

Test of Lepton Universality using $B^+ \rightarrow K^+ \ell^+ \ell^-$ decays

by

Ka Tung Lau

B.Sc. in Physics, the Chinese University of Hong Kong, 2017

Sc.M. in Physics, Brown University, 2019

*A dissertation submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy in the Department of Physics at Brown University*

Providence, Rhode Island

May 2023

© Copyright 2023 by Ka Tung Lau

This dissertation by Ka Tung Lau is accepted in its present form by the Department of Physics as satisfying the dissertation requirement for the degree of Doctor of Philosophy.

Date:
Greg Landsberg, Advisor

Recommended to the Graduate Council

Date:
David Cutts, Reader

Date:
Jiji Fan, Reader

Approved by the Graduate Council

Date:
Thomas A. Lewis,
Dean of the Graduate School

Curriculum Vitae

Education

2017 - 2023	Ph.D. in Physics, Brown University Advisor: Prof. Greg Landsberg
2013 - 2017	B.Sc. in Physics, The Chinese University of Hong Kong Advisor: Prof. Ming Chung Chu

Fellowships and Awards

2022	Physics Merit Fellowship
2016	Professor Dennis Yam Kuen Lo Physics Award

Publications

IN PREPARA- TION	CMS collaboration, Test of lepton flavor universality in $B^+ \rightarrow K^+ \ell^+ \ell^-$ decays using CMS parked data
---------------------	---

Teaching and Mentoring

2018	Teaching Assistant, Introduction to Relativity and Quantum Physics
2017	Teaching Assistant, Foundations of Mechanics

Preface and Acknowledgments

I would like to express my heartfelt gratitude to the individuals who have supported me during the completion of my Ph.D. This journey would not have been possible without their guidance, encouragement, and support.

First and foremost, I would like to thank my advisor, Prof. Greg Landsberg, for his constant support throughout my entire Ph.D. Without his guidance and mentorship, it would have been impossible to complete my thesis project. I am particularly grateful for his support during the pandemic, which presented unexpected challenges and uncertainties. It is an honor to have had the opportunity to work with him. I would also like to extend my deepest appreciation to the members of my thesis committee, Prof. David Cutts, and Prof. Jiji Fan, for their time spent reading my thesis and providing valuable feedback.

I would also like to express my sincere thanks to my collaborators, Robert Bainbridge, Oliver Buchmueller, Vincenzo Cacchio, Francesca Cavallari, Sara Fiorendi, Georgios Karathanasis, Riccardo Manzoni, Nancy Marinelli, Arabella Martelli, Georgios Melachroinos, Chiara Rovelli, and Paris Sphicas. We have worked on this project for over five years, overcoming numerous challenges and difficulties. I am grateful for their dedication and hard work, and I am glad that we managed to finish the project together. Without their collaboration, this

project would not have been possible. I have learned so much from each one of them, and I am honored to have had the opportunity to work with such a talented and passionate team.

Finally, I would like to thank my wife, Sze Ning Mak, for her support, love, and understanding throughout this journey. Thank you for bringing joy and inspiration into my life. I would also like to express my gratitude to my parents for their unwavering support, encouragement, and belief in me throughout my academic journey. I am forever grateful for their presence in my life.

Abstract

Test of Lepton Universality using $B^+ \rightarrow K^+ \ell^+ \ell^-$ decays

by Ka Tung Lau, Ph.D., Brown University, May 2023

Several hints of potential B physics anomalies have been observed in recent years. Some of the observables, such as R_K , the ratio of branching fractions of $B^+ \rightarrow K^+ \mu^+ \mu^-$ to that of $B^+ \rightarrow K^+ e^+ e^-$, have exhibited deviations from the standard model prediction. This could be a sign of new physics involving lepton non-universality. In 2018, the CMS experiment implemented a novel trigger strategy called B -parking, which enabled the collection of an unbiased sample of $\mathcal{O}(10^{10})$ B hadron decays. This dataset provides an opportunity to measure and search for rare B decays, such as the measurement of R_K . This dissertation details the use of the data parking and tag-and-probe trigger strategy to collect an enormous and unbiased sample of B hadron decays, followed by a thorough analysis of the R_K measurement using this data. To increase the sensitivity of the analysis for soft electrons, a new low- p_T electron reconstruction algorithm is developed to improve the reconstruction efficiency in the low-energy region for electrons. Dedicated electron identifications are also developed for this analysis to reduce contamination from fake electrons. A machine-learning-based classifier is developed and applied to the event selection process to extract rare signals from

the overwhelming background. Finally, the analysis strategy is validated through a series of cross-checks, and the first measurement of R_K at CMS is reported.

Contents

Preface and Acknowledgments	v
Abstract	vii
1 Introduction	1
2 The Standard Model	3
2.1 The Elementary Particles of the Standard Model	4
2.1.1 Bosons	4
2.1.2 Fermions	6
2.2 Theoretical framework of the Standard Model	7
2.2.1 Quantum Electrodynamics (QED)	7
2.2.2 Quantum Chromodynamics (QCD)	9
2.2.3 Electroweak theory	10
2.2.4 Higgs mechanism	12
3 Overview of the test of Lepton Universality in B-physics sector	16
3.1 The $b \rightarrow s\ell^+\ell^-$ transition as a probe of New Physics	16

3.2	Effective Hamiltonian for $b \rightarrow s\ell^+\ell^-$ transition	17
3.3	Lepton Universality in the Standard Model	20
3.4	B flavor anomalies and Lepton-Flavor-Universality Violation	22
3.5	R_K as an observable for the test of Lepton Universality	25
4	The Compact Muon Solenoid	28
4.1	The Large Hadron Collider	28
4.2	The CMS detector	32
4.2.1	Coordinate System	32
4.2.2	The Inner Tracker	34
4.2.3	Electromagnetic Calorimeter (ECAL)	36
4.2.4	Hadron Calorimeter (HCAL)	37
4.2.5	Muon System	39
4.3	Trigger System	40
4.4	Physics Objects Reconstruction	42
4.4.1	Tracks	44
4.4.2	Electrons and Photons	45
4.4.3	Muons	45
5	Test of Lepton Universality using $B^+ \rightarrow K^+\ell^+\ell^-$ decays	47
5.1	Introduction	47
5.1.1	Measurement strategy	49
5.2	Data parking	51
5.3	Triggers and data	51

5.3.1	Trigger purity	54
5.3.2	Pile-up and luminosity	57
5.3.3	Event simulation	57
5.4	Analysis strategy	59
5.5	Object reconstruction	61
5.5.1	Muons	61
5.5.2	Electrons	65
5.5.3	Kaons	78
5.5.4	Kinematic fit	79
5.6	Background sources	83
5.7	The muon channel	85
5.7.1	Trigger selection and scale factors	85
5.7.2	Event preselection	94
5.7.3	Event selection with a BDT in the muon channel	95
5.7.4	Mass fit	99
5.7.5	Acceptance and efficiency	117
5.7.6	Checks for the muon channel	123
5.7.7	Systematic uncertainties for the muon channel	129
5.8	The electron channel	135
5.8.1	Event preselection	135
5.8.2	Event selection with a BDT in the electron channel	137
5.8.3	Mass fit	141
5.8.4	Acceptance and efficiency	161

5.8.5	Checks for the electron channel	164
5.8.6	Systematic uncertainties for the electron channel	176
5.9	Cross-checks	194
5.9.1	Cross-check with $R_K^{\psi(2S)}$	194
5.9.2	Cross-check with $R_K^{J/\psi}$	197
5.9.3	Cross-check with single ratio $R_{\psi(2S)}^{\text{single ratio}}$	201
5.10	Results	203
5.11	Conclusions	206
A	Distributions for the check of mass sculpting in the electron channel	207
B	Distributions for the check of misidentified background in the electron channel	214
C	Alternative templates used for the study of the choice of fit parameterization in the muon channel	219
D	Alternative templates used for the study of the choice of fit parameterization in the electron channel	222

List of Tables

2.1	Summary of the coupling strengths of fundamental interactions in the SM.	5
4.1	Designed LHC beam parameters [35].	30
5.1	Minimum and maximum lepton invariant mass values for the R_K measurement ranges considered in this analysis.	50
5.2	Summary of the loosest muon trigger requirements imposed by the L1 and HLT algorithms in each instantaneous luminosity scenario: the L1 and HLT muon p_T thresholds, and the HLT muon impact parameter significance IP_{sig} . Also shown are the trigger purity and peak trigger rate. All values are shown as a function of the peak instantaneous luminosity L_{inst} . The table refers to the trigger configuration that was online starting from LHC Fill 7132.	52
5.3	Summary of the tag-side muon trigger L1 seeds at different peak instantaneous luminosity values. Mu10 stands for muon $p_T > 10$ GeV, and “er1p5” for $ \eta < 1.5 $. Each column represents one of the different settings that were online throughout the year, together with the number of the first LHC Fill in which they were implemented. The majority of the data has been collected using the settings in the last two columns.	53

5.4	List of variables and their corresponding cuts, to reduce the background in ditrack pairs. Prob(vtx) is the probability of the secondary vertex, and L_{xy} the absolute displacement from the Beam Spot.	55
5.5	Central developed muon identification at CMS.	63
5.6	Variables used in the identification BDT for LP electrons.	68
5.7	Variables used in the identification BDT for PF electrons.	70
5.8	Signal Crystal Ball mean and sigma values for fits to MC and data for PF-PF.	73
5.9	Signal Crystal Ball mean and sigma values for fits to MC and data for PF-LP.	74
5.10	Variables used in tag-and-probe pair selection.	88
5.11	A set of preselection cuts for the muon channel.	94
5.12	Variables used in the analysis BDT in the muon channel.	97
5.13	Blinded yields for signal and background sources in the dimuon low- q^2 bin.	101
5.14	Summary of the fit for the inclusive low- q^2 for the $B^\pm \rightarrow \mu\mu K^\pm$ mode.	104
5.15	Summary of the fit for the J/ψ bin for the $B^\pm \rightarrow J/\psi K^\pm \rightarrow \mu\mu K^\pm$ mode.	106
5.16	Yields for signal and background sources in the J/ψ bin.	106
5.17	Summary of the fit for the $\psi(2S)$ bin for the $B^\pm \rightarrow \psi(2S) K^\pm \rightarrow \mu\mu K^\pm$ mode.	111
5.18	Yields for signal and background sources for the muon channel in the $\psi(2S)$ bin.	111
5.19	Blinded yields for signal and background sources in the high q^2 bin.	114
5.20	Summary of the fit for the inclusive high- q^2 for the $B^\pm \rightarrow \mu\mu K^\pm$ mode.	114
5.21	Acceptance and efficiency factorized and inclusive for all final states. Uncer- tainties are only statistical.	121
5.22	Difference in R_K^μ from the variation of the HLT SF.	129

5.23	Difference in R_K^μ from the variation of the PU SF.	130
5.24	Difference in R_K^μ from the variation of the BDT SF.	130
5.25	Summary of the systematic uncertainties from the choice of fit parameteriza- tion in the J/ψ region for the muon channel.	132
5.26	Summary of the systematic uncertainties from the choice of fit parameteriza- tion in the low- q^2 region for the muon channel.	132
5.27	Summary of systematic uncertainty for the R_K measurement in the muon channel.	134
5.28	Variables used in the analysis BDT in the electron channel.	139
5.29	Summary of the fit in the low- q^2 bin.	142
5.30	Yields for signal and background sources in low- q^2	143
5.31	Summary of the fit in the J/ψ bin.	147
5.32	Yields for signal and background sources in J/ψ	147
5.33	Summary of the fit in the $\psi(2S)$ bin.	152
5.34	Yields for signal and background sources in $\psi(2S)$	152
5.35	Summary of the fit in the high- q^2 bin.	157
5.36	Acceptance and efficiency factorized and inclusive for all final states for PF- PF. Uncertainties are only statistical.	163
5.37	Acceptance and efficiency factorized and inclusive for all final states for PF- LP. Uncertainties are only statistical.	163
5.38	Summary of the expected number of misidentified background in whole q^2 region without applying anti- D^0 cut.	174

5.39	Summary of the expected number of misidentified background in the low- q^2 region with anti- D^0 cut applied.	175
5.40	Summary of $R_{\text{trigger}}^{\text{MC}}$ in different q^2 region at different working points for PF-PF. The proposed systematic uncertainty is 1%.	180
5.41	Summary of $R_{\text{trigger}}^{\text{MC}}$ in different q^2 region at different working points for PF-LP. The proposed systematic uncertainty is 4%.	180
5.42	Summary of the systematic uncertainties from the choice of fit parameterization in the J/ψ region for the electron channel. The total systematic uncertainties from the choice of fit parameterization are 0.8% and 1.5% for PF-PF and PF-LP, respectively.	188
5.43	Summary of the systematic uncertainties from the choice of fit parameterization in the low- q^2 region for the electron channel. The numbers show here is averaged over 1000 MC toys. The total systematic uncertainties from the choice of fit parameterization are 4.9% and 5.0% for PF-PF and PF-LP, respectively.	188
5.44	Summary of the signal yields and the efficiency with different q^2 windows in the J/ψ bin for PF-PF.	190
5.45	Summary of the signal yields and the efficiency with different q^2 windows in the J/ψ bin for PF-LP.	190
5.46	Summary of systematics for the R_K measurement in the electron channel . .	193
5.47	Summary of the parameters for simultaneous fit of $R_K^{\psi(2S)}$	195
5.48	Summary of the parameters for simultaneous fit of $R_K^{J/\psi}$	198
5.49	Summary of the parameters for simultaneous fit of R_K	205

List of Figures

2.1	Particle content of the Standard Model [6].	5
2.2	The Higgs potential [13].	13
3.1	Feynman diagrams contributing to the $b \rightarrow s\ell^+\ell^-$ transition	17
3.2	Diagrams for $\mathcal{O}_7$ (a) and $\mathcal{O}_{9\ell,10\ell}$ (b) operators.	19
3.3	A schematic diagram of the differential branching fractions as a function q^2 for $B^+ \rightarrow K^+\ell^+\ell^-$ and the chromium resonances.	19
3.4	Combined R_D and R_{D^*} results compare with the SM prediction [25].	24
3.5	Latest R_K and R_{K^*} results released by the LHCb [29].	25
4.1	A schematic diagram of the CERN accelerator complex [34].	29
4.2	Peak instantaneous luminosity (a) and total integrated luminosity (b) recorded at the CMS since 2010 [36].	31
4.3	A cutaway view of the CMS detector [37].	33
4.4	Sketch of one quarter of the Phase-1 CMS tracking system in r - z view. The pixel detector is shown in green, while single-sided and double-sided strip modules are depicted as red and blue segments, respectively [38].	34
4.5	Schematic cross section through the CMS tracker [39].	35

4.6	Layout of the ECAL at CMS [42].	37
4.7	A schematic diagram showing a quarter of the HCAL during Run 2 [43]. . .	38
4.8	A schematic diagram showing a quarter of the muon system at CMS [47]. . .	40
4.9	Overview of the CMS L1 trigger system [49].	41
4.10	A sketch of the specific particle interactions in a transverse slice of the CMS detector, which illustrates the particle reconstruction at CMS [50].	43
5.1	Simulated generated muons' p_T (left) and pseudorapidity- η (right) spectra using $B^+ \rightarrow K^+ \mu^+ \mu^-$ and $B^+ \rightarrow K^+ J/\psi \rightarrow K^+ \mu^+ \mu^-$	50
5.2	Rate of the CMS L1 trigger (blue data points), as a function of time, during an LHC fill in (a) 2017 and (b) 2018. The time intervals cover 13–15 hours. Changes in the run number and settings (prescale column) are indicated by vertical grey and red dashed lines, respectively. Figure (c) shows the trigger rates of the CMS HLT system as a function of time during an LHC fill in 2018, for the physics (black data points) and B parking (blue data points) streams.	55
5.3	Mass of the track pairs. The masses are assigned based on the charge of tracks and μ	56
5.4	Mass difference between the K , π , π_s and K , π combinations. Only a subset of parked events have been used. The right and wrong sign refer to the sign of K : right sign is the same as the μ , wrong is the opposite of the μ . Wrong sign events give a reasonably accurate description of the background under the peak.	57
5.5	Distribution of the number of true interactions in the events collected by the B -parking data sample.	58

5.6	A schematic diagram showing a $B^+ \rightarrow K^+ \ell^+ \ell^-$ decay in the probe side and a displaced trigger muon in the tag side.	60
5.7	Invariant mass of each of the reconstructed μ in the triplet and the track. A peak is observed around the D^0 mass.	62
5.8	(a) Invariant mass of a candidate μ and a track, calculated assuming that the muon has the K mass and the track the π mass; (b) Invariant mass of a candidate μ and a track, calculated assuming that the muon has the π mass and the track the K mass.	63
5.9	Invariant mass of each of the reconstructed muons in the triplet and the track, when the candidate passes medium (in black) and soft (in red) ID.	64
5.10	Distributions of the transverse momenta of the final state particles for the $B^+ \rightarrow K^+ \ell^+ \ell^-$ decay. The last bin contains the overflow.	65
5.11	Efficiency vs. mistag rate for electrons calculated using different working points of the seeding BDT. The red square indicates the PF electron working point, the blue curve shows the "kinematically agnostic" BDT that exploits tracking and calorimeter information, and the model-dependent "kinematically aware" BDT (green curve) that also utilises the p_T , η , and the track impact parameter of an electron candidate.	67
5.12	(left) Distribution of the LP electron ID BDT for low- p_T electrons matched to signal electrons (green) and for fakes (blue); (right) selection efficiency for signal LP electrons versus efficiency on fakes based on cuts on the LP electron ID BDT (black). Also shown is the corresponding curve for the model in the Jan/2020 processing (red).	69

5.13	ROC curves with cross-validation for the re-trained PF electron ID in kinematic region of (a) $p_T < 5$ GeV and (b) $p_T > 5$ GeV.	71
5.14	Selection efficiency for signal PF electrons versus efficiency on fakes based on cuts on the re-trained PF electron ID BDT (red). Also shown are the corresponding curves for the default BDTs provided by the EGamma POG (blue and green). Electrons with a) $p_T < 5$ GeV and b) $5 < p_T < 10$ GeV . . .	72
5.15	(left) Fit to the two electrons invariant mass distribution for events with a (left) probe LP electron in the barrel with $1.5 < p_T < 2$ GeV and (right) probe PF electron in the barrel with $2 < p_T < 5$ GeV. Data (dots) are fitted with a double crystal ball (red) as a signal component and an exponential (green) as a background component.	73
5.16	Distribution of the LP electron ID BDT for LP electrons in data (dots) and in MC (solid line); the <i>sPlot</i> technique is used to remove the background in data. Probe electrons in the barrel are binned in p_T : (a) $1 < p_T < 1.5$ GeV, (b) $1.5 < p_T < 2$ GeV, (c) $2 < p_T < 5$ GeV, (d) $p_T > 5$ GeV	75
5.17	Distribution of the PF electron ID BDT for PF electrons in data (dots) and in Monte Carlo (solid line); the <i>sPlot</i> technique is used to remove the background in data. Probe electrons in the barrel are binned in p_T : (a) $2 < p_T < 5$ GeV, (b) $p_T > 5$ GeV	76
5.18	Distribution of the LP electron ID BDT for LP electrons in data (dots) and in Monte Carlo (solid line); the <i>sPlot</i> technique is used to remove the background in data. Probe electrons in the endcaps are binned in p_T : (a) $2 < p_T < 5$ GeV, (b) $p_T > 5$ GeV	76

5.19	Distribution of the PF electron ID BDT for PF electrons in data (dots) and in Monte Carlo (solid line); the <i>sPlot</i> technique is used to remove the background in data. Probe electrons in the endcaps are binned in p_T : (a) $2 < p_T < 5$ GeV, (b) $p_T > 5$ GeV	77
5.20	Invariant mass of each of the reconstructed μ in the triplet and the track (assuming a μ mass for the particle generating the track) as observed in data. Both the J/ψ and $\psi(2S)$ peaks are visible in the distribution.	78
5.21	(a) p_T distribution for tracks in the barrel region; (b) p_T distribution for tracks in the endcap region.	80
5.22	Comparison between fitted and raw mass of $m(e^+e^-)$ (a), and $m(K^+e^+e^-)$ (b), with PF-PF (top) and PF-LP (bottom) in the J/ψ bin.	81
5.23	Comparison between fitted and raw mass of $m(e^+e^-)$ (a), and $m(K^+e^+e^-)$ (b), with PF-PF (top) and PF-LP (bottom) in the $\psi(2S)$ bin.	82
5.24	Muon η distribution, for muons that fired the HLT_Mu9_IP6, using a subset of the available data. The drop around 1.5 is due to the seed η restriction.	86
5.25	Distribution of tag μ in data and simulation (MC). Clearly, in simulation most of the events are triggered by different trigger than in data.	87
5.26	Mass distribution of the muon pair in two η regions.	88
5.27	Fit for the passing probes (left) and fit for the failing probes (right).	89
5.28	Efficiency in data for $ \eta < 0.4$ (a); Efficiency in simulation for $ \eta < 0.4$ (b); Scale factor for $ \eta < 0.4$ (c).	90
5.29	Efficiency in data for $0.4 < \eta < 0.8$ (a); Efficiency in simulation for $0.4 < \eta < 0.8$ (b); Scale factor for $0.4 < \eta < 0.8$ (c).	90

5.30	Efficiency in data for $0.8 < \eta < 1.5$ (a); Efficiency in simulation for $0.8 < \eta < 1.5$ (b); Scale factor for $0.8 < \eta < 1.5$ (c).	90
5.31	Trigger status of the leading muon in all 4 dimuon mass bins: low- q^2 bin (a), high- q^2 bin (b), J/ψ bin (c) and $\psi(2S)$ bin (d).	92
5.32	Trigger status of the subleading muon in all 4 dimuon mass bins: low- q^2 bin (a), high- q^2 bin (b), J/ψ bin (c) and $\psi(2S)$ bin (d).	93
5.33	Invariant mass of B candidates, with dashed lines indicating the left and right sidebands (SB(L) and SB(R) respectively) (a). Invariant mass of muon pairs, with dashed lines indicating the different q^2 bins (b).	95
5.34	Signal efficiency vs Signal and Background training events for a tight BDT cut (a). Time required for training as a function of signal and background events (b).	96
5.35	BDT output for signal (MC simulation in red) and background (Same-sign data in black).	98
5.36	Significance as a function of BDT working points for low- q^2	98
5.37	Triplet invariant mass models in the low q^2 bin for the signal $B^\pm \rightarrow K^\pm \mu^+ \mu^-$ (a), $B^\pm \rightarrow \pi^\pm \mu^+ \mu^-$ (b), $B \rightarrow K^*(K\pi) \mu^+ \mu^-$ where the π is mistaken for the K (c), $B \rightarrow K^* \mu^+ \mu^-$ where the K is used in the triplet (d), combinatorial background (e), other b decays (f).	102
5.38	Invariant mass distribution for the B candidates in the muon channel in the low- q^2 bin. The fit is overlaid to the data. The bottom panel shows the pull distribution.	103

5.39	Triplet invariant mass models in the J/ψ bin for the signal $B^\pm \rightarrow K^\pm J/\psi(\mu^+\mu^-)$	
	(a), $B^\pm \rightarrow \pi^\pm J/\psi(\mu^+\mu^-)$ (b), $B \rightarrow K^*(K\pi)J/\psi(\mu^+\mu^-)$ where the π is mis-	
	taken for the K (c), $B \rightarrow K^*J/\psi(\mu^+\mu^-)$ where the K is used in the triplet (d),	
	combinatorial background (e), other b decays (f).	107
5.40	Invariant mass distribution for the B candidates in the muon channel in the	
	J/ψ bin. The fit is overlaid to the data. The bottom panel shows the pull	
	distribution.	108
5.41	Triplet invariant mass models in the $\psi(2S)$ bin for the signal $B^\pm \rightarrow K^\pm \psi(2S)(\mu^+\mu^-)$	
	(a), $B^\pm \rightarrow \pi^\pm \psi(2S)(\mu^+\mu^-)$ (b), $B \rightarrow K^*(K\pi)\psi(2S)(\mu^+\mu^-)$ where the π is mis-	
	taken for the K (c), $B \rightarrow K^*\psi(2S)(\mu^+\mu^-)$ where the K is used in the triplet	
	(d), combinatorial background (e), other b decays (f).	110
5.42	Invariant mass distribution for the B candidates in the muon channel in the	
	$\psi(2S)$ bin. The fit is overlaid to the data. The bottom panel shows the pull	
	distribution.	112
5.43	Triplet invariant mass models in the high q^2 bin for the signal $B^\pm \rightarrow K^\pm \mu^+\mu^-$	
	(a), $B^\pm \rightarrow \pi^\pm \mu^+\mu^-$ (b), $B \rightarrow K^*(K\pi)\mu^+\mu^-$ where the π is mistaken for the	
	K (c), $B \rightarrow K^*\mu^+\mu^-$ where the K is used in the triplet (d), combinatorial	
	background (e), other b decays (f).	115
5.44	Invariant mass distribution for the B candidates in the muon channel in the	
	high- q^2 bin. The fit is overlaid to the data. The bottom panel shows the pull	
	distribution.	116

5.45	(a) Transverse momentum of the B meson; (b) B meson rapidity. The blue points represent the FONLL, and the red points the soft QCD. Bottom panels show the ratio of FONLL and Soft QCD distributions.	119
5.46	PU profile in MC and data.	120
5.47	Comparison of generated $\frac{dB}{dq^2}$ distribution by PYTHIA with the one predicted by the FLAVIO package for $B^+ \rightarrow K^+ \mu^+ \mu^-$. The ratio points represent the FLAVIO SFs.	121
5.48	$m(\mu\mu)$ distribution from $B^\pm \rightarrow X(\rightarrow \mu\mu)K^\pm$ candidates. In order to determine the dimuon mass distribution from the data we have cut in the analysis BDT working point (BDT output>4) and around the B mass peak ($5.15 < m(\mu\mu K) < 5.45$ GeV) in order to reduce the contamination from partially reconstructed background that might has a different dimuon mass distribution with respect to our signal. The same cuts have been applied in the MC. In order to determine the red distribution from MC, all the corrections and SFs have been removed, while in order to determine the green distribution from MC only the weights from FLAVIO/PYTHIA comparison have been applied.	122
5.49	BDT output calculated with the <i>sPlot</i> technique in the 4 dimuon mass bins: low- q^2 bin (a), high- q^2 bin (b), J/ψ bin (c) and $\psi(2S)$ bin (d).	124
5.50	Comparison of the K^*/K ratios between data and MC for various BDT working points for low- q^2 bin (a), J/ψ bin (b) and $\psi(2S)$ bin (c). The statistical uncertainties from the fit are shown in red, while the systematic branching ratio and statistical uncertainties are shown in green and cyan respectively.	127

5.51	Comparison of the K^*/K ratios between data and MC for various BDT working points for the high- q^2 bin. The statistical uncertainties from the fit are shown in red, while the systematic branching ratio and statistical uncertainties are shown in green and cyan respectively. The K^*/K ratio is presented for different ranges of the $m(\mu\mu K)$ spectrum: (a) $m(\mu\mu K) \in (4.7, 5.7)$ GeV, (b) $m(\mu\mu K) \in (4.9, 5.7)$ GeV, (c) $m(\mu\mu K) \in (4.95, 5.7)$ GeV, (d) $m(\mu\mu K) \in (5.0, 5.7)$ GeV, (e) $m(\mu\mu K) \in (5.05, 5.7)$ GeV and (f) $m(\mu\mu K) \in (5.10, 5.7)$ GeV.	128
5.52	Feature importance of the trained BDT for the PF-PF (a) and PF-LP (b) categories.	138
5.53	Comparing the BDT score distribution of training data against testing data for the PF-PF (a) and PF-LP (b) categories. The distribution shows agreement between training and testing.	140
5.54	Significance as a function of BDT cut for low- q^2 for PF-PF (a) and PF-LP (b).	141
5.55	Fit of simulated $B^+ \rightarrow K^+ e^+ e^-$ (a), same-sign data (b), and simulated $B^+ \rightarrow K^+ J/\psi(e^+ e^-)$ (c) in the low- q^2 bin for PF-PF.	143
5.56	Fit of simulated $B^+ \rightarrow K^+ e^+ e^-$ (a), same-sign data (b), and simulated $B^+ \rightarrow K^+ J/\psi(e^+ e^-)$ (c) in the low- q^2 bin for PF-LP.	144
5.57	Fit to data (top) and pull (bottom) in the low- q^2 bin for the PF-PF (left) and PF-LP (right) category.	145
5.58	Fit of simulated $B^+ \rightarrow K^+ J/\psi(e^+ e^-)$ (a), simulated $B^+ \rightarrow \pi^+ J/\psi(e^+ e^-)$ (b), same-sign data (c), simulated $B^0 \rightarrow K^*(892) J/\psi(e^+ e^-)$ (d), simulated $B^+ \rightarrow K^*(892)^+ J/\psi(e^+ e^-)$ (e), and simulated other B background in J/ψ bin for PF-PF.	148

5.59	Fit of simulated $B^+ \rightarrow K^+ J/\psi(e^+e^-)$ (a), simulated $B^+ \rightarrow \pi^+ J/\psi(e^+e^-)$ (b), same-sign data (c), simulated $B^0 \rightarrow K^*(892)J/\psi(e^+e^-)$ (d), simulated $B^+ \rightarrow K^*(892)^+ J/\psi(e^+e^-)$ (e), and simulated other B background in J/ψ bin for PF-LP.	149
5.60	Fit to data (top) and pull (bottom) in the J/ψ bin for the PF-PF (left) and PF-LP (right) category.	150
5.61	Fit of simulated $B^+ \rightarrow K^+ \psi(2S)(e^+e^-)$ (a), same-sign data (b), simulated other B background (c), simulated $B^0 \rightarrow K^*(892)\psi(2S)(e^+e^-)$ (d), and simulated $B^+ \rightarrow K^*(892)^+ \psi(2S)(e^+e^-)$ (e) in $\psi(2S)$ bin for PF-PF.	153
5.62	Fit of simulated $B^+ \rightarrow K^+ \psi(2S)(e^+e^-)$ (a), same-sign data (b), simulated other B background (c), simulated $B^0 \rightarrow K^*(892)\psi(2S)(e^+e^-)$ (d), and simulated $B^+ \rightarrow K^*(892)^+ \psi(2S)(e^+e^-)$ (e) in $\psi(2S)$ bin for PF-LP.	154
5.63	Fit to data (top) and pull (bottom) in the $\psi(2S)$ bin for the PF-PF (left) and PF-LP (right) category.	155
5.64	Fit of simulated $B^+ \rightarrow K^+ e^+ e^-$ (a), same-sign data (b), simulated $B^0 \rightarrow K^*(892)\psi(2S)(e^+e^-)$ (c), and simulated $B^+ \rightarrow K^+ \psi(2S)(e^+e^-)$ (d) in high- q^2 bin for PF-PF.	158
5.65	Fit of simulated $B^+ \rightarrow K^+ e^+ e^-$ (a), same-sign data (b), simulated $B^0 \rightarrow K^*(892)\psi(2S)(e^+e^-)$ (c), and simulated $B^+ \rightarrow K^+ \psi(2S)(e^+e^-)$ (d) in high- q^2 bin for PF-LP.	159
5.66	Fit to data (top) and pull (bottom) in the high- q^2 bin for the PF-PF (left) and PF-LP (right) category.	160

5.67	Comparison between the BDT output in data and MC for the PF-PF (a) and PF-LP (b) categories in the J/ψ bin. The <i>sPlot</i> technique is used to extract the signal component in data. The dotted line corresponds to the cut applied in the analysis.	165
5.68	BDT distribution in data for the PF-PF (a) and PF-LP (b) category.	165
5.69	Likelihood function from the fit profiled as a function of $R_K^{\psi(2S),e}$. The left, middle, and right shows the profiled likelihood using PF-PF, PF-LP, and combined data.	166
5.70	A comparison of the 68% CL of $R_K^{\psi(2S),e}$ with the SM prediction (black). The left, middle, and right shows the measurements using PF-PF, PF-LP, and combined data. The combined measurement agrees with the SM prediction within 1σ	167
5.71	K^*/K^+ ratio comparison in the J/ψ bin (a), and $\psi(2S)$ bin (b), with PF-PF (top) and PF-LP (bottom) as a function of BDT cut.	168
5.72	A comparison of the 68% CL of $R_K^{\psi(2S),e}$ with the SM prediction (black). The yield of the partially reconstructed background in the fits are fixed to the expected value. The left, middle, and right shows the measurements using PF-PF, PF-LP, and combined data. The fits in the lower sideband have negligible effect on the measurement.	169
5.73	An example of reweighting the subleading electron p_T distribution of $B^+ \rightarrow K^+ J/\psi (\rightarrow e^+ e^-)$ to match that of $B^+ \rightarrow K^+ e^+ e^-$ for the PF-PF (a) and PF-LP (b) category.	177

5.74	Fit after applying weight to the data according to the subleading electron p_T for the PF-PF (a) and PF-LP (b) category in the J/ψ bin.	178
5.75	$R_K^{e, \text{pseudo}}$ as a function of the BDT cut for the PF-PF (a) and PF-LP (b) category in the J/ψ bin.	178
5.76	$R_{\text{trigger}}^{\text{MC}}$ of different HLT paths for PF-PF in full- q^2 (a), low- q^2 (b) and high- q^2 (c). Three rows correspond to working point at preselection (top), loose BDT WP (middle) and tight BDT WP (bottom).	181
5.77	$R_{\text{trigger}}^{\text{MC}}$ of different HLT paths for PF-LP in full- q^2 (a), low- q^2 (b) and high- q^2 (c). Three rows correspond to working point at preselection (top), loose BDT WP (middle) and tight BDT WP (bottom).	182
5.78	$R_{\text{trigger}}^{\text{data}}$ of different HLT paths for PF-PF (a) and PF-LP (b). No significant deviation from unity is observed. The uncertainty from the normalization trigger path HLT_Mu7_IP4 is considered and added to the uncertainties of other paths.	183
5.79	Fits to the B-candidate mass distribution in the J/ψ bin for PF-PF (top) and PF-LP (bottom) category with nominal BDT cut (left) and $BDT > 0$ (right).	184
5.80	Simulated BDT spectrum in the low- q^2 bin and in the J/ψ bin after reweighting, for for PF-PF (left) and PF-LP (right) category.	185
5.81	Fits to the B -candidate mass distribution in the J/ψ bin after weighting as discussed in the text to match the low- q^2 bin, for the for PF-PF (top) and PF-LP (bottom) category with nominal BDT cut (left) and $BDT > 0$ (right).	186
5.82	Distribution of the $ \Delta N_{\text{sig}} /N_{\text{sig}}$ of the 1000 MC toys in the low- q^2 study for PF-PF (a) and PF-LP (b).	189

5.83	Fits in the J/ψ region in the q^2 window of $2.7 < m(e^+e^-) < 3.2$ GeV (left), and $2.5 < m(e^+e^-) < 3.2$ GeV (right) for PF-PF (top) and PF-LP (bottom) categories.	191
5.84	Examples of the MC toys generated for estimating the systematic uncertainty of constraining the J/ψ leakage yields in the low- q^2 region of the electron channel.	193
5.85	Likelihood function from the fit profiled as a function of $R_K^{\psi(2S)}$. The mea- surement agrees with the SM prediction within 1σ	196
5.86	B-candidates mass distribution for events which pass the HLT_Mu9_IP6 trig- ger in the J/ψ bin for PF-PF (a) and PF-LP (b). The fit is overlaid in red.	198
5.87	Likelihood function from the fit profiled as a function of $R_K^{J/\psi}$. The measure- ment agrees with the expected value within 1σ	199
5.88	B-candidates mass distribution for events which pass the HLT_Mu9_IP6 trig- ger in the $\psi(2S)$ bin for PF-PF (a) and PF-LP (b). The fit is overlaid in red.	202
5.89	Likelihood function from the fit profiled as a function of $R_{\psi(2S)}^{\text{single ratio}}$. The expected value is 1.2σ higher than the measured value.	202
5.90	Likelihood function from the fit profiled as a function of R_K^{-1} . The measure- ment agrees with the expected value within 1σ	204
A.0.1	Check of mass sculpting using same-sign data in working point 6.0 (a), 7.0 (b), 7.5 (c), 8.0 (d), 8.5 (e), and 9.0 (f) for PF-PF.	208

A.0.2	Check of mass sculpting using same-sign data in working point 6.0 (a), 7.0 (b), 7.5 (c), 8.0 (d), 8.5 (e), and 9.0 (f) for PF-LP.	209
A.0.3	Check of mass sculpting using combinatorics in $B^+ \rightarrow K^+ J/\psi(e^+e^-)$ MC sample in working point 6.0 (a), 7.0 (b), 7.5 (c), 8.0 (d), 8.5 (e), and 9.0 (f) for PF-PF.	210
A.0.4	Check of mass sculpting using combinatorics in $B^+ \rightarrow K^+ J/\psi(e^+e^-)$ MC sample in working point 6.0 (a), 7.0 (b), 7.5 (c), 8.0 (d), 8.5 (e), and 9.0 (f) for PF-LP.	211
A.0.5	Check of mass sculpting using combinatorics in the generic B decays MC sample in working point 6.0 (a), 7.0 (b), 7.5 (c), 8.0 (d), 8.5 (e), and 9.0 (f) for PF-PF.	212
A.0.6	Check of mass sculpting using combinatorics in the generic B decays MC sample in working point 6.0 (a), 7.0 (b), 7.5 (c), 8.0 (d), 8.5 (e), and 9.0 (f) for PF-LP.	213
B.0.1	Distributions of $m(K^+e^+e^-)$ (a), BDT score (b), $m(e^+e^-)$ (c), and $\min(m(K^+e_{[\rightarrow\pi^-]}^-), m(K_{[\rightarrow\pi^+]}^+e_{[\rightarrow K^-]}^-)$ (d), leading electron ID (e), and subleading electron ID (f) of $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-)\pi^+$ simulated events for PF-PF after pre-selection. .	215
B.0.2	Distributions of $m(K^+e^+e^-)$ (a), BDT score (b), $m(e^+e^-)$ (c), and $\min(m(K^+e_{[\rightarrow\pi^-]}^-), m(K_{[\rightarrow\pi^+]}^+e_{[\rightarrow K^-]}^-)$ (d), leading electron ID (e), and subleading electron ID (f) of $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-)\pi^+$ simulated events for PF-LP after pre-selection. .	216

B.0.3	Distributions of $m(K^+e^+e^-)$ (a), BDT score (b), $m(e^+e^-)$ (c), and $\min(m(K^+e_{[\rightarrow\pi^-]}^-)$, $m(K_{[\rightarrow\pi^+]}^+e_{[\rightarrow K^-]}^-)$ (d), leading electron ID (e), and subleading electron ID (f) of $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-)\pi^+$ simulated events for PF-LP after pre-selection without applying the electron ID cut.	217
B.0.4	Distribution of $m(K^+e^+e^-)$ of $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-)\pi^+$ simulated events for PF-PF (a), and PF-LP (b) after nominal BDT cut.	218
B.0.5	Distribution of $m(K^+e^+e^-)$ (a) and anti- D^0 variable of generated $B^+ \rightarrow$ $K^+\pi^-\pi^+$ simulated events	218
C.0.1	Fit performed with an alternative signal function that uses four Gaussians.	220
C.0.2	Fit performed with an alternative background function that uses a sum of Gaussian and polynomial.	220
C.0.3	Fit performed with alternative signal function using two gaussians and one double-sided Crystal Ball in MC (a) and in Data (b) in the region $q^2 \in$ $(1.1, 6.0)$ GeV ²	221
C.0.4	Fit performed with alternative background function using one gaussian and a first order polynomial in the region $q^2 \in (1.1, 6.0)$ GeV ²	221
D.0.1	Alternative fitting shape in the J/ψ region for signal (a), combinatorics (b), K^{*0} partially reconstructed background (c), K^{*+} partially reconstructed back- ground (d), other B components (e), and $J/\psi\pi$ contamination (f) for PF-PF.	223
D.0.2	Alternative fitting shape in the J/ψ region for signal (a), combinatorics (b), K^{*0} partially reconstructed background (c), K^{*+} partially reconstructed back- ground (d), other B components (e), and $J/\psi\pi$ contamination (f) for PF-LP.	224

D.0.3	Fits in the J/ψ region with alternative fitting shape for signal (a), combinatorics (b), partially reconstructed background (c), other B components (d), and $J/\psi\pi$ contamination (e) for PF-PF.	225
D.0.4	Fits in the J/ψ region with alternative fitting shape for signal (a), combinatorics (b), partially reconstructed background (c), other B components (d), and $J/\psi\pi$ contamination (e) for PF-LP.	226
D.0.5	An example of a MC toy generated in the low- q^2 region for PF-PF (a), and PF-LP (b) for the study of fit parameterization systematics.	227
D.0.6	An example of the alternative fitting shape in the low- q^2 region for signal (a), combinatorics (b), and J/ψ leakage for PF-PF.	227
D.0.7	An example of the alternative fitting shape in the low- q^2 region for signal (a), combinatorics (b), and J/ψ leakage for PF-LP.	228
D.0.8	An example of the fits in the low- q^2 region with nominal fit (a), alternative fitting shape for signal (b), combinatorics (c), and J/ψ leakage (d) for PF-PF.	229
D.0.9	An example of the fits in the low- q^2 region with nominal fit (a), alternative fitting shape for signal (b), combinatorics (c), and J/ψ leakage (d) for PF-LP.	230

Chapter 1

Introduction

The Standard Model of particle physics [1–3] combines the results of decades of research, which describes elementary particles and their interactions in an elegant mathematical framework. It has been tested with exceptional precision in the past few decades, including the Higgs boson discovery in 2012 [4, 5]. Despite the success of the Standard Model, it leaves a series of unexplained phenomena, including several anomalies in the B -physics sector, which may hint at a possible violation of the Lepton Universality principle.

Lepton Universality is an accidental symmetry of the Standard Model, in which the coupling of the electroweak gauge bosons to the leptons is independent of the lepton flavor. This can be tested using the decays of B hadrons involving a loop-level transition of $b \rightarrow s\ell^+\ell^-$. In recent years, several observables have shown discrepancies with the Standard Model predictions, known as B flavor anomalies. This may be a sign of New Physics beyond the Standard Model.

For the CMS experiment to establish a role in exploring potential Lepton Universality violation, a novel trigger strategy with data parking was proposed and deployed in 2018,

named "B-parking." The B-parking data-taking period collected $\mathcal{O}(10^{10})$ unbiased B hadron decays, making possible to perform several rare B decay searches reaching branching fractions down to $\mathcal{O}(10^{-7})$, including the R_K measurement. This dissertation describes the details of the B-parking trigger strategy. Afterward, the test of lepton universality using $B^+ \rightarrow K^+ \ell^+ \ell^-$ using B-parking data at CMS is presented.

This thesis is organized as follows: Chapter 2 gives a brief review of the Standard Model, providing a theoretical overview of the Standard Model. Chapter 3 describes the motivation of the test of Lepton Universality using $b \rightarrow s \ell^+ \ell^-$ transition. Chapter 4 introduces the CMS experiment and its apparatus. Chapter 5 discusses the B-parking trigger strategy as well as the analysis of R_K measurement.

Chapter 2

The Standard Model

The Standard Model of particle physics is a theoretical framework that describes the fundamental particles and interactions in the universe. It is a quantum field theory built upon the gauge symmetry principles of the electroweak and strong interaction. The Standard Model consists of three main forces: the strong force, the weak force, and the electromagnetic force. These forces are mediated by the exchange of gauge bosons, specifically the gluons, $W^\pm$, and Z^0 bosons, and photons. The Standard Model also includes the Higgs field, which is responsible for giving particles mass through the Higgs mechanism. Six types of quarks and leptons are the basic building blocks that make up matter in the universe.

In this chapter, we will overview the particle content of the Standard Model. Then, we will review the theory in the electroweak sector and quantum chromodynamics. Finally, we will go over the Higgs bosons and the Higgs mechanism in the Standard Model.

2.1 The Elementary Particles of the Standard Model

The Standard Model has 17 types of fundamental particles with 118 degrees of freedom including color and spin degrees of freedom, as well as antiparticles ($2 \times [6 \text{ quarks} \times 3 \text{ colors} \times 2 \text{ helicities} + 3 \text{ charged leptons} \times 2 \text{ helicities} + 3 \text{ neutrinos}] + \text{gluon} \times 8 \text{ colors} \times 2 \text{ helicities} + \text{photon} \times 2 \text{ helicities} + 3 \text{ heavy gauge bosons} \times 3 \text{ helicities} + \text{Higgs boson} = 2 \times [36 + 6 + 3] + 16 + 2 + 9 + 1 = 2 \times 45 + 28 = 118$). They are the building blocks of the universe. Fig. 2.1 summarizes the particle content of the Standard Model. These particles are divided into two main categories: fermions and bosons. Fermions of the Standard Model are spin-1/2 particles that are the building blocks of matter and include quarks and leptons. The Standard Model bosons either have spin-1 known as gauge boson that are responsible for force mediation, or spin-0 known as the Higgs boson that plays an essential role in mass generation.

2.1.1 Bosons

A boson is a type of subatomic particle characterized by its integer-valued spin, which follows the Bose-Einstein statistics. In the Standard Model, there are two types of fundamental particles that are bosons: gauge bosons and the Higgs boson.

Gauge bosons are spin-1 particles, and they are responsible for transmitting fundamental forces. The Standard Model describes three fundamental forces in the universe, electromagnetic force, weak force, and strong force. A photon is a massless and chargeless particle that is the electromagnetic force carrier. $W^\pm$ and Z^0 bosons are the force carriers of the weak interaction. They are massive, where $W^\pm$ bosons have an electron charge of ± 1 and the Z^0

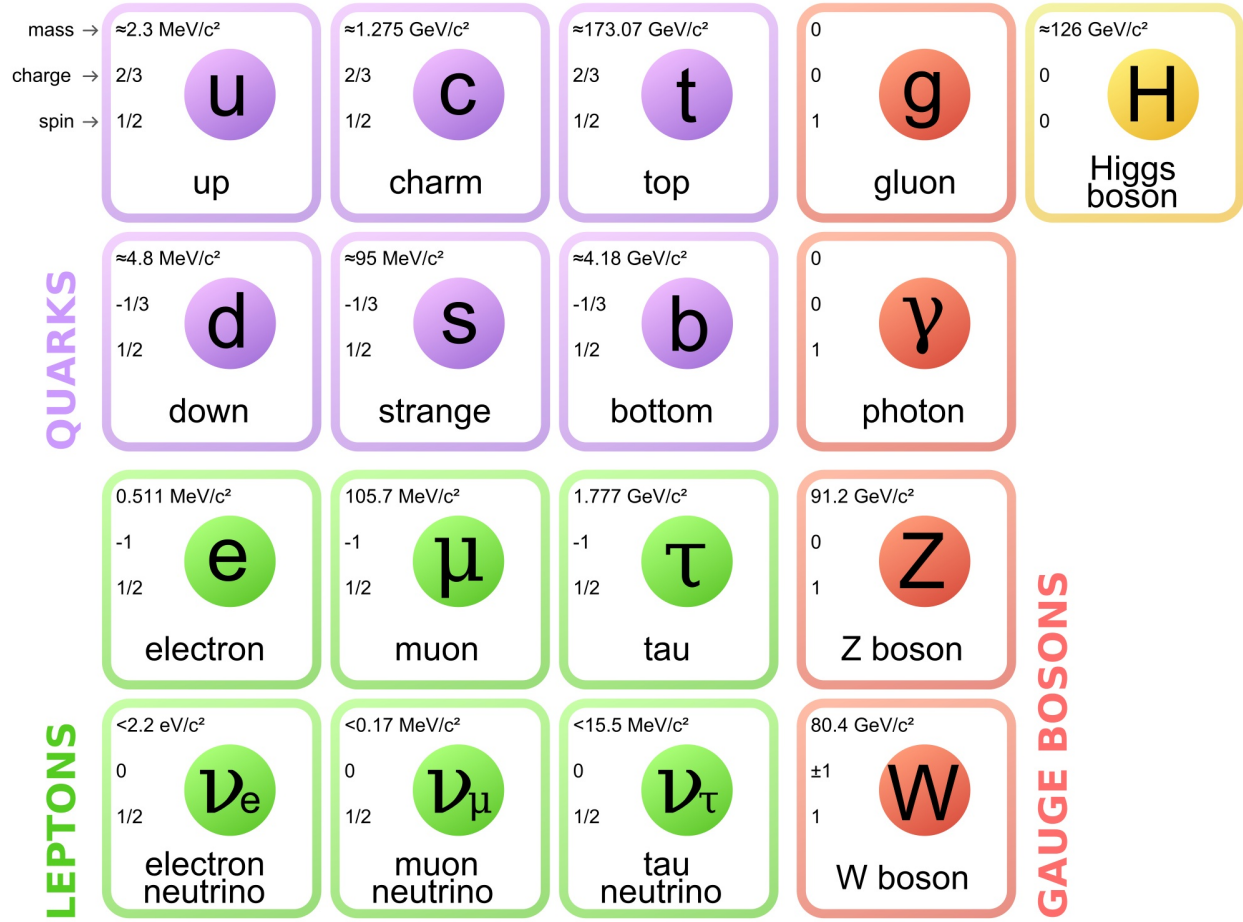


Figure 2.1: Particle content of the Standard Model [6].

boson is neutral. Gluons are massless gauge bosons and with a neutral electric charge. However, they carry a superposition of color charges allowing them to carry strong force through the exchange of color charges and self-interactions. There are 8 different gluons depending on the combination of color charges. The coupling strengths of the forces are summarized in Table 2.1.

Interaction	Force carriers	Coupling strength
Electromagnetic	photon	$1/137$
Weak	$W^\pm$ and Z	10^{-6}
Strong	gluons	0.1

Table 2.1: Summary of the coupling strengths of fundamental interactions in the SM.

The Higgs boson was first proposed in 1960s by theorists [7–12] and was discovered in 2012 by ATLAS and CMS [4, 5] at the LHC. The Higgs boson is a scalar particle that plays a role of giving masses to both weak gauge bosons, quarks, and leptons through the interactions with the Higgs field. The $W^\pm$ and Z^0 bosons acquire masses from spontaneously symmetry breaking in the electroweak sector. The quarks and the leptons acquire masses via the Yukawa coupling with the Higgs field.

2.1.2 Fermions

Fermions are particles with half-integer spin, which follow the Fermi-Dirac statistics. In the Standard Model, there are two types of fundamental particles that are fermions: quarks and leptons.

Quarks are fundamental particles that make up mesons and baryons, including protons, neutrons, kaons, and pions. Mesons consist of quark-antiquark pairs, while baryons are composed of three quarks. Quarks come in six different "flavors" divided into three generations: up (u), down (d), charm (c), strange (s), top (t), and bottom (b). Quarks have fractional electric charges based on their flavor: $+2/3$ for u , c , t , and $-1/3$ for d , s , b . Quarks also have a color charge, including red (r), blue (b), green (g), and their anti-color counterparts. Quarks interact through the electromagnetic, weak, and strong forces. Weak interactions allow quarks to change flavors, which are parameterized by the Cabibbo-Kobayashi-Maskawa (CKM) matrix. Quarks experience color confinement, meaning they must combine with other quarks (and gluons) to form color-neutral hadrons. Therefore, free quarks are not observed in nature. Asymptotic freedom is another property of quarks, where the strong

force weakens as quarks come closer together. This effect makes the strong force effectively short-ranged, and QCD calculations are only perturbative at high energies when the coupling is small.

There are three generations of six types of leptons: electron (e), muon (μ), tau (τ), electron neutrino (ν_e), muon neutrino (ν_μ), and tau neutrino (ν_τ). Unlike quarks, leptons do not have color charge, which means they do not interact with the strong force. The electric charge of e , μ , and τ is 1. Leptons interact through the electromagnetic and weak forces. Neutrinos are electrically neutral and only interact through the weak force. Unlike quarks and charged leptons, neutrinos have only one possible polarization, which is left-handed (and antineutrinos are only right-handed) due to their coupling with weak interaction and the violation of parity. A phenomenon not explained by the Standard Model is that neutrinos have a very small but non-zero mass, which results in neutrino oscillation. The flavor states of neutrinos do not align with their mass states, meaning that the lepton flavor of a neutrino can change between its creation and measurement. The mixing of neutrinos is parameterized by the Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix.

2.2 Theoretical framework of the Standard Model

2.2.1 Quantum Electrodynamics (QED)

Quantum electrodynamics (QED) is a relativistic quantum field theory that describes how electromagnetic forces arise from the exchange of photons between charged particles, and how these forces affect the behavior of matter and energy. It is an abelian gauge theory with

the symmetry group $U(1)$.

The Lagrangian density for QED is given by

$$\mathcal{L}_{\text{QED}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\gamma^\mu D_\mu - m)\psi, \quad (2.1)$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field tensor, $D_\mu = \partial_\mu + ieA_\mu$ is the gauge covariant derivative, A_μ is the electromagnetic vector potential, ψ is the bispinor field of the particle, e is the coupling constant equal to the charge of the particle, and m is the particle's mass.

The QED Lagrangian can be understood as a sum of the Lagrangian of Dirac scalar field $\mathcal{L}_{\text{Dirac}}$, the Lagrangian of classical electromagnetic $\mathcal{L}_{\text{Maxwell}}$, and the Lagrangian of interaction $\mathcal{L}_{\text{int}}$, where

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi, \quad \mathcal{L}_{\text{Maxwell}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \quad \mathcal{L}_{\text{int}} = -e\bar{\psi}\gamma^\mu\psi A_\mu. \quad (2.2)$$

Note that a standalone Dirac field does not satisfy local gauge symmetry. The gauge field and its interaction with the particle are always required such that the QED Lagrangian satisfies local $U(1)$ symmetry, where it is gauge invariant under the transformation gauge

$$\psi \rightarrow e^{i\alpha}\psi, \quad (2.3)$$

and

$$A_\mu \rightarrow A_\mu - \frac{1}{e}\partial_\mu\alpha, \quad (2.4)$$

where α is an arbitrary function.

2.2.2 Quantum Chromodynamics (QCD)

Quantum chromodynamics (QCD) is the theory of the strong interaction between quarks mediated by gluons. It is a non-abelian gauge theory with the symmetry group $SU(3)$.

The Lagrangian density for QCD is given by

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu} + \bar{\psi}_i (i\gamma^\mu (D_\mu)_{ij} - m\delta_{ij}) \psi_j, \quad (2.5)$$

where $G_{\mu\nu}^a = \partial_\mu \mathcal{A}_\nu^a - \partial_\nu \mathcal{A}_\mu^a + gf^{abc} \mathcal{A}_\mu^b \mathcal{A}_\nu^c$ is the gluon field strength tensor, $(D_\mu)_{ij} = \partial_\mu \delta_{ij} - ig(T_a)_{ij} \mathcal{A}_\mu^a$ is the gauge covariant derivative, $\mathcal{A}_\mu$ is the gluon potential, T_a are the generators of $SU(3)$, f^{abc} are the structure constants of $SU(3)$ indexed by a, b and c running from 1 to 8, ψ is the quark field, g is the coupling constant, and m is the particle's mass.

The QCD Lagrangian obeys the local gauge invariant in color space. The quark fields ψ are indexed by i and j running from 1 to 3, representing the color charges of r, g, b . We can use Gell-Mann matrices λ_a as an explicit representation of the generators of $SU(3)$ T_a , defined by

$$T_a = \frac{1}{2}\lambda_a, \quad a = 1, \dots, 8, \quad (2.6)$$

where T_a satisfy the commutation relations

$$[T_a, T_b] = if_{abc}T_c, \quad (2.7)$$

and therefore the QCD Lagrangian satisfy $SU(3)$ gauge invariance with transformation of

$$\psi \rightarrow e^{-i\lambda \cdot \mathbf{a}} \psi. \quad (2.8)$$

A crucial property of QCD is asymptotic freedom. From the renormalization equation, we can see that the coupling constant (α_s) of strong force decreases as the energy scale increases, which is given by

$$\alpha_s(Q) = \frac{2\pi}{(11 - (2/3)n_f) \ln \frac{Q}{\Lambda_{\text{QCD}}}}, \quad (2.9)$$

where Q is the energy scale, $n_f = 6$ is the number of quark flavors, Λ_{QCD} is the energy scale of QCD. Therefore, the strong force is short-ranged, and the QCD calculations at low energies are challenging as the QCD expansions converge slowly.

2.2.3 Electroweak theory

Unlike electromagnetic and strong forces, the weak force carriers $W^\pm$ and Z^0 bosons are massive. However, massive gauge bosons do not preserve local gauge invariance. Sheldon Glashow, Abdus Salam, and Steven Weinberg [1–3] introduced a theory that unites the electromagnetic and weak interactions and gives mass to the gauge bosons through spontaneous symmetry breaking by adding a scalar field called the Higgs field. The electroweak interaction is described by a non-abelian gauge theory with $SU(2)_L \times U(1)$ symmetry group. Below the unification energy at around 200 GeV, the $W^\pm$ and Z^0 bosons and the photon are produced from the spontaneous breaking of the electroweak symmetry to $U(1)_{em}$ symmetry

with the presence of the Higgs field.

The electroweak Lagrangian before electroweak symmetry breaking contains 4 parts

$$\mathcal{L}_{\text{EW}} = \mathcal{L}_g + \mathcal{L}_f + \mathcal{L}_h + \mathcal{L}_y. \quad (2.10)$$

The Lagrangian $\mathcal{L}_g$ describes the gauge fields

$$\mathcal{L}_g = -\frac{1}{4}W_a^{\mu\nu}W_{\mu\nu}^a - \frac{1}{4}B^{\mu\nu}B_{\mu\nu}, \quad (2.11)$$

where $W^{a\mu\nu}$ and $B^{\mu\nu}$ are the field strength tensors for the weak isospin and weak hypercharge gauge fields.

The Lagrangian $\mathcal{L}_f$ describes interaction of the gauge bosons and the fermions

$$\mathcal{L}_f = \bar{Q}_j i\gamma^\mu D_\mu Q_j + \bar{u}_j i\gamma^\mu D_\mu u_j + \bar{d}_j i\gamma^\mu D_\mu d_j + \bar{L}_j i\gamma^\mu D_\mu L_j + \bar{e}_j i\gamma^\mu D_\mu e_j, \quad (2.12)$$

where the subscript j sums over the three generations of fermions, Q_j are the left-handed doublet quark fields, u_j are the right-handed singlet up quark fields, d_j are the right handed singlet down quark fields, L_j are the left-handed doublet charged lepton fields, e_j are the right-handed singlet charged lepton fields, and the covariant derivative is defined as

$$D_\mu = \partial_\mu - i\frac{g}{2}\sigma_j W_\mu^j - i\frac{g'}{2}Y B_\mu, \quad (2.13)$$

where Y is the weak hypercharge, σ_j are the Pauli matrices, W_μ^j and B_μ are, respectively, the $SU(2)$ and $U(1)$ gauge bosons. The g and g' are the coupling constants of $SU(2)$ and

$U(1)$ respectively, which are related by the electroweak mixing angle θ_W by

$$g' \cos(\theta_W) = g \sin(\theta_W). \quad (2.14)$$

The Lagrangian $\mathcal{L}_h$ describes the Higgs field ϕ kinetic and the potential energies

$$\mathcal{L}_h = |D_\mu \phi|^2 - \mu^2 \phi^\dagger \phi + \frac{\lambda}{4} (\phi^\dagger \phi)^2, \quad (2.15)$$

where $V(\phi) = -\mu^2 \phi^\dagger \phi + \frac{\lambda}{4} (\phi^\dagger \phi)^2$ denotes the potential energy of the Higgs field.

The final Lagrangian $\mathcal{L}_y$ is the Yukawa interaction term, which describes

$$\mathcal{L}_y = -y_{uij} \epsilon^{ab} \phi_b^\dagger \bar{Q}_{ia} u_j^c - y_{dij} \phi \bar{Q}_i d_j^c - y_{eij} \phi \bar{L}_i e_j^c + h.c. , \quad (2.16)$$

where y_{uij} , y_{dij} , and y_{eij} are the Yukawa couplings between the Higgs and the fermions, which generate the fermions' masses.

2.2.4 Higgs mechanism

The Higgs mechanism is essential for mass generation in the Standard Model. The electroweak theory incorporates both local gauge invariance and spontaneous symmetry breaking such that the gauge vector bosons acquire masses but are still constrained by the underlying gauge symmetry. Below the unification energy, the Higgs field causes spontaneous symmetry breaking during interactions, triggering the gauge bosons to interact with it to have mass.

Consider the Higgs field as defined as the $SU(2)$ doublet:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \quad (2.17)$$

where ϕ^+ and ϕ^- are complex fields.

The Lagrangian of the Higgs term $\mathcal{L}_h$ consists of the kinetic term and the potential term:

$$\mathcal{L}_h = |D_\mu \phi|^2 - V(\phi) \quad (2.18)$$

$$= |D_\mu \phi|^2 - \mu^2 \phi^\dagger \phi + \frac{\lambda}{4} (\phi^\dagger \phi)^2. \quad (2.19)$$

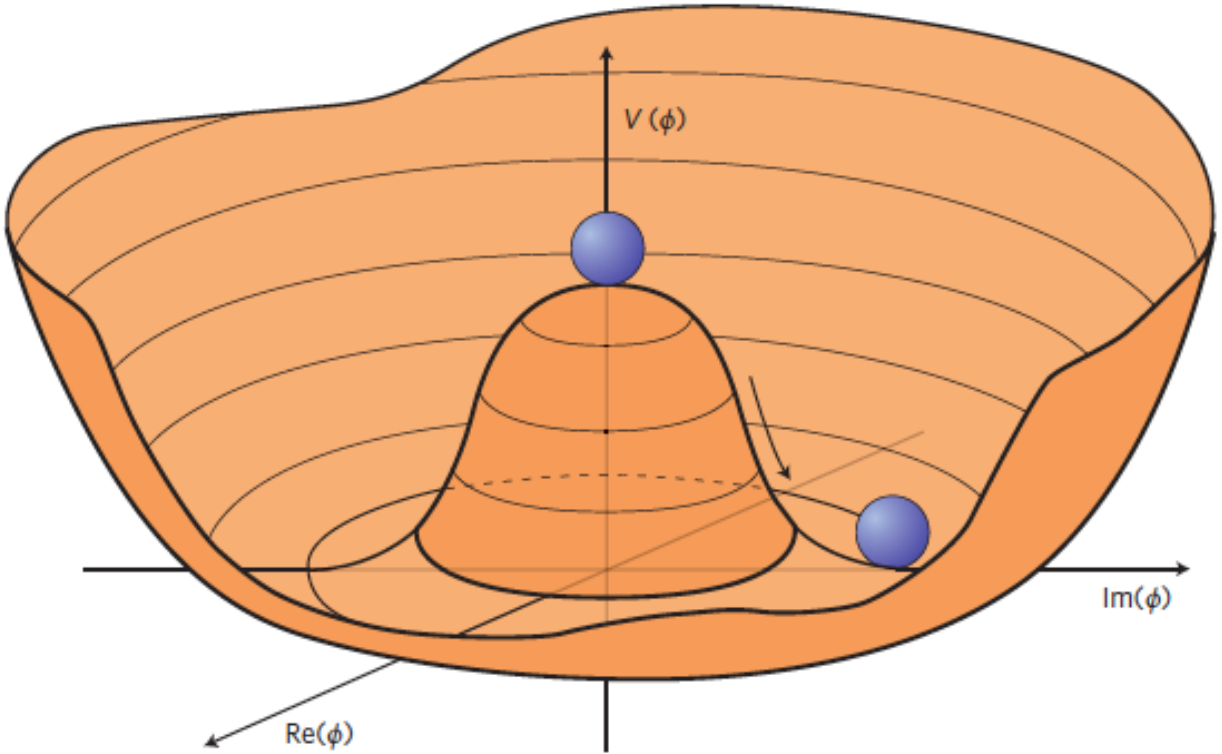


Figure 2.2: The Higgs potential [13].

If $\mu^2 < 0$, the potential has a minimum at 0. But if $\mu^2 > 0$, it takes the form of the

”Mexican hat” potential. Figure 2.2 illustrates the shape of the potential. The Higgs field acquires a non-zero vacuum expectation value (VEV) at the minimum

$$\langle \phi \rangle = \frac{1}{2} \begin{pmatrix} 0 \\ \nu \end{pmatrix}, \quad \nu = \frac{2\mu}{\lambda}, \quad (2.20)$$

and the global gauge symmetry is spontaneously broken.

To study the Lagrangian around VEV, we parametrize the Higgs field around the minimum ν as

$$\phi = \frac{1}{\sqrt{2}} e^{i\theta_a \sigma^a / \nu} \begin{pmatrix} 0 \\ \nu + h \end{pmatrix}, \quad (2.21)$$

where θ_a is the Goldstone boson and h is the Higgs boson. We recognize that $e^{i\theta_a \sigma^a / \nu}$ is just the $SU(2)_L$ gauge transform. We can choose the Unitary gauge for convenient such that

$$\phi' = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu + h \end{pmatrix}. \quad (2.22)$$

The kinetic term of Eq. 2.19 becomes

$$|D_\mu \phi|^2 = \frac{g^2 \nu^2}{4} W_\mu^+ W^{-\mu} + \frac{(g^2 + g'^2) \nu^2}{8} Z_\mu Z^\mu \quad (2.23)$$

where $W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2)$ and $Z_\mu = \frac{1}{\sqrt{g^2 + g'^2}}(gW_\mu^3 - g'B_\mu)$. We can then identify the mass terms for the $W^\pm$ and Z^0 bosons which yield the mass of

$$M_W = \frac{g\nu}{2} \quad (2.24)$$

$$M_Z = \frac{\sqrt{(g^2 + g'^2)}\nu}{2} = \frac{M_W}{\cos \theta_W}. \quad (2.25)$$

The potential term of Eq. 2.19 becomes

$$V(\phi) = \frac{\mu^4}{4\lambda} + \mu^2 h^2 + \lambda \nu h^3 + \frac{1}{4} \lambda h^4, \quad (2.26)$$

which gives the mass of the Higgs boson as $M_H = \sqrt{2}|\mu|$.

Chapter 3

Overview of the test of Lepton Universality in B -physics sector

This thesis focuses on the analysis of the test of Lepton Universality using $B^+ \rightarrow K^+ \ell^+ \ell^-$ decays. In this chapter, we will first discuss the motivation behind this analysis. Then, we will go over the theoretical framework used in the B -physics sector. Finally, we will summarize the recent B flavor anomalies found in the experiments.

3.1 The $b \rightarrow s \ell^+ \ell^-$ transition as a probe of New Physics

Despite the success of the Standard Model, it leaves many unexplained phenomena. For example, the observed value of the Higgs mass gives rise to sizeable fine-tuning at an energy scale lower than the Planck scale, known as the hierarchy problem. The Standard Model does not describe or explain dark matter. Neutrino oscillation shows neutrinos have non-zero masses, while the Standard Model predicts massless neutrinos. Also, the Standard Model

does not explain gravitation. Therefore, the Standard Model is considered as incomplete and scientists are looking for signs of New Physics (NP).

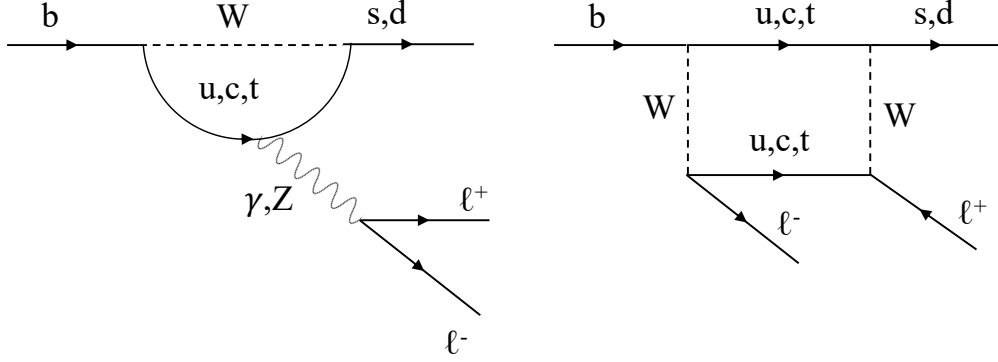


Figure 3.1: Feynman diagrams contributing to the $b \rightarrow s\ell^+\ell^-$ transition

Heavy Flavor physics is a strong indirect probe for signs of New Physics. Among all processes, the $b \rightarrow s\ell^+\ell^-$ transition is a promising channel for searching for NP. Figure 3.1 shows the Feynman diagrams contributing to the $b \rightarrow s\ell^+\ell^-$ transition. It is a Flavor Changing Neutral Current (FCNC), which has no diagram at the tree level. Second, this process is Glashow-Iliopoulos-Maiani (GIM) suppressed at the loop level. Any NP effects from tree-level diagram can easily be spotted. Third, the large value of the b -quark mass above the QCD scale gives precise predictions from the SM for the observables.

3.2 Effective Hamiltonian for $b \rightarrow s\ell^+\ell^-$ transition

Effective field theory provides a model-independent framework to parameterize the observables in B physics. The effective field theory generalizes the Fermi theory of weak interactions, which handles the short-distance and long-distance effects in terms of Wilson coefficients and operators, respectively.

The effective Hamiltonian for FCNC transitions takes the form of

$$\mathcal{H}_{\text{eff}}(b \rightarrow s\ell^+\ell^-) = -\frac{4G_F}{\sqrt{2}}V_{tb}V_{ts}^* \sum_i \mathcal{C}_i \mathcal{O}_i, \quad (3.1)$$

where G_F is the Fermi's coupling constant, V_{ij} are the CKM matrix elements, $\mathcal{O}_i$ are the operators, and $\mathcal{C}_i$ are the Wilson coefficients which contain all information about short-distance physics above the renormalization scale μ .

The relevant operators in weak decays can be divided into six classes [14]: Operators $\mathcal{O}_1$ and $\mathcal{O}_2$ are current-current operators, operators $\mathcal{O}_3, \dots, \mathcal{O}_6$ are QCD-penguin operators, operators $\mathcal{O}_7$ and $\mathcal{O}_8$ are electro- and chromo-magnetic penguin operators, and $\mathcal{O}_9$ and $\mathcal{O}_{10}$ are semi-leptonic operators. The dominant operators in the SM relevant to $b \rightarrow s\ell^+\ell^-$ decays are [15]

$$\mathcal{O}_7 = \frac{e}{16\pi^2} m_b (\bar{s} \sigma_{\mu\nu} P_R b) F^{\mu\nu}, \quad (3.2)$$

$$\mathcal{O}_{9\ell} = \frac{e^2}{16\pi^2} (\bar{s} \gamma_\mu P_L b) (\bar{\ell} \gamma^\mu \ell), \quad (3.3)$$

$$\mathcal{O}_{10\ell} = \frac{e^2}{16\pi^2} (\bar{s} \gamma_\mu P_L b) (\bar{\ell} \gamma^\mu \gamma_5 \ell), \quad (3.4)$$

where $\sigma_{\mu\nu} = \frac{i}{2}[\gamma_\mu, \gamma_\nu]$, $P_{L,R} = \frac{1 \mp \gamma_5}{2}$ are the left- and right-handed chiral projectors, and $F^{\mu\nu}$ is the electromagnetic tensor. Operator $\mathcal{O}_7$ describes the $b \rightarrow s\gamma$ transition with an on-shell photon. Operators $\mathcal{O}_{9\ell}$ and $\mathcal{O}_{10\ell}$ are the vectorial and axial operators describing $b \rightarrow s\ell^+\ell^-$ transition. The corresponding diagrams are shown in Fig. 3.2.

The effective Hamiltonian framework provides a convenient way to describe a high-energy theory by parameterizing it using Wilson coefficients. We can conduct experiments to con-

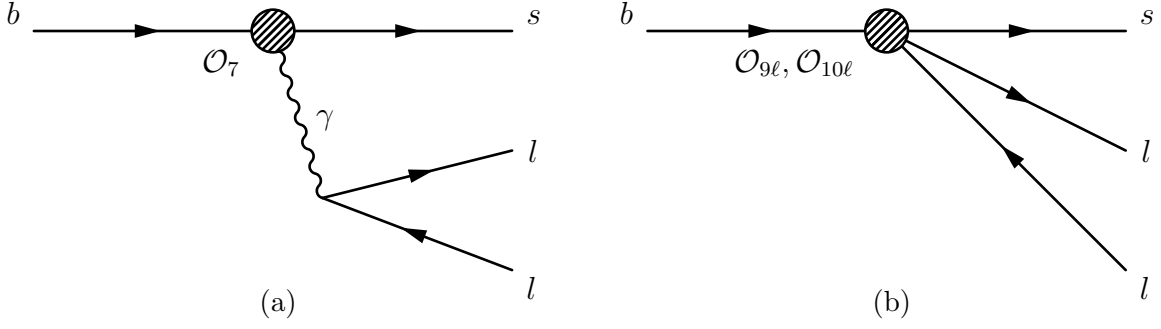


Figure 3.2: Diagrams for $\mathcal{O}_7$ (a) and $\mathcal{O}_{9\ell,10\ell}$ (b) operators.

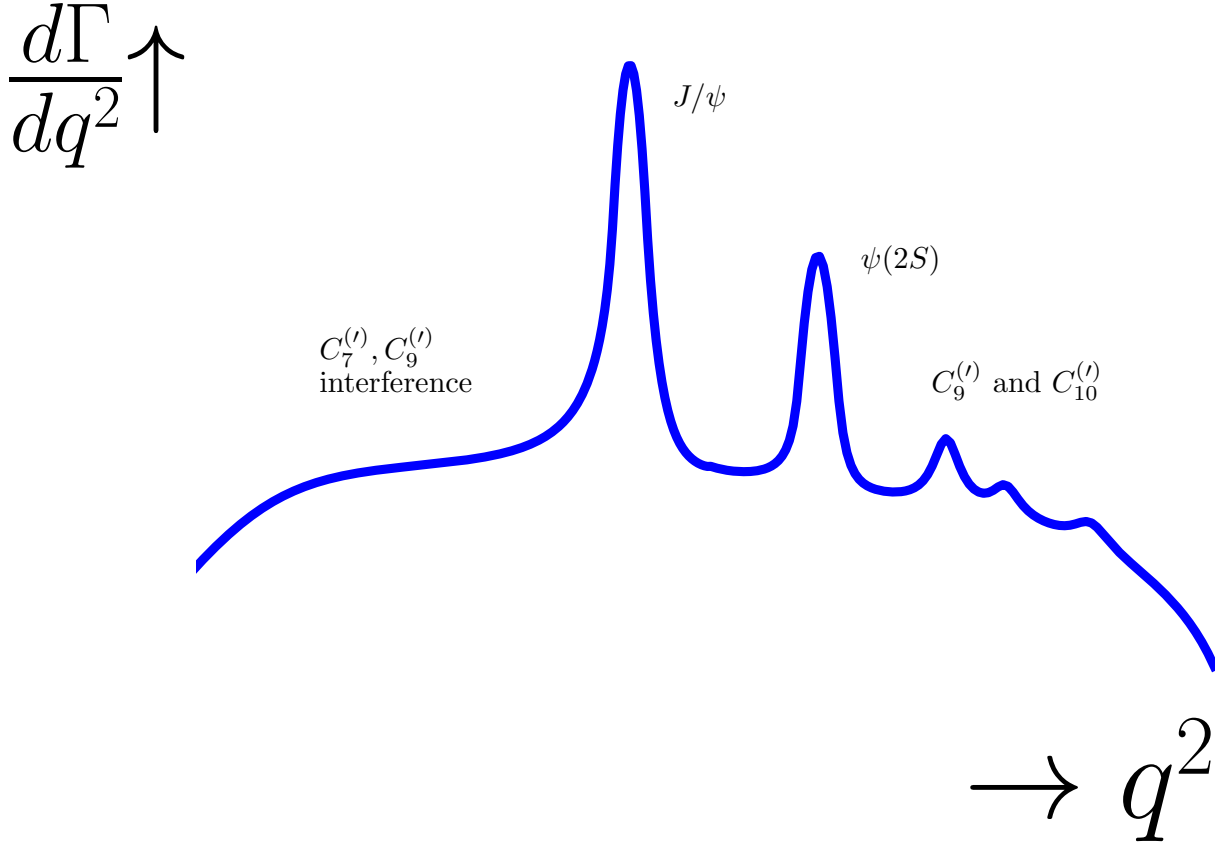


Figure 3.3: A schematic diagram of the differential branching fractions as a function q^2 for $B^+ \rightarrow K^+ \ell^+ \ell^-$ and the charmonium resonances.

strain the Wilson coefficients and study if there are any derivations from the SM. The $b \rightarrow s \ell^+ \ell^-$ transition is sensitive to $\mathcal{O}_7$, $\mathcal{O}_9$ and $\mathcal{O}_{10}$, and the importance of each of the operators depends on the transferred momentum q^2 in the $b \rightarrow s$ transition. Figure 3.3 shows the differential rate of the $B^+ \rightarrow K^+ \ell^+ \ell^-$ decay as well as the charmonium peaks. In the

region below the J/ψ resonance, denoted as low- q^2 , the rate is dominated by the interference between Wilson coefficients $\mathcal{C}_7$ and $\mathcal{C}_9$. In the region above the $\psi(2S)$ resonance, which is known as high- q^2 , it is mainly characterized by Wilson coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$. This region is contaminated by several higher charmonium resonances such as $\psi(3770)$ and $\psi(4040)$.

