

# LHCb上粲夸克偶素极化的测量

(申请清华大学理学博士学位论文)

培养单位：工程物理系

学 科：物理学

研究 生：张 艳 席

指 导 教 师：高 原 宁 教 授



# **Measurement of Charmonium Polarization with the LHCb Detector**

Dissertation Submitted to  
**Tsinghua University**  
in partial fulfillment of the requirement  
for the degree of  
**Doctor of Philosophy**

in

**Physics**

by

**ZHANG Yanxi**

Dissertation Supervisor : Professor Gao Yuanning

**June, 2013**



# 关于学位论文使用授权的说明

本人完全了解清华大学有关保留、使用学位论文的规定，即：

清华大学拥有在著作权法规定范围内学位论文的使用权，其中包括：（1）已获学位的研究生必须按学校规定提交学位论文，学校可以采用影印、缩印或其他复制手段保存研究生上交的学位论文；（2）为教学和科研目的，学校可以将公开的学位论文作为资料在图书馆、资料室等场所供校内师生阅读，或在校园网上供校内师生浏览部分内容；（3）根据《中华人民共和国学位条例暂行实施办法》，向国家图书馆报送可以公开的学位论文。

本人保证遵守上述规定。

（保密的论文在解密后应遵守此规定）

作者签名：\_\_\_\_\_

导师签名：\_\_\_\_\_

日 期：\_\_\_\_\_

日 期：\_\_\_\_\_



## 摘要

粒子物理中，量子色动力学（QCD）是用来描述带色荷的粒子的强相互作用的理论。重夸克偶素是由重夸克和其反重夸克形成的束缚态，它们在高能物理实验上的产生截面和极化可以用来检验基于QCD的理论模型。关于重夸克偶素产生截面，色单态机制（CSM）计算的结果远低于实验测量值，而非相对论量子力学（NRQCD）模型的计算结果可以很好的描述实验数据。但是非相对论量子力学模型预言在大横动量区间S态的重夸克偶素具有很大的横向极化，这与实验观测不符。

LHCb实验是大型强子对撞机（LHC）上的实验之一，目的是为了精确测量底夸克和粲夸克相关物理。LHCb探测器是一个前向谱仪，覆盖了2–5的赝快度区间，具有良好的粒子重建和鉴别系统。在2011年，LHCb探测器采集了大约 $1.0 \text{ fb}^{-1}$ 的质心能量为7 TeV质子质子对撞的数据。利用其中约 $370 \text{ pb}^{-1}$ 的数据，分别在螺旋度（helicity）参考系和Collins-Soper参考系，在不同的横动量和快度区间，测量了瞬时产生的重夸克偶素 $J/\psi$ 粒子的三个极化参数 $\lambda_\theta, \lambda_{\theta\phi}$ 以及 $\lambda_\phi$ 。在运动学区间 $2 < p_T < 15 \text{ GeV}/c$ ,  $2.0 < y < 4.5$ 内，在螺旋度参考系中 $\lambda_\theta \simeq -0.2$ ，具有轻微的纵向极化并且随着 $J/\psi$ 横动量和快度的增加稍微减小，而 $\lambda_{\theta\phi}$ 和 $\lambda_\phi$ 在误差范围内近似为0。利用全部 $1.0 \text{ fb}^{-1}$ 的数据，还在螺旋度和Collins-Soper参考系中测量了瞬时 $\psi(2S)$ 的极化参数。测量结果表明，在误差范围内，在大部分运动学区间 $\psi(2S)$ 的三个极化参数都近似0，而在某些区域 $\psi(2S)$ 有轻微负的极化， $\lambda_\theta$ 处于-0.2到0之间。

分析在新的能量尺度和独特的快度区间给出了关于 $J/\psi$ 和 $\psi(2S)$ 极化的丰富的信息，并且 $\psi(2S)$ 的极化测量是所有实验分析中最精确的。与以前的实验一致，在LHCb上 $J/\psi$ 和 $\psi(2S)$ 并不具有很大的横向或者纵向极化。在LHCb和ALICE探测器共同覆盖的运动学区间内， $J/\psi$ 极化的测量结果和ALICE结果在误差范围内是一致的。LHCb关于 $J/\psi$ 和 $\psi(2S)$ 的极化测量不支持次领头阶CSM或者NRQCD的理论预言。

**关键词：**量子色动力学；重夸克偶素；极化；大型强子对撞机

## Abstract

In particle physics, quantum chromodynamics (QCD) is the theory used to describe the interaction of colored particles. Heavy quarkonium is the bound state of heavy quark and its anti-quark, and its production cross section and polarization can be used to test the theory models in the framework of QCD. The computation of the heavy quarkonium production cross section by color singlet mechanism (CSM) underestimates the experimental measurements, while results from the calculation of non-relativistic QCD (NRQCD) can describe experimental data very well. However, the NRQCD predicts that the  $S$  wave heavy quarkonium is heavily transversely polarized in the large transverse momentum region, which is contrary to experimental observations.

LHCb, dedicated for precision measurement in bottom and charm physics, is one of the experiments located at the Large Hadron Collider (LHC). The LHCb detector, which is a forward region spectrometer covering the pseudo rapidity range 2–5, has fine particle reconstruction and identification systems. In the year 2011, the LHCb detector collected about  $1.0 \text{ fb}^{-1}$   $pp$  collision data in a center-of-mass energy of 7 TeV. With integrated luminosity of about  $370 \text{ pb}^{-1}$ , the three polarization parameters  $\lambda_\theta$ ,  $\lambda_{\theta\phi}$  and  $\lambda_\phi$  of prompt  $J/\psi$  have been measured as a function of transverse momentum and rapidity in both the helicity and the Collins-Soper frame. In the helicity frame, in the kinematic range  $2 < p_T < 15 \text{ GeV}/c$ ,  $2.0 < y < 4.5$ ,  $\lambda_\theta \simeq -0.2$  decreases (in absolute value) with the increase of  $J/\psi$  transverse momentum and rapidity, which means that  $J/\psi$  is slightly longitudinally polarized, while  $\lambda_{\theta\phi}$  and  $\lambda_\phi$  are consistent with zero within errors. The  $\psi(2S)$  polarization parameters have been measured with all  $1.0 \text{ fb}^{-1}$  data. The results show that in both the helicity and the Collins-Soper frame, in most of the kinematic region, the three parameters are consistent with zero within errors, while in some bins the  $\psi(2S)$  has slightly negative polarization with  $\lambda_\theta$  between -0.2 and zero.

The analysis provide detailed information of  $J/\psi$  and  $\psi(2S)$  polarization in the new energy scale and unique rapidity range, and the  $\psi(2S)$  polarization result is the most precise one among all experiments. LHCb does not see strong transversely or longitudinally polarized  $J/\psi$  or  $\psi(2S)$  events, confirming previous measurements. The  $J/\psi$  polarization measurements at LHCb are consistent with the results from ALICE in the overlapping

---

Abstract

---

kinematic region. The  $J/\psi$  and  $\psi(2S)$  polarization results disfavor the calculations of CSM or NRQCD at NLO.

**Key words:** QCD; Quarkonium; Polarization; LHC

## Contents

<b>Chapter 1 Introduction</b> .....	1
1.1 A brief history of particle physics .....	1
1.1.1 The Standard Model .....	5
1.2 Heavy quarkonium physics .....	6
1.2.1 The characteristics of QCD .....	6
1.2.2 Heavy quarkonium spectroscopy .....	7
1.2.3 Heavy quarkonium production .....	8
1.3 Quarkonium production cross section and polarization: status and puzzles....	18
1.3.1 Quarkonium polarization in QCD models .....	18
1.3.2 Experimental tests of quarkonium production cross section and polarization model. ....	21
1.3.3 Quarkonium studies at LHCb .....	25
1.4 Summary .....	25
<b>Chapter 2 The LHCb Experiment</b> .....	27
2.1 The Large Hadron Collider at CERN .....	27
2.2 The LHCb detector .....	28
2.2.1 The vertex locator .....	30
2.2.2 The RICH detectors .....	33
2.2.3 The main tracking detectors .....	39
2.2.4 The LHCb calorimeter .....	46
2.2.5 The Muon system .....	49
2.2.6 The LHCb trigger .....	54
2.3 Summary .....	60
<b>Chapter 3 Determination of quarkonium polarization from experimental data</b> ..	61
3.1 From theory to experiment .....	61
3.2 Polarization extraction from data .....	64
3.2.1 Construction of likelihood estimator .....	64
3.2.2 Determination of normalization from Monte Carlo .....	65
3.2.3 Minimization with TMinuit .....	67
3.2.4 Fit validation using fully simulated samples .....	67

<b>Chapter 4 Prompt <math>J/\psi</math> polarization measurement .....</b>	72
4.1 Data sets and Selections.....	72
4.1.1 Selections .....	73
4.2 Polarization Fit .....	79
4.2.1 Efficiency .....	79
4.2.2 Likelihood estimator using event weights .....	81
4.3 Uncertainties on the polarization .....	83
4.3.1 Background subtraction fluctuation.....	84
4.3.2 Fluctuation of the normalization .....	89
4.3.3 Tracking efficiency .....	92
4.3.4 Monte Carlo validation .....	97
4.3.5 Background subtraction .....	104
4.3.6 Selections .....	107
4.3.7 Residual background of $J/\psi$ from $b$ decay .....	115
4.3.8 MagUp and MagDown symmetry .....	116
4.3.9 Summary of the uncertainties .....	118
4.4 Results .....	119
<b>Chapter 5 Prompt <math>\psi(2S)</math> polarization measurement .....</b>	128
5.1 Data sets and Selections.....	128
5.1.1 Data sets.....	128
5.1.2 Selections .....	130
5.2 Polarization Fit .....	133
5.3 Uncertainties on the polarization parameters .....	135
5.3.1 Fluctuations of the normalization .....	136
5.3.2 Tracking efficiency .....	137
5.3.3 Monte Carlo validation .....	140
5.3.4 Background subtraction .....	141
5.3.5 Binning effect .....	145
5.3.6 Residual background of $\psi(2S)$ from $b$ decay .....	148
5.3.7 Muon PID selection .....	149
5.3.8 $\tau_S$ selection .....	151
5.3.9 Detector resolution .....	153
5.3.10 MagUp and MagDown symmetry .....	153
5.3.11 Summary of the measurement uncertainties.....	153
5.4 Results .....	157
<b>Chapter 6 Summary .....</b>	164

---

## Contents

---

<b>Reference</b> .....	169
<b>Acknowledgement</b> .....	179
<b>声 明</b> .....	180
<b>Appendix A Crosscheck studies</b> .....	181
A.1 Efficiency .....	181
A.2 Background subtraction validation .....	189
A.3 DLL method versus <i>s</i> Weight method .....	191
A.4 The likelihood method goodness .....	193
<b>Appendix B Appendicular tables</b> .....	197
B.1 Fitting the invariant mass spectrum .....	197
B.2 Normalization parameters .....	200
B.3 Summary of all $J/\psi$ and $\psi(2S)$ polarization results in the HX and the CS frames .....	203
<b>Resume and publications</b> .....	224

## **List of main acronyms**

SM	Standard Model
QCD	Quantum chromodynamics
QED	Quantum electrodynamics
EW	Electro-weak
NRQCD	Non-relativistic QCD
SU(n)	Special unitary (group) of degree n
CEM	Color Evaporation Mechanism
CSM	Color Singlet Mechanism
COM	Color Octet Mechanism
DIS	Deep inelastic scattering
PDF	Parton distribution function
LO	Leading order
NLO	Next to leading order
NNLO	Next to next to leading order
LDE	Long distance effect
SDE	Short distance effect
LDME	Long distance matrix element
ME	Matrix element
$p_T$	Transverse momentum
$y$	Rapidity
$\eta$	Pseudo-rapidity
TeV/GeV/MeV	tera/giga/mega electron volt
LHC	Large Hadron Collider
LHCb	Large Hadron Collider for Bottom
CDF	Collider Detector at Fermilab
D0	DZero Experiment
LEP	Large Electron Position Collider
PV	Primary vertex
IP	Impact parameter
FOI	Field of interest
GEC	Global Event Cut

---

**List of main acronyms**

---

PID	Particle identification
HLT	High Level Trigger
DST	Data Summary Tape

## Chapter 1 **Introduction**

### 1.1 A brief history of particle physics

Human beings are known to have tried to understand the world around them dating back to the time civilization began. In ancient China, the universe is thought to be made up of five elements, the Metal, the Wood, the Water, the Fire and the Earth; Objects evolve by changing the relative amount of the five elements. Almost at the same period, the Greek Leucippus and Democritus raised the idea that matter are made up of atoms, the smallest indivisible particles. Early in the nineteenth century, John Dalton extended the concept of atoms to explain why materials always react with definite ratios.

It is generally agreed that modern particle physics was born with the discovery of the electron in 1897 by Thomson who also introduced a model of atoms—in an atom the electrons were suspended in a heavy positively charged paste. However Rutherford's  $\alpha$  particle scattering experiment in 1911 showed that the positive charge and most of the mass is concentrated in a tiny core called nucleus. The lightest nucleus, or the hydrogen nucleus is named as proton. The other constituent of the nucleus, the neutron, was discovered in 1932 by Chadwick. Up till then, people thought that matter is made up of the three elementary particles: the proton, the neutron and the electron.

In the early twentieth century, two great theories were discovered, relativity and quantum physics. Relativity, established by Einstein<sup>[1]</sup>, unifies space and time and is used to describe the motions of high speed particles. Quantum theory, discovered and developed by Planck<sup>[2]</sup>, Bohr<sup>[3]</sup>, Heisenberg<sup>[4]</sup>, Schrödinger<sup>[5]</sup>, Dirac<sup>[6]</sup> etc., successfully describes the dynamics of sub-atomic systems. Developed from the combination of relativity and the quantum physics, the first quantum field theory, quantum electrodynamics (QED), was introduced by Dirac<sup>[6]</sup>, Schwinger<sup>[7]</sup>, Feynman<sup>[8,9]</sup> and Tomonaga<sup>[10]</sup>. QED makes extremely precise predictions about electromagnetic interactions in the concept of fields, the creation and annihilation of particles in the vacuum<sup>[11]</sup>.

Almost at the same time when the electromagnetic field theory was being developed for sub-atomic particles, a lot more particles were found in cosmic ray and accelerator experiments: the neutrino, the pion family, the muon, strange particles etc.. Particle physicists tried to find the relationship between different particles and the similar charac-

teristics during their interactions with other particles, the symmetries. The isospin symmetry (by Heisenberg<sup>[12]</sup> and Wigner<sup>[13]</sup>) is applied to similar hadrons, for example the proton-neutron system (isospin 1/2), the pion family (isospin 1), the kaon family (isospin 1/2) etc. The Eightfold Way (by Murray Gell-Mann<sup>[14]</sup>) extends the isospin symmetry by putting more hadrons in the same multiplet. Particles in the same multiplet share the same quantum numbers, parity (P), spin and their hadronic interactions (cross sections) with other particles behave similarly. The Eightfold Way remarkably predicted the  $\Omega^-$  particle<sup>[15]</sup>, which is needed to make one of the multiplet complete. Behind the successful Eightfold Way is the quark model, which says that hadrons are made up of quarks (three by then), and the symmetry between different hadrons is in fact the symmetry between the three flavors of quarks up ( $u$ ), down ( $d$ ) and strange ( $s$ ). Quarks are spin 1/2 fermions just like the electron and muon, with charge in absolute value one ( $d, s$ ) or two ( $u$ ) times the third of electric charge. Quarks also have three color degree of freedom, called Red ( $R$ ), Green ( $G$ ) and Blue ( $B$ ) respectively. Up till 1994, the year when the top quark was found by CDF collaboration at Tevatron<sup>[16]</sup>, six flavors of quarks had been discovered, namely  $u, d, s, c$  (charm),  $b$  (bottom or beauty) and  $t$  (top) quark and each of them has three color states,  $R, G, B$ .

There is a large fraction of interactions in the universe that evolve weakly, and they dominate the transition of one type of particle to another. The decays of nuclei, the interaction of neutrinos with material fall into this kind. In the 1960s, Glashow<sup>[17]</sup>, Weinberg<sup>[18]</sup> and Salam<sup>[19]</sup> introduced an  $SU(2) \times SU(1)$  gauge theory to describe the weak interactions. QED is automatically included as the  $U(1)$  interaction (after the symmetry is broken). The weak and electro-magnetic interactions are thus unified as electro-weak (EW) interaction. To cope with the experimental findings that parity and charge conjugation (C) are not conserved in weak interactions, the left handed and righted fermions contribute to the electro-weak Lagrangian differently, the right handed fermions don't interact weakly at all. The electro-weak Lagrangian reads:

$$\begin{aligned}
 \mathcal{L}_{EW} = & \mathcal{L}_g + \mathcal{L}_f + \mathcal{L}_h \\
 = & -\frac{1}{4}W^{a\mu\nu}W_{\mu\nu}^a - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} \ (\equiv \mathcal{L}_g) \\
 & + \bar{Q}_j iD_L Q_j + \bar{u}_{Rj} iD_R u_{Rj} + \bar{d}_{Rj} iD_R d_{Rj} + \bar{L}_j iD_L L_j + \bar{e}_{Rj} iD_R e_{Rj} \ (\equiv \mathcal{L}_f) \\
 & + |D_\mu h|^2 - \lambda \left( |h|^2 - \frac{v^2}{2} \right)^2 \ (\equiv \mathcal{L}_h),
 \end{aligned} \tag{1-1}$$

where  $\mathcal{L}_g$  is the kinetic term describing the propagation and the interactions of the four gauge bosons  $W^a$  ( $a=1,2,3$ ) and  $B$ .  $\mathcal{L}_f$  is the kinetic term for fermions (quarks and leptons).  $D_L \equiv \gamma^\mu \partial_\mu + i\frac{g}{2}\tau^a \gamma^\mu W_\mu^a + i\frac{g'}{2}\gamma^\mu B_\mu$  and  $D_R \equiv \gamma^\mu \partial_\mu + ig' \gamma^\mu B_\mu$  are the covariant derivatives for left handed fermions and right handed fermions respectively, where  $g$  and  $g'$  are the coupling constants. With the covariant derivatives the interactions (vertices) between gauge bosons and fermions are introduced. The subscript  $j$  in  $\mathcal{L}_f$  runs over three generation of fermions (see later):  $Q, L$  are left handed doublet for quarks ( $\begin{pmatrix} u_L \\ d_L \end{pmatrix}$ ) and leptons ( $\begin{pmatrix} e_L \\ \nu_L \end{pmatrix}$ ) respectively, while  $u_R, d_R, e_R$  are right handed singlet for quarks and leptons. The  $\mathcal{L}_h$ , where  $D_\mu \equiv \partial_\mu + i\frac{g}{2}\tau^a W_\mu^a + i\frac{g'}{2}B_\mu$ , is Higgs field term describing Higgs self interaction and its interaction with gauge bosons.

In Higgs field term  $\mathcal{L}_h$ , the vacuum, which is the field state(s) when the energy minimizes, is not zero, but randomly chooses one of field states with lowest possible energy. The choice obeys SU(2) symmetry, which means that one choice can be converted to another one by an SU(2) rotation. For any choice of the vacuum field state, the SU(2) symmetry is broken, called spontaneous symmetry breaking (SSB), or Higgs mechanism for the special case of electro-weak SSB. After the symmetry breaking, the EW Lagrangian still behaves  $U(1)$  symmetry, which is identified as the QED. At the same time charged  $W$  bosons acquire masses, and so does the  $Z$  boson—the mixture of the neutral  $W$  boson with the  $B$  vector boson, while the other admixture  $A$  which is the quantization of the electro-magnetic field (photon) is still massless.

Since the establishment of electro-weak unification model, it has been tested by many experiments at high precision. In the year 1983, the  $W$  and the  $Z$  were discovered by the UA1 and UA2 collaborations<sup>[20–23]</sup>. The Large Electron Positron Collider (LEP) with its four experiments ALEPH, DELPHI, OPAL and L3 made remarkable measurements and tests of the electro-weak model. They precisely measured the  $Z$  boson mass and its decay width<sup>[24]</sup>. They also precisely measured the lepton forward-backward asymmetries of  $Z$  boson decay<sup>[24]</sup>, exploiting the signature of weak interactions.

In the late 1960s and early 1970s, a series of experimental apparatus was constructed to study the structure of hadrons, the strongly interacting particles. Among them, the Deep Inelastic Scattering (DIS) experiments used high energy electrons to bombard the nucleon targets. The DIS data revealed that the structure function of the proton scales as a combined variable of  $x = \frac{Q^2}{2Mv}$ , which is called Bjorken scaling variable<sup>[25]</sup>. The Bjorken scaling requires that the colliding targets inside the proton are point-like particles (or va-

lence quarks), which are spin 1/2 particles as shown by further experimental analysis<sup>[26]</sup>. The analysis with data also revealed the fact that the particles inside proton are weakly bounded at high colliding energies-a quark just carries a fraction of proton's momentum with some probability<sup>[27]</sup>, which is later described as parton distribution function (PDF). The weakly bound behavior of quarks inside the hadron is called asymptotic free. Many more DIS experiments were carried out later on, and they showed that, however, when the colliding quark carries only a small or very large fraction of the target's momentum in infinite-momentum frame the scaling is violated<sup>[28]</sup>. The scaling violation is explained by the idea that there are also sea quarks—virtual quark-antiquark pair—produced inside the target. The sea quarks are created and destroyed around the valence quark (the strength depending on the momentum transfer), making the interaction valence quark not point-like, and thus the scaling is violated. Besides, further results showed that quarks only carry about 50% of the nucleon momentum and theories had to include also the gluons which are created by the quarks as a basic component to explain this puzzle. Quarks and gluons make up of the nucleons, so they are also called partons. Later on, neutrino measurements<sup>[29]</sup> confirmed that quarks are spin 1/2 and found that they carry fractional electric charge, consistent with the quark model assumptions. By 1973 people knew as much as possible to construct the theory dominating the strong interactions between quarks and gluons, the gauge theory Quantum chromodynamics (QCD). The QCD Lagrangian reads:

$$\mathcal{L}_{\text{QCD}} = \bar{\psi}_i \left( i(\gamma^\mu D_\mu)_{ij} - m \delta_{ij} \right) \psi_j - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu} \quad (1-2)$$

where  $\psi_i(x)$  is the 4-component spinor quark field with color index  $i$ , and  $G_\mu^a(x)$  are the eight gluon field indexed by  $a$ .  $D_\mu$  is the covariant derivative, containing terms of gluon field:

$$D_\mu = \partial_\mu - \frac{1}{2} g G_\mu^a \Gamma^a \quad (1-3)$$

in which  $\Gamma^a$  are the three-dimensional generators of the SU(3) group indexed by  $a$  ( $a=1,2,\dots,8$ ). The latter term in  $D_\mu$ ,  $\frac{1}{2} g G_\mu^a \Gamma^a$ , introduces interactions between gluons and quarks. The kinetic term of gluons fields is the contraction of gauge field strength tensor which reads:

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g f^{abc} G_\mu^b G_\nu^c. \quad (1-4)$$

Within the gluon kinetic term, there are three-gluon and four-gluon interactions. Because of these gluon self interactions, the coupling constant  $\alpha_s = g^2/4\pi$  decreases when interaction energy scale increases:  $\alpha_s(Q) \rightarrow 0$  as  $Q \rightarrow \infty$ , and thus QCD can explain the asymptotic free phenomenon<sup>[30]</sup> in strong interactions. The gluons are later confirmed by the appearance of three jets events by JADE Collaboration<sup>[31]</sup>.

In particle physics, theory people try to create new physics, by introducing new symmetries or by studying the possible behavior of physics in high energy scale. There are still several open questions in particles physics. In nature there is another force, the gravitation, that has not well established in quantum field theory. There are active experimental work trying find the the force-carrier particle, the graviton. In the late 20th century, it is found that the baryon matter only make up of 4% of the universe material, while there are dark matter and dark energy whose nature are still beyond our scope. Besides, as we know that our universe is dominated by matter, however the  $CP$  asymmetries known in particle physics are unable to explain the size of the matter anti-matter asymmetries. Experimentally, two alterative methods of studying unknown physics are used. In direct search way, people try to find new particles or new phenomena in a much high energy frontier, while in the indirect searches, people try to find data that can not be described by the existing theories. The experimental particles physics with two complementary methods have established a lot of structures in fundamental physics and enriched a lot of our knowledge.

### 1.1.1 The Standard Model

In summary, particle physics is described by the Standard Model, which is an  $SU(3) \times SU(2) \times U(1)$  theory describing the properties of particles and their electromagnetic, weak and strong interactions. It gives us a picture of what the universe is made up of (the matter particles), what binds the matter together (the force carrier particles), and how they bound together (the interactions). According to their mass hierarchy and interactions in the electro-weak sector, the matter particles are classified into three generations, the first generation which includes up ( $u$ ), down ( $d$ ) quarks and  $e, \nu_e$  leptons; the second generation made up of charm ( $c$ ), strange ( $s$ ) quarks and  $\mu, \nu_\mu$  leptons and the third generation including top ( $t$ ), bottom ( $b$ ) quarks and  $\tau, \nu_\tau$  leptons. The force carrier particles include the photon which mediates electromagnetic interactions,  $W^\pm, Z$  bosons which are responsible for weak interaction, and the gluons which carry the strong force. All quarks have

colors, electric charge and weak charge so they can both interact strongly and join the electromagnetic and weak interactions, while the leptons will only participate in the electromagnetic and weak interactions. In Figure 1.1, the Standard Model particles and the possible interactions (represented by a curve) among them are shown.

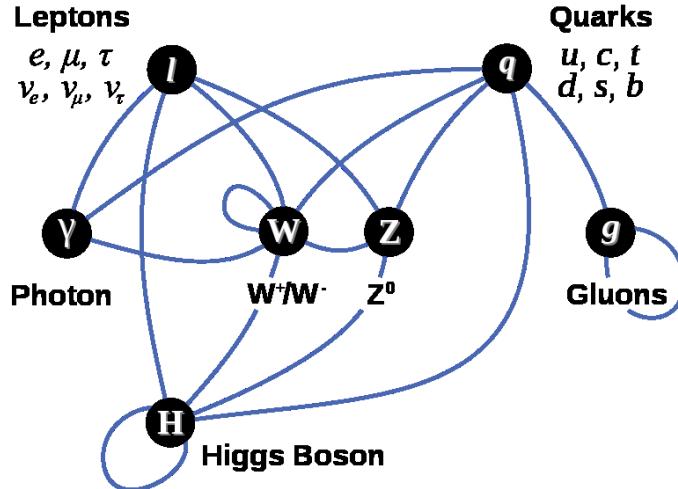


Figure 1.1 The gauge bosons  $W^\pm, Z, g, \gamma$ , the leptons and quarks are shown. The curves among them mean that they can interact directly. Figure reproduced from wikipedia.

In the summer of 2012, the two experiments ATLAS and CMS, with more than  $10 \text{ fb}^{-1}$  accumulated data respectively, independently found a new boson with mass around  $125 \text{ GeV}/c^2$ <sup>[32,33]</sup>. If this particle is finally confirmed to be the Standard Model Higgs, it will be another outstanding success of the Standard Model.

## 1.2 Heavy quarkonium physics

### 1.2.1 The characteristics of QCD

Although the electro-weak sector of the Standard Model has been well tested and measured by dedicated collaborations<sup>[34]</sup>, the strong interaction, as the SU(3) part of the SM, has not yet been understood very well<sup>[35-41]</sup>. The tests of QCD are generally limited to measure the differential production cross section or kinematics of hadrons, jets or multi-jets and make comparisons between experimental results and different models. Usually the theory models developed in the framework of QCD have to be tuned to reproduce data. The problems with precision tests of QCD come from theoretical aspects and from experimental practices. The common origin of these problems is the characteristics of QCD itself. The coupling constants of QCD decreases with the increase of the energy

scale of the QCD processes. So the QCD dynamics has two distinct phenomena: at one end it is asymptotic free (short distance effect, SDE), the partons interact weakly; while at the other end the color confinement (long distance effect, LDE), the partons interact fiercely. In perturbative theory, the calculations are expanded in powers (orders) of coupling constants, and higher order contributions are small compared to lower order ones on condition that the couplings are small enough, so the theoretical predictions can be as accurate as possible by using desired orders<sup>[42]</sup>. Due to the properties of QCD, the general method of the quantum field theory based on the perturbative theory can be no longer used for the QCD process: in SDE region of QCD, the perturbative calculations are valid, while in LDE region they fail, just because the coupling constant  $\alpha_s(q^2)$  decreases well below one at SDE region but increases strongly at LDE region, contrary to the QED constant  $\alpha$ , which is 1/137 at low energies and increases slowly with the interaction energy scale. Because of this special property, gluons or quarks can not be isolated from the collection of matter, but only color singlet hadrons. Besides, the processes of forming hadrons from partons and the hadronic bound states wave functions are not calculable because they involve the LDE. So the prediction power is limited by using data to fix the LDE and the cutoff between LDE and SDE will also introduce ambiguities to theoretical predictions. Even for hard processes (SDE), the calculations in higher orders are much more difficult for QCD because the diagrams increases highly in numbers. From experimental point of view, one should properly chooses the observables to establish the correspondence between observables obtained at the partonic (calculations) and the hadronic (measurements) levels. More interesting, during the last decades, many exotic hadrons, violating quantum numbers predicted by the SU(n) quark models for hadrons are observed<sup>[43–45]</sup>. The QCD theory is not sure of their nature, which urge us to dig more the behavior of QCD dynamics.

### 1.2.2 Heavy quarkonium spectroscopy

Quarkonium is the bound state of a quark and its anti-quark, similar to positronium—a system consisting of an electron and a positron<sup>[46]</sup>. By heavy quarks they mean charm ( $c$ ), bottom ( $b$ ) and top ( $t$ ) quarks, which are much heavier than the QCD confinement scale  $\Lambda$ <sup>[47]</sup>. The top quark is so heavy that it decays promptly with a lifetime ( $\approx 5 \times 10^{-25}$  s) too small to form hadrons ( $\approx \times 10^{-24}$  s)<sup>[48]</sup>. So typically by heavy quarkonia they mean bound states of  $c\bar{c}$  and  $b\bar{b}$  with a variety of  $J^{PC}$  quantum numbers determined by the total

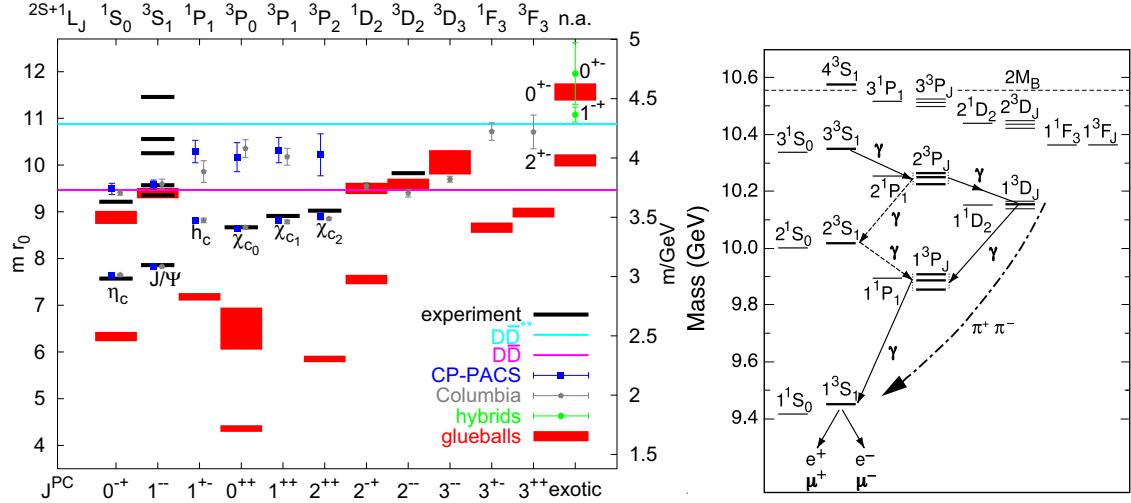


Figure 1.2 The charmonium and bottomonium spectroscopies together with their quantum numbers respectively. They are copied from [49,50] and [51] respectively.

spin  $S$  and angular momentum  $L$  of the two constituent heavy quarks. The quantum numbers run until the mass of the quarkonium exceeds double open charm (bottom) hadron threshold  $M(D\bar{D})$  ( $M(B\bar{B})$ ), at which the  $c\bar{c}$  ( $b\bar{b}$ ) will fragment into two open charm (bottom) hadrons easily. In Figure 1.2, the charmonium and bottomonium spectra are shown. In the following the capital  $Q$  denotes heavy quarks  $b$  or  $c$  unless otherwise specified.

The heavy quarkonium physics started with the revolutionary event when, in the year 1974, collaborations working at the Brookhaven National Laboratory (BNL)<sup>[52]</sup> and Stanford Linear Accelerator Center (SLAC)<sup>[53]</sup> respectively announced the discovery of the  $J/\psi$  meson, the first experimental evidence of the charm quark. It is now known that  $J/\psi$  is the bound state of  $c\bar{c}$  with  $S = 1$  and  $L = 0$ . The bottomonium state  $\Upsilon(1S)$  is discovered at Fermilab in the year 1977 in the decay  $\Upsilon \rightarrow \mu^+\mu^-$ <sup>[54]</sup>. Soon after the discovery of the  $1S$  ground states, the excited  $2S$  state  $\psi(2S)$  and  $2P$  wave states  $\chi_{cn}$  were discovered<sup>[55–58]</sup>, and studies of their properties and the search for abundant excited states were performed by further experiments<sup>[59–63]</sup> fruitfully.

### 1.2.3 Heavy quarkonium production

It is believed that the physical processes (production and decay) of heavy quarkonium are described by QCD, and they probe all the energy regimes of QCD, from the hard region (SDE dominates) at the parton level, where an expansion in the coupling constant  $\alpha_s(q^2)$  is applicable, to the low-energy region (LDE dominates), where nonperturbative effects

dominate. Heavy quarkonium bound states thus provide an ideal, and to some degree, unique laboratory where the understanding of nonperturbative QCD and its interplay with perturbative QCD in a controlled way<sup>[49]</sup> can be tested. In the past three decades, many theoretical works have been devoted to the task of applying QCD to the physics of heavy quarkonia. Together with the progress in experimental side, people have moved forward a lot in understanding the QCD. The theoretical calculations and experimental analysis are done on the production of heavy quarkonia. However many doubts are still up in the air. In high energy colliders, there are three sources of inclusive quarkonium production: direct production where a quarkonium is produced directly from initial partons, feed down production (if mass hierarchy allows) where a quarkonium is produced from the decay of heavier quarkonium states and  $b$  hadron decay (for charmonium only). The first two sources of quarkonium are also called prompt production with fraction from 90% in low  $p_T$  to 50% in  $p_T$  range above 20 GeV/c for  $J/\psi$  in hadronic production<sup>[64]</sup>, and of which about 60% are produced directly and 40% by feed down respectively<sup>[65]</sup>.

### 1.2.3.1 Quarkonium potential model

There are three scales associated with heavy quarkonium,  $M_Q$ ,  $M_Qv$ , and  $M_Qv^2$ , where  $v$  is the relative velocity of the two heavy quarks inside the heavy quarkonium and  $M_Q$  is the mass of the heavy quark  $Q$ , 1.2 GeV/c<sup>2</sup> for  $c$  and 4.2 GeV/c<sup>2</sup> for  $b$ , which sets the scale of the hard processes and is considered much heavier than the QCD scale  $\Lambda$ .  $M_Qv$  is the relative momentum of the two heavy quarks in the rest frame of the heavy quarkonium, which determines the size of the heavy quarkonium ( $\sim 1/(M_Qv)$ ).  $M_Qv^2$  is the scale of soft processes<sup>[66]</sup>. Due to the heavy mass  $M_Q$ , the square of relative speed of the two heavy quark,  $v^2$  is small ( $v^2 \approx 0.3$  for charmonium and  $v^2 \approx 0.1$  for bottomonium), namely the scale of the relativistic effects are relatively small. So the non-relativistic potential model is a good approximation and successfully describe the structure of the quarkonia<sup>[67–69]</sup>. The popular formula for the potential reads:

$$V(r) = -\frac{\kappa}{r} + \frac{r}{a^2}. \quad (1-5)$$

The potential in Equation 1-5 uses Coulomb type force in the  $r \rightarrow 0$  limit and uses linear growth trend to suggest the long distance confinement. There are other potentials that use logarithm like long distance component for better description of the moderate  $r$  region. Within the potential model, the quarkonium spectrum and interesting decay widths can

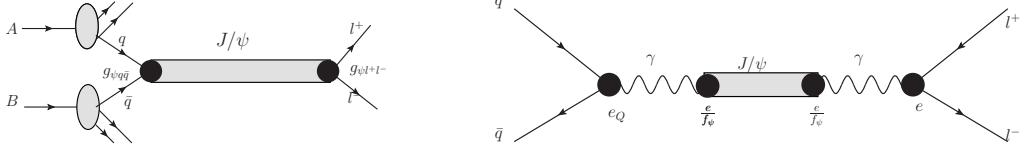


Figure 1.3 The Drell-Yan quark-antiquark annihilation diagram for  $J/\psi$  production and decay into lepton pairs. In the right plot, the coupling of quark (lepton) pair is thought to be through virtual massive photons.

be calculated and compared with experimentally measured mass spectrum splittings and the lepton pair decay width<sup>[70]</sup>. As an example, the decay width of  $J/\psi$  into lepton pair is theoretically determined by the formula:

$$\Gamma(\psi \rightarrow l\bar{l}) = \frac{16\pi e_c \alpha^2}{M^2} |\psi(0)|^2 \quad (1-6)$$

where  $e_c = 2/3$  is the electric charge of charm quark,  $M$  is the mass of the  $J/\psi$  and  $\alpha$  is the QED coupling. The wave function at the origin  $\psi(0)$  can be obtained by fitting to experimental data on lepton pair partial width. Then  $\psi(0)$  can be used for predictions of other decays and production.

The potential models mentioned above are simplified from the detailed perturbative computation of exact Wilson loop integral and these models take the main feature of the calculated quarkonium potential<sup>[71]</sup> in QCD. Higher orders, exploring the relativistic and quantum effects, can be introduced to refine the results. With all the details investigated carefully, the quarkonium model successfully describes the quarkonium spectroscopy and decay widths<sup>[49,72,73]</sup>.

### 1.2.3.2 QED process in heavy quarkonia

The  $J/\psi$  was considered to contribute to the Drell-Yan<sup>[74]</sup> quark-antiquark annihilation process (Fig.1.3) in hadronic collisions as:

$$q + \bar{q} \rightarrow J/\psi \rightarrow e^+ + e^- \quad (1-7)$$

in which  $q, \bar{q}$  are light quarks  $u, d$  or  $s$ . If one assumes that the process 1-7 is mediated by virtual photon as illustrated in Figure 1.3, and the  $\gamma \leftrightarrow J/\psi$  coupling is denoted by  $e/f_\psi$ , one simply has<sup>[75]</sup>

$$\Gamma(J/\psi \rightarrow \gamma^* \rightarrow e^+ e^-) = \left( \frac{e}{f_\psi} \right)^2 \frac{\alpha m_\psi}{3} \quad (1-8)$$

and

$$\Gamma(J/\psi \rightarrow \gamma^* \rightarrow q\bar{q}) = N_C e_q^2 \Gamma(J/\psi \rightarrow \gamma^* \rightarrow e^+ e^-), \quad (1-9)$$

where  $N_c$  is the number of color degrees of freedom which equals 3 in QCD. The electromagnetic production cross section of the  $J/\psi$  is calculated from the inverse process of Equation 1-9, and can be expressed as

$$\sigma_{q\bar{q}}^{J/\psi}(s) = \left(\frac{e}{f_\psi}\right)^2 \sigma_{q\bar{q}}^\gamma(s, m^{*2} = m_\psi^2) = \frac{4\pi^2}{3m_\psi^2} \frac{\Gamma(J/\psi \rightarrow q\bar{q})}{m_\psi} \delta\left(1 - \frac{m_\psi^2}{s}\right), \quad (1-10)$$

where  $s$  is the invariant mass of the colliding light quarks, which is constrained to be the heavy quarkonium mass. In the last equality, the fact  $\sigma_{q\bar{q}}^\gamma(s, m^{*2}) = \frac{e_q^2}{N_c} \frac{4\pi^2 \alpha}{m^{*2}} \delta\left(1 - \frac{m^{*2}}{s}\right)$  is used.

The  $J/\psi$  production cross section calculated from equation 1-10 is too low compared with experimental observations. There must be other mechanisms other than the  $\gamma \leftrightarrow J/\psi$  transition and they dominate  $J/\psi$  hadroproduction.

### 1.2.3.3 Color Evaporation Model

In the factorization models, the production of heavy quarkonium is through two separate steps, first a heavy  $Q\bar{Q}$  pair is produced, which is perturbative, and then the pair hadronizes to the final quarkonium, which is non-perturbative. The assumption is valid because the scale of producing heavy  $Q\bar{Q}$  ( $m_Q$ ) and that for  $Q\bar{Q}$  to form quarkonium ( $m_Q v^2$ ) are very different. Usually, in the first step, a heavy quark  $Q$  produced in high energy collisions fragments into hadrons, most often a meson  $Q\bar{q}$ . However, for a heavy quark pair  $Q\bar{q}$  produced in a color singlet state, this can happen only if the invariant mass exceeds twice the mass of the lightest  $Q\bar{q}$ . Below it, the  $Q\bar{Q}$  pair can only form discrete resonances, the quarkonium. This idea has been generalized for  $Q\bar{Q}$  with arbitrary color state and parity, spin states.

In the Color Evaporation Model, the cross section of bound heavy quark ( $Q\bar{Q}$ ) systems is obtained by integrating the cross section of open  $Q$ -quark production from the kinematical lower limit  $2m_c$  up to the threshold for open  $Q$ -meson ( $D\bar{D}$  for charm sector or  $B\bar{B}$  for bottomium family) production. The production of  $Q\bar{Q}$  can be calculated at any desired order in  $\alpha_s$  via hard parton processes. Since there are many quarkonium states (see section 1.2.2) in the allowed interval of the invariant mass of the  $Q\bar{Q}$  system, the transition fraction  $F_i$  for each specific state is free, assumed to be constant and has to be

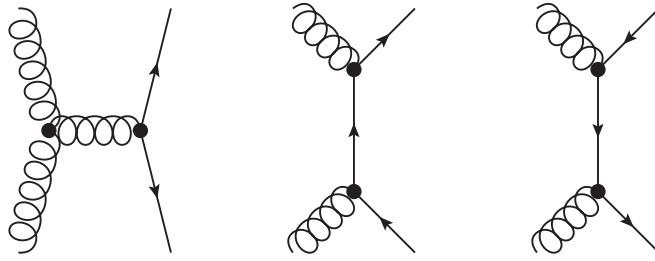


Figure 1.4 Gluon fusion diagrams that produce  $Q\bar{Q}$  at leading order.

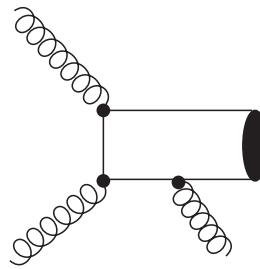


Figure 1.5 Gluon fusion diagrams that produce  $J/\psi$  with emitting a third gluon.

determined from data. It turned out that the  $F_i$  depends on the onium type, the process, the center of mass energy and the transverse momentum of quarkonium. Because of the inability to calculate relative fraction  $F_i$  for each quarkonium, and  $F_i$  is not universal, one can't avoid detailed calculation for the individual bound state production, which violate the assumption of CEM itself.

#### 1.2.3.4 Color Singlet Mechanism

In Color Single Mechanism (CSM), the  $Q\bar{Q}$  pair produced in the first step of the factorization must be in colorless state in order that they can form quarkonium. As gluons carry about half of the proton's momentum, the gluon fusion may be the predominately source of heavy quarkonium production<sup>[76]</sup> in hadronic environment. To obey the Landau-Yang theorem<sup>[77]</sup>, the  $J/\psi$  has to couple to at least 3 real gluons, while its partner  $\eta_c(^1S_0)$  and  $\chi_{cJ}(^3P_{0,2})$  can couple to 2 gluons at leading order (LO)  $\alpha_s^2$  with process  $g + g \rightarrow \eta_c(\chi_{c0,2})$  shown in Figure 1.4. One single gluon can't fragment into quarkonium entirely, because one gluon has color while the final state quarkonium is colorless. So it was estimated that the  $J/\psi$ , at LO in  $\alpha_s$ , is not produced directly in hadron-hadron collisions, but arises from strong or electromagnetic decay of charmed quarkonia that can couple to two gluons.

At next to LO (NLO) in  $\alpha_s$ , the  $J/\psi$  can be produced directly through the two gluon fusion process by emitting a third gluon (see Figure 1.5). Generally at the NLO there are

three diagrams that can produce heavy quarkonia in hadron colliders:

$$q + \bar{q} \rightarrow^{2S+1} L_J + g, \quad (1-11)$$

$$g + g \rightarrow^{2S+1} L_J + g, \quad (1-12)$$

and

$$g + q \rightarrow^{2S+1} L_J + q, \quad (1-13)$$

where  $q$  is a light hadron and  $L$  is the orbital angular momentum. When the heavy  $Q\bar{Q}$  pair is produced, their formation probability into quarkonium is related to  $\psi^{(L)}(0)$ —the  $L$ th order derivative of the quarkonium wave function at the origin. The  $\psi^{(L)}(0)$  can be extracted from the quarkonium decay-width similar to the production in Drell-Yan process.

In summary, considering the orders in  $\alpha_s$  and the branching fraction of  $\chi_{cJ}$  decays into  $J/\psi\gamma$ , which is about 1%, 35% and 20% for  $J = 0, 1, 2$  respectively, it is estimated that, for the production of  $J/\psi$  :

$$\begin{aligned} \sigma(\chi_{c2}) \times \text{Br}(\chi_{c2} \rightarrow \gamma J/\psi) &\gg \sigma(gg \rightarrow gJ/\psi) \\ &\gg \sigma(\chi_{c1}) \times \text{Br}(\chi_{c1} \rightarrow \gamma J/\psi) \\ &\approx \sigma(\chi_{c0}) \times \text{Br}(\chi_{c0} \rightarrow \gamma J/\psi) \end{aligned} \quad (1-14)$$

When the collision energy is much higher, the  $b$ -hadron decays will be an important source of  $J/\psi$  production, even comparable to the size from  $Q\bar{Q}$  processes in the early estimation<sup>[78]</sup>.

However the CDF analysis<sup>[79]</sup> later contradicted this picture of quarkonium production. They measured  $J/\psi$  and  $\psi$  production cross section in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.8$  TeV, which indicates that not only the normalization but also the  $p_T$  dependence is wrong. Besides, fixed target experiment<sup>[80,81]</sup> showed that the fraction of  $J/\psi$  coming from  $\chi_{cJ}$  decays is about 30%, and  $\sigma_{\chi_{c1}}/\sigma_{\chi_{c2}} \approx 1$ , in contrary to the gluon fusion model which says  $\chi_{c1}$  production starts at higher order in  $\alpha_s$ .

The order in  $\alpha_s$  is probably not the only thing that determines the cross section. Calculations showed that quarkonium produced from gluon fragmentation, higher order in  $\alpha_s$ , is enhanced by a factor of  $(E/m_Q)^2$  coming from the virtual gluon propagator<sup>[82]</sup>.

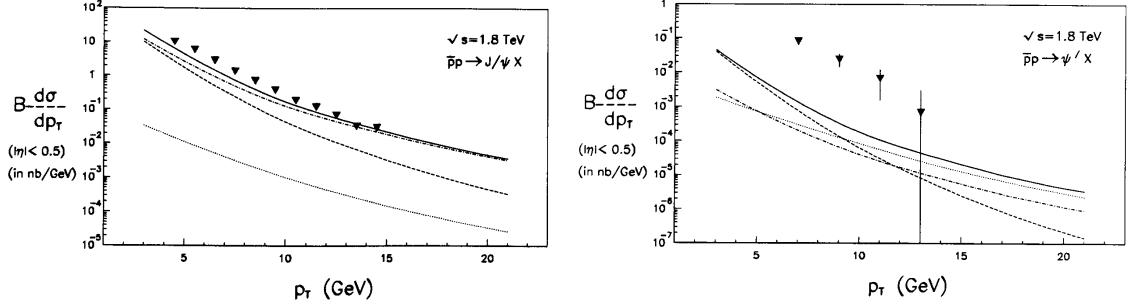


Figure 1.6 Cross section  $d\sigma/dp_T$  for inclusive  $J/\psi$  (left) and  $\psi(2S)$  (right). Data are from CDF 1.8 TeV run<sup>[?]</sup>. The different curves correspond to the direct production via fusion (dashed line), the gluon fragmentation contribution (dashed-dotted line), the charm quark fragmentation term (dotted line) and the sum of all contributions (solid line). For  $J/\psi$ , the calculations also include the decays from  $\chi_{cJ}$  states.

Here  $E$  is the energy of the virtual gluon with a value of the size of the large transverse momentum  $p_T$  in the large  $p_T$  small rapidity region (central detector region). The LO for production of  $nS$  state via fragmentation is  $g^* \rightarrow \psi gg$  with order  $\alpha_s^4$ . Fragmentation process can help to explain the enhancement of  $Z \rightarrow \psi c\bar{c}$  over  $Z \rightarrow \psi gg$  at the  $Z$  pole. The former one is through intermediate decay  $Z \rightarrow c\bar{c}$ , and then the  $c$  ( $\bar{c}$ ) fragments into  $\psi$  and an additional  $c$  ( $\bar{c}$ ) quark which is enhanced by  $M_Z^2/M_\psi^2$ <sup>[83]</sup>. Calculations of gluon fragmentation into  $P$ -wave charmonium in LO reveal that  $\chi_{c1}$  and  $\chi_{c2}$  can be produced almost equally at large transverse momentum<sup>[84]</sup>. In Figure 1.6,  $J/\psi$  and  $\psi(2S)$  production calculations<sup>[85]</sup> including gluon fusion, gluon fragmentation and charm quark fragmentation are compared with data. The gluon fragmentation dominates gluon fusion at large transverse momentum, which makes the  $J/\psi$  results more consistent between theoretical computations and data, however for  $\psi(2S)$  the theoretical calculation is rather below the data points, a factor of 30 inconsistency. So CSM at LO calculations can't be all the story for the quarkonium production.

### 1.2.3.5 Color Octet Mechanism

There are situations that the  $Q\bar{Q}$  produced in color octet states can't be avoided in CSM. For example, in the CSM calculations, the gluon fragmentation functions related to transition  $g \rightarrow c\bar{c}(^3P_J, \underline{1})$  are logarithmically singular.  $c\bar{c}$  pair produced in color octet states have to be introduced to cancel the infrared divergence<sup>[84,86]</sup>. Bodwin<sup>[87]</sup>, Braaten<sup>[88]</sup> and Lepage<sup>[89]</sup> generalized the color octet contribution in the quarkonium reactions in the Non-Relativistic QCD (NRQCD) framework for all quarkonium productions.

NRQCD consists of nonrelativistic Schrödinger field theory for the heavy quark and antiquark, while the theory for light quarks and gluons is still relativistic, the same as the usual QCD. NRQCD can be made as precise as possible to full QCD through the addition of relativistic corrections in orders of the heavy quark velocity  $v$  (in quarkonium rest frame). The velocity is small for heavy quarkonium,  $Mv^2 \ll M$  as  $v^2 \ll 1$ . So in computations of quarkonium production, calculations will be organized both in perturbative orders of  $\alpha_s$  and in relativistic correction orders of  $v$ , and both expansions are vital.

To extract the NRQCD Lagrangian from full QCD shown in Equation 1-2, in which heavy quarks are described by 4-component Dirac spinor fields, first an ultraviolet momentum cutoff  $\Lambda$  which is much smaller  $M_Q$  is introduced. The cutoff is appropriate to the analysis of heavy quarkonium because inside the quarkonium the involved momenta are of order  $Mv$  or less. By the cutoff the relativistic states of heavy quarks are discarded, to compensate this, local interactions are added, as the relativistic effects can only happen locally because high virtuality state can only travel a short distance. In the second step, 4-component spinors for the heavy quarks are block-diagonalized resulting in two-component Pauli spinor fields for heavy quarks and for heavy antiquarks respectively. The resulting field theory is the NRQCD. The NRQCD Lagrangian for heavy quarks (kinetic term), which is nonrelativistic explicitly reads:

$$\mathcal{L}_{\text{heavy}} = \psi^\dagger \left( iD_t + \frac{\mathbf{D}^2}{2M} \right) \psi + \chi^\dagger \left( iD_t - \frac{\mathbf{D}^2}{2M} \right) \chi \quad (1-15)$$

where  $\psi$  ( $\chi$ ) is the 2-component spinor field that annihilates (creates) a heavy quark (antiquark), and  $D_t$  and  $\mathbf{D}$  are the time and space components of the covariant derivative  $D^\mu$  (see Equation 1-3).

The relativistic correction terms have the form  $\frac{c}{M^{d_i}} \psi^\dagger f(\mathbf{D}, G^{\mu\nu}, \sigma^i) \psi + c.c.$ , where  $c$  is function of  $\alpha_s$ , and  $M^{d_i}$  a power of mass to make  $c$  dimensionless, while  $f(\mathbf{D}, G^{\mu\nu}, \sigma^i)$  is function of space component of covariant derivative  $\mathbf{D}$ , gluon fields strength tensor  $G^{\mu\nu}$  and spin matrix  $\sigma^i$ . The relative importance of such corrections are ordered by the heavy quark velocity  $v$ <sup>[89]</sup>, which is called  $v$ -scaling.

The annihilation and creation of  $Q\bar{Q}$  are described in NRQCD by the 4-fermion interactions constructed with two quark fields and two antiquark fields which generally look like:

$$\mathcal{L}_{\text{4-fermion}} = \chi^\dagger \mathcal{K}_n \psi \psi^\dagger \mathcal{K}'_n \chi \quad (1-16)$$

where  $\mathcal{K}_n$  are products of a color matrix, a spin matrix, a polynomial of space component of the covariant derivative and other fields, and there are infinite many 4-fermion terms with various spin-color state (matrix) configuration. All these interaction terms can contribute to the production (annihilation) of quarkonium through  $Q\bar{Q}$  pair, with the relative importance of these terms are also characterized by their powers in velocity  $v$ .

At the same time, the quarkonium can be expanded in a series of Fock-states ordered by the powers of velocity and labeled by the color-spin quantum number. The LO  $O(1)$  expansion for quarkonium  $H(nJ^{PC})$  will be color singlet (1)  $Q\bar{Q}$  with the same spin-orbital angular momentum  $|Q\bar{Q}[^{2S+1}L_J, 1]\rangle$ , and the Fock-state  $|Q\bar{Q}g\rangle$  has an amplitude of  $O(v)$  with color octet  $Q\bar{Q}$  in total spin  $S$  and orbital angular momentum  $L \pm 1$  (E1 transition), and so on. In general, the quarkonium is made up of superposition of infinite many Fock-states ordered in powers of  $v^{[90]}$ :

$$\begin{aligned}
 |H(nJ^{PC})\rangle = & \quad O(1)|Q\bar{Q}[^{2S+1}L_J^{(1)}]\rangle \\
 & + O(v)|Q\bar{Q}[^{2S+1}(L \pm 1)_{J'}^{(8)}g]\rangle \\
 & + O(v^2)|Q\bar{Q}[^{2S+1}L_J^{(1)}gg]\rangle + \dots (gg + Q\bar{Q}[\text{color - spin}]) \\
 & + O(v^3)\dots
 \end{aligned} \tag{1-17}$$

In which the color states are labeled as (1) or (8) for color singlet and octet respectively. The amplitude in Equation 1-17 for Fock-state with  $n_g$  gluons are suppressed by orders of  $v^{n_g}$  or higher. Here  $g$  is the dynamic gluon whose effect cannot be incorporated into an instantaneous potential and whose typical energy is  $M_Q v^2$ <sup>[91]</sup>. The LO term  $|Q\bar{Q}[^{2S+1}L_J^{(1)}]\rangle$  has exactly the same color-spin configuration as the quarkonium, so in the  $v \rightarrow 0$  limit, when higher order terms vanish, the expansion reduces to the Color Singlet Model result.

NRQCD also uses the factorization of short distance part and long distance part to predict quarkonium production:

$$\begin{aligned}
 \sigma(H) &= \sum_n \frac{F_n(\Lambda)}{M^{d_n-4}} \langle 0 | \chi^\dagger \mathcal{K}_n \psi \sum_{X, m_J} |H + X\rangle \langle H + X| \psi^\dagger \mathcal{K}'_n \chi |0\rangle \\
 &\equiv \sum_n \frac{F_n(\Lambda)}{M^{d_n-4}} \langle 0 | O_n^H(\Lambda) |0\rangle
 \end{aligned} \tag{1-18}$$

where short distance coefficients  $F_n$  determine the production of  $Q\bar{Q}$  in a specific state  $n$ , and they are independent of the final quarkonium state  $H$ . The  $F_n$  are calculated perturbatively using Feynman diagram method in powers of  $\alpha_s$ . The long distance matrix element (LDME)  $\langle 0 | O_n^H(\Lambda) |0\rangle$  determines the transition of  $Q\bar{Q}$  in a specific state into final

state quarkonium  $H$  with strength depending on  $H$  and  $Q\bar{Q}$  color-spin state but not on the detailed production processes. Equation 1-18 sums over all the color-spin state of  $Q\bar{Q}$ ,  $2J + 1$  spin states of quarkonium  $H$  and all other accompanying final state particles  $X$  (so valid for  $H$  inclusive production only). There are infinitely many long distance terms, however at any given order in  $v$  only a small number of them contribute and the orders in  $v$  come both from the Fock-state expansion and 4-fermion quark-antiquark operators. In the NRQCD factorization of quarkonium production (Equation 1-18), the nonperturbative matrix elements  $\langle 0|O_n^H(2S+1)L_J|0\rangle$  can be obtained by fitting the predicted result to experimental data. However, the LO matrix elements are related to the CSM nonperturbative factors (up to order  $v^4$ ), which can be extracted from quarkonium leptonic decays width. For example,

$$\langle 0|O_1^\psi(^3S_1)|0\rangle = \frac{N_c}{2\pi}|R(0)|^2 \quad (1-19)$$

$$\langle 0|O_1^{\chi_J}(^3P_J)|0\rangle = \frac{3N_c}{2\pi}(2J+1)|R'(0)|^2 \quad (1-20)$$

where  $R(0)$  is the quarkonium wave function at the origin as usual.

Concerning the production of  $nS$  quarkonium ( $H = \psi(nS)$  for example), at lower orders of  $v$ , a few color octet matrix elements are involved: the LDME  $\langle 0|O_8^H(^3S_1)|0\rangle$  takes the Fock-state  $|Q\bar{Q}(^3S_1^{(8)})gg\rangle$  ( $v^2 \times v^2$  order) and the LO 4-fermion ( $v^6$  order) interaction, and so it is  $v^4$  suppressed compared to color singlet matrix element ( $v^6$  order);  $\langle 0|O_8^H(^3P_J)|0\rangle$  takes the Fock-state  $|Q\bar{Q}(^3P_J^{(8)})g\rangle$  ( $v \times v$  order) and the  $v^8$  order 4-fermion interaction, and so it is also  $v^4$  suppressed.  $\langle 0|O_8^H(^1S_0)|0\rangle$  takes the Fock-state  $|Q\bar{Q}(^1S_0^{(8)})g\rangle$  ( $v^2 \times v^2$  order) and also the LO 4-fermion ( $v^6$  order) interaction, and so it is  $v^4$  suppressed too. Although these long distance CO matrix elements are suppressed in powers of  $v$  compared to CS matrix elements, the short distance factors  $F_n$  is enhanced in powers of  $\alpha_s$ .

Calculation had been done in the framework of NRQCD factorization by including the LO CO matrix elements<sup>[92–94]</sup>. The  $p_T$  spectrums associated with various matrix element transitions are different in high transverse momentum range:

$$\langle 0|O_1^H(^3S_1)|0\rangle : 1/p_T^8 \quad (1-21)$$

$$\langle 0|O_8^H(^3S_1)|0\rangle : 1/p_T^4 \quad (1-22)$$

$$\langle 0|O_8^H(^3P_J)|0\rangle : 1/p_T^6 \quad (1-23)$$

$$\langle 0|O_8^H(^1S_0)|0\rangle : 1/p_T^6 \quad (1-24)$$

The predictions are compared with CDF measurement, with the matrix elements extracted by fitting the data (see Figure 1.7). The CO matrix elements extracted are found to be consistent with the  $v^4$  suppression<sup>[92,94]</sup> compared to the CS one:

$$\langle 0 | O_1^{J/\psi}(^3S_1) | 0 \rangle \approx 5.7(0.8) \times 10^{-1} \text{GeV}^3 \quad (1-25)$$

$$\langle 0 | O_8^{J/\psi}(^3S_1) | 0 \rangle \approx 1.1(0.1) \times 10^{-2} \text{GeV}^3 \quad (1-26)$$

$$\frac{\langle 0 | O_8^{J/\psi}(^3P_0) | 0 \rangle}{M_c^2} + \frac{\langle 0 | O_8^{J/\psi}(^1S_0) | 0 \rangle}{3.5} \approx 1.3(0.3) \times 10^{-2} \text{GeV}^3 \quad (1-27)$$

Because matrix elements  $\langle 0 | O_8^{J/\psi}(^3P_J) | 0 \rangle$  and  $\langle 0 | O_8^{J/\psi}(^1S_0) | 0 \rangle$  predict almost the same transverse momentum distribution, they can not be discriminated through the fitting to  $p_T$  spectrum, only the combination is sensitive. As has been shown in the CSM, in the hadronic production of  $S$  state quarkonium, gluon fragmentation process—color octet at LO in  $\alpha_s$ —is enhanced by  $(p_T/M_Q)^2$ , so it will dominate when the large transverse momentum is large enough. Even with only gluon fragmentation at leading order, and only  $\langle 0 | O_8^H(^3S_1) | 0 \rangle$  taken into account, the predictions of  $J/\psi$  and  $\psi(2S)$  production cross section can fit CDF measurements very well<sup>[95–97]</sup>. This can be seen from Figure 1.7, at large transverse momentum, the  $\langle 0 | O_8^{J/\psi}(^3S_1) | 0 \rangle$  dominates other contributions.

The NRQCD has also been applied in quarkonium processes in fixed target colliders and photo-production. By including the CO matrix elements, prediction of  $\Upsilon$  and  $\psi$  cross section at LEP agrees well with experimental limits<sup>[98]</sup>. While at Hadron-Elektron-Ring-Anlage (HERA) experiments, quarkonium is produced from  $ep$  colliders, the CS plus CO contribution can reproduce the differential cross section as a function of  $p_T$ <sup>[99–101]</sup>, and the CO component is less important at low  $p_T$ <sup>[102]</sup> compared to CS contribution. By inclusion of the CO contribution, the theoretical computation of the center-of-mass energy dependent total cross section of  $\psi$ , which is dominated by gluon fusion processes, in fixed target experiments can describe data very well<sup>[103,104]</sup>.

### 1.3 Quarkonium production cross section and polarization: status and puzzles

#### 1.3.1 Quarkonium polarization in QCD models

The mechanisms for quarkonium can also calculate the polarized quarkonium production,  $S$  state particularly (see section 3.1). In the theoretical computation for polarized production,  $Q\bar{Q}$  with different helicity states are produced at short distance, and in the

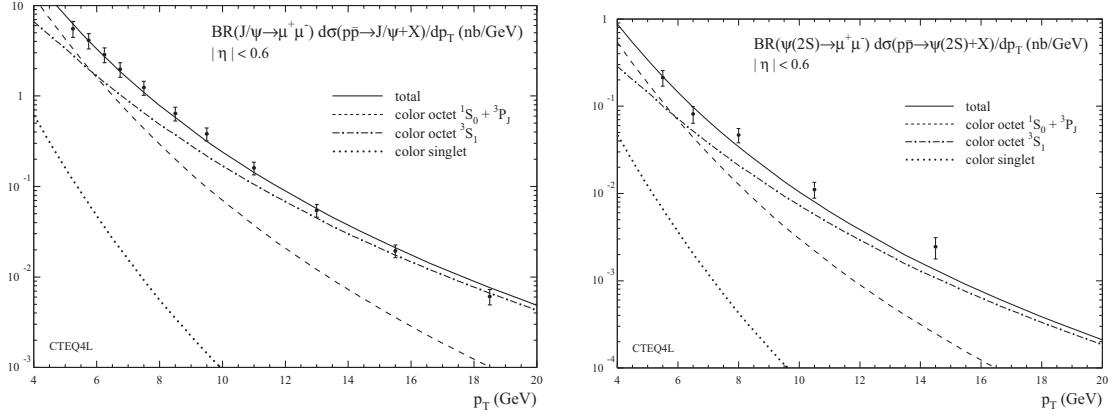


Figure 1.7  $J/\psi$  (left) and  $\psi(2S)$  (right) production cross section. Theoretical calculations including only gluon fragmentation in NRQCD framework are compared to CDF measurements. The long distance matrix elements are extracted by a  $\chi^2$  fitting to CDF data. The dashed curve depicts the direct CS contribution, the dot-dashed curve illustrates the  $\langle 0 | \mathcal{O}_8^H ({}^3S_1) | 0 \rangle$  cross section and the dotted curve denotes the combined  $\langle 0 | \mathcal{O}_8^H ({}^3P_J) | 0 \rangle$  and  $\langle 0 | \mathcal{O}_8^H ({}^1S_0) | 0 \rangle$  distributions. The plots are referenced from [94].

process of long distance transition of  $Q\bar{Q}$  into quarkonium, the helicity is conserved at LO in  $v^2$  because of heavy quark spin symmetry<sup>[105,106]</sup>. Quarkonium polarization provides more information to determine the production mechanism. The CSM and COM predict different polarization of  $\psi$  state at  $e^+e^-$  collider<sup>[107]</sup> and hadron colliders<sup>[108]</sup>. In the COM, the polarized production involves interference between contributions of different intermediate states<sup>[106]</sup>, which introduces additional matrix elements, making the prediction of polarized production not trivial.

At large transverse momentum the direct quarkonium production is dominated by gluon fragmentation into color octet  ${}^3S_1$  heavy quark pairs. The fragmenting gluons are effectively on shell, and so the intermediate heavy quark pairs are transversely polarized. The transition of these polarized intermediate states into quarkonium will predominantly preserve the polarization. Corrections such as spin symmetry breaking chromomagnetic interactions, and high order gluonic radiation will only depolarize the  $Q\bar{Q}$  at order of 10% – 15%<sup>[109]</sup>. The main source of depolarization comes from the contribution of color octet  ${}^1S_0$  and  ${}^3P_J$  components. At  $\mathcal{O}(v^4)$  in the velocity expansion, the polarization yield from all production channels can be calculated unambiguously in NRQCD in terms of the non-perturbative matrix elements that have been determined from the unpolarized cross section<sup>[110]</sup>.

The NRQCD predicts that  $\psi(nS)$  and  $\Upsilon(nS)$  are transversely polarized increasing

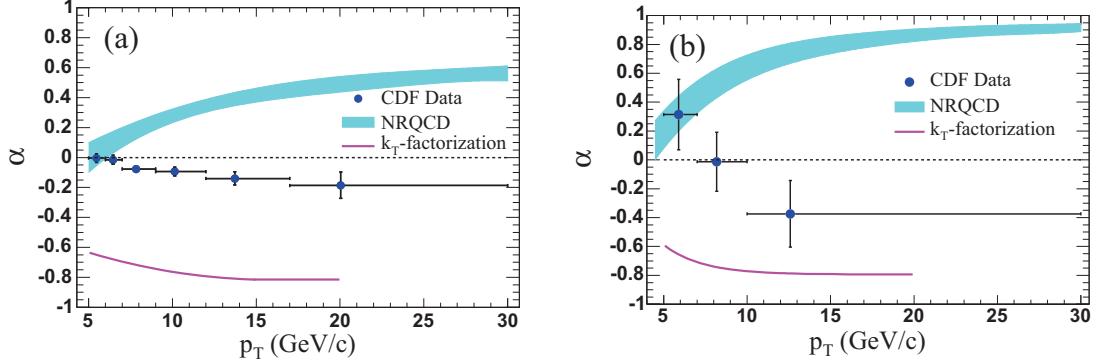


Figure 1.8 Prompt  $J/\psi$  (left) and  $\psi(2S)$  (right) polarization as functions of  $p_T$  at CDF. The  $J/\psi$  events used contain the feed down from excited states. The band (line) is the prediction of NRQCD ( $k_T$  factorization) model.

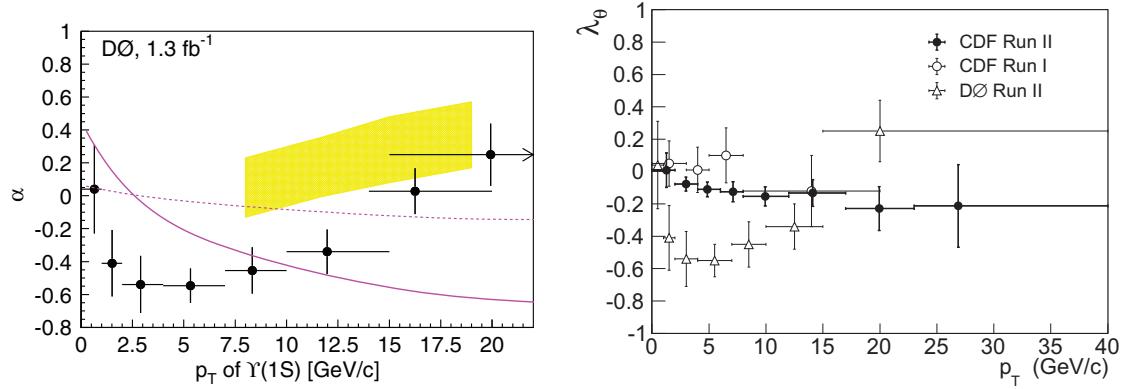


Figure 1.9  $\Upsilon$  polarization at D0 (left) and CDF (right) as functions  $p_T$ . In the left plot the D0 measurements are compared to NRQCD (band) and  $k_T$  factorization (line) model prediction. In the right plot, the CDF Run I, CDF Run II and D0 run II results are compared.

with transverse momentum at the Tevatron energy<sup>[94,111,112]</sup>. At Tevatron, both the CDF and the D0 groups measure the prompt quarkonium polarization with abundant events, however the  $\psi$ <sup>[113]</sup> and  $\Upsilon$ <sup>[114,115]</sup> polarization disfavor the strong transverse polarization at lower  $p_T$  or when the transverse momentum is quite larger than the quarkonium mass. As can be seen from Figures 1.8 and 1.9, the quarkonium polarization are almost zero or slightly longitudinal. Besides, for  $\Upsilon$  polarization, the data points and the dependence on transverse momentum of CDF and D0 are not quite consistent, which reveals the difficulty of the angular analysis.

### 1.3.2 Experimental tests of quarkonium production cross section and polarization model.

The underlying model for quarkonium production has to correctly address the inclusive cross section and polarization simultaneously in various production environments. The analysis by the ZEUS<sup>[99,116]</sup> and H1<sup>[100,101]</sup> collaborations at the HERA *ep* accelerator show that both the NLO CS and NRQCD can predict the  $p_T$  dependent cross section of  $J/\psi$ , however the  $p_T$  dependent  $J/\psi$  polarization is not consistent with NLO CS prediction<sup>[117,118]</sup>, and the inelasticity  $z$  (which is defined as  $\vec{P}_{J/\psi} \cdot \vec{p} / \vec{P}_{\gamma^*} \cdot \vec{p}$ ) dependent cross section at large  $p_T$  is contrary to NRQCD computation<sup>[100]</sup>. The COM and the LO C-SM calculations, which predict highly transversely polarized quarkonium at large  $p_T$ , fail to address the  $\psi(nS)$  and  $\Upsilon(nS)$  polarizations at hadronic colliders<sup>[113–115]</sup>. The  $J/\psi$  hadronic production at lower transverse momentum region ( $p_T < 5$  GeV/ $c$ ) are studied at Relativistic Heavy Ion Collider (RHIC) by the PHENIX<sup>[119]</sup> and the STAR<sup>[120]</sup> collaborations, and at HERA-B<sup>[121]</sup>, they measured slightly longitudinal polarization, which is consistent with the NRQCD prediction<sup>[120,122]</sup> but disfavors LO CS result.

The quarkonium production analysis — the (differential) cross section and polarization — at high  $p_T$  is crucial to understand the theory models. In the factorization procedure, the (differential) cross section is predicted by introducing several non-perturbative factors, which represent the main components that happen in the real production history. To make the transverse momentum distributions to mimic data, the fractions of various contributions can be determined during the fitting to data. Then the combination of different components with various ( $p_T$  dependent) polarization, with the interference correctly treated, the polarization can be predicted. However there are at least two things that have to be clarified, firstly because only the calculated  $p_T$  spectrum, which is derived at LO or NLO in  $\alpha_s$ , is in fact used to fit data, higher order short distance effects are involved if higher order computations give the same  $p_T$  distributions. However, the same calculations at LO or NLO in  $\alpha_s$  for polarization can be very different from those at higher orders. As can be seen from the color singlet contributions, even at the next to leading order (NLO), when the QCD corrections are included, the gap between LO CS calculation and data will be filled dramatically and the polarization can be quite different from those of LO calculations<sup>[123–125]</sup>. NLO CS produces longitudinally polarized  $\psi$  at large  $p_T$  as shown in Figure 1.10, while  $\psi$ -mesons are strongly transversely polarized at LO within CS framework. So the parametrization in the NRQCD calculation in lower orders

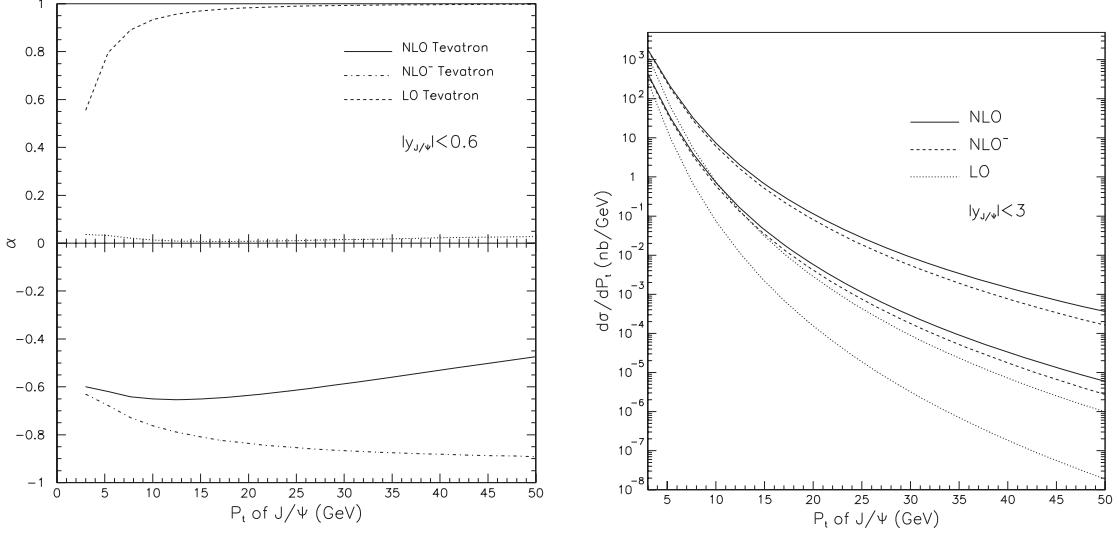


Figure 1.10  $J/\psi$  polarization (left) and transverse momentum distribution (right) in color singlet model at LO and NLO<sup>-</sup>. NLO<sup>-</sup> denotes result excluding contribution from subprocess  $gg \rightarrow J/\psi c\bar{c}$ . Figure referenced from[125].

of  $\alpha_s$  can produce the shape of quarkonium differential cross section as a function of  $p_T$ , but can not easily reproduce its polarization without detailed calculation of higher order effects.

Secondly, as seen by QCD correction calculation<sup>[123,125]</sup> for color singlet contribution, the  $p_T$  spectrum at NLO can be quite different as shown in Figure 1.10, two to three orders of magnitude enhanced at large  $p_T$  ( $\simeq 50$  GeV/c for  $J/\psi$ ), so the fractions (and also the NRQCD matrix elements) of various contributions, which generate different polarized quarkonium, will generally differ from the lower calculations. Thus the polarization predicted at lower order will not hold when higher effects have to be included. The NRQCD has also been incorporated in the  $k_T$  factorization framework<sup>[126]</sup>, where the unintegrated gluon distribution function is used to calculate the hard process of the quarkonium production. The  $k_T$  factorization procedure explores the higher order corrections due to soft gluon radiation. In the  $k_T$  factorization model, the initial gluons carry transverse momentum of order  $k_T$ , and thus will modify the  $p_T$  spectrum and fraction (long distance matrix elements) of various components. The  $k_T$  factorization can well reproduce the Tevatron and HERA quarkonium differential cross section at large  $p_T$  and also low  $p_T$  regions, however in the  $k_T$  framework at LO orders the  $\langle 0 | \mathcal{O}_8^H ({}^3S_1) | 0 \rangle$  matrix element almost vanishes, and thus directly produced  $\psi$  at Tevatron energy is highly longitudinally polarized, which is inconsistent with the measured prompt  $J/\psi$  polarization.

The  $\chi_{cJ}$  production and especially the  $\chi_{c2}$  over  $\chi_{c1}$  ratio is important in the test of various production mechanisms. Because the  $\langle 0 | \mathcal{O}_8^H ({}^3S_1) | 0 \rangle$  component dominates at high  $p_T$  at LO, the  $\chi_{c2}$  over  $\chi_{c1}$  ratio is thought to be 5/3—the ratio of the number of spin states  $(2J + 1)$ <sup>[127]</sup>. However the CDF collaboration measured the  $\chi_{c2}$  over  $\chi_{c1}$  ratio to be 0.75<sup>[128]</sup>, very much below the naive number from spin counting. On the other hand, the calculation<sup>[127]</sup> shows that NLO results can solve this problem. At NLO, the CS  $\langle 0 | \mathcal{O}_1^H ({}^3P_J) | 0 \rangle$  channel also scales as  $1/p_T^4$ , with short distance strength associated with  $\chi_{c1}$  slightly larger than the short distance strength of the  $\chi_{c2}$  channel, and by properly choosing the two relevant matrix elements, the CDF measurements of the  $\chi_{c2}$  over  $\chi_{c1}$  ratio and the  $\chi_{cJ}$  differential cross section as a function of  $J/\psi p_T$  can be well reproduced.

Double quarkonium production also challenges the various models. The exclusive production cross section of double charmonium in  $e^+e^- \rightarrow J/\psi H$  or inclusive production cross section  $e^+e^- \rightarrow J/\psi c\bar{c}X$  at  $\sqrt{s} = 10.6$  GeV measured by Belle<sup>[129–131]</sup> and BaBar<sup>[132]</sup> are larger than the LO calculations in NRQCD or CS by more than a factor of five<sup>[133–135]</sup>. However by introducing the NLO corrections<sup>[136,137]</sup> and the pure QED transitions (double  $\gamma^*$  processes)<sup>[138]</sup> and also the relativistic corrections<sup>[139]</sup>, the NRQCD can give result very close to the measurements of inclusive  $J/\psi c\bar{c}X$  and exclusive  $J/\psi \eta_c$  cross section.

### 1.3.2.1 Quarkonium at the Large Hadron Collider

At the Large Hadron Collider, the quarkonia are studied by the four experiments ATLAS, CMS, ALICE and LHCb with  $pp$  collisions at a center-of-mass energy of 7 TeV, 2.76 TeV or 8 TeV.

At ATLAS, the inclusive cross section of  $J/\psi$  is measured in bins of rapidity and transverse momentum, in the rapidity coverage range  $|y| < 2.4$  and the transverse momentum range  $1 < p_T < 70$  GeV/ $c$ <sup>[140]</sup>. The CEM model does not describe the shape of  $p_T$  distribution, while the calculation of direct  $J/\psi$  with NLO CSM model underestimates the cross section greatly, however partially NLO CSM computation agrees with data much better both in the global normalization of the cross section and in the shape of the  $p_T$ . The  $\Upsilon(1S)$  differential cross section in bins of rapidity and transverse momentum is also analyzed in the range  $p_T < 25$  GeV/ $c$ <sup>[141]</sup>. Again, the NLO CSM calculation for direct  $\Upsilon$  production underestimate the cross section significantly, while the LO NRQCD result, without including the uncertainties of the calculations, can't reproduce the

Table 1.1 Inclusive quarkonium cross section measurement at LHCb

Measurement	Kinematic range	Comparison with theory
$J/\psi$	$p_T < 14 \text{ GeV}/c$	Data is consistent with NLO NRQCD
	$2.0 < y < 4.5$	NLO CEM can't produce the $p_T$ spectrum
$\psi(2S)$	$p_T < 16 \text{ GeV}/c$	Data is consistent with NLO NRQCD
	$2.0 < y < 4.5$	and partially (main) NNLO CSM
$\Upsilon(nS)$	$p_T < 15 \text{ GeV}/c$	Data is consistent with NLO NRQCD
	$2.0 < y < 4.5$	and partially (main) NNLO CSM

normalization and the  $p_T$  dependence of the cross section data simultaneously.

At CMS, the  $J/\psi$  and the  $\psi(2S)$  differential cross sections as a function of rapidity and transverse momentum has been extracted in the rapidity range  $|y| < 2.4$  and the transverse momentum region  $p_T < 50 \text{ GeV}/c$ <sup>[142,143]</sup>. The NLO NRQCD calculations coincide with data reasonably well, while the CEM can't produce the shape of the  $p_T$  spectrum. The differential cross section of  $\Upsilon(nS)$  states are also measured in the rapidity range  $|y| < 2.0$  and the transverse momentum region  $p_T < 30 \text{ GeV}/c$ <sup>[144]</sup>. The LO NRQCD prediction overestimates the normalization of the cross section by a factor of two.

ATLAS and CMS agree with each other in the overlap kinematic regions for the quarkonium cross section measurement.

The ALICE experiment measures the  $J/\psi$  differential cross section in the forward rapidity region with  $p_T < 8 \text{ GeV}/c$  and  $2.5 < y < 4$  with 7 TeV data<sup>[145]</sup> and with 2.76 TeV data<sup>[146]</sup>. Their results agree with NLO NRQCD calculations and the LHCb measurements in the overlapping regions.

The LHCb experiment also works in the forward rapidity region. The inclusive cross section of  $J/\psi$ <sup>[147]</sup>,  $\psi(2S)$ <sup>[148]</sup> and  $\Upsilon(nS)$ <sup>[149]</sup> are measured at LHCb and compared with theoretical calculations as shown in Table 1.1.

The production cross section of  $J/\psi$  pairs are extracted at LHCb<sup>[150]</sup>, and is found to be consistent with theoretical calculations based on NRQCD framework<sup>[151,152]</sup> within statistical errors.

The  $\chi_{c2}/\chi_{c1}$ <sup>[153]</sup> and  $\chi_{c1,2}/J/\psi$  ratios<sup>[154]</sup> are also measured at LHCb as a function of  $J/\psi$  transverse momentum, and the results are found to be in good agreement with NLO

NRQCD calculations but not LO CSM.

The  $J/\psi$  polarization<sup>[155]</sup> and the  $\Upsilon(nS)$  polarization<sup>[156]</sup> are also measured by the ALICE and the CMS collaboration respectively. Both experiments find that the quarkonium polarization consistent with zero within experimental uncertainties—no evidence of large transverse or large longitudinal polarization observed.

### 1.3.3 Quarkonium studies at LHCb

The experiments, from fixed target to collider, with center-of-mass energy from tens of GeV to 7 TeV, do not agree with the theoretical calculations in both production cross section and polarization of heavy quarkonium simultaneously. However it should be noted that, the previous experimental studies for polarization are either one dimensional analysis by integrating over the other variable or performed by integrating over the rapidity in a quite large range. The one dimensional analysis can introduce additional systematic uncertainties if the experimental detection efficiency depends on all variables. The measurement performed in large kinematic bins reduces the ability of comparison with theoretical calculations and with results from other experiments, because different experiments have different kinematic coverage. Besides, the measurements for  $J/\psi$  polarization are for prompt or inclusive  $J/\psi$  respectively. The former one includes  $J/\psi$  from feed down of excited states, while the latter one also includes the  $J/\psi$  from  $b$ -hadron decay. However most theoretical calculations are for  $J/\psi$  mesons that are produced directly from partons.

At LHCb, the LHCb detector collects about  $3 \text{ fb}^{-1}$   $pp$  collision data in the three year data taking from 2010 to 2012, with a center of mass energy of 7 TeV or 8 TeV. With a cross section of about  $10.5 \text{ }\mu\text{b}$ <sup>[147]</sup> ( $1.44 \text{ }\mu\text{b}$ <sup>[148]</sup>) in the LHCb fiducial coverage, the  $J/\psi$  ( $\psi(2S)$ ) is produced abundantly. With large statistics, full angular analysis can be performed to extract the  $J/\psi$  ( $\psi(2S)$ ) polarization as a function of the quarkonium transverse momentum and rapidity in the unique LHCb kinematic region. Our measurement will supply abundant information for the test of quarkonium production mechanisms.

