

December 2014

## Determination of the X(3872) Meson Quantum Numbers

Bin Gui  
*Syracuse University*

Follow this and additional works at: <https://surface.syr.edu/etd>



Part of the [Physical Sciences and Mathematics Commons](#)

---

### Recommended Citation

Gui, Bin, "Determination of the X(3872) Meson Quantum Numbers" (2014). *Dissertations - ALL*. 189.  
<https://surface.syr.edu/etd/189>

This Dissertation is brought to you for free and open access by the SURFACE at SURFACE. It has been accepted for inclusion in Dissertations - ALL by an authorized administrator of SURFACE. For more information, please contact [surface@syr.edu](mailto:surface@syr.edu).

# *Abstract*

The Large Hadron Collider beauty (LHCb) is one of the several experiments located at the ring of the Large Hadron Collider (LHC) in Geneva. The LHCb detector is a single arm forward spectrometer and is designed to perform high precision measurements of Charge Parity (CP) violation parameters and rare decays of the beauty and charm hadrons. The detector was successfully operated at a center-of-mass energy of 7 TeV in 2010 and 2011 and at 8 TeV in 2012. Over  $3 \text{ fb}^{-1}$  of data has been collected by the LHCb. The LHCb experiment is also well suited for studies on hadron spectroscopy.

Besides the well established mesons consisting of quark-antiquark pairs ( $q\bar{q}$ ), it has been proposed that “exotic”  $qq\bar{q}\bar{q}$  mesons could exist. One of the candidates for a four-quark state is the charmonium-like state  $X(3872)$  which was first observed by the Belle experiment in 2003. This narrow state has a mass of about 3872 MeV which is located in a region of excited charmonium states ( $c\bar{c}$ ). However its mass does not match to any theoretically predicted charmonium state. In order to investigate the nature of this anomalous state, we analyze its quantum number which is the key for its interpretation. The  $X(3872)$  events are reconstructed from  $B^+ \rightarrow X(3872)K^+$ , where  $X(3872) \rightarrow \pi^+\pi^-J/\psi$ ,  $J/\psi \rightarrow \mu^+\mu^-$  based on  $1 \text{ fb}^{-1}$  of 2011 data collected by LHCb detector<sup>1</sup>. We implement a method which is guaranteed by statistics to be the most powerful way to discriminate between spin hypotheses; namely unbinned likelihood ratio test using full angular phase-space. The 5-dimensional analysis shows that  $1^{++}$  hypothesis is preferred with overwhelming significance. The only alternative assignment allowed by the previous measurements,  $J^{PC} = 2^{-+}$ , is rejected with a confidence level equivalent to more than eight Gaussian standard deviations. This result favors exotic explanations of the  $X(3872)$  state, such as a mesonic molecule or a tetraquark.

---

<sup>1</sup>We use  $1.3 \text{ fb}^{-1}$  of 2012 data for a cross-check

DETERMINATION OF THE  $X(3872)$   
MESON QUANTUM NUMBERS

BIN GUI

B.S., UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA, 2007

A DISSERTATION  
SUBMITTED IN PARTIAL FULFILLMENT OF THE  
REQUIREMENTS FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY IN PHYSICS

ADVISER: TOMASZ SKWARNICKI

SYRACUSE UNIVERSITY

DECEMBER 2014

©Copyright by Bin Gui 2014

All Rights Reserved



*Dedicated to my Mother, Father and Wife for their supports and love.*

*“You cannot teach a man anything; you can only find it within yourself.”*

- Galileo Galilei

*“I’ve never played for a draw in my life.”*

- Sir Alex Ferguson

*“Science is the attempt to make the chaotic diversity of our sense-experience correspond to a logically uniform system of thought.”*

- Albert Einstein

---

## *A c k n o w l e d g m e n t s*

Eventually, it comes to the end of the journey of pursuing a Ph.D. It is a lonely as well as rewarding experience for an international student. I gratefully acknowledge the support and help from experimental high energy physics group in the past five years. First and foremost I want to express my sincere gratitude to my advisor, professor **Tomasz Skwarnicki** for his support, guidance and encouragement. He teaches me not only the skills of physics analysis but also the attitude toward the scientific research. It is an unforgettable time when we were working together on the measurement of X(3872) quantum number. His intelligent ideas, stringent attitude and great personality set a good example of physicist for me. Besides my advisor, I would like to thank professor Sheldon Stone. His immense knowledge in heavy flavor physics and particle detector development makes him a great leader of our experimental high energy physics group. I want to thank him for providing me opportunities to present the analysis results in “US LHC users organization annual meeting 2011” and “The 2013 Phenomenology Symposium”. I am very grateful to professor Marina Artuso who is experienced in both physics analysis and detector R&D. She is always willing to provide technical support and professional suggestions. My sincere thanks also goes to professor J.C. Wang and professor Raymond Mountain who have strong background in experimental techniques of high energy physics. I have been working with both of them on electronics and mechanical through my Ph.D study. Their advices on research as well as on my career have been invaluable. I want to express my sincere appreciation to Dr. Liming Zhang now a professor at Tsinghua University who is an extremely experienced researcher. He has been selflessly sharing his knowledge and experience in physics analysis with us. I also want to thank all the other group members, professor Steven Blusk, Thomas Briton and others. Their

---

suggestions and comments are priceless to me and to my research. Last but not the least, I am very much indebted to my lovely wife and my caring parents for their encouragement and help at every stage of my life. Their love and support are my advancing motivation.

*Syracuse, New York*

*Nov., 2014*

# Contents

List of Figures	xi
List of Tables	xxi
<b>1 Physics Introduction</b>	<b>1</b>
1.1 The Standard Model of Particle Physics	1
1.1.1 Standard Model Particles	2
1.1.2 Standard Model Interactions	5
1.2 Charmonium States	7
1.3 The $X(3872)$	8
1.3.1 The Discovery of A Narrow $X(3872)$ State	9
1.3.2 The Nature of the $X(3872)$ : Theoretical Interpretations	11
1.3.3 The Early Measurements of the $X(3872)$ Quantum Numbers	12
<b>2 The LHCb Experiment</b>	<b>17</b>
2.1 The CERN	17
2.2 The LHC	18
2.3 The LHCb	21
2.3.1 Introduction	22
2.3.2 Magnet	24
2.3.3 Tracking System	26
2.3.3.1 VELO	26

2.3.3.2	Silicon Tracker . . . . .	32
2.3.3.3	Outer Tracker . . . . .	36
2.3.3.4	Summary of Tracking Detectors . . . . .	38
2.3.4	Particle Identification . . . . .	38
2.3.4.1	Ring Imaging Cherenkov Detector . . . . .	39
2.3.4.2	The Scintillator Pad Detector and the Preshower Detector . . . . .	45
2.3.4.3	Electromagnetic Calorimeter . . . . .	48
2.3.4.4	Hadronic Calorimeter . . . . .	51
2.3.4.5	Muon System . . . . .	54
2.3.4.6	Summary of PID Devices . . . . .	58
2.3.5	Trigger . . . . .	58
2.3.5.1	The Level-0 Trigger . . . . .	58
2.3.5.2	The High Level Trigger . . . . .	62
2.3.6	The LHCb Software . . . . .	63
2.3.7	Data Processing . . . . .	65
2.3.8	Summary . . . . .	67
3	Measurement . . . . .	69
3.1	Data Selection . . . . .	69
3.2	Angular Matrix Element . . . . .	82
3.3	Discrimination between the two $J^{PC}$ hypotheses . . . . .	87
3.4	Verification of the likelihood analysis on EVTGEN MC . . . . .	93
3.5	Results from the likelihood analysis on the real data. . . . .	95
3.5.1	The results for 2011 and 2012 data compared to the expectations from the MC . . . . .	95
3.5.2	Significance of $2^{-+}$ rejection and goodness-of-fit for $1^{++}$ using toy experiments . . . . .	99

3.5.3	Event-by-event distribution of the likelihood ratio . . . . .	104
3.5.4	Goodness-of-fits using 1D and 2D projections . . . . .	111
3.6	Stability of result with background level in the data . . . . .	116
3.7	Stability of result with changes of signal efficiency . . . . .	121
3.8	Cross-check of likelihood analysis on $B^+ \rightarrow \psi(2S)K^+$ signal . . . . .	124
3.9	Using $\psi(2S)$ signal for efficiency corrections in $X(3872)$ fit. . . . .	127
3.10	Testing other $J^{PC}$ assignments to $X(3872)$ . . . . .	128
3.11	Performing analysis in less than full phase-space . . . . .	129
3.12	Summary . . . . .	134
4	<b>Conclusion</b>	137
	References	139
	Vita	145

# List of Figures

1.1	The elementary particles of standard model [3]. . . . .	2
1.2	The diagram of established Hadrons: $q\bar{q}$ Meson (left), Baryon (right) [4]. . . . .	4
1.3	The interactions between elementary particles of standard model [3]. . . . .	6
1.4	The mass spectrum of charmonium states. The observed states are labeled with solid horizontal line. . . . .	7
1.5	Distribution of $M(\pi^+\pi^-J/\psi) - M(J/\psi)$ observed by Belle. In addition to the well-known $\psi(2S)$ (peaking at $\approx 0.6$ GeV), a second peak was seen at $\approx 0.8$ GeV. . . . .	10
1.6	Distribution of $M(J/\psi\pi^+\pi^-)$ observed by LHCb. The masses of $\psi(2S)$ (left) and the X(3872) (right) were measured to be $3686.12 \pm 0.06_{stat} \pm 0.10_{syst}$ MeV and $3871.95 \pm 0.48_{stat} \pm 0.12_{syst}$ MeV, respectively. . . . .	10
1.7	The comparisons of $1^{++}$ and $2^{-+}$ hypotheses. The solid black points are the data; the dashed blue histogram indicates the background determined from the events in the scaled $M(J/\psi\pi^+\pi^-)$ sidebands; the solid blue histogram is the sum of the simulated MC $X(3872) \rightarrow J/\psi\rho$ events generated with $1^{++}$ (left) / $2^{-+}$ (right) hypothesis and the background (dashed histogram) and normalized to the observed signal. . . . .	14
2.1	An overall view of the LHC experiments. . . . .	19
2.2	The LHC's injection chain. . . . .	20
2.3	The first $p\bar{p}$ collision at $\sqrt{s} = 3.5$ TeV at the LHCb detector, 2010. . . . .	21
2.4	LHCb collaboration in front of LHCb detector. . . . .	22
2.5	Side view of LHCb detector. . . . .	23
2.6	The polar angle distribution with respect to the beam line (b and $\bar{b}$ ) at $\sqrt{s} = 7$ TeV. . . . .	23



2.7	The LHCb dipole magnet (units in mm). . . . .	25
2.8	An overview of the VELO. Top is the schematic view of the twenty-one stations of the VELO along the beam direction; bottom are the closed and open positions of the VELO sensors. . . . .	27
2.9	An overview of the VELO vacuum vessel. . . . .	28
2.10	The VELO tracking efficiency for the 2011 data (black) and reweighed simulation (red) as a function of the momentum (top left), the pseudorapidity (top right), the azimuthal angle $\phi$ (bottom left) and the total number of the tracks in the event (bottom right). . . . .	29
2.11	Left plot shows the VELO resolution for two different project angles; right plot shows the VELO PV resolution for x and y directions. . . . .	29
2.12	Schematic illustration of $R$ -measuring and $\phi$ -measuring sensors. The routing lines (green) orientate perpendicular and parallel to the silicon strips (black), respectively. . . . .	30
2.13	Schematics of the VELO hybrid design (left) and the VELO module with the key components (right). . . . .	31
2.14	A schematic of the VELO front-end electronics. . . . .	32
2.15	Layout of the four detection layers of the TT. . . . .	33
2.16	Schematic of a TT half module. . . . .	34
2.17	Left: view of the IT boxes arranged around the beam pipe. Right: layout of the $x$ -layer in the second IT station. . . . .	35
2.18	Data process from one Beetle chip of the ST. . . . .	35
2.19	Arrangement of the OT straw-tube modules in layers and stations. . . . .	36
2.20	Cross section of a straw-tube module of the OT. . . . .	37
2.21	Schematic of various types of tracks through the LHCb detector: long, upstream, downstream, VELO and T tracks. . . . .	38
2.22	Schematics of the RICH1 and RICH2. . . . .	39
2.23	Cherenkov angles as a function of momentum for the RICH radiators. . . . .	40

2.24	Left: schematics of the pixel Hybrid Photon Detector (HPD). Right: HPD array of the RICH1 detector. . . . .	41
2.25	Event display of detected photoelectrons in RICH1 (left) and RICH2 (right). The circles are the fit results. . . . .	42
2.26	A drawing of the MDCS system of RICH1. Dotted circles represent the HPDs. The light bar is in parked position at the right edge of the box. . . . .	43
2.27	Distributions of the difference in log-likelihood between kaon and pion hypotheses for kaons (top) and pions (bottom). Two different requirements are implemented on the sample of $B_s^0 \rightarrow D_s^- K^+$ events. . . . .	44
2.28	The efficiency of kaon identification and the rate of muon misidentification as a function of momentum. . . . .	45
2.29	Side view of the SPD/PS detectors (between the MS1 and ECAL). . . . .	46
2.30	Energy deposition of 50 GeV electrons and pions in the PS. . . . .	47
2.31	General view of the ECAL (left) and HCAL (right). . . . .	48
2.32	Illustration of the segmentations of the SPD/PS, ECAL (left) and HCAL (right). One quarter of the detector front face is shown. . . . .	49
2.33	The energy resolution measured in the outer region of the ECAL (sources are electrons). . . . .	50
2.34	Scintillator (left) and WLS fiber (right) degradation and annealing effect after irradiation at LEP Injector Linac (LIL) up to 5 Mrad dose. . . . .	51
2.35	Schematic of the HCAL (left) and ECAL (right) structures. . . . .	52
2.36	The energy resolution of the HCAL as a function of energy. . . . .	53
2.37	Side view of the muon system. . . . .	54
2.38	Front view of one quadrant of the first muon station (M1). . . . .	55
2.39	Exploded schematic of a four-gap MWPC. . . . .	56
2.40	Left: exploded schematic of the GEM. Right: illustration of the GEM foil. . . . .	57
2.41	General view of the Level-0 trigger. . . . .	59

2.42	Left: basic principle of detecting vertices for the L0 Pile-Up trigger system. Right: histogram of the values of $z_v$ calculated from the hits of the two planes. . . . .	60
2.43	Tracking finding in the muon stations. . . . .	61
2.44	Flow-diagram of the data through the triggers. . . . .	62
2.45	Object diagram of the GAUDI architecture. . . . .	63
2.46	The computing logical data flow model of the LHCb. . . . .	66
2.47	The delivered and recorded luminosity of LHCb detector, 2010 ~ 2012. . . . .	67
2.48	The LHCb computing model. . . . .	68
3.1	Fraction of events after the preselection passing a cut on $\chi_{\text{IP}}^2(h) > X$ , where $X$ is given on the horizontal axis, for $B^+ \rightarrow \psi(2S)K^+$ Monte Carlo (solid black line with squares), $B^+ \rightarrow \psi(2S)K^+$ signal in the data (blue dotted line) and background events in the data from the far-sideband (solid red line with triangles). . . . .	72
3.2	Fraction of events after the preselection passing a cut on $\chi_{\text{IP}}^2(B^+) < X$ , where $X$ is given on the horizontal axis, for $B^+ \rightarrow \psi(2S)K^+$ Monte Carlo (solid black line with squares), $B^+ \rightarrow \psi(2S)K^+$ signal in the data (blue dotted line) and background events in the data from the far-sideband (solid red line with triangles). . . . .	72
3.3	Fraction of events after the preselection passing a cut on $\chi_{\text{vtx}}^2(B^+) < X$ , where $X$ is given on the horizontal axis, for $B^+ \rightarrow \psi(2S)K^+$ Monte Carlo (solid black line with squares), $B^+ \rightarrow \psi(2S)K^+$ signal in the data (blue dotted line) and background events in the data from the far-sideband (solid red line with triangles). . . . .	73
3.4	Fraction of events after the preselection passing a cut on cosine of the largest opening angle between the $\psi$ and the charged hadron candidates, $\cos(\text{hadron}, J/\psi) > X$ , where $X$ is given on the horizontal axis, for $B^+ \rightarrow \psi(2S)K^+$ Monte Carlo (solid black line with squares), $B^+ \rightarrow \psi(2S)K^+$ signal in the data (blue dotted line) and background events in the data from the far-sideband (solid red line with triangles). . . . .	73

- 
- 3.5 Fraction of events after the preselection passing a cut on  $DLL_{\text{sig/bkg}} < X$ , where  $X$  is given on the horizontal axis, for  $B^+ \rightarrow \psi(2S)K^+$  Monte Carlo (solid black line with squares),  $B^+ \rightarrow \psi(2S)K^+$  signal in the data (blue dotted line) and background events in the data from the far-sideband (solid red line with triangles). The vertical line indicates the actual requirement used in our analysis. . . . . 74
- 3.6 Expected dependence of  $\text{signal}/\sqrt{\text{signal} + \text{background}}$  for  $B^+ \rightarrow X(3872)K^+$  in a function of a cut on  $DLL_{\text{sig/bkg}} < X$ , where  $X$  is given on the horizontal axis. The vertical line indicates the actual requirement used in our analysis. . . . . 75
- 3.7 Observed  $M(J/\psi \pi^+ \pi^- K^+)$  mass distribution in the data. The fit results are superimposed (total-blue, signal-red, background-dashed). The vertical bars illustrate the range used for angular analysis. . . . . 76
- 3.8 Distribution of  $M(\pi^+ \pi^- J/\psi) - M(J/\psi)$  for  $B^+ \rightarrow J/\psi K^+ \pi^+ \pi^-$  events. Peaks from  $B^+ \rightarrow \psi(2S)K^+$  and  $B^+ \rightarrow X(3872)K^+$  are clearly visible. . . . . 76
- 3.9 Distribution of  $M(\pi^+ \pi^- J/\psi) - M(J/\psi)$  for  $B^+ \rightarrow J/\psi K^+ \pi^+ \pi^-$  events. The fit of the  $\psi(2S)$  signal is displayed. The blue, red and green blue lines represent the total fit, signal component and background component, respectively. . . . . 77
- 3.10 Distribution of  $M(\pi^+ \pi^- J/\psi) - M(J/\psi)$  for  $B^+ \rightarrow J/\psi K^+ \pi^+ \pi^-$  events (2011 data). The fit of the  $X(3872)$  signal is displayed. The blue, red and green blue lines represent the total fit, signal component and background component, respectively. . . . . 77
- 3.11 Distribution of  $M(\pi^+ \pi^- J/\psi) - M(J/\psi)$  for  $B^+ \rightarrow J/\psi K^+ \pi^+ \pi^-$  events (2012 data). The fit of the  $X(3872)$  signal is displayed. The blue, red and green blue lines represent the total fit, signal component and background component, respectively. . . . . 78

3.12	Distribution of $M(\pi^+\pi^-)$ for $B^+ \rightarrow X(3872)K^+$ , $X(3872) \rightarrow \pi^+\pi^- J/\psi$ events. The top plot shows the data in the $X(3872)$ peak region ( $\pm 3\sigma_M$ ) with no sideband subtraction. The bottom plot shows the $X(3872)$ region used in the angular analysis (723 – 823 MeV) after the sideband subtraction using the sWeights (blue points with error bars) compared to the simulations assuming $X(3872) \rightarrow \rho(770)J/\psi$ decay (solid black histogram). . . . .	79
3.13	Distribution of rapidity for $B^+ \rightarrow X(3872)K^+$ events in 2011 data (red points with squares) and in EVTGEN $1^{++}$ MC (blue points with triangles). . . . .	80
3.14	Distribution of $X(3872)$ $p_T$ for $B^+ \rightarrow X(3872)K^+$ events in 2011 data (red points with squares) and in EVTGEN $1^{++}$ MC (blue points with triangles). . . . .	81
3.15	Distribution of $X(3872)$ momentum for $B^+ \rightarrow X(3872)K^+$ events in 2011 data (red points with squares) and in EVTGEN $1^{++}$ MC (blue points with triangles). . . . .	81
3.16	Definition of the helicity angles (top) and of the azimuthal angles between the decay planes (bottom) used in the description of angular distribution in $B^+ \rightarrow X(3872)K^+$ , $X(3872) \rightarrow \rho J/\psi$ , $\rho \rightarrow \pi^+\pi^-$ , $J/\psi \rightarrow \mu^+\mu^-$ . . . . .	83
3.17	Distribution of $M(\pi^+\pi^- J/\psi) - M(J/\psi)$ for $B^+ \rightarrow X(3872)K^+$ , $X(3872) \rightarrow \rho(770)J/\psi$ , $\rho(770) \rightarrow \pi^+\pi^-$ , $J/\psi \rightarrow \mu^+\mu^-$ phase-space Monte Carlo events. The fit of the $X(3872)$ signal (symmetric Crystal Ball shape, with $\alpha_M = 1.41$ , $n = 10.7$ ) and linear background is superimposed. The fit gives $41,789 \pm 205$ signal events peaking at $775.40 \pm 0.02$ MeV with resolution $\sigma_M = 4.17 \pm 0.02$ MeV and $129 \pm 20$ background events. The latter is so small that the background polynomial is not visible. . . . .	91
3.18	Contours of $\sqrt{\chi_L^2(2^{-+}, \alpha) - \chi_L^2(2^{-+})_{min}}$ . . . . .	98
3.19	Distribution of $\Delta\chi_L^2$ for the simulated experiments (points with error bars), with $J^{PC} = 2^{-+}$ and $\alpha$ determined by the angular fit to the data, and a value of $\Delta\chi_L^2$ for the real data (vertical solid-bar). A fit of the Gaussian to the simulated distribution is superimposed (blue solid line). The bottom plot is the same as the top plot, except for the logarithmic vertical scale. . . . .	102

- 
- 3.20 Distribution of  $\Delta\chi_L^2$  for the simulated experiments (red points with error bars), with  $J^{PC} = 1^{++}$ , and a value of  $\Delta\chi_L^2$  for the real data (vertical solid-bar). . . . . 103
- 3.21 Distribution of  $\chi_L^2(1^{++})$  for the simulated experiments with  $J^{PC} = 1^{++}$  (red points with error bars) and a value of  $\chi_L^2(1^{++})$  for the real data (vertical solid-bar). . . . . 103
- 3.22 Distribution of  $n_\sigma^{toy MC^*}$  for the simulated experiments with  $J^{PC} = 1^{++}$  (points with error bars) and a value of  $n_\sigma^{toy MC^*}$  for the real data (vertical solid-bar). . . . . 104
- 3.23 Distribution of  $LLR$  with  $\alpha = (0.671, 0.280)$  for the 2011 data (black points with squares) compared to the  $LLR$  distribution for the simulated experiments with  $J^{PC} = 1^{++}$  (red points with triangles) and with  $2^{-+}$ ,  $\alpha = (0.671, 0.280)$  (blue points with open circles) after the background subtraction using sWeights. The simulated distributions are normalized to the number of signal events observed in the data. Bin content and its error are divided by bin width because of unequal bin sizes. . . . . 106
- 3.24 Distribution of  $LLR$  with  $\alpha = (0.68, 0.32)$  for the background-free EVTGEN simulation of  $1^{++}$  (blue points with open circles), compared to the  $LLR$  distributions for the  $1^{++}$  samples with the background (red points with triangles), before (top) and after (bottom) the background subtraction using sWeights. The plots on the left (right) show the simulations with the  $B^+ \rightarrow J/\psi K_1(1270)^+$  ( $B^+ \rightarrow J/\psi K^+ \pi^+ \pi^-$  phase-space) background model. The simulated distributions are normalized to the number of signal events observed in the data. Bin content and its error are divided by bin width because of unequal bin sizes. . . . . 107

3.25 Distribution of  $LLR$  with  $\alpha = (0.68, 0.32)$  for the background-free EVTGEN simulation of  $2^{-+}$   $\alpha = (0.68, 0.32)$  (blue points with open circles), compared to the  $LLR$  distributions for the  $2^{-+}$   $\alpha = (0.68, 0.32)$  samples with the background (red points with triangles), before (top) and after (bottom) the background subtraction using sWeights. The plots on the left (right) show the simulations with the  $B^+ \rightarrow J/\psi K_1(1270)^+$  ( $B^+ \rightarrow J/\psi K^+ \pi^+ \pi^-$  phase-space) background model. The simulated distributions are normalized to the number of signal events observed in the data. Bin content and its error are divided by bin width because of unequal bin sizes. . . . . 108

3.26 Distribution of  $LLR$  with  $\alpha = (0.671, 0.280)$  without the background subtraction for the 2011 data (black points with squares) compared to the  $LLR$  distributions for the simulated experiments with  $J^{PC} = 1^{++}$  (red points with triangles) and the two background models;  $B^+ \rightarrow J/\psi K_1(1270)^+$  (top) and  $B^+ \rightarrow J/\psi K^+ \pi^+ \pi^-$  phase-space (bottom). The simulated distributions are normalized to the number of events in the data. Bin content and its error are divided by bin width because of unequal bin sizes. . . . . 109

3.27 Distribution of  $LLR$  for the 2011 data (black points with squares) compared to the  $LLR$  distributions for the EVTGEN simulation of  $1^{++}$  (red points with triangles) and of  $2^{-+}$  (blue points with open circles), with  $\alpha = (0.68, 0.32)$  (top),  $(0.5, 0.0)$  (bottom left) and  $(1.0, 0.0)$  (bottom right). The value of  $\alpha$  used in the  $LLR$  definition changes and is set to the value of  $2^{-+}$  MC sample used in the comparison. The distributions are normalized to the number of signal events observed in the data. Bin content and its error are divided by bin width because of unequal bin sizes. . . . . 110

3.28 Distribution of  $\cos \theta_X$  for the data (black points with squares) and for the  $1^{++}$  (red points with triangles) and  $2^{-+}$  (blue points with open circles) fit projections on the left, and for the phase-space MC on the right. . . . . 113

3.29	Distribution of $\cos \theta_{\pi\pi}$ for the data (black points with squares) and for the $1^{++}$ (red points with triangles) and $2^{-+}$ (blue points with open circles) fit projections on the left, and for the phase-space MC on the right. . . . .	113
3.30	Distribution of $\cos \theta_{J/\psi}$ for the data (black points with squares) and for the $1^{++}$ (red points with triangles) and $2^{-+}$ (blue points with open circles) fit projections on the left, and for the phase-space MC on the right. . . . .	113
3.31	Distribution of $\phi_X - \phi_{\pi\pi}$ for the data (black points with squares) and for the $1^{++}$ (red points with triangles) and $2^{-+}$ (blue points with open circles) fit projections on the left, and for the phase-space MC on the right. . . . .	114
3.32	Distribution of $\phi_X - \phi_{J/\psi}$ for the data (black points with squares) and for the $1^{++}$ (red points with triangles) and $2^{-+}$ (blue points with open circles) fit projections on the left, and for the phase-space MC on the right. . . . .	114
3.33	Distribution of $\cos \theta_X$ for the data (black points with squares) and for the $1^{++}$ (red points with triangles) and $2^{-+}$ (blue points with open circles) fit projections on the left, and for the phase-space MC on the right, for 5 different $\cos \theta_{\pi\pi}$ bins. . . . .	115
3.34	Distribution of $\cos \theta_X$ with the $ \cos \theta_{\pi\pi}  > 0.6$ requirement for the data (black points with squares) and for the projections of the 5-dimensional fit of the $1^{++}$ hypothesis (red points with triangles) and $2^{-+}$ hypothesis (blue points with open circles). Compare to figure 3.28. . . . .	117
3.35	Distribution of $M(\pi^+\pi^-J/\psi) - M(J/\psi)$ for $B^+ \rightarrow J/\psi K^+\pi^+\pi^-$ events with the requirement $Q < 100$ MeV. The fit of the $X(3872)$ signal is displayed. The blue, red and green blue lines represent the total fit, signal component and background component, respectively. Compare to figure 3.10 which has no $Q$ cut. . . . .	118
3.36	Distribution of $\cos \theta_X$ for the data in the $X(3872)$ sidebands scaled to the expected background within the $X(3872)$ mass peak ( $\pm 3\sigma$ ) for the selection with no $Q$ requirement (red points with squares) and for $Q < 100$ MeV (blue points with triangles). . . . .	119



3.37	Distribution of $\cos \theta_{\pi\pi}$ for the data in the $X(3872)$ sidebands scaled to the expected background within the $X(3872)$ mass peak ( $\pm 3\sigma$ ) for the selection with no $Q$ requirement (red points with squares) and for $Q < 100$ MeV (blue points with triangles). . . . .	119
3.38	Distribution of $\cos \theta_{J/\psi}$ for the data in the $X(3872)$ sidebands scaled to the expected background within the $X(3872)$ mass peak ( $\pm 3\sigma$ ) for the selection with no $Q$ requirement (red points with squares) and for $Q < 100$ MeV (blue points with triangles). . . . .	120
3.39	Distribution of $\phi_X - \phi_{\pi\pi}$ for the data in the $X(3872)$ sidebands scaled to the expected background within the $X(3872)$ mass peak ( $\pm 3\sigma$ ) for the selection with no $Q$ requirement (red points with squares) and for $Q < 100$ MeV (blue points with triangles). . . . .	120
3.40	Distribution of $\phi_X - \phi_{J/\psi}$ for the data in the $X(3872)$ sidebands scaled to the expected background within the $X(3872)$ mass peak ( $\pm 3\sigma$ ) for the selection with no $Q$ requirement (red points with squares) and for $Q < 100$ MeV (blue points with triangles). . . . .	121
3.41	Distribution of $M(\pi^+\pi^-)$ for two versions of $B^+ \rightarrow X(3872)K^+$ , $X(3872) \rightarrow \pi^+\pi^-J/\psi$ phase-space MC events; $X(3872) \rightarrow \rho(770)J/\psi$ 2-body phase-space decay (black solid histogram) and $X(3872) \rightarrow \pi^+\pi^-J/\psi$ 3-body phase-space decay (blue dashed histogram). . . . .	123
3.42	Distribution of $\cos \theta_X$ for the data (black points with squares) and for the $1^{++}$ (red points with triangles) and $2^{-+}$ (blue points with open circles) 1D fits to the unbinned $\cos \theta_X$ data. . . . .	133
3.43	Distribution of $\cos \theta_X$ for the $1^{++}$ EVTGEN MC (red points with triangles) and for the $2^{-+}$ $\alpha = (0.5, 0.0)$ EVTGEN MC (blue points with open circles). . . . .	134

# List of Tables

1.1	List of the twelve Fermions. . . . .	3
1.2	List of the fundamental interactions. . . . .	5
1.3	Names of the states in charmonium spectrum. . . . .	8
1.4	Charmonium states near the mass of X(3872). The masses of predicted but unobserved states have no errors assigned. . . . .	11
1.5	The result of X(3872) angular analysis obtained by CDF. From left hand side to right hand side, the five columns are state, the decay mode, the $L$ and $S$ quantum numbers of $J/\psi(\pi\pi)$ system, the $\chi^2$ with 11 degrees of freedom and the $\chi^2$ probability. . . . .	15
3.1	Data selection requirements. . . . .	70
3.2	Values of $A_{\lambda_{J/\psi}, \lambda_{\pi\pi}}$ . . . . .	87
3.3	The results from $1^{++}/2^{-+}$ hypotheses testing on EVTGEN generated Monte Carlo data using 5D likelihood ratio test. . . . .	95
3.4	The results from $1^{++}/2^{-+}$ hypotheses testing on the real data and on the $1^{++}$ EVTGEN MC using 5D likelihood ratio test. The results for the MC are scaled to the 2011 signal statistics, except for the errors on $\alpha$ (scaling the errors gives $\pm 0.04$ for the expected errors in the data). . . . .	97
3.5	Tests performed on 1- and 2-dimensional projections of the data and of the fits. The last column shows predictions for an average expected $n_{\sigma}^{proj}$ obtained by fitting the large statistics EVTGEN $1^{++}$ MC and scaling down to the actual data sample. . . . .	116

- 3.6 Results of the 5D likelihood ratio test between the  $1^{++}$  and  $2^{-+}$   $J^{PC}$  hypotheses for various cuts on  $Q$  (i.e.  $M(\pi^+\pi^-)$ ) for the 2011 data. The first row corresponds to our nominal results. The MC predictions included in the table are based on high statistics EVTGEN  $1^{++}$  MC sample, with the results scaled down to the signal statistics ( $N_s$ ) observed in the data and multiplied by the  $s_w$  factor to account for the background subtraction error. . . . . 118
- 3.7 Results of the 5D likelihood ratio test between the  $1^{++}$  and  $2^{-+}$   $J^{PC}$  hypotheses for various cuts on  $p_T(\pi)$  for the 2011 data. The first row corresponds to our nominal results. The MC predictions included in the table are based on high statistics EVTGEN  $1^{++}$  MC sample, with the results scaled down to the signal statistics ( $N_s$ ) observed in the data and multiplied by the  $s_w$  factor to account for the background subtraction error. . . . . 122
- 3.8 The results from  $1^{++}/2^{-+}$  hypotheses testing on the  $X(3872)$  2011 data using 5D likelihood ratio test. The two rows differ by which data were used to perform normalization integral of the  $\mathcal{P}$ s. The nominal fit to the data uses phase-space MC with  $X(3872) \rightarrow \rho(770)J/\psi$  decay (1st row). As an alternative we used phase-space MC with  $X(3872) \rightarrow \pi^+\pi^-J/\psi$  decay (2nd row). . . . . 124
- 3.9 Tests performed on 1- and 2-dimensional projections of the  $\psi(2S)$  data and of the  $1^{--}$  fit. . . . . 126
- 3.10 The results from  $1^{--}/J^{P-}$  hypotheses testing on the real  $\psi(2S)$  data and on the  $1^{--}$  EVTGEN MC using 5D likelihood ratio test. The results for the MC are scaled to the 2011 signal statistics, except for the errors on  $\alpha_{2S}$  (scaling the errors gives  $\pm 0.029$  for the expected  $\alpha_{2S}^{1^{--}}$  errors in the data). . . . . 126

---

3.11	The results from $1^{++}/2^{-+}$ hypotheses testing on the $X(3872)$ 2011 data using 5D likelihood ratio test. The two rows differ by which data were used to perform normalization integral of the $\mathcal{P}$ . The nominal fit to the data uses high statistics phase-space MC (1st row). As an alternative (2nd row) we used the $\psi(2S)$ signal in the data, reweighted to the phase-space distribution by an inverse of the matrix element squared given by equation 3.30 for $J^{PC} = 1^{--}$ with the value of $\alpha_{2S}^{1^{--}}$ from the fit to the $\psi(2S)$ data (see table 3.10). . . . .	128
3.12	The results from $1^{++}/J^{P-}$ hypotheses testing on the real $X(3872)$ 2011 data and on the $1^{++}$ EVTGEN MC using full-phase likelihood ratio test. The results for the MC are scaled to the 2011 signal statistics. . . . .	129
3.13	The results from $1^{++}/2^{-+}$ hypotheses testing on the $X(3872)$ 2011 data and on the EVTGEN MC using the likelihood ratio test at different dimensions (5D represents the nominal analysis). The MC results have been scaled to the statistics and background in the data (except for the error on $\alpha$ ). . . . .	132

# Chapter 1

## Physics Introduction

This chapter briefly introduces the standard model of particle physics and the charmonium states, then reviews the discovery of the charmonium-like state  $X(3872)$ . Since the properties of the  $X(3872)$  are not in good agreement with the traditional quark model, several theories have been proposed for its nature, such as a mesonic molecule, charmonium-gluon hybrid or tetraquark. Determination of quantum numbers of this particle is a key to its interpretation.

### 1.1 The Standard Model of Particle Physics

The Standard Model (SM) of elementary particle physics is a theoretical framework to describe the fundamental particles and their interactions. The theory was formulated in the 1970s, and the name of standard model was also given in that period. The standard model describes the world as built up by a very few elementary particles whose existences have been proven by experiments, which include the recently discovered Higgs boson announced at CERN on July 4, 2012 [1, 2]. It also explains three out of the four elementary forces in nature. So sometimes the standard model is regarded as a “theory of almost everything”.

### 1.1.1 Standard Model Particles

In the standard model, there are seventeen fundamental particles which can be divided into two types. One group of fundamental particles, which make up matter, are called fermions and are spin  $1/2$ . There are twelve fermions obeying the Pauli exclusion principle. The others are force carriers, which mediate the strong, weak, and electromagnetic fundamental interactions, called bosons, with integer spin. The five bosons follow Bose statistics. All seventeen fundamental particles have been experimentally observed by now, and are shown in figure 1.1.

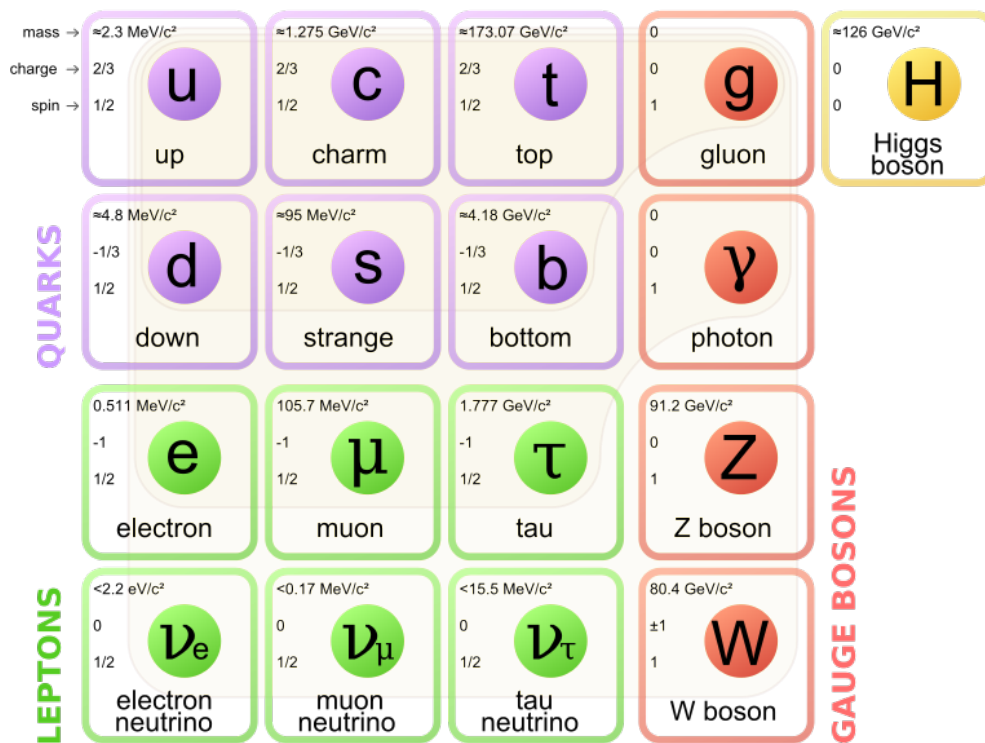


Figure 1.1: The elementary particles of standard model [3].

The twelve fermions of the standard model are categorized into two groups of six: quarks and leptons. Each fermion has its own flavor quantum numbers and a corresponding antiparticle with the same mass, but opposite quantum numbers. The twelve fermions are arranged in three generations, each comprising a  $SU(2)$  weak doublet of quarks and a doublet of

leptons. The first generation includes up and down quarks, electron and electron neutrino; the second charm and strange quarks, muon and muon neutrino; and the third bottom, which is sometimes referred to as beauty, and top quarks, tau and tau neutrino; listed in table 1.1. The masses of fermions increase from the first generation to the third. The heavier charged fermions rapidly decay into the first generation charged fermions (electron, up and down quarks). Due to this, electron and the lightest baryon made out of up and down quarks are stable and the most common in nature, whereas muons, taus, muon and tau leptons, charm, strange, top, and bottom quarks can only be created in high energy collisions such as proton-proton collisions at CERN.

Table 1.1: List of the twelve Fermions.

Name	Symbol	Mass MeV	Electric Charge $Q_e$
<i>1<sup>st</sup> generation</i>			
“Up”	$u$	$2.3^{+0.7}_{-0.5}$	$\frac{2}{3}$
“Down”	$d$	$4.8^{+0.5}_{-0.3}$	$-\frac{1}{3}$
“Electron”	$e$	0.511	-1
“Electron neutrino”	$\nu_e$	$< 2.2 \times 10^{-6}$	0
<i>2<sup>nd</sup> generation</i>			
“Charm”	$c$	$1275 \pm 25$	$\frac{2}{3}$
“Strange”	$s$	$95 \pm 5$	$-\frac{1}{3}$
“Muon”	$\mu$	105.7	-1
“Muon neutrino”	$\nu_\mu$	$< 0.17$	0
<i>3<sup>rd</sup> generation</i>			
“Top”	$t$	$\sim 1.73 \times 10^5$	$\frac{2}{3}$
“Bottom”	$b$	$\sim 4.18 \times 10^3$	$-\frac{1}{3}$
“Tau”	$\tau$	1777	-1
“Tau neutrino”	$\nu_\tau$	$< 15.5$	0

Quarks are the only particles whose electric charges are not integer multiples of the elementary charge  $Q_e$ , which equals to  $1.6 \times 10^{-19}$  Coulomb. Each quark doublet includes one “up” quark with electric charge  $+2/3 Q_e$  and one “down” quark with electric charge  $-1/3 Q_e$ , where “up” and “down” stand for the weak isospin (an “up” quark carries  $+1/2$  weak isospin, a “down”

quark carries  $-1/2$  weak isospin). Quarks possess another intrinsic property called color charge. There are three types of color charge, arbitrarily labeled blue, green and red. Each of them is complemented by an anticolor: antiblue, antigreen and antired. Each quark carries one color, while every antiquark carries an anticolor. Quarks are known as the only elementary particles in the standard model which experience all four fundamental interactions, since quarks are the unique particles who possess electric charge, weak charge and color charge. Because of a phenomenon known as color confinement, quarks are never directly observed or found in isolation; they can be found only within hadrons which are color neutral. The hadron family contains mesons which comprise pairs of quark and antiquark, and baryons which are the bound states of three quarks or three antiquarks. The structures of hadrons are shown in figure 1.2.

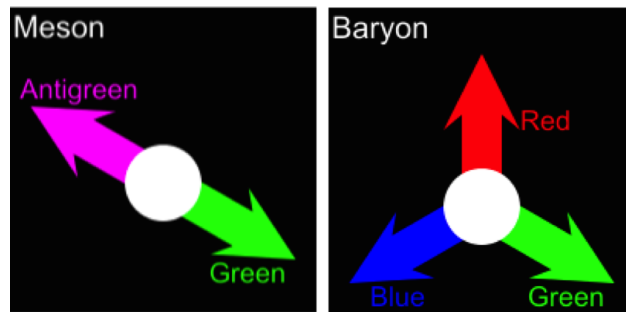


Figure 1.2: The diagram of established Hadrons:  $q\bar{q}$  Meson (left), Baryon (right) [4].

Each lepton doublet contains one charged lepton with electric charge  $-Q_e$  and weak isospin  $-1/2$ , and one neutral lepton known as neutrino with weak isospin  $+1/2$ . Charged leptons can combine with other particles to form composite bound states, while neutrinos rarely interact with the rest of the particles, and are exceptionally difficult to detect. In contrast to quarks, leptons are colorless, thus leptons do not undergo strong interactions. However leptons have leptonic numbers which are conserved under the standard model<sup>1</sup>. The first generation (electron and electron neutrino) has an electronic number  $+1$ , the second (muon and muon neutrino) has a muonic number  $+1$ , the third (tau and tau neutrino) has a tauonic number

<sup>1</sup>Neutrino oscillations are known to violate the conservation of the individual leptonic numbers. Such a violation is beyond the original standard model. Therefore, nowadays the standard model has been extended to include this phenomenon.



+1. In the original standard model, the neutrinos were assumed to be massless. However, the experiments of neutrino oscillations have indicated that neutrinos have some very small mass.

### 1.1.2 Standard Model Interactions

There are four fundamental interactions in nature: weak interaction, strong interaction, electromagnetism and gravitation, the first three interactions are described by the standard model while we do not yet have a good theory of quantum gravity. The elementary particles interact with each other by the exchange of gauge bosons which play the role of force carriers. The four fundamental interactions are summarized in table 1.2.

*Table 1.2: List of the fundamental interactions.*

Interaction	Relative Strength	Force Carrier		
		Name	Symbol	Mass (GeV)
Strong	$10^{38}$	Gluon	$g$	0
Electromagnetism	$10^{36}$	Photon	$\gamma$	0
Weak	$10^{25}$	W bosons	$W^\pm$	80.4
		Z Boson	$Z^0$	91.2
Gravitation	1	Graviton (hypothetical)	$G$	0

The strong interaction, which is also known as color force, happens through the exchange of a massless gluon which is a vector gauge boson. The strong interaction is the force that holds quarks together to form hadrons such as protons and neutrons. It is also the force that binds the protons and neutrons together to form atomic nuclei. These interactions between quarks and gluons are described by quantum chromodynamics (QCD). The weak interaction is mediated by vector bosons:  $W^\pm$  and  $Z^0$ . Due to their large masses,  $W^\pm$  and  $Z^0$  bosons are short-lived elementary particles and therefore the weak interaction does not act at large distance. The weak interaction is the only process that couples to all the fermions in the standard model and causes certain forms of radioactive decay such as a quark changing into

another quark, or a lepton to another lepton by emission & absorption of the  $W^\pm$  bosons. The electromagnetic force is mediated by the massless photon and acts on electrically-charged particles with infinite range. The interaction is explained by quantum electrodynamics (QED). The electromagnetism and the weak interaction are unified into a single electroweak force above the unification energy (100 GeV). The interactions between elementary particles in the standard model are illustrated in figure 1.3.

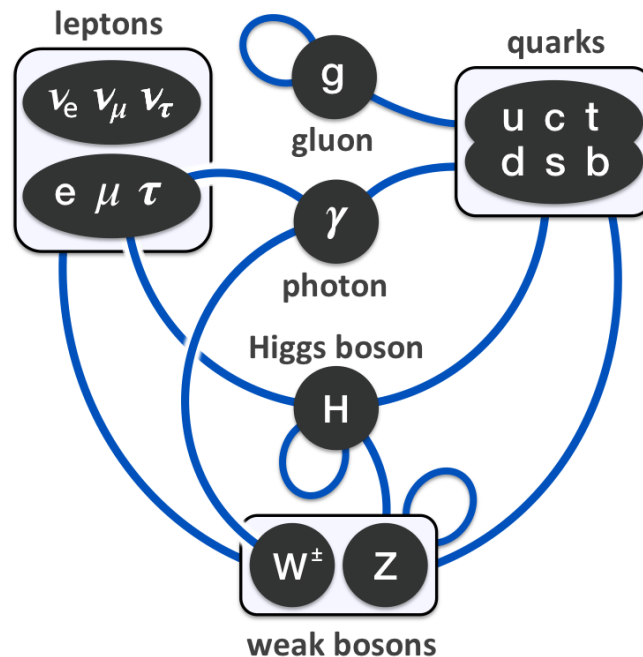


Figure 1.3: The interactions between elementary particles of standard model [3].

The symmetry of electroweak interaction requires the gauge bosons to be massless, but the weak interactions' gauge bosons ( $W^\pm$  and  $Z^0$ ) have been experimentally observed with large masses. This phenomenon is explained by Higgs mechanism: symmetry could be “broken” if a special field (Higgs field) happened to exist and the gauge bosons would receive masses. The observation of Higgs boson [1, 2], which is a matching particle associated with the Higgs field, has proven the existence of the Higgs field.

## 1.2 Charmonium States

Charmonium is a bound state of a charmed quark and its antiquark ( $c\bar{c}$ ). Due to the similar structure to positronium, charmonium is expected to have similar excited states. The first charmonium state,  $J/\psi$  was simultaneously observed at BNL [5] and SLAC [6] on November 11, 1974. Ten days later the second charmonium resonance,  $\psi(2S)$  was found [7]. However, unlike positronium, which is bound by the Coulomb force, the quark pair is tied by the strong interaction which determines the properties of charmonium. Thus analysis of the charmonium is a way to study the strong interaction and hadron dynamics.

Two methods can be used to predict the mass spectrum of charmonium states: theoretical and phenomenological. The theoretical approach intends to calculate the masses from QCD (e.g., direct lattice QCD calculation). The phenomenological approach focuses on modeling the prospective states with effective theory, like e.g. the potential approach. A more detailed description of the two methods can be found in Ref. [8].

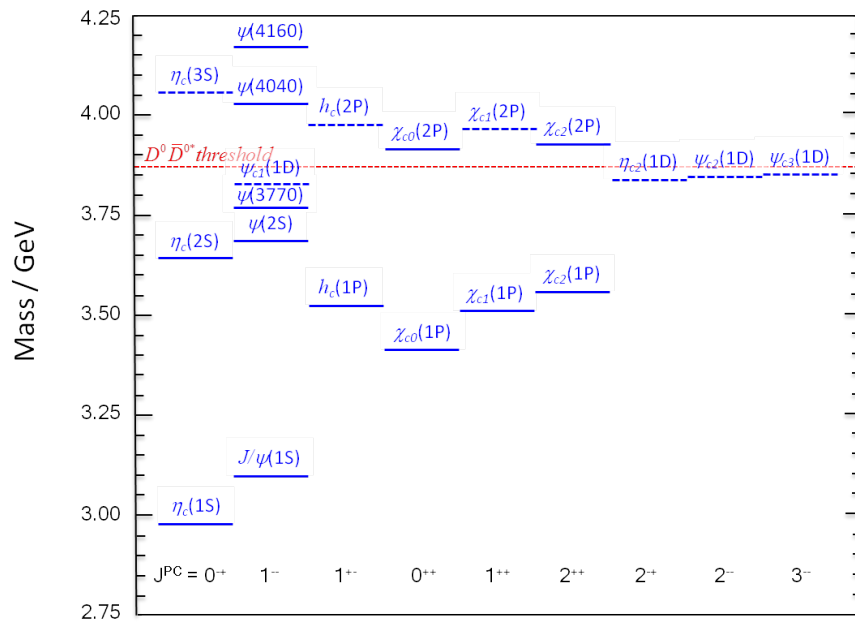


Figure 1.4: The mass spectrum of charmonium states. The observed states are labeled with solid horizontal line.

Figure 1.4 shows the predicted mass spectrum of charmonium states. The states are characterized by the radial quantum number  $n$ , the orbital angular momentum,  $L = 0, 1, 2, 3, \dots$  (values denoted as  $S, P, D, F, \dots$ ), the total spin of the quark pair,  $S = 0$  or  $1$  and the total angular momentum,  $J = L \oplus S$ . The spectroscopic notation is:

$$(n + 1)^{2S+1} L_J$$

The name assigned to each state follows the convention of Particle Data Group (PDG) (see table 1.3).

Table 1.3: Names of the states in charmonium spectrum.

Symbol	$L$	$S$
$\eta$	even	0 (singlet)
$\psi$	even	1 (triplet)
$h$	odd	0 (singlet)
$\chi$	odd	1 (triplet)

In the mass spectrum, every state is labeled by  $J^{PC}$ , where  $P$  and  $C$  are parity symmetry and charge conjugation symmetry, respectively. The quantum number  $P$  is determined by  $L$ :  $P = (-1)^{L+1}$ ; the other quantum number  $C$  is determined by  $L$  and  $S$  combined:  $C = (-1)^{L+S}$ . For example, the first observed charmonium state  $J/\psi$  ( $1^3S_1$ ) and the second  $\psi(2S)$  ( $2^3S_1$ ) have the same quantum numbers,  $J^{PC} = 1^{--}$ . So both of them can be produced as resonances in  $e^+e^-$  annihilation. The isospin of the charmonium state is always 0, because there are no light quarks in the  $c\bar{c}$ .

### 1.3 The $X(3872)$

The expected charmonium spectrum is composed of all  $c\bar{c}$  resonances observed below the threshold for decays to  $D\bar{D}$  meson pairs successfully predicted by QCD-based calculations.

Many states with masses above  $2m_D$  have been observed as well. Most of them are in agreement with the expected charmonium spectrum. However, there are some observed states which do not fit these predictions. Among them, the  $X(3872)$  state was the first to be discovered.

### 1.3.1 The Discovery of A Narrow $X(3872)$ State

A narrow state was discovered in<sup>2</sup>  $B^+ \rightarrow X(3872)K^+$ ,  $X(3872) \rightarrow \pi^+\pi^-J/\psi$ ,  $J/\psi \rightarrow l^+l^-$  decays ( $l = e$  or  $\mu$ ) by the Belle experiment in 2003 [9] (see figure 1.5). Distribution of invariant  $\pi^+\pi^-$  mass is consistent with isospin violating  $X(3872) \rightarrow \rho^0 J/\psi$  decay. The mass and width of the particle were measured to be

$$M_{X(3872)} = 3872.0 \pm 0.6_{stat} \pm 0.5_{syst} \text{ MeV},$$

and

$$\Gamma_{X(3872)} < 2.3 \text{ MeV}.$$

The existence of the  $X(3872)$  was confirmed by the CDF [10], DØ[11] and BaBar [12] experiments. It has also been observed by LHCb in prompt production in  $p\bar{p}$  collisions [13] (see figure 1.6). More than 500  $X(3872)$  events were reconstructed from  $X(3872) \rightarrow J\psi\pi^+\pi^-$ , and the mass of  $X(3872)$  was measured.

The masses observed by these experiments give rise to the world average value of the  $X(3872)$ 's mass [14]:

$$M_{X(3872)} = 3871.68 \pm 0.17 \text{ MeV}.$$

---

<sup>2</sup>In this thesis the charge conjugate is always implied, unless explicitly stated.

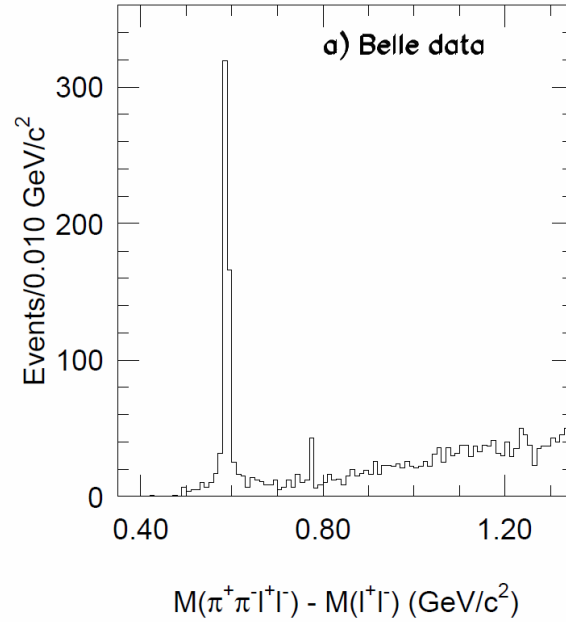


Figure 1.5: Distribution of  $M(\pi^+\pi^-J/\psi) - M(J/\psi)$  observed by Belle. In addition to the well-known  $\psi(2S)$  (peaking at  $\approx 0.6$  GeV), a second peak was seen at  $\approx 0.8$  GeV.

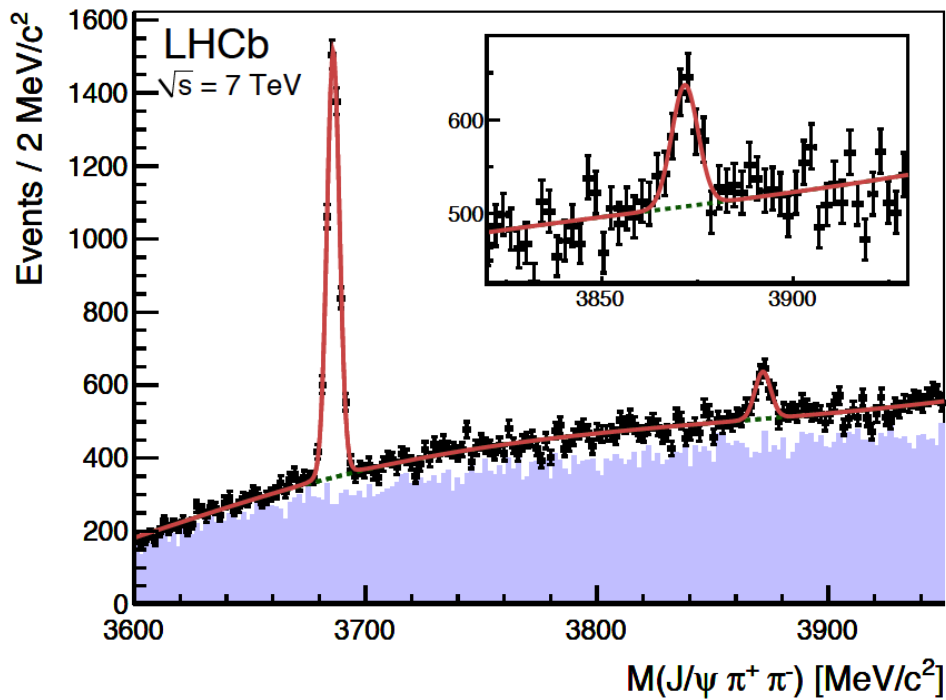


Figure 1.6: Distribution of  $M(J/\psi\pi^+\pi^-)$  observed by LHCb. The masses of  $\psi(2S)$  (left) and the  $X(3872)$  (right) were measured to be  $3686.12 \pm 0.06_{stat} \pm 0.10_{syst}$  MeV and  $3871.95 \pm 0.48_{stat} \pm 0.12_{syst}$  MeV, respectively.