3.3 Lepton Universality in the Standard Model

In the Standard Model, the electroweak interaction strengths for the three lepton generations are equal. The three lepton generations are essentially the same except for their masses. This accidental symmetry in the Standard Model is known as Lepton Universality (LU). Violation of LU can be a sign of New Physics, where the NP particles couple to the leptons with non-universal interactions.

LU has been tested extensively for decades and no deviations from the Standard Model have been observed. For example, the measurements of the $Z \rightarrow e^+e^-$, $Z \rightarrow \mu^+\mu^-$, and $Z \rightarrow \tau^+\tau^-$ partial widths show excellent agreement with each other [16]:

$$\frac{\Gamma_{Z \rightarrow \mu^+\mu^-}}{\Gamma_{Z \rightarrow e^+e^-}} = 1.0009 \pm 0.0028, \quad (3.5)$$

$$\frac{\Gamma_{Z \rightarrow \tau^+\tau^-}}{\Gamma_{Z \rightarrow e^+e^-}} = 1.0019 \pm 0.0032. \quad (3.6)$$

The partial widths of $W^- \rightarrow e^-\bar{\nu}_e$, $W^- \rightarrow \mu^-\bar{\nu}_\mu$ and $W^- \rightarrow \tau^-\bar{\nu}_\tau$ also show excellent

agreement to the Standard Model predictions [17]:

$$\frac{\Gamma_{W^- \rightarrow \tau^- \bar{\nu}_\tau}}{\Gamma_{W^- \rightarrow e^- \bar{\nu}_e}} = 1.063 \pm 0.027, \quad (3.7)$$

$$\frac{\Gamma_{W^- \rightarrow \tau^- \bar{\nu}_\tau}}{\Gamma_{W^- \rightarrow \mu^- \bar{\nu}_\mu}} = 1.070 \pm 0.026. \quad (3.8)$$

Leptonic decays of pseudoscalar mesons also provide a strong test of LU. For example, the ratio of the partial decay widths of $K^- \rightarrow e^- \bar{\nu}_e$ and $K^- \rightarrow \mu^- \bar{\nu}_\mu$ is measured to be [18]

$$\frac{\Gamma_{K^- \rightarrow e^- \bar{\nu}_e}}{\Gamma_{K^- \rightarrow \mu^- \bar{\nu}_\mu}} = (2.488 \pm 0.009) \times 10^{-5}, \quad (3.9)$$

which agrees with the SM prediction of $(2.477 \pm 0.001) \times 10^{-5}$ [19]. The same ratio for pion has been measured to be [20]

$$\frac{\Gamma_{\pi^- \rightarrow e^- \bar{\nu}_e}}{\Gamma_{\pi^- \rightarrow \mu^- \bar{\nu}_\mu}} = (1.230 \pm 0.004) \times 10^{-4}, \quad (3.10)$$

which also shows good agreement with the SM prediction of $(1.2352 \pm 0.0001) \times 10^{-4}$ [19].

The pure leptonic decays of the τ lepton can also be used for the test of LU. The Standard Model predicts that the decays of $\tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau$ and $\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$ are identical up to corrections due to the difference of mass. The ratios of the charged-current couplings g_l between generations are measured to be [15, 21]

$$g_\mu/g_e = 1.0018 \pm 0.0014, \quad (3.11)$$

$$g_\tau/g_\mu = 1.0011 \pm 0.0015, \quad (3.12)$$

$$g_\tau/g_e = 1.0030 \pm 0.0015. \quad (3.13)$$

The experimental results show excellent agreement with the SM. This is one of the most stringent tests of LU today [15].

Another type of decay that can be used for a test of LU is the leptonic decay of quarkonia resonances. The most precise test is obtained from the ratio of the partial widths of $J/\psi \rightarrow e^+e^-$ and $J/\psi \rightarrow \mu^+\mu^-$ [18]

$$\frac{\Gamma_{J/\psi \rightarrow e^+e^-}}{\Gamma_{J/\psi \rightarrow \mu^+\mu^-}} = 1.0016 \pm 0.0031, \quad (3.14)$$

which agrees with the SM prediction of unity up to 0.31%.

3.4 B flavor anomalies and Lepton-Flavor-Universality Violation

Despite many excellent agreements between the experimental results and the SM predictions, there are some discrepancies with the predictions of the SM in several observables involving $b \rightarrow c\ell^-\bar{\nu}_\ell$ and $b \rightarrow s\ell^+\ell^-$ transitions in recent years [22], which are known as B flavor anomalies. The first observation of the Flavor Changing Charged Current (FCCC) $b \rightarrow c\ell^-\bar{\nu}_\ell$

anomaly was in 2012 by the BaBar Collaboration in the measurement of $R_{D^{(*)}}$ [23], and found a discrepancy with the SM prediction at a level of 3.4σ . The first observation of the FCNC $b \rightarrow s\ell^+\ell^-$ anomaly was in 2013 by the LHCb Collaboration in the measurement of P'_5 [24], and found a disagreement with the SM at a level of 3.7σ . From then, more reported observables show discrepancies with the SM predictions. Among those anomalies, several of them measure Lepton-Flavor-Universality violation (LFUV). Below we will go over the states of some of the recent B flavor anomalies that hint at LFUV.

The LFUV observables showing discrepancies with the SM that involve the $b \rightarrow c\ell^-\bar{\nu}_\ell$ transition are

$$R_{D^{(*)}} = \frac{\mathcal{B}(\bar{B} \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau)}{\mathcal{B}(\bar{B} \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell)}, \quad \ell = e, \mu, \quad R_{J/\psi} = \frac{\mathcal{B}(B_c \rightarrow J/\psi\tau\nu_\tau)}{\mathcal{B}(B_c \rightarrow J/\psi\mu\nu_\mu)}. \quad (3.15)$$

The latest results of $R_{D^{(*)}}$ by the LHCb Collaboration are [25]

$$R_D = 0.441 \pm 0.060 \pm 0.066, \quad R_{D^*} = 0.281 \pm 0.018 \pm 0.024, \quad (3.16)$$

which have a tension with the SM at a level of 1.9σ . The combined R_D and R_{D^*} results differ from the SM by 3.2σ as shown in Fig. 3.4.

The latest result of $R_{J/\psi}$ by the LHCb Collaboration is [26]

$$R_{J/\psi} = 0.71 \pm 0.17 \pm 0.18, \quad (3.17)$$

which shows around 2σ discrepancy from the SM.

The LFUV observables show discrepancies with the SM that involving $b \rightarrow s\ell^+\ell^-$ tran-

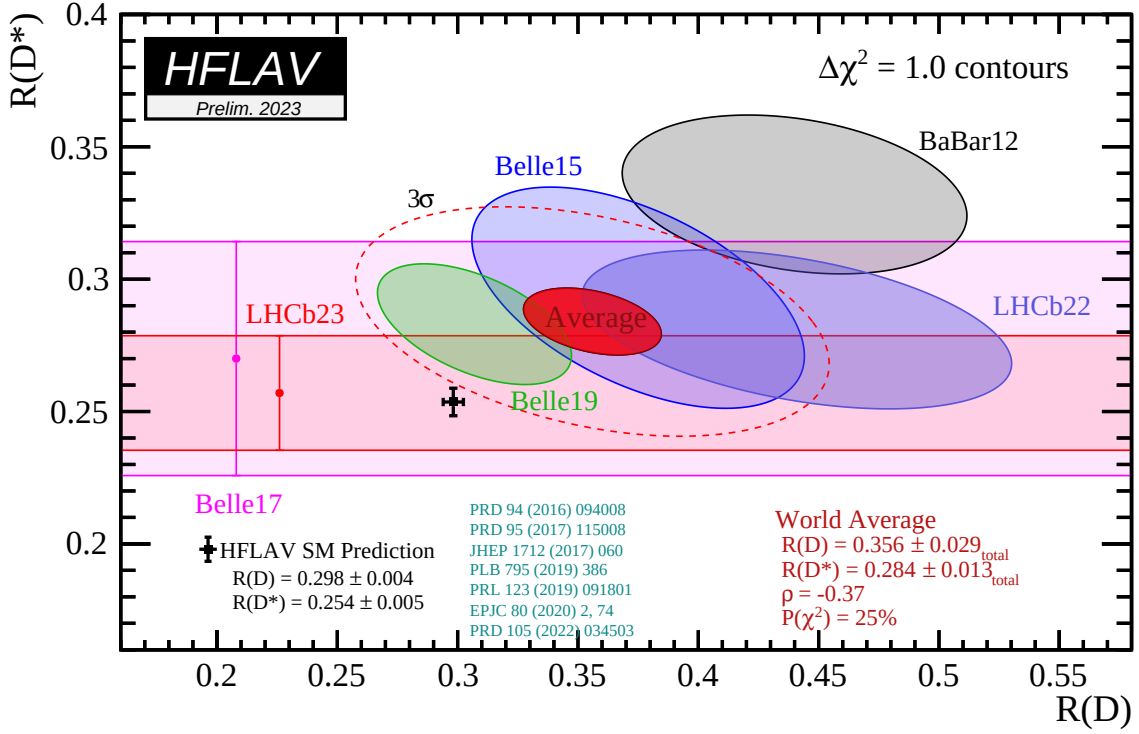


Figure 3.4: Combined R_D and R_{D^*} results compare with the SM prediction [25].

sitions that are

$$R_{K/K^*} = \frac{\mathcal{B}(B^{(+,0)} \rightarrow K^{(+,0)} \mu^+ \mu^-)}{\mathcal{B}(B^{(+,0)} \rightarrow K^{(+,0)} e^+ e^-)}, \quad (3.18)$$

There were tensions between the measurements of R_K and R_{K^*} by the LHCb Collaboration with the SM [27, 28], and found to be

$$R_K = 0.846^{+0.042+0.013}_{-0.039-0.012}, \quad R_{K^*} = 0.69^{+0.11}_{-0.07} \pm 0.05. \quad (3.19)$$

The R_K showed a tension with the SM at a level of 3.1σ while R_{K^*} differed from the SM at 2.4σ level.

Recently, the LHCb Collaboration released new results of a simultaneous analysis of R_K

and R_{K^*} [29], as shown in Fig 3.5, which are compatible with the SM at 0.2σ level. There are more LFV observables being measured, such as R_{pK} , $R_{K^{*0}}$ and $R_{K_S^0}$. The experimental results show that they are consistent with the SM predictions within 2σ level [30,31].

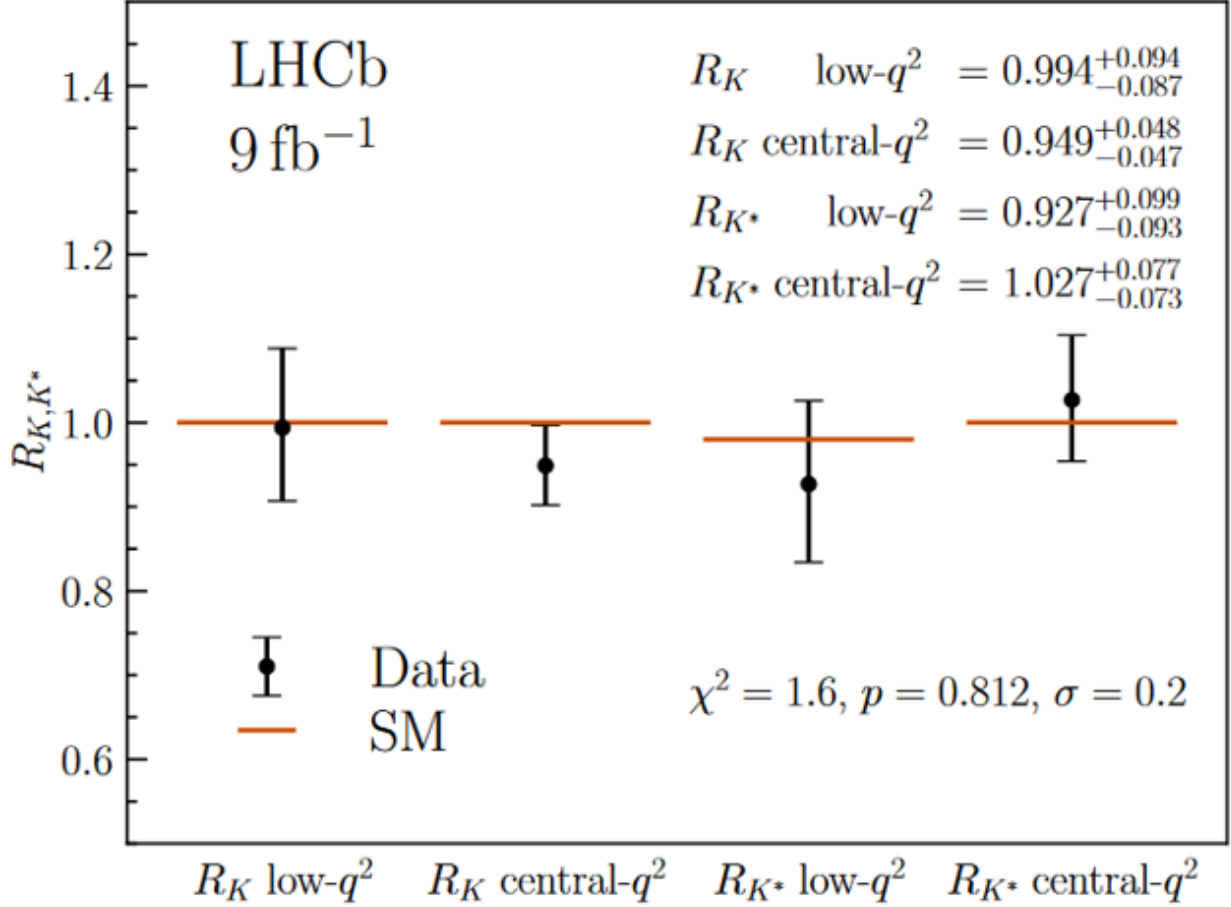


Figure 3.5: Latest R_K and R_{K^*} results released by the LHCb [29].

3.5 R_K as an observable for the test of Lepton Universality

This thesis focuses on the test of LU using $B^+ \rightarrow K^+ \ell^+ \ell^-$ decays, in particular, the R_K measurement. Decay of $B^+ \rightarrow K^+ \ell^+ \ell^-$, which involves the $b \rightarrow s \ell^+ \ell^-$ transition, is an ideal

channel for the test of LU. Since this decay involves only loop-level diagrams, it is highly suppressed in the SM and sensitive to any tree-level contributions from NP.

To remove the theoretical uncertainties in the branching fractions, the observable for a test of LU is constructed in a ratio of branching fractions. The ratio R_K for the decay of $B^+ \rightarrow K^+ \ell^+ \ell^-$ is defined as [32]

$$R_K = \frac{\int_{q_{\min}^2}^{q_{\max}^2} \frac{d\Gamma(B^+ \rightarrow K^+ \mu^+ \mu^-)}{dq^2} dq^2}{\int_{q_{\min}^2}^{q_{\max}^2} \frac{d\Gamma(B^+ \rightarrow K^+ e^+ e^-)}{dq^2} dq^2}, \quad (3.20)$$

where q^2 is the invariant mass squared of the dilepton mass system integrated between $q_{\min}^2$ and $q_{\max}^2$, and $\frac{d\Gamma}{dq^2}$ is the differential branching fractions as a function of q^2 . The lower cut is usually chosen above $4m_\mu^2$ in order to remove phase space effects. This is usually chosen to be 1.1 GeV to keep away from the ϕ resonance. The upper cut is often chosen to be 6 GeV such that it is below the J/ψ mass. The Standard Model prediction of R_K is calculated to be

$$R_K^{\text{SM}} = 1 + \mathcal{O}(m_\mu^2/m_b^2), \quad (3.21)$$

where hadronic uncertainties are largely canceled out with QED corrections controlled at the level of $\sim 1\%$ [33]. This allows us to do a precise test of LU with the R_K observable.

In the real measurement, the ratio R_K is often defined as a double ratio of branching fractions instead of a single ratio:

$$R_K = \frac{\mathcal{B}(B^+ \rightarrow K^+ \mu^+ \mu^-)}{\mathcal{B}(B^+ \rightarrow K^+ J/\psi(\rightarrow \mu^+ \mu^-))} \bigg/ \frac{\mathcal{B}(B^+ \rightarrow K^+ e^+ e^-)}{\mathcal{B}(B^+ \rightarrow K^+ J/\psi(\rightarrow e^+ e^-))}. \quad (3.22)$$

As mentioned in Sec. 3.3, the decay $J/\psi \rightarrow \ell^+ \ell^-$ is measured to be lepton universal within

0.4%. The use of $B^+ \rightarrow K^+ J/\psi(\rightarrow \ell^+ \ell^-)$ decay as a normalization channel largely reduces the systematic uncertainties induced by the different reconstruction of muons and electrons since the effects are canceled out in the double ratio.

Chapter 4

The Compact Muon Solenoid

4.1 The Large Hadron Collider

The European Council for Nuclear Research (CERN) operates the Large Hadron Collider (LHC), which is the largest particle accelerator for high-energy physics experiments in the world. Located 100 meters underground on the French-Swiss border, it is a circular hadron collider with a circumference of 26.7 kilometers. During the Run 2 operation, the LHC produced proton-proton collisions at a center-of-mass energy of 13 TeV by accelerating each proton beam in opposite directions to 6.5 TeV and colliding them with a bunch spacing of 25 nanoseconds, and delivered an instantaneous luminosity of $2 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$.

The acceleration of protons at the LHC involves several stages, as shown in Fig. 4.1. Firstly, the protons are extracted from hydrogen atoms by removing electrons. They are then accelerated to 50 MeV at Linear Accelerator 2 (LINAC2). The protons are then injected into the Proton Synchrotron Booster (PSB), Proton Synchrotron (PS), and Super Proton Synchrotron (SPS), which increase their energy to 450 GeV before they are finally

CERN's Accelerator Complex

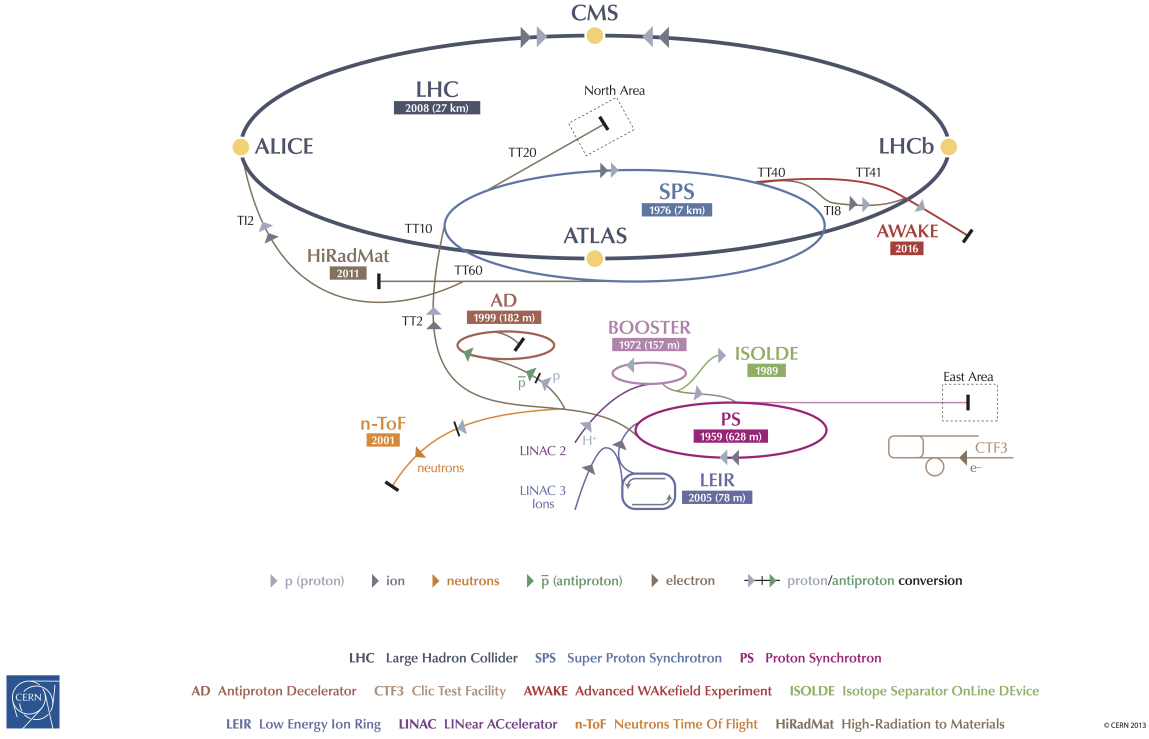


Figure 4.1: A schematic diagram of the CERN accelerator complex [34].

injected into the LHC. At the LHC, the protons are boosted to 6.5 TeV and collide in four detectors, namely the A Toroidal LHC Apparatus (ATLAS), A Large Ion Collider Experiment (ALICE), Compact Muon Solenoid (CMS), and Large Hadron Collider beauty (LHCb).

The total number of events for a specific physical process N is given by

$$N = \sigma \times \int L_{\text{int}} dt, \quad (4.1)$$

where σ is the cross-section of the process, and L_{int} is the instantaneous luminosity given by the equation

$$L_{\text{int}} = \frac{N_p^2 \cdot n_b \cdot f}{4\pi \cdot \sigma_x \cdot \sigma_y} \times F(\theta_c, \sigma_x, \sigma_y), \quad (4.2)$$

where N_p is the number of particles per bunch, n_b is the number of bunches, f is the frequency of the beam revolution, σ_x , σ_y , and σ_z are the transverse and longitudinal beam widths characterized by the beam optics, respectively, and F is a geometric luminosity reduction factor related to the crossing angle at the interaction point θ_c , given by

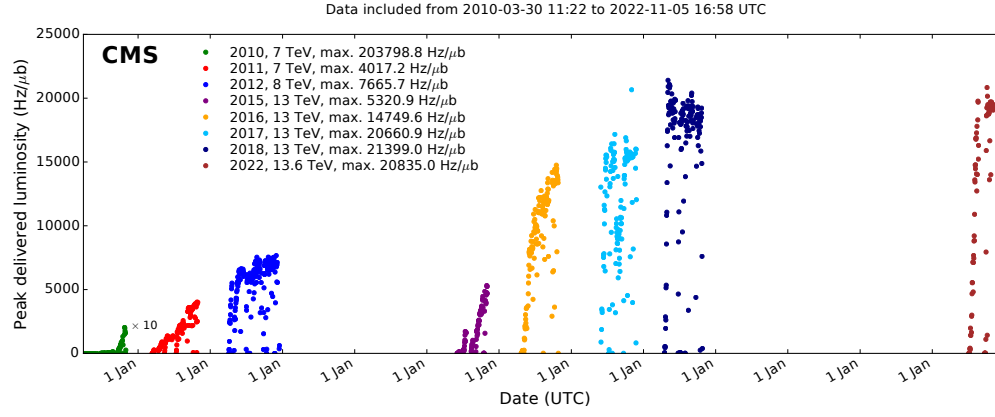
$$F(\theta_c, \sigma_x, \sigma_y) = \left(1 + \left(\frac{\sigma_z}{\sigma_x} \right)^2 \left(\frac{\theta_c}{2} \right) \right)^{-\frac{1}{2}}. \quad (4.3)$$

The designed beam parameters are summarized in Table 4.1.

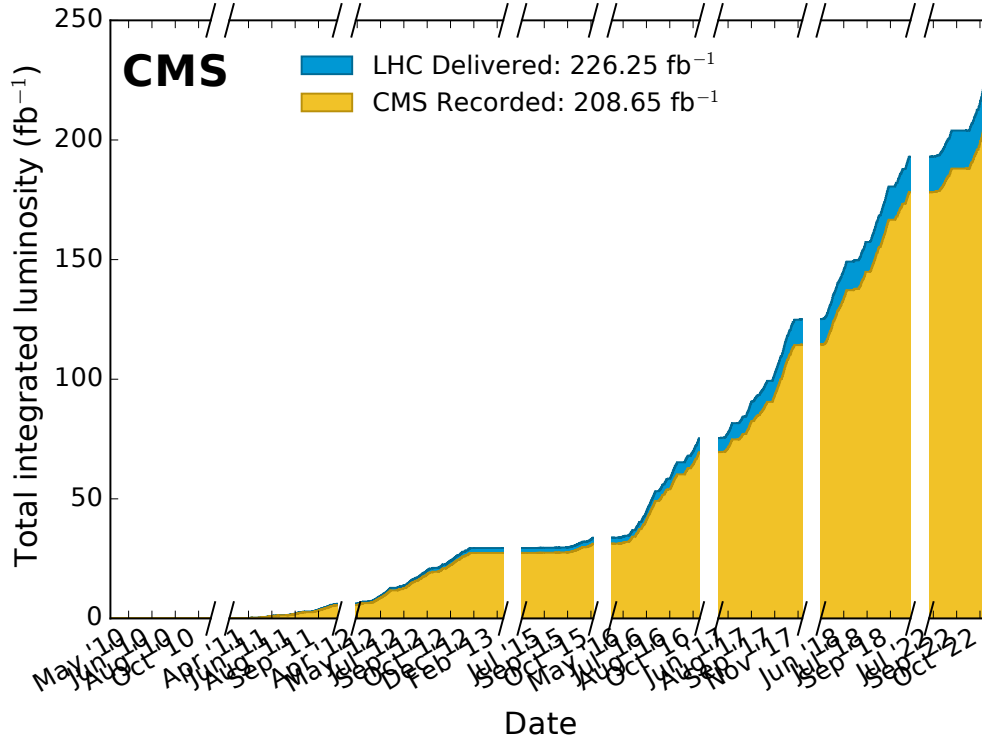
Symbol	Parameter	Designed value
N_p	Number of protons per bunch	1.15×10^{11}
n_b	Number of bunches	2808
f	frequency of the beam revolution	11.245 Hz
$\sigma_{x,y}$	RMS of transverse beam width at interaction point	$16.7 \mu\text{m}$
σ_z	RMS of longitudinal bunch length	7.55 cm
F	Geometric luminosity reduction factor	0.836
θ_c	Beam crossing angle at interaction point	$285 \mu\text{ rad}$

Table 4.1: Designed LHC beam parameters [35].

The LHC began Run 1 operation in 2010 at $\sqrt{s} = 7$ TeV and later at $\sqrt{s} = 8$ TeV and successfully delivered about a total luminosity of 30 fb^{-1} by the end of Run 1 in 2012. The machine then underwent the 2 years Long Shutdown (LS1) for maintenance and upgrades and started Run 2 operation in 2015. During the operation period of Run 2 between 2015 and 2018, LHC collided protons at $\sqrt{s} = 13$ TeV with a bunch spacing of 25 ns and delivered a total luminosity of 160 fb^{-1} . Another Long Shutdown (LS2) began in 2019, and the machine was back online for Run 3 operation in 2022. Figure 4.2 shows the peak instantaneous luminosity and the total integrated luminosity recorded at the CMS since 2010.



(a)



(b)

Figure 4.2: Peak instantaneous luminosity (a) and total integrated luminosity (b) recorded at the CMS since 2010 [36].

4.2 The CMS detector

The CMS detector, located at Point 5 of the LHC, is a general-purpose particle detector with a cutaway view shown in Fig. 4.3. It has a height of 15 meters, a length of 21 meters, and weighs 14000 tonnes. The detector comprises layers of sub-detectors, including a silicon tracker system in the core, an electromagnetic calorimeter (ECAL), a hadron calorimeter (HCAL), and an outermost Muon system. Each sub-detector is divided into two parts, the barrel region, and the end-cap region, providing complete coverage for the cylindrical geometry. A superconducting solenoid with a 6-meter internal diameter and a magnetic field of 3.8 T is located between the HCAL and the Muon systems. The solenoid enables measuring the momentum of charged particles. The CMS detector is designed to measure the momentum, energy, and charge of all particles from the interaction point. Its scientific goals range from studying the Standard Model to searching for particles beyond the Standard Model.

During the Run 2 operation period, CMS successfully recorded 151 fb^{-1} of data from proton-proton collisions at 13 TeV, $2397 \mu\text{b}^{-1}$ of data from lead-lead collisions at 5.02 TeV, and 39.2 pb^{-1} of data from proton-lead collisions at 8.16 TeV. This data set allows CMS to conduct a wide range of physics analyses.

4.2.1 Coordinate System

The coordinate system used by the CMS detector has its origin at the nominal interaction point, with the z -axis aligned with the beam pipe and the x - y plane representing the transverse plane perpendicular to the z -axis. To describe the particles, the CMS detec-

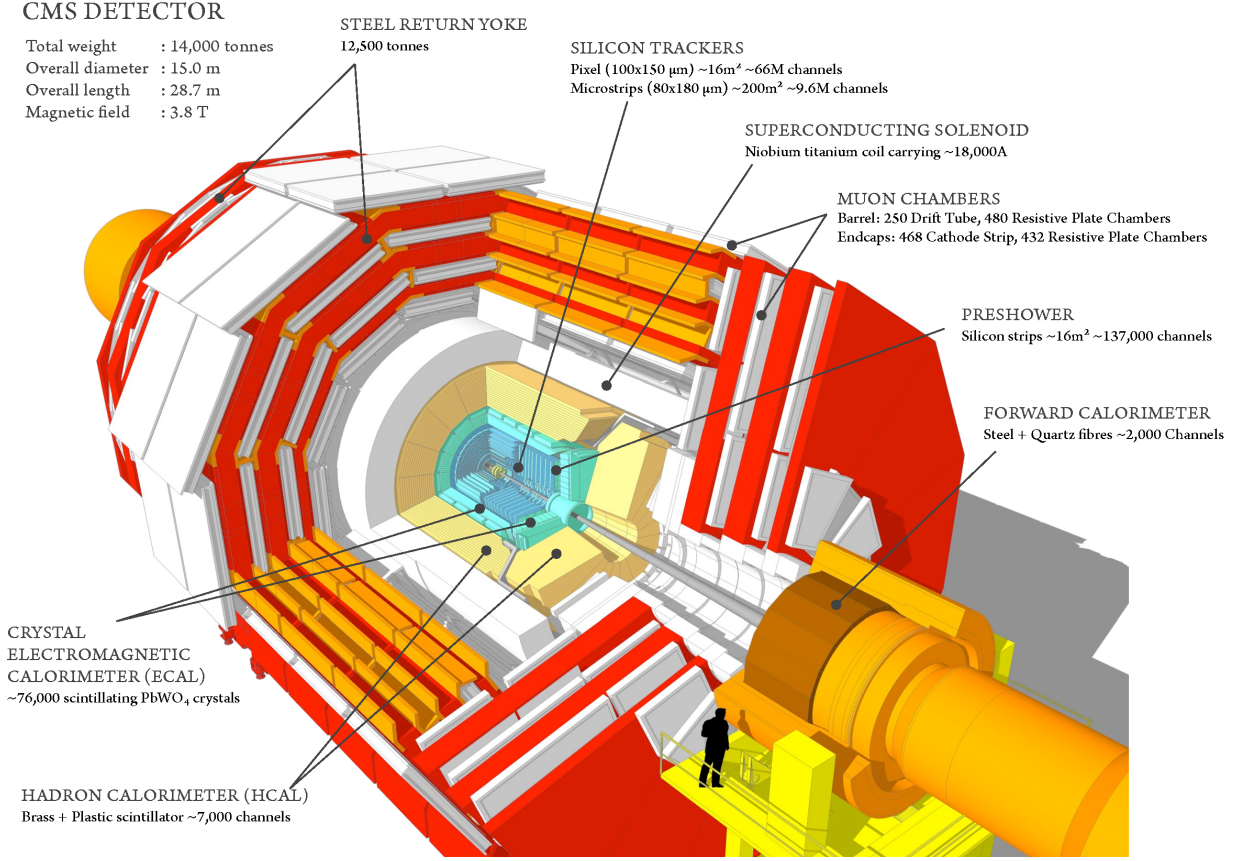


Figure 4.3: A cutaway view of the CMS detector [37].

tor uses pseudorapidity and azimuthal angle instead of the Cartesian coordinate system. The azimuthal angle ϕ is measured in the transverse plane from the positive x -axis. The pseudorapidity η is defined in terms of the polar angle θ measured from the z -axis as

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right]. \quad (4.4)$$

The production rate of highly relativistic particles is constant as a function of η , and the difference $\Delta\eta$ is Lorentz invariant under boosts along the beam axis. The separation between

two particles is often described by ΔR , defined by

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \quad (4.5)$$

4.2.2 The Inner Tracker

The inner tracker is the CMS detector's innermost subdetector, equipped with high-granularity silicon-based detectors. It is 2.5 m wide and 5.8 m long, covering an area of over 200 m². The primary purpose of the inner tracker is to measure the trajectories of charged particles and locate secondary vertices created by long-lifetime particles, which is essential for B physics analysis. The inner tracker consists of two types of sensors: the pixel detector and the silicon strip tracker. The pixel detector is placed closest to the collision point, with the silicon strip tracker wrapped outside. Figure 4.4 shows a sketch of the CMS tracking system.

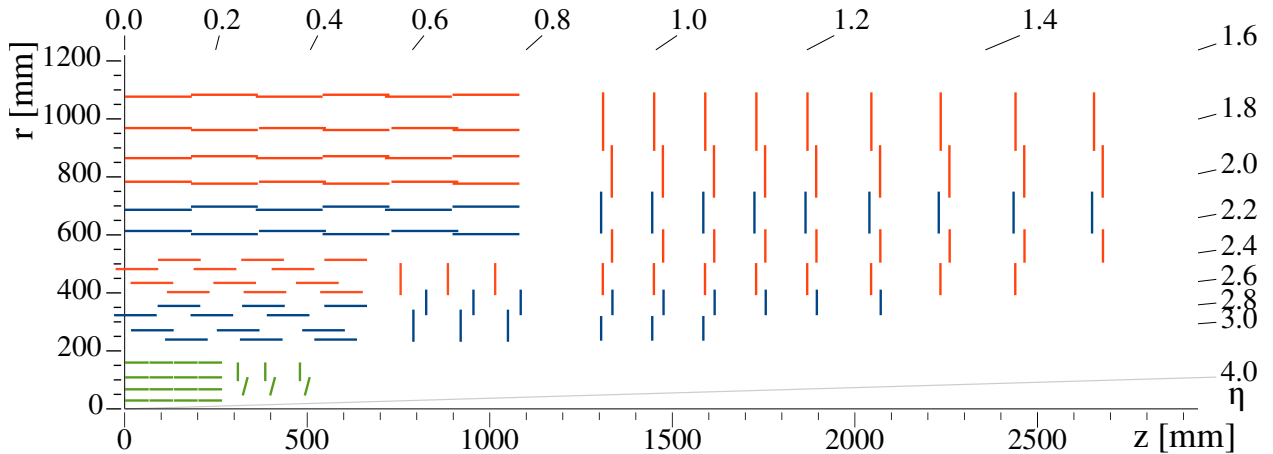


Figure 4.4: Sketch of one quarter of the Phase-1 CMS tracking system in r - z view. The pixel detector is shown in green, while single-sided and double-sided strip modules are depicted as red and blue segments, respectively [38].

The tracker system's inner portion comprises the pixel detector, which contains 1440 modules with a total of 66 million pixels measuring $100 \times 150 \mu\text{m}^2$. These modules are

arranged in three cylindrical barrel layers (BPix) and two end-cap layer disks (FPix). The BPix layers are composed of 768 modules with a total of 48 million pixels and are located at radii of 4.4 cm, 7.3 cm, and 10.2 cm. The FPix layers, located 34.5 cm and 46.5 cm from the collision point, consist of 18 million pixels. The pixel detector spans a pseudorapidity range of $|\eta| < 2.5$.

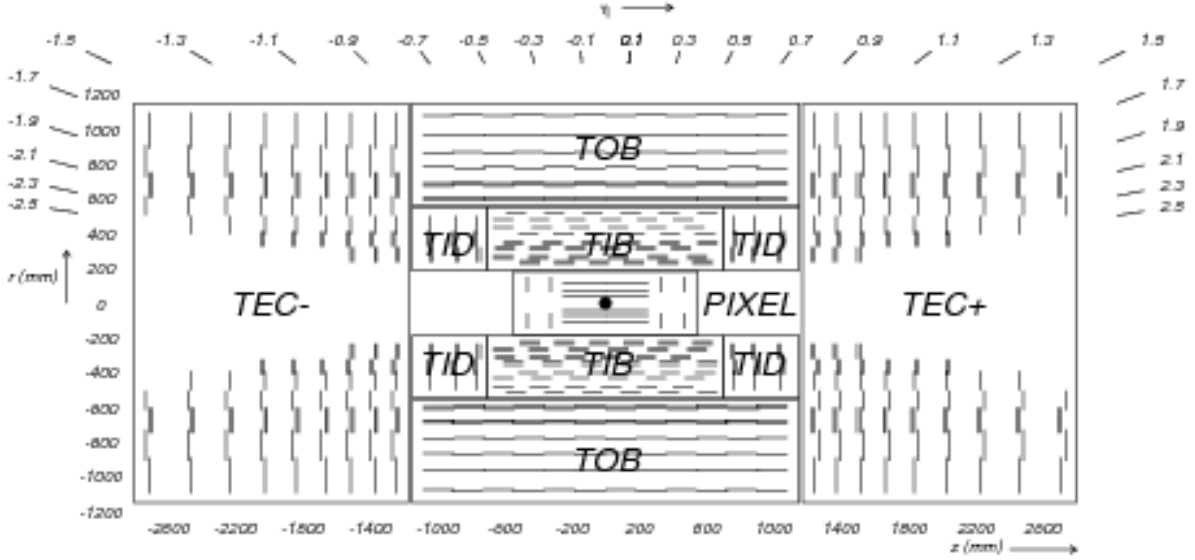


Figure 4.5: Schematic cross section through the CMS tracker [39].

The outer section of the tracker system is made up of the silicon strip tracker, which includes 9.3 million strips covering an area of around 200 m^2 . The main purpose of the silicon strip tracker is to extrapolate the trajectory from the pixel tracker to determine the momentum of charged particles. The silicon strip detector is divided into four parts: Tracker Inner Barrel (TIB), Tracker Outer Barrel (TOB), Track Inner Disks (TID), and Tracker Endcaps (TEC). The TIB comprises 4 layers and the TOB comprises 6 layers, making a total of 10 layers of detector modules in the barrel region. In the forward region, TIB is made up of three disks with three rings of sensors at both ends of TIB. TEC is comprised of

nine disks, each having four to seven rings. Figure 4.5 shows the layout of the CMS tracker system.

4.2.3 Electromagnetic Calorimeter (ECAL)

The Electromagnetic Calorimeter (ECAL) of CMS is a homogeneous, hermetic, fine-grained calorimeter consisting of lead tungstate (PbWO_4) crystals. It is designed to measure the energy of electrons and photons as they pass through the dense crystal and create an electromagnetic shower. The ECAL of CMS comprises a total of 75,848 PbWO_4 scintillating crystals [40]. Figure 4.6 illustrates the layout of the ECAL. PbWO_4 is selected for its short radiation length, small Molière radius, and fast scintillation decay time, which is in sync with the LHC bunch crossing time. Additionally, it is easy to produce from readily available raw materials [41]. The energy resolution of the ECAL depends on the energy as

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2, \quad (4.6)$$

where S is stochastic term, N is the noise term, and C is the constant term. The stochastic term includes fluctuations in shower containment as well as a contribution from photostatistics. The noise term contains the contributions from electronic noise and pileup energy [41]. They are determined to be $S = 2.8\%$, $N = 0.12\%$, and $C = 0.3\%$ from the beam-test run.

The ECAL barrel (EB) consists of 61,200 scintillating crystals and has a pseudorapidity coverage of $|\eta| < 1.48$. The granularity of the EB is $\Delta\eta \times \Delta\phi = 0.0175 \times 0.0175$, with a typical crystal size of $21.8 \times 21.8 \times 230 \text{ mm}^3$ and a radiation length depth of $25.8X_0$. The crystals are organized into 36 supermodules spanning 20 degrees in ϕ . The ECAL endcaps (EE) contain

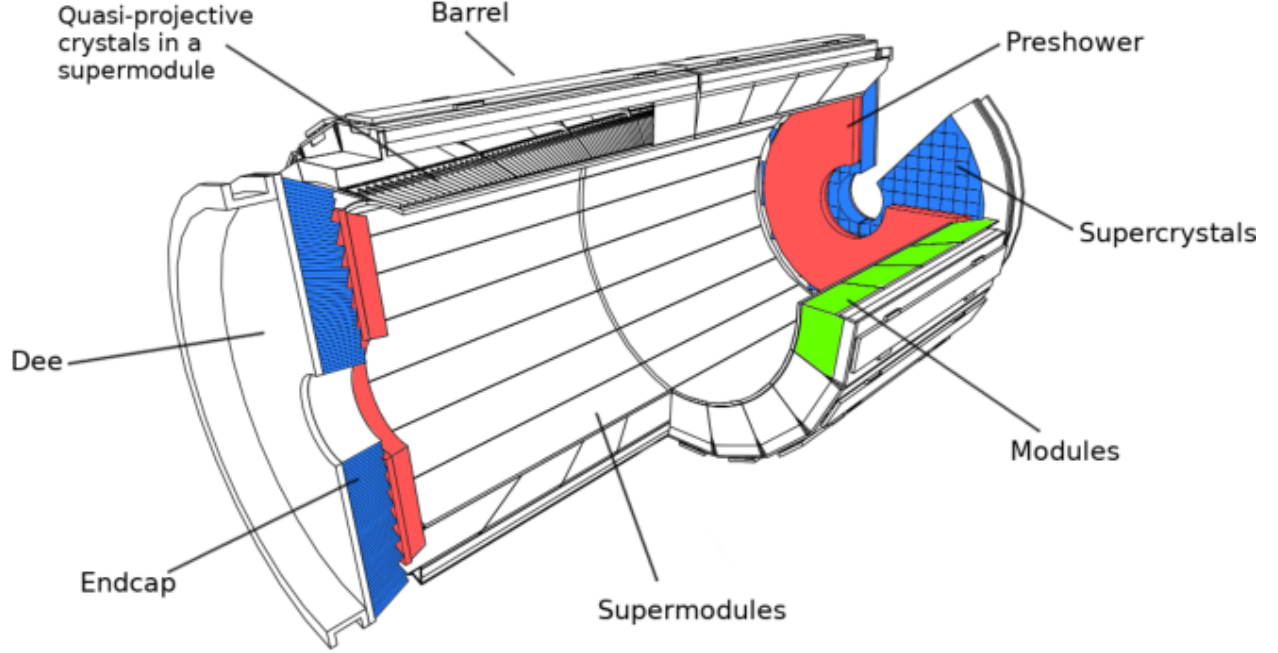


Figure 4.6: Layout of the ECAL at CMS [42].

7,324 scintillating crystals each and cover the pseudorapidity range of $1.48 < |\eta| < 3.0$. The granularity of the EE ranges from $\Delta\eta \times \Delta\phi = 0.0175 \times 0.0175$ to 0.05×0.05 , with a crystal size of $24.7 \times 24.7 \times 220 \text{ mm}^3$ and a radiation length depth of 24.7 times X_0 . In front of the EE, preshower detectors cover the range of $1.653 < |\eta| < 2.6$ to distinguish between photons and π^0 mesons.

4.2.4 Hadron Calorimeter (HCAL)

The HCAL at CMS is a sampling calorimeter with four main sections designed to measure the energy of charged and neutral hadrons. These sections are the HCAL Barrel (HB), HCAL Endcap (HE), HCAL Outer (HO), and HCAL Forward (HF). The HCAL provides coverage over a wide pseudorapidity range up to $|\eta| < 5.2$. Figure 4.7 illustrates the layout of the HCAL during Run 2.

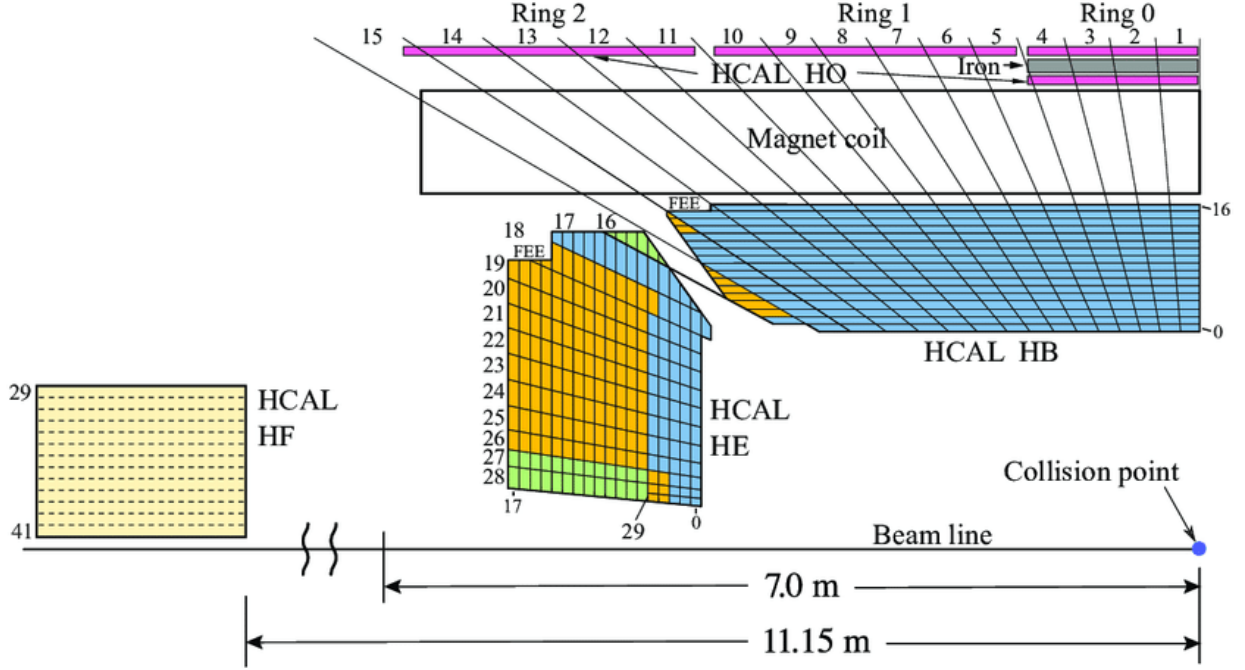


Figure 4.7: A schematic diagram showing a quarter of the HCAL during Run 2 [43].

The HB is a sampling calorimeter that uses brass absorbers and plastic scintillators as the active material. Light from the plastic scintillator is wavelength-shifted and captured in WLS fibers, which are fused to clear optical fibers for transport to the phototransducers and front-end electronics [44]. The layers are made up of 5 cm brass and 3.7 mm scintillators, which provide an interaction length (λ_I) of 5.82 at $|\eta| = 0$ and $10.6\lambda_I$ at $|\eta| = 1.3$. The HB is divided into 16 η segments, covering $|\eta| < 1.3$ with a granularity of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$.

The HE is a sampling calorimeter composed of brass and plastic scintillators similar to the HB, which covers a pseudorapidity range of $1.3 < |\eta| < 3$. The granularity of HE is $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ for $|\eta| < 1.6$ and $\Delta\eta \times \Delta\phi = 0.17 \times 0.17$ for $|\eta| > 1.6$. During the Phase-1 upgrade, photodetectors in HE were upgraded with silicon photomultipliers (SiPMs), which improved the photo-detection efficiency and the η -segmentations [45].

The HO is located outside the magnet covering $|\eta| < 1.3$, similar to the HB, and the

granularity matches that of the HB. It functions as a tail-catcher for hadronic showers and is useful for muon identification. Instead of brass, HO uses the steel return yoke and magnet material of CMS as an absorber [46].

The HF is a Cherenkov calorimeter based on a steel absorber and quartz fibers, which run longitudinally through the absorber and collect Cherenkov light [44]. Using quartz fibers helps the HF withstand the high radiation in the forward region. The HF is placed very close to the beamline, covering a pseudorapidity range of $2.9 < |\eta| < 5.0$. The granularity of HF is $\Delta\eta \times \Delta\phi = 0.175 \times 0.175$ for $|\eta| < 4.7$ and $\Delta\eta \times \Delta\phi = 0.175 \times 0.35$ for $|\eta| > 4.7$.

4.2.5 Muon System

The muon system is located on the outermost part of the CMS detector and consists of drift tubes (DT) in the barrel, cathode strip chambers (CSCs) in the endcaps, and resistive plate chambers (RPCs) for position measurement and redundant triggering. Its main purpose is to provide efficient muon identification, improve the resolution of transverse momentum measurement (p_T) for high- p_T muons, and offer a reliable stand-alone muon trigger for bunch crossing assignment [45].

In the barrel region, DT is chosen as the tracking detector due to the low expected rate and relatively low intensity of the local magnetic field [48]. The DT system is composed of four stations, covering a pseudorapidity range of $|\eta| < 1.2$. The basic unit of the DT system is a drift cell, which contains a gold-plated anode wire in a medium of 85% argon and 15% carbon dioxide gas. Additionally, there are two layers of RPCs in the two inner stations and one layer in the two outer stations, which provide very fast time response.

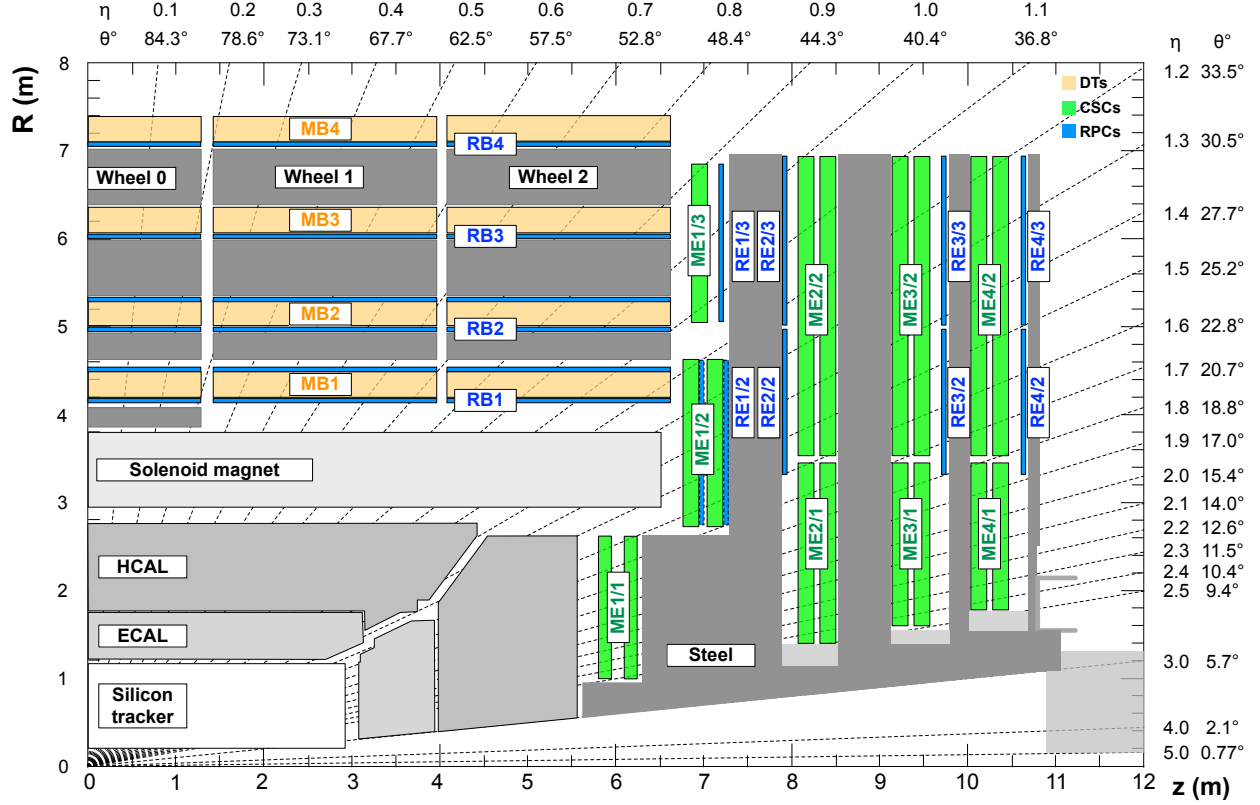


Figure 4.8: A schematic diagram showing a quarter of the muon system at CMS [47].

In the endcaps, the muon detectors cover $0.9 < |\eta| < 2.4$, overlapping with the barrel region. CSC is chosen due to its ability to provide precise space and time information in the presence of a high magnetic field and particle rate [48]. The CSC system consists of four stations, each comprising six layers of wires sandwiched between cathode panels. The CSC chambers are filled with a gas mixture of 50% CO_2 , 40% Ar, and 10% CF_4 . Each chamber can provide six measurements in ϕ and six measurements in r .

4.3 Trigger System

During Run 2, the LHC proton bunch crossing interval is 25 ns, or 40 MHz, in frequency. The CMS experiment relies on a complex trigger system to select and save the most interesting

particle collision events for further analysis. The system must be able to identify and record interesting events in real time, while discarding unimportant events in order to keep up with the high rate of particle collisions at the LHC. To achieve this, the CMS triggering system has been implemented in two stages, each designed to make a progressively more accurate and refined selection of events, the Level-1 (L1) trigger and the High-Level Trigger (HLT).

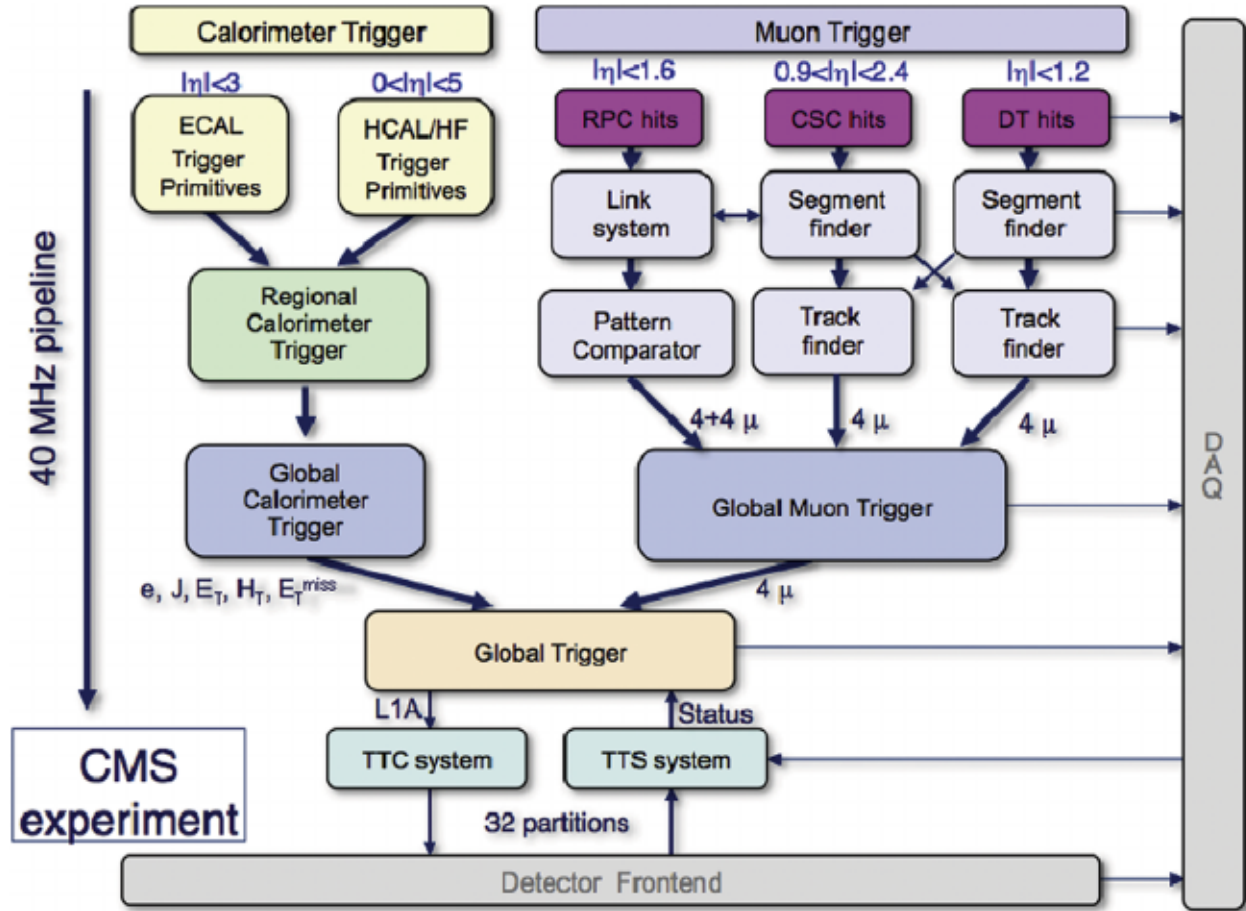


Figure 4.9: Overview of the CMS L1 trigger system [49].

The L1 trigger system is implemented in hardware and has a time constraint of about $4 \mu\text{s}$ to select the most interesting events for further processing. The trigger primitives (TP) from the calorimeters and the muon detectors are first sent to L1 system. The selected events are then passed on to the next level of the trigger system, the "global trigger" (GT), where

a more detailed analysis is performed using software algorithms to evaluate the combined event information. Figure 4.9 shows how the information are combined to make the decision. The event rate drops from 40 MHz down to 100 kHz after L1 trigger filtering.

After the L1 trigger stage, the events are sent to the software-based HLT system for further selection. The HLT selection uses the full precision of the data from the detector and the data is processed with offline-quality reconstruction algorithms. For each event, objects such as electrons, muons, jets, and missing energy are reconstructed. The event is then passed through a series of predefined selection criteria made on those objects, known as HLT paths, to determine whether to keep the event or not. At this stage, a special data stream can be collected and not analyzed until after the run is over. This procedure is known as data parking. The event rate after the HLT selection decreases down to around 1000 Hz and the data is transferred for long-term storage.

4.4 Physics Objects Reconstruction

Physics objects reconstruction is a critical step in analyzing high-energy particle collision data. The reconstruction aims to identify, track, and characterize the individual particles or physics objects produced in a collision event. Protons travel along the beamline, and the interaction point is called the primary vertex. The particles are created at the vertex and travel outward to different parts of the detector. The particles first enter the tracker, and the charged particles leave hits in it, from which tracks can be formed. The magnetic field bends the charged particles' trajectories and allows their electric charges and momenta to be determined. The particles are then entered into the ECAL. Electrons and Photons

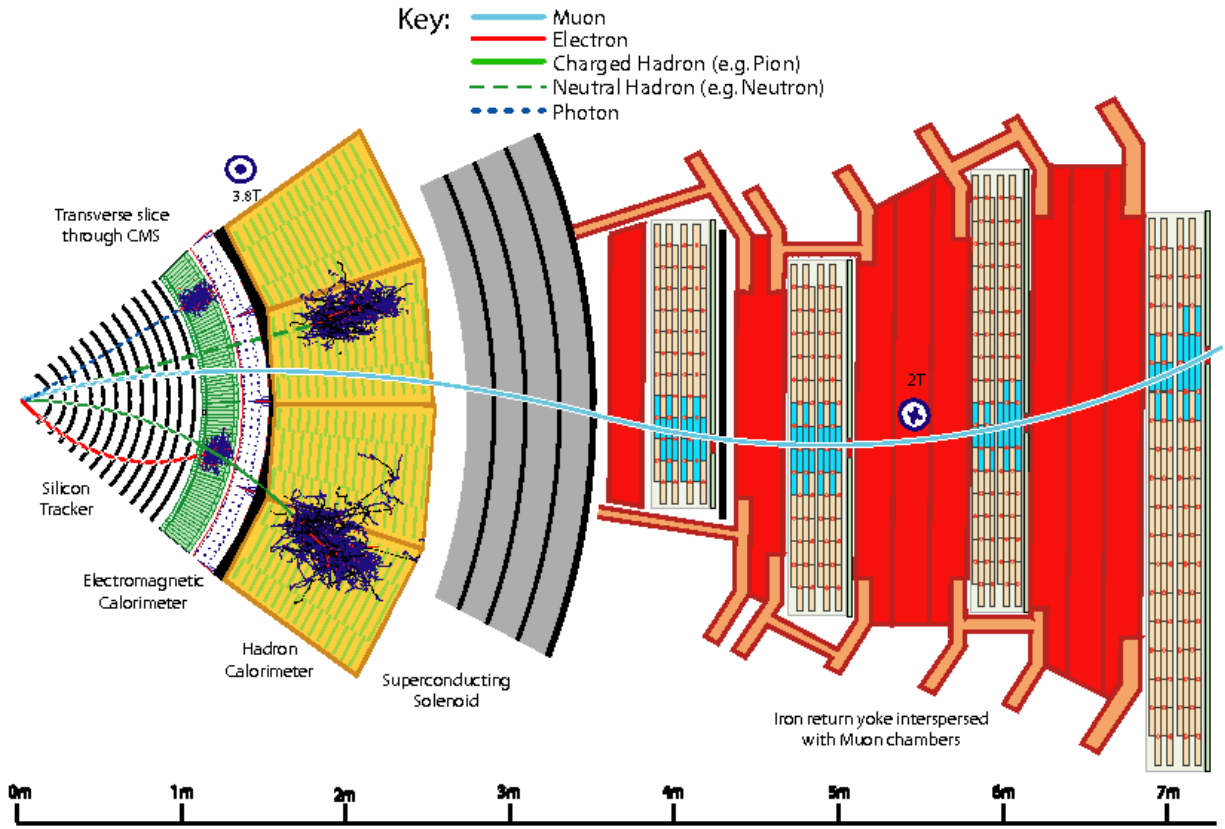


Figure 4.10: A sketch of the specific particle interactions in a transverse slice of the CMS detector, which illustrates the particle reconstruction at CMS [50].

interact with the materials and start showering. The ECAL absorbs the energy and allows the energy and direction to be measured. Charged and neutral hadrons enter the HCAL and create hadronic showers. The showering creates clusters of particles that form jets, and the HCAL measures their energy and direction. While muons traverse the calorimeters with little interaction, the muon detector outside the calorimeters provides tracking information for the muon. Figure 4.10 summarizes how different particles interact with the detector, illustrating particle identification at CMS.

It is best to utilize the information from all subdetectors for particle reconstruction. For example, linking the information between the tracker and the ECAL helps to differentiate

electrons from photons, and linking between the tracker and the HCAL helps to distinguish charged hadrons from neutral hadrons. CMS uses the particle-flow (PF) algorithm [50] for particle reconstruction. The algorithm takes tracks from the tracker, clusters from the ECAL and HCAL, and tracks from the muon detector as the inputs, known as the PF elements. The PF elements are then connected through a linking algorithm to form PF blocks. After that, a sequence of identification and reconstruction is performed and outputs a list of reconstructed particles, such as photon, electron, muon, charged hadron, and neutral hadron, as the PF candidates. PF candidates can be used in input to higher-level algorithms, such as jet clustering, missing-transverse-energy (MET) calculation, etc. Below we will further discuss the reconstruction of tracks, electrons, and muons that are being used in the R_K analysis.

4.4.1 Tracks

Track reconstruction is necessary to create meaningful track objects from the hits of the tracker. The reconstruction is divided into four stages: seeding, building, fitting, and selection. Hits in the pixel detector are used with Cellular Automaton (CA) technique to create track seeds. The building of the track is based on a combinatorial Kalman-filter (KF) method and iterative tracking, which finds the compatible hits at each layer of the tracker and updates the track candidate and its parameters. The final track parameters are estimated from the full set of hits associated with the same charged particle, and those poorly reconstructed tracks are removed.

4.4.2 Electrons and Photons

Electrons and photons primarily deposit their energy in the ECAL, with electrons also leaving tracks in the tracker due to their charge. Reconstruction of electrons and photons uses information from the tracker, ECAL, and HCAL. A crucial characteristic of these particles is their interaction with material as they pass through the ECAL. Electrons emit bremsstrahlung photons while photons convert into electron-positron pairs, creating electromagnetic showers in the ECAL. A clustering algorithm is used to collect all energy from the shower to form a single physics object. Multiple ECAL clusters are combined into superclusters (SC) to account for photon conversions and bremsstrahlung losses. For electrons, a dedicated tracking algorithm based on the Gaussian sum filter (GSF) is used to estimate track parameters to include bremsstrahlung losses, instead of the KF-based tracking algorithm [51].

4.4.3 Muons

Muon reconstruction relies on the tracker and the muon detector, as muons deposit little energy in the calorimeters. There are two types of muon reconstruction algorithms, global reconstruction and tracker reconstruction. For a global muon, hits and segments are first reconstructed locally in each muon subdetector. Then, tracking reconstruction starts from seeds made up of groups of DT or CSC segments and uses the KF tracking algorithm to form standalone muon tracks. Each standalone muon track is searched to match those reconstructed in the inner tracking system and combined to form global muons. A final fit is performed for the “tracker track” - “standalone muon” pair using the KF technique for track parameter estimation. On the other hand, the tracker muon starts from a tracker track

and interpolates it to the muon stations. The track is registered as a tracker muon if it is matched with one or more hits in the muon detector. The two independent muon collections are cross-cleaned to remove the duplicates and are used as input to the PF algorithm.

Chapter 5

Test of Lepton Universality using

$B^+ \rightarrow K^+ \ell^+ \ell^-$ decays

5.1 Introduction

The Standard Model (SM) of particle physics [1–3] successfully describes many phenomena in nature. One property in this model is called Lepton Universality (LU), whereby charged leptons (electrons, muons, and taus) have identical properties, except for their masses, because of identical couplings to the gauge bosons. Numerous experiments on $W^\pm$ and Z^0 boson decays have been conducted to verify LU, as described in 3.3, and all have consistently confirmed the theory’s predictions.

Rare B meson decays provide a good place for the test of LU. These decays are particularly interesting when a beauty quark changes into a strange quark and releases a pair of leptons, which is known as Flavor-Changing-Neutral-Current (FCNC) decay. These decays are not allowed at the tree level because of the Cabibbo suppression and need higher-order

Feynman diagrams with penguin loops. These loops can either be a neutral gauge boson emitting the lepton pair, or a box diagram with two W bosons, as shown in Fig. 3.1. As a result, the branching ratio of such decays is very low, around 10^{-7} . Neglecting the mass of leptons, and hence the difference in available phase space for the decay, the ratio of the decay amplitude with muons to that with electrons is unity in the SM. However, if any new physics is involved, it could alter the branching ratio for different lepton species if the coupling of the leptons to new particles differs.

The LHCb experiment reported searches of LU violation in various decay channels involving the FCCC $b \rightarrow c\ell\bar{\nu}_\ell$ and FCNC $b \rightarrow s\ell^+\ell^-$ decays [25–28, 30, 31, 52–54]. One of the measured observables is the ratio R_K , which is the ratio of the branching fractions for $B^+ \rightarrow K^+\mu^+\mu^-$ to that of $B^+ \rightarrow K^+e^+e^-$. The measured branching ratio to muons is smaller than expected, indicating a tension between the theoretical prediction of the SM and the experimental measurement. Recently, the LHCb experiment reported a new joint measurement on R_K and R_{K^*} [29], which showed agreement with the SM. These results give a good motivation to conduct similar measurements at CMS for the test of LU.

This thesis summarizes the work of the measurement of R_K using data collected by the CMS experiment at the LHC in 2018. A special trigger and storage strategy were used to collect unbiased beauty meson decay samples.

5.1.1 Measurement strategy

The quantity R_K is defined as a ratio of branching fractions between $B^+ \rightarrow K^+ \mu^+ \mu^-$ to $B^+ \rightarrow K^+ e^+ e^-$:

$$R_K = \frac{\mathcal{B}(B^+ \rightarrow K^+ \mu^+ \mu^-)}{\mathcal{B}(B^+ \rightarrow K^+ e^+ e^-)}. \quad (5.1)$$

As mentioned in Sec. 3.5, R_K is usually measured in a double ratio, normalized to the corresponding $B^+ \rightarrow K^+ J/\psi$ channel. In terms of measured quantities, R_K can be written as

$$R_K = \left(\frac{N_{K^+ \mu^+ \mu^-}}{N_{K^+ e^+ e^-}} \right) \left(\frac{N_{K^+ J/\psi(e^+ e^-)}}{N_{K^+ J/\psi(\mu^+ \mu^-)}} \right) \left(\frac{\epsilon_{K^+ e^+ e^-}}{\epsilon_{K^+ \mu^+ \mu^-}} \right) \left(\frac{\epsilon_{K^+ J/\psi(\mu^+ \mu^-)}}{\epsilon_{K^+ J/\psi(e^+ e^-)}} \right), \quad (5.2)$$

where N_X is the observed yield in the final state X , and ϵ_X is the acceptance and efficiency of that final state.

The use of a normalization channel reduces the systematic uncertainty due to the electron and muon acceptances and efficiencies. The branching fraction of the decay $J/\psi \rightarrow \mu^+ \mu^-$ is identical to that of the decay $J/\psi \rightarrow e^+ e^-$ in the SM, and confirmed by the measurement within the uncertainty of 0.4% [18]. Therefore, the SM prediction of R_K remains as unity after normalizing with J/ψ channel. On the other hand, Fig. 5.1 shows that the transverse momentum of the muons from the continuous and the resonant B decay are similar, while the pseudorapidity spectra are identical, reflecting that they have very similar kinematics. Most of the systematic uncertainties cancel in the double ratio.

The R_K measurement is typically performed in the lepton invariant mass $m(\ell^+ \ell^-)$ range between the Φ mass (1020 MeV) and just below the J/ψ mass (3097 MeV), which is a range that is relatively free of resonances. Although the higher mass range, above the $\psi(2S)$

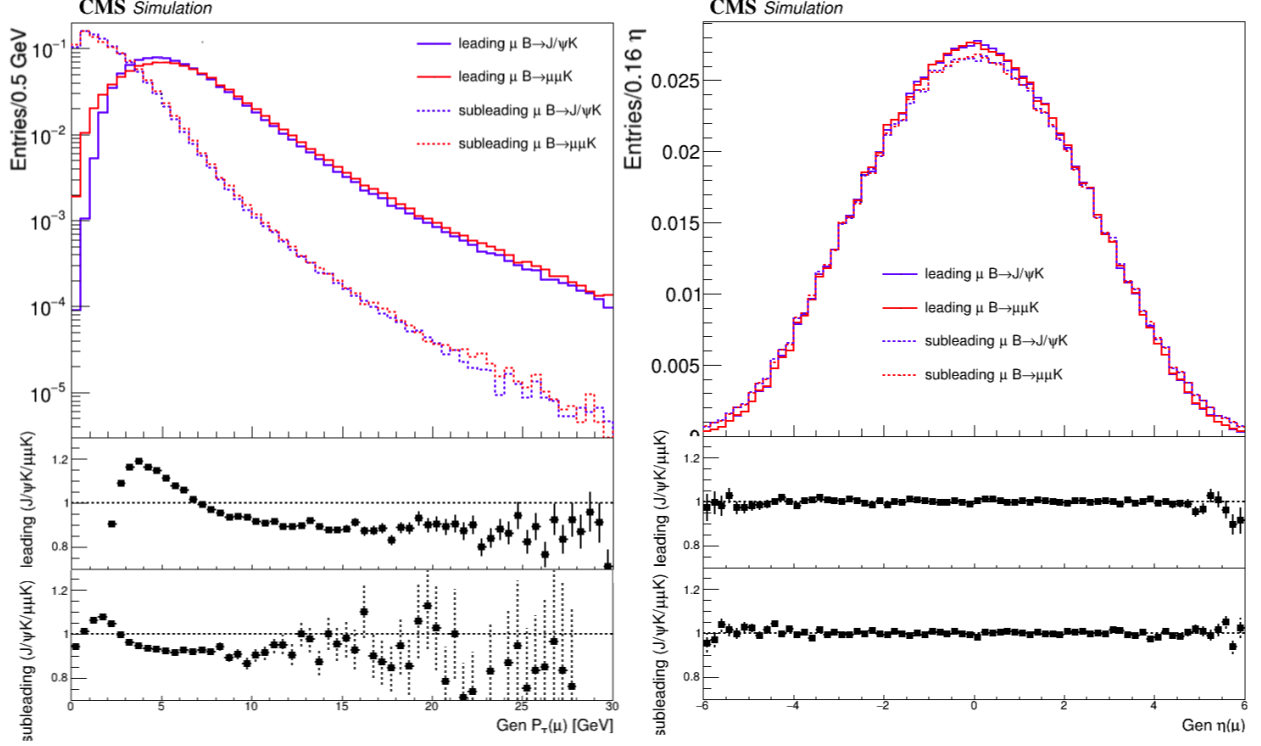


Figure 5.1: Simulated generated muons' p_T (left) and pseudorapidity- η (right) spectra using $B^+ \rightarrow K^+ \mu^+ \mu^-$ and $B^+ \rightarrow K^+ J/\psi \rightarrow K^+ \mu^+ \mu^-$.

mass, contains many resonances, it has been demonstrated that these resonances contribute only about 15% to the branching ratio. Therefore, it is interesting to perform the R_K measurement in this high-mass bin as well. The mass range of the $\psi(2S)$ is used as a sanity check of the measurement, leveraging the fact that, like the J/ψ , its branching fraction to electron and muon pairs is known to be the same at the measurement uncertainty level of 8% [18]. Table 5.1 summarizes the lepton invariant mass range used in this analysis.

Measurement range	min. $m(\ell^+ \ell^-)$ [GeV]	max. $m(\ell^+ \ell^-)$ [GeV]
low- q^2	1.05	2.45
J/ψ	2.9	3.2
$\psi(2S)$	3.55	3.8
high- q^2	4.0	4.8

Table 5.1: Minimum and maximum lepton invariant mass values for the R_K measurement ranges considered in this analysis.

5.2 Data parking

LHC collides proton bunches at a frequency of 40 MHz. It is impossible to store all of this massive amount of data. CMS adopts a two-level trigger system, which reduces the event rate from 40 MHz to a manageable amount of 1 kHz, as described in Sec. 4.3. The events accepted by HLT are transferred promptly to the CERN Tier-0 data center for reconstruction and storage and are generally available within 48 hours for physics. The data volume is constrained by the bandwidth of the link between Point 5 and the CERN data center and the speed of prompt reconstruction without significant backlog.

An effective way to bypass this limitation is a technique called data parking. The data are temporarily stored in local tape storage at Point 5, skipping prompt reconstruction. These data streams are processed at a later point in time. This trick allows CMS to record additional data on top of the standard physics data stream. In 2018, CMS deployed triggers and utilized the data parking strategy for B physics analyses, known as the B -parking campaign. The B -parking data set enables CMS to record $\mathcal{O}(10^{10})$ unbiased b hadron decays. This huge data set allows CMS to probe B decays reaching branching fractions as low as 10^{-7} , and therefore allowing CMS to conduct the R_K measurement.

5.3 Triggers and data

To obtain an unbiased sample of b hadron decays, the trigger strategy for B -parking data samples was designed to take advantage of the fact that b hadrons are mostly produced in pairs during proton-proton collisions at the LHC. The trigger logic was used to select events in which one of the two b hadrons ("tag-side") decayed semi-leptonically, emitting a high-

transverse-momentum muon. The "probe-side" b hadrons decays were not biased by the trigger logic requirements imposed on the "tag-side" decay. The presence of a single muon in the logic accounted for about 20% of all b hadron decays . The tag-side and probe-side provide the muonic and electronic final states, respectively, for the R_K measurement.

The trigger strategy is tuned to record b semileptonic decays with a purity of around 80%. The trigger strategy aims to maximize the number of events without affecting the ability of the CMS trigger and DAQ systems for the main physics program. The trigger strategy exploits the surplus of idle online computing resources resulting from the decrease in instantaneous luminosity L_{inst} during each LHC fill. As L_{inst} decreases, the L1 and HLT trigger rates decrease, allowing us to fit more B -parking events in the limited bandwidth with a looser HLT trigger.

Table 5.2: Summary of the loosest muon trigger requirements imposed by the L1 and HLT algorithms in each instantaneous luminosity scenario: the L1 and HLT muon p_T thresholds, and the HLT muon impact parameter significance IP_{sig} . Also shown are the trigger purity and peak trigger rate. All values are shown as a function of the peak instantaneous luminosity L_{inst} . The table refers to the trigger configuration that was online starting from LHC Fill 7132.

Settings	Peak L_{inst} [$10^{34} \text{ cm}^{-2}\text{s}^{-1}$]	L1 p_T^μ thresh. (GeV)	HLT p_T^μ thresh. (GeV)	HLT μ IP_{sig} thresh.	Purity	Peak rate [kHz]
1	1.7	12	12	6	0.92	1.5
2	1.5	10	9	6	0.87	2.8
3	1.3	9	9	5	0.86	3.0
4	1.1	8	8	5	0.83	3.7
5	0.9	7	7	4	0.59	5.4

Table 5.2 summarizes the requirement for the tag-side trigger muon imposed by the L1 and HLT algorithms. The L1 trigger either requires the trigger muon with a p_T greater than 22 GeV or 27 GeV, or with a lower p_T but having $|\eta| < 1.5$. Both the L1 and HLT requirements are relaxed through a series of settings that progressively increase the rate at

Table 5.3: Summary of the tag-side muon trigger L1 seeds at different peak instantaneous luminosity values. Mu10 stands for muon $p_T > 10$ GeV, and “er1p5” for $|\eta| < 1.5$. Each column represents one of the different settings that were online throughout the year, together with the number of the first LHC Fill in which they were implemented. The majority of the data has been collected using the settings in the last two columns.

