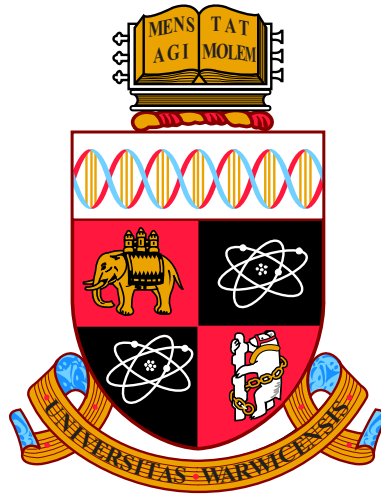




Testing lepton flavour universality with
 $B_s^0 \rightarrow \phi \ell^+ \ell^-$ decays and angular analysis of the
 $B_s^0 \rightarrow \phi e^+ e^-$ decay with the LHCb experiment

by

Lorenzo Paolucci

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy in Physics

Department of Physics

April 2025



Contents

List of Tables	v
List of Figures	vii
Acknowledgments	xiii
Declarations	xv
Abstract	xvi
Chapter 1 Introduction	1
Chapter 2 Theoretical background	3
2.1 The Standard Model of particle physics	3
2.1.1 From symmetries to interactions	5
2.1.2 A mass(ive) issue	7
2.1.3 The matter content	9
2.1.4 Intermezzo: Feynman diagrams and amplitudes	11
2.2 Flavour physics phenomenology	13
2.2.1 CP violation	14
2.2.2 Flavour changing processes	15
2.2.3 Neutral meson mixing	15
2.3 Limits of the Standard Model	17
2.3.1 Flavour anomalies	18
2.4 A model independent approach to $b \rightarrow s \ell^+ \ell^-$ processes	20
2.4.1 Writing down the amplitudes	22
2.4.2 Total angular decay rate	27
Chapter 3 The LHCb detector at the Large Hadron Collider	31
3.1 The Large Hadron Collider	31
3.2 Overview of the LHCb detector	34
3.3 Particle tracking	36

3.3.1	The Vertex Locator	37
3.3.2	The tracking stations	39
3.3.3	The dipole magnet	40
3.3.4	Vertex and track reconstruction	41
3.4	Particle identification	45
3.4.1	The RICH detectors	45
3.4.2	The calorimeters	48
3.4.3	The muon system	50
3.4.4	PID performance	52
3.5	The LHCb trigger system and data flow	54
3.6	The upgraded LHCb detector	57
3.6.1	The upgraded Vertex Locator	59
3.6.2	Operation of the upgraded Vertex Locator	61
3.6.3	Tracking	62
3.6.4	Particle identification	64
3.6.5	Trigger system	64
3.7	Monitoring the upgraded muon detector's performance	65
Chapter 4 Selection of candidates and simulation of $B_s^0 \rightarrow \phi \ell^+ \ell^-$ decays		70
4.1	Trigger selection	71
4.2	Electron reconstruction	72
4.3	Offline selection and background rejection	73
4.4	Definition of regions of dilepton mass squared	76
4.5	Definition of the blinding regions	76
4.6	Calibration of simulated samples	77
4.6.1	Particle identification	77
4.6.2	Electron track reconstruction	78
4.6.3	Beauty meson kinematics and event multiplicity	79
4.6.4	Trigger	79
4.6.5	Vertexing	80
4.6.6	Dielectron mass	80
4.6.7	Beauty meson lifetime	81
4.7	Multivariate selection	81
4.7.1	Combinatorial BDT	82
4.7.2	High q^2 BDT	83
4.7.3	Selection optimisation	84
4.8	Residual background contamination	86
4.8.1	Semileptonic double cascades	86
4.8.2	Charmonium leakage	87
4.8.3	Misidentified backgrounds	87

Chapter 5	Test of lepton flavour universality using $B_s^0 \rightarrow \phi \ell^+ \ell^-$ decays	90
5.1	Measurement strategy	91
5.2	The high- q^2 bin	92
5.3	Efficiency computation	99
5.4	Cross checks and validation	101
5.4.1	Measurement of $r_{J/\psi}$ and $R_{\psi(2S)}$	101
5.4.2	Pseudoexperiment studies	104
5.5	Mass fits	107
5.6	Systematic uncertainties	109
5.7	Results	112
Chapter 6	Angular analysis of the rare decay $B_s^0 \rightarrow \phi e^+ e^-$	115
6.1	Measurement strategy	116
6.2	Modelling the angular acceptance and resolution	118
6.2.1	Model description	118
6.2.2	Integration of the angular acceptance	121
6.2.3	Modelling semileptonic vetoes	122
6.2.4	Angular and q^2 resolution	124
6.3	Component modelling	127
6.4	Validation of the angular acceptance using control modes	130
6.4.1	Result for the $B_s^0 \rightarrow J/\psi (e^+ e^-) \phi$ sample	131
6.4.2	Result for the $B_s^0 \rightarrow \psi(2S) (e^+ e^-) \phi$ sample	132
6.4.3	Conversion to transversity amplitudes	133
6.5	Fit strategy validation	136
6.5.1	Angular fits of signal Monte Carlo	136
6.5.2	Fit validation with pseudoexperiments	139
6.6	Systematic uncertainties	142
6.6.1	Summary of systematic uncertainties	147
6.7	Results	148
Chapter 7	The LHCb Early Career, Gender and Diversity office	155
7.1	Structure of the LHCb collaboration and the ECGD office	155
7.2	Activities of the ECGD office	156
7.2.1	LHC wide activities	157
Chapter 8	Conclusions	159
Appendix A	Normalisation of the correction weights for efficiency computation	162
Appendix B	Projections of the $B_s^0 \rightarrow \phi e^+ e^-$ angular model including acceptance effects	164

Appendix C	Time-integrated observables as functions of transversity amplitudes	167
Appendix D	Estimating parameter confidence intervals	169
Appendix E	Detailed breakdown of systematic uncertainties on the angular observables	172
Bibliography		176

List of Tables

2.1	Quantum numbers of gauge fields before symmetry breaking.	10
2.2	Quantum numbers of boson fields after symmetry breaking.	10
2.3	Quantum numbers of fermion fields.	11
3.1	Different L0 trigger thresholds used in each year of Run 2 of data-taking. . .	56
4.1	Definitions of the q^2 -regions for the rare $B_s^0 \rightarrow \phi \ell^+ \ell^-$ decays and the resonant $B_s^0 \rightarrow J/\psi (\ell^+ \ell^-) \phi$ and $B_s^0 \rightarrow \psi(2S) (\ell^+ \ell^-) \phi$ decays.	76
4.2	Definitions of the blinded ranges in the different q^2 regions.	77
5.1	Reconstruction and selection efficiency in permille for signal simulation samples in the L0I trigger category.	100
5.2	Reconstruction and selection efficiency in permille for signal simulation samples in the L0L! trigger category.	100
5.3	Definitions of the $m_{J/\psi}(K^+ K^- \ell^+ \ell^-)$ mass ranges used for the fits to the data.	101
5.4	Definitions of the $m(K^+ K^- \ell^+ \ell^-)$ mass ranges used for the fits to the data in the three q^2 bins.	107
5.5	Measured yields of $B_s^0 \rightarrow \phi \mu^+ \mu^-$ and $B_s^0 \rightarrow \phi e^+ e^-$ decays in the low-, central-, and high- q^2 bins.	109
5.6	Summary of the sources of systematic uncertainty on the R_ϕ^{-1} measurement in the low-, central- and high- q^2 bins.	110
6.1	Predicted Standard Model values for the time- and q^2 -integrated values of F_L and S_3 , including the effects of q^2 -smearing.	118
6.2	Reference values of the squared magnitude of the transversity values measured by LHCb.	131
6.3	Fitted values of the angular observables from the mass-angle fits for $B_s^0 \rightarrow J/\psi (e^+ e^-) \phi$ candidates.	132
6.4	Fitted values of the angular observables from the mass-angle fits for $B_s^0 \rightarrow \psi(2S) (e^+ e^-) \phi$ candidates.	134

6.5	Comparison of the values of the transversity amplitudes $ A_0 ^2$ and $ A_\perp ^2$ as measured in this work with the reference $B_s^0 \rightarrow J/\psi (\mu^+ \mu^-) \phi$ measurement. .	136
6.6	Comparison of the values of the transversity amplitudes $ A_0 ^2$ and $ A_\perp ^2$ as measured in this work with the reference $B_s^0 \rightarrow \psi(2S) (\mu^+ \mu^-) \phi$ measurement.	137
6.7	Comparison of generator and reconstruction level angular fits.	139
6.8	Values of the angular observables used to generate the pseudoexperiments. . .	140
6.9	Summary table of the systematic uncertainties on each angular observable. .	148
6.10	Constraints entering the final angular fit.	150
6.11	Parameters determined from the $(\cos \theta_K, \cos \theta_e)$ fit to $B_s^0 \rightarrow \phi e^+ e^-$ data. . . .	150
6.12	Resulting parameters for the Φ fit to $B_s^0 \rightarrow \phi e^+ e^-$ data.	151
6.13	Measured values of the angular observables in the three q^2 bins.	151
E.1	Summary of the sources of systematic uncertainties in the low- q^2 bin.	173
E.2	Summary of the sources of systematic uncertainties in the central- q^2 bin. . . .	174
E.3	Summary of the sources of systematic uncertainties in the high- q^2 bin.	175

List of Figures

2.1	Feynman diagrams for Bhabha scattering.	12
2.2	Feynman diagrams for the $b \rightarrow s\ell^+\ell^-$ transition.	15
2.3	Feynman diagrams for the mixing of neutral B mesons.	16
2.4	Experimental status of LFU ratios in $b \rightarrow c\ell^-\bar{\nu}_\ell$ and $b \rightarrow s\ell^+\ell^-$ decays.	19
2.5	Current tensions, with the SM prediction, of the angular observable P'_5 and the differential branching fraction of $B_s^0 \rightarrow \phi\mu^+\mu^-$ decays.	20
2.6	Feynman diagram for the muon decay in the Standard Model and in the effective field theory description.	21
2.7	Result of global fit to $b \rightarrow s\ell^+\ell^-$ observables for the Wilson coefficients $\mathcal{C}_{9,10}$	23
2.8	Definition of the decay angles in the $B_s^0 \rightarrow \phi\ell^+\ell^-$ decay.	25
2.9	Feynman diagram for the decay $B_s^0 \rightarrow \phi\ell^+\ell^-$, decorated with effects from QCD.	27
3.1	The CERN accelerator complex as of 2020.	33
3.2	Production of $b\bar{b}$ pairs for simulated pp collisions at centre-of-mass energy $\sqrt{s} = 14$ TeV.	35
3.3	Example of the instantaneous luminosity profile for the ATLAS, CMS, and LHCb detectors.	36
3.4	View of the LHCb detector and its subcomponents.	36
3.5	Schematics of the Vertex Locator.	38
3.6	Exploded view of the VELO supports and RF box.	38
3.7	Scheme for the silicon strips in one layer of the Tracker Turicensis.	39
3.8	Scheme for the silicon sensors in one layer of the Inner Tracker.	40
3.9	Technical drawing of the Outer Tracker's layers and modules.	40
3.10	Technical drawing of the LHCb dipole magnet and its support structure.	41
3.11	Profile of the LHCb dipolar magnetic field along the z axis.	42
3.12	Tracking efficiency measured by LHCb with 2011 and 2012 control samples, as a function of kinematic variables.	44
3.13	Primary vertex resolution measured by LHCb in 2012, as a function of track multiplicity.	44

3.14	Measured Cherenkov angle for isolated tracks in one of the Ring Cherenkov detectors, as a function of the track's momentum.	46
3.15	Schematics of the Ring Cherenkov detectors.	47
3.16	Display of a typical collision at LHCb as reconstructed by the Ring Cherenkov detector.	47
3.17	Layout of the LHCb calorimeter system.	48
3.18	Electron and pion separation in the energy released in the preshower detector.	49
3.19	Schematics of the cells in the electromagnetic and hadronic calorimeters.	50
3.20	Side view of the muon system.	51
3.21	Schematics of a muon station, and multiwire chamber layout.	51
3.22	Readout scheme in a chamber of the second muon station.	52
3.23	Comparison of kaon identification and misidentification performance in data and simulation.	53
3.24	Muon identification and misidentification performance using the IsMuon binary variable.	54
3.25	Scheme of the trigger system that was implemented in Run 2 of LHCb.	55
3.26	LHCb data and simulation flow in Run2, with individual steps and software shown.	57
3.27	Relative yields of key physics channels as a function of instantaneous luminosity.	58
3.28	Layout of the upgraded LHCb detector.	59
3.29	Layout of the upgraded VELO modules and of the active sensors on a single module.	60
3.30	Layout and X-ray image of the microchannels used to cool the upgraded VELO modules.	61
3.31	Picture of one upgraded VELO side before installation into the vacuum vessel.	61
3.32	The upgraded VELO's Run Control panel.	63
3.33	Layouts of the Upstream Tracker and SciFi trackers.	63
3.34	Comparison of trigger efficiencies between the original LHCb hardware trigger and the upgraded software-based trigger system.	65
3.35	Example page of the MONET monitoring presenter, showing the (x, y) hit distribution in the muon stations M2 and M3.	67
3.36	Example distribution of hits in the muon detector for time alignment events.	68
3.37	Example output of the detection efficiency monitor for a pp run collected by LHCb in 2025.	69
4.1	Comparison of the efficiency of the L0Muon and L0Electron trigger lines.	72
4.2	Sketch of the bremsstrahlung radiation emitted by an electron in the LHCb detector.	73
4.3	Two dimensional distributions of the swap-mass of the kaon-lepton system and its combined PID response, for simulated signal and background decays.	75

4.4	Scheme of the correction chain to improve agreement between simulated samples and data.	77
4.5	Receiver operating characteristic curves for the combinatorial BDT trained for the muon and electron datasets, under 2017–2018 data-taking conditions.	83
4.6	Receiver operating characteristic curve for the high- q^2 BDT trained for electron data.	84
4.7	Result of the optimisation of the combinatorial and high- q^2 BDT selections for electron data taken in the 2015–2016 period.	86
4.8	Two dimensional distribution of reconstructed B_s^0 mass and dielectron mass squared for simulated $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ and $B_s^0 \rightarrow \psi(2S) (e^+e^-) \phi$ decays.	87
5.1	Observed two-dimensional distribution of the reconstructed dilepton mass squared and B_s^0 mass in LHCb data.	93
5.2	Result of unbinned maximum likelihood fits to the $m(K^+K^-\ell^+\ell^-)$ distributions of simulated signal muon and electron candidates in the high- q^2 bin.	94
5.3	Sketch of the prescription to build a proxy combinatorial-like sample made of recombination of random tracks.	95
5.4	Invariant B_s^0 mass distribution for fake candidates built from the recombination of random tracks from fully selected $B_s^0 \rightarrow \phi e^+e^-$ data.	96
5.5	Result of the maximum likelihood fit to the OS data q^2 spectrum in the electron upper mass sideband $m(K^+K^-e^+e^-) \in [5500, 6500] \text{ MeV}/c^2$	96
5.6	Illustration of the generated proxy sample in $(q^2, m(K^+K^-e^+e^-))$	97
5.7	Result of unbinned maximum likelihood fits to same-sign (SS) collision data for the muon (left) and electron (right) modes with the empirical model.	98
5.8	Mass acceptance functions for combinatorial muon and electron pseudodata.	98
5.9	Result of the maximum likelihood fits to SS muon and electron data with the acceptance corrected exponential.	99
5.10	Correlation matrix of the computed efficiencies, across all q^2 bins, trigger categories, and lepton final states, for events simulated in 2017–2018 data-taking conditions.	101
5.11	Result of maximum likelihood fits to the $m_{J/\psi}(K^+K^-\ell^+\ell^-)$ distributions for $B_s^0 \rightarrow J/\psi (\ell^+\ell^-) \phi$ candidates collected in Run 2p2. On the left, the muon mode; on the right, the electron mode.	102
5.12	Result of maximum likelihood fits to the $m_{\psi(2S)}(K^+K^-\ell^+\ell^-)$ distributions for $B_s^0 \rightarrow \psi(2S) (\ell^+\ell^-) \phi$ candidates collected in Run 2p2.	103
5.13	Differential measurement of $r_{J/\psi}$ in bins of the opening angle between the two electrons, and the minimum p_T of the electrons.	104
5.14	Evolution of the measurement of $r_{J/\psi}$ and $R_{\psi(2S)}$ with different corrections w applied to the simulation datasets used to compute the selection efficiencies.	105
5.15	Distribution of fitted values of R_ϕ^{-1} and pull distribution over an ensemble of 500 pseudoexperiments generated in the high- q^2 bin.	106

5.16	Fits to the $B_s^0 \rightarrow \phi\mu^+\mu^-$ and $B_s^0 \rightarrow \phi e^+e^-$ invariant mass distributions in the low- q^2 bin.	107
5.17	Fits to the $B_s^0 \rightarrow \phi\mu^+\mu^-$ and $B_s^0 \rightarrow \phi e^+e^-$ invariant mass distributions in the central- q^2 bin.	108
5.18	Fits to the $B_s^0 \rightarrow \phi\mu^+\mu^-$ and $B_s^0 \rightarrow \phi e^+e^-$ invariant mass distributions in the high- q^2 bin.	109
5.19	Profile log-likelihood scan of R_ϕ^{-1} in the low-, central-, and high- q^2 bins. . . .	112
5.20	Scan of the measured value of R_ϕ^{-1} in the low- q^2 bin for different BDT working points.	113
5.21	Differential branching fraction of the $B_s^0 \rightarrow \phi e^+e^-$ decay, compared to the Standard Model prediction.	114
6.1	Result of the parametrisation of the angular acceptance obtained for simulated $B_s^0 \rightarrow \phi e^+e^-$ candidates.	120
6.2	Result of the parametrisation of the angular acceptance obtained for simulated $B_s^0 \rightarrow \phi e^+e^-$ candidates, in three separate q^2 bins.	121
6.3	Effect of the $D_s^\pm$ veto applied in the nominal selection in the q^2 - $\cos\theta_e$ plane, for $\cos\theta_e > 0$	123
6.4	Result of the fits to the efficiency distribution in $\cos\theta_e$ of the $D_s^\pm$ vetos. . . .	123
6.5	Example fits to measure the angular resolution in a bin of q^2	124
6.6	Distributions of the angular resolution as a function of q^2	124
6.7	Distributions of the angular resolution as function of the reconstructed angle, in three q^2 bins.	125
6.8	Example resolution fit in the low- q^2 bin near the edge of the θ_e range. . . .	126
6.9	Migration probability density matrix $M(q^2 q_{\text{true}}^2)$	126
6.10	Parameter boundary for the angular decay rate in $\cos\theta_K$ and $\cos\theta_e$	128
6.11	Comparison of the combinatorial angular parameters as fitted for Opposite Sign and Same Sign candidates.	130
6.12	Mass and angle projections of the $B_s^0 \rightarrow J/\psi(e^+e^-)\phi$ upper mass sideband fit.	132
6.13	Result of the maximum likelihood fit to the mass and angle distribution of $B_s^0 \rightarrow J/\psi(e^+e^-)\phi$ candidates.	133
6.14	Mass and angle projections of the $B_s^0 \rightarrow \psi(2S)(e^+e^-)\phi$ upper mass sideband fit.	134
6.15	Result of the maximum likelihood fit to the mass and angle distribution of $B_s^0 \rightarrow \psi(2S)(e^+e^-)\phi$ candidates.	135
6.16	Result of the fit to simulated $B_s^0 \rightarrow \phi e^+e^-$ decays to model the decay time acceptance.	136
6.17	Result of the χ^2 fit to the fitted values of $\langle F_L \rangle$ and $\langle S_3 \rangle$, to measure $ A_0 ^2$ and $ A_\perp ^2$	137
6.18	Result of the generator level angular fits of simulated $B_s^0 \rightarrow \phi e^+e^-$ candidates.	138

6.19	Result of the reconstruction level angular fits of simulated $B_s^0 \rightarrow \phi e^+ e^-$ candidates.	138
6.20	Pull distribution from the angular pseudoexperiment studies in the central- q^2 bin.	141
6.21	Correlation between angular observables as observed in pseudoexperiments. .	142
6.22	Result of the angular fits to $B_s^0 \rightarrow \phi e^+ e^-$ data.	149
6.23	Result of the Feldman–Cousins procedure for the angular observables.	152
6.24	Result of the Feldman–Cousins procedure in two dimensions for the angular observables $\langle F_L \rangle$ and $\langle A'_6 \rangle$	153
6.25	Result of the Feldman–Cousins procedure in two dimensions for the angular observables $\langle S_3 \rangle$ and $\langle A_9 \rangle$	153
6.26	Result in bins of q^2 of the measurement of the time-integrated observables $\langle F_L \rangle$, $\langle S_3 \rangle$, $\langle A'_6 \rangle$, and $\langle A_9 \rangle$	154
7.1	Number of participants in the CERN mentoring schemes in the 2018–2025 period.	158

*A Nonna,
con cui avrei voluto tanto celebrare questo traguardo.*

Acknowledgments

Here we are, at the end of this long journey. While fluent in two languages, it is hard to put to words the gratitude I feel towards the people who surrounded me throughout the years. I shall try here my very best.

I think it is difficult to understate how lucky I have been to have been supervised by Tim and Tom. Your breadth of knowledge has never ceased to amaze me, and yet I have never felt less than because of it. You taught me rigor (both in my methods and in my scientific writing!), showed me unwavering support, and for that I will always be thankful.

Moving to the United Kingdom for this PhD was a big and scary step, leaving home for the first time and all the people that I care about. I want to thank all of the friends who I have met on this journey and who have made me feel at home even when I was not. To my Warwick cohort, Emir, Mo, Matthew, and Eliot: thank you for all the climbing, dinners, walks, teas, drinks and everything in between, and for endlessly listening to me complain about work. Anja, thank you so much for being unapologetically you. You might be the happiest PhD student in history, but I would like to think I have taken second place, and it is in no small part because of you. To the rest of my Warwick EPP friends, Aidan, Eva, Alex, and Matthew, thank you for the pub nights and dinners, chatting, gossiping and discussing everything and anything. Silvia, Veronica, Neil, I am so deeply grateful to have met you after that rainy campus tour in first year, for always offering a needed breather from the world of physics, and for all the soirées full of wine and cheese. To Trent, I'm thankful for all coffees we have had to break from work, which have definitely helped me keep sane, especially through writing. I'd also like to thank the friends I have met during my LTA at CERN, the Parisian crew, Chiara, Gaëlle and Renaud, and the queer girlies, Joe, Alina, Charlotte, Victoria and Nathan.

Vorrei fare anche dei ringraziamenti in italiano. Inizio ringraziando Barbara, che ormai ha preso un po' il ruolo di mia "fata madrina" della fisica. Mi hai seguito fin da prima del dottorato, e hai offerto supporto, consiglio, e continui spunti di riflessione; ti devo moltissimo. Mamma, Papà, Andrew, Zia Mariateresa, grazie per essermi sempre stati vicini in questo viaggio, anche se spesso separati da centinaia di chilometri di distanza; spero di avervi resi fieri. A tutti i miei amici romani, sapete già quanto sia grato della vostra amicizia. Chiara, Sitzietta, Giordano, Alice, Giulia, Sara, la mia Joyeuse Famille, grazie per non avermi mai fatto percepire i mesi passati tra una visita a Roma e l'altra. A Federico, grazie per avermi sopportato tutti questi anni. Grazie a Giulio, Matteo, Zeno e Carmine, i miei oomfs per eccellenza. Infine, grazie a Giulia e ad Ilaria, le mie rocce, senza le quali sarei completamente perso.

I was anxious of getting to the point of writing this part of the thesis and not knowing what to say, and yet now that I am rereading it, it feels just right. Now, onto the science!

Declarations

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy in Physics. It has been composed by myself and has not been submitted in any previous application for any degree.

The data used for the analyses presented in Chapters 5 and 6 were collected between 2011 and 2018 by the LHCb experiment, at the European Organisation for Nuclear Research, in Geneva, Switzerland. The work presented in Chapter 4 was mostly carried out by collaborators before I began my PhD. It is presented here as it is necessary to explain the offline processing of the data used in the following chapters. All simulated samples used in this thesis were also produced by collaborators. The work in Chapter 5 was partially carried out by me within a larger collaborative effort during the first two years of my PhD. In particular, I worked on the region of high-dilepton mass squared, on part of the efficiency computation and on the sensitivity studies with pseudoexperiments. I am directly responsible for all of the content presented in Chapter 6, which was produced under the direction of my supervisors.

The content of Chapter 3 contains a review of the upgraded LHCb detector, which was planned and constructed before I joined the collaboration. I contributed to the commissioning and early operations of the upgraded detector. I am the author and maintainer of the monitoring algorithms presented in the chapter. My work as one of the four officers of the Early Career, Gender, and Diversity office, presented in Chapter 7, was part of a collaborative effort.

Finally, the contents of Chapters 5 and 6 are published in the following:

- LHCb collaboration, R. Aaij *et al.*, *Test of lepton flavour universality with $B_s^0 \rightarrow \phi \ell^+ \ell^-$ decays*, Phys. Rev. Lett. **134** (2025) 121803, [arXiv:2410.13748](#).
- LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the decay $B_s^0 \rightarrow \phi e^+ e^-$* , JHEP **07** (2025) 069, [arXiv:2504.06346](#).

Abstract

This thesis describes two measurements of properties of the decay process $B_s^0 \rightarrow \phi \ell^+ \ell^-$, where $\ell^\pm$ is either an electron or muon, using data collected by the LHCb experiment at the European Organisation for Nuclear Research in Geneva (CERN), Switzerland. The first is the first measurement of the branching fraction ratio $R_\phi = \mathcal{B}(B_s^0 \rightarrow \phi \mu^+ \mu^-) / \mathcal{B}(B_s^0 \rightarrow \phi e^+ e^-)$, in three bins of dilepton mass squared. Such a ratio tests an accidental symmetry of the Standard Model of particle physics, known as lepton flavour universality. This symmetry states that all lepton flavours have equal interaction strengths in the Standard Model, except for effects related to their different masses. The results of the analysis are found to be compatible with lepton flavour universality, which means that the $B_s^0 \rightarrow \phi e^+ e^-$ branching fractions are mostly below the Standard Model expectations, in the same way as those for $B_s^0 \rightarrow \phi \mu^+ \mu^-$. The second measurement is the first angular analysis of the $B_s^0 \rightarrow \phi e^+ e^-$ decay. Angular observables are measured by fitting the angular distribution, in the same three bins of dilepton mass squared as in the measurement of R_ϕ . The measured observables are compared to previous LHCb measurements performed on the muon channel, to further test lepton flavour universality. The results are found to be compatible with the Standard Model as well as with previous measurements of the $B_s^0 \rightarrow \phi \mu^+ \mu^-$ decay. Furthermore, this thesis presents work done towards the commissioning of the upgraded LHCb detector, which was installed between 2018 and 2022. In particular, the thesis focuses on the upgraded vertex locator and muon detectors. Finally, the thesis includes work done as part of the LHCb Early Career, Gender and Diversity office, which aims at the improvement of the working environment of all LHCb members. The author filled the role of Early Career representative for two years, between 2023 and 2025, and has contributed to many activities both within the office and in collaboration with the other main LHC experiments.

Introduction

The author's recommendation

How Big, How Blue, How Beautiful

by **Florence and The Machine**

2015, Island

This thesis is the culmination of my PhD journey, which has lasted about three and a half years, from October 2021 to April 2025. I have specialised in the field of particle physics, the study of the most basic building blocks of the Universe. In particular, I have concentrated my research efforts on a special type of particle known as a *beauty quark*. Ever since its discovery in 1977, this particle has been the focus of multiple experiments, and studied by thousands of researchers across the globe. Its existence represents another piece in the large puzzle that led to the formulation the Standard Model, our most complete theory yet of fundamental particles and their interactions.

Like many PhD theses, the work presented here covers a highly specific area of research. In particular, it concerns itself with the decay of a B_s^0 meson, a particle containing a beauty quark, into a ϕ meson and a pair of leptons, $\ell^+\ell^-$ (either electrons or their more massive cousins, muons). These are a type of decays which are particularly rare in the Standard Model, and which can provide important insight towards the discovery of new physics phenomena. An introduction to the mathematical structure of the Standard Model, as well as to the phenomenology of these rare processes, is presented in Chapter 2.

A core principle of the field of particle physics has always been the collaboration between researchers. As the designs for instruments required to study fundamental particles became more and more complicated, the number of people involved in a single experiment grew exponentially. My research is based on data collected by the LHCb experiment, which is one of the four main experiments located on the Large Hadron Collider at the European Organisation for Nuclear Research – the biggest particle physics laboratory in the world. The LHCb collaboration, as of 2025, is composed of more than 1700 scientists, who collect and analyse the data. The detector which collected the data I have used in this thesis, as it operated in the 2011–2018 period, is described in Chapter 3.

The LHCb detector underwent significant upgrades between 2018 and 2022, aimed at maximising the physics output of the experiment. During my PhD I had the chance to contribute to the operation and commissioning of the new detector components, focussing in particular on two sub-detectors: the vertex locator, and the muon detector. In Chapter 3 I will also cover my contributions to the upgraded LHCb detector.

In Chapter 4 I will introduce how, among the billions of events that make up the data collected by LHCb, it is possible to efficiently select candidates for the $B_s^0 \rightarrow \phi \ell^+ \ell^-$ process. I will also describe the important procedures needed to calibrate the simulated samples used in this thesis. Chapter 5 features the first analysis I made significant contributions to. The goal is to study the relative frequency of the $B_s^0 \rightarrow \phi e^+ e^-$ and $B_s^0 \rightarrow \phi \mu^+ \mu^-$ processes, to test an important accidental symmetry of the Standard Model, lepton flavour universality. This symmetry states that the behaviour of electrons and muons in the Standard Model should be identical, apart from their difference in mass. In Chapter 6 I will present my main analysis project. Following from the experience gained from the first analysis, the aim is to perform the first ever *angular analysis* of the $B_s^0 \rightarrow \phi e^+ e^-$ mode, in an effort to further test whether the Standard Model treats electrons and muons differently in any way.

Finally, I had the privilege of contributing to something other than physics within the LHCb collaboration: I was part of Early Career, Gender and Diversity office, a fundamental body of LHCb which aims at monitoring and improving the well-being of all LHCb members. I took the role of Early Career representative, one of the four major figures in the office, from 2023 to 2025. In Chapter 7 I will give a brief overview of all of the activities I was a part of during my term.

A final note to the reader: you might notice that in place of the classic tradition of preceding a chapter with famous (or not so famous) quotes, I opted for a *multi-media* experience. I decided to assign to each chapter an album that I either have a strong connection to, or that I simply greatly appreciate. Some of the albums I have selected have been part of my “personal soundtrack” while working on the contents of this thesis, others I have been enjoying since before I even started my PhD. I hope the reader can appreciate them and, hopefully, discover new great pieces of music whilst reading about my work.

Chapter 2

Theoretical background

The author's recommendation

Confessions on a Dance Floor (non-stop mix)

by **Madonna**

2005, Warner Bros.

With this chapter I introduce the scientific theory which describes the physics behind the measurements presented in this thesis. Our knowledge of the fundamental building blocks of matter has improved drastically over the last two centuries, starting from the discovery of the first subatomic particle, the electron, at the end of the 1800s. From there, a series of groundbreaking results throughout the first half of the last century, both theoretical and experimental, started to indicate the existence of a common underlying pattern. The pieces of this large puzzle were slowly brought together over decades by many researchers, until what we call the Standard Model of particle physics was born in the 1970s. Since its first conception, its predictions have been confirmed experimentally time and time again, to astounding levels of precision. Its crowning achievement, the discovery of the Higgs boson in 2012, brought with it a new challenge. The Standard Model is known to be incomplete, yet with only some current experimental deviations from its predictions, the picture of what might supercede it is still murky. In the chapter I will first cover the basis of the Standard Model, and will then focus on the heavy flavour sector. Some of the known shortcomings of the model, and current discrepancies in flavour physics, will also be presented. Finally, I will introduce a model-independent approach useful to put the SM predictions to the test, and in particular I will apply it to the case of the $B_s^0 \rightarrow \phi \ell^+ \ell^-$ decays studied in the later chapters. Natural units will be used in all equations in this chapter, $\hbar = c = 1$, and Einstein's summation convention is also used where appropriate.

2.1 The Standard Model of particle physics

The Standard Model of particle physics, SM for short, is built within a theoretical framework known as quantum field theory (QFT) [1, 2]. QFT is a synthesis of quantum mechanics and

special relativity, applied, rather than on systems of particles, on systems of fields. Under this framework, particles are interpreted as excited states of an underlying quantum field. Following from classical field theory, the fields obey the Euler-Lagrange equations of motion, which for a field $\psi(x)$, where x is a four vector, associated to a Lagrangian density $\mathcal{L}(\psi, \partial_\mu \psi)$, from here on referred to simply as the Lagrangian, take the form

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi)} \right) - \frac{\partial \mathcal{L}}{\partial \psi} = 0, \quad (2.1)$$

where ∂_μ is the four-gradient operator. For the equations of motion to respect special relativity, one must impose that their form is invariant under Lorentz transformations. Looking at Eq. (2.1), this means that the Lagrangian $\mathcal{L}$ must be a Lorentz scalar.

The field $\psi(x)$ can be defined by how it changes under a Lorentz transformation. For example, a field is said to be scalar if its numerical value at any point in spacetime is invariant under a change of reference frame $\psi(x) \rightarrow \psi'(x') = \psi(x)$. Similarly, a vector field V^μ is defined as transforming like a Lorentz four-vector $V^\mu(x) \rightarrow V'^\mu(x') = \Lambda^\mu{}_\nu V^\nu(x)$, where $\Lambda^\mu{}_\nu$ is the Lorentz transformation matrix, and so on.

A fundamental property of quantum fields, which is closely tied to how they behave under Lorentz transformations, is spin. Spin is an intrinsic characteristic of a field, like the mass, and it can take either integer or half-integer values. Particles with integer spin are called bosons, while particles with half-integer spin are called fermions. As will be explained in the following, the SM contains both elementary fermions of spin-1/2, which are the fields that make up matter, and bosons, which represent the forces between fermions. A revolutionary result from Dirac in 1928 [3], which is at the basis of the SM, is the derivation of the equation of motion for free fermion fields of spin-1/2. In its modern form, this can be derived from the Lagrangian,

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi, \quad (2.2)$$

where ψ is a four-component field, of mass m , called a Dirac spinor, and γ^μ are the 4×4 Dirac matrices. The Dirac spinor can also be expressed in terms of two two-dimensional Weyl spinors, which are distinguished by a property known as chirality. For spin-1/2 particles, chirality comes in two states, left- and right-handedness. Thus the Dirac spinor can be decomposed as

$$\begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}. \quad (2.3)$$

The distinction of left- and right-handed particle states is a key feature of the SM, as will be shown later on. Chirality is related to another variable which will be useful later in the chapter, the helicity h . Helicity is defined as the projection of the spin of a particle along its momentum vector. A positive helicity is known as right-handed, while a negative helicity is known as left-handed. Unlike chirality, helicity is not a Lorentz invariant for massive

particles, meaning one can always boost to a reference frame where the sign of the helicity flips. For massless particles, helicity and chirality coincide.

2.1.1 From symmetries to interactions

Symmetries, such as Lorentz invariance, are an extremely powerful tool in theoretical physics. Noether's theorem [4] states that for every symmetry of the Lagrangian, there exists an associated conserved quantity. For example, to impose that the Lagrangian is invariant under space translations, results in a field theory which conserves linear momentum. Each component of the momentum that is conserved, along x, y, z , corresponds to each of the directions that one can translate the physical system into without changing its description.

The SM falls under a special type of field theory, known as gauge theory. In a gauge theory, the Lagrangian is invariant under local phase transformations, which can be expressed as

$$\psi \rightarrow e^{ig\theta(x)} \cdot \psi, \quad (2.4)$$

where $\theta(x)$ is a locally varying parameter, and g is known as the coupling constant. The exact form of the transformation and numbers of degrees of freedom of $\theta(x)$ are defined by the symmetry imposed on the Lagrangian. Taking the case of a single degree of freedom, where the exponential is simply a complex number, it follows that any term in the Lagrangian containing $\partial_\mu \psi$, will break gauge invariance. This can be solved by “promoting” the four-gradient to a covariant derivative operator,

$$\partial_\mu \mapsto D_\mu \equiv \partial_\mu - igA_\mu, \quad (2.5)$$

where a new field A_μ , called gauge field, is introduced. Thanks to the introduction of the covariant derivative, $D_\mu \psi$ transforms like ψ , and it is possible to write gauge-invariant terms containing the derivative of the field. The new gauge field is defined by its behaviour under the local gauge transformation $A_\mu \rightarrow A_\mu + \partial_\mu \theta$. In general, for each degree of freedom in the local transformation, there is an associated gauge field. A more detailed description of the group theory aspects of this description can be found in Ref. [2].

The gauge fields arise as direct manifestations of the local symmetries of the physical system. If the fields are quantised, then each gauge field corresponds to a particle, known as a gauge boson. In the field theory, the bosons act as the force carriers. Indeed, replacing the derivative with the covariant derivative for a single gauge field in the Dirac Lagrangian one obtains

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi + gA_\mu \gamma^\mu \bar{\psi}\psi, \quad (2.6)$$

where the product of ψ and A_μ is an interaction term between the two fields. A more complete gauge-invariant Lagrangian¹ describing a Dirac field also includes a kinetic term

¹The most complete Lagrangian would also include a term of the form $\theta \epsilon^{\alpha\beta\mu\nu} F_{\alpha\beta}^a F_{\mu\nu}^a$, where θ is a dimensionless constant and $\epsilon^{\alpha\beta\mu\nu}$ is the Levi-Civita symbol. Current experimental evidence constrains such this additional term to be very small, *i.e.* $\theta \sim 0$. As this goes beyond the scope of this thesis, it will not be

for the each associated gauge field A_μ^a (where the index a indicates all degrees of freedom of the imposed symmetry),

$$\mathcal{L}_{\text{gauge}} = \mathcal{L}_{\text{Dirac}} - \frac{1}{4} F_{\mu\nu}^a F^{a\mu\nu}, \quad (2.7)$$

where $F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - igf^{abc}A_\mu^b A_\nu^c$ is the field strength tensor for the field A_μ^a . The f^{abc} are parameters known as structure constants, and are unique to each symmetry that is imposed on the system.

The SM describes three fundamental interactions, which arise from local gauge symmetry under three groups, $SU(3) \times SU(2)_L \times U(1)$. The special unitary group $SU(n)$ includes all $n \times n$ unitary matrices with unit determinant; it has $n^2 - 1$ degrees of freedom. The group $U(1)$ contains all complex numbers that fall on the unit circle. Summing all degrees of freedom, from the three gauge symmetries of the SM arise $8 + 3 + 1 = 12$ gauge fields.

The strong nuclear force, $SU(3)$

The symmetry group $SU(3)$ describes the strong nuclear force, which is the fundamental interaction responsible for binding nucleons together in atomic nuclei. With eight degrees of freedom, it generates eight unique gauge fields. The particles corresponding to the excited states of these fields are called gluons. For a fermion to interact with the gluon fields it must carry a charge, similarly to how electric charge defines a particle's ability to feel the electromagnetic force. There are two sets of three possible charges, one for particles and one for antiparticles. In analogy with colour theory, these are called colour charges, and the two sets are named after the primary colours: red, blue, and green; cyan (anti-red), yellow (anti-blue), and magenta (anti-green). The composition of the three colours, or the three anti-colours, results in a net colour charge of zero. Similarly this happens when combining a colour with its corresponding anti-colour. This analogy gives also the modern name of the QFT that describes the $SU(3)$ sector of the SM, quantum chromodynamics (QCD).

QCD exhibits two interesting properties. Firstly, the gluons themselves carry a colour charge (more specifically, they carry a colour charge and an anticolour charge), meaning that they can interact between themselves. This is different than for the more familiar photon which, having no electrical charge, does not self-interact. Secondly, it is not possible to observe single colour states in QCD. This is due to what is known as colour confinement: fermions and bosons carrying colour charge can only be found as composite structures, held together by gluons, which have no colour charge in total. Colour confinement is a result of the dependence of the QCD coupling constant on the energy scale of a process: at low energies it becomes increasingly large, strongly binding coloured particles together. In this regime QCD can not be reliably predicted using perturbative methods, as other algorithms are needed to model of the dynamics within a bound structure. Such dynamics will be of interest later in the chapter, when discussing QCD form factors and their effect on the theoretical predictions of decays of bound particles.

included here.

The electroweak force, $SU(2)_L \times U(1)$

The combination of the $SU(2)_L \times U(1)$ symmetries gives rise to four gauge bosons and charges, which describes the unified model of the electromagnetic and weak nuclear interactions. From the $U(1)$ group there is a single associated charge, called the weak hypercharge. The gauge field associated to this symmetry does not carry itself a hypercharge, and hence no self-interaction terms are present.

The charges from the $SU(2)_L$ symmetry are components of three-dimensional objects called weak isospin, and each of the gauge fields couples to its specific component. However, the isospin operators, similarly to the angular momentum operators in quantum mechanics, do not commute. This results in only two independently observable quantities, which can be expressed as the square magnitude of the isospin vector, and its third component. The three gauge fields of the symmetry carry non-zero weak isospin, so they can also self-interact. The subscript L indicates that the gauge bosons from the $SU(2)$ group only couple to particles with left-handed chirality (and antiparticles with right-handed chirality). This is a well-known property of the weak nuclear force, and it has been extensively verified experimentally [5,6]. The consequences of the handedness of weak interactions are discussed below.

Discrete symmetries

While continuous symmetries are at the basis of the SM, discrete symmetries also have an important role. Three symmetries are of particular interest in QFT, namely parity, charge conjugation, and time reversal. Parity is the transformation P that maps the four-position $x^\mu = (x^0, \vec{x})$ to $x_\mu = (x^0, -\vec{x})$, *i.e.* it flips the directions of the space coordinates. Under parity, chirality changes as well $L \leftrightarrow R$. Charge conjugation is the operation C which swaps all particles with the corresponding antiparticles, but preserves the same momentum direction and chirality. Time reversal, the operation T changing the sign of the time coordinate, flips the momentum direction but preserves chirality.

Out of the interactions described from gauge theories presented above, only the $SU(2)_L$ group violates these discrete symmetries. In particular, both parity and charge conjugation map a field which couples to the $SU(2)_L$ gauge fields to one that does not. However, the combination of C and P maps a left-handed particle to a right-handed antiparticle, and is thus a symmetry of the theory.

These three discrete symmetries have been shown to be inextricably tied to Lorentz invariance: in fact, it is not possible to write down a QFT that is Lorentz invariant without having it also be invariant under the combined action of CPT .

2.1.2 A mass(ive) issue

The description given so far can be summarised by a total Lagrangian, which includes both fermion fields for matter, and the $SU(3) \times SU(2)_L \times U(1)$ gauge fields for the interactions between them. Consider an example case, the Lagrangian for a single fermion field, after

imposing the local gauge symmetry,

$$\mathcal{L} = i\bar{\psi}_L\gamma^\mu D_{L,\mu}\psi_L + i\bar{\psi}_R\gamma^\mu D_{R,\mu}\psi_R - \frac{1}{4}F_{\mu\nu}^a F^{a\mu\nu}, \quad (2.8)$$

where there are two distinct covariant derivatives $D_{L,\mu}, D_{R,\mu}$, to reflect that the gauge fields of the $SU(2)_L$ symmetry only couple with the left-handed component of the fermion field, and the $F_{\mu\nu}^a$ are the field strength tensors for the twelve gauge fields, $a = 0, \dots, 11$.

A fundamental shortcoming arises here, because both the fermions and gauge bosons described above are massless, which is in clear contrast with experimental evidence. Indeed, the fermion mass term, written in terms of the chiral fields,

$$\mathcal{L}_{\text{mass, Dirac}} = im_\psi(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R), \quad (2.9)$$

is not symmetric under a $SU(2)$ transformation which only couples to left-handed chirality. Similarly, a gauge field mass term,

$$\mathcal{L}_{\text{mass, gauge}} = \frac{1}{2}m_A A_\mu A^\mu, \quad (2.10)$$

breaks any local gauge invariance. The electromagnetic and strong force carriers, the photon and gluons, are massless, so the corresponding terms in the Lagrangian vanish and do not break gauge symmetry. However, an important property of the weak nuclear force is that it acts only on short-range distances, smaller than atomic radii. In QFT terms, this implies that the gauge fields that describe the interaction are massive. It is then fundamental to determine the mechanism which can introduce the mass terms into the theory while still respecting gauge invariance.

The solution came from a breakthrough in 1964 from three independent groups [7–9], demonstrating how spontaneous symmetry breaking could solve these issues. The idea of spontaneous symmetry breaking was originally developed within the theory of superconductivity [10, 11], and later applied to QFTs [12] for global symmetries. Spontaneous symmetry breaking is achieved when the ground state of a physical system breaks the symmetry which was originally imposed on the Lagrangian. This process happens when the QFT possesses, instead of a unique ground state which respects the symmetry of the Lagrangian, a family of degenerate ground states, which transform into each other under the symmetry transformation. Since there can only be one true ground state, by imposing the choice of a unique vacuum on the system, the symmetry is said to be spontaneously broken.

The key insight was to apply previous results on spontaneous breaking of global symmetries to local gauge theories. In the SM, this is colloquially known as the Higgs mechanism, and it is responsible for spontaneously breaking the $SU(2)_L \times U(1)$ gauge symmetry to generate the masses of the bosons of the weak interaction. The mechanism is driven by the introduction of a new scalar field φ , with a Lagrangian

$$\mathcal{L}_{\text{Higgs}} = (D_\mu\varphi)^\dagger(D^\mu\varphi) - V(\varphi), \quad (2.11)$$

where φ couples to the $SU(2)_L$ gauge fields $A_\mu^1, A_\mu^2, A_\mu^3$, and to the $U(1)$ gauge field B_μ , with coupling constants g and g' , respectively. The field φ takes the form of a complex-valued weak isospin doublet. The potential $V(\varphi)$ has the form,

$$V(\varphi) = \frac{1}{2}\lambda^2(|\varphi|^2 - v^2)^2, \quad (2.12)$$

where v is a constant known as the vacuum expectation value, and λ is a dimensionless constant. The potential is invariant under gauge transformations, and, because it only depends on the magnitude of the scalar field doublet, it has a continuous set of minima, which are degenerate up to a transformation under $SU(2)_L$. This freedom allows an arbitrary choice of the ground state, thus achieving spontaneous symmetry breaking. One can then expand the Lagrangian around the ground state of φ ; defining the physical fields corresponding to the $W^\pm$ and Z bosons, as linear combination of the gauge fields,

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}}(A_\mu^1 \mp iA_\mu^2) \\ Z_\mu &= \frac{1}{\sqrt{g^2 + g'^2}}(gA_\mu^3 - g'B_\mu), \end{aligned} \quad (2.13)$$

these gain mass terms with $m_W = gv/2$ and $m_Z = \sqrt{g^2 + g'^2}v/2$. The photon, γ , is also defined as a linear combination of the original gauge fields,

$$A_\mu = \frac{1}{\sqrt{g^2 + g'^2}}(g'A_\mu^3 + gB_\mu), \quad (2.14)$$

but it remains massless. With the Higgs mechanism, the SM then exists within two regimes: above the electroweak scale, $\mathcal{O}(200 \text{ GeV})$, the $SU(2)_L \times U(1)$ symmetry holds. Below this energy scale, the field φ translates into a new ground state, spontaneously breaking $SU(2)_L \times U(1)$ into a $U(1)_Q$ symmetry.

The surviving symmetry represents the familiar gauge symmetry of the electromagnetic interaction, which results in the conservation of the electric charge Q . The electric charge is related to the third component of the weak isospin I_3 and the weak hypercharge Y by the relation $Q = I_3 + Y/2$. Following this, it is then possible to define the correct quantum numbers for all fermion and boson fields to achieve the correct observed charges. For example, the Higgs field, h , has no electric charge, and following the description above, its ground state possesses an isospin of $1/2$ and a third component of $-1/2$. Thus, its hypercharge must be $Y = 1$. Tables 2.1 and 2.2 report the quantum numbers of the gauge fields before and after symmetry breaking.

2.1.3 The matter content

All the ingredients are now there to fully describe the matter fields in the SM. Two distinct families of elementary fermions exist: quarks and leptons. Each family exists across three generations of particles, all of which share the same quantum numbers, but have increasing

Table 2.1: Quantum numbers of colour c , weak hypercharge Y , weak isospin I and its third component I_3 , of the gauge fields in the SM, before spontaneous symmetry breaking.

Symmetry	Gauge field	c	Y	I	I_3
$SU(3)$	$g^a, a = 0...7$	8 combinations of r, g, b and c, m, y	0	0	0
$SU(2)_L$	W^1	0	0	1	1
	W^2	0	0	1	-1
	W^3	0	0	1	0
$U(1)$	B	0	0	0	0

Table 2.2: Quantum numbers of colour c , weak hypercharge Y , third component of the weak isospin I_3 , and electric charge Q of the boson fields and in the SM, after spontaneous symmetry breaking.

Symmetry	Field	c	Y	I_3	Q
$SU(3)$	$g^a, a = 0...7$	8 combinations of r, g, b and c, m, y	0	0	0
$U(1)$	γ	0	0	0	0
	W^+	0	0	1	1
	W^-	0	0	-1	-1
	Z	0	0	0	0
	h	0	1	-1/2	0

masses (the third generations containing the heaviest). Six types, or flavours, of quarks exist, paired up in the three generations: up and down, u, d , charm and strange, c, s , top and beauty, t, b . The up, charm, and top are also known as up-type quarks, whilst the other three are known as down-type quarks. Similarly, there exist six leptons, three of which have an electric charge and are massive, the electron, muon, and tau, e, μ, τ , and three of which are neutral and are massless,² the electron, muon, and tau neutrinos, ν_e, ν_μ, ν_τ .

Table 2.3 reports the quantum numbers of all known elementary fermions. Leptons interact only through the electromagnetic and weak nuclear forces. The quarks interact with all of the three fundamental forces, and in particular they carry colour charge, which makes them form bound states of two or more quarks called hadrons. Among hadrons, one can distinguish mesons, which are quark-antiquark states, and baryons, which contain three quarks. Candidates of tetraquark and pentaquark states, containing four and five quarks each, have been observed but the exact nature of many of these states is still uncertain. For an overview of the current known exotic states, see Ref. [13].

The masses of quarks and charged leptons are generated by adding the following gauge-invariant terms to the Lagrangian, describing the so-called Yukawa coupling between the

²While the non-zero masses of neutrinos have been established experimentally, this is not relevant for this thesis; neutrinos are thus treated as massless in the description of the SM presented here.

Table 2.3: Quantum numbers of colour c , weak hypercharge Y , weak isospin I and its third component I_3 , and electric charge Q of the fermion fields, distinguished between left-handed and right-handed chirality states. The quantum numbers are the same across the three fermion generations.

Field	Chirality	c	Y	I	I_3	Q
Up-type quarks	L	r, g, b	$1/3$	$1/2$	$1/2$	$2/3$
	R		$4/3$	0	0	
Down-type quarks	L	r, g, b	$1/3$	$1/2$	$-1/2$	$-1/3$
	R		$-2/3$	0	0	
Neutrinos	L	0	-1	$1/2$	$1/2$	0
	R	Does not exist in the SM				
Charged leptons	L	0	-1	$1/2$	$-1/2$	-1
	R		-2	0	0	

fermions ψ and the Higgs field,

$$\mathcal{L}_{\text{Yukawa,lepton}} = -g_\ell \left((\bar{\Psi}_L \cdot \varphi) \psi_R - \bar{\psi}_R (\varphi^\dagger \cdot \Psi_L) \right) \quad (2.15)$$

$$\mathcal{L}_{\text{Yukawa,quark}} = -g_u \left((\bar{\Psi}_L \cdot \varphi) \psi_R^u - \bar{\psi}_R^u (\varphi^\dagger \cdot \Psi_L) \right) - g_d \left((\bar{\Psi}_L \cdot \varphi) \psi_R^d - \bar{\psi}_R^d (\varphi^\dagger \cdot \Psi_L) \right). \quad (2.16)$$

Here, g_ℓ , g_u , and g_d are the coupling of the lepton, up-type quark and down-type quark fields to the Higgs field, respectively. They are unique to each fermion and are directly proportional to their mass. Because the Higgs field φ is a weak isospin doublet, then for the Yukawa term to be invariant the left-handed fermions must also be paired in doublets Ψ_L , so that the dot-product $(\bar{\Psi}_L \cdot \varphi)$ is well defined.

The pairing is done for each generation of fermions: up-quarks and down-quarks form three doublets, while neutrinos and their corresponding charged leptons form another three doublets. After spontaneous symmetry breaking, the Higgs field is in a non-zero ground state, which when substituted in Eq. (2.16) gives rise to the fermion masses $m_f = g_f v/2$. Note that since in the SM there are no right-handed neutrinos, only Yukawa terms for the charged leptons are possible.

2.1.4 Intermezzo: Feynman diagrams and amplitudes

With a complete description of the matter and gauge fields, it is possible to compute probabilities for physical processes. Consider an initial m -particle state $|i\rangle$ evolving into a final n -particle state $|f\rangle$. The differential rate of this process can be expressed as

$$d\Gamma_{i \rightarrow f} \sim |\mathcal{M}_{i \rightarrow f}|^2 (2\pi)^4 \delta^4(P_i^\mu - P_f^\mu) \underbrace{\prod_{f=1}^n \frac{d^3 p_f}{(2\pi)^3 2E_f}}_{\text{phase space factor } d\Phi^n}, \quad (2.17)$$

where $\mathcal{M}_{i \rightarrow f}$ is known as the amplitude; the phase-space factor $d\Phi^n$ contains a Dirac-delta function which imposes energy and momentum conservation between the initial and final state, and a product of normalisation factors for each final state particle.

The normalisation factor missing in the expression for $d\Gamma_{i \rightarrow f}$ depends on the initial state. In the case of a single particle of mass M decaying into n particles, such as the processes that are covered in this thesis, the previous expression takes the exact form,

$$d\Gamma_{1 \rightarrow n} = \frac{|\mathcal{M}_{1 \rightarrow n}|^2}{2M} d\Phi^n. \quad (2.18)$$

In QFT the standard way to compute $\mathcal{M}_{1 \rightarrow n}$ is to consider the interaction that mediates the process as a perturbation of the free field Lagrangian, $\mathcal{L} = \mathcal{L}_{\text{free}} + \lambda \mathcal{L}_{\text{int}}$, where $\lambda \ll 1$ is a dimensionless coupling constant. These perturbative methods are well suited to the environment of high energy particle collisions. The decay amplitude is expanded as an infinite sum of terms containing increasing powers of λ . Since $\lambda \ll 1$, the higher power terms become increasingly sub-leading, thus one can decide, for practical purposes, to truncate the sum at a finite order. The first order that has a non zero amplitude is known as leading-order (LO), the following is known as next-to-leading order (NLO), and so on.

Each term contributing to the infinite sum can be represented pictorially in what is known as a Feynman diagram. A diagram is formed by a series of segments, representing particles, and vertices, representing interaction points. From a given Lagrangian, one can use a prescription to define “Feynman rules”, which dictate how each particle is drawn and what each part of the diagram corresponds to in the infinite sum. There can be multiple diagrams per order considered. Figure 2.1 shows two of the LO diagrams for $e^+e^- \rightarrow e^+e^-$ scattering, as well as one of the NLO diagrams.

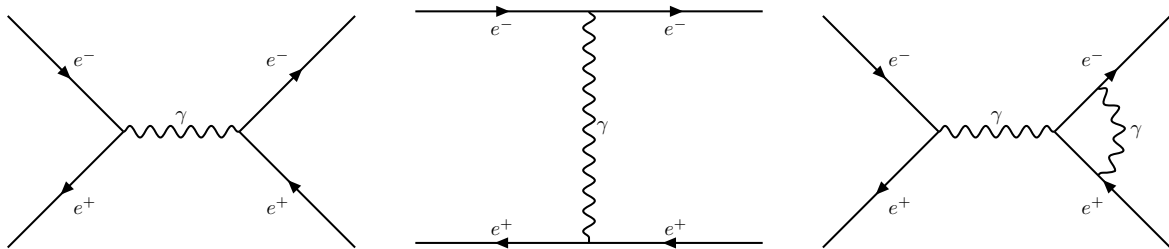


Figure 2.1: Feynman diagrams for the scattering process $e^+e^- \rightarrow e^+e^-$, mediated with the exchange of a photon, at leading order (left, centre) and next-to-leading order (right).

Any intermediate particles appearing in the diagram, *i.e.* ones which propagate between two vertices, are known as “virtual”. In general, higher order diagrams will contain loops of virtual particles; these correspond mathematically to integrals over all possible exchanges of momentum. These integrals are typically divergent, and need to be treated with care. The effect of these divergent integrals can be systematically accounted for, in what is nowadays a common practice in QFT known as renormalisation. The SM has been proven to be a renormalisable theory [14], meaning that all diverging integrals can be accounted for with a finite number of counterterms, for all orders of perturbation theory.

2.2 Flavour physics phenomenology

In the description of the quark masses, there are no symmetries preventing mixing between the three generations. As such, one should write the Yukawa term in a more general form, where the couplings g_f are substituted by 3×3 matrices $\lambda_{u,d}^{ij}$, with the i, j indices corresponding to the three generations of up-type and down-type quarks. In this more general description, it is possible to define two bases with which to describe the quark interactions: the “flavour” basis, and the “mass” basis. The flavour basis corresponds to the original description of the SM Lagrangian, where the interaction terms of up-type and down-type quarks with the gauge fields are diagonal. Take as an example the coupling term between the quarks and the W^+ field,

$$\mathcal{L}_{q\bar{q}W^+} = \frac{g}{\sqrt{2}} W_\mu^+ \bar{u}_L^i \gamma^\mu d_L^i, \quad (2.19)$$

where u_L^i and d_L^i are the left-handed up-type and down-type quark fields, with $i = 1, 2, 3$ indicating the three generations.

The mass basis corresponds instead to the basis in which the matrices $\lambda_{u,d}^{ij}$ are diagonal, which implies that the mass matrices $M_{u,d}^{ij} = v/2 \cdot \lambda_{u,d}^{ij}$ are diagonal. By defining the unitary matrices $U_{u,d}^{ij}$, which rotate the flavour basis into the mass basis,

$$u_L^i \rightarrow U_u^{ij} u_L'^j, \quad d_L^i \rightarrow U_d^{ij} d_L'^j, \quad (2.20)$$

then Eq. (2.19) can be rewritten in the mass basis, where a mixing matrix appears from the combination of the two transformations $V^{ij} = (U_d^\dagger U_u)^{ij}$. This matrix is known as the Cabibbo-Kobayashi-Maskawa matrix [15, 16], or CKM matrix for short. The magnitudes of the CKM matrix entries have been measured to be [17]

$$\begin{aligned} |V^{ij}| &= \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} \\ &= \begin{pmatrix} 0.97367 \pm 0.00032 & 0.22431 \pm 0.00085 & 0.00382 \pm 0.00020 \\ 0.221 \pm 0.004 & 0.975 \pm 0.006 & 0.0411 \pm 0.0012 \\ 0.0086 \pm 0.0002 & 0.0415 \pm 0.0009 & 1.010 \pm 0.027 \end{pmatrix}. \end{aligned} \quad (2.21)$$

The matrix is almost symmetric and diagonal. This implies that the weak interaction preferably couples quarks within the same generation, but the non-zero diagonal elements allow for transitions between up- and down-type quarks of any generation.

There are a series of important features of the flavour structure of the SM which are important to the work presented in this thesis.

2.2.1 CP violation

The 18 parameters of the CKM matrix, 9 magnitudes and 9 complex phases, are not all independent. Because it is defined as the product of two unitary matrices, the matrix is itself unitary. This imposes 9 relations between the parameters, halving the number of free parameters,

$$V^{ik}V^{*jk} = \delta^{ij}, \quad (2.22)$$

where δ^{ij} is the Kronecker delta. Furthermore, it is possible to absorb five complex phases into the definition of the quark fields, bringing the number of free parameters to four. These can be expressed as three Euler angles and one residual complex phase δ ,

$$V^{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_{23} & \sin \theta_{23} \\ 0 & -\sin \theta_{23} & \cos \theta_{23} \end{pmatrix} \begin{pmatrix} \cos \theta_{13} & 0 & \sin \theta_{13} e^{-i\delta} \\ 0 & 0 & 0 \\ -\sin \theta_{13} e^{i\delta} & 0 & \cos \theta_{13} \end{pmatrix} \begin{pmatrix} \cos \theta_{12} & \sin \theta_{12} & 0 \\ -\sin \theta_{12} & \cos \theta_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.23)$$

The presence of a complex phase in the CKM matrix constitutes the only source of CP violation in the SM. This is manifest, for example, in particle decays where multiple amplitudes are summed coherently. To see the effect it is helpful to write the amplitudes separating the phases that do not change sign under CP , the strong phases, δ_i , and the ones that do, the weak phases, ϕ_i . The weak phases in the SM correspond to the phases from the CKM matrix. The strong phases appear from any possible intermediate scattering states in the decay amplitude. For example, in the $B^\pm \rightarrow \pi^\pm K^+ K^-$ decay, there is a significant contribution from intermediate rescattering of the dikaon pair $K^+ K^- \leftrightarrow \pi^+ \pi^-$ [18]. This rescattering causes an overall shift in phase in the decay amplitude, but because it originates from a CP -conserving interaction, it does not change sign under CP .

To illustrate how nonzero weak phases can lead to observable CP -violating effects, consider $\mathcal{M}_f$ and $\overline{\mathcal{M}}_{\bar{f}}$ as the total amplitudes of a particle decay to a final state f and its CP conjugate, respectively, composed of two interfering amplitudes $A_{1,2}$,

$$\begin{aligned} \mathcal{M}_f &= |A_1| e^{i(\delta_1 + \phi_1)} + |A_2| e^{i(\delta_2 + \phi_2)} \\ \overline{\mathcal{M}}_{\bar{f}} &= |A_1| e^{i(\delta_1 - \phi_1)} + |A_2| e^{i(\delta_2 - \phi_2)}. \end{aligned} \quad (2.24)$$

Note that the phases depend on the convention used, however any phase difference, like $\phi_2 - \phi_1$, is physical and independent of convention. It's easy to see that the difference in the squared amplitudes, and thus in the decay rates, is

$$|\mathcal{M}_f|^2 - |\overline{\mathcal{M}}_{\bar{f}}|^2 \propto |A_1| \cdot |A_2| \sin(\delta_2 - \delta_1) \sin(\phi_2 - \phi_1), \quad (2.25)$$

implying that the complex weak phases generate CP violating processes.³

³It might seem as though CP violation is only possible when both strong and weak phases are present. That is not entirely the case, and CP violation can still be present regardless of a difference in strong phases,

2.2.2 Flavour changing processes

A significant feature of the SM is that the flavour-changing interactions are only mediated by the charged $W^\pm$ bosons. Indeed, in the Lagrangian the interactions with the photon and Z boson do not mix up- and down-type quarks; hence the rotation from the flavour-basis to the mass basis cancels out in those terms as $(U_{u,d}^\dagger U_{u,d})^{ij} = \delta^{ij}$.

The decay process studied in this thesis, the $b \rightarrow s\ell^+\ell^-$ quark transition, falls into a category known as flavour changing neutral currents (FCNCs). These involve transitions where there is no charge difference between the initial and final state quark. Because of the flavour-diagonal structure of the SM Lagrangian, these can only happen in diagrams with loops, which makes their rates naturally suppressed. Figure 2.2 shows two Feynman diagrams at LO which describe the $b \rightarrow s\ell^+\ell^-$ transition.

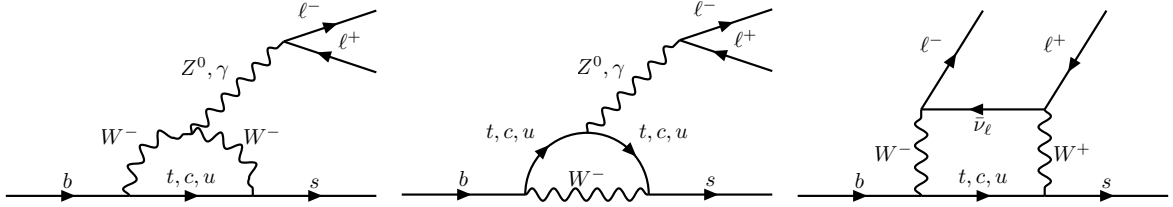


Figure 2.2: The leading order Feynman diagrams for the $b \rightarrow s\ell^+\ell^-$ quark transition. Left, centre: loop “penguin” diagrams. Right: a “box” diagram, mediated via the emission of two virtual W bosons and the exchange of a virtual anti-neutrino $\bar{\nu}_\ell$.

What is obvious in the diagrams is that to achieve a $b \rightarrow s$ transition in the SM one must make two intermediate transitions, going through a virtual up-type quark. The transition is however not mediated with equal contributions for the three up-type flavours, u, c, t . The resulting amplitude contains a term $m_i^2/m_W^2 V_{ib} V_{is}^*$, with $i = u, c, t$, which results in a natural suppression of the lighter quarks because of their mass difference with the $W^\pm$ bosons. The u quark diagrams are further suppressed by the product of two off-diagonal terms of the CKM matrix, compared to only one off-diagonal term in the case of the c and t diagrams.

2.2.3 Neutral meson mixing

The mixing of neutral mesons $M^0 \leftrightarrow \bar{M}^0$ has been a well established process since the 1950s, and was first observed for the neutral kaon system [19, 20]. Nowadays it is understood as a direct consequence of the flavour mixing induced by the weak interaction, and the CKM matrix. Figure 2.3 shows the diagrams at LO for the mixing of $B_q^0 - \bar{B}_q^0$ mesons. The process only happens through diagrams with loops, and is another example of FCNCs.

The mixing and its time dependence are described with an effective 2×2 Hamiltonian for example in the mixing of neutral mesons.

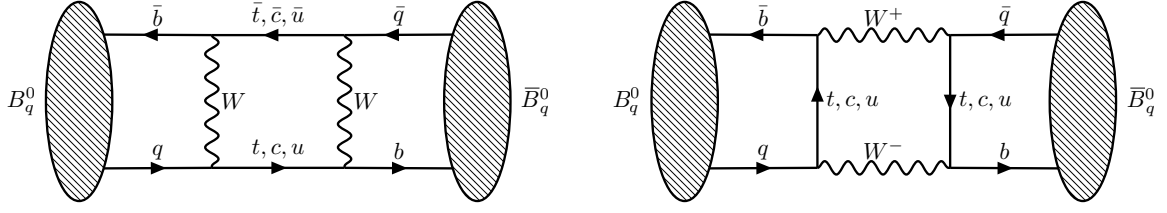


Figure 2.3: The leading order box diagrams for the mixing of neutral $B_q^0 - \bar{B}_q^0$ mesons. Each diagram represents three possible diagrams, mediated through three up-type quarks, u, c, t .

expressed in the form [17],

$$\mathcal{H} = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} = \begin{pmatrix} M_{00} - \frac{i}{2}\Gamma_{00} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & M_{00} - \frac{i}{2}\Gamma_{00} \end{pmatrix}, \quad (2.26)$$

where $\mathbf{M}$ and $\mathbf{\Gamma}$ are unitary matrices, representing the contributions of short-distance effects (like the diagrams in Fig. 2.3), and long-distance rescattering processes, respectively. Short- and long-distance processes differ in that in the former, one can apply perturbation theory to compute their contribution to the total amplitude, while in the latter effects from QCD require the use non-perturbative methods. Invariance under CPT imposes the diagonal terms to be equal, $\mathcal{H}_{00} = \mathcal{H}_{11}$.

If the off-diagonal terms are non-vanishing, the flavour states $|M^0\rangle, |\bar{M}^0\rangle$ are no longer eigenstates of the Hamiltonian, and are able to mix into one another. The eigenstates of the Hamiltonian have well-defined masses and decay widths, and are named in terms of their difference in mass, as light, $|M_L\rangle$, and heavy, $|M_H\rangle$. One can write the mass eigenstates in terms of the flavour-specific states,

$$\begin{aligned} |M_L\rangle &= p|M^0\rangle + q|\bar{M}^0\rangle \\ |M_H\rangle &= p|M^0\rangle - q|\bar{M}^0\rangle, \end{aligned} \quad (2.27)$$

where p, q are complex parameters, related by a normalisation condition $|p|^2 + |q|^2 = 1$. Another two important quantities are the differences in mass and decay width between the light and heavy states, Δm and $\Delta\Gamma$. Denoting their eigenvalues by $\lambda_{L,H}$, then the differences correspond to,

$$\begin{aligned} \Delta m &\equiv \Re(\lambda_H - \lambda_L) = m_H - m_L \\ \Delta\Gamma &\equiv 2\Im(\lambda_L - \lambda_H) = \Gamma_L - \Gamma_H, \end{aligned} \quad (2.28)$$

where the $\Delta\Gamma$ is a signed quantity, defined in the convention typically used for B_q^0 mesons [17], which follows from the SM prediction that $\Gamma_L > \Gamma_H$. The quantity Δm is positive by definition. One can solve the eigenvalue problem $(\mathcal{H} - \lambda_{L,H}\mathbb{1})|M_{L,H}\rangle = 0$ to obtain expressions for p, q and the eigenvalues $\lambda_{L,H}$. They are not relevant for the purposes of this thesis, and

can be found in Ref. [17].

Finally, one can write the time evolution of the flavour states, starting from initially pure states of either $|M^0\rangle$ or $|\bar{M}^0\rangle$, to obtain

$$\begin{aligned} |M^0(t)\rangle &= g_+(t)|M^0\rangle + \frac{q}{p}g_-(t)|\bar{M}^0\rangle \\ |\bar{M}^0(t)\rangle &= g_+(t)|\bar{M}^0\rangle + \frac{p}{q}g_-(t)|M^0\rangle, \end{aligned} \quad (2.29)$$

where $g_{\pm}(t)$ are functions of the proper time, containing combinations of the eigenvalues of the light and heavy states,

$$g_{\pm}(t) = \frac{1}{2} \left(e^{(-im_H - \Gamma_H/2)t} \pm e^{(-im_L - \Gamma_L/2)t} \right). \quad (2.30)$$

The functions can be used to compute the probabilities of detecting one of the flavour states at a certain time t after production,

$$\begin{aligned} p(\bar{M}^0 \rightarrow \bar{M}^0)(t) &= |g_+(t)|^2 \\ p(M^0 \rightarrow \bar{M}^0)(t) &= |q/p \cdot g_-(t)|^2 \\ p(\bar{M}^0 \rightarrow M^0)(t) &= |p/q \cdot g_-(t)|^2. \end{aligned} \quad (2.31)$$

The $g_{\pm}(t)$ functions can be re-expressed in terms of the mass and width differences, additionally introducing $m = (m_L + m_H)/2$ and $\Gamma = (\Gamma_L + \Gamma_H)/2$,

$$g_+(t) = e^{-imt} e^{-\Gamma/2t} \left\{ \cosh \frac{\Delta\Gamma t}{4} \cos \frac{\Delta m t}{2} - i \sinh \frac{\Delta\Gamma t}{4} \sin \frac{\Delta m t}{2} \right\} \quad (2.32)$$

$$g_-(t) = e^{-imt} e^{-\Gamma/2t} \left\{ -\sinh \frac{\Delta\Gamma t}{4} \cos \frac{\Delta m t}{2} + i \cosh \frac{\Delta\Gamma t}{4} \sin \frac{\Delta m t}{2} \right\}. \quad (2.33)$$

These expressions will be important when writing down the time evolution of the decay amplitudes for the $B_s^0 \rightarrow \phi \ell^+ \ell^-$ process.

2.3 Limits of the Standard Model

The SM can describe a plethora of known processes in high energy physics, to astounding precision. Nonetheless there are experimental observations which remain unaddressed by the model. The existence of such limitations has prompted decades of theoretical and experimental research into what the current extension of the SM might be. Any extension falls under the umbrella term of New Physics (NP), also known as physics beyond the Standard Model (BSM).

The biggest conceptual limit of the SM is arguably its inability to describe the gravitational interaction. The mathematical language of QFT, built on the principles of special relativity, is incompatible with the widely accepted theory of gravity, general relativity. Attempts have been made to quantise Einstein's theory, but so far there is no consensus on what

the theory of “quantum gravity” should look like. Furthermore, cosmological observations, at scales where general relativity has proven to be successful time and time again, indicate that our Universe is not composed of merely ordinary matter. Instead it contains large quantities of dark matter and dark energy [21, 22], for which there are no feasible particle candidates within the SM. These observations have also proven a large scale matter-antimatter asymmetry in the Universe, which is only possible with CP violation [23]. However, the SM is not enough to explain such an asymmetry [24, 25].

There are a series of properties of the SM and its free parameters which are deemed “unnatural”. The model has a large number of parameters, amounting to 19. That includes 6 quark masses, 3 charged lepton masses, 4 CKM parameters, 3 gauge couplings, the Higgs field’s mass and vacuum expectation value, and the strong interaction’s CP violating parameter θ_{QCD} , which is vanishingly small in the SM. Their values raise some questions about their origin. For example, the masses of the elementary fermions are ordered hierarchically [17], increasing with each generation, and span many orders of magnitude, reaching the scale of the electroweak and Higgs bosons. Similarly, the SM provides no explanation for the almost-diagonal structure of the CKM matrix. Finally, it contains no mechanism which can explain why neutrino masses are so much smaller than all other elementary fermions.

There are additional problems of fine-tuning in the value of some of the SM parameters. The mass of the Higgs boson in the SM is subject to quantum corrections from loop diagrams, with the largest contributions coming from virtual $t\bar{t}$ pairs. Many of these corrections are much larger than the observed value of the Higgs mass, and thus require precise and almost complete cancellation when they are summed up. The value of the CP -violating parameter of the strong interaction, $\theta_{\text{QCD}} \simeq 0$, also poses a question of fine-tuning, as there is no higher symmetry imposing the $SU(3)$ sector of the SM to be CP -conserving.