### 1.3.2 The Nature of the $X(3872)$ : Theoretical Interpretations

The final state of the decay  $X(3872) \rightarrow \pi^+\pi^- J/\psi$  contains the first discovered charmonium state  $J/\psi$ , so it is safe to assume that  $X(3872)$  contains a  $c\bar{c}$  pair inside. However the mass of the  $X(3872)$  does not fit any expected charmonium state. Investigation of the nature of this state has received extensive theoretical and experimental attention.

The predicted and observed charmonium states[8] near the  $X(3872)$  mass are listed in table 1.4.

Table 1.4: Charmonium states near the mass of  $X(3872)$ . The masses of predicted but unobserved states have no errors assigned.

State	$J^{PC}$	Mass (MeV)
$h_c(2P)$	$2^1P_1$	$1^{+-}$ $\sim 3956$
$\chi_{c0}(2P)$	$2^3P_0$	$0^{++}$ $3918.4 \pm 1.9$
$\chi_{c1}(2P)$	$2^3P_1$	$1^{++}$ $\sim 3953$
$\chi_{c2}(2P)$	$2^3P_2$	$2^{++}$ $3927.2 \pm 2.6$
$\eta_{c2}(1D)$	$1^1D_2$	$2^{-+}$ $\sim 3837$
$\psi_{c1}(1D)$	$1^3D_1$	$1^{--}$ $\sim 3819$
$\psi_{c2}(1D)$	$1^3D_2$	$2^{--}$ $\sim 3838$
$\psi_{c3}(1D)$	$1^3D_3$	$3^{--}$ $\sim 3849$

The predicted masses of  $2^1P_1$  and  $2^3P_1$  are  $\sim 80$  MeV higher than the mass of the  $X(3872)$ . The observed  $2^3P_0$  mass which should be lower than  $2^3P_1$  mass is still  $\sim 50$  MeV above the  $X(3872)$  mass. The  $2^3P_0$  and  $2^3P_2$  states have broad widths which are  $20 \pm 5$  MeV and  $24 \pm 6$  MeV, respectively [14]. However  $2^3P_1$  is expected to be narrower. The width of the  $X(3872)$  is less than 1.2 MeV at 90% confidence level [14]. The predicted masses of  $1^1D_2$ ,  $1^3D_1$ ,  $1^3D_2$  and  $1^3D_3$  are  $\sim 20$  MeV lower than the  $X(3872)$ . The singlet  $1^1D_2$  ( $J^{PC} = 2^{-+}$ ) state could be narrow and is the closest in mass to the  $X(3872)$  state.

Recently LHCb measured the mass of the  $D^0$  and calculated the mass difference between

$D^0\bar{D}^{0*}$  and the  $X(3872)$  [15]:

$$M_{D^0\bar{D}^{0*}} - M_{X(3872)} = 0.16 \pm 0.26 \text{ MeV}.$$

Due to the proximity of the  $X(3872)$ 's mass to the  $D^0\bar{D}^{0*}$  threshold, the  $X(3872)$  has been suggested to be a loosely bound deuteron-like  $D^0\bar{D}^{0*}$  molecule [16]. A  $D^0\bar{D}^{0*}$  molecule would explain the apparent breaking of isospin symmetry in the decay channel  $X(3872) \rightarrow J/\psi\rho^0$ . The molecular explanation requires that the  $J^{PC}$  of the  $X(3872)$  has to be  $0^{-+}$  or  $1^{++}$  [16].

Other theoretical proposals use tetraquark model to explain  $X(3872)$  [17]. A tetraquark is a “diquark-antidiquark” system bound by the color forces of gluon exchange, where each diquark is built up with a light quark ( $u$  or  $d$ ) and a heavy quark ( $c$  or  $s$ ). The interpretation of tetraquark model implies that two neutral  $X(3872)$  mesons should be observed. A quark model can accommodate the  $X(3872)$  mass with  $J^{PC} = 1^{++}$  assignment [18].

### 1.3.3 The Early Measurements of the $X(3872)$ Quantum Numbers

In order to understand the origin of the  $X(3872)$ , measurement of its spin and parity is crucially important.

In 2005, Belle observed the decay mode  $X(3872) \rightarrow \gamma J/\psi$  with a statistical significance which is greater than  $4\sigma$ , and demonstrated that the  $X(3872) \rightarrow \pi^+\pi^-\pi^0 J/\psi$  decay was consistent with the  $X(3872) \rightarrow \omega J/\psi$  decay from the  $\pi^+\pi^-\pi^0$  invariant mass distribution [19]. These results demonstrate that the charge conjugation parity of the  $X(3872)$  is  $+1$  ( $C_{X(3872)} = C_{J/\psi}C_{\gamma/\omega}$ ).

The CDF experiment analyzed the three-body decay channel  $X(3872) \rightarrow \pi^+\pi^- J/\psi$ , where  $J/\psi \rightarrow \mu^+\mu^-$  [20]. The  $X(3872)$  events were reconstructed, with majority of the signal events coming from the prompt production in  $p\bar{p}$  collisions. They used the helicity formalism, in which



the  $X(3872)$  decay is described by three helicity angles of the  $X(3872)(\theta_X)$ , the  $J/\psi(\theta_{J/\psi})$ , the  $\pi\pi$  system ( $\theta_{\pi\pi}$ ), and two azimuthal angles between the decay planes of the  $X(3872)$  and its two daughters,  $\phi_X - \phi_{J/\psi}$  and  $\phi_X - \phi_{\pi\pi}$ . CDF reconstructed the events coming from both the B meson decay and the  $p\bar{p}$  interaction. This gave them a large signal statistics,  $2292 \pm 113$  events, but the  $J^{PC}$  sensitivity suffered from the lack of  $X(3872)$  polarization and from the large background subtraction. Due to the lack of  $X(3872)$  polarization, only  $\cos\theta_{J/\psi}$ ,  $\cos\theta_{\pi\pi}$  and  $\phi_{J/\psi} - \phi_{\pi\pi}$  were useful in  $J^{PC}$  analysis. In the  $X(3872) \rightarrow \pi^+\pi^-J/\psi$  decay, there is more than one combination of L (total orbital angular momentum) and S (total spin) to form J (total angular momentum). The minimal L was assumed to be dominant. CDF performed a fit of different  $J^{PC}$  hypotheses to 3-dimensional (3D) binned distribution in these angles. Out of 13 possible  $J^{PC}$  configurations, they ruled out 11 spin-parity assignments except for two  $C$ -parity even candidates,  $J^{PC}$  equal to  $1^{++}$  or  $2^{-+}$ , which could not be resolved (see table 1.5). Most likely exotic explanations of the  $X(3872)$ , molecular or tetraquark models, demand it to be a  $1^{++}$  state. The  $2^{-+}$  hypothesis would favor  $1^1D_2$  charmonium interpretation.

The Belle collaboration analyzed the angular distributions in  $173 \pm 16$   $X(3872)$  events in  $B \rightarrow X(3872)K$  ( $K = K^\pm$  or  $K_S^0$ ), where  $X(3872) \rightarrow \pi^+\pi^-J/\psi$ ,  $J/\psi \rightarrow l^+l^-$  [21]. Reconstruction of this full decay chain results in a small background and polarized  $X(3872)$ . Thus, all 5 angles carry useful information for  $J^{PC}$  determination. The Belle collaboration decided they did not have enough signal statistics to bin them in many dimensions, thus they analyzed only 1-dimensional distributions in 3 different angles. They concluded that their data were equally well described by the  $1^{++}$  hypothesis and by the  $2^{-+}$  hypothesis with a complex parameter  $\alpha$  being equal to  $0.64 + i0.27$ . The latter is defined as  $B_{11}/(B_{11} + B_{12})$ , where  $B_{LS}$  are the two possible amplitudes for the  $2^{-+}$  hypothesis in which the lowest orbital angular  $L$  between the  $\pi\pi$  and  $J/\psi$  is assumed to dominate ( $L = 1$ ).  $S$  is the sum of the  $J/\psi$  spin (1) and of the  $\pi\pi$  spin (also 1) and can be equal 1 or 2 in this case.

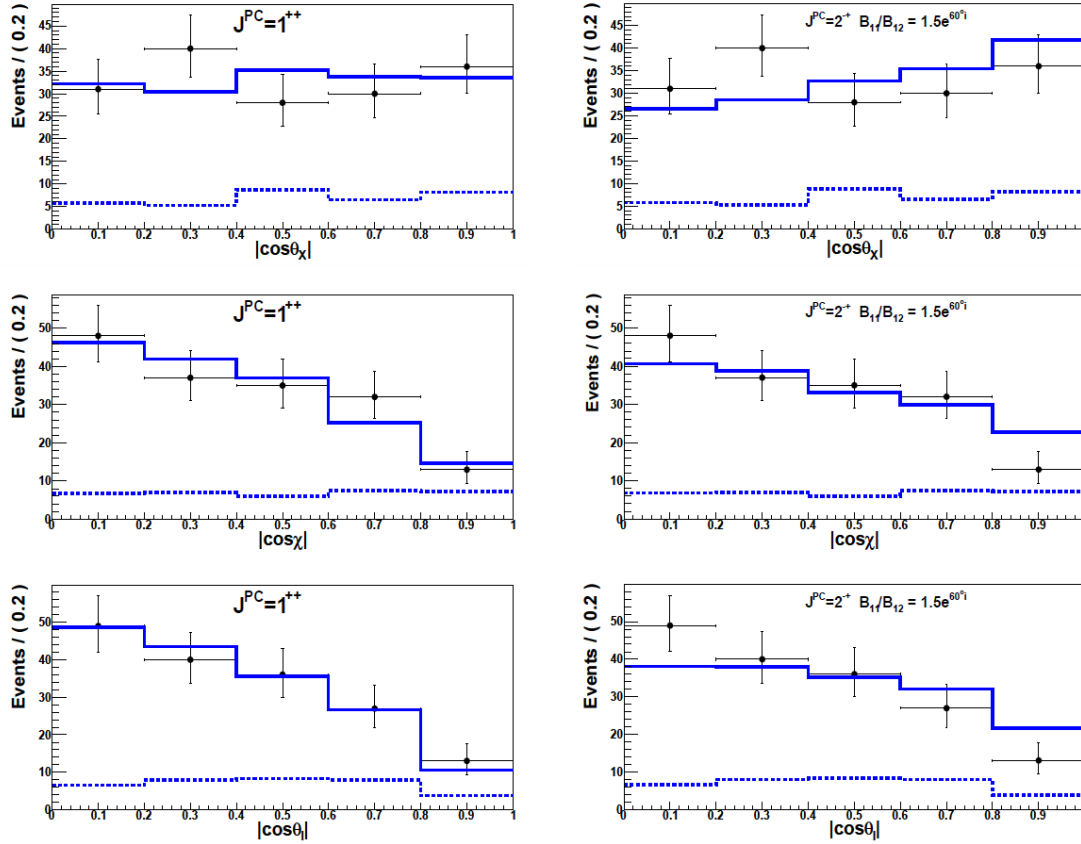


Figure 1.7: The comparisons of  $1^{++}$  and  $2^{-+}$  hypotheses. The solid black points are the data; the dashed blue histogram indicates the background determined from the events in the scaled  $M(J/\psi\pi^+\pi^-)$  sidebands; the solid blue histogram is the sum of the simulated MC  $X(3872) \rightarrow J/\psi\rho$  events generated with  $1^{++}$  (left) /  $2^{-+}$  (right) hypothesis and the background (dashed histogram) and normalized to the observed signal.

Table 1.5: The result of  $X(3872)$  angular analysis obtained by CDF. From left hand side to right hand side, the five columns are state, the decay mode, the  $L$  and  $S$  quantum numbers of  $J/\psi(\pi\pi)$  system, the  $\chi^2$  with 11 degrees of freedom and the  $\chi^2$  probability.

$J^{PC}$	decay	$LS$	$\chi^2$ (11 d.o.f.)	$\chi^2$ prob.
$1^{++}$	$J/\psi\rho^0$	01	13.2	0.28
$2^{-+}$	$J/\psi\rho^0$	11,12	13.6	0.26
$1^{--}$	$J/\psi(\pi\pi)_S$	01	35.1	$2.4 \times 10^{-4}$
$2^{+-}$	$J/\psi(\pi\pi)_S$	11	38.9	$5.5 \times 10^{-5}$
$1^{+-}$	$J/\psi(\pi\pi)_S$	11	39.8	$3.8 \times 10^{-5}$
$2^{--}$	$J/\psi(\pi\pi)_S$	21	39.8	$3.8 \times 10^{-5}$
$3^{+-}$	$J/\psi(\pi\pi)_S$	31	39.8	$3.8 \times 10^{-5}$
$3^{--}$	$J/\psi(\pi\pi)_S$	21	41.0	$2.4 \times 10^{-5}$
$2^{++}$	$J/\psi\rho^0$	02	43.0	$1.1 \times 10^{-5}$
$1^{-+}$	$J/\psi\rho^0$	10,11,12	45.4	$4.1 \times 10^{-6}$
$0^{-+}$	$J/\psi\rho^0$	11	104	$3.5 \times 10^{-17}$
$0^{+-}$	$J/\psi(\pi\pi)_S$	11	129	$\leq 1 \times 10^{-20}$
$0^{++}$	$J/\psi\rho^0$	00	163	$\leq 1 \times 10^{-20}$

The BaBar experiment observed  $34 \pm 7$   $X(3872)$  events in  $X(3872) \rightarrow \omega J/\psi$ , where  $\omega \rightarrow \pi^+\pi^-\pi^0$  [22]. The observed distribution of  $\omega$  mass was more consistent with being shaped by the angular momentum barrier factor for the  $2^{-+}$  hypothesis (CL= 68%) than for the  $1^{++}$  hypothesis, but the latter could not be ruled out (CL= 7%).

Analysis of 1-dimensional distributions of the 5 angles describing the  $B^+ \rightarrow X(3872)K^+$ ,  $X(3872) \rightarrow \pi^+\pi^- J/\psi$ ,  $J/\psi \rightarrow \mu^+\mu^-$  decay chain was simulated in the LHCb [23, 24] detector before. Only real values of  $\alpha$  were considered. It was concluded that with a few hundred of events it would be possible to distinguish the  $1^{++}$  hypothesis and the  $2^{-+}$  hypotheses with  $\alpha = 1.0$  using  $\cos\theta_X$  distribution; however, the  $2^{-+}$  with  $\alpha = 0.5$  and  $1^{++}$  hypotheses could not be distinguished using this distribution alone.

In this thesis, we implement 5-dimensional analysis in the full angular phase space for  $B^+ \rightarrow X(3872)K^+$ ,  $X(3872) \rightarrow \pi^+\pi^- J/\psi$ ,  $J/\psi \rightarrow \mu^+\mu^-$  decay, which allows us to determine the  $J^{PC}$  of the  $X(3872)$  with no ambiguity for the first time.



# Chapter 2

## The LHCb Experiment

This chapter introduces the experiment of LHCb, standing for Large Hadron Collider beauty, which focuses on the study of heavy flavor physics. LHCb detector is located on the Large Hadron Collider (LHC) ring at CERN. The LHCb experiment is designed to precisely measure the parameters of CP violation and rare decays of beauty and charm hadrons from proton-proton collisions.

### 2.1 The CERN

CERN (European Organization for Nuclear Research) was founded in 1954 by twelve European countries. The name CERN came from French, stood for Conseil Européen pour la Recherche Nucléaire (European Council for Nuclear Research). The acronym remains even though the name of the council changed to Organisation Européenne pour la Recherche Nucléaire. As one of Europe's first joint ventures, CERN was established to set up a laboratory for world-class fundamental physics research. In the early years, the research concentrated on the atomic nucleus. With a better understanding of the inside of the atom, the main area of work shifted to study the fundamental constituents of matter and the forces acting between them.

Several important achievements have been obtained in the experiments at CERN, among

them:

- 1965, the first observations of antinuclei.
- 1973, the discovery of neutral currents in Gargamelle bubble chamber.
- 1983, the discovery of the  $W^\pm$  and  $Z^0$  particles.
- 1989 - 2000, precision studies of  $Z^0$  formed in  $e^+e^-$  collisions.
- 2012, a new particle with mass around  $125 \sim 126$  GeV consistent with long-sought Higgs boson.

CERN is also known to be the birthplace of the World Wide Web.

Now CERN has more than 100 nations and 600 universities as members. Its main function nowadays is operating the world's largest and most powerful particle collider - LHC.

## 2.2 The LHC

The LHC, Large Hadron Collider, near Geneva is the highest energy collider in the world. It is a two-ring-superconducting-hadron accelerator and collider built in the old Large Electron-Positron Collider (LEP) tunnel which is located between 45 meters and 170 meters underground (see figure 2.1). Four major experiments are located along the 27 kilometers LHC circle:

- ATLAS (A Toroidal LHC ApparatuS), is one of the two general purpose detectors with central angular coverage. The experiment is designed to search for highly massive particles e.g. Higgs boson and direct searches for new physics.
- CMS (Compact Muon Solenoid), is the other general purpose detector. The goal of this experiment is the same as that of ATLAS.

- ALICE (A Large Ion Collider Experiment), measures the properties of quark-gluon plasma produced in heavy ion collisions.
- LHCb (LHC Beauty), focuses on the measurements of CP violation in b-physics experiments. It is also used to study heavy flavor spectroscopy. Other studies include measurements of production cross sections and electroweak physics in the forward region. More details of the LHCb experiment are introduced in section 2.3.

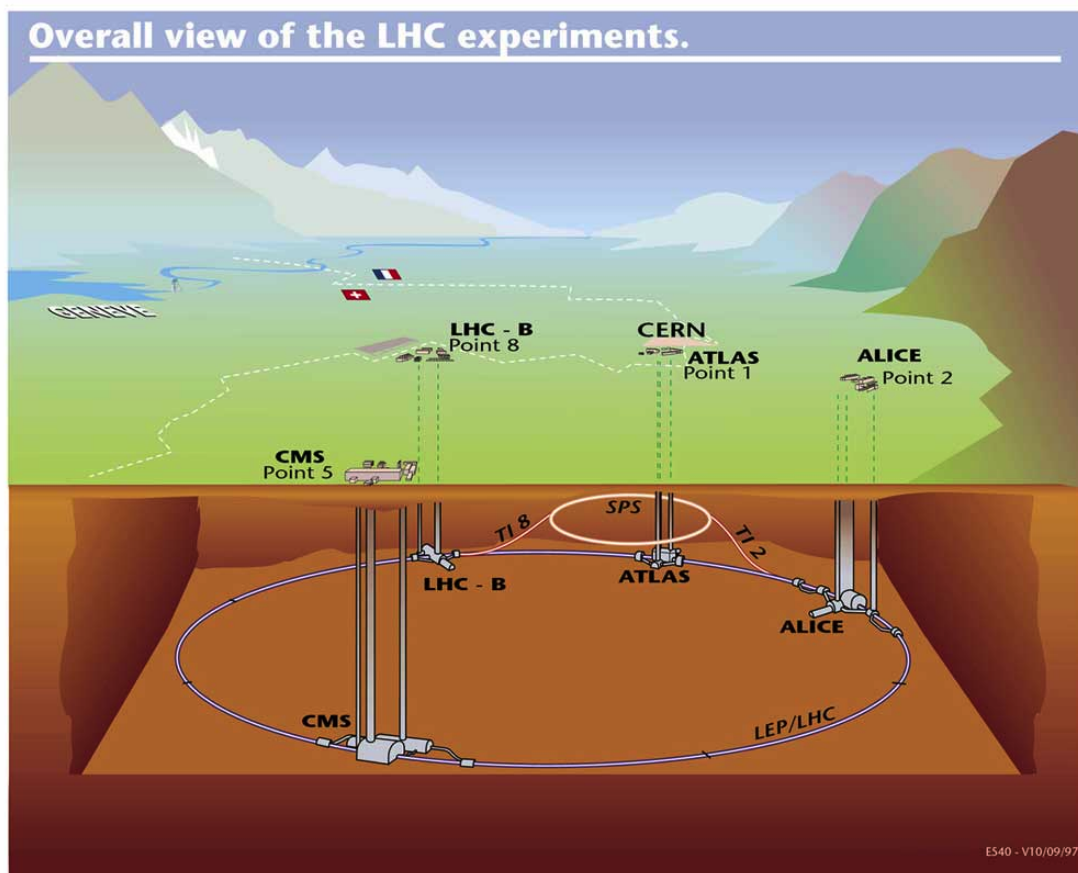


Figure 2.1: An overall view of the LHC experiments.

Each experiment receives two proton beams, traveling at close to the speed of light, directed in opposite paths. The collision of the high-energy particle beams, at center-of-mass energy  $\sqrt{s} = 7 - 8 \text{ TeV}$ <sup>1</sup> which is the highest energy achieved in the world's particle colliders.

<sup>1</sup>After 18-months shutdown, the data taking will be resumed in 2015 with center-of-mass energy of 13 - 14 TeV

Figure 2.2 shows the injection chain of the LHC. The starting point in this chain is the linear particle accelerator (Linac2), which replaced Linac1 in 1978 and will be replaced by Linac4 in 2017 or 2018. The hydrogen gas, which is stored in a bottle at one end of Linac2, passes through an electric field to produce protons. When the protons arrive at the other end of Linac2, the energy of protons reaches 50 MeV and the mass increases by 5%. To meet the LHC's requirement, linear accelerator produces a beam of 180 mA with pulse length of 100  $\mu$ s. Then the beam enters the Proton Synchrotron Booster (PSB), which is made up of four superimposed synchrotron rings. The original PS Booster was designed and built in the 1960s and then upgraded to allow the beam to be accelerated to 1.4 GeV instead of 800 MeV. The protons are moved by the four PSB rings and recombined into one string of bunches and transferred to the Proton Synchrotron (PS).

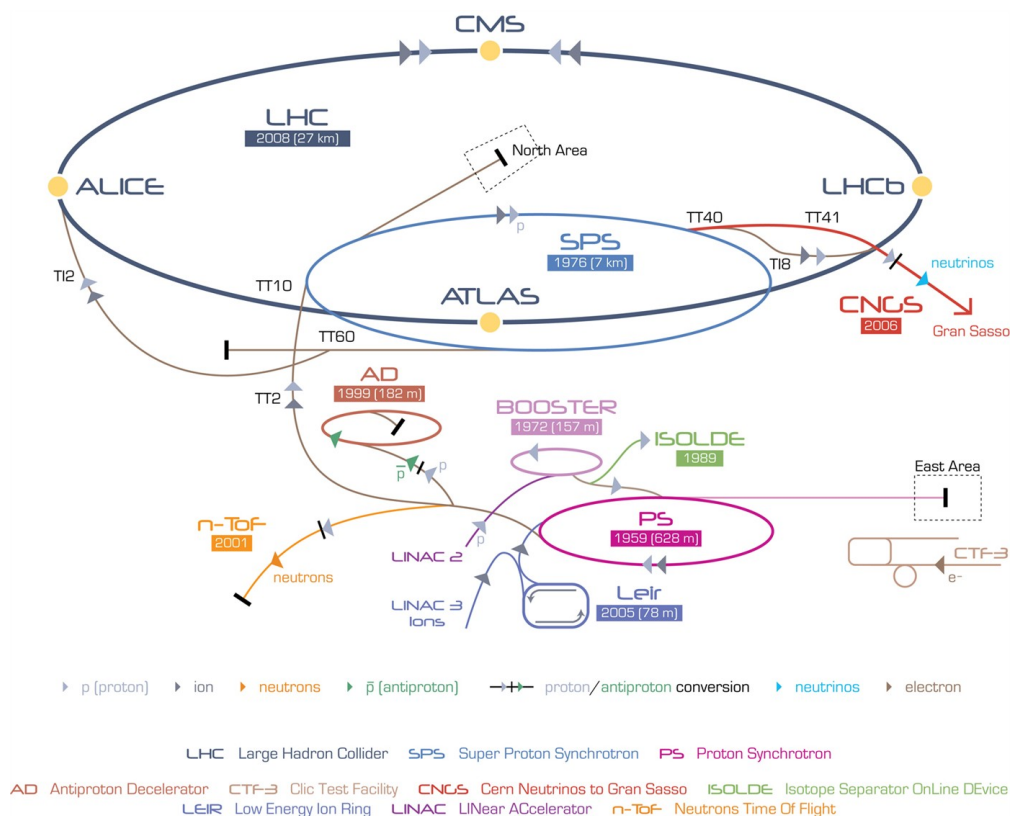


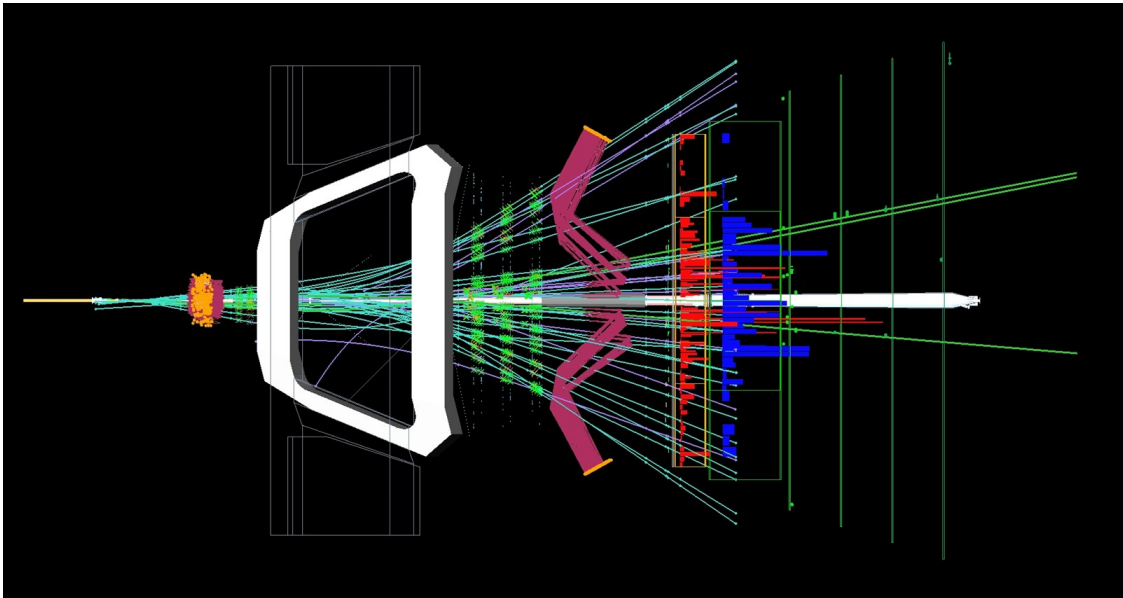
Figure 2.2: The LHC's injection chain.

As an old particle accelerator, shooting out the first beam in 1959, the PS still plays a critical role in the CERN's accelerator complex. It is a circle accelerator with a circumference of 628



meters. 277 conventional electromagnets bend and accelerate the beam round the ring. There are 6 PSB bunches, each bunch contains  $1.15 \times 10^{11}$  protons, captured in the PS. The protons are accelerated to an energy of 25 GeV and sent to Super Proton Synchrotron (SPS) with a bunch spacing of 25 ns and 72 bunches per batch. The SPS is the second-largest accelerator in the CERN. The protons are accelerated to an Energy of 450 GeV in the SPS. After proper synchronization, the protons are injected into the LHC.

The transfer lines TI 2 brings the beam towards ALICE , and TI 8 delivers the beam towards LHCb. The two beams are running in opposite directions and colliding at each experiment. Figure 2.3 shows the first  $p\bar{p}$  collision at  $\sqrt{s} = 3.5$  TeV at the LHCb detector. More detailed description of the process of LHC beam can be found in Ref. [25].



*Figure 2.3: The first  $p\bar{p}$  collision at  $\sqrt{s} = 3.5$  TeV at the LHCb detector, 2010.*

## 2.3 The LHCb

Large Hadron Collider beauty, LHCb, experiment was approved in 1996 at the LHC. The experiment is primarily designed to look for indirect evidence of new physics in CP violation and rare decays of the particles containing either beauty or charm quarks. Figure 2.4 shows

the photo of LHCb collaboration in front of the LHCb detector.



Figure 2.4: LHCb collaboration in front of LHCb detector.

### 2.3.1 Introduction

The LHCb detector is a single arm forward spectrometer, where the collisions happen at the front of the detector. The polar angular coverage is from approximately 10 milliradians (mrad) to 300 mrad in the horizontal bending plane and 250 mrad in the vertical plane. The layout of the LHCb detector is shown in figure 2.5. The LHCb possesses a unique pseudorapidity<sup>2</sup> range:  $2.0 < \eta < 4.5$ <sup>3</sup>. Thus, the LHCb can exploit the production of the  $b\bar{b}$ , which is flying along the axis of the beams, from  $p\bar{p}$  collisions as illustrated in figure 2.6.

<sup>2</sup>The pseudorapidity is defined as:

$$\eta = \frac{1}{2} \ln \frac{E + P_z}{E - P_z},$$

where  $E$  is the energy of the particle and  $P_z$  is the component of the particle momentum along the z-axis (LHCb system uses a right-handed coordinate: the x goes through figure 2.5, the y is directed away from the Earth and the z is parallel to the beam).

<sup>3</sup>In the real detector a small amount of events can be reconstructed outside this range, the real range of pseudorapidity is  $1.9 < \eta < 4.9$ .

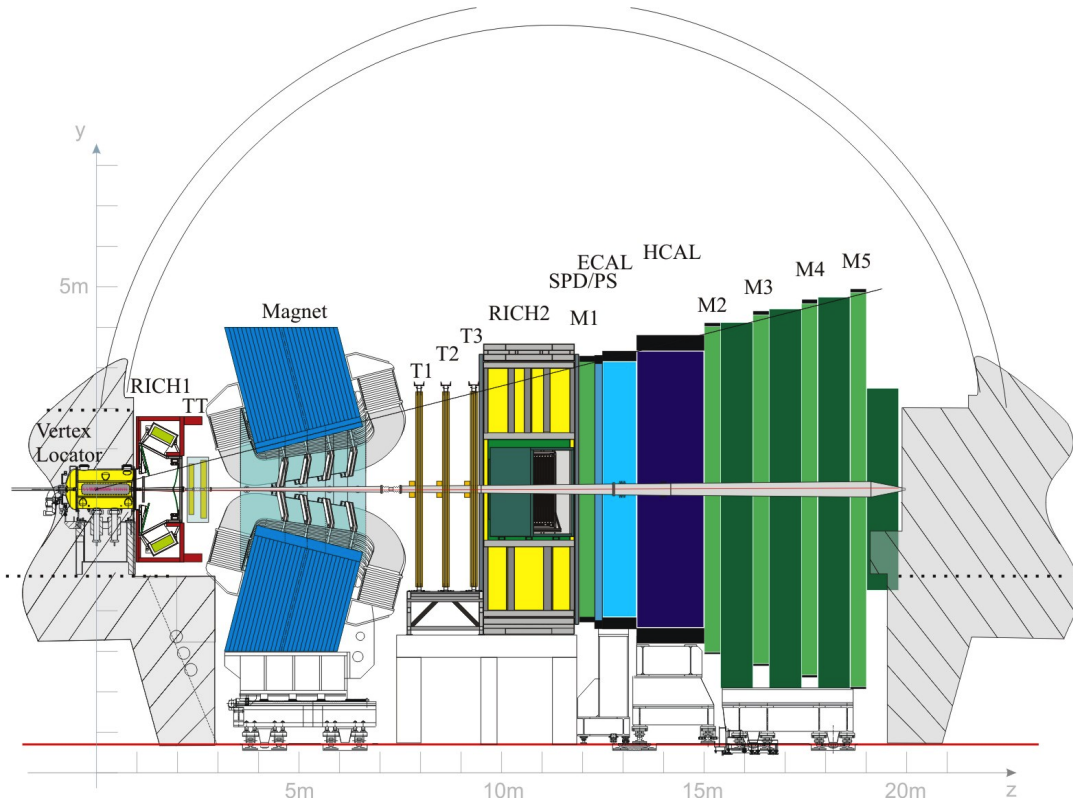


Figure 2.5: Side view of LHCb detector.

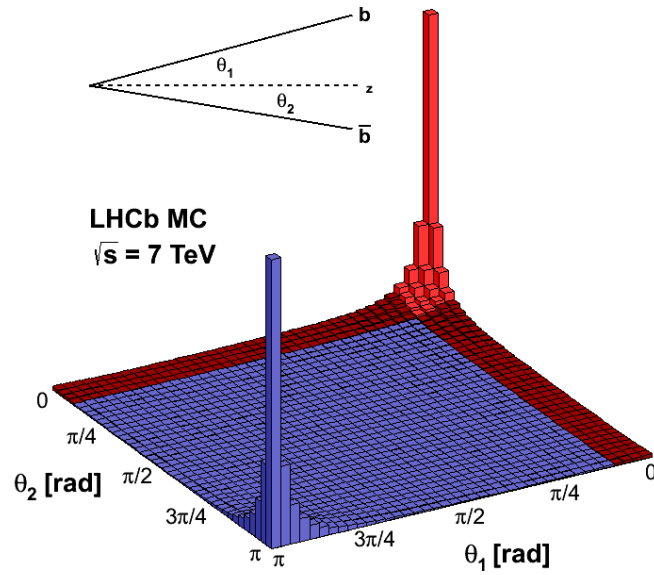


Figure 2.6: The polar angle distribution with respect to the beam line ( $b$  and  $\bar{b}$ ) at  $\sqrt{s} = 7 \text{ TeV}$ .

The LHCb system consists of several sub-detectors:

- VERTex LOcator (VELO) and four planar tracking stations belong to the tracking system (see section 2.3.3).
- Two Ring Imaging CHerenkov detectors (RICH1 and RICH2), Calorimeters (ECAL and HCAL) and Muon chambers build up the particle identification system (see section 2.3.5).

In addition to the detector sensors, electronic equipment is another important part of the LHCb experiment (see section 2.3.4). The front-end electronics are installed on or close to the sub-detectors, while the readout and trigger electronics are placed in a radiation protected counting house [26].

## 2.3.2 Magnet

The dipole magnet is located between TT and T1-T3 stations. Together they are used to measure the momentum of charged particles with a precision of about 0.4% for momenta up to 200 GeV [27]. Charged particles are bent in the magnet by the strong magnetic field<sup>4</sup> when they travel through the detector. The magnet field is created by a warm magnet with saddle-shaped coils in a window-frame yoke with sloping poles. Compared to a superconducting magnet, a warm magnet has several advantages: significantly lower cost, fast construction, low risks and rapid ramping-up of the field. The structure of the dipole magnet is illustrated in figure 2.7.

The two trapezoidal coils are bent at 45° and situated mirror-symmetrically to each other inside an iron yoke. The trapezoidal shape is chosen such that the two bent coils do not touch each other at their smaller gap side but are still relatively close to each other at their larger gap side [27]. Each coil is composed of 15 individual mono-layer pancakes. The two coils have

---

<sup>4</sup>The magnet is able to generate a 4 Tm integrated field for tracks originating near the interaction region.

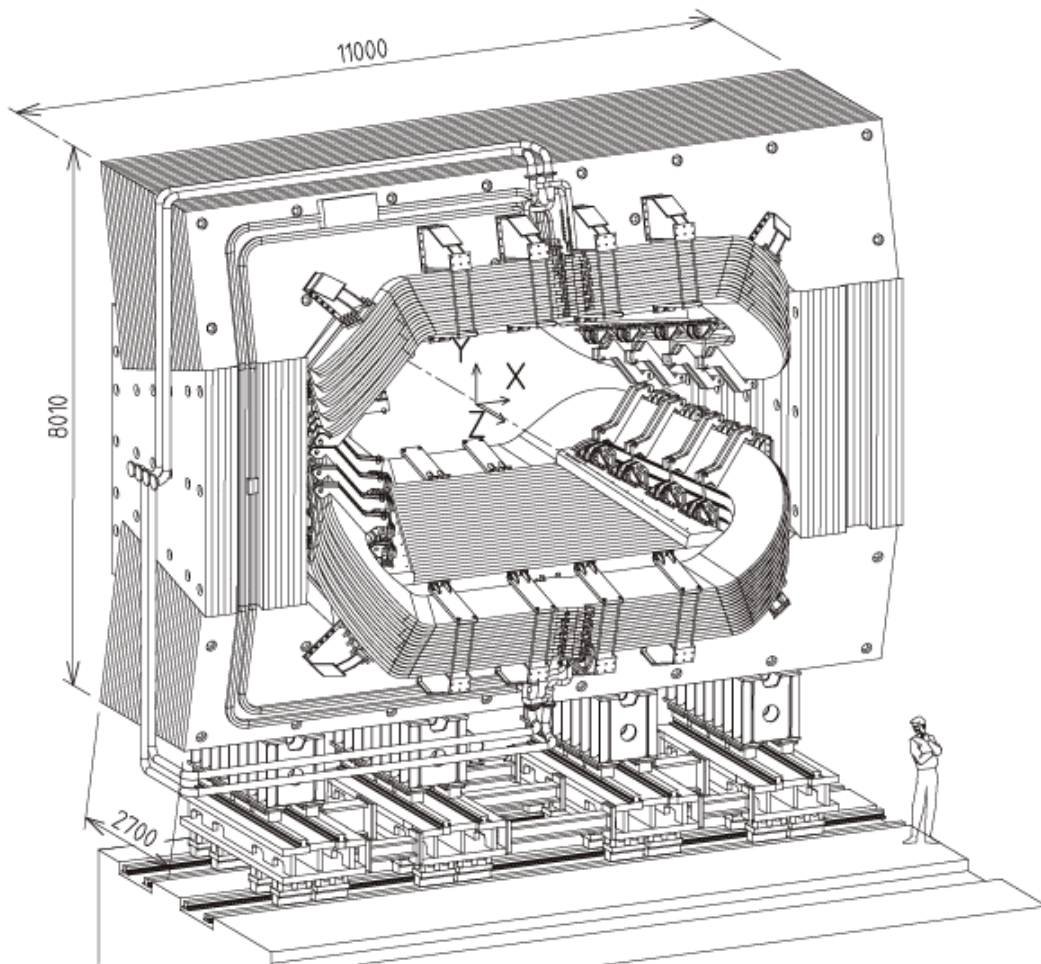


Figure 2.7: The LHCb dipole magnet (units in mm).



a total weight of 54 tons. Within the upper and lower coils is the water cooling channel of 25 mm diameter.

The function of the iron yoke is to shape and guide the magnetic flux generated by the two warm coils. The weight of the yoke is 1500 tons. It is comprised of two identical horizontal parts which are arranged orthogonal to the plane of the coils, and two identical vertical parts to close the flux return. The yoke is installed on top of a 100 mm thick plate made of laminated low carbon steel.

### 2.3.3 Tracking System

The LHCb tracking system includes the VERtex LOcator (VELO), the TT silicon planes and T1-T3 which consist of Inner Tracker (IT, silicon microstrip detectors which are used in the region close to the beam pipe) and Outer Tracker (OT, straw-tubes which are utilized in the outer region of the stations). The TT and the IT were developed in a common project called the Silicon Tracker (ST). The tracks of the passing particles are reconstructed from the positions of the particles measured by each detector of the Tracking stations (T-stations, refer to the ST and the OT together).

#### 2.3.3.1 VELO

The VELO is a silicon strip vertex detector. As its name implies, VELO is surrounding the interaction region as illustrated in figure 2.8 and able to precisely measure the positions of vertices which are the places where particles decay into other particles. Due to the close position and extensive beams, the VELO is operated in an extreme radiation environment. For this reason, the detector is required to have high tolerance to radiation<sup>5</sup>.

---

<sup>5</sup>The worst irradiation damage to the silicon is equivalent to that of 1 MeV neutrons with a flux of  $1.3 \times 10^{14} n_{eq}/cm^2$  per year. The detector is required to sustain three years of nominal LHCb operation.

The VELO consists of twenty-one stations arranged along the beam direction, where each station has the ability to measure the  $R$  and  $\phi$  coordinates by  $R$ -measuring sensor and  $\phi$ -measuring sensor respectively. Each station has two halves which can be moved apart during the period of beam injection and closed once the beams are stable (see figure 2.8). To minimize the material between the interaction region and the detectors, the silicon sensors are installed in an independent aluminum vessel which is able to preserve vacuum around the sensors [26] (see figure 2.9). In addition to the twenty-one VELO stations, there are four  $R$  sensors perpendicular to the beam direction and known as pile-up veto stations which are used in the Level-0 Trigger (L0) and are described in section 2.3.5.

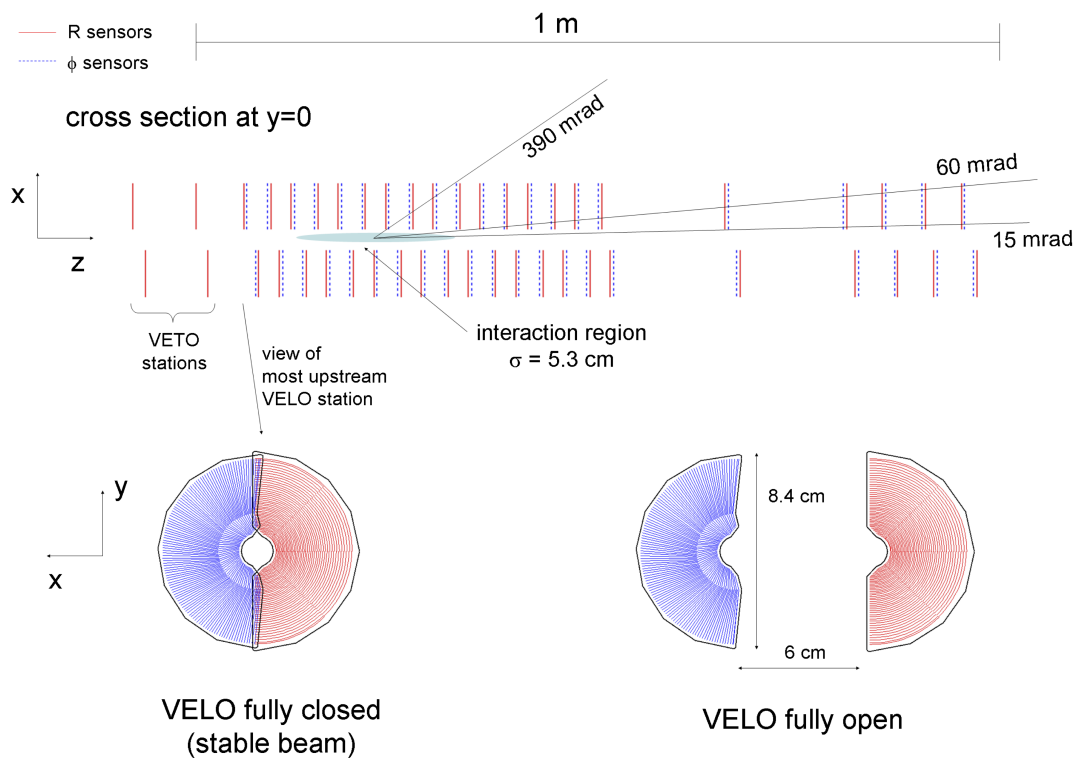


Figure 2.8: An overview of the VELO. Top is the schematic view of the twenty-one stations of the VELO along the beam direction; bottom are the closed and open positions of the VELO sensors.

To cover the full angular acceptance of the downstream detectors, the VELO is designed to be able to detect particles in the pseudorapidity range  $1.6 < \eta < 4.9$  and emerge from primary vertices (PV) within  $10.6$  cm. To reconstruct tracks, at least three hits are required

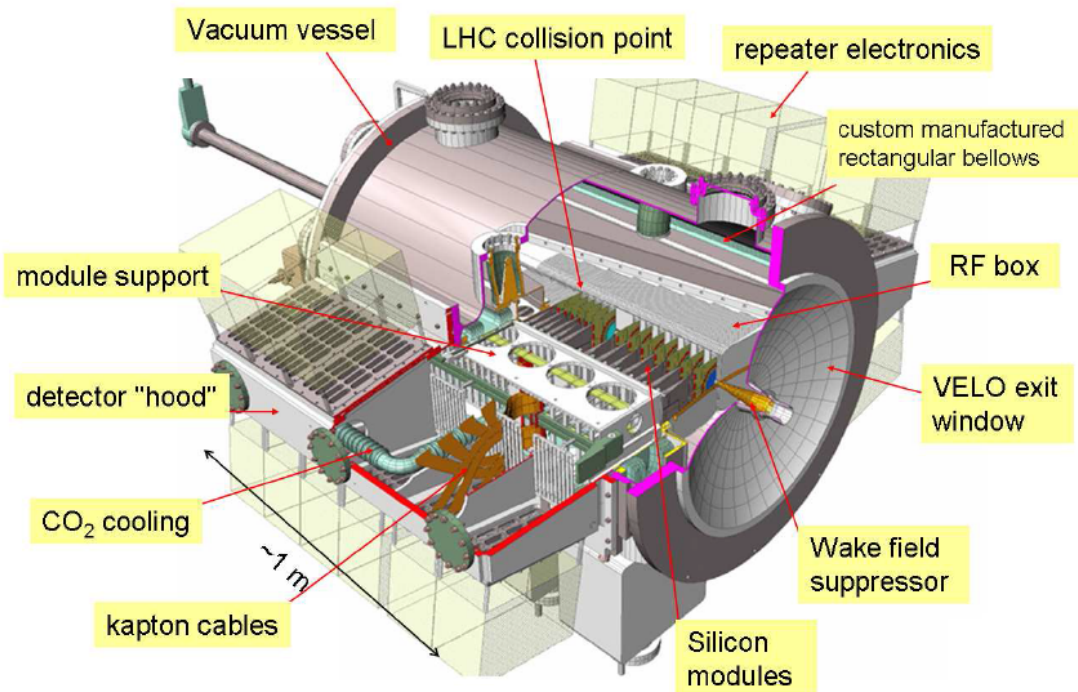


Figure 2.9: An overview of the VELO vacuum vessel.

to be measured per track. Other performance requirements for vertices reconstruction are:

- The signal to noise ratio (S/N) is greater than 14 to ensure efficient trigger performance.
- The overall channel efficiency is better than 99% with a signal to noise cut:  $S/N > 5$ .  
Figure 2.10 shows the VELO tracking efficiency [28]
- The spatial resolution is  $\sim 4 \mu m$  for 100 mrad tracks in the smallest strip pitch<sup>6</sup> region.  
Figure 2.11 shows the VELO resolution as a function of the strip pitch (left) or the track multiplicity (right) [28].

To tolerate significant radiation doses,  $n^+n$  (n-strip detectors on m-bulk material) silicon strip sensors with AC coupling and polysilicon biasing are chosen and implemented in the VELO. Each sensor is a semicircle patterned with 2048 azimuthal ( $R$  measuring) or quasiradial ( $\phi$

<sup>6</sup>The pitch is defined as the space among the strips of the sensor. It varies linearly from the center towards the edge. For  $R$ -measuring sensor, the pitch range is  $40 \mu m \sim 102 \mu m$ ; for  $\phi$ -measuring sensor, the pitch range is  $38 \mu m \sim 96 \mu m$ .



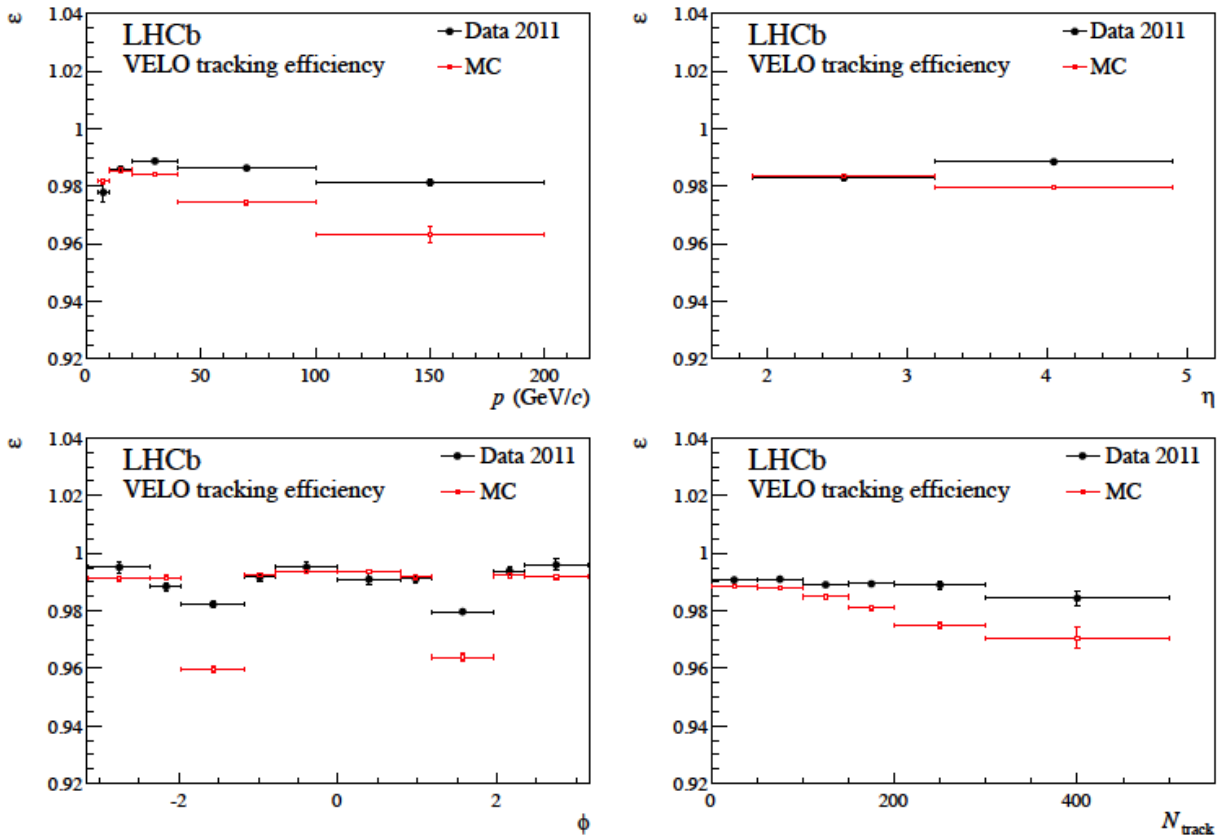


Figure 2.10: The VELO tracking efficiency for the 2011 data (black) and reweighed simulation (red) as a function of the momentum (top left), the pseudorapidity (top right), the azimuthal angle  $\phi$  (bottom left) and the total number of the tracks in the event (bottom right).

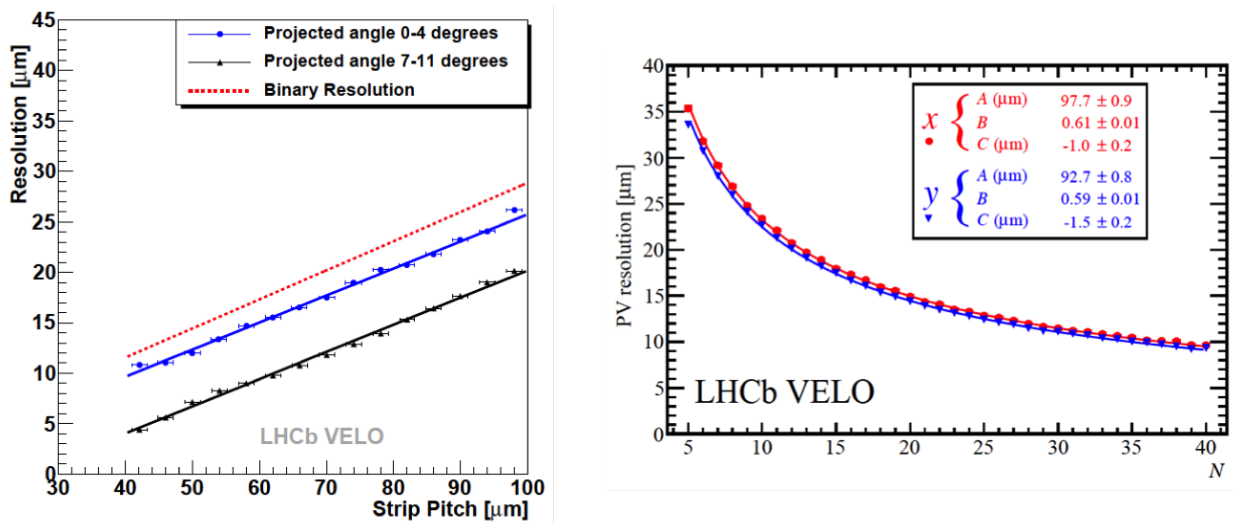


Figure 2.11: Left plot shows the VELO resolution for two different project angles; right plot shows the VELO PV resolution for  $x$  and  $y$  directions.

measuring) silicon strips. The strips on the  $R$ -measuring sensor are divided into four  $45^\circ$  sectors at all radii, where this design has advantages for the uniformity of the capacitance per channel, the execution time of the track reconstruction, and the number of clone and ghost tracks [29]. The strips on the  $\phi$ -measuring sensor which is built to readout the orthogonal coordinate to the  $R$ -measuring sensor are split into two regions with equivalent occupancy: 683 inner strips and 1365 outer strips [29]. Both  $R$ -measuring and  $\phi$ -measuring sensors are designed to be that the minimum pitch is at the innermost radius in order to optimize the vertex resolution [26]. The pitch on the sensors is a function of the radius  $r$  ( $\mu\text{m}$ ) [26]:

$$R - \text{measuring sensor} : 40 + (101.6 - 40) \times \frac{r - 8190}{41949 - 8190}$$

$$\phi - \text{measuring sensor} : \begin{cases} 37.7 + (79.5 - 37.7) \times \frac{r - 8170}{17250 - 8170} & \text{if } r < 17250 \\ 39.8 + (96.9 - 39.8) \times \frac{r - 17250}{42000 - 17250} & \text{if } r > 17250 \end{cases}$$

The design of the two sets of sensors is illustrated in figure 2.12.

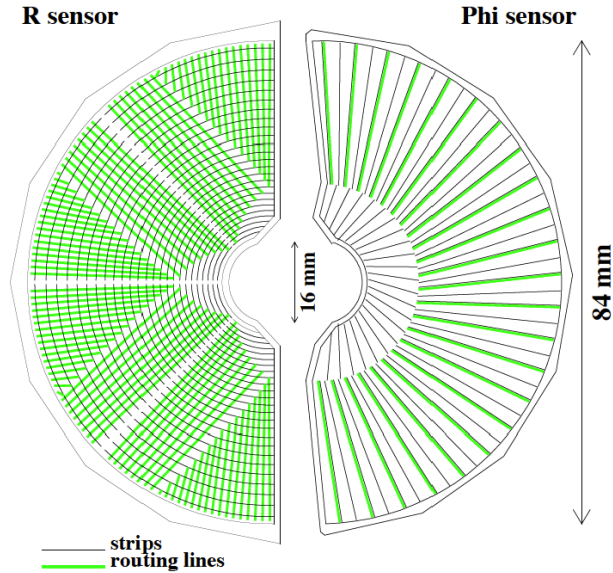


Figure 2.12: Schematic illustration of  $R$ -measuring and  $\phi$ -measuring sensors. The routing lines (green) orientate perpendicular and parallel to the silicon strips (black), respectively.