Peak L_{inst} [$10^{34}cm^{-2}s^{-1}$]	Loosest L1 seed	Loosest L1 seed	Loosest L1 seed	Peak L_{inst} [$10^{34}cm^{-2}s^{-1}$]	Loosest L1 seed
1.6	Mu22	Mu22	Mu12er1p5	1.7	Mu12er1p5
1.4	Mu10er1p5	Mu10er1p5	Mu10er1p5	1.5	Mu10er1p5
1.2	Mu10er1p5	Mu9er1p5	Mu9er1p5	1.3	Mu9er1p5
1.0	Mu8er1p5	Mu8er1p5	Mu8er1p5	1.1	Mu8er1p5
0.8	Mu7er1p5	Mu7er1p5	Mu7er1p5	0.9	Mu7er1p5
First Fill	6672	6693	6763		7132

which the CMS trigger system results in a positive decision with only a moderate reduction in purity. The HLT settings and the corresponding L1 seeds are summarized in Table 5.3. The B -Parking sample uses a mixture of HLT triggers, with different thresholds in p_T and different thresholds on the significance of the impact parameter with respect to the beam spot on the trigger muon. This information is shown in the name of the HLT paths. For example, HLT_Mu9IP6 refers to the HLT path requiring the trigger muon with a $p_T > 9$ GeV and impact parameter significance $IP_{sig} > 6$ with respect to the beam spot.

The peak trigger rates of the L1 and HLT systems are approximately 50 kHz and 5.4 kHz, respectively. The highest rates are observed towards the end of an LHC fill, resulting in a pileup (PU) value of around 29 when averaged over the entire collected dataset. This value is lower than that typically observed for the standard physics data streams of CMS. Figure 5.2 shows the trigger rate of the CMS L1 system as a function of time during an LHC fill in 2017 (a) and 2018 (b), highlighting the decrease in the total rate with time due to decreasing L_{inst} during the 2017 fill. In contrast, the total rate in 2018 is maintained close to the optimal value of 90 kHz by progressively increasing the number of B -parking triggers

as the luminosity decreases, as shown in Table 5.2. Figure 5.2(c) shows the trigger rates of the CMS HLT system as a function of time during the 2018 LHC fill for both the standard physics and B -parking streams. The rate for the standard physics data streams decreases monotonically, while the rate for the B -parking increases sharply throughout the LHC fill as the setting evolves.

The whole B -parking campaign collected 12 billion events from June to November 2018, containing a total of 10 billion unbiased b hadron decays. This data set enables CMS to conduct rich analyses in the B -physics sector.

5.3.1 Trigger purity

The trigger purity is the fraction of events triggered by muons that come from a B meson, defined as

$$P = \frac{N(B \rightarrow X\mu)}{N(\mu)}. \quad (5.3)$$

The trigger purity is found to be in the range of 0.59–0.92 from simulations, with an average of around 75%. To validate the purity in data, the decay of $B^0 \rightarrow D^{*+}\mu^-\nu_\mu$ is used, following the D decays in $D^{*+} \rightarrow D^0\pi_s^+ \rightarrow K^-\pi^+\pi_s^+$ mode, where a soft pion is denoted as π_s . This decay mode is used for validation because of its large branching fraction as well as the presence of μ , K , π , and π_s in the final state.

In order to reconstruct B^0 candidates in data, two tracks with opposite charges are first selected and fitted to a common vertex. The tracks are required to have $p_T > 1.5$ GeV, pass the "High Purity" ID, have normalized $\chi^2 < 2.5$, and leave at least 3 hits in the tracker. The track with the same sign as the triggering muon is assigned the K mass, and the oppositely

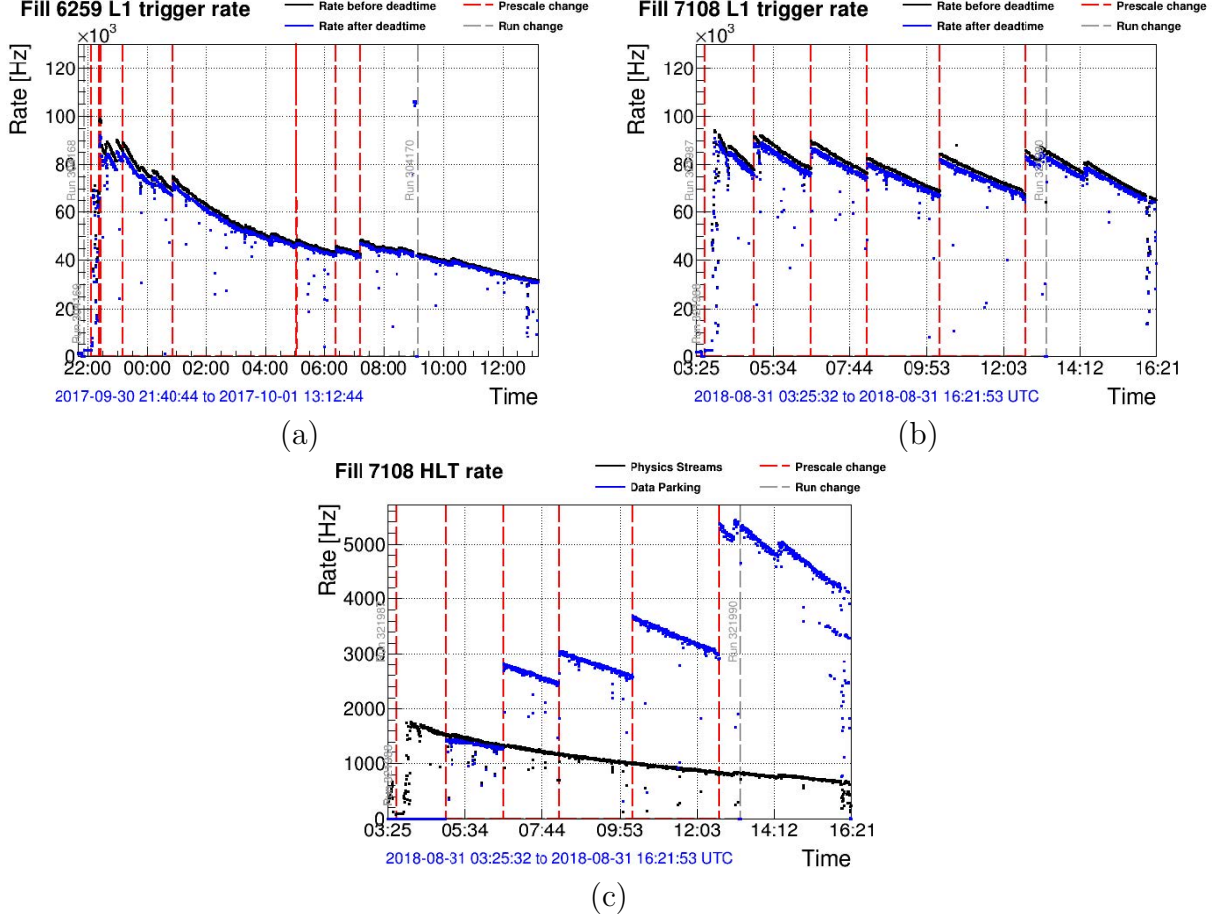


Figure 5.2: Rate of the CMS L1 trigger (blue data points), as a function of time, during an LHC fill in (a) 2017 and (b) 2018. The time intervals cover 13–15 hours. Changes in the run number and settings (prescale column) are indicated by vertical grey and red dashed lines, respectively. Figure (c) shows the trigger rates of the CMS HLT system as a function of time during an LHC fill in 2018, for the physics (black data points) and B parking (blue data points) streams.

charged track has the π mass. The candidates are then required to pass the selections listed in Table 5.4.

Table 5.4: List of variables and their corresponding cuts, to reduce the background in ditrack pairs. Prob(vtx) is the probability of the secondary vertex, and L_{xy} the absolute displacement from the Beam Spot.

Variables	Cuts
Prob(vtx)	> 0.02
L_{xy}	> 0.1 mm
$p_T(\text{trk } 1)$	> 5 GeV
$p_T(\text{trk } 2)$	> 3 GeV

Figure 5.3 shows the reconstructed D^0 peak in data after applying the selections. A mass window of ± 30 MeV from the nominal D^0 mass (i.e., 1.86 GeV) is then applied to the track pair to construct D^0 candidates. The D^0 candidates are combined with a soft track ($p_T > 300$ MeV) with the opposite charge as that of the μ . The mass difference between the K, π, π_s and K, π produces a very narrow peak around the mass of π , which is shown in the 5.4.

The mass distribution is fitted with a double Gaussian for the signal and a second-order polynomial for the background to extract the signal yield. The trigger purity is measured by taking into account the branching fractions and the acceptance and reconstruction efficiency. The result shows the trigger purity in the data is about 0.75, which agrees with the result from the simulation.

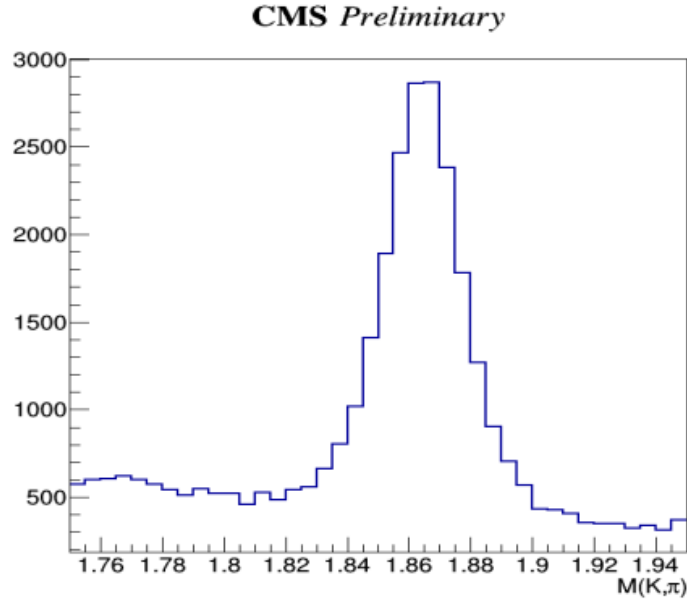


Figure 5.3: Mass of the track pairs. The masses are assigned based on the charge of tracks and μ .

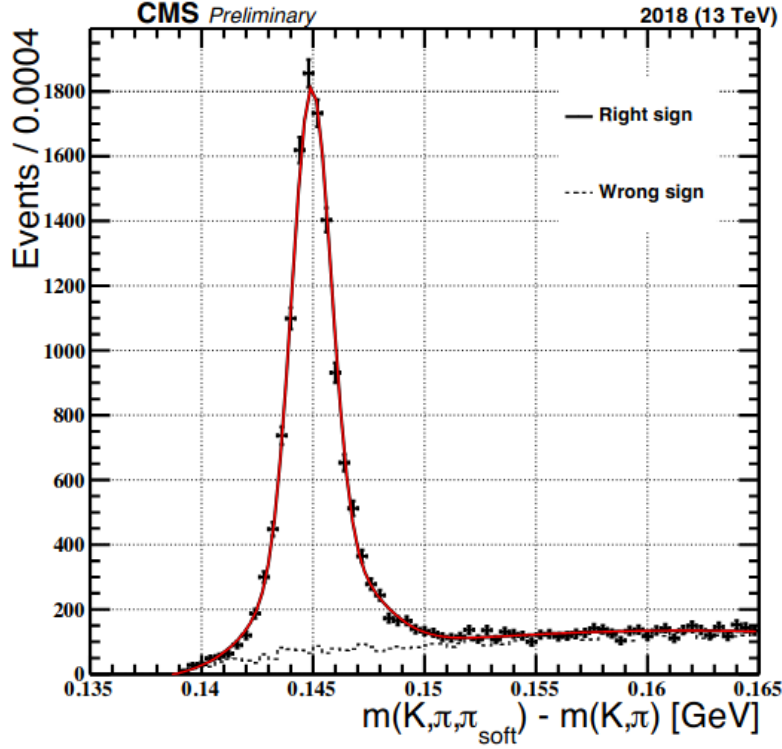


Figure 5.4: Mass difference between the K, π, π_s and K, π combinations. Only a subset of parked events have been used. The right and wrong sign refer to the sign of K : right sign is the same as the μ , wrong is the opposite of the μ . Wrong sign events give a reasonably accurate description of the background under the peak.

5.3.2 Pile-up and luminosity

Figure 5.5 shows the distribution of the number of true pileup interactions in the data collected in the B -parking sample. The integrated luminosity of the B -parking data is measured to be $41.5 \pm 1.0 \text{ fb}^{-1}$.

5.3.3 Event simulation

The Monte Carlo (MC) simulations used in this analysis are generated PYTHIA 8 [55], and include parton showering, fragmentation, and hadronization. The B meson decay simulations are performed with EvtGen [56]. The underlying events are modeled by PYTHIA with the

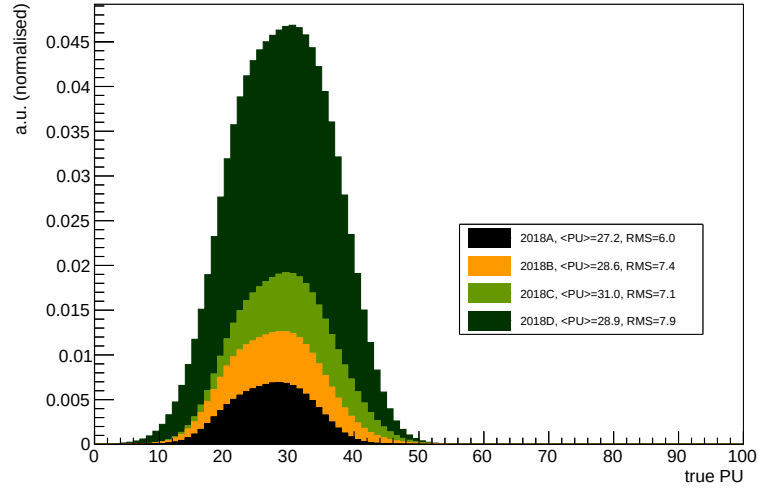


Figure 5.5: Distribution of the number of true interactions in the events collected by the *B*-parking data sample.

CP5 [57] tunes. The parton distribution functions (PDFs) are taken from the NNPDF 3.1 [58].

5.4 Analysis strategy

To measure R_K , the signal yields and efficiencies of four decay modes, namely $B^+ \rightarrow K^+ \mu^+ \mu^-$, $B^+ \rightarrow K^+ e^+ e^-$, $B^+ \rightarrow K^+ J/\psi(\rightarrow \mu^+ \mu^-)$, and $B^+ \rightarrow K^+ J/\psi(\rightarrow e^+ e^-)$, are required to obtain from the data and simulations. The analysis is divided into two channels: the muon channel and the electron channel, corresponding to the final states with muons and electrons, respectively. The B -parking data set is used to provide the desired events for the analysis. Specifically, the muon channel uses the tag-side of the B -parking trigger, which contains a displaced energetic muon, while the probe-side is used for the electron channel, consisting of unbiased B decays. Figure 5.6 is a schematic diagram showing the event topology of $B^+ \rightarrow K^+ \ell^+ \ell^-$ decay in the B -parking trigger setting.

The reconstruction of a B^+ candidate starts with the selection of two oppositely charged leptons with a track to serve as the K^+ meson candidate, forming a triplet. The charged leptons are reconstructed using the standard PF algorithm in CMS, with an additional object identification requirement. To improve electron reconstruction performance in the low-energy region, a new low- p_T electron reconstruction algorithm was developed specifically for this analysis, in addition to retraining the PF electron identification to enhance its performance in the low-energy regime. The charged leptons and track are then fitted to a common vertex, as the B^+ is a long-lifetime particle that forms a secondary vertex, distinct from the primary vertex. To distinguish the signals from the background, separate machine learning-based classifiers are developed for different channels to select the events.

Once the event selection is completed, a mass peak should become visible above the background. A fit is performed on the triplet mass spectrum to extract the signal yields

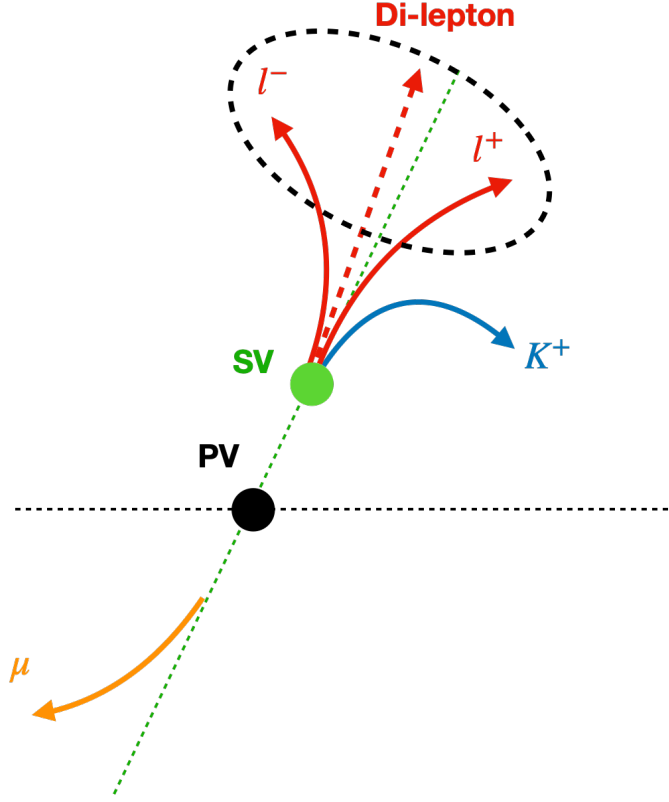


Figure 5.6: A schematic diagram showing a $B^+ \rightarrow K^+ \ell^+ \ell^-$ decay in the probe side and a displaced trigger muon in the tag side.

from the data. The efficiencies of the corresponding decays are calculated from simulation. To combine the muon and electron channels, a likelihood is developed, that takes into account the signal and background shapes and their corresponding efficiencies. The measurement of R_K is then obtained by maximizing the likelihood function and the corresponding confident levels (CL) are obtained by profiling the likelihood.

5.5 Object reconstruction

The following section discusses the reconstruction and selection of muons, electrons and tracks in the R_K measurement.

5.5.1 Muons

Sec. 4.4.3 describes the muon reconstruction algorithm in CMS. The muon collection reconstructed from the PF algorithm is called the PF muon. Additional quality criteria, known as muon identification (muon ID), are required for PF muons for them to be used as the muon candidates in this analysis.

Muon identification

Five muon IDs are developed in CMS, summarized in Table 5.5. They are tuned to give optimal performance in different kinds of analyses. Most B -physics analyses use the Soft ID, which is tuned to give high efficiency for slightly displaced and low p_T muon. However, this analysis can also be impacted by in-flight decays of heavy-flavored hadrons, where long-lived hadrons decay to a muon and leave hits in the muon chambers.

Figure 5.7 shows the invariant mass spectrum of track and a muon from the triplet in data from the region of the D^0 rest mass, suggesting that there are many D^0 mesons that could potentially contaminate our signal. In order to study the impact of in-flight decays in the analysis, especially from the decay mode of $D^0 \rightarrow K\pi$, we reconstruct the D^0 mass from the triplet and see how well different muon IDs could suppress it. The track and a muon from the triplet are used with the two mass hypotheses of K and π assigned to them. The

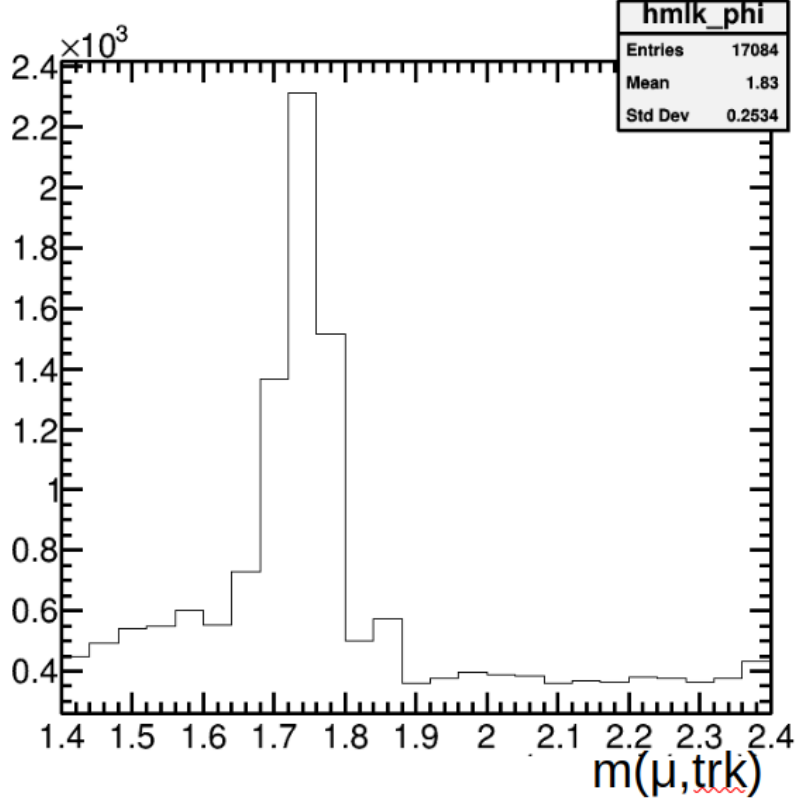


Figure 5.7: Invariant mass of each of the reconstructed μ in the triplet and the track. A peak is observed around the D^0 mass.

two hypotheses are compared in Fig. 5.8 (a) and (b), showing that the D^0 peak is sharper when the reconstructed muon is assigned to the K mass and the track is assigned to the π mass. It is because the decay mode of the K from $D^0 \rightarrow K\pi$ decay has a probability of 65% decaying into a muon, which causes the in-flight decays background in the analysis.

This background can be suppressed effectively with a proper choice of muon ID. The D^0 mass vanishes when the medium muon ID is required, as shown in Fig. 5.9. The tighter muon ID requirement costs 4% loss in $B^+ \rightarrow K^+\mu^+\mu^-$ signal efficiency, which is affordable in this analysis given that the background is reduced by an order of magnitude. Thus, the medium muon ID is chosen for the R_K measurement.

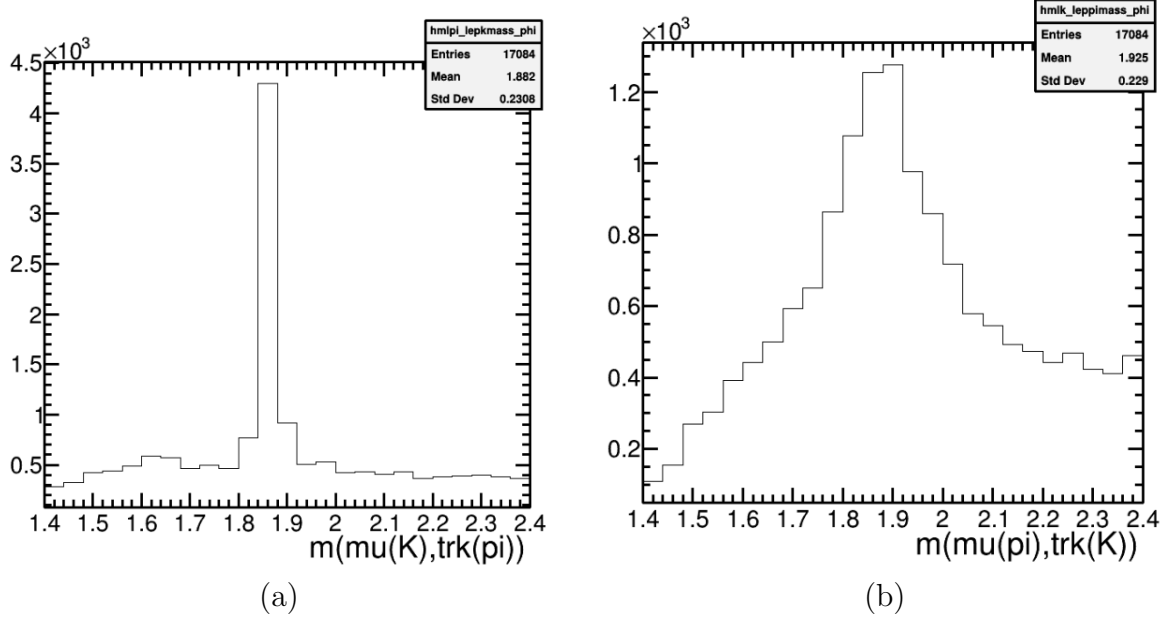


Figure 5.8: (a) Invariant mass of a candidate μ and a track, calculated assuming that the muon has the K mass and the track the π mass; (b) Invariant mass of a candidate μ and a track, calculated assuming that the muon has the π mass and the track the K mass.

Table 5.5: Central developed muon identification at CMS.

Name	Description
Loose ID	Requires a μ to be a PF - should be used in combination with more cuts
Soft ID	Tuned to give high efficiency in low p_T μ
Medium ID	Tuned to remove in μ from decays in-flight
Tight ID	Tuned for μ from gauge boson decays
HighPt ID	Tuned for high p_T μ ; used for Z' , W'

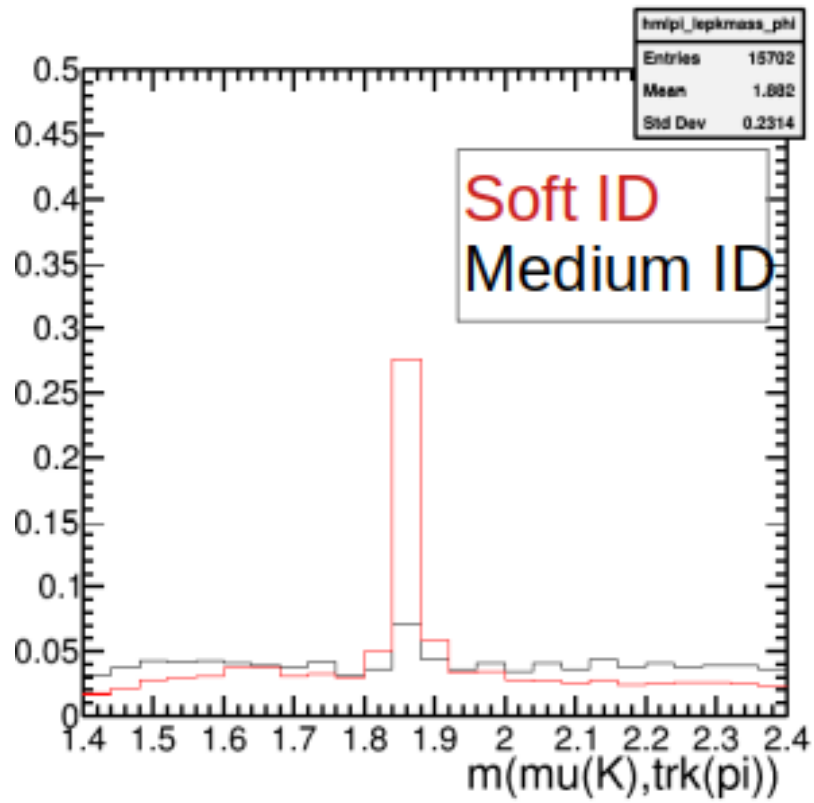


Figure 5.9: Invariant mass of each of the reconstructed muons in the triplet and the track, when the candidate passes medium (in black) and soft (in red) ID.

5.5.2 Electrons

Sec. 4.4.2 describes the standard electron reconstruction algorithm in CMS. In addition to the standard electron collection from the PF algorithm, known as PF electrons, a new electron reconstruction algorithm, called the low- p_T (LP) electron reconstruction algorithm, is developed specifically targeting better reconstruction efficiency at the low- p_T regime.

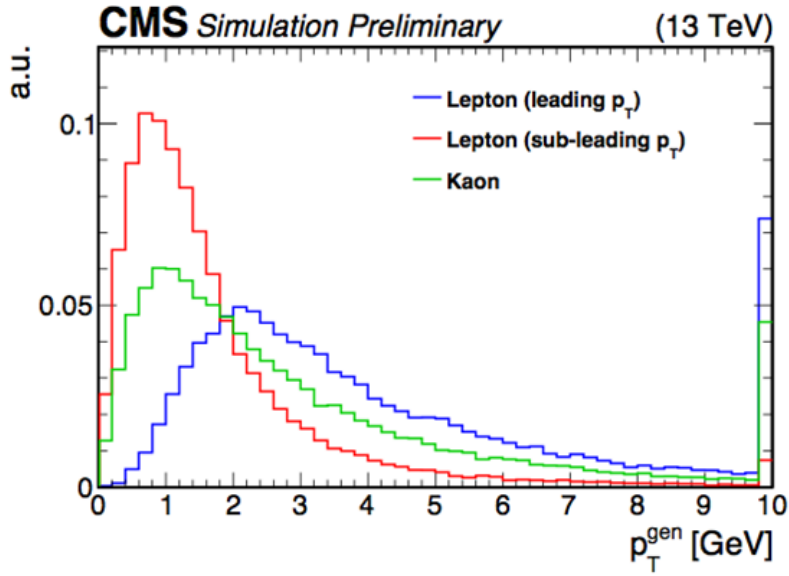


Figure 5.10: Distributions of the transverse momenta of the final state particles for the $B^+ \rightarrow K^+ \ell^+ \ell^-$ decay. The last bin contains the overflow.

Electron reconstruction

In the R_K measurement analysis, it is crucial to identify electrons down to very low transverse momenta efficiently. Figure 5.10 shows the p_T distributions for the particles from $B^+ \rightarrow K^+ \ell^+ \ell^-$ decays. The transverse momenta of all particles are very soft, with those for the kaon and subleading electron peaking at 1 GeV. The standard PF electron reconstruction algorithm is tuned for high efficiency at the W and Z boson leptonic decays energies, i.e., from transverse momenta higher than 15–20 GeV. In order to increase efficiency in the low- p_T

regime, a low- p_T electron reconstruction algorithm is developed for this analysis.

For the standard CMS electron algorithm and the LP algorithm, determining charged-particle track parameters for electron candidates involves using the Gaussian sum filter (GSF) approach to account for Bremsstrahlung energy loss. However, since the GSF tracking stage is computationally expensive, a more efficient logic is used to identify potential electron candidates. The trajectory of each GSF track is then utilized to find a compatible "seed" cluster of energy in the CMS electromagnetic calorimeter. The seed cluster is then combined with additional clusters of energy, which follow the Bremsstrahlung energy loss pattern of the electron candidate, to form a "super cluster". By using the tracking information and the super cluster, it is possible to identify genuine electron candidates with high efficiency and purity.

Two independent models based on boosted decision trees (BDT) are used for the seeding logic. The first BDT implements a "kinematically agnostic" approach, which discriminates signal from background based only on tracking and calorimeter information. The second BDT is a "kinematically aware" model, which uses additional variables such as p_T , η , and track impact parameter of an electron candidate for discrimination. Figure 5.11 displays the ROC curves for the two BDTs, obtained using simulated $B^+ \rightarrow K^+ e^+ e^-$ events. A loose working point is defined for each BDT, which yields a 10% mistag rate and approximately doubles the efficiency over the baseline seeding logic of the standard CMS algorithm. The loose working points from the two BDTs are combined using a logical OR to identify electron candidates and seed the reconstruction of LP electrons.

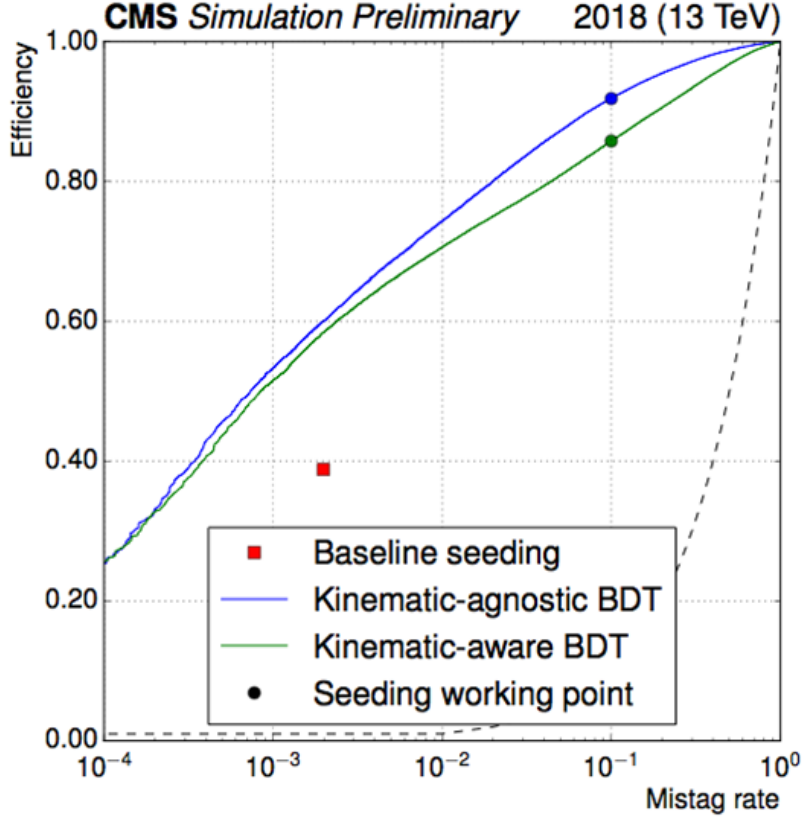


Figure 5.11: Efficiency vs. mistag rate for electrons calculated using different working points of the seeding BDT. The red square indicates the PF electron working point, the blue curve shows the "kinematically agnostic" BDT that exploits tracking and calorimeter information, and the model-dependent "kinematically aware" BDT (green curve) that also utilises the p_T , η , and the track impact parameter of an electron candidate.

Electron energy measurement

The CMS electromagnetic calorimeter (ECAL) measures the energies of electrons with low transverse momentum (p_T), while the silicon tracker measures their momenta. The calorimetric energy resolution is highest at high energies, whereas the tracker momentum measurement is most accurate at lower energies. In this analysis, the electron's momentum is determined by the "mode" of the GSF track momentum, which provides the best resolution.

Electron identification

Electron identification is very important in this analysis to discriminate soft electrons from substantial backgrounds. Two separate electron identifications for PF and LP electrons are developed based on BDT. Features, including both track-related quantities and calorimetric shower shapes, features related to the matching of the extrapolated track to the calorimeter cluster, and the fractional difference between the track momentum at the innermost tracker layer and at the outermost tracker layer, f_{brem} , are fed to the BDT for training.

For the LP identification, electron candidates from a sample of simulated $B^+ \rightarrow K^+ J/\psi (\rightarrow e^+ e^-)$ events are used to train the BDT with the XGBoost algorithm [59]. LP electrons matched within $\Delta R < 0.03$ to a generated electron are used as signals in training. Fake LP electrons that are not matched within $\Delta R < 0.1$ to a generated electron are used as background in training. The background p_T and η distribution of the training sample have been reweighted to reproduce the signal distribution to avoid biases. The list of input variables in the LP electron ID is reported in Table 5.6.

Table 5.6: Variables used in the identification BDT for LP electrons.

LP variables
R_9 , $SC - \eta$ width, $SC - \phi$ width, H/E
N_{KF} hits, χ_{KF}^2 , χ_{GSF}^2 , N_{GSF} hits,
E/p , f_{brem} , $shFracHits$,
$ \Delta\eta_{in} $, $ \Delta\phi_{in} $, $ \Delta\eta_{seed} $,
η_{SC} , E_{SC} , p_{KF} , p_{GSF} , ρ ,
$ \Delta R(ele, KF) $, $ \Delta R(ele, GSF) $,
$Clus1$: E , $N_{crystals}$, $ \Delta\eta $, $\Delta\phi$, E/p ,
$Clus2$: E , $\Delta\eta$, $\Delta\phi$, E/p ,
$N_{clusters}$, seeding BDT output

The default hyper-parameters are used in training the LP electron ID except for "depth" and "number of trees" to prevent overtraining. In Fig. 5.12 (left), the distribution of the

LP electron BDT is presented for both signal electrons and fakes, demonstrating effective separation between the two categories. The simulated sample is split into a training set and a validation set. The latter is used to evaluate the algorithm's performance, as depicted in Fig. 5.12 (right), where the efficiency of electrons versus fakes is plotted. The training sample is excluded from the analysis.

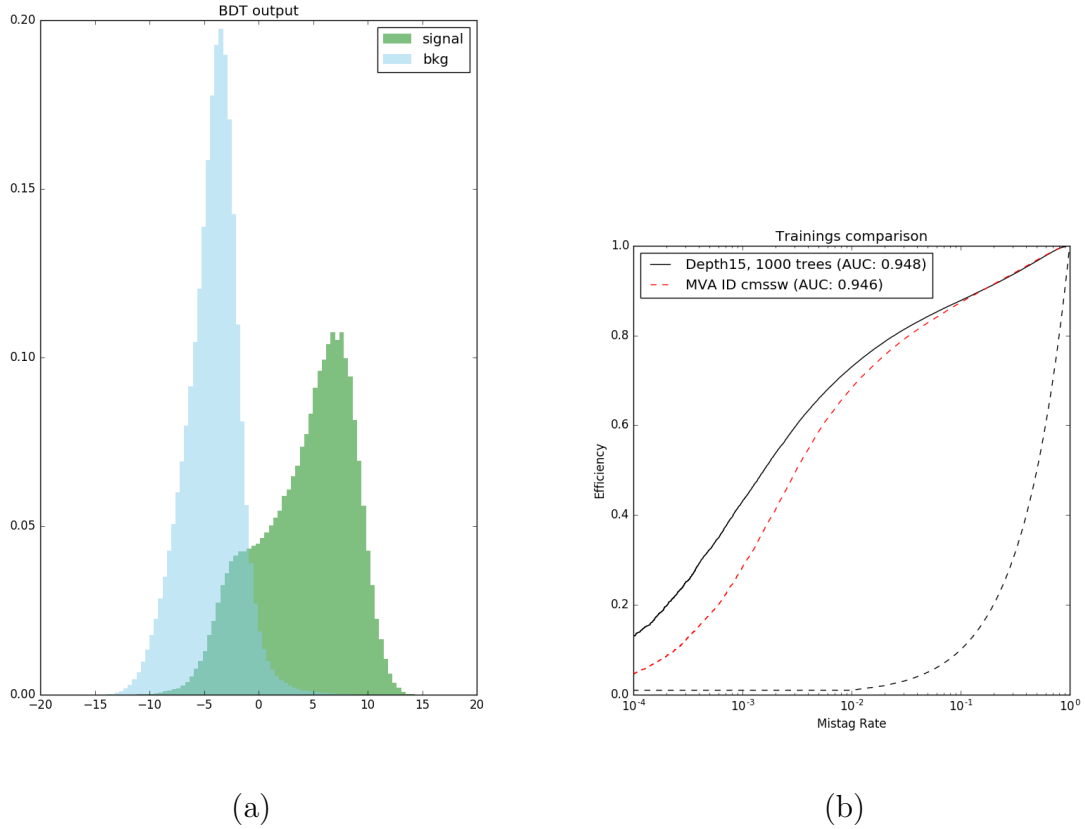


Figure 5.12: (left) Distribution of the LP electron ID BDT for low- p_T electrons matched to signal electrons (green) and for fakes (blue); (right) selection efficiency for signal LP electrons versus efficiency on fakes based on cuts on the LP electron ID BDT (black). Also shown is the corresponding curve for the model in the Jan/2020 processing (red).

The standard PF identification algorithm based on a BDT, as described in [60], was trained for electrons with energies down to 5 GeV. However, the performance of the BDT outside of the training phase space can be unpredictable. Therefore, improving identification

performance for low-energy electrons is necessary. To achieve this, the PF electron ID BDT is re-trained using a simulated sample of $B^+ \rightarrow K^+ J/\psi (\rightarrow e^+ e^-)$ electrons with energies down to 2 GeV. The training sample's signal p_T and η distribution were reweighted to replicate the background distribution, preventing biases. Two separate BDTs were used for two p_T regions: $p_T < 5$ GeV and $p_T > 5$ GeV. For the lower energy bin $p_T < 5$ GeV, fewer variables are used, while for the higher energy bin, more variables are used, as listed in Table 5.7.

Table 5.7: Variables used in the identification BDT for PF electrons.

PF variables ($p_T < 5$ GeV)	PF variables ($p_T > 5$ GeV)
$\sigma_{i\eta i\eta}, \sigma_{i\phi i\phi}$, circularity, R_9 , $SC\eta$ width, $SC\phi$ width, H/E , N_{KF} hits, χ_{KF}^2 , χ_{GSF}^2 , N_{GSF} hits, N_{inner}^{exp} hits, E/p , E/p_{out} $ \Delta\eta_{in} $, $ \Delta\phi_{in} $, $ \Delta\eta_{seed} $, ρ	$\sigma_{i\eta i\eta}, \sigma_{i\phi i\phi}$, circularity R_9 , $SC\eta$ width, $SC\phi$ width, H/E N_{KF} hits, χ_{KF}^2 , χ_{GSF}^2 , f_{brem} , N_{GSF} hits, N_{inner}^{exp} hits, E/p , E/p_{out} , $1/E - 1/p$, $ \Delta\eta_{in} $, $ \Delta\phi_{in} $, $ \Delta\eta_{seed} $, $I_\gamma, I_{ChHad}, I_{NeutralHad}, \rho$

The BDT training configuration used in this analysis was exactly the same as that used by the EGamma POG, with the exception of the changes listed previously. Additionally, early stopping was used to prevent overtraining. The ROC curves of the re-trained BDT using cross-validation are shown in Fig. 5.13, which indicates that the model has low variance and is not subjected to overtraining. Figure 5.14 shows the electron efficiency vs. fake rate for the PF electron ID BDT used in this analysis (red curve) compared to the default PF electron ID BDT used in CMS (blue and green curves). The performance was compared for electrons with $p_T < 5$ GeV and $5 < p_T < 10$ GeV, and in both cases, the performance improved compared to the default.

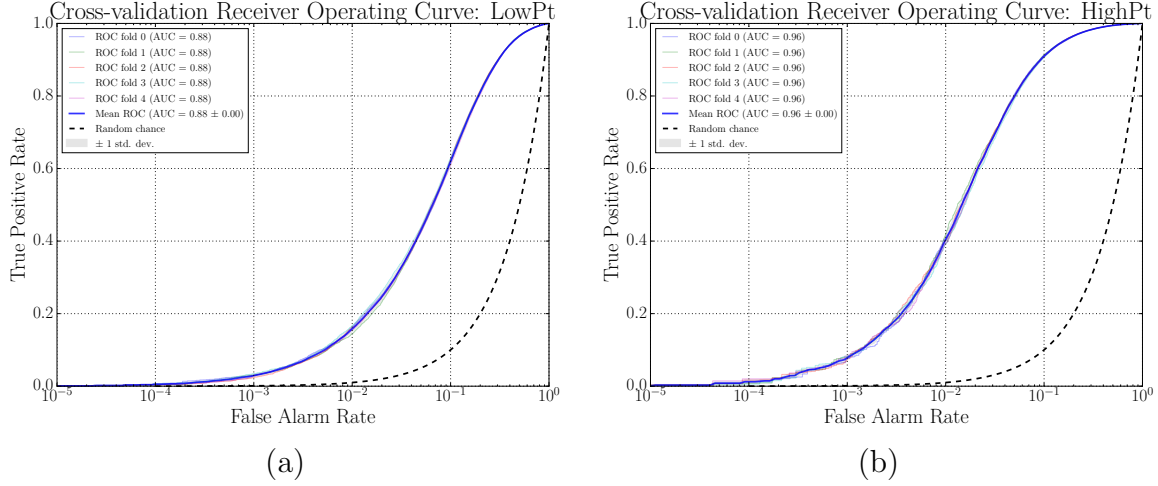


Figure 5.13: ROC curves with cross-validation for the re-trained PF electron ID in kinematic region of (a) $p_T < 5$ GeV and (b) $p_T > 5$ GeV.

Electron identification cross-checks

To evaluate the performance and reliability of the electron identification BDTs, several tests are conducted. The agreement between the BDT outputs in data and simulation is verified using the tag-and-probe technique, utilizing the J/ψ as a reference. In data, events enriched in $B^+ \rightarrow K^+ J\psi(\rightarrow e^+e^-)$ decays are selected using a cut-based selection requiring electrons (kaons) with $p_T > 1$ GeV (0.7 GeV) and $|\eta| < 2.4$. The B candidates must satisfy $p_T > 5$ GeV, vertex probability > 0.1 , $\cos \alpha > 0.99$, and $4.7 < m(K^+e^+e^-) < 5.6$ GeV. Furthermore, the invariant mass of the two electrons must be close to the J/ψ mass, which is ensured by imposing $2.6 < m(e^+e^-) < 3.4$ GeV. A relatively tight selection is then applied to one of the two electrons (tag) to reduce background without cutting on the electron to be tested (probe). Specifically, when considering LP probe electrons, the tag is required to be a PF electron with PF electron ID > 0 , and when considering PF probe electrons, the tag is required to be an LP electron with LP electron ID > 4 .

After the events are selected, the identification BDT output for signal electrons is vali-

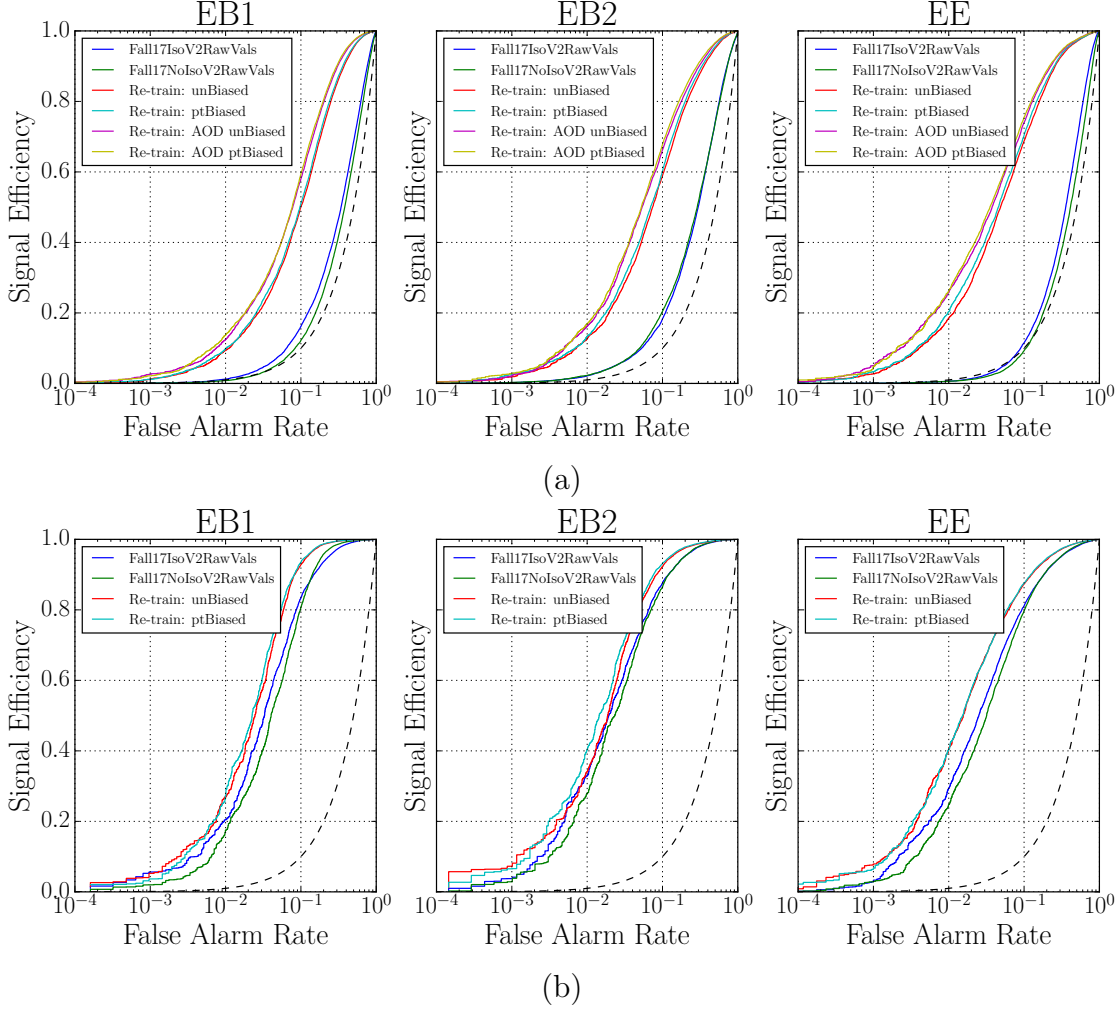


Figure 5.14: Selection efficiency for signal PF electrons versus efficiency on fakes based on cuts on the re-trained PF electron ID BDT (red). Also shown are the corresponding curves for the default BDTs provided by the EGamma POG (blue and green). Electrons with a) $p_T < 5$ GeV and b) $5 < p_T < 10$ GeV

dated in data using the *sPlot* technique [61] for background subtraction, which is performed in bins of the probe electron η and p_T . A maximum likelihood fit is made to the $m(e^+e^-)$ distribution in data to calculate an event-by-event signal probability (s-weight), which is then used to extract the BDT distribution for signal electrons. The signal model used in the fit is a Double Crystal-ball function [62], and the background model is an exponential function. The exponential slope is left free in all fits. For the signal, in the barrel, all parameters are

initialized to the results of a fit to MC and then floated in the fit to data. However, for the endcaps, the signal statistics are poor, and many parameters have to be constrained in the fit. As an example, Fig. 5.15 shows the fit to data for events with a probe LP electron in the barrel with $1.5 < p_T < 2$ GeV (left) and for events with a probe PF electron in the barrel with $2 < p_T < 5$ GeV (right). These fits refer to a previous version of the analysis ntuples, where an energy regression was applied to electrons. However, it has been verified that the electron identification distributions and the data/MC level of agreement are not affected by the regression. Finally, Table 5.8 and 5.9 compare the Crystal-ball mean and sigma values in data and MC for these fits in the barrel case.

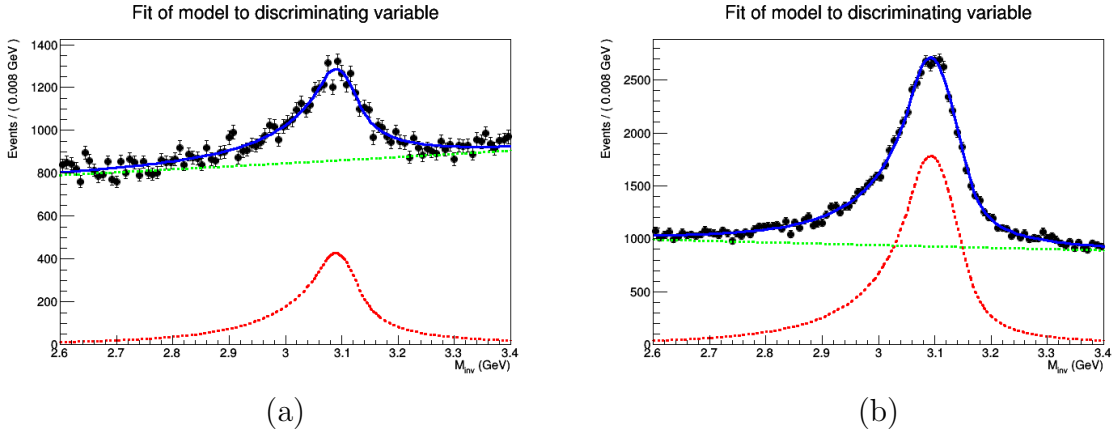


Figure 5.15: (left) Fit to the two electrons invariant mass distribution for events with a (left) probe LP electron in the barrel with $1.5 < p_T < 2$ GeV and (right) probe PF electron in the barrel with $2 < p_T < 5$ GeV. Data (dots) are fitted with a double crystal ball (red) as a signal component and an exponential (green) as a background component.

Table 5.8: Signal Crystal Ball mean and sigma values for fits to MC and data for PF-PF.

	MC mean	Data mean	MC sigma	Data sigma
$2.0 < p_T < 5.0$ GeV	3.0978 ± 0.0004	3.0964 ± 0.0002	0.0337 ± 0.0002	0.0452 ± 0.0008
$p_T > 5.0$ GeV	3.0869 ± 0.0009	3.0866 ± 0.0005	0.0478 ± 0.0010	0.0519 ± 0.0007

The BDT distribution of the signal in data is compared to the MC predictions for LP and PF electrons separately in the barrel, as shown in Fig. 5.16 and Fig. 5.17, respectively.

Table 5.9: Signal Crystal Ball mean and sigma values for fits to MC and data for PF-LP.

	MC mean	Data mean	MC sigma	Data sigma
$1.0 < p_T < 1.5$ GeV	3.0867 ± 0.0024	3.0920 ± 0.0002	0.0399 ± 0.0028	0.0569 ± 0.0027
$1.5 < p_T < 2.0$ GeV	3.0895 ± 0.0003	3.0891 ± 0.0022	0.0358 ± 0.0015	0.0419 ± 0.0026
$2.0 < p_T < 5.0$ GeV	3.0922 ± 0.0002	3.0894 ± 0.0006	0.0444 ± 0.0014	0.0478 ± 0.0008
$p_T > 5.0$ GeV	3.0909 ± 0.0013	3.0892 ± 0.0006	0.0468 ± 0.0018	0.0578 ± 0.0077

In both cases, a satisfactory agreement between the simulation and data distributions is observed. For electrons in the endcaps, the statistics are limited, but Fig. 5.18 and Fig. 5.19 show that a good agreement is also achieved.

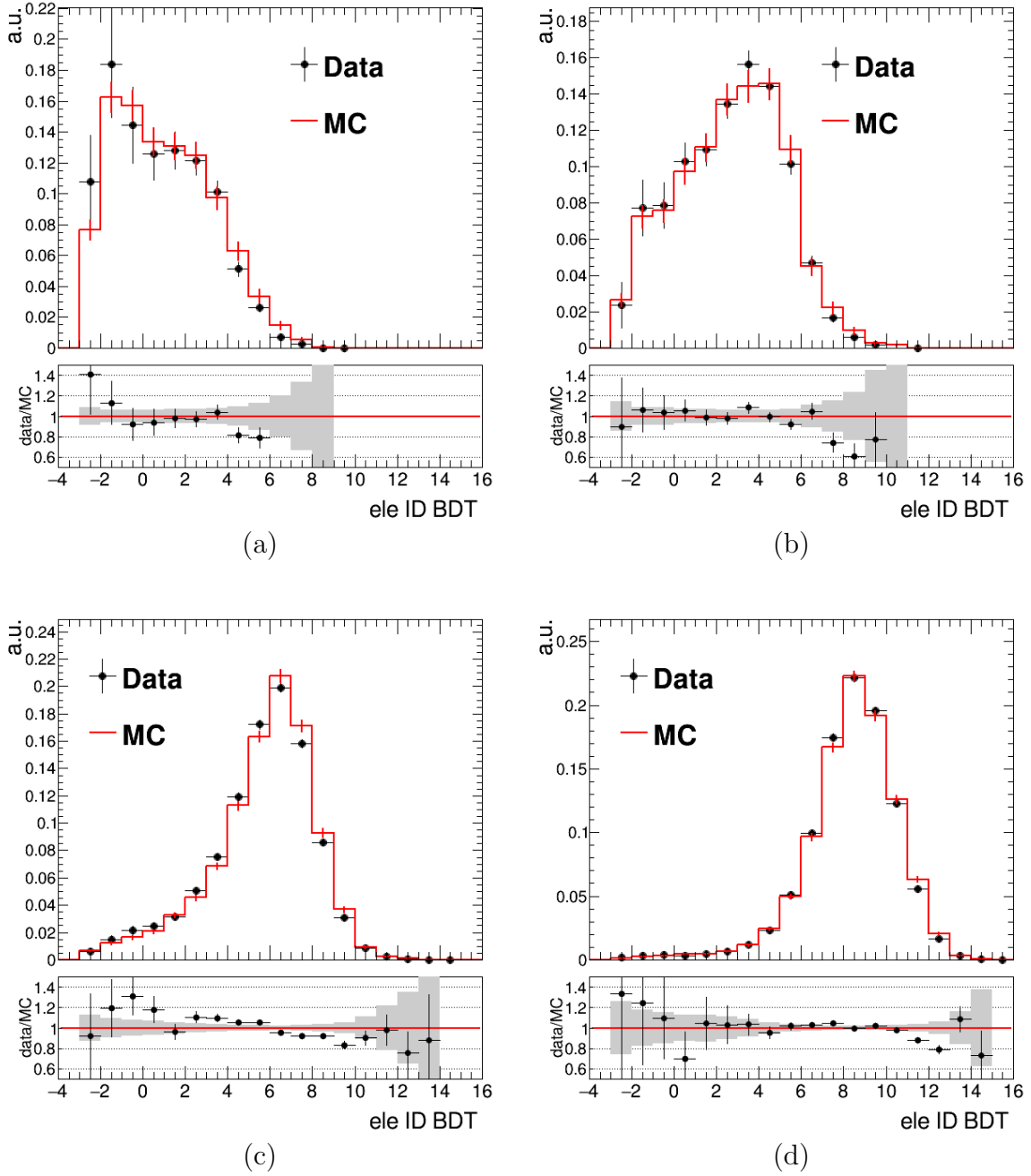


Figure 5.16: Distribution of the LP electron ID BDT for LP electrons in data (dots) and in MC (solid line); the *sPlot* technique is used to remove the background in data. Probe electrons in the barrel are binned in p_T : (a) $1 < p_T < 1.5$ GeV, (b) $1.5 < p_T < 2$ GeV, (c) $2 < p_T < 5$ GeV, (d) $p_T > 5$ GeV

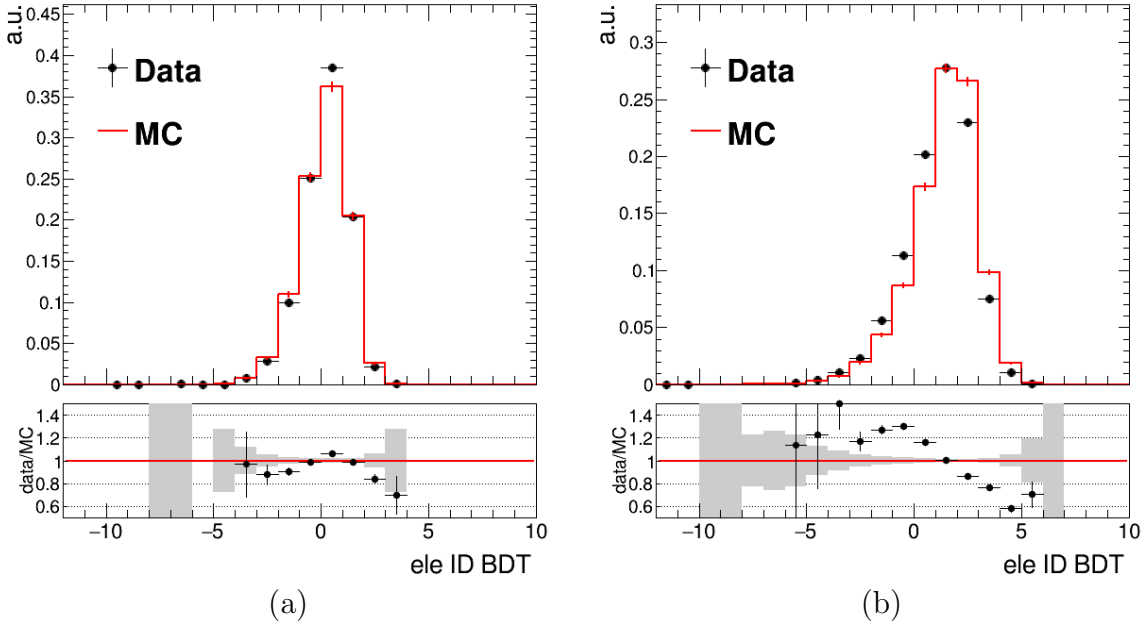


Figure 5.17: Distribution of the PF electron ID BDT for PF electrons in data (dots) and in Monte Carlo (solid line); the *sPlot* technique is used to remove the background in data. Probe electrons in the barrel are binned in p_T : (a) $2 < p_T < 5$ GeV, (b) $p_T > 5$ GeV

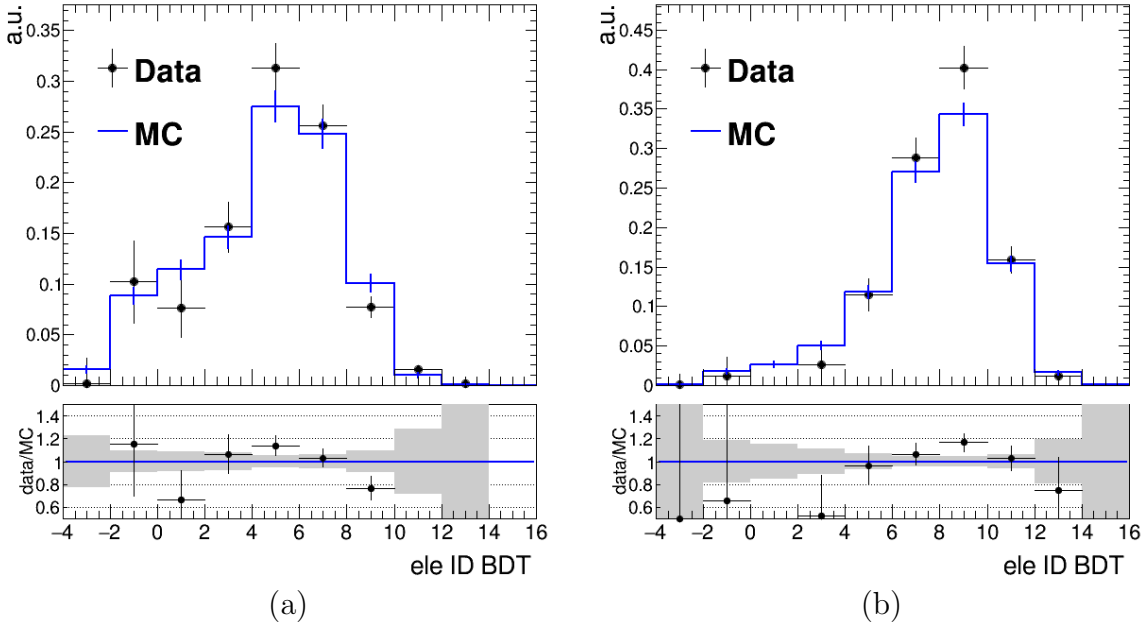


Figure 5.18: Distribution of the LP electron ID BDT for LP electrons in data (dots) and in Monte Carlo (solid line); the *sPlot* technique is used to remove the background in data. Probe electrons in the endcaps are binned in p_T : (a) $2 < p_T < 5$ GeV, (b) $p_T > 5$ GeV

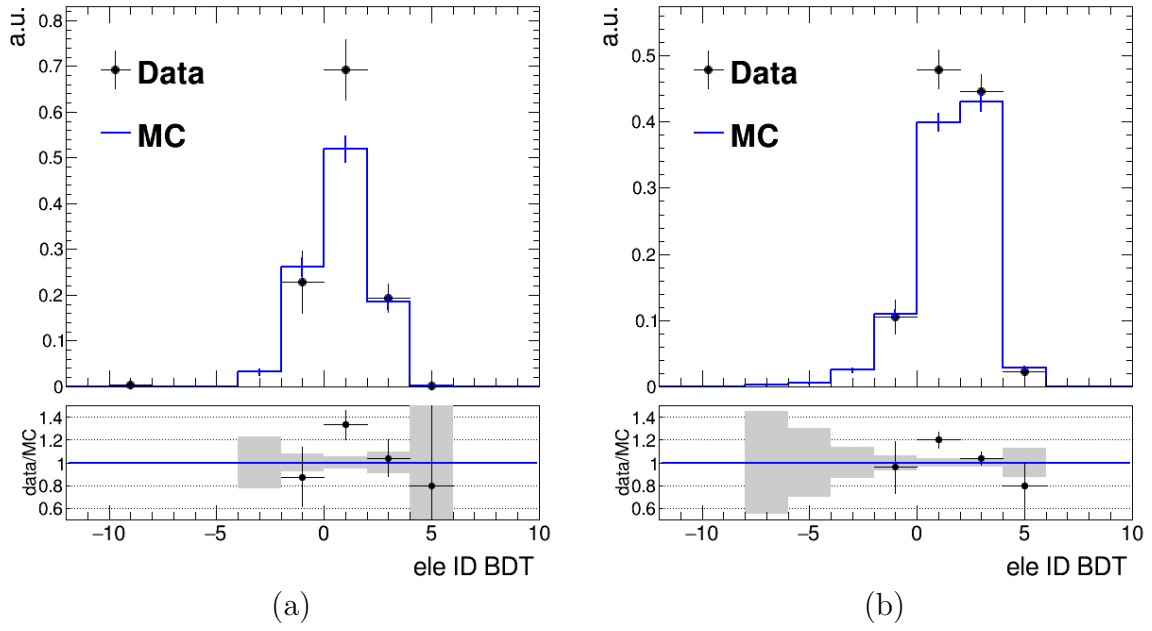


Figure 5.19: Distribution of the PF electron ID BDT for PF electrons in data (dots) and in Monte Carlo (solid line); the *sPlot* technique is used to remove the background in data. Probe electrons in the endcaps are binned in p_T : (a) $2 < p_T < 5$ GeV, (b) $p_T > 5$ GeV

5.5.3 Kaons

Since CMS lacks a particle identification system, it is not possible to distinguish between charged hadrons such as π and K . As a result, in this analysis, the only viable approach for reconstructing B candidates is to combine di-lepton pairs with each track in the event that serves as a candidate for K . In order to mitigate the effects of duplicate and poorly reconstructed tracks, all tracks are required to pass the "High Purity" identification criteria, as recommended by the Tracking POG. Additionally, to enable fitting to a common vertex, all tracks must possess a valid and accurate covariance matrix.

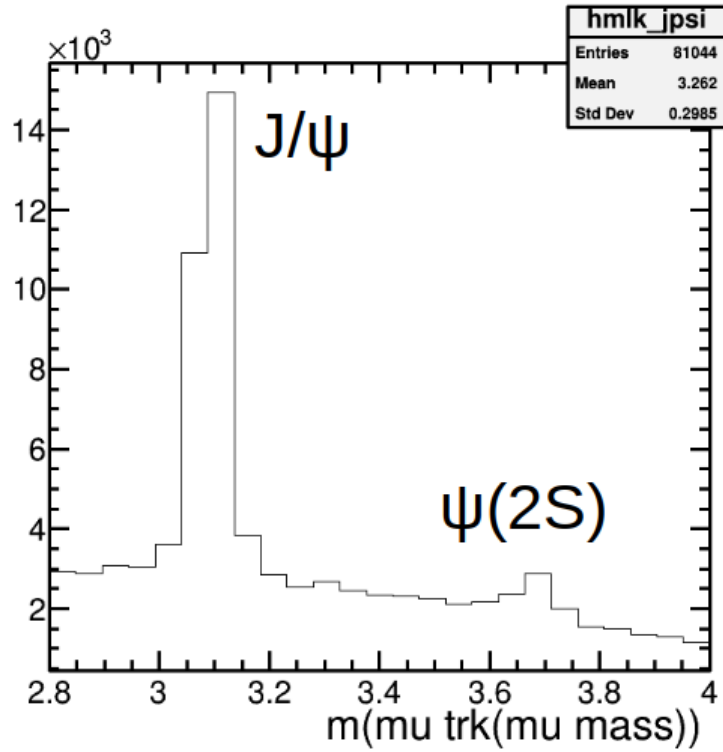


Figure 5.20: Invariant mass of each of the reconstructed μ in the triplet and the track (assuming a μ mass for the particle generating the track) as observed in data. Both the J/ψ and $\psi(2S)$ peaks are visible in the distribution.

One of the major sources of contamination in the track collection is from soft muons

with low transverse momentum. Muons are reconstructed with high efficiency in the tracker, and in situations where the muon chambers are not accessible, the muon is often mistakenly identified as a charged hadron. To study this phenomenon, triplets are selected that contain same-sign muons and opposite-sign tracks, which effectively suppresses the known SM backgrounds. The invariant mass of each track (which is explicitly verified to not belong to a reconstructed muon) and each of the reconstructed muons in the triplet exhibit distinctive peaks at the masses of J/ψ and $\psi(2S)$, as shown in Fig. 5.20. The track is assigned with muon mass hypothesis in this test.

In order to confirm that the track collection consists of fakes from soft muons, the p_T distribution of the track in the barrel and endcap regions are plotted as shown in Fig. 5.21. The p_T distribution in the barrel has a significant drop above 3.5 GeV, which agrees with the minimum momentum needed for a muon to reach the muon stations in the barrel. On the other hand, no kinks are observed in the distribution of the endcap region since muons need only p_T greater than 1 GeV to reach the muon stations. In this analysis, we cannot increase the p_T threshold on track to suppress these fakes as this will cause 40% decrease in signal efficiency. Instead, we exclude the triplet candidates in which the invariant mass of the track and the muon are close to the J/ψ resonance.

5.5.4 Kinematic fit

To form B^+ candidates, we combine a candidate K^+ meson with two oppositely charged leptons. The tracks of the two leptons and the hadron are then fitted to a common vertex using the KinematicFit package in CMSSW. This algorithm uses Least Means Square min-

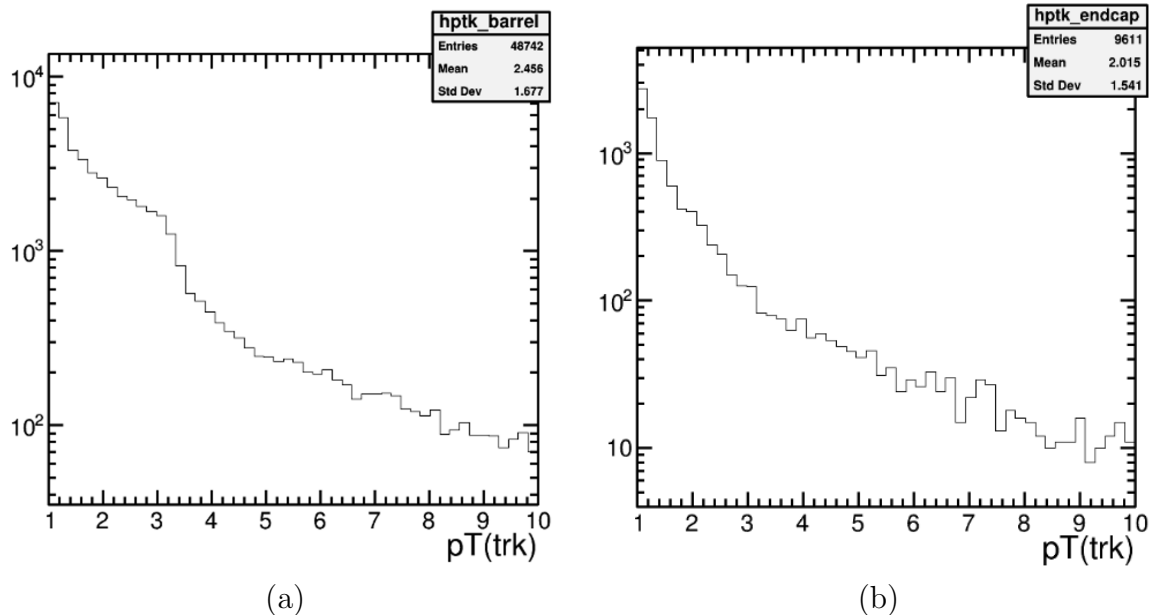


Figure 5.21: (a) p_T distribution for tracks in the barrel region; (b) p_T distribution for tracks in the endcap region.

imisation to constrain the tracks of the B^+ daughter particles to come from a single vertex, and provides us with the vertex position, covariance matrix, and χ^2 . A detailed description of the algorithm can be found in Ref. [63].

After the vertexing step, the particle trajectories are refitted using the vertex as an additional constraint. This allows for a recalculation of their momenta, which are used in the subsequent analysis. For electrons, the track from the GSF algorithm is provided to the fitter. Figures 5.22 and 5.23 present a comparison between the fitted mass and the raw mass using MC events in J/ψ and $\psi(2S)$, respectively. The kinematic fit results in a better resolution of the mass compared to the raw mass, resulting in more events being located at the center of the mass peak.

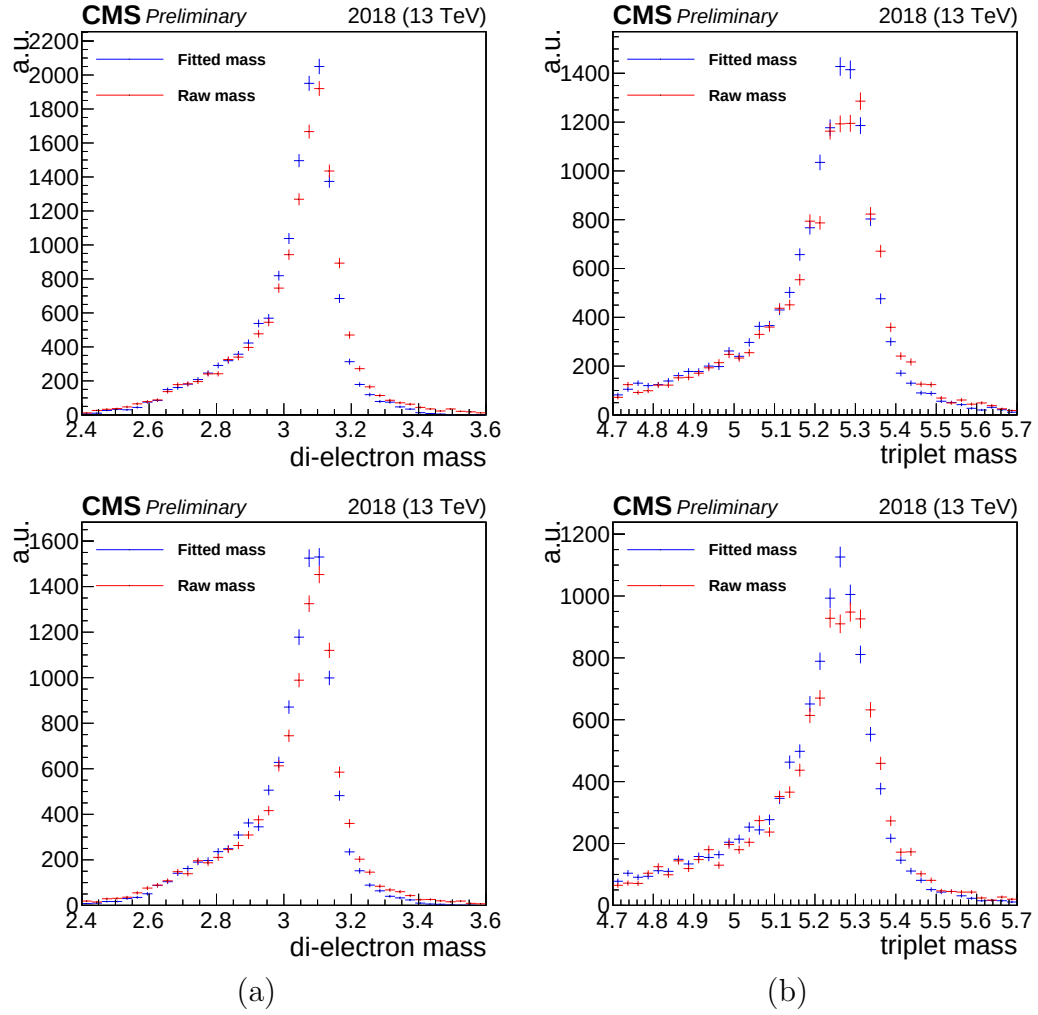


Figure 5.22: Comparison between fitted and raw mass of $m(e^+e^-)$ (a), and $m(K^+e^+e^-)$ (b), with PF-PF (top) and PF-LP (bottom) in the J/ψ bin.