## 1.4 Summary

This chapter described the theoretical background and experimental studies of heavy quarkonium physics, emphasizing on the theoretical models that predict the production cross section and polarization in the framework of QCD. This chapter also emphasizes on

the experimental tests of models by measuring the heavy quarkonium differential production cross section and polarization. None of the model can describe the heavy quarkonium production and polarization simultaneously in the experiments before LHC and at the LHC. At LHC, the production cross section has been measured for various heavy quarkonia, and the higher order CSM or NRQCD models can describe the data very well. In this dissertation, the  $J/\psi$  and  $\psi(2S)$  polarization are measured, providing supplementary information to the test of heavy quarkonium production mechanisms at the energy scale of the LHC.

## Chapter 2 The LHCb Experiment

### 2.1 The Large Hadron Collider at CERN

The Large Hadron Collider (LHC), operated by the European Organization for Nuclear Research (CERN), is a two-ring superconducting accelerator installed in the 27 km long Large Electron Positron (LEP) Collider tunnel, aiming at the discovery of the Higgs particle and the study of rare events with a designed center-of-mass energy up to 14 TeV in the proton-proton ( $pp$ ) collisions<sup>[157]</sup>. It can also accelerate heavy ions used for the study of quark matter physics. In the designed running conditions, the LHC ring has 2808 bunches separated by multiples of 25 ns in each beam, with about  $1.15 \times 10^{11}$  protons in each bunch.

The protons are prepared by the CERN accelerator complex that successively increases their energy as illustrated in Figure 2.1. Initially 50 MeV protons are generated at the linear particle accelerator (LINAC2), then the protons are fed to the Proton Synchrotron Booster (PSB) at which the protons are accelerated to 1.4 GeV and injected into the Proton Synchrotron (PS), where the protons are accelerated to 26 GeV. Finally the Super Proton Synchrotron (SPS) is used to further increase the proton energy to 450 GeV. The LHC accepts these protons and accelerate them up to the energy needed for physics studies. There are six experiments located along the LHC ring, ATLAS<sup>[158]</sup>, CMS<sup>[159]</sup>, ALICE<sup>[160]</sup>, LHCb<sup>[161]</sup>, TOTEM<sup>[162]</sup>, LHCf<sup>[163]</sup> and MoEDAL. The ATLAS and CMS experiments with a designed peak luminosity of  $10^{34} \text{cm}^{-2} \text{s}^{-1}$  intend to search the Higgs boson and to study the terascale physics directly. The ALICE experiment explores the quark-gluon matter physics using high energy ion collisions with nominal peak luminosity of  $10^{27} \text{cm}^{-2} \text{s}^{-1}$ . LHCb, operated with a low peak luminosity  $10^{32} \text{cm}^{-2} \text{s}^{-1}$  aims at precision studies of flavor physics in charm and bottom hadrons. The TOTEM experiment studies forward particles focusing on physics that is not accessible in the general-purpose experiments, for example the transverse size of the proton beam. The LHCf experiment uses forward particles created inside the LHC as a source to study cosmic rays in laboratory conditions. The MoEDAL experiment tries to search for the Magnetic Monopole and other highly ionizing Stable Massive Particles at the LHC.

In November 2009, beams were successfully injected into the LHC at the energy

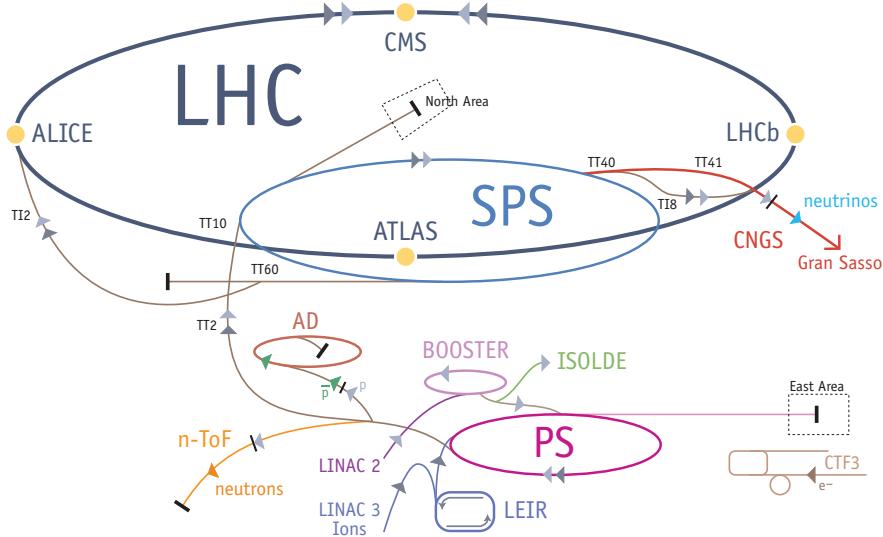


Figure 2.1 The CERN accelerator complex. The injection chain of protons and ions from LINAC to LHC are shown. The four large experiments are also shown at the interaction points on the LHC ring.

of 450 GeV and collided at each of the interaction points. In the year 2010 and 2011, the LHC collided protons at  $\sqrt{s} = 7$  TeV with increasing proton bunches in each beam, and about  $50 \text{ pb}^{-1}$  data in the year 2010,  $6 \text{ fb}^{-1}$  data in the year 2011 were collected by both ATLAS and CMS experiments, while LHCb collected about  $40 \text{ pb}^{-1}$  and  $1 \text{ fb}^{-1}$  respectively. In the year 2012, both the ATLAS and CMS experiments collected more than  $20 \text{ fb}^{-1}$   $pp$  data at  $\sqrt{s} = 8$  TeV, while LHCb took about  $2 \text{ fb}^{-1}$ . The increase of data in storage with time in the year 2011 and 2012 for the four large experiments are shown in Figure 2.2 respectively.

## 2.2 The LHCb detector

LHCb is an experiment dedicated to heavy flavor physics at the LHC. Its primary goal is to measure the Standard Model parameters precisely with high statistics and to look for indirect evidence of new physics in CP violation and rare decays of beauty and charm hadrons<sup>[161]</sup>. The size of CP violation that is already known in the Standard Model weak interactions is too small to explain the amount of matter anti-matter asymmetry in the universe. New sources of CP violation in or beyond the Standard Model is therefore needed to solve the problem. With much improved precision due to high statistics, the effect of new sources might be seen in heavy flavor physics. With the large  $c\bar{c}$  cross section ( $\approx 3.5 \text{ mb}$ ) and  $b\bar{b}$  cross section ( $\approx 500 \text{ } \mu\text{b}$ ) in  $pp$  collisions at center-of-mass

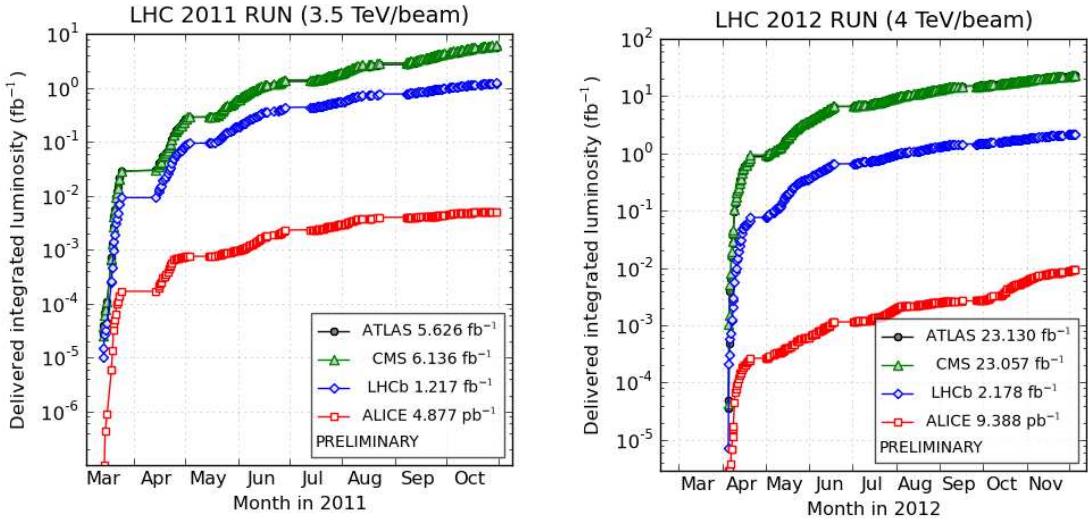


Figure 2.2 The integrated luminosity collected by the four experiments at LHC for the year 2011 (left) and year 2012 (right).

energy of 14 TeV, the LHC is the most copious source of charm and bottom hadrons in the world. With a modest peak luminosity of  $2 \times 10^{32} \text{cm}^{-2}\text{s}^{-1}$  for LHCb,  $10^{13} c\bar{c}$  and  $10^{12} b\bar{b}$  pairs would be produced yearly ( $10^7$  s). The peak luminosity of LHCb is lower than at ATLAS and CMS ( $10^{34} \text{cm}^{-2}\text{s}^{-1}$ ) for the advantages: the average number of  $pp$  interactions per bunch crossing is smaller which make it simpler to separate particles from different  $pp$  collisions, and the occupancy in the detector remains low and radiation damage is reduced.

The LHCb detector, shown in Figure 2.4, uses a right handed coordinate system with the origin located at the interaction point (IP), the  $z$ -axis pointing downstream the LHCb detector,  $y$ -axis pointing upward and  $x$ -axis completing the right-handed system. At LHC the direction of the  $b$  and  $\bar{b}$  hadrons produced from high energies  $pp$  collisions are correlated: predominately in the same forward or backward cone as can be seen from Figure 2.3. Because of this, the LHCb detector is designed as a single-arm spectrometer with a forward angular coverage from approximately 10 mrad to 300 mrad in the bending ( $x - z$ ) plane and 10 mrad to 250 mrad in the non-bending ( $y - z$ ) plane. The LHCb detector accepts 40% of the  $b\bar{b}$  cross sections while it covers only 4% of the  $4\pi$  solid angle. The LHCb detector consists of several sub detectors that record information like hit positions, energy deposits, etc. which are used to reconstruct  $pp$  collision events. Close to the interaction point, there is the Vertex Locator<sup>[164]</sup> (VELO) placed around the interaction region. The first Ring Imaging Cherenkov<sup>[165]</sup> (RICH1) detector

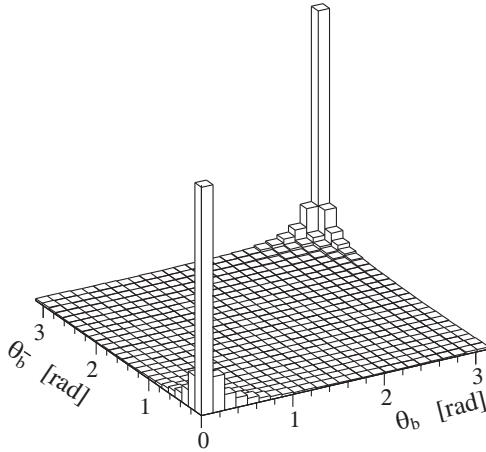


Figure 2.3 Simulated distribution of polar angles of  $b$  and  $\bar{b}$  hadrons with regards to the direction of the beam.

lies behind the VELO, allowing for separation of low momentum charged  $\pi - K$  particles. After that there are three tracking systems, the tracker turicensis<sup>[166]</sup> (TT), inner tracker<sup>[166]</sup> (IT) and outer tracker<sup>[167]</sup> (OT) positioned upstream and downstream of the dipole magnet<sup>[168]</sup>, providing momentum measurement of charged particles. They are followed by the RICH2 detector, which can discriminate charged particles with momentum up to 100  $\text{GeV}/c$ . The calorimeter system<sup>[169]</sup> providing energy deposit clusters for the reconstruction of photons,  $\pi^0$ , neutral hadrons and identification (ID) of electrons comprises a Scintillator Pad Detector (SPD), a Preshower (PS), and an electromagnetic (ECAL) and a hadronic (HCAL) calorimeter placed one after another. The five muon chambers<sup>[170]</sup>, one placed before and the other four after the calorimeter system, provide muon reconstruction and identification. The information supplied by the tracking system including VELO, the calorimeters and the muon systems is also used for the trigger decision during the online data taking.

### 2.2.1 The vertex locator

One of the signatures of a  $b$ -hadron decay at LHC is a displaced vertex (secondary vertex) from their production vertex—the primary  $pp$  collision vertex (PV). To explore this feature, excellent vertex resolution is needed to separate the secondary vertices from the primary vertices and this is accomplished by the VELO at LHCb. It requires that the VELO can measure the hit position precisely and as close as possible to the interaction region, to reduce the uncertainties when the trajectories are extrapolated from the hits to

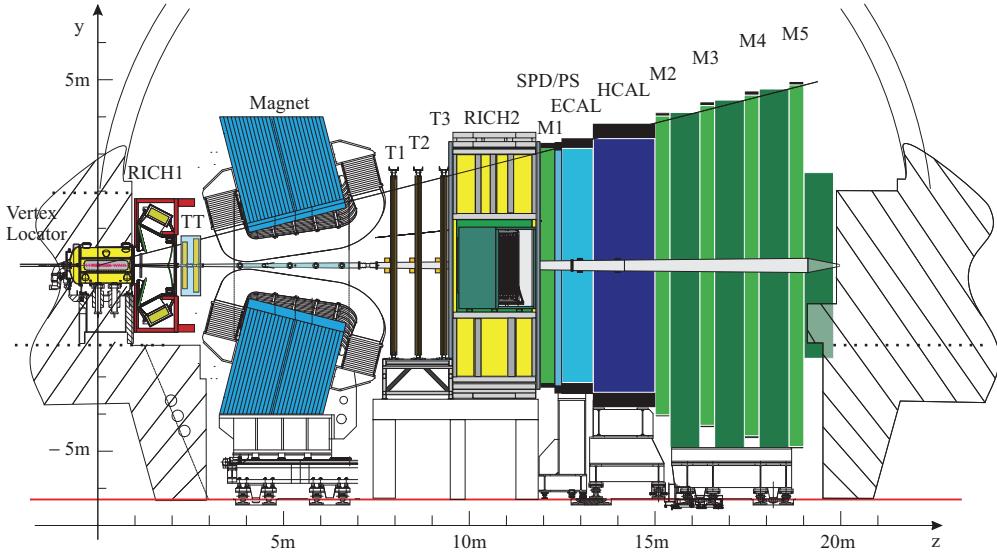


Figure 2.4 LHCb detector layout, showing the Vertex Locator (VELO), the dipole magnet, the two RICH detectors, the four tracking stations TT and T1-T3, the Scintillating Pad Detector (SPD), Preshower (PS), Electromagnetic (ECAL) and Hadronic (HCAL) calorimeters, and the five muon stations M1-M5 in the non-bending plane,  $y - z$  plane. The direction of the  $y$  and  $z$  coordinate axes are also shown; the  $x$  axis makes the right-handed coordinate system complete.

the reconstructed vertices. The VELO, shown in Figure 2.5, consists of two almost identical retractable detector halves and each half is equipped with 21 silicon strip tracking modules arranged along the beam. A silicon module is made up of two half disc sensors with azimuthally segmented strips ( $R$ -sensor) and radially segmented strips ( $\phi$ -sensor) respectively, and each sensor has 2048 strips and a radius of 42 mm and a thickness of 300  $\mu$ m. The  $R$ -sensor illustrated as the right half of the disk in Figure 2.6 provides a measurement of the radial distance of the hits from the beam axis, while the  $\phi$ -sensor illustrated as the left half the disk in Figure 2.6 measures the hit azimuthal coordinates around the beam. A cylindrical geometry ( $R - \phi$  geometry) is preferable as it permits faster reconstruction of tracks and vertices in the LHCb trigger compared to a rectilinear scheme<sup>[161]</sup>. The  $R - \phi$  sensors can measure the transverse hit position. Together with the knowledge of the position ( $z$  coordinate) of the sensors, they can provide 3-dimension coordinates of hits. There are two additional  $R$  sensor planes located upstream of the sensors called the pile-up veto system which allows to roughly determine the number of  $pp$  interactions for the trigger decision.

The strips of the VELO sensors were designed to optimize the vertex/track reconstruction quality and background suppression power. In order to minimize the per-strip

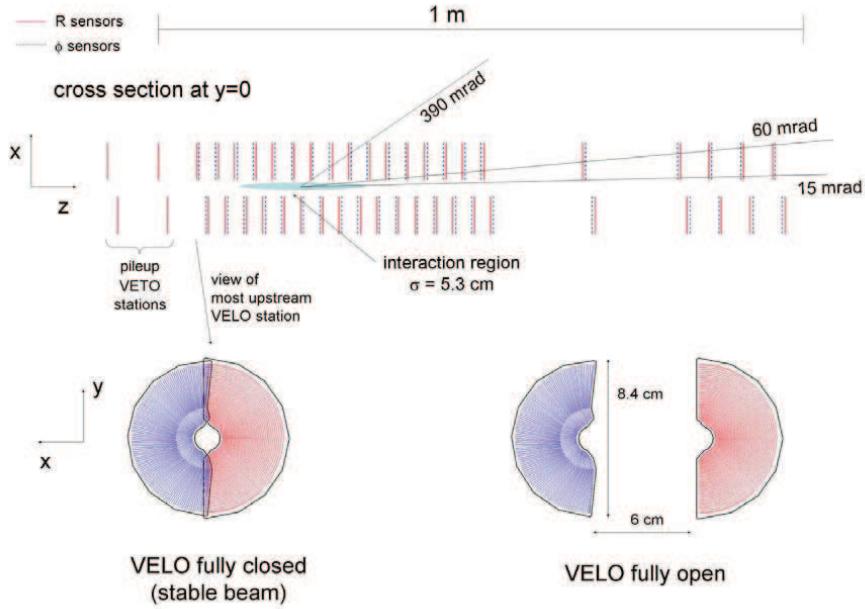


Figure 2.5 Cross section view of VELO silicon sensors in the  $(x,z)$  plane at  $y = 0$ , with the detector in the fully closed position. The front face of the first modules is also illustrated in both the closed and open positions. The two pile-up veto stations are located upstream of the VELO sensors.

occupancy and to reduce the strip capacitance, each  $R$ -sensor strip is subdivided into four  $\pi/4$  regions. The pitch at the innermost radius is about  $40 \mu\text{m}$  and increases linearly to about  $102 \mu\text{m}$  at the outermost radius, so that each VELO hit used to reconstruct the track contributes almost equally to the impact parameter precision. The  $\phi$ -sensors, to avoid high occupancies and too large strip width at the outer region, is subdivided into inner region and outer region separated at the radius of  $17.25 \text{ mm}$ . The size of the strips in various regions can be found in Figure 2.6. The strips in the  $\phi$ -sensors are also skewed to improve the pattern recognition, and the strips in successive  $\phi$ -sensor are skewed in opposite direction.

To protect the VELO from the damage by beam particles, the two VELO halves are retracted  $3 \text{ cm}$  (VELO open) from the beam during LHC injection. When the beams are stable, both halves of the VELO are moved towards the measured  $x - y$  position of the beam (VELO closed), reaching a distance of only  $7 \text{ mm}$  to the beam line<sup>[171]</sup> as can be seen by the cross section view of the VELO in Figure 2.5.

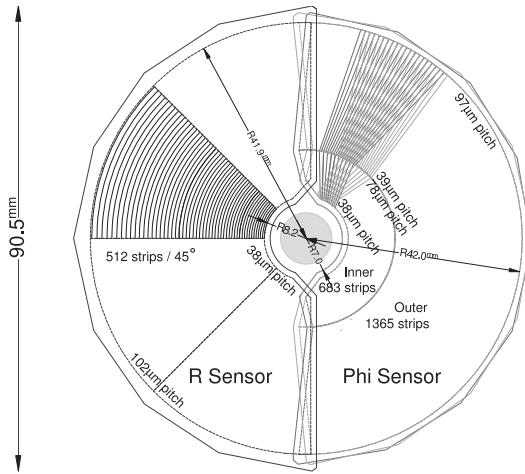


Figure 2.6 The  $R\phi$  geometry of the VELO sensors. A fraction of  $R$  strips in the  $R$  sensor (left half) and  $\phi$  strips in the  $\phi$  sensor (right half) are also illustrated. In the  $\phi$  sensor, the strips on two neighboring modules are included to highlight the stereo angle.

### 2.2.1.1 Performance of VELO tracking/vertexing

The VELO uses analogue readout, providing the hit position with resolution better than the strip size. The single hit resolution measures the fluctuations of the distances between the intercept of tracks with the sensors and the measured positions of corresponding VELO clusters. The resolution, shown in Figure 2.7, is from a few  $\mu\text{m}$  to tens of  $\mu\text{m}$  depending on the charge sharing among adjacent strips, and eventually depending on the size of the strip pitch and the projection angle, which is the angle between the track and the strip layout. As a result the primary vertex position resolution, as shown in Figure 2.8, is found to be about 10  $\mu\text{m}$  in the  $x - y$  direction and 50  $\mu\text{m}$  in the  $z$  direction for a typical PV with 40 tracks. The impact parameter (IP) resolution is below 20  $\mu\text{m}$  for the typical tracks with transverse momentum of a few  $\text{GeV}/c$  as can be seen in Figure 2.9. The PV and IP resolutions are a little bit worse in data than in the simulation, which can be due to the input parameters in Monte Carlo, for example the multiple scattering modeling and the material description<sup>[172]</sup>. With vertex resolution of about 200  $\mu\text{m}$  and flight distance of about 8 mm (in  $z$ -direction in two body decays),  $b$ -vertices can be clearly resolved.

### 2.2.2 The RICH detectors

For a precision measurement in charm and bottom physics, particle identification (PID) is a fundamental requirement for LHCb to suppress the background and to provide an efficient kaon tagging. The LHCb experiment uses Cherenkov radiation effect to identify

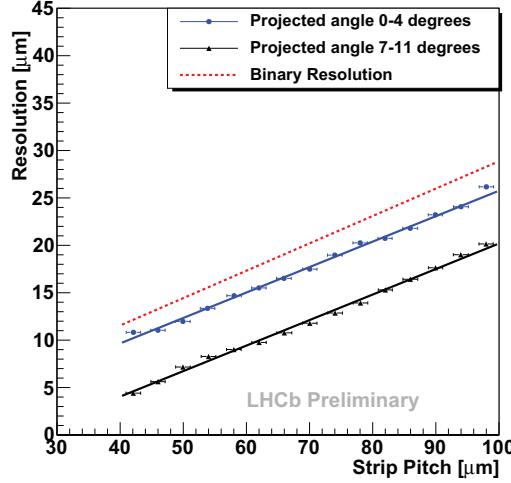


Figure 2.7 Single hit resolution of VELO sensors. Reproduced from reference [172].

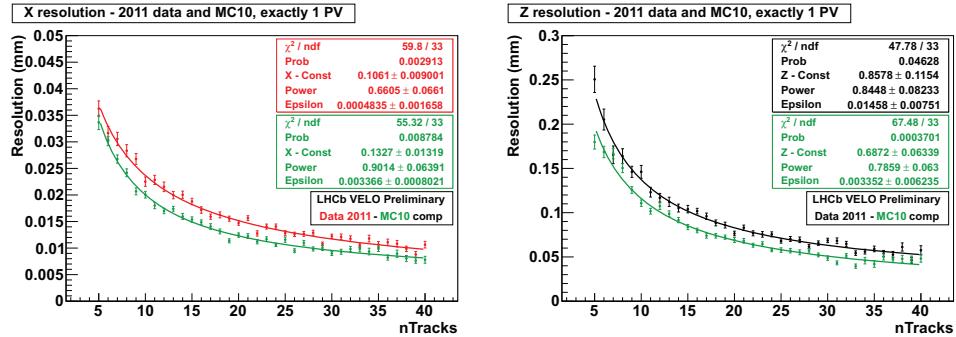


Figure 2.8 PV resolution in the  $x$  direction (left) and the  $z$  direction (right) as a function of number of tracks in the PV in real data (2011) and Monte Carlo (2010). Reproduced from reference [172].

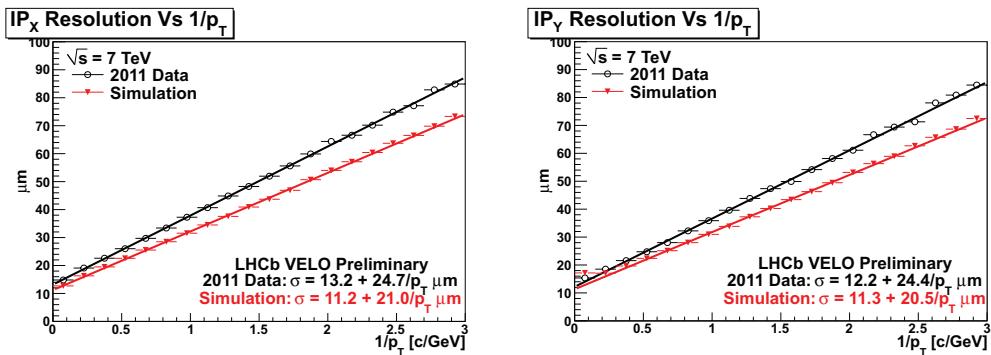


Figure 2.9 Impact parameter resolution in the  $x$  direction (left) and the  $y$  direction (right) as a function of the inverse of transverse momentum in real data (2011) and Monte Carlo (2010). The intercept with  $y$ -axis depends mainly on the single hit resolution, while the slope is determined by the differences of multiple scattering for different  $p_T$ . Reproduced from reference [172].

charged particles, predominately pions and kaons. The speed of light in the material with refractive index  $n$  is  $c/n$ . A charged particle traversing that material (radiator) with velocity  $v$  larger than  $c/n$  will radiate photons in a cone along the flight direction of the particle. The angle between the direction of radiation photons and the track momentum is the Cherenkov angle, which is:

$$\cos \theta_c = \frac{c}{nv}. \quad (2-1)$$

At LHCb the Cherenkov light radiated by charged particles is focused by a combination of spherical and flat mirrors and directed out of the detector acceptance where the photons are collected by Hybrid Photon Detectors (HPDs) (see Figure 2.10), which have high granularity ( $2.5 \times 2.5 \text{ mm}^2$  at the photo-cathode at the entrance window) and work at high speed. Because the photons are emitted symmetrically in the azimuthal direction around the trajectory of the particle, the photons collected by the HPDs will fall onto a ring—the Cherenkov ring. The size of the Cherenkov ring is a direct measure of the Cherenkov angle, which together with the material refractive index  $n$  will give a determination of the velocity. The estimated velocity and the momentum measured by the tracking system are used to give a hypothesis of the particle PID (mass).

LHCb has to discriminate pions and kaons in the vast momentum range from a few GeV to around 100 GeV and in the full angular acceptance. The refractive index  $n$  and the resolution of the HPDs determine the fiducial momentum range in which the types of hadrons can be discriminated as shown in Figure 2.11. To cover the large range of momentum, LHCb uses two Ring Imaging Cherenkov detectors (RICH) with different radiators. The RICH1 (RICH2) detector placed upstream (downstream) the main tracking system using aerogel and  $\text{C}_4\text{F}_{10}$  ( $\text{CF}_4$ ) radiators cover the low (high) momentum range  $\approx 1\text{--}60 \text{ GeV}/c$  (from  $\approx 15$  up to and beyond  $100 \text{ GeV}/c$ ). RICH1 has a wide acceptance of the full LHCb angular coverage from 25 mrad to 300 mrad (250 mrad) in the bending (non-bending) plane, while RICH2 has a limited angular acceptance from 15 mrad to 120 mrad (100 mrad) in the horizontal (vertical) plane but accepts most of the high momentum particles. To protect the HPD from the disturbance of the magnetic field, the RICH detectors are placed under an iron shield. The structures of the RICH detectors are shown in Figure 2.12.

During a typical event, as shown in Figure 2.13, there would be measurements (hits) by the RICH HPDs corresponding to real Cherenkov photons or noise. The hits falling

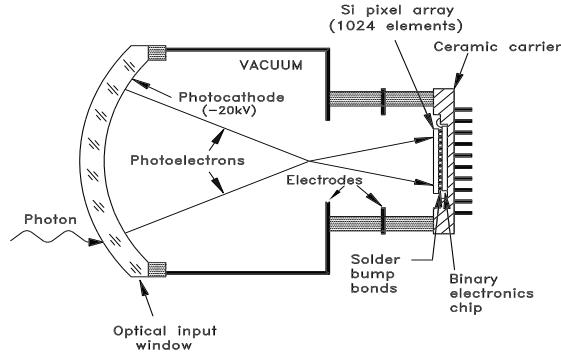


Figure 2.10 The detection of photons by HPD. The silicon photon-electron detector is segmented into 1024 pixels.

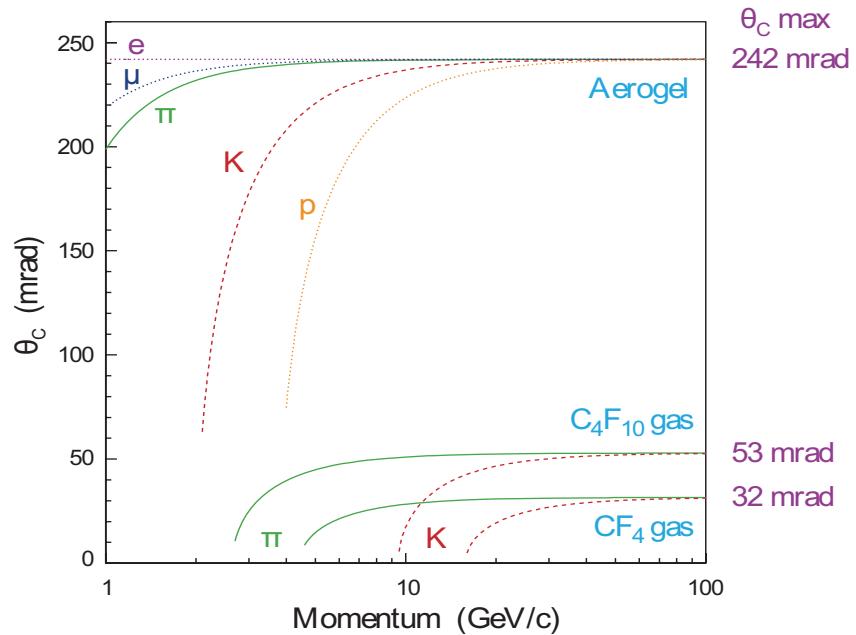


Figure 2.11 Cherenkov angles as a function of momentum for various charged particles and radiators.

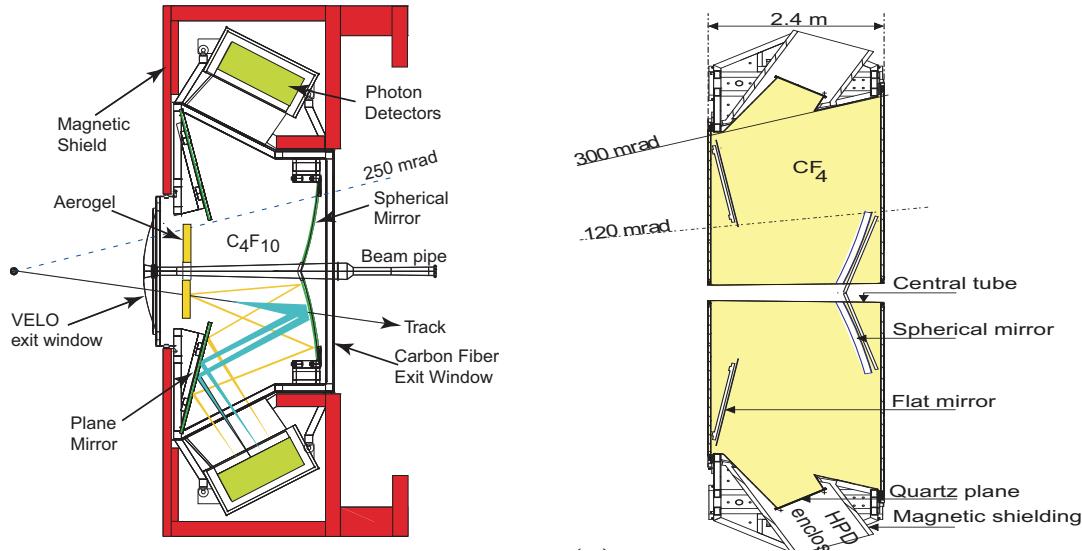


Figure 2.12 Schematic view of RICH1 (left) and RICH2 (right) detectors. The angular coverage, the spherical mirror, the plane mirror, the photon detectors and the magnetic shield are also illustrated.

onto the same ring are related to the same charged track. The coordinates of the passage of the charged particle through the LHCb detector including the RICH radiator volumes is provided by the tracking system. The emission point of each track is taken as the mid-point of the trajectory in the radiator. The candidate photons for each track are determined by combining the photon emission point with the measured photon hit positions. Once the photon candidates have been assigned, the Cherenkov angle can be computed. The Cherenkov angle resolution is measured to be 1.62 mrad for the RICH1  $C_4F_{10}$  gas, 5.6 mrad for the aerogel gas and 0.68 mrad for the RICH2  $CF_4$  gas, allowing very good pion/kaon separation at high momentum range (see Figure 2.14).

### 2.2.2.1 RICH PID and performance

Each charged track is assigned various PID hypothesis ( $e, \mu, \pi, K$  or  $p$ ) and for each hypothesis a likelihood ( $\mathcal{L}$ ) is computed from the quality of the matching between the track and the associated photon hits. The change of the (log-)likelihood ( $\Delta \log \mathcal{L}$ ) when the mass hypothesis is changed from one type of particle hypothesis to another can be used to identify particle types. As demonstrated in Figure 2.15, the particle identification by the RICH-system using the  $\Delta \log \mathcal{L}$  method will give reasonably high kaon efficiency (kaons identified as kaons) and keeping the pion misidentification (pions misidentified as kaons) fraction under control. For each track, requiring the likelihood with the kaon hypothesis

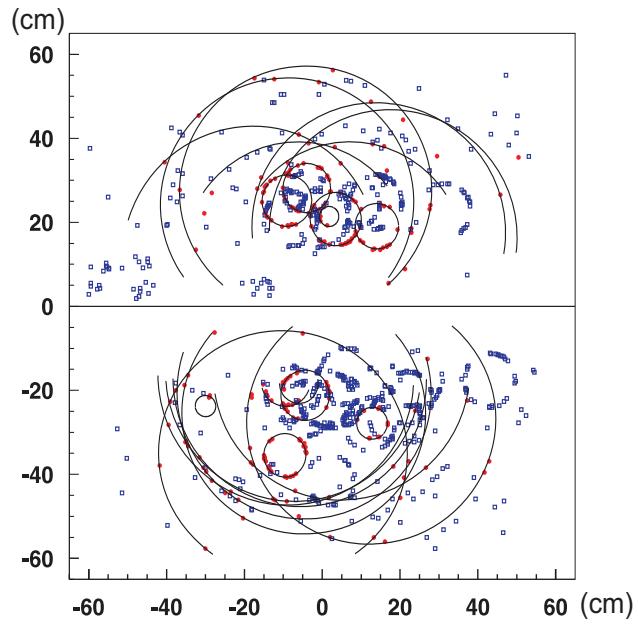


Figure 2.13 Display of RICH measurements in a typical LHCb event.

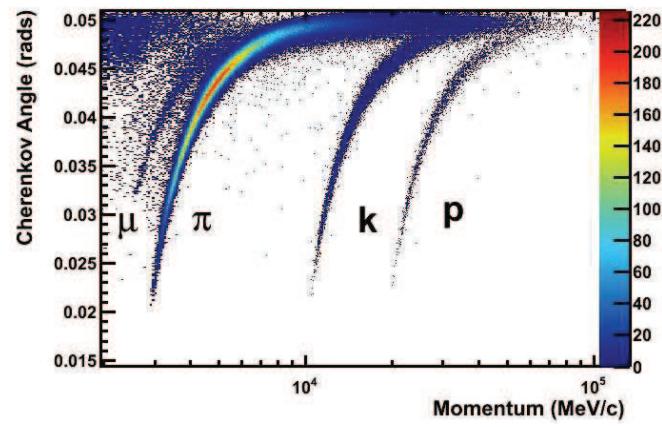


Figure 2.14 Reconstructed Cherenkov angle with respect to the track momentum for various particles in the  $\text{C}_4\text{F}_{10}$  radiator. The scattered bands for each particle corresponds to the resolution.

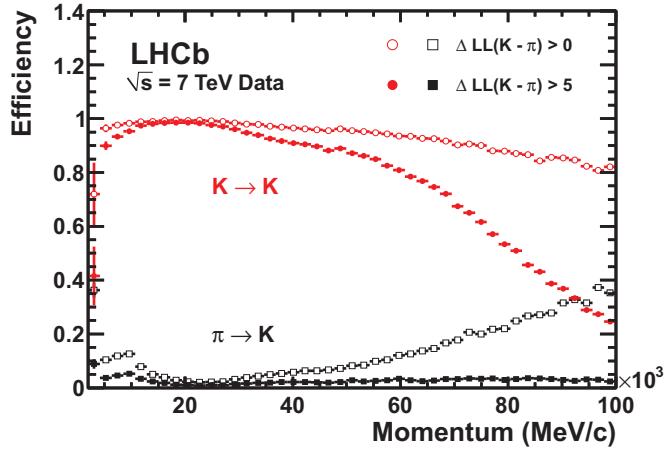


Figure 2.15 The kaon identification efficiency and pion misidentification rate in data as a function of track momentum. Two different  $\Delta \log \mathcal{L}_{K/\pi}$  requirements have been used.

to be larger than that with the pion hypothesis, which is  $\Delta \log \mathcal{L}_{K/\pi} > 0$ , and averaging over the momentum range 2 - 100 GeV/c, the kaon efficiency (pion misidentification fraction) are found to be  $\approx 95\%$  ( $\approx 10\%$ ).

### 2.2.3 The main tracking detectors

The main tracking system consists of the dipole magnet, the Tracker Turicensis (TT) located upstream of the magnet and the Tracking stations (comprised of IT and OT) installed downstream of the magnet. The primary goal of the tracking system is to supply efficient reconstruction of charged particle trajectories and precise measurements of their momenta. Besides, the reconstructed tracks are also used in the RICH ring reconstruction and the trigger decisions.

#### 2.2.3.1 The dipole magnet

The LHCb experiment uses a warm dipole magnet to provide magnetic field for the momentum measurement of charged particles in the forward angular range of  $\pm 250$  mrad vertically ( $\pm 300$  mrad horizontally). The yoke and coil geometry of the magnet is shown in Figure 2.16. The magnet consists of two coils placed face to face in almost the  $x - z$  plane and bent at  $45^\circ$  on the two transverse sides. The arrangement of coils and yokes produces magnet field mainly in the  $y$  direction. The current in the coil, with a nominal value of 5.85 kA, can be switched to make the direction (polarity) of the magnetic field inverted, from MagDown ( $B_y < 0$ ) to MagUp ( $B_y > 0$ ) or vice versa. The integrated

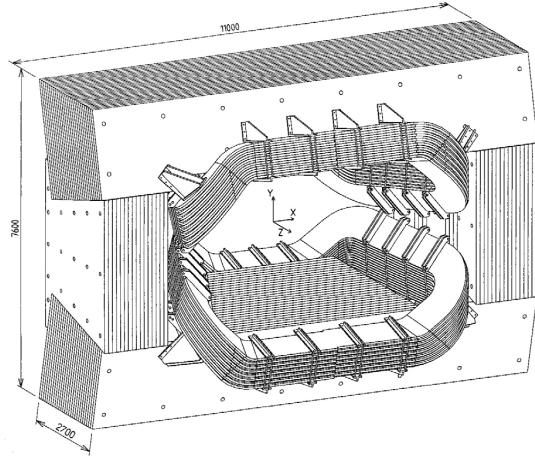


Figure 2.16 The perspective view of LHCb dipole magnet. The LHCb coordinate system is also shown. The dimensions are in mm.

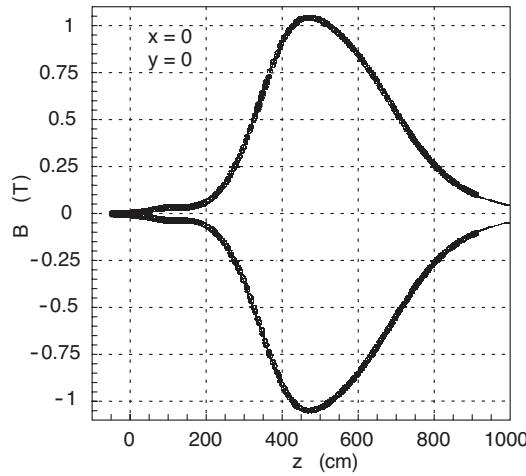


Figure 2.17 Magnetic field along the  $z$  axis at  $x = y = 0$  for MagUp and MagDown respectively.

magnetic field  $\int B dl$  for tracks of 10 m length reaches 4 Tm. The magnetic field ( $4 \times 10^{-4}$  relatively) has been precisely measured in all the tracking volumes—inside the magnet, in the main tracking detectors, and also in the region of the VELO and inside the magnetic shielding for the RICH1 and RICH2 photon detectors. The measured field  $B_y$  at different  $z$  coordinates with coordinates  $x = y = 0$  is shown in Figure 2.17.

### 2.2.3.2 The silicon tracker

The silicon tracker (ST) comprises the TT and IT, which are made of silicon microstrip sensors with a strip pitch of about 200  $\mu\text{m}$ . The TT, 150 cm wide 130 cm high in dimension, is a planar tracking station covering the full acceptance of the experiment. The IT, placed in the center of the three tracking stations, covers a 120 cm wide and 40 cm

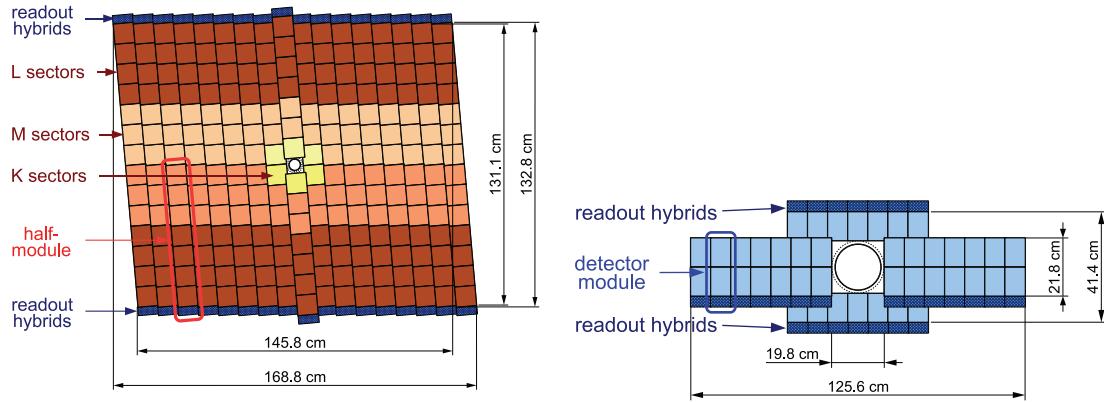


Figure 2.18 Layout of the third (v) TT detection layer (left) and IT detection layer (right). The dimensions of the layers in  $x - y$  plane are also shown.

high cross shaped region downstream of the magnet. Every ST station has four layers ( $x - u - v - x$ ), and each layer is segmented vertically, but with the strips in the second and the third layers rotated by a stereo angle of  $-5^\circ$  and  $+5^\circ$  respectively.

The active area of the TT covers the nominal acceptance of the LHCb spectrometer, 300 mrad in the horizontal bending plane and 250 mrad in the vertical plane. The layout of one the four TT layers is illustrated in the left plot of Figure 2.18. It is divided into two halves, upper half and lower half with each consisting of a row of seven silicon sensors (modules) organized into two (L and M sectors) or three (L, M and K sectors for modules close to the beam pipe) readout sectors. A silicon sensor is 500  $\mu\text{m}$  thick and about  $10 \times 10 \text{ cm}^2$  in  $x - y$  dimension, carrying 512 readout strips with each strip pitch around 200  $\mu\text{m}$ .

Each of the three IT stations, as shown in the right plot of Figure 2.18, contains four individual detector boxes arranged around the beam pipe, and each box contains four detection layers with each layer consisting of seven modules. One module has one (for boxes above and below the beam pipe) or two (for boxes on the right and left side of the beam pipe) sensors/readout hybrids. Each sensor carries 384 readout strips with a strip pitch of around 200  $\mu\text{m}$  and thickness of 320  $\mu\text{m}$  (410  $\mu\text{m}$ ) for the one-sensor modules (two-sensor modules).

Adjacent modules within one detection layer of the ST are staggered by a few mm in  $z$  and overlap by a few mm in  $x$  to avoid acceptance gaps and help in the alignment of the modules.

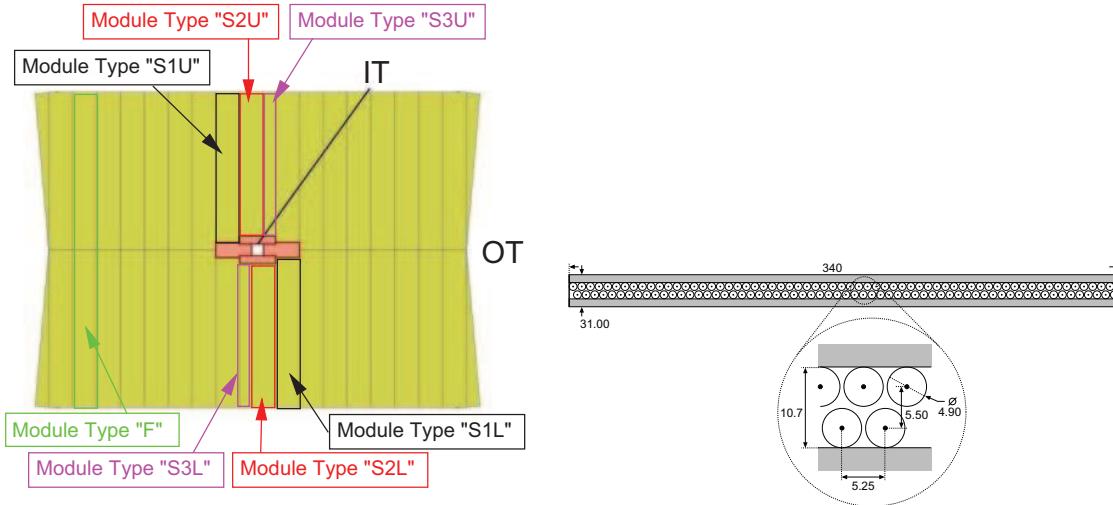


Figure 2.19 Layout of the a OT layer with various sizes of modules (*left*) and the cross section view of a straw-tubes module (*right*). The IT arrangement is also shown on the left, and a zoom view of arrangement of the tubes is also shown on the right.

### 2.2.3.3 The outer tracker

The out tracker (OT) consists of drift-time detectors placed at the same coordinates as the ITs, but covers the outer regions (up to LHCb outer acceptance boundaries) in the  $x - y$  plane and thus can accept particles with lower momenta or higher angles. The OT is designed as arrays of straw-tube modules with different dimensions, as shown in Figure 2.19. Each module contains two staggered layers of drift-tubes with inner diameters of 4.9 mm. The detector modules are separated into three stations, T1, T2 and T3. Each station consists of four layer of modules, arranged also in  $x - u - v - x$  layer geometry with tubes in the  $x$  layers installed vertically while the tubes in the  $u, v$  layers tiled by  $\pm 5^\circ$  with respect to the vertical direction. The tubes are filled with drift gas which is a mixture of Argon and CO<sub>2</sub> to guarantee a fast drifting time (below 50 ns) and sufficient drift-coordinate resolution (200  $\mu$ m). The signals collected by the tubes are read out from the outer end.

### 2.2.3.4 The LHCb tracking

Charged particles traversing the LHCb detector will generate hits in the VELO and the main tracking detectors. The hits are used to reconstruct the trajectories of the particles. Depending on their paths through the spectrometer, several track types are defined, as illustrated in Figure 2.20:

- **Long tracks** traverse the full tracking system. They have hits in both the VELO and the T stations, and possibly also in the TT. Because they traverse through the full tracking systems including the magnetic field, they have the most precise momentum estimation and therefore are the most important set of tracks for the physics analysis.
- **Upstream tracks** traverse only the VELO and TT stations. In general they are low momentum tracks and are bent outside of LHCb (OT) acceptance by the magnet. However, they pass through the RICH1 detector and will generate Cherenkov photons if they have sufficient velocity. The upstream tracks are therefore used to understand backgrounds in the particle-identification algorithm of the RICH1 detector.
- **Downstream tracks** traverse only the TT and T stations. The long lived particles ( $K_S^0$ ,  $\Lambda$ , etc.) will produce secondary particles of this kind when they decay outside of the acceptance of the VELO sensors.
- **VELO tracks** traverse only through the VELO detector and are typically large angle or backward tracks, which are useful for the primary vertex reconstruction.
- **T tracks** traverse only through the T stations. They are typically produced in secondary interactions, but are still useful for the global pattern recognition in the RICH2 detector.

The first step of track reconstruction is to search for the track ‘seeds’—the track segments—as candidates for tracks. A seed can be reconstructed in the VELO region (VELO-seeds) and the T stations (T-seeds) because the magnetic field in these regions is low and a search for almost straight line segments are possible.

The long track reconstruction starts with a search for VELO seeds. Then, there are two algorithms to promote these VELO tracks (segments) to long tracks. In the first algorithm, a VELO seed is combined with a single T-station hit, and a rough guess of the track’s trajectory and also momentum can be given. Around this trajectory, further hits in the T-stations are then searched. When enough hits have been collected for the candidates, a long track candidate is found. In the second algorithm, called track matching, the VELO seeds are combined with T-seeds by requiring that they have position and slope parameters matched.

The downstream tracking algorithm starts with T-seeds, extrapolates them through the magnetic field and searches for corresponding hits in the TT. Upstream tracks are

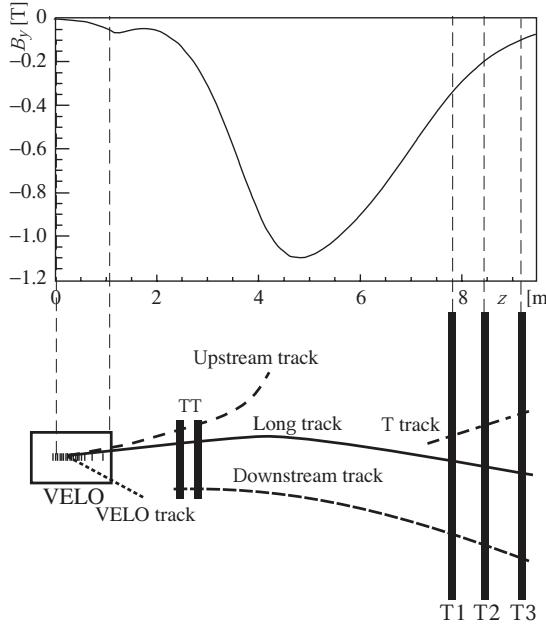


Figure 2.20 Illustration of the LHCb track types: long tracks, upstream tracks, downstream tracks, VELO tracks and T tracks. The main  $B$ -field component ( $B_y$ ) is plotted above as a function of the  $z$  coordinate at  $x = y = 0$ .

found by extrapolating the VELO seeds to the TT, selecting and adding the hits in TT to the track candidates.

Finally, tracks which share many of the same hits are removed (clone killing) while the remaining tracks are fitted using a Kalman fitter<sup>[173]</sup>. The fitting procedure retrieves the LHCb detector material description in order to take into account the multiple scattering and to correct for energy loss  $dE/dx$  due to ionization. The  $\chi^2$  of the fit can be used to monitor the quality of the reconstructed tracks. A typical event with reconstructed tracks and assigned hits are shown in Figure 2.21.

### 2.2.3.5 Performance of LHCb tracking

The performance of the track reconstruction is expressed by mainly the efficiency of the tracking finding and the precision of the reconstructed momentum.

The efficiency expresses the ability to reconstruct trajectories of charged particles that have passed through the tracking system. The (long track) efficiency can be measured on two body decays ( $J/\psi \rightarrow \mu^+ \mu^-$  for example) using a tag-and-probe technique, in which one of the daughter particles, the "tag" leg, is fully reconstructed, while the other particle, the "probe" leg, is only partially reconstructed. The tracking efficiency is determined

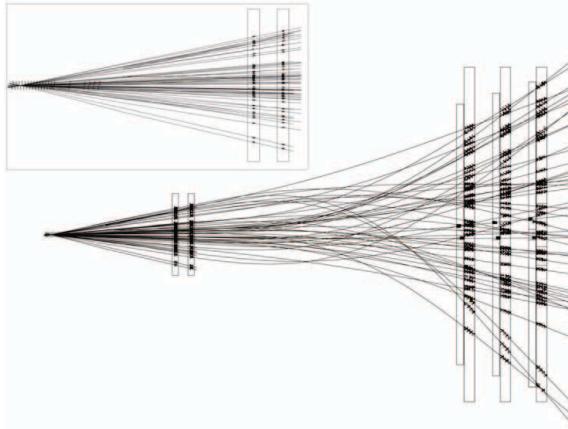


Figure 2.21 Display of the reconstructed tracks and assigned hits in each tracking detector. The insert shows a zoom into the VELO and TT region.

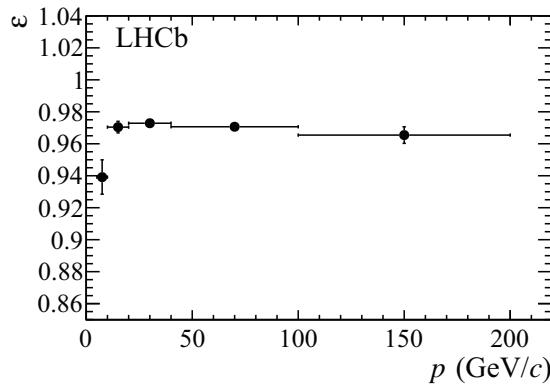


Figure 2.22 Long tracking reconstruction efficiency as a function of the track momentum in data (2011).

by the probability of the matching between the partially reconstructed probe leg to a fully reconstructed long track. The long track reconstruction efficiency depends on the momentum of the tracks as shown in Figure 2.22. The average efficiency for 2011 data taking is above 96% fulfilling the goal of the design of LHCb tracking system.

The mass resolution is an important parameter in the discrimination of signals and suppression of background events. Good mass resolution is achieved by fine momentum resolution of the tracks. The relative long track resolution  $\delta p/p$  is shown in figure 2.23 as a function of the momentum. From the figure it can be seen that, the momentum resolution is about 0.5% for particles below 20 GeV and about 1.0% for particles with momentum above around 150 GeV.

In Figure 2.24 the mass resolution for the various dimuon channels are shown. It can

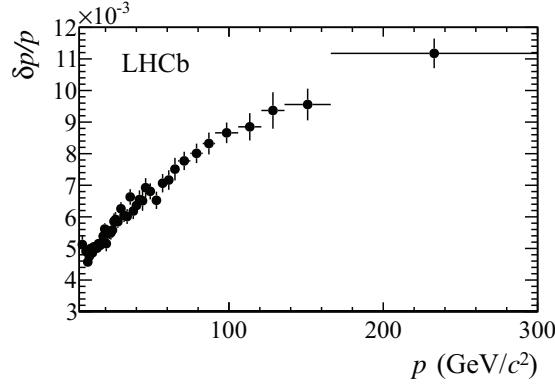


Figure 2.23 Relative momentum resolution as a function of momentum for long tracks in data (2011).

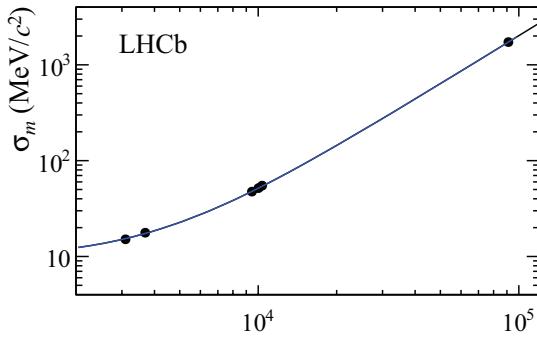


Figure 2.24 The mass resolution for  $J/\psi$ ,  $\psi(2S)$ ,  $\Upsilon(nS)$  and  $Z^0$  in the  $\mu^+\mu^-$  decay channel.

be seen that the relative mass resolution,  $\sigma_m/m$ , is about 0.5% up to the  $\Upsilon(nS)$  masses for the two body final state channels.

## 2.2.4 The LHCb calorimeter

The main purpose of the LHCb calorimeter is to identify electrons and hadrons and to provide measurements of their energies and positions, which are required for various trigger algorithms and offline physics analysis. Furthermore, the calorimeter is essential to reconstruct neutral particles (photons,  $\pi^0$ , neutrons etc.) for interesting physical channels. The LHCb calorimeter, covering the full LHCb acceptance consists of the Scintillating Pad Detector (SPD), the Preshower (PS), the Electromagnetic (ECAL) and Hadronic (HCAL) calorimeters placed one after another behind the tracking systems and before the last four muon stations.

The most important purpose of the calorimeter is to identify electrons for the first

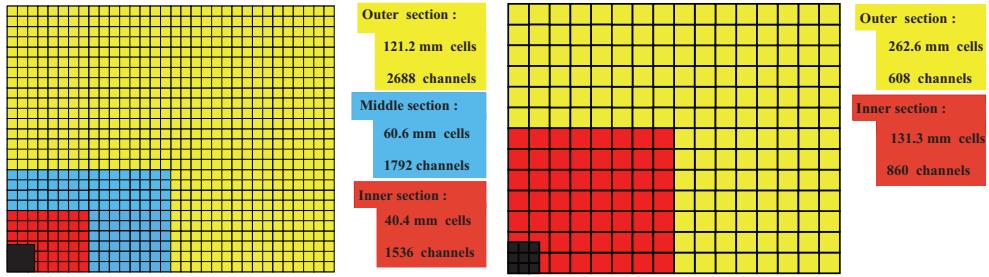


Figure 2.25 Segmentation of the SPD/PS and ECAL (*left*) and the HCAL (*right*) in the  $x - z$  plane. One quarter of the detector front face is shown.

level electron trigger, which is required to reject 99% of the inelastic  $pp$  interactions while providing an enrichment factor of at least 15 in  $b$  events. This is accomplished by the selection of electrons with large transverse energy  $E_T$ . To reject the high hadronic background of charged pions, the electromagnetic shower has to be segmented longitudinally, which can be fulfilled by a PS followed by the main section of the ECAL. The electron trigger can also be polluted by the background of  $\pi^0$ s with high  $E_T$ . Such contamination can be rejected by the introduction, in front of the PS, of the SPD plane used to select charged particles but transparent to the photons. The thickness of the ECAL (HCAL) is chosen to be 25 radiation lengths  $X_0$  (5.6 interaction lengths  $\lambda_l$ ).

Since the hit density varies fiercely from the inner region to outer region over the calorimeter surface, the PS/SPD, ECAL and HCAL adopt a variable lateral segmentation as shown in Figure 2.25. The ECAL is segmented into three different sections and the SPD/PS is projectively granulated with respect to the ECAL. Because the dimensions of the hadronic showers are quite large, the HCAL is segmented into two zones with larger cell sizes compared to the ECAL.

The primary signals from the calorimeter showers are scintillation photons, which are then collected and transmitted to the Photo-Multiplier (PMT) readouts by wavelength-shifting (WLS) fibres. The electric signals from the PMT are processed for further use, the trigger for example.

The SPD and PS detectors are two almost identical planes of 15 mm thick rectangular scintillator pads of high granularity with a 15 mm ( $2.5 X_0$ ) thick lead converter placed parallel in between.

The ECAL is a sampling scintillator/lead structure readout by plastic WLS fibres. In total, the ECAL contains 66 scintillator/lead layers, with a 4 mm thick scintillator and a 2

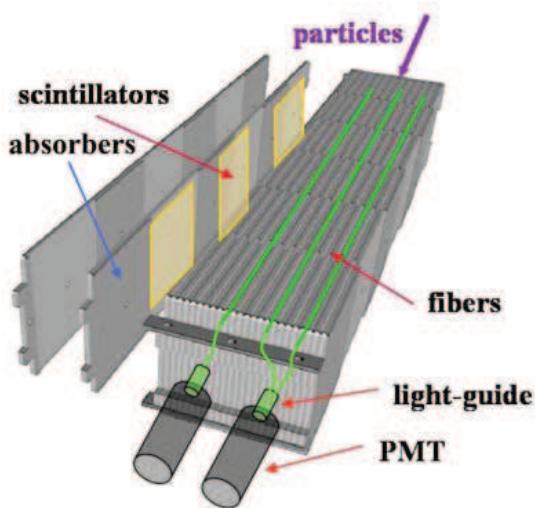


Figure 2.26 A schematic of the HCAL cell structure. The exploded view of two scintillator-absorber layers illustrates the elementary periodic structure of a HCAL module.

mm thick lead plate in each layer. The ECAL give a measure of the shower energy with relative resolution  $\delta E/E = [(8 - 10)/\sqrt{E} \oplus 0.9]\%$  ( $E$  in GeV)<sup>[174]</sup>.

The HCAL is also a sampling device made from iron and scintillating tiles as absorber and active material respectively. The 3 mm thick scintillating tiles are placed parallel to the beam axis as shown in Figure 2.26. In the lateral direction, the tiles are interspersed with 1 cm thick iron plates. The longitudinal length of scintillators and iron absorbers corresponds to the hadron interaction length  $5.6 \lambda_l$  in steel. The lights in these HCAL structures are collected by WLS fibres along the detector towards the back side where PMTs are housed. The energy of the shower given by the HCAL has a relative resolution of  $\delta E/E = (69/\sqrt{E} \oplus 9)\%$  ( $E$  in GeV)<sup>[174]</sup>.

#### 2.2.4.1 Performance of the LHCb calorimeter

Energy deposits in ECAL cells are clusterized applying a  $3 \times 3$  cell pattern around a local maximum which has the largest energy deposit compared to its direct neighboring cells. If one cell is shared among several reconstructed clusters, the energy of the cell is redistributed among the clusters proportionally to their total cluster energies. After the clustering, the total energy, the barycenter position and spread of the cluster can be calculated. The cluster energy is transformed into the particle energy with correction on the leakages. The energy resolution of the electromagnetic shower results in a  $B$  mass resolution of about 40 MeV/ $c^2$  for the  $B_s^0 \rightarrow J/\psi \eta(\gamma\gamma)$  decay<sup>[175]</sup> with high- $E_T$  photons,

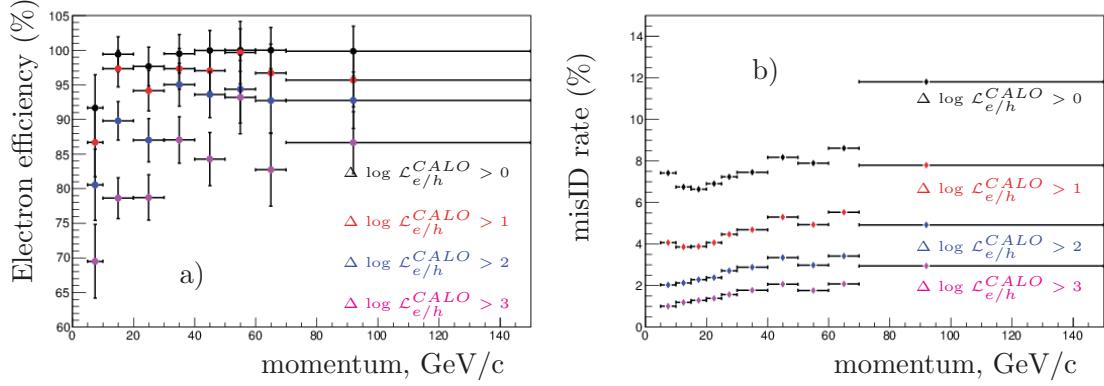


Figure 2.27 Electron PID efficiency (left) and pion misID rate into electron (right) as a function of the track momentum.

and the resolution of  $\pi^0$  is measured to be around  $8 \text{ MeV}/c^2$ .

The electron particle identification is done mainly by the calorimeter system relying on the information derived from the ECAL, SPD/PS and HCAL. The major ECAL estimator of electron ID uses the quality of the matching between the tracks and the energy deposit cluster. The electron ID can be refined with the properties of the calorimeter shower. The ratios of energy deposit of electrons in the ECAL over the momentum of the electron track are centered around unity, while hadrons only leave a fraction of their energy in the ECAL. Electrons will also deposit relatively more (less) energy in the P-S (HCAL) detector than the hadrons. Based on these factors, the difference between log-likelihood of electron hypothesis and hadron hypotheses of the matched track-cluster can be extracted. As can be seen in Figure 2.27, by using only the calorimeter information, the electrons have very high identification efficiency (above 95% for tracks larger than 20 GeV), while the  $\pi$  mis-identification rate is kept as low as 5% after requiring the  $\Delta \log \mathcal{L}_{e/h}^{\text{CALO}} > 2$ . The electron hadron discrimination can be refined when the RICH PID is also included.

## 2.2.5 The Muon system

Muons are of crucial importance in flavor physics, as they are present in the final state of many interesting  $b$ -physics decay channels, for example the decay of  $b$ -hadron into  $J/\psi$  and the rare decay of  $B_s^0$  to muon pairs. The channels involving muons will almost always mean low background contamination compared to purely hadronic decays. Triggers on high  $p_T$  muons will significantly reject the QCD background  $pp$  interactions but enrich

the fraction of charm/bottom events. The LHCb experiment uses the muon stations to provide reconstruction of muons used for online triggering and offline muon identification.

The LHCb muon system, shown in Figure 2.28 consists of five muon stations of rectangular shape, M1-M5, with M1 placed in front of the calorimeter and the other four placed behind the calorimeter. The muon stations covers the angular acceptances between 16 (20) and 258 (306) mrad in the non-bending (bending) plane. Each M2-M5 station is followed by a 80 cm thick muon absorber to select penetrating muons. The muon detectors provide space measurements of the tracks, which is binary (yes/no) information, to the trigger processor and to the Data Acquisition (DAQ) system. The muon position information is obtained by partitioning the detector into rectangular logical pads whose dimensions define the  $x - y$  resolution. Due to the variation of the flux density from inner region to outer region, each muon station is divided into four rectangular regions, R1-R4, with increasing distance from the beam pipe, and decreasing granularity (increasing logical pad size) as shown in Figure 2.29. The dimensions of the pads are chosen such that their contribution to the transverse momentum resolution (estimated by muon station alone) is approximately equal to the multiple-scattering degradation, which increases with the track polar angles.

The muon stations use the ionization effect to detect muons. Multi-wire proportional chambers (MWPC) are used for all regions in the five stations except the inner region of M1, where triple-GEM detectors are used for the consideration of aging facing the dense particle flux in M1 inner area. In stations M2 to M5 the MWPCs consist of four equal gas gaps superimposed, and two contiguous gas gaps have their readout in logic OR to form a double gap layer, resulting in two independent readouts. While in station M1, the MWPCs have only two gas gaps with independent readout to minimize the material in front of the electromagnetic calorimeter. In region M1R1 two superimposed triple-GEM (Gas Electron Multiplier) chambers are used in the logic OR state. The gas used for the MWPC and GEM are mixture of Ar/CO<sub>2</sub>/CF<sub>4</sub> with carefully selected proportions for a better time resolution (a few ns).

### 2.2.5.1 The Muon ID and Performance

High muon-identification efficiency while keeping misidentification from pions low is essential for rare decays, such as  $B_s^0 \rightarrow \mu^+ \mu^-$ . Muons are identified at LHCb by ex-

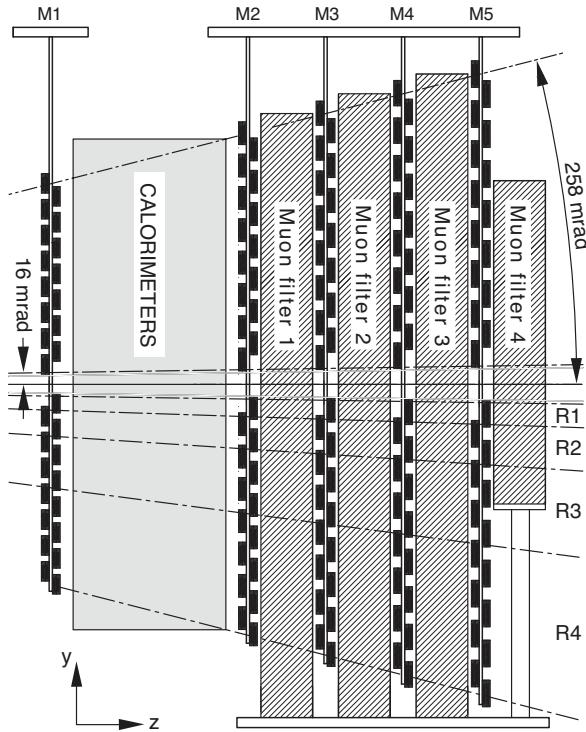


Figure 2.28 Side view of the LHCb muon system layout. The acceptance of the muon stations is also shown.

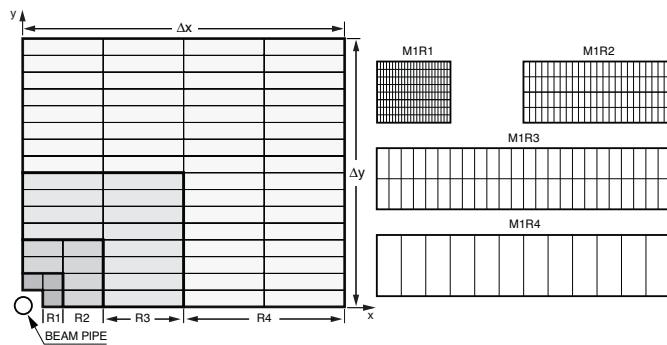


Figure 2.29 Front view of muon chambers (illustrated as small rectangular) in the four regions for a quadrant of a muon station (left) and the logical pad segmentation of chambers in each of the four regions of M1 (right). In each corresponding region of stations M2-M3 (M4-M5) the number of pad columns per chamber is two times (one half of) the number in M1.

extrapolating well reconstructed tracks into the muon stations. The tracks are extrapolated linearly starting from a track state at a  $z$  position downstream of the magnet. A track is identified as a muon (tagged as *IsMuon*) if hits in the muon stations are found inside the rectangular fields of interest (FOI) centered around the track extrapolation. The lateral dimensions of the FOI at each station are parameterized depending on the track momentum and are different for the four muon system regions in order to maximize the muon ID efficiency and keep the mis-ID rate tolerable. Since the total absorber thickness, including the calorimeters, is approximately 20 interaction lengths, the minimum momentum for a muon to go through the five stations is approximately 6  $\text{GeV}/c$  (or 3  $\text{GeV}/c$  for muons passing through only M1-M3), so the muon stations that are required to have corresponding hits for a potential muon track depend on the track momentum:

- M2+M3 for  $p < 6 \text{ GeV}/c$
- M2+M3+M4 or M5 for  $6 < p < 10 \text{ GeV}/c$
- M2+M3+M4+M5 for  $p > 10 \text{ GeV}/c$

The signal channel  $J/\psi \rightarrow \mu^+ \mu^-$  is used to study the muon identification efficiency, and the channels  $\Lambda \rightarrow p^+ \pi^-$  and  $D^0 \rightarrow K^- \pi^+$  are used to investigate the misidentification rates of protons, pions and kaons into muons. The (mis)identification efficiency is calculated as the fraction of event passing the *IsMuon* requirement. In Figure 2.30, the efficiencies as a function of the track momentum for different  $p_T$  ranges are shown. The muon ID efficiency is weakly dependent on momentum and transverse momentum, being always above 95% for typical muon tracks from  $b$  decay with  $p_T > 1.0 \text{ GeV}/c$  and momentum of tens of  $\text{GeV}/c$ . There are mainly two reasons for a hadron to be misidentified as a muon, firstly, the accidental coincidence of background hits or the hits produced from a real muon in the muon stations can be matched to the hadron track occasionally; secondly, hadron decays in flight will produce a real muon that will probably match the initial track. The misID of protons is only due to the first reason, resulting a misID rate quite below 1% for sufficiently large  $p_T$  and  $p$ . The misID of pions/kaons into muons is due to a combination of the two causes and the rate is below 1% for tracks with momentum above 20  $\text{GeV}/c$ . The decreasing of muon ID and hadron misID rate as a function of momentum is because the dimensions of the FOI decrease with the momentum ( $\text{FOI} = a + b \exp^{-cp}$ ).

The muon identification can also be improved by including the identification from RICH and Calorimeters to form the combined likelihood. From the combined likelihood,

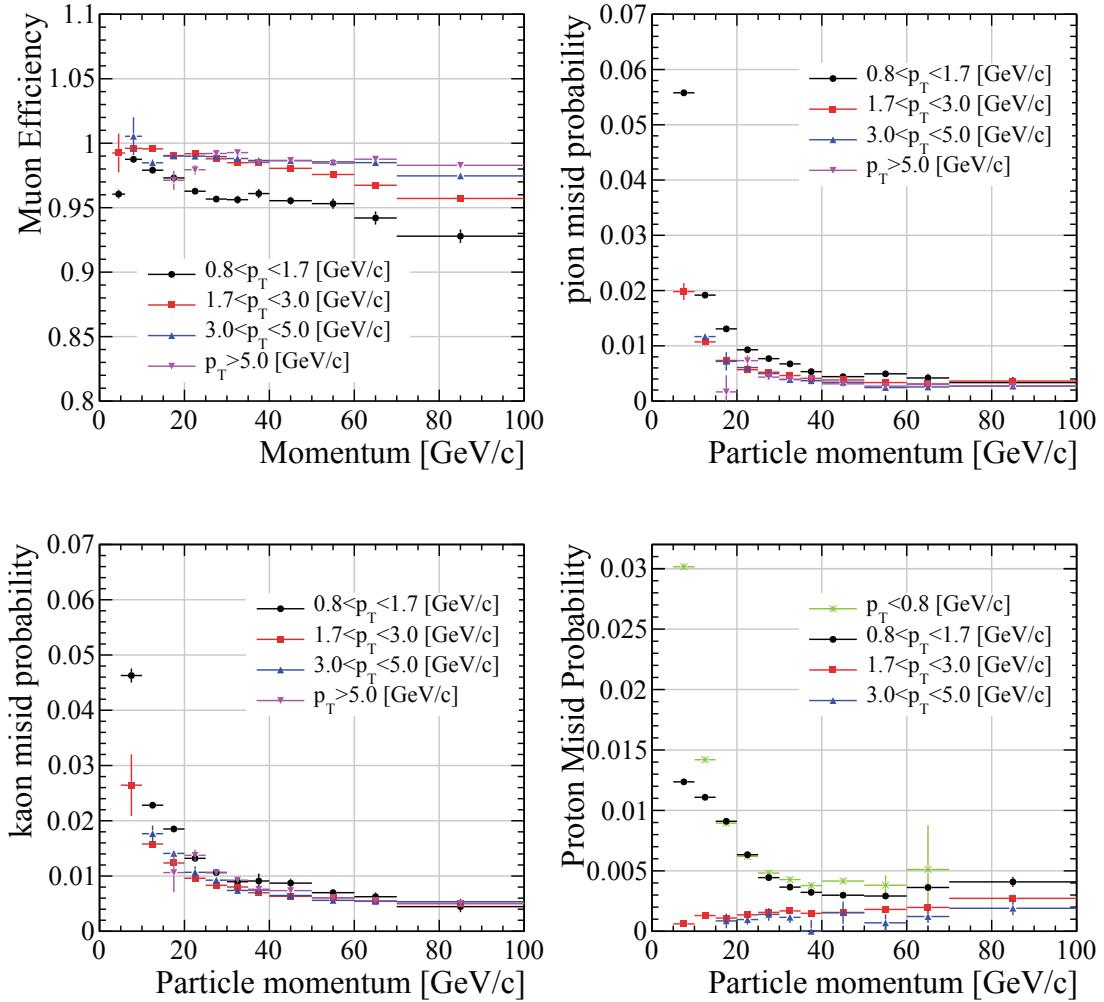


Figure 2.30 The efficiency of muon identification (*top left*), the misidentification rate of pions into muons (*top right*), the misidentification rate of kaons into muons (*bottom left*) and the misidentification rate of protons into muons (*bottom right*) as a function of the track momentum for different transverse momentum ranges in data (2011).

the muon hypothesis against other particle hypothesis ( $\Delta \log \mathcal{L}$ ) can be calculated for offline selections.

### 2.2.6 The LHCb trigger

The LHCb experiment is designed to operate at an average peak luminosity of  $2 \times 10^{32} \text{cm}^{-1}\text{s}^{-1}$ . With this low luminosity compared to the ATLAS and CMS and the LHC bunch structure (25 ns spacing), the frequency with visible interactions—interactions that have at least two charged particles with sufficient hits in the VELO and T-stations to be reconstructable as long tracks—is about 10 MHz, which has to be reduced to 2 kHz by the trigger in order to be sent to the storage for offline analysis. With a cross section of 0.5 mb, the peak luminosity  $2 \times 10^{32} \text{cm}^{-1}\text{s}^{-1}$  corresponds to about 100 kHz of  $b\bar{b}$  events, within which only 15% will have at least one  $B$  meson with all its decay products falling into the LHCb acceptance. The fraction of events where all the final state tracks of one  $b$  can be reconstructed is even smaller. At the same time, the branching fractions of interesting  $b$  decays are typically less than  $10^{-3}$ . Besides, the offline analysis always looks for some characteristics of the  $b$  events to enrich the signal over background. So within the 2kHz bandwidth the LHCb trigger can help to achieve the highest efficiency for the events selected in the offline analysis while reject on the uninteresting background as strongly as possible. The LHCb trigger consists of two trigger levels, the Level-0 (L0) and the High Level Trigger (HLT), as shown in Figure 2.31. The L0 is a hardware trigger implemented on the specially made electronics, operating synchronously with the bunch crossing, while the HLT is a software trigger executed on processor farms.

#### 2.2.6.1 L0 Trigger

L0 is a fast trigger implemented in hardware, based on the calorimeter and the muon systems. It reduces the rate of bunch crossings with interactions to below 1.1 MHz at which the whole detector information can be read out. L0 is divided into three independent triggers, the L0-Calorimeter trigger, L0-Muon trigger and the L0-PileUp trigger. The first two triggers use the  $b$  decay features—large transverse momentum ( $p_T$ ) and energy ( $E_T$ ) respectively. The L0-PileUp trigger estimates the number of primary  $pp$  interactions in each bunch crossing (event). The L0 Decision Unit (L0DU) collects information from all the three L0 components to form the final L0 trigger decision, a single L0-decision. This L0-decision is passed to the Readout Supervisor (RS) which checks the state of the

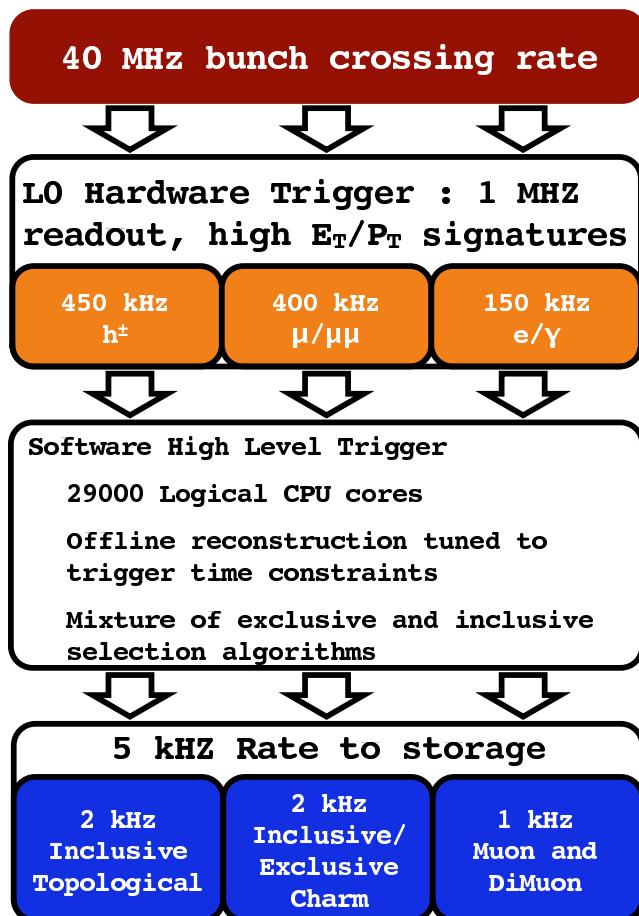


Figure 2.31 LHCb trigger scheme, showing the two trigger levels and the main information used for trigger decision.

front-end (FE) and read-out board buffers. The RS then decides to throttle the L0 trigger or accept it and transmits it to the FE electronics. The time between a  $pp$  interaction and the arrival of the L0 trigger decision at the front-end electronics is fixed to 4  $\mu\text{s}$ , which includes the time-of-flight of the particles, cable delays and all delays in the FE electronics.

**L0 calorimeter trigger** The purpose of the calorimeter triggers is to select and identify particles ( $\pi^0, \gamma$ , electrons and hadrons) with high  $E_{\text{T}}$  ( $E \sin \theta$ ) deposit in the calorimeters: ECAL and HCAL. The ECAL and HCAL signals are read out and processed in the Front-End boards, and each Front-End board is responsible for  $8 \times 4$  (32 in total) calorimeter cells<sup>[176]</sup>. Firstly the trigger processing produces elementary trigger clusters for each Front-End board, and only the highest  $E_{\text{T}}$  of the 32 sums of  $2 \times 2$  cells is selected to minimize the number of candidates to be processed. The combination of  $2 \times 2$  cells is large enough to contain most of one single shower energy, and at the same time the  $2 \times 2$  cells are small enough to avoid contamination from nearby showers. After the highest  $E_{\text{T}}$  cluster is selected in each board, the PS/SPD information, which is the logic decision of whether there are corresponding hits, is introduced to identify the electron and photon. In all, three types of candidates are built for the final trigger decision:

- Hadron candidate *L0Hadron*: an HCAL cluster. The  $E_{\text{T}}$  associated to the hadron candidate is the sum of the  $E_{\text{T}}$  of the HCAL cluster and the ECAL cluster in front if any.
- Electron candidate *L0Election*: an ECAL cluster with 1 or 2 PS cells (up to 4 cells in inner region) having hit in front of it and at least one SPD cell hit in front of the PS cells. The  $E_{\text{T}}$  of the candidate is only taken from the ECAL cluster.
- Photon candidate *L0Photon*: an ECAL cluster with 1 or 2 PS cells (up to 4 cells in inner region) having a hit in front of it but no hits in the corresponding SPD cells. The  $E_{\text{T}}$  of the candidate is only the  $E_{\text{T}}$  deposited in the ECAL.

The  $E_{\text{T}}$  of the candidates are compared to a fixed threshold. Events containing at least one candidate above threshold are accepted by the L0. In the calorimeter trigger the total number of hits in the SPD is also determined, which is used to veto very busy events that would take a disproportionately large fraction of the available processing time in the HLT. The typical  $E_{\text{T}}$  threshold for photon/electron (hadron) is 2.5 (3.5) GeV and the number of SPD hits is required to be below 600 in the typical 2011 data taking. In Figure 2.32, the

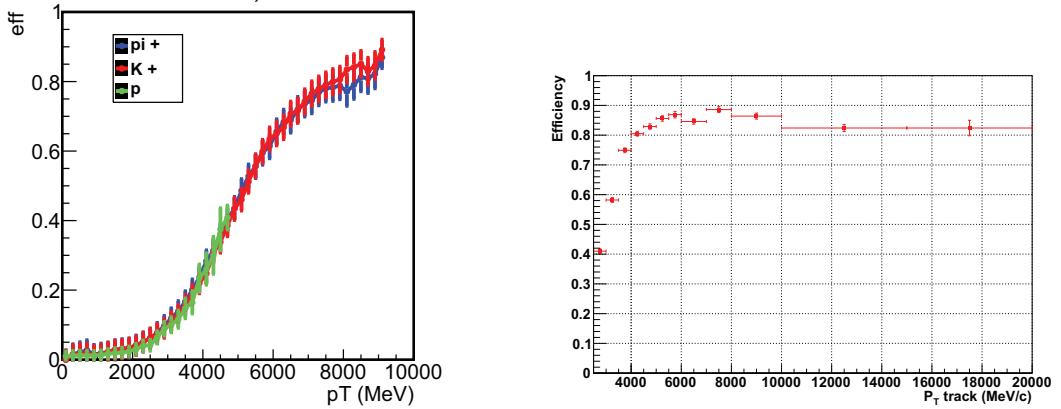


Figure 2.32 Efficiency of the *L0Hadron* trigger (left) and *L0Electron* trigger (right) as a function of the track  $p_T$  in data (2011).

efficiencies of *L0Hadron* and *L0Electron* are shown as a function of the  $p_T$  of the tracks.

**L0 muon trigger** The muon chambers allow stand-alone muon reconstruction where the track finding is performed on the logical pad layout. It assumes muon tracks coming from the interaction point with a single kick from the magnet. Hits in M3 are selected by the track finding algorithm as seeds. For each logical pad hit in M3, a straight line connecting the pad hit and the interaction point is drawn, and the straight line is subsequently extrapolated to M2, M4 and M5 where hits are searched in the FOI centered approximately on the extrapolated positions of the straight line. If for each M2, M4 and M5 station, at least one hit is found inside the FOI, a muon track is flagged and the pad hit in M2 closest to the track extrapolation is selected. When the straight line from M3 and M2 is extrapolated to M1, the track position in M1 can be determined. The position of the track in M1 and M2 allows the determination of the track  $p_T$  (by lookup-tables) with a resolution of  $\approx 20\%$ . Each quadrant of the muon stations work independently, hence muons traversing quadrant boundaries cannot be reconstructed in the trigger. The L0 muon trigger selects the two muons with the highest  $p_T$  for each quadrant of the muon detector and their information is sent to the L0DU. The final trigger decision sets a single threshold on either the largest  $p_T$  of the eight candidates (*L0Muon*), or a threshold on  $p_T^{\text{largest}} \times p_T^{\text{2nd largest}}$  (*L0DiMuon*). For the *L0Muon* (*L0DiMuon*), the number of SPD hits is required to be lower than 600 (900).