All of these observations motivate the search for a more complete model, endowed with additional symmetries, capable of predicting (some of) the parameters’ values.

2.3.1 Flavour anomalies

The term flavour anomalies generally refers to a series of experimental observations, made in the b quark sector, which have shown tensions with the SM predictions. This section offers but a cherry-picked selection of some of the most relevant known anomalies, based on Refs. [26] and [27], which offer a more complete review of the theoretical and experimental status.

Decays of b -hadron decays have been the subject of both theoretical and experimental interest for many years. Rare decays mediated through FCNCs, like the $b \rightarrow s\ell^+\ell^-$ transition, are considered to be an ideal probe for BSM effects [28], as their natural loop suppression in the SM leaves plenty of room for new particles to contribute to their amplitudes. Radiative $b \rightarrow s\gamma$ decays have also been extensively investigated, and have been key in making strong constraints on the space of possible NP effects. Decays mediated through the charged $W^\pm$ bosons, like the $b \rightarrow c\ell^-\bar{\nu}_\ell$ transition, have proven to be fundamental to determining and

overconstraining the elements of the CKM matrix.

Anomalies in both $b \rightarrow c\ell^-\bar{\nu}_\ell$ and $b \rightarrow s\ell^+\ell^-$ transitions have been observed, in differential branching fractions, branching fraction ratios, and angular observables. Branching fraction ratios of decays involving leptons from different generations are of particular interest, as they test a property of the SM known as lepton flavour universality (LFU). This is an accidental symmetry of the SM, *i.e.* it does not emerge from the underlying $SU(3) \times SU(2)_L \times U(1)$ symmetry, which results in the electroweak couplings to the three charged leptons being equal. Under LFU, the only differences between observables involving different lepton flavours should arise from the differences in their masses.

A longstanding tension has existed for the decays $\bar{B} \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell$ where $D^{(*)}$ indicates a D^0 , D^{*+} or D^{*0} meson, and ℓ^- is either a tau or a muon. In particular, the ratios of branching fractions $R(D^{(*)}) = \mathcal{B}(\bar{B} \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B} \rightarrow D^{(*)}\mu^-\bar{\nu}_\mu)$, obtained from a combination of measurements from the BaBar, Belle, Belle II, and LHCb collaborations, are in tension with the SM predictions at a level larger than three standard deviations. Similar branching fraction ratios have been measured for $B \rightarrow M\ell^+\ell^-$ decays, where B is a $B^\pm$ or B^0 meson, M is either a spin-0 or spin-1 meson containing an s quark, and ℓ is either a muon or an electron. The most recent measurement, a simultaneous determination of the ratios $R(K) = \mathcal{B}(B^+ \rightarrow K^+\mu^+\mu^-)/\mathcal{B}(B^+ \rightarrow K^+e^+e^-)$ and $R(K^*) = \mathcal{B}(B^0 \rightarrow K^{*0}\mu^+\mu^-)/\mathcal{B}(B^0 \rightarrow K^{*0}e^+e^-)$ [29, 30], shows good agreement with the SM, at the level of one standard deviation. Figure 2.4 shows the current experimental status for the LFU ratios $R(D^{(*)})$, and the result of the most recent measurement of $R(K)$ and $R(K^*)$.

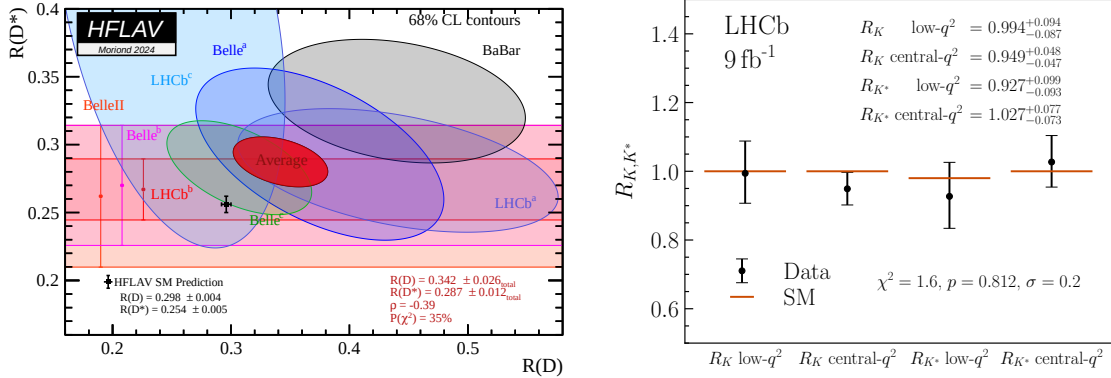


Figure 2.4: Experimental status of LFU ratios in $b \rightarrow c\ell^-\bar{\nu}_\ell$ and $b \rightarrow s\ell^+\ell^-$ decays. Left: current experimental average of the ratios $R(D^{(*)})$, computed by the Heavy Flavour Averaging Group (HFLAV). The resulting combination, reported as a 68% confidence level contour (red shaded area), is compared to the SM prediction (black marker). Right: most recent determination of the ratios $R(K)$ and $R(K^*)$, in bins of the dilepton mass squared, q^2 , measured by the LHCb experiment. The results are compared to the SM expectations which are close to unity. Figures taken from Refs. [27] and [29].

Other interesting tensions have been reported in the rates and angular observables of the

$B \rightarrow M\ell^+\ell^-$ decays. In particular, an anomaly for the angular observable P'_5 (defined as $S_5/\sqrt{F_L(1-F_L)}$, where the observables S_5 and F_L will be introduced in Sec. 2.4.2) has been reported by LHCb in $B^0 \rightarrow K^{*0}\mu^+\mu^-$ decays [31], with consistently anomalous results also reported by the ATLAS, CMS, and Belle collaborations [32–35]. While P'_5 is not directly accessible in untagged $B_s^0 \rightarrow \phi\ell^+\ell^-$ decays, an anomaly has been reported in the differential decay rate [36]. Figure 2.5 shows the observed tensions in P'_5 and in the differential decay rate of $B_s^0 \rightarrow \phi\ell^+\ell^-$ decays as measured by the LHCb experiment.

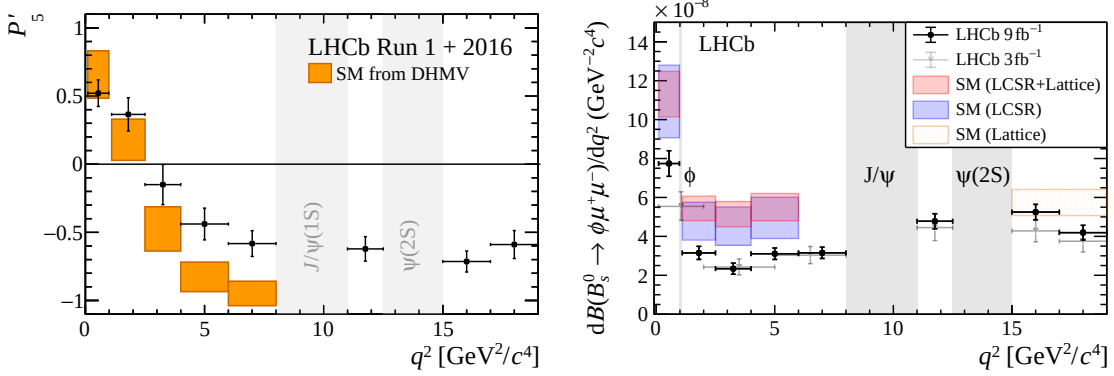


Figure 2.5: Current tensions, with the SM prediction, of the angular observable P'_5 (left), for $B^0 \rightarrow K^{*0}\mu^+\mu^-$ decays, and the differential branching fraction of $B_s^0 \rightarrow \phi\mu^+\mu^-$ decays (right), measured by the LHCb experiment as functions of the dilepton mass squared, q^2 . Figures taken from Refs. [31] and [36].

The picture emerging from these anomalies is still unclear, as more measurements and better theoretical understanding are needed. The work presented in this thesis aims to contribute to the current experimental picture by testing LFU for the first time with B_s^0 meson decays, as well as with the first angular analysis of the decay $B_s^0 \rightarrow \phi e^+e^-$.

2.4 A model independent approach to $b \rightarrow s\ell^+\ell^-$ processes

The Fermi description of the weak interactions [37,38], which was very effective when describing processes at energy scales much smaller than the mass of the $W^\pm$ boson, is considered to be a low-energy approximation of the $SU(2) \times U(1)$ electroweak gauge theory. The arguments of incompleteness and unnaturalness outlined in the previous section have resulted in a similar perspective on the SM, which is understood today as the low-energy limit of a more fundamental QFT.

The language of effective field theory (EFT) [39] comes in handy when describing a physical system without detailed knowledge of its fundamental structure. By setting the energy scale λ relevant to the phenomena one wishes to describe, an EFT is built by “integrating out” the degrees of freedom whose mass M is much larger than the interesting scale, $M \gg \lambda$. The effect of the heavier degrees of freedom is then encoded in dimensionless parameters, known as Wilson coefficients $\mathcal{C}_i^{(n)}$, whose value depends on the chosen scale. The EFT Lagrangian is composed of a sum over a series of point-interaction operators *à la Fermi*, $\mathcal{O}_i^{(n)}$,

each multiplying a corresponding Wilson coefficient,

$$\mathcal{L}_{\text{EFT}} = \mathcal{L}_{d \leq 4} + \sum_{d=5}^{\infty} \frac{1}{M^{d-4}} \sum_i \mathcal{C}_i^{(d)} \mathcal{O}_i^{(d)}, \quad (2.34)$$

where the sum has been split to explicitly separate the operators of (energy) dimension d less than or equal to 4. The first term in the sum contains all terms which can be written from the “light” fields, whose mass is at or below the relevant energy scale λ . The second term is a sum of higher-dimension operators, also built out of the light fields, written as an expansion in terms of powers of $(1/M)$, which results in their explicit suppression at the scale $\lambda \ll M$. To obtain the values of the Wilson coefficients, a matching procedure is performed by computing and comparing the amplitudes of the same process in the SM and in the EFT description. An example of an EFT diagram, compared to corresponding SM diagram, is shown in Fig. 2.6 for the decay $\mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu$.

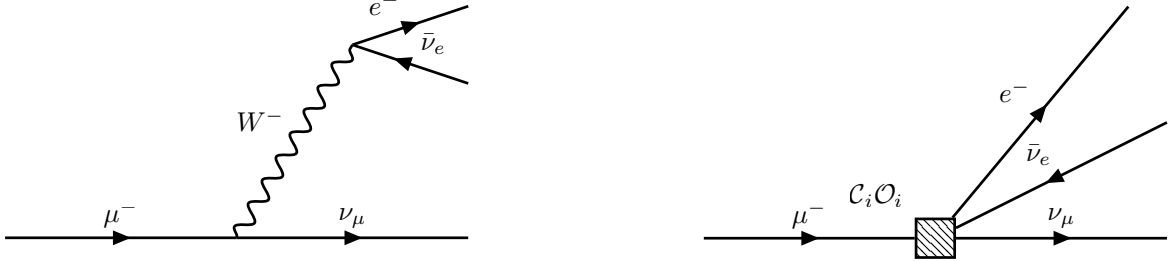


Figure 2.6: Feynman diagram for the muon decay $\mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu$ in the SM (left), mediated by a W^- boson, and in the EFT description (right), where the heavy degree of freedom is replaced by a dimension-6 four-fermion operator $\mathcal{C}_i \mathcal{O}_i$.

The EFT relevant for the decays of b hadrons is known as weak effective field theory (WET), and its typical energy scale is set by the mass of the b quark, $m_b \sim 5 \text{ GeV}$. The WET Lagrangian is expanded in powers of the Fermi constant $G_F = \sqrt{2}g^2/(8M_W^2)$, where g is the $SU(2)_L$ electroweak coupling constant, and M_W is the mass of the W boson [40, 41],

$$\mathcal{L}_{\text{WET}} = \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \sum_i \mathcal{C}_i^{(6)} \mathcal{O}_i^{(6)}, \quad (2.35)$$

where the lowest order which respects all of the symmetries of the SM is dimension $d = 6$. There are six⁴ relevant operators for the $b \rightarrow s \ell^+ \ell^-$ transition, contributing to the effective

⁴There are four-quark operators $\mathcal{O}_{1-6}$ which also contribute, but their effect can be generally absorbed into effective coefficients $\mathcal{C}_{7,9,10}^{\text{eff}}$. There are also scalar and pseudoscalar operators, but these are highly suppressed in the SM and are treated as negligible.

Lagrangian,

$$\begin{aligned}
\mathcal{O}_{7^{(\prime)}} &= \frac{e}{(4\pi)^2} m_b [\bar{s}_{R,L} \sigma^{\mu\nu} b_{R,L}] F_{\mu\nu} \\
\mathcal{O}_{9^{(\prime)}} &= \frac{e^2}{(4\pi)^2} [\bar{s}_{L,R} \gamma^\mu b_{L,R}] [\bar{\ell} \gamma_\mu \ell] \\
\mathcal{O}_{10^{(\prime)}} &= \frac{e^2}{(4\pi)^2} [\bar{s}_{L,R} \gamma^\mu b_{L,R}] [\bar{\ell} \gamma_\mu \gamma_5 \ell],
\end{aligned} \tag{2.36}$$

where s, b, ℓ represent the quark and lepton fields, $F_{\mu\nu}$ is the electromagnetic field tensor, γ^μ, γ^5 are Dirac matrices, and $\sigma^{\mu\nu} = i/2[\gamma^\mu, \gamma^\nu]$. The operator $\mathcal{O}_7$ corresponds to the transition $b \rightarrow s\gamma$, where the emitted virtual photon is converted to a dilepton pair. The operators $\mathcal{O}_{9,10}$ instead correspond to the vector and axial-vector components of the Z boson mediated process in Fig. 2.2. The primed operators correspond to the same transitions but with opposite chiralities for the quarks and are thus very suppressed in the SM.

The flavour measurements presented in the previous section can be interpreted with WET to build a global picture of the current anomalies. Global fits have been performed [42–44], pointing to a deviation in the value of the Wilson coefficient $\mathcal{C}_9$. Figure 2.7 shows the result of one of these global fits, which fit for the deviation from the SM values of the Wilson coefficients $\mathcal{C}_i^{\text{NP}} = \mathcal{C}_i - \mathcal{C}_i^{\text{SM}}$. The input to the fits are experimental measurements of various observables related to the $b \rightarrow s\ell^+\ell^-$ transition by the LHCb, ATLAS, CMS, and Belle II collaborations.

2.4.1 Writing down the amplitudes

This section will now approach in more detail the angular decay rate for the $B_s^0 \rightarrow \phi\ell^+\ell^-$ process, which is especially relevant for Chap. 6. The general approach to the description of the decay is to decompose its total amplitude in simpler terms. Since the initial state is a spin-0 meson, the total amplitude squared will be equal to a incoherent sum of the squared amplitudes, each corresponding to unique helicity states for the two leptons $\lambda_{+,-} \in \{1/2, -1/2\}$,

$$|\overline{\mathcal{M}}|^2 = \sum_{\lambda_+, \lambda_-} |\mathcal{M}_{\lambda_+, \lambda_-}|^2, \tag{2.37}$$

where the overline indicates that the amplitude is spin-averaged. The amplitudes can be first decomposed into the two subsequent decays $B_s^0 \rightarrow \phi\ell^+\ell^-$ and $\phi \rightarrow K^+K^-$, adding a sum over the helicity states of the ϕ meson,⁵ $\lambda_\phi = 0, \pm 1$,

$$\mathcal{M}_{\lambda_+, \lambda_-} = \sum_{\lambda_\phi} \mathcal{M}_{\lambda_+, \lambda_-, \lambda_\phi}^{B_s^0 \rightarrow \phi\ell^+\ell^-} \mathcal{M}_{\lambda_\phi}^{\phi \rightarrow K^+K^-}. \tag{2.38}$$

The $B_s^0 \rightarrow \phi\ell^+\ell^-$ amplitude can be expressed using the WET Lagrangian of Eq. (2.35),

⁵This description is restricted to the case of a resonant $\phi \rightarrow K^+K^-$ decay. The non-resonant, spin-0 case will only be considered for systematic studies in the analysis presented in Chap. 6.

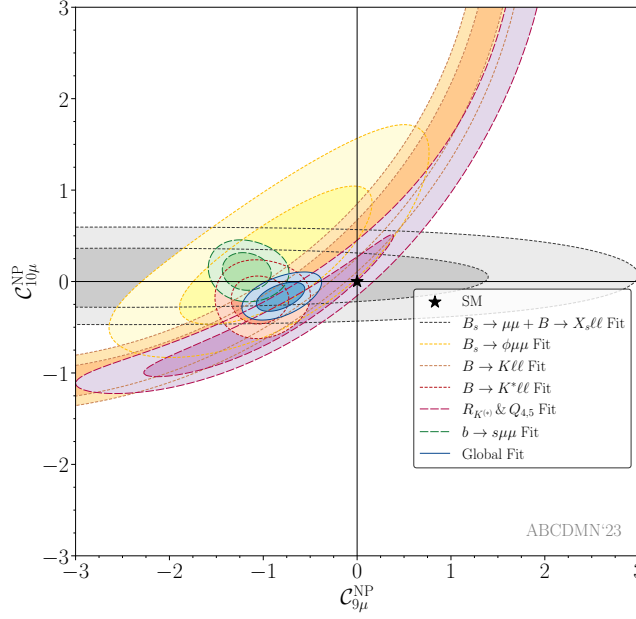


Figure 2.7: Result of the global fit to observables related to the $b \rightarrow s \ell^+ \ell^-$ transition in the framework of WET. The light and dark shaded areas show the 68.3% and 95.4% confidence regions for the deviation of Wilson coefficients from their SM predictions $\mathcal{C}_i^{\text{NP}} = \mathcal{C}_i - \mathcal{C}_i^{\text{SM}}$. A global fit is shown (blue contour) as well as partial fits using the inputs from each decay mode separately (short-dashed contours), with only the LFU observables (burgundy contour), or only the muon modes (green contour). Figure taken from Ref. [42].

which allows for the hadronic and leptonic parts to be factorised,

$$\begin{aligned} \mathcal{M}_{\lambda_+, \lambda_-, \lambda_\phi}^{B_s^0 \rightarrow \phi \ell^+ \ell^-} &= \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \sum_{i=7,9,10} \langle \ell^+ \ell^- \phi | \mathcal{C}_i \mathcal{O}_i | B_s^0 \rangle \\ &= \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \sum_{i=7,9,10} \langle \ell^+ \ell^- | \Gamma_{i,\text{lep.}}^\mu | 0 \rangle \langle \phi | \Gamma_{i,\text{had.}}^\nu | B_s^0 \rangle \eta_{\mu\nu}, \end{aligned} \quad (2.39)$$

where $\Gamma_{i,\text{lep.}}^\mu$ and $\Gamma_{i,\text{had.}}^\nu$ are the Dirac bilinears acting on the lepton and quark fields, and $\eta_{\mu\nu} = \text{diag}(1, -1, -1, -1)$ is the metric tensor. In the right hand side of the last expression, the indices indicating the helicity are suppressed for simplicity for the states $|\ell^+ \ell^- \rangle$ and $|\phi \rangle$.

A further simplification is possible if one considers the $B_s^0 \rightarrow \phi \ell^+ \ell^-$ decay to be the result of two two-body decays, $B_s^0 \rightarrow V^* \phi$, $V^* \rightarrow \ell^+ \ell^-$, where V^* is a virtual boson of either spin 0 or 1. One can then turn the sum over the indices μ, ν into a sum over the helicity states of V^* , $\lambda_V = 0, \pm 1, t$, where t indicates the time-like polarisation,

$$\mathcal{M}_{\lambda_+, \lambda_-, \lambda_\phi}^{B_s^0 \rightarrow \phi \ell^+ \ell^-} = \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \sum_{\lambda_V} \sum_{i=7,9,10} \underbrace{\langle \ell^+ \ell^- | \Gamma_{i,\text{lep.}}^\mu | 0 \rangle \epsilon_\mu(\lambda_V)}_{\mathcal{M}_{\lambda_+, \lambda_-, \lambda_V, i}^{V^* \rightarrow \ell^+ \ell^-}} \underbrace{\langle \phi | \Gamma_{i,\text{had.}}^\nu | B_s^0 \rangle \epsilon_\nu^*(\lambda_V)}_{\mathcal{M}_{\lambda_\phi, \lambda_V, i}^{B_s^0 \rightarrow V^* \phi}} \eta_{\lambda_V \lambda_V}, \quad (2.40)$$

and $\epsilon_\mu(\lambda_V)$ is the polarisation vector of the virtual boson. This final result allows to treat the three amplitudes, $\mathcal{M}_{\lambda_\phi, \lambda_V, i}^{B_s^0 \rightarrow V^* \phi}$, $\mathcal{M}_{\lambda_+, \lambda_-, \lambda_V, i}^{V^* \rightarrow \ell^+ \ell^-}$, and $\mathcal{M}_{\lambda_\phi}^{\phi \rightarrow K^+ K^-}$ in their respective centre-of-mass

frames.

Angular definitions and helicity frames

The rest frames of the B_s^0 , K^+K^- , and $\ell^+\ell^-$ systems are defined in Fig. 2.8. The vertical x axes are chosen to be aligned in the three frames. The angle θ_ℓ is the angle between the e^+ flight direction and the opposite of the B_s^0 direction in the $\ell^+\ell^-$ rest frame (the frame of reference indicated with superscripts)

$$\cos \theta_e = \frac{(\vec{p}_{e^+})^{\ell^+\ell^-} \cdot (-\vec{p}_{B_s^0})^{\ell^+\ell^-}}{|\vec{p}_{e^+}|^{\ell^+\ell^-} |\vec{p}_{B_s^0}|^{\ell^+\ell^-}} = \frac{(\vec{p}_{e^+})^{\ell^+\ell^-} \cdot (\vec{p}_{\ell^+\ell^-})^{B_s^0}}{|\vec{p}_{e^+}|^{\ell^+\ell^-} |\vec{p}_{\ell^+\ell^-}|^{B_s^0}}. \quad (2.41)$$

Similarly, θ_K is the angle between the K^+ flight direction and the opposite of the B_s^0 direction in the K^+K^- rest frame

$$\cos \theta_K = \frac{(\vec{p}_{K^+})^{K^+K^-} \cdot (-\vec{p}_{B_s^0})^{K^+K^-}}{|\vec{p}_{K^+}|^{K^+K^-} |\vec{p}_{B_s^0}|^{K^+K^-}} = \frac{(\vec{p}_{K^+})^{K^+K^-} \cdot (\vec{p}_{K^+K^-})^{B_s^0}}{|\vec{p}_{K^+}|^{K^+K^-} |\vec{p}_{K^+K^-}|^{B_s^0}}. \quad (2.42)$$

The angle $\Phi \equiv \Phi_\ell + \Phi_K$ is defined with the following sign convention, starting from the normal vectors to the $\ell^+\ell^-$ and K^+K^- decay planes, in the B_s^0 rest frame, $(\hat{n}_{\ell^+\ell^-})^{B_s^0}$ and $(\hat{n}_{K^+K^-})^{B_s^0}$,

$$\cos \Phi = (\hat{n}_{K^+K^-})^{B_s^0} \cdot (\hat{n}_{\ell^+\ell^-})^{B_s^0}, \quad \sin \Phi = \left[(\hat{n}_{K^+K^-})^{B_s^0} \times (\hat{n}_{\ell^+\ell^-})^{B_s^0} \right] \cdot \frac{(\vec{p}_{K^+K^-})^{B_s^0}}{|\vec{p}_{K^+K^-}|^{B_s^0}}. \quad (2.43)$$

Within the three rest frames, one can define helicity frames, which are special reference frames where the z axis is aligned to a particle's momentum. These frames are useful to simplify the computation of the decay amplitudes. The helicity states defined in Eq. (2.40) are not however rotationally-invariant. One must then apply appropriate transformations to rotate the helicity states from each decay. The rotation of a particle state with total spin J and projection onto the z axis m , $|J, m\rangle$, onto a new frame with a rotated axis z' , where the projection is m' , is described by the Wigner D-matrices [45]. These are unitary matrices of dimension $2J + 1$, and are defined as

$$D_{m'm}^J(\alpha, \beta, \gamma) = \langle J, m' | \mathcal{R}(\alpha, \beta, \gamma) | J, m \rangle, \quad (2.44)$$

where $\mathcal{R}(\alpha, \beta, \gamma)$ is the rotation matrix defined by three Euler angles α, β, γ . In this description only two angles are needed to rotate each z axis into the relevant momentum direction, hence one can arbitrarily set the additional degree of freedom in the rotation. A common choice is to set $\alpha = -\gamma$. The matrices are then written in terms of real-valued d-matrix elements which only depend on the second rotation angle, $d_{m'm}^J(\beta)$,

$$D_{m'm}^J(\alpha, \beta, -\alpha) = e^{i\alpha(m-m')} d_{m'm}^J(\beta). \quad (2.45)$$

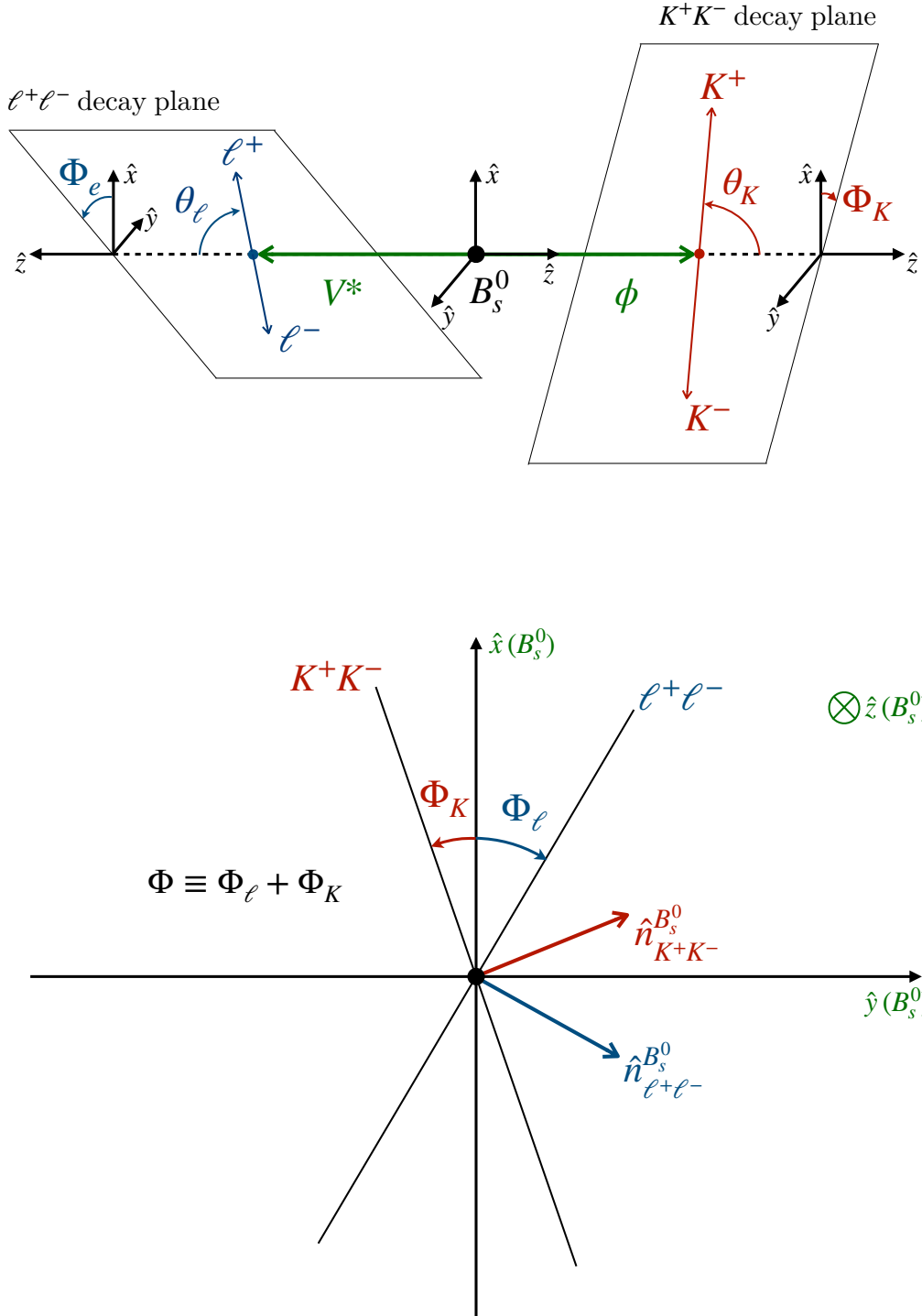


Figure 2.8: Top: Definition of the B_s^0 , K^+K^- , and $\ell^+\ell^-$ rest frames, as well as the relevant decay angles, θ_ℓ , θ_K , and $\Phi \equiv \Phi_\ell + \Phi_K$. Bottom: lateral view in the B_s^0 rest frame, showing the normal vectors to the K^+K^- , and $\ell^+\ell^-$ decay planes, $\hat{n}_{K^+K^-}^{B_s^0}$, $\hat{n}_{\ell^+\ell^-}^{B_s^0}$, which define the angles Φ_ℓ , Φ_K , and their sum $\Phi \equiv \Phi_\ell + \Phi_K$.

The approach to compute the decay amplitudes is to rotate the rest frames of the decay particle into the helicity frame of one of the children. In the $B_s^0 \rightarrow V^* \phi$ decay, the B_s^0 has no polarisation, which allows an arbitrary choice of the orientation of its rest frame so that the z axis coincides with the direction of the ϕ momentum. Thus, the only condition that arises from the conservation of angular momentum is that the helicities of the V^* and ϕ are equal, $\lambda_V = \lambda_\phi$, and no further rotations are needed. For the decay $\phi \rightarrow K^+ K^-$, the amplitude is computed in the helicity frame of the positive kaon, thus one must perform a rotation $D_{\lambda_\phi, 0}^1(\Phi_K, \theta_K, -\Phi_K)^*$, where the second index is zero because kaons are spin-0 particles. Finally, the amplitude of the decay of the virtual boson $V^* \rightarrow \ell^+ \ell^-$ is computed in the helicity frame of the positron, and the matrix element is $D_{\lambda_\phi, \lambda_+ - \lambda_-}^1(\Phi_\ell, \theta_\ell, -\Phi_\ell)^*$.

Combining Eqs. (2.40) and 2.38, and introducing the helicity amplitudes, corresponding to the helicity frames defined above, $\mathcal{H}_{\lambda_\phi}^{B_s^0 \rightarrow V^* \phi}$, $\mathcal{H}_{\lambda_\phi, \lambda_+, \lambda_-}^{V^* \rightarrow \ell^+ \ell^-}$ and $\mathcal{H}_{\lambda_\phi}^{\phi \rightarrow K^+ K^-}$, results in

$$\begin{aligned} \mathcal{M}_{\lambda_+, \lambda_-} = \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* e^{-i(\lambda_+ - \lambda_-)\Phi_\ell} \sum_{\lambda_\phi} \eta_{\lambda_\phi \lambda_\phi} \left(e^{i\lambda_\phi(\Phi_K + \Phi_\ell)} d_{\lambda_\phi, \lambda_+ - \lambda_-}^1(\theta_\ell) d_{\lambda_\phi, 0}^1(\theta_K) \right. \\ \left. \cdot \mathcal{H}_{\lambda_\phi}^{\phi \rightarrow K^+ K^-} \cdot \sum_i \mathcal{H}_{\lambda_\phi, i}^{B_s^0 \rightarrow V^* \phi} \cdot \mathcal{H}_{\lambda_\phi, \lambda_+, \lambda_-, i}^{V^* \rightarrow \ell^+ \ell^-} \right). \end{aligned} \quad (2.46)$$

A noteworthy feature of the last expression is the overall phase in front, which disappears in the incoherent sum over the electron helicities states. Thus there is no individual dependence on Φ_ℓ or Φ_K , the only dependence left is on their sum, Φ .

Hadronic form factors

The hadronic contributions of the helicity amplitudes will contain terms which mix the quark-level weak interactions with the QCD effects which arise from bound states of quarks, such as $\langle \phi | \bar{s}_L \gamma^\mu b_L | B_s^0 \rangle$. Such effects can be visualised through a more realistic Feynman diagram of the $B_s^0 \rightarrow \phi \ell^+ \ell^-$ decay, which is decorated by an arbitrary amount of gluon exchanges between the quarks present, like shown in Fig. 2.9.

In this picture, the s quark in the diagram is no longer a “spectator” of the decay, but rather it can interact and affect its dynamics. The effects are included in the calculation of the element $\langle \phi | \bar{s}_L \gamma^\mu b_L | B_s^0 \rangle$ by writing it as a sum of the possible combination of four-vectors which characterise the kinematics of the decay. In the case of the $B_s^0 \rightarrow \phi$ transition, these are the momenta of the ϕ meson, p^μ , and of the lepton pair, q^μ . Each term of the sum multiplies a function, called a form factor, which only depends on the squared magnitude of the dilepton momentum, q^2 . The q^2 variable is particularly important as it quantifies the energy-momentum transfer in the hadronic transition.

There exist many different theoretical approaches to the computation of hadronic form factors, which apply to different dynamical regimes of a given decay. For the $B_s^0 \rightarrow \phi \ell^+ \ell^-$ decay, the relevant methods are the light cone sum rules [46], which are valid for the region of large hadronic recoil, $q^2 \lesssim 6 \text{ GeV}^2/c^4$, and lattice QCD [47], which covers the region of

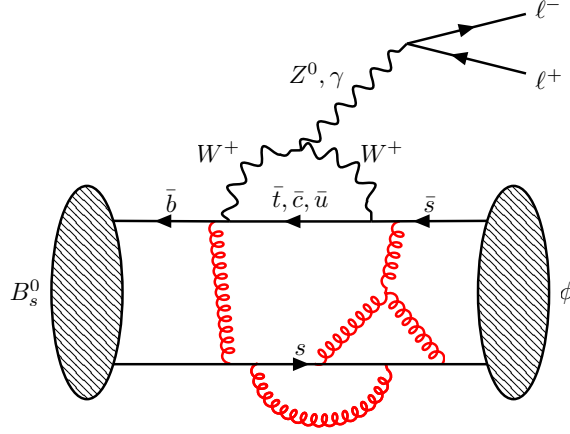


Figure 2.9: Feynman diagram for the decay $B_s^0 \rightarrow \phi \ell^+ \ell^-$, decorated with a number of gluon exchanges (red), to show the effects of QCD on the bound states of the B_s^0 and ϕ mesons.

low recoil, $q^2 \gtrsim 15 \text{ GeV}^2/c^4$. Form factor predictions are the leading source of theoretical uncertainty in the calculation of decay rates, because of the fundamentally non-perturbative nature of QCD at low-energies.

Observables built out of ratios of other observables sharing the same hadronic contribution, can greatly benefit from a cancellation of uncertainties in the ratio. A class of observables relevant to this thesis are LFU ratios, R_H , which were briefly introduced in Sec 2.3.1. These involve the ratio of branching fractions of a b hadron, H_b to a s hadron, H_s , and pairs of charged leptons of two different flavours.⁶ They are typically defined over a range $q_1^2 < q^2 < q_2^2$,

$$R_H(q_1^2, q_2^2) = \frac{\int_{q_1^2}^{q_2^2} \frac{d\Gamma}{dq^2}(H_b \rightarrow H_s \ell \ell) dq^2}{\int_{q_1^2}^{q_2^2} \frac{d\Gamma}{dq^2}(H_b \rightarrow H_s \ell' \ell') dq^2}, \quad (2.47)$$

where ℓ and ℓ' indicate the two different lepton flavours. If $H_b = B_s^0$ and $H_s = \phi$ one gets the ratio R_ϕ , which is predicted to be close to unity in the SM with very small uncertainty [48]. The first ever measurement of R_ϕ is presented in Chap. 5.

2.4.2 Total angular decay rate

A useful tool when writing down the decay rate of the $B_s^0 \rightarrow \phi(\rightarrow K^+ K^-) \ell^+ \ell^-$ decay is to introduce the transversity basis. The expression for the product of the different helicity amplitudes appearing in Eq. (2.46) can be rearranged to again explicitly factorise the hadronic and leptonic components, in particular distinguishing left- and right-handed lepton contributions, to form a total helicity amplitude $M_{\lambda_\phi} \hat{h}_{\lambda_+, \lambda_-}^{L,R}$. The transversity amplitudes

⁶Similar ratios can be built for b hadron decays involving a $b \rightarrow c \ell^- \bar{\nu}_\ell$ transition.

$A_i^{L,R}$, where $i = 0, \parallel, \perp, t$, are then defined as

$$A_{\parallel}^{L,R} = \frac{M_{+1} + M_{-1}}{\sqrt{2}}, \quad A_{\perp}^{L,R} = \frac{M_{+1} - M_{-1}}{\sqrt{2}}, \quad A_0 = M_0, \quad A_t = M_t, \quad (2.48)$$

and have well defined CP properties, A_0 and $A_{\parallel}$ being CP -even, and $A_{\perp}$ being CP -odd.

Squaring the total amplitude in Eq. (2.46), summing over the electron helicity states, and plugging it into the expression for the differential decay rate in Eq. (2.18), yields an expression of the form,

$$\frac{d\Gamma}{dq^2 d\cos\theta_K d\cos\theta_{\ell} d\Phi} = \sum_i J_i(q^2) f_i(\cos\theta_K, \cos\theta_{\ell}, \Phi), \quad (2.49)$$

where $J_i(q^2)$ are the angular coefficients, with $i = 1c, 1s, 2c, 2s, 3, 4, 5, 6c, 7, 8, 9$, containing bilinear combinations of the transversity amplitudes, and $f_i(\cos\theta_K, \cos\theta_{\ell}, \Phi)$ are functions of the decay angles, which appear from products of the d-matrices.

Since the $B_s^0 \rightarrow \phi \ell^+ \ell^-$ final state is flavour-symmetric, the more relevant quantity experimentally is the flavour-averaged decay rate at $t = 0$, indicating the $\bar{B}_s^0$ decay rate as $\bar{\Gamma}$,

$$\begin{aligned} \frac{1}{d(\Gamma + \bar{\Gamma})/dq^2} \frac{d(\Gamma + \bar{\Gamma})}{dq^2 d\cos\theta_K d\cos\theta_{\ell} d\Phi} &= \frac{9}{32\pi} \left[\frac{3}{4} (1 - S_{1c}) \sin^2\theta_K \left(1 + \frac{1}{3} \cos 2\theta_{\ell} \right) \right. \\ &\quad + S_{1c} \cos^2\theta_K (1 - \cos 2\theta_{\ell}) + S_3 \sin^2\theta_K \sin^2\theta_{\ell} \cos 2\Phi \\ &\quad + S_4 \sin 2\theta_K \sin 2\theta_{\ell} \cos \Phi + A_5 \sin 2\theta_K \sin \theta_{\ell} \cos \Phi \\ &\quad + A_{6s} \sin^2\theta_K \cos \theta_{\ell} + S_7 \sin 2\theta_K \sin \theta_{\ell} \sin \Phi \\ &\quad \left. + A_8 \sin 2\theta_K \sin 2\theta_{\ell} \sin \Phi + A_9 \sin^2\theta_K \sin^2\theta_{\ell} \sin 2\Phi \right], \end{aligned} \quad (2.50)$$

where $S_i(q^2) = (J_i + \bar{J}_i)/\frac{d(\Gamma + \bar{\Gamma})}{dq^2}$ are CP -averaged quantities, and $A_i(q^2) = (J_i - \bar{J}_i)/\frac{d(\Gamma + \bar{\Gamma})}{dq^2}$ are CP asymmetries. The appearance of both CP symmetries and asymmetries in the decay rate is the result of the use of the same convention for the definition of θ_K , θ_{ℓ} and Φ for both B_s^0 and $\bar{B}_s^0$ meson. A transformation under CP swaps particles with the antiparticles, corresponding to the transformation $\theta_K \rightarrow \theta_K - \pi$, $\theta_{\ell} \rightarrow \theta_{\ell} - \pi$ and $\sin \Phi \rightarrow -\sin \Phi$, which changes the sign for the coefficients with index $i = 5, 6s, 8, 9$.

Some of the angular observables have a direct physical interpretation. For example, the observable S_{1c} corresponds to the longitudinal polarisation fraction of the ϕ meson, $F_L = (|A_0^L|^2 + |A_0^R|^2)/(|A_0^L|^2 + |A_0^R|^2 + |A_{\parallel}^L|^2 + |A_{\parallel}^R|^2 + |A_{\perp}^L|^2 + |A_{\perp}^R|^2)$. The observable A_{6s} is related to the forward-backward CP -asymmetry in the lepton system, $A_{6s} = 4/3 A_{FB}^{CP}(q^2) =$

$4/3(A_{\text{FB}}(q^2) - \bar{A}_{\text{FB}}(q^2))$, where

$$A_{\text{FB}}(q^2) \equiv \left(\int_0^1 \frac{d\Gamma}{dq^2 d\cos\theta_\ell} d\cos\theta_\ell - \int_{-1}^0 \frac{d\Gamma}{dq^2 d\cos\theta_\ell} d\cos\theta_\ell \right) / \frac{d\Gamma}{dq^2} = J_6 / \frac{d\Gamma}{dq^2}, \quad (2.51)$$

and $\bar{A}_{\text{FB}}(q^2)$ is similarly defined for the $\bar{B}_s^0$ decay. The total decay rate as a function of q^2 is

$$\frac{d(\Gamma + \bar{\Gamma})}{dq^2} = \frac{3}{4} \left(2J_{1s}(q^2) + J_{1c}(q^2) \right) - \frac{1}{4} \left(2J_{2s}(q^2) + J_{2c}(q^2) \right). \quad (2.52)$$

In the limit of massless leptons ($m_\ell^2/q^2 \rightarrow 0$) one gets the following expressions for the $J_i(q^2)$ coefficients

$$\begin{aligned} J_{1s} &= \frac{3}{4} \left(|A_\perp^L|^2 + |A_\parallel^L|^2 + (L \rightarrow R) \right), & J_{1c} &= |A_0^L|^2 + (L \rightarrow R), \\ J_{2s} &= \frac{1}{4} \left(|A_\perp^L|^2 + |A_\parallel^L|^2 + (L \rightarrow R) \right), & J_{2c} &= -|A_0^L|^2 + (L \rightarrow R), \\ J_3 &= \frac{1}{2} \left(|A_\perp^L|^2 - |A_\parallel^L|^2 + (L \rightarrow R) \right), & J_4 &= \frac{1}{\sqrt{2}} \left[\Re \left(A_0^L A_\parallel^{L*} + (L \rightarrow R) \right) \right], \\ J_5 &= \sqrt{2} \left[\Re \left(A_0^L A_\perp^{L*} - (L \rightarrow R) \right) \right], & J_{6s} &= 2 \left[\Re \left(A_\parallel^L A_\perp^{L*} - (L \rightarrow R) \right) \right], \\ J_7 &= \sqrt{2} \left[\Im \left(A_0^L A_\parallel^{L*} - (L \rightarrow R) \right) \right], & J_8 &= \frac{1}{\sqrt{2}} \left[\Im \left(A_0^L A_\perp^{L*} + (L \rightarrow R) \right) \right], \\ J_9 &= \left[\Im \left(A_\perp^L A_\parallel^{L*} + (L \rightarrow R) \right) \right], \end{aligned} \quad (2.53)$$

where $(L \rightarrow R)$ indicates the same expressions but with the chiralities flipped. In this limit, not all coefficients are independent, and one gets the relations $J_{1s} = 3J_{2s}$ and $J_{1c} = -J_{2c}$. Furthermore, the dependence on the time-like polarisation state is suppressed by a factor m_ℓ^2/q^2 .

Angular observables in the presence of mixing

Following the description of B meson mixing in Sec. 2.2, the time-dependent transversity amplitudes are given by,

$$A_i(t) = g_+(t)A_i + \frac{q}{p}g_-(t)\tilde{A}_i \quad (2.54)$$

$$\tilde{A}_i(t) = g_+(t)\tilde{A}_i + \frac{p}{q}g_-(t)A_i, \quad (2.55)$$

where the CP -conjugate $\tilde{A}_i = \eta_i \bar{A}_i$ contains a factor $\eta_i = 1$ for the CP -even amplitudes, $A_\parallel, A_0, A_t$, and $\eta_i = -1$ for the CP -odd amplitude $A_\perp$, and $g_\pm(t)$ are the time-evolution functions from Eq. (2.30). The time-dependent angular coefficients $J_i(t)$ (and their CP -conjugates) are obtained by the substitution $A_i \rightarrow A_i(t)$. The time-dependent CP -averaged

decay rate will then contain terms of the form

$$J_i(t) + \tilde{J}_i(t) = e^{-\Gamma t} \left[\left(J_i(t=0) + \tilde{J}_i(t=0) \right) \cosh \frac{\Delta\Gamma t}{2} - h_i \sinh \frac{\Delta\Gamma t}{2} \right], \quad (2.56)$$

where $\tilde{J}_i = \zeta_i \bar{J}_i$, with the factor ζ_i being 1 for $i = 1c, 1s, 2c, 2s, 3, 4, 7$ and -1 for $i = 5, 6c, 8, 9$, to reflect the change in sign from the choice of angular convention. The expression contains a new set of coefficients, h_i , containing bilinear combinations of the transversity amplitudes at $t = 0$. The h_i coefficients create a “dilution” effect on the time-integrated angular observables $\langle J_i + \tilde{J}_i \rangle$ and decay rate $\langle d(\Gamma + \bar{\Gamma})/dq^2 \rangle$,

$$\langle J_i + \tilde{J}_i \rangle \equiv \int_0^\infty \left(J_i(t) + \tilde{J}_i(t) \right) dt = \frac{1}{\Gamma(1-y^2)} \left[J_i(t=0) + \tilde{J}_i(t=0) - y h_i \right], \quad (2.57)$$

$$\left\langle \frac{d(\Gamma + \bar{\Gamma})}{dq^2} \right\rangle = \frac{1}{\Gamma_s(1-y^2)} \left[\frac{3}{4} (2J_{1s}(t=0) + J_{1c}(t=0) - y(2h_{1s} + h_{1c})) \right. \\ \left. - \frac{1}{4} (2J_{2s}(t=0) + J_{2c}(t=0) - y(2h_{2s} + h_{2c})) \right], \quad (2.58)$$

where $y = \Delta\Gamma_s/(2\Gamma_s) \simeq 0.06$, and the q^2 dependence has been omitted for brevity. This dilution effect will be important in the angular analysis of the decay $B_s^0 \rightarrow \phi e^+ e^-$, presented in Chap. 6. The following coefficients h_i are relevant for the analysis

$$h_{1s}(q^2) = \frac{3}{2} \Re \left[e^{i\phi} \left(-A_\perp^{L*} \bar{A}_\perp^L + A_\parallel^{L*} \bar{A}_\parallel^L - A_\perp^{R*} \bar{A}_\perp^R + A_\parallel^{R*} \bar{A}_\parallel^R \right) \right], \quad (2.59)$$

$$h_{1c}(q^2) = 2 \Re \left[e^{i\phi} \left(A_0^{L*} \bar{A}_0^L + A_0^{R*} \bar{A}_0^R \right) \right], \quad (2.60)$$

$$h_{2s}(q^2) = \frac{1}{2} \Re \left[e^{i\phi} \left(-A_\perp^{L*} \bar{A}_\perp^L + A_\parallel^{L*} \bar{A}_\parallel^L - A_\perp^{R*} \bar{A}_\perp^R + A_\parallel^{R*} \bar{A}_\parallel^R \right) \right], \quad (2.61)$$

$$h_{2c}(q^2) = -2 \Re \left[e^{i\phi} \left(A_0^{L*} \bar{A}_0^L + A_0^{R*} \bar{A}_0^R \right) \right], \quad (2.62)$$

$$h_3(q^2) = -\Re \left[e^{i\phi} \left(A_\perp^{L*} \bar{A}_\perp^L + A_\parallel^{L*} \bar{A}_\parallel^L + A_\perp^{R*} \bar{A}_\perp^R + A_\parallel^{R*} \bar{A}_\parallel^R \right) \right], \quad (2.63)$$

where ϕ is related to the mixing parameters in the B_s^0 system, $e^{i\phi} = q/p$.

The LHCb detector at the Large Hadron Collider

The author's recommendation

Imaginal Disk
by **Magdalena Bay**
2024, Mom+Pop Music

The work described in this thesis was carried out using data collected between 2011 and 2018 by the Large Hadron Collider beauty (LHCb) experiment. LHCb is one of the four main experiments housed at the Large Hadron Collider (LHC), situated at the European Organization for Nuclear Research (CERN) in Geneva, Switzerland. In this chapter I introduce the main features of the LHC machine and of the LHCb detector. The software and trigger aspects of the data flow will be also covered. The description of the LHC and LHCb presented is as they were operating up until the most recent Long Shutdown, which lasted from 2019 to 2022. Towards the end of the chapter I will give a brief overview of the upgraded LHCb detector, which was installed during the last LHC Long Shutdown and has been collecting data since 2022. I will give a glimpse into the operation of the upgraded Vertex Locator, for which I have been an on-call expert since the second year of my PhD. I will also describe the revamped online monitoring system for the muon detector, of which I have been the main author and maintainer.

3.1 The Large Hadron Collider

The LHC [49] is a two-ring hadron accelerator, installed in the tunnel of 26.7 km circumference that was originally constructed to house the Large Electron-Positron collider [50] at CERN. The LHC mainly collides protons, but it has a rich heavy ion programme, performing collisions of $p\text{Pb}$ and PbPb , as well as xenon-xenon, and with plans for oxygen-oxygen; for the purpose of this thesis, only pp collisions will be discussed.

To collide particles, two rings with ultra-high vacuum are needed, where the beams circulate in opposite directions. The rings are divided in eight quadrants, and are stored

underground, at a variable depth which ranges between 40 and 170 m below the surface. They intersect in four points along the circumference, where particle collisions can occur. The four LHC experiments, the ATLAS, ALICE, CMS, and LHCb detectors, are located at these interaction points (Point 1, 2, 5, and 8, respectively).

The complex includes an injection chain, needed to accelerate the beams to an energy of up to 6.5 TeV each,¹ with the LHC being the last acceleration step before letting the beams collide. The different stages of acceleration, in the context of CERN’s accelerator complex, are shown in Fig. 3.1. The first accelerator in the chain is a linear accelerator. The linear accelerator 2 (Linac 2) takes protons extracted from an hydrogen ion source and accelerates them up to an energy of 50 MeV. Then the protons enter three increasingly larger circular accelerators, and after each stage the beam energy grows significantly: the Proton Synchrotron Booster (up to 1.4 GeV), the Proton Synchrotron (up to 25 GeV), and the Super Proton Synchrotron (SPS). After this last step the protons have reached an energy of 450 GeV and are finally ready to enter the LHC, through transfer lines connecting the SPS to the main rings, located near Point 2 and 8. After injecting the beams, the LHC goes through a series of stages [51, 52], to accelerate and adjust the beams for physics collisions. The acceleration is performed with radio-frequency cavities, operating in straight sections of the ring, near Point 4. Superconducting dipole magnets are used to curve the beam’s trajectory along the ring, while higher order magnets are used for focussing, as well correcting its orbit, or other properties of the beams, like the protons’ momentum distribution. In the final stage, the CERN Control Centre declares the state of “Stable Beams”, and the conditions of the beams are deemed good enough for physics production.

The particles in the LHC circulate in groups of $\approx 10^{11}$ particles called bunches [53]. During the collision time period, also known as fill, there are nominally around 2800 colliding bunches circulating in each ring. The filling pattern, or scheme, determines the bunch spacing for the two beams. They are organised in batches with a bunch spacing of 25 ns, known as bunch trains; each train is separated by a longer spacing, to accomodate the rise times of the SPS and LHC kicker magnets which are used at the injection stage. At end of a train follows the longest gap, which allows to safely extract the beams from the main rings within a very short time (order of a few μ s). When the beams are extracted, or “dumped”, they are redirected to an area situated at Point 6, where they hit beam absorber blocks made of graphite, which are able to withstand the destructive power of the beams. In normal conditions, during the second run of data-taking, LHC fills lasted about 10 hr with a turnaround time (from a dump to stable beams) of about 2–3 hr.

To gauge the physics performance of the LHC, one must take into account two variables: the centre-of-mass energy $\sqrt{s}$, which determines the number and relative proportion of accessible final states in the detector, and the instantaneous luminosity $\mathcal{L}$, which determines for a given time of data-taking how many times a certain process can be produced by the machine. For head on collisions of two beams, each with energy E_{beam} , the centre-of-mass

¹The design energy is 7.0 TeV, but the LHC has run at different energies over the years: 3.5 TeV in 2011, 4.0 TeV in 2012, 6.5 TeV in 2015–2018, and 6.8 TeV since restarting in 2022.

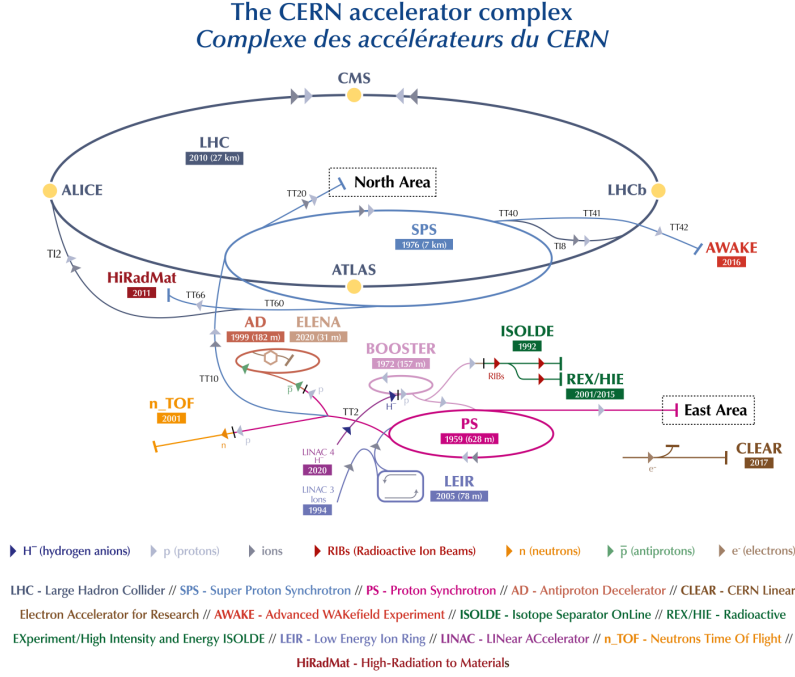


Figure 3.1: The CERN accelerator complex as of 2020, and the many experiments that operate in it. Note that the Linac 2, used for the Run 1 and 2 of data-taking, has been superseded by the new Linac 4 [54]. Figure taken from Ref. [55].

energy at the interaction point can be written as

$$\sqrt{s} = 2E_{\text{beam}}, \quad (3.1)$$

which represents the total available energy from the collision for producing a given final state. Each final state has some probability of being produced, that is usually represented by the cross-section σ : this variable has dimensions of an area [m^2], and it can be intuitively related to the intersecting transverse area of the two colliding beams. In high energy physics, a natural unit of measure for the cross-section is the barn, b , which corresponds to an area of 10^{-28} m^2 . The cross-section for most processes has a dependence on the total available energy, and in the case of particle collisions it can be expressed as a dependence on $\sqrt{s}$. For what concerns the work presented in this thesis, the production of b hadrons in pp collisions has been extensively studied by the LHCb collaboration [56–62], and it is found to be in good agreement with the theoretical predictions.

The instantaneous luminosity of a collision is defined as

$$\mathcal{L} = \frac{1}{\sigma} \frac{dN}{dt}, \quad (3.2)$$

where dN/dt is the number of processes per second generated during collisions (described by the given cross-section σ). The luminosity can be expressed in terms of properties of the

beam,

$$\mathcal{L} = \frac{N_b (N_p^2 f) \gamma}{4\pi\epsilon_n\beta^*} F, \quad (3.3)$$

where N_b is the number of bunches in each beam; N_p is the number of protons per bunch, assumed to be the same in each beam; f is the revolution frequency around the LHC ring; ϵ_n is the beam's normalised emittance, which is related to the area that the beam occupies in the position-and-momentum phase space; β^* is the beta function (at the collision point), which relates to the beam's transverse size; F is a geometrical factor which accounts for the overlap of the beams when colliding. The total number of occurrences of a given final state that can be produced by the machine, for example $b\bar{b}X$, where X is any set of additional particles, is obtained by integrating the previous expression,

$$N = \sigma(pp \rightarrow b\bar{b}X, \sqrt{s}) \int \mathcal{L} dt, \quad (3.4)$$

where the dependence of the cross-section on the centre-of-mass-energy has been explicitly written, and the instantaneous luminosity is integrated over the time of data-taking. As the beams collide, the number of available protons per bunch, and thus the instantaneous luminosity, decreases over a fill, following an exponential curve. A stable luminosity profile can be achieved by changing the beam parameters over a fill, either F or β^* in Eq. (3.3), to counteract this natural decay. This is known as luminosity levelling.

The LHC has completed two runs of data-taking: Run 1 in 2011–2012 at $\sqrt{s} = 7-8$ TeV, and Run 2 in 2015–2018 at $\sqrt{s} = 13$ TeV. The LHC is currently in its third run of physics production, at $\sqrt{s} = 13.6$ TeV, which started in 2022 and will last until the end of 2026. The changes to the machine conditions which are relevant to LHCb will be briefly discussed in the next chapter.

3.2 Overview of the LHCb detector

The principal goal of LHCb is to study at high precision the “heavy-flavour” sector of the Standard Model, which includes all particles containing b and/or c quarks, focussing in particular on charge-parity violation and searches of rare or forbidden decays. The LHCb detector [63] is very different in layout from the other three main LHC experiments, being a forward single-arm spectrometer. Its design naturally follows from the $b\bar{b}$ production kinematics at the LHC, shown in Fig. 3.2, which are heavily boosted in the beams' directions (approximately aligned with the z axis in the LHCb coordinate frame). It can detect particles in a range of pseudorapidity η of 2–5.

LHCb operated during the 2011–2012 and 2015–2018 periods, collecting data amounting to 9 fb^{-1} of integrated luminosity. It ran at a lower instantaneous luminosity than ATLAS and CMS, reaching $\mathcal{L} = 4 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ in Run 2. This design choice was aimed to reduce the number of proton-proton interactions that happen with each colliding bunch. Through a luminosity levelling procedure [65, 66], the beams collide with an offset, so that

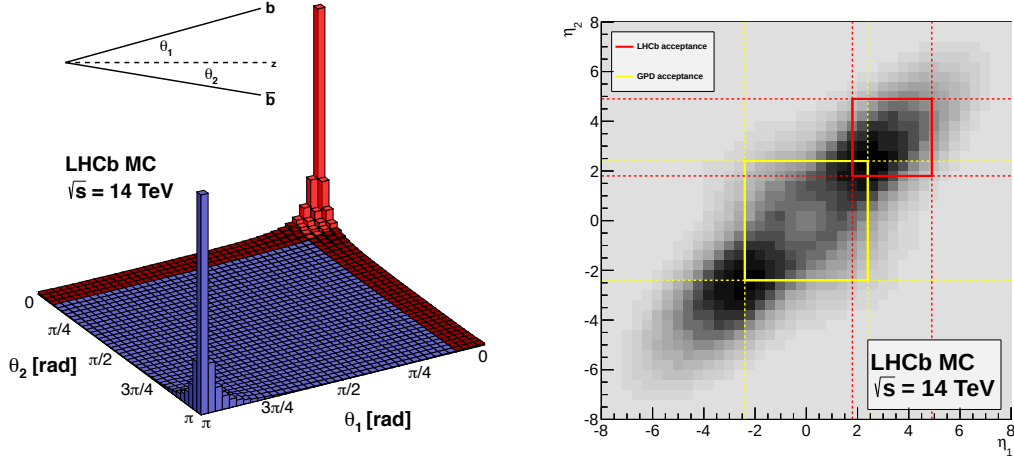


Figure 3.2: Production of $b\bar{b}$ pairs for simulated pp collisions at $\sqrt{s} = 14$ TeV: on the left, the distribution of the $\theta_{1,2}$ angles from the beam axis; on the right, the pseudorapidity distribution $\eta_{1,2}$, and the LHCb acceptance outlined in red. Figures taken from Ref. [64].

their overlap region is reduced in area. With this technique, the number of interactions per colliding bunch goes from approximately 32 to just one, on average. This has important advantages: the detector occupancy, a measure of the rate at which particles interact with each individual detector component, remains constant throughout a fill, reducing the overall radiation damage, as well as potential systematic effects in physics analyses; furthermore, it allows excellent reconstruction performance of the positions of decaying b and c hadrons with respect to the collision point they originated from. Figure 3.3 shows the profile of the instantaneous luminosity $\mathcal{L}$ over a particularly long Run 1 fill, for the ATLAS, CMS, and LHCb experiments.

The LHCb detector is made up of many sub-systems, and the purpose of each component can be categorised into one of three main groups. First, one must be able to reconstruct the “tracks”, *i.e.* the trajectories of the charged particles that pass through the LHCb detector. Then, one must infer the type of the particle that is being tracked, by means of measuring the ways it has interacted with the different detector components. One must also be able to identify neutral particles which do not leave a track in the detector. Lastly, a trigger is needed, which is the algorithm deciding if a particular pp collision is interesting enough to be recorded, by taking key information from the sub-detectors. This last element is fundamental to any successful analysis done at the LHC: at the extremely high collision frequency of up to 40 MHz,² it's not possible to keep a record of all the collisions that happen inside LHCb, because of a physical limit both in storage capacity and in the rate at which data can be recorded. Hence, the trigger can be tuned for example to select collisions where a b hadron is likely to have been produced, and to discard ones where no such particles are present. A

²The average non-empty bunch-crossing rate will depend on the filling scheme and total number of filled proton bunches.

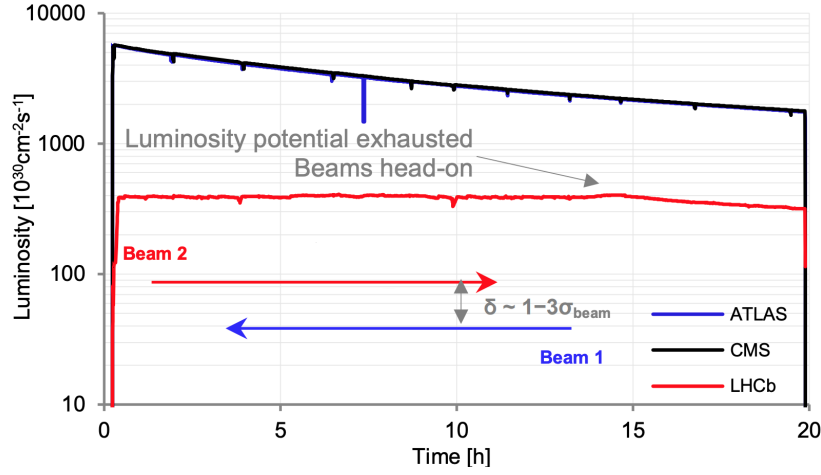


Figure 3.3: ATLAS, CMS, and LHCb instantaneous luminosities during fill 2651. LHCb maintains a constant value of $\mathcal{L}$ by letting the beams collide with a tuned transverse off-set; in ATLAS and CMS head-on collisions result in a luminosity that decreases exponentially with time. Figure taken from Ref. [65].

scheme of the LHCb detector and of its subcomponents is shown in Fig. 3.4. The figure also illustrates the coordinate system defined for the detector. The origin of the reference frame is positioned within the Vertex Locator, and represents the (ideal) beam collision point.

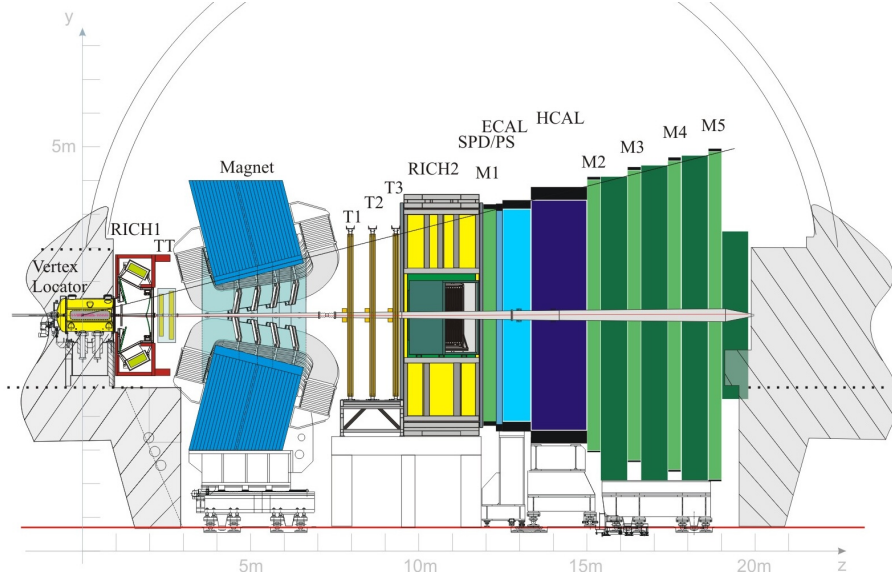


Figure 3.4: View of the LHCb detector and its subcomponents. The y and z axes of the LHCb frame are shown. The x axis points into the page. Figure taken from Ref. [63].

3.3 Particle tracking

The tracking system of the LHCb detector consists of the Vertex Locator (VELO), a dipole magnet, and four tracking stations, the first upstream of the magnet, and the other three

downstream of the magnet. The VELO system allows for high precision reconstruction of the interaction point and of the decay positions of particles originating from the collision. The tracking stations can measure the deflection of the direction of a particle that passes through the dipolar magnetic field, which can be used to calculate the particle’s momentum. When charged particles fly through the tracking detectors, they leave energy deposits which are reconstructed as “hits”, points where they interacted with the detector’s sensitive materials. By finely segmenting each tracking detector with many sensors, of which the positions are known precisely, the three-dimensional coordinates of a hit can be accurately determined. Tracking is then achieved with reconstruction software, which associates the hits to the track of a charged particle.

3.3.1 The Vertex Locator

The VELO is a silicon microstrip detector located in the region surrounding LHCb’s pp collision point. It consists of 42 silicon modules, arranged symmetrically along the beam interaction region, that provide measurements in their polar coordinates (r, ϕ) of hits produced by charged particles originating from the collision. Because of the large Lorentz boost, b or c hadrons can travel in LHCb’s reference frame for a few centimeters before decaying. The proximity of the VELO to the beams allows to precisely reconstruct and distinguish the collision point, known as the primary vertex (PV), and the secondary vertices, which identify where the decays have occurred.

The detector is mounted in a vessel which encloses the interaction region and is kept under the primary LHC vacuum. The modules on each side are placed in secondary vacuums encased by RF-boxes, made of aluminium, which protect the sensors from electromagnetic interferences which may be induced by the passing beams. Keeping the modules in a vacuum minimises the particles’ interactions as they travel from the interaction vertex.

The vessel allows for the modules on the two sides of the interaction region to be displaced from it. This is needed because, when the LHC is preparing the proton beams, these will have generally a larger transverse area, so they can pose a danger to the sensitive components of the VELO. Under these conditions, the detector is kept in an open configuration for safety, at 30 mm away from the beams. Once stable beams are declared, the VELO is brought into its closed position, at only 8.2 mm away from the centre of the interaction region. Each time the VELO is closed, it is recentred in the $x - y$ transverse plane, to allow for shifts in the beam position. In this state, the two halves of the detector overlap to achieve maximum proximity to the interaction point, and to cover the full azimuthal acceptance. Figure 3.5 (left) shows both configurations.

The side of each RF-box which faces the beamline, known as the RF-foil, is corrugated to allow for the overlap of the modules in the closed position. The foil is made to be about 0.3 mm thick, which reduces the material budget of the detector, while still offering the needed RF shielding. The thinness of the foil is achieved by reducing the pressure differential between the primary and secondary vacuums. Figure 3.6 shows a model of a VELO half,

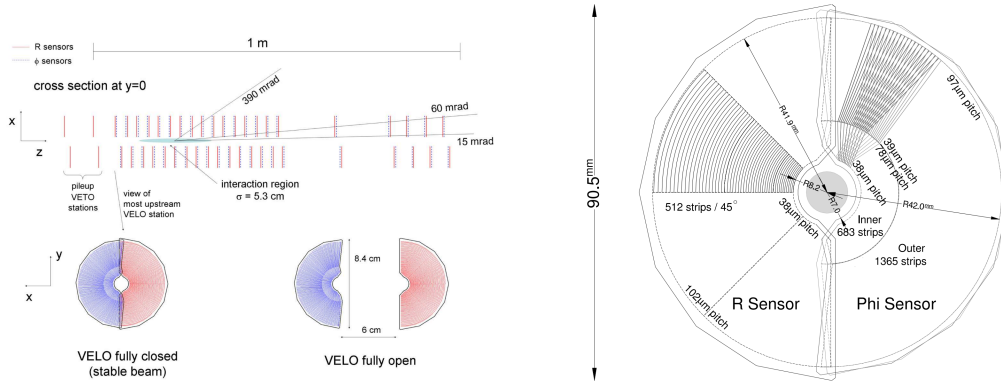


Figure 3.5: On the left, cross-section in the $y = 0$ plane with the positions of the 42 VELO modules. The vertical cross-section of two modules, in the closed and opened configurations, are also shown. A more detailed sketch of the (r, ϕ) configuration of a single sensor is shown on the right. Figures taken from Ref. [63].

with the RF box included.

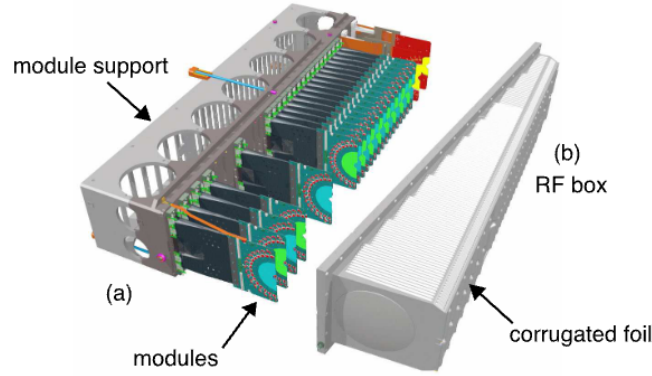


Figure 3.6: Exploded view of the module support and the modules (a), and the RF box (b). The face of the box facing the beam is corrugated, to allow the two halves to overlap when in the closed position. Figure taken from Ref. [63].

Each sensor is composed of two parts: the first measures the radial distance r from the beam pipe, with concentric strips; the second measures the ϕ angle with radial strips, arranged to be perpendicular to the r strips. For both types of strips around the inner region, closer to the beamline, the segmentation is denser to deal with the higher detector occupancy. To further reduce overall occupancy, the radial strips are divided in arcs each spanning an angle of 45 deg, while the ϕ strips are divided into an inner and outer region. The third spatial dimension z is provided by the information of the position of each module within the experiment. Figure 3.5 (right) features a detailed scheme of the strips in all three dimensions, as well as the size of the various components.

3.3.2 The tracking stations

The first tracking station is upstream of the magnet and is known as the Tracker Turicensis (TT). The TT is made of silicon strips arranged in four layers. The first and last layers have vertically aligned strips, whilst the two in the middle have them rotated by a stereo angle of $+5^\circ$ and -5° , for the second and third layer respectively; this configuration is known as $(x-u-v-x)$. The layout of one of the detector layers is shown in Fig. 3.7. The fundamental building blocks for a layer are the half-modules, rows of seven silicon sensors, organised in two or three sectors. A full module, made by joining two half modules end-to-end, covers the full height of the LHCb acceptance. The readout boards, as well as the cooling system and detector supports, are placed at the end of each half-module; being below and above the sensitive area of the detector, they are completely out of its acceptance.

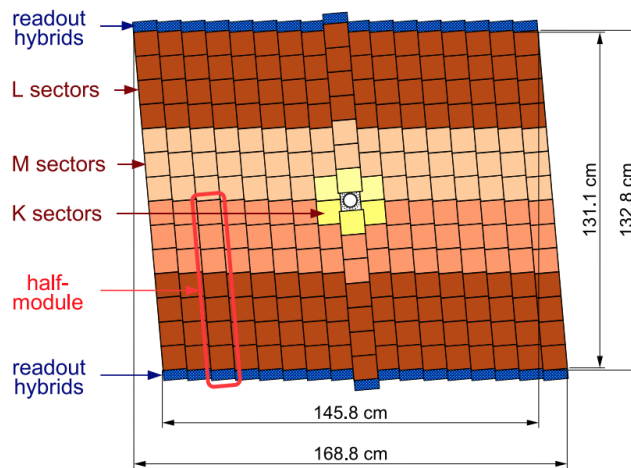


Figure 3.7: Scheme of the TT's third layer of silicon strips. Figure taken from Ref. [63].

The other three tracking stations, T1-2-3, are situated downstream of the magnet and are each made of an Inner Tracker (IT), made of silicon sensors similarly to the TT, and an Outer Tracker (OT), which is made of straw-tube detectors. The stations are also made of four layers arranged in the same $(x-u-v-x)$ geometry as the TT layers.

Figure 3.8 shows the module arrangement for one of the layers of the IT. A single detector module consists, similarly to the TT, of either one or two (vertically aligned) silicon sensors. The readout boards are placed at the top or bottom of the module, depending on its position in the layer. Each module is mounted on a support plate which provides the needed electrical insulation as well as a direct path for heat to flow from the front-end chips to the cooling system. Two rows containing seven single sensor modules are placed above and below the beam pipe; a row of seven two sensor modules is placed on each side of the beam pipe, with a slight overlap with the top and bottom rows, to ensure no acceptance gaps are present.