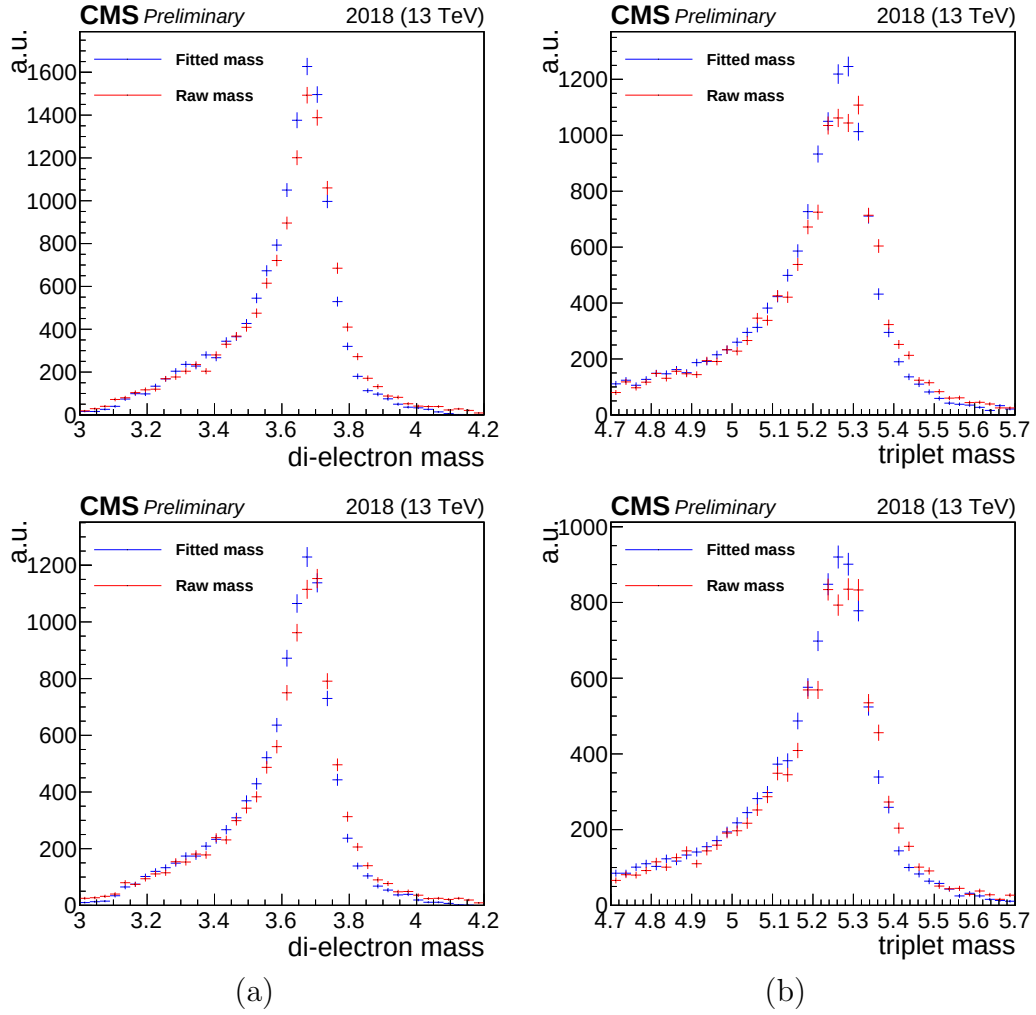


Figure 5.23: Comparison between fitted and raw mass of $m(e^+e^-)$ (a), and $m(K^+e^+e^-)$ (b), with PF-PF (top) and PF-LP (bottom) in the $\psi(2S)$ bin.

5.6 Background sources

This section describes the background sources encountered in the R_K measurement. Any triplet that falls within the B mass range can potentially contaminate the signal mass fit. The following outlines the origin of these background sources and how to handle them.

Combinatorial background is the main source of background, which originates from particles selected from multiple hadron decays. The particles are uncorrelated, creating a falling mass distribution without any structure. The combinatorial background is suppressed by the analysis BDT's selection.

A double peak structure appears in the lower sideband of the $m(K^+\ell^+\ell^-)$ distribution as a result of partially reconstructed decays of $B^{0,+} \rightarrow K^*(892)^{0,+}\ell^+\ell^-$ and $B^{0,+} \rightarrow K^*(892)^{0,+}J/\psi(\rightarrow \ell^+\ell^-)$, where the $K^*(892)^{0,+}$ subsequently decays into $K^+\pi^{-,0}$. The origin of the two peaks is the choice of whether the triplet includes the kaon or the pion. The partially reconstructed background is irreducible but sufficiently well modeled by dedicated simulations to be considered in the mass fit.

Cascade backgrounds originate from such as semileptonic decays with the form of $H_b \rightarrow H_c(\rightarrow K^+\ell^-\bar{\nu}X)\ell^+\nu Y$, where H_b is a beauty hadron (B^+ , B^0 , B_s^0 or Λ_b^0), H_c is a charm hadron (D^0 , D^+ , D_s^+ , Λ_c^+), and X, Y are other particles that are not reconstructed, can contaminate the mass spectrum. This background is greatly suppressed by requiring $m(K^+\ell^-_{[\rightarrow\pi^-]}) > 2$ GeV and $m(K^+_{[\rightarrow\pi^+]}\ell^-_{[\rightarrow K^-]}) > 2$ GeV, known as the anti- D^0 cut. The oppositely signed kaon and lepton pair is assigned to a kaon mass or a pion mass labeled by the subscript $[\rightarrow K/\pi]$, respectively. The cut of 2 GeV is chosen to be around the invariant mass of D^0 , to remove the resonances from the charm hadrons.

Cabibbo suppressed decays of $B^+ \rightarrow \pi^+ \ell^+ \ell^-$ and $B^+ \rightarrow \pi^+ J/\psi (\rightarrow \ell^+ \ell^-)$ can create a mass peak under the signal peak. These backgrounds have little contribution due to their small branching fractions, although they are irreducible. Dedicated simulations are used to model the shape of these backgrounds in the mass fit.

A potential peaking background could arise from hadrons misidentified as leptons in the all-hadronic decays of the type $B^+ \rightarrow K^+ h_1 h_2$, dominated by the $B^+ \rightarrow K^+ \pi^+ \pi^-$ decay. Extensive studies show that the required muon and electron identification suppress these backgrounds effectively, described in Sec. 5.8.5.

Other exclusive b -hadron decays that enter this analysis is called "Other B backgrounds". Dedicated generic B decays simulation is generated with the form of $H_b \rightarrow X \ell^+ \ell^-$. After excluding the known decay modes as described above, the residual concentrated in the lower sideband is modelled and considered in the mass fit.

5.7 The muon channel

This section describes the muon part of the R_K measurement, including event selection, fit strategy, systematics evaluation, and sanity checks.

5.7.1 Trigger selection and scale factors

The tag side of the B -parking event is used for the muon part of the R_K analysis. Among the HLT paths being used in B -parking, events collected from HLT_Mu9_IP6 contribute to most data, while others contribute only marginally from their exclusive rate. Therefore, we use only events triggered from HLT_Mu9_IP6 for the muon part of the R_K analysis, in which we benefit from simplifying the analysis by using single trigger path, as well as not affecting the statistical precision of the result.

Figure 5.24 shows the η distribution for muons that fired the HLT_Mu9_IP6 path. The drop around $|\eta| = 1.5$ comes from the η restriction of the L1 seed.

Since our trigger requires a large displacement online, the trigger scale factors provided centrally cannot be applied. Instead, dedicated trigger scale factors need to be measured by following the instructions of the Muon POG. The "tag-and-probe" method is the most commonly used method for measuring the trigger efficiency, and it is extensively used in CMS for different objects and analyses. Although there are numerous packages available, we use the Muon POG package for this study because it is optimized for this type of measurement and centrally maintained.

The tag-and-probe method is used to measure trigger efficiency in an unbiased manner by excluding the object that fired the path. This is achieved by selecting the muon that fired

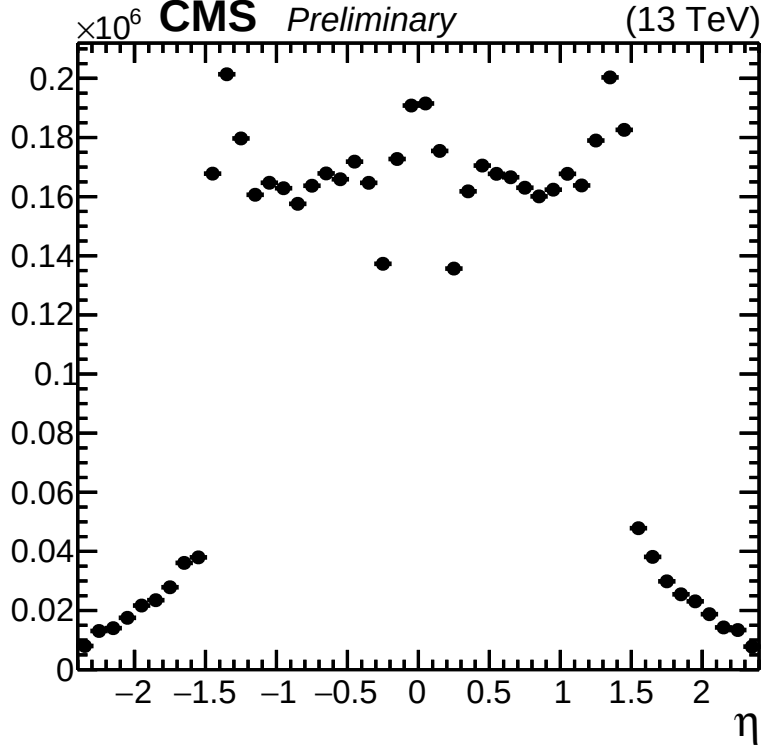


Figure 5.24: Muon η distribution, for muons that fired the HLT_Mu9_IP6, using a subset of the available data. The drop around 1.5 is due to the seed η restriction.

the trigger and passes tight quality selection (known as the "tag"), and combining it with an oppositely charged muon (known as the "probe"). The invariant mass of the pair must be close to a known resonance, which is the J/ψ resonance in our case. J/ψ produces muons in the spectrum of interest and is copiously produced at the LHC. To reduce the number of random combinatorial pairs, the di-muon pair must pass a set of selections. Additionally, a selection is applied to address the fact that in simulation, all triggers are unprescaled (which is equivalent to having all events triggered by the loosest path), while this is not the case in data. This issue is shown in Fig. 5.25, which shows that the p_T distribution for muons in simulation begins at lower values than in data. If not addressed, this difference would result in slightly different topology and incorrect results. As a result, cuts on the tag muon are

imposed to simulate the same topology. Table 5.10 provides a complete list of selections. The probes of the chosen pairs are effectively unbiased and can be used to calculate trigger efficiency. Depending on whether the probe muon can be matched to an HLT muon that passed the trigger under study, it is classified as passing or failing.

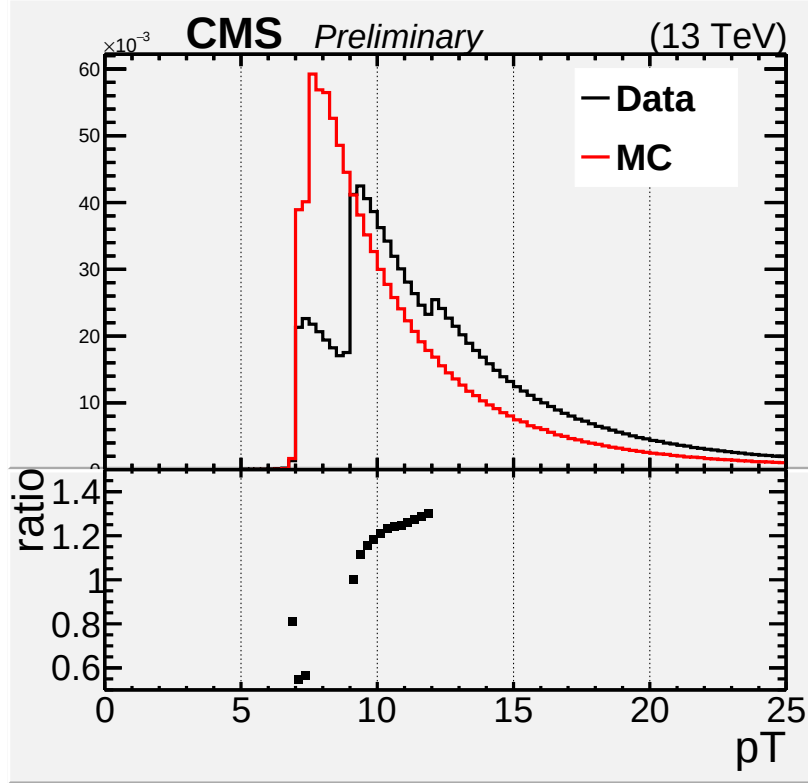


Figure 5.25: Distribution of tag μ in data and simulation (MC). Clearly, in simulation most of the events are triggered by different trigger than in data.

The trigger path we use explicitly involves two variables: the p_T of the muon and its displacement. Therefore, the efficiency must be determined as a function of these variables. We have studied the implicit dependence on other variables, and the only non-negligible one is the mass resolution's dependence on η as shown in Fig. 5.26. As a result, the efficiencies are divided into three η regions to account for this effect.

To determine the efficiency in data, we perform a fit to the mass distributions of failing

Table 5.10: Variables used in tag-and-probe pair selection.

Variables	Cuts	Reason
$\text{Prob}(\mu, \mu \text{ vtx})$	> 0.01	Background suppression
Δz	$< 0.5 \text{ cm}$	Background suppression
d_{xy}/σ (tag μ)	> 8	Data-MC agreement
$p_T(\text{tag } \mu)$	> 10	Data-MC agreement
p_T (L1 tag μ)	$> 10 \text{ GeV}$	Data-MC agreement
$\Delta R(\text{tag } \mu, \text{L1tag } \mu)$	< 0.05	Data-MC agreement
Match tag μ with HLT μ	-	Data-MC agreement
$\Delta R(\text{tag } \mu, \text{probe } \mu)$	> 0.15	Ensure probe and tag are not the same
Muon ID (tag muon)	tight ID	
Muon ID (probe muon)	medium ID	

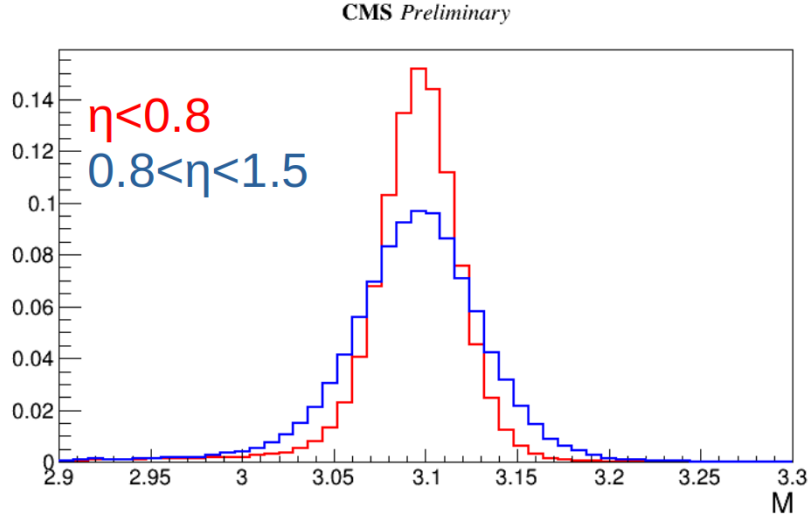


Figure 5.26: Mass distribution of the muon pair in two η regions.

and passing tag-and-probe pairs in bins of η , p_T , and d_{xy}/σ . The fit yields the estimated J/ψ yield. We use MC templates to describe the J/ψ peak, and a second-order polynomial to model the background. An example of these fits is presented in Fig. 5.27. The efficiency is calculated using the following formula:

$$\epsilon = \frac{N_{\text{passing probes}}}{N_{\text{passing probes}} + N_{\text{failing probes}}} \quad (5.4)$$

where $N_{\text{passing probes}}$ is the number of J/ψ with passing probes and $N_{\text{failing probes}}$ is the number of J/ψ with failing probes.

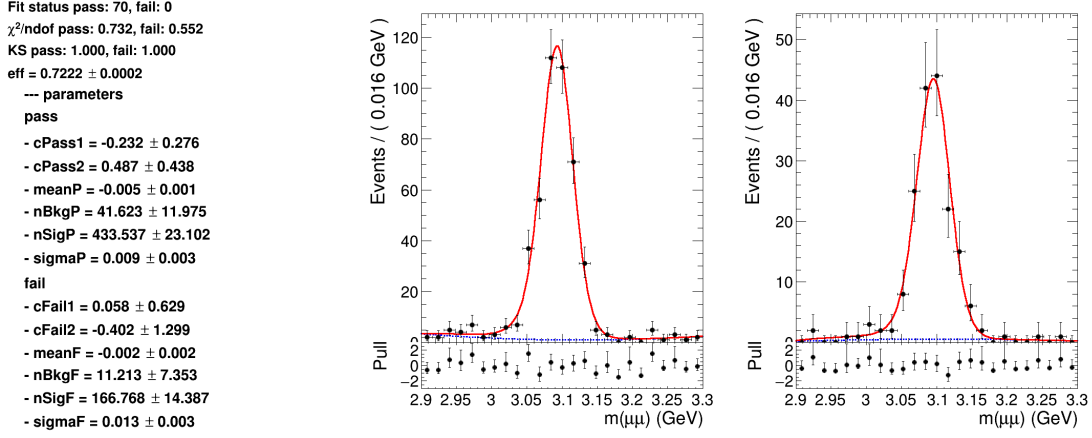


Figure 5.27: Fit for the passing probes (left) and fit for the failing probes (right).

In simulation, tag-and-probe pairs use the same selection as in data, but the efficiency is computed by matching with the generated truth information instead of fitting. The efficiencies are constructed as two-dimensional histograms, and the scale factor is defined as the ratio between data and simulation. In some bins, a large scale factor is obtained because of the limited statistics. These bins correspond to values well below the trigger thresholds (in p_T or d_{xy}/σ), and hence the number of passing probes is insufficient to obtain a proper fit. In the region of interest ($d_{xy}/\sigma > 6$ and $p_T > 9$ GeV), the scale factors are close to 1, which agrees with the scale factors computed in the context of other analyses with a similar trigger. The efficiencies and scale factors are presented in Fig. 5.28, 5.29, and 5.30 for different pseudorapidity bins. The uncertainty includes both statistical and systematic uncertainties, with the systematic uncertainty evaluated by changing the signal and background fit functions.

We apply the Trigger Scale Factors (SFs) to all B candidates by setting their weight ε

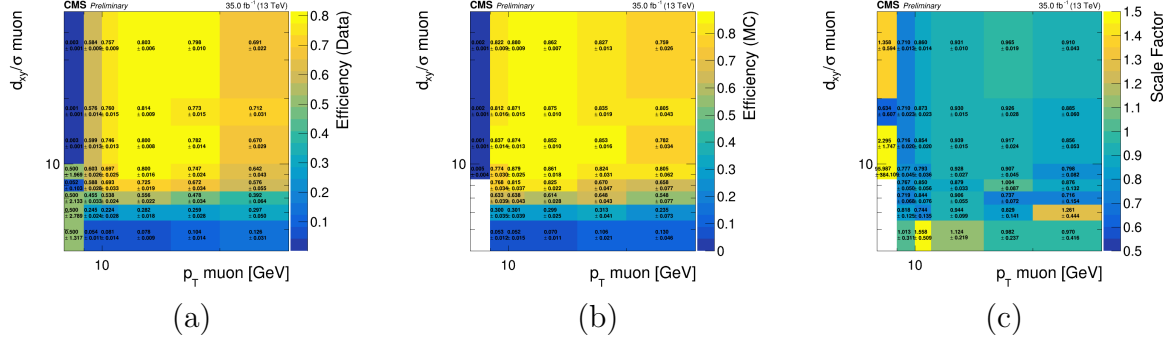


Figure 5.28: Efficiency in data for $|\eta| < 0.4$ (a); Efficiency in simulation for $|\eta| < 0.4$ (b); Scale factor for $|\eta| < 0.4$ (c).

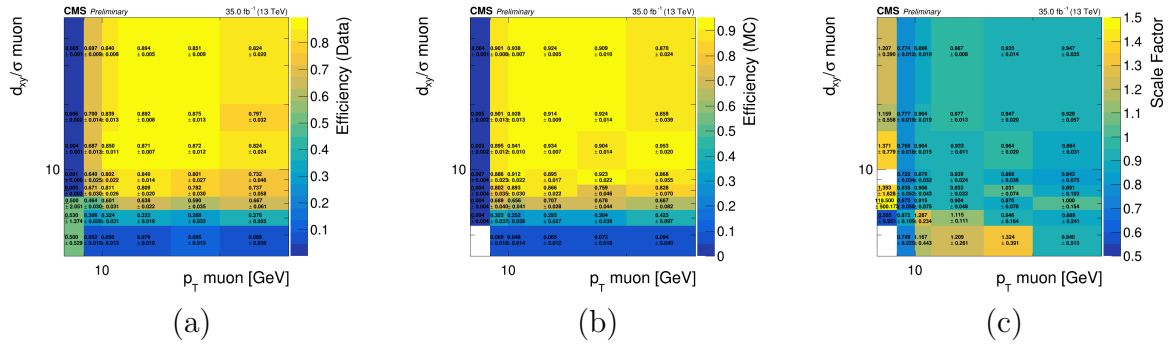


Figure 5.29: Efficiency in data for $0.4 < |\eta| < 0.8$ (a); Efficiency in simulation for $0.4 < |\eta| < 0.8$ (b); Scale factor for $0.4 < |\eta| < 0.8$ (c).

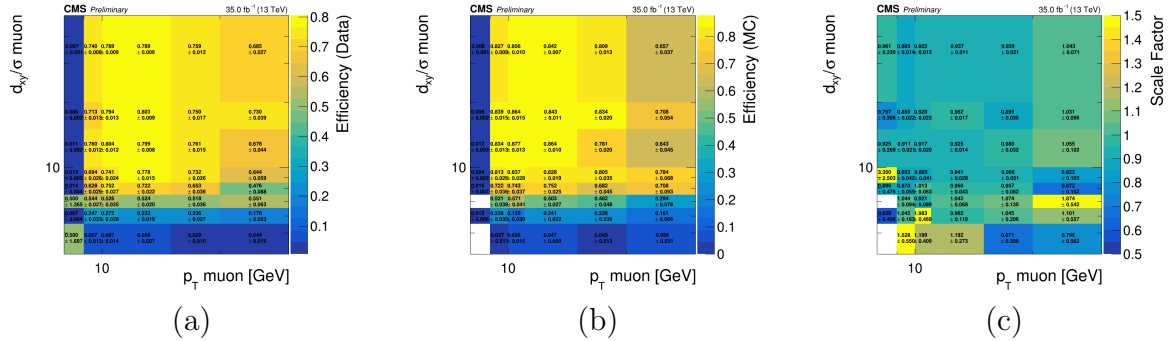


Figure 5.30: Efficiency in data for $0.8 < |\eta| < 1.5$ (a); Efficiency in simulation for $0.8 < |\eta| < 1.5$ (b); Scale factor for $0.8 < |\eta| < 1.5$ (c).

to $\varepsilon_1 + \varepsilon_2 - \varepsilon_1 \times \varepsilon_2$, where ε_i is the Trigger SF for the i -th muon in each triplet. A muon has non-zero ε_i only if it is matched with an HLT muon (and therefore it is considered as tag) and satisfies the trigger online criteria ($p_T > 9$ GeV, $|\eta| < 1.5$ and $|d_{xy}|/\sigma > 6$). Otherwise, the SF ε is zero. We do this to correct the difference between data and MC for a fraction of

subleading muons that are tag. Since the MC has no prescales, a larger fraction of subleading muons are labeled as tags in MC compared to data. This happens because the loosest HLT path is always activated in MC, while in data, the exact trigger is activated later when the L1 rate is low enough. We expect the formula $\varepsilon = \varepsilon_1 + \varepsilon_2 - \varepsilon_1 \times \varepsilon_2$ to correct the difference between data and MC when both muons fire the same trigger. In cases where μ_1 has fired HLT_Mu9_IP6, and μ_2 has fired HLT_MuX_IPY, the second muon is not considered as a tag in both MC and data samples.

Figures 5.31 and 5.32 present the fractions of leading and subleading muons considered as tags in data and MC, respectively. The MC has been weighted with the formula mentioned above, and no further corrections have been applied. The figures show the fractions of B candidates that have the leading/subleading muon labeled as triggering in data, superimposed with the fractions in MC with and without trigger SFs, given the muons passed the online trigger selection criteria. To reduce background, we have cut close to the B mass ($5.15 < m(B) < 5.35$ GeV) and required all candidates to pass the BDT analysis cut (BDT output > 4). The ratio plots show that in all cases, the effect of the HLT SFs is towards the correct direction, as the ratio points are closer to one.

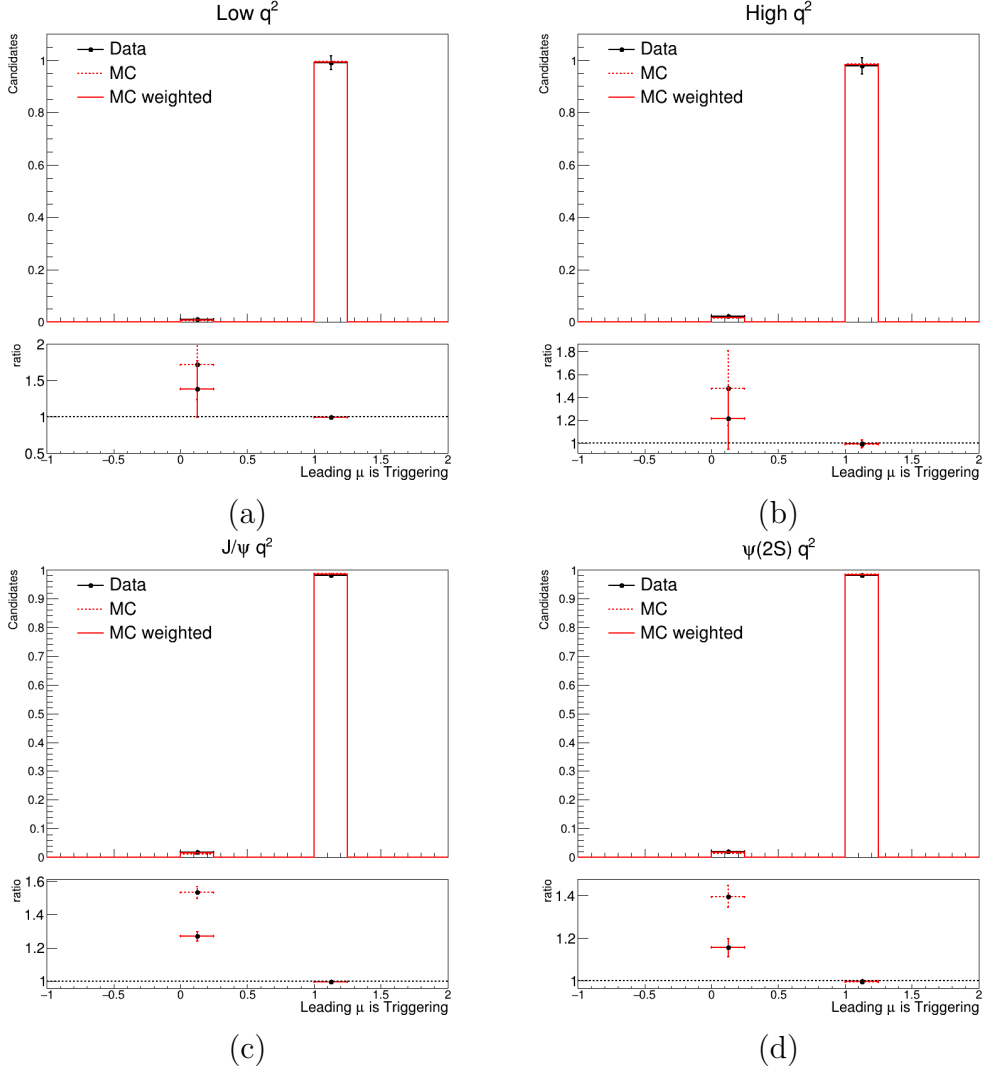


Figure 5.31: Trigger status of the leading muon in all 4 dimuon mass bins: low- q^2 bin (a), high- q^2 bin (b), J/ψ bin (c) and $\psi(2S)$ bin (d).

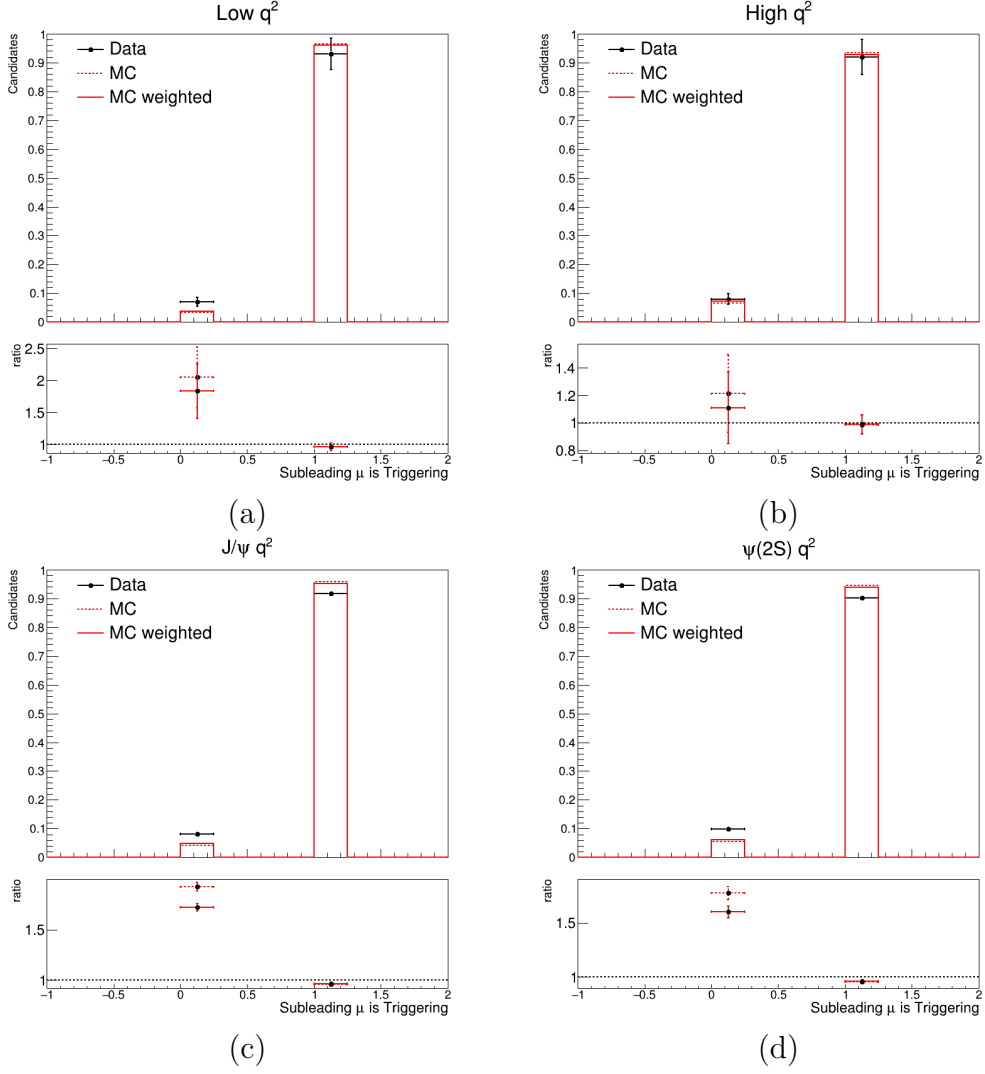


Figure 5.32: Trigger status of the subleading muon in all 4 dimuon mass bins: low- q^2 bin (a), high- q^2 bin (b), J/ψ bin (c) and $\psi(2S)$ bin (d).

5.7.2 Event preselection

To form a $B^+ \rightarrow K^+ \mu^+ \mu^-$ candidate, two oppositely charged muons are combined with a track that is hypothesized to be a K^+ meson. A set of preselection cuts is applied to the B candidates to remove the background, which is summarized in Table 5.11.

Table 5.11: A set of preselection cuts for the muon channel.

Variable	Cut
p_T (B cand.)	$> 3 \text{ GeV}$
Δz (triggering μ , track)	$< 1.0 \text{ cm}$
Δz (triggering μ , μ_2)	$< 1.0 \text{ cm}$
p_T (μ_2)	$> 2 \text{ GeV}$
muon ID	Medium
p_T (trk)	$> 1.0 \text{ GeV}$
Prob (B vertex)	$> 10^{-5}$
$\cos\alpha$ (B cand.)	> 0.90
L_{xy}/σ (B cand.)	> 1.0
mass(μ , track)	$> 2.0 \text{ GeV}$

Loose cuts are applied to other quantities of the B candidate, in addition to the selections applied to the muons. Specifically, these requirements are applied on Prob (B vertex), which is a measure of the quality of the kinematic fit of the B candidate vertex, $\cos\alpha$, which is the cosine between the momentum vector of the B candidate and its direction of flight and L_{xy}/σ , which is the significance of displacement of the B candidate vertex from the Beamspot.

All the preselection cuts are chosen to be very loose before they are fed into a BDT for final selection.

5.7.3 Event selection with a BDT in the muon channel

In order to extract signals from backgrounds, a Boosted Decision Tree (BDT) based on XGBoost package [59] is trained for event selection.

For the BDT training, the simulated $B^+ \rightarrow K^+ \mu \mu$ events are used as a proxy for the signals. While for the background proxy, real data from sidebands ($\pm 6\sigma - \pm 4\sigma$ from the B peak) are used since the background from simulated events may not reflect the background in real data because of the limited knowledge of all the heavy state decays. Both the signal and background are required to pass the preselection and fall in the low- q^2 window. Figure 5.33 shows the ranges of mass distributions used for the selection of training events of background. The hyper-parameters of training are chosen using Bayesian optimization.

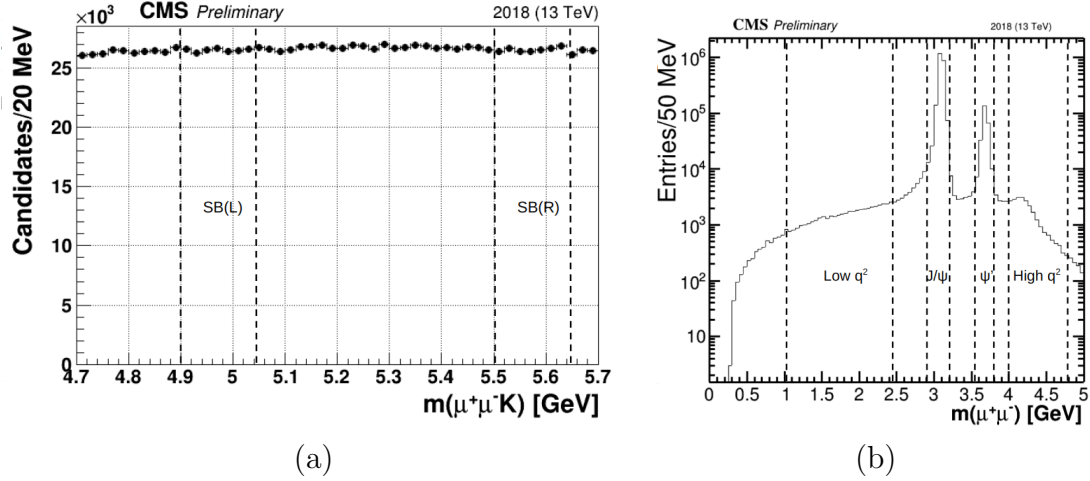


Figure 5.33: Invariant mass of B candidates, with dashed lines indicating the left and right sidebands (SB(L) and SB(R) respectively) (a). Invariant mass of muon pairs, with dashed lines indicating the different q^2 bins (b).

Cross-validation technique

Since we use the data-driven method to get the background proxy, it is important to exclude the data used in training for the final mass fit. However, we would like to utilize the whole data set instead of sacrificing part of the sample. Therefore, we use a technique called cross-validation, where the whole data set is partitioned into subsets, training on one subset while performing the analysis on the another subset.

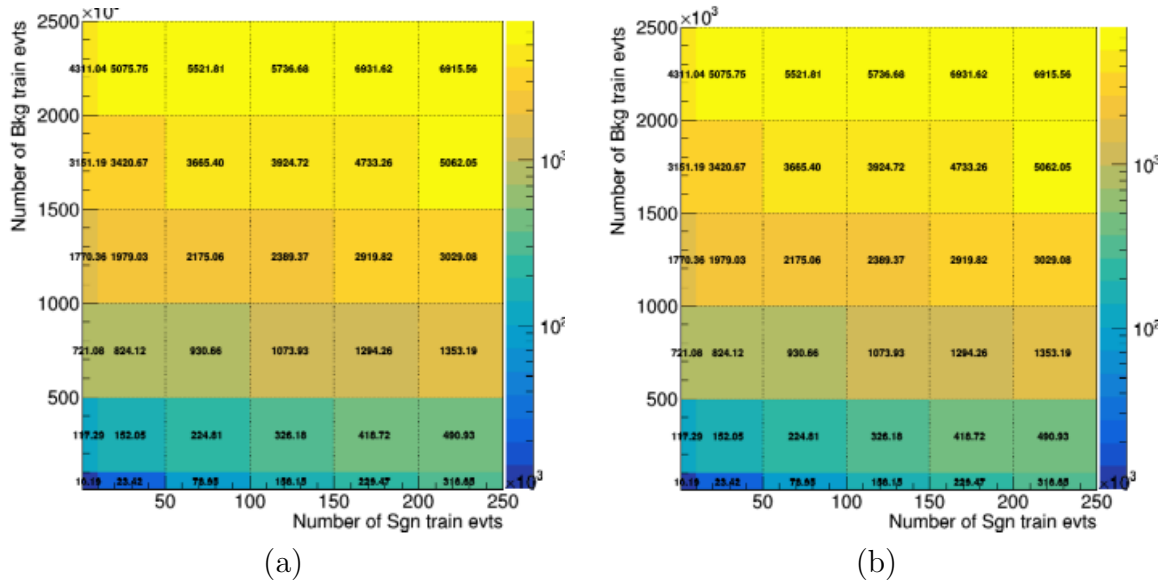


Figure 5.34: Signal efficiency vs Signal and Background training events for a tight BDT cut (a). Time required for training as a function of signal and background events (b).

As shown in Fig. 5.34, there is no improvement in BDT performance beyond $N_{\text{sig}} > 10^5$ and $N_{\text{bkg}} > 1.5 \times 10^6$, but the training time continues to increase. Hence, we set $N_{\text{bkg}} = 1.5 \times 10^6$ and $N_{\text{sig}} = 10^5$ as the number of background and signal training events, respectively. To achieve this, the background sample is divided into eight segments, each containing the same number of events, which are then used as background for different BDTs with the same hyperparameter tuning and input variables as the signal. The performance difference between BDTs is due solely to the stochastic aspect of the training, and is considered as a

systematic error.

Feature selection

Careful selection of features is crucial for achieving good discriminatory power in the BDT. The features are generated based on triplet kinematics, topology, vertex quality, particle identification, and isolation, and are listed in Table 5.12 along with detailed descriptions. Recursive feature exclusion method is carried out to ensure that the selected features effectively improve BDT performance. To simplify the model, unnecessary variables are eliminated.

Table 5.12: Variables used in the analysis BDT in the muon channel.

Variables	Descriptions
$p_T(K^+/B^+)$	Transverse momentum of the triplet of K^+ and B^+ , respectively
min. $dz(\mu, K^+)$	Minimum longitudinal distance between the muons and the kaon
$prob(SV)$	Secondary vertex fit χ^2 probability
$L_{xy}/\sigma_{L_{xy}}$	Secondary vertex transverse displacement significance
$\cos \alpha_{2D}$	Cosine angle in the transverse plane between the B momentum and the separation between the secondary vertex and the beamspot.
min. $\Delta R(\mu, K^+)$	Minimum ΔR between the muons and the kaon
$I_{\Delta R=0.4}^{\text{rel}}(\mu_1)$	Relative PF isolation around the cone $\Delta R < 0.4$ of μ_1

Training result

The BDT distribution of the trained BDT for signal and background is shown in Fig. 5.35. The working point of the BDT for the analysis is chosen at the point such the it maximizes the significance in the low q^2 bin, which corresponds to a selection cut $BDT > 4$. Figure 5.36

shows the significance as a function of BDT working points.

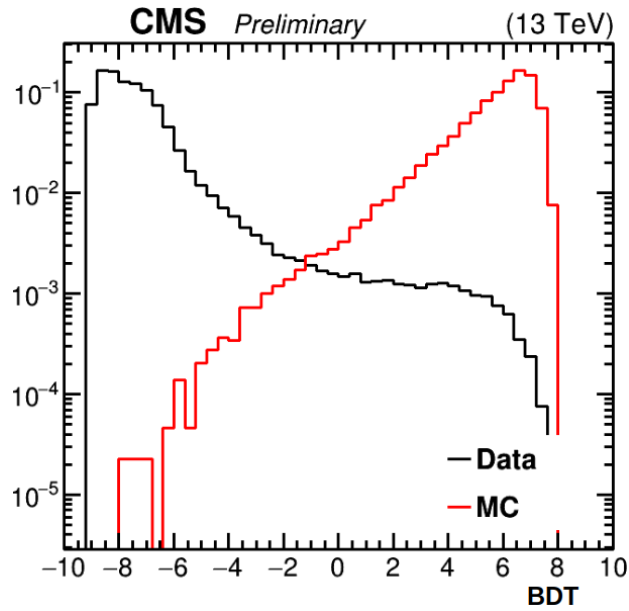


Figure 5.35: BDT output for signal (MC simulation in red) and background (Same-sign data in black).

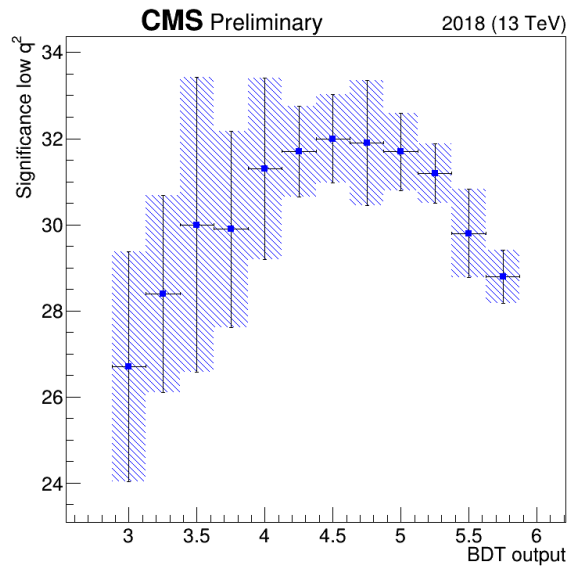


Figure 5.36: Significance as a function of BDT working points for low- q^2 .

5.7.4 Mass fit

The number of signal events in different q^2 ranges is determined through an unbinned maximum likelihood fit performed on the data. The probability density functions (PDF) describing the signal shapes and different backgrounds are fixed using data and MC and are known as templates. During the fit, the normalizations of the templates are allowed to vary to obtain the numbers of signals and backgrounds from the data.

Fit in the low- q^2 bin

For the low- q^2 bin, the shapes use in the triplet invariant mass distribution fit are obtained from simulated events that satisfy the preselection, invariant mass, and BDT cuts. The approach involves utilizing dedicated MC to derive the signal and background component shapes, and fixing certain shape parameters for the final fit to data.

Beginning with the signal shape, a combination of a double-sided Crystal Ball and a Gaussian is employed to describe the distribution. It has been verified that the Gaussian component of the fit is dominant. The mean and width of the double-sided Crystal Ball and Gaussian are allowed to vary independently. Figure 5.37 (a) displays the fit along with its corresponding pull distribution. The pull is defined as the discrepancy between the fit function values and data, divided by the uncertainty.

For each background source, a different MC sample is independently fitted. Starting with the Cabibbo suppressed mode $B^+ \rightarrow \mu^+ \mu^- \pi^+$, the pion is treated as a Kaon since there is no particle identification system to distinguish the two hadrons. Therefore, the structure has a peaking shape with a mean above the B meson mass, as shown in Fig. 5.37 (b). In

the low q^2 bin, a double-sided crystal ball is used to parameterize the Cabibbo suppressed mode.

The $\mu\mu K^*$ backgrounds generate two kinds of triplets: those created with the π and those with the K. The former ones, due to the wrong mass hypothesis (pion mass is assigned with Kaon mass) and the mass window applied on $m(\mu, \mu, K)$, follow an exponential curve. On the other hand the latter type of triplets create a bump-like structure. Therefore, the former is fitted with an exponential as shown in Fig. 5.37 (c), and the latter with a double-sided Crystal Ball as in Fig. 5.37 (d). As explained in the section of background analysis, this background covers three partially reconstructed decays: $B^+ \rightarrow \mu\mu K^{*+}$, followed by the decays of K^{*+} , $K^{*+} \rightarrow K^+\pi^0$ and $K^{*+} \rightarrow K^0\pi^+$, and $B^0 \rightarrow \mu\mu K^{*0} \rightarrow \mu\mu K^+\pi^-$. Since all the shapes and $\alpha \times \varepsilon$ numbers are the same, with the only difference being the relative BF, a common shape is used in the final fit for all those background sources. Furthermore, the relative yield of the pion's component is fixed with respect to the Kaon's component, taking into account the corresponding branching fractions, the isospin decomposition coefficients for the K^* mesons decays and the efficiencies.

The shape of the combinatorial background is modelled by using triplets with same charge muons. Figure 5.37 (e) shows that the shape is described well by using an exponential plus a constant.

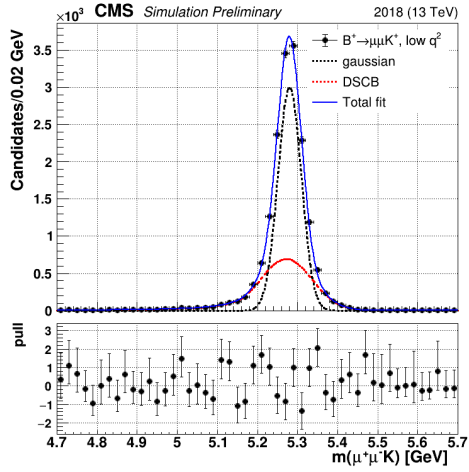
The last background source is the double semi-leptonic B decays that consists of partially reconstructed B decays involving c-mesons, and are inclusively treated in a single shape named "other B ". For this background a dedicated simulation of generic B decays is used. However the BDT is very efficient in cutting this type of background, thus only few candidates are passing the full selection. This is shown in Fig. 5.37 (f). Therefore, this background

source is parameterized empirically using a free exponential which is fitted directly to the data.

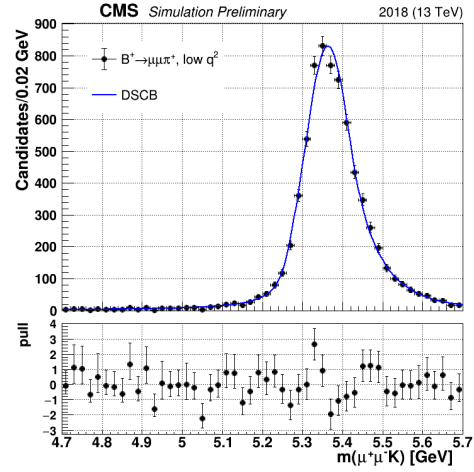
The inclusive fit to the data and its pull are shown in Fig. 5.38. For this fit all shapes are fixed, as discussed above, except the mean of the signal Gaussian and the only parameter that is left free to vary is the mean of the signal Gaussian. The yield of each background component of the fit is summarized in Table 5.13. Table 5.14 summarizes the fit for the inclusive low- q^2 bin.

Table 5.13: Blinded yields for signal and background sources in the dimuon low- q^2 bin.

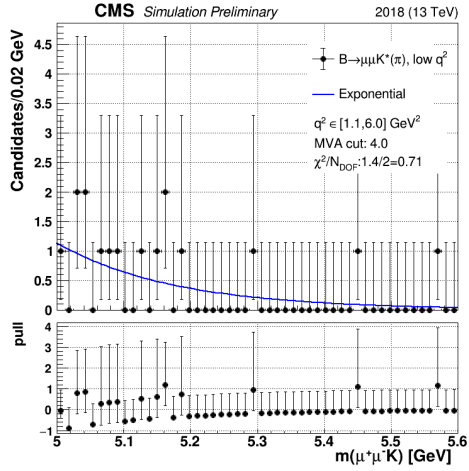
Source	Yield
$B^\pm \rightarrow \mu^+ \mu^- K^\pm$ (Signal)	1298 ± 97
$B^\pm \rightarrow \mu^+ \mu^- \pi^\pm$	21.26 ± 0.80
$B^{0/\pm} \rightarrow \mu^+ \mu^- K^{*(0/\pm)}$	292.32 ± 60.61
Combinatorial	732.15 ± 110.49
Other B	620.91 ± 169.07



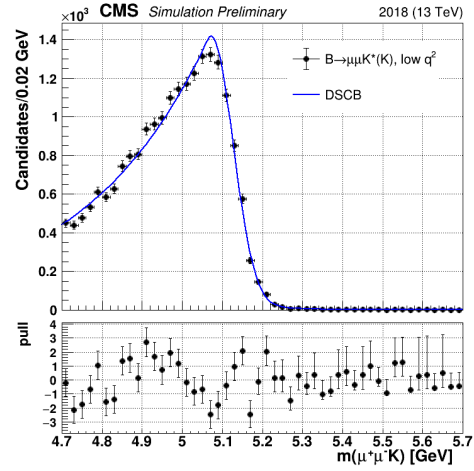
(a)



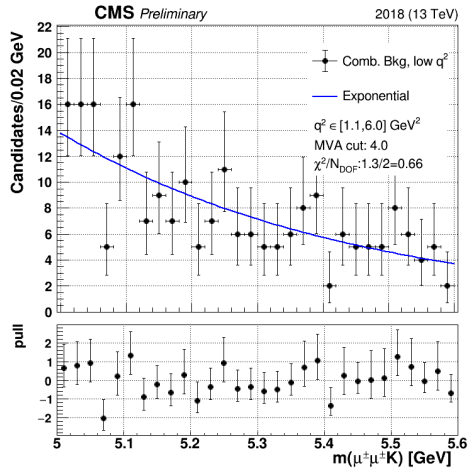
(b)



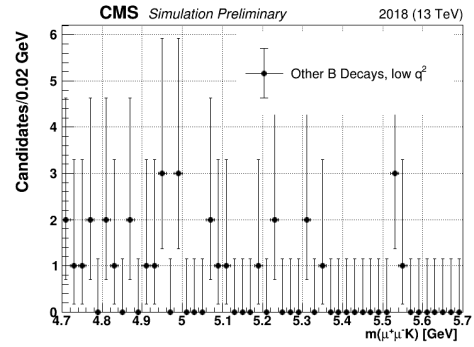
(c)



(d)



(e)



(f)

Figure 5.37: Triplet invariant mass models in the low q^2 bin for the signal $B^\pm \rightarrow K^\pm \mu^+ \mu^-$ (a), $B^\pm \rightarrow \pi^\pm \mu^+ \mu^-$ (b), $B \rightarrow K^*(K\pi) \mu^+ \mu^-$ where the π is mistaken for the K (c), $B \rightarrow K^* \mu^+ \mu^-$ where the K is used in the triplet (d), combinatorial background (e), other b decays (f).

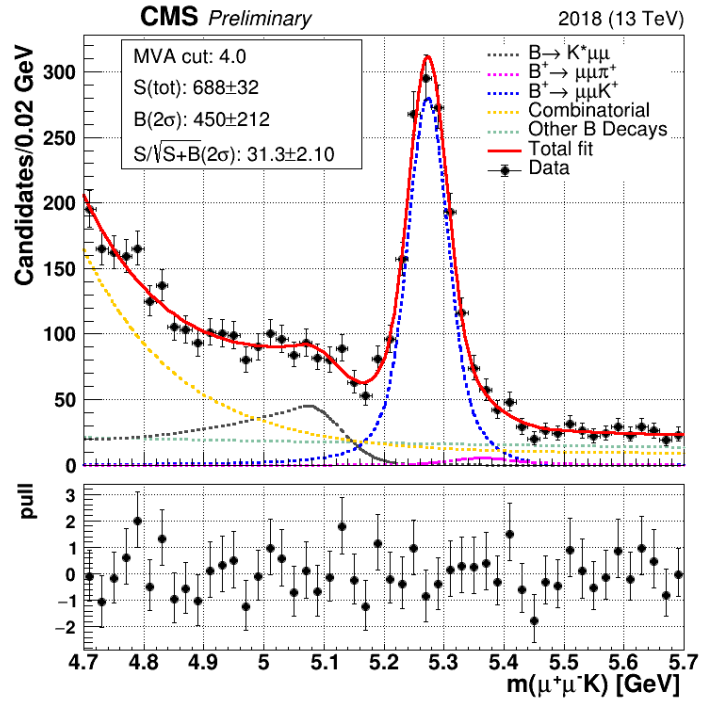


Figure 5.38: Invariant mass distribution for the B candidates in the muon channel in the low- q^2 bin. The fit is overlaid to the data. The bottom panel shows the pull distribution.

Table 5.14: Summary of the fit for the inclusive low- q^2 for the $B^\pm \rightarrow \mu\mu K^\pm$ mode.

Process	Model	Shape	Normalization
Combinatorial Background	Exponential + Zero order Polynomial	Slope for exponential and relative amplitude: Fixed from data with same-sign muons.	Floating
Other B Background $B^\pm \rightarrow \mu\mu\pi^\pm$	Exponential Double-sided CB	Slope for exponential: Floating Fixed from MC	Floating Fixed with respect to $B^\pm \rightarrow \mu\mu K^\pm$ yield using the relative ratio of $\alpha \times \varepsilon$ and relative branching ratios
$B^{0/(\pm)} \rightarrow \mu\mu K^{*(0/\pm)}(\pi^\pm)$	Exponential	Slope for exponential: Fixed from MC	Fixed wrt to $B^{0/(\pm)} \rightarrow \mu\mu K^{*(0/\pm)}(K^\pm)$ in order to include contribution from both neutral and charged B/K mesons.
$B^{0/(\pm)} \rightarrow \mu\mu K^{*(0/\pm)}(K^\pm)$ $B^\pm \rightarrow \mu\mu K^\pm$	Double-sided Crystal Ball Double-sided Crystal Ball + Gaussian	Fixed from MC DSCB: Fixed from MC Gaussian width: Fixed from MC Gaussian Mean: Free Relative yield: Fixed from MC	Floating Floating

Fit in the J/ψ bin

The fits in the J/ψ bin are done with the same procedure as described for the low- q^2 bin.

The signal in the J/ψ bin is modelled by the combination of three Gaussians, as illustrated in Fig. 5.39 (a). The function is effective up to 5.5 GeV, beyond which statistical fluctuations dominate the distribution. However, we examined the measured yield obtained from the fit and compared it with the total number of signal events. The discrepancy between the two is insignificant and is entirely covered by uncertainty of the fit.

A double-sided crystal ball is used to parameterize the Cabibbo suppressed mode $B^+ \rightarrow J/\psi\pi^+$, as depicted in Fig. 5.39 (b). Fig. 5.39 (c) illustrates that the π -component of the $B \rightarrow J/\psi K^*$ background is modeled by a single exponential. On the other hand, the K-component is described by a double-sided crystal ball, and the fit is limited to 5.35 GeV, as shown in Fig. 5.39 (d). The combinatorial background is described by an exponential plus a constant as used in the low q^2 bin, as shown in Fig. 5.39 (e).

In the J/ψ region, the semi-muonic B decays are not particularly significant, but the decays of $B \rightarrow J/\psi X$ are crucial. A simple exponential function is suitable for describing the shape of this region, as demonstrated in Fig. 5.39 (e). Given that both the this decay mode and the combinatorial background are characterized by an exponential function, we decided to combine them into a single exponential function.

Figure 5.40 shows the final fit to the data and its associated pull. The estimated yield for the signal and background components obtained from the fit is provided in Table 5.16. The quality of the fit has been evaluated across several BDT working points, and the results remain consistent. Notably, the fit yields are also consistent between binned and unbinned

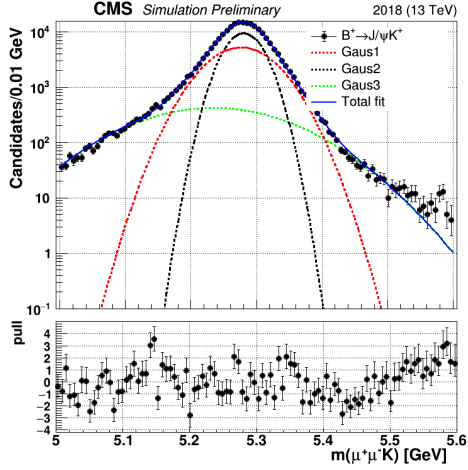
analyses, with only minor differences in the p-value. The summary of the J/ψ bin fit can be found in Table 5.15.

Table 5.15: Summary of the fit for the J/ψ bin for the $B^\pm \rightarrow J/\psi K^\pm \rightarrow \mu\mu K^\pm$ mode.

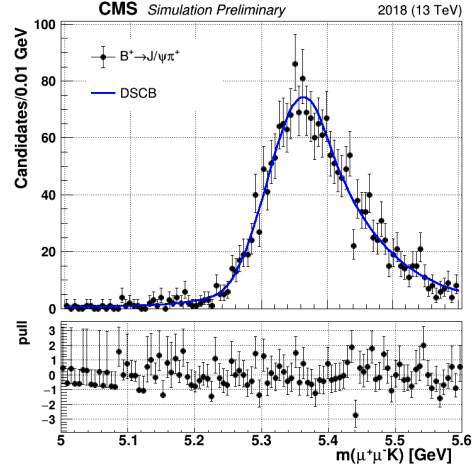
Process	Model	Shape	Normalization
Combinatorial + Other B Back- ground $B^\pm \rightarrow J/\psi \pi^\pm$	Exponential Double-sided CB	Slope for exponential: Floating Fixed from MC, mean shifted to the left from MC by 5 MeV.	Floating Fixed with respect to $B^\pm \rightarrow J/\psi K^\pm$ yield using the relative ratio of $\alpha \times \varepsilon$ and relative branching ratios
$B^{0/(\pm)} \rightarrow J/\psi K^{*(0/\pm)}(\pi^\pm)$	Exponential	Fixed from MC	Fixed wrt to $B^{0/(\pm)} \rightarrow J/\psi K^{*(0/\pm)}(K^\pm)$ in order to include contribution from both neutral and charged B/K mesons.
$B^{0/(\pm)} \rightarrow J/\psi K^{*(0/\pm)}(K^\pm)$ $B^\pm \rightarrow J/\psi K^\pm$	Double-sided Crystal Ball 3 Gaussians	Fixed from MC Everything is being constrained from MC	Floating Floating

Table 5.16: Yields for signal and background sources in the J/ψ bin.

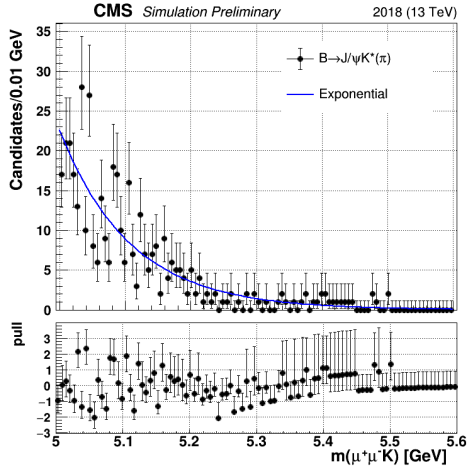
Source	Yield
$B^\pm \rightarrow J/\psi K^\pm$ (Signal)	727984 ± 1113
$B^\pm \rightarrow J/\psi \pi^\pm$	27124 ± 39
$B^{0/\pm} \rightarrow J/\psi K^{*(0/\pm)}$	124588 ± 853
Combinatorial/Other B	61134 ± 1099



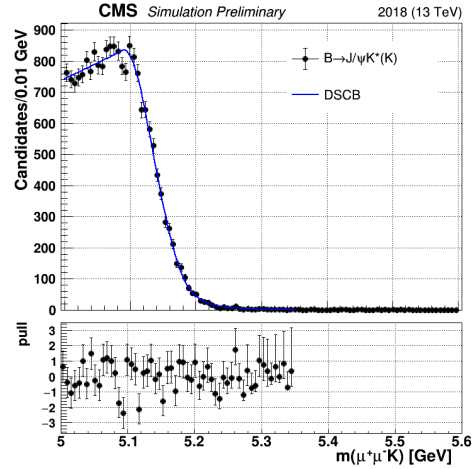
(a)



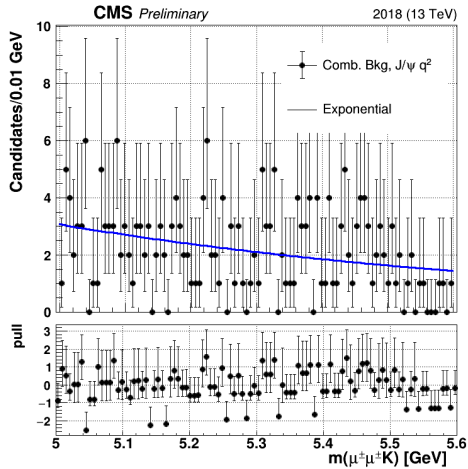
(b)



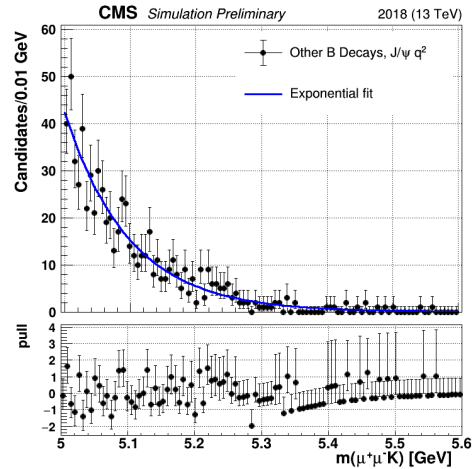
(c)



(d)



(e)



(f)

Figure 5.39: Triplet invariant mass models in the J/ψ bin for the signal $B^\pm \rightarrow K^\pm J/\psi(\mu^+\mu^-)$ (a), $B^\pm \rightarrow \pi^\pm J/\psi(\mu^+\mu^-)$ (b), $B \rightarrow K^*(K\pi)J/\psi(\mu^+\mu^-)$ where the π is mistaken for the K (c), $B \rightarrow K^*J/\psi(\mu^+\mu^-)$ where the K is used in the triplet (d), combinatorial background (e), other b decays (f).

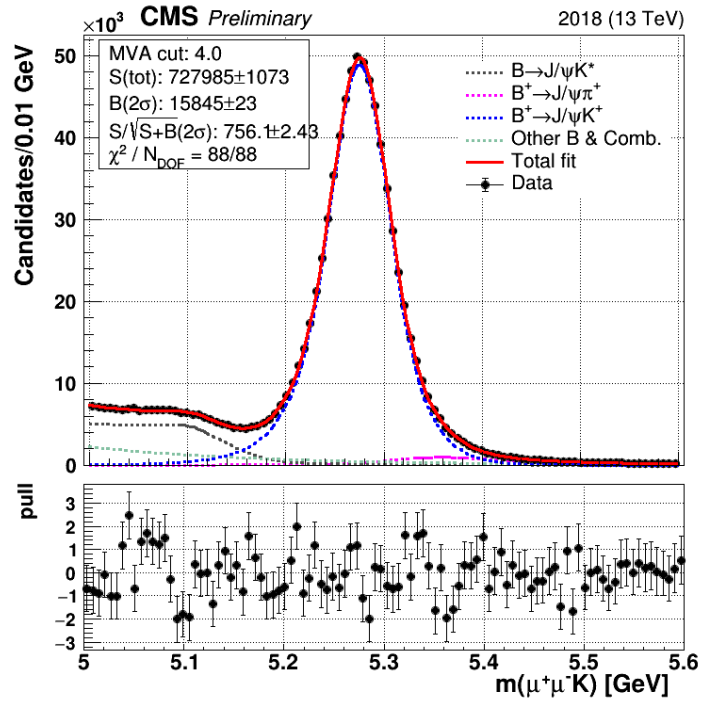
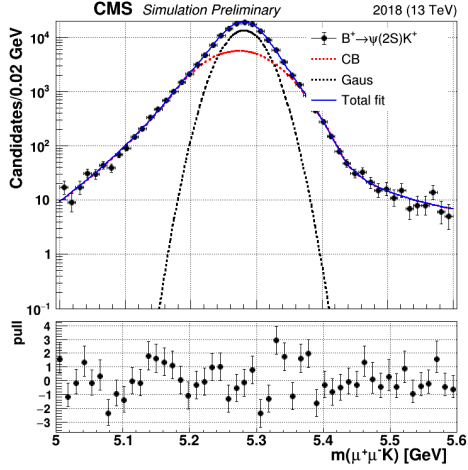


Figure 5.40: Invariant mass distribution for the B candidates in the muon channel in the J/ψ bin. The fit is overlaid to the data. The bottom panel shows the pull distribution.

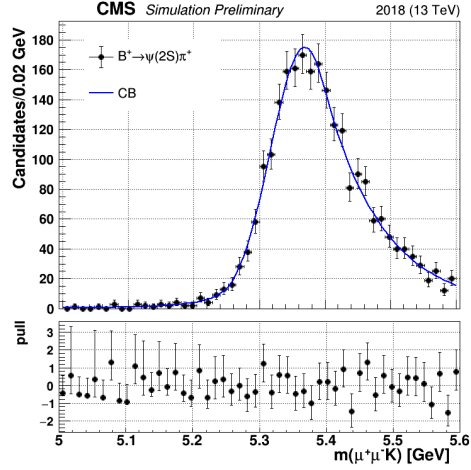
Fit in the $\psi(2S)$ bin

The fitting in the $\psi(2S)$ bin is similar to that in the J/ψ bin, except for the signal fit. The signal fit in the $\psi(2S)$ bin employs a double-sided crystal ball combined with a single Gaussian, while the background functions are identical to those in the J/ψ bin. The fit in the $\psi(2S)$ bin is used as an sanity check for the R_K measurement.

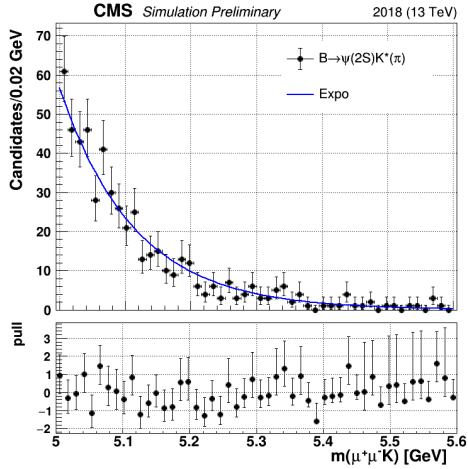
Figure 5.41 shows the template fits performed on the MC. Figure 5.42 shows the final fit performed on the data, along with its corresponding pull. The quantities estimated from the fitting, namely the signal and background yields, are listed in Table 5.18. Table 5.17 provides a summary of the fits conducted in the $\psi(2S)$ bin.



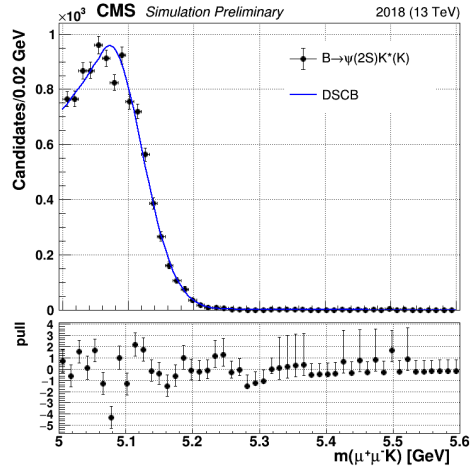
(a)



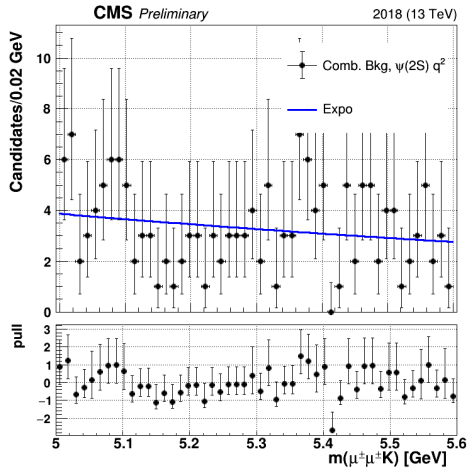
(b)



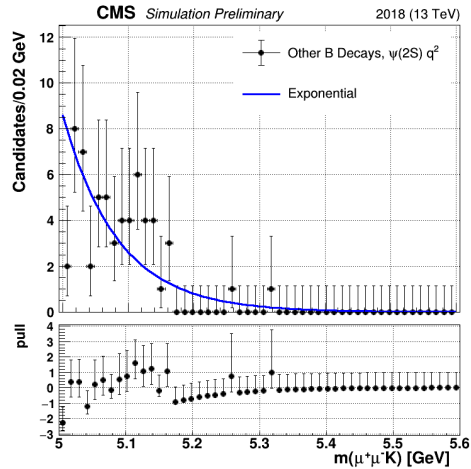
(c)



(d)



(e)



(f)

Figure 5.41: Triplet invariant mass models in the $\psi(2S)$ bin for the signal $B^\pm \rightarrow K^\pm \psi(2S)(\mu^+ \mu^-)$ (a), $B^\pm \rightarrow \pi^\pm \psi(2S)(\mu^+ \mu^-)$ (b), $B \rightarrow K^*(K\pi)\psi(2S)(\mu^+ \mu^-)$ where the π is mistaken for the K (c), $B \rightarrow K^*\psi(2S)(\mu^+ \mu^-)$ where the K is used in the triplet (d), combinatorial background (e), other b decays (f).

Table 5.17: Summary of the fit for the $\psi(2S)$ bin for the $B^\pm \rightarrow \psi(2S)K^\pm \rightarrow \mu\mu K^\pm$ mode.

Process	Model	Shape	Normalization
Combinatorial + Other B Back- ground $B^\pm \rightarrow \psi(2S)\pi^\pm$	Exponential Double-sided CB	Slope for exponen- tial: Floating Fixed from MC	Floating Fixed with respect to $B^\pm \rightarrow \psi(2S)K^\pm$ yield using the relative ratio of $\alpha \times \varepsilon$ and relative branching ratios
$B^{0/(\pm)} \rightarrow$ $\psi(2S)K^{*(0/\pm)}(\pi^\pm)$	Exponential	Fixed from MC	Fixed wrt to $B^{0/(\pm)} \rightarrow$ $\psi(2S)K^{*(0/\pm)}(K^\pm)$ in order to in- clude contribution from both neutral and charged B/K mesons.
$B^{0/(\pm)} \rightarrow$ $\psi(2S)K^{*(0/\pm)}(K^\pm)$ $B^\pm \rightarrow \psi(2S)K^\pm$	Double-sided Crys- tal Ball Double-sided Crys- tal Ball + Gaussian	Fixed from MC DSCB: Power Law and endpoints of power laws fixed from MC. Means, widths and relative yield: Constrained from MC	Floating Floating

Table 5.18: Yields for signal and background sources for the muon channel in the $\psi(2S)$ bin.

Source	Yield
$B^\pm \rightarrow \psi(2S)K^\pm$ (Signal)	68301 ± 315
$B^\pm \rightarrow \psi(2S)\pi^\pm$	2331 ± 11
$B^{0/\pm} \rightarrow \psi(2S)K^{*(0/\pm)}$	9936 ± 412
Combinatorial/Other B	5919 ± 539

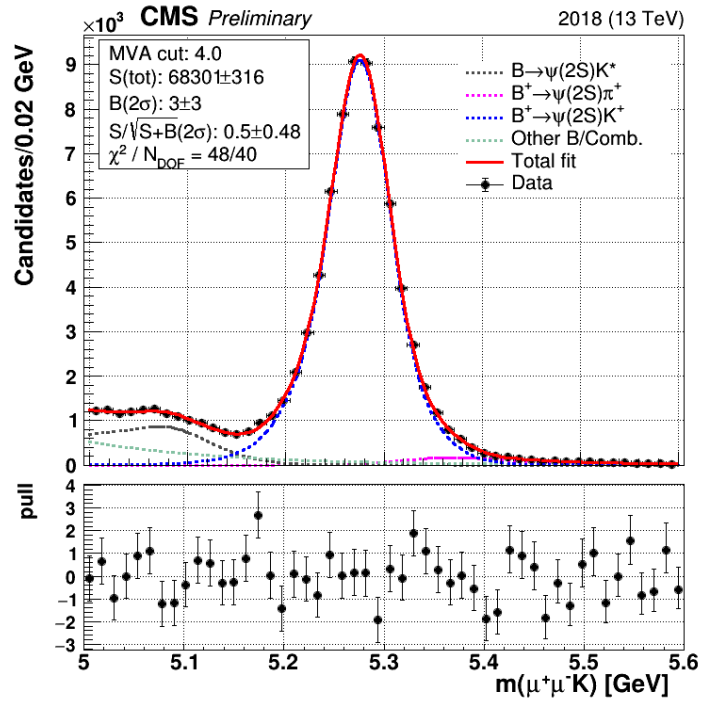


Figure 5.42: Invariant mass distribution for the B candidates in the muon channel in the $\psi(2S)$ bin. The fit is overlaid to the data. The bottom panel shows the pull distribution.

Fit in the high q^2 bin

The fitting procedure in the high- q^2 bin is similar to that in other q^2 bins but with some changes as described below.

Figure 5.43 (a) shows the template fit for the signal $B^\pm \rightarrow \mu\mu K^\pm$ in the high q^2 bin, which is modeled through the use of a double-sided crystal ball and a Gaussian, as in the other bins, to maintain consistency.

Figure 5.43 (b) shows the template fit of the Cabibbo suppressed mode $B^\pm \rightarrow \mu^+\mu^-\pi^\pm$, which is handled similarly in the high q^2 bin as in the low q^2 bin. Figure 5.43 (c) shows that the π -component of the K^* background in the high q^2 bin fitted with a double-sided crystal ball, as the structure has shifted to the right, following the pattern of the $\psi(2S)$ bin, due to the different kinematics. Moreover, the K -component of the K^* background in the high q^2 bin is now fitted with a more complex function, using both a double-sided and a single left-sided crystal ball, as demonstrated in Fig. 5.43 (d). A third-order polynomial is used to describe the shape of the combinatorial background in this bin, as presented in Fig. 5.43 (e).

The decays of $b \rightarrow c\mu\nu \rightarrow s\mu\mu\nu\nu$ are not expected to contribute significantly due to the masses of the B and D mesons and the dimuon mass window. In addition, as shown in Fig. 5.43, the generic dimuon sample that has proven to be efficient in other decays is not as useful for this q^2 bin. If there is any possible contamination from double semi-muonic B decays in the high q^2 bin, it will be absorbed in the fit of an exponential.

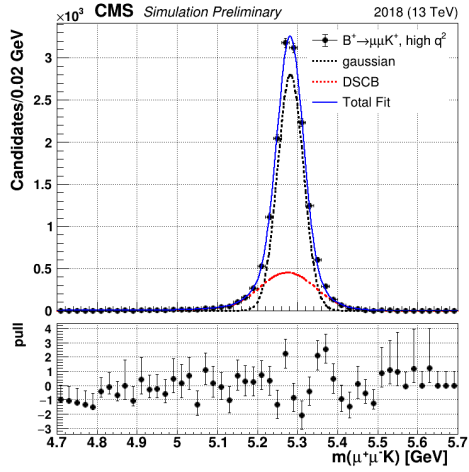
Figure 5.44 shows the final fit of the data along with its pull. Table 5.19 provides a summary of the yields for signal of each background component of the fit.

Table 5.19: Blinded yields for signal and background sources in the high q^2 bin.

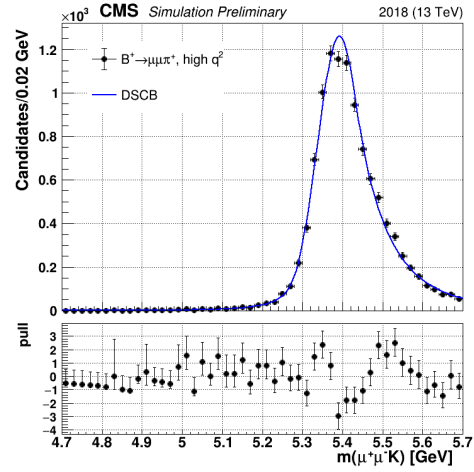
Source	Yield
$B^\pm \rightarrow \mu^+ \mu^- K^\pm$ (Signal)	1001 ± 43
$B^\pm \rightarrow \mu^+ \mu^- \pi^\pm$	19.78 ± 0.73
$B^{0/\pm} \rightarrow \mu^+ \mu^- K^{*(0/\pm)}$	298.10 ± 16.28
Combinatorial	304.21 ± 21.35
Other B	0.0003 ± 24.73

Table 5.20: Summary of the fit for the inclusive high- q^2 for the $B^\pm \rightarrow \mu\mu K^\pm$ mode.