**L0 Pile-Up trigger** A set of two planes of silicon strip detectors placed upstream of the VELO system is used to determine the number of primary interactions within one

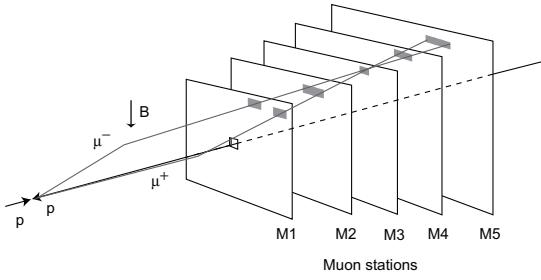


Figure 2.33 Track finding by the L0 muon trigger. Grey areas in each muon station illustrate the field of interests used by the track finding algorithm.

bunch crossing. The Pile-Up plates are  $R$ -sensors with strips at constant radii. With two hits from the two sensors  $A$  and  $B$ , assuming a track originating from the PV, the  $z$  coordinate of the PV can be reconstructed as  $z_{\text{PV}} = \frac{R_b Z_a - R_a Z_b}{(R_b - R_a)}$ , where  $R_a, R_b$  are the  $R$  coordinates of the two hits, and  $Z_a, Z_b$  are the  $z$  coordinates of the two hits, which are also the  $z$ -coordinates of the sensors. By calculating all the hit pairs from the two plates, a histogram with various  $z$  of PV can be produced, and the peak in the histogram is related to the PV. After the first PV is found, the entries of the  $z$  positions from hits coming from the first PV are removed from the histogram for further PV search.

#### 2.2.6.2 HLT trigger

The HLT is a C++ application running on an Event Filter Farm (EFF) composed of several thousands of CPU nodes. The HLT itself is divided into two parts: HLT1 and HLT2. HLT1 reduces its input rate from L0 to about 40 kHz using partial reconstruction of the event. At the HLT2 level, events are reconstructed and selected by a set of inclusive and exclusive algorithms. The reconstruction performed in HLT2 is as similar as possible to the one performed offline. Given the input L0 rate of about 1 MHz and the number of CPU available, the HLT should process an event in the limited time of about 30 ms.

**General Features of HLT1** The HLT1 partially reconstructs the event, such as the PV, the tracks with large IP etc.

The VELO reconstruction software is fast enough to allow a full 3D pattern recognition for all tracks. The VELO tracks are used to construct vertices with at least 5 tracks originating from them, and those vertices within a radius of 300 mm around  $\text{PV}_{xy}^{\text{mean}}$  are considered to be primary vertices (PV). In the HLT1, VELO tracks which are more signal

like are selected to look for the corresponding hits in the tracking stations. In addition cuts are applied to the quality of the VELO track and the difference between the number of hits assigned to the VELO track and the number of hits expected.

In HLT1, for events triggered by *L0Muon*, *L0DiMuon*, a fast muon identification is performed by extrapolating the VELO tracks to M3, where hits are searched in a window. The VELO track is combined with the M3 hits to form a candidate track, and extra hits are searched in other muon stations in the FOI around the track. The VELO track is accepted as a muon candidate if at least one additional hit is found in the stations M2, M4 and M5. Then a track (a muon candidate) is reconstructed combining the selected VELO tracks and the T-station hits, and the momentum and IP can be determined for further cuts. For tracks tagged as the muon candidates, the off-line muon identification algorithm is applied to the tracks to assign a muon ID. Using the identified muons, selections for single muon candidate and dimuon vertices are performed respectively.

The detailed settings in the triggers change according to the data taking conditions, and usually they are different for physics data taking and testing beams, and different for different peaking luminosity. The Trigger Configuration Keys (TCKs) specify a list of triggers applied and their detailed selections.

### 2.2.6.3 The LHCb Event Processing

At LHCb a series of software packages<sup>[177]</sup> are maintained to process the data, both Monte Carlo and Real Data, at different phases. Various steps normally follow each other in a sequential manner—the output of one stage is the input of another. The first step in Monte Carlo is the physics simulation, which simulates the  $pp$  collisions, the decays of resonances (Generation) and the propagation of secondary particles through the LHCb detector and their interactions with detector material (Simulation). The Gauss package handles the simulation phases integrating several software toolkits, PYTHIA<sup>[178]</sup> and EvtGen<sup>[179]</sup> for Generation and GEANT4<sup>[180]</sup> for Simulation. The outputs of Gauss are the Monte Carlo Truth history and hits in the detector volume for the next phase, digitalization which is processed by the Boole software package. Boole applies the detector response to these ‘hits’ for each sub detector, including the read-out electronics, as well as the L0 trigger hardware. The output of Boole, the Raw data, has the same format as the real data in the normal data taking. The Raw data, either from Monte Carlo Simulation or Real Data taking, must then be reconstructed in order to provide physical quantities:

the calorimeter clusters to provide the energy of electromagnetic and hadronic showers, the tracker hits for the reconstruction of charged particle tracks, the Particle ID in appropriate sub-systems of the tracks for physics analysis. The Brunel package is responsible for the LHCb reconstruction, producing the complete DST (Data Summary Tape) format for end-user analysis or reduced DST (rDST) for further data processing. The rDST is used for the production of streams of inclusive events by the DaVinci package for further individual analysis. This step is called the stripping. Each stream of the stripping comprises several preselected physics channels (lines) sharing the same decay products or having similar decay topology or analysis tools. The output of the stripping is also in DST format. From the DST, the end-user can extract various kinds of information for individual physics analysis easily via the DaVinci toolkit.

The packages are updated with a version number/tag according to the updating of various inputs of the software framework, for example: the event models, the description of beam or detector conditions, the convention of analysis tools/flows. The software framework uses unique tags to insure the consistency between various packages during the flow of data processing for the same dataset.

## 2.3 Summary

This chapter gives detailed description of LHCb detector, including the goals of various sub-detectors and their setups, configurations and real-time performance during the data taking. The LHCb triggers and data handling have also been described. In all, the LHCb detector works excellently, and can provide us high quality data for the heavy quarkonium analysis.

## Chapter 3 Determination of quarkonium polarization from experimental data

### 3.1 From theory to experiment

The polarization of a particle with spin is determined by its spin density matrix<sup>[181]</sup>, which defines the amplitude when the particle has mixing spin states. The  $J/\psi$  and  $\psi(2S)$  are spin one (vector) massive particles with three spin eigen states,  $\pm 1, 0$  states. The spin density matrix for a massive vector boson is of dimension three with elements denoted as  $\rho_{\lambda\lambda'}$ , where  $\lambda' = \pm 1, 0$ .  $\rho_{\lambda\lambda'}$  is given by the polarized production cross section  $\sigma_{\lambda\lambda'}$ <sup>[182]</sup>:

$$\rho_{\lambda\lambda'} \propto \sigma_{\lambda\lambda'} = A[\psi(\lambda)]A^*[\psi(\lambda')] \quad (3-1)$$

where  $A[\psi(\lambda)]$  is the production amplitude for a  $\psi$  produced with helicity  $\lambda$ . So the  $\psi$  polarization is completely determined by the production mechanism. The spin density matrix is Hermitian, satisfying  $\rho_{-1,-1} + \rho_{+1,+1} + \rho_{0,0} = 1$ . In the dimuon decay channel, if the  $z$ -axis is chosen in the production plane (as is done in this analysis), conservation of parity imposes the following constraints:

$$\begin{aligned} \rho_{-1,-1} &= \rho_{+1,+1} \\ \rho_{-1,+1} &= \rho_{+1,-1} \\ \rho_{-1,0} &= -\rho_{+1,0} \\ \rho_{0,-1} &= -\rho_{0,+1} \end{aligned} \quad (3-2)$$

and the number of independent matrix elements reduces to four:  $\rho_{-1,-1}$ ,  $\rho_{1,-1}$ ,  $\text{Re}(\rho_{0,1})$  and  $\text{Im}(\rho_{0,1})$ .

In the channel  $\psi \rightarrow \mu^+\mu^-$ , the helicity amplitude analysis<sup>[183]</sup> shows that the double differential cross section as a function of  $\mu^+$  (or  $\mu^-$ ) angular variables in the rest frame of  $\psi$  can be expressed as:

$$P(\Omega, \Lambda) \equiv \frac{d^2\sigma}{d\cos\theta d\phi} \propto (1 + \lambda_\theta \cos^2\theta + \lambda_{\theta\phi} \sin 2\theta \cos\phi + \lambda_\phi \sin^2\theta \cos 2\phi) \quad (3-3)$$

where  $\Omega \equiv (\cos\theta, \phi)$  are muon angular variables with  $\theta$  defined with regard to the  $z$ -axis and  $\phi$  is the azimuthal angle defined in a predefined coordinate system, while  $\Lambda \equiv (\lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi)$  are called polarization parameters, and  $\lambda_\theta$  is usually called  $\alpha$  in the

one dimensional analysis. The  $\Lambda$  parameters are related to the spin density matrix as<sup>[182]</sup>:

$$\begin{aligned}\lambda_\theta &= \frac{\rho_{+1,+1} - \rho_{0,0}}{\rho_{+1,+1} + \rho_{0,0}}, \\ \lambda_{\theta\phi} &= \frac{\sqrt{2}\rho_{+1,0}}{\rho_{+1,+1} + \rho_{0,0}}, \\ \lambda_\phi &= \frac{2\rho_{+1,-1}}{\rho_{+1,+1} + \rho_{0,0}},\end{aligned}\tag{3-4}$$

where, to make the expressions more symmetric,  $\rho_{+1,+1}$  is written explicitly. In fact  $\rho_{+1,+1} = (1 - \rho_{0,0})/2$  as can be seen above. Therefore the  $\psi$  polarization can be extracted from the angular distribution of muons.

If the  $\psi$  is produced in pure helicity 0 states ( $\pm 1$ ), which means  $\rho_{0,0} = 1$  ( $\rho_{0,0} = 0$ ), the polarization parameters will be  $\lambda_\theta = -1$  (+1), and  $\lambda_{\theta\phi} = \lambda_\phi = 0$ , which is called totally longitudinal (transverse) polarization. In the case  $\lambda_\theta = \lambda_{\theta\phi} = \lambda_\phi = 0$ , the quarkonium is unpolarized, and the muons fly isotropically in the rest frame of the quarkonium, and in this case the muon  $\cos \theta - \phi$  two-dimensional distribution is uniform.

Theoretically, the  $\psi$  polarization parameters ( $\lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi$ ) depend on the definition of the coordinate system, and at the same time, to analyze the muon angular distribution, a coordinate system must be chosen experimentally<sup>[183]</sup> to project the muon flight direction in collider experiments. There were several choices of frames and they have the same direction for  $y$ -axis which is the normal of the production plane. The production plane is formed by the direction of the colliding beam and the direction of quarkonium in the rest frame of the two colliding beams. Various frames are thus characterized by the choice of  $z$ -axis. Three commonly used coordinate systems are: the helicity frame (HX), which uses the flight direction of the quarkonium itself in the center-of-mass of the colliding beams as the quantization axis ( $z$ -axis); the Gottfried-Jackson frame<sup>[184]</sup> (GJ), which has the direction of the momentum of one of the two colliding beams in quarkonium rest frame as  $z$ -axis, and the Collins-Soper frame<sup>[185]</sup> (CS), which takes the bisector of the angle between one beam and the opposite of the other beam in quarkonium rest frame as the  $z$ -axis. Figure 3.1 illustrates the three definitions of the quantization frames. Different frames are connected by a pure space rotation around the common  $y$ -axis, and in the limit of zero quarkonium transverse momentum  $p_T$ , the three coordinate frames become coincident. Because of the pure space rotation, the three polarization parameters measured in one frame are combinations of the parameters in another frame, the relation in terms of

the rotation angle  $\delta$  (the angle between the two quantization axes in two frames) reads:

$$\begin{aligned}\lambda'_\theta &= \frac{\lambda_\theta - 3\Lambda}{1 + \Lambda} \\ \lambda'_\phi &= \frac{\lambda_\phi + \Lambda}{1 + \Lambda} \\ \lambda'_{\theta\phi} &= \frac{\lambda_{\theta\phi} \cos 2\delta - (\lambda_\theta - \lambda_\phi) \sin 2\delta/2}{1 + \lambda}\end{aligned}\quad (3-5)$$

where  $\Lambda = [(\lambda_\theta - \lambda_\phi) \sin^2 \delta - \lambda_{\theta\phi} \sin 2\delta]/2$ . The angle  $\delta$  varies from event to event because the  $\psi$  momentum is not fixed. There exists some frame invariant quantities that are definable in terms of  $\lambda_\theta$ ,  $\lambda_{\theta\phi}$  and  $\lambda_\phi$  in the form:

$$\mathcal{F}_{\{c_i\}} = \frac{(3 + \lambda_\theta) + c_1(1 - \lambda_\phi)}{c_2(3 + \lambda_\theta) + c_3(1 - \lambda_\phi)}$$

where  $c_i$  are arbitrary constant numbers. The popularly used one is:

$$\mathcal{F}_{-3,0,1} = \frac{\lambda_\theta + 3\lambda\phi}{1 - \lambda_\phi} \quad (3-6)$$

with  $c_1 = -3$ ,  $c_2 = 0$ ,  $c_3 = 1$

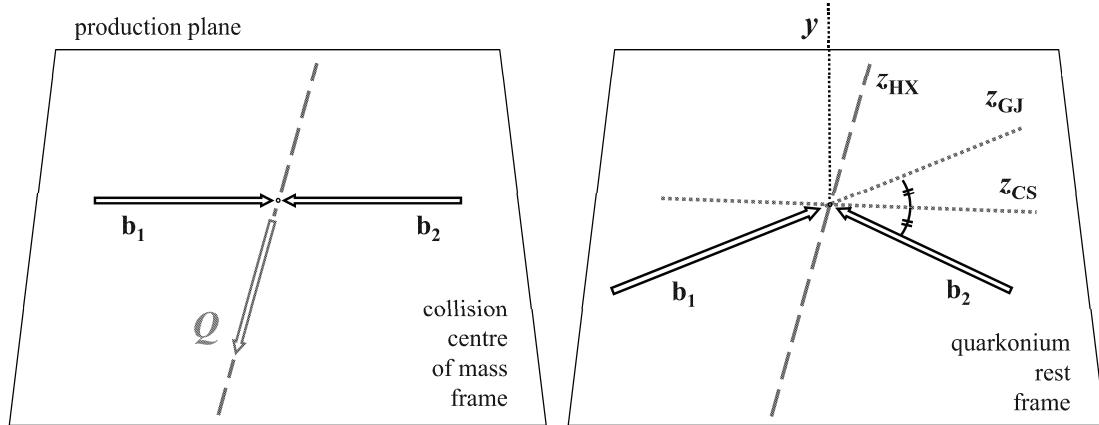


Figure 3.1 Illustration of the production plane (*left*) and three different definitions of the polarization axis  $z$  with respect to the directions of motions of the colliding beams ( $b_1$ ,  $b_2$ ) and of the quarkonium (*right*).

## 3.2 Polarization extraction from data

### 3.2.1 Construction of likelihood estimator

The polarization measurement is to extract the polarization parameters  $\lambda_\theta$ ,  $\lambda_{\theta\phi}$  and  $\lambda_\phi$  by fitting the two dimensional polarization angular distribution in Equation 3-3 to data. However, in general the global detection efficiency  $\epsilon$  as a function of  $(\cos \theta, \phi)$  is not constant, so the distribution observed is  $P(\Omega | \Lambda) \times \epsilon(\Omega)/N(\Lambda)$ , where  $\Omega = (\cos \theta, \phi)$ ,  $\Lambda = (\lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi)$  and  $N(\Lambda)$  is the normalization factor, which depends on the three polarization parameters through the polarization angular distribution and efficiency. The normalization is studied in Monte Carlo.

The log-likelihood function (estimator) is then constructed as:

$$\begin{aligned} \log \mathcal{L} &= \log \prod_{i=1}^{N_s} \left[ \frac{P(\cos \theta_i, \phi_i | \lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi) \times \epsilon_{\text{tot}}(\cos \theta_i, \phi_i)}{\text{Norm}(\lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi)} \right] \\ &= \sum_{i=1}^{N_s} \log \left[ \frac{P(\cos \theta_i, \phi_i | \lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi) \times \epsilon_{\text{tot}}(\cos \theta_i, \phi_i)}{\text{Norm}(\lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi)} \right] \end{aligned} \quad (3-7)$$

where the summation  $\sum_{i=1}^{N_s}$  runs over the number of signal events  $N_s$ , that is, the likelihood function is only valid for signal events because firstly  $P(\Omega|\Lambda)$  is parameterized only for signals and secondly the efficiency is calculated from signal events (calculated in Monte Carlo using signal events). So for real data where background events can not be separated from signals event by event, the background contamination has to be subtracted properly. To do this, a weighted likelihood estimator is used, where the weight is chosen to be function of the reconstructed  $\mu^+ \mu^-$  invariant mass. The weighted likelihood in general is written as:

$$\log \mathcal{L} = \sum_{i=1}^{N_{\text{tot}}} w(m_i) \times \log \left[ \frac{P(\cos \theta_i, \phi_i | \lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi) \times \epsilon_{\text{tot}}(\cos \theta_i, \phi_i)}{\text{Norm}(\lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi)} \right] \quad (3-8)$$

where  $m_i$  is the mass of the candidate for  $i$ th event and the summation runs over all events in data. The weighting function is chosen such that:

- the total weights for signal events should be equal to the number of signal events, which means that the procedure of subtracting the background should not change the contribution of signals;
- the total weights for background events should equal to zero within statistical error, so the weights for some of the events are positive, while others are negative;

- the angular distribution of those events with positive weights should be the same as the angular distribution of the events with negative weights.

With these assumptions, in the weighted likelihood the contribution from background events cancels automatically while the contribution from signal events is the same as in the unweighted likelihood constructed with only signal events.

The goal is to extract the polarization parameters  $\lambda_\theta$ ,  $\lambda_{\theta\phi}$  and  $\lambda_\phi$ , however from Equation 3-9 it can be seen that the total efficiency  $\epsilon_{\text{tot}}$  in the numerator does not contain any polarization parameter dependence, so it can be eliminated because it does not have any effect on the maximization of the likelihood function with regard to the three  $\lambda$ .

$$\begin{aligned}
 \log \mathcal{L} &= \sum_{i=1}^{N_{\text{tot}}} w(m_i) \times \log \left[ \frac{P(\cos \theta_i, \phi_i \mid \lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi)}{\text{Norm}(\lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi)} \right] + \sum_{i=1}^{N_{\text{tot}}} w(m_i) \times \log [\epsilon_{\text{tot}}(\cos \theta_i, \phi_i)] \\
 &= \sum_{i=1}^{N_{\text{tot}}} w(m_i) \times \log \left[ \frac{P(\cos \theta_i, \phi_i \mid \lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi)}{\text{Norm}(\lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi)} \right] + \text{const.}
 \end{aligned} \tag{3-9}$$

### 3.2.2 Determination of normalization from Monte Carlo

The normalization factor  $\text{Norm}(\lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi)$  in Equation 3-9 is calculated from unpolarized Monte Carlo.  $\text{Norm}(\lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi)$  is defined explicitly as:

$$\begin{aligned}
 \text{Norm}(\lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi) &\equiv \int d\Omega P(\cos \theta, \phi \mid \lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi) \times \epsilon_{\text{tot}}(\cos \theta, \phi) \\
 &= \int d\Omega [\epsilon_{\text{tot}}(\Omega) + \lambda_\theta \cos^2 \theta \times \epsilon_{\text{tot}}(\Omega) + \lambda_{\theta\phi} \sin 2\theta \cos \phi \times \epsilon_{\text{tot}}(\Omega) + \lambda_\phi \sin^2 \theta \cos 2\phi \times \epsilon_{\text{tot}}(\Omega)] \\
 &= \int d\Omega \epsilon_{\text{tot}}(\Omega) + \lambda_\theta \int d\Omega \cos^2 \theta \epsilon_{\text{tot}}(\Omega) \\
 &\quad + \lambda_{\theta\phi} \int d\Omega \sin 2\theta \cos \phi \epsilon_{\text{tot}}(\Omega) + \lambda_\phi \int d\Omega \sin^2 \theta \cos 2\phi \epsilon_{\text{tot}}(\Omega),
 \end{aligned} \tag{3-10}$$

where  $\Omega \equiv (\cos \theta, \phi)$  as before. Equation 3-9 shows that the estimator will give the same  $\lambda_\theta$ ,  $\lambda_{\theta\phi}$  and  $\lambda_\phi$  when  $\text{Norm}$  is multiplied by a constant factor, and this means that  $\epsilon_{\text{tot}}(\Omega)$  can be normalized to any value (unity, in the special case) in the two dimension  $(\cos \theta, \phi)$  space in Equation 3-10. The normalization  $\text{Norm}$  calculated this way is equivalent to the real  $\text{Norm}$ . If  $\epsilon_{\text{tot}}(\Omega)$  is normalized to unity, the  $\epsilon_{\text{tot}}(\Omega)$  can be considered as a two dimensional probability distribution function (PDF) as a function of the variables  $(\cos \theta, \phi)$ . Then the first term  $\int d\Omega \epsilon_{\text{tot}}(\Omega)$  in the last equality of Equation 3-10 just determines the

proper normalization while the second to fourth terms determine the expectation value of  $\cos^2 \theta$ ,  $\sin 2\theta \cos \phi$  and  $\sin^2 \theta \cos 2\phi$  over the PDF respectively.

If the quarkonium in Monte Carlo is generated unpolarized ( $\lambda_\theta = \lambda_{\theta\phi} = \lambda_\phi = 0$ ), the  $\Omega(\cos \theta, \phi)$  distribution is flat at generator level, and thus the  $\Omega$  distribution of the selected events is the same as the shape of the efficiency, but differs up to a global factor (see Equation 3-11). The global factor, however, does not make any trouble in the determination of the normalization.

$$\epsilon_{\text{tot}} = \frac{f(\Omega_{\text{selected}})}{f(\Omega_{\text{produced}})} = \frac{f(\Omega_{\text{selected}})}{C} \propto f(\Omega_{\text{selected}}) \quad (3-11)$$

The observed distribution of  $\Omega(\cos \theta, \phi)$  in unpolarized Monte Carlo can be normalized to be the PDF of the efficiency. Because the selected events in unpolarized Monte Carlo is a sample generated from the PDF, the integrals (expectations) in Equation 3-10 can be estimated from the sample in the standard way:

$$\begin{aligned} \int d\Omega \epsilon_{\text{tot}}(\Omega) &= \frac{\sum_i 1}{N_{\text{selected}}} \equiv 1 \\ \lambda_\theta \int d\Omega \cos^2 \theta \epsilon_{\text{tot}}(\Omega) &= \lambda_\theta \times \frac{\sum_i \cos^2 \theta_i}{N_{\text{selected}}} \equiv a\lambda_\theta \\ \lambda_{\theta\phi} \int d\Omega \sin 2\theta \cos \phi \epsilon_{\text{tot}}(\Omega) &= \lambda_{\theta\phi} \times \frac{\sum_i \sin 2\theta_i \cos \phi_i}{N_{\text{selected}}} \equiv b\lambda_{\theta\phi} \\ \lambda_\phi \int d\Omega \sin^2 \theta \cos 2\phi \epsilon_{\text{tot}}(\Omega) &= \lambda_\phi \times \frac{\sum_i \sin^2 \theta_i \cos 2\phi_i}{N_{\text{selected}}} \equiv c\lambda_\phi \end{aligned} \quad (3-12)$$

in which, the index  $i$  runs over all the selected events in the Monte Carlo sample, and  $N_{\text{selected}}$  is the number of events selected and is also the normalization of the efficiency PDF. The normalization  $\text{Norm}(\lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi)$  in the likelihood estimator is totally determined by the three normalization constants ( $a, b, c$ ) in Equation 3-12.

Finally the likelihood estimator reads:

$$\log \mathcal{L} = \sum_{i=1}^{N_{\text{tot}}} w(m_i) \times \log \left[ \frac{P(\cos \theta_i, \phi_i | \lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi)}{1 + a\lambda_\theta + b\lambda_{\theta\phi} + c\lambda_\phi} \right] \quad (3-13)$$

It should be noted that the normalization ( $a, b$  and  $c$  actually) can also be extracted through polarized Monte Carlo where the polarization is known in advance. In the polarized Monte Carlo, the observed  $\Omega(\cos \theta, \phi)$  distribution will be different from the efficiency as a function of  $\Omega(\cos \theta, \phi)$ . The observed  $\Omega(\cos \theta, \phi)$  distribution will be actually  $P(\cos \theta, \phi | \lambda'_\theta, \lambda'_{\theta\phi}, \lambda'_\phi) \times \epsilon_{\text{tot}}(\cos \theta, \phi)$ , where the three  $\lambda'$  are the input polarization in Monte Carlo. So the observed distribution can not be taken as efficiency  $\epsilon_{\text{tot}}(\cos \theta, \phi)$

directly. However, the real efficiency can be reproduced by multiplying the observed  $\Omega(\cos \theta, \phi)$  by the function  $1./P(\cos \theta, \phi \mid \lambda'_\theta, \lambda'_{\theta\phi}, \lambda'_\phi)$ . So if during the summation in Equation 3-12, every event is given a weight, which is  $1./P(\cos \theta, \phi \mid \lambda'_\theta, \lambda'_{\theta\phi}, \lambda'_\phi)$ , the constants  $(a, b, c)$  can be correctly extracted from the polarized Monte Carlo.

### 3.2.3 Minimization with TMinuit

To summarize, in Equation 3-13 the normalization constants are calculated from Monte Carlo, while  $w(m_i)$  and  $P(\cos \theta_i, \phi_i)$  take the values of reconstructed  $\mu^+ \mu^-$  ( $J/\psi$  or  $\psi(2S)$  candidate) mass and muon angles from each event in data, and by summing over all events in data the likelihood estimator is constructed. The weighted log-likelihood function is multiplied by a minus sign and then can be minimized by the TMinuit package integrated in ROOT<sup>[186]</sup>. The set of parameters that minimize the logarithm likelihood function is the best fit result. At the same time the TMinuit will return an error for each parameter, which is called likelihood error (uncertainty) in the following.

### 3.2.4 Fit validation using fully simulated samples

Before the method is applied to extract polarization in data, the method itself is tested with pure  $J/\psi$  Monte Carlo to see its validity. During the tests, the Monte Carlo samples play two roles: one sample is used to extract the efficiency, while the other one behaves just like data in the likelihood estimator, where the  $w(m_i)$  and  $P(\cos \theta_i, \phi_i)$  are calculated for each event. For the sample used to mimic data, both unpolarized Monte Carlo and totally transversely polarized Monte Carlo are produced respectively, and another totally longitudinally polarized Monte Carlo sample is also generated by artificially weighting the unpolarized Monte Carlo. It should be noted that, in the Monte Carlo each event has a  $J/\psi$  generated, so the level of background is rather low.

The results of tests on unpolarized Monte Carlo (with polarization zero) are shown in Figures 3.2, 3.3 and 3.4 for the three parameters respectively in various  $J/\psi$  kinematic bins. From these tests it can be seen that the method can recover the input zero polarization within the statistical errors. During the test on unpolarized Monte Carlo, the role of two Monte Carlo samples can be exchanged, which means that one sample is used to extract the efficiency while the other one used as toy real data (case I) or vice versa (case II), and the result ‘polarization’ parameters (due to statistical fluctuation) for these two cases should deviate from zero in the opposite direction with the same amount, because

in the first case the estimator is equivalently fitting:

$$\frac{f_{\text{selected}(1)}(\Omega)}{f_{\text{selected}(2)}(\Omega)} \propto 1 + \lambda_\theta \cos^2 \theta + \dots$$

where  $f_{\text{selected}(1)}(\Omega)$  and  $f_{\text{selected}(2)}(\Omega)$  is the muon angular distribution of the selected events for the two Monte Carlo samples respectively, and in the second case the fitting is

$$\frac{f_{\text{selected}(2)}(\Omega)}{f_{\text{selected}(1)}(\Omega)} \propto \frac{1}{1 + \lambda_\theta \cos^2 \theta + \dots} \approx 1 - \lambda_\theta \cos^2 \theta - \dots$$

The inverse pattern of the polarization parameters in the two alternative tests are confirmed by the fit results in Figures 3.2, 3.3 and 3.4, where the results in the plots on the right are extracted when the two Monte Carlo samples are swapped. Moreover, it is a proof that the fit method does not introduce any bias, because if there is any bias that shifts the true results, the sum of the two alternative results for each parameter will deviate from zero significantly, however from the results it can be seen that the sum is very close to zero.

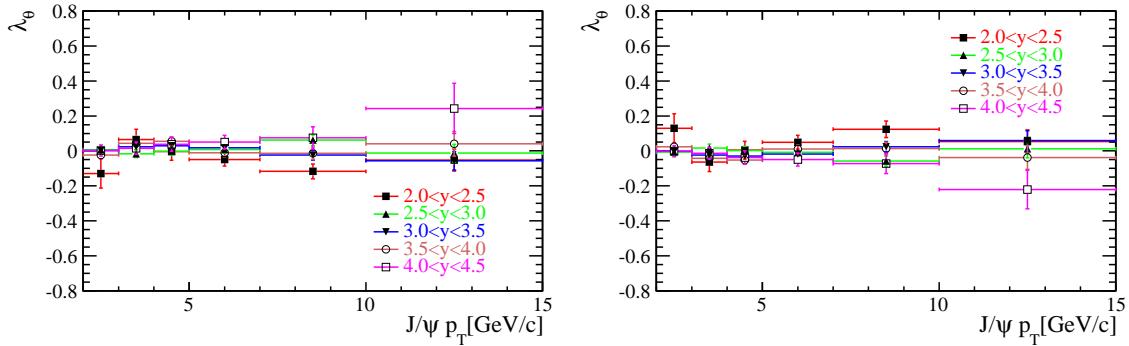


Figure 3.2 Results of  $\lambda_\theta$  in the test with unpolarized Monte Carlo in each transverse momentum and rapidity bin. The roles of the two unpolarized Monte Carlo can be swapped, so two alternative results are extracted. Only statistical errors plotted.

Concerning the test on the totally transversely polarized sample, the results are shown in the Figures 3.5, 3.6 and 3.7, together with the distribution of the pull defined as  $(\lambda_{\text{fit}} - \lambda_{\text{true}})/\delta_{\lambda_{\text{fit}}}$ . In an unbiased fit, the mean of the pull distribution will be zero. From the results and the pull distributions, it can be seen that the measurements are consistent with transverse polarization within the statistical fluctuations.

The results of test on the totally longitudinally polarized sample are shown in Figures 3.8, 3.9 and 3.10. The longitudinal polarization is produced by weighting one of the two unpolarized Monte Carlo samples according the polarization angular distribution

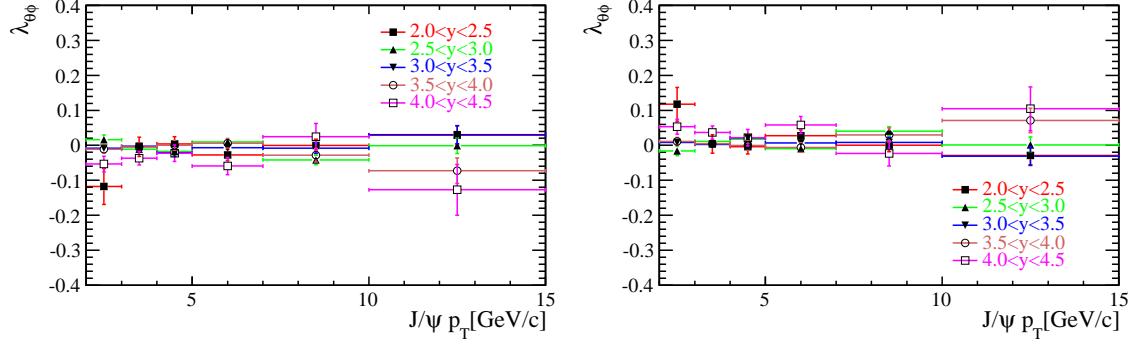


Figure 3.3 Results of  $\lambda_{\theta\phi}$  in the test with unpolarized Monte Carlo in each transverse momentum and rapidity bin. The roles of the two unpolarized Monte Carlo can be swapped, so two alternative results are extracted. Only statistical errors plotted.

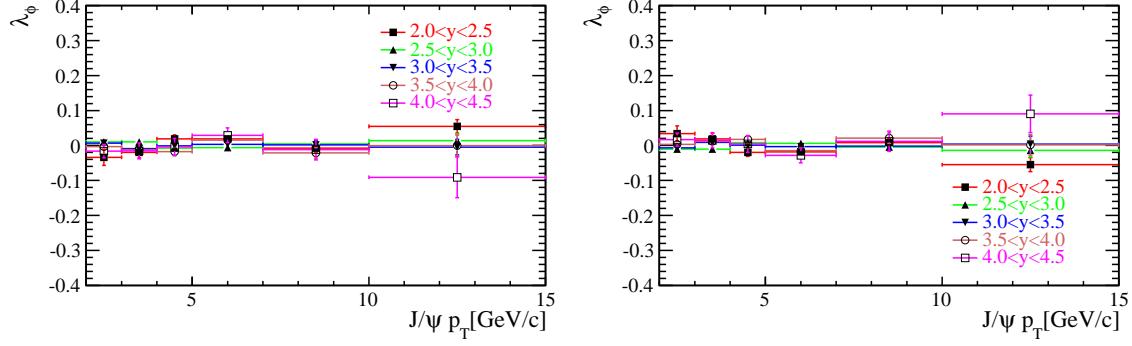


Figure 3.4 Results of  $\lambda_\phi$  in the test with unpolarized Monte Carlo in each transverse momentum and rapidity bin. The roles of the two unpolarized Monte Carlo can be swapped, so two alternative results are extracted. Only statistical errors plotted.

with parameters  $\lambda_\theta = -1$  and  $\lambda_{\theta\phi} = \lambda_\phi = 0$  and the other Monte Carlo sample is used to calculate the normalization. The roles of the two Monte Carlo can be swapped, yielding two sets of polarization parameters. From the results, it can be seen that the two measurements deviate from the true values in opposite direction with the same size, which is a proof that there is no bias.

In conclusion, the full angular likelihood method can give results that are consistent with the input values. The tests of the method are done particularly in the HX frame, however the estimator only uses quantities in a predefined reference frame and can reproduce the input polarization parameters within that frame, so the validity of the method itself does not depend on the frame choice as long as the same frame is used for Monte Carlo (for efficiency determination) and (toy) data.

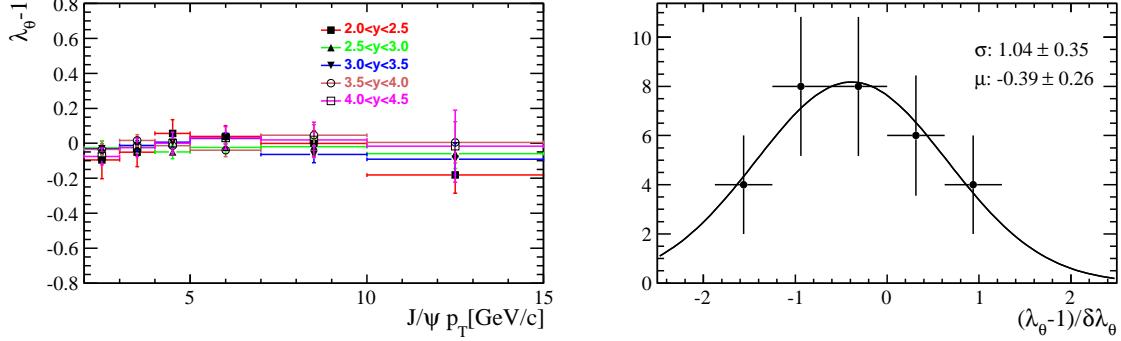


Figure 3.5 Results of  $\lambda_\theta$  in the test with totally transversely polarized Monte Carlo (left) and the distribution of the pull of the fitted results (right). The gaussian fit to the pull distribution is superimposed. Only statistical errors included.

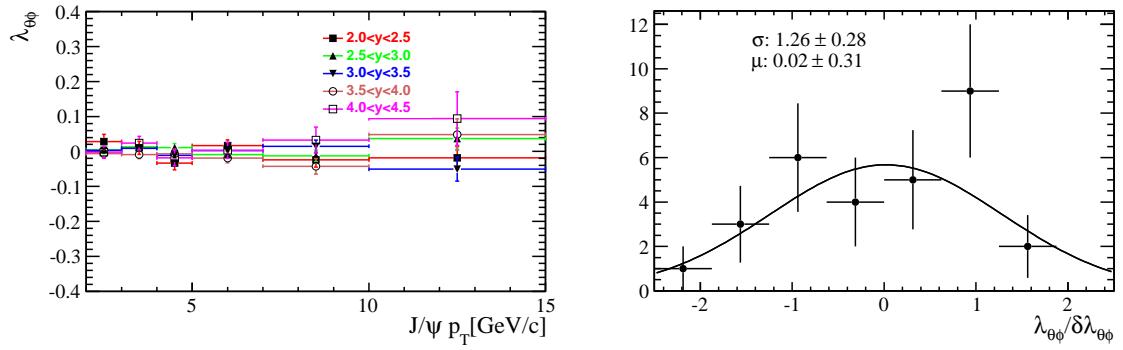


Figure 3.6 Results of  $\lambda_{\theta\phi}$  in the test with totally transversely polarized Monte Carlo (left) and the distribution of the pull of the fitted results (right). Only statistical errors included.

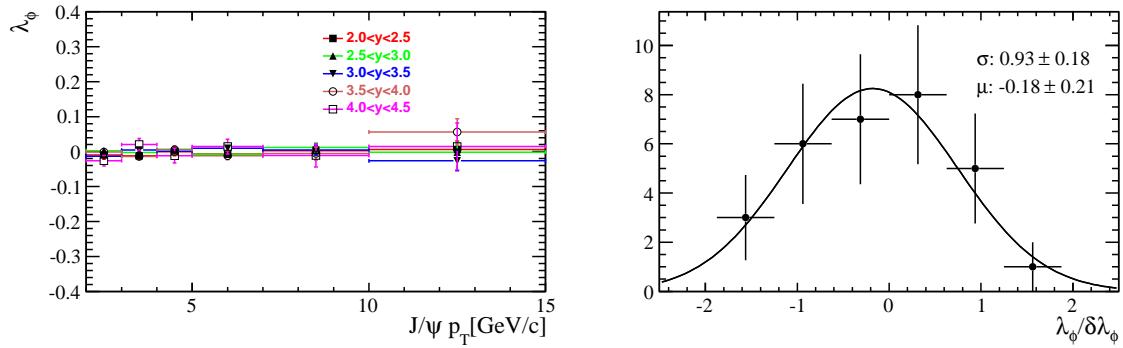


Figure 3.7 Results of  $\lambda_\phi$  in the test with totally transversely polarized Monte Carlo (left) and the distribution of the pull of the fitted results (right). The gaussian fit to the pull distribution is superimposed. Only statistical errors included.

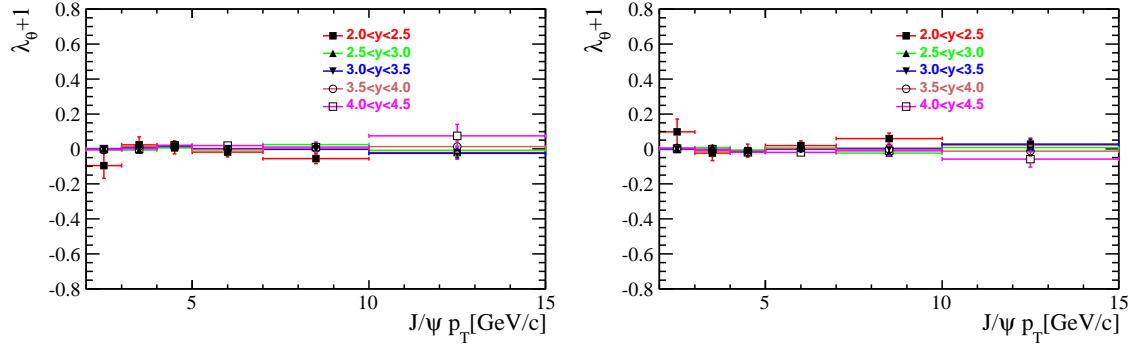


Figure 3.8 Results of  $\lambda_\theta$  in the test with artificially produced longitudinally polarized Monte Carlo. The roles of the two Monte Carlo can be swapped, so two alternative results are extracted. Only statistical errors included.

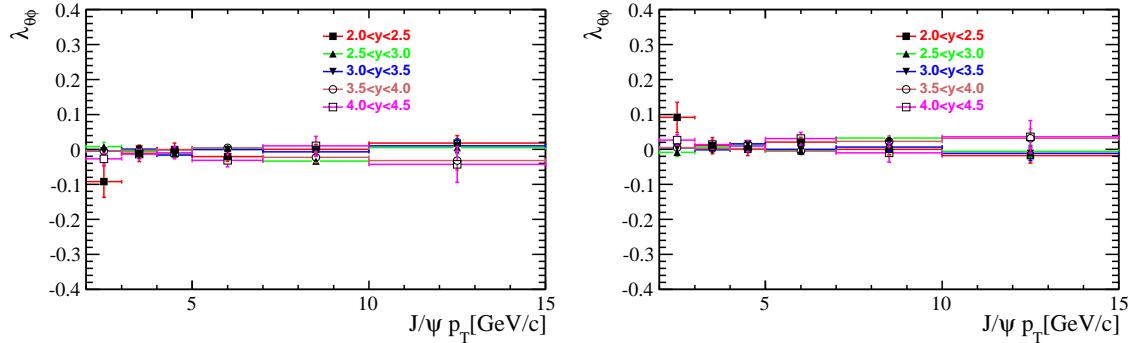


Figure 3.9 Results of  $\lambda_{\theta\phi}$  in the test with totally longitudinally polarized Monte Carlo. The roles of the two Monte Carlo can be swapped, so two alternative results are extracted. The gaussian fit to the pull distribution is superimposed. Only statistical errors included.

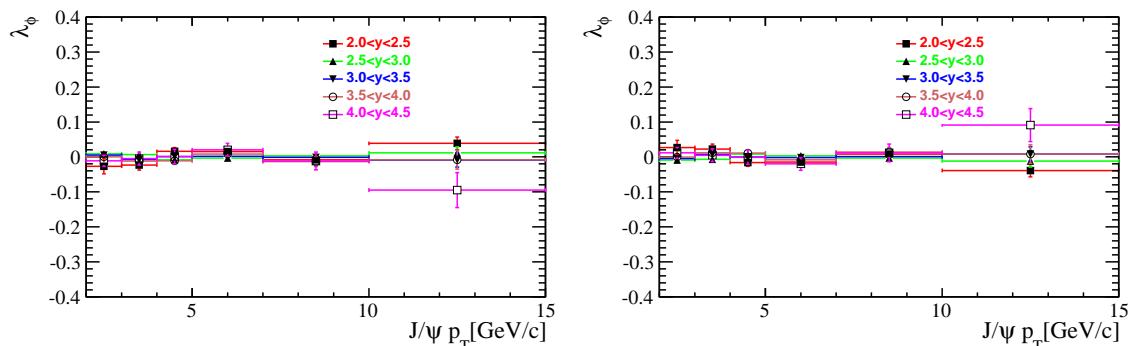


Figure 3.10 Results of  $\lambda_\phi$  in the test with totally longitudinally polarized Monte Carlo. The roles of the two Monte Carlo can be swapped, so two alternative results are extracted. Only statistical errors included.

## Chapter 4 **Prompt $J/\psi$ polarization measurement**

In this chapter the detailed analysis of prompt  $J/\psi$  polarization measurement is presented, including the data set preparation, the fitting procedures, the systematic uncertainty evaluation and finally the results and their comparisons with other measurements and theoretical calculations. The analysis is performed in both the HX frame and the CS frame, and generally the analysis procedure in the two frames is identical, so unless otherwise noticed, the discussions (tables and numbers) are for the HX frame.

### 4.1 Data sets and Selections

The  $J/\psi$  polarization measurement is performed with LHCb 7 TeV data with integrated luminosity of about  $370 \text{ pb}^{-1}$  taken in the early stage of the 2011 run — before the summer machine development/technical stop. The data was acquired with both magnet polarities, MagUp and MagDown. The  $J/\psi$  candidates are reconstructed by the standard LHCb stripping processing *Reco10 – Stripping13b* in the line *MicroDSTDiMuonDiMuon-IncLine* in the *Leptonic* stream. Because of the increasing peak luminosity during the data taking period, the data acquisition environment changed with time, reflected in the various TCKs used. The events are filtered with three TCKs, in which the configurations (see Table 4.2) are identical concerning the signal  $J/\psi \rightarrow \mu^+ \mu^-$ . The three TCK cover about 96% of all the  $370 \text{ pb}^{-1}$  luminosity. The used TCKs and the corresponding luminosities respectively are shown in Table 4.1.

A sample of 20 million (20 M)  $J/\psi \rightarrow \mu^+ \mu^-$  events (10M MagUp + 10M MagDown) is generated to study the LHCb detection efficiency in the Monte Carlo. In the Monte Carlo sample, the proton beam and LHCb detector conditions are described with

Table 4.1 The luminosities for the TCKs selected for the  $J/\psi$  polarization analysis.

TCK	Total ( $\text{pb}^{-1}$ )	MagUp ( $\text{pb}^{-1}$ )	MagDown ( $\text{pb}^{-1}$ )
0x005A0032	67.1	38.6	28.4
0x006D0032	100.3	-	100.3
0x00730035	195.6	133.7	61.8

the LHCb *MC11a-Sim05a* simulation condition tag which also defines a specified list of LHCb event processing packages used (see section 2.2.6.3). The TCK 0x40760037 is applied to mimic the data taking trigger. In this sample the  $J/\psi$  is forced to be unpolarized, which makes it easy to determine the relative detection efficiency as a function of muon angular variables taking advantage of the fact that the generated muon angular distribution is uniform.

In addition, to study the agreement between the data and Monte Carlo for the description of the detection efficiency, the exclusive  $B^+ \rightarrow J/\psi K^+$  channel is also investigated in data (2011) and at the same time a Monte Carlo sample of about 8 M events is generated with the *MC11a-Sim05* condition for both two magnet polarities.

#### 4.1.1 Selections

The  $J/\psi \rightarrow \mu^+ \mu^-$  candidates are selected from *MicroDSTDiMuonDiMuonIncLine* stripping line, which reconstructs  $J/\psi$  with some loose selections applied. For the inputs of the stripping line, two charged tracks are selected asking for the *IsMuon* (see section 2.2.5.1) requirement, and then a  $J/\psi$  vertex is fitted from the two muons. The  $J/\psi$  candidates with the following conditions fulfilled are kept in the stripping line (when the cut is applied on the muon, it is on both muons):

- $\mu$  track transverse momentum:  $p_T(\mu) > 650 \text{ MeV}/c$
- $\mu$  track quality:  $\chi^2/\text{ndof} < 5$
- $J/\psi$  candidate vertex quality:  $\chi^2/\text{ndof} < 20$
- $J/\psi$  candidate mass:  $M(\mu^+ \mu^-) > 2900 \text{ MeV}/c^2$

In the offline analysis, tighter cuts are applied to the  $J/\psi$  candidates to suppress the background and to select signal events with high quality. The cuts are chosen to be similar to the ones used in the cross section measurement of  $J/\psi$  <sup>[147]</sup> at LHCb. The cuts, which are generally still loose are:

- The events have at least one reconstructed primary vertex used to compute the pseudo lifetime
- $\mu$  track transverse momentum:  $p_T(\mu) > 750 \text{ MeV}/c$
- $\mu$  track quality:  $\chi^2/\text{ndof} < 4$
- Combined muon identification:  $\Delta \log \text{PID}_{\mu/\pi} > 0$
- Clone killing cuts: keep only one candidate if  $\cos \theta(\mu_1^+, \mu_2^+) > 0.9999$  and  $\cos \theta(\mu_1^-, \mu_2^-) > 0.9999$  for the two  $J/\psi$  candidates in the same events

Table 4.2 The cuts applied in the trigger lines selected for the  $J/\psi$  polarization analysis.

Trigger level	Trigger line	specific requirements
L0	<i>DiMuon</i>	$p_T(\mu_1) \times p_T(\mu_2) > 1.68 \text{ GeV}^2/c^2$
	<i>Muon</i>	$p_T(\mu) > 1.48 \text{ GeV}/c$
		Pass L0 <i>Muon</i> or L0 <i>DiMuon</i>
HLT1	<i>Hlt1DiMuonHighMass</i>	$p_T(\mu) > 0.5 \text{ GeV}/c$ and $P(\mu) > 6 \text{ GeV}/c$ for both muons $M_{\mu^+\mu^-} > 2700 \text{ MeV}/c^2$
HLT2	<i>Hlt2DiMuonJPsi</i>	Track quality $\chi^2/\text{ndof} < 5$ for both muons $M_{J/\psi} - 120 \text{ MeV}/c^2 < M_{\mu^+\mu^-} < M_{J/\psi} + 120 \text{ MeV}/c^2$ Di-muon vertex quality: $\chi^2/\text{ndof} < 25$

- $J/\psi$  candidate vertex quality:  $\text{Prob}(J/\psi \text{ vertex } \chi^2/\text{ndof}) > 0.5\%$

The clone killing is based on the idea that, if a muon track is the clone of another muon, they share very large fraction of the hits, so generally the flight direction of the two muons will be very close. A  $J/\psi$  candidate is the clone of another one if both of its decay product muons are clones of muons from another  $J/\psi$  decay. The result shows that the clone rate of  $J/\psi$  is around 1% in Monte Carlo and 2% in data. The clone rate in data is larger than in Monte Carlo due to higher track multiplicity.

It is further required that  $J/\psi$  are triggered by the dedicated muon triggers: L0 *Muon* or *DiMuon*, *Hlt1DiMuonHighMass* and *Hlt2DiMuonJPsi*. The triggers on muons generally filter muon(s) with large transverse momentum, and the Hlt2 trigger on  $J/\psi$  selects good  $J/\psi$  candidates in the specified mass range. The detailed requirements of the used triggers are listed in Table 4.2.

#### 4.1.1.1 Distinction between prompt $J/\psi$ and $J/\psi$ from $b$ -hadron decays

As discussed in section 1.2.3, apart from those produced from the PV and excited charmonium states,  $J/\psi$  can also be produced from  $b$ -hadron decays ( $J/\psi$  from  $b$ ) with typical fraction from a few percent at low  $p_T$  to above 50% in sufficiently large  $p_T$  regions. The  $c\tau$  of weakly decayed  $b$ -hadrons is about 0.45 mm and  $J/\psi$  decays shortly after production, so the  $J/\psi$  coming from  $b$  decay has a displaced vertex overlaying with the  $b$  vertex which is a few millimeters (averagely about 8 mm in LHCb acceptance) away from the PV. Thanks to the LHCb VELO detector, the secondary vertices can be reconstructed

with good resolution ( $\approx 0.2$  mm for  $J/\psi$ ), and they are thus significantly separated from the PV. To select prompt  $J/\psi$  candidates, the pseudo proper time  $t_z$  is constructed for each candidate to discriminate prompt candidate from those from  $b$ -hadron decay.  $t_z$  is calculated from the distance between the  $J/\psi$  decay vertex and the associated PV with the  $J/\psi$  momentum projected in the  $z$  direction of LHCb coordinate system as:

$$t_z = \frac{(z_{J/\psi \text{ vertex}} - z_{J/\psi \text{ PV}}) \times M(J/\psi)}{p_z(J/\psi)}, \quad (4-1)$$

where  $z_{J/\psi \text{ vertex}}$  and  $z_{J/\psi \text{ PV}}$  are the  $z$  component of the positions of the reconstructed  $J/\psi$  decay vertex and its associated PV respectively, while  $M(J/\psi)$  is the nominal  $J/\psi$  mass and  $p_z(J/\psi)$  is the  $J/\psi$  momentum along the  $z$  direction. Each  $J/\psi$  vertex has a  $7 \times 7$  covariance matrix associated defining the correlations and errors of the position ( $x, y, z$ ) and 4-momentum of the  $J/\psi$  candidate. It is assumed that the correlation between the PV position and the secondary vertex is negligible. From the covariance matrix, with Equation 4-1, the error of  $t_z$  can be calculated as  $\delta t_z$  by error propagation. For a prompt  $J/\psi$  event,  $\delta t_z$  mainly comes from the error of the  $z_{J/\psi \text{ vertex}}$  ( $\approx 0.2$  mm), while the contributions from the errors of  $z_{J/\psi \text{ PV}}$  ( $\approx 0.01$  mm) and  $p_z(J/\psi)$  ( $\delta p/p \sim 0.5\%$ ) are rather small. From the  $t_z$  and its error  $\delta t_z$ , the pseudo proper time significance  $\tau_S$  is defined as:

$$\tau_S = \frac{t_z}{\delta t_z}. \quad (4-2)$$

From this definition, the  $\tau_S$  can have both positive and negative values just as the  $t_z$  does for the  $J/\psi$  that decays shortly after the primary event production. Large  $|\tau_S|$  means that the  $J/\psi$  has a manifest displaced vertex.  $J/\psi$  mesons coming from  $b$  decay can have huge  $|\tau_S|$  value while prompt  $J/\psi$  usually has small  $|\tau_S|$ . However a prompt  $J/\psi$  meson can have very large  $|\tau_S|$  when the  $J/\psi$  meson is associated with a wrong PV due to the fact that the PV that the  $J/\psi$  really originates can not be reconstructed because of lack of enough tracks. These  $J/\psi$  mesons are called tail events because they fall onto the tails of the  $t_z$  distribution. The  $\tau_S$  distribution for signal events in data is shown in Figure 4.1. The prompt and delayed components are clearly identified by the peak centered around 0 ps and by the long tail in the positive  $\tau_S$  region respectively. On both sides, the tail component extends to very large  $\tau_S$  values. The fraction of events, in which the true PV is lost, is found to be below 1%<sup>[147]</sup>. Prompt  $J/\psi$  mesons are selected with the requirement  $|\tau_S| < 4$ . The study from Monte Carlo truth shows that this selection removes only 1-2%

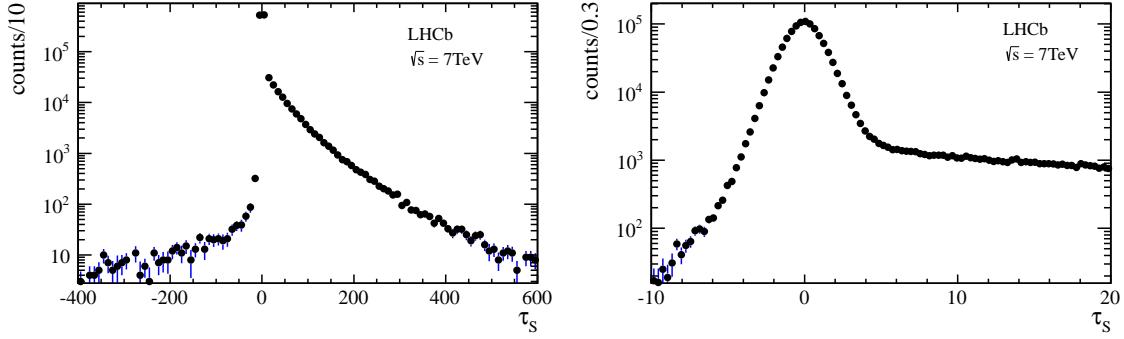


Figure 4.1  $\tau_s$  distribution in data (background subtracted) in a representative  $J/\psi$  kinematic bin. Plot on the right is a zoomed distribution around 0 ps.

of the prompt events and keeps only about 15% of the  $J/\psi$  produced in  $b$  decay.

#### 4.1.1.2 Fiducial region cut

It is found that the Monte Carlo doesn't simulate data very well in the outer boundary regions of the LHCb detector where muons with low momentum will fly outside of the LHCb acceptance quite easily. The pseudo rapidity ( $\eta = \frac{1}{2} \ln \left( \frac{p+p_z}{p-p_z} \right)$ ) distribution of muons are quite different between data and Monte Carlo in these regions. The region where Monte Carlo does not simulate the efficiency very well is discarded.

To identify the fiducial region, the distribution of track's transverse distance ( $r_{xy} = \sqrt{x^2 + y^2}$ ) from the beam in Monte Carlo is compared with the distribution in data. Here the track state means the position at any give point along the track trajectory. In the region before the LHCb dipole magnet the transverse distance at fixed  $z$  coordinate position is a measure of the pseudo rapidity. Figure 4.2 shows the transverse distance distribution at  $z = 1$  m just behind VELO for Monte Carlo and data, and their ratio. From the figure it can be seen that data is less efficient in the large  $r_{xy}$  area. The comparisons of the transverse distance distributions at other  $z$  coordinates have similar behavior. It is decided that  $J/\psi$  candidates with either of the two muons falling into the region  $r_{xy} > 220$  mm at  $z = 1$  m, which is almost equivalent to  $\eta < 2.2$ , are rejected. This cut will remove 50% of the signal  $J/\psi$  with rapidity less than 2.5, about 10% in the rapidity range  $2.5 < y < 3.0$  and almost does nothing to  $J/\psi$  with rapidity larger than 3, where the rapidity is defined as  $y = \frac{1}{2} \ln \left( \frac{E+p_z}{E-p_z} \right)$ .

With all these requirements, a sample of  $J/\psi$  events for the polarization measurements is selected. The events are divided into several bins of  $J/\psi$   $p_T$  and rapidity to study the dependence of polarization as a function of  $J/\psi$  kinematics. The binning scheme is

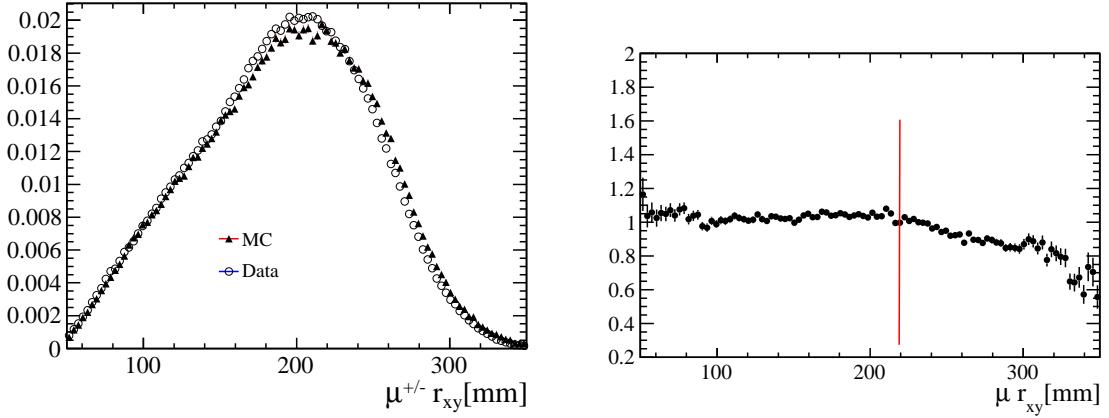


Figure 4.2 Track state transverse distance distribution for data and Monte Carlo (*left*) and the ratio of data over Monte Carlo (*right*) at  $z = 1$  m.

chosen such that in each bin the efficiency (see the discussions in appendix A.1) is approximately uniform as a function of  $J/\psi$   $p_T$  and rapidity, which requires the bin size is not too large. On the other hand, the statistics in each bin should be sufficient to perform a reasonable analysis. It is also known that the global efficiency in the region  $p_T < 2$  GeV/ $c$  is very low because of the  $p_T$  dependent triggers on muons, and the efficiency is especially low if the  $J/\psi$  rapidity is small as well because one of the muons will have rather small momentum or even fly in the opposite of LHCb acceptance in the laboratory frame. Besides, as can be seen from Figure 4.3, the events in this region accumulate in a small phase space around  $\cos \theta = 0$  in the muon polar angle distribution in  $J/\psi$  rest frame, and the events with  $\cos \theta = 0$  do not contribute to polarization statistically significantly. In addition, the background level is much higher in the low  $p_T$  region, resulting in much larger uncertainties from background subtraction. So the region where  $p_T$  is less than 2 GeV/ $c$  is discarded for the analysis. Finally, thirty bins—six bins in transverse momentum and five bins in rapidity—are used with the following bin schemes:

$$\begin{aligned} p_T : & (2, 3], (3, 4], (4, 5], (5, 7], (7, 10], (10, 15] \text{ GeV}/c, \\ y : & (2.0, 2.5], (2.5, 3.0], (3.0, 3.5], (3.0, 4.0], (4.0, 4.5]. \end{aligned} \quad (4-3)$$

#### 4.1.1.3 Background subtraction

In Figure 4.4, the mass distribution for the selected events in the  $J/\psi$  kinematic bin  $5 < p_T < 7$  GeV/ $c$ ,  $3.0 < y < 3.5$  is shown, and a fit result to the distribution is superimposed. In the fit the signal component of the mass distribution is parameterized by a Crystal Ball function (CB), composed of a gaussian function describing the mass

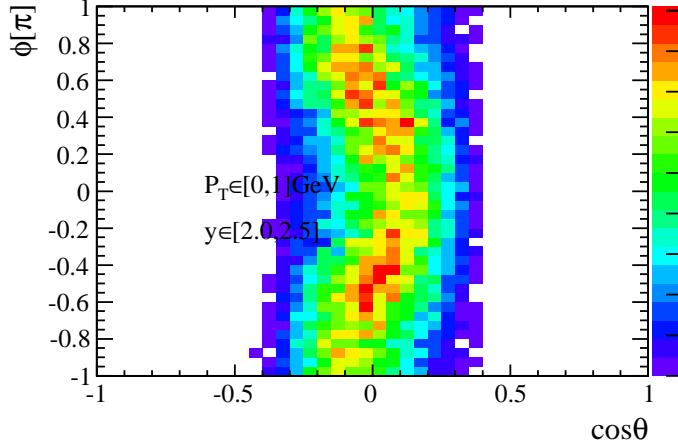


Figure 4.3 Two dimension distribution of  $\cos \theta - \phi$  in the low  $p_T$  small rapidity region.

resolution and a power-law tail describing the radiative decay of  $J/\psi$  with final state soft photons not reconstructed. The CB function is parameterized as:

$$F_{CB}(x; \mu, \sigma, \alpha, n) = \begin{cases} \frac{\left(\frac{n}{|\alpha|}\right)^n \exp(-\frac{1}{2}\alpha^2)}{\left(\frac{n}{|\alpha|} - |\alpha| - \frac{x-\mu}{\sigma}\right)^n} & \text{if } \frac{x-\mu}{\sigma} < -|\alpha| \\ \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] & \text{otherwise} \end{cases} \quad (4-4)$$

where  $\mu$  and  $\sigma$  define the central value (mean) and resolution of the gaussian component,  $\alpha$  determines where is the connection between the power-law tail and the gaussian, and  $n$  describes the shape of mass distribution due to missing photon energy. The parameters  $n$ ,  $\sigma$  and  $\mu$  of the Crystal Ball shape are let free in the fit while the  $\alpha$  parameter is chosen to be a quadratic function of  $\sigma$ :  $\alpha = 2.22 + 0.004\sigma - 0.0004\sigma^2$  where the coefficients are extracted from Monte Carlo studies. The combinatorial background is described by an exponential function, which turns to be approximately linear. The fitted parameters for the CB and exponential functions can be found in Table B.1 in appendix B.1.

In order to subtract the background, from the mass distribution, the signal region, which is  $[\mu - 3\sigma, \mu + 3\sigma]$  and two sideband regions, which is  $[\mu - 7\sigma, \mu - 4\sigma]$  (left sideband) and  $[\mu + 4\sigma, \mu + 7\sigma]$  (right sideband), are defined. The resolution  $\sigma$  is between 15  $\text{MeV}/c^2$  and 20  $\text{MeV}/c^2$  for most of the bins. In the sideband regions the events are predominantly combinatorial background, while in the signal region the events are mixture of signal events and background events with number of the same level — the number of signal events is several times of the number of background events in the signal region depending on the bins. It should be noted that the widths of the sideband regions adds up to be the same as the width of signal region.

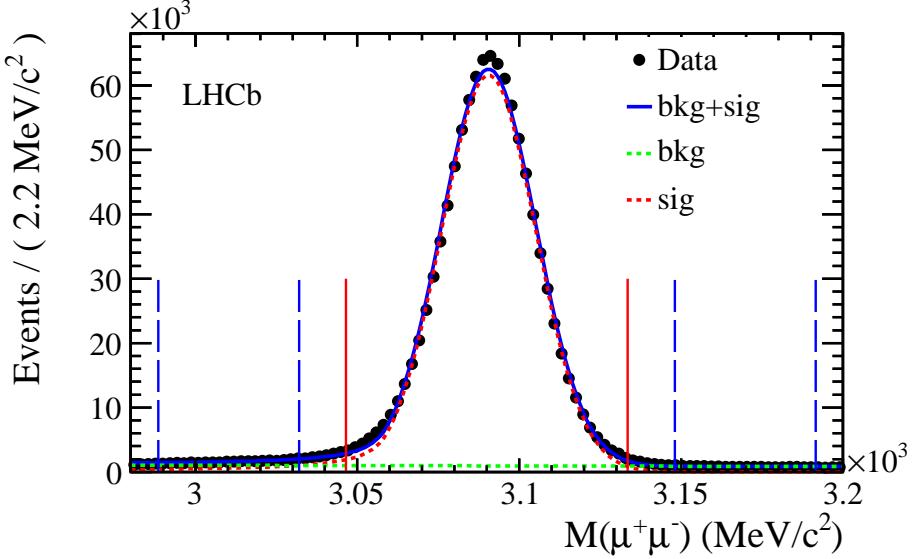


Figure 4.4 Mass distribution of the  $J/\psi$  candidates in the bin  $5 < p_T < 7 \text{ GeV}/c$ ,  $3.0 < y < 3.5$ . The parametrization of the distribution with a CB plus exponential is superimposed. The signal region (region between solid vertical lines) and two sideband regions (regions between dashed vertical lines) are also illustrated.

## 4.2 Polarization Fit

### 4.2.1 Efficiency

Following the studies in the  $J/\psi$  cross section measurement at LHCb<sup>[147]</sup>, the detection efficiencies can be divided into three components: the geometrical acceptance ( $\epsilon_{\text{geo}}$ ), the reconstruction and selection efficiency ( $\epsilon_{\text{rec\&sel}}$ ) and the trigger efficiency ( $\epsilon_{\text{trg}}$ ), and their definitions are expressed as:

$$\begin{aligned}
 \epsilon_{\text{tot}} &= \epsilon_{\text{geo}} \times \epsilon_{\text{rec\&sel}} \times \epsilon_{\text{trg}} \\
 \epsilon_{\text{geo}} &= \frac{N_{J/\psi} \text{ with both } \mu \text{ in LHCb acceptance}}{N_{J/\psi} \text{ generated}} \\
 \epsilon_{\text{rec\&sel}} &= \frac{N_{J/\psi} \text{ reconstructed and selected}}{N_{J/\psi} \text{ with both } \mu \text{ in LHCb acceptance}} \\
 \epsilon_{\text{trg}} &= \frac{N_{J/\psi} \text{ selected and triggered by dedicated trigger lines}}{N_{J/\psi} \text{ reconstructed and selected}}
 \end{aligned} \tag{4-5}$$

where the LHCb acceptance is the [10, 400] mrad region with regard to the beam in the  $+z$  direction for charged particles. The selections and trigger lines have been described in section 4.1.1.

The efficiencies for the three components are determined using the simulated unpo-

larized  $J/\psi$  sample, by counting the number of events for generated, geometry accepted, reconstructed/selected and triggered  $J/\psi$  in each  $p_T$  and rapidity bin. The geometry acceptance, selection and trigger efficiencies are not uniformly distributed as a function of the muon angles  $\cos \theta$  and  $\phi$ .  $\epsilon_{\text{tot}}$  generally depends on both angles,  $\epsilon_{\text{tot}} = \epsilon_{\text{tot}}(\cos \theta, \phi)$ . Besides, the shape of the efficiency as a function of  $(\cos \theta, \phi)$  is also different in different bins of  $J/\psi p_T$  and rapidity. The non-uniformity of the efficiencies results from the fact that at least one of the decay product muons is less easily falling into LHCb acceptance or fulfilling the  $p_T$  requirement at the selection or trigger stage in some area of the  $(\cos \theta, \phi)$  phase space. In Figure 4.5 the global efficiencies in the HX frame as a function of  $\cos \theta$  and  $\phi$  for the two  $p_T$  bins  $3 < p_T < 4 \text{ GeV}/c$  and  $7 < p_T < 10 \text{ GeV}/c$  for various rapidity bins are shown. It can be seen that the detection efficiency is lower when  $\cos \theta \approx \pm 1$  because in this case one of the two muons will fly backwards in  $J/\psi$  rest frame, and has relatively smaller momentum in the laboratory frame. By definition the  $x$ -axis of the HX frame points outside LHCb acceptance as can be see from Figure 4.6, so when  $\phi \approx 0$  or  $\pi$ , one of the muons leaves the LHCb acceptance easily, and thus the efficiency is relatively lower in these regions, especially for the lowest rapidity and the highest rapidity bins (in the same  $p_T$  bin) where the one of the muons flies near the 400 mrad and 10 mrad boundaries respectively. Appendix A.1 discusses how the geometry acceptance, the reconstruction&selections and the triggers affect the shapes of the efficiency as a function of  $(\cos \theta, \phi)$  in detail.

As a result, the total efficiency is actually studied in four kinematic variables:  $p_T$  and  $y$  of the  $J/\psi$  and  $\cos \theta$  and  $\phi$  of the muon in  $J/\psi$  rest frame. For a  $J/\psi$  decay, there are five degrees of freedom— $J/\psi p_T$ , rapidity and azimuthal angle in laboratory frame, the  $\cos \theta$  and  $\phi$  of the muon in  $J/\psi$  rest frame—that completely determines the global efficiency with or without polarization. However the efficiency dependence on azimuthal angle of  $J/\psi$  is almost uniform, so the  $J/\psi$  azimuthal angle is integrated. Besides, because the analysis is performed with the binning in  $J/\psi p_T$  and rapidity, and in the leading order it is assumed that the binning is so small that the shape of the efficiency as a function of  $\Omega(\cos \theta, \phi)$ , and the size of the polarization are the same for different  $p_T$  and  $y$  values in the same bin. So in the end, within one particular  $J/\psi$  kinematic bin, the efficiency is only a function of the muon angular variables  $\cos \theta$  and  $\phi$  with or without polarization.

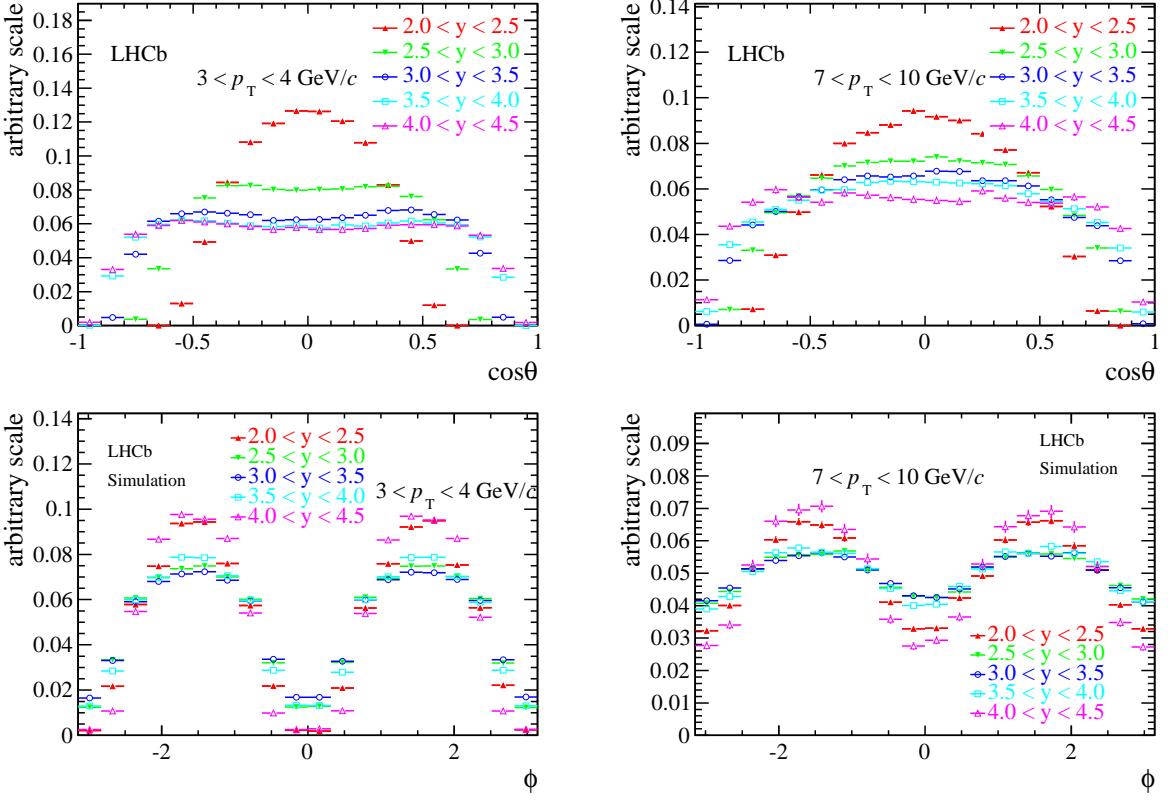


Figure 4.5 Global detection efficiency as a function of  $\mu^+ \cos \theta$  (up) and  $\phi$  (down) for  $3 < p_T < 4 \text{ GeV}/c$  (left) and for  $7 \text{ GeV}/c < p_T < 10 \text{ GeV}/c$  (right). Different rapidity bins are shown with different markers.

#### 4.2.2 Likelihood estimator using event weights

Following the way described in section 3.2, the weighted logarithm likelihood estimator is constructed as:

$$\log \mathcal{L} = \sum_{i=1}^{N_{\text{tot}}} w(m_i) \times \log \left[ \frac{P(\cos \theta_i, \phi_i | \lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi)}{\text{Norm}(\lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi)} \right], \quad (4-6)$$

where

$$P(\cos \theta, \phi | \lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi) = (1 + \lambda_\theta \cos^2 \theta + \lambda_{\theta\phi} \sin 2\theta \cos \phi + \lambda_\phi \sin^2 \theta \cos 2\phi)$$

and

$$\begin{aligned} \text{Norm}(\lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi) = & \int d\Omega \epsilon_{\text{tot}}(\Omega) + \lambda_\theta \int d\Omega \cos^2 \theta \epsilon_{\text{tot}}(\Omega) \\ & + \lambda_{\theta\phi} \int d\Omega \sin 2\theta \cos \phi \epsilon_{\text{tot}}(\Omega) + \lambda_\phi \int d\Omega \sin^2 \theta \cos 2\phi \epsilon_{\text{tot}}(\Omega). \end{aligned} \quad (4-7)$$

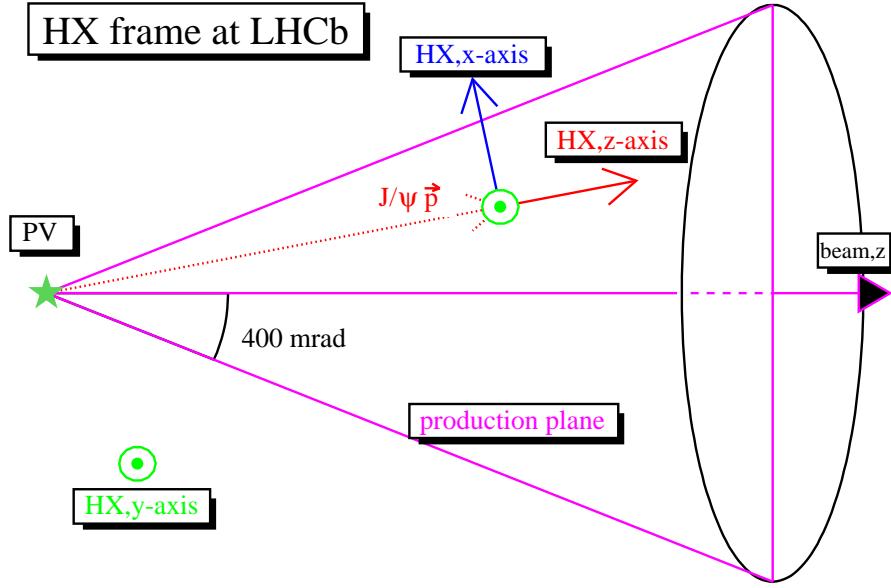


Figure 4.6 Schematic figure illustrating the definition of the HX frame at LHCb. The figure shows the production plane (lies in the screen), the direction of the  $x$ -axis (pointing outside of LHCb acceptance in the production plane),  $y$ -axis (pointing outside of the plane),  $z$ -axis (the direction of  $J/\psi$  momentum) of the HX frame, and the LHCb 400 mrad geometry acceptance.

To subtract the background, the weight  $w$  is chosen to be  $+1$  ( $-1$ ) if the  $J/\psi$  mass falls into signal (sideband) region(s), and zero otherwise. Here uniform weights for events in signal region ( $+1$ ) and background regions ( $-1$ ) are used, and to make sure the background contribution is correctly subtracted the following assumptions should be satisfied:

1. The combinational background mass distribution is linear. So the number of  $J/\psi$  background candidates with  $+1$  weights will be the same as the number of  $J/\psi$  background candidates with  $-1$  weights when the width of signal region is identical to the total width of sideband regions which is the case in the analysis. Otherwise, another weight value for sideband events should be introduced to make the number of events in the sideband regions normalized to the number of background events in signal region. If a common weight value is used for the events in signal region, to make sure the total weights for signal events equal to the number of signal events, the weight has to be  $+1$ :

$$\sum_{i=1}^{N_s} w(m_i) = N_s \Rightarrow w(m_i) = 1$$

From Table B.1 it can be seen that the parameter for the exponential is small enough

to justify the linear background assumption.

2. The  $\Omega(\cos \theta, \phi)$  distribution of the background events in the signal region equals to the distribution in sideband regions. A careful study shows that  $\Omega(\cos \theta, \phi)$  distribution of the background events is weakly dependent on the reconstructed invariant mass. This point has been deeply investigated and is discussed in details in appendix A.2. In conclusion, the  $\cos \theta$  ( $\phi$ ) distribution of background events in signal region lies between the distributions of background events in the left sideband and right sideband when they are normalized to the same area, and thus, by combining the left and right sidebands, the background distribution in signal region can be well reproduced. In the systematic uncertainty studies, only the left or the right sideband (with proper weights in the likelihood function) is used to subtract the background (see section 4.3.5), and it will give a conservative estimator of the bias of this assumption.

The normalization in Equation 4-6 is characterized by three constants  $(a, b, c)$  (see Equation 3-12) and they are determined from the unpolarized Monte Carlo sample. The calculated sets of  $(a, b, c)$  in each  $J/\psi$  kinematic bin can be found in Table B.3 in Appendix B. Some remarks can be made on the normalization constants. The factor  $a$  represents the average value of  $\cos^2 \theta$ , and because in lower  $p_T$  bins the events are more centered around  $\cos \theta \approx 0$  as can be seen from Figure 4.5,  $a$  is smaller in this region than in higher  $p_T$  bins. Since  $\phi$  accumulates around  $\pm \frac{\pi}{2}$  where  $\cos(\pm \frac{\pi}{2} \times 2) < 0$ ,  $c$  is always less than zero.

The constructed estimator is maximized with regard to the three polarization parameters  $\lambda_\theta$ ,  $\lambda_{\theta\phi}$  and  $\lambda_\phi$  with the TMinuit program. The parameters that maximize the estimator describe the  $J/\psi$  polarization.

### 4.3 Uncertainties on the polarization

In this section the uncertainties associated with each of the parameters will be evaluated in detail. The methods of extracting the uncertainties are common for both the HX frame and the CS frame, however for simplicity only the numbers calculated in the HX frame are shown in this section. For results in the CS frame, refer to the Tables B.10, B.11, B.12, B.13 and B.14 for details. *In the following, for the numbers in the summary tables, although for most time two or three digits will be proper, four digits are adopted uniformly to show values that are small and to make it easier to compare results in different bins.*

### 4.3.1 Background subtraction fluctuation

When TMinuit is used to maximize the estimator, the package will generate a set of parameters and their errors, which will be called likelihood uncertainties in the following. The uncertainties are calculated by investigating the property of the likelihood estimator as a function of the parameters around the best values. The best fit values ( $\hat{\lambda}$ ) are given by the equation<sup>[187]</sup>:

$$\frac{\partial \log \mathcal{L}}{\partial \lambda_i} = 0$$

and the errors are read from the covariance matrix which is the inverse of the Hessian matrix multiplied by -1. The Hessian matrix element ( $H_{ij}$ ) is the second derivative of log-likelihood  $\log \mathcal{L}$  with regards to  $\lambda$  at the best fit  $\hat{\lambda}$ :

$$H_{ij} = \frac{\partial^2 \log \mathcal{L}}{\partial \lambda_i \partial \lambda_j} \Big|_{\lambda_i = \hat{\lambda}_i, \lambda_j = \hat{\lambda}_j}, \quad (4-8)$$

where  $\lambda_i$  is  $\lambda_\theta$ ,  $\lambda_{\theta\phi}$  or  $\lambda_\phi$  in the analysis.

These likelihood uncertainties are statistical errors. However, considering this particular analysis, the errors returned by the likelihood estimator do not include all the statistical fluctuations correctly. In the following the fit with estimator using the standard Monte Carlo and real data is called as baseline fit or nominal fit. In general, the estimation of the parameters' statistical uncertainties by a weighted likelihood through Equation 4-8 is not correct; for example, one would expect that  $2 \times \log \mathcal{L}$  will give the same parameter errors as  $\log \mathcal{L}$  does. However if the errors are given by the Hessian matrix in Equation 4-8, they will be different. Equation 4.3.1 and 4-8 also show that if linear terms as a function of the parameters are added to the logarithm likelihood, the best fit values of parameters will change accordingly, however the Hessian matrix and thus the parameter errors will stay the same, because the second derivative of linear functions are zero.

The likelihood estimator of Equation 4-6 is a combination of two parts, the likelihood with signal events and the likelihood with background events:

$$\begin{aligned} \log \mathcal{L} &= \log \mathcal{L}_{\text{events in signal region}} - \log \mathcal{L}_{\text{events in sideband regions}} \\ &= \log \mathcal{L}_{\text{signal}} + \log \mathcal{L}_{\text{background in signal region}} - \log \mathcal{L}_{\text{background in sideband regions}} \quad (4-9) \\ &\equiv \log \mathcal{L}_{\text{signal}} + \log \mathcal{L}_{\text{background}} \end{aligned}$$

The latter one,  $\log \mathcal{L}_{\text{background}}$ , which is constructed from background events in signal region with weight +1 and background events in sideband regions with weight -1, will

generate polarization consistent with zero within statistical errors, provided that the two assumptions for the combinatorial background hold as described in section 4.2.2.

To study the behavior of  $\log \mathcal{L}_{\text{background}}$  around the best fit  $\lambda$  parameters, a toy background sample is randomly generated following the  $\cos \theta - \phi$  distribution of the events in the sideband regions of data, and this toy sample is used to replace the *background in signal region* component in the  $\log \mathcal{L}_{\text{background}}$ . Figure 4.7 shows that  $\log \mathcal{L}_{\text{background}}$  is quite linear around the best fitted parameters: the coefficients of those terms with high orders in  $\lambda$  ( $\lambda_\theta^2$  for example) are very small. It is also found that the coefficients of the terms that mix different  $\lambda$  parameters ( $\lambda_\theta * \lambda_\phi$  for example), which can produce off-diagonal Hessian matrix element are very small too. Following the discussions above, because the  $\log \mathcal{L}_{\text{background}}$  is a linear function of the parameters, it does not contribute at all to the estimation of the parameters uncertainties by way of the Hessian matrix—the errors returned by the estimator of Equation 4-6 only correspond to the signal contribution.

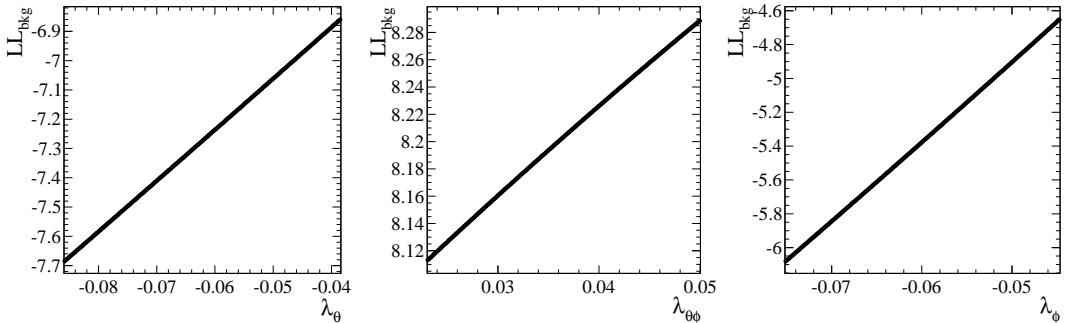


Figure 4.7 –  $-\log \mathcal{L}_{\text{background}}$  as a function of  $\lambda_\theta$  (left),  $\lambda_{\theta\phi}$  (middle) and  $\lambda_\phi$  (right) around the best fit values.