Both  $R$ -measuring and  $\phi$ -measuring sensors are  $300 \mu\text{m}$  thick. They are read out through the

routing line which is deposited as a second metal layer isolated from the AC-coupled diode strips by approximately  $3 \mu\text{m}$  of chemically vapor deposited (CVD)  $\text{SiO}_2$ . Each sensor is attached to the VELO hybrid which holds the sensor rigidly and perpendicular to the beam direction. The VELO hybrid has a size of approximately  $120 \times 170 \times 1 \text{ mm}^2$ . As an important part of the VELO module, the VELO hybrid provides electronic and mechanical support for the sensor. The VELO hybrid consists of a layer of Kapton on top of a  $400 \mu\text{m}$  thick thermal pyrolytic graphite (TPG) with  $250 \mu\text{m}$  of carbon fiber (CF) encapsulated on both side (see the left plot of figure 2.13). The front-end electronics (FEE) are glued on the thin Kapton sheet. The other parts of the VELO module are the cooling block which provides the thermal linkage with the cooling system, the paddle and the base. The design of the VELO module is illustrated by the right plot of figure 2.13.

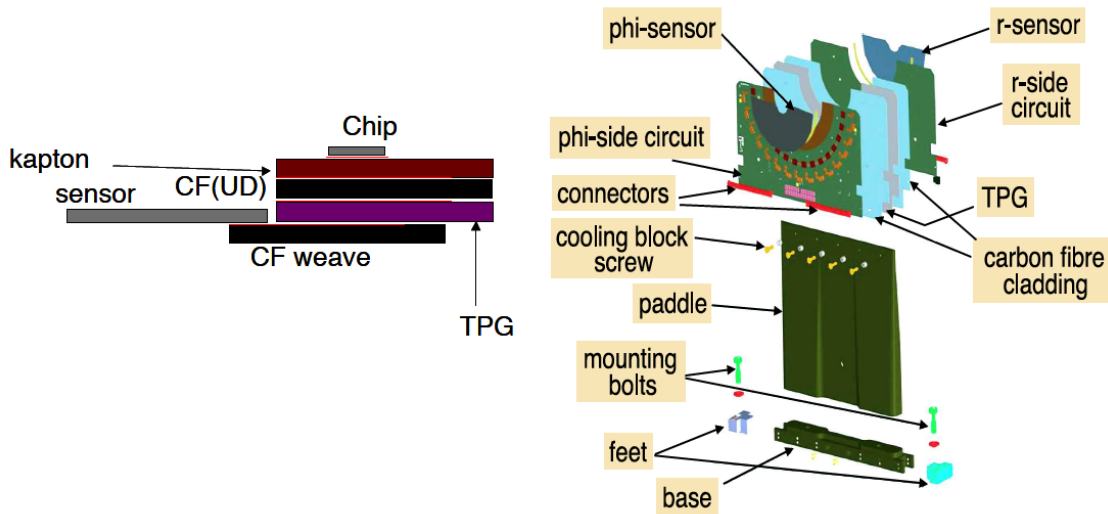


Figure 2.13: Schematics of the VELO hybrid design (left) and the VELO module with the key components (right).

The front-end electronics of the sensors are known as beetle chips. The chip is designed with a sampling frequency of 40 MHz bunch crossing rate to meet the LHCb requirements. Each silicon sensor connects to 16 beetle chips, thus each chip reads out 128 channels. The chip can be used as an analog or a binary pipeline. The readout signal is analog and transferred to the repeater board (RPT) which is located directly outside of the vacuum tank through

the Kapton cable which is chosen for its good flexibility and radiation tolerance. The readout chain is shown in figure 2.14. The high voltage and low voltage power supplies are installed behind the shilling wall which provides a radiation safe environment.

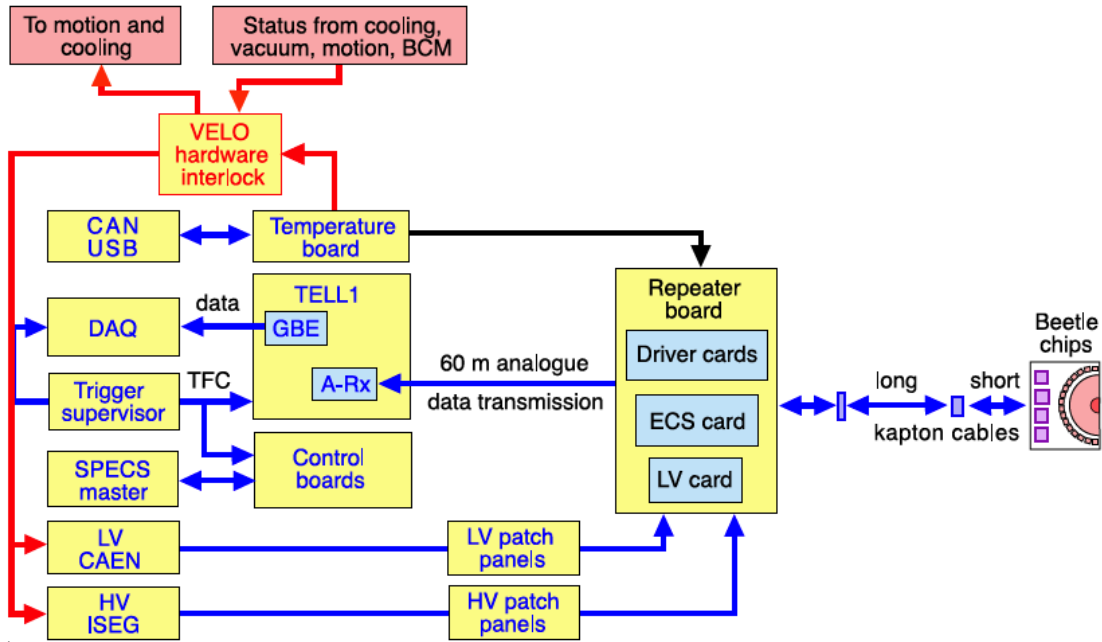


Figure 2.14: A schematic of the VELO front-end electronics.

### 2.3.3.2 Silicon Tracker

The Silicon Tracker (ST) comprises two detectors: the TT silicon planes and the Inner Tracker (IT). Both TT and IT use the silicon microstrip sensors with  $\sim 200 \mu\text{m}$  strip pitch of up to 38 cm in length. The TT is located in between RICH1 and the dipole magnet. It is approximately 150 cm wide and 130 cm high. The IT is located downstream of the magnet. It covers 120 cm in width and 40 cm in height. Each ST station has four detection layers in an  $(x - u - v - x)$  arrangement. Both the TT and the IT have a single hit resolution of about  $50 \mu\text{m}$ .

The TT sensors are installed in a thermally and electrically insulated detector volume with the temperature maintained below  $5^\circ\text{C}$ . The TT has approximately  $8.4 \text{ m}^2$  active area with 143360 readout strips. It is able to cover the full angular acceptance of the LHCb. The TT

is used in the offline analysis to reconstruct the tracks of long-lived neutral particles which decay outside of the VELO and of low-momentum particles which are bent out of the LHCb acceptance before reaching T1-T3 [29]. It also reduces false fraction among high momentum tracks.

The first and last layers of the TT have vertical readout strips measuring the  $x$ -direction. The second and third layers have strips rotated by a stereo angle of  $+5^\circ$  ( $u$ -axis) and  $-5^\circ$  ( $v$ -axis) with respect to the  $x$ -axis, respectively [26]. The four detection layers are arranged in two groups:  $(x, u)$  and  $(v, x)$ . The two groups are separated by approximately 27 cm along the beam direction. The four detection layers of the TT are illustrated in the figure 2.15.

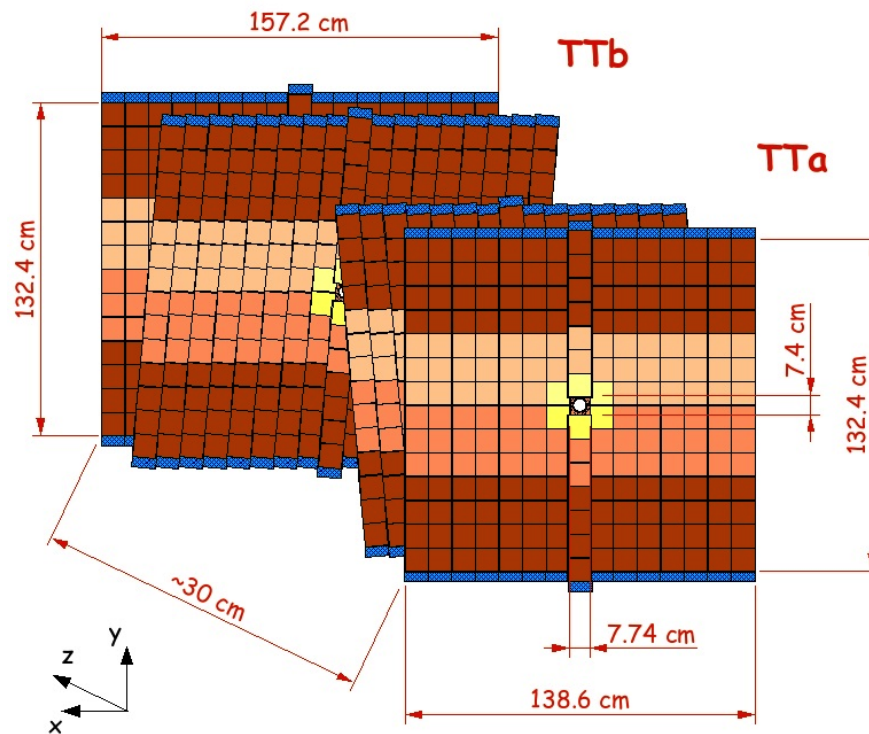


Figure 2.15: Layout of the four detection layers of the TT.

Each detection layer of the TT is built up of several half modules which cover half the height of the LHCb acceptance. The half module comprises a row of seven silicon sensors with the readout hybrid mounted at the end far from the beam (see figure 2.16). The seven silicon sensors are categorized into two (4-3) or three (4-2-1) readout sectors. Both 4-3 type and 4-2-1

type half modules have L sector formed by the four sensors closest to the readout. These four sensors are bounded together and directly connected to the lower-most readout hybrid. For 4-3 type half module, the remaining sector is M sector which is composed of three sensors. These three sensors are connected to a second readout hybrid mounted on top of the L hybrid via a Kapton flex cable with a length of 39 cm. For 4-2-1 type half module, the M sector consists two sensors next to L sector. The Last single sensor closest to the beam forms the K sector. It is connected to a third readout hybrid via a 58 cm long Kapton flex cable. Each pair of the two half modules are joined together end-to-end to cover the full height of the LHCb acceptance.

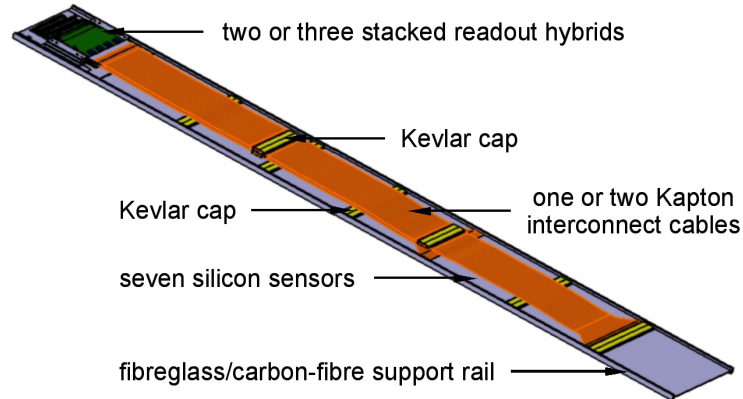


Figure 2.16: Schematic of a TT half module.

The IT consists of three stations and has approximately  $4.0 \text{ m}^2$  active area with 129024 readout strips. Each station comprises four individual detector boxes which are arranged around the beam pipe. Each detector box has four detection layers and each detection layer has seven one-sensor or two-sensor detector modules (see figure 2.17). Detector modules in the boxes above and below the beam pipe comprise a single silicon sensor and a readout hybrid, while the modules in the left and right boxes consist of two silicon sensors and a readout hybrid. Due to close position to the beam pipe, the expected radiation dose for the IT is similar to that expected for the VELO. Thus the material used in the IT has to tolerate high radiation damage. Both the one-sensor and two-sensor modules use  $11 \text{ cm} \times 7.6 \text{ cm}$  single-sided  $p^+n$  silicon sensors carry 384 readout strips with a strip pitch of  $198 \text{ }\mu\text{m}$ . The

only difference between the two types of modules is thickness, where the one-sensor is  $320\ \mu\text{m}$  thick and the two-sensor module is  $400\ \mu\text{m}$  thick.

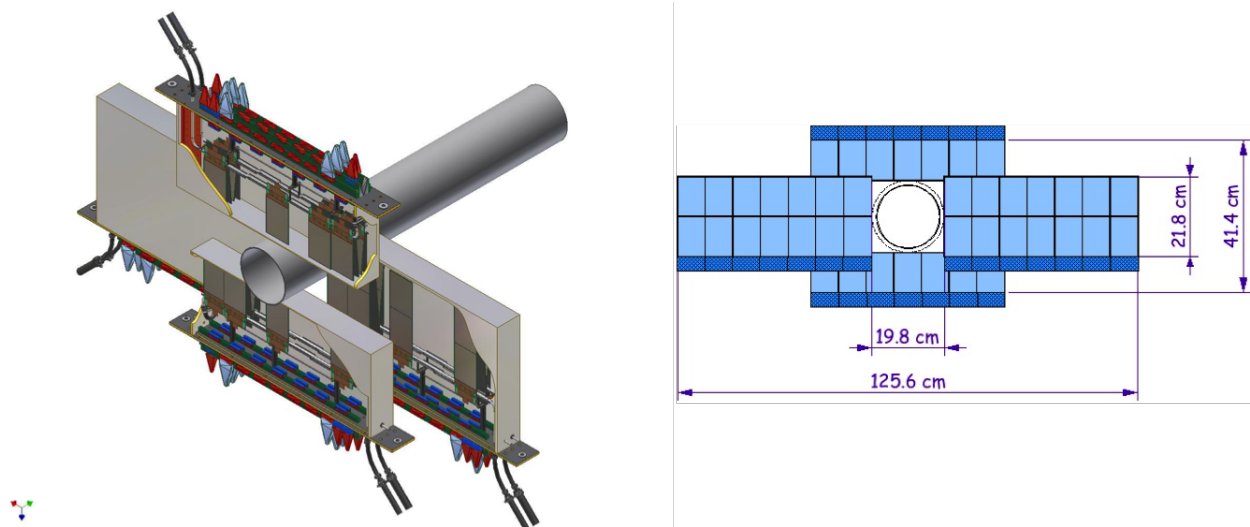


Figure 2.17: Left: view of the IT boxes arranged around the beam pipe. Right: layout of the  $x$ -layer in the second IT station.

The IT uses the same technology as the TT: the first and the last layers of each tracking station have vertical readout strips measuring the  $x$ -direction, while the second and third layers of the each tracking station have strips rotated by a stereo angle of  $+5^\circ$  ( $u$ -axis) and  $-5^\circ$  ( $v$ -axis), respectively. To keep leakage currents to a level where shot noise does not significantly deteriorate the S/N performance of the detector during several years of operation, the the silicon sensors of the IT are kept at a temperature below  $5^\circ$  [30].

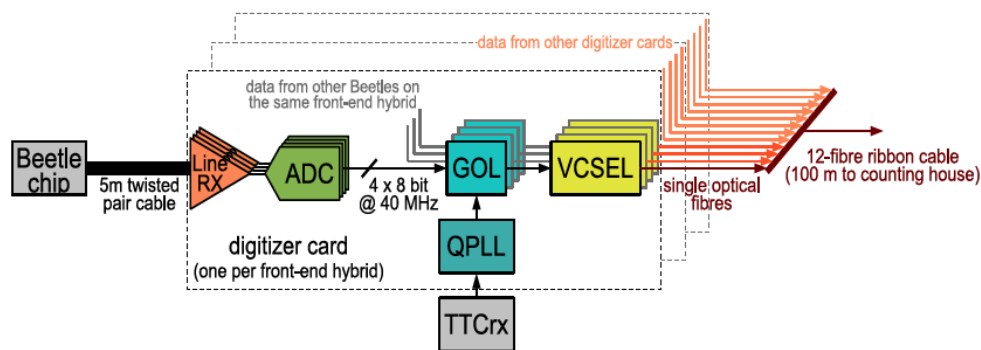


Figure 2.18: Data process from one Beetle chip of the ST.

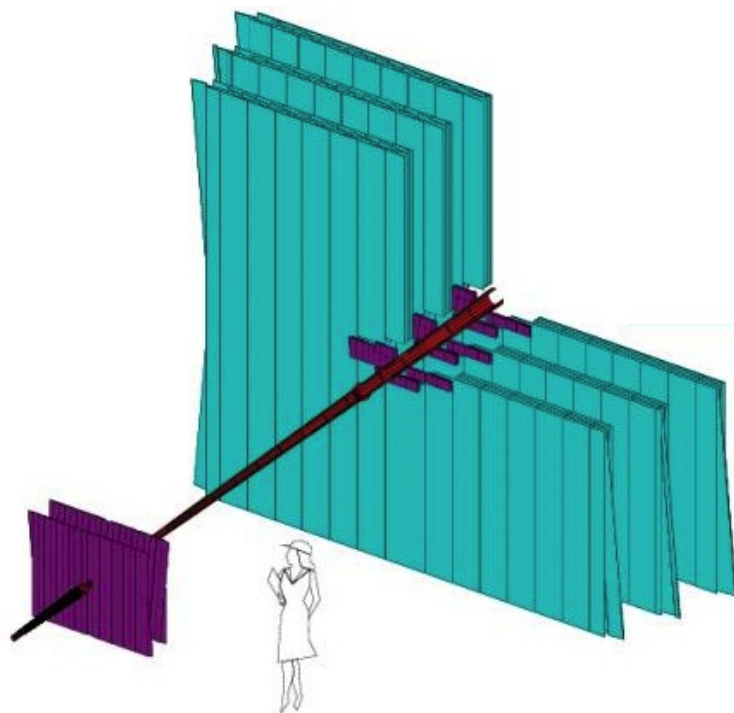
Both the TT and the IT use the Beetle FEE chip which is located on the front-end readout



hybrid. Each Beetle chip is connected to 128 readout strips. The Beetle chips sample the detector signals at the frequency of 40 MHz and store the sampled data in an analog pipeline. Then the signals from one front-end readout hybrid are transmitted from the detector boxes to the service boxes via a shielded 68-wire twisted-pair cable. The signal process from one Beetle chip of the ST is illustrated in figure 2.18.

### 2.3.3.3 Outer Tracker

The Outer Tracker (OT) is a drift-time detector comprised of three tracking stations. The OT stations are mounted around the IT stations (see figure 2.19). It covers approximately  $30 \text{ m}^2$  active area which allows the OT to track the charged particles and measure their momentum over a large acceptance. Each of the OT stations has four detection layers with the same arrangement ( $x-u-v-x$ ) as used in the ST.



*Figure 2.19: Arrangement of the OT straw-tube modules in layers and stations.*

Excellent momentum resolution is essential for a precise determination of the invariant mass



of the reconstructed b hadrons. Thus the OT is required to have a sufficient resolution of  $200 \mu\text{m}$  in the x-coordinate. Rather than the silicon sensors, the OT makes use of the gas-tight straw-tube modules [31]. Each module comprises two staggered layers (monolayers) of drift-tubes. The tubes are 12 m long with 4.9 mm inner diameters. In order to guarantee a fast drift time below 50 ns, the tubes are filled with a gas mixture of Argon (70%), CO<sub>2</sub> (28.5%) and Oxygen (1.5%) [32]. The wire along the center of the tube is the anode wire which is made of gold plated tungsten of  $25 \mu\text{m}$  diameter and set to +1550 V. The outside cylinder is the cathode which is kept at an equally negative voltage, or ground. The cathode is made of  $40 \mu\text{m}$  thick carbon-doped polyamide foil wound simultaneously with a  $20 \mu\text{m}$  thick Kapton laminated with aluminum of  $12.5 \mu\text{m}$  thickness. The tubes are glued to the panels which are sealed with  $400 \mu\text{m}$  thick carbon fiber sidewalls. The cross section of a straw-tube module is illustrated in figure 2.20.

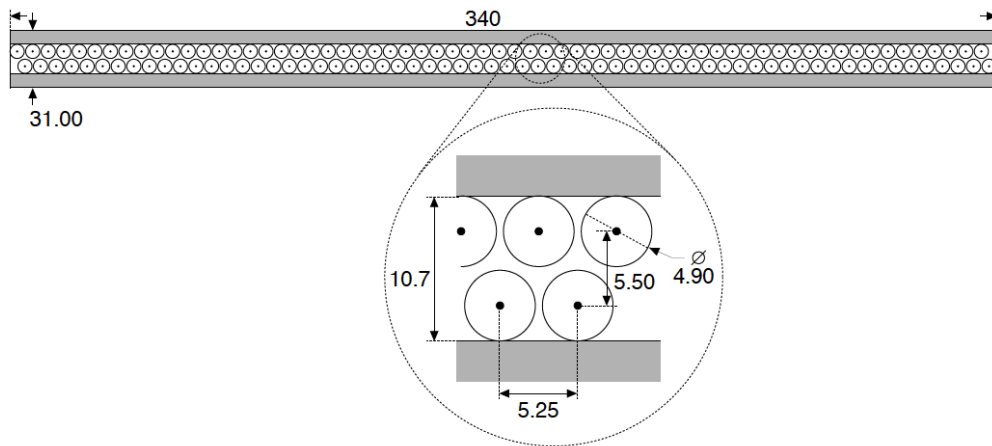


Figure 2.20: Cross section of a straw-tube module of the OT.

The FEE measures the drift times of the ionization clusters induced by charged particles traversing the straw-tubes with respect to the beam crossing signal. The drift times are digitized for every 25 ns and stored in a digital pipeline to await the Level-0 decision. On a positive L0 decision, the digitized data in a 75 ns window is transmitted via optical links to the TELL1 boards [32].

### 2.3.3.4 Summary of Tracking Detectors

The LHCb tracking system is used to efficiently reconstruct the trajectories of charged particles and measure their momentum. The track reconstruction is done with the entire tracking system. It starts with a search for track seeds in the VELO region and the T-stations where the magnetic field is low. Then the “observed” tracks are refitted with a Kalman filter which takes into account multiple scattering with detector material and corrects for  $dE/dx$  energy loss [26]. Figure 2.21 illustrates the different type of tracks.

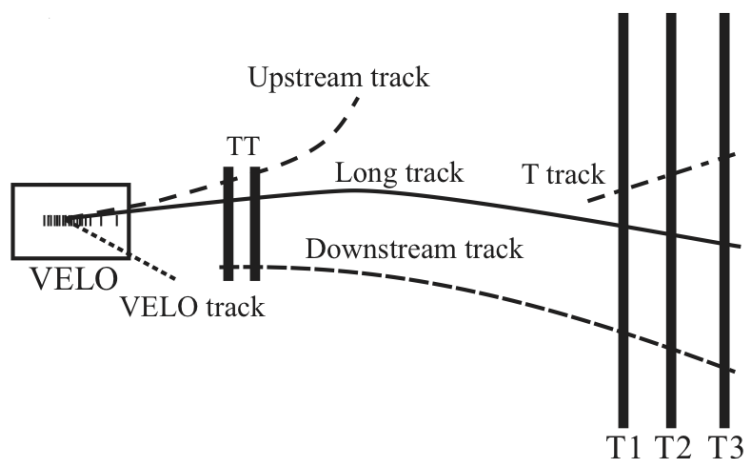


Figure 2.21: Schematic of various types of tracks through the LHCb detector: long, upstream, downstream, VELO and T tracks.

### 2.3.4 Particle Identification

Particle identification (PID) is a fundamental requirement for the LHCb. The LHCb PID system consists of two Ring Imaging Cherenkov detectors (RICH1 and RICH2), Electromagnetic Calorimeter (ECAL), Hadronic Calorimeter (HCAL) and muon system. The two RICH detectors are used to identify particles with different range of momentum. The calorimeters provide the measurements of the energy of electrons, photons and hadrons. The muon system is used to identify and trigger on muons.

### 2.3.4.1 Ring Imaging CHerenkov Detector

The Ring Imaging CHerenkov detectors (RICH) are built for identification of charged particles. Both the RICH detectors are aligned vertically to the beam pipe. The first (RICH1) is located between the VELO and the magnet and occupies the region  $990 < z < 2165$  mm. The second (RICH2) is placed downstream of the T-stations with its front face positioned at 9500 mm from the interaction point and with a depth of 2332 mm. The alignment of the two RICH detectors is illustrated in figure 2.22. The base components of the RICH detector are mirrors, photon detectors and radiators which are the mediums filling up the cavity of the RICH detector.

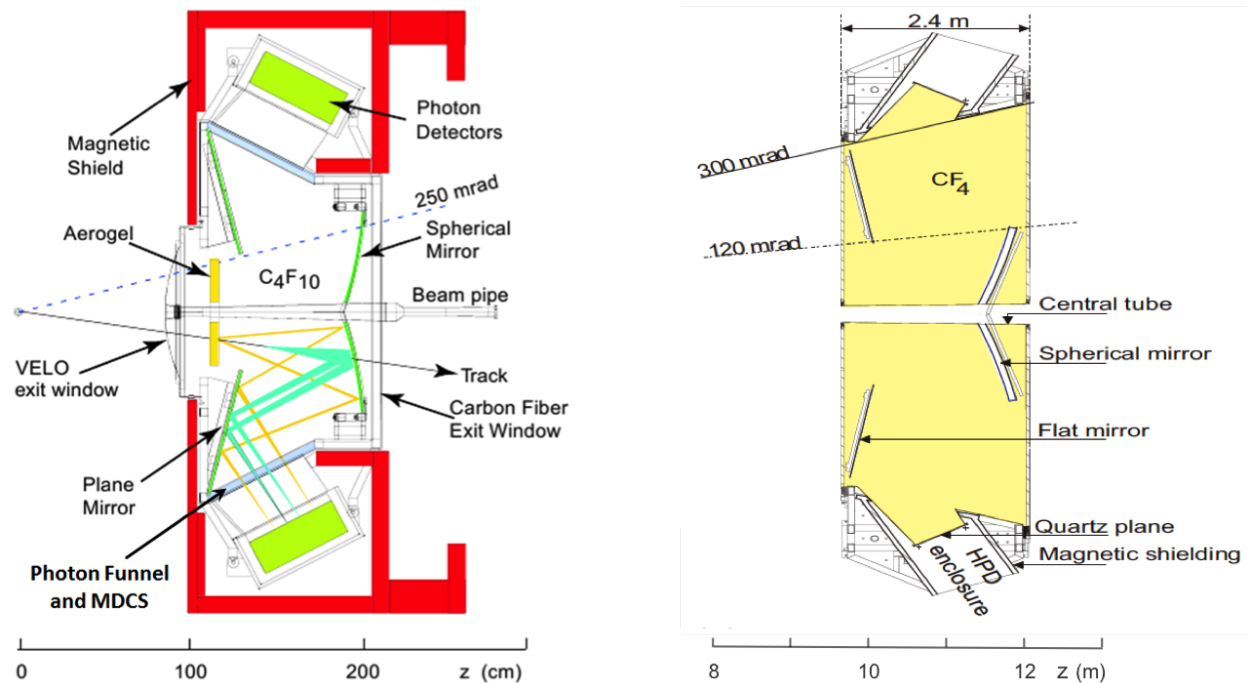


Figure 2.22: Schematics of the RICH1 and RICH2.

When a charged particle travels through a medium at a speed higher than the phase velocity of light in that medium, a faint radiation is produced. This phenomenon is known as Cherenkov radiation which is utilized by the RICH detector to identify the charged particles. The identification is achieved by measuring the angle of emission of the Cherenkov radiation, which is related to the speed of the charged particle. The relation between the speed of the

charged particle ( $v$ ) and the emission angle ( $\theta_c$ ) is given in:

$$\cos\theta_c = \frac{c}{nv} \quad (2.1)$$

where  $n$  is the refractive index of the medium and  $c$  is the speed of light. Figure 2.23 shows the relation between emission angle and particle momentum in different mediums.

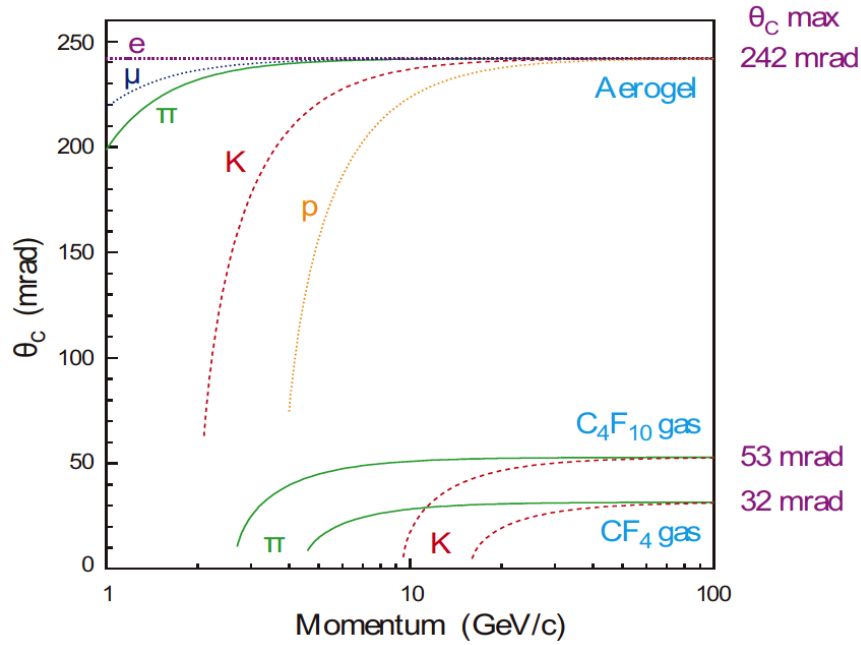


Figure 2.23: Cherenkov angles as a function of momentum for the RICH radiators.

The upstream RICH detector (RICH1) covers the full acceptance of the experiment and identifies the charged particle with low momentum of 1 GeV to 60 GeV, while the downstream RICH2 detector has a limited angular acceptance of  $\sim \pm 15$  mrad to  $\pm 120$  mrad in the horizontal and  $\pm 100$  mrad in the vertical plane, and detects the charged particles with high momentum of 15 GeV to 100 GeV. For the RICH1 detector, silica aerogel and perfluorobutane ( $C_4F_{10}$ ) gas are chosen as the radiators in which the Cherenkov photons are generated uniformly along the length of each track. The RICH2 detector is filled with tetrafluoromethane ( $CF_4$ ) gas ( $n = 1.0005$ ) which has the smallest refractive index among the three radiators used in the RICH detectors and works well for tracks with high momentum. Silica aerogel ( $n = 1.03$ )

is suitable for the tracks with low momentum, while the intermediate region is well matched to  $C_4F_{10}$  gas ( $n = 1.0014$ ) [33]. The silica aerogel is designed to be a thin layer with 5 cm thickness and placed in the forward region of the RICH1 detector. The fluorobutane gas is chosen for its low dispersion, and is sealed by the Gas enclosure which also provides a light-tight and mechanically stable platform for all optical components. The Gas enclosure is a six-sided aluminum alloy tooling plate box which can sustain the pressure differential between the fluorobutane gas and the outside atmospheric environment.

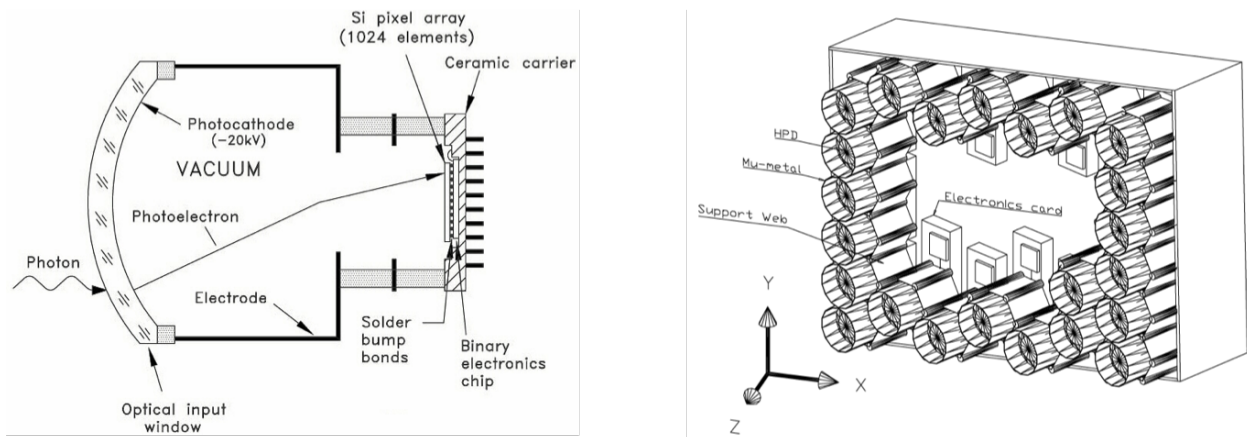


Figure 2.24: Left: schematics of the pixel Hybrid Photon Detector (HPD). Right: HPD array of the RICH1 detector.

The photons are emitted in the shape of a cone along each track in the radiators, while the focusing of the light is accomplished using spherical mirrors. They are tilted and located within the spectrometer acceptance. The spherical mirrors and flat mirrors work together and bring the image to the photon detectors which are mounted out of the acceptance to avoid degrading the tracking. This structure is also designed to shorten the overall length of the detector. The spherical mirrors for RICH1 are constructed from a Carbon Fiber Reinforced Polymer (CFRP) to reduce the amount of scattering, while the spherical mirrors for RICH2 are made of glass. The CFRP mirrors are coated with  $Al$  (80 nm) and  $MgF_2$  (160 nm). Flat mirrors use glass substrates since they are located outside of the acceptance. The flat mirrors are assembled into two planes in the RICH detector, one above the beam pipe, the other below. Each plane consists of eight rectangular mirrors.  $Al+SiO_2+HfO_2$  coating is deposited

on the flat mirrors.

The photons are detected and measured by pixel Hybrid Photon Detectors (HPD) which use silicon detector anode inside the vacuum tube. Each tube comprises 1024 pixels arranged as a  $32 \times 32$  matrix, while each pixel is  $500 \times 500 \mu\text{m}^2$  in size. The schematic of the HPD is illustrated in figure 2.24. A photoelectron, released from the conversion in a photocathode of an incident photon, is accelerated by a high voltage of typically 10 to 20 kV towards the silicon sensor. This kinetic energy is then dissipated near the silicon surface, which results in the creation of  $\sim 5000$  electron-hole pairs at an average yield of one for every 3.6 eV of deposited energy in the silicon sensor. There are total of 484 tubes  $\{196$  for RICH-1 and  $288$  for RICH-2  $\}$  to cover the four detection surfaces. They are arranged in arrays and positioned on a hexagonal lattice (see figure 2.24). With the suitably focusing optical system, the ring images are reconstructed on the HPDs (see figure 2.25). The radius of the ring is independent of the emission point along the particle track and is an approach to measure the Cherenkov emission angle. Then the velocity of the particle can be obtained from equation 2.1.

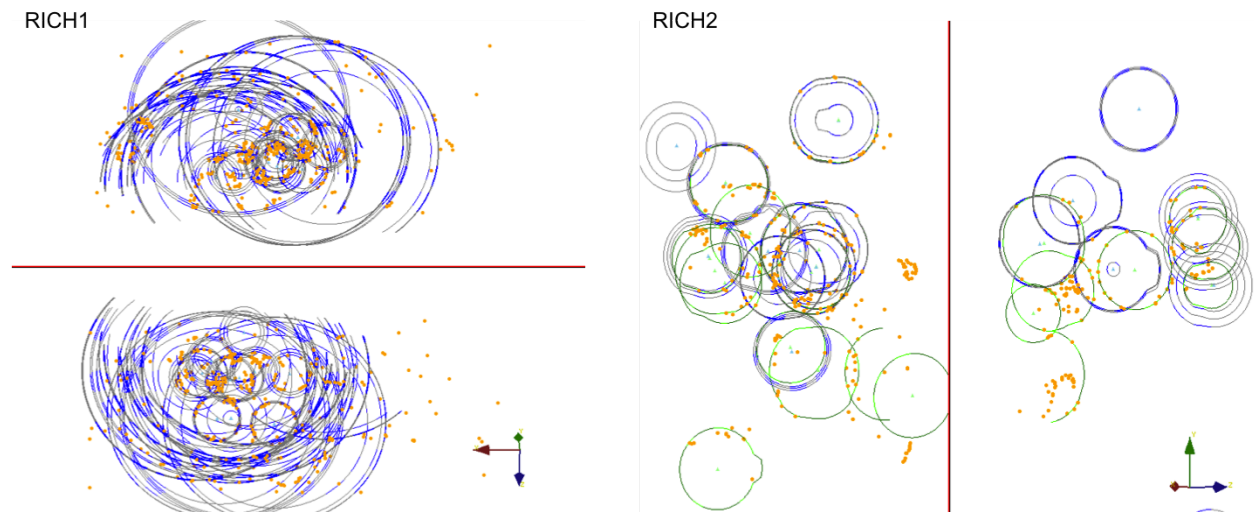


Figure 2.25: Event display of detected photoelectrons in RICH1 (left) and RICH2 (right). The circles are the fit results.

Both detectors must be operated at a low magnetic field environment to preserve a good

resolution. Since in the strong magnetic field, the tracks curve appreciably while passing through the medium and the images recorded by HPDs become distorted. To reduce the magnetic field, the RICH detectors are surrounded by iron shields and the HPDs are individually shielded by  $Ni-Fe$  cylinders. To solve the problem of distortion, two identical Magnetic Distortion Calibration Systems (MDCS) are independently applied to the upper and lower HPD enclosures of RICH1 [34]. Figure 2.26 shows the design of the MDCS system.

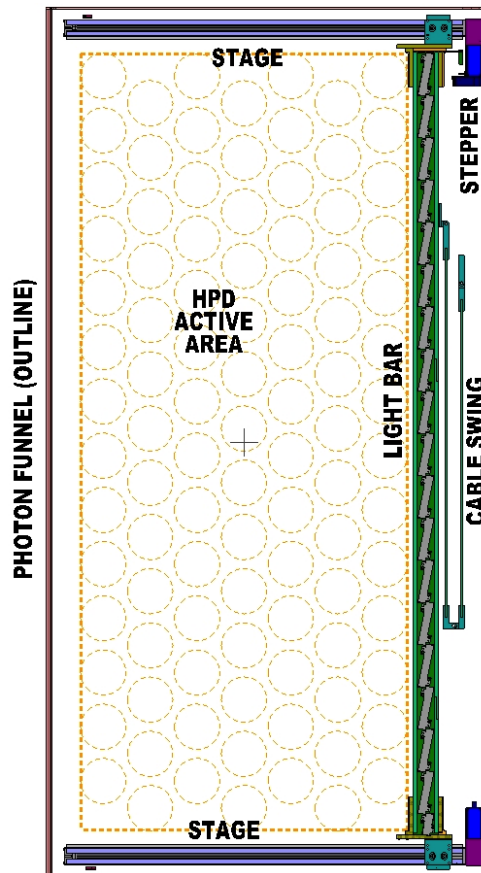


Figure 2.26: A drawing of the MDCS system of RICH1. Dotted circles represent the HPDs. The light bar is in parked position at the right edge of the box.

The readout system of the RICH consists of: binary FEE pixel chip, Level-0 adapter board and Level-1 readout board. The FEE pixel chip is encapsulated inside the HPD tube. It is designed with special techniques to enhance its resistance to radiation. The Level-0 adapter board is mounted on detector. It drives the optical data links to the off-detector electronics and distributes clocks and triggers to the FEE chips. The Level-1 readout board is located

approximately 100 m away from the detector.

By combining the information of velocity with the momentum measured from tracking system, the probability likelihood distribution is determined for each type of particle and compared to a probability likelihood of pions in the RICH. This difference in log-likelihood is defined as:

$$\text{DLL}_{X\pi} = \log\left(\frac{P_X}{P_\pi}\right) \quad (2.2)$$

where  $P$  is momentum and  $X$  represents the particle to be identified. A plot of difference in log-likelihood is shown in figure 2.27 for tracks that have been matched to true kaons and pions [35]. In the plot,  $\text{DLL}_{K\pi}$  the value for kaons tends to be positive, while that for pions tends to be negative. There is a double-peaked structure for kaons which is due to the momentum-dependence of the  $\pi$ -K separation of the RICH system.

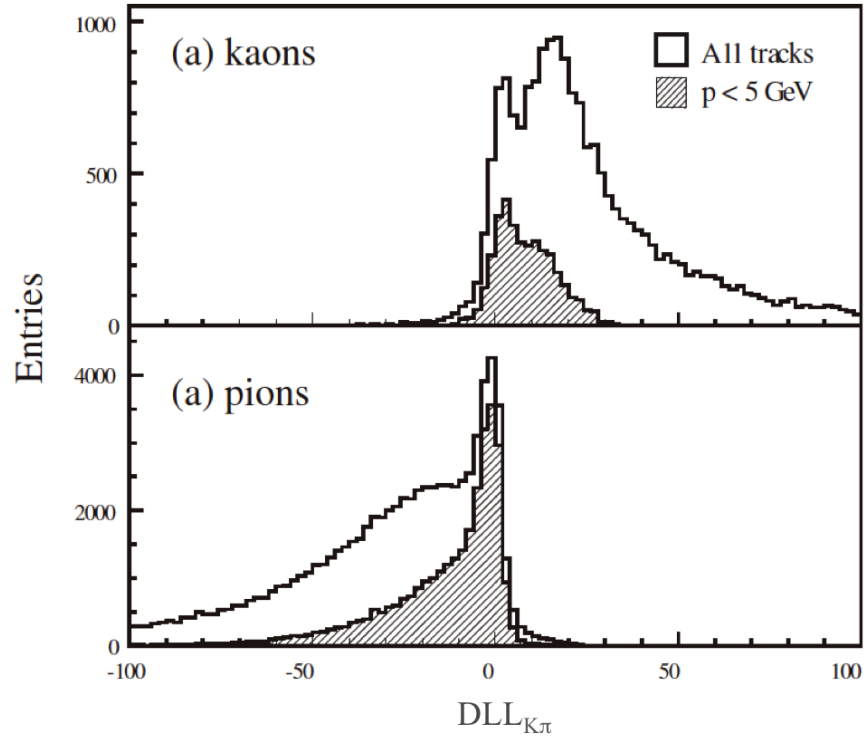


Figure 2.27: Distributions of the difference in log-likelihood between kaon and pion hypotheses for kaons (top) and pions (bottom). Two different requirements are implemented on the sample of  $B_s^0 \rightarrow D_s^- K^+$  events.



To test the PID performance of the RICH, a data-driven approach is implemented, which requires large statistics of pure samples. The RICH PID power depends strongly on the particle momentum, so each sample covers the full momentum range of 2 GeV to 100 GeV. Furthermore, the selection of such control samples has to be independent of PID information. The samples of  $K^\pm$ ,  $\pi^\pm$  events are reconstructed from:  $K_s^0 \rightarrow \pi^+\pi^-$ ,  $\Lambda \rightarrow p\pi^-$  and  $D^{*+} \rightarrow D(K^-\pi^+)\pi^+$ . The figure 2.28 demonstrates the kaon efficiency (kaons identified as kaons) and pion misidentification (pions misidentified as kaons) rate as a function of momentum. Two different PID requirements are applied in the test. When  $\Delta LL_{K\pi}$  is required to be positive, the efficiency of the kaon identification is  $\sim 95\%$ , while the rate of the pion misidentification is  $\sim 10\%$ . Alternatively, the tighter PID requirement can reduce the misidentification rate significantly.

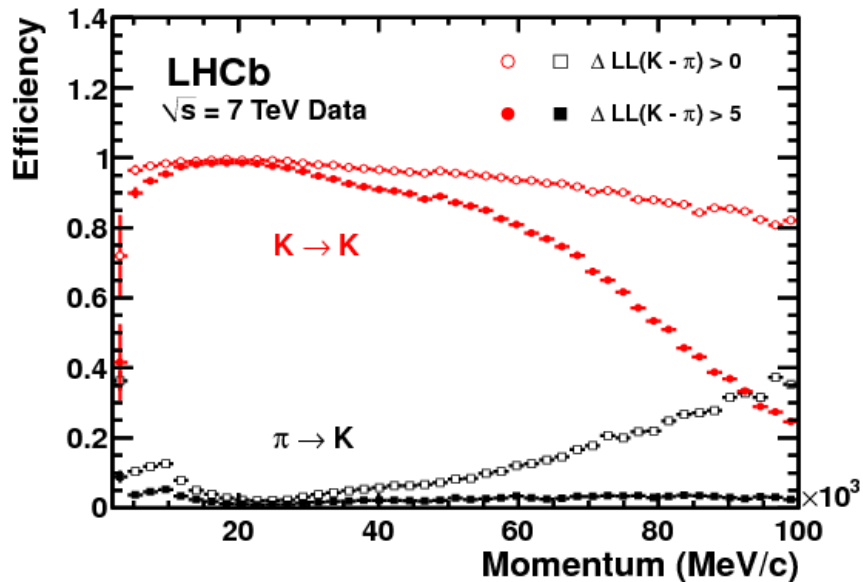


Figure 2.28: The efficiency of kaon identification and the rate of muon misidentification as a function of momentum.

#### 2.3.4.2 The Scintillator Pad Detector and the Preshower Detector

The Scintillator Pad Detector (SPD) and the Preshower Detector (PS) are situated between the first muon chamber and ECAL. The SPD/PS detectors use scintillator pad readout by

wavelength-shifting (WLS) fibers coupled to multi-anode photo-multiplier tubes (MAPMT) via clear plastic fibers, and cover  $7.6 \times 6.2 \text{ m}^2$  active area [26]. They are designed for better electron photon identification and hadron discrimination, and used at the trigger level in association with the calorimeters. The layout of the SPD/PS detectors is shown in figure 2.29.

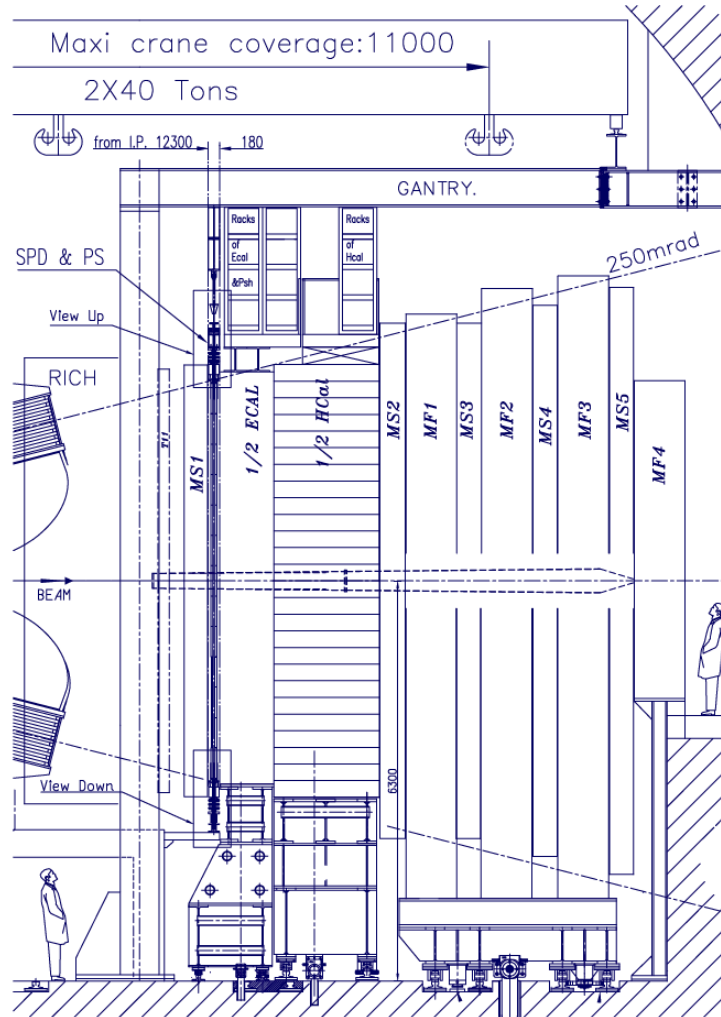


Figure 2.29: Side view of the SPD/PS detectors (between the MS1 and ECAL).

The SPD/PS detectors consist of two almost identical rectangular scintillator planes of high granularity with 12032 channels (cells) of scintillator pads, while a lead converter of 15 mm thickness is between the two planes. Each plane is made of two halves which can slide independently on horizontal rails. Furthermore, each plane is divided into three sections: inner (3072 cells), middle (3584 cells) and outer (5376 cells). This is done to achieve a one

to one projective correspondence with ECAL segmentation. The detection cells are packed in boxes (detector units) with the size of  $\sim 48 \times 48 \text{ cm}^2$ . Each 26 boxes are grouped into a supermodule. The PMT tubes which are connected to the detector units via optical fibers, are mounted on both the top and bottom ends of the supermodule outside the spectrometer acceptance.

The electron/pion separation is studied to test the performance of the SPD/PS. The electrons and pions between 10 GeV and 50 GeV momentum are injected to the SPD/PS detectors. The energy deposited in the PS for 50 GeV electrons and pions is shown in figure 2.30 [26].

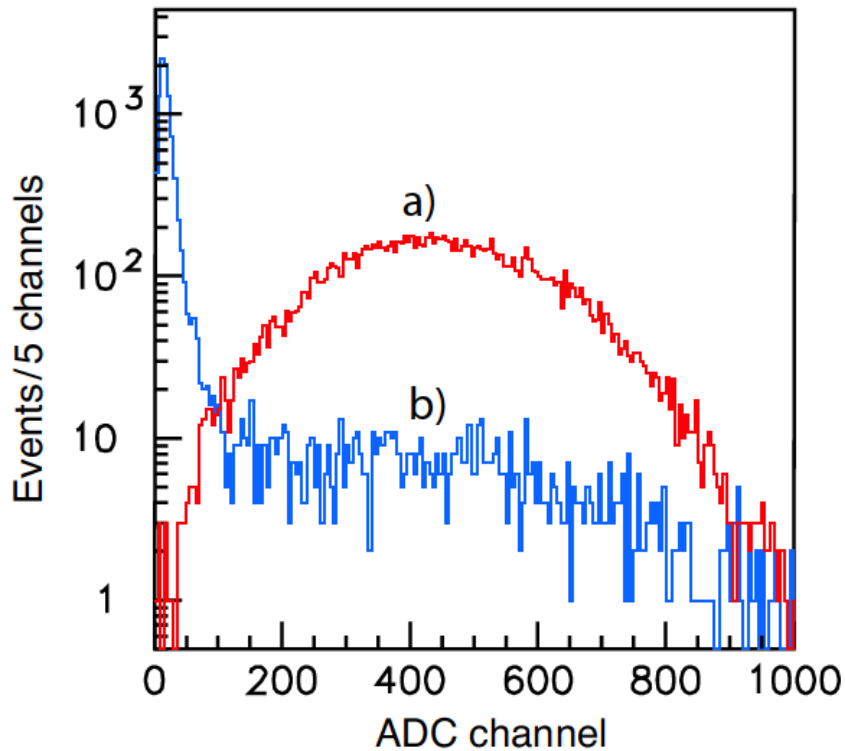


Figure 2.30: Energy deposition of 50 GeV electrons and pions in the PS.

To separate photons and pions, the information from the SPD is used. The measurements show that the probability of photon misidentification due to interactions in the SPD scintillator is  $(0.8 \pm 0.3)\%$ , when applying a threshold of 0.7 MIPs (short for Minimum Ionizing Particle). The probability to pass this threshold due to backward moving charged particles was measured to be  $(0.9 \pm 0.6)\%$  and  $(1.4 \pm 0.6)\%$  for 20 and 50 GeV photons, respectively.

### 2.3.4.3 Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) is located downstream of the SPD/PD planes, 12.5 m away from the interaction region. The overall dimensions of the ECAL are  $7.76 \times 6.30$  m, covering an angular acceptance of  $\pm 25$  mrad to  $\pm 300$  mrad horizontally and  $\pm 250$  mrad vertically. The ECAL is designed to identify photons and neutral pions for trigger and offline analysis as well as measure the information of energies and positions. A general view of the ECAL is shown in figure 2.31.

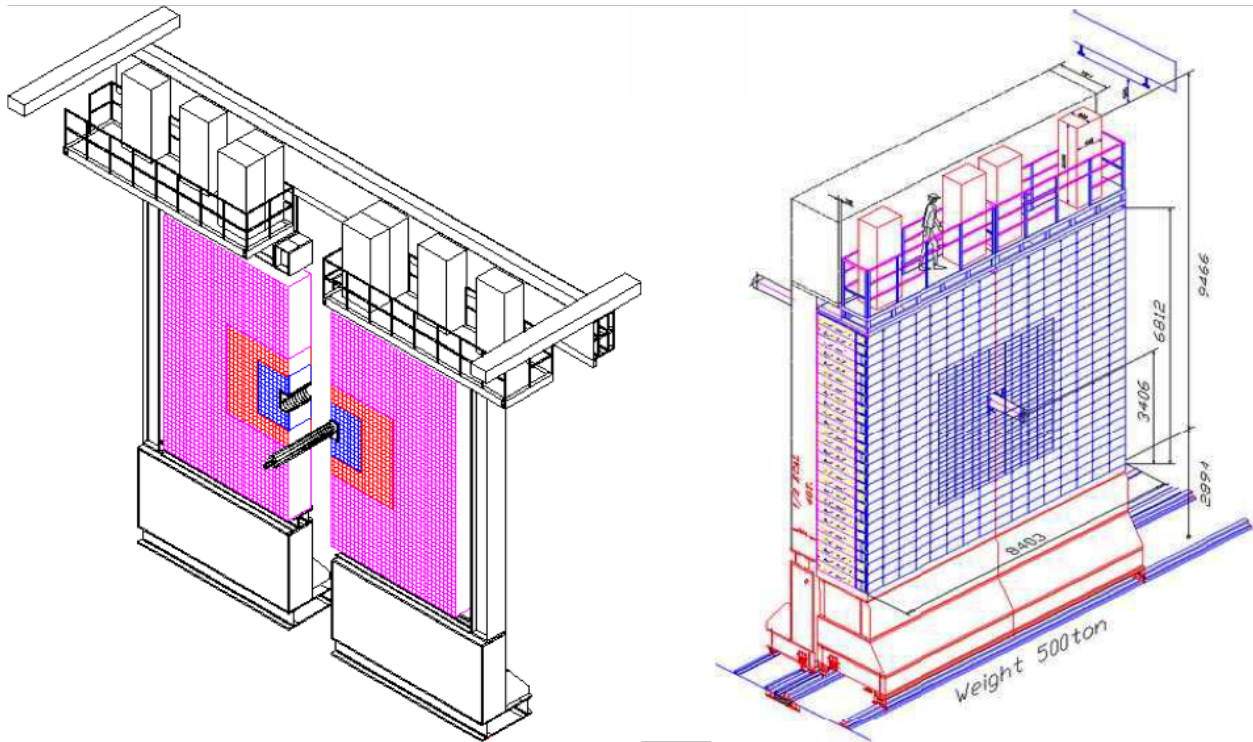


Figure 2.31: General view of the ECAL (left) and HCAL (right).

The ECAL employs the “shashlik” technology of alternating scintillating tiles and lead plates with the readout of plastic WLS fibers. The hit density on the ECAL surface is a steep function of the distance from the beam pipe, and varies by two orders of magnitude, so the ECAL is divided into three regions with three different cell sizes, respectively (see figure 2.32). The cell size is  $4 \times 4 \text{ cm}^2$  for the inner region,  $6 \times 6 \text{ cm}^2$  for the middle region and  $12 \times 12 \text{ cm}^2$  for the outer region [36]. The cell granularity corresponds to that of the SPD/PS, aiming at a

combined use in separation of photons and electrons. Each cell module consists of alternating layers of 2 mm thick lead, a foil of white reflecting paper and 4 mm thick scintillator tiles. All layers are perpendicular to the beam pipe.

