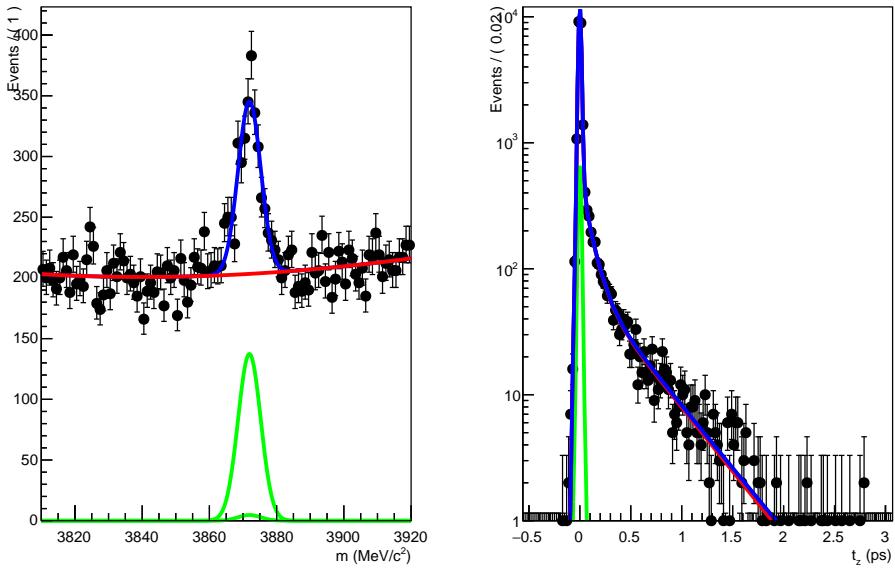


Figure 77: Simultaneous fit to the mass and proper time spectra for  $X(3872)$ , in the event activity range  $0 < N_{\text{tracks}}^{\text{VELO}} < 40$ .

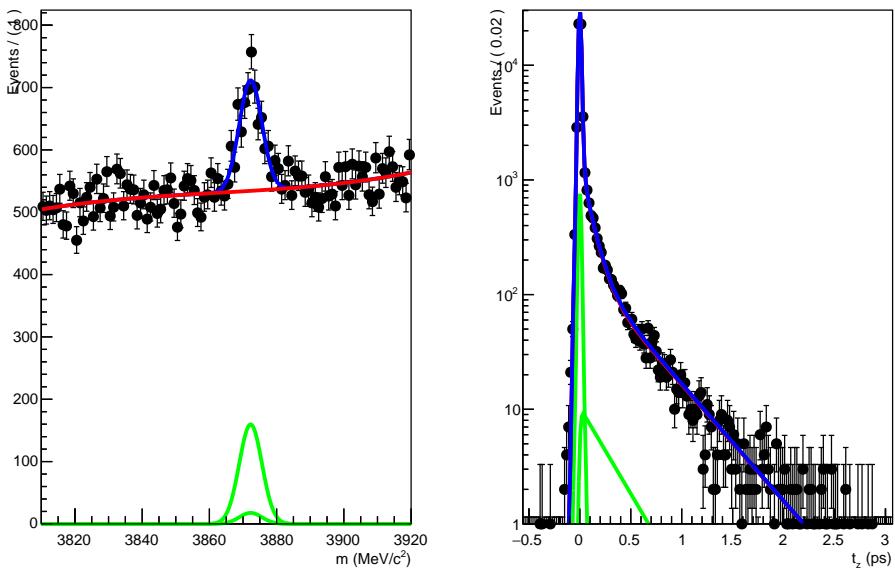


Figure 78: Simultaneous fit to the mass and proper time spectra for  $X(3872)$ , in the event activity range  $40 < N_{\text{tracks}}^{\text{VELO}} < 60$ .