However, the background subtraction does introduce fluctuations to the estimation of the parameters, because the  $\cos \theta - \phi$  distribution of background events in the signal region are statistically different from those in the sideband regions because of statistical fluctuation.  $\log \mathcal{L}_{\text{background}}$  is the contribution of one background sample (in sideband regions) subtracted by the other (in signal region), so two background samples can be created artificially, with +1 weight and -1 weight respectively, to mimic the influence of  $\log \mathcal{L}_{\text{background}}$  to the uncertainties of the estimation of parameters. The two toy background samples, named ‘background1’ and ‘background2’, are generated following the muon angular distribution in the sidebands using a toy Monte Carlo technique (the ran-

dom number seeds of the two backgrounds are different). The angular ( $\cos \theta, \phi$ ) distributions in both the additional samples are statistically the same as the shape of ( $\cos \theta, \phi$ ) of background events. The two additional components ( $\log \mathcal{L}_{\text{background1}}$  and  $\log \mathcal{L}_{\text{background2}}$ ) constructed from the two toy Monte Carlo background samples are added to the nominal likelihood, and the two additional terms form the term  $\log \mathcal{L}'_{\text{background}}$  in the modified likelihood:

$$\begin{aligned}\log \mathcal{L}' &= \log \mathcal{L}_{\text{nominal}} + \log \mathcal{L}_{\text{background1}} - \log \mathcal{L}_{\text{background2}} \\ &\equiv \log \mathcal{L}_{\text{nominal}} + \log \mathcal{L}'_{\text{background}}.\end{aligned}\quad (4-10)$$

With the additional terms the best fit values of the parameters will change. The deviations of new results from the nominal values come from the additional background subtraction term  $\log \mathcal{L}'_{\text{background}}$  and they are similar to the fluctuations in the real background subtraction  $\log \mathcal{L}_{\text{background}}$ . Because the standard errors given by the maximization program (through the Hessian matrix way) does not take into account the background subtraction fluctuation, which is also true for the modified likelihood  $\log \mathcal{L}'$ , the errors given by the modified estimator of Equation 4-10 do not include the fluctuations due to the  $\log \mathcal{L}'_{\text{background}}$ , i.e. they should be the same as the ones given by the standard estimator (the estimator without  $\log \mathcal{L}'_{\text{background}}$ ).

The whole procedure—constructing the background likelihood  $\log \mathcal{L}'_{\text{background}}$  from toy Monte Carlo sample and extracting the polarization with the modified estimator of Equation 4-10—is done 100 times for each  $J/\psi$   $p_T$  and rapidity bin, and each time a new set of best fit parameters are extracted. In Figure 4.8, the 100 results for each  $\lambda$  parameter for one particular  $J/\psi$  kinematic bin are shown. It can be seen that the 100 new parameters follow a gaussian distribution, and the  $\sigma$  of gaussian is quoted as background subtraction uncertainty for each parameter. Table 4.3 lists these errors for all  $J/\psi$   $p_T$  and  $y$  bins.

The errors of each parameter in the 100 fits can also be investigated by comparing them with the nominal errors. In Figure 4.9 their relative difference, calculated as the difference of the errors divided by the nominal error, is plotted for each parameter. It can be seen that the errors given by the modified likelihood estimator with additional terms are very similar (less than 1% different) to the ones returned by the nominal estimator. This confirms that the fluctuations due to background subtraction are not included in the parameter errors given directly by the program. Because the signal to background ratio in data is larger than one (see Figure 4.4), the fluctuations coming from background subtrac-

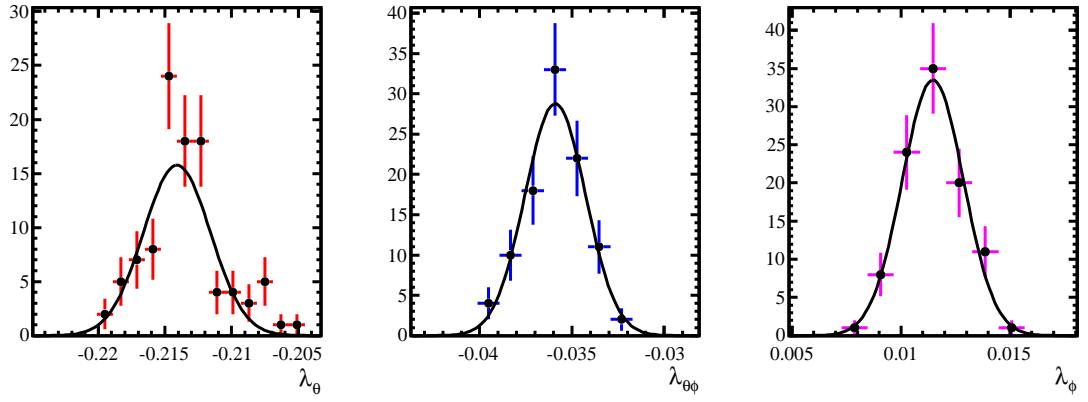


Figure 4.8 Distributions of  $\lambda_\theta$  (left),  $\lambda_{\theta\phi}$  (middle) and  $\lambda_\phi$  (right) parameters from the 100 fits with modified estimator (Equation 4-10). A gaussian fit to each distribution is superimposed. One arbitrary bin of  $J/\psi$  kinematics is shown without losing generality.

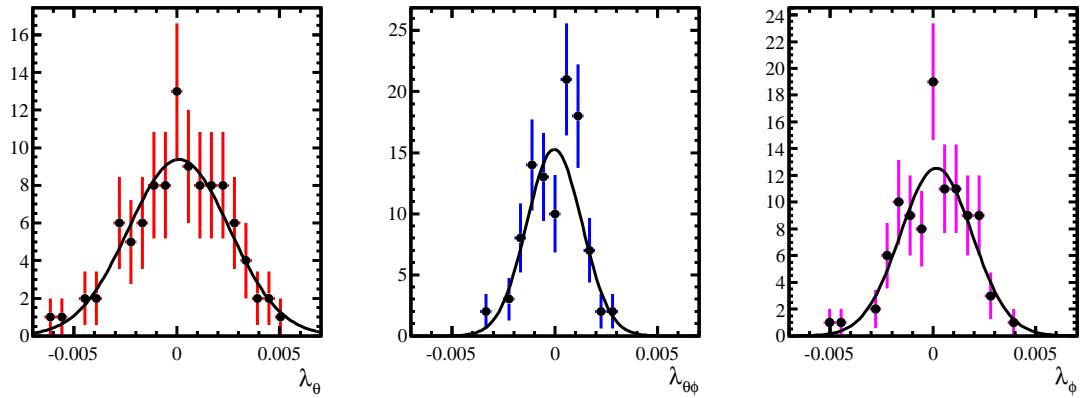


Figure 4.9 The relative difference of the errors between nominal fit and the 100 fits with modified estimators for  $\lambda_\theta$  (left),  $\lambda_{\theta\phi}$  (middle) and  $\lambda_\phi$  (right) respectively. A gaussian fit to each distribution is superimposed. One arbitrary bin of  $J/\psi$  kinematics is shown without losing generality.

Table 4.3 The statistical error coming from background subtraction for each polarization parameter in each  $J/\psi$  kinematic bin in the HX frame.

$p_T$ (GeV/c)		y bin			
$2 < p_T < 3$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0451	0.0050	0.0037	0.0025	0.0027
$\lambda_{\theta\phi}$	0.0344	0.0029	0.0016	0.0015	0.0025
$\lambda_\phi$	0.0110	0.0016	0.0010	0.0012	0.0015
$3 < p_T < 4$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0368	0.0042	0.0016	0.0017	0.0033
$\lambda_{\theta\phi}$	0.0244	0.0021	0.0012	0.0014	0.0015
$\lambda_\phi$	0.0091	0.0016	0.0010	0.0011	0.0016
$4 < p_T < 5$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0200	0.0033	0.0025	0.0016	0.0037
$\lambda_{\theta\phi}$	0.0170	0.0017	0.0013	0.0011	0.0015
$\lambda_\phi$	0.0064	0.0009	0.0008	0.0010	0.0014
$5 < p_T < 7$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0092	0.0028	0.0021	0.0022	0.0047
$\lambda_{\theta\phi}$	0.0070	0.0010	0.0011	0.0013	0.0017
$\lambda_\phi$	0.0032	0.0007	0.0006	0.0007	0.0016
$7 < p_T < 10$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0087	0.0025	0.0023	0.0025	0.0055
$\lambda_{\theta\phi}$	0.0057	0.0012	0.0011	0.0014	0.0030
$\lambda_\phi$	0.0021	0.0007	0.0007	0.0009	0.0016
$10 < p_T < 15$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0086	0.0030	0.0078	0.0083	0.0178
$\lambda_{\theta\phi}$	0.0066	0.0021	0.0021	0.0033	0.0081
$\lambda_\phi$	0.0020	0.0012	0.0014	0.0021	0.0061

tion are always smaller than the likelihood errors. Their ratio decreases with increased  $p_T$  and rapidity because the background fractions are much smaller in higher  $p_T$  and rapidity regions.

### 4.3.2 Fluctuation of the normalization

In the likelihood estimator, the uncertainty of the normalization will introduce a bias to the polarization determination. The normalization constants, as shown in Equation 3-12, are calculated from the Monte Carlo — starting from one Monte Carlo sample, three particular values of the normalization constants  $a$ ,  $b$  and  $c$  can be obtained. The limited Monte Carlo statistics introduces uncertainties to these three constants. To study how the normalization constants fluctuate, many toy Monte Carlo samples are generated randomly, and for each sample new normalization constants  $a$ ,  $b$  and  $c$  are computed for a new polarization fit. The constants  $a$ ,  $b$  and  $c$  vary from one sample to another, so are the polarization parameters. In summary, in each bin of  $J/\psi$   $p_T$  and rapidity, the following steps are performed:

- A Fill the two-dimensional distribution  $\cos \theta - \phi$  of the standard Monte Carlo (the base Monte Carlo) into a histogram;
- B Generate  $N$  events randomly. Each event has the  $\cos \theta - \phi$  value sampled from the 2-D histogram above, where  $N$  is the number of events in the standard Monte Carlo. By this step a toy Monte Carlo sample is produced;
- C Use the toy Monte Carlo sample to calculate the normalization constants  $a$ ,  $b$  and  $c$  and then to extract new numbers of polarization parameters. In this step, real data is the same as in the standard fit.

The whole procedure is repeated 100 times resulting in a group of polarization parameters for each  $J/\psi$  kinematics bin. In Figure 4.10, for a particular bin, the new fit results divided by the standard errors with the mean shifted to zero are shown for the three parameters. In Figure 4.11, the mean of the 100 fit results subtracted by the nominal polarization results is shown for each bin of  $J/\psi$  kinematics. It can be seen that the distributions of the 100 fit results closely follow a gaussian shape and their means are very close to the nominal results. It is also found that the likelihood errors of each parameter in the 100 fits are almost the same as the error in the nominal fit (less than 0.5% difference), which means that the variations of the normalization constants do not change the estimation of the statistical errors in the likelihood estimator.

To estimate the uncertainties of the measurements coming from the fluctuations of the normalization, a gaussian function is fitted to the 100 results, and the sigma of the gaussian is quoted. In Table 4.4 these uncertainties are listed for each  $p_T$  and rapidity bin. The size of the fluctuations is directly related to statistics in the Monte Carlo. Because the statistics of the Monte Carlo is usually one fourth of the signal statistics in data (depending on  $J/\psi$  kinematic bins), these uncertainties are usually two times of the statistical errors returned directly by the estimators, as suggested by intuition.

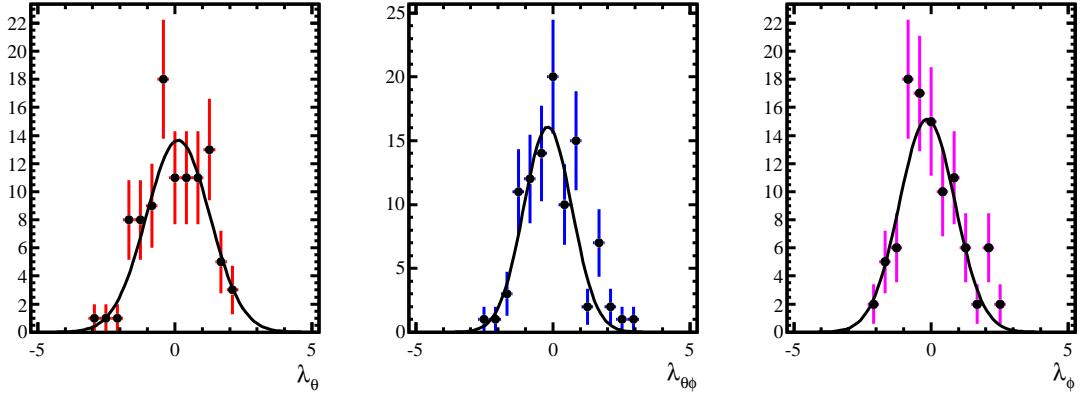


Figure 4.10 The pull distribution of  $\lambda_\theta$  (left),  $\lambda_{\theta\phi}$  (middle) and  $\lambda_\phi$  (right). The pull is calculated as the fit result from toy Monte Carlo subtracted by the result in the standard fit and then divided by the likelihood error for each parameter. A gaussian fit to each distribution is superimposed. One arbitrary bin of  $J/\psi$  kinematics is shown without losing generality.

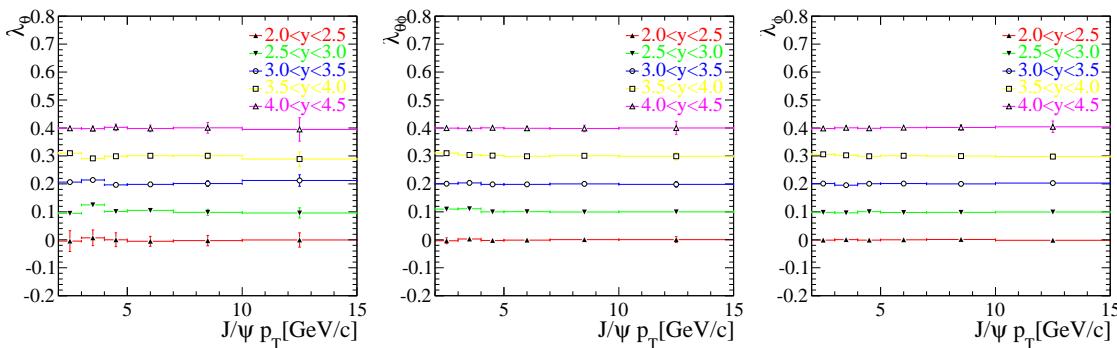


Figure 4.11 The mean values of the 100 fits in the toy Monte Carlo subtracted by the standard fit result are plotted for  $\lambda_\theta$  (left),  $\lambda_{\theta\phi}$  (middle) and  $\lambda_\phi$  (right) in each  $J/\psi$   $p_T$  and rapidity bin. To make the values in different bins clearly visible, they are offset according to the  $y$  bin as 0.1 for the second  $y$  bin, 0.2 for the third  $y$  bin, etc. The error for each value is the likelihood uncertainty in the standard fit.

Table 4.4 The uncertainty coming from the fluctuations of the normalization for each polarization parameters in each  $J/\psi$  kinematic bin in the HX frame.

$p_T$ (GeV/c)		$y$			
$2 < p_T < 3$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.1269	0.0097	0.0060	0.0051	0.0150
$\lambda_{\theta\phi}$	0.0576	0.0043	0.0019	0.0031	0.0107
$\lambda_\phi$	0.0212	0.0020	0.0020	0.0028	0.0096
$3 < p_T < 4$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.1196	0.0104	0.0071	0.0077	0.0125
$\lambda_{\theta\phi}$	0.0629	0.0055	0.0039	0.0037	0.0088
$\lambda_\phi$	0.0213	0.0032	0.0024	0.0034	0.0147
$4 < p_T < 5$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0824	0.0158	0.0108	0.0126	0.0218
$\lambda_{\theta\phi}$	0.0549	0.0065	0.0061	0.0094	0.0124
$\lambda_\phi$	0.0206	0.0042	0.0040	0.0061	0.0126
$5 < p_T < 7$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0495	0.0123	0.0126	0.0123	0.0222
$\lambda_{\theta\phi}$	0.0436	0.0059	0.0053	0.0060	0.0116
$\lambda_\phi$	0.0194	0.0036	0.0039	0.0060	0.0113
$7 < p_T < 10$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0638	0.0185	0.0191	0.0202	0.0286
$\lambda_{\theta\phi}$	0.0347	0.0105	0.0077	0.0100	0.0243
$\lambda_\phi$	0.0163	0.0073	0.0083	0.0107	0.0167
$10 < p_T < 15$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0719	0.0294	0.0359	0.0416	0.0447
$\lambda_{\theta\phi}$	0.0477	0.0179	0.0144	0.0241	0.0297
$\lambda_\phi$	0.0265	0.0120	0.0158	0.0162	0.0354

### 4.3.3 Tracking efficiency

The efficiency which is used to compute the normalization is totally determined from the Monte Carlo simulation as discussed in section 3.2.2. At LHCb many studies show that Monte Carlo does not simulate the tracking — the charged track finding and reconstruction — in data perfectly. The overall tracking efficiency in data and Monte Carlo agrees within  $100 \pm 2\%$  for early 2011 data — the data set used for the analysis. The difference in muon tracking efficiency will introduce bias to the polarization measurement if and only if the difference is not uniform as a function of muon kinematics. In fact from the studies of the tracking group at LHCb, the difference of the tracking efficiency between Monte Carlo and data really depends on muon kinematics (see Figure 4.12).

The difference of efficiency in different muon kinematic bins provided by the tracking group can be used to study how the tracking in simulation can bias the measurements. Instead a similar bias, which is not uniform as a function of muon kinematics and of same size as the measured difference between Monte Carlo and data, can be artificially introduced to the nominal Monte Carlo tracking efficiency by weighting on muons kinematics. As there are two muons,  $\mu^+$  and  $\mu^-$  for each  $J/\psi$ , so the weight for each  $J/\psi$  is the product of the weight for the positive muon and the weight for the negative muon. In the second way when the bias between data and Monte Carlo is artificially created, the weights are chosen to be dependent on muon pseudo rapidity or momentum. In the following, several models (shapes) which describe how the difference of tracking efficiency between data and simulation depends on muon kinematics are described.

#### 4.3.3.1 Weighting according to efficiency difference table

By the LHCb tracking group, the relative tracking efficiency (ratio) between data and Monte Carlo is calculated in different muon  $P - \eta$  bins, as shown in Figure 4.12. With these numbers provided, the Monte Carlo is weighted when the normalization constants are calculated, and the polarization parameters are extracted again using the weighted Monte Carlo. It is found that the deviations of the new results from the nominal ones are rather small for most of the bins. However the study is not sufficient to assess the bias due to the tracking effect because the numbers provided by the tracking group have relatively large errors and they do not cover all the muon kinematic range in our analysis.

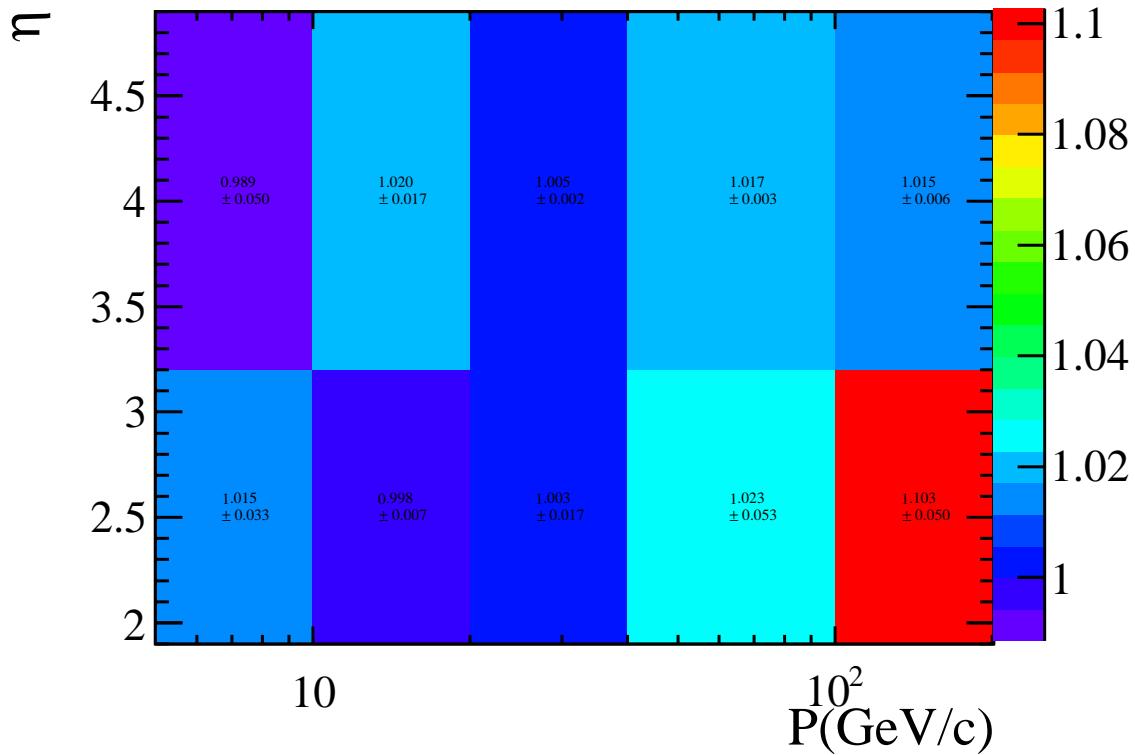


Figure 4.12 The tracking efficiency in data divided by the tracking efficiency in Monte Carlo for different muon momentum and pseudo rapidity bins.

#### 4.3.3.2 Weighting on muon pseudo rapidity

In this section, the weight for each muon is chosen as a function of muon pseudo rapidity ( $\eta$ ). Three models are used for the function, linear shape, parabolic shape and  $\Lambda$  shape. Since the LHCb detector has almost a geometry coverage between 10 mrad and 400 mrad which corresponds to a range between  $\eta = 2$  and  $\eta = 5$ , so the weighting functions take muon  $\eta$  in the range  $2.0 < \eta < 5.0$  as the argument.

**Linear function** For the linear dependence case, the weight for each muon as a function of its pseudo rapidity is defined as:

$$c_1 * (1 + c_2 * \eta_\mu),$$

where  $c_1$  and  $c_2$  are two constant numbers which determine the size of the bias introduced. Because only non uniform weighting functions change the measurements while a global weighting factor is meaningless, by properly selecting  $c_1$  and  $c_2$  the value of the function (weight) is fixed to 1 at  $\eta = 5$  and the weight at  $\eta = 2.0$  is 0.9, 0.96 or 0.98 to study

the relative tracking efficiency bias of 5%, 2% or 1% level respectively. The situation when the weight is chosen to be 1 at  $\eta = 5$  and 0.96 at  $\eta = 2$  is equivalent (for our measurement) to the case when the weight is chosen to be 1.02 at  $\eta = 5$  and 0.98 at  $\eta = 2$ , so this function generates at most 2% bias due to the tracking efficiency, and that is why the weighting functions with the three particular choices of  $c_1$  and  $c_2$  correspond to 5%, 2% or 1% tracking efficiency bias respectively.

The event by event weighted Monte Carlo is used to calculate the normalization constants, and with the new polarization constants the polarization parameters are extracted again with the nominal data set. In Figure 4.13, the polarization parameters extracted with the weighted Monte Carlo with the 5%, 2% and 1% level of bias on the tracking efficiency together with the nominal results are shown for a  $J/\psi$  rapidity bin. Compared with the nominal results, it can be seen that the deviations are small. So if the difference of the tracking efficiency between data and Monte Carlo only depends on muon pseudo rapidity linearly, only small biases can be introduced to the measurements of the  $\lambda$  parameters.

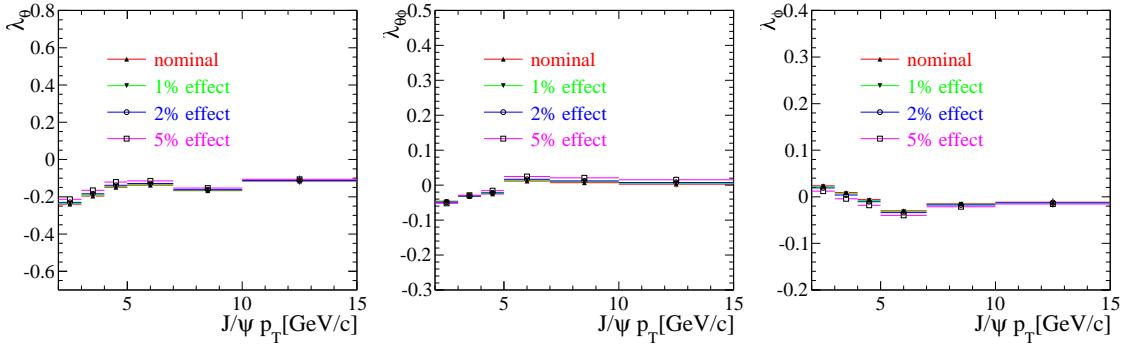


Figure 4.13 Polarization parameters  $\lambda_\theta$  (left),  $\lambda_{\theta\phi}$  (middle) and  $\lambda_\phi$  (right) extracted with weighted Monte Carlo, where the weight for each muon is chosen to be linear function of its pseudo rapidity as  $c_1 \times (1 + c_2 \times \eta_\mu)$ . Three different levels of biases, 5%, 2% and 1% are used. The nominal values are shown for comparison. The figures are for  $J/\psi$  rapidity in  $3.0 < y < 3.5$  bin.

**Parabolic and  $\Lambda$  shape** The situation when the difference of the tracking efficiency between Monte Carlo and data follows parabolic or  $\Lambda$  shapes is similar.

The coefficients in the parabola are properly used such that the parabola maximizes at  $\eta = 3.5$ —the middle of the LHCb  $\eta$  range—with maximum value equal to 1, and the values of the parabola at  $\eta = 2$  and 5 are 0.90, 0.96 or 0.98 respectively, and thus the tracking efficiency bias of levels 5%, 2% and 1% are studied respectively.

The  $\Lambda$  shape is a combination of two linear functions with one function starting at

$\eta = 2$  and ending at  $\eta = 3$  (or  $\eta = 4$  alternatively) and the other function joining the first one at  $\eta = 3$  (or  $\eta = 4$ ) and ending at  $\eta = 5$ . The 2% level tracking efficiency bias effect is studied for the  $\Lambda$  shape, where the function has values 0.96 at  $\eta = 2$  and  $\eta = 5$  while it takes value 1 at the turn point ( $\eta = 3$  or  $\eta = 4$ ).

The weighted Monte Carlo is used to calculate the normalization again for the polarization fit for each case, and new results are extracted. In Figure 4.14, the new polarization parameters together with the nominal ones are shown for various  $J/\psi$  kinematic bin in a  $J/\psi$  rapidity bin. Similar to the studies with linear weighting model, the results show that the deviations are also small in the parabola or  $\Lambda$  shape model.

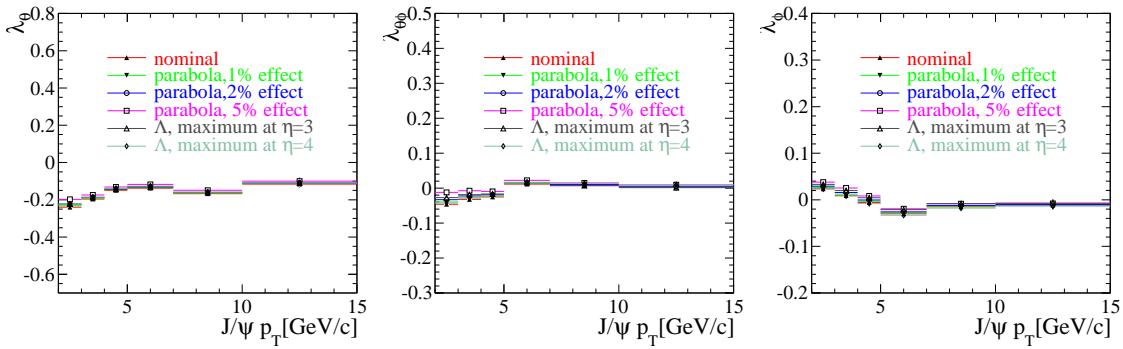


Figure 4.14 Polarization parameters  $\lambda_\theta$  (left),  $\lambda_{\theta\phi}$  (middle) and  $\lambda_\phi$  (right) extracted with weighted Monte Carlo, where the weight for each muon is chosen to be parabola or  $\Lambda$  shape function of its pseudo rapidity. Several different sizes of weights are used. The nominal values are shown for comparison. The figures are for  $J/\psi$  rapidity in  $3.0 < y < 3.5$  bin.

#### 4.3.3.3 Weighting on muon momentum

In this section, the weight for each  $J/\psi$  is defined according to the momenta of the two muons. The muon momentum globally covers a large range, from  $6 \text{ GeV}/c$  to several hundreds of  $\text{GeV}/c$ , but in the low  $p_T$  and small rapidity  $J/\psi$  kinematic bins, the muon momentum falls into very narrow range, from  $6 \text{ GeV}/c$  to several tens of  $\text{GeV}/c$ , so a linear weight function that gives 2% (for example) bias globally will generate  $\approx 0.2\%$  bias in the low momentum bins, which is probably not enough. As a result the weight is chosen to be a *logarithm* function of the muon momentum:

$$c_1 + c_2 \times \log\left(\frac{P_\mu}{\text{GeV}/c}\right),$$

where  $P_\mu$  is the momentum of muons, and the coefficients  $c_1$  and  $c_2$  can be tuned to represent 2% and 5% level of bias in tracking efficiency. The slope of the logarithm shape as a function of  $P_\mu$  is large at low  $P_\mu$  and small at large  $P_\mu$  and because muons in low  $p_T$  and low rapidity  $J/\psi$  bins cover small momentum range, while muons in large  $p_T$  and large rapidity range covers large momentum range, the logarithm shape can thus generate similar size of bias in low  $p_T$  and low rapidity regions and high  $p_T$  and high rapidity regions.

For the 2% and 5% weighting situation, the polarization parameters are extracted again with the weighted Monte Carlo. In Figure 4.15, the newly extracted polarization parameters in different  $J/\psi$   $p_T$  bins in a  $y$  bin are shown. From the plot it can be seen that, even for the 10% case, the discrepancy is quite below or at the same level of the statistical fluctuation. In the procedures only models that are increasing functions of muon

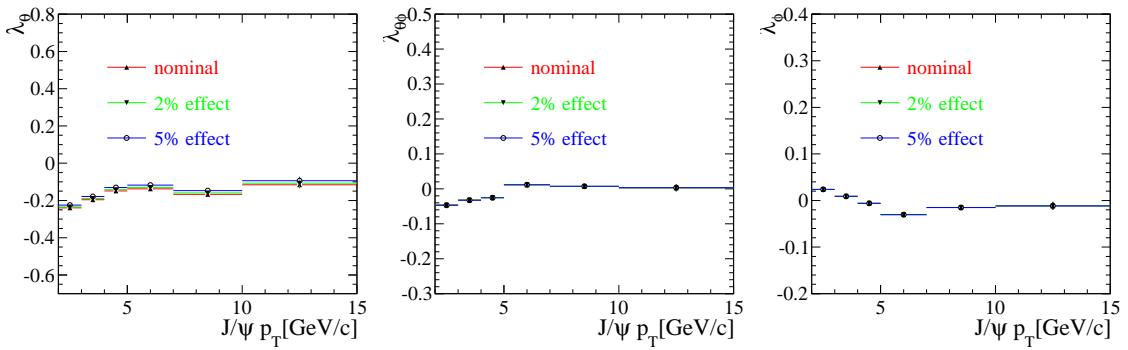


Figure 4.15 Polarization parameters  $\lambda_\theta$  (left),  $\lambda_{\theta\phi}$  (middle) and  $\lambda_\phi$  (right) extracted with weighted Monte Carlo, where the weight for each muon is chosen to be logarithm function of its momentum. 2% and 5% bias are used. The nominal values are shown for comparison. The figures for  $J/\psi$  rapidity in  $3.0 < y < 3.5$  bin.

kinematics are considered, however for the decreasing functions the deviation of the new results from nominal ones will only change sign; the absolute amounts of the deviations will be the same. It should also be noted that only continuous functions are used for the weights — it is assumed that there is no sharp discrepancy of tracking efficiency between data and simulation in a specific narrow muon pseudo rapidity or momentum interval. The hit density distribution of muons in different detector regions in data has been compared with Monte Carlo and no significant discrepancy is observed, which confirms that the continuous dependence models are valid.

As the discrepancy of tracking efficiency between data and Monte Carlo is up to 2% in some muon kinematic region, the track weighting scenarios with 2% variation are used

as the reference for the tracking systematic uncertainty. For each 2% level bias weighting model, the differences between the newly fitted parameters and the nominal ones are calculated, and the largest one among them is chosen to be the systematic uncertainty. In Table 4.5, the numbers are listed for all the  $J/\psi$  kinematic bins. The results show that the tracking systematic effect is small or of the same order compared to the statistical errors for some high statistics bins. As discussed above, in different  $J/\psi$  kinematic bins the same weighting function is used to mimic the underlying tracking bias, so the tracking systematic uncertainties are correlated across  $J/\psi$  kinematic bins.

#### 4.3.4 Monte Carlo validation

The studies in section 4.3.3.1 focus on the quality of the Monte Carlo simulation on tracking systems with a pre-defined tracking algorithm, so the studies are mainly about the reconstruction/selection efficiency. However there are some other issues that can be different between simulation and data, but are not taken into account by the difference of tracking efficiency — the geometry acceptance efficiency and the trigger efficiency for example. Generally the discrepancies of these kinds of efficiency between data and Monte Carlo depend on the kinematics of the final state muons in the laboratory frame, so the discrepancies depend the angular distributions in the  $J/\psi$  rest frame, and as a result the polarization measurement will be biased.

To verify the agreement between the efficiency in data and the efficiency estimated from the Monte Carlo sample, the  $J/\psi$  from the exclusive sample of  $B^+ \rightarrow J/\psi K^+$  is studied. The  $B^+ \rightarrow J/\psi K^+$  has the largest reconstruction rate because it has the fewest possible final state charged particles with a  $J/\psi$  as intermediate state. At the same time, the choice of this particular sample is motivated by the fact that the polarization of the  $J/\psi$  coming from the  $B^+$  is fixed and known, independent of the production environment of the spinless  $B^+$  meson. In fact, due to the helicity conservation, the  $J/\psi$  must be totally longitudinally polarized in the  $B^+$  rest frame, and in the HX frame of  $J/\psi$  there will be a small residual polarization, which will be the **same** in data and in Monte Carlo.

Since the polarization is the same for  $J/\psi$  from  $B^+ \rightarrow J/\psi K^+$  in the HX frame in data and simulation, different muon angular distributions between data and simulation can be originated only by not perfect detector description in the Monte Carlo simulation program. In other words, the comparison of the muon angular distribution between data and Monte Carlo allows us to check the reliability of the simulation in the detector

Table 4.5 The systematic uncertainties coming from the 2 percent bias in the Monte Carlo tracking efficiency using the weighting technique for each  $p_T$  and rapidity bin of  $J/\psi$  bin in the HX frame.

$p_T$ (GeV/c)		$y$			
$2 < P_T < 3$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0078	0.0124	0.0111	0.0164	0.0285
$\lambda_{\theta\phi}$	0.0188	0.0221	0.0196	0.0201	0.0221
$\lambda_\phi$	0.0159	0.0072	0.0093	0.0109	0.0133
$3 < P_T < 4$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0072	0.0053	0.0050	0.0190	0.0278
$\lambda_{\theta\phi}$	0.0075	0.0167	0.0137	0.0127	0.0137
$\lambda_\phi$	0.0165	0.0081	0.0108	0.0131	0.0149
$4 < P_T < 5$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0101	0.0046	0.0038	0.0192	0.0266
$\lambda_{\theta\phi}$	0.0030	0.0114	0.0085	0.0061	0.0075
$\lambda_\phi$	0.0131	0.0078	0.0104	0.0126	0.0139
$5 < P_T < 7$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0078	0.0042	0.0038	0.0179	0.0234
$\lambda_{\theta\phi}$	0.0084	0.0073	0.0051	0.0060	0.0076
$\lambda_\phi$	0.0091	0.0062	0.0092	0.0101	0.0114
$7 < P_T < 10$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0037	0.0041	0.0025	0.0164	0.0193
$\lambda_{\theta\phi}$	0.0101	0.0076	0.0066	0.0075	0.0073
$\lambda_\phi$	0.0082	0.0042	0.0069	0.0063	0.0080
$10 < P_T < 15$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0055	0.0061	0.0048	0.0151	0.0140
$\lambda_{\theta\phi}$	0.0106	0.0072	0.0068	0.0070	0.0062
$\lambda_\phi$	0.0066	0.0024	0.0046	0.0030	0.0051

description.

The 2011  $B^+ \rightarrow J/\psi K^+$  exclusive sample has been selected with the *Bu2JpsiKDetached* stripping line. The  $B^+ \rightarrow J/\psi K^+$  Monte Carlo is generated with the MC11a condition, and the trigger condition applied is TCK 0x40760037, which has the same configurations on the trigger lines as those applied in data concerning our signal. In the reconstruction process, the  $J/\psi$  is reconstructed by looking at its decay in two muons. To make the  $J/\psi$  from the  $B^+ \rightarrow J/\psi K^+$  as comparable to the inclusive  $J/\psi$  as possible, the same offline selections used in the polarization analysis are applied to the reconstructed muon tracks and on the  $J/\psi$  decay vertex. In addition, to suppress the background, the following cuts are applied to the  $K^+$  tracks and to the  $B^+$ :

- $\chi^2(K^+ \text{ track})/\text{ndof} < 4$ ;
- $\Delta \log \text{PID}_{K/\pi} > 0$  (5);
- $B^+ \text{ vertex } \chi^2/\text{nDoF} < 10$ ;
- $\tau(B^+) > 0.3 \text{ ps}$ .

Since only the  $J/\psi$  is of interest, to reduce the bias (if any) from the kaon selection, two different cuts on kaon PID are used, a loose one ( $K^+ \text{ PID}_{K/\pi} > 0$ ) and a tight one ( $K^+ \text{ PID}_{K/\pi} > 5$ ), and the two cases are analyzed independently.

Those events that fail L0 *Muon* and L0 *DiMuon* are rejected. At the Hlt trigger level, the events are required to pass the *Hlt1DiMuonHighMass* and *Hlt2DiMuonJPsiHighPT* requirements. The L0 and Hlt1 triggers are the same as those used to select inclusive  $J/\psi$ , while *Hlt2DiMuonJPsiHighPT* differs from the one *Hlt2DiMuonJPsi*. The trigger line *Hlt2DiMuonJPsiHighPT* uses a narrower mass window ( $100 \text{ MeV}/c^2$ ), and requires the transverse momentum of  $J/\psi$  candidate to be larger than  $2 \text{ GeV}/c$ , which is also applied to the inclusive  $J/\psi$  sample offline. As a result, there are about 800 000 signal  $B^+ \rightarrow J/\psi K^+$  left for the study from real data.

Since the difference of the muon angular distribution can be introduced by the difference of the efficiency as well as by the difference of the  $J/\psi$  kinematics, the latter difference has to be removed first. To do that, a four-step weighting is used, firstly the  $B^+ p_T - y$  distribution in Monte Carlo is weighted to data according to a two-dimensional histogram which is the ratio of the  $B^+ p_T - y$  (transverse momentum and rapidity) distribution in data over the distribution in Monte Carlo, secondly the simulated  $K^+ p_T - \eta$  (transverse momentum and pseudo-rapidity in laboratory frame) distribution is weighted to data using the same way as the first step, thirdly the simulated  $J/\psi$  angular ( $\theta - \phi$ ) dis-

tribution the in  $B^+$  rest frame is weighted to the distribution in data and the final weighting makes the  $J/\psi$   $p_T - y$  distribution (in laboratory frame) in Monte Carlo to data. A successive weighting can make the previous weighting not perfect, so after all the weighting, a second iteration of the weighting process is applied. However, it is found that the second round makes only small corrections to the first round weighting (around 0.1% correction). In Figures 4.16, 4.17, 4.18 and 4.19 the  $B^+$   $p_T$  and rapidity distributions, the  $K^+$   $p_T$  and rapidity distributions, the  $J/\psi$  angular  $\cos \theta$  and  $\phi$  distributions in  $B^+$  rest frame and the  $J/\psi$   $p_T$  and rapidity distributions for data and for Monte Carlo before and after the weighting are shown respectively. From these comparison plots it can be seen that before the weighting the  $B^+$  kinematics are slightly different before the weighting, however the agreement becomes much better between the weighted Monte Carlo and data.

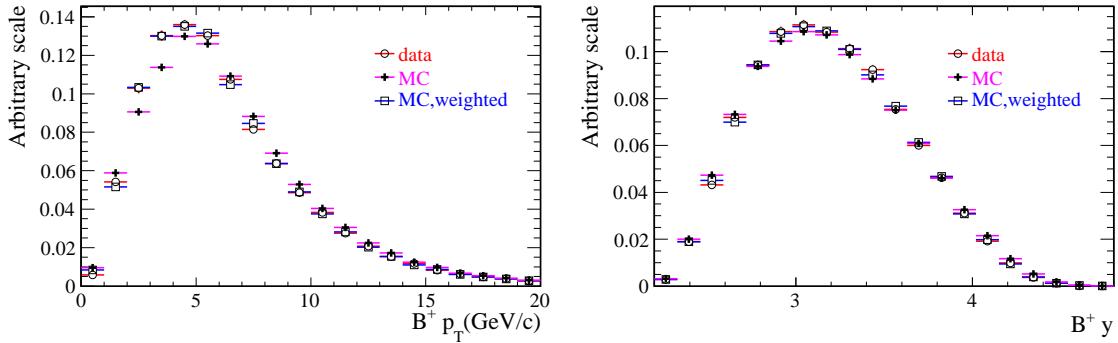


Figure 4.16 The  $B^+$  transverse momentum (left) and rapidity (right) distributions for data (open circle) and Monte Carlo before (cross) and after the weighting (open box).

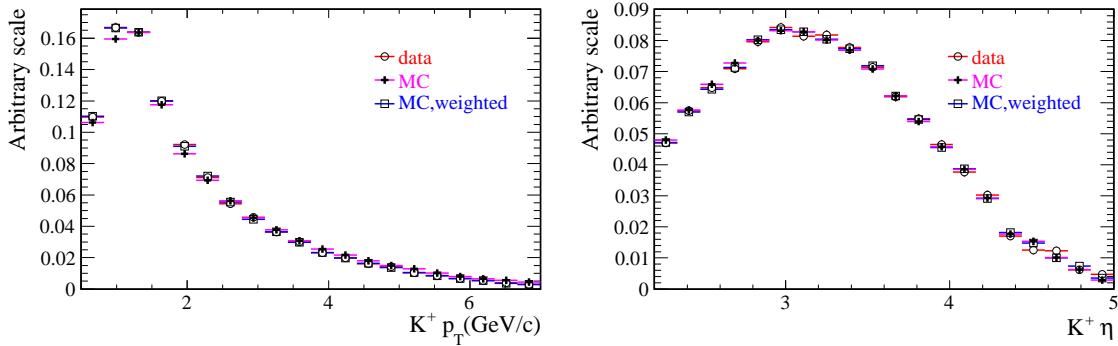


Figure 4.17 The  $K^+$  transverse momentum (left) and pseudo-rapidity (right) distributions for data (open circle) and Monte Carlo before (cross) and after the weighting (open box).

In the  $B^+$  system, the muon angular variables can also be defined in the  $B^+$  rest frame, where the  $z$ -axis is chosen as the  $J/\psi$  flight direction in the  $B^+$  rest frame while the  $y$ -axis is defined as the normal to the plane formed by the direction of  $J/\psi$  momentum

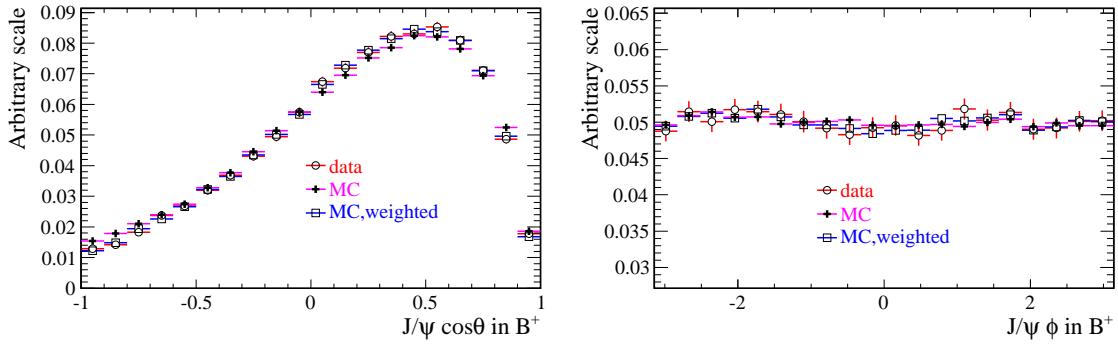


Figure 4.18 The distributions of  $J/\psi$  angular variable  $\cos \theta$  (left) and  $\phi$  (right) in  $B^+$  rest frame for data (open circle) and Monte Carlo before (cross) and after the weighting (open box).

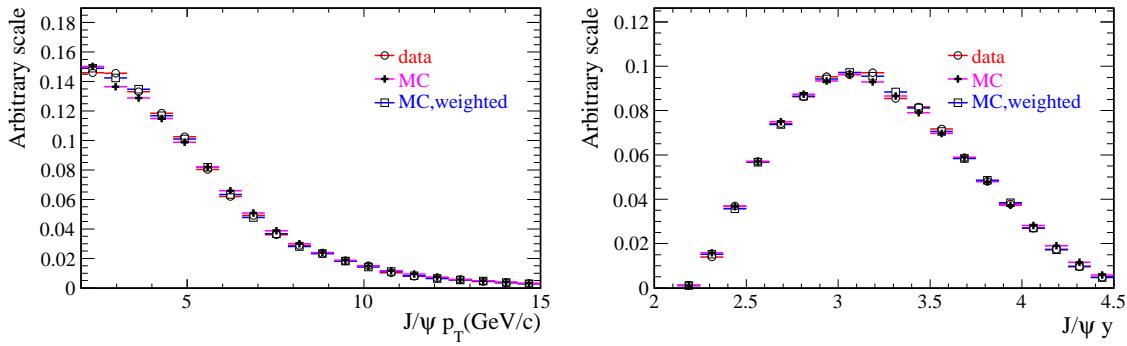


Figure 4.19 The  $J/\psi$  transverse momentum (left) and rapidity (right) distributions for data (open circle) and Monte Carlo before (cross) and after the weighting (open box).

in  $B^+$  rest frame and the direction of  $B^+$  momentum in laboratory. In Figures 4.20 and 4.21, the muon  $\cos \theta$  distributions in the  $B^+$  rest frame and the HX frame are shown respectively. From the figures it can be seen that muon angular distribution in the  $B^+$  rest frame is almost the same in data and weighted Monte Carlo, while in the HX frame data and Monte Carlo are not consistent, which means that the efficiencies in data and Monte Carlo do not agree perfectly. In laboratory frame, the difference of the muon  $p_T - \eta$  (or  $p - \eta$ ) distribution between data and the weighted Monte Carlo is a measurement of the difference of the efficiency. In Figure 4.22, the muon  $p_T - \eta$  and  $p - \eta$  distribution in data over the distributions in the weighted Monte Carlo are shown. When the ratio is calculated the two original histograms are both normalized to unity.

The ratio table (Figure 4.22) is used to weight the inclusive  $J/\psi$  Monte Carlo sample for the prompt  $J/\psi$  polarization measurement. With the weighted Monte Carlo and the nominal data set, a new set of polarization parameters is extracted, and the differences between the new results and the nominal ones are considered to come from the inconsistency of the efficiency between data and Monte Carlo. As the efficiency ratio table is

derived from exclusive  $B^+ \rightarrow J/\psi K^+$  decay, to study how the selection of kaons biases the estimation of the efficiency ratio, two different kaon PID cuts are used —  $\Delta \log \text{PID}_{K/\pi}$  larger than zero or five respectively. However the results are very similar. It is also checked that the weight on kaon kinematics have only a very small effect on the muon kinematic distributions. So as a result, the bias of the kaon reconstruction and selection to the efficiency ratio table is negligible. To reduce the uncertainty of the efficiency table due to the binning in muon kinematics  $p_T - \eta$  or  $p - \eta$ , several binning choices are used, and for each case the inclusive Monte Carlo is weighted to calculate the normalization for a polarization fit independently. The results are averaged, and the deviation of the average value from the nominal one is quoted as acceptance systematic uncertainties.

In Table 4.6, the systematic uncertainties are listed for various  $J/\psi$   $p_T$  and rapidity bins. The acceptance systematic uncertainties are of the same level as the statistical errors in the low statistical bins, while they are much larger than the statistical fluctuations for high statistics bins. The average values are about 0.06 for  $\lambda_\theta$ , and 0.02 for  $\lambda_{\theta\phi}$  and  $\lambda_\phi$  in the HX frame, dominating the systematic uncertainties. It should also be noted that the studies in this section and the studies of the tracking efficiency are not totally independent, however as the efficiency table provided by the tracking group is calculated with slightly different selections on the  $J/\psi$  candidates and the kinematic range of the muons in the two methods is also a little different. Since the uncertainties coming from difference of the offline tracking efficiency are rather small, the analysis still supposes that they are two independent sources of systematic uncertainties.

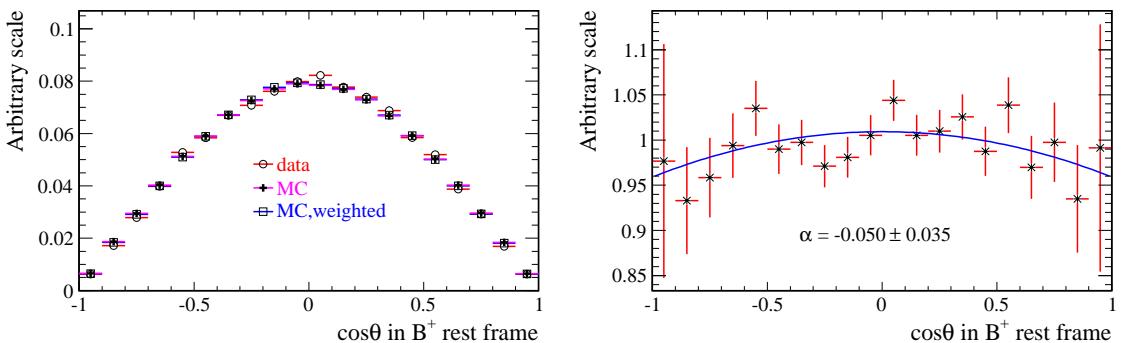


Figure 4.20 The distributions of muon angular variable  $\cos \theta$  (*left*) in  $B^+$  rest frame for data (open circle) and Monte Carlo before (cross) and after the weighting (open box) and the ratio of the distribution in data over the one in weighted Monte Carlo (*right*). A fit to the ratio distribution with the function  $c \times (1 + \alpha \cos^2 \theta)$  is superimposed.

Table 4.6 The acceptance systematic uncertainties for each  $p_T$  and rapidity bin of  $J/\psi$  in the HX frame .

$p_T$ (GeV/c)		$y$			
$2 < P_T < 3$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0996	0.0728	0.0710	0.0228	0.0368
$\lambda_{\theta\phi}$	0.0662	0.0030	0.0078	0.0592	0.0978
$\lambda_\phi$	0.0038	0.0087	0.0148	0.0170	0.0439
$3 < P_T < 4$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0768	0.1575	0.0230	0.0262	0.0369
$\lambda_{\theta\phi}$	0.0135	0.0116	0.0029	0.0500	0.0671
$\lambda_\phi$	0.0058	0.0187	0.0113	0.0211	0.0457
$4 < P_T < 5$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0181	0.1851	0.0003	0.0148	0.0962
$\lambda_{\theta\phi}$	0.0083	0.0038	0.0118	0.0181	0.0416
$\lambda_\phi$	0.0092	0.0085	0.0026	0.0287	0.0450
$5 < P_T < 7$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.1645	0.1054	0.0203	0.0285	0.0454
$\lambda_{\theta\phi}$	0.0243	0.0068	0.0045	0.0164	0.0447
$\lambda_\phi$	0.0103	0.0070	0.0045	0.0283	0.0454
$7 < P_T < 10$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.2242	0.0465	0.0054	0.0653	0.0650
$\lambda_{\theta\phi}$	0.0015	0.0090	0.0223	0.0094	0.0515
$\lambda_\phi$	0.0051	0.0050	0.0095	0.0137	0.0263
$10 < P_T < 15$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0052	0.0740	0.0756	0.0462	0.0732
$\lambda_{\theta\phi}$	0.0328	0.0003	0.0150	0.0048	0.0528
$\lambda_\phi$	0.0049	0.0004	0.0068	0.0056	0.0065

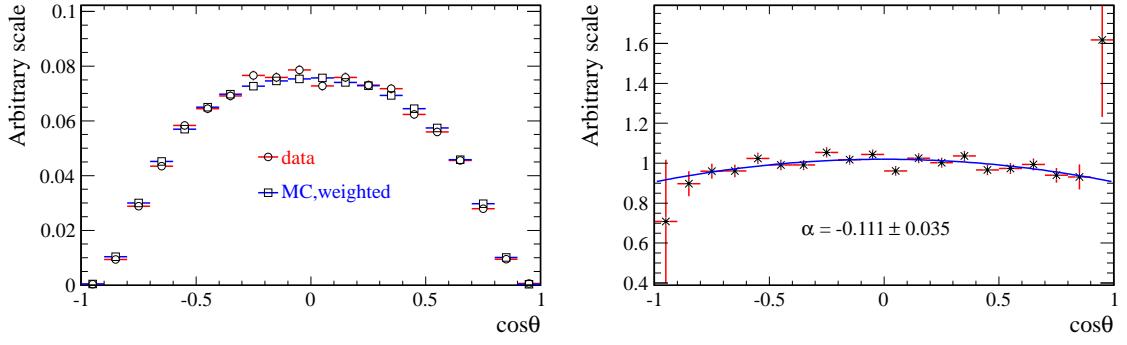


Figure 4.21 The distributions of muon angular variable  $\cos\theta$  (*left*) in the HX frame for data (open circle) and Monte Carlo before (cross) and after the weighting (open box) and the ratio of the distribution in data over the one in weighted Monte Carlo (*right*). A fit to the ratio distribution with the function  $c \times (1 + \alpha \cos^2 \theta)$  is superimposed.

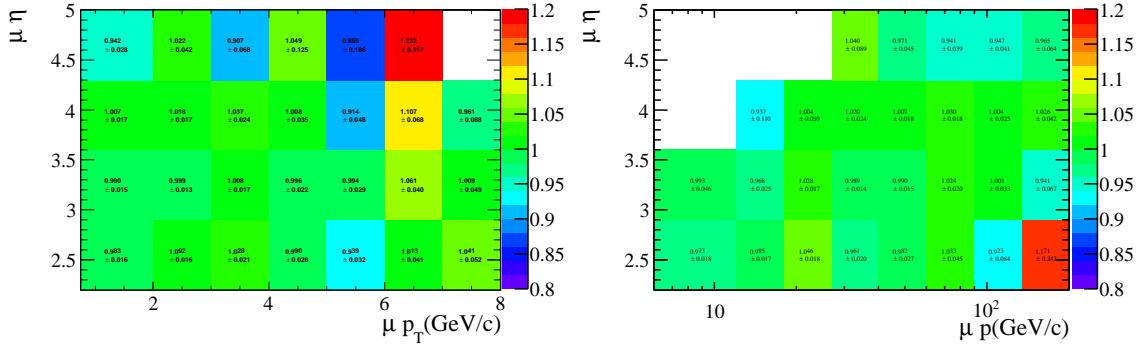


Figure 4.22 The ratio of the muon  $p_T - \eta$  (*left*) and  $p - \eta$  (*right*) distributions in data over the distributions in the weighted Monte Carlo.

### 4.3.5 Background subtraction

In the construction of the likelihood estimator, two assumptions are made (see discussions in section 4.2.2). In this section the situations when the real background distribution deviates from these two assumptions are investigated.

The first assumption is that, concerning the mass spectrum, the background distribution is linear, so when the total widths of the sidebands are chosen to be equal to the width of signal region, a weight of -1 should be assigned to events in the sidebands and for events in signal region the weight is +1. The background can also be assumed to follow exponential shape, which is used in the cross section measurement paper<sup>[147]</sup>, and in this case the weights for sideband events are chosen to normalize the number of background events in sideband regions to the number of background events in the signal region. The number of events is calculated from the integral of the exponential function over the sideband or signal regions. However it is found that the weights for sideband

events from exponential background function are only slightly (below 0.2%) different from -1, the differences are very much below the statistical fluctuation of background events ( $\propto 1/\sqrt{N_{\text{background}}}$ ). With these new weights for sideband events, the polarization parameters are extracted again, and as expected the differences between the newly fitted parameters and the nominal ones are negligible.

It is also assumed that the distributions of  $\cos\theta - \phi$  are similar for background events in sideband regions and for background events in signal regions — independent of the mass regions. In Figure A.9, the one dimensional  $\cos\theta$  and  $\phi$  distributions are plotted for the left sideband and the right sideband respectively, and it can be seen that background angular distributions do depends on the reconstructed mass, which means that the second assumption is not valid perfectly. However the dependence is smooth and linear: the background  $\cos\theta$  ( $\phi$ ) distribution in the signal region will always lie between the distributions in the left sideband and the right sideband. So by combining the left and right sideband the background distribution in signal regions can be reproduced approximately. Extreme cases are used to study the imperfectness of our second assumption; in the extreme cases only the left sideband or the right sideband is used to subtract the background, and the weights for the sideband events are properly chosen to normalize the number of background events in the specified (left or right) sideband to the number of background events in the signal region, assuming the background mass distribution follows a linear function. With these two cases of background subtraction, another two sets of polarization results are extracted. Because the left or right sideband do differ from background in signal region, in  $J/\psi$  kinematic bins where the signal to background ratio is not large enough, drastically different polarization parameters are observed in the two fits with extreme background subtraction.

In each  $p_T$  and  $y$  bin, for each polarization parameter, the two variations, when only left or right sideband is used to subtract the background, are calculated, and the larger one is quoted as background subtraction systematic uncertainties, which are listed in Table 4.7. The systematic uncertainties evaluated this way are quite conservative and usually larger than (in high statistical bins) or of the same order (in low statistical bins) of statistical errors. Since the differences of muon angular distributions between background events in the signal and sideband regions are similar across  $J/\psi$  kinematic bins — the background subtraction systematic uncertainties are highly correlated between bins.

Table 4.7 The background subtraction systematic uncertainties for each  $p_T$  and rapidity bin of  $J/\psi$  bin in the HX frame.

$p_T$ (GeV/c)		$y$			
$2 < P_T < 3$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0988	0.0295	0.0273	0.0069	0.0099
$\lambda_{\theta\phi}$	0.0554	0.0242	0.0150	0.0254	0.0311
$\lambda_\phi$	0.0554	0.0380	0.0170	0.0047	0.0055
$3 < P_T < 4$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0453	0.0417	0.0186	0.0131	0.0017
$\lambda_{\theta\phi}$	0.0313	0.0077	0.0059	0.0085	0.0165
$\lambda_\phi$	0.0110	0.0005	0.0033	0.0016	0.0044
$4 < P_T < 5$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0246	0.0079	0.0071	0.0064	0.0040
$\lambda_{\theta\phi}$	0.0014	0.0006	0.0009	0.0028	0.0063
$\lambda_\phi$	0.0027	0.0001	0.0003	0.0019	0.0030
$5 < P_T < 7$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0025	0.0027	0.0017	0.0007	0.0132
$\lambda_{\theta\phi}$	0.0015	0.0002	0.0006	0.0013	0.0008
$\lambda_\phi$	0.0016	0.0007	0.0010	0.0026	0.0029
$7 < P_T < 10$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0050	0.0037	0.0067	0.0084	0.0242
$\lambda_{\theta\phi}$	0.0011	0.0006	0.0035	0.0030	0.0015
$\lambda_\phi$	0.0008	0.0001	0.0020	0.0023	0.0049
$10 < P_T < 15$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0076	0.0108	0.0077	0.0340	0.0348
$\lambda_{\theta\phi}$	0.0009	0.0005	0.0031	0.0085	0.0135
$\lambda_\phi$	0.0008	0.0006	0.0033	0.0055	0.0007

### 4.3.6 Selections

Several cuts are applied to select good candidates and to suppress the background fraction in data. The distributions of these selection variables could be different between data and Monte Carlo, so cutting at the same value selects different set of events in data and in Monte Carlo, thus biases can be introduced in the polarization measurement. The possible sources of such systematic effects are analyzed in this section.

#### 4.3.6.1 Binning effect

The kinematics of a  $J/\psi$  decay event is completely determined by the four variables  $J/\psi$  transverse momentum,  $J/\psi$  rapidity and  $\mu^+ (\mu^-) \cos \theta$  and  $\phi$ , averaging over  $J/\psi$  azimuthal angle. The analysis divides events in  $J/\psi$   $p_T$  and rapidity bin as discussed in Equation 4-3, and the efficiency as a function of  $\cos \theta - \phi$  is assumed to be independent of  $J/\psi$   $p_T$  and rapidity within a particular bin. However, this assumption is not completely true, especially in the bins where muons can easily go outside of the LHCb acceptance (low  $p_T$  and small rapidity bins for example) or fail some selection cuts, because in these regions whether the muons can be reconstructed and selected is strongly related to the  $J/\psi$  kinematics. In the bins where the efficiencies change drastically as functions of  $J/\psi$   $p_T$  and  $y$ , and if the  $J/\psi$   $p_T$  and  $y$  spectrums are different between Monte Carlo and data, the measurement will be biased due to the large binning size.

To study the systematic effect, finer bins are made in each  $p_T$  and rapidity bin to reduce the efficiency's dependence on  $J/\psi$  kinematics. Practically two equal size sub bins in  $y$  and two equal size sub bins in  $p_T$  — four in total — are used and in each sub bin the polarization parameters are extracted separately, resulting in four sets of parameters. For  $\lambda_\theta$ ,  $\lambda_{\theta\phi}$  and  $\lambda_\phi$ , the consistency of the four new values are checked, and their weighted average are compared with the nominal values and the differences between the average value and the nominal values are quoted as the binning systematic uncertainties, which are listed in Table 4.8. Because our choice of binning is already small enough and the variations of efficiency as a function of muon angular variables within one bin are usually small and smooth, the systematic uncertainties are very small for most of the bins, and well below the statistical errors for high statistics bins. Only for some low rapidity and low  $p_T$  bins, where the efficiencies change strongly, the binning systematic uncertainties can be the dominant systematic effect. The binning systematic uncertainties are bin dependent, as the difference of  $J/\psi$  kinematics spectrum between data and Monte Carlo

Table 4.8 The binning systematic uncertainties for each  $p_T$  and rapidity bin of  $J/\psi$  bin in the HX frame.

$p_T$ (GeV/c)		$y$				
$2 < P_T < 3$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$	
$\lambda_\theta$	0.1647	0.0384	0.0067	0.0126	0.0219	
$\lambda_{\theta\phi}$	0.0285	0.0096	0.0185	0.0214	0.0071	
$\lambda_\phi$	0.0247	0.0083	0.0067	0.0072	0.0239	
$3 < P_T < 4$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$	
$\lambda_\theta$	0.1089	0.0464	0.0028	0.0053	0.0096	
$\lambda_{\theta\phi}$	0.0347	0.0057	0.0032	0.0040	0.0032	
$\lambda_\phi$	0.0121	0.0015	0.0030	0.0018	0.0036	
$4 < P_T < 5$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$	
$\lambda_\theta$	0.0743	0.0272	0.0031	0.0014	0.0111	
$\lambda_{\theta\phi}$	0.0301	0.0073	0.0005	0.0003	0.0049	
$\lambda_\phi$	0.0137	0.0006	0.0015	0.0010	0.0073	
$5 < P_T < 7$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$	
$\lambda_\theta$	0.0516	0.0177	0.0024	0.0014	0.0045	
$\lambda_{\theta\phi}$	0.0327	0.0062	0.0023	0.0007	0.0003	
$\lambda_\phi$	0.0141	0.0002	0.0026	0.0014	0.0021	
$7 < P_T < 10$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$	
$\lambda_\theta$	0.0360	0.0051	0.0047	0.0017	0.0005	
$\lambda_{\theta\phi}$	0.0215	0.0029	0.0013	0.0010	0.0002	
$\lambda_\phi$	0.0127	0.0002	0.0015	0.0003	0.0078	
$10 < P_T < 15$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$	
$\lambda_\theta$	0.0223	0.0009	0.0055	0.0069	0.0431	
$\lambda_{\theta\phi}$	0.0172	0.0006	0.0005	0.0007	0.0003	
$\lambda_\phi$	0.0069	0.0004	0.0001	0.0022	0.0071	

varies across bins.

#### 4.3.6.2 $\tau_S$ selection

In order to select the prompt  $J/\psi$  events, the pseudo proper time significance ( $\tau_S$ , see Equation 4-2) is defined and  $|\tau_S|$  is required to be smaller than 4. By doing this, a little fraction (< 5%) of  $b$ -decay events is still present in the "prompt" sample and another little

fraction ( $\approx 2\%$ ) of real prompt events is excluded. In the Monte Carlo, an alternative way can be used to select real prompt  $J/\psi$  event; the whole production chain where the  $J/\psi$  originates can be investigated, and if any of the  $J/\psi$  ancestors is a long lived hadron, the  $J/\psi$  is excluded, and finally a pure prompt  $J/\psi$  Monte Carlo is selected.

A real prompt  $J/\psi$  is produced at the PV, so both the reconstructed  $t_z$  and its error come from the detector resolution. The detector resolution depends on  $J/\psi$  and muon kinematics, so  $\tau_S$  depends on the  $\mu$  angular variables  $\cos\theta - \phi$ . Figure 4.23 plots the  $\cos\theta$  distribution for events that pass the  $\tau_S$  cuts and for events that are excluded. The plot shows that  $\tau_S$  cut really biases the muon angular distribution. From Figure 4.24, it

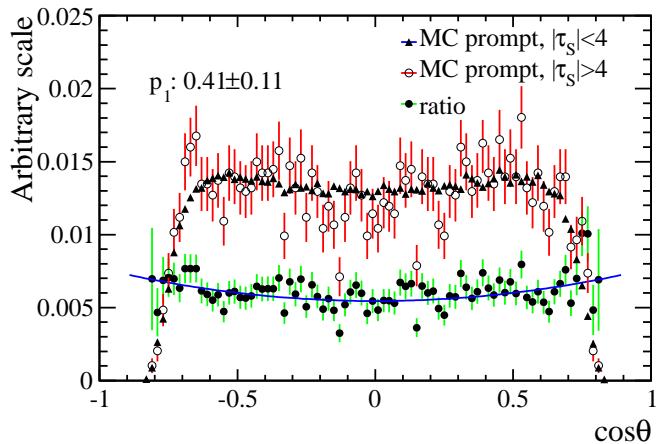


Figure 4.23  $\cos\theta$  distributions for prompt events that pass (circle) the  $\tau_S$  cut and those excluded (triangle) by the cut. The ratio (dot) of the two are also plotted, and a fit to the ratio distribution with the function  $p_0 * (1 + p_1 \cos^2 \theta)$  is superimposed. An arbitrary  $J/\psi$  kinematic bin is used for an example.

can be seen that the  $\tau_S$  distribution is different in data and in simulation, so the  $\tau_S$  cut will select different sets of events for Monte Carlo and for data, and thus a bias can be introduced. However it is also discovered that if  $\tau_S$  in real prompt Monte Carlo is scaled by a properly chosen factor, the  $\tau_S$  distributions in Monte Carlo and data agree reasonably well as shown in Figure 4.24.

The retention (rejection) rate related to  $\tau_S$  selection in real prompt Monte Carlo can be easily calculated from the number of events before and after the cut. The scaled prompt Monte Carlo that reproduces the prompt  $\tau_S$  distribution in data is used to calculate the retention (rejection) rate for data.

To be strict, the  $b$  component will make the  $\tau_S$  comparison between Monte Carlo and

data a little more complicated, because prompt events can not be separated from  $b$  events in data event by event, and the fraction of  $J/\psi$  from  $b$  in data is not correctly simulated by Monte Carlo. However the contribution of the  $b$  component in the  $|\tau_S| < 4$  region is only about 3% in most of the bins as can be seen from Figure 4.25, and the contamination is negligible ( $< 0.1\%$ ) in the left hand side of  $\tau_S$  distribution ( $\tau_S < 0$ ), so the scale factor is chosen in the way that the  $\tau_S$  distribution in data and Monte Carlo overlap in  $\tau_S < 0$  region.

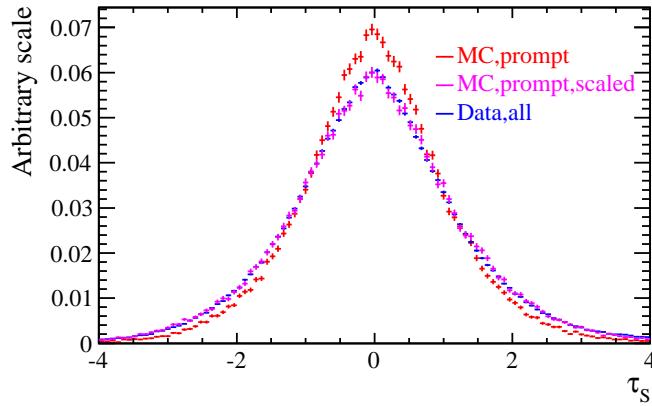


Figure 4.24  $\tau_S$  distributions for data (blue) real prompt Monte Carlo (red) and scaled Monte Carlo (purple).

With the normalization constants calculated from pure prompt Monte Carlo sample, while the data is selected with the nominal  $\tau_S$  cut ( $|\tau_S| < 4$ ), a new set of polarization parameters is extracted. The difference between the new fit results and nominal results is just because the  $\cos \theta - \phi$  distribution in the events removed by the  $\tau_S$  cut is different from the distribution in the events that pass the cut, when the  $b$  contamination is neglected for the moment.

In the new fit the rejection rate of the  $\tau_S$  selection in Monte Carlo is zero, and it is  $r_{MC}$  in the nominal fit, however the rejection rate in data stays to be  $r_{data}$  in the two fits, so the difference of rejection rate between data and Monte Carlo changes by an amount of  $r_{MC}$ . The difference of the polarization results ( $\Delta$ ) between the two fits corresponds to the variation of the difference of the rejection rate, which is  $r_{MC}$ . However in the nominal fit itself, the difference of rejection rate between data and Monte Carlo is only  $|r_{MC} - r_{data}|$ , so the systematic uncertainty due to  $\tau_S$  in the nominal fit is computed as  $\Delta \times |r_{MC} - r_{data}|/r_{MC}$ . An alternative way to estimate the systematic uncertainty is to

change the  $\tau_S$  cut in Monte Carlo such that the retention rate of  $\tau_S$  in Monte Carlo is the same as data, and new polarizations can be extracted with this selection, while the difference between the new result and the nominal one is the systematic uncertainty. The two methods are consistent. If the rejection rate is the same for data and simulation, the  $\tau_S$  cut bias will be zero; this is as expected because the cut then selects the same events for Monte Carlo and data. The  $\tau_S$  systematic uncertainties are summarized in Table 4.9.

Because the  $\tau_S$  cut at 4 includes more than 98% percent of the prompt events for most of the bins, the rejection rate itself and the difference between data and Monte Carlo is rather small (at most 2%), so only very small bias due to the cut is observed—negligible compared to the statistical errors and other systematic uncertainties. As  $\tau_S$  cut systematic uncertainties come from the fact that the muon angular distribution in  $J/\psi$  rest frame depends on  $\tau_S$ , and the trend of the dependence is similar across  $J/\psi$  kinematic bins, the  $\tau_S$  cut systematic uncertainties are highly correlated across  $J/\psi$  kinematic bins.

#### 4.3.6.3 Muon PID selection

Concerning the systematic uncertainty due to the selection on the muon PID variable, the effect of this selection is strictly related to the systematics introduced by the background subtraction, since the effect of the selection is mainly a consistent reduction of the background. It is verified that, for  $p_T > 2 \text{ GeV}/c$ , the signal retention of this cut ( $\text{PID}_{\mu/\pi} > 0$ ) is the same (inside the statistical sensitivity) on data and Monte Carlo. This conclusion is also supported by studies of the muon PID performance made by the other groups that use muons to reconstruct their signal.

#### 4.3.6.4 Detector resolution

Due to detection resolution, a  $J/\psi$  generated in one  $p_T$  rapidity bin could fall into a different bin when reconstructed. Besides, the calculated muon polar angle is also a little bit different from the generated one. This effect is called bin migration. By looking at the Monte Carlo truth, it is found that the event migration only happens between neighboring bins and mostly between neighboring  $p_T$  bins: the number of events going across the borders of neighboring rapidity bins is only 1/5 (or smaller) of the number of events across  $p_T$  bins. The fraction of migration events is related to the size of resolution, and the Monte Carlo shows that in most bins the fractions of migrated events are less than 1%, while within very few bins the values are about 2%.

Table 4.9 The  $\tau_S$  systematic uncertainties for each  $p_T$  and rapidity bin of  $J/\psi$  bin in the HX frame.

$p_T$ (GeV/c)		$y$			
$2 < P_T < 3$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0144	0.0003	0.0013	0.0008	0.0002
$\lambda_{\theta\phi}$	0.0030	0.0003	0.0004	0.0000	0.0000
$\lambda_\phi$	0.0002	0.0009	0.0007	0.0003	0.0001
$3 < P_T < 4$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0016	0.0002	0.0004	0.0007	0.0001
$\lambda_{\theta\phi}$	0.0022	0.0006	0.0002	0.0002	0.0000
$\lambda_\phi$	0.0008	0.0004	0.0006	0.0001	0.0000
$4 < P_T < 5$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0001	0.0017	0.0021	0.0010	0.0001
$\lambda_{\theta\phi}$	0.0000	0.0008	0.0002	0.0002	0.0000
$\lambda_\phi$	0.0003	0.0007	0.0005	0.0001	0.0000
$5 < P_T < 7$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0009	0.0017	0.0007	0.0006	0.0001
$\lambda_{\theta\phi}$	0.0009	0.0006	0.0003	0.0002	0.0000
$\lambda_\phi$	0.0009	0.0013	0.0006	0.0003	0.0001
$7 < P_T < 10$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0013	0.0003	0.0014	0.0009	0.0000
$\lambda_{\theta\phi}$	0.0009	0.0010	0.0003	0.0001	0.0000
$\lambda_\phi$	0.0015	0.0014	0.0011	0.0002	0.0000
$10 < P_T < 15$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0047	0.0009	0.0013	0.0013	0.0004
$\lambda_{\theta\phi}$	0.0024	0.0017	0.0004	0.0003	0.0013
$\lambda_\phi$	0.0049	0.0017	0.0003	0.0003	0.0001

In data, the resolution can not be determined event by event since the true kinematics is unknown. However, the  $J/\psi$  mass resolution is a clear measure of the kinematic resolution. The relative resolution difference between data and Monte Carlo can be estimated from the resolutions of the reconstructed invariant mass as:

$$\frac{|\sigma_{\text{data}} - \sigma_{\text{MC}}|}{\sigma_{\text{MC}}}.$$

To study the effect of  $\cos\theta - \phi$  resolution and  $J/\psi$  kinematic resolution in the Monte Carlo, the Monte Carlo truth variables —  $J/\psi$   $p_{\text{T}}$  and  $y$ , muon four momentum — are used to determine the  $J/\psi$  kinematic binning and to calculate  $\mu$  polar angles. The Monte Carlo with these new binning and new muon angles is used to extract the normalization constants, which are subsequently used to extract polarization from data. The deviation of these fitted results from the nominal ones is taken as resolution bias  $\Delta$ . However, the size of the deviation corresponds to the entire amount of resolution  $\sigma_{\text{MC}}$  and must be propagated according to the difference of the resolution between data and Monte Carlo in the way:

$$\Delta \times \frac{|\sigma_{\text{data}} - \sigma_{\text{MC}}|}{\sigma_{\text{MC}}}$$

The propagated deviations are quoted as bin migration (resolution) systematic uncertainties as listed in Table 4.10.

There is also another possible source of systematic uncertainty coming from bin migration effect. Since the data is polarized, events coming in and going out from a particular bin could carry different polarization information, so the polarization can be biased by exchanging a few events with a neighboring bin. From the measured polarization, it can be seen that the differences between neighboring bins are at most 0.1 for  $\lambda_{\theta}$  and  $\lambda_{\theta\phi}$ , and 0.05 for  $\lambda_{\phi}$ , so through exchanging a fraction (less than 2%) of events, the bin migration will in the worst case bias  $\lambda_{\theta}$  and  $\lambda_{\theta\phi}$  by an amount of 0.002 ( $0.1 \times 2\%$ ) and 0.001 for  $\lambda_{\phi}$ . As the migration events mostly lie near the boundaries between the neighboring bins, this estimate is safe enough.

Because the fraction of migrated events is so small (so  $\Delta$  is small), and the resolutions in data and Monte Carlo are not too different, the resolution effect is quite negligible compared to the statistical errors and other systematic uncertainties.

Table 4.10 The bin migration systematic uncertainties for each  $p_T$  and rapidity bin of  $J/\psi$  bin in the HX frame.

$p_T$ (GeV/c)		$y$			
$2 < P_T < 3$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0013	0.0010	0.0001	0.0011	0.0020
$\lambda_{\theta\phi}$	0.0007	0.0004	0.0001	0.0003	0.0011
$\lambda_\phi$	0.0000	0.0004	0.0004	0.0006	0.0002
$3 < P_T < 4$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0002	0.0006	0.0016	0.0011	0.0013
$\lambda_{\theta\phi}$	0.0005	0.0002	0.0007	0.0004	0.0008
$\lambda_\phi$	0.0003	0.0006	0.0007	0.0001	0.0002
$4 < P_T < 5$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0037	0.0004	0.0003	0.0003	0.0022
$\lambda_{\theta\phi}$	0.0004	0.0002	0.0002	0.0015	0.0005
$\lambda_\phi$	0.0012	0.0003	0.0003	0.0013	0.0015
$5 < P_T < 7$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0000	0.0011	0.0021	0.0005	0.0028
$\lambda_{\theta\phi}$	0.0006	0.0007	0.0012	0.0000	0.0006
$\lambda_\phi$	0.0002	0.0004	0.0005	0.0000	0.0001
$7 < P_T < 10$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0025	0.0005	0.0011	0.0024	0.0002
$\lambda_{\theta\phi}$	0.0009	0.0008	0.0008	0.0009	0.0003
$\lambda_\phi$	0.0003	0.0005	0.0001	0.0004	0.0005
$10 < P_T < 15$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0081	0.0013	0.0019	0.0003	0.0039
$\lambda_{\theta\phi}$	0.0019	0.0001	0.0011	0.0013	0.0016
$\lambda_\phi$	0.0009	0.0010	0.0006	0.0010	0.0028

### 4.3.7 Residual background of $J/\psi$ from $b$ decay

By the cut  $|\tau_S|$  at 4, there are still a few percent of  $J/\psi$  coming from long lived  $b$ -hadron decays survived in the prompt sample. In the data sample after the  $|\tau_S|$  cut, the fraction of  $J/\psi$  from  $b$  decay is estimated in the following way:

- Firstly the fraction of  $J/\psi$  from  $b$  ( $f_b$ ) in data before the  $\tau_S$  cut is extracted by fitting the  $t_Z$  distribution following the procedure in the cross section measurement<sup>[147]</sup>;
- Secondly the retention rate related to the  $\tau_S$  cut for both prompt  $J/\psi$  ( $r_p$ ) and  $J/\psi$  from  $b$  ( $r_b$ ) are calculated by looking at  $\tau_S$  distribution of the two component separately in Monte Carlo. Here again the generator level information is used to select pure prompt and pure  $b$  decay Monte Carlo sample. As has been discussed before, the  $\tau_S$  in Monte Carlo is properly scaled to reproduce  $\tau_S$  distribution of data.
- The fraction of  $J/\psi$  from  $b$  in data after  $\tau_S$  cut is calculated as the number of survived  $b$  decay events divided by the sum of the numbers of survived  $b$  decay events and survived prompt events:  $f_b \times r_b / ((1 - f_b) \times r_p + f_b \times r_b)$ .

The Figure 4.25 summarizes the fraction of  $b$  contamination after the selection for all the bins. It can be seen that for most of the bins the fraction is below 5%. To study the

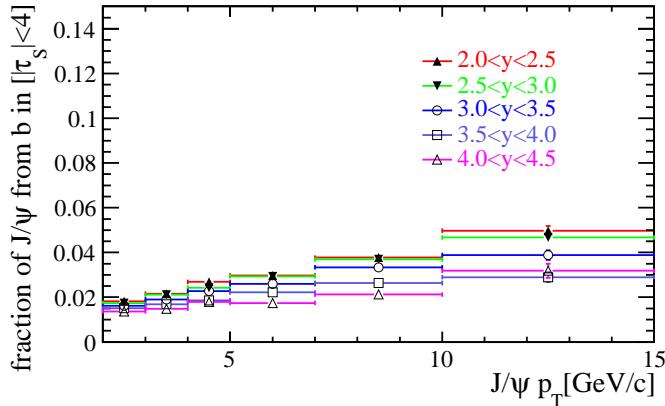


Figure 4.25 The fraction of  $J/\psi$  from long lived  $b$  hadron decay in data after  $|\tau_S| < 4$  cut.

possible bias introduced by  $b$  contamination, the  $\tau_S$  cut value is enlarged from 4 to 10, and by doing this more  $J/\psi$  from  $b$  events are accepted and as a result the retention rate of  $J/\psi$  from  $b$  increases to 1.6 – 2.4 times the rate with  $\tau_S$  cut at 4. The polarization parameters are extracted again with the new  $\tau_S$  cut, and the deviations ( $\Delta$ ) of the results from the nominal results, which are shown Figure 4.26 for one  $J/\psi$  rapidity bin, are expected

to be introduced by the extra  $b$  contamination, which corresponds to the events with  $\tau_S$  values between 4 and 10. However what is desired to know is the bias introduced by the  $b$  decay contamination when  $\tau_S$  is cut at 4 in the nominal fit, so the deviations above must be propagated to the case when the  $\tau_S$  is at 4. Suppose the contamination fraction in the prompt sample increases from  $f_1$  (about 3% averagely) to  $f_2$  (about 6% averagely) when the cut value of  $\tau_S$  changes from 4 to 10, and because the extra fraction—( $f_2 - f_1$ )—of  $J/\psi$  from  $b$  decay introduces polarization deviation  $\Delta$ , the propagated deviation is simply calculated as  $f_1/(f_2 - f_1) \times \Delta$ . The propagated deviation is quoted as systematic uncertainty due to  $b$  decay contamination and is listed in Table 4.11 for each bin of  $J/\psi p_T$  and rapidity.

As the fraction of  $J/\psi$  from  $b$  contamination is small (< 5%), the systematic uncertainties are not large, roughly of the same size (in high statistics bins) or less than the statistical errors. Because the  $b$  fraction increases with  $J/\psi$  transverse momentum, the size of the uncertainties increases with  $p_T$  (also affected by the difference of prompt polarization and polarization in  $J/\psi$  from  $b$ ). The residual polarization of  $J/\psi$  from  $b$  decay will alter the polarization of prompt  $J/\psi$  in the same way in different  $p_T$  and rapidity bins, so the systematic uncertainties coming from  $b$  contamination are correlated across bins.

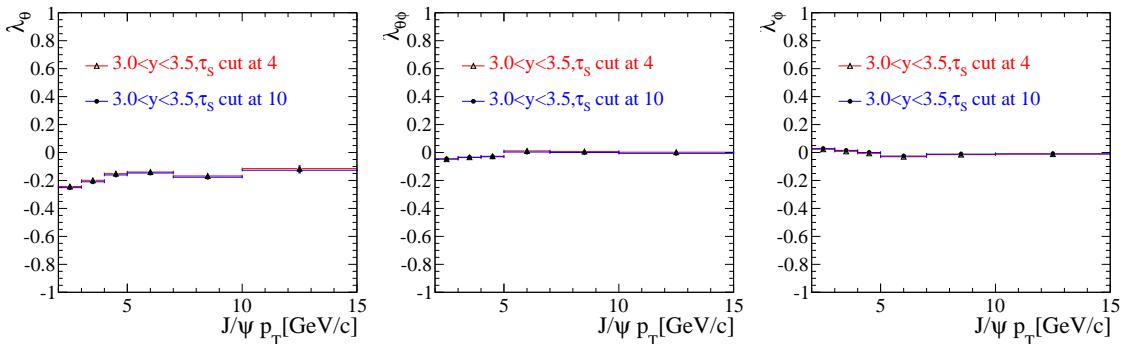


Figure 4.26 Polarization parameters  $\lambda_\theta$  (left),  $\lambda_{\theta\phi}$  (middle) and  $\lambda_\phi$  (right) extracted with  $\tau_S$  cut at 4 (triangle) and 10 (circle) respectively. Results in one rapidity bin is shown here for example.