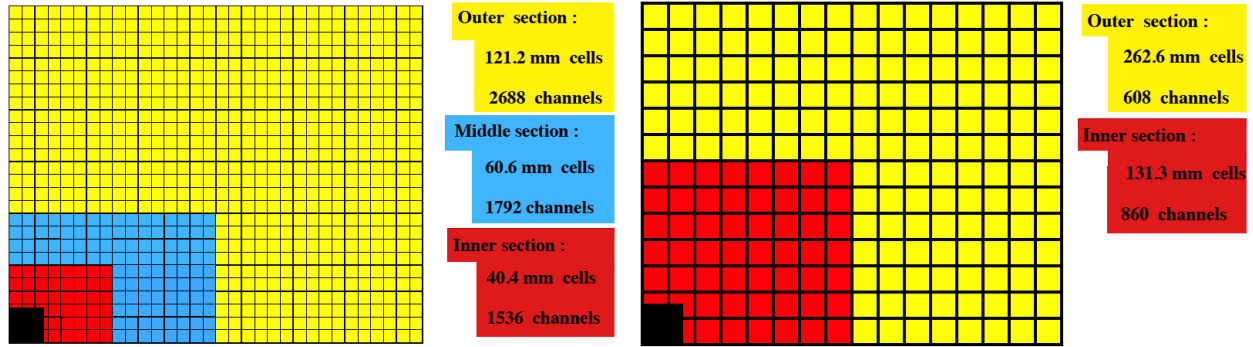


Figure 2.32: Illustration of the segmentations of the SPD/PS, ECAL (left) and HCAL (right). One quarter of the detector front face is shown.

The basic principle of the calorimeters is: a charged particle loses energy through electromagnetic interaction with Coulomb fields in traversing a transparent material. The molecules of the material then absorb the energy and become excited state. In order to the ground state, the excited molecules will release a small fraction of this energy as photons. This process is known as scintillation. Then the emitted photons are absorbed, re-emitted and transported to the PMTs by WLS fibers. The PMT is a vacuum photodetector which is extremely sensitive to the light in the ultraviolet, visible, and near-infrared ranges of the electromagnetic spectrum. It consists of a photocathode, focusing electrode, several dynode stages and an anode. The photons are converted into photoelectrons by the photocathode. These photoelectrons are directed by the focusing electrode toward the multiplier, where electrons are multiplied by the process of secondary emission. Finally this large number of photoelectrons reaches the anode, resulting in a sharp current pulse.

The energy resolution  $(\sigma(E)/E)$  of the ECAL module has been determined at the test beam. The result shows that the energy resolution satisfies the LHCb requirement, and can be parametrized as  $10\%/\sqrt{E} \oplus 1.5\%$ , where  $E$  is energy in GeV and  $\oplus$  means sum in quadrature (see figure 2.33) [26].

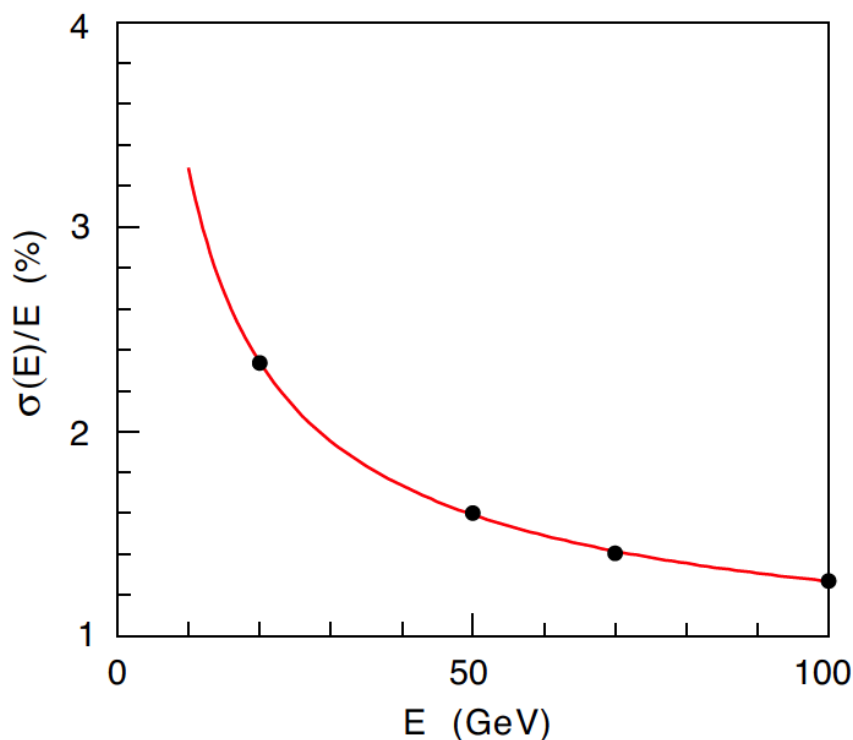


Figure 2.33: The energy resolution measured in the outer region of the ECAL (sources are electrons).

A comprehensive study has been performed to determine the degradation of the ECAL resolution under irradiation. In order to obtain a longitudinal dose profile as close as possible to the one expected under LHC running conditions, special modules made by 20 layers of 6 mm thick lead plates and 4 mm thick scintillating tiles were irradiated with a total dose of 5 Mrad at a dose delivery rate of 10 Rad/s. Figure 2.34 shows the measured degradation in light yield and transmission for irradiated scintillating tiles as a function of the distance to the PMT [26]. This delivery rate is 200 times higher than the rate expected at LHCb. And the total irradiation dose is equivalent to the radiation dose for the ECAL over 10 years of LHCb operation. However, taking into account the uncertainties on the expected irradiation dose, the ECAL is designed such that the inner most modules closest to the beam pipe can be replaced when necessary.

The readout system of the ECAL consists of 192 front-end boards (FEB) located in 14 crates

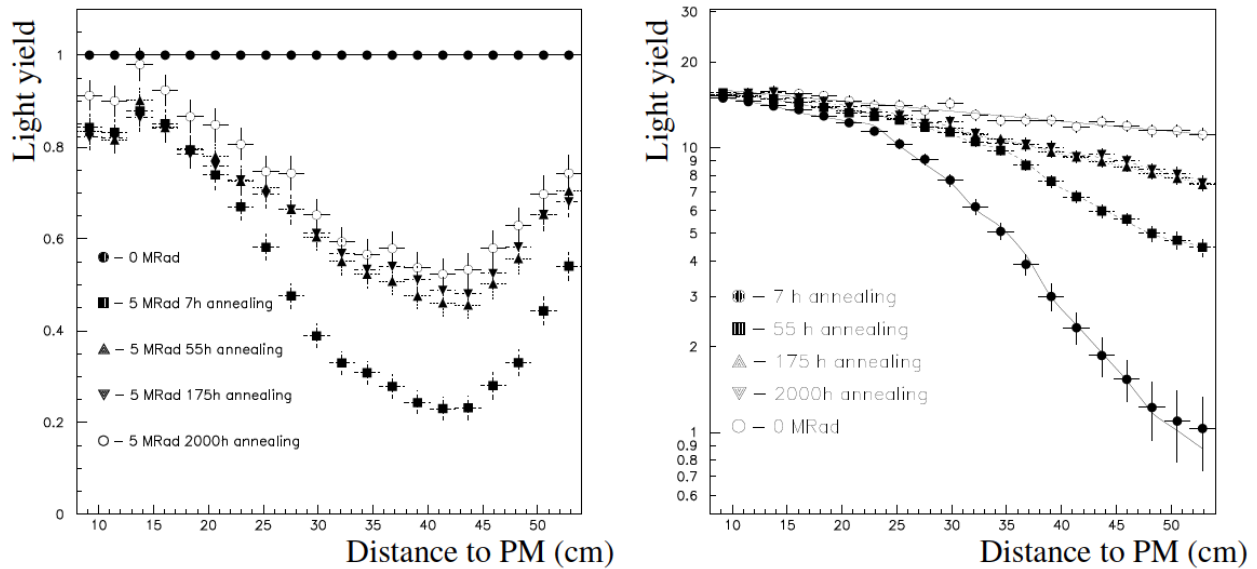


Figure 2.34: Scintillator (left) and WLS fiber (right) degradation and annealing effect after irradiation at LEP Injector Linac (LIL) up to 5 Mrad dose.

on the top of the detector. Each FEB reads out up to 32 channels of the ECAL. The PMT pulses are integrated by shaper integrator and digitized by 40 MHz two stage bipolar flash analog-to-digital converters (ADC). Then the digital data is sent to a trigger validation card and also pipelined and stored on the FEB. Finally, the data in FEB format is sent to the data acquisition (DAQ) if the event is accepted for further processing.

#### 2.3.4.4 Hadronic Calorimeter

The Hadronic Calorimeter (HCAL) is located downstream of the ECAL. The general view is shown in figure 2.31. The HCAL is a sampling device made from thin iron plates interspaced with scintillating tiles. One big difference between the HCAL and ECAL is the structure, the scintillating tiles and iron plates are arranged parallel to the beam pipe in the HCAL. In the longitudinal direction the length of tiles and iron spacers corresponds to the nuclear interaction length ( $\lambda$ ) in the steel. The light is collected by WLS fibers running along the detector towards the back side where the PMTs are situated (see figure 2.35). The HCAL consists of two identical halves, each of which is built from 26 stacked horizontal modules

piled on top of each other and positioned on movable platforms. Similar to the ECAL, the HCAL has inner and outer regions with the size of  $13 \times 13 \text{ cm}^2$  and  $26 \times 26 \text{ cm}^2$ , respectively. The lateral segmentation of the HCAL is shown in figure 2.35).

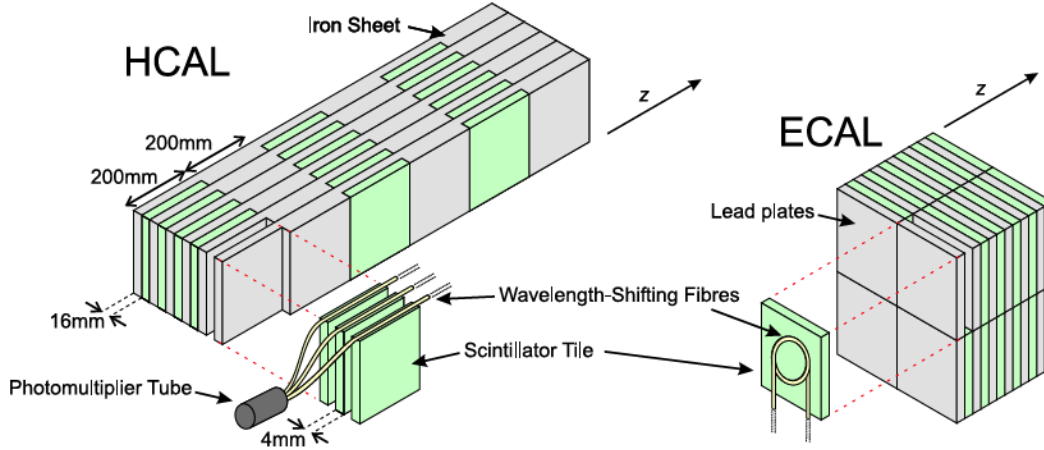


Figure 2.35: Schematic of the HCAL (left) and ECAL (right) structures.

The HCAL is the last station for  $p$ ,  $\pi^\pm$  and  $K^\pm$ . These nucleons and hadrons deposit their remaining energies in the form of hadronic showers via the strong nuclear force. The hadronic showers are similar to the electromagnetic showers but more complex. In traversing a material, the hadrons interact with the nucleus of that material, resulting in the production of multiple particles such as pions and nucleons. These secondary particles continue interacting with the material until the energies are expended. The hadronic multiplication process is dominated by a succession of inelastic hadronic interactions. This process is measured at the scale of the nuclear interaction length which is defined as:

$$\lambda = \frac{A}{\rho \sigma_I N_A} \quad (2.3)$$

where  $A$  is atomic weight,  $\rho$  is density,  $\sigma_I$  is the inelastic cross section and  $N_A$  is Avogadro's number. The nuclear interaction length is independent of energy. At high energies, the



maximum shower length can be described by a simple parameterization in terms of  $\lambda$

$$L_{max}(\lambda) \simeq 0.2\ln(E) + 2.5\lambda E^{0.13} + 0.7 \quad (2.4)$$

where  $E$  is energy in GeV.

In the LHCb, the main purpose of the HCAL is to provide hadron trigger for high transverse momentum hadrons. The required energy resolution of  $80\%/\sqrt{E}$  is quite moderate. By fitting the energy spectrum with a gaussian distribution around  $\pm 2.5\sigma$ , the energy resolution of the HCAL is  $(69 \pm 5)\%/\sqrt{E} \oplus (9 \pm 2)\%$  ( $E$  in GeV) as shown in the figure 2.36 [26].

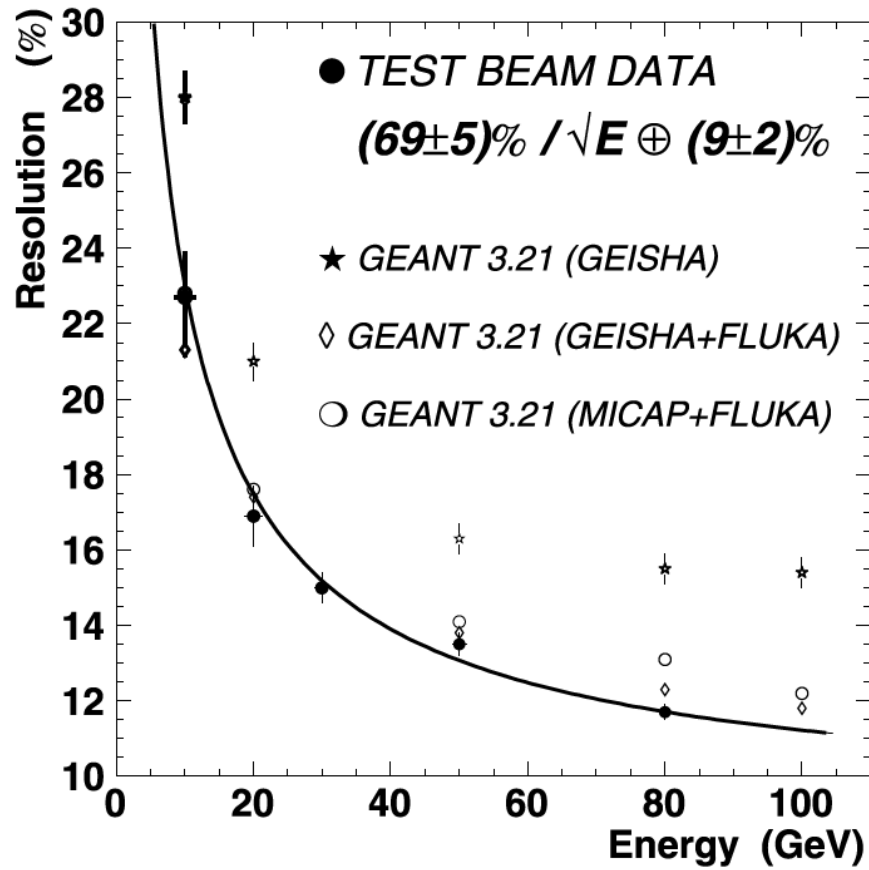


Figure 2.36: The energy resolution of the HCAL as a function of energy.

### 2.3.4.5 Muon System

The muon system comprises five stations of rectangular shape. They are situated along the beam pipe and separated by iron filters. The first muon station (M1) is located between RICH2 and ECAL, at 12.1 m from the interaction region. It is important for the  $p_T$  measurement of the muon track used in the Level-0 trigger. The remaining four stations (M2-M5) are placed downstream of the HCAL. The system covers an acceptance of  $\pm 20$  mrad to  $\pm 306$  mrad horizontally and  $\pm 16$  mrad to  $\pm 258$  mrad vertically. A side view of the muon system is shown in figure 2.37.

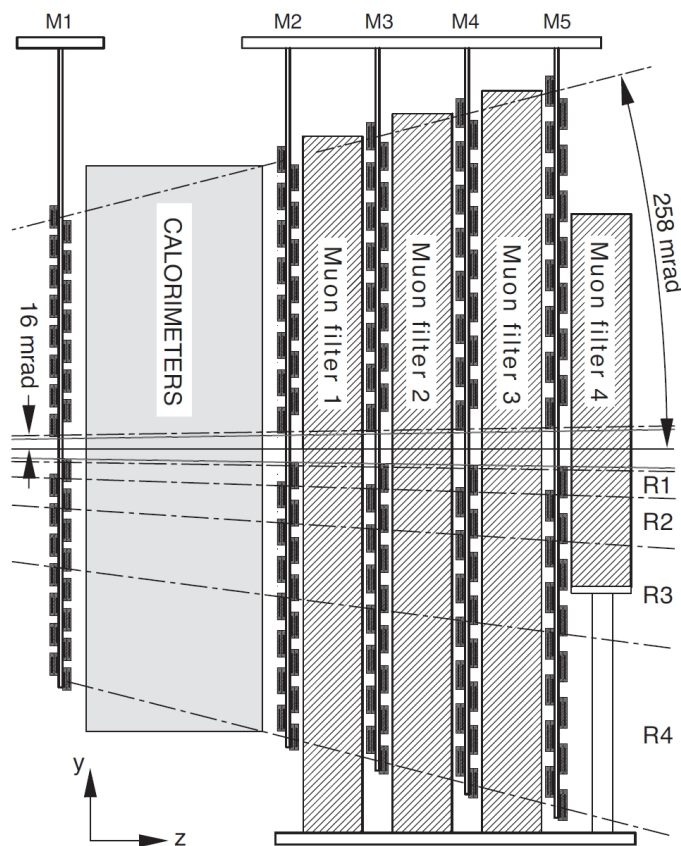


Figure 2.37: Side view of the muon system.

The muon system is designed to provide the measurements of muons for Level-0 trigger and identification of muons for the high-level trigger (HLT) and offline analysis. Muon

identification is important since they are present in the final stages of many B meson decays, such as the “gold-plated” decays  $B_d^0 \rightarrow J/\psi(\mu^+\mu^-)K_s^0$  and  $B_s^0 \rightarrow J/\psi(\mu^+\mu^-)\phi$ , where the muons play a major role in CP asymmetry and oscillation measurements, and the decay  $B^+ \rightarrow X(3872)(J/\psi(\mu^+\mu^-)\rho(\pi^+\pi^-))K^+$  which is analyzed to determine the X(3872) quantum number in this thesis.

In order to have a uniform occupancy over the detector, each station of the muon system is divided into four regions, R1 to R4 with different logical pad dimensions [37]. Their logical pad segmentations scale in the ratio 1:2:4:8 (see figure 2.38). The logical layouts in the five muon stations are projective vertically to the interaction point which is illustrated in figure 2.37.

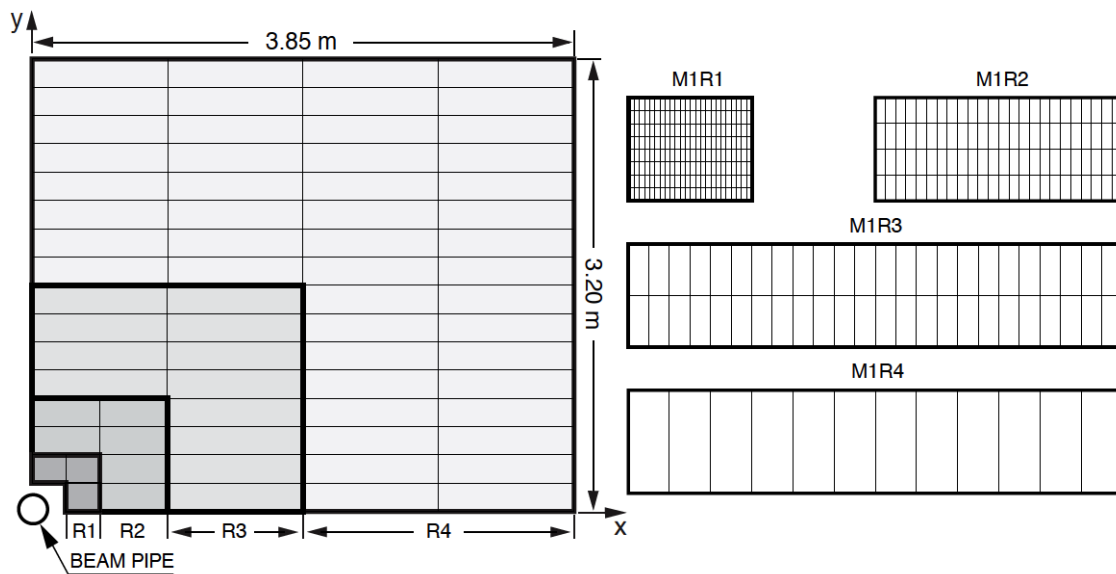


Figure 2.38: Front view of one quadrant of the first muon station (M1).

The trigger algorithm requires a five-fold coincidence among all the stations, so the efficiency of each station must exceed 95% within a time window smaller than 25 ns, which requires the detector to have excellent time resolution and redundancy [38]. To satisfy this requirement, Multi-Wire Proportional Chambers (MWPC) and triple Gas Electron Multiplier detectors (triple-GEM) are chosen for the muon system. There are a total of 1368 MWPCs used for all regions except the R1 region of M1 station, and 12 sets of triple-GEMs used in region M1R1 where the particle flux is highest.

The MWPC is a type of proportional counter constructed with alternating planes of high voltage wires and sense wires which are grounded or connected to negative voltage. The chamber is filled with fast, non-flammable gas mixtures of  $Ar/CO_2/CF_4$  with the ratio 40:55:5. Prototype studies show that a time resolution of about 5 ns can be achieved in a gas gap with 2 mm wire spacing and 5 mm gas gap [39]. There are three different types of panels: Outer Cathode Panel (OCP), Inner Cathode Panel (ICP) and Central Cathode Panel (CCP). The ICP is wired on both sides while OCP and CCP are not wired. A typical chamber is made of four equal gas gaps (2 OCPs, 2 ICPs and 1 CCP) as shown in figure 2.39, while the MWPC equipped on M1 has two gaps (2 OCPs and 1 ICP).

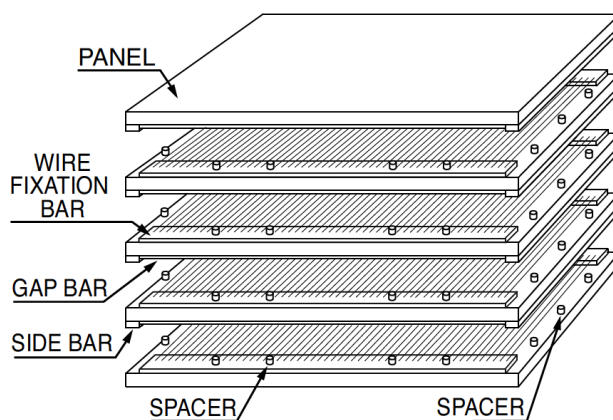


Figure 2.39: Exploded schematic of a four-gap MWPC.

When a charged particle passes through the chamber, it ionizes the gas molecules. The freed electrons are accelerated by the electric field around the wire, ionizing more of the gas. Finally the localized cascade of ionization is collected on the wire and results in a negative signal proportional to the energy of the detected particle.

The 12 pairs of Triple-GEMs are installed in the innermost region R1 of the muon station M1. Rather than the MWPC, triple-GEM is chosen for its outstanding tolerance of the irradiation. The triple-GEM consists of three copper-cladded Kapton foils of  $50 \mu m$  thickness sandwiched between anode and cathode planes (see figure 2.40). Each foil has a large density of holes arranged in a triangular pattern with a pitch of  $p = 140 \mu m$ . Each hole has a bi-conical

structure with internal diameter of  $d = 50 \mu\text{m}$  and external diameter of  $D = 70 \mu\text{m}$  [40]. The bi-conical shape can minimize the effect of charging-up of the Kapton inside the holes. A schematic drawing of the GEM foil is illustrated in figure 2.40. Like the MWPC, the triple-GEM is also filled with the gas mixtures of  $\text{Ar}/\text{CO}_2/\text{CF}_4$  but with a different ratio 45:15:40. The triple-GEMs operate on a similar principle to the MWPCs. The multiplied electrons are collected by the anode plane which is connected to the readout electronics.

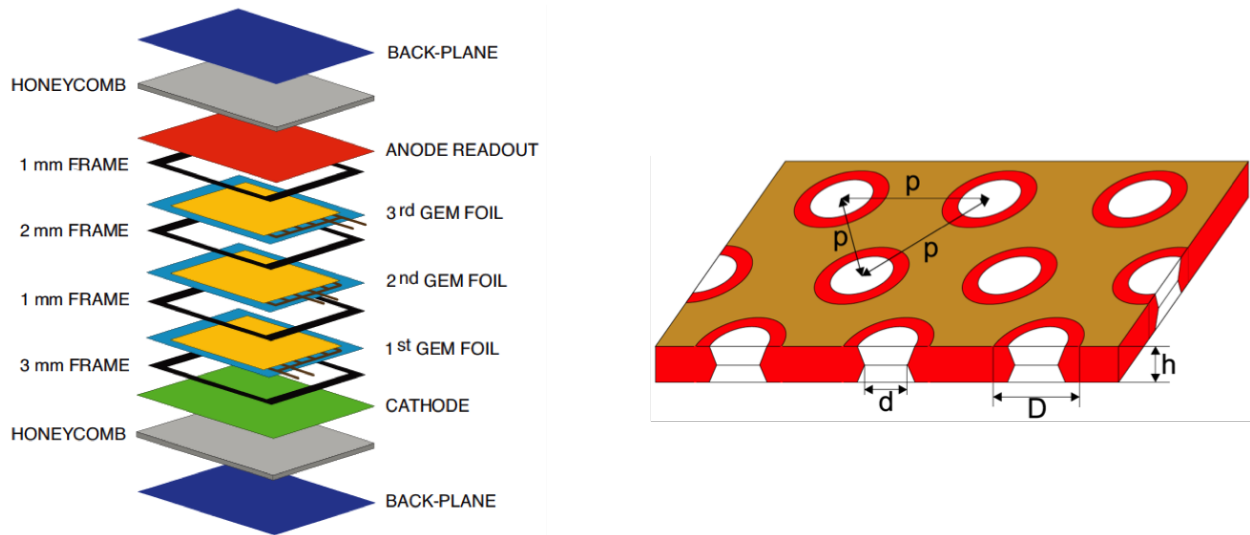


Figure 2.40: Left: exploded schematic of the GEM. Right: illustration of the GEM foil.

The readout electronics of the muon system are mainly used in preparation of the information for the Level-0 muon trigger and transmission of the data. The readout is performed via the FEBs where the fast amplification, shaping and discrimination of the signals are implemented. Then the logical-channel signals generated by suitable logical ORs of the physical channels are sent into the Off Detector Electronics (ODE) boards. The signals then are tagged with the number of the bunch crossing (BX identifier) and dispatched to the Level-0 muon trigger processors. The signals are also transmitted to TELL1 board and from TELL1 to the DAQ system with fine time information.

### 2.3.4.6 Summary of PID Devices

Particle identification is crucial for the LHCb experiment as it reduces background and separates particles. The PID is performed with the whole system: the two RICH detectors, the SPD/PS detectors, the calorimeters and the muon system. Their information is combined for optimal identification of charged particles types ( $e$ ,  $\mu$ ,  $\pi$ ,  $K$ ,  $p$ ). Photons and neutral pions ( $\gamma$ ,  $\pi^0$ ) are identified using the SPD/PS detectors and ECAL.

### 2.3.5 Trigger

The LHCb experiment was operated at an average luminosity of  $2 - 4 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ , with 40 MHz frequency of bunch crossings at the interaction point inside LHCb. But only 10 MHz of events are visible by the spectrometer. However, due to the limited offline computing capacity, this rate has to be reduced by the trigger to about 2 - 3 kHz. And it is organized in two trigger levels: Level-0 (L0) and the High Level Trigger (HLT). The L0 is a hardware level trigger implemented in custom-designed electronics, while the HLT is a software level trigger executed on a farm of commodity processors.

#### 2.3.5.1 The Level-0 Trigger

The Level-0 (L0) trigger is used in reducing the 40 MHz LHC beam crossing rate to below 1.1 MHz with which the entire detector can be read out. It attempts to reconstruct photon, electron and hadron with the highest transverse energy ( $E_T$ ) in the calorimeters, and two muons with the highest transverse momentum ( $p_T$ ) in the muon system. The L0 consists of three independent triggers: the L0-PileUp detector, the L0-Calorimeter trigger and the L0-Muon trigger (see figure 2.41). Each component is connected to detector and to the L0 Decision Unit (DU) which evaluates the final decision by collecting all information calculated

by the trigger systems.

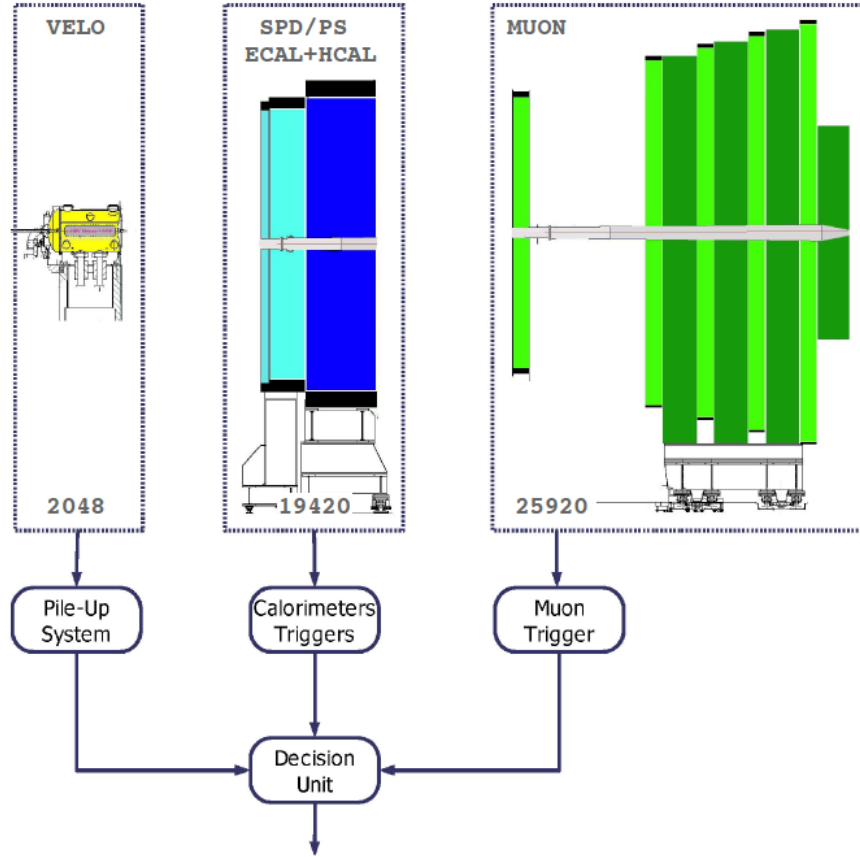


Figure 2.41: General view of the Level-0 trigger.

The L0-PileUp trigger system comprises two planes perpendicular to the beam pipe located at the very front of the detector. Each  $300 \mu\text{m}$  thick plane has two overlapping R sensors measuring the radial position of tracks. The L0-PileUp trigger system aims at determining the number of primary  $p\bar{p}$  interactions in each bunch crossing. The system also provides the position of the primary vertices along the beam pipe. The two track hits  $r_a, r_b$  are measured by the two silicon planes as illustrated in figure 2.42. The hits belonging to the tracks from the same origin have the simple relation  $k = r_a/r_b$ , giving:

$$z_v = \frac{kz_a - z_b}{k - 1} \quad (2.5)$$

where  $z_a$  and  $z_b$  are the sensor positions and  $z_v$  is the vertex position on the beam axis [41].

The  $z_v$  values calculated from the hits in the same octant of both planes are projected into a histogram, in which a peak search is performed (see figure 2.42). Due to the effect of multiple scattering and the hit resolution of the R sensor, the resolution of  $z_v$  is limited to around 3 mm.

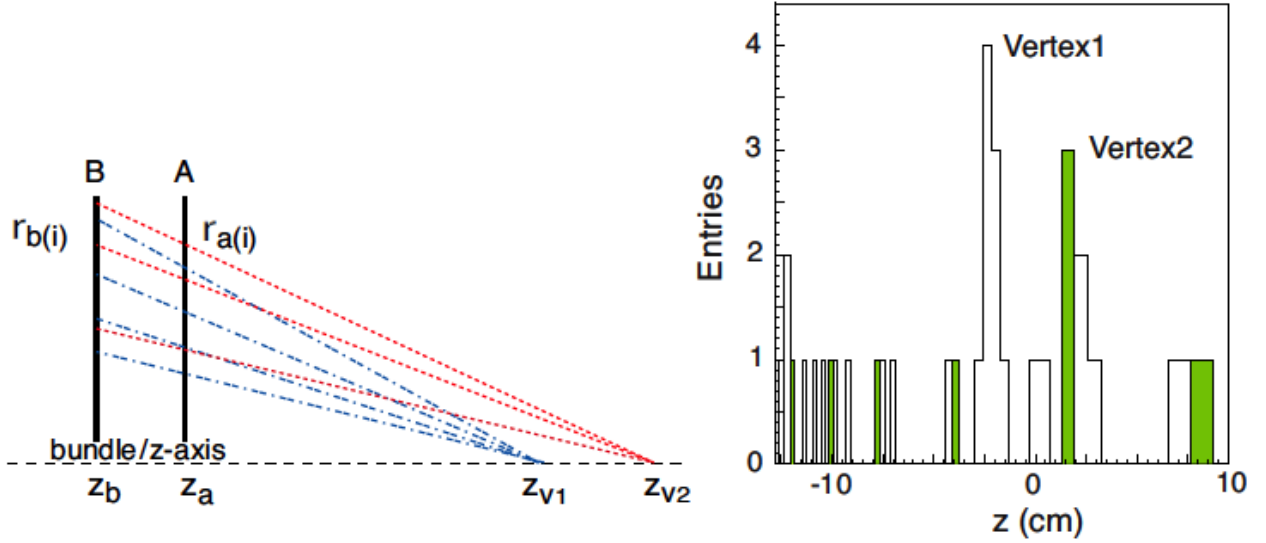


Figure 2.42: Left: basic principle of detecting vertices for the L0 Pile-Up trigger system. Right: histogram of the values of  $z_v$  calculated from the hits of the two planes.

The L0-Calorimeter trigger system uses information from the SPD/PS detectors and two calorimeters to select and identify particles with high transverse energy ( $E_T$ ) deposit in the calorimeters, and to estimate the number of tracks based on the number of hits in the SPD detector. The  $E_T$  is defined as:

$$E_T = \sum E_i \sin \theta_i \quad (2.6)$$

where  $E_i$  is the energy deposited in cell  $i$  of the  $2 \times 2$  cluster and  $\theta_i$  is the angle between the beam axis and a neutral particle assumed to be coming from the mean position of the interaction envelope hitting the centre of the cell [42]. The basic idea of the L0-Calorimeter trigger system is to search for high  $E_T$  particles: photons, electrons, neutral pions or hadrons. This is performed in three steps:

- Selection of high  $E_T$  deposits. This step is performed on the FE card.



- Identification of the type of electromagnetic candidate. Only one candidate is selected and sent to the next stage for each type.
- Selection of the candidate per type with the highest  $E_T$ . This step is implemented on the Selection Crate.

The L0-Muon trigger system is designed to search for muons with a large transverse momentum ( $p_T$ ): a typical signature of a b-hadron decay. The system consists of five muon stations (M1-M5), of which each quadrant is connected to a L0 muon processor. The L0 muon processor selects two muon tracks with the largest and second largest  $p_T$  for each quadrant. The track finding is performed on the logical pad layout. It looks for hits defining a straight line through the five muon stations and pointing towards the interaction point (see figure 2.43). The position of a track in the first two stations allows the determination of its  $p_T$ .

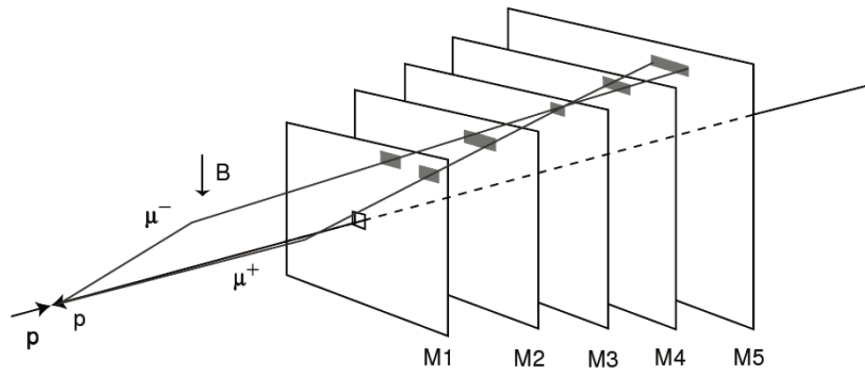


Figure 2.43: Tracking finding in the muon stations.

The L0 Decision Unit (DU) collects all information from the three sub detectors at 40 MHz to form the trigger decision by performing simple logic. The trigger decision is sent to the readout supervisor to determine to accept or reject the event. The readout supervisor can generate and time-in all types of self-triggers. It is also able to control the trigger rate by taking into account the status of the different components to avoid buffer overflows.

### 2.3.5.2 The High Level Trigger

The High Level Trigger (HLT) is the second and last level of trigger. The HLT is a program (called “Moore”) written in C++ and python, and 26110 copies of it run on the Event Filter Farm (EFF) of 1000 16-core computers. The HLT application has access to full detector information and aims at reducing the rate of events from  $\sim 1$  MHz passing the L0 trigger to  $\sim 2$  kHz. The HLT is divided into two sub-levels: HLT1 and HLT2. Figure 2.44 illustrates the data flow through the triggers of L0 and HLT.

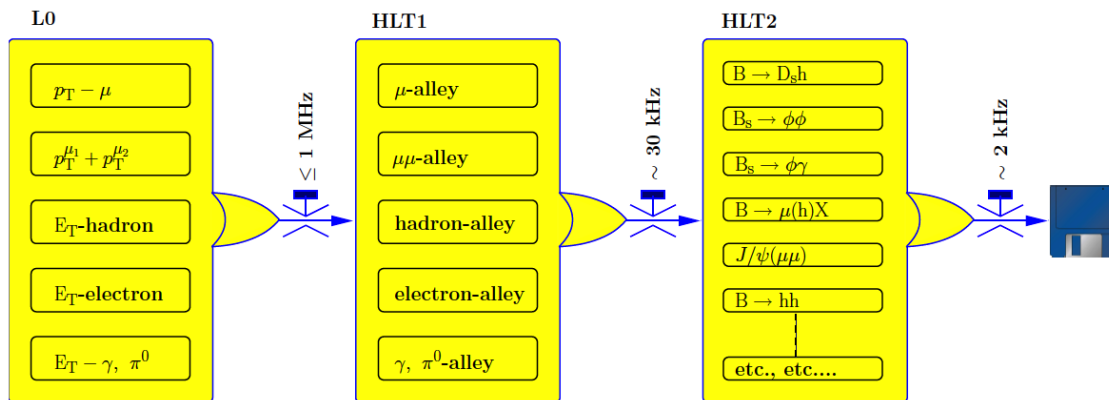


Figure 2.44: Flow-diagram of the data through the triggers.

The HLT1 performs a partial reconstruction of the particles in the VELO and T-stations, and determines the position of the Primary Vertices (PV) in the event. According to the properties of B mesons, a combination of selections such as Impact Parameter (IP), transverse momentum ( $p_T$ ) and invariant mass is implemented to the L0 candidate particles in the HLT1. This is done to reduce the rate from the  $\sim 1$  MHz output of L0 to  $\sim 30$  kHz.

Since the rate has been significantly reduced by the HLT1, the HLT2 performs a complete reconstruction of the particles and searches for the vertices away from the primary interaction. The HLT2 consists of a set of inclusive and exclusive selections with the cuts looser than in the offline selections. Inclusive selections are designed to select generic B decays of resonances which are useful for calibration, while exclusive selections aim to provide the highest possible efficiency on specific B decay channels, using all available information. At this stage the rate

of event is reduced to 2 kHz which is upper limit that can be handled and saved on the disk.

TIS and TOS are the two main categories to assort the trigger lines. If the event is triggered by those trigger objects that are independent to the signal, the event is classified as TIS (Trigger Independent Signal). If the event is triggered by those trigger objects that are associated with the signal are sufficient to trigger the event, the event is classified as TOS (Trigger On Signal). All the information needed for such a classification is recorded by the trigger.

### 2.3.6 The LHCb Software

The LHCb online and offline software is based a well-structured C++ framework known as Gaudi. This framework provides event simulation, reconstruction, visualization and a set of tools which facilitate the writing of analysis programs. There are five main applications developed within the Gaudi framework: Gauss, Boole, Brunel, DaVinci and Moore. The Gaudi architecture is schematically illustrated in figure 2.45.

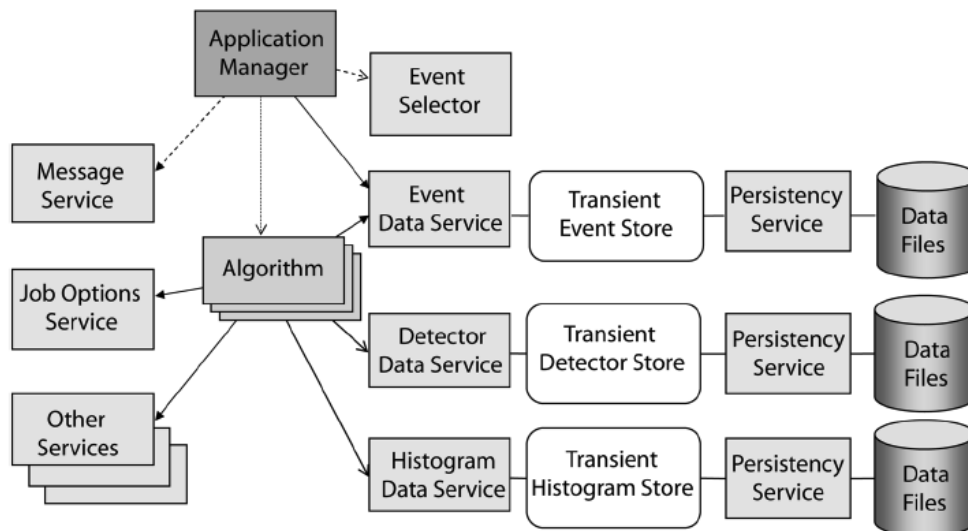


Figure 2.45: Object diagram of the GAUDI architecture.

Gauss is the simulation software of the LHCb experiment. It consists of a generator phase where the  $p\bar{p}$  collisions are generated and the decay of the particles are produced, and a

simulation phase where the particles are propagated through the LHCb detector. The main event generator of Gauss is the Pythia which produces both the minimum-bias and the signal events. The B meson decay is managed by the EVTGEN which is another external package used in the generation phase. The simulation process is achieved using GEANT4. The geometry and material of the detector are precisely described in the GEANT4 to simulate the physics processes occurring in the experimental setup.

Boole is the digitization program of the LHCb experiment. It is a part of the simulation process. Boole reads in the output of detector simulation, adds additional hits from Spillover events and LHC background. It also implements the simulation of the detector response and of the readout electronics. The output is digitized data in the same format as the real data coming from the real detector.

Brunel is the reconstruction program of the LHCb experiment. It can process either the digitized data from Boole or read data from the LHCb DAQ system. Brunel uses pattern recognition and performs reconstruction of the track and identification of particle. The output consists of all the information for every triggered event and is stored in the DST<sup>7</sup> files.

DaVinci is the physics analysis software for the LHCb experiment. It supports reconstruction of the decay sequence from processing of the DST data. It provides tools for the manipulation and analysis of the physics event objects, and for the evaluation of the physics performance of the code are provided to enable study and comparison with the Monte Carlo truth information. The output of DaVinci can be purely statistical or event data with the physics objects for further processing. The output of DaVinci can also be a reduced DST with the events satisfying certain conditions [43].

Moore can process identically the results of digitized data from Boole and real data from the LHCb DAQ system. It is used to execute high level software trigger algorithm.

---

<sup>7</sup>DST stands for Data Summary Tape, though these files are often stored on disks

Distributed Infrastructure with Remote Agent Control (DIRAC) provides a uniform access to the computing and storage resources for LHCb users. It consists of the Workload Management System (WMS) and Data Management System (DMS). DIRAC is used to ensure that the LHCb virtual organization (VO) uses its resources efficiently and to enforce job prioritization [44].

Ganga is a job management system for the LHCb experiment. The package is based on the language of python. It is used for job submission and management on the computing Grid using Dirac.

### 2.3.7 Data Processing

The LHCb data, as well as the simulations undergo several stages of processing. The logical data flow is illustrated in figure 2.46. The raw data from the detector are transferred to the Tier-0 for further processing and archiving. Then the physical quantities of the data are obtained from partial reconstruction and saved as a reduced Data Summary Tape (rDST). The rDST includes the information of the particle identification, energy of electromagnetic and hadronic showers and trajectory in the T-stations. And it is used for determination of the four-momentum vectors of the measured particles, location of the primary and secondary vertices and reconstruction of the composite particles. The rDST is analyzed in a production-type mode to produce streams of selected events for further individual analysis. This activity is known as stripping. Full reconstruction is applied to the events passing the selection criteria in order to generate full DST. Before being stored, the events passing the selection criteria will have their raw data added in order to have as detailed information as needed for the analysis [26]. An event tag containing a brief summary of each event's characteristics is created for selected events. The tag is stored independent of the DST. The processing stages for simulated data are the same.

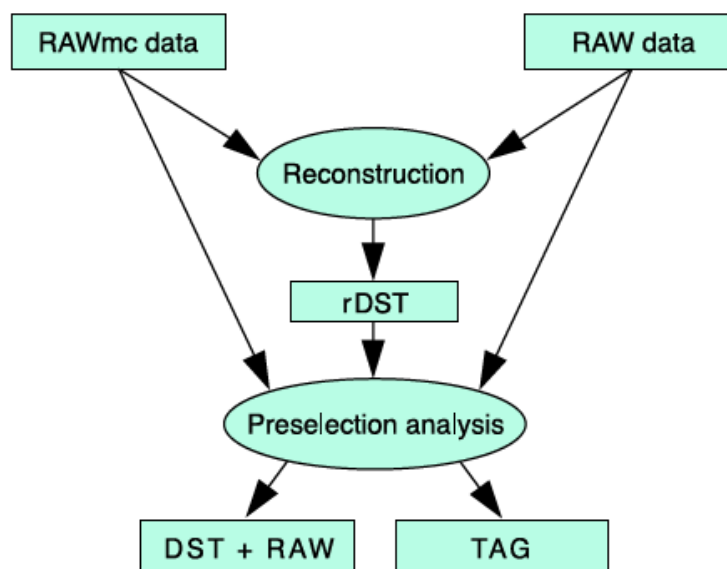


Figure 2.46: The computing logical data flow model of the LHCb.

Over  $3 \text{ fb}^{-1}$  data has been collected by LHCb since 2010 (see figure 2.47). The raw data are stored in Tier-1 sites which are also the computing centers located in Europe. There are total of seven Tier-1 sites: CERN, CNAF (Italy), FZK (Germany), IN2P3 (France), NIKHEF (The Netherlands), PIC (Spain) and RAL (United Kingdom). As the central production site, CERN keeps an entirety copy of the raw data and has the responsibility of distributing another copy of the raw data in quasi-real time across the other six Tier-1 sites. A schematic of the LHCb computing model is shown in figure 2.48. The majority of the distributed analyses are performed at the Tier-1 sites (including CERN) which are also in charge of all the production processing phases associated with the real data. There is also a number of Tier-2 computing sites which primarily takes care of the Monte Carlo production.

The reconstruction and the first stripping of data at the Tier-1 sites are expected to follow the production in quasi real-time, with a maximum delay of a few days. The stripping productions remain on disk for analysis and are distributed to all other Tier-1 sites, while the raw and rDST are migrated to the Mass Storage System (MSS). The re-processing of the data happens over a two-month period. During this process the raw data need to be accessed from the MSS both at the Tier-1 sites. The CPU resources available at the pit allow a significant fraction of

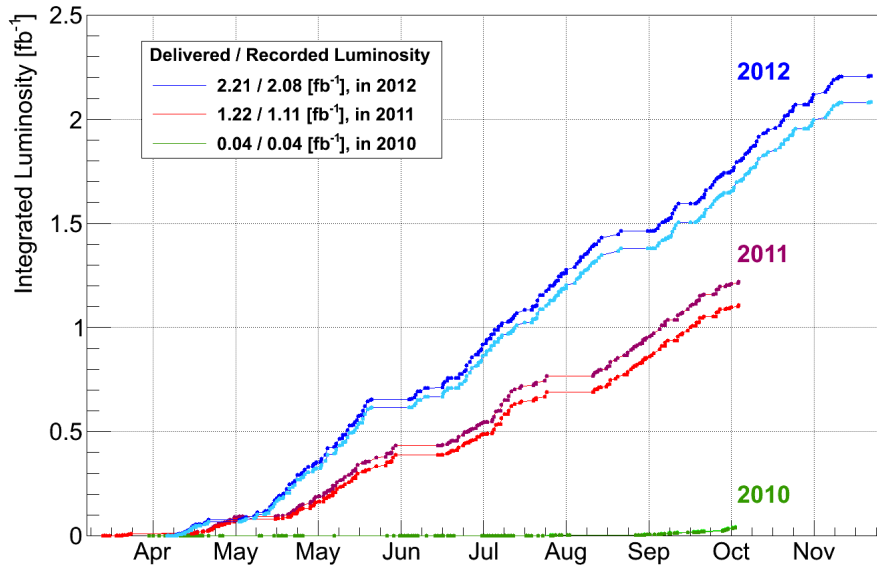


Figure 2.47: The delivered and recorded luminosity of LHCb detector, 2010 ~ 2012.

the total re-processing and perhaps the subsequent stripping to be performed there. Hence at CERN there is an additional complication that the raw data also have to be transferred to the pit; similarly, the produced rDST has to be transferred back to the CERN computing centre. To enable later stripping, it is necessary to distribute a fraction of the rDST produced at CERN during this re-processing to the other Tier-1 sites; this is a consequence of the large contribution from the online farm.

The stripping productions outside of the reconstruction or of the re-processing of the data are performed over a one-month period. Both the raw and the rDST need to be accessed from the MSS to perform this production. The produced stripped DSTs are distributed to all production sites.

### 2.3.8 Summary

The LHC has delivered  $p\bar{p}$  collisions at a center-of-mass energy of  $\sqrt{s} = 7$  TeV (2010, 2011) and  $\sqrt{s} = 8$  TeV (2012). The LHCb spectrometer has been operated in a high irradiation

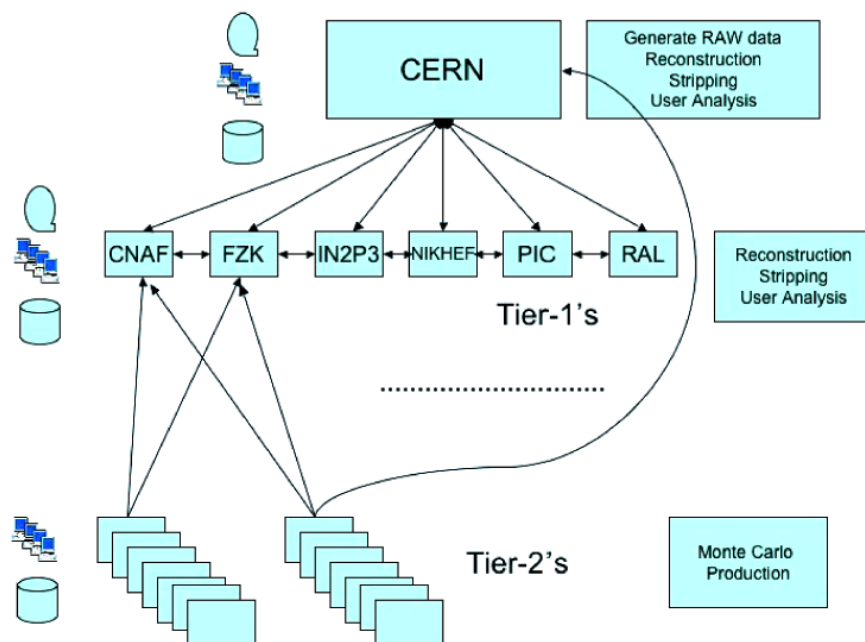


Figure 2.48: The LHCb computing model.

environment since 2010 with a small fraction of channels failed. The integrated luminosity recorded by LHCb is more than  $3 \text{ fb}^{-1}$ . We used first  $1 \text{ fb}^{-1}$  recorded by LHCb to determine  $X(3872)$  quantum numbers, as the remaining  $2 \text{ fb}^{-1}$  was in early stages of data processing when this analysis was performed.



# Chapter 3

## Measurement

Comparing to the signals obtained by Belle, we almost double the signal statistics using the 2011 data sample by employing the multivariate data selection previously developed for the published analyses of  $B_c^+ \rightarrow J/\psi \pi^+ \pi^- \pi^+$  [45] and  $B^+ \rightarrow J/\psi K^+ K^- K^+$  [46] decays. More importantly, we analyze 5-dimensional correlations in the full angular phase space for  $B^+ \rightarrow X(3872)K^+$ ,  $X(3872) \rightarrow \pi^+ \pi^- J/\psi$ ,  $J/\psi \rightarrow \mu^+ \mu^-$  decay, which allows us to resolve the  $1^{++}$  and  $2^{-+}$  with arbitrary complex value of  $\alpha^1$  with overwhelming significance, and therefore, we determine  $J^{PC}$  of  $X(3872)$  with no ambiguity for the first time.

### 3.1 Data Selection

The data selection is almost the same as used in the published  $B_c^+ \rightarrow J/\psi \pi^+ \pi^- \pi^+$  [45] and  $B^+ \rightarrow J/\psi K^+ K^- K^+$  [46] analyses and described in the previous note [47]. The  $J/\psi \rightarrow \mu^+ \mu^-$  candidates are selected with the following criteria:  $p_T(\mu) > 900$  MeV<sup>2</sup>,  $p_T(\mu^+ \mu^-) > 1500$  MeV,  $\chi^2$  per degree of freedom for the two muons to form a common vertex,  $\chi_{\text{vtx}}^2(\mu^+ \mu^-)/\text{NDOF} < 9$ , a mass consistent with the  $J/\psi$  meson (3040 – 3140 MeV), and the separation of the  $J/\psi$

---

<sup>1</sup> $\alpha$  is defined as  $B_{11}/(B_{11} + B_{12})$ , where  $B_{LS}$  are the two possible amplitudes for the  $2^{-+}$  hypothesis in which the lowest orbital angular  $L$  between the  $\pi\pi$  and  $J/\psi$  is assumed to dominate ( $L = 1$ ).

<sup>2</sup> $p_T$  is the component of momentum in the transverse plane.

decay vertex from the nearest primary vertex (PV) must be at least three standard deviations. Combinations of  $K^+\pi^-\pi^+$  candidates that are consistent with originating from a common vertex with  $\chi_{\text{vtx}}^2(K^+\pi^-\pi^+)/\text{NDOF} < 9$ , with each charged hadron ( $h$ ) separated from all PVs ( $\chi_{\text{IP}}^2(h) > 9$ ) and having  $p_{\text{T}}(h) > 250$  MeV, are selected. The quantity  $\chi_{\text{IP}}^2(h)$  is defined as the difference between the  $\chi^2$  of the PV reconstructed with and without the considered particle. Kaon and pion candidates are required to satisfy  $\text{Ln}[L(K)/L(\pi)] > 0$  and  $< 5$ , respectively, where  $L$  is the particle identification likelihood [35]. If both same-sign hadrons in this combination meet the kaon requirement, only the particle with higher  $p_{\text{T}}$  is considered a kaon candidate. We combine  $J/\psi$  candidates with  $K^+\pi^-\pi^+$  candidates to form  $B^+$  candidates, which must satisfy  $\chi_{\text{vtx}}^2(J/\psi K^+\pi^-\pi^+)/\text{NDOF} < 9$ ,  $p_{\text{T}}(B^+) > 2000$  MeV and have decay time greater than 0.25 ps. The  $J/\psi K^+\pi^-\pi^+$  mass is calculated using the known  $J/\psi$  mass and the  $B$  vertex as constraints. The selection requirements are summarized in table 3.1.

Table 3.1: Data selection requirements.