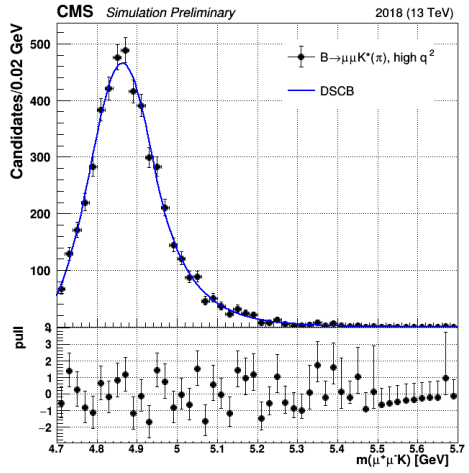
Process	Model	Shape	Normalization
Combinatorial Background	Third order Polynomial	Fixed from data with same-sign muons.	Floating
Other B Background $B^\pm \rightarrow \mu\mu\pi^\pm$	Exponential Double-sided CB	Slope for exponential: Floating Fixed from MC	Floating Fixed with respect to $B^\pm \rightarrow \mu\mu K^\pm$ yield using the relative ratio of $\alpha \times \varepsilon$ and relative branching ratios
$B^{0/(\pm)} \rightarrow \mu\mu K^{*(0/\pm)}(\pi^\pm)$	Double-Sided Crystal Ball	Fixed from MC	Fixed wrt to $B^{0/(\pm)} \rightarrow \mu\mu K^{*(0/\pm)}(K^\pm)$ in order to include contribution from both neutral and charged B/K mesons.
$B^{0/(\pm)} \rightarrow \mu\mu K^{*(0/\pm)}(K^\pm)$	Double-sided Crystal Ball + Single-Sided Crystal Ball	Fixed from MC	Floating
$B^\pm \rightarrow \mu\mu K^\pm$	Double-sided Crystal Ball + Gaussian	DSCB: Fixed from MC Gaussian width: Fixed from MC Gaussian Mean: Free Relative yield: Fixed from MC	Floating



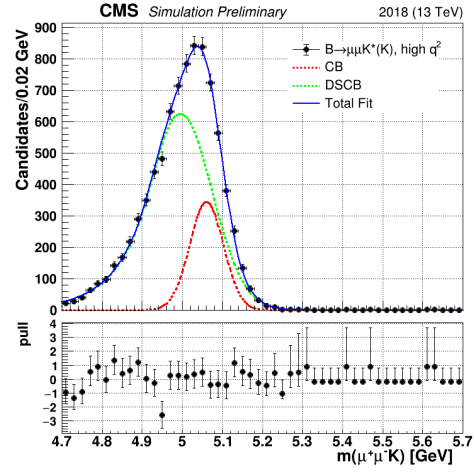
(a)



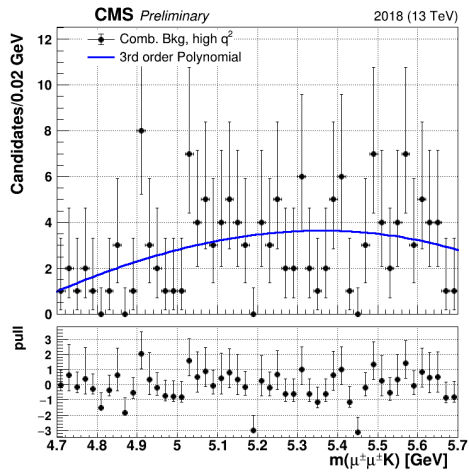
(b)



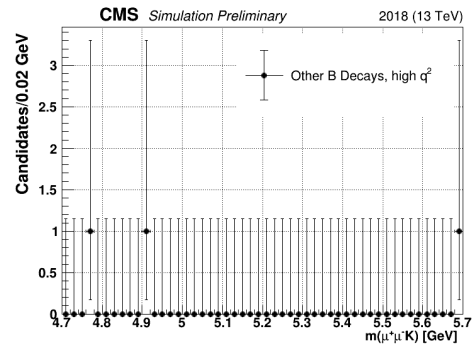
(c)



(d)



(e)



(f)

Figure 5.43: Triplet invariant mass models in the high q^2 bin for the signal $B^\pm \rightarrow K^\pm \mu^+ \mu^-$ (a), $B^\pm \rightarrow \pi^\pm \mu^+ \mu^-$ (b), $B \rightarrow K^*(K\pi) \mu^+ \mu^-$ where the π is mistaken for the K (c), $B \rightarrow K^* \mu^+ \mu^-$ where the K is used in the triplet (d), combinatorial background (e), other b decays (f).

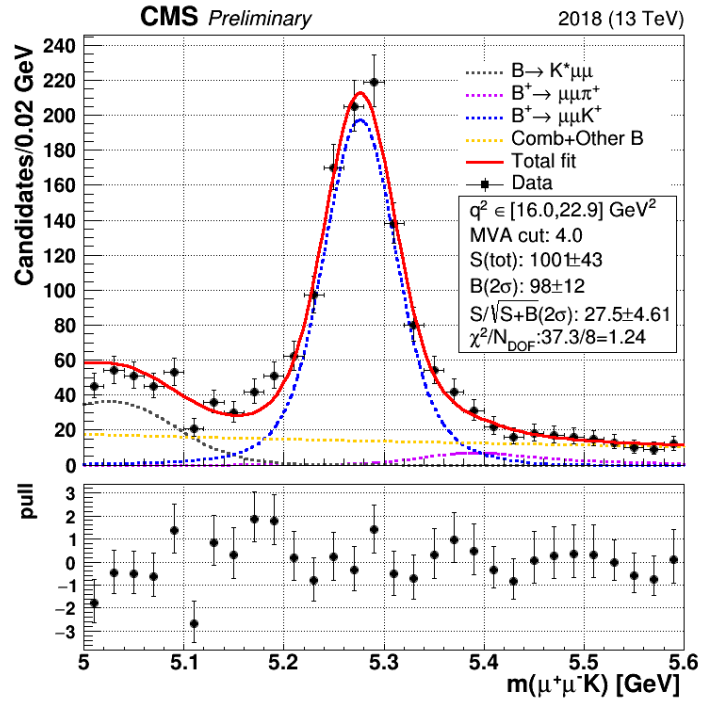


Figure 5.44: Invariant mass distribution for the B candidates in the muon channel in the high- q^2 bin. The fit is overlaid to the data. The bottom panel shows the pull distribution.

5.7.5 Acceptance and efficiency

The calculation of acceptance and efficiency is essential for the R_K measurement. Acceptance refers to the fraction of the particles passed from the generator because of the acceptance of the detector, such as in p_T and η , while efficiency refers to the selection criteria applied to reconstructed objects in order to distinguish signal from background. Both acceptance and efficiency can be measured independently and can be factorized. To minimize statistical uncertainty and reduce the number of generated events, we use different MC samples, each optimized for a specific factor of $\alpha \times \epsilon$. These samples are designed to achieve the best possible statistical uncertainty.

For each B decay, we have produced two dedicated MC samples:

- Unfiltered MC: It is used to measure the efficiency of the generator filter. This sample is produced without a generator filter, so the generated particles can be produced outside of CMS geometrical boundaries or with a very soft momentum. Therefore, this sample is more computational costly in producing it.
- Filtered MC: The production of this sample involves applying cuts at the generator level to filter out certain events. Specifically, the generator cuts require that one of the muons must be within the CMS acceptance range and have a transverse momentum of greater than 6 GeV. These cuts were selected to ensure that at least one of the muons can trigger and be used to measure the efficiency factor ($\alpha \times \epsilon$) of all other factors and the final selection, based on the filtered muon.

The acceptance and efficiency are factorized in detail as follows:

$$\alpha \times \epsilon = \alpha(\text{filter}) \times \alpha(\text{trg}\mu) \times \epsilon(\text{trg}\mu) \times \alpha(\mu 2, K) \times \epsilon(\mu 2, K) \times \epsilon(\text{preselection}) \times \epsilon(m(\mu, \mu)) \times \epsilon(\text{BDT}), \quad (5.5)$$

where $\alpha(\text{filter})$ is the acceptance of the filtered muon, $\alpha(\text{trg}\mu)$ is the acceptance of the triggering muon, $\epsilon(\text{trg}\mu)$ is the reconstruction efficiency of the triggering muon, $\alpha(\mu 2, K)$ and $\epsilon(\mu 2, K)$ are the acceptance and reconstruction efficiencies of the other two objects, $\epsilon(\text{preselection})$ the effect of preselection cuts, $\epsilon(\text{mass})$ the efficiency of the dimuon invariant mass window selection, and $\epsilon(\text{BDT})$ the BDT efficiency.

Corrections on simulated events in the muon channel

To improve the agreement between the simulated events and the data, Scale Factors (SF) are applied as corrections. First, a trigger SF is applied on an event-by-event basis using tag-and-probe, which is privately measured (refer to section 5.7.1), taking into account that both muons can trigger the event. The muon identification SF is measured centrally by the MUO POG and applied to all events. Our studies revealed that the B $p_T(B)$ in simulation does not provide an accurate description of the observed $p_T(B)$. Hence, the Fixed Order + Next-to-Leading Logarithm (FONLL) SF is applied to the simulated events. To compute the specific decays, we will compare the totally unbiased Soft QCD simulated $p_T(B)$ [GeV] and rapidity(B) spectrum with the one produced by the FONLL calculator. To derive the unbiased spectrum by the FONLL calculator, no decay options have been applied except for the B hadron production (instead of bare quarks) and the collision energy of the LHC. Figure 5.45 (b) shows that there is no clear trend in the rapidity $y(B)$ distribution between

the generated spectrum by Soft QCD using PYTHIA and the derived spectrum by the FONLL calculator. The only deviations between the two spectra are observed in the case of transverse momentum $p_T(B)$ spectrum, as shown in Fig. 5.45 (a). As a result, all the produced MC samples will be reweighted according to the ratio plot of the transverse momentum spectrum.

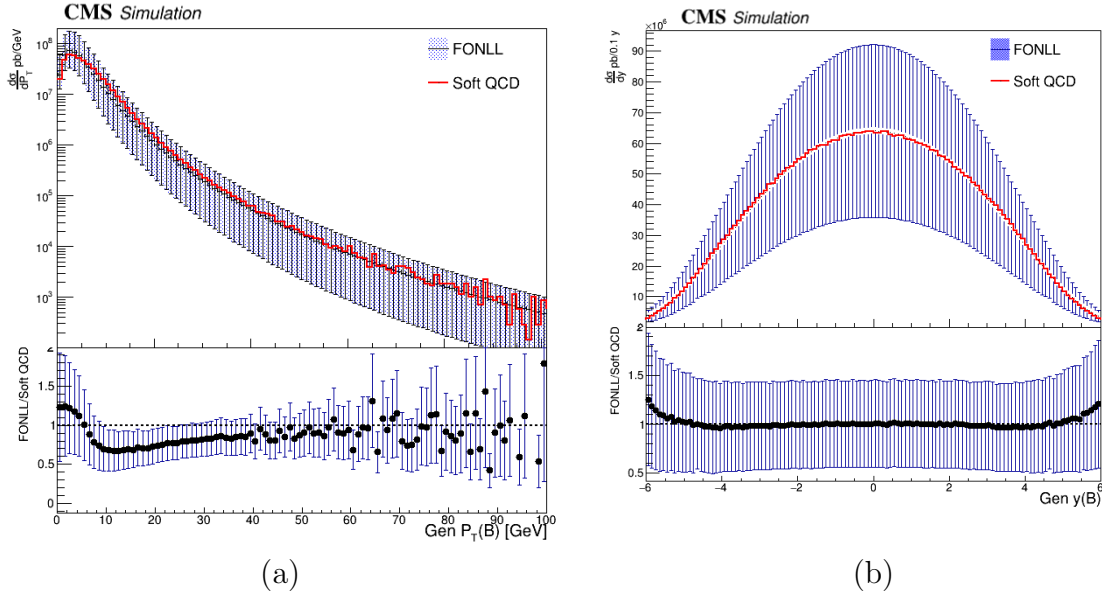


Figure 5.45: (a) Transverse momentum of the B meson; (b) B meson rapidity. The blue points represent the FONLL, and the red points the soft QCD. Bottom panels show the ratio of FONLL and Soft QCD distributions.

Although the correct PU profile was used for the MC samples, the trigger prescale in data depends on PU, which is not present in the MC. Therefore, the events that pass our trigger have a slightly different PU profile in data, as shown in Fig. 5.46. To minimize this effect, all MC events are reweighted to account for the PU profile in data.

The final SF we applied to MC is the BDT SF. The BDT SF for muons is determined in data by utilizing the $B \rightarrow J/\psi K$ control region and repeating the fit for multiple BDT cuts. The efficiency for the working point (WP) X is defined as the signal yield in that WP

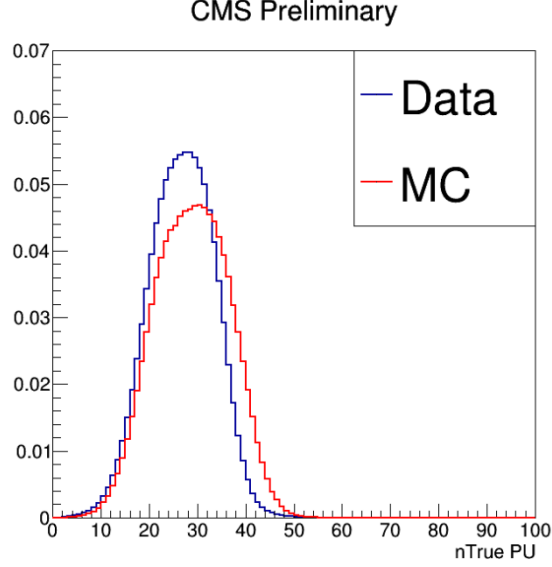


Figure 5.46: PU profile in MC and data.

divided by the yield for a very loose WP (in our case BDT_{i0}). The selection of a very loose WP is made because it has a negligible amount of signal reduction. The same definition of efficiency is used in simulation, and the efficiencies are then divided to obtain the SF. To avoid double-counting of effects, all previous SF are already applied to the MC.

The dimuon invariant mass distribution from the theoretical calculation of Pythia has a discrepancy from the data. Figure 5.48 shows the dimuon mass distributions between data and MC. The theoretical dimuon distributions from FLAVIO [64] and Pythia are shown in 5.47. This can affect the efficiency measurement of the dimuon window. In order to make our measurement independent of the theory, the acceptance and efficiency calculations in the low- q^2 region are measured on top of the generated q^2 falling in the low- q^2 mass window, i.e. requiring $1.1 < q^2 < 6.0 \text{ GeV}^2$ in the denominator. This changes the definition of R_K to be measuring only on low- q^2 , instead of the full q^2 range. This eliminates the need to calculate the efficiency of the q^2 window (ϵ_{q^2}), making this measurement independent of differential branching fraction from different theories.

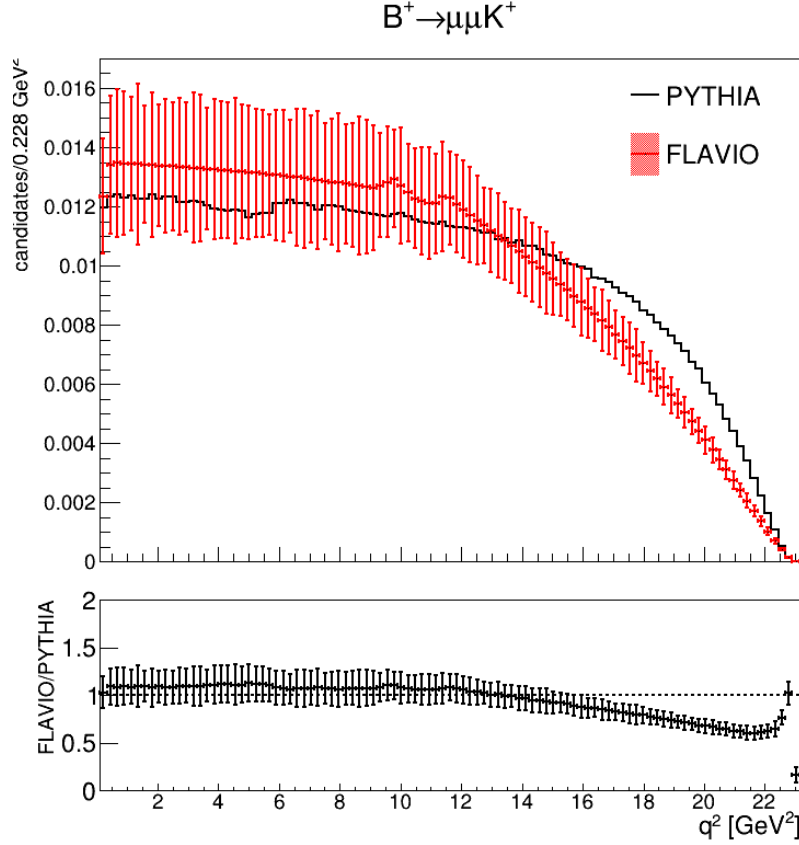


Figure 5.47: Comparison of generated $\frac{dB}{dq^2}$ distribution by PYTHIA with the one predicted by the FLAVIO package for $B^+ \rightarrow K^+ \mu^+ \mu^-$. The ratio points represent the FLAVIO SFs.

The finalized corrected acceptance and efficiency is shown in the Table 5.21.

Table 5.21: Acceptance and efficiency factorized and inclusive for all final states. Uncertainties are only statistical.

Efficiency	$B \rightarrow \mu\mu K$ (low q^2)	$B \rightarrow J/\psi K$	$B \rightarrow \psi(2S)K$
$\alpha(\text{filter})$	4.291×10^{-2}	6.171×10^{-2}	7.757×10^{-2}
$\alpha \times \epsilon(\text{trg})$	6.050×10^{-2}	5.956×10^{-2}	6.000×10^{-2}
$\alpha \times \epsilon(\mu 2, K) \times \epsilon(\text{presel.})$	3.614×10^{-1}	2.741×10^{-1}	4.240×10^{-1}
$\epsilon(\text{BDT})$	8.796×10^{-1}	8.743×10^{-1}	8.710×10^{-1}
$\alpha \times \epsilon$ (total) [$\times 10^{-4}$]	8.253 ± 0.06	9.690 ± 0.03	11.124 ± 0.04

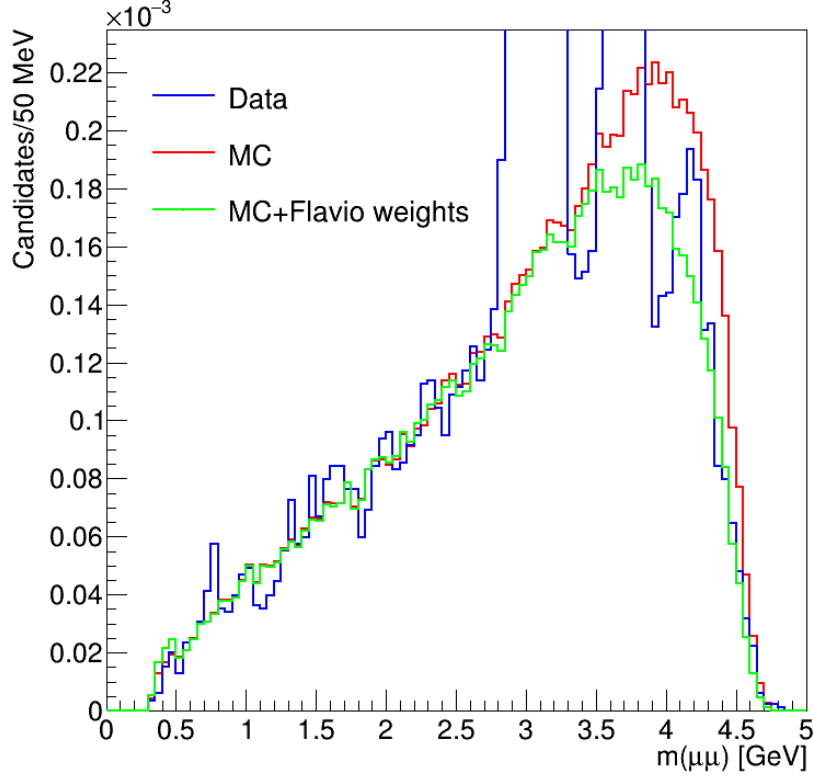


Figure 5.48: $m(\mu\mu)$ distribution from $B^\pm \rightarrow X(\rightarrow \mu\mu)K^\pm$ candidates. In order to determine the dimuon mass distribution from the data we have cut in the analysis BDT working point (BDT output > 4) and around the B mass peak ($5.15 < m(\mu\mu K) < 5.45$ GeV) in order to reduce the contamination from partially reconstructed background that might have a different dimuon mass distribution with respect to our signal. The same cuts have been applied in the MC. In order to determine the red distribution from MC, all the corrections and SFs have been removed, while in order to determine the green distribution from MC only the weights from FLAVIO/PYTHIA comparison have been applied.

5.7.6 Checks for the muon channel

In order to test the validity of the analysis approach used in the muon channel, including the invariant mass fits, efficiency calculations, and the use of the BDT, we performed several cross checks as described below.

Data-MC agreement with *sPlot* in the muon channel

To evaluate the consistency of the BDT output between the data and MC, we used the *sPlot* technique [61]. The discriminating variable was the reconstructed triplet mass, which was fitted using the same functions as described in Section 5.7.4. We obtained *sWeights* from the fit, which enabled us to estimate the likelihood of an event being "signal-like" and to perform a statistical background subtraction. Figure 5.49 compares the signal distribution of data and MC using the *sPlot* technique in different q^2 bins. The results show good data-MC agreement.

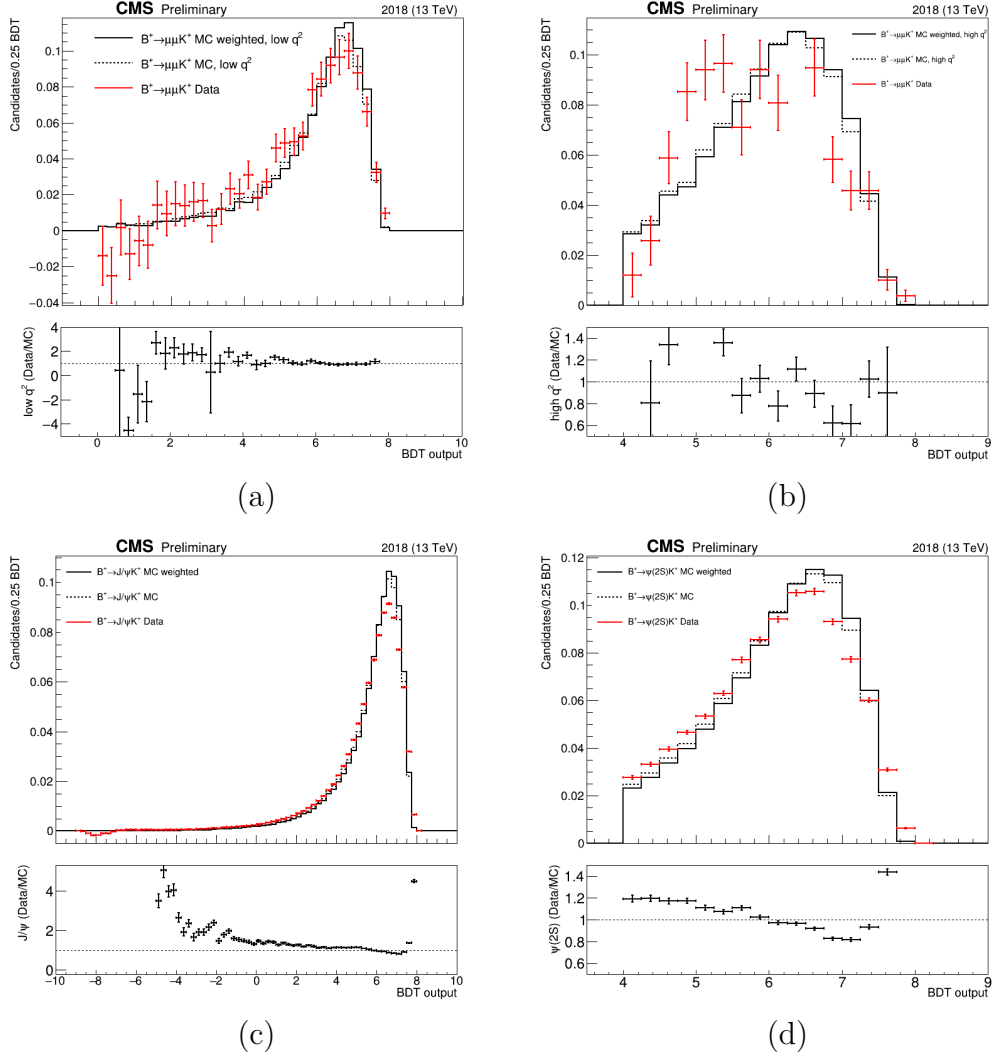


Figure 5.49: BDT output calculated with the *sPlot* technique in the 4 dimuon mass bins: low- q^2 bin (a), high- q^2 bin (b), J/ψ bin (c) and $\psi(2S)$ bin (d).

$R_K^{\psi(2S)}$ measurement in the muon channel

The $\psi(2S)$ bin serves as a sanity check because of its proximity to the signal region and because it is known to agree with the SM with great precision. To perform this cross-check, the ratio of the branching fractions $R_K^{\psi(2S),\mu}$ is determined by measuring the number of $B^+ \rightarrow K^+ \psi(2S)(\rightarrow \mu^+ \mu^-)$ to that of $B^+ \rightarrow K^+ J/\psi(\rightarrow \mu^+ \mu^-)$, and the obtained value is compared with the expected value.

$$R_K^{\psi(2S),\mu} = \frac{N_{K^+ \psi(2S)(\mu^+ \mu^-)}}{N_{K^+ J/\psi(\mu^+ \mu^-)}} \frac{\epsilon_{K^+ J/\psi(\mu^+ \mu^-)}}{\epsilon_{K^+ \psi(2S)(\mu^+ \mu^-)}}, \quad (5.6)$$

where N is the number of signal measured in the data, and ϵ is the efficiency.

The measured value is then compared with the SM prediction calculated from the branching fractions:

$$R_K^{\psi(2S),\mu} = \frac{\mathcal{B}(B^+ \rightarrow K^+ \psi(2S)(\rightarrow \mu^+ \mu^-))}{\mathcal{B}(B^+ \rightarrow K^+ J/\psi(\rightarrow \mu^+ \mu^-))}, \quad (5.7)$$

where the branching fractions are taken from the latest PDG report [18].

The literature [18] reports that the accepted value is $R_K^{\psi(2S),\mu} = 0.0826 \pm 0.0072$. Using our data, we measured the value to be $R_K^{\psi(2S),\mu} = 0.0803 \pm 0.0010$, which is compatible with the accepted value, before considering the systematic uncertainties. This test measurement is an excellent cross-check of the robustness of this analysis.

K^*/K^+ ratio test in the muon channel

The background originating from the partially reconstructed decays of $B^{0/+} \rightarrow K^*(892)^{0/+} \psi(\rightarrow \mu^+ \mu^-)$ makes a considerable contribution to the J/ψ and $\psi(2S)$ regions. The partially reconstructed background yields obtained from the fit can be compared to the expected yields

from the branching fractions and efficiencies to perform a sanity check. This comparison validates the reliability of the mass fit.

The ratio of K^* to K^+ , is defined as

$$r_{K^*/K^+} = \frac{N_{K^{*0}\psi} + N_{K^{*+}\psi}}{N_{K^+\psi}}, \quad (5.8)$$

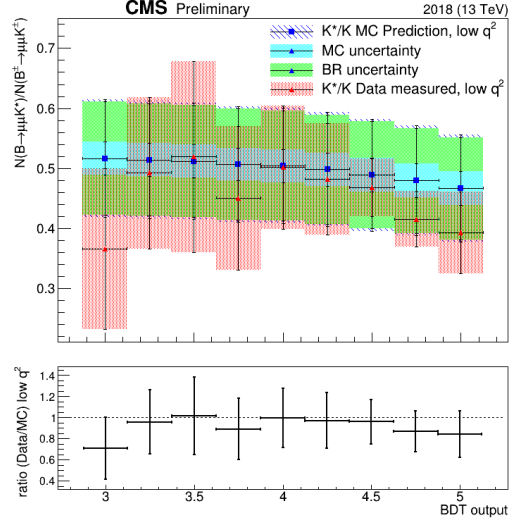
where N_{q^2} is the number of signal or partially reconstructed background measured in the q^2 bin from the mass fit. This ratio is compared to the ratio of branching fractions

$$r_{K^*/K^+} = \frac{\mathcal{B}(B^0 \rightarrow K^{*0}\psi(\rightarrow \mu^+\mu^-)) \cdot \epsilon_{K^{*0}\psi} + \mathcal{B}(B^+ \rightarrow K^{*+}\psi(\rightarrow \mu^+\mu^-)) \cdot \epsilon_{K^{*+}\psi}}{\mathcal{B}(B^+ \rightarrow K^+\psi(\rightarrow \mu^+\mu^-)) \cdot \epsilon_{K^+\psi}}, \quad (5.9)$$

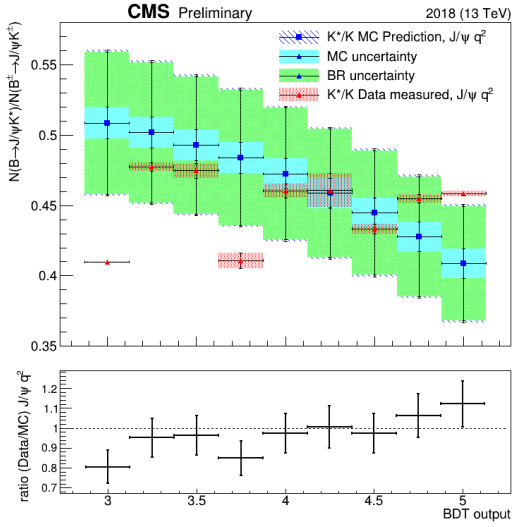
where ϵ is the efficiency calculated using MC.

Figure 5.50 shows the measured r_{K^*/K^+} compared to the expected r_{K^+/K^*} value in different q^2 bins. The measured ratios are in general within 1σ deviation from the predicted values for all BDT working points for all the q^2 bins. The few points that deviate in the low q^2 J/ψ bins can be explained by poor fits. Overall, the predicted yield is consistent with expectations within the uncertainties.

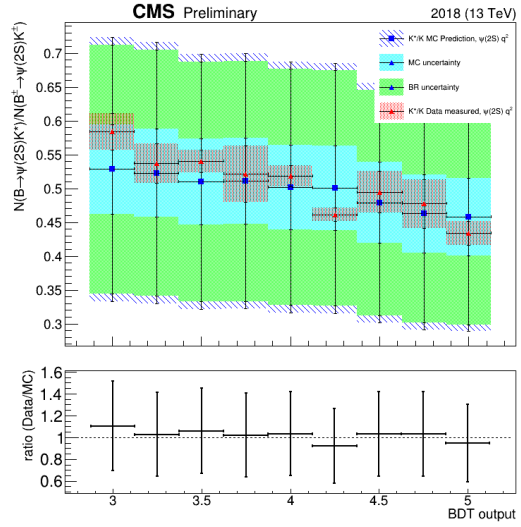
The measured ratio in the high- q^2 bin is higher than the predicted range, which could be due to the higher $c\bar{c}$ resonances that are not included in the fitting process. The K^*/K ratios for the high q^2 ratios are calculated for various windows in the $m(K^+\mu^+\mu^-)$ spectrum. The results for the high- q^2 bin are presented in Fig. 5.51.



(a)



(b)



(c)

Figure 5.50: Comparison of the K^*/K ratios between data and MC for various BDT working points for low- q^2 bin (a), J/ψ bin (b) and $\psi(2S)$ bin (c). The statistical uncertainties from the fit are shown in red, while the systematic branching ratio and statistical uncertainties are shown in green and cyan respectively.

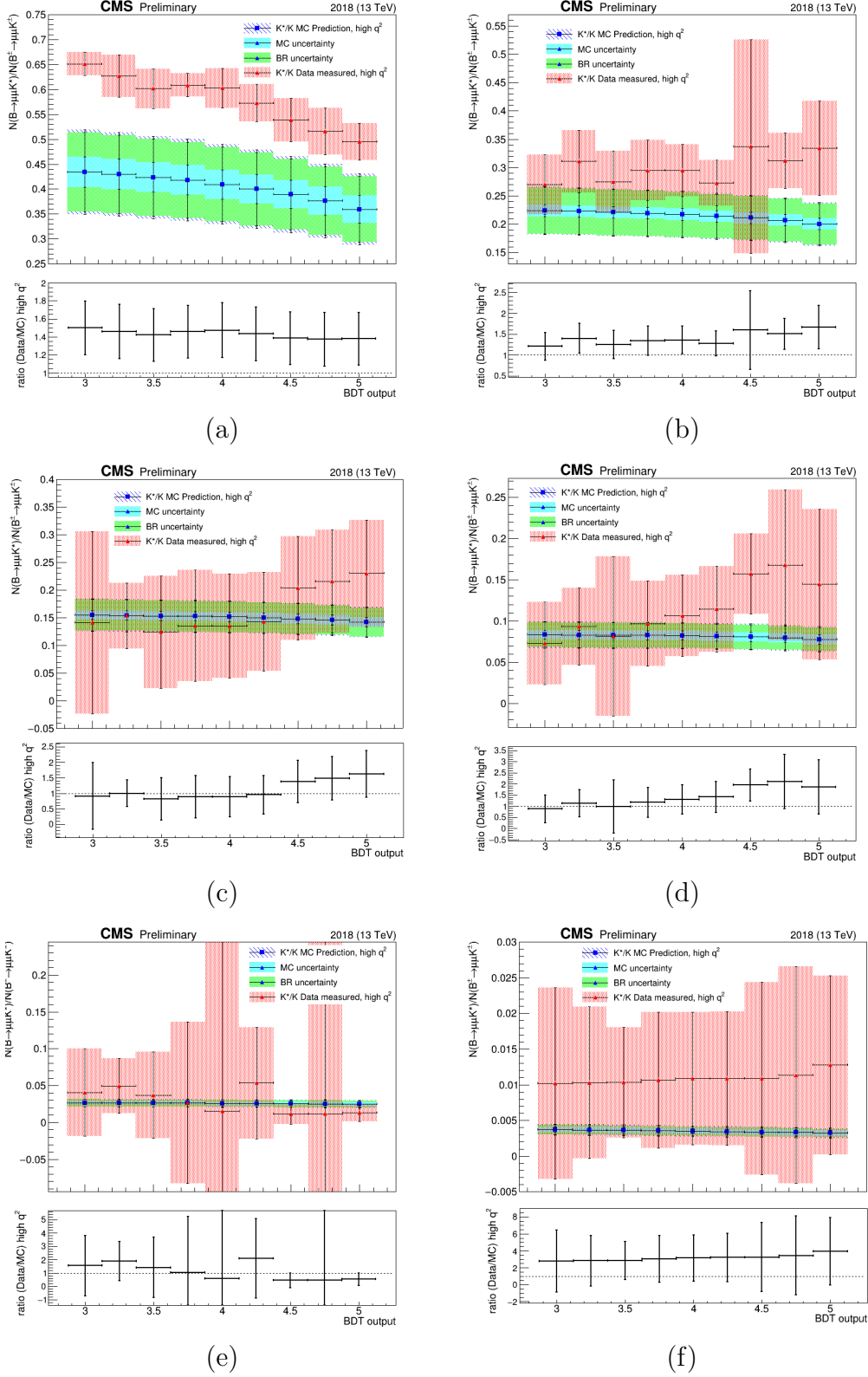


Figure 5.51: Comparison of the K^*/K ratios between data and MC for various BDT working points for the high- q^2 bin. The statistical uncertainties from the fit are shown in red, while the systematic branching ratio and statistical uncertainties are shown in green and cyan respectively. The K^*/K ratio is presented for different ranges of the $m(\mu\mu K)$ spectrum: (a) $m(\mu\mu K) \in (4.7, 5.7)$ GeV, (b) $m(\mu\mu K) \in (4.9, 5.7)$ GeV, (c) $m(\mu\mu K) \in (4.95, 5.7)$ GeV, (d) $m(\mu\mu K) \in (5.0, 5.7)$ GeV, (e) $m(\mu\mu K) \in (5.05, 5.7)$ GeV and (f) $m(\mu\mu K) \in (5.10, 5.7)$ GeV.

5.7.7 Systematic uncertainties for the muon channel

Muon HLT Scale Factor uncertainty

The Muon High-Level Trigger Scale Factors (Muon HLT SF) are obtained using the tag and probe method with the J/ψ resonance, and their measurement is subject to some uncertainty that includes both statistical and systematic sources. To evaluate the effect of these uncertainties, the SFs are adjusted by $\pm 1 \sigma$, and the resulting impact on the R_K^μ measurement is assessed. This adjustment is made simultaneously for all bins, and the maximum deviation from the nominal R_K^μ value is taken as the uncertainty, in a conservative approach.

$$\delta R_K^\mu = \max(R_K^{\mu,+1\sigma} - R_K^{\mu,\text{nom}}, R_K^{\mu,-1\sigma} - R_K^{\mu,\text{nom}}) \quad (5.10)$$

Table 5.22 summarizes the results

Table 5.22: Difference in R_K^μ from the variation of the HLT SF.

Variation	R_K^μ
SF-1 σ	3.474×10^{-3}
SF+1 σ	3.471×10^{-3}

The systematic uncertainty is determined by taking the maximum difference from the nominal value, which in this case is 0.003×10^{-3} . This value corresponds to a relative uncertainty of 0.08%.

Pile Up Scale Factor uncertainty

The Pile Up Scale Factors (PU SF) are obtained by utilizing the true PU shape, which is calculated using Eq 5.11. This equation includes the total inelastic cross section of the p-p

collisions:

$$N_{\text{true}} = \frac{L \times \sigma(\text{pp})}{f_{\text{LHC}}}, \quad (5.11)$$

where L is the instantaneous luminosity collected with the B -parking HLT, $\sigma(\text{pp})$ is the total inelastic cross section and f_{LHC} the orbital frequency of the LHC beam. The maximum difference that $\pm 1 \sigma$ variation of $\sigma(\text{pp})$ induces in the R_K^μ is taken as the uncertainty.

Table 5.23 shows the result of the PU SF uncertainty.

Table 5.23: Difference in R_K^μ from the variation of the PU SF.

Variation	R_K^μ
SF-1 σ	3.468×10^{-3}
SF+1 σ	3.469×10^{-3}

The systematic uncertainty is determined by taking the maximum difference from the nominal value, which is found to be 0.005×10^{-3} in this case. This corresponds to a relative uncertainty of 0.13%.

BDT Scale Factor uncertainty

The BDT SFs are obtained by dividing the BDT efficiency in the data by the corresponding efficiency in the MC. The SF values are varied by $\pm 1 \sigma$ in a similar manner to the other SF-related uncertainties. Then, modified R_K^μ values are calculated, and the maximum difference from $R_K^{\mu, \text{nominal}}$ is considered as the uncertainty. Table 5.24 shows the result of BDT SF uncertainty.

Table 5.24: Difference in R_K^μ from the variation of the BDT SF.

Variation	R_K^μ
SF-1 σ	3.474×10^{-3}
SF+1 σ	3.473×10^{-3}

The systematic uncertainty is determined by taking the maximum difference from the nominal value, which is found to be 0.001×10^{-3} in this case. This corresponds to a relative uncertainty of 0.02%.

Cross-validation uncertainty

To utilize the full dataset, we employ a cross-validation technique. The approach involves dividing the dataset into eight parts, with each part serving as the training set for a different BDT. These BDTs are then utilized to evaluate events in a section of the dataset different from the one on which they were trained. After this process, eight independent BDTs are produced, and the reported "BDT efficiency" is the mean efficiency. To address potential variations in the efficiencies of the BDTs, R_K^μ is calculated using each BDT independently, and the average deviation from $R_K^{\mu, \text{nominal}}$ is taken as the uncertainty. By taking the average difference as systematic, the result is 0.0058×10^{-3} which corresponds to a relative uncertainty of 0.17%.

Operating below the trigger plateau uncertainty

It is common practice in analyses to employ trigger object cuts that are more restrictive than the online selection to reach the plateau of the trigger efficiency. However, due to the slow turn-on of the $d_{xy}(\mu)$ part of the trigger, a considerable portion of events had to be excluded to achieve the plateau. This would have affected the competitiveness of the outcome. As a result, it was decided to adopt a threshold that is less strict than the plateau and set offline cuts to match those online. To address the systematic associated with this decision, R_K^μ is remeasured using tighter cuts on the triggering muon, and the deviation from $R_K^{\mu, \text{nominal}}$ is

taken as the uncertainty. To prevent the fitting uncertainty from being counted twice, a cut and count method is employed instead of rerunning the fitting code. The ΔR_K^μ is determined to be 0.09×10^{-3} , corresponding to a relative uncertainty of 1.0%.

Choice of parameterization in the fit

Using different fitting functions to extract the signal yield can affect the measurement. To assess the impact of the choice of fitting function, alternative functions are employed, and the variation in the signal yield is determined. A variety of functional forms for both the signal and background are included in the selection of alternative fitting functions. The examination is conducted in both the J/ψ bin and the low- q^2 bin.

Figures C.0.1, C.0.2, C.0.3 and C.0.4 show the alternative templates used and the total fits for the study of the choice of fit parameterization. Tables 5.25 and 5.26 summarize the choice of the alternative function and the impact. The relative systematic uncertainty from the choice of fit parameterization in the muon channel is 2.4%.

Table 5.25: Summary of the systematic uncertainties from the choice of fit parameterization in the J/ψ region for the muon channel.

Type	Nominal shape	Alternative shape	$\Delta N_{\text{sig}}/N_{\text{sig}}$
Signal	3 Gaussian	4 Gaussian	9.0×10^{-4}
Combinatorics	Exponential	Gaussian + 1 st -order Poly	6.6×10^{-3}
Total			6.7×10^{-3}

Table 5.26: Summary of the systematic uncertainties from the choice of fit parameterization in the low- q^2 region for the muon channel.

Type	Nominal shape	Alternative shape	$\Delta N_{\text{sig}}/N_{\text{sig}}$
Signal	double-sided CB + Gaussian	double-sided CB + 2 Gaussian	2.8×10^{-3}
Combinatorics	Exponential	Gaussian + 1 st -order Ploy	2.3×10^{-2}
Total			2.3×10^{-2}

Systematic uncertainty from the J/ψ radiative tail

One potential source of systematic uncertainty is the potential mismodeling of the radiative tail, which can result in incorrect efficiency calculations. Unlike other uncertainties, the mismodeling of the radiative tail only affects the denominator. To quantify this source of uncertainty, we increase the lower bound of the J/ψ q^2 bin to 2.8 GeV, including a larger portion of the tail. We then measure a modified R_K^μ and calculate the deviation from $R_K^{\mu, \text{nominal}}$, which is taken as the uncertainty. The ΔR_K^μ is determined to be 0.019×10^{-3} , corresponding to a relative uncertainty of 0.4%.

Rare mode contribution on the resonant region

The rare decay mode of $B^+ \rightarrow K^+ \mu^+ \mu^-$ can contribute to the signal yield of $B^+ \rightarrow K^+ J/\psi (\rightarrow \mu^+ \mu^-)$. This effect is small compared to the much larger branching fraction of the resonant decay. However, to be conservative, we estimate the rare decay yield in this region and calculate a modified R_K^μ . We then determine the difference between this value and $R_K^{\mu, \text{nominal}}$ and take it as the systematic uncertainty. The resulting ΔR_K^μ is determined to be 0.003×10^{-3} , with a relative error of 0.07%.

Summary of the systematic uncertainties in the muon channel

Table 5.27 summarizes the systematic uncertainties and their relative impact on R_K^μ . The sources of systematic uncertainties are uncorrelated and the total is sum in quadrature.

Table 5.27: Summary of systematic uncertainty for the R_K measurement in the muon channel.

Source	$\Delta R_K^\mu / R_K^\mu$
HLT SF	0.0008
Pile Up SF	0.0013
BDT SF	0.0002
BDT Cross-validation	0.0017
Trigger plateau	0.0010
J/ψ radiative tail	0.0040
Choice of fit parameterization	0.0240
Rare contribution in resonant region	0.0007
Total	0.0245

5.8 The electron channel

This section describes the electron part of the R_K ratio, including event selection, fit strategy, evaluation of systematic uncertainties, and sanity checks.

5.8.1 Event preselection

To form a $B^+ \rightarrow K^+ e^+ e^-$ candidate, two oppositely charged electrons are combined with a track that is hypothesized to be a K^+ meson from the probe side of the event. The candidate is accepted only if both electrons and the K^+ meson are close to a trigger muon within $|\Delta z| < 0.1$ cm, but far enough to remove contamination from the trigger muon ($\Delta R > 0.03$). Additionally, each particle must have a transverse momentum greater than 0.5 GeV and an absolute pseudorapidity less than 2.4.

Two oppositely charged electrons are combined to form a dielectron system, which is categorized based on the type of electrons:

- PF-PF: The dielectron system is constructed by two PF electrons.
- PF-LP: The dielectron system is constructed by a PF electron and a LP electron, where the LP electron is not overlapped with any PF electron in the event.
- LP-LP: The dielectron system is constructed by two LP electrons and, where both LP electrons are not overlapped with any PF electron in the event.

Due to the low purity of the LP-LP category, it is not possible to extract a significant amount of signal from the large background, hence only the PF-PF and PF-LP categories are used in this analysis.

The leading electron is required to have $p_T(e_1) > 1.3$ GeV. Then, the two electrons are required to have a longitudinal separation within $|\Delta z| < 0.1$ cm, a separation greater than $\Delta R > 0.03$ to prevent overlap, and an invariant mass below $m(e^+e^-) < 5$ GeV. To further reduce the fake rate, an additional cut on the LP electron ID is imposed, requiring it to be greater than -2.5 . Afterwards, a kinematic fit is performed on the dielectron system to obtain a common vertex, as described in Section 5.5.4. The two-track vertex fit χ^2 probability must be at least 0.001% to be considered for analysis.

A B^+ candidate is formed by attaching the K^+ meson to the dielectron system. To avoid overlapping, the K^+ meson must have a separation $\Delta R > 0.03$ from the electrons. The B^+ candidate must have $p_T > 1.75$ GeV and a pre-fit invariant mass of $4 < m(K^+e^+e^-)_{\text{pre-fit}} < 7$ GeV. The electrons and K^+ meson are then fitted to a common vertex, as described in Sec.5.5.4. The invariant mass of the B^+ candidate obtained from the kinematic fit must be between 4.5 and 6.0 GeV.

Using the same adaptive grid search approach as detailed in section 5.7.2, additional selections are applied to further reduce the background. These selections are:

PF-PF

- $4.7 < m(K^+e^+e^-) < 5.7$ GeV
- $\cos \alpha_{2D} > 0.95$
- $|d_{3D}(K^+, e^+e^-)| < 0.06$
- $prob(SV) > 10^{-5}$

PF-LP

- $4.7 < m(K^+e^+e^-) < 5.7$ GeV

- $\cos \alpha_{2D} > 0.95$
- $|d_{3D}(K^+, e^+e^-)| < 0.06$
- $\text{prob}(\text{SV}) > 10^{-5}$
- $p_T(e_2) > 1 \text{ GeV}$
- $\text{ID}(e_1) > -2$
- $\text{ID}(e_2) > 0$

Table 5.1 defines the four dilepton invariant mass bins.

5.8.2 Event selection with a BDT in the electron channel

Following the preselection, a significant amount of combinatorial background remains in the data sample. To extract the signal from this background, a multivariate discriminator is used based on the XGBoost algorithm [59].

The decays from D meson can contribute significantly to the background in this analysis, resulting a distinctive D mass peak. To mitigate this, anti- D^0 cuts of $m(K^+e^-_{[\rightarrow\pi^-]}) > 2 \text{ GeV}$ and $m(K^+_{[\rightarrow\pi^+]}e^-_{[\rightarrow K^-]}) > 2 \text{ GeV}$ are applied to veto the contamination from D meson decays, where the subscripts indicate the mass given to the candidates. This further cleaning of background facilitates the BDT training and thus improves the performance of the BDT.

Different BDTs are trained independently for different dielectron categories using simulated $B^+ \rightarrow K^+e^+e^-$ candidates as signal proxies, and candidates from the side-bands ($4.7 < m(K^+e^+e^-) < 5.0 \text{ GeV}$ or $5.5 < m(K^+e^+e^-) < 5.7 \text{ GeV}$) of the data as background proxies. The training data must satisfy $1.05 < m(e^+e^-) < 2.45 \text{ GeV}$ to optimize the discrimination power in the low- q^2 region. The full dataset is unitized using a cross-validation

technique during training as described in Sec. 5.7.3, preventing the same data being used for training and fitting. The BDTs' hyper-parameters are chosen based on Bayesian optimization, and the BDTs are trained to maximize the predicted significance in the low- q^2 region.

Feature selection and feature importance

To achieve good discriminatory power, it is crucial to carefully select features for the BDT. These features are generated based on the triplet kinematics, topology, vertex quality, particle identification, and isolation, and are listed in Table 5.28 with detailed descriptions. Normalizing the transverse momenta by the triplet mass is necessary to prevent mass sculpting.

The selected features undergo rigorous testing to ensure their effectiveness in improving the BDT performance. Unnecessary variables are eliminated to simplify the model. The importance of the features for PF-PF and PF-LP can be observed in the feature importance plots of the BDT, as shown in Fig. 5.52.

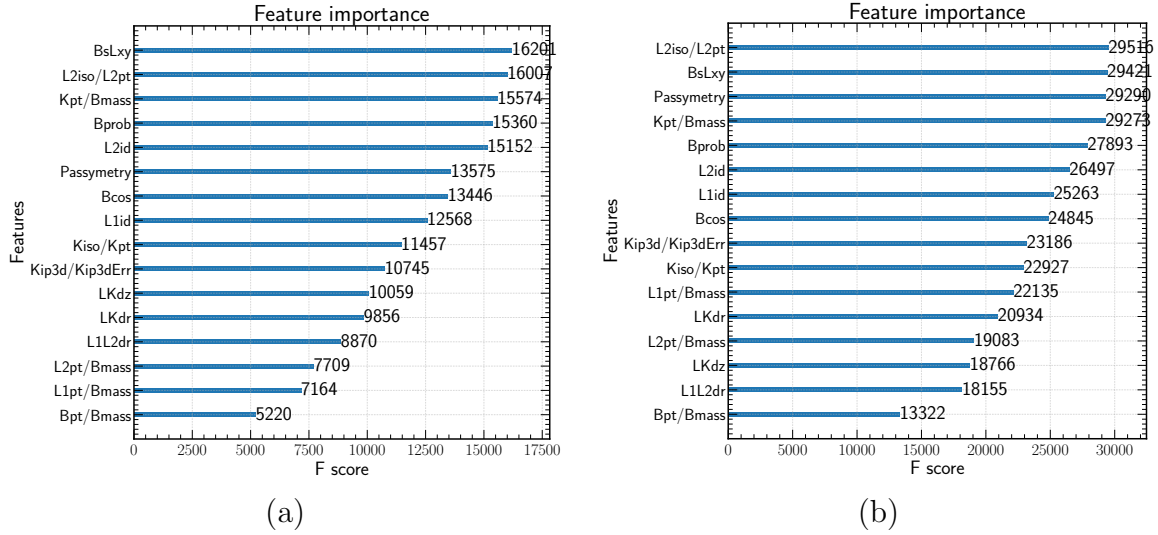


Figure 5.52: Feature importance of the trained BDT for the PF-PF (a) and PF-LP (b) categories.

Table 5.28: Variables used in the analysis BDT in the electron channel.

Variables	Descriptions
$p_T(e_1/e_2/K^+/B^+)/m(K^+e^+e^-)$	Transverse momentum divided by the invariant mass of the triplet of e_1 , e_2 , K^+ and B^+ , respectively
$dz(e, K^+)$	Longitudinal distance between the electron and the kaon
$d_{3D}(K^+, e^+e^-)/\sigma_{d_{3D}}$	Kaon 3D impact parameter significance w.r.t the dielectron secondary vertex
$prob(SV)$	Secondary vertex fit χ^2 probability
$L_{xy}/\sigma_{L_{xy}}$	Secondary vertex transverse displacement significance
$\cos \alpha_{2D}$	Cosine angle in the transverse plane between the B momentum and the separation between the secondary vertex and the beamspot.
$\Delta R(e^+, e^-)$	ΔR between two electrons
$\Delta R(e, K^+)$	ΔR between the electron, and the kaon
$\frac{ \mathbf{p}(e^+e^-) \times \mathbf{r} - \mathbf{p}(K^+) \times \mathbf{r} }{ \mathbf{p}(e^+e^-) \times \mathbf{r} + \mathbf{p}(K^+) \times \mathbf{r} }$	Asymmetry of the transverse momentum w.r.t. B trajectory of the dielectron momentum and that of the K^+ momentum
$ID(e_1/e_2)$	Electron ID of e_1 and e_2 , respectively
$I_{\Delta R=0.4}^{\text{rel}}(e_1/e_2/K^+)$	Relative track isolation around the cone $\Delta R < 0.4$ of e_1 , e_2 and K^+ , respectively

Checks for overtraining

Overtraining occurs when the BDT performs well on the training data but fails to generalize well on new, unseen testing data. To avoid overtraining, L2 regularization is incorporated into the loss function to penalize the model's complexity. Additionally, careful feature selection, including dropping unnecessary variables, can also help reduce model complexity. Figure 5.53 illustrates the BDT score distribution for both the training and testing data, indicating good agreement between the two datasets. In this case, overtraining is not observed.

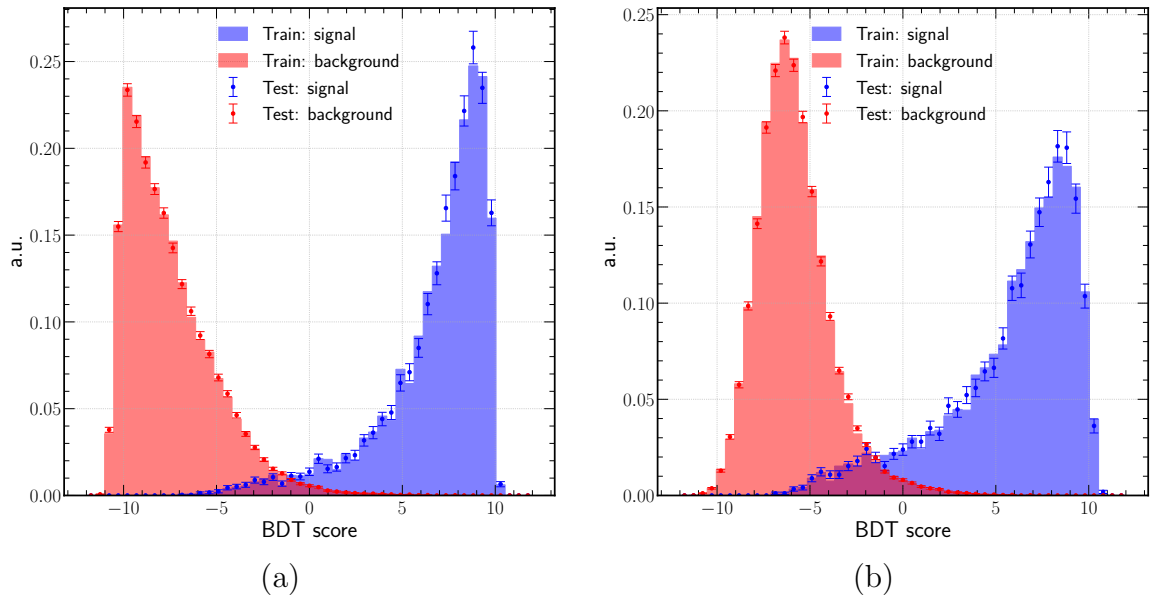


Figure 5.53: Comparing the BDT score distribution of training data against testing data for the PF-PF (a) and PF-LP (b) categories. The distribution shows agreement between training and testing.

5.8.3 Mass fit

An unbinned maximum likelihood fit is performed on the data to determine the number of signal events in various q^2 ranges. As described by probability density functions (PDF), the signal shapes and various backgrounds are fixed using data and MC, known as the templates. The normalizations of the templates are allowed to vary during the fit to obtain the numbers of signals and backgrounds from the data.

The BDT cuts are chosen to maximize the significance of the expected signal in the low- q^2 bin. Figure 5.54 shows the electron channel's significance as a function of the BDT cut.

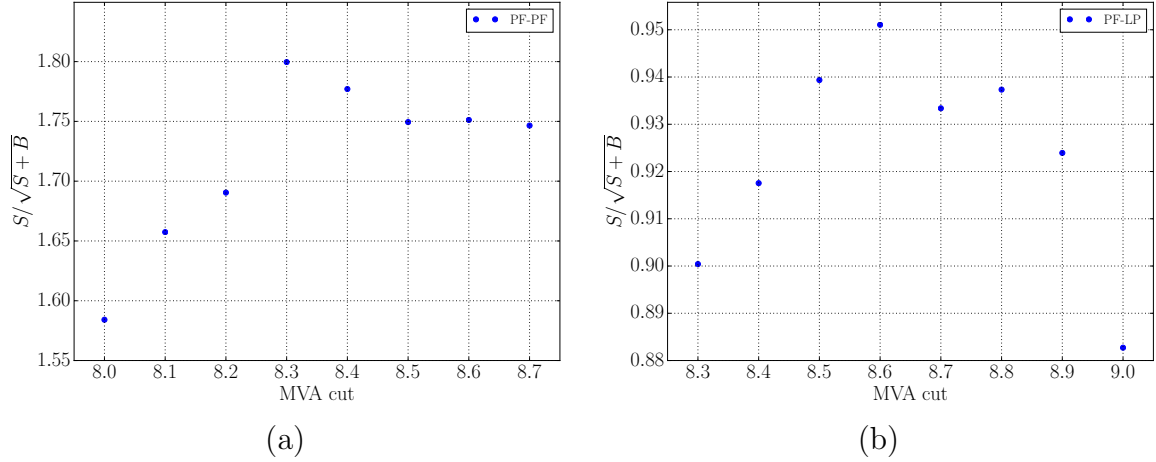


Figure 5.54: Significance as a function of BDT cut for low- q^2 for PF-PF (a) and PF-LP (b).

Fit in the low- q^2 bin

There are two types of background present in the low- q^2 bin: J/ψ contamination and combinatorial background.

The J/ψ contamination is caused by the long radiative tail of $B^+ \rightarrow K^+ J/\psi(\rightarrow e^+e^-)$ entering into the low- q^2 region, affecting the lower sideband of the fit. The shape of this background is modeled using $B^+ \rightarrow K^+ J/\psi(\rightarrow e^+e^-)$ MC simulation and described using a kernel density estimator (KDE). The template is constructed in an extended mass range down to $m(K^+e^+e^-) = 4.5$ GeV to capture the shape better. The combinatorial background affects the fit in the entire $m(K^+e^+e^-)$ region and is modeled by an exponential function.

The signal is modeled using $B^+ \rightarrow K^+e^+e^-$ MC simulation and described by a double-sided Crystal-ball function. The lower mass exponential tail of the double-sided Crystal-ball function captures the radiative tail due to Bremsstrahlung, while the upper mass exponential tail captures the excessive energy recovery from the GSF algorithm.

Table 5.29 summarizes the templates and models used in the low- q^2 region. The fits of the templates are displayed in Fig. 5.55 and 5.56, while the fit to the data is shown in Fig. 5.57. The signal and background yields are listed in Table 5.30.

Table 5.29: Summary of the fit in the low- q^2 bin.

Process	Sample	Model	Shape	Normalization
Signal	$B^+ \rightarrow K^+e^+e^-$ MC	Double-sided CB	Fixed	Floating
J/ψ Contamination	$B^+ \rightarrow K^+ J/\psi(\rightarrow e^+e^-)$ MC	Exponential	Fixed	Fixed
Combinatorics	same-sign data	KDE	Fixed	Floating

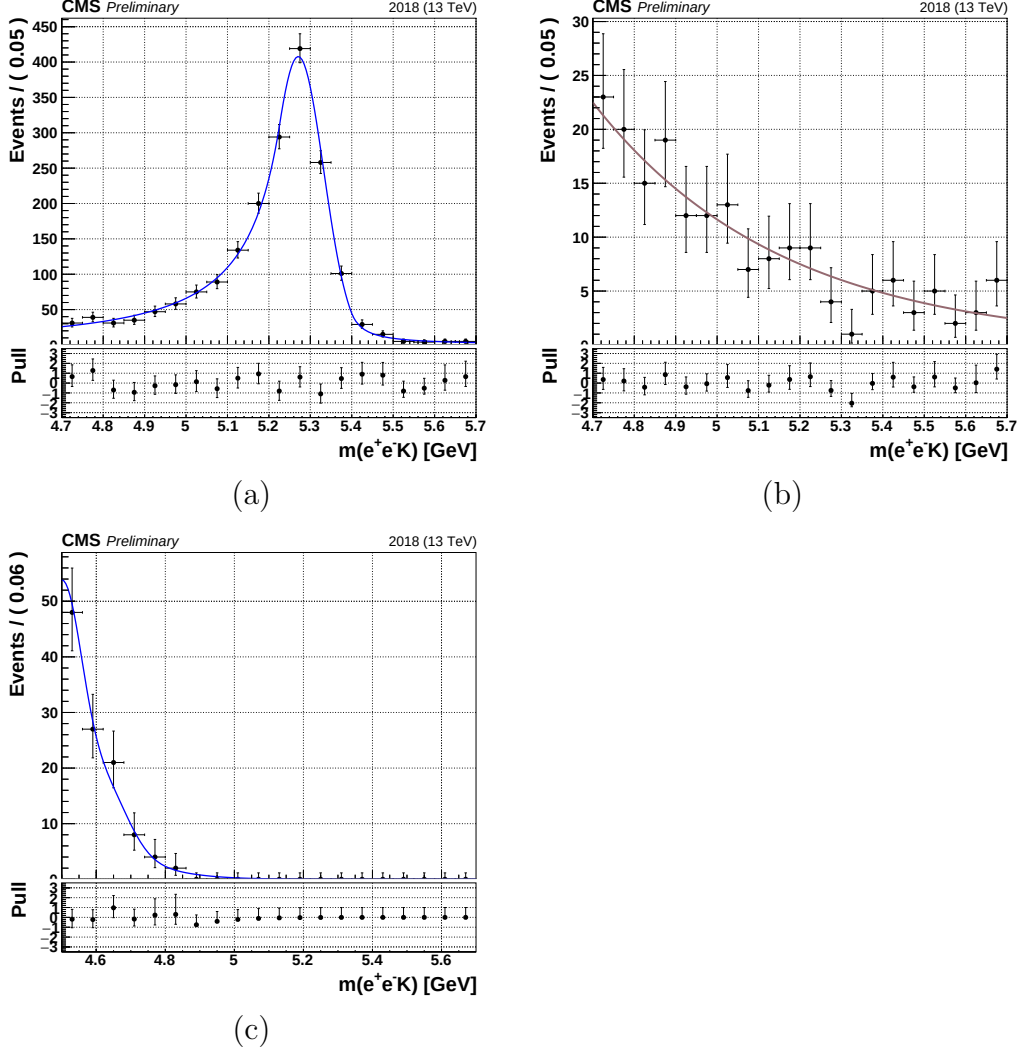
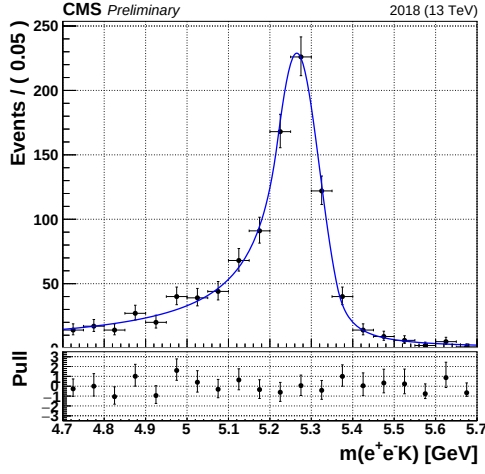


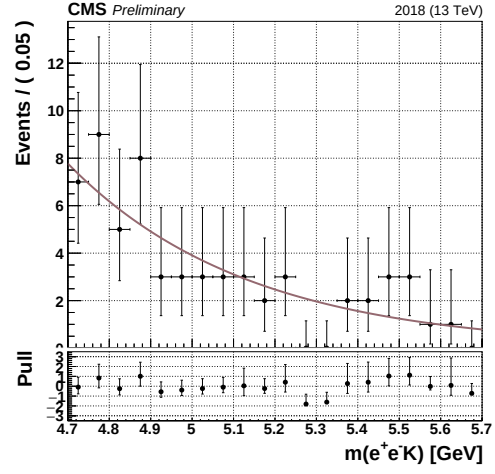
Figure 5.55: Fit of simulated $B^+ \rightarrow K^+e^+e^-$ (a), same-sign data (b), and simulated $B^+ \rightarrow K^+J/\psi(e^+e^-)$ (c) in the low- q^2 bin for PF-PF.

Table 5.30: Yields for signal and background sources in low- q^2 .

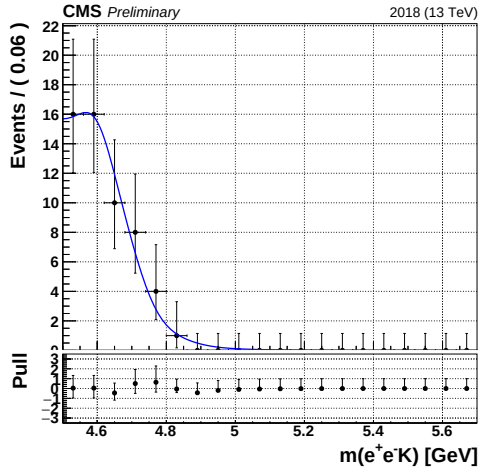
Source	PF-PF Yield	PF-LP Yield
$B^+ \rightarrow K^+e^+e^-$ (Signal)	17.9 ± 7.2	3.0 ± 5.9
Combinatorics	59.7 ± 9.9	60.8 ± 9.7



(a)



(b)



(c)

Figure 5.56: Fit of simulated $B^+ \rightarrow K^+e^+e^-$ (a), same-sign data (b), and simulated $B^+ \rightarrow K^+J/\psi(e^+e^-)$ (c) in the low- q^2 bin for PF-LP.

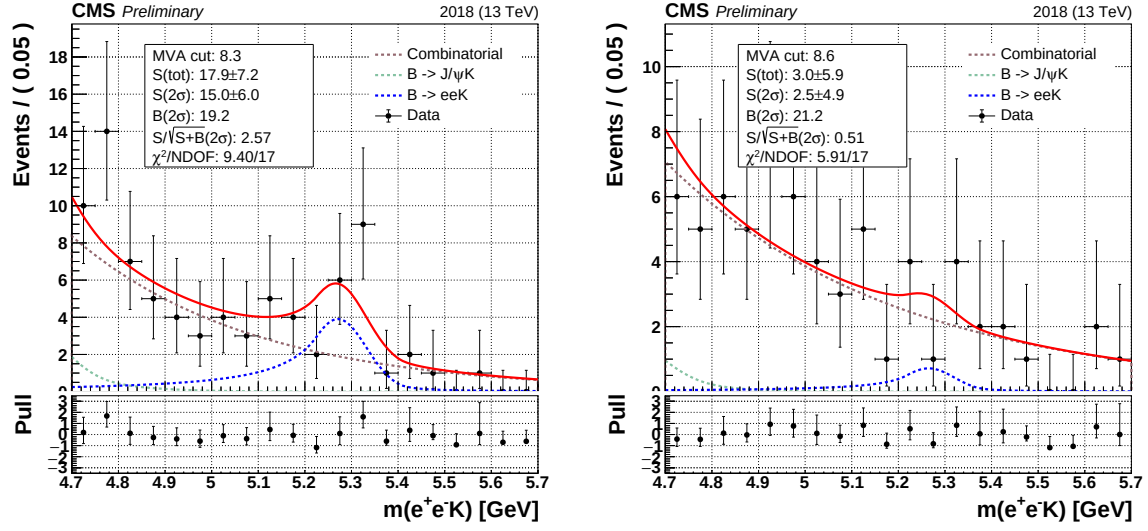


Figure 5.57: Fit to data (top) and pull (bottom) in the low- q^2 bin for the PF-PF (left) and PF-LP (right) category.

Fit in the J/ψ bin

In the J/ψ bin, four types of background are present: the partially reconstructed background, contamination from $B^+ \rightarrow \pi^+ J/\psi(\rightarrow e^+e^-)$, background from other B components, and combinatorial background.

The partially reconstructed background is due to partially reconstructed triplets from $B^{0/+} \rightarrow K^*(892)^{0/+} J/\psi(\rightarrow e^+e^-)$. It has two components: a kaon component where the pion is not reconstructed, and a pion component where the kaon is not reconstructed. The shape of this background is modeled using $B^{0/+} \rightarrow K^*(892)^{0/+} J/\psi(\rightarrow e^+e^-)$ MC simulation and described by a KDE. The relative normalization between the two components is fixed using MC. The contamination from $B^+ \rightarrow \pi^+ J/\psi(\rightarrow e^+e^-)$ in the J/ψ region is modeled using the generic B MC simulation, selecting the relevant process. Its shape is described by a Crystal-ball function. The background from other B components is due to B decays other than $B^{0/+} \rightarrow K^*(892)^{0/+} J/\psi(\rightarrow e^+e^-)$ and $B^+ \rightarrow \pi^+ J/\psi(\rightarrow e^+e^-)$. It is modeled using generic $B \rightarrow X e^+e^-$ MC, excluding certain decay modes. The shape is described by a KDE. The combinatorial background affects the entire $m(K^+e^+e^-)$ region and is modeled by an exponential function. Its shape is taken from the same-sign event distribution in data, where two electrons have the same charge.

The signal is modeled using $B^+ \rightarrow K^+ J/\psi(\rightarrow e^+e^-)$ MC simulation and described by a double-sided Crystal-ball function plus a Gaussian. The lower mass exponential tail in the double-sided Crystal-ball function captures the radiative tail due to Bremsstrahlung, and the upper mass exponential tail captures the excessive energy recovery from GSF algorithm. An extra Gaussian function is added to the lower mass region to model the distorted tail

due to the q^2 cut.

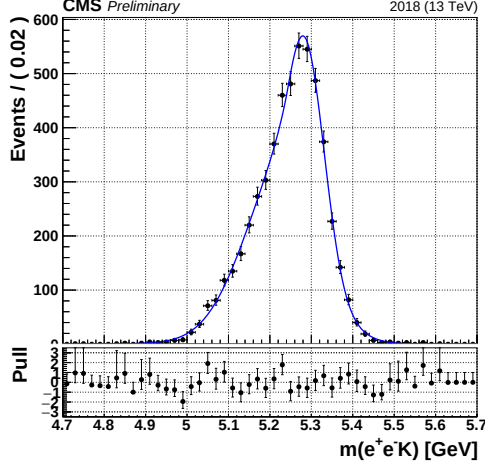
Table 5.31 summarizes the templates and models used in the J/ψ bin. Figures 5.58 and 5.59 show the fits to the templates, and Fig. 5.60 shows the fit to the data. The yields of the signal and background are listed in Table 5.32.

Table 5.31: Summary of the fit in the J/ψ bin.

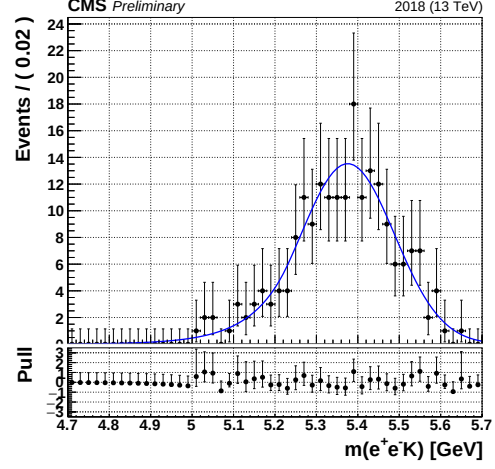
Process	Sample	Model	Shape	Normalization
Signal	$B^+ \rightarrow K^+ J/\psi (\rightarrow e^+ e^-)$ MC	CB + Gaussian	Fixed	Floating
Partially reco. bkg	$B^{+0} \rightarrow K^*(892)^{+0} J/\psi (\rightarrow e^+ e^-)$ MC	KDE	Fixed	Floating
Other B components	generic B MC	KDE	Fixed	Floating
$B^+ \rightarrow \pi^+ J/\psi (\rightarrow e^+ e^-)$ Contamination	generic B MC	CB	Fixed	Floating
Combinatorics	same-sign data	Exponential	Fixed	Floating

Table 5.32: Yields for signal and background sources in J/ψ .

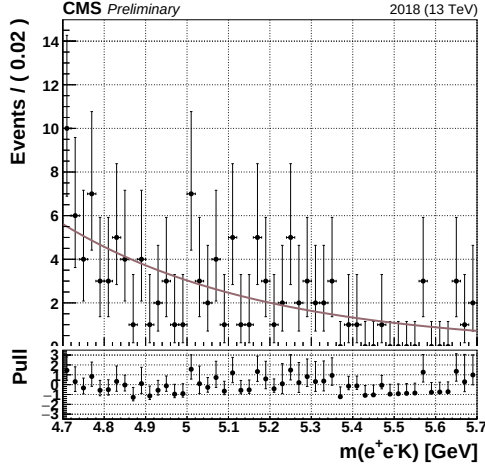
Source	PF-PF Yield	PF-LP Yield
$B^+ \rightarrow K^+ J/\psi (\rightarrow e^+ e^-)$ (Signal)	4856.9 ± 83.7	2097.8 ± 58.1
Partially reco. bkg	2307.4 ± 88.9	786.2 ± 50.6
Other B components	1231.5 ± 81.9	485.6 ± 43.8
$B^+ \rightarrow \pi^+ J/\psi (\rightarrow e^+ e^-)$	104.9 ± 46.8	46.2 ± 32.4
Combinatorics	261.5 ± 107.8	24.2 ± 16.2



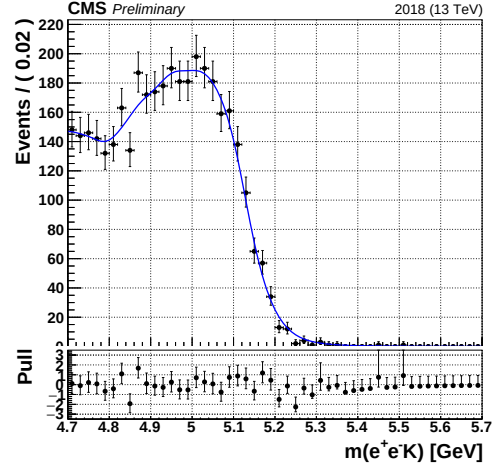
(a)



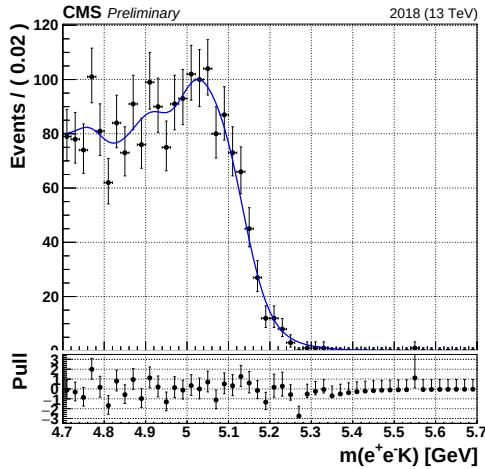
(b)



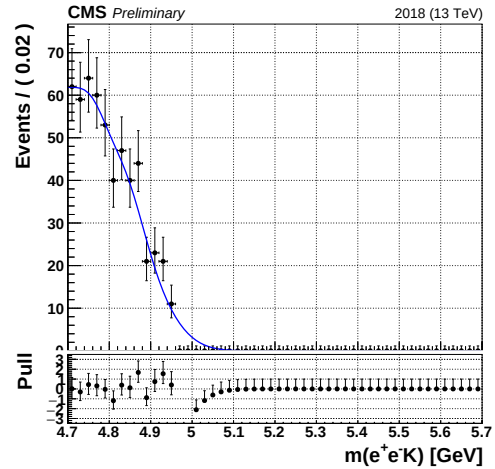
(c)



(d)

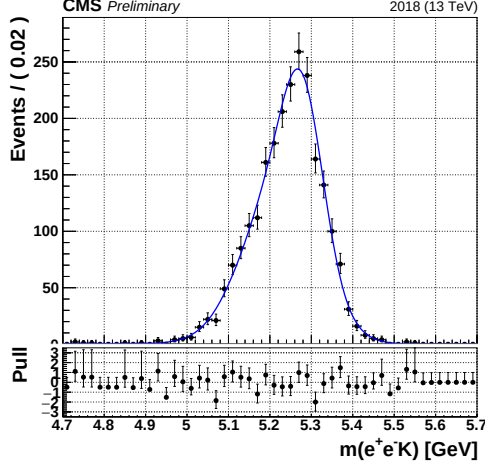


(e)

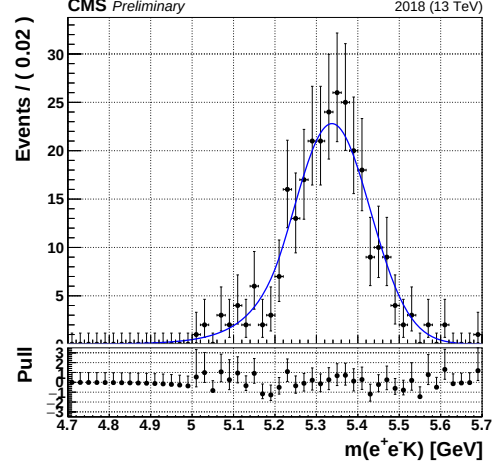


(f)

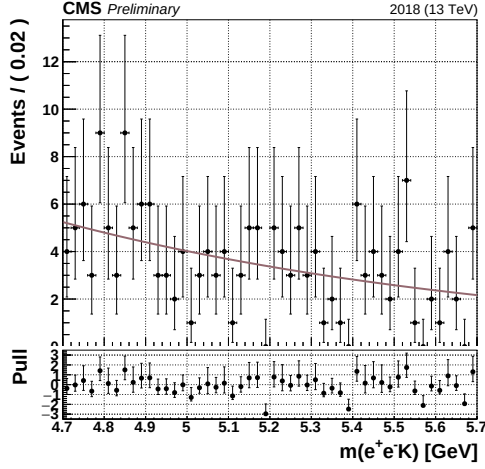
Figure 5.58: Fit of simulated $B^+ \rightarrow K^+ J/\psi(e^+e^-)$ (a), simulated $B^+ \rightarrow \pi^+ J/\psi(e^+e^-)$ (b), same-sign data (c), simulated $B^0 \rightarrow K^*(892) J/\psi(e^+e^-)$ (d), simulated $B^+ \rightarrow K^*(892)^+ J/\psi(e^+e^-)$ (e), and simulated other B background in J/ψ bin for PF-PF.



