

Amplitude analysis of B^0 decays to $J/\psi\pi^-K^+$ and $\psi(2S)\pi^-K^+$

by Andy Beiter

B.S., Canisius College 2015

Dissertation

Submitted in partial fulfillment of the requirements for a degree of

Doctor of Philosophy in Physics

Syracuse University

January 2023

Approved:

Tomasz Skwarnicki
Prof. Tomasz Skwarnicki

Date:

January 27, 2023



Abstract

Amplitude analyses of $B^0 \rightarrow J/\psi(1S)\pi^-K^+$ and $B^0 \rightarrow \psi(2S)\pi^-K^+$ decays are presented. The data correspond to the full LHCb Run 1 and Run 2 samples, which are much larger than previously analyzed by the Belle and LHCb experiments in similar analyses. Many more exotic $\psi(nS)\pi^-$ states are required for a good description of the data than disclosed by the previous amplitude analyses. For the first time, significant contributions from exotic $\psi(nS)K^+$ states are observed. The $J^P = 1^+ Z(4200)^-$ state, previously established only in the $J/\psi(1S)\pi^-$ mode, is confirmed and observed with the consistent mass and width in the $\psi(2S)\pi^-$ mode. However, there are also a number of inconsistencies between the exotic states required by the two data sets which point to theoretical limitations of modelling these decays as a collection of resonances decaying only to $K^+\pi^-$, $\psi(nS)\pi^-$ or $\psi(nS)K^+$. Possible future directions in amplitude analyses of these decays are discussed.

Copyright © Andy Beiter 2023
All Rights Reserved

Graduate school has been a long journey, and I could not have finished this work without the help and support that so many have given me. I cannot thank enough my advisor, Tomasz Skwarnicki. From him, I learned far more than just the physics presented in this thesis. Predicting questions asked about this analysis before presentations, helping me shape a good narrative for interesting talks, sincerely asking about my family and my life, and having casual conversations about current events: Tomasz helped prepare me to perform a thoughtful and introspective analysis, while at the same time being a caring and compassionate mentor. I will never forget my first year in the LHCb group when Patty Whitmore told me she was happy that Tomasz was my advisor, saying that we were a good match. On my end, that has certainly been true, as I have been very lucky to have Tomasz guiding me through this analysis.

I also want to thank the professors at Syracuse University, especially those within the Syracuse LHCb group: Marina Artuso, Steve Blusk, Matt Rudolph, Ray Mountain, and, of course, the late Sheldon Stone. You have all been great sources of advice, always having your doors open for quick questions (and long questions), initiating great discussions in group meetings, and helping me hit the ground running when I first started doing research. Working in the lab and on the UT especially has taught me a lot about collaborating on a project and about problem solving and critical thinking. I am very proud of the challenges the group has overcome to ultimately complete the UT, and I am glad to have had a hand in the work done.

I'd like to thank the JPAC theorists as well, in particular Adam Szczepaniak, Mikhail Mikhasenko, Alessandro Pilloni, and Cesar Fernandez-Ramirez. The discussions of the formalisms used in amplitude analyses and of resonance representations used in this analysis were incredibly insightful and interesting.

I also greatly appreciate the work done by our collaborators Liming Zhang, Zhihong Shen, and Mengzhen Wang, as they laid the groundwork in migrating many calculations to GPUs, which tremendously increased the speed of fit executions. I really enjoyed learning how to integrate the fit program with GPU code, and I benefited a lot from their work.

Two former students of Tomasz also paved the way for my work, and I cannot thank enough Nathan Jurik and Thomas Britton. Their programs served as the starting point for my own, and algorithms they wrote were still used in the final version of my fitter. Thank you both for your work and letting me stand on the shoulders of giants.

To the faculty at Syracuse University, whether in the machine shop, in the main office, or in the halls of the Physics Building, you have all been so kind and understanding. From the cheer you all bring and the laughs we have shared over the many years, I have truly appreciated being at Syracuse University. Thank you for being great people to be around day in and day out.

Many friends have helped me so much throughout the years, through the ups and downs. I want to name especially fellow students at Syracuse University: Harris Bernstein, Aravindhan Venkateswaran, Brandon Melcher, Eric King, Avinay Bhat, Michael Wilkinson, Matt Kelsey, and Joe Shupperd. Whether it's been trivia, softball, grabbing lunch, complaining about submitting jobs to the grid, or just hanging out, you have all made for a fun experience in graduate school, and I am truly glad to have met you all. So many other people have been so great to me in the past many years while I went through this journey, and I could write another thesis detailing everyone and everything I appreciate about them. I thank you all for being great friends to me, but I do want to thank especially two great friends who have made the end of graduate school so much easier to bear: Dave Pickett and Jasmine Howell. You both have been there to pick me up at my lowest, and you both have been people I can always talk to and always have a good time with. I feel like words are not enough for you two, but thank you so much for lifting me up and helping carry me to the end.

Lastly, I would like to thank my family. My brothers: we all quote the same lines from every movie and show we've all seen, make the same cracks at each other, and all try to say the most clever joke we can in the moment. These joys I've missed for a long time, and being home at the end of this journey has reminded me of all the fun things we've done together. To my parents, you have been the best role models a son could ask for. You knew

my potential and encouraged me every step of the way to reach it. You have always been so supportive and wise, and every day I am thankful for the opportunities you both let me have. I love you both. To my grandparents, “Grumpy” and Grandma Angie, you have always been there for me, always showing support, always cheering me on, always giving me too much food. Grumpy, if I have the honor of being a grandparent one day, I cannot think of a better role model than you. You taught me well to be kind and caring to those close to me. Grandma, you were a wonderful, kind, and caring soul. I hope that I inherited even a fraction of your generosity and selflessness. I wish you could have read this, and not a day goes by that I don’t wonder what you’d say to me. I can picture you saying you’re proud of me, but one thing I do know is that I am very proud to have had you as my grandmother. I love you both, and thank you for everything.

Contents

Lists of Illustrative Material	xi
List of Tables	xi
List of Figures	xv
1 The Quark Model	1
2 Detector Description	6
2.1 The Large Hadron Collider	6
2.2 The LHCb Experiment	7
2.3 Tracking System	9
2.3.1 Vertex Locator	10
2.3.2 Dipole Magnets	11
2.3.3 Silicon Tracker	12
2.3.4 Outer Tracker	13
2.4 Particle identification	14
2.4.1 RICH	14
2.4.2 Calorimeters	16
2.4.3 Muon system	17
2.5 Trigger System	19
2.5.1 L0 Trigger	20
2.5.2 HLT	22
2.6 Data Flow	26
3 Tetraquark candidates observed in $B^0 \rightarrow \psi(nS)\pi^- K^+$ decays	28
4 Data Selection	31
4.1 Preselection	31

4.2	The multivariate final selection	33
5	Kinematic re-weighting of Monte Carlo	40
6	Amplitude Analysis	44
6.1	Fitted kinematic variables	44
6.2	The signal PDF	47
6.3	Background treatment	48
7	Helicity formalism and matrix element	61
7.1	Helicity formalism for the K^* decay chain	61
7.2	Helicity formalism for the Z decay chain	62
7.3	Helicity formalism for the Z_K decay chain	63
7.4	Amplitude dependence on the invariant masses	64
7.4.1	Single channel K-matrix	64
7.4.2	Model Independent Contribution	66
8	Significance and fit quality calculations	67
8.1	Calculation of significance of adding states	67
8.2	Multidimensional χ^2 s	68
9	K^{*0} model	70
10	K^* Only Amplitude Model	73
11	Previous $\psi(2S)K\pi$ Model on Current Dataset	81
11.1	Fits with the previous K^{*0} -only model	81
11.2	Fits with the old model including $Z(4430)^-$	85
11.3	$Z(4430)^-$ using a Model Independent Approach	90
12	Resonant Z models with $\psi(2S)$ channel	95

12.1	Models with one $Z^- \rightarrow \psi(2S)\pi^-$	95
12.2	Models with two $Z^- \rightarrow \psi(2S)\pi^-$	100
12.3	Models with three $Z^- \rightarrow \psi(2S)\pi^-$	104
12.4	Models with four $Z^- \rightarrow \psi(2S)\pi^-$	108
12.5	Models with five $Z^- \rightarrow \psi(2S)\pi^-$	113
13	Resonant Z_K models with $\psi(2S)$ channel	114
13.1	Plots of $m_{\psi(2S)K}$ with no $Z_K^+ \rightarrow \psi(2S)K^+$	114
13.2	Models with one $Z_K^+ \rightarrow \psi(2S)K^+$	116
13.3	Models with two $Z_K^+ \rightarrow \psi(2S)K^+$	122
13.4	Models with three $Z_K^+ \rightarrow \psi(2S)K^+$	132
14	Resonant Z models with J/ψ channel	133
14.1	Models with one $Z^- \rightarrow J/\psi\pi^-$	133
14.2	Models with two $Z^- \rightarrow J/\psi\pi^-$	138
14.3	Models with three $Z^- \rightarrow J/\psi\pi^-$	142
14.4	Models with four $Z^- \rightarrow J/\psi\pi^-$	146
14.5	Models with five $Z^- \rightarrow J/\psi\pi^-$	150
14.6	Models with six $Z^- \rightarrow J/\psi\pi^-$	154
14.7	Models with seven $Z^- \rightarrow J/\psi\pi^-$	158
14.8	Models with eight $Z^- \rightarrow J/\psi\pi^-$	162
14.9	Models with nine $Z^- \rightarrow J/\psi\pi^-$	166
14.10	Models with ten $Z^- \rightarrow J/\psi\pi^-$	172
15	Resonant Z_K models with J/ψ channel	173
15.1	Plots of $m_{J/\psi K}$ with no $Z_K^+ \rightarrow J/\psi K^+$	173
15.2	Models with one $Z_K^+ \rightarrow J/\psi K^+$	176
15.3	Models with two $Z_K^+ \rightarrow J/\psi K^+$	186
15.4	Fits with fixed-shape $Z_{cs}(4000)$ and $Z_{cs}(4220)$	186

16 Comparison of resonant models between $\psi(2S)$ and J/ψ channels	188
17 Model Independent Fit of 1^+ Z Wave with $\psi(2S)$ Channel	194
18 Model Independent Fit of 1^+ Z Wave with J/ψ Channel	203
19 Summary	210
A PID Corrected Monte Carlo	213
B Fit displays for Legendre moments of $\cos\theta_{K^*}$ helicity angle	215
References	238
Vita	241

List of Tables

1	Thresholds used for L0 trigger in Run 1 and Run 2. The values represent a minimum threshold for E_T or p_T in GeV and a maximum threshold for SPD hits.	21
2	Stripping campaigns and luminosity by year.	31
3	Data selection requirements.	32
4	Fit shapes used for each kinematic/event variable for each channel.	40
5	Known K^{*0} resonances decaying to $K^+\pi^-$, with their spectroscopic classification and measured properties [1].	72
6	Fit quality metrics comparing 2014 K^{*0} -only fit with 2014 model fit on current dataset.	85
7	Mass and width of the $Z(4430)^-$ compared to results from 2014 using the model from LHCb in 2014.	85
8	All resonance fit fractions compared to results from 2014 using the model from LHCb in 2014.	86
9	Fit quality metrics comparing 2014 K^{*0} -only fit with 2014 model independent fit on current dataset.	91
10	Goodness-of-fit metrics for adding the first Z with the given J^P to the $\psi(2S)$ channel.	95
11	Mass and width of the Z added to the $\psi(2S)$ channel model.	96
12	Goodness-of-fit metrics for adding a second Z with the given J^P to the $\psi(2S)$ channel.	100
13	Masses and widths of the Zs in the 2 Z model for the $\psi(2S)$ channel.	100
14	Goodness-of-fit metrics for adding a third Z with the given J^P to the $\psi(2S)$ channel.	104
15	Masses and widths of the Zs in the 3 Z model for the $\psi(2S)$ channel.	104

16	Goodness-of-fit metrics for adding a fourth Z with the given J^P to the $\psi(2S)$ channel.	108
17	Masses and widths of the Z s in the 4 Z model for the $\psi(2S)$ channel.	109
18	Masses and widths of the K^* s in the 4 Z model for the $\psi(2S)$ channel.	109
19	All resonance fit fractions in the 4 Z fit for the $\psi(2S)$ channel.	109
20	Goodness-of-fit metrics for adding a fifth Z with the given J^P to the $\psi(2S)$ channel.	113
21	Goodness-of-fit metrics for adding one Z_K with the given J^P to the $\psi(2S)$ channel.	116
22	Masses and widths of the Z s in the 1 Z_K model for the $\psi(2S)$ channel.	116
23	Masses and widths of the K^* s in the 1 Z_K model for the $\psi(2S)$ channel.	117
24	Masses and widths of the Z_K in the 1 Z_K model for the $\psi(2S)$ channel.	117
25	Goodness-of-fit metrics for adding two Z_K s with the given J^P to the $\psi(2S)$ channel.	122
26	Masses and widths of the Z s in the 2 Z_K model for the $\psi(2S)$ channel.	122
27	Masses and widths of the K^* s in the 2 Z_K model for the $\psi(2S)$ channel.	123
28	Masses and widths of the Z_K s in the 2 Z_K model for the $\psi(2S)$ channel.	123
29	All resonance fit fractions in the 2 Z_K fit for the $\psi(2S)$ channel.	124
30	Goodness-of-fit metrics for adding a third Z_K with the given J^P to the $\psi(2S)$ channel.	132
31	Goodness-of-fit metrics for adding the first Z with the given J^P to the J/ψ channel.	133
32	Mass and width of the $Z(4200)$ in the 1 Z model for the J/ψ channel.	134
33	Goodness-of-fit metrics for adding the second Z with the given J^P to the J/ψ channel.	138
34	Masses and widths of the Z s in the 2 Z model for the J/ψ channel.	138

35	Goodness-of-fit metrics for adding the third Z with the given J^P to the J/ψ channel.	142
36	Masses and widths of the Zs in the 3 Z model for the J/ψ channel.	142
37	Goodness-of-fit metrics for adding the fourth Z with the given J^P to the J/ψ channel.	146
38	Masses and widths of the Zs in the 4 Z model for the J/ψ channel.	146
39	Goodness-of-fit metrics for adding the fifth Z with the given J^P to the J/ψ channel.	150
40	Masses and widths of the Zs in the 5 Z model for the J/ψ channel.	150
41	Goodness-of-fit metrics for adding the sixth Z with the given J^P to the J/ψ channel.	154
42	Masses and widths of the Zs in the 6 Z model for the J/ψ channel.	154
43	Goodness-of-fit metrics for adding the seventh Z with the given J^P to the J/ψ channel.	159
44	Masses and widths of the Zs in the 7 Z model for the J/ψ channel.	160
45	Goodness-of-fit metrics for adding the eighth Z with the given J^P to the J/ψ channel.	162
46	Masses and widths of the Zs in the 8 Z model for the J/ψ channel.	162
47	Goodness-of-fit metrics for adding the ninth Z with the given J^P to the J/ψ channel.	166
48	Masses and widths of the Zs in the 9 Z model for the J/ψ channel.	167
49	Masses and widths of the K^* s in the 9 Z model for the J/ψ channel.	167
50	All resonance fit fractions in the 9 Z fit for the J/ψ channel.	168
51	Goodness-of-fit metrics for adding the tenth Z with the given J^P to the J/ψ channel.	172
52	Goodness-of-fit metrics for adding one Z_K with the given J^P to the J/ψ channel.	176
53	Masses and widths of the Zs in the 1 Z_K model for the J/ψ channel.	177

54	Masses and widths of the K^* s in the 1 Z_K model for the J/ψ channel.	177
55	Masses and widths of the Z_K in the 1 Z_K model for the J/ψ channel.	177
56	All resonance fit fractions in the 1 Z_K fit for the J/ψ channel.	178
57	Goodness-of-fit metrics for adding a second Z_K with the given J^P to the J/ψ channel.	186
58	Goodness-of-fit metrics comparing the 1 Z_K model presented in this analysis with the published LHCb Z_{cs} states with fixed masses and widths.	187
59	Masses and widths of the Z s in the default K-matrix models for both the $\psi(2S)$ and J/ψ channels. While we tried to associate poles between the $\psi(2S)$ and J/ψ channels with each other by their mass proximity, their widths often differ a lot. Therefore, these associations, marked by the same contribution label, should be taken with caution. Comparison to the world average values for $Z(4430)$ and $Z(4200)$ are also included [2] but should be taken with caution since our values are K-matrix poles, while the PDG values are from Breit-Wigner form.	192
60	Masses and widths of the K^* s in the default K-matrix models for both the $\psi(2S)$ and J/ψ channels compared to the world average values [2]. The comparison should be taken with some caution since our values are K-matrix poles, while the PDG values are not precisely defined.	192
61	All resonance fit fractions in the default K-matrix models in this work for both the $\psi(2S)$ and J/ψ channels compared to the fit fractions from the previously published Breit-Wigner models.	193
62	Fit quality metrics comparing default K-matrix Z model fit with model independent Z fit on current dataset.	195
63	Masses and widths of the K^* s and $J^P = 2^-$ Z in the model independent fit for the $\psi(2S)$ channel.	196
64	All resonance fit fractions in the model independent Z fit for the $\psi(2S)$ channel.	196

65	Fit quality metrics comparing 9 Z K-matrix model fit with model independent Z fit on current dataset.	204
66	Masses and widths of the K^* s and Zs (not 1^+) in the model independent fit for the J/ψ channel.	205
67	All resonance fit fractions in the model independent Z fit for the J/ψ channel.	205
68	χ^2 s for default resonant Z model fits for uncorrected and PID-corrected $\psi(2S)$ MC.	213
69	Fit results for floating masses and widths in default resonant Z model fits for uncorrected and PID-corrected $\psi(2S)$ MC.	214
70	Fit fractions for default resonant Z model fits for uncorrected and PID-corrected $\psi(2S)$ MC.	214

List of Figures

1	The Standard Model. [3]	3
2	Arrangement of CERN's subaccelerators and experiments. [4]	7
3	A cross section of the LHCb detector [5]	8
4	Production angles for $b\bar{b}$ pairs with respect to the beam line. The LHCb acceptance is in red.	8
5	The components of the LHCb Tracking System. Different types of tracks are shown and distinguished by what subdetector components see them. [6] . . .	10
6	Depictions of the VELO detector.	11
7	Depiction of the LHCb dipole magnet with dimensions in mm.	12
8	Layers within the different ST stations.	13
9	The three OT stations (light blue) and the components of the ST (purple). .	14
10	RICH detector measurements of opening angles versus particle momenta. . .	15
11	RICH detector layouts.	16

12	The segmentation of one-quarter of the SPD/PS and ECAL (left) and the HCAL (right).	17
13	Layout of the muon system.	18
14	LHCb trigger schema for Run 1 and Run 2.	20
15	Process through which the HLT1 performs track and vertex reconstruction. .	24
16	Process through which the HLT2 performs track and vertex reconstruction. .	25
17	Flowchart depicting the data flow within LHCb for Run 2.	27
18	Fraction of the $B \rightarrow J/\psi K\pi$ signal events passing a cut $DLL < x$ for the data (black), the signal training sample (blue), and fraction of the background events (B sideband) passing this cut (red).	34
19	Fraction of the $B \rightarrow J/\psi K\pi$ signal events passing a cut $DLL < x$ for the data (black) and fraction of the signal MC events passing this cut (blue).	35
20	Fraction of the $B \rightarrow \psi(2S)K\pi$ signal events passing a cut $DLL < x$ for the data (black) and fraction of the background events (B sideband) passing this cut (red).	36
21	Fraction of the $B \rightarrow \psi(2S)K\pi$ signal events passing a cut $DLL < x$ for the data (black) and fraction of the signal MC events passing this cut (blue).	36
22	Mass distribution of the $B \rightarrow \psi(2S)\pi K$ candidates in Run 1 and Run 2 data. The fit (blue line) of a double-sided asymmetric Crystal Ball signal shape and quadratic background (green line) are superimposed. The fit results are $m(B^0) = 5280.2 \pm 0.02$ MeV, $\sigma(B^0) = 5.22 \pm 0.03$ MeV, $m(B_s^0) = 5367.7 \pm 0.25$ MeV, $\sigma(B_s^0) = 5.01 \pm 0.24$ MeV, $\alpha_{left} = 1.56 \pm 0.03$, $\alpha_{right} = 1.43 \pm 0.04$, $N_{sig} = 140,344 \pm 435$ events, $N_{B_s^0} = 1,538 \pm 72$ events, and $N_{bkg} = 2232 \pm 277$ events. Vertical red lines show the signal region used in the default amplitude fit. Smaller vertical black lines indicate boundaries of the sidebands used for the background parameterization in the fit. A log scale is used to better show the B_s^0 peak.	38

23	Mass distribution of the $B \rightarrow J/\psi\pi K$ candidates in Run 1 and Run 2 data. The fit (blue line) of a double-sided asymmetric Crystal Ball signal shape and quadratic background (green line) are superimposed. The fit results are $m(B^0) = 5280.3 \pm 0.007$ MeV, $\sigma = 7.20 \pm 0.009$ MeV, $m(B_s^0) = 5367.6 \pm 0.10$ MeV, $\sigma(B_s^0) = 6.49 \pm 0.10$ MeV, $\alpha_{left} = 1.45 \pm 0.008$, $\alpha_{right} = 1.44 \pm 0.009$, $N_{sig} = 2,278,460 \pm 1,786$ events, $N_{B_s^0} = 20,597 \pm 341$ events, and $N_{bkg} = 36,606 \pm 131$ events. Vertical red lines show the signal region used in the default amplitude fit. Smaller vertical black lines indicate boundaries of the sidebands used for the background parameterization in the fit. A log scale is also shown to better show the B_s^0 peak.	39
24	Plots of the fits (in red) used to re-weight $\psi(2S)$ MC. Weights are applied in order (left to right, top to bottom): the number of tracks, the transverse momentum of B, the momentum of the kaon, and the momentum of the pion.	41
25	Comparison of MC (red) and data (black) for the 4 kinematic variables used to re-weight MC after complete re-weighting for $\psi(2S)$	42
26	Plots of the fits (in red) used to re-weight J/ψ MC. Weights are applied in order (left to right, top to bottom): the number of tracks, the transverse momentum of B, the momentum of the kaon, and the momentum of the pion.	42
27	Comparison of MC (red) and data (black) for the 4 kinematic variables used to re-weight MC after complete re-weighting for J/ψ	43
28	Definition of the helicity angles. See text for details.	45
29	Parameterized efficiency (ϵ_1) in the rectangular Dalitz plane ($m_{K\pi}, \cos\theta_{K^*}$) (top) and in the traditional Dalitz plane ($m_{K\pi}^2, m_{J/\psi\pi}^2$) (bottom) for the $\psi(2S)$ channel. The normalization arbitrarily corresponds to an average efficiency of 1 over the phase-space.	51
30	Parameterized efficiency function $\epsilon_2(\cos\theta_\psi - m_{K\pi})$ for the $\psi(2S)$ channel. By construction it integrates to 1.0 at each $m_{K\pi}$ value.	52

31	Parameterized efficiency function $\epsilon_3(\phi—m_{K\pi})$ for the $\psi(2S)$ channel. By construction it integrates to 1.0 at each $m_{K\pi}$ value.	52
32	Parameterized background distributions in the rectangular Dalitz plane $(m_{K\pi}, \cos\theta_{K^*})$ (top) and in the traditional Dalitz plane $(m_{K\pi}^2, m_{J/\psi\pi}^2)$ (bottom) for the $\psi(2S)$ channel. The normalization arbitrarily corresponds to an average efficiency of 1 over the phase-space.	53
33	Parameterized background distribution for $(\cos\theta_\psi m_{K\pi})$ for the $\psi(2S)$ channel. By construction it integrates to 1.0 at each $m_{K\pi}$ value.	54
34	Parameterized background distribution for $(\phi_\psi m_{K\pi})$ for the $\psi(2S)$ channel. By construction it integrates to 1.0 at each $m_{K\pi}$ value.	54
35	Comparison of the background from the data (black) and the parameterized background (red) for $m_{K\pi}$, $\cos\theta_{K^*}$, $\cos\theta_\psi$, and ϕ_ψ for the $\psi(2S)$ channel.	55
36	Parameterized efficiency (ϵ_1) in the rectangular Dalitz plane $(m_{K\pi}, \cos\theta_{K^*})$ (top) and in the traditional Dalitz plane $(m_{K\pi}^2, m_{J/\psi\pi}^2)$ (bottom) for the J/ψ channel. The normalization arbitrarily corresponds to an average efficiency of 1 over the phase-space.	56
37	Parameterized efficiency function $\epsilon_2(\cos\theta_\psi—m_{K\pi})$ for the J/ψ channel. By construction it integrates to 1.0 at each $m_{K\pi}$ value.	57
38	Parameterized efficiency function $\epsilon_3(\phi—m_{K\pi})$ for the J/ψ channel. By construction it integrates to 1.0 at each $m_{K\pi}$ value.	57
39	Parameterized background distributions in the rectangular Dalitz plane $(m_{K\pi}, \cos\theta_{K^*})$ (top) and in the traditional Dalitz plane $(m_{K\pi}^2, m_{J/\psi\pi}^2)$ (bottom) for the J/ψ channel. The normalization arbitrarily corresponds to an average efficiency of 1 over the phase-space.	58
40	Parameterized background distribution for $(\cos\theta_\psi m_{K\pi})$ for the J/ψ channel. By construction it integrates to 1.0 at each $m_{K\pi}$ value.	59

41	Parameterized background distribution for $(\phi_\psi m_{K\pi})$ for the J/ψ channel. By construction it integrates to 1.0 at each $m_{K\pi}$ value.	59
42	Comparison of the background from the data (black) and the parameterized background (red) for $m_{K\pi}$, $\cos\theta_{K^*}$, $\cos\theta_\psi$, and ϕ_ψ for the J/ψ channel.	60
43	The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) using a model with only the K^{*0} s listed in PDG. Masses and widths of the states are fixed to the PDG values except for those mentioned in Sec. 9. The widths are fixed to the PDG values or left free for the predicted states.	75
44	Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the K^{*0} -only 4D amplitude fit (red) onto the $m_{\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), at $K^*(892)$ (top right), in between $K^*(892)$ and $K_2^*(1430)$ (middle left), at $K_2^*(1430)$ (middle right), and above $K_2^*(1430)$ (bottom).	76
45	Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the K^{*0} -only 4D amplitude fit (red) onto the angular variables of the fit: cosine of the K^{*0} helicity angle (top), cosine of the ψ helicity angle (middle), and the angle between the K^{*0} and ψ decay planes (ϕ).	77
46	The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi K\pi$ data (black) and the fit (red) using a model with only the K^{*0} s listed in PDG. Masses and widths of the states are fixed to the PDG values except for those mentioned in Sec. 9. The widths are fixed to the PDG values or left free for the predicted states.	78
47	Projections of the $B^0 \rightarrow J/\psi K\pi$ data (black points) and of the K^{*0} -only 4D amplitude fit (red) onto the $m_{\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), at $K^*(892)$ (top right), in between $K^*(892)$ and $K_2^*(1430)$ (middle left), at $K_2^*(1430)$ (middle right), above $K_2^*(1430)$ but below the $m_{K\pi}$ phase space limit in $B^0 \rightarrow \psi(2S)K\pi$ decays (bottom left), and above the $m_{K\pi}$ phase space limit in $B^0 \rightarrow \psi(2S)K\pi$ decays (bottom right).	79

48	Projections of the $B^0 \rightarrow J/\psi K\pi$ data (black points) and of the K^{*0} -only 4D amplitude fit (red) onto the angular variables of the fit: cosine of the K^{*0} helicity angle (top), cosine of the ψ helicity angle (middle), and the angle between the K^{*0} and ψ decay planes (ϕ).	80
49	The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) using the 2014 LHCb model with only K^* contributions.	82
50	The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) using the 2014 LHCb model with only K^* contributions.	83
51	Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the 2014 LHCb model 4D amplitude fit (red) with only K^* contributions onto the $m_{\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), at $K^*(892)$ (top right), in between $K^*(892)$ and $K_2^*(1430)$ (middle left), at $K_2^*(1430)$ (middle right), and above $K_2^*(1430)$ (bottom).	84
52	The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) using the 2014 LHCb model with $Z(4430)^-$	86
53	The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) using the 2014 LHCb model with $Z(4430)^-$	87
54	Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the 2014 LHCb model 4D amplitude fit (red) with $Z(4430)^-$ onto the $m_{\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), at $K^*(892)$ (top right), in between $K^*(892)$ and $K_2^*(1430)$ (middle left), at $K_2^*(1430)$ (middle right), and above $K_2^*(1430)$ (bottom).	88
55	Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the 2014 LHCb model 4D amplitude fit (red) with $Z(4430)^-$ onto the angular variables of the fit: cosine of the K^{*0} helicity angle (top), cosine of the ψ helicity angle (middle), and the angle between the K^{*0} and ψ decay planes (ϕ).	89
56	The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) using the 2014 LHCb model independent approach.	91

57	The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) using the 2014 LHCb model independent approach.	92
58	Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the 2014 LHCb model 4D amplitude fit (red) with $Z(4430)^-$ as model independent contribution onto the $m_{\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top) and in between $K^*(892)$ and $K_2^*(1430)$ (bottom).	93
59	Argand plot for model independent Z contribution from 2014 analysis [7] (top) and current analysis (bottom). The red circle represents a Breit-Wigner for the $Z(4430)^-$. Each complex amplitude is labeled with its $m_{\psi\pi}$ bin center in the Argand plot from the current analysis.	94
60	The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) which adds a $J^P = 1^+$ Z.	97
61	The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) which adds a $J^P = 1^+$ Z.	98
62	Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the 4D amplitude fit (red) with $Z(4430)^-$ as a $J^P = 1^+$ resonance onto the $m_{\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $K_2^*(1430)$ (bottom).	99
63	The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) which adds a second $J^P = 1^+$ Z.	101
64	The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) which adds a second $J^P = 1^+$ Z near $m_{\psi(2S)\pi} = 4.1\text{GeV}$	102
65	Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the 4D amplitude fit (red) with 2 $J^P = 1^+$ Zs in a K-matrix onto the $m_{\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $K_2^*(1430)$ (bottom).	103

66	The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) which adds a $J^P = 2^-$ Z into the model with two 1^+ Zs in a K-matrix.	105
67	The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) which adds a $J^P = 2^-$ Z near $m_{\psi(2S)\pi} = 4.4\text{GeV}$ into the model with two 1^+ Zs in a K-matrix.	106
68	Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the 4D amplitude fit (red) with 2 $J^P = 1^+$ Zs in a K-matrix and 1 $J^P = 2^-$ Z as a separate resonance onto the $m_{\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $K_2^*(1430)$ (bottom).	107
69	The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) which adds a third $J^P = 1^+$ Z into the model with two 1^+ Zs in a K-matrix and a separate 2^- Z resonance.	110
70	The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) which adds a third $J^P = 1^+$ Z near $m_{\psi(2S)\pi} = 4.8\text{GeV}$ into the model with two 1^+ Zs in a K-matrix and a separate 2^- Z resonance.	111
71	Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the 4D amplitude fit (red) with 3 $J^P = 1^+$ Zs in a K-matrix and 1 $J^P = 2^-$ Z as a separate resonance onto the $m_{\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $K_2^*(1430)$ (bottom).	112
72	The distribution of $m_{\psi(2S)K}$ for the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black) and the fit (red) which has three $J^P = 1^+$ Zs in a K-matrix and a separate 2^- Z resonance.	114
73	Projections of the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$, $Z(4430)^-$, and $Z(4800)^-$ as a $J^P = 1^+$ K-matrix and $Z(4470)^-$ as a $J^P = 2^-$ resonance onto the $m_{\psi(2S)\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $K_2^*(1430)$ until the end of $\psi(2S)$ phase space (bottom).	115

- 74 The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 1^- Z_K$ into the default model which already has 4 $Z \rightarrow \psi(2S)\pi$ resonances. 117
- 75 The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 1^- Z_K$ into the default model which already has 4 $Z \rightarrow \psi(2S)\pi$ resonances. 118
- 76 Projections of the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$, $Z(4430)^-$, and $Z(4800)^-$ as a $J^P = 1^+$ K-matrix, $Z(4470)^-$ as a $J^P = 2^-$ resonance, and $Z_{cs}(4480)^-$ as a $J^P = 1^- \psi(2S)K$ resonance. onto the $m_{\psi(2S)\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $K_2^*(1430)$ until the end of $\psi(2S)$ phase space (bottom). 119
- 77 The distribution of $m_{\psi(2S)K}$ for the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 1^- Z_K$ into the default model which already has 4 $Z \rightarrow \psi(2S)\pi$ resonances. 120
- 78 Projections of the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$, $Z(4430)^-$, and $Z(4800)^-$ as a $J^P = 1^+$ K-matrix, $Z(4470)^-$ as a $J^P = 2^-$ resonance, and $Z_{cs}(4480)^-$ as a $J^P 1^+ \psi(2S)K$ resonance. onto the $m_{\psi(2S)K}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $K_2^*(1430)$ until the end of $\psi(2S)$ phase space (bottom). 121
- 79 The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+ Z_K$ into the 1 Z_K model. 125
- 80 The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+ Z_K$ into the 1 Z_K model. 126

- 81 Projections of the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$, $Z(4430)^-$, and $Z(4800)^-$ as a $J^P = 1^+$ K-matrix, $Z(4470)^-$ as a $J^P = 2^-$ resonance, $Z_{cs}(4480)^-$ as a $J^P = 1^- \psi(2S)K$ resonance, and $Z_{cs}(4520)^-$ as a $J^P = 1^+ \psi(2S)K$ resonance onto the $m_{\psi(2S)\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $K_2^*(1430)$ until the end of $\psi(2S)$ phase space (bottom). 127
- 82 The distribution of $m_{\psi(2S)K}$ for the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+ Z_K$ into the 1 Z_K model. 128
- 83 Projections of the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$, $Z(4430)^-$, and $Z(4800)^-$ as a $J^P = 1^+$ K-matrix, $Z(4470)^-$ as a $J^P = 2^-$ resonance, $Z_{cs}(4480)^-$ as a $J^P = 1^- \psi(2S)K$ resonance, and $Z_{cs}(4520)^-$ as a $J^P = 1^+ \psi(2S)K$ resonance onto the $m_{\psi(2S)K}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $K_2^*(1430)$ until the end of $\psi(2S)$ phase space (bottom). 129
- 84 Dalitz plot with fit pulls in each bin used to calculate 2-D χ^2 . White line draws phase space boundary for the $B^0 \rightarrow \psi(2S)\pi^-K^+$ decay. 130
- 85 Distributions of the pulls from the 2-D and 4-D pulls from the multidimensional χ^2 s for the default $B^0 \rightarrow \psi(2S)\pi^-K^+$ model. The distribution is fit with a Gaussian (red line). 131
- 86 The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds a $J^P = 1^+ Z$ 135
- 87 The distribution of $m_{J/\psi\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds a $J^P = 1^+ Z$ 136
- 88 Projections of the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$ as a $J^P = 1^+$ resonance onto the $m_{J/\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom). 137

- 89 The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds a $J^P = 0^-$ Z on top of a model with a separate 1^+ Z resonance. . . . 139
- 90 The distribution of $m_{J/\psi\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds a $J^P = 0^-$ Z on top of a model with a separate 1^+ Z resonance. . . . 140
- 91 Projections of the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$ as a $J^P = 1^+$ resonance and a broad $Z(4200)^-$ as a $J^P = 0^-$ resonance onto the $m_{J/\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom). 141
- 92 The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds an additional $J^P = 1^+$ Z into a K-matrix with one other 1^+ Z and a separate 0^- Z resonance. 143
- 93 The distribution of $m_{J/\psi\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds an additional $J^P = 1^+$ Z into a K-matrix with one other 1^+ Z and a separate 0^- Z resonance. 144
- 94 Projections of the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$ and $Z(4430)^-$ as a $J^P = 1^+$ K-matrix and a broad $Z(4200)^-$ as a $J^P = 0^-$ resonance onto the $m_{J/\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom). 145
- 95 The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 2^-$ Z into the model with a K-matrix of two 1^+ Zs and a separate 0^- Z resonance. 147
- 96 The distribution of $m_{J/\psi\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 2^-$ Z into the model with a K-matrix of two 1^+ Zs and a separate 0^- Z resonance. 148

- 97 Projections of the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$ and $Z(4430)^-$ as a $J^P = 1^+$ K-matrix, a broad $Z(4200)^-$ as a $J^P = 0^-$ resonance, and a broad $Z(4700)^-$ as a $J^P = 2^-$ resonance onto the $m_{J/\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom). 149
- 98 The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds an additional $J^P = 0^-$ Z into a K-matrix with one other 0^- Z, two 1^+ Zs, and a separate 2^- Z resonance. 151
- 99 The distribution of $m_{J/\psi\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds an additional $J^P = 0^-$ Z into a K-matrix with one other 0^- Z, two 1^+ Zs, and a separate 2^- Z resonance. 152
- 100 Projections of the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$ and $Z(4430)^-$ as a $J^P = 1^+$ K-matrix, a broad $Z(4200)^-$ and $Z(3900)^-$ as a $J^P = 0^-$ K-matrix, and $Z(4700)^-$ as a $J^P = 2^-$ resonance onto the $m_{J/\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom). 153
- 101 The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 1^-$ Z into the model with two $J^P = 0^-$ Zs, two 1^+ Zs, and a separate 2^- Z resonance. 155
- 102 The distribution of $m_{J/\psi\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 1^-$ Z into the model with two $J^P = 0^-$ Zs, two 1^+ Zs, and a separate 2^- Z resonance. 156

- 103 Projections of the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$ and $Z(4430)^-$ as a $J^P = 1^+$ K-matrix, a broad $Z(4200)^-$ and $Z(3900)^-$ as a $J^P = 0^-$ K-matrix, $Z(4700)^-$ as a $J^P = 2^-$ resonance, and $Z(4300)^-$ as a $J^P = 1^-$ resonance. onto the $m_{J/\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom). 157
- 104 The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+$ Z into the model with two other $J^P = 1^+$ Zs in a K-matrix, two $J^P = 0^-$ Zs, one 1^- Z, and a separate 2^- Z resonance. 159
- 105 The distribution of $m_{J/\psi\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+$ Z into the model with two other $J^P = 1^+$ Zs in a K-matrix, two $J^P = 0^-$ Zs, one 1^- Z, and a separate 2^- Z resonance. 160
- 106 Projections of the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(3900)^-$, $Z(4200)^-$, and $Z(4430)^-$ as a $J^P = 1^+$ K-matrix, a broad $Z(4200)^-$ and $Z(3900)^-$ as a $J^P = 0^-$ K-matrix, $Z(4700)^-$ as a $J^P = 2^-$ resonance, and $Z(4300)^-$ as a $J^P = 1^-$ resonance. onto the $m_{J/\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom). 161
- 107 The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+$ Z into the model with three other $J^P = 1^+$ Zs in a K-matrix, two $J^P = 0^-$ Zs, one 1^- Z, and a separate 2^- Z resonance. 163
- 108 The distribution of $m_{J/\psi\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+$ Z into the model with three other $J^P = 1^+$ Zs in a K-matrix, two $J^P = 0^-$ Zs, one 1^- Z, and a separate 2^- Z resonance. 164

- 109 Projections of the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(3900)^-$, $Z(4200)^-$, $Z(4430)^-$, and $Z(4660)^-$ as a $J^P = 1^+$ K-matrix, a broad $Z(4200)^-$ and $Z(3900)^-$ as a $J^P = 0^-$ K-matrix, $Z(4700)^-$ as a $J^P = 2^-$ resonance, and $Z(4300)^-$ as a $J^P = 1^-$ resonance. onto the $m_{J/\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom). 165
- 110 The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+$ Z into the model with four other $J^P = 1^+$ Zs in a K-matrix, two $J^P = 0^-$ Zs, one 1^- Z, and a separate 2^- Z resonance. 169
- 111 The distribution of $m_{J/\psi\pi}$ for the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+$ Z into the model with four other $J^P = 1^+$ Zs in a K-matrix, two $J^P = 0^-$ Zs, one 1^- Z, and a separate 2^- Z resonance. 170
- 112 Projections of the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(3900)^-$, $Z(4200)^-$, $Z(4400)^-$, $Z(4430)^-$, and $Z(4800)^-$ as a $J^P = 1^+$ K-matrix, a broad $Z(4200)^-$ and $Z(3900)^-$ as a $J^P = 0^-$ K-matrix, $Z(4700)^-$ as a $J^P = 2^-$ resonance, and $Z(4300)^-$ as a $J^P = 1^-$ resonance. onto the $m_{J/\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom). 171
- 113 The distribution of $m_{J/\psi K}$ for the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black) and the fit (red) which has five $J^P = 1^+$ Zs in a K-matrix, two $J^P = 0^-$ Zs, one 1^- Z, and a 2^- Z resonance. 174
- 114 Projections of the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(3900)^-$, $Z(4200)^-$, $Z(4400)^-$, $Z(4430)^-$, and $Z(4800)^-$ as a $J^P = 1^+$ K-matrix, a broad $Z(4200)^-$ and $Z(3900)^-$ as a $J^P = 0^-$ K-matrix, $Z(4700)^-$ as a $J^P = 2^-$ resonance, and $Z(4300)^-$ as a $J^P = 1^-$ resonance. onto the $m_{J/\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom). 175

- 115 The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+ Z_K$ into the default model which already has 9 $Z \rightarrow J/\psi\pi$ resonances. 179
- 116 The distribution of $m_{J/\psi\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+ Z_K$ into the default model which already has 9 $Z \rightarrow J/\psi\pi$ resonances. 180
- 117 Projections of the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(3900)^-$, $Z(4200)^-$, $Z(4400)^-$, $Z(4430)^-$, and $Z(4800)^-$ as a $J^P = 1^+$ K-matrix, a broad $Z(4200)^-$ and $Z(3900)^-$ as a $J^P = 0^-$ K-matrix, $Z(4700)^-$ as a $J^P = 2^-$ resonance, $Z(4300)^-$ as a $J^P = 1^-$ resonance, and $Z_{cs}(4470)^-$ as a $J^P = 1^+ J/\psi K$ resonance. onto the $m_{J/\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom). 181
- 118 The distribution of $m_{J/\psi K}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+ Z_K$ into the default model which already has 9 $Z \rightarrow J/\psi\pi$ resonances. 182
- 119 Projections of the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(3900)^-$, $Z(4200)^-$, $Z(4400)^-$, $Z(4430)^-$, and $Z(4800)^-$ as a $J^P = 1^+$ K-matrix, a broad $Z(4200)^-$ and $Z(3900)^-$ as a $J^P = 0^-$ K-matrix, $Z(4700)^-$ as a $J^P = 2^-$ resonance, $Z(4300)^-$ as a $J^P = 1^-$ resonance, and $Z_{cs}(4470)^-$ as a $J^P = 1^+ J/\psi K$ resonance. onto the $m_{J/\psi K}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom). 183
- 120 Dalitz plot with fit pulls in each bin used to calculate 2-D χ^2 . White line draws phase space boundary for the $B^0 \rightarrow J/\psi\pi^-K^+$ decay. 184

121	Distributions of the pulls from the 2-D and 4-D pulls from the multidimensional χ^2 s for the default $B^0 \rightarrow J/\psi\pi^-K^+$ model. The distribution is fit with a Gaussian (red line).	185
122	The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) using the model independent approach.	197
123	The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) using the model independent approach.	198
124	Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the model independent amplitude fit (red) onto the $m_{\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), at $K^*(892)$ (top right), in between $K^*(892)$ and $K_2^*(1430)$ (middle left), at $K_2^*(1430)$ (middle right), and above $K_2^*(1430)$ (bottom).	199
125	The distribution of $m_{\psi(2S)K}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) using the model independent approach.	200
126	Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the model independent amplitude fit (red) onto the $m_{\psi K}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), at $K^*(892)$ (top right), in between $K^*(892)$ and $K_2^*(1430)$ (middle left), at $K_2^*(1430)$ (middle right), and above $K_2^*(1430)$ (bottom).	201
127	Argand plots for model independent 1^+ Z contribution (black points) for helicity-0 (top) and helicity- ± 1 (bottom). Each complex amplitude is labeled with its $m_{\psi\pi}$ bin center. The Argand plots generated from the default 1^+ Z K-matrix for both helicities are shown in red for comparison.	202
128	The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi K\pi$ data (black) and the fit (red) using the model independent approach.	206
129	The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow J/\psi K\pi$ data (black) and the fit (red) using the model independent approach.	207

- 130 Projections of the $B^0 \rightarrow J/\psi K\pi$ data (black points) and of the model independent amplitude fit (red) onto the $m_{\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), at $K^*(892)$ (top right), in between $K^*(892)$ and $K_2^*(1430)$ (middle left), at $K_2^*(1430)$ (middle right), and above $K_2^*(1430)$ (bottom). 208
- 131 Argand plots for model independent 1^+ Z contribution for helicity-0 (top) and helicity- ± 1 (bottom). Each complex amplitude is labeled with its $m_{J/\psi\pi}$ bin center. Red graph represents 1^+ K-matrix from 9 Z fit drawn Argand-style in the same binning as used in the model independent fits. 209
- 132 Unnormalized Legendre polynomial moment $\langle P_1^U \rangle$ of $\cos\theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 4 Zs and 2 Z_K s (bottom). The plots below the moment displays shows the fit pull distributions. 218
- 133 Unnormalized Legendre polynomial moment $\langle P_1^U \rangle$ of $\cos\theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 9 Zs and 1 Z_K s (bottom). The plots below the moment displays shows the fit pull distributions. 219
- 134 Unnormalized Legendre polynomial moment $\langle P_2^U \rangle$ of $\cos\theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 4 Zs and 2 Z_K s (bottom). The plots below the moment displays shows the fit pull distributions. 220
- 135 Unnormalized Legendre polynomial moment $\langle P_2^U \rangle$ of $\cos\theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 9 Zs and 1 Z_K (bottom). The plots below the moment displays shows the fit pull distributions. 221

- 136 Unnormalized Legendre polynomial moment $\langle P_3^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 4 Z s and 2 Z_K s (bottom). The plots below the moment displays shows the fit pull distributions. 222
- 137 Unnormalized Legendre polynomial moment $\langle P_3^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 9 Z s and 1 Z_K (bottom). The plots below the moment displays shows the fit pull distributions. 223
- 138 Unnormalized Legendre polynomial moment $\langle P_4^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 4 Z s and 2 Z_K s (bottom). The plots below the moment displays shows the fit pull distributions. 224
- 139 Unnormalized Legendre polynomial moment $\langle P_4^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 9 Z s and 1 Z_K (bottom). The plots below the moment displays shows the fit pull distributions. 225
- 140 Unnormalized Legendre polynomial moment $\langle P_5^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 4 Z s and 2 Z_K s (bottom). The plots below the moment displays shows the fit pull distributions. 226
- 141 Unnormalized Legendre polynomial moment $\langle P_5^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 9 Z s and 1 Z_K (bottom). The plots below the moment displays shows the fit pull distributions. 227

- 142 Unnormalized Legendre polynomial moment $\langle P_6^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 4 Z s and 2 Z_K s (bottom). The plots below the moment displays shows the fit pull distributions. 228
- 143 Unnormalized Legendre polynomial moment $\langle P_6^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 9 Z s and 1 Z_K (bottom). The plots below the moment displays shows the fit pull distributions. 229
- 144 Unnormalized Legendre polynomial moment $\langle P_7^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 4 Z s and 2 Z_K s (bottom). The plots below the moment displays shows the fit pull distributions. 230
- 145 Unnormalized Legendre polynomial moment $\langle P_7^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 9 Z s and 1 Z_K (bottom). The plots below the moment displays shows the fit pull distributions. 231
- 146 Unnormalized Legendre polynomial moment $\langle P_8^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 4 Z s and 2 Z_K s (bottom). The plots below the moment displays shows the fit pull distributions. 232
- 147 Unnormalized Legendre polynomial moment $\langle P_8^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 9 Z s and 1 Z_K (bottom). The plots below the moment displays shows the fit pull distributions. 233

- 148 Unnormalized Legendre polynomial moment $\langle P_9^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 4 Z s and 2 Z_K s (bottom). The plots below the moment displays shows the fit pull distributions. 234
- 149 Unnormalized Legendre polynomial moment $\langle P_9^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 9 Z s and 1 Z_K (bottom). The plots below the moment displays shows the fit pull distributions. 235
- 150 Unnormalized Legendre polynomial moment $\langle P_{10}^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 4 Z s and 2 Z_K s (bottom). The plots below the moment displays shows the fit pull distributions. 236
- 151 Unnormalized Legendre polynomial moment $\langle P_{10}^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 9 Z s and 1 Z_K (bottom). The plots below the moment displays shows the fit pull distributions. 237

1 The Quark Model

The atomic nucleus was first discovered by Ernest Rutherford in the early 20th century based on the results of a series of experiments conducted by his colleague Hans Geiger and undergraduate student Ernest Marsden. Rutherford determined that based on the scattering of α particles through gold foil that a massive, positively charged center must exist within an atom. During additional experiments of striking various light elements with α particles some years later, Rutherford recognized the presence of hydrogen atoms in his scintillators, proving that “hydrogen nuclei” are present in the nuclei of other elements. This “hydrogen nuclei” was in fact a proton, and this result is viewed as its discovery.

In the 1960s, electron beams with high energies (for the time) were used at the Stanford Linear Accelerator (SLAC) for deep inelastic scattering experiments with a liquid hydrogen target, and these experiments revealed point-like constituents inside the proton, dubbed “partons”¹ at the time. The discovery that the proton, and ultimately the neutron as well, was not an elementary particle served as evidence for the quark model, which was independently proposed by Murray Gell-Mann and George Zweig in 1964 to classify the numerous amount of hadrons (i.e. particles experiencing strong interactions) that were being discovered at the time. Hadrons, such as the proton, were determined to be made of *valence quarks*, gluons, and *sea quarks*. The quark model catalogues hadrons according to their valence quark and anti-quark content, which determine the quantum numbers of hadrons. Sea quarks are virtual quark and anti-quark pairs that form when gluons split inside a hadron and annihilate each other to form a gluon, and since they are much less stable than valence quarks, they do not contribute to the quantum number determination of the hadron. Because of this, further discussion on quark context will only refer to the valence quarks.

Hadrons are particles that contain quarks which are held together by gluons, which is the *gauge boson*² for the strong force. Strong interactions are well described by a quantum field

¹The term *parton* nowadays refers to quarks, anti-quarks, and gluons collectively.

²A gauge boson is a particle that acts as an intermediary for a fundamental force.

theory called Quantum Chromodynamics (QCD). Much like electric charge in relation to the electromagnetic force, there exists a type of charge in relation to QCD known as *color charge*. Quarks are color charged and can be red, green, or blue, while anti-quarks take on the anti-colors of anti-red, anti-green, and anti-blue. To form bound states, color charges need to be added in a way to form a colorless or color neutral state. This can be done in the simplest manner by adding red, green, and blue (or the three anti-colors) or by adding a color and its corresponding anti-color such as red and anti-red. A property called color confinement exists in QCD: due to the strength of the strong force, when two color charges are separated, rather than the bond between them be broken and the color charges be free, a corresponding color-anticolor pair is created from the energy added into the system to form bonds with the previous color charges. Thus, no free quarks can exist under normal conditions and can only be found in colorless combinations with quarks and anti-quarks.

In the Standard Model, there are six known flavors of quarks: up (u), down (d), charm (c), strange (s), top (t), and bottom (b), as can be seen in Figure 1. All quarks have a spin of $1/2$ and have positive parity. The u , c , and t quarks have an electric charge of $2/3e$, while the d , s , and b quarks have an electric charge of $-1/3e$. The u , d , and s quarks constitute the group called the light quarks, since their masses are much smaller than the remaining quarks, and thus the c , b , and t quarks are called heavy quarks. Anti-quarks possess the same quantum numbers but with opposite sign. These quantities along with the flavor of the quarks are all conserved in strong interactions.

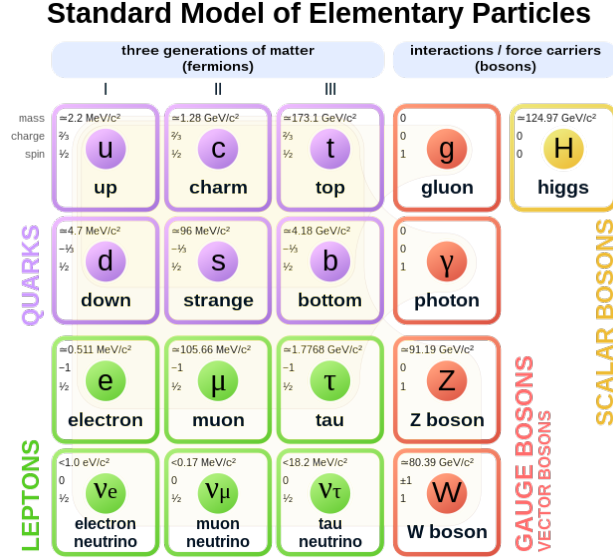


Figure 1: The Standard Model. [3]

A quantity conserved in all decays is *baryon number* (\mathcal{B}). All quarks have $\mathcal{B}=+1/3$, while anti-quarks have $\mathcal{B}=-1/3$. Thus, the color neutral states formed from three quarks (qqq) have a baryon number of 1 and are aptly called *baryons*. Their corresponding antiparticles, $\bar{q}\bar{q}\bar{q}$, have a baryon number of -1. The states formed from a quark and an anti-quark ($q\bar{q}$) have $\mathcal{B}=0$ and are known as *mesons*. These are the simplest states that form a color neutral configuration, and often are referred to as conventional hadrons. However, there are many other ways to create a colorless states besides conventional mesonic and baryonic ones. As discussed later in Section 3, mesonic molecules ($q\bar{q})(q\bar{q})$, tightly bound tetraquarks ($qq\bar{q}\bar{q}$), loosely bound baryon-meson molecules ($qqq)(q\bar{q}$), tightly bound pentaquarks ($qqqq\bar{q}$), or even gluonic states such as (gg) gluonia or hybrid states ($q\bar{q}g$) are all ways to form color singlets. These various states are collectively called *exotic hadronic states*, as they possess more than 3 valence quarks or have valence gluons. Nothing in QCD expressly forbids the existence of these states, and as such they could very well occur. However, such exotic hadrons might decay so quickly that their masses may not be well defined via the Heisenberg's Uncertainty Principle (this mass ambiguity is known as particle width). Very large particles widths would make them theoretically ill defined and experimentally difficult to identify. Even for

reasonably narrow unstable hadrons, known as resonances, production rates in experimentally accessible environments can be very small. Resonances are identified by detection of their decay products to stable particles directly observable in particle detectors. Invariant mass of the decay products serve as measurements of their mass and width. Angular correlations between the decay products serve as determination of their intrinsic angular momentum (spin) and parity (property under mirror reflection). Amplitude analysis is an experimental method which allows for the analysis of resonance masses and of angular correlations at the same time.

For a long time, it was not clear if exotic hadrons in experimentally identifiable form existed at all. Recent results from Belle and LHCb have provided strong evidence for resonances with heavy quarks, inconsistent with conventional hadrons, that are tetraquark and pentaquark candidates, and many analyses are ongoing to identify additional exotic states. This thesis deals with investigation of tetraquark candidates in decays of B^0 mesons ($\bar{b}d$ states) to a vector ($J^P = 1^-$, where J is spin and P is parity) charmonium mesons ($c\bar{c}$ states) denoted as ψ particles, charged pion meson π^- ($d\bar{u}$ ground state) and charged kaon meson of opposite electric charge K^+ (ground $u\bar{s}$ state). In this thesis, charged conjugate states are automatically implied, thus we study both $B^0 \rightarrow \psi\pi^-K^+$ and $\bar{B}^0 \rightarrow \psi\pi^+K^-$ decays. Such decays are due to weak decays of bottom quarks, $b \rightarrow c\bar{c}s$. Charged pions and kaons are stable with respect to strong interactions and have sufficiently long lifetimes such that, once produced, they usually go through many particle detector layers, in which their momentum and charge are measured. Also their mass (distinguishing pions from kaons and from other charged stable particles) is identified with dedicated detector layers. We investigate two vector charmonium states corresponding to the first ($n = 1$) and second ($n = 2$) radial excitations of the triplet (quark spins aligned to total quark spin of $S_{c\bar{c}} = 1$) S-wave (angular momentum between quarks $L_{c\bar{c}} = 0$) $c\bar{c}$ states. Such particles, denoted as J/ψ (we sometimes denote it as $J/\psi(1S)$) and $\psi(2S)$, have suppressed strong decays and therefore decay with reasonably large branching fractions via electromagnetic annihilation

to dimuon pairs, $\psi(nS) \rightarrow \mu^+\mu^-$. Muons are heavier versions of electrons. Their momenta and charges are measured in the same way as other charged particles, but their identity is established by the muon's ability to penetrate a significant amount of dense material, as muons lack strong interactions.

The studied decays of B^0 mesons are dominated by the production of intermediate neutral kaon excitations ($d\bar{s}$ states), generically denoted K^{*0} , $B^0 \rightarrow \psi(nS)K^{*0}$, $K^{*0} \rightarrow K^+\pi^-$. Strong decays of K^{*0} resonances are associated with the generation of $u\bar{u}$ pairs out of their excitation energy. However, there is also a possibility of observing resonances in $\psi(nS)\pi^-$ or $\psi(nS)K^+$ subsystems. Such resonances are necessarily exotic, as their minimum quark content is $c\bar{c}d\bar{u}$ and $c\bar{c}u\bar{s}$, respectively. We denote them as Z^- and Z_K^+ resonances (other common conventions to denote them are Z_c^- and Z_{cs}^+ states). Extra light quarks (u or \bar{u}) needed to form them are easily generated via strong interactions accompanying the weak decays in B^0 disintegration (e.g. in $B^0 \rightarrow Z^-K^+$ intermediate decay).

We first describe the experimental set-up, before we turn to a more detailed discussion of the selected physics topic and the amplitude analysis of the data.

2 Detector Description

2.1 The Large Hadron Collider

The Large Hadron Collider, the largest and highest energy particle collider, is located on the border between France and Switzerland and is housed by CERN - the European Organization for Nuclear Research. The LHC was built to collide beams of protons with a center-of-mass energy up to $\sqrt{s} = 14$ TeV. An overview of the 27-kilometer circumference collider is provided in Figure 2 and details the many subaccelerators that feed to the LHC as well as the current experiments and their relative positions along the ring. Providing 50 MeV of energy, the linear particle accelerator LINAC2 accelerates hydrogen ions into the Proton Synchrotron Booster (PBS), which further accelerates the protons to 1.4 GeV. During the long shutdown 2 (LS2), LINAC2 was replaced with LINAC4, which provides 160 MeV protons, and the PBS accelerates the protons to 2.0 GeV. Following the PBS is the Proton Synchrotron (PS) and then the Super Proton Synchrotron (SPS), where protons achieve energies of 26 GeV and 450 GeV, respectively. After the SPS, the protons are finally injected into the main LHC rings, where two proton beams are accelerated up to their final energies of a few TeV in opposite directions using separate beampipes. These beampipes are held at extremely high vacuum pressure ($\sim 10^{-7}$ Pa) to avoid collisions with gas molecules, and the pressure is even lower near the Interaction Points ($\sim 10^{-9}$ Pa), making it the emptiest space in the Solar System. The protons are accelerated and guided in the beampipes by strong magnetic fields, which are maintained by superconducting magnets that must be held at cryogenic temperatures (~ 5 K). The proton beams are actually not continuous but consist of discrete bunches of 10^{11} protons each, separated by ~ 25 ns corresponding to a bunch-crossing rate of 40 MHz. The beams circulate and collide at the predesignated interaction points - where the ALICE, ATLAS, CMS, and LHCb experiments' detectors are located - for approximately 10-20 hours until the beams are depleted and need to be dumped and replenished.

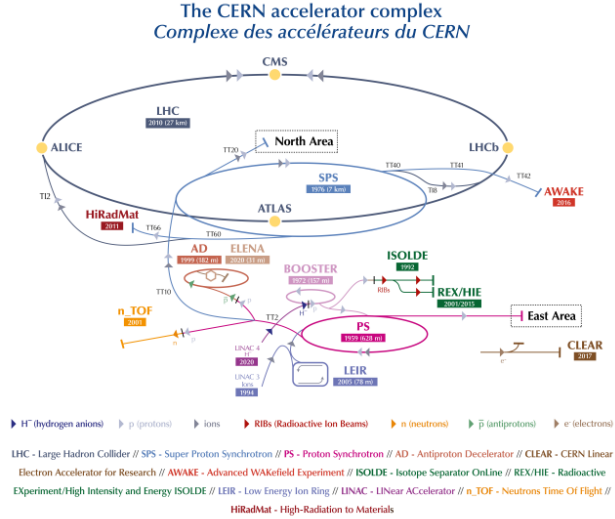


Figure 2: Arrangement of CERN’s subaccelerators and experiments. [4]

2.2 The LHCb Experiment

The LHCb detector is a single-arm forward spectrometer, with a primary goal of precision measurements of CP violation and studies of the decays of bottom and charm quarks. The layout of the detector is shown in Fig. 3, and descriptions of the subdetectors are in the following sections. The angular acceptance of the LHCb detector is 10-300 mrad in the magnet bending plane and up to 250 mrad in the vertical plane, which corresponds to a pseudorapidity³ range of $1.8 < \eta < 4.9$. The choice of this forward geometry is due to the tendency of the directions of b and \bar{b} quarks produced at LHC energies to be along the beam line, as can be seen in Fig 4, which shows the polar angles of the b and \bar{b} -hadrons produced in $\sqrt{14}$ TeV collisions as simulated in PYTHIA.

The LHC was able to operate at a design maximum instantaneous luminosity, $\mathcal{L} \approx 10^{34} \text{cm}^{-2} \text{s}^{-1}$, which was two orders of magnitude above the operational instantaneous luminosity of the design of the LHCb detector, which was $2 \times 10^{32} \text{cm}^{-2} \text{s}^{-1}$. The beams were

³Pseudorapidity is defined as $\eta \equiv -\ln \left[\tan \left(\frac{\theta}{2} \right) \right]$, where θ is the polar angle of the particle momentum with respect to the beam. At hadron colliders, particle production from beam fragments is approximately constant in η . Large η values correspond to small values of θ , thus so called ”forward” direction close to the beam direction.

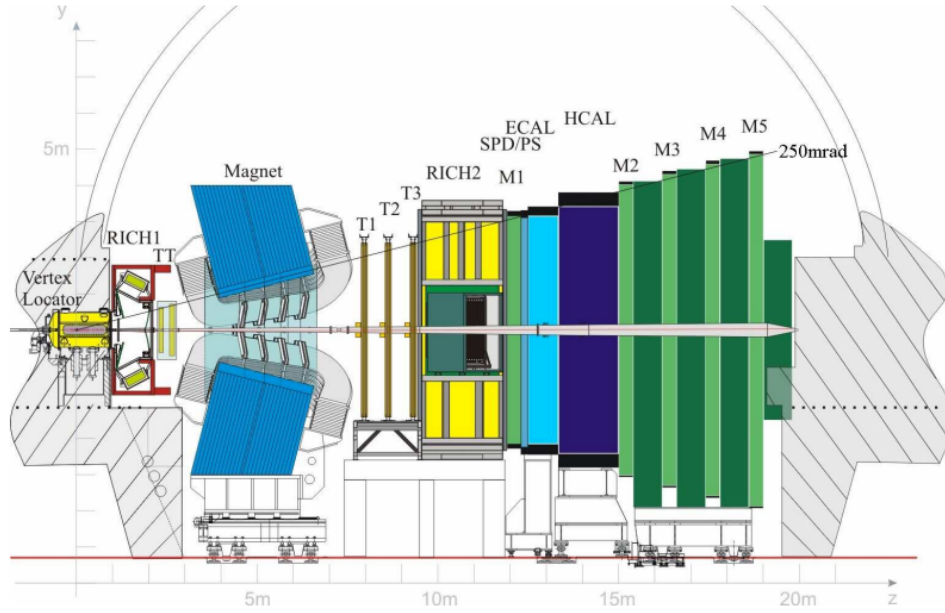


Figure 3: A cross section of the LHCb detector [5]

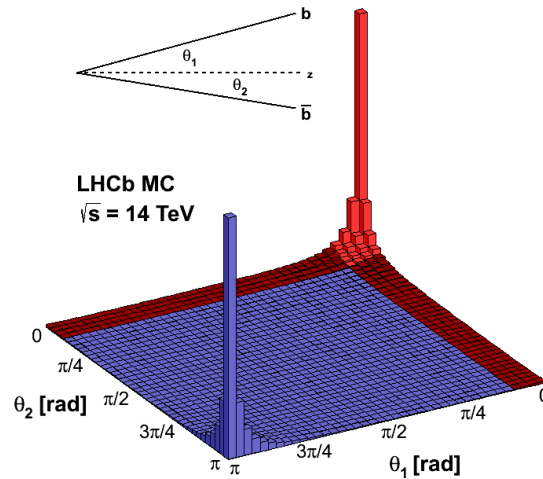


Figure 4: Production angles for $b\bar{b}$ pairs with respect to the beam line. The LHCb acceptance is in red.

[8]

defocused at the LHCb interaction point to produce instantaneous luminosities of up to $4 \times 10^{32} \text{cm}^{-2} \text{s}^{-1}$. Higher instantaneous luminosities correspond to more interactions per second, resulting on average in more proton-proton collisions per bunch crossing (since the

bunch crossing frequency is fixed at 40 MHz), while lower instantaneous luminosities represent less busy events (on average).

During Run 1 (at which the LHC operated at $\sqrt{s}=7$ TeV in 2011 and 8 TeV in 2012)⁴, the LHCb experiment collected data equivalent to a time integrated luminosity ($\mathcal{L}_{int} = \int \mathcal{L} dt$) of 3 fb^{-1} ($1 \text{ fb}^{-1} = 10^{39} \text{ cm}^{-2}$), while in Run 2, the experiment recorded data samples with an integrated luminosity of 6 fb^{-1} , during which the LHC operated at $\sqrt{s}=13$ TeV in 2015-2018. Integrated luminosity multiplied by cross-section (specially normalized probability) for a given reaction gives a number of events containing such reaction produced by the collider in the detector. The cross-section for the production of $b\bar{b}$ quark pairs increases with the beam energy and is of the order of $10^2 \mu\text{b}$ ($1 \text{ b}=10^{-24} \text{ cm}^2$) in the LHCb pseudorapidity range [9]. Less than a third of \bar{b} quarks create B^0 mesons. An even smaller fraction of B^0 mesons decay to the channels under study here, $BF(B^0 \rightarrow J/\psi\pi^-K^+) = (1.15 \pm 0.05) \times 10^{-3}$ and $BF(B^0 \rightarrow \psi(2S)\pi^-K^+) = (0.58 \pm 0.04) \times 10^{-3}$ [2]. The J/ψ channel is further favored over the $\psi(2S)$ channel, because of the larger dimuon branching fraction, $BF(J/\psi \rightarrow \mu^+\mu^-) = (5.96 \pm 0.03) \times 10^{-2}$ as compared to $BF(\psi(2S) \rightarrow \mu^+\mu^-) = (0.80 \pm 0.06) \times 10^{-2}$ [2]. Only a fraction of the produced events containing the channels we analyze produce the final state particles in a configuration in which they can be properly reconstructed in the LHCb detector and selected for physics analysis (this fraction is called efficiency). Data selection is necessary to find events of interest (“signal”) among more frequently produced background events, and occurs at the data collection stage (“triggering”), off-line standard filtering (“stripping”), and our own data selection, as described below.

2.3 Tracking System

The tracking system of LHCb is composed of the Vertex Locator (VELO), the dipole magnet, and the tracking stations. There are four tracking stations: TT, T1, T2, and T3, and the layout of the tracking stations can be seen in Figure 5 along with the different types of tracks

⁴Here, s is the center-of-mass collision energy squared. \sqrt{s} is twice the beam energy.

seen in LHCb. The TT, or Tracker Turicensis, is located directly upstream from the dipole magnet, while the T1, T2, and T3 tracking stations are downstream of the magnet. The tracking system is crucial in the LHCb detector to determine the locations of primary and secondary vertices as well as measuring the momentum and charge of a particle based on how much it bends in the magnetic field. Descriptions of the various subdetectors and components that make up the tracking system are in the following subsections.

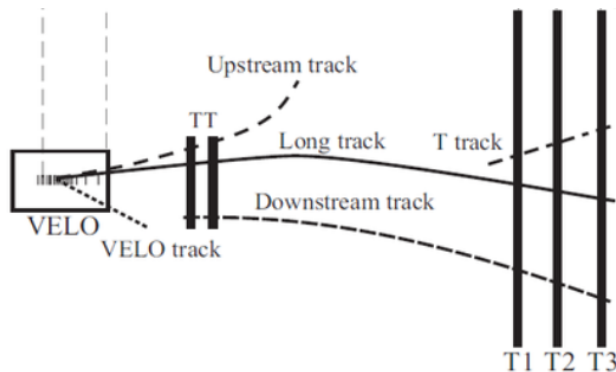


Figure 5: The components of the LHCb Tracking System. Different types of tracks are shown and distinguished by what subdetector components see them. [6]

2.3.1 Vertex Locator

The Vertex Locator (VELO) encompasses the interaction point where the proton beams collide and provides precise measurements on the positions of charged tracks near the interaction region. Given the lifetimes of b and c-hadrons, the VELO is vital to reconstruct the production and decay vertices of such hadrons, as they can travel several millimeters in the lab frame before decaying, and the VELO is also able to measure the impact parameter of particles with respect to the primary vertices. The determination of detached vertices assists in the performance of the High Level Trigger as discussed in Sec. 2.5.2. The VELO consists of 2×21 silicon modules which enclose the beampipe in several stations: half of the modules are responsible for measuring the radial position (r) of tracks, while the other half measure the angular position (ϕ) of tracks. The r modules and ϕ modules are mounted back-to-back, and the position of each station provides the z-coordinate to provide a full three dimensional

measurement of each hit. The aperture needed for the beam injection is larger than the radial distance between the beampipe and the VELO, thus VELO is designed to be retractable such that the beam can be focused over time. These open and closed configurations can be seen in Fig. 6a. A wall of corrugated aluminum shields the VELO modules from the radio frequency radiation generated by the beams, and this wall is designed to be as thin as possible to minimize multiple scattering while still being vacuum-tight.

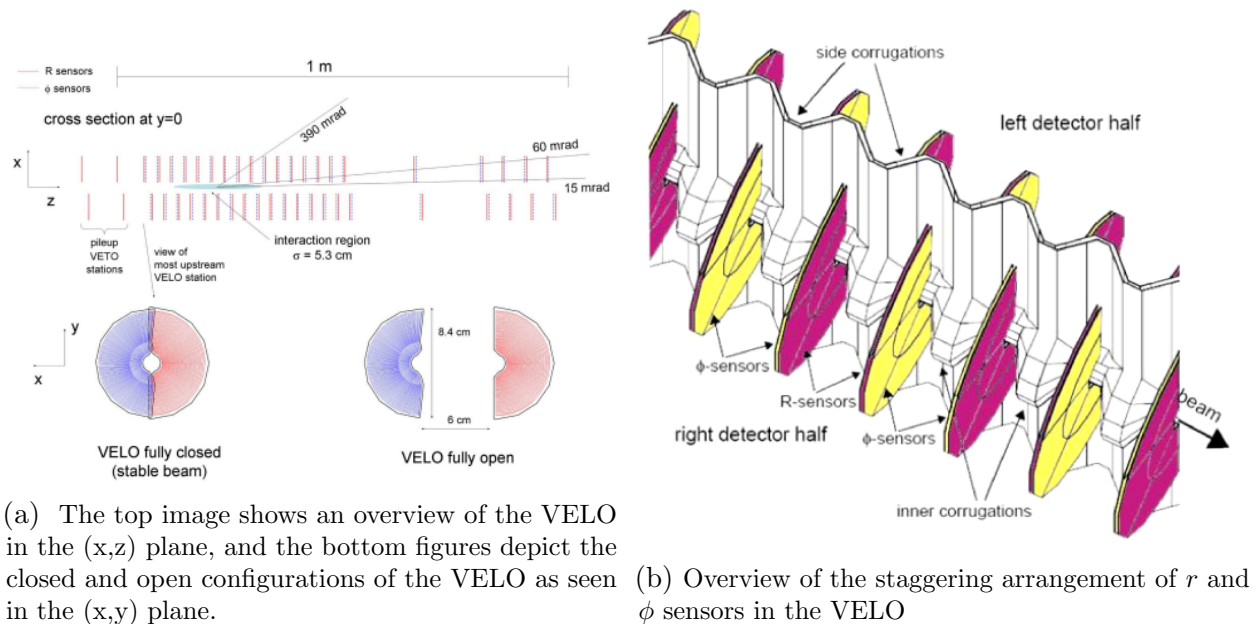


Figure 6: Depictions of the VELO detector.

2.3.2 Dipole Magnets

The dipole magnets are used to measure the momentum of charged particles in the LHCb detector. Two saddle-shaped aluminum coils are mounted symmetrically as seen in Figure 7 in a window-frame yoke. The magnetic field is vertically oriented in the y -direction, so as charged particles traverse the magnetic field, their trajectories bend based on their momenta. The integrated magnetic field for tracks of 10 m in length the lab frame is 4 Tm. The magnet polarity within LHCb can be reversed, and this is done periodically during data taking to study detector asymmetry which impacts CP violation measurements. The magnet thus

spends equal time in the “up” and “down” polarities to avoid systematic biases in these studies.

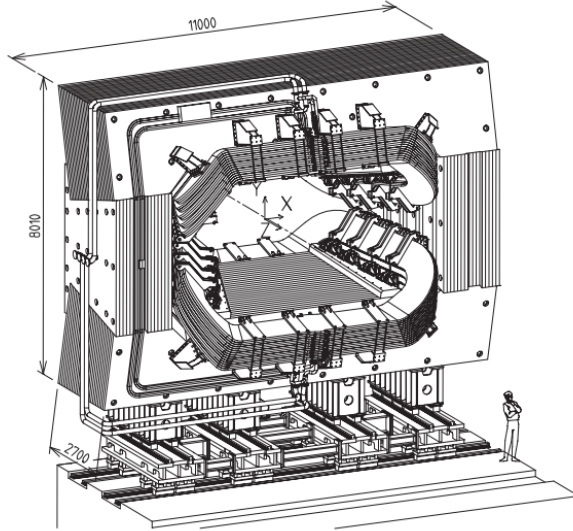


Figure 7: Depiction of the LHCb dipole magnet with dimensions in mm.

2.3.3 Silicon Tracker

The Silicon Tracker (ST) is comprised of the Tracker Turicensis (TT) and the Inner Tracker (IT). They both consist of silicon microstrip sensors, which have a strip pitch of approximately $200\ \mu\text{m}$, giving a spatial resolution of $50\ \mu\text{m}$. The TT’s acceptance matches that of the full detector, and the IT covers only a 120 cm wide and 40 cm high area in the center of the three downstream tracking stations. Each station within the ST contains four layers arranged in a $(x - u - v - x)$ configuration, in which the outer layers have vertically aligned silicon strips and the two inner layers have strips rotated by a stereo angle of -5° and $+5^\circ$, respectively.

The TT consists of half modules that cover half of the acceptance of the detector and combined form full modules that give the full acceptance, as seen in Figure 8a which shows the v layer of the TT. Each half module has seven silicon sensors which are arranged into two or three readout sectors, depending on how far from the beampipe the module is located. The first two layers consist of 15 full modules, and the last two layers have 17 full modules.

To minimize gaps in the acceptance, the neighboring modules are staggered in z and overlap in x . The readout electronics are mounted at the end of the modules, which minimizes the multiple scattering within the active area of the detector.

The IT consists of three stations, each of which has four detector boxes around the beampipe. Each box has four layers of silicon strips with a pitch of about $200 \mu\text{m}$ and the same $(x - u - v - x)$ configuration with a stereo angle of 5° . Each layer has seven modules, with adjacent modules again staggered in z and overlapping in x . The top and bottom modules have a single silicon sensor each while the modules in the side boxes have two silicon sensors, and an example of a layer within the IT is seen in Figure 8b.

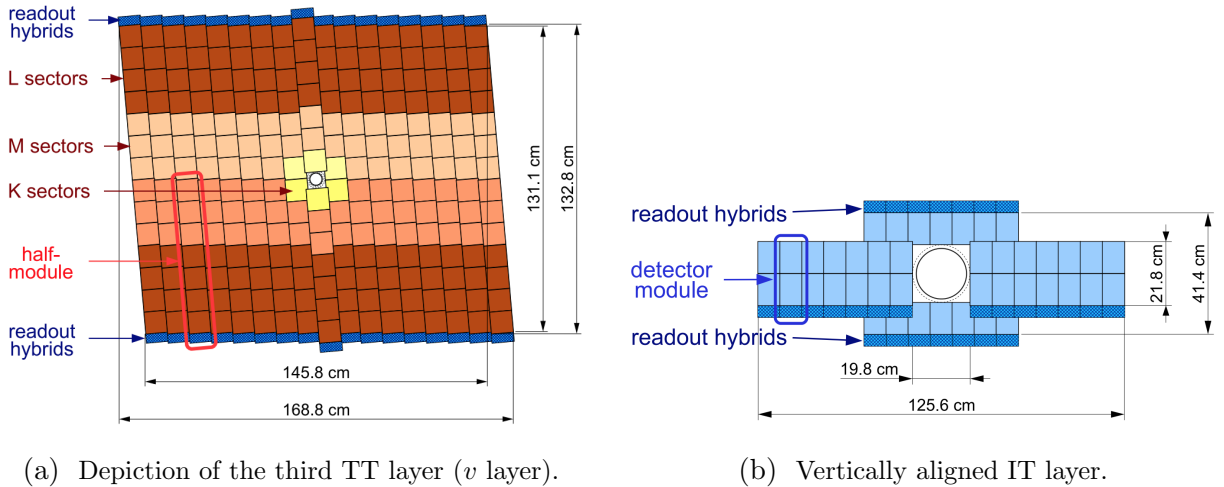


Figure 8: Layers within the different ST stations.

2.3.4 Outer Tracker

The Outer Tracker (OT) is a series of drift time detectors located at the T1, T2, and T3 positions around the IT to cover the remaining acceptance of the detector. Each station consists of gas-tight straw-tube modules of inner diameter 4.9 mm containing a mixture of 70% Argon and 30% CO_2 and is arranged in four layers, following the same $(x - u - v - x)$ configuration as in the ST. The gas mixture provides a maximum drift time of 50 ns and a drift coordinate resolution of $200 \mu\text{m}$. The arrangement of the OT stations and their relative

positions compared to the ST is shown in Figure 9.

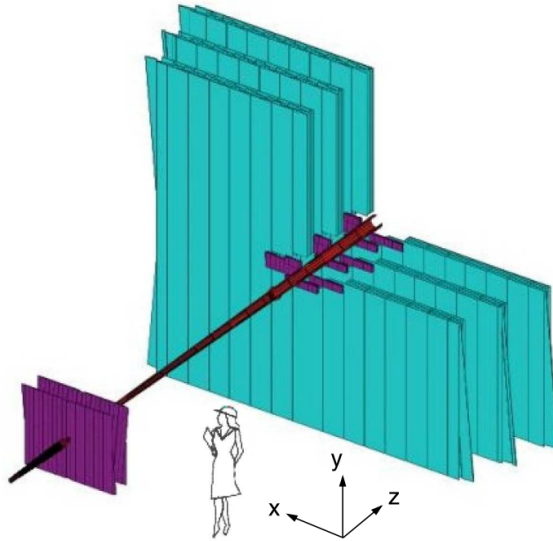


Figure 9: The three OT stations (light blue) and the components of the ST (purple).

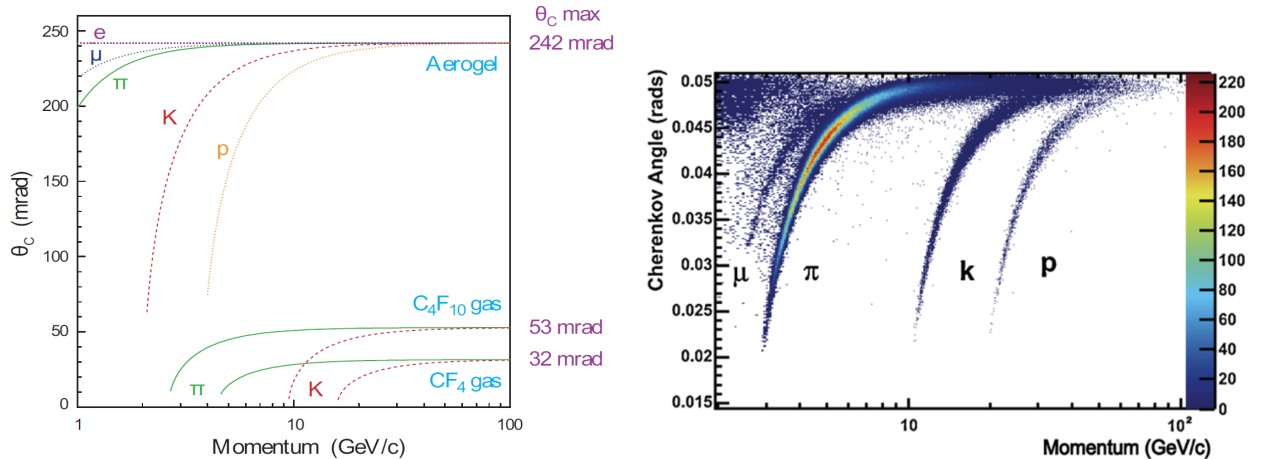
2.4 Particle identification

Particle identification is performed by two Ring-Imaging Cherenkov (RICH) detectors and a muon system and is a critical part of many analyses at LHCb. By combining the information collected by these detectors, the probability of being a specific particle is calculated, allowing for better discrimination between signal and background.

2.4.1 RICH

The RICH detectors enable LHCb to separate hadrons from one another. When a charged particle moves through a dielectric medium at a speed greater than the phase velocity of light in that medium, it emits electromagnetic radiation in a cone with an opening angle, θ_c . This angle depends on the speed of light, c , the refractive index, n , of the dielectric material, and the velocity of the particle, v , by the relationship $\cos \theta_c = \frac{c}{nv}$. Using the measurements of the trajectories and momenta from the tracking system combined with the measured θ_c from the RICH detectors, the mass and charge of a particle can be calculated, and with these, the

likely identity of a particular particle can be determined. Since the momentum spectrum of particles depends on the polar angle of the cone, covering a large range of momenta improves the ability to distinguish particle identities. This is why the LHCb uses two RICH detectors with different radiator materials to cover different areas of the momenta spectrum. RICH1 uses C_4F_{10} as well as aerogel, with the goal of mainly targeting low momenta particles as RICH1 lies upstream of the magnet. A plot showing the polar angle vs the momentum of the particle in different radiator media is found in Figure 10a, and the polar angle as a function of the momentum for RICH1 can be found in Figure 10b. It covers the full angular acceptance of LHCb and a momentum range of 1 – 60 GeV. Downstream of the magnet lies RICH2, and it covers a smaller angular acceptance ($\approx \pm 15$ mrad to ± 120 mrad horizontally and ± 100 mrad vertically). Using a CF_4 radiator, RICH2 covers a high momentum range of $\approx 15 - 100$ GeV.



(a) The opening angle v. the particle momentum in different radiator media shown for multiple particles. (b) The opening angle vs. the particle momentum in the RICH1.

Figure 10: RICH detector measurements of opening angles versus particle momenta.

RICH1 and RICH2 both use spherical and flat mirrors to focus the Cherenkov radiation and reflect the image out of the acceptance of the main detector. Hybrid Photon Detectors (HPDs) are used by both RICH detectors to measure the Cherenkov radiation in a wavelength range of 200 – 600 nm and form the cone that originates from the charged track. RICH1

uses a vertical optical layout, while RICH2 uses a horizontal optical layout, as can be seen in Figure 11.

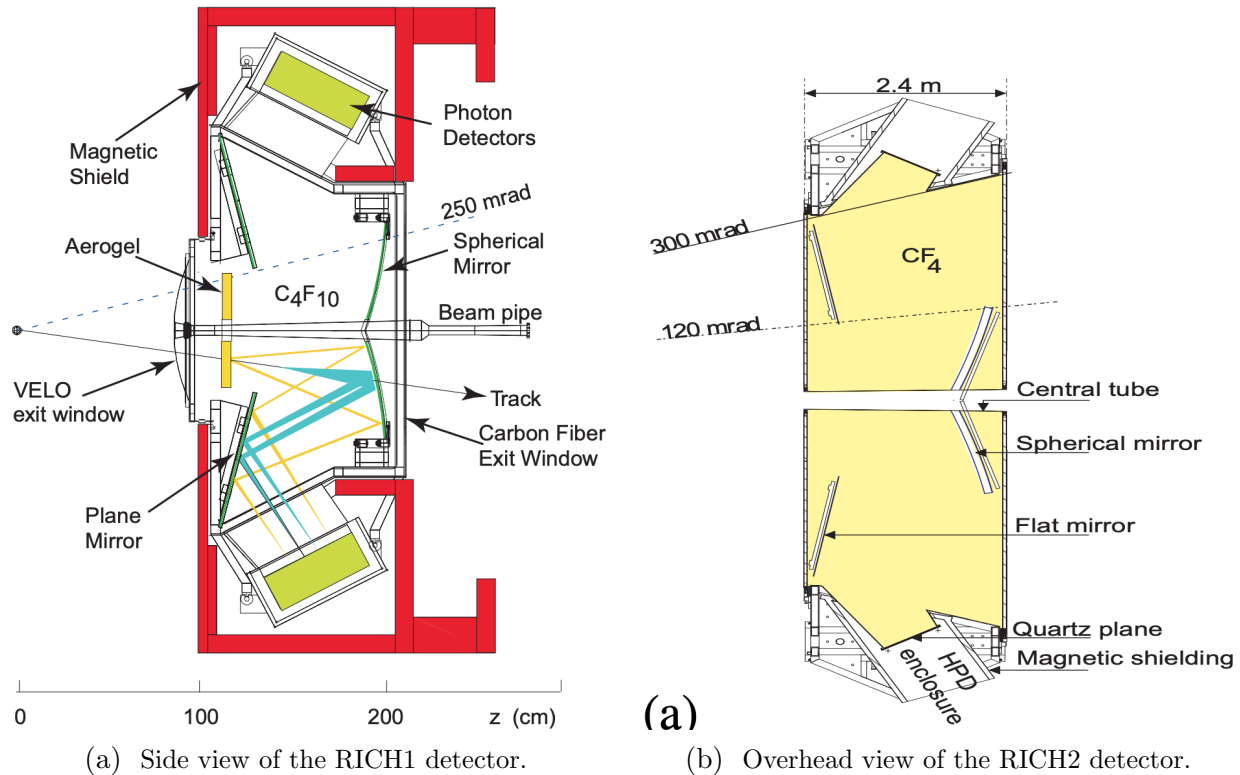


Figure 11: RICH detector layouts.

2.4.2 Calorimeters

Downstream from the tracking stations is the calorimeter system in LHCb. The two main calorimeters are the Electron Calorimeter (ECAL) and the Hadron Calorimeter (HCAL), and upstream of these rest two more detectors: the Scintillating Pad Detector (SPD) and the PreShower detector (PS). These detectors all use scintillation material to detect showers of particles as they pass through. When particles pass through these scintillators, they lose energy and emit light. The scintillation light produced is transmitted via wavelength-shifting (WLS) fibers to photomultipliers, which have an increased efficiency for longer wavelength photons. The detectors are segmented in the $x - y$ plane with increased segmentation near the beampipe, as seen in Figure 12.

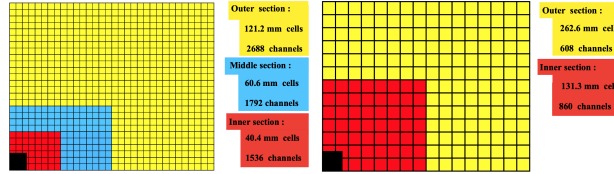


Figure 12: The segmentation of one-quarter of the SPD/PS and ECAL (left) and the HCAL (right).