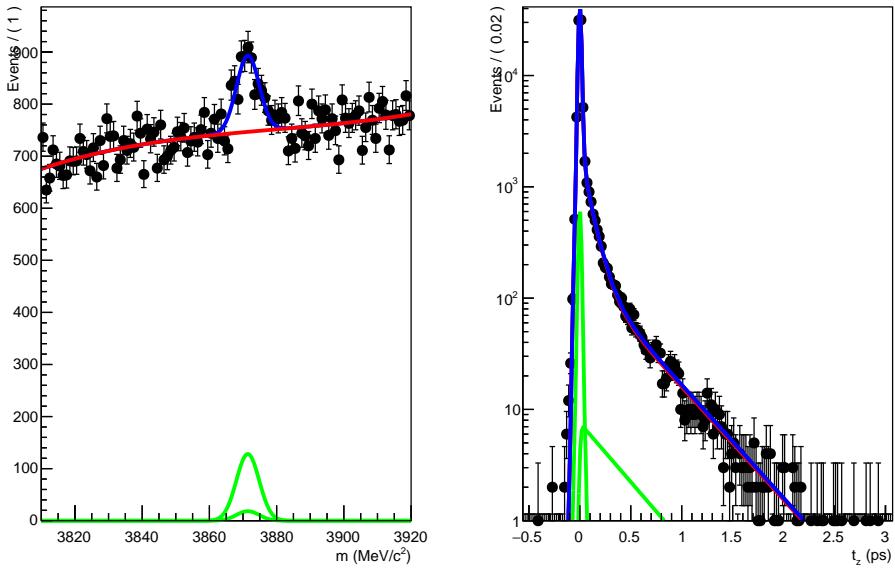


Figure 79: Simultaneous fit to the mass and proper time spectra for  $X(3872)$ , in the event activity range  $60 < N_{tracks}^{VELO} < 80$ .

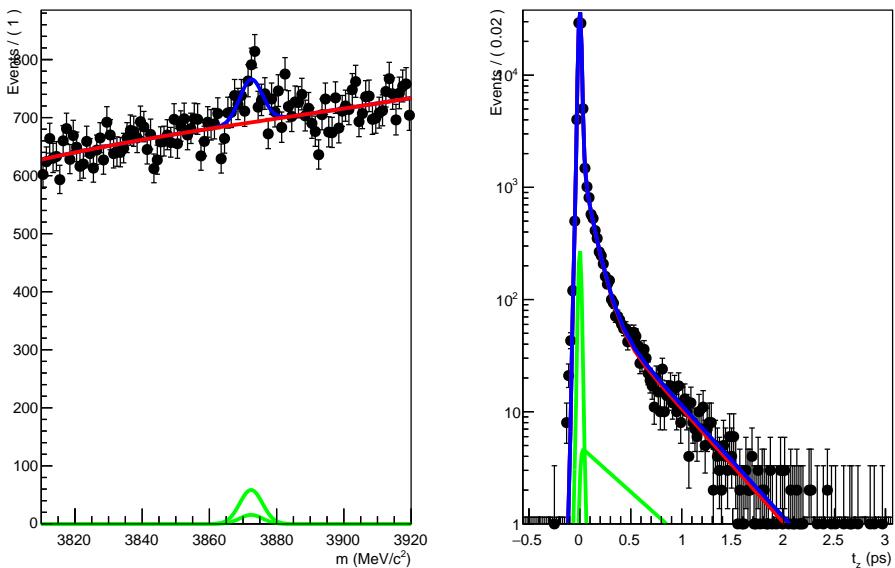


Figure 80: Simultaneous fit to the mass and proper time spectra for  $X(3872)$ , in the event activity range  $80 < N_{tracks}^{VELO} < 100$ .