The OT layers are designed as an array of individual straw-tube modules. The layers, shown in Fig. 3.9, consist of two halves which are independently retractable on each side of the beam pipe. Each module contains two staggered layers of 64 drift-tubes each, containing

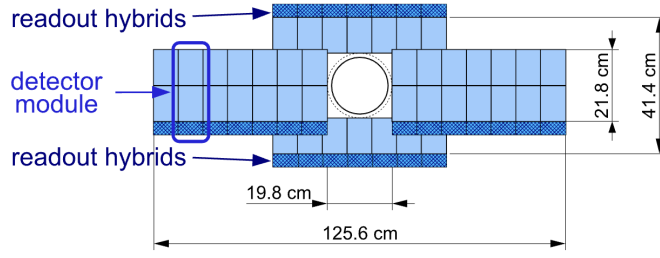


Figure 3.8: Scheme of the silicon sensors of the first layer of the second station of the IT. Figure taken from Ref. [63].

a 70–30 mixture of Argon and CO_2 . On each side, in each layer seven modules cover the vertical acceptance of the detector; the full-length modules are split in the middle into two sections and each has separate readout boards placed at the top and bottom of the layer. Additionally, each layer has four half-length modules, located above and below the beampipe; these are read out only on the outer end of the module.

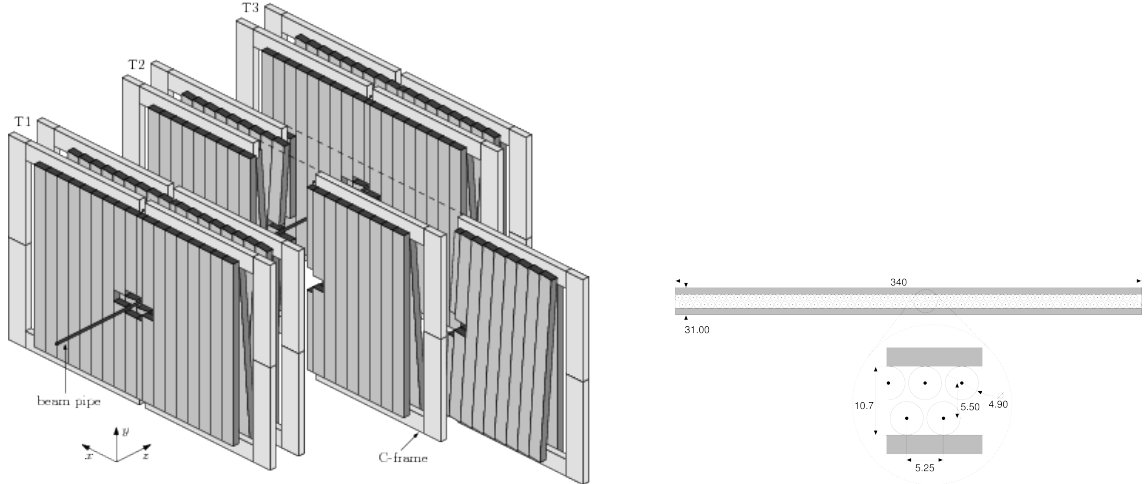


Figure 3.9: Left: scheme of the OT modules in layers and stations, with the second station, T2, shown with one half open. The space cut out around the beam pipe is made to fit the IT modules (not shown in the figure). Right: cross-section of a single straw-tube module. Figures taken from Ref. [67].

3.3.3 The dipole magnet

As introduced in the beginning of this section, the LHCb tracking system uses a dipole magnet to bend charged particles traversing the detector. The bending of a track is measured with the combined information from the upstream and downstream trackers, and is used to provide a momentum measurement. The dipole is a conventional warm magnet made with two saddle-shaped coils, made of aluminium, and placed symmetrically above and below the beam pipe. The coils are inserted in a window-frame yoke with a slope to match the detector's conical acceptance. Figure 3.10 shows a technical scheme of the magnet and its support structure.

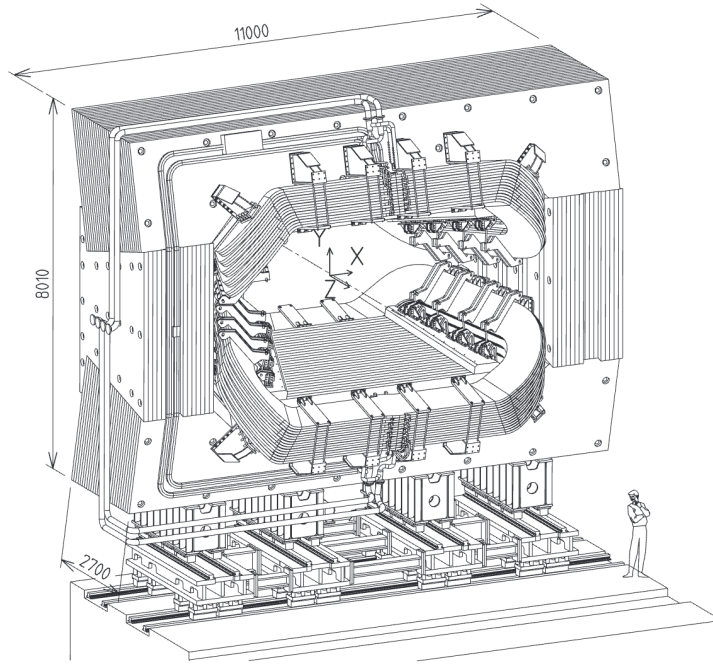


Figure 3.10: Technical drawing of the LHCb dipole magnet and its support structure. Figure taken from Ref. [63].

The magnet was designed to have a field as high as possible in the region between the upstream and downstream trackers, to achieve a target momentum resolution of $0.5 - 1.0\%$, for tracks with momentum in the range of $5 - 100 \text{ GeV}/c$. This means producing an integrated field over a length of 10 m (from the interaction point) of 4 T m . At the same time, the Ring Imaging Cherenkov detectors require a low field for their photon detectors to work properly. This final requirement was achieved by adding additional magnetic shielding which envelops these detectors. The profile of the magnetic field was scanned in a volume that covers most of the LHCb acceptance, spanning from the nominal interaction point to the RICH 2 detector, using sensitive Hall probes. Figure 3.11 shows the profile of the magnetic field along the z axis, in the $x = 0, y = 0$ plane.

The polarity of the magnet was regularly inverted during data-taking, so that about half of the total dataset was collected with one polarity and half with the other. This is especially useful in studies of CP violation, where averaging over the two opposite magnet polarities strongly suppresses charge detection asymmetries and associated systematic effects.

3.3.4 Vertex and track reconstruction

A track can be classified depending on the path it followed inside LHCb and which sub-detector provided hits in its reconstruction [68]:

- **Long tracks** go through all tracking detectors, and are defined by having hits in the VELO and T stations, and optionally in the TT detector. These are the most impor-

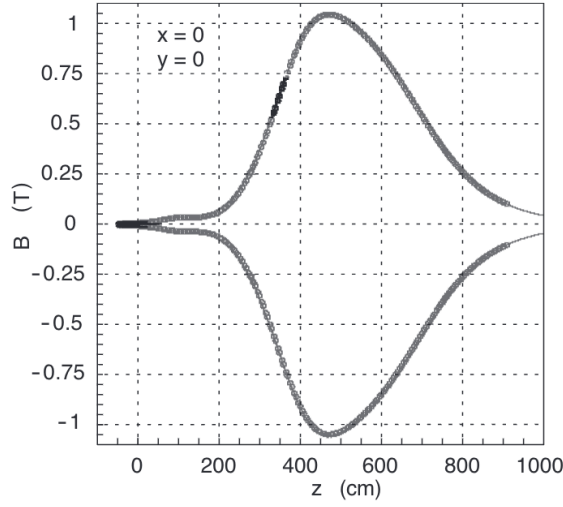


Figure 3.11: Profile of the magnetic field, in both polarities, along the z axis, in the $x = 0, y = 0$ plane. Figure taken from Ref. [63].

tant tracks for physics analyses in LHCb, as they have the most precise momentum measurement.

- **Upstream tracks** only have associated hits in the VELO and TT detectors, and do not usually have high enough momentum to traverse the magnet fully.
- **Downstream tracks** pass only through the TT and T stations, and are important in the reconstruction of long-lived particles like K_S^0 mesons and Λ baryons, which decay out of the VELO acceptance.
- **VELO tracks** are ones that only have hits in the VELO, such as large-angle tracks or ones that go in the backwards ($-z$) direction. They are very useful in reconstructing the PV of the collision.
- **T tracks**, that only pass through the T stations and originate typically from secondary interactions inside the spectrometer.

For long tracks, which will be the ones considered throughout this thesis, the tracking algorithm combines all information coming from the VELO, TT and T stations, and tries to fit them to the most probable track. The algorithm first looks for hits in the VELO that line up in a straight segment, requiring at least three hits in both the r and ϕ sensors. These track segments are then matched to hits in the T stations with two distinct algorithms.

The first one, called forward tracking [69], looks for a compatible hit in one of the T stations, and then searches for hits in the other stations in a window around the plausible trajectory. The information from the VELO segment and the first compatible T station hit is enough to fully determine the direction and momentum of the track. The second algorithm, called track matching [70], tries to combine the VELO information with other

straight segments found in the T stations. The T station tracks are required to have at least one hit in the x layer and one in the stereo layers, in each of the three stations. The two algorithms are then compared and duplicates are removed, usually the ones with less associated hits. Additional hits found in the TT that are compatible with the track are added to improve the momentum resolution.

The final momentum and direction determination is performed with a Kalman fitter [71, 72]. The fit takes into account additional multiple scattering in the detector material, as well as correcting for energy losses from ionisation. The reconstructed track is characterised by a state vector, for a given z position, $(x, y, dx/dz, dy/dz, q/p)$, where q is the track's charge.

Sometimes “ghost” tracks are reconstructed, meaning fictitious tracks that do not correspond to any real particle. Most of them originate from wrong associations of the VELO and T station segments, and in collisions with high track multiplicity they can take up a large fraction of the reconstructed tracks, as much as 20%. A neural network is used to greatly reduce the fake track rate, at the cost of a small reduction in the total tracking efficiency.

The tracking efficiency is estimated with a tag-and-probe method, using a control sample of $J/\psi \rightarrow \mu^+ \mu^-$ candidates.³ In this method, one track is fully reconstructed, the so-called “tag”, whilst the second track, the “probe”, is only partially reconstructed. The probe should have sufficient information to reconstruct the invariant mass of the parent J/ψ . The probe is then matched to a fully reconstructed long track: if such track is found, then the reconstruction was efficient. No trigger or offline requirements are made on the probe track, to avoid biases on the measured efficiency. Plots of the tracking efficiency as a function of momentum, pseudorapidity, the number of tracks and the number of primary vertices in the collision, are shown in Fig. 3.12.

Measurements of rare decays of b hadrons, like the ones described in this thesis, require excellent PV and decay vertex reconstruction, as well as accurate measurements of the tracks' impact parameters. The PV resolution can be measured by comparing two independent measurements of the vertex position from the same collision. This can be achieved by splitting the set of tracks from a collision in two and reconstructing the PV from each subset. The resolution is extracted from the distribution of the difference in the two measured PV positions (corrected with a factor of $\sqrt{2}$, to account for the dependence of the resolution on the total number of tracks). The number of tracks used to reconstruct a vertex goes from a minimum of 5 up to about 150, so with this procedure the resolution can be measured with up to around 65 tracks. The result for the PV resolution in the x and y coordinates is shown in Fig. 3.13.

The impact parameter (IP) of a track is defined as its distance of closest approach, relative to the primary vertex. Because b and c hadrons travel a non-negligible distance from the PV before decaying, the particles originating from their decays tend to have a larger IP than those of particles produced at the collision point (so-called prompt production). As will be

³In general the tracking efficiency can depend on the particle species, as in the case of electrons, which suffer from larger energy losses as they travel through the detector. The differences in the reconstruction of muons and electrons are discussed in more detail in Chap. 4.

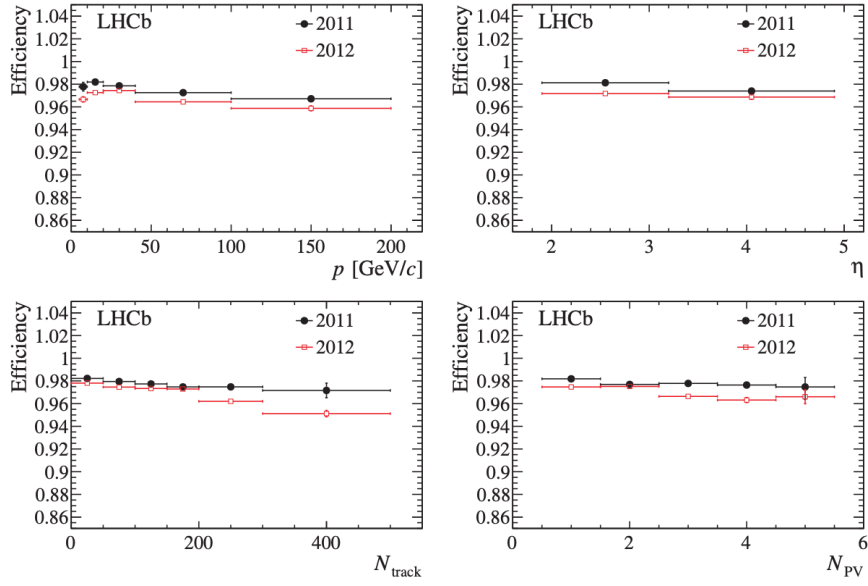


Figure 3.12: Tracking efficiency measured with 2011 and 2012 control samples, as a function of the momentum p , the pseudorapidity η , the number of tracks from the collision N_{track} , and the number of reconstructed primary vertices N_{PV} . The error bars account for the statistical uncertainty. The slightly worse performance in 2012 is due to the higher hit multiplicity at higher $\sqrt{s}$. Figure taken from Ref. [68].

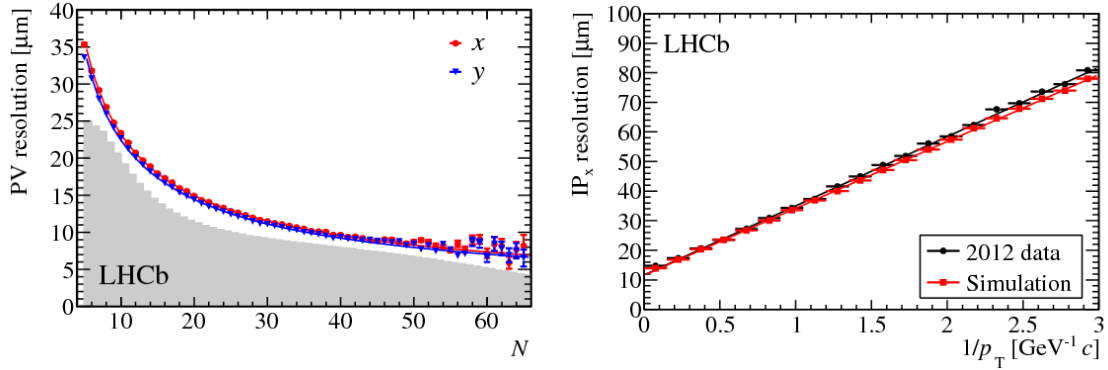


Figure 3.13: Left: primary vertex resolution in the x (red) and y (blue) coordinates, measured with data collected by LHCb in 2012, as a function of track multiplicity. The data points are fit with $1/\sqrt{N}$ curves. The superimposed histogram shows the distribution of the number of tracks per reconstructed PV, for all collisions selected by the high level trigger. Right: the impact parameter resolution in the x direction (black) as a function of the inverse of the transverse momentum, compared to the resolution determined with simulation (red). The data points are fit with a linear function of $1/p_T$. Figure taken from Ref. [68].

explained further in Chap. 4, the analysis of B_s^0 meson decays relies on the variables IP and IP χ^2 to discriminate genuine signal from prompt backgrounds. The IP χ^2 is defined as the difference in the χ^2 resulting from fitting the PV when including or not including the track under consideration. Figure 3.13 shows the resolution of the x component of the IP vector in the plane transverse to the beam direction, IP_x , as a function of $1/p_T$. The projection of

the IP vector is defined as

$$\begin{aligned} \text{IP}_x &= x - x_{\text{PV}} - (z - z_{\text{PV}}) t_x, \\ \text{IP}_y &= y - y_{\text{PV}} - (z - z_{\text{PV}}) t_y, \end{aligned} \quad (3.5)$$

where $(x, y, z)_{\text{PV}}$ is the position of the PV, (x, y, z) is the track's point of closest approach to the PV, and $(t_x, t_y, 1)$ is the direction vector of the track. The IP is related to the x, y projections and the direction vector of the track by

$$\text{IP} = \frac{\sqrt{\text{PV}_x^2 + \text{PV}_y^2}}{\sqrt{1 + t_x^2 + t_y^2}}. \quad (3.6)$$

3.4 Particle identification

Particle Identification (PID) is a crucial step in reconstructing a collision: it allows selection of the particle species that are relevant for a particular analysis and reconstruction of the selected b or c hadron decays. The main detectors that contribute to PID in LHCb are: the Ring Imaging Cherenkov detectors (RICH), that identify long-lived particles such as pions, kaons and protons; the electromagnetic and hadronic calorimeters (ECAL, HCAL), that measure the energy released by photons, electrons, and hadrons; the muon stations, which identify and measure muons' momenta.

3.4.1 The RICH detectors

There are two RICH detectors, one situated upstream of the TT, and the other immediately downstream of the T1-2-3 stations (see Fig. 3.4). Both rely on the Cherenkov effect to measure a particle's velocity; combined with the momentum information provided by the tracking system, it is possible to determine the particle's mass, thus providing PID information on a given track.

The Cherenkov effect [73] is what happens when a charged particle traverses a dielectric medium faster than the speed of light in that medium; the threshold speed is determined by $v_{\text{light}} = c/n$, where c is the speed of light in vacuum and n is the refractive index of the medium under consideration, with $n > 1$. A charged particle with a velocity larger than c/n will emit light in a cone around its direction of motion, with an opening angle known as Cherenkov angle, θ_C . The Cherenkov angle can be written as a function of the particle's velocity $\vec{v}$ and the refractive index of the medium,

$$\cos \theta_C = \frac{c}{|\vec{v}|n}. \quad (3.7)$$

If the photons emitted are detected on a sensitive area perpendicular to the particle's direction of flight, then, in the limit of the active medium being a thin surface, they will form a ring. The Cherenkov angle θ_C can be determined if one measures the cone emission point and the radius of the projected ring. However, since the medium is actually distributed over

a three-dimensional volume, the Cherenkov light can be emitted from multiple points along the particle's trajectory, leading to concentric rings being formed. This makes detection of high multiplicity collisions more difficult. With the use of spherical mirrors, the light can be focussed back into a ring, and then redirected with flat mirrors to the sensitive area for detection. The photons are then ray traced back through the optical system, and θ_C is determined assuming the emission point is midway through the track trajectory in the medium. As the true emission point is randomly distributed along the particle trajectory, this assumption results in a smearing effect on the reconstructed angle. The optical systems of both RICH detectors have been optimised, with the aid of simulation, to reduce this smearing effect to an acceptable level. Figure 3.14 shows a scatter plot of the reconstructed Cherenkov angle as a function of the particle's momentum, as measured by RICH 1, using isolated tracks (tracks where no overlapping ring was found).

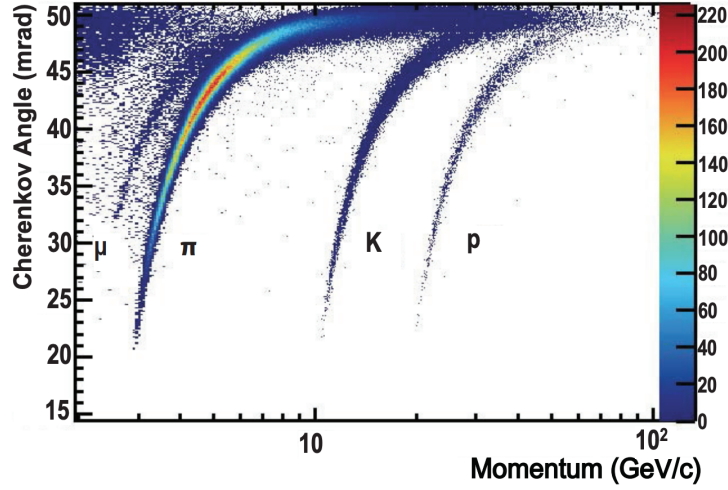


Figure 3.14: Measured Cherenkov angle for isolated tracks as a function of the track's momentum, as reconstructed by the RICH 1 detector. The muon, pion, kaon, and proton bands are clearly visible. Figure taken from Ref. [74].

The two RICH detectors are designed differently, and with different active mediums, to cover a wide range of track momenta and angles. Figure 3.15 shows a technical sketch of the two detectors. RICH 1 covers particles with low momentum, in the range 1–60 GeV/ c , using C₄F₁₀ as radiator, whilst RICH 2 covers the higher momentum range from 15 GeV/ c up to 100 GeV/ c and beyond, using a CF₄ radiator. RICH 1 has a wide angular coverage, covering the full LHCb geometrical acceptance from ± 25 mrad to ± 300 mrad (horizontal) and ± 250 mrad (vertical); RICH 2 covers a more limited angular range, from ± 15 mrad to ± 120 mrad (horizontal) and ± 100 mrad (vertical), as higher momentum particles are boosted in the forward direction. Only the active media and spherical mirrors are within the acceptance, which reduces the overall material budget of the detectors. The Cherenkov light, after being focussed and redirected out of the detector's acceptance by the optical system, is detected using hybrid photon detectors. These use photocathodes, which, when hit by an

incident photon, can produce photoelectrons. The photoelectrons are accelerated in a strong electric field onto a silicon pixel detector, where they are readout. The pixel detectors provide information about the incident angle of the photon, which allows to trace its trajectory through the detector's optical system. As the photon detectors need to be shielded from the field of LHCb's dipole magnet, they are encased in iron shield boxes. Figure 3.16 shows an example of a typical collision at LHCb as reconstructed by RICH 1.

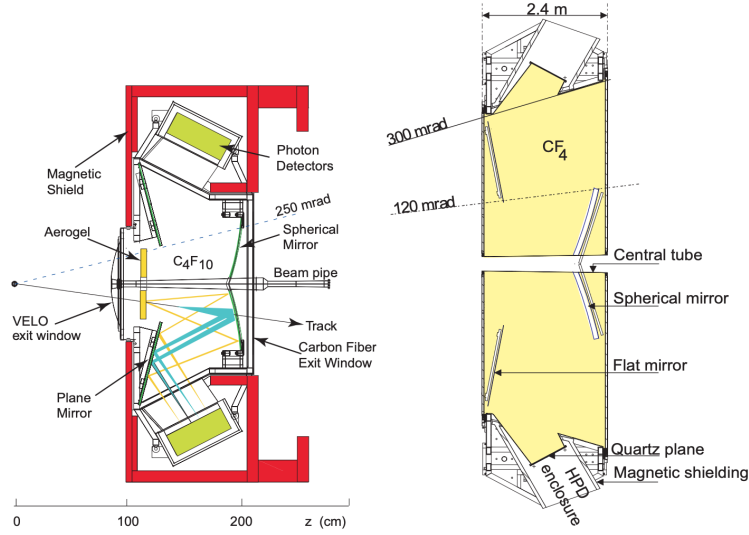


Figure 3.15: On the left, a side view of RICH 1. On the right, a top-down view of RICH 2. The aerogel used in RICH 1 was removed after the first run of data taking. Figures taken from Ref. [63].

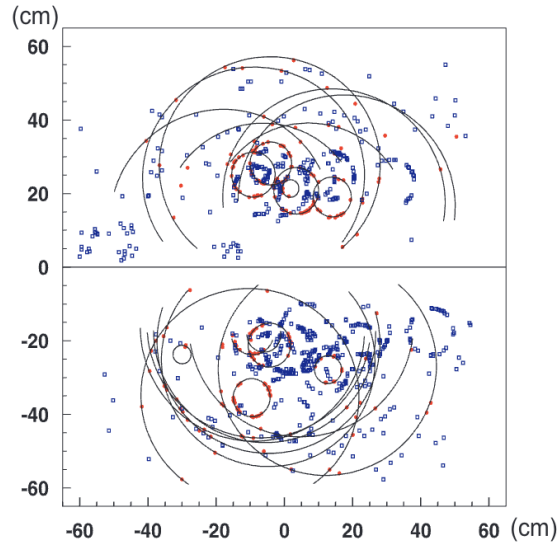


Figure 3.16: A display of a typical collision at LHCb as reconstructed by RICH 1. Figure taken from Ref. [63].

3.4.2 The calorimeters

The calorimeters are used to measure the energies of photons, electrons, and hadrons, and also provide PID information on the tracks that traverse them. The logic of the system is to induce incident particles to shower, *i.e.* to produce a cascade of secondary particles from the interactions with the detector's passive material. The cascade evolves longitudinally, and ends when the showering secondary particles no longer have enough energy to propagate the shower further, and are fully absorbed in the detector material. The passive material, also called the absorber, is made with high Z elements, like lead, increasing the probability of interaction; it is interleaved with a lighter active material which can detect the traversing particles and provide a measurement of the shower direction and deposited energy. By fully containing a shower, the calorimeter provides a measurement of the initial energy of the incident particle.

The LHCb calorimetry system is situated between RICH 2 and the muon system, and it is composed of four detectors: the electromagnetic calorimeter (ECAL), which is preceded by the scintillator pad detector (SPD) and the preshower detector (PS); the hadronic calorimeter (HCAL). The SPD and PS are positioned right before the ECAL to provide vetoes on high E_T neutral and charged pions. Figure 3.17 shows a layout of the complete system.

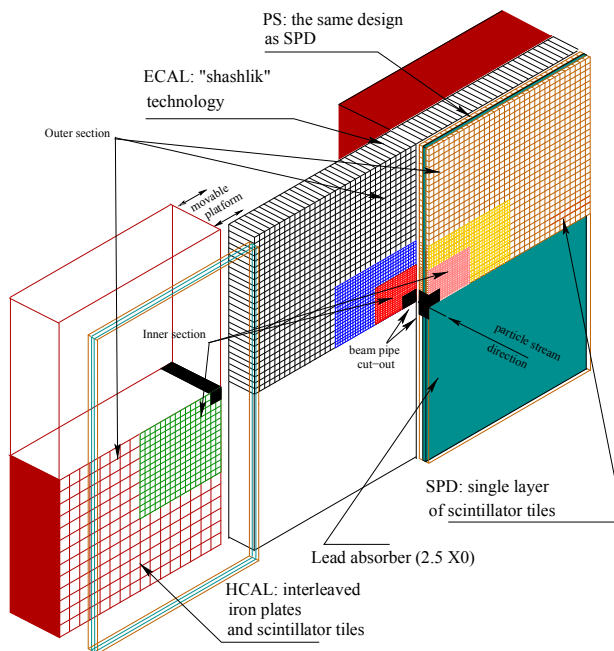


Figure 3.17: Layout of the LHCb ECAL, HCAL, PS, and SPD systems. Figure taken from Ref. [75].

The SPD/PS system is composed of a lead converter, used to induce a pre-shower, sandwiched between two planes of scintillator pads with high granularity. The segmentation of the active planes is made to have a one-to-one projective match with the ECAL cells. By

positioning the SPD layer upstream of the first absorber, it is possible to distinguish showers initiated by charged and neutral particles, since only the former can deposit energy in the SPD. The PS layer can help to discriminate electrons against other charged hadrons (most commonly pions), by comparing the distributions of the deposited energy of the pre-shower induced by the lead converter. Figure 3.18 shows the energy distributions of electrons and pions with momenta of $50 \text{ GeV}/c$, as measured by the PS during a test beam at CERN.

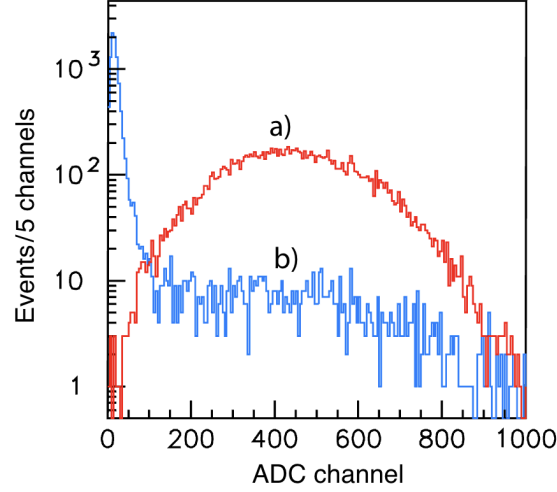


Figure 3.18: Separation of electrons (blue) and pions (red) with momenta of $50 \text{ GeV}/c$, as measured by the PS during a test beam at CERN. The ADC counts can be converted into a measurement of the deposited energy after calibration. Figure taken from Ref. [63].

The ECAL [75] was chosen to be 25 radiation lengths thick, which ensures it can fully contain the electromagnetic showers produced by high energy photons and electrons. It is designed in a “shashlik” layout, each cell consisting of alternating passive and active detector layers. The signal coming from the scintillators in the active layers, which is collected by fibre optic cables embedded longitudinally along the cell, and is then readout at the back of each cell, with photomultiplier tubes. The passive layer is made of lead. Figure 3.19 shows a schematic of an ECAL cell.

The HCAL does not have similar strict requirements on the shower containment, as the information coming from the tracking and RICH systems is enough to measure the energies of charged hadrons; its thickness is thus limited to 5.6 interaction lengths. The design of an HCAL cell is known as “sampling”, where alternating layers of iron absorbers and scintillators are placed parallel to the beamline. Similarly to the ECAL, scintillating light is collected by fibres running along the length of the detector and readout by photomultipliers housed at the back of each cell. Both ECAL and HCAL have segmented layouts, with smaller cells placed closer to the beamline to deal with the higher multiplicity.

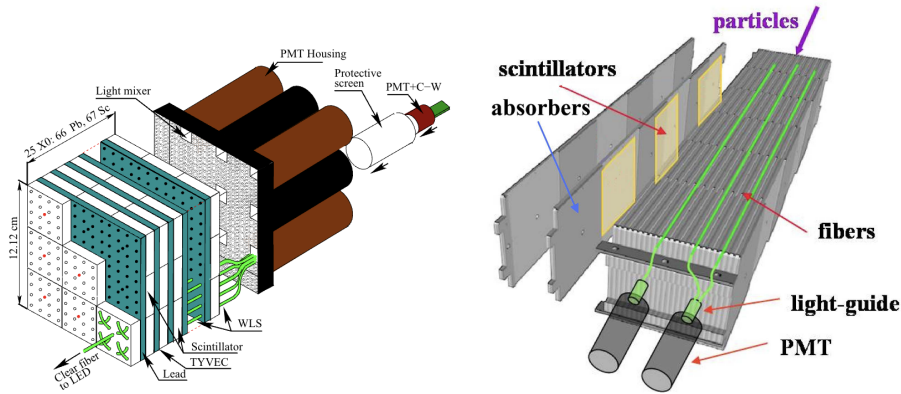


Figure 3.19: Schematics of ECAL (left) and HCAL (right) cells. Figures taken from Refs. [75] and [63].

3.4.3 The muon system

The muon system is a fundamental part of the LHCb detector, both for its triggering capabilities, and for its offline muon identification. It is composed of five individual stations, M1 to M5: the first one is positioned between the RICH 2 and ECAL detector, whilst the other four are situated downstream of the calorimeters, over 15 meters away from the interaction point. Taking advantage of the fact that muons are highly penetrating particles, the four stations M2–M5 are interleaved with passive iron absorbers; including the absorbers which are part of the calorimeter system, this amounts to a total absorber thickness after RICH 2 of about 20 interaction lengths. This allows to shield the muon system from long-lived hadrons (namely kaons, protons, and pions), making it highly efficient in identifying muons. The minimum momentum for a muon to traverse all stations is $6 \text{ GeV}/c$. A side view of the five stations and their relative positions in the detector is shown in Fig. 3.20. The purpose of the first station, being placed upstream of the calorimetry system, is to improve the momentum determination for high p_T muons detected by the hardware level trigger, which will be further described in Sec. 3.5.

All muon stations use multi-wire proportional chambers containing a gas mixture of $\text{Ar-CO}_2\text{-CF}_4$ (in a 40–55–5 ratio), except for the innermost region of the first station in which triple gas-electron multipliers have been deployed; the latter can deal with the higher multiplicity and are able to withstand greater radiation damage. The stations' geometry is projective, meaning that their dimensions scale with the increasing distance from the IP, and each station is divided into four regions which surround the beamline, R1 to R4. These regions have increasing distance from the beam axis, and their segmentation in chambers scales according to a 1:2:4:8 ratio, which helps in having on average the same particle flux and chamber occupancy over the area of a given station. The chambers are arranged in a staggered manner, to allow for overlap between the sensitive areas of neighbouring chambers, thus avoiding dead spots in the detector acceptance. Figure 3.21 shows the segmentation scheme of a muon station.

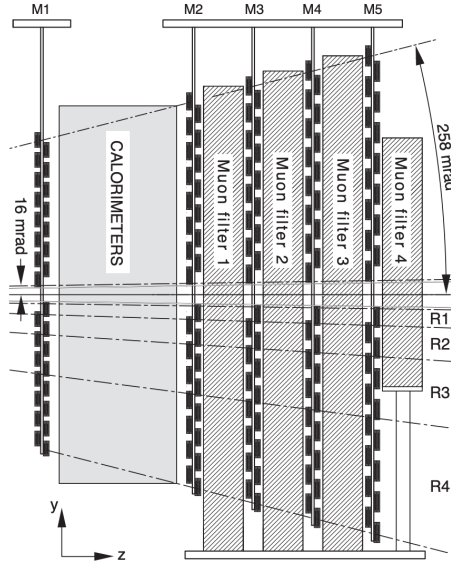


Figure 3.20: Side view of the muon system. Figure taken from Ref. [63].

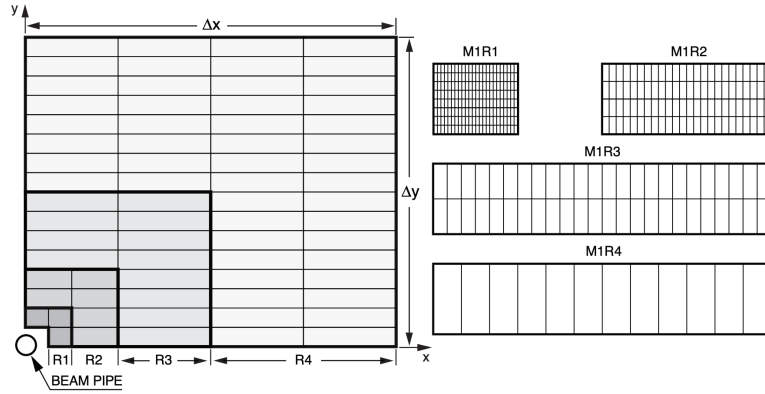


Figure 3.21: (Left) Front view of a quadrant of a muon station, each rectangle representing a chamber. (Right) Granularity of the logical pads in the M1 station, for each of the four R1-R4 regions. In stations M2-M3 (M4-M5) the number of pad columns per chamber is double (half) that of the corresponding region in M1; the number of pad rows per chamber is the same. Figure taken from Ref. [63].

The muon chambers have a hybrid readout system, each front end-chip reading either the anode wires or the cathode pads. The wire or cathodes are also referred to as physical pads. Figure 3.22 shows an example of the hybrid readout in a M2R1 chamber. The wires provide a measurement in the x dimension, while the pads provide a measurement in the perpendicular direction. Depending on the position in the detector, multiple wires can be read out together by the same front end; this happens for example in the outer R4 regions where the need for high spatial resolution is less stringent. Logical pads are then defined by the OR of adjacent horizontal wires and vertical pads. Conversely, in the R1-R2 regions of the second and third stations, where the particle multiplicity imposes a high granularity, the single wire and pad channels are read out, and a logical pad is defined by the logical AND

of the two. The logical pads in each station provide binary information on the hits left by a traversing muon.

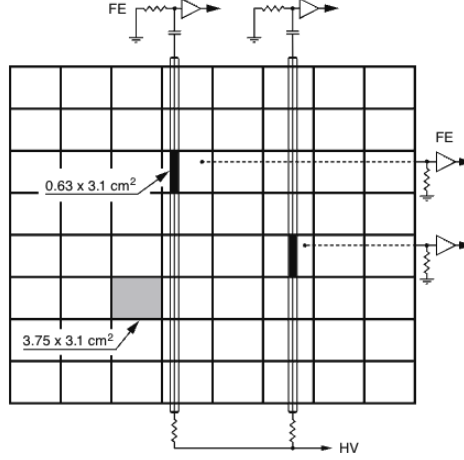


Figure 3.22: Readout scheme in a chamber of the station M2R1. Two anode wire and cathode pad readouts are shown, and their signal is taken in coincidence to make a logical AND (shaded black). Figure taken from Ref. [63].

3.4.4 PID performance

Charged particle identification is achieved in LHCb by relying on likelihood functions $\mathcal{L}$ [68], which take into account the signatures left by a track in the RICH, calorimeter, and muon systems. For a given track, likelihoods are computed assuming different mass hypotheses, corresponding to each possible particle species. At first each track is assumed to be a pion; then the other mass hypotheses, electron, muon, kaon, and proton, are sequentially tested and the corresponding likelihoods are computed. A delta-log-likelihood (DLL) test is used to compare the alternatives to the pion hypothesis:

$$\text{DLL}_{x\pi} = \log \mathcal{L}(x) - \log \mathcal{L}(\pi), \quad (3.8)$$

where x can be a proton, kaon, muon, or electron. The DLL variables are constructed by summing linearly DLL values coming from the individual sub-detectors. These PID variables can then be used to select a given particle species, by requiring only candidates with a large enough value of the corresponding DLL.

A second set of PID variables is obtained using multivariate techniques [76], combining the information from the three PID sub-detectors, as well as information from the trackers. This ensures additional correlations between hits across multiple sub-detectors are taken into account, which is not done in the DLL variables. The multivariate PID variables are referred to as ProbNN_x , where again x indicates the particle species hypothesis under consideration. They are normalised to have an output which ranges from zero to unity. In this work, both PID variable sets are used to select the $B_s^0 \rightarrow \phi \ell^+ \ell^-$ final states. In Chapter 4, the PID

selection is outlined, and its signal efficiency and background rejection power is studied in Chap. 5.

The PID performance of LHCb can be studied using high purity control samples [76, 77]. These samples are collected with minimal requirements on the final state, as to avoid any bias on the PID performance. This can be achieved by choosing certain decay channels which have very clear kinematic signatures, by means of a tag-and-probe procedure. For example, high purity kaons are selected through the decay chain $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$. The soft pion from the $D^{*+} \rightarrow D^0\pi^+$ decay, as well as the hard pion from the subsequent D^0 decay, can be used as tags, while the additional kaon is only required to be a well reconstructed long track. The decay is further identified by the combined invariant mass of the kaon and hard pion, which is required to be close to the known D^0 mass, and by the mass difference of the D^{*+} and D^0 candidates, which for signal peaks around $145 \text{ MeV}/c^2$. Similarly, the muon and electron PID performances can be studied by selecting tag-and-probe samples of $J/\psi \rightarrow \mu^+\mu^-$ and $J/\psi \rightarrow e^+e^-$ decays, respectively. One lepton is used to tag the decay, along with the combined invariant mass.

It is then possible to use the obtained high purity samples to gauge the performance of a given PID selection. As an example, Fig. 3.23 shows the efficiency of kaon identification and pion to kaon misidentification probability, for different requirements on the $\text{DLL}_{K\pi}$ variable, as computed with this method on 2011 data or simulation.

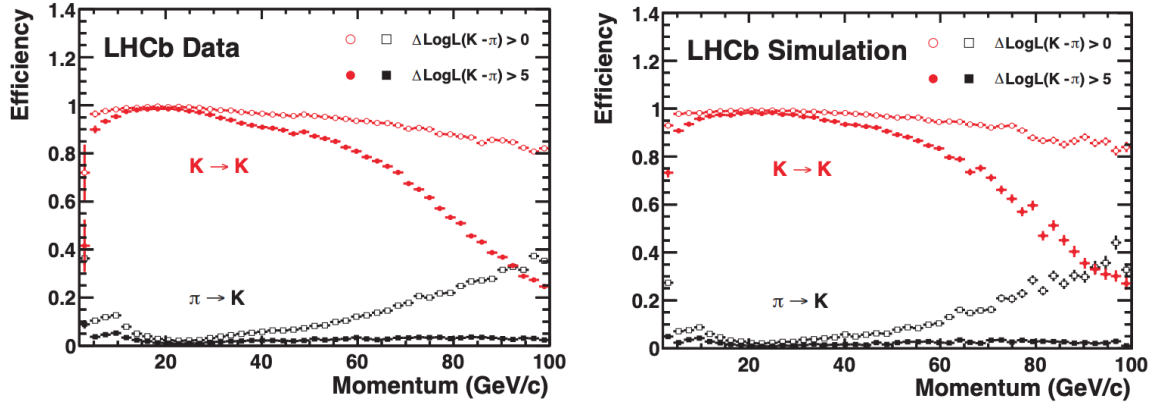


Figure 3.23: Kaon identification efficiency (red) and pion misidentification probability (black) as measured with 2011 data (left) or with simulation (right) of kaons selected from the decay mode $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$, as functions of the particle's momentum. Two requirements on the PID variable $\text{DLL}_{K\pi}$ are applied (open and filled markers) on the same sample to show to variation in efficiency and misid probability. Figure taken from Ref. [68].

Muons can be additionally identified by a third PID variable, called IsMuon [78]. This is a binary variable which is computed for a given track on the basis of its momentum and number of associated hits in the muon detector. For low momentum particles, IsMuon returns a positive output if the track has associated hits in the M2 and M3 stations; for higher momentum, the tracks are required to have hits also in M4 and M5. Because of the shielding

offered by the calorimeters and the passive iron absorbers interleaved with each station, the IsMuon variable achieves excellent PID performance, with an average efficiency above 90% and very low misidentification probability. Figure 3.24 shows the measured IsMuon performance using Run 1 data.

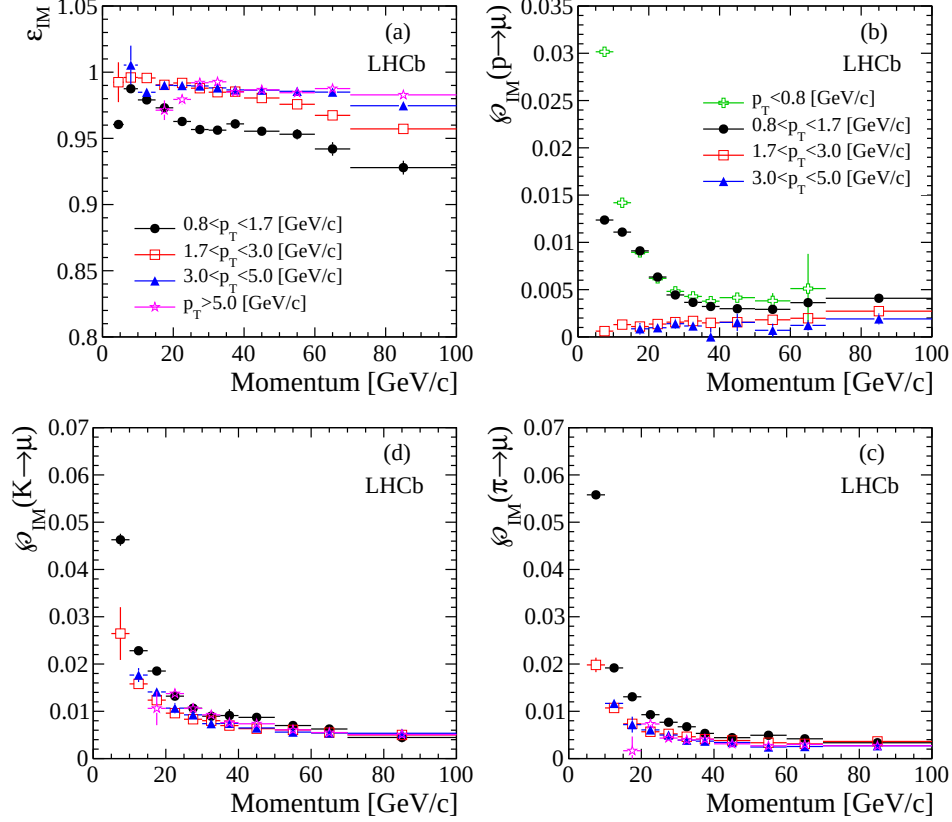


Figure 3.24: IsMuon efficiency (upper left) and proton (upper right), kaon (lower left), and pion (lower right) misidentification probabilities, as a function of momentum, in ranges of transverse momentum. The performance is computed using calibration samples collected in the first run of LHCb data-taking. Figure taken from Ref. [78].

3.5 The LHCb trigger system and data flow

The trigger [79] is the system which decides if the signal left by the products of each of the pp collisions, hereafter referred to as “event”, is to be recorded on physical storage or not. As hinted at the beginning of this chapter, it is impossible to store all the data coming from the LHCb detectors at the rate of collision up to 40 MHz. It is therefore fundamental to have a system in place which, between one collision and the next, determines if the event is relevant for the physics interests of LHCb. An overview of the scheme of the trigger system is shown in Fig. 3.25. The pipeline, which effectively reduces the event rate from 40 MHz down to a much more manageable rate of 12.5 kHz, is composed of three sequential steps.

- **The Level 0 trigger**, or L0 for short, is the hardware based trigger: it operates at

the collision rate, and reduces the event rate down to 1 MHz.

- **The High Level Trigger 1**, or HLT1, the first level of the software-based trigger: it applies an inclusive filter of events which brings the event rate down to 150 kHz.
- **The High Level Trigger 2**, or HLT2, the second level of the software-based trigger: it combines both inclusive and exclusive selections for b and c hadrons, and further reduces the event rate down to 12.5 kHz.

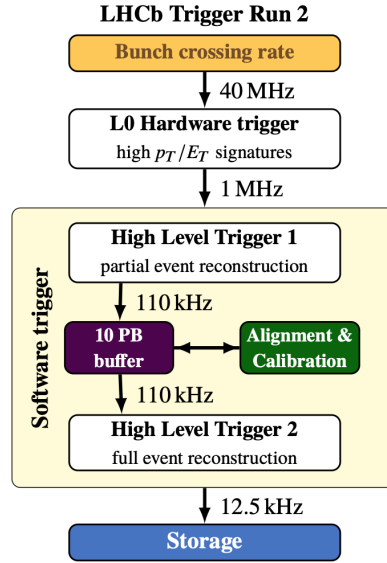


Figure 3.25: The trigger scheme that was implemented in Run 2 of LHCb. Figure taken from Ref. [79].

The L0 trigger takes information directly at the sub-detector level to be able to operate at the bunch-crossing rate of up to 40 MHz. It is made with custom electronics, based on field-programmable gate arrays, that interface directly with the calorimetry and muon systems. The selection of L0 is loose, and based on two trigger lines, originating from the calorimeter and muon systems: L0CALO and L0MUON. These two lines determine a rough estimate of the transverse energy E_T and of the transverse momentum p_T of the particles that traversed the calorimeters and the muon detector. This is because generally b and c hadrons, due to their high masses, decay into products which have high E_T and p_T relative to background of soft particles coming directly from the pp collision.

The L0CALO trigger line has its decision based on three trigger sub-lines, L0Hadron, L0Photon and L0Electron, which are the combined in a logical OR. L0Hadron measures the maximum transverse energy deposited in 2x2 cells of the HCAL, and sums it up with the transverse energy deposited in the ECAL if found. The lines L0Photon and L0Electron base their decision on the energy deposited in the ECAL, and the hits in the SPD and PD: as explained earlier, the presence of hits in the SPD is used to distinguish between neutral and charged signatures.

The L0MUON trigger line looks for straight line segments in the five muon stations. Assuming that the muon candidate originated from the interaction point, and that its trajectory was bent only because of the field of the dipole magnet, the hits lined up in a segment are used to estimate the its transverse momentum, with a relative resolution of around 25%, when averaged over the relevant p_T range. The trigger decision selects the two muon candidates with the highest p_T ; the L0Muon sub-line the one with the largest p_T to be above certain threshold, while the L0DiMuon sub-line requires the product of the p_T of the two muons to be above threshold.

In all trigger lines, apart from the high- p_T L0Muon one, requirements on the maximum number of SPD hits recorded are also added, to select only events with lower overall multiplicity. Compared to a strategy which only uses the E_T and p_T information of the event, this does not result in a significant difference in the overall signal efficiency, but it retains events which are faster to reconstruct in HLT. Table 3.1 reports the different values of thresholds used in the 2015–2018 data-taking period.

Table 3.1: The L0 thresholds for the various trigger lines, for each year of Run 2 data-taking. In 2018 the same thresholds were applied as in 2017. The threshold of the L0DiMuon line is applied to the product of the highest and second highest transverse momenta muon candidates. Table from Ref. [79].

L0 trigger	E_T/p_T threshold			SPD threshold (number of hits)
	2015	2016	2017	
Hadron	> 3.6 GeV	> 3.7 GeV	> 3.46 GeV	< 450
Photon	> 2.7 GeV	> 2.78 GeV	> 2.47 GeV	< 450
Electron	> 2.7 GeV	> 2.4 GeV	> 2.11 GeV	< 450
Muon	> 2.8 GeV/ c	> 1.8 GeV/ c	> 1.35 GeV/ c	< 450
Muon (high p_T)	> 6.0 GeV/ c	> 6.0 GeV/ c	> 6.0 GeV/ c	–
DiMuon	> 1.69 GeV ² / c^2	> 2.25 GeV ² / c^2	> 1.69 GeV ² / c^2	< 900

The output of the L0 trigger selection is transferred to the Event Filter Farm, which can apply more sophisticated algorithms. The first level of the software trigger, HLT1, operates a partial reconstruction, and looks for specific topological signatures, applying an inclusive selection, and then writes the output on a 10 PB buffer. This buffer allows for processing and re-calibration of the detector in-between two LHC fills; it is also large enough to serve as a safety mechanism to delay the HLT2 processing in case of detector or calibration problems. The output from HLT1 is processed by HLT2 which operates the full offline reconstruction, using a combination of around 500 trigger lines, both inclusive and exclusive. Given the reduced input event rate of 150 kHz, HLT2 is able to fully reconstruct tracks and give the most precise estimates of particle momenta, run the most advanced neutral clustering algorithms and the full particle identification. The HLT1 and HLT2 algorithms run in an application called MOORE, based on the GAUDI framework [80].

The outputs of the software triggers are further processed before they can be used for a physics analysis. The software tools used for this are also based on the GAUDI framework.

The reconstruction is processed in the BRUNEL package, which converts the raw data coming from the detector hits into clusters and tracks. The output of BRUNEL is already suitable for data analysis, however a common procedure is to run additional selections, known as stripping, which are tailored for the specific need of an analysis. The stripping is run by the DAVINCI application. The final output is in the ROOT [81] ntuple format, and is ready to be analysed.

The simulation of the events produced at LHCb runs in parallel with this data flow, and is managed by the GAUSS package [82]. GAUSS runs three steps: the generation of events in simulated pp collisions (using the PYTHIA package [83]); the decays of the generated particles (using the EVTGEN package [84]); and finally, the propagation of the final state particles through the detector (with the GEANT4 package [85]). The output of the last step is then digitised by the BOOLE tool, which transforms the simulated output into a mimic of the electronic readout of the real detector, and runs an emulated version of the L0 trigger. The simulated events then enter the nominal data flow, being first processed by MOORE, and following the same processing steps as real LHCb data. The LHCb data and simulation flow is visualised in Fig. 3.26.

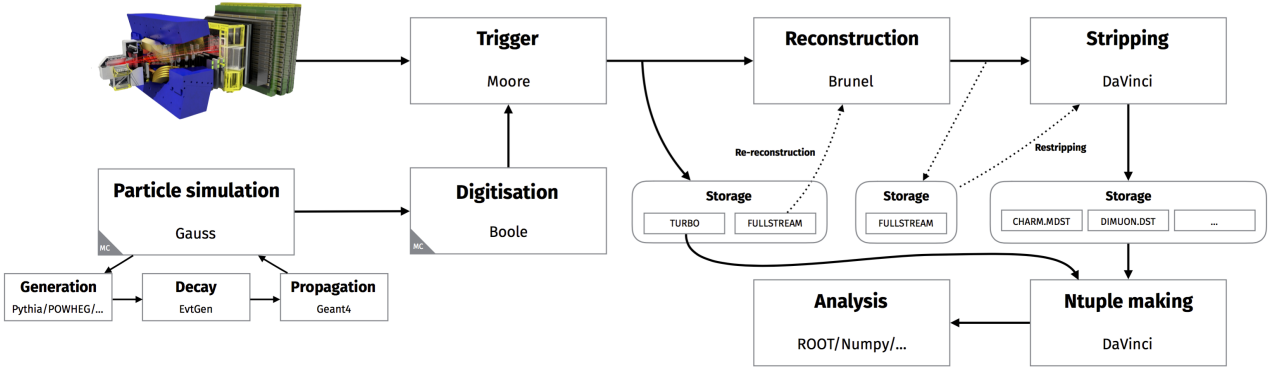


Figure 3.26: The LHCb data and simulation flow in Run2, with individual steps and software shown. Figure taken from Ref. [86].

3.6 The upgraded LHCb detector

In the previous sections the LHCb detector is described as it was during the 2011–2018 data-taking period. During the last LHC Long Shutdown period (2018–2022), the LHCb detector went through a major overhaul, upgrading more than half of its sub-detector components.

The ultimate goal of the LHCb experiment is to make high-precision measurements of the beauty and charm sector of the SM, in the search for BSM effects. In order to be competitive with the precision achieved in theoretical predictions, a large dataset is needed [87, 88]. A dataset corresponding to 50 fb^{-1} of integrated luminosity, more than five times the amount of the combined Run 1 and 2 datasets, would bring the uncertainty on many key flavour observables close to or smaller than the size of the theoretical uncertainty on their predictions. The original LHCb experiment, as it ran in the 2011–2018 period, is limited by the instan-

taneous luminosity $\mathcal{L}$ at which it operates. The upgraded LHCb detector is thus designed to operate at an instantaneous luminosity five times larger than that of Run 2, in order to be able to collect the desired 50 fb^{-1} of integrated luminosity within 10 years of data-taking, during the third and fourth physics runs of the LHC (2022–2033).

Another limitation in the original LHCb design comes with its readout electronics. The original system is limited to a readout rate of about 1 MHz. This is achieved through the hardware trigger, which reduces the input beam-crossing frequency of up to 40 MHz to the required rate. Its selection criteria, as explained in Sec. 3.5, are based on the transverse momenta and energies of particles, as estimated by the muon and calorimeter systems. The efficiency of the hardware trigger is however not uniform across different physics channels: in particular fully hadronic final states are the most affected by these selections. Running at higher $\mathcal{L}$ requires increasing the thresholds in L0 to keep the event rate at 1 MHz, which causes increasingly large inefficiencies in the achievable yields of many interesting physics channels. This effect is visible in Fig. 3.27, which shows the yields of key decays, relative to those observed in Run 1, as a function of the instantaneous luminosity. Any gain in fully hadronic final states from running at higher $\mathcal{L}$ quickly plateaus, showing the clear need for a change in the trigger strategy for such channels. For these reasons, the hardware trigger is removed in the upgraded LHCb detector, in favour of a fully software-based trigger, capable of making more targeted selections. The totality of all detectors' electronic readout systems are upgraded to be able to deal with the higher instantaneous luminosity and to interface with the new software trigger.

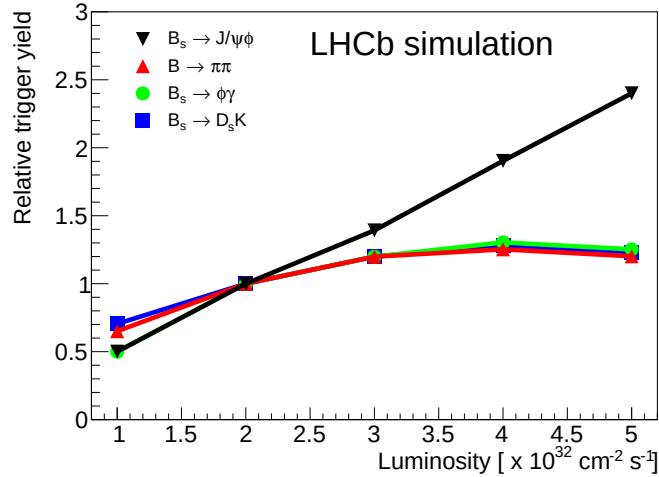


Figure 3.27: Relative trigger yields of key physics channels as a function of instantaneous luminosity. The yields are normalised to those observed at the instantaneous luminosity of Run 1 data-taking, $\mathcal{L} = 2 \cdot 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$. The only channel which does not plateau at higher $\mathcal{L}$ is the $B_s^0 \rightarrow J/\psi \phi$ decay, which benefits from the high efficiency of the L0 muon triggers. The other channels, which trigger on the charged hadrons or neutral signatures in the final state, are limited by the high thresholds of the ECAL and HCAL. Figure taken from Ref. [89].

The following sections will overview the various systems which have been upgraded in

the LHCb detector. They are mainly based on Ref. [89], which offers a much more detailed account of the changes that happened between 2018 and 2022, but goes however beyond the scope of this thesis. Figure 3.28 shows the layout of the upgraded LHCb detector, as it has been operating since 2022. The detector features a completely revamped tracking system, made of a new VELO and two new trackers upstream and downstream of the magnet: the Upstream Tracker (UT), and the scintillating fiber (SciFi) tracker. The RICH detectors are also upgraded and their photomultiplier sensors completely replaced. The SPD, PS, as well as the first station of the muon detector have been removed due to their more limited role in the new trigger compared to the original L0 trigger.

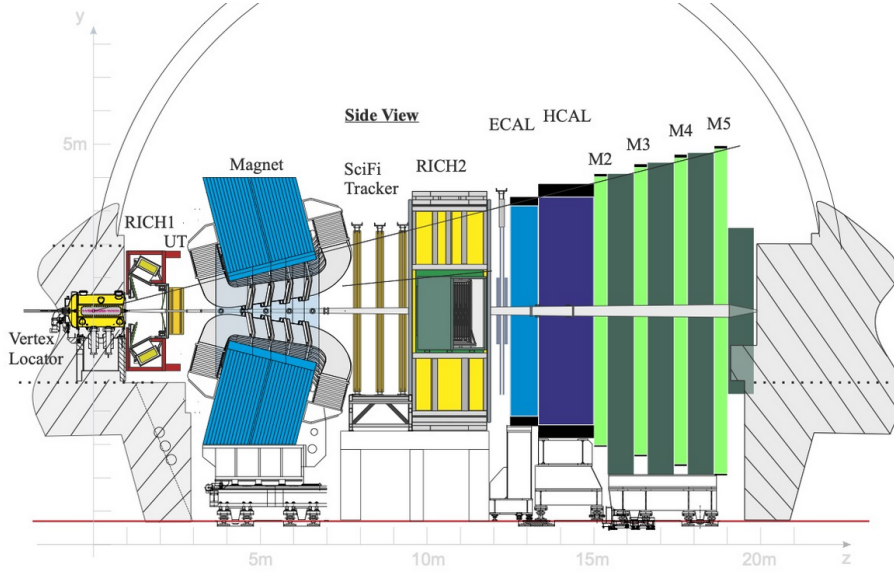


Figure 3.28: Layout of the upgraded LHCb detector and its subcomponents in the $x = 0$ plane. Figure taken from Ref. [89].

3.6.1 The upgraded Vertex Locator

The upgraded VELO is designed to achieve a performance which is as at least as good as the original, in terms of IP resolution and track finding efficiency, within the harsher environment brought by the higher instantaneous luminosity. The silicon strip detectors of the original VELO are replaced with pixelated hybrid silicon detectors, arranged into modules and cooled down by CO_2 passing through microchannels embedded in the module substrate. The primary vacuum vessel and motion system are kept as they were, however the RF foil was replaced with a thinner version, matching the geometry of the pixel modules, to reduce the overall material budget and to allow for the upgraded VELO to be closer to the beamline. Now, in its closed position, the VELO modules are only 5.1 mm away from the centre of the interaction region. The thickness of the RF boxes at the interface with the beamline is reduced to $150 \mu\text{m}$.⁴

⁴Unfortunately due to an incident with the vacuum systems in early 2023, the upgraded RF foil suffered irreparable damage; a replacement was installed at the following end-of-year technical stop of the LHC, albeit

The number of modules is increased to 52, arranged symmetrically around the expected collision point on each side, similarly to the original design. Each module is made up of four silicon pixel tiles, half of which are mounted on one side of the module, and the other half on the other side. The sensors are laid out so that there is a slight overlap between them on the two sides and there are no dead spots. Figure 3.29 shows the layout of the 52 modules, as well as the layout of the tiles on a pair of modules, in the closed position.

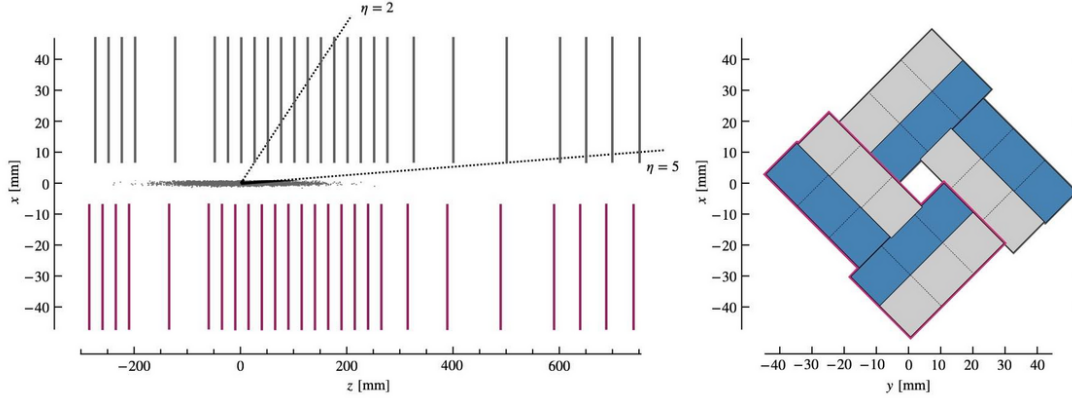


Figure 3.29: Left: cross-section in the $y = 0$ plane of the layout of the 52 upgraded VELO modules, showing also the interaction region and the nominal pseudorapidity acceptance $2 < \eta < 5$. Right: layout of the twelve VELO ASICs on a pair of modules (a single module highlighted by the red contour, note the rotated x and y axes) on each side of the interaction region, in the VELO closed position. In grey, the ASICs which are on the upstream side; in blue, the ASICs on the downstream side. Figures taken from Ref. [89].

The analog signal from each sensor on a tile is read by three custom Application-Specific Integrated Circuits (ASICs), which process the signal and digitise it. Each ASIC is a matrix made of 256×256 square pixels with side $55 \mu\text{m}$. They are capable of withstanding the extremely high flux of particles that hit at a bunch-crossing frequency of up to 40 MHz and with an average of 5 interactions per bunch crossing, being able to handle up to 900 million hits/s. A significant amount of heat is generated in the operation of the ASICs. Active cooling is achieved by means of silicon microchannels, which use evaporating liquid CO_2 to conduct heat away from each module. The liquid/gas is continuously circulated inside the microchannels, passing under the detector elements, and keeping the ASICs at a temperature of around -30°C . The microchannel substrate is at the core of the module, so that tiles on each side can be simultaneously cooled down. Figure 3.30 shows an illustration of the microchannels on one module, as well as an X-ray image of the channels on an actual module before installation of the tiles. Finally, Fig. 3.31 shows one full assembled side of the upgraded VELO, before installation into the vacuum vessel.

one with a thicker aluminium body, impacting the achievable IP resolution.

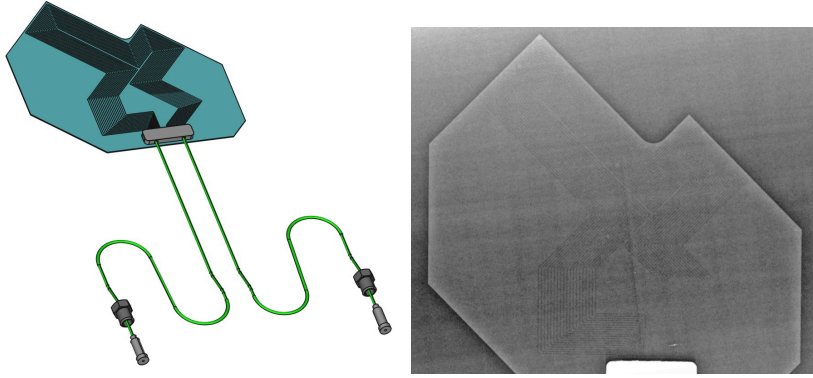


Figure 3.30: Left: illustration of the path of the microchannels through one VELO module, and their connector. Right: X-ray image of a set of microchannels as etched into one of the VELO modules. Figures taken from Ref. [89].

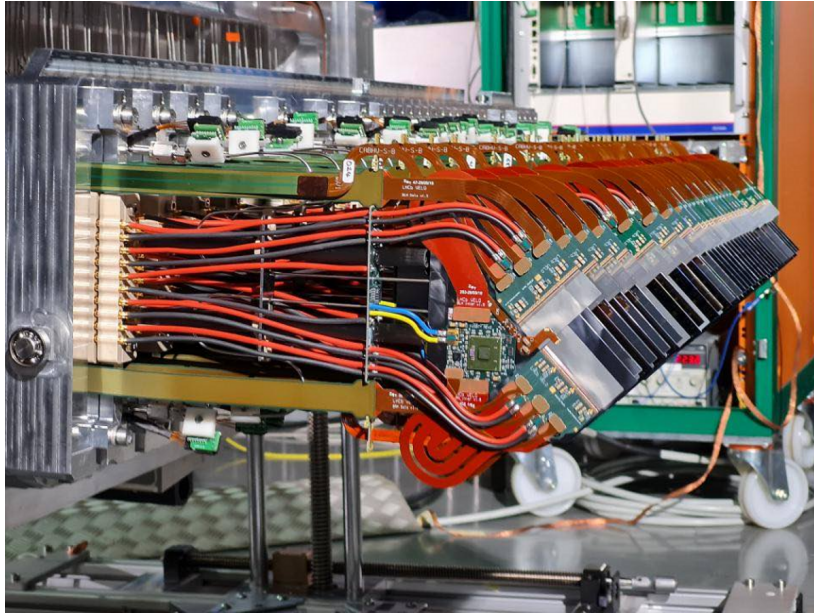


Figure 3.31: Picture of one upgraded VELO side, corresponding to 26 modules, before installation into the vacuum vessel. Figure taken from Ref. [89].

3.6.2 Operation of the upgraded Vertex Locator

The operation of the VELO happens at two distinct levels. During the detector commissioning period, it is operated locally, *i.e.* independently of the global operation of LHCb, by detector experts. The bulk of the commissioning of the new detector has been performed during the first part of Run 3, between 2022 and 2023. Additionally, after each winter LHC shutdown, the VELO goes through a recommissioning period, to ensure that all of its components are calibrated and performing optimally for data-taking. Global commissioning is also performed to ensure the consistency of VELO data with the other LHCb subsystems, such as its time and spatial alignment. Once the detector is fully commissioned, and global data-taking begins, the main operation of the VELO is handed over to the central shift crew.

The data-taking of LHCb is overseen by two shifters, who are supported by a team of on-call experts, or piquets, who provide technical support for each sub-detector. This separation of layers of operation of VELO is achieved through a partitioning of the data acquisition (DAQ) resources [90], which allows them to be allocated independently to each sub-detector, and to run them in parallel to those allocated to the other sub-systems. Partitioning is key for commissioning, but is also a useful tool for swift debugging during global data-taking.

The VELO piquet is the intermediary figure in this operating structure. The piquet offers day-to-day support to both the central shift crew and to the detector experts, ensuring that the VELO is safely operated, and that both the detector-specific plans as well as the global LHCb plans are achieved. The piquet also monitors and logs the daily performance and behaviour of the VELO and all of its sub-systems.

The VELO is operated through the Experiment Control System (ECS), which is an LHCb-wide system in charge of the control and monitoring of all LHCb components. The ECS framework is set as a hierarchical and tree-like structure, which represents the structure of the detector and each of its components. To each hardware component a corresponding device unit is assigned, which tracks its state and has direct access to its readings. Control units are defined to control multiple devices at once (for example, to control all components of a single VELO module). The interface to the ECS contains a series of panels which show the current state of a given set of device or control units as well as the options to send commands to each of them. The most fundamental panel is the Run Control, which accesses the main VELO components: the DAQ system and its infrastructure, the high voltage controls, and the detector control system (which monitors voltages, temperatures, pressures and so on). The VELO's Run Control panel is shown in Fig. 3.32. The piquet can operate the Run Control locally to collect data needed to commission the detector, such as noise runs and time-alignment scans. When operated globally, the VELO is controlled by the central LHCb Run Control panel.

3.6.3 Tracking

The upgraded LHCb tracking system is made up of two sub-detectors, the UT and SciFi, together with the VELO, and the dipole magnet. The magnet has been kept the same as it was in Run 1 and 2 of data-taking. The UT is located between RICH 1 and the magnet, and is made of four layers of silicon microstrip sensors, organised in two stations. The four layers are shown in Fig. 3.33 (left). Its geometry is like that of the original TT, with its $(x-u-v-x)$ configuration.

The SciFi tracker is responsible for the tracking of particles downstream of the magnet. It is designed to achieve a momentum resolution and track efficiency comparable to that of the IT and OT in Run 1 and 2, while operating at higher occupancy. There are twelve detection layers, organised in three stations. Within each station, the layers are set according to the same $(x-u-v-x)$ configuration as the UT. Each layer is made of vertical modules which contain the scintillating fibers. Figure 3.33 (right) shows the layout of a SciFi station.



Figure 3.32: The upgraded VELO's Run Control panel. Each sub-system represents a control unit controlling a major part of the VELO detector, which can be accessed from this top panel. The state of the different control units can be seen as the coloured blocks. The panel also contains options to configure a local data-taking run.

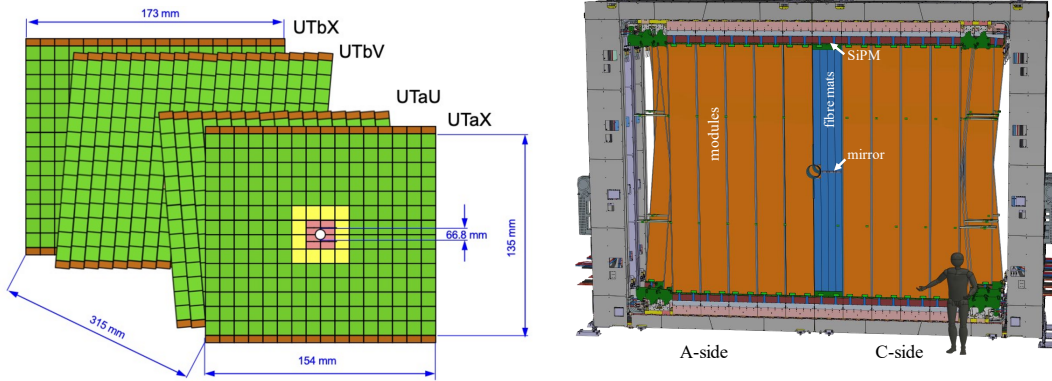


Figure 3.33: On the left, layout of the four layers of a UT station, arranged in the $(x-u-v-x)$ configuration. In green, yellow, and pink are indicated the different types of sensors needed to obtain a finer segmentation closer to the beamline. On the right, front view of a SciFi detector station. The front view shows the division in vertical modules and detailed layout of the scintillating fibers and the photomultiplier readout (SiPM). Figures taken from Ref. [89].

3.6.4 Particle identification

While the overall concept of the RICH detectors remains unchanged with respect to their original designs, important changes have been made to make sure that they are able to cope with the busier environment at higher $\mathcal{L}$. The hybrid photon detectors used in Run 1 and 2 had front-end electronics which were limited to a 1 MHz output rate. They are replaced by multi-anode photomultiplier tubes, equipped with new electronics to allow for a higher output rate. The readout channel density is optimised and split into two regions with different granularity. The optical layout and mechanics of RICH 1 are also redesigned, to reduce the occupancy in the central regions of the detectors.

The detector components of the calorimeter system are left unchanged, but its readout electronics are completely upgraded to comply with the high output rate imposed by the new trigger strategy. Ageing effects due to long term exposure to radiation from the LHC operations were assessed prior to the upgrade. The performance of the inner cells of the ECAL, which receive the highest exposure to ionising radiation, is found to be satisfactory up to about 20fb^{-1} of integrated luminosity [91]. For this reason, plans have been made to replace the inner cells in the third LHC long shutdown, which will happen between 2026 and 2029 [92]. In the case of the inner HCAL cells, their performance was found to already be quickly degrading after Run 1. Because of the higher technical challenge of a replacement during an LHC run, they are replaced by tungsten absorber blocks, which help mitigate the low energy particle background in the muon stations M2 and M3. The effect of this replacement on the overall LHCb physics performance is found to be negligible, as the information from HCAL inner cells has little impact in the reconstruction performed by the software trigger.

Similarly, the muon system detector components, its multiwire proportional chambers, are left as they were in the original LHCb. No significant signs of ageing have been observed after nine years of operations [93]. The first station, M1, is removed, along with the PS/SPD systems. Its usefulness in providing a quick momentum estimate of muon candidates at the lowest level of the trigger quickly degrades in an higher occupancy environment, and is not needed with a software trigger. In its place additional neutron shielding is installed to minimise the radiation exposure of the SciFi's silicon photomultipliers. Unlike in the other sub-detectors, the front-end electronics from the original design are already compliant with the upgraded running conditions, and are kept as they are. The off-detector electronics are instead completely replaced, with new dedicated ASICs.

3.6.5 Trigger system

The new software trigger is made of two stages, the High Level Trigger 1 and 2 (HLT1, HLT2). The HLT1 stage performs inclusive selections based on partial reconstruction of the collision event, running at the beam-crossing rate of up to 40 MHz. The most cost effective solution to process the full detector readout at this frequency is to run HLT1 on graphics processing units. Its selections are based on the reconstruction of tracks and of

displaced secondary vertices, associated with decays of b and c hadrons. It is able to achieve percent-level precision on the measured track momenta.