The SPD/PS detector system contains two planes of scintillator pads with a lead converter layer in between. Since the showers are generated when particles pass through the lead layer, the SPD only detects charged tracks, while the PS can detect the existence of neutral particles when the electromagnetic showers pass through the detector. This allows backgrounds from neutral pions to be reduced by checking for hits in the SPD. Additionally, the backgrounds of charged tracks are reduced by the longitudinal segmentation of the electromagnetic shower detection in the PS. Both the SPD and PS assist the ECAL in the first trigger level (discussed more in Section 2.5.1 by providing the transverse energies of hadrons, electrons, and photons in addition to providing additional positional information about these candidates.

The ECAL and HCAL are sampling calorimeters: the ECAL is mainly responsible for discriminating between electrons, photons, and neutral pions by recording the remaining electromagnetic shower after the PS, and the HCAL identifies hadrons through hadronic cascades. Both detectors use alternating scintillator and absorbing layers. The ECAL uses 4 mm thick scintillator tiles and 2 mm thick lead layers over a 42 cm distance, and the HCAL uses 4 mm thick scintillators and a 16 mm thick iron layer. The HCAL's thickness is limited to 5.6 nuclear interaction lengths due to space constraints and thus can only provide an estimation of the hadron energies, since the HCAL does not absorb the full hadronic shower.

2.4.3 Muon system

The muon system is located the furthest upstream in the LHCb detector, designed to identify muons as they are the only charged particle able to pass through the whole Calorimeter system. It consists of 5 stations (M1-M5) with 80 cm thick iron absorber walls between

the M2-M5 stations, which necessitate that a muon has momentum of at least 6 GeV to penetrate all the stations. A layout of the muon system can be seen in Figure 13. The M1 station is location upstream of the calorimeter system, while M2-M5 are downstream of the HCAL. These stations are responsible for improving the measurement of high p_T muon triggers (to be discussed in Section 2.5.1) and for muon identification information for the HLT (Section 2.5.2) and in offline analysis.

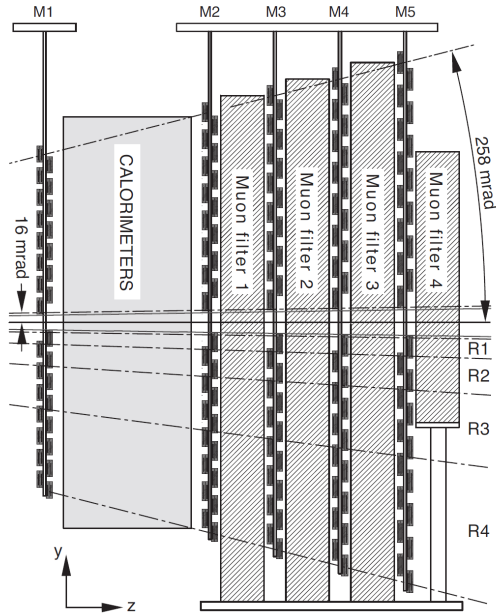


Figure 13: Layout of the muon system.

The muon stations are each made up of four regions (R1-R4), with each region representing a range of distances from the beampipe, where R1 is the closest. Every region uses Multi-Wire Proportional Chambers (MWPCs) except for R1 of M1, which instead uses triple Gas Electron Multiplier (GEM) detectors due to their high radiation tolerances. As R1 of M1 is the most radiation intense region of the muon system, it is necessary to use such detectors instead in this region. The MWPCs use two cathode planes with a plane of wires in between, and the space between is filled with a gaseous mixture of Ar/CO₂/CF₄ at a ratio of 40:55:5, respectively. When an ionizing particle passes through the chamber, electrons and other ions are created and pulled towards the wires. With a sufficiently strong enough electric

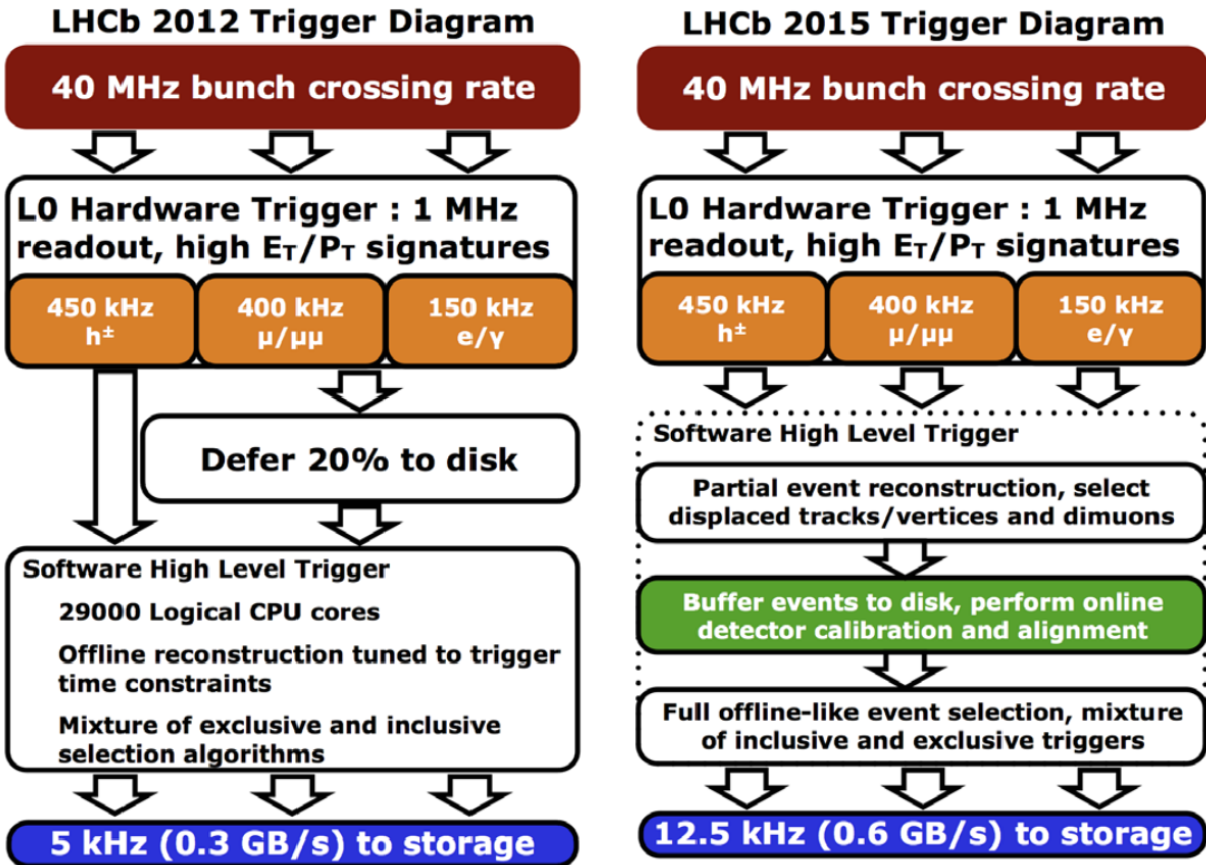
field, these ions and electrons accelerate to high enough energies to further ionize the gas and create additional electrons and ions via a Townsend avalanche, which ultimately creates an electrical signal to identify which chamber was ionized by a particle and thus where the hit took place.

The GEM detectors are made of 50 μm thick insulating kapton foil with a 5 μm thin copper layer on each side. The foil is chemically perforated to create a high density of hourglass-shaped holes, with an external diameter of 70 μm , an internal diameter of 50 μm , and a pitch of 140 μm . These holes improve the gain of the detector by applying a voltage difference between the two copper layers, generating a high electric field within the holes. As these are triple GEM detectors, three of these foils are stacked and sandwiched between the cathode and anode plates, allowing for even higher detector gain. Thus, when a particle passes through the gaseous mixture with the triple GEM detector, the ions and electrons generated can reach high enough energies to trigger an avalanche similar to the ones in the MWPCs. The GEM detectors also use a mixture of Ar/CO₂/CF₄ but at a ratio of 45:15:40.

2.5 Trigger System

The rate at which events with enough measured information take place in the LHCb detector is too high to record them all on disk. These so-called visible interactions occur when there are at least two charged tracks seen by the tracking system. While pp collisions occur within LHCb at a rate of 40 MHz, visible interactions happened at a rate of approximately 10-15 MHz in run 1, and this rate was roughly double in run 2. It is the role of the trigger system to reduce the rate to a manageable level to save, including filtering out uninterested decays which are not useful for physics analyses. In Run 1, this rate was 2-5 kHz, but this increased to 12.5 kHz in Run 2 due to improvements in the ability to store larger number of events for further analysis. To achieve this rate, the initial reduction comes from the level-0 (L0) hardware trigger, which bases its decisions using information from the muon system and the calorimeters, outputting data at 1 MHz. Next, the High Level Trigger (HLT) operates

asynchronously with the collisions taking place by running on a processor farm and reduces the rate further down to the final rate for storage. An overview of the schema used for the triggers for run 1 and run 2 can be seen in Figure 14 and will be explained in more detail in the following sections.



(a) Run 1 trigger scheme.

(b) Run 2 trigger scheme.

Figure 14: LHCb trigger schema for Run 1 and Run 2.

2.5.1 L0 Trigger

The L0 hardware trigger is responsible for the first reduction in the data rate by looking for high transverse momentum (p_T) particles within the calorimeter and high p_T muons in the muon system, both of which are decent indications of a heavy particle decaying. This is done through three independent triggers: the L0-Calorimeter trigger, the L0-Muon trigger, and

the L0-PileUp trigger. The L0-PileUp trigger uses two pile-up modules located upstream of the VELO and was designed to reject events that had too many visible interactions, but since LHCb experience a higher average number of visible interactions per event, this trigger is only used for the determination of luminosity.

The L0-Calorimeter trigger collects information from the entire calorimeter system to determine if an event should be recorded. The transverse energy, E_T , is computed by looking at the deposits from incident particles in 2×2 cell clusters, and from these, three different candidate species are constructed: L0Hadron, which is the highest E_T cluster from the HCAL and also contains the corresponding ECAL cluster energy; L0Photon, which is the highest E_T ECAL cluster that has 1 or 2 PS hits and no corresponding hits in the SPD; and L0Electron, which is the same as L0Photon except that there is at least one hit in the SPD that corresponds to the PS hits. A E_T threshold exists for each candidate type, and if that threshold is met for any candidate type, then the L0 trigger is fired. Similarly, a maximum threshold of SPD hits is used to veto events that would take too long to process the event in the HLT. These values are in Table 1 for both Run 1 and Run 2.

Table 1: Thresholds used for L0 trigger in Run 1 and Run 2. The values represent a minimum threshold for E_T or p_T in GeV and a maximum threshold for SPD hits.

L0 Trigger	E_T or p_T threshold					SPD threshold	
	2011	2012	2015	2016	2017	Run 1	Run 2
Hadron	3.50	3.70	3.60	3.70	3.46	600	450
Electron	2.50	3.00	2.70	2.40	2.11	600	450
Photon	2.50	3.00	2.70	2.78	2.47	600	450
Muon	1.48	1.76	2.80	1.80	1.35	600	450
Muon high p_T	-	-	6.00	6.00	6.00	-	none
DiMuon ($p_{T1} \times p_{T2}$)	$(1.30)^2$	$(1.60)^2$	$(1.69)^2$	$(2.25)^2$	$(1.69)^2$	900	900

Similar to how the L0-Calorimeter trigger works, the L0-Muon trigger checks for the two highest p_T tracks in each quadrant of the muon system. The p_T of muon candidates is estimated based on the direction of the track with the assumptions that the muon came from the interaction point and that the path of the muon was bent by the magnetic field once. The p_T resolution is approximately 25% averaged over the relevant p_T range. The trigger is

fired either if the largest p_T muon is above the L0Muon threshold or if the product of the two largest p_T values is above the L0DiMuon threshold. A special high p_T single muon line exists that checks for a higher threshold than the default single muon line, but it has no maximum threshold for SPD hits. This trigger is used in analyses of electroweak production, and the SPD threshold is removed to not have systematic uncertainties dealing with determining the efficiency of the SPD threshold. These values are also found in Table 1 for both Run 1 and Run 2.

The L0DecisionUnit (L0DU) is passed the information from the various subdetector systems at the collision rate of 40 MHz and renders the final decision of the L0 for an event. Based on beforehand knowledge of the number of collisions LHCb will experience, the thresholds for the different trigger decisions are tuned such that following a logical OR of the trigger decisions made by the L0DU, the L0 output rate is approximately 1 MHz.

2.5.2 HLT

The HLT is a two-stage software trigger that runs on the output from the L0 trigger. It runs on the Event Filter Farm (EFF), which processes the events in parallel on 1,700 nodes with 27,000 CPU cores as of Run 2. The goal of the HLT is to reduce the data rate of 1 MHz from the L0 trigger to a value that can be reasonably recorded onto disk. Given that the HLT has a limited amount of computing resources, this goal is accomplished by first rejecting a large amount of uninteresting events using partial reconstruction (done in the HLT1) and then a more complex veto after performing typically full reconstruction on the event (done in the HLT2).

The HLT1 does its partial reconstruction by performing a precise reconstruction of the primary vertices and of the trajectories of charged particles that go through the entire LHCb detector (so-called "long tracks") that have a $p_T > 500$ MeV. To form long tracks, a sequence of algorithms is performed, as can be seen in Figure 15. First, VELO hits are reconstructed into straight lines that point towards the beam line to create VELO tracks. Then using

three hits in the TT that appear to form a straight line from the VELO, upstream tracks are created. Since the TT is in the fringe area of the magnet, the momentum can be calculated with a resolution of approximately 20%, which allows for the rejection of low p_T tracks. By matching long tracks with TT hits, fake VELO tracks are able to be identified and thus reduced in number. VELO tracks that have a low impact parameter are not able to deposit energy in the TT and thus are passed in the HLT without requiring hits in the TT. Upstream tracks are searched for in a window that is defined by the maximum deflection of a charged particle with $p_T > 500$ MeV for Run 2. This value is lower than in Run 1, because in Run 2, the search is limited to one side of the straight line extrapolation as the charge of the upstream track is able to be estimated. Based on the track upstream and a hit in the downstream tracking stations, the momentum is determined, and hits are projected into a common plane, and then finally by checking for hit clusters, a long track can be identified. After tracks are formed, they are all fit with a Kalman filter that is the same as used in offline reconstruction and uses an improved simplified geometry description of the detector from Run 1 to Run 2. This optimizes parameter estimation of the tracks and takes into account multiple scattering from the material budget of the detector. The fake track rejection has reduced the rate of events passing through this step in the HLT1 by 4%.

The HLT1 also identifies muons beginning with the fully fitted tracks. The geometry of the detector makes the muon identification algorithm momentum dependent: tracks that have momentum $p < 3$ GeV cannot be identified as muons as they would not have enough energy to reach the muon system. The next momentum range is between 3 and 6 GeV, in which hits in M2 and M3 are required to identify a muon. Above 6 GeV and below 10 GeV, a hit in either M4 or M5 is needed in addition to the hits in M2 and M3. Above 10 GeV, hits are required in M2-M5 to identify a muon. This algorithm is also used in HLT2 and offline, but in HLT1, the track reconstruction is only done for tracks with $p_T > 500$ MeV. For particles with p_T below this, an additional muon identification algorithm is used. Upstream tracks are extrapolated to M1-M5, and hits are searched near the extrapolation. If hits are

found in the muon system, then the VELO-TT tracks are also extrapolated with a bend from the magnetic field based on the momentum estimation of the track. Hits that do not already belong to reconstructed long tracks are checked to see if they correspond to the upstream-muon system tracks segments. This allows for muon identification down to a $p_T \approx 80$ MeV at only a small computational cost to the trigger system.

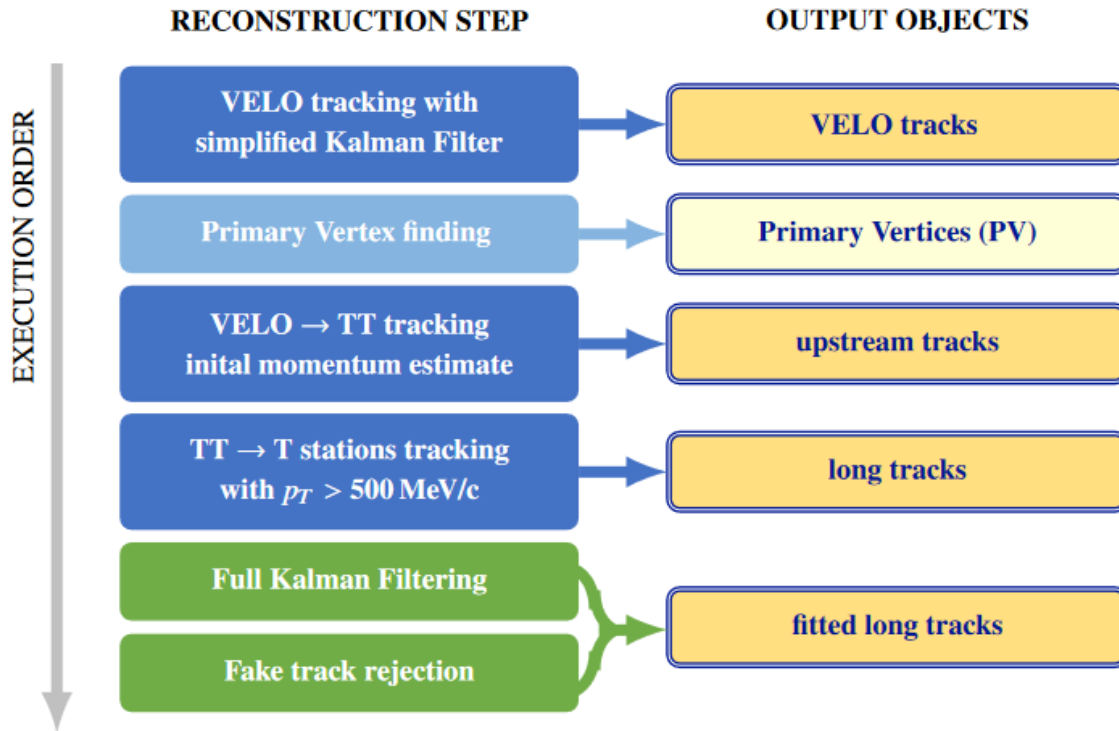


Figure 15: Process through which the HLT1 performs track and vertex reconstruction.

Roughly 55 kB per event pass the HLT1, so with a disk buffer on the EFF of approximately 10 PB, the HLT2 can operate on two weeks of HLT1 data. The HLT2 performs full event reconstruction to further reduce the data rate of about 150 kHz coming out of the HLT1. The process begins with track reconstruction of charged particles, using the full information from the tracking system and finishing any pattern recognition steps omitted from HLT1 due to time constraints. This provides high-quality long and downstream tracks with very precise momentum estimation, and similarly, neutral particle reconstruction algorithms are also executed. Lastly the HLT2 performs particle identification using the RICH detectors,

the calorimeter system, and the muon system. This enables the HLT2 to render decisions on many decision lines, reducing the data rate to a final 12.5 kHz for storage.

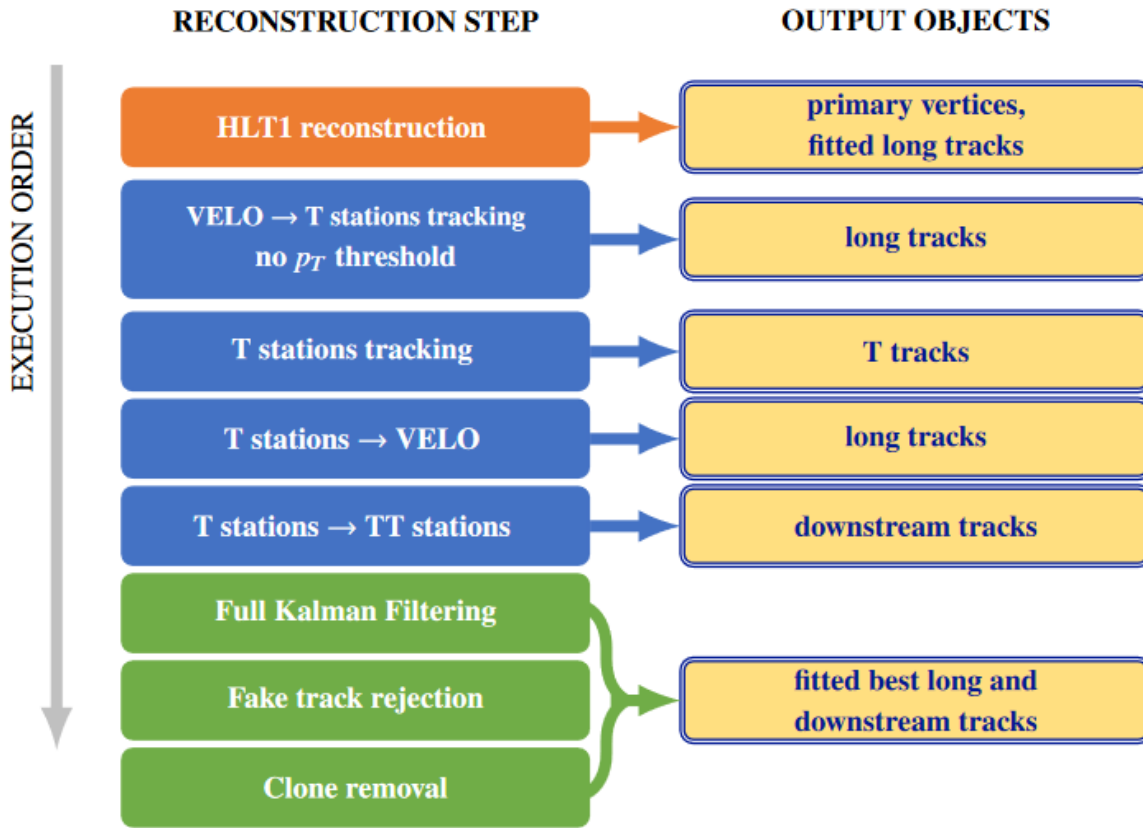


Figure 16: Process through which the HLT2 performs track and vertex reconstruction.

Many decision lines exist in the HLT: 20 in the HLT1 and 500 in the HLT2. The HLT1 renders a decision on two inclusive lines that look for high p_T tracks, one with a single displaced track and the other with a displaced two-track vertex. It also has lines dedicated to select muonic decays of b and c hadrons and muons from W and Z boson decays, such as a displaced muon with high p_T line and a high p_T muon line with no displacement requirement. Lines also exist in the HLT1 to calibrate the trigger, select events with low multiplicity, and select a few exclusive lines such as two-body beauty and charm hadron decays. The HLT2 has many trigger lines, both inclusive and exclusive. The bandwidth of the HLT2 is dominated by the topological b -hadron, inclusive c -hadron, and dimuon trigger lines as well as a large number of exclusive c -hadron TURBO trigger lines.

2.6 Data Flow

Both real data stored by the HLT2 from the detector and simulated Monte Carlo data require additional processing before they are used in offline analyses, as can be seen in Figure 17. This processing is done by a suite of packages developed on the GAUDI framework in order to provide physicists data files that only contain the information relevant to their analysis. For simulated data, the GAUSS project uses PYTHIA and EVTGEN to generate pp collisions and simulate the decays of the produced particles, and GEANT4 is used to simulate the detector geometry and the propagation of the physics generated by PYTHIA and EVTGEN through the detector. Following this, the GAUSS output is run through the BOOLE project to emulate the detector's electronics, converting the geometrical hits into signals that are very similar to the output generated by the detector itself, which allows for simulated data at this point to follow the same pipeline for data processing. The next step for real data or Monte Carlo data is the HLT stage (as discussed previously in Section 2.5.2), which is handled by the MOORE package. To create tracks, vertices, and hit clusters, data is then run through the BRUNEL package which performs a full reconstruction of the triggered data and contains all the information from an event. Data is further processed and filtered through the *stripping* process, which applies loose cuts specific to certain decays and studies. Each specific set of cuts corresponds to a *stripping line*, and the stripping lines for decay channels that are similar and come from similar physics processes are grouped into *streams*. This process is conducted by the DAVINCI project and is done by the collaboration to ease the processing needed each time a user reconstructs data.. Finally, also using the DAVINCI project, data can be acquired from these streams, providing the information needed for offline physics analysis.

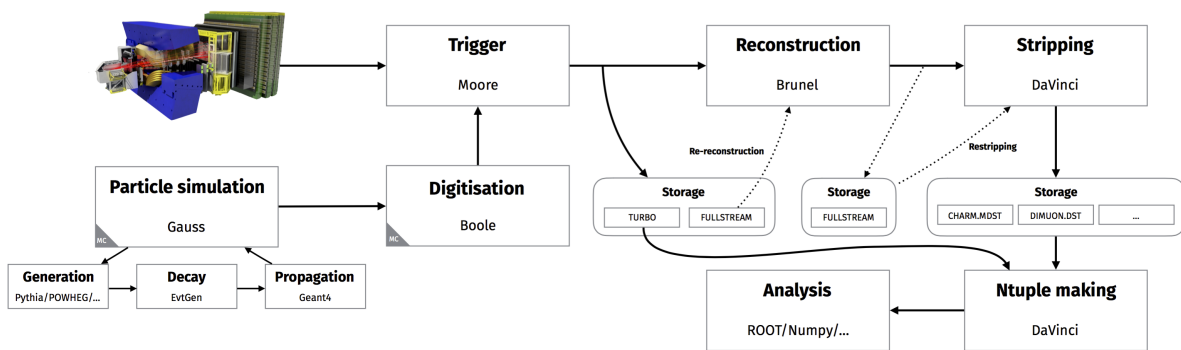


Figure 17: Flowchart depicting the data flow within LHCb for Run 2.

3 Tetraquark candidates observed in $B^0 \rightarrow \psi(nS)\pi^- K^+$ decays

Belle observed the first charmonium-like charged state, $Z(4430)^-$, in the $\psi(2S)\pi^-$ mass spectrum in 2007 analyzing $B \rightarrow \psi(2S)\pi K$ decays [10]. The minimal quark content of such a state is $(c\bar{c}d\bar{u})$. Using a two-dimensional amplitude analysis (*i.e.* “Dalitz analysis”), Belle updated their result in 2009 using a fit of a decay amplitude to $m_{K\pi}^2$ and $m_{\psi(2S)\pi}^2$ distributions with known K^{*0} resonances and the exotic $Z(4430)^-$ resonance [11]. They claimed preference of a model with $Z(4430)^-$ over a model with only K^{*0} resonances by 6.5σ . The mass of the $Z(4430)^-$ was determined to be $4443_{-12}^{+15} {}_{-13}^{+19}$ MeV, and its width was determined to be $107_{-43}^{+86} {}_{-56}^{+74}$ MeV. In 2013, Belle updated their analysis further with a 4D amplitude fit, using the masses, the $\psi(2S)$ helicity angle, and an angle between the decay planes of the intermediate resonances [12]. They measured the mass and width of the $Z(4430)^-$ to be $4485 \pm 22_{-11}^{+28}$ MeV and $200_{-46}^{+41} {}_{-35}^{+26}$ MeV, respectively. Given the construction of the helicity amplitudes, they were also able to probe the quantum numbers and found a preference of $J^P(Z) = 1^+$ over other hypotheses by 3.4σ .

BaBar analyzed their $B \rightarrow \psi(2S)\pi K$ and $B \rightarrow J/\psi\pi K$ data with a different approach, determining the angular moments of $\cos\theta_{K^*}$ (K^{*0} helicity angle) as a function of the $K^+\pi^-$ mass [13]. For a given $m_{K\pi}$ value, there is a one-to-one correspondence between the $\cos\theta_{K^*}$ and $m_{\psi\pi}$ variables. The angular moments for their $B \rightarrow \psi(2S)\pi K$ and $B \rightarrow J/\psi\pi K$ data appeared consistent within the errors, thus they assumed they should be the same for a given $m_{K\pi}$ value. They had no statistically significant peaks in the $m_{J/\psi\pi}$ distribution, so they also assumed that the $B \rightarrow J/\psi\pi K$ data were free of $J/\psi\pi^-$ resonances. Using the angular moments determined in the higher statistics $B \rightarrow J/\psi\pi K$ data and the observed $m_{K\pi}$ distribution in $B \rightarrow \psi(2S)\pi K$ data, they predicted the $m_{\psi(2S)\pi}$ mass distribution in the $B \rightarrow \psi(2S)\pi K$ sample. BaBar claimed that there was no need for a $Z(4430)^-$ resonance to explain their data within the errors of their analysis approach. Their upper limit, however,

did not contradict Belle’s claim.

In 2014, LHCb confirmed presence of the $Z(4430)^- \rightarrow \psi(2S)\pi^-$ resonance (with significance of 14σ) in $B^0 \rightarrow \psi(2S)\pi^- K^+$ decays using Run 1 data (3 fb^{-1}) and the 4D amplitude analysis with ~ 10 times larger signal statistics than available in Belle or BaBar [7]. The mass and width of $Z(4430)^-$ were determined to be $4475 \pm 7_{-25}^{+15}$ MeV and $172 \pm 13_{-34}^{+37}$ MeV, respectively. Its quantum numbers were established to be $J^P = 1^+$ at 9.7σ . Resonant character of $Z(4430)^-$ was demonstrated on an Argand diagram. The analysis also produced evidence for a significant structure at lower $m_{\psi(2S)\pi}$ mass, ~ 4240 MeV, with a large width, $200 - 700$ MeV, which could be attributed to either 0^- or 1^+ partial wave.

LHCb also demonstrated, with a high significance, presence of exotic hadron structures in the $B^0 \rightarrow \psi(2S)\pi^- K^+$ decays using semi-model-independent approach [7, 14] similar to the method used by BaBar [13]. This approach makes assumptions only about the highest spin of the $\pi^- K^+$ partial waves contributing to the data at a given $m_{K\pi}$. In quark model, high spin resonances require a high value of relative angular momentum between constituent $s\bar{d}$ quarks inside an excited kaon, which necessarily means a high mass value of the resonance. Therefore, lower $m_{K\pi}$ values are expected to be populated by low-spin partial waves. In this method, unlike in the amplitude analysis, it is not necessary to specify the exact number and parameters of the contributing K^{*0} resonances. An outcome of such an analysis is to assess consistency of the data with the hypothesis that only conventional K^{*0} resonances can contribute. In case of inconsistency, the semi-model-independent approach is not able to characterize composition of the exotic structures, or even distinguish between $\psi\pi^-$ or ψK^+ exotics. This approach is also sensitive to only relatively narrow exotic structures, as broader structures contribute to low-spin $K^+\pi^-$ partial waves and “self-subtract” [15].

In 2014, Belle analyzed the $B \rightarrow J/\psi\pi K$ data with the 4D amplitude fit [16]. They claimed to observe (6.2σ) a $J/\psi\pi^-$ resonance at 4196_{-29-13}^{+31+17} MeV, with a width of $370 \pm 70_{-132}^{+70}$ MeV and $J^P = 1^+$ determined at 6.1σ (so called $Z(4200)^-$ resonance). They also claimed to observe 4.0σ evidence for $Z(4430)^- \rightarrow J/\psi\pi^-$ decays, albeit by fixing the mass and width

of that resonance to the values determined in the $B \rightarrow \psi(2S)\pi K$ amplitude fits. It should be noted that their amplitude analysis did not find evidence for the narrow $1^+ Z_c(3900)^-$ resonance discovered in $J/\psi\pi^-$ decays by BESIII [17] and Belle [18] in $e^+e^- \rightarrow J/\psi\pi^-\pi^+$ data.

LHCb analyzed $B^0 \rightarrow J/\psi\pi^- K^+$ Run 1 data with the 4D amplitude fit (much larger data sample than available in Belle) [19, 20] but did not publish these results. The semi-model-independent analysis of the $B^0 \rightarrow J/\psi\pi^- K^+$ Run 1 data by LHCb demonstrated with a very high significance presence of exotic $J/\psi\pi^-$ mass structures [21]. However, as discussed above, such approach cannot characterize exotic resonances.

The most popular interpretation of $Z_c(3900)$ resonance, also observed in the neutral version, is a loosely bound $D\bar{D}^*$ molecular state, since the state is narrow ($\Gamma = 28.4 \pm 2.6$ MeV), has a mass near the $D\bar{D}^*$ threshold, has a large branching fraction for the $D\bar{D}^*$ decays, and has the quantum numbers of the $D\bar{D}^*$ pair in S-wave [2]. The $Z(4430)^-$ and $Z(4200)^-$ do not fit all the properties expected in the molecular model, thus could be tightly bound tetraquarks. In fact, such states should be relatively wide, since there is no obvious mechanism to suppress $\psi\pi^-$ decay widths, given the substantial phase-space for such decays. If the tightly bound tetraquark picture is correct, there should be many more $\psi\pi^-$ states of different quantum numbers corresponding to various quark spin configurations and radial or orbital angular momentum excitations.

The goal of our analysis is to perform the 4D amplitude fits to the LHCb Run 1 and Run 2 data (9 fb^{-1}) in $B^0 \rightarrow J/\psi\pi^- K^+$ and $B^0 \rightarrow \psi(2S)\pi^- K^+$ channels. These data samples are much larger than available in the Belle experiments and larger than previously analyzed in LHCb. The increase is not only due to the larger integrated luminosity, but also higher proton-proton collision energies during Run 2 of the LHC, which increased B^0 production cross-section. We also increase the data selection efficiency by improving the selection criteria. By performing back-to-back analyses of the two channels, we hope to shed new light into $\psi\pi^-$ exotic hadron structures.

4 Data Selection

The analyzed data consist of 3 fb^{-1} for Run 1 (collected at $\sqrt{s} = 7 - 8 \text{ TeV}$) and 5.9 fb^{-1} for Run 2 (collected at $\sqrt{s} = 13 \text{ TeV}$) for both $B \rightarrow \psi(2S)\pi K$ and $B \rightarrow J/\psi\pi K$ channels. A breakdown of the stripping versions and luminosity by year is presented in Table 2.

Table 2: Stripping campaigns and luminosity by year.

Year	Stripping Campaign	Luminosity (fb^{-1})
2011	21r1	1.11
2012	21	2.08
2015	24r0p1	0.33
2016	28	1.67
2017	29r2	1.71
2018	34	2.19

B candidates from the B2XMuMu stripping line are used as starting point and some cuts are tightened or added. The selection criteria are summarized in Table 3.

4.1 Preselection

Only good quality charged tracks (χ^2/ndf) with low track ghost probability (GhostProb) are used for B^0 meson reconstruction. Muons are identified with help of Muon detector (IsMuon flag, log-likelihood difference, hereafter DLL, between muon and pion hypothesis: $\text{PID}_\mu \equiv \text{DLL}(\mu-\pi)$) and are required to have a transverse momentum to the beam (p_T) of at least 0.3 GeV. Oppositely charged muon candidates are required to form a good quality vertex (χ^2/ndf). The J/ψ or $\psi(2S)$ candidates are selected by applying cuts on the dimuon mass. After that the muon four-momenta are constrained to reproduce the known J/ψ and $\psi(2S)$ masses [22] exactly. Charged hadron candidates must have $p_T > 0.1 \text{ GeV}$ and must miss the primary pp interaction points by at least 3 standard deviations ($\chi_{IP}^2 > 9$). Together with the muon candidates, they must form a good quality B^0 vertex (χ^2/ndf), well separated from the primary pp interaction vertex (cut on B^0 lifetime) By design, only loose hadron

Table 3: Data selection requirements.

Particle	Quantity	Requirement
All tracks	Track quality: χ^2/ndf	< 3 (< 4 for Run 2)
All tracks	Ghost Probability: GhostProb	< 0.35 (< 0.4 for Run 2)
μ	p_T	> 0.300 GeV
μ	IsMuon	True
μ	PID_μ (DLL(μ - π))	> -0.3
Di- μ	Vertex quality: χ^2/ndf	< 9
Di- μ	p_T	> 0.100 GeV
J/ ψ	Mass window	[3.040,3.140] GeV
$\psi(2S)$	Mass window	[3.631,3.731] GeV
π	p_T	> 0.100 GeV
π	χ_{IP}^2	> 9
π	PID_K (DLL(K- π))	< 5 && $\neq 0$
K	p_T	> 0.100 GeV
K	χ_{IP}^2	> 9
K	PID_K (DLL(K- π))	> -5 && $\neq 0$
K, π	$\text{ProbNN}_K(\text{K}) * (1 - \text{ProbNN}_\pi(\text{K}))$ - $\text{ProbNN}_K(\pi) * (1 - \text{ProbNN}_\pi(\pi))$	> 0
K $\pi\mu\mu$ ($\pi \rightarrow \text{K}$ mass)	KK $\mu\mu$ mass veto window	$(m_{KK} \subset [1.0125, 1.0265]$ GeV $ \pi_{\text{ProbNN}_K} > 0.2)$ & $m_{KK\mu\mu} \subset [5.355, 5.380]$ GeV
K $\pi\mu\mu$ (K $\rightarrow \pi$ mass)	$\pi\pi\mu\mu$ mass veto window	[5.355, 5.380] GeV & $K_{\text{ProbNN}_\pi} > 0.2$
K $\pi\mu\mu$ ($\pi \rightarrow \text{p}$ mass)	pK $\mu\mu$ mass veto window	[5.605, 5.635] GeV & $\pi_{\text{ProbNN}_p} > 0.2$
B^0	Mass window	[5.1, 5.6] GeV
B^0	Vertex quality: χ^2/ndf	< 9
B^0	Lifetime: τ	> 0.250 ps
B^0	Trigger	L0, Hlt1, Hlt2 TOS
B^0	DLL (multivariate - Sec. 4.2)	< 2.2

identification cuts are used. The RICH detectors are the main contributors to the related DLL values. Since the accepted DLL(K- π) ranges for kaons and pions have an overlap region, we choose for a kaon, the track which is more likely to be a kaon among the two hadrons in B^0 candidate (a combination of ProbNN variables is used, which are also determined mostly from the RICH detectors).

The selected B^0 candidate must be among the reasons for why the event was saved during the data taking for offline analysis (so called Triggered-On-Signal requirement - TOS).

The detailed trigger criteria are for the three decision stages (L0, Hlt1 and Hlt2) are:

(L0Muon_{TOS}|L0DiMuon_{TOS})&
 (Hlt1TrackAllL0_{TOS}|Hlt1TrackMuon_{TOS}|Hlt1TrackMVA_{TOS}|Hlt1DiMuonHighMass_{TOS})&
 (Hlt2DiMuonPsi2S(JPsi)_{TOS}|(Hlt2Topo(2,3,4)BBDT_{TOS}|(Hlt2TopoMu(2,3,4)BBDT_{TOS}|
 Hlt2DiMuonPsi2S(JPsi)HighPT_{TOS}|Hlt2DiMuonDetachedPsi2S(JPsi)_{TOS}|
 Hlt2DiMuonDetachedHeavy_{TOS}).

Decays of other neutral b -quark hadrons to $\psi \rightarrow \mu^+\mu^-$ and oppositely charged hadron pair can fake $B^0 \rightarrow \psi K^+\pi^-$ candidates, when a hadron is misidentified. In particular, $B_s^0 \rightarrow \psi K^+K^-$ (among which $\phi \rightarrow K^+K^-$ is most prevalent) can easily become B^0 candidate when K^- is taken as π^- candidate. To veto such cases the b -hadron mass is recalculated under $\pi^- \rightarrow K^-$ reassignment. The B^0 candidate is eliminated if such mass falls into the B_s^0 mass window. To reduce loss of signal efficiency, this is done only if either K^+K^- mass is within the ϕ mass window or probability of the π^- candidate to be kaon (π_{ProbNN_K}) is reasonably high. While $B_s^0 \rightarrow \psi\pi^+\pi^-$ decays cannot contribute to the B^0 mass peak when π^+ is mistaken as K^+ , they can peak in the upper B^0 sideband, thus such decays are also eliminated. Similarly we veto $\bar{\Lambda}_b^0 \rightarrow \psi K^+\bar{p}$ decays, with \bar{p} mistaken as π^- .

4.2 The multivariate final selection

Background is further suppressed by selecting events based on a likelihood ratio. The total likelihood is a product of the probability density functions ($PDFs$) of four sensitive variables (x_i):

- the minimum impact parameter (IP) χ^2 of K or π with respect to the closest primary vertex (PV),
- $\chi_{\text{vtx}}^2/\text{ndf}$ of the B candidate,
- the B candidate IP significance (χ_{IP}^2),

- and minimum transverse momentum of K or π .

The signal PDF s ($PDF_{sig}(x_i)$) are obtained from the real $B \rightarrow J/\psi\pi K$ data using 10% of the run 2 statistics. The background PDF s ($PDF_{bkg}(x_i)$) come from the B signal's side-band region in the same sample of data (5207 – 5251 | 5309 – 5353 MeV).

We form the logarithm of the ratio between the signal and background PDF s, $DLL_{sig/bkg} = -2 \sum_{i=1}^4 \ln(PDF_{sig}(x_i)/PDF_{bkg}(x_i))$. Figure 18 shows the good agreement between the data and the training sample on the DLL distribution as well as the discrimination between signal and background of $B \rightarrow \psi(2S)\pi K$. We require $DLL_{sig/bkg} < 2.2$ to maximize $N_{sig}/\sqrt{N_{sig} + N_{bkg}}$, where N_{sig} is the expected signal yield and N_{bkg} is the background yield in the region of the B peak. The normalization of N_{sig} and N_{bkg} is obtained from a fit to the B mass distribution.⁵

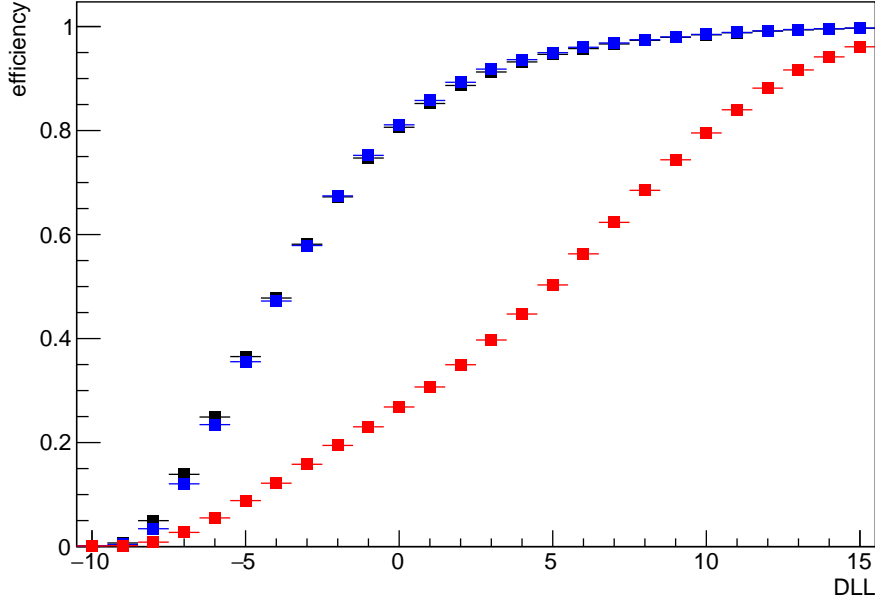


Figure 18: Fraction of the $B \rightarrow J/\psi K\pi$ signal events passing a cut $DLL < x$ for the data (black), the signal training sample (blue), and fraction of the background events (B sideband) passing this cut (red).

⁵The B^0 peak region here is defined as $\pm 2.0\sigma$, where σ is the observed mass resolution given later in the text.

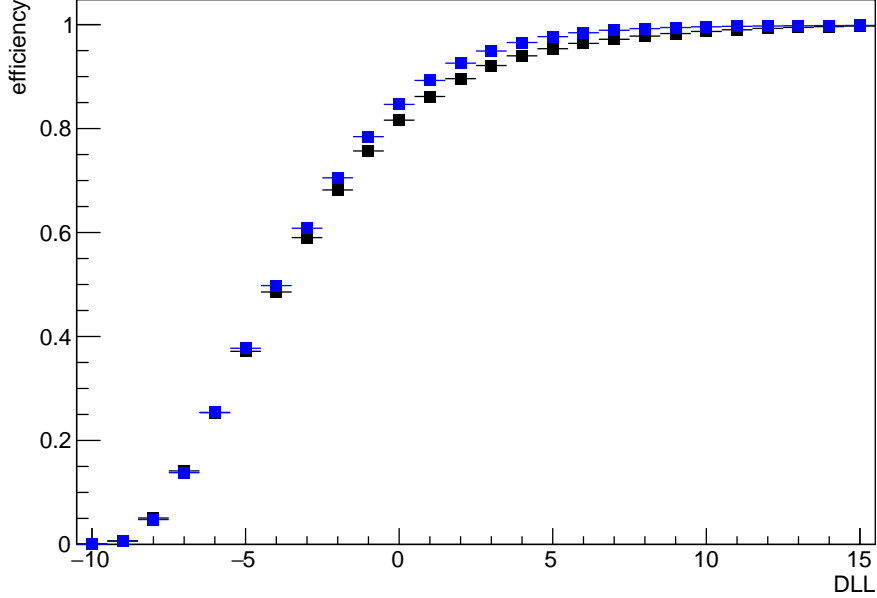


Figure 19: Fraction of the $B \rightarrow J/\psi K \pi$ signal events passing a cut $DLL < x$ for the data (black) and fraction of the signal MC events passing this cut (blue).

The signal PDF produced from the real $B \rightarrow J/\psi \pi K$ data was also used for calculating the DLL variable in $B \rightarrow \psi(2S) \pi K$.

The mass of the $\psi(2S)$ and J/ψ candidates are constrained to their nominal respective masses, and the $\psi \pi K$ candidate is constrained to point to the nearest primary vertex when calculating the mass of the B^0 candidates. The $\psi \pi K$ candidate is also constrained to the world average B^0 mass, before events are used in the amplitude fitter.

The mass distribution of the $B \rightarrow \psi(2S) \pi K$ candidates passing the full selection is shown in Fig. 22. The distribution is fit with a double-sided asymmetric Crystal Ball function for the signal B^0 , another double-sided asymmetric Crystal Ball function for the B_s^0 peak, and a quadratic function for the background. The Crystal Ball shapes used for both the signal B^0 and the background B_s^0 are forced to share the same tails, i.e. the same α_{left} , α_{right} , n_{left} , and n_{right} are used for each peak, but each peak has its own mean, width, and amplitude. The fit gives $140\,344 \pm 435$ $B \rightarrow \psi(2S) \pi K$ events as compared to $2\,255 \pm 50$ in Belle’s 2D analysis, $2\,021 \pm 53$ in BaBar’s, and $25\,176 \pm 174$ in LHCb’s 2014 analysis. The

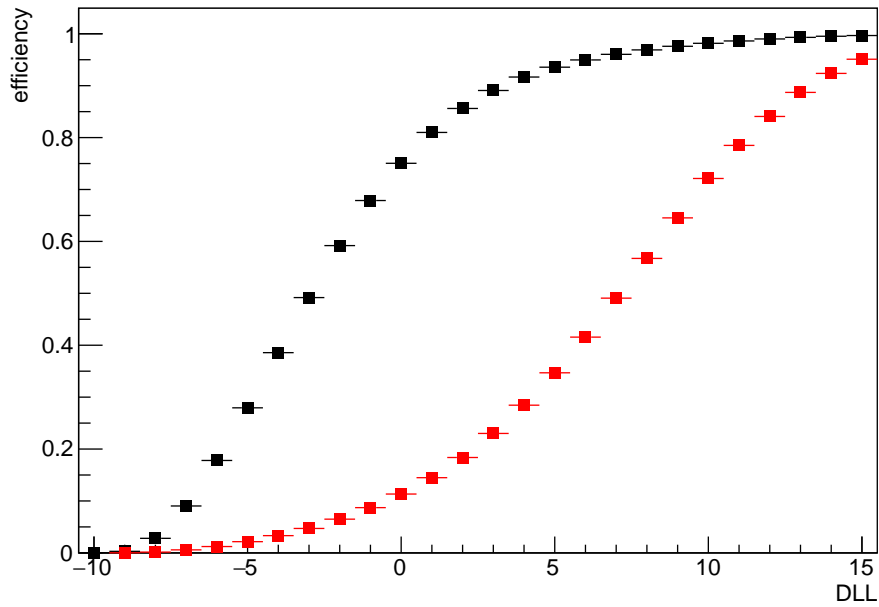


Figure 20: Fraction of the $B \rightarrow \psi(2S)K\pi$ signal events passing a cut $DLL < x$ for the data (black) and fraction of the background events (B sideband) passing this cut (red).

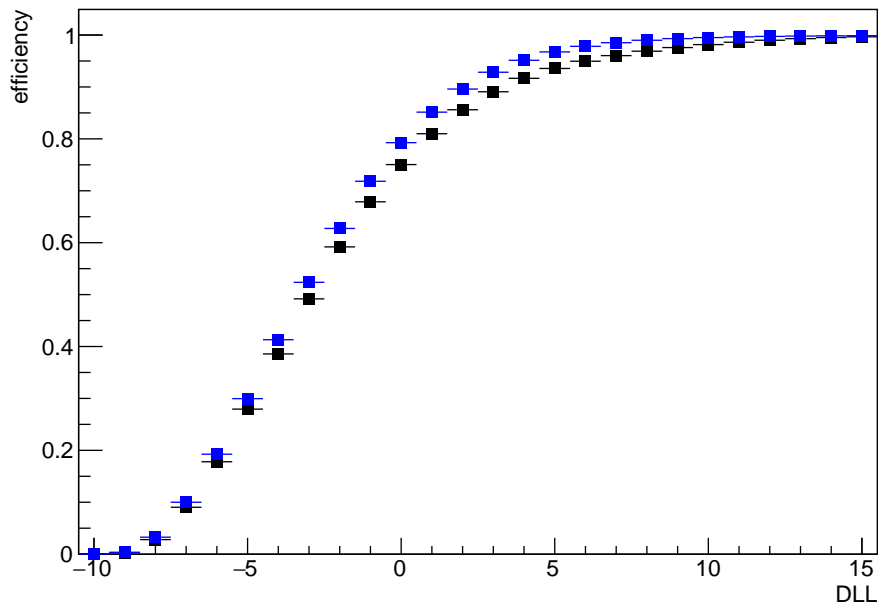


Figure 21: Fraction of the $B \rightarrow \psi(2S)K\pi$ signal events passing a cut $DLL < x$ for the data (black) and fraction of the signal MC events passing this cut (blue).

fitted B^0 mass resolution is 5.22 ± 0.03 MeV. We define a $B \rightarrow \psi(2S)\pi K$ signal region as $5270 < m(\psi(2S)K^+\pi^-) < 5290$ MeV ($\pm 2\sigma$ around the peak) and two mass sidebands as $5175 - 5215$ MeV and $5390 - 5430$ MeV to probe the background. The background level (from B_s^0 and combinatorial) in the signal region is 1.55% (hereafter denoted as f_{bkg}).

Likewise, the mass distribution of the $B \rightarrow J/\psi\pi K$ candidates passing the full selection is shown in Fig. 23. The distribution is also fit with a double-sided asymmetric Crystal Ball function for the signal and a quadratic function for the background. The fit gives $2\,278\,460 \pm 1\,786$ $B \rightarrow J/\psi\pi K$ events as compared to $29\,990 \pm 190$ in Belle's analysis. The fitted B^0 mass resolution for the J/ψ channel is 7.20 ± 0.01 MeV. We define a $B \rightarrow J/\psi\pi K$ signal region as $5265 < m(\psi(2S)K^+\pi^-) < 5295$ MeV ($\pm 2\sigma$ around the peak) and two mass sidebands as $5175 - 5205$ MeV and $5400 - 5430$ MeV to probe the background. The background level in the signal region is $f_{\text{bkg}} = 2.32\%$.

The Monte Carlo sample contains 583 466 events in the signal range for $B^0 \rightarrow \psi(2S)\pi^- K^+$ and 1 814 640 events for $B^0 \rightarrow J/\psi\pi^- K^+$.

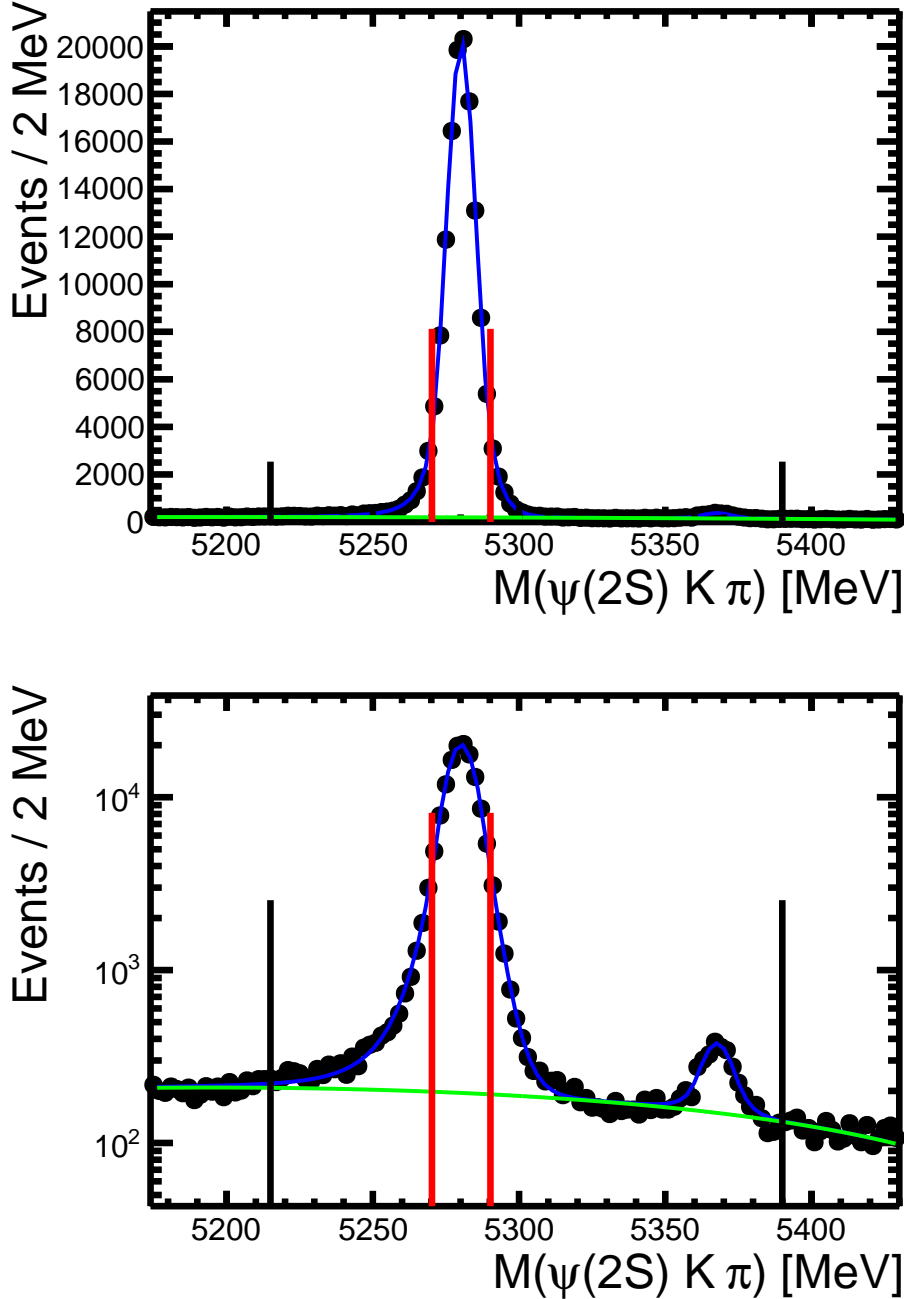


Figure 22: Mass distribution of the $B \rightarrow \psi(2S)\pi K$ candidates in Run 1 and Run 2 data. The fit (blue line) of a double-sided asymmetric Crystal Ball signal shape and quadratic background (green line) are superimposed. The fit results are $m(B^0) = 5280.2 \pm 0.02$ MeV, $\sigma(B^0) = 5.22 \pm 0.03$ MeV, $m(B_s^0) = 5367.7 \pm 0.25$ MeV, $\sigma(B_s^0) = 5.01 \pm 0.24$ MeV, $\alpha_{left} = 1.56 \pm 0.03$, $\alpha_{right} = 1.43 \pm 0.04$, $N_{sig} = 140,344 \pm 435$ events, $N_{B_s^0} = 1,538 \pm 72$ events, and $N_{bkg} = 2232 \pm 277$ events. Vertical red lines show the signal region used in the default amplitude fit. Smaller vertical black lines indicate boundaries of the sidebands used for the background parameterization in the fit. A log scale is used to better show the B_s^0 peak.

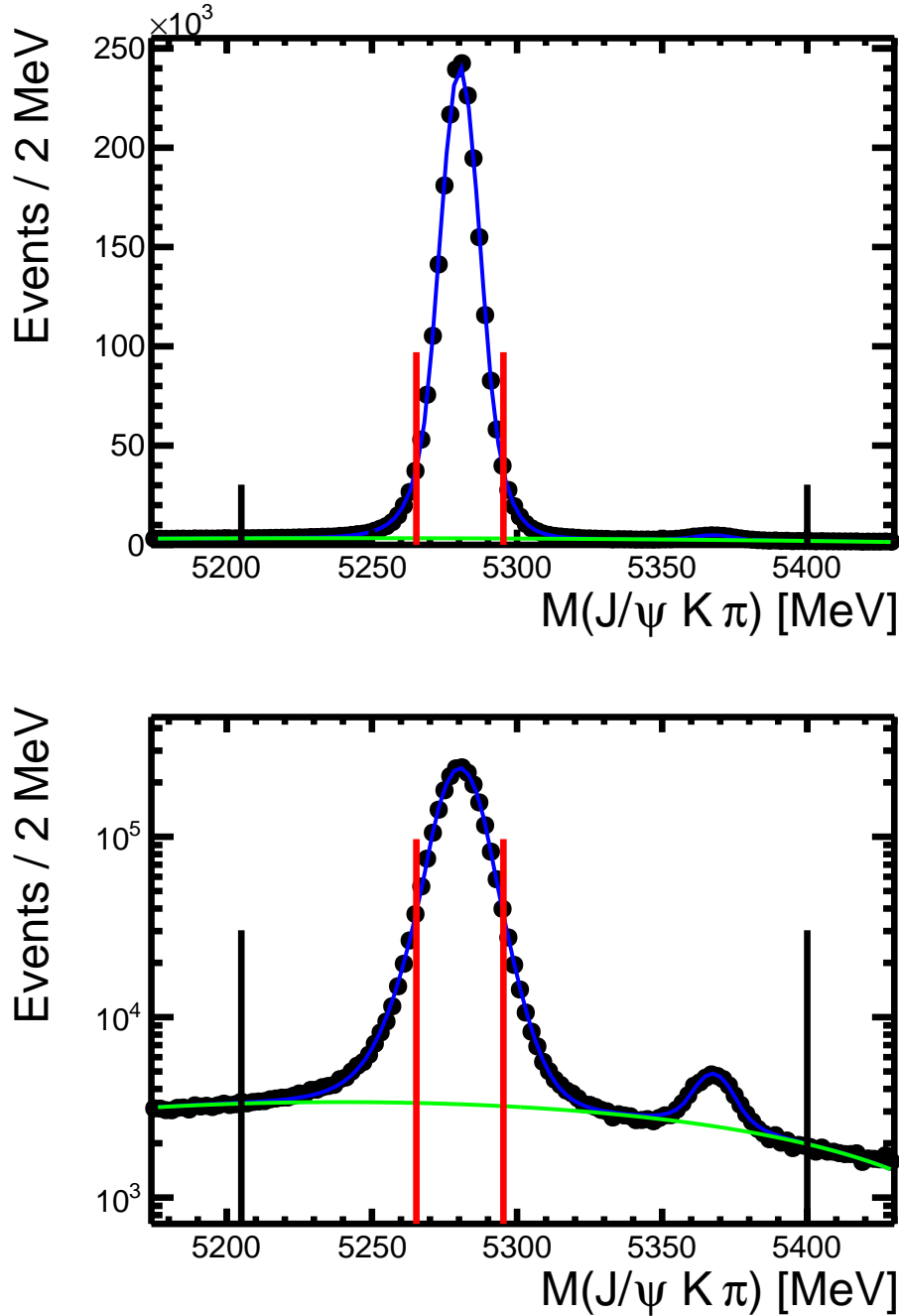


Figure 23: Mass distribution of the $B \rightarrow J/\psi \pi K$ candidates in Run 1 and Run 2 data. The fit (blue line) of a double-sided asymmetric Crystal Ball signal shape and quadratic background (green line) are superimposed. The fit results are $m(B^0) = 5280.3 \pm 0.007$ MeV, $\sigma = 7.20 \pm 0.009$ MeV, $m(B_s^0) = 5367.6 \pm 0.10$ MeV, $\sigma(B_s^0) = 6.49 \pm 0.10$ MeV, $\alpha_{left} = 1.45 \pm 0.008$, $\alpha_{right} = 1.44 \pm 0.009$, $N_{sig} = 2,278,460 \pm 1,786$ events, $N_{B_s^0} = 20,597 \pm 341$ events, and $N_{bkg} = 36,606 \pm 131$ events. Vertical red lines show the signal region used in the default amplitude fit. Smaller vertical black lines indicate boundaries of the sidebands used for the background parameterization in the fit. A log scale is also shown to better show the B_s^0 peak.

5 Kinematic re-weighting of Monte Carlo

To correct for possible inaccuracies in the modeling the B^0 production within the LHCb acceptance and in the modeling of the detector efficiency, we give Monte Carlo events weights (w^{MC}) based on the relative yields of the data and MC in certain kinematic or event variables, following all previous cuts and corrections.

The variables corrected for in this study are: the number of tracks in the event ($nTracks$), transverse momentum of the B^0 ($p_T(B)$), the momentum of the kaon ($p(K)$), and the momentum of the pion ($p(\pi)$). The overall weight applied to a Monte Carlo event is the product of the individual weights for each variable:

$$w^{\text{MC}} = w_{p_T(B)} \cdot w_{nTracks} \cdot w_{p(K)} \cdot w_{p(\pi)}. \quad (1)$$

Data in the signal region is background subtracted using the distribution in the sideband region. To produce a given weight, a normalized plot of both Monte Carlo and the background subtracted data was produced. A binning was chosen based on trends in discrepancy as well as statistical errors. This leads to a finer binning at lower variable values (*e.g.* low momentum, low multiplicity, etc.) and a coarser binning at higher values (*e.g.* high momentum, high multiplicity, etc.). Once an appropriate binning has been established, a ratio of data to Monte Carlo is found for each bin. The plot of each ratio is then fit with an appropriate functional shape, which can be found in Table 4.

Table 4: Fit shapes used for each kinematic/event variable for each channel.

Variable	$\psi(2S)$	J/ψ
nTracks	4th order polynomial	4th order polynomial
$p_T(B)$	Gaussian + constant	Gaussian + constant
$p(K)$	Crystal Ball + 2nd order polynomial	Crystal Ball + constant
$p(\pi)$	2nd order polynomial	Crystal Ball + constant

Then, each Monte Carlo event is weighted by the function value corresponding to that

event before similar procedure is repeated for the next kinematic weight variable. After weighting, all kinematic variables of both data and MC are once again plotted and compared to ensure proper agreement. Following this process, the Monte Carlo set and data reasonably agree with each other. The individual weights used in this analysis are obtained from the plots in Fig. 24 for $\psi(2S)$ and Fig. 26 for J/ψ , while the final comparison between the data and MC is shown in Fig. 25 for $\psi(2S)$ and Fig. 27 for J/ψ .

Analyzing the data while using unweighted Monte Carlo events in the amplitude fitter produces surprisingly little change in the fit results.

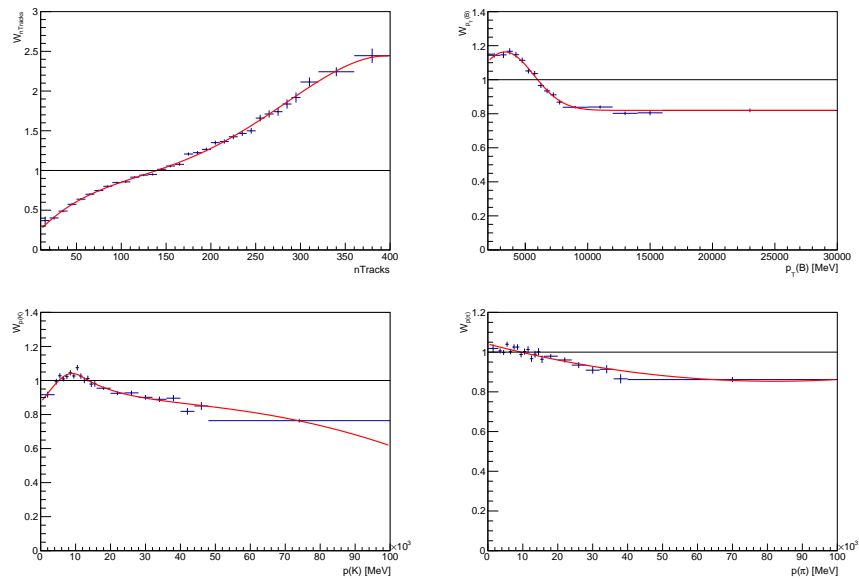


Figure 24: Plots of the fits (in red) used to re-weight $\psi(2S)$ MC. Weights are applied in order (left to right, top to bottom): the number of tracks, the transverse momentum of B, the momentum of the kaon, and the momentum of the pion.

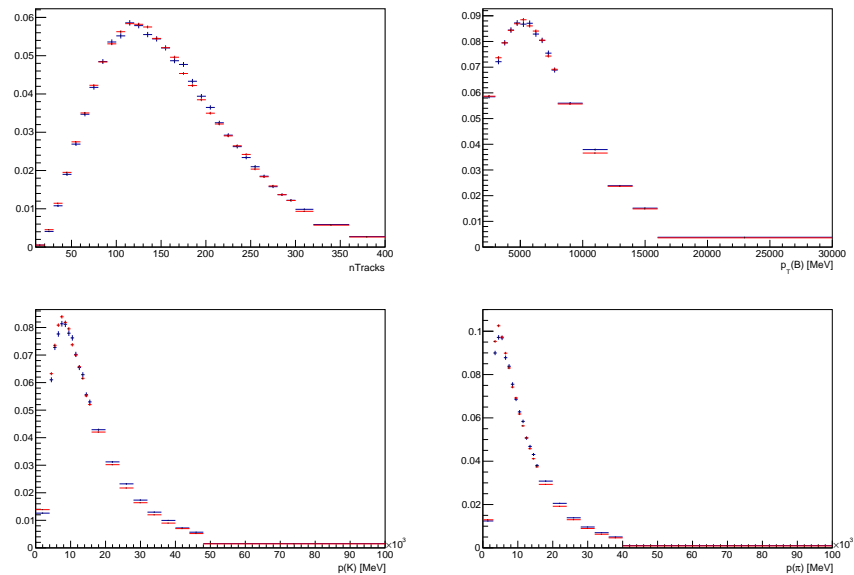


Figure 25: Comparison of MC (red) and data (black) for the 4 kinematic variables used to re-weight MC after complete re-weighting for $\psi(2S)$.

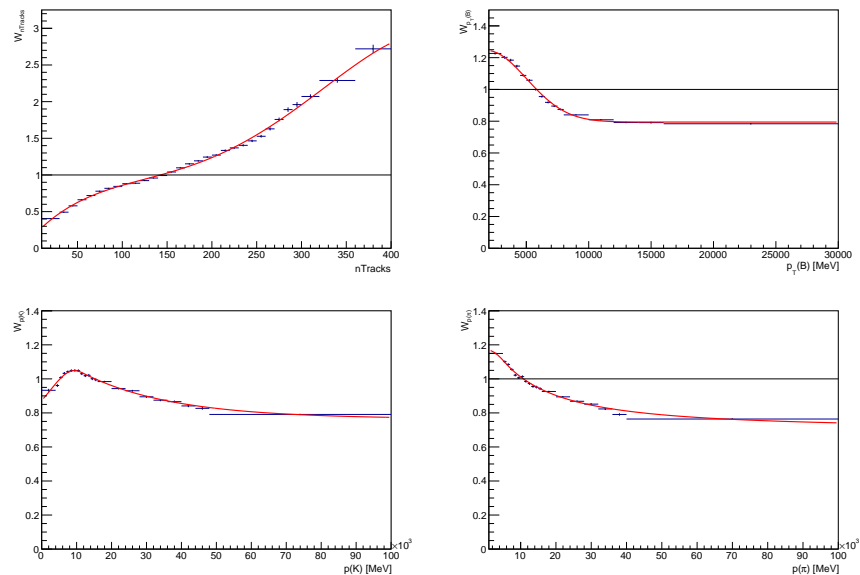


Figure 26: Plots of the fits (in red) used to re-weight J/ψ MC. Weights are applied in order (left to right, top to bottom): the number of tracks, the transverse momentum of B, the momentum of the kaon, and the momentum of the pion.

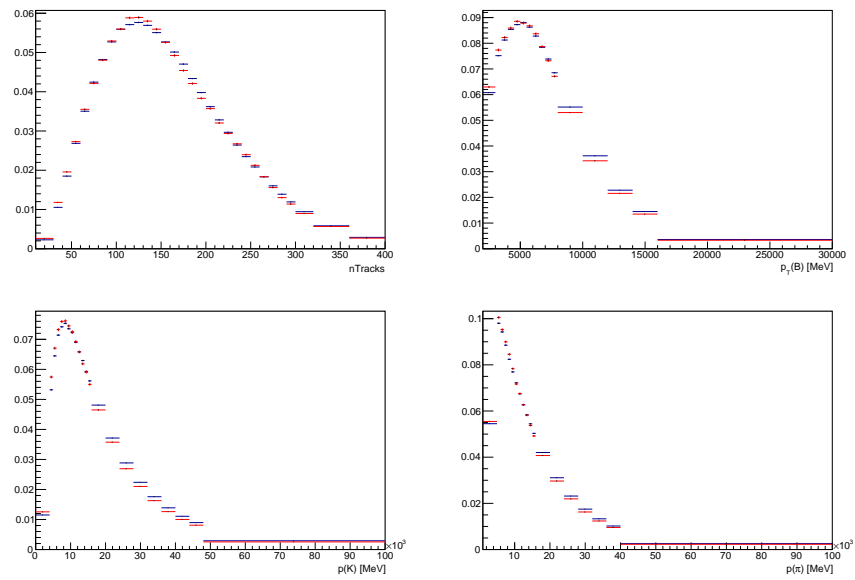


Figure 27: Comparison of MC (red) and data (black) for the 4 kinematic variables used to re-weight MC after complete re-weighting for J/ψ .

6 Amplitude Analysis

The goal of this analysis is to search for exotic charmonium-like Z states by fitting a model of the $B \rightarrow \psi(2S)\pi K$ or $B \rightarrow J/\psi\pi K$ amplitude to the data. First, we evaluate how well the data can be described without Z contributions, *i.e.* with the $K^{*0} \rightarrow K^+\pi^-$ contributions alone. We then add a Z contribution to the amplitude model, allowing us to obtain the significance of the Z signal from the change of the fit quality, repeating the process for every additional Z contribution. Since the model depends on the Z spin and parity, we repeat this exercise for $J^P(Z) = 0^-, 1^+, 1^-, 2^-$ and 2^+ . Higher spin values are not plausible, and 0^+ hypothesis is forbidden by parity conservation in the $Z \rightarrow \psi\pi$ decay. Thus, if a significant Z contribution is found, we can in principle also identify its quantum numbers, in addition to determining its mass and width. The fit fraction (fraction of the integral of the amplitude squared) corresponding to each Z is determined as well. This process will be repeated for $Z_K \rightarrow K\pi$ as well. We evaluate the K^{*0} model dependence and other systematic effects.

For the best use of the information contained in the data sample, we use an unbinned maximum likelihood method. In this section, we describe the fit formalism.

6.1 Fitted kinematic variables

Let us first consider the $B^0 \rightarrow \psi K^{*0}$, $\psi \rightarrow \mu^+\mu^-$, $K^{*0} \rightarrow K^+\pi^-$ decay chain, which we call the “ K^{*0} decay chain”. There are four independent variables which completely describe the kinematics of this decay:

$m_{K\pi}$, the $K^+\pi^-$ invariant mass, *i.e.* the K^{*0} mass;

$\cos\theta_{K^*}$, the cosine of the K^{*0} helicity angle, which is the angle between the K^- direction and the B^0 direction in the K^{*0} rest frame; equivalently, one can think about this degree of freedom as $m_{\psi\pi}$, since for a given value of $m_{K\pi}$, there is a one-to-one transformation between $\cos\theta_{K^*}$ and $m_{\psi\pi}$;

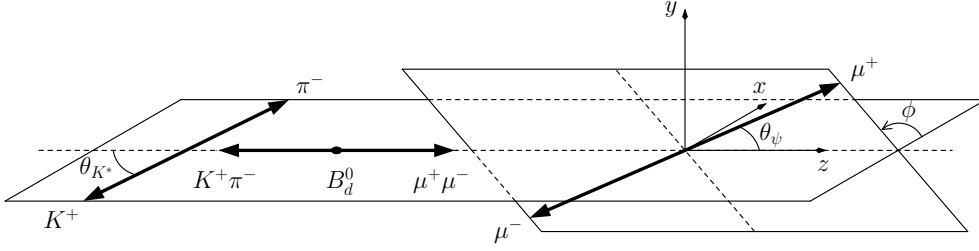


Figure 28: Definition of the helicity angles. See text for details.

$\cos \theta_{J/\psi}$, the cosine of the ψ helicity angle, which is the angle between the μ^+ direction and the B^0 direction in the ψ rest frame;

ϕ , the angle between the decay planes of the K^{*0} and ψ mesons.

The angles are defined in Fig. 28. For the parameterization of the amplitude of the $B^0 \rightarrow ZK^+$, $Z \rightarrow \psi\pi^-$, $\psi \rightarrow \mu^+\mu^-$ decay chain (which we call the “ Z decay chain”), the following four different variables are more convenient:

$m_{\psi\pi}$, the $\psi\pi^-$ invariant mass, *i.e.* the Z mass;

$\cos \theta_Z$, the cosine of the Z helicity angle, which is the angle between the π^- direction and the B^0 direction in the Z rest frame; equivalently, one can think about this degree of freedom as $m_{K\pi}$, since for a given value of $m_{\psi\pi}$ there is a one-to-one transformation between $\cos \theta_Z$ and $m_{K\pi}$;

$\cos \theta_{J/\psi}^Z$, the cosine of the ψ helicity angle, which is the angle between the μ^+ direction and the Z direction in the ψ rest frame;

ϕ^Z , the angle between the decay planes of the Z and ψ mesons.

These four variables can be determined with no ambiguities from the four variables used to describe the K^{*0} decay chain, thus they do not represent independent degrees of freedom.

The ψ helicity frame used to describe the Z decay chain is not the same as the ψ helicity frame used to describe the K^{*0} decay chain. To properly describe the interference between the K^{*0} and Z amplitudes, a rotation is needed, which again can be calculated from the 4 variables describing the K^{*0} decay.

Similarly, to parameterize the amplitude of the $B^0 \rightarrow Z\pi^-, Z \rightarrow \psi K^+, \psi \rightarrow \mu^+\mu^-$ decay chain (which we call the “ Z_K decay chain”), these four different variables are more convenient:

$m_{\psi K}$, the ψK^+ invariant mass, *i.e.* the Z_K mass;

$\cos\theta_{Z_K}$, the cosine of the Z_K helicity angle, which is the angle between the K^+ direction and the B^0 direction in the Z_K rest frame; equivalently, one can think about this degree of freedom as $m_{K\pi}$, since for a given value of $m_{\psi K}$ there is a one-to-one transformation between $\cos\theta_{Z_K}$ and $m_{K\pi}$;

$\cos\theta_{J/\psi}^{Z_K}$, the cosine of the ψ helicity angle, which is the angle between the μ^+ direction and the Z_K direction in the ψ rest frame;

ϕ^{Z_K} , the angle between the decay planes of the Z_K and ψ mesons.

Much like those in the Z decay chain, these four variables can be determined with no ambiguities from the four variables used to describe the K^{*0} decay chain, and thus they also do not represent independent degrees of freedom.

Again, the ψ helicity frame used to describe the Z_K decay chain is not the same as the ψ helicity frame used to describe the K^{*0} decay chain or the Z decay chain. To properly describe the interference between the amplitudes in the other chains, a rotation is needed, which again can be calculated from the 4 variables describing the K^{*0} decay.

The prescription to calculate the angles defined above, including treatment of \bar{B}^0 vs. B^0 candidates, can be found in the appendix of Ref. [12].⁶

To maximize the statistical and systematic sensitivity of our analysis, we perform four-dimensional (4D) fits to the data. We will denote the fitted variables as \vec{v} . It is a technical

⁶In Belle’s notation $\tilde{\phi} = \phi^Z$.

detail which four independent variables are chosen in the formulation of the matrix element. For example, they can be chosen as $\vec{v} = (m_{K\pi}^2, m_{\psi\pi}^2, \cos\theta_{J/\psi}, \phi)$ (choice A) or $\vec{v} = (m_{K\pi}, \cos\theta_{K^*}, \cos\theta_{J/\psi}, \phi)$ (choice B). Choice A corresponds to phase-space variables, of which the probability density function (PDF) is given by⁷

$$P(\vec{v}) \propto |\mathcal{M}(\vec{v})|^2, \quad (2)$$

where \mathcal{M} is the matrix element for the decay. With no dynamics and only spinless particles (which is not the case here), the PDF would be constant for this choice. For choice B, known as “rectangular phase-space” variables,

$$P(\vec{v}) \propto |\mathcal{M}(\vec{v})|^2 p(\vec{v}) q(\vec{v}), \quad (3)$$

where p (q) is the daughter momentum in the B^0 (K^{*0}) rest frame. In this case, phase-space distribution is proportional to $p(\vec{v}) q(\vec{v})$.

6.2 The signal PDF

The signal probability density function, P_{sig} , is a four-dimensional function that depends on $m_{K\pi}$ and the set of independent angular variables in the K^* decay chain labeled Ω . The combined set of the four independent variables is labeled as \vec{v} . The PDF also depends on the fit parameters, referred to as ω , that consists of the helicity couplings, masses and widths of resonances, etc. The signal PDF is given by

$$P_{\text{sig}}(\vec{v}|\vec{\omega}) = \frac{1}{I(\vec{\omega})} |\mathcal{M}(\vec{v}|\vec{\omega})|^2 \Phi(m_{K\pi}) \epsilon(\vec{v}), \quad (4)$$

where $\Phi(m_{K\pi}) = pq$ is the phase space function (with p as the momentum of the $K\pi$ system in the B^0 rest frame and q as the momentum of the K in the K^* rest frame). $\epsilon(\vec{v})$ is

⁷See e.g. PDG write-up on kinematics in three-body decays.

the signal efficiency, and $I(\vec{\omega})$ is the normalization integral:

$$I(\vec{\omega}) = \int P_{\text{sig}}(\vec{v}|\vec{\omega})d\vec{v} \propto \frac{\sum_j w_j^{\text{MC}} |\mathcal{M}(\vec{v}_j|\vec{\omega})|^2}{\sum_j w_j^{\text{MC}}}, \quad (5)$$

where the sum is over the fully simulated Monte Carlo events, which are generated uniformly over the entire phase-space. In the case of $B^0 \rightarrow J/\psi\pi^-K^+$, a large portion of MC was produced with non-uniform $m_{K\pi}$ distribution, enhancing rates in the $K^*(892)$ and $K_2(1430)$ peak regions (so called "XLL model"). We reweight this MC sample to being uniformly spread in phase space. This procedure folds in the efficiency corrections without a need to parameterize them, as can be seen in the equation for the log-likelihood sum over the N data events:

$$\begin{aligned} \ln L(\vec{\omega}) &= \sum_i \ln P_{\text{sig}}(\vec{v}_i|\vec{\omega}) \\ &= \sum_i \ln |\mathcal{M}(\vec{v}_i|\vec{\omega})|^2 - N \ln I(\vec{\omega}) + \sum_i \ln[\Phi(m_{K\pi,i})\epsilon(\vec{v}_i)]. \end{aligned} \quad (6)$$

Since the last term has no dependence on $\vec{\omega}$, it can be dropped.

6.3 Background treatment

The backgrounds in the data are small, but not completely negligible. As such, we construct a probability density function for the background events (P_{bkg}) using the sidebands of the B^0 mass peak and add it to the probability density for the signal events to construct the total

PDF (so called "cFit" procedure). Thus, the log-likelihood is defined as:

$$\begin{aligned}
-\ln L(\vec{\omega}) &= \sum_i \ln[(1 - f_{\text{bkg}})P_{\text{sig}}(\vec{v}_i|\vec{\omega}) + f_{\text{bkg}}P_{\text{bkg}}(\vec{v}_i)] \\
&= -\sum_i \ln \left[(1 - f_{\text{bkg}}) \frac{|\mathcal{M}(\vec{v}_i|\vec{\omega})|^2 \Phi(m_{K\pi,i}) \epsilon(\vec{v}_i)}{I(\vec{\omega})} + f_{\text{bkg}} \frac{P_{\text{bkg}}^u(\vec{v}_i)}{I_{\text{bkg}}} \right] \\
&= -\sum_i \ln \left\{ \frac{(1 - f_{\text{bkg}}) \Phi(m_{K\pi,i}) \epsilon(\vec{v}_i)}{I(\vec{\omega})} \left[|\mathcal{M}(\vec{v}_i|\vec{\omega})|^2 + \frac{f_{\text{bkg}} I(\vec{\omega})}{(1 - f_{\text{bkg}}) I_{\text{bkg}}} \frac{P_{\text{bkg}}^u(\vec{v}_i)}{\Phi(m_{K\pi,i}) \epsilon(\vec{v}_i)} \right] \right\} \\
&= -\sum_i \ln \left[|\mathcal{M}(\vec{v}_i|\vec{\omega})|^2 + \frac{f_{\text{bkg}} I(\vec{\omega})}{(1 - f_{\text{bkg}}) I_{\text{bkg}}} \frac{P_{\text{bkg}}^u(\vec{v}_i)}{\Phi(m_{K\pi,i}) \epsilon(\vec{v}_i)} \right] + N \ln I(\vec{\omega}) + \text{constant},
\end{aligned} \tag{7}$$

where N is the number of events, f_{bkg} is the background fraction in the peak region (as described in Sec. 4.2), $I(\vec{\omega})$ is the normalization integral of the signal PDF (Eq. 5), $P_{\text{bkg}}^u(m_{K\pi}, \Omega)$ is the unnormalized background density proportional to the density of the side-band events, with its normalization determined by:

$$I_{\text{bkg}} \equiv \int P_{\text{bkg}}^u(\vec{v}) d\vec{v} \propto \frac{\sum_i w_i^{\text{MC}} \frac{P_{\text{bkg}}^u(\vec{v}_i)}{\Phi(m_{K\pi,i}) \epsilon(\vec{v}_i)}}{\sum_i w_i^{\text{MC}}}. \tag{8}$$

The background term is then efficiency-corrected so it can be added to the efficiency-independent signal probability expressed by the matrix element squared, $|\mathcal{M}|^2$. Thus, the signal efficiency parameterization, $\epsilon(\vec{v})$, becomes only a part of the background parameterization, and therefore only affects a small part of the total PDF ($< 3\%$ for both channels).

We assume that the efficiency used in the background parameterization approximately factorizes as

$$\epsilon(m_{K\pi}^2, m_{\psi\pi}^2, \cos \theta_{J/\psi}, \phi) = \epsilon_1(m_{K\pi}^2, m_{\psi\pi}^2) \cdot \epsilon_2(\cos \theta_{J/\psi} | m_{K\pi}) \cdot \epsilon_3(\phi | m_{K\pi}). \tag{9}$$

To avoid dealing with inconvenient kinematic boundaries in the $(m_{K\pi}^2, m_{\psi\pi}^2)$ Dalitz plane, we use the "rectangular" Dalitz plane $(m_{K\pi}, \cos \theta_{K^*})$ in the parameterization of the first term:

$$\epsilon_1(m_{K\pi}^2, m_{\psi\pi}^2) = \epsilon'_1(m_{K\pi}, \cos \theta_{K^*}). \tag{10}$$

To obtain ϵ'_1 , we make a 2D histogram in $(m_{K\pi}, \cos\theta_{K^*})$ of the fully simulated phase-space MC passed through the analysis procedure. At the generator level, the phase-space is flat in $\cos\theta_{K^*}$ and has a $p \cdot q$ dependence on $m_{K\pi}$. Therefore, we weight each MC event by $1/(p \cdot q)$ to obtain the efficiency variation. We smooth the efficiency histogram, and then interpolate between the bin centers using bi-cubic interpolation. This efficiency is displayed in Fig. 29 and Fig. 36 for $\psi(2S)$ and J/ψ channels, respectively, in both the $(m_{K\pi}, \cos\theta_{K^*})$ and $(m_{K\pi}^2, m_{\psi\pi}^2)$ planes. The efficiencies for ϵ_2 and ϵ_3 are displayed in Figs. 30 and 31 for the $\psi(2S)$ channel and in Figs. 37 and 38 for the J/ψ channel.

The background PDF, $P_{\text{bkg}}^u(\vec{v})/\Phi(m_{K\pi})$, is parameterized using exactly the same approach as $\epsilon(\vec{v})$, except that the histograms are not smoothed because of the presence of a narrow $K^{*0}(892)$ peak in the data:

$$\frac{P_{\text{bkg}}^u(\vec{v})}{\Phi(m_{K\pi})} = P_{\text{bkg}1}^u(m_{K\pi}, \cos\theta_{K^*}) \cdot P_{\text{bkg}2}^u(\cos\theta_{J/\psi}|m_{K\pi}) \cdot P_{\text{bkg}3}^u(\phi|m_{K\pi}). \quad (11)$$

The individual terms are displayed in Figs. 32, 33, and 34 for the $\psi(2S)$ channel, while for the J/ψ channel, the individual background parameterization distributions are displayed in Figs. 39, 40, and 41. Comparisons between the background from data and the parameterized background for each of the variables used in the efficiency functions are represented in Figs. 35 and 42 for $\psi(2S)$ and J/ψ channels, respectively.