### 4.3.8 MagUp and MagDown symmetry

In the analysis two magnet polarity data are combined because the behavior of  $\mu^+$  in the MagDown data is the same as  $\mu^-$  in MagUp data in the ideal world and it can be shown that swapping the  $\mu^+$  with a  $\mu^-$  for the same event, in the  $J/\psi$  rest frame,  $\theta$  becomes  $\pi - \theta$

Table 4.11 The  $J/\psi$  from  $b$  contamination systematic uncertainties for each  $p_T$  and rapidity bin of  $J/\psi$  bin in the HX frame.

$p_T$ (GeV/c)		$y$			
$2 < P_T < 3$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0017	0.0118	0.0096	0.0123	0.0100
$\lambda_{\theta\phi}$	0.0014	0.0008	0.0031	0.0013	0.0003
$\lambda_\phi$	0.0047	0.0058	0.0063	0.0056	0.0033
$3 < P_T < 4$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0172	0.0129	0.0119	0.0121	0.0103
$\lambda_{\theta\phi}$	0.0052	0.0011	0.0018	0.0019	0.0010
$\lambda_\phi$	0.0029	0.0067	0.0075	0.0068	0.0024
$4 < P_T < 5$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0082	0.0122	0.0121	0.0138	0.0154
$\lambda_{\theta\phi}$	0.0001	0.0051	0.0052	0.0034	0.0033
$\lambda_\phi$	0.0051	0.0062	0.0083	0.0063	0.0046
$5 < P_T < 7$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0142	0.0117	0.0089	0.0109	0.0121
$\lambda_{\theta\phi}$	0.0061	0.0066	0.0096	0.0069	0.0035
$\lambda_\phi$	0.0024	0.0057	0.0064	0.0065	0.0021
$7 < P_T < 10$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0106	0.0144	0.0126	0.0111	0.0149
$\lambda_{\theta\phi}$	0.0032	0.0058	0.0104	0.0097	0.0045
$\lambda_\phi$	0.0010	0.0029	0.0051	0.0053	0.0033
$10 < P_T < 15$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\lambda_\theta$	0.0193	0.0177	0.0129	0.0164	0.0177
$\lambda_{\theta\phi}$	0.0106	0.0060	0.0098	0.0066	0.0071
$\lambda_\phi$	0.0009	0.0002	0.0032	0.0017	0.0011

and  $\phi$  becomes  $\pi + \phi$ , and the polarization formula (see Equation 3-3) is invariant under this transformation.

However in the real case, the detector may not follow such symmetry. The two magnet polarity data can be analyzed separately, and for each data set independent polarization parameters can be extracted. In Figures 4.27, the results for each magnet polarity are shown, together with the nominal results. It can be seen that the measurements are consistent between data with the two magnet polarities.

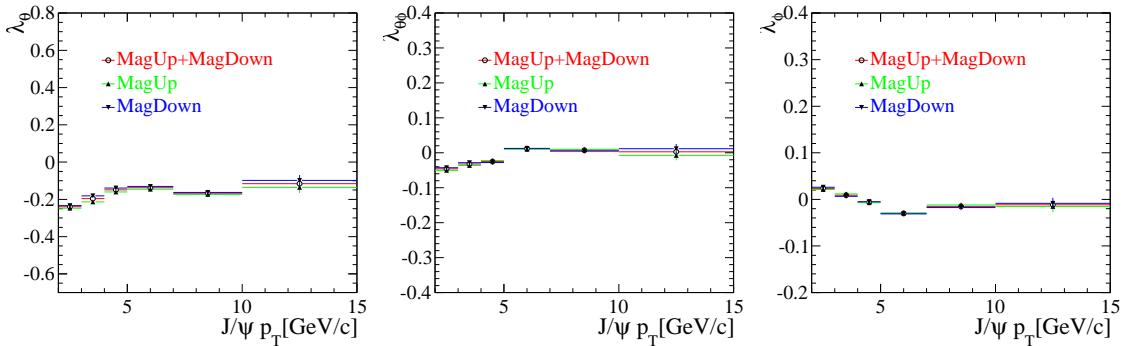


Figure 4.27 Polarization parameters  $\lambda_\phi$  (left),  $\lambda_{\theta\phi}$  (middle) and  $\lambda_\phi$  (right) extracted with MagDown (triangle down) and MagUp (triangle up) data separately. The nominal fit (circle) when MagDown and MagUp data combined are also shown. The figures for  $J/\psi$  rapidity in  $3.0 < y < 3.5$  bin.

### 4.3.9 Summary of the uncertainties

The uncertainties of the parameters in the HX frame, studied in details in previous sections, are summarized in Tables B.5, B.6, B.7, B.8 and B.9. In Tables B.10, B.11, B.12, B.13 and B.14, the uncertainties of the parameters in the CS frame are summarized.

The uncertainties in general have the following features:

- Likelihood uncertainties: the statistical uncertainties due to the fluctuation in signal events, and they depend on the rapidity and transverse momentum bin.
- Normalization systematic uncertainties: the fluctuations due to normalization of the polarization angular distribution multiplied by the efficiency. They are directly related to the finite size of Monte Carlo. Because Monte Carlo has fewer signal events than data, the errors from Monte Carlo are generally 1.5-2.0 times the errors given by the estimator.
- Background subtraction systematic uncertainties: the effect is quite negligible at higher  $p_T$  and higher  $y$  bins where the background fraction is smaller. The size of

the uncertainties are of the order of statistical errors in low transverse momentum and rapidity bins.

- Cuts/selections systematic uncertainties: each one of the fluctuations due to the cuts/selections (binning effect, detection resolution effect,  $J/\psi$  from  $b$  contamination and  $\tau_S$  cut effect) is quite below the statistical errors in almost all rapidity and  $p_T$  bins. However for some bins, especially for low  $p_T$  small  $y$  bins, the binning effect dominates the systematic uncertainties.
- Tracking systematic uncertainty: the effect is below or at the same order of the statistical fluctuations.
- Acceptance systematic uncertainty: the effect is the largest one for most of the bins. For low statistics bins, it introduces systematic uncertainties with similar size as statistical errors, while for bins with high statistics the systematic uncertainties can be several times the statistical errors.

In Table 4.12 and Table 4.13, the various uncertainties including the statistical errors from the estimators are shown for  $\lambda_\theta$  in the HX frame and the CS frame respectively with the minimum, maximum and average values among all the  $J/\psi$   $p_T$  and rapidity bins listed. In the HX frame, typically the errors for  $\lambda_{\theta\phi}$  and  $\lambda_\phi$  are much smaller (1/3 or 1/2) than for  $\lambda_\theta$ . While in the CS frame, the three polarization parameters are combination of the polarization parameters in the HX frame, and the errors for  $\lambda_{\theta\phi}$  and  $\lambda_\phi$  are found to be of similar size as  $\lambda_\theta$ .

Concerning the comparison of the uncertainties of the parameters in the HX frame and the CS frame, the statistical errors — coming from the estimator and sideband subtraction — are very similar as they are measures of the numbers of signal and background events in data. On the other hand, other systematic uncertainties differ clearly as expected because the systematic factors affect the measurement differently in different frames, and this is one of the reasons that this analysis presents results in different frames. However it is also checked that the invariant parameter defined in Equation 3-6 is almost the same for most of the bins with the differences very much below the statistical fluctuations.

## 4.4 Results

In Figure 4.28, Figure 4.29 and Figure 4.30, the fitted polarization parameters  $\lambda_\theta$ ,  $\lambda_{\theta\phi}$  and  $\lambda_\phi$  for prompt  $J/\psi$  as a function of the transverse momentum  $p_T$  are shown for various rapidity bins in both the HX frame and the CS frame. The uncertainty for each point is

Table 4.12 List of the main contributions to the uncertainties of the parameter  $\lambda_\theta$  in the HX frame. The absolute uncertainty is reported. Since the measurement is made on many transverse momentum and rapidity bins, the minimum and maximum values are shown together with the average.

Source	average	min.	max.	comment
statistical error from estimator	0.010	0.005	0.083	Bin dependent
MC efficiency (acceptance)	0.060	0.001	0.224	Correlated between bins
Binning effect	0.018	0.001	0.165	Bin dependent
Normalization from MC	0.015	0.005	0.127	Bin dependent
Sideband subtraction	0.016	0.001	0.099	Correlated between bins
$b$ -hadrons contamination	0.012	0.002	0.019	Correlated between bins
Tracking	0.012	0.003	0.029	Correlated between bins
bkg subtraction statistical error	0.004	0.002	0.045	Bin dependent

Table 4.13 List of the main contributions to the uncertainties of the parameter  $\lambda_\theta$  in the CS frame. The absolute uncertainty is reported. The minimum and maximum values are shown together with the average among all the  $J/\psi$  kinematic bins.

Source	average	min.	max.	comment
statistical error from estimator	0.019	0.004	0.096	Bin dependent
MC efficiency (acceptance)	0.027	0.005	0.071	Correlated between bins
Binning effect	0.016	0.001	0.129	Bin dependent
Normalization from MC	0.031	0.007	0.170	Bin dependent
Sideband subtraction	0.029	0.001	0.183	Correlated between bins
$b$ -hadrons contamination	0.006	0.002	0.029	Correlated between bins
Tracking	0.021	0.003	0.051	Correlated between bins
bkg subtraction statistical error	0.008	0.002	0.070	Bin dependent

the quadratical sum of the statistical uncertainty and systematic uncertainties.

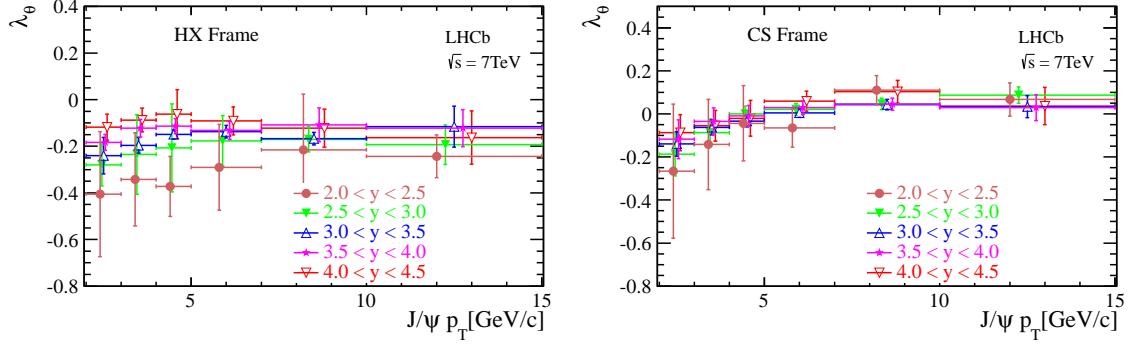


Figure 4.28  $\lambda_\theta$  in different  $p_T$  bins for the five rapidity bins in the HX frame (left) and the CS frame (right) respectively, the uncertainties are summed quadratically.

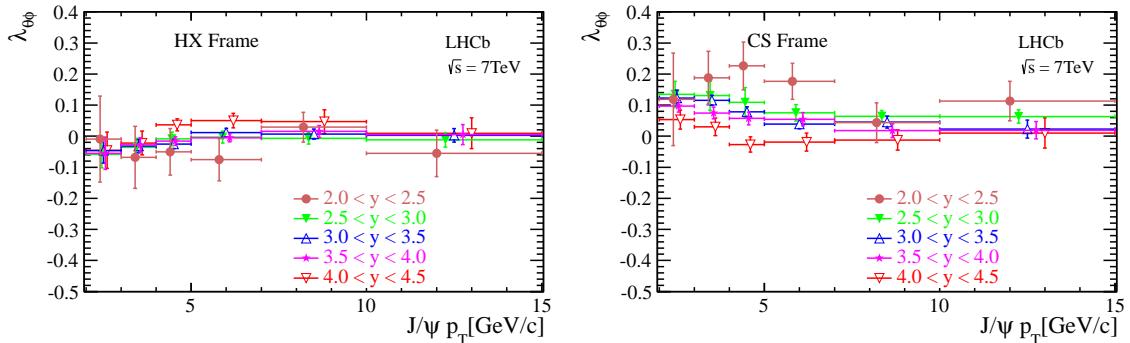


Figure 4.29  $\lambda_{\theta\phi}$  in different  $p_T$  bins for the five rapidity bins in the HX frame (left) and the CS frame (right) respectively, the uncertainties are summed quadratically.

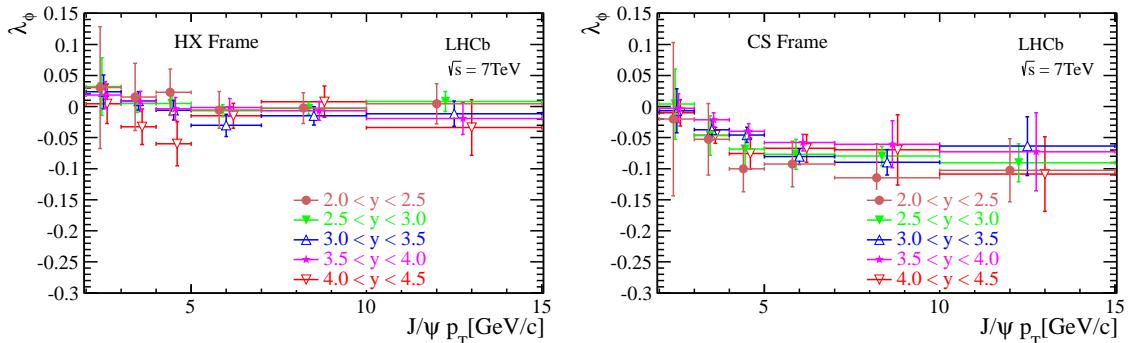


Figure 4.30  $\lambda_\phi$  in different  $p_T$  bins for the five rapidity bins in the HX frame (left) and the CS frame (right) respectively, the uncertainties are summed quadratically.

In Tables B.5 ⋯ B.9 (for the HX frame) and Tables B.10 ⋯ B.14 (for the CS frame) all the polarization parameters and the calculated uncertainties are listed.

In Figure 4.31, Figure 4.32 and Figure 4.33 the results for the three parameters are shown integrating over the rapidity range between 2.0 and 4.5. The error bars represent the statistical uncertainties while the boxes are the statistical and systematics uncertainties added in quadrature.

The first consideration that can be made is that  $\lambda_\phi$  and  $\lambda_{\theta\phi}$  in the HX frame is consistent with zero, and the errors for the two parameter are considerably low. This is a remarkable fact and has an important consequence. In other words,  $\lambda_\phi \sim \lambda_{\theta\phi} \sim 0$  tells us that the measured  $\lambda_\theta$  parameter is a direct measurement of the  $J/\psi$  polarization. In fact the invariant parameter:

$$\lambda_{inv} = \frac{\lambda_\theta + 3\lambda_\phi}{1 - \lambda_\phi} \quad (4-11)$$

is essentially equal to  $\lambda_\theta$  in the HX frame case. While in the CS frame, the  $\lambda_\phi$  starts with zero at low  $p_T$  and goes to about  $-0.1$  at higher  $p_T$ , and  $\lambda_\theta$  has values around zero, increasing slightly with  $p_T$ .

Although with a small value,  $\lambda_\theta$  in the HX frame results to be negative ( $\approx -0.2$ ), so the  $J/\psi$  has a slight longitudinal polarization. This polarization (absolute value) decreases from low  $p_T$  and low  $y$  high  $p_T$  to rapidity  $y$ . The results can be compared with the ones obtained by other experiments. For what concerns the  $\lambda_\theta$  parameter CDF, PHENIX and HERA-B also show a slightly longitudinal polarization. For what concerns the other parameters only HERA-B measured  $\lambda_\phi$  and  $\lambda_{\theta\phi}$ , obtaining results very closed to zero, consistent with the LHCb values. However one should note that the kinematic ranges are different between LHCb and the other three experiments.

At LHC the ALICE collaborations have also studied the  $J/\psi$  polarization in  $pp$  collisions at  $\sqrt{s} = 7$  TeV, looking at its decay into a muon pair. The polar and azimuthal angles  $\theta$  and  $\phi$  have been obtained in the HX frame and the CS frame and their 1-D distributions have been studied to extract the  $\lambda_\theta$  and  $\lambda_\phi$  parameters (since only the  $\cos \theta$  and  $\phi$  single variable distributions have been studied in the ALICE analysis, the parameter  $\lambda_{\theta\phi}$  has not been measured). The ALICE analysis does not discriminate prompt  $J/\psi$  from those from  $b$  decays. The measurement has been performed in bins of  $J/\psi$  transverse momentum and integrating over the rapidity in a kinematic range very similar to LHCb, being  $2.5 < y < 4.0$  and  $2 \text{ GeV}/c < p_T < 8 \text{ GeV}/c$ . Thus a comparison with the LHCb results is possible and an overall good agreement is found between the two measurements for both the  $\lambda_\theta$  and  $\lambda_\phi$  parameters in the HX frame and the CS frame, respectively. The

comparisons are shown in Figures 4.34 and 4.35 for the  $\lambda_\theta$  and  $\lambda_\phi$  parameters respectively.

Theoretical predicitons<sup>[188]</sup> for  $J/\psi$  polarization in color singlet and NRQCD approaches in LHCb kinematic region are superimposed with the measurements in Figure 4.36, 4.37, and Figure 4.38 in the HX frame. From these plots, it can be seen that neither the size of the polarization parameters nor their  $p_T$  dependence coincide with either the NLO CS or the NLO NRQCD computations, which have quite large transverse (CS) or longitudinal (NRQCD) polarization parameter  $\lambda_\theta$ . However, one should note that our measurements are performed with prompt  $J/\psi$  events, while the theoretical calculations do not include feed-downs from excited states.

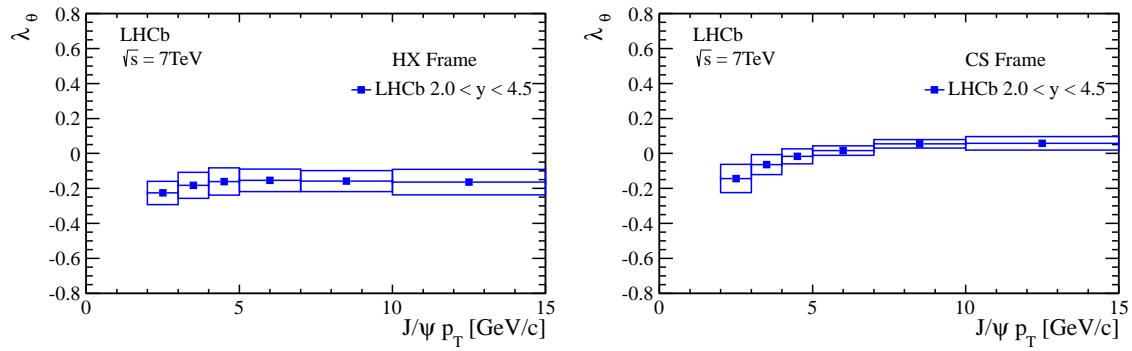


Figure 4.31  $\lambda_\theta$  in different  $p_T$  bins integrating over the rapidity range  $[2.5, 4.0]$  for the HX frame (left) and the CS frame (right) respectively. Error bars represent the statistic uncertainties while the boxes are the statistic and systematics uncertainties added in quadrature.

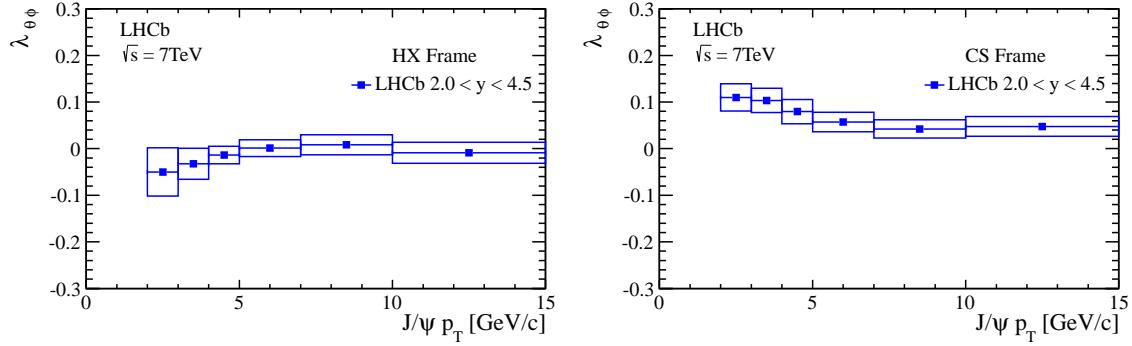


Figure 4.32  $\lambda_{\theta\phi}$  in different  $p_T$  bins integrating over the rapidity range  $[2.5, 4.0]$  for the HX frame (left) and the CS frame (right) respectively. Error bars represent the statistic uncertainties while the boxes are the statistic and systematics uncertainties added in quadrature.

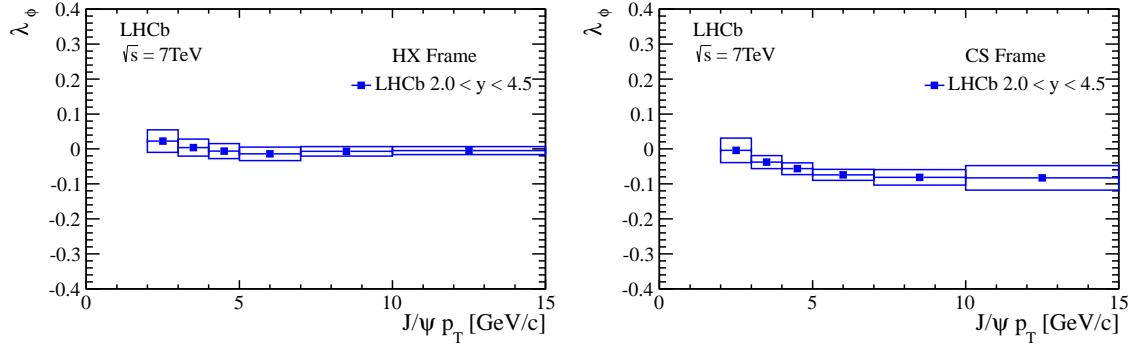


Figure 4.33  $\lambda_\phi$  in different  $p_T$  bins integrating over the rapidity range  $[2.5, 4.0]$  for the HX frame (left) and the CS frame (right) respectively. Error bars represent the statistic uncertainties while the boxes are the statistic and systematics uncertainties added in quadrature.

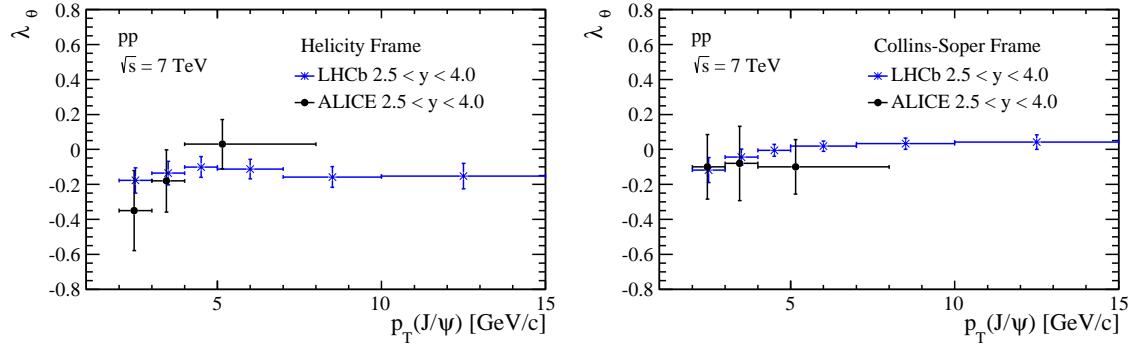


Figure 4.34 LHCb (cross) and ALICE (circle) results for  $\lambda_\theta$  in different  $p_T$  bins integrating over the rapidity range for the HX frame (left) and the CS frame (right) respectively. Statistical and systematic uncertainties are added in quadrature.

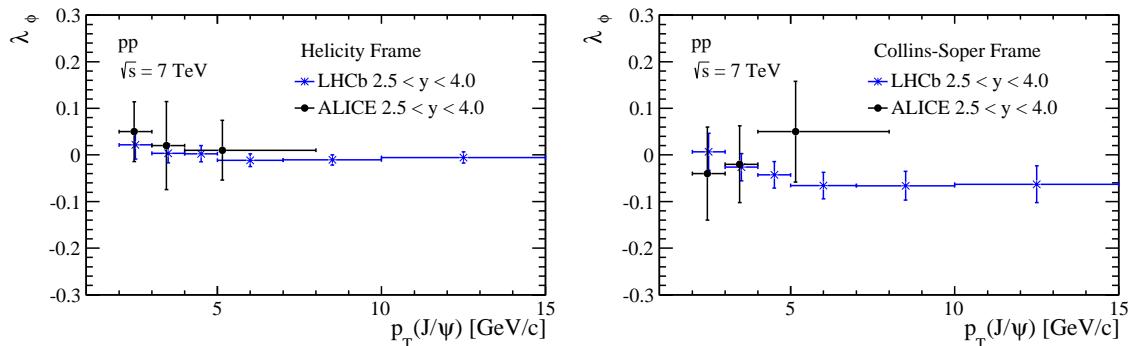


Figure 4.35 LHCb (cross) and ALICE (circle) results for  $\lambda_\theta$  in different  $p_T$  bins integrating over the rapidity range for the HX frame (left) and the CS frame (right) respectively. Statistical and systematic uncertainties are added in quadrature.

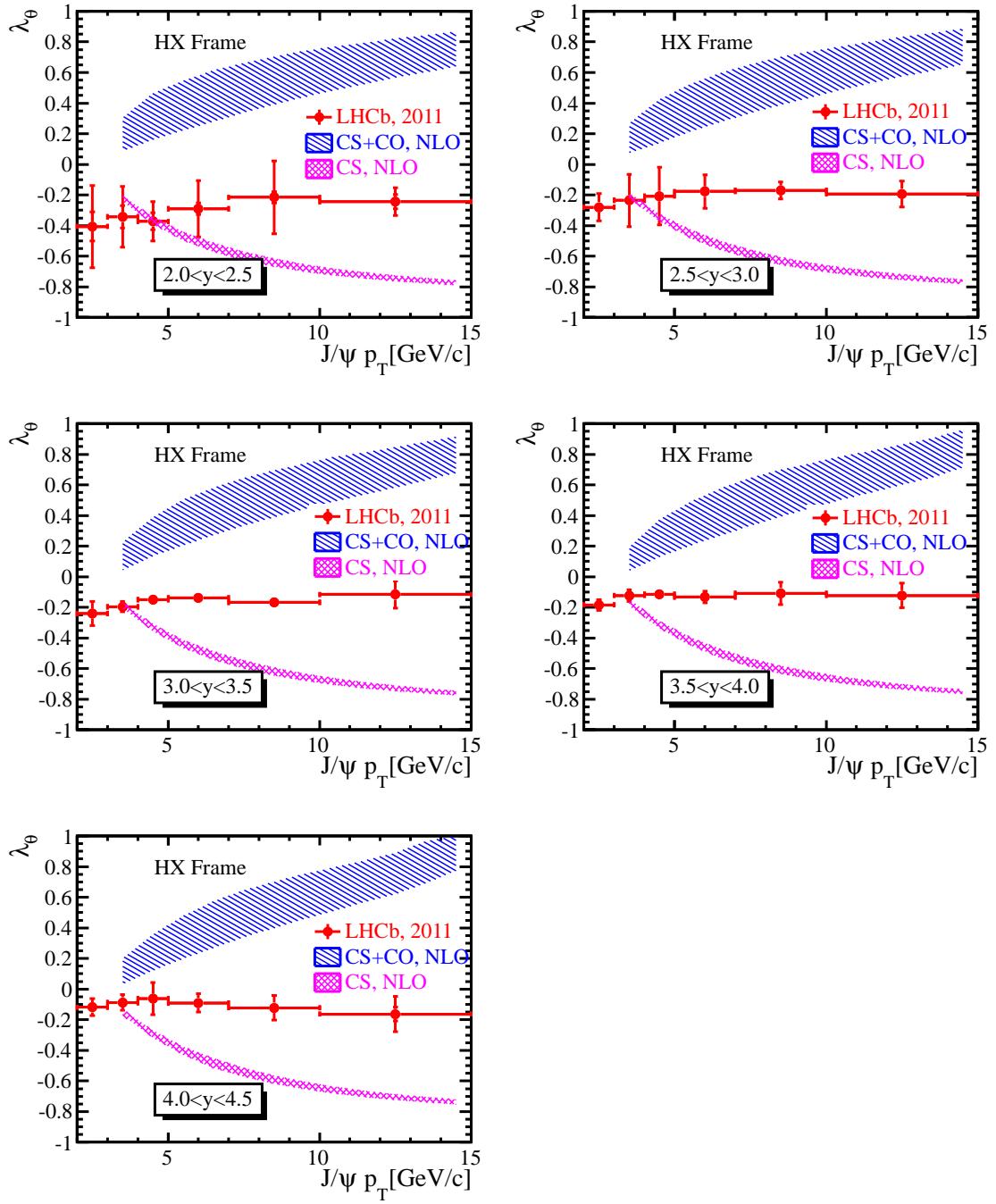


Figure 4.36 Prompt  $J/\psi$  polarization parameter  $\lambda_\theta$  measured at LHCb (points with error bars) compared with direct NLO color singlet (shaded purple) and NLO NRQCD (shaded blue) predictions.

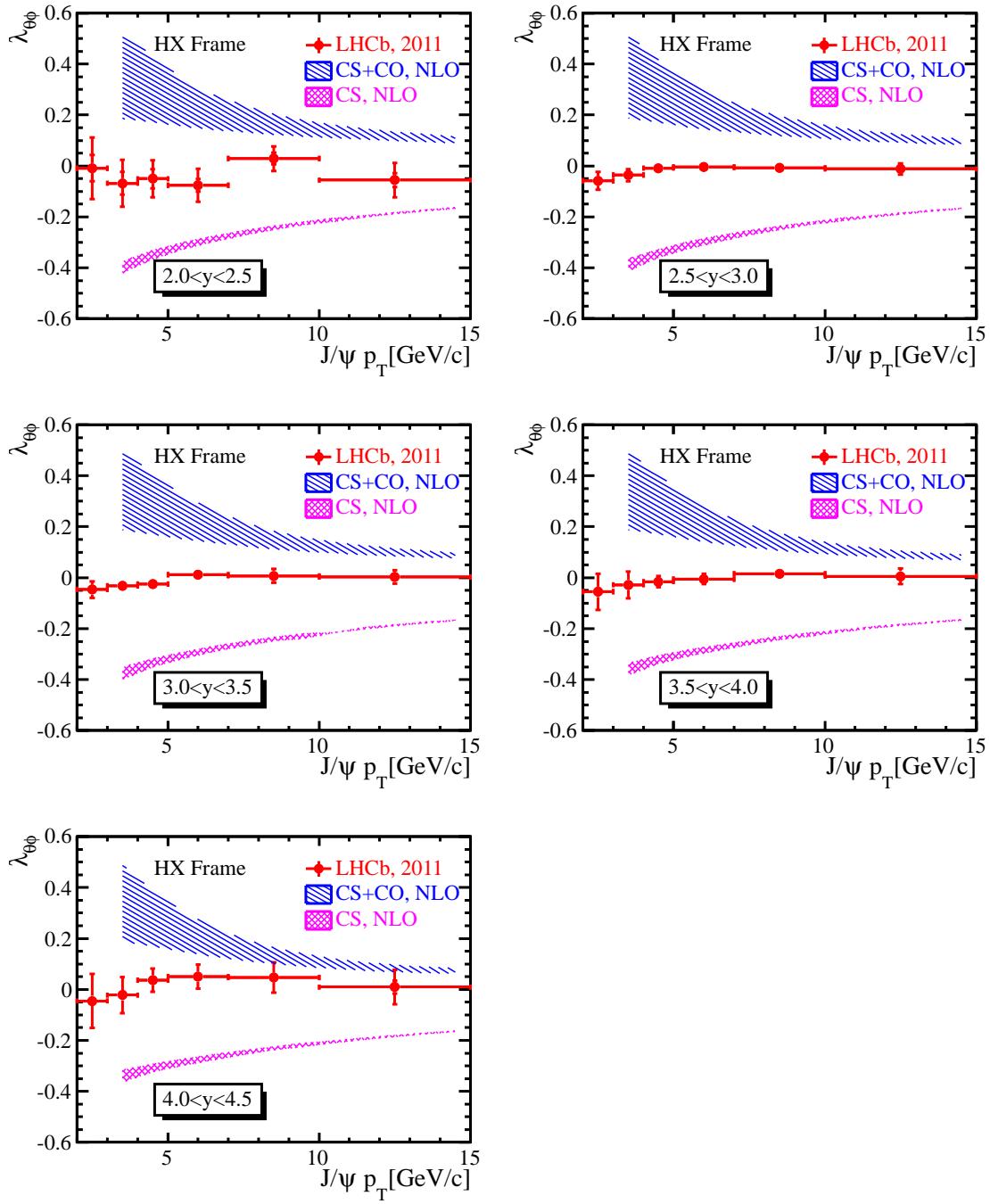


Figure 4.37 Prompt  $J/\psi$  polarization parameter  $\lambda_{\theta\phi}$  measured at LHCb (points with error bars) compared with direct NLO color singlet (shaded purple) and NLO NRQCD (shaded blue) predictions.

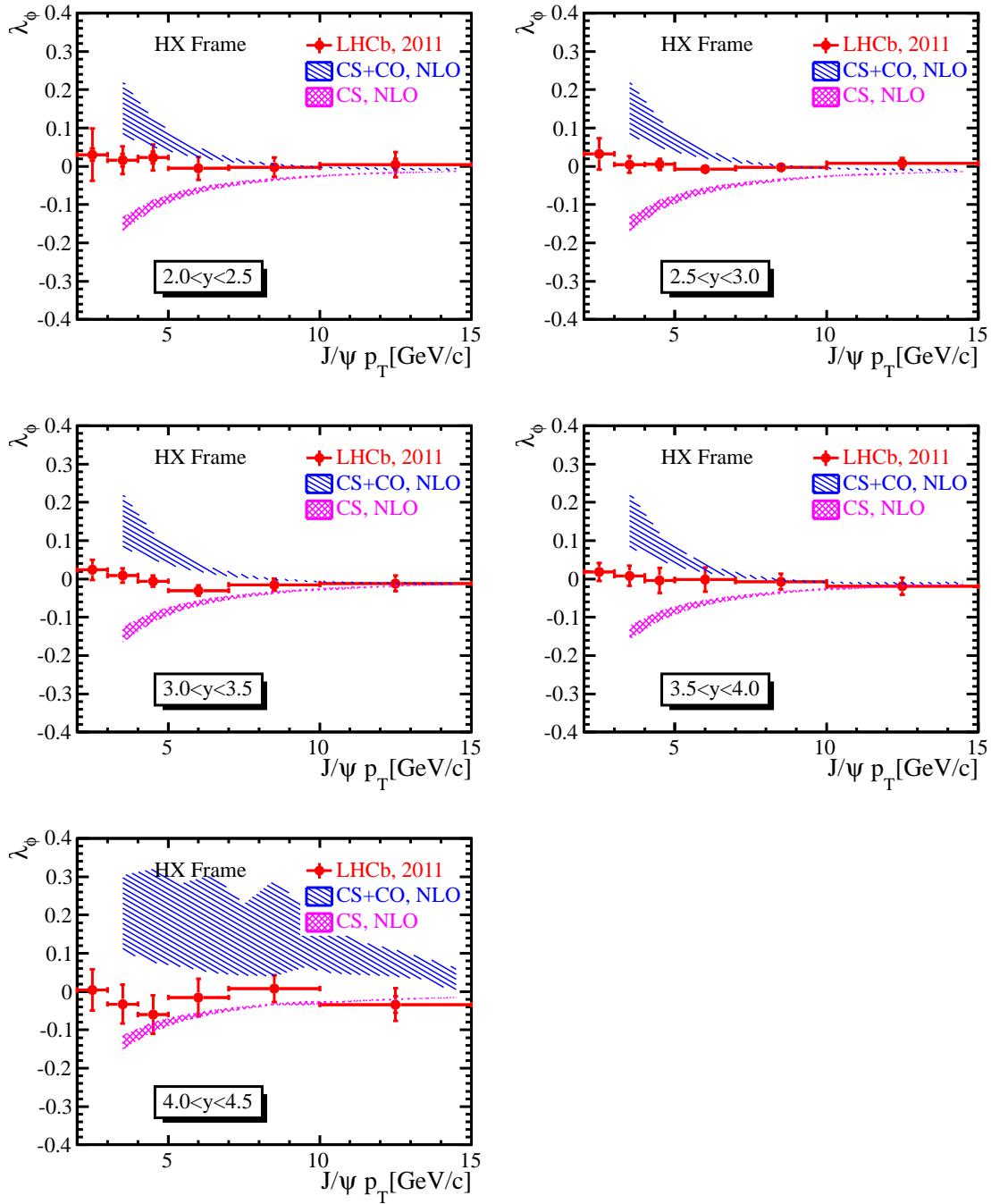


Figure 4.38 Prompt  $J/\psi$  polarization parameter  $\lambda_\phi$  measured at LHCb (points with error bars) compared with direct NLO color singlet (shaded purple) and NLO NRQCD (shaded blue) predictions.

## Chapter 5 **Prompt $\psi(2S)$ polarization measurement**

This chapter describes the analysis of the  $\psi(2S)$  polarization measurement in the decay  $\psi(2S) \rightarrow \mu^+\mu^-$ . Firstly, the data set and selections are discussed, followed by the determination of the systematic uncertainties, and in the end the results will be presented. In general the procedure to extract the polarization parameters and methods of evaluating the systematic uncertainties are the same as for the  $J/\psi$  polarization measurement presented in the Chapter 4.

### 5.1 Data sets and Selections

#### 5.1.1 Data sets

The production cross section for  $\psi(2S)$  <sup>[148]</sup> is about one order of magnitude smaller than the  $J/\psi$  cross section; since the decay branching fraction of  $\psi(2S) \rightarrow \mu^+\mu^-$  ( $\approx 7.7 \times 10^{-3}$ ) is much smaller than the decay of  $J/\psi \rightarrow \mu^+\mu^-$  ( $\approx 5.9 \times 10^{-2}$ ), the number of signal  $\psi(2S)$  events is about 2% of the number of  $J/\psi$  events in the dimuon channel. However, as the mass of  $\psi(2S)$  ( $\sim 3686 \text{ MeV}/c^2$ ) and the mass of  $J/\psi$  ( $\sim 3097 \text{ MeV}/c^2$ ) does not differ too much, the combinatorial background will be of similar level. So generally, for the  $\psi(2S)$  candidates, the signal to background ratio is much smaller, so tighter cuts are applied to suppress the background pollution.

The  $\psi(2S) \rightarrow \mu^+\mu^-$  decay in all the LHCb 2011  $pp$  collision data corresponding to integrated luminosity of around  $1.0 \text{ fb}^{-1}$  is analyzed; the data are taken with a center of mass energy of 7 TeV with both magnet polarities. The  $\psi(2S)$  events are reconstructed from the line *MicroDSTDiMuonDiMuonIncLine* of the Leptonic stream in the LHCb stripping processing *Reco12 – Stripping17*, and in the stripping some pre-selection cuts are applied to the  $\psi(2S)$  candidates.

As the instantaneous luminosity changed with time during the data taking, generally two intrinsically different sets of TCKs are used, the TCKs before the summer technical stop which corresponds to the same  $370 \text{ pb}^{-1}$  data used in the  $J/\psi$  polarization measurement and the TCKs after the technical stop which covers the other  $700 \text{ pb}^{-1}$ . The first TCK group is the same as described in Table 4.1, and the second one contains the TCKs listed in 5.1. The TCKs within each group have homogenous setting concerning

Table 5.1 The TCKs and the corresponding luminosities in the second half of 2011 data taking.

TCK	MagUp (pb <sup>-1</sup> )	MagDown (pb <sup>-1</sup> )	Total (pb <sup>-1</sup> )
0x00760037	107.1	191.6	298.7
0x00790037	40.3	-	40.3
0x00790038	154.0	209.4	363.4

 Table 5.2 The cuts in the two Hlt2 lines used in the selections of  $\psi(2S)$ .

Hlt2 line	representative TCK	specific requirements
<i>Hlt2DiMuonPsi2S</i>	0x00730035	$M_{\psi(2S)} - 120 \text{ MeV}/c^2 < M_{\mu^+\mu^-}$ $M_{\mu^+\mu^-} < M_{\psi(2S)} + 120 \text{ MeV}/c^2$ Vertex quality: $\chi^2/\text{ndof} < 25$ Track quality $\chi^2/\text{ndof} < 5$ for both muons
<i>Hlt2DiMuonPsi2S</i>	0x00730035	$M_{\psi(2S)} - 100 \text{ MeV}/c^2 < M_{\mu^+\mu^-}$ $M_{\mu^+\mu^-} < M_{\psi(2S)} + 100 \text{ MeV}/c^2$ $p_T(\psi(2S)) > 3.5 \text{ GeV}/c$ Vertex quality: $\chi^2/\text{ndof} < 25$ Track quality $\chi^2/\text{ndof} < 5$ for both muons

our signals. About the trigger lines,  $\psi(2S)$  are required to be triggered by L0 *Muon* or L0 *DiMuon* of the Level-0 triggers, and by *Hlt1DiMuonHighmass* of the Hlt1 trigger. The selections in the trigger lines L0 *Muon*, L0 *DiMuon* and *Hlt1DiMuonHighmass* are described in Table 4.2. For the Hlt2 trigger, the  $\psi(2S)$  is required to be triggered by the *Hlt2DiMuonPsi2S* for the data taken in the first half of the year, and by *Hlt2DiMuonHPT* for rest. Comparing the two Hlt2 trigger lines, the *Hlt2DiMuonPsi2SHighPT* line has an extra selection on  $\psi(2S)$   $p_T$ . The details of the requirements in the two Hlt2 lines can be found in Table 5.2. When the reconstructed dimuon mass is not far away from the PDG value, in the region where the  $\psi(2S)$  has transverse momentum larger than 3.5 GeV/ $c$ , the two sets of TCKs have identical effect on the selections. So the two data sets are combined in the region  $p_T(\psi(2S)) > 3.5 \text{ GeV}/c$  for the analysis.

To study the efficiency, a Monte Carlo sample with 10 million events — 5 million with MagUp and 5 million with MagDown magnet polarity setup — is generated with the simulation condition tag MC11a-Sim05c. The TCK 0x40760037 is used in the

Monte Carlo. The selections in TCK 0x40760037 is identical to the TCK 0x00760037 concerning our signal. The  $\psi(2S)$  polarization in the Monte Carlo is set to zero ( $\lambda_\theta = \lambda_{\theta\phi} = \lambda_\phi = 0$ ), so the angular ( $\cos \theta - \phi$ ) distribution of muons in the  $\psi(2S)$  rest frame is uniform.

### 5.1.2 Selections

In the stripping line *MicroDSTDiMuonDiMuonIncLine*, the following cuts are applied to the vertex reconstructed from two muon tracks with opposite charge:

- $\mu$  track transverse momentum:  $p_T(\mu) > 650 \text{ MeV}/c$
- $\mu$  track quality:  $\chi^2(\mu \text{ track})/\text{ndof} < 5$
- vertex quality:  $\chi^2/\text{ndof} < 20$
- vertex mass:  $3000 \text{ MeV}/c^2 < M(\mu^+\mu^-) < 4000 \text{ MeV}/c^2$
- vertex transverse momentum:  $p_T > 3 \text{ MeV}/c$

The invariant mass distribution of  $\psi(2S)$  candidates after the pre-selection (stripping selection) is shown in Figure 5.1, in which it can be seen that the background level is very high. To reduce the combinatorial background reconstructed from fake muons, which are

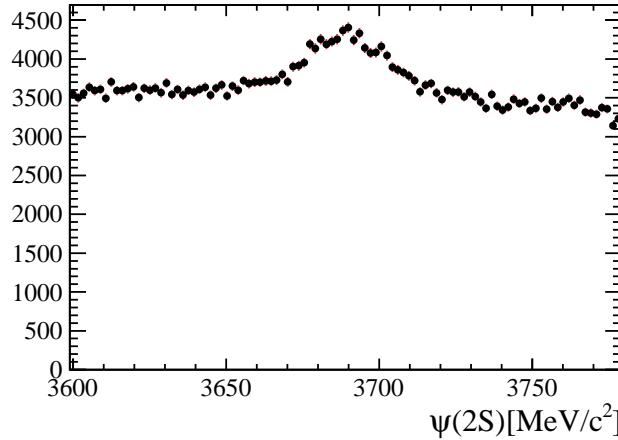


Figure 5.1 The invariant mass distribution of  $\psi(2S)$  candidates after the stripping cuts.

mostly ghost tracks and muons from  $\pi/K$  decay in flight, muon tracks are further required to be less ghost-like, and tighter cuts on muon transverse momentum and PID are also applied. The  $p_T$  of the muons is required to be larger than  $1 \text{ GeV}/c$ ; with this selection, the background events are heavily filtered, while the number of signal events and the muon angular distribution in  $\psi(2S)$  rest frame are not significantly changed, as shown in

Table 5.3 The offline cuts used to select  $\psi(2S)$ .

Quantity	Requirement
$\mu$ transverse momentum	$p_T(\mu) > 1000 \text{ MeV}/c$
$\mu$ track quality	$\chi^2(\mu \text{ track})/\text{ndof} < 4$
$\mu$ PID	$\Delta \log \text{PID}_{\mu/\pi} > 3$
Clone killing	keep only one candidate if $\cos \theta(\mu_1^\pm, \mu_2^\pm) > 0.9999$
$\psi(2S)$ vertex quality	$\text{prob}(J/\psi \text{ vertex } \chi^2/\text{ndof}) > 0.5\%$

Figure 5.2. The difference of muon hypothesis (logarithm value) against pion hypothesis (logarithm value), or  $\Delta \log \text{PID}_{\mu/\pi}$ , is required to be larger than 3; this value almost maximizes the signal significance defined as  $S / \sqrt{S + B}$ , which is the number of signals over the fluctuation of all observed events. The effect of the PID cut on the muon angular distribution is quite small (see Figure 5.2), so no significant systematic uncertainties will be introduced. With the transverse momentum and PID selections on muons, about 80% background events are removed, while 70% signals are kept. The requirements used offline are listed in Table 5.3.

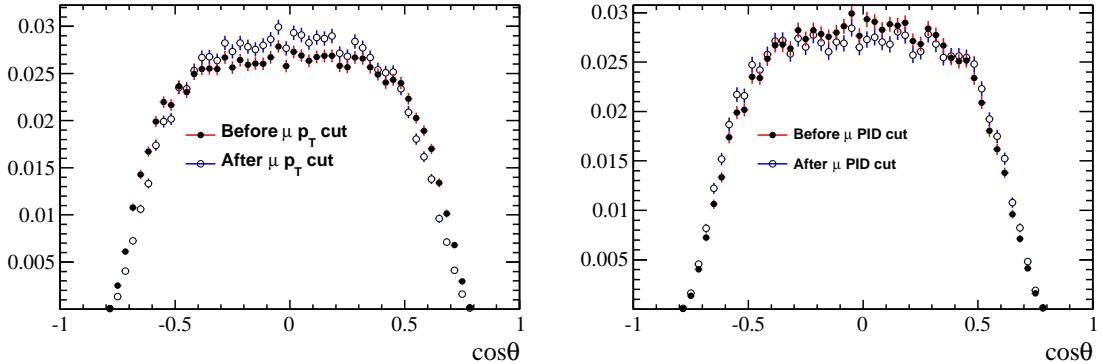


Figure 5.2 The  $\cos \theta$  distribution before (filler circle) and after (open circle) the muon transverse momentum (left) and PID (right) cut in Monte Carlo.

In the Monte Carlo sample, the LHCb detector is not well simulated in the outer boundary region (with  $\eta \approx 2$ ); in Figure 5.3 the muon pseudo rapidity ( $\eta$ ) distributions in data and in Monte Carlo are compared, and it can be seen that data is less efficient in the area  $\eta < 2.2$ , so in the analysis candidates with at least one muon  $\eta < 2.2$  are discarded to avoid the problematic area. With this requirement, about half of the  $\psi(2S)$  candidates with rapidity smaller than 2.5 are removed, while  $\psi(2S)$  with rapidity larger than 3.0 are

almost not affected (rejection rate less than 1%).

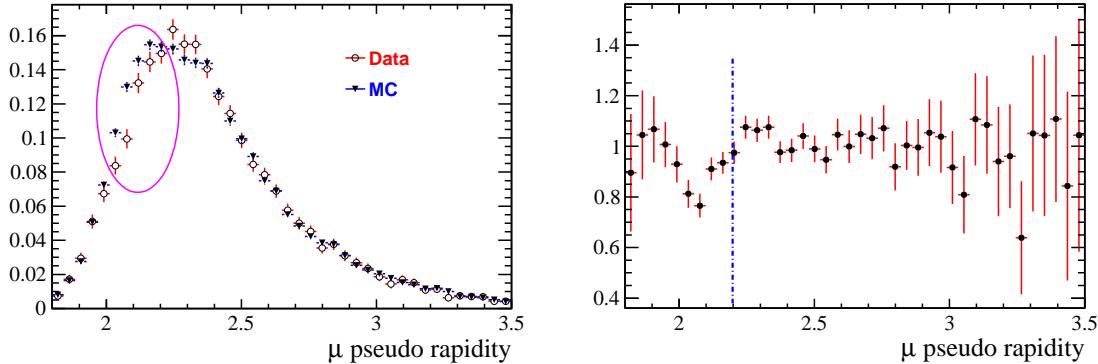


Figure 5.3 The muon pseudo rapidity distribution in data and Monte Carlo (*left*) and their ratio (*right*). The events are taken from the  $\psi(2S)$  kinematics region  $4 \text{ GeV}/c < p_T < 5 \text{ GeV}/c$ ,  $2.0 < y < 2.5$ .

To select the prompt  $\psi(2S)$  candidates, the lifetime significance ( $\tau_S$ , see section 4.1.1.1) is required to be smaller than 4 (absolute value). Monte Carlo shows that this selection keeps more than 98% of the prompt  $\psi(2S)$  mesons, on the other hand, about 90% of the  $\psi(2S)$  from  $b$ -hadron (detached  $\psi(2S)$  ) decay are rejected. As there is about 20% to 30%  $\psi(2S)$  from  $b$  in real data, the fraction of detached  $\psi(2S)$  in the sample after the  $\tau_S$  is around 2 – 3%.

Because the efficiency depends on the angular variables, after all these cuts the muon angular distribution is not uniform anymore, which is similar to the case in  $J/\psi$  polarization measurement. The shape of the efficiency as a function of muon  $\cos \theta - \phi$  is not uniform in different  $\psi(2S)$  kinematic region, so to reduce the systematic uncertainties due to the discrepancy of  $\psi(2S)$  kinematics and to study the  $p_T - y$  dependent polarization, the following binning schemes similar to the  $J/\psi$  polarization analysis are used:

$$\begin{aligned} p_T &: (3.5, 4], (4, 5], (5, 7], (7, 10], (10, 15] \text{ GeV}/c \\ y &: (2.0, 2.5], (2.5, 3.0], (3.0, 3.5], (3.0, 4.0], (4.0, 4.5] \end{aligned} \quad (5-1)$$

Only  $\psi(2S)$  with transverse momentum larger than  $3.5 \text{ GeV}/c$  can be analyzed because of the transverse momentum requirement/selection on  $\psi(2S)$  in the Hlt2 trigger. With the choice of the bin width above, the shape of the efficiency within one particular bin can be taken constant in different area of  $\psi(2S)$  kinematics phase space; the efficiency is approximately only a function of  $\cos \theta - \phi$  but independent of  $\psi(2S)$   $p_T$  or  $y$ . Besides, within

one particular bin the dependence of the polarization parameters on  $\psi(2S)$  kinematics (if any) is also ignored; only the average polarization within each bin is measured.

## 5.2 Polarization Fit

To extract polarization in data, the weighted logarithm likelihood estimator introduced in section 3.2 is used:

$$\log \mathcal{L} = \sum_{i=1}^{N_{\text{tot}}} w(m_i) \times \log \left[ \frac{P(\cos \theta_i, \phi_i | \lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi)}{\text{Norm}(\lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi)} \right], \quad (5-2)$$

where

$$P(\cos \theta, \phi) = (1 + \lambda_\theta \cos^2 \theta + \lambda_{\theta\phi} \sin 2\theta \cos \phi + \lambda_\phi \sin^2 \theta \cos 2\phi)$$

and

$$\begin{aligned} \text{Norm}(\lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi) = & \int d\Omega \epsilon_{\text{tot}}(\Omega) + \lambda_\theta \int d\Omega \cos^2 \theta \epsilon_{\text{tot}}(\Omega) \\ & + \lambda_{\theta\phi} \int d\Omega \sin 2\theta \cos \phi \epsilon_{\text{tot}}(\Omega) + \lambda_\phi \int d\Omega \sin^2 \theta \cos 2\phi \epsilon_{\text{tot}}(\Omega). \end{aligned} \quad (5-3)$$

The likelihood is constructed from the three variables:  $\cos \theta, \phi$  and invariant mass of each candidate. In the  $J/\psi$  polarization analysis, signal region and sideband regions in the invariant mass distribution are defined to determine the weight of each event. However, due to the tight mass cut in the Hlt2 trigger line (100 MeV around the PDG  $\psi(2S)$  mass) and the relatively low  $\psi(2S)$  mass resolution ( $\sim 20 \text{ MeV}/c^2$ ), proper sideband and signal regions can not be well defined to subtract the background as in the  $J/\psi$  polarization analysis. Instead the *sFit*<sup>[189]</sup> technique is used to subtract the background. In the *sFit* method, the weight ( $w(m_i)$ ) is taken to be the signal *sPlot*<sup>[190]</sup> which is calculated for each event from the distributions of the discriminating variable (invariant mass in this analysis) of signal and background in data. *sPlot* is a general method to unfold the control variable distributions (for example, muon angular distribution in this analysis) for different sources (background and signal in the analysis) using the discriminating variable (invariant mass) whose distribution is known for each source. In the following when the distributions of variables (except mass distribution) in data for only signal or background are used, the distributions have already been unfolded for the signal or background using the *sPlot* method. Equation 5-2 is the sum of event by event likelihood which is constructed from the discriminating variables  $\cos \theta$  and  $\phi$ , so it is also a discriminating variable, and thus

the signal *sPlot* weight  $w(m_i)$  will make the background contribution into the estimator disappear effectively.

To use the *sFit* method, firstly the mass spectrum should be fitted carefully. The  $\psi(2S)$  mass spectrum (pure signal) in Monte Carlo shows that two Crystal Ball functions (see Equation 4-4) are needed and sufficient to fit the signal very well. The two CB functions share the common mean ( $\mu$ ) for the gaussian component. The fraction of the CB function with larger width ( $\sigma$ ) for the gaussian is around 30% from Monte Carlo study and thus fixed to 0.3 during the fit to data. Besides, from Figure 5.4, it can be seen that the parameter  $\alpha$  describing the power law tail is correlated to the gaussian  $\sigma$ , and the two  $\sigma$  of the two CB functions are also correlated, so during the parametrization of the signal mass distribution in data, these relationships are fixed according to the Monte Carlo:

$$\begin{aligned}\alpha &= 1.91 + 0.013 \times \sigma \\ \sigma_2 &= 2.11 + 1.46 \times \sigma_1\end{aligned}\tag{5-4}$$

The background mass distribution in data is modeled by a first order polynomial. The  $\psi(2S)$  mass distributions are fitted with a combination of two CB functions for the signal and the polynomial for the background; the fitting is performed in each bin of  $p_T$  and rapidity of  $\psi(2S)$  with the unbinned maximum likelihood method. In Figure 5.5, the mass distribution and the best fit function are shown for the bin  $5 < p_T < 7$  GeV/ $c$  and  $3.0 < y < 3.5$ .

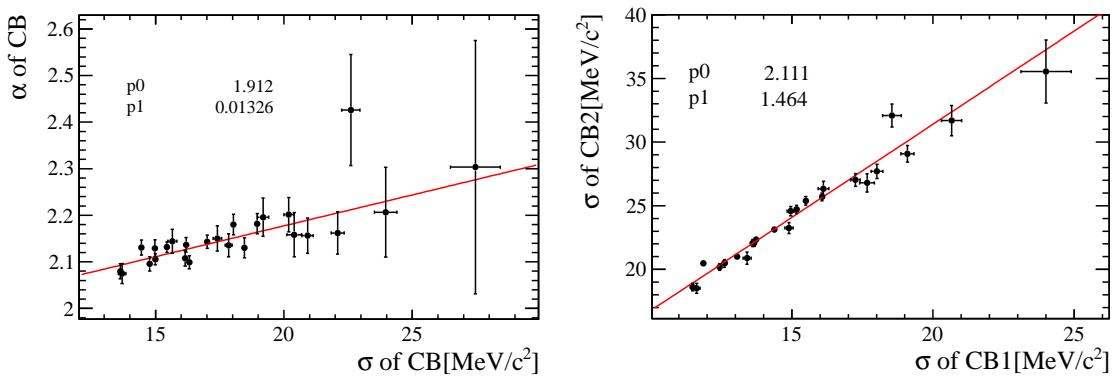


Figure 5.4 The relationship between  $\alpha$  and  $\sigma$  of the CB function (*left*) and the two  $\sigma$  of the two CB functions (*right*).

With the CB functions and the polynomial, the *sPlot* for the signal, which is assigned to  $w(m_i)$  in equation 5-2, can be constructed. On the other hand, to correctly estimate the

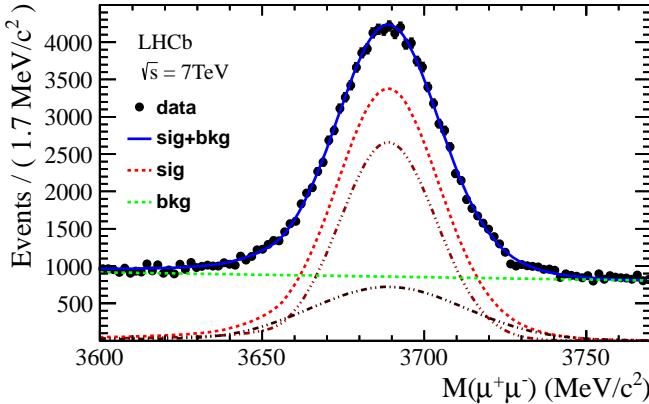


Figure 5.5 The invariant mass distribution of  $\psi(2S)$  (data points); the parametrization with two CB functions plus polynomial is superimposed. The plot are for the  $\psi(2S)$  kinematic bin  $5 \text{ GeV}/c < p_T < 7 \text{ GeV}/c, 3.0 < y < 3.5$ .

statistical error in the weighted maximum likelihood, a global scale<sup>[191]</sup> is introduced as

$$\alpha = \frac{\sum_i^{N_{\text{tot}}} w(m_i)}{\sum_i^{N_{\text{tot}}} w^2(m_i)}$$

to the likelihood as a scale factor. With this factor, the best fit values for each parameter are identical to the nominal one, however the error of each parameter increases roughly to  $1/\sqrt{\alpha}$  times the nominal one. Finally the logarithm likelihood reads:

$$\log \mathcal{L} = \alpha \sum_{i=1}^{N_{\text{tot}}} w(m_i) \times \log \left[ \frac{P(\cos \theta_i, \phi_i | \lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi)}{\text{Norm}(\lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi)} \right] \quad (5-5)$$

The normalization  $\text{Norm}(\lambda_\theta, \lambda_{\theta\phi}, \lambda_\phi)$ , which is characterized by three constants constants  $a$ ,  $b$ , and  $c$  can be calculated from the unpolarized Monte Carlo. The estimator is maximized with regard to the three polarization parameters  $\lambda_\theta$ ,  $\lambda_{\theta\phi}$  and  $\lambda_\phi$  using the TMinuit program. The parameters that maximize the estimator are the measured  $\psi(2S)$  polarization parameters.

### 5.3 Uncertainties on the polarization parameters

In the following section, the various systematic uncertainties associated with each parameter are described. The methods to extract the uncertainties are common for both HX frame and CS frame, and are similar or identical to the way used in the  $J/\psi$  polarization analysis. The methods will be illustrated in the HX frame.

### 5.3.1 Fluctuations of the normalization

The uncertainties of the normalization constants, which are determined from unpolarized Monte Carlo, can introduce bias to the polarization determination. However from one Monte Carlo sample only one set of the normalization constants can be extracted, so the fluctuations of the normalization (the normalization constants) is not taken into account automatically. In each  $\psi(2S)$  kinematic bin, to assign fluctuations to (a, b, c), toy Monte Carlo following the muon angular distributions in the baseline Monte Carlo is generated using the procedure described in 4.3.2. The normalization constants calculated from the toy Monte Carlo fluctuate with regard to the baseline (a, b, c). With each set of normalization constants computed from the toy Monte Carlo, new polarization parameters  $\lambda_\theta$ ,  $\lambda_{\theta\phi}$  and  $\lambda_\phi$  can be determined with the same data. The process is performed many times, and a list of polarization parameters for each bin can be obtained. They follow roughly a gaussian distribution with the mean very close to the baseline fit, as shown in Figure 5.6 for a randomly chosen  $\psi(2S)$  kinematic bin; the width ( $\sigma$ ) of the gaussian fit is quoted as the systematic uncertainty from fluctuation due of the normalization, which is generally determined by the number of events in the Monte Carlo. In Table 5.4 these uncertainties in HX frame are listed for each  $p_T$  and rapidity bin. The number of Monte Carlo events is almost twice the number of signal events in data, so the uncertainties arising from normalization fluctuations are smaller ( $\approx 2/3$ ) than the likelihood uncertainties (statistical uncertainties).

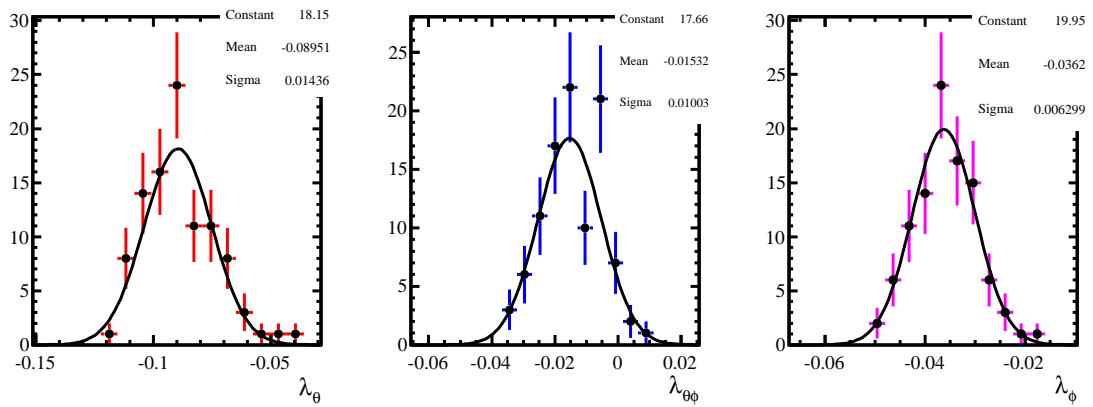


Figure 5.6 The distributions of  $\lambda_\theta$  (left),  $\lambda_{\theta\phi}$  (middle) and  $\lambda_\phi$  (right) extracted with normalization constants calculated from toy Monte Carlo. The study in the  $\psi(2S)$  kinematic bin  $5 \text{ GeV}/c < p_T < 7 \text{ GeV}/c$ ,  $3.0 < y < 3.5$  is shown for example. Only statistical uncertainties from the fit are shown.

Table 5.4 The uncertainty coming from the fluctuations of the normalization for each polarization parameters in each  $\psi(2S)$  kinematic bin bin in the HX frame.

$p_T$ (GeV/c)		y bin				
3.5 < $p_T$ < 4	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5	
	$\lambda_\theta$	0.2250	0.0442	0.0264	0.0267	0.0505
	$\lambda_{\theta\phi}$	0.1380	0.0147	0.0122	0.0136	0.0319
	$\lambda_\phi$	0.0461	0.0111	0.0115	0.0162	0.0373
4 < $p_T$ < 5	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5	
	$\lambda_\theta$	0.1210	0.0278	0.0175	0.0279	0.0616
	$\lambda_{\theta\phi}$	0.1044	0.0125	0.0109	0.0125	0.0214
	$\lambda_\phi$	0.0440	0.0090	0.0102	0.0140	0.0237
5 < $p_T$ < 7	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5	
	$\lambda_\theta$	0.0864	0.0232	0.0143	0.0219	0.0390
	$\lambda_{\theta\phi}$	0.0588	0.0120	0.0102	0.0113	0.0207
	$\lambda_\phi$	0.0286	0.0057	0.0058	0.0065	0.0172
7 < $p_T$ < 10	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5	
	$\lambda_\theta$	0.0772	0.0256	0.0214	0.0279	0.0545
	$\lambda_{\theta\phi}$	0.0520	0.0122	0.0113	0.0140	0.0284
	$\lambda_\phi$	0.0234	0.0069	0.0094	0.0126	0.0245
10 < $p_T$ < 15	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5	
	$\lambda_\theta$	0.0774	0.0285	0.0287	0.0548	0.0763
	$\lambda_{\theta\phi}$	0.0476	0.0136	0.0159	0.0192	0.0384
	$\lambda_\phi$	0.0224	0.0121	0.0119	0.0175	0.0406

### 5.3.2 Tracking efficiency

As has been discussed in section 4.3.3, studies by the LHCb tracking group show that the tracking efficiency (track finding and reconstruction) is slightly different between Monte Carlo and data. The tracking efficiency in data reconstructed with stripping 13b over the efficiency in Monte Carlo is about from 98% to 1.02%, depending on the momentum and pseudo rapidity (eta) of muons. The discrepancy is smaller for the stripping 17 data and the MC11a Monte Carlo, which are the data and Monte Carlo used for this analysis. To be conservative, it is still assumed that inconsistency of the tracking efficiency between Monte Carlo and data is at the 2% level, and the analysis studies how the polarization is

biased by this discrepancy.

The global decrease or increase in the ratio of data over Monte Carlo efficiency will not change the polarization result, however the measurement will be different if the efficiency ratio changes as a function of muon angular variables in the  $\psi(2S)$  rest frame or muon kinematics in the laboratory frame. During calculation of the normalization constants, the Monte Carlo is weighted as a function of the muon pseudo rapidity or momentum, and the polarizations are determined with normalization calculated from the weighted Monte Carlo again. The weight on each event is the product of the weight for  $\mu^+$  and the weight for  $\mu^-$ . For the shape of weighting on muon kinematics, the following scenarios (functions) are considered:

- Linear function of muon  $\eta$ :  $\text{Wt}(\eta) = p_0 + p_1 \times \eta$
- Parabola function of muon  $\eta$ :  $\text{Wt}(\eta) = p_0 + p_1 \times (\eta - 3.5)^2$
- $\Lambda$  shape function of muon  $\eta$ :  $\text{Wt}(\eta) = p_0 + p_1 \times \eta$  for  $\eta \leq 3$ , and  $q_0 + q_1 \times \eta$  for  $\eta \geq 3$
- $\Lambda$  shape function of muon  $\eta$ :  $\text{Wt}(\eta) = p_0 + p_1 \times \eta$  for  $\eta \leq 4$ , and  $q_0 + q_1 \times \eta$  for  $\eta \geq 4$
- Logarithm shape function of muon momentum:  $\text{Wt}(p) = p_0 + p_1 \times \log(p)$

In each case, the coefficients  $p_0$  and  $p_1$  are chosen such that the maximal value of each function is 1.02 while the minimal value is 0.98 in the LHCb pseudo rapidity range  $2.0 < \eta < 5.0$  and momentum coverage from a few  $\text{GeV}/c$  to hundreds of  $\text{GeV}/c$ . The two  $\Lambda$  shapes are constructed from two linear functions connected at  $\eta = 3$  or  $\eta = 4$ , respectively. For the logarithm function, the weights change rapidly at low momentum and slowly at higher momentum, which is reasonable because the larger the track momentum the narrower the  $\eta$  coverage, and tracks with similar momentum will pass through similar detector material.

In Figure 5.7, the results for various weighting scenarios are shown together with the nominal fits. The difference between the nominal result and the results from the weighted Monte Carlo is considered to be the systematic uncertainty due to the track efficiency, and the largest deviation amongst the various weighting scenarios is quoted. In Table 5.5, these uncertainties for all  $\psi(2S)$   $p_T$  and  $y$  bins in HX frame are listed. The tracking systematic errors are of level 0.02 or smaller, quite small compared to the statistical errors.

Table 5.5 The systematic uncertainties coming from the 2 percent bias in the Monte Carlo tracking efficiency using the weighting technique for each  $p_T$  and rapidity bin of  $\psi(2S)$  bin in the HX frame.

$p_T$ (GeV/c)		y bin			
3.5 < $p_T$ < 4	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5
$\lambda_\theta$	0.0023	0.0103	0.0068	0.0177	0.0286
$\lambda_{\theta\phi}$	0.0241	0.0183	0.0161	0.0148	0.0163
$\lambda_\phi$	0.0198	0.0078	0.0098	0.0128	0.0159
4 < $p_T$ < 5	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5
$\lambda_\theta$	0.0110	0.0078	0.0053	0.0188	0.0284
$\lambda_{\theta\phi}$	0.0156	0.0151	0.0129	0.0104	0.0119
$\lambda_\phi$	0.0201	0.0081	0.0101	0.0136	0.0154
5 < $p_T$ < 7	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5
$\lambda_\theta$	0.0183	0.0055	0.0054	0.0187	0.0251
$\lambda_{\theta\phi}$	0.0022	0.0100	0.0074	0.0040	0.0065
$\lambda_\phi$	0.0143	0.0071	0.0099	0.0116	0.0130
7 < $p_T$ < 10	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5
$\lambda_\theta$	0.0090	0.0020	0.0032	0.0168	0.0230
$\lambda_{\theta\phi}$	0.0084	0.0074	0.0061	0.0067	0.0073
$\lambda_\phi$	0.0091	0.0051	0.0081	0.0077	0.0096
10 < $p_T$ < 15	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5
$\lambda_\theta$	0.0041	0.0051	0.0039	0.0152	0.0120
$\lambda_{\theta\phi}$	0.0101	0.0077	0.0066	0.0067	0.0070
$\lambda_\phi$	0.0078	0.0029	0.0056	0.0042	0.0062

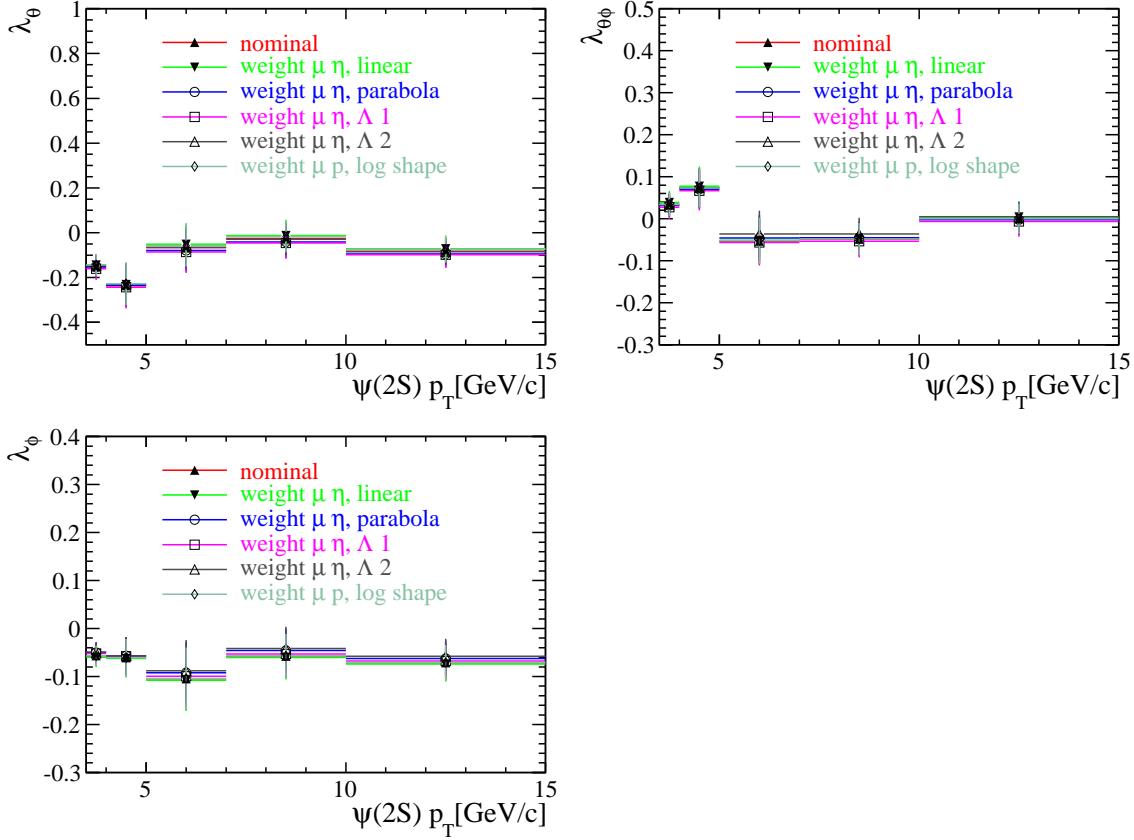


Figure 5.7 The polarization parameters  $\lambda_\theta$  (top left),  $\lambda_{\theta\phi}$  (top right) and  $\lambda_\phi$  (bottom left) determined with the weighted Monte Carlo; the baseline fit is also shown for comparison. The studies for HX frame in the rapidity bin  $3.0 < y < 3.5$  are presented for an example. Only statistical uncertainties are shown.

### 5.3.3 Monte Carlo validation

The Monte Carlo is used to estimate the efficiency in data. Any inconsistency between the efficiency in data and the efficiency in the Monte Carlo simulation will introduce a bias to the polarization measurement. In section 4.3.4, the consistency between the efficiency in Monte Carlo and data is studied using the exclusive  $B^+ \rightarrow J/\psi K^+$  sample, because the  $J/\psi$  polarization in this sample is well defined independent of the production mechanism of the spinless  $B^+$ . Thus any difference in the final state muon kinematic distribution between data and Monte Carlo is coming from the imperfectness of the simulation, assuming that the  $B^+$  kinematics is the same in data and Monte Carlo, and that Monte Carlo reconstruct/select the same  $K^+$  as data. The muon  $p_T - \eta$  (or  $P - \eta$ ) in the  $B^+ \rightarrow J/\psi K^+$  Monte Carlo is compared to data, and a ratio histogram is extracted to weight the inclusive  $J/\psi$  sample which is used to calculate the normalization in the

likelihood estimator. Because the efficiency and the ratio of the efficiency between data and Monte Carlo is expressed in muon kinematics, the ratio table can also be applied to the muons from the  $\psi(2S)$  decay. Therefore, the same method and ratio histogram in the study of  $J/\psi$  acceptance systematic uncertainties are used to extract  $\psi(2S)$  acceptance systematic uncertainties. Several binning schemes have been tried for muon  $p_T - \eta$  or muon  $P - \eta$ , and for each case a new set of polarization parameters is extracted, and the average one is compared to the nominal fits. In Figure 5.8, the difference between the averaged result and the nominal ones is shown in different  $p_T$  and rapidity bins for the HX frame. In Table 5.6, the systematic uncertainty is listed for each parameter in each kinematic bin for the HX frame. The systematic uncertainty due to acceptance, which has an average value of 0.06 for  $\lambda_\theta$  and is much smaller for  $\lambda_{\theta\phi}$  and  $\lambda_\phi$ , is at the same level as the statistical uncertainties for the bins with relatively large statistics.

### 5.3.4 Background subtraction

In the estimator, only signal events in data effectively contribute; the background is subtracted automatically because of the weights (signal *sPlot*) as a function of the mass. For the bin  $5 < p_T < 7$  GeV/ $c$ ,  $3.0 < y < 3.5$  the signal *sPlot* is shown in Figure 5.9, from which it can be seen that in the region where background events dominate the *sPlot* is less than zero, while around the  $\psi(2S)$  mass peak the *sPlot* is positive. With the signal *sPlot*, some of the background events have positive weights, others negative, and the sum of the weights for background events is zero (statistically). To use the *sPlot* unfolding method properly, the control variable ( $\cos \theta - \phi$  in this case) and the discriminative variable (invariant mass in this case) should be independent, and the distributions of the discriminative variable for various components (background and signal in this case) should be well known (or parameterized).

In the baseline analysis, two Crystal Ball functions for the signal mass and a first order polynomial for the background are used, and in the two CB functions, the  $\alpha$  is an empirical function of sigma ( $\sigma$ ), the  $\sigma$  of the first CB function is an empirical function of the  $\sigma$  of the second CB function and the fraction of the CB function with larger width is fixed to be 0.7. To study how well the mass spectrum is parameterized, other ways to describe the mass distribution have been considered:

- A) the two  $\sigma$  of the two CB functions are both free parameters;
- B) the fraction of the narrower CB function is free parameter;

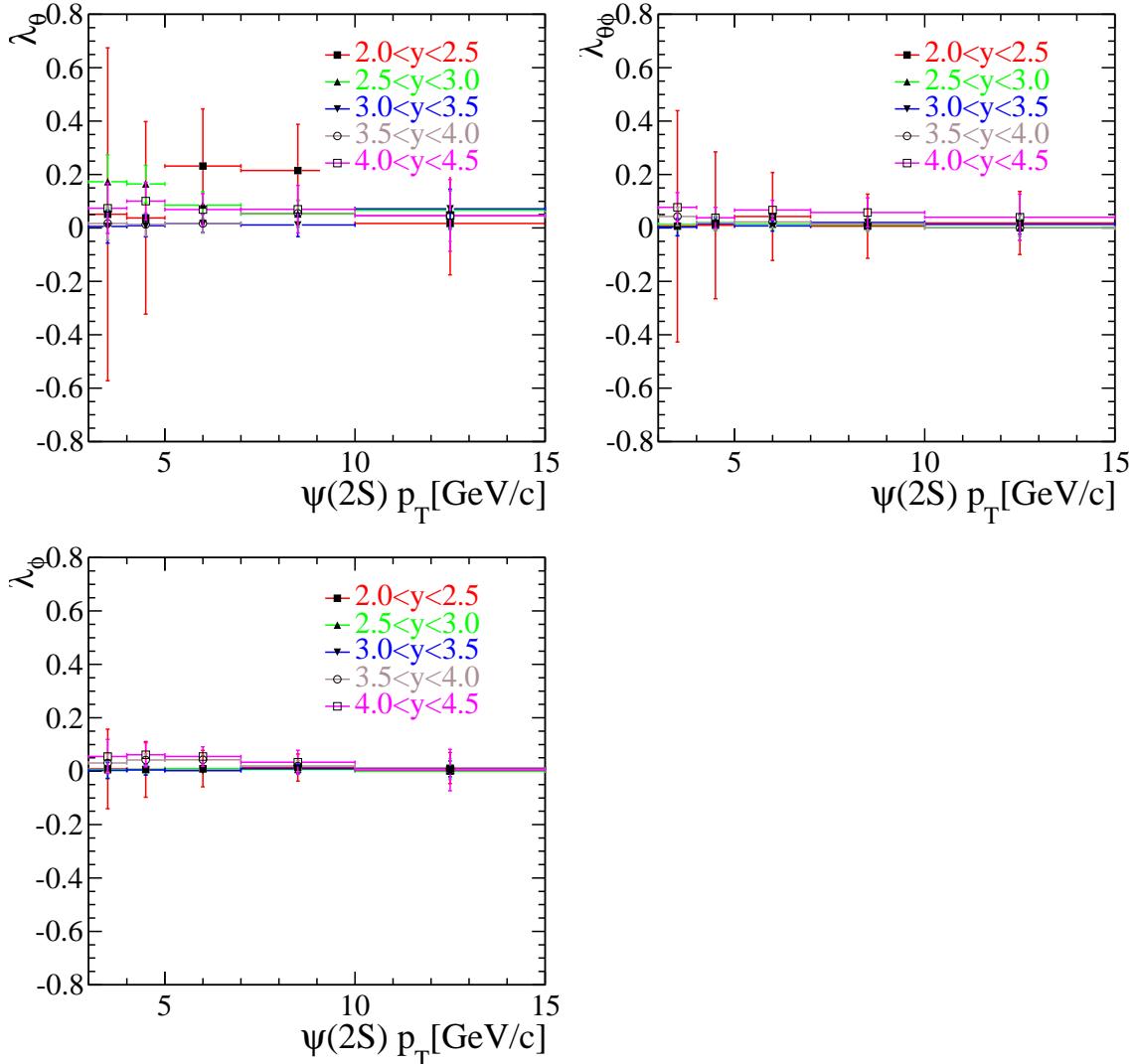


Figure 5.8 The difference between the polarization parameters  $\lambda_\theta$  (top left),  $\lambda_{\theta\phi}$  (top right) and  $\lambda_\phi$  (bottom left) extracted with the weighted Monte Carlo and the nominal ones in different  $p_T$  and rapidity bins for the HX frame. Only statistical uncertainties from the fit are shown.

Table 5.6 The acceptance systematic uncertainties for each  $p_T$  and rapidity bin of  $\psi(2S)$  in the HX frame.

$p_T$ (GeV/c)		y bin			
3.5 < $p_T$ < 4	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5
$\lambda_\theta$	0.0514	0.1729	0.0055	0.0169	0.0738
$\lambda_{\theta\phi}$	0.0063	0.0132	0.0030	0.0425	0.0770
$\lambda_\phi$	0.0081	0.0059	0.0026	0.0310	0.0548
4 < $p_T$ < 5	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5
$\lambda_\theta$	0.0376	0.1648	0.0087	0.0134	0.1002
$\lambda_{\theta\phi}$	0.0100	0.0198	0.0159	0.0188	0.0381
$\lambda_\phi$	0.0068	0.0066	0.0047	0.0421	0.0611
5 < $p_T$ < 7	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5
$\lambda_\theta$	0.2318	0.0857	0.0170	0.0170	0.0690
$\lambda_{\theta\phi}$	0.0428	0.0140	0.0072	0.0226	0.0670
$\lambda_\phi$	0.0098	0.0092	0.0020	0.0430	0.0545
7 < $p_T$ < 10	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5
$\lambda_\theta$	0.2154	0.0520	0.0109	0.0554	0.0701
$\lambda_{\theta\phi}$	0.0067	0.0122	0.0208	0.0150	0.0582
$\lambda_\phi$	0.0136	0.0062	0.0097	0.0192	0.0332
10 < $p_T$ < 15	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5
$\lambda_\theta$	0.0170	0.0657	0.0720	0.0447	0.0470
$\lambda_{\theta\phi}$	0.0190	0.0009	0.0129	0.0002	0.0404
$\lambda_\phi$	0.0114	0.0002	0.0090	0.0078	0.0040

- C) the  $\alpha$  parameters of the CB functions are free parameters;
- D) only one CB function is used to fit the signal;
- E) use an exponential function to describe the background.

For each variation of the mass spectrum description, the new polarization is calculated, and the difference between the new parameters and the nominal ones are considered as systematic uncertainties.

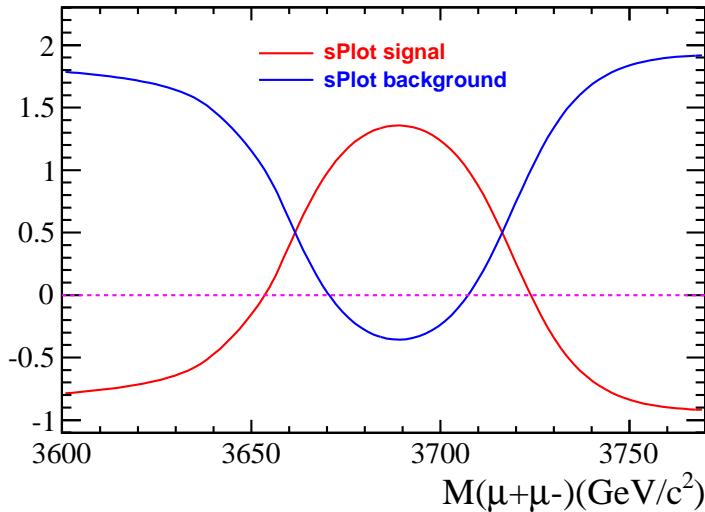


Figure 5.9 The signal  $sPlot$  (red) and background  $sPlot$  (blue) as a function of the  $\psi(2S)$  mass. The  $sPlot$  in the bin  $5 \text{ GeV}/c < p_T < 7 \text{ GeV}/c$ ,  $3.0 < y < 3.5$  is shown.

In Figure 5.10, the  $\cos\theta$  and  $\phi$  distributions for events in the mass region  $[3600, 3630] \text{ MeV}/c^2$  (left sideband) and  $[3740, 3770] \text{ MeV}/c^2$  (right sideband), and in Monte Carlo are shown for the HX frame. The fraction of signal events in the left sideband is about 5%, and 2% in the right sideband. The distributions show that the angular distributions in the two sidebands are almost the same but not identical within errors. Therefore the angular distribution (control variable) is not totally independent of the mass (discriminative variable).