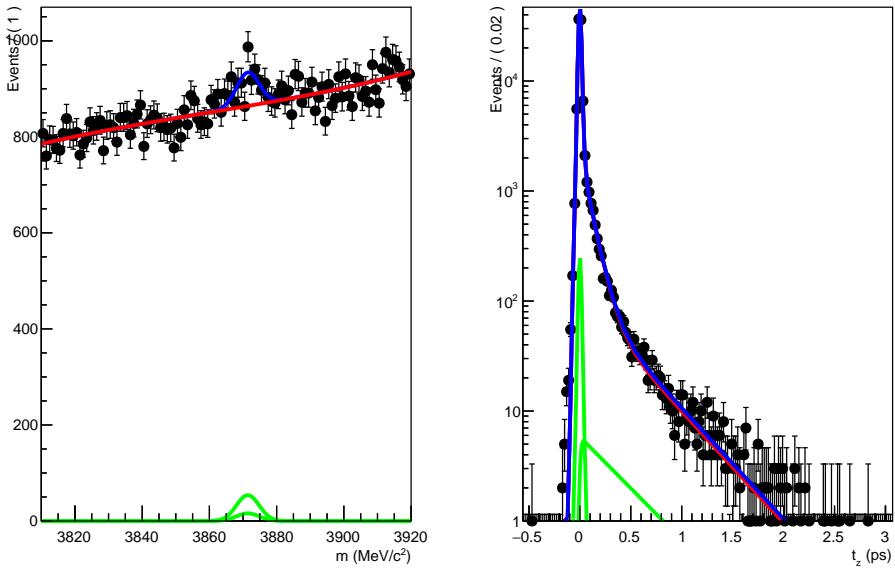


Figure 81: Simultaneous fit to the mass and proper time spectra for  $X(3872)$ , in the event activity range  $100 < N_{tracks}^{VELO} < 200$ .

923    **B.4     $X(3872)$  with Sideband Histogram as Proper Time Back-**  
 924    **ground**

| nVeloTracks bin center | $f_{prompt}$             | Inclusive $N_X$       |
|------------------------|--------------------------|-----------------------|
| 20                     | $0.924593 \pm 0.0135976$ | $1204.41 \pm 65.5481$ |
| 50                     | $0.894007 \pm 0.01709$   | $1490.17 \pm 87.1734$ |
| 70                     | $0.851982 \pm 0.0234985$ | $1226.63 \pm 97.5684$ |
| 90                     | $0.710925 \pm 0.0381282$ | $623.809 \pm 94.4885$ |
| 150                    | $0.703291 \pm 0.042304$  | $587.317 \pm 112.303$ |

Table 26: Values returned from the simultaneous fit to the  $X(3872)$  mass and proper time spectra, using the mass sideband histogram as background.

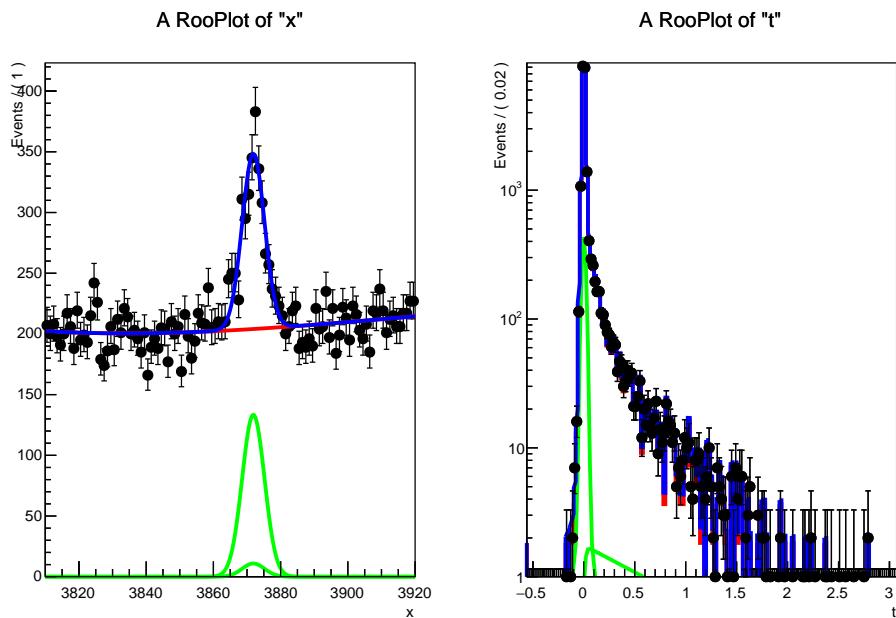


Figure 82: Simultaneous fit to the mass and proper time spectra for  $X(3872)$ , in the event activity range  $0 < N_{tracks}^{VELO} < 40$ .