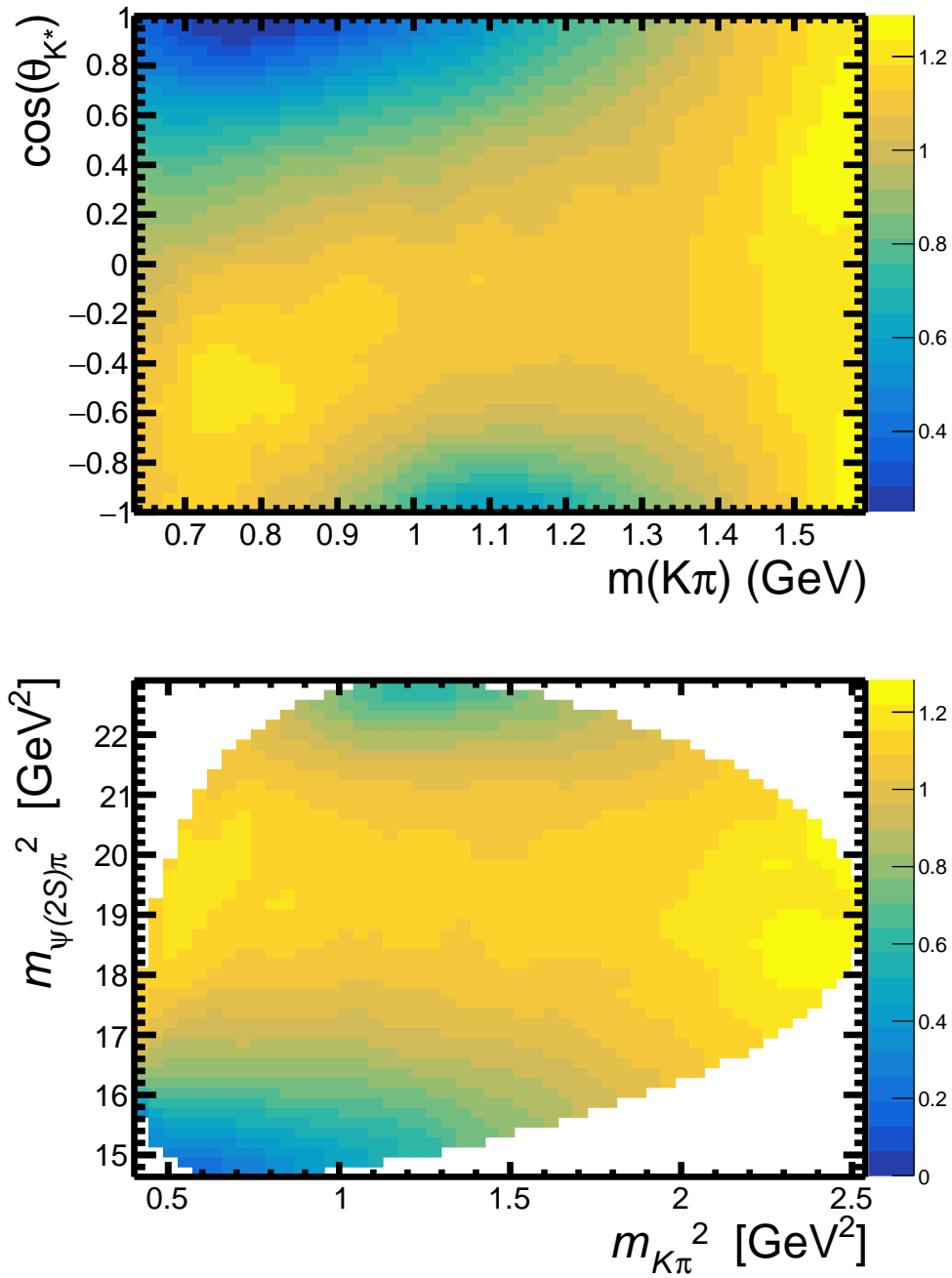


Figure 29: Parameterized efficiency (ϵ_1) in the rectangular Dalitz plane ($m_{K\pi}, \cos\theta_{K^*}$) (top) and in the traditional Dalitz plane ($m_{K\pi}^2, m_{J/\psi\pi}^2$) (bottom) for the $\psi(2S)$ channel. The normalization arbitrarily corresponds to an average efficiency of 1 over the phase-space.

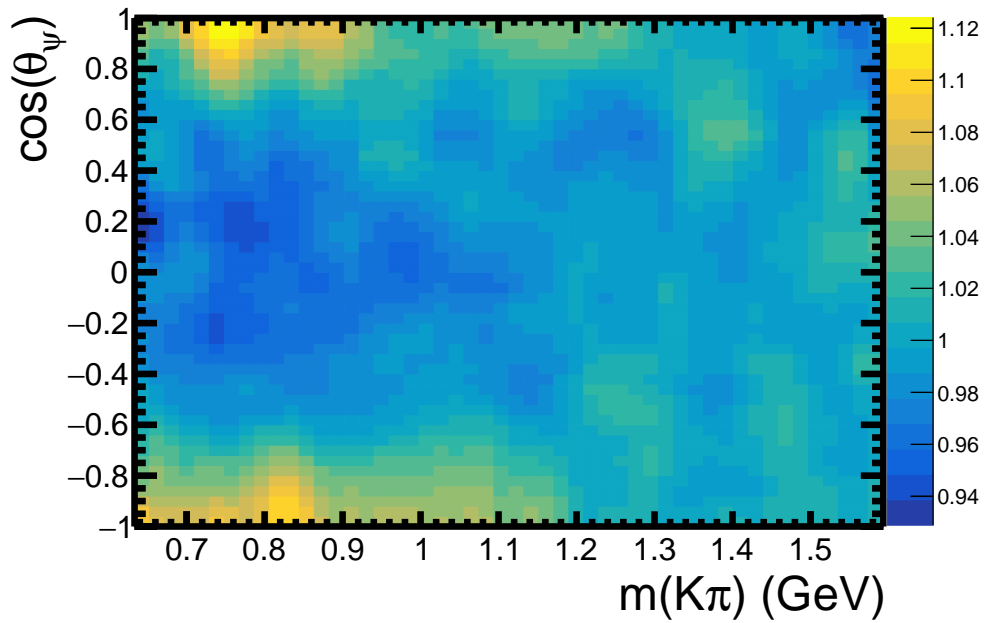


Figure 30: Parameterized efficiency function $\epsilon_2(\cos\theta_\psi - m_{K\pi})$ for the $\psi(2S)$ channel. By construction it integrates to 1.0 at each $m_{K\pi}$ value.

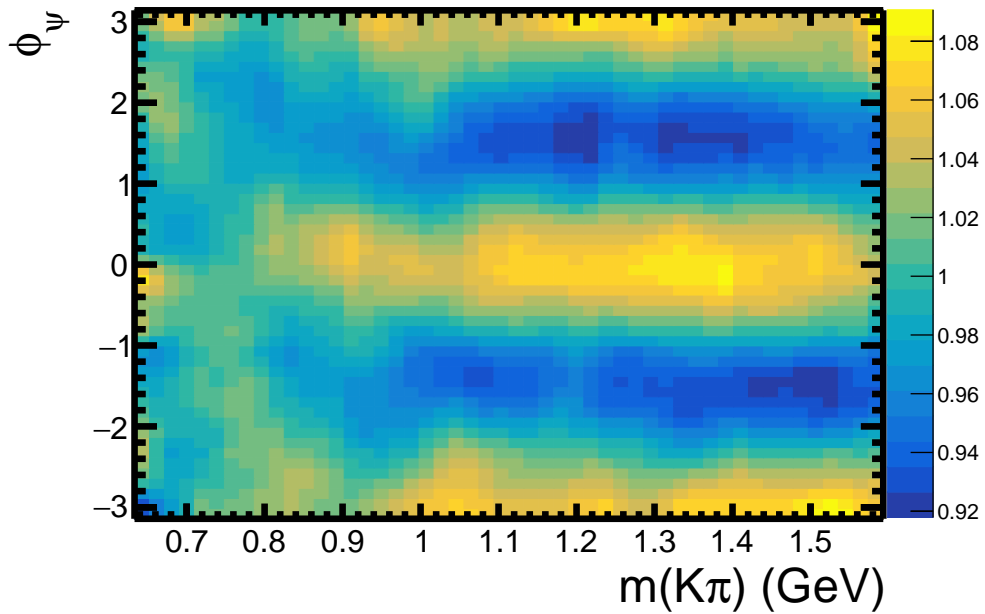


Figure 31: Parameterized efficiency function $\epsilon_3(\phi - m_{K\pi})$ for the $\psi(2S)$ channel. By construction it integrates to 1.0 at each $m_{K\pi}$ value.

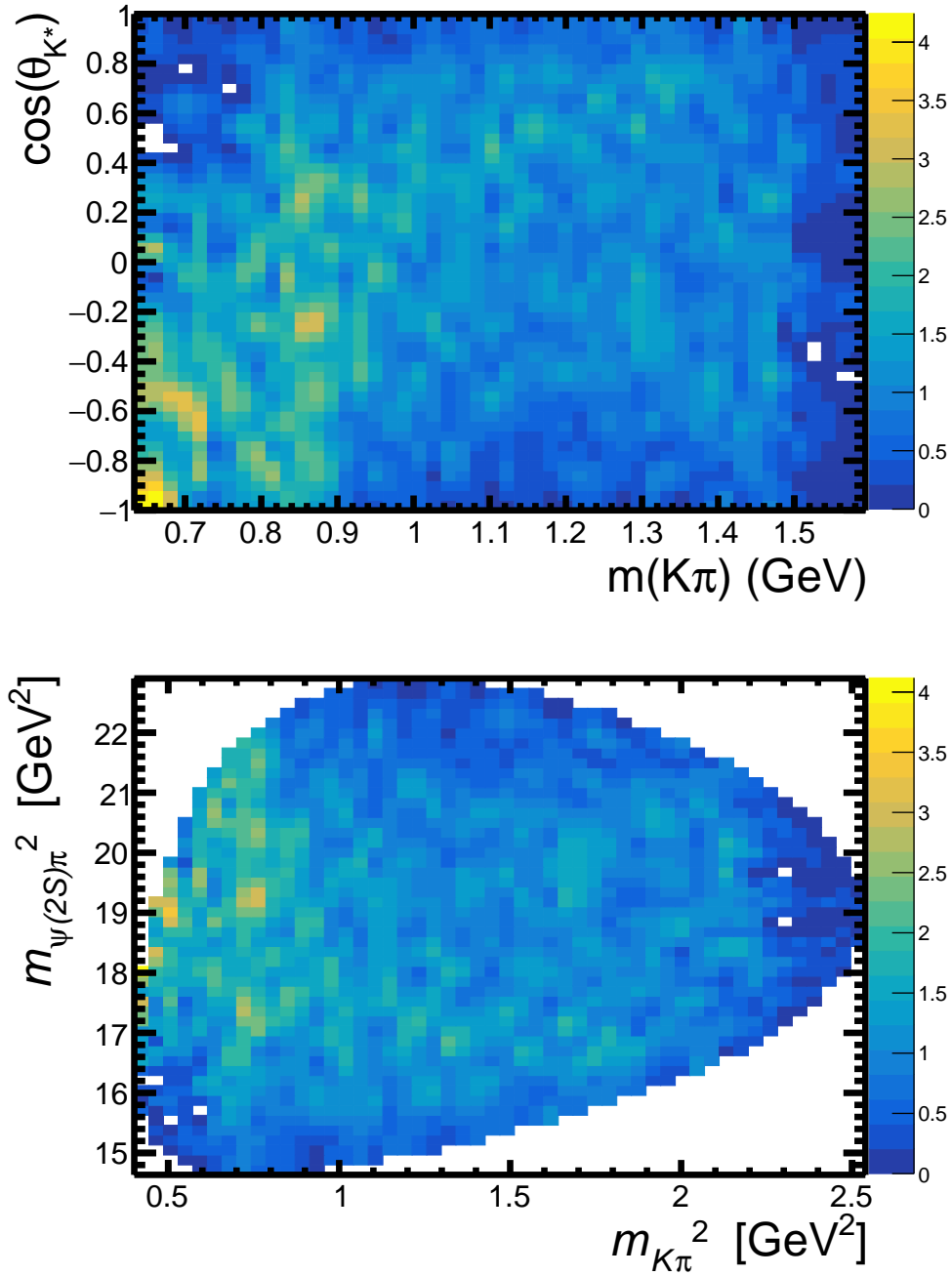


Figure 32: Parameterized background distributions in the rectangular Dalitz plane ($m_{K\pi}, \cos\theta_{K^*}$) (top) and in the traditional Dalitz plane ($m_{K\pi}^2, m_{J/\psi\pi}^2$) (bottom) for the $\psi(2S)$ channel. The normalization arbitrarily corresponds to an average efficiency of 1 over the phase-space.

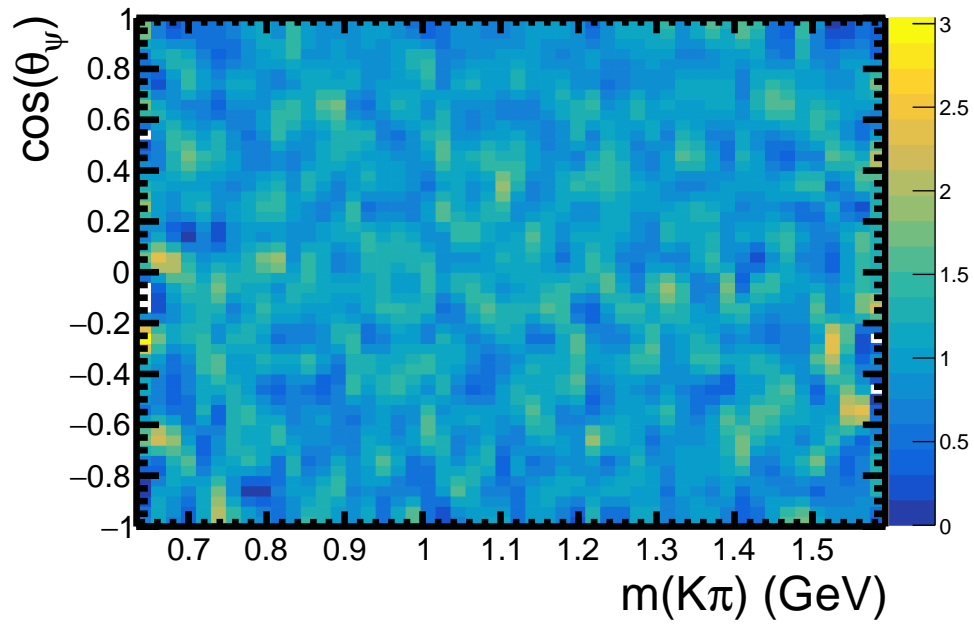


Figure 33: Parameterized background distribution for $(\cos\theta_\psi|m_{K\pi})$ for the $\psi(2S)$ channel. By construction it integrates to 1.0 at each $m_{K\pi}$ value.

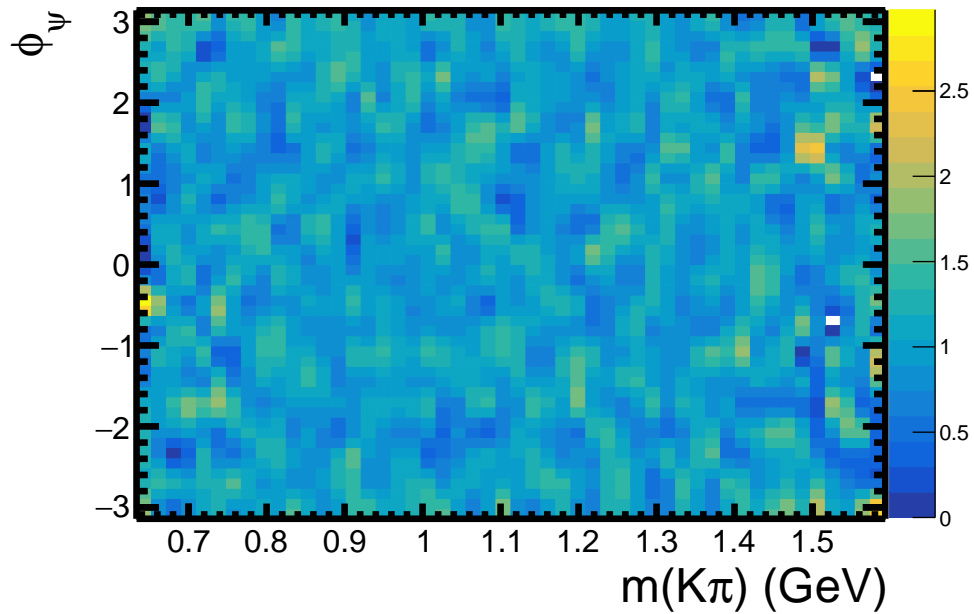


Figure 34: Parameterized background distribution for $(\phi_\psi|m_{K\pi})$ for the $\psi(2S)$ channel. By construction it integrates to 1.0 at each $m_{K\pi}$ value.

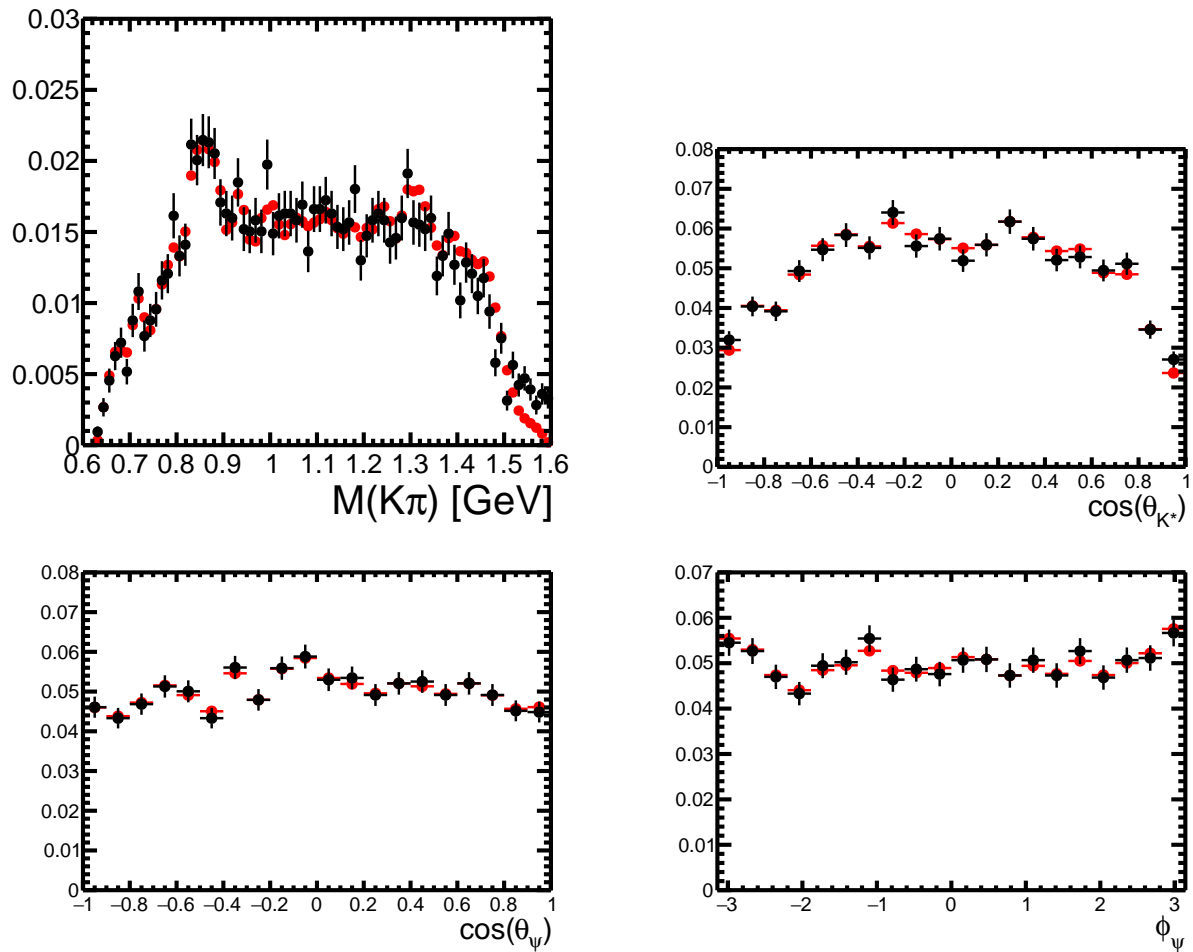


Figure 35: Comparison of the background from the data (black) and the parameterized background (red) for $m_{K\pi}$, $\cos\theta_{K^*}$, $\cos\theta_\psi$, and ϕ_ψ for the $\psi(2S)$ channel.

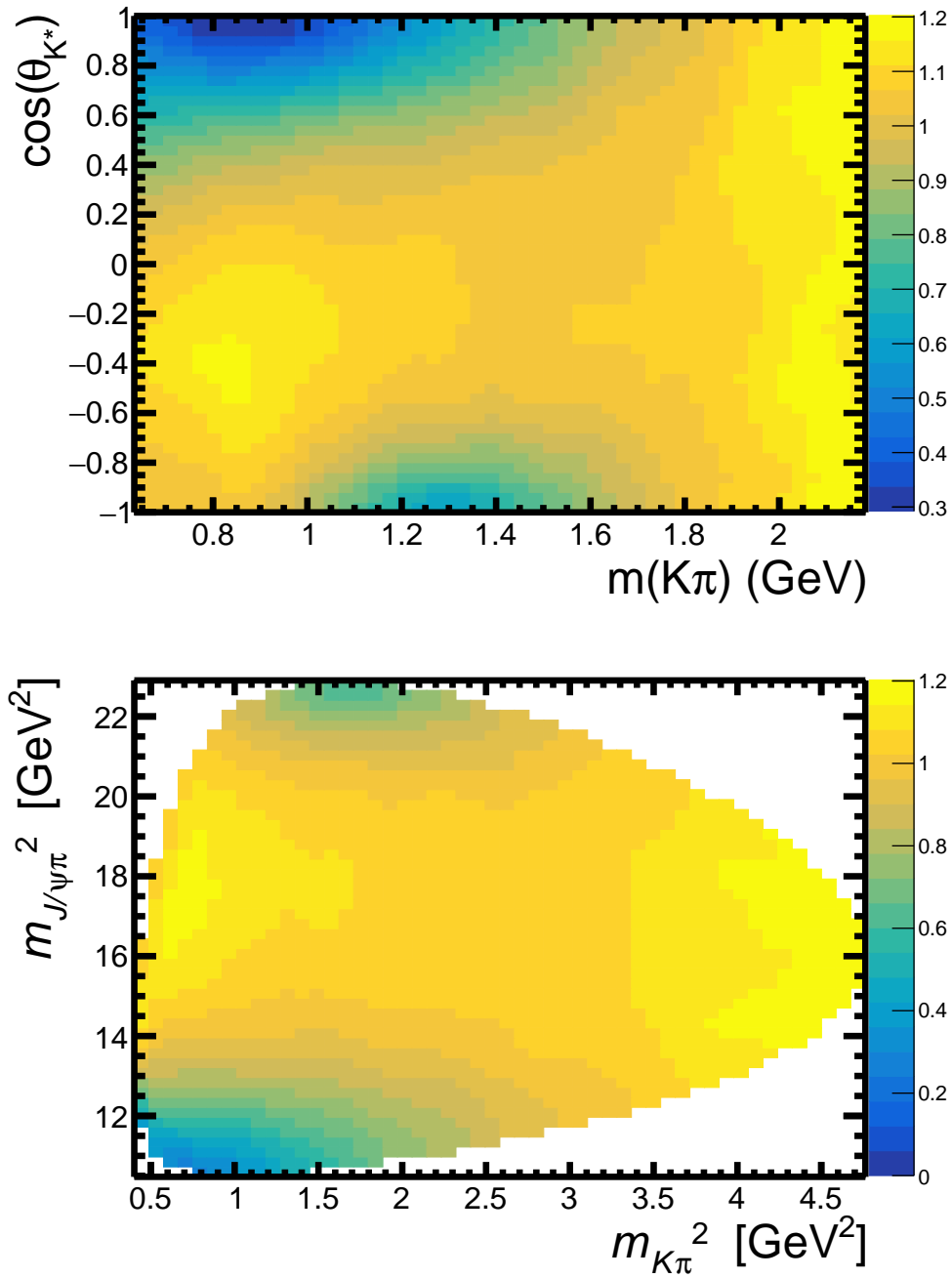


Figure 36: Parameterized efficiency (ϵ_1) in the rectangular Dalitz plane ($m_{K\pi}, \cos\theta_{K^*}$) (top) and in the traditional Dalitz plane ($m_{K\pi}^2, m_{J/\psi\pi}^2$) (bottom) for the J/ψ channel. The normalization arbitrarily corresponds to an average efficiency of 1 over the phase-space.

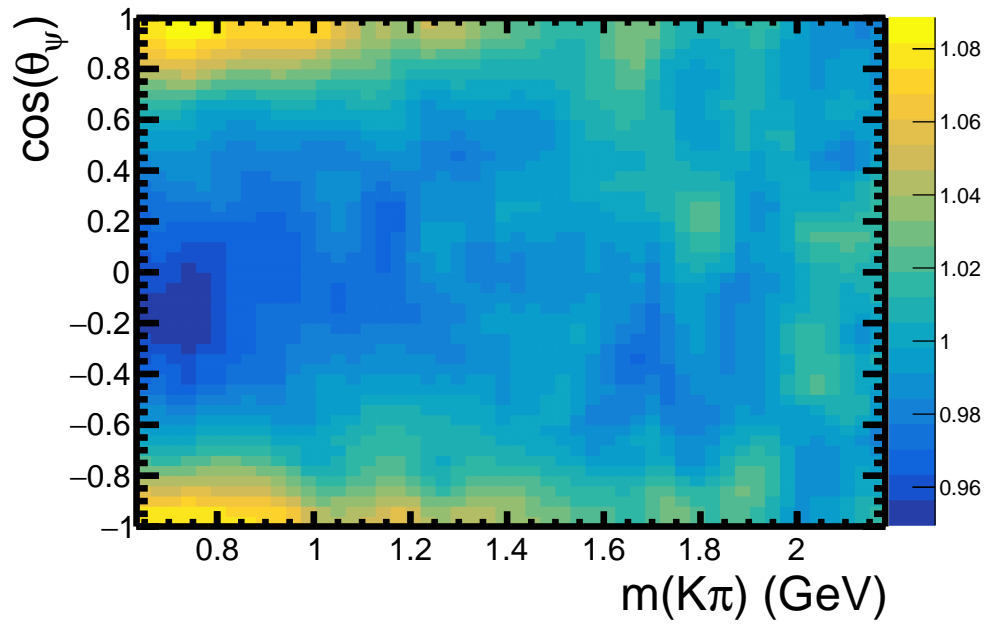


Figure 37: Parameterized efficiency function $\epsilon_2(\cos\theta_\psi - m_{K\pi})$ for the J/ψ channel. By construction it integrates to 1.0 at each $m_{K\pi}$ value.

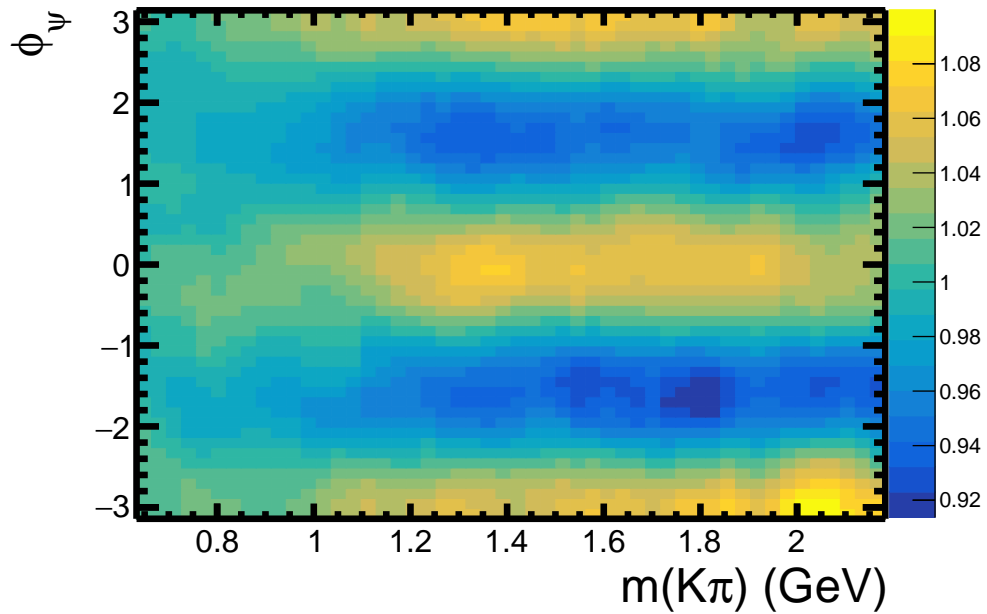


Figure 38: Parameterized efficiency function $\epsilon_3(\phi - m_{K\pi})$ for the J/ψ channel. By construction it integrates to 1.0 at each $m_{K\pi}$ value.

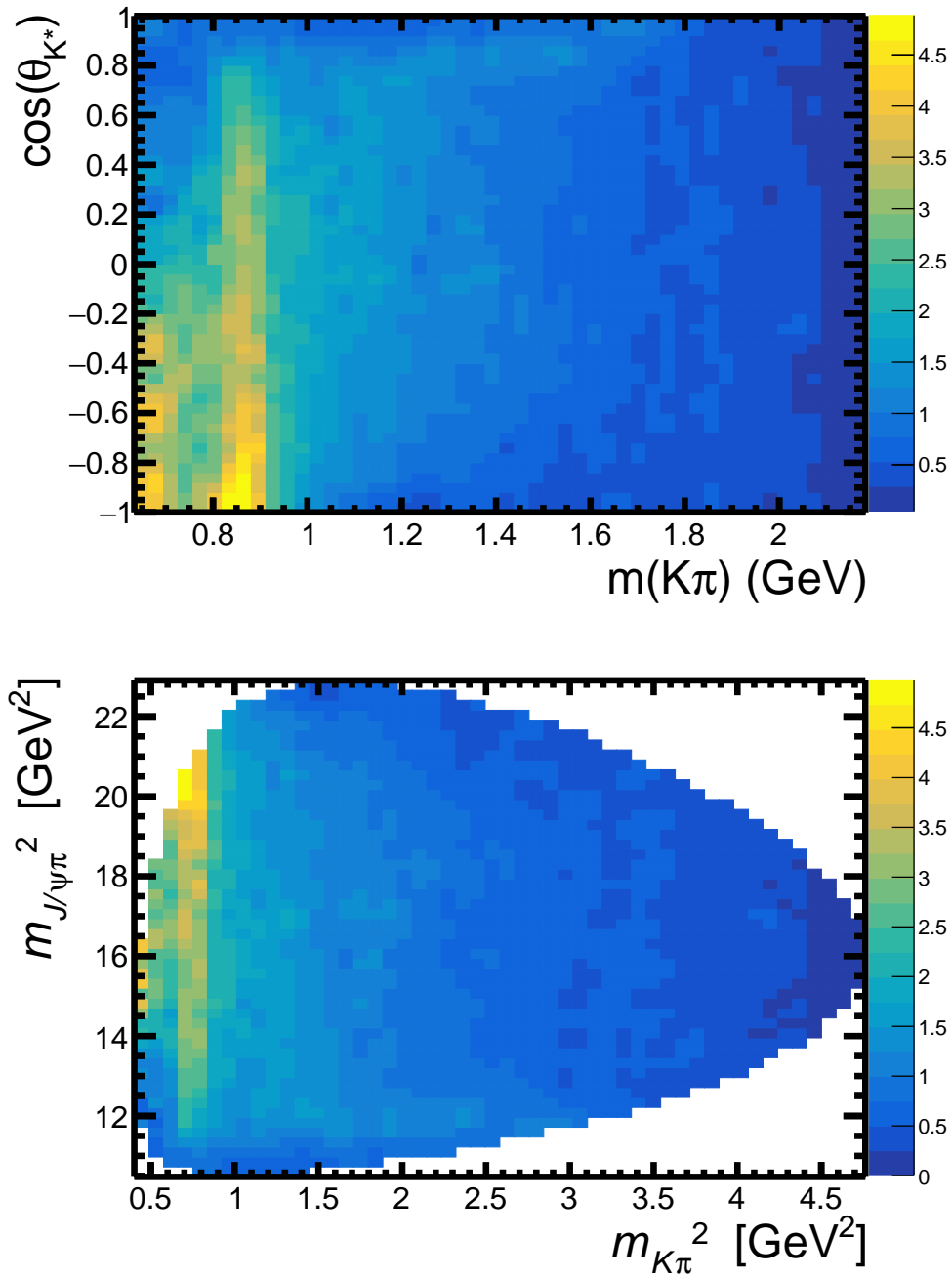


Figure 39: Parameterized background distributions in the rectangular Dalitz plane ($m_{K\pi}, \cos\theta_{K^*}$) (top) and in the traditional Dalitz plane ($m_{K\pi}^2, m_{J/\psi\pi}^2$) (bottom) for the J/ψ channel. The normalization arbitrarily corresponds to an average efficiency of 1 over the phase-space.

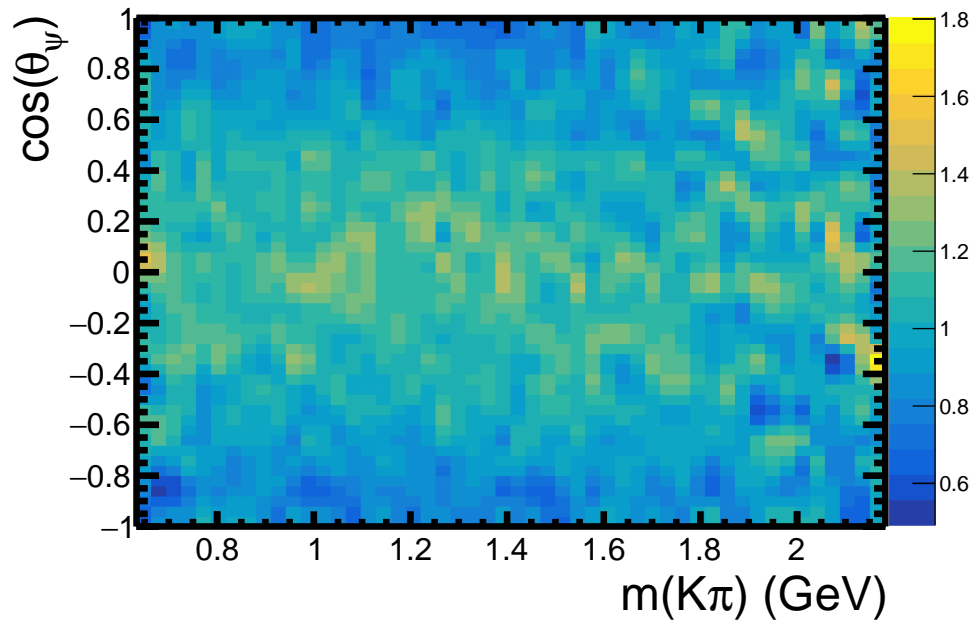


Figure 40: Parameterized background distribution for $(\cos\theta_\psi|m_{K\pi})$ for the J/ψ channel. By construction it integrates to 1.0 at each $m_{K\pi}$ value.

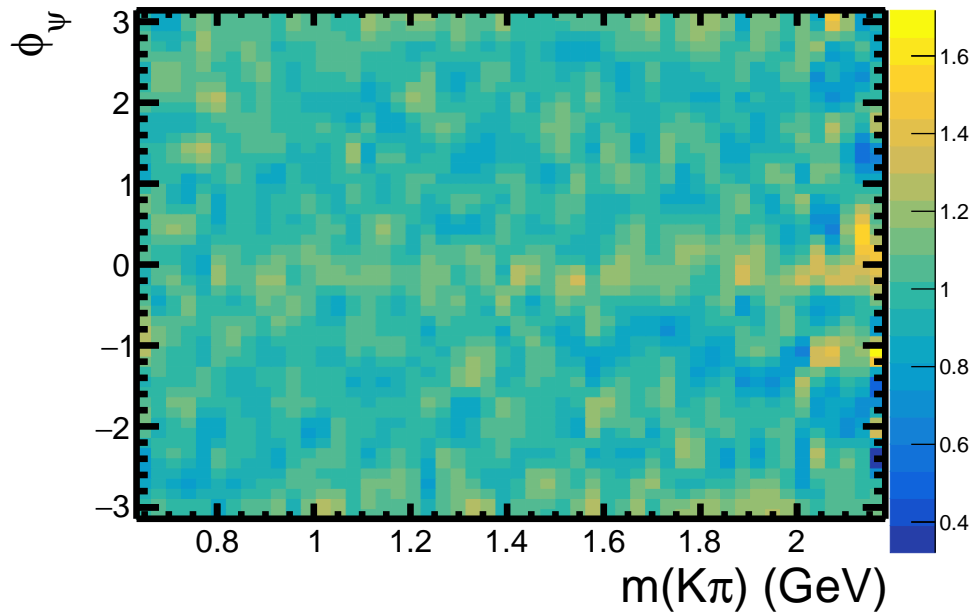


Figure 41: Parameterized background distribution for $(\phi_\psi|m_{K\pi})$ for the J/ψ channel. By construction it integrates to 1.0 at each $m_{K\pi}$ value.

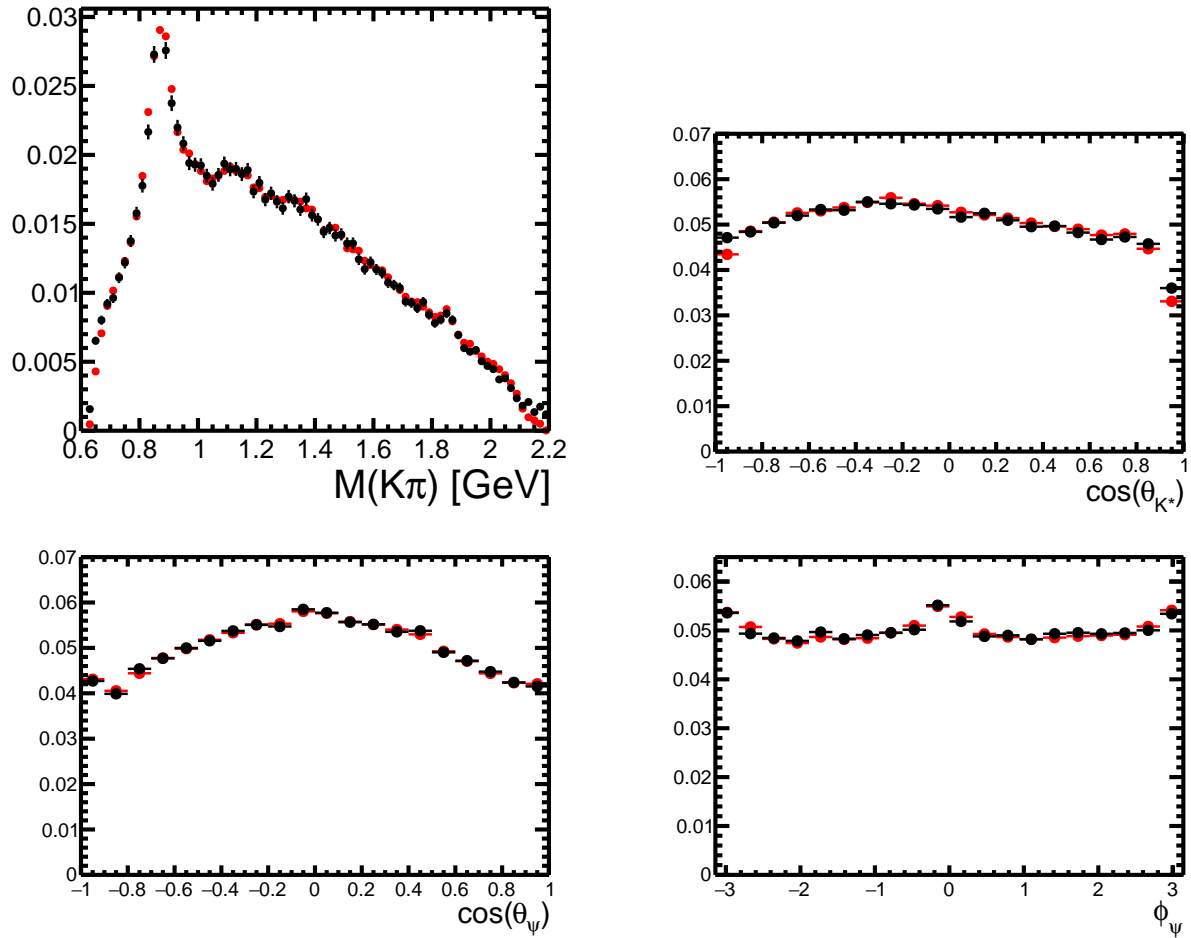


Figure 42: Comparison of the background from the data (black) and the parameterized background (red) for $m_{K\pi}$, $\cos\theta_{K^*}$, $\cos\theta_\psi$, and ϕ_ψ for the J/ψ channel.

7 Helicity formalism and matrix element

The matrix element formulation is the same as used in the 2014 analysis [19] and by Belle in their 4D [12] analysis.

We use the helicity basis for the decay amplitudes, and here we will describe how this is applied to the K^* , Z , and Z_K decay chains and how the complete matrix element is constructed such that the different decay chains can interfere.

7.1 Helicity formalism for the K^* decay chain

In the K^* decay chain, B decays proceed via $B \rightarrow \psi K^*$, $K^* \rightarrow \pi K$, with ψ decaying via $\psi \rightarrow \mu\mu$. The K^* resonances that are expected to contribute are discussed later.

The full 4D matrix element for the K^* decay chain is given by:

$$|\mathcal{M}^{K^*}|^2 \equiv \sum_{\Delta\lambda_\mu=-1,1} \left| \sum_{\lambda_\psi=-1,0,1} \sum_k \underline{A}_{k,\lambda_\psi} R(m_{K\pi}|m_{0k}, \Gamma_{0k}) d_{\lambda_\psi,0}^{J_k}(\theta_{K^*}) e^{i\lambda_\psi\phi} d_{\lambda_\psi,\Delta\lambda_\mu}^1(\theta_\psi) \right|^2, \quad (12)$$

where k indexes specific K^* contributions (i.e. $K^*(892)^0$, $K_2^*(1430)^0$, \dots). J_k is the spin of the given K^* resonance. λ is helicity of the particle (projection of the particle spin onto its momentum in the rest frame of its parent) and $\Delta\lambda_\mu \equiv \lambda_{\mu^+} - \lambda_{\mu^-}$. $d_{\lambda_1,\lambda_2}^J(\theta)$ denotes Wigner d -functions. $R(m_{K\pi})$ is the mass dependence of the contribution, which is discussed in greater detail later. The fitted helicity amplitudes are complex numbers,

$$\underline{A}_{k,\lambda_\psi} = A_{k,\lambda_\psi} \exp(i\phi_{k,\lambda_\psi}) = \text{Re}\underline{A}_{k,\lambda_\psi} + i\text{Im}\underline{A}_{k,\lambda_\psi}, \quad (13)$$

therefore, each amplitude must be represented by two free parameters in the fit, for which we use the real and imaginary components. For K^* resonances with $J_k = 0$, only $\lambda_\psi = 0$ is allowed, as $\lambda_\psi = \lambda_{K^*}$ as $J_B = 0$, thus there is only one complex amplitude to fit. For K^* resonances with $J_k > 0$, there are 3 complex amplitudes to fit, totaling 6 free parameters. To

fix the magnitude and phase conventions, we set

$$\text{Re}\underline{A}_{K^*(892)^0,0} = 1, \quad \text{Im}\underline{A}_{K^*(892)^0,0} = 0. \quad (14)$$

7.2 Helicity formalism for the Z decay chain

In the Z decay chain, B decays proceed via $B \rightarrow ZK$, $Z \rightarrow \psi\pi$, with ψ decaying via $\psi \rightarrow \mu\mu$. The matrix element for the Z decay chain is formed in a similar way as in the K^* chain. The Z decay chain also picks up a rotation to align the muon frames to the ones used in the K^* decay chain and allows for the proper description of interference between the decay chains. The full 4D matrix element for the Z chain is given by:

$$|\mathcal{M}^Z|^2 \equiv \sum_{\Delta\lambda_\mu=-1,1} \left| \sum_{\lambda_\psi^Z=-1,0,1} \sum_n \underline{A}_{n,\lambda_\psi^Z} R(m_{\psi\pi}|m_{0n}, \Gamma_{0n}) d_{0,\lambda_\psi^Z}^{J_n}(\theta_Z) e^{i\lambda_\psi^Z \phi^Z} d_{\lambda_\psi^Z, \Delta\lambda_\mu}^1(\theta_\psi^Z) e^{i\Delta\lambda_\mu \alpha} \right|^2, \quad (15)$$

where n indexes specific Z contributions. J_n is the spin of the given Z resonance. α is the angle between the muon helicity frames produced via K^* and Z decay chains.

In the $B \rightarrow ZK$ decay, the Z helicity must be zero since both B and K have spin $J = 0$. When $J_n = 0$, the ψ helicity must also be zero, so there will be only one complex amplitude to fit for a spinless Z contribution. When $J_n \neq 0$, the ψ can take values of -1, 0, or +1, thus the ψ helicity labels the amplitudes; however, the strong Z decay conserves parity, thus these amplitudes are not independent:

$$\underline{A}_{n,-\lambda_\psi^Z} = P_\psi P_\pi P_n (-1)^{J_\psi + J_\pi - J_n} \underline{A}_{n,\lambda_\psi^Z}, \quad (16)$$

where P_n is the parity of a specific Z contribution. Given $P_\psi = P_\pi = -1$, which are the ψ and π parities, $J_\psi = 1$, and $J_\pi = 0$, this becomes

$$\underline{A}_{n,-\lambda_\psi^Z} = -P_n (-1)^{J_n} \underline{A}_{n,\lambda_\psi^Z}. \quad (17)$$

This presents a few scenarios. First, when $J_n = 0$, we see that $\underline{A}_{n,0} = -P_n \underline{A}_{n,0}$, so the Z parity must be odd, otherwise $\underline{A}_{n,0} = -\underline{A}_{n,0}$, meaning $\underline{A}_{n,0} = 0$. Similarly, when $J_n^P = 1^-, 2^+$, the helicity-zero amplitude is forbidden, and $\underline{A}_{n,-1} = -\underline{A}_{n,+1}$. Thus for $J_n = 0^-, 1^-, 2^+$, there is only one complex amplitude to fit. For $J_n^P = 1^+, 2^-$, the helicity-zero amplitude is allowed, and $\underline{A}_{n,-1} = \underline{A}_{n,+1}$, therefore there are only two complex amplitudes to fit for these J_n^P values.

Thus combining the K^* and Z chains matrix elements, we get the full 4D matrix element:

$$|\mathcal{M}|^2 \equiv \sum_{\Delta\lambda_\mu=-1,1} \left| \sum_{\lambda_\psi=-1,0,1} \sum_k \underline{A}_{k,\lambda_\psi} R(m_{K\pi}|m_{0k}, \Gamma_{0k}) d_{\lambda_\psi,0}^{J_k}(\theta_{K^*}) e^{i\lambda_\psi\phi} d_{\lambda_\psi,\Delta\lambda_\mu}^1(\theta_\psi) \right. \\ \left. + \sum_{\lambda_\psi^Z=-1,0,1} \sum_n \underline{A}_{n,\lambda_\psi^Z} R(m_{\psi\pi}|m_{0n}, \Gamma_{0n}) d_{0,\lambda_\psi^Z}^{J_n}(\theta_Z) e^{i\lambda_\psi^Z\phi^Z} d_{\lambda_\psi^Z,\Delta\lambda_\mu}^1(\theta_\psi^Z) e^{i\Delta\lambda_\mu\alpha} \right|^2, \quad (18)$$

7.3 Helicity formalism for the Z_K decay chain

The Z_K chain follows the same construction as the Z chain for its contribution to the matrix element. It uses the same logic for helicity values and applies the same parity constraint for helicity amplitudes. The difference lies in that variables used for the Z_K chain are in a different rest frame than those in the Z chain. Thus, the full 4D matrix element for the Z_K chain is:

$$|\mathcal{M}^{Z_K}|^2 \equiv \sum_{\Delta\lambda_\mu=-1,1} \left| \sum_{\lambda_\psi^{Z_K}=-1,0,1} \sum_l \underline{A}_{l,\lambda_\psi^{Z_K}} R(m_{\psi K}|m_{0l}, \Gamma_{0l}) d_{0,\lambda_\psi^{Z_K}}^{J_l}(\theta_{Z_K}) \right. \\ \left. \times e^{i\lambda_\psi^{Z_K}\phi^{Z_K}} d_{\lambda_\psi^{Z_K},\Delta\lambda_\mu}^1(\theta_\psi^{Z_K}) e^{i\Delta\lambda_\mu\alpha_K} \right|^2, \quad (19)$$

where l indexes specific Z_K contributions. J_l is the spin of the given Z_K resonance. α_K is the angle between the muon helicity frames produced via K^* and Z_K decay chains.

Thus the full 4D matrix element when including all three decay chains is as follows:

$$\begin{aligned}
|\mathcal{M}|^2 \equiv & \sum_{\Delta\lambda_\mu=-1,1} \left| \sum_{\lambda_\psi=-1,0,1} \sum_k \underline{A}_{k,\lambda_\psi} R(m_{K\pi}|m_{0k}, \Gamma_{0k}) d_{\lambda_\psi,0}^{J_k}(\theta_{K^*}) e^{i\lambda_\psi\phi} d_{\lambda_\psi,\Delta\lambda_\mu}^1(\theta_\psi) \right. \\
& + \sum_{\lambda_\psi^Z=-1,0,1} \sum_n \underline{A}_{n,\lambda_\psi^Z} R(m_{\psi\pi}|m_{0n}, \Gamma_{0n}) d_{0,\lambda_\psi^Z}^{J_n}(\theta_Z) e^{i\lambda_\psi^Z\phi^Z} d_{\lambda_\psi^Z,\Delta\lambda_\mu}^1(\theta_\psi^Z) e^{i\Delta\lambda_\mu\alpha} \\
& \left. + \sum_{\lambda_\psi^{ZK}=-1,0,1} \sum_l \underline{A}_{l,\lambda_\psi^{ZK}} R(m_{\psi K}|m_{0l}, \Gamma_{0l}) d_{0,\lambda_\psi^{ZK}}^{J_l}(\theta_{ZK}) e^{i\lambda_\psi^{ZK}\phi^{ZK}} d_{\lambda_\psi^{ZK},\Delta\lambda_\mu}^1(\theta_\psi^{ZK}) e^{i\Delta\lambda_\mu\alpha^K} \right|^2, \quad (20)
\end{aligned}$$

7.4 Amplitude dependence on the invariant masses

Each contribution to the matrix element comes with its own $R(m_A)$ function, which gives its dependence on the invariant mass of the intermediate resonance in the decay chain A ($A = K^*$ or Z).

7.4.1 Single channel K-matrix

Historically, Breit-Wigner sums were used in the Isobar model approach to calculate the full matrix element for the K^* decay chain. In either decay chain though, B^0 can decay to multiple states in the same partial wave, i.e. $B^0 \rightarrow \psi K_1^*(892)$, $B^0 \rightarrow \psi K_1^*(1410)$..., which can have strong overlaps and thus interfere with each other. To account for these overlaps and interferences, we use a single-channel K-matrix to model the resonant contributions from multiple states in the same partial wave, combined with Blatt-Weisskopf functions:

$$R(m|M_0, \Gamma_0) = B'_{L_B}(p, p_0, d) \left(\frac{p}{m_B} \right)^{L_B} KM(m|M_0, \Gamma_0) B'_{L_A}(q, q_0, d) \left(\frac{q}{q_0} \right)^{L_A}, \quad (21)$$

where

$$KM(m|M_0, \Gamma_0) = \frac{\frac{\sqrt{k}}{M_0^2 - m^2}}{1 - i(\sum_j \frac{M_{0j}\Gamma_j(m)}{M_{0j}^2 - m^2} + \rho(m)f_{SC})}, \quad (22)$$

is the single-channel K-matrix amplitude including the mass dependent width,

$$\Gamma(m) = \Gamma_0 \left(\frac{q}{q_0} \right)^{2L_A+1} \frac{M_0}{m} B'_{L_A}(q, q_0, d)^2. \quad (23)$$

The sum over j is a sum over all the resonances in the same partial wave for that decay chain, where f_{SC} is a real constant (which is free to float in the fit) representing the non-resonant scattering in that partial wave (we use it only in some S-wave components). $\rho(m) = 2q/m$ is a phase space factor for this non-resonant term. Here, p is the momentum of the resonance A (K^* or Z) in the rest frame of B^0 . q is the momentum of the A resonance's daughter in the rest frame of the A resonance. The symbols p_0 and q_0 are used to indicate values of these quantities at the resonance peak mass ($m = M_0$). The orbital angular momentum in B decay is denoted as L_B , while in the decay of the resonance A as L_A . The orbital angular momentum barrier factors, $p^L B'_L(p, p_0, d)$, involve the Blatt-Weisskopf functions:

$$B'_0(p, p_0, d) = 1, \quad (24)$$

$$B'_1(p, p_0, d) = \sqrt{\frac{1 + (p_0 d)^2}{1 + (p d)^2}}, \quad (25)$$

$$B'_2(p, p_0, d) = \sqrt{\frac{9 + 3(p_0 d)^2 + (p_0 d)^4}{9 + 3(p d)^2 + (p d)^4}}, \quad (26)$$

$$B'_3(p, p_0, d) = \sqrt{\frac{225 + 45(p_0 d)^2 + 6(p_0 d)^4 + (p_0 d)^6}{225 + 45(p d)^2 + 6(p d)^4 + (p d)^6}}, \quad (27)$$

$$B'_4(p, p_0, d) = \sqrt{\frac{11025 + 1575(p_0 d)^2 + 135(p_0 d)^4 + 10(p_0 d)^6 + (p_0 d)^8}{11025 + 1575(p d)^2 + 135(p d)^4 + 10(p d)^6 + (p d)^8}}, \quad (28)$$

$$B'_5(p, p_0, d) = \sqrt{\frac{893025 + 99225(p_0 d)^2 + 6300(p_0 d)^4 + 315(p_0 d)^6 + 15(p_0 d)^8 + (p_0 d)^{10}}{893025 + 99225(p d)^2 + 6300(p d)^4 + 315(p d)^6 + 15(p d)^8 + (p d)^{10}}}, \quad (29)$$

which account for the difficulty in creating the orbital angular momentum (L) and depends on the momentum of the decay products in the rest frame of the decaying particle (p) as well as the size of the decaying particle given by the constant d . In this analysis, we set this parameter to a nominal value of $d = 3.0 \text{ GeV}^{-1}$.

In the helicity approach, each helicity state is a mixture of many different L values. We

follow the usual approach of using in the functions above the minimal L value allowed by the quantum numbers of the given resonance A .

The k constant follows suit from the formula for a relativistic Breit-Wigner amplitude. It is conventional and is introduced here in an attempt to decouple the numerical values of the helicity couplings to be fit from M_0 and Γ_0 parameters of the resonance:

$$k \equiv \frac{M_0 \Gamma_0 \gamma}{\sqrt{M_0^2 + \gamma}} \quad (30)$$

with $\gamma \equiv \sqrt{M_0^2 (M_0^2 + \Gamma_0^2)}$ (see Wikipedia for the Relativistic Breit-Wigner).

7.4.2 Model Independent Contribution

An alternate approach to representing a partial wave for Zs is to replace the K-matrix (which assumes a specific resonance shape) with a number of complex amplitudes equally spaced in $m_{\psi\pi}$. The real and imaginary parts of the amplitude are interpolated between the bin centers using a cubic spline based on the nearest four bins. We try this approach only for the dominant $J^P = 1^+$ Z wave. Since it has two sets of helicity couplings, we allocate two sets of these complex amplitudes: one for each helicity coupling. Both helicity couplings float in the fit, so in order to minimize correlated errors, we fix one mass bin (the one which has the greatest magnitude from an Argand plot produced from the nominal resonant model fits) to set the convention for the model independent amplitudes. The complex amplitudes for the remaining mass bins are free to float in the fit, which allows us to probe for $\psi\pi$ contributions without making assumptions regarding the complex phase motion of the contribution.

8 Significance and fit quality calculations

In this section, metrics used to evaluate fit models are described. The calculation of significance from the log-likelihoods and the change in two-dimensional and four-dimensional χ^2 s are all fit quality measures that are used to assess the validity of adding additional states to the model as well as determine spin-parity assignments of additional states. The results of these tests are reported in future sections.

8.1 Calculation of significance of adding states

The log-likelihoods for each fit are calculated according to Equation 7 in Section 6.3. When adding new states with a fixed resonance shape, a log-likelihood difference is used as a test statistic to evaluate the hypothesis of these new states. To account for the change in parameters when adding new shapes, Wilks' theorem is employed, which states that as the sample size approaches infinity, the test statistic $\Delta(-2\log(L))$ asymptotically approaches the χ^2 distribution with a number of degrees of freedom equal to the number of constrained parameters in the null hypothesis (no resonant contribution present) relative to the alternative hypothesis (the resonant contribution is present). This allows for calculation of the p-value of the null hypothesis, which is converted to number of standard deviations in a Gaussian distribution. Wilks' theorem applies only if the shape of the resonant contribution is fixed *i.e.* its mass and width are fixed. The previous $Z(4430)^-$ amplitude analysis conducted extensive statistical simulations of pseudo-experiments and determined that $\Delta(-2\ln L)$ distribution is much better approximated by χ^2 with the number of degrees of freedom equal to twice the difference in the number of free parameters between the hypotheses when the $Z(4430)^-$ mass and width are free to float. As such, the calculation of significance of a new state with mass and width free to float is:

$$n_\sigma(\Delta(-2\ln L)) = \sqrt{2} \operatorname{erfc}^{-1} [P(\Delta(-2\ln L), 2\Delta(\text{nPar}))], \quad (31)$$

where n_σ is the number of Gaussian standard deviations that represents the significance of the new state, and $P(\Delta(-2 \ln L), 2\Delta(\text{nPar}))$ is the χ^2 probability with number of degrees of freedom equal to twice the difference of the number of free parameters between the null hypothesis fit and the alternative hypothesis fit. erfc^{-1} is the inverse error function used to convert the p-value into Gaussian standard deviations.

8.2 Multidimensional χ^2 s

While the pulls of mass or angle projections are useful to identify missing contributions to the fit model, these fits are multidimensional, and therefore it is also informative to measure the quality of the fits in more than one dimension. The fit quality is measured both in the two-dimensional Dalitz plane ($m_{K\pi}^2$ vs $m_{\psi\pi}^2$) as well as in the full, four-dimensional phase space with $\cos\theta_\psi$ and ϕ_ψ being the two additional dimensions combined with the Dalitz plane. To obtain the χ^2 values, the data is binned, and the pulls between the data and the fit prediction for all bins are summed in quadrature, $\chi^2 = \sum_{j=1}^{N_{\text{bins}}} \delta_j^2$, where δ_j is the difference between the data and the fit value in the j -th bin divided by the expected error.

When the count of a data bin is sufficiently large, the χ^2 variable follows the χ^2 distribution to a good approximation with ndf value in the range $[N_{\text{bins}} - 1 - n_{\text{par}}, N_{\text{bins}} - 1]$, where n_{par} is the number of free parameters in the fit. To bin the data, all entries are first put into one bin, and then the data count is split in half based on new bin boundaries according to the number of divisions set for each variable. Therefore, the bin size is adaptive trying to maintain equal number of events in all bins. Bins are not divided if the bin content falls below 10. For the 2-D (Dalitz plane) χ^2 , each mass variable is divided 5 times, alternating between the mass variables for each division. This produces 1024 bins for the 2-D χ^2 . For the full 4-D χ^2 , each variable is divided twice, cycling through the variables for each division (first $m_{K\pi}^2$, then $m_{\psi\pi}^2$, $\cos\theta_\psi$, and finally ϕ_ψ), and after, the mass variables are divided twice more each, producing 4096 bins. Smaller χ^2 values indicate that an alternative hypothesis improves either the 2-D or 4-D phase space compared to the null hypothesis, and thus they

are used to justify the inclusion of additional states as will be seen in the coming sections.

Change of χ^2 value between the null and alternative hypotheses can also be used to assess significance of the resonance added to the amplitude model. However, since there is a loss of information when binning the data, this approach is less sensitive than the log-likelihood difference. Since the mass variables discriminate better between various resonant contributions than the angular distributions, the 4D χ^2 approach is often less sensitive than the 2D χ^2 approach as it has coarser bins in masses.

9 K^{*0} model

In $B^0 \rightarrow \psi K\pi$ decays, the $m_{K\pi}$ distribution features several resonant structures, most notably $K^*(892)$ and $K_2^*(1430)$. Listed in Table 5, the K^{*0} resonances known to decay to $K\pi$ are listed along with their properties. The lightest K^{*0} is $K_0^*(800)^0$ (often called κ in the literature); it has spin 0, is very broad and its existence was questioned in the past. In recent years, both experimental and theoretical evidences support its existence [23]. From a theoretical point of view, this state is not a traditional $s\bar{d}$ state, but rather a state predicted by the chiral Lagrangian for lightest mesons, which explains its very large width. The other K^{*0} resonances find their interpretation in classical $s\bar{d}$ spectroscopy [24] and correspond to various radial ($n = 1, 2, \dots$), angular momentum ($L = S, P, D, F, \dots$, not to be confused with the angular momenta between the final state particles in the decay) and quark total spin ($S = 0, 1$) configurations. Their likely spectroscopic assignments ($n^{2S+1}L_J$) are given in Table 5. The total spin of the meson J ($\equiv J_k$) is of particular importance for this analysis, since it determines the number of complex amplitudes to be included in the fit and impacts their $m_{\psi\pi}$ distribution.

For $B^0 \rightarrow \psi(2S)K\pi$ decays, the mass of the $K\pi$ system is limited to less than 1593 MeV, below which lie an additional spin-0 resonance, $K_0^*(1430)^0$, two spin-1 resonances, $K^*(892)^0$ and $K_1^*(1410)^0$, and a spin-2 resonance, $K_2^*(1430)^0$. Above the kinematic mass limit lie several more K^* s, including those of higher spin. There is an additional spin-0 resonance, $K_0^*(1950)^0$, a spin-1 resonance, $K_1^*(1680)^0$, a spin-2 resonance, $K_2^*(1980)^0$, a spin-3 resonance, $K_3^*(1780)^0$, a spin-4 resonance, $K_4^*(2045)^0$, and a spin-5 resonance, $K_5^*(2380)^0$. The $K_2^*(1980)^0$ was not used in the previous $B^0 \rightarrow \psi K\pi$ amplitude analysis, but it was since included in the amplitude analysis for $B^+ \rightarrow J/\psi\phi K^+$ [25] as part the model of ϕK^+ structures. All of these higher mass states can contribute their resonance tails to the $K\pi$ mass spectrum in $B^0 \rightarrow \psi(2S)K\pi$ decays, but the heavier the resonance, the lower the probability for it to show up due to phase-space suppression. Also the higher the spin the lower the probability,

since $L_B^{min} = |J - 1|$, and high angular momenta are efficiently suppressed by the $|(p/m_B)^{L_B}|^2$ term in the matrix element squared. Relatively to a spin-1 resonance for which $L_B^{min} = 0$, this term for a mid-range $m_{K\pi}$ (1.1 GeV) is 0.03, 0.001, 0.00004 and 0.000001 for $J = 0$ or 2, 3, 4 and 5, respectively. Therefore, spin-3 is not likely to contribute much to the decay width, and spin 4 and 5 contributions are extremely unlikely to be present.

For most K^{*0} resonances, we fix the masses and widths to the PDG world average values, but because of the dominant mass peaks of $K^*(892)^0$ and $K_2^*(1430)$, small variations in their masses and widths can greatly impact the goodness-of-fit measures. As such, we allow their masses and widths to float freely in the fit. We also allow the mass and width of $K_0^*(800)^0$ to float freely, as it contributes strongly to the K^{*0} S-wave, and the flexibility in its fit shape improves goodness-of-fit measures in a similar manner as $K^*(892)^0$ and $K_2^*(1430)$.

Since the J/ψ mass is lower than the $\psi(2S)$ mass, the kinematic limit for the $m_{K\pi}$ spectrum is at 2183 MeV for $B^0 \rightarrow J/\psi K\pi$ decays. This means that known K^{*0} s up to spin $J = 4$ are within phase space. The only known K^{*0} above the $m_{K\pi}$ kinematic limit for the J/ψ channel is $K_5^*(2380)^0$, which is also the first (and only known) spin $J = 5$ K^{*0} . Much like the K^{*0} s above the mass limit for the $B^0 \rightarrow \psi(2S)K\pi$ channel, this resonance can also contribute via its resonance tail. As with the $\psi(2S)$ channel, $K_0^*(800)^0$, $K^*(892)^0$, and $K_2^*(1430)$ have freely floating masses and widths. In addition to those, we allow the $K_0^*(1430)$'s mass and width to float freely, since it improves the goodness-of-fit measures in the larger data sample in the $B^0 \rightarrow J/\psi K\pi$ channel.

Table 5: Known K^{*0} resonances decaying to $K^+\pi^-$, with their spectroscopic classification and measured properties [1].

Resonance	J^P	Likely $n^{2S+1}L_J$	Mass (MeV)	Width (MeV)	$\mathcal{B}(K^{*0} \rightarrow K^+\pi^-)$
$K_0^*(800)^0$ (κ)	0^+	—	845 ± 17	468 ± 30	$\sim 100\%$
$K^*(892)^0$	1^-	1^3S_1	895.55 ± 0.20	47.3 ± 0.5	$\sim 100\%$
$K_0^*(1430)^0$	0^+	1^3P_0	1425 ± 50	270 ± 80	$(93 \pm 10)\%$
$K_1^*(1410)^0$	1^-	2^3S_1	1414 ± 15	232 ± 21	$(6.6 \pm 1.3)\%$
$K_2^*(1430)^0$	2^+	1^3P_2	1432.4 ± 1.3	109 ± 5	$(49.9 \pm 1.2)\%$
$B^0 \rightarrow \psi(2S)K\pi$ phase space limit			1593		
$K_1^*(1680)^0$	1^-	1^3D_1	1718 ± 18	322 ± 110	$(38.7 \pm 2.5)\%$
$K_3^*(1780)^0$	3^-	1^3D_3	1779 ± 8	161 ± 17	$(18.8 \pm 1.0)\%$
$K_0^*(1950)^0$	0^+	2^3P_0	1944 ± 18	100 ± 40	$(52 \pm 14)\%$
$K_2^*(1980)^0$	2^+	2^3P_2	1994_{-50}^{+60}	348_{-30}^{+50}	-
$K_4^*(2045)^0$	4^+	1^3F_4	2048_{-9}^{+8}	199_{-19}^{+27}	$(9.9 \pm 1.2)\%$
$B^0 \rightarrow J/\psi K\pi$ phase space limit			2183		
$K_5^*(2380)^0$	5^-	1^3G_5	2382 ± 24	178 ± 50	$(6.1 \pm 1.2)\%$

10 K^* Only Amplitude Model

We present here the results of 4D fits that use the full K^{*0} model without any $\psi\pi$ resonances. Each partial wave was represented with the K-matrix formulation discussed in Sec 7. For the $B^0 \rightarrow J/\psi K\pi$ channel only, we include a non-resonant term in the K^{*0} S-wave K-matrix. This term was not necessary for the $B^0 \rightarrow \psi(2S)K\pi$ dataset or for other partial waves in the K^{*0} model.

We show the projections for the masses in Fig. 43 for the $B^0 \rightarrow \psi(2S)K\pi$ dataset and in Fig. 46 for the $B^0 \rightarrow J/\psi K\pi$ dataset. The fits to both datasets describe $m_{K\pi}$ fairly well, but there is a significant disagreement for both channels below $K^*(892)$ (see the fit pull distributions). For $B^0 \rightarrow J/\psi K\pi$ decays, there are additional disagreements near $K^*(1430)$ and at high $m_{K\pi}$ mass (well above the $B^0 \rightarrow \psi(2S)K\pi$ phase space limit).

The $m_{\psi\pi}$ projections for both channels are not well represented by the fits. For $B^0 \rightarrow \psi(2S)K\pi$, there is a large disagreement at $m_{\psi\pi} > 4.4$ GeV. In $B^0 \rightarrow J/\psi K\pi$, there is a large disagreement at $m_{\psi\pi} > 4.4$ GeV too, and in addition, there is a disagreement between $m_{\psi\pi} = 4.0$ GeV and 4.3 GeV.

We also show the projections for the $m_{\psi\pi}$ distribution in different slices of $m_{K\pi}$, which are in Fig. 44 for the $B^0 \rightarrow \psi(2S)K\pi$ dataset and Fig.47 for the $B^0 \rightarrow J/\psi K\pi$ dataset. In the $\psi(2S)$ channel, the $m_{\psi\pi}$ projections for slices of $m_{K\pi}$ near a known K^{*0} resonance, i.e. $K^*(892)$ and $K^*(1430)$ mass slices, are fairly well described by the fits, but in regions of $m_{K\pi}$ near no obvious $K\pi$ structures, the $m_{\psi\pi}$ projections have disagreements at $m_{\psi\pi} > 4.4$ GeV and also below $m_{\psi\pi} = 4.2$ GeV. In the J/ψ channel, the $m_{\psi\pi}$ projections show several disagreements throughout phase space, even in regions close to prominent K^{*0} resonances.

Similarly, the various angular distributions are shown in Fig. 45 for the $B^0 \rightarrow \psi(2S)K\pi$ dataset and in Fig. 48 for the $B^0 \rightarrow J/\psi K\pi$ dataset. While the $\cos(\theta_\psi)$ and ϕ_ψ projections have minimal disagreements, the $\cos(\theta_\psi)$ projection for both channels is not well represented with large disagreements between the fit and the data.

Given the numerous discrepancies between the fit results and data, it is clear that the data cannot be described by the K^{*0} -only model, which confirms the finding in the previous analyses by LHCb and Belle (see Sec. 3).

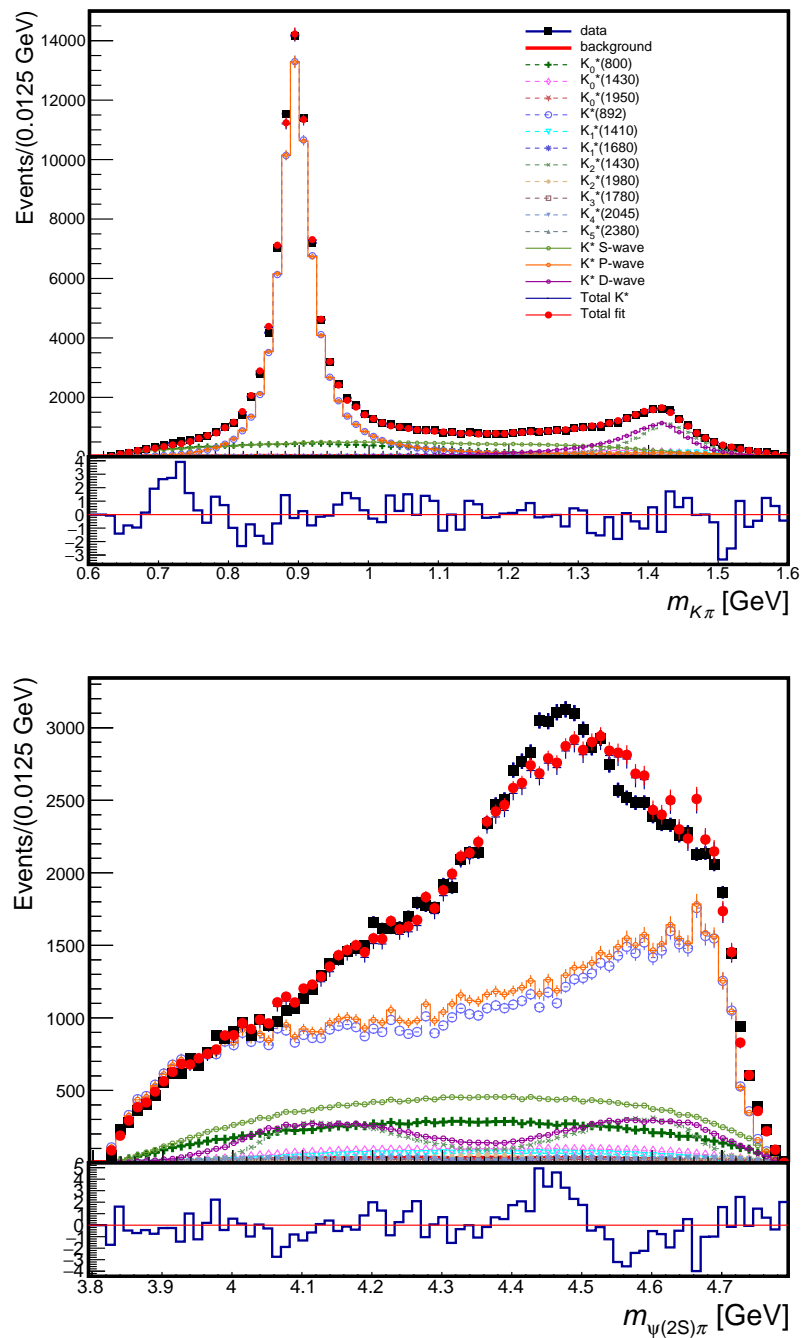


Figure 43: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) using a model with only the K^{*0} s listed in PDG. Masses and widths of the states are fixed to the PDG values except for those mentioned in Sec. 9. The widths are fixed to the PDG values or left free for the predicted states.

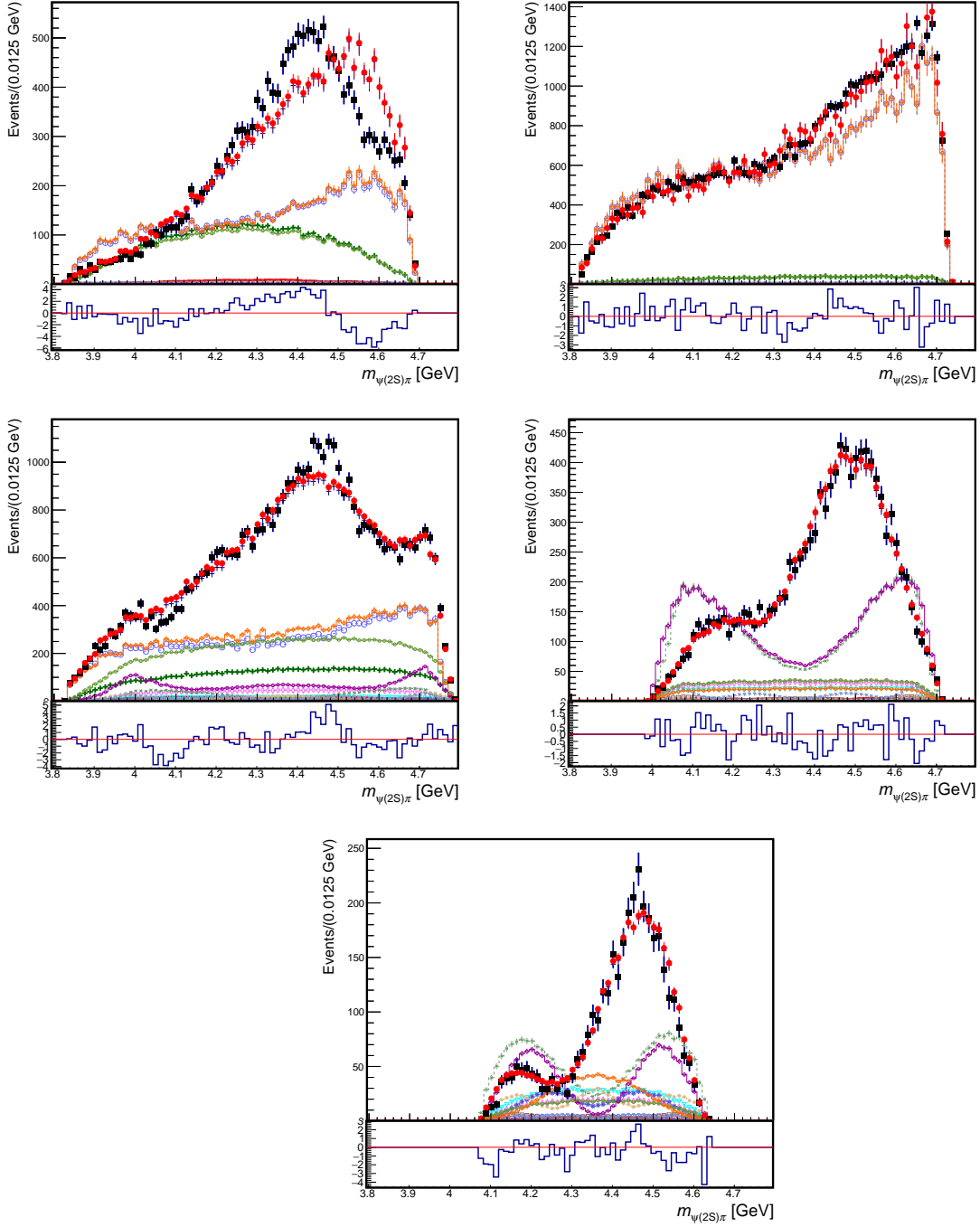


Figure 44: Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the K^{*0} -only 4D amplitude fit (red) onto the $m_{\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), at $K^*(892)$ (top right), in between $K^*(892)$ and $K_2^*(1430)$ (middle left), at $K_2^*(1430)$ (middle right), and above $K_2^*(1430)$ (bottom).

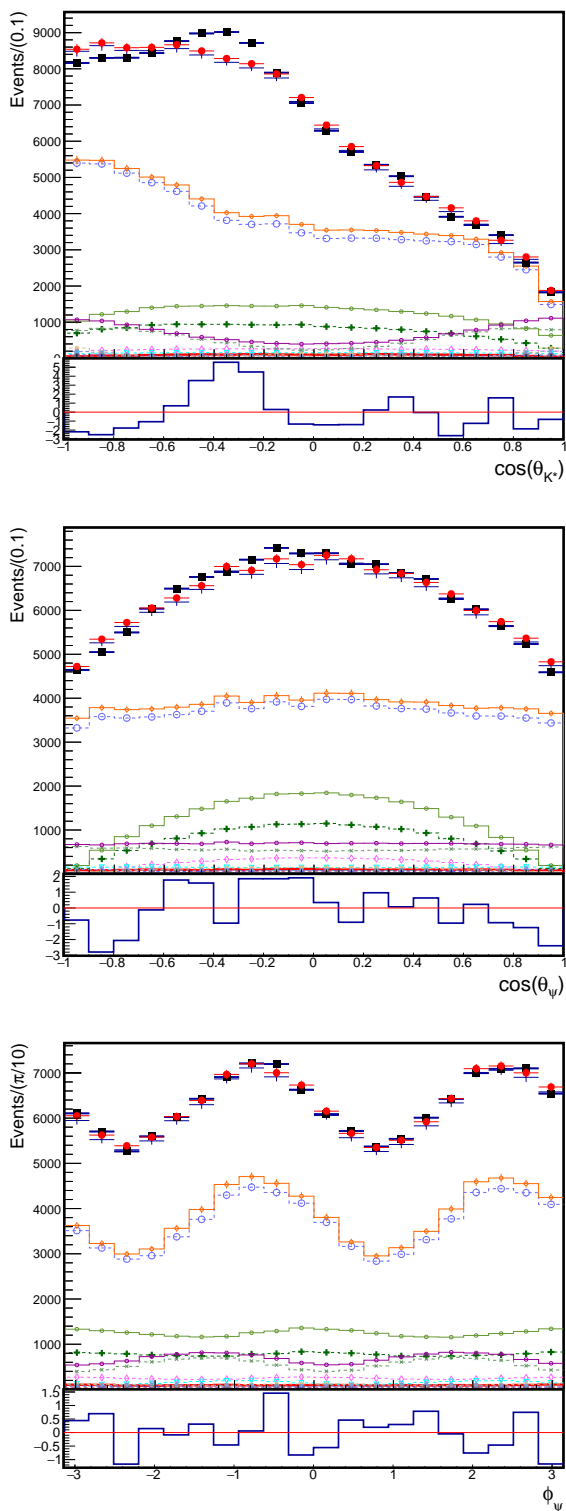


Figure 45: Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the K^{*0} -only 4D amplitude fit (red) onto the angular variables of the fit: cosine of the K^{*0} helicity angle (top), cosine of the ψ helicity angle (middle), and the angle between the K^{*0} and ψ decay planes (ϕ).

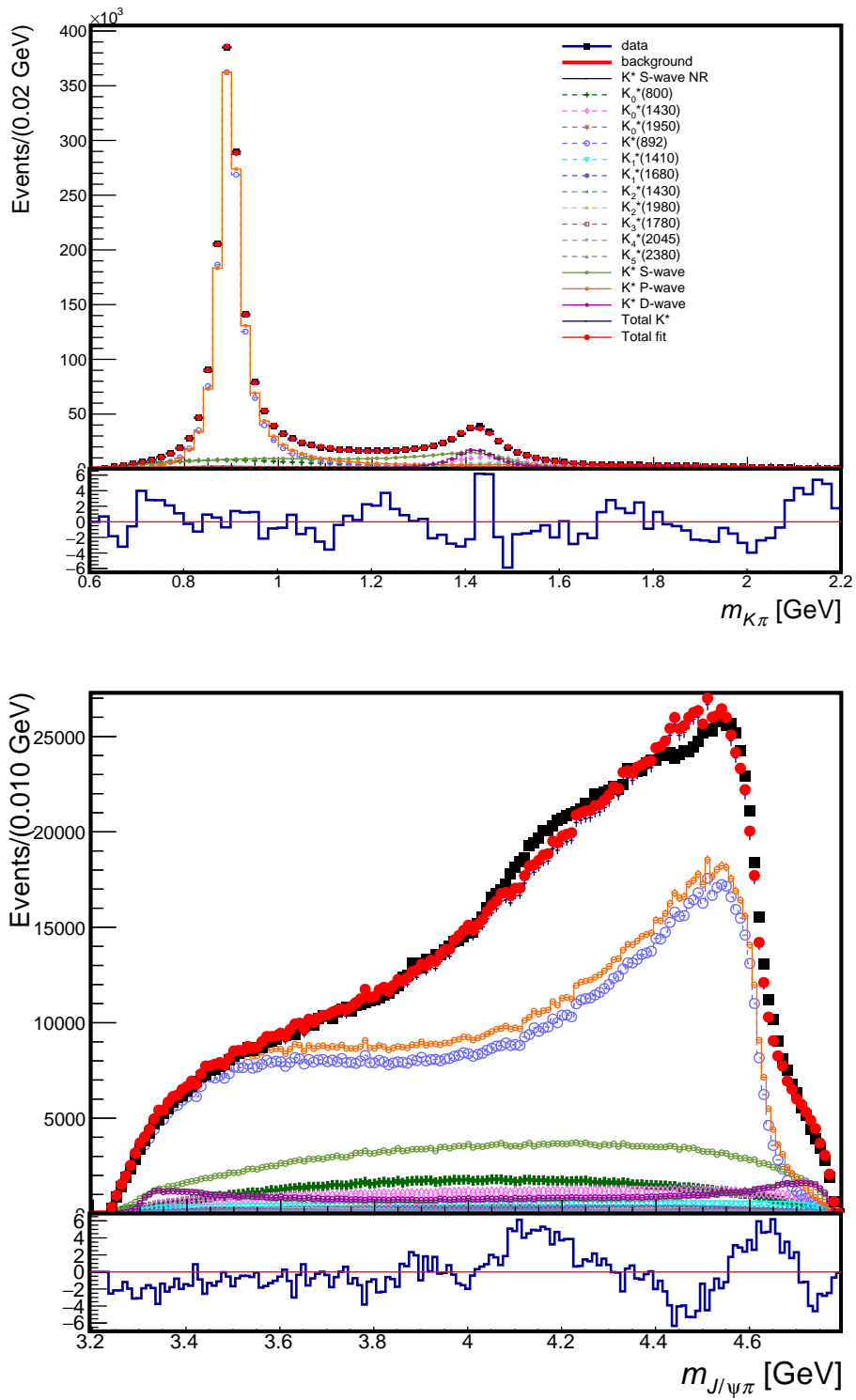


Figure 46: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi K\pi$ data (black) and the fit (red) using a model with only the K^{*0} s listed in PDG. Masses and widths of the states are fixed to the PDG values except for those K^* mentioned in Sec. 9. The widths are fixed to the PDG values or left free for the predicted states.

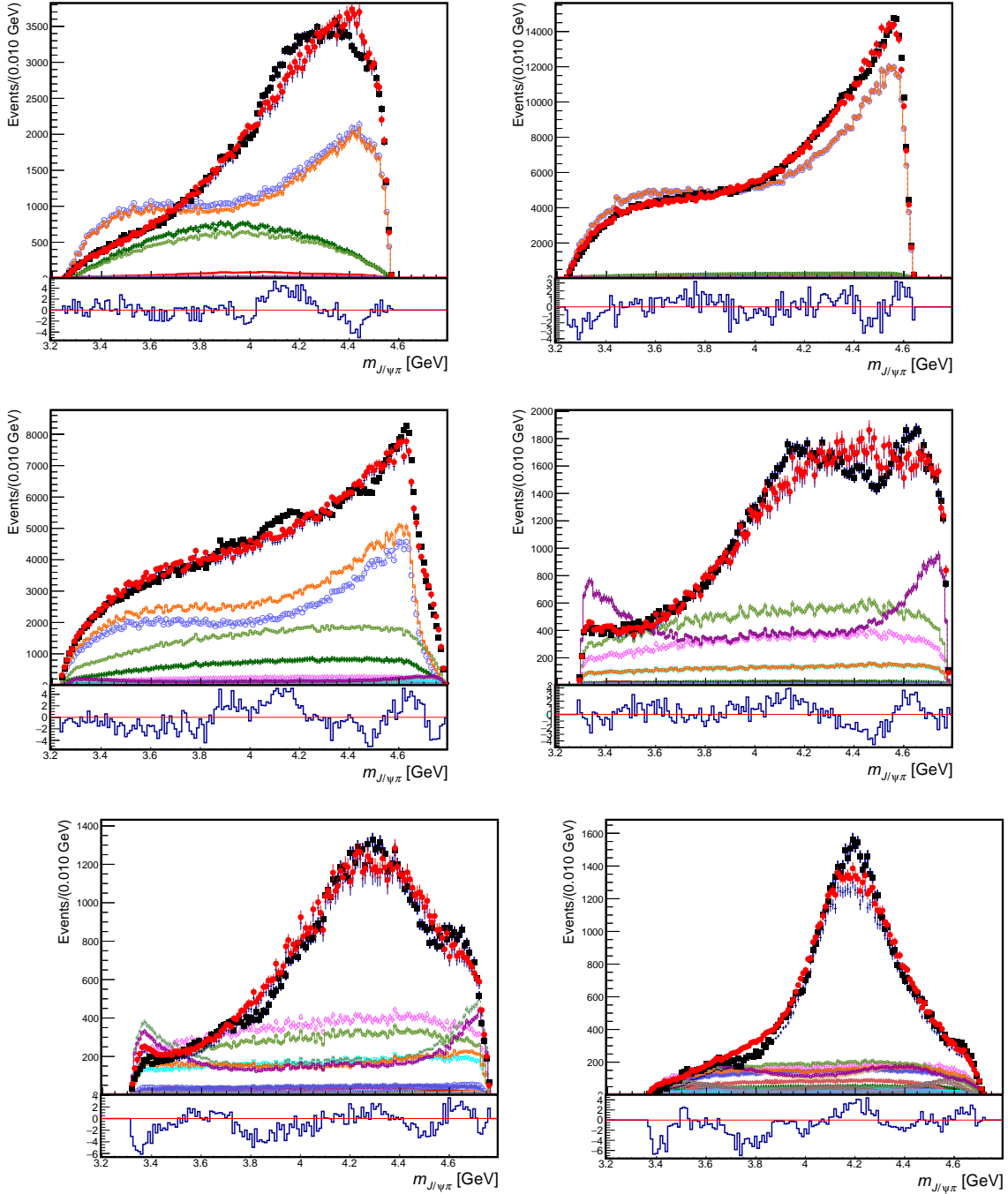


Figure 47: Projections of the $B^0 \rightarrow J/\psi K\pi$ data (black points) and of the K^{*0} -only 4D amplitude fit (red) onto the $m_{\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), at $K^*(892)$ (top right), in between $K^*(892)$ and $K_2^*(1430)$ (middle left), at $K_2^*(1430)$ (middle right), above $K_2^*(1430)$ but below the $m_{K\pi}$ phase space limit in $B^0 \rightarrow \psi(2S)K\pi$ decays (bottom left), and above the $m_{K\pi}$ phase space limit in $B^0 \rightarrow \psi(2S)K\pi$ decays (bottom right).