Particle	Quantity	Criteria
All tracks	Track quality: $\chi^2/\text{nDoF}$	$< 4$
All tracks	Clone Dist.	$> 5000$
$\mu$	$p_{\text{T}}$	$> 0.9$ GeV
$\mu$	IsMuon	True
$\mu$	$\text{PID}_{\mu}(\text{DLL}(\mu-\pi))$	$> 0$
Di- $\mu$	Vertex quality: $\chi^2/\text{nDoF}$	$< 9$
Di- $\mu$	$p_{\text{T}}$	$> 1.5$ GeV
$J/\psi$	Mass window	[3.040,3.150] GeV
$\pi$	$p_{\text{T}}$	$> 0.25$ GeV
$\pi$	$\chi_{\text{IP}}^2$	$> 9$
$\pi$	$\text{PID}_{\pi}(\text{DLL}(K-\pi))$	$< 5$
$K$	$p_{\text{T}}$	$> 0.25$ GeV
$K$	$\chi_{\text{IP}}^2$	$> 9$
$K$	$\text{PID}_K(\text{DLL}(K-\pi))$	$> 0$
$\bar{B}$	Mass window	[5.261, 5.300] GeV
$\bar{B}$	Vertex quality: $\chi^2/\text{nDoF}$	$< 9$
$\bar{B}$	$p_{\text{T}}$	$> 2.0$ GeV
$\bar{B}$	Lifetime: $\tau$	$> 0.25$ ps
$\bar{B}$	DLL	$< 1$

Further background suppression is provided by an event selection based on a likelihood ratio. In the case of uncorrelated input variables, this provides the most efficient discrimination between signal and background. The overall likelihood is a product of the probability density functions (PDFs),  $\mathcal{P}(x_i)$ , for the four sensitive variables ( $x_i$ ): smallest impact parameter (IP) with respect to closest PV  $\chi_{\text{IP}}^2(h)$  among the  $K$  and  $\pi$ s,  $\chi_{\text{vtx}}^2(J/\psi K^+\pi^+\pi^-)/\text{NDOF}$ ,  $B^+$  candidate IP significance,  $\chi_{\text{IP}}^2(B^+)$ , and cosine of the largest opening angle between the  $\psi$  and the charged hadron candidates in the plane transverse to the beam. The latter peaks at positive values for the signal as the  $B^+$  meson has a high transverse momentum. Background events that combine particles from two different  $B$  mesons peak at negative values, whilst background events that include random combinations of tracks are uniformly distributed. We develop PDFs by selecting  $\psi(2S) \rightarrow \pi^+\pi^- J/\psi$  signal region,  $|M(J/\psi \pi^+\pi^-) - M(J/\psi) - 589.2| < 16.3$  MeV. This decay is kinematically similar to  $X(3872) \rightarrow \pi^+\pi^- J/\psi$  and offers high statistics  $B^+ \rightarrow \psi(2S)K^+$  control signal in the data. This approach produced a slightly better signal-to-background ratio for the  $X(3872)$  signal than developing PDFs on generic  $B^+ \rightarrow J/\psi K^+\pi^+\pi^-$  decays. The signal PDFs,  $\mathcal{P}_{\text{sig}}(x_i)$ , are obtained from a Monte Carlo simulation of  $B^+ \rightarrow \psi(2S)K^+$ ,  $\psi(2S) \rightarrow \pi^+\pi^- J/\psi$ ,  $J/\psi \rightarrow \mu^+\mu^-$  decays. The background PDFs,  $\mathcal{P}_{\text{bkg}}(x_i)$ , are obtained from the data with a  $(J/\psi \pi^+\pi^-)K^+$  invariant mass in the range 4.85 – 5.10 GeV or 5.45 – 6.50 ( $B^+$  far-sidebands). The signal and background PDFs for the four variables used in the definition of  $DLL_{\text{sig/bkg}}$  are shown in the integral form in figures 3.1-3.4. To check how well the simulations agree with the data we compare the signal PDFs between the  $B^+ \rightarrow \psi(2S)K^+$  MC and the  $B^+ \rightarrow \psi(2S)K^+$  signal in the data. To get an agreement in  $\chi_{\text{IP}}^2$  and  $\chi_{\text{vtx}}^2$  we have smeared these variables in the simulations. We apply the same smearing to the  $B^+ \rightarrow X(3872)K^+$  simulations.

We form the logarithm of the ratio of the signal and background PDFs,  $DLL_{\text{sig/bkg}} = -2 \sum_{i=1}^4 \ln(\mathcal{P}_{\text{sig}}(x_i)/\mathcal{P}_{\text{bkg}}(x_i))$ . The discrimination between the  $B^+ \rightarrow \psi(2S)K^+$  signal and the background in the far-sideband is illustrated in figure 3.5, which also shows a good agreement between the  $B^+ \rightarrow \psi(2S)K^+$  signal in the data and in the simulations. To

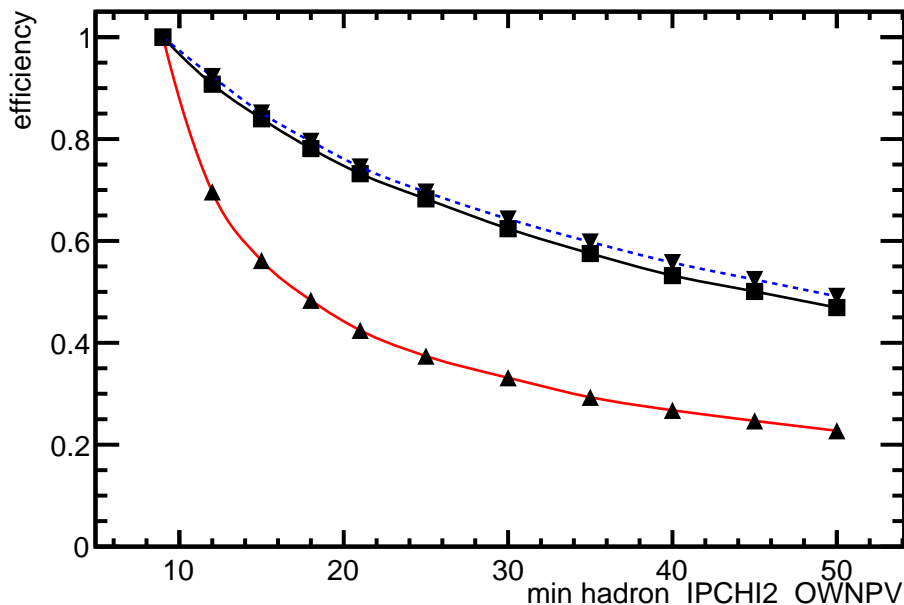


Figure 3.1: Fraction of events after the preselection passing a cut on  $\chi_{\text{IP}}^2(h) > X$ , where  $X$  is given on the horizontal axis, for  $B^+ \rightarrow \psi(2S)K^+$  Monte Carlo (solid black line with squares),  $B^+ \rightarrow \psi(2S)K^+$  signal in the data (blue dotted line) and background events in the data from the far-sideband (solid red line with triangles).

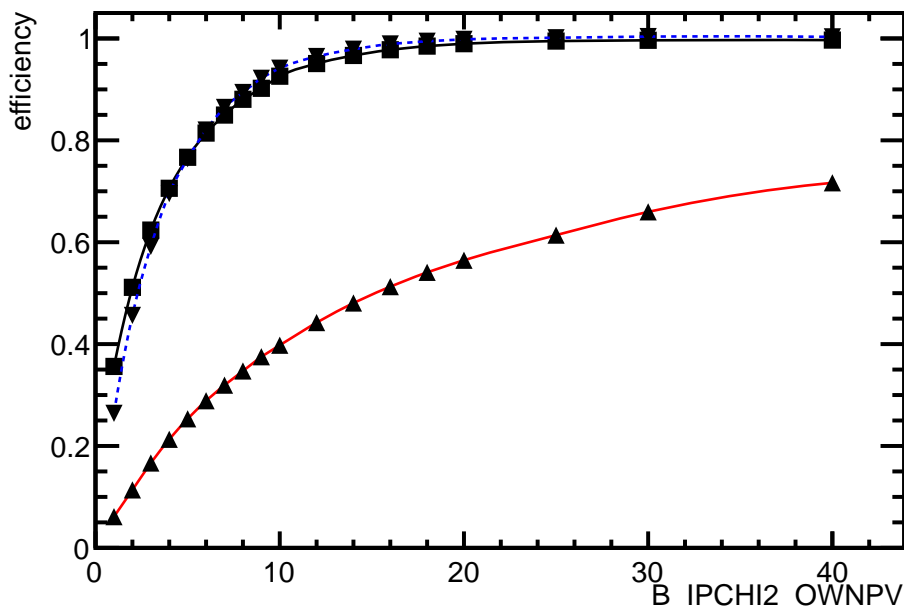


Figure 3.2: Fraction of events after the preselection passing a cut on  $\chi_{\text{IP}}^2(B^+) < X$ , where  $X$  is given on the horizontal axis, for  $B^+ \rightarrow \psi(2S)K^+$  Monte Carlo (solid black line with squares),  $B^+ \rightarrow \psi(2S)K^+$  signal in the data (blue dotted line) and background events in the data from the far-sideband (solid red line with triangles).

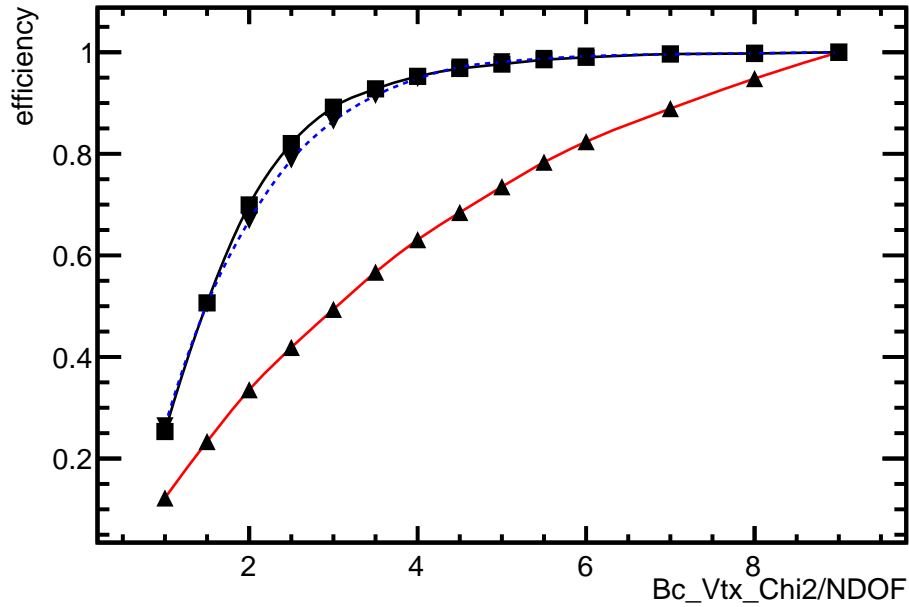


Figure 3.3: Fraction of events after the preselection passing a cut on  $\chi_{\text{vtx}}^2(B^+) < X$ , where  $X$  is given on the horizontal axis, for  $B^+ \rightarrow \psi(2S)K^+$  Monte Carlo (solid black line with squares),  $B^+ \rightarrow \psi(2S)K^+$  signal in the data (blue dotted line) and background events in the data from the far-sideband (solid red line with triangles).

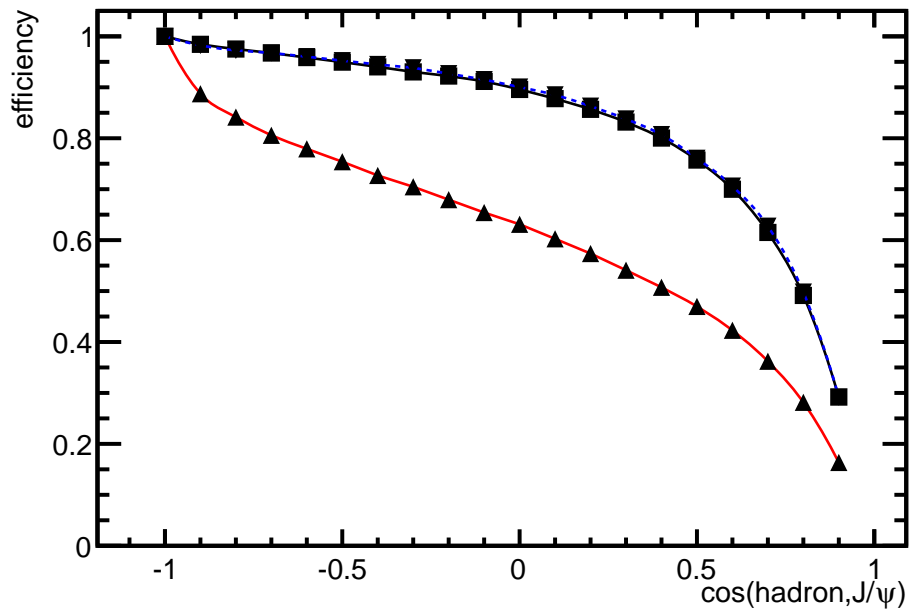


Figure 3.4: Fraction of events after the preselection passing a cut on cosine of the largest opening angle between the  $\psi$  and the charged hadron candidates,  $\cos(\text{hadron}, J/\psi) > X$ , where  $X$  is given on the horizontal axis, for  $B^+ \rightarrow \psi(2S)K^+$  Monte Carlo (solid black line with squares),  $B^+ \rightarrow \psi(2S)K^+$  signal in the data (blue dotted line) and background events in the data from the far-sideband (solid red line with triangles).

optimize a cut on  $DLL_{\text{sig/bkg}}$ , we switch to use of  $B^+ \rightarrow X(3872)K^+$ ,  $X(3872) \rightarrow \pi^+\pi^- J/\psi$ ,  $J/\psi \rightarrow \mu^+\mu^-$  signal Monte Carlo<sup>3</sup> to predict how the expected signal yield ( $N_{\text{sig}}$ ) changes with  $DLL_{\text{sig/bkg}}$ . We require the  $J/\psi \pi^+\pi^- K^+$  mass to be in the  $B^+$  peak region in this study. We normalized this yield to the  $X(3872)$  yield in the real data obtained with the  $DLL_{\text{sig/bkg}} < 0$  requirement, and determined by a fit to the  $M(\pi^+\pi^- J/\psi) - M(J/\psi)$  distribution. To predict relative change of the background ( $N_{\text{bkg}}$ ) under the  $X(3872)$  mass peak we count events in the real data in the  $900 < M(\pi^+\pi^- J/\psi) - M(J/\psi) < 1200$  MeV range ( $X(3872)$  far-sideband). We normalize this background yield to the one determined by the  $X(3872)$  fit, with the  $DLL_{\text{sig/bkg}} < 0$  requirement, and integrated in  $\pm 2.5$  sigma of the  $X(3872)$  peak. We maximize  $N_{\text{sig}}/\sqrt{N_{\text{sig}} + N_{\text{bkg}}}$  as illustrated in figure 3.6. We require  $DLL_{\text{sig/bkg}} < 1$  for further analysis. This requirement is  $94.1 \pm 2.6$  % efficient for the  $1^{++}$   $X(3872)$  signal and  $94.8 \pm 2.6$  % efficiency for the  $2^{-+}$   $\alpha = (0.68, 0.32)$  signal model.

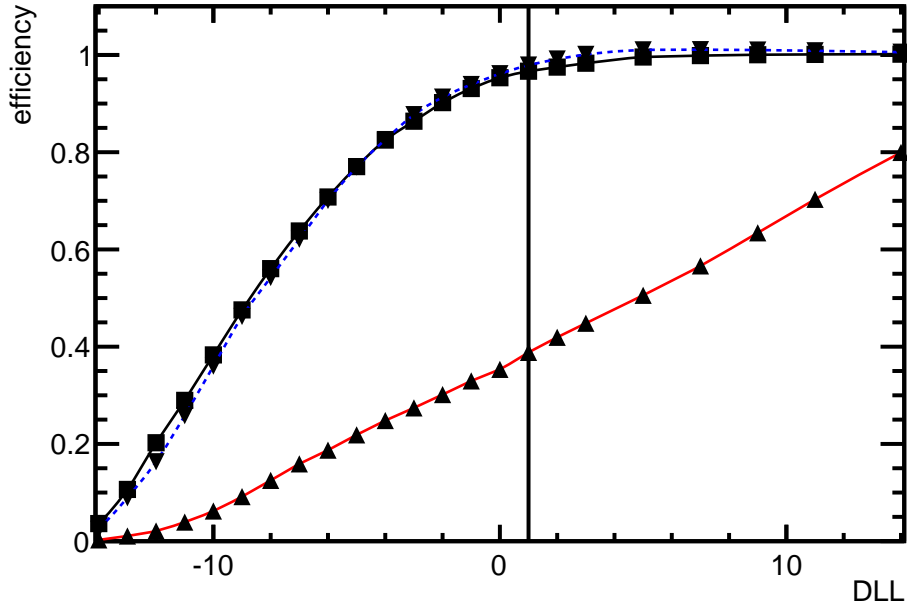


Figure 3.5: Fraction of events after the preselection passing a cut on  $DLL_{\text{sig/bkg}} < X$ , where  $X$  is given on the horizontal axis, for  $B^+ \rightarrow \psi(2S)K^+$  Monte Carlo (solid black line with squares),  $B^+ \rightarrow \psi(2S)K^+$  signal in the data (blue dotted line) and background events in the data from the far-sideband (solid red line with triangles). The vertical line indicates the actual requirement used in our analysis.

<sup>3</sup>We assumed  $J^{PC} = 1^{++}$  for  $X(3872)$ .

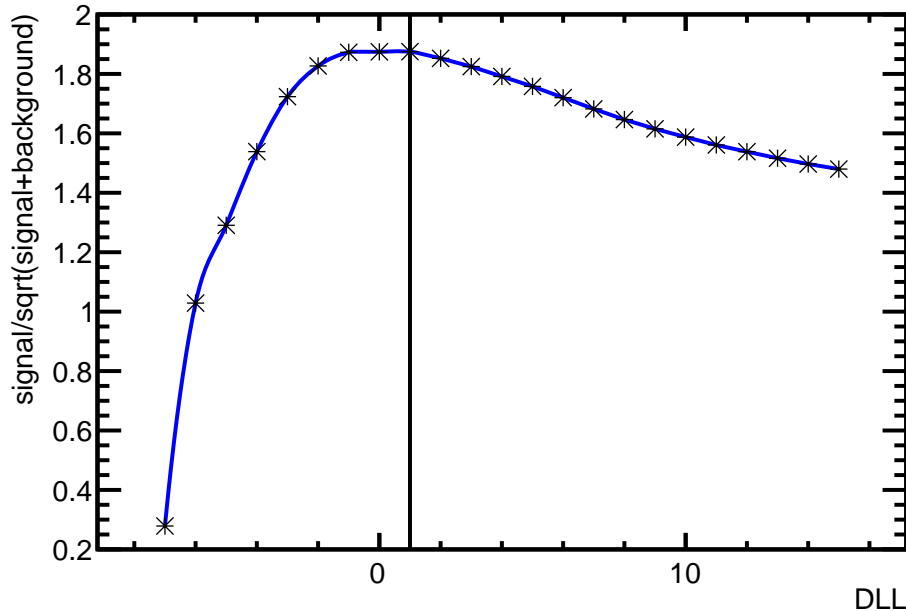


Figure 3.6: Expected dependence of  $\text{signal}/\sqrt{\text{signal} + \text{background}}$  for  $B^+ \rightarrow X(3872)K^+$  in a function of a cut on  $DLL_{\text{sig}/\text{bkg}} < X$ , where  $X$  is given on the horizontal axis. The vertical line indicates the actual requirement used in our analysis.

To maximize the signal statistics we use events accepted by any trigger lines. As a systematic check, we also analyze the data by selecting events with TOS (Trigger On Signal) on trigger lines which were not prescaled.

After the  $DLL_{\text{sig}/\text{bkg}} \text{ cut} < 1$ , the mass distribution for  $J/\psi \pi^+ \pi^- K^+$  is shown in figure 3.7 with the fit of symmetric Crystal Ball function for the signal (see definition below), and a linear function for the background, superimposed. The mass range selected for further analysis, 5261 – 5300 MeV, is indicated. According to the fit, in this mass range, there are  $38,241 \pm 235$   $B^+$  events and the background level is 10.9%.

Distribution of  $M(\pi^+ \pi^- J/\psi) - M(J/\psi)$  is shown in figure 3.8. Peaks from  $B^+ \rightarrow \psi(2S)K^+$  and  $B^+ \rightarrow X(3872)K^+$  are clearly visible. We use symmetric Crystal Ball shape to describe the signal shapes. It consists of a Gaussian shape turning into symmetrical power-law tails at  $\pm \alpha_M \sigma_M$  of the peak value. Extended maximum likelihood fits to the unbinned data for the  $\psi(2S)$  and  $X(3872)$  signals are shown in figures 3.9-3.10. The background is

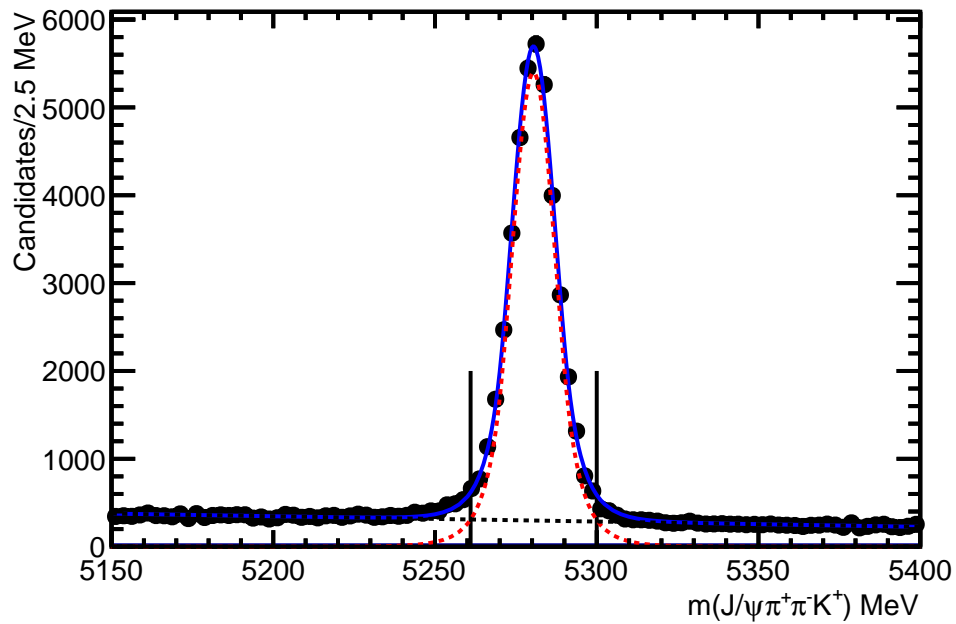


Figure 3.7: Observed  $M(J/\psi \pi^+ \pi^- K^+)$  mass distribution in the data. The fit results are superimposed (total-blue, signal-red, background-dashed). The vertical bars illustrate the range used for angular analysis.

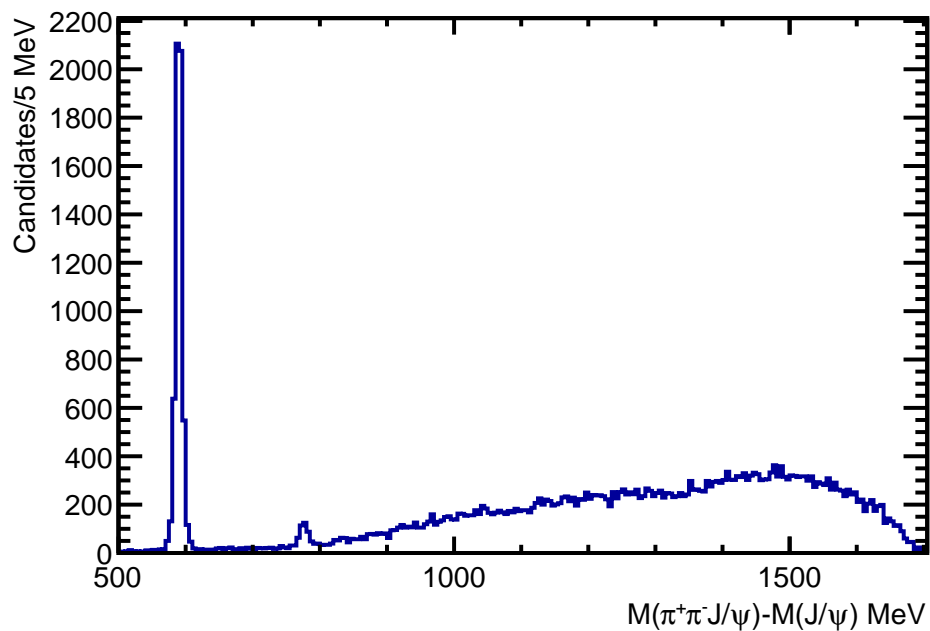


Figure 3.8: Distribution of  $M(\pi^+ \pi^- J/\psi) - M(J/\psi)$  for  $B^+ \rightarrow J/\psi K^+ \pi^+ \pi^-$  events. Peaks from  $B^+ \rightarrow \psi(2S)K^+$  and  $B^+ \rightarrow X(3872)K^+$  are clearly visible.



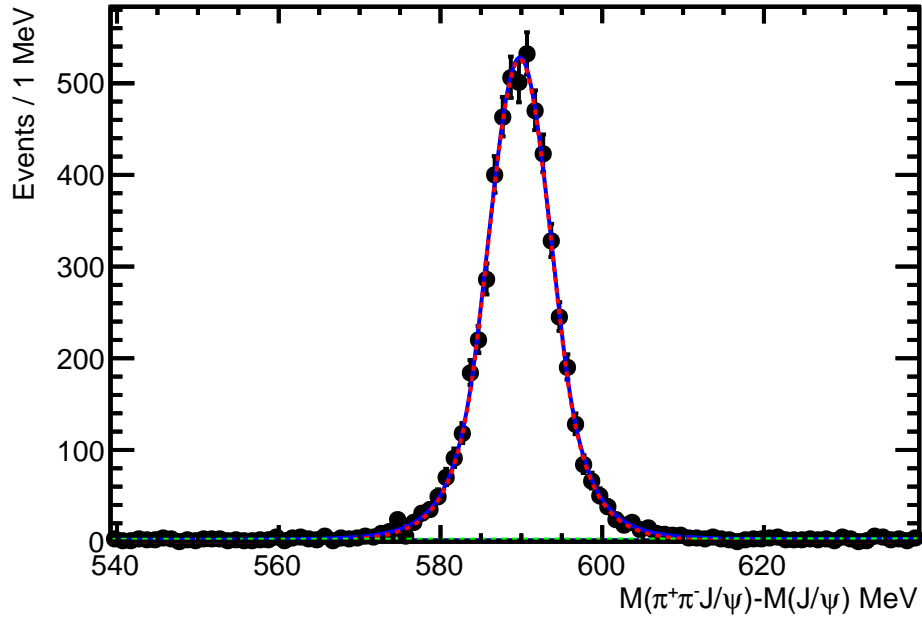


Figure 3.9: Distribution of  $M(\pi^+\pi^-J/\psi) - M(J/\psi)$  for  $B^+ \rightarrow J/\psi K^+ \pi^+ \pi^-$  events. The fit of the  $\psi(2S)$  signal is displayed. The blue, red and green blue lines represent the total fit, signal component and background component, respectively.

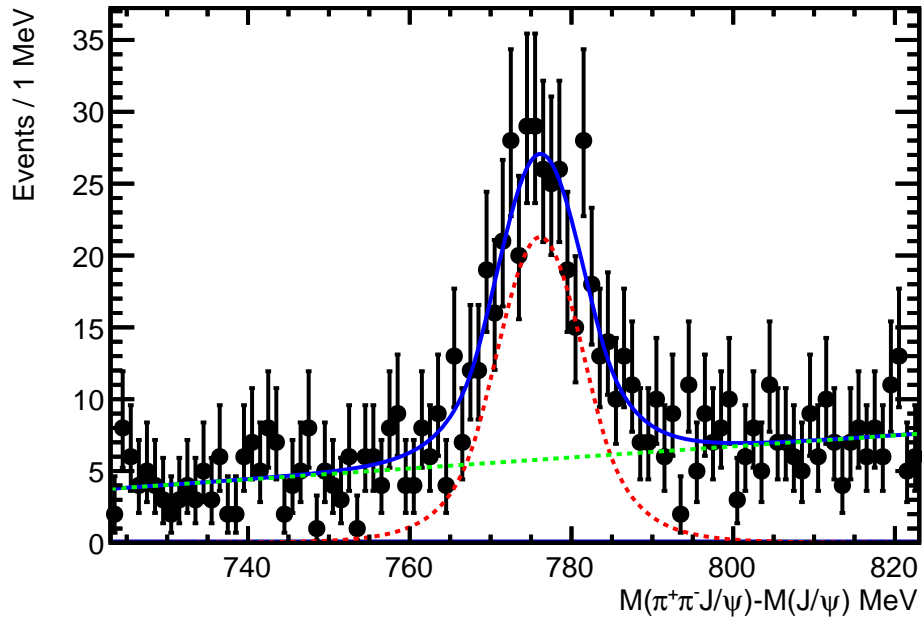


Figure 3.10: Distribution of  $M(\pi^+\pi^-J/\psi) - M(J/\psi)$  for  $B^+ \rightarrow J/\psi K^+ \pi^+ \pi^-$  events (2011 data). The fit of the  $X(3872)$  signal is displayed. The blue, red and green blue lines represent the total fit, signal component and background component, respectively.

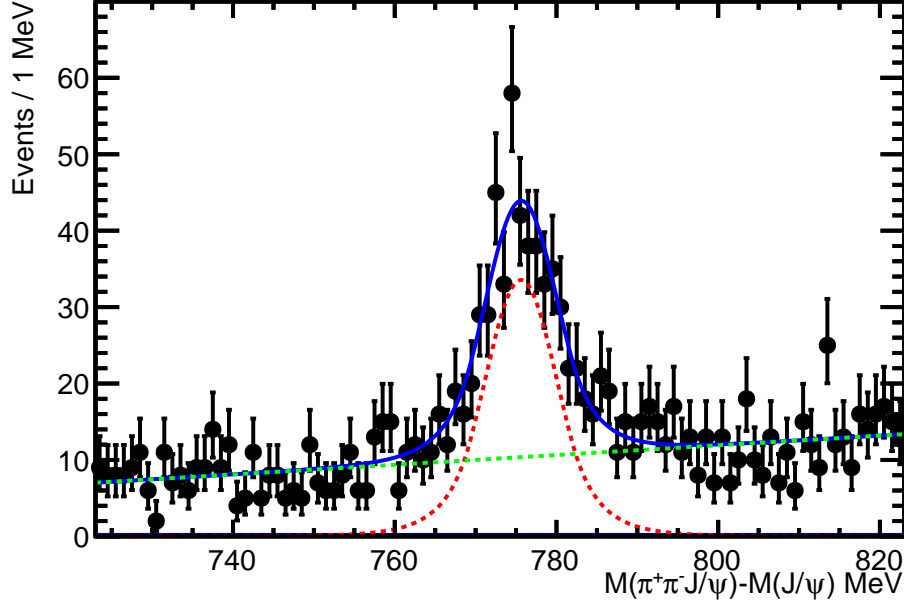


Figure 3.11: Distribution of  $M(\pi^+\pi^-J/\psi) - M(J/\psi)$  for  $B^+ \rightarrow J/\psi K^+\pi^+\pi^-$  events (2012 data). The fit of the  $X(3872)$  signal is displayed. The blue, red and green blue lines represent the total fit, signal component and background component, respectively.

assumed to be linear. We observe  $N_s = 5642 \pm 76$   $B^+ \rightarrow \psi(2S)K^+$  events, peaking at  $M(\pi^+\pi^-J/\psi) - M(J/\psi) = 589.81 \pm 0.06$  MeV with a resolution of  $\sigma_M = 3.99 \pm 0.05$  MeV and the tail parameters  $\alpha_M = 1.41 \pm 0.09$ ,  $n = 10.7 \pm 4.4$  ( $n$  is the power). There are  $N_b = 230 \pm 21$  background events in the fitted range (539.2 – 639.2 MeV), which corresponds to 0.8% background in the peak region (defined as  $\pm 2.5\sigma_M$ ). We observe  $N_s = 313 \pm 26$   $B^+ \rightarrow X(3872)K^+$  events, peaking at  $M(\pi^+\pi^-J/\psi) - M(J/\psi) = 776.1 \pm 0.5$  MeV with a resolution of  $\sigma_M = 5.5 \pm 0.5$  MeV, where the tail parameters are fixed to the values obtained in the  $\psi(2S)$  fit.<sup>4</sup> There are  $N_b = 568 \pm 31$  background events in the fitted range (723 – 823 MeV), which corresponds to 32% background in the peak region. The dominant source of background under the  $X(3872)$  peak is from  $B^+ \rightarrow J/\psi K^*$ ,  $K^* \rightarrow K^+\pi^+\pi^-$  decays,  $K_1(1270)$  in particular. Fraction of multiple entries per event in the fitted range is very small, 0.45%. The mass ranges used in these fits are also used in the analysis of angular distributions. The fits shown here are used to assign sWeights to events,  $w_i$ . The sWeight is positive in the signal

<sup>4</sup>The same tail parameters describe the signal  $X(3872)$  MC distribution very well as illustrated in figure 3.17.

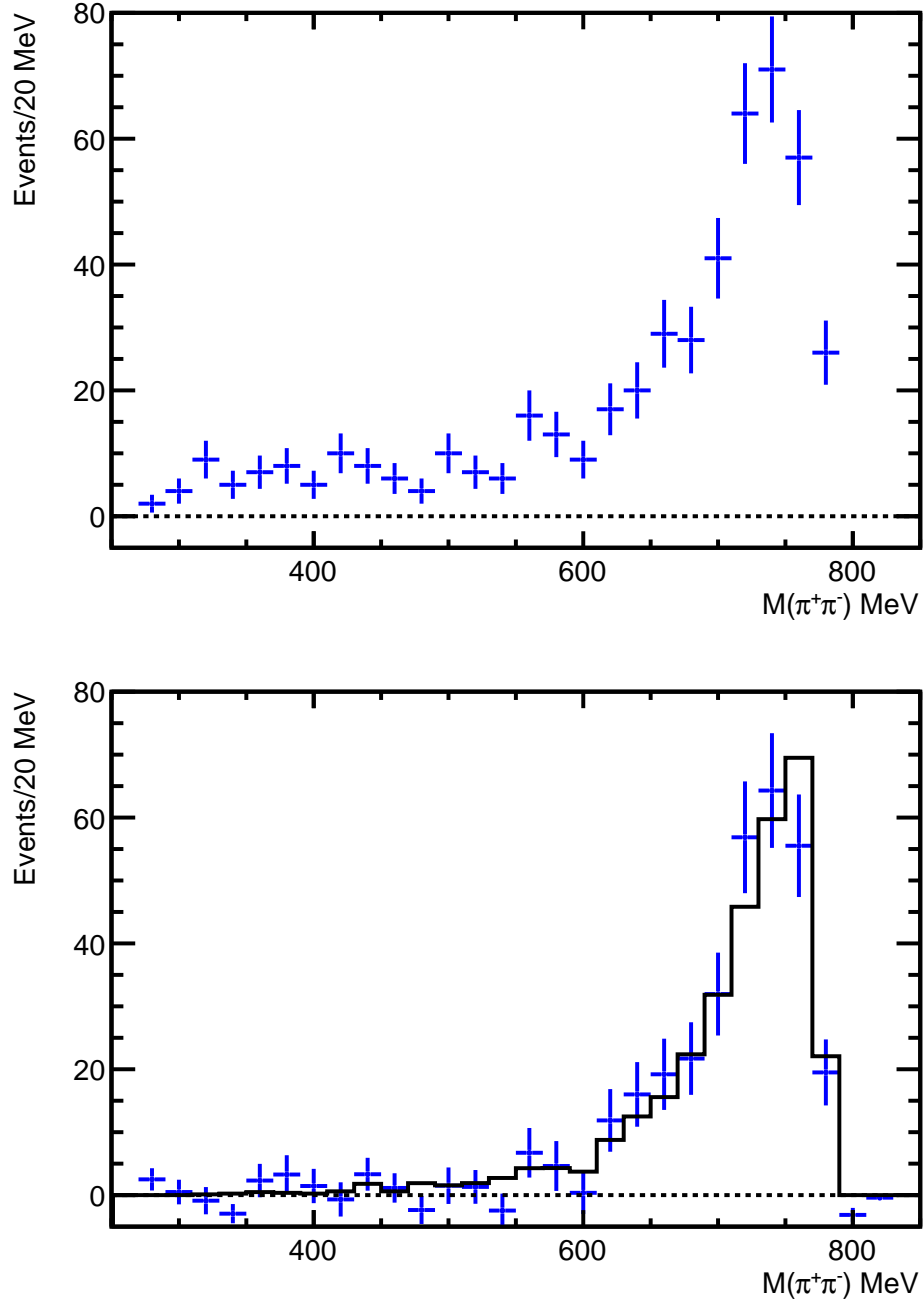


Figure 3.12: Distribution of  $M(\pi^+\pi^-)$  for  $B^+ \rightarrow X(3872)K^+$ ,  $X(3872) \rightarrow \pi^+\pi^- J/\psi$  events. The top plot shows the data in the  $X(3872)$  peak region ( $\pm 3\sigma_M$ ) with no sideband subtraction. The bottom plot shows the  $X(3872)$  region used in the angular analysis (723 – 823 MeV) after the sideband subtraction using the  $sWeights$  (blue points with error bars) compared to the simulations assuming  $X(3872) \rightarrow \rho(770)J/\psi$  decay (solid black histogram).

region and negative in the sidebands. When summed up sWeights give the number of signal events determined in the fit. We use these sWeights to subtract the background in the angular analysis. Therefore, no explicit background component term needs to be included in the fitted angular PDFs. The background level can be drastically reduced without dramatic loss of the  $X(3872)$  signal by eliminating low  $M(\pi^+\pi^-)$  region, as the  $X(3872) \rightarrow \pi^+\pi^- J/\psi$  decays are consistent with  $X(3872) \rightarrow \rho(770) J/\psi$  decays, with the  $\rho(770)$  mass being very close to the kinematic boundary. This is illustrated in figure 3.12. We explore cuts on  $M(\pi^+\pi^-)$  when studying stability of our results.

The distributions of rapidity, transverse momentum and momentum for the  $X(3872)$  signal selected in the data and in the simulations<sup>5</sup> are shown in figures 3.13-3.15. The backgrounds have been subtracted using the sWeights. The MC distributions have been normalized to the number of signal events observed in the data. The agreement between the data and the simulations is very good.

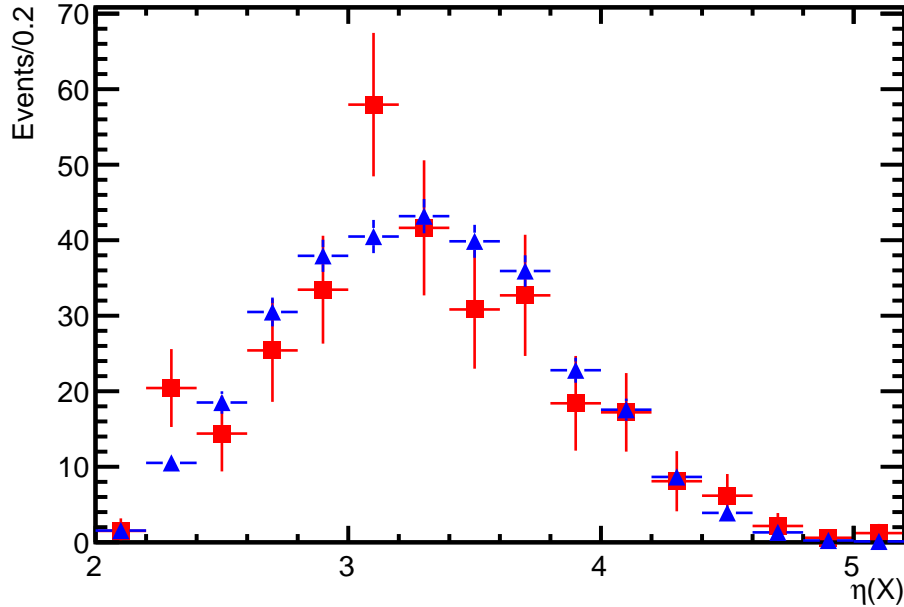


Figure 3.13: Distribution of rapidity for  $B^+ \rightarrow X(3872)K^+$  events in 2011 data (red points with squares) and in EVTGEN 1<sup>++</sup> MC (blue points with triangles).

<sup>5</sup>We have used EVTGEN 1<sup>++</sup> MC sample here (EventType 12145003).

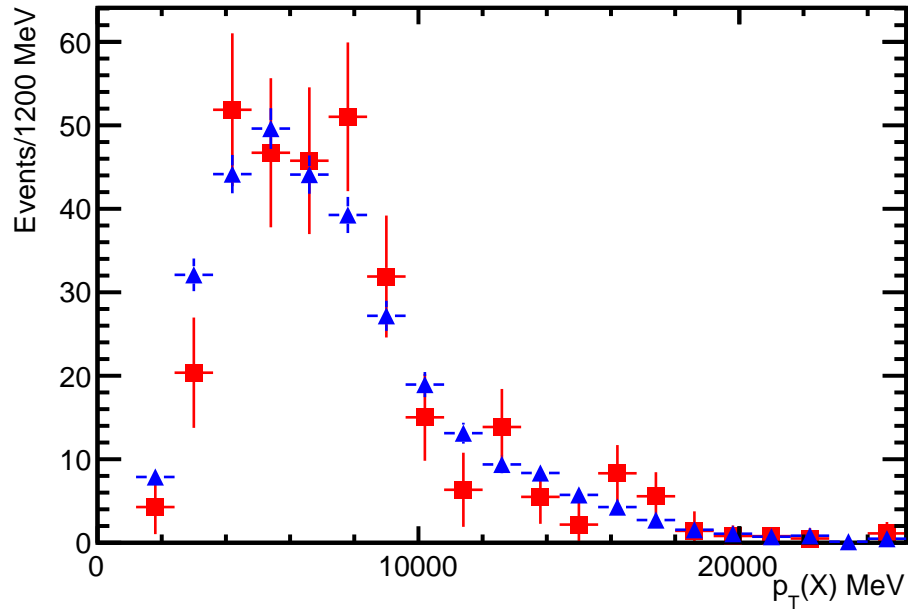


Figure 3.14: Distribution of  $X(3872)$   $p_T$  for  $B^+ \rightarrow X(3872)K^+$  events in 2011 data (red points with squares) and in EVTGEN 1<sup>++</sup> MC (blue points with triangles).

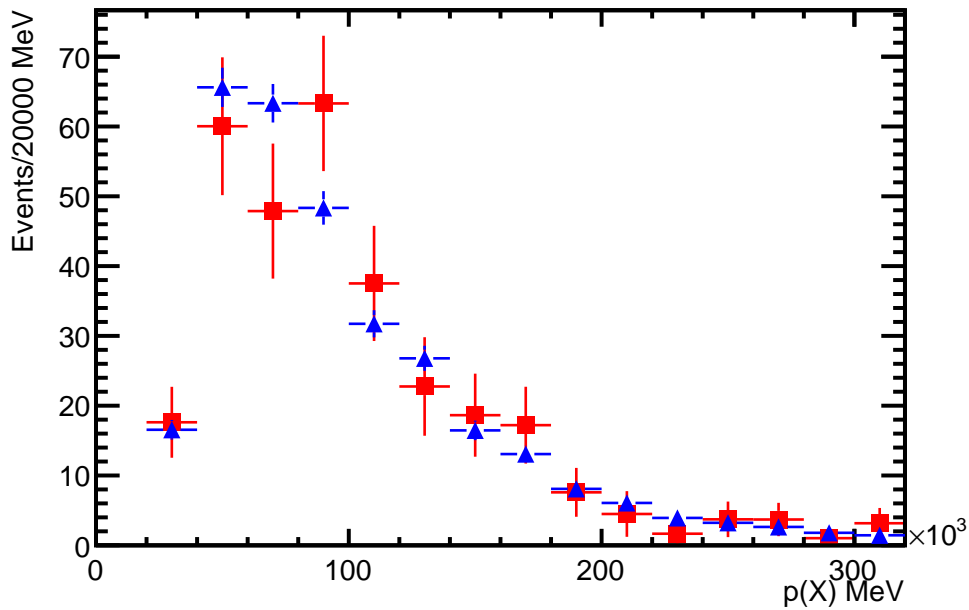


Figure 3.15: Distribution of  $X(3872)$  momentum for  $B^+ \rightarrow X(3872)K^+$  events in 2011 data (red points with squares) and in EVTGEN 1<sup>++</sup> MC (blue points with triangles).

We base our analysis on 2011 data. However, we use  $1.3 \text{ fb}^{-1}$  of 2012 data (Stripping19) for a cross-check. We have selected these data without proper retuning of the signal/background likelihood. The fit to  $M(\pi^+\pi^-J/\psi) - M(J/\psi)$  distribution in the  $X(3872)$  region is shown in figure 3.11. The number of signal events is  $403 \pm 30$ , with the total number of background events in the fitted range  $1020 \pm 39$  MeV. The peak value is  $775.6 \pm 0.4$  MeV and  $\sigma_M = 4.5 \pm 0.4$  MeV.

### 3.2 Angular Matrix Element

We use the helicity formalism [48] to predict the angular distributions in the  $B^+ \rightarrow X(3872)K^+$ ,  $X(3872) \rightarrow \pi^+\pi^-J/\psi$ ,  $J/\psi \rightarrow \mu^+\mu^-$  decay depending on  $J^{PC}$  of the  $X(3872)$ <sup>6</sup>. Application of this formalism to this particular sequential decay was discussed at length in the CDF thesis[49], which was the basis for the CDF publication on this topic[20], and in the LHCb theses[23, 24].

Using the helicity approach, the angular distribution in a two-body decay  $a \rightarrow 1 + 2$  can be expressed as:

$$\frac{dN}{d \cos \theta d\phi} = D_{\lambda_a, \lambda_1-\lambda_2}^{J_a}(\phi, \theta, -\phi) A_{\lambda_1, \lambda_2} = d_{\lambda_a, \lambda_1-\lambda_2}^{J_a}(\theta) e^{i\phi(\lambda_a - (\lambda_1 - \lambda_2))} A_{\lambda_1, \lambda_2} \quad (3.1)$$

where  $\lambda_a, \lambda_1, \lambda_2$  are helicities of the mother and daughters (i.e. projection of their spins onto the direction of their momenta),  $D_{m', m}^J$  and  $d_{m', m}^J$  are the large and small Wigner functions,  $A_{\lambda_1, \lambda_2}$  is the helicity coupling (i.e. amplitude),  $\theta, \phi$  are the polar and azimuthal angles of the daughter particle 1 in the rest frame of the mother particle, with the  $z$  axis along the boost direction from the grandmother particle rest frame to the mother particle rest frame (i.e. the quantization axis for the mother particle spin). The angle  $\theta$  is usually called ‘‘helicity

<sup>6</sup>The  $X(3872)$  decay is described by 3 helicity angles of the  $X(3872)$  ( $\theta_X$ ), the  $J/\psi$  ( $\theta_{J/\psi}$ ), the  $\pi\pi$  system ( $\theta_{\pi\pi}$ ), and 2 azimuthal angles between the decay planes of the  $X(3872)$  and its two daughters,  $\phi_X - \phi_{J/\psi}$  and  $\phi_X - \phi_{\pi\pi}$  (see figure 3.16).

angle of the particle  $a$ ", thus we denote it as  $\theta_a$ . For consistency of notation we also give the  $\phi$  angle the same subscript,  $\phi_a$ . The azimuthal angle  $\phi_a$  depends on arbitrary choice of the  $x$ -axis in the rest frame of  $a$ . The dependence on this angle becomes observable when dealing with a sequential decay e.g.  $2 \rightarrow 3 + 4$ , since the product of the two terms describing each decay depends on the difference between the two azimuthal angles in each decay,  $\phi_a - \phi_2$ , i.e. the angle between the decay planes of the mother and the daughter. In practice, only these angular differences need to be determined from the data.

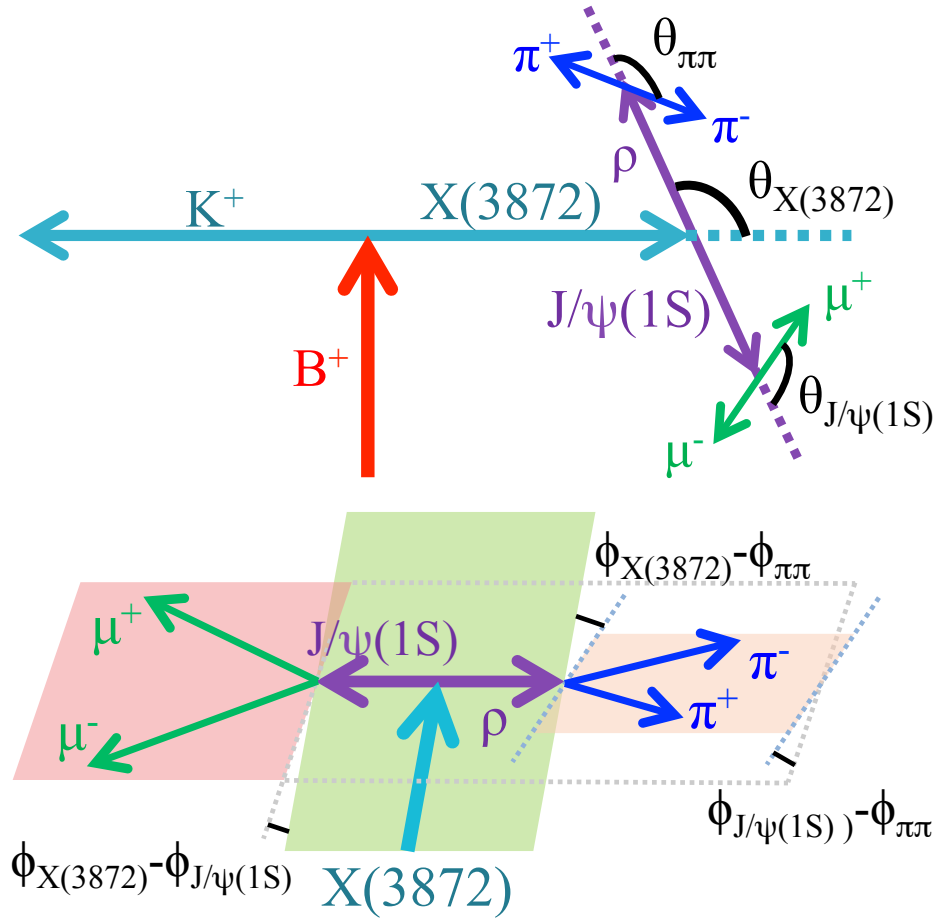


Figure 3.16: Definition of the helicity angles (top) and of the azimuthal angles between the decay planes (bottom) used in the description of angular distribution in  $B^+ \rightarrow X(3872)K^+$ ,  $X(3872) \rightarrow \rho J/\psi$ ,  $\rho \rightarrow \pi^+\pi^-$ ,  $J/\psi \rightarrow \mu^+\mu^-$ .

Since  $B^+$  has no spin,  $B^+ \rightarrow X(3872)K^+$  decay results in a flat angular distribution and requires  $\lambda_X - \lambda_{K^+} = 0$ . Since  $K^+$  is spinless too, the  $X(3872)$  is produced completely polarized in  $\lambda_X = 0$  state. There is only one helicity coupling in this decay.

The subsequent  $X(3872) \rightarrow \pi^+\pi^- J/\psi$  decay, is treated as two-body decay,  $J/\psi + \pi^+\pi^-$ . Thus it introduces a  $D_{0, \lambda_{J/\psi} - \lambda_{\pi\pi}}^{J_X}(\phi_X, \theta_X, -\phi_X) A_{\lambda_{J/\psi}, \lambda_{\pi\pi}}$  term. Since  $J_{J/\psi} = 1$ , there are 3 possible values of  $\lambda_{J/\psi}$ ,  $-1, 0$  and  $+1$ . Since this is a strong decay,  $C$ -parity is conserved, thus  $C_X = C_{J/\psi} C_{\pi\pi} = -C_{\pi\pi}$ . For the di-pion system,  $C_{\pi\pi} = (-1)^{L_{\pi\pi}}$  and  $L_{\pi\pi} = J_{\pi\pi}$ . The two important spin hypotheses tested in this note,  $J_X^{PC} = 1^{++}$  and  $2^{-+}$ , both have  $C_X = +1$  and require  $C_{\pi\pi} = -1$ , which can be provided by  $L_{\pi\pi} = 1, 3, \dots$ . Values of  $L_{\pi\pi} = 3$  and higher can be safely neglected, not only because of the angular momentum barrier, but also because the di-pion mass is limited to be smaller than 775 MeV. There are no known  $\pi^+\pi^-$  resonances with  $J_{\pi\pi} \geq 3$  in this mass region. Therefore, the  $\pi\pi$  system must be in  $1^{--}$  state ( $P_{\pi\pi} = (-1)^{L_{\pi\pi}}$ ). In fact, the di-pion mass distribution in the  $X(3872) \rightarrow \pi^+\pi^- J/\psi$  decay is known to be consistent with isospin-violating  $X(3872) \rightarrow \rho(770)J/\psi$  decay. Our data are also consistent with the dominance by  $\rho(770)$  as shown in figure 3.12. In our signal simulations we generate the di-pion mass according to  $X(3872) \rightarrow \rho(770)J/\psi$ . Since we do not include di-pion mass in fitting the spin hypotheses to the data, the di-pion mass distribution assumed in the MC matters only via a  $2^{nd}$  order effect; it enters via the simulation of angular efficiency. For  $J_{\pi\pi} = 1$ , there are 3 possible values of  $\lambda_{\pi\pi}$  ( $-1, 0, +1$ ). Therefore, there are  $3 \times 3 = 9$  helicity couplings contributing to this decay.

The subsequent  $\rho \rightarrow \pi^+\pi^-$  has only one helicity coupling since the pions have no spin. This decay contributes  $D_{\lambda_{\pi\pi}, 0}^1(\phi_{\pi\pi}, \theta_{\pi\pi}, -\phi_{\pi\pi})$  term.

The decay  $J/\psi \rightarrow \mu^+\mu^-$  introduces  $D_{\lambda_{J/\psi}, \Delta\lambda_\mu}^1(\phi_{J/\psi}, \theta_{J/\psi}, -\phi_{J/\psi})$  term, where  $\Delta\lambda_\mu = \lambda_{\mu^+} - \lambda_{\mu^-}$  can assume only  $-1, 1$  values, since this decay is mediated by a photon. Furthermore, the  $C$ -parity conservation makes the two helicity couplings in this decay equal.