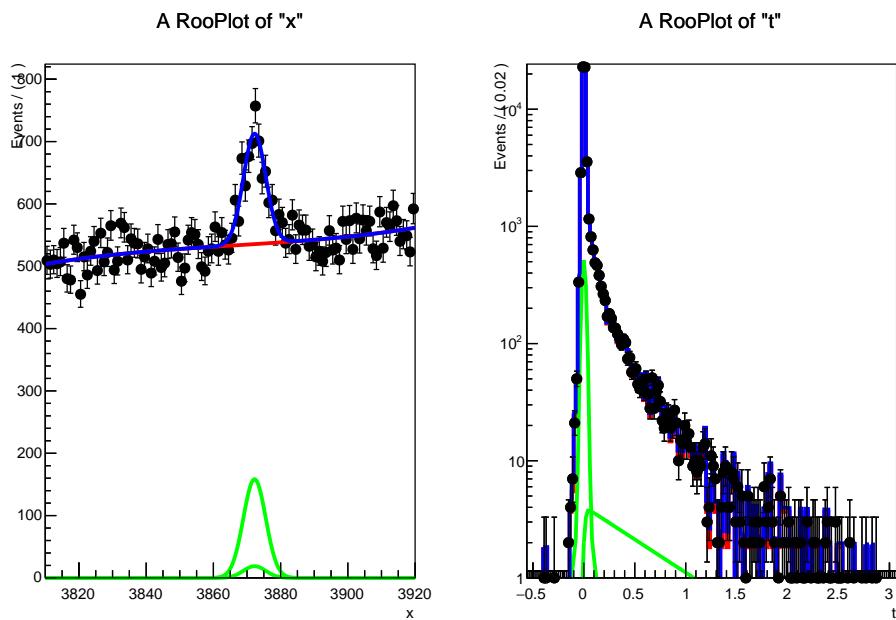


Figure 83: Simultaneous fit to the mass and proper time spectra for  $X(3872)$ , in the event activity range  $40 < N_{tracks}^{VELO} < 60$ .

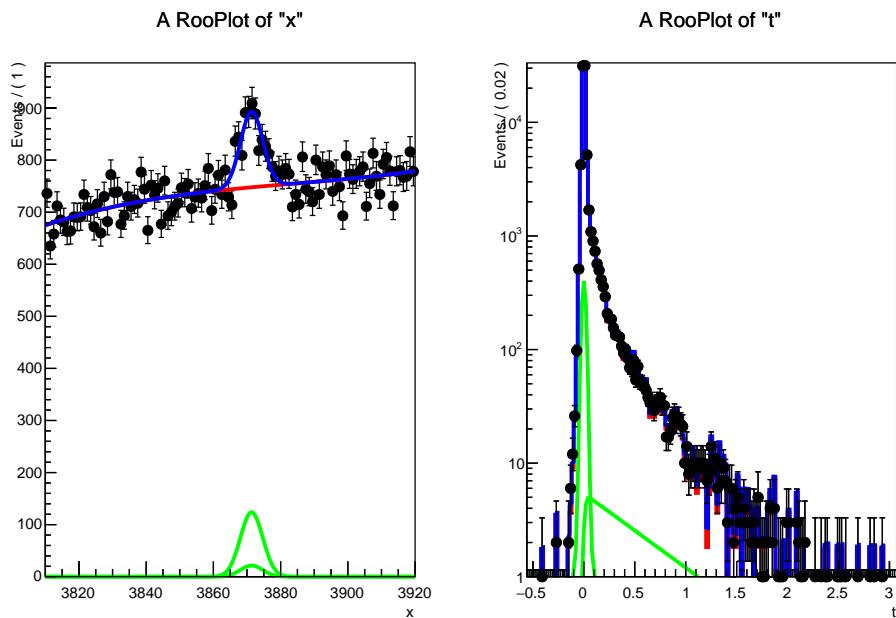


Figure 84: Simultaneous fit to the mass and proper time spectra for  $X(3872)$ , in the event activity range  $60 < N_{tracks}^{VELO} < 80$ .