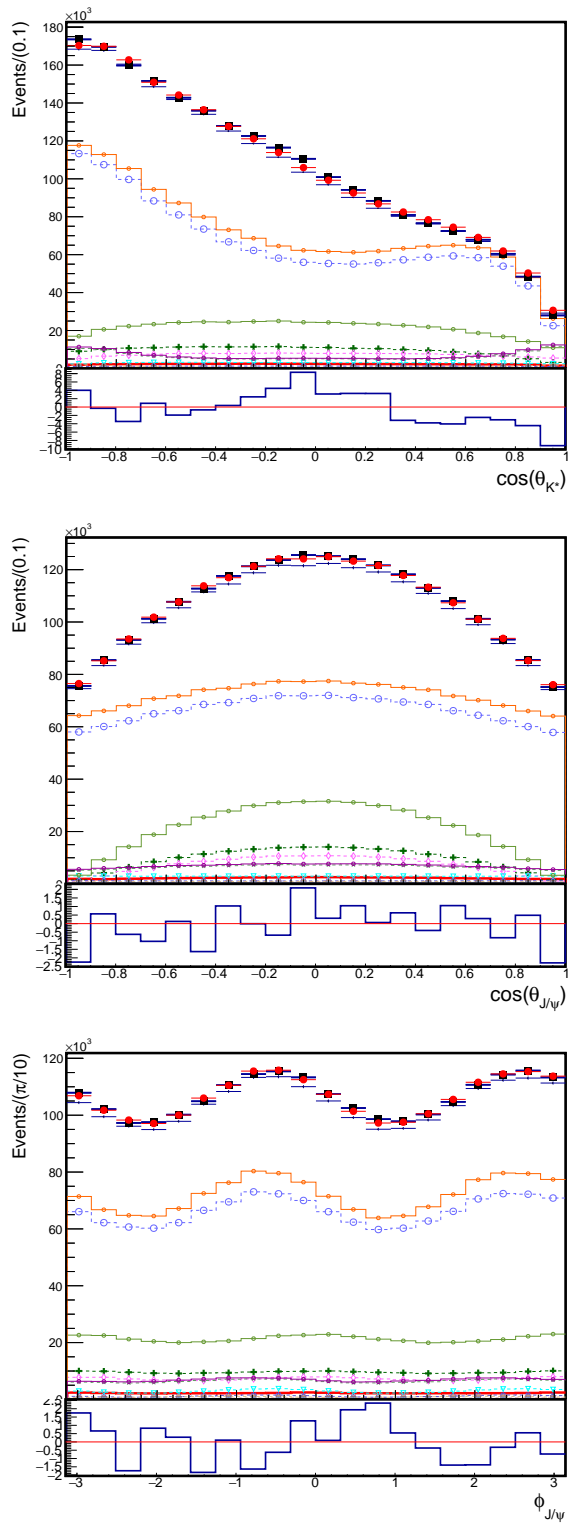


Figure 48: Projections of the $B^0 \rightarrow J/\psi K\pi$ data (black points) and of the K^{*0} -only 4D amplitude fit (red) onto the angular variables of the fit: cosine of the K^{*0} helicity angle (top), cosine of the ψ helicity angle (middle), and the angle between the K^{*0} and ψ decay planes (ϕ).

11 Previous $\psi(2S)K\pi$ Model on Current Dataset

The present Run 1 and 2 data set has about 6 times more signal events than used in the previously published LHCb amplitude analysis of the $B^0 \rightarrow \psi(2S)\pi^-K^+$ channel [7], which was based on Run 1 data and on less efficient data selection. To check compatibility of the new data set with the previously published results, we fit it with the amplitude model used previously. The new amplitude model uses single-channel K-matrix approach. It reduces to the Breit-Wigner representation in the case that there is only one resonance in the partial wave, however, it adds multiple resonances of the same quantum numbers differently than the Breit-Wigner sum used previously. Therefore, in this section we revert to the Breit-Wigner sum approach.

11.1 Fits with the previous K^{*0} -only model

The resonant content is set to be the same as in the default model used in Ref. [7], which includes K^{*0} resonances up to spin 2, but without resonances with pole masses beyond the phase-space limit (*i.e.* no $K_0^*(1950)^0$ or $K_2^*(1980)^0$). The previous analysis floated, but constrained, $K^*(892)$ and $K_0^*(800)$ masses and widths. Since we have much more data, we just float them. We also do not use the cut on the Dalitz plot boundary, which excluded 4% of the maximal m^2 range (reducing statistics by 12%). Obviously, such K^{*0} -only model cannot describe the new data as shown in Figs. 49-51.

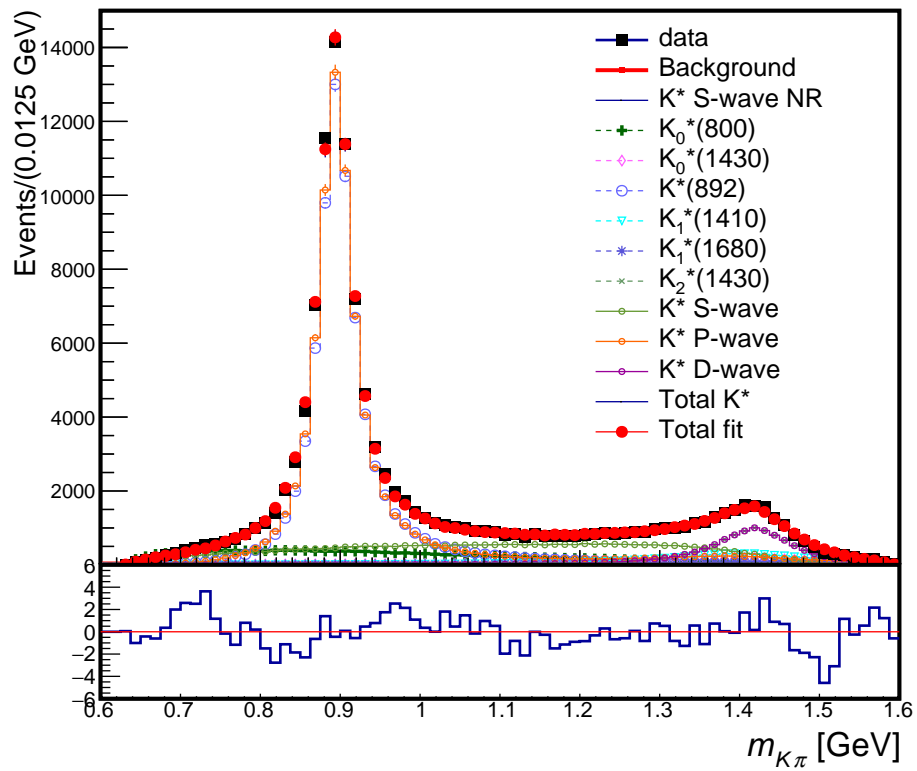


Figure 49: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) using the 2014 LHCb model with only K^* contributions.

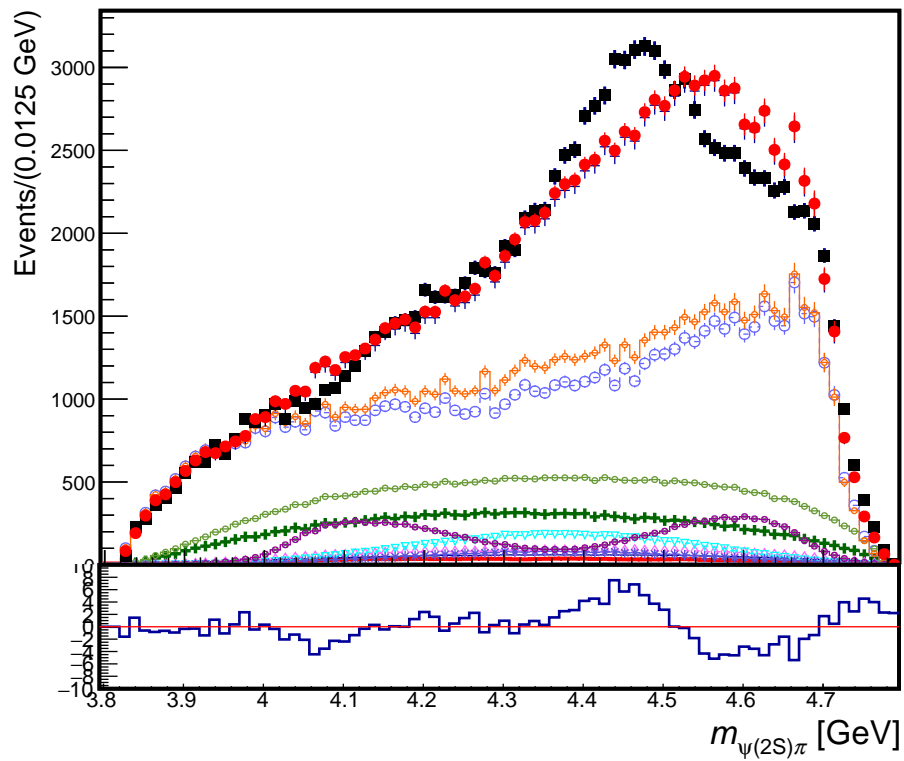


Figure 50: The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) using the 2014 LHCb model with only K^* contributions.

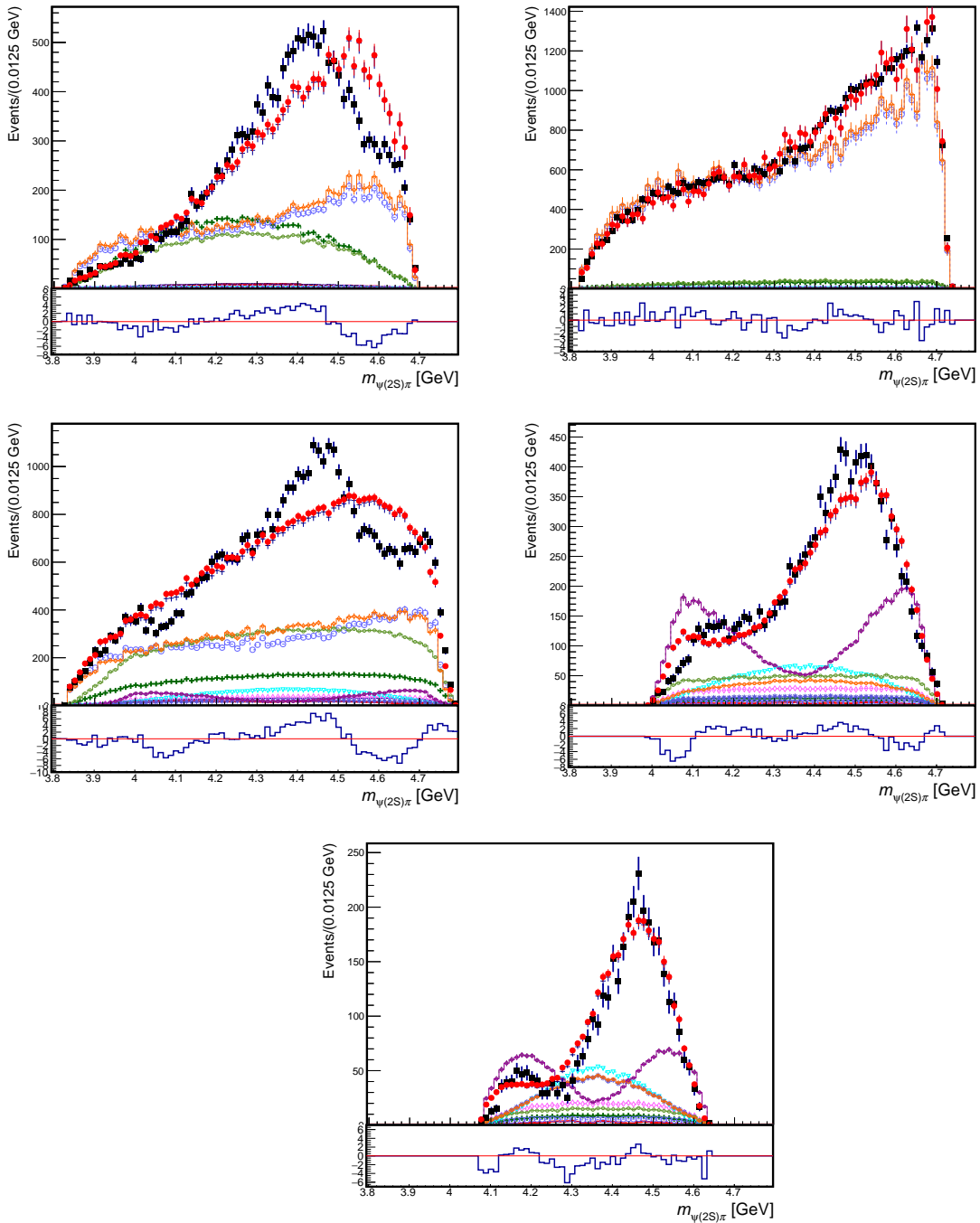


Figure 51: Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the 2014 LHCb model 4D amplitude fit (red) with only K^* contributions onto the $m_{\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), at $K^*(892)$ (top right), in between $K^*(892)$ and $K_2^*(1430)$ (middle left), at $K_2^*(1430)$ (middle right), and above $K_2^*(1430)$ (bottom).

11.2 Fits with the old model including $Z(4430)^-$

Like in the previous default model, we now add just one $1^+ \psi(2S)\pi^-$ resonance, assumed to decay in S-wave, which imposes a constraint on its helicity couplings ($\underline{A}_{Z,1} = \underline{A}_{Z,-1} = \underline{A}_{Z,0}$). The fit quality improves dramatically as shown in Figs. 52-55, and quantified by multiple metrics summarized in Tab. 6. We see huge improvement in the negative log-likelihood, 2D, and 4D χ^2 s from the fit.

Table 6: Fit quality metrics comparing 2014 K^{*0} -only fit with 2014 model fit on current dataset.

Metric	2014 K^{*0} -only model	2014 $Z(4430)^-$ model
NDF	34	38
$\log L$ ($\Delta(-2 \log L)$)	-100,113	-101,625 (3,024)
2D χ^2 1024 bins ($\Delta\chi^2$)	3510	1794 (1,716)
4D χ^2 4096 bins ($\Delta\chi^2$)	6571	4984 (1,587)

Parameters of the $Z(4430)^-$ resonances are in agreement with the published results as shown in Tab. 7.

Table 7: Mass and width of the $Z(4430)^-$ compared to results from 2014 using the model from LHCb in 2014.

Parameter	2014 Data	Current Data
$M(Z(4430)^-)$	$4475 \pm 7_{-25}^{+15}$ MeV	4475 ± 2 MeV
$\Gamma(Z(4430)^-)$	$172 \pm 13_{-34}^{+37}$ MeV	187 ± 5 MeV

Remarkably, individual fit fractions of all model components are also very similar as shown in Tab. 8.

In spite of huge improvement in the fit quality, the old default model has deficiencies even in the description of the $K\pi$ mass spectrum, as noticeable in the 1350-1550 MeV region (Fig. 52), which are mostly removed with more K^{*0} components, even without exotic contributions in the model (Fig. 43).

As already recognized in the previously published analysis even bigger disagreements between the data and the model are seen in $\psi(2S)\pi$ mass distribution, in particular in the

Table 8: All resonance fit fractions compared to results from 2014 using the model from LHCb in 2014.

Contribution	2014 Data	Current Data
Z(4430)	$5.9 \pm 0.9\%$	8.0%
K^* Total S-wave	$10.8 \pm 1.3\%$	13.2%
K^* S-wave NR	$0.3 \pm 0.8\%$	0.02%
$K_0^*(800)$	$3.2 \pm 2.2\%$	6.0%
$K_0^*(1430)$	$3.6 \pm 1.1\%$	3.6%
$K^*(892)$	$59.1 \pm 0.9\%$	57.6%
$K_1^*(1410)$	$1.7 \pm 0.8\%$	2.0%
$K_1^*(1680)$	$4.0 \pm 1.5\%$	2.8%
$K_2^*(1430)$	$7.0 \pm 0.4\%$	6.3%

$K\pi$ mass slice in between $K^*(892)$ and $K_2^*(1430)$ resonances (Fig. 54). Upgrades to the model of exotic contributions are discussed in the section which follow.

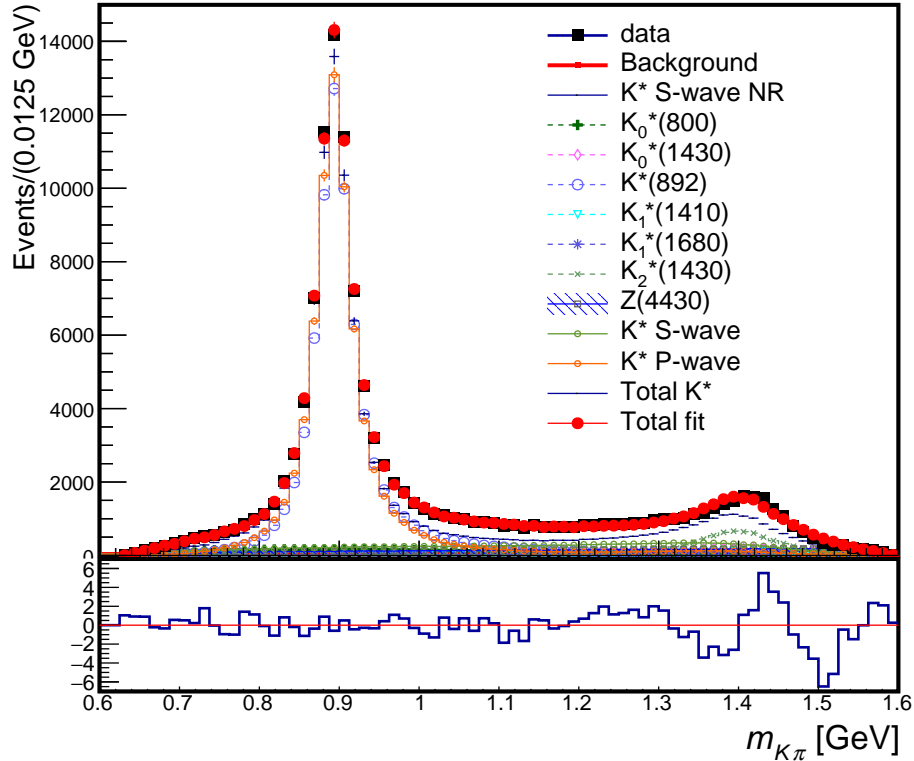


Figure 52: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) using the 2014 LHCb model with $Z(4430)^-$.

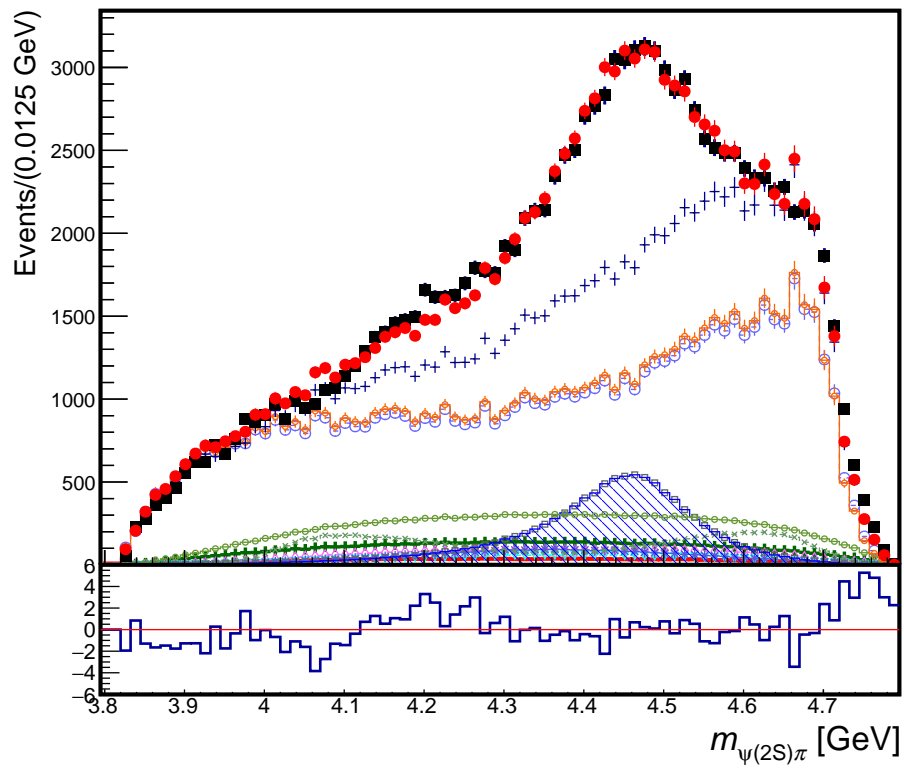


Figure 53: The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) using the 2014 LHCb model with $Z(4430)^-$.

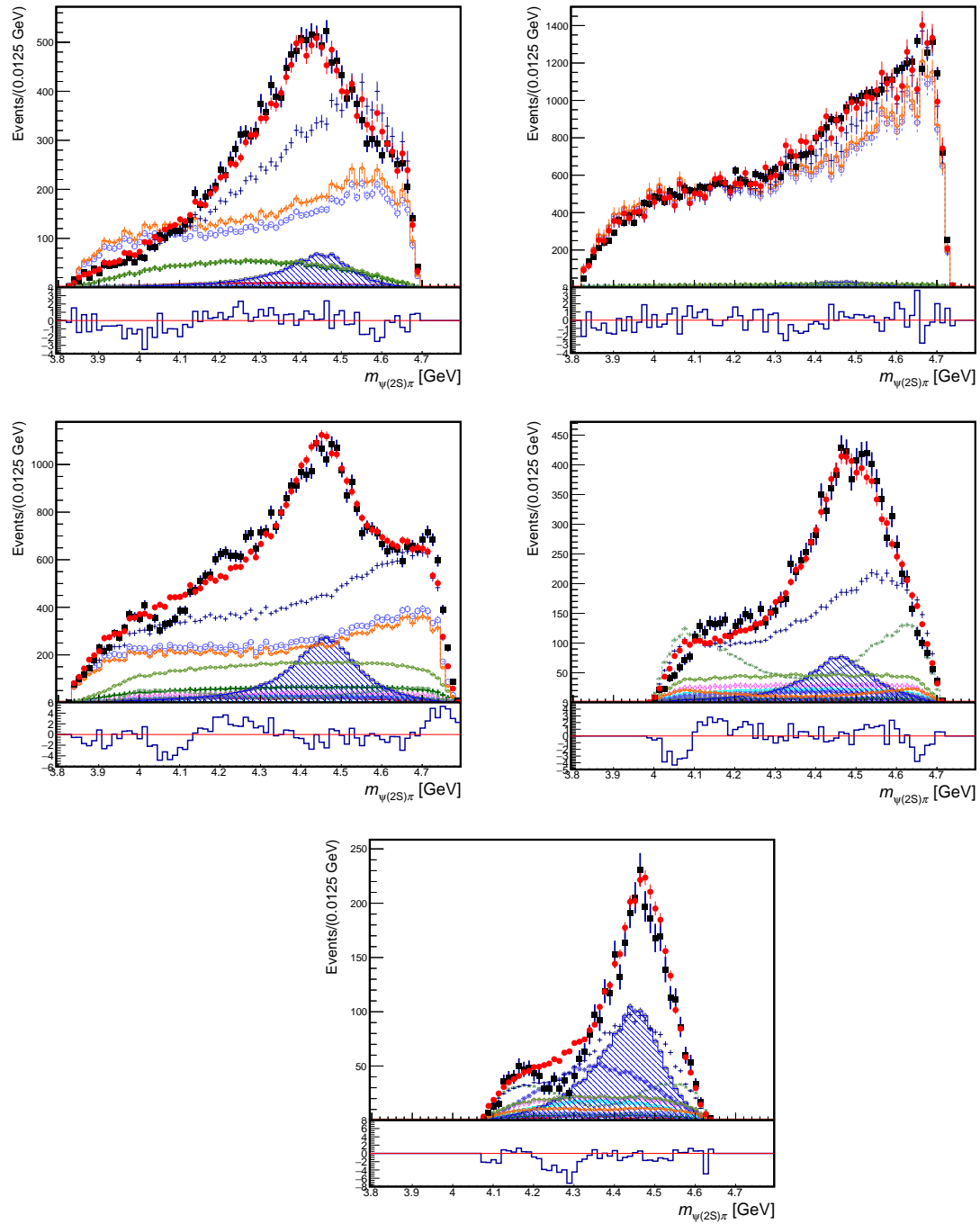


Figure 54: Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the 2014 LHCb model 4D amplitude fit (red) with $Z(4430)^-$ onto the $m_{\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), at $K^*(892)$ (top right), in between $K^*(892)$ and $K_2^*(1430)$ (middle left), at $K_2^*(1430)$ (middle right), and above $K_2^*(1430)$ (bottom).

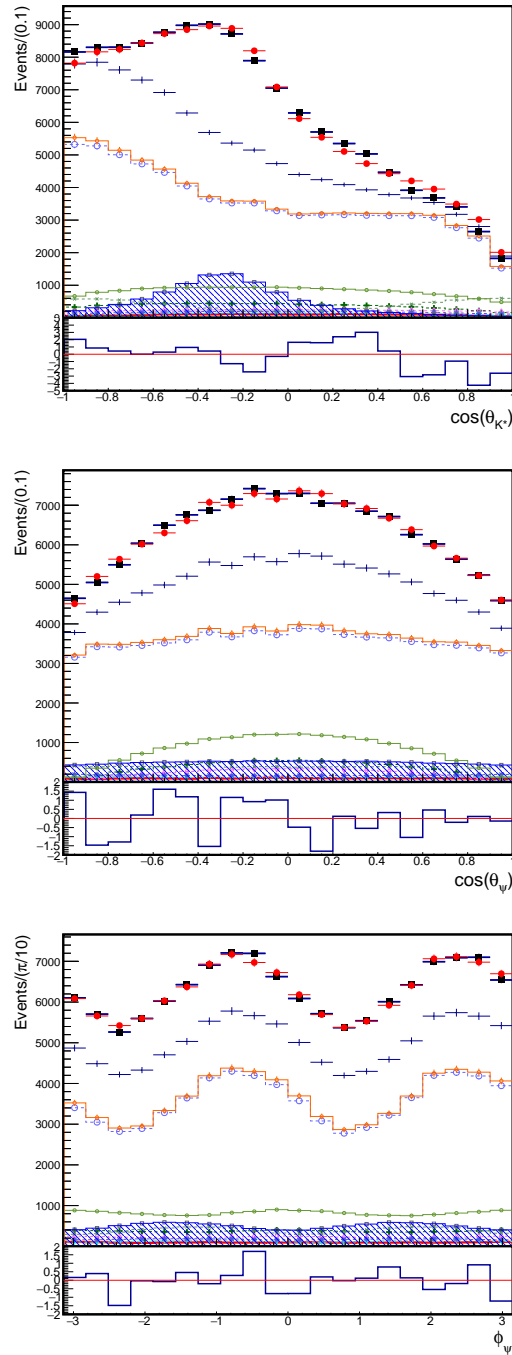


Figure 55: Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the 2014 LHCb model 4D amplitude fit (red) with $Z(4430)^-$ onto the angular variables of the fit: cosine of the K^{*0} helicity angle (top), cosine of the ψ helicity angle (middle), and the angle between the K^{*0} and ψ decay planes (ϕ).

11.3 $Z(4430)^-$ using a Model Independent Approach

An important part of the previous LHCb publication was a demonstration of resonant phase-motion of $J^P = 1^+ \psi(2S)\pi$ amplitude in the $Z(4430)^-$ region. This was accomplished by replacing the Breit Wigner $Z(4430)^-$ shape with 6 complex amplitudes, equally spaced in $m_{\psi(2S)\pi}^2$ bins in the region 18.0 - 21.5 GeV², and these complex amplitudes were free to float in the fit. This follows a similar approach as described later in Sec. 7, except that like previously, we fit here only one set of helicity couplings stemming from the S-wave constraint on $Z^- \rightarrow \psi(2S)\pi$ decay. The fit projections are shown in Fig. 56- 58. They exhibit similar agreements and disagreements with the new data as discussed in Sec. 11.2. The obtained Argand plot is presented in Fig. 59, where it is compared to the one previously published. The amplitude scale convention is different than in the previous analysis, thus absolute values of amplitude magnitudes should not be compared. Overall phase of the diagram is also different, which is likely due to small differences in the size of various K^{*0} components. However, the counter-clockwise relative phase running with $\psi(2S)\pi$ mass, and nearly circular shape of the amplitude is the same, with much smaller errors than previously obtained.

In conclusion, when treating the new dataset with the previous approach we arrive at the same conclusions as published by LHCb in 2014 [7]. However, the increased data sample also reveals that the old model is insufficient in describing all features in the data, therefore the model requires additional components. Additional K^{*0} components have been added as already discussed in Sec. 10. Following sections discuss improvements to the model of the exotic hadron sector, as well as the amplitude analysis of the $B^0 \rightarrow J/\psi\pi^-K^+$ data, which has never been published by LHCb.

Table 9: Fit quality metrics comparing 2014 K^{*0} -only fit with 2014 model independent fit on current dataset.

Metric	2014 K^{*0} -only model	plus $Z(4430)^-$ in model independent form
NDF	34	48
$\log L$ ($\Delta(-2 \log L)$)	-100,113	-101,527 (2,828)
2D χ^2 1024 bins ($\Delta\chi^2$)	3510	1908 (1,602)
4D χ^2 4096 bins ($\Delta\chi^2$)	6571	5062 (1,509)

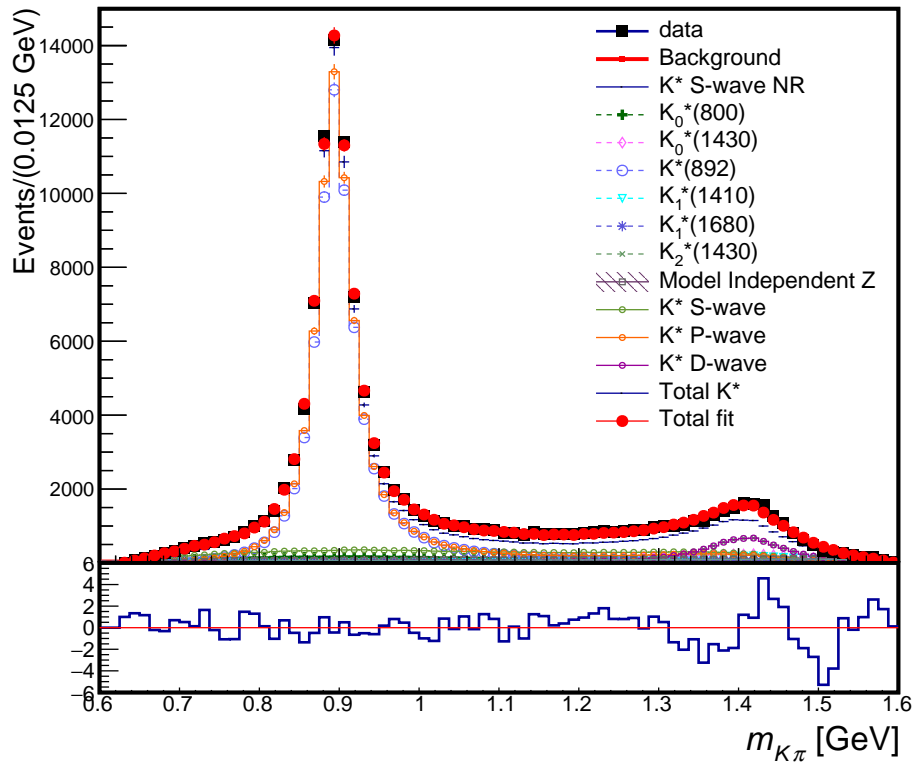


Figure 56: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) using the 2014 LHCb model independent approach.

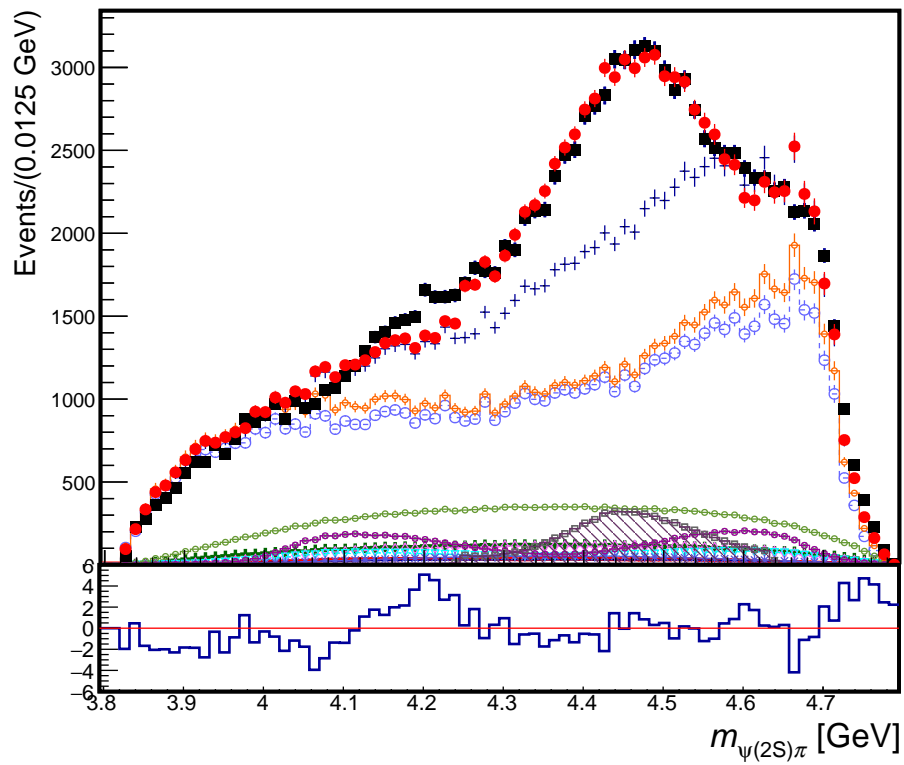


Figure 57: The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) using the 2014 LHCb model independent approach.

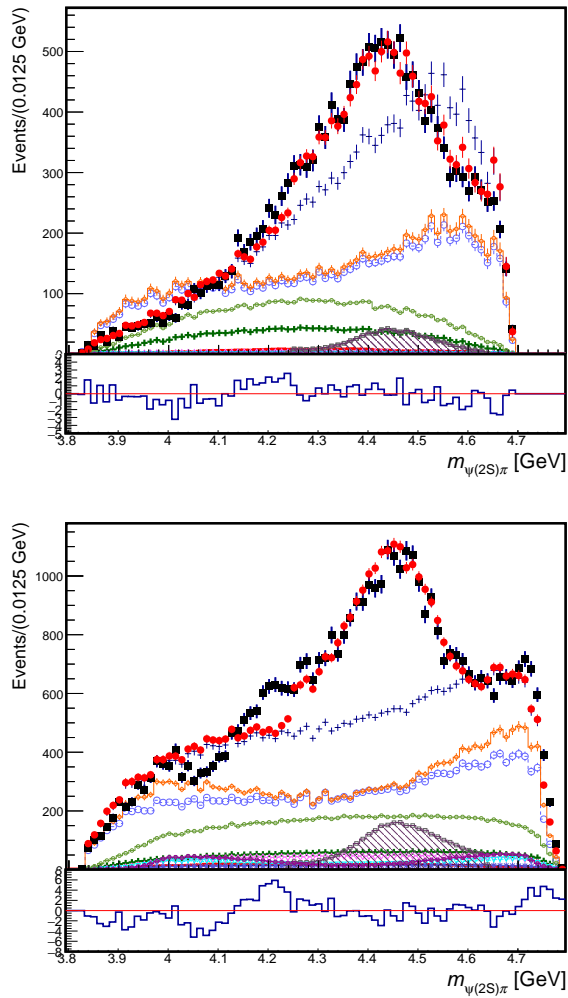


Figure 58: Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the 2014 LHCb model 4D amplitude fit (red) with $Z(4430)^-$ as model independent contribution onto the $m_{\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top) and in between $K^*(892)$ and $K_2^*(1430)$ (bottom).

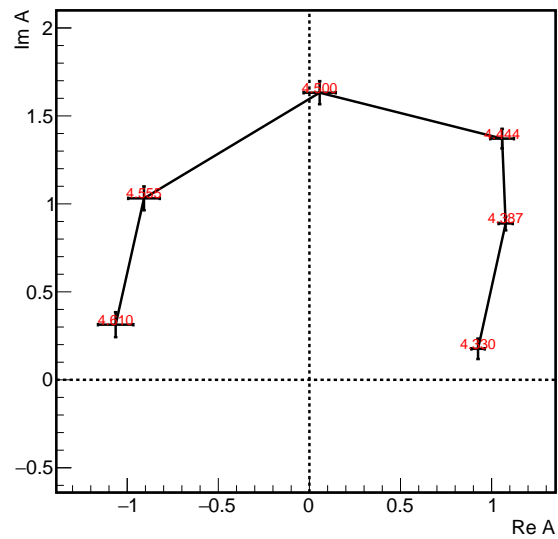
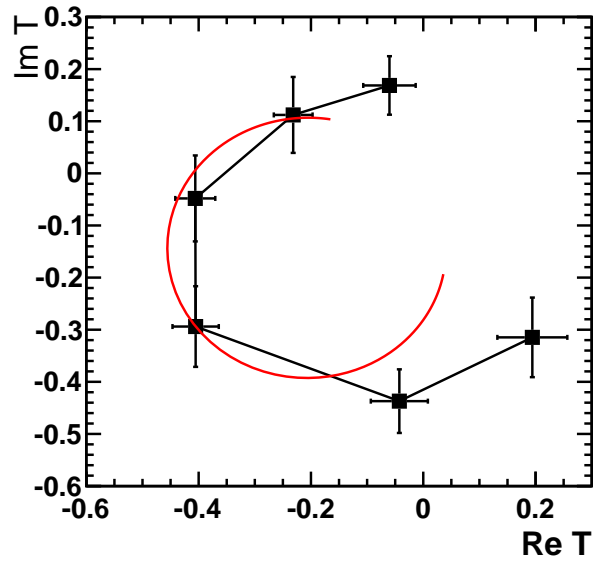


Figure 59: Argand plot for model independent Z contribution from 2014 analysis [7] (top) and current analysis (bottom). The red circle represents a Breit-Wigner for the $Z(4430)^-$. Each complex amplitude is labeled with its $m_{\psi\pi}$ bin center in the Argand plot from the current analysis.

12 Resonant Z models with $\psi(2S)$ channel

Since the 2014 LHCb model [7] is insufficient to represent the current dataset in all detail, here we explore improving a Z model to reconcile the disagreements between these earlier fits and the data. Z s are added to the model one by one, testing all possible J^P contributions for up to spin $J = 2$. Based on a combination of significance from goodness-of-fit metrics and improvements in projections of kinematic variables, we continue to add Z resonances until no further improvements can be made.

12.1 Models with one $Z^- \rightarrow \psi(2S)\pi^-$

Fits were performed, each adding a Z with unique J^P to the K^* -only model in Section 10 for $B^0 \rightarrow \psi(2S)K\pi$. This is similar to the $Z(4430)^-$ spin analysis performed previously [7] with the smaller data set. However, the K^{*0} model has been extended since then as discussed in Sec. 9-10. We also no longer impose the S -wave constraint on Z^- helicity couplings.

The fit quality metrics for each of these fits are presented in Table 10. The best such fit is adding a $J^P = 1^+$ Z , which is the same conclusions as reached in the previous amplitude analyses of this channel [7, 12]. The significance of this contribution is 36.8σ , which is as expected much higher than 18.7σ obtained in the previous LHCb analysis [7]. The mass, 4477 ± 4 MeV, and width, 226 ± 8 MeV, are consistent within the systematic errors with the previous LHCb results [7], $4475 \pm 7_{-25}^{+15}$ MeV and $172 \pm 13_{-34}^{+37}$ MeV.

Table 10: Goodness-of-fit metrics for adding the first Z with the given J^P to the $\psi(2S)$ channel.

Variable	K^* -only	$0^- Z$	$1^- Z$	$1^+ Z$	$2^- Z$	$2^+ Z$
nPar	58	62	62	64	64	62
$\log L$	-101529	-102075	-101579	-102238	-101866	-101560
$\Delta(-2 \log L)$	-	1091	98	1418	674	62
2-D χ^2 (1024 bins)	1907	1458	1800	1285	1540	1833
2-D $\Delta\chi^2$	-	449	106	621	366	73
4-D χ^2 (4096 bins)	5137	4746	5038	4581	4839	5075
4-D $\Delta\chi^2$	-	391	99	556	298	62

Table 11: Mass and width of the Z added to the $\psi(2S)$ channel model.

Contribution	J^P	M (GeV)	Γ (GeV)
Z(4430)	1^+	4.477 ± 0.004	0.226 ± 0.008
Ref. [7]	1^+	$4.475 \pm 0.007^{+0.015}_{-0.025}$	$0.172 \pm 0.013^{+0.037}_{-0.034}$

The significance of the Z(4430) is apparent in the mass projections. With the K^* -only model, there is a noticeable disagreement at low $m_{K\pi}$ (below $K^*(892)$), but this disagreement is resolved after adding the Z(4430) as seen in Figure 60. Even more evident is the improvement in the $\psi(2S)\pi$ distribution in Figure 61, which shows much better agreement with the data near the Z(4430) pole when compared to the K^* -only model.

Despite these improvements to the overall mass distributions, the $m_{\psi(2S)\pi}$ projections in the different $m_{K\pi}$ slices still show disagreement at lower mass ($m_{\psi(2S)\pi} < 4.2$ GeV). This is best seen in Figure 62, which shows the $m_{\psi(2S)\pi}$ projection in the $m_{K\pi}$ slices below $K^*(892)$ and between $K^*(892)$ and $K^*(1430)$. Thus, we need to investigate if this model can be improved by included another $\psi(2S)\pi$ resonance.

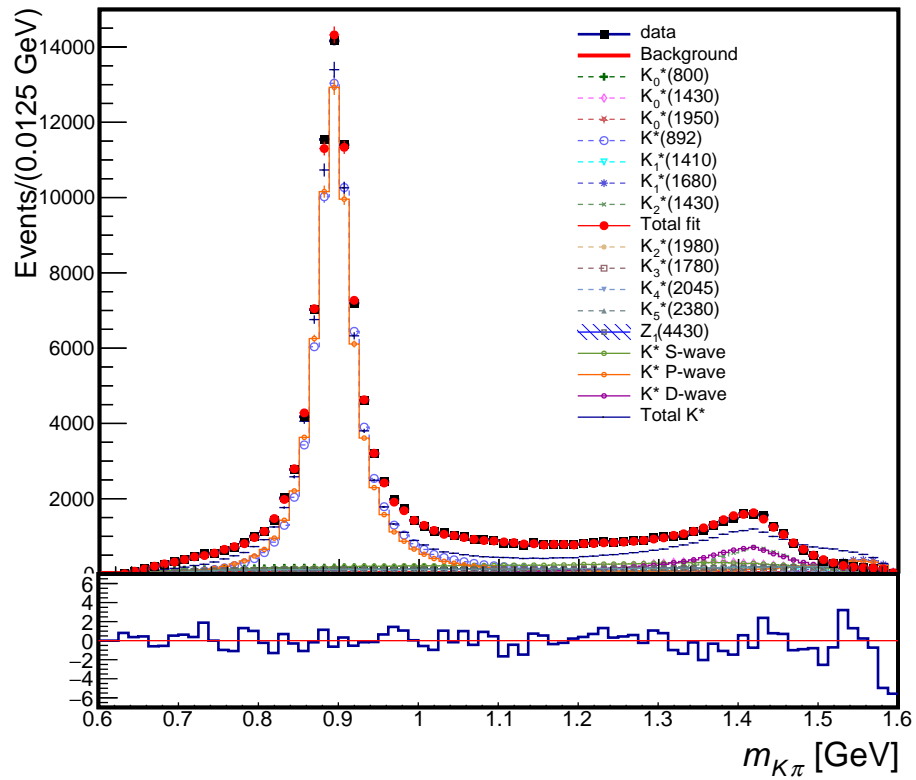


Figure 60: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) which adds a $J^P = 1^+$ Z.

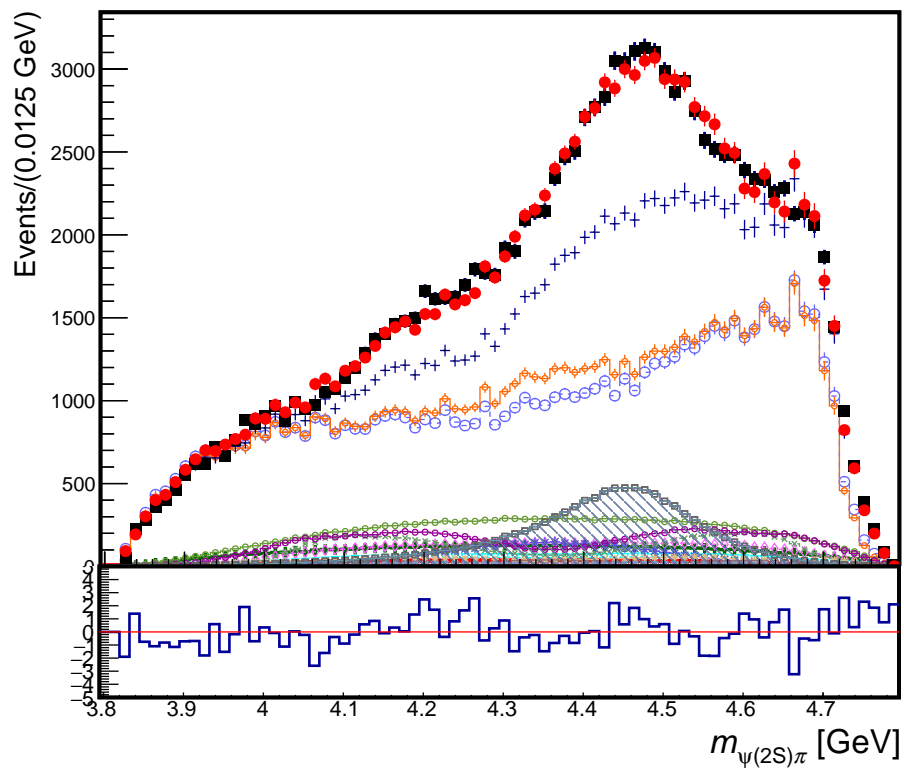


Figure 61: The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) which adds a $J^P = 1^+ Z$.

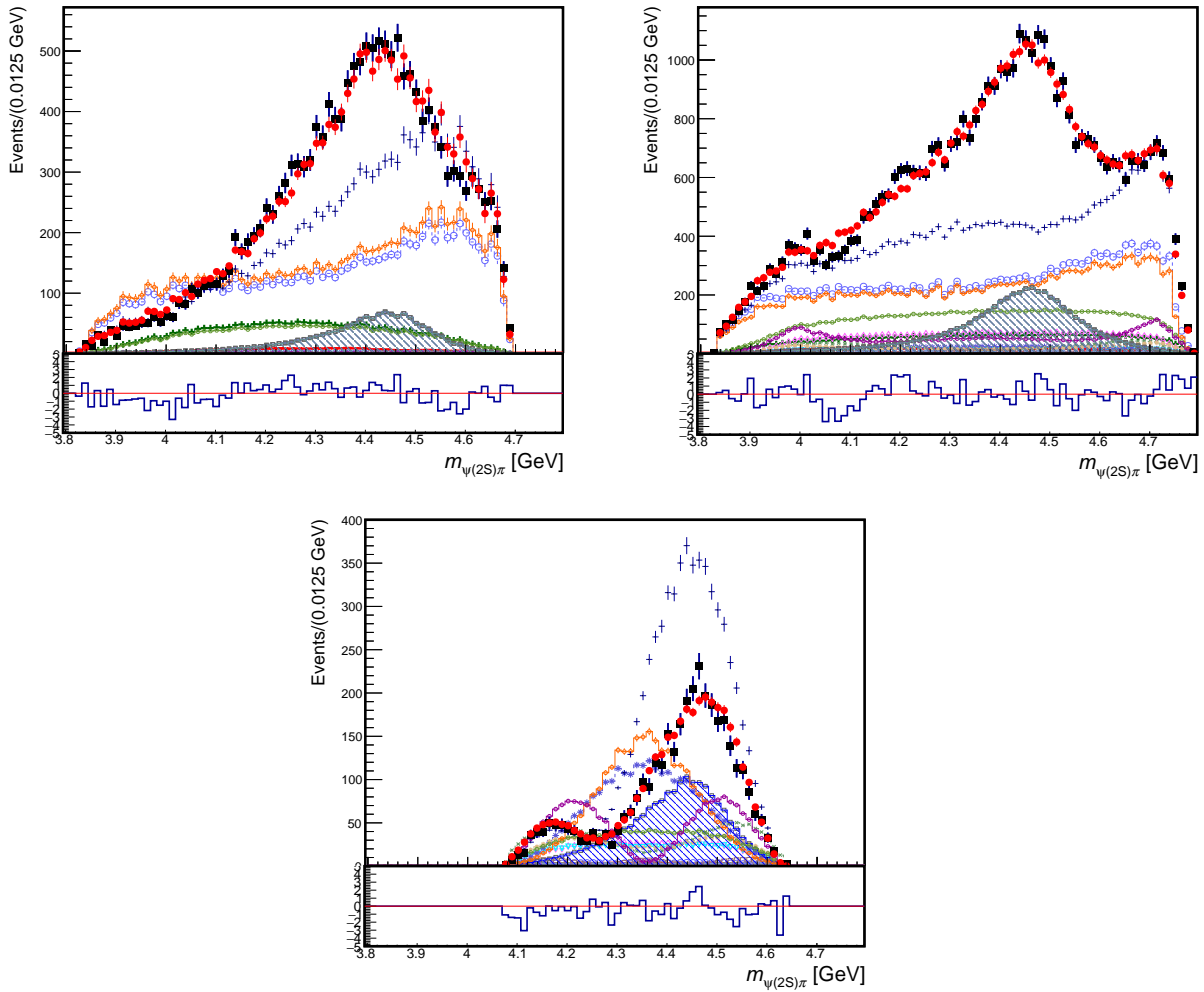


Figure 62: Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the 4D amplitude fit (red) with $Z(4430)^-$ as a $J^P = 1^+$ resonance onto the $m_{\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $K_2^*(1430)$ (bottom).

12.2 Models with two $Z^- \rightarrow \psi(2S)\pi^-$

To improve upon the 1 Z model, more fits were performed, once again adding a Z with various J^P values to the model for $B^0 \rightarrow \psi(2S)\pi^- K^+$. The fit quality metrics for each of these fits are presented in Table 12, which finds adding another $J^P = 1^+$ Z significant at 17.6σ in terms of log-likelihood, with significant improvements in the 2-D and 4-D χ^2 s, as well. This second 1^+ Z was added to the K-matrix with the Z(4430) from the 1 Z model, with both Zs masses and widths allowed to float, and the fit results for these parameters are in Table 13. The new Z resonance, which we call $Z(4200)^-$, has mass at 4.120 ± 0.006 GeV and a width of 0.183 ± 0.012 GeV, and the fit now prefers the $Z(4430)^-$ pole at $m=4.539 \pm 0.006$ GeV but maintains the width at $\Gamma=0.226 \pm 0.010$ GeV.

Table 12: Goodness-of-fit metrics for adding a second Z with the given J^P to the $\psi(2S)$ channel.

Variable	1 Z Only	0^- Z	1^- Z	1^+ Z	2^- Z	2^+ Z
nPar	64	68	68	70	70	68
$\log L$	-102238	-102311	-102276	-102417	-102342	-102244
$\Delta(-2 \log L)$	-	145	75	358	208	11
2-D χ^2 (1024 bins)	1285	1213	1259	1151	1253	1288
2-D $\Delta\chi^2$	-	73	26	134	32	-3
4-D χ^2 (4096 bins)	4581	4526	4561	4464	4541	4585
4-D $\Delta\chi^2$	-	55	20	117	40	-4

Table 13: Masses and widths of the Zs in the 2 Z model for the $\psi(2S)$ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
Z(4430)	1^+	4.539 ± 0.006	0.226 ± 0.010
Z(4200)	1^+	4.120 ± 0.006	0.183 ± 0.012

The overall mass projections for $K\pi$ and $\psi(2S)\pi$ remain overall pretty good in terms of their pull plots as seen in Figures 63 and 64, but the addition of the second 1^+ Z shows improvement in the $\psi(2S)\pi$ mass distribution in the different $m_{K\pi}$ slices. The disagreement between the fit and the data from the Z model in the range $4.0 < m_{\psi(2S)\pi} < 4.25$ GeV is no longer present in these slices when adding the second Z, as seen in Figure 65, most notably in the "middle" slice of $m_{K\pi}$ that falls between $K^*(892)$ and $K^*(1430)$.

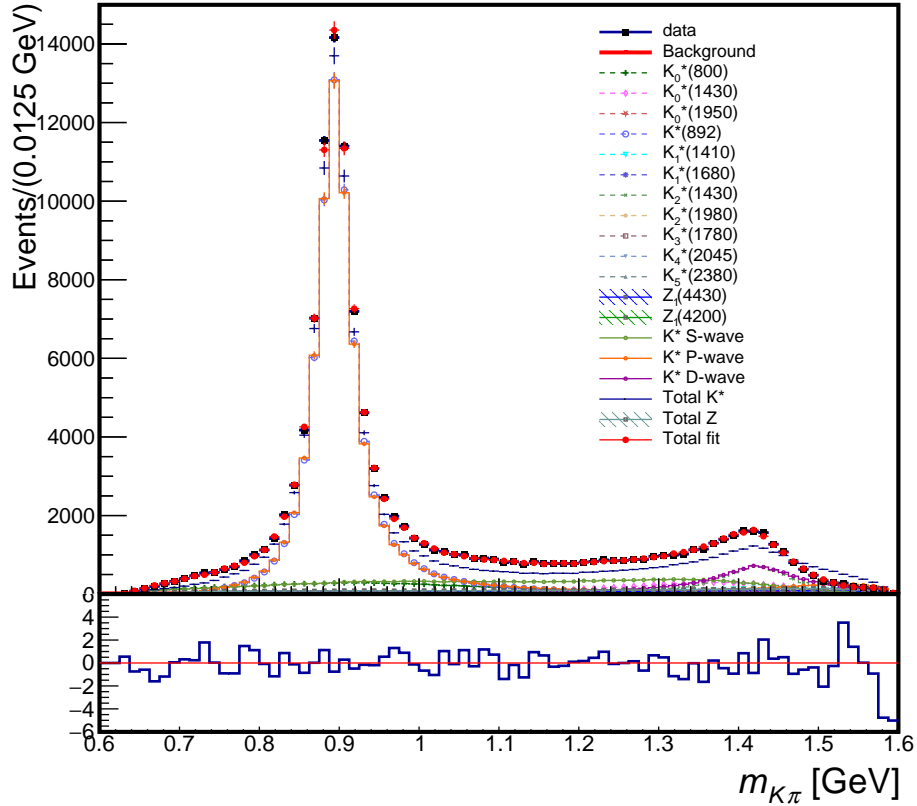


Figure 63: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) which adds a second $J^P = 1^+$ Z.

The need for a second $Z^- \rightarrow \psi(2S)\pi^-$ resonance at the lower mass was already reported in the previous LHCb analysis of the $B^0 \rightarrow \psi(2S)\pi^-K^+$ channel [7]. However, with the smaller data set, both 0^- and 1^+ resonances were giving similar fit qualities, with 1σ preference for the 0^- hypothesis. The previous amplitude analysis was adding two 1^+ resonances in Breit-Wigner form, while the present model adds two 1^+ poles into the same K -matrix term. The latter prescription improves analytical properties of the amplitude, therefore it is better motivated theoretically. As discussed previously, also the K^{*0} model has been improved. With these changes and with much larger data set, our analysis prefers the 1^+ hypothesis by significant margins: 12.2σ over 2^- , 14.6σ over 0^- , 16.8σ over 1^- , and 18.6σ over 2^+ , as estimated by taking the square-root of the $-2\log L$ differences as appropriate for testing the hypotheses which are not nested in each other [7].

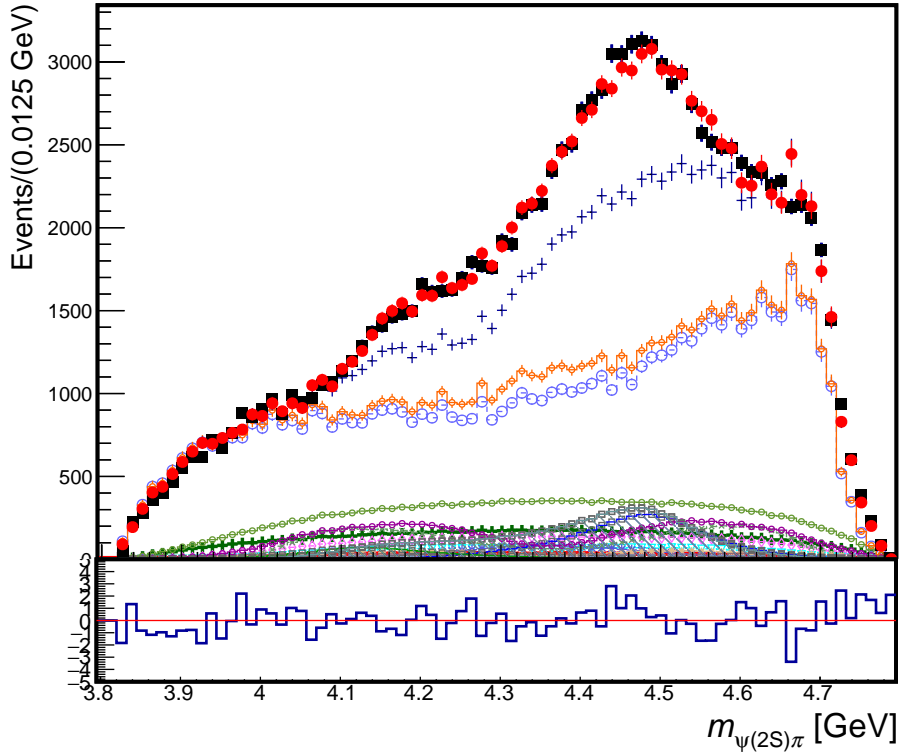


Figure 64: The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) which adds a second $J^P = 1^+$ Z near $m_{\psi(2S)\pi} = 4.1\text{GeV}$.

The two $1^+ Z^- \rightarrow \psi(2S)\pi^-$ model is reminiscent of the two $1^+ Z^- \rightarrow J/\psi\pi^-$ model deduced by the Belle experiment from the amplitude analysis of the $B^0 \rightarrow J/\psi\pi^-K^+$ data [16], in which the lower mass Breit-Wigner resonances was claimed at $4.196^{+0.031}_{-0.029} {}^{+0.017}_{-0.013}$ GeV, with a width of $0.370^{+0.070}_{-0.070} {}^{+0.070}_{-0.132}$ GeV while also requiring the presence of the higher mass $1^+ Z(4430)^- \rightarrow J/\psi\pi^-$ resonance with the mass and width fixed at their results from the analysis of the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data. Since Belle did not use the K -matrix formulation, the masses and widths cannot be rigorously compared to our results. However, qualitatively the models are similar.

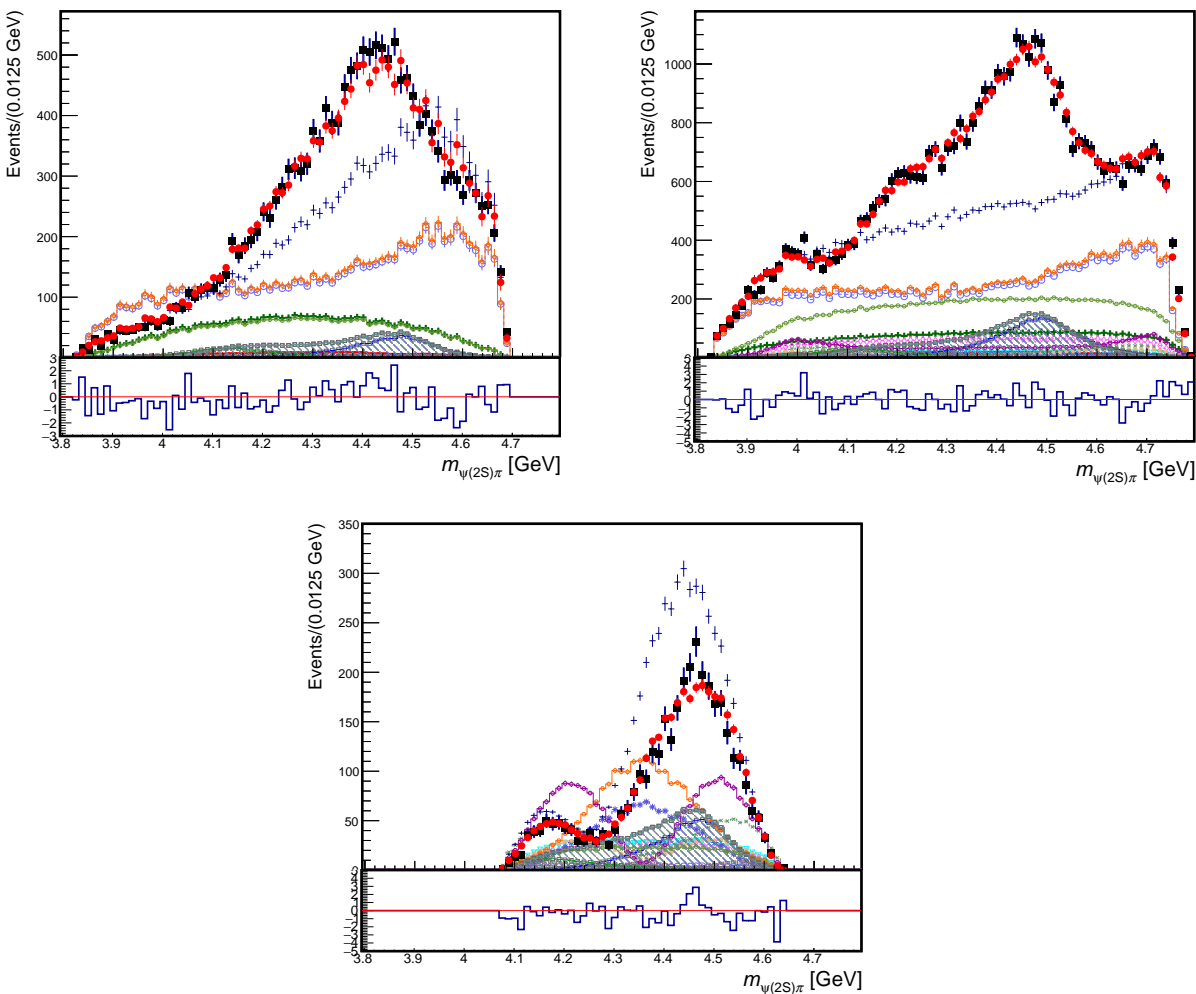


Figure 65: Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the 4D amplitude fit (red) with $2 J^P = 1^+$ Zs in a K-matrix onto the $m_{\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $K_2^*(1430)$ (bottom).

12.3 Models with three $Z^- \rightarrow \psi(2S)\pi^-$

Fits are performed again to improve upon the 2 Z model for $B^0 \rightarrow \psi(2S)\pi^-K^+$. The fit quality metrics for each of these fits are presented in Table 14, which finds adding a $J^P = 2^-$ Z significant at 11.3σ in terms of log-likelihood, and a third Z modestly improves the 2-D and 4-D χ^2 s. These 2^- quantum numbers are preferred over the other spin-parity hypotheses by 9σ . This is the first significant evidence that exotic $\psi\pi^-$ resonances, other than in 1^+ wave exist. The new Z is added as a separate resonance from the two 1^+ Zs in a K-matrix together, with all Zs masses and widths allowed to float, and the fit results for these parameters are in Table 15. The new Z resonance, which we call $Z(4470)^-$, has mass at 4.465 ± 0.006 GeV and a width of 0.125 ± 0.014 GeV, and the masses and widths of the other Zs remain largely the same.

Table 14: Goodness-of-fit metrics for adding a third Z with the given J^P to the $\psi(2S)$ channel.

Variable	2 Z Only	0^- Z	1^- Z	1^+ Z	2^- Z	2^+ Z
nPar	70	74	74	76	76	74
$\log L$	-102417	-102451	-102431	-102460	-102501	-102423
$\Delta(-2 \log L)$	-	68	28	84	166	11
2-D χ^2 (1024 bins)	1151	1127	1153	1123	1115	1149
2-D $\Delta\chi^2$	-	24	-2	28	36	2
4-D χ^2 (4096 bins)	4464	4458	4456	4440	4438	4459
4-D $\Delta\chi^2$	-	6	8	24	27	5

Table 15: Masses and widths of the Zs in the 3 Z model for the $\psi(2S)$ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
Z(4430)	1^+	4.556 ± 0.006	0.234 ± 0.009
Z(4200)	1^+	4.117 ± 0.004	0.164 ± 0.009
Z(4470)	2^-	4.465 ± 0.006	0.125 ± 0.014

The overall mass projection for $K\pi$ (Figure 66) improved at high mass (> 1.5 GeV), as a small disagreement between the 2 Z fit and data is not present in the 3 Z fit. The $\psi(2S)\pi$ mass projection in Fig. 67 remains mostly the same, and both of the overall projections look good. There is also improvement in the $\psi(2S)\pi$ mass distribution in the different $m_{K\pi}$ slices

in Fig. 68. The pulls improve near the peak of the new Z in both the slice below $K^*(892)$ and in the slice above $K_2^*(1430)$. In addition, there is much smaller interference between the K^* and Z waves than before as seen in the slice above $K_2^*(1430)$, as the total K^* wave is no longer well above the fit in the plot.

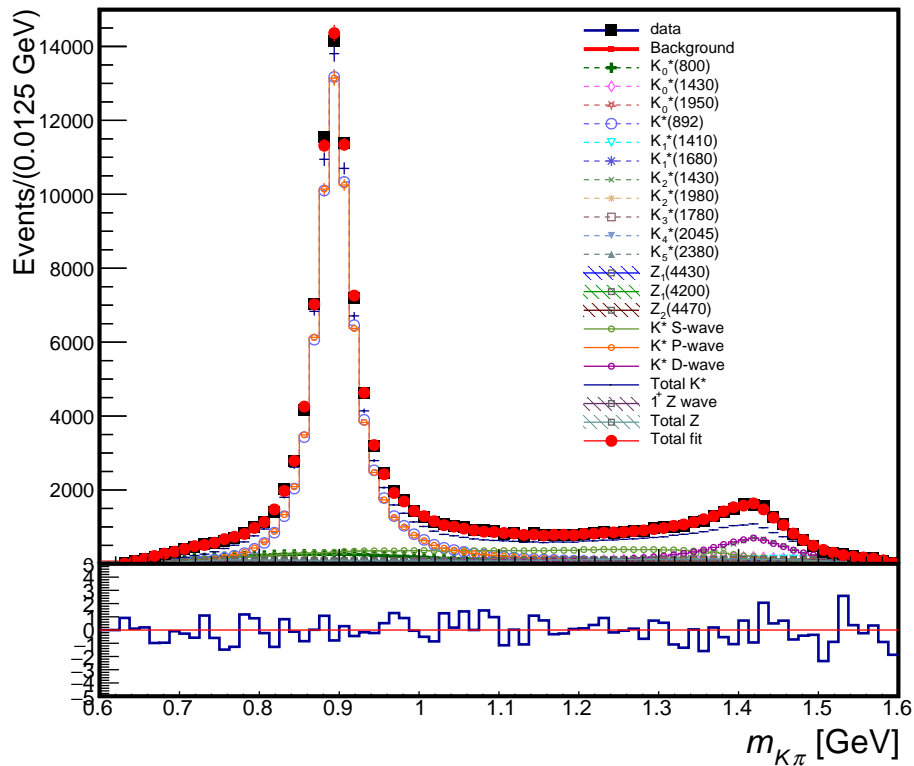


Figure 66: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) which adds a $J^P = 2^-$ Z into the model with two 1^+ Zs in a K-matrix.

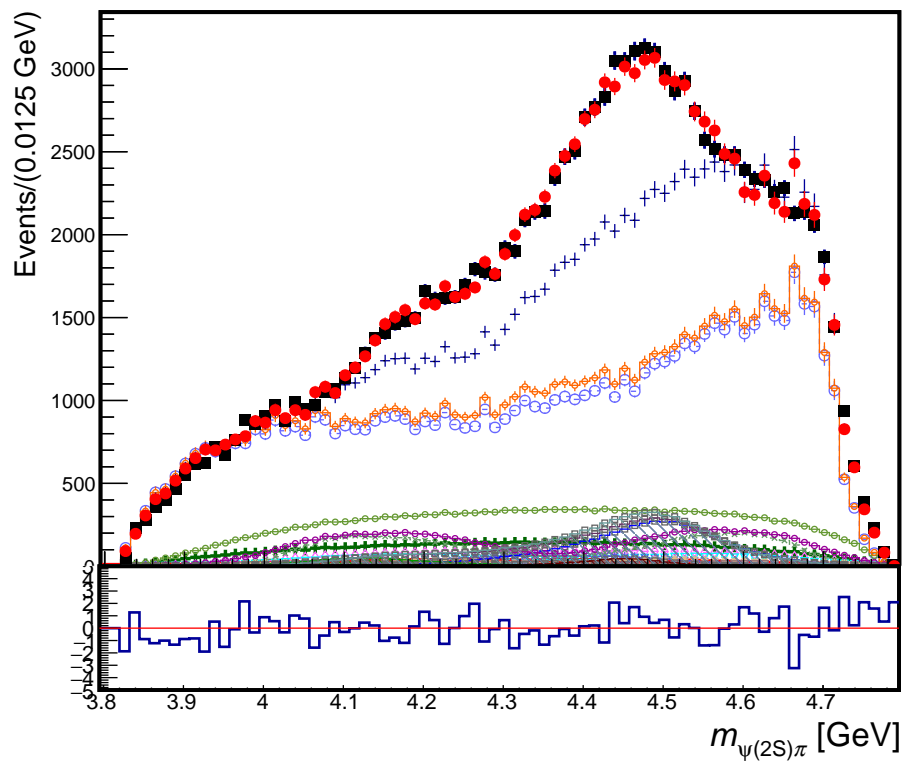


Figure 67: The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) which adds a $J^P = 2^-$ Z near $m_{\psi(2S)\pi} = 4.4$ GeV into the model with two 1^+ Zs in a K-matrix.

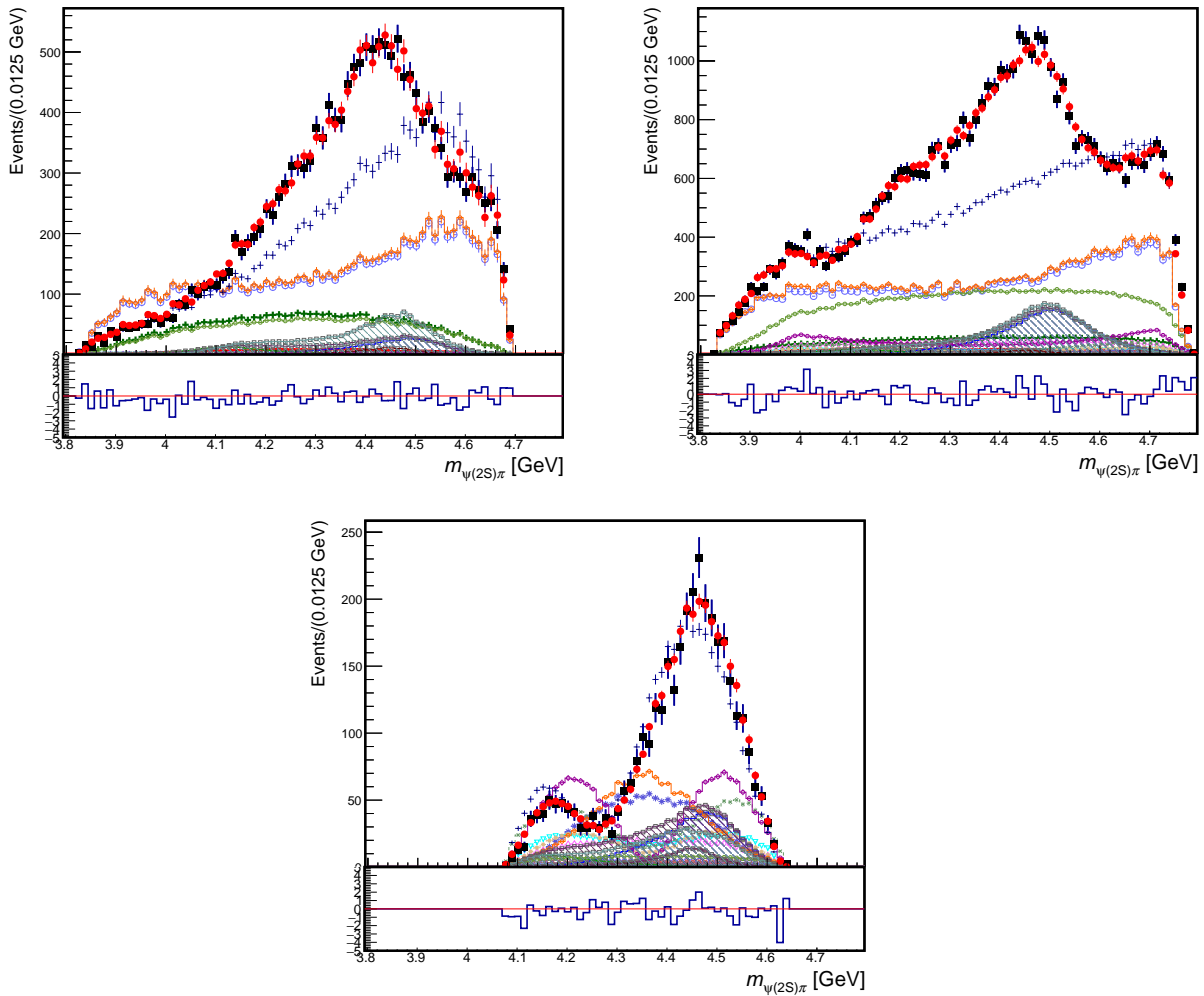


Figure 68: Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the 4D amplitude fit (red) with 2 $J^P = 1^+$ Zs in a K-matrix and 1 $J^P = 2^-$ Z as a separate resonance onto the $m_{\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $K_2^*(1430)$ (bottom).

12.4 Models with four $Z^- \rightarrow \psi(2S)\pi^-$

Fits with a fourth Z were performed on $B^0 \rightarrow \psi(2S)\pi^-K^+$ data, and the fit quality metrics for these are shown in Table 16. Adding a third 1^+ Z to the K-matrix with the $Z(4430)$ and $Z(4200)$ proves to be the most significant at 8.0σ in terms of log-likelihood, and it also slightly improves the 2-D and 4-D χ^2 s. The J^P preference is not as significant as when adding the other Z resonances. The preference over 2^- hypothesis (the second pole in this wave) is 2.8σ and $> 4.1\sigma$ for the other J^P values. The numbers given here include only statistical factors, with systematic effects neglected. We allow all masses and widths of these Zs to float, and their results are found in Table 17. The new Z resonance, which we call Z(4800), has the mass at 4.804 ± 0.007 GeV, which is near the upper kinematic bound, and a relatively narrow width of 0.139 ± 0.025 GeV. The masses and widths of the other Zs remain largely the same, save for Z(4430), which is slightly less broad than it was in the 3 Z fit. The masses and widths of the K^* s that are floating can be found in Table 18, and are consistent with the world average values [22]. The fit fractions for all resonances and partial waves are found in Table 19.

Table 16: Goodness-of-fit metrics for adding a fourth Z with the given J^P to the $\psi(2S)$ channel.

Variable	3 Z Only	0^- Z	1^- Z	1^+ Z	2^- Z	2^+ Z
nPar	76	80	80	82	82	80
$\log L$	-102501	-102530	-102533	-102542	-102538	-102533
$\Delta(-2 \log L)$	-	59	66	83	75	66
2-D χ^2 (1024 bins)	1115	1085	1097	1081	1096	1097
2-D $\Delta\chi^2$	-	30	18	34	19	18
4-D χ^2 (4096 bins)	4438	4404	4420	4395	4409	4420
4-D $\Delta\chi^2$	-	34	18	43	29	18

The overall mass projection for $K\pi$ (Figure 69) remains the same and good, while the overall mass projection for $\psi(2S)\pi$ (Figure 70) improves slightly at high mass near the new Z but largely remains the same and good. There is also improvement in the $\psi(2S)\pi$ mass distribution in the $m_{K\pi}$ slice between $K^*(892)$ and $K_2^*(1430)$ in Figure 71. There is a small

Table 17: Masses and widths of the Zs in the 4 Z model for the $\psi(2S)$ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
Z(4430)	1^+	4.522 ± 0.009	0.199 ± 0.010
Z(4200)	1^+	4.119 ± 0.007	0.164 ± 0.008
Z(4470)	2^-	4.471 ± 0.005	0.116 ± 0.010
Z(4800)	1^+	4.804 ± 0.007	0.139 ± 0.025

Table 18: Masses and widths of the K^* s in the 4 Z model for the $\psi(2S)$ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
$K^*(892)$	1^-	0.889 ± 0.0001	0.045 ± 0.0003
$K_0^*(800)$	0^+	0.801 ± 0.009	0.317 ± 0.011
$K_2^*(1430)$	2^+	1.416 ± 0.001	0.104 ± 0.002

improvement at high mass near the new Z pole.

Table 19: All resonance fit fractions in the 4 Z fit for the $\psi(2S)$ channel.

Contribution	Fit fraction
K^* Total S-wave	17.2%
$K_0^*(800)$	4.6%
$K_0^*(1430)$	5.5%
$K_0^*(1950)$	1.2%
K^* Total P-wave	58.9%
$K^*(892)$	55.9%
$K_1^*(1410)$	2.5%
$K_1^*(1680)$	1.7%
K^* Total D-wave	7.4%
$K_2^*(1430)$	6.4%
$K_2^*(1980)$	1.0%
$K_3^*(1780)$	0.4%
$K_4^*(2045)$	0.2%
$K_5^*(2380)$	0.08%
Z Total 1^+ wave	5.3%
Z(4430)	2.8%
Z(4200)	1.2%
Z(4800)	0.3%
Z(4470) 2^- wave	0.5%

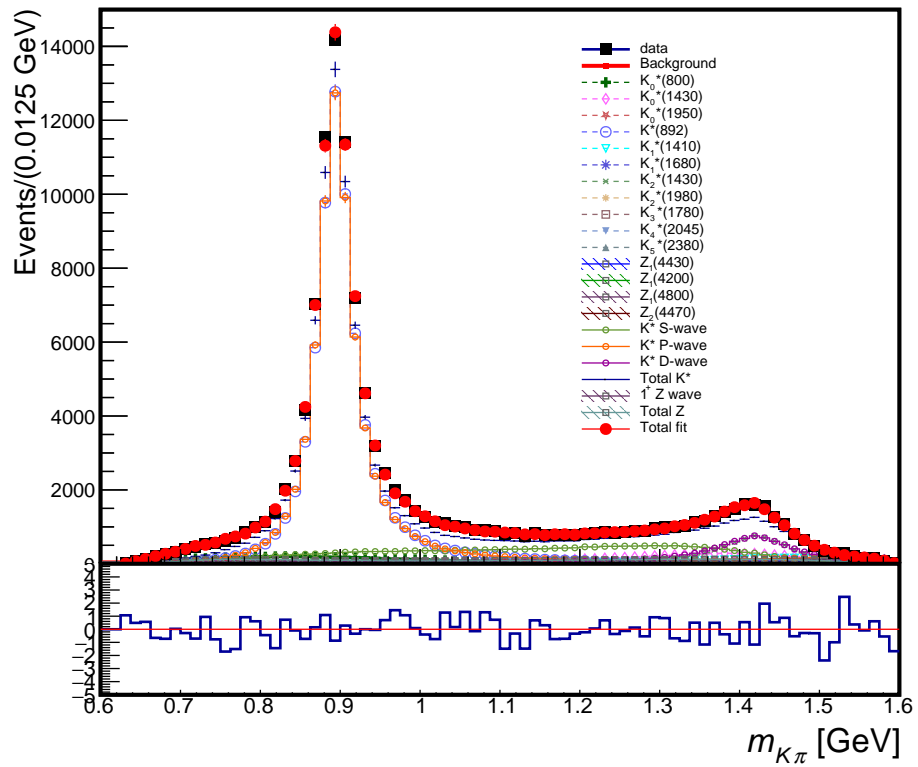


Figure 69: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) which adds a third $J^P = 1^+$ Z into the model with two 1^+ Zs in a K-matrix and a separate 2^- Z resonance.

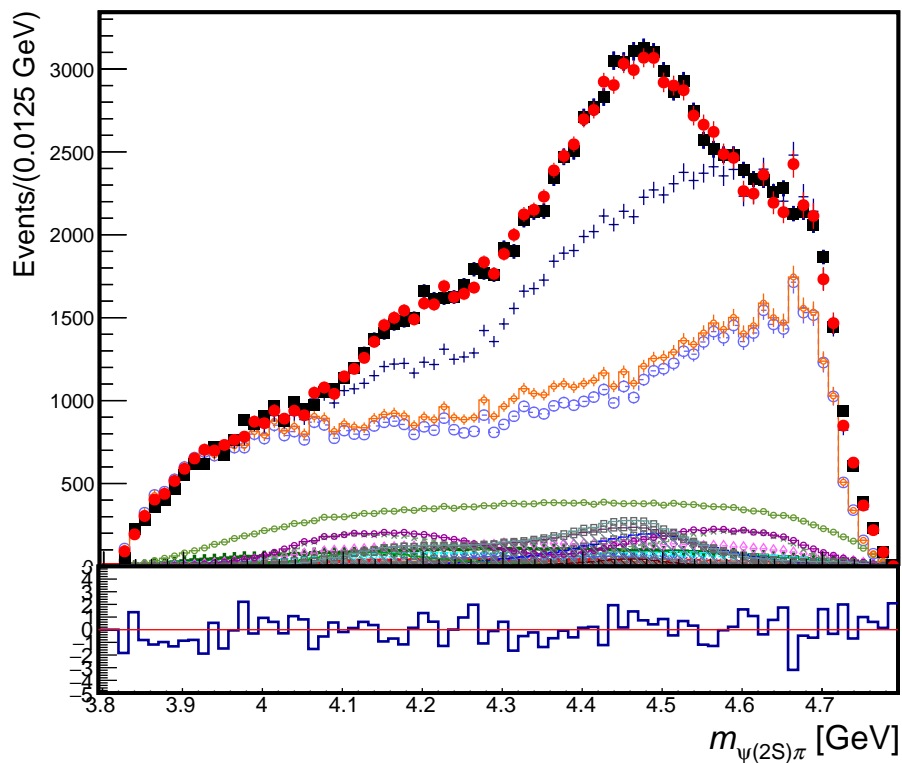


Figure 70: The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) which adds a third $J^P = 1^+$ Z near $m_{\psi(2S)\pi} = 4.8\text{GeV}$ into the model with two 1^+ Zs in a K-matrix and a separate 2^- Z resonance.

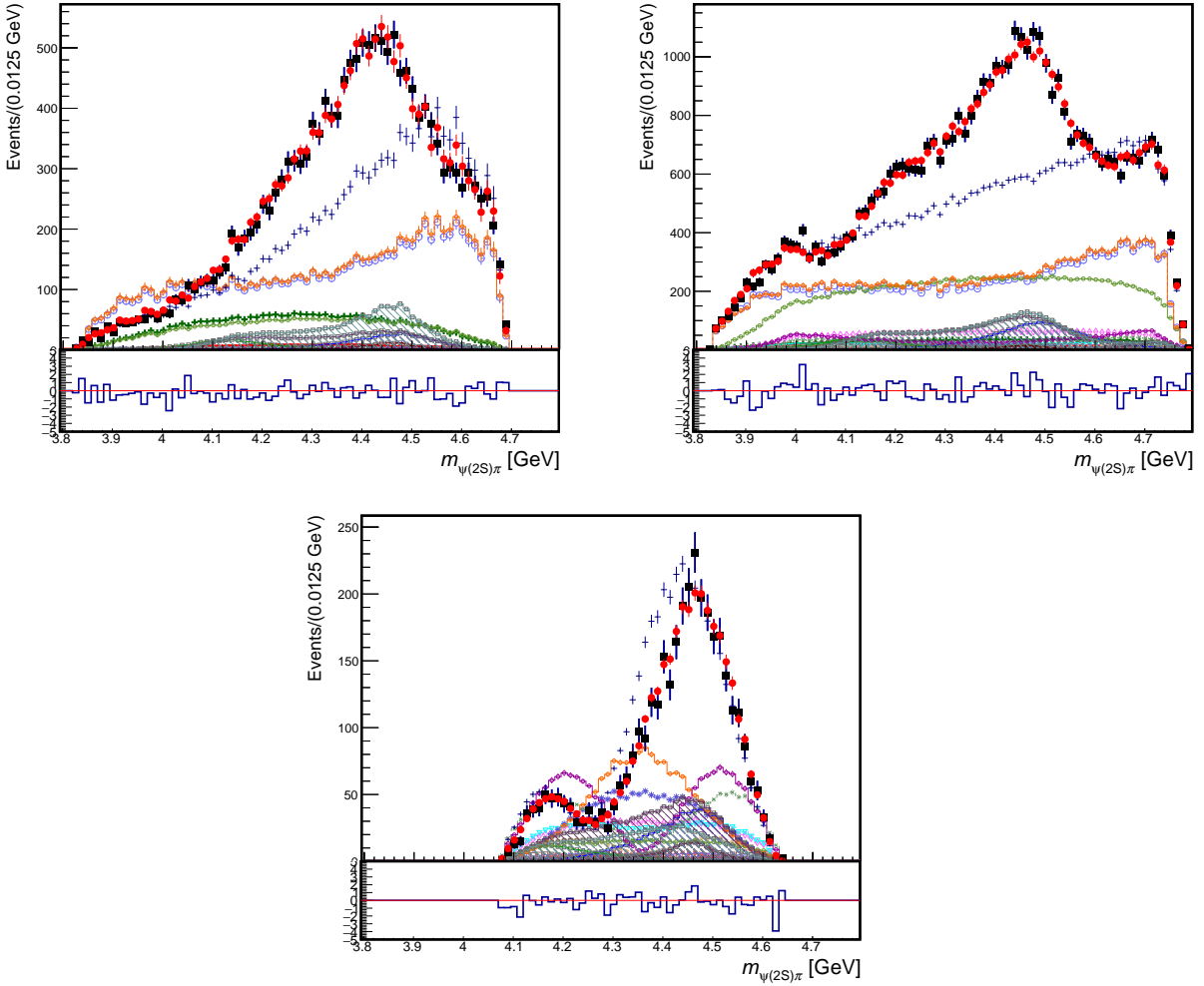


Figure 71: Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the 4D amplitude fit (red) with 3 $J^P = 1^+$ Zs in a K-matrix and 1 $J^P = 2^-$ Z as a separate resonance onto the $m_{\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $K_2^*(1430)$ (bottom).

12.5 Models with five $Z^- \rightarrow \psi(2S)\pi^-$

Fits with a fifth Z were performed on $B^0 \rightarrow \psi(2S)K\pi$, and the fit quality metrics for these are in Table 20. While in terms of log-likelihood the fifth Z component is still 9.7σ significant for the second 2^- pole, the changes in the 2-D χ^2 values calculated on the binned Dalitz plot no longer provide supporting evidence. Its value hardly changes, and it prefers a fourth 1^+ pole instead. Furthermore, the model has reached a complexity where fits converge with warnings of machine accuracy limits. Therefore, we conclude that while the data may contain even more Z resonant poles, most likely in 2^- or 1^+ waves, we cannot reliably extract their parameters.

Table 20: Goodness-of-fit metrics for adding a fifth Z with the given J^P to the $\psi(2S)$ channel.

Variable	4 Z Only	0^- Z	1^- Z	1^+ Z	2^- Z	2^+ Z
nPar	82	86	86	88	88	86
$\log L$	-102542	-102557	-102550	-102576	-102608	-102533
$\Delta(-2\log L)$	-	29	15	67	132	47
2-D χ^2 (1024 bins)	1081	1081	1079	1061	1078	1074
2-D $\Delta\chi^2$	-	0	2	20	3	7
4-D χ^2 (4096 bins)	4395	4391	4388	4378	4360	4393
4-D $\Delta\chi^2$	-	4	7	17	35	2

13 Resonant Z_K models with $\psi(2S)$ channel

This section will explore adding $Z_K^+ \rightarrow \psi(2S)K^+$ resonances into the $B^0 \rightarrow \psi(2S)\pi^-K^+$ model already equipped with the $K^{*0} \rightarrow K^+\pi^-$ contributions and four $Z^- \rightarrow \psi(2S)\pi^-$ states.