The second stage, HLT2, performs the full offline-quality reconstruction of the event, making use of the real-time alignment and calibration of the detector. This calibration is achieved by delaying the running of HLT2, temporarily storing the output of HLT1 on a 50 PB buffer and using it as input to the calibration. The HLT2 stage includes more than 1000 exclusive lines, each tailored and tuned to select specific topologies or physics channels.

The performance of the new trigger strategy has been validated with data collected by the upgraded LHCb detector in 2024. Figure 3.34 shows a comparison of the HLT1 Trigger on Signal (TOS) electron efficiency as measured in 2024 with the combined L0Electron and HLT1 efficiency as measured in Run 2. Both are determined using $B^0 \rightarrow J/\psi (e^+e^-) K^{*0}$ candidates. The figure also shows a similar comparison but for the L0Hadron and HLT1 trigger efficiency, determined using $B^+ \rightarrow \bar{D}^0(\rightarrow K^+\pi^-)\pi^+$ decays. In both cases, the upgraded trigger system results in a significantly higher efficiency at low b hadron p_T , which corresponds to the kinematic region where most of the events in the LHCb acceptance lie.

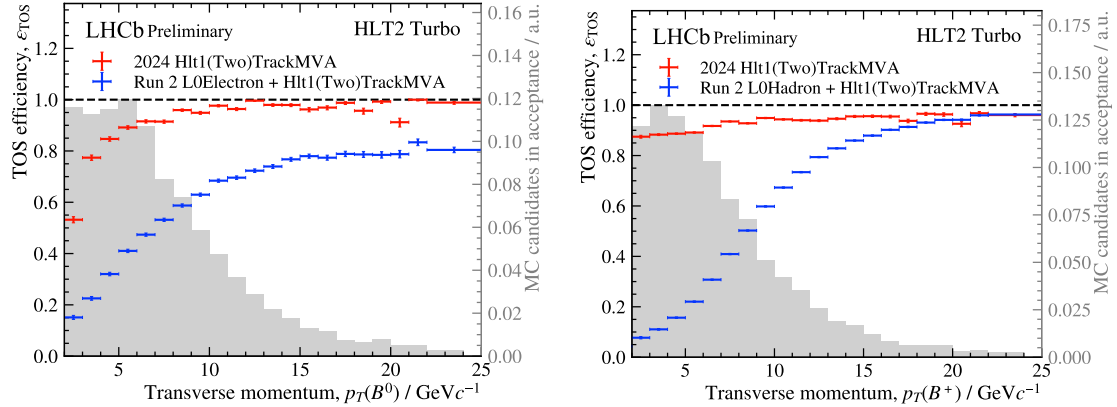


Figure 3.34: Comparison of trigger on signal (TOS) efficiencies between the original LHCb hardware trigger (blue) and the upgraded software-based trigger system (red). On the left, the TOS efficiency for electron final states, measured with $B^0 \rightarrow J/\psi (e^+e^-) K^{*0}$ decays. On the right, the TOS efficiency for fully hadronic final states, determined with $B^+ \rightarrow \bar{D}^0(K^+\pi^-)\pi^+$ decays. In grey, the MC truth-level distributions of the transverse momentum of the b hadrons that lie in the LHCb acceptance. Figures taken from Ref. [94].

3.7 Monitoring the upgraded muon detector's performance

An important aspect of the operation of LHCb is to ensure the highest quality of the data that it collects. For this purpose, a variety of monitoring systems are run, in parallel to data-taking, to actively check the performance of each of its subcomponents. A monitoring system is also key in identifying any issues with the hardware and/or data-taking conditions in real time, and offers information to understand them and perform the necessary fixes.

There are multiple components to the monitoring system for the muon detector. The

ECS provides constant monitoring of many parameters, like temperature and voltages of the different hardware and electronic components. The ECS contains automatic alarms that can be set to trigger whenever a parameter is out of the accepted range for normal data-taking, as well as automatic actions that follow, such as a pause of data-taking. These alarms are picked up by the data-taking crew, who can relay the information to the piquet to deal with the issue at hand.

A second, higher-level monitor is in place to check in real time the quality of the readout from the muon detector. Such a monitor is implemented as a C++ algorithm, within the GAUDI framework, that receives as input the muon detector's raw readout. It makes use of the same decoder used by the reconstruction algorithm, to convert the raw output of the detector into a collection of hits. For each hit, the monitor uses thread-safe counters, known as accumulators, to produce histograms of a plethora of variables. The algorithm runs in parallel to HLT1 on special computing nodes, receiving data from the detector at a rate of about 10 kHz. As this monitor does not have any particular requirement on the type of particle traversing the muon detector, the selected data are chosen by a trigger line with minimum bias, *i.e.* a trigger line which only requires a minimum amount of activity in the detector to select an event.

There are two types of histograms that the monitor provides: histograms meant for the LHCb shift crew, which provide easy-to-digest information about the current status and performance of the detector, and expert-level histograms, which provide more detector-specific information, used by the piquet and the other detector experts. The information of the monitor is handled by the Distributed Information Management (DIM) system [95], and distributed through a presenter. The presenter, MONET [96], is a web-based interface which allows the user to navigate the output histograms of any running LHCb monitor. MONET also preserves a database of the output of the monitors in ROOT format, which can be consulted offline for further data quality checks. The interface is organised in folders containing pages corresponding to a selection of histograms. Each page contains information about the content of the histograms, as well as instructions for the shift crew on what to look out for and check during data-taking. Figure 3.35 shows an example page from a LHCb PbPb run taken in 2024, which shows the (x, y) distribution of hits in the muon stations M2 and M3. Any non-uniformities or empty regions in the distribution can be an indication of a malfunction somewhere in the detector or its readout electronics.

The monitor can also receive as input a special type of events known as Time Alignment Events (TAE). A detector is said to be time-aligned when it is correctly in phase with the LHC clock, *i.e.* its data are correctly read out within the given collision window. This is known as coarse time alignment. A fine time alignment can be achieved by ensuring that all readout channels are synchronous within the same collision window, which increases the overall hit efficiency of a detector. A TAE is typically chosen to trigger on an isolated bunch in the beam, *i.e.* a bunch that is not part of a bunch train but is instead surrounded by empty slots in the filling scheme. The LHC can provide isolated bunches at the request of the experiments. The TAE includes data from the isolated bunch, as well as from the

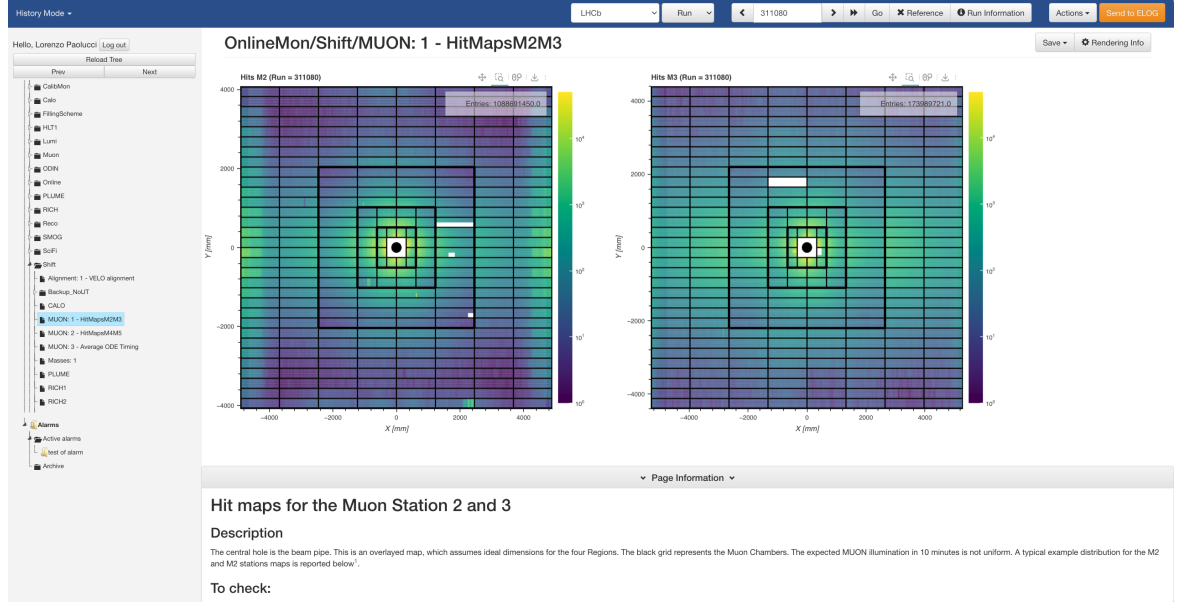


Figure 3.35: Example page of the MONET monitoring presenter, showing the (x, y) hit distribution in the muon stations M2 and M3 for a PbPb run collected by the LHCb experiment in 2024. The page contains a description of the information contained in the histograms, as well as a list of checks for the LHCb shift crew. Small inefficiencies corresponding to empty regions in the hit distributions are most likely due to issues in the detector hardware, but have a small impact on the overall performance.

previous and following three (or more) empty bunch slots.⁵ In this way, the monitor can provide information on the time alignment of the muon detector: if large activity is observed in an empty bunch slot, then that means that the detector, or part of it, is out of phase with the LHC clock. Figure 3.36 shows a typical distribution of TAE events for the muon detector when it is correctly time-aligned.

The final layer of monitoring of the muon detector is a C++ algorithm that makes use of the output of the HLT2 reconstruction to provide information about the detection efficiency of the detector. This information, along with the output of similar monitors for the other sub-detectors, is used to assess the overall quality of the data collected by the LHCb detector before it is prepared for use by analysts. To do so, it makes use of a tag-and-probe trigger line of $J/\psi \rightarrow \mu^+\mu^-$ decays, with the J/ψ signal selected using PID information only for the tag muon. Probe muons are selected to be long tracks with a minimum momentum of $10\text{ GeV}/c$, which ensures that they traverse all stations of the detector. The detection efficiency of a given region of the detector is computed by selecting, out of the collection of probe muons, ones that have associated hits in the other three stations of the detector. For example, to compute the detection efficiency of one of the regions in M4, that means selecting probe muons which have associated hits in M2, M3, and M5. The association of

⁵Since 2024, the LHC stopped providing isolated bunches regularly, because of their impact on the maximum achievable instantaneous luminosity for the ATLAS and CMS experiments. Most of the TAE data are then collected by scanning the empty bunch slots preceding and following a bunch train (known as leading and trailing bunches).

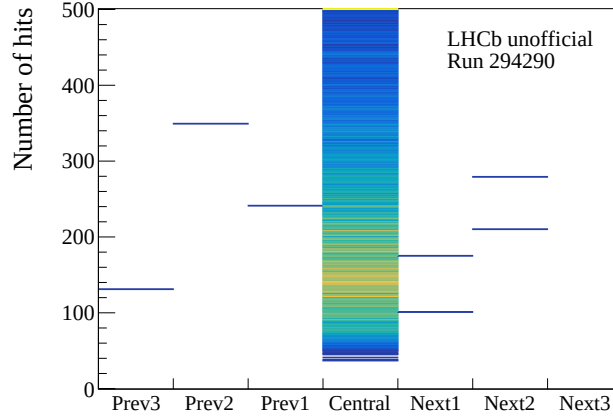


Figure 3.36: Example distribution of hits in the muon detector for TAE data collected in 2024 during a pp fill, as displayed by the muon monitor. Clear activity can be observed in the central bunch-crossing slot, which shows that the detector is correctly time-aligned with the global LHC clock. Very little activity is observed in the empty slots around the isolated colliding bunch pair (indicated as “Prev” or “Next”).

hits is done by computing the distance between the extrapolated trajectory of the probe muon within the detector and the closest hit found in a given station, normalised by the logical pad size. The manual association is needed as HLT2 reconstruction does not use the information of the muon system for particle tracking. Once the subset of probe muons has been selected, the algorithm checks whether they have an associated hit in the region under analysis: if they do, the probes are said to be efficient. Figure 3.37 shows the output of the detection efficiency monitor for the muon detector, for a pp run collected in 2025. The search window used to associate hits to probe muons needs to be finely optimised, to include effects from the edges of the detector acceptance and multiple scattering. If the window is not well tuned, the efficiency computed for a given region can be lower than expected. The region M3R1 is particularly sensitive to the search window definition, and future work will be needed to further optimise it.

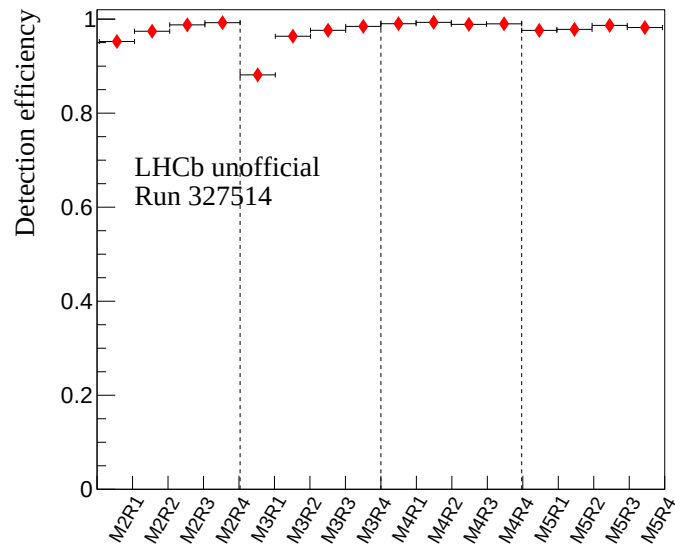


Figure 3.37: Example output of the detection efficiency monitor in all muon regions (R1–4) and stations (M2–5) for a pp run collected by the LHCb experiment in 2025. The statistical uncertainty is too small to be visible.

Chapter 4

Selection of candidates and simulation of $B_s^0 \rightarrow \phi \ell^+ \ell^-$ decays

The author's recommendation

Vespertine

by **Björk**

2001, One Little Indian

In Chap. 3 I introduced the LHCb detector and its components, as well as the data flow which leads from the pp collisions all the way to the datasets which can be used for a physics analysis. The trigger strategy makes the challenge of selecting interesting candidates within these datasets, rather than the proverbial needle in a haystack, about looking for the *right* needle among the many present. With this chapter I will outline the strategies employed to select both $B_s^0 \rightarrow \phi \mu^+ \mu^-$ and $B_s^0 \rightarrow \phi e^+ e^-$ candidates. The selection of the muon final state is based on the experience of previous LHCb analyses of the same channel [36, 97–99]. These include the trigger and stripping lines, meant to select specific signatures of these final states, as well as additional offline requirements, aimed at improving the overall signal purity of the final sample. The signal efficiency of these selections will be discussed Chap. 5. Furthermore, the surviving background components, which will be modelled in the following chapter, will be introduced. As will become clear, the reconstruction of the electron final state brings additional challenges with respect to the muon one. These challenges will be discussed and the dedicated work to address them will be presented. Finally, both measurements presented in this thesis make use of simulation, which is known to be imperfect, to compute selection efficiencies. This chapter will cover the steps needed to improve the agreement of real collision data and simulated datasets. This procedure is based on an iterative correction chain which was first introduced in the context of an earlier measurement of lepton flavour universality from LHCb, using $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ and $B^+ \rightarrow K^+ \ell^+ \ell^-$ decays [29, 30].

4.1 Trigger selection

In order for collision data to be stored by LHCb, they have to have passed the initial selections imposed by its hardware and software trigger system, first introduced in Chap. 3. A trigger line “fires” whenever a part of the event satisfies its specific conditions. The information of which lines fired is attached to all reconstructed particles in the event as part of the data flow, and stored offline in the form of binary flags. The trigger flags are named after each existing line, from L0 to HLT2. Since in a typical event multiple lines are fired by the particles in the event, this additional information is useful in distinguishing processes from one another.

During offline processing, after the reconstruction of the signal candidate, two binary flags are attached, Trigger On Signal (TOS), and Trigger Independent of Signal (TIS) [100]. For any given line, the TOS flag determines whether the signature left by the candidate in the detector was sufficient to fire a positive trigger decision. The TIS flag instead checks whether the “rest” of the event, *i.e.* all tracks which do not make up the signal candidate, was sufficient to fire the trigger line. Note that under this logic, it is possible for an event to be both TIS and TOS, if both the signal and the rest of the event are independently sufficient to fire the trigger line. A third flag, known as Trigger On Both (TOB), is also attached: these are events where neither the signal nor the rest of the event have independently fired the trigger line, but rather the presence of both was necessary to generate a positive decision. These account for a very small fraction of all events, and are not examined further here.

With a sufficiently efficient trigger, one would simply need to require a positive TOS flag on the lines of interest to select signal candidates. This would be the case for the $B_s^0 \rightarrow \phi\mu^+\mu^-$ final state, where the candidate muons leave a signature in the detector that is clear enough for the L0 muon trigger lines to fire with high efficiency. On the contrary, triggering on final states with electrons is more challenging. At the L0 trigger stage the decision is based on the deposited transverse energy E_T in the ECAL; to reduce the overall rate of triggered events to an acceptable level, the E_T threshold has to be quite high. The high threshold rejects a substantial amount of signal candidates, meaning that requiring a TOS flag at the L0 level would result in a low trigger efficiency for the $B_s^0 \rightarrow \phi e^+e^-$ final state. Figure 4.1 shows a comparison of the trigger efficiency of the L0Muon and L0Electron trigger lines.

In order to maximise the electron yield for this work, candidates with a positive TIS flag, are also retained. In particular, two mutually exclusive trigger categories are defined, based on the lepton L0 responses:

- L0I: requires that the B_s^0 candidate is TIS with respect to the global L0 response.
- L0e!: requires that one of the electron candidates fired the L0Electron line (TOS), and that the B_s^0 candidate is *not* L0Global TIS.

For consistency across the datasets, the same strategy is employed for $B_s^0 \rightarrow \phi\mu^+\mu^-$ candidates, by defining an equivalent L0 μ ! category, based on the decision of the L0Muon

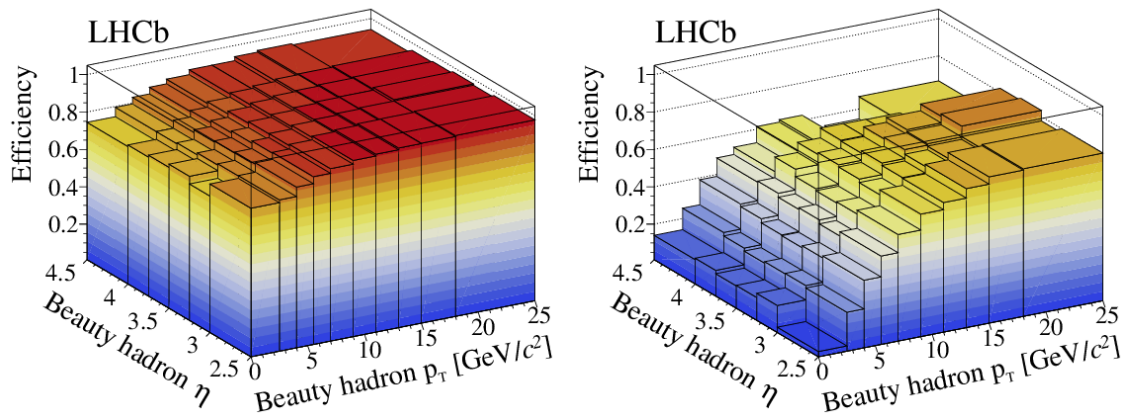


Figure 4.1: Comparison of the selection efficiency of the L0Muon (left) and L0Electron (right) trigger lines, as a function of the b hadron’s pseudorapidity η and transverse momentum p_T . Figure taken from Ref. [79].

line. The label L0L! will be used here to mean both lepton trigger categories. The choice to make the two categories exclusive is to avoid any double counting of candidates, which is particularly important when computing the trigger efficiencies.

At the software trigger stage, the selection of candidates is based on the topology of a typical b hadron decay [101–103]. In particular, at the HLT1 stage, candidates are retained if they have at least one track with high p_T significantly displaced from all PVs. At the next trigger step, HLT2, the tracks selected are combined with either one or two other tracks from the same event, to form a decay vertex. Multiple two- and three-body HLT2 lines are present, which are distinguished by the lepton flavour they trigger on, and whether they require one or two high p_T leptons. Signal candidates are selected to be TOS for the HLT1 single high- p_T track line, and TOS for at least one of the topological HLT2 lines.

4.2 Electron reconstruction

An additional challenge is present in the reconstruction of the dielectron pairs, compared to the dimuon counterpart. Electrons, having a mass much smaller than that of the muon, can radiate a significant amount of energy as they fly through the detector material, in the form of bremsstrahlung photons [73]. Depending on where this energy is radiated along the electron’s trajectory, this can have an important impact on the track reconstruction. In particular, photons emitted before the dipole magnet can change the direction of the flying electron, creating “kinks” in the trajectory, which in turn worsen the quality of the reconstructed track. Furthermore, if enough energy is lost to radiation before the magnet, the electron can be swept out of the downstream acceptance, reducing the overall reconstruction efficiency [104]. On the other hand, the ones emitted in the downstream region have a higher likelihood of being reconstructed in the same calorimeter cluster as the parent electron, since the photons are emitted mostly collinearly to the electron’s flight direction. This therefore has little impact on the reconstructed momentum. A sketch of these effects is shown in

Fig. 4.2.

A large fraction of the bremsstrahlung emission happens due to material interactions in the VELO and TT detectors. A dedicated recovery procedure extrapolates the upstream directions of candidate electrons into the ECAL, to look for compatible neutral clusters. Any cluster found within a fiducial region around the extrapolated direction, and with deposited energy larger than 75 MeV, is associated to the candidate electron and its energy is added back to the reconstructed momentum. The recovery procedure makes a considerable improvement in the electron momentum resolution. However, because of wrong or missed cluster associations, as well as the worse ECAL energy resolution compared to the tracking resolution, the resulting B_s^0 mass resolution is still appreciably worse compared to the case with muons in the final state.

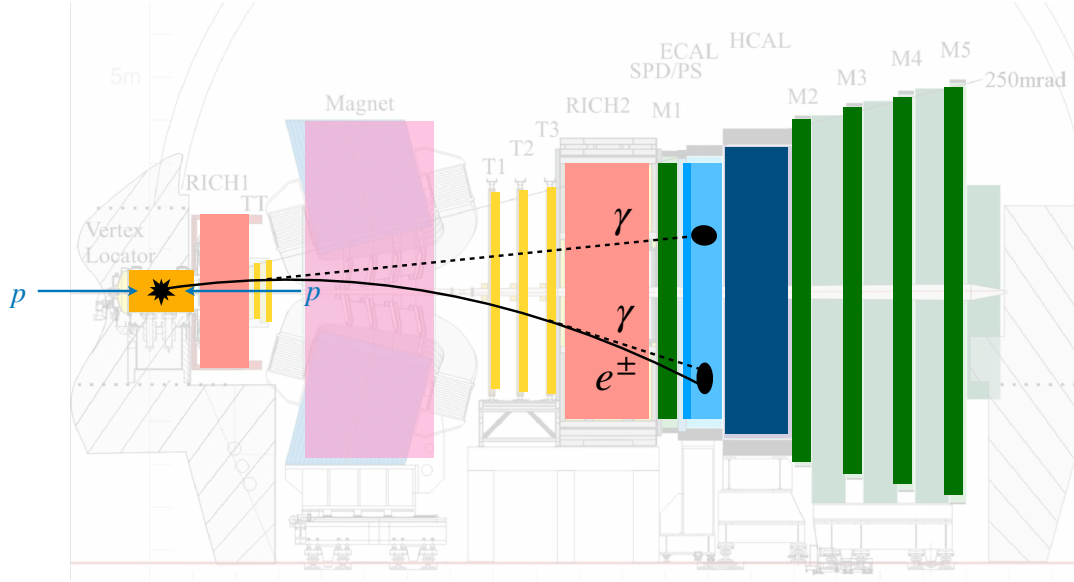


Figure 4.2: Sketch of the bremsstrahlung radiation emitted by an electron in the LHCb detector. The electron is shown emitting a photon both in the upstream region before the dipole magnet, as well as in the downstream region. Photons emitted downstream have a higher chance of being reconstructed in the same cluster as the parent electron, leading to little effect on the reconstruction.

4.3 Offline selection and background rejection

Following the online HLT selection, $B_s^0 \rightarrow \phi \ell^+ \ell^-$ candidates are formed offline during stripping, by combining a pair of two kaons of opposite charge and two leptons of opposite charge. The four tracks are all required to have a large IP and to originate from a common well-reconstructed decay vertex, displaced from all PVs. The B_s^0 candidate is required to have a small IP with respect to one of the PVs, and that its flight direction is well aligned with the displacement vector of the decay vertex from that PV. The combined four-body

invariant masses are required to be in the ranges $5000 < m(K^+K^-\mu^+\mu^-) < 6000$ and $4600 < m(K^+K^-e^+e^-) < 6600 \text{ MeV}/c^2$. The wider range for electrons is to account for the worse mass resolution due to radiative losses. Finally, a global cut on the number of SPD hits of $N_{\text{SPD}} < 600$ (Run 1) and $N_{\text{SPD}} < 450$ (Run 2) is applied, which ensures consistency between the different lepton trigger lines. This improves the control over potential systematic effects which can arise in the computation of the selection efficiencies, which are needed to perform the measurement described in Chap. 5.

After stripping, a series of fiducial requirements, targeting specific characteristics of the $B_s^0 \rightarrow \phi\ell^+\ell^-$ final state, are applied to improve the purity of the selected sample of candidates. The most important of these is to require the combined invariant mass of the kaon pair to be within only $12 \text{ MeV}/c^2$ of the known ϕ mass [17]. Such a tight selection is only possible because of the excellent mass resolution and the narrow natural width of the ϕ resonance. It is highly efficient in retaining signal $B_s^0 \rightarrow \phi\ell^+\ell^-$ decays, whilst largely suppressing backgrounds. In particular, backgrounds where either 1: one of the kaons is actually a misidentified particle of another species; 2: one or more particles have not been reconstructed, or 3: a random particle has been added to form the four-body final state, are very unlikely to yield a dikaon mass so close to the ϕ mass, unless a true ϕ is present. This requirement makes the $B_s^0 \rightarrow \phi\ell^+\ell^-$ decays experimentally cleaner when compared to its non-strange counterparts, $B^+ \rightarrow K^+\ell^+\ell^-$ and $B^0 \rightarrow K^{*0}\ell^+\ell^-$. Contributions from non-resonant $B_s^0 \rightarrow K^+K^-\ell^+\ell^-$ decays in the allowed $m(K^+K^-)$ range are expected to be small [36, 105] and are neglected.

In an effort to further suppress backgrounds with one or more misidentified particles, candidates are further selected by their PID responses. Though the $m(K^+K^-)$ selection is already quite stringent, the kaon candidates are required to fulfill a minimum ProbNN_K . Furthermore, a combined cut on the proton and kaon PID response, ProbNN_p and ProbNN_K , is used to reject the abundant $\Lambda_b^0 \rightarrow pK^-\ell^+\ell^-$ decays.¹ Similarly a minimum value on the ProbNN_μ response is applied to muon candidates on top of the stripping requirement of a positive IsMuon response. In the case of electrons, a combined selection using both $\text{DLL}_{e\pi}$ and ProbNN_e variables is applied. To ensure that the PID response of all final state particles can be well calibrated, there are additional constraints on the tracks. Candidate kaons and muons are required to be within the muon detector’s geometrical acceptance, as well as to have associated hits in the RICH detectors. Similarly, candidate electrons are retained only when having hits in both RICH detectors and the ECAL.

Beauty hadron decays into a J/ψ or $\psi(2S)$ and two charged hadrons can pollute the datasets when a double misidentification occurs, where one of the hadrons is reconstructed as a lepton and vice versa. The hadron-lepton “swaps” include the decays $B^0 \rightarrow J/\psi(\ell^+\ell^-)K^{*0}$ (with $\pi \rightarrow \ell$ and $\ell \rightarrow K$), $\Lambda_b^0 \rightarrow J/\psi pK^-$ (with $p \rightarrow \ell$ and $\ell \rightarrow K$), and $B_s^0 \rightarrow J/\psi\phi$ (with $K \leftrightarrow \ell$), as well as the corresponding decays with a $\psi(2S)$ in the final state. In the case of muons, a veto is applied to the two-body masses $m(K^\pm \rightarrow_{\mu^\pm} \mu^\mp)$, where the subscript indicates that the mass is computed by exchanging the mass hypothesis of the

¹Here implying also resonant $\Lambda_b^0 \rightarrow J/\psi(\ell^+\ell^-)pK^-$ and $\Lambda_b^0 \rightarrow \psi(2S)(\ell^+\ell^-)pK^-$ decays.

kaon candidate to that of a muon. For true swaps, the two-body mass should peak near the J/ψ or $\psi(2S)$ mass [17]; moreover, one also expects one of the kaon candidates to be muon-like and the same-charged muon candidate to be kaon like. Thus candidates with low values of $\text{ProbNN}_\mu(\mu^\pm) \cdot (1 - \text{ProbNN}_\mu(K^\pm))$, as well as a two-body mass $m(K^\pm \rightarrow \mu^\pm \mu^\mp)$ close to the known J/ψ and $\psi(2S)$ masses, are rejected. Figure 4.3 (left) shows the effect of the veto on simulated $B_s^0 \rightarrow \phi \mu^+ \mu^-$ and $B_s^0 \rightarrow J/\psi (\mu^+ \mu^-) \phi$ decays.

The electron veto is based on the four-body masses $m_{J/\psi, \psi(2S)}(K^\pm \rightarrow e^\pm K^\mp e^\pm \rightarrow K^\pm e^\mp)$,² constraining the swapped kaon-electron pair to originate from a J/ψ or $\psi(2S)$ [106]. The constraint, indicated with the subscript $m_{J/\psi, \psi(2S)}$, is used to improve the four-body mass resolution. Similarly to muons, the veto additionally uses information from a combined PID variable $(1 - \text{ProbNN}_x(e^\pm)) \cdot (1 - \text{ProbNN}_e(K^\pm))$, where $x \in [K, \pi, p]$. Note that for each possible electron-hadron swap there is a corresponding veto. Figure 4.3 (right) shows the effect of one of these vetoes on simulated $B_s^0 \rightarrow \phi e^+ e^-$ and $B^0 \rightarrow J/\psi (e^+ e^-) K^{*0}$ decays. In the case of an electron-hadron swap, one expects the electron candidate to be hadron-like (corresponding to values of $\text{ProbNN}_x(e^\pm)$ close to unity) and the same-charged kaon candidate to be electron-like (corresponding to values of $\text{ProbNN}_e(K^\pm)$ close to unity): thus the background will populate the region of low values of the combined PID variable. Studies on simulated decays show that the vetoes are able to effectively make the pollution from hadron-lepton swaps negligible, while keeping almost the totality of signal candidates.

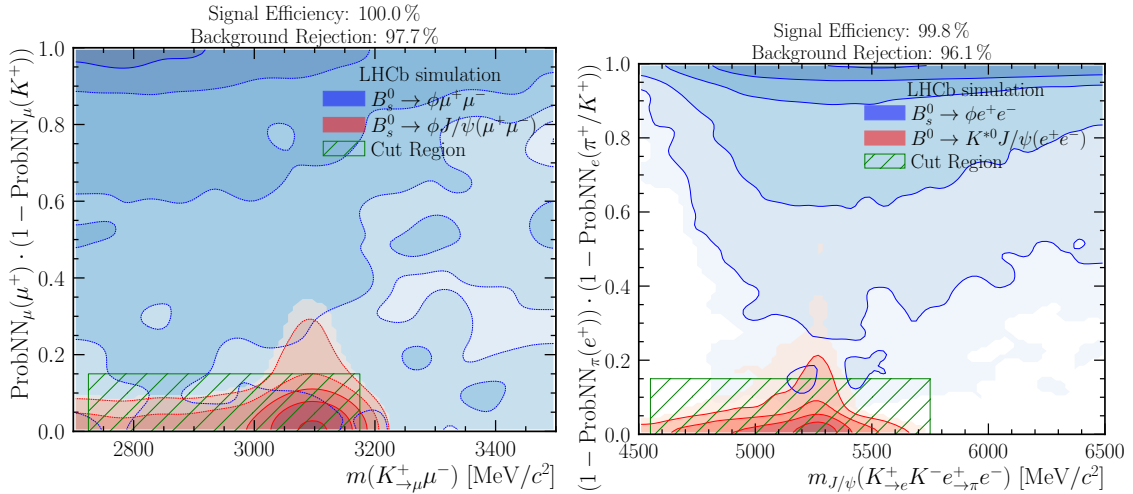


Figure 4.3: Left: Two dimensional distribution of the swap-mass of the kaon-muon system and its combined PID response, for simulated signal $B_s^0 \rightarrow \phi \mu^+ \mu^-$ (blue) and background $B_s^0 \rightarrow J/\psi (\mu^+ \mu^-) \phi$ (red) decays. Right: Two dimensional distribution of the four-body swap-masses and the combined hadron-electron PID response, for simulated signal $B_s^0 \rightarrow \phi e^+ e^-$ (blue) and background $B^0 \rightarrow J/\psi (e^+ e^-) K^{*0}$ (red) decays. The vetoed regions are highlighted.

A final set of mass vetoes is applied to reject misidentified hadronic and semileptonic B_s^0 decays. In particular, the decays³ $B_s^0 \rightarrow D_s^{(*)-} \ell^+ \nu_\ell$ and $B_s^0 \rightarrow D_s^{(*)-} \pi^+$ are highly

²No bremsstrahlung information is used here to avoid potential biasing from wrongly attached photons.

³Where a D_s^{*-} is present its radiative decay $D_s^{*-} \rightarrow D_s^- \gamma$ is implied, with the photon unreconstructed.

abundant [17], and are of particular concern if the charmed meson subsequently decays as $D_s^- \rightarrow K^+ K^- \pi^-$. If a single or double $\pi \rightarrow \ell$ misidentification occurs, for the semileptonic or hadronic decays, respectively, such processes are likely to pass all of the previous selections. Their contribution is completely suppressed by requiring that the three-body masses are not close to the known $D_s^\pm$ mass $m(K^+ K^- \ell_{\text{track}}^\pm) \notin [1920, 2000] \text{ MeV}/c^2$, where the subscript indicates that the mass is computed excluding any correction of the particle's momentum from bremsstrahlung recovery. These vetoes have a high signal efficiency.

4.4 Definition of regions of dilepton mass squared

In this work, five regions are defined based on the square of dilepton invariant mass, referred as $q^2 = m^2(\ell^+ \ell^-)$, and are summarised in Table 4.1. The low-, central-, and high- q^2 regions are relevant for the measurements presented in this thesis, and offer sensitivity to different potential signatures of New Physics. In particular, the low- q^2 region offers the highest sensitivity to the Wilson coefficients $\mathcal{C}_7$ and $\mathcal{C}_{7'}$, the central- q^2 region to the coefficients $\mathcal{C}_9$ and $\mathcal{C}_{9'}$, and the high- q^2 region to the coefficients $\mathcal{C}_{10}$ and $\mathcal{C}_{10'}$. The J/ψ - and $\psi(2S)$ - q^2 regions are where the resonant $B_s^0 \rightarrow J/\psi (\ell^+ \ell^-) \phi$ and $B_s^0 \rightarrow \psi(2S) (\ell^+ \ell^-) \phi$ decays dominate, and they are used to cross-check the analysis strategies of Chapters 5 and 6.

Table 4.1: Definitions of the q^2 -regions for the rare $B_s^0 \rightarrow \phi \ell^+ \ell^-$ decays and the resonant $B_s^0 \rightarrow J/\psi (\ell^+ \ell^-) \phi$ and $B_s^0 \rightarrow \psi(2S) (\ell^+ \ell^-) \phi$ decays. Note that for $B_s^0 \rightarrow J/\psi (\mu^+ \mu^-) \phi$ and $B_s^0 \rightarrow \psi(2S) (\mu^+ \mu^-) \phi$ decays the definition is based on the dimuon mass, which is required to be within $100 \text{ MeV}/c^2$ of the known J/ψ and $\psi(2S)$ mass, respectively [17].

Region	Selection
low- q^2	$q^2 \in [0.1, 1.1] \text{ GeV}^2/c^4$
central- q^2	$q^2 \in [1.1, 6.0] \text{ GeV}^2/c^4$
J/ψ - q^2 , $e^+ e^-$	$q^2 \in [6.0, 11] \text{ GeV}^2/c^4$
$\psi(2S)$ - q^2 , $e^+ e^-$	$q^2 \in [11, 15] \text{ GeV}^2/c^4$
J/ψ - q^2 , $\mu^+ \mu^-$	$ m(\mu^+ \mu^-) - m_{J/\psi}^{\text{PDG}} < 100 \text{ MeV}/c^2$
$\psi(2S)$ - q^2 , $\mu^+ \mu^-$	$ m(\mu^+ \mu^-) - m_{\psi(2S)}^{\text{PDG}} < 100 \text{ MeV}/c^2$
high- q^2	$q^2 \in [15, 19] \text{ GeV}^2/c^4$

4.5 Definition of the blinding regions

The analysis described in Chap. 5 is performed blinded. This is of particular interest for the $B_s^0 \rightarrow \phi e^+ e^-$ mode, which had never been observed prior to this work. To avoid any experimenter's bias, certain parts of the selected datasets are thus not looked at until the analysis strategy is finalised and has passed all cross-checks. Table 4.2 reports the blinding

mass regions for each q^2 region, which are chosen as the smallest possible interval that would include 90% of simulated $B_s^0 \rightarrow \phi \ell^+ \ell^-$ decays.

Table 4.2: Blinded ranges in the different q^2 regions. The ranges are chosen to include roughly 90 % of simulated $B_s^0 \rightarrow \phi \ell^+ \ell^-$ decays. The invariant mass $m(K^+ K^- \ell^+ \ell^-)$ is denoted with $m(B_s^0)$ for simplicity.

$\ell^+ \ell^-$	$\mu^+ \mu^-$	$e^+ e^-$
low- q^2	$m(B_s^0) \in [5340, 5400] \text{ MeV}/c^2$	$m(B_s^0) \in [5100, 5500] \text{ MeV}/c^2$
central- q^2	$m(B_s^0) \in [5330, 5400] \text{ MeV}/c^2$	$m(B_s^0) \in [5100, 5450] \text{ MeV}/c^2$
high- q^2	$m(B_s^0) \in [5330, 5400] \text{ MeV}/c^2$	$m(B_s^0) \in [5150, 5550] \text{ MeV}/c^2$

4.6 Calibration of simulated samples

It is well understood that simulated samples and real collision data show discrepancies for several reconstructed quantities. These differences can have an important impact on key aspects of the work presented in this thesis, such as training the multivariate classifiers, computing selection efficiencies and expected background contamination, and hence risk biasing the final results. For this reason, data-driven corrections for these effects are applied, to better match simulated samples to data. Following the approach originally developed in Refs. [29, 30], weights are computed from ratios of the relevant variables as observed in simulated samples and data, and applied to the simulated samples. The corrections are computed in an iterative procedure, each step building on top of the previous correction, to account for correlations between various effects being corrected. The correction chain is shown in Fig. 4.4, and it covers the simulation of the PID response, tracking, event multiplicity and B candidate kinematics, trigger response, vertex reconstruction, q^2 resolution, and B candidate lifetime.

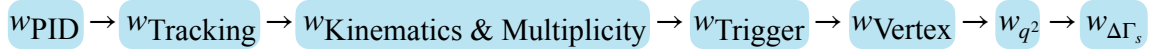


Figure 4.4: Scheme of the iterative correction chain to improve agreement between simulated samples and data. Each step is computed on top of all previous corrections.

4.6.1 Particle identification

Particle identification in LHCb requires input from almost all sub-detectors, and is thus among the more difficult variables to simulate accurately. To correctly estimate the efficiency of a given PID selection and for a given particle species, high purity calibration samples have been collected [107], either by using tag-and-probe methods or by choosing decays with

kinematic properties that allow for unambiguous identification of the children particles. The following samples are available, categorised by particle species:

- kaons and pions from $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ decays;
- protons from $\Lambda \rightarrow p\pi^-$ decays;
- muons from detached $J/\psi \rightarrow \mu^+\mu^-$ decays;
- electrons from $B^+ \rightarrow K^+J/\psi(\rightarrow e^+e^-)$ decays.

The PIDCALIB tool [77] relies on the *sPlot* method [108] to perform background subtraction on the calibration samples. It then computes the efficiency of a given PID requirement as a function of signal track kinematics, specifically the pseudorapidity η and the transverse momentum p_T , and the number of reconstructed tracks in the event N_{tracks} . The efficiency is determined as the ratio of the histogram of events that passed the selection, and the histogram of all considered candidates.

In the case of electrons, potential correlations between the control variable used for the background subtraction and the track kinematics, break one of the assumptions needed for the *sPlot* method to be effective. For this reason, a simpler fit-and-count method is employed, by means of a simultaneous fit to the datasets of events that pass and those that fail the electron PID requirement. Similarly to before, the PID efficiency is determined as a function of the particle's η , p_T , and N_{tracks} .

For muons and kaons, the efficiency determined with PIDCALIB is applied directly to simulated candidates as a per-event weight, effectively replacing the nominal PID cut, $w_{\text{PID}}^{\mu,K} = \epsilon_{\mu,K}(\eta, p_T, N_{\text{tracks}})$. For electrons, the nominal PID cut is applied normally, and the per-event weight is determined as the ratio of the PID efficiency as determined in simulation and in real data, $w_{\text{PID}}^e = (\epsilon_e^{\text{data}}/\epsilon_e^{\text{sim}})(\eta, p_T, N_{\text{tracks}})$. This is needed as it correctly takes into account known correlations in the PID responses of the electron pairs.

4.6.2 Electron track reconstruction

This correction addresses the mismodelling of particle interactions with the detector materials, which affects the computation of the track reconstruction efficiency in simulation. The reconstruction efficiency of electron long tracks can be estimated with $B^+ \rightarrow K^+J/\psi(\rightarrow e^+e^-)$ data, as described in Ref. [104], and used to correct the value estimated in simulation. The correction is a per-track weight of the form

$$w_{\text{TRK}} = \frac{\epsilon(\text{long-track}|\text{VELO})^{\text{data}}}{\epsilon(\text{long-track}|\text{VELO})^{\text{sim}}}, \quad (4.1)$$

where $\epsilon(\text{long-track}|\text{VELO})$ indicates the reconstruction efficiency of long tracks that have been reconstructed in the VELO, for data and simulation, respectively. The efficiency is parametrised as a function of the pseudorapidity η , the transverse momentum p_T , and the azimuthal angle φ . The latter is used to separate regions where the VELO's RF-shielding

is present, as those tracks traverse up to an additional extra 10% of material (expressed in radiation lengths), compared to other tracks.

4.6.3 Beauty meson kinematics and event multiplicity

The standard event generator used for simulation, PYTHIA, is known to have some limitations in its accuracy, in particular in the kinematic spectrum of B_s^0 mesons produced in simulated pp collisions, as well as the overall event multiplicity. Furthermore, there are inconsistencies in the simulation of secondary particles, such as final state radiation (performed by PHOTOS), which is produced in great quantities for electrons. Corrections for these shortcomings can be made using a reweighter with a multivariate approach, following the method outlined in Ref. [109]. The data sample used to compute these corrections is $B_s^0 \rightarrow J/\psi (\mu^+ \mu^-) \phi$ decays selected in the LOL! trigger category, using both collision data and simulation. The *sPlot* method is used to statistically remove the background pollution in data. The correction weights are computed as a ratio of data and simulation, as a function of the momentum, transverse momentum, and pseudorapidity of the B_s^0 candidate, as well as N_{tracks} (which is used as a proxy for the event multiplicity).

4.6.4 Trigger

This section covers the corrections that tackle any mismodelling in the trigger response in simulation, both at the hardware level (L0) and software level (HLT). These corrections are computed using the TISTOS method [100] with $B_s^0 \rightarrow J/\psi (\ell^+ \ell^-) \phi$ data. It is a tag-and-probe method that uses independent trigger decisions to measure the efficiency of the trigger selection itself, as a function of relevant kinematic variables. The TISTOS method can be applied to both data and simulation, and correction weights are then computed as the ratio of the two sets of efficiencies.

At the L0 stage, the per-event weight depends on the trigger categories defined in Sec. 4.1. For the L0I category, the weight is computed as the simple ratio

$$w_{\text{L0I}} = \frac{\epsilon_{\text{L0I}}^{\text{data}}}{\epsilon_{\text{L0I}}^{\text{MC}}}, \quad (4.2)$$

where $\epsilon_{\text{L0I}}^{\text{data}}$ is the data/simulation efficiency for a positive trigger decision. This trigger decision is by definition independent of the signal candidate, however it does show dependency on the B_s^0 candidate's transverse momentum, as well as the number of tracks in the event N_{tracks} . This behaviour can be easily explained when considering that the B_s^0 candidate's transverse momentum originates from the initial pp collision, as does the rest of the event; whatever fired L0 trigger, is likely to be kinematically correlated with the signal candidate. Additionally, as the multiplicity of the event increases, one can naïvely expect the chance of the rest of the event firing the L0 trigger to increase as well. Hence a correlation with a multiplicity proxy such as N_{tracks} can be expected. This correction is independent of the lepton flavour, so it was computed using $B_s^0 \rightarrow J/\psi (\mu^+ \mu^-) \phi$ decays, as they are much cleaner

and with a larger sample size. It was tested that the weights obtained are compatible if recomputed using $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ decays.

Using the efficiency for the L0I category, the corrections for the exclusive L0L! trigger categories are computed as

$$w_{\text{L0L!}} = \frac{\epsilon_{\text{L0L}}^{\text{data}}}{\epsilon_{\text{L0L}}^{\text{MC}}} \cdot \frac{1 - \epsilon_{\text{L0I}}^{\text{data}}}{1 - \epsilon_{\text{L0I}}^{\text{MC}}}. \quad (4.3)$$

For L0Muon the weight is a function of the muon's p_T and η . For L0Electron, the weight is a function of the transverse energy E_T and ECAL region (inner, middle, outer).

The corrections for HLT also employ the TISTOS method, to determine the trigger efficiencies in data and simulation. Just as for the L0 trigger, it is possible to build correction weights from the ratios of such efficiencies. The correction weights are computed independently for the L0I and L0L! categories,

$$w_{\text{L0I,L0L!}}(\text{HLT}) = \frac{\epsilon_{\text{L0I,L0L!}}^{\text{data}}(\text{HLT})}{\epsilon_{\text{L0I,L0L!}}^{\text{sim}}(\text{HLT})}. \quad (4.4)$$

4.6.5 Vertexing

Variables related to the event reconstruction are known to be imperfectly modelled in simulation. Similarly to the correction for the B_s^0 kinematics, a correction can be applied using a reweighter with a multivariate approach, following the method outlined in Ref. [109]. The weights are computed as functions of the χ_{IP}^2 , and the chi-square per degree-of-freedom of the fit to determine the B_s^0 decay vertex. As these variables are known to be correlated with the discriminating mass variable, the *sPlot* method is not used to subtract the background from the data sample of $B_s^0 \rightarrow J/\psi (\ell^+\ell^-) \phi$ decays. Instead, a more stringent selection is applied, by restricting the candidates to a narrow window around the known B_s^0 mass, using the J/ψ mass-constrained variable $|m_{J/\psi}(K^+K^-\ell^+\ell^-) - m(B_s^0)^{\text{PDG}}| < 60 \text{ MeV}/c^2$. This preselection achieves very good background rejection. Since the corrections are observed to depend on the trigger category, they are computed separately for the L0I and L0L! selections.

4.6.6 Dielectron mass

In simulation, the $m(e^+e^-)$ resolution is usually overestimated compared to data, due to misalignment of the material budget in simulation with respect to the real detector. The resolution is strongly dependent on the process of bremsstrahlung recovery. Three categories can be defined, based on the number of attached photons to the dielectron pair:

- Bremsstrahlung category 0 γ : No bremsstrahlung photons are added to either electron.
- Bremsstrahlung category 1 γ : There was exactly one bremsstrahlung photon recovered.
- Bremsstrahlung category 2 γ : Two or more bremsstrahlung photons were recovered.

The corrections are applied to the simulated dielectron mass, m^{sim} , to shift and scale it into a smeared mass, m^{corr} using the following expression

$$m^{\text{corr}} = m^{\text{true}} + \mathfrak{s}_\sigma(m^{\text{sim}} - m^{\text{true}}) + \Delta\mu + (1 - \mathfrak{s}_\sigma)(\mu^{\text{sim}} - m_{J/\psi}^{\text{PDG}}). \quad (4.5)$$

where m^{true} is the generator-level dielectron mass. The scale parameter $\mathfrak{s}_\sigma$ accounts for the difference in detector resolution between data and simulation. The shift parameter $\Delta\mu$ aligns the reconstructed mass peak position between data and simulation, while the $(\mu^{\text{sim}} - m_{J/\psi}^{\text{PDG}})$ term aligns the difference between the reconstructed peak and the “true” value given by the Particle Data Group average of the J/ψ mass $m_{J/\psi}^{\text{PDG}}$ [17]. The correction is computed per candidate and for each bremsstrahlung category.

The $m(e^+e^-)$ distribution of simulated $B_s^0 \rightarrow J/\psi(e^+e^-)\phi$ decays is first fitted to determine all parameters of the lineshape, notably, the mean μ^{sim} and the width σ^{sim} . Simulated events are weighted with all previous correction weights. Then, these parameters are fixed in a following fit to $B_s^0 \rightarrow J/\psi(e^+e^-)\phi$ data, where the mean shift $\Delta\mu$ and width scale $\mathfrak{s}_\sigma$ are left floating. The data candidates are chosen with the nominal selection; to further reduce background, a cut on the J/ψ constrained B_s^0 mass is applied.

4.6.7 Beauty meson lifetime

The EVTGEN model used to generate the signal $B_s^0 \rightarrow \phi\ell^+\ell^-$ simulated samples neglects the lifetime difference between the two B_s^0 meson mass eigenstates, *i.e.* it uses the approximation $\Delta\Gamma_s = 0$. This does not match other simulation samples, notably that for the control channel $B_s^0 \rightarrow J/\psi(\ell^+\ell^-)\phi$, where the lifetime difference is set to be non zero $\Delta\Gamma_s/\Gamma_s = 0.166$.

A correction is built per event, at the generator level, using the five dimensional decay rate $\frac{d\Gamma}{dq^2 d\tau d\cos\theta_K d\cos\theta_\ell d\Phi}$, expressed in terms of the proper decay time τ , q^2 , and the decay angles $\cos\theta_K$, $\cos\theta_\ell$, and Φ as described in Refs. [40, 41]. The weight is computed as the ratio of the decay rate with and without including the effects of a non-zero $\Delta\Gamma_s$,

$$w_{\Delta\Gamma_s} = \frac{\frac{d\Gamma}{dq^2 d\tau d\cos\theta_K d\cos\theta_\ell d\Phi}(\Delta\Gamma_s/\Gamma_s = 0.166)}{\frac{d\Gamma}{dq^2 d\tau d\cos\theta_K d\cos\theta_\ell d\Phi}(\Delta\Gamma_s/\Gamma_s = 0)}. \quad (4.6)$$

4.7 Multivariate selection

The selection strategy described so far has been based on rectangular selections, based on one or two variables. However, some backgrounds require a more complex approach, as one or two variables might not discriminate sufficiently between them and the $B_s^0 \rightarrow \phi\ell^+\ell^-$ signal. This is the case for what is known as combinatorial background. This background arises from the nature of particle reconstruction in the detector. Tracks are reconstructed and matched to a particular particle hypothesis, based on the signature left in the various LHCb sub-detectors, and are then combined to form a signal candidate. In any given collision

event many tracks will be found, so the number of possible four-track combinations with a total invariant mass within the interesting mass window can be very high. This increases the chances of forming fake candidates from tracks which are actually not correlated to one another. The combinatorial background encompasses any such wrong combination.

The process of bremsstrahlung recovery can sometimes combine random hard photons with the electron candidates, which makes the tails of the charmonium resonances leak into the high- q^2 region, as will be further explained in the next section. In the case of $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ decays, such large overreconstruction, needed to get to a q^2 larger than $15 \text{ GeV}^2/c^4$, leads to candidates which populate a region of very high B_s^0 candidate mass, far from the signal peak. However, partially reconstructed $B \rightarrow J/\psi \phi X$ decays, where B is any beauty meson, can end up back in a region of $m(K^+K^-\ell^+\ell^-)$ much closer to the signal peak, after bremsstrahlung recovery. This is in contrast with partially reconstructed $B_s^0 \rightarrow \psi(2S)\phi X$ decays, which, since the $\psi(2S)$ can leak into high- q^2 with softer photons, would fall out of the selected mass window.

For each of these backgrounds the separation with real signal candidates is done with a multivariate approach, by training two classifiers known as boosted decision trees (BDT). A BDT uses single decision trees, which perform one-dimensional selections of the input data based on the given discriminating variables. Two boosting algorithms, known as categorical boosting, or CATBOOST [110,111], and gradient boosting, are used to reduce the effects of statistical fluctuations of the single trees and to improve classification performance. Each classifier is trained on two datasets, which act as signal and background proxies. The training is performed by splitting the input datasets into a training dataset and a testing dataset, in an 80:20 ratio. The BDT, once trained, returns a score for each analysed event, normalised between zero (background-like) and unity (signal-like). The performance of the trained classifier is assessed on the testing dataset by the receiver operating characteristic (ROC) curve, which plots the resulting signal efficiency against the background rejection, as function of the chosen threshold for the BDT score. To flag any possible overfitting or selection bias, the assessment is done with a tenfold cross-validation: the input datasets are split into the training and testing datasets ten times, with each “fold” selecting a unique and equal sized random subset of candidates. A BDT is trained for each fold and its performance assessed with the ROC curve. Once the performance has been optimised and no significant overfitting or selection bias is found, the BDT trained on the first fold is used to select real data candidates based on their output score.

4.7.1 Combinatorial BDT

The input variables used by this BDT are based on the quality of the candidate reconstruction and the decay kinematics, which offer the greatest separation power. In particular, the list includes the quality of the fit to reconstruct the decay chain [112], the χ_{IP}^2 of the B_s^0 , dikaon and dilepton pairs, and the separation between the B_s^0 decay vertex and the PV. The BDT uses also the following kinematic variables: the p_T of the B_s^0 , the dikaon and dilepton pairs,

as well as the angle between the reconstructed B_s^0 flight direction and the vector connecting the PV and decay vertex.

For the signal proxy, simulated $B_s^0 \rightarrow \phi \ell^+ \ell^-$ decays are used, selected with all the criteria described above. As for the background proxy, real collision data is used, selecting candidates which populate the $m(K^+ K^- \ell^+ \ell^-)$ upper mass sidebands, which are known to be dominated by combinatorial background. The upper mass sidebands are defined as $m(K^+ K^- \mu^+ \mu^-) > 5400 \text{ MeV}/c^2$ for muons and $m(K^+ K^- e^+ e^-) > 5600 \text{ MeV}/c^2$ for electrons, with the upper limit defined by the stripping selection. To increase the size of the background training set, the $m(K^+ K^-)$ mass window is relaxed to be $\pm 50 \text{ MeV}/c^2$ within the known ϕ mass. Three unique BDTs are trained for each lepton flavour, corresponding to three data-taking periods: Run 1 (2011–2012), the first half of Run 2 (2015–2016), and the second half of Run 2 (2017–2018). The split follows from the fit strategy which is outlined in the next chapter.

Figure 4.5 shows the ROC curves, averaged over the ten validation folds, for the electron and muon BDTs trained with 2017–2018 data-taking conditions. The BDT achieves excellent background rejection while maintaining high signal efficiency.

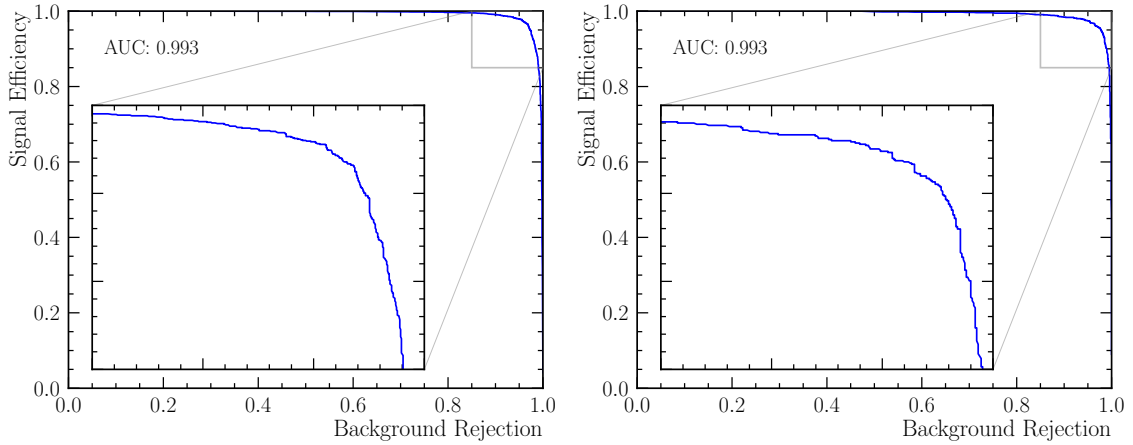


Figure 4.5: ROC curves, averaged over the ten validation folds, for the combinatorial BDT trained for the muon (left) and electron (right) datasets, under 2017–2018 data-taking conditions. The BDT achieves excellent background rejection and high signal efficiency.

4.7.2 High q^2 BDT

The second classifier is chosen to have no overlap in input variables with the combinatorial BDT. The choice of variables is based on ones which are characteristic to the process of bremsstrahlung recovery, such as the uncorrected track momenta of the electron candidates, $p_{\text{track}}^{e\pm}$, the transverse momenta of their associated bremsstrahlung photons, p_{T}^{γ} , as well as the ratio $p^{\gamma}/p_{\text{track}}^{e\pm}$. Other kinematic variables include the total transverse energy deposited in the ECAL, the uncorrected dielectron mass, the pseudorapidity of the electron candidates, and the opening angle between the two electron tracks. Finally, the classifier uses the ratio of the ϕ and dielectron momentum components transverse to the B_s^0 direction. Momentum

conservation would make the ratio unity, with any wrongly attached photons breaking this balance. If approximating real bremsstrahlung photons to be collinear with the dielectron flight direction, the ratio can be used to compute a corrected B_s^0 mass, which is also used as input to the BDT. This alternative method to correct electron momenta for radiative losses was first introduced in Ref. [113]. While the collinearity approximation holds better for lower values of q^2 , the addition of these variables still improves the discriminating power of the classifier.

The high- q^2 BDT is trained on simulated $B_s^0 \rightarrow \phi e^+ e^-$ decays, which represent the signal, and simulated $B_s^0 \rightarrow J/\psi (e^+ e^-) \phi$, $B^+ \rightarrow J/\psi (e^+ e^-) K^+$, and $B^0 \rightarrow J/\psi (e^+ e^-) K^{*0}$ decays as a background proxy. As expected, since it is all related to the reconstruction of bremsstrahlung photons, the three background simulated samples share the same distributions for all input variables. To maximise the size of the background training set, only one BDT is trained for all data-taking periods. Figure 4.6 shows the resulting ROC curve, averaged over the ten validation folds. The BDT achieves good background rejection and signal efficiency.

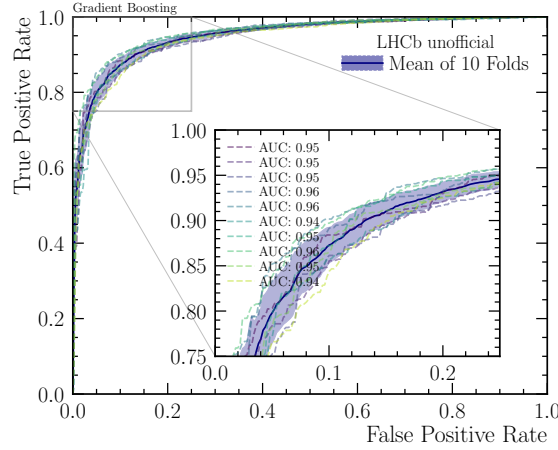


Figure 4.6: ROC curve (shown here as signal efficiency against the false positive rate) for the high- q^2 BDT trained for electron data. The average curve as well as the single curves obtained in the tenfold cross-validation are shown. The BDT achieves good background rejection and signal efficiency.

4.7.3 Selection optimisation

The optimisation of the selection criteria on the BDT scores is performed differently for the muon and electron samples. For the former, the selection is optimised using a figure of merit (FoM), the significance, defined as

$$\text{FoM}(q^2, x) = \frac{S(q^2, x)}{\sqrt{S(q^2, x) + B(q^2, x)}}, \quad (4.7)$$

where S and B are the estimated signal and background yields for a given value x of the BDT threshold, in each q^2 bin. The expected signal yield is computed from the measured

yield in the $B_s^0 \rightarrow J/\psi\phi$ control channel, $N(B_s^0 \rightarrow J/\psi(\mu^+\mu^-)\phi)(x)$, scaled by the ratio of selection efficiencies between the resonant and rare mode $\epsilon_{\mu^+\mu^-}(x)/\epsilon_{J/\psi}(x)$, as well as the known branching ratios of each mode. Thus one can write S as

$$S(q^2, x) = N(B_s^0 \rightarrow J/\psi(\mu^+\mu^-)\phi)(x) \cdot \frac{\mathcal{B}(B_s^0 \rightarrow \phi\mu^+\mu^-)}{\mathcal{B}(B_s^0 \rightarrow J/\psi\phi)\mathcal{B}(J/\psi \rightarrow \ell^+\ell^-)} \cdot \frac{\epsilon_{\ell^+\ell^-}(q^2, x)}{\epsilon_{J/\psi}(x)}. \quad (4.8)$$

As described in the next chapter, the $B_s^0 \rightarrow J/\psi\phi$ yield is obtained from an unbinned maximum likelihood fit to the mass distributions of muon candidates selected in the J/ψ - q^2 region, and the efficiencies are determined from simulation. The background yield is measured by performing a blind unbinned maximum likelihood fit to the muon candidates selected in the low-, central-, and high- q^2 bins. The blind regions are defined in Table 4.2. The optimisation is performed for each q^2 region and in the three data-taking periods, corresponding to each uniquely trained BDT.

In the case of electrons, the same optimisation strategy is used for the combinatorial BDT in the low- and central- q^2 bins. For the high- q^2 bin, a more realistic optimisation is performed using pseudoexperiments. Using pseudoexperiments provides a more accurate estimate of the signal significance. It also allows to consider fit stability in the question of finding the optimal working point. To account for a possible correlation in the responses of the combinatorial and high- q^2 BDT, the optimisation is performed in two dimensions.

At any given BDT selection point ($\text{BDT}_{\text{comb.}}, \text{BDT}_{\text{high-}q^2}$), 500 pseudoexperiments are generated from sideband maximum likelihood fits to the $m(K^+K^-e^+e^-)$ distribution, assuming the expected number $B_s^0 \rightarrow \phi e^+e^-$ candidates as determined using Eq. (4.8).⁴ The shapes to generate signal and background events are described in the next chapter, and are assumed to not change significantly when varying the BDT selection. The fits to the pseudoexperiments are used to extract the FoM as the ratio of the uncertainty on the signal yield $\sigma(N)$ and N .

Small convergence rates indicate issues with the overall stability of the fit strategy, and can potentially result in biases to the estimated FoM. Thus, the optimal point is chosen by maximising the FoM while simultaneously requiring a minimum convergence rate of 80%. Figure 4.7 shows the significance as determined from pseudoexperiments for the 2015–2016 data-taking period, in the two-dimensional plane for the combinatorial and high- q^2 BDT selections. It is found that two different optimal points are needed for the 2011–2016 and 2017–2018 periods. This is done in order to accommodate the smaller fit convergence rates observed for candidates in the 2011–2016 datasets, which are due to the more limited sample size compared to 2017–2018 data. Thus, a looser working point of (0.2, 0.4) is chosen for the 2011–2016 period, and a harder one of (0.6, 0.6) is chosen for 2017–2018. This achieves a convergence rate of 95%; further detail on the pseudoexperiment studies performed to study the overall sensitivity to R_ϕ^{-1} will be given in Sec. 5.4.2.

After the optimisation, some combinatorial background is still expected to contaminate

⁴Since at the time of this work no measurement of the $B_s^0 \rightarrow \phi e^+e^-$ branching ratio was available, the same as the muon mode is assumed.

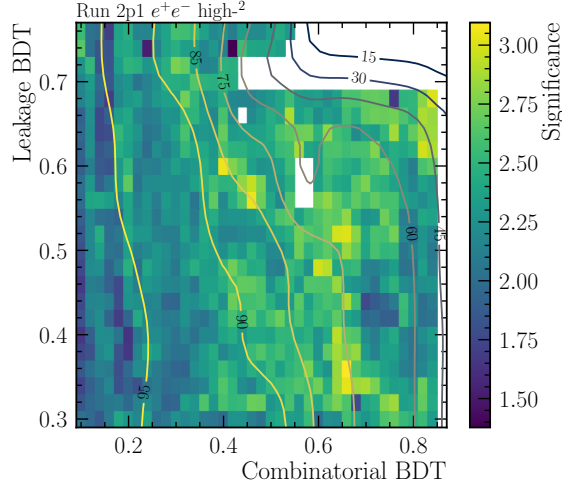


Figure 4.7: Result of the optimisation of the combinatorial and high- q^2 BDT selections for electron data taken in the 2015–2016 period, using pseudoexperiments. The statistical significance is estimated as the ratio of the uncertainty on the fitted signal yield $\sigma(N)$ and N . Curves for points with the same convergence rate are overlayed.

the dataset, hence it is modelled as a component in the fits described in the following chapters.

4.8 Residual background contamination

After the selection strategy described in the previous sections, some background contamination is still leftover. Because no other efficient selection requirements can be applied to remove these backgrounds, they will be modelled as individual components in the maximum likelihood fits described in Chaps. 5 and 6.

4.8.1 Semileptonic double cascades

The mass vetoes described in Sec. 4.3 efficiently remove semileptonic B_s^0 decays where the D_s^- decays hadronically. However, the D_s^- can also decay semileptonically, creating a “double-cascade” $B_s^0 \rightarrow D_s^{(*)-} (\rightarrow K^+ K^- \ell^- \bar{\nu}_\ell(\gamma)) \ell^+ \nu_\ell$. Since both neutrinos are, for all intents and purposes, invisible to LHCb, this creates the same visible final state as signal $B_s^0 \rightarrow \phi \ell^+ \ell^-$ decays. This background contaminates the lower $m(K^+ K^- \ell^+ \ell^-)$ sideband in the low- and central- q^2 bins. For the muon final state, it is possible to restrict the mass fit region enough to completely suppress this contamination. Because of the worse mass resolution, this is not feasible for the electron final state. With simulated $B_s^0 \rightarrow D_s^{(*)-} (\rightarrow K^+ K^- e^- \bar{\nu}_e(\gamma)) e^+ \nu_e$ decays, it is possible to estimate the expected background contamination. In the relevant mass fit ranges, after the full selection, the means of the expected yields are 22.0 ± 2.3 semileptonic double cascade candidates in the low- q^2 bin, and 75.8 ± 6.9 candidates in the central- q^2 bin.

4.8.2 Charmonium leakage

As already mentioned previously, the bremsstrahlung recovery process is not perfect, either because of missed energy or wrongly assigned photons; this, in combination with the limited ECAL resolution, worsens the q^2 resolution for the electron final state, compared to muons. In cases of under- or overreconstruction, the charmonium resonances $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ and $B_s^0 \rightarrow \psi(2S) (e^+e^-) \phi$ can “leak” in the neighbouring q^2 bins. This contamination happens in both directions, towards lower and higher reconstructed q^2 . The effect, for the combination of the 1γ and 2γ bremsstrahlung categories, is shown in Fig. 4.8.

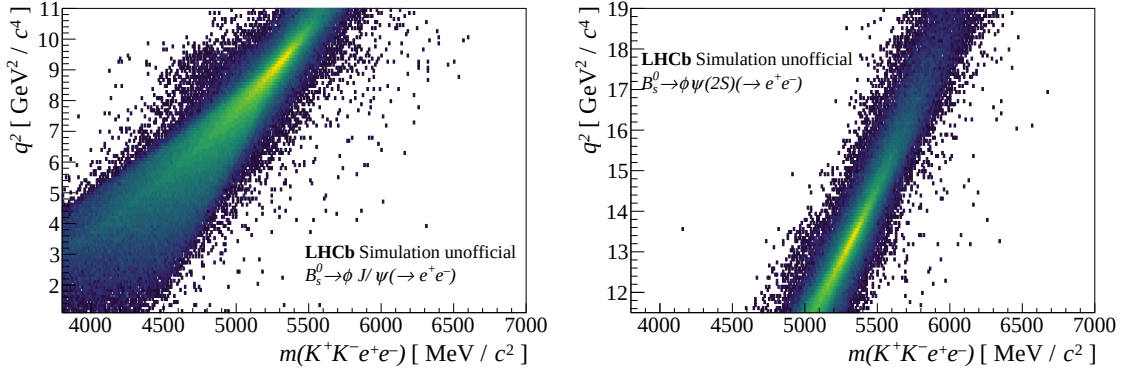


Figure 4.8: Left: two dimensional distribution of reconstructed $m(K^+K^-e^+e^-)$ and q^2 for simulated $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ decays, showing how the radiative tails leak from the J/ψ - q^2 bin ($q^2 \in [6.0, 11]$ GeV^2/c^4) into the central- q^2 bin ($q^2 \in [1.1, 6.0]$ GeV^2/c^4). Right: two dimensional distribution of reconstructed $m(K^+K^-e^+e^-)$ and q^2 for simulated $B_s^0 \rightarrow \psi(2S) (e^+e^-) \phi$ decays, showing how the radiative tails leak from the $\psi(2S)$ - q^2 bin ($q^2 \in [11, 15]$ GeV^2/c^4) into the high- q^2 bin ($q^2 \in [15, 19]$ GeV^2/c^4).

These backgrounds will be referred to as leakage backgrounds. The tail from missed or underreconstructed bremsstrahlung of $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ candidates pollutes the central- q^2 bin. In the high- q^2 bin, wrongly assigned hard photons lead to significant overreconstruction of the candidate’s q^2 , so that both $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ and $B_s^0 \rightarrow \psi(2S) (e^+e^-) \phi$ candidates leak from their respective bins. While the high- q^2 BDT greatly reduces the contribution of $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ and $B \rightarrow J/\psi \phi X$ decays, where B is any beauty meson, at the optimal working point there is still enough contamination that these components need to be taken into account.

With simulation, it is possible to estimate the expected background contamination. For $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ decays one expects 192.8 ± 6.3 candidates in the central- q^2 bin, and 12.7 ± 1.7 candidates in the high- q^2 bin. For $B_s^0 \rightarrow \psi(2S) (e^+e^-) \phi$ decays the expectation is 49.0 ± 4.2 candidates in the high- q^2 bin. Finally, one expects partially reconstructed $B \rightarrow J/\psi \phi X$ decays to contribute to 7.2 ± 1.1 candidates in the high- q^2 bin.

4.8.3 Misidentified backgrounds

While the PID selection described above is quite stringent, backgrounds with one or more particles misidentified as electrons are expected to make up a significant fraction of the selected

dataset. Misidentified decays which are fully reconstructed, such as $B^0 \rightarrow K^+ K^- h^+ h'^-$, where $h^{(\prime)}$ can be either pions or kaons, can create structures which may peak under or close to the signal $B_s^0 \rightarrow \phi e^+ e^-$ peak. Furthermore, misidentified partially reconstructed decays such as (but not limited to) $B^{0,\pm} \rightarrow K^+ K^- h^+ h'^- X$, where X can be any missing charged or neutral particle (most commonly pions), as well as over-reconstructed decays such as $B^+ \rightarrow K^+ K^- K^+$, where an additional electron or misidentified hadron is wrongly attached to the candidate, can also pass the selection and populate a wide region in $m(K^+ K^- e^+ e^-)$, forming more complicated structures. The challenge of simulating an accurate mixture of all possible misidentified decays is nearly impossible, as many of the contributing processes are not well known. Indeed, these many-body decays can have rich decay structures, with various allowed intermediate resonances. To correctly simulate them would need extensive studies, which would go far beyond the scope of the work presented here.

A recently published measurement of lepton flavour universality from LHCb [29, 30], using $B^+ \rightarrow K^+ \ell^+ \ell^-$ and $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ decays, introduced a data driven way to estimate the inclusive contribution from misidentified backgrounds. The strategy uses control regions which are enriched with candidates where at least one electron is actually a misidentified hadron. Using the label Pass (P) for the electron candidates that pass the nominal PID requirement, introduced in Sec. 4.3, and Fail (F) for those that don't, three inverted regions can be defined:

- Pass-Fail (PF): where the e^+ passes the PID selection, but the e^- doesn't.
- Fail-Pass (FP): where the e^+ fails the PID selection, but the e^- passes it.
- Fail-Fail (FF): where both electrons fail the PID requirement.