To put all terms together, we multiply them by each other and sum up over  $\lambda_{J/\psi}$ ,  $\lambda_{\pi\pi}$  (coherently) and  $\Delta\lambda_\mu$  (incoherently) and obtain dependence of the matrix elements squared



on the decay angles:

$$|M(\theta_X, \theta_{J/\psi}, \theta_{\pi\pi}, \phi_X - \phi_{J/\psi}, \phi_X - \phi_{\pi\pi})|^2 \propto \quad (3.2)$$

$$\sum_{\Delta\lambda_\mu=1,+1} \left| \sum_{\lambda_{J/\psi}=1,0,+1} \sum_{\lambda_{\pi\pi}=1,0,+1} D_{\lambda_{J/\psi}, \Delta\lambda_\mu}^1(\phi_{J/\psi}, \theta_{J/\psi}, -\phi_{J/\psi}) \times \quad (3.3)$$

$$D_{\lambda_{\pi\pi}, 0}^1(\phi_{\pi\pi}, \theta_{\pi\pi}, -\phi_{\pi\pi}) \times \quad (3.4)$$

$$D_{0, \lambda_{J/\psi} - \lambda_{\pi\pi}}^{J_X}(\phi_X, \theta_X, -\phi_X) A_{\lambda_{J/\psi}, \lambda_{\pi\pi}} \right|^2 \quad (3.5)$$

Number of independent helicity couplings,  $A_{\lambda_{J/\psi}, \lambda_{\pi\pi}}$ , can be reduced by about a factor of 2 by imposing  $P$ -parity conservation

$$A_{\lambda_{J/\psi}, \lambda_{\pi\pi}} = P_{J/\psi} P_{\pi\pi} P_X (-1)^{J_{J/\psi} + J_{\pi\pi} - J_X} A_{-\lambda_{J/\psi}, -\lambda_{\pi\pi}} \quad (3.6)$$

which for the  $1^{++}$  and  $2^{-+}$  hypotheses leads to antisymmetrical couplings

$$A_{\lambda_{J/\psi}, \lambda_{\pi\pi}} = -A_{-\lambda_{J/\psi}, -\lambda_{\pi\pi}} \quad (3.7)$$

Fortunately, their number can be further reduced by considering couplings in the  $LS$  instead of the helicity basis.  $\vec{L}$  stands for a relative angular momentum between  $J/\psi$  and  $\pi\pi$  systems, and  $S$  is their total spin,  $\vec{S} = \vec{J}_{J/\psi} + \vec{J}_{\pi\pi}$ . Of course, they have to add up to  $X(3872)$  spin  $\vec{J}_X = \vec{L} + \vec{S}$ . The spin quantization axis for  $\vec{L}$ ,  $\vec{S}$  and  $\vec{J}$  is fixed, unlike in the helicity approach in which the quantization axis changes between the mother and the daughters. The number of independent  $LS$  amplitudes equals the number of independent helicity amplitudes. Furthermore, since both approaches offer a complete angular momentum bases, there is a linear transformation from the  $LS$  amplitudes,  $B_{LS}$ , to the helicity amplitudes, with

Clebsch-Gordan coefficients completely defining the transformation:

$$A_{\lambda_{J/\psi}, \lambda_{\pi\pi}} = \sum_L \sum_S B_{LS} \times \left( \begin{array}{cc|c} J_{J/\psi} & J_{\pi\pi} & S \\ \lambda_{J/\psi} & -\lambda_{\pi\pi} & \lambda_{J/\psi} - \lambda_{\pi\pi} \end{array} \right) \times \left( \begin{array}{cc|c} L & S & J_X \\ 0 & \lambda_{J/\psi} - \lambda_{\pi\pi} & \lambda_{J/\psi} - \lambda_{\pi\pi} \end{array} \right) \quad (3.8)$$

Possible values of  $L$  are constrained by the  $P$ -parity conservation:

$$P_X = P_{J/\psi} P_{\pi\pi} (-1)^L = (-1)^L \quad (3.9)$$

Thus, for the  $1^{++}$  hypothesis,  $L = 0, 2, \dots$ , whereas for the  $2^{-+}$  hypothesis  $L = 1, 3, \dots$ . Since the energy released in the  $X(3872) \rightarrow \rho(770)J/\psi$  decay is very small, the orbital angular momentum barrier is expected to be very effective and the lowest value of  $L$  will dominate, especially that  $L$  can only advance in units of 2. For  $J_X = 1$  and  $L = 0$ ,  $S$  must be equal to 1. Therefore, for the  $1^{++}$  hypothesis there is only one  $B_{LS}$  amplitude,  $B_{01}$ . All 9  $A_{\lambda_{J/\psi}, \lambda_{\pi\pi}}$  amplitudes can be related to it via the Clebsch-Gordan coefficients given above. For  $J_X = 2$  and  $L = 1$ , possible values of  $S$  are 1 or 2, therefore, there are two  $B_{LS}$  amplitudes,  $B_{11}$  and  $B_{12}$ . Again 9  $A_{\lambda_{J/\psi}, \lambda_{\pi\pi}}$  amplitudes can be expressed as their combination. Overall normalization of the matrix element is arbitrary, since the normalization is adjusted to the data. Also one complex phase is unobservable. Thus, the matrix element for the  $1^{++}$  hypothesis has no free parameters. The matrix element for the  $2^{-+}$  hypothesis has two free parameters, which following the previous LHCb convention, we choose to be a complex ratio:

$$\alpha = \frac{B_{11}}{B_{11} + B_{12}} \quad (3.10)$$

Real and imaginary parts of  $\alpha$  are nuisance parameters for the  $2^{-+}$  hypothesis, determined by the fit to our data. They are completely arbitrary.<sup>7</sup> The helicity amplitudes,  $A_{\lambda_{J/\psi}, \lambda_{\pi\pi}}$ , for  $1^{++}$  and  $2^{-+}$  hypotheses are summarized in table 3.2.

<sup>7</sup>In the previous work in LHCb,  $\alpha$  was assumed to be real and in between 0 and 1. This assumption has no theoretical motivation, thus we abandon it. We also note, that the magnitude of  $\alpha$  can be larger than 1, thus  $\alpha$  can have any value in a complex plane.

Table 3.2: Values of  $A_{\lambda_{J/\psi}, \lambda_{\pi\pi}}$ .

$\lambda_{J/\psi}$	$\lambda_{\pi\pi}$	$A_{\lambda_{J/\psi}, \lambda_{\pi\pi}}^{1^{++}}$	$A_{\lambda_{J/\psi}, \lambda_{\pi\pi}}^{2^{-+}}$
-1	-1	$-1/\sqrt{2}$	$-\alpha/\sqrt{3}$
-1	0	$-1/\sqrt{2}$	$-\alpha/2 + (1 - \alpha)/\sqrt{12}$
-1	+1	0	$+(1 - \alpha)\sqrt{2/3}$
0	-1	$-1/\sqrt{2}$	$-\alpha/2 - (1 - \alpha)/\sqrt{12}$
0	0	0	0
0	+1	$+1/\sqrt{2}$	$+\alpha/2 + (1 - \alpha)/\sqrt{12}$
+1	-1	0	$-(1 - \alpha)\sqrt{2/3}$
+1	0	$+1/\sqrt{2}$	$+\alpha/2 - (1 - \alpha)/\sqrt{12}$
+1	+1	$+1/\sqrt{2}$	$+\alpha/\sqrt{3}$

### 3.3 Discrimination between the two $J^{PC}$ hypotheses

We discriminate between two different values of  $J^{PC}$  using the likelihood-ratio test [50]. We use sWeighted unbinned likelihoods (sFit) in full 5-dimensional angular phase-space. The details of this technique are explained below.

The matrix element squared, multiplied by the detection efficiency, provides angular Probability Density Function ( $\mathcal{P}$ ) which depends on the  $J^{PC}$  hypothesis and the nuisance parameter  $\alpha$  in case of the  $2^{-+}$  hypothesis:

$$\mathcal{P}(\theta_X, \theta_{J/\psi}, \theta_{\pi\pi}, \phi_X - \phi_{J/\psi}, \phi_X - \phi_{\pi\pi} | J_X, \alpha) = \quad (3.11)$$

$$\frac{|M(\theta_X, \theta_{J/\psi}, \theta_{\pi\pi}, \phi_X - \phi_{J/\psi}, \phi_X - \phi_{\pi\pi} | J_X, \alpha)|^2 \epsilon(\theta_X, \theta_{J/\psi}, \theta_{\pi\pi}, \phi_X - \phi_{J/\psi}, \phi_X - \phi_{\pi\pi})}{\mathcal{P}_{norm}(J_X, \alpha)} \quad (3.12)$$

where  $\mathcal{P}_{norm}(J_X, \alpha)$  is the normalization factor

$$\mathcal{P}_{norm}(J_X, \alpha) = \int d \cos \theta_X d \phi_X d \cos \theta_{J/\psi} d \phi_{J/\psi} d \cos \theta_{\pi\pi} d \phi_{\pi\pi} \quad (3.13)$$

$$|M(\theta_X, \theta_{J/\psi}, \theta_{\pi\pi}, \phi_X - \phi_{J/\psi}, \phi_X - \phi_{\pi\pi} | J_X, \alpha)|^2 \epsilon(\theta_X, \theta_{J/\psi}, \theta_{\pi\pi}, \phi_X - \phi_{J/\psi}, \phi_X - \phi_{\pi\pi}) \quad (3.14)$$

The most sensitive variable discriminating between the two spin hypotheses is the likelihood-ratio,  $L(1^{++})/L(2^{-+}, \alpha)_{max}$ , where the likelihood  $L$  is maximized with respect to  $\alpha$  for the  $2^{-+}$  hypothesis [50]. The likelihood itself is given by:

$$L(J_X, \alpha) = \prod_{i=1}^N \mathcal{P}(\theta_X^i, \theta_{J/\psi}^i, \theta_{\pi\pi}^i, \phi_X^i - \phi_{J/\psi}^i, \phi_X^i - \phi_{\pi\pi}^i | J_X, \alpha) \quad (3.15)$$

where the product is over  $N$  signal events in the data sample. It is more convenient to operate on  $-2 \log L$ , which we denote as  $\chi_L^2$ , since near its maximum the likelihood is expected to have the Gaussian distribution:

$$\chi_L^2(J_X, \alpha) \equiv -2 \log L(J_X, \alpha) = -2 \sum_{i=1}^N \log \mathcal{P}_i \quad (3.16)$$

where  $\mathcal{P}_i = \mathcal{P}(\theta_X^i, \theta_{J/\psi}^i, \theta_{\pi\pi}^i, \phi_X^i - \phi_{J/\psi}^i, \phi_X^i - \phi_{\pi\pi}^i | J_X, \alpha)$ . The data contains backgrounds under the  $X(3872)$  mass peak. To subtract the backgrounds, we fit the  $M(J/\psi \pi^+ \pi^-) - M(J/\psi)$  distribution in 723 – 823 MeV range with the  $X(3872)$  signal and with linear background as described in Sec. 3.1 (see figure 3.10). We then assign sWeight to each event,  $w_i$ , and use sWeighted log-likelihood [51]

$$\chi_L^2(J_X, \alpha) = -2 \sum_{i=1}^N w_i \log \mathcal{P}_i \quad (3.17)$$

Use of sWeights is the only exact background subtraction method in 5-dimensions. To take into account the effect of the background subtraction on the fit errors, we also multiply the

log-likelihood by the constant scale factor<sup>8</sup> [52],

$$s_w = \sum_i w_i / \sum_i w_i^2 \quad (3.18)$$

equal to 0.574 for our data:

$$\chi_L^2(J_X, \alpha) = -2s_w \sum_{i=1}^N w_i \log \mathcal{P}_i(J_X, \alpha) \quad (3.19)$$

We discriminate between the  $1^{++}$  and  $2^{-+}$  hypotheses using:

$$\Delta\chi_L^2 \equiv \chi_L^2(2^{-+}) - \chi_L^2(1^{++}) = -2s_w \sum_{i=1}^N w_i \log \left( \frac{\mathcal{P}_i(2^{-+}, \alpha)}{\mathcal{P}_i(1^{++})} \right) \quad (3.20)$$

where  $\alpha$  is the value which minimizes  $\chi_L^2(2^{-+}, \alpha)$  (i.e. maximizes  $2^{-+}$  likelihood). In the publication draft for this analysis, we denote this test statistic,  $t$ .

Inserting equations 3.12 and 3.19 to equation 3.20, proves that the value of efficiency for  $i^{\text{th}}$  event,  $\epsilon(\theta_X^i, \theta_{J/\psi}^i, \theta_{\pi\pi}^i, \phi_X^i - \phi_{J/\psi}^i, \phi_X^i - \phi_{\pi\pi}^i)$ , drops out. This cancellation relies on the efficiency not depending on  $J_X$  or  $\alpha$ . Independence on  $\alpha$  is satisfied exactly, since this parameter has to do with total spin of the  $\rho$  plus the  $J/\psi$ ,  $S$ , which does not effect mass distribution shapes, while all five angular degrees of freedom are included in the PDFs. There is a slight dependence of the  $M(\pi\pi)$  distribution on  $J_X$  via different value of orbital angular momentum between  $\rho$  and  $J/\psi$ ;  $L = 0$  for  $J_X = 1$ ,  $L = 1$  for  $J_X = 2$ . The dependence on  $L$  enters the

---

<sup>8</sup>This scale factor was studied previously in Ref. [52]. It can also be derived using simple arguments. The log-likelihood  $\sum_{i=1}^N \log \mathcal{P}_i$  for unweighted (i.e. background free) events can be rewritten as  $N_s \langle \log \mathcal{P} \rangle$ , where  $N_s$  is the number of signal events ( $N_s = N$  with no background) and  $\langle \log \mathcal{P} \rangle$  is the average value of the  $\mathcal{P}$  in the data sample,  $\langle \log \mathcal{P} \rangle = \sum_{i=1}^N \log \mathcal{P}_i / N$ . For weighted events,  $\langle \log \mathcal{P} \rangle = \sum_{i=1}^N w_i \log \mathcal{P}_i / \sum_{i=1}^N w_i$ . For sWeighted sample it would be wrong to replace  $N_s$  by  $\sum_{i=1}^N w_i$ , since this would neglect fluctuations in the background subtraction. Replacing  $N_s$  by  $N$ , which counts both signal and background events, would also be wrong since the background events cancel via sWeighting in the weighted average. The right choice is to replace  $N_s$  by the number of equivalent signal events,  $N_{equ}$ , which is the number of unweighted (here signal free) events which would produce the same relative statistical fluctuations as  $N$  weighted events. It can be shown that  $N_{equ} = (\sum_{i=1}^N w_i)^2 / \sum_{i=1}^N w_i^2$ . Multiplying this formula by the weighted average leads to equations 3.18-3.19.

decay amplitude via the Blatt-Weiskopf factor,  $B_L(p)$ , where  $p$  is the daughter momentum in the  $X(3872)$  rest frame, which in turn depends on  $M(\pi\pi)$ . The Blatt-Weiskopf factor contains a somewhat arbitrary parameter related to the size of the decaying particle. Even though CDF included  $M(\pi\pi)$  in their fits (it helped to discriminate the signal from the large background in their inclusive analysis), they explicitly neglected the small  $L$  dependence in order to leave the  $J_X$  determination to the angular degrees of freedom alone, and these avoid any model dependent factors in the  $M(\pi\pi)$  parameterization [20]. We follow a similar strategy and neglect this effect. Since the backgrounds in our exclusive analysis are much smaller, we do not include  $M(\pi\pi)$  in the amplitude parametrization. Since in this approach the efficiency ratio cancels in the likelihood ratio, it is not necessary to determine the efficiency on event-to-event basis. However, the efficiency still enters via the normalization integral in equation 3.14. This integral can be determined numerically with help of Monte-Carlo events generated with uniform angular distributions<sup>9</sup> and passed through the event reconstruction and selection:

$$\mathcal{P}_{norm}(J_X, \alpha) \propto \sum_{i=1}^{N_{MC}} w_i^{MC} |M(\theta_X^i, \theta_{J/\psi}^i, \theta_{\pi\pi}^i, \phi_X^i - \phi_{J/\psi}^i, \phi_X^i - \phi_{\pi\pi}^i | J_X, \alpha)|^2 \quad (3.21)$$

where,  $w_i^{MC}$  is the sWeight assigned to  $i^{th}$  MC events. We allowed for a background component in the fit to the signal MC distribution of  $M(\pi^+\pi^- J/\psi) - M(J/\psi)$ , and therefore for MC sWeights, in order to accommodate any self-subtraction of the signal in case of inadequate signal shape parameterization. Our nominal signal shape describes the MC data very well and the fitted background is extremely small (see figure 3.17). To make statistical errors from the MC fluctuations in the normalization integral negligible we have generated a large sample of  $B^+ \rightarrow X(3872)K^+$ ,  $X(3872) \rightarrow \rho(770)J/\psi$ ,  $\rho(770) \rightarrow \pi^+\pi^-$ ,  $J/\psi \rightarrow \mu^+\mu^-$  events, with each decay simulated according to the phase-space model. Number of MC events after all selection requirements is  $N_{MC} = 41,789$ . Use of phase-space MC in the normalization

---

<sup>9</sup>EventType 121455005. We have used 41,770 reconstructed signal events. This MC assumes intermediate  $\rho(770)$  resonance in  $X(3872)$  decay.

integral is the only exact efficiency correction method in 5-dimensions.

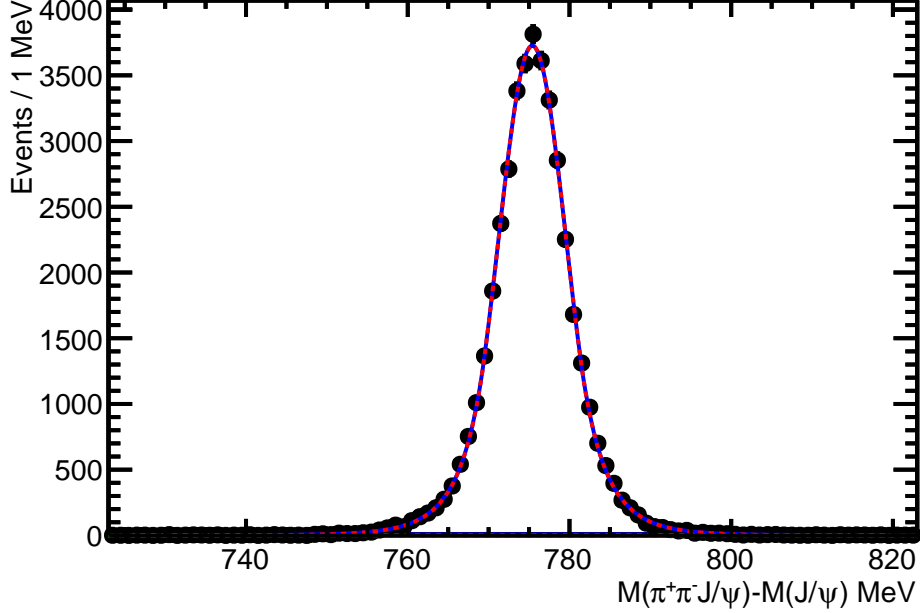


Figure 3.17: Distribution of  $M(\pi^+\pi^-J/\psi) - M(J/\psi)$  for  $B^+ \rightarrow X(3872)K^+$ ,  $X(3872) \rightarrow \rho(770)J/\psi$ ,  $\rho(770) \rightarrow \pi^+\pi^-$ ,  $J/\psi \rightarrow \mu^+\mu^-$  phase-space Monte Carlo events. The fit of the  $X(3872)$  signal (symmetric Crystal Ball shape, with  $\alpha_M = 1.41$ ,  $n = 10.7$ ) and linear background is superimposed. The fit gives  $41,789 \pm 205$  signal events peaking at  $775.40 \pm 0.02$  MeV with resolution  $\sigma_M = 4.17 \pm 0.02$  MeV and  $129 \pm 20$  background events. The latter is so small that the background polynomial is not visible.

For nested hypotheses, and  $N \rightarrow \infty$ , Wilks' theorem predicts  $\Delta\chi_L^2$  to have  $\chi^2$ -distribution with degrees of freedom equal to the difference in the number of free parameters in the hypotheses. The latter is 2 in case of discriminating between the  $1^{++}$  and  $2^{-+}$  hypotheses. The spin hypotheses cannot be considered nested. Nevertheless, we apply it as an approximate measure of significance with unknown accuracy and express the  $\Delta\chi_L^2$  probability in number of standard deviation in the normal distribution yielding the same probability

$$n_\sigma = \sqrt{2} \text{ROOT}::\text{TMath}::\text{ErfcInverse} \left( \text{ROOT}::\text{TMath}::\text{Prob}(\Delta\chi_L^2, 2) \right)^{10} \quad (3.22)$$

<sup>10</sup>ROOT::TMath::Prob is used for computation of the probability for a certain  $\chi^2$  and number of degrees of freedom; ROOT::TMath::ErfcInverse will return the inverse of the complementary error function.

We also calculate another approximate measure of significance

$$n_\sigma^* = \sqrt{|\Delta\chi_L^2 - 2|} \quad (3.23)$$

which gives slightly higher values than the Wilks' method. Both estimators approximately follow  $\sqrt{N}$  scaling with the signal statistics. While accuracy of both of these significance measures is unknown a priori, they serve as rough estimates and provide means of comparing the result on the real data with the simulations.

Unbiased estimate of the significance of the separation between the two spin hypotheses can be obtained by simulating many experiments with the same signal and background statistics as determined in the data. For example, if data prefers hypothesis  $J_A$  over  $J_B$  by  $\Delta\chi_L^2{}^{data}$  then a large number of experiments is generated according to  $\mathcal{P}(J_B)$ . Each MC experiment is analyzed the same way as the data, i.e. the MC data are fit to  $2^{-+}$  hypothesis by minimizing  $\chi_L^2(2^{-+}, \alpha)$  with respect to  $\alpha$ , then  $\Delta\chi_L^2{}^{J_B MC}$  is determined. A fraction of MC experiments with  $\Delta\chi_L^2{}^{J_B MC} > \Delta\chi_L^2{}^{data}$  gives a probability of rejecting  $J_B$  hypothesis by a statistical fluctuation,  $P_{J_B}$ . The latter can be expressed as number of standard deviations

$$n_\sigma^{toy MC} = \sqrt{2} \text{ROOT} :: \text{TMath} :: \text{ErfcInverse}(P_{J_B}) \quad (3.24)$$

In case this fraction is zero, only a lower limit on the significance can be set with this method. However, we observe that  $\Delta\chi_L^2{}^{J_B MC}$  distribution closely follows the normal distribution. Therefore, we fit the  $\Delta\chi_L^2{}^{J_B MC}$  distribution to a Gaussian, obtain the mean and RMS values, and calculate

$$n_\sigma^{toy MC*} = \frac{\Delta\chi_L^2{}^{data} - \text{mean}(\Delta\chi_L^2{}^{J_B MC})}{\text{RMS}(\Delta\chi_L^2{}^{J_B MC})} \quad (3.25)$$

Generating many MC experiments under the hypothesis preferred by the data allows determination of goodness-of-fit to the  $J_A$  hypothesis, via either fraction of MC experiments with  $\chi_L^2(J_A)^{J_A MC} < \chi_L^2(J_A)^{data}$  ( $CL_{\chi_L^2}(J_A)$ ) or fraction of MC experiments with



$\Delta\chi_L^2{}^{J_A MC} < \Delta\chi_L^2{}^{data}$  ( $CL_{\Delta\chi_L^2}(J_A)$ ). The first method is known to be often self-fulfilling and is not considered to be a sensitive goodness-of-fit test. Sensitivity of the second method can also be limited if the  $J_A$  hypothesis has free parameters.

In addition to the above goodness-of-fit test based on the sum of the log of the likelihood ratio over entire sample, we also look at the shape of the distribution of the log of the likelihood ratio calculated for individual events. We construct goodness-of-fit test by binning this distribution and calculating a  $\chi^2$  value between the observed and expected distribution for the  $J_A$  hypothesis.

We also evaluate goodness-of-fit by projecting the sWeighted data and the fits onto binned 1- and 2-dimensional histograms of the angles involved in the fit and calculating  $\chi^2$ -value between the data and the fit ( $\chi_{proj}^2$ ). We turn the  $\chi_{proj}^2$  values into confidence levels ( $CL_{proj}$ ) assuming that they follow the  $\chi^2$ -statistics with number of degrees of freedom equal to the number of bins of the histogram (minus 2 for the  $2^{-+}$  hypothesis). We also test discrimination between the two hypothesis by calculating the difference between the two hypotheses,  $\Delta\chi_{proj}^2 \equiv \chi_{proj}^2(2^{-+}) - \chi_{proj}^2(1^{++})$ , and turn it to a significance using the Wilks' theorem (equation 3.22),  $n_{\sigma}^{proj}$ , or using equation 3.23,  $n_{\sigma}^{proj*}$ . Obviously, these significances are much smaller than the ones based on 5-dimensional likelihood since the information from the other angles is lost. There is also loss of information due to binning, which is rather coarse to keep number of entries in each bin relatively large to justify use of  $\chi^2$ -statistics when calculating the confidence levels. We calculate these significances to gain insight which variables, or correlations, provide the best discrimination.

### 3.4 Verification of the likelihood analysis on EVTGEN MC

To verify our code implementing the matrix element formulae (equation 3.5 and table 3.2) and to verify the likelihood fitting and hypotheses discrimination described in the previous section,

we generate Monte Carlo events using EVTGEN to produce desired angular distributions. We use the HELAMP decay model to dial desired helicity structure for  $1^{++}$  (EventType 12145003) and  $2^{-+}$  hypotheses in the  $X(3872) \rightarrow \rho(770)J/\psi$  decay. The EVTGEN code is completely independent of ours, thus providing an excellent debugging opportunity. We generate three different values of  $\alpha$  for the  $2^{-+}$  hypothesis:

(1.0, 0.0) This is the setting which used to be a default in the LHCb version of  $2^{-+}$   $X(3872)$  MC. This is a model which is particularly easy to discriminate against the  $1^{++}$  hypothesis using 1D analysis of  $\cos \theta_X$  (EventType 12145004). [24].

(0.5, 0.0) This is a model which is essentially indistinguishable from the  $1^{++}$  hypothesis using 1D analysis of  $\cos \theta_X$  (EventType 12145006).

(0.68, 0.32) This is a model which is among hardest to distinguish from the  $1^{++}$  hypothesis using 5D likelihood analysis (EventType 12145007).

The events have been passed through full simulations and selection procedure. There are about 2,600 signal events in each MC data set passed to the 5D likelihood analysis. The results are given in table 3.3. From the sign of  $\chi_L^2$  we see that the right hypothesis is always preferred. With such large signal statistics, the wrong hypothesis would be rejected with significance of about  $n_\sigma^* \approx 40\sigma$ . It is interesting to scale this number to the signal statistics observed in the actual data (313 signal events). We assume that the significance should scale as  $\sqrt{N_s}$ . The analyzed MC data samples have no background. The real data have  $S/N$  ratio of about 2. To factor the dilution from the background subtraction in, we also apply additional scaling factor of  $\sqrt{s_w} = \sqrt{0.547} = 0.74$  (see equations 3.18-3.19). From this scaling to the 2011 data we expect to separate  $1^{++}$  and  $2^{-+}$  under all scenarios with a significance of about  $n_\sigma^* \approx 10\sigma$ . In case the data are  $2^{-+}$ , we can determine the helicity structure by obtaining a value of  $\alpha$ . We see that the fitted values of real and imaginary parts of  $\alpha$  for the  $2^{-+}$  MC fit to  $2^{-+}$  hypothesis (the last column in the last three rows in table 3.3) are reasonably close to

Table 3.3: The results from  $1^{++}/2^{-+}$  hypotheses testing on EVTGEN generated Monte Carlo data using 5D likelihood ratio test.

analyzed MC			$J^{PC}$	$\Delta\chi_L^2$	$n_\sigma^*$		Fit results for $\alpha$
$J^{PC}$	$\alpha$	$N_s$	pref- erred		act- ual	scal- ed	in the $2^{-+}$ fit $\alpha$
$1^{++}$	(—,—)	2626	$1^{++}$	+1633	$40.4\sigma$	$10.5\sigma$	$(0.650 \pm 0.011, 0.294 \pm 0.012)$
$2^{-+}$	(1.00, 0.00)	2611	$2^{-+}$	-1835	$42.9\sigma$	$11.2\sigma$	$(0.970 \pm 0.018, 0.143 \pm 0.036)$
$2^{-+}$	(0.50, 0.00)	2589	$2^{-+}$	-2159	$46.5\sigma$	$12.2\sigma$	$(0.498 \pm 0.007, 0.026 \pm 0.036)$
$2^{-+}$	(0.68, 0.32)	2563	$2^{-+}$	-1553	$39.4\sigma$	$10.5\sigma$	$(0.667 \pm 0.014, 0.333 \pm 0.013)$

the values used in the MC generation (the second column for the same rows).

The main purpose of this section is to verify that our likelihood-ratio test performs well on large statistics background-free samples. The scaling of the expected sensitivity to the actual 2011 data sample is only approximate. In section 3.5.2 we present more accurate simulations in which we generate many MC experiments with the exact signal statistics as observed in the data and we also simulate the background together with its subtraction. The expected sensitivity predicted that way is slightly lower than presented here. For example, for the  $1^{++}$  data sample the  $2^{-+}$  hypothesis is expected to be rejected on average at  $8.8\sigma$  as compared to  $10.5\sigma$  obtained here. The other simulation allows us to also predict experiment-to-experiment fluctuations in the  $2^{-+}$  rejection level (RMS=  $0.93\sigma$ ).

### 3.5 Results from the likelihood analysis on the real data.

#### 3.5.1 The results for 2011 and 2012 data compared to the expectations from the MC

The log-likelihood difference,  $\Delta\chi_L^2$ , obtained on the 2011 data is positive which means that the data favor  $1^{++}$  hypothesis. Using Wilks' theorem to translate it to a significance of  $2^{-+}$  rejection, we obtain  $n_\sigma = 9.7\sigma$ . An estimate based on  $\sqrt{\Delta\chi_L^2 - 2}$  approach is only slightly

higher,  $n_\sigma^* = 9.9\sigma$ . The results on 2012 data provide even stronger rejection  $n_\sigma = 11.1\sigma$ , as shown in table 3.4, which is expected for the higher statistics signal. Combining the 2011 and 2012 data together gives  $14.8\sigma$  rejection. Since the results involving 2012 data use MC11a for the efficiency correction (wrong  $\sqrt{s}$ ), they should be taken only as a cross-check. For further analysis, we concentrate on the 2011 data.

To compare to the expectations discussed in the previous section we also include in table 3.4 the results for  $1^{++}$  EVTGEN MC data scaled to 2011 signal statistics, as previously described. The result on the data is in good agreement with these projections. More accurate simulations which include explicit simulations of the background subtraction and don't require any scaling are discussed in the next subsection (Sec. 3.5.2).

The value of  $\alpha$  which maximizes  $2^{-+}$  likelihood on the 2011 data is  $(0.671 \pm 0.046, 0.280 \pm 0.046)$ . Contours of square root of the difference of the  $2^{-+}$  log-likelihood around this value are shown in figure 3.18. The angular distributions are symmetrical under change of sign of the imaginary part, which is reflected in the likelihood contours. The likelihood has only one maximum (modulo sign of the imaginary part). The 2012 data gives the consistent result for  $\alpha$ . The value of  $\alpha$  obtained from the data is also consistent with the value obtained by the fit to 25k background-free  $1^{++}$  EVTGEN MC signal events,  $(0.646 \pm 0.004, 0.289 \pm 0.004)$ . It is also worth noting that the value of  $\alpha$  obtained in our 5-dimensional analysis is similar to the value of  $\alpha$  which the Belle experiment found to describe all their 1-dimensional angular distributions,  $(0.64, 0.27)$  [21].

Our default selection accepts all triggers. In section 3.9 we replace use of Monte Carlo for the efficiency corrections in the fit and use reweighted  $B^+ \rightarrow \psi(2S)K^+$  signal in the data instead. In this check the  $\psi(2S)$  data probe all triggers in a proper way. The result ( $n_\sigma = 9.4\sigma$ ) is reasonably consistent with the one obtained using MC ( $n_\sigma = 9.7\sigma$ ). As an explicit check on systematics of trigger simulations, we also analyze a subsample of  $X(3872)$  events selected with specific trigger requirements, which preserves  $(89 \pm 2)\%$  signal events ( $278 \pm 25$ ). The

Table 3.4: The results from  $1^{++}/2^{-+}$  hypotheses testing on the real data and on the  $1^{++}$  EVTGEN MC using 5D likelihood ratio test. The results for the MC are scaled to the 2011 signal statistics, except for the errors on  $\alpha$  (scaling the errors gives  $\pm 0.04$  for the expected errors in the data).

analyzed		$J^{PC}$	$\Delta\chi_L^2$	$n_\sigma$	$n_\sigma^*$	Fit result for $\alpha$
data		pref-				in the $2^{-+}$ fit
$N_s$		erred				$\alpha$
2011	$313 \pm 26$	$1^{++}$	+99	$9.7\sigma$	$9.9\sigma$	$(0.671 \pm 0.046, 0.280 \pm 0.046)$
2012	$403 \pm 30$	$1^{++}$	+126	$11.0\sigma$	$11.1\sigma$	$(0.602 \pm 0.035, 0.256 \pm 0.035)$
2011+12	$718 \pm 40$	$1^{++}$	+223	$14.8\sigma$	$14.9\sigma$	$(0.631 \pm 0.028, 0.268 \pm 0.029)$
$1^{++}$ MC	$2626 \rightarrow 313$	$1^{++}$	+112		$10.5\sigma$	$(0.650 \pm 0.011, 0.294 \pm 0.012)$
2011 TOS	$278 \pm 25$	$1^{++}$	+110	$10.2\sigma$	$10.4\sigma$	$(0.671 \pm 0.049, 0.281 \pm 0.049)$
2011 Gauss.sig.	$296 \pm 24$	$1^{++}$	+100	$9.8\sigma$	$9.9\sigma$	$(0.667 \pm 0.046, 0.284 \pm 0.046)$
2011 flat bkg	$323 \pm 27$	$1^{++}$	+97	$9.6\sigma$	$9.8\sigma$	$(0.672 \pm 0.046, 0.281 \pm 0.046)$
2011 quad.bkg	$315 \pm 26$	$1^{++}$	+99	$9.7\sigma$	$9.9\sigma$	$(0.671 \pm 0.046, 0.280 \pm 0.046)$

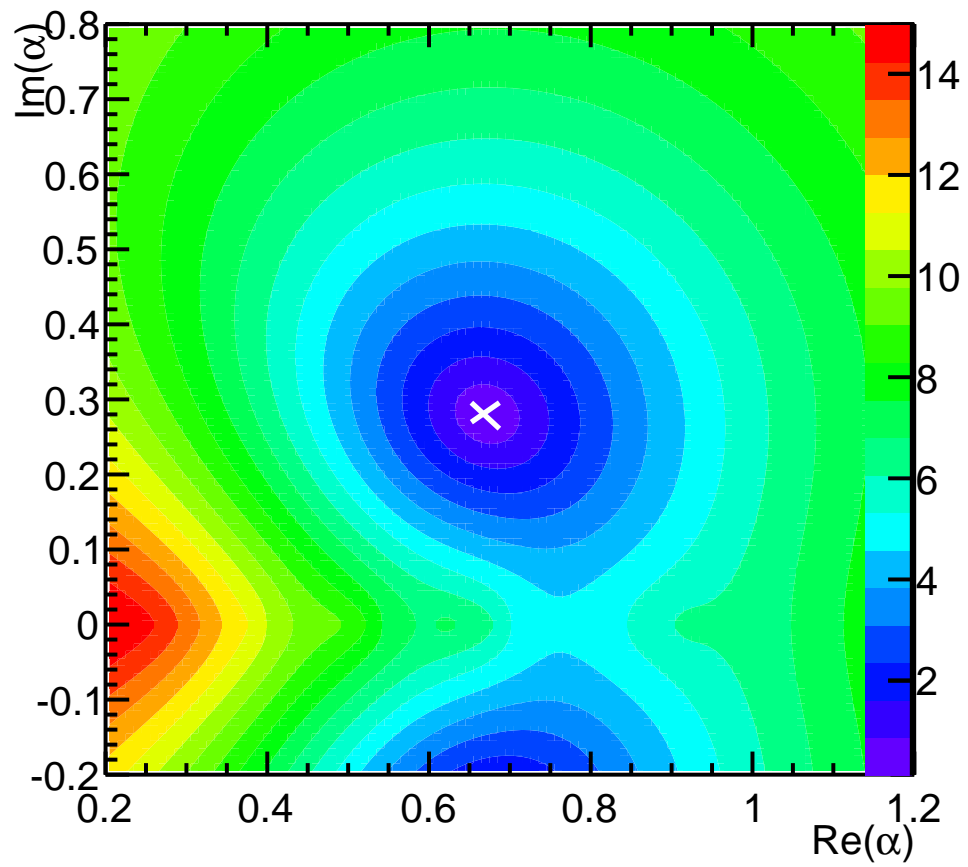


Figure 3.18: Contours of  $\sqrt{\chi_L^2(2^{-+}, \alpha) - \chi_L^2(2^{-+})_{\min}}$ .

rejection of  $2^{-+}$  hypothesis remains very large ( $n_\sigma = 10.2\sigma$ ), thus the trigger simulations don't bias the results against the  $2^{-+}$  hypothesis.

To investigate dependence of the results on the assumed signal shape in the fit to  $M(\pi^+\pi^-J/\psi) - M(J/\psi)$  distribution we have replaced the symmetric Crystal Ball shape with a simple Gaussian. The rejection of the  $2^{-+}$  hypothesis hardly moves ( $n_\sigma = 9.8\sigma$ ). The results are also very insensitive to the background shape assumed in this fit. The default fit assumes linear background. The results for flat and quadratic background are included in the table.

### 3.5.2 Significance of $2^{-+}$ rejection and goodness-of-fit for $1^{++}$ using toy experiments

Since a priori it is difficult to know how accurate the approximate estimates of significance are, it is important to calibrate values of log-likelihood-ratio with simulated experiments. To estimate significance of rejection of the  $2^{+-}$  hypothesis, we simulate 2,100,000 experiments with the same number of events as in the 2011 data, with angular distributions corresponding to  $2^{+-}$  and  $\alpha = (0.671, 0.280)$ . The latter are the values of  $\alpha$  obtained in the  $2^{-+}$  fit to the 2011 data. To properly simulate the efficiency effects for the signal events we re-weight the phase-space MC events using our formulae for the matrix element. The reweighting is done by picking randomly an event from the  $B^+ \rightarrow X(3872)K^+$  phase-space MC passed through the selection, and accepting it only if a random number drawn from the uniform distribution in the  $[0, 1]$  interval is smaller than the value of the matrix element squared (equation 3.5) evaluated for this event divided by the maximal value of the matrix element squared. The generation of the signal sample stops when 313 signal events are accepted. To properly simulate the background subtraction present in the analysis of the data, we also generate 568 background events for each experiment. The dominant background under the  $X(3872)$  mass peak is from  $B^+ \rightarrow J/\psi K^{*+}$  events,  $K_1(1270)^+$  in particular. We have generated  $B^+ \rightarrow J/\psi K_1(1270)^+$

Monte Carlo, with  $K_1(1270)^+$  decaying to  $K^+\pi^-\pi^+$  via various channel according to the PDG branching fractions.<sup>11</sup> As alternative background simulation we have also generated  $B^+ \rightarrow J/\psi K^+\pi^+\pi^-$  events according to three-body phase-space. More discussion related to the background model can be found in the next section (Sec. 3.5.3). In the analysis of a toy experiment, we include the procedure of determination of sWeights. Since the mass resolution is somewhat better in the signal MC than observed for the real data we replace values of  $M(\pi^+\pi^-J/\psi) - M(J/\psi)$  in the signal MC events by values regenerated from the signal shape obtained in the fit to the real data. Also values of  $M(\pi^+\pi^-J/\psi) - M(J/\psi)$  in the background MC events are replaced by the ones regenerated from the background shape obtained in the fit to the real data. We then fit  $M(\pi^+\pi^-J/\psi) - M(J/\psi)$  distribution for each toy experiment, assign event sWeights and determine the  $s_w$  factor. We then perform the likelihood analysis to obtain  $\alpha$  and  $\Delta\chi_L^2$  values. The  $\Delta\chi_L^2$  values are plotted for the  $2^{-+}$  toy experiments in figure 3.19, where they are also compared to the value of  $\Delta\chi_L^2$  obtained on the 2011 data,  $\Delta\chi_L^2{}^{data} = 99$ . No toy experiment comes even close to the value in the data, thus probability of obtaining  $\Delta\chi_L^2 > \Delta\chi_L^2{}^{data}$  from  $2^{-+}$  distribution is less than  $1/2100000$ , which corresponds to a significance of  $n_\sigma^{toy MC} > 5.0\sigma$ . It will not be feasible to verify significance to  $10\sigma$  level this way. However, we observe that a Gaussian fits the  $\Delta\chi^2$  distribution very well as illustrated in figure 3.19. Furthermore, the Gaussian has a more pronounced tail towards the data point than the simulated  $\Delta\chi^2$  distribution for the  $2^{-+}$  hypothesis. Therefore, we can use the Gaussian approximation to get a more accurate estimate of  $2^{-+}$  rejection while staying on conservative side. From the Gaussian method,  $\Delta\chi_L^2{}^{data}$  is  $n_\sigma^{toy MC*} = (8.434 \pm 0.006)\sigma$  away from the mean  $2^{-+}$  value. We have also generated 2.0M toy  $2^{-+}$  experiments with the value of  $\alpha$  obtained by the fit to the high statistics EVTGEN  $1^{++}$  MC sample (25,854 signal events), (0.646, 0.289) and we have obtained the consistent result,  $n_\sigma^{toy MC*} = (8.363 \pm 0.006)\sigma$ . This is only a slightly lower estimate of the significance than obtained using the approximate measures ( $9.7 - 9.9\sigma$ ). When using the phase-space  $B^+ \rightarrow J/\psi K^+\pi^+\pi^-$  MC instead of

<sup>11</sup>We have used EventType 12245000 to generate these decays which attempts to simulate various  $K_1(1270)^+$  decay modes with proper angular distributions.



$B^+ \rightarrow J/\psi K_1(1270)^+$  MC for the background model we obtain  $n_\sigma^{toy MC*} = (8.06 \pm 0.15)\sigma$ , which shows that  $n_\sigma^{toy MC*}$  value is rather insensitive to the background model.

To conclude various significance measures. We can reject  $2^{-+} J^{PC}$  hypothesis for  $X(3872)$  with significance of at least  $5\sigma$  without any assumptions. Using the Gaussian approximation the  $2^{-+}$  hypothesis is ruled out at  $8.4\sigma$ , which is still likely to be an underestimate given that the Gaussian has a more pronounced tail towards the data point than the simulated distribution of the log-likelihood-ratio. Since all angular distributions vary smoothly over the phase-space, free of any narrow peaks or singularities, it is unconceivable that there is a corner of angular-phase space that has not been yet probed by the 2.1M sample of the simulated  $2^{-+}$  experiments that could produce a tail in the log-likelihood-ratio with a different trend than the one observed with the present simulation. The  $8.4\sigma$   $2^{-+}$  rejection is further backed up by the approximate methods which give  $9.7 - 9.9\sigma$ .

Since we test  $2^{-+}$  hypothesis with respect to  $1^{++}$  it is also important to quantify consistency of our data with the  $1^{++}$  assignment. The value of  $\Delta\chi_L^2$  for the data is well within the range predicted by the  $1^{++}$  toy experiments - see figure 3.20.<sup>12</sup> A fraction of the simulated experiments with the  $\Delta\chi_L^2$  below the data value is  $CL_{\Delta\chi_L^2}(1^{++}) = 34\%$ . This is a highly non-trivial goodness-of-fit test which probes all five dimensions of the data, in an optimal way to detect any biases away from the right hypothesis towards the other hypothesis. Because the  $1^{++}$  hypothesis has no free parameters, this goodness-of-fit test is completely unbiased. The distribution of values of  $\chi_L^2(1^{++})$  for the  $1^{++}$  toy experiments is shown in figure 3.21. Again the value obtained on the real data is well within the expected range. The confidence level of  $1^{++}$  hypothesis is  $CL_{\chi_L^2(1^{++})}(1^{++}) = 51\%$  from this test. The goodness-of-fit evaluated on the likelihood for the preferred hypothesis (in our case on  $\chi_L^2(1^{++})$ ) is known to be often self-fulfilling, thus this test is not as meaningful as the one on the likelihood-ratio. The distribution shown in figure 3.20 can be also presented in a form of the expected  $2^{-+}$

---

<sup>12</sup>We used  $B^+ \rightarrow J/\psi K_1(1270)^+$  as background model here.

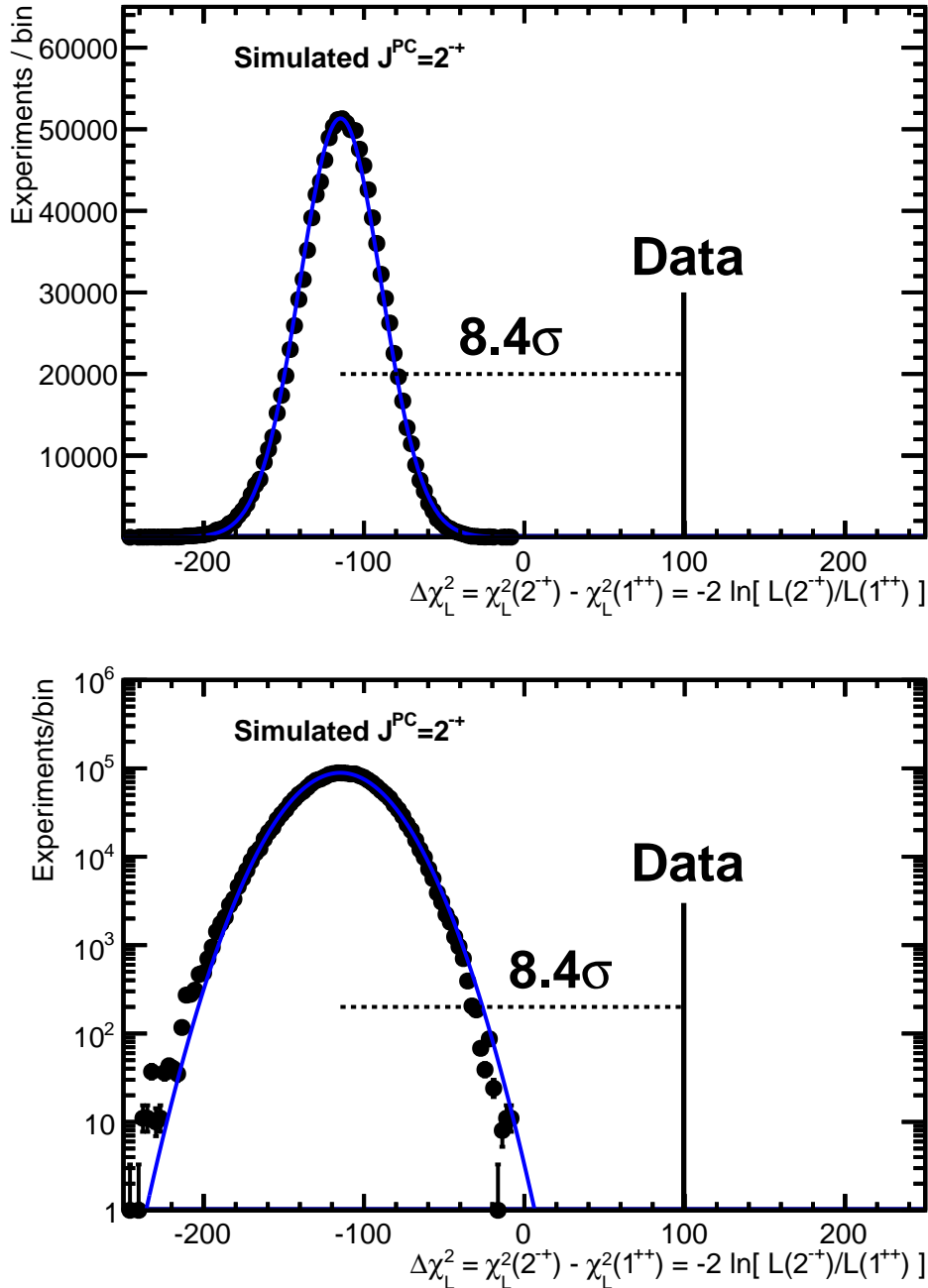


Figure 3.19: Distribution of  $\Delta\chi_L^2$  for the simulated experiments (points with error bars), with  $J^{PC} = 2^{-+}$  and  $\alpha$  determined by the angular fit to the data, and a value of  $\Delta\chi_L^2$  for the real data (vertical solid-bar). A fit of the Gaussian to the simulated distribution is superimposed (blue solid line). The bottom plot is the same as the top plot, except for the logarithmic vertical scale.

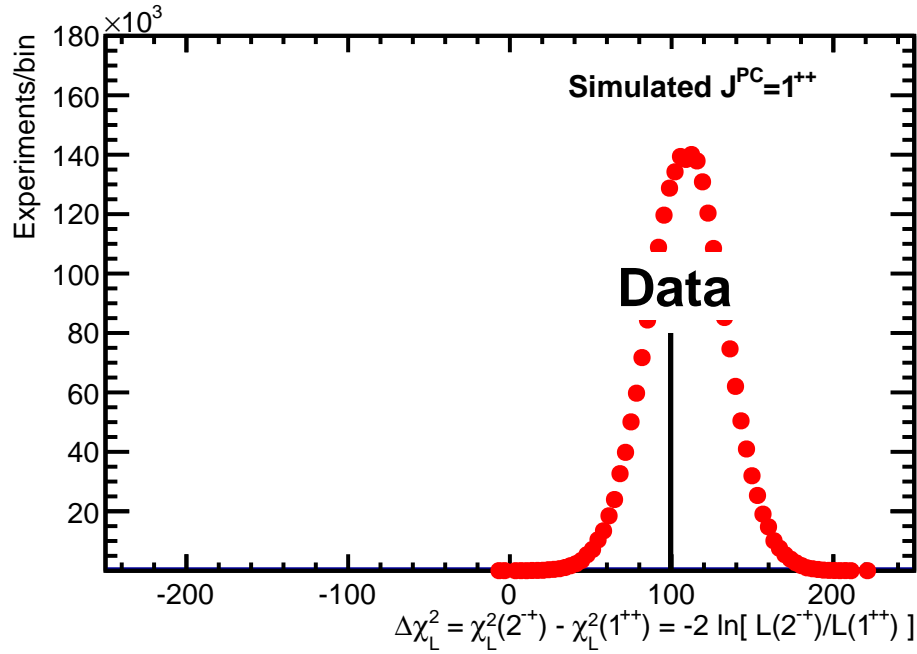


Figure 3.20: Distribution of  $\Delta\chi_L^2$  for the simulated experiments (red points with error bars), with  $J^{PC} = 1^{++}$ , and a value of  $\Delta\chi_L^2$  for the real data (vertical solid-bar).

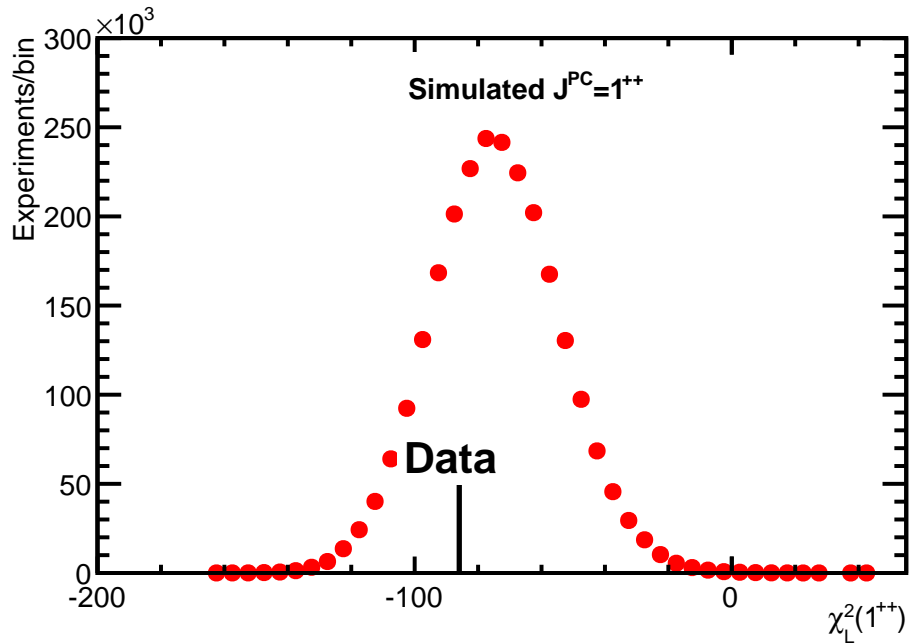


Figure 3.21: Distribution of  $\chi_L^2(1^{++})$  for the simulated experiments with  $J^{PC} = 1^{++}$  (red points with error bars) and a value of  $\chi_L^2(1^{++})$  for the real data (vertical solid-bar).

hypothesis rejection for the  $1^{++}$  toy experiments. This is shown in figure 3.22). The mean value of expected  $n_{\sigma}^{\text{toy } MC^*}$  is  $8.81\sigma$  with a RMS of  $0.93\sigma$ . The result for the data,  $8.4\sigma$ , is  $0.4\sigma$  from the average expected significance. Different goodness-of-fit measures are presented in the next two sections.

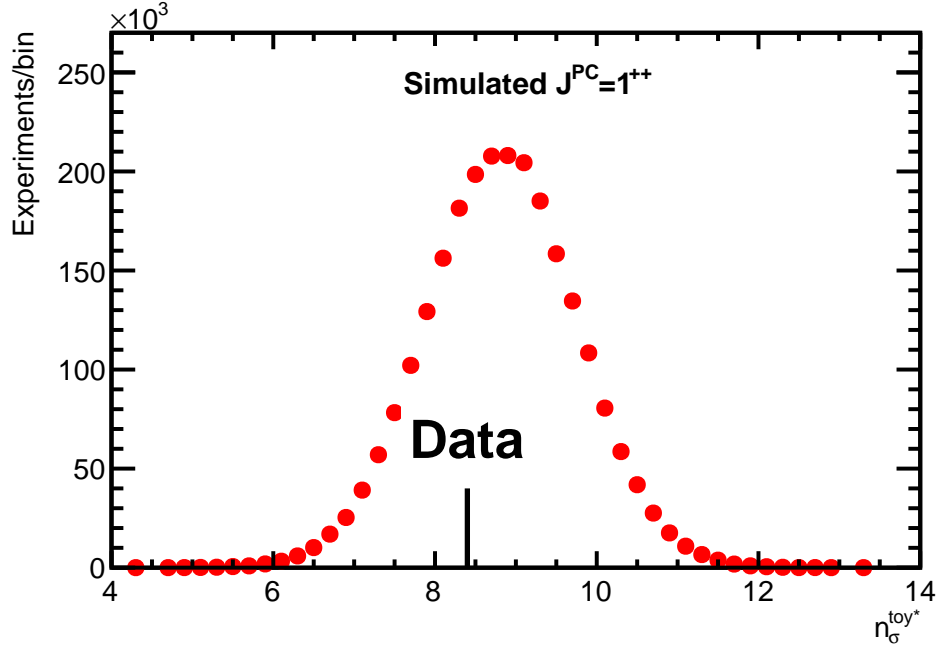


Figure 3.22: Distribution of  $n_{\sigma}^{\text{toy } MC^*}$  for the simulated experiments with  $J^{PC} = 1^{++}$  (points with error bars) and a value of  $n_{\sigma}^{\text{toy } MC^*}$  for the real data (vertical solid-bar).

### 3.5.3 Event-by-event distribution of the likelihood ratio

The likelihood ratio test which we use to discriminate between the spin hypotheses, relies on the log of the ratio of the likelihoods for each spin hypothesis summed over all events in the data sample (equations 3.17-3.20):

$$\Delta\chi_L^2 = -2s_w \sum_{i=1}^N w_i \log \left( \frac{\mathcal{P}_i(2^{-+}, \alpha)}{\mathcal{P}_i(1^{++})} \right) \quad (3.26)$$

where  $\alpha$  has the value which maximizes the  $2^{-+}$  likelihood on our data sample. Modulo constant factors, this test variable is an average of the logarithm of the likelihood ratio.

This test is proven on general grounds [50] to be the most sensitive way to distinguish between two hypotheses. In the previous section we showed that the value of  $\Delta\chi_L^2$  for the data is inconsistent with the  $2^{-+}$  hypothesis at  $8.4\sigma$  level. We also showed, that when used as goodness-of-fit measure, the value of  $\Delta\chi_L^2$  is consistent with the  $1^{++}$  hypothesis with a p-value of 34%. In this section we look at event-by-event distribution of the logarithm of the likelihood ratio,

$$LLR_i = -\log \frac{\mathcal{P}_i(2^{-+}, \alpha)}{\mathcal{P}_i(1^{++})}. \quad (3.27)$$

We fill histogram of  $LLR_i$  with sWeights ( $w_i$ ) to subtract the background. We then compare the shape of the distribution observed in the data with the shape of  $LLR$  predicted using  $1^{++}$  simulations. We calculate a  $\chi^2$  value between these two distributions and turn it to a p-value, which serves as additional goodness-of-fit measure, which checks that not only the average value, but also the distribution of the likelihood-ratio is consistent with the preferred hypothesis. The distributions of  $LLR$  with  $\alpha = (0.671, 0.280)$  for the data and for the  $1^{++}$  simulations are shown in figure 3.23. We also include distribution simulated for the  $2^{-+}$  hypothesis with this value of  $\alpha$ . The simulations here are performed the same way as in the previous section i.e. we generate many toy experiments with the signal and background statistics as observed in the data and go through the sWeighting procedure for each experiment. We generate enough experiments, to make statistical errors in the simulated distributions much smaller than in the data. The data match the  $1^{++}$  distribution very well and disagree with the  $2^{-+}$  distribution. p-value for the  $1^{++}$  hypothesis is 21%. p-value for the  $2^{-+}$  hypothesis is  $5.3 \cdot 10^{-12}\%$ , which corresponds to  $7.5\sigma$  disagreement. Since there is a loss of information due to the binning of the data, it is not surprising that this is slightly lower than in the unbinned test discussed in the previous section.