13.1 Plots of $m_{\psi(2S)K}$ with no $Z_K^+ \rightarrow \psi(2S)K^+$

To compare with models that add $Z_K^+ \rightarrow \psi(2S)K^+$ states, presented here are the $m_{\psi(2S)K}$ distributions both overall (Fig 72) and in different slices of $m_{K\pi}$ (Fig 73) from the fits done with 4 $Z^- \rightarrow \psi(2S)\pi^-$ states without any Z_K states. Fit quality metrics are in Sec. 12.4.

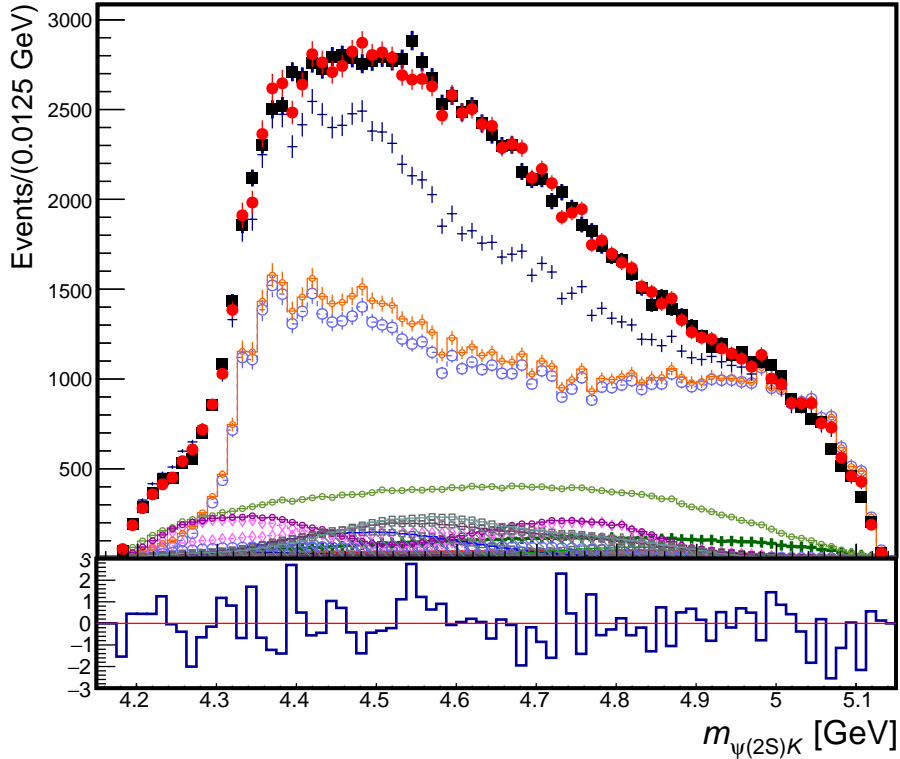


Figure 72: The distribution of $m_{\psi(2S)K}$ for the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black) and the fit (red) which has three $J^P = 1^+$ Zs in a K-matrix and a separate 2^- Z resonance.

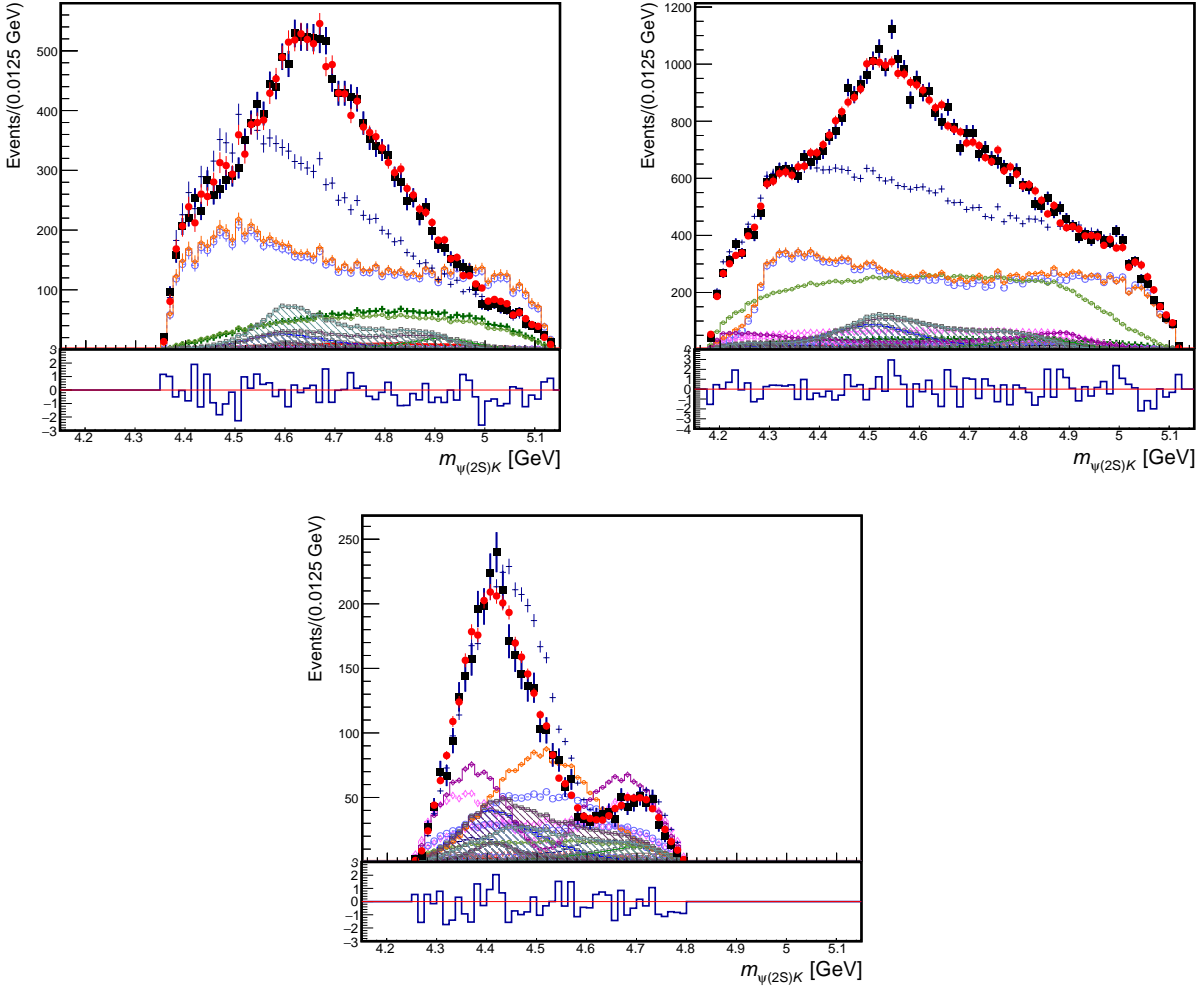


Figure 73: Projections of the $B^0 \rightarrow \psi(2S)\pi^- K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$, $Z(4430)^-$, and $Z(4800)^-$ as a $J^P = 1^+$ K-matrix and $Z(4470)^-$ as a $J^P = 2^-$ resonance onto the $m_{\psi(2S)\pi}$ axis in different slices of $m_{K^*\pi^-}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $K_2^*(1430)$ until the end of $\psi(2S)$ phase space (bottom).

13.2 Models with one $Z_K^+ \rightarrow \psi(2S)K^+$

Fits with a single Z_K^+ state added to the best K^{*0} and Z^- model were performed on $B^0 \rightarrow \psi(2S)\pi^-K^+$ data, and the fit quality metrics for these are given in Table 21.

Adding a $1^- Z_K^+$ is the most significant in terms of the likelihood (7.4σ), and it also slightly improves the 2-D and 4-D χ^2 s. We allow the mass and width of the Z_K^+ to float, and the result is found in Table 24. The new Z_K^+ resonance, which we label $Z_K(4480)$, has mass at 4.498 ± 0.014 GeV and a width of 0.183 ± 0.029 GeV. The masses and widths of the $Z^- \rightarrow \psi(2S)\pi^-$ states do not change much except for Z(4430) which becomes a little bit narrower (see Table 22). The masses and widths of the K^{*} s that are floating can be found in Table 23.

Table 21: Goodness-of-fit metrics for adding one Z_K with the given J^P to the $\psi(2S)$ channel.

Variable	No Z_K	$0^- Z_K$	$1^- Z_K$	$1^+ Z_K$	$2^- Z_K$	$2^+ Z_K$
nPar	82	86	86	88	88	86
$\log L$	-102542	-102568	-102581	-102577	-102564	-102575
$\Delta(-2 \log L)$	-	51	78	70	44	65
2-D χ^2 (1024 bins)	1081	1070	1072	1075	1068	1078
2-D $\Delta\chi^2$	-	11	9	6	13	4
4-D χ^2 (4096 bins)	4396	4382	4377	4381	4377	4381
4-D $\Delta\chi^2$	-	14	18	15	18	14

Table 22: Masses and widths of the Zs in the 1 Z_K model for the $\psi(2S)$ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
Z(4430)	1^+	4.526 ± 0.011	0.130 ± 0.042
Z(4200)	1^+	4.124 ± 0.010	0.166 ± 0.015
Z(4470)	2^-	4.473 ± 0.007	0.119 ± 0.013
Z(4800)	1^+	4.816 ± 0.011	0.130 ± 0.042

The overall mass projections for $K\pi$ (Figure 74) and $\psi(2S)\pi$ (Figure 75) do not change much after adding a Z_K resonance to the 4 Z model. The $K\pi$ slices for the $\psi(2S)\pi$ mass projections also are pretty similar to those in the 4 Z model (Figure 76). The overall mass projection for $\psi(2S)K$ does improve slightly near the new pole (Fig. 77), where a slight

Table 23: Masses and widths of the K^* 's in the 1 Z_K model for the $\psi(2S)$ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
$K^*(892)$	1^-	0.889 ± 0.0002	0.045 ± 0.0003
$K_0^*(800)$	0^+	0.804 ± 0.003	0.306 ± 0.003
$K_2^*(1430)$	2^+	1.415 ± 0.001	0.098 ± 0.003

Table 24: Masses and widths of the Z_K in the 1 Z_K model for the $\psi(2S)$ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
$Z_K(4480)$	1^-	4.498 ± 0.014	0.183 ± 0.029

improvement in the pulls can be seen. The $K\pi$ slices for the $\psi(2S)K$ mass projections do not change much when adding a Z_K as seen in Figure 78.

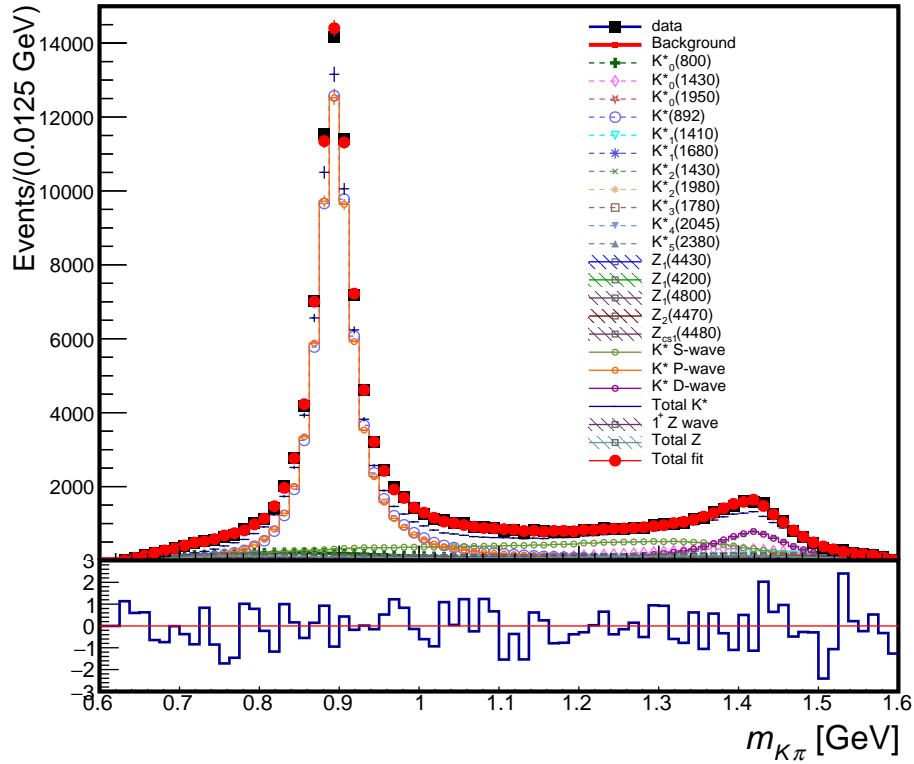


Figure 74: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 1^- Z_K$ into the default model which already has 4 $Z \rightarrow \psi(2S)\pi$ resonances.

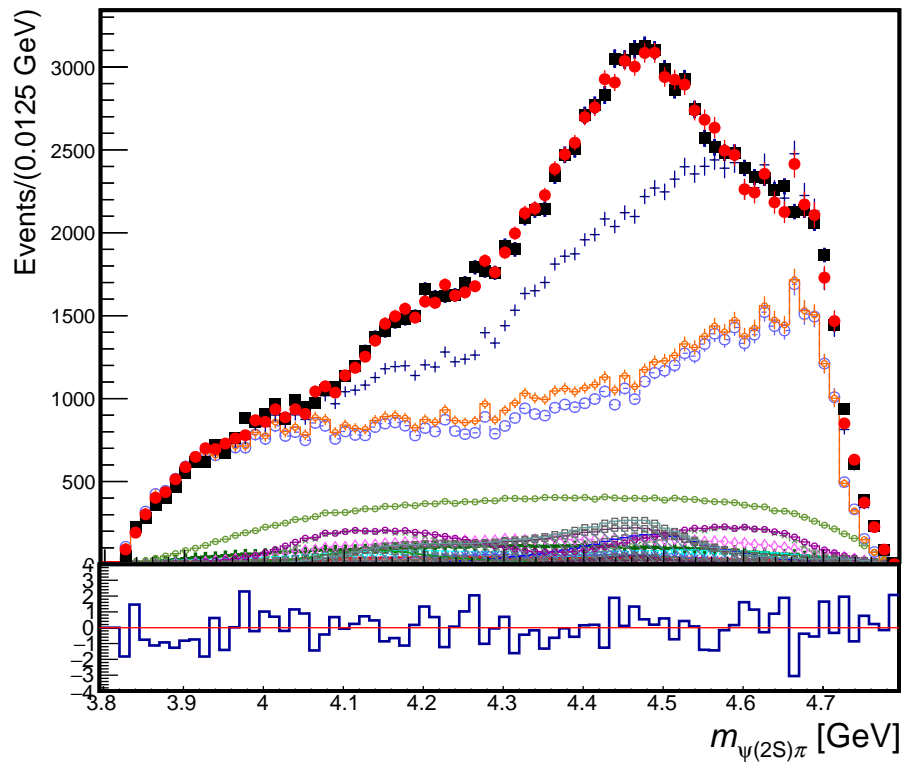


Figure 75: The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^- K^+$ data (black) and the fit (red) which adds a new $J^P = 1^- Z_K$ into the default model which already has 4 $Z \rightarrow \psi(2S)\pi$ resonances.

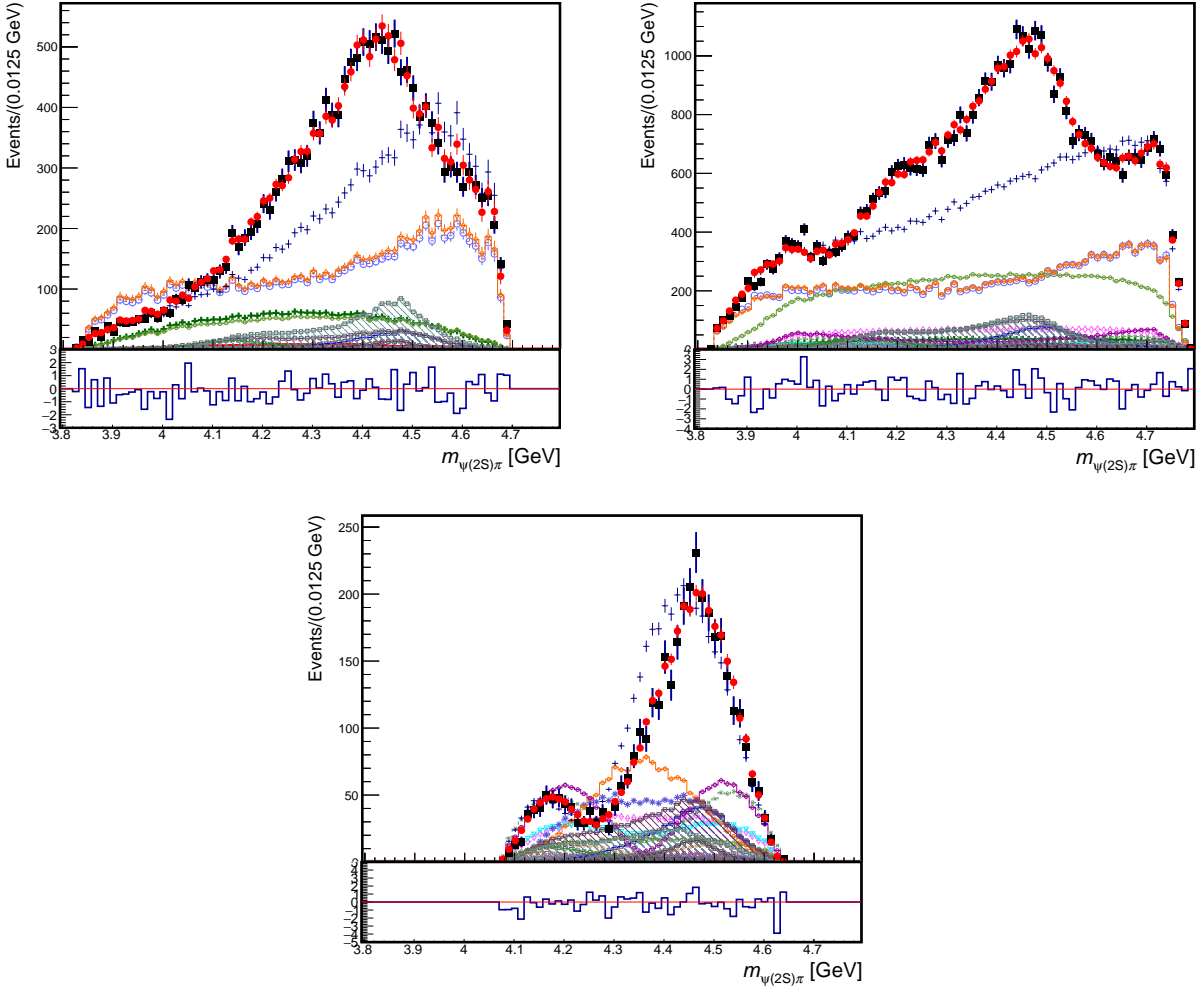


Figure 76: Projections of the $B^0 \rightarrow \psi(2S)\pi^- K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$, $Z(4430)^-$, and $Z(4800)^-$ as a $J^P = 1^+$ K-matrix, $Z(4470)^-$ as a $J^P = 2^-$ resonance, and $Z_{cs}(4480)^-$ as a $J^P = 1^- \psi(2S)K$ resonance. onto the $m_{\psi(2S)\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $K_2^*(1430)$ until the end of $\psi(2S)$ phase space (bottom).

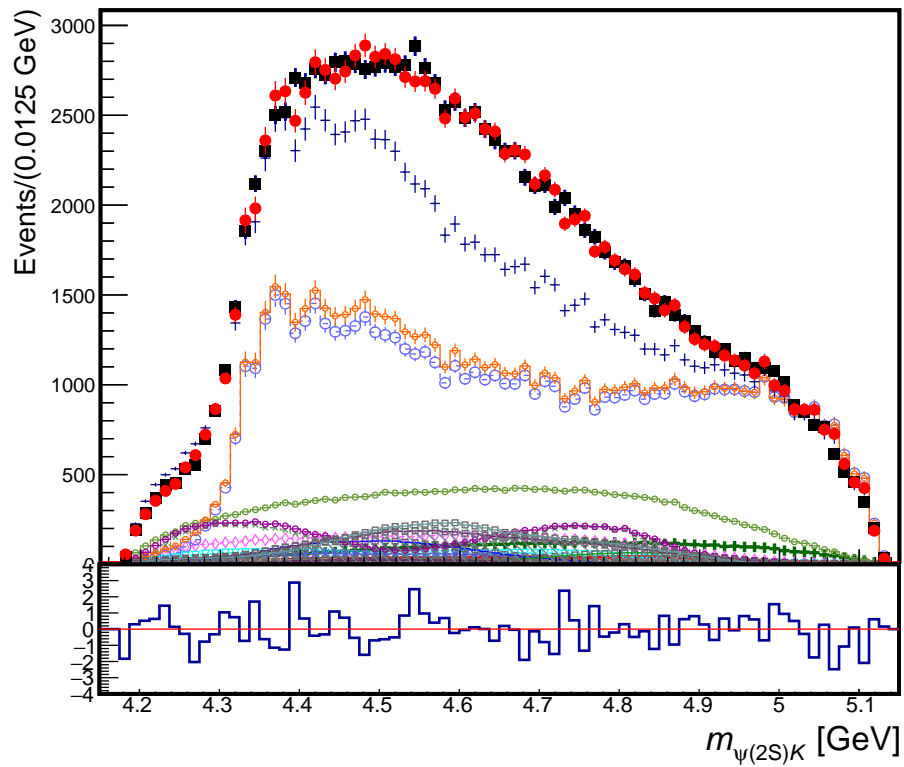


Figure 77: The distribution of $m_{\psi(2S)K}$ for the $B^0 \rightarrow \psi(2S)\pi^- K^+$ data (black) and the fit (red) which adds a new $J^P = 1^- Z_K$ into the default model which already has 4 $Z \rightarrow \psi(2S)\pi$ resonances.

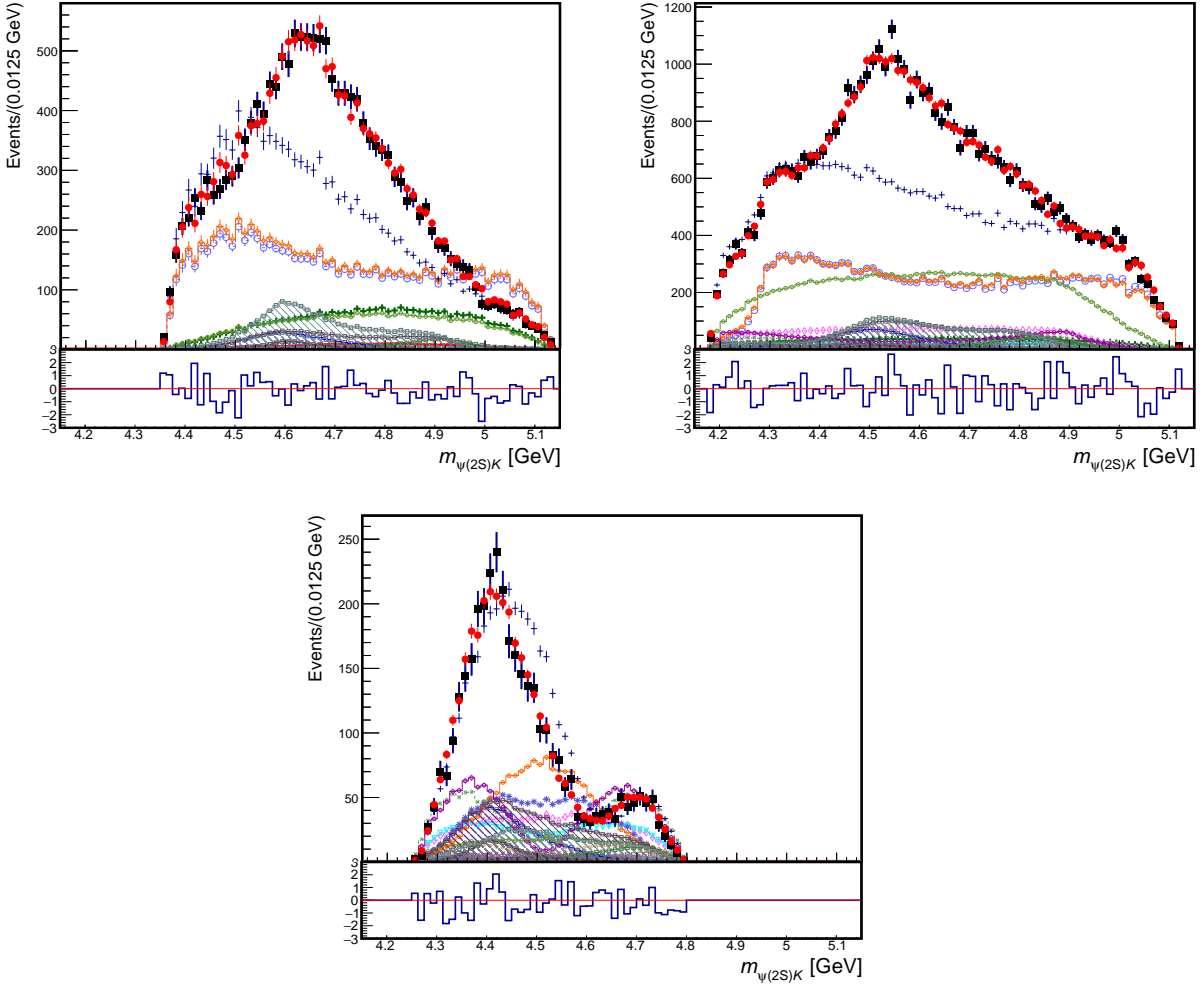


Figure 78: Projections of the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$, $Z(4430)^-$, and $Z(4800)^-$ as a $J^P = 1^+$ K-matrix, $Z(4470)^-$ as a $J^P = 2^-$ resonance, and $Z_{cs}(4480)^-$ as a $J^P 1^+$ $\psi(2S)K$ resonance. onto the $m_{\psi(2S)K}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $K_2^*(1430)$ until the end of $\psi(2S)$ phase space (bottom).

13.3 Models with two $Z_K^+ \rightarrow \psi(2S)K^+$

Fits with two Z_K^+ s were performed on $B^0 \rightarrow \psi(2S)\pi^-K^+$ data, and the fit quality metrics for these are in Table 25.

Adding a $1^+ Z_K^+$ is the most significant in terms of the likelihood (6.1σ), and it also improves the 2-D and 4-D χ^2 s. The plot of the pulls for each bin in the Dalitz plane for the 2-D χ^2 is in Figure 84, and the distribution of the 2-D and 4-D pulls fit to a Gaussian are in Figure 85, which feature a mean close to 0 (more so for the 2-D distribution) and a width close to 1. We allow the mass and width of the Z_K s to float, and the results are found in Table 28. The new Z_K resonance, which we label $Z_K(4520)$, has mass at 4.521 ± 0.007 GeV and a width of 0.056 ± 0.012 GeV. The masses and widths of the Zs ($\psi(2S)\pi$ states) do not change much except for Z(4430) which returns from the 1 Z_K fit to being a bit more broad, and their values can be found in Table 26. The masses and widths of the K^* s that are floating can be found in Table 27, and the fit fractions for all resonances and partial waves are found in Table 29.

Table 25: Goodness-of-fit metrics for adding two Z_K s with the given J^P to the $\psi(2S)$ channel.

Variable	1 Z_K Only	$0^- Z_K$	$1^- Z_K$	$1^+ Z_K$	$2^- Z_K$	$2^+ Z_K$
nPar	86	90	90	92	92	90
$\log L$	-102581	-102597	-102592	-102615	-102592	-102588
$\Delta(-2 \log L)$	-	32	21	67	21	14
2-D χ^2 (1024 bins)	1072	1060	1073	1057	1068	1067
2-D $\Delta\chi^2$	-	13	-1	15	5	5
4-D χ^2 (4096 bins)	4377	4365	4378	4365	4371	4376
4-D $\Delta\chi^2$	-	13	-1	12	6	1

Table 26: Masses and widths of the Zs in the 2 Z_K model for the $\psi(2S)$ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
Z(4430)	1^+	4.526 ± 0.012	0.213 ± 0.012
Z(4200)	1^+	4.113 ± 0.009	0.169 ± 0.017
Z(4470)	2^-	4.471 ± 0.006	0.106 ± 0.013
Z(4800)	1^+	4.818 ± 0.013	0.141 ± 0.044

Table 27: Masses and widths of the K^* s in the 2 Z_K model for the $\psi(2S)$ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
$K^*(892)$	1^-	0.889 ± 0.0001	0.045 ± 0.0003
$K_0^*(800)$	0^+	0.799 ± 0.042	0.319 ± 0.019
$K_2^*(1430)$	2^+	1.415 ± 0.002	0.109 ± 0.003

Table 28: Masses and widths of the Z_K s in the 2 Z_K model for the $\psi(2S)$ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
$Z_K(4480)$	1^-	4.487 ± 0.013	0.155 ± 0.025
$Z_K(4520)$	1^+	4.521 ± 0.007	0.056 ± 0.012

The overall mass projections for $K\pi$ (Figure 79) and $\psi(2S)\pi$ (Figure 80) again do not change much after adding a second Z_K^+ resonance to the 1 Z_K^+ model. The $K\pi$ slices for the $\psi(2S)\pi$ mass projections do not change much, but there is more interference required from the total K^* wave that is necessary to get a good fit in the $K\pi$ mass slice above $K^*(1430)$ for $m_{\psi\pi}$ in Figure 81. The overall mass projection for $\psi(2S)K$ in Figure 82 does not change much when adding a second Z_K , while the $K\pi$ slices for the $\psi(2S)K$ mass projections slightly improve near the new pole, as seen in the slice below $K^*(892)$, and again more interference is seen in the slice above $K^*(1430)$ from the total K^* wave (Figure 83).

As it will become clear from the next subsection, the model presented here becomes our nominal pick for the $\psi(2S)$ channel, which we later use in the comparison to the J/ψ model (Sec. 16). We include additional plots showing the fit quality on the Dalitz plane and in all four fitted dimensions of the data (Figs. 84-85). We also show comparison between this model and the data on various moments of the K^{*0} helicity angle, which is another way to test fit quality on the Dalitz plane (see Appendix B).

Table 29: All resonance fit fractions in the 2 Z_K fit for the $\psi(2S)$ channel.

Contribution	Fit fraction
K^* Total S-wave	$18.4 \pm 0.9\%$
$K_0^*(800)$	$4.7 \pm 1.6\%$
$K_0^*(1430)$	$6.9 \pm 0.5\%$
$K_0^*(1950)$	$1.0 \pm 0.4\%$
K^* Total P-wave	$58.6 \pm 1.1\%$
$K^*(892)$	$55.2 \pm 1.1\%$
$K_1^*(1410)$	$2.8 \pm 0.3\%$
$K_1^*(1680)$	$1.5 \pm 0.2\%$
K^* Total D-wave	$7.9 \pm 0.3\%$
$K_2^*(1430)$	$7.1 \pm 0.3\%$
$K_2^*(1980)$	$0.9 \pm 0.2\%$
$K_3^*(1780)$	$0.6 \pm 0.1\%$
$K_4^*(2045)$	$0.2 \pm 0.05\%$
$K_5^*(2380)$	$0.1 \pm 0.04\%$
Z Total 1^+ wave	$5.3 \pm 0.1\%$
Z(4430)	$2.5 \pm 0.05\%$
Z(4200)	$1.3 \pm 0.03\%$
Z(4800)	$0.4 \pm 0.01\%$
Z Total 2^- wave	$0.4 \pm 0.01\%$
$Z_2(4470)$	$0.4 \pm 0.01\%$
Z_K Total 1^- wave	$0.4 \pm 0.01\%$
$Z_K(4480)$	$0.4 \pm 0.01\%$
Z_K Total 1^+ wave	$0.1 \pm 0.002\%$
$Z_K(4520)$	$0.1 \pm 0.002\%$

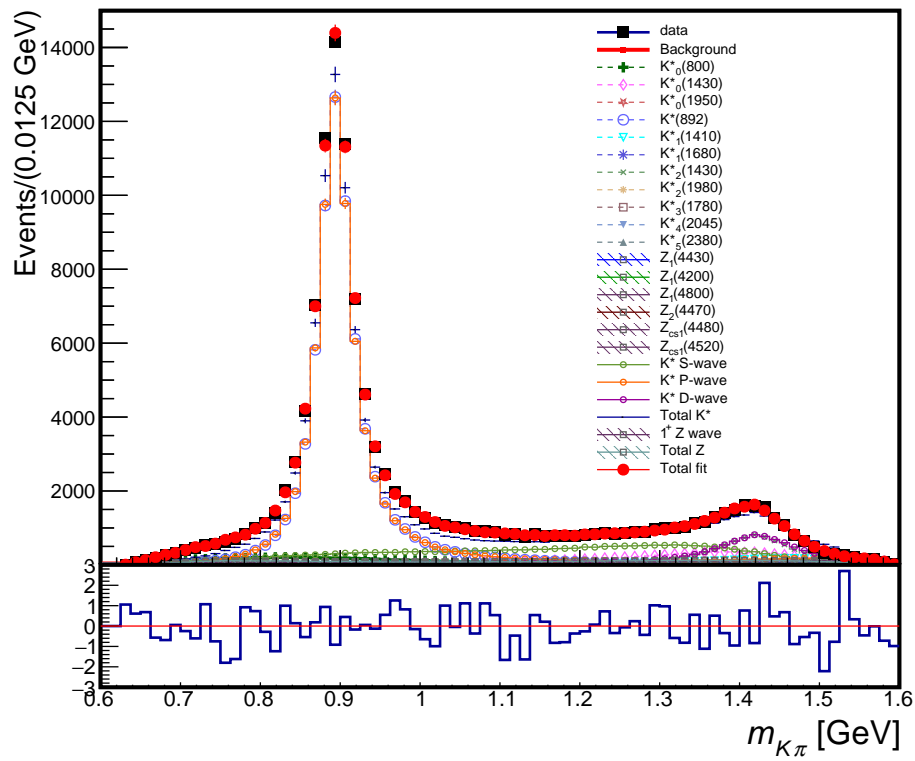


Figure 79: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^- K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+$ Z_K into the 1 Z_K model.

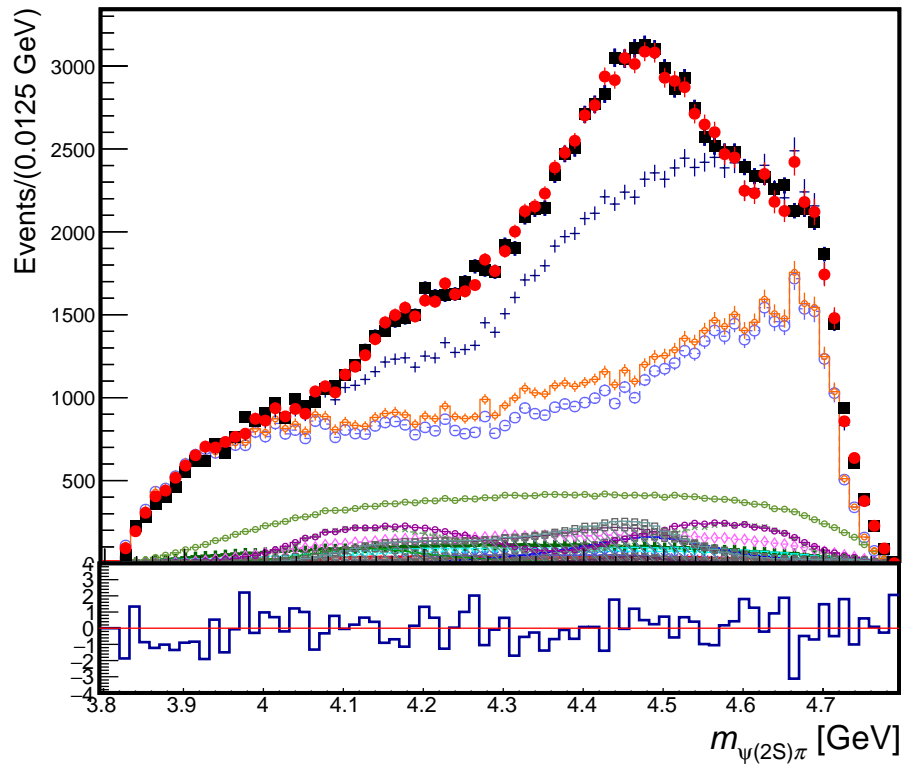


Figure 80: The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^- K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+$ Z_K into the 1 Z_K model.

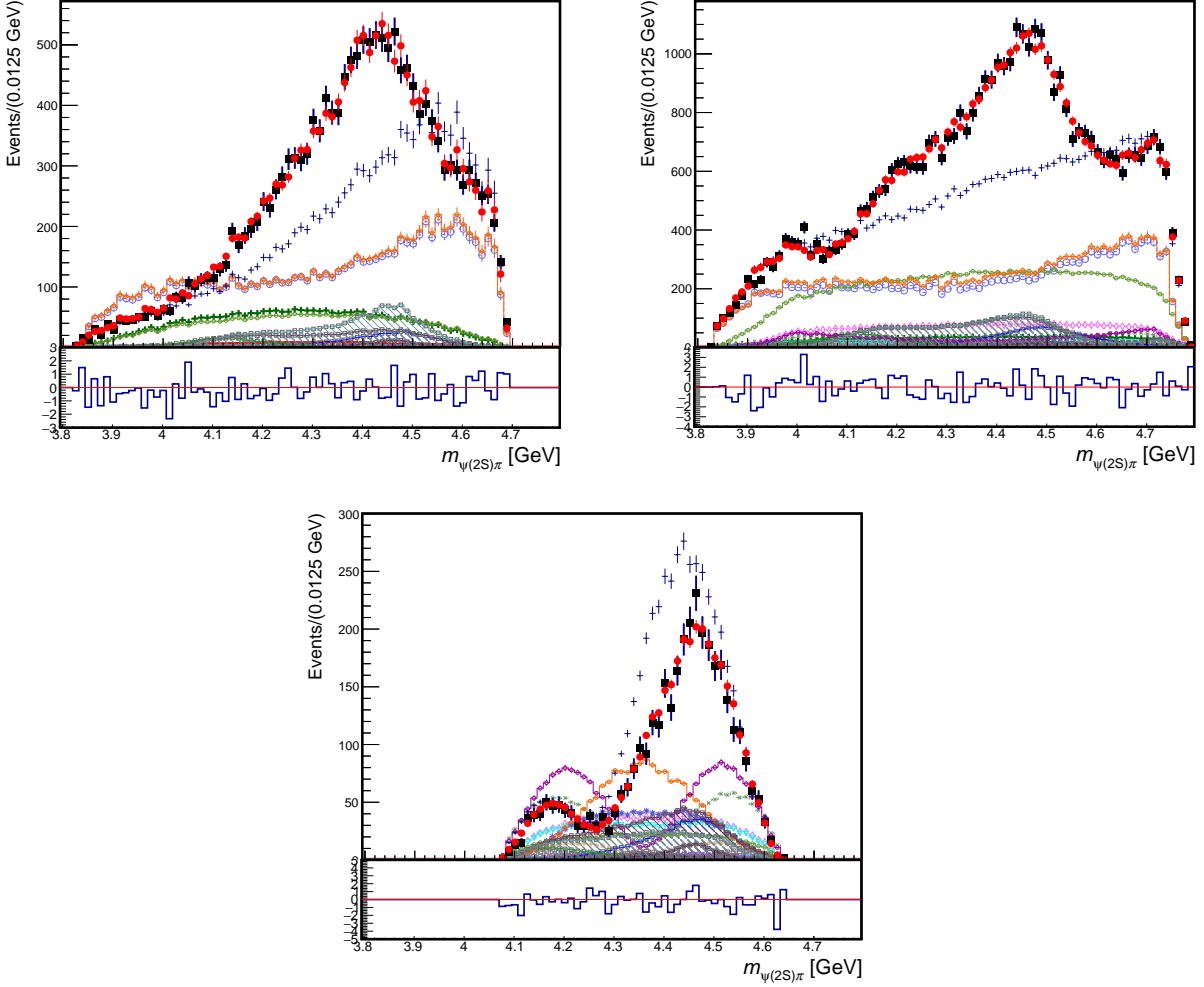


Figure 81: Projections of the $B^0 \rightarrow \psi(2S)\pi^- K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$, $Z(4430)^-$, and $Z(4800)^-$ as a $J^P = 1^+$ K-matrix, $Z(4470)^-$ as a $J^P = 2^-$ resonance, $Z_{cs}(4480)^-$ as a $J^P = 1^- \psi(2S)K$ resonance, and $Z_{cs}(4520)^-$ as a $J^P = 1^+ \psi(2S)K$ resonance onto the $m_{\psi(2S)\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $K_2^*(1430)$ until the end of $\psi(2S)$ phase space (bottom).

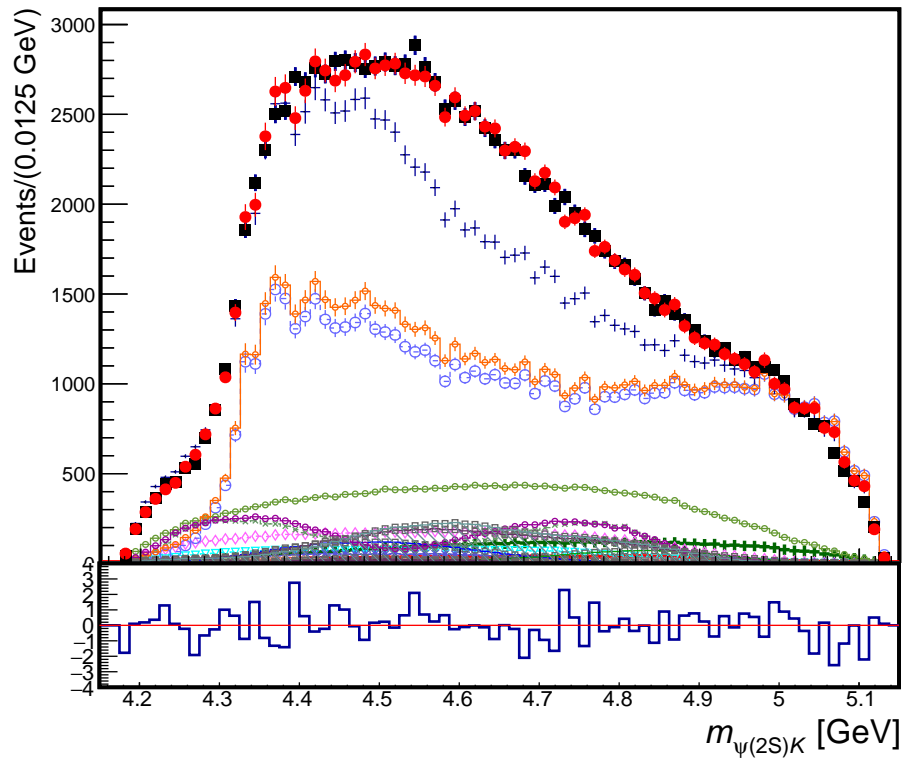


Figure 82: The distribution of $m_{\psi(2S)K}$ for the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+$ Z_K into the 1 Z_K model.

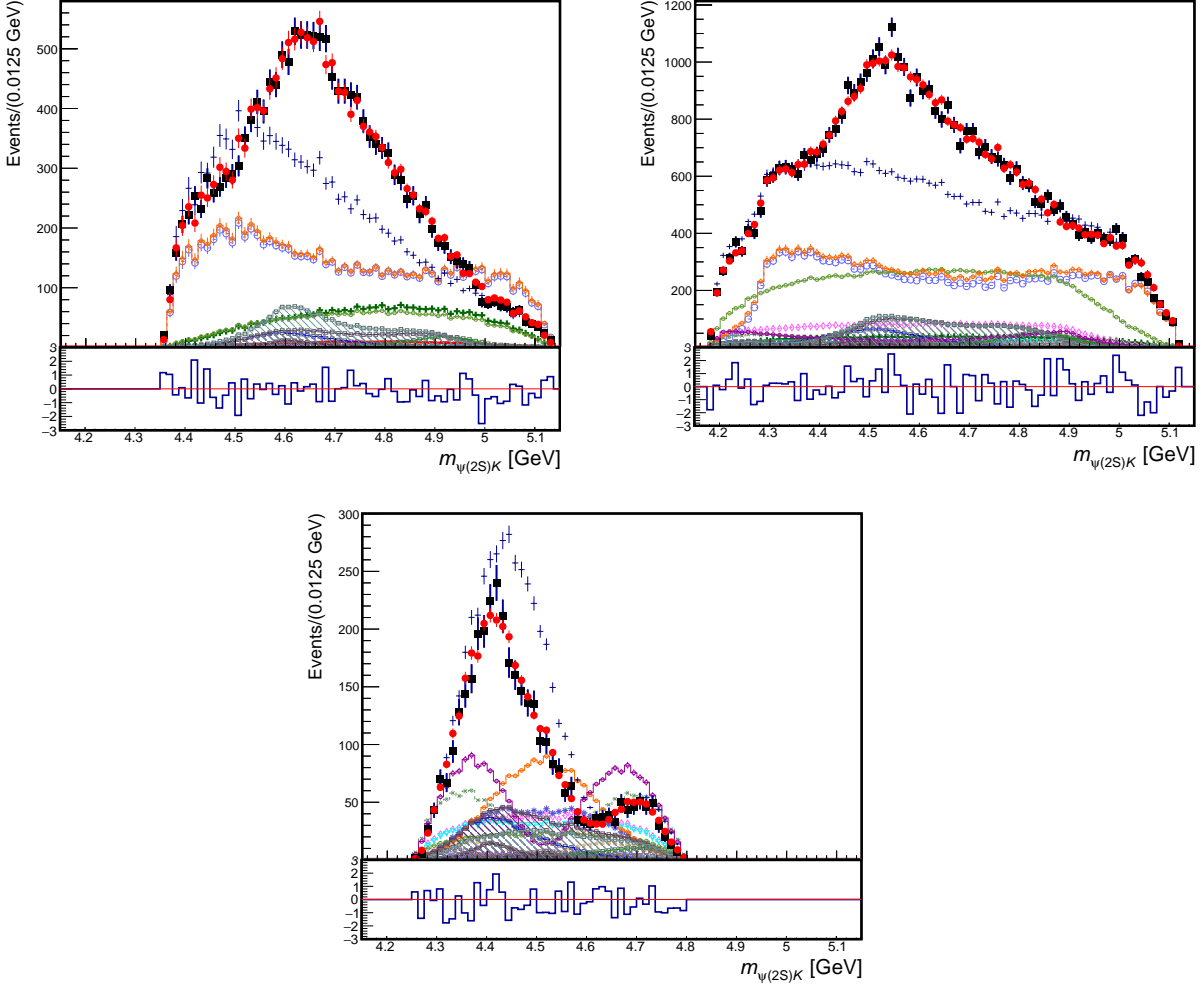


Figure 83: Projections of the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$, $Z(4430)^-$, and $Z(4800)^-$ as a $J^P = 1^+$ K-matrix, $Z(4470)^-$ as a $J^P = 2^-$ resonance, $Z_{cs}(4480)^-$ as a $J^P = 1^- \psi(2S)K$ resonance, and $Z_{cs}(4520)^-$ as a $J^P = 1^+ \psi(2S)K$ resonance onto the $m_{\psi(2S)K}$ axis in different slices of $m_{K\pi^-}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $K_2^*(1430)$ until the end of $\psi(2S)$ phase space (bottom).

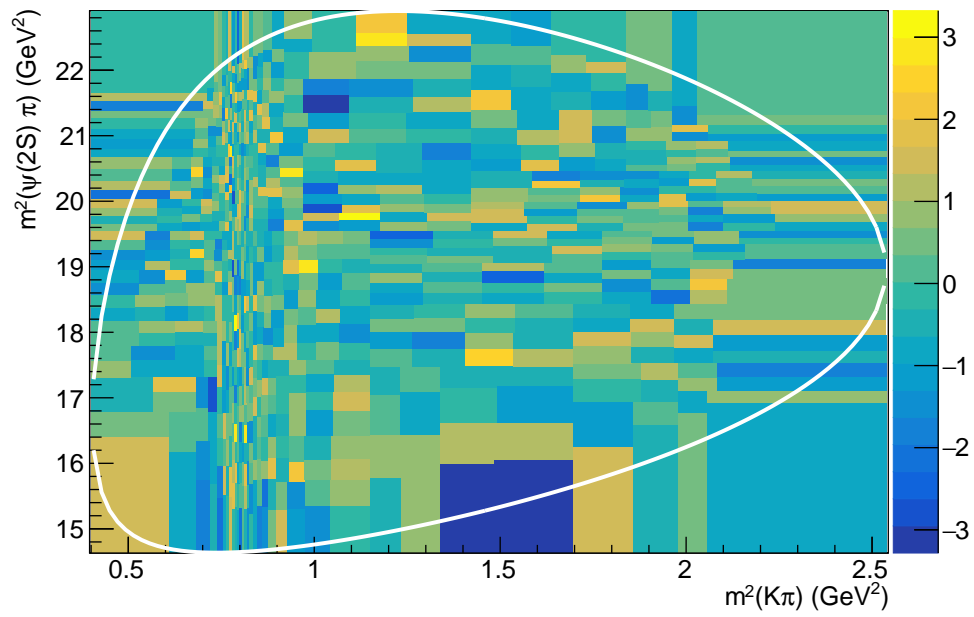


Figure 84: Dalitz plot with fit pulls in each bin used to calculate 2-D χ^2 . White line draws phase space boundary for the $B^0 \rightarrow \psi(2S)\pi^-K^+$ decay.

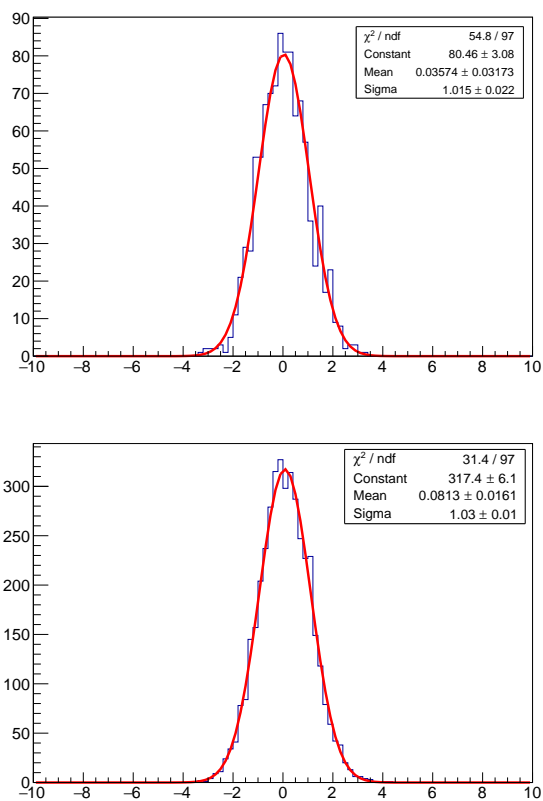


Figure 85: Distributions of the pulls from the 2-D and 4-D pulls from the multidimensional χ^2 s for the default $B^0 \rightarrow \psi(2S)\pi^-K^+$ model. The distribution is fit with a Gaussian (red line).

13.4 Models with three $Z_K^+ \rightarrow \psi(2S)K^+$

Fits with a third Z_K^+ state were performed on $B^0 \rightarrow \psi(2S)\pi^-K^+$ data, and the fit quality metrics for these are in Table 30. These fits do not produce significant log-likelihood improvements, thus a third Z_K^+ is not introduced into the default model. As such, the most reliable Z_K^+ resonance model for the $\psi(2S)$ channel is the one in the 3 Z_K^+ fit model.

Table 30: Goodness-of-fit metrics for adding a third Z_K with the given J^P to the $\psi(2S)$ channel.

Variable	2 Z_K only	0 ⁻ Z_K	1 ⁻ Z_K	1 ⁺ Z_K	2 ⁻ Z_K	2 ⁺ Z_K
nPar	92	96	96	98	98	96
log L	-102615	-102624	-102621	-102632	-102631	-102625
$\Delta(-2 \log L)$	-	19	14	35	33	20
2-D χ^2 (1024 bins)	1057	1055	1058	1049	1050	1060
2-D $\Delta\chi^2$	-	2	-1	8	7	-3
4-D χ^2 (4096 bins)	4365	4364	4365	4358	4361	4369
4-D $\Delta\chi^2$	-	1	0	7	4	-4

14 Resonant Z models with J/ψ channel

This section will follow the same procedure as the last but instead deals with the $B^0 \rightarrow J/\psi\pi^-K^+$ dataset, which has much higher statistics and enlarged phase-space for conventional and exotic resonances at the higher and lower masses, respectively. The modeling will begin again with one $Z^- \rightarrow J/\psi\pi^-$ resonance of various J^P added to the K^* -only model, and then sequentially adding more Z resonances.

14.1 Models with one $Z^- \rightarrow J/\psi\pi^-$

Fits were performed each adding a Z with various J^P to the K^* -only model discussed in Section 10 for $B^0 \rightarrow J/\psi\pi^-K^+$. In Table 31, the goodness-of-fit metrics are listed for these fits, with the best being the one adding a $J^P = 1^+$ Z. The mass and width of the Z is allowed to float in the fit, which gives $m(Z) = 4.145 \pm 0.003$ GeV and $\Gamma(Z) = 0.214 \pm 0.004$ GeV. We label this pole Z(4200), because its parameters are not far from the mass and width of the state first claimed by Belle in the amplitude analysis of the $B \rightarrow J/\psi\pi K$ channel [16]. Our 1 Z model is favored over the K^* -only model by $> 37\sigma$ in terms of log-likelihood values. It also improves the 2-D and 4-D χ^2 s by very large numbers.

Table 31: Goodness-of-fit metrics for adding the first Z with the given J^P to the J/ψ channel.

Variable	K^* -only	$0^- Z$	$1^- Z$	$1^+ Z$	$2^- Z$	$2^+ Z$
nPar	63	67	67	69	69	67
$\log L$	-3048980	-3049165	-3049314	-3051810	-3051076	-3049159
$\Delta(-2 \log L)$	-	370	668	5660	4192	358
2-D χ^2 (1024 bins)	4476	4401	4118	3206	3109	4325
2-D $\Delta\chi^2$	-	75	358	1270	1367	151
4-D χ^2 (4096 bins)	6893	6827	6622	5798	5880	6759
4-D $\Delta\chi^2$	-	66	271	1095	1013	134

The high significance of the Z(4200) is clear in the different mass projections. With the K^* -only model, there is a large disagreement at high $m_{K\pi}$ near the end of $K\pi$ phase space, but this disagreement is improved upon after adding the Z(4200) as seen in Figure 86. The

Table 32: Mass and width of the Z(4200) in the 1 Z model for the J/ψ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
Z(4200)	1^+	4.145 ± 0.003	0.214 ± 0.004

$J/\psi\pi$ distribution in Figure 87 improves tremendously near the pole of the new Z (4.1 GeV), resolving a very large disagreement present in K^* -only model.

The $m_{J/\psi\pi}$ projections in the different $m_{K\pi}$ slices also improve a lot near the new Z pole, as can be seen in Figure 88 Better agreement between the fit and data is best seen in the slice between $K^*(892)$ and $K^*(1430)$ and in the slice above $\psi(2S)$ phase space. Despite these improvements, there remain large disagreements in the mass distributions at high $J/\psi\pi$ mass (near 4.5 GeV) and also near 3.8 GeV either in the overall projection or in the $m_{K\pi}$ slices. In the subsequent subsections, additional $J/\psi\pi$ resonances are added to attempt to improve this 1 Z fit.

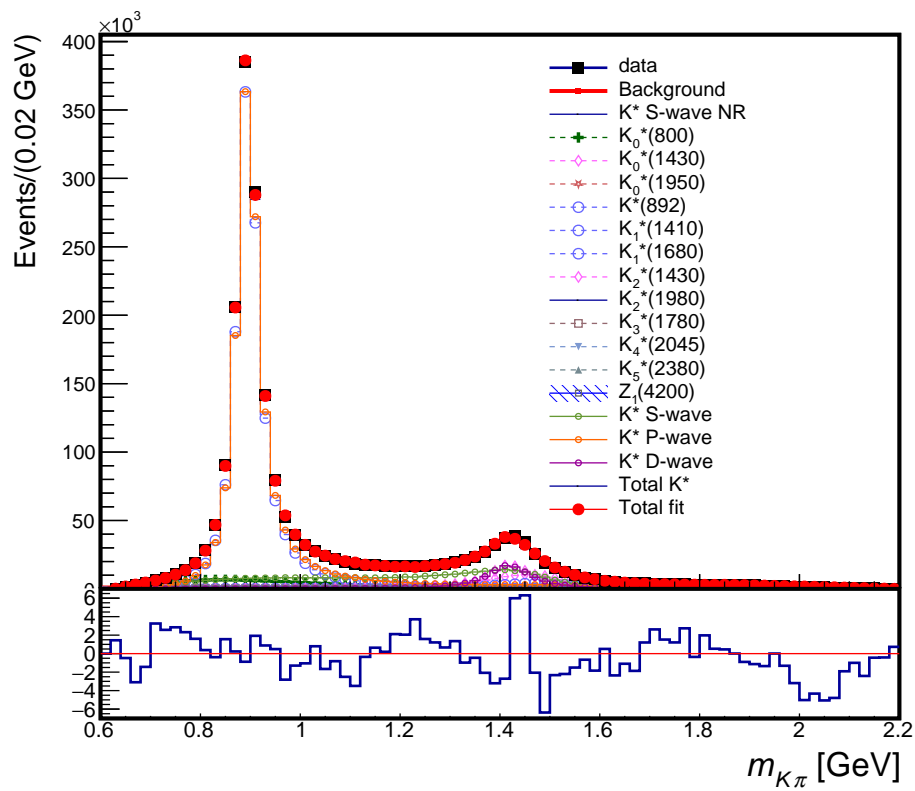


Figure 86: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black) and the fit (red) which adds a $J^P = 1^+ Z$.

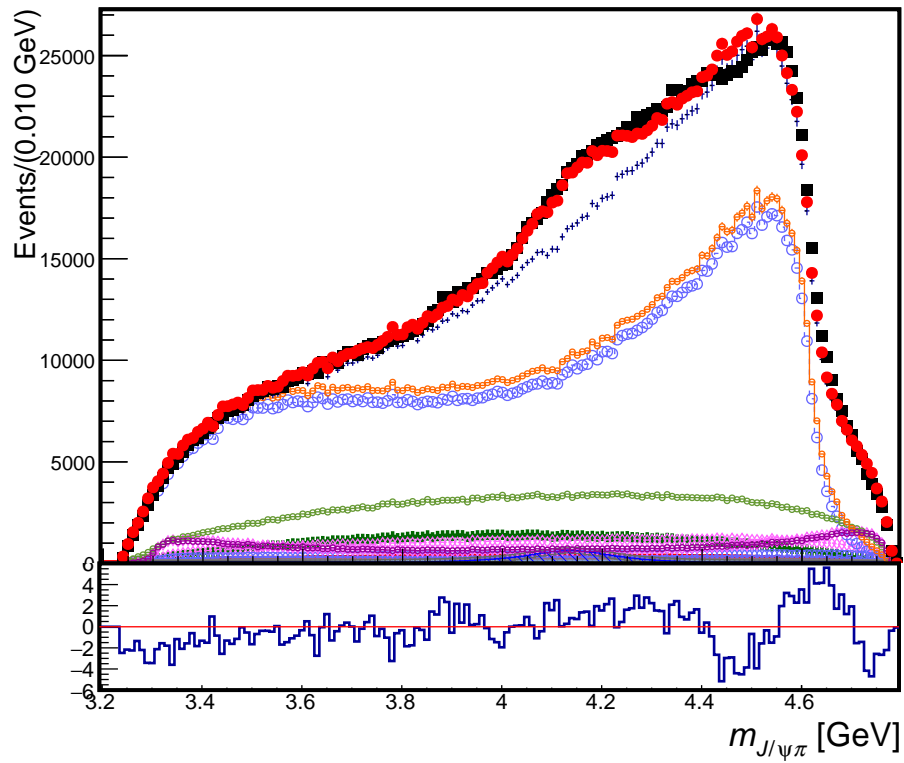


Figure 87: The distribution of $m_{J/\psi\pi}$ for the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black) and the fit (red) which adds a $J^P = 1^+ Z$.

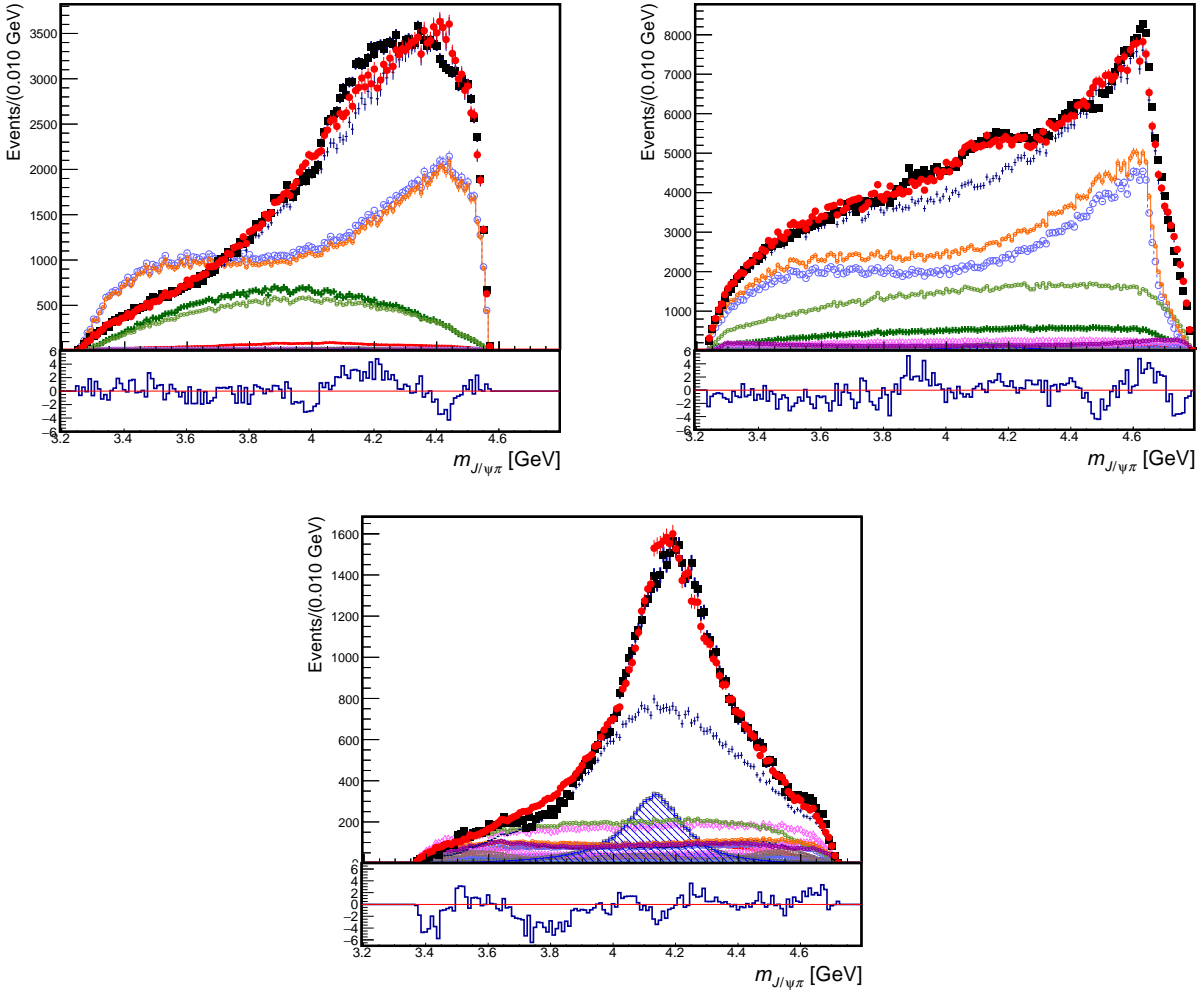


Figure 88: Projections of the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$ as a $J^P = 1^+$ resonance onto the $m_{J/\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom).

14.2 Models with two $Z^- \rightarrow J/\psi\pi^-$

Given the number of disagreements between the data and the 1 Z model fit, two Z fits were performed on $B^0 \rightarrow J/\psi\pi^- K^+$. The fit quality metrics for these are in Table 33, which shows including a very broad 0^- Z as the most significant in terms of the all the fit quality metrics. It has a significance of $> 37\sigma$ based on the log-likelihood difference. The masses and widths of both Zs are free to float in the fit, with their values are listed in Table 34. The new broad 0^- Z resonant pole has a mass at 4.201 ± 0.008 GeV and a width of 0.826 ± 0.015 GeV, and while the 1^+ Z(4200)'s mass is similar to the value from the 1 Z fit, it becomes significantly broader in the 2 Z fit.

Table 33: Goodness-of-fit metrics for adding the second Z with the given J^P to the J/ψ channel.

Variable	1 Z only	0^- Z	1^- Z	1^+ Z	2^- Z	2^+ Z
nPar	69	73	73	75	75	73
$\log L$	-3051810	-3053797	-3051912	-3053467	-3052510	-3051841
$\Delta(-2 \log L)$	-	3974	204	3314	1400	62
2-D χ^2 (1024 bins)	3206	2434	3092	2504	2836	3181
2-D $\Delta\chi^2$	-	772	113	701	369	25
4-D χ^2 (4096 bins)	5798	5181	5743	5181	5535	5771
4-D $\Delta\chi^2$	-	617	55	618	263	27

Table 34: Masses and widths of the Zs in the 2 Z model for the J/ψ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
Z(4200)	1^+	4.111 ± 0.003	0.306 ± 0.006
$Z_0(4200)$	0^-	4.201 ± 0.008	0.826 ± 0.015

In Figure 89, the overall mass projection for $K^+\pi^-$ improves very significantly, finding much better agreement at low mass (below $K^*(892)$), near the $K^*(1430)$ peak, and above the $K^*(1430)$ peak until the end of $K^+\pi^-$ phase space. The overall mass projection for $J/\psi\pi^-$ (Figure 90) does not change much from the 1 Z fit, still featuring the same disagreements above 4.2 GeV. However, the $J/\psi\pi^-$ mass distributions in the different $m_{K\pi}$ slices in Figure 91 improve a lot, as all slices shown exhibit many improvements throughout $J/\psi\pi^-$ phase space. Nonetheless large disagreements remain near 3.8 GeV and 4.5 GeV in the $m_{J/\psi\pi}$ spectrum.

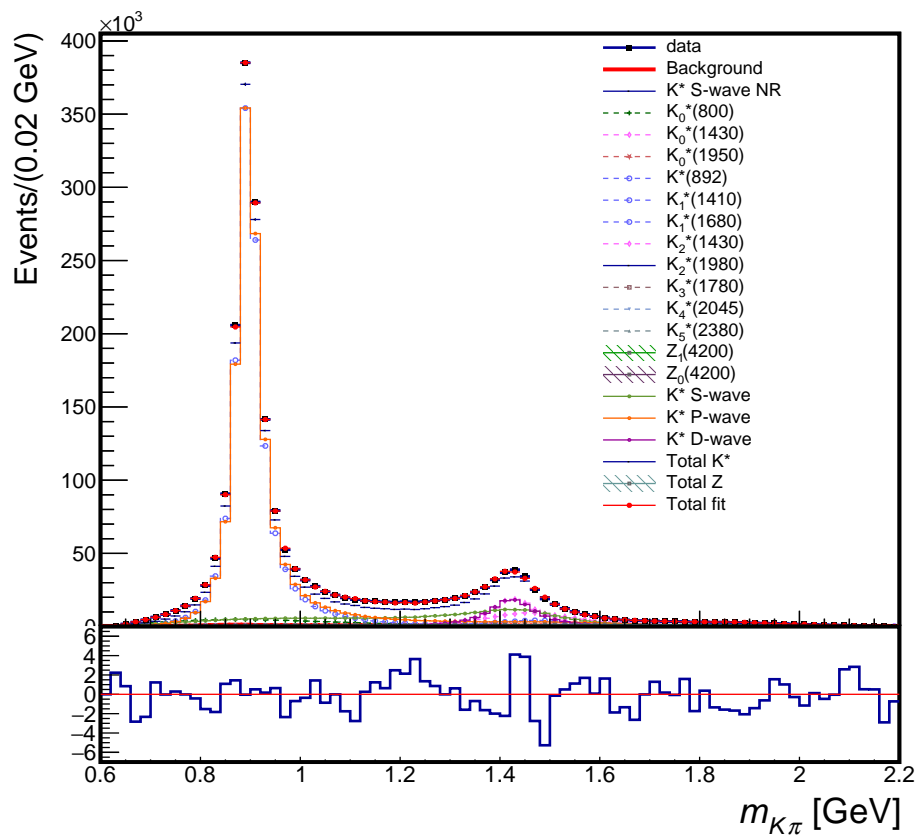


Figure 89: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black) and the fit (red) which adds a $J^P = 0^-$ Z on top of a model with a separate 1^+ Z resonance.

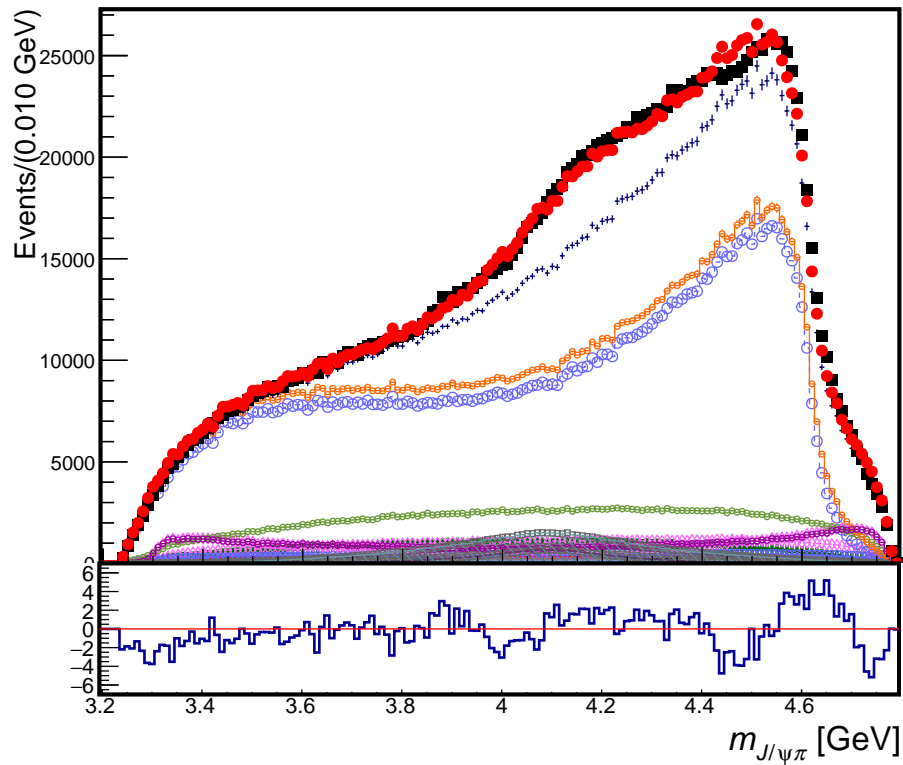


Figure 90: The distribution of $m_{J/\psi\pi}$ for the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black) and the fit (red) which adds a $J^P = 0^-$ Z on top of a model with a separate 1^+ Z resonance.

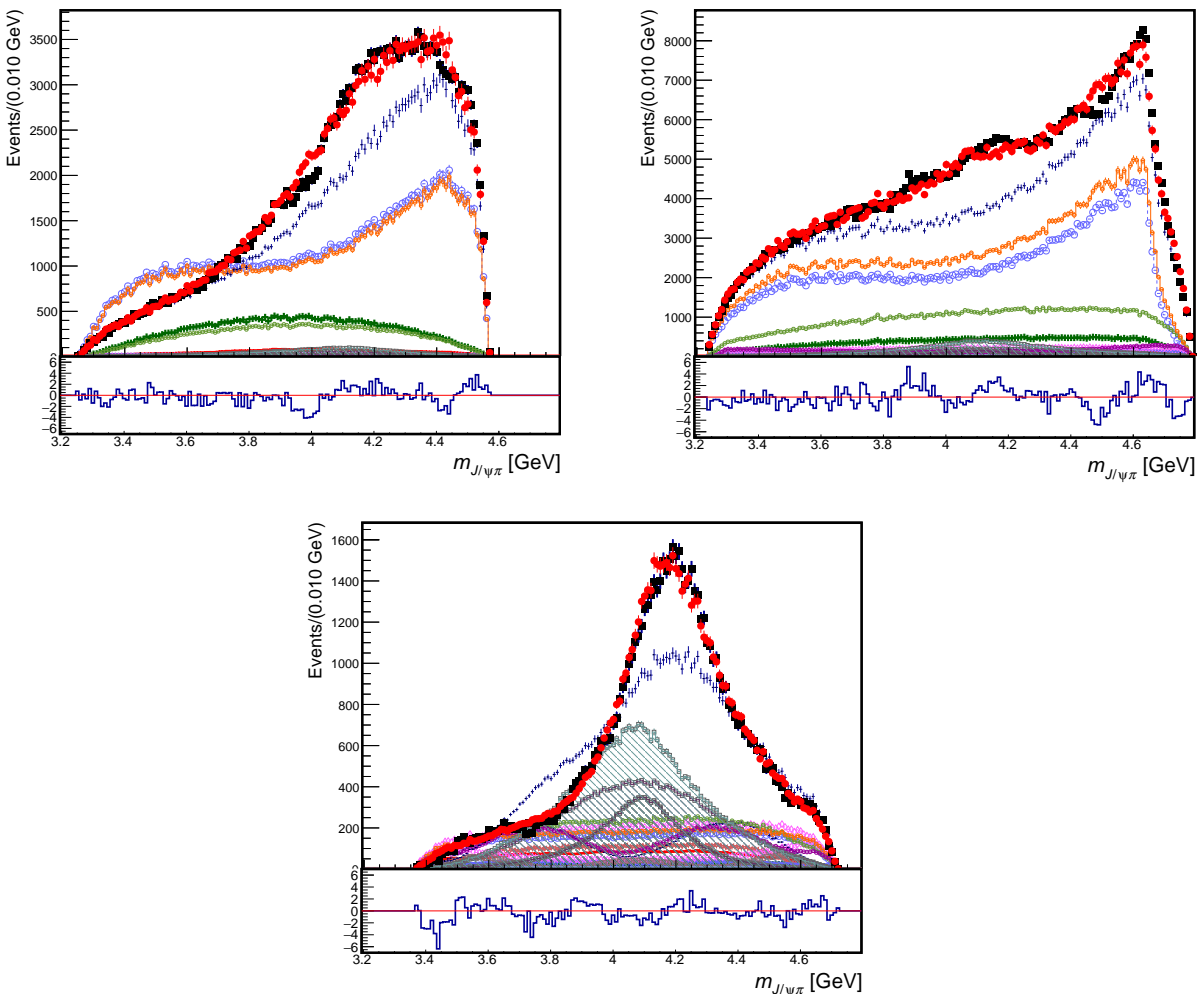


Figure 91: Projections of the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$ as a $J^P = 1^+$ resonance and a broad $Z(4200)^-$ as a $J^P = 0^-$ resonance onto the $m_{J/\psi \pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom).

14.3 Models with three $Z^- \rightarrow J/\psi\pi^-$

To improve on the 2 Z model, three Z fits were performed on $B^0 \rightarrow J/\psi\pi^-K^+$, and the fit quality metrics for these are in Table 35. Including a second 1^+ Z to the K-matrix with Z(4200) proves to be the most significant in terms of the all the fit quality metrics. This term is $> 37\sigma$ significant based on the log-likelihood difference. The masses and widths of the Zs are free to float in the fit, and their values are listed in Table 36. The second 1^+ Z resonance is labeled Z(4430), since its mass, 4.553 ± 0.003 GeV, and width, 0.165 ± 0.005 GeV, are in the range of values obtained for the $Z(4430)^- \rightarrow \psi(2S)\pi^-$ resonance. The masses and widths of the 1^+ Z(4200) and the broad 0^- $Z_0(4200)$ change very little between fits, with only the width of the very broad 0^- $Z_0(4200)$ getting even slightly larger.

The need for at least two 1^+ $Z^- \rightarrow J/\psi\pi^-$ resonances, one near 4.2 GeV, and the other near the $Z(4430)^-$ pole, to explain $B^0 \rightarrow J/\psi\pi^-K^+$ data was first claimed by Belle [16], albeit they fixed the $Z(4430)^-$ mass and width to their $\psi(2S)\pi^-$ values. Our data confirm the need for a higher mass 1^+ resonance without such constraint. However, they point also to even more urgent need to include a broad 0^- $Z^- \rightarrow J/\psi\pi^-$ state in the model.

Table 35: Goodness-of-fit metrics for adding the third Z with the given J^P to the J/ψ channel.

Variable	2 Z only	0^- Z	1^- Z	1^+ Z	2^- Z	2^+ Z
nPar	73	77	77	79	79	77
$\log L$	-3053797	-3054148	-3054035	-3055331	-3054170	-3054054
$\Delta(-2 \log L)$	-	702	476	3068	746	514
2-D χ^2 (1024 bins)	2434	2228	2278	1899	2222	2399
2-D $\Delta\chi^2$	-	206	156	535	212	35
4-D χ^2 (4096 bins)	5181	5018	5029	4628	4917	5105
4-D $\Delta\chi^2$	-	163	152	553	264	76

Table 36: Masses and widths of the Zs in the 3 Z model for the J/ψ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
Z(4200)	1^+	4.100 ± 0.003	0.352 ± 0.006
$Z_0(4200)$	0^-	4.265 ± 0.011	1.014 ± 0.020
Z(4430)	1^+	4.553 ± 0.003	0.165 ± 0.005

In Figure 92, the overall mass projection for $K^+\pi^-$ does not change much. The overall mass projection for $J/\psi\pi^-$ improves greatly near the newly added $Z(4430)$ pole, as can be easily seen in Fig. 93. There is also great improvement in the $J/\psi\pi^-$ mass distributions in the $m_{K\pi}$ slices as seen in Figure 94. Both the slice between $K^*(892)$ and $K^*(1430)$, and the slice above the phase space limit reachable in the $\psi(2S)$ channel, show much better agreement in the pull plots near the new pole.

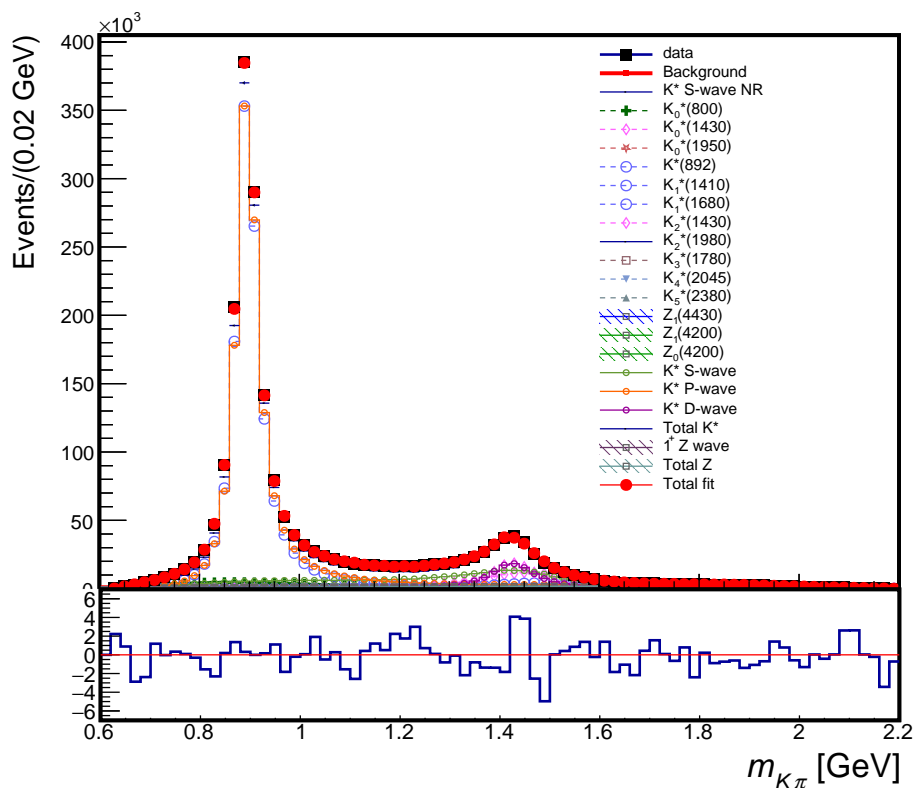


Figure 92: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds an additional $J^P = 1^+$ Z into a K-matrix with one other 1^+ Z and a separate 0^- Z resonance.

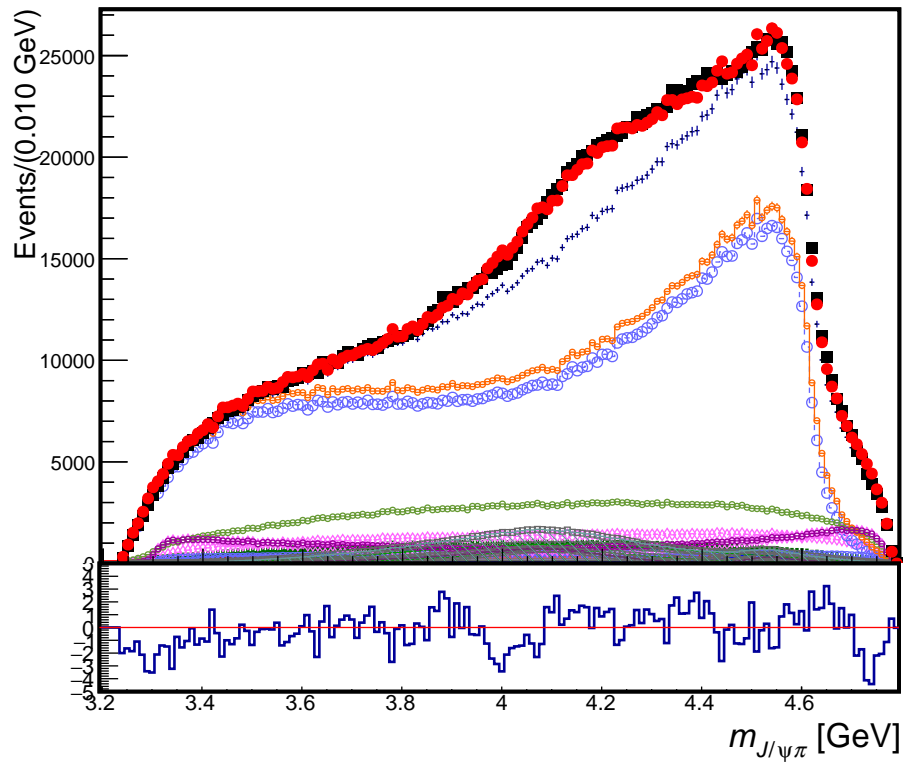


Figure 93: The distribution of $m_{J/\psi\pi}$ for the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black) and the fit (red) which adds an additional $J^P = 1^+$ Z into a K-matrix with one other 1^+ Z and a separate 0^- Z resonance.

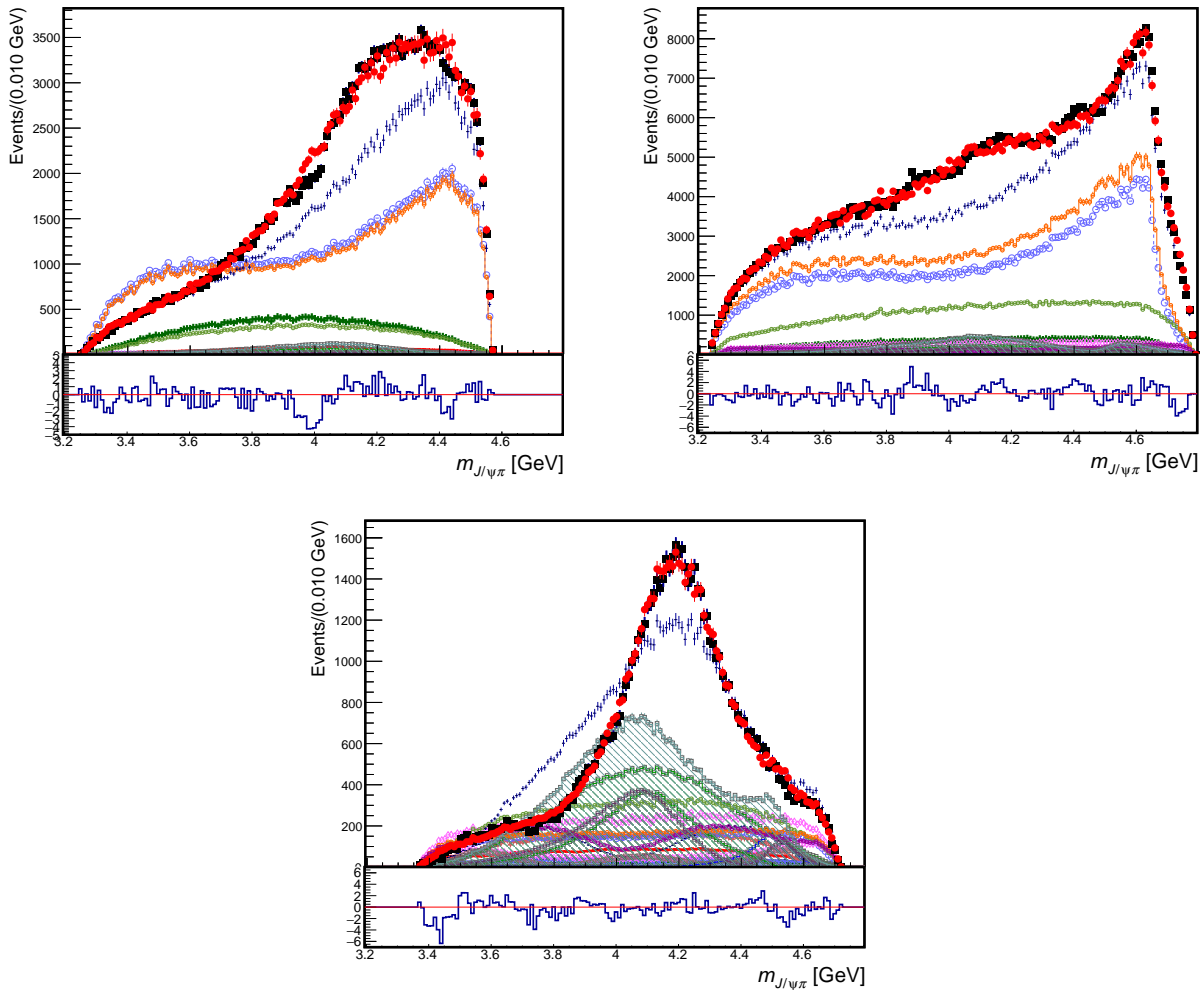


Figure 94: Projections of the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$ and $Z(4430)^-$ as a $J^P = 1^+$ K-matrix and a broad $Z(4200)^-$ as a $J^P = 0^-$ resonance onto the $m_{J/\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom).

14.4 Models with four $Z^- \rightarrow J/\psi\pi^-$

Four Z fits were performed on $B^0 \rightarrow J/\psi\pi^- K^+$, and the fit quality metrics for these are in Table 37. Adding a 2^- Z is the most significant in terms of log-likelihood and improves the multidimensional χ^2 s. From the likelihood change, such contribution is 22.8σ significant. We allow all masses and widths of these Zs to float, and their results are found in Table 38. The new 2^- Z resonance has a pole mass at 4.730 ± 0.015 GeV and a width of 0.610 ± 0.068 GeV. We label it as Z(4700).

With the addition of Z(4700), the masses and widths of the other Zs change very little overall. The overall mass projection for $K\pi$ (Figure 95) remains the same, but the mass projections for $J/\psi\pi$ improve near the new pole as seen in Figures 96 and 97, with disagreements as seen in the pulls becoming smaller.

Table 37: Goodness-of-fit metrics for adding the fourth Z with the given J^P to the J/ψ channel.