Because of the requirements at the stripping stage, all regions are defined with $\text{DLL}_{e\pi} > 0$. The total yield of the PP region can be related to the yields in the inverted regions by

$$\mathcal{N}_{\text{PP}} = \mathcal{N}_{e^+e^-} + \mathcal{N}_{h^+e^-} + \mathcal{N}_{e^+h^-} + \mathcal{N}_{h^+h^-} = \mathcal{N}_{e^+e^-} + \mathcal{N}_{\text{misid}}, \quad (4.9)$$

where $\mathcal{N}_{ij}$ is the yield of the “real” final state (i, j) , $\mathcal{N}_{ij} = N_{ij} \cdot \epsilon_{i \rightarrow e} \cdot \epsilon_{j \rightarrow e}$, with i, j being an electron or any charged hadron. One can similarly write the yields in the other three regions, where the selection efficiency for the Fail category will just be $(1 - \epsilon_{i \rightarrow e})$. Defining the following transfer functions,

$$T_h = \frac{\epsilon_{h \rightarrow e}}{1 - \epsilon_{h \rightarrow e}}, \quad (4.10)$$

and applying them to the three inverted regions (applying it twice to the FF category, assuming that the efficiencies are the same for both lepton charges), one finds that

$$\begin{aligned} \mathcal{N}_{h^+e^-} + \mathcal{N}_{e^+h^-} + \mathcal{N}_{h^+h^-} &= (\mathcal{N}_{\text{FP}} - \mathcal{N}_{\text{FP}, e^+e^-}) \cdot T_h + (\mathcal{N}_{\text{PF}} - \mathcal{N}_{\text{PF}, e^+e^-}) \cdot T_h \\ &\quad - (\mathcal{N}_{\text{FF}} - \mathcal{N}_{\text{FF}, e^+e^-}) \cdot T_h^2 \end{aligned} \quad (4.11)$$

This means that applying the transfer functions with the signature FP+PF–FF, one obtains the total misidentified yield in the signal region, with a small contamination due to real dielectron candidates which end up in one of the inverted regions. Note here that the contamination from real dielectron pairs should also include contributions from the leakage backgrounds, in the central- and high- q^2 bins. This contamination can be removed using simulated $B_s^0 \rightarrow \phi e^+ e^-$, $B_s^0 \rightarrow J/\psi (e^+ e^-) \phi$, and $B_s^0 \rightarrow \psi(2S) (e^+ e^-) \phi$ decays.

The signal efficiency $\epsilon_{e \rightarrow e}$ and misidentification rates $\epsilon_{h \rightarrow e}$ are also derived with a fully data-driven approach [76, 77], as previously described in Sec. 3.4.4. For the work described in Chap. 5, only misidentification of kaons and pions is considered. The transfer functions are determined in bins of the transverse momenta and pseudorapidity, and are applied using a binary decision based on the PID response of the candidate: if $\text{ProbNN}_K > 0.1$, then the kaon transfer function is applied; the pion transfer function is applied otherwise. In Chap. 6, the method outlined here is expanded to include contributions from misidentified protons.

In practice, the transfer functions are applied as per-event weights, which then allows to obtain both an estimate for the expected number of background candidates, as well as a way to determine the background's $m(K^+ K^- e^+ e^-)$ distribution. The expected yields are determined as the signal-subtracted weighted sum of all candidates in the three inverted regions, and the uncertainty on the expectation is the square root of the sum of the squared weights. In total, the background is expected to contribute to 27.0 ± 3.8 candidates in the low- q^2 bin, 105.8 ± 7.2 candidates in the central- q^2 bin, and 45.5 ± 5.2 candidates in the high- q^2 bin.

Test of lepton flavour universality using $B_s^0 \rightarrow \phi \ell^+ \ell^-$ decays

The author's recommendation

RENAISSANCE

by **Beyoncé**

2022, Parkwood Entertainment/Columbia

This chapter is dedicated to the first test of lepton flavour universality (LFU) with $B_s^0 \rightarrow \phi \ell^+ \ell^-$ decays, performed with data collected by the LHCb experiment in its first and second data-taking runs [114]. As introduced in Chap. 2, ratios of branching fractions offer an ideal test of LFU, its theoretical predictions largely free of the systematic uncertainty associated to the hadronic dynamics of the decay. I will first introduce the experimental strategy to measure the ratio $R_\phi = \mathcal{B}(B_s^0 \rightarrow \phi \mu^+ \mu^-) / \mathcal{B}(B_s^0 \rightarrow \phi e^+ e^-)$ in bins of the dilepton mass squared, q^2 . A novelty in this measurement is the inclusion of a bin at high dielectron mass squared, making this the first LFU ratio measured by LHCb in this region. The inclusion of this bin was possible thanks to the smaller background contamination in the $B_s^0 \rightarrow \phi \ell^+ \ell^-$ final state, compared to its non-strange counterpart, $B^0 \rightarrow K^{*0} \ell^+ \ell^-$. The unique challenges introduced in this regime will be discussed. This chapter will also briefly cover the computation of the selection efficiencies, which enter the determination of R_ϕ . The analysis uses the $B_s^0 \rightarrow J/\psi \phi$ and $B_s^0 \rightarrow \psi(2S) \phi$ decays as control modes to validate the fit strategy and simulation correction chain. The modelling of the signal and background mass distribution will be introduced, as well as the final $B_s^0 \rightarrow \phi \ell^+ \ell^-$ mass fits, reporting the first observation of the $B_s^0 \rightarrow \phi e^+ e^-$ mode. After an overview of the sources of systematic uncertainty affecting this measurement, the final results and their compatibility with the Standard Model will be discussed.

5.1 Measurement strategy

The R_ϕ ratio is measured in three bins of dilepton mass squared, q^2 , as defined in Table 4.1. The theoretical definition reads, in a q^2 bin $[q_1^2, q_2^2]$,

$$R_\phi(q_1^2, q_2^2) = \frac{\int_{q_1^2}^{q_2^2} \frac{d\Gamma}{dq^2}(B_s^0 \rightarrow \phi \mu^+ \mu^-) dq^2}{\int_{q_1^2}^{q_2^2} \frac{d\Gamma}{dq^2}(B_s^0 \rightarrow \phi e^+ e^-) dq^2}. \quad (5.1)$$

There are known differences, outlined in the previous chapter, between the electron and muon reconstruction. Experimentally, these can become sources of significant systematic uncertainty. By measuring R_ϕ as a double ratio, introducing the branching fractions of the control modes $B_s^0 \rightarrow J/\psi (\mu^+ \mu^-) \phi$ and $B_s^0 \rightarrow J/\psi (e^+ e^-) \phi$, one can ensure the cancellation of many of these sources of uncertainty. The introduction of the second ratio has a negligible effect on the value of R_ϕ , as the branching fraction of the control mode is known to respect LFU at the per-mille level [17].

Furthermore, because of their lower selection and reconstruction efficiency, the expected number of electron signal candidates is small and more sensitive to statistical fluctuations. Having the electron mode in the denominator would result in a statistically less well-behaved likelihood function. Thus, the more preferable observable is the inverse of the ratio,

$$R_\phi^{-1}(q_1^2, q_2^2) = \frac{\int_{q_1^2}^{q_2^2} \frac{d\Gamma}{dq^2}(B_s^0 \rightarrow \phi e^+ e^-) dq^2}{\int_{q_1^2}^{q_2^2} \frac{d\Gamma}{dq^2}(B_s^0 \rightarrow \phi \mu^+ \mu^-) dq^2} \cdot \frac{\mathcal{B}(B_s^0 \rightarrow J/\psi (\mu^+ \mu^-) \phi)}{\mathcal{B}(B_s^0 \rightarrow J/\psi (e^+ e^-) \phi)}. \quad (5.2)$$

One can measure the ratio using the observed yields N and selection efficiencies $\epsilon_{\ell^+ \ell^-}$ (with an additional subscript to indicate the selection efficiency of the control mode), of each final state. Thus experimentally,

$$R_\phi^{-1}(q_1^2, q_2^2) = \frac{N(B_s^0 \rightarrow \phi e^+ e^- | q_1^2, q_2^2)}{N(B_s^0 \rightarrow \phi \mu^+ \mu^- | q_1^2, q_2^2)} \cdot \frac{N(B_s^0 \rightarrow J/\psi (\mu^+ \mu^-) \phi)}{N(B_s^0 \rightarrow J/\psi (e^+ e^-) \phi)} \cdot \frac{\epsilon_{e^+ e^-}^{J/\psi}}{\epsilon_{\mu^+ \mu^-}^{J/\psi}} \cdot \frac{\epsilon_{\mu^+ \mu^-}(q_1^2, q_2^2)}{\epsilon_{e^+ e^-}(q_1^2, q_2^2)}, \quad (5.3)$$

where the explicit dependence on the definition of the q^2 bin is shown. Unless otherwise specified, R_ϕ^{-1} will be used to mean the observable in each of the q^2 bins. The yields are determined through extended maximum likelihood fits to the $m(K^+ K^- \ell^+ \ell^-)$ distributions, while the efficiency ratios are determined using simulated $B_s^0 \rightarrow \phi \ell^+ \ell^-$ and $B_s^0 \rightarrow J/\psi (\ell^+ \ell^-) \phi$ decays, with corrections applied as described in the previous chapter.

The control modes can be used to define similar branching fraction ratios, which are used

to validate the analysis strategy. The single ratio entering the definition of R_ϕ^{-1} ,

$$r_{J/\psi} = \frac{N(B_s^0 \rightarrow J/\psi (\mu^+ \mu^-) \phi)}{N(B_s^0 \rightarrow J/\psi (e^+ e^-) \phi)} \cdot \frac{\epsilon_{e^+ e^-}^{J/\psi}}{\epsilon_{\mu^+ \mu^-}^{J/\psi}}, \quad (5.4)$$

is expected to be unity [17], but does not benefit from the same cancellation of systematic uncertainties as in the double ratio. Thus, it is used to validate the simulation corrections that enter the computation of the efficiency ratios. The double ratio

$$R_{\psi(2S)} = \frac{N(B_s^0 \rightarrow \psi(2S) (\mu^+ \mu^-) \phi)}{N(B_s^0 \rightarrow \psi(2S) (e^+ e^-) \phi)} \cdot \frac{\epsilon_{e^+ e^-}^{\psi(2S)}}{\epsilon_{\mu^+ \mu^-}^{\psi(2S)}} \cdot r_{J/\psi}^{-1}, \quad (5.5)$$

is also expected to be unity [17], and is used to further validate the fit strategy, and to show the robustness of the double ratio against systematic uncertainties related to reconstruction.

The analysis is performed blind, to avoid experimenter's bias. In particular, since the $B_s^0 \rightarrow \phi e^+ e^-$ mode has never been observed before, the yields of the $B_s^0 \rightarrow \phi \ell^+ \ell^-$ modes are kept hidden. Furthermore, the signal mass regions defined in Table 4.2 are removed from the dataset. The result of the final fits are only inspected once all aspects of the analysis are finalised.

5.2 The high- q^2 bin

The high- q^2 region, defined as $q^2 \in [15, 19] \text{ GeV}^2/c^4$, above the $\psi(2S)$ resonance, presents unique challenges. At very high q^2 , relativistic kinematics constrain the allowed phase space, such that, within the considered fit ranges in $m(K^+ K^- \ell^+ \ell^-)$, the signal and background distributions are subject to a sculpting effect at a low masses. Integrating over q^2 , not accounting for resolution effects, this should correspond to an observed cutoff in the candidate B_s^0 mass distribution at $m(K^+ K^- \ell^+ \ell^-) \simeq \sqrt{q_{\min}^2} + m_\phi$, where $q_{\min}^2$ is the lower edge of the q^2 bin and m_ϕ is the ϕ meson's rest mass [17]. This corresponds, for $q_{\min}^2 = 15 \text{ GeV}^2/c^4$, to an expected cutoff at $m(K^+ K^- \ell^+ \ell^-) \simeq 4.9 \text{ GeV}/c^2$. Figure 5.1 illustrates the phase-space constraints in the two-dimensional distributions of the candidate B_s^0 mass and q^2 , for the dilepton final states.

The cutoff sculpts the signal and background distributions, which need to be carefully modelled. The signal distribution peaks around the known B_s^0 mass at $5366 \text{ MeV}/c^2$, far from the expected cutoff. Thus, only the lower mass tail of the distribution is sculpted by the phase-space effects. The core distribution is modelled with double-sided Crystal Ball function. This distribution is a modified version of the Crystal Ball function as described in Ref. [115], which includes two power-law tails and an asymmetric Gaussian core, and is

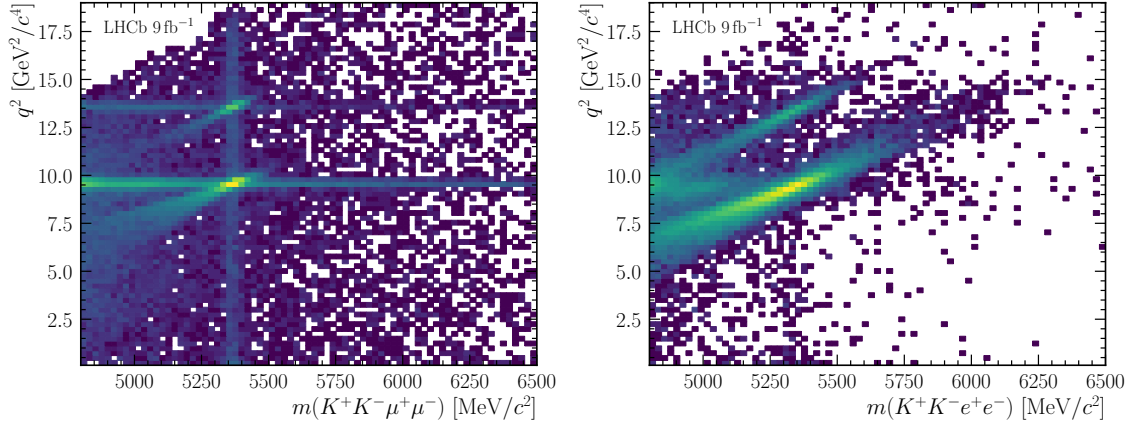


Figure 5.1: The observed two-dimensional distributions of the dilepton mass squared q^2 and $m(K^+K^-\ell^+\ell^-)$ mass in LHCb data, for the muon (left) and electron (right) final states. The $B_s^0 \rightarrow J/\psi (\ell^+\ell^-) \phi$ and $B_s^0 \rightarrow \psi(2S) (\ell^+\ell^-) \phi$ resonances are visible (the diagonal bands), as well as combinatorial background containing real J/ψ or $\psi(2S)$ at low $m(K^+K^-\ell^+\ell^-)$ mass (the horizontal bands). The vertical band visible in the left plot are non-resonant $B_s^0 \rightarrow \phi\mu^+\mu^-$ decays. The kinematic phase-space boundary is visible in the upper left corner of the plots. Figures taken from Ref. [114].

expressed as (minus a normalisation factor),

$$f(m|\mu, \sigma_L, \sigma_R, \alpha_L, n_L, \alpha_R, n_R) = \begin{cases} A_L \cdot \left(B_L - \frac{m-\mu}{\sigma_L}\right)^{-n_L} & \frac{m-\mu}{\sigma_L} < -\alpha_L \\ \exp\left(-\frac{(m-\mu)^2}{2\sigma_L^2}\right) & -\alpha_L < \frac{m-\mu}{\sigma_L} < 0 \\ \exp\left(-\frac{(m-\mu)^2}{2\sigma_R^2}\right) & 0 < \frac{m-\mu}{\sigma_R} < \alpha_R \\ A_R \cdot \left(B_R - \frac{m-\mu}{\sigma_R}\right)^{n_R} & \frac{m-\mu}{\sigma_R} > \alpha_R \end{cases}, \quad (5.6)$$

where μ , $\sigma_{L,R}$ are the mean and widths of the Gaussian core, and $\alpha_{L,R}$ are the turn-on points for the power-law tails with powers $n_{L,R}$. The factors $A_{L,R}$ and $B_{L,R}$ ensure that the function is continuous and differentiable for all values of m .

To model the change in slope of the lower mass tail, an empirical function is implemented to modulate the Crystal Ball function. The modulation is chosen to be a sigmoid function, which multiplies Eq. (5.6),

$$f_{\text{modulated}}(m) = \frac{1}{1 + e^{-c(m-m_0)}} \cdot f(m|\mu, \sigma_L, \sigma_R, \alpha_L, n_L, \alpha_R, n_R), \quad (5.7)$$

where the parameter m_0 is the turn-on parameter of the modulation (value at which the non-modulated PDF is reduced by a factor of a half), and c expresses how sharply the tail is modulated. Example results of the unbinned maximum likelihood fits to simulated candidates, for both signal modes, are shown in Fig. 5.2. Additionally, it is cross-checked, for the electron mode, that the parameters for the modulating sigmoid do not depend on

the bremsstrahlung reconstruction. This is verified by separately fitting three subsets of the simulated dataset, corresponding to the three bremsstrahlung categories defined in Sec. 4.6.6. The parameters m_0 and c are found to be compatible in the three fits, which is expected since this effect is only due to the overall phase-space constraints.

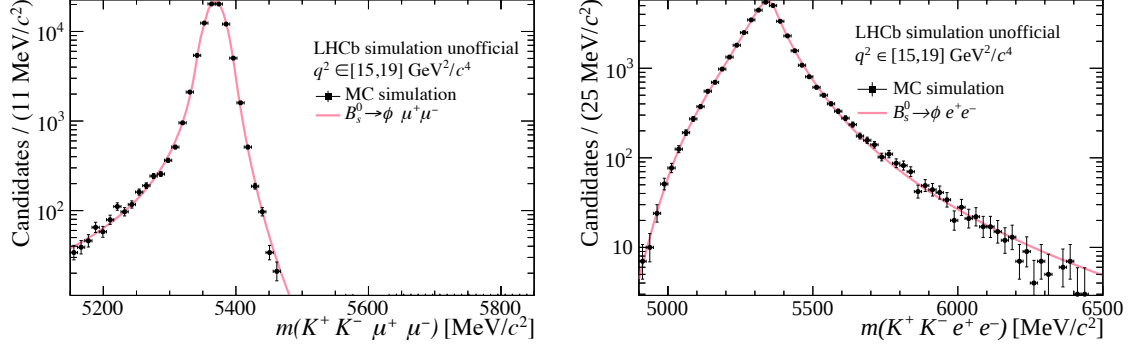


Figure 5.2: Result of unbinned maximum likelihood fits to the $m(K^+K^-\ell^+\ell^-)$ distributions of simulated signal muon (left) and electron (right) candidates. The plot is in logarithmic scale on the y axis to show the modulation of the lower mass tail due to the phase-space effects.

Proxies for the combinatorial background

The most sculpted component is the combinatorial background. In the low- and central- q^2 bins, its mass distribution follows a falling exponential curve. At high- q^2 , because of the phase-space cutoff, the exponential curve is instead suppressed: at low $m(K^+K^-\ell^+\ell^-)$ values, one should then expect a turn-on behaviour, starting from the cutoff point; then, going higher in $m(K^+K^-\ell^+\ell^-)$, the curve should slow its rise and reach a peak, after which the falling exponential should become dominant. To investigate and understand these different regimes, appropriate proxy datasets are needed, which are affected by the same phase-space constraints. This is fundamental because, as it will become clear when fitting the full mass spectrum in Sec. 5.5, it isn't possible at high- q^2 to isolate a clean sideband with only combinatorial-like candidates.

Three proxies have been used in this work: the first contains collision data which has been selected to contain a same-sign (SS) final state, *i.e.* with either a lepton pair with the same charge, $K^+K^-\ell^\pm\ell^\pm$, or a kaon pair with the same charge, $K^\pm K^\pm\ell^+\ell^-$; as these final states violate charge conservation, they can only be reconstructed if some of the particles are actually coming from separate decays, and have been randomly combined in the reconstruction process. Through dedicated stripping lines SS candidates are selected, and are passed through the same nominal selection described in the previous chapter, with the exception of the combinatorial BDT. The selection in this case is relaxed to maximise the size of the available sample. This proxy is the closest of the three to the real distribution for OS data, and it is used to check the validity of the models derived from the other two proxy datasets.

The second proxy consists of track recombinations. This dataset is built starting from

fully selected opposite-sign (OS) data, randomly recombining tracks from different events, with the following prescription:

- Select at random four events from the dataset, across the full q^2 range.
- Select from each event one of the four charged tracks (K^+ , K^- , e^+ , e^-).
- Rotate the four tracks so that the momenta of the original four B_s^0 candidates align.
- Combine the selected tracks into a new combinatorial-like fake B_s^0 candidate and recompute all kinematic variables of interest.
- Finally, retain only candidates with a recomputed q^2 in the range $[15, 19] \text{ GeV}^2/c^4$.

Figure 5.3 shows a sketch of this method.

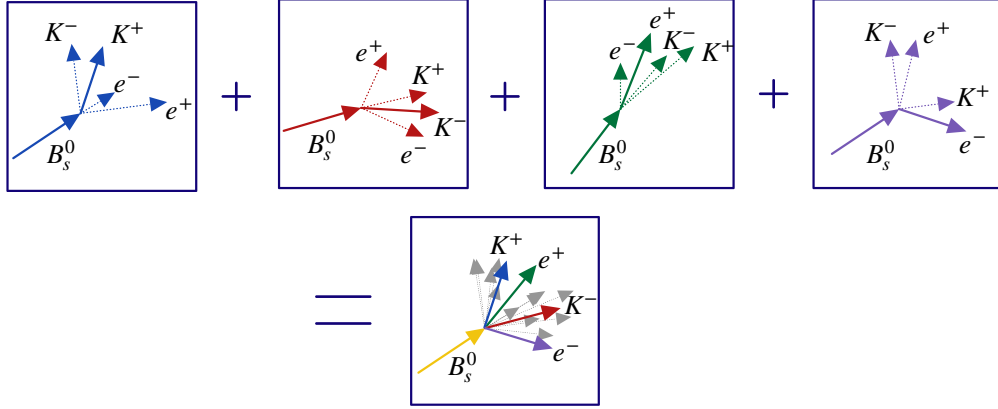


Figure 5.3: Sketch of the prescription to build a proxy combinatorial-like sample made of recombination of random tracks.

As the number of possible recombinations of tracks scales rapidly as the binomial coefficient ($N_{\text{candidates}}^4$), it is easy to obtain a very large sample of this kind. Figure 5.4 shows the result of the recombination of electron mode candidates.

The third proxy is pseudodata generated in two dimensions ($q^2, m(K^+K^-\ell^+\ell^-)$), according to the distributions observed in OS data. The distributions are determined by performing maximum likelihood fits to the lepton upper mass sidebands, including the contribution from the other polluting backgrounds. It is assumed that the two dimensions are uncorrelated for the combinatorial component. As previously mentioned, the $m(K^+K^-\ell^+\ell^-)$ shape is well described by an exponential curve in the upper mass sideband. For the q^2 shape instead three different distributions are trialed: an exponential curve, a uniform distribution, and a truncated Gaussian function. The exponential is chosen as nominal while the other two are used for the study of the systematic uncertainties associated to the acceptance model (see Sec. 5.6). Figure 5.5 shows the result of the nominal fit to the q^2 distribution in the electron upper mass sideband. After generating events in $(q^2, m(K^+K^-\ell^+\ell^-))$, a weight is applied

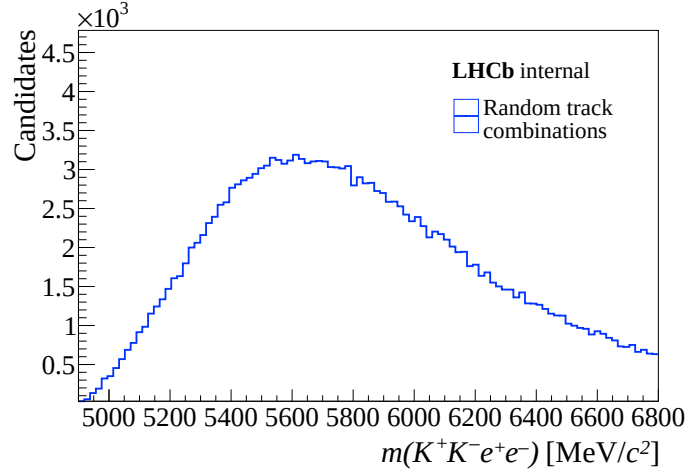


Figure 5.4: Invariant B_s^0 mass distribution for fake candidates built from the recombination of random tracks from fully selected $B_s^0 \rightarrow \phi e^+ e^-$ data.

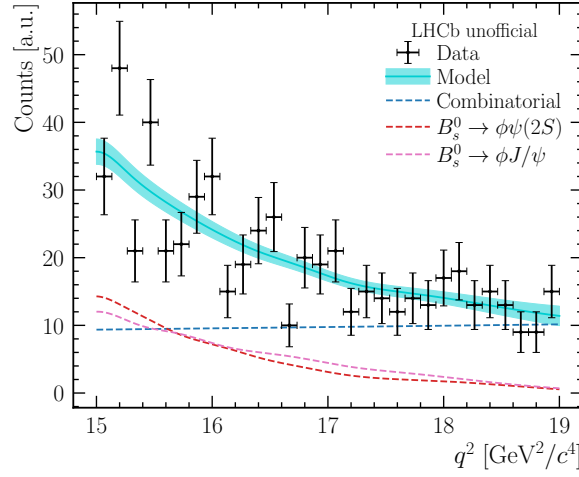


Figure 5.5: Result of the maximum likelihood fit to the OS data electron q^2 spectrum in the upper mass sideband $m(K^+ K^- e^+ e^-) \in [5500, 6500]$ MeV/c^2 . The pollution from leakage $B_s^0 \rightarrow J/\psi (e^+ e^-) \phi$ and $B_s^0 \rightarrow \psi(2S) (e^+ e^-) \phi$ decays is constrained to the expectation value computed from simulation.

to simulate the effect of the physical boundary. The weight function is applied per event and has the form

$$w_i = \frac{1}{1 + \exp \left(\frac{q^2 - (m(K^+ K^- \ell^+ \ell^-) - m_{\min}(K^+ K^-))^2}{\sigma} \right)}, \quad (5.8)$$

where $m_{\min}(K^+ K^-)$ is the lower bound of the selected dikaon mass spectrum, and a constant $\sigma = 1 \text{ MeV}/c^2$ reflects the finite dikaon mass resolution. Figure 5.6 illustrates the generated proxy sample before and after the application of the phase-space weights.

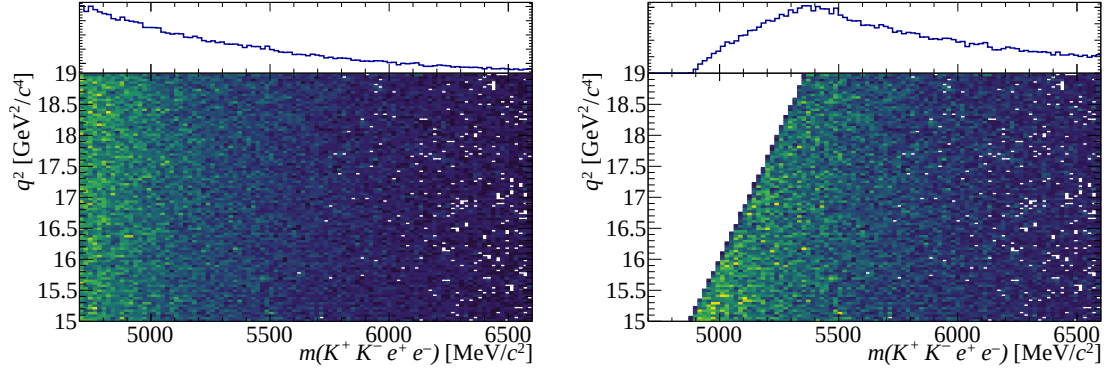


Figure 5.6: Illustration of the generated proxy sample in $(q^2, m(K^+ K^- e^+ e^-))$. On the left plot, the unweighted distribution; on the right, the distribution after the application of the weights which simulate the effect of the kinematic phase-space boundary. The resulting $m(K^+ K^- e^+ e^-)$ projections are shown in blue.

Combinatorial models

The sample of recombined tracks is not fully representative of real combinatorial events, which are not just a mix of purely random combinations of charged tracks. However, since it's possible to generate a very large sample, it still serves as a proxy to understand the effects of the constraints due to phase space, and can be used as a starting point to get a handle on the general shape. The function chosen, which best captures the varying regimes and features across $m(K^+ K^- \ell^+ \ell^-)$, is a product of polynomials,

$$f(m|m_0, n, k) \propto \begin{cases} (m - m_0)^n \cdot m^{-k} & m > m_0 \\ 0 & m \leq m_0 \end{cases} \quad (5.9)$$

where the parameter m_0 is the physical boundary point at which the PDF turns on, and the combination of the powers n, k determines the slope of the rise and consequently the position of the peaking shoulder. Figure 5.7 shows the result of an unbinned maximum likelihood fit to the $m(KK\ell\ell)$ distributions for SS muon and electron data. The empirical model agrees well with the observed distribution.

The empirical model, while versatile, has one disadvantage, which lies in the number of its free parameters. With the available sample size, and the background composition of the high- q^2 bin, it is not possible to float freely its three parameters in the mass fit. Instead, one needs to add additional constraints to the model, for example coming from the SS datasets, to ensure a well-behaved likelihood fit. An alternative approach tries to address this shortcoming by modelling the effects of the phase-space boundary as an acceptance function, which sculpts the expected exponential behaviour of combinatorial events,

$$f(m|\lambda, q_{\min}^2) \sim \epsilon(m|q_{\min}^2) \cdot e^{-\lambda m}. \quad (5.10)$$

The acceptance function $\epsilon(m|q_{\min}^2)$ accounts for the phase-space suppression and should only

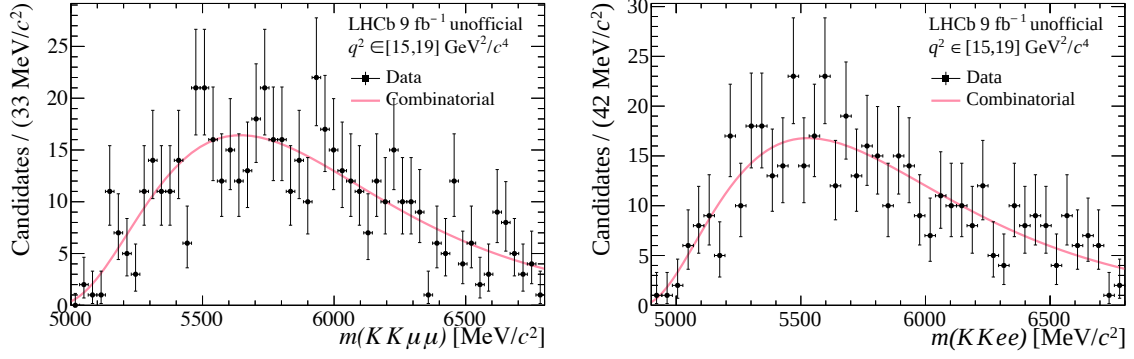


Figure 5.7: Result of unbinned maximum likelihood fits to same-sign (SS) collision data for the muon (left) and electron (right) modes, using an empirical polynomial model. The candidates are selected after relaxing the combinatorial BDT selection.

ideally be dependent on the chosen threshold for the q^2 bin, $q_{\min}^2$, and the underlying q^2 distribution. The parametrisation of $\epsilon(m|q_{\min}^2)$ can then be fixed in the fit to the $m(K^+K^-\ell^+\ell^-)$ distributions, leaving effectively only one parameter floating, the slope of the exponential.

The acceptance function is determined by taking the ratio of the weighted and unweighted distributions of the third proxy, and projecting onto $m(K^+K^-\ell^+\ell^-)$. The function which best describes the observed acceptance is parametrised as

$$\epsilon(m|q_{\min}^2, a, b, x_0) = \begin{cases} 0, & \text{for } m \leq x_0 \\ b \cdot (1 - e^{(-a \cdot (x - x_0))}), & \text{for } m > x_0 \\ 1, & \text{for } b \cdot (1 - e^{(-a \cdot (m - x_0))}) \geq 1 \end{cases} \quad (5.11)$$

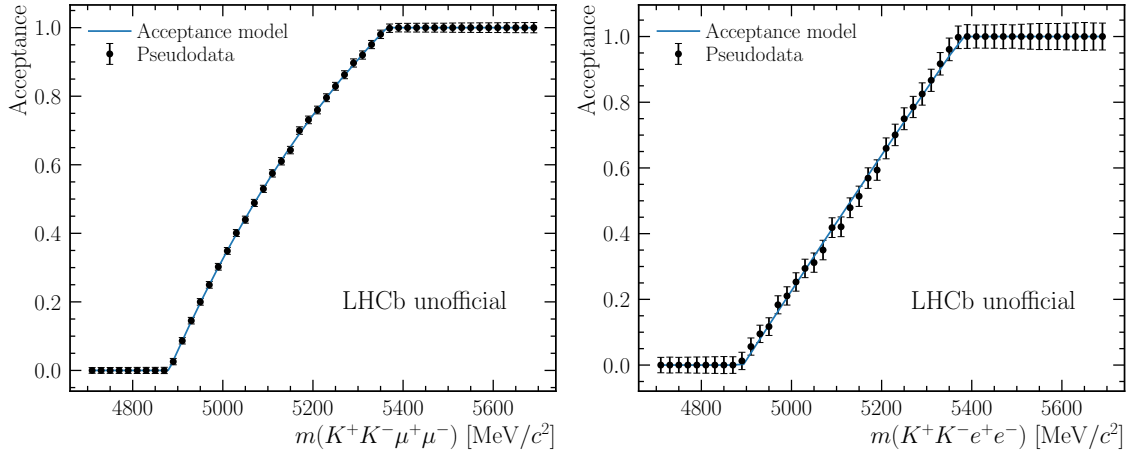


Figure 5.8: Acceptance functions obtained from a least squares fits to the ratios of weighted and unweighted generated $m(K^+K^-\mu^+\mu^-)$ (left) and $m(K^+K^-e^+e^-)$ (right) distributions.

Figure 5.8 shows the results of least squares fits to the ratios of distributions for the two lepton modes, and the resulting acceptance functions. The obtained acceptance models are tested against the SS datasets. Figure 5.9 shows the results of maximum likelihood fits

to the $m(KK\ell\ell)$ distributions. The resulting shapes are found to model well the SS mass distributions.

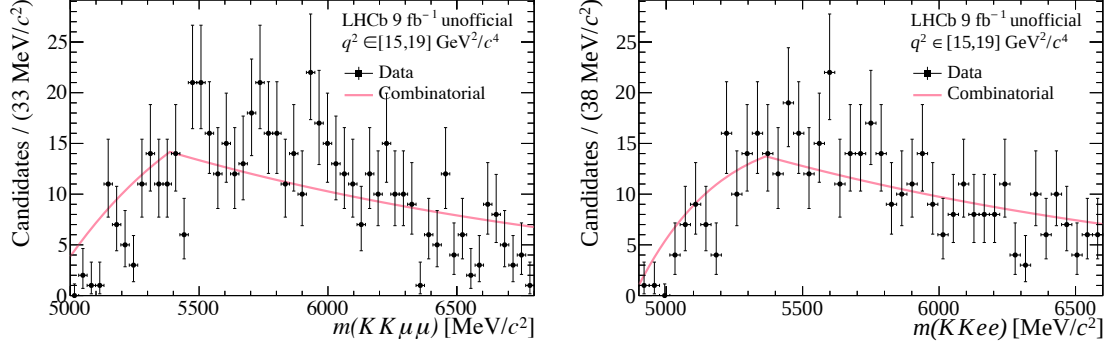


Figure 5.9: Result of the maximum likelihood fits to SS muon (right) and electron (left) data with the acceptance corrected exponential. The acceptance is fixed in the fit, the exponential slope is floating freely.

As it will be shown in Sec. 5.4, pseudoexperiments are used to determine which of the two models, the empirical one based on polynomials, or the acceptance modulated exponential, returns the best expected sensitivity to R_ϕ^{-1} in the high- q^2 bin.

5.3 Efficiency computation

As described in Sec. 5.1, R_ϕ^{-1} is measured experimentally as a double ratio of efficiency-corrected yields. The selection and reconstruction efficiencies are computed using simulated $B_s^0 \rightarrow \phi\ell^+\ell^-$ decays, after having applied the correction weights described in Chap. 4.

To take into account the possible migration of events across the q^2 range, an additional selection based on the true- q^2 of each candidate is applied: the efficiencies are only computed for candidates which had, at generator level, a true- q^2 falling in the desired interval $[0.1, 19] \text{ GeV}^2/c^4$. The true- q^2 variable is computed from the difference of the B_s^0 and ϕ four-momenta, to exclude effects from final state radiation. The variable used to obtain the efficiency for each q^2 bin for the electron candidates is the smeared q_{smeared}^2 , so that the data-driven corrections to the dielectron mass are also accounted for. Computing efficiencies relative to these requirements ensures R_ϕ is well defined over each q^2 bin.

Efficiencies are generally computed as a ratio of the number of candidates passing the selection requirements N_{pass} , and the total number of candidates which were generated N_{gen} ,

$$\epsilon = \frac{N_{\text{pass}}}{N_{\text{gen}}}. \quad (5.12)$$

When making this calculation on weighted candidates, like in the case of the correction weights applied to the simulated samples, one must be careful that said weights are properly normalised to number of candidates. For example, the weights computed as the output of a BDT are not normalised to the number of simulated events available. Thus, replacing the

numerator in Eq. (5.12) with a weighted sum over the selected candidates would not return the correct selection efficiency. The normalisations of each set of weights are treated in more detail in App. A.

Tables 5.1 and 5.2 report the obtained efficiencies for simulated $B_s^0 \rightarrow \phi \ell^+ \ell^-$ decays per data-taking period and trigger category. The efficiencies for the $B_s^0 \rightarrow \psi(2S) (\ell^+ \ell^-) \phi$ mode, as well as for the expected backgrounds, are computed similarly and used to obtain the expectation values for their yields in each q^2 bin.

Table 5.1: Reconstruction and selection efficiency in permille for signal simulation samples in the L0I trigger category.

Decay Mode	Run 1	Run 2p1	Run 2p2
$B_s^0 \rightarrow \phi \mu^+ \mu^-$ (low- q^2)	9.69 ± 0.17	14.23 ± 0.19	14.17 ± 0.15
$B_s^0 \rightarrow \phi e^+ e^-$ (low- q^2)	4.14 ± 0.08	6.38 ± 0.12	6.33 ± 0.07
$B_s^0 \rightarrow \phi \mu^+ \mu^-$ (central- q^2)	10.34 ± 0.12	14.69 ± 0.14	14.94 ± 0.11
$B_s^0 \rightarrow \phi e^+ e^-$ (central- q^2)	4.67 ± 0.06	7.12 ± 0.09	7.28 ± 0.06
$B_s^0 \rightarrow \phi \mu^+ \mu^-$ (high- q^2)	8.26 ± 0.12	11.87 ± 0.13	11.84 ± 0.10
$B_s^0 \rightarrow \phi e^+ e^-$ (high- q^2)	3.09 ± 0.04	4.76 ± 0.06	5.04 ± 0.04
$B_s^0 \rightarrow J/\psi (\mu^+ \mu^-) \phi$	12.00 ± 0.02	16.98 ± 0.07	17.13 ± 0.03
$B_s^0 \rightarrow J/\psi (e^+ e^-) \phi$	4.26 ± 0.01	7.26 ± 0.03	7.29 ± 0.02

Table 5.2: Reconstruction and selection efficiency in permille for signal simulation samples in the LOL! trigger category.

Decay Mode	Run 1	Run 2p1	Run 2p2
$B_s^0 \rightarrow \phi \mu^+ \mu^-$ (low- q^2)	21.63 ± 0.25	27.24 ± 0.26	30.91 ± 0.22
$B_s^0 \rightarrow \phi e^+ e^-$ (low- q^2)	3.16 ± 0.07	5.17 ± 0.11	5.52 ± 0.07
$B_s^0 \rightarrow \phi \mu^+ \mu^-$ (central- q^2)	23.21 ± 0.18	28.63 ± 0.19	32.91 ± 0.16
$B_s^0 \rightarrow \phi e^+ e^-$ (central- q^2)	3.85 ± 0.05	6.47 ± 0.08	6.62 ± 0.05
$B_s^0 \rightarrow \phi \mu^+ \mu^-$ (high- q^2)	21.62 ± 0.19	26.03 ± 0.19	29.05 ± 0.16
$B_s^0 \rightarrow \phi e^+ e^-$ (high- q^2)	3.09 ± 0.05	4.76 ± 0.07	5.04 ± 0.04
$B_s^0 \rightarrow J/\psi (\mu^+ \mu^-) \phi$	29.27 ± 0.04	35.76 ± 0.09	40.79 ± 0.04
$B_s^0 \rightarrow J/\psi (e^+ e^-) \phi$	4.16 ± 0.01	7.15 ± 0.03	7.38 ± 0.02

Because some of the corrections are shared across the different q^2 bins and datasets, the efficiencies are expected to be correlated. This can be quantified by a bootstrapping approach [116]. For each of the different samples used to compute the correction weights, 100 sets of per-event weights are extracted from a Poisson distribution of unit mean. For each set of weights, also called a bootstrap, the corrections are recomputed, as well as the efficiencies. Then, by looking at the distribution of each efficiency over the ensemble of bootstraps, it is possible to obtain the correlation matrix across the various datasets and q^2 bins. Figure 5.10 shows the correlation matrix obtained for all samples simulated in 2017–2018 data-taking conditions. The matrix is used in the fit model to build a multidimensional Gaussian constraint of the efficiencies, which ensures that the correlations are taken into

account.

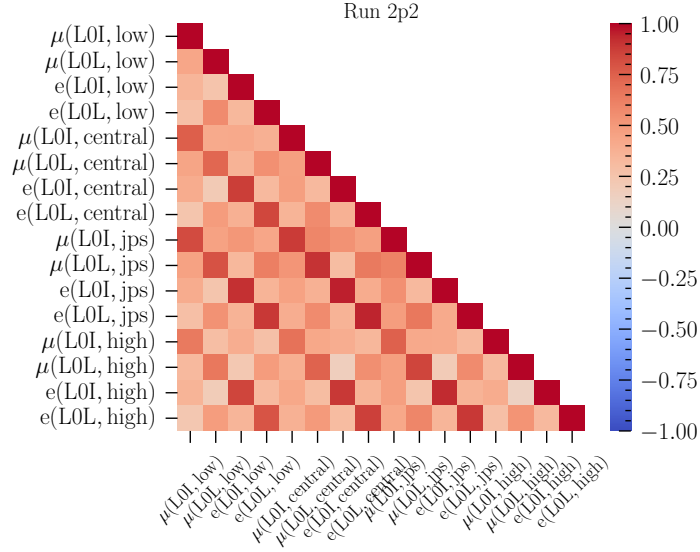


Figure 5.10: Correlation matrix of the computed efficiencies, across all q^2 bins, trigger categories, and lepton final states, for events simulated in 2017–2018 data-taking conditions.

5.4 Cross checks and validation

The validation of the fitting strategy outlined in Sec. 5.1 is done by performing unbinned maximum likelihood fits to the mass distributions of the resonant $B_s^0 \rightarrow J/\psi (\ell^+ \ell^-) \phi$ and $B_s^0 \rightarrow \psi(2S) (\ell^+ \ell^-) \phi$ modes, and by generating pseudoexperiments to assess the statistical properties of the fit. This section presents the measurement of $r_{J/\psi}$ and $R_{\psi(2S)}$, as well as the result of the pseudoexperiments.

5.4.1 Measurement of $r_{J/\psi}$ and $R_{\psi(2S)}$

The maximum likelihood fits to the $B_s^0 \rightarrow J/\psi (\ell^+ \ell^-) \phi$ and $B_s^0 \rightarrow \psi(2S) (\ell^+ \ell^-) \phi$ mass distributions are performed using the constrained four-body masses $m_{J/\psi}(K^+ K^- \ell^+ \ell^-)$ and $m_{\psi(2S)}(K^+ K^- \ell^+ \ell^-)$. The subscripts indicate that the mass has been recomputed by requiring the dilepton system to have either the mass of the J/ψ or $\psi(2S)$ meson [17], respectively. The additional information improves the overall mass resolution. Table 5.3 reports the fit ranges for the two q^2 bins.

Table 5.3: Definitions of the $m_{J/\psi}(K^+ K^- \ell^+ \ell^-)$ mass ranges used for the fits to the data.

q^2 bin	$\mu^+ \mu^-$	$e^+ e^-$
J/ψ - q^2	$m_{J/\psi}(K^+ K^- \mu^+ \mu^-) \in [5150, 5850] \text{ MeV}/c^2$	$m_{J/\psi}(K^+ K^- e^+ e^-) \in [5150, 5850] \text{ MeV}/c^2$
$\psi(2S)$ - q^2	$m_{\psi(2S)}(K^+ K^- \mu^+ \mu^-) \in [5150, 5850] \text{ MeV}/c^2$	$m_{\psi(2S)}(K^+ K^- e^+ e^-) \in [5150, 5850] \text{ MeV}/c^2$

The fits are performed simultaneously to the muon and electron datasets and for both q^2 bins; each dataset is split in three run periods, Run 1 (2011–2012), Run 2p1 (2015–2016) and

Run 2p2 (2017–2018). The electron yields are expressed as a composite parameter depending on the muon yields, by inverting Eqs. (5.4) and 5.5,

$$N(B_s^0 \rightarrow J/\psi (e^+e^-) \phi) = N(B_s^0 \rightarrow J/\psi (\mu^+\mu^-) \phi) \cdot \frac{\epsilon_{e^+e^-}^{J/\psi}}{\epsilon_{\mu^+\mu^-}^{J/\psi}} \cdot r_{J/\psi}^{-1}, \quad (5.13)$$

$$N(B_s^0 \rightarrow \psi(2S) (e^+e^-) \phi) = N(B_s^0 \rightarrow \psi(2S) (\mu^+\mu^-) \phi) \cdot \frac{\epsilon_{e^+e^-}^{\psi(2S)}}{\epsilon_{\mu^+\mu^-}^{\psi(2S)}} \cdot r_{J/\psi}^{-1} \cdot R_{\psi(2S)}^{-1}, \quad (5.14)$$

leaving $r_{J/\psi}$ and $R_{\psi(2S)}$ free to float. The parameter $r_{J/\psi}$ is shared across the J/ψ - and $\psi(2S)$ - q^2 regions.

Figure 5.11 shows the result of the fit in the J/ψ - q^2 bin, for the Run 2p2 dataset. The signal peak is described for the muon mode as a double-sided Crystal Ball function, as defined by Eq. (5.6). For the electron final state, the signal peak is described by the sum of three modulated double-sided Crystal Ball functions, with a sigmoid modulation as in Eq. (5.7), one for each of the bremsstrahlung categories; the parameters of these functions and the relative fractions are determined from simulation. A shift in the mean and a scale factor for the width, for each lepton mode, are included as floating parameters, to allow for any residual data-simulation difference in the signal description.

In the J/ψ - q^2 bin, contamination from combinatorial background is modelled with an exponential function, with its yield and slope floating freely in the fit; additionally, the small contamination from backgrounds with a misidentified hadron, $B^0 \rightarrow J/\psi (\ell^+\ell^-) K^{*0}$ and $\Lambda_b^0 \rightarrow J/\psi (\ell^+\ell^-) pK^-$ decays, as well as from $B^0 \rightarrow J/\psi (\ell^+\ell^-) \phi$ decays, is accounted for. Their mass shapes are fixed from the corresponding simulated samples; their yields, relative to the signal modes, are constrained to their expectations, which are computed from known branching fractions [17] and the reconstruction and selection efficiencies.

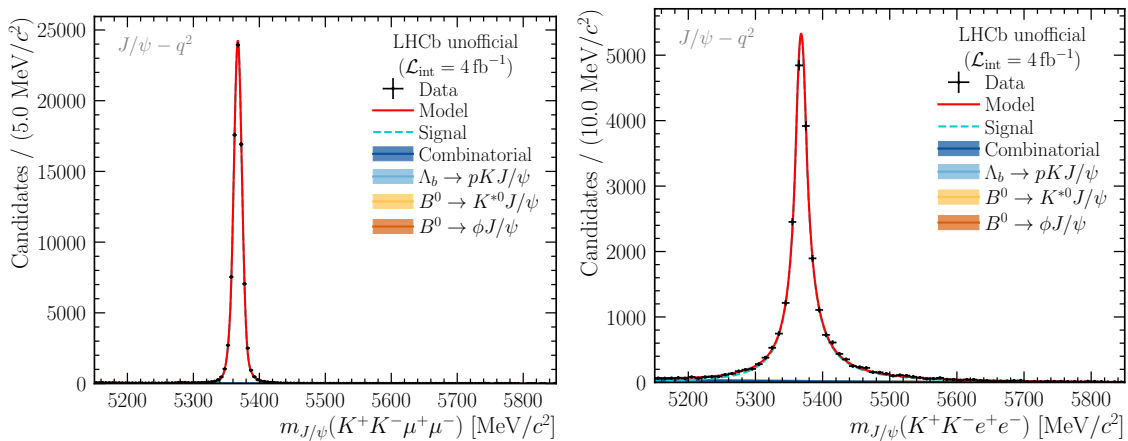


Figure 5.11: Result of maximum likelihood fits to the $m_{J/\psi}(K^+K^-\ell^+\ell^-)$ distributions for $B_s^0 \rightarrow J/\psi (\ell^+\ell^-) \phi$ candidates collected in Run 2p2. On the left, the muon mode; on the right, the electron mode.

Figure 5.12 shows the result of the fit in the $\psi(2S)$ - q^2 bin, for the Run 2p2 dataset. The

shapes describing the muon and electron peaks are the same as for the $B_s^0 \rightarrow J/\psi (\ell^+ \ell^-) \phi$ modes, with the parameters obtained from simulated $B_s^0 \rightarrow \psi(2S) (\ell^+ \ell^-) \phi$ decays. The shift and scale factor parameters for the peaks' mean and width are shared in the fit with the J/ψ - q^2 bin.

Similarly to the fit of the J/ψ - q^2 bin, contamination from combinatorial background is modelled with an exponential curve, with its slope floating freely in the fit; additionally, a small contamination in the electron dataset from $B_s^0 \rightarrow J/\psi (e^+ e^-) \phi$ and partially reconstructed $B_s^0 \rightarrow \phi \psi(2S) (J/\psi (e^+ e^-) \pi^+ \pi^-)$ decays, is accounted for. Their mass shapes are fixed from the corresponding simulated samples; the $B_s^0 \rightarrow \phi \psi(2S) (J/\psi (e^+ e^-) \pi^+ \pi^-)$ yield, relative to the signal mode, is constrained to its expectation, while the $B_s^0 \rightarrow J/\psi (e^+ e^-) \phi$ yield is allowed to float freely in the fit.

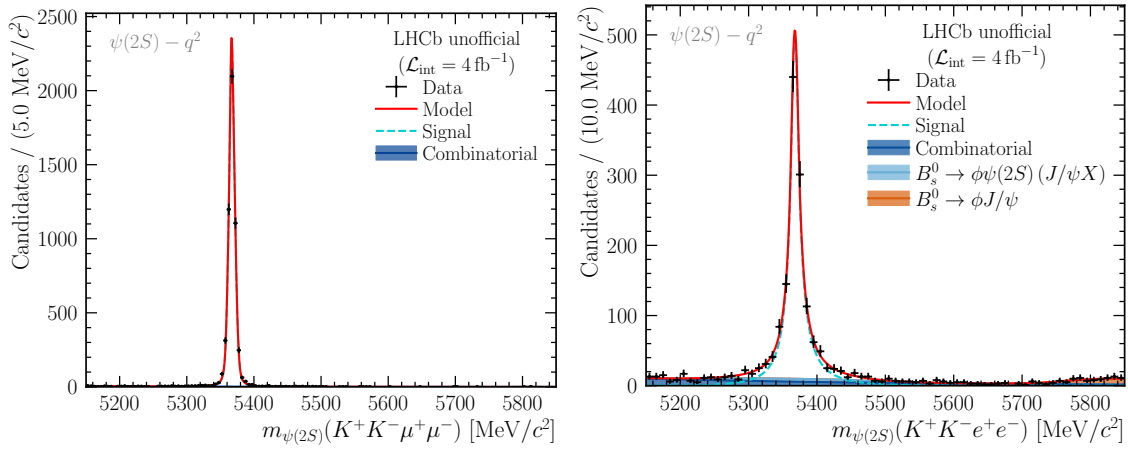


Figure 5.12: Result of maximum likelihood fits to the $m_{\psi(2S)}(K^+ K^- \ell^+ \ell^-)$ distributions for $B_s^0 \rightarrow \psi(2S) (\ell^+ \ell^-) \phi$ candidates collected in Run 2p2.

The measurements of $r_{J/\psi}$ and $R_{\psi(2S)}$ obtained from the result of the fit are

$$r_{J/\psi} = 0.997 \pm 0.013, \quad (5.15)$$

$$R_{\psi(2S)} = 1.010 \pm 0.026, \quad (5.16)$$

where the uncertainties combine both the statistical uncertainty and the systematic uncertainties related to the simulation corrections, described in Sec. 5.6. The systematic uncertainties are included by constraining the efficiencies to the values obtained from simulation, using the correlation matrix as obtained in the previous section in a multidimensional Gaussian constraint. Both ratios are found to be in agreement with the expectation of unity. The results are also stably consistent with unity when fitting the datasets of the three run periods separately, as well as when splitting by trigger category. The control over the simulation corrections that enter the efficiency ratio needed to measure $r_{J/\psi}$ can be further tested by performing the measurement in different regions of the kinematic phase space. The chosen kinematic variables cover ones that are relevant to the correction steps described in Chap. 4, or to the training of the combinatorial BDT. The binning scheme for each variable is opti-

mised to form uniformly populated bins. Figure 5.13 shows the variation of $r_{J/\psi}$ as a function of the opening angle between the electron pair, and the minimum p_T of the electrons, for data collected in the Run 2p2 period. The differential measurements validate the stability of the simulation corrections.

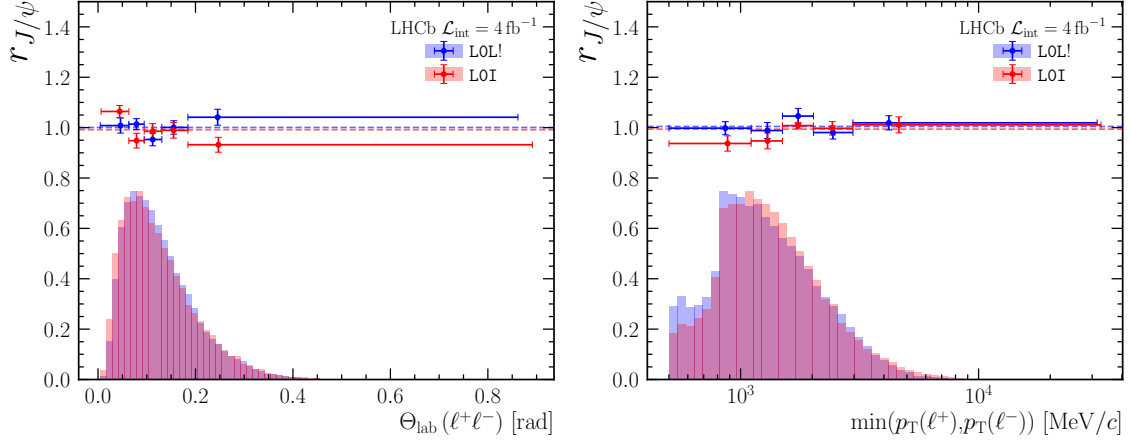


Figure 5.13: Differential measurement of $r_{J/\psi}$ in bins of the opening angle between the two electrons (left), and the minimum p_T of the electrons, in data collected in the 2017–2018 data-taking period. The red and blue markers indicate the measurement of $r_{J/\psi}$ when done using subsets of data corresponding to the LOI and LOL! trigger categories, respectively. The underlying kinematic distributions are also shown as the filled histograms. Figures taken from Ref. [114].

Figure 5.14 shows the measured values of $r_{J/\psi}$ and $R_{\psi(2S)}$ at each stage of application of the simulation corrections. The plots show the impact of the simulation corrections to this measurement, as well as the power of the method of the double ratio: the corrections in $r_{J/\psi}$ compensate for a bias of more than 30% (where no corrections are applied); the value of $R_{\psi(2S)}$ is not sensitive to this and its value is instead varying by at most a few percent, as the effects which the corrections address cancel out in the double ratio.

5.4.2 Pseudoexperiment studies

In high energy physics, Monte Carlo generated pseudoexperiments, commonly referred to as toy datasets, are a very useful tool. Starting from an assumed model for the signal and background components, as well as from the expected yields for each, one can generate a large ensemble of pseudoexperiments. Under the assumption that the model and background composition are an appropriate description of real data, these can be used to test for the presence of biases, assess the sensitivity to the parameters of interest, and to look for other possible pathologies in the chosen fitting strategy.

The toy datasets are generated and fitted with the same model used to fit real collision data, which will be described in more detail in the following section. The yields are generated from Poisson distributions with means equal to the expected yield of each component. The constraints used in the fit are also varied according to a Gaussian distribution of mean and

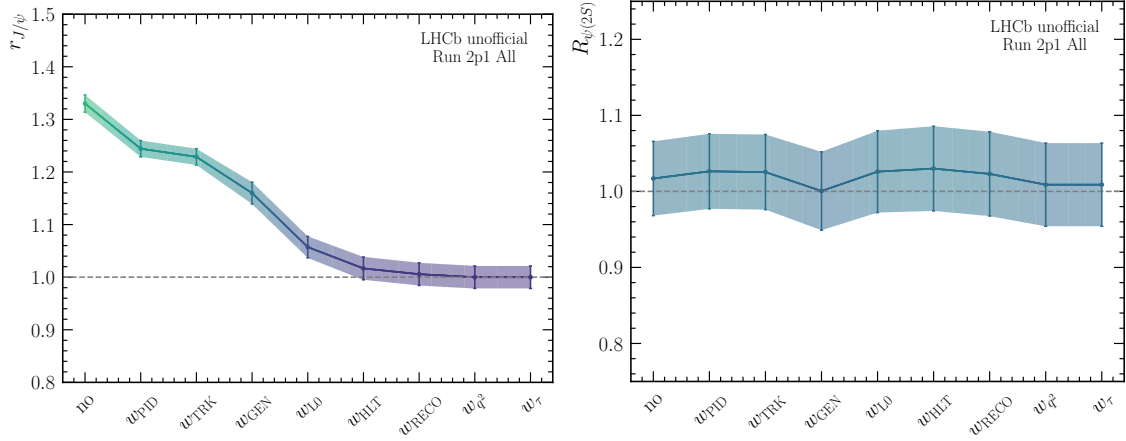


Figure 5.14: Evolution of the measurement of $r_{J/\psi}$ (left) and $R_{\psi(2S)}$ (right) with data collected in the 2015–2016 period, with different corrections w applied to the simulation datasets used to compute the selection efficiencies. The corrections are applied cumulatively from left to right: particle identification (PID), track reconstruction (TRK), B_s^0 kinematics (GEN), trigger (L0 and HLT), vertex reconstruction (RECO), dielectron mass (q^2), and B_s^0 decay model (τ).

width equal to the value and uncertainty of the constraint, respectively.

The statistical properties of the fit strategy and model can be inferred from the distributions, over the ensemble of pseudoexperiments, of the fitted values and uncertainties of the floating parameters. If the size of the generated samples is large enough, by the Central Limit Theorem the distribution of fitted values of a given parameter should follow a Gaussian distribution. The width of the distribution can be used as an estimate for the expected statistical sensitivity to that parameter. Furthermore, it is possible to define a measure of the distance of the fitted values from what is used for the generation of the ensemble [117]. This distance is known as pull, and it is defined for a parameter x as

$$p_i = \frac{x_i - \mu}{\sigma_i}, \quad (5.17)$$

where $x_i \pm \sigma_i$ is the fitted value and uncertainty for the i th toy, and μ is the generating value. If, in the large number limit, x is Gaussian-distributed, then in the absence of biases one can expect the distribution of p_i over the ensemble of pseudoexperiments to also be Gaussian-distributed, with a mean zero and unit width. If the mean of the pull distribution for a given parameter in the fit model deviates from zero, then there are biases present in the fit strategy. If the width is smaller (larger) than unity, then one can infer that the uncertainties estimated in the maximum likelihood fit are being over-(under-)estimated.

For each q^2 bin separately, 500 pseudoexperiments are generated, under the SM assumption of $R_\phi = 1$. For each one, both muon and electron datasets are generated and fitted simultaneously. In the low- q^2 bin, 97.6% of the performed likelihood fits converged properly, indicating a high stability of the setup. The pull distribution is found to agree with a Gaussian distribution of mean zero and unit width. The expected statistical sensitivity, relative

to R_ϕ^{-1} , is 22%.

Similarly, in the central- q^2 bin the fit setup is found to be very stable, with 99% of the performed likelihood fits converging. The pull distribution indicates a small bias on the fitted value of R_ϕ^{-1} , about 13% of the standard error. The width of the pull distribution is found to be compatible with unity. The expected relative statistical sensitivity is 17%.

In the high- q^2 bin two fit setups are tested with independent ensembles of pseudoexperiments, each corresponding to the models introduced in Sec. 5.2 for the combinatorial background: one being the empirical model based on polynomials, and the other the phase-space model. Figure 5.15 shows the comparison of the distribution of the fitted values of R_ϕ^{-1} in the two setups, as well as a comparison of the corresponding pull distributions.

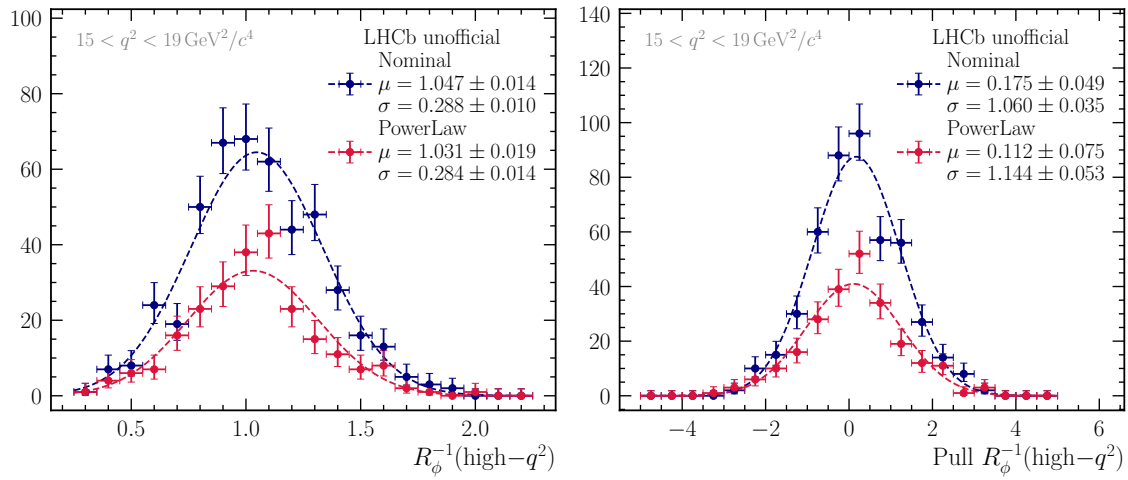


Figure 5.15: Comparison of the distribution of fitted values of R_ϕ^{-1} (left) and pull distribution (right) over an ensemble of 500 pseudoexperiments generated according to the expected signal and background composition for the high- q^2 bin. In blue, the result of pseudoexperiments generated and fitted using the phase-space model for the description of the combinatorial background; in red, the result of pseudoexperiments generated and fitted using the alternative empirical model. Note the different height of the distributions is due to the lower convergence rate of the fit setup with the empirical model.

Because of the larger number of free parameters present in the empirical model, only about 60% of the maximum likelihood fits are found to converge properly. A higher convergence rate is observed for the phase-space model, which only has one floating parameter, of about 95%. A similar sensitivity is achieved of about 28%, and similar biases are observed in the pull distributions. The fit setup with the empirical model features however a width larger than unity, indicating an underestimation of the uncertainty on R_ϕ^{-1} . The stability of the first setup can be improved greatly by introducing constraints on the parameters of the combinatorial model, obtained from fits to the SS datasets. However, the underestimation of the uncertainty on R_ϕ^{-1} remains even after these improvements, and for this reason, the phase-space model is chosen to be the nominal combinatorial model.

The pseudoexperiment studies overall confirm that the stability and reliability of the nominal fit strategy. The residual observed biases will be accounted for as a source of systematic

uncertainty to the final value R_ϕ^{-1} , as explained in Sec. 5.6.

5.5 Mass fits

As introduced in Sec. 5.1, R_ϕ^{-1} is measured by performing unbinned maximum likelihood fits to the invariant $m(K^+K^-\ell^+\ell^-)$ mass distributions for $B_s^0 \rightarrow \phi\mu^+\mu^-$ and $B_s^0 \rightarrow \phi e^+e^-$ candidates. The fit is performed simultaneously to the muon and electron datasets in each q^2 bin, for the three run periods, Run 1, Run 2p1 and Run 2p2; similarly to the fit setup to measure $r_{J/\psi}$ and $R_{\psi(2S)}$, the electron yields are expressed as a composite parameter,

$$N(B_s^0 \rightarrow \phi e^+e^-|q^2) = N(B_s^0 \rightarrow \phi\mu^+\mu^-|q^2) \cdot \frac{\epsilon_{e^+e^-}(q^2)}{\epsilon_{\mu^+\mu^-}(q^2)} \cdot r_{J/\psi}^{-1} \cdot R_\phi^{-1}(q^2), \quad (5.18)$$

where $N(B_s^0 \rightarrow \phi\mu^+\mu^-|q^2)$ and $R_\phi^{-1}(q^2)$ are left floating freely. The single ratio $r_{J/\psi}$ is also floating in the fit, constrained to the value and uncertainty obtained from the fit to $B_s^0 \rightarrow J/\psi(\ell^+\ell^-)\phi$ decays; similarly the efficiencies $\epsilon_{e^+e^-}(q^2), \epsilon_{\mu^+\mu^-}(q^2)$ are constrained to their expectation and their correlations are taken into account. By including the constraints, the statistical uncertainty on these parameters is accounted for in the double ratio. Table 5.4 shows the fit $m(K^+K^-\ell^+\ell^-)$ ranges for each q^2 bin.

Table 5.4: Definitions of the $m(K^+K^-\ell^+\ell^-)$ mass ranges used for the fits to the data in the three q^2 bins.

q^2 bin	$\mu^+\mu^-$	e^+e^-
low- q^2	$m(K^+K^-\mu^+\mu^-) \in [5150, 5850] \text{ MeV}/c^2$	$m(K^+K^-e^+e^-) \in [4600, 6120] \text{ MeV}/c^2$
central- q^2	$m(K^+K^-\mu^+\mu^-) \in [5150, 5850] \text{ MeV}/c^2$	$m(K^+K^-e^+e^-) \in [4800, 6120] \text{ MeV}/c^2$
high- q^2	$m(K^+K^-\mu^+\mu^-) \in [5000, 6000] \text{ MeV}/c^2$	$m(K^+K^-e^+e^-) \in [4900, 6600] \text{ MeV}/c^2$

Figure 5.16 shows the result of the unbinned maximum likelihood fit performed in the low- q^2 bin. The signal component is described by Crystal Ball functions for both the muon

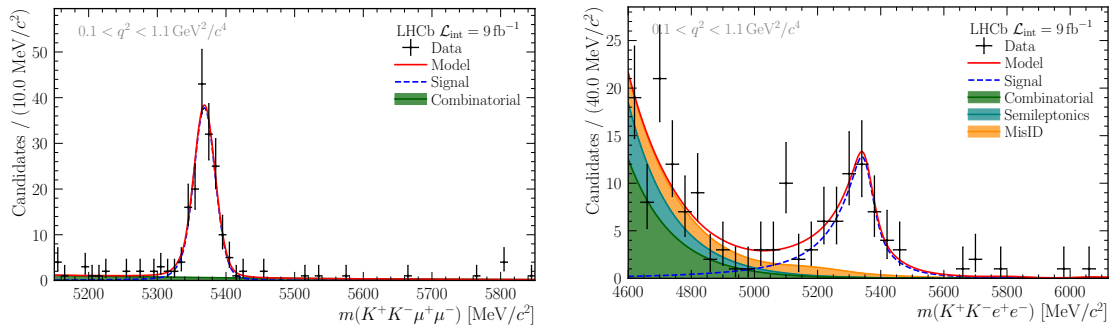


Figure 5.16: Fits to the $B_s^0 \rightarrow \phi\mu^+\mu^-$ and $B_s^0 \rightarrow \phi e^+e^-$ invariant mass distributions in the low- q^2 bin, for data collected by the LHCb detector in the 2011–2018 period. Figures taken from Ref. [114].

and electron modes, with the parameters of each fixed from fits to the respective simulated

datasets. Combinatorial background pollutes both datasets, and is modelled by a falling exponential, with its yields and slope parameters free to float. In the electron dataset, additional contamination from semileptonic decays and decays where one or both electrons have been misidentified is included in the fit. The semileptonic component is modelled with an exponential distribution, its slope parameter fixed from simulation. The yield is floating but constrained to the expected values (see Sec. 4.8.1). The misidentified backgrounds are modelled by a non-parametric kernel density estimation, using the method described in Sec. 4.8.3. Its yield is also floating but constrained to the expected value.

Figure 5.17 shows the result of the unbinned maximum likelihood fit performed in the central- q^2 bin.

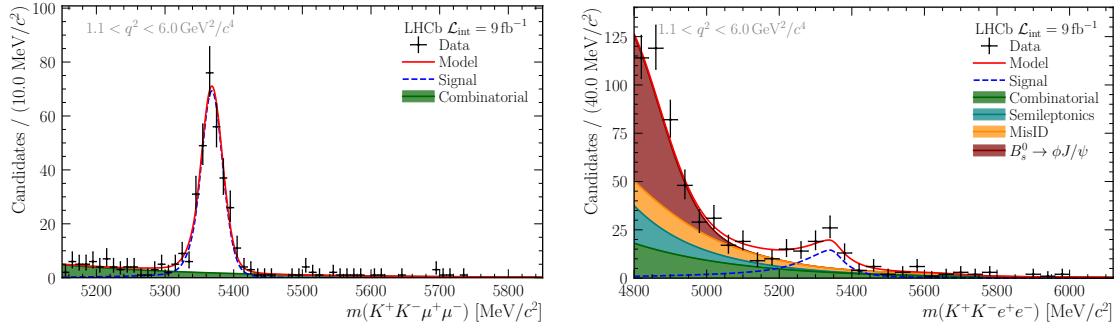


Figure 5.17: Fits to the $B_s^0 \rightarrow \phi \mu^+ \mu^-$ and $B_s^0 \rightarrow \phi e^+ e^-$ invariant mass distributions in the central- q^2 bin, for data collected by the LHCb detector in the 2011–2018 period. Figures taken from Ref. [114].

The signal component is described as before by Crystal Ball functions for both the muon and electron modes, with all parameters fixed from simulation. Combinatorial background is also similarly modelled by falling exponentials as for the previous bin. In the electron dataset, additional contamination from semileptonic decays, leakage $B_s^0 \rightarrow J/\psi (e^+ e^-) \phi$ decays, and misidentified backgrounds is accounted for in the model. The semileptonic component and misidentified backgrounds are treated in the same way as in the low- q^2 bin. The leakage $B_s^0 \rightarrow J/\psi (e^+ e^-) \phi$ background is modelled by a truncated Gaussian function, with its mean and width fixed from simulated decays. Its yield is also floating but constrained to the expected value (see Sec. 4.8.2).

Figure 5.18 shows the result of the unbinned maximum likelihood fit performed in the high- q^2 bin. As explained in Sec. 5.2, the signal component for both lepton modes is described by Crystal Ball functions which are modulated with a sigmoid function, accounting for phase-space effects on the lower mass tail. All parameters are fixed from simulation. Combinatorial background is modelled by the product of the phase-space acceptance and an exponential function, whose slope is free to float. In the electron dataset, additional contamination from leakage $B_s^0 \rightarrow \psi(2S) (e^+ e^-) \phi$ and $B_s^0 \rightarrow J/\psi (e^+ e^-) \phi$ decays, misidentified backgrounds, and partially reconstructed $B_s^0 \rightarrow \psi(2S) (J/\psi (e^+ e^-) X) \phi$ decays is accounted for in the model. The misidentified backgrounds are treated in the same way as in the low- q^2 bin. The leakage $B_s^0 \rightarrow \psi(2S) (e^+ e^-) \phi$ background is modelled by a Johnson- S_U distribution,

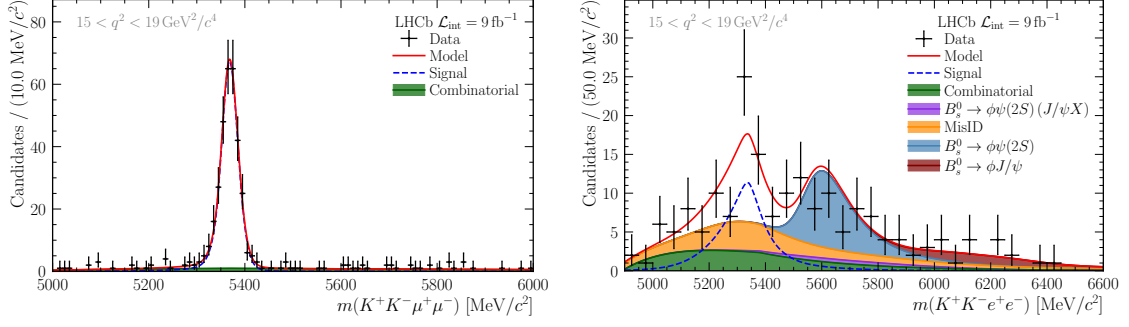


Figure 5.18: Fits to the $B_s^0 \rightarrow \phi \mu^+ \mu^-$ and $B_s^0 \rightarrow \phi e^+ e^-$ invariant mass distributions in the high- q^2 bin, for data collected by the LHCb detector in the 2011–2018 period. Figures taken from Ref. [114].

with its parameters fixed from simulated decays. Leakage $B_s^0 \rightarrow J/\psi (e^+ e^-) \phi$ and partially reconstructed $B_s^0 \rightarrow \psi(2S) (J/\psi (e^+ e^-) X) \phi$ decays are modelled by non-parametric kernel density estimations. All background yields except for the combinatorial background are constrained to their expected values.

Table 5.5 reports the measured yields of the $B_s^0 \rightarrow \phi e^+ e^-$ and $B_s^0 \rightarrow \phi \mu^+ \mu^-$ decays in the three q^2 bins. The statistical significance of the $B_s^0 \rightarrow \phi e^+ e^-$ mode is determined using Wilks’ theorem [118], and is found to be of 6.8, 5.4, and 3.6 standard deviations in the low-, central-, and high- q^2 bins, respectively. This thus constitutes the first observation of this decay mode.