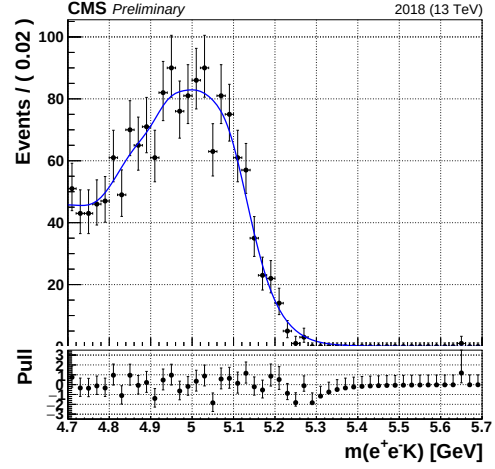
(a)



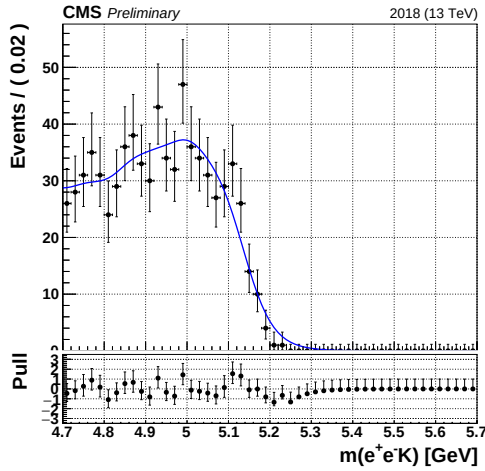
(b)



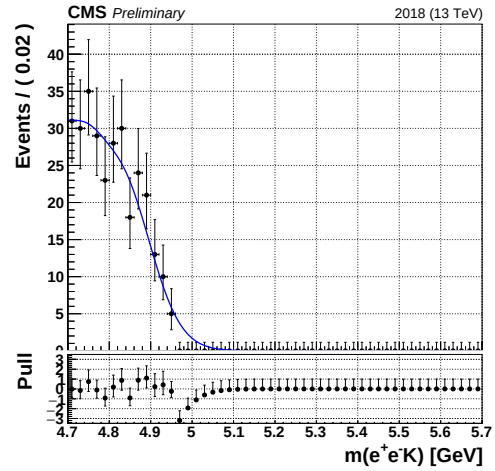
(c)



(d)



(e)



(f)

Figure 5.59: Fit of simulated $B^+ \rightarrow K^+ J/\psi(e^+e^-)$ (a), simulated $B^+ \rightarrow \pi^+ J/\psi(e^+e^-)$ (b), same-sign data (c), simulated $B^0 \rightarrow K^*(892) J/\psi(e^+e^-)$ (d), simulated $B^+ \rightarrow K^*(892)^+ J/\psi(e^+e^-)$ (e), and simulated other B background in J/ψ bin for PF-LP.

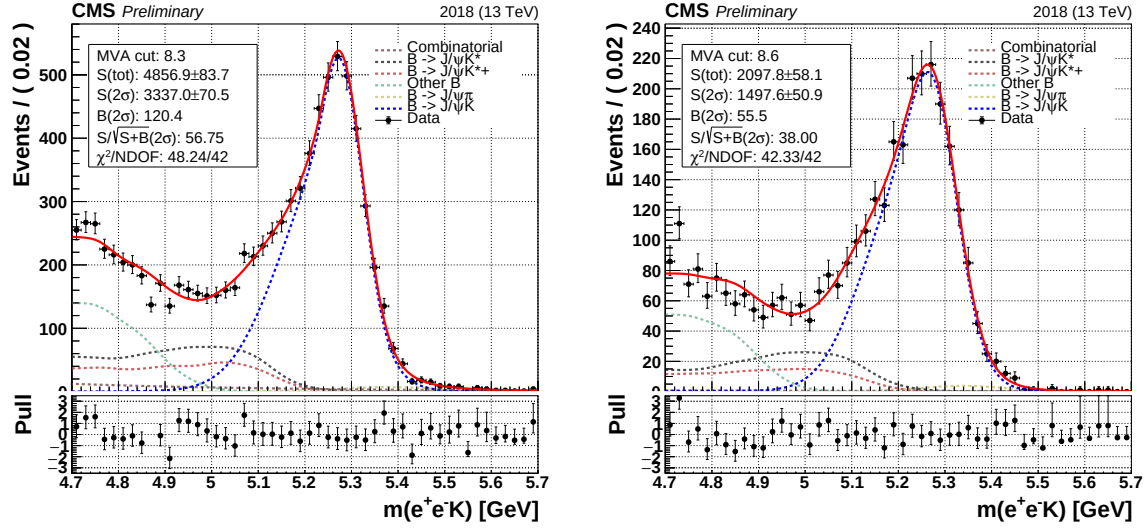


Figure 5.60: Fit to data (top) and pull (bottom) in the J/ψ bin for the PF-PF (left) and PF-LP (right) category.

Fit in the $\psi(2S)$ bin

In the $\psi(2S)$ bin, there are three types of background: partially reconstructed background, backgrounds including other B components, and combinatorial background.

The partially reconstructed background is caused by partially reconstructed triplets from $B^{0/+} \rightarrow K^*(892)^{0/+}\psi(2S)(\rightarrow e^+e^-)$. As in the J/ψ bin, it consists of a kaon component and a pion component, and its shape is modeled using $B^{0/+} \rightarrow K^*(892)^{0/+}\psi(2S)(\rightarrow e^+e^-)$ Monte Carlo (MC) simulation. The shape of the background is described by a KDE, and the relative normalization between the two components is determined using MC. The other B components background is caused by B decays other than $B^{0/+} \rightarrow K^*(892)^{0/+}\psi(2S)(\rightarrow e^+e^-)$. It is modeled using generic $B \rightarrow Xe^+e^-$ MC, excluding decay modes already included in the fit. The shape is described by a KDE. The combinatorial background affects the fit in the entire $m(K^+e^+e^-)$ region. It is modeled using an exponential function, and its shape is taken from the same-sign event distribution in data, where two leptons are required to have the same charge.

The signal is modeled using $B^+ \rightarrow K^+\psi(2S)(\rightarrow e^+e^-)$ MC simulation and described by a double-sided Crystal-ball function plus a Gaussian. The lower mass exponential tail in the double-sided Crystal-ball function accounts for the radiative tail due to Bremsstrahlung, and the upper mass exponential tail captures the excessive energy recovery from GSF algorithm. An additional Gaussian function is added to the lower mass region to model the distorted tail due to the q^2 cut.

The templates and models used in the $\psi(2S)$ bin are summarized in Table 5.33. The fits to the templates are shown in Fig. 5.61 and 5.62, and the fit to the data is shown in Fig.

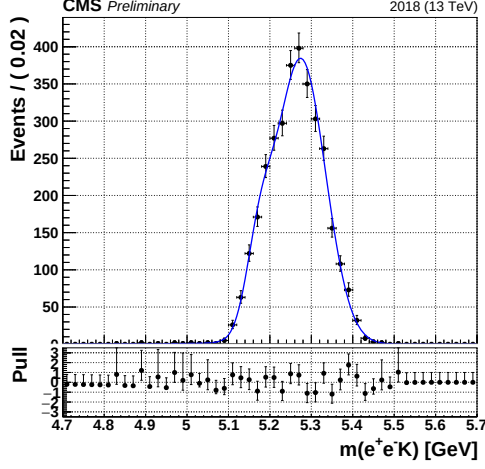
5.63. The yields of the signal and background are listed in Table 5.34.

Table 5.33: Summary of the fit in the $\psi(2S)$ bin.

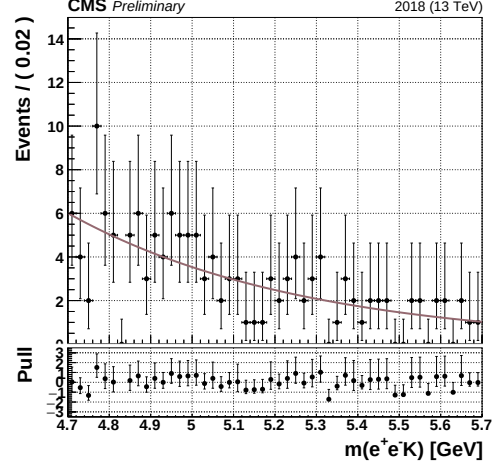
Process	Sample	Model	Shape	Normalization
Signal	$B^+ \rightarrow K^+ \psi(2S)(\rightarrow e^+ e^-)$ MC	CB + Gaussian	Fixed	Floating
Partially reco. bkg	$B^{+0} \rightarrow K^*(892)^{+0} \psi(2S)(\rightarrow e^+ e^-)$ MC	KDE	Fixed	Floating
Combinatorics	same-sign data	Exponential	Fixed	Floating

Table 5.34: Yields for signal and background sources in $\psi(2S)$.

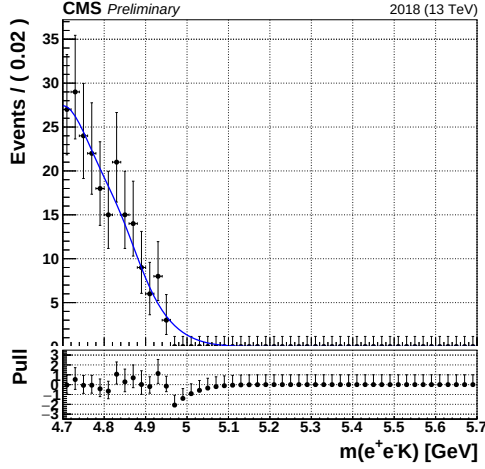
Source	PF-PF Yield	PF-LP Yield
$B^+ \rightarrow K^+ \psi(2S)(\rightarrow e^+ e^-)$ (Signal)	320.2 ± 20.4	93.9 ± 10.9
Partially reco. bkg	115.7 ± 20.0	32.7 ± 8.7
Other B components	25.8 ± 16.6	3.8 ± 7.2
Combinatorics	142.3 ± 33.7	31.6 ± 12.8



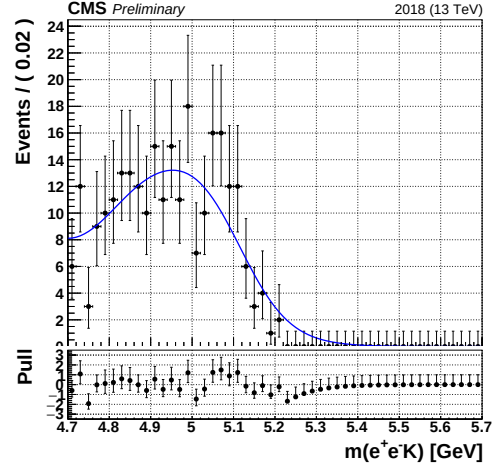
(a)



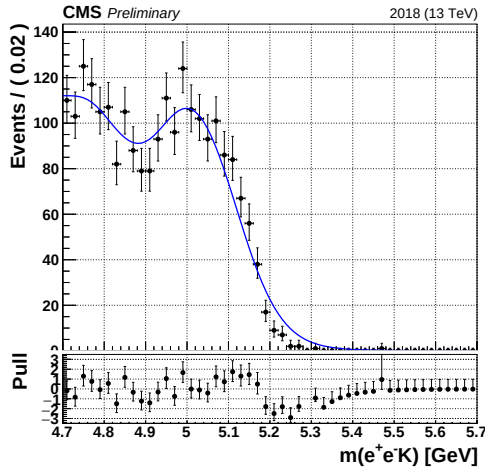
(b)



(c)

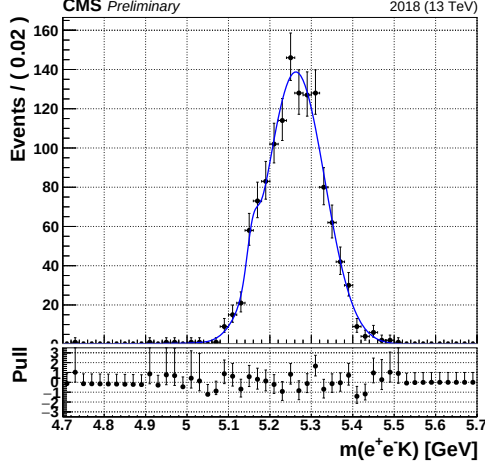


(d)

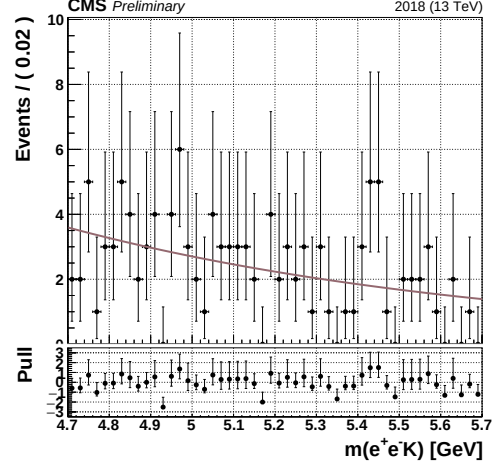


(e)

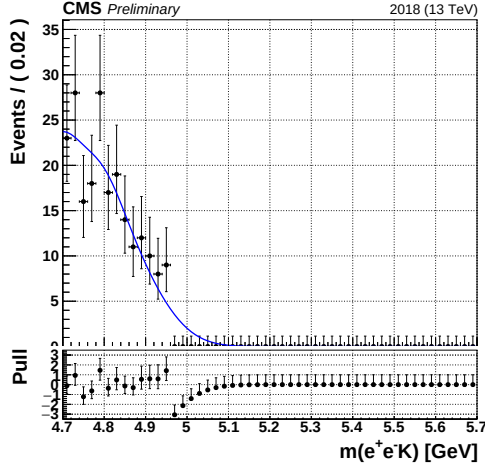
Figure 5.61: Fit of simulated $B^+ \rightarrow K^+\psi(2S)(e^+e^-)$ (a), same-sign data (b), simulated other B background (c), simulated $B^0 \rightarrow K^*(892)\psi(2S)(e^+e^-)$ (d), and simulated $B^+ \rightarrow K^*(892)^+\psi(2S)(e^+e^-)$ (e) in $\psi(2S)$ bin for PF-PF.



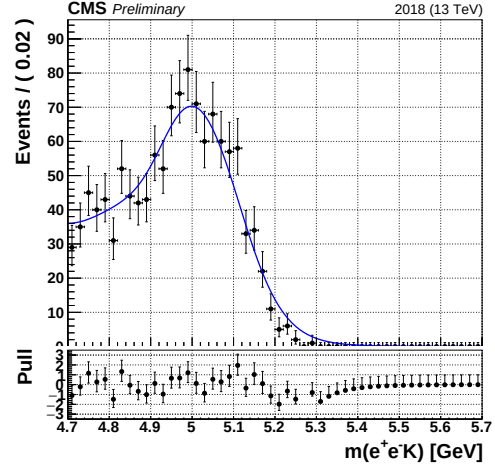
(a)



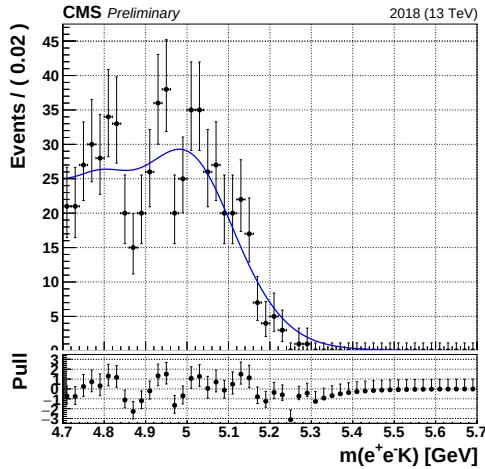
(b)



(c)



(d)



(e)

Figure 5.62: Fit of simulated $B^+ \rightarrow K^+\psi(2S)(e^+e^-)$ (a), same-sign data (b), simulated other B background (c), simulated $B^0 \rightarrow K^*(892)\psi(2S)(e^+e^-)$ (d), and simulated $B^+ \rightarrow K^*(892)^+\psi(2S)(e^+e^-)$ (e) in $\psi(2S)$ bin for PF-LP.

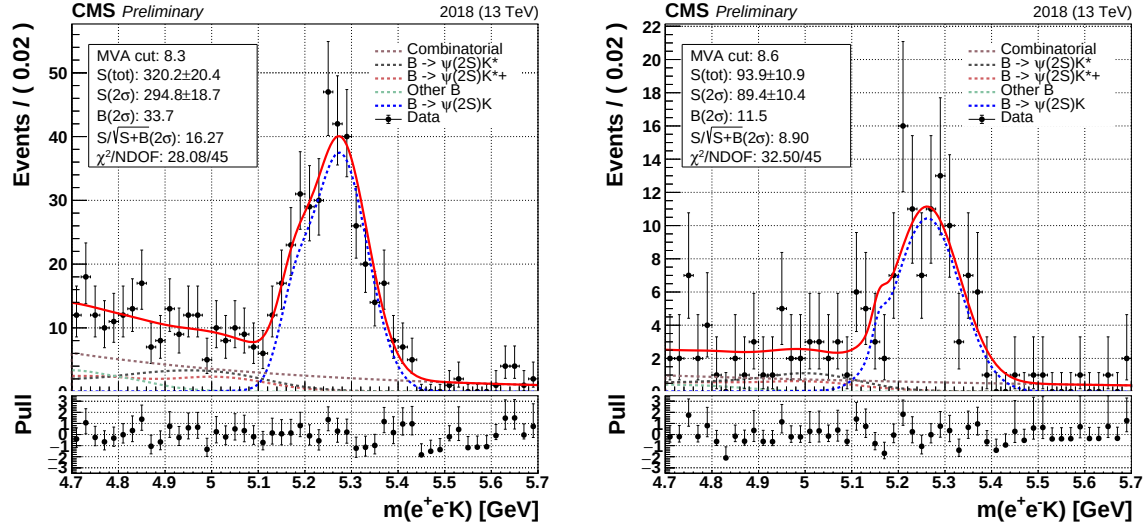


Figure 5.63: Fit to data (top) and pull (bottom) in the $\psi(2S)$ bin for the PF-PF (left) and PF-LP (right) category.

Fit in the high- q^2 bin

In the high- q^2 bin, there are four types of background: contamination from partially reconstructed $\psi(2S)$ background, contamination from $B^+ \rightarrow K^+\psi(2S)(\rightarrow e^+e^-)$ background, other B components background, and combinatorial background.

The contamination from partially reconstructed $\psi(2S)$ background is caused by the long tail of $B^{0/+} \rightarrow K^*(892)^{0/+}\psi(2S)(\rightarrow e^+e^-)$ decays entering the high- q^2 region. This background affects the entire $m(K^+e^+e^-)$ range and is modeled using $B^{0/+} \rightarrow K^*(892)^{0/+}\psi(2S)(\rightarrow e^+e^-)$ MC simulation, with its shape described by a KDE. The contamination from $B^+ \rightarrow K^+\psi(2S)(\rightarrow e^+e^-)$ background is also caused by the long tail of $B^+ \rightarrow K^+\psi(2S)(\rightarrow e^+e^-)$ decays entering the high- q^2 region. This background affects only the upper sideband of the fit and is modeled using $B^+ \rightarrow K^+\psi(2S)(\rightarrow e^+e^-)$ MC simulation, with its shape described by a KDE. The other B components background is caused by the other B decays, and its shape is modeled by an exponential function. The combinatorial background affects the entire $m(K^+e^+e^-)$ range and is modeled by an exponential function, with its shape taken from the same-sign event distribution in data where two electrons have the same charge.

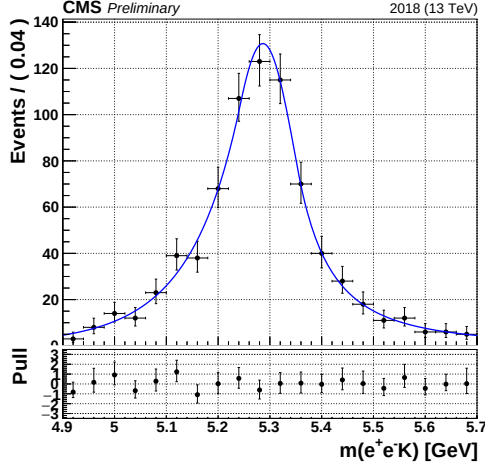
The signal is modeled by $B^+ \rightarrow K^+e^+e^-$ MC simulation and is described by a double-sided Crystal-ball function. The lower mass exponential tail captures the radiative tail due to Bremsstrahlung, and the upper mass exponential tail captures the excessive energy recovery from the GSF algorithm.

The templates and models used in the high- q^2 region are summarized in Table 5.35. The fits of the templates are shown in Fig. 5.64 and 5.65, and the fit to the data is shown in Fig. 5.66.

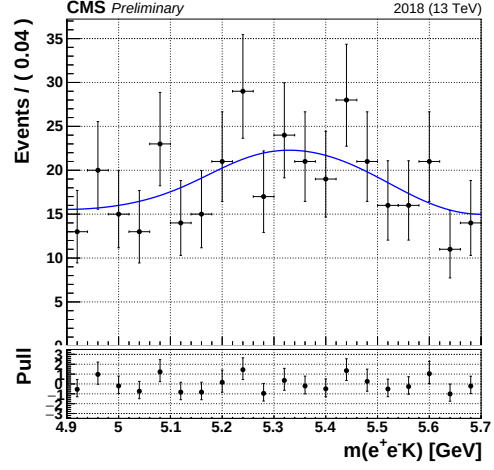
Note that in the electron channel, there is no sensitivity to extract any signal in the high- q^2 region, and this information is included for completeness. The signal region is kept blinded at this moment.

Table 5.35: Summary of the fit in the high- q^2 bin.

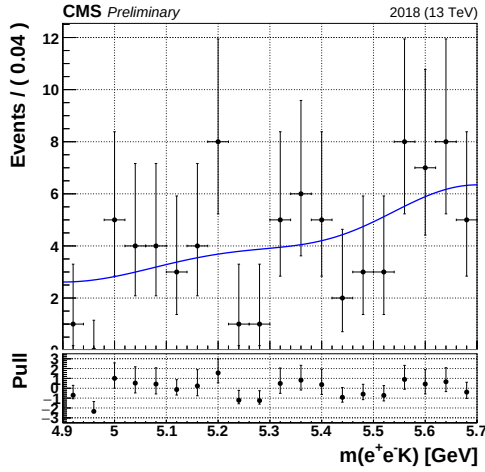
Process	Sample	Model	Shape	Normalization
Signal	$B^+ \rightarrow K^+ e^+ e^-$ MC	Double-sided CB	Fixed	Floating
$\psi(2S)$ Contamination	$B^+ \rightarrow K^+ \psi(2S)(\rightarrow e^+ e^-)$ MC	KDE	Fixed	Floating
Contamination from $\psi(2S)$ partially reco. bkg	$B^0 \rightarrow K^{*}(892)^0 \psi(2S)(\rightarrow e^+ e^-)$ MC	KDE	Fixed	Floating
Other B compinents	N/A	Exponential	Floating	Floating
Combinatorics	same-sign data	KDE	Fixed	Floating



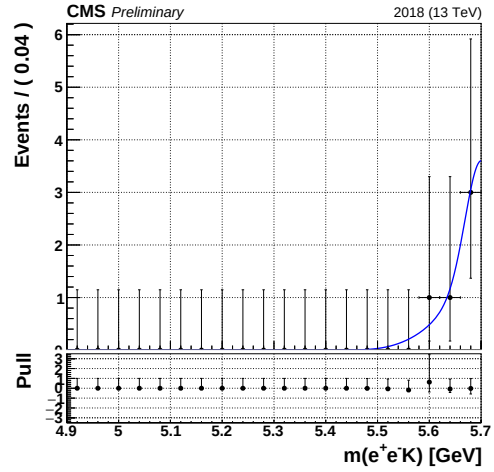
(a)



(b)

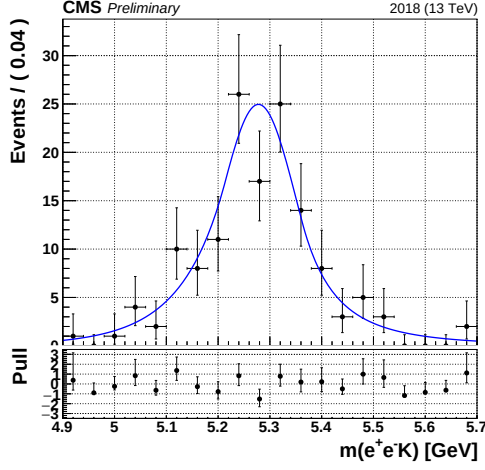


(c)

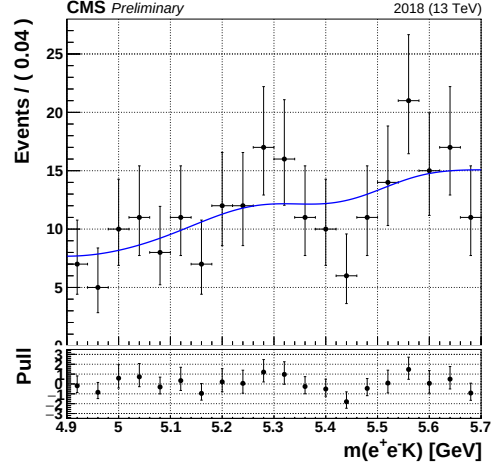


(d)

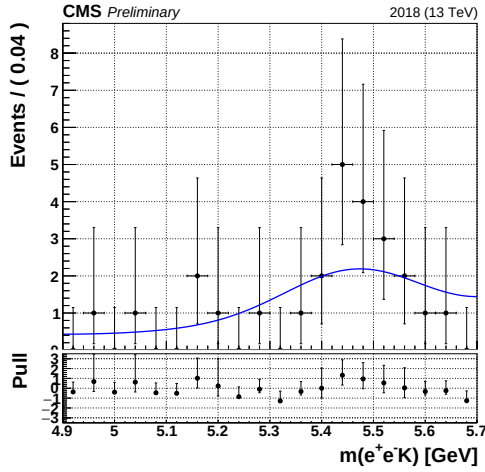
Figure 5.64: Fit of simulated $B^+ \rightarrow K^+e^+e^-$ (a), same-sign data (b), simulated $B^0 \rightarrow K^*(892)\psi(2S)(e^+e^-)$ (c), and simulated $B^+ \rightarrow K+\psi(2S)(e^+e^-)$ (d) in high- q^2 bin for PF-PF.



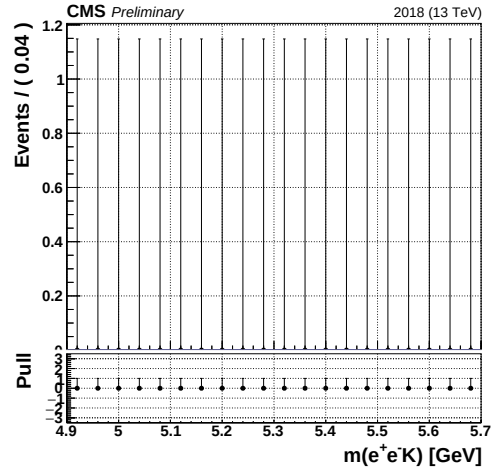
(a)



(b)



(c)



(d)

Figure 5.65: Fit of simulated $B^+ \rightarrow K^+e^+e^-$ (a), same-sign data (b), simulated $B^0 \rightarrow K^*(892)\psi(2S)(e^+e^-)$ (c), and simulated $B^+ \rightarrow K+\psi(2S)(e^+e^-)$ (d) in high- q^2 bin for PF-LP.

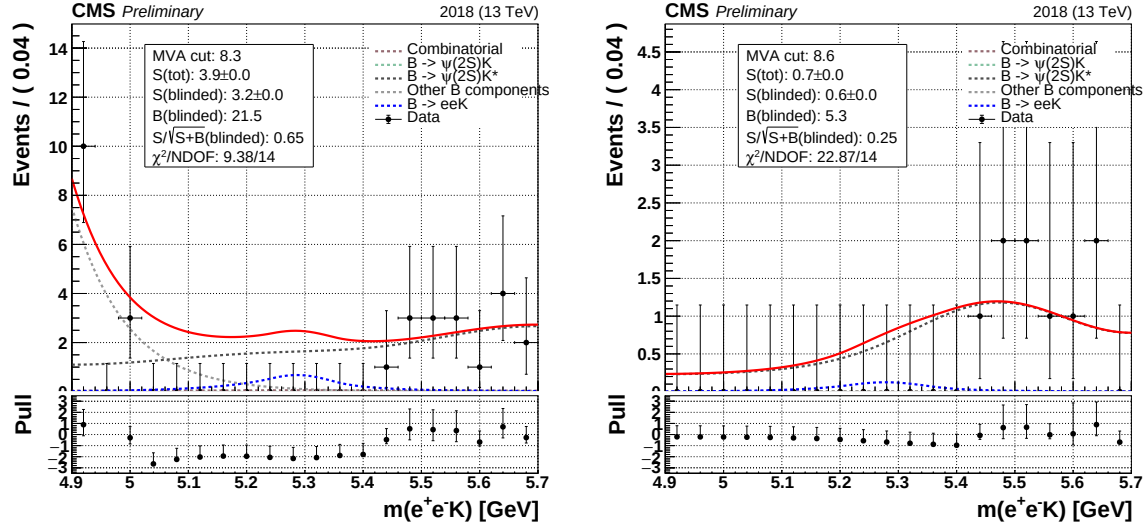


Figure 5.66: Fit to data (top) and pull (bottom) in the high- q^2 bin for the PF-PF (left) and PF-LP (right) category.

5.8.4 Acceptance and efficiency

The acceptance and efficiency are calculated using the signal MC. The efficiencies are factorized as follow:

$$\alpha \times \epsilon = \alpha_{\text{detector}} \times \epsilon_{\text{reco}} \times \epsilon_{\text{nano}} \times \epsilon_{\text{pre-sel}} \times \epsilon_{q^2} \times \epsilon_{\text{anti-}D_0} \times \epsilon_{\text{BDT}}, \quad (5.12)$$

where α_{detector} is the detector acceptance, ϵ_{reco} is the reconstruction efficiency, ϵ_{nano} is the nanoAOD selection efficiency, $\epsilon_{\text{pre-sel}}$ is the pre-selection efficiency, ϵ_{q^2} is the q^2 window efficiency, $\epsilon_{\text{anti-}D_0}$ is the efficiency of the anti- D_0 cut, and ϵ_{BDT} is the efficiency of the BDT. All these efficiencies are defined with respect to the events passing the previous selection step. The detailed definitions are listed below:

- α_{detector} : All the generator-level particles from B decay, i.e. electrons and kaon, are required to satisfy the acceptance requirement of $p_T^{\text{gen}} > 0.5$ GeV and $|\eta^{\text{gen}}| < 2.5$.
- ϵ_{reco} : The reconstructed physics objects are provided by the Particle-Flow and low- p_T electron algorithm, matched to the generator-level daughters, and are required to satisfy the kinematic requirements of $p_T > 0.5$ GeV and $|\eta| < 2.5$.
- ϵ_{nano} : The triplet candidates satisfy a minimal set of pre-selection at nanoAOD production level, as described in Sec. 5.8.1.
- $\epsilon_{\text{pre-sel}}$: The triplet candidates satisfy further pre-selection before feeding into a BDT, as described in Sec. 5.8.1.
- ϵ_{q^2} : The triplet candidates satisfy the $q^2 = m(e^+e^-)^2$ mass window.

- $\epsilon_{\text{anti-}D_0}$: The triplet candidates satisfy anti- D_0 cut of $m(K^+e_{[\rightarrow\pi^-]}^-) > 2.0$ GeV and $m(K_{[\rightarrow\pi^+]}^+e_{[\rightarrow K^-]}^-) > 2.0$ GeV. Note that this cut is highly correlated with q^2 .
- ϵ_{BDT} : The triplet candidates satisfy the final BDT selection.

The corrections provided by several SFs, such as PU SF, FONLL SF, and BDT SF, have been found to cancel out between the low- q^2 bin and the J/ψ bin in the muon channel. A study we conducted revealed that implementing these SFs would result in a 1.14% change in the value of R_K . Thus, we chose not to apply these SFs in the electron channel, but the associated systematic uncertainties will be taken into account.

As mentioned in Sec. 5.7.5, there is a discrepancy between data and MC in the generated $m(\ell^+\ell^-)$ spectrum, which can affect the R_K measurement, as the J/ψ normalization channel does not cancel it out. To adopt a model-independent approach, we redefine the definition of R_K to be restricted to the low- q^2 region only, rather than the entire q^2 range. This eliminates the need to calculate the efficiency of the q^2 window (ϵ_{q^2}), making the measurement independent of the theory.

Table 5.36 and Table 5.37 summarize the corrected efficiencies, with uncertainties reflecting only the statistical uncertainty of the number of MC events.

Table 5.36: Acceptance and efficiency factorized and inclusive for all final states for PF-PF. Uncertainties are only statistical.

Efficiency	$B^+ \rightarrow K^+ e^+ e^-$ (low- q^2)	$B^+ \rightarrow K^+ J/\psi(\rightarrow e^+ e^-)$	$B^+ \rightarrow K^+ \psi(2S)(\rightarrow e^+ e^-)$
α_{detector}	0.5504	0.5828	0.5859
ϵ_{reco}	0.1309	0.1455	0.1593
ϵ_{nano}	0.9366	0.9480	0.9110
$\epsilon_{\text{pre-sel}}$	0.9116	0.8797	0.8727
ϵ_{q^2}	N/A	0.7789	0.655
$\epsilon_{\text{anti-}D_0}$	0.9524	0.8785	0.7916
ϵ_{BDT}	0.5036	0.5121	0.4622
$\alpha \times \epsilon$	0.0295 ± 0.0007	0.0248 ± 0.0003	0.0182 ± 0.0002

Table 5.37: Acceptance and efficiency factorized and inclusive for all final states for PF-LP. Uncertainties are only statistical.

Efficiency	$B^+ \rightarrow K^+ e^+ e^-$ (low- q^2)	$B^+ \rightarrow K^+ J/\psi(\rightarrow e^+ e^-)$	$B^+ \rightarrow K^+ \psi(2S)(\rightarrow e^+ e^-)$
α_{detector}	0.5504	0.5828	0.5859
ϵ_{reco}	0.1594	0.1698	0.1877
ϵ_{nano}	0.9216	0.9276	0.9110
$\epsilon_{\text{pre-sel}}$	0.6445	0.5995	0.5907
ϵ_{q^2}	N/A	0.7641	0.6495
$\epsilon_{\text{anti-}D_0}$	0.9432	0.8844	0.8050
ϵ_{BDT}	0.3067	0.3058	0.2198
$\alpha \times \epsilon$	0.0150 ± 0.0005	0.0108 ± 0.0001	0.0068 ± 0.0001

5.8.5 Checks for the electron channel

In order to test the validity of the analysis approach used in the electron channel, including the invariant mass fits, efficiency calculations, and the use of the BDT, we performed several cross checks as described below.

Data-MC agreement with *sPlot* in the electron channel

The *sPlot* technique [61] is used to examine the agreement of the BDT output between the data and MC. The reconstructed triplet mass serves as the discriminating variable, and it is fitted using the same functions as described in Section 5.8.3. The *sWeights* are then extracted from the fit to estimate the probability of an event being "signal-like" and to perform a statistical background subtraction. The resulting BDT distribution for signal-like events in the J/ψ bin data is compared to the one in MC for the PF-PF and PF-LP categories in Fig. 5.67.

The *sPlot* is applied to the data with a BDT working point of $BDT > 0$, a looser cut w.r.t. the cut used in the final analysis. This working point allows us to spot differences in the distributions while not too loose, including other backgrounds not described by our fit. This working point retains 97% of the signal. The plots show good data-MC agreement. The BDT distribution in data without background subtraction is shown in Fig. 5.68 for reference.

$R_K^{\psi(2S)}$ measurement in the electron channel

$R_K^{\psi(2S),e}$ can be measured by using the signal yield from $B^+ \rightarrow K^+ J/\psi(\rightarrow e^+e^-)$ and $B^+ \rightarrow K^+ \psi(2S)(\rightarrow e^+e^-)$ as an important sanity check for the electron channel. This measurement involves determining the value of $R_K^{\psi(2S),e}$ using the signal yield from the fits

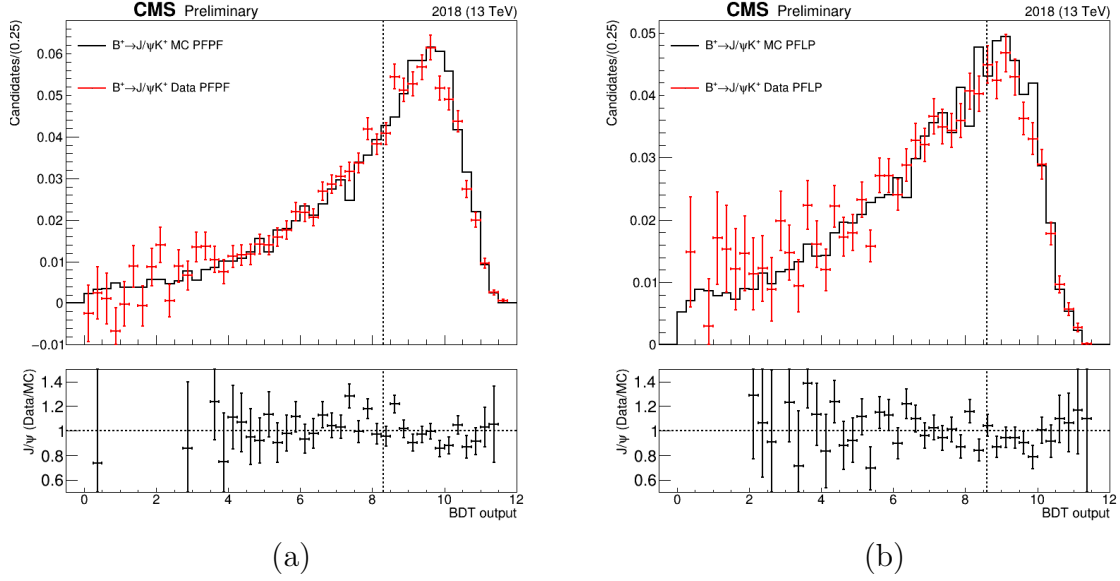


Figure 5.67: Comparison between the BDT output in data and MC for the PF-PF (a) and PF-LP (b) categories in the J/ψ bin. The $sPlot$ technique is used to extract the signal component in data. The dotted line corresponds to the cut applied in the analysis.

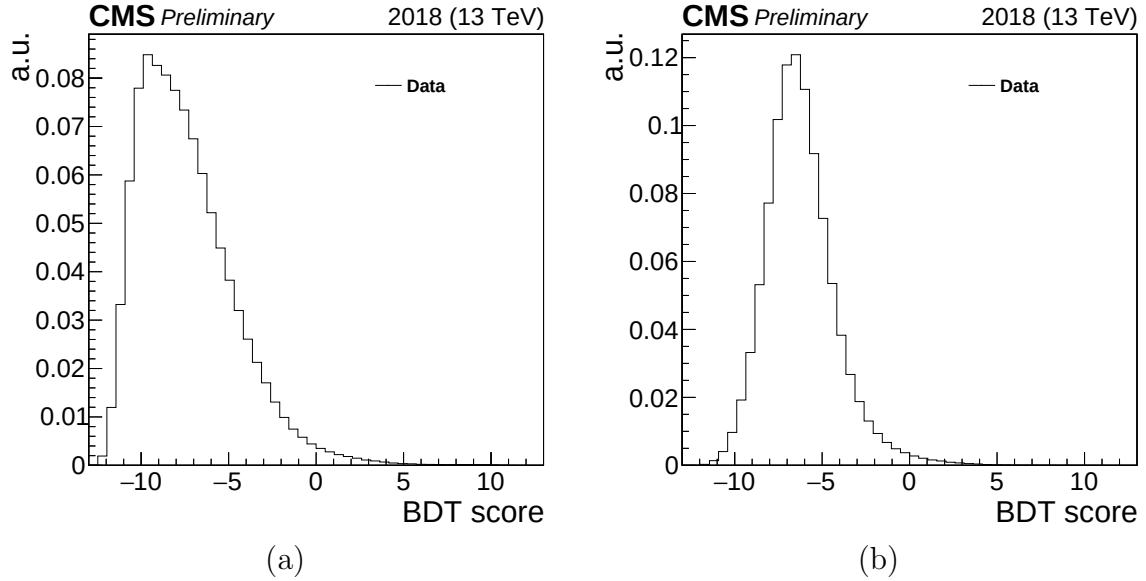


Figure 5.68: BDT distribution in data for the PF-PF (a) and PF-LP (b) category.

and the efficiencies from MC, which is defined as

$$R_K^{\psi(2S),e} = \frac{N_{K^+\psi(2S)(e^+e^-)}}{N_{K^+J/\psi(e^+e^-)}} \frac{\epsilon_{K^+J/\psi(e^+e^-)}}{\epsilon_{K^+\psi(2S)(e^+e^-)}}, \quad (5.13)$$

where N is the number of signal measured in the data, and ϵ is the efficiency.

The measured value is then compared with the SM prediction calculated from the branching fractions:

$$R_K^{\psi(2S),e} = \frac{\mathcal{B}(B^+ \rightarrow K^+ \psi(2S)(\rightarrow e^+ e^-))}{\mathcal{B}(B^+ \rightarrow K^+ J/\psi(\rightarrow e^+ e^-))}, \quad (5.14)$$

where the branching fractions are taken from the latest PDG report [18].

A simultaneous fit on J/ψ and $\psi(2S)$ bins is performed to measure the $R_K^{\psi(2S),e}$. The fits are performed using PF-PF only, PF-LP only and combining both measurement from PF-PF and PF-LP by making use of a joint likelihood. The profile likelihood for the fit is shown in Fig. 5.69. The combined value of $R_K^{\psi(2S),e}$ is measured to be 0.0842 ± 0.0049 . Compared to the PDG value of 0.0812, our measured value is compatible within 1σ . A comparison of the 68% CL of $R_K^{\psi(2S),e}$ with the PDG prediction is shown in Fig. 5.70.

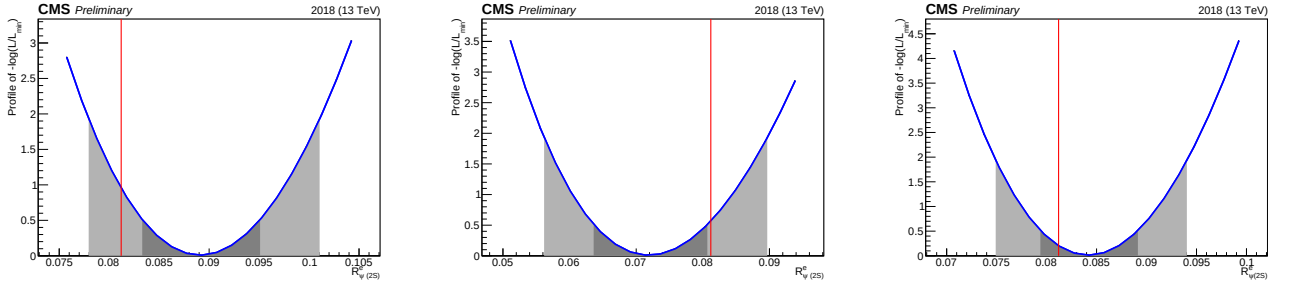


Figure 5.69: Likelihood function from the fit profiled as a function of $R_K^{\psi(2S),e}$. The left, middle, and right shows the profiled likelihood using PF-PF, PF-LP, and combined data.

K^*/K^+ ratio test in the electron channel

The partially reconstructed background coming from the decays of $B^{0/+} \rightarrow K^*(892)^{0/+} \psi(\rightarrow e^+ e^-)$ contribute significantly in the J/ψ and $\psi(2S)$ regions. The amounts of the partially reconstruction background from the fit can be used to compare with the expected yields

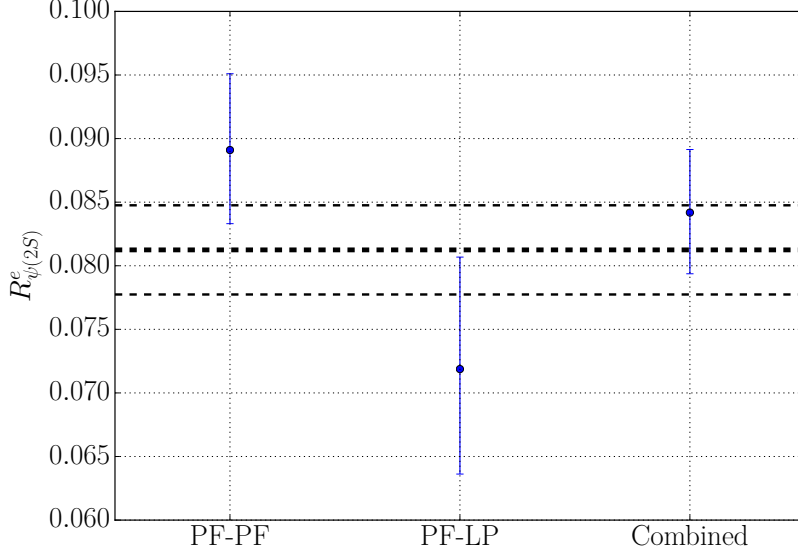


Figure 5.70: A comparison of the 68% CL of $R_K^{\psi(2S),e}$ with the SM prediction (black). The left, middle, and right shows the measurements using PF-PF, PF-LP, and combined data. The combined measurement agrees with the SM prediction within 1σ .

from the branching fractions and efficiencies as a sanity check, which validates the reliability of the mass fit. The ratio of K^* to K^+ , is defined as

$$r_{K^*/K^+} = \frac{N_{K^{*0}\psi} + N_{K^{*+}\psi}}{N_{K^+\psi}}, \quad (5.15)$$

where N_{q^2} is the number of signal or partially reconstructed background measured in the q^2 bin from the mass fit. This ratio is compared to the ratio of branching fractions

$$r_{K^*/K^+} = \frac{\mathcal{B}(B^0 \rightarrow K^{*0}\psi(\rightarrow e^+e^-)) \cdot \epsilon_{K^{*0}\psi} + \mathcal{B}(B^+ \rightarrow K^{*+}\psi(\rightarrow e^+e^-)) \cdot \epsilon_{K^{*+}\psi}}{\mathcal{B}(B^+ \rightarrow K^+\psi(\rightarrow e^+e^-)) \cdot \epsilon_{K^+\psi}}, \quad (5.16)$$

where ϵ is the efficiency calculated using MC.

In Fig. 5.71, we present the measured r_{K^*/K^+} compared to the expected r_{K^+/K^*} values in the J/ψ and $\psi(2S)$ regions for various BDT working points. No significant deviation from

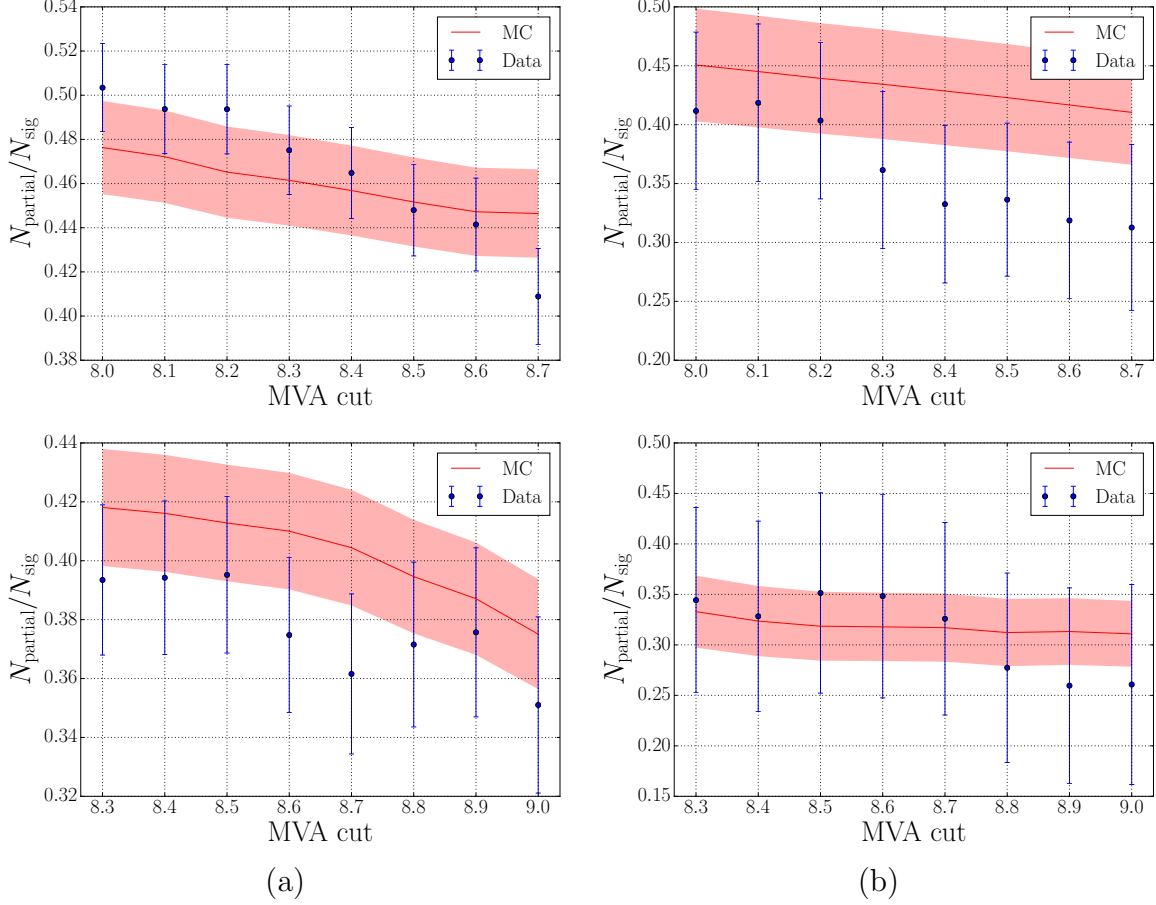


Figure 5.71: K^*/K^+ ratio comparison in the J/ψ bin (a), and $\psi(2S)$ bin (b), with PF-PF (top) and PF-LP (bottom) as a function of BDT cut.

the expected values is observed.

To investigate the impact of the K^*/K^+ ratio on the signal yield, we conducted a test where the yield of partially reconstructed background was fixed to the expected value calculated using the efficiency during the fit, instead of allowing it to float. We then recomputed the $R_K^{\psi(2S),e}$ value to assess any effect on the measurement.

The computed $R_K^{\psi(2S),e}$ value with the fixed partially reconstructed yield is 0.0834 ± 0.0047 , which is compared to the floating background yield of $R_K^{\psi(2S),e} = 0.0842 \pm 0.0049$. The values have a negligible difference. The $R_K^{\psi(2S),e}$ values with fixed background yields are summarized in Fig. 5.72. This exercise demonstrates that the fits in the lower sideband have a minimal

impact on the measurement.

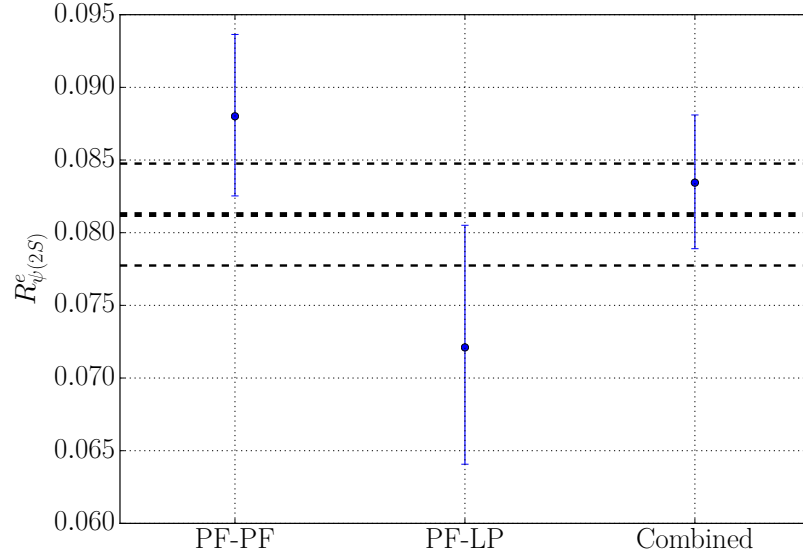


Figure 5.72: A comparison of the 68% CL of $R_K^{\psi(2S),e}$ with the SM prediction (black). The yield of the partially reconstructed background in the fits are fixed to the expected value. The left, middle, and right shows the measurements using PF-PF, PF-LP, and combined data. The fits in the lower sideband have negligible effect on the measurement.

Checks for mass sculpting in the electron channel

Mass sculpting refers to the effect where the BDT learns the mass from the input variables, which creates a fake signal peak in the measurement. To avoid this, the angular variables (η and ϕ) are excluded from the input variables so that the BDT cannot reconstruct the full four momentum. Furthermore, to eliminate any correlation with the mass, all transverse momenta are normalized by the triple mass $m(K^+e^+e^-)$ to make them dimensionless.

In order to check for any signs of mass sculpting with our BDT, events from three different types of samples are used: combinatorics in same-sign data, combinatorics in $B^+ \rightarrow K^+ J/\psi(\rightarrow e^+e^-)$ MC sample, and combinatorics in the generic B decays MC sample. Same-sign data events comprise of two electrons with the same charge sign. Combinatorics in $B^+ \rightarrow K^+ J/\psi(\rightarrow e^+e^-)$ MC samples consist of triplets in the MC sample where the electrons are gen-matched but the track is unmatched with any generated kaon. Combinatorics in the generic B decays MC sample are triplets in the generic B MC sample that exclude all signal and background decays ($B^+ \rightarrow K^+ J/\psi(e^+e^-)$, $B^{(0/+)} \rightarrow K^{*(0/+)} J/\psi(e^+e^-)$, $B^+ \rightarrow K^+ \psi(2S)(e^+e^-)$, $B^{(0/+)} \rightarrow K^{*(0/+)} \psi(2S)(e^+e^-)$, etc), as well as events in which electrons originate from J/ψ or $\psi(2S)$. The particles composed the triplet are expected to be uncorrelated. The $m(K^+e^+e^-)$ distributions from these samples are expected to be falling smoothly, without any visible mass peaks. Any mass structures observed in these samples can be a sign of mass sculpting.

Figures A.0.1 and A.0.2 show the mass distribution using same-sign data for PF-PF and PF-LP, respectively. Figures A.0.3 and A.0.4 show the mass distribution using combinatorics in $B^+ \rightarrow K^+ J/\psi(\rightarrow e^+e^-)$ MC sample for PF-PF and PF-LP, respectively. Figures A.0.5

and A.0.6 show the mass distribution using combinatorics in the generic B decays MC sample for PF-PF and PF-LP, respectively. No visible mass peaks are observed. There are also no correlated bumps between three different samples. Thus, the BDT doesn't learn the mass during training.

Checks for misidentified background in the electron channel

The latest result on R_K from LHCb [29] shows a significant contribution from a "misidentified background" in the low- q^2 region. In this process, hadrons are misidentified as electrons, which leads to a peaking background from fully reconstructed decays of $B^+ \rightarrow K^+ h_1 h_2$, where $h_{1,2}$ represent kaons or pions. To determine the impact of this background for this analysis, we used $B^+ \rightarrow \bar{D}^0(\rightarrow K^+ \pi^-) \pi^+$ MC sample as a proxy to assess how well we can distinguish this type of background.

The distributions of the $B^+ \rightarrow \bar{D}^0(\rightarrow K^+ \pi^-) \pi^+$ MC for PF-PF after preselection are presented in Fig. B.0.1. The peaking structure of this background can be seen in the mass distribution. The electron ID can effectively veto this type of background. The anti- D^0 variable $\min(m(K^+ e^- [\rightarrow \pi^-]), m(K^+ [\rightarrow \pi^+] e^-_{[\rightarrow K^-]}))$ can effectively eliminate this decay channel since this decay originates from the resonance of D^0 . However, in this study, we only used this specific decay channel as a proxy to examine the effectiveness of vetoing $B^+ \rightarrow K^+ h_1 h_2$ decays. Since some of the decays may not originate from D^0 resonance, we did not apply an anti- D^0 cut in this study. The acceptance and efficiency of $B^+ \rightarrow \bar{D}^0(\rightarrow K^+ \pi^-) \pi^+$ for PF-PF after pre-selection is 7.7×10^{-5} .

Figure B.0.2 shows the distributions of the $B^+ \rightarrow \bar{D}^0(\rightarrow K^+ \pi^-) \pi^+$ MC for PF-LP after preselection. In the preselection stage for the PF-LP category, an electron ID cut of $\text{ID}(e_1) > -2$ and $\text{ID}(e_2) > 0$ has already been applied. Figure B.0.3 shows another set of plots without the electron ID cut to demonstrate the effect of electron ID. The outcomes reveal that our electron ID can effectively veto the misidentified background. The acceptance and efficiency of PF-LP after preselection, without applying the electron ID cut, are 1.7×10^{-4} .

and 1.1×10^{-5} , respectively.

The $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-)\pi^+$ MC sample includes 2173669 events that passed the HLT_Mu7_IP4 requirement. After applying the nominal BDT cut (8.3 for PF-PF and 8.6 for PF-LP), there are only 2 and 3 reconstructed events for PF-PF and PF-LP, respectively. This corresponds to an acceptance and efficiency of 9.2×10^{-7} for PF-PF and 1.3×10^{-6} for PF-LP. The mass distributions after the nominal BDT cut for PF-PF and PF-LP are presented in Fig. B.0.4.

The number of $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-)\pi^+$ in data is estimated using the signal yield of $B^+ \rightarrow K^+J/\psi(\rightarrow e^+e^-)$, with the following equation,

$$N^{\bar{D}^0\pi^+} = N^{K^+J/\psi} \cdot \frac{\mathcal{B}(B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-)\pi^+) \cdot \epsilon_{\bar{D}^0\pi^+}}{\mathcal{B}(B^+ \rightarrow K^+J/\psi(\rightarrow e^+e^-)) \cdot \epsilon_{K^+J/\psi}}, \quad (5.17)$$

where $\mathcal{B}$ is the branching fraction, and ϵ is the efficiency.

After calculations, the expected number of $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-)\pi^+$ decays in the entire q^2 region, without applying an anti- D^0 cut, is 0.6 for PF-PF and 0.8 for PF-LP. This indicates that the contribution of this background is insignificant. Other decay modes, such as $B^+ \rightarrow K^+\pi^+\pi^-$, $B^+ \rightarrow K^0(\rightarrow K_S^0 \rightarrow \pi^+\pi^-)\pi^+$, and $B^+ \rightarrow \pi^+\pi^+\pi^-$, may produce misidentified background. The expected amount of these backgrounds can be estimated from the ratio of branching fractions between each decay and $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-)\pi^+$. The expected number of background events is summarized in Table 5.38. The estimated number of events from misidentified background in the entire q^2 range is less than 1.

The primary concern is the impact of this background on the low- q^2 region, and it is important to determine the efficiency of the low- q^2 window in conjunction with the anti- D^0 cut. To estimate the efficiency for resonant channels (decays involving D^0), we utilize

the $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-)\pi^+$ MC. However, due to the limited statistics in MC events that pass the nominal BDT cut, we assume that q^2 and the BDT score are independent, and we calculate the low- q^2 window efficiency before the BDT cut by using MC events that passed the preselection as the denominator. The efficiency for the low- q^2 window is determined to be 0.273 ± 0.034 for PF-PF and 0.458 ± 0.102 for PF-LP. Upon applying the anti- D^0 cut in addition to the low- q^2 window, the efficiency decreases further to 0.238 ± 0.033 and 0.333 ± 0.096 for PF-PF and PF-LP, respectively.

To estimate the efficiency of the low- q^2 window and anti- D^0 cut for non-resonant channels (decays without D^0), we generated a small private $B^+ \rightarrow K^+\pi^+\pi^-$ MC sample. However, the sample had a very small number of reconstructed events, so we used the generated particles to estimate the efficiency of the low- q^2 window and anti- D^0 cut instead. The $m(e^+e^-)$ and anti- D^0 distributions of the generated events from the $B^+ \rightarrow K^+\pi^+\pi^-$ MC are shown in Fig. B.0.5. We calculated the efficiency of the low- q^2 window and anti- D^0 cut to be 0.285.

By combining the efficiency of the low- q^2 window and the anti- D^0 cut with the numbers in Table 5.38, we estimate that there will be around 0.3 events from misidentified background in the low- q^2 region for each of the categories. These numbers are summarized in Table 5.39. We conclude that the misidentified background does not contaminate our signal.

Process	Branching fraction	Yield for PF-PF	Yield for PF-LP
$B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-)\pi^+$	1.9×10^{-4}	0.6	0.8
$B^+ \rightarrow K^+\pi^+\pi^-$	5.1×10^{-5}	0.2	0.2
$B^+ \rightarrow K^+K^-\pi^+$	3.4×10^{-5}	0.1	0.1
$B^+ \rightarrow K^0(\rightarrow K_S^0 \rightarrow \pi^+\pi^-)\pi^+$	8.0×10^{-6}	0.03	0.03
$B^+ \rightarrow \pi^+\pi^+\pi^-$	1.5×10^{-5}	0.04	0.06

Table 5.38: Summary of the expected number of misidentified background in whole q^2 region without applying anti- D^0 cut.

Process	Yield for PF-PF	Yield for PF-LP
$B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-)\pi^+$	0.14	0.27
$B^+ \rightarrow K^+\pi^+\pi^-$	0.06	0.06
$B^+ \rightarrow K^+K^-K^+$	0.04	0.04
$B^+ \rightarrow K^0(\rightarrow K_S^0 \rightarrow \pi^+\pi^-)\pi^+$	0.01	0.01
$B^+ \rightarrow \pi^+\pi^+\pi^-$	0.01	0.01
Total	0.26	0.39

Table 5.39: Summary of the expected number of misidentified background in the low- q^2 region with anti- D^0 cut applied.

5.8.6 Systematic uncertainties for the electron channel

The R_K measurement is largely limited by the statistics in the electron channel. The systematic uncertainty studies are limited to those which may give a large contribution to the measurement.

BDT cut variation stability

The optimal point for the BDT cut was selected by maximizing the significance in the low q^2 bin. The stability of the measurement was investigated by extending the range of the BDT cut such that variation of significance relative to the optimal point is around 10%.

The following quantity as a function of BDT cut is used to evaluate the BDT stability

$$R_K^{e, \text{pseudo}} = \frac{N_{K+J/\psi}^{e, \text{weighted}}}{N_{K+J/\psi}^e} \cdot \frac{\epsilon_{\text{BDT}}^{\text{low-}q^2}}{\epsilon_{\text{BDT}}^{J/\psi}}, \quad (5.18)$$

where $N_{K+J/\psi}^e$ is the signal yield in J/ψ bin, $N_{K+J/\psi}^{e, \text{weighted}}$ is the signal yield in J/ψ bin after reweighting, and ϵ_{BDT} is the BDT efficiency.

The reason for computing this quantity is that there are no sufficient statistics in the low- q^2 bin to estimate stability. Instead, the stability is estimated using the J/ψ bin. The subleading electron p_T is expected to be the main difference between $B^+ \rightarrow K^+ e^+ e^-$ and $B^+ \rightarrow K^+ J/\psi (\rightarrow e^+ e^-)$. Thus, the subleading electron p_T distribution of $B^+ \rightarrow K^+ J/\psi (\rightarrow e^+ e^-)$ was reweighted to match that of $B^+ \rightarrow K^+ e^+ e^-$, and a refit was performed on the weighted data to obtain $N_{K+J/\psi}^{e, \text{weighted}}$. This procedure allows for the use of J/ψ bin information to mimic non-resonant events in the data.

The reweighting is done at the nominal BDT working point using MC. Figure 5.74 shows

an example of reweighting the subleading electron p_T distribution of $B^+ \rightarrow K^+ J/\psi(\rightarrow e^+ e^-)$ to match that of $B^+ \rightarrow K^+ e^+ e^-$ for PF-PF and PF-LP categories. The resulting weights are applied to the data, and a refit is performed using the reweighted data. An example of the fit after reweighting for PF-PF and PF-LP categories is shown in Fig. 5.73.

The stability of $R_K^{e, \text{pseudo}}$ as a function of the BDT cut is shown in Fig. 5.75 for PF-PF and PF-LP categories. The maximum relative variation of $R_K^{e, \text{pseudo}}$ in the BDT cut range is 2.05% for the PF-PF category and 4.70% for the PF-LP category.

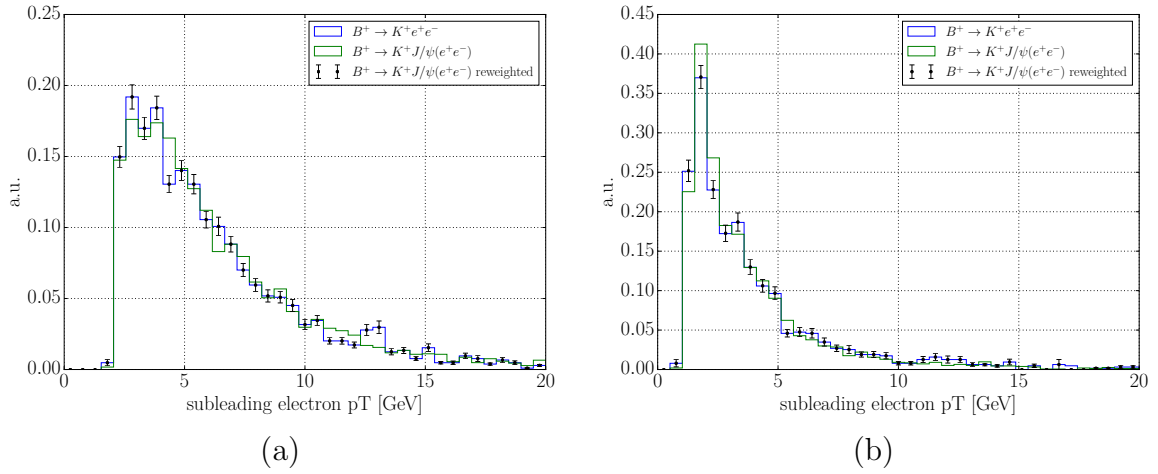


Figure 5.73: An example of reweighting the subleading electron p_T distribution of $B^+ \rightarrow K^+ J/\psi(\rightarrow e^+ e^-)$ to match that of $B^+ \rightarrow K^+ e^+ e^-$ for the PF-PF (a) and PF-LP (b) category.

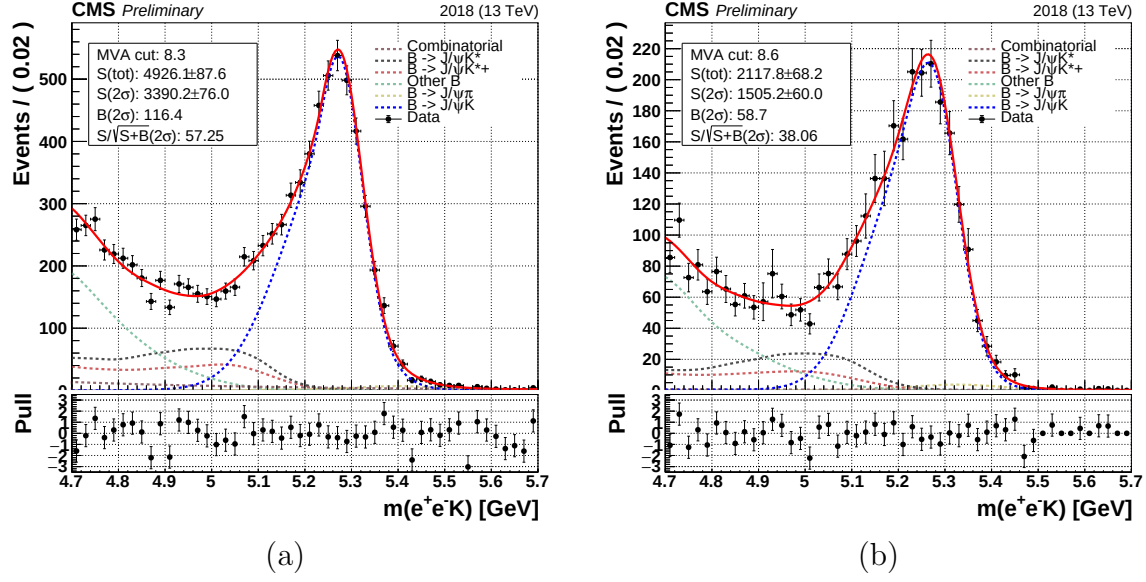


Figure 5.74: Fit after applying weight to the data according to the subleading electron p_T for the PF-PF (a) and PF-LP (b) category in the J/ψ bin.

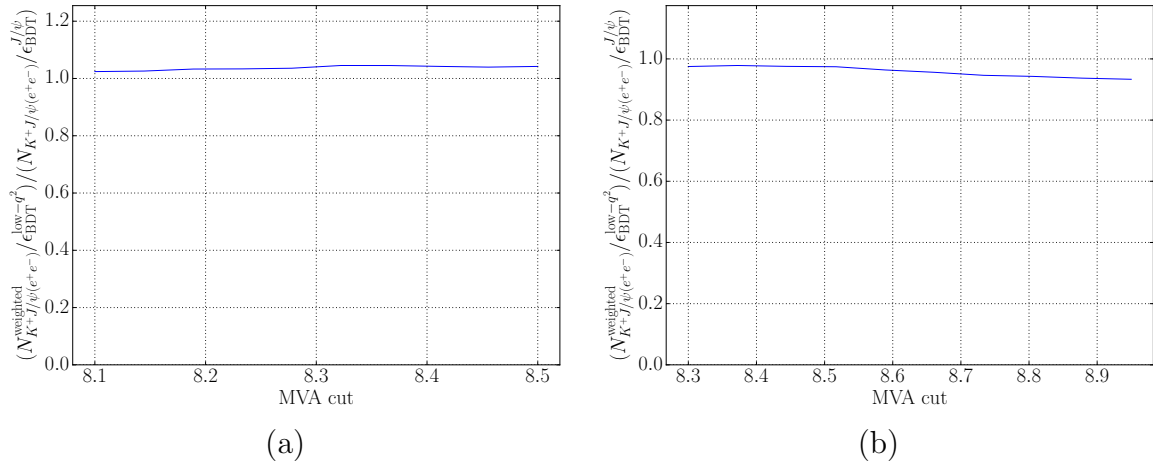


Figure 5.75: $R_K^{e, \text{pseudo}}$ as a function of the BDT cut for the PF-PF (a) and PF-LP (b) category in the J/ψ bin.

Pile Up Scale Factor uncertainty

The method for calculating the pile up SF for electrons is the same as that used for the muon channel. To determine the uncertainty of this quantity, the effect of varying $\pm 1\sigma$ on R_K^e is assessed to estimate the error. By calculating new values of R_K^e with these variations, the maximum difference from the nominal value is taken as the uncertainty. The percentage uncertainties for the PF-PF and PF-LP categories of the PU SF are found to be 0.44% and 0.017%, respectively.

Trigger path stability

In this analysis, the electron channel uses various L1 and HLT paths with different p_T thresholds for the triggering muon, resulting in different kinematics. As the kinematics between the tag side and the probe side are weakly correlated, the kinematics of B^+ candidates can be different according to the trigger paths. However, since we measure R_K in double ratio, the effect of trigger paths is mostly eliminated. This systematic investigation aims to evaluate the extent to which the trigger effect is eliminated and to estimate the associated systematic uncertainty. The study is divided into two parts, one using MC simulations and the other using data to validate the findings.

For the MC checking, we compute a ratio $R_{\text{trigger}}^{\text{MC}}$ across different HLT paths, defined as

$$R_{\text{trigger}}^{\text{MC}}(\text{HLT}) = \frac{N_{\text{non-resonant}}(\text{HLT and L1_SingleMu7er1p5})/N_{\text{non-resonant}}^{\text{total}}}{N_{J/\psi}(\text{HLT and L1_SingleMu7er1p5})/N_{J/\psi}^{\text{total}}}, \quad (5.19)$$

where $N(\text{HLT and L1_SingleMu7er1p5})$ is the number of MC events that passed that particular HLT and L1.SingleMu7er1p5, and N^{total} is the total number of MC events as a

normalization factor and is to get the uncertainty for $R_{\text{trigger}}^{\text{MC}}$.

If the ratio $R_{\text{trigger}}^{\text{MC}}$ is stable across different HLT paths, then it implies that all the systematic uncertainty due to the trigger paths is completely cancelled out and does not enter to the measurement of R_K . $R_{\text{trigger}}^{\text{MC}}$ is analyzed for various q^2 regions and selection working points, as presented in Fig. 5.76 and Fig. 5.77 for PF-PF and PF-LP, respectively. The outcomes are summarized in Table 5.40 for PF-PF and Table 5.41 for PF-LP. Systematic uncertainties of 1% for PF-PF and 4% for PF-LP are proposed to cover the trend observed in the study.

Table 5.40: Summary of $R_{\text{trigger}}^{\text{MC}}$ in different q^2 region at different working points for PF-PF. The proposed systematic uncertainty is 1%.

q^2 region	Working point	Central value of $R_{\text{trigger}}^{\text{MC}}$	Proposed systematic	Percentage
full- q^2	preselection	1.010	0.005	0.5%
	loose BDT WP	0.955	0.005	0.5%
	tight BDT WP	0.900	0.005	0.5%
low- q^2	preselection	0.256	0.001	0.4%
	loose BDT WP	0.237	0.001	0.4%
	tight BDT WP	0.238	0.001	0.4%
high- q^2	preselection	0.252	0.003	1.0%
	loose BDT WP	0.224	0.002	1.0%
	tight BDT WP	0.182	0.002	1.0%

Table 5.41: Summary of $R_{\text{trigger}}^{\text{MC}}$ in different q^2 region at different working points for PF-LP. The proposed systematic uncertainty is 4%.

q^2 region	Working point	Central value of $R_{\text{trigger}}^{\text{MC}}$	Proposed systematic	Percentage
full- q^2	preselection	0.990	0.005	0.5%
	loose BDT WP	0.920	0.005	0.6%
	tight BDT WP	0.855	0.010	1.0%
low- q^2	preselection	0.258	0.005	2.0%
	loose BDT WP	0.247	0.002	0.8%
	tight BDT WP	0.269	0.002	0.7%
high- q^2	preselection	0.215	0.010	4.0%
	loose BDT WP	0.167	0.001	0.6%
	tight BDT WP	0.108	0.004	4.0%

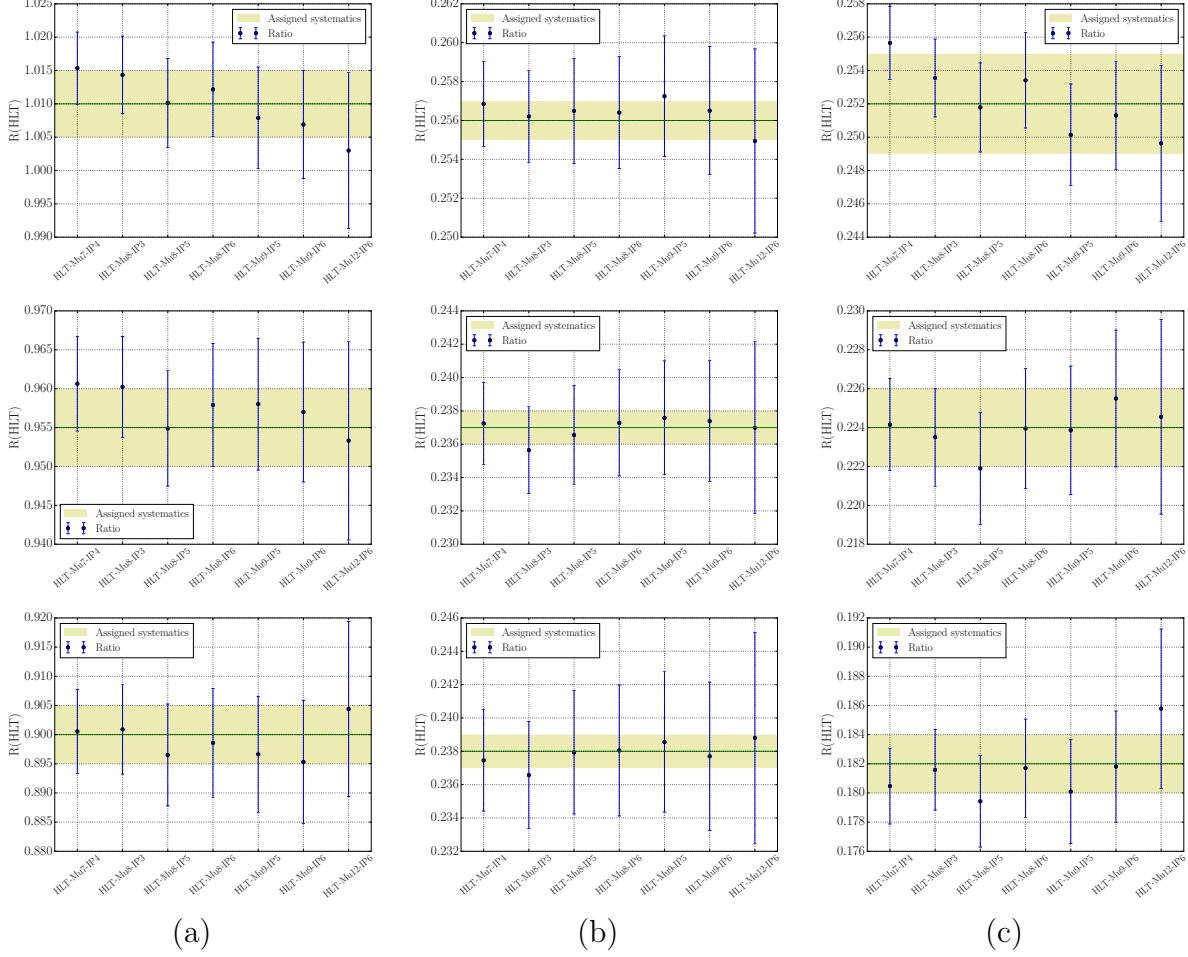


Figure 5.76: $R_{\text{trigger}}^{\text{MC}}$ of different HLT paths for PF-PF in full- q^2 (a), low- q^2 (b) and high- q^2 (c). Three rows correspond to working point at preselection (top), loose BDT WP (middle) and tight BDT WP (bottom).