To study the dependence of the mass shape as a function of the angular  $(\cos\theta - \phi)$  distribution, the signal and background mass distributions in data are extracted in different bins of  $\cos\theta - \phi$ . Four intrinsically different regions in the  $\cos\theta - \phi$  two dimensional phase space are selected taking the symmetry into account, and the signal and background mass shapes in each region are used to construct the global  $sPlot$  for another polarization fit. In all four new sets of polarization parameters are extracted, and the difference between the

new polarization parameters and the nominal ones are systematic uncertainties.

In summary, several situations have been studied to estimate systematic uncertainties due to the background subtraction, and for each bin and each polarization parameter the largest difference with respect to the nominal value among all the situations is quoted as the systematic uncertainty. In Table 5.7, the subtraction systematic uncertainties are listed for each  $p_T$  and rapidity bin of  $\psi(2S)$ .

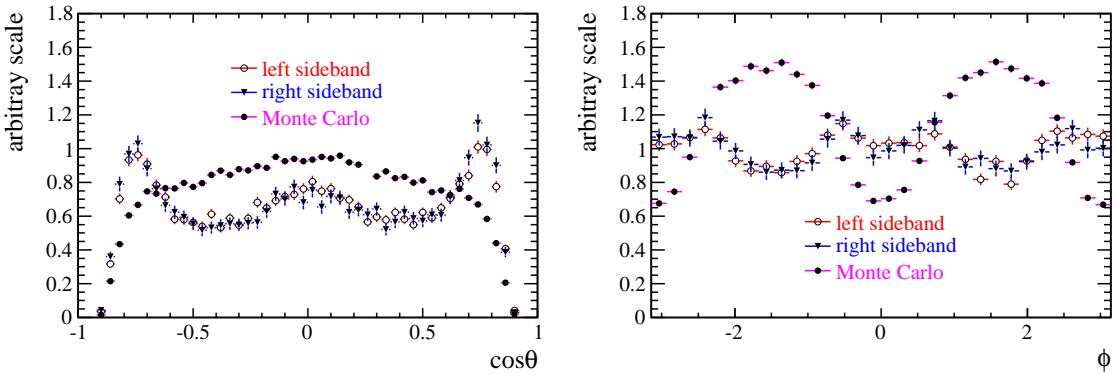


Figure 5.10 The  $\cos \theta$  (left) and  $\phi$  (right) distribution in the left sideband (open circle) and right sideband (filled triangle) of  $\psi(2S)$  mass distribution, and in the unpolarized Monte Carlo (dot).

### 5.3.5 Binning effect

In the analysis, the  $\psi(2S)$  kinematic spectrum in Monte Carlo is assumed to be the same as data in each  $p_T$  and rapidity bin. The comparison of  $\psi(2S)$   $p_T$  and rapidity between simulation and data for a typical bin is shown in Figure 5.11, from which it can be seen that the  $\psi(2S)$  kinematics are a little different. Since the efficiency as a function  $\cos \theta$  and  $\phi$  depends on the  $\psi(2S)$  kinematics, the difference of  $\psi(2S)$  kinematic spectrum can introduce systematic uncertainties to the polarization measurement. The  $\psi(2S)$   $p_T$  and rapidity distributions in Monte Carlo are weighted to the distributions in data. Because the weighting on  $p_T$  will only marginally change the rapidity distribution and vice versa, the total weight is the product of the weight for  $p_T$  and the weight for rapidity. With the weighted Monte Carlo, the polarization parameters can be extracted again, the difference between the nominal results and results from this weighted Monte Carlo is considered as the systematic uncertainties due to the kinematic spectrum difference. In Table 5.8, the systematic uncertainties in HX frame are listed.

Table 5.7 The background subtraction systematic uncertainties for each  $p_T$  and rapidity bin of  $\psi(2S)$  bin in the HX frame.

$p_T$ (GeV/c)		y bin			
3.5 < $p_T$ < 4	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5
$\lambda_\theta$	0.0513	0.0039	0.0038	0.0192	0.0370
$\lambda_{\theta\phi}$	0.0614	0.0247	0.0076	0.0094	0.0303
$\lambda_\phi$	0.0159	0.0079	0.0059	0.0056	0.0209
4 < $p_T$ < 5	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5
$\lambda_\theta$	0.0040	0.0022	0.0064	0.0116	0.0218
$\lambda_{\theta\phi}$	0.0094	0.0029	0.0021	0.0057	0.0124
$\lambda_\phi$	0.0082	0.0083	0.0043	0.0052	0.0098
5 < $p_T$ < 7	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5
$\lambda_\theta$	0.0089	0.0022	0.0056	0.0076	0.0198
$\lambda_{\theta\phi}$	0.0146	0.0018	0.0028	0.0023	0.0024
$\lambda_\phi$	0.0065	0.0023	0.0034	0.0048	0.0043
7 < $p_T$ < 10	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5
$\lambda_\theta$	0.0041	0.0052	0.0072	0.0150	0.0311
$\lambda_{\theta\phi}$	0.0059	0.0022	0.0021	0.0033	0.0073
$\lambda_\phi$	0.0015	0.0014	0.0013	0.0035	0.0030
10 < $p_T$ < 15	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5
$\lambda_\theta$	0.0130	0.0061	0.0171	0.0232	0.1379
$\lambda_{\theta\phi}$	0.0027	0.0032	0.0045	0.0036	0.0224
$\lambda_\phi$	0.0017	0.0008	0.0019	0.0012	0.0063

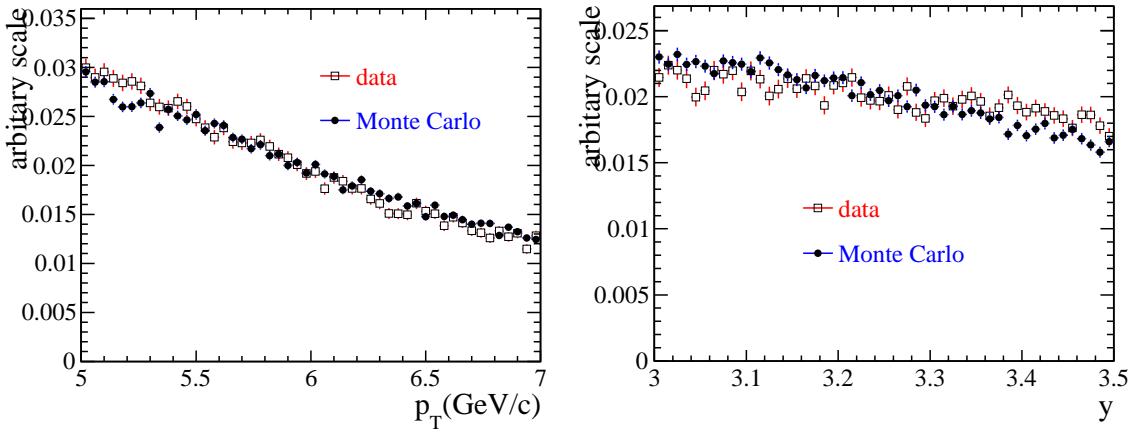


Figure 5.11 The transverse momentum (*left*) and rapidity (*right*) distributions in Monte Carlo (dot) and data (box) for the bin  $5 \text{ GeV}/c < p_T < 7 \text{ GeV}/c$ ,  $3.0 < y < 3.5$  is shown.

Table 5.8 The binning systematic uncertainties for each  $p_T$  and rapidity bin of  $\psi(2S)$ .

$p_T$ (GeV/c)			y bin			
$3.5 < p_T < 4$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$	
$\lambda_\theta$	0.1283	0.0345	0.0093	0.0052	0.0024	
$\lambda_{\theta\phi}$	0.1387	0.0116	0.0000	0.0028	0.0124	
$\lambda_\phi$	0.0526	0.0039	0.0016	0.0045	0.0078	
$4 < p_T < 5$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$	
$\lambda_\theta$	0.0059	0.0662	0.0097	0.0066	0.0141	
$\lambda_{\theta\phi}$	0.0118	0.0292	0.0007	0.0022	0.0196	
$\lambda_\phi$	0.0071	0.0053	0.0015	0.0058	0.0182	
$5 < p_T < 7$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$	
$\lambda_\theta$	0.0269	0.0403	0.0113	0.0047	0.0046	
$\lambda_{\theta\phi}$	0.0324	0.0247	0.0037	0.0021	0.0033	
$\lambda_\phi$	0.0191	0.0006	0.0052	0.0070	0.0093	
$7 < p_T < 10$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$	
$\lambda_\theta$	0.0196	0.0032	0.0009	0.0001	0.0036	
$\lambda_{\theta\phi}$	0.0539	0.0140	0.0073	0.0101	0.0049	
$\lambda_\phi$	0.0300	0.0030	0.0047	0.0067	0.0115	
$10 < p_T < 15$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$	
$\lambda_\theta$	0.0950	0.0027	0.0188	0.0050	0.0145	
$\lambda_{\theta\phi}$	0.0157	0.0103	0.0122	0.0104	0.0001	
$\lambda_\phi$	0.0056	0.0025	0.0022	0.0038	0.0138	

### 5.3.6 Residual background of $\psi(2S)$ from $b$ decay

The fraction of  $\psi(2S)$  from  $b$  hadron decay ( $\psi(2S)$  from  $b$ , detached  $\psi(2S)$ ) without any lifetime biased selection is extracted from the simultaneous fit to the pseudo-proper time  $t_z$  and  $\psi(2S)$  invariant mass distribution in data. The functions for signal (two CB functions) and background (first order polynomial) component of mass have been described previously. Concerning the proper time distribution, the signal is described by a delta function at  $t_z = 0$  for the prompt  $\psi(2S)$  and an exponential function for the detached  $\psi(2S)$ :

$$f_{\text{signal}}(t_z; n_s, f_p, \tau_b) = n_s(f_p \delta(t_z) + e^{-t_z/\tau_b}), \quad (5-6)$$

where  $n_s$  is the number of signal  $\psi(2S)$  events,  $f_p$  is the fraction of prompt  $\psi(2S)$  in the signal component and  $\tau_b$  is the pseudo lifetime of  $b$ -hadrons. The signal proper time function is convolved with a resolution function to describe the detection effect; the resolution function is combination of two Gaussian functions with the common mean:

$$f_{\text{signal resolution}}(t_z; \mu, \sigma_1, \sigma_2, f) = \frac{f}{\sqrt{2\pi}\sigma_1} e^{-\frac{(t_z-\mu)^2}{2\sigma_1^2}} + \frac{1-f}{\sqrt{2\pi}\sigma_2} e^{-\frac{(t_z-\mu)^2}{2\sigma_2^2}}. \quad (5-7)$$

The background  $t_z$  distribution can be described by the sum of a delta function, two exponentials with positive  $t_z$  and another two exponentials with negative  $t_z$

$$f_{\text{bkg}}(t_z; n_b, f_{p1}, f_{p2}, f_{m1}, f_{m2}, \tau_{p1}, \tau_{p2}, \tau_{m1}, \tau_{m2}) = n_b(\delta(t_z) + f_{p1}e^{-t_z/\tau_{p1}} + \dots). \quad (5-8)$$

The background proper time is convoluted with a Gaussian function for the resolution:

$$f_{\text{bkg resolution}}(t_z; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t_z-\mu)^2}{2\sigma^2}} \quad (5-9)$$

In Figure 5.12 (left plot), the extracted fraction of  $\psi(2S)$  from  $b$  is shown. In the simulation, the prompt  $\psi(2S)$  candidates and those from  $b$  decay can be separated from Monte Carlo truth. From these two separated samples, the retention fraction for prompt  $\psi(2S)$  ( $r_p$ ) and  $\psi(2S)$  from  $b$  ( $r_b$ ) can be calculated. Furthermore the fraction of detached  $\psi(2S)$  after the  $\tau_S$  selection is computed as:  $f_b * r_b / [(1 - f_b) * r_p + f_b * r_b]$ , where the  $f_b$  is the fraction of  $\psi(2S)$  from  $b$ . In the right plot in Figure 5.12, the fraction of  $\psi(2S)$  from  $b$  in the sample after  $\tau_S$  cut is shown.

The polarization of  $\psi(2S)$  from  $b$ -hadron decays can be different from the prompt  $\psi(2S)$  polarization, so the contamination of  $\psi(2S)$  from  $b$  in the prompt sample can bias the prompt polarization measurement. By changing the  $\tau_S$  cut value from 4 to 9, the

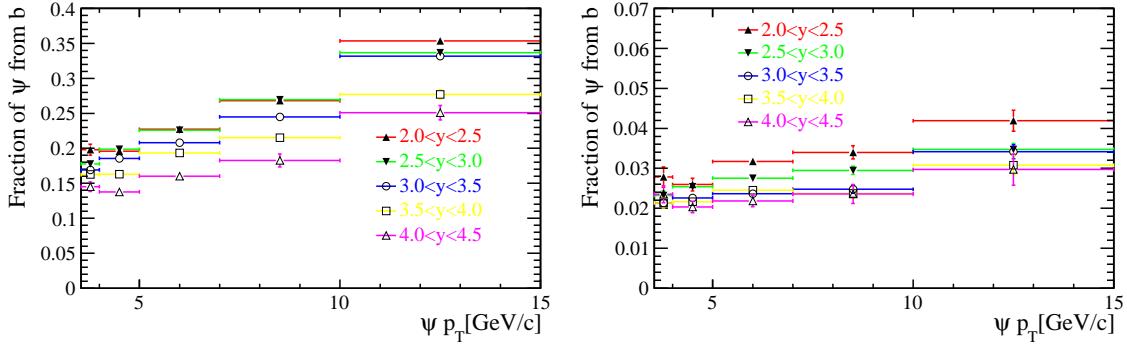


Figure 5.12 The fraction of  $\psi(2S)$  from  $b$  in different  $p_T$  and rapidity bins in data before (left) and after (right) the  $\tau_S$  cut.

fraction of  $\psi(2S)$  from  $b$  is doubled, and by studying the polarization bias due to the detached  $\psi(2S)$  with  $4 < \tau_S < 9$ , the influence of the  $\psi(2S)$  from  $b$  in the prompt sample with  $|\tau_S| < 4$  can be extrapolated. The difference between the polarization parameters determined with  $|\tau_S| < 9$  and the nominal ones are considered as systematic uncertainties coming from  $b$  decay contamination. In Table 5.9, the values for each  $p_T$  and rapidity bin are listed. The results show that in some bins, the uncertainties are especially large because by releasing the  $\tau_S$  cut, many more background events are also included, so the systematic uncertainties extracted in this way include some statistical fluctuations, i.e. it is a conservative estimate of the uncertainties. The systematic uncertainties for  $\lambda_{\theta\phi}$  and  $\lambda_\phi$  are about 0.01, while the value is 0.02 for  $\lambda_\theta$  in average.

### 5.3.7 Muon PID selection

To suppress the combinatorial background the logarithm value of the muon PID hypothesis against pion PID hypothesis for each track is required to be larger than 3. However, by comparing the muon PID distribution in simulation with data, it is found that their distributions are slightly different, hence the retention fraction of the PID selection also differs in data and in simulation. The polarization results with three different muon PID cut values (in the nominal case, the cut value is 3) are shown in Figure 5.13 for one rapidity bin; the plot shows that the difference is small, which is consistent with the fact that  $\cos \theta$  and  $\phi$  distributions are only slightly affected by the PID cut. In the toy studies, the PID selection value in Monte Carlo is changed so that the retention rate in Monte Carlo is the same as that in data for when the selection value is fixed at 3. For the low rapidity bins, the Monte Carlo PID cut value will change to be 2.7 in the low rapidity bins, and goes to 3.05 in the high rapidity bins. With these PID cuts, the polarization parameters

Table 5.9 The  $\psi(2S)$  from  $b$  contamination systematic uncertainties for each  $p_T$  and rapidity bin of  $J/\psi$  bin in the HX frame.

$p_T$ (GeV/c)		y bin			
3.5 < $p_T$ < 4	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5
$\lambda_\theta$	0.0775	0.0101	0.0240	0.0235	0.0057
$\lambda_{\theta\phi}$	0.0947	0.0096	0.0068	0.0009	0.0103
$\lambda_\phi$	0.0323	0.0013	0.0183	0.0132	0.0080
4 < $p_T$ < 5	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5
$\lambda_\theta$	0.0331	0.0348	0.0314	0.0143	0.0011
$\lambda_{\theta\phi}$	0.0915	0.0042	0.0025	0.0036	0.0051
$\lambda_\phi$	0.0435	0.0047	0.0144	0.0054	0.0075
5 < $p_T$ < 7	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5
$\lambda_\theta$	0.0108	0.0242	0.0146	0.0069	0.0382
$\lambda_{\theta\phi}$	0.0147	0.0013	0.0061	0.0062	0.0048
$\lambda_\phi$	0.0012	0.0031	0.0113	0.0086	0.0103
7 < $p_T$ < 10	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5
$\lambda_\theta$	0.0359	0.0118	0.0135	0.0179	0.0516
$\lambda_{\theta\phi}$	0.0203	0.0017	0.0003	0.0027	0.0041
$\lambda_\phi$	0.0001	0.0068	0.0062	0.0091	0.0047
10 < $p_T$ < 15	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5
$\lambda_\theta$	0.0428	0.0279	0.0072	0.0048	0.0797
$\lambda_{\theta\phi}$	0.0012	0.0014	0.0199	0.0256	0.0025
$\lambda_\phi$	0.0001	0.0083	0.0015	0.0134	0.0022

are extracted again, and the difference between those and the nominal ones are considered as systematic uncertainties, which are listed in Table 5.10. In the low rapidity and low  $p_T$  bins, the muon particle identification is not as good as in the high rapidity and high  $p_T$  bins, and the retention rate difference between Monte Carlo and data is also large, therefore the PID systematic uncertainties are large in the first rapidity bins.

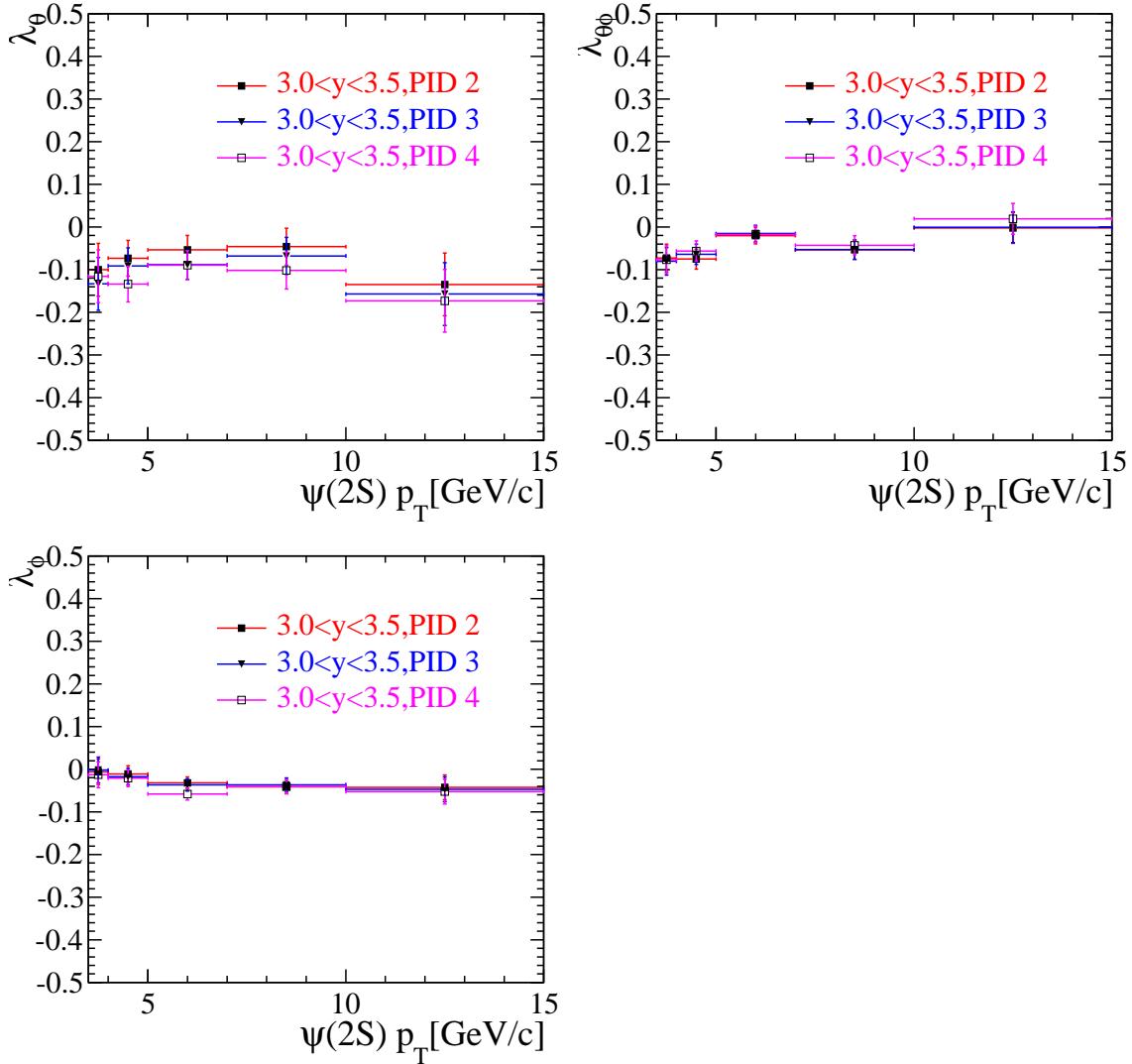


Figure 5.13 The polarization parameters  $\lambda_\theta$  (top left),  $\lambda_{\theta\phi}$  (top right) and  $\lambda_\phi$  (bottom left) extracted with three different muon PID cut values in the HX frame for the bins  $3.0 < y < 3.5$ . Only statistical errors from the fit shown.

### 5.3.8 $\tau_S$ selection

The  $\tau_S$  selection is used to select prompt  $\psi(2S)$  candidates; the  $\tau_S$  distribution is different in data and Monte Carlo, so the requirement  $|\tau_S| < 4$  can introduce bias to the polarization

Table 5.10 The statistical error coming from background subtraction for each polarization parameter in each  $\psi(2S)$  kinematic bin bin in the HX frame.

$p_T$ (GeV/c)		y bin				
3.5 < $p_T$ < 4	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5	
$\lambda_\theta$	0.0426	0.0278	0.0126	0.0055	0.0026	
$\lambda_{\theta\phi}$	0.0303	0.0086	0.0028	0.0021	0.0030	
$\lambda_\phi$	0.0009	0.0002	0.0012	0.0056	0.0009	
4 < $p_T$ < 5	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5	
$\lambda_\theta$	0.0654	0.0270	0.0123	0.0023	0.0010	
$\lambda_{\theta\phi}$	0.0399	0.0061	0.0010	0.0011	0.0041	
$\lambda_\phi$	0.0101	0.0004	0.0028	0.0046	0.0008	
5 < $p_T$ < 7	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5	
$\lambda_\theta$	0.0905	0.0262	0.0048	0.0004	0.0059	
$\lambda_{\theta\phi}$	0.0133	0.0035	0.0011	0.0022	0.0021	
$\lambda_\phi$	0.0013	0.0000	0.0012	0.0015	0.0001	
7 < $p_T$ < 10	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5	
$\lambda_\theta$	0.0409	0.0101	0.0017	0.0052	0.0070	
$\lambda_{\theta\phi}$	0.0040	0.0018	0.0005	0.0028	0.0030	
$\lambda_\phi$	0.0010	0.0026	0.0005	0.0018	0.0004	
10 < $p_T$ < 15	2.0 < $y$ < 2.5	2.5 < $y$ < 3.0	3.0 < $y$ < 3.5	3.5 < $y$ < 4.0	4.0 < $y$ < 4.5	
$\lambda_\theta$	0.0361	0.0048	0.0063	0.0026	0.0030	
$\lambda_{\theta\phi}$	0.0177	0.0023	0.0008	0.0013	0.0027	
$\lambda_\phi$	0.0042	0.0002	0.0009	0.0012	0.0013	

measurement. However, Monte Carlo truth shows that the requirement keeps more than 99% of prompt  $\psi(2S)$  events for most of the bins. A pure prompt  $\psi(2S)$  sample can be selected by using the Monte Carlo truth, and the polarization parameters extracted from this pure prompt sample is almost the same as the nominal results; the difference is less than 0.01 for  $\lambda_\theta$ , and even smaller for  $\lambda_{\theta\phi}$  and  $\lambda_\phi$  for most of the bins, quite negligible compared to the statistical uncertainties.

### 5.3.9 Detector resolution

The reconstructed kinematic variables are a little different from the simulated ones because of the detector resolution which is characterized by the resolution of  $\psi(2S)$  mass. The resolution has several effects: firstly the binning is different because the event generated in one bin can migrate to the neighboring bins; secondly the  $\cos\theta - \phi$  variables are different from the simulated one; and thirdly the events migrated from one bin to the other can carry different polarization.

Monte Carlo truth is used to calculate the angular variables  $\cos\theta - \phi$  and to define the binning. With the variables in Monte Carlo defined this way, the polarization is extracted again with the same data set. The difference between the two results is a conservative estimation of the systematic uncertainties. The studies show that the bias is negligible.

As the polarization parameters between neighboring bins is not too different (at most 0.2 for  $\cos\theta$ ), and the fraction of migrating events is very small (2%), the bias due to the third effect is smaller than 0.004 ( $0.2 \times 2\%$ ) for  $\cos\theta$  and even smaller for the other two polarization parameters. So again the resolution systematic uncertainties are negligible.

### 5.3.10 MagUp and MagDown symmetry

In Figure 5.14, the polarization parameters extracted from magnet up and magnet down data respectively are shown. The plots show that the results are consistent between the two magnet polarities.

### 5.3.11 Summary of the measurement uncertainties

The  $\psi(2S)$  polarization parameter uncertainties in the HX frame, studied in detail in the previous sections, are summarized in Tables B.15, B.16, B.17, B.18 and B.19. In Tables B.20, B.21, B.22, B.23 and B.24, the parameter uncertainties in the CS frame are summarized.

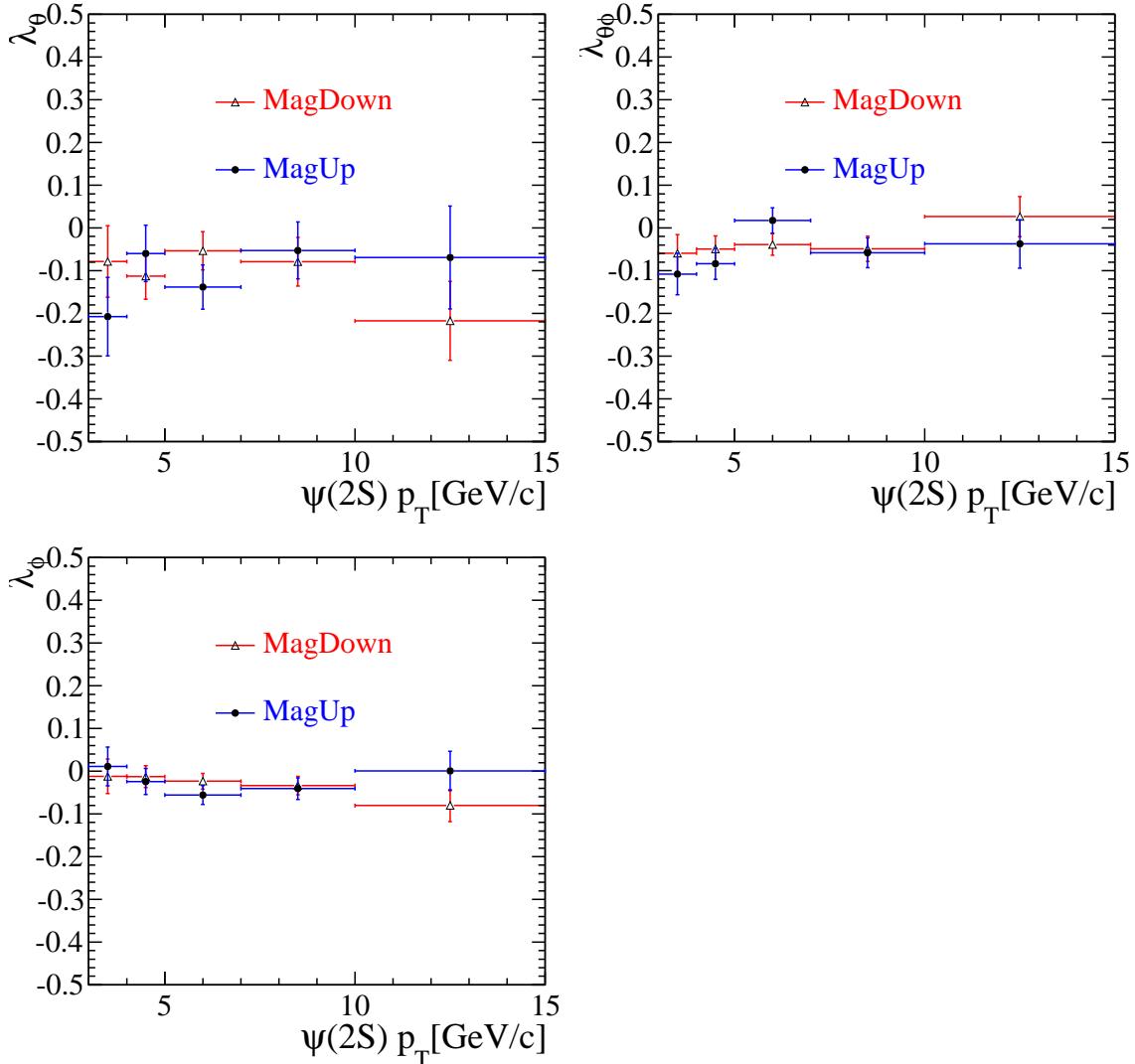


Figure 5.14 The polarization parameters  $\lambda_\theta$  (top left),  $\lambda_{\theta\phi}$  (top right) and  $\lambda_\phi$  (bottom left) extracted for the magnet polarities respectively in HX frame for the bins  $3.0 < y < 3.5$ . Only statistical errors from the fit shown.

Table 5.11 List of the main contributions to the uncertainties on the parameter  $\lambda_\theta$  in HX frame. The absolute error is reported. The minimum and maximum values are shown together with the average for each component.

Source	average	min.	max.	comment
statistical uncertainties from estimator	0.141	0.034	0.624	Bin dependent
MC efficiency (acceptance)	0.069	0.006	0.232	Correlated between bins
Binning effect	0.027	0.001	0.128	Bin dependent
Normalization from MC	0.057	0.014	0.225	Bin dependent
Sideband subtraction	0.011	0.002	0.138	Correlated between bins
$b$ -hadrons contamination	0.025	0.001	0.080	Correlated between bins
Tracking	0.011	0.002	0.029	Correlated between bins
PID cut	0.029	0.000	0.091	Correlated between bins

The statistical uncertainties — the uncertainties from the fit — depend on the number of events in each bin. The systematic uncertainties coming from the normalization of the angular distribution are related to the number of simulated events in Monte Carlo. They are generally smaller than the statistical errors from the fit, so they dominates the systematic uncertainties in low statistics bins. The systematic uncertainties due to the difference of the  $\psi(2S)$  kinematics spectrum between data and Monte Carlo, due to the PID selection and the contamination from detached  $\psi(2S)$  are of similar size, with typical values of 0.02 for  $\lambda_\theta$ , 0.01 or smaller for the other two parameters in HX frame. The largest systematic uncertainties comes from the acceptance, which has average value 0.06 for  $\lambda_\theta$  in the HX frame.

In Table 5.11 and Table 5.12, the various uncertainties including the statistical errors from the estimators are shown for  $\lambda_\theta$  in the HX frame and the CS frame. With plenty of bins, for each error, only the minimum, maximum and average values among all the  $\psi(2S)$   $p_T$  and rapidity bins are presented. Just as for the  $J/\psi$  polarization results, in the HX frame, the  $\lambda_\theta$  uncertainties are typically two to three times those for  $\lambda_{\theta\phi}$  and  $\lambda_\phi$ , while in the CS frame, the uncertainties of all three polarization parameters are found to be of similar size.

The statistical uncertainties for the  $\lambda_\theta$  parameters in both frames are very similar, especially in bins where the angular distribution in the two frames are close, as they are measures of the number of events. However, other systematic uncertainties differ clearly

Table 5.12 List of the main contributions to the errors on the parameter  $\lambda_\theta$  in CS frame. The absolute error is reported. The minimum and maximum values are shown together with the average for each component.

Source	average	min.	max.	comment
statistical uncertainties from estimator	0.151	0.024	0.878	Bin dependent
MC efficiency (acceptance)	0.027	0.022	0.134	Correlated between bins
Binning effect	0.037	0.001	0.273	Bin dependent
Normalization from MC	0.060	0.008	0.297	Bin dependent
Sideband subtraction	0.020	0.002	0.158	Correlated between bins
$b$ -hadrons contamination	0.034	0.001	0.175	Correlated between bins
Tracking	0.022	0.005	0.049	Correlated between bins
PID cut	0.010	0.000	0.044	Correlated between bins

just as expected, because  $\lambda_\theta$  in one frame is combination of the three parameters in the other frame, and the systematic factors affect the three polarization parameters differently. To compare the measurements in various frames, the frame invariant parameter  $\lambda_{\text{inv}}$ , can help us to check whether there is a significant bias to our results, and it has been verified that the difference of  $\lambda_{\text{inv}}$  between two frames is below 0.005 for most of bins, only in the first rapidity bin, the difference can be as large as 0.08 for one particular bin, but still significantly smaller than the combined uncertainties. The difference of the invariant parameter  $\lambda_{\text{inv}}$  in the HX frame and CS frame are shown in Figure 5.15 for each  $\psi(2S)$   $p_T$  and rapidity bin.

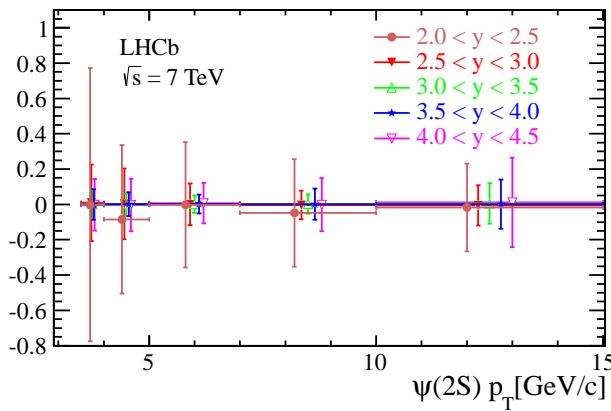


Figure 5.15 The difference of  $\lambda_{\text{inv}}$  between HX and CS frames in various  $\psi(2S)$   $p_T$  and  $y$  bins. Only statistical uncertainties are plotted.

## 5.4 Results

In Figure 5.16, Figure 5.17 and Figure 5.18, the fitted polarization parameters  $\lambda_\theta$ ,  $\lambda_{\theta\phi}$  and  $\lambda_\phi$  together with the quadratical sum of the uncertainties measured for prompt  $\psi(2S)$  as a function of the transverse momentum  $p_T$  are shown for various rapidity bins in both the HX frame and the CS frame. In Figures 5.19, 5.20 and 5.19, the  $\lambda_\theta$ ,  $\lambda_{\theta\phi}$  and  $\lambda_\phi$  integrated over the rapidity range from 2.5 to 4.0, avoiding the LHCb acceptance boundaries, are shown as a function of  $\psi(2S)$   $p_T$  respectively.

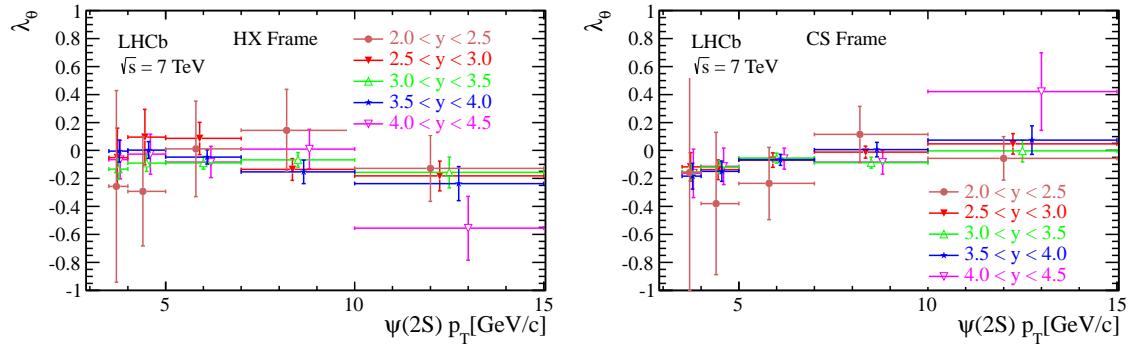


Figure 5.16 Measured  $\psi(2S)$  polarization  $\lambda_\theta$  in different  $p_T$  bins for the five rapidity bins in HX frame (*left*) and CS frame (*right*) respectively, uncertainties added quadratically.

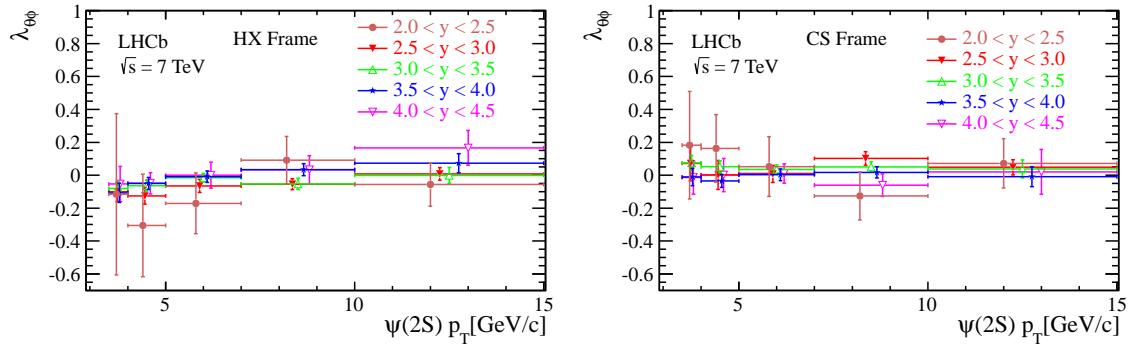


Figure 5.17 Measured  $\psi(2S)$  polarization  $\lambda_{\theta\phi}$  in different  $p_T$  bins for the five rapidity bins in HX frame (*left*) and CS frame (*right*) respectively, uncertainties added quadratically.

In Tables B.15 ··· B.19 (for HX frame) and Tables B.20 ··· B.24 (for CS frame) all the polarization parameters and the calculated uncertainties are listed.

From the results it can be concluded that in most of the kinematic region analyzed the three polarization parameters in both the HX frame and the CS frame are consistent with zero within errors, which have values that vary in the range 0.05-0.5 (0.05-0.8) for  $\lambda_\theta$ , 0.03-0.40 (0.03-0.30) for  $\lambda_{\theta\phi}$  and 0.02-0.15 (0.02-0.20) for  $\lambda_\phi$  in the HX frame (CS

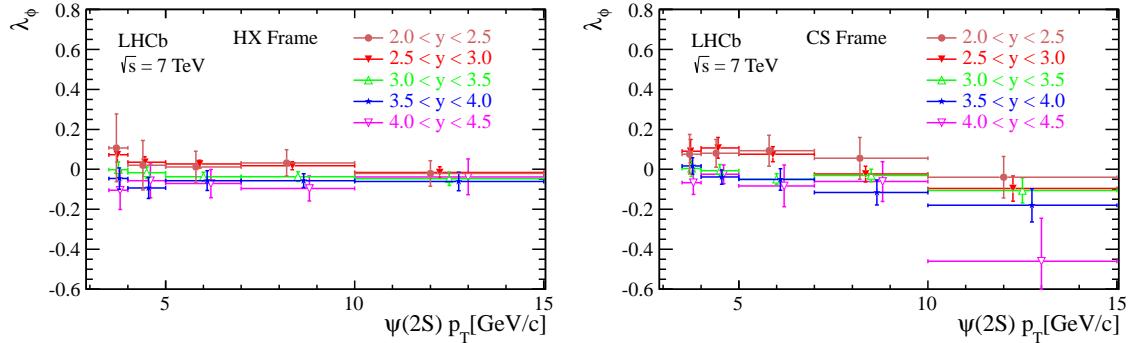


Figure 5.18 Measured  $\psi(2S)$  polarization  $\lambda_\phi$  in different  $p_T$  bins for the five rapidity bins in HX frame (left) and CS frame (right) respectively, uncertainties added quadratically.

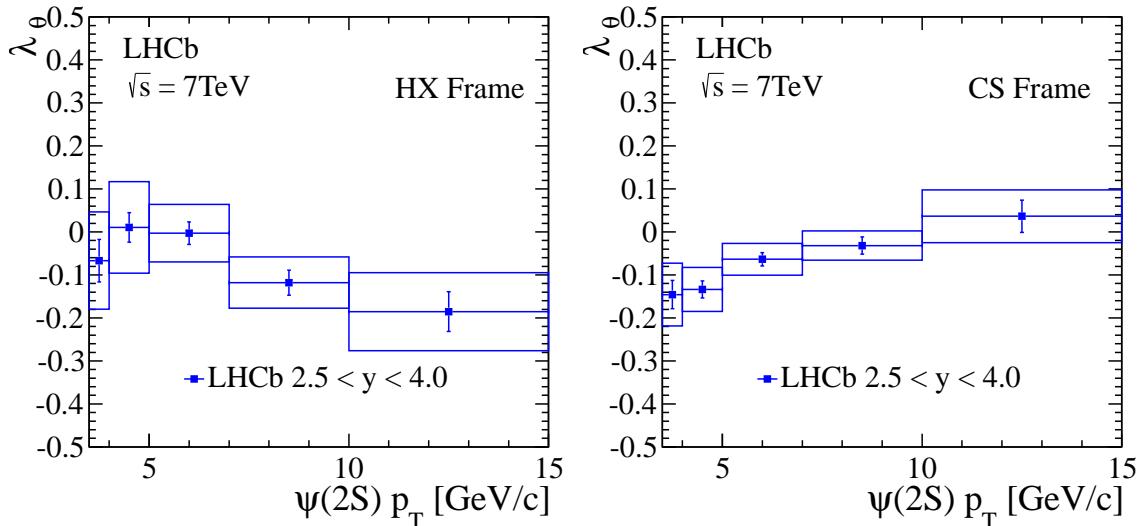


Figure 5.19 The polarization parameter  $\lambda_\theta$  integrated over the rapidity range  $2.5 < y < 4.0$  as a function of  $\psi(2S)$   $p_T$  in HX frame (left) and CS frame (right) respectively. Error bars represent the statistic uncertainties while the boxes are the statistic and systematics uncertainties added in quadrature.

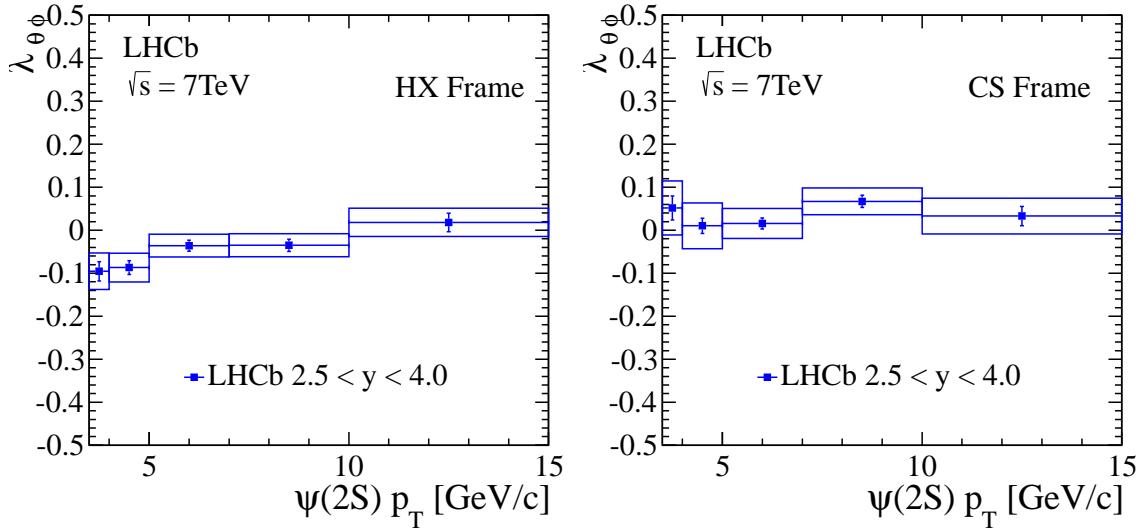


Figure 5.20 The polarization parameter  $\lambda_{\theta\phi}$  integrated over the rapidity range  $2.5 < y < 4.0$  as a function of  $\psi(2S) p_T$  in HX frame (left) and CS frame (right) respectively. Error bars represent the statistic uncertainties while the boxes are the statistic and systematics uncertainties added in quadrature.

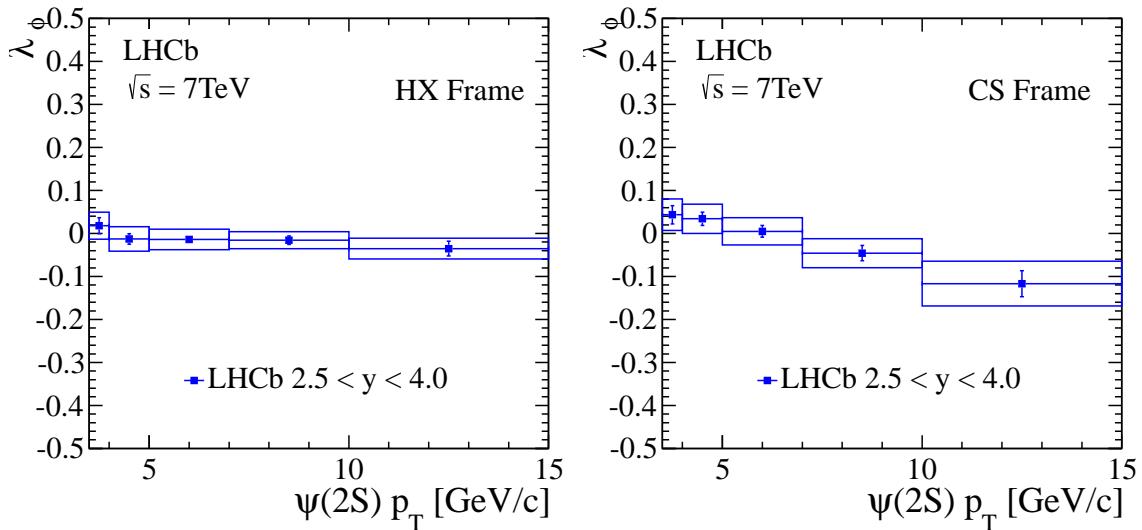


Figure 5.21 The polarization parameter  $\lambda_\phi$  integrated over the rapidity range  $2.5 < y < 4.0$  as a function of  $\psi(2S) p_T$  in HX frame (left) and CS frame (right) respectively. Error bars represent the statistic uncertainties while the boxes are the statistic and systematics uncertainties added in quadrature.

frame). In some bins the polarization parameter  $\lambda_\theta$  integrated over the rapidity range  $2.5 < y < 4.0$  in both the helicity frame and the CS frame shows slightly longitudinal polarization between -0.2 and zero.

At Tevatron, CDF<sup>[113]</sup> also measured the  $\psi(2S)$  polarization parameter  $\lambda_\theta$ , and their results are consistent with ours, although the statistical uncertainties of CDF result are quite large and their geometry coverage ( $\eta < 0.6$ ) is different from the LHCb detector.

In Figure 5.22, 5.23, and Figure 5.24, theoretical calculations<sup>[188]</sup> in the framework of Color Singlet and NRQCD are compared our measurements in HX frame. These plots for  $\lambda_\theta$  show that neither the NLO CS nor NLO NRQCD computations, which have quite large transverse (CS) or longitudinal (NRQCD) values, agree with our analysis. Concerning the  $\lambda_{\theta\phi}$  and  $\lambda_\phi$ , the NLO NRQCD calculations are consistent with our measurement.

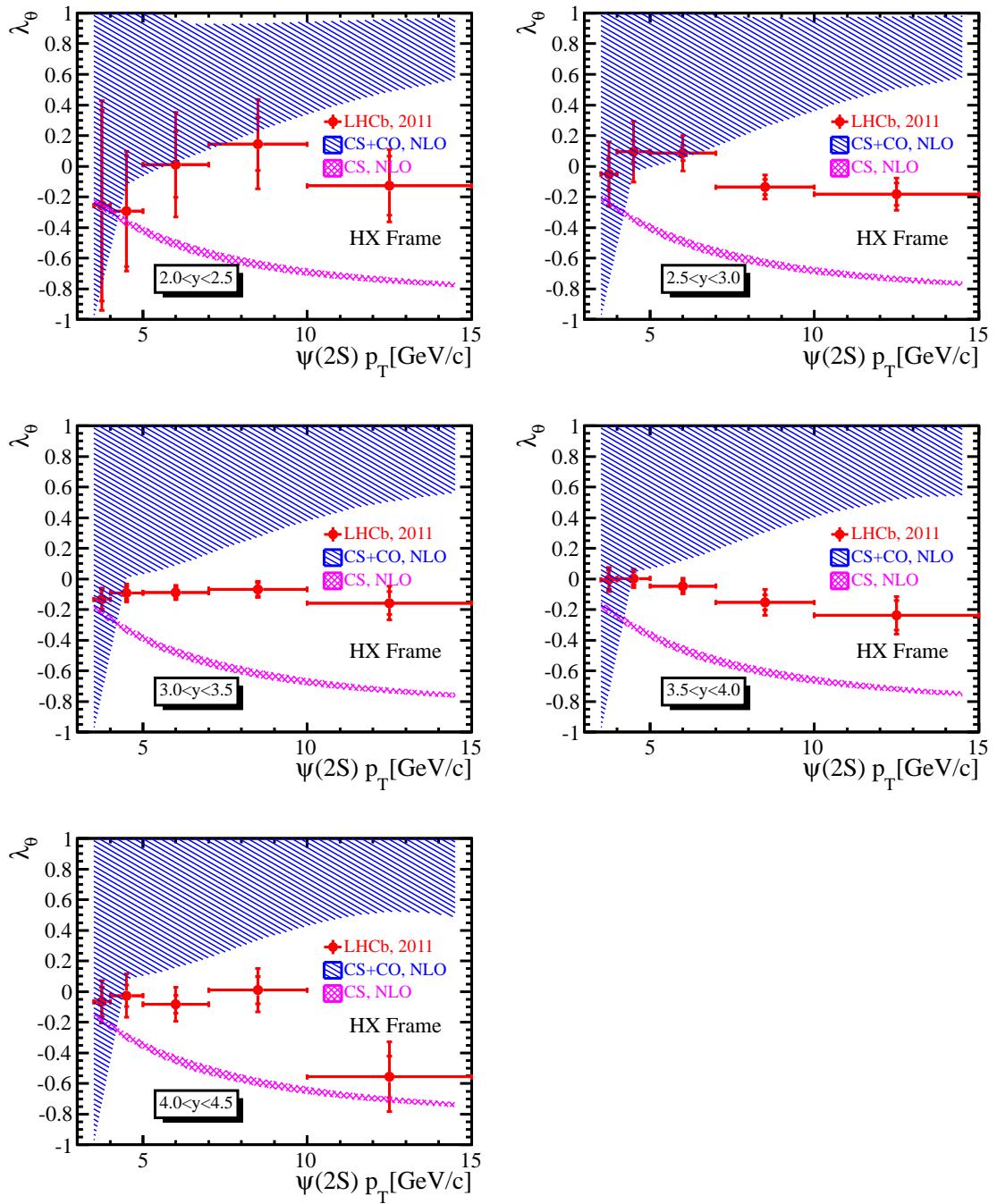


Figure 5.22 Prompt  $\psi(2S)$  polarization parameter  $\lambda_\theta$  measured at LHCb (points with error bars) compared with direct NLO color singlet (shaded purple) and NLO NRQCD (shaded blue) predictions.

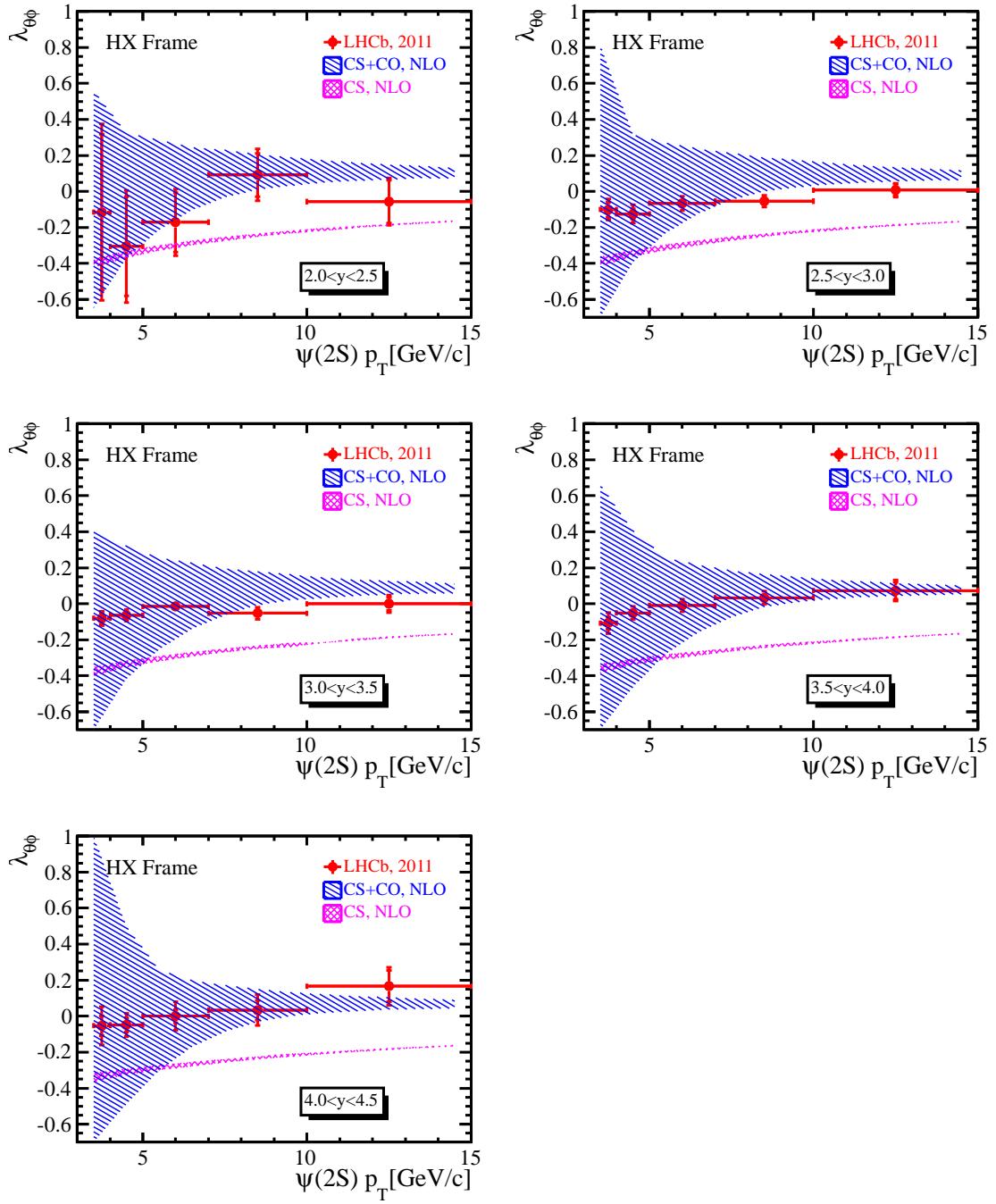


Figure 5.23 Prompt  $\psi(2S)$  polarization parameter  $\lambda_{\theta\phi}$  measured at LHCb (points with error bars) compared with direct NLO color singlet (shaded purple) and NLO NRQCD (shaded blue) predictions.

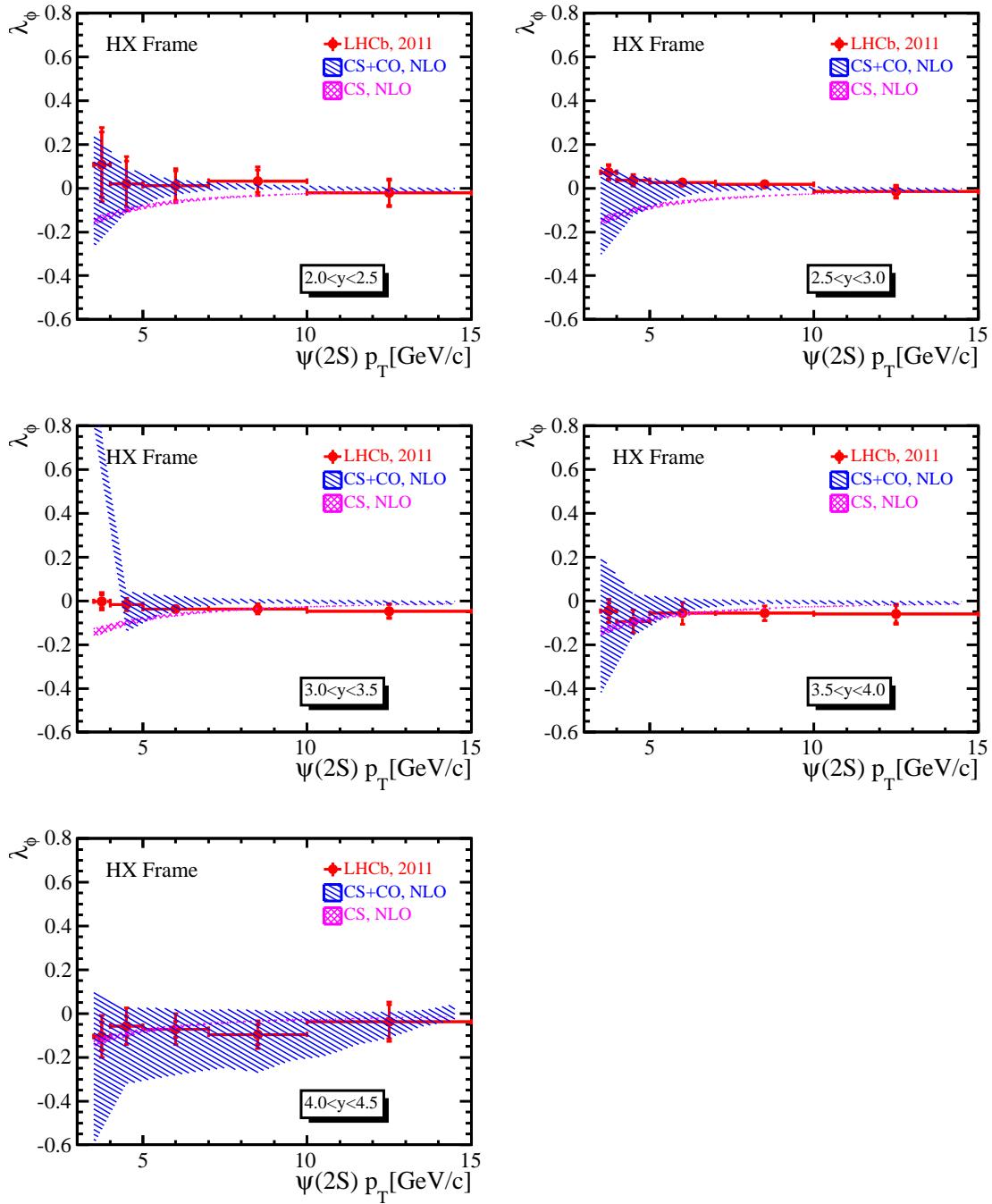


Figure 5.24 Prompt  $\psi(2S)$  polarization parameter  $\lambda_\phi$  measured at LHCb (points with error bars) compared with direct NLO color singlet (shaded purple) and NLO NRQCD (shaded blue) predictions.

## Chapter 6 Summary

The production of heavy quarkonium provides ideal place to study the mechanics of QCD. In the short distance, QCD is asymptotic, while in the long distance, QCD has the feature called color confinement. The generation of intermediate heavy quark pair  $Q\bar{Q}$  involves short distance of QCD, which is calculable perturbatively, while the transitions of  $Q\bar{Q}$  into heavy quarkonia fall into the long distance region and their rates are model dependent. There are two popular models: the CSM assumes that the  $Q\bar{Q}$  must have identical quantum numbers in order to produce the final state quarkonium, while in the NRQCD framework, which introduces color octet intermediate states,  $Q\bar{Q}$  with all kinds of quantum numbers can produce the quarkonium and the rate expanded in power of  $v$ —the velocity of heavy quarks in the rest frame of quarkonium—depending the quantum numbers of  $Q\bar{Q}$  and final state quarkonium. The leading order calculation of quarkonium production cross section in CSM fall below experimental data and the transverse momentum dependence is also wrong, however the NLO and NNLO results significantly reduce the gap between experimental data and LO computations and the transverse momentum dependence agrees with data much better. The NRQCD calculations, in which matrix elements are determined by fitting CDF data (most of the calculations) can describe the transverse momentum dependence very well, and they can also describe the LHC data. However both CSM and COM fail to describe the CDF quarkonium polarization measurements. The NLO CSM favors heavy longitudinal polarization which NRQCD gives significant transverse polarization, however at CDF they only measures slightly longitudinal polarization. The dissertation extend the polarization measurement in the new energy scale with  $pp$  collision at 7 TeV and in the forward region to test the validation between theoretical calculations.

The LHCb, dedicated for precision measurements in charm and bottom physics to probe new physics indirectly, is one of the four large experiments at CERN. The LHCb detector includes a high precision tracking system consisting of the silicon-strip VELO detector surrounding the  $pp$  interaction region, the silicon-strip TT detectors located upstream of a dipole magnet with a bending power of about 4 Tm, and three stations of silicon-strip IT detectors and straw drift-tube OT detectors placed downstream. The com-

combined tracking system provides a precise momentum and impact parameter measurement. At LHCb charged hadrons are identified using the RICH detectors, the photon, electron and hadron candidates are identified by the calorimeter system consisting of SPD, PS, ECAL and HCAL and the muons are identified by the muon system. Two-level trigger is used by the LHCb data taking, firstly candidate events are first required to pass the L0 hardware trigger exploring the large  $p_T$  or  $E_T$  of particles in interested events. The HLT trigger, composed of HLT1 and HLT2, using the full detector information to confirm the L0 triggers and to reconstruct and filter physical events respectively. In the year 2011, LHCb collected about  $1.0 \text{ fb}^{-1}$   $pp$  collision data at  $\sqrt{S} = 7 \text{ TeV}$ , and the polarization analysis is performed with this data set. By maximizing the likelihood the polarization parameters can be determined.

In this dissertation, a maximum likelihood method is introduced for the full angular analysis of di-muons in  $J/\psi$  and  $\psi(2S)$  systems in the  $\mu^+\mu^-$  channel. The maximum likelihood uses only the parameterized angular distribution of signal events, however the background events are also included in the likelihood construction. By giving a weight, which is a function of the reconstructed mass, for each event, the background event contamination is subtracted automatically. Because the angular distribution in the background does not have to be estimated, this method is easy and more reliable. The efficiency in the likelihood is estimated through unpolarized Monte Carlo, taking the advantage of the uniformity of the generated angular distribution. The method has been tested with totally transverse and longitudinal Monte Carlo events, and the input  $\pm 1$  polarization can be reproduced respectively.

Several systematic uncertainties have been investigated for the  $J/\psi$  and  $\psi(2S)$  polarization analysis. The acceptance, which comes from possible difference between efficiency in data and in the simulation is studied by the control sample  $B^+ \rightarrow J/\psi K^+$ . With average value of 0.06 for  $\lambda_\theta$  in the helicity frame, this is the dominate systematic uncertainties. The uncertainty of the normalization introduces another systematic uncertainty which is determined by the number of statistics in the simulations. The background subtraction can also introduce uncertainties if the background distribution and the mass distribution are correlated, however as almost symmetric (with regard to the  $J/\psi$  or  $\psi(2S)$  mass) mass distribution is used in the likelihood, the background events with positive weights almost have the same angular distribution as those with negative weights, and thus the background subtraction systematics uncertainties are quite small. We also

studied the systematic uncertainties associated the cuts, the binning, the pollution from quarkonium from  $b$ -hadron decay and the detection resolution, and all these uncertainties are smaller than or of the same size of the statistical uncertainties averagely.

With the maximum likelihood method, the analysis of prompt  $J/\psi$  polarization has been presented in the decay  $J/\psi \rightarrow \mu^+\mu^-$ . In the likelihood the weights have been chosen as  $+1(-1)$  when the events fall in signal region (sideband) regions, while the angular distribution is the muon polar and azimuthal angle two dimensional function with polarization parameters. The study is performed with about  $370 \text{ pb}^{-1}$   $pp$  collisions collected by the LHCb experiment at CERN in the early 2011 data taking period at a center of mass energy of 7 TeV. The  $J/\psi$  polarization parameters  $\lambda_\theta$ ,  $\lambda_{\theta\phi}$  and  $\lambda_\phi$  have been extracted in the helicity frame and the Collins-Soper frame respectively in five bins of  $J/\psi$  rapidity  $y$  and six bins of  $J/\psi$  transverse momentum  $p_T$  in the kinematic range  $2.0 < y < 4.5$  and  $2 < p_T < 15 \text{ GeV}/c$ . In each bin, the various sources of systematic uncertainties are studied in details.

The invariant polarization parameter  $\lambda_{\text{inv}}$  is almost the same in the helicity frame and the Collins-Soper frame. The results in the helicity frame for  $\lambda_\theta$ , which is about  $\simeq -0.2$  show a slightly longitudinal polarization which decreases in absolute value with the increasing  $J/\psi$   $p_T$  and rapidity, while the  $\lambda_{\theta\phi}$  and  $\lambda_\phi$  are consistent with zero. These results confirm the measurements performed by other experiments, such as CDF, PHENIX and HERA-B in such a way that no significant polarization is observed, though a direct comparison is not possible because of the different kinematic range of the experiment acceptance. A good agreement has also been found with the ALICE results in both the helicity and the Collins-Soper frame, where ALICE performed a measurement of the  $J/\psi$  polarization in a  $p_T$  and rapidity range very similar to the one explored by LHCb. Our prompt  $J/\psi$  polarization results are in contradiction to both the CSM and NRQCD NLO predictions for the direct  $J/\psi$  production, both in the size of the polarization parameters and their transverse momentum dependence. The NLO CSM predicts that at sufficient large transverse region  $J/\psi$  is strongly longitudinally polarized and  $\lambda_\theta$  has sizable values, while in the NLO NRQCD calculations the  $J/\psi$  is highly transversely polarized.

The prompt  $\psi(2S)$  polarization analysis at LHCb has been presented in the decay  $\psi(2S) \rightarrow \mu^+\mu^-$  with the weighted maximum likelihood method, in which the *sPlot* method is used to determine the weight for each event. All of the  $1.0 \text{ fb}^{-1}$  LHCb  $pp$  collision data taken in the year 2011 has been analyzed. The  $\psi(2S)$  polarization parame-

ters  $\lambda_\theta$ ,  $\lambda_{\theta\phi}$  and  $\lambda_\phi$  are extracted and their various sources of systematic uncertainties are studied in both the helicity frame and the Collins-Soper frame respectively in five bins of  $J/\psi$  rapidity  $y$  and five bins of  $J/\psi$  transverse momentum  $p_T$  in the kinematic range  $2.0 < y < 4.5$  and  $3.5 < p_T < 15$  GeV/c.

The  $\psi(2S)$  polarization results show that the three  $\lambda$  parameters are consistent with zero within errors in the HX frame for almost all of the bins, only in some bins, the  $\psi(2S)$  is slightly negatively polarized. The measurement confirms the result for  $\lambda_\theta$  at CDF; that no significantly polarized  $\psi(2S)$  is observed. The prompt  $\psi(2S)$  polarization results do not favor the computations by either NLO CSM or NLO NRQCD. Both models fail to describe the size of the  $\psi(2S)$  polarization and its  $p_T$  dependence.

Both the  $J/\psi$  and  $\psi(2S)$  polarization show that although the NRQCD, but no the lower order CSM, can describe the production cross section, it is not the final story of the quarkonium production. From the LO to NLO, the polarization of quarkonium in the framework of NRQCD almost does not change, however the polarization in CSM calculations change a lot, and production cross section changes a lot too. If this picture is correct, the NRQCD will probably fail to explain the polarization even in the higher orders, however the CSM will possibly describe both the polarization and cross section. So in the end, higher order results are of crucial importance, and we expect these calculations will come out soon.

The polarization measurements can be used to significantly reduce the systematic uncertainties in the  $J/\psi$ <sup>[147]</sup> and  $\psi(2S)$ <sup>[148]</sup> cross section measurements, in which efficiency correction strongly depends on the size of the polarization. The polarization can also change the cross section results in each kinematic bin of  $\psi$  mesons, which in turn changes the inputs for theoretical calculations. Neither the NLO CSM nor the NLO NRQCD computations can describe our data, probably higher order diagrams should be included for the calculation of quarkonium productions.

The polarization is a powerful and popular tool to test various theoretical models complementary to other variables, and sometimes the polarization provides us more information than the cross section, because the polarization is less affected by the tuning of the parameters in theory models with regard to measurements. The polarization is extremely important to discriminate models that can generate similar production cross section but different polarization. At LHC, the lead-lead and proton-lead collisions are collected. In the ion collisions, quarkonium production is an important tool to study the

high density nuclear material — the production cross section and polarization of prompt charmonium and bottomonium, and their comparison with those measured in  $pp$  collision, can be used to investigate their production processes.

## Reference

- [1] A. Einstein. *Zur Elektrodynamik bewegter Körper.* Annalen der Physik, 1905, 332(10):891–921.
- [2] M. Planck. *On the Law of Distribution of Energy in the Normal Spectrum.* Annalen der Physik, 1901, 4.
- [3] N. Bohr. *On the constitution of atoms and molecules.* The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 1913, 26.
- [4] W. Heisenberg. *Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik.* Zeitschrift für Physik A Hadrons and Nuclei, 1927, 43(3):172–198.
- [5] E. Schrödinger. *Quantisierung als eigenwertproblem.* Annalen der physik, 1926, 385(13):437–490.
- [6] P. A. Dirac. *Quantum theory of emission and absorption of radiation.* Proc.Roy.Soc.Lond., 1927, A114.
- [7] J. S. Schwinger. *On Quantum electrodynamics and the magnetic moment of the electron.* Phys.Rev., 1948, 73.
- [8] R. Feynman. *Mathematical formulation of the quantum theory of electromagnetic interaction.* Phys.Rev., 1950, 80.
- [9] R. Feynman. *Space - time approach to quantum electrodynamics.* Phys.Rev., 1949, 76.
- [10] S. Tomonaga. *On a relativistically invariant formulation of the quantum theory of wave fields.* Prog.Theor.Phys., 1946, 1.
- [11] J. Bailey et al. *Final Report on the CERN Muon Storage Ring Including the Anomalous Magnetic Moment and the Electric Dipole Moment of the Muon, and a Direct Test of Relativistic Time Dilation.* Nucl.Phys., 1979, B150.
- [12] H. W. *über den Bau der Atomkerne. I.* Zeitschrift fr Physik, 1932, 77.
- [13] E. Wigner. *On the Consequences of the Symmetry of the Nuclear Hamiltonian on the Spectroscopy of Nuclei.* Phys.Rev., 1937, 51.
- [14] M. Gell-Mann. *The Eightfold Way: A Theory of strong interaction symmetry.* CTS-20, TID-12608.
- [15] V. Barnes, P. Connolly, D. Crennell, B. Culwick, W. Delaney, et al. *Observation of a Hyperon with Strangeness -3.* Phys.Rev.Lett., 1964, 12.
- [16] F. Abe et al. (CDF Collaboration). *Observation of Top Quark Production in pp Collisions with the Collider Detector at Fermilab.* Physical Review Letters, 1995, 74(114):2626-2631.
- [17] S. Glashow. *Partial Symmetries of Weak Interactions.* Nucl.Phys., 1961, 22.
- [18] S. Weinberg. *A Model of Leptons.* Phys.Rev.Lett., 1967, 19.
- [19] A. Salam and J. C. Ward. *Electromagnetic and weak interactions.* Technical report, DTIC Document, 1964.

- [20] G. Arnison *et al.* (UA1 Collaboration). *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at  $\sqrt{s} = 540$  GeV*. Phys. Lett. B, 1983, 122(31):103-116.
- [21] M. Banner *et al.* (UA2 Collaboration). *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN  $\bar{p}p$  Collider*. Phys. Lett. B, 1983, 122(15):476-485.
- [22] G. Arnison *et al.* (UA1 Collaboration). *Experimental observation of lepton pairs of invariant mass around 95 GeV/c<sup>2</sup> at the CERN SPS collider*. Phys. Lett. B, 1983, 126(5):103-116.
- [23] M. Banner *et al.* (UA2 Collaboration). *Evidence for  $Z^0 \rightarrow e^+e^-$  At the CERN  $\bar{p}p$  Collider*. Phys. Lett. B, 1983, 129(12):130-140.
- [24] *Combination procedure for the precise determination of Z boson parameters from results of the LEP experiments*. CERN-EP-2000-153. [arXiv:hep-ex/0101027](https://arxiv.org/abs/hep-ex/0101027).
- [25] J. Bjorken. *Asymptotic Sum Rules at Infinite Momentum*. Phys.Rev., 1969, 179.
- [26] J. Bjorken and E. A. Paschos. *Inelastic Electron Proton and gamma Proton Scattering, and the Structure of the Nucleon*. Phys.Rev., 1969, 185.
- [27] R. Feynman. *Photon-hadron interactions*. ISBN-9780201360745.
- [28] J. Beringer *et al.* *Review of Particle Physics (RPP)*. Phys.Rev., 2012, D86.
- [29] J. de Groot, T. Hansl, M. Holder, J. Knobloch, J. May, et al. *Inclusive Interactions of High-Energy Neutrinos and anti-neutrinos in Iron*. Z.Phys., 1979, C1.
- [30] G. Altarelli and G. Parisi. *Asymptotic Freedom in Parton Language*. Nucl.Phys., 1977, B126.
- [31] W. Bartel *et al.* *Observation of Planar Three Jet Events in  $e^+e^-$  Annihilation and Evidence for Gluon Bremsstrahlung*. Phys.Lett., 1980, B91.
- [32] G. Aad *et al.* *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*. Phys.Lett., 2012, B716. [arXiv:1207.7214](https://arxiv.org/abs/1207.7214).
- [33] S. Chatrchyan *et al.* *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*. Phys.Lett., 2012, B716. [arXiv:1207.7235](https://arxiv.org/abs/1207.7235).
- [34] the LEP, Tevatron, SLD Electroweak Working Group. *Precision electroweak measurements and constraints on the Standard Model*. FERMILAB-TM-2480-PPD, CERN-PH-EP-2010-095, SLAC-PUB-14301.
- [35] J. Huston. *QCD: Experimental Review*. [arXiv:1110.3284](https://arxiv.org/abs/1110.3284).
- [36] L. Sawyer. *QCD at D0: A review of recent results*. AIP Conf.Proc., 2011, 1348.
- [37] H. Stenzel. *Review of QCD at LEP*. MPI-PHE-2000-28.
- [38] G. White. *Tests of perturbative and non-perturbative QCD from identified proton, kaon and pion studies in deep inelastic scattering ep interactions at HERA*.
- [39] D. Beecher. *PDF and QCD Effects in the Precision Measurement of the W Boson at CDF*. FERMILAB-THESIS-2011-18.
- [40] S. Lami. *Recent QCD results from CDF*. FERMILAB-CONF-04-261-E.
- [41] P. Demine. *Recent QCD results from the CDF and D0 experiments at the Tevatron*. Acta Phys.Polon., 2004, B35.

- [42] Michael E. Peskin. *An Introduction to Quantum Field Theory*. Stanford Linear Accelerator Center: Westview Press, 1995: 87,405,422,425,552.
- [43] A. G. Mokhtar. *Searches for exotic X, Y, and Z- states with BABAR*. SLAC-PUB-14779. **arXiv:0810.1073**.
- [44] C. Shen. *Search for the exotic states at Belle*. **arXiv:1202.6105**.
- [45] F. Nerling. *New results on the search for spin-exotic mesons with COMPASS*. volume EPS-HEP2011, 2011. 303.
- [46] C. Smith. *What could be learnt from positronium for quarkonium?* Int.J.Mod.Phys., 2004, A19. **arXiv:hep-ph/0308082**.
- [47] Y. P. Goncharov. *Quark confinement mechanism and the scale  $\Lambda_{QCD}$* . Int.J.Theor.Phys., 2012, 51. **arXiv:1201.4327**.
- [48] M. Gallinaro. *Top quark physics: From a few to a few millions*. **arXiv:1210.0786**.
- [49] N. Brambilla et al. *Heavy quarkonium physics*. FERMILAB-FN-0779, CERN-2005-005. **arXiv:hep-ph/0412158**.
- [50] G. S. Bali. *Lattice calculations of hadron properties*. Eur.Phys.J., 2004, A19. **arXiv:hep-lat/0308015**.
- [51] N. Brambilla, S. Eidelman, B. Heltsley, R. Vogt, G. Bodwin, et al. *Heavy quarkonium: progress, puzzles, and opportunities*. Eur.Phys.J., 2011, C71. **arXiv:1010.5827**.
- [52] J. Aubert et al. *Experimental Observation of a Heavy Particle* J. Phys.Rev.Lett., 1974, 33.
- [53] J. Augustin et al. *Discovery of a Narrow Resonance in  $e^+e^-$  Annihilation*. Phys.Rev.Lett., 1974, 33.
- [54] S. Herb, D. Hom, L. Lederman, J. Sens, H. Snyder, et al. *Observation of a Dimuon Resonance at 9.5 GeV in 400 GeV Proton-Nucleus Collisions*. Phys.Rev.Lett., 1977, 39.
- [55] G. Abrams, D. Briggs, W. Chinowsky, C. Friedberg, G. Goldhaber, et al. *The Discovery of a Second Narrow Resonance in  $e^+e^-$  Annihilation*. Phys.Rev.Lett., 1974, 33.
- [56] W. Braunschweig et al. *Observation of the Two Photon Cascade  $3.7\text{-GeV} \rightarrow 3.1\text{ GeV} + \text{gamma gamma via an Intermediate State } p\text{-Charm}$* . Phys.Lett., 1975, B57.
- [57] G. Abrams, A. Boyarski, M. Breidenbach, F. Bulos, W. Chinowsky, et al. *Radiative Decays of the psi-prime (3684) to New High Mass States*. SLAC-PUB-1659, LBL-4293.
- [58] W. M. Tanenbaum, J. S. Whitaker, G. Abrams, A. Boyarski, M. Breidenbach, et al. *Observation of an Intermediate State in psi-prime (3684) Radiative Cascade Decay*. Phys.Rev.Lett., 1975, 35.
- [59] R.-G. Ping. *Charmonium results from BES*. Chin.Phys., 2010, C34.
- [60] H. Vogel. *Charmonium spectroscopy and decay at CLEO-c*. Chin.Phys., 2010, C34.
- [61] A. Kuzmin. *Heavy quarkonium and quarkonium-like states at Belle and BaBar*. volume DIS2010, 2010. 118.
- [62] C. Patrignani. *E835 at FNAL: Charmonium spectroscopy in anti- $p\ p$  annihilations*. **arXiv:hep-ex/0410085**.

- [63] M. Ambrogiani, S. Argiro, S. Bagnasco, W. Baldini, D. Bettoni, et al. *Study of the gamma gamma decays of the  $\chi_{c2}(1^3P_2)$  and  $\chi_{c0}(1^3P_0)$  charmonium resonances*. Phys.Rev., 2000, D62.
- [64] D. Acosta et al. *Measurement of the  $J/\psi$  meson and  $b$ -hadron production cross sections in  $p\bar{p}$  collisions at  $\sqrt{s} = 1960$  GeV*. Phys.Rev., 2005, D71. [arXiv:hep-ex/0412071](https://arxiv.org/abs/hep-ex/0412071).
- [65] F. Abe et al. *Production of  $J/\psi$  mesons from  $\chi_c$  meson decays in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.8$  TeV*. Phys.Rev.Lett., 1997, 79.
- [66] G. T. Bodwin, E. Braaten, and G. P. Lepage. *Rigorous QCD analysis of inclusive annihilation and production of heavy quarkonium*. Phys.Rev., 1995, D51. [arXiv:hep-ph/9407339](https://arxiv.org/abs/hep-ph/9407339).
- [67] E. Eichten, K. Gottfried, T. Kinoshita, J. B. Kogut, K. Lane, et al. *The Spectrum of Charmonium*. Phys.Rev.Lett., 1975, 34.
- [68] E. Eichten, K. Gottfried, T. Kinoshita, K. Lane, and T.-M. Yan. *Charmonium: The Model*. Phys.Rev., 1978, D17.
- [69] E. Eichten, K. Gottfried, T. Kinoshita, K. Lane, and T.-M. Yan. *Charmonium: Comparison with Experiment*. Phys.Rev., 1980, D21.
- [70] E. J. Eichten and C. Quigg. *Quarkonium wave functions at the origin*. Phys.Rev., 1995, D52. [arXiv:hep-ph/9503356](https://arxiv.org/abs/hep-ph/9503356).
- [71] A. Billoire. *How Heavy Must Be Quarks in Order to Build Coulombic  $q$  anti- $q$  Bound States*. Phys.Lett., 1980, B92.
- [72] X.-H. Yuan, H.-W. Ke, Y.-B. Ding, and X.-Q. Li. *Re-study on the contribution of scalar potential and spectra of  $c\bar{c}$ ,  $b\bar{b}$  and  $b\bar{c}$  ( $\bar{b}c$ ) families*. Chin.Phys., 2012, C36. [arXiv:1012.0478](https://arxiv.org/abs/1012.0478).
- [73] X.-H. Yuan, H.-W. Ke, and X.-Q. Li. *Phenomenological study on the significance of the scalar potential and Lamb shift*. Chin.Phys., 2011, C35. [arXiv:1011.4629](https://arxiv.org/abs/1011.4629).
- [74] S. Drell and T.-M. Yan. *Partons and their Applications at High-Energies*. Annals Phys., 1971, 66.
- [75] G. A. Schuler. *Quarkonium production and decays*. Phys.Rept., 1994, . [arXiv:hep-ph/9403387](https://arxiv.org/abs/hep-ph/9403387).
- [76] M. Einhorn and S. Ellis. *Hadronic Production of the New Resonances - Are Gluons Important?* Phys.Rev.Lett., 1975, 34.
- [77] C.-N. Yang. *Selection Rules for the Dematerialization of a Particle Into Two Photons*. Phys.Rev., 1950, 77.
- [78] F. Abe et al. *Inclusive  $\chi_c$  and  $b$  quark production in  $\bar{p}p$  collisions at  $\sqrt{s} = 1.8$  TeV*. Phys.Rev.Lett., 1993, 71.
- [79] F. Abe et al. *Inclusive  $J/\psi$ ,  $\psi(2S)$  and  $b$  quark production in  $\bar{p}p$  collisions at  $\sqrt{s} = 1.8$  TeV*. Phys.Rev.Lett., 1992, 69.
- [80] D. A. Bauer, T. Graff, P. Lukens, J. Wilson, G. Alverson, et al. *DIFFERENCES BETWEEN PROTON AND PI- INDUCED PRODUCTION OF THE CHARMONIUM CHI STATES*. Phys.Rev.Lett., 1985, 54.
- [81] L. Antoniazzi et al. *Production of  $J/\psi$  via  $\psi'$  and  $\chi$  decay in 300 GeV/c proton and pi+ nucleon interactions*. Phys.Rev.Lett., 1993, 70.