Variable	3 Z only	0^- Z	1^- Z	1^+ Z	2^- Z	2^+ Z
nPar	79	83	83	85	85	83
$\log L$	-3055331	-3055528	-3055487	-3055511	-3055605	-3055522
$\Delta(-2 \log L)$	-	394	312	360	548	382
2-D χ^2 (1024 bins)	1899	1720	1919	1691	1843	1918
2-D $\Delta\chi^2$	-	179	-21	208	56	-19
4-D χ^2 (4096 bins)	4628	4493	4615	4463	4543	4591
4-D $\Delta\chi^2$	-	135	13	164	85	36

Table 38: Masses and widths of the Zs in the 4 Z model for the J/ψ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
Z(4200)	1^+	4.135 ± 0.003	0.352 ± 0.006
Z(4200)	0^-	4.266 ± 0.007	0.951 ± 0.017
Z(4430)	1^+	4.579 ± 0.002	0.172 ± 0.005
Z(4700)	2^-	4.730 ± 0.015	0.610 ± 0.068

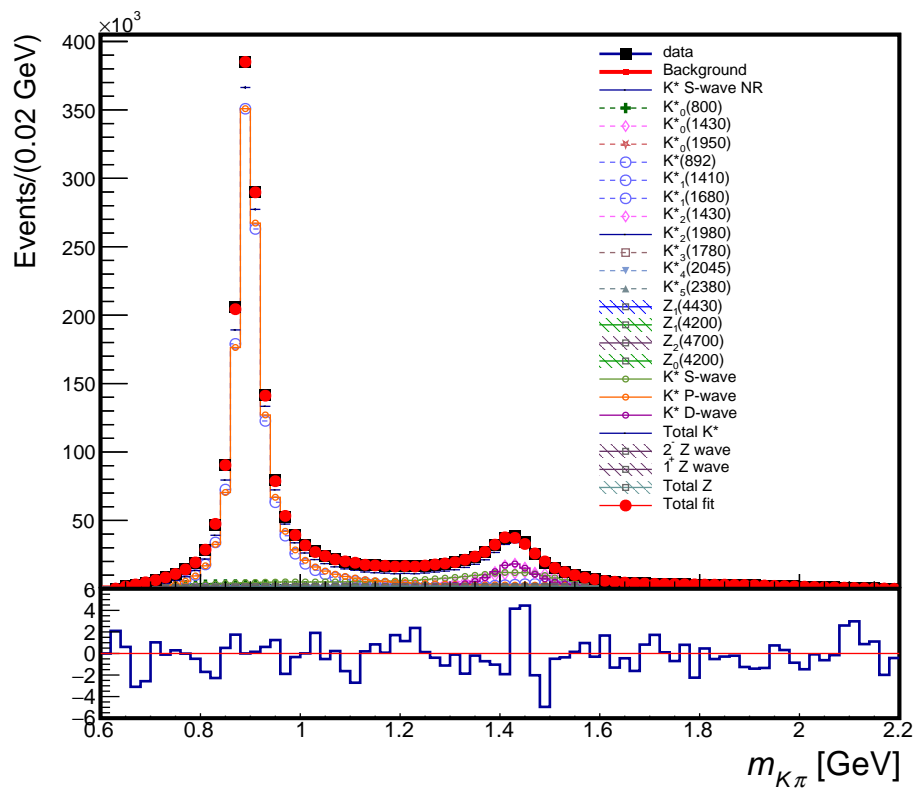


Figure 95: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black) and the fit (red) which adds a new $J^P = 2^- Z$ into the model with a K-matrix of two 1^+ Zs and a separate $0^- Z$ resonance.

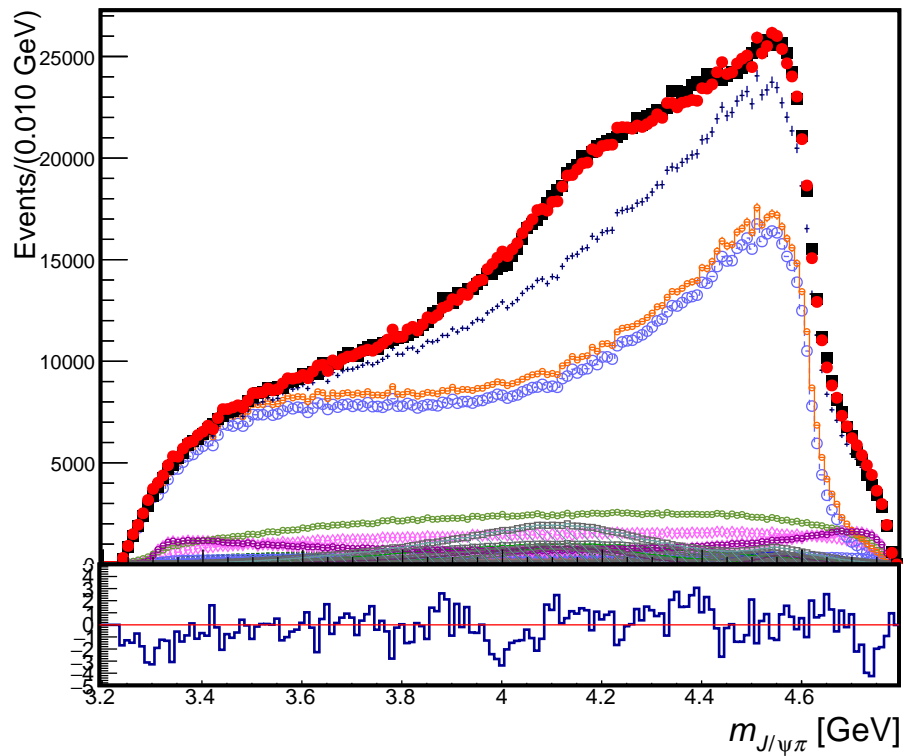


Figure 96: The distribution of $m_{J/\psi\pi}$ for the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black) and the fit (red) which adds a new $J^P = 2^- Z$ into the model with a K-matrix of two 1^+ Zs and a separate $0^- Z$ resonance.

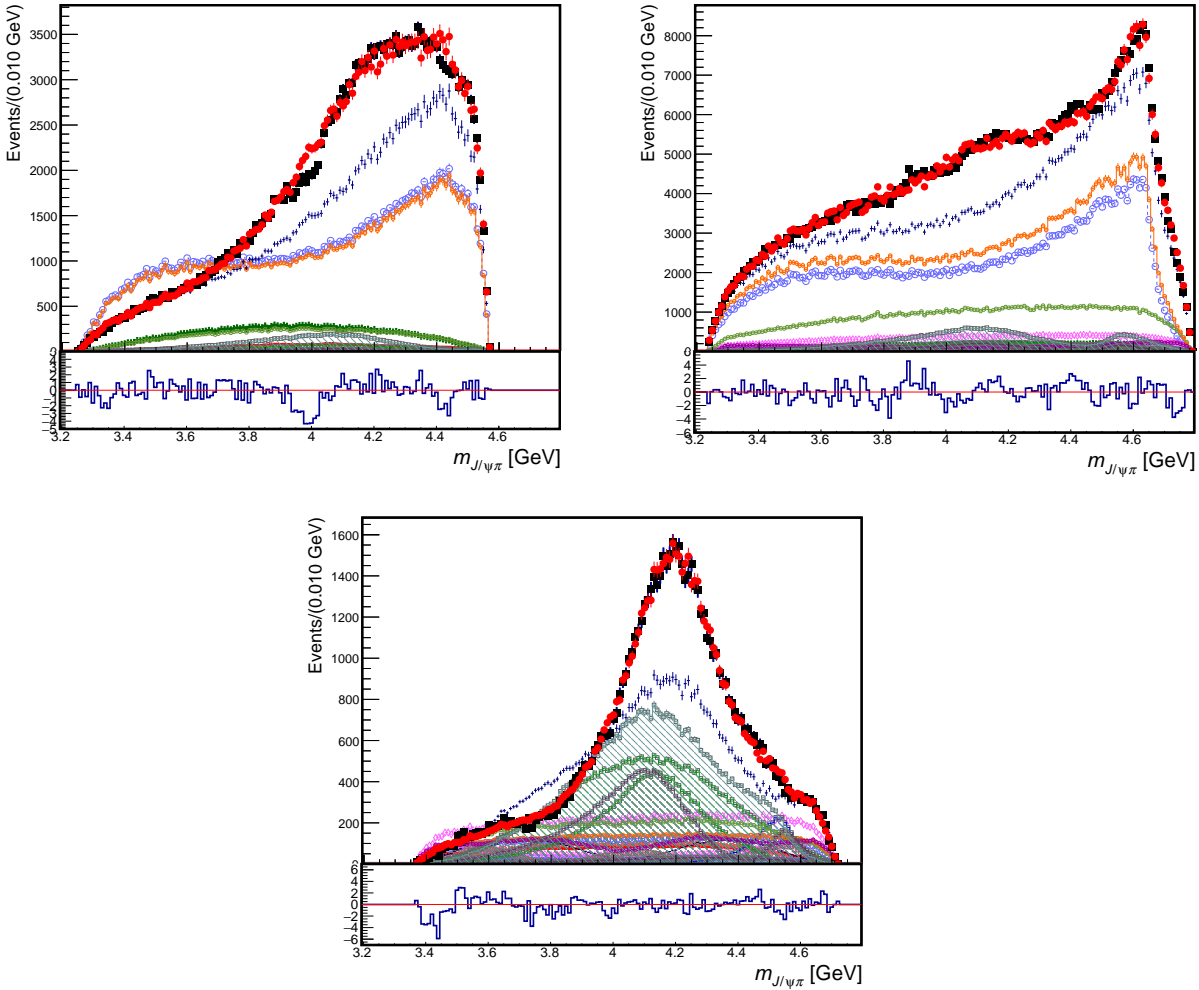


Figure 97: Projections of the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$ and $Z(4430)^-$ as a $J^P = 1^+$ K-matrix, a broad $Z(4200)^-$ as a $J^P = 0^-$ resonance, and a broad $Z(4700)^-$ as a $J^P = 2^-$ resonance onto the $m_{J/\psi \pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom).

14.5 Models with five $Z^- \rightarrow J/\psi\pi^-$

Fits with a fifth Z were performed on $B^0 \rightarrow J/\psi\pi^- K^+$, and the fit quality metrics for these are in Table 39.

Adding a 0^- Z to the K-matrix with Z(4200) proves to be the most significant in terms of the likelihood, and it improves the χ^2 s as well. We allow all masses and widths of these Zs to float, and their results are found in Table 40. The new Z resonance, which we label Z(3900), has mass at 3.905 ± 0.004 GeV and a width of 0.169 ± 0.006 GeV, and the masses and widths of the other Zs remain largely the same, except for the $J^P = 0^-$ Z(4200) which becomes less broad, but it still remains largely broad.

Table 39: Goodness-of-fit metrics for adding the fifth Z with the given J^P to the J/ψ channel.

Variable	4 Z only	0^- Z	1^- Z	1^+ Z	2^- Z	2^+ Z
nPar	85	89	89	91	91	89
$\log L$	-3055605	-3055806	-3055778	-3055765	-3055711	-3055747
$\Delta(-2 \log L)$	-	402	346	320	212	284
2-D χ^2 (1024 bins)	1843	1671	1827	1663	1751	1840
2-D $\Delta\chi^2$	-	172	16	180	92	3
4-D χ^2 (4096 bins)	4543	4416	4533	4400	4450	4528
4-D $\Delta\chi^2$	-	127	10	143	93	15

Table 40: Masses and widths of the Zs in the 5 Z model for the J/ψ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
Z(4200)	1^+	4.128 ± 0.002	0.329 ± 0.005
Z(4200)	0^-	4.298 ± 0.006	0.677 ± 0.016
Z(4430)	1^+	4.575 ± 0.002	0.161 ± 0.004
Z(4700)	2^-	4.745 ± 0.019	0.635 ± 0.101
Z(3900)	0^-	3.905 ± 0.004	0.169 ± 0.006

The overall mass projection for $K\pi$ (Figure 98) remains the same, while the overall mass projection for $J/\psi\pi$ (Figure 99) improves in the middle of $J/\psi\pi$ phase space, near the pole mass of the new Z. There is also improvement in the $J/\psi\pi$ mass distribution in the $m_{K\pi}$ slices, notably below $K^*(892)$ and between $K^*(892)$ and $K_2^*(1430)$ in Figure 100 near the new pole as well.

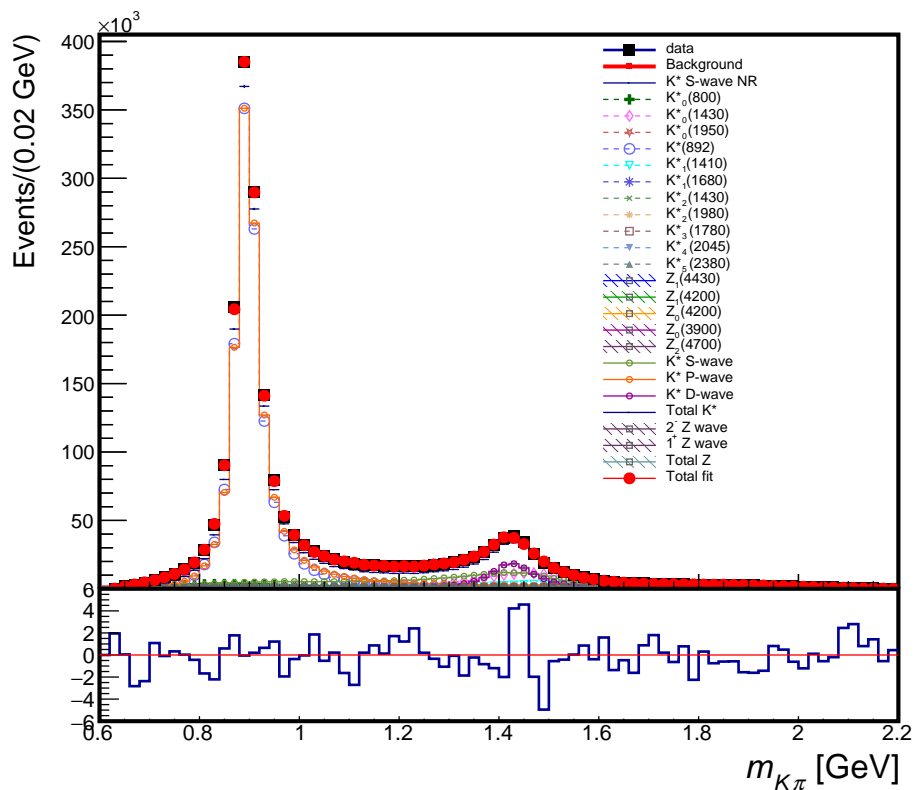


Figure 98: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black) and the fit (red) which adds an additional $J^P = 0^-$ Z into a K-matrix with one other 0^- Z, two 1^+ Zs, and a separate 2^- Z resonance.

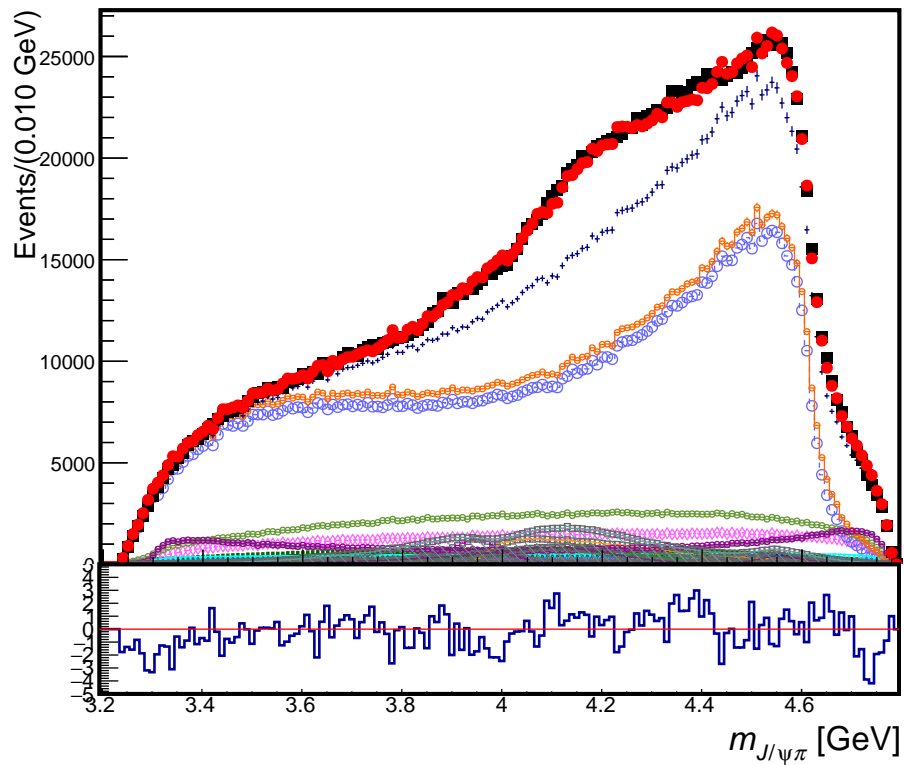


Figure 99: The distribution of $m_{J/\psi\pi}$ for the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black) and the fit (red) which adds an additional $J^P = 0^-$ Z into a K-matrix with one other 0^- Z, two 1^+ Zs, and a separate 2^- Z resonance.

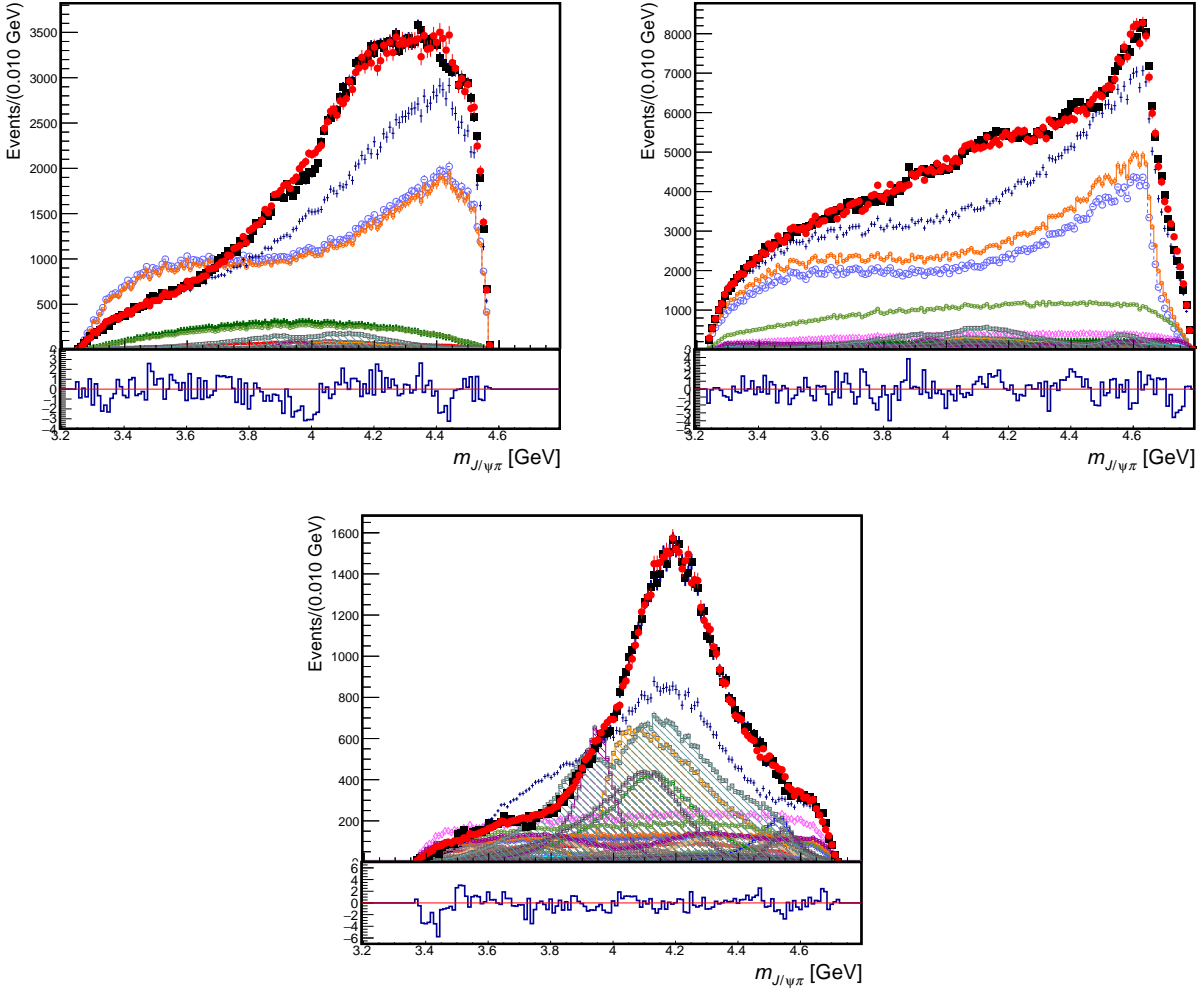


Figure 100: Projections of the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$ and $Z(4430)^-$ as a $J^P = 1^+$ K-matrix, a broad $Z(4200)^-$ and $Z(3900)^-$ as a $J^P = 0^-$ K-matrix, and $Z(4700)^-$ as a $J^P = 2^-$ resonance onto the $m_{J/\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom).

14.6 Models with six $Z^- \rightarrow J/\psi\pi^-$

Fits with a sixth Z were performed on $B^0 \rightarrow J/\psi\pi^-K^+$, and the fit quality metrics for these are in Table 41.

Adding a 1^- Z as a separate resonance to the other K-matrices proves to be the most significant in terms of the likelihood, and it slightly improves the χ^2 s as well. We allow all masses and widths of these Zs to float, and their results are found in Table 42. The new Z resonance, which we label Z(4300), has mass at 4.349 ± 0.013 GeV and a width of 0.412 ± 0.047 GeV, and the masses and widths of the other Zs remain mostly the same, except for $J^P = 2^-$ Z(4700) which becomes half as broad.

Table 41: Goodness-of-fit metrics for adding the sixth Z with the given J^P to the J/ψ channel.

Variable	5 Z only	0^- Z	1^- Z	1^+ Z	2^- Z	2^+ Z
nPar	89	93	93	95	95	93
$\log L$	-3055806	-3055810	-3055965	-3055914	-3055942	-3055923
$\Delta(-2 \log L)$	-	8	318	216	272	234
2-D χ^2 (1024 bins)	1671	1673	1658	1592	1614	1669
2-D $\Delta\chi^2$	-	-3	13	79	57	2
4-D χ^2 (4096 bins)	4417	4417	4404	4357	4363	4406
4-D $\Delta\chi^2$	-	0	13	59	54	10

Table 42: Masses and widths of the Zs in the 6 Z model for the J/ψ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
Z(4200)	1^+	4.128 ± 0.003	0.324 ± 0.006
Z(4200)	0^-	4.311 ± 0.007	0.726 ± 0.015
Z(4430)	1^+	4.573 ± 0.002	0.156 ± 0.004
Z(4700)	2^-	4.761 ± 0.013	0.345 ± 0.045
Z(3900)	0^-	3.909 ± 0.005	0.145 ± 0.008
Z(4300)	1^-	4.349 ± 0.013	0.412 ± 0.047

The overall mass projection for $K\pi$ (Figure 101) remains the same, and the overall mass projection for $J/\psi\pi$ (Figure 102) also does not change much from the 5 Z fit. Similarly, there is not much difference in the $K\pi$ slices for the $J/\psi\pi$ distribution, seen in Figure 103.

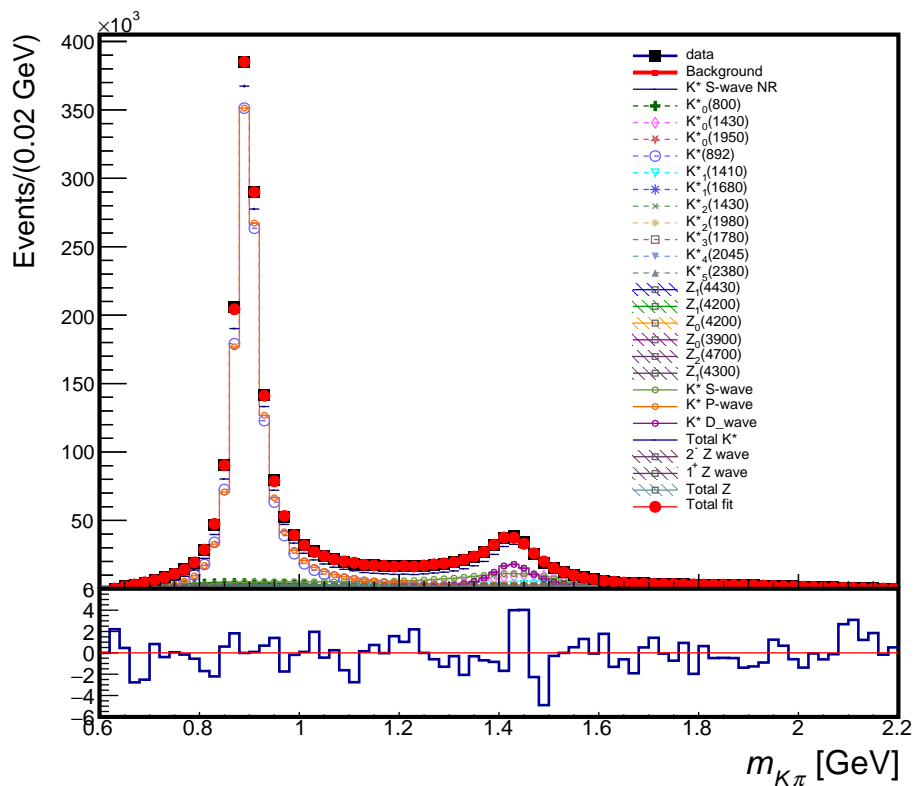


Figure 101: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black) and the fit (red) which adds a new $J^P = 1^- Z$ into the model with two $J^P = 0^- Z$ s, two 1^+ Zs, and a separate $2^- Z$ resonance.

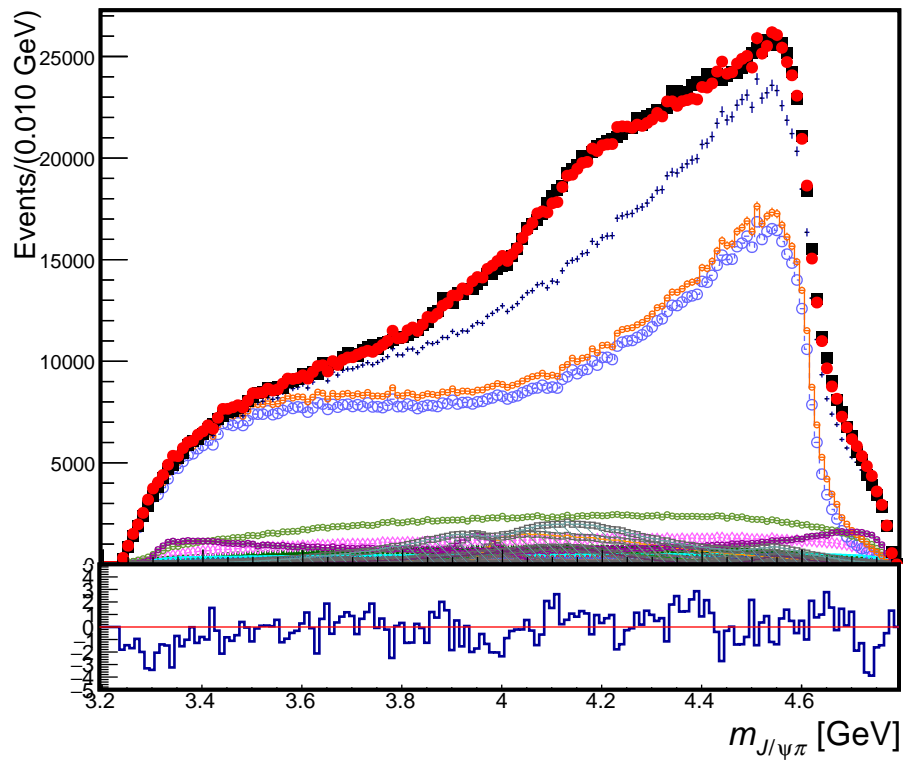


Figure 102: The distribution of $m_{J/\psi\pi}$ for the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black) and the fit (red) which adds a new $J^P = 1^- Z$ into the model with two $J^P = 0^- Z$ s, two 1^+ Zs, and a separate $2^- Z$ resonance.

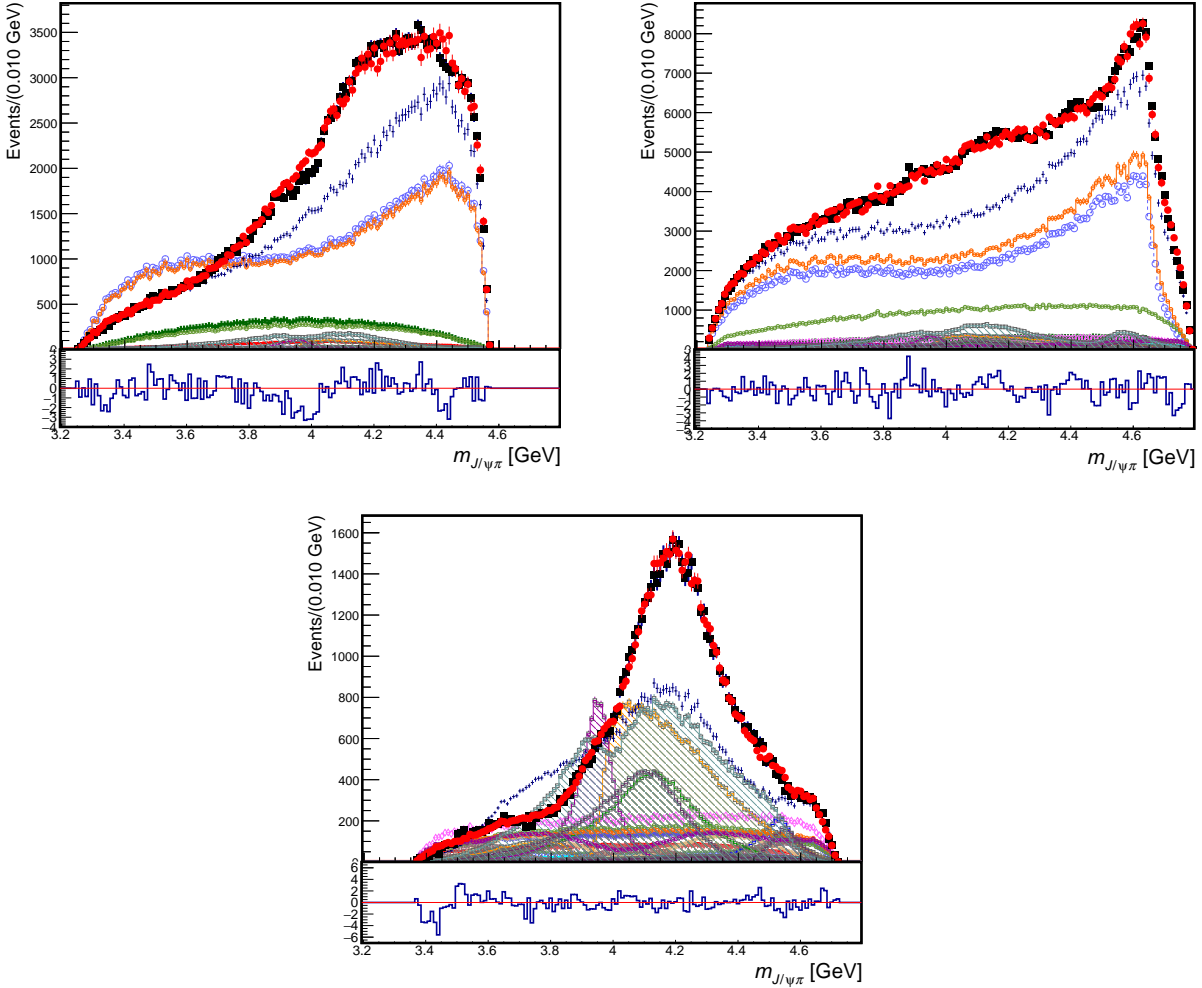


Figure 103: Projections of the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(4200)^-$ and $Z(4430)^-$ as a $J^P = 1^+$ K-matrix, a broad $Z(4200)^-$ and $Z(3900)^-$ as a $J^P = 0^-$ K-matrix, $Z(4700)^-$ as a $J^P = 2^-$ resonance, and $Z(4300)^-$ as a $J^P = 1^-$ resonance. onto the $m_{J/\psi \pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom).

14.7 Models with seven $Z^- \rightarrow J/\psi\pi^-$

Fits with a seventh Z were performed on $B^0 \rightarrow J/\psi\pi^-K^+$, and the fit quality metrics for these are in Table 43.

Adding an additional 1^+ Z into the K-matrix with the others is the most significant in terms of the likelihood (15.7σ), and it improves the 2-D and 4-D χ^2 s. We allow all masses and widths of these Zs to float, and their results are found in Table 44. The new Z resonance, which we label Z(3900), has mass at 3.911 ± 0.008 GeV and a width of 0.216 ± 0.009 GeV, and the masses and widths of the other Zs remain mostly similar, except for $J^P = 1^-$ Z(4300) which becomes twice as broad.

Our relatively broad 1^+ Z(3900) resonance should not be confused with the narrow $Z(3900)^- \rightarrow J/\psi\pi^-$ state discovered by the BESIII [17] and Belle [18] in $e^+e^- \rightarrow J/\psi\pi^-\pi^+(\gamma)$ data, seen also in other channels, which has a very similar mass, 3.8871 ± 0.0026 GeV, but an order of magnitude narrower width, 0.0284 ± 0.0026 GeV [2]. The narrow $Z_c(3900)$ state is often interpreted as isovector molecule made out of $D\bar{D}^*$ pairs, as such constituent masses sum up to 3.872-3.880 GeV. The narrow width is expected in the molecular binding, as the decays to the constituents are suppressed by phase-space, while decays to $J/\psi\pi$ require c and \bar{c} to be found in close proximity overcoming the spacial separation by a near confinement in two different charm mesons. The large width observed in this analysis, does not fit the molecular interpretation. However, it would be naturally expected in compact tetraquark models, in which all four quarks, including c and \bar{c} , are confined by direct color interactions into a volume not much larger than a regular charmonium state like J/ψ . The situation could be similar to the one observed near the $D\bar{D}_s^*$ threshold [26], where a narrow state is observed in e^+e^- annihilation, and a broad state is observed in B decays [27].

The overall mass projections for $K\pi$ (Figure 104) and $J/\psi\pi$ (Figure 105) do not change with the inclusion of the seventh Z. The mass projection for $J/\psi\pi$ improves in the $K\pi$ slices a bit, most notably in the slice below $K^*(892)$, seen in Figure 106.

Table 43: Goodness-of-fit metrics for adding the seventh Z with the given J^P to the J/ψ channel.

Variable	6 Z only	0^- Z	1^- Z	1^+ Z	2^- Z	2^+ Z
nPar	93	97	97	99	99	97
$\log L$	-3055965	-3056010	-3056015	-3056111	-3056054	-3055986
$\Delta(-2 \log L)$	-	90	100	292	178	42
2-D χ^2 (1024 bins)	1658	1673	1646	1574	1584	1649
2-D $\Delta\chi^2$	-	-16	11	84	74	9
4-D χ^2 (4096 bins)	4404	4413	4417	4363	4325	4398
4-D $\Delta\chi^2$	-	-10	-14	40	79	6

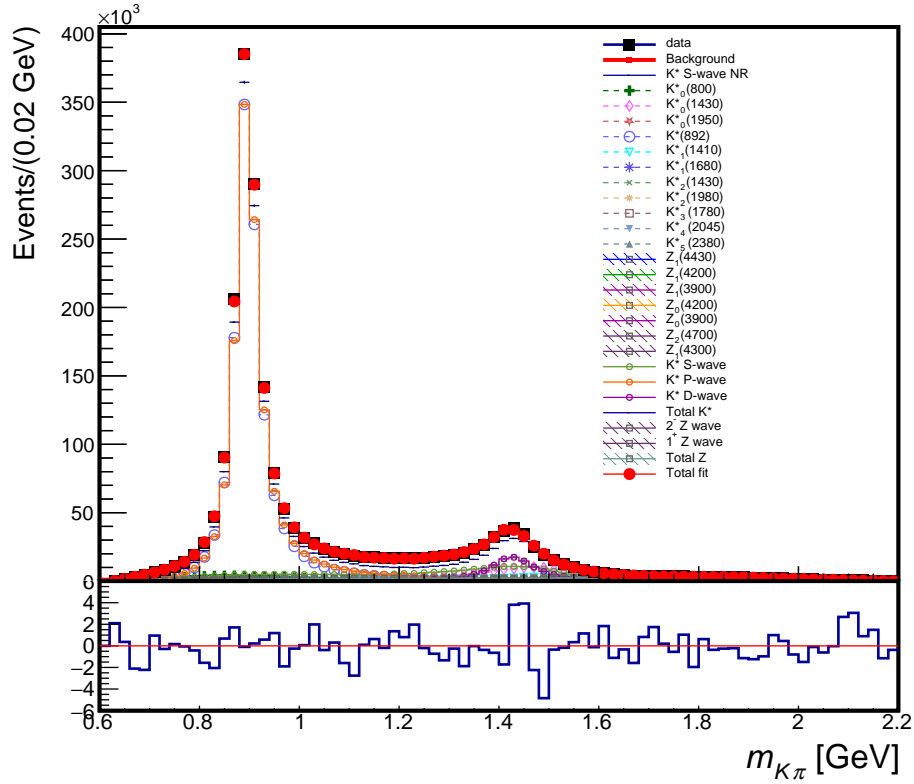
Figure 104: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+$ Z into the model with two other $J^P = 1^+$ Zs in a K-matrix, two $J^P = 0^-$ Zs, one 1^- Z, and a separate 2^- Z resonance.

Table 44: Masses and widths of the Zs in the 7 Z model for the J/ψ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
Z(4200)	1^+	4.139 ± 0.005	0.290 ± 0.005
Z(4200)	0^-	4.360 ± 0.008	0.915 ± 0.068
Z(4430)	1^+	4.592 ± 0.004	0.200 ± 0.006
Z(4700)	2^-	4.740 ± 0.015	0.321 ± 0.030
Z(3900)	0^-	3.871 ± 0.005	0.114 ± 0.006
Z(4300)	1^-	4.281 ± 0.020	0.915 ± 0.068
Z(3900)	1^+	3.911 ± 0.008	0.216 ± 0.009

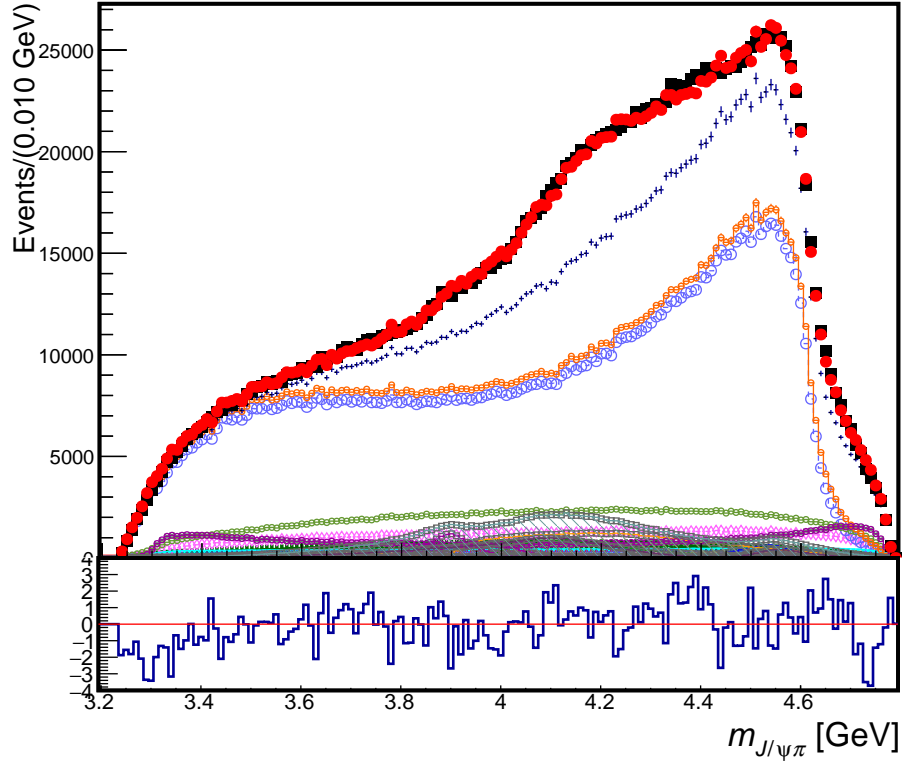


Figure 105: The distribution of $m_{J/\psi\pi}$ for the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+$ Z into the model with two other $J^P = 1^+$ Zs in a K-matrix, two $J^P = 0^-$ Zs, one 1^- Z, and a separate 2^- Z resonance.

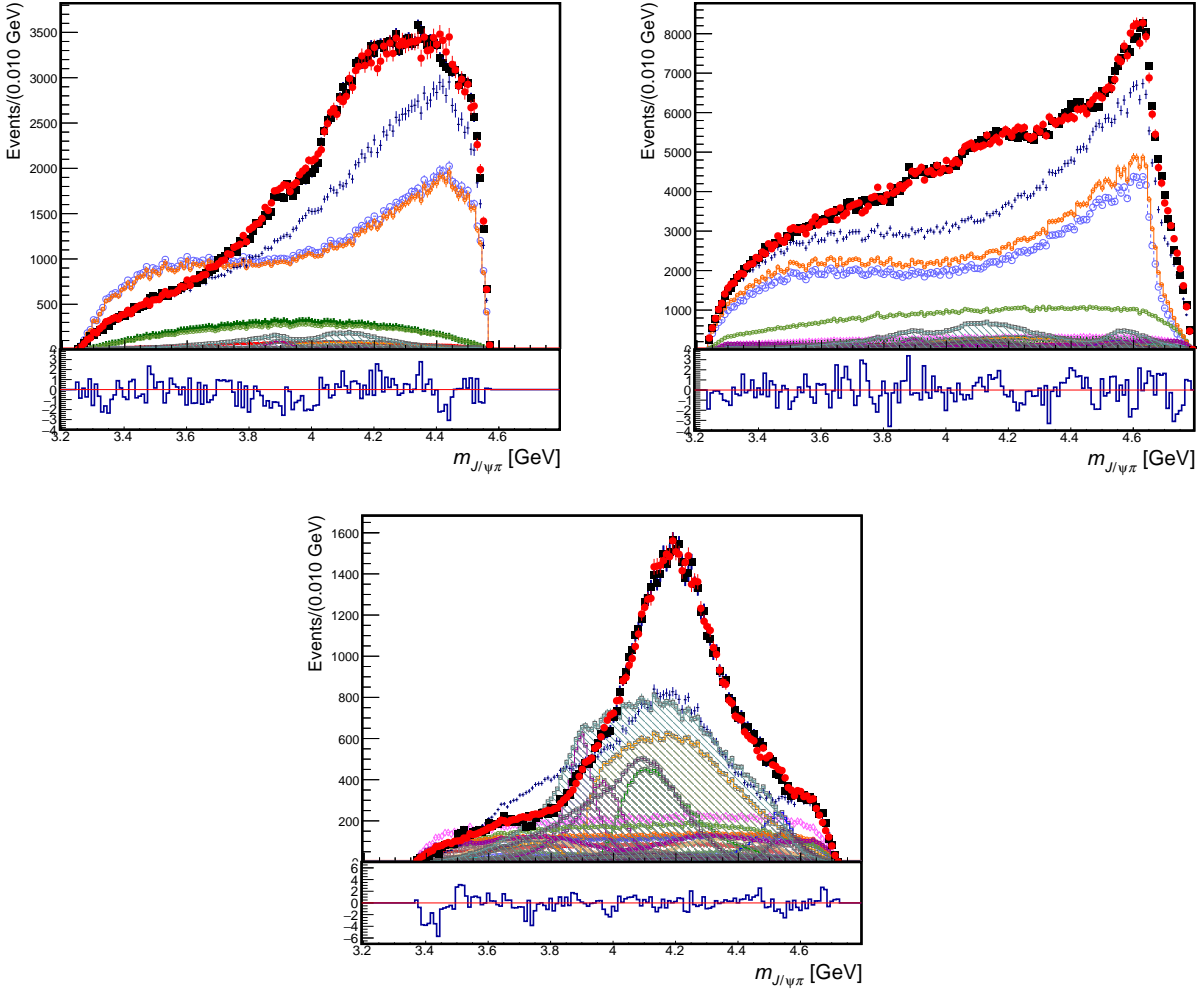


Figure 106: Projections of the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(3900)^-$, $Z(4200)^-$, and $Z(4430)^-$ as a $J^P = 1^+$ K-matrix, a broad $Z(4200)^-$ and $Z(3900)^-$ as a $J^P = 0^-$ K-matrix, $Z(4700)^-$ as a $J^P = 2^-$ resonance, and $Z(4300)^-$ as a $J^P = 1^-$ resonance. onto the $m_{J/\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom).

14.8 Models with eight $Z^- \rightarrow J/\psi\pi^-$

Fits with a eighth Z were performed on $B^0 \rightarrow J/\psi\pi^-K^+$, and the fit quality metrics for these are in Table 45.

Adding yet another 1^+ Z into the K-matrix with the others is the most significant in terms of the likelihood (15.3σ), and it again improves the 2-D and 4-D χ^2 s. We allow all masses and widths of these Zs to float, and their results are found in Table 46. The new Z resonance, which we label Z(4660), has mass at 4.665 ± 0.007 GeV and a width of 0.223 ± 0.013 GeV. The masses and widths of the other Zs do change significant in this model. Z(4430) has a lower pole mass and becomes a lot more narrow, Z(4200) also has a lower pole mass and becomes more narrow (though to a lesser extent than Z(4430)), while the rest of the Zs do not change much in terms of mass and width.

Table 45: Goodness-of-fit metrics for adding the eighth Z with the given J^P to the J/ψ channel.

Variable	7 Z only	0^- Z	1^- Z	1^+ Z	2^- Z	2^+ Z
nPar	99	103	103	105	105	103
$\log L$	-3056111	-3056106	-3056117	-3056251	-3056133	-3056132
$\Delta(-2 \log L)$	-	-10	12	280	44	42
2-D χ^2 (1024 bins)	1574	1698	1572	1528	1566	1568
2-D $\Delta\chi^2$	-	-124	2	46	8	5
4-D χ^2 (4096 bins)	4363	4375	4362	4322	4353	4359
4-D $\Delta\chi^2$	-	-12	2	41	10	4

Table 46: Masses and widths of the Zs in the 8 Z model for the J/ψ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
Z(4200)	1^+	4.159 ± 0.006	0.225 ± 0.005
Z(4200)	0^-	4.362 ± 0.009	0.737 ± 0.017
Z(4430)	1^+	4.472 ± 0.005	0.065 ± 0.005
Z(4700)	2^-	4.746 ± 0.008	0.297 ± 0.025
Z(3900)	0^-	3.870 ± 0.005	0.118 ± 0.008
Z(4300)	1^-	4.291 ± 0.016	0.943 ± 0.075
Z(3900)	1^+	3.912 ± 0.012	0.200 ± 0.009
Z(4660)	1^+	4.665 ± 0.007	0.223 ± 0.013

The overall mass projection for $K\pi$ (Figure 107) remains pretty much the same after

including an eighth Z , while the overall $J/\psi\pi$ mass projection (Figure 108) improves near 4.4 GeV. A similar improvement is seen in the mass projection for $J/\psi\pi$ in the $K\pi$ slice between $K^*(892)$ and $K^*(1430)$, and the slice above the $B^0 \rightarrow \psi(2S)K\pi$ phase space limit for $m(K\pi)$ improves near the mass of the new pole, both seen in Figure 109.

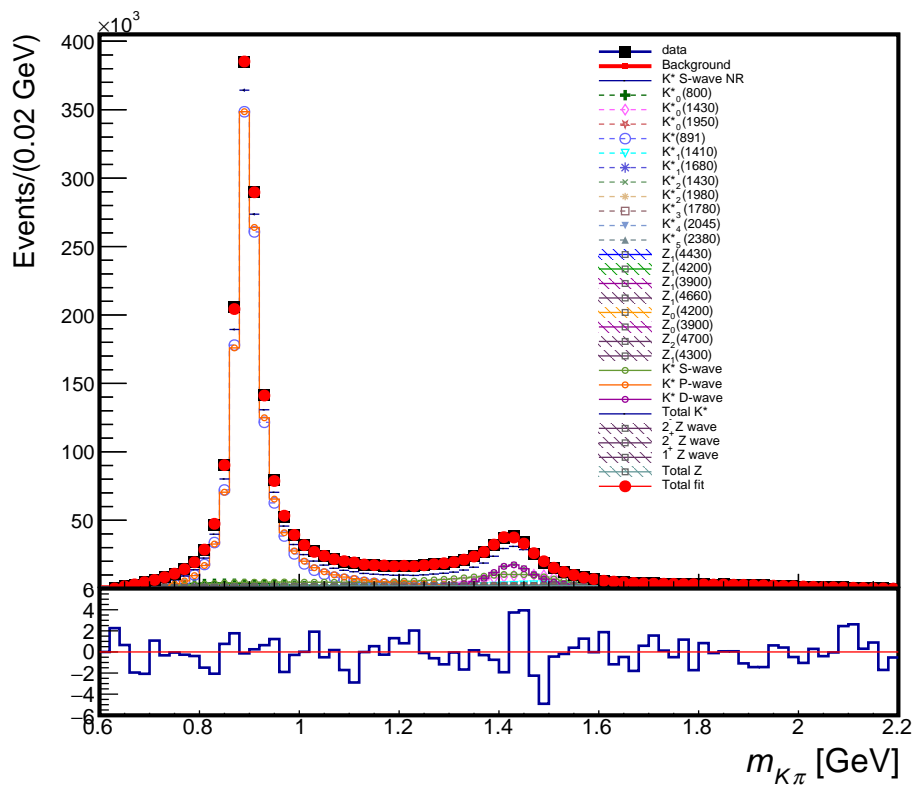


Figure 107: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+$ Z into the model with three other $J^P = 1^+$ Z s in a K-matrix, two $J^P = 0^-$ Z s, one 1^- Z , and a separate 2^- Z resonance.

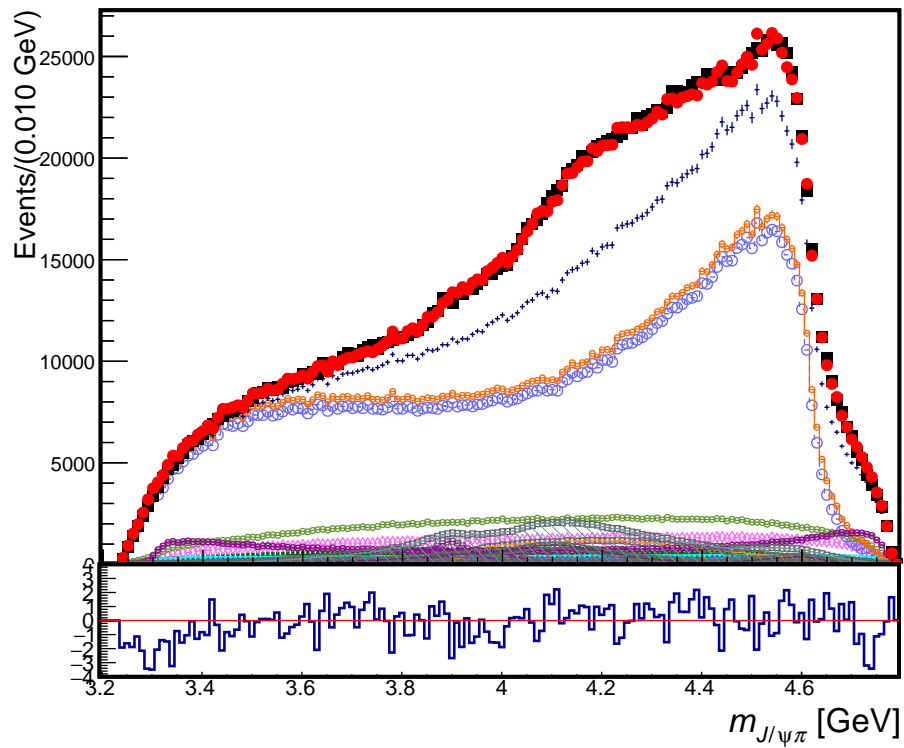


Figure 108: The distribution of $m_{J/\psi\pi}$ for the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+$ Z into the model with three other $J^P = 1^+$ Zs in a K-matrix, two $J^P = 0^-$ Zs, one 1^- Z, and a separate 2^- Z resonance.

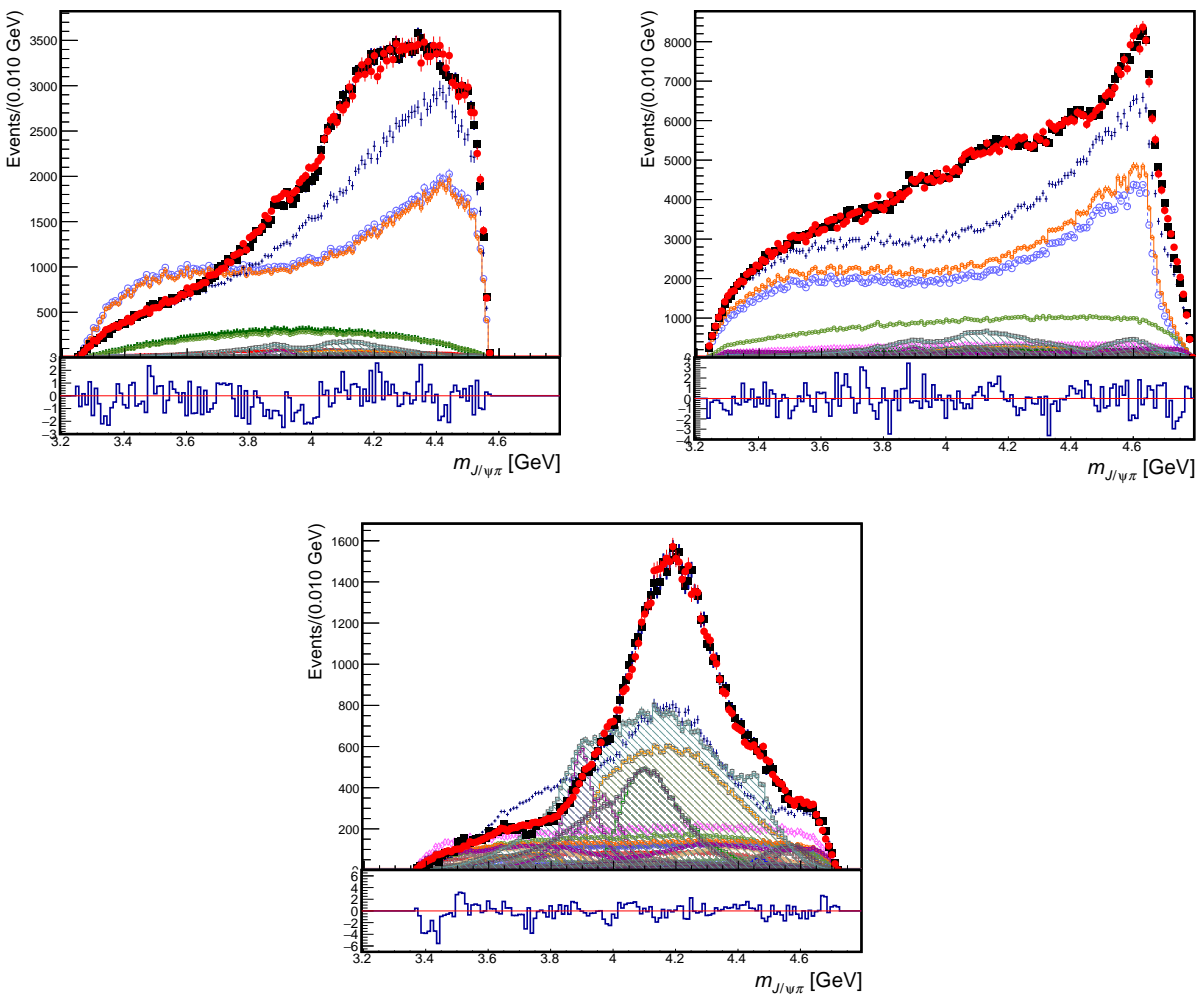


Figure 109: Projections of the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(3900)^-$, $Z(4200)^-$, $Z(4430)^-$, and $Z(4660)^-$ as a $J^P = 1^+$ K-matrix, a broad $Z(4200)^-$ and $Z(3900)^-$ as a $J^P = 0^-$ K-matrix, $Z(4700)^-$ as a $J^P = 2^-$ resonance, and $Z(4300)^-$ as a $J^P = 1^-$ resonance. onto the $m_{J/\psi \pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom).

14.9 Models with nine $Z^- \rightarrow J/\psi\pi^-$

Fits with a ninth Z were performed on $B^0 \rightarrow J/\psi\pi^-K^+$, and the fit quality metrics for these are in Table 47.

Adding a fifth 1^+ Z into the K-matrix with the others is again the most significant in terms of the likelihood (10.4σ), and it also moderately improves the 2-D and 4-D χ^2 s. We allow all masses and widths of these Zs to float, and their results are found in Table 48. The new Z resonance, which we label Z(4800), has mass at 4.817 ± 0.012 GeV and a width of 0.410 ± 0.036 GeV. The masses and widths of the other Zs do change significantly in this model. In the 1^+ Z K-matrix for the 8 Z model, there was a less broad pole at 4.47 GeV (then labeled Z(4430)) and a moderately broad pole at 4.67 GeV (labeled Z(4660)) along with Z(4200) and Z(3900) as usual. However here we get a pole again at 4.56 GeV, which was typical for Z(4430) in earlier fits and in the $\psi(2S)$ channel, another pole at 4.4 GeV, and the typical Z(4200) and Z(3900) poles along with the new Z(4800) pole. The masses and widths of Zs of other quantum numbers change very little with the addition of a ninth Z. The masses and widths of the K^* s that are floating can be found in Table 49, and the fit fractions for all resonances and partial waves are found in Table 50.

Table 47: Goodness-of-fit metrics for adding the ninth Z with the given J^P to the J/ψ channel.

Variable	8 Z only	0^- Z	1^- Z	1^+ Z	2^- Z	2^+ Z
nPar	105	109	109	111	111	109
$\log L$	-3056251	-3056282	-3056258	-3056324	-3056275	-3056272
$\Delta(-2 \log L)$	-	62	14	146	48	42
2-D χ^2 (1024 bins)	1528	1519	1528	1506	1520	1523
2-D $\Delta\chi^2$	-	9	0	23	8	5
4-D χ^2 (4096 bins)	4322	4315	4322	4295	4311	4318
4-D $\Delta\chi^2$	-	7	0	27	11	5

The overall mass projections for $K\pi$ (Figure 110) and $J/\psi\pi$ (Figure 111) do not change much after adding a ninth Z. Also the $K\pi$ slices for the $J/\psi\pi$ mass projections remain largely the same in the new 9 Z model, shown in Figure 112.

Table 48: Masses and widths of the Zs in the 9 Z model for the J/ψ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
Z(4200)	1^+	4.121 ± 0.004	0.176 ± 0.007
Z(4200)	0^-	4.363 ± 0.007	0.737 ± 0.015
Z(4430)	1^+	4.562 ± 0.005	0.152 ± 0.006
Z(4700)	2^-	4.746 ± 0.011	0.296 ± 0.028
Z(3900)	0^-	3.870 ± 0.005	0.118 ± 0.008
Z(4300)	1^-	4.309 ± 0.026	0.939 ± 0.073
Z(3900)	1^+	3.881 ± 0.008	0.291 ± 0.012
Z(4400)	1^+	4.405 ± 0.006	0.093 ± 0.009
Z(4800)	1^+	4.817 ± 0.012	0.410 ± 0.036

Table 49: Masses and widths of the K^* s in the 9 Z model for the J/ψ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
$K^*(892)$	1^-	0.890 ± 0.00003	0.045 ± 0.00007
$K_0^*(800)$	0^+	0.814 ± 0.005	0.293 ± 0.004
$K_0^*(1430)$	0^+	1.457 ± 0.002	0.222 ± 0.002
$K_2^*(1430)$	2^+	1.417 ± 0.0003	0.098 ± 0.0006

Table 50: All resonance fit fractions in the 9 Z fit for the J/ψ channel.

Contribution	Fit fraction
K^* Total S-wave	13.8%
K^* S-wave NR	0.7%
$K_0^*(800)$	3.4%
$K_0^*(1430)$	7.7%
$K_0^*(1950)$	0.4%
K^* Total P-wave	65.3%
$K^*(892)$	61.3%
$K_1^*(1410)$	1.9%
$K_1^*(1680)$	1.1%
K^* Total D-wave	6.5%
$K_2^*(1430)$	6.2%
$K_2^*(1980)$	0.5%
$K_3^*(1780)$	0.3%
$K_4^*(2045)$	0.03%
$K_5^*(2380)$	0.012%
Z Total 1^+ wave	1.7%
Z(4200)	0.9%
Z(4430)	0.3%
Z(3900)	0.7%
Z(4400)	0.01%
Z(4800)	0.02%
Z Total 0^- wave	3.23%
$Z_0(4200)$	2.7%
$Z_0(3900)$	0.7%
Z Total 1^- wave	0.26%
$Z_1(4300)$	0.26%
Z Total 2^- wave	0.13%
$Z_2(4700)$	0.13%

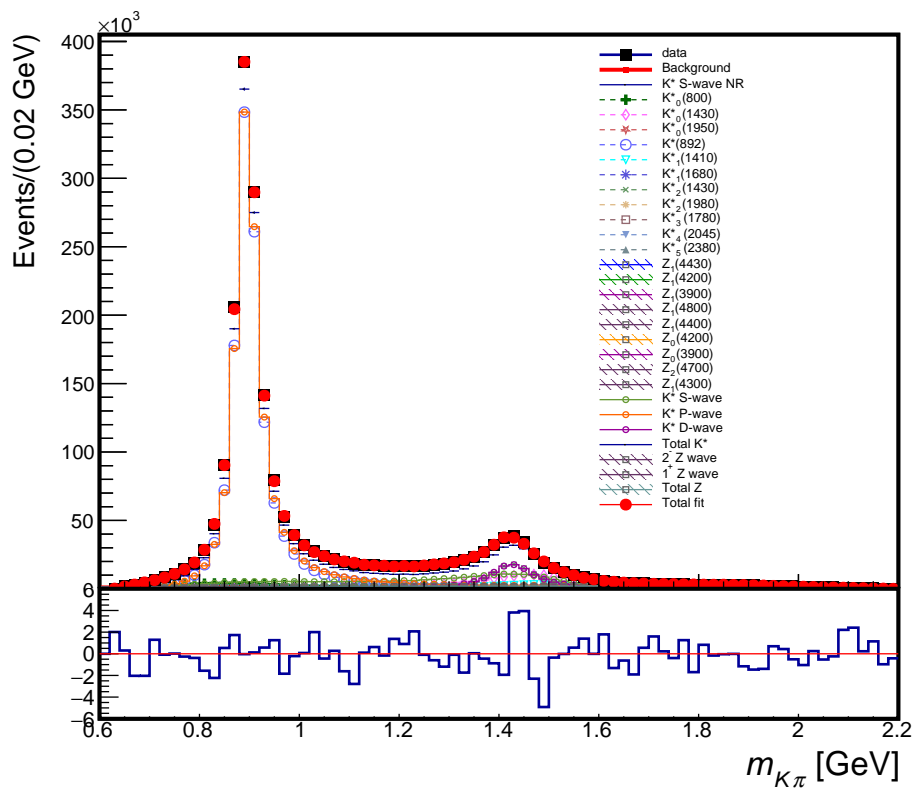


Figure 110: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+$ Z into the model with four other $J^P = 1^+$ Zs in a K-matrix, two $J^P = 0^-$ Zs, one 1^- Z, and a separate 2^- Z resonance.

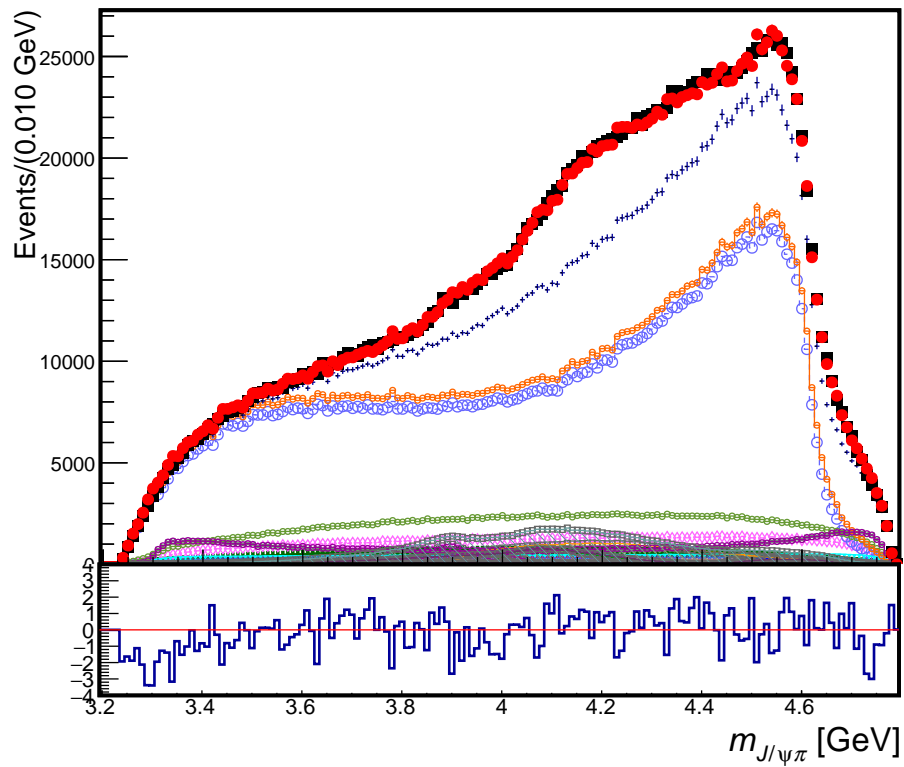


Figure 111: The distribution of $m_{J/\psi\pi}$ for the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+$ Z into the model with four other $J^P = 1^+$ Z s in a K -matrix, two $J^P = 0^-$ Z s, one 1^- Z , and a separate 2^- Z resonance.

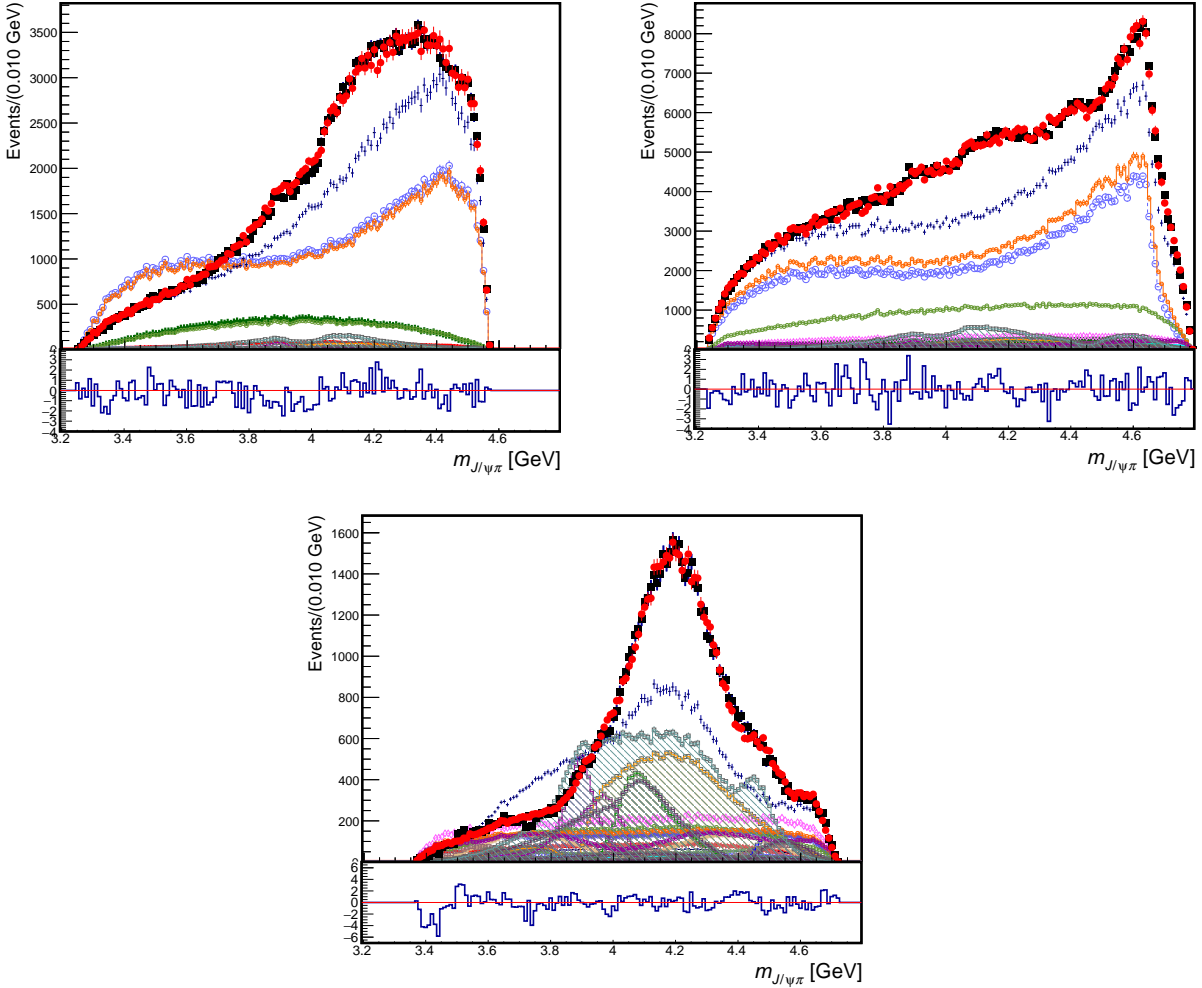


Figure 112: Projections of the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(3900)^-$, $Z(4200)^-$, $Z(4400)^-$, $Z(4430)^-$, and $Z(4800)^-$ as a $J^P = 1^+$ K-matrix, a broad $Z(4200)^-$ and $Z(3900)^-$ as a $J^P = 0^-$ K-matrix, $Z(4700)^-$ as a $J^P = 2^-$ resonance, and $Z(4300)^-$ as a $J^P = 1^-$ resonance. onto the $m_{J/\psi \pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom).

14.10 Models with ten $Z^- \rightarrow J/\psi\pi^-$

Fits with a tenth Z were performed on $B^0 \rightarrow J/\psi\pi^-K^+$, and the fit quality metrics for these are in Table 51. While in terms of changes in log-likelihood value these fits are significant, much like the $\psi(2S)$ channel, these fits only converge with warnings about machine accuracy limits, and they have not positive definite error matrices. They also fail to produce significant change in 2-D and 4-D χ^2 values. As such, the most reliable $Z^- \rightarrow J/\psi\pi^-$ resonance model for the J/ψ channel is the one in the 9 Z resonances described in the previous subsection.

Table 51: Goodness-of-fit metrics for adding the tenth Z with the given J^P to the J/ψ channel.

Variable	9 Z only	$0^- Z$	$1^- Z$	$1^+ Z$	$2^- Z$	$2^+ Z$
nPar	111	115	115	117	117	115
$\log L$	-3056324	-3056417	-3056329	-3056335	-3056338	-3056343
$\Delta(-2 \log L)$	-	186	10	22	28	38
2-D χ^2 (1024 bins)	1506	1511	1505	1506	1503	1499
2-D $\Delta\chi^2$	-	-5	0	-1	2	7
4-D χ^2 (4096 bins)	4295	4289	4294	4294	4289	4291
4-D $\Delta\chi^2$	-	6	1	1	6	4

15 Resonant Z_K models with J/ψ channel

This section will again try to add resonances into the $B^0 \rightarrow J/\psi\pi^-K^+$ model but in the $Z_K^+ \rightarrow J/\psi K^+$ decay chain. The modeling will begin with one $Z_K^+ \rightarrow J/\psi K^+$ resonance of various J^P added to the default 9 Z model from the previous section and then sequentially trying to add more Z_K resonances.

15.1 Plots of $m_{J/\psi K}$ with no $Z_K^+ \rightarrow J/\psi K^+$

To compare with models that add $Z_K^+ \rightarrow J/\psi K^+$ states, presented here are the $m_{J/\psi K}$ distributions both overall (Fig 113) and in different slices of $m_{K\pi}$ (Fig 114) from the fits done with 9 $Z^- \rightarrow J/\psi\pi^-$ states without any Z_K states. Fit quality metrics are in the last section (14.9).

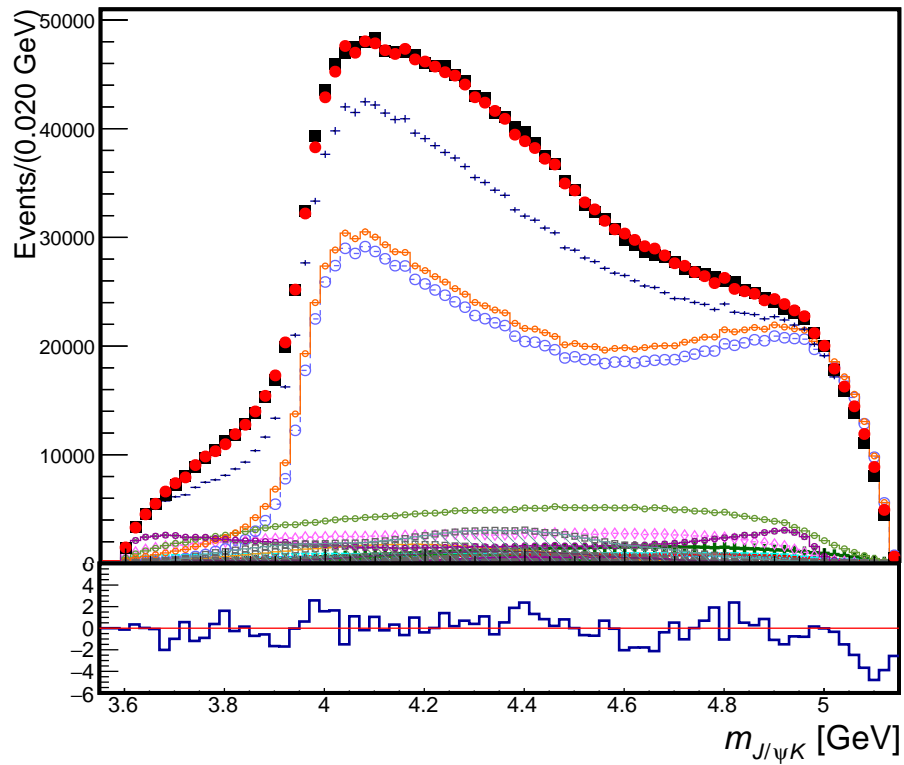


Figure 113: The distribution of $m_{J/\psi K}$ for the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black) and the fit (red) which has five $J^P = 1^+$ Zs in a K-matrix, two $J^P = 0^-$ Zs, one 1^- Z, and a 2^- Z resonance.

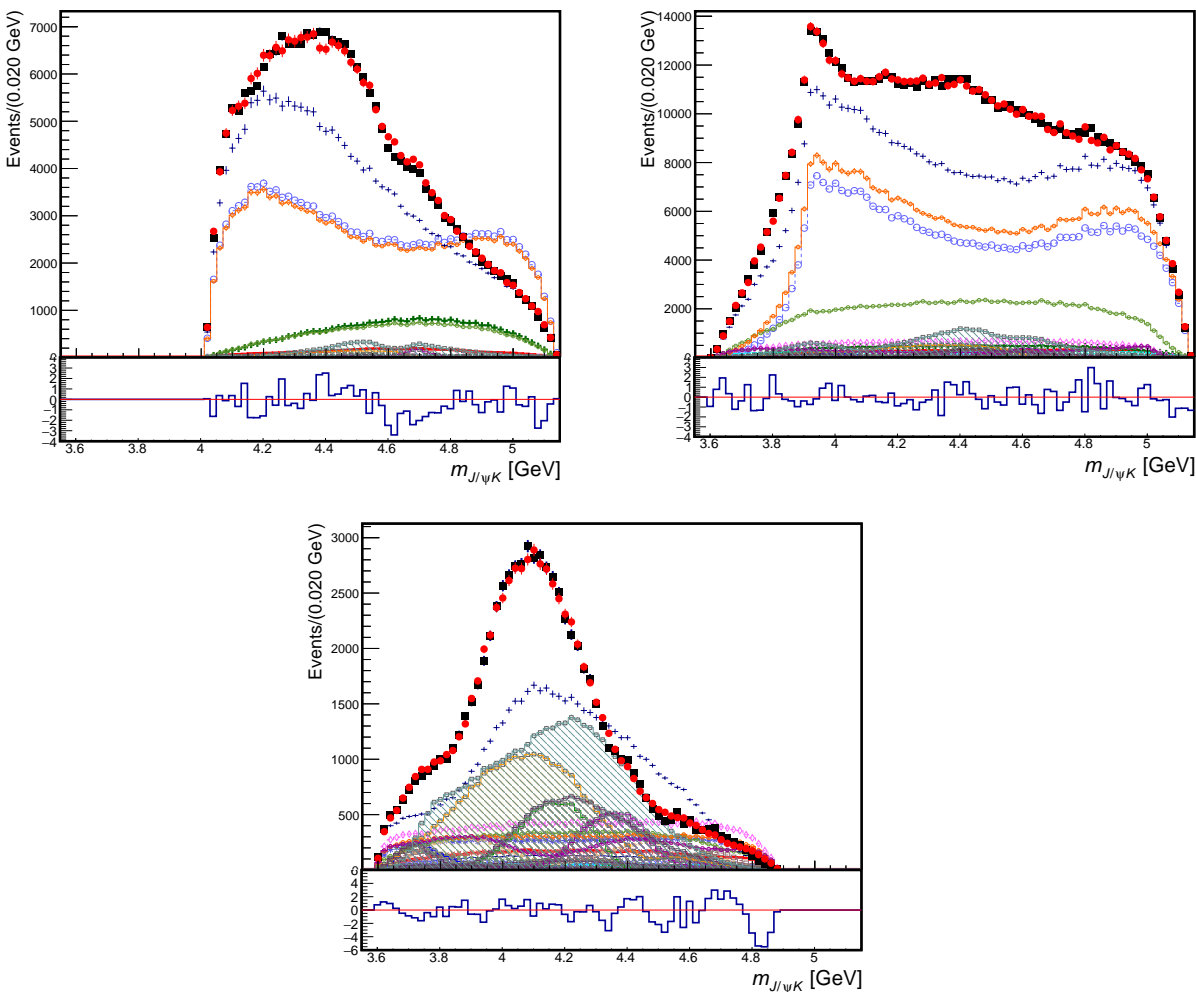


Figure 114: Projections of the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(3900)^-$, $Z(4200)^-$, $Z(4400)^-$, $Z(4430)^-$, and $Z(4800)^-$ as a $J^P = 1^+$ K-matrix, a broad $Z(4200)^-$ and $Z(3900)^-$ as a $J^P = 0^-$ K-matrix, $Z(4700)^-$ as a $J^P = 2^-$ resonance, and $Z(4300)^-$ as a $J^P = 1^-$ resonance. onto the $m_{J/\psi \pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom).

15.2 Models with one $Z_K^+ \rightarrow J/\psi K^+$

Fits with a single Z_K were performed on $B^0 \rightarrow J/\psi \pi^- K^+$, and the fit quality metrics for these are in Table 52.

Adding a 1^+ Z_K is the most significant in terms of the likelihood, and it also improves the 2-D and 4-D χ^2 s. The plot of the pulls for each bin in the Dalitz plane for the 2-D χ^2 is in Figure 120, and the distribution of the 2-D and 4-D pulls fit to a Gaussian are in Figure 121, which feature a mean consistent with 0 for the 2-D distribution. The width of the 2-D distribution is a bit larger than 1, which is not too surprising given the 2-D χ^2 fit quality. The 4-D distribution has a mean close to 0 and a width close to 1. We allow the mass and width of the Z_K to float, and the result is found in Table 55. The new Z_K resonance, which we label $Z_K(4470)$, has mass at 4.469 ± 0.013 GeV and a width of 0.204 ± 0.015 GeV. The masses and widths of the Z s ($J/\psi \pi$ states) do not change much, and their values can be found in Table 53. The masses and widths of the K^* s that are floating can be found in Table 54, and the fit fractions for all resonances and partial waves are found in Table 56.

Table 52: Goodness-of-fit metrics for adding one Z_K with the given J^P to the J/ψ channel.

Variable	No Z_K	$0^- Z_K$	$1^- Z_K$	$1^+ Z_K$	$2^- Z_K$	$2^+ Z_K$
nPar	111	115	115	117	117	115
$\log L$	-3056324	-3056385	-3056390	-3056472	-3056341	-3056420
$\Delta(-2 \log L)$	-	122	132	296	34	192
2-D χ^2 (1024 bins)	1506	1493	1492	1457	1494	1506
2-D $\Delta\chi^2$	-	12	13	49	12	-1
4-D χ^2 (4096 bins)	4295	4292	4275	4248	4289	4278
4-D $\Delta\chi^2$	-	3	20	47	6	17

The overall mass projections for $K\pi$ (Figure 115) and $J/\psi\pi$ (Figure 116) do not change much after adding a Z_K resonance to the 9 Z model. The $K\pi$ slices for the $J/\psi\pi$ mass projections do improve slightly most notably in the middle of $J/\psi\pi$ phase space in the $m_{K\pi}$ slice below $K^*(892)$, seen in Figure 117. The overall mass projection for $J/\psi K$ improves slightly near the new pole (Fig. 118). The $K\pi$ slices for the $J/\psi K$ mass projections improve

Table 53: Masses and widths of the Z s in the 1 Z_K model for the J/ψ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
Z(4200)	1^+	4.118 ± 0.005	0.174 ± 0.007
Z(4200)	0^-	4.337 ± 0.007	0.651 ± 0.015
Z(4430)	1^+	4.554 ± 0.005	0.152 ± 0.006
Z(4700)	2^-	4.749 ± 0.013	0.291 ± 0.029
Z(3900)	0^-	3.880 ± 0.006	0.134 ± 0.008
Z(4300)	1^-	4.310 ± 0.002	0.908 ± 0.075
Z(3900)	1^+	3.899 ± 0.008	0.270 ± 0.012
Z(4400)	1^+	4.400 ± 0.007	0.087 ± 0.009
Z(4800)	1^+	4.798 ± 0.010	0.340 ± 0.032

Table 54: Masses and widths of the K^* s in the 1 Z_K model for the J/ψ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
$K^*(892)$	1^-	0.890 ± 0.00003	0.045 ± 0.00007
$K_0^*(800)$	0^+	0.805 ± 0.005	0.306 ± 0.004
$K_0^*(1430)$	0^+	1.455 ± 0.002	0.218 ± 0.002
$K_2^*(1430)$	2^+	1.417 ± 0.0003	0.098 ± 0.0006

Table 55: Masses and widths of the Z_K in the 1 Z_K model for the J/ψ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
$Z_K(4470)$	1^+	4.469 ± 0.013	0.204 ± 0.015

considerably as well near the new pole, best seen in the slices below $K^*(892)$ and between $K^*(892)$ and $K^*(1430)$ in Figure 119.

As it will become clear from the next subsection, the model presented here becomes our nominal pick for the J/ψ channel, which we later use in the comparison to the $\psi(2S)$ model (Sec. 16). We include additional plots showing the fit quality on the Dalitz plane and in all four fitted dimensions of the data (Figs. 120-121). We also show comparison between this model and the data on various moments of the K^{*0} helicity angle, which is another way to test fit quality on the Dalitz plane (see Appendix B).

Table 56: All resonance fit fractions in the 1 Z_K fit for the J/ψ channel.

Contribution	Fit fraction
K^* Total S-wave	$14.1 \pm 0.2\%$
K^* S-wave NR	$0.7 \pm 0.04\%$
$K_0^*(800)$	$3.4 \pm 0.1\%$
$K_0^*(1430)$	$7.8 \pm 0.1\%$
$K_0^*(1950)$	$0.5 \pm 0.02\%$
K^* Total P-wave	$65.7 \pm 0.2\%$
$K^*(892)$	$61.6 \pm 0.2\%$
$K_1^*(1410)$	$1.9 \pm 0.04\%$
$K_1^*(1680)$	$1.1 \pm 0.04\%$
K^* Total D-wave	$6.3 \pm 0.05\%$
$K_2^*(1430)$	$6.1 \pm 0.04\%$
$K_2^*(1980)$	$0.4 \pm 0.02\%$
$K_3^*(1780)$	$0.2 \pm 0.01\%$
$K_4^*(2045)$	$0.02 \pm 0.003\%$
$K_5^*(2380)$	$0.007 \pm 0.002\%$
Z Total 1^+ wave	$1.7 \pm 0.006\%$
Z(4200)	$0.9 \pm 0.003\%$
Z(4430)	$0.3 \pm 0.001\%$
Z(3900)	$0.8 \pm 0.003\%$
Z(4400)	$0.02 \pm 0.0001\%$
Z(4800)	$0.004 \pm 0.00001\%$
Z Total 0^- wave	$3.07 \pm 0.01\%$
$Z_0(4200)$	$2.6 \pm 0.01\%$
$Z_0(3900)$	$0.7 \pm 0.002\%$
Z Total 1^- wave	$0.26 \pm 0.001\%$
$Z_1(4300)$	$0.26 \pm 0.001\%$
Z Total 2^- wave	$0.14 \pm 0.0005\%$
$Z_2(4700)$	$0.14 \pm 0.0005\%$
Z_K Total 1^+ wave	$0.046 \pm 0.0002\%$
$Z_K(4470)$	$0.046 \pm 0.0002\%$

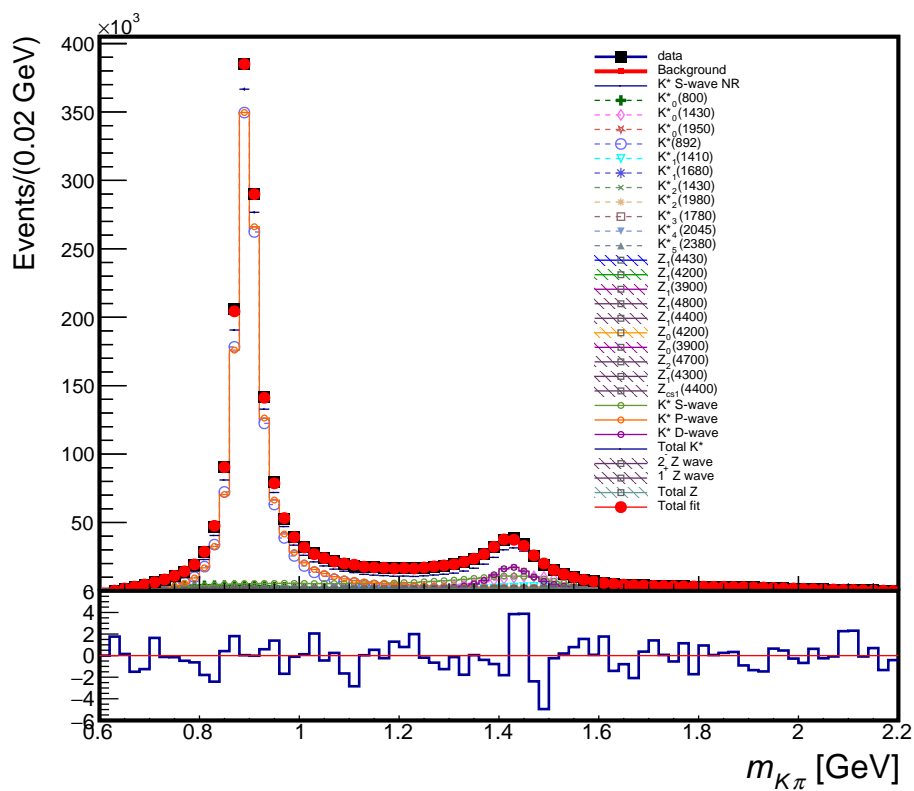


Figure 115: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi\pi^-K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+ Z_K$ into the default model which already has 9 $Z \rightarrow J/\psi\pi$ resonances.