To examine whether the MC is accurately modeled, it is necessary to verify it against data. The J/ψ region is used for this comparison as it provides sufficient statistics. In actual data collection, various L1 and HLT paths are sequentially tuned on as the instantaneous luminosity decreases, as described in Sec. 5.3. For instance, at the beginning of a fill, L1_Mu12er1p5 and HLT_Mu12_IP6 are initially tuned on. However, towards the end of the fill, looser trigger settings of L1_SingleMu7er1p5 and HLT_Mu7_IP4 are tuned on while the previously activated triggers are still active. Therefore, there exist numerous trigger configurations, making it impossible to check them all. To address this, an inclusive representation

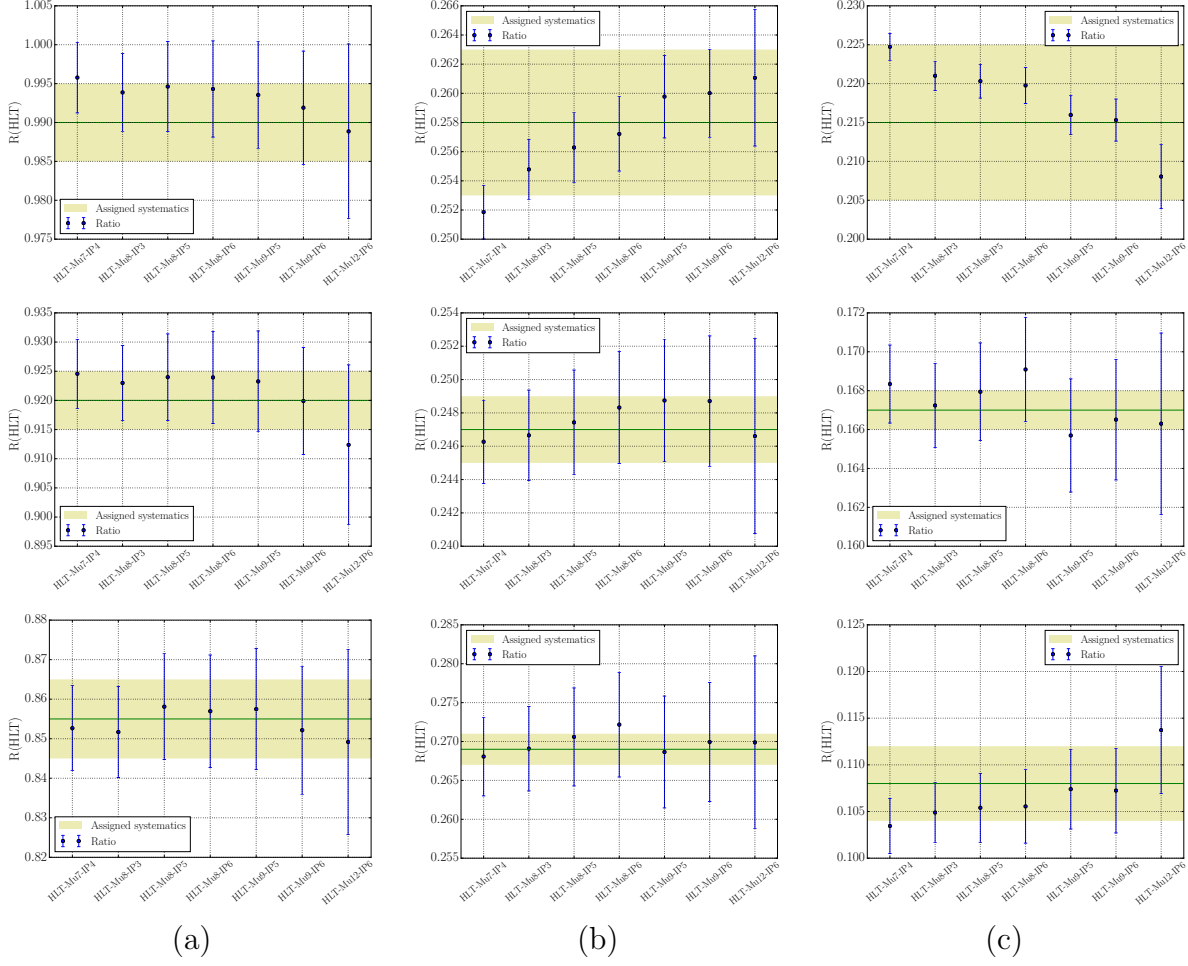


Figure 5.77: $R_{\text{trigger}}^{\text{MC}}$ of different HLT paths for PF-LP in full- q^2 (a), low- q^2 (b) and high- q^2 (c). Three rows correspond to working point at preselection (top), loose BDT WP (middle) and tight BDT WP (bottom).

is employed in this study.

For the MC, the events are required to pass L1.SingleMu7er1p5 and HLT_Mu7_IP4.

While for the data, only lumi sections, where L1.SingleMu7er1p5, HLT_Mu7_IP4, HLT_Mu8_IP5,

HLT_Mu9_IP5 and HLT_Mu12_IP6 are tuned on, are considered. For the data, we compute

a ratio R_{data} across different HLT paths, defined as

$$R_{\text{trigger}}^{\text{data}}(\text{HLT}) = \frac{\epsilon_{J/\psi}^{\text{data}}}{\epsilon_{J/\psi}^{\text{MC}}}, \text{ with } \epsilon_{J/\psi} = \frac{N_{J/\psi}(\text{HLT and HLT_Mu7_IP4 and L1.SingleMu7er1p5})}{N_{J/\psi}(\text{HLT_Mu7_IP4 and L1.SingleMu7er1p5})}, \quad (5.20)$$

where $N_{J/\psi}(\text{triggers})$ is the number of J/ψ events that passed the triggers. The number of events in data is measured by fitting of $m(K^+e^+e^-)$ and that in MC is obtained by counting the number of events. $R_{\text{trigger}}^{\text{data}} = 1$ indicates no data-MC discrepancy.

Figure 5.78 shows the $R_{\text{trigger}}^{\text{data}}$ across different HLT. No significant deviation from unity is observed. We propose not to add additional systematic uncertainty from $R_{\text{trigger}}^{\text{data}}$.

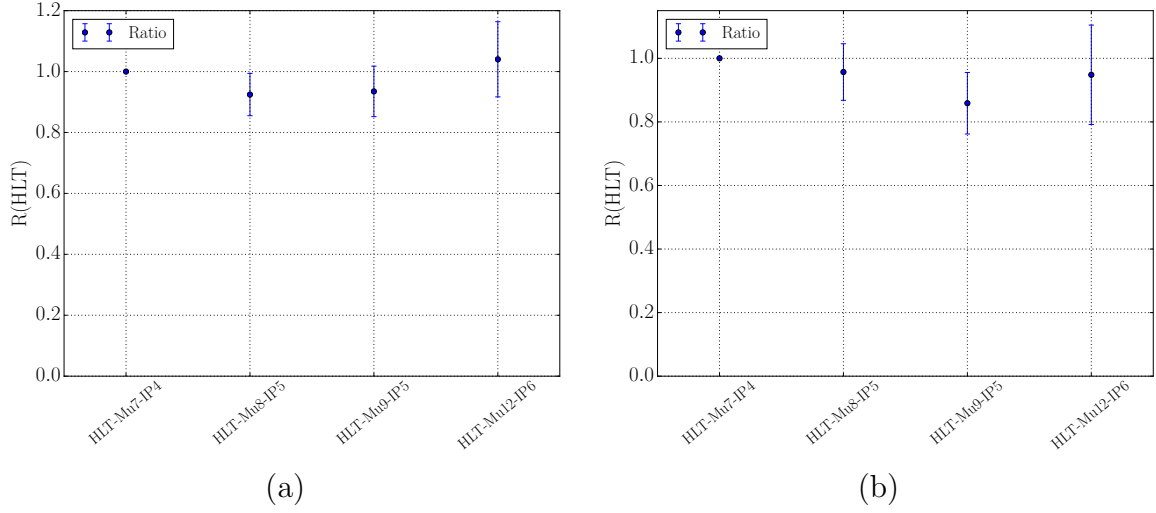


Figure 5.78: $R_{\text{trigger}}^{\text{data}}$ of different HLT paths for PF-PF (a) and PF-LP (b). No significant deviation from unity is observed. The uncertainty from the normalization trigger path HLT_Mu7_IP4 is considered and added to the uncertainties of other paths.

BDT Scale Factor uncertainty

As illustrated in Sec 5.8.5, the BDT output is well described by the simulation using the *sPlot* technique. To account for the residual BDT differences between data and MC, we evaluate the efficiency of the nominal BDT cut relative to $BDT > 0$ in data, determining the signal yield with the fit, and in MC via a cut and count method.

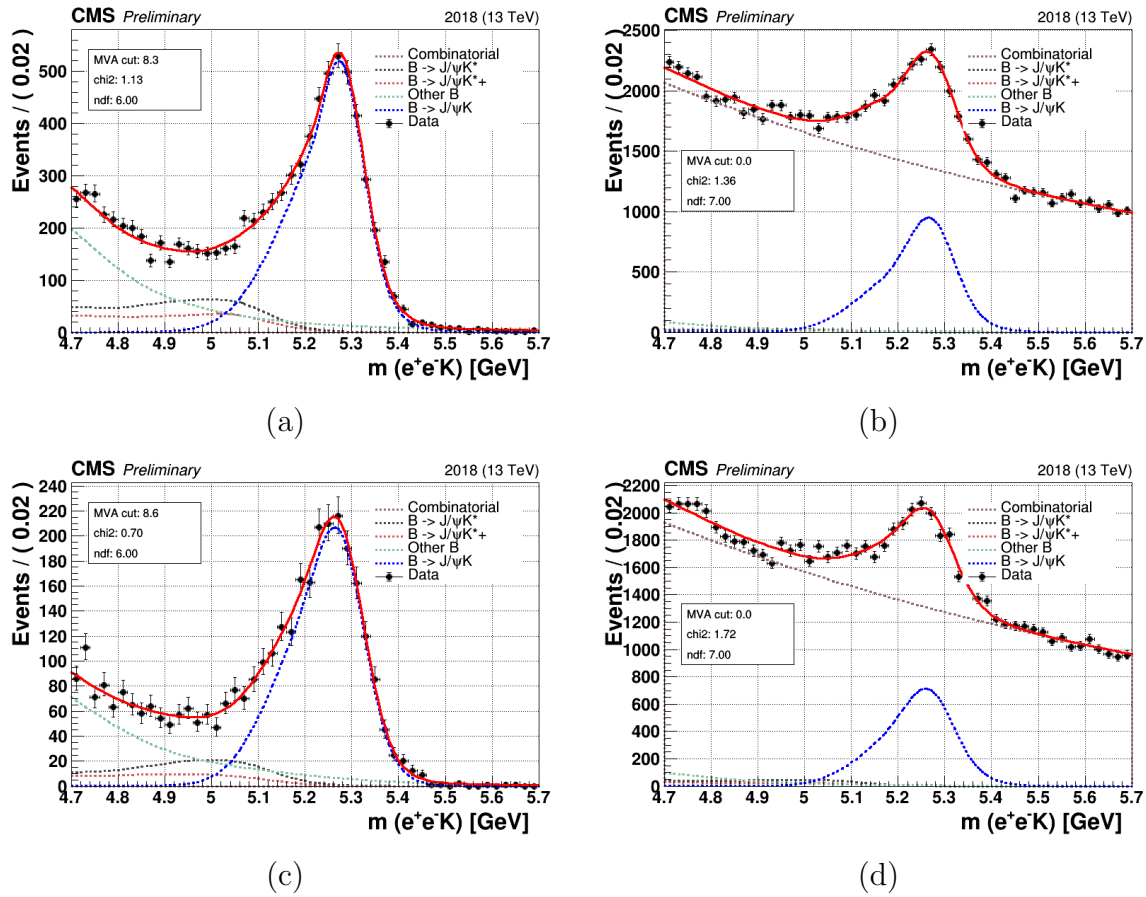


Figure 5.79: Fits to the B-candidate mass distribution in the J/ψ bin for PF-PF (top) and PF-LP (bottom) category with nominal BDT cut (left) and $BDT > 0$ (right).

Fig. 5.80 shows the BDT output distribution in the simulation in the low- q^2 bin and in the J/ψ bin after reweighting for the two categories. In the PF-PF category, after weighting, the BDT cut efficiency is 0.518 ± 0.003 in the low- q^2 bin and 0.514 ± 0.003 in the J/ψ bin. In

contrast, it is 0.333 ± 0.003 and 0.325 ± 0.002 in the PF-LP category. The level of non-closure determined using this method serves as an additional systematic in the low- q^2 bin.

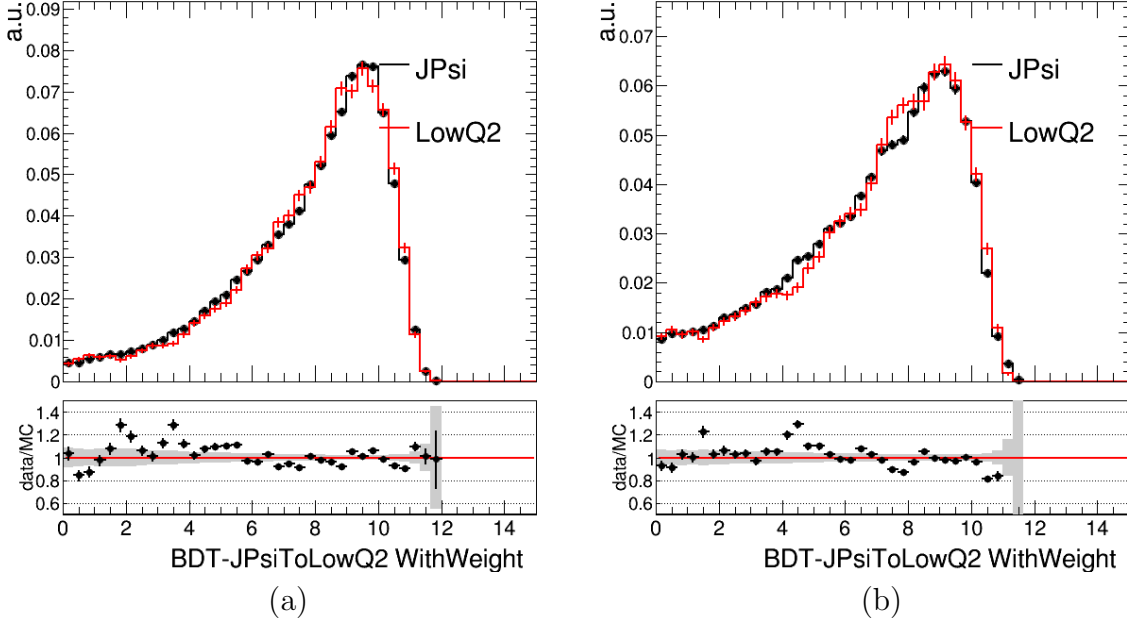


Figure 5.80: Simulated BDT spectrum in the low- q^2 bin and in the J/ψ bin after reweighting, for for PF-PF (left) and PF-LP (right) category.

The weights are subsequently implemented to the data, followed by a refit with the reweighted data. Figure 5.81 shows the fits after reweighting for both the PF-PF and PF-LP categories. Using the weighted spectra, we determine the efficiency ratio in data and MC simulations for the low- q^2 bin is:

$$\varepsilon(\text{data}, J/\psi, \text{PF} - \text{PF})/\varepsilon(\text{MC}, J/\psi, \text{PF} - \text{PF}) = (1.006 \pm 0.012) \quad (5.21)$$

$$\varepsilon(\text{data}, J/\psi, \text{PF} - \text{LP})/\varepsilon(\text{MC}, J/\psi, \text{PF} - \text{LP}) = (0.943 \pm 0.019). \quad (5.22)$$

We determine the efficiency ratio $R\epsilon = \epsilon(\text{low-}q^2)/\epsilon(J/\psi)$ and calculate the corresponding

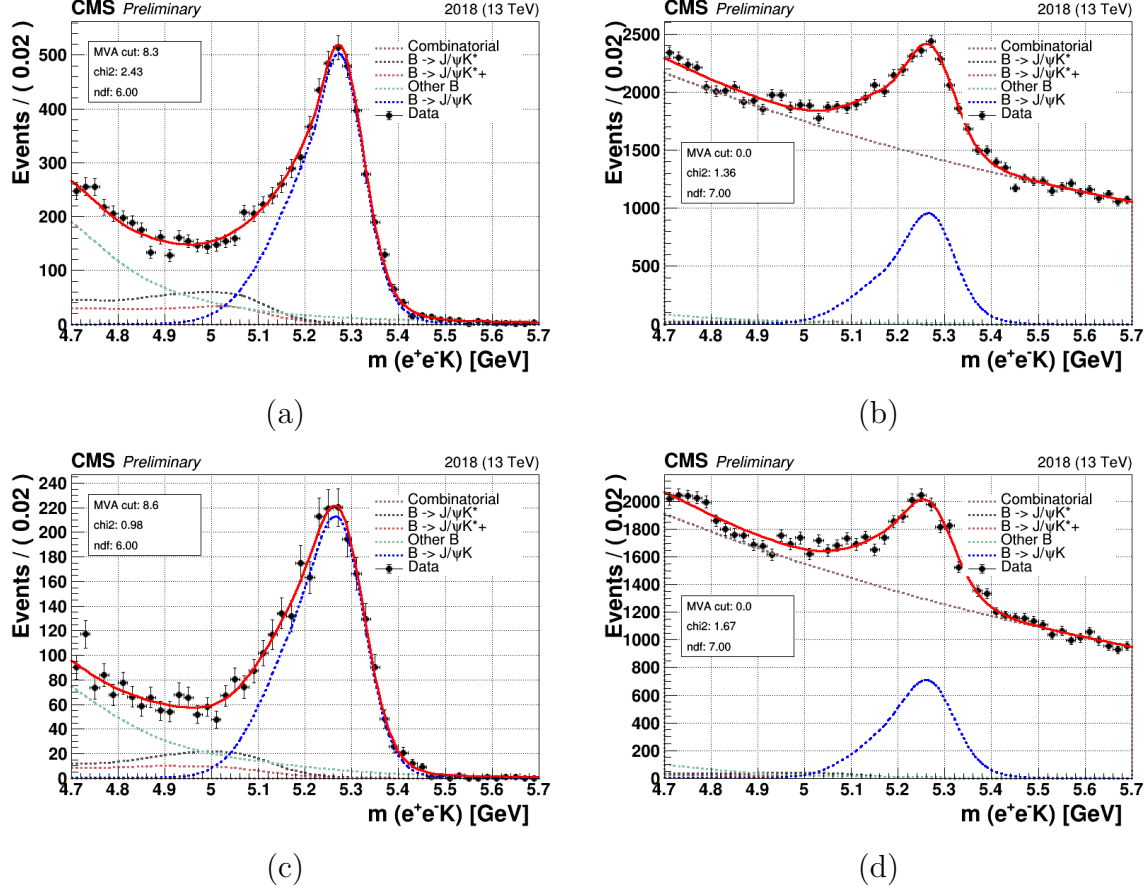


Figure 5.81: Fits to the B -candidate mass distribution in the J/ψ bin after weighting as discussed in the text to match the low- q^2 bin, for the for PF-PF (top) and PF-LP (bottom) category with nominal BDT cut (left) and BDT > 0 (right).

nominal values and $\pm 1\sigma$ uncertainties. We use the efficiency ratio in data and MC simulations for the J/ψ bin as the total uncertainty. For the low- q^2 bin, we consider the efficiency ratio plus the non-closure of the reweighting procedure on MC. The maximum difference between $|R\epsilon(nominal) - R\epsilon(+1\sigma)|$ and $|R\epsilon(nominal) - R\epsilon(-1\sigma)|$ is taken into account. The percentage uncertainty of the BDT data/MC systematics is found to be $\Delta R\epsilon/R\epsilon = 0.7\%$ in the PF-PF category and 1.5% in the PF-LP category.

Choice of parameterization in the fit

The choice of fitting function can have an impact on the extracted signal yield and, thus, the measurement. To estimate the effect of this choice on the measurement, alternative fitting functions are used and the difference in the signal yield is evaluated. The alternative fitting functions are chosen to cover a wide range of possible functional forms for the signal and background. The study is done in J/ψ bin and the low- q^2 bin.

For the J/ψ bin, Fig. D.0.1 and D.0.2 show the shapes of the alternative fitting functions for PF-PF and PF-LP, respectively. The fits with the alternative functions are shown in Fig. D.0.3 and D.0.4. The impact of the choice of fitting function in the J/ψ bin is summarized in Table 5.42. The total systematic uncertainties concerning the choice of fit parameterization J/ψ bin for PF-PF and PF-LP are 0.8% and 1.5%, respectively.

The study in the low- q^2 bin is conducted before unblinding. We generate MC toys using the background fit function from the sideband, and then add a signal peak with yields equal to the expected yield using the signal fit function. To estimate the systematics, we average the differences of signal yields from the fit over 1000 MC toys. Figure D.0.5 shows an example of an MC toy generated for this study in the low- q^2 region. Figures D.0.6 and D.0.7 show examples of the shape of the alternative fitting templates for PF-PF and PF-LP, respectively, using one of the MC toys. Figures D.0.8 and D.0.9 show examples of the fit with alternative fitting templates for PF-PF and PF-LP, respectively. The distribution of the $|\Delta N_{\text{sig}}|/N_{\text{sig}}$ of the 1000 MC toys in the low- q^2 study is shown in Fig. 5.82. TABLE 5.43 summarizes the choice of parameterization in the fit and the impact, where the numbers are averaged over 1000 MC toys. The total systematic uncertainties due to the choice of fit parameterization

in the low- q^2 bin for PF-PF and PF-LP are 4.9% and 5.0%, respectively.

The combined systematics from the choice of fit parameterization in the J/ψ bin and the low- q^2 bin for PF-PF and PF-LP results in a total of 5.0% and 5.2%, respectively.

Table 5.42: Summary of the systematic uncertainties from the choice of fit parameterization in the J/ψ region for the electron channel. The total systematic uncertainties from the choice of fit parameterization are 0.8% and 1.5% for PF-PF and PF-LP, respectively.

Type	Nominal shape	Alternative shape	$\frac{\Delta N_{\text{sig}}}{N_{\text{sig}}} \text{ (PF-PF)}$	$\frac{\Delta N_{\text{sig}}}{N_{\text{sig}}} \text{ (PF-LP)}$
Signal	CB + Gaussian	CB + 2 Gaussian	4.1×10^{-5}	9.8×10^{-3}
Combinatorics	Exponential	2 nd -order Chebyshev	2.7×10^{-4}	1.9×10^{-4}
Partially reco.	KDE	CB + Gaussian	4.7×10^{-3}	9.1×10^{-3}
Other B	KDE	Gaussian	6.2×10^{-3}	7.4×10^{-3}
$J/\psi\pi$	CB	Gaussian	1.1×10^{-3}	8.6×10^{-4}
Total			7.9×10^{-3}	1.5×10^{-2}

Table 5.43: Summary of the systematic uncertainties from the choice of fit parameterization in the low- q^2 region for the electron channel. The numbers show here is averaged over 1000 MC toys. The total systematic uncertainties from the choice of fit parameterization are 4.9% and 5.0% for PF-PF and PF-LP, respectively.

Type	Nominal shape	Alternative shape	$\frac{\Delta N_{\text{sig}}}{N_{\text{sig}}} \text{ (PF-PF)}$	$\frac{\Delta N_{\text{sig}}}{N_{\text{sig}}} \text{ (PF-LP)}$
Signal	double-sided CB	CB + Gaussian	8.0×10^{-3}	2.1×10^{-2}
Combinatorics	Exponential	Power law	4.8×10^{-2}	4.4×10^{-2}
J/ψ Leakage	KDE	Exponential	4.1×10^{-3}	1.2×10^{-2}
Total			4.9×10^{-2}	5.0×10^{-2}

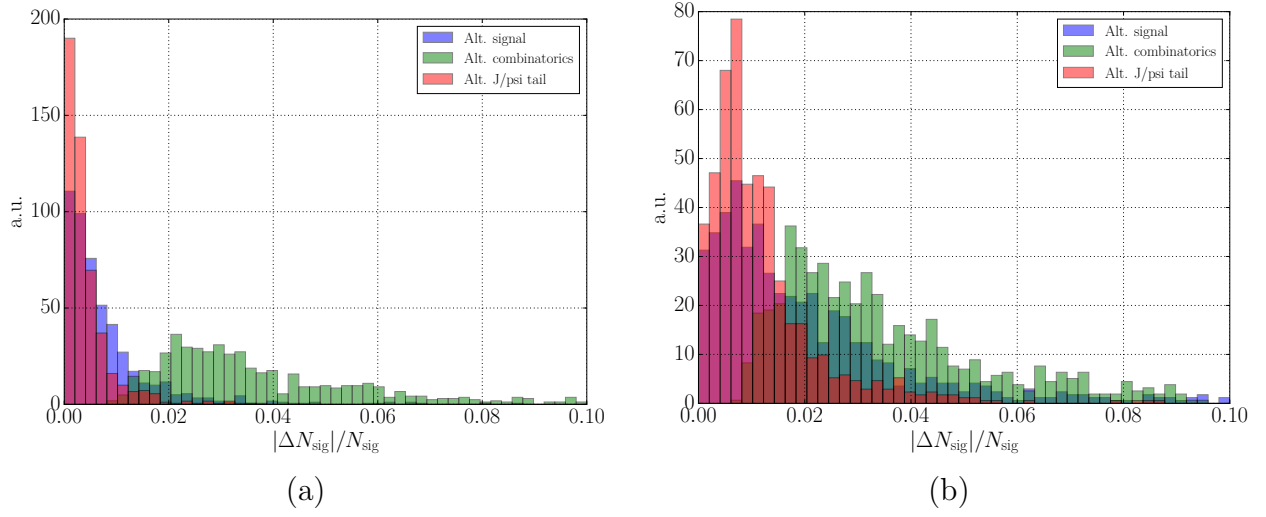


Figure 5.82: Distribution of the $|\Delta N_{\text{sig}}|/N_{\text{sig}}$ of the 1000 MC toys in the low- q^2 study for PF-PF (a) and PF-LP (b).

Systematic uncertainty from the J/ψ radiative tail

A long radiative tail is produced in the B peak due to bremsstrahlung from electrons. If the radiation tail is not modelled correctly, it can lead to an incorrect estimation of efficiency, resulting in systematic uncertainty for the R_K measurement. To examine the systematic uncertainty arising from the J/ψ radiative tail, we widened the q^2 window from the standard $2.9 < m(e^+e^-) < 3.2$ GeV to include the ranges $2.7 < m(e^+e^-) < 3.2$ GeV and $2.5 < m(e^+e^-) < 3.2$ GeV. The data in the J/ψ region were then refitted, and we compared the difference in signal yield to the nominal value.

The fits in the J/ψ region with widened q^2 windows are shown in Fig. 5.83. The number of signal N_{sig} and the corresponding acceptance times efficiency $\alpha \times \epsilon$ for different q^2 windows in the J/ψ bin are summarized in Table 5.44 and 5.45 for PF-PF and PF-LP, respectively. The mis-modelling of the J/ψ radiative tail results in systematic uncertainties of 1.3% and 0.7% for PF-PF and PF-LP, respectively.

Table 5.44: Summary of the signal yields and the efficiency with different q^2 windows in the J/ψ bin for PF-PF.

$m(e^+e^-)$ [GeV]	N_{sig}	$\alpha \times \epsilon$	$\frac{\Delta(N_{\text{sig}}/(\alpha \times \epsilon))}{N_{\text{sig}}/(\alpha \times \epsilon)}$
$2.9 < m(e^+e^-) < 3.2$	4857 ± 84	0.024787	N/A
$2.7 < m(e^+e^-) < 3.2$	5493 ± 97	0.028416	0.013
$2.5 < m(e^+e^-) < 3.2$	5767 ± 237	0.029477	0.002

Table 5.45: Summary of the signal yields and the efficiency with different q^2 windows in the J/ψ bin for PF-LP.

$m(e^+e^-)$ [GeV]	N_{sig}	$\alpha \times \epsilon$	$\frac{\Delta(N_{\text{sig}}/(\alpha \times \epsilon))}{N_{\text{sig}}/(\alpha \times \epsilon)}$
$2.9 < m(e^+e^-) < 3.2$	2098 ± 58	0.010848	N/A
$2.7 < m(e^+e^-) < 3.2$	2419 ± 64	0.012593	0.007
$2.5 < m(e^+e^-) < 3.2$	2533 ± 364	0.013058	0.003

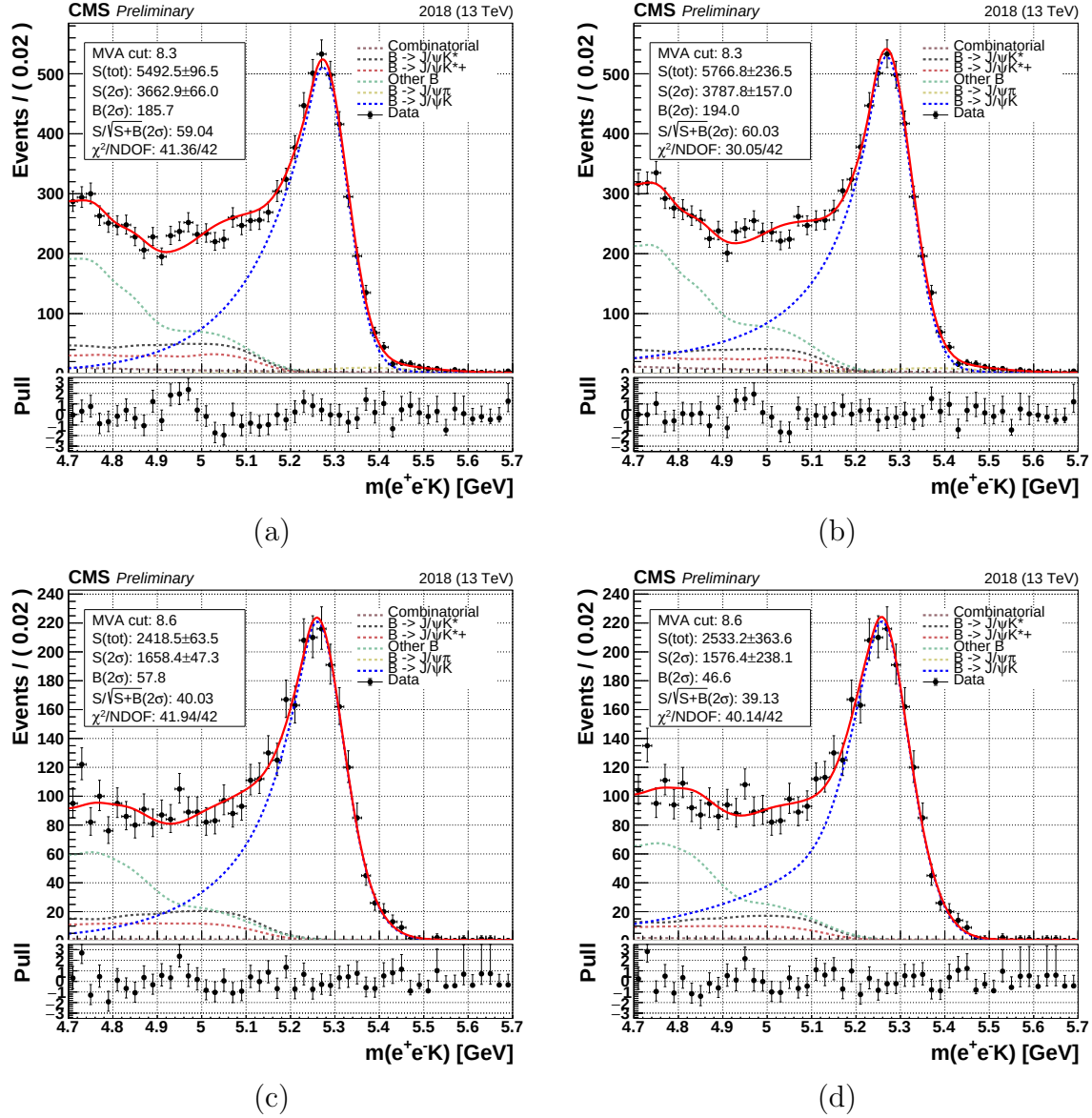


Figure 5.83: Fits in the J/ψ region in the q^2 window of $2.7 < m(e^+e^-) < 3.2$ GeV (left), and $2.5 < m(e^+e^-) < 3.2$ GeV (right) for PF-PF (top) and PF-LP (bottom) categories.

Systematic uncertainty from fixing J/ψ leakage yield in the low- q^2 bin

To prevent overfitting due to limited data points in the low- q^2 region in the electron channel, the normalization of J/ψ leakage in the low- q^2 bin is fixed to the expected yield. To estimate the systematic uncertainty of fixing the J/ψ leakage yield, we need to identify processes that can cause deviation from the expected yield and examine their effect on the signal yield N_{sig} . The study of J/ψ radiative tail shows a small 1.3% mis-modeling, and the statistical uncertainty of the J/ψ signal yield in the J/ψ bin used to estimate the expected J/ψ leakage yield is also negligible at 1.7%. The primary fluctuation in the J/ψ leakage yield $N_{J/\psi \text{ tail}}$ is due to statistical variation.

We generated 1000 MC toys, in which the signal and combinatorics follow the expected distribution, and J/ψ tail is added according to a Poisson distribution, as illustrated in 5.84. We then re-fitted each MC toy to obtain the signal yield and examine the fluctuation of the J/ψ tail's impact on the signal yield. The standard deviation of the refitted signal yield of those 1000 MC toys is taken to be the systematic uncertainty. The results demonstrate that the systematics of fixing J/ψ leakage in the low- q^2 bin are 3.67% and 9.30% in PF-PF and PF-LP, respectively.

Summary of the systematic uncertainties in the electron channel

Table 5.46 summarizes the systematic uncertainties and their relative impact on R_K^e for the PF-PF and PF-LP categories. The sources of systematic uncertainties are uncorrelated and the total is sum in quadrature. They are incorporated into the likelihood function by assuming they are following the Gaussian distribution.

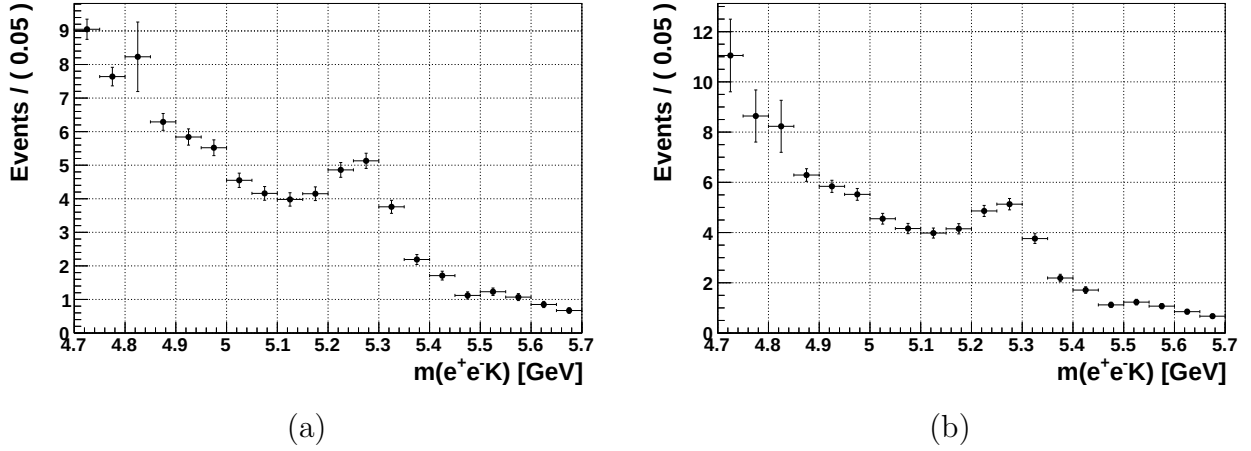


Figure 5.84: Examples of the MC toys generated for estimating the systematic uncertainty of constraining the J/ψ leakage yields in the low- q^2 region of the electron channel.

Table 5.46: Summary of systematics for the R_K measurement in the electron channel

Source	$\Delta R_K^e / R_K^e$ (PF-PF)	$\Delta R_K^e / R_K^e$ (PF-LP)
Trigger path	0.010	0.040
BDT Cross-validation & MC stat	0.017	0.026
BDT stability	0.021	0.047
BDT SF	0.010	0.015
Pile UP SF	0.004	0.002
Choice of fit parameterization	0.050	0.052
J/ψ radiative tail	0.013	0.007
J/ψ leakage	0.037	0.093
Total	0.071	0.127

5.9 Cross-checks

Cross-checks are essential for verifying the whole analysis procedure, which ensures that the results obtained are valid and reliable. We did several cross-checks in the R_K measurement to show that the analysis strategy gives expected and consistent results in the control channels.

5.9.1 Cross-check with $R_K^{\psi(2S)}$

Since the $\psi(2S) \rightarrow \ell^+\ell^-$ decay is well tested to be satisfying LU, the decay channel of $B^+ \rightarrow K^+\psi(2S)(\rightarrow \ell^+\ell^-)$ can be used as a cross-check to verify the analysis procedure. Similar to R_K measurement, the cross-check in $\psi(2S)$ bin is done by measuring a double ratio of branching fractions defined as

$$R_K^{\psi(2S)} = \frac{\mathcal{B}(B^+ \rightarrow K^+\psi(2S)(\rightarrow \mu^+\mu^-))}{\mathcal{B}(B^+ \rightarrow K^+J/\psi(\rightarrow \mu^+\mu^-))} \bigg/ \frac{\mathcal{B}(B^+ \rightarrow K^+\psi(2S)(\rightarrow e^+e^-))}{\mathcal{B}(B^+ \rightarrow K^+J/\psi(\rightarrow e^+e^-))}, \quad (5.23)$$

which is predicted to be unity in the SM because of LU.

In terms of measured quantities, $R_K^{\psi(2S)}$ is expressed as

$$R_K^{\psi(2S)} = \left(\frac{N_{K^+\psi(2S)(\mu^+\mu^-)}/N_{K^+J/\psi(\mu^+\mu^-)}}{N_{K^+\psi(2S)(e^+e^-)}/N_{K^+J/\psi(e^+e^-)}} \right) \left(\frac{\epsilon_{K^+\psi(2S)(e^+e^-)}/\epsilon_{K^+J/\psi(e^+e^-)}}{\epsilon_{K^+\psi(2S)(\mu^+\mu^-)}/\epsilon_{K^+J/\psi(\mu^+\mu^-)}} \right), \quad (5.24)$$

where N is the number of signal measured in the data, and ϵ is the efficiency.

A simultaneous fit is used in order to measure the value of $R_K^{\psi(2S)}$. The number of $B^+ \rightarrow K^+\psi(2S)$ in the muon mode is expressed as

$$N_{K^+\psi(2S)}^\mu = r_B^{\psi(2S)} \cdot N_{K^+J/\psi}^\mu \cdot \frac{\mathcal{B}(\psi(2S) \rightarrow \mu^+\mu^-)}{\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)} \cdot \frac{\epsilon_{K^+\psi(2S)}^\mu}{\epsilon_{K^+J/\psi}^\mu}, \quad (5.25)$$

and that in the electron mode is expressed as

$$N_{K^+\psi(2S)}^e = \frac{r_{\mathcal{B}}^{\psi(2S)}}{R_K^{\psi(2S)}} \cdot N_{K^+J/\psi}^e \cdot \frac{\mathcal{B}(\psi(2S) \rightarrow e^+e^-)}{\mathcal{B}(J/\psi \rightarrow e^+e^-)} \cdot \frac{\epsilon_{K^+\psi(2S)}^e}{\epsilon_{K^+J/\psi}^e}, \quad (5.26)$$

where $r_{\mathcal{B}}^{\psi(2S)} = \frac{\mathcal{B}(B^+ \rightarrow K^+\psi(2S))}{\mathcal{B}(B^+ \rightarrow K^+J/\psi)}$. The simultaneously fit is done across these quantities. The details of the parameters are summarized in Table 5.47. The measurement of $R_K^{\psi(2S)}$ and its confidence level can be obtained through profiling the likelihood function.

Type	Parameter	Note
parameter of interest	$R_K^{\psi(2S)}$	shared between PF-PF and PF-LP
parameter of interest	$r_{\mathcal{B}}^{\psi(2S)}$	shared between PF-PF and PF-LP
observable	$N_{K^+\psi(2S)}^e$	
observable	$N_{K^+\psi(2S)}^\mu$	shared between PF-PF and PF-LP
observable	$N_{K^+J/\psi}^e$	
observable	$N_{K^+J/\psi}^\mu$	shared between PF-PF and PF-LP
nuisance	$\mathcal{B}(\psi(2S) \rightarrow e^+e^-)$	shared between PF-PF and PF-LP
nuisance	$\mathcal{B}(J/\psi \rightarrow e^+e^-)$	shared between PF-PF and PF-LP
nuisance	$\mathcal{B}(\psi(2S) \rightarrow \mu^+\mu^-)$	shared between PF-PF and PF-LP
nuisance	$\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)$	shared between PF-PF and PF-LP
nuisance	$\epsilon_{K^+\psi(2S)}^e$	
nuisance	$\epsilon_{K^+\psi(2S)}^\mu$	shared between PF-PF and PF-LP
nuisance	$\epsilon_{K^+J/\psi}^e$	
nuisance	$\epsilon_{K^+J/\psi}^\mu$	shared between PF-PF and PF-LP

Table 5.47: Summary of the parameters for simultaneous fit of $R_K^{\psi(2S)}$.

Figure 5.85 shows the profile likelihood for the fit to $\psi(2S)$ decays. The non-Gaussian distribution of the profile likelihood is due to larger statistical uncertainties in the denominator of the ratio compared to the numerator. The measurement of $R_K^{\psi(2S)}$ is $0.966_{-0.066}^{+0.071}$, and the 68% and 95% confidence intervals are $[0.90001, 1.03727]$ and $[0.84081, 1.11233]$, respectively. This result is consistent with the standard model prediction at the 1σ level.

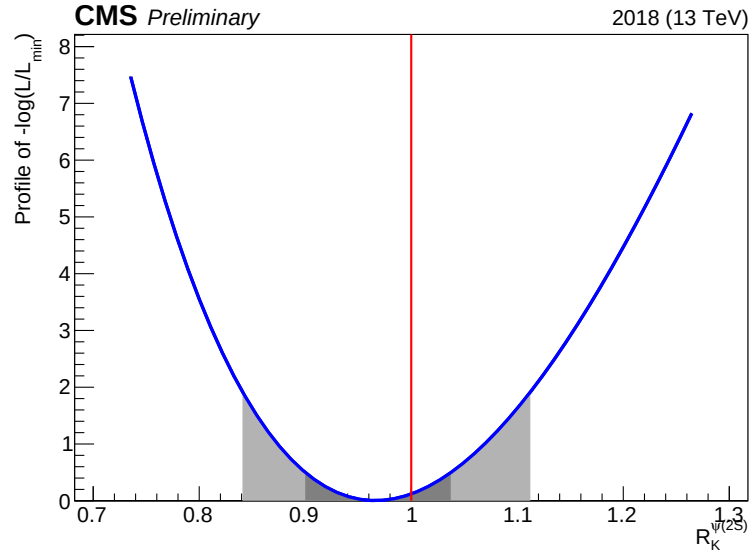


Figure 5.85: Likelihood function from the fit profiled as a function of $R_K^{\psi(2S)}$. The measurement agrees with the SM prediction within 1σ .

5.9.2 Cross-check with $R_K^{J/\psi}$

The decay of $J/\psi \rightarrow \ell^+ \ell^-$ is well tested to be lepton universal, we expected to see the same number of signals for $B^+ \rightarrow K^+ J/\psi (\rightarrow \mu^+ \mu^-)$ and $B^+ \rightarrow K^+ J/\psi (\rightarrow e^+ e^-)$ after taking acceptance and efficiency into account. The measurement of $R_K^{J/\psi}$ is another important cross-check to verify the analysis procedure. Unlike R_K or $R_K^{\psi(2S)}$ which are defined in a double ratio, $R_K^{J/\psi}$ is measured in a single ratio such that it does not benefit from the cancellation of systematic effects that the double ratio used. This is a test of the control of the efficiencies in the analysis. The ratio of branching fractions, $R_K^{J/\psi}$ is written as

$$R_K^{J/\psi} = \frac{\mathcal{B}(B^+ \rightarrow K^+ J/\psi (\rightarrow \mu^+ \mu^-))}{\mathcal{B}(B^+ \rightarrow K^+ J/\psi (\rightarrow e^+ e^-))}, \quad (5.27)$$

which is predicted to be unity in the SM due to LU.

In terms of measured quantities, $R_K^{J/\psi}$ is

$$R_K^{J/\psi} = \left(\frac{N_{J/\psi(\mu^+ \mu^-)K^+}}{N_{J/\psi(e^+ e^-)K^+}} \right) \left(\frac{\epsilon_{J/\psi(e^+ e^-)K^+}}{\epsilon_{J/\psi(\mu^+ \mu^-)K^+}} \right), \quad (5.28)$$

where N is the number of signal measured in the data, and ϵ is the efficiency.

This quantity is measured using a simultaneous fit on

$$N_{K^+ J/\psi}^e = \frac{N_{K^+ J/\psi}^\mu}{R_K^{J/\psi}} \cdot \frac{\epsilon_{K^+ J/\psi}^e}{\epsilon_{K^+ J/\psi}^\mu}, \quad (5.29)$$

where this quantity is simultaneously fitted in both PF-PF and PF-LP to get a combined result. The details of the parameters are summarized in Table 5.48.

Type	Parameter	Note
parameter of interest	$R_K^{J/\psi}$	shared between PF-PF and PF-LP
observable	$N_{K+J/\psi}^e$	
observable	$N_{K+J/\psi}^\mu$	shared between PF-PF and PF-LP
nuisance	$\epsilon_{K+J/\psi}^e$	
nuisance	$\epsilon_{K+J/\psi}^\mu$	shared between PF-PF and PF-LP

Table 5.48: Summary of the parameters for simultaneous fit of $R_K^{J/\psi}$.

Since $R_K^{J/\psi}$ is measured in a single ratio without normalization channel, we need to take care of the trigger efficiency in the electron side in order to compare with the muon channel. Explicitly, the efficiency consists of these terms

$$\epsilon_{K+J/\psi} = \epsilon_{\text{trigger}} \times \epsilon_{\text{post-trigger}}, \quad (5.30)$$

where $\epsilon_{\text{post-trigger}} = \alpha_{\text{detector}} \times \epsilon_{\text{reco}} \times \epsilon_{\text{pre-sel}} \times \epsilon_{q^2} \times \epsilon_{\text{BDT}}$ is the usual efficiency we used on the electron channel when measuring the double ratio like R_K and $R_K^{\psi(2S)}$.

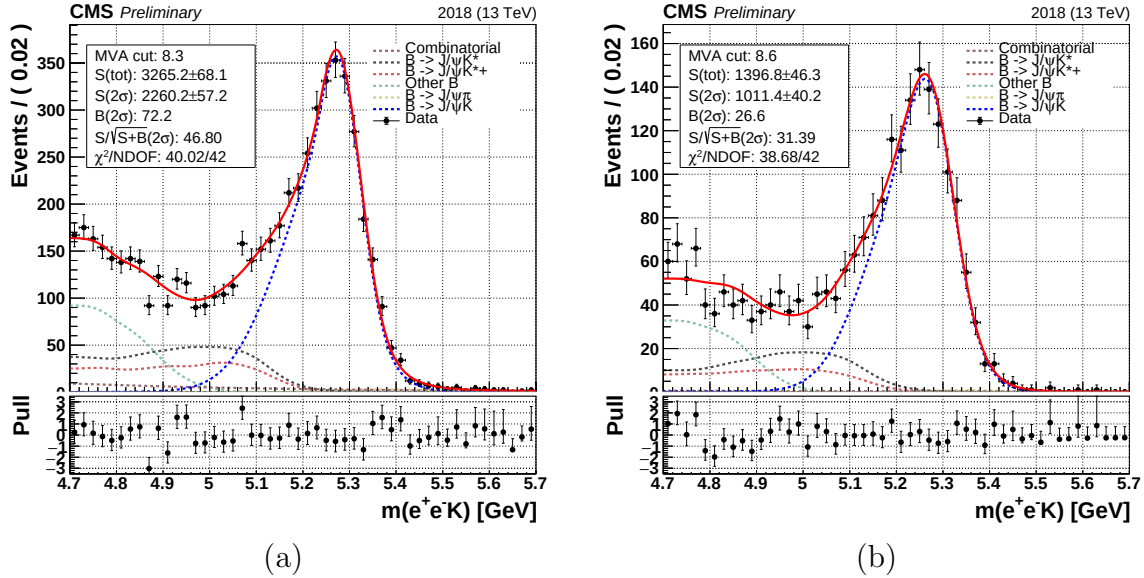


Figure 5.86: B-candidates mass distribution for events which pass the HLT_Mu9_IP6 trigger in the J/ψ bin for PF-PF (a) and PF-LP (b). The fit is overlaid in red.

In the R_K measurement, the muon channel uses only events that passed the HLT_Mu9_IP6 trigger, while the electron channel uses all events in the B-parking data. The trigger normalization between the two channels is eliminated when we make measurements using the double ratio. However, when we measure $R_K^{J/\psi}$ in a single ratio, we must consider the different HLT exposure periods in the B-parking data. To simplify the problem, we decided to use only the HLT_Mu9_IP6 data in the electron channel for $R_K^{J/\psi}$. This eliminates the need to address the different HLT exposure times and the dependence of efficiency on HLT. The fits to the HLT_Mu9_IP6 data in the J/ψ bin are shown in Fig 5.86.

The double ratio measurements, such as R_K and $R_K^{\psi(2S)}$, do not account for the trigger efficiency on the electron side, as the trigger efficiency on the probe side is entirely cancelled out in the ratio. However, when measuring $R_K^{J/\psi}$, it is necessary to consider the trigger efficiency on the probe side in order to compare with the muon channel. The probe side trigger efficiency of HLT_Mu9_IP6 has been measured to be $2.081 \pm 0.016 \times 10^{-4}$.

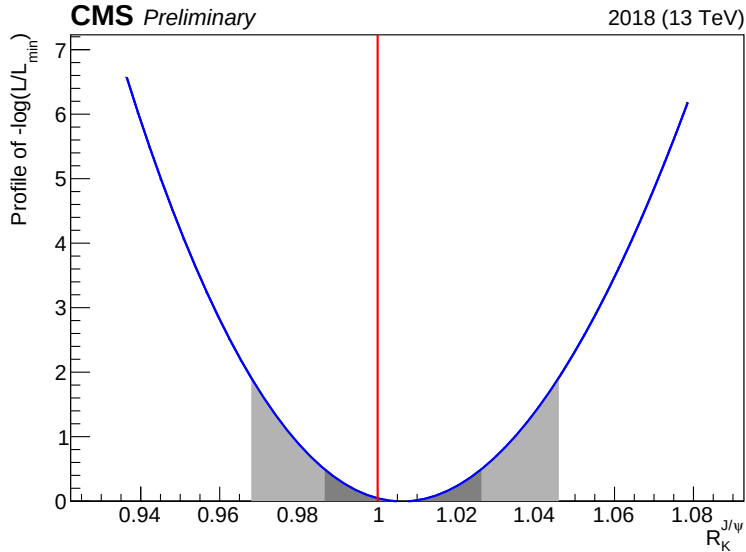


Figure 5.87: Likelihood function from the fit profiled as a function of $R_K^{J/\psi}$. The measurement agrees with the expected value within 1σ .

Figure 5.87 shows the profile likelihood for the fit to the J/ψ decays. The value of $R_K^{J/\psi}$ is determined to be $1.006_{-0.019}^{+0.020}$. The 68% and 95% confidence intervals are $[0.98659, 1.02628]$ and $[0.96800, 1.04590]$, respectively. The obtained result is in agreement with the expected value within 1σ .

5.9.3 Cross-check with single ratio $R_{\psi(2S)}^{\text{single ratio}}$

In Sec. 5.9.1, the cross-check of $R_K^{\psi(2S)}$ is performed using a double ratio, where the J/ψ region is utilized as a normalization channel. To further verify the results, we measured $R_{\psi(2S)}^{\text{single ratio}}$ using the same methodology as the $R_K^{J/\psi}$ cross-check, as described in Sec. 5.9.2. $R_{\psi(2S)}^{\text{single ratio}}$ is defined as a single ratio as an extra cross-checks, which is written as

$$R_{\psi(2S)}^{\text{single ratio}} = \frac{\mathcal{B}(B^+ \rightarrow K^+ \psi(2S)(\rightarrow \mu^+ \mu^-))}{\mathcal{B}(B^+ \rightarrow K^+ \psi(2S)(\rightarrow e^+ e^-))}, \quad (5.31)$$

and this quantity can be measured using a simultaneous fit on

$$N_{K^+ \psi(2S)}^e = \frac{N_{K^+ \psi(2S)}^\mu}{R_{\psi(2S)}^{\text{single ratio}}} \cdot \frac{\epsilon_{K^+ \psi(2S)}^e}{\epsilon_{K^+ \psi(2S)}^\mu}, \quad (5.32)$$

where this quantity is simultaneous fitted in both PF-PF and PF-LP to get a combined result.

Figure 5.88 shows the fits to HLT_Mu9_IP6 data in the $\psi(2S)$ bin. The profile likelihood for the fit to the $\psi(2S)$ decays can be seen in Fig. 5.89. The value of $R_K^{J/\psi}$ was determined to be $0.923_{-0.059}^{+0.066}$, with 68% and 95% confidence intervals of [0.86445, 0.98878] and [0.81292, 1.05855], respectively. The result is found to be around 1.2σ higher than the expected value.

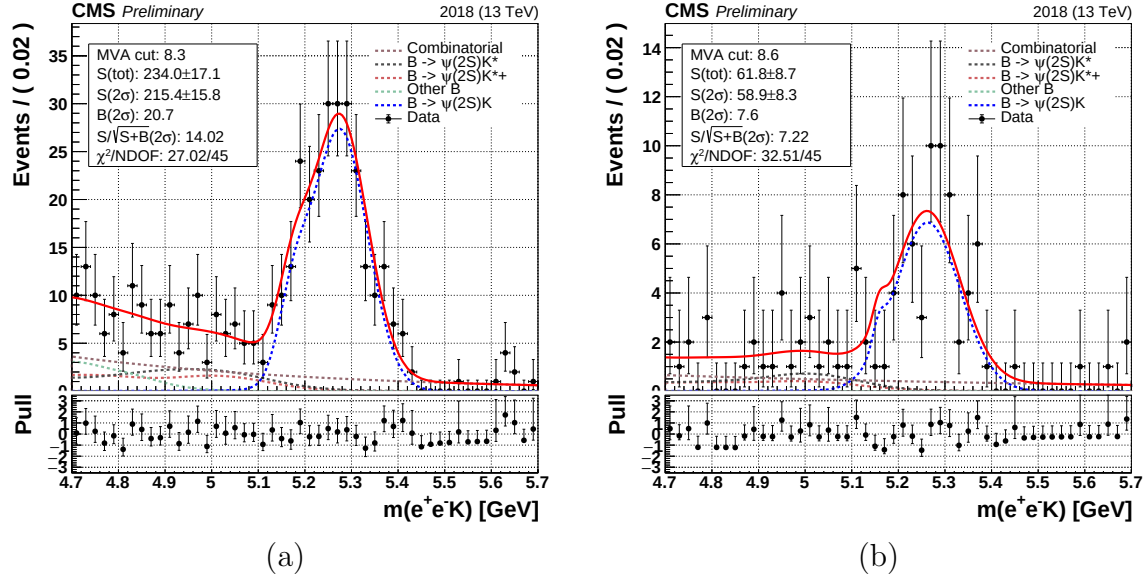


Figure 5.88: B-candidates mass distribution for events which pass the HLT_Mu9_IP6 trigger in the $\psi(2S)$ bin for PF-PF (a) and PF-LP (b). The fit is overlaid in red.

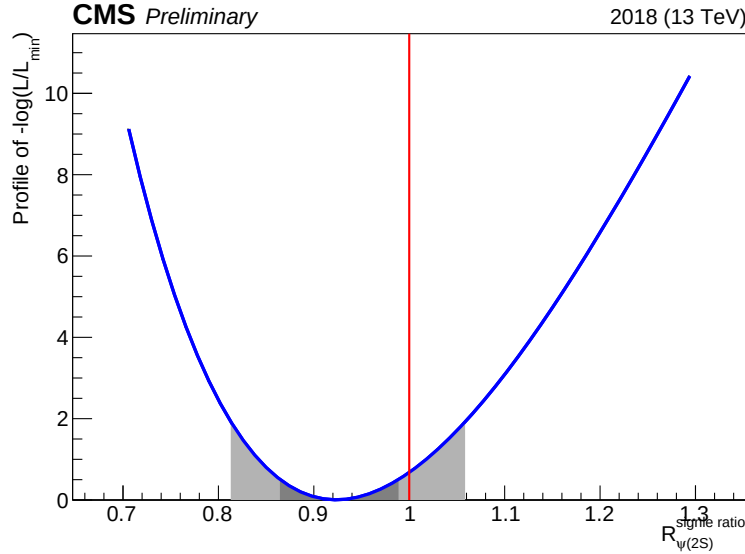


Figure 5.89: Likelihood function from the fit profiled as a function of $R_{\psi(2S)}^{\text{single ratio}}$. The expected value is 1.2σ higher than the measured value.

5.10 Results

A simultaneous fit is performed across the following quantities to measure the value of R_K :

$$N_{K^+\ell^+\ell^-}^\mu = r_{\mathcal{B}} \cdot \frac{N_{K^+J/\psi}^\mu}{\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)} \cdot \frac{\epsilon_{K^+\ell^+\ell^-}^\mu}{\epsilon_{K^+J/\psi}^\mu}, \quad (5.33)$$

$$N_{K^+\ell^+\ell^-}^e = R_K^{-1} \cdot r_{\mathcal{B}} \cdot \frac{N_{K^+J/\psi}^e}{\mathcal{B}(J/\psi \rightarrow e^+e^-)} \cdot \frac{\epsilon_{K^+\ell^+\ell^-}^e}{\epsilon_{K^+J/\psi}^e}, \quad (5.34)$$

where $r_{\mathcal{B}} = \frac{\mathcal{B}(B^+ \rightarrow K^+\mu^+\mu^-)}{\mathcal{B}(B^+ \rightarrow K^+J/\psi)}$.

The likelihood function of R_K is defined as

$$\mathcal{L}^i = \mathcal{L}_{\text{low-}q^2}^i(N_{K^+\ell^+\ell^-}^i) \times \mathcal{L}_{J/\psi}^i(N_{K^+J/\psi}^i) \times G(\epsilon_{K^+\ell^+\ell^-}^i) \times G(\epsilon_{J/\psi}^i) \times G(\mathcal{B}_{J/\psi}^i), \quad (5.35)$$

where $i = e/\mu$, $\mathcal{L}_{q^2}$ is the mass fit likelihood function used in the q^2 bin, and G is the Gaussian function. The systematic uncertainties are incorporated into the likelihood, assuming they follow Gaussian distributions. The details of the parameters are summarized in Table 5.49. The measurement of R_K and its confidence level can be obtained through profiling the likelihood function of Eq. 5.35.

It is important to note that in the likelihood, the parameter of interest used is R_K^{-1} instead of R_K because R_K appears in the denominator of Eq. 5.33, resulting in a highly non-Gaussian likelihood. Selecting R_K^{-1} makes the likelihood distribution behave much more like a Gaussian, thereby making it easier to converge. The profile likelihood for the fit to the non-resonant decays is illustrated in Fig. 5.90.

The value of R_K^{-1} is found to be $R_K^{-1} = 1.275$ from maximum likelihood. The profile

likelihood yields the 68% CL for R_K^{-1} with and without systematic uncertainty as $[0.80233, 1.80591]$ and $[0.79711, 1.81624]$, respectively. For the 95% CL of R_K^{-1} with and without systematic uncertainty, the values are $[0.39053, 2.37266]$ and $[0.39016, 2.40685]$, respectively. The value of R_K is measured to be $R_K = 0.784^{+0.470}_{-0.234}(\text{total})$, or $R_K = 0.784^{+0.462}_{-0.230}(\text{stat})^{+0.087}_{-0.038}(\text{syst})$, which is consistent with the SM prediction within 1σ .

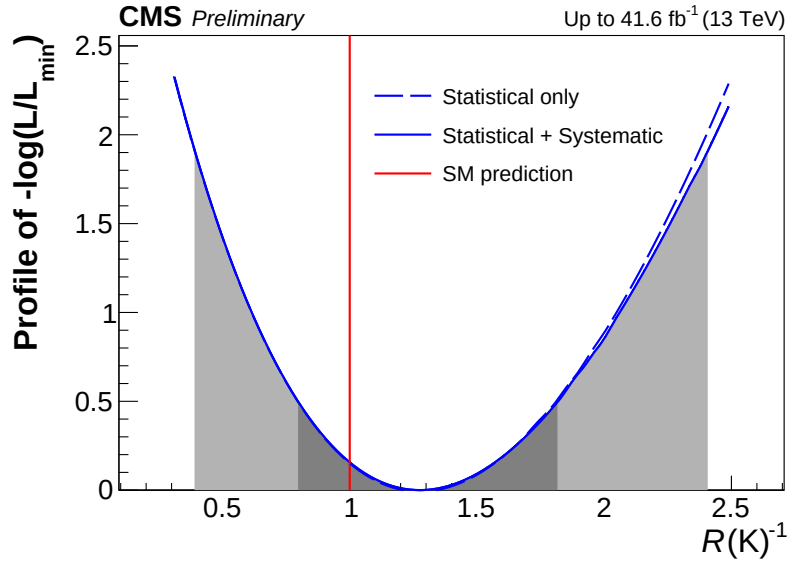


Figure 5.90: Likelihood function from the fit profiled as a function of R_K^{-1} . The measurement agrees with the expected value within 1σ .

Type	Parameter	Note
parameter of interest	R_K^{-1}	shared between PF-PF and PF-LP
parameter of interest	$r_{\mathcal{B}}$	shared between PF-PF and PF-LP
observable	N_{K+l+l-}^e	
observable	N_{K+l+l-}^μ	shared between PF-PF and PF-LP
observable	$N_{K+J/\psi}^e$	
observable	$N_{K+J/\psi}^\mu$	shared between PF-PF and PF-LP
nuisance	$\mathcal{B}(J/\psi \rightarrow e^+e^-)$	shared between PF-PF and PF-LP
nuisance	$\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)$	shared between PF-PF and PF-LP
nuisance	ϵ_{K+l+l-}^e	
nuisance	ϵ_{K+l+l-}^μ	shared between PF-PF and PF-LP
nuisance	$\epsilon_{K+J/\psi}^e$	
nuisance	$\epsilon_{K+J/\psi}^\mu$	shared between PF-PF and PF-LP

Table 5.49: Summary of the parameters for simultaneous fit of R_K .

5.11 Conclusions

This thesis reports on a test of lepton universality using the decays of $B^+ \rightarrow K^+ \mu^+ \mu^-$ and $B^+ \rightarrow K^+ e^+ e^-$ with data from the CMS experiment at the LHC. The corresponding J/ψ channel was used to normalize the measurements and reduce systematic uncertainties.

The test was performed using the data set collected during the LHC Run II of proton-proton collisions at $\sqrt{s} = 13$ TeV in 2018, using a special data-taking mode triggered by a low transverse momentum muon. The analysis of both the muon and electron channels was checked for consistency, and the value of R_K was obtained from the low- q^2 bin with a measurement of $R_K = 0.784^{+0.470}_{-0.234}$, which is consistent with the SM prediction within 1σ . The statistical uncertainty in the electron channel dominated the overall uncertainty in R_K .

Appendix A

Distributions for the check of mass sculpting in the electron channel

This section shows the distributions for the check of mass sculpting in the electron channel.

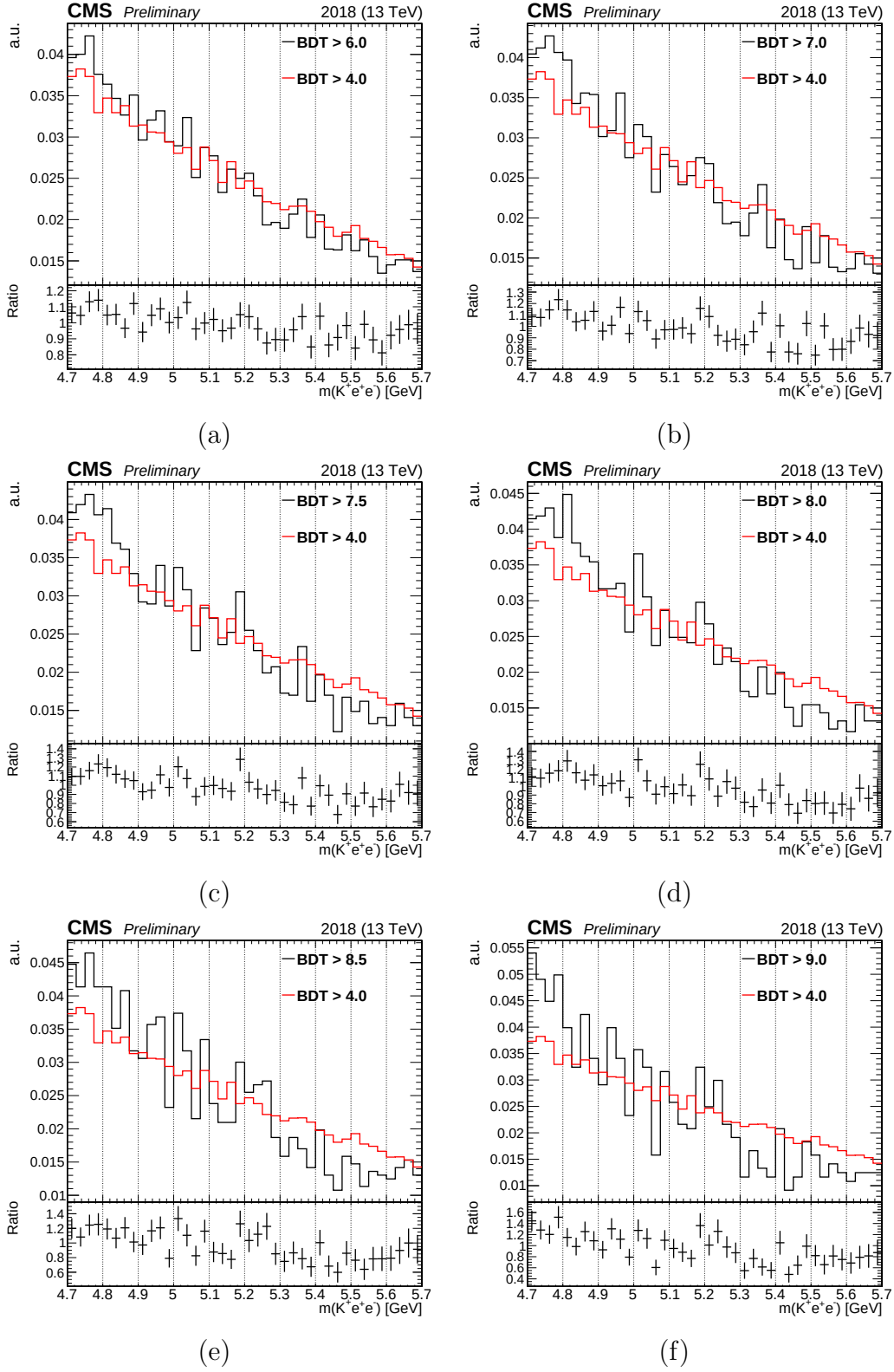
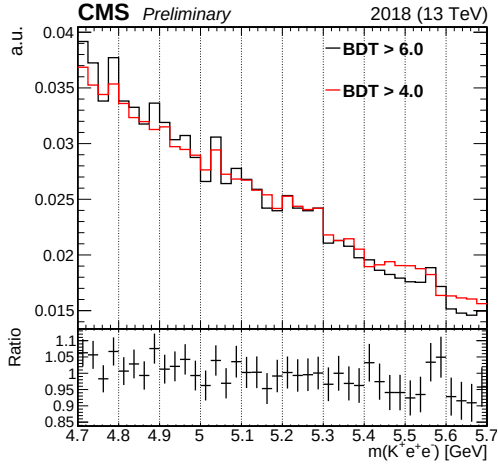
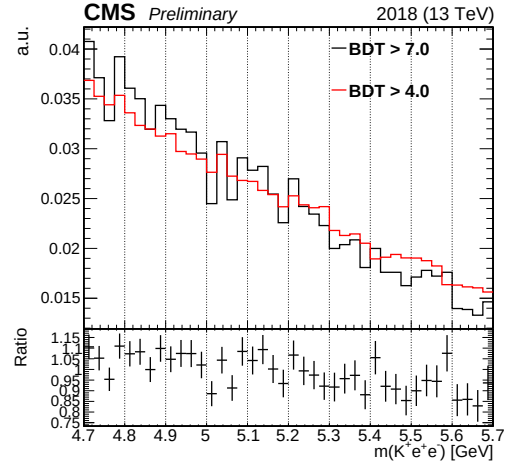


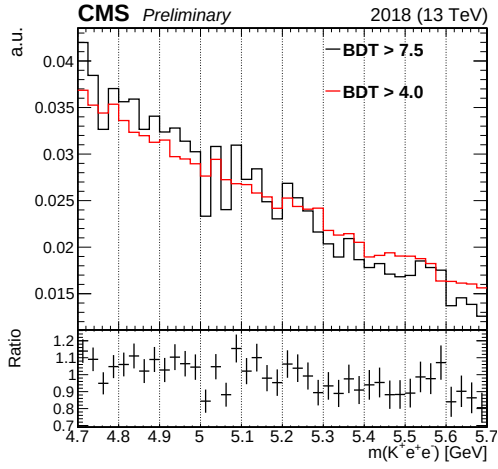
Figure A.0.1: Check of mass sculpting using same-sign data in working point 6.0 (a), 7.0 (b), 7.5 (c), 8.0 (d), 8.5 (e), and 9.0 (f) for PF-PF.



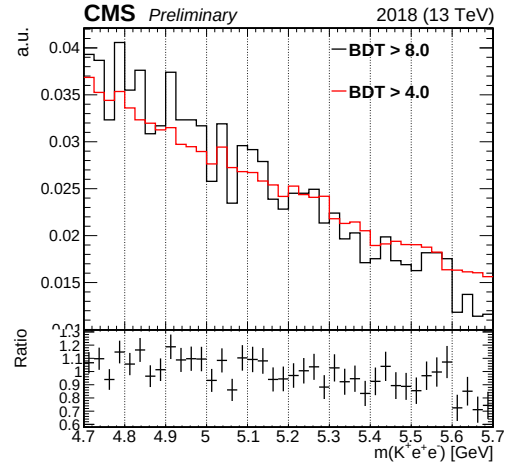
(a)



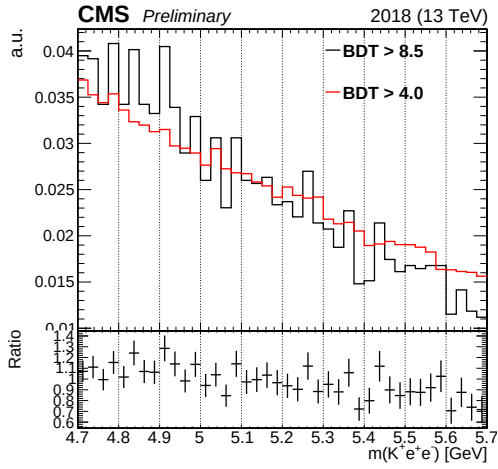
(b)



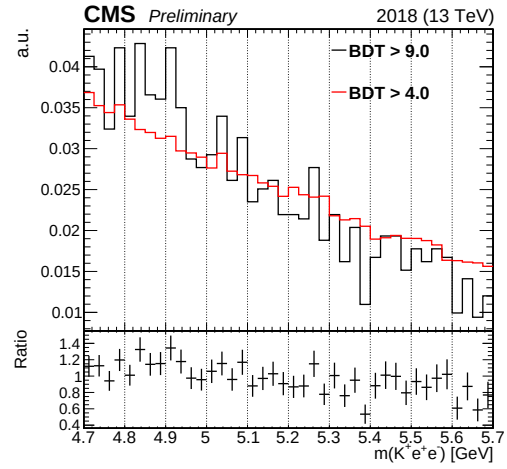
(c)



(d)



(e)



(f)

Figure A.0.2: Check of mass sculpting using same-sign data in working point 6.0 (a), 7.0 (b), 7.5 (c), 8.0 (d), 8.5 (e), and 9.0 (f) for PF-LP.