As an additional cross-check on the background subtraction in the log-likelihoods using sWeights, we compare the  $1^{++}$  and  $2^{-+}$  simulations which include the background with the background-free EVTGEN MC samples for  $1^{++}$  and  $2^{-+}$   $\alpha = (0.68, 0.32)$ . The value of  $\alpha$  used

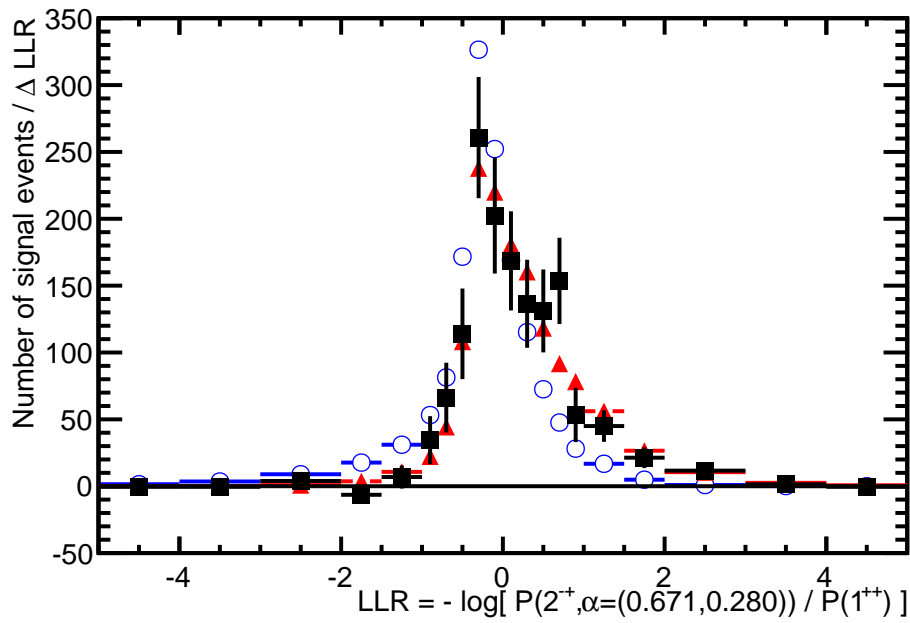


Figure 3.23: Distribution of LLR with  $\alpha = (0.671, 0.280)$  for the 2011 data (black points with squares) compared to the LLR distribution for the simulated experiments with  $J^{PC} = 1^{++}$  (red points with triangles) and with  $2^{-+}$ ,  $\alpha = (0.671, 0.280)$  (blue points with open circles) after the background subtraction using  $sWeights$ . The simulated distributions are normalized to the number of signal events observed in the data. Bin content and its error are divided by bin width because of unequal bin sizes.

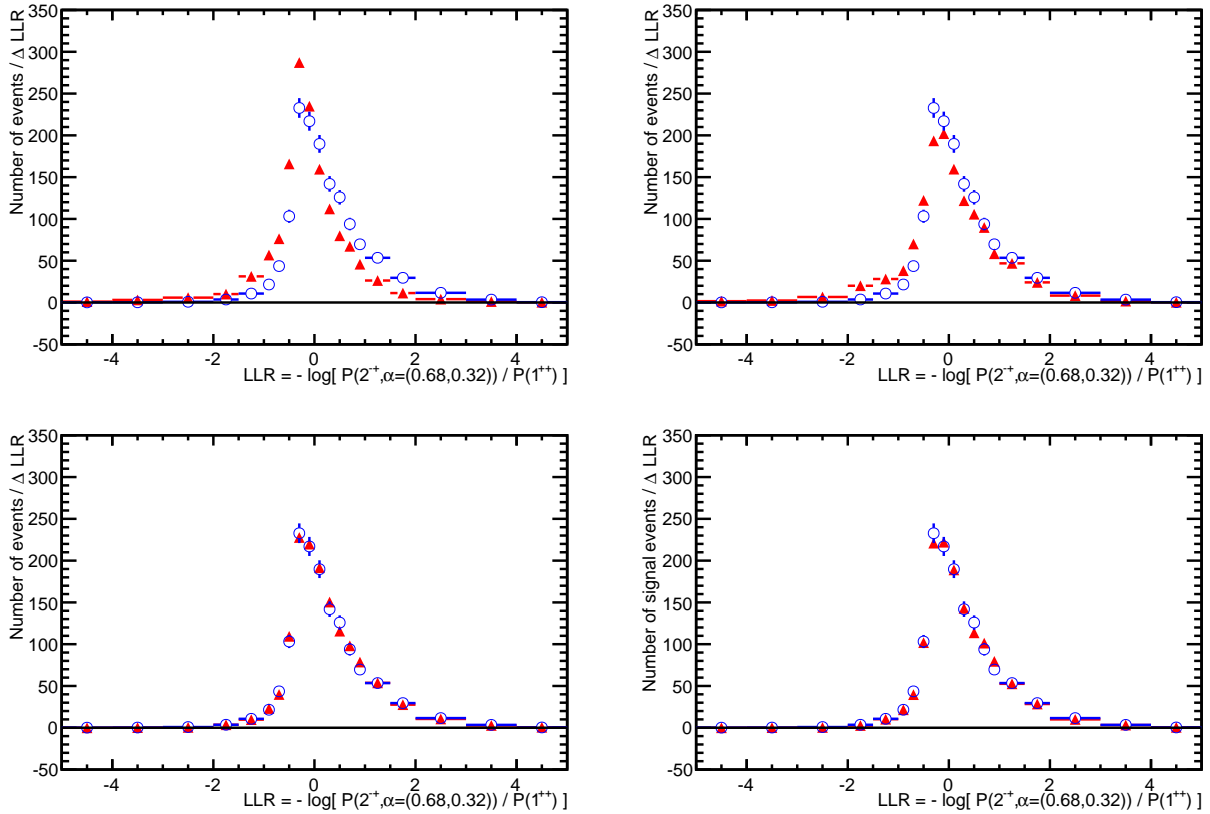


Figure 3.24: Distribution of LLR with  $\alpha = (0.68, 0.32)$  for the background-free EVTGEN simulation of  $1^{++}$  (blue points with open circles), compared to the LLR distributions for the  $1^{++}$  samples with the background (red points with triangles), before (top) and after (bottom) the background subtraction using  $sWeights$ . The plots on the left (right) show the simulations with the  $B^+ \rightarrow J/\psi K_1(1270)^+$  ( $B^+ \rightarrow J/\psi K^+ \pi^+ \pi^-$  phase-space) background model. The simulated distributions are normalized to the number of signal events observed in the data. Bin content and its error are divided by bin width because of unequal bin sizes.

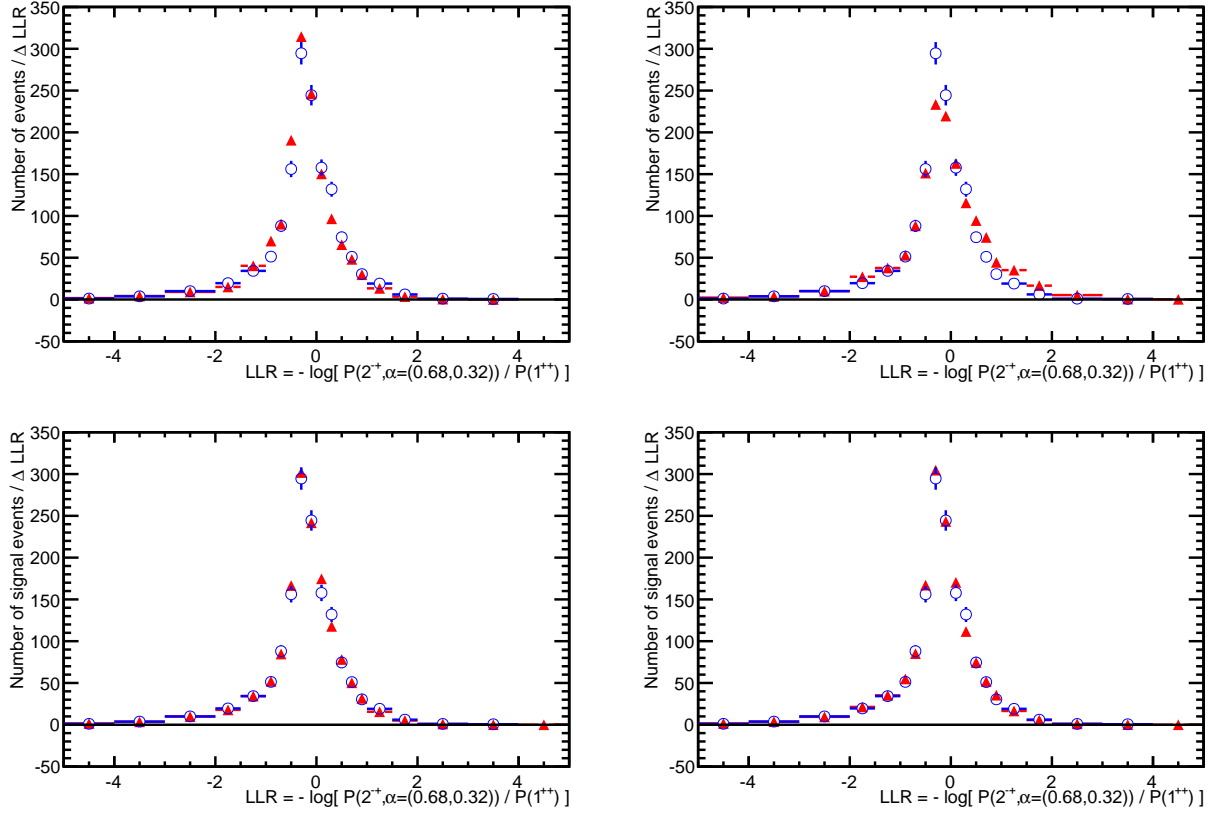


Figure 3.25: Distribution of LLR with  $\alpha = (0.68, 0.32)$  for the background-free EVTGEN simulation of  $2^{-+}$   $\alpha = (0.68, 0.32)$  (blue points with open circles), compared to the LLR distributions for the  $2^{-+}$   $\alpha = (0.68, 0.32)$  samples with the background (red points with triangles), before (top) and after (bottom) the background subtraction using *sWeights*. The plots on the left (right) show the simulations with the  $B^+ \rightarrow J/\psi K_1(1270)^+$  ( $B^+ \rightarrow J/\psi K^+ \pi^+ \pi^-$  phase-space) background model. The simulated distributions are normalized to the number of signal events observed in the data. Bin content and its error are divided by bin width because of unequal bin sizes.



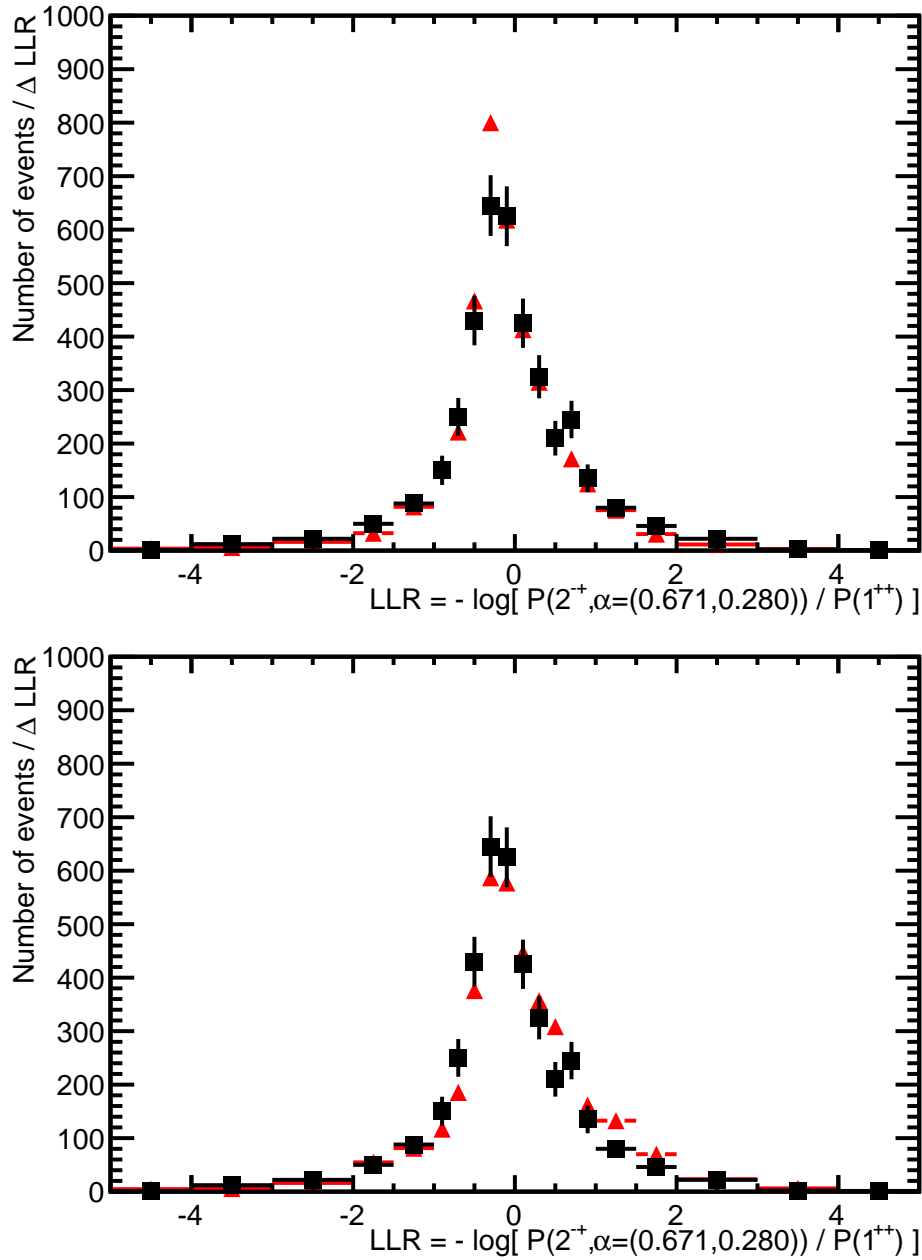


Figure 3.26: Distribution of LLR with  $\alpha = (0.671, 0.280)$  without the background subtraction for the 2011 data (black points with squares) compared to the LLR distributions for the simulated experiments with  $J^{PC} = 1^{++}$  (red points with triangles) and the two background models;  $B^+ \rightarrow J/\psi K_1(1270)^+$  (top) and  $B^+ \rightarrow J/\psi K^+ \pi^+ \pi^-$  phase-space (bottom). The simulated distributions are normalized to the number of events in the data. Bin content and its error are divided by bin width because of unequal bin sizes.

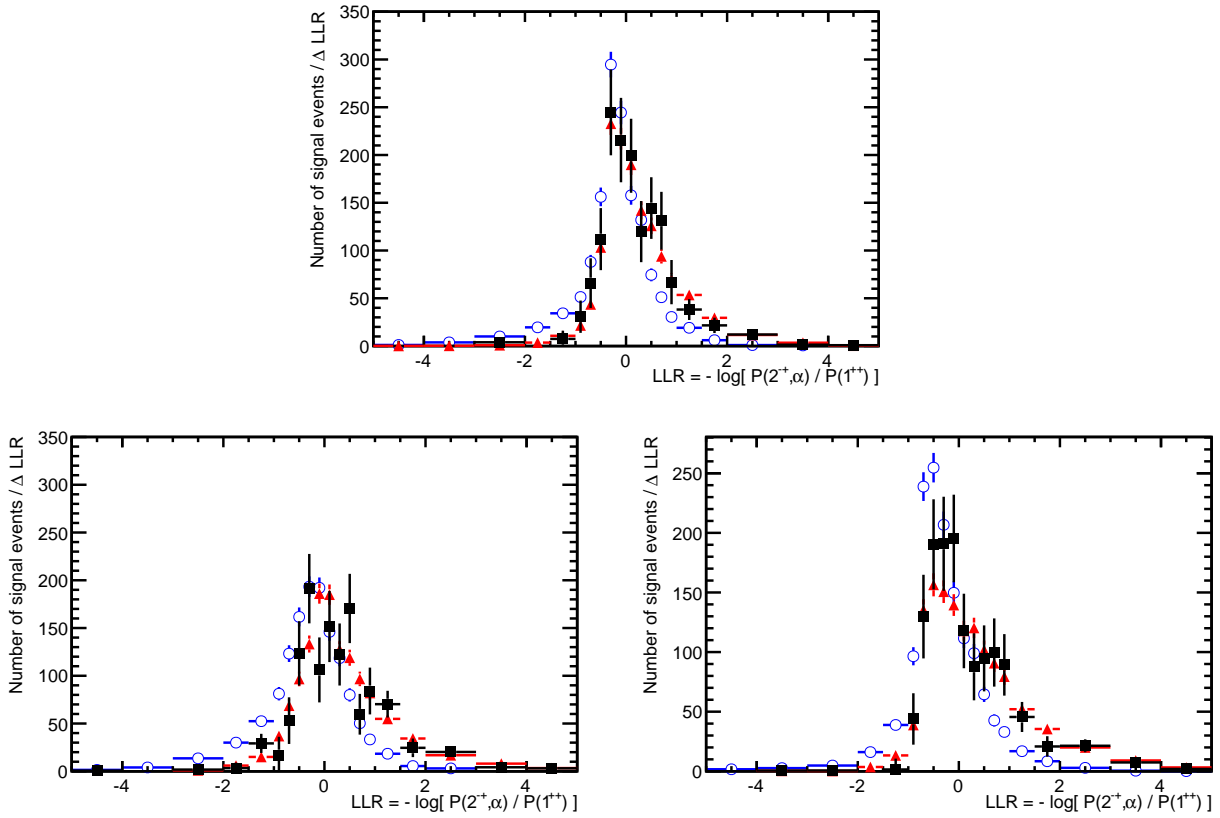


Figure 3.27: Distribution of LLR for the 2011 data (black points with squares) compared to the LLR distributions for the EVTGEN simulation of  $1^{++}$  (red points with triangles) and of  $2^{-+}$  (blue points with open circles), with  $\alpha = (0.68, 0.32)$  (top),  $(0.5, 0.0)$  (bottom left) and  $(1.0, 0.0)$  (bottom right). The value of  $\alpha$  used in the LLR definition changes and is set to the value of  $2^{-+}$  MC sample used in the comparison. The distributions are normalized to the number of signal events observed in the data. Bin content and its error are divided by bin width because of unequal bin sizes.

in equation 3.27 is also set to  $(0.68, 0.32)$ . These comparisons are shown in figures 3.24-3.25 without and with sWeighting the samples. We also show simulations with the two background models; the nominal  $B^+ \rightarrow J/\psi K_1(1270)^+$  and the alternative  $B^+ \rightarrow J/\psi K^+ \pi^+ \pi^-$  phase-space. Without the background subtraction (i.e. without weighting events), the distributions disagree as expected (top rows). The distribution of  $LLR$  in our nominal background model differs more from the  $1^{++}$  signal distribution (top left in figures 3.24) than from the  $2^{-+}$   $\alpha = (0.68, 0.32)$  signal distribution (top left in figures 3.25). The alternative background model makes the difference smaller for  $1^{++}$  and larger for  $2^{-+}$  (top right plots). After the background subtraction (bottom rows), the distributions agree within the errors. This is just a demonstration that sWeighting 5-dimensional log-likelihoods works as expected. Comparison of the  $1^{++}$  simulations which include the background to the 2011 data without the background subtraction reveals that the nominal background model,  $B^+ \rightarrow J/\psi K_1(1270)^+$ , does a fairly reasonable job describing the background in the data (top plot in figure 3.26). The alternative background model shows larger deviations (bottom plot in figure 3.26). Since the background subtraction works very well, the exact background model used in the simulations is of secondary concern.

As an illustration how separation between the  $1^{++}$  and  $2^{-+}$  hypotheses grows in  $LLR$  distribution when  $\alpha$  moves away from the one obtained by the  $2^{-+}$  fit to the data, we show comparison of the distributions of the 2011 data to the distributions of the EVTGEN samples for  $1^{++}$  and  $2^{-+}$  with  $\alpha = (0.68, 0.32)$ ,  $(0.5, 0.0)$  and  $(1.0, 0.0)$  (figure 3.27). For each of these comparisons, the value of  $\alpha$  used in the  $LLR$  definition (equation 3.27) is set to the value of  $\alpha$  used in the  $2^{-+}$  data generation. The data follow the  $1^{++}$  simulations.

### 3.5.4 Goodness-of-fits using 1D and 2D projections

In the previous two subsections, we have shown that the data are very consistent with the expectations for the  $1^{++}$  hypothesis on the value of the log of the likelihood ratio summed over

all events ( $\Delta\chi_L^2 \propto \sum_i w_i LLR_i$ ), as well as shape of  $LLR_i$  distribution for the signal. These goodness-of-fit measures probe complete 5-dimensional angular correlations. In this section, we present further goodness-of-fit measures based on binned projections of the data and of the 5-dimensional  $1^{++}$  fit onto the angular variables entering the matrix element. We also show projections of the  $2^{-+}$  fit and of the phase-space MC distribution to indicated approximate variation of the efficiency. One dimensional projections are shown in figures 3.28-3.31. We test goodness-of-fit calculating a  $\chi^2$  value between the data and the fit, and turning it to a confidence level using  $\chi^2$ -statistics (with number of degrees of freedom equal the number of bins minus 1)<sup>13</sup>. The results are summarized in table 3.5. In addition to goodness-of-fit of the  $1^{++}$  hypothesis ( $CL(1^{++})$ ), we also show confidence level for the  $2^{-+}$  hypothesis ( $CL(2^{++})$ ), and significance of rejection of the  $2^{-+}$  hypothesis based on the  $\chi^2$  differences between the two hypotheses and Wilks' theorem ( $n_\sigma^{proj}$ ). For comparison, also an average expected  $n_\sigma^{proj}$  is shown. The latter is obtained by fitting the large statistics EVTGEN  $1^{++}$  MC and scaling down to the actual data sample. In addition to 1-dimensional projections we also investigate 2-dimensional projections onto various pairs of the angular variables. Such 2-dimensional histogram for  $\cos\theta_X$  vs.  $\cos\theta_{\pi\pi}$  is shown in figure 3.33, where it is disassembled to five  $\cos\theta_X$  histograms for various  $\cos\theta_{\pi\pi}$  bins. Our signal statistics is too small to bin the data in 3-dimensions and expect the  $\chi^2$  statistics to be obeyed.

All confidence levels for the  $1^{++}$  hypothesis are good (table 3.5). Some confidence levels for the  $2^{-+}$  hypothesis are low, for the  $\cos\theta_X$  vs.  $\cos\theta_{\pi\pi}$  correlation in particular (0.02%). This 2D projection also gives the best rejection of the  $2^{-+}$  hypothesis when using the  $\chi^2$  difference between the hypotheses and the Wilks' theorem,  $n_\sigma^{proj} = 4.5\sigma$ . Not surprisingly, this is lower than the rejection based on the 5-dimensional log-likelihoods. The 2D  $\cos\theta_X$  vs.  $\cos\theta_{\pi\pi}$  distribution is illustrated in figure 3.33, which shows  $\cos\theta_X$  distributions (5 bins each) for 5 different slices in  $\cos\theta_{\pi\pi}$  (25 2-dimensional bins altogether). It is clear that the first and the last slice have much larger differences between the two spin hypotheses than

<sup>13</sup>We subtract 3 degrees of freedom for the  $2^{-+}$  hypothesis.

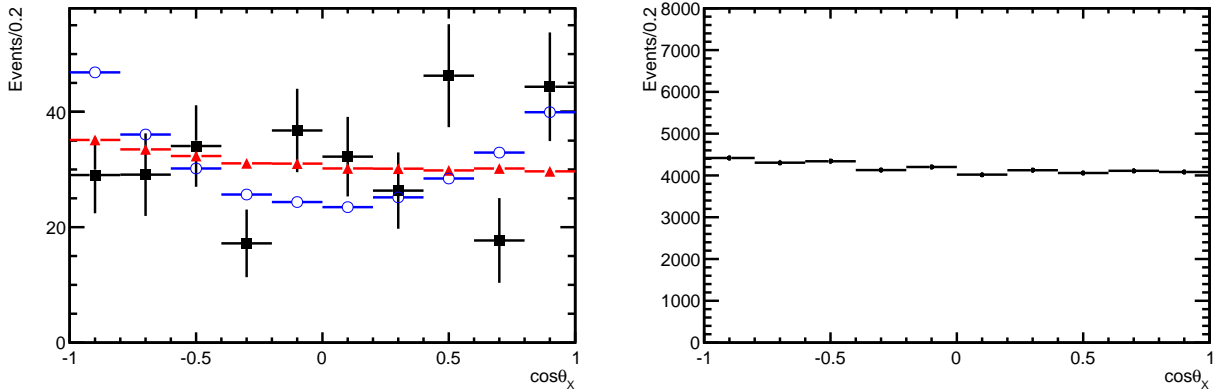


Figure 3.28: Distribution of  $\cos\theta_X$  for the data (black points with squares) and for the  $1^{++}$  (red points with triangles) and  $2^{-+}$  (blue points with open circles) fit projections on the left, and for the phase-space MC on the right.

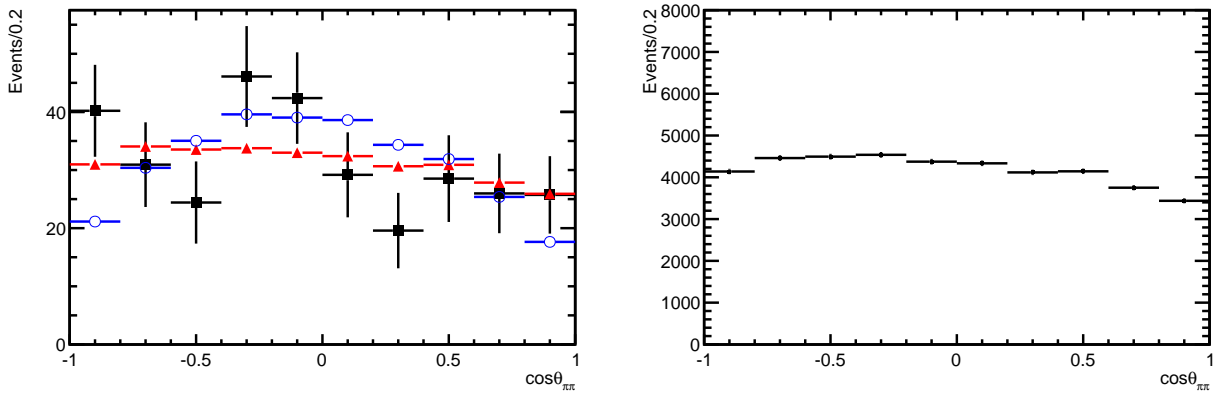


Figure 3.29: Distribution of  $\cos\theta_{\pi\pi}$  for the data (black points with squares) and for the  $1^{++}$  (red points with triangles) and  $2^{-+}$  (blue points with open circles) fit projections on the left, and for the phase-space MC on the right.

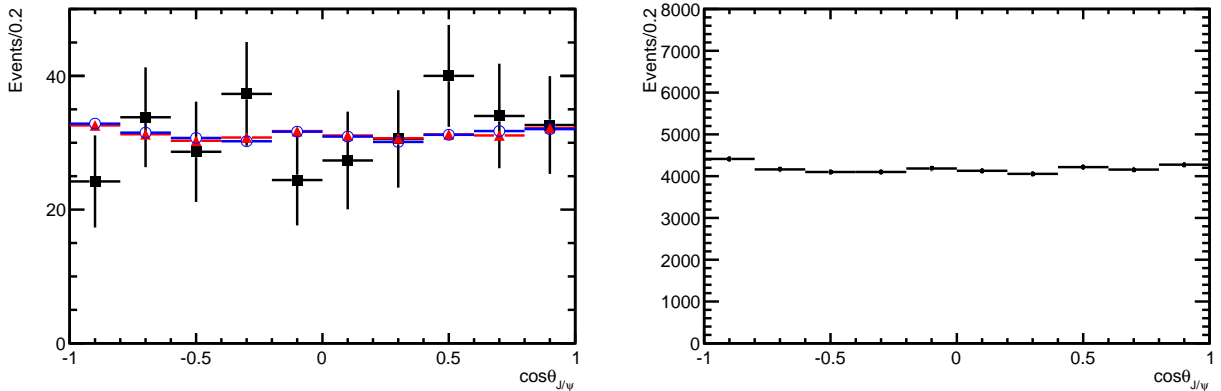


Figure 3.30: Distribution of  $\cos\theta_{J/\psi}$  for the data (black points with squares) and for the  $1^{++}$  (red points with triangles) and  $2^{-+}$  (blue points with open circles) fit projections on the left, and for the phase-space MC on the right.

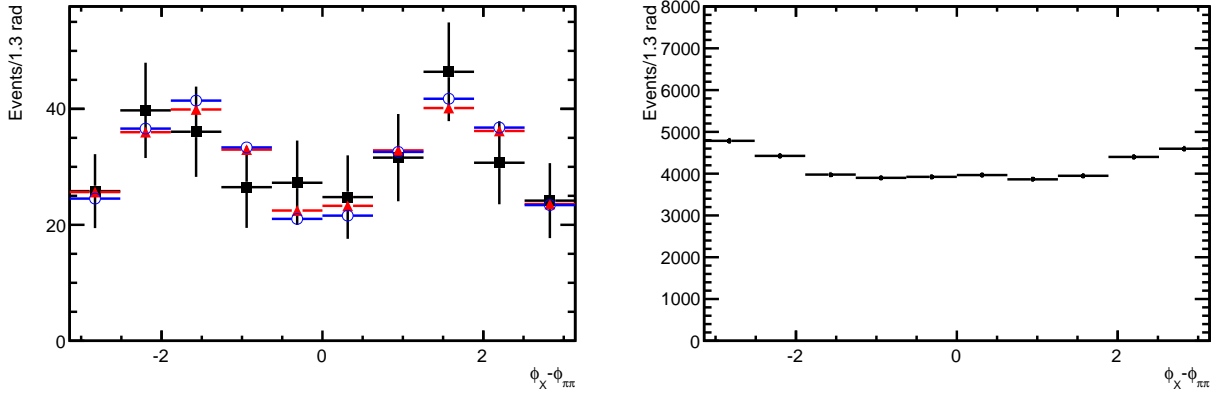


Figure 3.31: Distribution of  $\phi_X - \phi_{\pi\pi}$  for the data (black points with squares) and for the  $1^{++}$  (red points with triangles) and  $2^{-+}$  (blue points with open circles) fit projections on the left, and for the phase-space MC on the right.

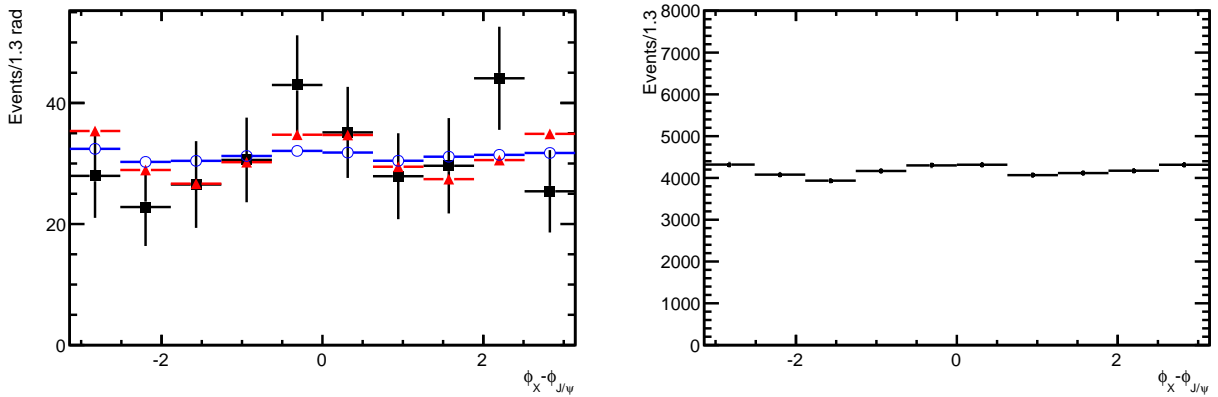


Figure 3.32: Distribution of  $\phi_X - \phi_{J/\psi}$  for the data (black points with squares) and for the  $1^{++}$  (red points with triangles) and  $2^{-+}$  (blue points with open circles) fit projections on the left, and for the phase-space MC on the right.

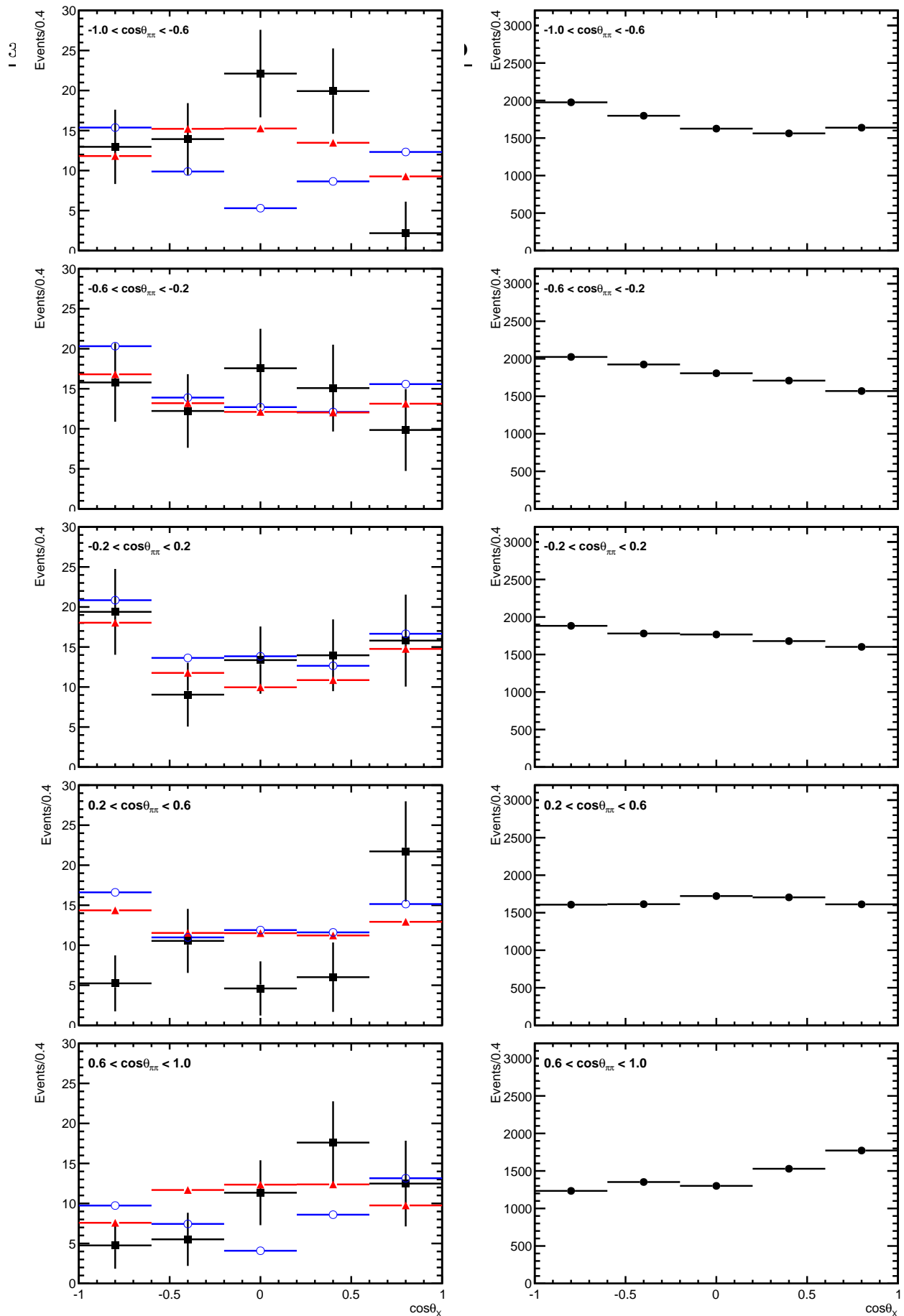


Figure 3.33: Distribution of  $\cos\theta_X$  for the data (black points with squares) and for the  $1^{++}$  (red points with triangles) and  $2^{++}$  (blue points with open circles) fit projections on the left, and for the phase-space MC on the right, for 5 different  $\cos\theta_{\pi\pi}$  bins.

Table 3.5: Tests performed on 1- and 2-dimensional projections of the data and of the fits. The last column shows predictions for an average expected  $n_{\sigma}^{proj}$  obtained by fitting the large statistics EVTGEN  $1^{++}$  MC and scaling down to the actual data sample.

Distribution	Data			$1^{++}$ MC $< n_{\sigma}^{proj} >$
	$CL_{proj}(1^{++})$	$CL_{proj}(2^{-+})$	$n_{\sigma}^{proj}$	
$\cos \theta_X$	6%	0.14%	$2.15\sigma$	$1.97\sigma$
$\cos \theta_{J/\psi}$	82%	61.73%	$0.09\sigma$	$0.25\sigma$
$\cos \theta_{\pi\pi}$	36%	1.60%	$2.24\sigma$	$2.44\sigma$
$\phi_X - \phi_{J/\psi}$	58%	40.70%	$0.00\sigma$	$0.84\sigma$
$\phi_X - \phi_{\pi\pi}$	97%	82.79%	$0.35\sigma$	$0.36\sigma$
$\cos \theta_X$ vs. $\cos \theta_{J/\psi}$	89%	10.49%	$3.40\sigma$	$2.03\sigma$
$\cos \theta_X$ vs. $\cos \theta_{\pi\pi}$	18%	0.02%	$4.49\sigma$	$3.87\sigma$
$\cos \theta_{J/\psi}$ vs. $\cos \theta_{\pi\pi}$	91%	46.32%	$2.10\sigma$	$2.37\sigma$
$\phi_X - \phi_{J/\psi}$ vs. $\phi_X - \phi_{\pi\pi}$	72%	56.50%	$0.03\sigma$	$0.74\sigma$
$\cos \theta_X$ vs. $\phi_X - \phi_{J/\psi}$	7%	0.16%	$2.80\sigma$	$1.67\sigma$
$\cos \theta_X$ vs. $\phi_X - \phi_{\pi\pi}$	31%	1.09%	$2.96\sigma$	$1.68\sigma$
$\cos \theta_{J/\psi}$ vs. $\phi_X - \phi_{J/\psi}$	63%	49.78%	$0.00\sigma$	$0.14\sigma$
$\cos \theta_{J/\psi}$ vs. $\phi_X - \phi_{\pi\pi}$	91%	80.72%	$0.23\sigma$	$0.00\sigma$
$\cos \theta_{\pi\pi}$ vs. $\phi_X - \phi_{J/\psi}$	99%	77.67%	$1.78\sigma$	$2.16\sigma$
$\cos \theta_{\pi\pi}$ vs. $\phi_X - \phi_{\pi\pi}$	64%	18.44%	$1.92\sigma$	$2.10\sigma$

the three middle slices. These two slices added together, i.e. the data for  $|\cos \theta_{\pi\pi}| > 0.6$ , are shown in figure 3.34. This distribution shows much stronger separation between the spin hypotheses in  $\cos \theta_X$  than without any  $\cos \theta_{\pi\pi}$  requirement (figure 3.28), illustrating importance of employing the angular correlations in the spin analysis. The 2011 data clearly prefer the  $1^{++}$  hypothesis in figure 3.34. Rejection of the  $2^{-+}$  hypothesis for all variables and correlations follows the pattern predicted by the  $1^{++}$  MC. The deviations from the expectations are well within the expected fluctuations, which we studied on toy experiments.

### 3.6 Stability of result with background level in the data

Since we did not want to bias  $X(3872)$  sample in any way, our nominal data selection does not impose any requirements on  $M(\pi^+\pi^-)$ . Because of strong peaking of the  $X(3872)$  signal at high  $M(\pi^+\pi^-)$  values it is easy to reduce the background level in our sample



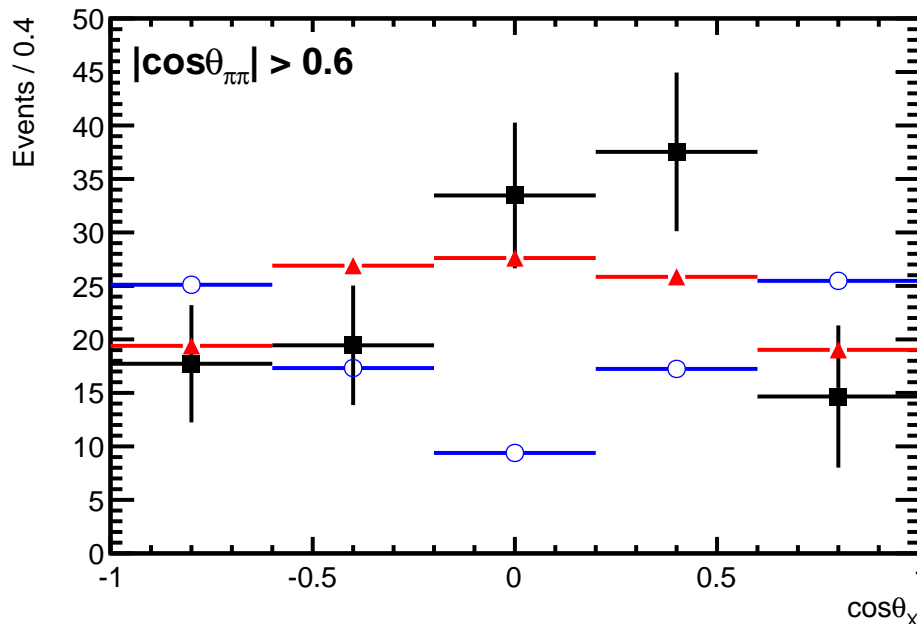


Figure 3.34: Distribution of  $\cos\theta_X$  with the  $|\cos\theta_{\pi\pi}| > 0.6$  requirement for the data (black points with squares) and for the projections of the 5-dimensional fit of the  $1^{++}$  hypothesis (red points with triangles) and  $2^{-+}$  hypothesis (blue points with open circles). Compare to figure 3.28.

without sacrificing much of the signal yield - see figure 3.12. It is traditional to cut on  $Q \equiv M(J/\psi \pi^+ \pi^-) - M(J/\psi) - M(\pi^+ \pi^-)$  instead of  $M(\pi^+ \pi^-)$ . The results of the 5D likelihood analysis performed with different values of  $Q$  cut are given in table 3.6. The background under the  $X(3872)$  peak is cut by a factor of 4 in this variation, while the signal decreases only by about 20%. The  $M(\pi^+ \pi^- J/\psi) - M(J/\psi)$  mass spectrum for the hardest cut is shown in figure 3.35, to be compared to figure 3.8 in the nominal analysis. The background changes not only in the magnitude but also in shape, as illustrated by angular distributions observed for the background events in the  $X(3872)$   $M(\pi^+ \pi^- J/\psi) - M(J/\psi)$  sidebands; figures 3.36-3.40.

The results, for the significance of  $2^{-+}$  rejection and  $\alpha$  in the  $2^{-+}$  fit are very stable. The significance is the best with a moderate  $Q$  cut, for both the data and for the  $1^{++}$  MC. Since the gain is fairly small and qualitatively completely negligible, we have decided to stick to the  $M(\pi^+ \pi^-)$  unbiased selection for the nominal results.

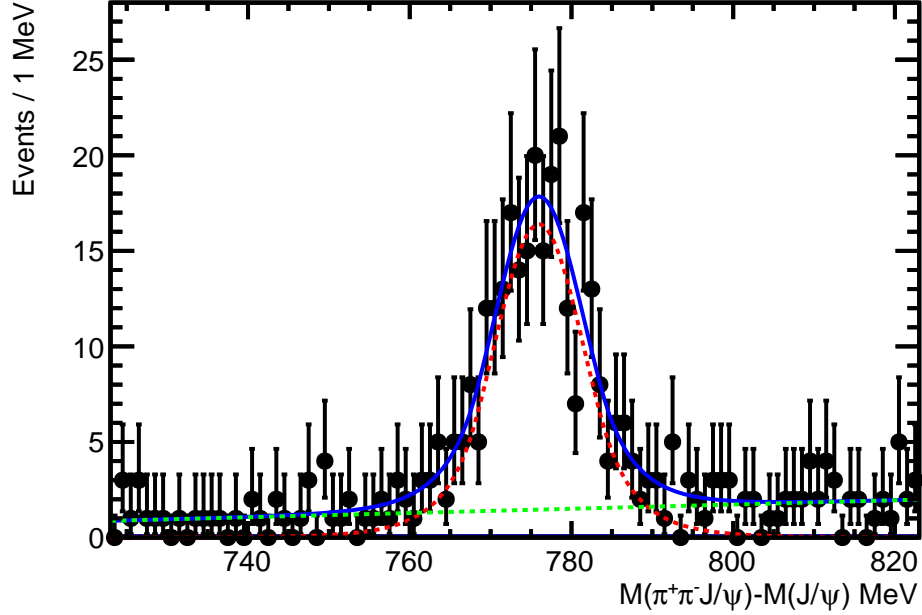


Figure 3.35: Distribution of  $M(\pi^+\pi^-J/\psi) - M(J/\psi)$  for  $B^+ \rightarrow J/\psi K^+\pi^+\pi^-$  events with the requirement  $Q < 100$  MeV. The fit of the  $X(3872)$  signal is displayed. The blue, red and green blue lines represent the total fit, signal component and background component, respectively. Compare to figure 3.10 which has no  $Q$  cut.

Table 3.6: Results of the 5D likelihood ratio test between the  $1^{++}$  and  $2^{-+}$   $J^{PC}$  hypotheses for various cuts on  $Q$  (i.e.  $M(\pi^+\pi^-)$ ) for the 2011 data. The first row corresponds to our nominal results. The MC predictions included in the table are based on high statistics EVTGEN  $1^{++}$  MC sample, with the results scaled down to the signal statistics ( $N_s$ ) observed in the data and multiplied by the  $s_w$  factor to account for the background subtraction error.

Q cut MeV	Signal			Background		2 <sup>-+</sup> rejection			α in 2 <sup>-+</sup> fit	
	$N_s$ data	$N_s/N_s^0$ data	MC	$N_b$ data	$\frac{N_b}{N_b^0}$ data	$n_\sigma^*$ data	MC	$n_\sigma$ data	data	
no cut	313 ± 26	1.00	1.00	568 ± 31	1.00	9.9σ	10.5σ	9.7σ	(.671 ± .046, .280 ± .046)	
< 250	295 ± 23	0.94	0.97	361 ± 24	0.64	10.5σ	11.0σ	10.3σ	(.688 ± .044, .269 ± .047)	
< 150	280 ± 21	0.90	0.90	217 ± 19	0.38	10.4σ	11.3σ	10.2σ	(.699 ± .046, .290 ± .046)	
< 100	245 ± 19	0.78	0.79	143 ± 16	0.25	9.9σ	11.0σ	9.7σ	(.689 ± .048, .299 ± .048)	

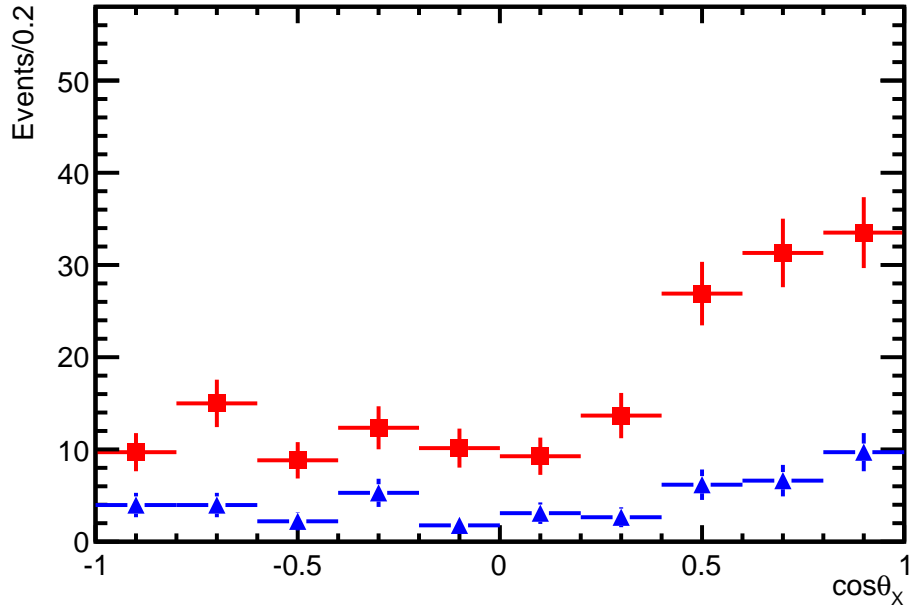


Figure 3.36: Distribution of  $\cos\theta_X$  for the data in the  $X(3872)$  sidebands scaled to the expected background within the  $X(3872)$  mass peak ( $\pm 3\sigma$ ) for the selection with no  $Q$  requirement (red points with squares) and for  $Q < 100$  MeV (blue points with triangles).

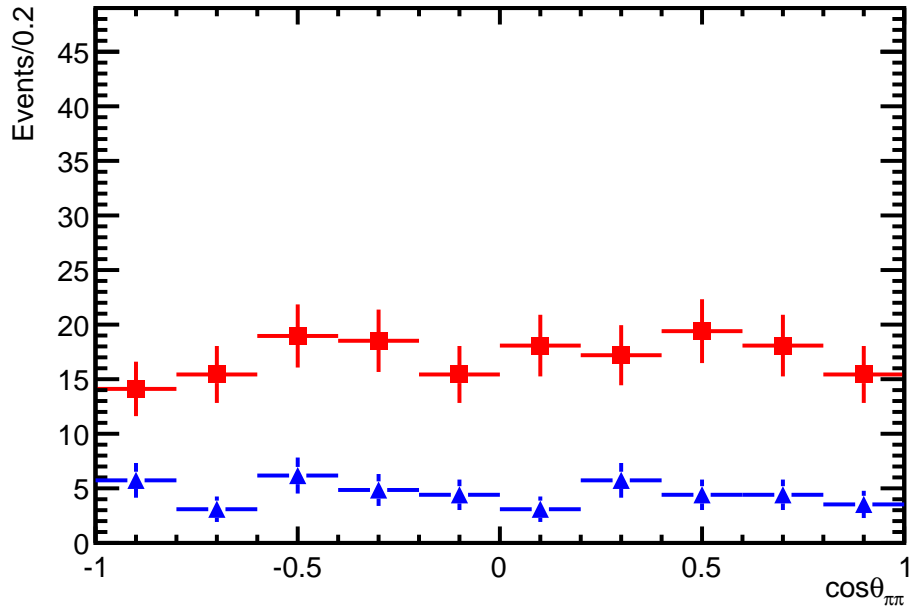


Figure 3.37: Distribution of  $\cos\theta_{\pi\pi}$  for the data in the  $X(3872)$  sidebands scaled to the expected background within the  $X(3872)$  mass peak ( $\pm 3\sigma$ ) for the selection with no  $Q$  requirement (red points with squares) and for  $Q < 100$  MeV (blue points with triangles).

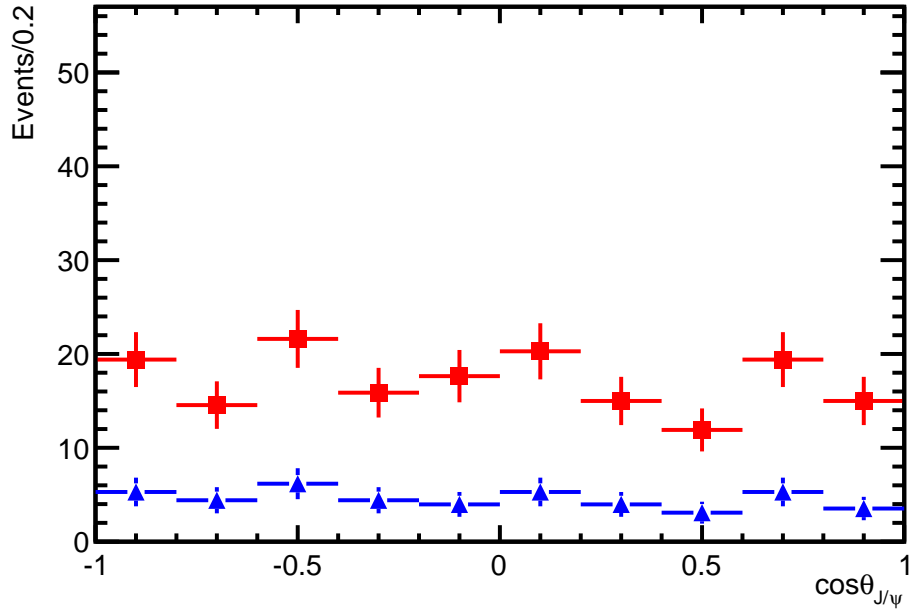


Figure 3.38: Distribution of  $\cos\theta_{J/\psi}$  for the data in the  $X(3872)$  sidebands scaled to the expected background within the  $X(3872)$  mass peak ( $\pm 3\sigma$ ) for the selection with no  $Q$  requirement (red points with squares) and for  $Q < 100$  MeV (blue points with triangles).

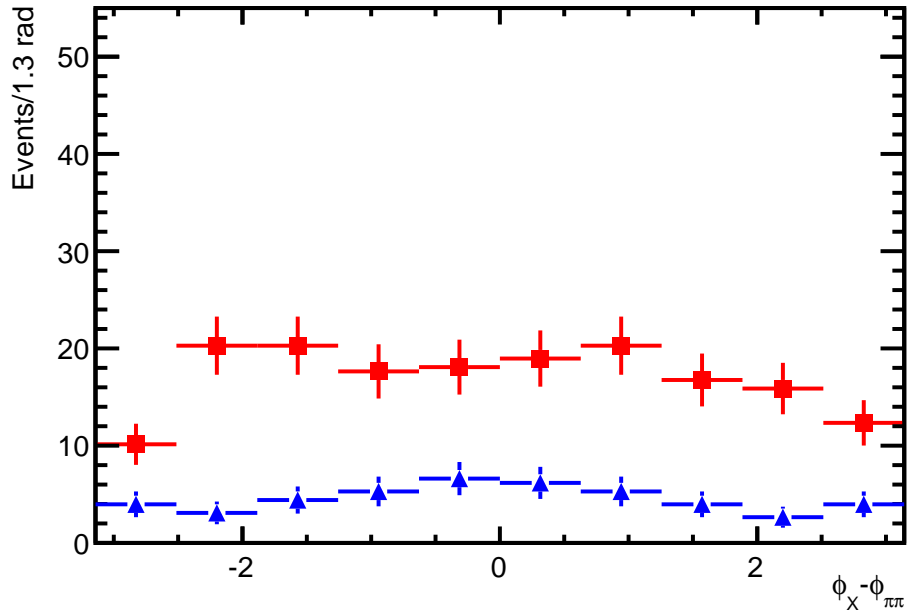


Figure 3.39: Distribution of  $\phi_X - \phi_{\pi\pi}$  for the data in the  $X(3872)$  sidebands scaled to the expected background within the  $X(3872)$  mass peak ( $\pm 3\sigma$ ) for the selection with no  $Q$  requirement (red points with squares) and for  $Q < 100$  MeV (blue points with triangles).

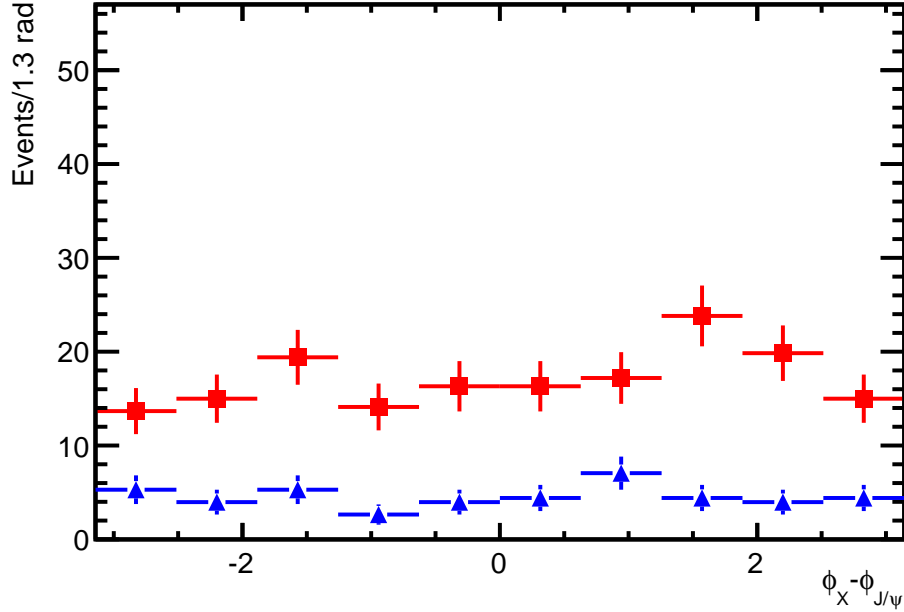


Figure 3.40: Distribution of  $\phi_X - \phi_{J/\psi}$  for the data in the  $X(3872)$  sidebands scaled to the expected background within the  $X(3872)$  mass peak ( $\pm 3\sigma$ ) for the selection with no  $Q$  requirement (red points with squares) and for  $Q < 100$  MeV (blue points with triangles).