- [82] E. Braaten and T. C. Yuan. *Gluon fragmentation into heavy quarkonium*. Phys.Rev.Lett., 1993, 71. [arXiv:hep-ph/9303205](https://arxiv.org/abs/hep-ph/9303205).
- [83] V. D. Barger, K.-m. Cheung, and W.-Y. Keung. *Z BOSON DECAYS TO HEAVY QUARKONIUM*. Phys.Rev., 1990, D41.
- [84] E. Braaten and T. C. Yuan. *Gluon fragmentation into P wave heavy quarkonium*. Phys.Rev., 1994, D50. [arXiv:hep-ph/9403401](https://arxiv.org/abs/hep-ph/9403401).
- [85] D. Roy and K. Sridhar. *Fragmentation contribution to quarkonium production in hadron collision*. Phys.Lett., 1994, B339. [arXiv:hep-ph/9406386](https://arxiv.org/abs/hep-ph/9406386).
- [86] J. Ma. *Gluon fragmentation into P wave triplet quarkonium*. Nucl.Phys., 1995, B447. [arXiv:hep-ph/9503346](https://arxiv.org/abs/hep-ph/9503346).
- [87] G. T. Bodwin, E. Braaten, and G. P. Lepage. *Rigorous QCD analysis of inclusive annihilation and production of heavy quarkonium*. Phys.Rev., 1995, D51. [arXiv:hep-ph/9407339](https://arxiv.org/abs/hep-ph/9407339).
- [88] E. Braaten. *Introduction to the NRQCD factorization approach to heavy quarkonium*. OHSTPY-HEP-T-97-004. [arXiv:hep-ph/9702225](https://arxiv.org/abs/hep-ph/9702225).
- [89] G. P. Lepage, L. Magnea, C. Nakhleh, U. Magnea, and K. Hornbostel. *Improved nonrelativistic QCD for heavy quark physics*. Phys.Rev., 1992, D46. [arXiv:hep-lat/9205007](https://arxiv.org/abs/hep-lat/9205007).
- [90] G. A. Schuler. *Production of heavy quarks and heavy quarkonia*. Z.Phys., 1996, C71. [arXiv:hep-ph/9504242](https://arxiv.org/abs/hep-ph/9504242).
- [91] S. Fleming and I. Maksymyk. *Hadronic psi production calculated in the NRQCD factorization formalism*. Phys.Rev., 1996, D54. [arXiv:hep-ph/9512320](https://arxiv.org/abs/hep-ph/9512320).
- [92] P. L. Cho and A. K. Leibovich. *Color octet quarkonia production. 2*. Phys.Rev., 1996, D53. [arXiv:hep-ph/9511315](https://arxiv.org/abs/hep-ph/9511315).
- [93] S. M. Tkaczyk. *Quarkonium production in  $p\bar{p}$  collisions at the Tevatron*. FERMILAB-CONF-96-425-E, C96-07-25. [arXiv:hep-ex/9611009](https://arxiv.org/abs/hep-ex/9611009).
- [94] M. Beneke and . Kramer, Michael. *Direct  $J/\psi$  and  $\psi'$  polarization and cross-sections at the Tevatron*. Phys.Rev., 1997, D55. [arXiv:hep-ph/9611218](https://arxiv.org/abs/hep-ph/9611218).
- [95] M. Cacciari, M. Greco, M. L. Mangano, and A. Petrelli. *Charmonium production at the Tevatron*. Phys.Lett., 1995, B356. [arXiv:hep-ph/9505379](https://arxiv.org/abs/hep-ph/9505379).
- [96] E. Braaten and S. Fleming. *Color octet fragmentation and the psi-prime surplus at the Tevatron*. Phys.Rev.Lett., 1995, 74. [arXiv:hep-ph/9411365](https://arxiv.org/abs/hep-ph/9411365).
- [97] P. L. Cho and A. K. Leibovich. *Color octet quarkonia production*. Phys.Rev., 1996, D53. [arXiv:hep-ph/9505329](https://arxiv.org/abs/hep-ph/9505329).
- [98] P. L. Cho. *Prompt Upsilon and Psi production at LEP*. Phys.Lett., 1996, B368. [arXiv:hep-ph/9509355](https://arxiv.org/abs/hep-ph/9509355).
- [99] S. Chekanov et al. *Measurements of inelastic  $J/\psi$  and  $\psi'$  photoproduction at HERA*. Eur.Phys.J., 2003, C27. [arXiv:hep-ex/0211011](https://arxiv.org/abs/hep-ex/0211011).
- [100] C. Adloff et al. *Inelastic leptoproduction of  $J/\psi$  mesons at HERA*. Eur.Phys.J., 2002, C25. [arXiv:hep-ex/0205065](https://arxiv.org/abs/hep-ex/0205065).

[101] C. Adloff et al. *Inelastic photoproduction of  $J/\psi$  mesons at HERA*. Eur.Phys.J., 2002, C25. [arXiv:hep-ex/0205064](https://arxiv.org/abs/hep-ex/0205064).

[102] M. Cacciari and . Kramer, Michael. *Prospects for quarkonium physics at HERA*. [arXiv:hep-ph/9609500](https://arxiv.org/abs/hep-ph/9609500).

[103] S. Gupta and K. Sridhar. *Color octet contributions to  $J/\psi$  hadroproduction at fixed target energies*. Phys.Rev., 1996, D54. [arXiv:hep-ph/9601349](https://arxiv.org/abs/hep-ph/9601349).

[104] M. Beneke and I. Rothstein. *Hadroproduction of quarkonia in fixed target experiments*. Phys.Rev., 1996, D54. [arXiv:hep-ph/9603400](https://arxiv.org/abs/hep-ph/9603400).

[105] M. B. Wise. *Heavy flavor theory: Overview*. AIP Conf.Proc., 1994, 302. [arXiv:hep-ph/9311212](https://arxiv.org/abs/hep-ph/9311212).

[106] E. Braaten and Y.-Q. Chen. *Helicity decomposition for inclusive  $J/\psi$  production*. Phys.Rev., 1996, D54. [arXiv:hep-ph/9604237](https://arxiv.org/abs/hep-ph/9604237).

[107] S. Baek, P. Ko, J. Lee, and H. Song. *Color octet mechanism and  $J/\psi$  polarization at LEP*. Phys.Rev., 1997, D55. [arXiv:hep-ph/9701208](https://arxiv.org/abs/hep-ph/9701208).

[108] M. Beneke. *Quarkonium polarization as a test of nonrelativistic effective theory*. CERN-TH-97-355. [arXiv:hep-ph/9712298](https://arxiv.org/abs/hep-ph/9712298).

[109] M. Beneke and I. Rothstein. *Psi-prime polarization as a test of color octet quarkonium production*. Phys.Lett., 1996, B372. [arXiv:hep-ph/9509375](https://arxiv.org/abs/hep-ph/9509375).

[110] M. Kramer. *Quarkonium production at high-energy colliders*. Prog.Part.Nucl.Phys., 2001, 47. [arXiv:hep-ph/0106120](https://arxiv.org/abs/hep-ph/0106120).

[111] P. L. Cho and M. B. Wise. *Spin symmetry predictions for heavy quarkonia alignment*. Phys.Lett., 1995, B346. [arXiv:hep-ph/9411303](https://arxiv.org/abs/hep-ph/9411303).

[112] E. Braaten and J. Lee. *Polarization of  $\Upsilon(nS)$  at the Tevatron*. Phys.Rev., 2001, D63. [arXiv:hep-ph/0012244](https://arxiv.org/abs/hep-ph/0012244).

[113] A. Abulencia et al. *Polarization of  $J/\psi$  and  $\psi_{2S}$  mesons produced in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96\text{-TeV}$* . Phys.Rev.Lett., 2007, 99. [arXiv:0704.0638](https://arxiv.org/abs/0704.0638).

[114] T. Aaltonen et al. *Measurements of Angular Distributions of Muons From  $\Upsilon$  Meson Decays in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.96\text{ TeV}$* . Phys.Rev.Lett., 2012, 108. [arXiv:1112.1591](https://arxiv.org/abs/1112.1591).

[115] V. Abazov et al. *Measurement of the polarization of the  $\Upsilon(1S)$  and  $\Upsilon(2S)$  states in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96\text{ TeV}$* . Phys.Rev.Lett., 2008, 101. [arXiv:0804.2799](https://arxiv.org/abs/0804.2799).

[116] S. Chekanov et al. *Measurement of  $J/\psi$  helicity distributions in inelastic photoproduction at HERA*. JHEP, 2009, 0912. [arXiv:hep-ex/0906.1424](https://arxiv.org/abs/hep-ex/0906.1424).

[117] P. Artoisenet, J. M. Campbell, F. Maltoni, and F. Tramontano.  *$J/\psi$  production at HERA*. Phys.Rev.Lett., 2009, 102. [arXiv:0901.4352](https://arxiv.org/abs/0901.4352).

[118] C.-H. Chang, R. Li, and J.-X. Wang.  *$J/\psi$  polarization in photo-production up-to the next-to-leading order of QCD*. Phys.Rev., 2009, D80. [arXiv:0901.4749](https://arxiv.org/abs/0901.4749).

[119] A. Adare et al. *Transverse momentum dependence of  $J/\psi$  polarization at midrapidity in  $p+p$  collisions at  $\sqrt{s} = 200\text{GeV}$* . Phys.Rev., 2010, D82. [arXiv:0912.2082](https://arxiv.org/abs/0912.2082).

[120] B. Trzeciak. *J/ψ polarization in p+p collisions at  $\sqrt{s} = 200$  GeV in STAR*. Acta Phys.Polon.Supp., 2012, 5. [arXiv:1202.0897](#).

[121] I. Abt et al. *Angular distributions of leptons from J/ψ's produced in 920 GeV fixed-target proton-nucleus collisions*. Eur.Phys.J., 2009, C60. [arXiv:0901.1015](#).

[122] H. S. Chung, C. Yu, S. Kim, and J. Lee. *Polarization of prompt J/ψ in proton-proton collisions at RHIC*. Phys.Rev., 2010, D81. [arXiv:0911.2113](#).

[123] J. M. Campbell, F. Maltoni, and F. Tramontano. *QCD corrections to J/ψ and Upsilon production at hadron colliders*. Phys.Rev.Lett., 2007, 98. [arXiv:hep-ph/0703113](#).

[124] P. Artoisenet, J. M. Campbell, J. Lansberg, F. Maltoni, and F. Tramontano. *Υ Production at Fermilab Tevatron and LHC Energies*. Phys.Rev.Lett., 2008, 101. [arXiv:0806.3282](#).

[125] B. Gong and J.-X. Wang. *QCD corrections to polarization of J/ψ and Υ at Tevatron and LHC*. Phys.Rev., 2008, D78. [arXiv:0805.2469](#).

[126] S. P. Baranov. *Highlights from the  $k_T$  factorization approach on the quarkonium production puzzles*. Phys.Rev., 2002, D66.

[127] Y.-Q. Ma, K. Wang, and K.-T. Chao. *QCD radiative corrections to  $\chi_{cJ}$  production at hadron colliders*. Phys.Rev., 2011, D83. [arXiv:1002.3987](#).

[128] A. Abulencia et al. *Measurement of  $\sigma_{\chi_{c2}} \mathcal{B}(\chi_{c2} \rightarrow J/\psi \gamma) / \sigma_{\chi_{c1}} \mathcal{B}(\chi_{c1} \rightarrow J/\psi \gamma)$  in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$ -TeV*. Phys.Rev.Lett., 2007, 98. [arXiv:hep-ex/0703028](#).

[129] K. Abe et al. *Observation of double  $c\bar{c}$  production in  $e^+e^-$  annihilation at  $\sqrt{s}$  approximately 10.6 GeV*. Phys.Rev.Lett., 2002, 89. [arXiv:hep-ex/0205104](#).

[130] P. Pakhlov. *Further study of double charmonium production in  $e^+e^-$  annihilation at Belle*. [arXiv:hep-ex/0412041](#).

[131] P. Pakhlov et al. *Measurement of the  $e^+e^- \rightarrow J/\psi c\bar{c}$  cross section at  $\sqrt{s} \approx 10.6$  GeV*. Phys.Rev., 2009, D79. [arXiv:0901.2775](#).

[132] B. Aubert et al. *Measurement of double charmonium production in  $e^+e^-$  annihilations at  $\sqrt{s} = 10.6$  GeV*. Phys.Rev., 2005, D72. [arXiv:hep-ex/0506062](#).

[133] G. T. Bodwin, J. Lee, and E. Braaten.  *$e^+e^-$  annihilation into  $J/\psi + J/\psi$* . Phys.Rev.Lett., 2003, 90. [arXiv:hep-ph/0212181](#).

[134] E. Braaten and J. Lee. *Exclusive double charmonium production from  $e^+e^-$  annihilation into a virtual photon*. Phys.Rev., 2003, D67. [arXiv:hep-ph/0211085](#).

[135] K.-Y. Liu, Z.-G. He, and K.-T. Chao. *Production of  $J/\psi + c\bar{c}$  through two photons in  $e^+e^-$  annihilation*. Phys.Rev., 2003, D68. [arXiv:hep-ph/0305084](#).

[136] Y.-J. Zhang and K.-T. Chao. *Double charm production  $e^+e^- \rightarrow J/\psi c\bar{c}$  at B factories with next-to-leading order QCD correction*. Phys.Rev.Lett., 2007, 98. [arXiv:hep-ph/0611086](#).

[137] Y.-J. Zhang, Y.-j. Gao, and K.-T. Chao. *Next-to-leading order QCD correction to  $e^+e^- \rightarrow J/\psi \eta_c$  at  $\sqrt{s} = 10.6$  GeV*. Phys.Rev.Lett., 2006, 96. [arXiv:hep-ph/0506076](#).

[138] K.-Y. Liu, Z.-G. He, and K.-T. Chao. *Search for excited charmonium states in  $e^+e^-$  annihilation at  $\sqrt{s} = 10.6$  GeV*. Phys.Rev., 2008, D77. [arXiv:hep-ph/0408141](#).

[139] G. T. Bodwin, D. Kang, T. Kim, J. Lee, and C. Yu. *Relativistic Corrections to  $e^+e^- \rightarrow J/\psi\eta_c$  in a Potential Model*. AIP Conf.Proc., 2007, 892. [arXiv:hep-ph/0611002](https://arxiv.org/abs/hep-ph/0611002).

[140] G. Aad et al. *Measurement of the differential cross-sections of inclusive, prompt and non-prompt  $J/\psi$  production in proton-proton collisions at  $\sqrt{s} = 7$  TeV*. Nucl.Phys., 2011, B850. [arXiv:1104.3038](https://arxiv.org/abs/1104.3038).

[141] G. Aad et al. *Measurement of the  $\Upsilon(1S)$  Production Cross-Section in  $pp$  Collisions at  $\sqrt{s} = 7$  TeV in ATLAS*. Phys.Lett., 2011, B705. [arXiv:1106.5325](https://arxiv.org/abs/1106.5325).

[142] V. Khachatryan et al. *Prompt and non-prompt  $J/\psi$  production in  $pp$  collisions at  $\sqrt(s) = 7$  TeV*. Eur.Phys.J., 2011, C71. [arXiv:1011.4193](https://arxiv.org/abs/1011.4193).

[143] S. Chatrchyan et al.  *$J/\psi$  and  $\psi_{2S}$  production in  $pp$  collisions at  $\sqrt{s} = 7$  TeV*. JHEP, 2012, 1202. [arXiv:1111.1557](https://arxiv.org/abs/1111.1557).

[144] V. Khachatryan et al. *Measurement of the Inclusive Upsilon production cross section in  $pp$  collisions at  $\sqrt(s)=7$  TeV*. Phys.Rev., 2011, D83. [arXiv:1012.5545](https://arxiv.org/abs/1012.5545).

[145] K. Aamodt et al. *Rapidity and transverse momentum dependence of inclusive  $J/\psi$  production in  $pp$  collisions at  $\sqrt{s} = 7$  TeV*. Phys.Lett., 2011, B704. [arXiv:1105.0380](https://arxiv.org/abs/1105.0380).

[146] B. Abelev et al. *Inclusive  $J/\psi$  production in  $pp$  collisions at  $\sqrt{s} = 2.76$  TeV*. CERN-PH-EP-2012-055. [arXiv:1203.3641](https://arxiv.org/abs/1203.3641).

[147] R. Aaij et al. *Measurement of  $J/\psi$  production in  $pp$  collisions at  $\sqrt{s} = 7$  TeV*. Eur.Phys.J., 2011, C71. [arXiv:1103.0423](https://arxiv.org/abs/1103.0423).

[148] R. Aaij et al. *Measurement of  $\psi(2S)$  meson production in  $pp$  collisions at  $\sqrt(s)=7$  TeV*. Eur.Phys.J., 2012, C72. [arXiv:1204.1258](https://arxiv.org/abs/1204.1258).

[149] R. Aaij et al. *Measurement of Upsilon production in  $pp$  collisions at  $\sqrt{s} = 7$  TeV*. Eur.Phys.J., 2012, C72. [arXiv:1202.6579](https://arxiv.org/abs/1202.6579).

[150] R. Aaij et al. *Observation of  $J/\psi$  pair production in  $pp$  collisions at  $\sqrt{s} = 7$  TeV*. Phys.Lett., 2012, B707. [arXiv:1109.0963](https://arxiv.org/abs/1109.0963).

[151] C.-F. Qiao, L.-P. Sun, and P. Sun. *Testing Charmonium Production Mechanism via Polarized  $J/\psi$  Pair Production at the LHC*. J.Phys., 2010, G37. [arXiv:0903.0954](https://arxiv.org/abs/0903.0954).

[152] A. Berezhnoy, A. Likhoded, A. Luchinsky, and A. Novoselov. *Double  $J/\psi$  meson Production at LHC and 4c-tetraquark state*. Phys.Rev., 2011, D84. [arXiv:1101.5881](https://arxiv.org/abs/1101.5881).

[153] R. Aaij et al. *Measurement of the cross-section ratio  $\sigma(\chi_{c2})/\sigma(\chi_{c1})$  for prompt  $\chi_c$  production at  $\sqrt{s} = 7$  TeV*. Phys.Lett., 2012, B714. [arXiv:1202.1080](https://arxiv.org/abs/1202.1080).

[154] R. Aaij et al. *Measurement of the ratio of prompt  $\chi_c$  to  $J/\psi$  production in  $pp$  collisions at  $\sqrt{s} = 7$  TeV*. Phys.Lett., 2012, B718. [arXiv:1204.1462](https://arxiv.org/abs/1204.1462).

[155] B. Abelev et al.  *$J/\psi$  polarization in  $pp$  collisions at  $\sqrt(s)=7$  TeV*. Phys.Rev.Lett., 2012, 108. [arXiv:1111.1630](https://arxiv.org/abs/1111.1630).

[156] S. Chatrchyan et al. *Measurement of the  $\Upsilon(1S)$ ,  $\Upsilon(2S)$  and  $\Upsilon(3S)$  polarizations in  $pp$  collisions at  $\sqrt{s} = 7$  TeV*. CMS-BPH-11-023, CERN-PH-EP-2012-205. [arXiv:1209.2922](https://arxiv.org/abs/1209.2922).

[157] E. Bruning, Oliver S., E. Collier, P., E. Lebrun, P., E. Myers, S., E. Ostojic, R., et al. *LHC Design Report. 1. The LHC Main Ring*. CERN-2004-003-V-1, CERN-2004-003.

- [158] G. Aad et al. *The ATLAS Experiment at the CERN Large Hadron Collider*. JINST, 2008, 3.
- [159] S. Chatrchyan et al. *The CMS experiment at the CERN LHC*. JINST, 2008, 3.
- [160] K. Aamodt et al. *The ALICE experiment at the CERN LHC*. JINST, 2008, 3.
- [161] J. Alves, A. Augusto et al. *The LHCb Detector at the LHC*. JINST, 2008, 3.
- [162] G. Anelli et al. *The TOTEM experiment at the CERN Large Hadron Collider*. JINST, 2008, 3.
- [163] O. Adriani et al. *The LHCf detector at the CERN Large Hadron Collider*. JINST, 2008, 3.
- [164] the LHCb Collaboration. *LHCb VELO TDR: Vertex locator. Technical design report*. CERN-LHCC-2001-011.
- [165] the LHCb Collaboration. *LHCb: RICH technical design report*. CERN-LHCC-2000-037.
- [166] the LHCb Collaboration. *LHCb: Inner tracker technical design report*. CERN-LHCC-2002-029.
- [167] the LHCb Collaboration. *LHCb: Outer tracker technical design report*. CERN-LHCC-2001-024.
- [168] the LHCb Collaboration. *LHCb magnet: Technical design report*. CERN-LHCC-2000-007.
- [169] the LHCb Collaboration. *LHCb calorimeters: Technical design report*. CERN-LHCC-2000-036.
- [170] the LHCb Collaboration. *LHCb muon system technical design report*. CERN-LHCC-2001-010.
- [171] P. R. Perez. *The LHCb Vertex Locator performance and Vertex Locator upgrade*. **arXiv:1209.4845**.
- [172] P. Bjornstad. *Performance of the LHCb Vertex Locator*. JINST, 2011, 6.
- [173] R. Frühwirth. *Application of Kalman filtering to track and vertex fitting*. Nucl. Instrum. Meth. 1987, A262.
- [174] I. Machikhiliyan. *Current status and performance of the LHCb electromagnetic and hadron calorimeters*. J.Phys.Conf.Ser., 2011, 293.
- [175] R. Aaij et al. *Evidence for the decay  $B^0 \rightarrow J/\psi\omega$  and measurement of the relative branching fractions of  $B_s^0$  meson decays to  $J/\psi\eta$  and  $J/\psi\eta'$* . CERN-PH-EP2012-287, LHCb-PAPER-2012-022. **arXiv:1210.2631**.
- [176] A. Martin Sanchez, P. Robbe, and M.-H. Schune. *Performances of the LHCb L0 Calorimeter Trigger*. LHCb-PUB-2011-026, CERN-LHCb-PUB-2011-026.
- [177] N. Brook. *LHCb Computing Model*. LHCb-2004-119, CERN-LHCb-2004-119.
- [178] T. Sjostrand, L. Lonnblad, S. Mrenna, and P. Z. Skands. *Pythia 6.3 physics and manual*. FERMILAB-PUB-03-457, LU-TP-03-38. **arXiv:hep-ph/0308153**.
- [179] D. Lange. *The EvtGen particle decay simulation package*. Nucl.Instrum.Meth., 2001, A462.
- [180] S. Agostinelli et al. *GEANT4: A Simulation toolkit*. Nucl.Instrum.Meth., 2003, A506.
- [181] K. Blum. *Density Matrix Theory and Applications*. Springer, 2012: 32.
- [182] M. Beneke, M. Kramer, and M. Vantinen. *Inelastic photoproduction of polarized  $J/\psi$* . Phys.Rev., 1998, D57. **arXiv:hep-ph/9709376**.

[183] P. Faccioli, C. Lourenco, J. Seixas, and H. K. Wohri. *Towards the experimental clarification of quarkonium polarization*. Eur.Phys.J., 2010, C69. [arXiv:1006.2738](#).

[184] K. Gottfried and J. D. Jackson. *On the Connection between production mechanism and decay of resonances at high-energies*. Nuovo Cim., 1964, 33.

[185] J. C. Collins and D. E. Soper. *Angular Distribution of Dileptons in High-Energy Hadron Collisions*. Phys.Rev., 1977, D16.

[186] R. Brun and F. Rademakers. *ROOT: An object oriented data analysis framework*. Nucl.Instrum.Meth., 1997, A389.

[187] G. Cowan. *Statistical Data Analysis*. University of London: Oxford University Press, 1998.

[188] M. Butenschoen and B. A. Kniehl.  *$J/\psi$  production in NRQCD: A global analysis of yield and polarization*. Nucl.Phys.Proc.Suppl., 2012, 222-224. [arXiv:1201.3862](#).

[189] Y. Xie. *sFit: a method for background subtraction in maximum likelihood fit*. [arXiv:0905.0724](#).

[190] M. Pivk and F. R. Le Diberder. *SPlot: A Statistical tool to unfold data distributions*. Nucl.Instrum.Meth., 2005, A555. [arXiv:physics/0402083](#).

[191] Y. Xie. *sFit: concept, implementation and possible pitfalls*. , 2013, .

[192] M. Derrick et al. *Measurement of the  $F_2$  structure function in deep inelastic  $e + p$  scattering using 1994 data from the ZEUS detector at HERA*. Z.Phys., 1996, C72. [arXiv:hep-ex/9607002](#).

[193] S. Piec. *Measurement of the proton structure function  $FL(x, Q^2)$  with the H1 detector at HERA*. DESY-THESIS-2010-051, 2010, .

[194] B. T. Fleming et al. *A First measurement of low  $x$  low  $Q^{**2}$  structure functions in neutrino scattering*. Phys.Rev.Lett., 2001, 86. [arXiv:hep-ex/0011094](#).

[195] H. Abramowicz, F. Dydak, J. de Groot, J. Knobloch, J. May, et al. *Determination of the Gluon Distribution in the Nucleon from Deep Inelastic Neutrino Scattering*. Z.Phys., 1982, C12.

[196] F. Bergsma et al. *Experimental Study of the Nucleon Structure Functions and of the Gluon Distribution from Charged Current Neutrino and anti-neutrinos Interactions*. Phys.Lett., 1983, B123.

[197] C. Ward. *Precision tests of the Standard Model at LEP*.

[198] R. Van Royen and V. Weisskopf. *Hadron Decay Processes and the Quark Model*. Nuovo Cim., 1967, A50.

[199] B. Gong, X. Q. Li, and J.-X. Wang. *QCD corrections to  $J/\psi$  production via color octet states at Tevatron and LHC*. Phys.Lett., 2009, B673. [arXiv:0805.4751](#).

[200] E. Aslanides, J. Cachemiche, J. Cogan, B. Dinkespiler, S. Favard, et al. *The Level-0 muon trigger for the LHCb experiment*. Nucl.Instrum.Meth., 2007, A579. [arXiv:physics.ins-det/0705.0310](#).

## **Acknowledgement**

首先感谢导师高原宁教授和杨振伟副教授对本人的精心指导，无论是在工作之内还是工作外，他们的言语举止我无一敢忘，他们的言传身教将使我终生受益，没有他们，这个工作不能如期完成。

然后感谢我的父母和兄弟姐妹，虽然平时交流不多，但是他们的关心感觉尽在咫尺。

同时感谢我的妻子，她把生活琐事繁物整理的井井有条，使我能够把精力投入到工作中去。

本课题承蒙国家自然科学基金资助，特此致谢。

## 声 明

本人郑重声明：所呈交的学位论文，是本人在导师指导下，独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本学位论文的研究成果不包含任何他人享有著作权的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。

签 名： \_\_\_\_\_ 日 期： \_\_\_\_\_

## Appendix A Crosscheck studies

### A.1 Efficiency

In Figures A.1,A.2, A.3 and A.4 the one dimensional distributions of  $\cos \theta$  and  $\phi$  for selected  $J/\psi$  mesons in Monte Carlo in the HX frame and the CS frame are shown respectively. The efficiency distribution as a function of muon angular variables is similar for  $\psi(2S)$  and  $J/\psi$  events. Because the LHCb detector only records  $J/\psi$  in the forward region, the shape of the efficiency in the HX frame and the CS frame is quite similar, especially in the low transverse momentum  $J/\psi$  bins, where the angle between the two quantization  $z$  axis is quite small. As the selections and triggers for  $J/\psi$  and  $\psi(2S)$  are similar, and the two mesons have similar mass, it is expected that the distribution of the efficiency will be similar for the two mesons in the di-muon final states. At generation level the distribution is uniform due to the zero input polarization, so these plots reflect the dependence of the detection efficiency as a function of the muon polar angles  $\Omega(\cos \theta, \phi)$  for different  $J/\psi$  rapidity and transverse momentum bins. In the following, explanations to the generation of the non-uniformity of the shape of the efficiency distribution (in the HX frame) from the geometrical acceptance, the selections and trigger requirements are given.

Three  $J/\psi$  kinematic bins are used to demonstrate various factors that can lead to the final distribution, a low  $p_T$  and low rapidity bin, a middle  $p_T$  and middle rapidity bin and a large  $p_T$  and large rapidity bin. Figure 4.6 shows the scratch view of the HX frame at LHCb with the production plane parallel to the screen.

The LHCb geometrical acceptance [10, 400] mrad clearly has the following two influences on the  $J/\psi$  events:

- Because  $\theta$  is the angle between  $\mu^+$  in  $J/\psi$  rest frame and the  $J/\psi$  momentum in laboratory frame,  $\cos \theta \approx +1(-1)$  means that the  $\mu^-$  ( $\mu^+$ ) will fly almost in the opposite direction of  $J/\psi$  in the laboratory frame. If the  $J/\psi$  momentum is not large enough, the muon will have negative momentum in  $z$  direction or relatively smaller momentum. As the LHCb detector covers only the forward  $z > 0$  region, such muons have less probability to fly into the LHCb region.
- By definition, as the  $x$ -axis points outside of LHCb acceptance,  $\phi \approx 0$  or  $\pi$  (along

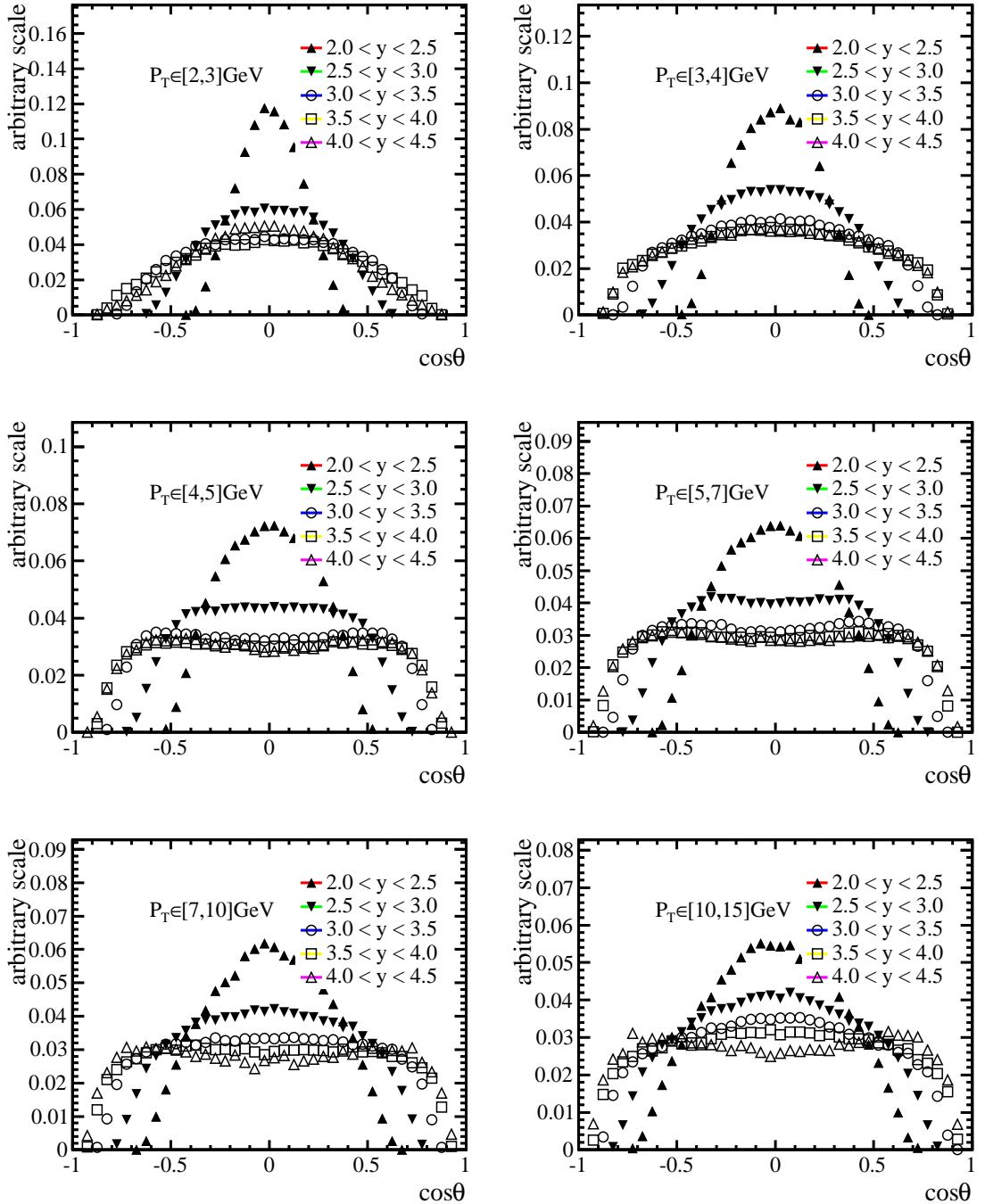


Figure A.1 Efficiency ( $\epsilon_{\text{tot}}$ ) as a function of  $\cos \theta$  in each  $p_T$  and rapidity of  $J/\psi$  is shown in the HX frame. Each curve is normalized to unity.

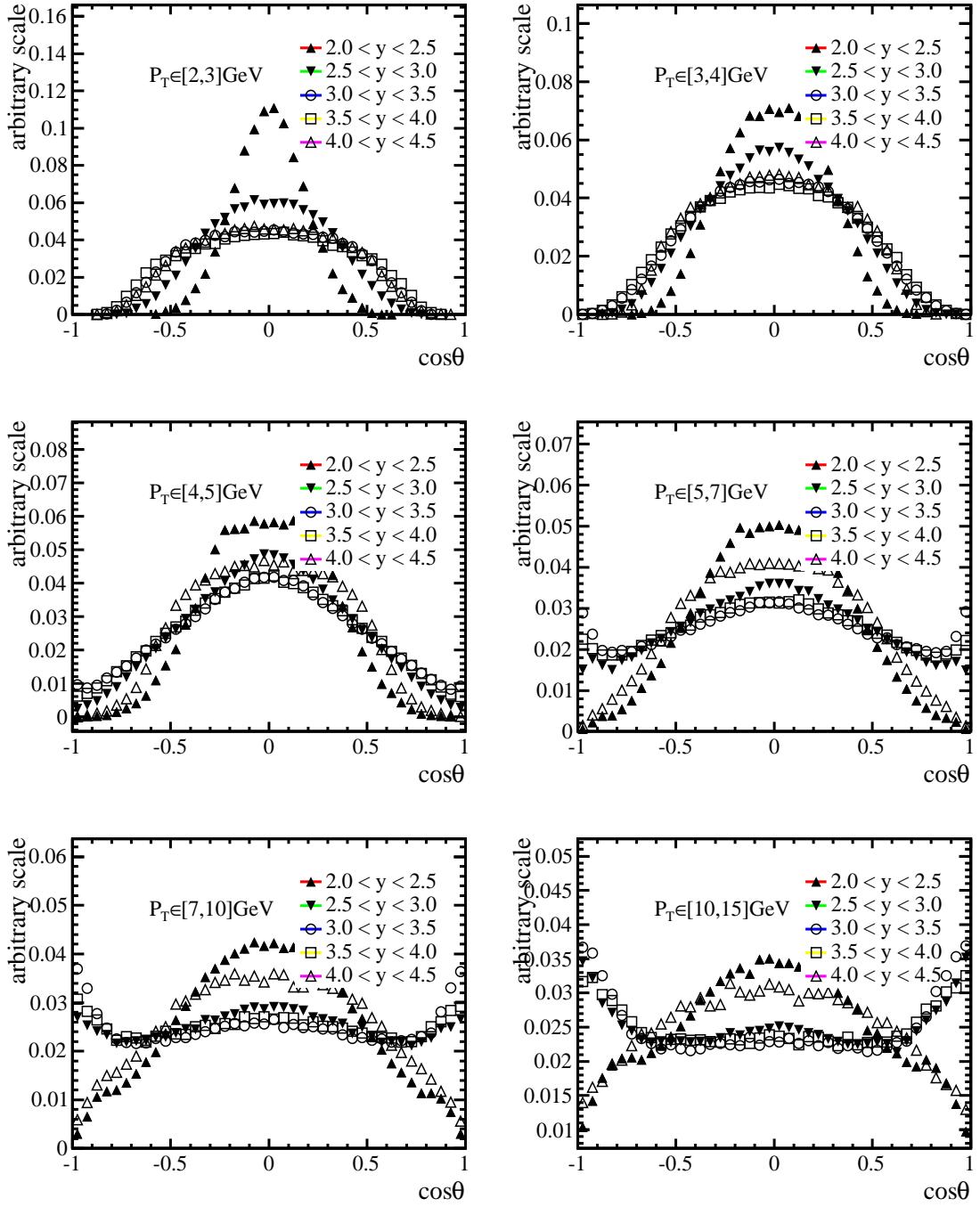


Figure A.2 Efficiency ( $\epsilon_{\text{tot}}$ ) as a function of  $\cos\theta$  in each  $p_T$  and rapidity of  $J/\psi$  is shown in the CS frame. Each curve is normalized to unity.

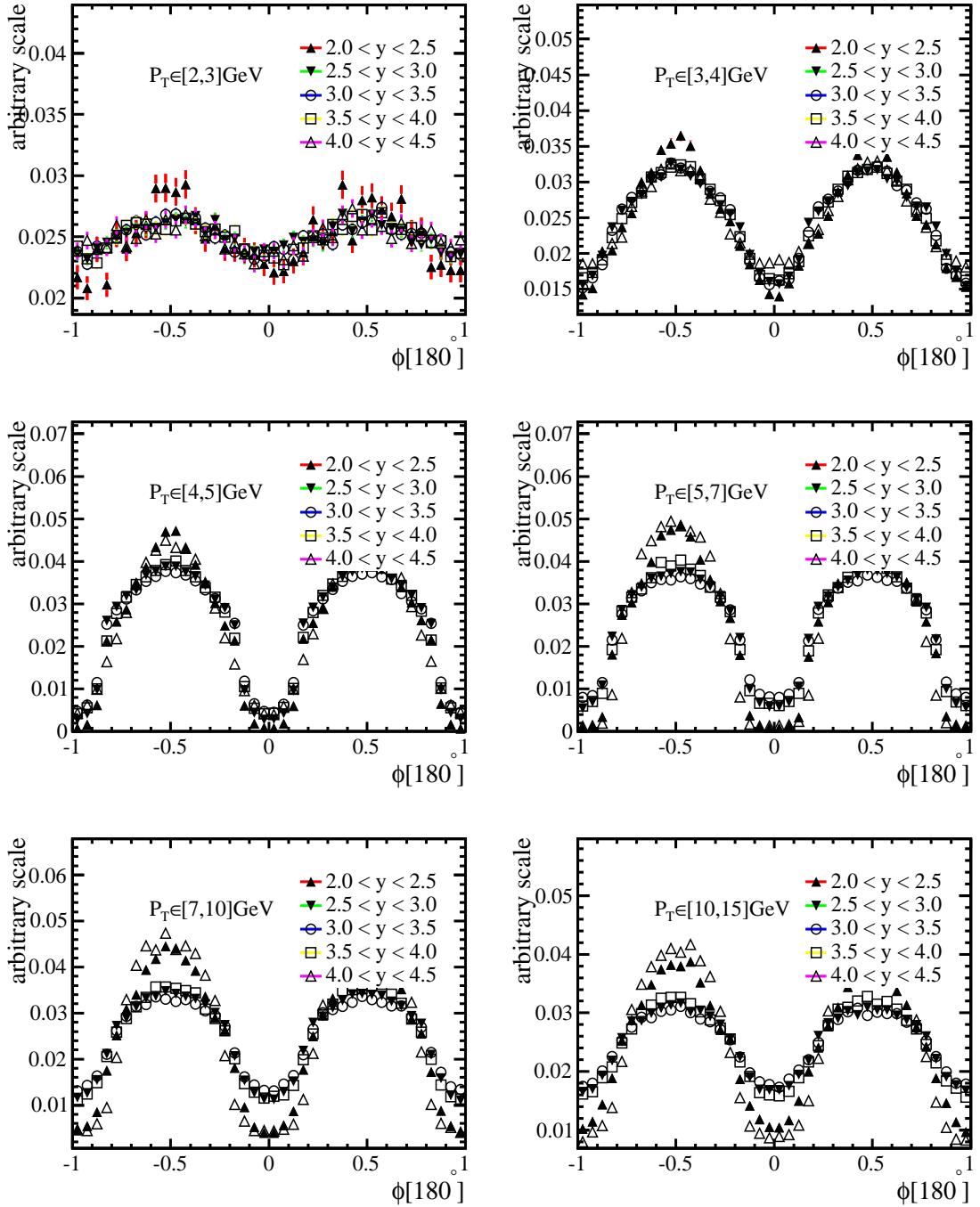


Figure A.3 Efficiency ( $\epsilon_{\text{tot}}$ ) as a function of  $\phi$  in each  $p_T$  and rapidity of  $J/\psi$  is shown in the HX frame. Each curve is normalized to unity.

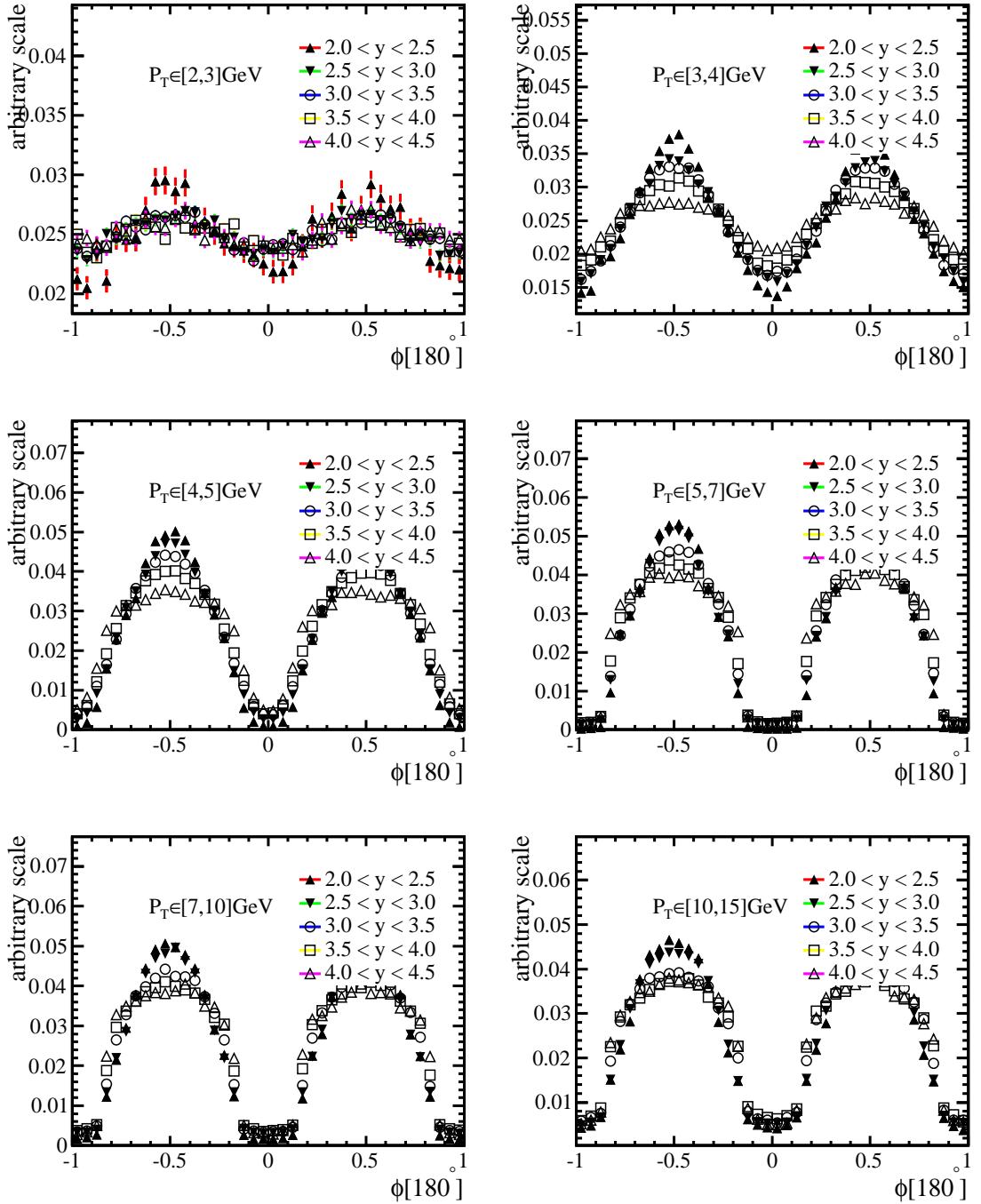


Figure A.4 Efficiency ( $\epsilon_{\text{tot}}$ ) as a function of  $\phi$  in each  $p_T$  and rapidity of  $J/\psi$  is shown in the CS frame. Each curve is normalized to unity.

the  $x$ -axis in the  $x - y$  plane) means that one muon flies predominantly out of LHCb acceptance region and the other one flies into LHCb detector almost parallel the production plane. On the other hand, the two cases  $\phi \approx \pi/2$  or  $3\pi/2$  mean that the  $J/\psi$  decay plane formed by the two muon momenta is almost perpendicular to production plane.

Figure A.5 shows distributions of  $\cos \theta$  and  $\phi$  after the geometry requirement (both muons in the LHCb acceptance). The number of events with  $\cos \theta \approx \pm 1$  are heavily filtered compared to those with  $\cos \theta \approx 0$  mainly because the muon flying backwards in  $J/\psi$  rest frame can't be boosted to the LHCb detector region easily. When the  $J/\psi$  momentum is large enough, from the low rapidity bins to the high rapidity bins, fewer events are lost by the acceptance. However when the momentum (large rapidity) is ultrahigh, the  $J/\psi$  will fly along the beam pipe (in the region less than 10 mrad acceptance), the muons with  $\cos \theta \approx \pm 1$  will easily go into the beam pipe, so that for the high  $p_T$  high rapidity bin, the efficiency drops at  $\cos \theta \approx \pm 1$  compared to middle momentum bins. For the  $J/\psi$  flying closely along the LHCb inner and outer boundary, the efficiency is lower at  $\phi \approx 0$  ( $\pi$ ).

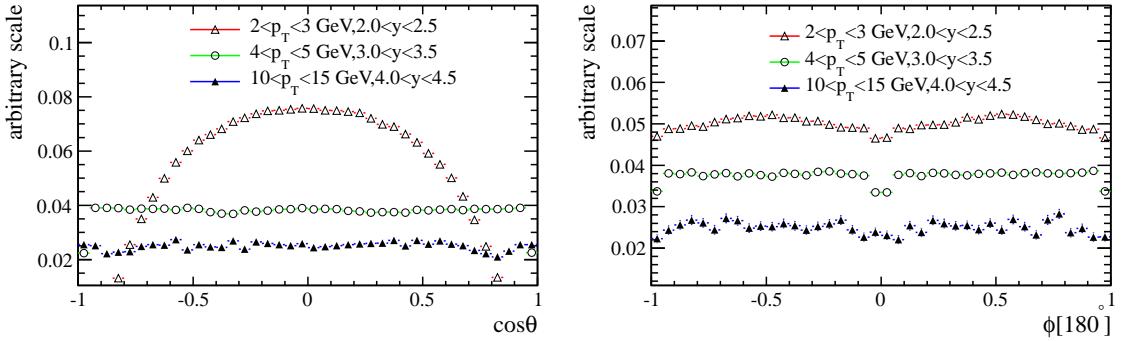


Figure A.5 The  $\cos \theta$  (left) and  $\phi$  (right) distributions after the geometry requirement.

It is known that muons flying along the (especially outer) boundary of LHCb detector have smaller reconstruction efficiency to be good tracks, predominantly because these charged tracks will be easily bent outside of the LHCb coverage by the magnetic field. Besides, tracks with lower momentum will probably have lower reconstruction efficiency because of multiple interaction with detector material and because they are easily affected by the magnetic field. So to have larger reconstruction efficiency for both muons (and thus the  $J/\psi$  meson),  $\cos \theta$  tends to be around 0 for both muons to have large enough momentum, and  $\phi$  tends to be around  $\pi/2$  or  $3\pi/2$  to avoid the situation when one muon easily flying outside LHCb coverage (around boundaries). Figure A.6 shows  $\cos \theta$  and  $\phi$

distribution for reconstructed events.

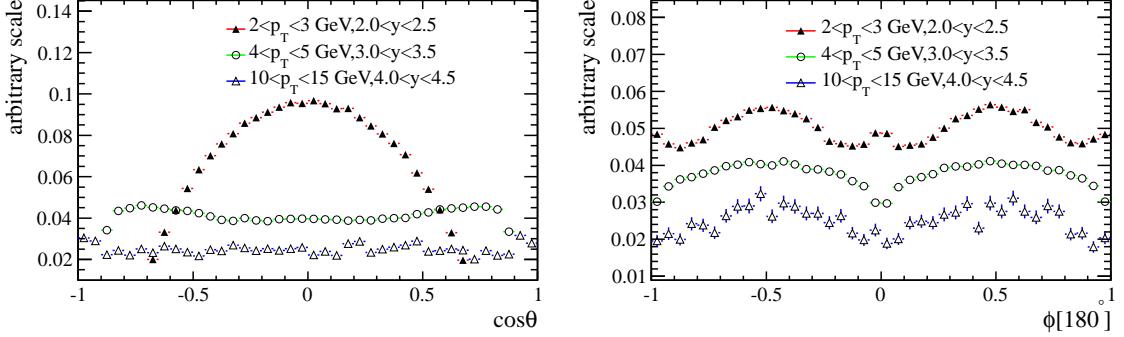


Figure A.6 The  $\cos \theta$  (left) and  $\phi$  (right) distributions at the reconstruction level before offline cuts.

Most of the selections (track quality, vertex quality for example) are very loose and have little effect on  $\cos \theta$  and  $\phi$  distributions, except for the muon  $p_T$  and trigger requirement and the muon PID cuts. Muon PID cut efficiency increase with muon momentum, and plateau above  $\approx 20$  GeV/ $c$ , so this cut will keep more events with  $\cos \theta \approx 0$ , because both muons will have a relatively large momentum in this case. Figure A.7 shows  $\cos \theta$  and  $\phi$  distribution for events after all selections including muon PID except the muon transverse momentum and trigger requirement.

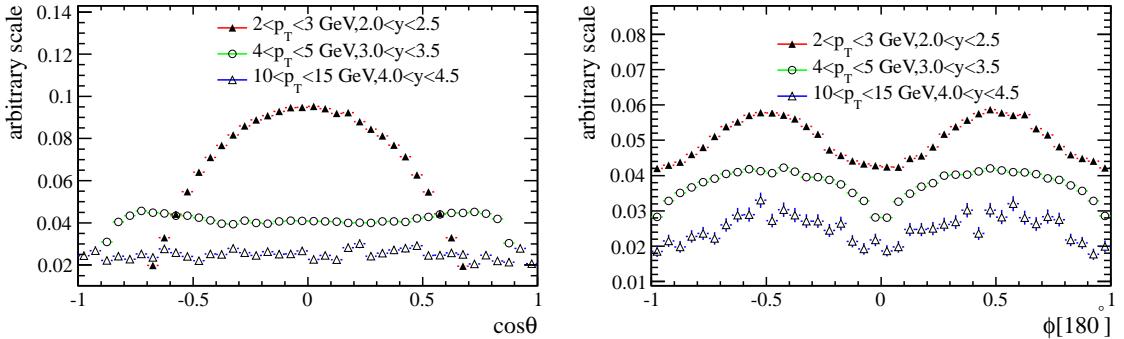


Figure A.7 The  $\cos \theta$  (left) and  $\phi$  (right) distributions after all selections but muon  $p_T$  and trigger requirements.

The requirement  $p_T > 750$  MeV/ $c$  on muons is used to select good muons, and will heavily filter  $J/\psi$  with lower  $p_T$  and rapidity, because in these bins the events with  $\cos \theta \approx \pm 1$  and  $\phi \approx 0(\pi)$  will probably produce one muon with lower  $p_T$ . However for  $J/\psi$  mesons with large rapidity and large  $p_T$ , both muons will carry a large fraction of the  $J/\psi$   $p_T$ , so the  $p_T$  cut on muons is easily fulfilled. So again, in the low  $p_T$  low rapidity bins, events with  $\cos \theta \approx \pm 1$  and  $\phi \approx 0(\pi)$  are more reduced.

The triggers (see Table 4.2) require the transverse momentum of both of the muons to be large enough, so it has similar effect as the offline  $p_T$  cut, but will affect strongly even high  $p_T$  and rapidity bins, shown in Figure A.8, as the cut is much tighter.

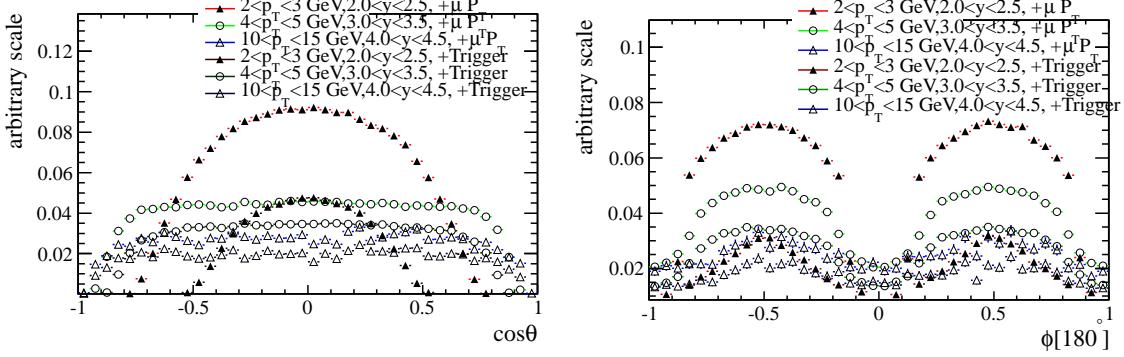


Figure A.8 The  $\cos \theta$  (left) and  $\phi$  (right) distributions after muons transverse momentum selection and trigger requirements.

In summary, the final  $\cos \theta$  shape of the efficiency is mainly produced by the geometrical acceptance, which requires both muons to fall in the LHCb full detector region. The trigger will also change the  $\cos \theta$  shape for low  $p_T$  and low rapidity bins by requiring the transverse momentum of muons to be large enough, and more events with  $\cos \theta \approx 0$  are kept. The final  $\phi$  shape is also partially produced by the geometry acceptance, because  $\phi \approx 0$  or  $\pi$  means one muon will easily fall outside of the LHCb geometry. However,  $\phi$  shape is mainly produced by the requirements related to muon  $p_T$  in the offline  $p_T$  cut and by the triggers.  $\phi \approx 0$  or  $\pi$  means that one of the two muons will have lower  $p_T$  and the other one have large  $p_T$ , and if both muon  $p_T$  are required to be large enough offline and during the triggers, the events are heavily rejected.

## A.2 Background subtraction validation

This section tries to demonstrate that the angular distribution for background events in signal region is (roughly) the same as the distribution for background events in signal region using the  $J/\psi$  measurement in the HX frame as an example. However they are not required to be exactly the same, and in the section 4.3.5 a conservative systematic uncertainty is assigned concerning the (small) inconsistency between the two distributions.

Since the background events in the signal region can't be separated event by event from the signal, it is not possible to extract the pure angular distribution for background events in the signal region to compare with those from the sidebands. The following studies try to justify that by combining the left and right sidebands, the distribution in signal region can be reproduced. The nominal sideband region (signal) are defined as within from  $4\sigma$  to  $7\sigma$  away from the mass peak (within  $3\sigma$  around the peak) where  $\sigma$  is about  $15 \text{ MeV}/c^2$ . In this study, another two definition of sidebands are introduced:

A: left  $[\mu - 100\text{MeV}/c^2, \mu - 85\text{MeV}/c^2]$ , right  $[\mu + 85\text{MeV}/c^2, \mu + 100\text{MeV}/c^2]$

B: left  $[\mu - 70\text{MeV}/c^2, \mu - 55\text{MeV}/c^2]$ , right  $[\mu + 55\text{MeV}/c^2, \mu + 70\text{MeV}/c^2]$

The choice of sideband [B] is nearer to the signal region. Figure A.9 shows the  $\cos\theta$  and  $\phi$  distribution for the left sideband in definition [A], the right sideband in definition [A] and the sum of the left sideband and right sideband in definition [B] (which is called (pseudo) signal region). The three distributions are very similar, but not identical within statistical fluctuations, and the shape changes smoothly from left sideband to the pseudo signal region and to the right sideband: for each  $\cos\theta$  or  $\phi$  bin, the points for the pseudo signal region almost always lie between the two points for the two sidebands in definition [A], which means that the sum of the two sidebands in definition [A] is almost identical to the sum of the two sidebands in definition [B]. In fact in Figure A.10, the ratio of  $\cos\theta$  ( $\phi$ ) distribution in the pseudo signal region (combination of sideband [A]) over the sum of the distribution in sideband [B] is plotted, from which it can be seen that the ratio is almost flat everywhere with a value 1.05 which is not 1 because of the contamination of signal events in the pseudo signal region (overlap with radiative tail region). The Figure A.11 illustrates the relative difference (pull distribution) of the entries in the two dimensional  $\cos\theta - \phi$  histograms of sideband [A] and sideband [B], and the figure tell us that the two distributions are consistent.

So, the two choices of sidebands [A] and [B] give similar  $\cos\theta - \phi$  distribution, which means that a combination of two background mass regions that are of equal width

and are symmetrical around the mass peak ( $\mu$ ) will produce similar background angular distribution regardless of the distance between the background region and the signal peak. As the nominal signal region is a special sideband region and are symmetrical around the signal peak, so the nominal signal region and nominal sideband region will produce similar  $\cos \theta - \phi$  distributions.

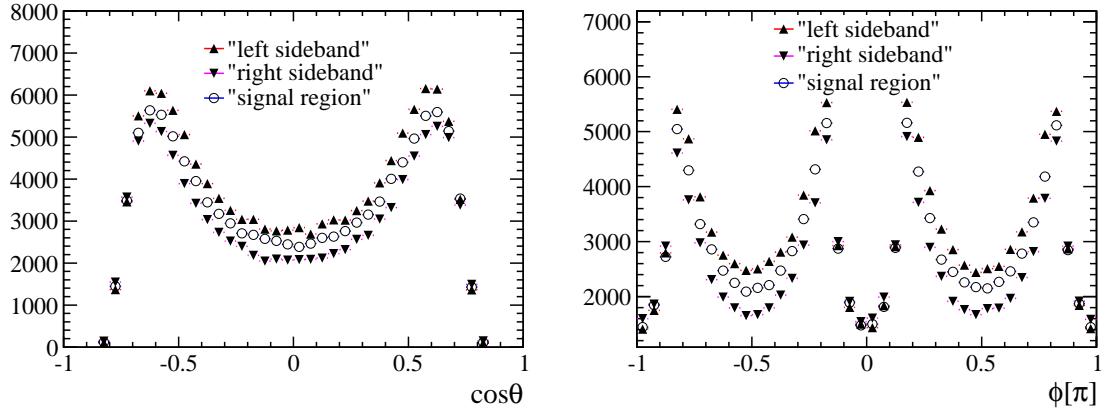


Figure A.9 The  $\cos \theta$  (left) and  $\phi$  (right) distributions for backgrounds in left sideband (upward triangle) and right sideband (upward triangle) of definition [A] and combination of sideband [B] (circle).

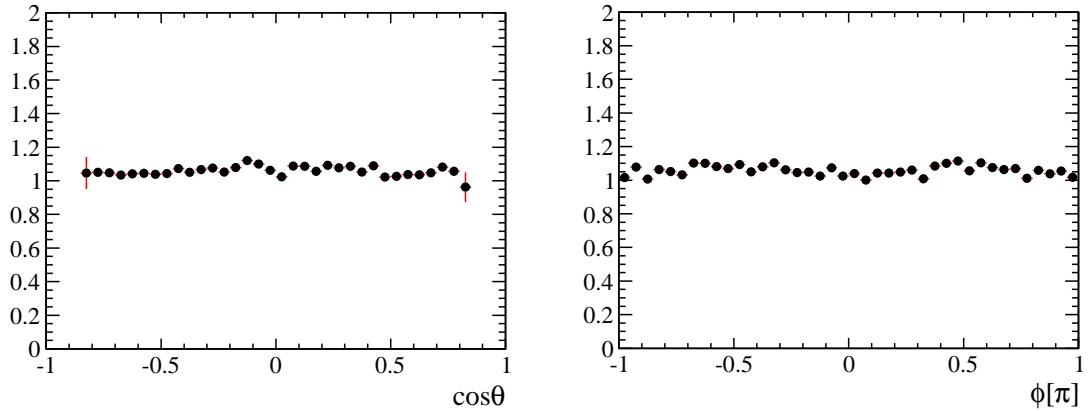


Figure A.10 The ratio of  $\cos \theta$  (left) and  $\phi$  (right) distributions in sideband [B] (left + right) over the distributions in sideband [A] (left + right).

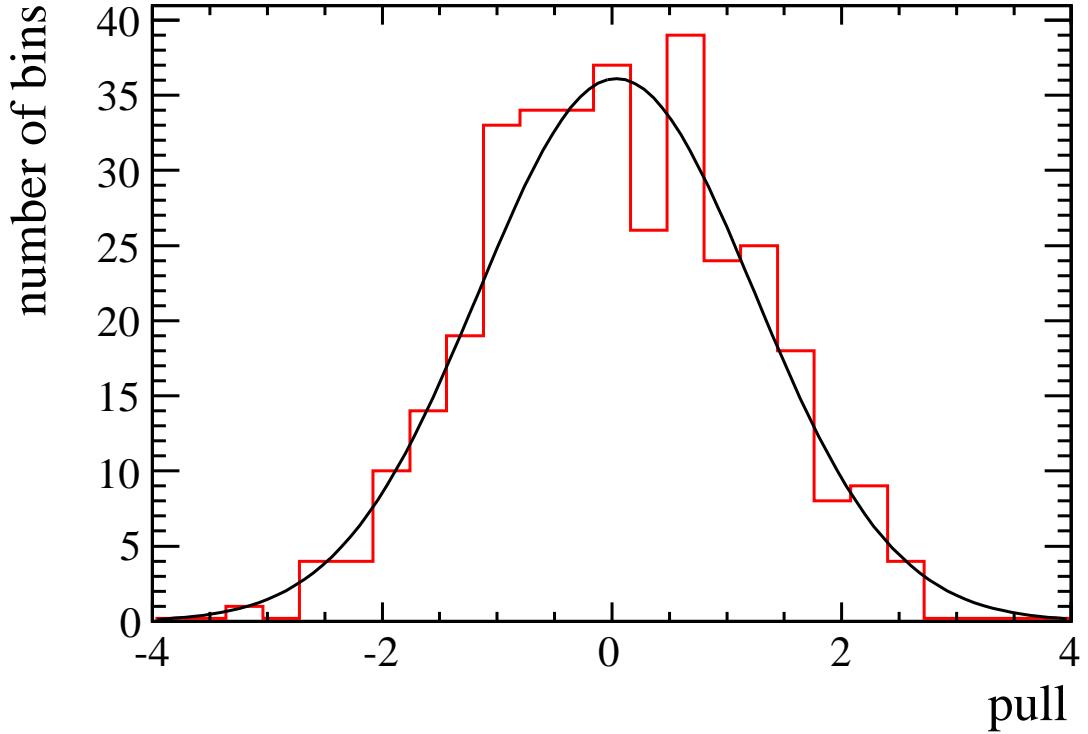


Figure A.11 The distribution of the relative differences of the entries in the two dimensional  $\cos \theta - \phi$  histograms for sideband [A] (left + right) and sideband [B] (left + right).

### A.3 DLL method versus *s*Weight method

The maximum likelihood (DLL estimator) constructed for the  $J/\psi$  analysis is similar to the *s*Weight method<sup>[189,190]</sup> for the  $\psi(2S)$  analysis; the only difference is that in the DLL method the background is subtracted with weights -1 for the sideband events and +1 for signal region events, while the *s*Weight method subtracts the background with more elegant way, by using a function as the weight for each event. In both cases, the weighting is a function of the mass of each event. In the following the results from the DLL estimator with the results with *s*Weight estimator are compared for the  $J/\psi$  analysis.

The  $J/\psi$  mass distribution is parameterized with a CB function for the signal plus exponential for the background in each  $J/\psi$  kinematic bin, and with the two p.d.f (signal and background) the signal *s*Weight function can calculated. In the *s*Weight method all the signal events contribute to the likelihood, while in the estimator of  $J/\psi$  analysis, a small fraction of signal events is also subtracted in the estimator because they fall into the sideband region, especially those events in the radiative tail. Because effective less signal events contributes to the  $J/\psi$  polarization estimator compared to the *s*Weight estimator,

in general the error returned by our estimator will be slightly larger than the error returned by the *s*Weight estimator. In Figure A.12 the results extracted from the two methods (in the first rapidity bins for example) are compared, showing very good consistency.

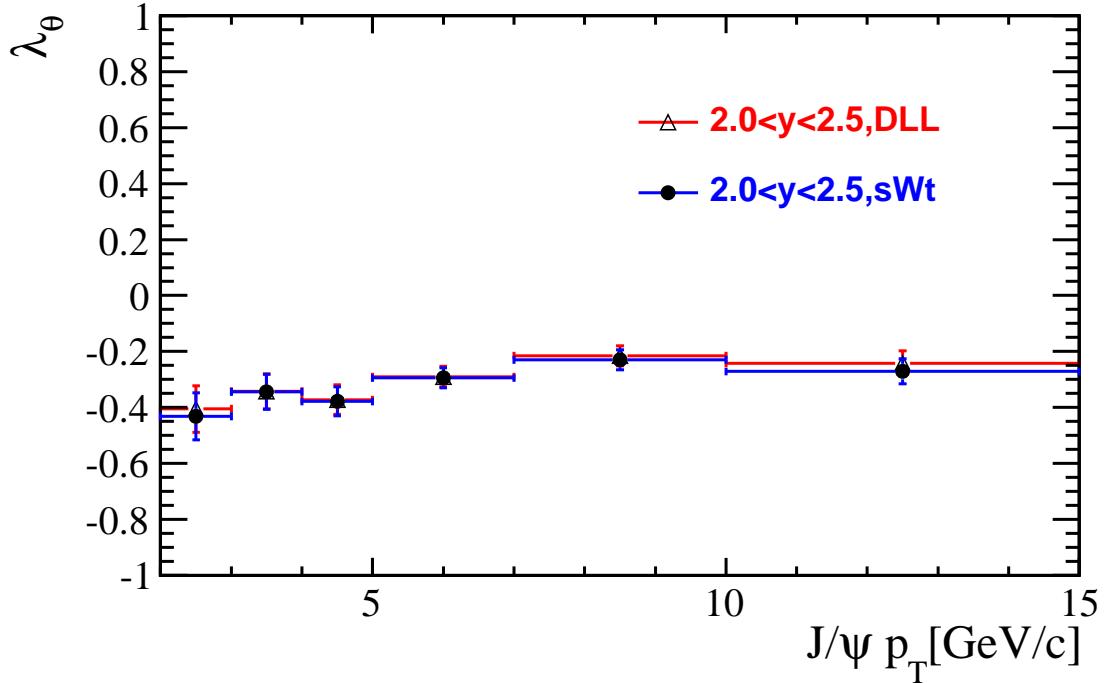


Figure A.12  $\lambda_\theta$  extracted with the DLL estimator (triangle) and the *s*Weight estimator (circle) in the first rapidity bin. The errors are only statistical errors returned by the estimators, namely without corrections and any systematic uncertainties.

## A.4 The likelihood method goodness

This section tries to describe the quality of the fits for  $J/\psi$  polarization analysis.

For this analysis, the efficiency distribution as a function of  $\cos \theta$  and  $\phi$  is not parameterized analytically, so the fit function which is the product of polarization angular distribution multiplied by the efficiency can not be determined analytically, and an unbinned goodness of fit test will not possible. On the hand, the goodness of fit is tested with the binned (histogram)  $\chi^2$  method by comparing the weighted  $\cos \theta - \phi$  distribution in Monte Carlo with data. For Monte Carlo, to each event a weight is given when producing the 2-dimensional  $\cos \theta - \phi$  histogram distribution. The weight is chosen to be the polarization formula with the polarization parameters fixed to the measured values in each  $J/\psi$  kinematic bin, so in this way the distribution of the weighted Monte Carlo should be consistent with the distribution in data (background subtracted). The consistency of the two histograms is tested with the  $\chi^2$  method.

In Figure A.13, the  $p$ -values of the tests in the 30 kinematic bins are presented. It can be seen that there are several entries accumulated in the  $p < 0.05$  region and some of them corresponds to the low statistics  $J/\psi$  bins, where the test itself will favor small  $p$ -values and some of low  $p$  values may be caused by systematic uncertainties, since only statistical fluctuations are considered in the tests.

During the test several binning choices for the histograms are tried, and for the high statistics  $J/\psi$   $p_T$  and rapidity bins the  $p$ -values are reasonably stable with regards to the number of bins, but for  $J/\psi$  bins with low statistics, the  $p$ -values can be worse or better with different number of bins in the two dimensional histogram. To check the consistency between the weighted Monte Carlo and data visibly, in Figure A.14 and Figure A.15 the one-dimensional  $\cos \theta$  and  $\phi$  distributions for two  $J/\psi$  kinematic bins with low and high statistics are shown respectively. From the plots it can be seen that the distributions in the weighted Monte Carlo agree reasonably well with data.

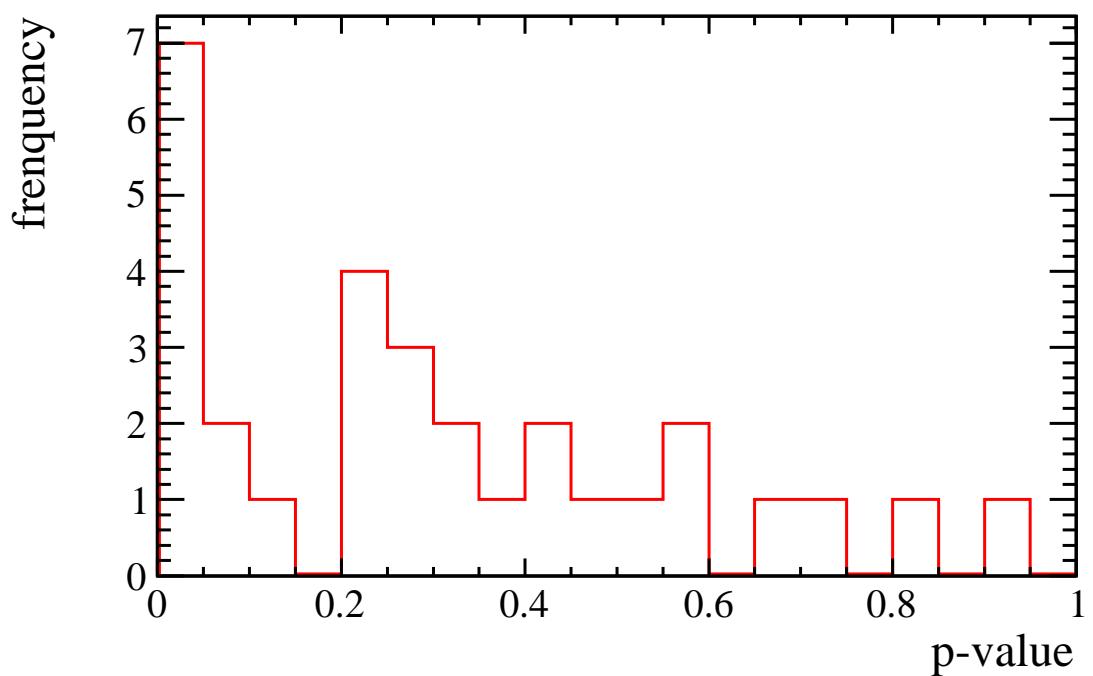


Figure A.13 The distribution of the  $p$ -values of the  $\chi^2$  test of the fit quality. In the ideal case the  $p$ -value is uniformly distributed between 0 and 1.

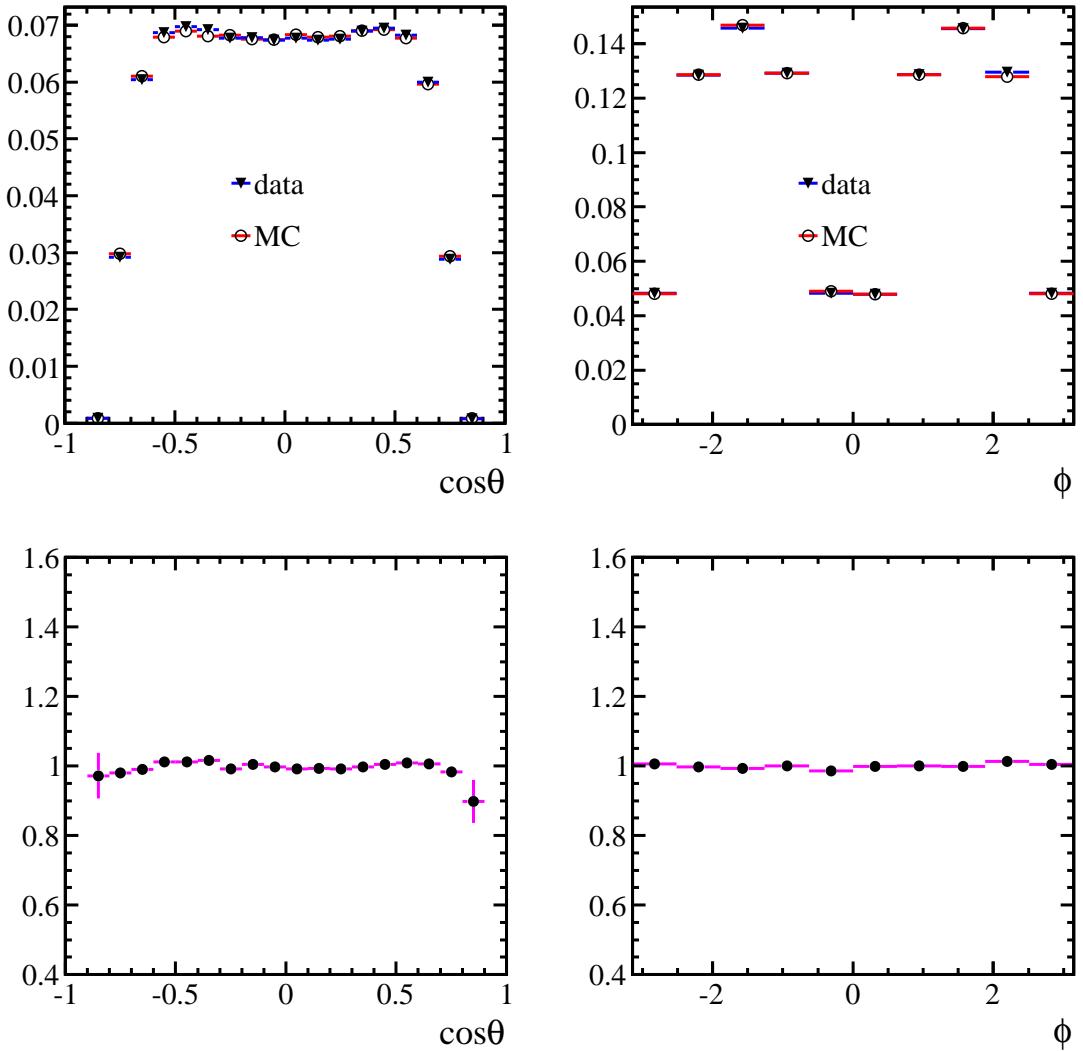


Figure A.14  $\cos\theta$  (upper left) and  $\phi$  (upper right) distribution in the weighted Monte Carlo (circle) are compared to data (triangle) respectively. The ratio distributions of the Monte Carlo over data are also plotted in two plots at the bottom. This figure is for a  $J/\psi$  kinematic bin with high statistics.

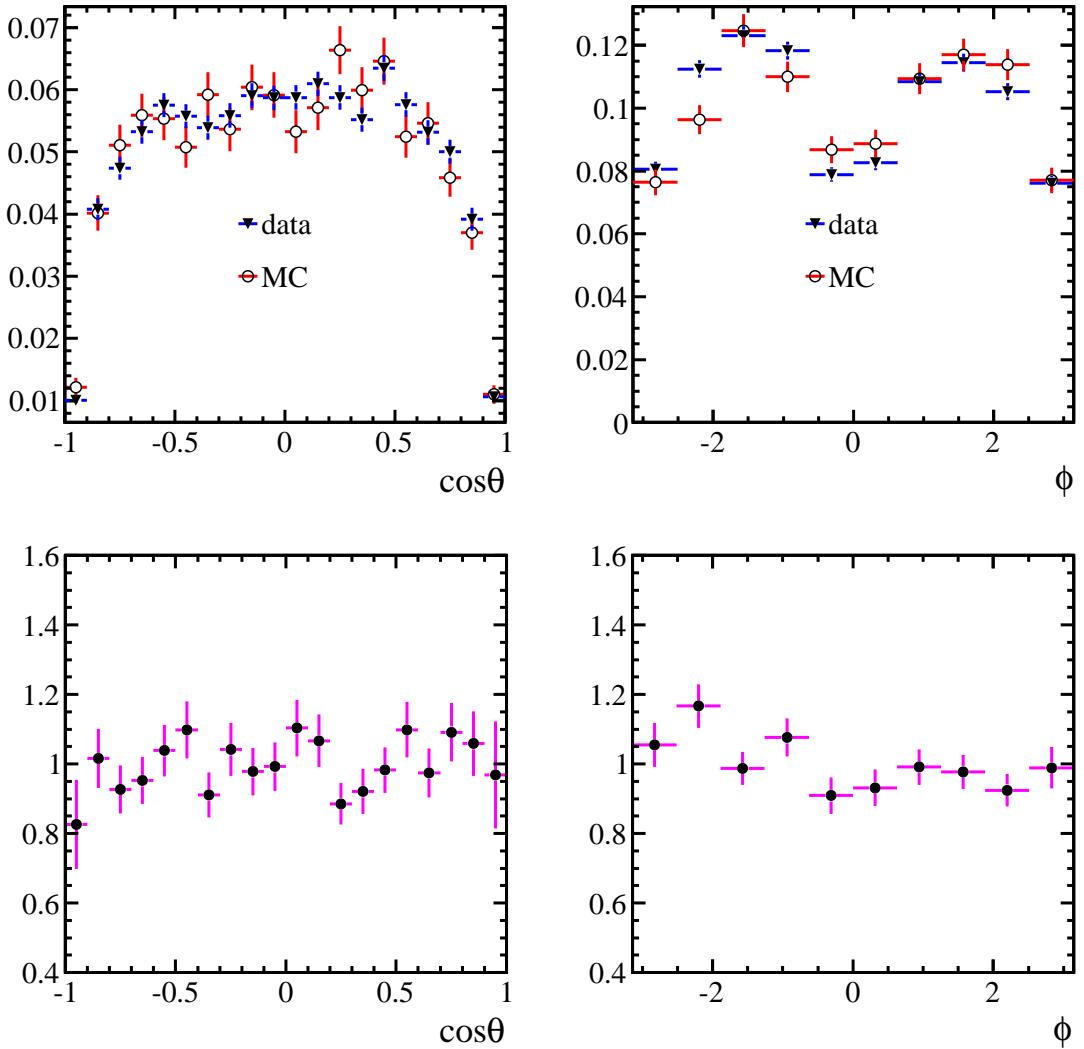


Figure A.15  $\cos\theta$  (upper left) and  $\phi$  (upper right) distribution in the weighted Monte Carlo (circle) are compared to data (triangle) respectively. The ratio distributions of the Monte Carlo over data are also plotted in two plots at the bottom. This figure is for a  $J/\psi$  kinematic bin with low statistics.

## Appendix B **Appendicular tables**

### B.1 Fitting the invariant mass spectrum

To define the signal region and the sideband regions for  $J/\psi$  candidate (see section 5.1), the resolution is quoted from the mass peak of the signal from the fit to the mass spectrum, which is parameterized with a Crystal Ball function (for signal component) plus an exponential (for background component). In Table B.1 the parameters  $\mu$  and  $\sigma$  of the Crystal Ball function and the  $\tau$  of the exponential are listed for each  $p_T$  and rapidity bin. For the  $\psi(2S)$  case, two Crystal Ball functions plus a first order polynomial are used to describe the mass spectrum, and the parameters for the CB and the polynomial are shown in Table B.2.

---

## Appendix B   **Appendicular tables**

---

Table B.1 The best fit parameters describing the mass spectrum in  $J/\psi$  data. In the table the  $\tau$  of the exponential is multiplied by 1000.

$p_T$ (GeV/c)		y bin			
$2 < p_T < 3$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\mu$ of CB	$3092.67 \pm 0.05$	$3091.90 \pm 0.05$	$3091.28 \pm 0.06$	$3090.77 \pm 0.06$	$3089.94 \pm 0.06$
$\sigma$ of CB	$10.69 \pm 0.05$	$11.65 \pm 0.05$	$13.07 \pm 0.05$	$15.21 \pm 0.05$	$18.12 \pm 0.05$
n of CB	$0.98 \pm 0.13$	$0.95 \pm 0.14$	$1.24 \pm 0.20$	$1.61 \pm 0.26$	$1.88 \pm 0.28$
$\tau$ of Exp	$-0.33 \pm 0.12$	$-0.72 \pm 0.12$	$-0.87 \pm 0.13$	$-1.24 \pm 0.15$	$-1.55 \pm 0.21$
$3 < p_T < 4$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\mu$ of CB	$3092.48 \pm 0.04$	$3091.66 \pm 0.04$	$3090.95 \pm 0.05$	$3090.59 \pm 0.05$	$3089.87 \pm 0.06$
$\sigma$ of CB	$11.16 \pm 0.04$	$12.12 \pm 0.04$	$13.50 \pm 0.04$	$15.79 \pm 0.05$	$18.93 \pm 0.05$
n of CB	$0.99 \pm 0.10$	$1.07 \pm 0.13$	$1.14 \pm 0.13$	$1.38 \pm 0.17$	$2.15 \pm 0.30$
$\tau$ of Exp	$-0.74 \pm 0.13$	$-0.80 \pm 0.13$	$-0.95 \pm 0.16$	$-1.33 \pm 0.21$	$-1.86 \pm 0.22$
$4 < p_T < 5$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\mu$ of CB	$3092.40 \pm 0.03$	$3091.52 \pm 0.04$	$3090.75 \pm 0.04$	$3090.25 \pm 0.05$	$3089.72 \pm 0.05$
$\sigma$ of CB	$11.57 \pm 0.03$	$12.52 \pm 0.04$	$13.97 \pm 0.04$	$16.37 \pm 0.04$	$19.71 \pm 0.05$
n of CB	$0.92 \pm 0.07$	$0.95 \pm 0.07$	$1.14 \pm 0.10$	$1.60 \pm 0.17$	$2.07 \pm 0.26$
$\tau$ of Exp	$-0.73 \pm 0.15$	$-0.65 \pm 0.19$	$-1.26 \pm 0.22$	$-1.74 \pm 0.25$	$-2.13 \pm 0.26$
$5 < p_T < 7$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\mu$ of CB	$3092.32 \pm 0.04$	$3091.30 \pm 0.04$	$3090.61 \pm 0.04$	$3090.25 \pm 0.05$	$3089.67 \pm 0.06$
$\sigma$ of CB	$12.21 \pm 0.03$	$13.05 \pm 0.03$	$14.65 \pm 0.04$	$17.02 \pm 0.04$	$20.81 \pm 0.05$
n of CB	$0.96 \pm 0.06$	$1.09 \pm 0.07$	$1.31 \pm 0.11$	$1.79 \pm 0.19$	$2.65 \pm 0.40$
$\tau$ of Exp	$-0.75 \pm 0.22$	$-1.33 \pm 0.24$	$-1.72 \pm 0.29$	$-2.59 \pm 0.31$	$-2.47 \pm 0.31$
$7 < p_T < 10$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\mu$ of CB	$3092.05 \pm 0.03$	$3090.96 \pm 0.03$	$3090.36 \pm 0.04$	$3090.20 \pm 0.05$	$3089.68 \pm 0.08$
$\sigma$ of CB	$13.05 \pm 0.03$	$13.84 \pm 0.03$	$15.61 \pm 0.04$	$18.30 \pm 0.05$	$22.64 \pm 0.07$
n of CB	$1.10 \pm 0.06$	$1.13 \pm 0.06$	$1.36 \pm 0.11$	$2.09 \pm 0.27$	$4.00 \pm 2.62$
$\tau$ of Exp	$-1.74 \pm 0.27$	$-2.00 \pm 0.31$	$-2.13 \pm 0.37$	$-2.91 \pm 0.37$	$-2.90 \pm 0.30$
$10 < p_T < 15$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\mu$ of CB	$3091.72 \pm 0.04$	$3090.67 \pm 0.04$	$3090.17 \pm 0.06$	$3090.20 \pm 0.09$	$3090.16 \pm 0.20$
$\sigma$ of CB	$14.25 \pm 0.04$	$15.25 \pm 0.04$	$17.17 \pm 0.05$	$20.33 \pm 0.08$	$25.71 \pm 0.19$
n of CB	$1.33 \pm 0.10$	$1.44 \pm 0.11$	$1.61 \pm 0.17$	$2.50 \pm 0.50$	$4.00 \pm 2.92$
$\tau$ of Exp	$-1.91 \pm 0.35$	$-2.37 \pm 0.33$	$-2.55 \pm 0.42$	$-2.09 \pm 0.51$	$-1.16 \pm 0.60$

Table B.2 The best parameters of the CB functions and the first order polynomial describing the mass spectrum in  $\psi(2S)$  data. The coefficient of the first order in polynomial is multiplied by 1000.