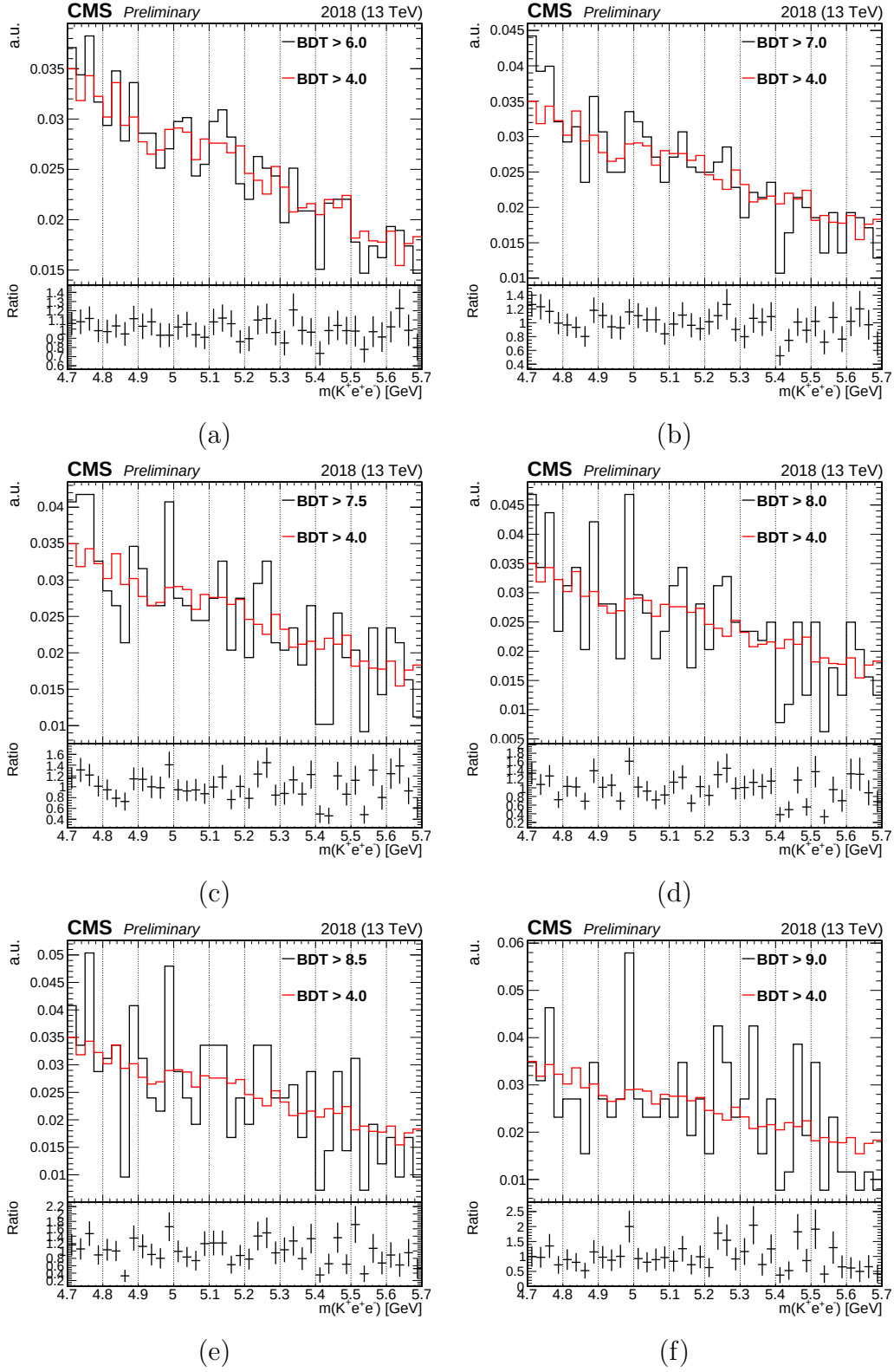


Figure A.0.3: Check of mass sculpting using combinatorics in $B^+ \rightarrow K^+ J/\psi(e^+e^-)$ MC sample in working point 6.0 (a), 7.0 (b), 7.5 (c), 8.0 (d), 8.5 (e), and 9.0 (f) for PF-PF.

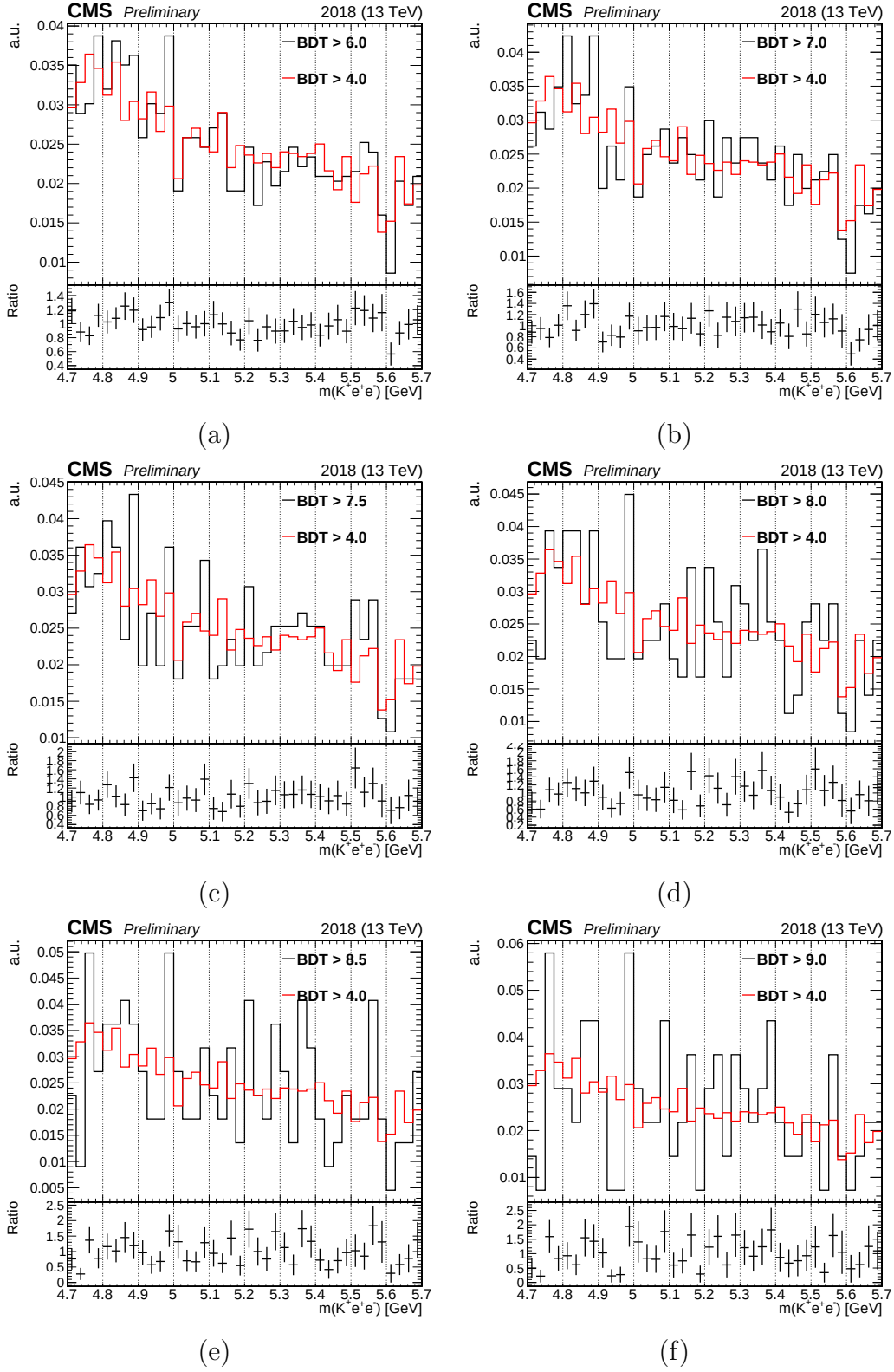


Figure A.0.4: Check of mass sculpting using combinatorics in $B^+ \rightarrow K^+ J/\psi(e^+e^-)$ MC sample in working point 6.0 (a), 7.0 (b), 7.5 (c), 8.0 (d), 8.5 (e), and 9.0 (f) for PF-LP.

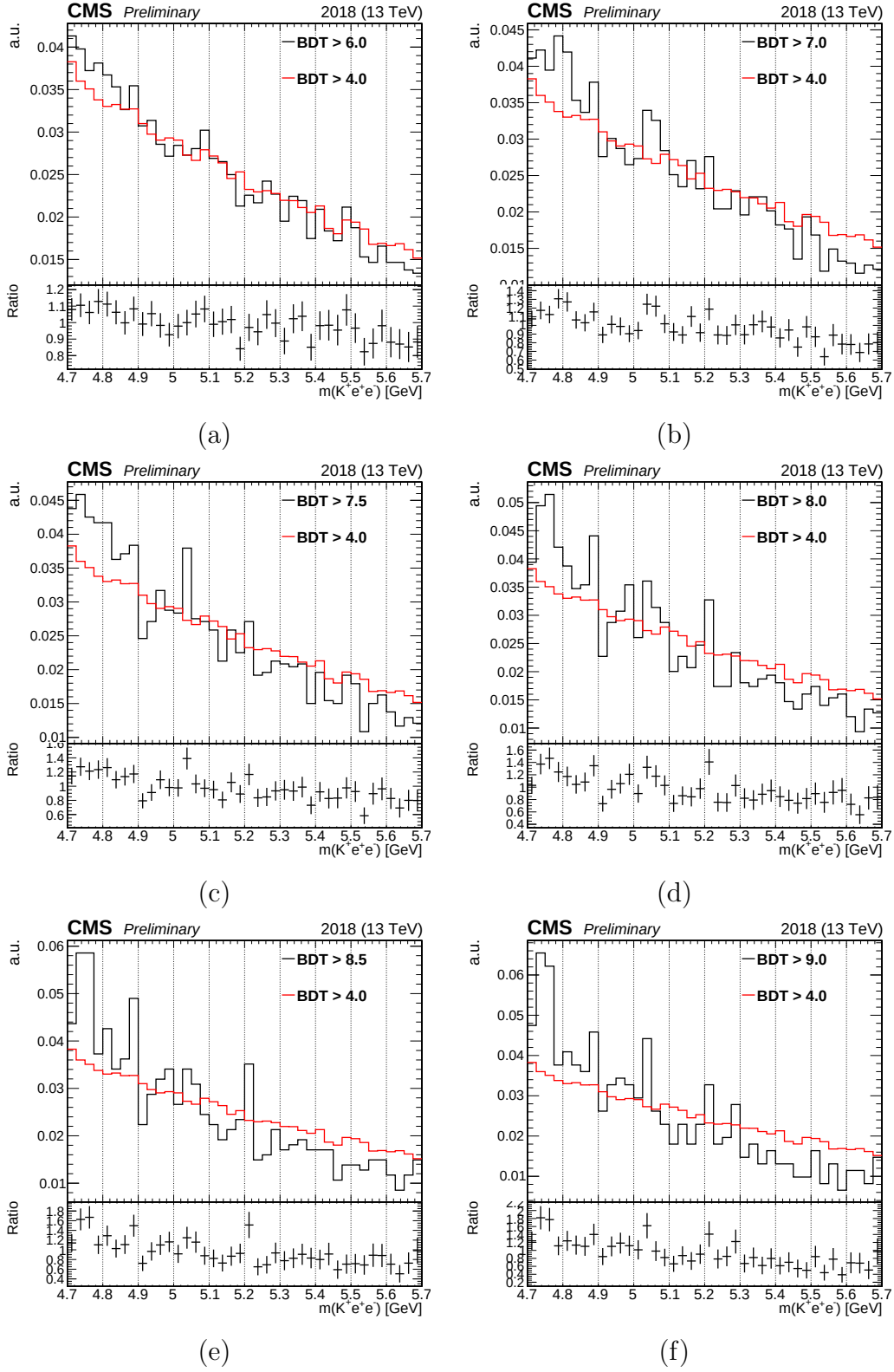


Figure A.0.5: Check of mass sculpting using combinatorics in the generic B decays MC sample in working point 6.0 (a), 7.0 (b), 7.5 (c), 8.0 (d), 8.5 (e), and 9.0 (f) for PF-PF.

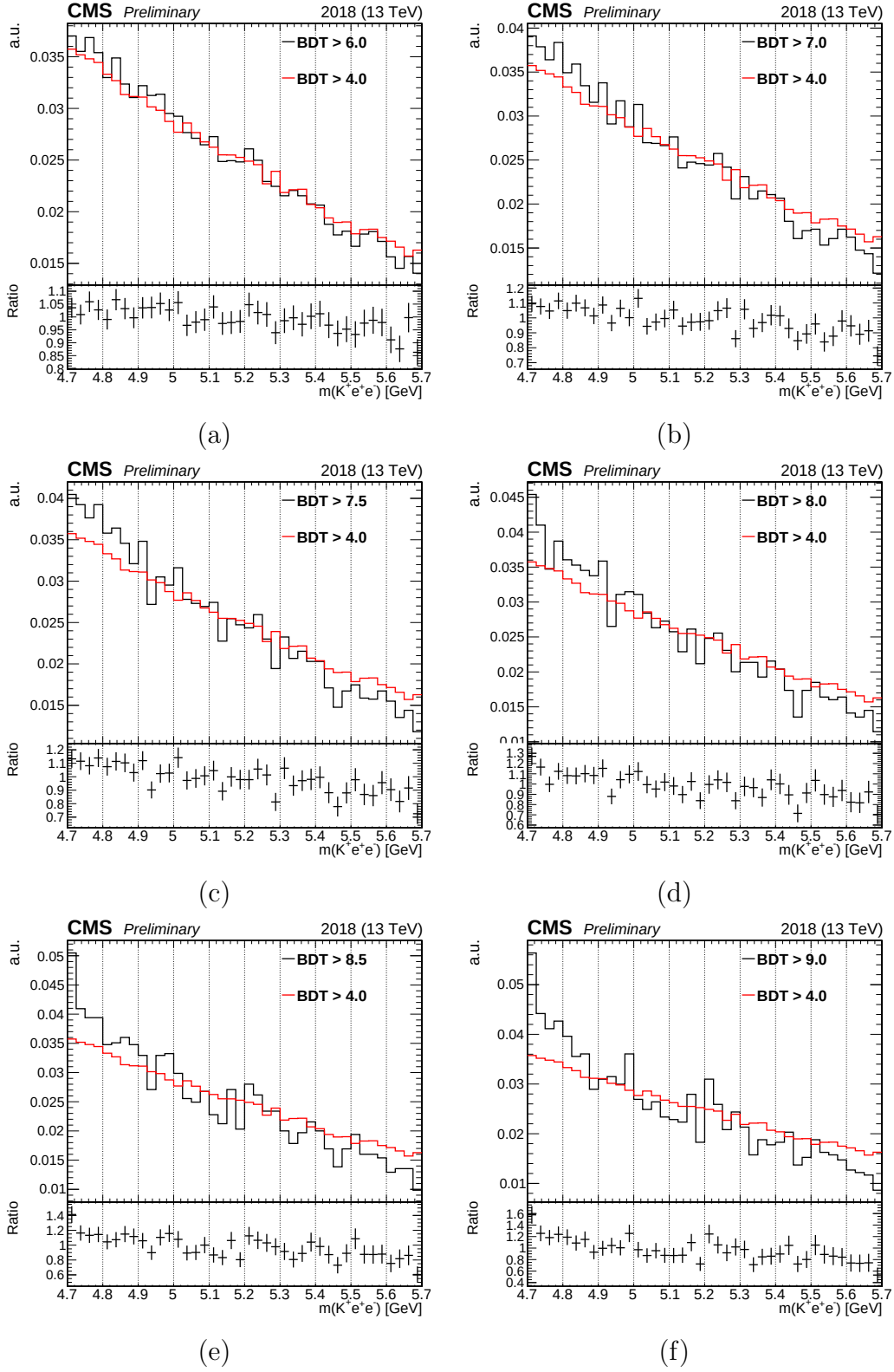


Figure A.0.6: Check of mass sculpting using combinatorics in the generic B decays MC sample in working point 6.0 (a), 7.0 (b), 7.5 (c), 8.0 (d), 8.5 (e), and 9.0 (f) for PF-LP.

Appendix B

Distributions for the check of misidentified background in the electron channel

This section shows the distributions for the check of misidentified background in the electron channel.

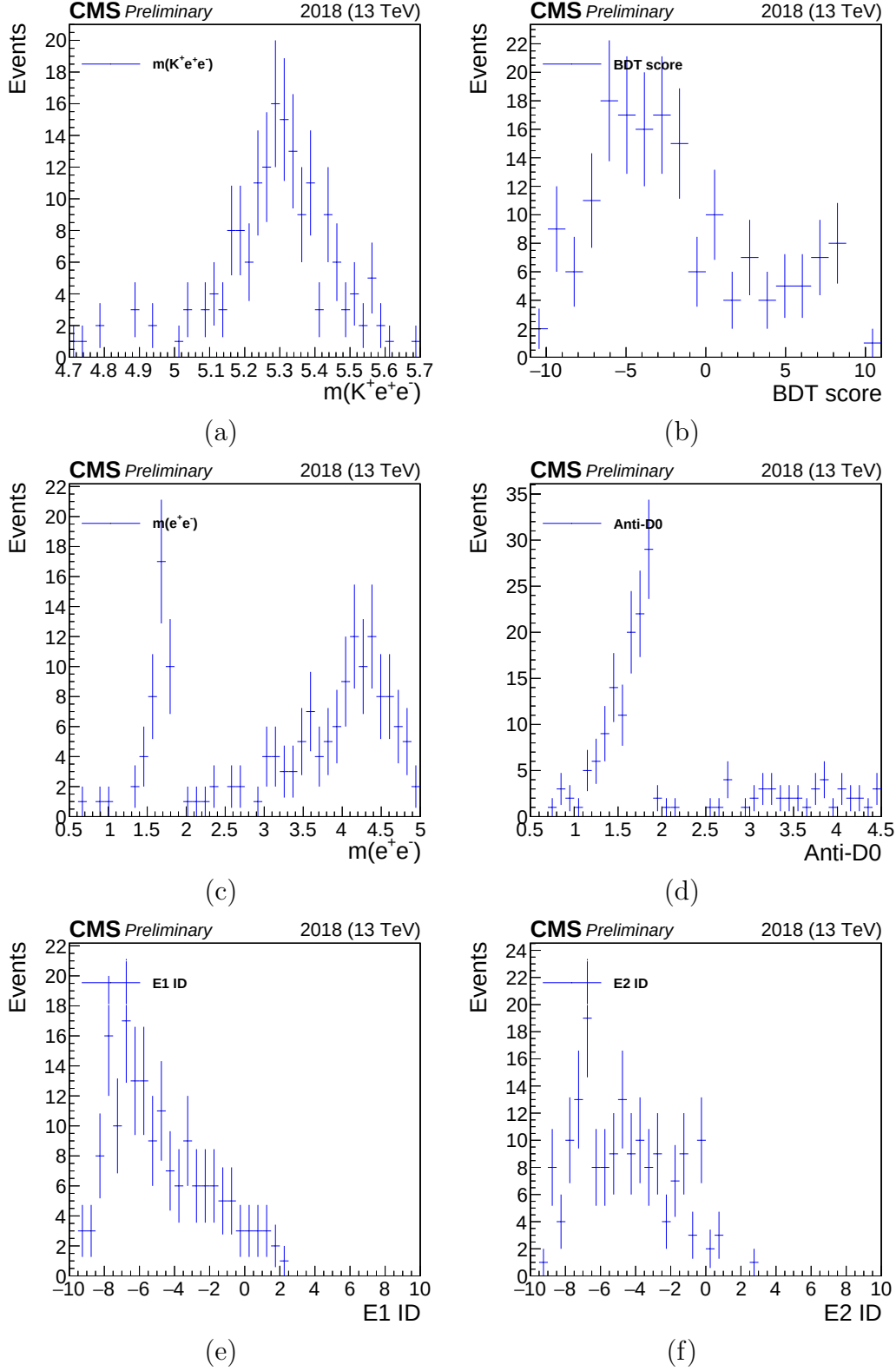


Figure B.0.1: Distributions of $m(K^+e^+e^-)$ (a), BDT score (b), $m(e^+e^-)$ (c), and $\min(m(K^+e_{[\rightarrow\pi^-]}^-), m(K_{[\rightarrow\pi^+]}^+e_{[\rightarrow K^-]}^-))$ (d), leading electron ID (e), and subleading electron ID (f) of $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-)\pi^+$ simulated events for PF-PF after pre-selection.

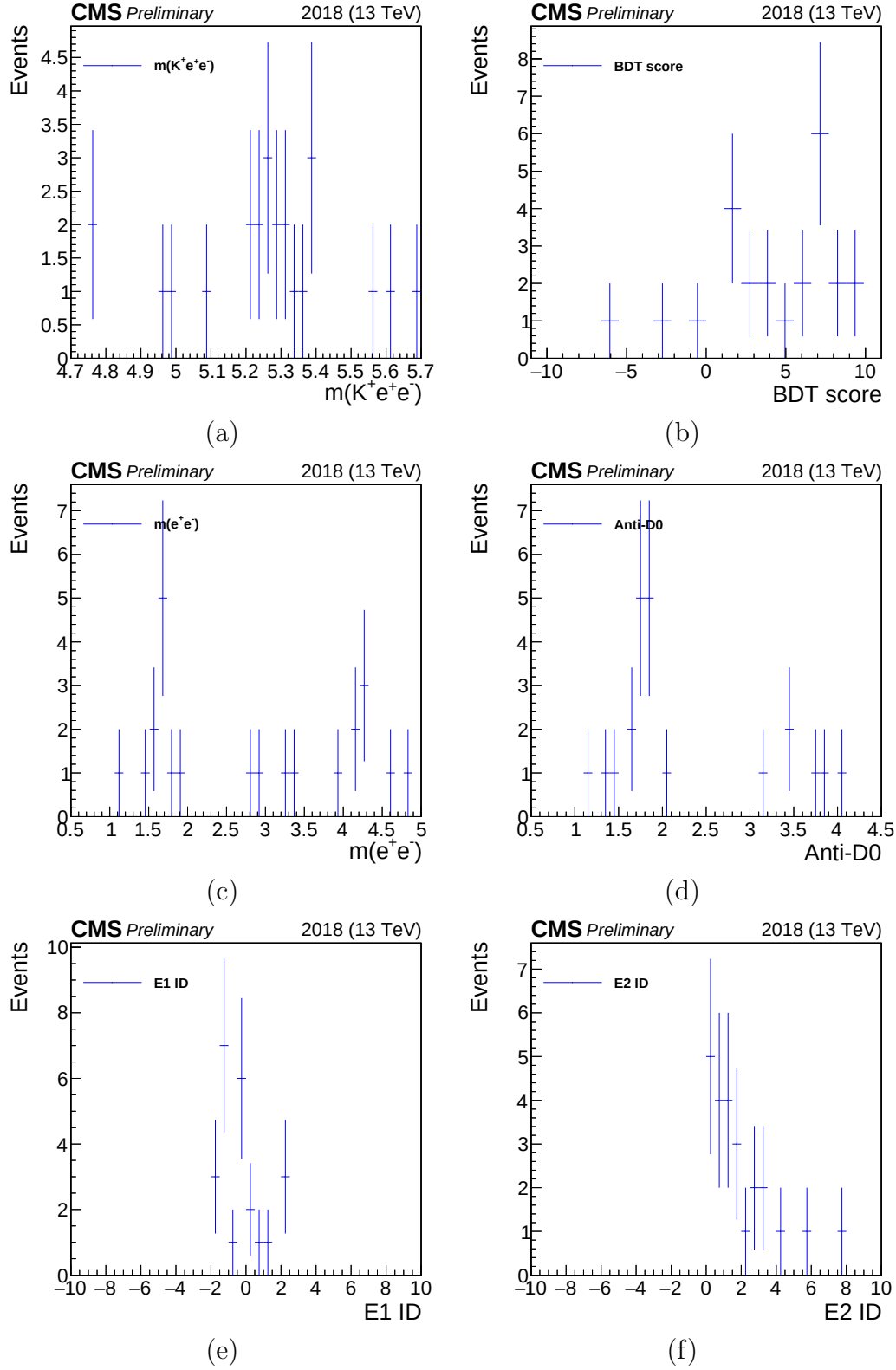


Figure B.0.2: Distributions of $m(K^+e^+e^-)$ (a), BDT score (b), $m(e^+e^-)$ (c), and $\min(m(K^+e_{[\rightarrow\pi^-]}^-), m(K_{[\rightarrow\pi^+]}^+e_{[\rightarrow K^-]}^-))$ (d), leading electron ID (e), and subleading electron ID (f) of $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-)\pi^+$ simulated events for PF-LP after pre-selection.

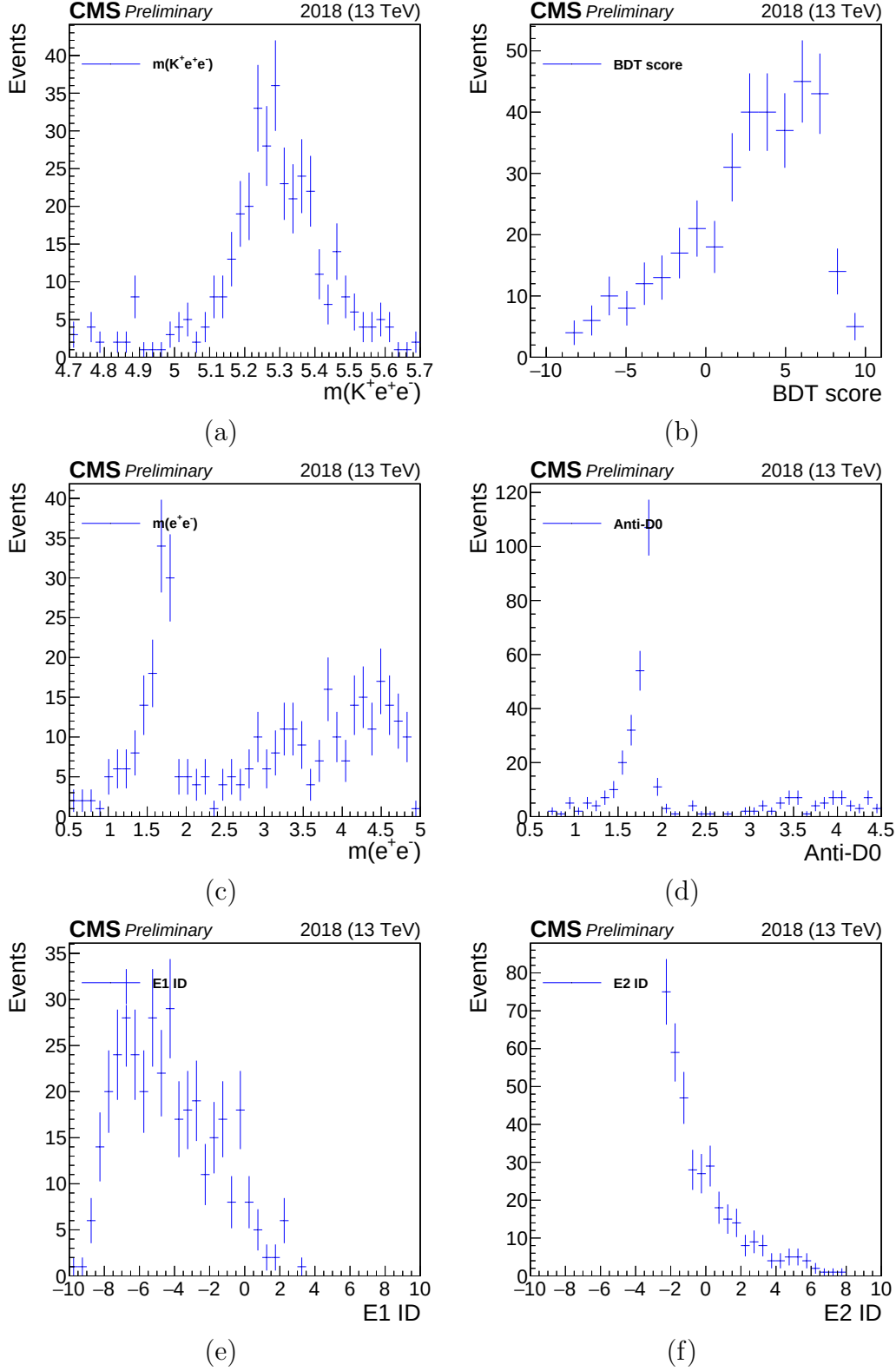


Figure B.0.3: Distributions of $m(K^+e^+e^-)$ (a), BDT score (b), $m(e^+e^-)$ (c), and $\min(m(K^+e_{[\rightarrow\pi^-]}^-), m(K_{[\rightarrow\pi^+]}^+e_{[\rightarrow K^-]}^-))$ (d), leading electron ID (e), and subleading electron ID (f) of $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-)\pi^+$ simulated events for PF-LP after pre-selection without applying the electron ID cut.

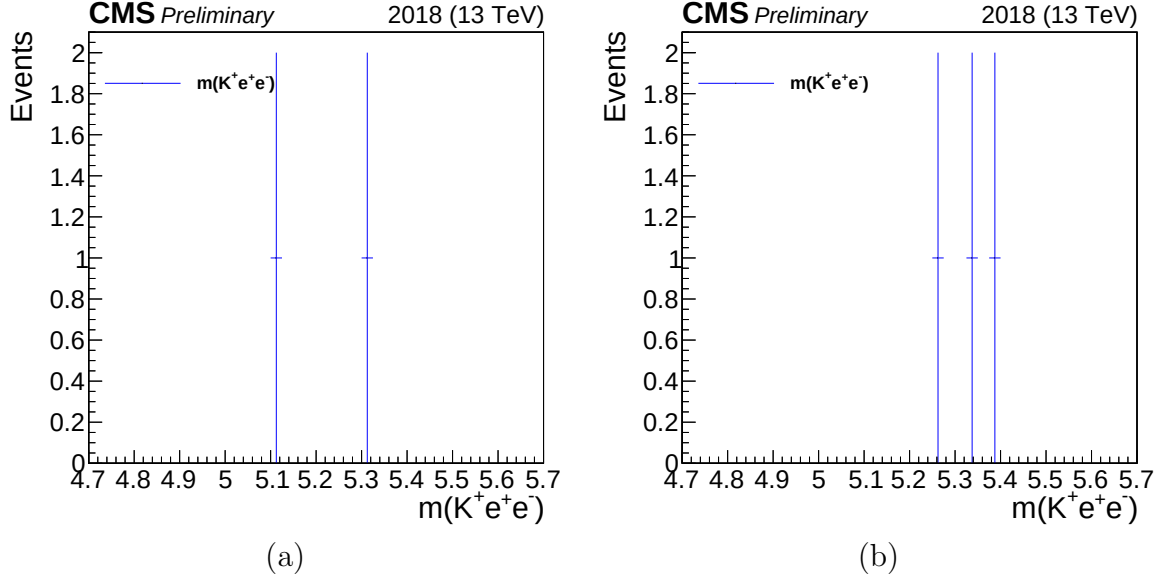


Figure B.0.4: Distribution of $m(K^+e^+e^-)$ of $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-)\pi^+$ simulated events for PF-PF (a), and PF-LP (b) after nominal BDT cut.

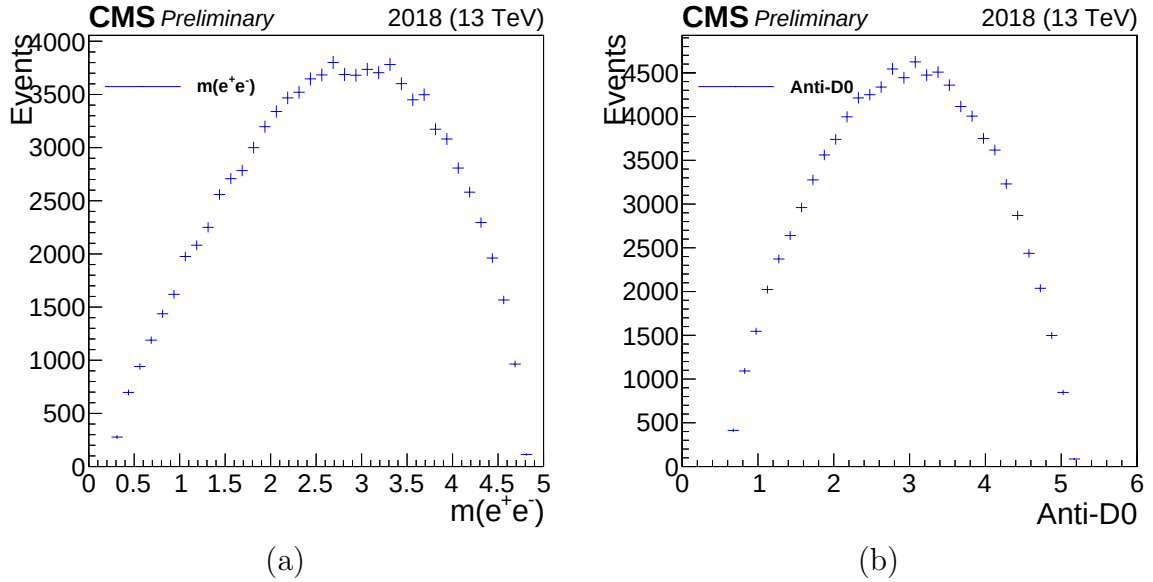


Figure B.0.5: Distribution of $m(K^+e^+e^-)$ (a) and anti- D^0 variable of generated $B^+ \rightarrow K^+\pi^-\pi^+$ simulated events .

Appendix C

Alternative templates used for the study of the choice of fit parameterization in the muon channel

This section shows the alternative templates used and the total fits for the study of the choice of fit parameterization in the electron channel.

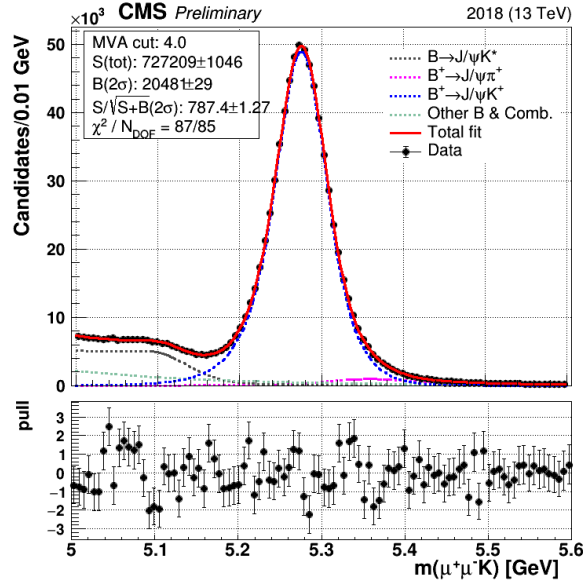


Figure C.0.1: Fit performed with an alternative signal function that uses four Gaussians.

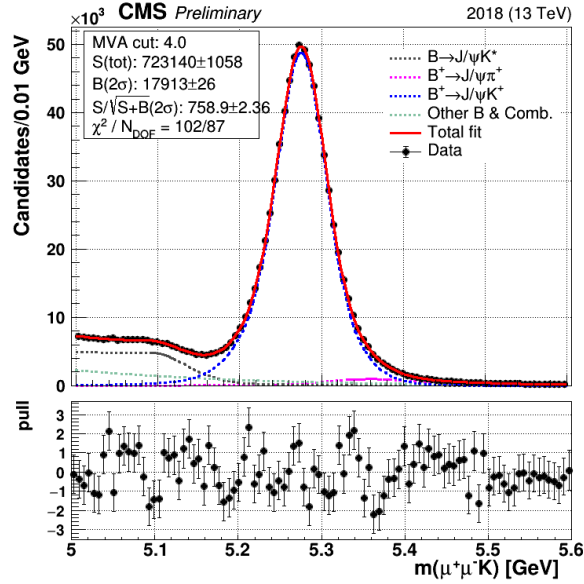
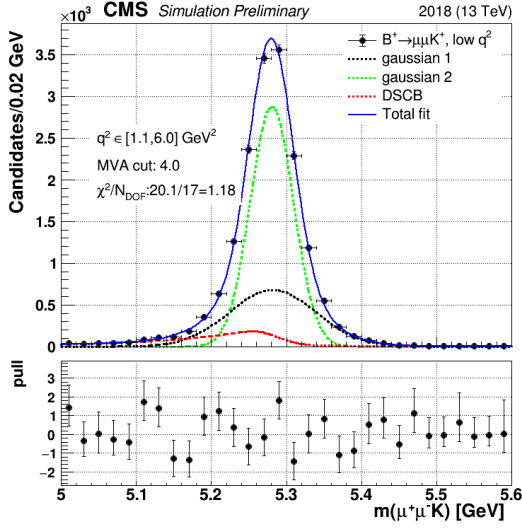
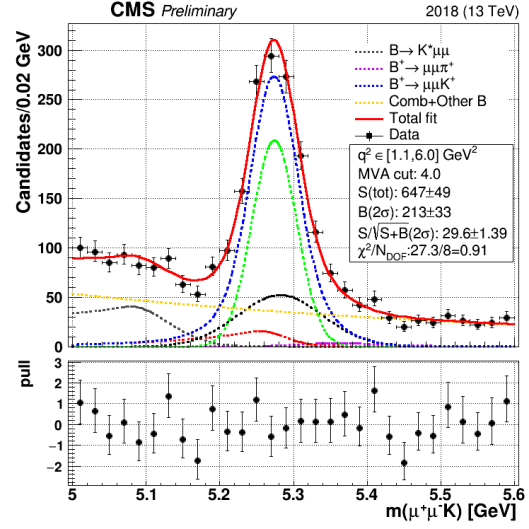


Figure C.0.2: Fit performed with an alternative background function that uses a sum of Gaussian and polynomial.



(a)



(b)

Figure C.0.3: Fit performed with alternative signal function using two gaussians and one double-sided Crystal Ball in MC (a) and in Data (b) in the region $q^2 \in (1.1, 6.0)$ GeV^2 .

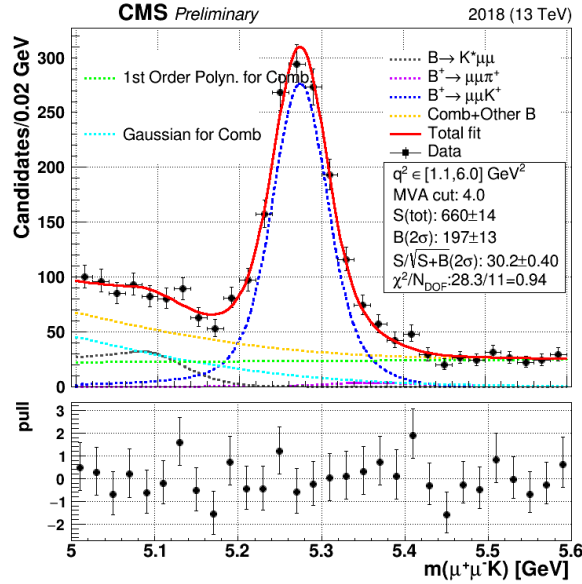
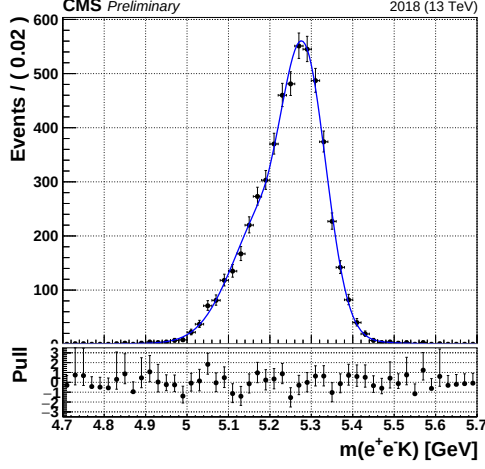


Figure C.0.4: Fit performed with alternative background function using one gaussian and a first order polynomial in the region $q^2 \in (1.1, 6.0)$ GeV^2 .

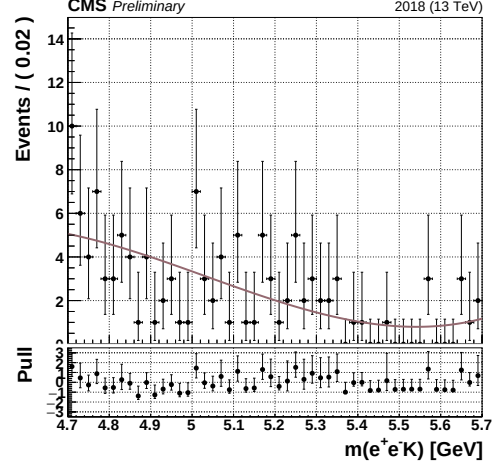
Appendix D

Alternative templates used for the study of the choice of fit parameterization in the electron channel

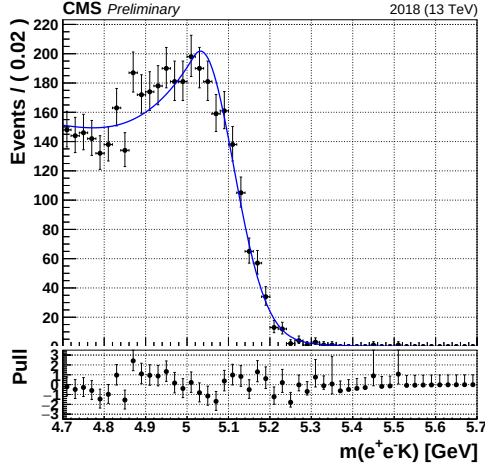
This section shows the alternative templates used and the total fits for the study of the choice of fit parameterization in the electron channel.



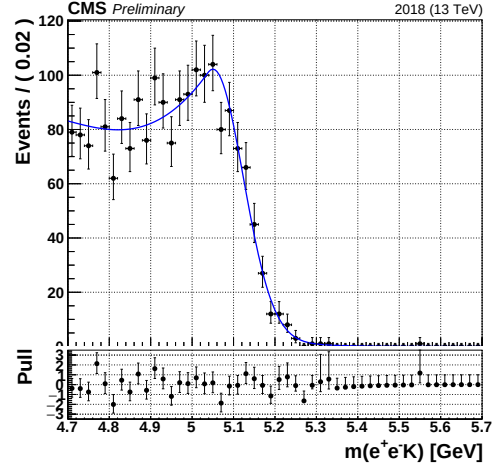
(a)



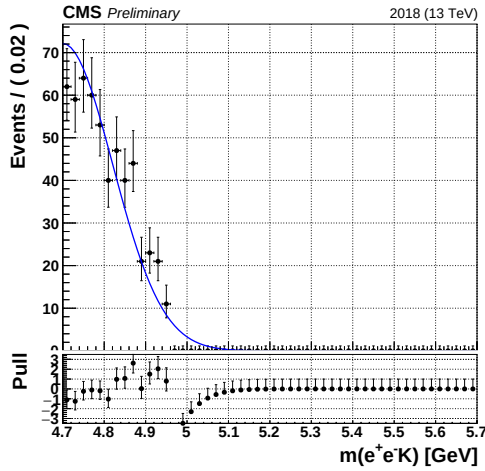
(b)



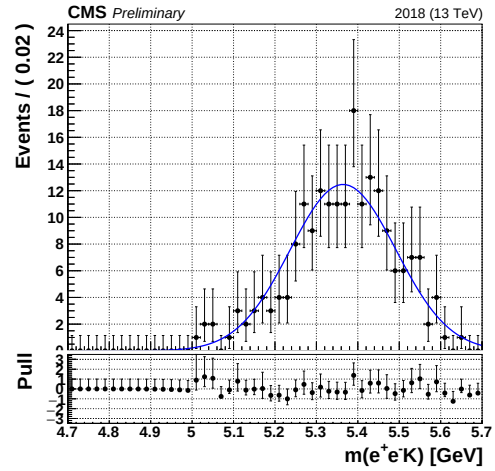
(c)



(d)

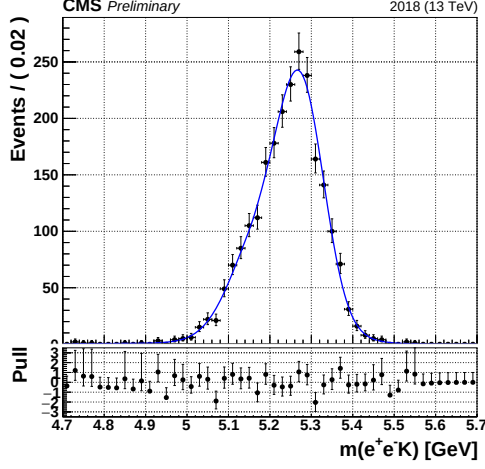


(e)

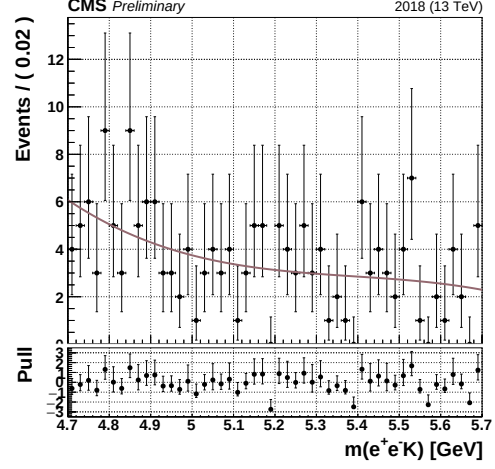


(f)

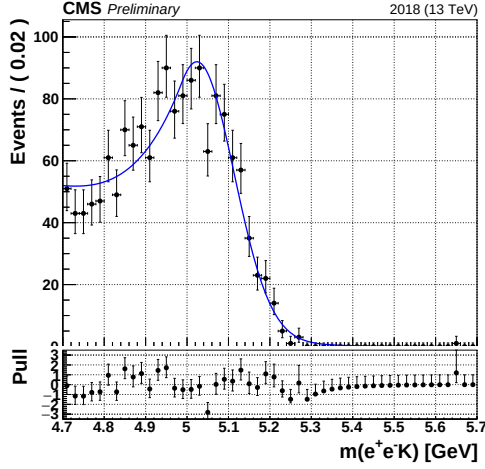
Figure D.0.1: Alternative fitting shape in the J/ψ region for signal (a), combinatorics (b), K^{*0} partially reconstructed background (c), K^{*+} partially reconstructed background (d), other B components (e), and $J/\psi\pi$ contamination (f) for PF-PF.



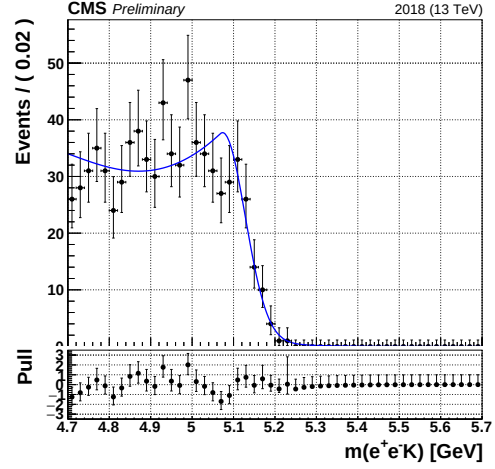
(a)



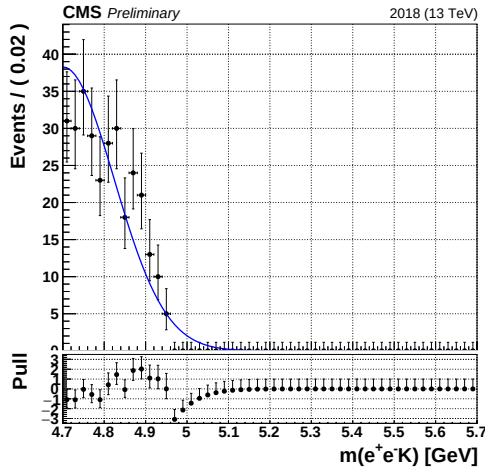
(b)



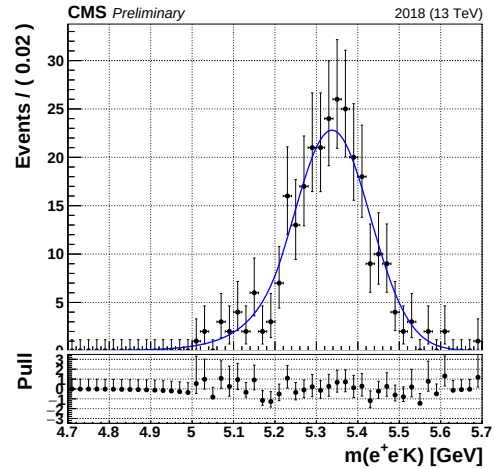
(c)



(d)

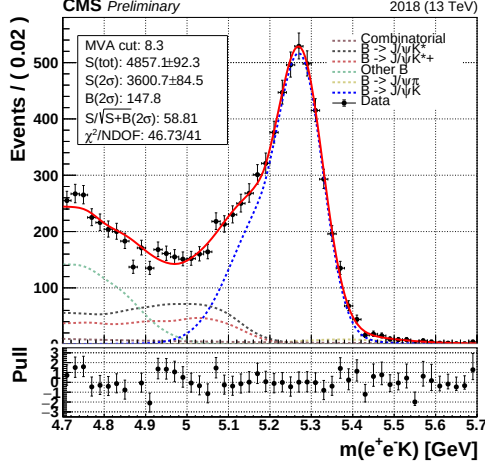


(e)

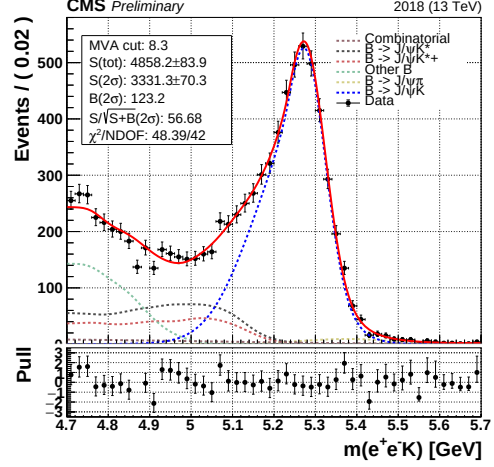


(f)

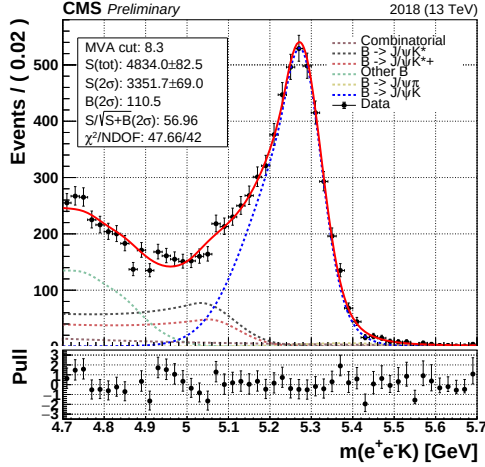
Figure D.0.2: Alternative fitting shape in the J/ψ region for signal (a), combinatorics (b), K^{*0} partially reconstructed background (c), K^{*+} partially reconstructed background (d), other B components (e), and $J/\psi\pi$ contamination (f) for PF-LP.



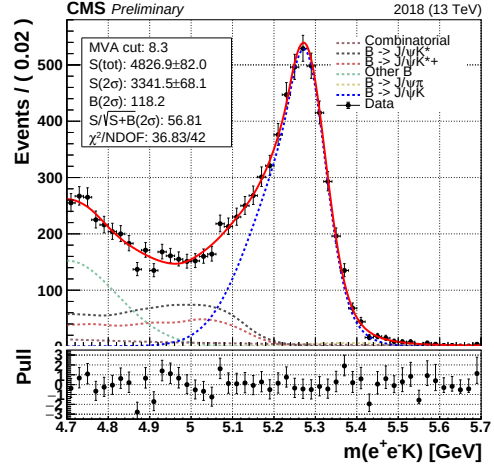
(a)



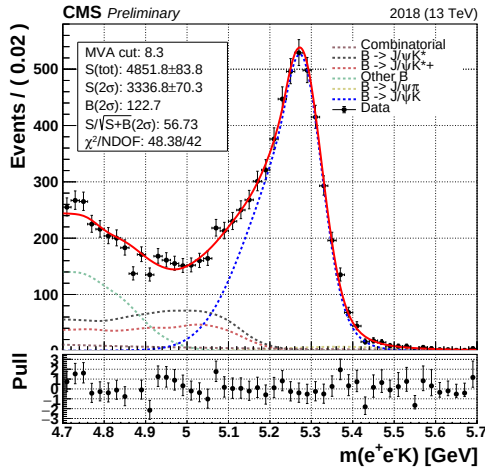
(b)



(c)

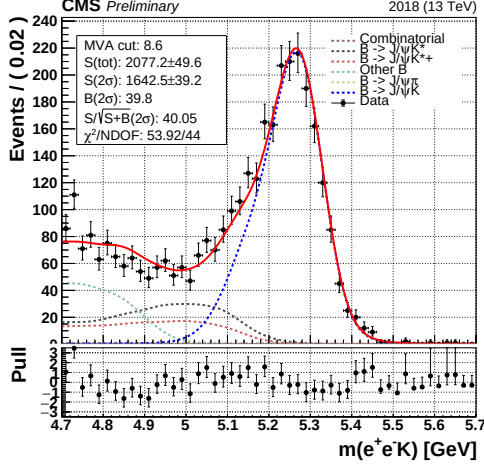


(d)

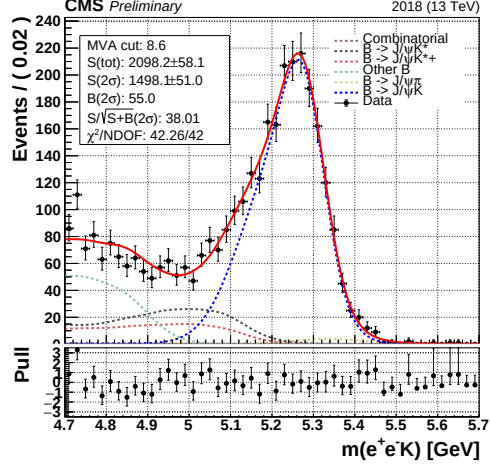


(e)

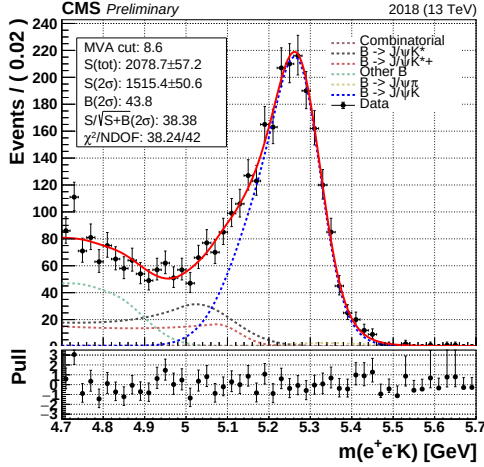
Figure D.0.3: Fits in the J/ψ region with alternative fitting shape for signal (a), combinatorics (b), partially reconstructed background (c), other B components (d), and $J/\psi\pi$ contamination (e) for PF-PF.



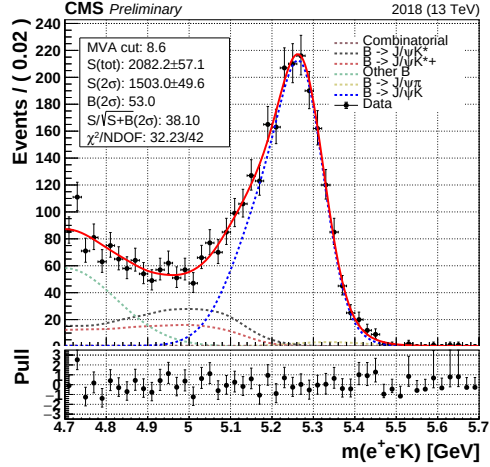
(a)



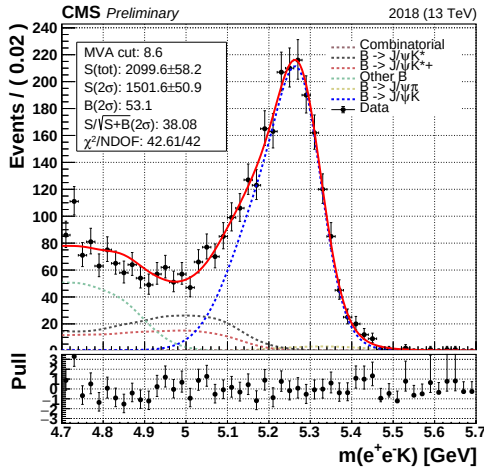
(b)



(c)

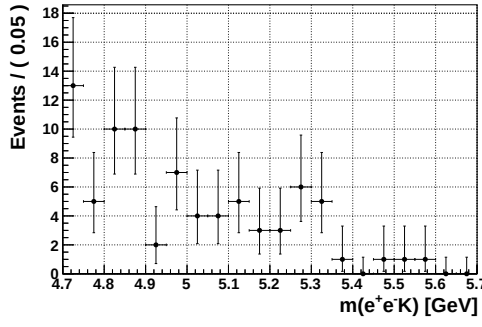


(d)

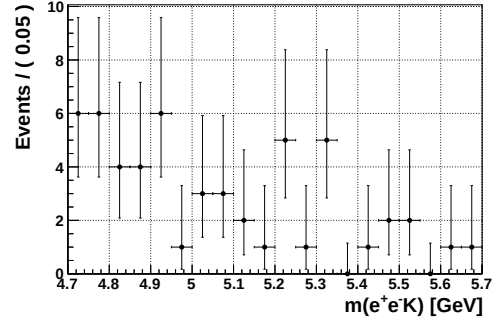


(e)

Figure D.0.4: Fits in the J/ψ region with alternative fitting shape for signal (a), combinatorics (b), partially reconstructed background (c), other B components (d), and $J/\psi\pi$ contamination (e) for PF-LP.

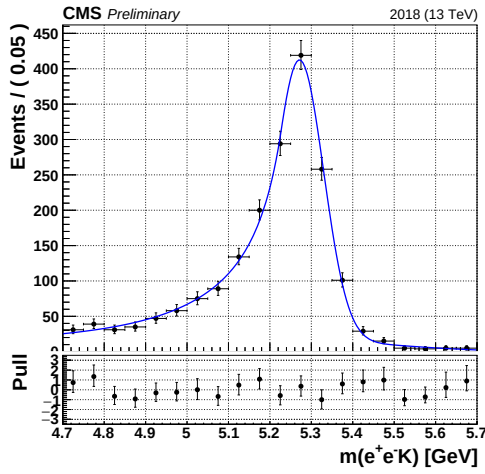


(a)

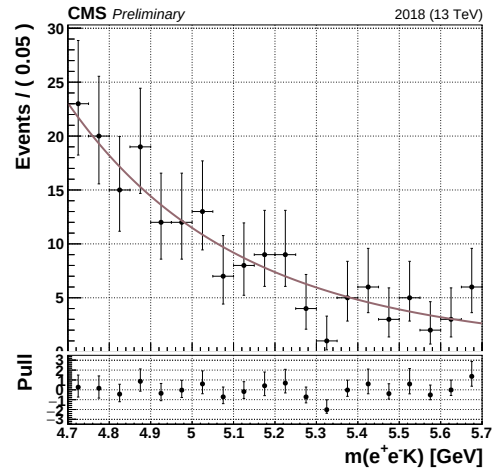


(b)

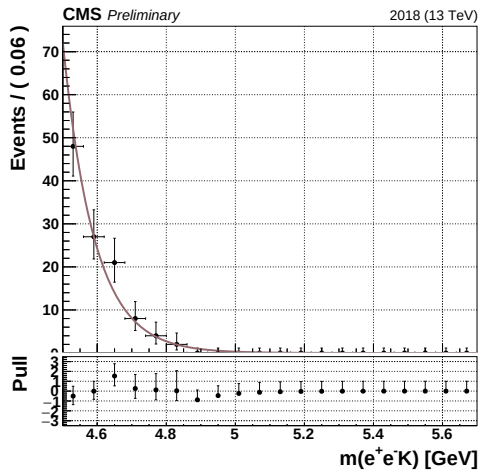
Figure D.0.5: An example of a MC toy generated in the low- q^2 region for PF-PF (a), and PF-LP (b) for the study of fit parameterization systematics.



(a)

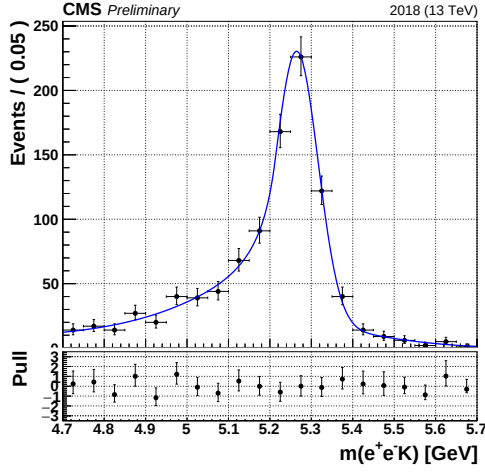


(b)

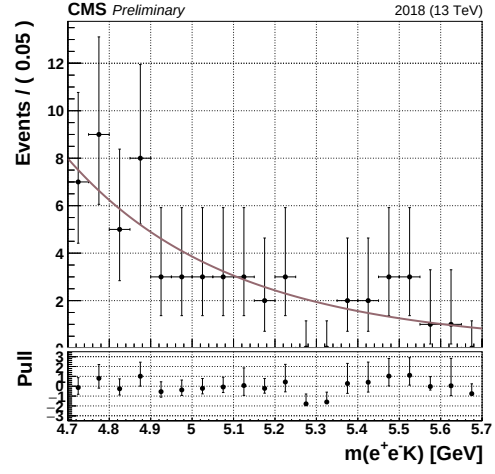


(c)

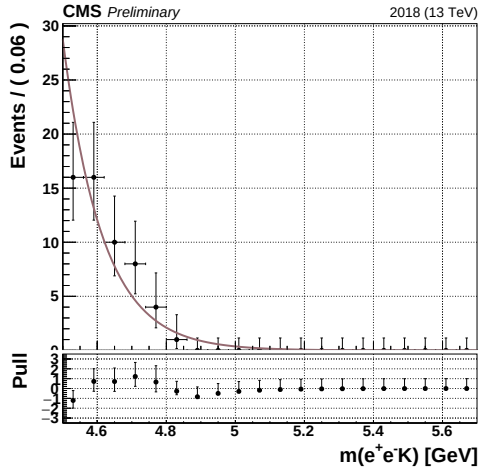
Figure D.0.6: An example of the alternative fitting shape in the low- q^2 region for signal (a), combinatorics (b), and J/ψ leakage for PF-PF.



(a)

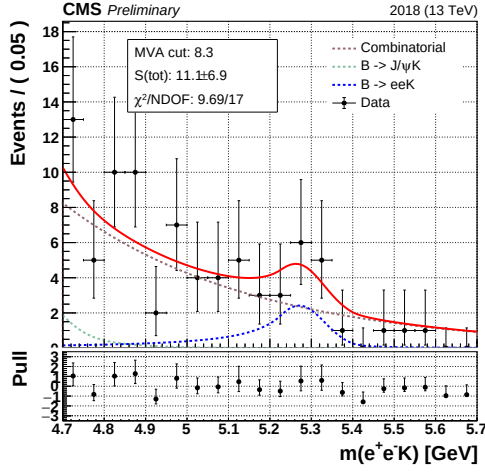


(b)

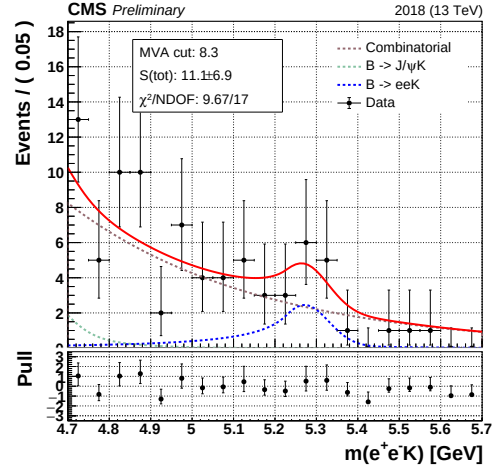


(c)

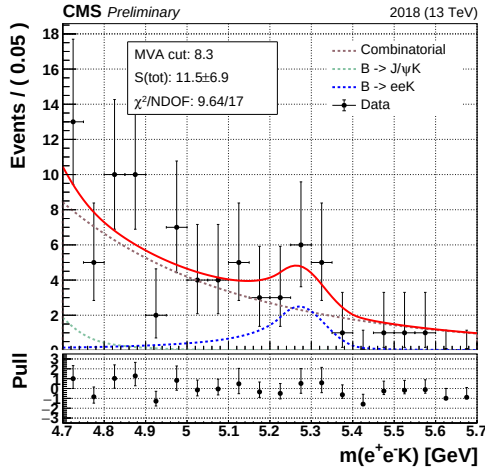
Figure D.0.7: An example of the alternative fitting shape in the low- q^2 region for signal (a), combinatorics (b), and J/ψ leakage for PF-LP.



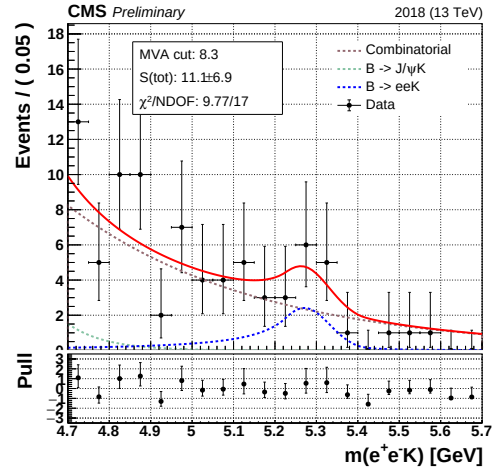
(a)



(b)

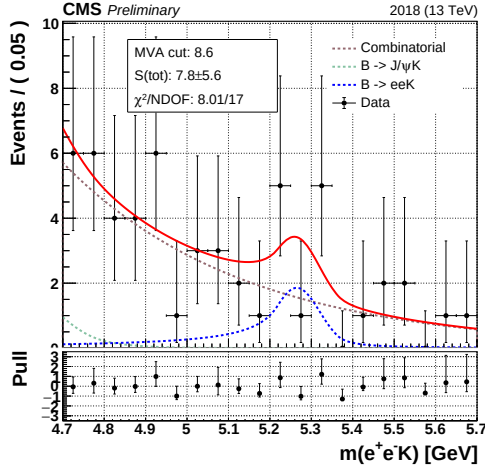


(c)

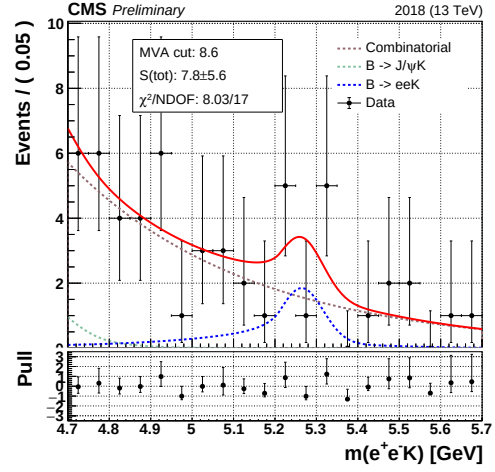


(d)

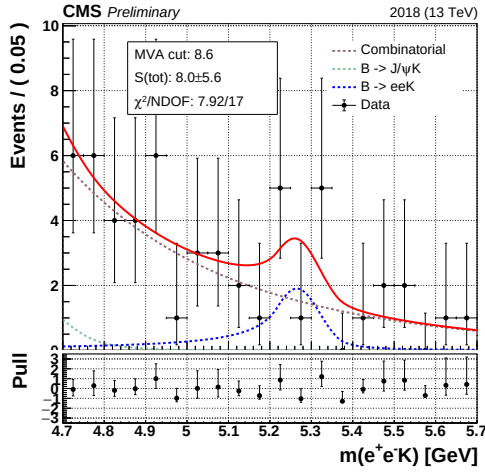
Figure D.0.8: An example of the fits in the low- q^2 region with nominal fit (a), alternative fitting shape for signal (b), combinatorics (c), and J/ψ leakage (d) for PF-PF.



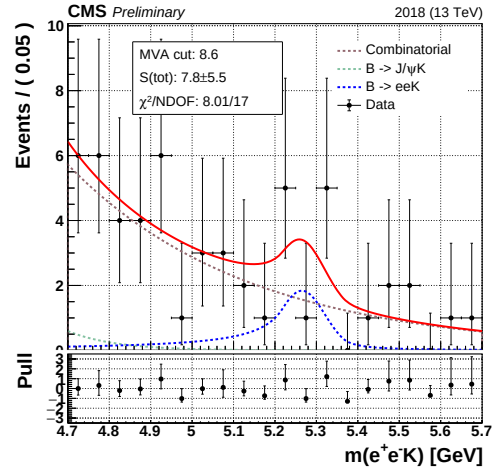
(a)



(b)



(c)



(d)

Figure D.0.9: An example of the fits in the low- q^2 region with nominal fit (a), alternative fitting shape for signal (b), combinatorics (c), and J/ψ leakage (d) for PF-LP.

Bibliography

- [1] S.L. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl. Phys.* **22** (1961) 579.
- [2] S. Weinberg, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (1967) 1264.
- [3] A. Salam, *Weak and Electromagnetic Interactions*, *Conf. Proc. C* **680519** (1968) 367.
- [4] CMS collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30 [1207.7235].
- [5] ATLAS collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1 [1207.7214].
- [6] https://www.quantumdiaries.org/wp-content/uploads/2014/03/2000px-Standard_Model_of_Elementary_Particles.svg_.jpg.
- [7] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321.
- [8] P.W. Higgs, *Broken symmetries, massless particles and gauge fields*, *Phys. Lett.* **12** (1964) 132.
- [9] P.W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508.
- [10] G.S. Guralnik, C.R. Hagen and T.W.B. Kibble, *Global Conservation Laws and Massless Particles*, *Phys. Rev. Lett.* **13** (1964) 585.
- [11] P.W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, *Phys. Rev.* **145** (1966) 1156.
- [12] S.R. Coleman and E.J. Weinberg, *Radiative Corrections as the Origin of Spontaneous Symmetry Breaking*, *Phys. Rev. D* **7** (1973) 1888.

- [13] J. Ellis, *Higgs Physics*, 2013. 10.5170/CERN-2015-004.117.
- [14] G. Buchalla, A.J. Buras and M.E. Lautenbacher, *Weak decays beyond leading logarithms*, *Rev. Mod. Phys.* **68** (1996) 1125 [[hep-ph/9512380](#)].
- [15] S. Bifani, S. Descotes-Genon, A. Romero Vidal and M.-H. Schune, *Review of Lepton Universality tests in B decays*, *J. Phys. G* **46** (2019) 023001 [[1809.06229](#)].
- [16] ALEPH, DELPHI, L3, OPAL, SLD, LEP ELECTROWEAK WORKING GROUP, SLD ELECTROWEAK GROUP, SLD HEAVY FLAVOUR GROUP collaboration, *Precision electroweak measurements on the Z resonance*, *Phys. Rept.* **427** (2006) 257 [[hep-ex/0509008](#)].
- [17] ALEPH, DELPHI, L3, OPAL, LEP ELECTROWEAK collaboration, *Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP*, *Phys. Rept.* **532** (2013) 119 [[1302.3415](#)].
- [18] PARTICLE DATA GROUP collaboration, *Review of Particle Physics*, *Phys. Rev. D* **98** (2018) 030001.
- [19] V. Cirigliano and I. Rosell, *Two-loop effective theory analysis of $\pi(K) \rightarrow e \text{ anti-}\nu/e[\gamma]$ branching ratios*, *Phys. Rev. Lett.* **99** (2007) 231801 [[0707.3439](#)].
- [20] PIENU collaboration, *Improved Measurement of the $\pi \rightarrow e\nu$ Branching Ratio*, *Phys. Rev. Lett.* **115** (2015) 071801 [[1506.05845](#)].
- [21] A. Pich, *Precision Tau Physics*, *Prog. Part. Nucl. Phys.* **75** (2014) 41 [[1310.7922](#)].
- [22] D. London and J. Matias, *B Flavour Anomalies: 2021 Theoretical Status Report*, *Ann. Rev. Nucl. Part. Sci.* **72** (2022) 37 [[2110.13270](#)].
- [23] BABAR collaboration, *Evidence for an excess of $\bar{B} \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ decays*, *Phys. Rev. Lett.* **109** (2012) 101802 [[1205.5442](#)].
- [24] LHCb collaboration, *Measurement of Form-Factor-Independent Observables in the Decay $B^0 \rightarrow K^{*0}\mu^+\mu^-$* , *Phys. Rev. Lett.* **111** (2013) 191801 [[1308.1707](#)].
- [25] LHCb collaboration, *Measurement of the ratios of branching fractions $\mathcal{R}(D^*)$ and $\mathcal{R}(D^0)$* , [2302.02886](#).
- [26] LHCb collaboration, *Measurement of the ratio of branching fractions $\mathcal{B}(B_c^+ \rightarrow J/\psi\tau^+\nu_\tau)/\mathcal{B}(B_c^+ \rightarrow J/\psi\mu^+\nu_\mu)$* , *Phys. Rev. Lett.* **120** (2018) 121801 [[1711.05623](#)].

- [27] LHCb collaboration, *Test of lepton universality in beauty-quark decays*, *Nature Phys.* **18** (2022) 277 [2103.11769].
- [28] LHCb collaboration, *Test of lepton universality with $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ decays*, *JHEP* **08** (2017) 055 [1705.05802].
- [29] LHCb collaboration, *Measurement of lepton universality parameters in $B^+ \rightarrow K^+ \ell^+ \ell^-$ and $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ decays*, 2212.09153.
- [30] LHCb collaboration, *Test of lepton universality with $\Lambda_b^0 \rightarrow p K^- \ell^+ \ell^-$ decays*, *JHEP* **05** (2020) 040 [1912.08139].
- [31] LHCb collaboration, *Tests of lepton universality using $B^0 \rightarrow K_S^0 \ell^+ \ell^-$ and $B^+ \rightarrow K^{*+} \ell^+ \ell^-$ decays*, *Phys. Rev. Lett.* **128** (2022) 191802 [2110.09501].
- [32] G. Hiller and F. Kruger, *More model-independent analysis of $b \rightarrow s$ processes*, *Phys. Rev. D* **69** (2004) 074020 [hep-ph/0310219].
- [33] M. Bordone, G. Isidori and A. Pattori, *On the Standard Model predictions for R_K and R_{K^*}* , *Eur. Phys. J. C* **76** (2016) 440 [1605.07633].
- [34] J. Haffner, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, .
- [35] O.S. Bruning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole et al., eds., *LHC Design Report Vol.1: The LHC Main Ring*, .
- [36] <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>.
- [37] T. Sakuma and T. McCauley, *Detector and Event Visualization with SketchUp at the CMS Experiment*, *J. Phys. Conf. Ser.* **513** (2014) 022032 [1311.4942].
- [38] <https://twiki.cern.ch/twiki/bin/view/CMSPublic/DPGResultsTRK>.
- [39] CMS collaboration, *Description and performance of track and primary-vertex reconstruction with the CMS tracker*, *JINST* **9** (2014) P10009 [1405.6569].
- [40] CMS collaboration, *The CMS Experiment at the CERN LHC*, *JINST* **3** (2008) S08004.
- [41] CMS collaboration, *CMS: The electromagnetic calorimeter. Technical design report*, .
- [42] CMS collaboration, *The CMS ECAL Trigger and DAQ system: electronics auto-recovery and monitoring*, 6, 2018 [1806.09136].

- [43] CMS collaboration, *Calibration of the CMS hadron calorimeters using proton-proton collision data at $\sqrt{s} = 13$ TeV*, *JINST* **15** (2020) P05002 [1910.00079].
- [44] CMS collaboration, J. Mans et al., eds., *CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter*, .
- [45] CMS collaboration, *Overview of the CMS Detector Performance at LHC Run 2*, *Universe* **5** (2019) 18.
- [46] CMS HCAL collaboration, *Design, performance, and calibration of the CMS Hadron-outer calorimeter*, *Eur. Phys. J. C* **57** (2008) 653.
- [47] CMS collaboration, *Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JINST* **13** (2018) P06015 [1804.04528].
- [48] CMS collaboration, *The CMS muon project: Technical Design Report*, .
- [49] CMS collaboration, *Upgrade of the trigger system of CMS*, *Nucl. Instrum. Meth. A* **718** (2013) 11.
- [50] CMS collaboration, *Particle-flow reconstruction and global event description with the CMS detector*, *JINST* **12** (2017) P10003 [1706.04965].
- [51] W. Adam, R. Fruhwirth, A. Strandlie and T. Todorov, *Reconstruction of electrons with the Gaussian sum filter in the CMS tracker at LHC*, *eConf C0303241* (2003) TULT009 [physics/0306087].
- [52] LHCb collaboration, *Search for lepton-universality violation in $B^+ \rightarrow K^+ \ell^+ \ell^-$ decays*, *Phys. Rev. Lett.* **122** (2019) 191801 [1903.09252].
- [53] LHCb collaboration, *Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \rightarrow D^{*+} \tau^- \bar{\nu}_\tau) / \mathcal{B}(\bar{B}^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu)$* , *Phys. Rev. Lett.* **115** (2015) 111803 [1506.08614].
- [54] LHCb collaboration, *Test of lepton universality using $B^+ \rightarrow K^+ \ell^+ \ell^-$ decays*, *Phys. Rev. Lett.* **113** (2014) 151601 [1406.6482].
- [55] T. Sjöstrand, S. Ask, J.R. Christiansen, R. Corke, N. Desai, P. Ilten et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [1410.3012].
- [56] D.J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth. A* **462** (2001) 152.

- [57] CMS collaboration, *Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements*, *Eur. Phys. J. C* **80** (2020) 4 [1903.12179].
- [58] NNPDF collaboration, *Parton distributions from high-precision collider data*, *Eur. Phys. J. C* **77** (2017) 663 [1706.00428].
- [59] T. Chen and C. Guestrin, *XGBoost: A Scalable Tree Boosting System*, 1603.02754.
- [60] CMS collaboration, *Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC*, *JINST* **16** (2021) P05014 [2012.06888].
- [61] M. Pivk and F.R. Le Diberder, *SPlot: A Statistical tool to unfold data distributions*, *Nucl. Instrum. Meth. A* **555** (2005) 356 [physics/0402083].
- [62] S. Das, *A simple alternative to the Crystal Ball function*, 1603.08591.
- [63] K. Prokofiev and T. Speer, *A kinematic and a decay chain reconstruction library*, in *14th International Conference on Computing in High-Energy and Nuclear Physics*, pp. 411–414, 2005.
- [64] D.M. Straub, *flavio: a Python package for flavour and precision phenomenology in the Standard Model and beyond*, 1810.08132.