Table 5.5: Measured yields of $B_s^0 \rightarrow \phi \mu^+ \mu^-$ and $B_s^0 \rightarrow \phi e^+ e^-$ decays in the low-, central-, and high- q^2 bins, as obtained from maximum likelihood fits to the invariant mass distributions $m(K^+ K^- \ell^+ \ell^-)$, for data collected by LHCb in the 2011–2018 period.

q^2 bin	$N(B_s^0 \rightarrow \phi \mu^+ \mu^- q^2)$	$N(B_s^0 \rightarrow \phi e^+ e^- q^2)$
low- q^2	159 ± 13	63 ± 9
central- q^2	300 ± 18	119 ± 14
high- q^2	312 ± 18	50 ± 13

5.6 Systematic uncertainties

There are many sources of systematic uncertainty which need to be considered when performing the measurement of R_ϕ^{-1} . Following Eq. (5.18), two main categories of systematic uncertainty can be identified: those related to the computation of the efficiency ratios $\epsilon_{e^+ e^-}(q^2)/\epsilon_{\mu^+ \mu^-}(q^2)$, and those relating to the maximum likelihood fit used to extract R_ϕ^{-1} . Table 5.6 summarises the various sources of systematic uncertainty considered and their sizes.

The systematic uncertainties in the computation of the selection efficiencies are all associated with the correction procedure applied to the simulated datasets, and are thus treated separately for each correction step. The systematic uncertainties are computed on the effi-

Table 5.6: Sources of systematic uncertainty in 10^{-2} on the R_ϕ^{-1} measurement in the low-, central- and high- q^2 bins. The total uncertainty is computed by combining the contributions from individual sources in quadrature. Correlations not taken into account when summing in quadrature are considered to be negligible when comparing to the statistical uncertainty, which is the dominant uncertainty in the measurement. Table taken from Ref. [114].

Source	low- q^2	central- q^2	high- q^2
Efficiency calibration	0.7	0.5	1.2
q^2 smearing	0.6	0.5	0.3
Decay model	0.1	0.3	0.1
$r_{J/\psi}$ variation	1.3	0.6	0.9
Fit bias	< 0.1	0.9	1.4
Normalisation	0.2	0.2	0.2
Signal lineshape	1.7	1.5	3.7
Misidentified bkg.	4.1	4.3	7.4
Combinatorial bkg.	–	–	3.9
Leakage bkg.	–	0.9	2.1
Semileptonic bkg.	1.3	1.2	–
Total	4.9	4.9	9.6

ciencies ratios, assuming that R_ϕ is unity, by either a bootstrapping procedure [116] (in the case of those related to the limited sample sizes) or by varying the method with which the corrections weights are computed.

For the PID correction stage, the sources considered include the limited sample size of the calibration samples used, the choice of kinematic binning, and the uncertainty coming from the *sPlot* method used to compute the kaon and muon efficiencies.

For the corrections to the electron track reconstruction, the main source considered is the limited size of the $B^+ \rightarrow J/\psi (e^+e^-) K^+$ dataset used to compute them. An additional systematic uncertainty, which estimates how much the background contamination biases the correction, is obtained by computing an alternative set of weights using $B^0 \rightarrow J/\psi (e^+e^-) K^{*0}$ decays.

The systematic uncertainty associated to the B_s^0 kinematics and vertex reconstruction corrections, both determined using a BDT reweighting approach, are determined by varying the training sets and input variables to the two BDTs.

The leading uncertainty on the trigger correction step is the limited sample size of the tag and probe sample of $B_s^0 \rightarrow J/\psi (\ell^+\ell^-) \phi$ decays used to extract the weights.

The sources considered for the correction applied to the dielectron mass are the limited sample size of $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ decays used to compute the correction parameters, and the assumption that the corrections are constant across the electron kinematic phase space. The latter is computed by computing and applying the corrections for the three separate bremsstrahlung categories, 0γ , 1γ , and 2γ .

The effect of the application of the B_s^0 lifetime correction weights is estimated conservatively by recomputing the efficiencies without applying the weights, and by taking the

difference with the nominal case.

Finally, a systematic related to any residual miscalibration of the $B_s^0 \rightarrow \phi \ell^+ \ell^-$ samples is obtained by taking into account the deviations from unity observed in the differential measurement of $r_{J/\psi}$. This systematic is estimated following the same procedure as in Ref. [29], using what is known as a flatness parameter, which quantifies the overall deviation from unity in $r_{J/\psi}$ across the different kinematic bins.

The sources of systematic uncertainty related to the fitting procedure are considered separately for each component of the fit. Each uncertainty is extracted by generating a set of pseudoexperiments with a single variation in one of the fit models, and then studying the bias induced on R_ϕ^{-1} . The biases are computed either relative to the generating value or, where an alternative model is used to fit the datasets, by performing the fit twice and computing a “toy-by-toy” difference against the nominal model.

The sources of uncertainty considered for the signal model are the possible mismodelling of the mass peak, and the relative fractions of the bremsstrahlung categories computed from simulation.

Since the combinatorial background shape is free to float in the low- and central- q^2 bins, only systematic uncertainties related to the model at high- q^2 are taken into account. These include the choice of the phase-space acceptance model, as well as the overall choice of model, which is estimated by using the alternative empirical shape, first introduced in Sec. 5.2.

For the leakage and semileptonic components, a systematic uncertainty for each is estimated, to account for any possible mismodelling in simulation. Additionally, the systematic uncertainties from the simulation corrections are propagated to the expectation values of the yields for these backgrounds. These uncertainties are included in the constraints used for the nominal mass fits described in the previous section.

Similarly an alternative parametrisation for the misidentified background component is used to estimate the systematic uncertainty associated to the chosen nominal model. Furthermore, variations are made to the method described in Sec. 4.8.3 used to model this component. These variations include using an alternative binning for the transfer maps, the use of an alternative PID control region, and the inclusion of a contribution from possible misidentified protons. These variations are propagated both to the mass model shape, and to the obtained expected yield. The systematic uncertainties on the latter are included directly in the constraint terms in the nominal mass fits.

A systematic uncertainty associated to the residual biases observed in the pseudoexperiments described in Sec. 5.4 is also included.

The last source of systematic uncertainty considered is related to the determination of the normalisation yields of the control modes $B_s^0 \rightarrow J/\psi (\ell^+ \ell^-) \phi$, which enter the double ratio to determine R_ϕ^{-1} . In particular, the effect of determining $r_{J/\psi}$ with a fit to the J/ψ constrained invariant mass $m_{J/\psi}(K^+ K^- \ell^+ \ell^-)$ is estimated by repeating the measurement fitting for the non constrained mass $m(K^+ K^- \ell^+ \ell^-)$. The difference observed in $r_{J/\psi}$ is then propagated to R_ϕ^{-1} .

The leading source of systematic uncertainty is related to the modelling of the misidenti-

fied background. Additionally, in the high- q^2 bin, the modelling of the signal, combinatorial, and leakage components are also significant sources of uncertainty. As expected, when measuring R_ϕ^{-1} as a double ratio, the systematic uncertainties related to the computed selection efficiencies are found to be subleading. The total systematic uncertainties are significantly smaller than the statistical uncertainty from the maximum likelihood fits.

5.7 Results

From the maximum likelihood fits shown in Sec. 5.5, the measured values are

$$\begin{aligned} R_\phi^{-1}(\text{low-}q^2) &= 1.57^{+0.28}_{-0.25} \pm 0.05 \\ R_\phi^{-1}(\text{central-}q^2) &= 0.91^{+0.20}_{-0.19} \pm 0.05 \\ R_\phi^{-1}(\text{high-}q^2) &= 0.85^{+0.24}_{-0.23} \pm 0.10, \end{aligned} \tag{5.19}$$

where the first uncertainty is statistical, and the second is systematic. The values have been corrected for the small biases observed in the pseudoexperiments studies. Figure 5.19 shows the profile log-likelihood scans of R_ϕ^{-1} in each of the q^2 bins. The values are found to be compatible with their Standard Model expectations.

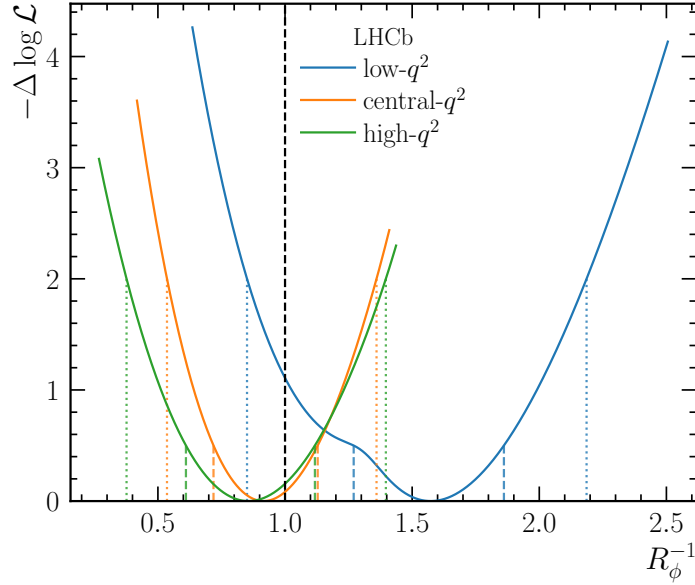


Figure 5.19: Profile log-likelihood scan of R_ϕ^{-1} in the low-, central-, and high- q^2 bins. Systematic uncertainties are included by convolving the likelihood function with a Gaussian of width equal to the total systematic uncertainty. The 1 and 2σ confidence intervals are reported with the dashed and dotted vertical lines, respectively. Figure taken from Ref. [114].

In the scan of the profile log-likelihood in the low- q^2 bin, a secondary local minimum is found for $R_\phi^{-1} \sim 1.3$. The presence of the second minimum is clear in Fig. 5.19, appearing as a significant deviation from the expected parabolic behaviour of the likelihood function near the global minimum. The two minima are characterised by different fitted values for the combinatorial component: at the local minimum point, the combinatorial slope is softer,

while at the global minimum it is much steeper. Given the observed behaviour in the likelihood, the stability of the result is investigated further.

To assess the effect of the combinatorial component on the fitted value of R_ϕ^{-1} , a scan is performed over the different combinatorial BDT selections. Figure 5.20 shows the result of the scan. Both when loosening the BDT requirement, thus allowing for more background pollution, and when tightening it, further suppressing the combinatorial component, the minimum at $R_\phi^{-1} \sim 1.3$ becomes the global minimum. A similar effect happens when the electron PID requirements are tightened, and when performing the fit over a wider mass range. Since the same minimum is found across all the scan points, the possibility of an issue with the description of the background components is excluded. Instead, it is most likely that the presence of two minima is induced from the small sample size of the low- q^2 dataset, which makes it more susceptible to statistical fluctuations.

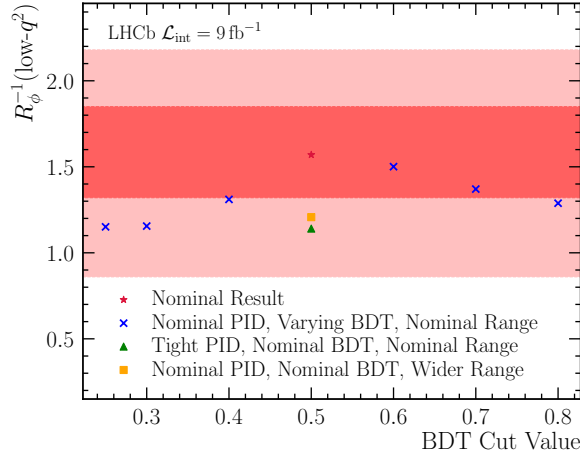


Figure 5.20: Scan of the measured value of R_ϕ^{-1} in the low- q^2 bin for different BDT and electron PID working points. An additional variant of the fit with a range extending to lower values of $m(K^+K^-e^+e^-)$ (“Wider Range”) are also shown. Figure taken from Ref. [114].

Using the measured values for R_ϕ^{-1} and the measured differential branching fraction of the $B_s^0 \rightarrow \phi\mu^+\mu^-$ decay [36], it is possible to obtain the differential branching fraction of the $B_s^0 \rightarrow \phi e^+e^-$ decay. Figure 5.21 reports the branching fraction in bins of q^2 , compared to the Standard Model expectation, which is computed with the FLAVIO package [48].

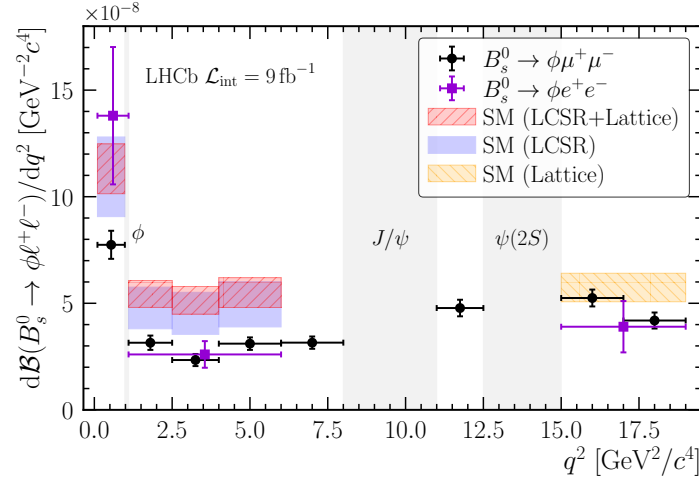


Figure 5.21: Differential branching fraction of the $B_s^0 \rightarrow \phi e^+ e^-$ decay, compared to the $B_s^0 \rightarrow \phi \mu^+ \mu^-$ differential branching fraction [36] and to the Standard Model prediction from the FLAVIO package [48], using form factors calculated with light-cone sum-rule [46] and lattice QCD techniques [119, 120]. Figure taken from Ref. [114].

Chapter 6

Angular analysis of the rare decay

$$B_s^0 \rightarrow \phi e^+ e^-$$

The author's recommendation

Desire, I Want to Turn Into You

by **Caroline Polachek**

2023, Sony Music/The Orchard/Perpetual Novice

The previous chapter the first observation of the decay mode $B_s^0 \rightarrow \phi e^+ e^-$ is reported, in the context of the measurement of the LFU ratio R_ϕ [114]. The measurement is in agreement with the SM at the level of one standard deviation. While LFU may be respected at the branching fraction level, the measurement of the electron angular observables can provide additional inputs to distinguish between NP and unaccounted for SM effects [121]. In this chapter I present the first angular analysis of the decay $B_s^0 \rightarrow \phi e^+ e^-$, using data collected with the LHCb experiment in the 2011–2018 period, with the ultimate goal of a direct comparison with the muon mode [122]. The LHCb collaboration has published an angular analysis of the $B_s^0 \rightarrow \phi \mu^+ \mu^-$ decay [99], benefitting from the efficient muon trigger and an almost background-free environment. In contrast, the unique experimental challenges of electron final states, introduced in Chap. 4, limit the statistical power of the current available dataset for $B_s^0 \rightarrow \phi e^+ e^-$. I will first introduce the measurement strategy, then move onto the modelling of the angular dependence of the selection and reconstruction efficiency. The angular models for the backgrounds first introduced in Chapters 4 and 5 will also be described. Then, the fitting strategy will be validated using the resonant $B_s^0 \rightarrow J/\psi (e^+ e^-) \phi$ and $B_s^0 \rightarrow \psi(2S) (e^+ e^-) \phi$ decays, as well as with simulated $B_s^0 \rightarrow \phi e^+ e^-$ decays and pseudoexperiments. Finally, the chapter will cover the appropriate statistical method used to estimate the confidence intervals of the measured angular observables, the major sources of systematic uncertainty, the final result and comparison with the SM and previous $B_s^0 \rightarrow \phi \mu^+ \mu^-$ measurements.

6.1 Measurement strategy

The analysis follows and builds on the measurement of R_ϕ [114]. In particular, it inherits the same candidate selection and simulation corrections.¹ The $m(K^+K^-e^+e^-)$ models used for the background distributions are also the same, with the exception of the misidentified background, which features an updated model including proton to muon misidentification, as described in Sec. 6.3. Furthermore, the small residual contamination from leakage from partially reconstructed $B_s^0 \rightarrow J/\psi\phi X$ in the high- q^2 bin is not included in the angular fit. This is because of the small sample size of the available simulated sample, which would introduce a significant source of systematic uncertainty associated the angular modelling of this component. Neglecting this background leads to a small systematic uncertainty, which is treated in Sec. 6.6.

The analysis is performed without tagging the flavour of the decaying B_s^0 meson, with the flavour-averaged angular decay rate described by Eq. (2.50). The $B_s^0 \rightarrow \phi e^+e^-$ signal yields, for the full LHCb Run 1-2 dataset, are quite low, as shown in Table 5.5. Combined with the significant background contamination, it is not feasible to fit the full three-dimensional distribution. Instead, projections of the decay rate can be used to reduce the total number of observables, as done in past angular analyses of $B \rightarrow V\mu^+\mu^-$ decays when the yields were comparably low [97, 123]. In particular, by integrating Eq. (2.50) in Φ one gets

$$\frac{1}{d(\Gamma + \bar{\Gamma})/dq^2} \frac{d(\Gamma + \bar{\Gamma})}{dq^2 d\cos\theta_K d\cos\theta_e} = \frac{9}{16} \left[\frac{3}{4} (1 - F_L) \sin^2\theta_K \left(1 + \frac{1}{3} \cos 2\theta_e \right) + F_L \cos^2\theta_K (1 - \cos 2\theta_e) + A_6 \sin^2\theta_K \cos\theta_e \right], \quad (6.1)$$

and similarly, by integrating Eq. (2.50) in $\cos\theta_K$ and $\cos\theta_e$,

$$\frac{1}{d(\Gamma + \bar{\Gamma})/dq^2} \frac{d(\Gamma + \bar{\Gamma})}{dq^2 d\Phi} = \frac{1}{2\pi} [1 + S_3 \cos 2\Phi + A_9 \sin 2\Phi], \quad (6.2)$$

where the cancellation of the angular terms is exact in the integration in the absence of experimental effects. The angle $\cos\theta_\ell$ introduced in Chap. 2 will be referred to as $\cos\theta_e$ from here on.

As explained in Chap. 2, there are effects due to the B_s^0 mixing, and in particular due to the non-zero width difference in the $B_s^0\text{--}\bar{B}_s^0$ system, $\Delta\Gamma_s$, to be taken into account. The analysis is not performed accounting for the B_s^0 decay time, *i.e.* this variable is effectively integrated over. What is actually measured then are the time-integrated angular observables,

¹With the exception of the B_s^0 lifetime correction, which is not applied. This is treated as a source of systematic uncertainty, as discussed in Sec. 6.6.

$$\langle S_i(q^2) \rangle = \frac{\int_0^\infty (J_i(q^2, t) + \bar{J}_i(q^2, t)) \epsilon(t) dt}{\int_0^\infty \frac{d(\Gamma + \bar{\Gamma})}{dq^2 dt} \epsilon(t) dt}, \quad \langle A_i(q^2) \rangle = \frac{\int_0^\infty (J_i(q^2, t) - \bar{J}_i(q^2, t)) \epsilon(t) dt}{\int_0^\infty \frac{d(\Gamma + \bar{\Gamma})}{dq^2 dt} \epsilon(t) dt}, \quad (6.3)$$

where $\epsilon(t)$ is the decay time acceptance, and the expression for the sum of the angular coefficients as a function of the decay time is given in Eq. (2.56). The braces $\langle \cdot \rangle$ are used to indicate the time-integrated angular observables. Finally, the observables are measured in bins of q^2 , so the previous expressions are modified to include a convolution with the resolution function $R(q^2 - q_{\text{true}}^2)$,

$$\langle S_i \rangle, \langle A_i \rangle = \frac{\int_0^\infty \int_{q_1^2}^{q_2^2} \int (J_i(q_{\text{true}}^2, t) \pm \bar{J}_i(q_{\text{true}}^2, t)) \epsilon(q_{\text{true}}^2, t) R(q^2 - q_{\text{true}}^2) dq_{\text{true}}^2 dq^2 dt}{\int_0^\infty \int_{q_1^2}^{q_2^2} \int \frac{d(\Gamma + \bar{\Gamma})}{dq_{\text{true}}^2 dt} \epsilon(q_{\text{true}}^2, t) R(q^2 - q_{\text{true}}^2) dq_{\text{true}}^2 dq^2 dt}, \quad (6.4)$$

where the integral over q_{true}^2 is performed over the kinematically allowed range, and now the observables $\langle S_i \rangle, \langle A_i \rangle$ are defined within a bin $[q_1^2, q_2^2]$. The q^2 resolution is discussed in Sec. 6.2.

The strategy is to measure $\langle F_L \rangle, \langle A_6 \rangle, \langle S_3 \rangle$, and $\langle A_9 \rangle$ in three q^2 bins, defined in Table 4.1 and aligned to the measurement of R_ϕ , by performing separate extended maximum likelihood fits to the $(\cos \theta_K, \cos \theta_e)$ and Φ distributions, each simultaneously with the B -candidate mass distribution. The $m(K^+ K^- e^+ e^-)$ variable is used to discriminate between signal and the various background components. The total angular model used in the fit can be written generally, for each q^2 bin, separating the signal and background contributions, as

$$\begin{aligned} \mathcal{P}(\vec{\Omega}, m | \vec{\Theta}, \vec{\lambda}) &\propto N_{\text{sig}} \cdot \epsilon(\vec{\Omega}) \cdot \frac{1}{d(\Gamma + \bar{\Gamma})/dq^2} \frac{d(\Gamma + \bar{\Gamma})}{dq^2 d\vec{\Omega}}(\vec{\Theta}) \cdot \mathcal{P}_{\text{sig}}(m) \\ &+ \sum_{i \in \{\text{bkgs}\}} N_{\text{bkg}, i} \cdot \mathcal{P}_{\text{bkg}, i}(\vec{\Omega}, m | \vec{\lambda}), \end{aligned} \quad (6.5)$$

where $\vec{\Omega}$ is either $(\cos \theta_K, \cos \theta_e)$ or Φ ; $\vec{\Theta}$ are the angular observables, either $[\langle F_L \rangle, \langle A_6 \rangle]$ or $[\langle S_3 \rangle, \langle A_9 \rangle]$, appearing in the expression for the signal angular decay rate; $\vec{\lambda}$ indicates the background model's nuisance parameters; $\epsilon(\vec{\Omega})$ is the parametrisation of the geometrical, selection, and reconstruction efficiency as a function of the decay angles, known as angular acceptance. The analysis is performed blind, by blinding the values of the angular observables as well as the angular projections.

The $B_s^0 \rightarrow J/\psi(e^+e^-)\phi$ and $B_s^0 \rightarrow \psi(2S)(e^+e^-)\phi$ modes, of which large yields are available, are treated as control channels and used to validate the angular fit strategy. In particular, they serve to validate the approach to model the angular acceptance. The validation is performed by comparing the results of fits to the $B_s^0 \rightarrow J/\psi(e^+e^-)\phi$ and

$B_s^0 \rightarrow \psi(2S) (e^+e^-) \phi$ modes to previously published measurements by LHCb of the equivalent muon decays [105, 124]. The effect of time integration is taken into account when performing the check with the control modes $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ and $B_s^0 \rightarrow \psi(2S) (e^+e^-) \phi$, as the reference LHCb analyses involve time-dependent fits to the angle and decay time distributions.

Table 6.1 reports the predicted Standard Model values for the time- and q^2 -integrated values of F_L and S_3 . These have been computed using the `flavio` package [48]. The table also reports the smeared predictions, folding in the effects related to q^2 migration (see Sec. 5.8). The CP asymmetries A_6 and A_9 are expected to be zero (within the theoretical uncertainties) in all q^2 bins.

Table 6.1: Predicted Standard Model values for the time- and q^2 -integrated values of F_L and S_3 . The first row for each observable reports the truth-level `flavio` predictions [48], while the second row reports the prediction after smearing them with the q^2 migration matrix shown in Fig. 6.9. The values are reported in three bins of q^2 : either true q^2 for the unsmeared prediction, or reconstructed q^2 for the smeared one. The uncertainties on the predictions are obtained by a bootstrapping procedure.

Observable		[0.1, 1.1] GeV ² /c ⁴	[1.1, 6.0] GeV ² /c ⁴	[15, 19] GeV ² /c ⁴
$\langle F_L \rangle$	Truth-level	0.434 ± 0.045	0.798 ± 0.022	0.342 ± 0.023
	Smeared	0.439 ± 0.045	0.766 ± 0.023	0.337 ± 0.025
$\langle S_3 \rangle$	Truth-level	0.030 ± 0.008	-0.008 ± 0.004	-0.207 ± 0.012
	Smeared	0.049 ± 0.008	-0.006 ± 0.005	-0.202 ± 0.011

6.2 Modelling the angular acceptance and resolution

The observed angular and q^2 distributions of the signal decay are expected to be distorted by the effects of the detector’s geometrical acceptance, as well as the trigger, reconstruction, and selection strategy. The combination of all of the above is commonly referred to as angular acceptance, and it needs to be properly taken into account when performing a fit to data.

6.2.1 Model description

The angular acceptance is modelled in four dimensions, namely q^2 , and the three decay angles $\vec{\Omega} = (\cos \theta_K, \cos \theta_e, \Phi)$, using Legendre polynomials and cosine functions. The q^2 variable is linearly transformed so that the function domain is between -1 and 1 ; the transformed variable is referred to as $\tilde{q}^2$,

$$q^2 \rightarrow \tilde{q}^2 = -1 + 2 \frac{q^2 - q_{\min}^2}{q_{\max}^2 - q_{\min}^2}, \quad (6.6)$$

where $q_{\min}^2$ and $q_{\max}^2$ are the minimum and maximum values of the selected q^2 range. The acceptance model is expressed as

$$\epsilon(q^2, \vec{\Omega}) = \sum_{k,l,m,n} C_{klmn} L_k(\tilde{q}^2) L_l(\cos \theta_K) L_m(\cos \theta_e) \cos(n\Phi), \quad (6.7)$$

where $L_i(x)$ are Legendre polynomials of order k , l , and m . Due to the flavour symmetry of the final state, there is no reason to expect the acceptance to be asymmetric in the decay angles,² hence only even orders in $\cos \theta_K$ and $\cos \theta_e$ are included, while the choice to model the Φ dependence with cosine functions automatically ensures such symmetry. For each angular function in Eq. (6.7), one can compute normalisation factors,

$$\int_{-1}^1 L_i(x) L_j(x) dx = \frac{2\delta_{ij}}{2i+1}, \quad (6.8)$$

$$\int_{-\pi}^{\pi} \cos(n\Phi) \cos(m\Phi) d\Phi = \begin{cases} \pi\delta_{nm} & n, m \neq 0 \\ 2\pi & n, m = 0 \end{cases}, \quad (6.9)$$

which also explicitly shows the orthogonality of the chosen angular basis. Thanks to this property, the coefficients can be obtained by taking the average of the term of order (k, l, m, n) over the allowed phase space, weighted by the angular acceptance,

$$C_{0000} = \frac{1}{16\pi} \quad (6.10)$$

$$C_{klmn} = \left(\frac{2k+1}{2}\right) \left(\frac{2l+1}{2}\right) \left(\frac{2n+1}{2}\right) \left(\frac{1}{\pi}\right) \times \int \epsilon(\tilde{q}^2, \vec{\Omega}) \cdot L_k(\tilde{q}^2) L_l(\cos \theta_K) L_m(\cos \theta_e) \cos(n\Phi) d\tilde{q}^2 d\vec{\Omega}, \quad (6.11)$$

where the first expression follows from the normalisation factors for each angular function, which appear explicitly in the second expression. The integral in the previous equation can be computed numerically through Monte Carlo integration, by replacing it with a sum over all selected $B_s^0 \rightarrow \phi e^+ e^-$ candidates of an appropriate simulated sample,

$$C_{klmn} = \frac{1}{N} \sum_{i=1}^N \left[w_i \left(\frac{2k+1}{2}\right) \left(\frac{2l+1}{2}\right) \left(\frac{2n+1}{2}\right) \left(\frac{1}{\pi}\right) \times L_k((\tilde{q}^2)_i) L_l((\cos \theta_K)_i) L_m((\cos \theta_e)_i) \cos(n\Phi_i) \right] \quad (6.12)$$

where w_i are the per-event simulation correction weights which, as described in Sec. 4.6, are normalised to the number of selected events N .

The simulated sample of $B_s^0 \rightarrow \phi e^+ e^-$ candidates used to extract the set of coefficients is generated to follow a uniform distribution in the four-dimensions. This is achieved at

²Any charge-dependent detection asymmetry between the K^+ and K^- mesons or between the e^+ and e^- leptons is expected to have minimal effect.

generation level using a dedicated EVTGEN model [84], known as **FLATQ2**. The sample is then passed through the same reconstruction procedure and selection strategy as data, and corrected through the methods outlined in Chap. 4. With this procedure, all the aforementioned acceptance effects can be captured, with the exception of the $D_s^\pm$ vetoes described in Sec. 4.3. These vetoes affect a region of the four-dimensional phase space where the efficiency varies sharply, and are difficult to model with the slowly-varying polynomials used in this description. The vetoes are removed in the determination of the coefficients C_{klmn} , and treated separately.

The acceptance parametrisation is determined by the maximum order allowed for each dimension. The lowest set of orders which is found to reasonably describe the acceptance shape, both across q^2 and in the three signal q^2 bins, is $k \leq 6$, $l \leq 4$, $m \leq 8$, and $n \leq 2$. A visual check of the quality of the acceptance model is shown in Fig. 6.1, which shows the one-dimensional projections of the acceptance model overlaid on the simulated **FLATQ2** sample.

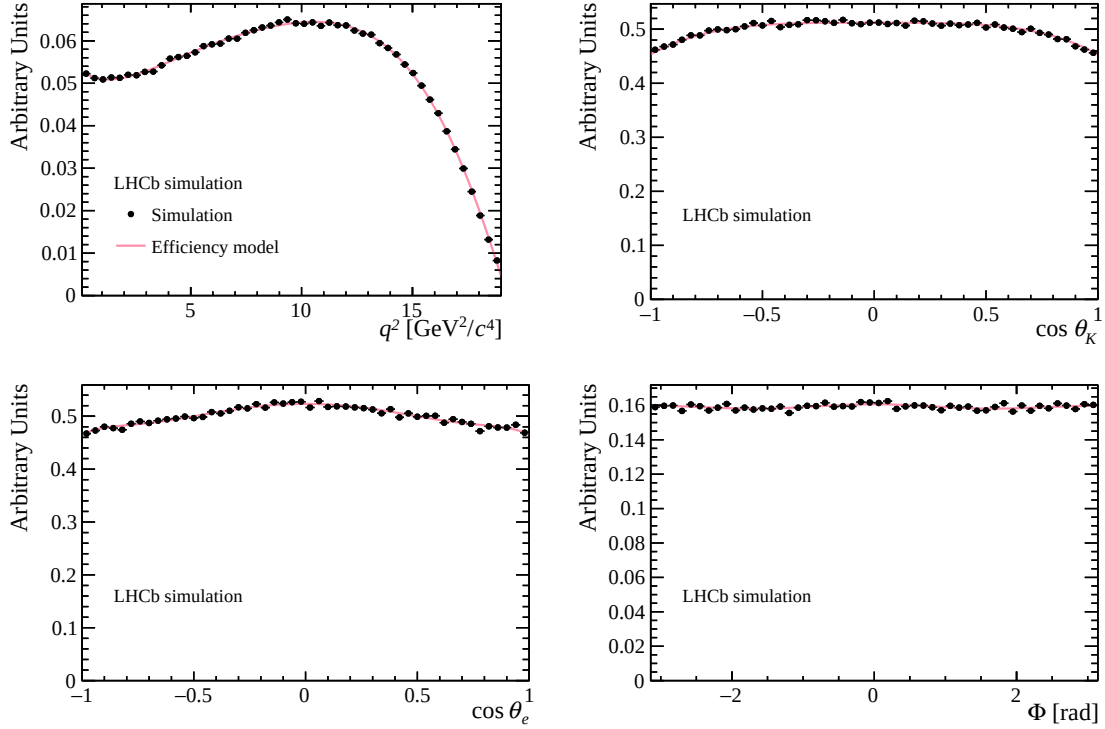


Figure 6.1: Angular and q^2 projections of the **FLATQ2** simulated $B_s^0 \rightarrow \phi e^+ e^-$ sample, after applying the nominal selection and simulation correction weights, and the acceptance model derived from it using Eq. (6.12). From the top left, clockwise: q^2 , $\cos \theta_K$, $\cos \theta_e$, and Φ projections. Figures taken from Ref. [122].

Figure 6.2 shows the one-dimensional projections of the acceptance model in the low-, central-, and high- q^2 bins. The agreement with the projections of the simulated dataset are acceptable for all cases. A further check of the procedure will be performed by applying the acceptance model in the angular fit to the control modes $B_s^0 \rightarrow J/\psi (e^+ e^-) \phi$ and

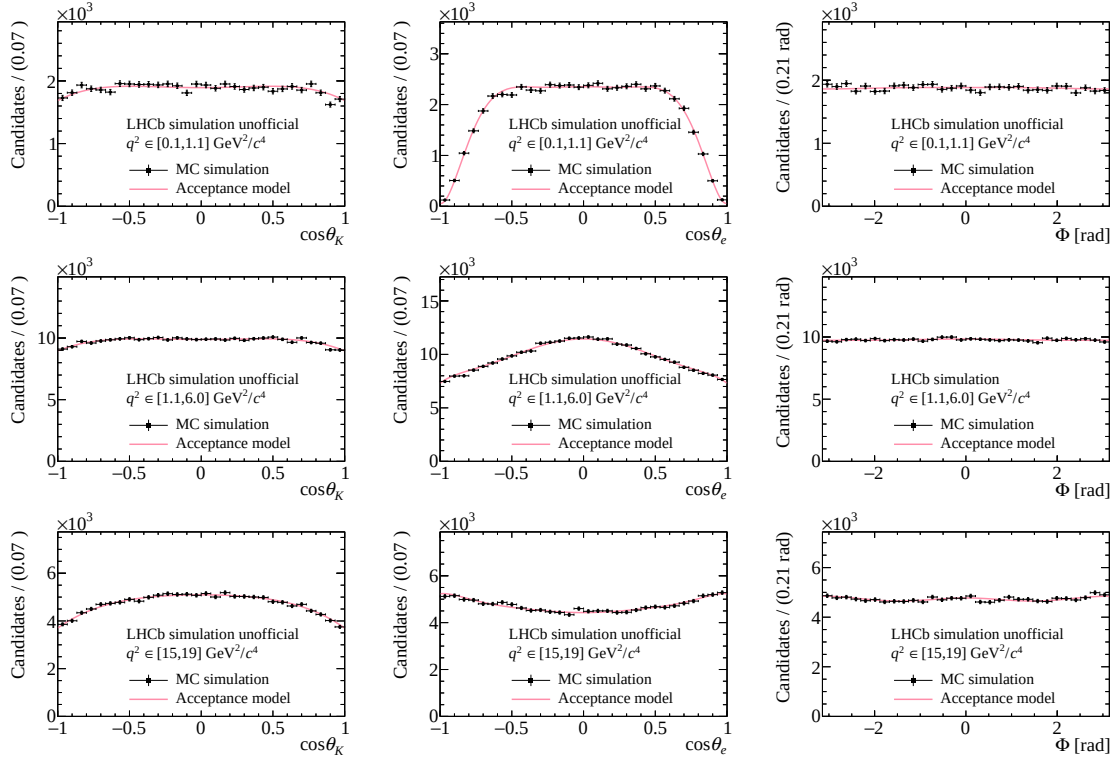


Figure 6.2: Angular projections of the FLATQ2 simulated $B_s^0 \rightarrow \phi e^+ e^-$ sample and the acceptance model in the (top to bottom) low-, central-, and high- q^2 bins. The left, centre, and right columns show the $\cos \theta_K$, $\cos \theta_e$ and Φ projections in each bin, respectively.

$$B_s^0 \rightarrow \psi(2S) (e^+ e^-) \phi.$$

6.2.2 Integration of the angular acceptance

As q^2 is not one of the fitted variables for the angular fit, but rather the fit is done in bins of q^2 , the “marginalisation” of the total acceptance must be handled with care. The choice adopted in this analysis is to evaluate the acceptance function at specific points q_0^2 ,

$$\epsilon(\vec{\Omega}) = \sum_{l,m,n} C_{lmn} L_l(\cos \theta_K) L_m(\cos \theta_e) \cos(n\Phi), \quad \text{where } C_{lmn} = \sum_k C_{klmn} L_k(\tilde{q}_0^2), \quad (6.13)$$

where the points q_0^2 are chosen to be the median points of the SM q^2 distribution, as determined from simulation, in each of the q^2 bins. It is important to note that the acceptance model is affected by the q^2 resolution, so that the acceptance model depends on the underlying q^2 distribution at generator level (here in particular, using FLATQ2 or a physics model). As shown in Sections 6.4 and 6.5 however, the control mode and signal MC fits indicate any bias due to q^2 resolution effects must be small. A systematic uncertainty associated to the arbitrary choice of q^2 point is estimated in Sec. 6.6.

The acceptance model $\epsilon(\cos \theta_K, \cos \theta_e, \Phi)$ can then be naturally included in the angular

fit by simply multiplying it with the signal angular decay rate,

$$\mathcal{P}_{\text{sig}}(\cos \theta_K, \cos \theta_e, \Phi) \propto \epsilon(\cos \theta_K, \cos \theta_e, \Phi) \cdot \frac{1}{d(\Gamma + \bar{\Gamma})/dq^2} \frac{d(\Gamma + \bar{\Gamma})}{dq^2 d\cos \theta_K d\cos \theta_e d\Phi}. \quad (6.14)$$

When performing the fit on a subset of the decay angles, the angular projections of the previous expression have a more complicated form than Eqs. (6.1) and (6.2). It can be shown (see App. B) that the $(\cos \theta_K, \cos \theta_e)$ projection can be written as

$$\begin{aligned} \mathcal{P}_{\text{sig}}(\cos \theta_K, \cos \theta_e) &\propto \frac{1}{2\pi} \epsilon(\cos \theta_K, \cos \theta_e) \cdot \frac{1}{d(\Gamma + \bar{\Gamma})/dq^2} \frac{d(\Gamma + \bar{\Gamma})}{dq^2 d\cos \theta_K d\cos \theta_e} \\ &+ \sum_{\substack{l,m \\ n \neq 0}} C_{lmn} \mathcal{I}_n(\cos \theta_K, \cos \theta_e) L_l(\cos \theta_K) L_m(\cos \theta_e), \end{aligned} \quad (6.15)$$

where the first term is the product of the individual projections of the acceptance model and signal decay rate, and the second term contains the residual terms which may not cancel out in the Φ integration, encapsulated in the integral $\mathcal{I}_n(\cos \theta_K, \cos \theta_e)$. The full expressions for these integrals are derived in App. B. Similarly, performing a double integration over $\cos \theta_K$ and $\cos \theta_e$, it's possible to write the Φ projection as

$$\begin{aligned} \mathcal{P}_{\text{sig}}(\Phi) &\propto \frac{1}{4} \epsilon(\Phi) \cdot \frac{1}{d(\Gamma + \bar{\Gamma})/dq^2} \frac{d(\Gamma + \bar{\Gamma})}{dq^2 d\Phi} \\ &+ \sum_{\substack{n \\ (l,m) \neq (0,0)}} C_{lmn} \mathcal{I}_{lm}(\Phi) \cos(n\Phi). \end{aligned} \quad (6.16)$$

Depending on the order of the angular function, some of the residual integrals vanish. However, some orders result in non-zero integrals, which means that some angular observables, which should vanish in the projection, survive the integration and “leak” as nuisance parameters. For example, the residual integral of order $n = 1$ in the $(\cos \theta_K, \cos \theta_e)$ projection contains the observables S_4 and A_5 . In the Φ case, both $\mathcal{P}_{\text{phys.}}(\Phi)$ and the residual integrals $\mathcal{I}_{lm}(\Phi)$ contain the observables S_3 and A_9 , and not including both terms of Eq. (6.16) could then potentially bias the fit.

The C_{lmn} (with $l, m, n \neq 0$) coefficients and the non-vanishing integrals $\mathcal{I}_n, \mathcal{I}_{lm}$ are numerically small, relative to the main term of the total projections. Indeed, maximum likelihood fits performed on the control modes, as well as on the signal $B_s^0 \rightarrow \phi e^+ e^-$ simulated sample, show that any biases from these terms are negligible when compared to the expected sensitivity of the observables.

6.2.3 Modelling semileptonic vetoes

The veto $m(K^+ K^- \ell_{\text{track}}^\pm) \notin [1920, 2000] \text{ MeV}/c^2$ removes narrow regions of the phase space in q^2 and $\cos \theta_e$, as shown in Fig. 6.3. In reality, because the vetoes are applied using the momentum of the candidate lepton prior to bremsstrahlung recovery (the “track” momentum),

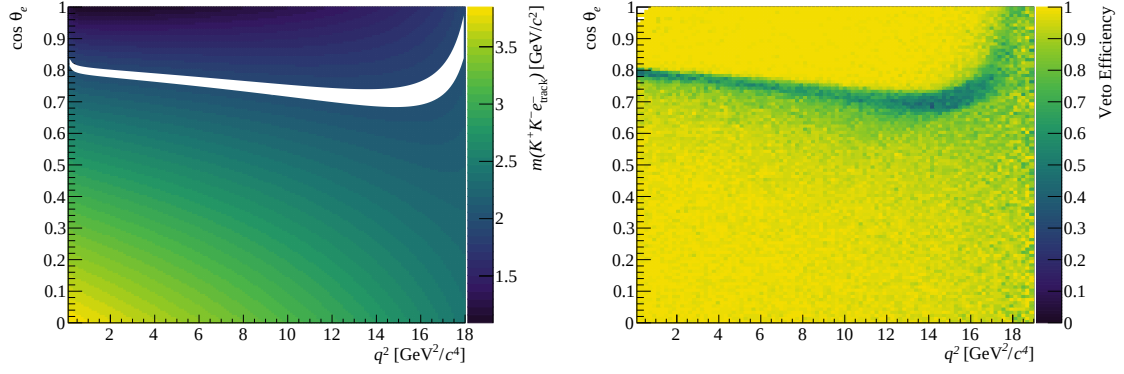


Figure 6.3: Effect of the $D_s^\pm$ veto applied in the nominal selection in the q^2 - $\cos\theta_e$ plane, for $\cos\theta_e > 0$. On the left, an illustration of the value of the 3-body mass $m(K^+ K^- e^-_{\text{track}})$, and the region that is removed when applying the veto. On the right, the efficiency of the veto for fully selected $B_s^0 \rightarrow \phi e^+ e^-$ simulated candidates, generated with the FLATQ2 model. Because of the flavour symmetry of the final state, the veto is symmetric around the axis $\cos\theta_e = 0$.

the regions affected still have a non-negligible efficiency. While the effects of the vetoes are two-dimensional, it is sufficient to model their projection in $\cos\theta_e$, in each q^2 bin, to obtain a good overall description. The dependence of the efficiency of the veto on $\cos\theta_e$ is parametrised as

$$\epsilon_{\text{veto}}(\cos\theta_e) \propto 1 - a \cdot f(|\cos\theta_e| | \mu, \sigma_L, \sigma_R, \alpha_L, \alpha_R, n_L, n_R), \quad (6.17)$$

where f is a double-sided Crystal Ball function (see Eq. (5.6)), which is evaluated for the absolute value of $\cos\theta_e$, to impose the expected symmetry. The values of the function's parameters are determined from a maximum likelihood fit to the same FLATQ2 sample, as used for the rest of the acceptance modelling. In particular, the fit is performed to the histogram ratio $\epsilon(\cos\theta_e) = \frac{h(\cos\theta_e|_{\text{passed}})}{h(\cos\theta_e|_{\text{total}})}$, where $h(\cos\theta_e|_{\text{total}})$ and $h(\cos\theta_e|_{\text{passed}})$ are the distributions for simulated events passing the full selection before and after applying the veto, respectively.

Figure 6.4 shows the best fitted model to the efficiency distribution, for the three q^2 bins. The veto shape then simply multiplies the angular acceptance in Eq. (6.5) to obtain the full signal model.

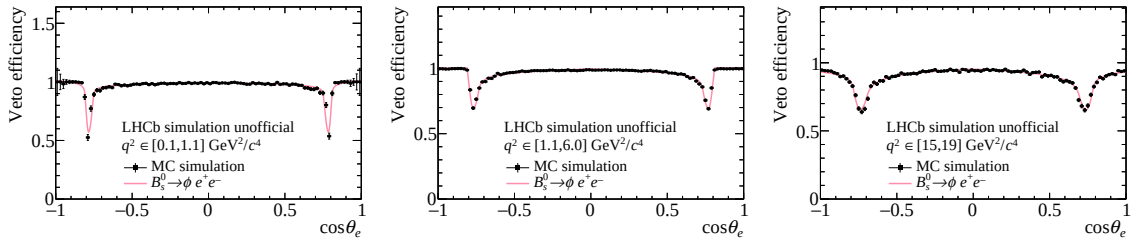


Figure 6.4: Result of the fits to the efficiency distribution in $\cos\theta_e$ of the $D_s^\pm$ vetos. The projections in the low-, central- and high- q^2 bins are shown left to right.

6.2.4 Angular and q^2 resolution

As explained in the previous section, the acceptance model captures the effects which distort the signal angular distribution, from the geometrical acceptance of the detector, to the reconstruction and selection. It is also important to study the angular resolution, as function of the angles and of q^2 . This can be achieved with Monte Carlo simulation of the signal decay $B_s^0 \rightarrow \phi e^+ e^-$, where both reconstructed and truth-level information is accessible, allowing to study the profiles of the variables $(\theta_{K,\text{true}} - \theta_K)$, $(\theta_{e,\text{true}} - \theta_e)$, $(\Phi_{\text{true}} - \Phi)$, where the subscript “true” indicates the generator level variable. From these distributions, the Full Width Half Maximum (FWHM) of the distribution is quoted as a measure of the resolution.

The angular resolution is studied first by subdividing the three q^2 bins: four bins in low- q^2 , five bins in central- q^2 , and four bins in high- q^2 . In each sub-bin, three resolution variables are fitted, and the FWHM computed from the fit result. The distributions are fitted with a Crystal Ball function which is made symmetric in both the core and power-law tails. Example fits are shown in Fig. 6.5. The FWHM is computed analytically, depending on the value of the α parameter:

$$\text{FWHM} = \begin{cases} 2\sqrt{2 \ln 2} \sigma & \alpha > \sqrt{2 \ln 2} \\ 2\sigma \left(2^{\frac{1}{n}} \frac{n}{\alpha} \exp\left(-\frac{\alpha^2}{2n}\right) - \frac{n}{\alpha} + \alpha \right) & \alpha < \sqrt{2 \ln 2} \end{cases} \quad (6.18)$$

The uncertainty on the FWHM is estimated by sampling the covariance matrix obtained from the fit result. Figure 6.6 shows the resolution of the three angles, as a function of q^2 .

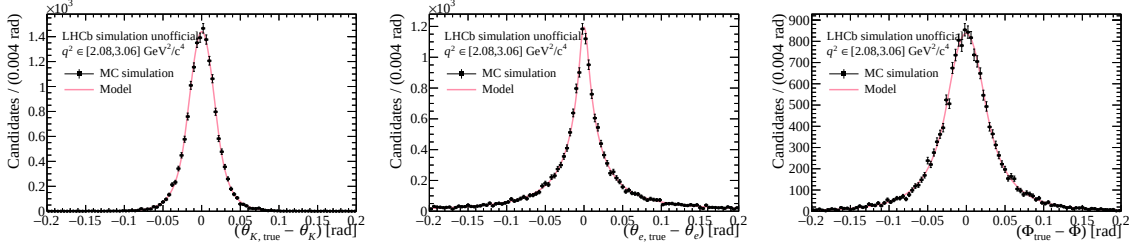


Figure 6.5: Example resolution fits in a sub-bin of the central- q^2 bin, $[2.08, 3.06] \text{ GeV}^2/c^4$, for the (left to right) θ_K , θ_e and Φ decay angles.

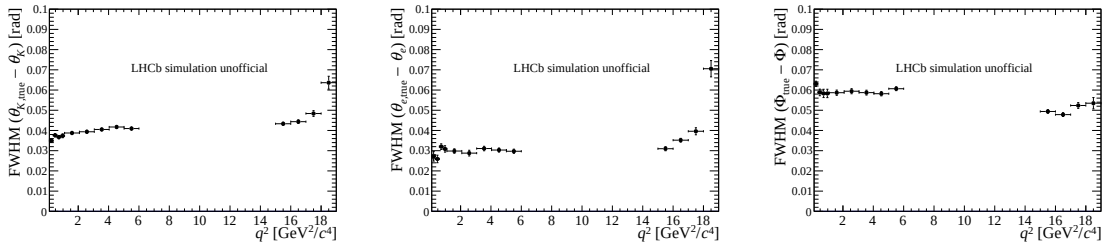


Figure 6.6: Distributions of the angular resolution (measured as FWHM) as determined with simulated $B_s^0 \rightarrow \phi e^+ e^-$ decays, in bins of q^2 for the (left to right) θ_K , θ_e and Φ decay angles.

The angular resolution is also studied as functions of the angles, by subdividing each of the three q^2 bins into eight bins of each angle. Figure 6.7 shows the resolution of the three angles, as a function of the corresponding reconstructed angle, in the three q^2 bins. It is important to note that the bins at the edges (for example at $\theta_e = 0, \pi$), the resolution function must be asymmetric. However, in the bins defined above, this is a small effect. As an example, Fig. 6.8 shows the $(\theta_{e,\text{true}} - \theta_e)$ distribution and fit at low- q^2 in the bin containing the edge $\theta_e = 0$.

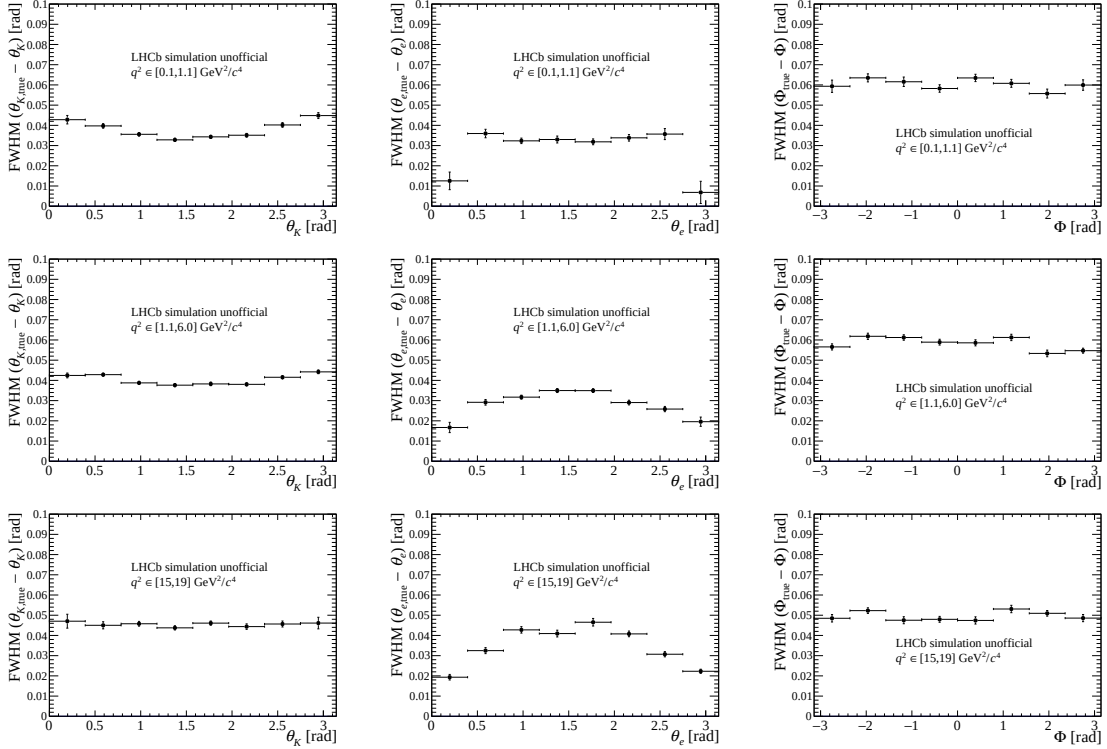


Figure 6.7: Distributions of the angular resolution (measured as FWHM) as determined with simulated $B_s^0 \rightarrow \phi e^+ e^-$ decays, in bins of the corresponding reconstructed angle. From left to right: θ_K , θ_e , and Φ . From top to bottom: low-, central-, and high- q^2 bin.

From these studies, it can be seen that the size of the resolution is generally small, compared to the relevant scale to determine the physics parameters from the angular distributions. Furthermore, as described in Sec. 6.5, validation fits are performed on fully reconstructed simulated signal candidates, using the acceptance model determined in the previous section. No bias is observed compared to the generator level observables. It is thus determined that the resolution is good enough to be neglected in the angular fit model; a systematic uncertainty is assigned to this choice in Sec. 6.6.

The effects of q^2 resolution can be quantified by a “migration” matrix, determined from simulation, which is a function of the truth- and reconstructed-level q^2 . Each bin represents the probability, for a given $q_{\text{true}}^2 \in [q_{\text{true},j}^2, q_{\text{true},j+1}^2]$, to be reconstructed at a value

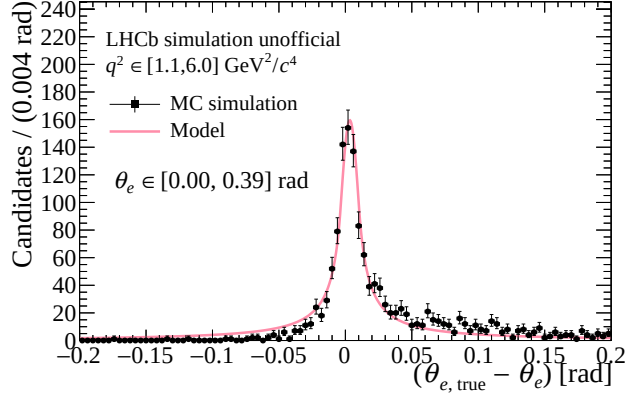


Figure 6.8: Example resolution fit in the low- q^2 bin for $\theta_e \in [0.0, 0.39]$ rad. At the edge, the distribution is more asymmetric, but the discrepancy with the fit model (which is imposed to be symmetric) is small.

$q^2 \in [q_i^2, q_{i+1}^2]$. The matrix is normalised so that

$$\sum_i p(q_i^2 | q_{\text{true},j}^2) = 1. \quad (6.19)$$

Figure 6.9 shows the migration matrix obtained from simulated signal $B_s^0 \rightarrow \phi e^+ e^-$ decays. The truth-level q^2 is determined without including effects from final state radiation, from the difference of the B_s^0 and ϕ four-momenta,

$$q_{\text{true}}^2 = (p_{B_s^0} - p_\phi)^\mu \cdot (p_{B_s^0} - p_\phi)_\mu. \quad (6.20)$$

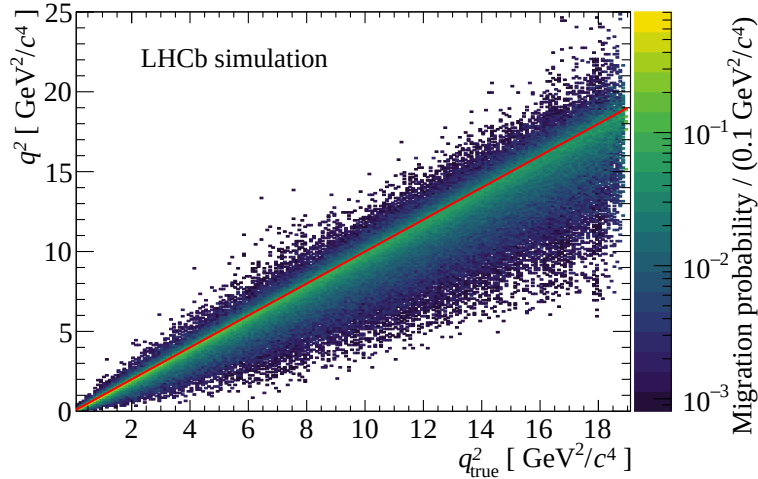


Figure 6.9: Migration probability density matrix $M(q^2 | q_{\text{true}}^2)$ as determined from simulated $B_s^0 \rightarrow \phi e^+ e^-$ decays. The red line corresponds to the diagonal reconstructed $q^2 = q_{\text{true}}^2$. Figure taken from Ref. [122].

6.3 Component modelling

This section covers the modelling of the various components that enter the mass-angle model described by Eq. (6.5). As first mentioned in Sec. 6.1, the signal and background mass models, with the exception of that for the misidentified background, are taken to be the same as for the R_ϕ measurement, and are described in Sec. 5.5.

Allowed signal parameter space

Before describing the background modelling, there is an important point concerning the physical ranges of the angular observables which needs to be considered. There are two ways in which the product of the acceptance projection and angular decay rates in Eqs. (6.1) and (6.2) might return unphysical, *i.e.* negative, values. Firstly, including higher orders of Legendre polynomials might result in regions of the allowed angular phase space where the acceptance is negative. This happens in the $\cos\theta_e$ projection of the low- q^2 region, where the acceptance sharply drops near the edges of the fitted range $\cos\theta_e = \pm 1$ (as shown in Fig. 6.2). The issue is resolved by reducing slightly the fitting range to $\cos\theta_e \in [-0.95, 0.95]$, with minimal impact on the signal efficiency.

Secondly, the angular decay rates have to be positive definite, which reduces the space of possible values which the angular observables can take. For the $(\cos\theta_K, \cos\theta_e)$ projection, the allowed region forms a triangular boundary,

$$\begin{aligned}\langle F_L \rangle &\in [0, 1] \\ \langle A_6 \rangle &\in [-1 + \langle F_L \rangle, 1 - \langle F_L \rangle].\end{aligned}\tag{6.21}$$

Preliminary studies with pseudoexperiments showed that with the available number of signal candidates, the fit is quite unstable, often reaching the boundary and ending in unphysical regions that the minimiser struggles to recover from. To improve the fit stability, an alternative parametrisation is employed instead,

$$\langle A_6 \rangle = (1 - \langle F_L \rangle) \langle A'_6 \rangle,\tag{6.22}$$

with $\langle A'_6 \rangle$ defined in $[-1, 1]$ (this observable is also known in the literature as $A_T^{\mathcal{R}eCP}$ [125]). Figure 6.10 shows a visual comparison of the two physical boundaries. The parametrisation using $\langle A'_6 \rangle$ has a rectangular boundary, which the fit minimiser can handle much better. The trade-off with this more stable parametrisation is a worse sensitivity to $\langle A'_6 \rangle$, in particular when the value of $\langle F_L \rangle$ is close to unity. This effect is described in more detail in the pseudoexperiment studies, in Sec. 6.5.2.

The Φ projection of the angular decay rate also features a non-trivial boundary. Using simple trigonometric relations in Eq. (6.2), imposing that the function is positive-definite

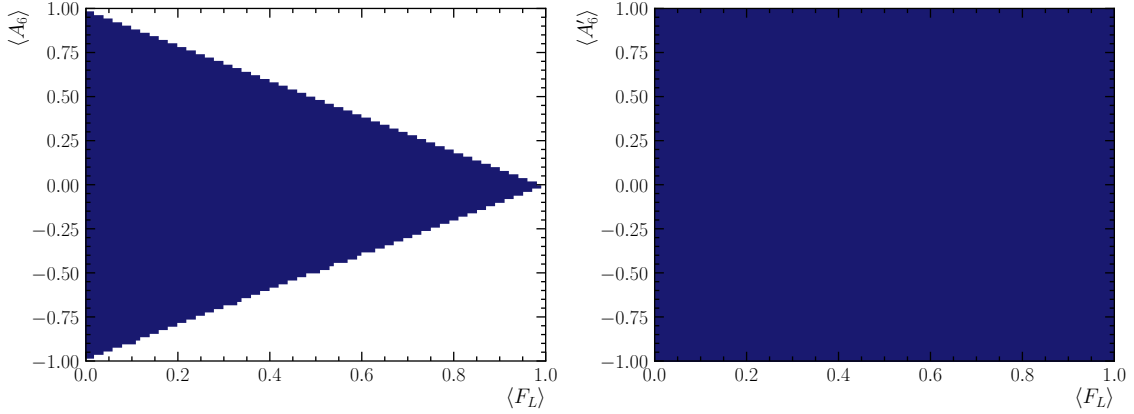


Figure 6.10: Scans of the region (in blue) for which the angular decay rate in $\cos \theta_K$ and $\cos \theta_e$ is positive-definite, when parametrised by the observables $[\langle F_L \rangle, \langle A_6 \rangle]$ (left) or $[\langle F_L \rangle, \langle A'_6 \rangle]$ (right).

results in the inequality

$$\text{sgn}(\langle S_3 \rangle) \sqrt{\langle S_3 \rangle^2 + \langle A_9 \rangle^2} \cos(2\Phi + \arctan(-\langle A_9 \rangle / \langle S_3 \rangle)) \geq -1, \quad (6.23)$$

which is satisfied if the inequality $\sqrt{\langle S_3 \rangle^2 + \langle A_9 \rangle^2} \leq 1$ holds. This means that the physical values of $\langle S_3 \rangle$ and $\langle A_9 \rangle$ lie within the unit circle. Pseudoexperiments show that this boundary does not affect the overall stability of the fit.

Misidentified background

The misidentified background is built upon the “Fail-Pass” method, firstly introduced in Sec. 4.8.3. The method results in a weighted dataset which can be used to estimate both the expected contamination of this component, as well its distributions. The transfer function weights applied however differ than the previous description, to include potential contributions from misidentified protons.

While the proton transfer function is built in the same way as those for pions and kaons, with three misidentified particle species it is not possible to apply the functions using a binary decision as described in Chap. 4. Instead, the three functions T_h are combined in a single transfer function, by weighting each with a probability that the candidate is of the corresponding species,

$$T = \sum_{h=\pi, K, p} p_h T_h. \quad (6.24)$$

The per-event probability p_h is the output of a multivariate classifier trained on the PID control samples described in Sec. 3.4.4. The mass and angle model for the misidentified component is then obtained from the new weighted dataset, and modelled in $m(K^+ K^- e^+ e^-)$, $\cos \theta_K$ and $\cos \theta_e$ with three one-dimensional kernel density estimations. The Φ model is a

second-order polynomial in cosine and sine functions,

$$\mathcal{P}(\Phi) \propto 1 + a_1 \cos \Phi + b_1 \sin \Phi + a_2 \cos 2\Phi + b_2 \sin 2\Phi, \quad (6.25)$$

where the parameters a_1 , b_1 , a_2 and b_2 are obtained from a fit to the weighted control dataset. The overall shape is thus fixed in the angular fit.

Combinatorial background

The angular shape of the combinatorial background is described with even-ordered Chebyshev polynomials (up to fourth order) in $\cos \theta_K$ and $\cos \theta_e$, and a linear sum of cosine and sines in Φ (up to first order),

$$\begin{aligned} \mathcal{P}(\cos \theta_K, \cos \theta_e, \Phi) \propto & [1 + a \cdot (2 \cos^2 \theta_K - 1) + b \cdot (8 \cos^4 \theta_K - 8 \cos^2 \theta_K + 1)] \\ & \cdot [1 + c \cdot (2 \cos^2 \theta_e - 1) + d \cdot (8 \cos^4 \theta_e - 8 \cos^2 \theta_e + 1)] \\ & \cdot [1 + a_1 \cdot \sin \Phi + b_1 \cdot \cos \Phi]. \end{aligned} \quad (6.26)$$

The parameters of the model are determined in the three q^2 bins using SS data as a proxy. The validity of SS data as a proxy for the angular distribution of this background is checked by comparing with OS data in the upper mass sidebands, after relaxing the BDT selection to a minimum value of 0.05. The validation is performed by fitting the angular distributions of SS and OS candidates and directly comparing the fitted values of the model's parameters. Comparisons in the low-, central-, J/ψ , and $\psi(2S)$ - q^2 bins of the six parameters, a , b , c , d , a_1 , and b_1 , are shown in Fig. 6.11. The determined parameters are generally in good agreement.

Such a comparison is not feasible in the high- q^2 bin. Because the misidentified background extends into a region of high $m(K^+K^-e^+e^-)$ mass, and because the large contamination from charmonium leakage appears only in the OS sample, it is not possible to isolate an appropriate sideband for such a comparison. However, given the good agreement with OS data across a large q^2 range, the results from SS data are still used in the full angular fit. To allow for unaccounted differences between the two samples, the combinatorial component is allowed to float in the angular fit, with its angular parameters constrained to the SS results.

Other backgrounds

The leakage and double semileptonic backgrounds are modelled using dedicated simulated samples. The angular shape is given by the product of one-dimensional kernel density estimates for $\cos \theta_K$ and $\cos \theta_e$, and a first-order polynomial of cosine and sine functions in Φ . The parameters are determined from simulation, and fixed in the angular fit.

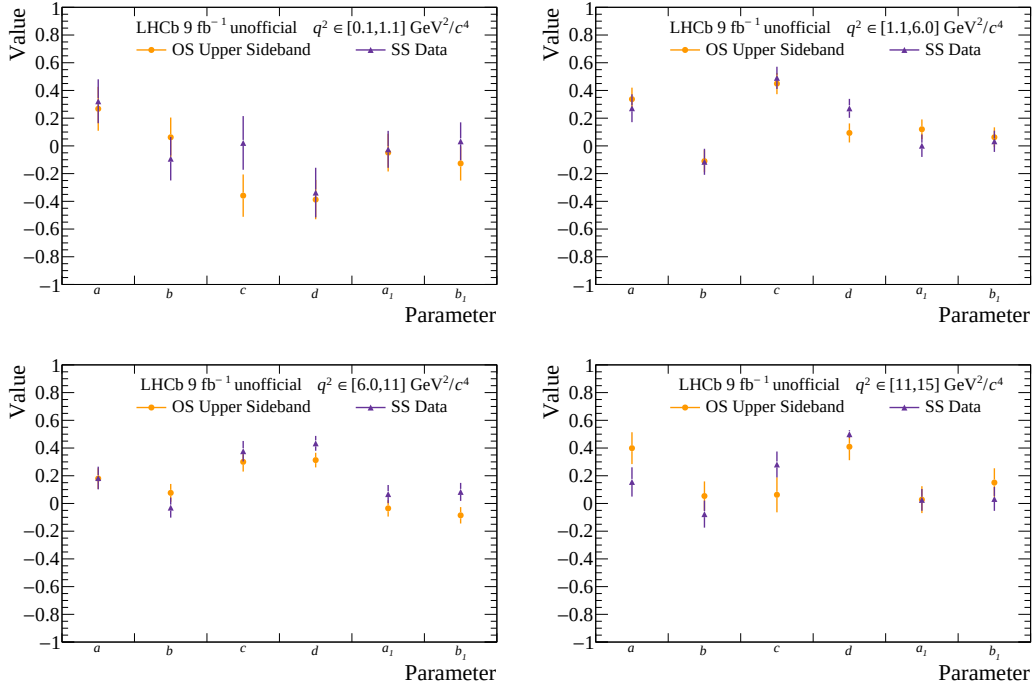


Figure 6.11: Comparison of the combinatorial parameters for the angular shape in $\cos\theta_K$, $\cos\theta_K$ and Φ , as fitted using SS data (purple) and the upper sideband in OS data after relaxing the BDT requirements (orange), in four different q^2 bins. From the top left, clockwise: low-, central, $\psi(2S)$ -, and J/ψ - q^2 .

6.4 Validation of the angular acceptance using control modes

To validate the angular acceptance model as obtained from simulation, fits to the control modes $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ and $B_s^0 \rightarrow \psi(2S) (e^+e^-) \phi$ are performed, and the results are compared to the measurements carried out in dedicated time-dependent amplitude analyses with muons in LHCb [105,124]. The reference values are reported in Table 6.2. The validation using the muon mode is possible as LFU is known to be respected at the per-mille level for J/ψ and $\psi(2S)$ decays to a lepton pair [17].

To achieve the highest signal-to-background ratio, the angular fit is performed in a narrow region around the mass peaks, $m(K^+K^-e^+e^-) \in [5260, 5380] \text{ MeV}/c^2$. This interval was chosen starting from the means and widths obtained from fits to the invariant mass distributions of simulated $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ and $B_s^0 \rightarrow \psi(2S) (e^+e^-) \phi$ candidates. The validation procedure is then as follows, for both resonant modes:

- First, the invariant mass distribution is fitted, to estimate the relative fractions of signal and background in the mass region of interest.
- Then, an angular fit is performed to the upper $m(K^+K^-e^+e^-)$ sidebands, with the BDT selection relaxed, to obtain a constraint on the shape of the background (assuming it does not vary with the invariant mass).
- Finally, the two angular fits are performed, in $(m(K^+K^-e^+e^-), \cos\theta_K, \cos\theta_e)$, and

$(m(K^+K^-e^+e^-), \Phi)$, in the reduced mass region. The fit fractions and mass shapes are fixed, and the background angular shape constrained, using the results of the previous two fits.

Table 6.2: Values for the squared magnitude of the transversity amplitudes A_0 and $A_\perp$, as measured by LHCb for $B_s^0 \rightarrow J/\psi (\mu^+\mu^-) \phi$ decays [105] and $B_s^0 \rightarrow \psi(2S) (\mu^+\mu^-) \phi$ decays [124]. The first reported uncertainty is statistical and the second is systematic.

Parameter	$B_s^0 \rightarrow J/\psi (\mu^+\mu^-) \phi$	$B_s^0 \rightarrow \psi(2S) (\mu^+\mu^-) \phi$
	PRL 132 (2024) 051802	PLB 762 (2016) 253
$ A_\perp ^2$	$0.2463 \pm 0.0023 \pm 0.0024$	$0.264^{+0.024}_{-0.023} \pm 0.002$
$ A_0 ^2$	$0.5179 \pm 0.0017 \pm 0.0032$	$0.422 \pm 0.014 \pm 0.003$

6.4.1 Result for the $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ sample

The fit to determine the relative fraction of signal and background in the B -candidate invariant mass region of $[5260, 5380] \text{ MeV}/c^2$ is performed using the unconstrained B_s^0 candidate mass $m(K^+K^-e^+e^-)$. This is in contrast to the nominal fit for the measurement of $r_{J/\psi}$, shown in Sec. 5.4, which uses the J/ψ -constrained variable $m_{J/\psi}(K^+K^-e^+e^-)$. This is done as for the latter setup the tails of the signal mass shape extend further at high mass, making it difficult to isolate a background-dominated upper mass sideband.

The signal peak is described by a Crystal Ball function, with an asymmetric Gaussian core, and with a sigmoid modulation to both its power law tails. The shape parameters are determined from simulation combining all data-taking conditions. Similarly to the nominal fit in the $r_{J/\psi}$ measurement, the mean and widths of the Gaussian core are allowed to vary, by means of freely floating scale factors. The combinatorial background is described by a falling exponential, with its slope freely floating in the fit. Residual background contamination from $B^0 \rightarrow J/\psi (e^+e^-) K^{*0}$, $B^0 \rightarrow \phi J/\psi (e^+e^-)$ and $\Lambda_b^0 \rightarrow J/\psi (e^+e^-) p K^-$ decays, which is less than 100 events in total in the signal window, is not included. The fit results in approximately 17,600 signal candidates in the reduced mass window, with a contamination of about 260 combinatorial events.

The upper mass sideband is chosen to be $m(K^+K^-e^+e^-) \in [5700, 6800] \text{ MeV}/c^2$. The upper sideband is fitted in the three decay angles $\cos \theta_K$, $\cos \theta_e$, Φ , with the nominal BDT selection threshold relaxed to a value of 0.05, to allow more combinatorial background in. Figure 6.12 shows the result of the maximum likelihood fit to this sample.

Figure 6.13 shows the result of the angular fit, in the signal-enriched region. The fitted values for the four angular observables $\langle F_L \rangle$, $\langle A_6 \rangle$, $\langle S_3 \rangle$ and $\langle A_9 \rangle$ are reported in Table 6.3. As expected, the measured value of $\langle F_L \rangle$ appears to be “biased”, if compared directly to the value of $|A_0|^2$ from the reference measurement in Table 6.2. This is due to the dilution effect introduced by decay-time integration, as first introduced in Chap. 2. The asymmetries $\langle A'_6 \rangle$ and $\langle A_9 \rangle$ are found to be compatible with zero.

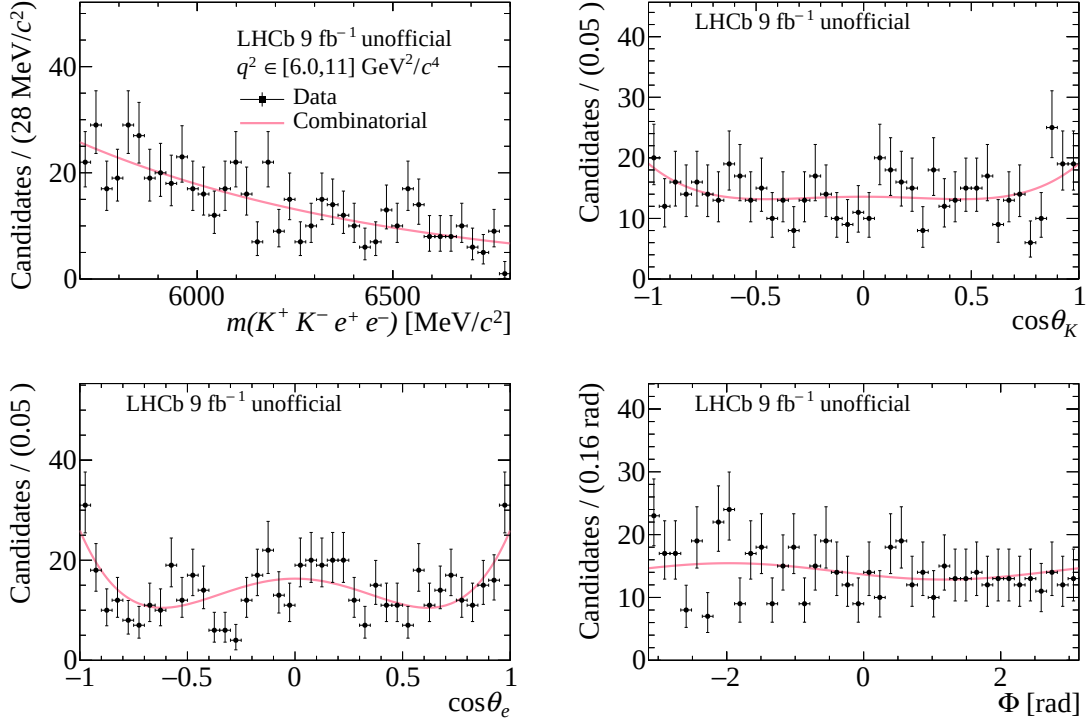


Figure 6.12: Mass and angle projections of the fit to $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ candidates selected in the upper mass sideband $m(K^+K^-e^+e^-) \in [5700, 6800] \text{ MeV}/c^2$, in the J/ψ - q^2 region, with the nominal BDT selection relaxed to 0.05.