### 3.7 Stability of result with changes of signal efficiency

To investigate stability of the results with respect to variation of signal efficiency we vary cuts on  $p_T$  for pions, muons and the kaon separately. We change a  $p_T$  cut for the pions (muons) from the nominal value of  $> 250$  ( $> 900$ ) MeV to  $> 500$  ( $> 2000$ ) MeV. We change a  $p_T$  cut for the kaon from the nominal value of  $> 250$  MeV to  $> 2000$  MeV. The results are shown in table 3.7. The efficiency and background change by roughly a factor of 2 in this variation. When varying the pion or kaon cut the significance of  $2^{-+}$  rejection drops following the trend predicted by the  $1^{++}$  MC. Even for the hardest cut, the significance is still large  $n_\sigma = 5.8\sigma$ . When changing the muon cut, the significance drops less than average drop expected from the MC. This can be accounted for by relatively large fluctuations in the achieved significance expected for a single experiment, as revealed by the toy experiments and illustrated in figure 3.22.

Table 3.7: Results of the 5D likelihood ratio test between the  $1^{++}$  and  $2^{-+}$   $J^{PC}$  hypotheses for various cuts on  $p_T(\pi)$  for the 2011 data. The first row corresponds to our nominal results. The MC predictions included in the table are based on high statistics EVTGEN  $1^{++}$  MC sample, with the results scaled down to the signal statistics ( $N_s$ ) observed in the data and multiplied by the  $s_w$  factor to account for the background subtraction error.

$p_T$ cut MeV	Signal			Background		$2^{-+}$ rejection			$\alpha$ in $2^{-+}$ fit
	$N_s$ data	$N_s/N_s^0$ data	MC	$N_b$ data	$\frac{N_b}{N_b^0}$ data	$n_\sigma^*$ data	MC	$n_\sigma$ data	data
$p_T$ cut on pions									
> 250	$313 \pm 26$	1.00	1.00	$568 \pm 31$	1.00	$9.9\sigma$	$10.5\sigma$	$9.7\sigma$	$(.671 \pm .046, .280 \pm .046)$
> 300	$298 \pm 25$	0.95	0.88	$470 \pm 28$	0.83	$9.9\sigma$	$10.3\sigma$	$9.8\sigma$	$(.685 \pm .049, .288 \pm .048)$
> 400	$227 \pm 22$	0.73	0.63	$340 \pm 25$	0.60	$7.9\sigma$	$8.5\sigma$	$7.7\sigma$	$(.697 \pm .061, .319 \pm .041)$
> 500	$168 \pm 19$	0.54	0.45	$245 \pm 20$	0.43	$6.7\sigma$	$7.0\sigma$	$6.6\sigma$	$(.649 \pm .065, .310 \pm .060)$
$p_T$ cut on muons									
> 900	$313 \pm 26$	1.00	1.00	$568 \pm 31$	1.00	$9.9\sigma$	$10.5\sigma$	$9.7\sigma$	$(.671 \pm .046, .280 \pm .046)$
> 1000	$299 \pm 25$	0.96	0.95	$532 \pm 30$	0.94	$10.1\sigma$	$10.4\sigma$	$10.0\sigma$	$(.685 \pm .035, .253 \pm .047)$
> 1500	$215 \pm 21$	0.69	0.68	$400 \pm 25$	0.70	$10.0\sigma$	$8.8\sigma$	$9.9\sigma$	$(.672 \pm .045, .303 \pm .059)$
> 2000	$133 \pm 16$	0.43	0.43	$262 \pm 20$	0.46	$8.4\sigma$	$6.7\sigma$	$8.2\sigma$	$(.725 \pm .063, .318 \pm .080)$
$p_T$ cut on kaon									
> 250	$313 \pm 26$	1.00	1.00	$568 \pm 31$	1.00	$9.9\sigma$	$10.5\sigma$	$9.7\sigma$	$(.671 \pm .046, .280 \pm .046)$
> 500	$304 \pm 25$	0.97	0.95	$486 \pm 28$	0.86	$9.4\sigma$	$10.5\sigma$	$9.2\sigma$	$(.654 \pm .035, .294 \pm .045)$
> 1000	$244 \pm 22$	0.78	0.78	$406 \pm 25$	0.71	$9.3\sigma$	$9.5\sigma$	$9.1\sigma$	$(.674 \pm .040, .284 \pm .049)$
> 1500	$198 \pm 20$	0.63	0.62	$314 \pm 23$	0.55	$7.3\sigma$	$8.3\sigma$	$7.2\sigma$	$(.714 \pm .051, .302 \pm .057)$
> 2000	$151 \pm 18$	0.48	0.47	$246 \pm 20$	0.43	$6.0\sigma$	$7.0\sigma$	$5.8\sigma$	$(.659 \pm .054, .315 \pm .063)$

In our analysis we assume that the efficiency in 5-dimensional angular phase-space,  $\epsilon(\theta_X, \theta_{J/\psi}, \theta_{\pi\pi}, \phi_X - \phi_{J/\psi}, \phi_X - \phi_{\pi\pi})$ , does not depend on  $X(3872)$  spin. As discussed in Sec. 3.3 small spin dependence of this efficiency can be introduced via modifications of  $\rho(770)$  Breit-Wigner shape via different values of orbital angular momentum between  $\rho(770)$  and  $J/\psi$  ( $L = 0$  for  $1^{++}$ ,  $L = 1$  for  $2^{-+}$ ). As an illustration of insensitivity of our results to  $M(\pi^+\pi^-)$  distribution, we perform the likelihood-ratio analysis of the 2011 data using phase-space  $X(3872)$  MC which decays  $X(3872)$  via 3-body phase-space model,  $X(3872) \rightarrow \pi^+\pi^- J/\psi$ , instead of the default phase-space MC in which  $X(3872)$  is decayed to  $\rho(770)J/\psi$  via 2-body phase-space model. In spite of a very big difference in the  $M(\pi^+\pi^-)$  distribution for these two MC samples (figure 3.41), the results hardly change as documented in table 3.8.

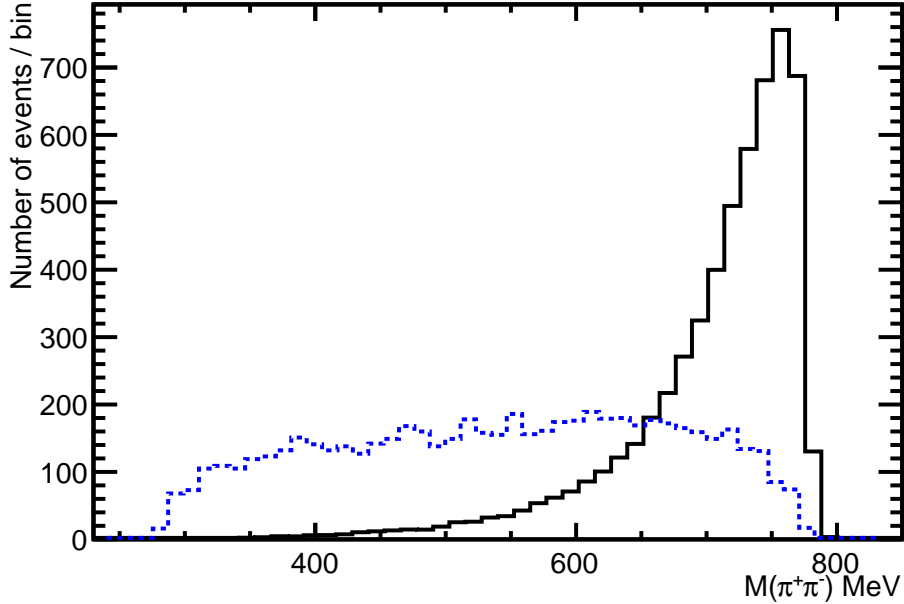


Figure 3.41: Distribution of  $M(\pi^+\pi^-)$  for two versions of  $B^+ \rightarrow X(3872)K^+$ ,  $X(3872) \rightarrow \pi^+\pi^- J/\psi$  phase-space MC events;  $X(3872) \rightarrow \rho(770)J/\psi$  2-body phase-space decay (black solid histogram) and  $X(3872) \rightarrow \pi^+\pi^- J/\psi$  3-body phase-space decay (blue dashed histogram).

Table 3.8: The results from  $1^{++}/2^{-+}$  hypotheses testing on the  $X(3872)$  2011 data using 5D likelihood ratio test. The two rows differ by which data were used to perform normalization integral of the  $\mathcal{P}$ s. The nominal fit to the data uses phase-space MC with  $X(3872) \rightarrow \rho(770)J/\psi$  decay (1st row). As an alternative we used phase-space MC with  $X(3872) \rightarrow \pi^+\pi^-J/\psi$  decay (2nd row).

phase-space MC used in $\mathcal{P}_{norm}$	$J^{PC}$	$\Delta\chi_L^2$	$n_\sigma$	$n_\sigma^*$	Fit result for $\alpha$ in the $2^{-+}$ fit $\alpha$
$X(3872) \rightarrow \rho J/\psi$	$1^{++}$	+99.4	$9.7\sigma$	$9.9\sigma$	$(.671 \pm .046, .280 \pm .046)$
$X(3872) \rightarrow \pi^+\pi^-J/\psi$	$1^{++}$	+105.1	$10.0\sigma$	$10.2\sigma$	$(.709 \pm .050, .294 \pm .050)$

### 3.8 Cross-check of likelihood analysis on $B^+ \rightarrow \psi(2S)K^+$ signal

The  $B^+ \rightarrow \psi(2S)K^+$ ,  $\psi(2S) \rightarrow \pi^+\pi^-J/\psi$  decays offer a high statistics signal in our data with similar kinematics to  $B^+ \rightarrow X(3872)K^+$ ,  $X(3872) \rightarrow \pi^+\pi^-J/\psi$  decays.  $J^{PC}$  of  $\psi(2S)$  is well established to be  $1^{--}$ . We apply our angular likelihood analysis to the  $\psi(2S)$  data, to cross-check the method and to validate efficiency simulations.

Since  $C$ -parity is negative, allowed values of angular momentum between the two pions is  $L_{\pi\pi} = 0, 2, \dots$  (see Sec. 3.2). We can safely neglect values of 2 and higher since the di-pion mass is limited to less than 560 MeV. Therefore, the di-pion system must be in  $0^{++}$  state. This simplifies the matrix element a lot since the helicity of the dipion system is fixed to zero. The amplitudes are now labeled by  $\lambda_{J/\psi}$  helicity only. Also because the system is spinless,



### 3.8. CROSS-CHECK OF LIKELIHOOD ANALYSIS ON $B^+ \rightarrow \psi(2S)K^+$ SIGNAL 125

the  $\cos \theta_{\pi\pi}$  and  $\phi_{pp}$  distributions are flat and carry no useful information:

$$|M(\theta_X, \theta_{J/\psi}, \phi_X - \phi_{J/\psi})|^2 \propto \quad (3.28)$$

$$\sum_{\Delta\lambda_\mu=1,+1} \left| \sum_{\lambda_{J/\psi}=1,0,+1} D_{\lambda_{J/\psi}, \Delta\lambda_\mu}^1(\phi_{J/\psi}, \theta_{J/\psi}, -\phi_{J/\psi}) \times \quad (3.29)$$

$$D_{0, \lambda_{J/\psi}}^{J_X}(\phi_X, \theta_X, -\phi_X) A_{\lambda_{J/\psi}} \right|^2 \quad (3.30)$$

where the subscript  $X$  should be understood as referring to the  $\psi(2S)$ . For the actual  $J^{PC}$  of  $\psi(2S)$ ,  $P$ -parity conservation (equation 3.6) requires the helicity coupling to be symmetric,

$$A_{\lambda_{J/\psi}} = A_{-\lambda_{J/\psi}}, \quad (3.31)$$

thus there are two couplings to consider  $A_1$  and  $A_0$ . We introduce a free complex parameter  $\alpha_{2S}^{1--} = A_1/A_0$ , which we fit to the data.

We test how well the  $\mathcal{P}(\theta_X, \theta_{J/\psi}, \phi_X - \phi_{J/\psi} | 1^{--})$  describes the  $\psi(2S)$  data by calculating  $CL_{proj}$  from the  $\chi^2$  values for the binned 1- and 2-dimensional histograms - table 3.9. These confidence level are acceptable, in spite of the large signal statistics, from which we conclude that the phase-space MC used for the efficiency corrections is accurate.

As a test of the spin discrimination method, we test the  $1^{--}$  hypothesis against the other spin assignments with odd  $C$ -parity. For  $2^{+-}$  the helicity couplings are also symmetric, thus this hypothesis has also a free complex parameter  $\alpha_{2S}^{2+-} = A_1/A_0$ . For  $1^{+-}$  and  $2^{--}$  hypotheses, the helicity couplings are antisymmetric,

$$A_{\lambda_{J/\psi}} = -A_{-\lambda_{J/\psi}}, \quad (3.32)$$

which implies  $A_0 = 0$ . Since there is only one coupling there are no free parameters in the fit.

We first test the likelihood analysis on high statistics fully simulated EVTGEN  $\psi(2S)$  MC

Table 3.9: Tests performed on 1- and 2-dimensional projections of the  $\psi(2S)$  data and of the  $1^{--}$  fit.

Distribution	$CL_{proj}(1^{--})$
$\cos \theta_X$	52%
$\cos \theta_{J/\psi}$	26%
$\cos \theta_{\pi\pi}$	32%
$\phi_X - \phi_{J/\psi}$	13%
$\phi_X - \phi_{\pi\pi}$	21%
$\cos \theta_X$ vs. $\cos \theta_{J/\psi}$	90%
$\cos \theta_X$ vs. $\cos \theta_{\pi\pi}$	50%
$\cos \theta_{J/\psi}$ vs. $\cos \theta_{\pi\pi}$	15%
$\phi_X - \phi_{J/\psi}$ vs. $\phi_X - \phi_{\pi\pi}$	32%
$\cos \theta_X$ vs. $\phi_X - \phi_{J/\psi}$	34%
$\cos \theta_X$ vs. $\phi_X - \phi_{\pi\pi}$	27%
$\cos \theta_{J/\psi}$ vs. $\phi_X - \phi_{J/\psi}$	46%
$\cos \theta_{J/\psi}$ vs. $\phi_X - \phi_{\pi\pi}$	4%
$\cos \theta_{\pi\pi}$ vs. $\phi_X - \phi_{J/\psi}$	10%
$\cos \theta_{\pi\pi}$ vs. $\phi_X - \phi_{\pi\pi}$	5%

Table 3.10: The results from  $1^{--}/J^{P-}$  hypotheses testing on the real  $\psi(2S)$  data and on the  $1^{--}$  EVTGEN MC using 5D likelihood ratio test. The results for the MC are scaled to the 2011 signal statistics, except for the errors on  $\alpha_{2S}$  (scaling the errors gives  $\pm 0.029$  for the expected  $\alpha_{2S}^{1^{--}}$  errors in the data).

analyzed		$J^{PC}$ preferred		$J^{PC}$ rejected			
data				$1^{+-}$	$2^{--}$	$2^{+-}$	
$N_s$		$\alpha_{2S}^{1^{--}}$		$n_\sigma^*$	$n_\sigma^*$	$n_\sigma^*$	$\alpha_{2S}^{2^{+-}}$
2011	$5642 \pm 76$	$1^{--}$	$(0.970 \pm 0.028, 0.989 \pm 0.027)$	$87.5\sigma$	$107.4\sigma$	$58.3\sigma$	$(.86 \pm .02, .84 \pm .02)$
$1^{--}$ -MC	$27407 \rightarrow 5642$	$1^{--}$	$(1.045 \pm 0.013, 1.025 \pm 0.013)$	$85.1\sigma$	$107.3\sigma$	$61.2\sigma$	$(.83 \pm .01, .89 \pm .01)$

sample. We then apply it to the real data. The results are given in the table 3.10. The  $1^{--}$  hypothesis is preferred over the other assignments by very large margins (from 57 to  $109\sigma$ ). The values of  $\alpha_{2S}$  and the significances of rejections of the other spin hypotheses are in good agreement between the data and the MC. From this we conclude that the multidimensional likelihood analysis we have developed passes the check on the  $\psi(2S)$  data very well. Naively scaling the  $\psi(2S)$   $2^{+-}$  rejection margin,  $57\sigma$ , to the  $X(3872)$  signal statistics and background ( $\sqrt{\frac{X(3872)}{s_w} \frac{N_s^{X(3872)}}{N_s^{\psi(2S)}}}$ ), yields  $9.9\sigma$  for an expected sensitivity in the  $X(3872)$  spin analysis. As the underlying angular distributions are different this should be taken only as an order of magnitude estimate. This estimate is close to the actual sensitivity achieved.

### 3.9 Using $\psi(2S)$ signal for efficiency corrections in $X(3872)$ fit.

The high statistics  $B^+ \rightarrow \psi(2S)K^+$ ,  $\psi(2S) \rightarrow \pi^+\pi^- J/\psi$  signal in the data offers opportunity to eliminate use of Monte Carlo for the efficiency corrections in the 5D likelihood analysis of the  $X(3872)$  data. We reweight the  $\psi(2S)$  data by an inverse of the matrix element squared given by equation 3.30 for  $J^{PC} = 1^{--}$  with the value of  $\alpha_{2S}^{1^{--}}$  from the fit to the  $\psi(2S)$  data (table 3.10). We then use these reweighted  $\psi(2S)$  data to carry out the numerical integration of  $X(3872)$   $\mathcal{P}$  given by equation 3.21. Results of the likelihood analysis on the 2011  $X(3872)$  data obtained with this approach are compared to our nominal results based on use of the phase-space MC to calculate  $\mathcal{P}_{norm}$ . The agreement is quite good, from which we conclude that there is no evidence that the MC efficiency corrections bias the likelihood analysis. We stick to use of the phase-space MC in the nominal fit, since the statistics is much larger. Also the  $\psi(2S)$  decays produce different  $M(\pi^+\pi^-)$  distribution than the  $X(3872)$  decays, which cannot be corrected for because of the different range. Since the difference in  $M(\pi^+\pi^-)$  distribution between the  $\psi(2S)$  and  $X(3872)$  decays is much larger than any difference in  $M(\pi^+\pi^-)$  distribution caused by different  $J_X$  values, the results in this section also probe for

Table 3.11: The results from  $1^{++}/2^{-+}$  hypotheses testing on the  $X(3872)$  2011 data using 5D likelihood ratio test. The two rows differ by which data were used to perform normalization integral of the  $\mathcal{P}$ . The nominal fit to the data uses high statistics phase-space MC (1st row). As an alternative (2nd row) we used the  $\psi(2S)$  signal in the data, reweighted to the phase-space distribution by an inverse of the matrix element squared given by equation 3.30 for  $J^{PC} = 1^{--}$  with the value of  $\alpha_{2S}^{1^{--}}$  from the fit to the  $\psi(2S)$  data (see table 3.10).

data used in $\mathcal{P}_{norm}$	$J^{PC}$	$\Delta\chi_L^2$	$n_\sigma$	$n_\sigma^*$	Fit result for $\alpha$ in the $2^{-+}$ fit
phase-space MC	$1^{++}$	+99.4	$9.7\sigma$	$9.9\sigma$	$(0.671 \pm 0.046, 0.280 \pm 0.046)$
reweighted $\psi(2S)$ data	$1^{++}$	+94.1	$9.4\sigma$	$9.6\sigma$	$(0.650 \pm 0.045, 0.272 \pm 0.045)$

systematics of neglecting the remnant  $J_X$  dependence in  $\epsilon(\theta_X, \theta_{J/\psi}, \theta_{\pi\pi}, \phi_X - \phi_{J/\psi}, \phi_X - \phi_{\pi\pi})$  discussed in Sec. 3.3.

### 3.10 Testing other $J^{PC}$ assignments to $X(3872)$ .

While the other spin assignments than  $1^{++}$  and  $2^{-+}$  were previously ruled out by the CDF, it is interesting to see how well they are excluded by our data. Since we have developed code to test odd  $C$ -parity assignments for the  $\psi(2S)$  fitting we can easily apply it to the  $X(3872)$  data as well. The results are summarized in table 3.12. The odd  $C$ -parity assignments are ruled out by our angular likelihood ratio tests at the level of  $13.2\sigma$  or higher, with the sensitivity pattern matching the predictions from the  $1^{++}$  EVTGEN MC. We have not developed code to test the other even  $C$ -parity spin assignments.

Table 3.12: The results from  $1^{++}/J^{P-}$  hypotheses testing on the real  $X(3872)$  2011 data and on the  $1^{++}$  EVTGEN MC using full-phase likelihood ratio test. The results for the MC are scaled to the 2011 signal statistics.

analyzed	$J^{PC}$	$J^{PC}$ rejected			
		$1^{--}$	$1^{+-}$	$2^{--}$	$2^{+-}$
data	pref- erred	$n_\sigma$	$n_\sigma$	$n_\sigma$	$n_\sigma$
2011	$1^{++}$	$13.2\sigma$	$15.5\sigma$	$18.4\sigma$	$13.7\sigma$
$1^{++}$ MC	$1^{++}$	$14.3\sigma$	$14.8\sigma$	$18.9\sigma$	$14.2\sigma$

### 3.11 Performing analysis in less than full phase-space

There are no benefits from performing the likelihood analysis in less than full angular phase-space, by neglecting some angular degrees of freedom. Sensitivity is expected to drop, as information in the neglected angular correlations is lost. Furthermore, it is difficult to perform exact efficiency corrections in the reduced dimensions, as the efficiency starts to depend on the fitted spin hypothesis and possibly its free parameter  $\alpha$ . We still apply the efficiency corrections via equation 3.21, however this method is only an approximation now, thus some systematic biases can occur. The likelihood calculation and the hypotheses testing remain the same, thus there is no simplification from dealing with smaller number of angles. A prejudice that analyses performed in smaller number of input angles is simpler is based on binned likelihood fits, in which dealing with many dimensions is cumbersome, or even prohibitive as the number of bins to consider explodes when adding dimensions. Complexity of unbinned likelihood fit is the same in 1- and 5-dimensions. We perform an exercise of reducing dimensionality only to gain an insight to which angular correlations are the most powerful and out of curiosity how sensitivity degrades with the dimensions.

The full phase-space matrix element is 5-dimensional and is given by equation 3.5. We

perform to 3–dimensional analysis, by neglecting either  $J/\psi$  or  $\pi\pi$  decay angles. By doing so we make either  $\lambda_{J/\psi}$  or  $\lambda_{\pi\pi}$  states incoherent, thus sum over these helicity values moves to the outside of the modulus of the matrix element squared. When dropping the  $J/\psi$  decay angles, we get:

$$|M(\theta_X, \theta_{\pi\pi}, \phi_X - \phi_{\pi\pi})|^2 \propto \quad (3.33)$$

$$\sum_{\Delta\lambda_\mu=1,+1} \sum_{\lambda_{J/\psi}=1,0,+1} \left| \sum_{\lambda_{\pi\pi}=1,0,+1} D_{\lambda_{\pi\pi},0}^1(\phi_{\pi\pi}, \theta_{\pi\pi}, -\phi_{\pi\pi}) \times \quad (3.34)$$

$$D_{0,\lambda_{J/\psi}-\lambda_{\pi\pi}}^{J_X}(\phi_X, \theta_X, -\phi_X) A_{\lambda_{J/\psi}, \lambda_{\pi\pi}} \right|^2 \quad (3.35)$$

When dropping the  $\pi\pi$  decay angles, we obtain:

$$|M(\theta_X, \theta_{J/\psi}, \phi_X - \phi_{J/\psi})|^2 \propto \quad (3.36)$$

$$\sum_{\Delta\lambda_\mu=1,+1} \sum_{\lambda_{\pi\pi}=1,0,+1} \left| \sum_{\lambda_{J/\psi}=1,0,+1} D_{\lambda_{J/\psi}, \Delta\lambda_\mu}^1(\phi_{J/\psi}, \theta_{J/\psi}, -\phi_{J/\psi}) \times \quad (3.37)$$

$$D_{0,\lambda_{J/\psi}-\lambda_{\pi\pi}}^{J_X}(\phi_X, \theta_X, -\phi_X) A_{\lambda_{J/\psi}, \lambda_{\pi\pi}} \right|^2 \quad (3.38)$$

To reduce 3–dimensions to 2, we numerically integrate the 3D formulae given above over the neglected angle. We also perform 1–dimensional analysis in  $\cos\theta_X$ , as this angle has been studied in LHCb before [23, 24]. Since in this approach, both  $J/\psi$  and  $\pi\pi$  decay angles are neglected both  $\lambda_{J/\psi}$  and  $\lambda_{\pi\pi}$  states are incoherent:

$$|M(\theta_X)|^2 \propto \quad (3.39)$$

$$\sum_{\Delta\lambda_\mu=1,+1} \sum_{\lambda_{J/\psi}=1,0,+1} \sum_{\lambda_{\pi\pi}=1,0,+1} \left| D_{0,\lambda_{J/\psi}-\lambda_{\pi\pi}}^{J_X}(\phi_X, \theta_X, -\phi_X) A_{\lambda_{J/\psi}, \lambda_{\pi\pi}} \right|^2 = \quad (3.40)$$

$$\sum_{\Delta\lambda_\mu=1,+1} \sum_{\lambda_{J/\psi}=1,0,+1} \sum_{\lambda_{\pi\pi}=1,0,+1} \left| d_{0,\lambda_{J/\psi}-\lambda_{\pi\pi}}^{J_X}(\theta_X) A_{\lambda_{J/\psi}, \lambda_{\pi\pi}} \right|^2 \quad (3.41)$$

Not surprisingly the dependence on  $\phi_X$  drops out.

The results from the likelihood tests performed in reduced dimensions are summarized in

table 3.13. We have included the results from the 5D analysis for an easy comparison. For each analysis we show verification on high statistics EVTGEN MC samples for both  $2^{-+}$  (with the three different values of  $\alpha$ ) and  $1^{++}$  analysis. The MC results for log-likelihood difference ( $\Delta\chi_L^2$ ) and for significance of the rejection of the other spin hypothesis ( $1^{++}$  for the  $2^{-+}$  MC,  $2^{-+}$  for the  $1^{++}$  MC) are scaled down to the signal statistics and the background subtraction error in the 2011 data (only the errors on fitted  $\alpha$  parameter are not scaled). The significances were calculated using the Wilks' theorem. The sign of  $\Delta\chi_L^2$  gives the preferred spin; negative values mean  $2^{-+}$  is preferred, positive values mean  $1^{++}$  is preferred.

From the two 3D analysis it is clear that correlations between the  $X(3872)$  and  $\pi\pi$  decay angles are by far more important than correlations between the the  $X(3872)$  and  $J/\psi$  decay angles. The former establishes  $1^{++}$  at  $7.2\sigma$  level, which is not far behind  $9.7\sigma$  from the nominal 5D analysis. The other 3D approach produces only  $1.8\sigma$  preference for  $1^{++}$ , thus for further reduction of dimensionality we pursue only combinations of the  $X(3872)$  and  $\pi\pi$  angles.

From the two 2D analysis we performed it is clear that  $\cos\theta_X$  vs.  $\cos\theta_{\pi\pi}$  correlation is by far more important than  $\cos\theta_X$  vs.  $\phi_X - \phi_{\pi\pi}$  correlation. The later has only  $0.5\sigma$  preference for the  $1^{++}$  hypotheses, whereas the former produces respectable  $5.5\sigma$  rejection of the  $2^{-+}$  spin.

When performing 1D likelihood fits to  $\cos\theta_X$ , we found out that this method has essentially no sensitivity to the imaginary part of  $\alpha$  in the  $2^{-+}$  fits. Therefore, we fix this part to zero, which makes this study directly comparable to the previous studies in LHCb, in which the  $\alpha$  was assumed to be real. From the fits to the  $2^{-+}$  MC samples, we see that  $2^{-+}$  with  $\alpha = 1.0$  is easy to distinguish from  $1^{++}$ . However,  $2^{-+}$  with  $\alpha = 0.5$  is completely indistinguishable from the  $1^{++}$  assignment. This is illustrated in figure 3.42 which shows that both  $1^{++}$  and  $2^{-+}$  fits describe the  $\cos\theta_X$  distribution equally all. In fact, the  $1^{++}$  and  $2^{-+}$  with  $\alpha = 0.5$  produce the same  $\cos\theta_X$  distribution which is also apparent from the EVTGEN MC samples - figure 3.43. For that reason, if the data are  $1^{++}$  (as our multidimensional analyses indicate) then  $2^{-+}$

Table 3.13: The results from  $1^{++}/2^{-+}$  hypotheses testing on the  $X(3872)$  2011 data and on the EVTGEN MC using the likelihood ratio test at different dimensions (5D represents the nominal analysis). The MC results have been scaled to the statistics and background in the data (except for the error on  $\alpha$ ).

Analyzed (generated $\alpha$ )	$\Delta\chi_L^2$	$n_\sigma$	$\alpha$ from $2^{-+}$ fit
5D $\cos\theta_X, \cos\theta_{J/\psi}, \cos\theta_{\pi\pi}, \phi_X - \phi_{J/\psi}, \phi_X - \phi_{\pi\pi}$			
$2^{-+}$ MC (1.00, 0.00)	-127.1	11.04 $\sigma$	(0.970 $\pm$ 0.018, 0.143 $\pm$ 0.036)
$2^{-+}$ MC (0.68, 0.32)	-109.4	10.21 $\sigma$	(0.667 $\pm$ 0.014, 0.333 $\pm$ 0.013)
$2^{-+}$ MC (0.50, 0.00)	-150.0	12.02 $\sigma$	(0.498 $\pm$ 0.007, 0.026 $\pm$ 0.036)
$1^{++}$ MC	+112.3	10.35 $\sigma$	(0.650 $\pm$ 0.011, 0.294 $\pm$ 0.012)
2011 data	+99.4	9.72 $\sigma$	(0.671 $\pm$ 0.046, 0.280 $\pm$ 0.046)
3D $\cos\theta_X, \cos\theta_{\pi\pi}, \phi_X - \phi_{\pi\pi}$			
$2^{-+}$ MC (1.00, 0.00)	-90.4	9.24 $\sigma$	(0.988 $\pm$ 0.023, 0.084 $\pm$ 0.094)
$2^{-+}$ MC (0.68, 0.32)	-60.4	7.48 $\sigma$	(0.660 $\pm$ 0.016, 0.337 $\pm$ 0.014)
$2^{-+}$ MC (0.50, 0.00)	-44.8	6.37 $\sigma$	(0.500 $\pm$ 0.008, 0.047 $\pm$ 0.025)
$1^{++}$ MC	+62.2	7.60 $\sigma$	(0.600 $\pm$ 0.011, 0.217 $\pm$ 0.012)
211 data	+56.3	7.20 $\sigma$	(0.638 $\pm$ 0.044, 0.209 $\pm$ 0.047)
3D $\cos\theta_X, \cos\theta_{J/\psi}, \phi_X - \phi_{J/\psi}$			
$2^{-+}$ MC (1.00, 0.00)	-58.1	7.32 $\sigma$	(1.081 $\pm$ 0.047, -0.277 $\pm$ 0.056)
$2^{-+}$ MC (0.68, 0.32)	-24.7	4.60 $\sigma$	(0.616 $\pm$ 0.025, 0.456 $\pm$ 0.033)
$2^{-+}$ MC (0.50, 0.00)	-38.7	5.88 $\sigma$	(0.430 $\pm$ 0.012, -0.123 $\pm$ 0.044)
$1^{++}$ MC	+2.5	1.07 $\sigma$	(0.487 $\pm$ 0.035, 0.758 $\pm$ 0.049)
2011 data	+5.1	1.76 $\sigma$	(0.477 $\pm$ 0.102, 0.593 $\pm$ 0.139)
2D $\cos\theta_X, \cos\theta_{\pi\pi}$			
$2^{-+}$ MC (1.00, 0.00)	-43.3	6.26 $\sigma$	(1.000 $\pm$ 0.036, 0.000 $\pm$ 0.208)
$2^{-+}$ MC (0.68, 0.32)	-21.4	4.24 $\sigma$	(0.668 $\pm$ 0.018, 0.363 $\pm$ 0.033)
$2^{-+}$ MC (0.50, 0.00)	-33.3	5.42 $\sigma$	(0.481 $\pm$ 0.012, 0.145 $\pm$ 0.028)
$1^{++}$ MC	+37.5	5.79 $\sigma$	(0.622 $\pm$ 0.025, 0.527 $\pm$ 0.039)
2011 data	+34.1	5.50 $\sigma$	(0.679 $\pm$ 0.112, 0.591 $\pm$ 0.168)
2D $\cos\theta_X, \phi_X - \phi_{\pi\pi}$			
$2^{-+}$ MC (1.00, 0.00)	-36.0	5.66 $\sigma$	(0.981 $\pm$ 0.031, 0.114 $\pm$ 0.087)
$2^{-+}$ MC (0.68, 0.32)	-6.6	2.09 $\sigma$	(0.653 $\pm$ 0.021, 0.361 $\pm$ 0.019)
$2^{-+}$ MC (0.50, 0.00)	-9.7	2.66 $\sigma$	(0.477 $\pm$ 0.009, 0.001 $\pm$ 0.084)
$1^{++}$ MC	+1.3	0.66 $\sigma$	(0.510 $\pm$ 0.015, 0.309 $\pm$ 0.020)
2011 data	-0.3	0.57 $\sigma$	(0.524 $\pm$ 0.059, 0.323 $\pm$ 0.073)
1D $\cos\theta_X$			
$2^{-+}$ MC (1.00, 0.00)	-33.9	5.82 $\sigma$	(1.206 $\pm$ 0.071, 0)
$2^{-+}$ MC (0.68, 0.32)	-5.4	2.32 $\sigma$	(0.713 $\pm$ 0.030, 0)
$2^{-+}$ MC (0.50, 0.00)	-0.2	0.50 $\sigma$	(0.458 $\pm$ 0.022, 0)
$1^{++}$ MC	-0.1	0.00 $\sigma$	(0.526 $\pm$ 0.022, 0)
2011 data	-0.2	0.00 $\sigma$	(0.538 $\pm$ 0.084, 0)



hypothesis with arbitrary  $\alpha$  cannot be ruled out. Our conclusions from the 1D study are the same as previously reached in the 1D simulations by [24] using the other discrimination method than the likelihood ratio test employed here, and later confirmed on the 2011 data [53].

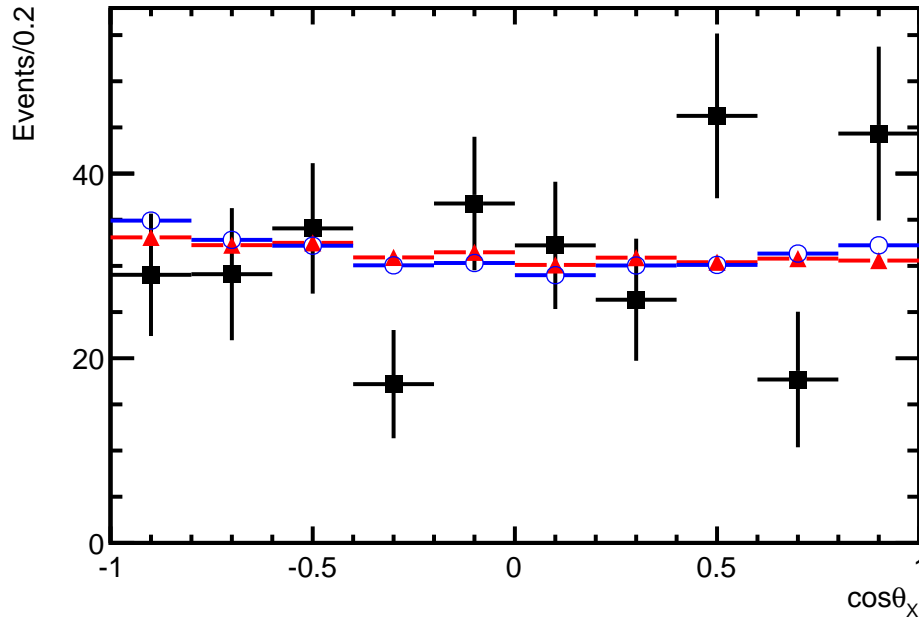


Figure 3.42: Distribution of  $\cos\theta_X$  for the data (black points with squares) and for the  $1^{++}$  (red points with triangles) and  $2^{-+}$  (blue points with open circles) 1D fits to the unbinned  $\cos\theta_X$  data.

The pattern of sensitivities uncovered in this section is roughly consistent with what can be deduced from the 1D and 2D projections of the 5D fit in comparison to the data - see Sec. 3.5.4 and table 3.5.

The Belle collaboration performed more advanced 1D studies in which they performed three 1D dimensional fits to  $\cos\theta_X$  and two other angles in transversity formalism. They concluded that they could not separate  $1^{++}$  and  $2^{-+}$  in such approach either [21]. Therefore, testing multidimensional correlations is a crucial step in gaining discrimination power between these two hypotheses. CDF did perform analysis in many dimensions [20], but because of the lack of  $X(3872)$  polarization they did not have access to the information provided to us via  $\cos\theta_X$  vs.  $\cos\theta_{\pi\pi}$  correlation. They fitted the correlations between the  $X(3872)$  and  $J/\psi$

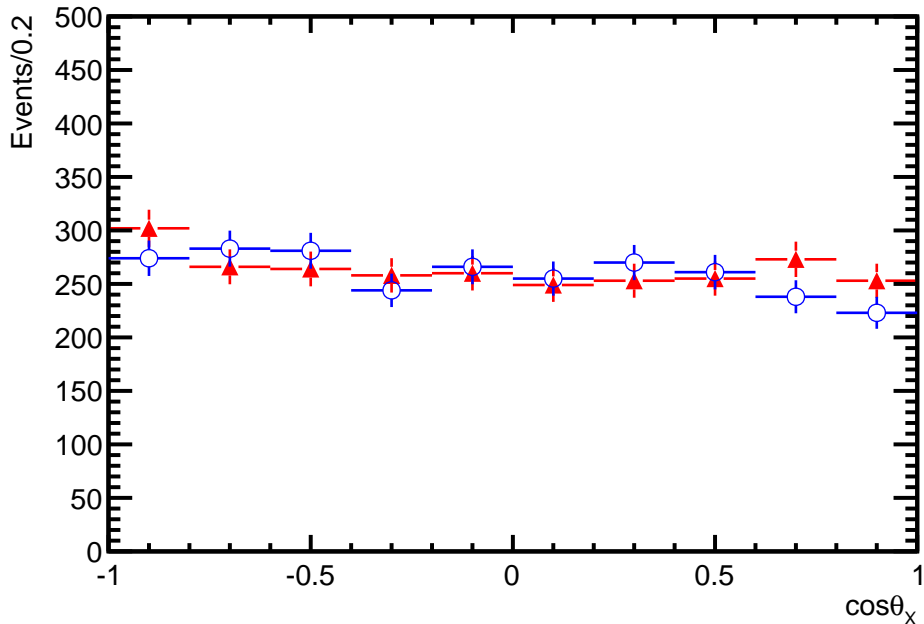


Figure 3.43: Distribution of  $\cos\theta_X$  for the  $1^{++}$  EVTGEN MC (red points with triangles) and for the  $2^{-+}$   $\alpha = (0.5, 0.0)$  EVTGEN MC (blue points with open circles).

decay angles. As our 3D results show, even with polarized  $X(3872)$  these correlations don't have great discrimination power between  $2^{-+}$  and  $1^{++}$  assignments.

Additional conclusion from this study is that the results on the data closely follow the predictions from the analysis of  $1^{++}$  EVTGEN MC sample. The actual  $2^{-+}$  rejection levels are consistent with the pattern predicted by the simulations. Also values of  $\alpha$  obtained in the  $2^{-+}$  fits in different dimensions agree between the data and  $1^{++}$  MC.

### 3.12 Summary

We have performed spin analysis of  $B^+ \rightarrow X(3872)K^+$ ,  $X(3872) \rightarrow \pi^+\pi^-J/\psi$ ,  $J/\psi \rightarrow \mu^+\mu^-$  signal observed with the LHCb detector in  $1 \text{ fb}^{-1}$  of pp collision data to distinguish  $X(3872)$   $J^{PC}$  of  $1^{++}$  and  $2^{-+}$  - the only two spin assignments which have not been ruled out by the other experiments. We have employed a method which is guaranteed by statistics to be the most powerful way to discriminate between the two hypotheses; namely unbinned

likelihood ratio test using full angular phase-space. The spin  $1^{++}$  hypothesis is preferred with overwhelming significance. Calibrating the log-likelihood difference on simulated  $2^{-+}$  experiments and using the Gaussian approximation to the shape of the log-likelihood-ratio distribution we quantify  $2^{-+}$  hypothesis rejection at  $8.4\sigma$ . Since the Gaussian has a more pronounced tail towards the data point than the simulated distribution, this estimate is conservative. Without use of the Gaussian for the extrapolation we exclude the  $2^{-+}$  hypothesis at  $> 5.0\sigma$  level.

The consistency of the data with the simulation of  $1^{++}$  hypothesis is excellent in all respect. The confidence level of the sum of the log of the ratio of  $2^{-+}$  and of  $1^{++}$  likelihoods is 34%. Also shape of event-by-event distribution of the log of the ratio of  $2^{-+}$  and of  $1^{++}$  likelihoods is consistent with the  $1^{++}$  hypothesis with a confidence level of 21% as obtained via  $\chi^2$  method applied to the binned data. The confidence levels between the data and the  $1^{++}$  fit projections evaluated using the  $\chi^2$  statistics applied to binned projections of the data and of the fits onto 5 angles parameterization the phase-space, and their 2–dimensional combinations are good. The data match the expectations from the simulations of the  $1^{++}$  data for the  $2^{-+}$  rejection power in the likelihood ratio tests in the nominal 5–dimensional analysis, as well as in 3–, 2– and 1–dimensional tests tried as cross-checks. The agreement extends to the results from the Wilks’ theorem applied to the  $\chi^2$  differences between the two hypotheses on the binned 1– and 2–dimensional projections of the data and of the fits. Some other disfavored spin assignments have been also tested in the full-phase approach and yielded the expected rejections, which are even higher than for the  $2^{-+}$  hypothesis. The fitted values of the ratio of the two  $LS$  amplitudes present in the  $2^{-+}$  matrix element also agrees with the  $1^{++}$  simulations in all variants of the analysis.

The simulations of efficiency have been further cross-checks by applying the likelihood analysis to the high statistics signal  $B^+ \rightarrow \psi(2S)K^+$ ,  $\psi(2S) \rightarrow \pi^+\pi^-J/\psi$ ,  $J/\psi \rightarrow \mu^+\mu^-$  selected with the same cuts. This decay is kinematically similar to the  $X(3872)$  signal. The right  $J^{PC}$  is

preferred ( $1^{--}$ ) when tested against several other spin assignments by margins which are in good agreement with the  $1^{--}$  simulation of the  $\psi(2S)$  signal. The confidence levels obtained from the  $\chi^2$  values from the binned 1- and 2-dimensional projections of the data and the  $1^{--}$  simulations are acceptable. We have eliminated use of the simulations for the efficiency corrections in the likelihood analysis of the  $X(3872)$  signal altogether by reweighting the  $\psi(2S)$  events to the phase-space distribution and using it to integrate the angular matrix elements squared for the  $X(3872)$  with the detector efficiency. We also varied the  $X(3872)$  efficiency by a factor of 2 by changing the  $p_T$  requirement for the charged pions, and maintained good rejection of the  $2^{-+}$  hypothesis at the levels following predictions of the  $1^{++}$  simulations.

We have tested immunity of the results to the background level in the sample by varying the cut on the  $Q$  value in the  $X(3872) \rightarrow \pi^+\pi^- J/\psi$  decay (i.e. cut on the  $M(\pi^+\pi^-)$  mass) and observed stable rejection levels in agreement with the  $1^{++}$  simulations, while the background was varied by a factor of four.

The likelihood analysis performed on the  $1.3 \text{ fb}^{-1}$  of the 2012 data, yields very similar results as on the 2011 data sample.

# Chapter 4

## Conclusion

The  $1^{++}$  assignment to the  $X(3872)$  state has been accomplished beyond any doubt. This assignment rules out the explanation of the  $X(3872)$  state as a conventional  $\eta_{c2}(1^1D_2)$  state. Among remaining possibilities are  $\chi_{c1}(2^3P_1)$  charmonium, strongly disfavored by the  $X(3872)$  mass, and several exotic models like a  $D^0\bar{D}^{0*}$  molecule or tetraquark state.



## References

- [1] Georges Aad et al. “Observation of a New Particle in the Search for the Standard Model Higgs boson with the ATLAS Detector at the LHC”. In: *Physics Letters B* 716.1 (2012), pp. 1–29 (cit. on pp. 1, 6).
- [2] Serguei Chatrchyan et al. “Observation of a New boson at a Mass of 125 GeV with the CMS Experiment at the LHC”. In: *Physics Letters B* 716.1 (2012), pp. 30–61 (cit. on pp. 1, 6).
- [3] Wikipedia. URL: [http://en.wikipedia.org/wiki/Standard\\_Model](http://en.wikipedia.org/wiki/Standard_Model) (cit. on pp. 2, 6).
- [4] Wikipedia. URL: <http://en.wikipedia.org/wiki/Hadron> (cit. on p. 4).
- [5] Jean-Jacques Aubert et al. “Experimental Observation of a Heavy Particle  $J$ ”. In: *Physical Review Letters* 33.23 (1974), p. 1404 (cit. on p. 7).
- [6] J-E Augustin et al. “Discovery of a Narrow Resonance in  $e^+ e^-$  Annihilation”. In: *Physical Review Letters* 33.23 (1974), p. 1406 (cit. on p. 7).
- [7] GS Abrams et al. “Discovery of a Second Narrow Resonance in  $e^+ e^-$  Annihilation”. In: *Physical Review Letters* 33.24 (1974), p. 1453 (cit. on p. 7).
- [8] N Brambilla et al. “Heavy Quarkonium Physics”. In: *arXiv preprint hep-ph/0412158* (2004) (cit. on pp. 7, 11).
- [9] S-K Choi et al. “Observation of a Narrow Charmoniumlike State in Exclusive  $B^\pm \rightarrow K^\pm \pi^+ \pi^- J/\psi$  Decays”. In: *Physical Review Letters* 91.26 (2003), p. 262001 (cit. on

- p. 9).
- [10] D Acosta et al. “Observation of the Narrow State  $X(3872) \rightarrow J/\psi\pi^+\pi^-$  in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.96$  T e V”. In: *Physical Review Letters* 93.7 (2004), p. 072001 (cit. on p. 9).
- [11] VM Abazov et al. “Observation and Properties of the  $X(3872)$  Decaying to  $J/\psi\pi^+\pi^-$  in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.96$  T e V”. In: *Physical Review Letters* 93.16 (2004), p. 162002 (cit. on p. 9).
- [12] Bernard Aubert et al. “Observation of the Decay  $B \rightarrow J/\psi\eta K$  and Search for  $X(3872) \rightarrow J/\psi\eta$ ”. In: *Physical Review Letters* 93.4 (2004), p. 041801 (cit. on p. 9).
- [13] R Aaij et al. “Observation of  $X(3872)$  Production in  $pp$  Collisions at  $\sqrt{s} = 7$  TeV”. In: *The European Physical Journal C* 72.5 (2012), pp. 1–9 (cit. on p. 9).
- [14] Juerg Beringer et al. “Review of Particle Physics”. In: *Physical Review D* 86.1 (2012) (cit. on pp. 9, 11).
- [15] R Aaij et al. “Precision Measurement of D Meson Mass Differences”. In: *Journal of High Energy Physics* 6 (2013) (cit. on p. 12).
- [16] Nils A Törnqvist. “Isospin Breaking of the Narrow Charmonium State of Belle at 3872 MeV as a Deuson”. In: *Physics Letters B* 590.3 (2004), pp. 209–215 (cit. on p. 12).
- [17] Luciano Maiani et al. “Diquark-antidiquark states with hidden or open charm and the nature of X (3872)”. In: *Physical Review D* 71.1 (2005), p. 014028 (cit. on p. 12).
- [18] D Ebert, RN Faustov, and VO Galkin. “Excited Heavy Tetraquarks with Hidden Charm”. In: *The European Physical Journal C-Particles and Fields* 58.3 (2008), pp. 399–405 (cit. on p. 12).
- [19] K Abe et al. “Evidence for  $X(3872) \rightarrow \gamma J/\psi$  and the Sub-threshold Decay  $X(3872) \rightarrow \omega J/\psi$ ”. In: *arXiv preprint hep-ex/0505037* (2005) (cit. on p. 12).
- [20] A Abulencia et al. “Analysis of the Quantum Numbers  $J^{PC}$  of the  $X(3872)$  Particle”. In: *Physical Review Letters* 98.13 (2007), p. 132002 (cit. on pp. 12, 82, 90, 133).
- [21] S-K Choi et al. “Bounds on the Width, Mass Difference and other Properties of  $X(3872) \rightarrow \pi^+\pi^-J/\psi$  Decays”. In: *Physical Review D* 84.5 (2011), p. 052004 (cit. on



- pp. 13, 96, 133).
- [22] Pablo del Amo Sanchez et al. “Evidence for the Decay  $X(3872) \rightarrow J/\psi\omega$ ”. In: *Physical Review D* 82.1 (2010), p. 011101 (cit. on p. 15).
- [23] N. Mangiafave, J. Dickens, and V. Gibson. “A Study of the Angular Properties of the  $X(3872) \rightarrow J\psi\pi^+\pi^-$  Decay”. In: *LHCb-PUB-2010-003* (2010) (cit. on pp. 15, 82, 130).
- [24] Nicola Mangiafave. “Measurements of Charmonia Production and a Study of the  $X(3872)$  at LHCb”. PhD thesis. 2012 (cit. on pp. 15, 82, 94, 130, 133).
- [25] Lyndon Evans and Philip Bryant. “LHC Machine”. In: *Journal of Instrumentation* 3.08 (2008), S08001 (cit. on p. 21).
- [26] A Augusto Alves Jr et al. “The LHCb Detector at the LHC”. In: *Journal of Instrumentation* 3.08 (2008), S08005 (cit. on pp. 24, 27, 30, 33, 38, 46, 47, 49, 50, 53, 65).
- [27] M. Adamus et al. “LHCb Magnet Technical Design Report”. In: *CERN/LHCC-2000-007* (2000) (cit. on p. 24).
- [28] R Aaij et al. “Performance of the LHCb Vertex Locator”. In: *LHCb-DP-2014-001* (2014) (cit. on p. 28).
- [29] M. Adamus et al. “LHCb Technical Design Report”. In: *CERN/LHCC-2003-030* (2003) (cit. on pp. 30, 33).
- [30] M. Adamus et al. “LHCb Inner Tracker Technical Design Report”. In: *CERN/LHCC-2002-029* (2002) (cit. on p. 35).
- [31] M. Adamus et al. “LHCb Outer Tracker Technical Design Report”. In: *CERN/LHCC-2001-024* (2001) (cit. on p. 37).
- [32] R Arink et al. “Performance of the LHCb Outer Tracker”. In: *LHCb-DP-2013-003* (2013) (cit. on p. 37).
- [33] M. Adamus et al. “LHCb RICH Technical Design Report”. In: *CERN/LHCC-2000-037* (2000) (cit. on p. 41).

- [34] A Borgia et al. “The Magnetic Distortion Calibration System of the LHCb RICH1 Detector”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 735 (2014), pp. 44–52 (cit. on p. 43).
- [35] M Adinolfi et al. “Performance of the LHCb RICH Detector at the LHC”. In: *The European Physical Journal C* 73.5 (2013), pp. 1–17 (cit. on pp. 44, 70).
- [36] M. Adamus et al. “LHCb Calorimeters Technical Design Report”. In: *CERN/LHCC-2000-036* (2000) (cit. on p. 48).
- [37] M. Adamus et al. “LHCb Muon System Technical Design Report”. In: *CERN/LHCC-2001-010* (2001) (cit. on p. 55).
- [38] A Augusto Alves Jr et al. “Performance of the LHCb Muon System”. In: *Journal of Instrumentation* 8.02 (2013), P02022 (cit. on p. 55).
- [39] Gaia Lanfranchi. “Time Resolution and Aging Properties of the MWPCs for the LHCb Muon System”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 535.1 (2004), pp. 221–225 (cit. on p. 56).
- [40] G Bencivenni et al. “A triple GEM detector with pad readout for high rate charged particle triggering”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 488.3 (2002), pp. 493–502 (cit. on p. 57).
- [41] M. Adamus et al. “LHCb Trigger System Technical Design Report”. In: *CERN/LHCC-2003-031* (2003) (cit. on p. 59).
- [42] R Aaij et al. “The LHCb trigger and its performance in 2011”. In: *Journal of Instrumentation* 8.04 (2013), P04022 (cit. on p. 60).
- [43] Gloria Corti et al. “Software for the LHCb experiment”. In: *Nuclear Science, IEEE Transactions on* 53.3 (2006), pp. 1323–1328 (cit. on p. 64).

- [44] Mike Williams, Ulrik Egede, Stuart Paterson, et al. “Distributed Analysis at LHCb”. In: *Journal of Physics: Conference Series*. Vol. 331. 7. IOP Publishing. 2011, p. 072037 (cit. on p. 65).
- [45] R Aaij et al. “First Observation of the Decay  $B_c^+ \rightarrow J/\psi \pi^+ \pi^- \pi^+$ ”. In: *Physical review letters* 108.25 (2012), p. 251802 (cit. on p. 69).
- [46] R Aaij et al. “Search for the  $X(4140)$  State in  $B^+ \rightarrow J/\psi \phi K^+$  Decays”. In: *Physical Review D* 85.9 (2012), p. 091103 (cit. on p. 69).
- [47] B. Gui, T. Skwarnicki, and J. C. Wang. “Analysis of  $B_{c,u}$  Decays to  $J/\psi$  and Three Charged Pions or Kaons”. In: *LHCb-ANA-2011-054* (2011) (cit. on p. 69).
- [48] Jeffrey D Richman. “An Experimenter’s Guide to the Helicity Formalism”. In: *CALT-68-1148* (1984) (cit. on p. 82).
- [49] Joachim Heuser. “Measurement of the Mass and the Quantum Numbers  $J^{PC}$  of the  $X(3872)$  State”. PhD thesis. 2008 (cit. on p. 82).
- [50] W. T. Eadie, D. Drijard, and F. E. James. *Statistical Methods in Experimental Physics*. Amsterdam: North-Holland, 1971 (cit. on pp. 87, 88, 105).
- [51] Yuehong Xie. “sFit: A Method for Background Subtraction in Maximum Likelihood Fit”. In: *arXiv preprint arXiv:0905.0724* (2009) (cit. on p. 88).
- [52] Yuehong Xie. “Some Physics and Technical Issues in Analysis of the Decay  $B_s \rightarrow J/\psi K^+ K^-$ ”. In: *LHCb-INT-2012-017* (2012) (cit. on p. 89).
- [53] Edgar Engel and Valerie Gibson. “Update on  $X(3872)$   $J^{PC}$  measurement”. In: *Onia Meeting* (2012) (cit. on p. 133).

## Vita

Author's Name: Bin Gui

Place of Birth: Shijiazhuang, Hebei, China

Date of Birth: Sep. 14, 1984

### Degrees Awarded:

Bachelor of Science, University of Science and Technology of China, 2003

### Professional Experience:

Teaching Assistant, Syracuse University Department of Physics, 2008-2009

Research Assistant, Syracuse University Department of Physics, 2009-2014

Syracuse High Energy Experimental Physics Group 2009-2014

CERN User, Geneva, Switzerland 2009-2014