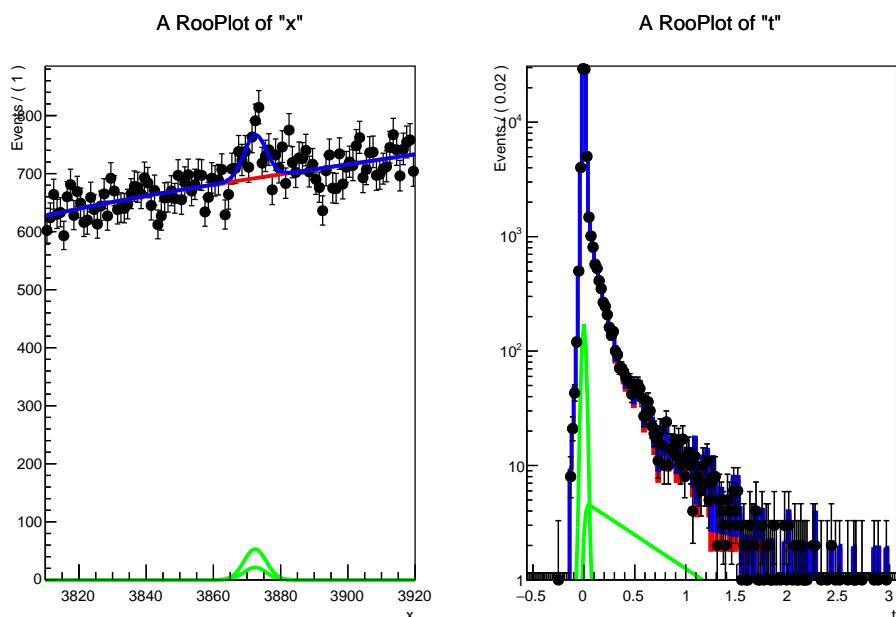


Figure 85: Simultaneous fit to the mass and proper time spectra for  $X(3872)$ , in the event activity range  $80 < N_{tracks}^{VELO} < 100$ .

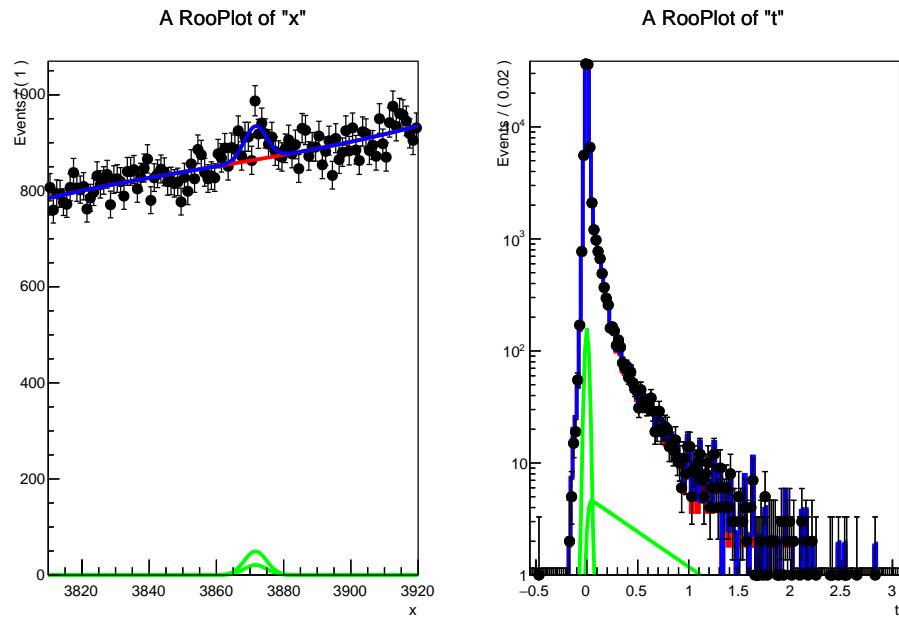


Figure 86: Simultaneous fit to the mass and proper time spectra for  $X(3872)$ , in the event activity range  $100 < N_{tracks}^{VELO} < 200$ .

## LHCb collaboration

925 A. N. Other<sup>1</sup>.

926 <sup>1</sup>*University of nowhere*