Table 6.3: Fitted values of the angular observables $\langle F_L \rangle$, $\langle A'_6 \rangle$, $\langle S_3 \rangle$ and $\langle A_9 \rangle$ from the mass-angle fits for $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ candidates selected in the signal region $m(K^+K^-e^+e^-) \in [5260, 5380] \text{ MeV}/c^2$, with the full Run 1-2 dataset. The reported uncertainties are statistical only.

Angular observable	Fit value
$\langle F_L \rangle$	0.4788 ± 0.0058
$\langle A'_6 \rangle$	0.008 ± 0.014
$\langle S_3 \rangle$	0.035 ± 0.011
$\langle A_9 \rangle$	-0.001 ± 0.011

6.4.2 Result for the $B_s^0 \rightarrow \psi(2S) (e^+e^-) \phi$ sample

Following the same approach as with the previous fit, the B_s^0 invariant mass distribution (without a $\psi(2S)$ mass constraint) is first fit to determine the fraction of background events in the window used for the following angular fit. In the $B_s^0 \rightarrow \psi(2S) (e^+e^-) \phi$ case, there is an additional contamination from leakage $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ candidates. This component is treated in the same way as in the measurement of $R_{\psi(2S)}$, by fixing its mass shape and constraining its yield using simulation. The fit results in approximately 1,200 signal candidates in the reduced mass window, with a contamination of about 96 combinatorial events; the contamination from leakage $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ decays in the reduced window is negligible.

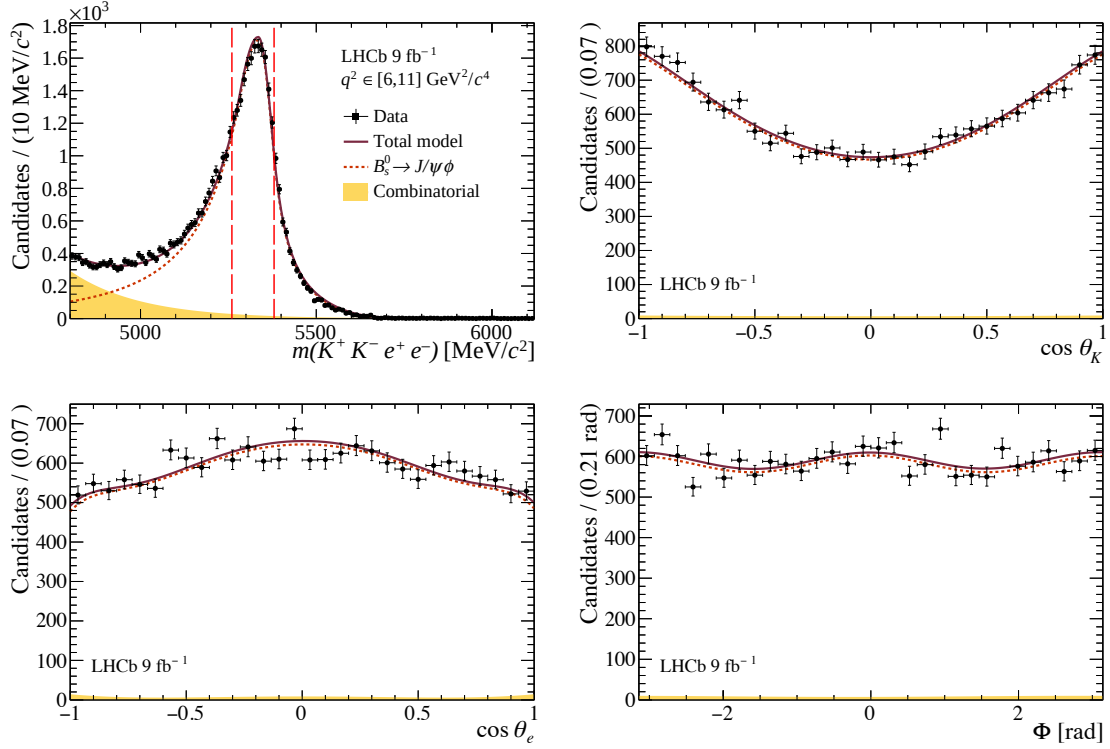


Figure 6.13: Result of the maximum likelihood fits to the $m(K^+K^-e^+e^-)$ -($\cos \theta_K, \cos \theta_e$) and $m(K^+K^-e^+e^-)$ - Φ distributions of Run 1-2 data, selected in the J/ψ - q^2 region. The region used for the angular fit is highlighted by the vertical red dashed lines. Figures taken from Ref. [122].

The upper mass sideband is chosen to be $m(K^+K^-e^+e^-) \in [5600, 6800] \text{ MeV}/c^2$, and is similarly fitted after relaxing the BDT selection to 0.05. In the sideband there is still a significant amount of leakage $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ events, whose angular shape is fixed from simulation. The relative contributions of the leakage component is constrained from simulation. The presence of this background component increases the overall uncertainty on the combinatorial angular shape. Figure 6.14 shows the result of the maximum likelihood fit. The final angular fit results are shown in Fig. 6.15, and the measured values of the angular observables $\langle F_L \rangle$, $\langle A_6 \rangle$, $\langle S_3 \rangle$ and $\langle A_9 \rangle$ are reported in Table 6.4. A similar bias in $\langle F_L \rangle$ is observed, when compared to the value of transversity amplitude $|A_0|^2$ in Table 6.2. The asymmetries $\langle A'_6 \rangle$ and $\langle A'_9 \rangle$ are found to be compatible with zero.

6.4.3 Conversion to transversity amplitudes

The reference analyses [105, 124] do not measure the angular observables, but rather they directly measure the magnitude and phases of the transversity amplitudes A_i , at $t = 0$.³ As shown in Eq. (2.56), where there is a non-vanishing width difference $\Delta\Gamma_s$ between the mass eigenstates, like in the case of the B_s^0 - $\bar{B}_s^0$ system, a dilution effect is present on the time-

³To access the amplitudes, a full tagged time-dependent analysis is needed, as in the reference measurements.

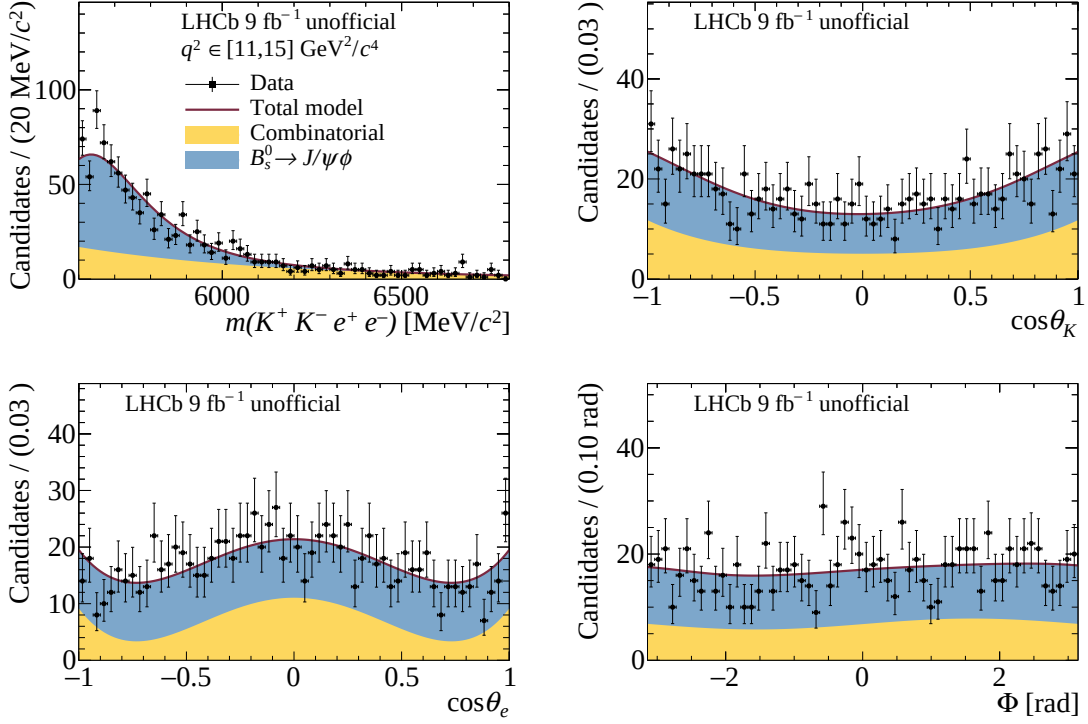


Figure 6.14: Mass and angle projections of the fit to $B_s^0 \rightarrow \psi(2S) (e^+e^-) \phi$ candidates selected in the upper mass sideband $m(K^+K^-e^+e^-) \in [5600, 6800] \text{ MeV}/c^2$, in the $\psi(2S)$ - q^2 region, with the nominal BDT selection relaxed to 0.05.

Table 6.4: Fitted values of the angular observables $\langle F_L \rangle$, $\langle A'_6 \rangle$, $\langle S_3 \rangle$ and $\langle A_9 \rangle$ from the mass-angle fits for $B_s^0 \rightarrow \psi(2S) (e^+e^-) \phi$ candidates selected in the signal region $m(K^+K^-e^+e^-) \in [5260, 5380] \text{ MeV}/c^2$, with the full Run 1-2 dataset. The reported uncertainties are statistical only.

Angular observable	Fit value
$\langle F_L \rangle$	0.398 ± 0.022
$\langle A'_6 \rangle$	-0.024 ± 0.049
$\langle S_3 \rangle$	0.063 ± 0.042
$\langle A_9 \rangle$	-0.024 ± 0.043

integrated angular observables. The equation has to be corrected to include experimental effects, which are encapsulated in the decay-time acceptance $\epsilon(t)$. Then, one can write

$$\int_0^\infty (J_i(t) + \tilde{J}_i(t)) \epsilon(t) dt = \int_0^\infty \epsilon(t) e^{-\Gamma_s t} \cosh \frac{\Delta\Gamma_s t}{2} dt (J_i(t=0) + \tilde{J}_i(t=0) - h_i \mathcal{D}), \quad (6.27)$$

where a new dilution factor is introduced,

$$\mathcal{D} = \int_0^\infty \epsilon(t) e^{-\Gamma_s t} \sinh \left(\frac{\Delta\Gamma_s t}{2} \right) dt / \int_0^\infty \epsilon(t) e^{-\Gamma_s t} \cosh \left(\frac{\Delta\Gamma_s t}{2} \right) dt. \quad (6.28)$$

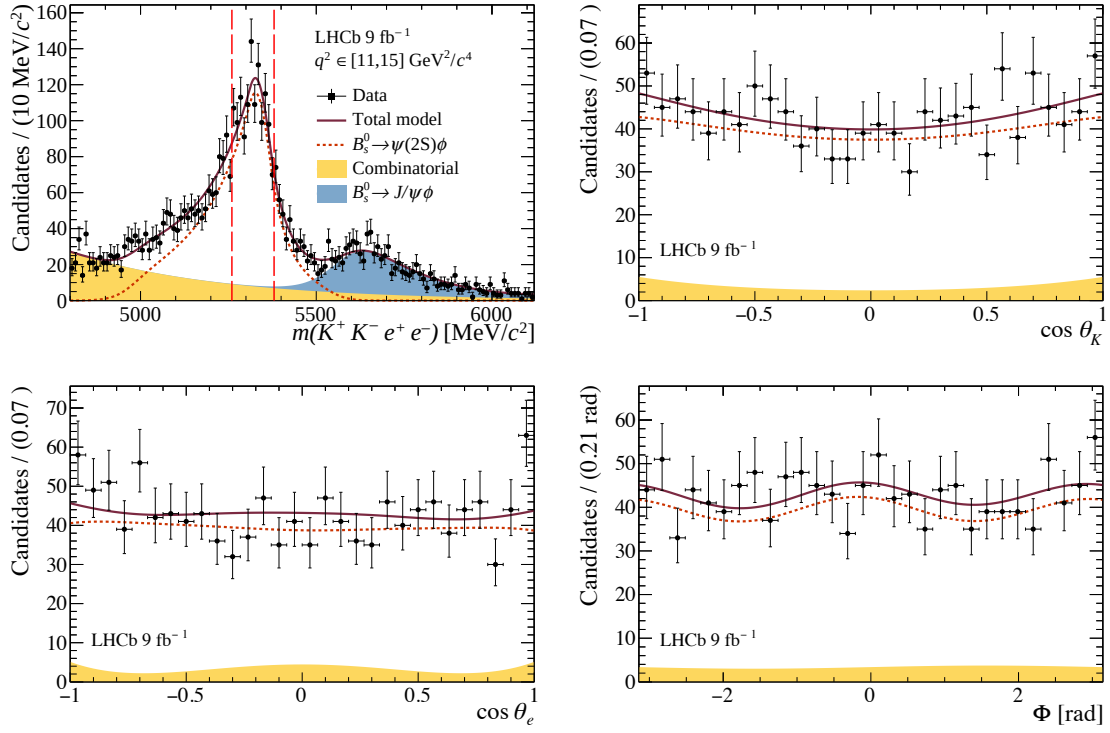


Figure 6.15: Result of the maximum likelihood fits to the $m(K^+K^-e^+e^-)$ -($\cos \theta_K, \cos \theta_e$) and $m(K^+K^-e^+e^-)$ - Φ distributions of Run 1-2 data, selected in the $\psi(2S)$ - q^2 region. The region used for the angular fit is highlighted by the vertical red dashed lines. Figures taken from Ref. [122].

The decay time acceptance is parametrised as

$$\epsilon(t) = \frac{a(t - t_0)^c}{1 + a(t - t_0)^c} (1 + bt). \quad (6.29)$$

Figure 6.16 shows the result of the fit to the FLATQ2 simulated sample of $B_s^0 \rightarrow \phi e^+ e^-$ decays to obtain the parameters of the model that describes $\epsilon(t)$. The candidates are weighted on a per-event basis by $w_i = 1/e^{-\Gamma_s^{\text{Gauss}} t_{\text{true}}}$, where t_{true} is the generator level decay time, and Γ_s^{Gauss} is the value of the B_s^0 decay width used by GAUSS in the generation of the events. From the values of t_0 , a , b and c , the dilution is determined to be $\mathcal{D} = 0.122 \pm 0.015$, where the uncertainty is propagated from the statistical uncertainty on the fitted parameters.

It can be shown that the time-integrated observables $\langle F_L \rangle$ and $\langle S_3 \rangle$ can be written in terms of the transversity amplitudes $|A_0|^2$ and $|A_\perp|^2$ at $t = 0$, and the dilution factor. The expressions, whose detailed derivation is presented in App. C, are

$$\langle F_L \rangle = \frac{|A_0|^2}{1 + \frac{2\mathcal{D}}{1 - \mathcal{D}} |A_\perp|^2}, \quad (6.30)$$

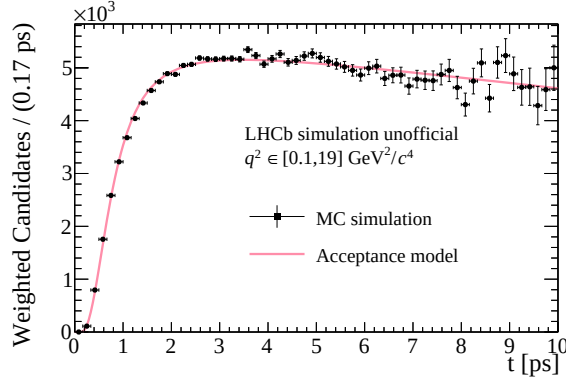


Figure 6.16: Result of the fit to the B_s^0 decay time distribution of the FLATQ2 $B_s^0 \rightarrow \phi e^+ e^-$ sample.

$$\langle S_3 \rangle = \frac{|A_0|^2 + \frac{2}{1-\mathcal{D}}|A_\perp|^2 - 1}{2 \left(1 + \frac{2\mathcal{D}}{1-\mathcal{D}}|A_\perp|^2 \right)}. \quad (6.31)$$

Using these expressions, it is possible to perform a two-dimensional χ^2 fit to obtain the values of $|A_0|^2$ and $|A_\perp|^2$ from the measured values $\langle F_L \rangle$ and $\langle S_3 \rangle$. Figure 6.17 shows the confidence level contours obtained from the χ^2 fit for the control modes $B_s^0 \rightarrow J/\psi (e^+ e^-) \phi$ and $B_s^0 \rightarrow \psi(2S) (e^+ e^-) \phi$, compared to the reference muon measurements. The values for the transversity amplitudes, reported in Tables 6.5 and 6.6, are found to be compatible with the reference values. The values provided in the tables, which do not include any systematic studies, but only the statistical uncertainties, are not meant to be taken as a measurement of these observables. Nonetheless, they provide strong validation of the angular acceptance model derived in Sec. 6.2.

Table 6.5: Comparison of the values of the transversity amplitudes $|A_0|^2$ and $|A_\perp|^2$ as measured in this work with $B_s^0 \rightarrow J/\psi (e^+ e^-) \phi$ candidates, and the reference measurement [105]. The uncertainties reported in this work are only statistical, and the values reported to not constitute a full measurement of the transversity amplitudes.

Parameter	$B_s^0 \rightarrow J/\psi (\mu^+ \mu^-) \phi$	This work, $B_s^0 \rightarrow J/\psi (e^+ e^-) \phi$
$ A_\perp ^2$	$0.2463 \pm 0.0023 \pm 0.0024$	0.239 ± 0.010
$ A_0 ^2$	$0.5179 \pm 0.0017 \pm 0.0032$	0.510 ± 0.006

6.5 Fit strategy validation

6.5.1 Angular fits of signal Monte Carlo

To further validate the angular fit strategy, fits to simulated $B_s^0 \rightarrow \phi e^+ e^-$ candidates are performed. The reconstructed level fits are compared to the values of the angular observables at generator level, which are determined with a separate sample generated without imposing

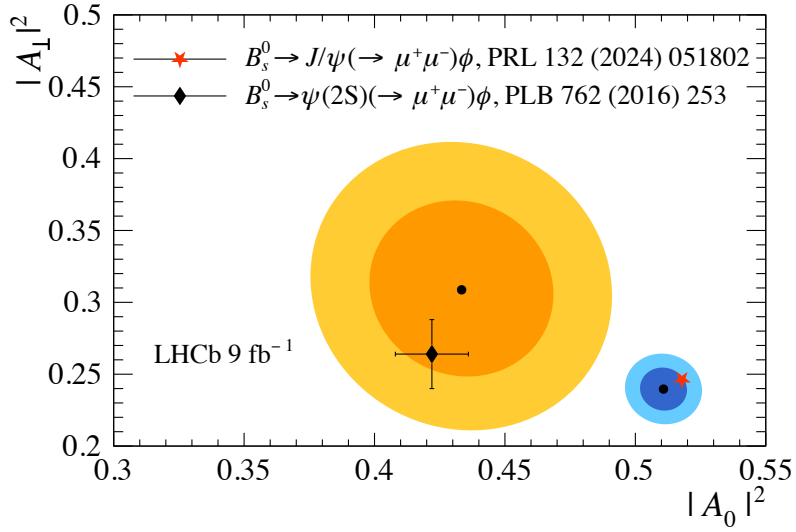


Figure 6.17: Result of the χ^2 fit to the fitted values of $\langle F_L \rangle$ and $\langle S_3 \rangle$, to measure $|A_0|^2$ and $|A_\perp|^2$. In blue, the $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ result, and in orange, the $B_s^0 \rightarrow \psi(2S) (e^+e^-) \phi$ result. The darker and lighter shaded areas correspond to intervals at 68% and 95% confidence level, respectively. The values of the squared transversity amplitudes for the corresponding muon decays from Refs. [105] and [124] are indicated by the markers. The uncertainties on the $B_s^0 \rightarrow J/\psi (\mu^+\mu^-) \phi$ measurement are too small to be visible. Figure taken from Ref. [122].

Table 6.6: Comparison of the values of the transversity amplitudes $|A_0|^2$ and $|A_\perp|^2$ as measured in this work with $B_s^0 \rightarrow \psi(2S) (e^+e^-) \phi$ candidates, and the reference measurement [124]. The uncertainties reported in this work are only statistical, and the values reported to not constitute a full measurement of the transversity amplitudes.

Parameter	$B_s^0 \rightarrow \psi(2S) (\mu^+\mu^-) \phi$	This work, $B_s^0 \rightarrow \psi(2S) (e^+e^-) \phi$
$ A_\perp ^2$	$0.264^{+0.024}_{-0.023} \pm 0.002$	0.309 ± 0.040
$ A_0 ^2$	$0.422 \pm 0.014 \pm 0.003$	0.433 ± 0.023

the geometrical acceptance of the LHCb detector. Figure 6.18 shows the projections of the generator level fits in the three bins. To avoid effects from final state radiation, which disproportionately affect the $\cos\theta_e$ distribution, the generator level values of $\langle F_L \rangle$ are determined by fitting only the $\cos\theta_K$ projection.

Figure 6.19 shows the result of the angular fits to the $(\cos\theta_K, \cos\theta_e)$ and Φ distributions of simulated $B_s^0 \rightarrow \phi e^+e^-$ which have passed the full selection. The results of the fits are reported in Table 6.7. Since the B_s^0 lifetime weights are not applied to the dataset, $\Delta\Gamma_s = 0$ and there are no dilution effects to be taken into account. Only the observable $\langle S_3 \rangle$ in the high- q^2 bin shows a significant bias of 0.022, or 3.1 times the statistical uncertainty from the simulated sample size, which can be attributed to the acceptance modelling, which varies rapidly with q^2 in this bin. The bias is nonetheless one order of magnitude smaller than the expected sensitivity for this observable, which is $\mathcal{O}(0.2)$. Furthermore, it is comparable in size to the estimated size of the systematic uncertainty on $\langle S_3 \rangle$ associated to the acceptance

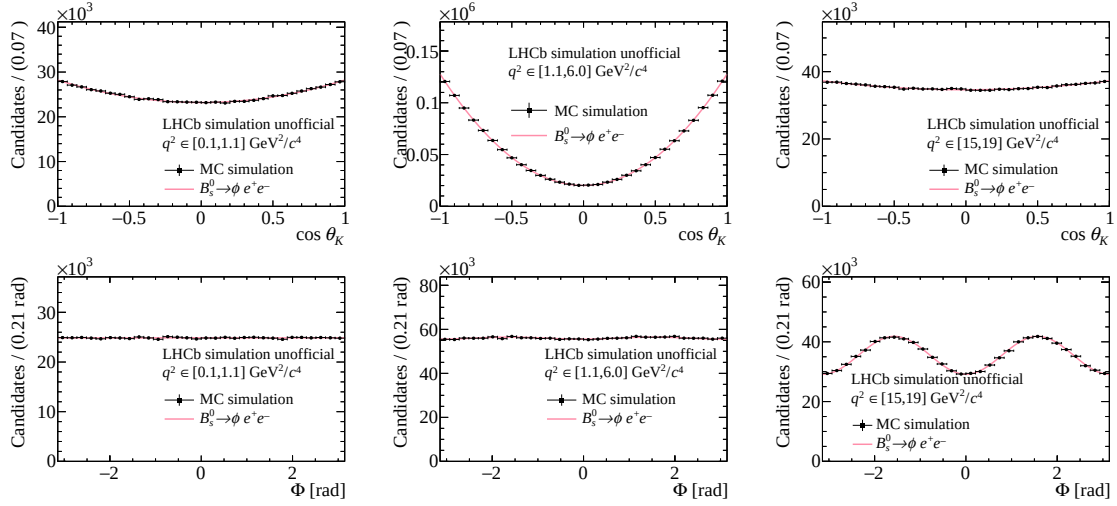


Figure 6.18: Result of the generator level angular fits to simulated $B_s^0 \rightarrow \phi e^+ e^-$ candidates in $\cos \theta_K$ (top row) and Φ (bottom row) for simulated $B_s^0 \rightarrow \phi e^+ e^-$ events selected in the low- (left), central- (middle), and high- q^2 (right) region.

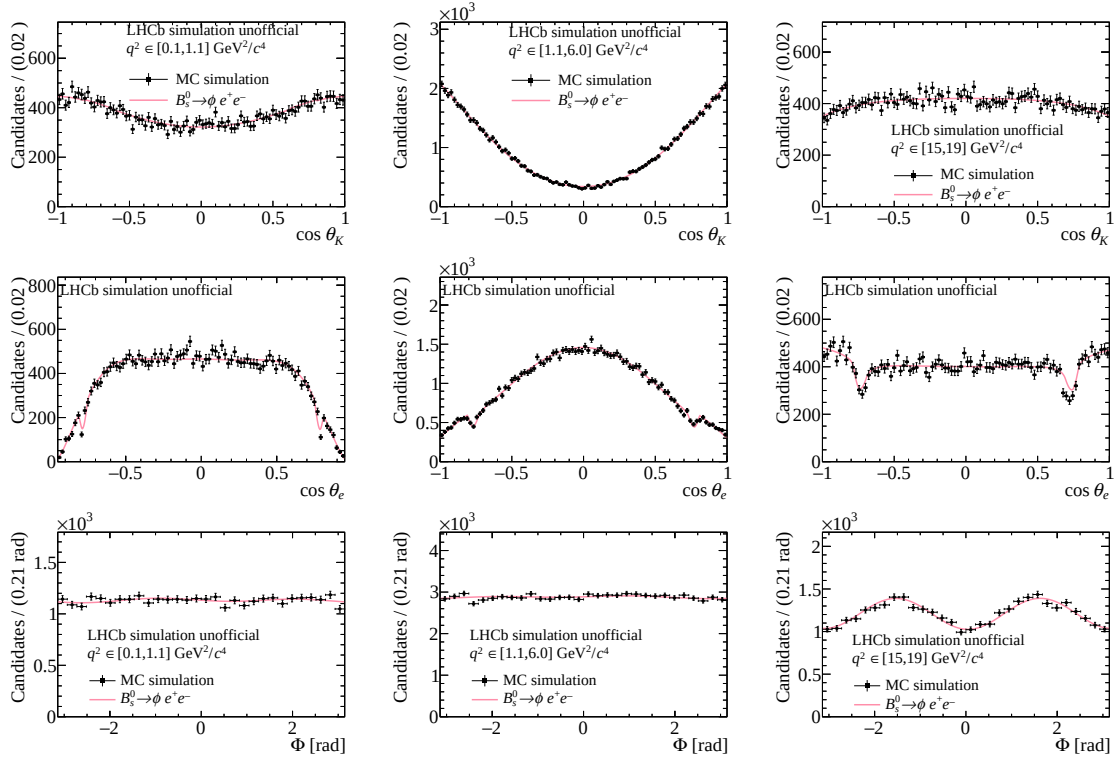


Figure 6.19: Result of the reconstruction level angular fits to simulated $B_s^0 \rightarrow \phi e^+ e^-$ candidates in $(\cos \theta_K, \cos \theta_e)$ (top and middle row) and Φ (bottom row) for simulated $B_s^0 \rightarrow \phi e^+ e^-$ events selected in the low- (left), central- (middle), and high- q^2 (right) region.

model, which is 0.016 (see Sec. 6.6). All other angular observables are found to be unbiased with respect to the generator-level values, within the uncertainty of the simulated sample size.

Table 6.7: Comparison of the generator and reconstruction level angular observables as fitted with simulated $B_s^0 \rightarrow \phi e^+ e^-$ candidates. At generator level, because of the very large sample size available, the statistical uncertainty is negligible. The generator-level value of the asymmetries $\langle A'_6 \rangle$ and $\langle A_9 \rangle$ is undistinguishable from zero.

q^2	Observable	Generator Level	Reconstructed Level
low- q^2	$\langle F_L \rangle$	0.378	0.376 ± 0.004
	$\langle A'_6 \rangle$	0.0	-0.015 ± 0.011
	$\langle S_3 \rangle$	0.0005	-0.0017 ± 0.0077
	$\langle A_9 \rangle$	0.0	-0.0013 ± 0.0077
central- q^2	$\langle F_L \rangle$	0.760	0.758 ± 0.002
	$\langle A'_6 \rangle$	0.0	-0.010 ± 0.011
	$\langle S_3 \rangle$	-0.0081	-0.0039 ± 0.0048
	$\langle A_9 \rangle$	0.0	0.0053 ± 0.0048
high- q^2	$\langle F_L \rangle$	0.350	0.353 ± 0.004
	$\langle A'_6 \rangle$	0.0	-0.010 ± 0.009
	$\langle S_3 \rangle$	-0.179	-0.157 ± 0.007
	$\langle A_9 \rangle$	0.0	-0.014 ± 0.007

6.5.2 Fit validation with pseudoexperiments

Similarly to the procedure for the measurement of R_ϕ , discussed in the previous chapter, pseudoexperiments are generated to investigate the statistical power of the nominal fit strategy, as well as assessing its stability and coverage. First, fits are performed to the invariant mass distributions in the low-, central-, and high- q^2 bins, to determine the yields of each component.⁴ Then, one thousand pseudodatasets are generated per q^2 bin, by generating events following the mass and angle models which have been described above, and varying the number of candidates in each dataset within the Poisson expectation of the known yields. The constraints used in the fit are also varied, each having a unique expectation value for each pseudoexperiment, from a Gaussian with mean and width set from the constraint itself. Where appropriate, the constraints are treated with a multivariate Gaussian. The signal component is generated with values for the angular observables, $\langle S_i \rangle$ and $\langle A_i \rangle$, as determined experimentally by the LHCb collaboration using $B_s^0 \rightarrow \phi \mu^+ \mu^-$ decays [99]. The values used in the generation are reported in Table 6.8.

The datasets are generated using the three-dimensional decay rate and angular acceptance model, and then the mass and angular projections are fitted with the same dataset following the nominal fit strategy outlined in Sec. 6.1. For ease of reading, the $m(K^+ K^- e^+ e^-)$ variable is omitted in this section when referring to the two fit configurations, which will only be referred to by the angular projections fitted. The yields of the leakage, misidentified, and semileptonic backgrounds are constrained to their expectation, as described in the previous section. Their mass and angle shapes are completely fixed in the fit. The combinatorial yield

⁴While the yields have been determined in Chap. 5, they are updated to include effects from the different model for the misidentification background.

Table 6.8: Values of the angular observables used to generate the pseudoexperiments in the three q^2 bins, taken from the most recent $B_s^0 \rightarrow \phi\mu^+\mu^-$ published angular analysis by LHCb [99].

Observable	low- q^2	central- q^2	high- q^2
$\langle F_L \rangle$	0.254	0.715	0.359
$\langle S_3 \rangle$	-0.004	-0.083	-0.247
$\langle S_4 \rangle$	0.213	-0.155	-0.208
$\langle A_5 \rangle$	0.043	-0.007	-0.025
$\langle A_6 \rangle$	0.091	0.008	-0.015
$\langle S_7 \rangle$	-0.178	0.020	0.003
$\langle A_8 \rangle$	-0.007	0.016	0.072
$\langle A_9 \rangle$	-0.030	0.009	0.021

is allowed to float freely, as is its exponential parameter for the invariant mass model. The angular model is left floating but its parameters are constrained to the result of the fits to SS data described in Sec. 6.3.

The estimated sensitivities are, in the low- q^2 bin, $\sigma_{\langle F_L \rangle} = 0.11$, $\sigma_{\langle A'_6 \rangle} = 0.33$, $\sigma_{\langle S_3 \rangle} = 0.26$, and $\sigma_{\langle A_9 \rangle} = 0.24$. The pull distributions show good coverage, and the bias is consistent with zero. Only about 2% of fits end up at the physical boundary for the $(\cos\theta_K, \cos\theta_e)$ fits. The observed failure rate for the $(\cos\theta_K, \cos\theta_e)$ fits is 3‰. No fit failures are observed for the Φ fits.

The distributions of the fitted values and pulls for the four angular observables in the central- q^2 bin are illustrated in Fig. 6.20. The estimated sensitivities are $\sigma_{\langle F_L \rangle} = 0.12$, $\sigma_{\langle A'_6 \rangle} = 0.57$, $\sigma_{\langle S_3 \rangle} = 0.24$, and $\sigma_{\langle A_9 \rangle} = 0.25$. The Φ fit shows no bias and good coverage, with no fit failures. A striking difference with the previous q^2 bin is seen in the distributions for $\langle A'_6 \rangle$, which have a much worse sensitivity and highly non-Gaussian pulls. This effect is expected because of the transformation which defines $\langle A_6 \rangle$ as a function of $\langle A'_6 \rangle$ and $\langle F_L \rangle$. When $\langle F_L \rangle$ is close to unity, most of the sensitivity to $\langle A'_6 \rangle$ is lost, and the likelihood becomes highly non-Gaussian around the minimum. Indeed, around 200 pseudoexperiments, out of the 1000 generated, converge with a value of $\langle A'_6 \rangle$ at its physical boundary. The trade-off in sensitivity results in a gain in fit stability, and the observed failure rate in the $(\cos\theta_K, \cos\theta_e)$ fits is only 8‰. The Feldman–Cousins method, described in App. D, will be used to recover correct coverage intervals, properly accounting for the presence of the physical boundary.

In the high- q^2 bin, the estimated sensitivities are $\sigma_{\langle F_L \rangle} = 0.13$, $\sigma_{\langle A'_6 \rangle} = 0.29$, $\sigma_{\langle S_3 \rangle} = 0.26$, and $\sigma_{\langle A_9 \rangle} = 0.25$. The pull distributions show good coverage, and the bias is consistent with zero. No fits are observed to reach the parameter boundaries. The observed failure rate is 2‰ for the $(\cos\theta_K, \cos\theta_e)$ fits; the failure rate for the Φ fits is 1.1%.

Finally, Fig. 6.21 shows the correlation matrices obtained from the pseudoexperiment ensembles, for the four angular observables, in each q^2 bin. Very little correlation is observed in each of the q^2 bins, and in particular none between the $(\cos\theta_K, \cos\theta_e)$ and Φ fit

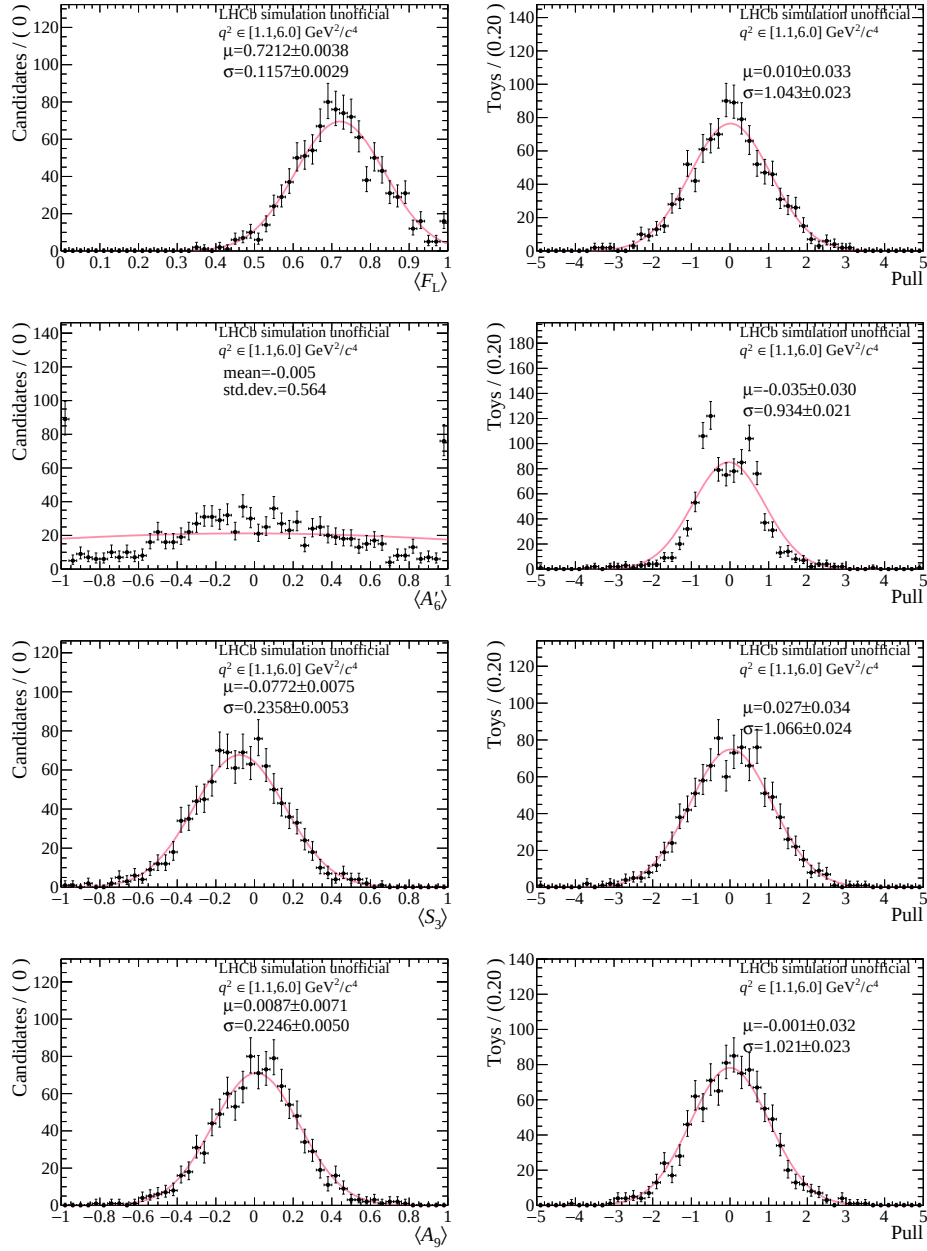


Figure 6.20: Distribution of the fitted values (left column) and pull distributions (right column) for the four angular observables, as fitted in pseudoexperiments for the central- q^2 bin. The distributions are fit with Gaussian functions to determine the sensitivity, bias, and coverage. From top to bottom, $\langle F_L \rangle$, $\langle A'_6 \rangle$, $\langle S_3 \rangle$, and $\langle A_9 \rangle$. In the case of $\langle A'_6 \rangle$, the large number of fitted values at the edges of the allowed parameter space makes such a fit unreliable, hence the standard deviation of the distribution is reported.

configurations.

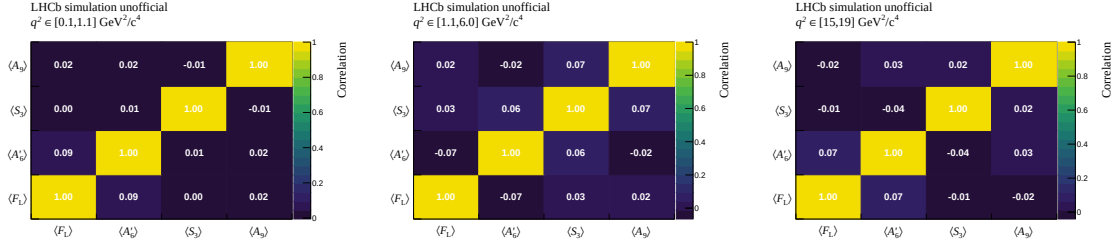


Figure 6.21: Correlation matrices between the fitted angular observables as determined from 1000 pseudoexperiments, in each q^2 bin. From left to right, the low-, central-, and high- q^2 bins. No significant correlation is observed.

6.6 Systematic uncertainties

There are many different sources of systematic uncertainty considered in this work. Unless specified otherwise, they are evaluated with high statistics pseudoexperiments. A dataset is generated for each q^2 bin, with a much larger size (a factor 1000 more) than the yields as determined in data. Using large pseudoexperiments ensures fit stability, and it ensures that the statistical uncertainty on the pseudoexperiments is not significant compared to the size of the observed biases. The generation includes a variation corresponding to the systematic effect being investigated, and the sample is then fitted back twice: first, with the alternative model used for generation, then with the nominal model.

The need to fit back twice and make a comparison between two fit configurations, rather than one compared to the generating value of each of the observables, is due to the strategy in generating the pseudoexperiments. Just like in Sec. 6.5, the pseudoexperiments are generated using the signal three-dimensional angular decay rate, including the acceptance model. As discussed in Sec. 6.7, when projecting onto $(\cos \theta_K, \cos \theta_e)$ or Φ , some residual integrals are not included in the nominal model. This can introduce an implicit bias in the fit strategy, which can become statistically relevant when dealing with very large sample sizes. The fit-by-fit difference for each of the observables, $\langle F_L \rangle$, $\langle A'_6 \rangle$, $\langle S_3 \rangle$, $\langle A_9 \rangle$, combined in quadrature with the corresponding statistical uncertainty due to the sample size, is then taken as the systematic uncertainty associated to a given source.

Among the sources considered, there are ones for which there is no alternative model to compare to, such as for the systematic uncertainty associated to the exclusion of a non-resonant $B_s^0 \rightarrow K^+ K^- e^+ e^-$ component in the fit. In such cases, the uncertainty is taken as the bias in the angular observables as extracted in the fit, compared to the generating value, again summed in quadrature with the sample size uncertainty.

Mass model description

The systematic uncertainties in this category account for the choices of parametrisation to model the mass distributions of the signal and background components. Alternative mass models are used to generate pseudoexperiments. The systematic uncertainty for each fit component is computed from a fit-by-fit comparison in pseudoexperiments, and combined in

quadrature with the alternative fit’s statistical uncertainty.

The default signal peak models are replaced by an alternative which features a different treatment of the tails. In the low- and central- q^2 bins, the alternative parametrisation features a modulation of the tails, and its parameters are determined from simulation. In the high- q^2 bin, where the modulation is present in the nominal model, the alternative model is simplified to include no modulation.

The alternative models for the leakage backgrounds in the central- and high- q^2 bins, as determined from simulation, are: a kernel density estimation for the $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ component in the central- q^2 bin, and for the $B_s^0 \rightarrow \psi(2S) (e^+e^-) \phi$ component in the high- q^2 bin; a Johnson- S_U function for the $B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ component in the high- q^2 bin. Similarly, the alternative mass model for the semileptonic cascades is a kernel density estimation determined from simulation.

In the case of the misidentified background, it is modelled alternatively with an exponential distribution in the low- and central- q^2 bins. In the high- q^2 bin, the alternative shape is the same as the polynomial model for the combinatorial background described in Sec. 5.2. The models’ parameters are determined from fits to the weighted control samples described in Sec. 4.8.3

Finally, a systematic uncertainty associated to the high- q^2 model for the combinatorial background is estimated by introducing a series of variations, and by using a power-law model as an alternative. The variations are the same as those outlined in Sec. 5.6.

Background angular model description

The systematic uncertainties in this category account for the arbitrary choices of parametrisation to model the angular distributions of the background components. Like before, alternative models are used to generate pseudoexperiments and the systematic uncertainties are obtained for each component from fit-by-fit comparisons.

In the case of the leakage, misidentified, and semileptonic cascade backgrounds, the $(\cos \theta_K, \cos \theta_e)$ distributions are modelled alternatively with even-ordered Chebychev polynomials, up to 4th order. The coefficients for each polynomial are determined with fits to the corresponding simulated or control sample. The Φ alternative models are kernel density estimations, again determined from the appropriate sample.

Limited sample size of background simulation samples

The uncertainties in this category estimate the effect of limited available sample sizes used for the description of the background components. They are estimated for the leakage, misidentified, and semileptonic cascade background, by a bootstrapping procedure. The samples used to determine each background model are bootstrapped, so that an ensemble of alternative models and expected yields are extracted. For each q^2 bin, pseudoexperiments are generated from each unique bootstrapped model and yield. The systematic uncertainties are then taken to be the mean of the distribution of the fit-by-fit difference between the

nominal and alternative model, over the generated ensemble, combined in quadrature with the fit's statistical uncertainty.

Assumption of mass-angle factorisation

An important assumption that is made in the description of the signal and background models is that there are no correlations between the mass and angle distributions. Experimental effects, like the bremsstrahlung recovery procedure, can however introduce correlations. Additionally, the combinatorial and misidentified backgrounds are composed of a varied mixture of physical backgrounds, and their composition can vary across the mass and angle ranges explored by the nominal fit. A systematic uncertainty associated to such effects is estimated with large pseudoexperiments, where correlations are introduced in the generation. The quoted systematic for each observable, and for each fit component, is the difference between the generating value and the value obtained in the nominal fit strategy, summed in quadrature with the statistical uncertainty from the fit.

In the case of the signal and semileptonic cascades, it is possible to generate pseudoexperiments by directly sampling their simulated samples, which will include any unaccounted for correlation. In the case of the misidentified and leakage backgrounds however, the sample sizes are not large enough to generate very large pseudoexperiments. Alternative models for these two backgrounds are built using the same angular moment analysis employed in Sec. 6.2 to model the angular acceptance. The use of angular moments ensures that correlations between the invariant mass and the three angles are taken into account in the model. These angular moments models are then used to generate the pseudoexperiments.

Angular acceptance model

The angular acceptance model is a key element entering the nominal fit, and many different sources associated to it are considered here. Firstly, a systematic uncertainty that accounts for the finite sample size of the simulation sample of $B_s^0 \rightarrow \phi e^+ e^-$ candidates generated with the FLATQ2 model is evaluated. This is computed in a similar manner as for the background models, by means of bootstrapping the sample and making an ensemble of alternative acceptance models.

The effect of the choice of maximum orders for the Legendre polynomials and cosine functions is estimated by creating an alternative model which includes one additional order for each dimension, *i.e.* $k \leq 7$ for q^2 , $l \leq 6$ for $\cos \theta_K$, $m \leq 10$ for $\cos \theta_e$, and $n \leq 3$ for Φ . The alternative acceptance model is used to generate pseudoexperiments, and the systematic uncertainty is estimated by the combination in quadrature of the fit-to-fit difference compared to the nominal model, and the statistical uncertainty from the fit.

A systematic uncertainty is also assigned to account for the effect of the simulation corrections on the acceptance model. To assess the change of the model due to these corrections, an alternative model is determined using the same FLATQ2 simulation sample, but without applying the corrections. This alternative model is used to generate pseudoexperiments and

estimate the uncertainty compared to the nominal model.

The EVTGEN model used to generate the $B_s^0 \rightarrow \phi e^+ e^-$ simulated sample with which the acceptance is modelled, FLATQ2, neglects the lifetime difference between the two B_s^0 meson mass eigenstates, *i.e.* $\Delta\Gamma_s = 0$, and assumes the lifetime to be the average of the light and heavy B_s^0 states. Since the B_s^0 lifetime distribution is incorrect, any effect on the acceptance from selections which shape the lifetime distribution may not be properly modelled. To assess the variation of the acceptance, two alternative models are computed, by reweighting the FLATQ2 sample to either have the light or heavy lifetimes

$$w_{L,H} = e^{-\Gamma_{L,H} t_{\text{TRUE}}} / e^{-\Gamma_s^{\text{Gauss}} t_{\text{TRUE}}}. \quad (6.32)$$

For each variation pseudoexperiments are generated, and the fit-by-fit difference compared to the nominal acceptance model is computed for each observable. Combining it in quadrature with the statistical uncertainty from each fit, one obtains the systematic uncertainty for the two variations. The individual uncertainties are then combined in quadrature.

Another important aspect is the choice of q^2 point at which to evaluate the angular acceptance in each q^2 bin. Because the q^2 efficiency is not perfectly flat across each bin, and the acceptance includes correlations between q^2 and the decay angles, a systematic uncertainty is estimated on the variation of the angular acceptance across the q^2 bins. The acceptance is recomputed at two alternative points, obtained by adding or subtracting the root mean square of the SM q^2 distribution to the median point, as determined from simulation. The two are used to generate the pseudoexperiments to estimate the variations in the angular observables. As the two variations are similar in size, the largest, combined in quadrature with the statistical uncertainty, is taken as the systematic uncertainty.

Semileptonic veto model

There are two sources of systematic uncertainty associated to the semileptonic veto model. The first is related to the choice of parametrisation as a function of $\cos\theta_e$. To assess this effect, an alternative model is built with kernel density estimation, in each q^2 bin. Pseudoexperiments are generated from this alternative model.

Additionally, the effect of q^2 migration on the shape of the veto is assessed. The migration of events due to the large emission of photons from final state electrons depends on the underlying q^2 distribution, at truth level. This means that potentially discrepancies between the q^2 distribution for simulated candidates and for real $B_s^0 \rightarrow \phi e^+ e^-$ events can bias the veto model. The effect is assessed by modelling the veto with the physics simulation sample, rather than the nominal FLATQ2. The alternative veto model is fit with the same lineshape, and used as variation to generate pseudoexperiments.

Just as before, from both sets of pseudoexperiments, the systematic uncertainties are estimated as fit-by-fit differences with the nominal model, combined in quadrature with the statistical uncertainty of the fit.

Other sources

This section covers other sources of systematic uncertainty which do not fall into any of the previous categories. Two sources are considered for the method used to obtain the weighted dataset used to model the misidentified background. In particular, the effect of the choice of kinematic binning and the definition of the “fail” region are assessed. An alternative binning variable is used, which uses the track’s momenta instead of their transverse momenta. An alternative smaller region is used, requiring that the PID selections based on $\text{DLL}_{e\pi}$ and ProbNN_e fail simultaneously, rather than taking the logical OR of the PID selection. From these variations, new transfer maps are computed and applied to the data to obtain alternative misidentified models. In both cases, the maps otherwise follow the nominal method, and is used to generate pseudoexperiments to estimate the systematic uncertainty on the angular observables.

An additional source of systematic uncertainty considered is the $D_s^\pm \rightarrow \phi$ form factor model used in EVTGEN when generating the double semileptonic decays [126]. This form factor model is outdated and in slight tension with the most recent experimental determinations [127]. An alternative sample was generated using form factors from a more reliable model [128], which is used to produce a set of generator-level weights as a function of the decay angle $\cos\theta_e$. The weights are then applied to the nominal simulation, and a second alternative angular model is derived, from which to generate pseudoexperiments, and compute the systematic uncertainty.

The resolution of the three decay angles is not taken into account in the nominal model. To estimate this contribution, pseudoexperiments are generated smearing the θ_K , θ_e , and Φ angles by the resolution function from Sec. 6.2, in each q^2 bin. The pseudoexperiments are then fitted back with the nominal model, and the systematic uncertainty is taken for each observable from the difference between the fitted and generating value.

In the high- q^2 bin, a residual contamination from partially reconstructed $B_s^0 \rightarrow \phi J/\psi X$ is expected, of around 6 events. This component is not included in the nominal fit. To estimate the systematic uncertainty related to this, pseudoexperiments are generated injecting the additional expected contribution, which is modelled with kernel density estimates from simulation, and fitted back with the nominal model. The systematic uncertainty is taken for each observable from the difference between the fitted and generating value, combined in quadrature with the statistical uncertainty from the fit.

In this analysis, the S-wave configuration of the di-kaon system (spin-0), corresponding to a non-resonant $B_s^0 \rightarrow K^+ K^- e^+ e^-$ component, is neglected in the fit. Such a component was measured in dedicated $B_s^0 \rightarrow J/\psi \phi$ analyses by LHCb [105]. Assuming it is of similar size to the resonant mode, using the results of Ref. [105], the expected S-wave fraction can be estimated in the window of $m(K^+ K^-)$ (relevant to this work) of $[1008, 1032]$ MeV/ c^2 , and it is found to be $F_S \approx 2\%$. To estimate the effect that neglecting the S-wave would have on the angular observables, pseudoexperiments are generated with the signal decay rate modified

to include an additional term,

$$\frac{1}{\frac{d(\Gamma+\bar{\Gamma})}{dq^2}} \frac{d(\Gamma+\bar{\Gamma})}{dq^2 d\cos\theta_K d\cos\theta_\ell d\Phi} \Big|_{S+P} = (1-F_S) \frac{d(\Gamma+\bar{\Gamma})}{dq^2 d\cos\theta_K d\cos\theta_\ell d\Phi} \Big|_P + \frac{3}{16\pi} F_S \sin\theta_e^2, \quad (6.33)$$

where the P-wave term corresponds to Eq. (2.50), and the interference terms in the S-wave component are taken to be zero. These are expected to cancel out in the flavour-averaged decay rate. The systematic uncertainty is taken for each observable as the difference between the fitted and generating value, combined in quadrature with the statistical uncertainty from the fit.

Finally, it is known that in a pp collider there is a small production asymmetry between the two $b\bar{b}$ flavours. This means there is an unequal amount of B_s^0 and $\bar{B}_s^0$ mesons at production and hence the CP -symmetric and CP -asymmetric observables mix, following the expression, at $t = 0$,

$$\begin{aligned} S_i^{\text{meas.}} &= S_i - \mathcal{A}_P A_i, \\ A_i^{\text{meas.}} &= A_i - \mathcal{A}_P S_i, \end{aligned} \quad (6.34)$$

where $\mathcal{A}_P$ is the production asymmetry expressed in terms of the production cross-sections $\sigma(B_s^0), \sigma(\bar{B}_s^0)$

$$\mathcal{A}_P = \frac{\sigma(\bar{B}_s^0) - \sigma(B_s^0)}{\sigma(\bar{B}_s^0) + \sigma(B_s^0)}. \quad (6.35)$$

This cross-contamination effect of the two observables is however “washed out” by the B_s^0 - $\bar{B}_s^0$ mixing, so that for the time integrated observables one has

$$\begin{aligned} \langle S_i^{\text{meas.}} \rangle &= \langle S_i \rangle - \mathcal{A}_P \kappa \langle A_i \rangle, \\ \langle A_i^{\text{meas.}} \rangle &= \langle A_i \rangle - \mathcal{A}_P \kappa \langle S_i \rangle, \end{aligned} \quad (6.36)$$

where an additional dilution factor κ is introduced, which is expressed as

$$\kappa = \int_0^\infty \epsilon(t) e^{-\Gamma_s t} \cos(\Delta m_s t) dt \Big/ \int_0^\infty \epsilon(t) e^{-\Gamma_s t} \cosh \frac{\Delta \Gamma_s t}{2} dt. \quad (6.37)$$

Using the time acceptance function as determined from $B_s^0 \rightarrow \phi e^+ e^-$ FLATQ2 simulation (see Sec. 6.4), as well as the current world averages for $\Delta \Gamma_s$, Γ_s , and Δm_s [17], the dilution factor is determined to be $\kappa = 4.75 \cdot 10^{-4}$. Given the size of the dilution, as well as the size of the measured asymmetry $\mathcal{A}_P = (1.1 \pm 2.6 \pm 0.7)\%$ [58], this effect is considered negligible and no systematic uncertainty is estimated.

6.6.1 Summary of systematic uncertainties

A detailed breakdown of the systematic uncertainties from each source discussed here is reported in App. E. Table 6.9 reports the systematic uncertainties partially combined for each component entering the fit. The two leading sources of systematic uncertainty are

related to the angular acceptance model, and to the assumption of mass-angle factorisation of the background distributions. In particular the latter (summed in quadrature over the different background components) reaches for $\langle F_L \rangle$ a size which is almost half of the statistical uncertainty, in all q^2 bins. It is important to note however that the sample sizes available to model these backgrounds are small, making it difficult to conclude much on the importance of the correlations. This source of systematic uncertainty includes a large statistical component which is expected to be reducible with a larger data set.

Table 6.9: Summary of the systematic uncertainties on $\langle F_L \rangle$, $\langle A'_6 \rangle$, $\langle S_3 \rangle$ and $\langle A_9 \rangle$ in units of 10^{-2} , in the three q^2 bins (separated by forward slashes). The last row shows the sum in quadrature of all sources. Correlations not taken into account when summing in quadrature are considered to be negligible when comparing to the statistical uncertainty, which is the dominant uncertainty in the measurement. Table taken from Ref. [122].

Source	$\langle F_L \rangle$	$\langle A'_6 \rangle$	$\langle S_3 \rangle$	$\langle A_9 \rangle$
Signal mass model	1.2 / 2.1 / 0.6	1.9 / 2.8 / 1.8	1.3 / 1.1 / 1.7	1.7 / 1.1 / 1.3
Acceptance model	1.3 / 1.9 / 1.9	2.6 / 5.0 / 2.3	2.1 / 1.9 / 1.9	1.8 / 1.7 / 1.7
S-wave contamination	0.5 / 1.6 / 0.8	1.3 / 1.8 / 0.8	1.3 / 0.8 / 1.4	0.9 / 0.9 / 0.7
Angular resolution	0.7 / 1.9 / 0.4	1.0 / 2.3 / 2.0	0.8 / 0.7 / 0.7	1.7 / 1.7 / 1.1
Misidentified background	5.6 / 2.8 / 3.1	2.3 / 4.0 / 2.3	4.3 / 5.4 / 7.1	1.9 / 1.5 / 2.4
Leakage background	0.0 / 1.6 / 1.7	0.0 / 3.2 / 2.4	0.0 / 2.1 / 2.9	0.0 / 1.4 / 2.2
Semileptonic background	1.4 / 1.1 / 0.0	2.0 / 3.8 / 0.0	1.4 / 3.0 / 0.0	1.4 / 1.4 / 0.0
Combinatorial background	2.0 / 3.2 / 1.5	1.2 / 4.0 / 1.9	4.3 / 0.9 / 1.1	1.0 / 1.3 / 2.3
Combined	6.0 / 6.0 / 5.0	5.0 / 9.0 / 5.0	7.0 / 7.0 / 8.0	4.0 / 4.0 / 5.0

6.7 Results

The fits are performed separately in $(m(K^+K^-e^+e^-), \cos \theta_e, \cos \theta_K)$ and $(m(K^+K^-e^+e^-), \Phi)$ in each q^2 bin, following the fit strategy outlined in Sec. 6.1. The yields of the misidentified, semileptonic, and leakage backgrounds are constrained to the expected values as computed from simulation. Additionally, the angular parameters for the combinatorial background, a, b, c, d for the $(\cos \theta_K, \cos \theta_e)$ distribution, and a_1, b_1 for the Φ distribution, are constrained to the values obtained from fits to SS data. Table 6.10 reports the constraints used in the fit. Figure 6.22 shows the result of the maximum likelihood fits to data, in the low-, central-, and high- q^2 bins. The best fit values for the floating parameters are reported in Tables 6.11 and 6.12.

The FC procedure is used to determined confidence intervals as described in App. D. The resulting intervals on individual parameters are shown in Fig. 6.23, along with two-dimensional intervals for $[\langle F_L \rangle, \langle A'_6 \rangle]$ and $[\langle S_3 \rangle, \langle A_9 \rangle]$ in Figs. 6.24 and 6.25, respectively.

Table 6.13 reports the final results for the decay time-integrated observables $\langle F_L \rangle$, $\langle A'_6 \rangle$, $\langle S_3 \rangle$, and $\langle A_9 \rangle$ in the three q^2 bins, reporting both the intervals obtained with the FC prescription and the total systematic uncertainties. A graphical comparison with the SM

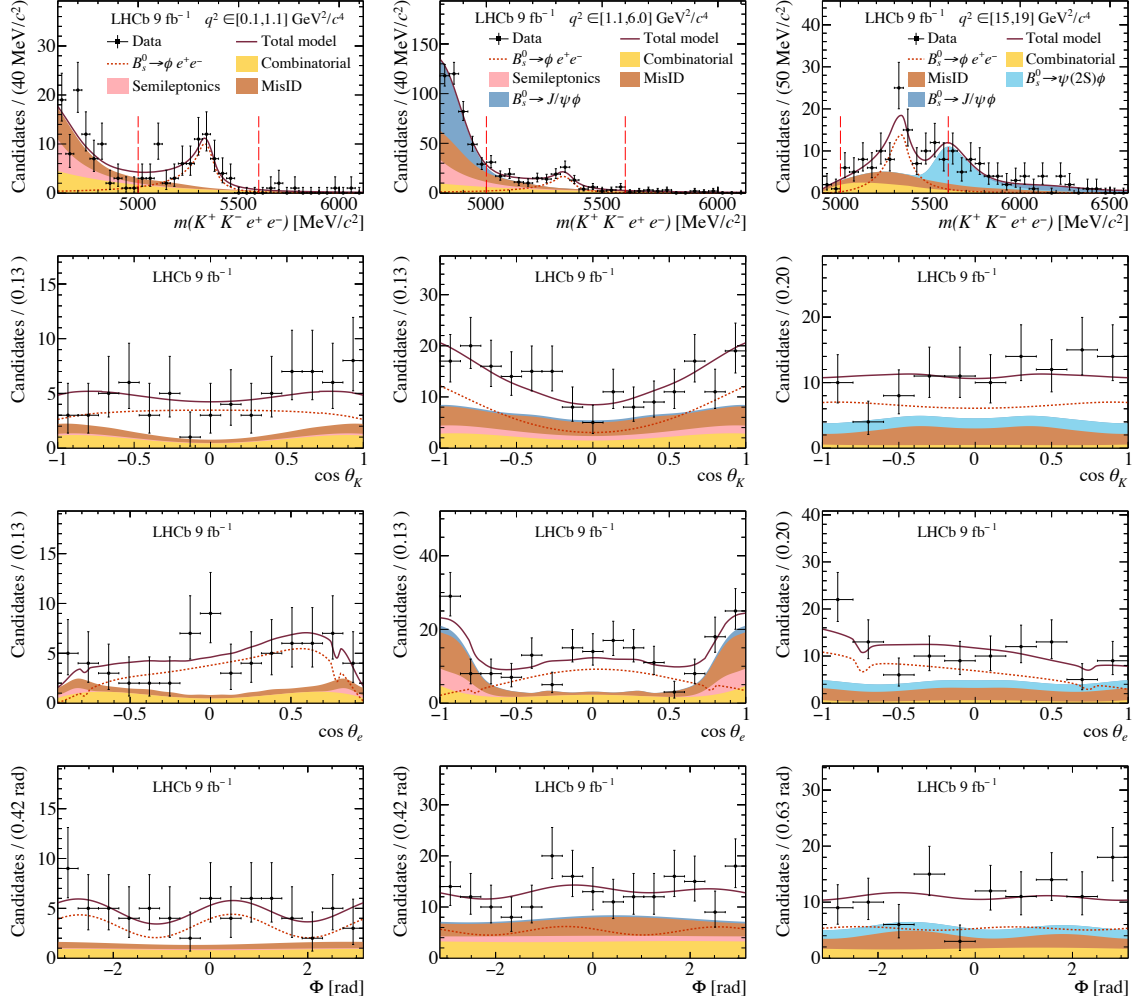


Figure 6.22: Result of the fits to the distribution of $B_s^0 \rightarrow \phi e^+ e^-$ data in the low-, central-, and high- q^2 bins (left to right) in $m(K^+ K^- e^+ e^-)$, $\cos \theta_K$, $\cos \theta_e$ and Φ (top to bottom). The mass distribution shows the result of the $m(K^+ K^- e^+ e^-)$, Φ fit as an example. The angular distributions are shown in reduced mass windows indicated by the vertical dashed red lines. Figures taken from Ref. [122].

Table 6.10: Expectation values and errors used to constrain the yields and shapes of various backgrounds components when performing the mass and angle fits to $B_s^0 \rightarrow \phi e^+ e^-$ data. Where indicated with “–”, background contributions are not modelled as they have been found to be negligible.

Yield	low- q^2	central- q^2	high- q^2
a	$(3.2 \pm 1.6) \cdot 10^{-1}$	$(2.7 \pm 1.0) \cdot 10^{-1}$	$(0.8 \pm 0.9) \cdot 10^{-1}$
b	$(-0.9 \pm 1.6) \cdot 10^{-1}$	$(-1.1 \pm 0.9) \cdot 10^{-1}$	$(0.7 \pm 0.7) \cdot 10^{-1}$
c	$(0.2 \pm 1.9) \cdot 10^{-1}$	$(4.9 \pm 0.8) \cdot 10^{-1}$	$(0.3 \pm 0.9) \cdot 10^{-1}$
d	$(-3.4 \pm 1.8) \cdot 10^{-1}$	$(2.7 \pm 0.7) \cdot 10^{-1}$	$(3.0 \pm 0.7) \cdot 10^{-1}$
a_1	$(-0.2 \pm 1.3) \cdot 10^{-1}$	$(0.0 \pm 0.8) \cdot 10^{-1}$	$(0.3 \pm 0.7) \cdot 10^{-1}$
b_1	$(0.3 \pm 1.4) \cdot 10^{-1}$	$(0.3 \pm 0.8) \cdot 10^{-1}$	$(0.8 \pm 0.7) \cdot 10^{-1}$
$N_{\text{semileptonic}}$	22.0 ± 2.1	75.8 ± 6.3	–
$N_{B_s^0 \rightarrow J/\psi(e^+e^-)\phi}$	–	192.8 ± 5.6	12.7 ± 0.9
$N_{B_s^0 \rightarrow \psi(2S)(e^+e^-)\phi}$	–	–	49.0 ± 4.1
N_{misID}	36.7 ± 4.3	148.9 ± 8.3	41.1 ± 5.1

Table 6.11: Parameters determined from the $(m(K^+K^-e^+e^-), \cos\theta_K, \cos\theta_e)$ maximum likelihood fits to $B_s^0 \rightarrow \phi e^+ e^-$ data in the low, central, and high q^2 bins. The reported uncertainties are statistical and are computed with HESSE [129]. The uncertainties for the yields of the semileptonic, misidentified, and leakage backgrounds are driven by the uncertainty on the constraint added to the fit. Where indicated with “–”, background contributions are not modelled as they have been found to be negligible.

Parameter	low- q^2	central- q^2	high- q^2
$\langle F_L \rangle$	0.25 ± 0.12	0.67 ± 0.09	0.43 ± 0.10
$\langle A'_6 \rangle$	0.60 ± 0.25	0.23 ± 0.40	-0.57 ± 0.23
$\lambda_{\text{combinatorial}}$	$(-1.7 \pm 0.7) \cdot 10^{-3}$	$(-1.5 \pm 0.6) \cdot 10^{-3}$	$(-9.5 \pm 6.8) \cdot 10^{-3}$
a	$(3.8 \pm 1.6) \cdot 10^{-1}$	$(2.8 \pm 1.0) \cdot 10^{-1}$	$(0.7 \pm 0.9) \cdot 10^{-1}$
b	$(-1.6 \pm 1.5) \cdot 10^{-1}$	$(-1.1 \pm 1.0) \cdot 10^{-1}$	$(0.6 \pm 0.7) \cdot 10^{-1}$
c	$(0.3 \pm 2.0) \cdot 10^{-1}$	$(4.6 \pm 0.8) \cdot 10^{-1}$	$(0.2 \pm 0.8) \cdot 10^{-1}$
d	$(-4.0 \pm 1.7) \cdot 10^{-1}$	$(3.1 \pm 0.7) \cdot 10^{-1}$	$(3.0 \pm 0.7) \cdot 10^{-1}$
$N_{B_s^0 \rightarrow \phi e^+ e^-}$	53 ± 10	101 ± 15	72 ± 12
$N_{\text{combinatorial}}$	32 ± 13	62 ± 21	11.7 ± 8.6
$N_{\text{semileptonic}}$	23.8 ± 2.1	79.2 ± 6.6	–
N_{misID}	37.4 ± 4.3	149.4 ± 8.0	41.0 ± 4.9
$N_{B_s^0 \rightarrow J/\psi\phi}$	–	195.9 ± 6.0	12.8 ± 1.6
$N_{B_s^0 \rightarrow \psi(2S)\phi}$	–	–	47.9 ± 3.8

predictions from `flavio` [48] and with the previous LHCb results using $B_s^0 \rightarrow \phi \mu^+ \mu^-$ decays [99], is shown in Fig. 6.26. The observables are compatible with the Standard Model predictions, as well as with the previous measurement using the muon mode.

Table 6.12: Parameters determined from the $(m(K^+K^-e^+e^-), \Phi)$ maximum likelihood fit to $B_s^0 \rightarrow \phi e^+e^-$ data in the low, central, and high q^2 bins. The reported uncertainties are statistical and are computed with HESSE [129]. The uncertainties for the yields of the semileptonic, misidentified, and leakage backgrounds are driven by the uncertainty on the constraint added to the fit. Where indicated with “–”, background contributions are not modelled as they have been found to be negligible.

Parameter	low- q^2	central- q^2	high- q^2
$\langle S_3 \rangle$	0.23 ± 0.23	0.07 ± 0.20	-0.01 ± 0.26
$\langle A_9 \rangle$	0.28 ± 0.22	-0.14 ± 0.21	0.06 ± 0.23
$\lambda_{\text{combinatorial}}$	$(-2.4 \pm 0.8) \cdot 10^{-3}$	$(-2.3 \pm 0.6) \cdot 10^{-3}$	$(-4.2 \pm 3.7) \cdot 10^{-3}$
a_1	$(-0.8 \pm 1.3) \cdot 10^{-1}$	$(-0.1 \pm 0.8) \cdot 10^{-1}$	$(0.2 \pm 0.6) \cdot 10^{-1}$
b_2	$(-0.1 \pm 1.3) \cdot 10^{-1}$	$(0.3 \pm 0.7) \cdot 10^{-1}$	$(0.9 \pm 0.7) \cdot 10^{-1}$
$N_{B_s^0 \rightarrow \phi e^+e^-}$	55 ± 10	88 ± 16	59 ± 18
$N_{\text{combinatorial}}$	44 ± 14	97 ± 29	28 ± 26
$N_{\text{semileptonic}}$	22.0 ± 2.3	76.5 ± 6.8	–
N_{misID}	37.5 ± 4.3	149.1 ± 8.2	40.4 ± 5.1
$N_{B_s^0 \rightarrow J/\psi \phi}$	–	193.7 ± 6.1	12.7 ± 1.6
$N_{B_s^0 \rightarrow \psi(2S)\phi}$	–	–	47.5 ± 4.2

Table 6.13: Measured values of the time-integrated observables in the three q^2 bins. The central value indicates the best-fit point, the first set of uncertainties indicate the range of the 68% confidence level interval about the best-fit point, as determined with the FC method, and the second correspond to the total systematic uncertainty. Table taken from Ref. [122].

Observable	low- q^2	central- q^2	high- q^2
$\langle F_L \rangle$	$0.25^{+0.12}_{-0.12} \pm 0.06$	$0.67^{+0.12}_{-0.13} \pm 0.06$	$0.43^{+0.11}_{-0.10} \pm 0.05$
$\langle A'_6 \rangle$	$0.60^{+0.23}_{-0.28} \pm 0.05$	$0.24^{+0.40}_{-0.42} \pm 0.09$	$-0.57^{+0.24}_{-0.25} \pm 0.05$
$\langle S_3 \rangle$	$0.23^{+0.24}_{-0.24} \pm 0.07$	$0.07^{+0.21}_{-0.21} \pm 0.07$	$-0.01^{+0.29}_{-0.28} \pm 0.08$
$\langle A_9 \rangle$	$0.28^{+0.23}_{-0.24} \pm 0.04$	$-0.14^{+0.23}_{-0.24} \pm 0.04$	$0.06^{+0.25}_{-0.25} \pm 0.05$

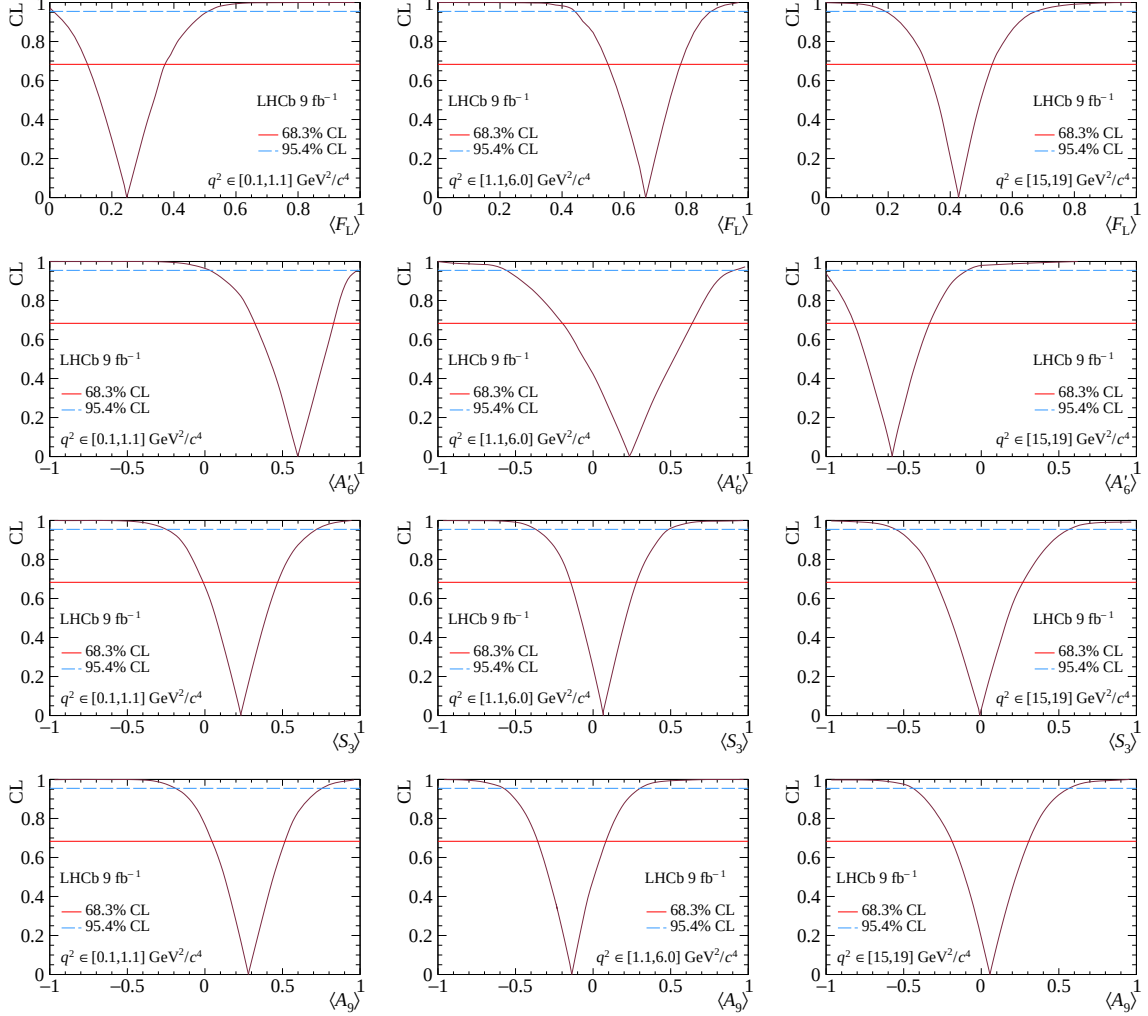


Figure 6.23: Result of the FC procedure for the observables $\langle F_L \rangle$, $\langle A'_6 \rangle$, $\langle S_3 \rangle$ and $\langle A_9 \rangle$ (top to bottom), in the low-, central, and high- q^2 bins (left to right). The 68.3% (95.4%) confidence level is drawn as the red solid (blue dashed) line. Figures taken from Ref. [122].