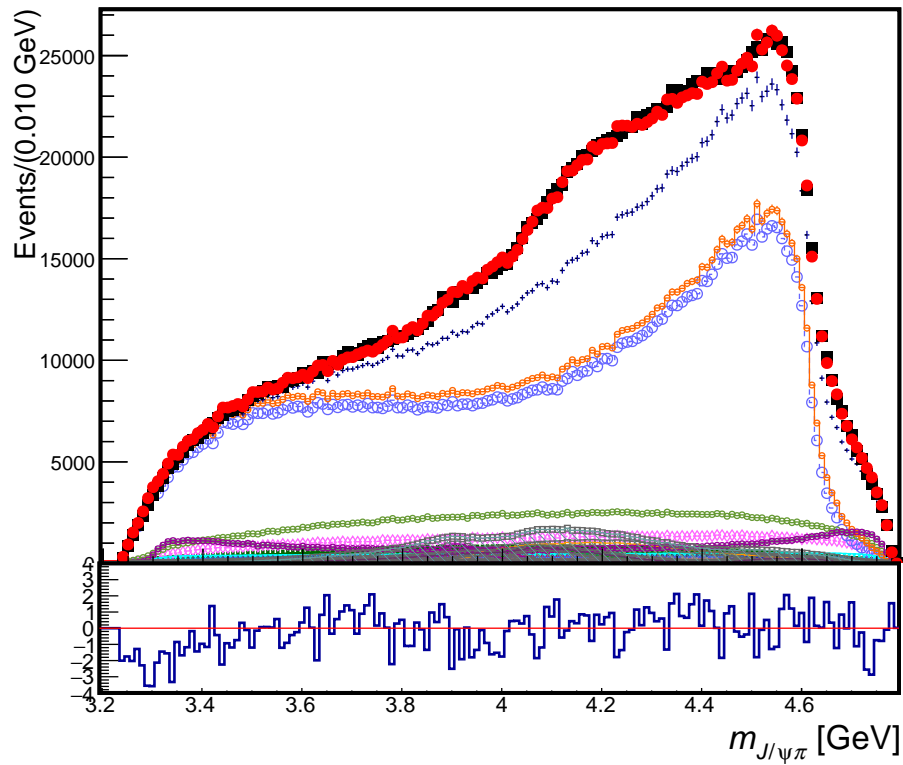


Figure 116: The distribution of $m_{J/\psi\pi}$ for the $B^0 \rightarrow J/\psi\pi^- K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+ Z_K$ into the default model which already has 9 $Z \rightarrow J/\psi\pi$ resonances.

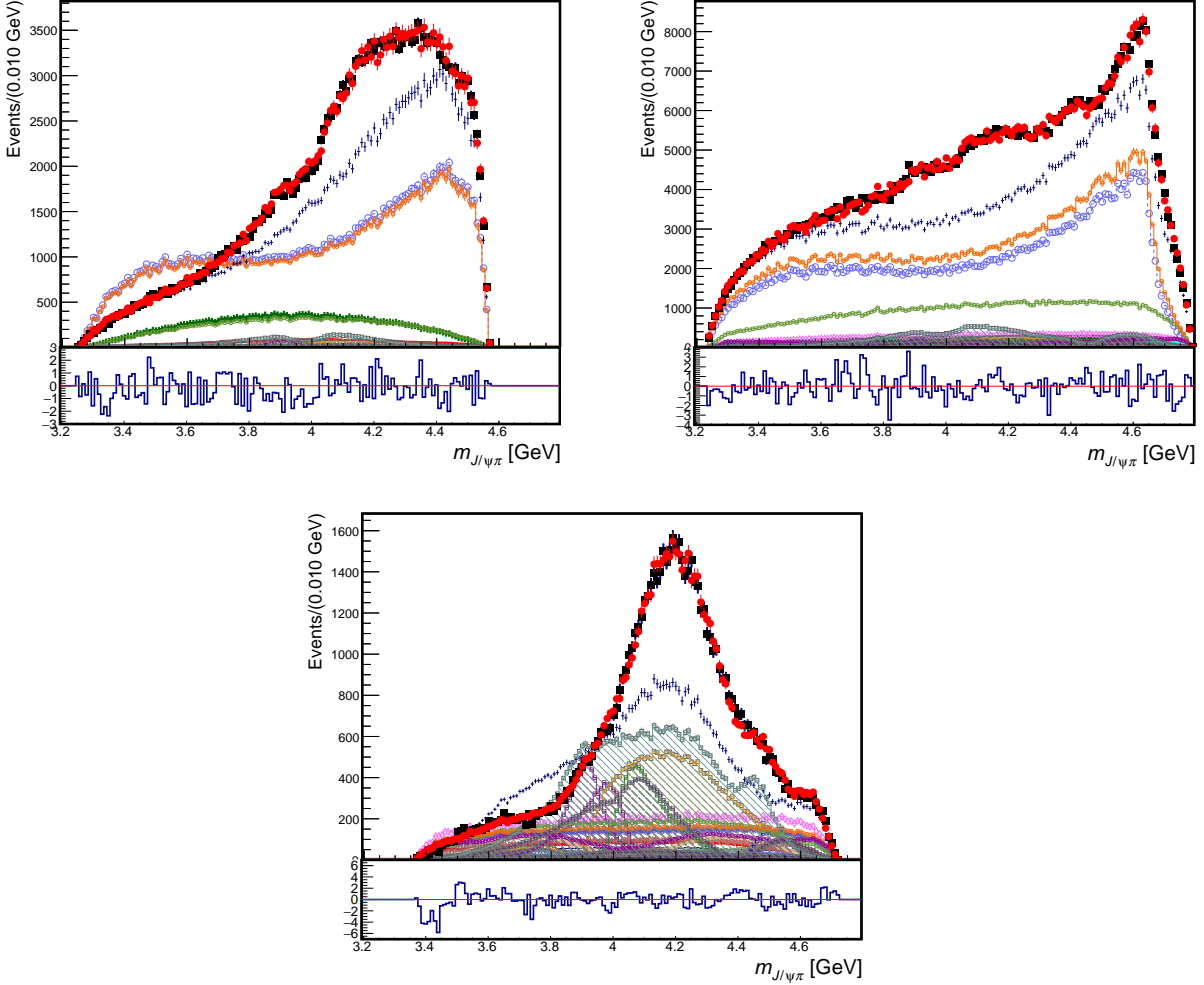


Figure 117: Projections of the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(3900)^-$, $Z(4200)^-$, $Z(4400)^-$, $Z(4430)^-$, and $Z(4800)^-$ as a $J^P = 1^+$ K-matrix, a broad $Z(4200)^-$ and $Z(3900)^-$ as a $J^P = 0^-$ K-matrix, $Z(4700)^-$ as a $J^P = 2^-$ resonance, $Z(4300)^-$ as a $J^P = 1^-$ resonance, and $Z_{cs}(4470)^-$ as a $J^P = 1^+$ $J/\psi K$ resonance. onto the $m_{J/\psi \pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom).

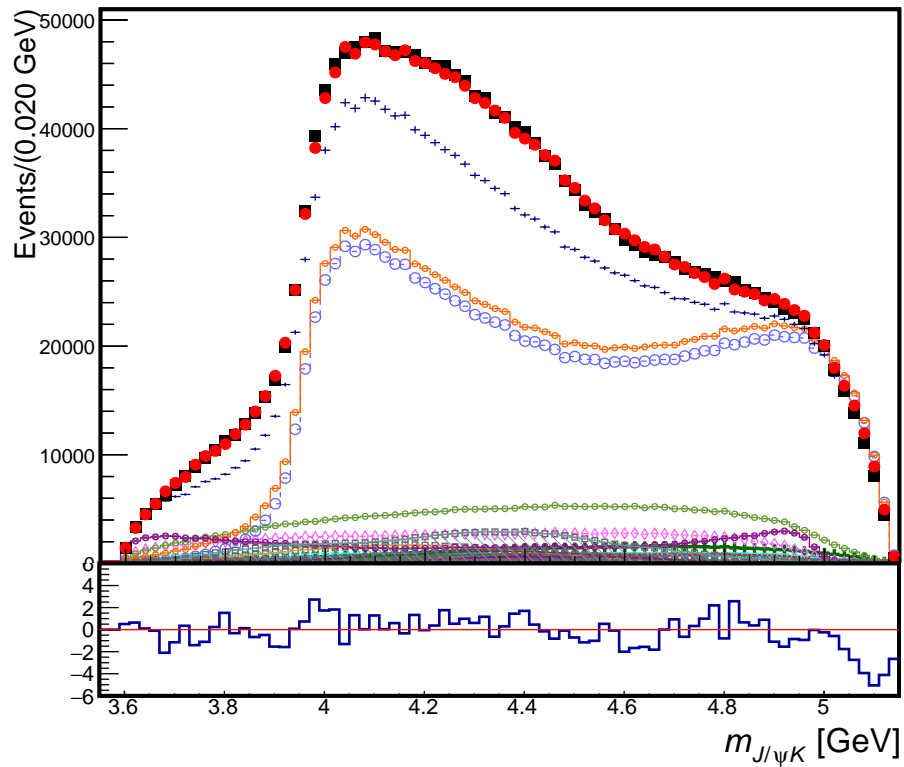


Figure 118: The distribution of $m_{J/\psi K}$ for the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black) and the fit (red) which adds a new $J^P = 1^+$ Z_K into the default model which already has 9 $Z \rightarrow J/\psi \pi$ resonances.

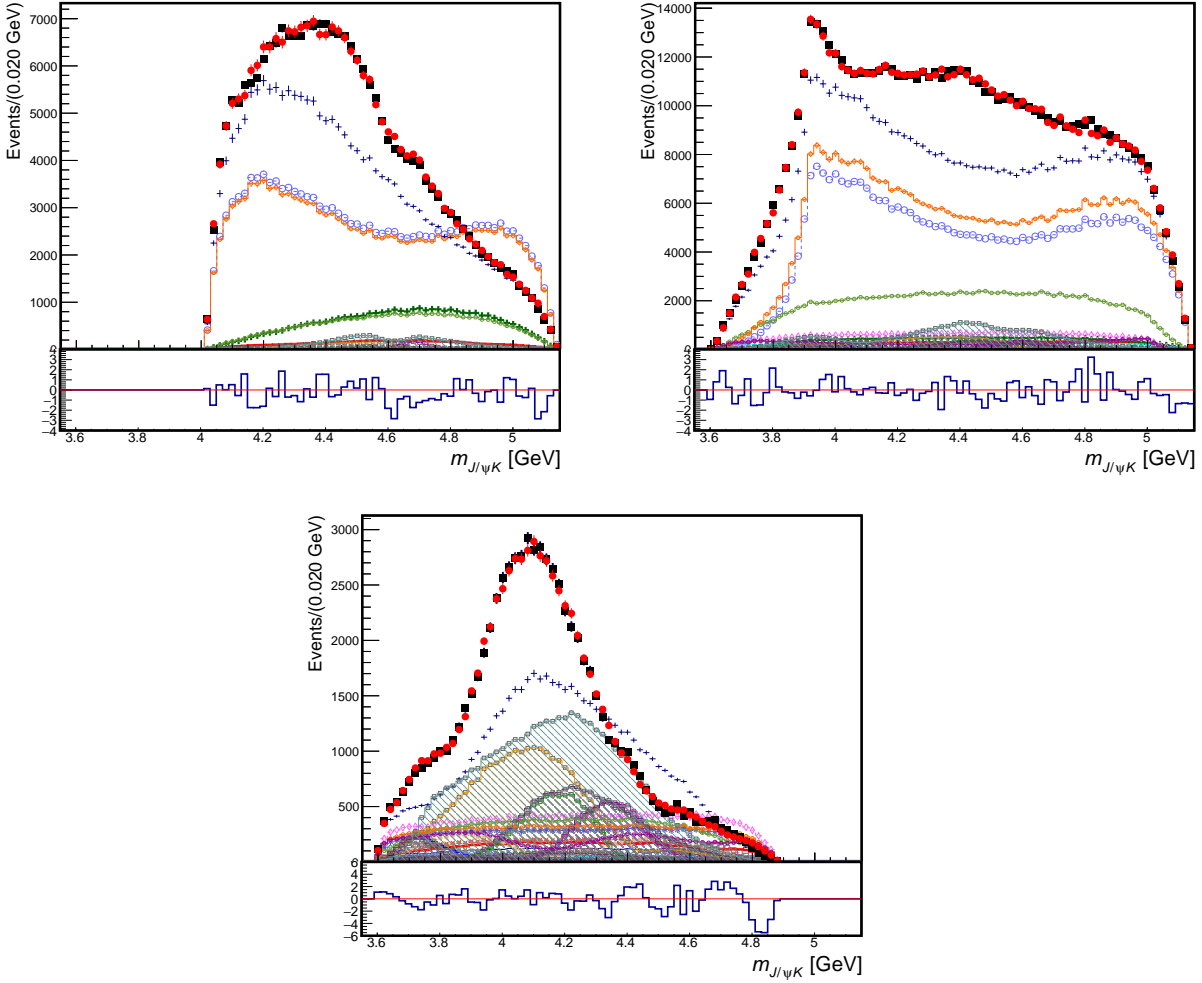


Figure 119: Projections of the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black points) and of the 4D amplitude fit (red) with $Z(3900)^-$, $Z(4200)^-$, $Z(4400)^-$, $Z(4430)^-$, and $Z(4800)^-$ as a $J^P = 1^+$ K-matrix, a broad $Z(4200)^-$ and $Z(3900)^-$ as a $J^P = 0^-$ K-matrix, $Z(4700)^-$ as a $J^P = 2^-$ resonance, $Z(4300)^-$ as a $J^P = 1^-$ resonance, and $Z_{cs}(4470)^-$ as a $J^P = 1^+$ $J/\psi K$ resonance. onto the $m_{J/\psi K}$ axis in different slices of $m_{K\pi^-}$: below $K^*(892)$ (top left), in between $K^*(892)$ and $K_2^*(1430)$ (top right), and above $\psi(2S)$ phase space (bottom).

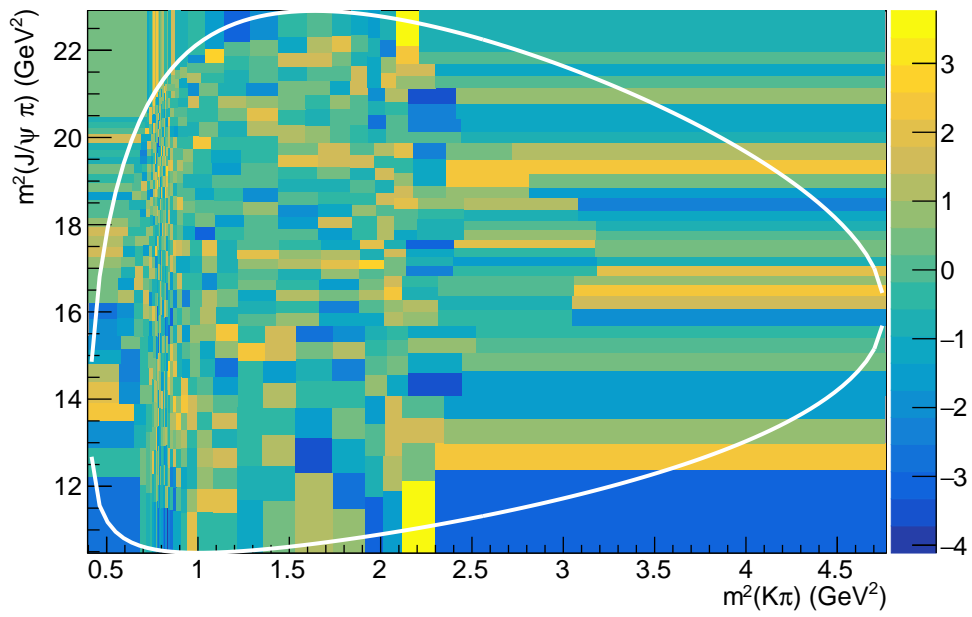


Figure 120: Dalitz plot with fit pulls in each bin used to calculate 2-D χ^2 . White line draws phase space boundary for the $B^0 \rightarrow J/\psi \pi^- K^+$ decay.

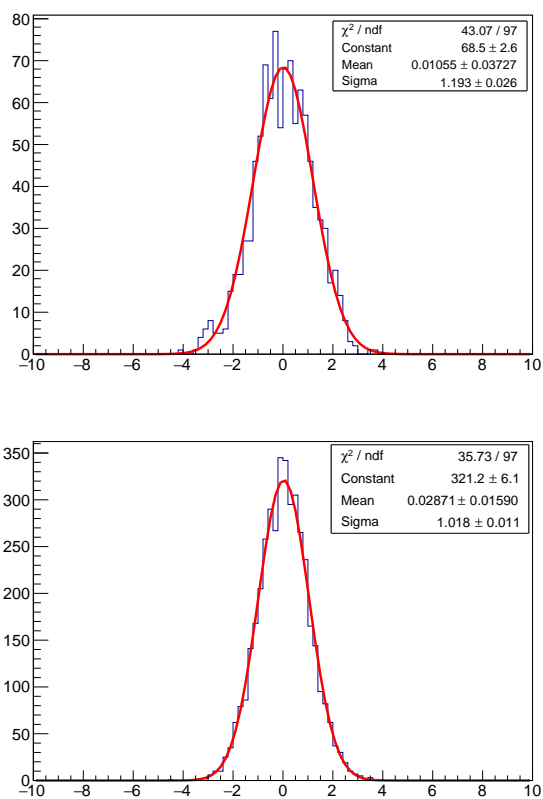


Figure 121: Distributions of the pulls from the 2-D and 4-D pulls from the multidimensional χ^2 s for the default $B^0 \rightarrow J/\psi\pi^-K^+$ model. The distribution is fit with a Gaussian (red line).

15.3 Models with two $Z_K^+ \rightarrow J/\psi K^+$

Fits with a second Z_K were performed on $B^0 \rightarrow J/\psi \pi^- K^+$, and the fit quality metrics for these are in Table 57. These fits, while yielding significant log-likelihood improvements, converge with warnings such as machine accuracy limits, and they have not positive definite error matrices. As such, the most reliable Z_K resonance model for the J/ψ channel is the one in the 1 Z_K fit model.

Table 57: Goodness-of-fit metrics for adding a second Z_K with the given J^P to the J/ψ channel.

Variable	1 Z_K only	$0^- Z_K$	$1^- Z_K$	$1^+ Z_K$	$2^- Z_K$	$2^+ Z_K$
nPar	117	121	121	123	123	121
$\log L$	-3056472	-3056522	-3056577	-3056532	-3056532	-3056617
$\Delta(-2 \log L)$	-	100	210	120	120	290
2-D χ^2 (1024 bins)	1457	1463	1432	1454	1438	1443
2-D $\Delta\chi^2$	-	-6	25	3	19	14
4-D χ^2 (4096 bins)	4248	4247	4223	4218	4237	4237
4-D $\Delta\chi^2$	-	0	25	30	10	11

15.4 Fits with fixed-shape $Z_{cs}(4000)$ and $Z_{cs}(4220)$

LHCb published in 2021 evidence of two exotic $J/\psi K$ states in the amplitude analysis of $B^+ \rightarrow J/\psi \phi K^+$ [27]. Fits to our data were attempted to include these states but with fixed masses and widths and free floating amplitudes. Both fits included $Z_{cs}(4000)$ as a $J^P = 1^+$ as it was determined to have such quantum numbers with high significance. Different quantum numbers are used for $Z_{cs}(4220)$ in the two fits, since the LHCb paper only preferred 1^+ over 1^- by 2σ while ruling out other quantum numbers. The fit qualities of both fits compared to the default 1 Z_K model in this analysis are presented in Table 58. Fits heavily prefer the single $Z_K(4470)$ model over the two resonances with quantum numbers, masses and widths taken from the $B^+ \rightarrow J/\psi \phi K^+$ data,

Table 58: Goodness-of-fit metrics comparing the 1 Z_K model presented in this analysis with the published LHCb Z_{cs} states with fixed masses and widths.

Variable	1 Z_K	$Z_{cs}(4000) + 1^+ Z_{cs}(4220)$	$Z_{cs}(4000) + 1^- Z_{cs}(4220)$
nPar	117	119	117
$\log L$	-3056472	-3056406	-3056381
$\Delta(-2 \log L)$	-	-132	-182
2-D χ^2 (1024 bins)	1457	1489	1493
2-D $\Delta\chi^2$	-	-32	-36
4-D χ^2 (4096 bins)	4248	4247	4268
4-D $\Delta\chi^2$	-	0	-21

16 Comparison of resonant models between $\psi(2S)$ and J/ψ channels

It is very interesting to compare the poles (Tables 59-60) and fit fractions (Table 61) from the best models of $\psi(2S)$ and J/ψ channels.

The fit fractions for the two channels show a similar pattern for the conventional $K^{*0} \rightarrow \pi^- K^+$ contributions, which dominate the $B^0 \rightarrow \psi(nS)\pi^- K^+$ decays. They are not expected to be exactly the same because of different phase-space factors but also because the internal structures of the $J/\psi(1S)$ and $\psi(2S)$ differ, with the $\psi(2S)$ state having an extra node in the radial wave function of the $c\bar{c}$ system and extending towards larger quark separation. Interplay between these differences and an internal structure of a companion hadron, both conventional or exotic, will make production couplings somewhat different. The $1^- K^+\pi^-$ wave (labeled P-wave, for one unit of angular momentum in its disintegration to K^+ and π^-) dominates. No angular momentum is required in a B^0 decay to produce it in a $\psi(nS)K^{*0}$ decay. The dominant contributor to this wave is the $K^*(892)^0$ state, which is a $1^3S_1 s\bar{d}$ state. The wave function similarity to J/ψ , which is $1^3S_1 c\bar{c}$ state, may be playing a role in enlarging its production rate in this channel. The production of S-wave (0^+) or D-wave (2^+) $K^+\pi^-$ requires one unit of angular momentum in B^0 decay. S-wave is more pronounced, perhaps since it involves no angular momentum in its decay. Higher spin states are progressively more suppressed by increasing angular momentum needed for production in the B^0 decay and subsequent disintegration. The phase space suppression also enters via increasing resonance mass. The K^* masses and widths which were allowed to float in the fits are fairly consistent with each other and consistent with the world average values as shown in Table 60, which is very reassuring. The biggest difference with the world average values is for the width of $K_0^*(800)$. However, this very broad state may well be a part of effective parametrization of non-resonant effects in $K\pi$ S-wave, rather than a true resonance. In fact, it does not belong to the $s\bar{d}$ quark model of kaon excitations. Non-resonant effects don't need to be universal in

various reactions.

The exotic hadron component in the $\psi(2S)$ channel is significantly more pronounced than in the J/ψ channel. Various models of exotic hadrons and their production mechanisms should take clues from this observation. Unlike for the conventional hadrons, there are also significant differences in the rate pattern (*i.e.* values of fit fractions) over various exotic waves.

In both channels, $1^+ \psi(nS)\pi^-$ wave is pronounced, which is not surprising since it requires no angular momentum in its disintegration to $\psi(nS)$ and π^- . Its production in B^0 decay requires one unit of angular momentum. The values of angular momentum in production and decay are reversed for the $0^- \psi(2S)\pi^-$ wave. Since it is produced in S -wave, its dominant exotic rate in the $J/\psi\pi^-$ channel is again not surprising. However, it is more difficult to understand why this wave does not register in the $\psi(2S)\pi^-$ channel.

The $1^+ \psi(nS)\pi^-$ wave has a rich pole structure in both channels. There are so many that it is hard to have full confidence in the exact pole decomposition. While one can find strong similarities in some pole positions, it would be difficult to claim that they are definitely the same within commonly accessible phase-space (which is smaller in $\psi(2S)\pi^-$ channel), since the widths often differ quite a bit. Also the number of the poles is different even within the same mass range. In fact, fitting the $\psi(2S)$ data with the masses and widths fixed to the J/ψ data model, or vice-versa, produces poor quality fits. For these reasons, we also present model independent fits to this wave in the following sections.

Among the $1^+ \psi(nS)\pi^-$ poles, we compared the dominant poles (*i.e.* with the largest fit fraction) to the $Z(4430)^- \rightarrow \psi(2S)\pi^-$ and $Z(4200)^- \rightarrow J/\psi\pi^-$ resonances claimed in the previously published analyses (see Table 59). The masses agree within 2.4σ and the widths agree within 1.3σ even without assigning systematic errors to our results, which is quite satisfactory given that in our models we use K-matrix formula with many interfering poles, while the previous analyses used a sum of Breit-Wigner resonances with just one or two $1^+ Z^-$ states. The results for $Z(4430)^- \rightarrow \psi(2S)\pi^-$ agree much better with the published LHCb analysis on a fraction of the present data [7] if the same fit model is used (see Sec. 11.2).

The $\psi(2S)$ and J/ψ data agree within the statistical errors on mass and width of the $Z(4200)^-$ pole. Averaging over the two channels we obtain 4116 ± 4 MeV for the pole mass and 174 ± 5 MeV for the width. This analysis should be taken as the confirmation of this state previously claimed in the $J/\psi\pi^-$ decay mode by Belle [16]. Our analysis also firmly establishes it in the $\psi(2S)\pi^-$ decay mode.

The 1^+ pole, that we associate with the $Z(4430)^-$ state, differs by 2.1σ in mass and by 4.5σ in width between the $\psi(2S)$ and J/ψ data. The statistical disagreement in width value is of concern. The previous attempt to include the $Z(4430)^-$ state in the amplitude analysis of $B^0 \rightarrow J/\psi\pi^-K^+$ by Belle [16], fixed its mass and width to the $B^0 \rightarrow \psi(2S)\pi^-K^+$ results, thus they did not perform a consistency check.

There is a $1^+ Z^-$ state near 4.8 GeV in our data, which has consistent mass but a width different by 3.7σ between the $\psi(2S)$ and J/ψ channels. It is also not clear why $1^+ Z(4400)^-$ and $Z(3900)^-$ states were required by the J/ψ but not the $\psi(2S)$ data. The $Z(3900)^- \rightarrow J/\psi\pi^-$ pole has a large width, 270 ± 12 MeV, which is inconsistent with the narrow 28 ± 3 MeV width for the $1^+ Z_c(3900)$ state [2] observed near the $D\bar{D}^*$ threshold by BESIII and BaBar. Therefore, it cannot be assumed to be the same state. The $\psi(nS)\pi^-$ poles with the other spin and parity are not consistent at all between the J/ψ and $\psi(2S)$ data.

Both J/ψ and $\psi(2S)$ channels require $\psi(nS)K^+$ components, which are much smaller than the $\psi(nS)\pi^-$ contributions. The poles are not the same between the two channels. None of them coincide with the $J/\psi K^+$ states observed in the amplitude analyses of $B^+ \rightarrow J/\psi K^+ \phi$ data by LHCb [27].

Already at this point, we can see that the only conclusion we can reach is that both channels have a very rich exotic hadron structure, far more complicated than previously believed based on the amplitude fits with less data.

Lack of universality in number and values of masses and widths of exotic hadrons between the J/ψ and $\psi(2S)$ channels likely point to incompleteness of our theoretical framework in

which data models are formulated. One known shortcoming of this and all previous amplitude fits to these channels is lack of accounting for coupled-channel effects. Coupled-channels are other decay modes accessible to the same hadronic systems. They exist for both K^{*0} and exotic hadrons. For example, K^{*0} resonances can decay also to $K^0\omega$, $K^0\eta$, $K^*(892)\pi$ etc. Z^- states can decay to various combinations of charmed meson-antimeson pairs. The best theoretical approaches integrate hadronic amplitudes in coupled-channels into one common description. In fact, K-matrix approach can handle it. Each decay mode contributes one dimension to the K-matrix. Since, in this analysis, we fit and describe only $\psi(nS)\pi^-K^+$ data, our K-matrix is not really a matrix but a one-dimensional formula. Resonant poles are expected to be the same in all the channels, while resonant width is not a fit parameter in multidimensional K-matrix. It instead emerges from a pattern of coupling constants to various decay channels. The coupled-channel amplitudes reflect themselves into other decay channels. To properly constrain such multi-dimensional coupled-channel models, one in principle needs to fit data from all the related decay channels. Unfortunately, this may be impossible to ever achieve for the exotic systems studied here, because of the experimental difficulties in reconstructing all relevant decay modes. An intermediate step would be to include all coupled-channel contributions from the expected dominant channels without constraining them to the data from the channels other than the one analyzed here, simply to evaluate their potential impact. This would be a natural next step for any future amplitude analyses of the data presented in this work.

Because we have no confidence that the present approach is a proper extraction of all the exotic resonant poles, there is also no motivation to study other systematic uncertainties to the masses and widths presented in this thesis. The results of our work are qualitative rather than quantitative. Nevertheless, they represent a significant progress over the previously published results, by revealing the complexity of exotic hadron effects in these channels. The diversity of various quantum-numbers and sheer number of resonant poles needed to describe the data point to rich spectroscopy underlying them. They are more likely a product of

direct color interactions of four-quarks rather than meson-meson forces in “molecular” models. Large widths of the states observed here also point in this direction.

Table 59: Masses and widths of the Z s in the default K-matrix models for both the $\psi(2S)$ and J/ψ channels. While we tried to associate poles between the $\psi(2S)$ and J/ψ channels with each other by their mass proximity, their widths often differ a lot. Therefore, these associations, marked by the same contribution label, should be taken with caution. Comparison to the world average values for $Z(4430)$ and $Z(4200)$ are also included [2] but should be taken with caution since our values are K-matrix poles, while the PDG values are from Breit-Wigner form.

State	J^P	$\psi(2S)$ Channel		J/ψ Channel		PDG [2]	
		M (GeV)	Γ (GeV)	M (GeV)	Γ (GeV)	M (GeV)	Γ (GeV)
Z(4430)	1^+	4.526 ± 0.012	0.213 ± 0.012	4.554 ± 0.005	0.152 ± 0.006	$4.478^{+0.015}_{-0.018}$	0.181 ± 0.031
Z(4200)	1^+	4.113 ± 0.009	0.169 ± 0.017	4.118 ± 0.005	0.174 ± 0.007	$4.196^{+0.035}_{-0.032}$	$0.370^{+0.100}_{-0.150}$
Z(4800)	1^+	4.818 ± 0.013	0.141 ± 0.044	4.798 ± 0.010	0.340 ± 0.032	-	-
Z(3900)	1^+	-	-	3.899 ± 0.008	0.270 ± 0.012	-	-
Z(4400)	1^+	-	-	4.400 ± 0.007	0.087 ± 0.009	-	-
$Z_2(4470)$	2^-	4.471 ± 0.006	0.106 ± 0.013	-	-	-	-
$Z_2(4700)$	2^-	-	-	4.749 ± 0.013	0.291 ± 0.029	-	-
$Z_0(4200)$	0^-	-	-	4.337 ± 0.007	0.651 ± 0.015	-	-
$Z_0(3900)$	0^-	-	-	3.880 ± 0.006	0.134 ± 0.008	-	-
$Z_1(4300)$	1^-	-	-	4.310 ± 0.002	0.908 ± 0.075	-	-
$Z_{K1}(4480)$	1^-	4.487 ± 0.013	0.155 ± 0.025	-	-	-	-
$Z_K(4520)$	1^+	4.521 ± 0.007	0.056 ± 0.012	-	-	-	-
$Z_K(4470)$	1^+	-	-	4.469 ± 0.013	0.204 ± 0.015	-	-

Table 60: Masses and widths of the K^* s in the default K-matrix models for both the $\psi(2S)$ and J/ψ channels compared to the world average values [2]. The comparison should be taken with some caution since our values are K-matrix poles, while the PDG values are not precisely defined.

Contribution	J^P	$\psi(2S)$ Channel		J/ψ Channel		PDG [2]	
		M (MeV)	Γ (MeV)	M (MeV)	Γ (MeV)	M (MeV)	Γ (MeV)
$K^*(892)$	1^-	889 ± 0.1	45 ± 0.3	890 ± 0.03	45 ± 0.07	895.5 ± 0.2	47.3 ± 0.5
$K_0^*(800)$	0^+	799 ± 42	319 ± 19	805 ± 5	306 ± 4	845 ± 17	468 ± 30
$K_2^*(1430)$	2^+	1415 ± 2	109 ± 3	1417 ± 0.3	98 ± 0.6	1432.4 ± 1.3	109 ± 5
$K_0^*(1430)$	0^+	-	-	1455 ± 2	218 ± 2	1425 ± 50	270 ± 80

Table 61: All resonance fit fractions in the default K-matrix models in this work for both the $\psi(2S)$ and J/ψ channels compared to the fit fractions from the previously published Breit-Wigner models.

Contribution	Fit fraction ($\psi(2S)$)		Fit fraction (J/ψ)	
	this work	Ref. [7]	this work	Ref. [16]
K^* Total S-wave	$(18.4 \pm 0.9)\%$		$(14.1 \pm 0.2)\%$	
K^* S-wave NR	-	$(0.3 \pm 0.8)\%$	$(0.7 \pm 0.04)\%$	
$K_0^*(800)$	$(4.7 \pm 1.6)\%$	$(3.2 \pm 2.2)\%$	$(3.4 \pm 0.1)\%$	$(7.1^{+0.7}_{-0.5})\%$
$K_0^*(1430)$	$(6.9 \pm 0.5)\%$	$(3.6 \pm 1.1)\%$	$(7.8 \pm 0.1)\%$	$(5.9^{+0.6}_{-0.4})\%$
$K_0^*(1950)$	$(1.0 \pm 0.4)\%$	-	$(0.5 \pm 0.02)\%$	$(0.1 \pm 0.1)\%$
K^* Total P-wave	$(58.6 \pm 1.1)\%$		$(65.7 \pm 0.2)\%$	
$K^*(892)$	$(55.2 \pm 1.1)\%$	$(59.1 \pm 0.9)\%$	$(61.6 \pm 0.2)\%$	$(69.0^{+0.6}_{-0.5})\%$
$K_1^*(1410)$	$(2.8 \pm 0.3)\%$	$(1.7 \pm 0.8)\%$	$(1.9 \pm 0.04)\%$	$(0.3^{+0.2}_{-0.1})\%$
$K_1^*(1680)$	$(1.0 \pm 0.4)\%$	$(4.0 \pm 1.5)\%$	$(1.1 \pm 0.04)\%$	$(0.3^{+0.2}_{-0.1})\%$
K^* Total D-wave	$(7.9 \pm 0.3)\%$		$(6.3 \pm 0.05)\%$	
$K_2^*(1430)$	$(7.1 \pm 0.3)\%$	$(3.6 \pm 1.1)\%$	$(6.1 \pm 0.04)\%$	$(6.3^{+0.3}_{-0.4})\%$
$K_2^*(1980)$	$(0.9 \pm 0.2)\%$	-	$(0.4 \pm 0.02)\%$	$(0.4 \pm 0.1)\%$
$K_3^*(1780)$	$(0.6 \pm 0.1)\%$	-	$(0.2 \pm 0.01)\%$	$(0.2 \pm 0.1)\%$
$K_4^*(2045)$	$(0.2 \pm 0.05)\%$	-	$(0.02 \pm 0.003)\%$	$(0.2 \pm 0.1)\%$
$K_5^*(2380)$	$(0.1 \pm 0.04)\%$	-	$(0.007 \pm 0.002)\%$	-
Z Total 1^+ wave	$(5.3 \pm 0.1)\%$	$(5.9^{+1.7}_{-3.4})\%$	$(1.7 \pm 0.006)\%$	
Z(4430)	$(2.5 \pm 0.05)\%$	$(5.9^{+1.7}_{-3.4})\%$	$(0.3 \pm 0.001)\%$	$(0.5^{+0.4}_{-0.1})\%$
Z(4200)	$(1.3 \pm 0.04)\%$	-	$(0.9 \pm 0.003)\%$	$(1.9^{+0.7}_{-0.5})\%$
Z(4800)	$(0.4 \pm 0.01)\%$	-	$(0.004 \pm 0.00001)\%$	-
Z(4400)	-	-	$(0.02 \pm 0.0001)\%$	-
Z(3900)	-	-	$(0.8 \pm 0.003)\%$	-
Z Total 0^- wave	-		$(3.1 \pm 0.01)\%$	-
$Z_0(4200)$	-	-	$(2.6 \pm 0.01)\%$	-
$Z_0(3900)$	-	-	$(0.7 \pm 0.002)\%$	-
$Z_2(4470)$	$(0.4 \pm 0.01)\%$	-	-	-
$Z_2(4700)$	-	-	$(0.14 \pm 0.005)\%$	-
$Z_1(4300)$	-	-	$(0.26 \pm 0.001)\%$	-
Z_K Total 1^- wave	$(0.4 \pm 0.01)\%$	-	-	-
$Z_{K1}(4480)$	$(0.4 \pm 0.01)\%$	-	-	-
Z_K Total 1^+ wave	$(0.1 \pm 0.002)\%$	-	$(0.046 \pm 0.0002)\%$	-
$Z_K(4520)$	$(0.1 \pm 0.002)\%$	-	-	-
$Z_K(4470)$	-	-	$(0.046 \pm 0.0002)\%$	-

17 Model Independent Fit of 1^+ Z Wave with $\psi(2S)$ Channel

In this section and the next, the K-matrix model of $J^P = 1^+$ Z wave, which exhibits large number of resonant poles, is replaced with a set of independent complex amplitudes in different mass bins to probe the resonant nature of the states without enforcing a particular mass shape. If there are significant coupled channel effects from the $D\bar{D}^*$ or $D^*\bar{D}^*$ modes in this wave, they are automatically absorbed into this model independent representation. For $B^0 \rightarrow \psi(2S)K\pi$, the 1^+ Z K-matrix is replaced with 14 independent complex amplitudes per helicity coupling, equally spaced in $m_{\psi(2S)\pi}$ phase space.

To set convention, Argand plots for each helicity are first generated using the default K-matrix model of 1^+ Z wave in the same mass binning, and the mass bin with the greatest magnitude is chosen to be fixed in the model independent representation to the K-matrix values, while still floating overall complex scale factors for the entire model independent sets of the helicity couplings. Fixing complex amplitudes for one of the mass bins allows the rest of helicity couplings to have less correlated fit errors as the overall scales/phases are factored out. Since the default model contains a separate $J^P = 2^-$ Z resonance, this Z is included in the fit along with the default K^* model and the two Z_K resonances, with the same masses and widths being free to float.

The fit quality metrics for this fit are presented in Table 62, and the default K-matrix model fit qualities are shown for comparison. Following Wilks' Theorem, the model independent fit is more significant in terms of log-likelihood change by 9.0σ , and it also outperforms the K-matrix model in the 2-D and 4-D χ^2 values. This can be interpreted as the K-matrix model of the 1^+ Z wave missing some significant aspects of our data.

The fit value of the masses and widths of K-matrix poles for other waves that floated in the fit are reported in Table 63, and they are fairly similar to the results from the all K-matrix model. The main differences are that the $K^{*0}(800)$ pole is at a higher mass position

and that the $J^P = 1^- Z_{K1}(4480)$ mass pole is slightly higher. Despite these differences, the fit fractions for the model independent $1^+ Z$ fit, as seen in Table 64, are comparable to those from the default all K-matrix fit model (Table 29). Much like the mass projections in the default all K-matrix fit model, the mass projections for the model independent $1^+ Z$ fit look very good and agree with data well (Figs. 122-126).

The Argand plots are presented in Figure 127. While the 2014 model independent fits both on the old and current dataset showed a single circular shape with counter-clockwise motion near the $Z(4430)$ pole in the Argand diagram, these plots over the entire $\psi(2S)\pi$ phase space in the decay show more complicated exotic activity. Shown alongside the Argand plots from this model independent fit are the Argand plots generated from the $1^+ Z$ K-matrix from Sec. 13.3. Qualitatively they are very similar with a lot of curling and counter-clockwise complex phase motion, thus the model independent parametrization confirms complicated resonant structure of $1^+ Z$ wave. There are some differences, for example the behavior of the helicity-0 Argand plot at low mass, in which the model independent fit curls in a loop-like manner while the K-matrix fit trends towards low magnitude. The helicity-1 Argand plots differ even more, especially at lower masses.

Table 62: Fit quality metrics comparing default K-matrix Z model fit with model independent Z fit on current dataset.

Metric	Default K-matrix Z	Model Independent Z
nPar	92	130
$\log L$	-102,615	-102,702
$\Delta(-2 \log L)$	-	174
2-D χ^2 (1024 bins)	1057	1008
2-D $\Delta\chi^2$	-	49
4-D χ^2 (4096 bins)	4365	4337
4-D $\Delta\chi^2$	-	28

Table 63: Masses and widths of the K^* s and $J^P = 2^-$ Z in the model independent fit for the $\psi(2S)$ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
$K^*(892)$	1^-	0.889 ± 0.0001	0.045 ± 0.0003
$K_0^*(800)$	0^+	0.842 ± 0.020	0.305 ± 0.019
$K_2^*(1430)$	2^+	1.416 ± 0.002	0.113 ± 0.003
$Z_2(4470)$	2^-	4.468 ± 0.006	0.089 ± 0.012
$Z_{K1}(4480)$	1^-	4.503 ± 0.015	0.164 ± 0.030
$Z_{K1}(4520)$	1^+	4.524 ± 0.008	0.057 ± 0.019

Table 64: All resonance fit fractions in the model independent Z fit for the $\psi(2S)$ channel.

Contribution	Fit fraction
K^* Total S-wave	18.0%
$K_0^*(800)$	5.4%
$K_0^*(1430)$	7.2%
$K_0^*(1950)$	0.8%
K^* Total P-wave	58.5%
$K^*(892)$	55.9%
$K_1^*(1410)$	2.6%
$K_1^*(1680)$	1.8%
K^* Total D-wave	7.5%
$K_2^*(1430)$	6.5%
$K_2^*(1980)$	1.0%
$K_3^*(1780)$	0.5%
$K_4^*(2045)$	0.2%
$K_5^*(2380)$	0.08%
Model Indep. 1^+ Z	6.7%
Z Total 2^- wave	0.4%
$Z_2(4470)$	0.4%
Z_K Total 1^- wave	0.3%
$Z_K(4480)$	0.3%
Z_K Total 1^+ wave	0.1%
$Z_K(4520)$	0.1%

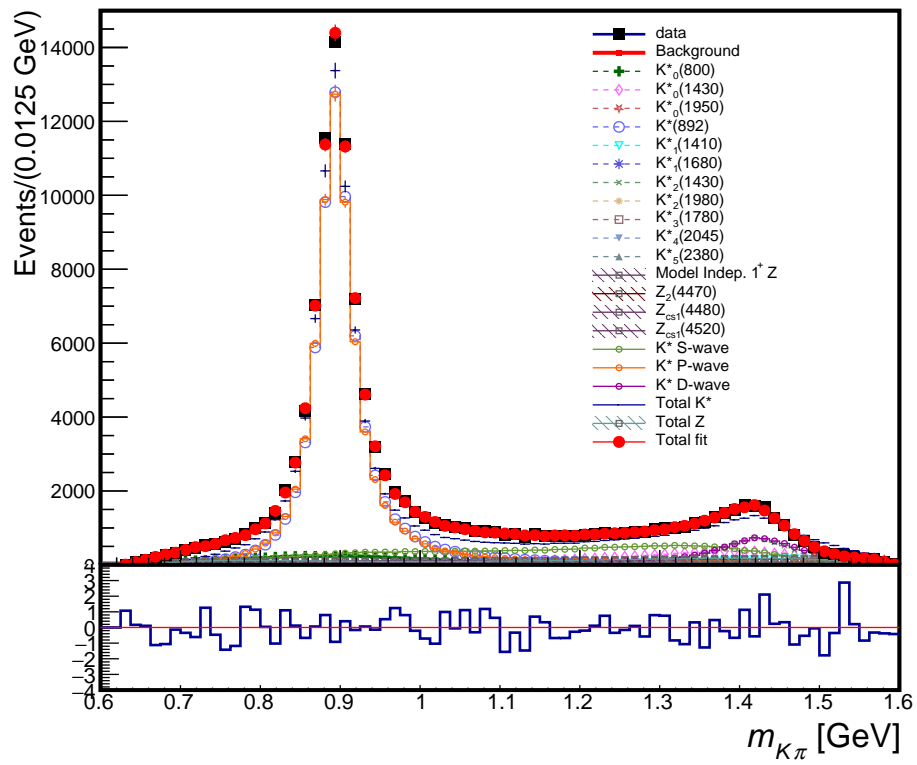


Figure 122: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) using the model independent approach.

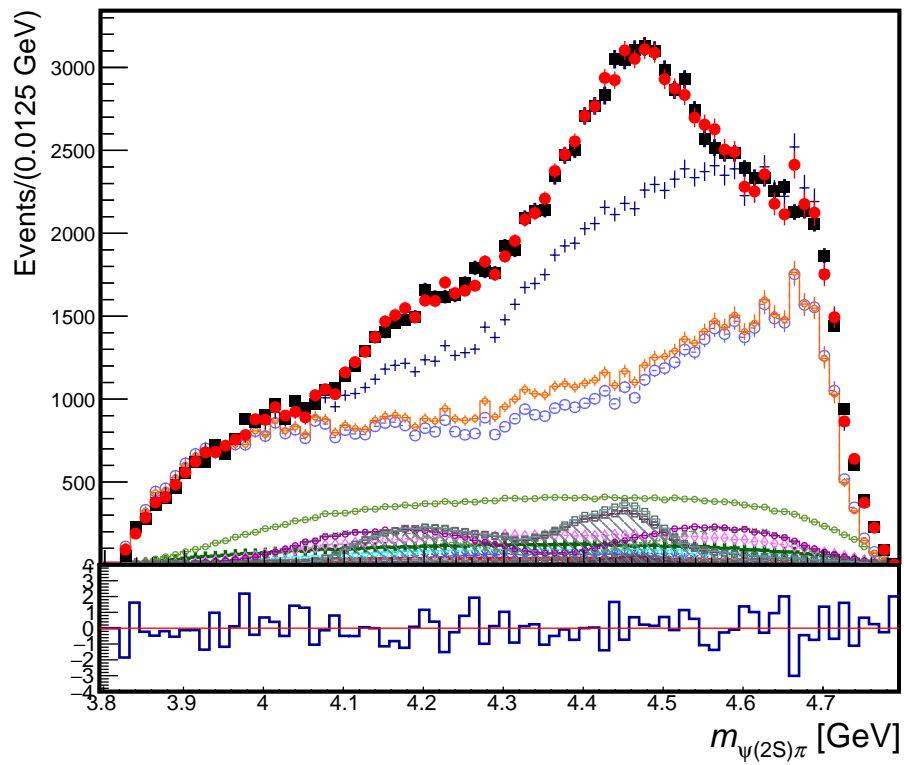


Figure 123: The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) using the model independent approach.

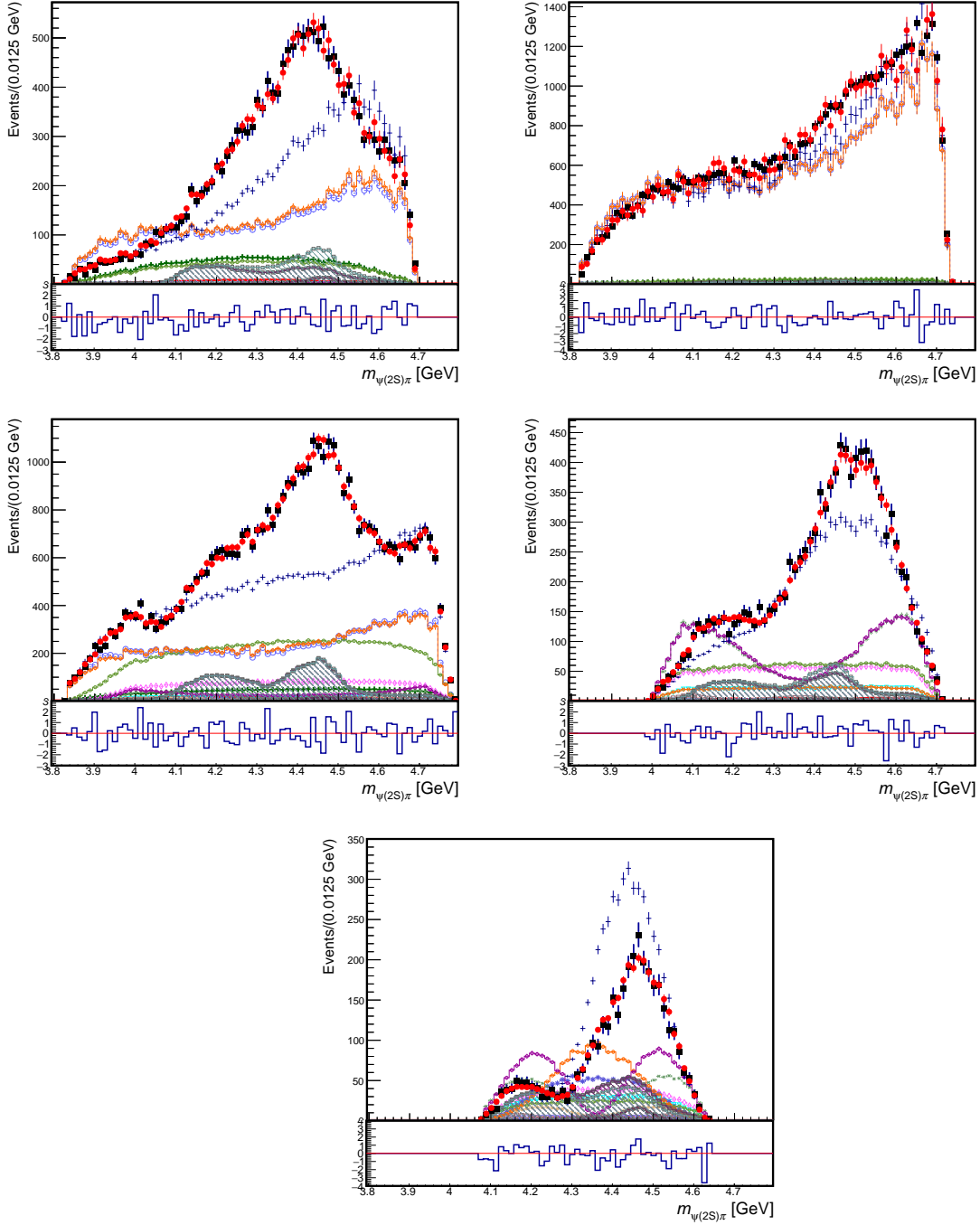


Figure 124: Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the model independent amplitude fit (red) onto the $m_{\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), at $K^*(892)$ (top right), in between $K^*(892)$ and $K_2^*(1430)$ (middle left), at $K_2^*(1430)$ (middle right), and above $K_2^*(1430)$ (bottom).

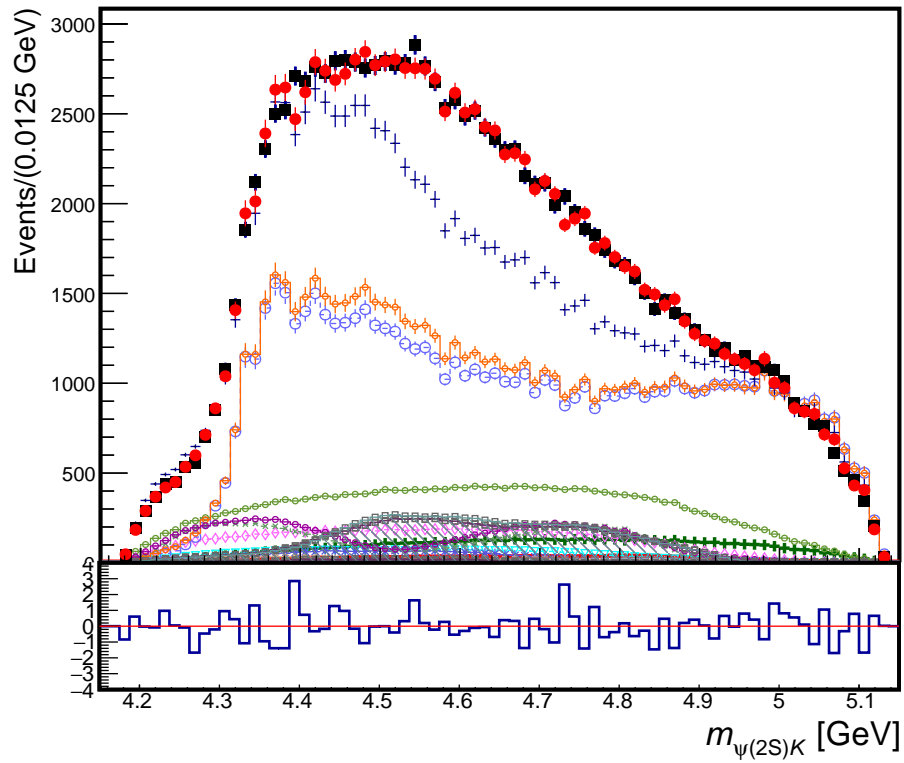


Figure 125: The distribution of $m_{\psi(2S)K}$ for the $B^0 \rightarrow \psi(2S)K\pi$ data (black) and the fit (red) using the model independent approach.

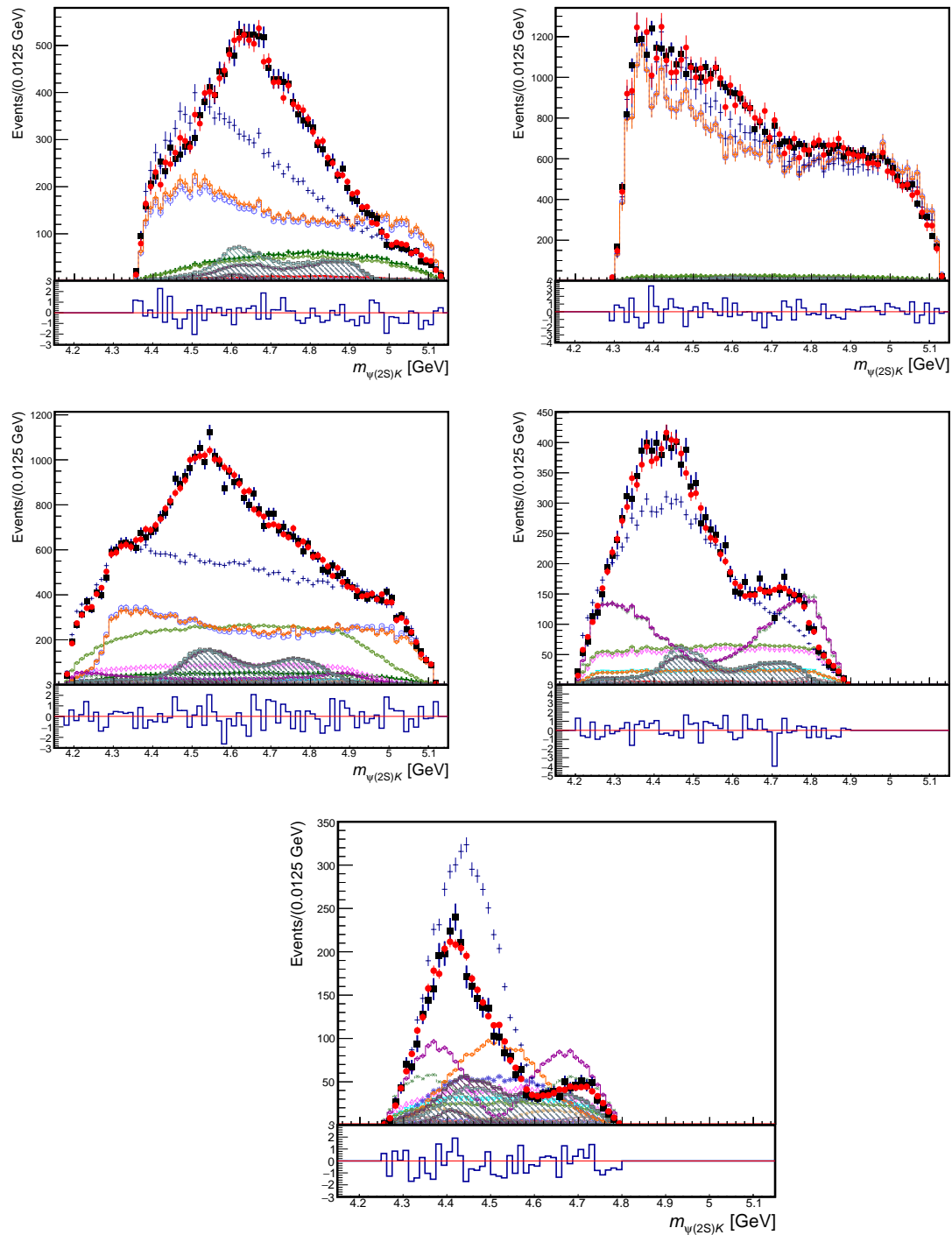


Figure 126: Projections of the $B^0 \rightarrow \psi(2S)K\pi$ data (black points) and of the model independent amplitude fit (red) onto the $m_{\psi K}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), at $K^*(892)$ (top right), in between $K^*(892)$ and $K_2^*(1430)$ (middle left), at $K_2^*(1430)$ (middle right), and above $K_2^*(1430)$ (bottom).

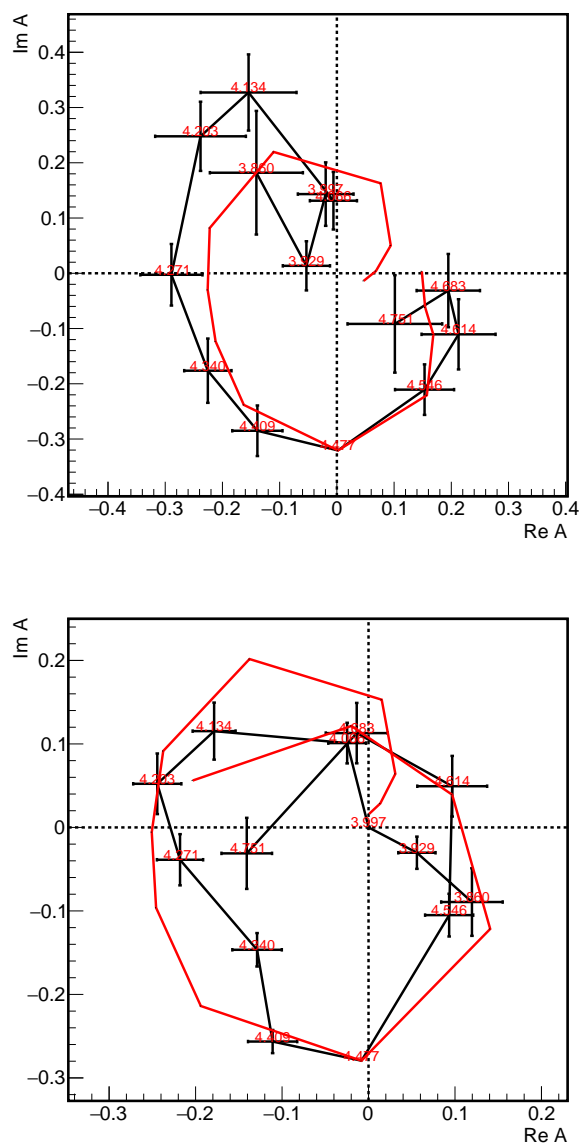


Figure 127: Argand plots for model independent $1^+ Z$ contribution (black points) for helicity-0 (top) and helicity- ± 1 (bottom). Each complex amplitude is labeled with its $m_{\psi\pi}$ bin center. The Argand plots generated from the default $1^+ Z$ K-matrix for both helicities are shown in red for comparison.

18 Model Independent Fit of 1^+ Z Wave with J/ψ Channel

Like the previous section with the $\psi(2S)$ channel, the $J^P = 1^+$ Z wave K-matrix is replaced with a set of independent complex amplitudes to probe the resonant nature of the states without enforcing a particular mass shape model. For $B^0 \rightarrow J/\psi K\pi$, 14 independent complex amplitudes per helicity coupling, equally spaced in $m_{J/\psi\pi}$ phase space, replace the 1^+ Z K-matrix parameterization. Similarly to $\psi(2S)$ channel, to set convention, Argand plots are generated using the default K-matrix model in the same binning, and the bin with the greatest magnitude is chosen to be fixed in the model independent fit to the K-matrix values, while overall complex scale factors for two helicities are allowed to float. The 9 Z J/ψ channel K-matrix model (Sec. 14.9) contains Z s of other quantum numbers besides $J^P = 1^+$, so these K-matrix Z poles are included in the fit along with the default K^* model. Model independent fits to the 1^+ Z wave with Z_K included in the model were unable to converge at various numbers of mass bins, thus $Z_K(4470)$ is omitted from the model independent 1^+ Z fit.

The fit quality metrics for this fit are presented in Table 65, and the default K-matrix model fit qualities are shown for comparison. The model independent fit is significant better in terms of log-likelihood (by 11.5σ using Wilks' Theorem), and the 2-D and 4-D χ^2 s also are considerably better compared to those from the all K-matrix fit. The final value of the masses and widths that floated in the fit are reported in Table 66, and for most poles, they are quite similar to the results from the 9 Z model fit. The main differences are both the $Z_0(4200)$ and $Z_1(4300)$ become quite a bit less broad compared to in the 9 Z model fit. The fit fractions for the model independent fit, as seen in Table 67, are pretty similar to those from the 9 Z fit model (Table 50). The mass projections for the model independent fit are similar to the mass projections from the 9 Z fit, but there is improvement in the $K\pi$ slices for the $m_{J/\psi\pi}$ mass projections, notably in the slice between $K^*(892)$ and $K^*(1430)$ and in the slice above $B \rightarrow \psi(2S)K\pi$ phase space (Figs. 128-130).

The Argand plots are presented in Figure 131, with Argand plots generated by the 1^+ K-matrix from the 9 Z fit for the J/ψ channel drawn in red. Much like with the Argand plots with the $\psi(2S)$ channel on the current dataset, these plots over the entire $J/\psi\pi$ phase space in the decay show a lot of complicated resonant activity. There are at least two loops in the Argand plot with counterclockwise motion. The Argand plots between the model independent fit and the K-matrix are mostly the same, they differ slightly at very low and very high $J/\psi\pi$ mass. The model independent fit prefers low $J/\psi\pi$ mass with low magnitude and high $J/\psi\pi$ mass with higher magnitude, and the opposite can be seen in the Argand plots for the K-matrix fits. Again, despite these differences at the ends of phase space, the multiple looping activity in both the K-matrix model fit and the model independent fit results suggest multiple resonances in the $J/\psi\pi$ distribution, matching the conclusion drawn from $\psi(2S)$ model independent fits.

Table 65: Fit quality metrics comparing 9 Z K-matrix model fit with model independent Z fit on current dataset.

Metric	K-matrix 1^+ Z Model	Model Independent 1^+ Z
nPar	111	137
$\log L$	-3,056,324	-3,056,457
$\Delta(-2\log L)$	-	266
2-D χ^2 (1024 bins)	1506	1403
2-D $\Delta\chi^2$	-	103
4-D χ^2 (4096 bins)	4295	4220
4-D $\Delta\chi^2$	-	75

Table 66: Masses and widths of the K^* s and Zs (not 1^+) in the model independent fit for the J/ψ channel.

Contribution	J^P	M (GeV)	Γ (GeV)
$K^*(892)$	1^-	0.890 ± 0.00002	0.045 ± 0.00004
$K_0^*(800)$	0^+	0.807 ± 0.0007	0.314 ± 0.001
$K_0^*(1430)$	0^+	1.455 ± 0.0003	0.212 ± 0.0008
$K_2^*(1430)$	2^+	1.417 ± 0.0001	0.098 ± 0.0003
Z(4200)	0^-	4.342 ± 0.001	0.589 ± 0.004
Z(4700)	2^-	4.795 ± 0.005	0.263 ± 0.007
Z(3900)	0^-	3.881 ± 0.001	0.149 ± 0.002
Z(4300)	1^-	4.345 ± 0.004	0.424 ± 0.017

Table 67: All resonance fit fractions in the model independent Z fit for the J/ψ channel.

Contribution	Fit fraction
K^* Total S-wave	13.0%
K^* S-wave NR	0.48%
$K_0^*(800)$	3.5%
$K_0^*(1430)$	6.9%
$K_0^*(1950)$	0.3%
K^* Total P-wave	66.0%
$K^*(892)$	62.0%
$K_1^*(1410)$	2.0%
$K_1^*(1680)$	1.0%
K^* Total D-wave	6.4%
$K_2^*(1430)$	6.2%
$K_2^*(1980)$	0.5%
$K_3^*(1780)$	0.3%
$K_4^*(2045)$	0.03%
$K_5^*(2380)$	0.007%
Model Indep. 1^+ Z	1.4%
Z Total 0^- wave	2.6%
Z ₀ (4200)	2.0%
Z ₀ (3900)	0.7%
Z Total 1^- wave	0.14%
Z ₁ (4300)	0.14%
Z Total 2^- wave	0.17%
Z ₂ (4700)	0.17%

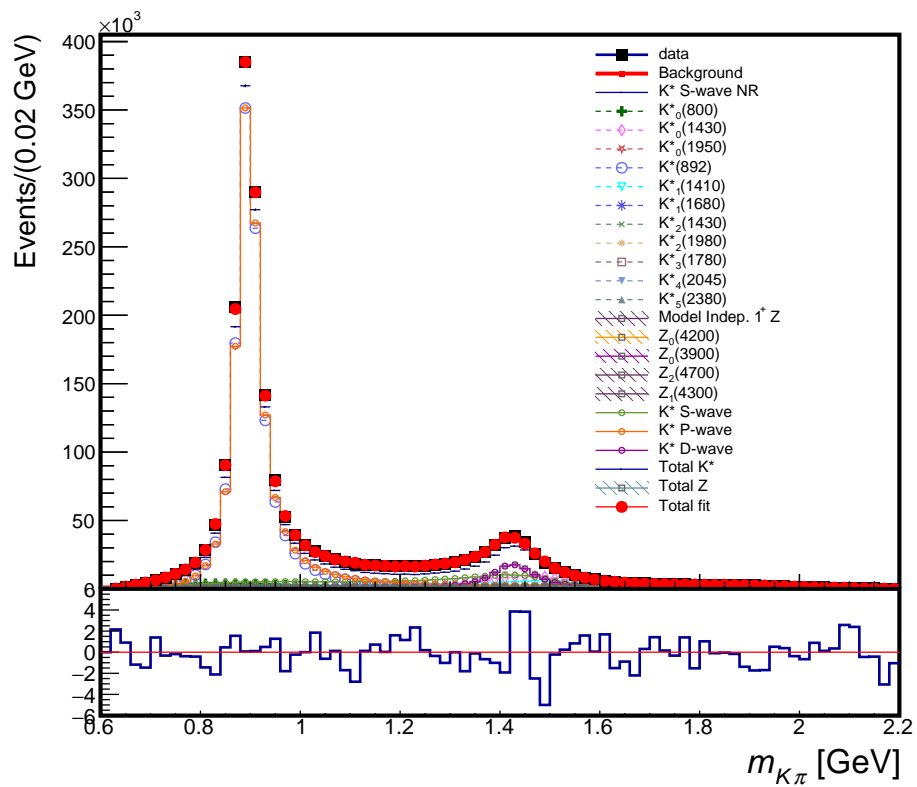


Figure 128: The distribution of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi K\pi$ data (black) and the fit (red) using the model independent approach.

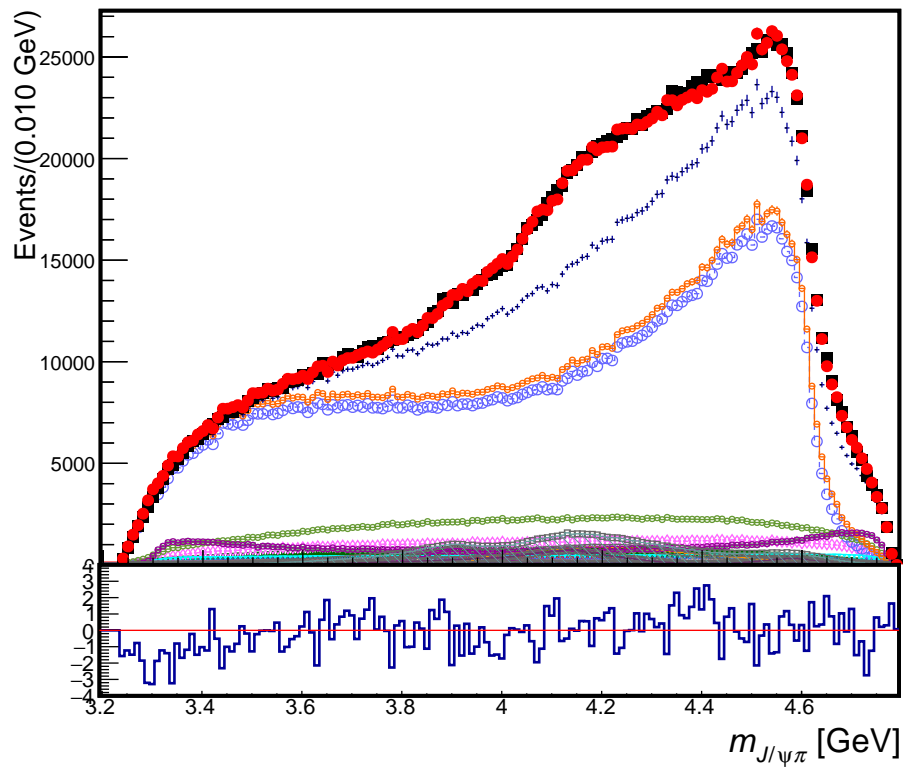


Figure 129: The distribution of $m_{\psi(2S)\pi}$ for the $B^0 \rightarrow J/\psi K\pi$ data (black) and the fit (red) using the model independent approach.

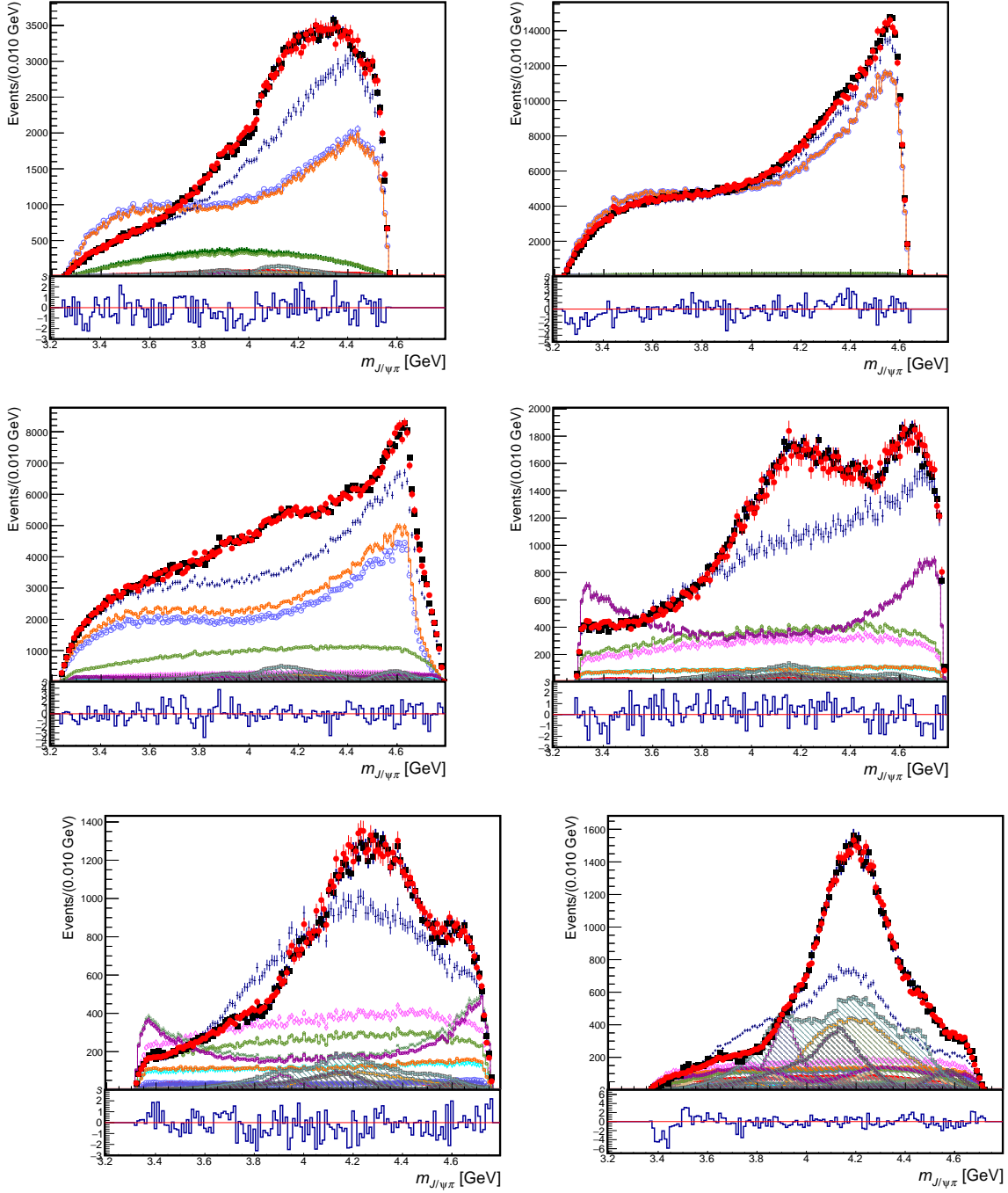


Figure 130: Projections of the $B^0 \rightarrow J/\psi K \pi$ data (black points) and of the model independent amplitude fit (red) onto the $m_{J/\psi\pi}$ axis in different slices of $m_{K\pi}$: below $K^*(892)$ (top left), at $K^*(892)$ (top right), in between $K^*(892)$ and $K_2^*(1430)$ (middle left), at $K_2^*(1430)$ (middle right), and above $K_2^*(1430)$ (bottom).

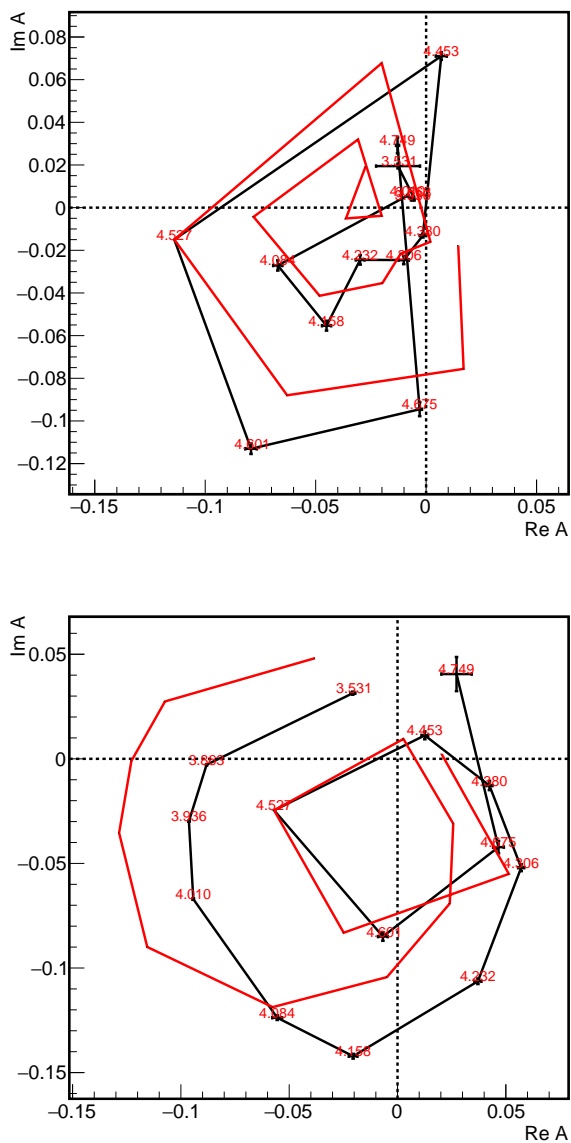


Figure 131: Argand plots for model independent 1^+ Z contribution for helicity-0 (top) and helicity- ± 1 (bottom). Each complex amplitude is labeled with its $m_{J/\psi\pi}$ bin center. Red graph represents 1^+ K-matrix from 9 Z fit drawn Argand-style in the same binning as used in the model independent fits.

19 Summary

Four-dimensional amplitude fits of $B^0 \rightarrow \psi(2S)\pi^-K^+$ and $B^0 \rightarrow J/\psi\pi^-K^+$ decays studied with the LHCb detector have been presented. The analyzed data samples are much larger than utilized in the previously published results from the analyses of that type: a factor of 76 increase over the Belle $B^0 \rightarrow J/\psi\pi^-K^+$ sample [16] and nearly a factor of 6 increase over the LHCb $B^0 \rightarrow \psi(2S)\pi^-K^+$ sample [7] (see Sec. 4).

The earlier results proved that even though both of these channels are dominated by production of conventional hadrons in the form of kaon excitations ($\bar{s}d$ states), $K^{*0} \rightarrow K^+\pi^-$, exotic tetraquark contributions, in the form of $Z^- \rightarrow \psi(nS)\pi^-$ ($c\bar{c}d\bar{u}$) states ($n = 1, 2$), were also necessary to explain the data. The dominant resonances were identified as 1^+ $Z(4430)^- \rightarrow \psi(2S)\pi^-$ [7, 12] and $Z(4200)^- \rightarrow J/\psi\pi^-$ [16], with some indication that both resonances were decaying to both $J/\psi(1S)\pi^-$ [16] and $\psi(2S)\pi^-$ [7].

The results presented here confirm that production of kaon excitations is dominant but grossly insufficient to describe the $B^0 \rightarrow J/\psi\pi^-K^+$ and $B^0 \rightarrow \psi(2S)\pi^-K^+$ decays (see Sec. 10) and that the dominant exotic hadron contributions are in the form of $Z^- \rightarrow \psi(nS)\pi^-$ states. The decay rate pattern is similar between our results and the previously published results, in spite of more advanced formalism used in our analysis (K-matrix) than in the previous approaches (sum over Breit-Wigner amplitudes) and with a much richer resonance model (see Table 61).

However, our results also show that exotic resonance spectrum is by far more abundant than previously disclosed. Instead of two $Z^- \rightarrow \psi(nS)\pi^-$ resonances utilized in the previous analyses, our fits point to at least four in the $\psi(2S)\pi^-$ channel and nine in the $J/\psi(1S)\pi^-$ channel with spin-parity numbers going beyond the 1^+ wave. Furthermore, two (one) $Z_K^+ \rightarrow \psi(nS)K^+$ resonances are needed for the best description of the $\psi(2S)$ ($J/\psi(1S)$) data. This is the first observation of such contributions in these decay modes.

None of these exotic resonances are narrow, with the dominant ones having widths of

the order of 10^2 MeV. Diversity of quantum numbers and width values point to direct color interactions of tetraquark configurations ($c\bar{c}d\bar{u}$ and $c\bar{c}u\bar{s}$) responsible for these states, rather than loose nuclear-type binding between meson-antimeson pairs for charmed mesons. In fact, our fits do not provide any evidence for production of the narrow Z_c^- and Z_{cs}^+ states observed by BESIII and Belle experiments at the $D\bar{D}^*$, $D^*\bar{D}^*$ and $D\bar{D}_s^* + D_s\bar{D}^*$ thresholds [2], which are good candidates for the latter (*i.e.* for "molecular states"). Qualitatively, our results lead to the same conclusions as previously obtained for the $J/\psi\phi$ and $J/\psi K^+$ states seen in amplitude analysis of the $B^+ \rightarrow J/\psi\phi K^+$ decays by the LHCb [27], in which a large number of relatively wide exotic states of different quantum numbers were required.

The largest number of exotic states is observed in the $J^P = 1^+ \psi(nS)\pi^-$ wave (see Table 59). Among them, a state matching $Z(4200)^-$ parameters from the previous analyses appears with consistent mass and width in both $\psi(2S)\pi^-$ and $J/\psi(1S)\pi^-$ decay modes. This is the first confirmation of this state previously claimed only by the Belle experiment in the $J/\psi\pi^-$ mode [16].

The dominant pole in the $1^+ \psi(2S)\pi^-$ wave appears in our data with the mass and width consistent with the previous Belle and LHCb results for the $Z(4430)^- \rightarrow \psi(2S)\pi^-$ state, with the differences attributable to the change in amplitude formalism from Breit-Wigner sum to K-matrix and the richer overall resonant model (see Sec. 16). While it also appears in the $J/\psi(1S)\pi^-$ decay mode with the consistent mass, the widths disagree by 4.5σ .

Both decay modes point to existence of previously unobserved $1^+ Z(4800)^-$ resonance. While the mass measurements are consistent between the two modes, the widths disagree by 3.7σ . The $J/\psi\pi^-$ data require two additional $1^+ Z^-$ states not supported by the $\psi(2S)\pi^-$ sample. One of them has the mass consistent with the $1^+ Z_c(3900)$ isospin-1 state observed by BESIII and Belle via a different production mechanism in various decay modes, including $J/\psi\pi^-$. However, the width observed in our data is a factor of 10 larger, thus it cannot be the same state.

Because of a large number of $1^+ Z^-$ poles extracted from our data, we have also performed

model independent fits to this wave, in which the K-matrix parameterization for it is replaced by a set of helicity zero and helicity ones amplitudes not constrained by any ansatz about their mass dependence. The obtained Argand plots confirm very rich resonant activity (see Secs. 17-18).

The spectrum of the other $\psi(nS)\pi^-$ and $\psi(nS)K^+$ states uncovered by our amplitude analyses do not agree on masses, widths, and J^P pattern between the $\psi(2S)$ and $J/\psi(1S)$ data. This is in contrast with a good consistency between the values of any free parameters in the $K^{*0} \rightarrow K^+\pi^-$ sector fit independently to the $\psi(2S)$ and $J/\psi(1S)$ data (see Table 60). Together, with the other inconsistencies mentioned above, this casts doubt into completeness of our modelling of the exotic amplitudes. This is the first time both $B^0 \rightarrow J/\psi\pi^-K^+$ and $B^0 \rightarrow \psi(2S)\pi^-K^+$ decays were simultaneously investigated with a sufficient statistics to perform such consistency tests. Therefore, we consider the observation of these inconsistencies to be an important outcome of our work, which calls for further improvements to modelling of relevant amplitudes. A natural next step would be to fit our data with models including coupled-channel effects, which have been neglected so far in all amplitude analyses performed for B meson decays to charmonium and light hadrons. This complicates amplitude parameterization by a lot and was beyond the scope of this thesis. See Sec. 16 for a more detailed discussion. The other possible source of the inconsistencies seen between the two data sets are genuine three-body decays, not proceeding via a sequence of two-body decays. Theoretical input is needed on possible parameterizations of such effects.

We approached computational limitations on complexity of the models, which we were successfully able to fit to our data (see *e.g.* Secs. 12.5, 14.10, 15.3 and 18). The LHCb detector is presently undergoing a major upgrade, with another major upgrade likely to follow after a decade of data taking. The upgrades will allow collection of much larger data samples than analyzed here. This should be helpful in overcoming these limitation.

Appendices

A PID Corrected Monte Carlo

We have investigated using particle-ID corrected Monte Carlo as input to our fitter.

The PID response in MC is corrected using `PIDCorr` in the `Meerkat` tool in the `PIDCalib` package, which transforms the PID variables in simulation based on calibration samples such that they are distributed like in data. The variables `ProbNNk`, `ProbNNpi`, and `PIDK` are transformed in this way for both the kaon and the pion, with the addition of `ProbNNp` being transformed for the pion only. After the variable transformation, the PID-corrected MC sample applies the same selection as in Sec. 4, using the same loose PID cuts.

With the PID-correct MC sample, we perform our default resonant Z model fit with $\psi(2S)$ and compare the results to the fit that uses a sample without MC PID correction. The results are presented in Tables 68, 69, and 70. Both fits performed well, but the multidimensional χ^2 s were worse in the PID corrected sample, while masses, widths, and fit fractions for each resonance were similar between the uncorrected and PID corrected MC samples within fit errors. Since we do not see improvement in our fits, we have elected to continue using an uncorrected PID sample.

Table 68: χ^2 s for default resonant Z model fits for uncorrected and PID-corrected $\psi(2S)$ MC.

Variable	Uncorrected	PID Corrected
nParameters	82	82
2-D χ^2 (1024 bins)	1078.14	1116.85
2-D χ^2 /nBins	1.05	1.09
4-D χ^2 (4096 bins)	4399.48	4447.74
2-D χ^2 /nBins	1.07	1.09

Table 69: Fit results for floating masses and widths in default resonant Z model fits for uncorrected and PID-corrected $\psi(2S)$ MC.

Contribution	J^P	Uncorrected		PID Corrected	
		M_0 (GeV)	Γ_0 (GeV)	M_0 (GeV)	Γ_0 (GeV)
K*(800)	0^+	0.823 ± 0.029	0.384 ± 0.024	0.812 ± 0.024	0.354 ± 0.021
K*(892)	1^-	0.889 ± 0.0001	0.045 ± 0.0003	0.889 ± 0.0001	0.045 ± 0.0003
K*(1430)	2^+	1.417 ± 0.001	0.104 ± 0.003	1.417 ± 0.001	0.104 ± 0.003
Z(4430)	1^+	4.535 ± 0.012	0.231 ± 0.014	4.535 ± 0.012	0.229 ± 0.015
Z(4200)	1^+	4.092 ± 0.009	0.152 ± 0.015	4.094 ± 0.010	0.163 ± 0.017
Z(4800)	1^+	4.820 ± 0.010	0.104 ± 0.050	4.822 ± 0.011	0.111 ± 0.049
Z(4470)	2^-	4.476 ± 0.008	0.130 ± 0.014	4.480 ± 0.009	0.136 ± 0.015

Table 70: Fit fractions for default resonant Z model fits for uncorrected and PID-corrected $\psi(2S)$ MC.

Contribution	Uncorrected	PID Corrected
$K_0^*(800)$	6.32	5.85
$K_0^*(1430)$	6.07	6.28
$K_0^*(1950)$	0.38	0.63
$K_1^*(892)$	56.57	56.43
$K_1^*(1410)$	2.77	2.91
$K_1^*(1680)$	2.03	1.95
$K_2^*(1430)$	6.56	6.59
$K_2^*(1980)$	1.07	1.04
$K_3^*(1780)$	0.46	0.43
$K_4^*(2045)$	0.21	0.23
$K_5^*(2380)$	0.10	0.10
Z(4430) 1^+	3.48	3.06
Z(4200) 1^+	0.34	0.42
Z(4800) 1^+	0.43	0.44
Z(4460) 2^-	0.54	0.52
K* S-wave	17.14	17.70
K* P-wave	59.42	59.44
K* D-wave	7.73	7.73
Total K*	87.59	88.17
Z 1^+	4.89	4.56
Total Z	5.41	5.07

B Fit displays for Legendre moments of $\cos \theta_{K^*}$ helicity angle

An interesting insight into $K\pi$ orbital-momentum structures in the data and the fits is provided by calculating moments of $\cos \theta_{K^*}$ distribution and plotting them in a function of $m_{K\pi}$. Such moments were the starting point for the semi-model-independent analysis of the decays considered here, as discussed in the introduction (Sec. 3). Unnormalized moment of l^{th} degree is defined as:

$$\langle P_l^U \rangle = \sum_{i=1}^{N_{\text{data}}} \frac{1}{\epsilon_i} P_l(\cos \theta_{K^* i}) \quad (32)$$

where $P_l(x)$ is a Legendre polynomial of order l , and ϵ_i is the efficiency for given event. We use parameterized efficiency to calculate ϵ_i as described in Sec. 6.3. We calculate moments for both the data, and the phase-space MC reweighted by the matrix element squared obtained from the default fit model.