$p_T$ (GeV/c)		y bin			
$3.5 < p_T < 4$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\mu$ of CB1	$3689.54 \pm 0.32$	$3689.08 \pm 0.16$	$3689.23 \pm 0.16$	$3689.16 \pm 0.18$	$3689.75 \pm 0.29$
$\sigma$ of CB1	$11.44 \pm 0.33$	$12.64 \pm 0.16$	$14.08 \pm 0.16$	$16.35 \pm 0.19$	$19.93 \pm 0.31$
c of poly.	$-0.1920 \pm 0.0110$	$-0.2030 \pm 0.0040$	$-0.1820 \pm 0.0060$	$-0.1900 \pm 0.0070$	$-0.1970 \pm 0.0160$
$4 < p_T < 5$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\mu$ of CB1	$3689.71 \pm 0.22$	$3689.13 \pm 0.11$	$3688.89 \pm 0.11$	$3688.85 \pm 0.13$	$3689.22 \pm 0.23$
$\sigma$ of CB1	$11.77 \pm 0.22$	$12.86 \pm 0.11$	$14.58 \pm 0.11$	$16.88 \pm 0.13$	$20.51 \pm 0.25$
c of poly.	$-0.2000 \pm 0.0080$	$-0.1920 \pm 0.0040$	$-0.1840 \pm 0.0050$	$-0.1870 \pm 0.0080$	$-0.1860 \pm 0.0190$
$5 < p_T < 7$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\mu$ of CB1	$3689.65 \pm 0.18$	$3688.97 \pm 0.09$	$3688.79 \pm 0.09$	$3688.94 \pm 0.12$	$3689.31 \pm 0.23$
$\sigma$ of CB1	$12.12 \pm 0.18$	$13.24 \pm 0.09$	$14.82 \pm 0.09$	$17.61 \pm 0.12$	$21.60 \pm 0.25$
c of poly.	$-0.1760 \pm 0.0180$	$-0.1760 \pm 0.0080$	$-0.1810 \pm 0.0080$	$-0.1860 \pm 0.0120$	$-0.1920 \pm 0.0220$
$7 < p_T < 10$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\mu$ of CB1	$3689.48 \pm 0.22$	$3688.88 \pm 0.11$	$3688.81 \pm 0.12$	$3688.64 \pm 0.18$	$3688.33 \pm 0.37$
$\sigma$ of CB1	$13.44 \pm 0.21$	$13.93 \pm 0.10$	$15.85 \pm 0.12$	$18.96 \pm 0.18$	$23.76 \pm 0.40$
c of poly.	$-0.0820 \pm 0.0920$	$-0.1670 \pm 0.0220$	$-0.0970 \pm 0.0540$	$-0.2070 \pm 0.0160$	$-0.2130 \pm 0.0250$
$10 < p_T < 15$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
$\mu$ of CB1	$3689.71 \pm 0.34$	$3688.93 \pm 0.20$	$3688.62 \pm 0.25$	$3688.82 \pm 0.36$	$3687.68 \pm 0.98$
$\sigma$ of CB1	$14.40 \pm 0.33$	$15.10 \pm 0.19$	$17.33 \pm 0.25$	$20.43 \pm 0.39$	$26.68 \pm 1.26$
c of poly.	$-0.0000 \pm 0.2510$	$-0.1590 \pm 0.0600$	$-0.1980 \pm 0.0310$	$-0.0000 \pm 0.2750$	$-0.1430 \pm 0.2130$

## B.2 Normalization parameters

In the construction of likelihood estimator, Monte Carlo are used to calculate the three constants determine the normalization of the angular distribution, which is the parameterized polarization angular distribution multiplied by the efficiency. In Table B.3 and Table B.4, these constants for the  $J/\psi$  and  $\psi(2S)$  polarization measurement are listed respectively.

Table B.3 Extracted normalization constants ( $a, b, c$ ) in the estimator for  $J/\psi$  polarization measurement.

$p_T$ (GeV/c)		y bin			
$2 < p_T < 3$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
a	$0.0616 \pm 0.0003$	$0.1110 \pm 0.0002$	$0.1879 \pm 0.0003$	$0.2235 \pm 0.0004$	$0.2249 \pm 0.0006$
b	$0.1865 \pm 0.0012$	$0.1011 \pm 0.0007$	$0.0425 \pm 0.0008$	$0.0746 \pm 0.0009$	$0.2089 \pm 0.0012$
c	$-0.4871 \pm 0.0027$	$-0.3444 \pm 0.0011$	$-0.2762 \pm 0.0009$	$-0.2883 \pm 0.0010$	$-0.3354 \pm 0.0014$
$3 < p_T < 4$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
a	$0.0774 \pm 0.0004$	$0.1219 \pm 0.0003$	$0.1993 \pm 0.0004$	$0.2368 \pm 0.0005$	$0.2463 \pm 0.0008$
b	$0.1682 \pm 0.0011$	$0.0066 \pm 0.0008$	$-0.1132 \pm 0.0009$	$-0.0979 \pm 0.0010$	$0.0354 \pm 0.0014$
c	$-0.5815 \pm 0.0022$	$-0.3552 \pm 0.0011$	$-0.2517 \pm 0.0011$	$-0.2847 \pm 0.0011$	$-0.4244 \pm 0.0013$
$4 < p_T < 5$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
a	$0.0953 \pm 0.0006$	$0.1296 \pm 0.0004$	$0.2008 \pm 0.0005$	$0.2391 \pm 0.0006$	$0.2639 \pm 0.0011$
b	$0.1709 \pm 0.0013$	$-0.0321 \pm 0.0009$	$-0.1600 \pm 0.0011$	$-0.1620 \pm 0.0012$	$-0.0697 \pm 0.0018$
c	$-0.5703 \pm 0.0025$	$-0.2650 \pm 0.0015$	$-0.1698 \pm 0.0014$	$-0.2055 \pm 0.0015$	$-0.3682 \pm 0.0018$
$5 < p_T < 7$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
a	$0.1147 \pm 0.0007$	$0.1415 \pm 0.0004$	$0.2039 \pm 0.0005$	$0.2376 \pm 0.0007$	$0.2704 \pm 0.0011$
b	$0.1806 \pm 0.0013$	$-0.0409 \pm 0.0009$	$-0.1536 \pm 0.0011$	$-0.1741 \pm 0.0013$	$-0.1402 \pm 0.0019$
c	$-0.4935 \pm 0.0026$	$-0.1614 \pm 0.0015$	$-0.0998 \pm 0.0014$	$-0.1318 \pm 0.0016$	$-0.2822 \pm 0.0021$
$7 < p_T < 10$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
a	$0.1413 \pm 0.0010$	$0.1670 \pm 0.0006$	$0.2179 \pm 0.0008$	$0.2367 \pm 0.0011$	$0.2669 \pm 0.0018$
b	$0.2060 \pm 0.0021$	$-0.0319 \pm 0.0015$	$-0.1198 \pm 0.0018$	$-0.1410 \pm 0.0022$	$-0.1663 \pm 0.0032$
c	$-0.3518 \pm 0.0040$	$-0.0774 \pm 0.0022$	$-0.0495 \pm 0.0022$	$-0.0666 \pm 0.0027$	$-0.1810 \pm 0.0037$
$10 < p_T < 15$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
a	$0.1743 \pm 0.0021$	$0.2035 \pm 0.0014$	$0.2378 \pm 0.0018$	$0.2488 \pm 0.0024$	$0.2673 \pm 0.0037$
b	$0.2069 \pm 0.0043$	$-0.0215 \pm 0.0032$	$-0.0778 \pm 0.0039$	$-0.0924 \pm 0.0050$	$-0.1503 \pm 0.0072$
c	$-0.2291 \pm 0.0071$	$-0.0323 \pm 0.0040$	$-0.0193 \pm 0.0044$	$-0.0275 \pm 0.0056$	$-0.0814 \pm 0.0082$

Table B.4 Extracted normalization constants ( $a, b, c$ ) in the estimator for  $\psi(2S)$  polarization measurement.

$p_T$ (GeV/c)		y bin			
$3.5 < p_T < 4$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
a	0.0908 $\pm$ 0.0010	0.1369 $\pm$ 0.0007	0.2196 $\pm$ 0.0009	0.2373 $\pm$ 0.0013	0.2190 $\pm$ 0.0020
b	0.1953 $\pm$ 0.0023	0.0245 $\pm$ 0.0016	-0.1289 $\pm$ 0.0020	-0.1150 $\pm$ 0.0023	0.0352 $\pm$ 0.0031
c	-0.5643 $\pm$ 0.0040	-0.4135 $\pm$ 0.0021	-0.3012 $\pm$ 0.0022	-0.3383 $\pm$ 0.0025	-0.4815 $\pm$ 0.0032
$4 < p_T < 5$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
a	0.1013 $\pm$ 0.0008	0.1415 $\pm$ 0.0005	0.2182 $\pm$ 0.0007	0.2399 $\pm$ 0.0010	0.2355 $\pm$ 0.0016
b	0.1921 $\pm$ 0.0016	-0.0180 $\pm$ 0.0012	-0.1723 $\pm$ 0.0015	-0.1647 $\pm$ 0.0018	-0.0288 $\pm$ 0.0026
c	-0.5704 $\pm$ 0.0028	-0.3549 $\pm$ 0.0018	-0.2393 $\pm$ 0.0017	-0.2794 $\pm$ 0.0020	-0.4522 $\pm$ 0.0026
$5 < p_T < 7$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
a	0.1186 $\pm$ 0.0008	0.1467 $\pm$ 0.0005	0.2113 $\pm$ 0.0007	0.2326 $\pm$ 0.0009	0.2539 $\pm$ 0.0017
b	0.1893 $\pm$ 0.0015	-0.0505 $\pm$ 0.0011	-0.1859 $\pm$ 0.0014	-0.1956 $\pm$ 0.0017	-0.1247 $\pm$ 0.0029
c	-0.5440 $\pm$ 0.0027	-0.2291 $\pm$ 0.0018	-0.1471 $\pm$ 0.0018	-0.1851 $\pm$ 0.0022	-0.3602 $\pm$ 0.0031
$7 < p_T < 10$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
a	0.1387 $\pm$ 0.0012	0.1610 $\pm$ 0.0007	0.2111 $\pm$ 0.0010	0.2272 $\pm$ 0.0014	0.2598 $\pm$ 0.0025
b	0.2014 $\pm$ 0.0021	-0.0526 $\pm$ 0.0016	-0.1563 $\pm$ 0.0020	-0.1738 $\pm$ 0.0027	-0.1910 $\pm$ 0.0045
c	-0.4132 $\pm$ 0.0041	-0.1192 $\pm$ 0.0025	-0.0660 $\pm$ 0.0027	-0.0929 $\pm$ 0.0035	-0.2238 $\pm$ 0.0053
$10 < p_T < 15$	$2.0 < y < 2.5$	$2.5 < y < 3.0$	$3.0 < y < 3.5$	$3.5 < y < 4.0$	$4.0 < y < 4.5$
a	0.1641 $\pm$ 0.0019	0.1872 $\pm$ 0.0012	0.2234 $\pm$ 0.0018	0.2342 $\pm$ 0.0028	0.2637 $\pm$ 0.0051
b	0.2108 $\pm$ 0.0036	-0.0348 $\pm$ 0.0029	-0.1013 $\pm$ 0.0038	-0.1130 $\pm$ 0.0056	-0.1629 $\pm$ 0.0096
c	-0.2574 $\pm$ 0.0064	-0.0416 $\pm$ 0.0039	-0.0179 $\pm$ 0.0046	-0.0301 $\pm$ 0.0068	-0.1243 $\pm$ 0.0113

### B.3 Summary of all $J/\psi$ and $\psi(2S)$ polarization results in the HX and the CS frames

In the following, the various systematic uncertainties in different kinematics bins in the HX frame and the CS frame for  $J/\psi$  and  $\psi(2S)$  are summarized respectively.

Table B.5 The measured  $J/\psi$  polarization parameters, and their statistical errors and systematic uncertainties in different  $p_T$  bins in rapidity bin  $2.0 < y < 2.5$  in the HX frame. LL stat are the statistical errors returned by likelihood estimator, SB stat are the statistical errors coming from background subtraction, MCsys are the systematic errors coming from fluctuations of the normalization computed from Monte Carlo, SB/Specsys are the background subtraction/unkown  $J/\psi$  spectrum (binning) related systematic uncertainties, tzs sys are systematic uncertainties coming from  $\tau_S$  cut. Trk/AccSys are systematic uncertainties related to the difference of efficiency between MC and data, it combines the tracking and acceptance systematic uncertainties quadratically. BConSys are systematic uncertainties coming from contamination of  $J/\psi$  from  $b$  decay to our prompt sample. stat sum is the quadratically added statistical errors while sys sum is the quadratically added systematic uncertainties. Bin migration systematic uncertainties are so small that they are neglected.

204

$p_T(\text{GeV}/c)$	$\lambda_s$	fit	LL stat	SB stat	MCsys	SB sys	Specsys	tzs sys	Trk/AccSys	BConSys	stat sum	sys sum
2 – 3	$\lambda_\theta$	-0.4054	0.0832	0.0451	0.1269	0.0988	0.1647	0.0144	0.0999	0.0017	0.0947	0.2514
	$\lambda_{\theta\phi}$	-0.0091	0.0387	0.0344	0.0576	0.0554	0.0285	0.0030	0.0688	0.0014	0.0518	0.1093
	$\lambda_\phi$	0.0305	0.0114	0.0110	0.0212	0.0554	0.0247	0.0002	0.0164	0.0047	0.0158	0.0665
3 – 4	$\lambda_\theta$	-0.3426	0.0630	0.0368	0.1196	0.0453	0.1089	0.0016	0.0771	0.0172	0.0730	0.1856
	$\lambda_{\theta\phi}$	-0.0680	0.0367	0.0244	0.0629	0.0313	0.0347	0.0022	0.0155	0.0052	0.0441	0.0801
	$\lambda_\phi$	0.0155	0.0136	0.0091	0.0213	0.0110	0.0121	0.0008	0.0175	0.0029	0.0163	0.0322
4 – 5	$\lambda_\theta$	-0.3724	0.0521	0.0200	0.0824	0.0246	0.0743	0.0001	0.0207	0.0082	0.0558	0.1158
	$\lambda_{\theta\phi}$	-0.0503	0.0330	0.0170	0.0549	0.0014	0.0301	0.0000	0.0088	0.0001	0.0371	0.0632
	$\lambda_\phi$	0.0229	0.0133	0.0064	0.0206	0.0027	0.0137	0.0003	0.0160	0.0051	0.0147	0.0300
5 – 7	$\lambda_\theta$	-0.2902	0.0362	0.0092	0.0495	0.0025	0.0516	0.0009	0.1647	0.0142	0.0374	0.1801
	$\lambda_{\theta\phi}$	-0.0758	0.0234	0.0070	0.0436	0.0015	0.0327	0.0009	0.0257	0.0061	0.0226	0.0606
	$\lambda_\phi$	-0.0054	0.0096	0.0032	0.0194	0.0016	0.0141	0.0009	0.0138	0.0024	0.0101	0.0278
7 – 10	$\lambda_\theta$	-0.2154	0.0359	0.0087	0.0638	0.0050	0.0360	0.0013	0.2242	0.0106	0.0370	0.2362
	$\lambda_{\theta\phi}$	0.0289	0.0219	0.0057	0.0347	0.0011	0.0215	0.0009	0.0102	0.0032	0.0226	0.0422
	$\lambda_\phi$	-0.0026	0.0100	0.0021	0.0163	0.0008	0.0127	0.0015	0.0096	0.0010	0.0103	0.0229
10 – 15	$\lambda_\theta$	-0.2431	0.0460	0.0086	0.0719	0.0076	0.0223	0.0047	0.0076	0.0193	0.0467	0.0786
	$\lambda_{\theta\phi}$	-0.0553	0.0260	0.0066	0.0477	0.0009	0.0172	0.0024	0.0345	0.0106	0.0268	0.0623
	$\lambda_\phi$	0.0045	0.0143	0.0020	0.0265	0.0008	0.0069	0.0049	0.0082	0.0009	0.0145	0.0290

Table B.6 The measured  $J/\psi$  polarization parameters, and their statistical errors and systematic uncertainties in different  $p_T$  bins in rapidity bin  $2.5 < y < 3.0$  in the HX frame. LL stat are the statistical errors returned by likelihood estimator, SB stat are the statistical errors coming from background subtraction, MCsys are the systematic errors coming from fluctuations of the normalization computed from Monte Carlo, SB/Specsys are the background subtraction/unkown  $J/\psi$  spectrum (binning) related systematic uncertainties, tzs sys are systematic uncertainties coming from  $\tau_S$  cut. Trk/AccSys are systematic uncertainties related to the difference of efficiency between MC and data, it combines the tracking and acceptance systematic uncertainties quadratically. BConSys are systematic uncertainties coming from contamination of  $J/\psi$  from  $b$  decay to our prompt sample. stat sum is the quadratically added statistical errors while sys sum is the quadratically added systematic uncertainties. Bin migration systematic uncertainties are so small that they are neglected.

$p_T(\text{GeV}/c)$	$\lambda_s$	fit	LL stat	SB stat	MCsys	SB sys	Specsys	tzs sys	Trk/AccSys	BConSys	stat sum	sys sum
2 – 3	$\lambda_\theta$	-0.2797	0.0091	0.0050	0.0097	0.0295	0.0384	0.0003	0.0738	0.0118	0.0104	0.0896
	$\lambda_{\theta\phi}$	-0.0583	0.0033	0.0029	0.0043	0.0242	0.0096	0.0003	0.0223	0.0008	0.0044	0.0346
	$\lambda_\phi$	0.0321	0.0021	0.0016	0.0020	0.0380	0.0083	0.0009	0.0113	0.0058	0.0027	0.0409
3 – 4	$\lambda_\theta$	-0.2345	0.0092	0.0042	0.0104	0.0417	0.0464	0.0002	0.1576	0.0129	0.0101	0.1703
	$\lambda_{\theta\phi}$	-0.0357	0.0035	0.0021	0.0055	0.0077	0.0057	0.0006	0.0203	0.0011	0.0041	0.0231
	$\lambda_\phi$	0.0047	0.0024	0.0016	0.0032	0.0005	0.0015	0.0004	0.0204	0.0067	0.0029	0.0218
4 – 5	$\lambda_\theta$	-0.2070	0.0099	0.0033	0.0158	0.0079	0.0272	0.0017	0.1852	0.0122	0.0104	0.1884
	$\lambda_{\theta\phi}$	-0.0089	0.0041	0.0017	0.0065	0.0006	0.0073	0.0008	0.0120	0.0051	0.0045	0.0163
	$\lambda_\phi$	0.0049	0.0026	0.0009	0.0042	0.0001	0.0006	0.0007	0.0116	0.0062	0.0028	0.0138
5 – 7	$\lambda_\theta$	-0.1770	0.0089	0.0028	0.0123	0.0027	0.0177	0.0017	0.1055	0.0117	0.0094	0.1084
	$\lambda_{\theta\phi}$	-0.0032	0.0038	0.0010	0.0059	0.0002	0.0062	0.0006	0.0100	0.0066	0.0051	0.0147
	$\lambda_\phi$	-0.0069	0.0024	0.0007	0.0036	0.0007	0.0002	0.0013	0.0094	0.0057	0.0025	0.0116
7 – 10	$\lambda_\theta$	-0.1702	0.0118	0.0025	0.0185	0.0037	0.0051	0.0003	0.0467	0.0144	0.0121	0.0526
	$\lambda_{\theta\phi}$	-0.0069	0.0050	0.0012	0.0105	0.0006	0.0029	0.0010	0.0118	0.0058	0.0051	0.0171
	$\lambda_\phi$	-0.0031	0.0036	0.0007	0.0073	0.0001	0.0002	0.0014	0.0065	0.0029	0.0036	0.0103
10 – 15	$\lambda_\theta$	-0.1929	0.0193	0.0030	0.0294	0.0108	0.0009	0.0009	0.0743	0.0177	0.0195	0.0825
	$\lambda_{\theta\phi}$	-0.0118	0.0084	0.0021	0.0179	0.0005	0.0006	0.0017	0.0073	0.0060	0.0087	0.0203
	$\lambda_\phi$	0.0080	0.0070	0.0012	0.0120	0.0006	0.0004	0.0017	0.0025	0.0002	0.0071	0.0124

Table B.7 The measured  $J/\psi$  polarization parameters, and their statistical errors and systematic uncertainties in different  $p_T$  bins in rapidity bin  $3.0 < y < 3.5$  in the HX frame. LL stat are the statistical errors returned by likelihood estimator, SB stat are the statistical errors coming from background subtraction, MCsys are the systematic errors coming from fluctuations of the normalization computed from Monte Carlo, SB/Specsys are the background subtraction/unkown  $J/\psi$  spectrum (binning) related systematic uncertainties, tzs sys are systematic uncertainties coming from  $\tau_S$  cut. Trk/AccSys are systematic uncertainties related to the difference of efficiency between MC and data, it combines the tracking and acceptance systematic uncertainties quadratically. BConSys are systematic uncertainties coming from contamination of  $J/\psi$  from  $b$  decay to our prompt sample. stat sum is the quadratically added statistical errors while sys sum is the quadratically added systematic uncertainties. Bin migration systematic uncertainties are so small that they are neglected.

$p_T(\text{GeV}/c)$	$\lambda_s$	fit	LL stat	SB stat	MCsys	SB sys	Specsys	tzs sys	Trk/AccSys	BConSys	stat sum	sys sum
2 – 3	$\lambda_\theta$	-0.2399	0.0049	0.0037	0.0060	0.0273	0.0067	0.0013	0.0719	0.0096	0.0061	0.0780
	$\lambda_{\theta\phi}$	-0.0462	0.0020	0.0016	0.0019	0.0150	0.0185	0.0004	0.0211	0.0031	0.0026	0.0321
	$\lambda_\phi$	0.0237	0.0018	0.0009	0.0020	0.0170	0.0067	0.0007	0.0175	0.0063	0.0021	0.0262
3 – 4	$\lambda_\theta$	-0.1959	0.0055	0.0016	0.0071	0.0186	0.0028	0.0004	0.0235	0.0119	0.0057	0.0332
	$\lambda_{\theta\phi}$	-0.0321	0.0026	0.0012	0.0039	0.0059	0.0032	0.0002	0.0140	0.0018	0.0029	0.0161
	$\lambda_\phi$	0.0091	0.0023	0.0010	0.0024	0.0033	0.0030	0.0006	0.0157	0.0075	0.0025	0.0181
4 – 5	$\lambda_\theta$	-0.1492	0.0068	0.0025	0.0108	0.0071	0.0031	0.0021	0.0038	0.0121	0.0072	0.0184
	$\lambda_{\theta\phi}$	-0.0252	0.0035	0.0013	0.0061	0.0009	0.0005	0.0002	0.0145	0.0053	0.0038	0.0167
	$\lambda_\phi$	-0.0059	0.0027	0.0008	0.0040	0.0003	0.0015	0.0005	0.0108	0.0083	0.0028	0.0143
5 – 7	$\lambda_\theta$	-0.1375	0.0070	0.0021	0.0126	0.0017	0.0024	0.0007	0.0207	0.0089	0.0073	0.0260
	$\lambda_{\theta\phi}$	0.0117	0.0035	0.0011	0.0053	0.0006	0.0023	0.0003	0.0068	0.0096	0.0052	0.0131
	$\lambda_\phi$	-0.0303	0.0026	0.0006	0.0039	0.0010	0.0026	0.0006	0.0103	0.0064	0.0027	0.0130
7 – 10	$\lambda_\theta$	-0.1675	0.0105	0.0023	0.0191	0.0067	0.0047	0.0014	0.0059	0.0126	0.0107	0.0250
	$\lambda_{\theta\phi}$	0.0070	0.0051	0.0011	0.0077	0.0035	0.0013	0.0003	0.0233	0.0104	0.0052	0.0269
	$\lambda_\phi$	-0.0150	0.0040	0.0007	0.0083	0.0020	0.0015	0.0011	0.0117	0.0051	0.0041	0.0155
10 – 15	$\lambda_\theta$	-0.1155	0.0206	0.0078	0.0359	0.0077	0.0055	0.0013	0.0758	0.0129	0.0220	0.0853
	$\lambda_{\theta\phi}$	0.0027	0.0098	0.0021	0.0144	0.0031	0.0005	0.0004	0.0165	0.0098	0.0100	0.0242
	$\lambda_\phi$	-0.0115	0.0085	0.0014	0.0158	0.0033	0.0001	0.0003	0.0082	0.0032	0.0086	0.0184

Table B.8 The measured  $J/\psi$  polarization parameters, and their statistical errors and systematic uncertainties in different  $p_T$  bins in rapidity bin  $3.5 < y < 4.0$  in the HX frame. LL stat are the statistical errors returned by likelihood estimator, SB stat are the statistical errors coming from background subtraction, MCsys are the systematic errors coming from fluctuations of the normalization computed from Monte Carlo, SB/Specsys are the background subtraction/unkown  $J/\psi$  spectrum (binning) related systematic uncertainties, tzs sys are systematic uncertainties coming from  $\tau_S$  cut. Trk/AccSys are systematic uncertainties related to the difference of efficiency between MC and data, it combines the tracking and acceptance systematic uncertainties quadratically. BConSys are systematic uncertainties coming from contamination of  $J/\psi$  from  $b$  decay to our prompt sample. stat sum is the quadratically added statistical errors while sys sum is the quadratically added systematic uncertainties. Bin migration systematic uncertainties are so small that they are neglected.

$p_T(\text{GeV}/c)$	$\lambda_s$	fit	LL stat	SB stat	MCsys	SB sys	Specsys	tzs sys	Trk/AccSys	BConSys	stat sum	sys sum
2 – 3	$\lambda_\theta$	-0.1836	0.0046	0.0025	0.0051	0.0069	0.0126	0.0008	0.0281	0.0123	0.0052	0.0342
	$\lambda_{\theta\phi}$	-0.0551	0.0023	0.0015	0.0031	0.0254	0.0214	0.0000	0.0625	0.0013	0.0027	0.0709
	$\lambda_\phi$	0.0185	0.0022	0.0012	0.0028	0.0047	0.0072	0.0003	0.0202	0.0056	0.0025	0.0229
3 – 4	$\lambda_\theta$	-0.1228	0.0056	0.0017	0.0077	0.0131	0.0053	0.0007	0.0323	0.0121	0.0059	0.0381
	$\lambda_{\theta\phi}$	-0.0281	0.0030	0.0014	0.0037	0.0085	0.0040	0.0002	0.0516	0.0019	0.0033	0.0526
	$\lambda_\phi$	0.0082	0.0029	0.0011	0.0034	0.0016	0.0018	0.0001	0.0249	0.0068	0.0031	0.0261
4 – 5	$\lambda_\theta$	-0.1138	0.0070	0.0015	0.0126	0.0064	0.0014	0.0010	0.0242	0.0138	0.0072	0.0313
	$\lambda_{\theta\phi}$	-0.0156	0.0041	0.0011	0.0094	0.0028	0.0003	0.0002	0.0191	0.0034	0.0043	0.0217
	$\lambda_\phi$	-0.0036	0.0034	0.0010	0.0061	0.0019	0.0010	0.0001	0.0314	0.0063	0.0036	0.0326
5 – 7	$\lambda_\theta$	-0.1322	0.0075	0.0022	0.0123	0.0007	0.0014	0.0006	0.0337	0.0109	0.0078	0.0375
	$\lambda_{\theta\phi}$	-0.0054	0.0043	0.0013	0.0060	0.0013	0.0007	0.0002	0.0175	0.0069	0.0069	0.0198
	$\lambda_\phi$	-0.0014	0.0034	0.0007	0.0060	0.0026	0.0014	0.0003	0.0301	0.0065	0.0034	0.0315
7 – 10	$\lambda_\theta$	-0.1078	0.0127	0.0025	0.0202	0.0084	0.0017	0.0009	0.0673	0.0111	0.0130	0.0717
	$\lambda_{\theta\phi}$	0.0152	0.0067	0.0014	0.0100	0.0030	0.0010	0.0001	0.0120	0.0097	0.0069	0.0186
	$\lambda_\phi$	-0.0070	0.0053	0.0009	0.0107	0.0023	0.0003	0.0002	0.0151	0.0053	0.0054	0.0194
10 – 15	$\lambda_\theta$	-0.1225	0.0265	0.0083	0.0416	0.0340	0.0069	0.0013	0.0486	0.0164	0.0277	0.0746
	$\lambda_{\theta\phi}$	0.0051	0.0133	0.0033	0.0241	0.0085	0.0007	0.0003	0.0085	0.0066	0.0137	0.0277
	$\lambda_\phi$	-0.0192	0.0115	0.0021	0.0162	0.0055	0.0022	0.0003	0.0064	0.0017	0.0117	0.0185

Table B.9 The measured  $J/\psi$  polarization parameters, and their statistical errors and systematic uncertainties in different  $p_T$  bins in rapidity bin  $4.0 < y < 4.5$  in the HX frame. LL stat are the statistical errors returned by likelihood estimator, SB stat are the statistical errors coming from background subtraction, MCsys are the systematic errors coming from fluctuations of the normalization computed from Monte Carlo, SB/Specsys are the background subtraction/unkown  $J/\psi$  spectrum (binning) related systematic uncertainties, tzs sys are systematic uncertainties coming from  $\tau_S$  cut. Trk/AccSys are systematic uncertainties related to the difference of efficiency between MC and data, it combines the tracking and acceptance systematic uncertainties quadratically. BConSys are systematic uncertainties coming from contamination of  $J/\psi$  from  $b$  decay to our prompt sample. stat sum is the quadratically added statistical errors while sys sum is the quadratically added systematic uncertainties. Bin migration systematic uncertainties are so small that they are neglected.

$p_T(\text{GeV}/c)$	$\lambda_s$	fit	LL stat	SB stat	MCsys	SB sys	Specsys	tzs sys	Trk/AccSys	BConSys	stat sum	sys sum
2 – 3	$\lambda_\theta$	-0.1180	0.0071	0.0027	0.0150	0.0099	0.0219	0.0002	0.0465	0.0100	0.0076	0.0554
	$\lambda_{\theta\phi}$	-0.0454	0.0052	0.0025	0.0107	0.0311	0.0071	0.0000	0.1003	0.0003	0.0058	0.1057
	$\lambda_\phi$	0.0043	0.0042	0.0015	0.0096	0.0055	0.0239	0.0001	0.0459	0.0033	0.0045	0.0530
3 – 4	$\lambda_\theta$	-0.0878	0.0093	0.0033	0.0125	0.0017	0.0096	0.0001	0.0462	0.0103	0.0099	0.0499
	$\lambda_{\theta\phi}$	-0.0222	0.0052	0.0015	0.0088	0.0165	0.0032	0.0000	0.0685	0.0010	0.0054	0.0711
	$\lambda_\phi$	-0.0326	0.0062	0.0016	0.0147	0.0044	0.0036	0.0000	0.0481	0.0024	0.0064	0.0506
4 – 5	$\lambda_\theta$	-0.0622	0.0119	0.0037	0.0218	0.0040	0.0111	0.0001	0.0998	0.0154	0.0124	0.1040
	$\lambda_{\theta\phi}$	0.0365	0.0067	0.0015	0.0124	0.0063	0.0049	0.0000	0.0423	0.0033	0.0069	0.0449
	$\lambda_\phi$	-0.0599	0.0076	0.0014	0.0126	0.0030	0.0073	0.0000	0.0471	0.0046	0.0078	0.0496
5 – 7	$\lambda_\theta$	-0.0908	0.0117	0.0047	0.0222	0.0132	0.0045	0.0001	0.0511	0.0121	0.0126	0.0587
	$\lambda_{\theta\phi}$	0.0501	0.0074	0.0017	0.0116	0.0008	0.0003	0.0000	0.0453	0.0035	0.0123	0.0470
	$\lambda_\phi$	-0.0150	0.0069	0.0016	0.0113	0.0029	0.0021	0.0001	0.0468	0.0021	0.0071	0.0483
7 – 10	$\lambda_\theta$	-0.1221	0.0192	0.0055	0.0286	0.0242	0.0005	0.0000	0.0678	0.0149	0.0199	0.0789
	$\lambda_{\theta\phi}$	0.0468	0.0119	0.0030	0.0243	0.0015	0.0002	0.0000	0.0520	0.0045	0.0123	0.0576
	$\lambda_\phi$	0.0079	0.0098	0.0016	0.0167	0.0049	0.0078	0.0000	0.0275	0.0033	0.0100	0.0336
10 – 15	$\lambda_\theta$	-0.1626	0.0420	0.0178	0.0447	0.0348	0.0431	0.0004	0.0745	0.0177	0.0456	0.1046
	$\lambda_{\theta\phi}$	0.0095	0.0237	0.0081	0.0297	0.0135	0.0003	0.0013	0.0532	0.0071	0.0251	0.0628
	$\lambda_\phi$	-0.0339	0.0205	0.0061	0.0354	0.0007	0.0071	0.0001	0.0082	0.0011	0.0214	0.0371

Table B.10 The measured  $J/\psi$  polarization parameters, and their statistical errors and systematic uncertainties in different  $p_T$  bins in rapidity bin  $2.0 < y < 2.5$  in the CS frame. LL stat are the statistical errors returned by likelihood estimator, SB stat are the statistical errors coming from background subtraction, MCsys are the systematic errors coming from fluctuations of the normalization computed from Monte Carlo, SB/Specsys are the background subtraction/unkown  $J/\psi$  spectrum (binning) related systematic uncertainties, tzs sys are systematic uncertainties coming from  $\tau_S$  cut. Trk/AccSys are systematic uncertainties related to the difference of efficiency between MC and data, it combines the tracking and acceptance systematic uncertainties quadratically. BConSys are systematic uncertainties coming from contamination of  $J/\psi$  from  $b$  decay to our prompt sample. stat sum is the quadratically added statistical errors while sys sum is the quadratically added systematic uncertainties. Bin migration systematic uncertainties are so small that they are neglected.

209

$p_T(\text{GeV}/c)$	$\lambda_s$	fit	LL stat	SB stat	MCsys	SB sys	Specsys	tzs sys	Trk/AccSys	BConSys	stat sum	sys sum
2 – 3	$\lambda_\theta$	-0.2656	0.0955	0.0701	0.1704	0.1828	0.1291	0.0038	0.0541	0.0286	0.1184	0.2879
	$\lambda_{\theta\phi}$	0.1184	0.0358	0.0263	0.0639	0.1220	0.0035	0.0001	0.0354	0.0042	0.0444	0.1423
	$\lambda_\phi$	-0.0203	0.0092	0.0067	0.0164	0.1181	0.0256	0.0004	0.0117	0.0053	0.0114	0.1227
3 – 4	$\lambda_\theta$	-0.1417	0.0744	0.0441	0.1217	0.1377	0.0247	0.0108	0.0444	0.0085	0.0864	0.1912
	$\lambda_{\theta\phi}$	0.1879	0.0360	0.0213	0.0589	0.0091	0.0148	0.0020	0.0380	0.0162	0.0418	0.0740
	$\lambda_\phi$	-0.0528	0.0122	0.0072	0.0200	0.0490	0.0010	0.0003	0.0177	0.0003	0.0142	0.0558
4 – 5	$\lambda_\theta$	-0.0432	0.0614	0.0293	0.0963	0.1210	0.0049	0.0008	0.0433	0.0048	0.0681	0.1608
	$\lambda_{\theta\phi}$	0.2262	0.0367	0.0175	0.0575	0.0120	0.0000	0.0001	0.0204	0.0199	0.0406	0.0653
	$\lambda_\phi$	-0.1003	0.0150	0.0072	0.0235	0.0216	0.0049	0.0009	0.0072	0.0011	0.0166	0.0331
5 – 7	$\lambda_\theta$	-0.0647	0.0351	0.0126	0.0538	0.0323	0.0231	0.0016	0.0452	0.0019	0.0373	0.0807
	$\lambda_{\theta\phi}$	0.1767	0.0268	0.0096	0.0411	0.0003	0.0075	0.0054	0.0214	0.0182	0.0276	0.0507
	$\lambda_\phi$	-0.0926	0.0133	0.0048	0.0204	0.0068	0.0009	0.0031	0.0252	0.0056	0.0141	0.0337
7 – 10	$\lambda_\theta$	0.1105	0.0305	0.0080	0.0457	0.0019	0.0184	0.0134	0.0321	0.0041	0.0315	0.0605
	$\lambda_{\theta\phi}$	0.0434	0.0267	0.0070	0.0401	0.0030	0.0244	0.0076	0.0315	0.0063	0.0276	0.0575
	$\lambda_\phi$	-0.1146	0.0172	0.0045	0.0257	0.0018	0.0167	0.0021	0.0415	0.0012	0.0177	0.0517
10 – 15	$\lambda_\theta$	0.0675	0.0358	0.0080	0.0527	0.0033	0.0133	0.0324	0.0046	0.0255	0.0367	0.0685
	$\lambda_{\theta\phi}$	0.1131	0.0326	0.0073	0.0479	0.0071	0.0149	0.0026	0.0138	0.0135	0.0334	0.0543
	$\lambda_\phi$	-0.1027	0.0254	0.0057	0.0373	0.0036	0.0141	0.0084	0.0124	0.0080	0.0260	0.0435

Table B.11 The measured  $J/\psi$  polarization parameters, and their statistical errors and systematic uncertainties in different  $p_T$  bins in rapidity bin  $2.5 < y < 3.0$  in the CS frame. LL stat are the statistical errors returned by likelihood estimator, SB stat are the statistical errors coming from background subtraction, MCsys are the systematic errors coming from fluctuations of the normalization computed from Monte Carlo, SB/Specsys are the background subtraction/unkown  $J/\psi$  spectrum (binning) related systematic uncertainties, tzs sys are systematic uncertainties coming from  $\tau_S$  cut. Trk/AccSys are systematic uncertainties related to the difference of efficiency between MC and data, it combines the tracking and acceptance systematic uncertainties quadratically. BConSys are systematic uncertainties coming from contamination of  $J/\psi$  from  $b$  decay to our prompt sample. stat sum is the quadratically added statistical errors while sys sum is the quadratically added systematic uncertainties. Bin migration systematic uncertainties are so small that they are neglected.

210

$p_T(\text{GeV}/c)$	$\lambda_s$	fit	LL stat	SB stat	MCsys	SB sys	Specsys	tzs sys	Trk/AccSys	BConSys	stat sum	sys sum
2 – 3	$\lambda_\theta$	-0.1869	0.0071	0.0052	0.0139	0.0894	0.0133	0.0023	0.0420	0.0029	0.0088	0.1007
	$\lambda_{\theta\phi}$	0.1344	0.0050	0.0037	0.0098	0.0174	0.0289	0.0034	0.0214	0.0076	0.0062	0.0420
	$\lambda_\phi$	0.0039	0.0023	0.0017	0.0045	0.0521	0.0182	0.0009	0.0102	0.0037	0.0029	0.0565
3 – 4	$\lambda_\theta$	-0.0866	0.0057	0.0033	0.0101	0.0504	0.0146	0.0015	0.0344	0.0075	0.0066	0.0640
	$\lambda_{\theta\phi}$	0.1303	0.0055	0.0032	0.0097	0.0003	0.0179	0.0026	0.0409	0.0098	0.0063	0.0468
	$\lambda_\phi$	-0.0467	0.0037	0.0022	0.0066	0.0220	0.0065	0.0002	0.0204	0.0024	0.0043	0.0315
4 – 5	$\lambda_\theta$	0.0006	0.0061	0.0026	0.0103	0.0169	0.0110	0.0032	0.0260	0.0037	0.0066	0.0348
	$\lambda_{\theta\phi}$	0.1086	0.0057	0.0024	0.0097	0.0014	0.0132	0.0038	0.0424	0.0121	0.0062	0.0472
	$\lambda_\phi$	-0.0683	0.0050	0.0021	0.0085	0.0070	0.0053	0.0003	0.0269	0.0032	0.0055	0.0297
5 – 7	$\lambda_\theta$	0.0216	0.0057	0.0017	0.0091	0.0036	0.0080	0.0051	0.0225	0.0062	0.0059	0.0270
	$\lambda_{\theta\phi}$	0.0752	0.0047	0.0014	0.0075	0.0032	0.0083	0.0028	0.0204	0.0114	0.0059	0.0263
	$\lambda_\phi$	-0.0767	0.0050	0.0015	0.0081	0.0011	0.0057	0.0001	0.0213	0.0026	0.0052	0.0236
7 – 10	$\lambda_\theta$	0.0466	0.0083	0.0020	0.0129	0.0007	0.0036	0.0064	0.0153	0.0140	0.0085	0.0255
	$\lambda_{\theta\phi}$	0.0640	0.0058	0.0014	0.0089	0.0028	0.0015	0.0032	0.0108	0.0087	0.0059	0.0171
	$\lambda_\phi$	-0.0797	0.0071	0.0017	0.0109	0.0015	0.0043	0.0007	0.0068	0.0032	0.0073	0.0140
10 – 15	$\lambda_\theta$	0.0879	0.0160	0.0035	0.0239	0.0073	0.0017	0.0069	0.0131	0.0184	0.0164	0.0344
	$\lambda_{\theta\phi}$	0.0634	0.0096	0.0021	0.0144	0.0031	0.0006	0.0017	0.0086	0.0073	0.0098	0.0186
	$\lambda_\phi$	-0.0905	0.0122	0.0027	0.0182	0.0062	0.0015	0.0031	0.0189	0.0058	0.0125	0.0278

Table B.12 The measured  $J/\psi$  polarization parameters, and their statistical errors and systematic uncertainties in different  $p_T$  bins in rapidity bin  $3.0 < y < 3.5$  in the CS frame. LL stat are the statistical errors returned by likelihood estimator, SB stat are the statistical errors coming from background subtraction, MCsys are the systematic errors coming from fluctuations of the normalization computed from Monte Carlo, SB/Specsys are the background subtraction/unkown  $J/\psi$  spectrum (binning) related systematic uncertainties, tzs sys are systematic uncertainties coming from  $\tau_S$  cut. Trk/AccSys are systematic uncertainties related to the difference of efficiency between MC and data, it combines the tracking and acceptance systematic uncertainties quadratically. BConSys are systematic uncertainties coming from contamination of  $J/\psi$  from  $b$  decay to our prompt sample. stat sum is the quadratically added statistical errors while sys sum is the quadratically added systematic uncertainties. Bin migration systematic uncertainties are so small that they are neglected.

211

$p_T(\text{GeV}/c)$	$\lambda_s$	fit	LL stat	SB stat	MCsys	SB sys	Specsys	tzs sys	Trk/AccSys	BConSys	stat sum	sys sum
2 – 3	$\lambda_\theta$	-0.1393	0.0036	0.0023	0.0072	0.0239	0.0298	0.0012	0.0404	0.0087	0.0043	0.0568
	$\lambda_{\theta\phi}$	0.1228	0.0031	0.0019	0.0062	0.0083	0.0017	0.0021	0.0215	0.0090	0.0036	0.0256
	$\lambda_\phi$	-0.0068	0.0018	0.0012	0.0037	0.0278	0.0152	0.0004	0.0146	0.0034	0.0022	0.0353
3 – 4	$\lambda_\theta$	-0.0619	0.0036	0.0017	0.0066	0.0099	0.0067	0.0011	0.0312	0.0059	0.0040	0.0345
	$\lambda_{\theta\phi}$	0.1155	0.0034	0.0016	0.0062	0.0042	0.0019	0.0019	0.0085	0.0105	0.0038	0.0157
	$\lambda_\phi$	-0.0373	0.0031	0.0015	0.0056	0.0127	0.0026	0.0003	0.0054	0.0036	0.0034	0.0156
4 – 5	$\lambda_\theta$	-0.0332	0.0046	0.0016	0.0079	0.0031	0.0000	0.0010	0.0238	0.0065	0.0049	0.0261
	$\lambda_{\theta\phi}$	0.0786	0.0037	0.0013	0.0064	0.0011	0.0007	0.0031	0.0061	0.0116	0.0040	0.0149
	$\lambda_\phi$	-0.0457	0.0043	0.0015	0.0073	0.0038	0.0006	0.0011	0.0057	0.0053	0.0046	0.0114
5 – 7	$\lambda_\theta$	0.0046	0.0051	0.0014	0.0083	0.0005	0.0024	0.0013	0.0173	0.0039	0.0053	0.0198
	$\lambda_{\theta\phi}$	0.0387	0.0035	0.0009	0.0057	0.0008	0.0010	0.0018	0.0065	0.0113	0.0052	0.0144
	$\lambda_\phi$	-0.0806	0.0046	0.0012	0.0075	0.0015	0.0035	0.0000	0.0056	0.0059	0.0047	0.0117
7 – 10	$\lambda_\theta$	0.0432	0.0086	0.0019	0.0133	0.0032	0.0027	0.0023	0.0159	0.0079	0.0088	0.0227
	$\lambda_{\theta\phi}$	0.0452	0.0051	0.0012	0.0079	0.0010	0.0002	0.0009	0.0112	0.0128	0.0052	0.0188
	$\lambda_\phi$	-0.0896	0.0069	0.0016	0.0108	0.0060	0.0030	0.0008	0.0142	0.0023	0.0071	0.0192
10 – 15	$\lambda_\theta$	0.0341	0.0174	0.0039	0.0259	0.0010	0.0081	0.0009	0.0388	0.0079	0.0178	0.0480
	$\lambda_{\theta\phi}$	0.0226	0.0098	0.0022	0.0146	0.0017	0.0014	0.0009	0.0198	0.0117	0.0100	0.0273
	$\lambda_\phi$	-0.0637	0.0128	0.0029	0.0191	0.0067	0.0054	0.0008	0.0403	0.0016	0.0131	0.0454

Table B.13 The measured  $J/\psi$  polarization parameters, and their statistical errors and systematic uncertainties in different  $p_T$  bins in rapidity bin  $3.5 < y < 4.0$  in the CS frame. LL stat are the statistical errors returned by likelihood estimator, SB stat are the statistical errors coming from background subtraction, MCsys are the systematic errors coming from fluctuations of the normalization computed from Monte Carlo, SB/Specsys are the background subtraction/unkown  $J/\psi$  spectrum (binning) related systematic uncertainties, tzs sys are systematic uncertainties coming from  $\tau_S$  cut. Trk/AccSys are systematic uncertainties related to the difference of efficiency between MC and data, it combines the tracking and acceptance systematic uncertainties quadratically. BConSys are systematic uncertainties coming from contamination of  $J/\psi$  from  $b$  decay to our prompt sample. stat sum is the quadratically added statistical errors while sys sum is the quadratically added systematic uncertainties. Bin migration systematic uncertainties are so small that they are neglected.

212

$p_T(\text{GeV}/c)$	$\lambda_s$	fit	LL stat	SB stat	MCsys	SB sys	Specsys	tzs sys	Trk/AccSys	BConSys	stat sum	sys sum
2 – 3	$\lambda_\theta$	-0.1171	0.0041	0.0020	0.0083	0.0320	0.0307	0.0001	0.0779	0.0028	0.0046	0.0901
	$\lambda_{\theta\phi}$	0.0968	0.0030	0.0015	0.0060	0.0023	0.0064	0.0012	0.0136	0.0099	0.0033	0.0192
	$\lambda_\phi$	-0.0026	0.0020	0.0010	0.0040	0.0179	0.0137	0.0005	0.0028	0.0018	0.0022	0.0232
3 – 4	$\lambda_\theta$	-0.0343	0.0044	0.0016	0.0079	0.0116	0.0045	0.0000	0.0603	0.0049	0.0046	0.0623
	$\lambda_{\theta\phi}$	0.0745	0.0036	0.0013	0.0065	0.0006	0.0005	0.0012	0.0144	0.0098	0.0038	0.0187
	$\lambda_\phi$	-0.0213	0.0033	0.0013	0.0060	0.0067	0.0030	0.0001	0.0046	0.0016	0.0036	0.0107
4 – 5	$\lambda_\theta$	-0.0083	0.0056	0.0017	0.0095	0.0056	0.0030	0.0001	0.0462	0.0067	0.0059	0.0480
	$\lambda_{\theta\phi}$	0.0570	0.0042	0.0012	0.0070	0.0007	0.0018	0.0011	0.0153	0.0105	0.0043	0.0199
	$\lambda_\phi$	-0.0400	0.0047	0.0014	0.0080	0.0020	0.0014	0.0003	0.0071	0.0002	0.0049	0.0110
5 – 7	$\lambda_\theta$	0.0298	0.0063	0.0016	0.0101	0.0012	0.0034	0.0001	0.0292	0.0061	0.0065	0.0317
	$\lambda_{\theta\phi}$	0.0544	0.0041	0.0010	0.0066	0.0012	0.0011	0.0005	0.0159	0.0116	0.0066	0.0208
	$\lambda_\phi$	-0.0581	0.0052	0.0013	0.0083	0.0027	0.0028	0.0001	0.0080	0.0020	0.0054	0.0123
7 – 10	$\lambda_\theta$	0.0461	0.0108	0.0025	0.0162	0.0030	0.0036	0.0004	0.0182	0.0065	0.0111	0.0256
	$\lambda_{\theta\phi}$	0.0178	0.0064	0.0015	0.0096	0.0004	0.0013	0.0009	0.0131	0.0136	0.0066	0.0212
	$\lambda_\phi$	-0.0607	0.0083	0.0019	0.0125	0.0066	0.0025	0.0004	0.0339	0.0001	0.0085	0.0369
10 – 15	$\lambda_\theta$	0.0295	0.0232	0.0059	0.0329	0.0136	0.0113	0.0005	0.0395	0.0102	0.0239	0.0553
	$\lambda_{\theta\phi}$	0.0192	0.0131	0.0033	0.0186	0.0015	0.0001	0.0005	0.0105	0.0106	0.0135	0.0239
	$\lambda_\phi$	-0.0728	0.0168	0.0043	0.0238	0.0236	0.0057	0.0001	0.0492	0.0066	0.0173	0.0602

Table B.14 The measured  $J/\psi$  polarization parameters, and their statistical errors and systematic uncertainties in different  $p_T$  bins in rapidity bin  $4.0 < y < 4.5$  in the CS frame. LL stat are the statistical errors returned by likelihood estimator, SB stat are the statistical errors coming from background subtraction, MCsys are the systematic errors coming from fluctuations of the normalization computed from Monte Carlo, SB/Specsys are the background subtraction/unkown  $J/\psi$  spectrum (binning) related systematic uncertainties, tzs sys are systematic uncertainties coming from  $\tau_S$  cut. Trk/AccSys are systematic uncertainties related to the difference of efficiency between MC and data, it combines the tracking and acceptance systematic uncertainties quadratically. BConSys are systematic uncertainties coming from contamination of  $J/\psi$  from  $b$  decay to our prompt sample. stat sum is the quadratically added statistical errors while sys sum is the quadratically added systematic uncertainties. Bin migration systematic uncertainties are so small that they are neglected.

213

$p_T(\text{GeV}/c)$	$\lambda_s$	fit	LL stat	SB stat	MCsys	SB sys	Specsys	tzs sys	Trk/AccSys	BConSys	stat sum	sys sum
2 – 3	$\lambda_\theta$	-0.0863	0.0104	0.0037	0.0209	0.0444	0.0004	0.0000	0.0668	0.0035	0.0110	0.0829
	$\lambda_{\theta\phi}$	0.0534	0.0044	0.0016	0.0088	0.0092	0.0092	0.0003	0.0249	0.0067	0.0047	0.0302
	$\lambda_\phi$	-0.0102	0.0028	0.0010	0.0056	0.0083	0.0160	0.0002	0.0093	0.0003	0.0030	0.0210
3 – 4	$\lambda_\theta$	-0.0547	0.0107	0.0034	0.0191	0.0195	0.0076	0.0000	0.0651	0.0029	0.0113	0.0711
	$\lambda_{\theta\phi}$	0.0301	0.0060	0.0019	0.0107	0.0080	0.0020	0.0001	0.0232	0.0051	0.0063	0.0273
	$\lambda_\phi$	-0.0462	0.0047	0.0015	0.0083	0.0023	0.0010	0.0001	0.0079	0.0026	0.0049	0.0120
4 – 5	$\lambda_\theta$	-0.0199	0.0119	0.0035	0.0196	0.0061	0.0217	0.0001	0.0762	0.0082	0.0124	0.0823
	$\lambda_{\theta\phi}$	-0.0267	0.0077	0.0022	0.0126	0.0048	0.0046	0.0003	0.0163	0.0088	0.0080	0.0234
	$\lambda_\phi$	-0.0756	0.0068	0.0020	0.0112	0.0028	0.0028	0.0001	0.0207	0.0031	0.0071	0.0240
5 – 7	$\lambda_\theta$	0.0596	0.0118	0.0033	0.0181	0.0028	0.0163	0.0000	0.0370	0.0074	0.0122	0.0450
	$\lambda_{\theta\phi}$	-0.0189	0.0079	0.0022	0.0121	0.0035	0.0046	0.0002	0.0234	0.0069	0.0128	0.0279
	$\lambda_\phi$	-0.0674	0.0077	0.0021	0.0118	0.0096	0.0012	0.0000	0.0137	0.0044	0.0080	0.0210
7 – 10	$\lambda_\theta$	0.1034	0.0191	0.0055	0.0271	0.0065	0.0153	0.0000	0.0350	0.0116	0.0198	0.0487
	$\lambda_{\theta\phi}$	-0.0123	0.0122	0.0036	0.0174	0.0047	0.0007	0.0000	0.0225	0.0082	0.0128	0.0300
	$\lambda_\phi$	-0.0696	0.0131	0.0038	0.0186	0.0173	0.0005	0.0000	0.0483	0.0060	0.0136	0.0549
10 – 15	$\lambda_\theta$	0.0369	0.0399	0.0142	0.0506	0.0218	0.0469	0.0001	0.0204	0.0079	0.0423	0.0756
	$\lambda_{\theta\phi}$	0.0103	0.0240	0.0086	0.0305	0.0076	0.0039	0.0015	0.0241	0.0109	0.0255	0.0413
	$\lambda_\phi$	-0.1086	0.0284	0.0101	0.0360	0.0218	0.0154	0.0011	0.0258	0.0050	0.0301	0.0520

Table B.15 The measured  $\psi(2S)$  polarization parameters, and their statistical errors and systematic uncertainties in different  $p_T$  bins in rapidity bin  $2.0 < y < 2.5$  in the HX frame. LL stat are the statistical errors returned by likelihood estimators, MCsys are the systematic errors coming from fluctuations of the normalization computed from Monte Carlo, SBsys are the background subtraction systematic uncertainties, Specsyst are systematic uncertainties due to different  $\psi(2S)$  spectrum between data and Monte Carlo, PIDsys are systematic uncertainties coming from the muon PID cut, Trk/AccSys are systematic uncertainties related to the difference between MC and data, it combines the tracking and Monte Carlo efficiency systematic uncertainty quadratically. BConSys are systematic uncertainties coming from contamination of  $\psi(2S)$  from  $b$  decay to our prompt sample. sys sum is the quadratically added systematic uncertainties.

$p_T(\text{GeV}/c)$	$\lambda s$	fit	LL stat	MCsys	SBsys	Specsys	PIDsys	Trk/AccSys	BConSys	sys sum
3 – 4	$\lambda_\theta$	-0.2559	0.6238	0.2250	0.0513	0.1283	0.0426	0.0515	0.0775	0.2831
	$\lambda_{\theta\phi}$	-0.1156	0.4333	0.1380	0.0614	0.1387	0.0303	0.0249	0.0947	0.2292
	$\lambda_\phi$	0.1072	0.1488	0.0461	0.0159	0.0526	0.0009	0.0214	0.0323	0.0816
4 – 5	$\lambda_\theta$	-0.2927	0.3608	0.1210	0.0040	0.0059	0.0654	0.0392	0.0331	0.1470
	$\lambda_{\theta\phi}$	-0.3054	0.2748	0.1044	0.0094	0.0118	0.0399	0.0186	0.0915	0.1464
	$\lambda_\phi$	0.0205	0.1041	0.0440	0.0082	0.0071	0.0101	0.0213	0.0435	0.0670
5 – 7	$\lambda_\theta$	0.0111	0.2142	0.0864	0.0089	0.0269	0.0905	0.2325	0.0108	0.2658
	$\lambda_{\theta\phi}$	-0.1710	0.1649	0.0588	0.0146	0.0324	0.0133	0.0429	0.0147	0.0833
	$\lambda_\phi$	0.0115	0.0681	0.0285	0.0065	0.0191	0.0013	0.0173	0.0012	0.0391
7 – 10	$\lambda_\theta$	0.1442	0.1724	0.0771	0.0041	0.0196	0.0409	0.2156	0.0359	0.2362
	$\lambda_{\theta\phi}$	0.0917	0.1203	0.0520	0.0059	0.0539	0.0040	0.0108	0.0203	0.0787
	$\lambda_\phi$	0.0321	0.0508	0.0234	0.0015	0.0300	0.0010	0.0164	0.0001	0.0415
10 – 15	$\lambda_\theta$	-0.1276	0.1927	0.0774	0.0130	0.0950	0.0361	0.0175	0.0428	0.1365
	$\lambda_{\theta\phi}$	-0.0568	0.1182	0.0476	0.0027	0.0157	0.0177	0.0215	0.0012	0.0574
	$\lambda_\phi$	-0.0211	0.0581	0.0224	0.0017	0.0056	0.0042	0.0138	0.0001	0.0273

Table B.16 The measured  $\psi(2S)$  polarization parameters, and their statistical errors and systematic uncertainties in different  $p_T$  bins in rapidity bin  $2.5 < y < 3.0$  in the HX frame. LL stat are the statistical errors returned by likelihood estimators, MCsys are the systematic errors coming from fluctuations of the normalization computed from Monte Carlo, SBsys are the background subtraction systematic uncertainties, Specsyst are systematic uncertainties due to different  $\psi(2S)$  spectrum between data and Monte Carlo, PIDsys are systematic uncertainties coming from the muon PID cut, Trk/AccSys are systematic uncertainties related to the difference between MC and data, it combines the tracking and Monte Carlo efficiency systematic uncertainty quadratically. BConSys are systematic uncertainties coming from contamination of  $\psi(2S)$  from  $b$  decay to our prompt sample. sys sum is the quadratically added systematic uncertainties.

$p_T(\text{GeV}/c)$	$\lambda s$	fit	LL stat	MCsys	SBsys	Specsys	PIDsys	Trk/AccSys	BConSys	sys sum
3 – 4	$\lambda_\theta$	-0.0505	0.1012	0.0442	0.0039	0.0345	0.0278	0.1732	0.0101	0.1845
	$\lambda_{\theta\phi}$	-0.1004	0.0423	0.0147	0.0247	0.0116	0.0086	0.0226	0.0096	0.0404
	$\lambda_\phi$	0.0727	0.0311	0.0112	0.0079	0.0039	0.0002	0.0098	0.0013	0.0173
4 – 5	$\lambda_\theta$	0.0957	0.0694	0.0278	0.0022	0.0662	0.0270	0.1650	0.0348	0.1853
	$\lambda_{\theta\phi}$	-0.1253	0.0292	0.0125	0.0029	0.0292	0.0061	0.0249	0.0042	0.0411
	$\lambda_\phi$	0.0357	0.0205	0.0090	0.0083	0.0053	0.0004	0.0104	0.0047	0.0176
5 – 7	$\lambda_\theta$	0.0859	0.0499	0.0232	0.0022	0.0403	0.0262	0.0859	0.0242	0.1040
	$\lambda_{\theta\phi}$	-0.0656	0.0224	0.0120	0.0018	0.0247	0.0035	0.0172	0.0013	0.0326
	$\lambda_\phi$	0.0264	0.0141	0.0057	0.0023	0.0006	0.0000	0.0116	0.0031	0.0136
7 – 10	$\lambda_\theta$	-0.1360	0.0500	0.0256	0.0052	0.0032	0.0101	0.0520	0.0118	0.0603
	$\lambda_{\theta\phi}$	-0.0546	0.0228	0.0122	0.0022	0.0140	0.0018	0.0143	0.0017	0.0236
	$\lambda_\phi$	0.0180	0.0153	0.0069	0.0014	0.0030	0.0026	0.0080	0.0068	0.0132
10 – 15	$\lambda_\theta$	-0.1820	0.0727	0.0285	0.0061	0.0027	0.0048	0.0659	0.0279	0.0775
	$\lambda_{\theta\phi}$	0.0071	0.0327	0.0136	0.0032	0.0103	0.0023	0.0078	0.0014	0.0192
	$\lambda_\phi$	-0.0159	0.0250	0.0120	0.0008	0.0025	0.0002	0.0029	0.0083	0.0151

Table B.17 The measured  $\psi(2S)$  polarization parameters, and their statistical errors and systematic uncertainties in different  $p_T$  bins in rapidity bin  $3.0 < y < 3.5$  in the HX frame. LL stat are the statistical errors returned by likelihood estimators, MCsys are the systematic errors coming from fluctuations of the normalization computed from Monte Carlo, SBsys are the background subtraction systematic uncertainties, Specsyst are systematic uncertainties due to different  $\psi(2S)$  spectrum between data and Monte Carlo, PIDsys are systematic uncertainties coming from the muon PID cut, Trk/AccSys are systematic uncertainties related to the difference between MC and data, it combines the tracking and Monte Carlo efficiency systematic uncertainty quadratically. BConSys are systematic uncertainties coming from contamination of  $\psi(2S)$  from  $b$  decay to our prompt sample. sys sum is the quadratically added systematic uncertainties.

$p_T(\text{GeV}/c)$	$\lambda s$	fit	LL stat	MCsys	SBsys	Specsys	PIDsys	Trk/AccSys	BConSys	sys sum
3 – 4	$\lambda_\theta$	-0.1333	0.0620	0.0264	0.0038	0.0093	0.0126	0.0088	0.0240	0.0401
	$\lambda_{\theta\phi}$	-0.0804	0.0324	0.0122	0.0076	0.0000	0.0028	0.0163	0.0068	0.0230
	$\lambda_\phi$	-0.0019	0.0303	0.0115	0.0059	0.0016	0.0012	0.0102	0.0183	0.0247
4 – 5	$\lambda_\theta$	-0.0910	0.0419	0.0175	0.0064	0.0097	0.0123	0.0102	0.0314	0.0410
	$\lambda_{\theta\phi}$	-0.0639	0.0237	0.0109	0.0021	0.0007	0.0010	0.0205	0.0025	0.0235
	$\lambda_\phi$	-0.0173	0.0196	0.0101	0.0043	0.0015	0.0028	0.0112	0.0144	0.0215
5 – 7	$\lambda_\theta$	-0.0885	0.0340	0.0143	0.0056	0.0113	0.0048	0.0178	0.0146	0.0303
	$\lambda_{\theta\phi}$	-0.0152	0.0194	0.0101	0.0028	0.0037	0.0011	0.0103	0.0061	0.0164
	$\lambda_\phi$	-0.0368	0.0141	0.0058	0.0034	0.0052	0.0012	0.0101	0.0113	0.0174
7 – 10	$\lambda_\theta$	-0.0678	0.0434	0.0214	0.0072	0.0009	0.0017	0.0114	0.0135	0.0287
	$\lambda_{\theta\phi}$	-0.0527	0.0226	0.0113	0.0021	0.0073	0.0005	0.0217	0.0003	0.0256
	$\lambda_\phi$	-0.0366	0.0163	0.0095	0.0013	0.0047	0.0005	0.0127	0.0062	0.0177
10 – 15	$\lambda_\theta$	-0.1571	0.0736	0.0287	0.0171	0.0188	0.0063	0.0721	0.0072	0.0822
	$\lambda_{\theta\phi}$	-0.0007	0.0361	0.0159	0.0045	0.0122	0.0008	0.0145	0.0199	0.0321
	$\lambda_\phi$	-0.0470	0.0291	0.0119	0.0019	0.0022	0.0009	0.0106	0.0015	0.0163

Table B.18 The measured  $\psi(2S)$  polarization parameters, and their statistical errors and systematic uncertainties in different  $p_T$  bins in rapidity bin  $3.5 < y < 4.0$  in the HX frame. LL stat are the statistical errors returned by likelihood estimators, MCsys are the systematic errors coming from fluctuations of the normalization computed from Monte Carlo, SBsys are the background subtraction systematic uncertainties, Specsyst are systematic uncertainties due to different  $\psi(2S)$  spectrum between data and Monte Carlo, PIDsys are systematic uncertainties coming from the muon PID cut, Trk/AccSys are systematic uncertainties related to the difference between MC and data, it combines the tracking and Monte Carlo efficiency systematic uncertainty quadratically. BConSys are systematic uncertainties coming from contamination of  $\psi(2S)$  from  $b$  decay to our prompt sample. sys sum is the quadratically added systematic uncertainties.

$p_T(\text{GeV}/c)$	$\lambda s$	fit	LL stat	MCsys	SBsys	Specsys	PIDsys	Trk/AccSys	BConSys	sys sum
3 – 4	$\lambda_\theta$	-0.0046	0.0624	0.0267	0.0192	0.0052	0.0055	0.0245	0.0235	0.0478
	$\lambda_{\theta\phi}$	-0.1072	0.0344	0.0136	0.0094	0.0028	0.0021	0.0450	0.0009	0.0481
	$\lambda_\phi$	-0.0463	0.0346	0.0162	0.0056	0.0045	0.0056	0.0335	0.0132	0.0406
4 – 5	$\lambda_\theta$	0.0035	0.0431	0.0278	0.0116	0.0066	0.0023	0.0231	0.0143	0.0412
	$\lambda_{\theta\phi}$	-0.0506	0.0259	0.0125	0.0057	0.0022	0.0011	0.0215	0.0036	0.0259
	$\lambda_\phi$	-0.0936	0.0233	0.0140	0.0052	0.0058	0.0046	0.0442	0.0054	0.0476
5 – 7	$\lambda_\theta$	-0.0467	0.0367	0.0219	0.0076	0.0047	0.0004	0.0252	0.0069	0.0353
	$\lambda_{\theta\phi}$	-0.0090	0.0225	0.0113	0.0023	0.0021	0.0022	0.0229	0.0062	0.0266
	$\lambda_\phi$	-0.0568	0.0170	0.0065	0.0048	0.0070	0.0015	0.0445	0.0086	0.0466
7 – 10	$\lambda_\theta$	-0.1532	0.0496	0.0279	0.0150	0.0001	0.0052	0.0579	0.0179	0.0686
	$\lambda_{\theta\phi}$	0.0338	0.0278	0.0140	0.0033	0.0101	0.0028	0.0164	0.0027	0.0244
	$\lambda_\phi$	-0.0568	0.0208	0.0126	0.0035	0.0067	0.0018	0.0207	0.0091	0.0270
10 – 15	$\lambda_\theta$	-0.2369	0.0952	0.0548	0.0232	0.0050	0.0026	0.0472	0.0048	0.0763
	$\lambda_{\theta\phi}$	0.0730	0.0467	0.0192	0.0036	0.0104	0.0013	0.0067	0.0256	0.0345
	$\lambda_\phi$	-0.0609	0.0393	0.0175	0.0012	0.0038	0.0012	0.0089	0.0134	0.0241

Table B.19 The measured  $\psi(2S)$  polarization parameters, and their statistical errors and systematic uncertainties in different  $p_T$  bins in rapidity bin  $4.0 < y < 4.5$  in the HX frame. LL stat are the statistical errors returned by likelihood estimators, MCsys are the systematic errors coming from fluctuations of the normalization computed from Monte Carlo, SBsys are the background subtraction systematic uncertainties, Specsyst are systematic uncertainties due to different  $\psi(2S)$  spectrum between data and Monte Carlo, PIDsys are systematic uncertainties coming from the muon PID cut, Trk/AccSys are systematic uncertainties related to the difference between MC and data, it combines the tracking and Monte Carlo efficiency systematic uncertainty quadratically. BConSys are systematic uncertainties coming from contamination of  $\psi(2S)$  from  $b$  decay to our prompt sample. sys sum is the quadratically added systematic uncertainties.

$p_T(\text{GeV}/c)$	$\lambda s$	fit	LL stat	MCsys	SBsys	Specsys	PIDsys	Trk/AccSys	BConSys	sys sum
3 – 4	$\lambda_\theta$	-0.0657	0.0940	0.0505	0.0370	0.0024	0.0026	0.0791	0.0057	0.1011
	$\lambda_{\theta\phi}$	-0.0534	0.0549	0.0319	0.0303	0.0124	0.0030	0.0787	0.0103	0.0916
	$\lambda_\phi$	-0.1047	0.0643	0.0373	0.0209	0.0078	0.0009	0.0571	0.0080	0.0722
4 – 5	$\lambda_\theta$	-0.0261	0.0703	0.0616	0.0218	0.0141	0.0010	0.1041	0.0011	0.1237
	$\lambda_{\theta\phi}$	-0.0489	0.0389	0.0214	0.0124	0.0196	0.0041	0.0399	0.0051	0.0513
	$\lambda_\phi$	-0.0577	0.0454	0.0238	0.0098	0.0182	0.0008	0.0630	0.0075	0.0708
5 – 7	$\lambda_\theta$	-0.0825	0.0592	0.0390	0.0198	0.0046	0.0059	0.0734	0.0382	0.0939
	$\lambda_{\theta\phi}$	-0.0002	0.0361	0.0207	0.0024	0.0033	0.0021	0.0673	0.0048	0.0707
	$\lambda_\phi$	-0.0715	0.0366	0.0172	0.0043	0.0093	0.0001	0.0560	0.0103	0.0604
7 – 10	$\lambda_\theta$	0.0102	0.0887	0.0545	0.0311	0.0036	0.0070	0.0738	0.0516	0.1100
	$\lambda_{\theta\phi}$	0.0332	0.0543	0.0284	0.0073	0.0049	0.0030	0.0587	0.0041	0.0660
	$\lambda_\phi$	-0.0960	0.0448	0.0245	0.0030	0.0115	0.0004	0.0346	0.0046	0.0443
10 – 15	$\lambda_\theta$	-0.5561	0.1348	0.0763	0.1379	0.0145	0.0030	0.0485	0.0797	0.1837
	$\lambda_{\theta\phi}$	0.1662	0.0866	0.0384	0.0224	0.0001	0.0027	0.0410	0.0025	0.0606
	$\lambda_\phi$	-0.0378	0.0777	0.0406	0.0063	0.0138	0.0013	0.0074	0.0022	0.0441

Table B.20 The measured  $\psi(2S)$  polarization parameters, and their statistical errors and systematic uncertainties in different  $p_T$  bins in rapidity bin  $2.0 < y < 2.5$  in the CS frame. LL stat are the statistical errors returned by likelihood estimators, MCsys are the systematic errors coming from fluctuations of the normalization computed from Monte Carlo, SBsys are the background subtraction systematic uncertainties, Specsyst are systematic uncertainties due to different  $\psi(2S)$  spectrum between data and Monte Carlo, PIDsys are systematic uncertainties coming from the muon PID cut, Trk/AccSys are systematic uncertainties related to the difference between MC and data, it combines the tracking and Monte Carlo efficiency systematic uncertainty quadratically. BConSys are systematic uncertainties coming from contamination of  $\psi(2S)$  from  $b$  decay to our prompt sample. sys sum is the quadratically added systematic uncertainties.

$p_T(\text{GeV}/c)$	$\lambda s$	fit	LL stat	MCsys	SBsys	Specsys	PIDsys	Trk/AccSys	BConSys	sys sum
3 – 4	$\lambda_\theta$	-0.1576	0.8780	0.2972	0.1580	0.2727	0.0378	0.0510	0.1747	0.4714
	$\lambda_{\theta\phi}$	0.1828	0.3010	0.1063	0.0186	0.0470	0.0247	0.0383	0.0237	0.1284
	$\lambda_\phi$	0.0756	0.0849	0.0345	0.0265	0.0093	0.0236	0.0057	0.0005	0.0506
4 – 5	$\lambda_\theta$	-0.3798	0.4464	0.1797	0.0144	0.0202	0.0444	0.0376	0.1537	0.2448
	$\lambda_{\theta\phi}$	0.1624	0.1874	0.0703	0.0183	0.0060	0.0308	0.0256	0.0083	0.0836
	$\lambda_\phi$	0.0803	0.0586	0.0223	0.0072	0.0011	0.0259	0.0124	0.0002	0.0370
5 – 7	$\lambda_\theta$	-0.2357	0.2334	0.0897	0.0230	0.0503	0.0114	0.0300	0.0178	0.1116
	$\lambda_{\theta\phi}$	0.0525	0.1299	0.0557	0.0064	0.0176	0.0343	0.1062	0.0015	0.1261
	$\lambda_\phi$	0.0931	0.0490	0.0232	0.0031	0.0011	0.0328	0.0434	0.0077	0.0597
7 – 10	$\lambda_\theta$	0.1145	0.1597	0.0783	0.0084	0.0845	0.0151	0.0309	0.0176	0.1218
	$\lambda_{\theta\phi}$	-0.1266	0.1157	0.0449	0.0039	0.0252	0.0116	0.0655	0.0294	0.0892
	$\lambda_\phi$	0.0550	0.0581	0.0215	0.0042	0.0015	0.0165	0.0821	0.0076	0.0870
10 – 15	$\lambda_\theta$	-0.0563	0.1406	0.0622	0.0076	0.0161	0.0081	0.0094	0.0115	0.0669
	$\lambda_{\theta\phi}$	0.0723	0.1309	0.0449	0.0044	0.0392	0.0254	0.0240	0.0178	0.0715
	$\lambda_\phi$	-0.0399	0.0888	0.0352	0.0044	0.0302	0.0149	0.0080	0.0214	0.0540

Table B.21 The measured  $\psi(2S)$  polarization parameters, and their statistical errors and systematic uncertainties in different  $p_T$  bins in rapidity bin  $2.5 < y < 3.0$  in the CS frame. LL stat are the statistical errors returned by likelihood estimators, MCsys are the systematic errors coming from fluctuations of the normalization computed from Monte Carlo, SBsys are the background subtraction systematic uncertainties, Specsyst are systematic uncertainties due to different  $\psi(2S)$  spectrum between data and Monte Carlo, PIDsys are systematic uncertainties coming from the muon PID cut, Trk/AccSys are systematic uncertainties related to the difference between MC and data, it combines the tracking and Monte Carlo efficiency systematic uncertainty quadratically. BConSys are systematic uncertainties coming from contamination of  $\psi(2S)$  from  $b$  decay to our prompt sample. sys sum is the quadratically added systematic uncertainties.

$p_T(\text{GeV}/c)$	$\lambda s$	fit	LL stat	MCsys	SBsys	Specsys	PIDsys	Trk/AccSys	BConSys	sys sum
3 – 4	$\lambda_\theta$	-0.1170	0.0681	0.0262	0.0522	0.0266	0.0050	0.0388	0.0117	0.0761
	$\lambda_{\theta\phi}$	0.0718	0.0551	0.0181	0.0075	0.0157	0.0139	0.0856	0.0056	0.0904
	$\lambda_\phi$	0.0910	0.0362	0.0143	0.0123	0.0063	0.0102	0.0387	0.0055	0.0451
4 – 5	$\lambda_\theta$	-0.1372	0.0379	0.0171	0.0035	0.0496	0.0040	0.0271	0.0044	0.0595
	$\lambda_{\theta\phi}$	0.0019	0.0345	0.0141	0.0030	0.0323	0.0108	0.0702	0.0168	0.0811
	$\lambda_\phi$	0.1072	0.0258	0.0118	0.0035	0.0086	0.0093	0.0414	0.0068	0.0456
5 – 7	$\lambda_\theta$	-0.0693	0.0274	0.0107	0.0039	0.0233	0.0056	0.0352	0.0060	0.0445
	$\lambda_{\theta\phi}$	0.0078	0.0238	0.0118	0.0024	0.0234	0.0097	0.0341	0.0111	0.0455
	$\lambda_\phi$	0.0754	0.0223	0.0109	0.0020	0.0041	0.0095	0.0255	0.0061	0.0303
7 – 10	$\lambda_\theta$	-0.0112	0.0331	0.0161	0.0021	0.0069	0.0036	0.0147	0.0142	0.0272
	$\lambda_{\theta\phi}$	0.1017	0.0250	0.0112	0.0004	0.0068	0.0067	0.0282	0.0065	0.0325
	$\lambda_\phi$	-0.0237	0.0286	0.0126	0.0042	0.0040	0.0021	0.0236	0.0019	0.0275
10 – 15	$\lambda_\theta$	0.0473	0.0564	0.0248	0.0049	0.0033	0.0043	0.0303	0.0267	0.0480
	$\lambda_{\theta\phi}$	0.0485	0.0356	0.0144	0.0017	0.0041	0.0005	0.0230	0.0103	0.0294
	$\lambda_\phi$	-0.0965	0.0456	0.0191	0.0045	0.0047	0.0032	0.0370	0.0117	0.0439

Table B.22 The measured  $\psi(2S)$  polarization parameters, and their statistical errors and systematic uncertainties in different  $p_T$  bins in rapidity bin  $3.0 < y < 3.5$  in the CS frame. LL stat are the statistical errors returned by likelihood estimators, MCsys are the systematic errors coming from fluctuations of the normalization computed from Monte Carlo, SBsys are the background subtraction systematic uncertainties, Specsyst are systematic uncertainties due to different  $\psi(2S)$  spectrum between data and Monte Carlo, PIDsys are systematic uncertainties coming from the muon PID cut, Trk/AccSys are systematic uncertainties related to the difference between MC and data, it combines the tracking and Monte Carlo efficiency systematic uncertainty quadratically. BConSys are systematic uncertainties coming from contamination of  $\psi(2S)$  from  $b$  decay to our prompt sample. sys sum is the quadratically added systematic uncertainties.

$p_T(\text{GeV}/c)$	$\lambda s$	fit	LL stat	MCsys	SBsys	Specsys	PIDsys	Trk/AccSys	BConSys	sys sum
3 – 4	$\lambda_\theta$	-0.1526	0.0395	0.0184	0.0126	0.0012	0.0014	0.0322	0.0163	0.0424
	$\lambda_{\theta\phi}$	0.0740	0.0374	0.0149	0.0009	0.0047	0.0072	0.0058	0.0217	0.0283
	$\lambda_\phi$	0.0043	0.0364	0.0155	0.0025	0.0012	0.0037	0.0029	0.0046	0.0170
4 – 5	$\lambda_\theta$	-0.1191	0.0270	0.0084	0.0082	0.0010	0.0019	0.0333	0.0111	0.0371
	$\lambda_{\theta\phi}$	0.0518	0.0235	0.0086	0.0005	0.0053	0.0073	0.0105	0.0234	0.0286
	$\lambda_\phi$	-0.0071	0.0266	0.0139	0.0054	0.0015	0.0022	0.0086	0.0002	0.0174
5 – 7	$\lambda_\theta$	-0.0542	0.0238	0.0109	0.0027	0.0016	0.0001	0.0237	0.0041	0.0266
	$\lambda_{\theta\phi}$	0.0351	0.0172	0.0092	0.0006	0.0054	0.0033	0.0111	0.0152	0.0219
	$\lambda_\phi$	-0.0494	0.0229	0.0115	0.0035	0.0007	0.0004	0.0047	0.0050	0.0139
7 – 10	$\lambda_\theta$	-0.0868	0.0301	0.0163	0.0030	0.0022	0.0007	0.0149	0.0096	0.0244
	$\lambda_{\theta\phi}$	0.0502	0.0192	0.0086	0.0012	0.0026	0.0008	0.0220	0.0078	0.0251
	$\lambda_\phi$	-0.0307	0.0269	0.0119	0.0044	0.0051	0.0009	0.0091	0.0015	0.0165
10 – 15	$\lambda_\theta$	-0.0034	0.0601	0.0290	0.0074	0.0145	0.0028	0.0283	0.0241	0.0499
	$\lambda_{\theta\phi}$	0.0377	0.0346	0.0164	0.0017	0.0003	0.0030	0.0361	0.0139	0.0422
	$\lambda_\phi$	-0.1058	0.0482	0.0214	0.0059	0.0148	0.0025	0.0281	0.0108	0.0403

Table B.23 The measured  $\psi(2S)$  polarization parameters, and their statistical errors and systematic uncertainties in different  $p_T$  bins in rapidity bin  $3.5 < y < 4.0$  in the CS frame. LL stat are the statistical errors returned by likelihood estimators, MCsys are the systematic errors coming from fluctuations of the normalization computed from Monte Carlo, SBsys are the background subtraction systematic uncertainties, Specsyst are systematic uncertainties due to different  $\psi(2S)$  spectrum between data and Monte Carlo, PIDsys are systematic uncertainties coming from the muon PID cut, Trk/AccSys are systematic uncertainties related to the difference between MC and data, it combines the tracking and Monte Carlo efficiency systematic uncertainty quadratically. BConSys are systematic uncertainties coming from contamination of  $\psi(2S)$  from  $b$  decay to our prompt sample. sys sum is the quadratically added systematic uncertainties.

$p_T(\text{GeV}/c)$	$\lambda s$	fit	LL stat	MCsys	SBsys	Specsys	PIDsys	Trk/AccSys	BConSys	sys sum
3 – 4	$\lambda_\theta$	-0.1843	0.0394	0.0143	0.0120	0.0052	0.0005	0.0802	0.0019	0.0825
	$\lambda_{\theta\phi}$	-0.0112	0.0368	0.0201	0.0034	0.0037	0.0053	0.0249	0.0178	0.0374
	$\lambda_\phi$	0.0166	0.0355	0.0181	0.0104	0.0007	0.0037	0.0029	0.0040	0.0218
4 – 5	$\lambda_\theta$	-0.1487	0.0284	0.0129	0.0097	0.0061	0.0020	0.0635	0.0063	0.0661
	$\lambda_{\theta\phi}$	-0.0351	0.0240	0.0136	0.0063	0.0033	0.0034	0.0257	0.0087	0.0313
	$\lambda_\phi$	-0.0386	0.0278	0.0134	0.0089	0.0011	0.0029	0.0152	0.0016	0.0224
5 – 7	$\lambda_\theta$	-0.0665	0.0269	0.0167	0.0041	0.0039	0.0016	0.0224	0.0011	0.0286
	$\lambda_{\theta\phi}$	0.0029	0.0190	0.0114	0.0032	0.0033	0.0015	0.0258	0.0097	0.0303
	$\lambda_\phi$	-0.0507	0.0254	0.0143	0.0034	0.0040	0.0023	0.0444	0.0057	0.0474
7 – 10	$\lambda_\theta$	0.0058	0.0405	0.0213	0.0047	0.0046	0.0028	0.0131	0.0203	0.0330
	$\lambda_{\theta\phi}$	0.0173	0.0251	0.0151	0.0024	0.0050	0.0007	0.0138	0.0095	0.0232
	$\lambda_\phi$	-0.1161	0.0355	0.0219	0.0075	0.0090	0.0051	0.0446	0.0050	0.0515
10 – 15	$\lambda_\theta$	0.0740	0.0874	0.0490	0.0078	0.0074	0.0007	0.0085	0.0069	0.0513
	$\lambda_{\theta\phi}$	-0.0093	0.0476	0.0193	0.0018	0.0063	0.0006	0.0172	0.0260	0.0373
	$\lambda_\phi$	-0.1805	0.0691	0.0290	0.0075	0.0093	0.0027	0.0286	0.0196	0.0468

Table B.24 The measured  $\psi(2S)$  polarization parameters, and their statistical errors and systematic uncertainties in different  $p_T$  bins in rapidity bin  $4.0 < y < 4.5$  in the CS frame. LL stat are the statistical errors returned by likelihood estimators, MCsys are the systematic errors coming from fluctuations of the normalization computed from Monte Carlo, SBsys are the background subtraction systematic uncertainties, Specsyst are systematic uncertainties due to different  $\psi(2S)$  spectrum between data and Monte Carlo, PIDsys are systematic uncertainties coming from the muon PID cut, Trk/AccSys are systematic uncertainties related to the difference between MC and data, it combines the tracking and Monte Carlo efficiency systematic uncertainty quadratically. BConSys are systematic uncertainties coming from contamination of  $\psi(2S)$  from  $b$  decay to our prompt sample. sys sum is the quadratically added systematic uncertainties.

$p_T(\text{GeV}/c)$	$\lambda s$	fit	LL stat	MCsys	SBsys	Specsys	PIDsys	Trk/AccSys	BConSys	sys sum
3 – 4	$\lambda_\theta$	-0.1636	0.0862	0.0491	0.0219	0.0216	0.0042	0.1381	0.0200	0.1511
	$\lambda_{\theta\phi}$	-0.0141	0.0654	0.0371	0.0078	0.0047	0.0017	0.0665	0.0002	0.0768
	$\lambda_\phi$	-0.0670	0.0478	0.0260	0.0080	0.0013	0.0015	0.0222	0.0021	0.0353
4 – 5	$\lambda_\theta$	-0.1143	0.0579	0.0340	0.0211	0.0385	0.0048	0.1020	0.0136	0.1170
	$\lambda_{\theta\phi}$	0.0013	0.0472	0.0388	0.0075	0.0114	0.0006	0.0793	0.0041	0.0894
	$\lambda_\phi$	-0.0252	0.0382	0.0248	0.0079	0.0003	0.0028	0.0080	0.0023	0.0274
5 – 7	$\lambda_\theta$	-0.0579	0.0495	0.0307	0.0060	0.0111	0.0031	0.0414	0.0196	0.0566
	$\lambda_{\theta\phi}$	0.0103	0.0396	0.0232	0.0124	0.0028	0.0017	0.0273	0.0196	0.0428
	$\lambda_\phi$	-0.0831	0.0398	0.0233	0.0173	0.0041	0.0033	0.0910	0.0115	0.0963
7 – 10	$\lambda_\theta$	-0.0834	0.0656	0.0356	0.0111	0.0055	0.0047	0.0369	0.0119	0.0543
	$\lambda_{\theta\phi}$	-0.0606	0.0489	0.0291	0.0050	0.0058	0.0005	0.0301	0.0228	0.0483
	$\lambda_\phi$	-0.0613	0.0573	0.0359	0.0181	0.0083	0.0037	0.0676	0.0182	0.0813
10 – 15	$\lambda_\theta$	0.4217	0.2150	0.0998	0.0714	0.0454	0.0028	0.0902	0.0738	0.1752
	$\lambda_{\theta\phi}$	0.0207	0.1133	0.0533	0.0212	0.0154	0.0025	0.0267	0.0352	0.0741
	$\lambda_\phi$	-0.4604	0.1568	0.0833	0.0697	0.0090	0.0048	0.0587	0.0790	0.1469

## **Resume and publications**

### **Resume**

- Jun. 1984, born in Shangqiu, Henan province, China
- Sep. 2007 - present, Ph.D. of Physics. Department of Engineering Physics, Tsinghua University, Beijing, China
- Sep. 2003 - Jul. 2007, Bachelor of Physics. Department of Physics, Tsinghua University, Beijing, China

### **Publications**

[1] Zhang Y. (on behalf of the LHCb Collaboration), Quarkonium studies at LHCb. PoS, 2011, EPS-HEP2011:314