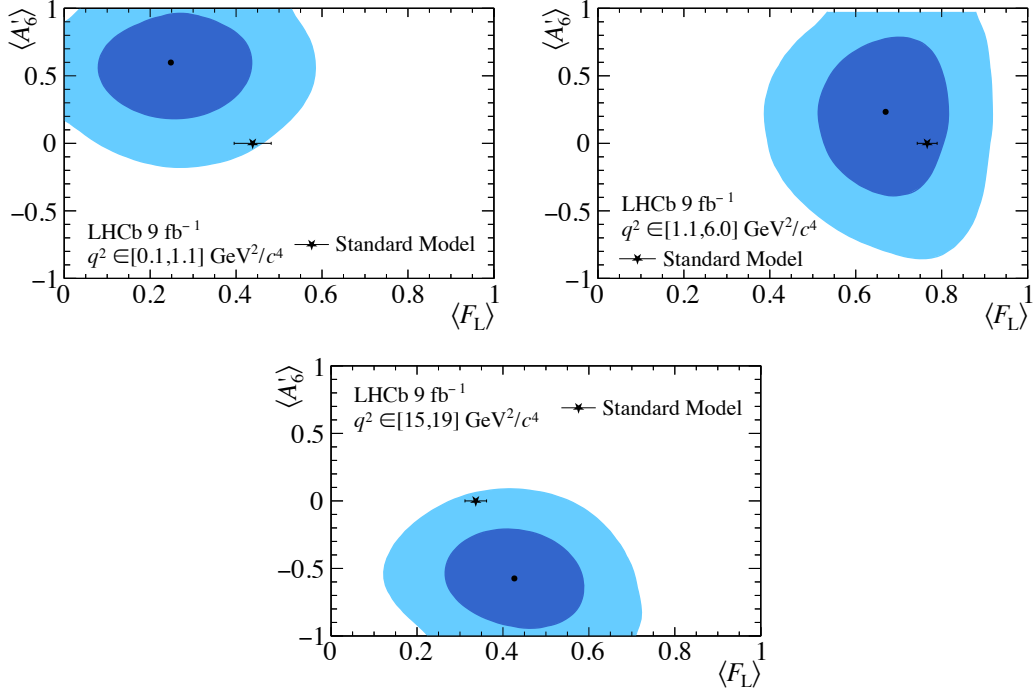


Figure 6.24: Two-dimensional 68.3% and 95.5% confidence level contours for the observables $\langle F_L \rangle$ and $\langle A'_6 \rangle$ in the low- (top left), central- (top right), and high- q^2 (bottom) bins. Figures taken from Ref. [122].

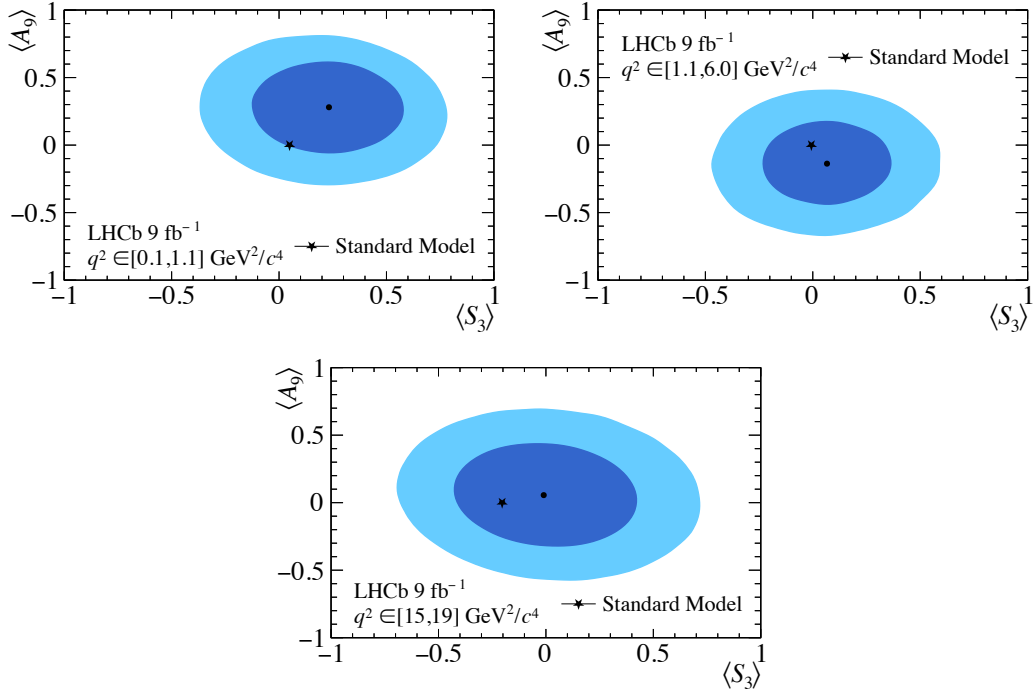


Figure 6.25: Two-dimensional 68.3% and 95.5% confidence level contours for the observables $\langle S_3 \rangle$ and $\langle A_9 \rangle$ in the low- (top left), central- (top right), and high- q^2 (bottom) bins. Figures taken from Ref. [122].

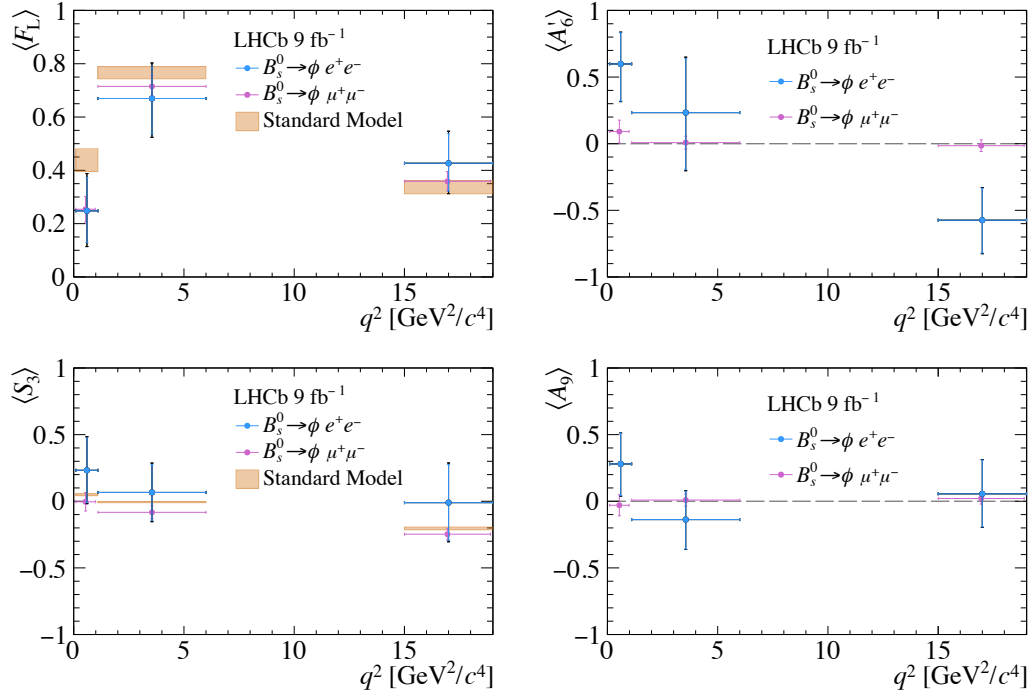


Figure 6.26: Measured values, in bins of q^2 , of the time-integrated angular observables $\langle F_L \rangle$, $\langle S_3 \rangle$, $\langle A'_6 \rangle$, and $\langle A_9 \rangle$, shown by light blue crosses, overlaid with the SM prediction computed with the `flavio` package [48] as the shaded orange boxes. The results are compared to the previous LHCb measurements of the same observables with the muon mode $B_s^0 \rightarrow \phi \mu^+ \mu^-$ [99], and are shown as pink crosses. Figures taken from Ref. [122].

The LHCb Early Career, Gender and Diversity office

The author's recommendation

Metropolis: The Chase Suite

by **Janelle Mon  e**

2007, Bad Boy

With this chapter I would like to highlight a part of the LHCb collaboration that I had the pleasure of being a part of during my PhD, parallel to my main research focus: the LHCb Early Career, Gender and Diversity office (ECGD). Particle physics has a long history of being a highly collaborative field, with thousands of researchers taking part in large scientific collaborations, like those at CERN. The ECGD office was created to support LHCb members and to ensure the best conditions for effective collaborative work. Having been one of the four members of the office from 2023 to 2025, in the role of Early Career representative, and chair of the LHC-wide early career fora from 2024 to 2025, I will give here a short overview of some of the activities I was involved in during my term.

7.1 Structure of the LHCb collaboration and the ECGD office

The LHCb collaboration, as of 2025, is composed of over 1700 members from more than 100 institutes across 24 countries [130]. Its organisational structure is set out by the LHCb Constitution. The fundamental authority within the collaboration is the Collaboration Board (CB), which is the body through which major decisions are taken. The executive body of the collaboration is the LHCb Management, which is responsible for organising the detector operations, its upgrades and physics analysis output, as well as its resources.

An important body which is defined through the LHCb Constitution is the Early Career, Gender and Diversity office. It came into existence after being ratified by the CB in 2014 [131], with the goal, quoting from the Constitution, to “oversee the well-being and working environment of all LHCb members”. LHCb was the first large collaboration in the

field of high-energy physics to create such a body. The office is composed of two ECGD officers, and two Early Career (EC) representatives. The ECGD officers are senior members of the collaboration, typically with permanent contracts, whose seniority ensures they can act independently and without fear of intimidation and/or retaliation. They are selected by the LHCb Management, and are appointed by the CB for two years. The role of the EC representatives is to listen to the needs and concerns of the LHCb EC community. They are elected for a two-year term by the LHCb EC community, which is defined by all members of the collaboration who do not hold a permanent position. Self-definition is also allowed for those who do hold a permanent position but still identify as being EC. The EC community makes up about half of all LHCb members.

7.2 Activities of the ECGD office

The ECGD office is available to all members of the collaboration, offering advice, in a confidential manner, in cases of harassment, discrimination, or other problems. It monitors the gender and diversity balance within the collaboration, as a function of career status, in leadership roles, and in the allocation of public talks given on behalf of LHCb; it additionally provides insight into the quality of the work environment of the collaboration. This monitoring is done through surveys, which have been carried out in 2013 (prior to the official ratification of the office), 2016, and 2019. The key results from the 2016 and 2019 surveys can be found in Refs. [131] and [132], respectively. The surveys highlight progress over the years in the fraction of female members of the collaboration,¹ which has increased especially for the early career members. At the same time, they report on the highly stressful environment in which many collaboration members find themselves in, especially ones at the beginning of their career.

The ECGD office over the years has started, or been part of, a series of initiatives which try to alleviate the difficulties which LHCb members might be going through. These are advertised on the main ECGD webpage [133], and in a weekly newsletter to the collaboration. At each LHCb collaboration week, which is a week-long event held four times a year, the office organises a plenary meeting, covering a wide range of topics related to ECGD matters. The sessions are typically attended by 100-200 members (in person and online), and are set up to encourage an open discussion of the topic at hand. Topics which have been covered during my term as EC representative include [134]:

- The experience of being a post-doctoral researcher in the collaboration.
- Work-related stress, and results of an internal survey on stress levels felt as a function of the career stage.
- The visibility and recognition of technical work within the collaboration.

¹The CERN membership databases currently only store binary gender information, which limits the information that can be gathered on the experience of non-binary members of the collaboration.

- The feedback-giving process of internal code review, and how it affects the user experience of contributing to the LHCb code base.
- Improving the accessibility of collaboration meetings through recordings and transcripts.
- The gender balance of senior-level roles of the collaboration.

Some of these topics have resulted in actions being taken by the collaboration to incite change. For example, the CB has ratified the creation of an internal committee which is investigating the challenges and barriers which might prevent a more diverse composition of LHCb’s management and other senior-level roles; the ECGD office, in conjunction with the LHCb Management, has also started trialling a standardised procedure to record online meetings. Through these actions the ECGD office helps to improve the work environment within LHCb, ultimately increasing the collaboration’s scientific productivity.

7.2.1 LHC wide activities

The ECGD office, through its EC representatives, also collaborates with the equivalent figures from the other three main experiments on the LHC ring. Together with the ALICE Junior representatives, the ATLAS Early Career Scientist Board, and the CMS Young Scientist Committee, it established in 2020 the LHC Early Career Scientist Fora (ECSF) [135]. The ECSF allows the four independent groups to join forces to coordinate actions to address issues regarding the LHC wide EC community. The four groups meet regularly (every 3 to 6 months) to discuss joint initiatives as well as to share insights on the internal efforts in each collaboration.

There are many initiatives organised jointly through the ECSF. The Soft Skills Workshop series [136] holds an event twice a year to help EC researchers develop a more complete set of skills, which they can use to further develop their career both in or out of the academic setting. Each workshop focuses on a particular topic and covers it either as a seminar-style talk, or through a panel of speakers from all four experiments. Topics have included CV writing, interview skills, job networking, and others.

The Job Matching Event [137] is an initiative held twice a year, which offers a platform for job recruiters and seekers to meet and discuss in moderated sessions. The sessions allow a recruiter to meet up to six interested job seekers at a time, and to advertise a job vacancy in one of the four experiments. Each seeker can “break the ice” by presenting their career history and previous experience in a rapid-fire presentation, after which the rest of the session is dedicated to a open discussion between the seekers and recruiter on the vacancy. An equivalent event, focused on job vacancies outside of the academic sector, is currently being planned by the ECSF.

The Mentoring@CERN scheme [138] was born in 2024 out of the merger of the LHC ECSF and Women in Technology (WIT) mentoring schemes, which have been running since 2020 and 2018 respectively. During its 2024 and 2025 rounds, it successfully matched more

than 150 mentor-mentee pairings, bringing the total number of matches across the three programmes to more than 400. Figure 7.1 shows the yearly participation in the each programme 2018–2025 period.

The Mentoring@CERN programme allows anyone with a current affiliation to CERN to become a potential mentor and/or mentee in the programme. The application process assesses the needs and preferences of the mentor/mentee, through a series of multiple choice and freeform questions. The organisers then use the information from each application to pair up and propose potential mentor-mentee matches. After a kick-off meeting on what mentoring is and how to efficiently set goals within the mentoring relationship, the pairings independently setup a schedule of regular meetings. Each mentoring round lasts about nine months. A survey conducted on participants from the 2024 round found high levels of satisfaction with the programme.

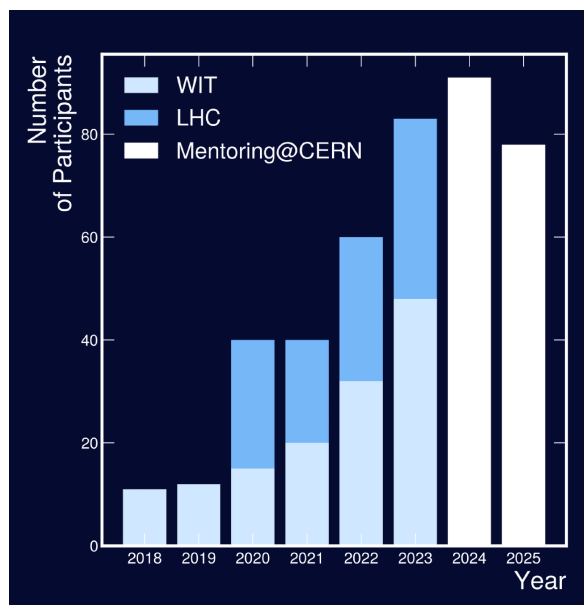


Figure 7.1: Participation in CERN mentoring schemes (WIT, LHC ECSF, and Mentoring@CERN) in the 2018–2025 period. The numbers correspond to the number of unique mentor-mentee pairings made in the yearly rounds of each mentoring scheme. Figure taken from Ref. [138].

Conclusions

The author's recommendation

What's Your Pleasure?

by **Jessie Ware**

2020, PMR / EMI

This thesis has presented two analyses of data collected by the LHCb experiment between 2011 and 2018, of the decay properties of the $B_s^0 \rightarrow \phi \ell^+ \ell^-$ process. The first is a test of lepton flavour universality, comparing the relative decay rates of the $B_s^0 \rightarrow \phi \mu^+ \mu^-$ and $B_s^0 \rightarrow \phi e^+ e^-$ decays. The analysis presents important challenges related to the reconstruction of electrons with the LHCb detector. The $B_s^0 \rightarrow \phi e^+ e^-$ final state, compared to its muon counterpart, is characterised by a lower reconstruction and selection efficiency, a worse invariant mass resolution, and a higher misidentification probability. All of these challenges have been tackled, building on the experience gained from previous similar LHCb analyses. The analysis has been successfully completed and published in Physics Review Letters. Its results are found to be compatible with the assumption of lepton flavour universality, within the reported uncertainties. This measurement also corresponds to the first ever observation of the $B_s^0 \rightarrow \phi e^+ e^-$ decay. While limited by the sample size of the Run 1 and 2 LHCb dataset, compared to the corresponding muon decay, the measurement is helpful in further constraining the available parameter space for many proposed extensions to the Standard Model.

The second measurement follows naturally from the first, and is the first angular analysis of the $B_s^0 \rightarrow \phi e^+ e^-$ decay, aimed at testing lepton flavour universality at the level of the angular observables. The available sample size, due to the small reconstruction and selection efficiency, poses again a challenge. Because of this, it is not yet feasible to perform a full analysis of the three-dimensional decay rate. Still, by performing separate fits to the angular projections, it is possible to measure four angular observables, in three bins of the dielectron invariant mass squared. The analysis is complete and its results have been published in the Journal of High Energy Physics. The measurements of each observable are compatible both with their Standard Model predictions, as well as with their previous determination

using $B_s^0 \rightarrow \phi \mu^+ \mu^-$ decays. The uncertainties on angular observables are dominated by the limited sample size of the dataset. The leading systematic uncertainties are also limited by the available control datasets used to model the various background components.

The LHCb Early Career, Gender and Diversity office was founded almost a decade ago. Since then, it has frequently monitored the quality of the work environment of LHCb members, as well as the gender balance within the collaboration, through a series of anonymous surveys. It is actively involved, through its Early Career representatives, in both LHCb-specific and LHC-wide activities, aimed at helping and guiding young researchers through the difficulties of the academic environment. The office has evolved since its original mandate in 2016, and is still evolving today to meet the needs of LHCb’s researchers. Helping to improve the working conditions of all LHCb members is crucial to achieve the best scientific output of the experiment.

The upgraded LHCb detector has been fully commissioned and is operating successfully in its third data-taking run, having already collected in 2024 more data than from Run 1 and Run 2 combined. The vertex locator, which has been completely redesigned, is achieving its design performance. The muon detector’s hardware has remained largely unchanged, while its data acquisition electronics have been completely upgraded to be able to interface with the new trigger. Its new monitoring system has been key during the commissioning of the new electronics, and is now running stably at Point 8. The performance of the detector, as monitored in real time, is still excellent despite its hardware having been operating for more than a decade.

Bremsstrahlung reconstruction will remain an important challenge of analyses of the $B_s^0 \rightarrow \phi e^+ e^-$ final state (as well as any other decay containing electrons). The upgraded tracking system features a reduced material budget, compared to the original LHCb trackers, which should reduce the amount of radiation emitted by an electron traversing the detector. At the same, the performance of the electromagnetic calorimeter is expected to have degraded with respect to the 2011–2018 period, because of ionising radiation damage and the higher track multiplicity. However, the future upgrade of the calorimeter inner cells, which is planned for the upcoming LHC Long Shutdown 3 (2026–2030), is expected to make significant improvements in the calorimeter’s performance. Given these differences, one can expect the overall performance of bremsstrahlung reconstruction over the lifetime of the upgraded LHCb detector (Run 3 and Run 4) to be similar that in Run 1 and Run 2.

The dataset which will be collected with the upgraded detector is expected to be more than four times as large as that collected in the 2011–2018 period, corresponding to a total integrated luminosity of 50 fb^{-1} by 2032. The fundamental change to a software-based trigger system has already been shown to have greatly improved the reconstruction efficiency of electron final states, which will further increase the size of the $B_s^0 \rightarrow \phi e^+ e^-$ dataset. This is expected to reduce statistical uncertainties on R_ϕ and the angular observables by almost a factor of four. At the same time, the larger dataset will allow to perform the first full three-dimensional fit of the angular decay rate of the $B_s^0 \rightarrow \phi e^+ e^-$ decay, doubling the number of measureable observables with respect to the analysis presented in this thesis. One can expect

also some of the systematic uncertainties related to the modelling of certain backgrounds to decrease: for example, the uncertainties associated to the combinatorial and misidentified backgrounds, which are mainly driven by the size of their respective control samples. Future LHCb $b \rightarrow se^+e^-$ analyses will greatly benefit from the upgraded detector and contribute to a further understanding of the flavour physics landscape, within the Standard Model and hopefully beyond.

Normalisation of the correction weights for efficiency computation

The author's recommendation

Run Away with Me

from the album **E·MO·TION**

by **Carly Rae Jepsen**

2015, 604 / School Boy / Interscope

Computing efficiencies for weighted datasets requires that the set of weights are properly normalised, so that when summed over all simulated candidates they correctly return the number of selected and generated events, N_{pass} and N_{gen} .

The weights for the PID corrections for muons and kaons do not need to be normalised, and, since they effectively replace the PID selection, the efficiency can be simply computed as the ratio

$$\epsilon = \frac{\sum_{i \in N_{\text{gen}}} w_i^{\text{PID}}}{N_{\text{gen}}}. \quad (\text{A.1})$$

The electron PID, track reconstruction and trigger correction weights, which are all expressed as data-simulation efficiency ratios, do not need additional normalisation. The efficiency is computed in a similar manner, with the numerator containing a weighted sum, with the difference that the sum is performed over the candidates which pass the corresponding selection,

$$\epsilon = \frac{\sum_{i \in N_{\text{pass}}} w_i}{N_{\text{gen}}} = \frac{\sum_{i \in N_{\text{pass}}} w_i}{N_{\text{pass}}} \cdot \frac{N_{\text{pass}}}{N_{\text{gen}}}, \quad (\text{A.2})$$

the standard unweighted efficiency is corrected to include the correct normalisation of the weights.

The B_s^0 kinematics, event multiplicity and vertex reconstruction correction weights are built from the output of a BDT reweighter, and are not normalised to the number of simulated candidates. The B_s^0 lifetime weights are also not correctly normalised as they alter the generator level decay rates. Thus the efficiency is computed applying the weights in both

the numerator and denominator,

$$\epsilon = \frac{\sum_{i \in N_{\text{pass}}} w_i}{\sum_{i \in N_{\text{gen}}} w_i}, \quad (\text{A.3})$$

to correctly scale the weights. The B_s^0 kinematics, event multiplicity, and B_s^0 lifetime weights can all be normalised at the stage of generation of the simulated events. This is not the case for the vertex reconstruction weights: to compute them one needs to have run the reconstruction algorithm on the simulated sample. This requires the introduction of an intermediate selection stage, at which enough information is available in the simulated sample to compute and normalise these weights. Since the weights are computed independently for each trigger category, the earliest they can be obtained is at the stripping stage. Hence, the total reconstruction and selection efficiency is computed as the product of two efficiencies,

$$\begin{aligned} \epsilon_{\text{selection}} &= \epsilon(\text{stripping}|\text{generated}, q_{\text{true}}^2) \cdot \epsilon(\text{selection}|\text{stripping}) \\ &= \frac{\sum_{i \in N_{\text{strip}}} \tilde{w}_i}{\sum_{i \in N_{\text{gen}}} \tilde{w}_i} \cdot \frac{\sum_{i \in N_{\text{sel}}} w_i}{\sum_{i \in N_{\text{strip}}} w_i}, \end{aligned} \quad (\text{A.4})$$

where $\tilde{w}_i$ is the product of the B_s^0 kinematics, event multiplicity, and B_s^0 lifetime weights; $\tilde{w}_i$ is the product of all weights except the ones for vertex reconstruction; w_i is the product of all correction weights. The sums are performed over the generated candidates which fall in the signal q_{true}^2 region N_{gen} , the candidates which pass the stripping selection N_{strip} , and the candidates which pass the full selection N_{sel} . As explained above, the only differences with the nominal selection when computing the efficiencies is the removal of the muon and kaon PID requirements, which are replaced by their respective PID weights, and the use of the q_{smear}^2 variable to select each q^2 bin.

Appendix B

Projections of the $B_s^0 \rightarrow \phi e^+ e^-$ angular model including acceptance effects

The author's recommendation

Città Vuota

from the album **Studio Uno**

by **Mina**

1965, Ri-Fi

Following Sec. 6.2.2, the unnormalised three-dimensional model is written as

$$\begin{aligned} \mathcal{P}(\cos \theta_K, \cos \theta_e, \Phi) &\propto \epsilon(\cos \theta_K, \cos \theta_e, \Phi) \cdot \frac{1}{d(\Gamma + \bar{\Gamma})/dq^2} \frac{d(\Gamma + \bar{\Gamma})}{dq^2 d\cos \theta_K d\cos \theta_e d\Phi} \\ &= \sum_{l,m,n} C_{lmn} L_l(\cos \theta_K) L_m(\cos \theta_e) \cos(n\Phi) \\ &\quad \times \mathcal{P}_{\text{phys.}}(\cos \theta_K, \cos \theta_e, \Phi), \end{aligned} \quad (\text{B.1})$$

where the notation $\mathcal{P}_{\text{phys.}}(\cos \theta_K, \cos \theta_e, \Phi)$ has been introduced for brevity to express the signal three-dimensional decay rate. The projection in $(\cos \theta_K, \cos \theta_e)$ is obtained integrating the full expression over Φ ,

$$\mathcal{P}(\cos \theta_K, \cos \theta_e) \propto \int_{-\pi}^{\pi} \sum_{l,m,n} C_{lmn} L_l(\cos \theta_K) L_m(\cos \theta_e) \cos(n\Phi) \cdot \mathcal{P}_{\text{phys.}}(\cos \theta_K, \cos \theta_e, \Phi) d\Phi. \quad (\text{B.2})$$

Splitting the sum over n , and exchanging the integral with the summation results in

$$\begin{aligned} \mathcal{P}(\cos \theta_K, \cos \theta_e) &\propto \sum_{l,m} C_{lm0} L_l(\cos \theta_K) L_m(\cos \theta_e) \int_{-\pi}^{\pi} \mathcal{P}_{\text{phys.}}(\cos \theta_K, \cos \theta_e, \Phi) d\Phi \\ &+ \sum_{\substack{l,m \\ n \neq 0}} C_{lmn} L_l(\cos \theta_K) L_m(\cos \theta_e) \int_{-\pi}^{\pi} \mathcal{P}_{\text{phys.}}(\cos \theta_K, \cos \theta_e, \Phi) \cos(n\Phi) d\Phi. \end{aligned} \quad (\text{B.3})$$

The integral in the first term is simply the $(\cos \theta_K, \cos \theta_e)$ projection of $\mathcal{P}_{\text{phys.}}$, and it doesn't depend on any order, so it can be factored out of the summation. Similarly, the first summation is (proportional to) the $(\cos \theta_K, \cos \theta_e)$ projection of the angular acceptance, since, following the orthogonality of the $\cos(n\Phi)$ term, only the $n = 0$ element results in a non-vanishing integral,

$$\begin{aligned} \int_{-\pi}^{\pi} \epsilon(\cos \theta_K, \cos \theta_e, \Phi) d\Phi &= \int_{-\pi}^{\pi} \sum_{l,m,n} C_{lmn} L_l(\cos \theta_K) L_m(\cos \theta_e) \cos(n\Phi) d\Phi \\ &= 2\pi \sum_{l,m} C_{lm0} L_l(\cos \theta_K) L_m(\cos \theta_e). \end{aligned} \quad (\text{B.4})$$

Defining the integral in the second summation as

$$\mathcal{I}_n(\cos \theta_K, \cos \theta_e) = \int_{-\pi}^{\pi} \mathcal{P}_{\text{phys.}}(\cos \theta_K, \cos \theta_e, \Phi) \cdot \cos(n\Phi) d\Phi, \quad (\text{B.5})$$

one obtains the final expression

$$\begin{aligned} \mathcal{P}(\cos \theta_K, \cos \theta_e) &\propto \frac{1}{2\pi} \epsilon(\cos \theta_K, \cos \theta_e) \cdot \mathcal{P}_{\text{phys.}}(\cos \theta_K, \cos \theta_e) \\ &+ \sum_{\substack{l,m \\ n \neq 0}} C_{lmn} \mathcal{I}_n(\cos \theta_K, \cos \theta_e) L_l(\cos \theta_K) L_m(\cos \theta_e). \end{aligned} \quad (\text{B.6})$$

One can follow a similar logic when computing the double integral over $(\cos \theta_K, \cos \theta_e)$,

$$\mathcal{P}(\Phi) \propto \int_{-1}^1 \int_{-1}^1 \sum_{l,m,n} C_{lmn} L_l(\cos \theta_K) L_m(\cos \theta_e) \cos(n\Phi) \cdot \mathcal{P}_{\text{phys.}}(\cos \theta_K, \cos \theta_e, \Phi) d\cos \theta_K d\cos \theta_e, \quad (\text{B.7})$$

which is split into

$$\begin{aligned}
\mathcal{P}(\Phi) &\propto \sum_n C_{00n} \cos(n\Phi) \int_{-1}^1 \int_{-1}^1 \mathcal{P}_{\text{phys.}}(\cos \theta_K, \cos \theta_e, \Phi) d \cos \theta_K d \cos \theta_e \\
&+ \sum_{\substack{n \\ (l,m) \neq (0,0)}} C_{lmn} \cos(n\Phi) \int_{-1}^1 \int_{-1}^1 \mathcal{P}_{\text{phys.}}(\cos \theta_K, \cos \theta_e, \Phi) L_l(\cos \theta_K) L_m(\cos \theta_e) d \cos \theta_K d \cos \theta_e .
\end{aligned} \tag{B.8}$$

The first term is again proportional to the product of the two projections, and defining the integral of the second term as

$$\mathcal{I}_{lm}(\Phi) = \int_{-1}^1 \int_{-1}^1 \mathcal{P}_{\text{phys.}}(\cos \theta_K, \cos \theta_e, \Phi) \cdot L_l(\cos \theta_K) L_m(\cos \theta_e) d \cos \theta_K d \cos \theta_e, \tag{B.9}$$

one obtains the final expression

$$\begin{aligned}
\mathcal{P}(\Phi) &\propto \frac{1}{4} \epsilon(\Phi) \cdot \mathcal{P}_{\text{phys.}}(\Phi) \\
&+ \sum_{\substack{n \\ (l,m) \neq (0,0)}} C_{lmn} \mathcal{I}_{lm}(\Phi) \cos(n\Phi) .
\end{aligned} \tag{B.10}$$

Residual integrals

When integrating over the Φ dimension the following integrals are non-vanishing,

$$\begin{aligned}
\mathcal{I}_1(\cos \theta_K, \cos \theta_e) &= \frac{9}{32} (A_5 + 2S_4 \cos \theta_e) \sin 2\theta_K \sin \theta_e , \\
\mathcal{I}_2(\cos \theta_K, \cos \theta_e) &= \frac{9}{32} S_3 \sin^2 \theta_K \sin^2 \theta_e .
\end{aligned} \tag{B.11}$$

Integrating over $(\cos \theta_K, \cos \theta_e)$,

$$\begin{aligned}
\mathcal{I}_{02}(\Phi) &= \frac{1}{20\pi} [1 - 3F_L - 2S_3 \cos(2\Phi) - 2A_9 \sin(2\Phi)] , \\
\mathcal{I}_{20}(\Phi) &= \frac{1}{10\pi} [-1 + 3F_L - S_3 \cos(2\Phi) - A_9 \sin(2\Phi)] , \\
\mathcal{I}_{22}(\Phi) &= \frac{1}{100\pi} [-1 - 3F_L + 2S_3 \cos(2\Phi) + 2A_9 \sin(2\Phi)] .
\end{aligned} \tag{B.12}$$

All integrals with $m \geq 4$ or $l \geq 4$ vanish.¹

¹The integrals are actually non-vanishing when including the effects of the $D_s^\pm$ veto $\epsilon_{\text{veto}}(\cos \theta_e)$, however the argument that they are numerically negligible is still valid.

Time-integrated observables as functions of transversity amplitudes

The author's recommendation

Everything is romantic
from the album **brat**
by **Charli xcx**
2024, Atlantic

For $B_s^0 \rightarrow J/\psi\phi$ and $B_s^0 \rightarrow \psi(2S)\phi$ decays, the expressions for the coefficients J_i and h_i can be simplified. Firstly, as the decays $J/\psi, \psi(2S) \rightarrow e^+e^-$ do not have a chirality preference, there are no distinctions between left- and right-handed amplitudes. Secondly, one can write to very good approximation $e^{-i\phi}\bar{A}_i/A_i = 1$ [17], where ϕ is the phase related to the $B_s^0\text{--}\bar{B}_s^0$ mixing. Thus the coefficients simplify to

$$J_{1s}(t=0) = \frac{3}{4} (|A_\perp|^2 + |A_\parallel|^2), \quad h_{1s} = \frac{3}{2} (-|A_\perp|^2 + |A_\parallel|^2), \quad (\text{C.1})$$

$$J_{1c}(t=0) = |A_0|^2, \quad h_{1c} = 2|A_0|^2, \quad (\text{C.2})$$

$$J_{2s}(t=0) = \frac{1}{4} (|A_\perp|^2 + |A_\parallel|^2), \quad h_{2s} = \frac{1}{2} (-|A_\perp|^2 + |A_\parallel|^2), \quad (\text{C.3})$$

$$J_{2c}(t=0) = -|A_0|^2, \quad h_{2c} = -2|A_0|^2, \quad (\text{C.4})$$

$$J_3(t=0) = \frac{1}{2} (|A_\perp|^2 - |A_\parallel|^2), \quad h_3 = -(|A_\perp|^2 + |A_\parallel|^2). \quad (\text{C.5})$$

One can substitute the coefficients for $i = 1s, 2s, 1c, 2c$ in Eq. (2.58) to write the time-

integrated differential decay rate,

$$\begin{aligned} \int_0^\infty \frac{d(\Gamma + \bar{\Gamma})}{dq^2 dt} \epsilon(t) dt = & \frac{3}{4} \mathcal{I} \left\{ 3 \left[|A_\perp|^2 + |A_\parallel|^2 - \mathcal{D} \left(-|A_\perp|^2 + |A_\parallel|^2 \right) \right] + 2|A_0|^2 (1 - \mathcal{D}) \right\} \\ & - \frac{1}{4} \mathcal{I} \left\{ \left[|A_\perp|^2 + |A_\parallel|^2 - \mathcal{D} \left(-|A_\perp|^2 + |A_\parallel|^2 \right) \right] - 2|A_0|^2 (1 - \mathcal{D}) \right\}, \end{aligned} \quad (\text{C.6})$$

where

$$\mathcal{I} = \int_0^\infty \epsilon(t) e^{-\Gamma_s t} \cosh \frac{\Delta \Gamma_s t}{2} dt. \quad (\text{C.7})$$

Grouping the various terms by the three amplitudes one finally gets, making use of the relation for the transversity amplitudes at $t = 0$, $|A_0|^2 + |A_\parallel|^2 + |A_\perp|^2 = 1$,

$$\int_0^\infty \frac{d(\Gamma + \bar{\Gamma})}{dq^2 dt} \epsilon(t) dt = 2(1 - \mathcal{D}) \mathcal{I} \left(1 + \frac{2\mathcal{D}}{1 - \mathcal{D}} |A_\perp|^2 \right). \quad (\text{C.8})$$

Substituting the expressions for the coefficients h_i in Eq. (6.27) for $i = 1c, 3$, one can write the time-integrated observables $\langle F_L \rangle$ and $\langle S_3 \rangle$, using Eq. (6.3), as functions of the amplitudes at $t = 0$,

$$\langle F_L \rangle = \frac{2|A_0|^2 - 2\mathcal{D}|A_0|^2}{2(1 - \mathcal{D}) \left(1 + \frac{2\mathcal{D}}{1 - \mathcal{D}} |A_\perp|^2 \right)} = \frac{|A_0|^2}{1 + \frac{2\mathcal{D}}{1 - \mathcal{D}} |A_\perp|^2}, \quad (\text{C.9})$$

$$\langle S_3 \rangle = \frac{\left(|A_\perp|^2 - |A_\parallel|^2 \right) + \mathcal{D} \left(|A_\perp|^2 + |A_\parallel|^2 \right)}{2(1 - \mathcal{D}) \left(1 + \frac{2\mathcal{D}}{1 - \mathcal{D}} |A_\perp|^2 \right)} = \frac{|A_0|^2 + \frac{2}{1 - \mathcal{D}} |A_\perp|^2 - 1}{2 \left(1 + \frac{2\mathcal{D}}{1 - \mathcal{D}} |A_\perp|^2 \right)}, \quad (\text{C.10})$$

where $\mathcal{I}$ has disappeared in the ratio, and having again made use of the $t = 0$ expression $|A_0|^2 + |A_\perp|^2 + |A_\parallel|^2 = 1$.

Estimating parameter confidence intervals

The author's recommendation

Ricomincio Da Qui

from the album **Grovigli**

by **Malika Ayane**

2010, Sugar Music

The pseudoexperiments used to validate the fit strategy in Sec. 6.5, indicate that in some cases, in particular in the central- q^2 bin, a significant fraction of the fits hits one of the parameter boundaries, which can lead to an incorrect estimation of the statistical coverage. To recover more reliable coverage intervals, the Feldman–Cousins (FC) procedure is used [139]. For consistency reasons, the method is used for all three q^2 bins, and for both sets of angular parameters $[\langle F_L \rangle, \langle A'_6 \rangle]$, and $[\langle S_3 \rangle, \langle A_9 \rangle]$.

The FC method uses likelihood ratios as an ordering principle to determine confidence intervals, defined in the frequentist sense: given a parameter μ of unknown value μ_{true} , for a given confidence level (CL) α the procedure will return an interval $[\mu_1, \mu_2]$ that's part of a set with the following property

$$P(\mu \in [\mu_1, \mu_2]) = \alpha, \tag{D.1}$$

for all allowed values of μ . If repeating the same experiment multiple times, this means that the unknown true value μ_{true} will be contained in a fraction α of the resulting intervals; the intervals are said to have coverage α . If for a given value of μ the probability of it being in the interval is less than α , then the interval is said to be undercovering. Similarly, if μ is contained with a probability higher than α , the interval is said to be overcovering.

To construct such intervals, a pseudoexperiment based prescription can be used, scanning the possible values of the parameter of interest μ , as follows [140]:

- From the known PDF $\mathcal{P}(x|\mu)$ in the observable x , used to construct a maximum likelihood estimator for the parameter μ , generate N pseudoexperiments assuming the value $\mu = \mu_i$. The yield of each generated dataset is extracted from a Poisson distribution with mean set to the measured yield in data.

- For each pseudoexperiment, compute the log-likelihood ratio

$$\lambda_{ij} = -2 \log \frac{\mathcal{L}_j(x|\mu_i)}{\mathcal{L}_j(x|\hat{\mu}_i)}, \quad (\text{D.2})$$

where the numerator is the likelihood computed for the j th dataset when fixing the parameter μ to the generated value μ_i , while the denominator is the value of the likelihood when maximised over μ (with the hat symbol to indicate the best fit value).

- Compute a similar ratio in data,

$$\lambda_{i,\text{data}} = -2 \log \frac{\mathcal{L}_{\text{data}}(x|\mu_i)}{\mathcal{L}_{\text{data}}(x|\hat{\mu})}, \quad (\text{D.3})$$

where now the likelihood is constructed over the data, and the best fit value $\hat{\mu}$ is what is measured.

- The $1 - \text{CL}$ (*i.e.* $1 - \alpha$) value is the fraction of pseudoexperiments that have a likelihood ratio larger than what is determined in data

$$1 - \text{CL} = \frac{N(\lambda_{i,\text{data}} < \lambda_{ij})}{N}. \quad (\text{D.4})$$

- Compute the $1 - \text{CL}$ value for a set of true values of the parameter μ , and obtain a curve as a function of μ .
- The confidence interval of coverage α_0 is then defined as the set

$$I_\mu = \{\mu | 1 - \alpha(\mu) > 1 - \alpha_0\}. \quad (\text{D.5})$$

When generating pseudoexperiments, one probes the distribution of the test statistic λ under the assumption that the true value of μ is μ_i . With a large enough number of these, computing the fraction in Eq. (D.4) is equivalent to computing the cumulative distribution function for the test statistic, *i.e.* the probability of getting a result that's at least as extreme as the result obtained in data. Then, one obtains an interval by including all values of μ_i for which such a probability is larger than that corresponding to the desired coverage, α_0 .

Nuisance parameters

The FC method, in the absence of nuisance parameters, has perfect coverage by construction. However, when introducing nuisance parameters $\vec{\lambda}$ to the problem, then the full procedure requires scanning the multidimensional parameter space, which is computationally expensive and most times not feasible. In this work, the so called “plug-in” method is used [141], which makes some modifications on the prescription described above:

- The pseudoexperiments are generated from the PDF $\mathcal{P}(x|\mu_i, \vec{\lambda} = \vec{\hat{\lambda}}_i)$, where the nuisance parameters are set to their best fit values $\vec{\hat{\lambda}}_i$, obtained by fitting the data when

fixing the parameter of interest to μ_i .

- The likelihood ratios are similarly modified

$$\begin{aligned}\lambda_{ij} &= -2 \log \frac{\mathcal{L}_j(x|\mu_i, \vec{\hat{\lambda}}_i)}{\mathcal{L}_j(x|\hat{\mu}_i, \vec{\hat{\lambda}}_i)}, \\ \lambda_{i,\text{data}} &= -2 \log \frac{\mathcal{L}_{\text{data}}(x|\mu_i, \vec{\hat{\lambda}}_i)}{\mathcal{L}_{\text{data}}(x|\hat{\mu}, \vec{\hat{\lambda}})},\end{aligned}\tag{D.6}$$

where now in the numerator the likelihood function is maximised over the nuisance parameters (and the best fit value is indicated by the two hats), but keeping μ fixed, and in the denominator the likelihood is maximised over all parameters (the best fit values indicated by the single hat).

It is easy to apply this prescription to a two-dimensional situation, like in this work (where two parameters are fitted simultaneously). The $1 - \text{CL}$ values are scanned over a two-dimensional grid, and the contours at the desired coverage are obtained by slicing the distribution at the level $1 - \alpha_0$. For both $(m(K^+K^-e^+e^-), \cos\theta_K, \cos\theta_e)$ and $(m(K^+K^-e^+e^-), \Phi)$ fits, this can be done to produce the coverage plots in the $[\langle F_L \rangle, \langle A'_6 \rangle]$ and $[\langle S_3 \rangle, \langle A_9 \rangle]$ planes. To quote the intervals on each individual parameter, the other is then treated as a nuisance parameter as described above.

Appendix E

Detailed breakdown of systematic uncertainties on the angular observables

The author's recommendation

Slow Burn

from the album **Golden Hour**

by **Kacey Musgraves**

2018, MCA Nashville

Tables E.1, E.2, and E.3 report summaries of the size of the systematic uncertainties for each individual source, in the low-, central-, and high- q^2 bins, respectively.

Table E.1: Summary of the sources of systematic uncertainties in the low- q^2 bin. The statistical uncertainty reported is estimated from the pseudoexperiments of Sec. 6.5.

Source	Variations	$\langle F_L \rangle$	$\langle A'_6 \rangle$	$\langle S_3 \rangle$	$\langle A_9 \rangle$
Signal	Mass model	0.003	0.010	0.007	0.007
	Model factorisation	0.012	0.016	0.011	0.016
Mis-ID background	Mass model	0.006	0.010	0.008	0.008
	Angular model	0.004	0.010	0.009	0.011
	Model factorisation	0.055	0.012	0.040	0.008
	Sample size	0.002	0.001	0.000	0.000
	Kinematic binning	0.004	0.010	0.008	0.008
	Fail region	0.004	0.009	0.007	0.007
Semileptonic background	Mass model	0.004	0.010	0.008	0.008
	Angular model	0.009	0.010	0.008	0.008
	Model factorisation	0.004	0.010	0.009	0.008
	Sample size	0.002	0.000	0.000	0.000
	Form factor model	0.009	0.010	—	—
Combinatorial background	Model factorisation	0.020	0.012	0.043	0.010
Angular acceptance	MC sample size	0.003	0.000	0.007	0.000
	Highest order	0.004	0.010	0.008	0.007
	MC corrections	0.004	0.009	0.008	0.007
	q^2 point	0.004	0.010	0.010	0.007
	$\Delta\Gamma_s = 0$ in MC	0.008	0.013	0.011	0.011
	Veto model	0.004	0.009	—	—
	q^2 migration on veto	0.004	0.009	—	—
Angular resolution		0.007	0.010	0.008	0.017
Exclusion of S-wave component		0.005	0.013	0.013	0.009
Total systematic uncertainty		0.06	0.05	0.07	0.04
Statistical uncertainty		0.12	0.33	0.26	0.24

Table E.2: Summary of the sources of systematic uncertainties in the central- q^2 bin. The statistical uncertainty reported is estimated from the pseudoexperiments of Sec. 6.5.

Source	Variations	$\langle F_L \rangle$	$\langle A'_6 \rangle$	$\langle S_3 \rangle$	$\langle A_9 \rangle$
Signal	Mass model	0.004	0.017	0.007	0.007
	Model factorisation	0.021	0.022	0.008	0.008
Leakage background	Mass model	0.004	0.017	0.007	0.007
	Angular model	0.014	0.020	0.007	0.007
	Model factorisation	0.006	0.019	0.018	0.010
	Sample size	0.003	0.000	0.002	0.002
Mis-ID background	Mass model	0.016	0.016	0.007	0.007
	Angular model	0.009	0.018	0.008	0.007
	Model factorisation	0.005	0.020	0.052	0.007
	Sample size	0.010	0.001	0.003	0.001
	Kinematic binning	0.006	0.018	0.007	0.007
	Fail region	0.017	0.016	0.007	0.006
Semileptonic background	Mass model	0.005	0.017	0.007	0.007
	Angular model	0.007	0.017	0.011	0.008
	Model factorisation	0.005	0.023	0.027	0.009
	Sample size	0.002	0.000	0.000	0.000
	Form factor model	0.004	0.018	—	—
Combinatorial background	Model factorisation	0.032	0.040	0.009	0.013
Angular acceptance	MC sample size	0.002	0.000	0.004	0.000
	Highest order	0.005	0.017	0.007	0.007
	MC corrections	0.004	0.018	0.007	0.007
	q^2 point	0.015	0.018	0.009	0.007
	$\Delta\Gamma_s = 0$ in MC	0.006	0.025	0.010	0.010
	Veto model	0.004	0.018	—	—
	q^2 migration on veto	0.004	0.017	—	—
Angular resolution		0.019	0.023	0.007	0.017
Exclusion of S-wave component		0.016	0.018	0.008	0.009
Total systematic uncertainty		0.06	0.09	0.07	0.04
Statistical uncertainty		0.12	0.57	0.24	0.25

Table E.3: Summary of the sources of systematic uncertainties in the high- q^2 bin. The statistical uncertainty reported is estimated from the pseudoexperiments of Sec. 6.5.

Source	Variations	$\langle F_L \rangle$	$\langle A'_6 \rangle$	$\langle S_3 \rangle$	$\langle A_9 \rangle$
Signal	Mass model	0.004	0.008	0.009	0.008
	Model factorisation	0.005	0.016	0.014	0.010
$B_s^0 \rightarrow J/\psi (e^+e^-) \phi$ leakage background	Mass model	0.004	0.008	0.008	0.007
	Angular model	0.004	0.008	0.007	0.007
	Model factorisation	0.007	0.009	0.017	0.009
	Sample size	0.001	0.000	0.002	0.002
$B_s^0 \rightarrow \psi(2S) (e^+e^-) \phi$ leakage background	Mass model	0.005	0.008	0.008	0.007
	Angular model	0.006	0.009	0.007	0.007
	Model factorisation	0.006	0.012	0.011	0.008
	Sample size	0.002	0.000	0.004	0.003
Mis-ID background	Mass model	0.008	0.008	0.009	0.007
	Angular model	0.013	0.008	0.016	0.007
	Model factorisation	0.023	0.016	0.067	0.019
	Sample size	0.011	0.000	0.005	0.001
	Kinematic binning	0.004	0.008	0.008	0.007
	Fail region	0.008	0.008	0.013	0.007
Combinatorial background	Mass model high- q^2	0.014	0.016	0.008	0.015
	Model factorisation	0.005	0.010	0.008	0.017
Angular acceptance	MC sample size	0.002	0.000	0.004	0.000
	Highest order	0.004	0.008	0.007	0.007
	MC corrections	0.004	0.008	0.007	0.007
	q^2 point	0.015	0.009	0.009	0.007
	$\Delta\Gamma_s = 0$ in MC	0.007	0.012	0.011	0.010
	Veto model	0.004	0.008	—	—
	q^2 migration on veto	0.005	0.008	—	—
Angular resolution		0.004	0.020	0.007	0.011
Excluded part. reco. background in high- q^2		0.010	0.008	0.014	0.011
Exclusion of S-wave component		0.008	0.008	0.014	0.007
Total systematic uncertainty		0.05	0.05	0.08	0.05
Statistical uncertainty		0.13	0.29	0.26	0.25

Bibliography

- [1] M. E. Peskin and D. V. Schroeder, *An introduction to quantum field theory*, Addison-Wesley, 1995.
- [2] M. Maggiore, *A modern introduction to quantum field theory*, vol. 12, Oxford University Press, 2005.
- [3] P. A. M. Dirac and R. H. Fowler, *The quantum theory of the electron*, Proceedings of the Royal Society of London **A117** (1928) 610.
- [4] E. Noether, *Invariante variationsprobleme*, Nachr. v. d. Ges. d. Wiss. zu Göttingen **1918** (1918) 235.
- [5] C. S. Wu *et al.*, *Experimental test of parity conservation in beta decay*, Phys. Rev. **105** (1957) 1413.
- [6] R. L. Garwin, L. M. Lederman, and M. Weinrich, *Observations of the failure of conservation of parity and charge conjugation in meson decays: the magnetic moment of the free muon*, Phys. Rev. **105** (1957) 1415.
- [7] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.
- [8] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.
- [9] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global conservation laws and massless particles*, Phys. Rev. Lett. **13** (1964) 585.
- [10] Y. Nambu, *Quasi-particles and gauge invariance in the theory of superconductivity*, Phys. Rev. **117** (1960) 648.
- [11] P. W. Anderson, *Plasmons, gauge invariance, and mass*, Phys. Rev. **130** (1963) 439.
- [12] J. Goldstone, *Field theories with “superconductor” solutions*, Nuovo Cimento **19** (1961) 154.

- [13] LHCb collaboration, *List of hadrons observed at the LHC*, LHCb-FIGURE-2021-001, 2021.
- [14] G. 't Hooft, *Renormalizable Lagrangians for massive Yang-Mills fields*, Nuclear Physics **B35** (1971) 167.
- [15] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.
- [16] M. Kobayashi and T. Maskawa, *CP-violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49** (1973) 652.
- [17] Particle Data Group, S. Navas *et al.*, *Review of particle physics*, Phys. Rev. **D110** (2024) 030001.
- [18] LHCb collaboration, R. Aaij *et al.*, *Amplitude analysis of $B^\pm \rightarrow \pi^\pm K^+ K^-$ decays*, Phys. Rev. Lett. **123** (2019) 231802, [arXiv:1905.09244](#).
- [19] M. Gell-Mann and A. Pais, *Behavior of neutral particles under charge conjugation*, Phys. Rev. **97** (1955) 1387.
- [20] K. Lande *et al.*, *Observation of long-lived neutral V particles*, Phys. Rev. **103** (1956) 1901.
- [21] Planck Collaboration, Aghanim, N. *et al.*, *Planck 2018 results - I. Overview and the cosmological legacy of Planck*, Astronomy & Astrophysics **641** (2020) A1, [arXiv:1807.06205](#).
- [22] Planck Collaboration, Aghanim, N. *et al.*, *Planck 2018 results - VI. Cosmological parameters (Corrigendum)*, Astronomy & Astrophysics **652** (2021) C4, [arXiv:1807.06209](#).
- [23] A. D. Sakharov, *Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.
- [24] M. B. Gavela, P. Hernandez, J. Orloff, and O. Pène, *Standard model CP violation and baryon asymmetry*, Mod. Phys. Lett. **A9** (1994) 795, [arXiv:hep-ph/9312215](#).
- [25] P. Huet and E. Sather, *Electroweak baryogenesis and Standard Model CP violation*, Phys. Rev. **D51** (1995) 379, [arXiv:hep-ph/9404302](#).
- [26] J. Albrecht, D. van Dyk, and C. Langenbruch, *Flavour anomalies in heavy quark decays*, Prog. Part. Nucl. Phys. **120** (2021) 103885, [arXiv:2107.04822](#).
- [27] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b -hadron, c -hadron, and τ -lepton properties as of 2021*, Phys. Rev. **D107** (2023) 052008, [arXiv:2206.07501](#), updated results and plots available at <https://hflav.web.cern.ch>.

- [28] A. Ali, G.-F. Giudice, and T. Mannel, *Towards a model-independent analysis of rare B decays*, Z. Phys. **C67** (1995) 417, [arXiv:hep-ph/9408213](#).
- [29] LHCb collaboration, R. Aaij *et al.*, *Measurement of lepton universality parameters in $B^+ \rightarrow K^+ \ell^+ \ell^-$ and $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ decays*, Phys. Rev. **D108** (2023) 032002, [arXiv:2212.09153](#).
- [30] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality in $b \rightarrow s \ell^+ \ell^-$ decays*, Phys. Rev. Lett. **131** (2023) 051803, [arXiv:2212.09152](#).
- [31] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP-averaged observables in the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay*, Phys. Rev. Lett. **125** (2020) 011802, [arXiv:2003.04831](#).
- [32] ATLAS collaboration, M. Aaboud *et al.*, *Angular analysis of $B_d^0 \rightarrow K^* \mu^+ \mu^-$ decays in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **10** (2018) 047, [arXiv:1805.04000](#).
- [33] CMS collaboration, A. M. Sirunyan *et al.*, *Measurement of angular parameters from the decay $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ in proton-proton collisions at $\sqrt{s} = 8$ TeV*, Phys. Lett. **B781** (2018) 517, [arXiv:1710.02846](#).
- [34] CMS collaboration, A. Hayrapetyan *et al.*, *Angular analysis of the $B^0 \rightarrow K^*(892)^0 \mu^+ \mu^-$ decay in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Phys. Lett. **B864** (2025) 139406, [arXiv:2411.11820](#).
- [35] Belle collaboration, S. Wehle *et al.*, *Lepton-flavor-dependent angular analysis of $B \rightarrow K^* \ell^+ \ell^-$* , Phys. Rev. Lett. **118** (2017) 111801, [arXiv:1612.05014](#).
- [36] LHCb collaboration, R. Aaij *et al.*, *Branching fraction measurements of the rare $B_s^0 \rightarrow \phi \mu^+ \mu^-$ and $B_s^0 \rightarrow f_2'(1525) \mu^+ \mu^-$ decays*, Phys. Rev. Lett. **127** (2021) 151801, [arXiv:2105.14007](#).
- [37] E. Fermi, *Tentativo di una teoria dei raggi β* , Il Nuovo Cimento **11** (1934) 1.
- [38] E. Fermi, *Versuch einer theorie der β -Strahlen. I*, Zeitschrift für Physik **88** (1934) 161.
- [39] G. Isidori, F. Wilsch, and D. Wyler, *The standard model effective field theory at work*, Rev. Mod. Phys. **96** (2024) 015006.
- [40] S. Descotes-Genon and J. Virto, *Time dependence in $B \rightarrow V \ell \ell$ decays*, JHEP **04** (2015) 045, Erratum *ibid.* **07** (2015) 049, [arXiv:1502.05509](#).
- [41] C. Bobeth, G. Hiller, and G. Piranishvili, *CP asymmetries in $\bar{B} \rightarrow \bar{K}^*(\rightarrow \bar{K} \pi) \bar{\ell} \ell$ and untagged $\bar{B}_s, B_s \rightarrow \phi(\rightarrow K^+ K^-) \bar{\ell} \ell$ decays at NLO*, JHEP **07** (2008) 106, [arXiv:0805.2525](#).
- [42] M. Algueró *et al.*, *To (b)e or not to (b)e: no electrons at LHCb*, Eur. Phys. J. **C83** (2023) 648, [arXiv:2304.07330](#).

- [43] M. Ciuchini *et al.*, *Constraints on lepton universality violation from rare B decays*, Phys. Rev. **D107** (2023) 055036, [arXiv:2212.10516](#).
- [44] T. Hurth, F. Mahmoudi, and S. Neshatpour, *b anomalies in the post- $R_{K^{(*)}}$ era*, Phys. Rev. **D108** (2023) 115037, [arXiv:2310.05585](#).
- [45] S. U. Chung, *Spin formalisms*, CERN-71-08, CERN, 1971. Academic Training Lectures, CERN, Geneva, doi: 10.5170/CERN-1971-008.
- [46] A. Bharucha, D. M. Straub, and R. Zwicky, *$B \rightarrow V \ell^+ \ell^-$ in the Standard Model from light-cone sum rules*, JHEP **08** (2016) 098, [arXiv:1503.05534](#).
- [47] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, *Lattice QCD calculation of form factors describing the rare decays $B \rightarrow K^* \ell^+ \ell^-$ and $B_s \rightarrow \phi \ell^+ \ell^-$* , Phys. Rev. **D89** (2014) 094501, [arXiv:1310.3722](#).
- [48] D. M. Straub, *flavio: a Python package for flavour and precision phenomenology in the Standard Model and beyond*, [arXiv:1810.08132](#).
- [49] L. Evans and P. Bryant, *LHC machine*, JINST **3** (2008) S08001.
- [50] T. Taylor and D. Treille, *The Large Electron Positron Collider (LEP): probing the Standard Model*, Adv. Ser. Direct. High Energy Phys. **27** (2017) 217.
- [51] M. S. Camillocci, *LHC nominal cycle*, in *6th Evian Workshop on LHC beam operation*, (Geneva), 45–48, CERN, 2016.
- [52] R. Alemany Fernandez, M. Lamont, and S. Page, *LHC modes*, LHC-OP-ES-0022, CERN, Geneva, 2019.
- [53] R. Bailey and P. Collier, *Standard filling schemes for various LHC operation modes*, LHC-PROJECT-NOTE-323, CERN, Geneva, 2003.
- [54] M. Vretenar *et al.*, *Linac4 design report*, CERN-2020-006, CERN, Geneva, 2020. doi: 10.23731/CYRM-2020-006.
- [55] E. Mobs, *The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019*, <https://cds.cern.ch/record/2684277>, 2019. General Photo - Accessed on 2024-09-08.
- [56] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B^\pm$ production cross-section in pp collisions at $\sqrt{s} = 7$ TeV*, JHEP **04** (2012) 093, [arXiv:1202.4812](#).
- [57] LHCb collaboration, R. Aaij *et al.*, *Measurement of B meson production cross-sections in proton-proton collisions at $\sqrt{s} = 7$ TeV*, JHEP **08** (2013) 117, [arXiv:1306.3663](#).
- [58] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $\bar{B}^0 - B^0$ and $\bar{B}_s^0 - B_s^0$ production asymmetries in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. **B739** (2014) 218, [arXiv:1408.0275](#).

- [59] LHCb collaboration, R. Aaij *et al.*, *Measurement of B_c^+ production in proton-proton collisions at $\sqrt{s}=8$ TeV*, Phys. Rev. Lett. **114** (2015) 132001, [arXiv:1411.2943](#).
- [60] LHCb collaboration, R. Aaij *et al.*, *Study of the productions of Λ_b^0 and $\bar{B}^0$ hadrons in pp collisions and first measurement of the $\Lambda_b^0 \rightarrow J/\psi p K^-$ branching fraction*, Chin. Phys. **C40** (2016) 011001, [arXiv:1509.00292](#).
- [61] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B^\pm$ production asymmetry and the CP asymmetry in $B^\pm \rightarrow J/\psi K^\pm$ decays*, Phys. Rev. **D95** (2017) 052005, [arXiv:1701.05501](#).
- [62] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B^\pm$ production cross-section in pp collisions at $\sqrt{s}=7$ and 13 TeV*, JHEP **12** (2017) 026, [arXiv:1710.04921](#).
- [63] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.
- [64] LHCb collaboration, C. Elsässer, *$\bar{b}b$ production angle plots*, https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html. Accessed on 2024-09-08.
- [65] F. Follin, R. Alemany-Fernandez, and R. Jacobsson, *Online luminosity optimization at the LHC*, in *14th International Conference on Accelerator and Large Experimental Physics Control Systems*, 2014.
- [66] B. Muratori and T. Pieloni, *Luminosity levelling techniques for the LHC*, in *ICFA Mini-Workshop on Beam-Beam Effects in Hadron Colliders*, CERN, Geneva, Switzerland, 18-22 Mar 2013, 177–181, 2014, [arXiv:1410.5646](#).
- [67] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, [arXiv:1311.3893](#).
- [68] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, [arXiv:1412.6352](#).
- [69] O. Callot and S. Hansmann-Menzemer, *The forward tracking: algorithm and performance studies*, LHCb-2007-015, CERN, Geneva, 2007.
- [70] M. Needham and J. Van Tilburg, *Performance of the track matching*, LHCb-2007-020, CERN, Geneva, 2007.
- [71] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. **A262** (1987) 444.
- [72] J. Van Tilburg, *Track simulation and reconstruction in LHCb*, CERN-THESIS-2005-040, PhD thesis, Vrije U. Amsterdam, 2005, Presented on 01 Sep 2005.
- [73] G. F. Knoll, *Radiation detection and measurement*, John Wiley & Sons, 2010.

- [74] S. Borghi *et al.*, *First spatial alignment of the LHCb VELO and analysis of beam absorber collision data*, Nucl. Instrum. Meth. **A618** (2010) 108 LHCb-DP-2010-001.
- [75] C. Abellán Beteta *et al.*, *Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC*, [arXiv:2008.11556](#).
- [76] L. Anderlini *et al.*, *Computing strategy for PID calibration samples for LHCb Run 2*, LHCb-PUB-2016-020, CERN, Geneva, 2016.
- [77] L. Anderlini *et al.*, *The PIDCalib package*, LHCb-PUB-2016-021, 2016.
- [78] F. Archilli *et al.*, *Performance of the muon identification at LHCb*, JINST **8** (2013) P10020, [arXiv:1306.0249](#).
- [79] R. Aaij *et al.*, *Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, JINST **14** (2019) P04013, [arXiv:1812.10790](#).
- [80] G. Barrand *et al.*, *GAUDI — A software architecture and framework for building HEP data processing applications*, Comput. Phys. Commun. **140** (2001) 45, CHEP2000.
- [81] R. Brun and F. Rademakers, *ROOT — An object oriented data analysis framework*, Nucl. Instrum. Meth. **A389** (1997) 81.
- [82] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.
- [83] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, [arXiv:0710.3820](#).
- [84] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.
- [85] Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.
- [86] *Changes to the data flow in Run 2 - LHCb Starterkit Lessons*, <https://lhcb.github.io/starterkit-lessons/first-analysis-steps/run-2-data-flow.html> Accessed on 2024-09-08.
- [87] LHCb collaboration, R. Aaij *et al.*, *Letter of Intent for the LHCb Upgrade*, CERN-LHCC-2011-001, LHCC-I-018, CERN, Geneva, 2011.
- [88] LHCb collaboration, R. Aaij *et al.*, and A. Bharucha *et al.*, *Implications of LHCb measurements and future prospects*, Eur. Phys. J. **C73** (2013) 2373, [arXiv:1208.3355](#).
- [89] LHCb collaboration, R. Aaij *et al.*, *The LHCb Upgrade I*, JINST **19** (2024) P05065, [arXiv:2305.10515](#).

- [90] LHCb collaboration, *LHCb Trigger and Online Upgrade Technical Design Report*, CERN-LHCC-2014-016, 2014.
- [91] LHCb collaboration, *LHCb calorimeters: Technical Design Report*, CERN-LHCC-2000-036, 2000.
- [92] LHCb collaboration, *LHCb Particle Identification Enhancement Technical Design Report*, CERN-LHCC-2023-005, 2023.
- [93] F. P. Albicocco *et al.*, *Long-term operation of the multi-wire-proportional-chambers of the LHCb muon system*, JINST **14** (2019) P11031, [arXiv:1908.02178](#).
- [94] LHCb collaboration, *HLT1 trigger efficiencies in 2024 data*, LHCb-FIGURE-2024-030, 2024.
- [95] C. Gaspar, M. Dönszelmann, and P. Charpentier, *Dim, a portable, light weight package for information publishing, data transfer and inter-process communication*, Comput. Phys. Commun. **140** (2001) 102, CHEP2000.
- [96] M. Adinolfi *et al.*, *LHCb data quality monitoring*, Journal of Physics: Conference Series **898** (2017) 092027.
- [97] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$* , JHEP **07** (2013) 084, [arXiv:1305.2168](#).
- [98] LHCb collaboration, R. Aaij *et al.*, *Angular analysis and differential branching fraction of the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$* , JHEP **09** (2015) 179, [arXiv:1506.08777](#).
- [99] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the rare decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$* , JHEP **11** (2021) 043, [arXiv:2107.13428](#).
- [100] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, *Data driven trigger efficiency determination at LHCb*, LHCb-PUB-2014-039, 2014.
- [101] M. Williams *et al.*, *The HLT2 topological lines*, LHCb-PUB-2011-002, CERN, Geneva, 2011.
- [102] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, JINST **8** (2013) P02013, [arXiv:1210.6861](#).
- [103] T. Likhomanenko *et al.*, *LHCb topological trigger reoptimization*, J. Phys. Conf. Ser. **664** (2015) 082025, [arXiv:1510.00572](#).
- [104] LHCb collaboration, R. Aaij *et al.*, *Measurement of the electron reconstruction efficiency at LHCb*, JINST **14** (2019) P11023, [arXiv:1909.02957](#).
- [105] LHCb collaboration, R. Aaij *et al.*, *Improved measurement of CP violation parameters in $B_s^0 \rightarrow J/\psi K^+ K^-$ decays in the vicinity of the $\phi(1020)$ resonance*, Phys. Rev. Lett. **132** (2024) 051802, [arXiv:2308.01468](#).

- [106] J. Amoraal *et al.*, *Application of vertex and mass constraints in track-based alignment*, Nucl. Instrum. Meth. **A712** (2013) 48, [arXiv:1207.4756](#).
- [107] R. Aaij *et al.*, *Selection and processing of calibration samples to measure the particle identification performance of the LHCb experiment in Run 2*, Eur. Phys. J. Tech. Instr. **6** (2019) 1, [arXiv:1803.00824](#).
- [108] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, [arXiv:physics/0402083](#).
- [109] A. Rogozhnikov, *Reweighting with boosted decision trees*, Journal of Physics: Conference Series **762** (2016) 012036.
- [110] A. V. Dorogush, V. Ershov, and A. Gulin, *CatBoost: gradient boosting with categorical features support*, [arXiv:1810.11363](#).
- [111] L. Prokhorenkova *et al.*, *CatBoost: unbiased boosting with categorical features*, in *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, (Red Hook, NY, USA), 6639–6649, 2018, [arXiv:1706.09516](#).
- [112] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, [arXiv:physics/0503191](#).
- [113] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality with $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ decays*, JHEP **08** (2017) 055, [arXiv:1705.05802](#).
- [114] LHCb collaboration, R. Aaij *et al.*, *Test of lepton flavour universality with $B_s^0 \rightarrow \phi \ell^+ \ell^-$ decays*, Phys. Rev. Lett. **134** (2025) 121803, [arXiv:2410.13748](#).
- [115] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, DESY-F31-86-02, PhD thesis, Institute of Nuclear Physics, Krakow, 1986.
- [116] B. Efron, *Bootstrap methods: Another look at the jackknife*, Ann. Statist. **7** (1979) 1.
- [117] L. Demortier and L. Lyons, *Everything you always wanted to know about pulls*, CD-F/ANAL/PUBLIC/5776, CDF, 2002.
- [118] S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, Ann. Math. Stat. **9** (1938) 60.
- [119] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, *Calculation of $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ and $B_s^0 \rightarrow \phi \mu^+ \mu^-$ observables using form factors from lattice QCD*, Phys. Rev. Lett. **112** (2014) 212003, [arXiv:1310.3887](#).
- [120] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, *Rare B decays using lattice QCD form factors*, PoS **LATTICE2014** (2015) 372, [arXiv:1501.00367](#).

- [121] M. Algueró, J. Matias, B. Capdevila, and A. Crivellin, *Disentangling lepton flavor universal and lepton flavor universality violating effects in $b \rightarrow sl + \ell$ - transitions*, Phys. Rev. **D105** (2022) 113007, [arXiv:2205.15212](#).
- [122] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the decay $B_s^0 \rightarrow \phi e^+ e^-$* , JHEP **07** (2025) 069, [arXiv:2504.06346](#).
- [123] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay $B^0 \rightarrow K^{*0} \mu^+ \mu^-$* , Phys. Rev. Lett. **108** (2012) 181806, [arXiv:1112.3515](#).
- [124] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP violating phase and decay-width difference in $B_s^0 \rightarrow \psi(2S) \phi$ decays*, Phys. Lett. **B762** (2016) 253, [arXiv:1608.04855](#).
- [125] S. Descotes-Genon, I. Plakias, and O. Sumensari, *On the impact of meson mixing on $B_s \rightarrow \phi e e$ angular observables at low q^2* , JHEP **02** (2023) 096, [arXiv:2210.11995](#).
- [126] D. Scora and N. Isgur, *Semileptonic meson decays in the quark model: An update*, Phys. Rev. **D52** (1995) 2783, [arXiv:hep-ph/9503486](#).
- [127] BaBar collaboration, B. Aubert *et al.*, *Study of the decay $D_s^+ \rightarrow K^+ K^- e^+ \nu_e$* , Phys. Rev. **D78** (2008) 051101, [arXiv:0807.1599](#).
- [128] D. Melikhov and B. Stech, *Weak form-factors for heavy meson decays: An update*, Phys. Rev. **D62** (2000) 014006, [arXiv:hep-ph/0001113](#).
- [129] M. Hatlo *et al.*, *Developments of mathematical software libraries for the LHC experiments*, IEEE Trans. Nucl. Sci. **52** (2005) 2818.
- [130] LHCb collaboration, *LHCb organisation*, https://lhcb.web.cern.ch/lhcb_page/collaboration/organization/default.html. Accessed on 2025-04-14.
- [131] J. Rademacker and B. Sciascia, *The early career, gender, and diversity actions within the LHCb Collaboration*, PoS **ICHEP2016** (2017) 318.
- [132] LHCb collaboration, E. A. Smith, *The early career, gender and diversity at LHCb*, PoS **ICHEP2020** (2021) 974.
- [133] LHCb collaboration, *LHCb Early Career, Gender and Diversity office*, https://lhcb.web.cern.ch/ECGD_Office/ECGD-intro.html. Accessed on 2025-04-14.
- [134] LHCb collaboration, *LHCb Early Career, Gender and Diversity office: ECGD Meetings*, https://lhcb.web.cern.ch/ECGD_Office/meetings.html. Accessed on 2025-04-14.
- [135] ALICE, ATLAS, CMS, LHCb collaboration, K. W. Coldham, *Report from the LHC experiments' Young Scientist Fora*, PoS **LHCP2024** (2025) 208.

- [136] LHC Early Career Scientist Fora, *Indico category: LHC Soft Skills Workshop*, <https://indico.cern.ch/category/17237/>. Accessed on 2025-04-14.
- [137] LHC Early Career Scientist Fora, *Indico category: LHC Job Matching Events*, <https://indico.cern.ch/category/17674/>. Accessed on 2025-04-14.
- [138] Women in Technology and LHC Early Career Scientist Fora, *Mentoring@CERN*, <https://mentoring-cern.web.cern.ch/>. Accessed on 2025-04-14.
- [139] G. J. Feldman and R. D. Cousins, *A unified approach to the classical statistical analysis of small signals*, Phys. Rev. **D57** (1998) 3873, [arXiv:physics/9711021](#).
- [140] T. M. Karbach, *Feldman–Cousins confidence levels – toy MC method*, [arXiv:1109.0714](#).
- [141] S. Bodhisattva, M. Walker, and M. Woodroffe, *On the unified method with nuisance parameters*, Statist. Sinica **19** (2009) 301.