A K^{*0} resonance of spin J can only contribute its amplitude in quadrature up to $l_{\text{max}} = 2J$. If it interferes with a resonance of spin J' , this will show up in moments up to $l_{\text{max}} = J + J'$. Only $J \leq 2$ K^{*0} resonances have large contributions to our data. For these resonances, we expect (see Eqs. 27–30 of Ref. [13]):

$\langle P_1^U \rangle$ to contain $S - P$ and $P - D$ interference terms;

$\langle P_2^U \rangle$ to contain $|P|^2$, $|D|^2$ and $S - D$ interference terms;

$\langle P_3^U \rangle$ to contain $P - D$ interference terms;

$\langle P_4^U \rangle$ to contain $|D|^2$ term.

Non-zero higher moments can be induced by higher spins of K^{*0} resonances, or exotic hadrons, or interplay of both. Reflections of exotic Z^- and Z_K^- resonances can contribute to any order of moments of $K^+\pi^-$ helicity angle. Vice-versa, reflections of K^{*0} resonances spread to any

order of $\psi\pi^-$ or ψK^+ helicity angles. Since the K^{*0} resonances dominate our data, inspecting moments of the latter helicity angles is less interesting. Regardless, comparison between the moments of the data and of the fitted model is a good visual way to access fit quality on the Dalitz plane. We present such comparisons for fits without (K^{*0} -only models as described in Sec. 10) and with exotic hadron contributions (best models as described in Sec. 13.3 for the $\psi(2S)$ and Sec. 15.2 for the J/ψ data), which illustrate which parts of the K^{*0} angular momentum structures in the data require exotic hadron contributions.

Figs. 132-133 show $\langle P_1^U \rangle$, with the structures driven mostly by interference of the S-wave with $K^*(892)$, $K_1^*(1410)$ and $K_1^*(1680)$ (P -wave) and $K_2^*(1430)$ (D -wave). Even though model without exotics reproduces the main features of the data, inspection of the pull distributions reveal significant drawbacks of the K^* -only models, in the region at and below the $K^*(892)$ resonance, and above 2 GeV for the J/ψ data. The exotic hadron contributions mitigate these shortcomings (notice the change in the range of the fit pull distributions between the top and bottom plots).

Figs. 134-135 show $\langle P_2^U \rangle$, which is dominated by P -wave amplitudes squared. $K_2^*(1430)$ (D -wave) can also contribute on its own and interfering with S -wave, but produces no peaking structure, except for the wiggle when crossing $K_2^*(1430)$ resonance. Effect of the exotic hadron contribution is limited in this moment. In contrast, they play a much bigger role in describing $\langle P_3^U \rangle$ moment, especially below the $K^*(892)$ region (Figs. 136-137). Interference of $K_2^*(1430)$ with the higher mass P -wave resonances explain the peaking structure near this resonance pole, but its tail does not reach below the $K^*(892)$ peak. The S -wave, which is significant at this mass range, cannot contribute to this moment. Therefore, significant value of this moment below the $K^*(892)$ is entirely due to the presence of exotic resonances (mostly Z^- states, which have higher rate than Z_K^+ states).

The exotic hadron contribution improve description of of the $\langle P_4^U \rangle$ moment in a wide range between the $K^*(892)$ and $K_2^*(1430)$ resonances, and at the high masses for the J/ψ data (Figs. 138-139).

The $\langle P_5^U \rangle$ and higher moments are driven by interference of the $K_2(1430)$ resonance with tails of the higher spin states. While impact of the exotic hadrons on this moment is minor for the $\psi(2S)$ data, they often significantly improve description of the J/ψ data (Figs. 140-151) at and below this resonances.

Overall agreement between the data and the fits which include the exotic hadrons is very good, especially for the $\psi(2S)$ data. The much bigger sample of the J/ψ data, which also spread to a bigger phase-space, is harder to describe within the small statistical errors. The statistical errors in the Monte Carlo sample for the J/ψ channel are bigger than in the data, which may be one of the limiting factors in construction of even more complex models of the J/ψ data. Overcoming such limitations in the future will not be trivial because of the computing time limitation.

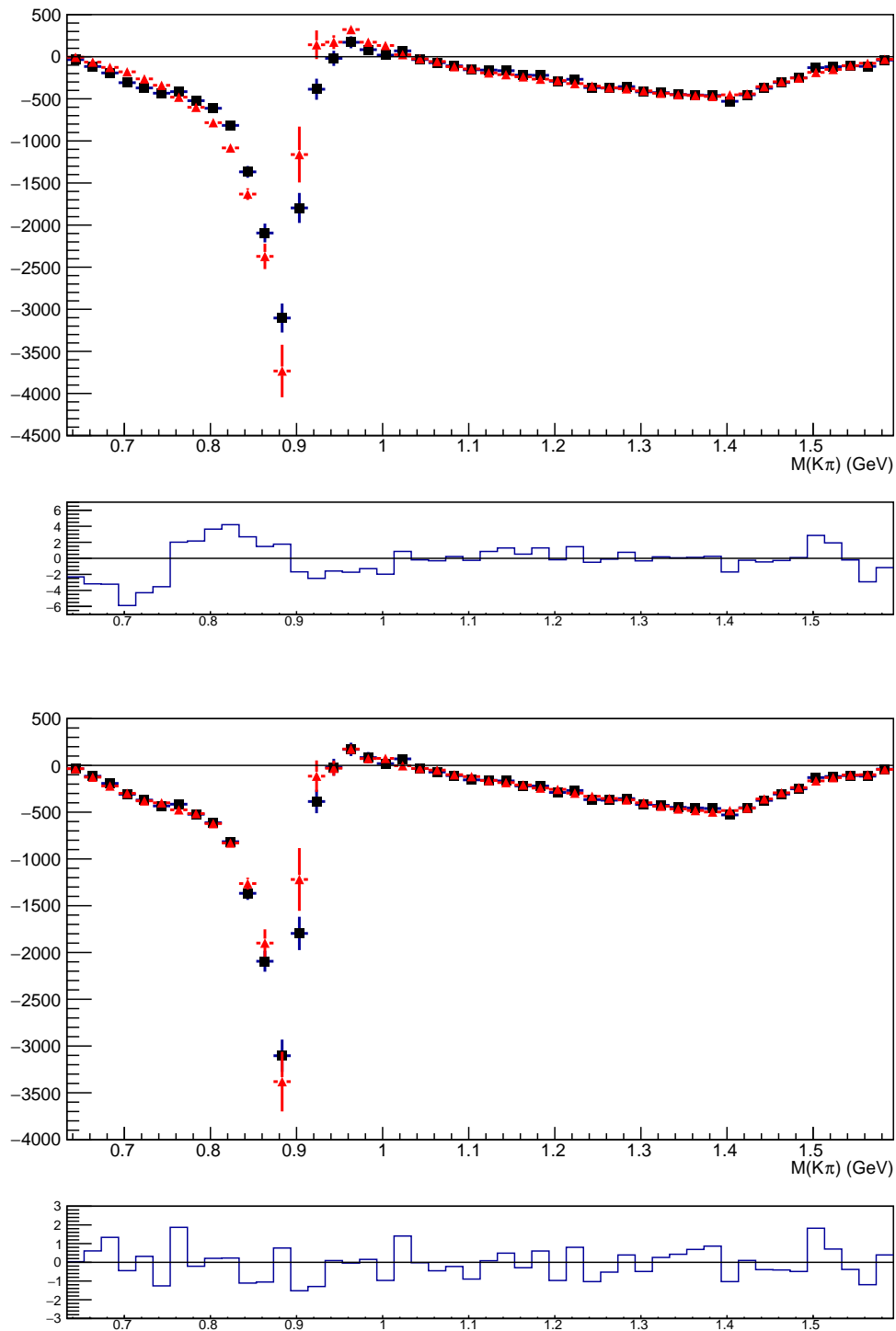


Figure 132: Unnormalized Legendre polynomial moment $\langle P_1^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 4 Z s and 2 Z_{K^*} s (bottom). The plots below the moment displays shows the fit pull distributions.

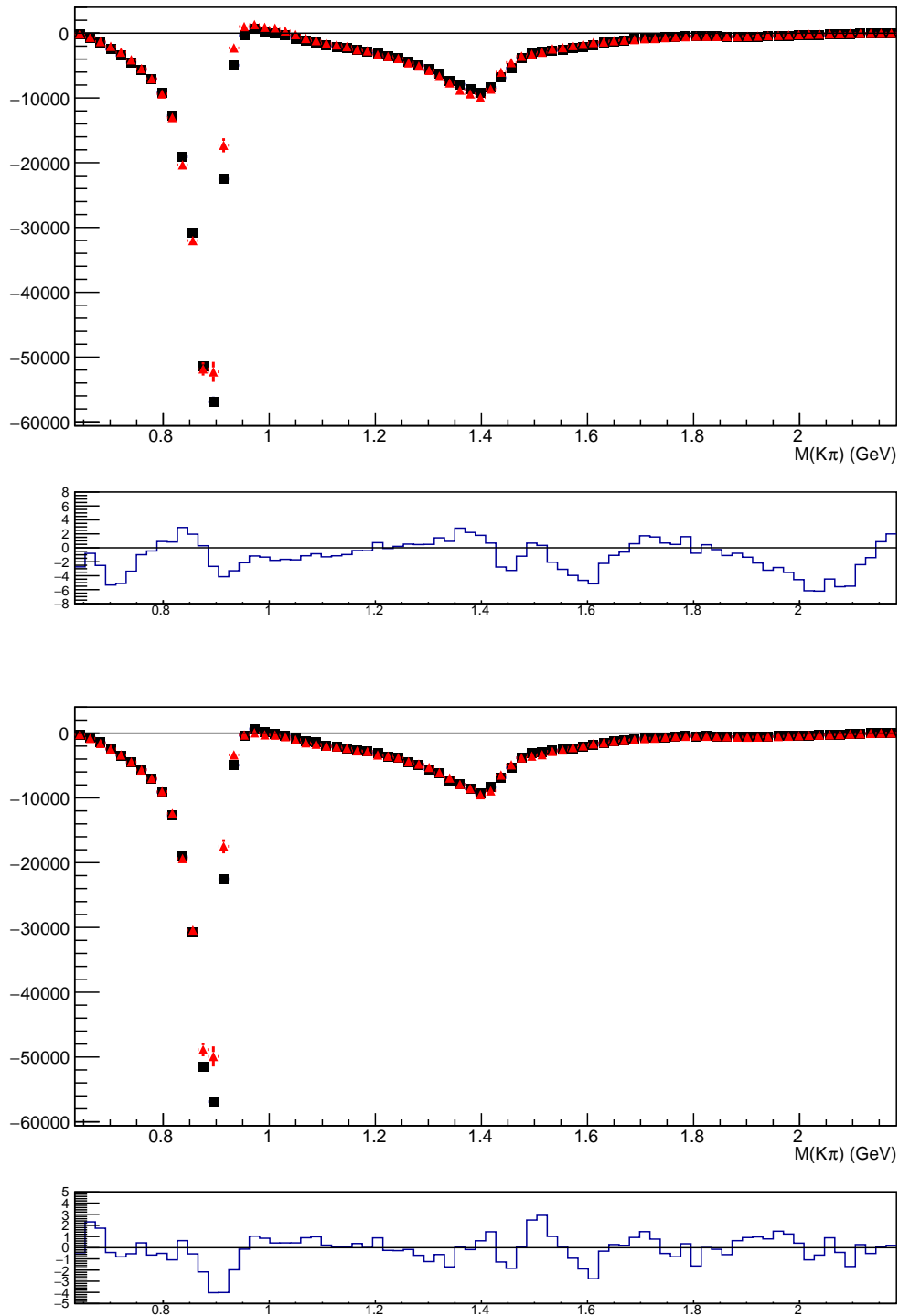


Figure 133: Unnormalized Legendre polynomial moment $\langle P_1^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 9 Z s and 1 Z_{K^*} s (bottom). The plots below the moment displays shows the fit pull distributions.

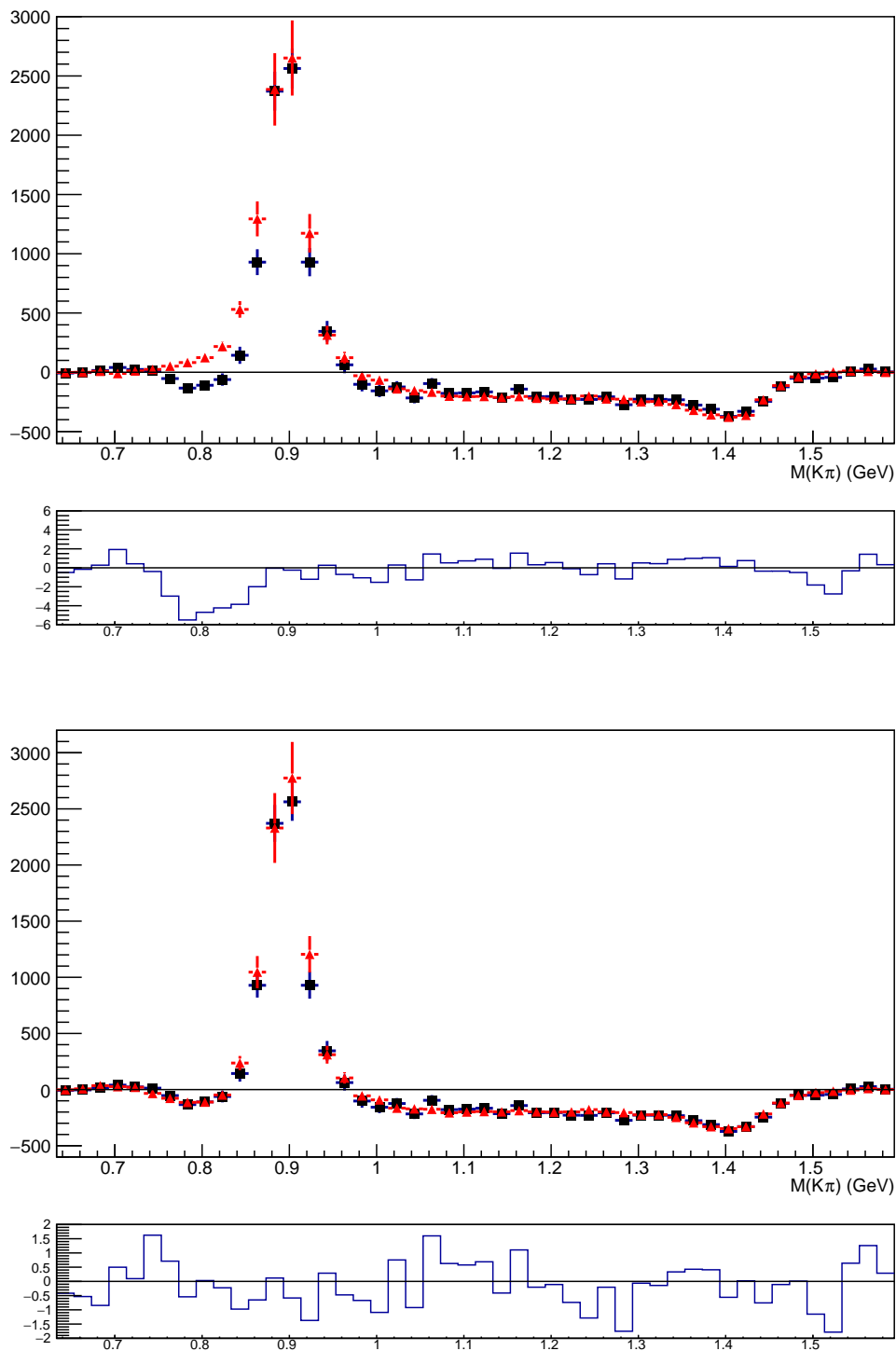


Figure 134: Unnormalized Legendre polynomial moment $\langle P_2^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 4 Z s and 2 Z_{K^*} s (bottom). The plots below the moment displays shows the fit pull distributions.

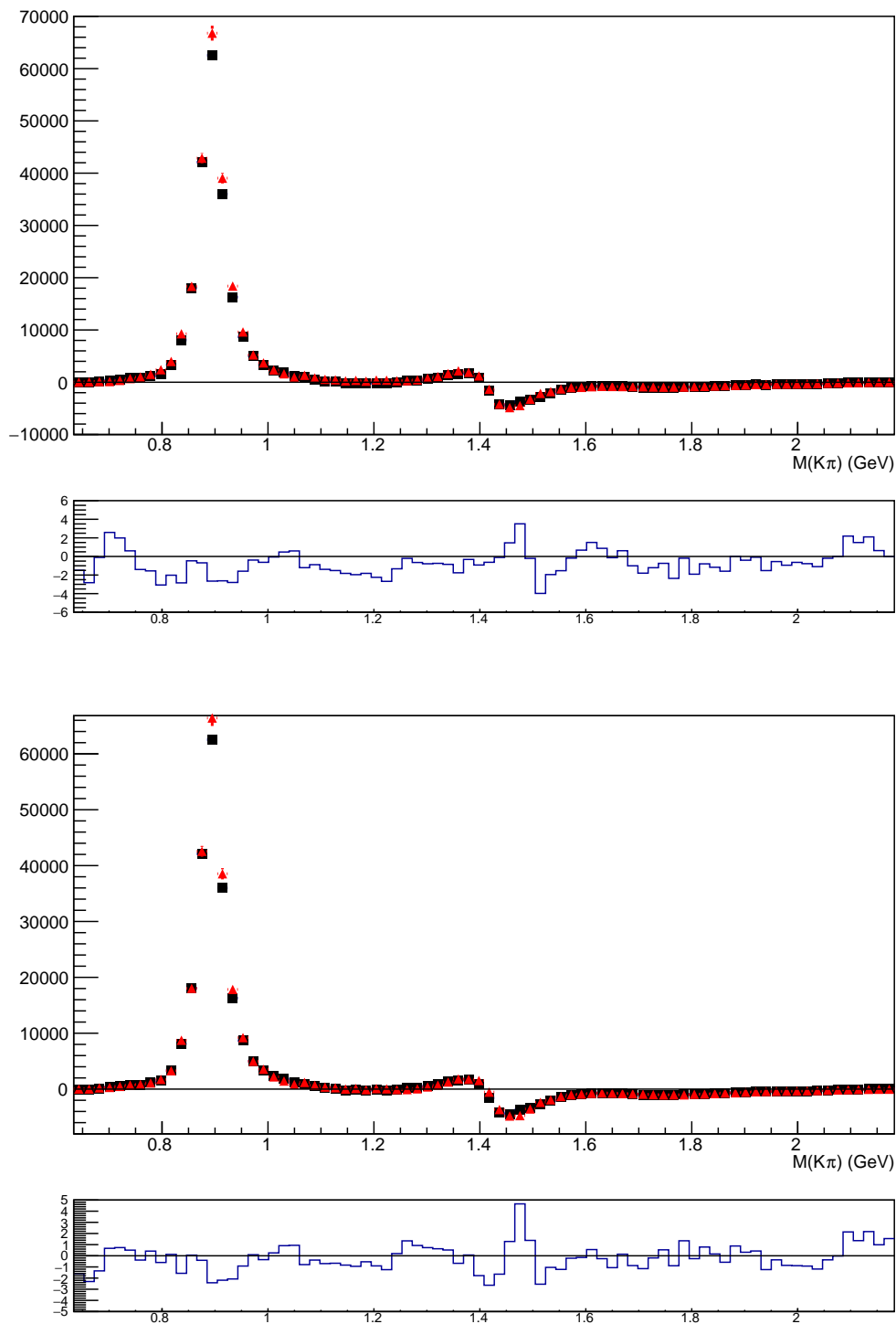


Figure 135: Unnormalized Legendre polynomial moment $\langle P_2^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 9 Z s and 1 Z_K (bottom). The plots below the moment displays shows the fit pull distributions.

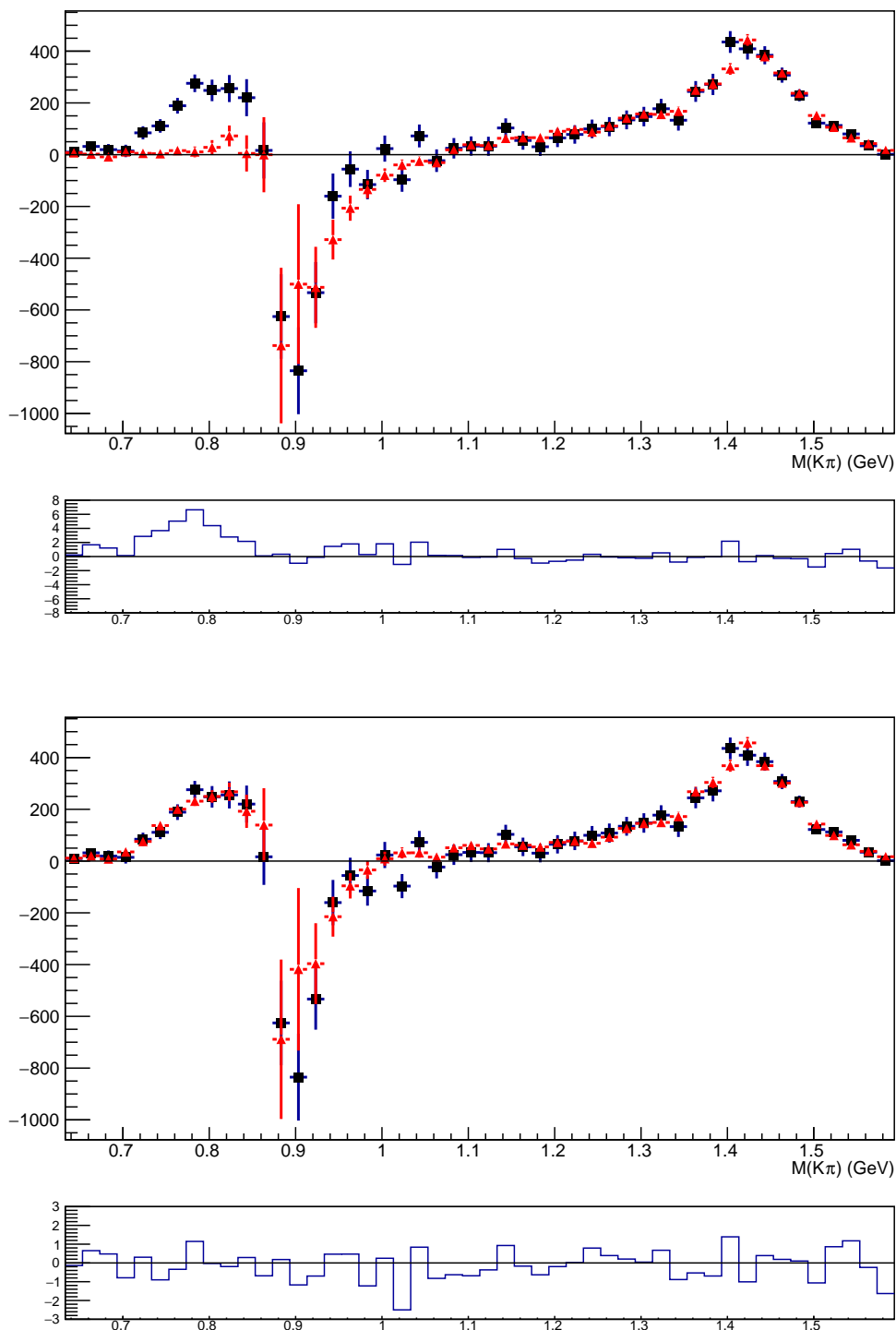


Figure 136: Unnormalized Legendre polynomial moment $\langle P_3^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 4 Z s and 2 Z_{K^*} s (bottom). The plots below the moment displays shows the fit pull distributions.

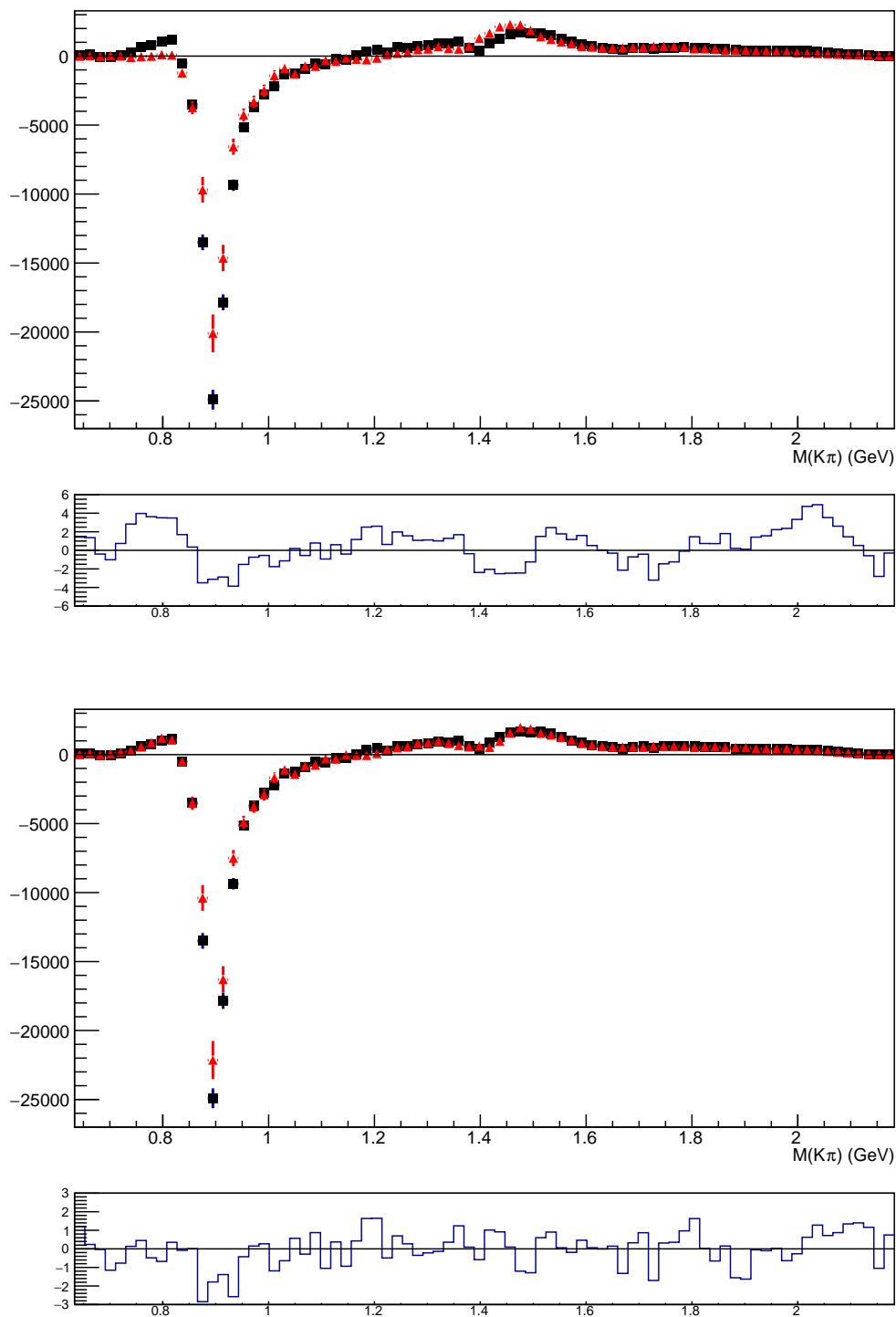


Figure 137: Unnormalized Legendre polynomial moment $\langle P_3^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 9 Z s and 1 Z_K (bottom). The plots below the moment displays shows the fit pull distributions.

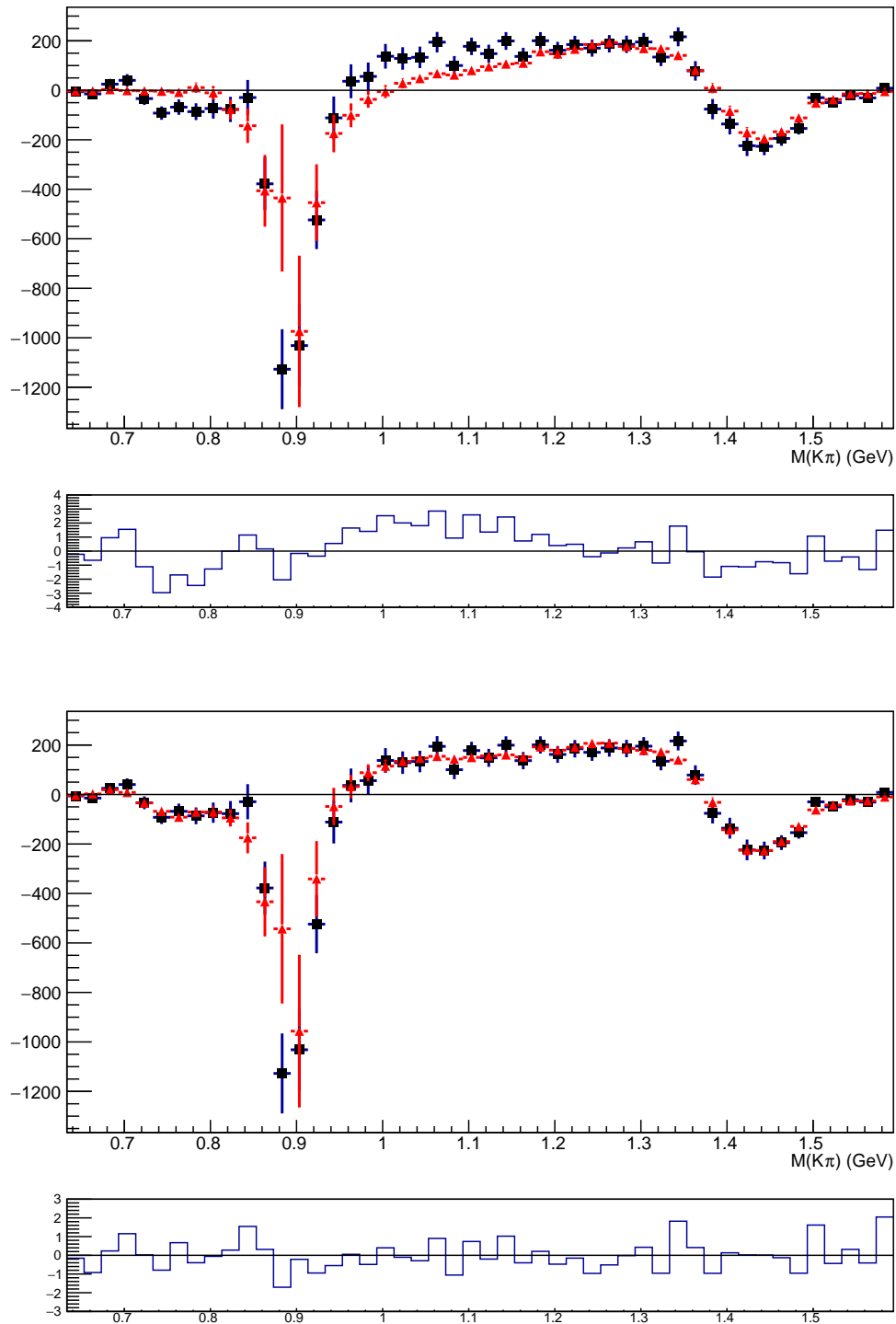


Figure 138: Unnormalized Legendre polynomial moment $\langle P_4^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 4 Z s and 2 Z_{K^*} s (bottom). The plots below the moment displays shows the fit pull distributions.

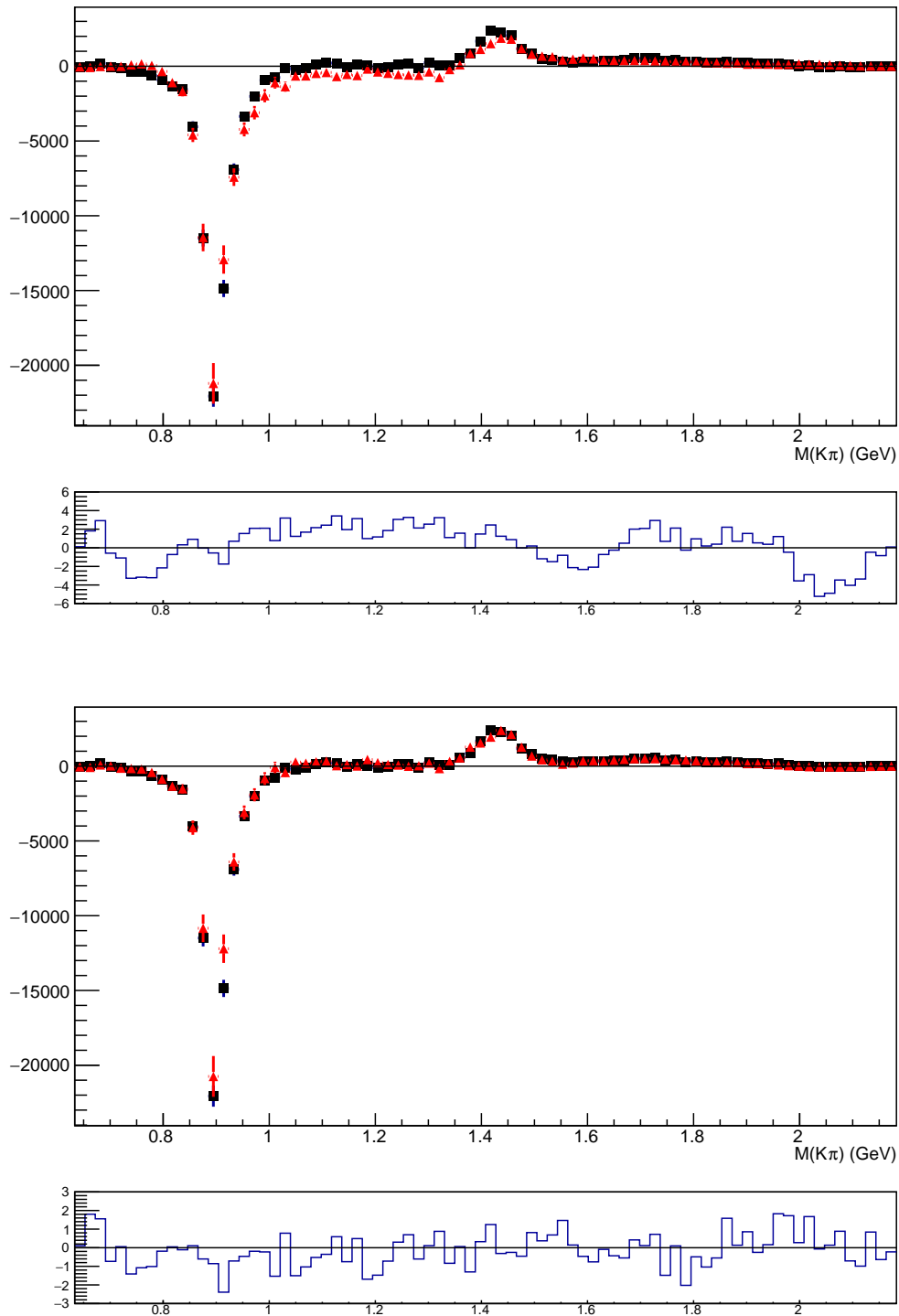


Figure 139: Unnormalized Legendre polynomial moment $\langle P_4^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 9 Z s and 1 Z_K (bottom). The plots below the moment displays shows the fit pull distributions.

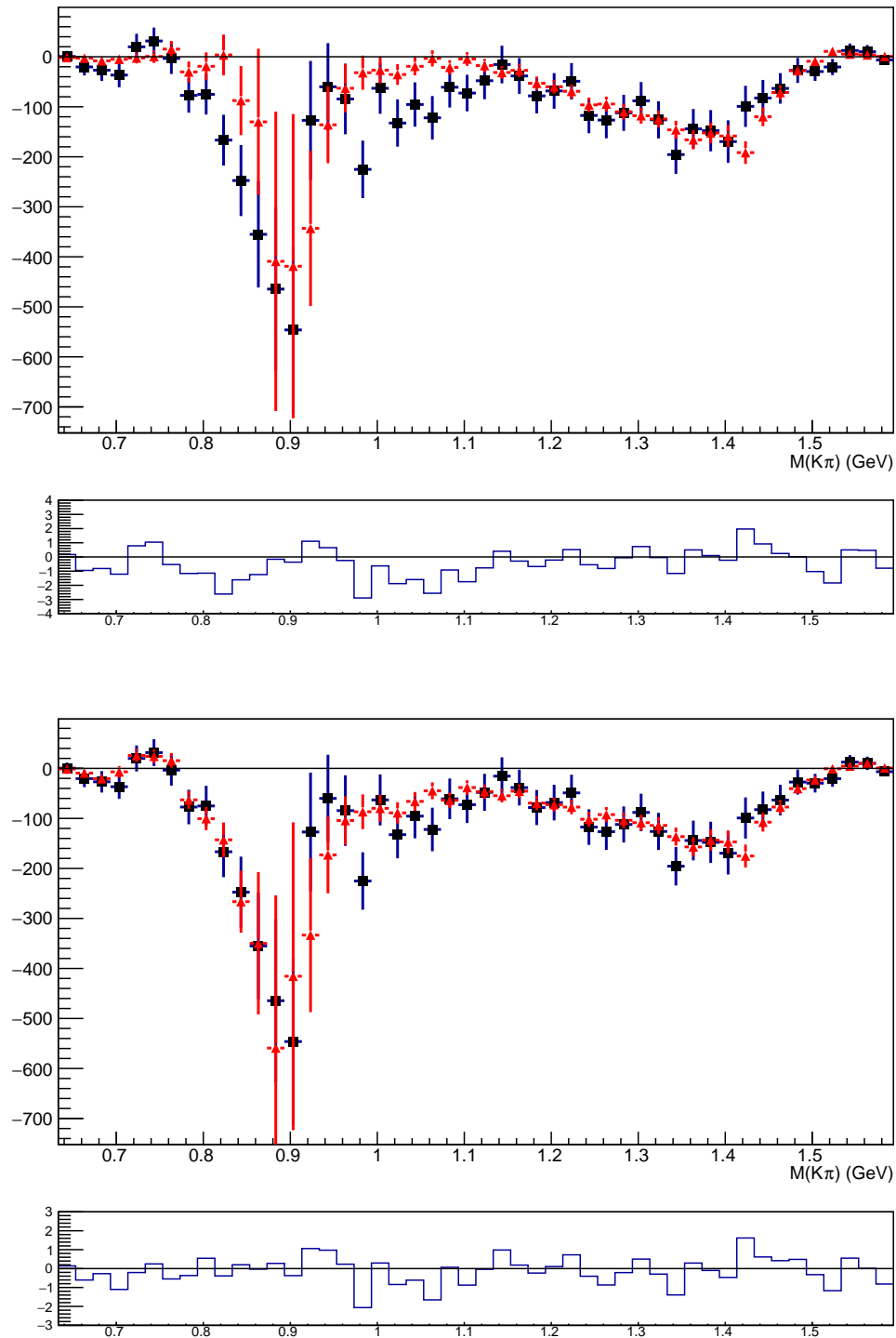


Figure 140: Unnormalized Legendre polynomial moment $\langle P_5^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 4 Z s and 2 Z_{K^*} s (bottom). The plots below the moment displays shows the fit pull distributions.

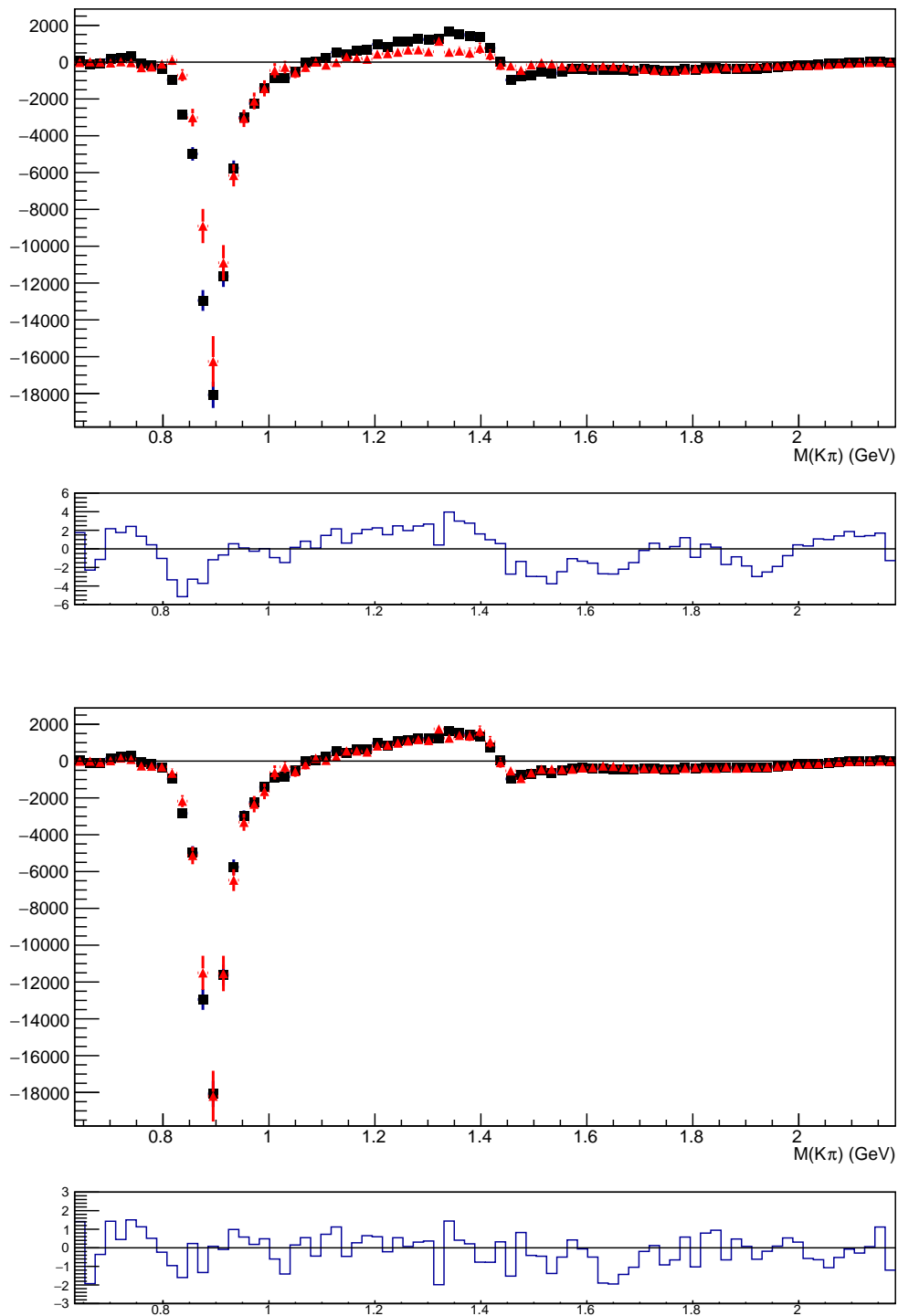


Figure 141: Unnormalized Legendre polynomial moment $\langle P_5^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 9 Z s and 1 Z_K (bottom). The plots below the moment displays shows the fit pull distributions.

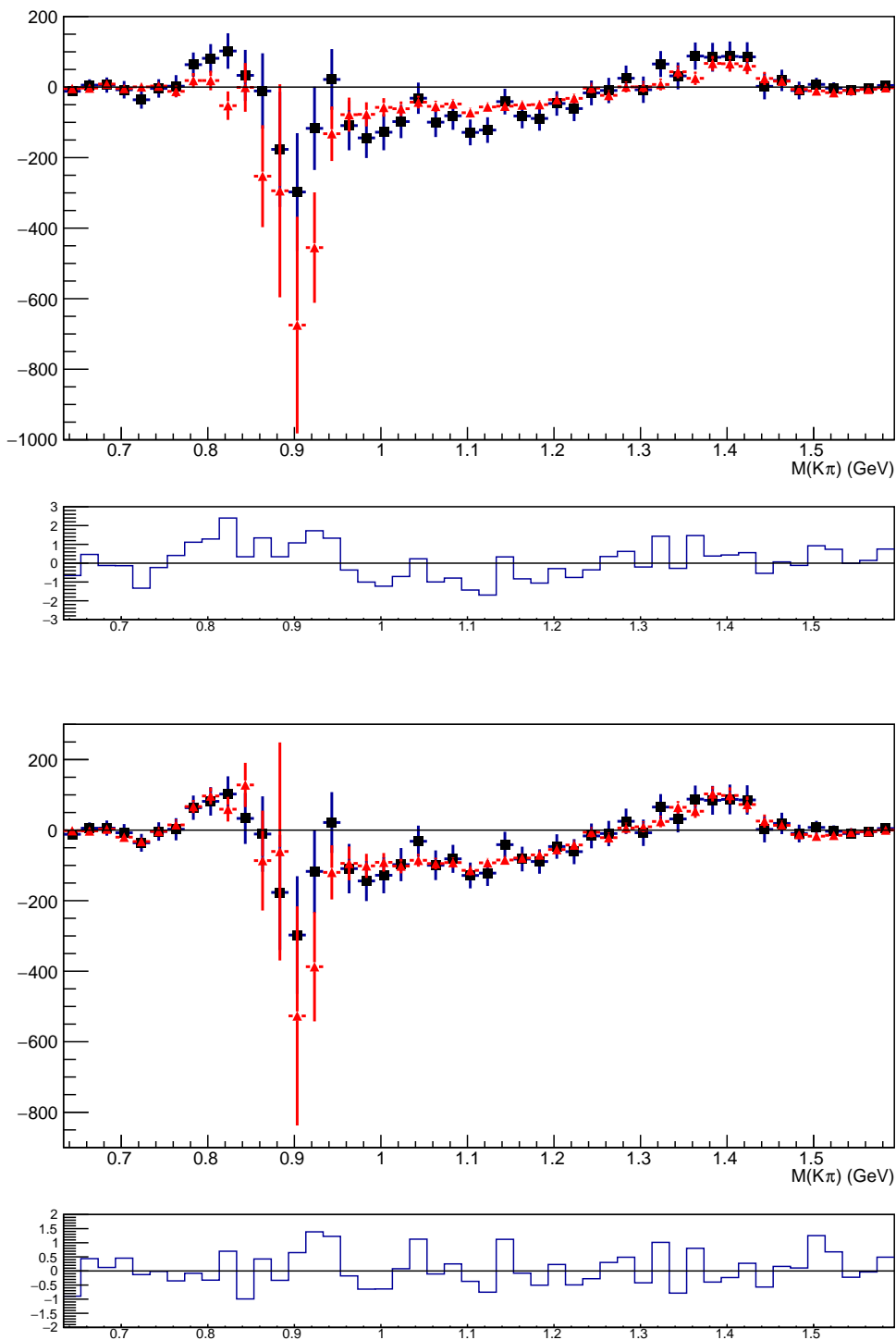


Figure 142: Unnormalized Legendre polynomial moment $\langle P_6^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 4 Z s and 2 Z_{K^*} s (bottom). The plots below the moment displays shows the fit pull distributions.

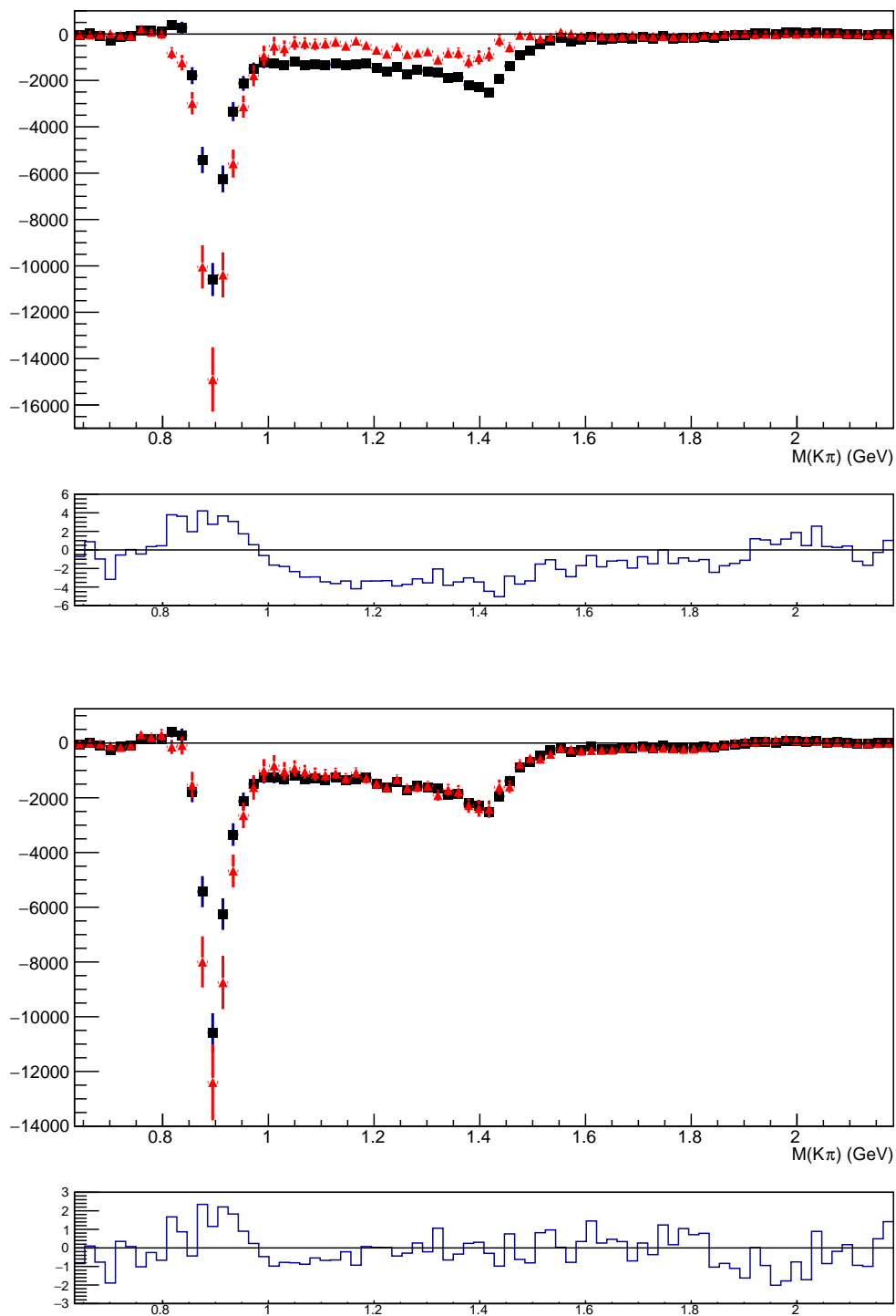


Figure 143: Unnormalized Legendre polynomial moment $\langle P_6^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 9 Z s and 1 Z_K (bottom). The plots below the moment displays shows the fit pull distributions.

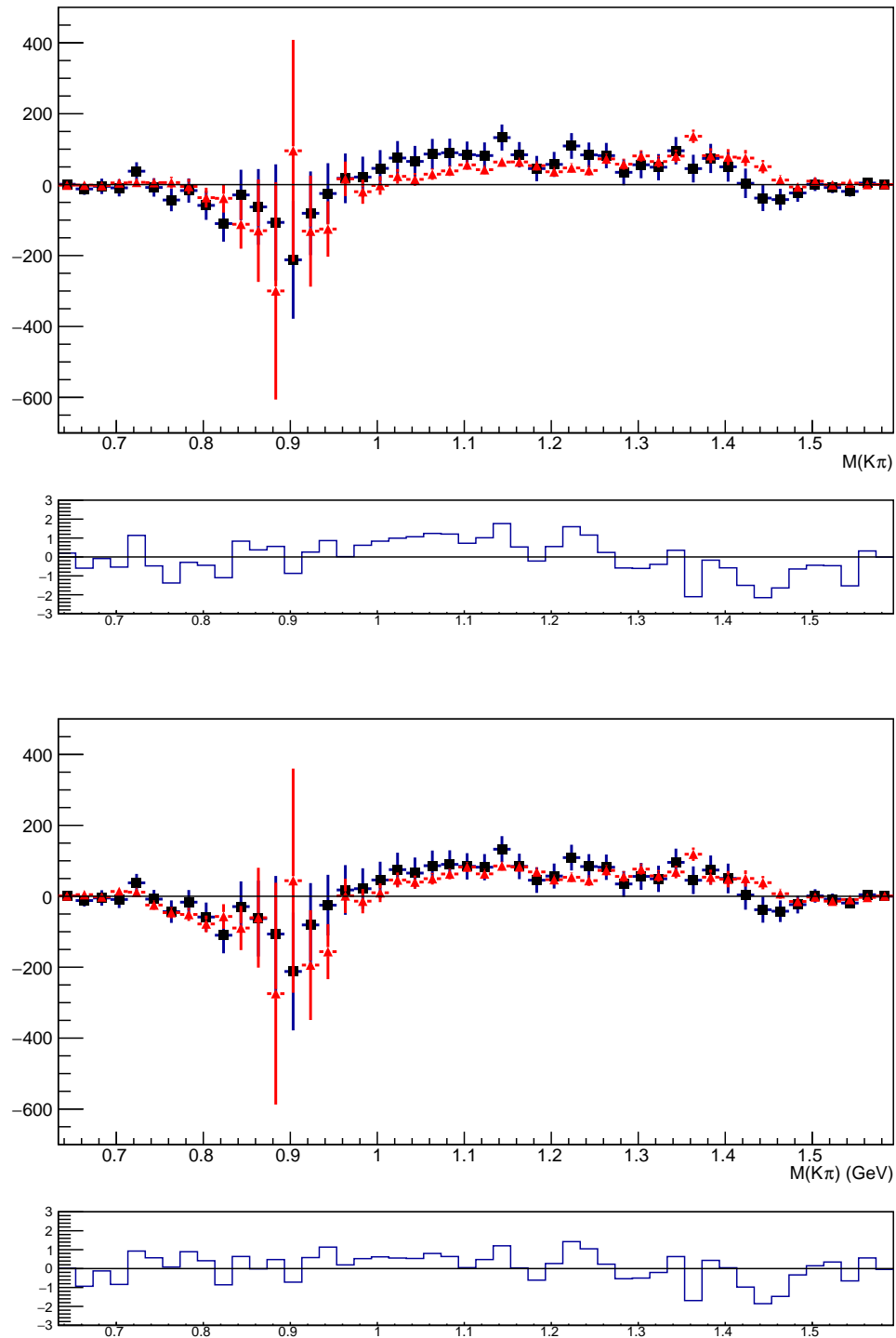


Figure 144: Unnormalized Legendre polynomial moment $\langle P_7^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 4 Z s and 2 Z_{K^*} s (bottom). The plots below the moment displays shows the fit pull distributions.

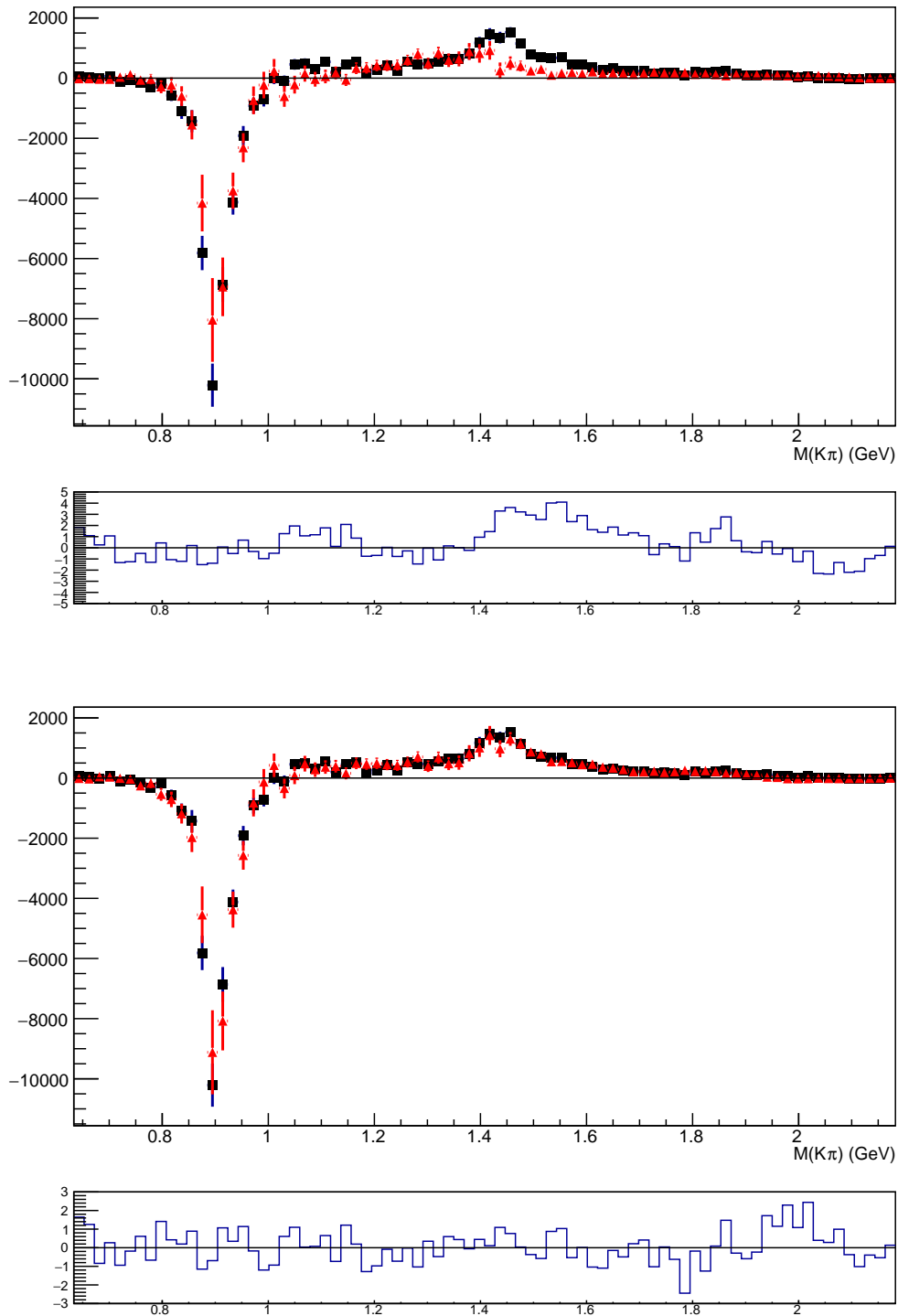


Figure 145: Unnormalized Legendre polynomial moment $\langle P_7^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 9 Z s and 1 Z_K (bottom). The plots below the moment displays shows the fit pull distributions.

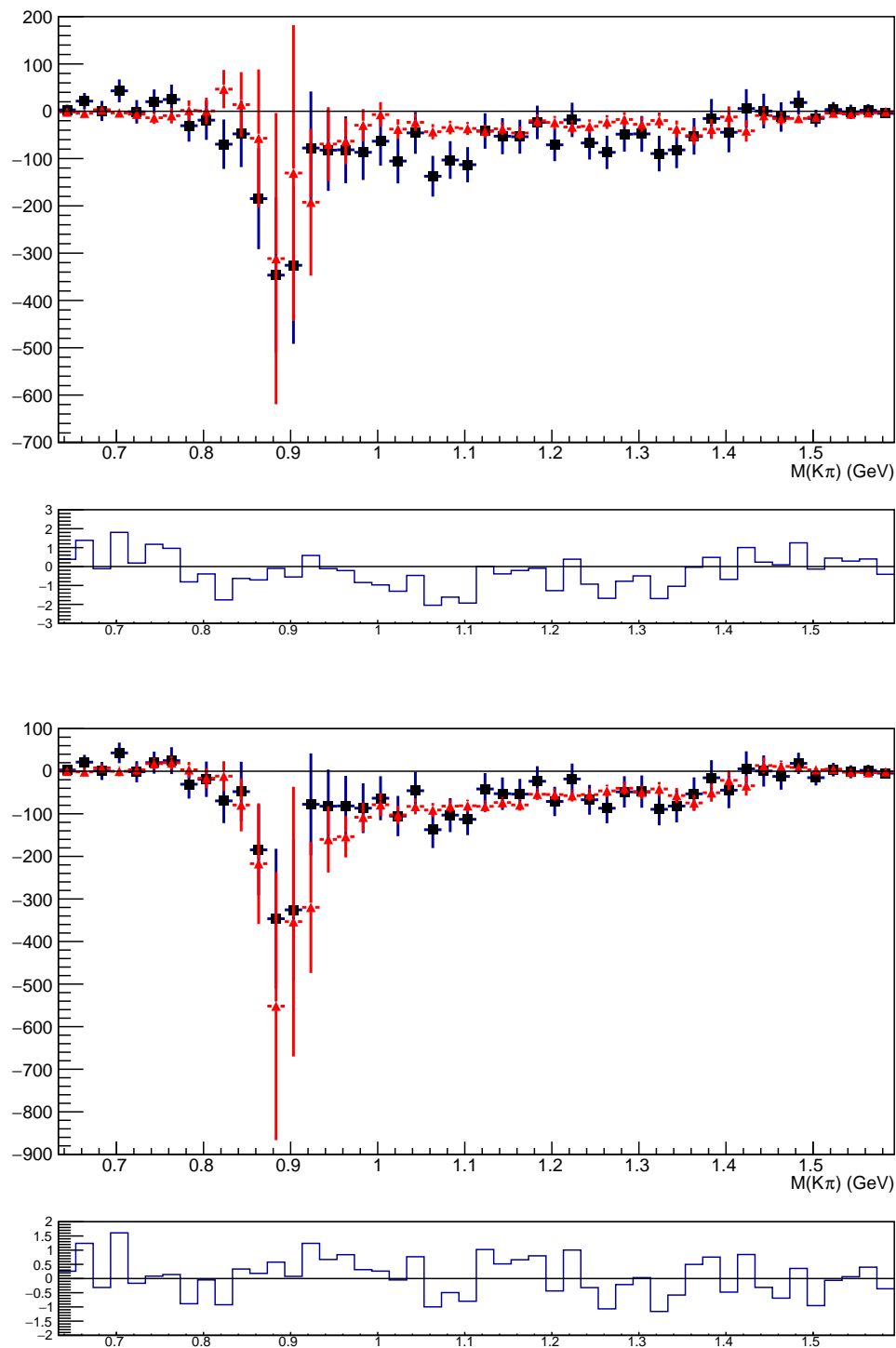


Figure 146: Unnormalized Legendre polynomial moment $\langle P_8^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^-K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 4 Z s and 2 Z_{K^*} s (bottom). The plots below the moment displays shows the fit pull distributions.

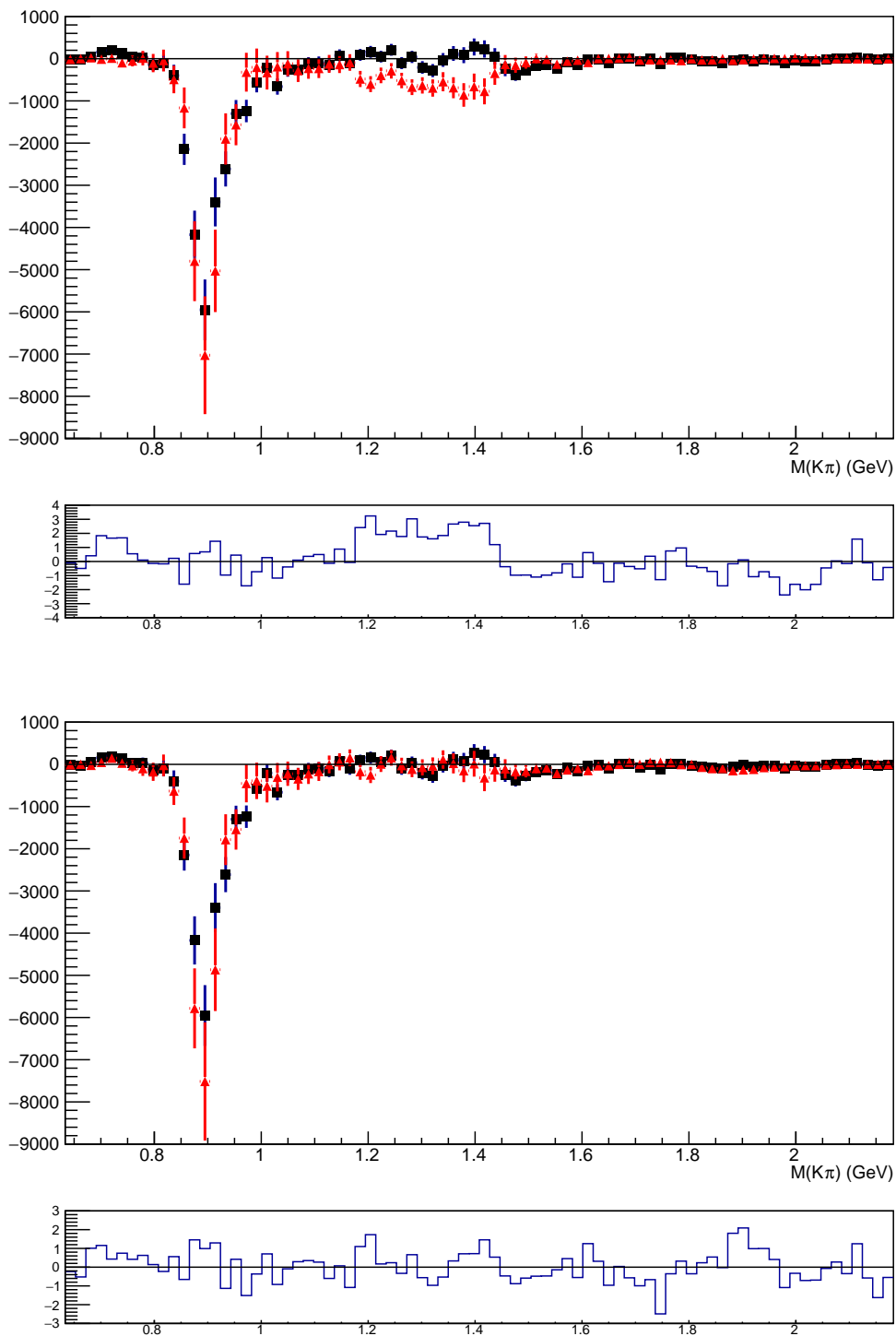


Figure 147: Unnormalized Legendre polynomial moment $\langle P_8^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 9 Z s and 1 Z_K (bottom). The plots below the moment displays shows the fit pull distributions.

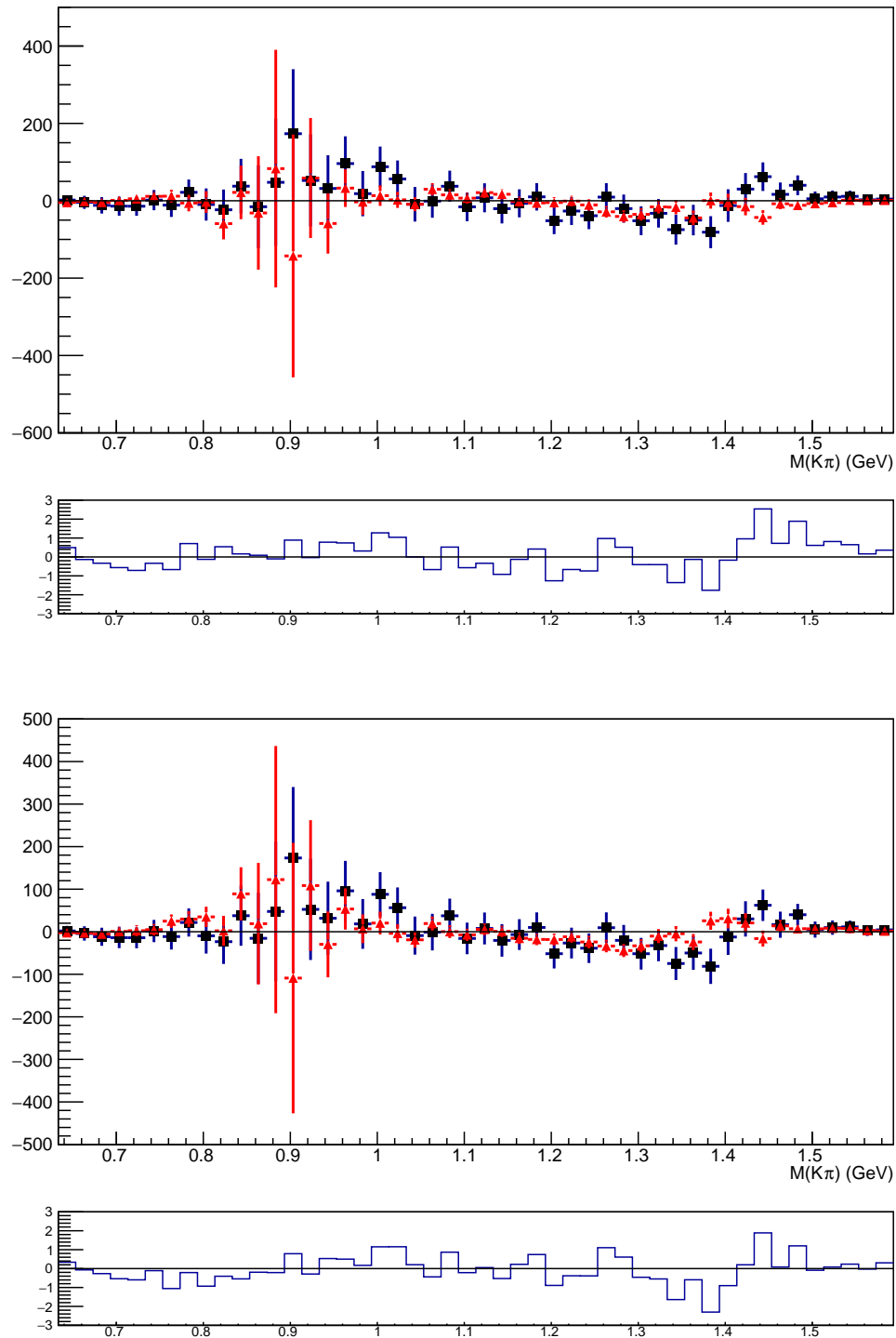


Figure 148: Unnormalized Legendre polynomial moment $\langle P_9^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 4 Z s and 2 Z_{K^*} s (bottom). The plots below the moment displays shows the fit pull distributions.

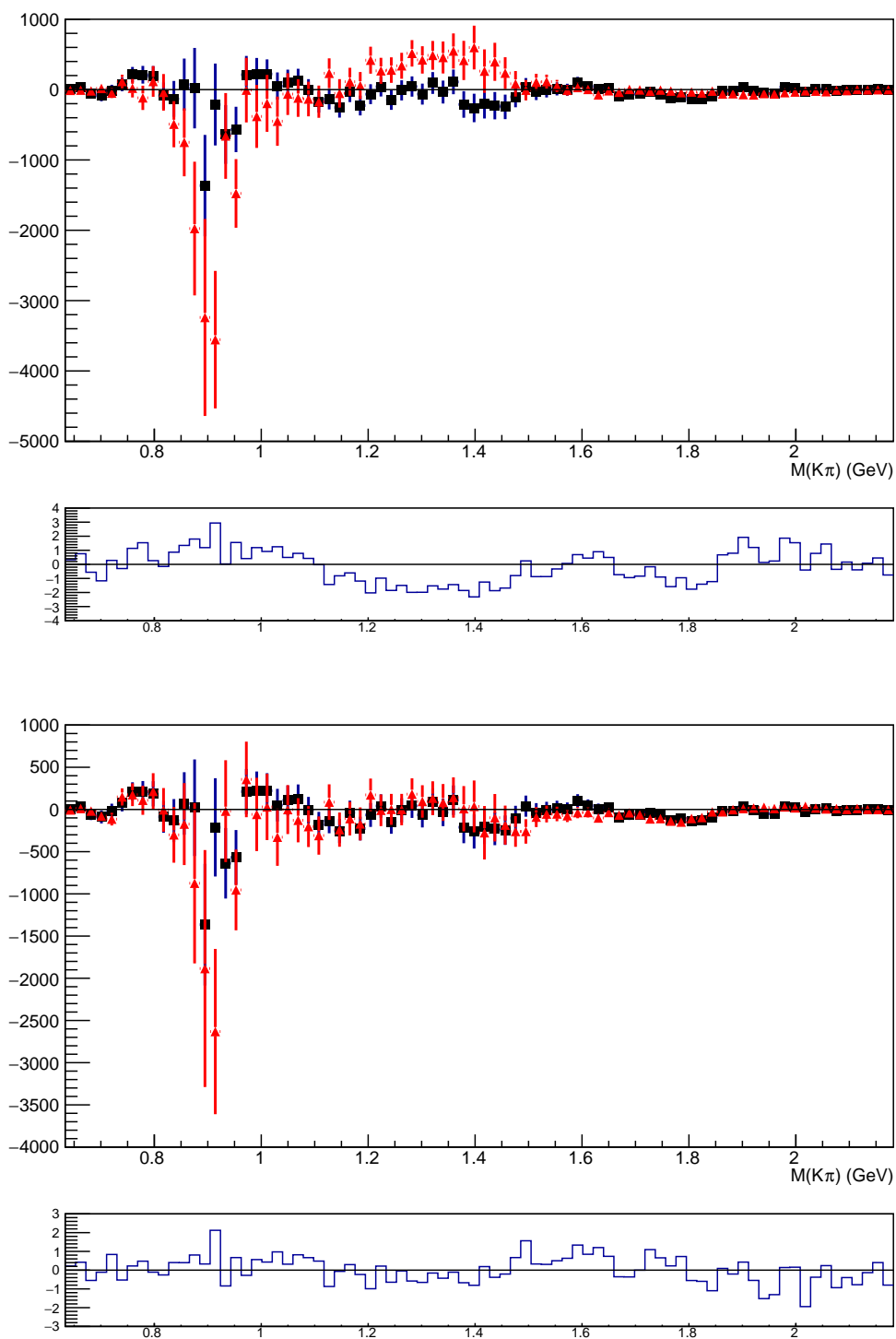


Figure 149: Unnormalized Legendre polynomial moment $\langle P_9^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 9 Z s and 1 Z_K (bottom). The plots below the moment displays shows the fit pull distributions.

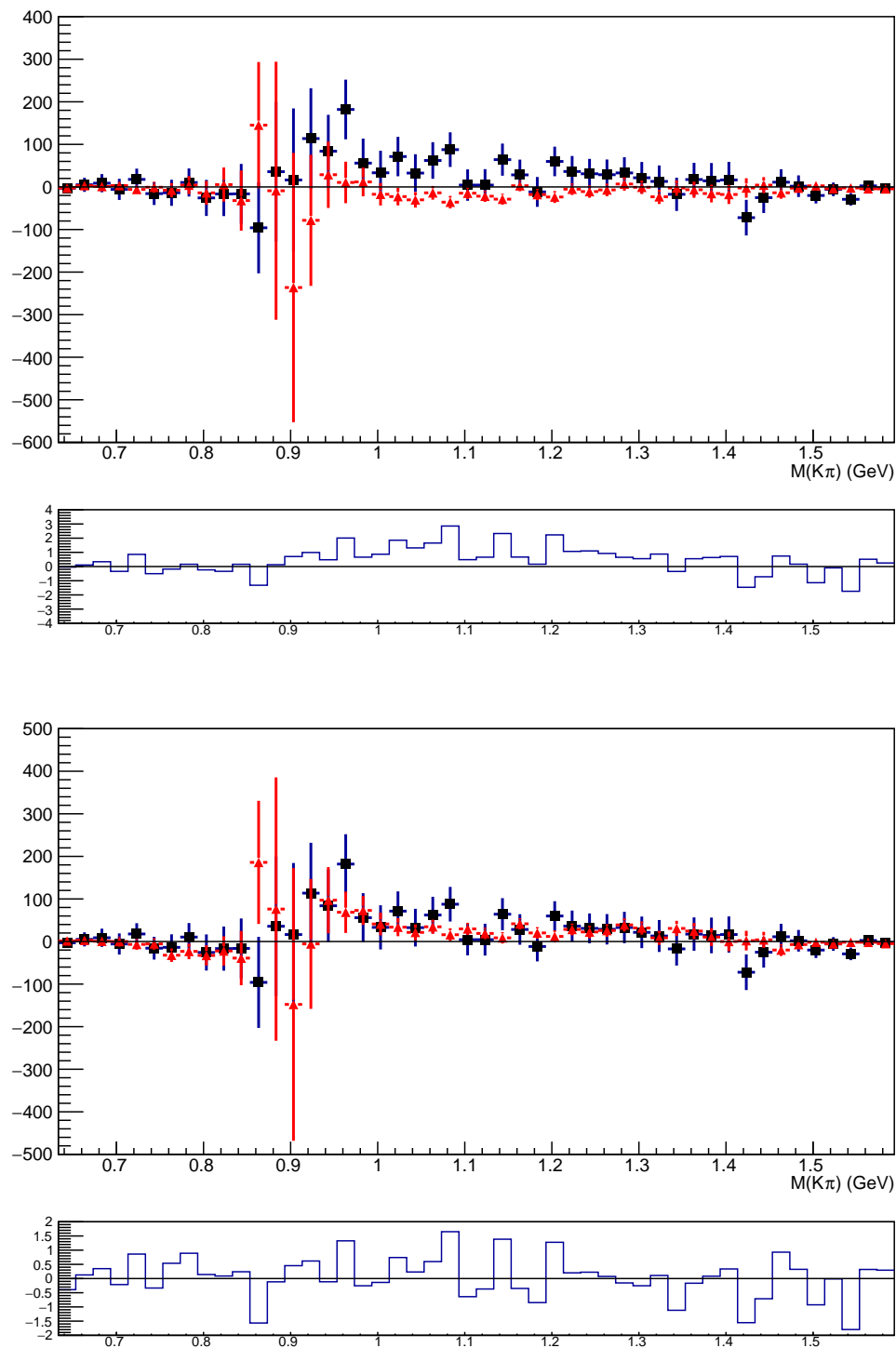


Figure 150: Unnormalized Legendre polynomial moment $\langle P_{10}^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow \psi(2S)\pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 4 Z s and 2 Z_{K^*} s (bottom). The plots below the moment displays shows the fit pull distributions.

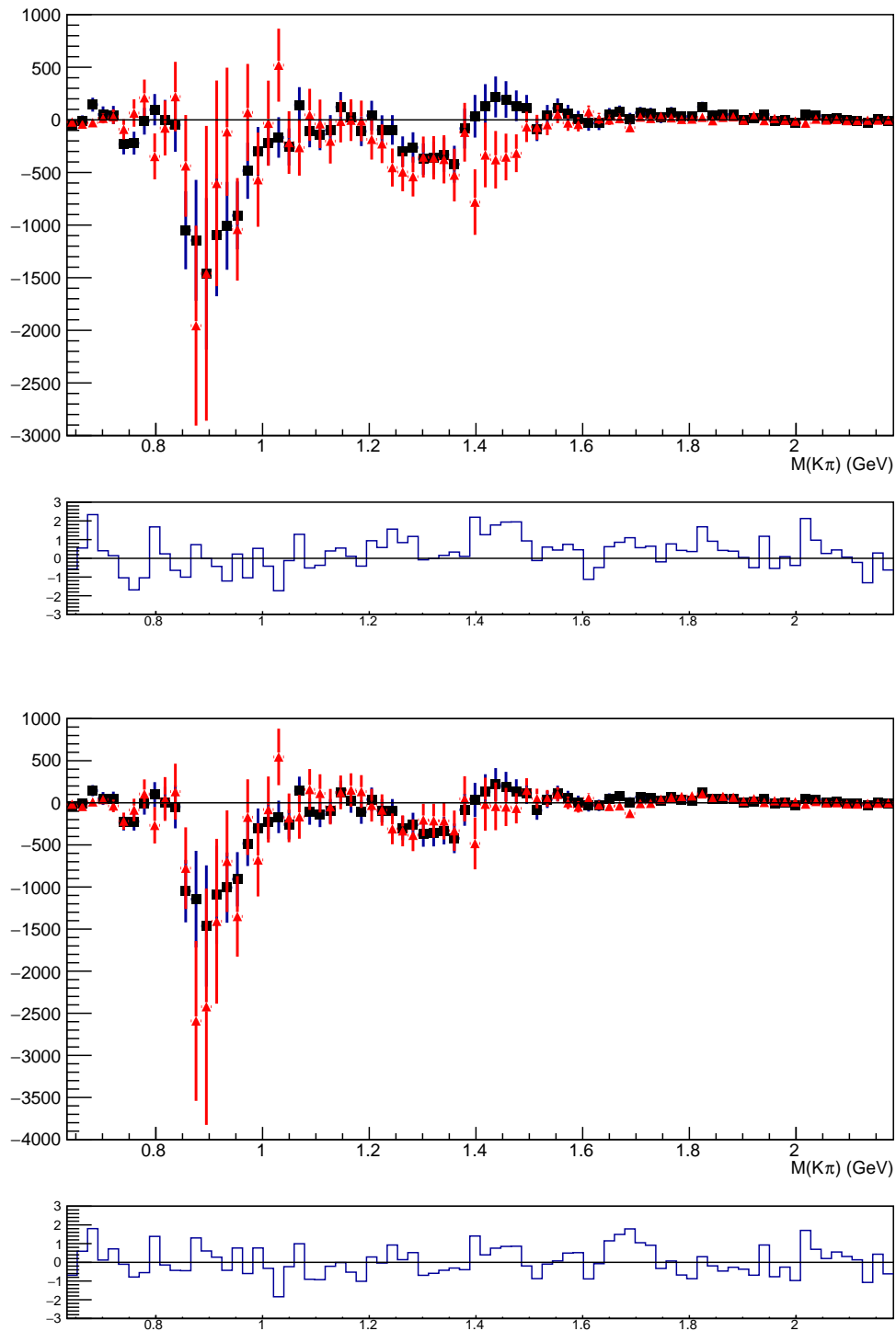


Figure 151: Unnormalized Legendre polynomial moment $\langle P_{10}^U \rangle$ of $\cos \theta_{K^*}$ in a function of $m_{K\pi}$ for the $B^0 \rightarrow J/\psi \pi^- K^+$ data (black points) and for the fits (red points) with the K^* -only model (top) and including 9 Z s and 1 Z_K (bottom). The plots below the moment displays shows the fit pull distributions.

References

- [1] Particle Data Group, R. L. Workman and Others, *Review of Particle Physics*, PTEP **2022** (2022) 083C01.
- [2] Particle Data Group, R. L. Workman and Others, *Review of Particle Physics*, PTEP **2022** (2022) 083C01.
- [3] *Standard model of elementary particles.*, https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg. Accessed: 2022-11-03.
- [4] E. Mobs, *The CERN accelerator complex in 2019. Complexe des accélérateurs du CERN en 2019*, , General Photo.
- [5] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.
- [6] *Lhcb facts*, <https://twiki.cern.ch/twiki/bin/view/Main/LHCb-Facts>. Accessed: 2022-02-17.
- [7] LHCb collaboration, R. Aaij *et al.*, *Observation of the resonant character of the $Z(4430)^-$ state*, Phys. Rev. Lett. **112** (2014) 222002, arXiv:1404.1903.
- [8] C. Elsasser, *Bb production angle plots*, .
- [9] LHCb collaboration, R. Aaij *et al.*, *Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions*, Phys. Rev. Lett. **118** (2017) 052002, Erratum ibid. **119** (2017) 169901, arXiv:1612.05140.
- [10] Belle collaboration, S. K. Choi *et al.*, *Observation of a resonance-like structure in the $\pi^\pm\psi'$ mass distribution in exclusive $B \rightarrow K\pi^\pm\psi'$ decays*, Phys. Rev. Lett. **100** (2008) 142001, arXiv:0708.1790.

- [11] Belle collaboration, R. Mizuk *et al.*, *Dalitz analysis of $B \rightarrow K\pi^+\psi'$ decays and the $Z(4430)^+$* , Phys. Rev. **D80** (2009) 031104, [arXiv:0905.2869](#).
- [12] Belle collaboration, K. Chilikin *et al.*, *Experimental constraints on the spin and parity of the $Z(4430)^+$* , Phys. Rev. **D88** (2013) 074026, [arXiv:1306.4894](#).
- [13] BABAR collaboration, B. Aubert *et al.*, *Search for the $Z(4430)^-$ at BABAR*, Phys. Rev. **D79** (2009) 112001, [arXiv:0811.0564](#).
- [14] LHCb collaboration, R. Aaij *et al.*, *Model-independent confirmation of the $Z(4430)^-$ state*, Phys. Rev. **D92** (2015) 112009, [arXiv:1510.01951](#).
- [15] LHCb collaboration, R. Aaij *et al.*, *Model-independent evidence for $J/\psi p$ contributions to $\Lambda_b^0 \rightarrow J/\psi p K^-$ decays*, Phys. Rev. Lett. **117** (2016) 082002, [arXiv:1604.05708](#).
- [16] Belle collaboration, K. Chilikin *et al.*, *Observation of a new charged charmoniumlike state in $\bar{B}^0 \rightarrow J/\psi K^- \pi^+$ decays*, Phys. Rev. D **90** (2014) 112009, [arXiv:1408.6457](#).
- [17] BESIII collaboration, M. Ablikim *et al.*, *Observation of a Charged Charmoniumlike Structure in $e^+e^- \rightarrow \pi^+\pi^- J/\psi$ at $\sqrt{s} = 4.26$ GeV*, Phys. Rev. Lett. **110** (2013) 252001, [arXiv:1303.5949](#).
- [18] Belle collaboration, Z. Q. Liu *et al.*, *Study of $e^+e^- \rightarrow \pi^+\pi^- J/\psi$ and Observation of a Charged Charmoniumlike State at Belle*, Phys. Rev. Lett. **110** (2013) 252002, [arXiv:1304.0121](#), [Erratum: Phys.Rev.Lett. 111, 019901 (2013)].
- [19] J. Bressieux, G. Cowan, M. Kreps, O. Schneider, T. Skwarnicki, and Q. Wenbin, *Evidence for the resonant character of the $Z(4430)^- \rightarrow \psi(2S)\pi^-$ mass peak observed in $B^0 \rightarrow \psi(2S)K^+\pi^-$ decays, and determination of the $Z(4430)^-$ spin-parity*, LHCb-ANA-2013-053, 2013.

- [20] T. Skwarnicki, $Z(4430)^+$ and other exotic charged meson candidates, Presented at the LHCb workshop on multi-body decays of B and D mesons, Rio de Janeiro, July 29, 2015, [Link to the slides](#).
- [21] LHCb collaboration, R. Aaij *et al.*, *Model-independent observation of exotic contributions to $B^0 \rightarrow J/\psi K^+ \pi^-$ decays*, Phys. Rev. Lett. **122** (2019) 152002, [arXiv:1901.05745](#).
- [22] Particle Data Group, P. A. Zyla *et al.*, *Review of particle physics*, Prog. Theor. Exp. Phys. **2020** (2020) 083C01.
- [23] E. Klempt and A. Zaitsev, *Glueballs, hybrids, multiquarks. Experimental facts versus QCD inspired concepts*, Phys. Rept. **454** (2007) 1, [arXiv:0708.4016](#).
- [24] S. Godfrey and N. Isgur, *Mesons in a relativized quark model with chromodynamics*, Phys. Rev. **D32** (1985) 189.
- [25] LHCb collaboration, R. Aaij *et al.*, *Amplitude analysis of $B^+ \rightarrow J/\psi \phi K^+$ decays*, Phys. Rev. **D95** (2017) 012002, [arXiv:1606.07898](#).
- [26] BESIII collaboration, M. Ablikim *et al.*, *Observation of a Near-Threshold Structure in the K^+ Recoil-Mass Spectra in $e^+e^- \rightarrow K^+(D_s^- D^{*0} + D_s^{*-} D^0)$* , Phys. Rev. Lett. **126** (2021) 102001, [arXiv:2011.07855](#).
- [27] LHCb collaboration, R. Aaij *et al.*, *Observation of new resonances decaying to $J/\psi K^+$ and $J/\psi \phi$* , Phys. Rev. Lett. **127** (2021) 082001, [arXiv:2103.01803](#).

Vita

NAME OF AUTHOR: Andy Beiter

PLACE OF BIRTH: Buffalo, NY, USA

DATE OF BIRTH: February 19, 1993

DEGREES AWARDED: Bachelors of Science in Physics and Computer Science. Canisius College, Buffalo, NY, USA